

Marcus Schröter

Historische Volltextdatenbanken

1. Einführung zu historischen Volltextdatenbanken

1.1 Einführung

Volltextdatenbanken bestimmen die digitalen Arbeitstechniken von HistorikerInnen grundlegend: Sie können aus genuin digital entstandenen Texten hervorgegangen sein, aber auch aus der Digitalisierung einzelner Bücher oder Kollektionen in Bibliotheken, Archiven und Museen. Dabei orientiert man sich bei der Zusammenführung des Materials konzeptionell oft anhand epochaler, regionaler oder thematischer Kriterien, die ja auch für die Unterscheidung der Guides dieses Handbuchs bestimmend sind. Als Ergebnis präsentiert sich dem Benutzer die historische Volltextdatenbank als eine Einzelbestände übergreifende Sammlung, folgerichtig ist sie auch hier in den Teil „Sammlungen“ eingebettet.

Für die Geschichtswissenschaften lassen sich zwei grundlegende Typen von Volltextdatenbanken unterscheiden: für historische Quellen in Gestalt von Texten und Objekten einerseits, für historische Forschungsliteratur andererseits. Historische Volltextdatenbanken haben sich insbesondere in den vergangenen beiden Jahrzehnten ausgesprochen dynamisch entwickelt. Ihre analoge Vorgeschichte indessen ist sehr viel älter und hat ihre Wurzeln vielfach in der Verfilmung und Verfichtung ursprünglich gedruckter Einzelwerke oder Sammlungen zu teilweise umfangreichen kumulierten Textcorpora. Ein wichtiges Ziel in analoger und digitaler Zeit war und ist dabei stets, thematisch oder hinsichtlich der Provenienz zusammengehörende Texte zusammenzuführen. Gelangten Bibliotheken früher durch Mikrofilm- und Mikrofichearchive in den Besitz von Textkopien, deren gedruckte Originale sie in einer solchen Breite und Vollständigkeit meistens nicht besaßen, so blieb dieser Schatz an Wissen jedoch häufig verborgen – lag doch eine gewisse Barriere für die Benutzung in der Notwendigkeit, wenig komfortable Lesegeräte verwenden zu müssen. Automatische Volltextrecherchen, wie wir sie heute in Volltextdatenbanken gewohnt sind, waren im Falle von Mikroformen

selbstverständlich nicht möglich. Mit der Entwicklung der elektronischen Medien und der Digitalisierung wurde schließlich damit begonnen, Mikrofilme- und Mikrofiches zu digitalisieren und die wertvollen Inhalte in Datenbanken zu überführen, um sie der Benutzung leichter zugänglich zu machen. So entwickelten sich Volltextdatenbanken zu immer komplexeren digitalen Bibliotheken, die – wie bereits zuvor die Mikrofilm- und Mikrofichearchive – in der individuellen Bibliothek gedruckt vorhandene Texte ergänzten.

Inzwischen gehören Volltextdatenbanken zu den wichtigsten Instrumenten historischen Arbeitens, sei es für Quellentexte, sei es für Forschungsliteratur. Dabei ermöglicht die Volltextsuche eine Formulierung von Fragen und die Gewinnung von Erkenntnissen, für die eine automatisierte Analyse großer Textmengen grundlegende Voraussetzung ist. Volltextdatenbanken sind somit weit mehr als nur eine neue Form der Archivierung zuvor gedruckten Wissens, indem sie differenzierte Instrumente der Erschließung und Analyse von Texten bereitstellen und strukturierte Fragestellungen und Antworten ermöglichen. Auch in den nächsten Jahren ist mit einer kontinuierlichen Zunahme von Volltextdatenbanken zu rechnen, da nicht nur immer größere Textmengen erstmals digitalisiert, sondern auch viele der bisher lediglich als Images vorliegenden digitalen Sammlungen mit automatischen Texterkennungsverfahren aufbereitet werden. Darüber hinaus wird die Qualität der Erschließung digitaler Sammlungen durch Anwendung computerlinguistischer Verfahren zunehmen. In den Geisteswissenschaften wird diese Entwicklung gegenwärtig mit dem Begriff der *Digital Humanities* umschrieben. Hierunter versteht man – weit über die Digitalisierung unseres Wissens, unserer Texte und kulturellen Artefakte hinausgehend – deren intelligente Vernetzung innerhalb einer zunehmend digitalen Wissenschaftskultur. Zahlreiche Instrumente, die sich diese Digitalisierung nutzbar machen und Vernetzungen vornehmen, befinden sich gegenwärtig in der Erprobung: Für den deutschen und europäischen Bereich seien stellvertretend lediglich die Projekte *DA-RIAH-DE: Digitale Forschungsinfrastruktur für die Geistes- und Kultur-*

wissenschaften¹ sowie *TextGrid – Virtuelle Forschungsumgebung für die Geisteswissenschaften*² genannt. Für die internationale Community der HistorikerInnen sind beide Konzepte attraktiv, da sie zahlreiche Tools zur zeit- und ortsunabhängigen Arbeit mit Quellen und Texten zur Verfügung stellen und beispielsweise für die Erstellung digitaler Editionen bisher noch nicht annähernd ausgeschöpfte Möglichkeiten bieten.

1.2 Definition und Diskussion

Eine eindeutige Definition des Begriffs „Volltextdatenbank“ scheint auf den ersten Blick schwierig, wie der Blick auf seine Verwendung in der geschichtswissenschaftlichen Praxis einerseits, in der bibliothekarischen und informationstechnologischen Fachliteratur andererseits zeigt: Im *Datenbank-Infosystem (DBIS)*³, dem wichtigsten Verzeichnis wissenschaftlicher Datenbanken im deutschen Sprachraum, zählen zu den etwa 1.800 Datenbanken für das Fach Geschichte knapp 400 Volltextdatenbanken, wobei kontinuierlich neue hinzukommen. Betrachtet man das darin gebotene Spektrum im Einzelnen, so wird deutlich, dass in DBIS als Volltextdatenbanken inhaltlich und formal sehr heterogene Produkte verstanden werden.

Inhaltlich lassen sich unter Volltextdatenbanken für die Geschichtswissenschaft ganz allgemein Datenbanken verstehen, in denen unterschiedliche, für HistorikerInnen relevante Sammlungen von Texten und historischen Quellen publiziert werden. Bei Texten kann es sich sowohl um Primärtexte, als auch um Forschungsliteratur handeln, wobei zu „Quellen“ nicht nur Texte innerhalb von Karten oder auf unterschiedlichen Trägern wie Stein, Papyrus oder Pergament gezählt werden sollen, sondern auch nicht-textliche Überlieferungszeugnisse, wie beispielsweise museale Artefakte. Für diesen Überblick wird daher mit Blick sowohl auf die Terminologie des Datenbank-Infosystems, als auch auf die alltägliche Praxis von HistorikerInnen eine bewusst weite Defini-

¹ <https://de.dariah.eu>

² <https://textgrid.de>

³ <http://rzblx10.uni-regensburg.de/dbinfo/fachliste.php?lett=1>

tion des Begriffs Volltextdatenbank gewählt: Nicht nur reine Textarchive zählen dazu, sondern auch Objektdatenbanken, in denen der Anteil an Texten relativ gering ist. Ferner spielt es für unser Verständnis keine Rolle, ob die Texte lediglich als Images digitalisiert sind oder ob eine OCR-Erkennung implementiert und damit eine Volltextsuche möglich ist. Diese weite Auffassung kann somit im übertragenen Sinne auch dem *material turn* innerhalb der Geschichtswissenschaften gerecht werden: Objekte und Artefakte sind für die Geschichtswissenschaft inzwischen so wichtig geworden wie Texte, denn auch sie vermögen Geschichte zu „erzählen“, wenn sie – genau wie Texte auch – von HistorikerInnen angemessen „entziffert“ werden.

Um unter inhaltlichen und formalen Aspekten zu einer terminologischen Präzisierung des Begriffes „Volltextdatenbank“ zu gelangen, soll zunächst von der jeweiligen Definition des Kompositums „Volltextdatenbank“ im informationstechnologischen, bibliothekarischen und konkreten Verständnis der Informationspraxis am Beispiel des Datenbank-Infosystems (DBIS) ausgegangen werden. Separiert in seine bedeutungstragenden Bestandteile, „Volltext“ und „Datenbank“, lässt sich formal unter „Volltext“ ein vollständiger, elektronischer oder digitalisierter Text verstehen. Dieser kann einerseits aus einem ursprünglich gedruckten, retrodigitalisierten Text gewonnen werden, der durch Verfahren der automatischen optischen Zeichenerkennung mittels *Optical Character Recognition (OCR)* im Zusammenhang mit dem Digitalisierungsprozess aufbereitet ist. Um digitalisierte Handschriften oder Druckwerke als Volltexte lesbar und durchsuchbar zu machen, müssen die als Bilder erfassten Buchstaben, Wörter und Sätze einer handschriftlichen oder gedruckten Buchseite als solche erkannt, als bedeutungstragende Zeichen interpretiert und in maschinenlesbare Texte⁴ umgesetzt werden. In Zukunft könnten Verfahren der Texterkennung und Volltextgenerierung durch OCR immer weiter verbessert werden, so dass nicht nur die vielfältigen Drucktypen vom 16.–18. Jahrhundert, sondern auch Inkunabeltypen und Handschriften automatisch lesbar werden.

⁴ In der Regel im XML-Format.

Insbesondere im Hinblick auf die Paläographie als historische Grundwissenschaft liegen hierin große Potenziale, wenn auf diese Weise dereinst beliebige handschriftliche Quellen nicht mehr manuell transkribiert werden müssen, sondern automatisch lesbar gemacht werden können. Gegenwärtig ist eine Volltextrecherche in digitalisierten handschriftlichen Originalquellen allerdings noch nicht möglich, da das Problem der Texterkennung paläographischer Besonderheiten und individueller Schriften noch nicht befriedigend gelöst werden konnte. Die Verfahren automatischer Texterkennung verwenden auch Methoden der Kontextanalyse, durch die falsche Lesungen identifiziert und „intelligent“ korrigiert werden, indem beispielsweise irrtümlich erkannte Zahlen innerhalb von Wortumgebungen zu Buchstaben korrigiert werden. Es ist davon auszugehen, dass in künftigen Digitalisierungsprojekten sowohl Metadaten als auch die in Images erfassten und durch OCR aufbereiteten Texte durchsuchbar sein werden.⁵

Andererseits kann ein elektronischer Volltext bereits elektronisch entstanden („digital born“) sein. Dies gilt für Quellen der jüngeren Zeitgeschichte, wie elektronisch erstellte Akten oder Emailkorrespondenzen, wie auch Forschungsliteratur, insbesondere für die dem Gedanken des Open Access verpflichteten Publikationen.

Ebenso wenig eindeutig wie eine Definition von „Volltext“ ist eine Definition von „Datenbank“. Eine Datenbank besteht aus einer spezifischen Datenbasis und einer dazu gehörigen Datenverwaltung, dem Datenbankmanagementsystem (DBMS). Dessen Aufgabe besteht in der Strukturierung, dauerhaften Archivierung und Analysierbarkeit der Daten sowie der Verwaltung der Metadaten, den Informationen über die in der Datenbank enthaltenen Daten.

Der definitorischen Unschärfe im informationstechnologischen Sinn steht eine ebensolche Unschärfe im bibliothekarischen Sinn gegenüber. Eine allgemeine bibliothekarische Definition von „Volltextdatenbank“ oder „Datenbank“ existiert, beispielsweise in

⁵ Vgl. Kämmerer, Carmen, Vom Image zum Volltext – Möglichkeiten und Grenzen des Einsatzes von OCR beim Alten Buch, in: Bibliotheksdienst 43 (2009), S. 628ff.

den Regeln für die *alphabetische Katalogisierung (RAK)*⁶, bemerkenswerterweise nicht.⁷


Datenbank-Infosystem (DBIS)																															
Gesamtbestand in DBIS																															
<p>Suche nach Datenbanken</p> <p>Schnelle Suche</p> <input type="text"/> <input type="button" value="Go!"/> <p>Erweiterte Suche</p> <p>Fachübersicht</p> <p>Alphabetische Liste</p> <p>Bibliotheksauswahl / Einstellungen</p> <p>Über DBIS</p> <p>Gefördert durch:</p> 	<p>Fachübersicht</p> <table border="1"> <thead> <tr> <th>Fachgebiete</th> <th>Anzahl</th> </tr> </thead> <tbody> <tr> <td>Allgemein / Fachübergreifend</td> <td>2277</td> </tr> <tr> <td>Allgemeine und vergleichende Sprach- und Literaturwissenschaft</td> <td>523</td> </tr> <tr> <td>Anglistik, Amerikanistik</td> <td>540</td> </tr> <tr> <td>Archäologie</td> <td>218</td> </tr> <tr> <td>Architektur, Bauingenieur- und Vermessungswesen</td> <td>394</td> </tr> <tr> <td>Biologie</td> <td>640</td> </tr> <tr> <td>Chemie</td> <td>445</td> </tr> <tr> <td>Elektrotechnik, Mess- und Regelungstechnik</td> <td>169</td> </tr> <tr> <td>Energie, Umweltschutz, Kerntechnik</td> <td>314</td> </tr> <tr> <td>Ethnologie (Volks- und Völkerkunde)</td> <td>250</td> </tr> <tr> <td>Geographie</td> <td>427</td> </tr> <tr> <td>Geowissenschaften</td> <td>247</td> </tr> <tr> <td>Germanistik, Niederländische Philologie, Skandinavistik</td> <td>738</td> </tr> <tr> <td>Geschichte</td> <td>1791</td> </tr> </tbody> </table>	Fachgebiete	Anzahl	Allgemein / Fachübergreifend	2277	Allgemeine und vergleichende Sprach- und Literaturwissenschaft	523	Anglistik, Amerikanistik	540	Archäologie	218	Architektur, Bauingenieur- und Vermessungswesen	394	Biologie	640	Chemie	445	Elektrotechnik, Mess- und Regelungstechnik	169	Energie, Umweltschutz, Kerntechnik	314	Ethnologie (Volks- und Völkerkunde)	250	Geographie	427	Geowissenschaften	247	Germanistik, Niederländische Philologie, Skandinavistik	738	Geschichte	1791
Fachgebiete	Anzahl																														
Allgemein / Fachübergreifend	2277																														
Allgemeine und vergleichende Sprach- und Literaturwissenschaft	523																														
Anglistik, Amerikanistik	540																														
Archäologie	218																														
Architektur, Bauingenieur- und Vermessungswesen	394																														
Biologie	640																														
Chemie	445																														
Elektrotechnik, Mess- und Regelungstechnik	169																														
Energie, Umweltschutz, Kerntechnik	314																														
Ethnologie (Volks- und Völkerkunde)	250																														
Geographie	427																														
Geowissenschaften	247																														
Germanistik, Niederländische Philologie, Skandinavistik	738																														
Geschichte	1791																														

Abbildung 1. Datenbank-Infosystem (DBIS)

Aufschlussreicher hingegen ist die bibliothekarische Praxis am Beispiel des Datenbank-Infosystems (DBIS). So heterogen die darin verzeichneten Datenbanken generell sind, so gibt es doch formale Selektionskriterien für die Aufnahme in dieses Verzeichnis: Umfang der Datenmenge, langfristige Verfügbarkeit, kontinuierliche Pflege und Aktualisierung, Wissenschaftlichkeit, Seriosität der Inhalte und Herausgeber.⁸

Als „Volltextdatenbank“ versteht das Datenbank-Infosystem (DBIS) grundsätzlich eine „Datenbank jeglicher Art mit direkten Zugriffen auf Volltexte“⁹, die „vollständig vorliegen und (...) ggf.

⁶ <http://d-nb.info/986402338/34>

⁷ Regeln für die alphabetische Katalogisierung in wissenschaftlichen Bibliotheken RAK-WB. 2., überarbeitete und ergänzte Auflage. Leipzig, Frankfurt am Main, Berlin: Deutsche Nationalbibliothek, 2007.

⁸ Vgl. Beschreibung zur Administration von DBIS, S. 16 (Stand: 28.04.2009).
http://rzblx10.uni-regensburg.de/dbinfo/admindoc/hilfe_admin.pdf

⁹ Ebd., S. 19.

durch Abstracts, Schlagwörter etc. ergänzt“¹⁰ sind – eine außerordentlich weit gefasste Definition.¹¹

Entsprechend dieser weiten Definition von „Volltextdatenbank“ im Datenbank-Infosystem (DBIS) fallen insgesamt etwa 400 geschichtswissenschaftliche Datenbanken und Internetressourcen unter diese Kategorie. Dazu zählen nicht nur „reine“ Volltextdatenbanken im engeren Sinne der Definition, sondern auch Portale, in denen Volltexte nur einen Teil des Angebotes darstellen, oder Datenbanken, die keine Texte, sondern beispielsweise audiovisuelle Medien beinhalten. Im Hinblick auf E-Books oder E-Journals spricht man erst dann von einer Volltextdatenbank, wenn es sich nicht nur um einzelne E-Books oder E-Journals – formal gemäß der Definition Volltexte – handelt, sondern um Kollektionen von E-Books (zum Beispiel *Early English Books Online*¹²) oder E-Journals (zum Beispiel *JSTOR*¹³), die strukturiert gespeichert und formal und inhaltlich differenziert erschlossen sind.

Insofern lassen sich im Datenbank-Infosystem (DBIS) folgende Gruppen grundsätzlich unterscheiden: Handschriftenbibliotheken (zum Beispiel *Manuscripta Mediaevalia*¹⁴), Urkundensammlungen (zum Beispiel *Monasterium – Das Virtuelle Urkundenarchiv pas*¹⁵), Editionen (zum Beispiel) *Monumenta Germaniae Historica*¹⁶), Sammlungen (zum Beispiel *Nineteenth Century Collections line*¹⁷), Zeitungsdatenbanken (zum Beispiel *America's Historical*

¹⁰ Ebd.

¹¹ Die Daten müssen allerdings strukturiert vorliegen und differenzierte Suchen erlauben. Nicht aufgenommen werden Linklisten, Literaturlisten im HTML- oder PDF-Format, E-Books, E-Journals und Bibliotheks-OPACs.

¹² <http://eebo.chadwyck.com/home>

¹³ <http://www.jstor.org>

¹⁴ <http://www.manuscripta-mediaevalia.de/#|4>

¹⁵ <http://icar-us.eu/cooperation/online-portals/monasterium-net>

¹⁶ <http://www.mgh.de>

¹⁷ <http://gale.cengage.co.uk/product-highlights/history/nineteenth-century-collections-online.aspx>

*Newspapers*¹⁸), Wissenschaftliche Zeitschriften (zum Beispiel *JSTOR*¹⁹), digitale Bibliotheken (zum Beispiel *Gallica*²⁰), Virtuelle Museen (zum Beispiel *LeMO – Lebendiges Museum Online*²¹), Landeskundliche Informationssysteme (zum Beispiel *LEO-BW – Landeskunde entdecken online*²²), Film-/ Mediendatenbanken (zum Beispiel *Wochenschau-Archiv*²³).

Trotz der skizzierten definitorischen Unschärfe der Termini „Volltext“ und „Datenbank“ lässt sich eine Volltextdatenbank jedoch eindeutig von einer bibliographischen Datenbank abgrenzen: Während eine auf Standards basierende bibliographische Datenbank exakte bibliographische Angaben wie Verfasser, Titel, Erscheinungsort, Erscheinungsjahr, Standardbuchnummern sowie Abstracts umfasst und damit den Charakter einer Verweisdatenbank besitzt, ist eine Volltextdatenbank sehr viel weniger normiert als Sammlung elektronischer Volltexte zu verstehen, in der neben den bibliographischen und weiteren Metadaten die Volltexte selbst enthalten sind.²⁴ Das Fehlen der Volltexte in bibliographischen Datenbanken kann indessen vielfach dadurch kompensiert werden, dass durch den Einsatz eines Linkresolvers aus einer bibliographischen Datenbank heraus automatisch geprüft werden kann, ob die lokale Bibliothek einen Zugriff auf den gesuchten Text innerhalb einer Volltextdatenbank besitzt. So können beispielsweise in der bibliographischen Datenbank *Internationale Bibliographie der Geistes- und Sozialwissenschaftlichen Zeitschriftenliteratur (IBZ)*²⁵ die Daten eines Artikels recherchiert werden, auf den

¹⁸ <http://www.readex.com/content/americas-historical-newspapers>

¹⁹ <http://www.jstor.org>

²⁰ <http://gallica.bnf.fr>

²¹ <https://www.dhm.de/lemo>

²² <http://www.leo-bw.de>

²³ <http://www.wochenschau-archiv.de>

²⁴ Metadaten sind „Daten über Daten“, da sie Informationen über andere Daten beinhalten, beispielsweise über Einzeldokumente oder größere Datensammlungen in Datenbanken.

²⁵ <http://www.degruyter.com/databasecontent?dbid=ibz&dbsource=%2Fdb%2Fibz>

man über einen Linkresolver etwa in der Volltextdatenbank JSTOR direkt zugreifen kann.

Unter welchen Voraussetzungen lassen sich Volltextdatenbanken von HistorikerInnen effizient nutzen? Auch wenn Volltextdatenbanken in der Regel einen bequemen Zugriff auf die Inhalte bieten, so gibt es doch große Unterschiede in ihrer Erschließung. Für den Benutzer erkennbar ist häufig zuallererst die unterschiedliche Technik der Aufbereitung und Erschließung der Dokumente und Daten. Liegen Texte und Quellen in genuin elektronischer Form vor („born digital“), sind die Voraussetzungen einer Volltextrecherche durch eine Volltextindexierung am günstigsten. Werden hingegen ursprünglich gedruckt publizierte Texte erst retrospektiv durch Verfahren einer Image-Digitalisierung als elektronische Faksimilies digitalisiert und in eine Volltextdatenbank aufgenommen, so hängt der Nutzen häufig davon ab, ob und wie gut eine automatisierte Texterkennung durch OCR umgesetzt wurde. Ohne eine solche Texterkennung bezieht sich die Volltextrecherche lediglich auf die Metadaten, die die einzelnen Original-Dokumente bibliographisch beschreiben, sofern die Texte nicht nach bestimmten Standards wie zum Beispiel der Text im Dokumentenformat Encoding Initiative (TEI)²⁶ intellektuell transkribiert und in eine strukturierte Textdatei überführt werden.

Ein zentrales Qualitätskriterium für Volltextdatenbanken ist somit die Volltextindexierung der darin enthaltenen Einzeltexte, welche keineswegs die Regel darstellt. Unter Volltextindexierung versteht man die automatische Erfassung aller Wörter eines Textes mit Ausnahme von so genannten Stoppwörtern, die für das Verständnis des Dokumentinhalts keine Relevanz besitzen. Stoppwörter sind bestimmte Artikel („der“, „die“, „das“), unbestimmte Artikel („einer“, „eine“, „ein“), Konjunktionen (zum Beispiel „und“, „oder“), Präpositionen (zum Beispiel „an“, „in“, „von“) oder die Negation „nicht“.

Neben der technischen Volltextindexierung ist die bibliothekarische Beschreibung nach einem strukturierten Metadatenschema

²⁶ <http://www.tei-c.org/index.xml>

Voraussetzung für die formale Identifizierung einzelner Dokumente. Eine effiziente inhaltliche Recherche nach einzelnen Texten ist darüber hinaus nur dann möglich, wenn die in einer Volltextdatenbank enthaltenen Dokumente wissenschaftlich differenziert erschlossen sind. Diese fachliche, das heißt in der Regel mit inhaltlichen Schlagwörtern, Thesauri oder Fachklassifikationen durchgeführte Erschließung ist allerdings keineswegs Standard. Da Volltextdatenbanken in der Regel eine große Anzahl an Einzeltextrn enthalten, ist ihre differenzierte inhaltliche Erschließung jedoch grundlegend und es ist zu wünschen, dass hierbei Fachwissenschaftler und fachwissenschaftlich gebildete Bibliothekare zusammen wirken.²⁷ Die Eignung von Volltextdatenbanken für historische Recherche und Interpretationsarbeit wird daher einerseits durch die technische Aufbereitung der retrodigitalisierten oder genuin elektronischen Dokumente, andererseits durch ihre bibliothekarisch-formale und intellektuelle Erschließung entscheidend beeinflusst.

Schließlich sollte man sich mit den konkreten Funktionalitäten historischer Volltextdatenbanken vertraut machen, um sie effizient nutzen zu können. Viele Volltextdatenbanken bieten sowohl eine Suche in den die Einzeldokumente beschreibenden Metadaten als auch in den Volltexten selbst. In den meisten Fällen wird in Volltextdatenbanken wie in bibliographischen Datenbanken zwischen einer „einfachen Suche“ („Basic Search“) und „erweiterter Suche“ („Advanced Search“) unterschieden. Während in der „einfachen Suche“ entweder nur ein einfacher Suchschlitz nach dem „Google-Prinzip“ oder eine einfache Recherchemaske mit Suchfeldern für Autor, Titel, Stichwörtern oder Schlagwörtern zur Verfügung steht, ermöglicht die „erweiterte Suche“ noch differenziertere Abfragemöglichkeiten wie beispielsweise die Eingrenzung nach exaktem Publikationsdatum, Sprache oder Druckort. Häufig sind hinter den einzelnen Suchfeldern Register oder Indices hin-

²⁷ Dieses Problem trifft sogar auf umfassende Volltextarchive von Verlagen wie Elsevier oder auf JSTOR zu. Unterschiedliche Versuche, Verfahren zur automatischen Inhaltserschließung zu entwickeln, konnten bisher keine optimalen Ergebnisse erzielen.

terlegt, in denen man gezielt nach bestimmten Begriffen oder Wortformen recherchieren kann. Vielfach kann man entscheiden, ob man nur im Titel oder innerhalb des gesamten Dokumentes suchen möchte, eine Option, die insbesondere bei Aufsatzdatenbanken (zum Beispiel *JSTOR*) oder Zeitungsdatenbanken (zum Beispiel *America's Historical Newspapers*) hilfreich ist. Für differenzierte Recherchen empfiehlt sich insbesondere die „erweiterte Suche“. Grundsätzlich gilt, dass jede Volltextdatenbank im Hinblick auf die Eigenart der darin enthaltenen Dokumente spezifische Funktionalitäten aufweist. Daher sollte man sich zunächst mit diesen technischen Besonderheiten vertraut machen, um sowohl die Möglichkeiten, als auch die Grenzen einer differenzierten Recherche zu erkennen.

Ein entscheidender Komfort von Volltextdatenbanken ist schließlich die Möglichkeit der Weiterverarbeitung von Texten und Textteilen. In der Regel lässt sich der gewünschte Text für wissenschaftliche Zwecke ausdrucken oder herunterladen, sofern keine urheberrechtlichen Bedenken bestehen.

Welche konkreten Nutzungsszenarien bieten historische Volltextdatenbanken für HistorikerInnen? Statt zeitintensiver Bibliotheks- und Archivreisen, um die Originalquellen aufzusuchen, statt aufwändiger Suche nach relevanten Textstellen in Quelleneditionen oder Forschungsliteratur ermöglichen Volltextdatenbanken den unmittelbaren Zugriff auf digitalisierte Originale oder die automatische Analyse großer Textcorpora für eine zielgenaue Identifizierung relevanter Quellen und Forschungsliteratur. Beispielsweise ist denkbar, aus historischen Primärquellen größerer Zeiträume Thesauri historischer Begriffe zu generieren und ihre Semantik differenzierter und umfassender zu analysieren, als es noch in den *Geschichtlichen Grundbegriffen* in vordigitaler Zeit gelang.²⁸ Es ist aber auch denkbar, auf der Grundlage digitaler Bib-

²⁸ Brunner, Otto u.a. (Hrsg.), *Geschichtliche Grundbegriffe: Historisches Lexikon zur politisch-sozialen Sprache in Deutschland*, Stuttgart 1972–1997. Hier werden insgesamt 122 Begriffe wie Adel, Demokratie, Herrschaft, Staat in einzelnen Artikeln analysiert. Da der Fokus auf der Zeit von 1700 bis zur Gegenwart liegt, bieten sich künftig entsprechende Analysen für Antike und Mittelalter an. Da in den Literaturwissenschaften Volltextdatenbanken ebenfalls eine zentrale Rolle

liotheken und intelligent vernetzter Volltextdatenbanken quantitative Analysen historischer Kulturen und Gesellschaften in einem Umfang vorzunehmen, der nicht mehr durch einen einzelnen Wissenschaftler, sondern nur durch computergestützte kollaborative Forschung zu leisten ist. Innerhalb des historischen Arbeitsprozesses steht dadurch mehr Zeit für die interpretatorische Arbeit auf der Grundlage sehr viel schneller als zuvor aufgefundenen Quellen und Texte zur Verfügung. Zugleich ist der Gefahr einer Überforderung zu begegnen, die durch die schiere Masse an verfügbarer wissenschaftlicher Information konzentriertes und fokussiertes interpretatorisches Arbeiten gefährden kann: Geschichtswissenschaftliche Fachkompetenz in Verbindung mit Informations- und Selektionskompetenz hingegen ermöglicht effizientes historisches Arbeiten in der digitalen Welt.



Abbildung 2. Aufbereitung und Erschließung von Texten in Volltextdatenbanken (vereinfacht)

spielen, lassen sich Analysen historischer Quellen vor dem Hintergrund literarischer Texte der jeweiligen Epoche durchführen.

1.3 Institutionelle Infrastrukturen

Historische Volltextdatenbanken können aus ganz unterschiedlichen Kontexten heraus entstanden sein: Sei es, dass Bibliotheken, Archive, Museen oder Forschungseinrichtungen Quellen (zum Beispiel Handschriften, Urkunden) oder urheberrechtsfreie Forschungsliteratur aus ihren eigenen Beständen digitalisieren und in Datenbanken oder digitalen Bibliotheken publizieren, sei es, dass Wissenschaftsverlage (zum Beispiel *Brepols*, *ProQuest*, *Gale Cengage*) von ihnen bereits verlegte Titel nach der gedruckten Publikation auch elektronisch veröffentlichen. Konzeption und Publikation von Volltextdatenbanken hängt also insbesondere mit öffentlichen Einrichtungen des Wissenschaftsbereiches oder kommerziellen Verlagen zusammen. Der grundlegende Unterschied liegt darin, dass öffentliche Einrichtungen die von ihnen erstellten Volltextdatenbanken meist kostenfrei publizieren, während Verlage für deren Nutzung Kauf- und Lizenzierungsmodelle anbieten.

Infolgedessen entstehen einerseits Strukturen staatlicher Wissenschaftsinstitutionen (Bibliotheken, Archive, Museen, Forschungsinstitute), die in Volltextdatenbanken nicht nur bibliographische Daten bereitstellen, sondern auch die Digitalisate der Texte frei anbieten. Andererseits finanzieren öffentliche Einrichtungen der Wissenschaft zugleich den Zugang zu kommerziellen elektronischen Produkten, zum Teil in länderübergreifenden Konsortien, für ihre jeweiligen NutzerInnen. Beispielsweise lizenzieren Universitäten und Forschungseinrichtungen ausschließlich für ihre eigenen Forschenden, Lehrenden und Studierenden historische Volltextdatenbanken. Landes- und Staatsbibliotheken dagegen lizenzieren diese Produkte für eine breitere, wissenschaftlich interessierte Öffentlichkeit, die nicht an Universitäten und Forschungseinrichtungen angebunden ist. Den weitaus größten Nutzen für wissenschaftliche und private Nutzer in Deutschland hat das Konzept der *Nationallizenzen*²⁹ gestiftet. Die Deutsche Forschungsgemeinschaft (DFG) finanzierte von 2004 bis 2010 den Erwerb von Lizenzen, um Studierenden, Wissenschaftlern und

²⁹ <https://www.nationallizenzen.de>

der wissenschaftlich interessierten Öffentlichkeit den freien Zugriff auf kostenpflichtige elektronische Verlagsprodukte zu ermöglichen. Unter diesen spielen neben bibliographischen Datenbanken, Elektronischen Zeitschriften und Nachschlagewerken insbesondere Volltextdatenbanken eine zentrale Rolle. Von den im Datenbank-Infosystem (DBIS) verzeichneten 153 Volltextdatenbanken aus der Kollektion der Nationallizenzen für alle Fächer betreffen allein 57 Produkte die Geschichtswissenschaften (Stand: August 2015).

Neben den ausschließlich von der Deutschen Forschungsgemeinschaft (DFG) finanzierten Nationallizenzen für abgeschlossene Produkte gibt es DFG-geförderte Allianz-Lizenzen für laufende Produkte und Allianz-Lizenzen für laufende Produkte ohne DFG-Förderung.

Am Beispiel dieser unterschiedlichen Formen der Lizenzierungen im Vergleich zu den von Bibliotheken, Archiven und Forschungseinrichtungen selbst erstellten, frei nutzbaren Volltextdatenbanken wird deutlich, dass die Zugriffsrestriktionen bei kommerziellen Produkten die Forschungsmöglichkeiten derjenigen Nutzer, die keine institutionelle Anbindung besitzen, deutlich einschränken.

2. Beispiele historischer Volltextdatenbanken

Die Geschichtswissenschaft gewinnt ihre Erkenntnisse einerseits aus den Quellen, die je nach historischer Epoche ganz unterschiedlich sein können, andererseits aus wissenschaftlichen Texten, die traditionell in Büchern und Zeitschriften publiziert worden sind. Bernhard Fabian hatte in einem prägnanten Bild die Rolle von Bibliotheken und Laboren für unterschiedliche Wissenschaftsfächer und -kulturen miteinander verglichen.³⁰ Seiner Auf-

³⁰ Fabian, Bernhard, *Buch, Bibliothek und geisteswissenschaftliche Forschung. Zu Problemen der Literaturversorgung und Literaturproduktion in der Bundesrepublik Deutschland*, Göttingen 1983. Die Ideen Fabians haben im deutschen Bibliothekswesen eine Entwicklung angestoßen, die die geisteswissenschaftliche Forschung inzwischen auf eine neue Grundlage stellen konnte: Mit der Gründung der *Arbeitsgemeinschaft Sammlung Deutscher Drucke (AG SDD)* 1989 zum Aufbau einer verteilten deutschen Nationalbibliothek begann die systematische

fassung nach verkörpern Bibliotheken für Geisteswissenschaftler das, was Labore für Naturwissenschaftler bedeuten. Da das wissenschaftliche Material, das Bibliotheken für die Geisteswissenschaften bereitstellen, primär aus Texten besteht, liegt eine grundlegende Herausforderung historischen Arbeitens darin, mit diesen Texten professionell umzugehen, sie zu lesen, zu analysieren und zu interpretieren. Historische Volltextdatenbanken, die digitalisierte Einzelquellen (zum Beispiel Handschriften, Urkunden, Münzen) enthalten, ermöglichen heutigen HistorikerInnen in bisher nicht gekannter Weise Analysen, die sich auf repräsentatives Material gründen und computergestützt effizient durchführbar sind. Ebenso verhält es sich bei Volltextdatenbanken, die Quellencorpora (zum Beispiel *Monumenta Germaniae Historica*³¹) enthalten, oder bei Volltextdatenbanken für Forschungsliteratur, seien es wissenschaftliche Zeitschriften (zum Beispiel *JSTOR*³²), seien es wissenschaftliche Monographien (zum Beispiel *OstDok*³³).

Nachdem im Abschnitt 1.2 erörtert wurde, auf welche Weise sich Volltextdatenbanken von HistorikerInnen nutzen lassen, werden im Folgenden herausragende Volltextdatenbanken für die Geschichtswissenschaft vorgestellt. Allein angesichts ihrer beständig steigenden Menge ist lediglich eine repräsentative Auswahl möglich. Daher liegt der Fokus auf exemplarischen Datenbanken für wichtige historische Epochen und Regionen. Es werden dabei sowohl freie als auch zugriffsbeschränkte Angebote vorgestellt.

Das zentrale Nachweisinstrument für die an 293 wissenschaftlichen Bibliotheken mittlerweile verfügbaren 11.106 Datenbanken,

retrospektive Erwerbung des gedruckten deutschen Kulturerbes durch sechs für einzelne Epochen ausgewiesene Bibliotheken Deutschlands. Darauf aufbauend werden die wichtigsten historischen Texte digitalisiert und in Portalen wie dem *Zentralen Verzeichnis digitalisierter Drucke (zvd)* sowie in den Datenbanken *Verzeichnis der Drucke des 16. Jahrhunderts (VD16)*, *Verzeichnis der Drucke des 17. Jahrhunderts (VD17)* und *Verzeichnis der Drucke des 18. Jahrhunderts (VD18)* publiziert (vgl. Kapitel 2.3 in diesem Guide).

³¹ <http://www.dmgh.de>

³² <http://www.jstor.org>

³³ <https://www.vifaost.de/ostdok>

von denen 4.488 frei im Internet zur Verfügung stehen (Stand: Dezember 2014), ist das bereits zuvor besprochene *Datenbank-Infosystem (DBIS)*³⁴. Neben einem Überblick über den Gesamtbestand sowie das Fächerspektrum, informiert die lokale Sicht über das jeweilige Angebot der Bibliothek vor Ort. Die eminente Bedeutung von Datenbanken für die geschichtswissenschaftliche Arbeit wird bereits daran deutlich, dass aus dem DBIS-Gesamtbestand etwa 1.800 Datenbanken dem Fach Geschichte zugeordnet sind. Neben Volltextdatenbanken gehören im engeren Sinn dazu: Fachbibliographien, Aufsatzdatenbanken, Bilddatenbanken, biographische Datenbanken, Wörterbücher, Enzyklopädien, Nachschlagewerke. Die im Folgenden besprochenen Volltextdatenbanken sowohl für Quellen und Quelleneditionen als auch für Forschungsliteratur (Monographien, Zeitschriftenartikel) sind einerseits im Datenbank-Infosystem (DBIS) verzeichnet. Andererseits bieten vielfach auch die Virtuellen Fachbibliotheken³⁵, die innerhalb der Geschichtswissenschaft eine wichtige Rolle spielen, Zugriff auf relevante Volltexte.

2.1 Große, fachlich übergreifende Institutionen

Die Geschichtswissenschaften in Deutschland werden durch vielfältige staatliche und nichtstaatliche Institutionen getragen. Häufig sind diese Institutionen zugleich Anbieter maßgeblicher Volltextdatenbanken – insbesondere, wenn diese auf der Digitalisierung eigener Sammlungen beruhen. Stellvertretend für fachlich übergreifende öffentliche Einrichtungen sollen hier zunächst die beiden zentralen historischen Museen der Bundesrepublik Deutschland gewürdigt werden – das *Deutsche Historische Museum*³⁶ (Berlin) und das *Haus der Geschichte*³⁷ (Bonn). Neben ihren jeweiligen Dauerausstellungen aus eigenen Beständen präsentie-

³⁴ http://rzblx10.uni-regensburg.de/dbinfo/index.php?bib_id=alle&colors=3&ocolors=40&ref=about#inh

³⁵ Eine Übersicht über die Virtuellen Fachbibliotheken bietet der Gemeinsame Bibliotheksverbund: <https://www.gbv.de/bibliotheken/vifa-olc-ssg>.

³⁶ <https://www.dhm.de>

³⁷ <https://www.hdg.de/stiftung>

ren beide Institutionen zahlreiche Sonderausstellungen zu aktuellen Themen der deutschen Geschichte seit dem 19. Jahrhundert.

Eine diesen Museen vergleichbare Bedeutung für das deutsche Archivwesen hat das *Bundesarchiv*³⁸, die für Sicherung und Auswertung des Archivgutes der Bundesrepublik Deutschland und seiner Vorgängerstaaten zuständige Behörde. Neben seinem zentralen Rechercheinstrument *invenio*, dem *digitalen Bildarchiv* und dem *digitalen Filmportal* bietet das Bundesarchiv zwei wichtige historische Volltextdatenbanken: die digitale Editionen *Akten der Reichskanzlei Weimarer Republik*³⁹ und die *Kabinettsprotokolle der Bundesregierung*⁴⁰.

Ein gemeinsames Projekt des Bundesarchivs mit dem Deutschen Historischen Museum und dem Haus der Geschichte ist das *Lebendige Museum Online (LeMO)*⁴¹. Hier werden (Volltext-) Datenbanken musealer Objekte mit Texten, Karten, Statistiken, Bildern sowie Film- und Tondokumenten miteinander verknüpft und in einem attraktiven Portal der Öffentlichkeit präsentiert.

Fachlich übergreifende Institutionen außerhalb Deutschlands werden in den regionalen Guides besprochen.

2.2 Portale und digitale Bibliotheken

Für die Geschichtswissenschaften spielen Kultur- und Wissenschaftsportale sowie digitale Bibliotheken, die das kulturelle und wissenschaftliche Erbe zugänglich machen, eine wichtige Rolle. Eine digitale Bibliothek verfolgt grundsätzlich dieselben Ziele wie eine traditionelle Bibliothek mit gedruckten Texten: Sammlung, Organisation, Strukturierung, Erschließung und Archivierung von Texten und Inhalten. Digitale Bibliotheken beinhalten nicht nur Metadaten zu Texten, Bildern, Objekten und sonstigen Informationsressourcen, sondern auch die dazu gehörigen Volltexte, Bild- und audiovisuelle Daten in digitaler Form.

³⁸ <http://www.bundesarchiv.de/index.html>

³⁹ <http://www.bundesarchiv.de/aktenreichskanzlei/1919-1933/0000/index.html>

⁴⁰ <http://www.bundesarchiv.de/cocoon/barch/0000/index.html>

⁴¹ <https://www.dhm.de/lemo>

Digitale Bibliotheken können zugleich auch Virtuelle Bibliotheken sein, die über keine eigenen Bestände verfügen, sondern als elektronisches Informationssystem Metadaten und Volltexte aus unterschiedlichen Quellen unter einer einheitlichen Oberfläche zusammenführen und recherchierbar machen. Gemeinsam mit digitalen Archiven und digitalen Museen sind sie für die Dokumentation und Erhaltung des digitalisierten wissenschaftlichen und kulturellen Erbes verantwortlich – und dieses geschieht häufig durch Erstellung von Volltextdatenbanken.

Die *Europeana Collections*⁴², das europäische Kultur- und Wissenschaftsportal, setzt auf den jeweiligen nationalen digitalen Bibliotheken auf. Da die Erschließung der heterogenen Quellen aus den unterschiedlichsten internationalen Institutionen jedoch nicht immer einheitlich ist, empfiehlt sich für qualitativ hochwertige Ergebnisse eine differenzierte Suche in den Quelldatenbanken selbst.

An der Stiftung Preußischer Kulturbesitz ist die *Deutsche Digitale Bibliothek*⁴³ angesiedelt, die als nationales Portal die digital verfügbaren Angebote der Kultur- und Wissenschaftseinrichtungen Deutschlands frei zugänglich zur Verfügung stellt. Dazu gehören Bücher, Archivalien, Bilder, Skulpturen, Tondokumente, Filme, Noten. Wie bei Europeana Collections stehen innerhalb dieses Kultur- und Wissenschaftsportals neben bibliographischen Informationen auch Volltexte zur Verfügung, die jedoch nicht immer mit OCR-Erkennung aufbereitet und somit durchsuchbar sind.

Neben dieser nationalen Sicht entstehen in Deutschland gegenwärtig auch auf Länderebene Kulturportale:

- Baden-Württemberg (*LEO-BW – Landeskunde entdecken online*⁴⁴)
- Bayern (*BLO – Bayerische Landesbibliothek Online*⁴⁵, *Bavarrikon – Kultur und Wissensschätze Bayerns*⁴⁶)

⁴² <http://www.europeana.eu/portal>

⁴³ <https://www.deutsche-digitale-bibliothek.de>

⁴⁴ <http://www.leo-bw.de>

- Hessen (LAGIS – Landesgeschichtliches *tem*⁴⁷)
- Niedersachsen (*Kulturerbe Niedersachsen*⁴⁸)
- Sachsen (*Sachsen.digital*⁴⁹).

Wie in Deutschland entstehen auch in anderen Ländern übergreifende Portale und spezifische Angebote der Nationalbibliotheken, die für die eigene nationale Geschichte relevante Quellen und Forschungsliteratur digitalisieren und in Volltextdatenbanken veröffentlichen. Für die historische Forschung lassen sich hier zahlreiche Entdeckungen machen. Selbst wenn viele dieser Bestände auch in übernationalen digitalen Bibliotheken wie der *Bibliotheca Europeana*⁵⁰ nachgewiesen sind, lohnt es sich, auch auf den jeweiligen Webseiten der Nationalbibliotheken zu recherchieren.

Nach dem Vorbild der Europeana Collections wird in den Vereinigten Staaten von Amerika die *Digital Public Library of America (DPLA)*⁵¹ mit mehreren Millionen Digitalisaten von Handschriften, Kunstwerken und Fotos aus zahlreichen Kultur- und Wissenschaftsinstitutionen der USA aufgebaut.

Von den maßgeblichen US-amerikanischen Bibliotheken hat die *Library of Congress (LoC)*⁵² in Washington für ihr Portal *American Memory*⁵³ zentrale Dokumente ihrer Americana-Sammlung digitalisiert: Dazu gehören Handschriften, seltene Drucke und Bücher, Karten, Noten, Ton- und Filmdokumente.

⁴⁵ <http://www.bayerische-landesbibliothek-online.de>

⁴⁶ <http://www.bavarikon.de>

⁴⁷ <http://www.lagis-hessen.de>

⁴⁸ <http://kulturerbe.niedersachsen.de/viewer/start>

⁴⁹ <http://www.sachsendigital.de/startseite>

⁵⁰ <http://www.europeana.eu/portal>

⁵¹ <http://dp.la>

⁵² <http://www.loc.gov>

⁵³ <http://memory.loc.gov/ammem/index.html>

Für Frankreich hat die *Bibliothèque nationale de France (BnF)*⁵⁴ in Paris mit dem Projekt *Gallica*⁵⁵ eine eigene digitale Bibliothek aufgebaut, die mehr als 2 Millionen Bücher, Karten, Handschriften, Bilder, Noten, Tonaufnahmen und mehr als eine Million Zeitschriften- und Zeitungsausgaben umfasst, die teilweise mit OCR bearbeitet wurden und somit im Volltext durchsucht werden können.

Die *British Library*⁵⁶ in London hat für ihr Projekt *Collect Britain*⁵⁷ etwa 100.000 Bilder, Fotos, Karten und Tondokumente zur britischen Regional- und Kulturgeschichte aus ihren Beständen digitalisiert, differenziert erschlossen und recherchierbar gemacht. Diese kontinuierlich wachsende digitale Bibliothek bietet einen guten Einstieg in die Recherche nach Quellen zur englischen Geschichte.

Neben diesen regionalen, nationalen und übernationalen Portalen und digitalen Bibliotheken gibt es weitere Sammlungen von Volltexten, die für historische Forschungen von Interesse sein können. Als älteste digitale Bibliothek gilt das bereits 1971 begründete *Project Gutenberg*⁵⁸. Es hat zum Ziel, urheberrechtsfreie Bücher im Internet frei zugänglich zu machen. Getragen von Freiwilligen wurden ausgewählte Bücher zunächst manuell abgetippt. Die kontinuierliche Verbesserung der Scan-Technik und die Entwicklung von Texterkennungsoftware ermöglichte eine immer schnellere Veröffentlichung dieser Volltexte. Die im Portal inzwischen etwa 50.000 verfügbaren Bücher, Filme, Bilder und Audiodateien können online gelesen oder in unterschiedlichen Formaten herunter geladen werden. Da im *Project Gutenberg* vorwiegend englischsprachige Bücher enthalten sind, bietet das *Projekt Gutenberg-DE*⁵⁹ urheberrechtsfreie deutschsprachige E-Books. Auch hier werden die gescannten Printeditionen von einer

⁵⁴ <http://www.bnf.fr/fr/acc/x.accueil.html>

⁵⁵ <http://gallica.bnf.fr>

⁵⁶ <http://www.bl.uk>

⁵⁷ <http://www.bl.uk/onlinegallery/index.html>

⁵⁸ <http://www.gutenberg.org>

⁵⁹ <http://www.projekt.gutenberg.de>

*Community*⁶⁰ kollaborativ korrigiert. Im Gegensatz zum Project Gutenberg ist das Herunterladen vollständiger Texte im Projekt Gutenberg-DE jedoch nur eingeschränkt möglich.

Von den kommerziellen Anbietern digitaler Bibliotheken ist *Google Books*⁶¹ sicher am bekanntesten. Enthalten sind retrodigitalisierte Bücher, die der US-Konzern systematisch gescannt hat. Da die Digitalisate nach Möglichkeit mit OCR bearbeitet wurden, besteht die komfortable Möglichkeit der Volltextsuche. Problematisch ist Google Books jedoch aus rechtlicher Sicht sowie der daraus resultierenden unterschiedlichen Zugriffsbeschränkung auf die im Rahmen von Verlagskooperationen digitalisierten Bücher. Urheberrechtlich geschützte Bücher können nur in kleinen Teilen (häufig die Inhaltsverzeichnisse und Indices) eingesehen werden, bei kostenfreier Registrierung sind weitere Seiten zugänglich. Große Teile sind jedoch grundsätzlich gesperrt, auch ist nach Ausschöpfung eines bestimmten Tageskontingents kein weiterer Zugriff auf Volltexte möglich. Für sein Projekt kooperiert Google Books mit namhaften wissenschaftlichen Bibliotheken in den USA (die Bibliotheken der Universitäten von Michigan, Harvard, Stanford, Virginia, Wisconsin-Madison, Princeton, California, Austin sowie die New York Public Library), Deutschland (Bayerische Staatsbibliothek), England (Bodleian Library, Oxford), Frankreich (Bibliothèque Municipale de Lyon), Österreich (Österreichische Nationalbibliothek), Spanien (Universitätsbibliothek Madrid, Nationalbibliothek von Katalonien).

Unter dem Aspekt der Sichtbarkeit des globalen kulturellen und wissenschaftlichen gedruckten Erbes innerhalb digitaler Bibliotheken ist zu beachten, dass eine Recherche mit der Suchmaschine Google unter Berücksichtigung von Google Books ein Ranking der Ergebnisse zeigt, in dem englischsprachige Literatur häufig dominiert. Daher empfiehlt sich für geschichtswissenschaftliche

⁶⁰ <http://www.gaga.net/pgdp/default.php>

⁶¹ <https://books.google.de/?hl=de>

Arbeit die gleichzeitige Konsultation der *Europeana Collections*⁶² oder der *Deutschen Digitalen Bibliothek*⁶³.

Eine grundsätzliche Problematik nicht nur der geschichtswissenschaftlichen Recherche innerhalb von *Google Books* liegt in der fehlenden fachlichen Selektionsmöglichkeit der digitalisierten Bücher sowie in deren mangelhafter intellektueller Erschließung durch Verzicht auf Anwendung von Fachklassifikationen bzw. Fachthesauri oder Verschlagwortung – Services, die Bibliothekskataloge und Fachdatenbanken in hoher Qualität bieten. Die Recherche innerhalb von *Google Books* muss daher mit einer Stichwortsuche auskommen, wobei die gewählten Stichworte in unterschiedlichen Sprachen formuliert werden müssen, um jeweils Zugriff auf internationale Forschungsliteratur zu erhalten. Auch die Metadaten sind mit genuin bibliothekarischen Angeboten nicht vergleichbar. Insofern bietet *Google Books* zwar eine interessante Ergänzung zu geschichtswissenschaftlichen Volltextdatenbanken, doch sollte man sich der Grenzen durchaus bewusst sein.

Nicht nur ein Portal für Volltexte, sondern zugleich ein Dienst, der Webseiten in unterschiedlichen Versionen speichert, ist das 1996 gegründete gemeinnützige *Internet Archive*⁶⁴. Obwohl es nicht möglich ist, das Internet insgesamt in allen Zuständen dauerhaft zu archivieren, bietet die Speicherung von „Momentaufnahmen“ dennoch einen breiten Zugriff auf Webseiten, die einmal im Internet verfügbar waren. Dieser Zugriff erfolgt mit Hilfe der *Wayback Machine*⁶⁵. Das Volumen des *Internet Archives* ist im Jahr 2015 auf etwa 500 Milliarden Webseiten angewachsen. Seine digitale Bibliothek umfasst inzwischen Millionen Bücher, Musik, Filme, Software, Konzertmitschnitte. Innerhalb des *Million Book Project*⁶⁶

⁶² <http://www.europeana.eu/portal>

⁶³ <https://www.deutsche-digitale-bibliothek.de>

⁶⁴ <https://archive.org>

⁶⁵ <http://archive.org/web>

⁶⁶ <https://archive.org/details/millionbooks>

werden die inzwischen gemeinfrei gewordenen Bücher digitalisiert und als Teil der *Open Library*⁶⁷ publiziert.

Die Open Library hat zum Ziel, kollaborativ jedes in einer digitalen Bibliothek bisher veröffentlichte Buch auf einer eigenen Webseite zu dokumentieren. Dabei kann zum gemeinfreien bibliographischen Nachweis auch der direkte Zugriff auf das Digitalisat mit dem Volltext treten. Durch diese Vision der mit etwa 20 Millionen öffentlich zugänglichen Datensätzen weltweit größten digitalen Bibliothek verfolgt die Open Library ähnliche Ziele wie Google Books: Während das Internet Archive und die Open Library ihren Fokus auf gemeinfreie Bücher legen, digitalisiert Google Books auch urheberrechtlich geschützte Literatur.

Ein gemeinsames Projekt von Google Books, Internet Archive und zahlreichen Bibliotheken und Forschungseinrichtungen aus der ganzen Welt mit dem Ziel, ein gemeinsames Repositorium zu errichten, um digitalisierte Inhalte für die Zukunft zu archivieren, ist die *HathiTrust Digital Library*⁶⁸. Der aus dem Hindi und Urdu stammende Name „Hathi „Elephant“, ein Tier, dem ein besonderes Gedächtnis nachgesagt wird. HathiTrust umfasst inzwischen mehr als 10 Millionen digitalisierte und im Volltext recherchierbare Dokumente.

Eine kleinere deutsche geisteswissenschaftliche Volltextbibliothek ist schließlich *Zeno.org*⁶⁹. Sie basiert auf den zwischen 1997 und 2007 von der Firma Directmedia Publishing veröffentlichten CD's und DVD's der Reihe *Digitale Bibliothek*. Seit 2009 stehen die Texte von Zeno.org als digitale Bibliothek im Repository der Virtuellen Forschungsumgebung für die Geisteswissenschaften *TextGrid*⁷⁰ zur Lektüre und weiteren Bearbeitung in digitalen Editionen zur Verfügung.

⁶⁷ <https://openlibrary.org>

⁶⁸ <https://www.hathitrust.org/home>

⁶⁹ <http://www.zeno.org>

⁷⁰ <https://textgrid.de/digitale-bibliothek>

2.3 Volltextdatenbanken für Quellen und Quelleneditionen

Volltextdatenbanken ermöglichen in besonderer Weise einen quellenorientierten Zugang zum geschichtswissenschaftlichen Arbeiten. Im Folgenden sollen für einzelne historische Epochen, Regionen oder Spezialthemen aus der stetig wachsenden Fülle neuer Angebote einige Beispiele ausgewählt werden. Dabei werden sowohl freie als auch lizenzpflichtige Angebote berücksichtigt.

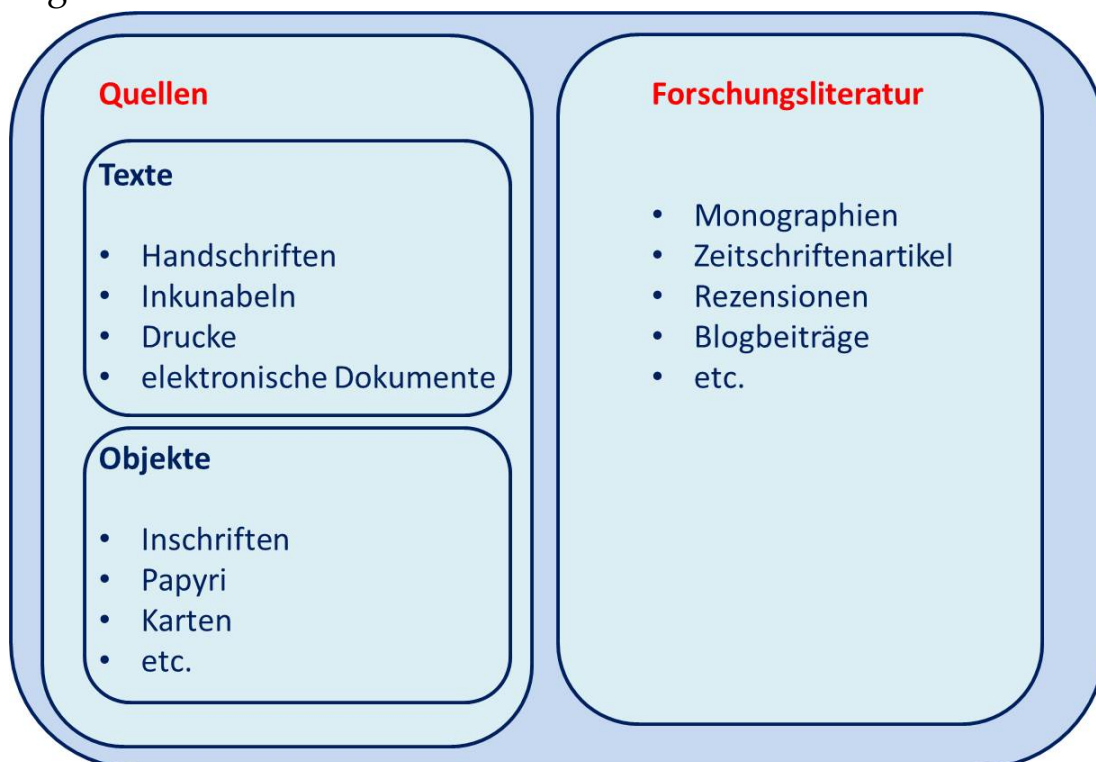


Abbildung 3. Volltextdatenbanken für Quellen und für Forschungsliteratur

Alte Geschichte

Für die Alte Geschichte spielt die schriftliche Überlieferung der Antike naturgemäß eine zentrale Rolle. *Perseus Digital Library*⁷¹ ist eine der ältesten online verfügbaren Textsammlungen für Geisteswissenschaften, die einen besonderen Schwerpunkt auf die antike Überlieferung legt. In der Kollektion *Greek and Roman Materi-*

⁷¹ <http://www.perseus.tufts.edu/hopper>

*als*⁷² bietet sie auf der Grundlage zitierfähiger Editionen mehr als 400 Volltexte klassischer griechischer und lateinischer Literatur, teilweise mit (englischer) Übersetzung. Insbesondere Analysen historischer Begrifflichkeiten oder philologische Fragestellungen zum Wortgebrauch lassen sich mit den Recherchemöglichkeiten in der Perseus Digital Library effizient durchführen.

Die Volltextdatenbank *Thesaurus Linguae Graecae (TLG)*⁷³ enthält klassische griechische Texte aus der Zeit vom 8. Jh. v. Chr. bis 600 n. Chr. sowie historiographische, lexikographische und scholastische griechische Texte des Mittelalters von 600 n. Chr. bis 1453 n. Chr. Die differenzierte Suchmaske erlaubt komplexe Recherchen nach Autor, Werk, Datierung, Gattung, aber auch eine Volltextsuche innerhalb der Werke eines oder mehrerer Autoren. Die Darstellung der Texte erfolgt wahlweise in griechischen oder transliteriert in lateinische Buchstaben.

Die Volltextdatenbank *Patrologia Graeca*⁷⁴ basiert auf der *Patrologia Graeco-Latina*, die von Jacques-Paul Migne zwischen 1857 und 1866 in 161 Bänden herausgegeben wurde.⁷⁵ Sie enthält die christlich-griechische Kirchenliteratur mit den wichtigsten Werken spätantiker und mittelalterlicher Theologie, Philosophie und Geschichte aus der Zeit von 100 n. Chr. bis 1478 und ist grundlegend für historische Forschungen zum frühen Christentum. Die Volltexterschließung mit unterschiedlichen Rechercheoptionen bietet einen deutlichen Mehrwert der Datenbank im Vergleich zur gedruckten Edition. Die Texte können sowohl ausgedruckt als auch heruntergeladen werden.

Im Zusammenhang mit der *Patrologia Graeca* steht die *Patrologia Latina*, die Edition des lateinischen Schrifttums der Kirche von

⁷² <http://www.perseus.tufts.edu/hopper/collection%3Fcollection%3DPerseus:collection:Greco-Roman>

⁷³ <https://stephanus.tlg.uci.edu/index.php>

⁷⁴ http://rzblx10.uni-regensburg.de/dbinfo/detail.php?bib_id=alle&colors=&ocolors=&lett=fs&tid=0&titel_id=4002

⁷⁵ [Patrologia cursus completus / Series Graeca] Patrologiae cursus completus: in qua prodeunt patres, doctores scriptoresque ecclesiae Graecae [...] accurante J. P. Migne. Paris, 1857–1866.

den Anfängen bis ins Hochmittelalter in insgesamt 217 Bänden, die Jacques-Paul Migne in zwei Reihen zwischen 1844 und 1855 publiziert hat.⁷⁶ Die elektronische Volltextausgabe *Patrologia Latina Database*⁷⁷ auf CD-ROM und Online-Datenbank beruht auf dieser Erstausgabe Mignes. Texte der *Patrologia Latina* sind auch Bestandteil der *Library of Latin Texts*.

Die *Library of Latin Texts – Series A (LLT-A)*⁷⁸ ist hinsichtlich der Breite der darin enthaltenen lateinischen Texte umfassend, indem sie von der Antike bis ins 20. Jahrhundert hinein reicht. Sie bietet hoch differenzierte Möglichkeiten der Analyse der Texte und der darin vorkommenden Wortformen, eignet sich daher ausgezeichnet für historische und philologische Fragestellungen. Die *Library of Latin Texts – Series B (LLT-B)*⁷⁹ ergänzt als Volltextdatenbank *LLT-A* mit dem Fokus auf die lateinische Literatur vom 2. bis zum 16. Jahrhundert.

Neben Texten sind für althistorische Forschungen auch andere Quellen zentral, die ebenfalls in – im Sinne des hier verwendeten weiten Verständnisses – Volltextdatenbanken zur Verfügung stehen. Als Beispiel sei die *Epigraphische Datenbank Heidelberg (EDH)*⁸⁰ gewählt. Sie hat das Ziel einer umfassenden Dokumentation der lateinischen und bilinguen (v.a. lateinisch-griechischen) Inschriften des Römischen Reiches. Unterschiedliche Sucheinstiege über zwei komplexe Suchmasken ermöglichen eine differenzierte Recherche innerhalb des umfangreichen Materials – insbesondere innerhalb der Inschriftentexte.

⁷⁶ [Patrologia cursus completus / Series Latina] Patrologiae cursus completus: seu bibliotheca universalis, integra, uniformis, commoda, oeconomica, omnium ss. Patrum [...] ecclesiasticorum, sive Latinorum, sive Graecorum [...] accurante J. P. Migne. Paris, 1844–1855.

⁷⁷ http://rzblx10.uni-regensburg.de/dbinfo/detail.php?bib_id=alle&colors=&ocolors=&lett=fs&tid=0&titel_id=1795

⁷⁸ <http://www.brepols.net/Pages/BrowseBySeries.aspx?TreeSeries=LLT-O>

⁷⁹ http://rzblx10.uni-regensburg.de/dbinfo/detail.php?bib_id=alle&colors=&ocolors=&lett=fs&tid=0&titel_id=9698

⁸⁰ <http://edh-www.adw.uni-heidelberg.de/home>

Geschichte des Mittelalters

Wie für die Alte Geschichte bilden auch für die Geschichte des Mittelalters einige klassische Printeditionen die Grundlage von Volltextdatenbanken und bieten komfortable Möglichkeiten des Zugriffs auf die Quellen durch Volltextrecherche. Unentbehrliche Hilfsmittel sind die *Monumenta Germaniae Historica (MGH)*⁸¹, die grundlegende Sammlung von Quelleneditionen (zum Beispiel Historiographie, Rechtstexte, Urkunden, Briefe, Dichtung) zur mittelalterlichen Geschichte aus dem Zeitraum 500 n. Chr. – 1500 n. Chr., die seit 1819 in mehr als 300 Bänden erschienen ist. Seit 2004 werden sämtliche gedruckten Editionen mit einer Moving wall von drei Jahren online als Teil der Volltextsammlung publiziert. Exzellente Suchoptionen ermöglichen einen umfassenden Zugriff auf diese wichtigste Quellenedition der mittelalterlichen Geschichte Deutschlands und Europas.

Ursprünglich als Vorarbeit zu den *Monumenta Germaniae Historica (MGH)* konzipiert, entstand mit den *Regesta Imperii (RI)*⁸² ein weiteres Monument der Wissenschaftsgeschichte, das seit seiner Begründung 1839 für Mittelalter HistorikerInnen unverzichtbar ist. Chronologisch geordnet verzeichnet es sämtliche urkundlichen und historiographischen Quellen einerseits der römisch-deutschen Herrscher von den Karolingern bis zu Maximilian I. (751 – 1519), andererseits der Päpste des frühen und hohen Mittelalters. Die Quellen werden dabei nicht historisch-kritisch ediert, sondern es wird lediglich ihr wesentlicher Inhalt knapp zusammengefasst. Nicht nur sämtliche gedruckten Regestenbände wurden digitalisiert, sondern ihre Inhalte in die Regestendatenbank übernommen. Als Volltextdatenbank bietet diese nicht nur differenzierte Rechercheoptionen, sondern darüber hinaus intelligente Vernetzungen mit der *Regesta Imperii*

⁸¹ <http://www.dmgh.de>

⁸² <http://www.regesta-imperii.de/startseite.html>

*Literaturdatenbank (RI-OPAC)*⁸³ oder dem *Lichtbildarchiv Marburg (LBA)*⁸⁴.

Doch nicht nur große Monumente der Wissenschaftsgeschichte wie die Editionen mittelalterlicher Quellen im Rahmen der *MGH* und der *RI* wurden systematisch digitalisiert, sondern auch die Quellen selbst. Für den Bereich der Handschriften ist für den deutschen Kulturraum *Manuscripta Mediaevalia*⁸⁵ das zentrale Portal für Mediävisten. Herausgegeben von der Staatsbibliothek zu Berlin – Preußischer Kulturbesitz, der Bayerischen Staatsbibliothek München sowie dem Deutschen Dokumentationszentrum für Kunstgeschichte – Bildarchiv Foto Marburg werden hier digitalisierte Handschriften und Handschriftenkataloge aus Bibliotheken Deutschlands, Österreichs und der Schweiz publiziert. Ziel ist es, in *Manuscripta Mediaevalia die gesamte handschriftliche Überlieferung in Deutschland*⁸⁶, die mit insgesamt 60.000 Handschriften beziffert wird, im Volltext digital bereit zu stellen. Unter den differenzierten Recherchemöglichkeiten ist besonders interessant der Index ikonographischer Begriffe zur systematischen Analyse der zahlreichen Bildhandschriften.

Auch Österreich besaß eine reiche mittelalterliche Klosterkultur und damit vielfältige Handschriftenüberlieferung. Um diese zu dokumentieren und zu erforschen baut das Institut für Mittelalterforschung der Österreichischen Akademie der Wissenschaften mit renommierten Partnern das Portal *manuscripta.at – mittelalterliche Handschriften in Österreich*⁸⁷ auf. In seinen Zielen ist das Projekt der Konzeption von *Manuscripta Mediaevalia* vergleichbar.

Für die Handschriftenüberlieferung der Schweiz sind insbesondere zwei Projekte zu nennen: Die *Codices Electronici Sangallensis*

⁸³ http://opac.regesta-imperii.de/lang_de

⁸⁴ <http://lba.hist.uni-marburg.de/lba>

⁸⁵ <http://www.manuscripta-mediaevalia.de/#|8>

⁸⁶ http://www.bsb-muenchen.de/fileadmin/imageswww/pdf-dateien/projekte/konzeptpapier_digitalisierung_2011_Veroeffentlichung.pdf

⁸⁷ <http://manuscripta.at/m1>

(CESG)⁸⁸, die Virtuelle Stiftsbibliothek St. Gallen, ist sicher das interessanteste Projekt, da die frühmittelalterliche Bibliothek als Teil des UNESCO-Weltkulturerbes mit ihren etwa 1000 mittelalterlichen Handschriften eine der bedeutendsten Handschriftensammlungen der Welt beherbergt. Sowohl die digitalisierten Handschriften als auch ihre wissenschaftlichen Beschreibungen sind frei zugänglich. Inzwischen ist die digitale Bibliothek des berühmten Bodenseeklosters eingebettet in das Nachfolgeprojekt *e-codices – Virtuelle Handschriftenbibliothek der Schweiz*⁸⁹. In dieser sollen sämtliche mittelalterlichen sowie die wichtigsten neuzeitlichen Handschriften der Schweiz frei zugänglich publiziert werden.

Wie für Handschriften gibt es auch für mittelalterliche und frühneuzeitliche Urkunden zentrale Portale. *Monasterium – Das Virtuelle Urkundenarchiv Europas*⁹⁰ bietet als Volltextdatenbank und Verzeichnis inzwischen Daten über mehr als 500.000 Urkunden aus 10 europäischen Ländern: Bilder, Archivregesten, wissenschaftliche Regesten, ältere oder neuere gedruckte Editionen. Über die Volltextsuche kann man sammlungsübergreifend recherchieren (zum Beispiel nach Personen).

Das 1928 an der Universität Marburg gegründete *Forschungsinstitut Lichtbildarchiv älterer Originalurkunden bis 1250 (LBA)*⁹¹ dokumentiert die originale urkundliche Überlieferung des mittelalterlichen Reiches und stellt die Abbildungen inzwischen in einer Datenbank zur Verfügung. Die differenzierte Recherchemaske erlaubt beispielsweise die Suche nach Ausstellern, Empfängern, Mitsiegler, Datierungen von Urkunden oder ihren gegenwärtigen Aufbewahrungsort. Wenn die gesuchte Urkunde innerhalb der Regesta Imperii (RI) ediert ist, gelangt man direkt in diese Datenbank und kann das Digitalisat mit dem Regest unmittelbar vergleichen.

⁸⁸ <http://www.cesg.unifr.ch/de/index.htm>

⁸⁹ <http://www.e-codices.unifr.ch/de>

⁹⁰ <http://icar-us.eu/cooperation/online-portals/monasterium-net>

⁹¹ <http://lba.hist.uni-marburg.de/lba/pages>

Eine weitere, für die Erforschung des Mittelalters wichtige Volltextdatenbank, die eine zentrale Quellengattung erschließt, sind die *Deutsche Inschriften Online (DIO)*⁹². Sie basiert auf dem als Gemeinschaftsprojekt der deutschen und österreichischen Akademien der Wissenschaften „Die deutschen Inschriften des Mittelalters und der Frühen Neuzeit“ mit dem Ziel der Sammlung und Edition aller lateinischen und deutschen Inschriften bis zum Jahr 1650. Aufgenommen werden sowohl die erhaltenen Originalinschriften als auch kopiael überlieferte Dokumente. Die Datenbank basiert zwar auf der Digitalisierung der gedruckten Edition, geht aber weit über diese hinaus, indem weiteres Material integriert ist und durch die Suchmaske Recherchen sowohl über einzelne Bände als auch über den Gesamtbestand dieser wichtigen Quellengattung erlaubt.

Geschichte der Frühen Neuzeit

Seitdem sich die Frühe Neuzeit als eigenständige Epoche innerhalb der Forschung fest etabliert hat, sind für dieses Arbeitsfeld zahlreiche Volltextdatenbanken entstanden, die zentrale Quellen komfortabel präsentieren. Da zu Beginn der Frühen Neuzeit der Buchdruck mit beweglichen Lettern ganz neue mediale Möglichkeiten schuf, stellen insbesondere die Druckerzeugnisse dieser Zeit eine für HistorikerInnen besonders wichtige Quellengattung dar. Für den deutschen Sprachraum haben sich sehr früh Projekte entwickelt, die im Zusammenhang mit der Initiative Bernhard Fabians⁹³ stehen und deren Ziel es ist, das Kulturgut Buch auch der digitalen Wissenschaft in größtmöglicher Vollständigkeit zur Verfügung zu stellen: Die nach Jahrhunderten segmentierte retrospektive Nationalbibliographie in Gestalt der Verzeichnisse der Drucke des 16., des 17. und des 18. Jahrhunderts. Das *Verzeichnis der Drucke des 16. Jahrhunderts (VD16)* erfasst die deutschsprachigen und im historischen deutschen Sprachgebiet gedruckten und verlegten fremdsprachigen Werke aus dem Zeitraum 1501-1600. In der zwischen 1969 und 1999 als Druckausgabe in 24 Bänden

⁹² <http://www.inschriften.net>

⁹³ Vgl. Anm. 30.

erschienenen Version gelang der Nachweis von 75.000 Drucken. Inzwischen als Datenbank publiziert, umfasst das VD16 nunmehr etwa 110.000 Titel und wird kontinuierlich ergänzt. Seit 2006 wird der VD16-Bestand der Bayerischen Staatsbibliothek sukzessive digitalisiert, so dass neben den bibliographischen Daten und digitalisierten Schlüsselseiten auch die gesamten Volltexte frei zugänglich sind.

Fortgeführt wird das VD16 durch das *Verzeichnis der Drucke des 17. Jahrhunderts (VD17)*⁹⁴ als retrospektive Nationalbibliographie für die Zeit von 1601–1700, dessen Publikation von Anfang an als Datenbank konzipiert war. Wie im VD16 ist es auch hier Ziel, die gegenwärtig etwa 300.000 katalogisierten Titel in mehr als 700.000 Exemplaren (Stand: Juli 2015) im Rahmen der Erschließung und Sicherung des gedruckten deutschen Kulturgutes vollständig zu digitalisieren. Im Unterschied zum VD16 werden im VD17 auch Einblattdrucke erfasst, die eine zentrale Quellengattung nicht nur für HistorikerInnen beispielsweise in der Erforschung der frühneuzeitlichen Kommunikation politischer Ereignisse darstellen, sondern gleichermaßen für theologische oder sprachhistorische Fragestellungen. Den Grundsätzen des VD16 und VD17 folgt schließlich auch das *Verzeichnis der Drucke des 18. Jahrhunderts (VD18)*⁹⁵, in dem die etwa 600.000 Drucke aus den Jahren 1701–1800 bibliographisch erfasst und im Volltext digitalisiert werden sollen.

Die innerhalb von VD16, VD17 und VD18 erfassten Dokumente werden schließlich einerseits nachgewiesen im *Zentralen Verzeichnis digitalisierter Drucke (zvdd)*⁹⁶, dem sammlungsunabhängigen Portal für den Nachweis der von deutschen Kultur- und Wissenschaftseinrichtungen erstellten Digitalisate von Drucken seit dem 15. Jahrhundert. Die Recherchemaske ermöglicht eine Suche innerhalb der Metadaten und der Volltexte der digitalisierten Wer-

⁹⁴ http://www.vd17.de/index.php?category_id=1&article_id=1&clang=0

⁹⁵ <http://vd18.de>

⁹⁶ <http://www.zvdd.de/startseite>

ke. Andererseits kooperieren diese Verzeichnisse mit der *Deutschen Digitalen Bibliothek*⁹⁷ sowie der *Europeana Collections*⁹⁸.

Für die englische Buchproduktion der Frühen Neuzeit ist die als Nationallizenz zur Verfügung stehende Volltextdatenbank *Early English Books Online (EEBO)*⁹⁹ maßgeblich. Sie umfasst etwa 125.000 Bücher aus der Zeit vom Beginn der englischen Buchproduktion bis zu Shakespeare und ist damit vergleichbar mit VD16 und VD17. Sämtliche Werke sind im Volltext durchsuchbar und können heruntergeladen oder ausgedruckt werden.

Ein weiteres breites Spektrum von Textquellen für die Erforschung des frühneuzeitlichen englischen Kulturraumes bietet die Volltextdatenbank *Eighteenth Century Collections Online (ECCO)*¹⁰⁰. Enthalten sind zwischen 1701 und 1800 in England erschienene Bücher, Pamphlete, Essays und Einblattdrucke, in denen mit Hilfe einer komplexen Suchmaske differenziert recherchiert werden kann. Insbesondere für vergleichende Forschungen zwischen dem deutschsprachigen und dem englischen Kulturraum bieten *Early English Books Online*¹⁰¹ und *Eighteenth Century Collections Online*¹⁰² einen reichen Schatz an Quellen, dessen systematische Erforschung erst am Anfang steht.

Ein zentrales Thema der Frühneuzeitforschungen zielt auf die Analyse religiösen Denkens beispielsweise im Zusammenhang mit der Reformation oder dem Dreißigjährigen Krieg. Für den Protestantismus bietet die *Digital Library of Classic Protestant Texts*¹⁰³ Volltexte von mehr als 1.500 Werken 325 protestantischer Autoren des 16. und 17. Jahrhunderts. Das Pendant für den Ka-

⁹⁷ <https://www.deutsche-digitale-bibliothek.de>

⁹⁸ <http://www.europeana.eu/portal>

⁹⁹ <http://eebo.chadwyck.com/home>

¹⁰⁰ <http://gale.cengage.co.uk/product-highlights/history/eighteenth-century-collections-online.aspx>

¹⁰¹ <http://eebo.chadwyck.com/home>

¹⁰² <http://quod.lib.umich.edu/e/ecco>

¹⁰³ <http://alexanderstreet.com/products/digital-library-classic-protestant-texts>

tholizismus ist die *Digital Library of the Catholic Reformation*¹⁰⁴, die etwa 2.000 Texte katholischer Autoren desselben Zeitraumes umfasst. In beiden Fällen ermöglicht eine komplexe Suchmaske differenzierteste Recherchen innerhalb der Volltexte.

In der Epoche der Frühen Neuzeit gewann nach der Entdeckung Amerikas die USA eine immer wichtigere Bedeutung innerhalb einer transatlantischen (Kultur-) Geschichte. Eine wichtige Quellengrundlage für die Erforschung der frühen Geschichte Amerikas und der Vereinigten Staaten bietet hierfür die Volltextdatenbank *Early American Imprints*¹⁰⁵. Innerhalb der *Series I*¹⁰⁶ finden sich mehr als 36.000 digitale Ausgaben der zwischen 1639 und 1800 in Nordamerika publizierten Bücher. Die *Series II*¹⁰⁷ schließt daran an und umfasst 47.000 digitalisierte Werke, die zwischen 1801 und 1819 in Nordamerika erschienen sind.

Neuere Geschichte und Zeitgeschichte

Für die Neuere Geschichte und Zeitgeschichte stehen angesichts der weiter expandierenden Schriftlichkeit immer mehr Quellen und Forschungsliteratur innerhalb von Volltextdatenbanken zur Verfügung. Begünstigend wirkt hierbei, dass Digitalisierung und OCR-Erkennung moderner Typographien seit dem 19. Jahrhundert weniger aufwändig ist als im Falle von Inkunabeln, früher Drucke oder unikaler Dokumente älterer Epochen der Schriftlichkeit.

Erweitert man den Blick auf die Geschichte der Vereinigten Staaten von Amerika im 19. Jahrhundert so stehen hier zahlreiche Primärquellen in Volltextdatenbanken zur Verfügung. Exemplarisch sei eine Volltextdatenbank für Zeitungen, das zentrale Kommunikationsmedium des 19. und 20. Jahrhunderts, genannt:

¹⁰⁴ <http://alexanderstreet.com/products/digital-library-catholic-reformation>

¹⁰⁵ <http://www.readex.com/content/early-american-imprints-series-i-evans-1639-1800>

¹⁰⁶ <http://www.readex.com/content/early-american-imprints-series-i-evans-1639-1800>

¹⁰⁷ <http://www.readex.com/content/early-american-imprints-series-ii-shaw-shoemaker-1801-1819>

*America's Historical Newspapers*¹⁰⁸ umfasst digitalisierte Zeitungen aus 50 Bundesstaaten aus der Zeit zwischen 1690 und 1922. Mit Hilfe der Volltextsuche lassen sich politische und gesellschaftliche Diskurse des gesamten nordamerikanischen Kontinents auch in einer diachronen Perspektive analysieren.

Unterschiedliche Archive internationaler Primärquellen aus Geschichte, Politik, Literatur und Kunst des 19. Jahrhunderts über Nordamerika hinaus bietet beispielsweise die Volltextdatenbank *Nineteenth Century Collections Online (NCCO)*¹⁰⁹. Ihre regionalen Schwerpunkte liegen auf Europa, Afrika, dem Mittleren Osten, Ost- und Südasiens sowie Mittelamerika.

Trotz vielfältiger Schwierigkeiten und Restriktionen bei der Erschließung und Digitalisierung der in osteuropäischen Bibliotheken und Archiven aufbewahrten Quellen existieren inzwischen doch wichtige Volltextdatenbanken zur Geschichte Osteuropas. *Integrum World Wide*¹¹⁰ ist die umfangreichste Volltextdatenbank Russlands und der GUS mit etwa 360 Millionen Dokumenten aus den Bereichen Politik, Kultur, Wirtschaft und Gesellschaft.

Für weitere Volltextdatenbanken zu den einzelnen Regionen und Epochen der Geschichte sei auf die thematisch einschlägigen Guides dieses Handbuchs verwiesen.

2.4 Volltextdatenbanken und Volltextsammlungen für Forschungsliteratur – Monographien

Entsprechend dem Überblick über Volltextdatenbanken für Quellen werden im Folgenden exemplarisch zentrale Volltextdatenbanken für Forschungsliteratur besprochen. Hierbei wird unterschieden zwischen Monographien und Zeitschriften respektive Zeitschriftenaufsätzen. Sofern Volltextdatenbanken als Nationallizenzen deutschlandweit zur Verfügung stehen, werden die einzelnen Texte unter einem gemeinsamen Portal zugänglich gemacht. Da es sich – wie oben dargelegt – bei den Inhalten histori-

¹⁰⁸ <http://www.readex.com/content/americas-historical-newspapers>

¹⁰⁹ <http://gale.cengage.co.uk/product-highlights/history/nineteenth-century-collections-online.aspx>

¹¹⁰ <http://www.integrumworld.com>

scher Volltextdatenbanken sowohl um Quellen als auch um Forschungsliteratur handeln kann, finden sich innerhalb dieses Portals beide Kategorien. Auf diese Weise ermöglicht das Portal *Nationallizenzen / Sammlungen – Monographien*¹¹¹ die parallele Recherche über 39 (Stand: Februar 2014) E-Book- respektive Textsammlungen.

Neben den auf nationaler Ebene von der Deutschen Forschungsgemeinschaft (DFG) aus Steuermitteln erworbenen kommerziellen Volltextdatenbanken entwickeln sich gegenwärtig Portale, die nach dem Prinzip des Open Access publizierte Volltexte systematisch erfassen. Für den Bereich Monographien ist das *Directory of Open Access Books (DOAB)*¹¹² das zentrale Portal, das inzwischen 3.284 Monographien aus 110 wissenschaftlichen Verlagen (Stand: August 2015) verzeichnet.

Schließlich darf ein wichtiges Instrument, (geschichts-) wissenschaftliche Forschungsergebnisse in Volltexten zu recherchieren, nicht unerwähnt bleiben: *BASE – Bielefeld Academic Search Engine*¹¹³ ist eine Suchmaschine, die den Zugriff auf etwa 75 Millionen frei zugängliche wissenschaftliche Texte in mehr als 3.500 Repositorien ermöglicht (Stand: Juni 2015). Da in diesen Repositorien, die inzwischen beinahe jede größere Universität oder Forschungseinrichtung weltweit aufgebaut hat, wissenschaftliche Publikationen der jeweiligen Institution (zum Beispiel Dissertationen) veröffentlicht werden, ist BASE – noch vor Google Scholar – ein zentraler Einstieg in die Recherche nach wissenschaftlicher Literatur in Volltexten.

Wie staatliche Wissenschaftsinstitutionen ihre Forschungsergebnisse immer häufiger Open Access zur Verfügung stellen, positionieren auch kommerzielle Verlage ihre Produkte. Beispielsweise werden Monographien oder Zeitschriften, die zuvor bei dem jeweiligen Verlag im Druck erschienen sind, retrodigitalisiert in Volltextdatenbanken kostenpflichtig angeboten. Ein Beispiel ist

¹¹¹ <http://gso.gbv.de/xslt/DB=1.50/LNG=DU/?COOKIE=U8000,K8000,I0,B1999+++++,SY,NVZG,D1.50,E8f5eedd5-1a,A,H,R132.230.239.175,FY>

¹¹² <http://www.doabooks.org>

¹¹³ <http://www.base-search.net>

Cambridge University Press mit der Volltextdatenbank *Cambridge Histories Online*^{114, 115}

Für die Geschichte Osteuropas umfassen die *Osteuropa-Dokumente Online (OstDok)*¹¹⁶ neben gemeinfreien auch urheberrechtlich geschützte wissenschaftliche Werke zur Osteuropafor- schung und bieten die Möglichkeit, relevante Forschungsergeb- nisse als Erst- oder Zweitveröffentlichung online zu publizieren. Eingebunden ist diese hochinteressante Volltextsammlung in die Virtuelle Fachbibliothek Osteuropa (ViFaOst).

Über den engeren Bereich der Geschichtswissenschaften hinaus geht schließlich die internationale, interdisziplinäre und epo- chenübergreifende Publikationsplattform für Geisteswissenschaf- ten *perspectivia.net*¹¹⁷. Sie steht für sämtliche wissenschaftliche Textsorten zur Verfügung und publiziert nicht nur zuvor im Druck erschienene Publikationen wie die retrodigitalisierte Zeit- schrift *Francia* oder ausgewählte *Publikationen der Max Weber Stif- tung – Deutsche Geisteswissenschaftliche Institute im Ausland*, son- dern auch primäre elektronische Texte.

2.5 Volltextdatenbanken und Volltextsammlungen für Forschungsliteratur – Zeitschriften und Zeitschriftenaufsätze

Ähnlich vielfältig wie Volltextdatenbanken oder Volltextsamm- lungen für wissenschaftliche Monographien ist das Angebot für in Periodika publizierte Forschungsliteratur. Das Portal *Nationalli- zenzen / Sammlungen – Zeitschriften*¹¹⁸ bietet analog zu dem Portal für Monographien für die 23 (Stand: Juli 2012) als Nationallizen- zen finanzierten Zeitschriften- und Zeitungsarchive mit etwa

¹¹⁴ <http://universitypublishingonline.org/cambridge/histories>

¹¹⁵ Aber auch andere Verlage wie Oxford University Press, de Gruyter, Cengage oder Blackwell bieten vergleichbare Produkte an.

¹¹⁶ <https://www.vifaost.de/ostdok>

¹¹⁷ <http://www.perspectivia.net>

¹¹⁸ <http://gso.gbv.de/DB=1.55/LNG=DU/?COOKIE=U8000,K8000,I0,B1999+++++,SY,NVZG,D1.55,E9ccb150f-2d,A,H,R132.230.239.127,FY&COOKIE=U8000,K8000,I0,B1999+++++,SY,NVZG,D1.55,E9ccb150f-2d,A,H,R95.115.0.8,FY>

16.000 Zeitschriften und 450 Zeitungen den Zugriff auf die Volltexte auf Artikelebene.

Für die nach dem Prinzip des open access publizierten Zeitschriften ist das *Directory of Open Access Journals (DOAJ)*¹¹⁹ das zentrale Verzeichnis. Gegenwärtig weist es 10.529 Zeitschriften mit 1.985.568 Artikeln (Stand: August 2015) nach.

Eine der wichtigsten lizenzpflichtigen Volltextdatenbanken für Zeitschriften aus dem gesamten Spektrum der Geistes-, Kultur- und Sozialwissenschaften ist das *Periodicals Archive Online (PAO)*¹²⁰. Es bietet über eine differenzierte Recherchemaske den Zugriff auf 1,9 Millionen durch Abstracts gut erschlossene Artikel aus 500 Zeitschriften. Als wichtige Ergänzung hierzu insbesondere für geschichtswissenschaftliche Fragestellungen bietet *Periodicals Index Online (PIO)*¹²¹ als historischer Zeitschriften-Index bibliographische Daten für Artikel aus etwa 6.000 Zeitschriften aus den Jahren 1665-2000. Sind die Volltexte im Periodicals Archive Online (PAO) enthalten, wird aus dem Periodicals Index Online (PIO) direkt auf diese verlinkt.

Unverzichtbar für die Recherche in internationalen Kernzeitschriften nicht nur der Geschichtswissenschaften ist die Volltextdatenbank *JSTOR*¹²². Enthalten sind Zeitschriften von ihrem ersten Jahrgang an, es fehlt jedoch der Zugriff auf die aktuellen Hefte aufgrund einer *moving wall* von bis zu 11 Jahren. Da die einzelnen Aufsätze nicht durch Schlagworte erschlossen sind, ist für die Volltextrecherche in JSTOR zu beachten, dass jeweils Suchbegriffe in unterschiedlichen Sprachen verwendet werden müssen – je nachdem, in welcher Sprache ein gesuchter Artikel publiziert ist.

Das deutsche Pendant zu JSTOR ist die Volltextdatenbank *Digi-zeitschriften*¹²³, in der etwa 200 deutsche Kernzeitschriften mit etwa 500.000 Artikeln aus dem Bereich der Geistes- und Sozialwis-

¹¹⁹ <https://doaj.org>

¹²⁰ http://www.proquest.com/products-services/periodicals_archive.html

¹²¹ <http://search.proquest.com/pio?accountid=10979>

¹²² <http://www.jstor.org>

¹²³ <http://www.digi-zeitschriften.de/startseite>

senschaften retrospektiv digitalisiert sind. Während Recherche-
maske und Inhaltsverzeichnisse frei zugänglich sind, müssen die
Volltexte überwiegend lizenziert werden. Für die Geschichtswis-
senschaft stehen 15 Zeitschriften mit 948 Bänden zur Verfügung.

Einen Schwerpunkt auf Ostmittel- und Südosteuropa besitzt die
*Central and Eastern European Online Library (CEEOL)*¹²⁴ mit Volltex-
ten wissenschaftlicher, politischer aber auch kultureller Zeitschrif-
ten. Durch die Erschließung der Texte durch englische Überset-
zungen, Abstracts und Schlagwörter ist eine differenzierte Re-
cherche möglich.

3. Zusammenfassung und Ausblick

Dieser knappe Survey über historische Volltextdatenbanken und
Volltextsammlungen will einen allerersten Einblick in die Beson-
derheiten dieses für geschichtswissenschaftliches Lernen, Lehren
und Forschen so zentralen Ressourcentypus gewähren sowie re-
levante Beispiele für Quellen und Forschungsliteratur im Hinblick
auf einzelne Epochen vorstellen. Ein wichtiges Ziel ist es dabei,
HistorikerInnen Kriterien zu benennen, anhand derer einzelne
Volltextdatenbanken in ihrer Eignung für spezifische historische
Fragestellungen beurteilt werden können. Als entscheidend wer-
den hierbei Fragen der technischen, formalen und inhaltlichen
Beschreibung der Texte durch OCR-Erkennung, Indexierung, bib-
liothekarische Katalogisierung und intellektuelle Erschließung
herausgearbeitet. Sofern diese Kriterien nicht ausreichend in der
Konzeption von Volltextdatenbanken berücksichtigt sind, stößt
eine Volltextsuche leicht an ihre Grenzen. Daher ist anzustreben,
künftig noch konsequenter auf eine einheitliche technische Aufbe-
reitung und Strukturierung der Daten einerseits, auf ihre formale
und wissenschaftliche Erschließung andererseits zu achten. Ma-
thematische Prozeduren und Algorithmen alleine genügen in den
meisten Fällen nicht, um differenzierte historische Fragestellun-
gen zu beantworten, ebenso wenig die bibliothekarische Erfas-
sung rein formaler Metadaten. Notwendig sind vielmehr fachspe-
zifische Thesauri und Vokabulare, die verknüpft sind mit der von

¹²⁴ <http://www.ceeol.com>

Bibliotheken verwendeten *Gemeinsamen Normdatei (GND)* für Personen, Körperschaften, Geographica, Sachschlagwörter.¹²⁵

Historische Volltextdatenbanken für Quellen und Forschungsliteratur werden künftig auch eine zentrale Rolle bei der Weiterentwicklung von Konzepten der Historischen Grundwissenschaften in Verbindung mit der weiteren Digitalisierung der Geschichtswissenschaft und der Entwicklung von Methoden und Werkzeugen der Digital Humanities spielen.¹²⁶ Nicht nur immer neue Volltextdatenbanken entstehen, sie werden auch immer intelligenter in unterschiedliche Rechercheszenarien eingebunden. So können beispielsweise Quellendatenbanken, die Teile des Handschriftenerbes dokumentieren, nicht nur mit wissenschaftlichen Editionen der in ihnen überlieferten Texte verknüpft werden, sondern auch mit der relevanten Forschungsliteratur. Konzepte des *Semantic Web* und der *Linked Open Data* ermöglichen über formale Erschließungsverfahren hinaus kontextsensitive Vernetzungen von Daten unterschiedlicher Herkunft.

¹²⁵ http://www.dnb.de/DE/Standardisierung/GND/gnd_node.html – Vgl. Moeller, Katrin, Grundwissenschaften als Masterdisziplin der Nachnutzung, in: H-Soz-Kult, 11.12.2015, <http://www.hsozkult.de/debate/id/diskussionen-2923>.

¹²⁶ Das Grundsatzpapier „Quellenkritik im digitalen Zeitalter. Die Historischen Grundwissenschaften als zentrale Kompetenz der Geschichtswissenschaft und benachbarter Fächer“ (Schlotheuber, Eva; Bösch, Frank, Quellenkritik im digitalen Zeitalter: Die Historischen Grundwissenschaften als zentrale Kompetenz der Geschichtswissenschaft und benachbarter Fächer, in: H-Soz-Kult, 16.11.2015. <http://www.hsozkult.de/debate/id/diskussionen-2866> regte eine intensive Diskussion an (Diskussionsforum: Historische Grundwissenschaften und die digitale Herausforderung, in: H-Soz-Kult, 15.11.2015, <http://www.hsozkult.de/text/id/texte-2890>.

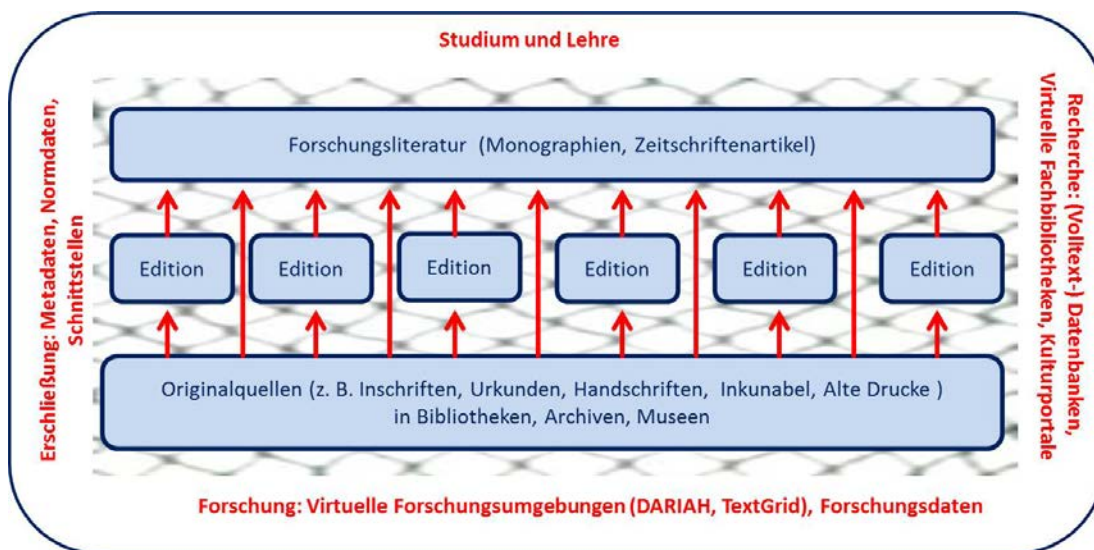


Abbildung 4. Volltextdatenbanken für Quellen – Editionen – Forschungsliteratur im Netz der Digital Humanities

Um diese komplexen digitalen Ressourcen und webbasierten Rechercheszenarien passgenau weiter zu entwickeln, ist es notwendig, dass die geschichtswissenschaftliche Fachcommunity gemeinsam mit Bibliotheken, Archiven, Museen und Wissenschaftsverlagen zukunftsfähige Konzepte erarbeitet. Den Historischen Grundwissenschaften und Digital Humanities fällt hierbei eine Schlüsselrolle zu, um die Geschichtswissenschaft im digitalen Raum handlungsfähig zu machen.

Da am Beispiel der Volltextdatenbanken die grundlegende Änderung des geschichtswissenschaftlichen Arbeits-, Schreib- und Publikationsprozesses bereits jetzt sichtbar wird, sollte vor allem der Vermittlung dieser Kenntnisse und Fähigkeiten in den geschichtswissenschaftlichen Curricula ein angemessener Platz eingeräumt werden. Hierbei kann den Historischen Grundwissenschaften eine besondere Rolle zufallen. Warum?

Da durch ihre Digitalisierung geschichtswissenschaftliche Quellen aus Bibliotheken, Archiven und Museen eine bisher nie dagewesene Sichtbarkeit besitzen (wie viel Zeit, Kosten und Mühen bedeuteten für HistorikerInnen in vordigitaler Zeit Archiv- und Bibliotheksreisen im In- und Ausland!), müssen die Kompetenzen für den Umgang mit diesen umso intensiver vermittelt werden. Die immer umfangreichere Digitalisierung von Quellen durch Bibliotheken, Archive und Museen erfordert somit einen systema-

tischen Ausbau der wissenschaftlichen Ausbildung, damit diese Quellen überhaupt dechiffriert werden können. Insofern würde die verstärkte Öffnung der Geschichtswissenschaft zur Digitalisierung und den Digital Humanities zugleich eine Wiedergeburt der Historischen Grundwissenschaften bedeuten, die innerhalb des Bologna-Prozesses an Bedeutung verloren haben.¹²⁷ Digitalisierung und Digital Humanities könnten aber nun nicht nur zu einer neuen Blüte der Historischen Grundwissenschaften führen, sondern sich zugleich als deren unverzichtbarer Bestandteil als „Quellenkunde des Digitalen“¹²⁸ als eigene Teildisziplin fest etablieren.¹²⁹ Bereits heute ist einerseits klar, dass die Historischen Grundwissenschaften künftig „ohne die Nutzung großer internetgestützter Datenressourcen nicht mehr auskommen,“¹³⁰ dass andererseits eine Rekonstruktion der Geschichte durch schlichte *data driven history*, die „nur alle möglichen Daten sammelt und nebeneinander stellt [...] im Romantizismus des 19. Jahrhunderts stecken“¹³¹ bleibe. So verstanden würden digitale Wende und Digital Humanities sogar Rückschritt bedeuten.

¹²⁷ Vogeler sieht darin eine „eigenartige Position“, dass erst die Digitalisierung den Bedarf an quellenkritischer Kompetenz gesteigert habe – da doch die digitalisierten Quellen bereits vorher da gewesen sind. Die „demokratische“ Verfügbarkeit der Quellen unserer Kultur durch Digitalisierung bedeute aber zugleich, dass auch die Kompetenzen ihrer Dechiffrierung für weitere Kreise von Interessierten erreichbar sein müssten. Vogeler, Georg, Digitale Quellenkritik in der Forschungspraxis, in: H-Soz-Kult, 28.11.2015, <http://www.hsozkult.de/debate/id/diskussionen-2893>.

¹²⁸ Müller, Harald, Schlüsselkompetenzen der Quellenkundigkeit, in: H-Soz-Kult, 20.11.2015, <http://www.hsozkult.de/debate/id/diskussionen-2892>.

¹²⁹ Hiltmann geht so weit, in den digitalen Techniken und Methoden „nicht nur eine Chance, sondern vielleicht die einzige Möglichkeit für eine sinnvolle Weiterentwicklung der Hilfswissenschaften“ zu erkennen. Hiltmann, Torsten, Hilfswissenschaften in Zeiten der Digitalisierung, in: H-Soz-Kult, 14.12.2015, <http://www.hsozkult.de/debate/id/diskussionen-2936>.

¹³⁰ Spickermann, Harald; Scheuermann, Leif, Grundwissenschaften in den Altertumswissenschaften, in: H-Soz-Kult, 05.12.2015, <http://www.hsozkult.de/debate/id/diskussionen-2907>.

¹³¹ Ebd.

Erkennt man jedoch die außerordentliche Chance der Digitalisierung für die „scheinbar verstaubten Hilfswissenschaften,“¹³² ihren traditionell eher deskriptiven Ansatz aufzugeben und die sich entwickelnden Werkzeuge aus dem Bereich der Digital Humanities aktiv anzuwenden – dann freilich könnte eine Situation entstehen, von der alle Beteiligten substanziell profitieren: Die bisherigen Digitalisierungsprojekte bieten einerseits eine bisher nie dagewesene Basis für die Ausbildung in den Historischen Grundwissenschaften, andererseits ist die erst durch das Studium dieser Grundwissenschaften zu erlangende Expertise für künftige Digitalisierungsprojekte unverzichtbar. Der innerhalb des Bologna-Prozesses geforderte Praxisbezug insbesondere historisch orientierter geisteswissenschaftlicher Disziplinen wird im Falle der Historischen Grundwissenschaften als Grundlagenfach mit unmittelbarem Anwendungspotenzial im Kontext der Digitalisierung und Digital Humanities vorbildlich eingelöst.

Das vom Verband der Historiker und Historikerinnen Deutschlands veröffentlichte Grundsatzpapier zum Status der Historischen Grundwissenschaften „Quellenkritik im digitalen Zeitalter: Die Historischen Grundwissenschaften als zentrale Kompetenz der Geschichtswissenschaft und benachbarter Fächer“¹³³ wird nicht von ungefähr im unmittelbaren Kontext mit der Digitalisierung der Quellen intensiv diskutiert.¹³⁴ Die zahlreichen Beiträge umreißen sehr prägnant das Feld digitaler Fachinformation, in das sich der vorliegende Guide zu historischen Volltextdatenbanken explizit einordnet. Die Kernproblematik wird darin gesehen, dass inzwischen zwar Millionen digitalisierter Dokumente und Objekte mit technischen und bibliothekarischen Metadaten vorliegen, dass inhaltliche Tiefenerschließung sowie die wissen-

¹³² Johrendt, Jochen, Digitalisierung als Chance, in: H-Soz-Kult, 02.12.2015, <http://www.hsozkult.de/debate/id/diskussionen-2902>.

¹³³ Schlothuber, Eva; Bösch, Frank, Quellenkritik im digitalen Zeitalter, wie Anm. 125.

¹³⁴ Diskussionsforum: Historische Grundwissenschaften und die digitale Herausforderung, wie Anm. 125. In den folgenden Bemerkungen wird lediglich auf signifikante Argumente Bezug genommen.

schaftliche Transkription und Kommentierung in der Regel jedoch fehle.¹³⁵

Auch wenn die Gegenstände der Historischen Grundwissenschaften durch die Digitalisierung inzwischen vor aller Augen liegen, hat sich dennoch an der Tatsache nichts geändert, dass Lektüre und Interpretation dieser Quellen Spezialwissen erfordert, das erst in den Historischen Grundwissenschaften erlernt werden muss. Dieser Omnipräsenz digitalisierter historischer Dokumente, die für Studierende zugleich eine hohe Attraktivität besitzen, steht jedoch nicht nur die vielfach beklagte „Marginalisierung der Grundwissenschaften an den Universitäten“¹³⁶ entgegen, sondern auch eine Abnahme entsprechender Kompetenzen in den wissenschaftlichen Bibliotheken. Die Fortführung der bibliothekarischen Digitalisierung und Erschließung historischer Quellen sollte daher noch konsequenter als bisher durch geschichtswissenschaftliche Expertise intensiv begleitet werden: „Noch nie standen die Bemühungen zur Sicherung und Verfügbarmachung des historischen Erbes in einem so krassen Missverhältnis zur Aufschlüsselbarkeit wie heute [...] Damit droht die Digitalisierungsoffensive in einen medialen, technisch hochwertigen und gut verfügbaren Illustrationsfundus zu münden, der letztlich nurmehr bestaunt, kaum noch bearbeitet werden kann.“¹³⁷ Ziel sollte also vielmehr die „Rückeroberung der quellenkritischen Kompetenzen durch Bibliotheken, Archive sowie Medien- und Informationswissenschaften auch im Digitalen“¹³⁸ sein. Und bereits in der Auswahl der für Digitalisierungsprojekte vorgesehe-

¹³⁵ Forum: Historische Grundwissenschaften und die digitale Herausforderung, in: H-Soz-Kult, 16.11.2015, <http://www.hsozkult.de/debate/id/diskussionen-2889>.

¹³⁶ Märkl, Claudia, Grundwissenschaften im Studium, in: H-Soz-Kult, 26.11.2015, <http://www.hsozkult.de/debate/id/diskussionen-2896>.

¹³⁷ Müller, Schlüsselkompetenzen der Quellenkundigkeit, wie Anm. 127.

¹³⁸ Pflanzelter, Eva, Historische Quellenkritik in Lehre und Forschung, in: H-Soz-Kult, 24.11.2015, <http://www.hsozkult.de/debate/id/diskussionen-2903>.

nen Quellen sollte nicht vornehmlich das technisch Machbare im Fokus stehen, sondern vielmehr die konkrete historische Frage.¹³⁹

Die künftigen Kompetenzen innerhalb der historischen Grundwissenschaften könnten dann aber nicht mehr nur in traditionell deskriptiver Arbeit verortet sein, sondern in konzeptioneller Aufbereitung der digitalen Daten, ihrer Strukturierung, Kontextualisierung, Analyse und vor allem in der Überlegung, mit welchen Werkzeugen der Digital Humanities sich alte historische Fragestellungen beantworten und neue stellen ließen.¹⁴⁰ Auf diese Weise wäre nicht so sehr das Ziel, die Digital Humanities als weitere Einzeldisziplin innerhalb des Spektrums der Historischen Hilfswissenschaften zu etablieren, sondern viel umfassender die „Historischen Hilfswissenschaften methodisch zu digitalisieren.“¹⁴¹

Da die ubiquitäre Verfügbarkeit digitalisierter Quellen und Literatur fundiertes Fachwissen nicht wird ersetzen können, ist die weitere Massendigitalisierung nur dann sinnvoll, wenn die grundwissenschaftlichen Kompetenzen ihrer Bearbeitung nicht nur innerhalb der geschichtswissenschaftlichen Curricula vermittelt werden, sondern auch innerhalb der Archive und Bibliotheken zuverlässig vorhanden sind. Würde in Zukunft die von Bibliotheken, Archiven und Museen durchgeführte Digitalisierung darüber hinaus intensiv durch den erneuten Fokus auf die Historischen Grundwissenschaften durch die historische Fachwissenschaft flankiert, könnten aus dieser Synergie heraus Techniken und Werkzeuge der Digital Humanities entwickelt werden. Erst dann kann ein wissenschaftlicher Mehrwert aus den digitalisierten Quellen und Texten entstehen, indem durch neue Analyseverfahren bisher ungelöste Fragen beantwortet und neue Fragen ge-

¹³⁹ Ebd. Digitalisierung bedeutet daher nicht den „fröhlichen Positivismus des ‚Einfach-mal-auf-den-Scanner-legen-irgendwer-wird-es-schon-brauchen-in-den-nächsten-100-Jahren‘“. Krajewski, Markus, Programmieren als Kulturtechnik, in: H-Soz-Kult, 30.11.2015, <http://www.hsozkult.de/debate/id/diskussionen-2901>.

¹⁴⁰ Vgl. Hiltmann, Torsten, Hilfswissenschaften in Zeiten der Digitalisierung, in: H-Soz-Kult, 14.12.2015, <http://www.hsozkult.de/debate/id/diskussionen-2936>.

¹⁴¹ Ebd.

stellt werden können. „Digital History war von Anfang an ein Mannschaftssport und kein individueller Leistungssport“¹⁴² – geschichtswissenschaftliche Arbeit im digitalen Raum bedeutet die Notwendigkeit nicht nur fachlicher Kooperation und Kollaboration wie beispielsweise mit der Informatik, sondern gleichermaßen institutioneller Kooperation mit Archiven, Bibliotheken und Museen.

Da „Informations- und Quellenkritik“¹⁴³ im Digitalen wie im Analogen gilt, sollte auch die „digitale Aufklärung“¹⁴⁴ nicht nur in der universitären Lehre¹⁴⁵, sondern gleichermaßen innerhalb der Lehrangebote der Bibliotheken zur Vermittlung fachwissenschaftlicher Informationskompetenz systematisch erfolgen. Da Digitalisierung und Datenbanken an sich noch keine historische Erkenntnis liefern, ermöglichen die Digital Humanities als Historische Grundwissenschaft „das so elementar wichtige Nutzenkönnen digitaler Methoden und Daten, wie die Paläographie uns das Lesenkönnen unserer Quellen sicherstellt.“¹⁴⁶ Informations- und Medienkompetenz verbinden sich auch in der Geschichtswissenschaft der Zukunft mit der klassischen Quellenkritik – daher dürfte „Programmieren als Kulturtechnik“¹⁴⁷ genauso wichtig werden wie traditionelle Kulturtechniken – gilt es doch einen „digitalen Analphabetismus“¹⁴⁸ zu vermeiden.

Dennoch – bei allen genannten Vorteilen von Digitalisierung, Volltextdatenbanken und Digital Humanities – wird geschichtswissenschaftliches Arbeiten auch künftig nicht auf das klassische Bibliotheks- und Archivstudium verzichten können. Insbesondere

¹⁴² Olsen, Jon, Digital History als Mannschaftssport, in: H-Soz-Kult, 23.11.2015, <http://www.hsozkult.de/debate/id/diskussionen-2894>.

¹⁴³ Rehbein, Malte, Digitalisierung braucht HistorikerInnen, die sie beherrschen, nicht beherrscht, in: H-Soz-Kult, 27.11.2015, <http://www.hsozkult.de/debate/id/diskussionen-2905>.

¹⁴⁴ Ebd.

¹⁴⁵ Vogeler, Georg, wie Anm. 126.

¹⁴⁶ Rehbein, Malte, wie Anm. 142.

¹⁴⁷ Krajewski, Markus, wie Anm. 138.

¹⁴⁸ Ebd.

spezialisierte Forschungsvorhaben werden immer auch auf Quellen zurückgreifen, die (noch) nicht im Fokus der Digitalisierung der Geschichtswissenschaft und unserer Kultur liegen, sondern auf ihre Entdeckung in Bibliotheken, Archiven und Museen warten. Eine vollständige, flächendeckende Volltextsuche innerhalb sämtlicher geschichtswissenschaftlich relevanter Quellen und Forschungsliteratur vom eigenen Arbeitsplatz aus bleibt Utopie, solange nicht alle Texte und Artefakte der menschlichen Kultur digitalisiert worden sind. Selbst wenn es möglich wäre, sämtliche relevanten Bibliotheken, Archive und Museen zu identifizieren und einzubeziehen, sind nicht allein die Kosten einer solchen Digitalisierung inklusive der Erschließung und Einbindung der Digitalisate in wissenschaftsrelevante Lehr-, Lern- und Forschungsumgebungen bereits erheblich. Noch viel höhere Kosten würden durch die dauerhafte Archivierung und Pflege der Daten entstehen – zumal das Problem der zuverlässigen Langzeitarchivierung digitaler Daten bis heute als nicht gelöst gilt. Insofern dürfte der Gedächtnisverlust der digitalen Epoche – gemessen an der Quantität der gegenwärtig und künftig produzierten Daten, Informationen, Texten und Medien – noch deutlich größer sein als der Gedächtnisverlust vergangener Epochen mit ihren analog gespeicherten Daten in Gestalt von Texten auf Stein, Metall, Papyrus, Pergament oder Papier.

Literaturhinweise

- Assmann, Bernhard; Sahle, Patrick, Digital ist besser. Die Monumenta Germaniae Historica mit den dMGH auf dem Weg in die Zukunft – eine Momentaufnahme, Norderstedt 2008.
- Brunner, Otto u.a. (Hrsg.), Geschichtliche Grundbegriffe: historisches Lexikon zur politisch – sozialen Sprache in Deutschland, Stuttgart 1972–1997.
- Brügger, Niels, Web history, New York u.a. 2010.
- Cohen, Daniel J.; Rosenzweig, Roy, Digital history: A guide to gathering, preserving, and presenting the past on the web, Philadelphia 2006.
- Danker, Uwe; Schwabe, Astrid, Geschichte im Internet, Stuttgart 2014.
- Diskussionsforum: Historische Grundwissenschaften und die digitale Herausforderung, in: H-Soz-Kult, 16.11.2015, <http://www.hsozkult.de/debate/id/diskussionen-2889>.
- Dougherty, Jack, Writing history in the digital age, Ann Arbor 2013.
- Fabian, Bernhard, Buch, Bibliothek und geisteswissenschaftliche Forschung. Zu Problemen der Literaturversorgung und Literaturproduktion in der Bundesrepublik Deutschland, Göttingen 1983.
- Gantert, Klaus, Elektronische Informationsressourcen für Historiker, Berlin 2011.
- Haber, Peter, Digital Past: Geschichtswissenschaften im digitalen Zeitalter, München 2011.
- Kann, Bettina; Hintersonleitner, Michael, Volltextsuche in historischen Texten. Erfahrungen aus den Projekten der Österreichischen Nationalbibliothek, in: Bibliothek – Forschung und Praxis 39 (2015), S. 73–79.
- Kämmerer, Carmen, Vom Image zum Volltext – Möglichkeiten und Grenzen des Einsatzes von OCR beim Alten Buch, in: Bibliotheksdienst 43 (2009), S. 626–659.
- Kemper, Alfons; Eickler, Andre, Datenbanksysteme: eine Einführung, 7., aktualisierte und erweiterte Auflage, München 2009.
- Oehlmann, Doina, Erfolgreich recherchieren – Geschichte, Berlin 2012.
- [Patrologia cursus completus / Series Graeca] Patrologiae cursus completus: in qua prodeunt patres, doctores scriptoresque ecclesiae Graecae [...] accurante J. P. Migne. Paris, 1857–1866.
- [Patrologia cursus completus / Series Latina] Patrologiae cursus completus: seu bibliotheca universalis, integra, uniformis, commoda, oeconomica, omnium ss. Patrum [...] ecclesiasticorum, sive Latinorum, sive Graecorum [...] accurante J. P. Migne. Paris, 1844–1855.

Schlotheuber, Eva; Bösch, Frank, Quellenkritik im digitalen Zeitalter: Die Historischen Grundwissenschaften als zentrale Kompetenz der Geschichtswissenschaft und benachbarter Fächer, in: H-Soz-Kult, 16.11.2015, <http://www.hsozkult.de/debate/id/diskussionen-2866>.

Weller, Toni, History in the digital age, London; New York 2013.

Dr. Marcus Schröter ist Fachreferent für Geschichte und Altertumswissenschaften an der Universitätsbibliothek Freiburg (im Breisgau). An der Bibliotheksakademie Bayern unterrichtet er das Fach „Vermittlung von Informationskompetenz“. Seine Arbeitsschwerpunkte sind Buch- und Bibliotheksgeschichte sowie die Didaktik der Informationskompetenz.

Zitation: Marcus Schröter, Historische Volltextdatenbanken, in: Clio Guide – Ein Handbuch zu digitalen Ressourcen für die Geschichtswissenschaften, Hrsg. von Laura Busse, Wilfried Enderle, Rüdiger Hohls, Gregor Horstkemper, Thomas Meyer, Jens Prellwitz, Annette Schuhmann, Berlin 2016 (=Historisches Forum, Bd. 19), <http://www.clio-online.de/guides/sammlungen/historische-volltextdatenbanken/2016>.