
Aufbau einer digitalen Forschungsbibliothek an der Universitätsbibliothek Mannheim

Christian Hänger, Irene Schumm, Stefan Weil

In den letzten Jahren ist ein zunehmendes Interesse von namhaften Ökonomen an historischen Themen und dementsprechend an historischen empirischen Daten wahrnehmbar. Beispielhaft kann hier die ifo Prussian Economic History Database (iPEHD) genannt werden, die statistische Daten für das Land Preußen während des 19. Jahrhunderts beinhaltet und eine wichtige Grundlage für darauf aufbauende wissenschaftliche Arbeiten ist.¹ Historische Börsendaten spielen für das DFG-Projekt „Empirische Untersuchung zur Ausschüttungspolitik deutscher Unternehmen“ von Erik Theissen eine zentrale Rolle. Der Mannheimer Wirtschaftshistoriker Jochen Streb untersucht mit Hilfe von wirtschaftshistorischen Daten die Rolle der Berliner Börse als Finanzquelle für innovative Unternehmen zwischen 1892 und 1913.² Ein weiteres Beispiel ist die immer wiederkehrende Auseinandersetzung mit Max Webers These von der protestantischen Ethik und dem Geist des Kapitalismus, die diese These auf der Basis von empirischen Daten zu belegen versucht.³

In den Geistes- und Kulturwissenschaften gewinnen die Digital Humanities an Bedeutung. Hierbei handelt es sich um einen interdisziplinären Forschungsansatz, in dem mit computergestützten Methoden und Werkzeugen geistes- und kulturwissenschaftliche Forschungsfragen gestellt, mit Algorithmen Dokumente ausgewertet und neue Erkenntnisse gewonnen werden.⁴ Ein Schwerpunkt liegt hier in der Begriffsgeschichte, da durch die algorithmische Auswertung von Volltexten in kurzer Zeit umfangreiche Begriffsnetzwerke analysiert werden können. In früheren Zeiten haben sich ähnliche Projekte oft über mehrere Jahre hingezogen.

¹ Wößmann 2014, S. 57–66.

² Lehmann-Hasemeyer und Streb 2016.

³ Cantoni 2015, S. 561–598.

Wichtige Quellen für die oben skizzierten Fragestellungen liegen in der UB Mannheim in gedruckter Form vor. Auch wenn die UB Mannheim sich als eine wissenschaftliche Gebrauchsbibliothek versteht, schließt das jedoch nicht aus, dass sie auch Verantwortung übernimmt für die Bewahrung und Pflege von Spezialbeständen, wie sie aus der historischen Überlieferung der Universität und der Stadt stammen. Einige dieser Bestände sind für die aktuelle wissenschaftliche Forschung – auch über den Standort Mannheim hinaus – von so großer Bedeutung, dass sie in digitaler Form aufbereitet und für die gesamte wissenschaftliche Community zur Verfügung gestellt werden sollen. So wird die Universitätsbibliothek Mannheim auch zu einer digitalen Forschungsbibliothek. Damit unterstützt die die UB Mannheim die Spitzenforschung explizit in den Wirtschaftswissenschaften und in einigen Bereichen der Geistes- und Kulturwissenschaften. In dieser Definition folgt die UB Mannheim der DFG, die einen umfassenden Begriff der Forschungsbibliothek präferiert, „der allgemein auf die wesentliche Funktion als Standort für die wissenschaftliche Arbeit mit herausragenden und umfangreichen Spezialbeständen abzielt und nicht ausschließlich auf die Einzigartigkeit der historischen Bestände“.⁵

Wir möchten in dem vorliegenden Beitrag die bisher in digitaler Form zur Verfügung gestellten Sammlungen vorstellen, auf technische Details der Präsentation eingehen und zukünftige Pläne für den weiteren Ausbau der digitalen Forschungsbibliothek skizzieren.

Hoppenstedt Aktienführer und Saling's Börsenpapiere

Saling's Börsenpapiere erschienen zwischen 1870 und 1942 und untergliedern sich in verschiedene Teile. Nach dem Zweiten Weltkrieg erschien der Aktienführer von 1956 bis 1999 jährlich in gedruckter Form – zunächst im Verlag

⁴ Definition nach DARIAH-DE 2014, S. 5.

⁵ DFG 2012. Vgl. auch Wehry 2013 mit seinem sehr umfangreichen, sich aber auf das Alte Buch fokussierenden Ansatz.

Saling, später im Verlag Hoppenstedt. Ab 1998 wurde der Aktienführer auf CD-ROM herausgegeben. Der Aktienführer enthält in standardisierter Form jährliche Berichte über deutsche und ausländische Aktiengesellschaften, die an einer deutschen Börse gehandelt werden bzw. wurden. In den Berichten finden sich u. a. Informationen und Daten zum Firmensitz und Vorstand, zu Tätigkeitsbereichen und Beteiligungen, zur Aktionärs- bzw. Kapitalstruktur und zu wesentlichen Positionen der Bilanz sowie Gewinn- und Verlustrechnung. Die konstant hohe Datenqualität sowie die lange Publikationshistorie des Aktienführers machen die Daten sehr wertvoll für die auf Deutschland bezogene empirische wirtschaftswissenschaftliche und unternehmenshistorische Forschung.

Solange die Daten nur in gedruckter Form vorlagen, mussten in allen Forschungsprojekten, die sich darauf bezogen, zunächst umfangreiche Datenerfassungsarbeiten geleistet werden. Dies geschah unkoordiniert und daher teilweise auch mehrfach parallel in verschiedenen Forschungsteams. Eine zentrale Digitalisierung und Präsentation war daher ein dringendes Desiderat nicht nur von Mannheimer Forschern. In der ersten Phase des DFG-Projekts „Aktienführer-Datenarchiv“ wurden daher zwischen 2013 und 2015 die Jahrgänge 1976 bis 1999 der gedruckten Bände digitalisiert. Neben einer reinen Erfassung des fortlaufenden Volltexts wurde dieser zusätzlich feingliedrig strukturiert und in dieser Struktur in einer Datenbank verfügbar gemacht, die sehr umfangreiche Recherche-, Export- und Auswertungsmöglichkeiten bietet. Mit dem Verlag wurde eine Vereinbarung getroffen, die es ermöglicht, die Daten Wissenschaftlern an deutschen Forschungseinrichtungen kostenfrei zur Verfügung zu stellen.

Die Rückmeldung der Wissenschaftler zu der neu entstandenen Datenbank⁶ waren äußerst positiv, jedoch wurde auch schnell klar, dass der Teilbestand von 1976 bis 1999 weder für historische noch für aktuelle oder besonders umfangreiche Auswertungen ausreichend war. Zwischenzeitlich konnten die in Mannheim vorhandenen Printbände des Aktienführers und der Saling's Börsenpapiere

⁶ <http://digi.bib.uni-mannheim.de/aktienfuehrer/>.

(Teil 2) mit Hilfe von Sondermitteln des Landes Baden-Württemberg digitalisiert werden. Im DFG-Projekt „Aktienführer-Datenarchiv II“ werden zudem von 2017 bis 2019 sowohl die nun noch fehlenden Printbände der Saling’s Börsenpapiere digitalisiert und die Daten strukturiert erfasst, als auch die Daten der CD-ROMs ab 1999 in die Datenbank integriert.

Aachener und Münchener Versicherung Aktiengesellschaft

Sitz: Aureliusstraße 2, Postfach 10, 5100 Aachen
 Telefon: (02 41) 45 61
 Telex: 8 32 872

Beteiligungen:
 I. Konzerngesellschaften
 Central Krankenversicherung AG, Köln
 Kapital: DM 16 000 000.-
 Dividenden ab 1972: 12,10,10,10,14,15 + St.G. 8,44 %
 COSMOS Allgemeine Versicherungs-AG, Aachen
 Kapital: DM 3 200 000.-
 Dividenden ab 1972: 0,0,0,0,0 %
 Cosmos Lebensversicherungs-AG, Aachen

	A	B	C	D	E	F	
	Profil-ID	Unternehmensname	Jahr	Name (Aufsichtsrat)	Vorname (Aufsichtsrat)	Ort (Aufsichtsrat)	Funktion
1	19.890.002	Aachener Rückversicherung	1989	Gies	Helmut	Aachen	Vors.
2	19.890.002	Aachener Rückversicherung	1989	Schmidt	Reimer	Aachen	stellv. Vo
3	19.890.002	Aachener Rückversicherung	1989	Heusch	Albert	Aachen/Düsseldorf	
4	19.890.002	Aachener Rückversicherung	1989	Kleyboldt	Claas	Köln	
5	19.890.002	Aachener Rückversicherung	1989	Meyhöfer	Günther	Eschweiler	
6	19.890.002	Aachener Rückversicherung	1989	Schneider	Denis	Paris	
7	19.890.002	Aachener Rückversicherung	1989	Frings	Manfred	Aachen	Arbeitneh
8	19.890.002	Aachener Rückversicherung	1989	Rosinsky	Johanna	Aachen	Arbeitneh
9	19.890.002	Aachener Rückversicherung	1989	Schayer	Marjan	Aachen	Arbeitneh
10	19.890.004	Aachener Straßenbahn	1989	Everartz	Wilhelm	Aachen	Vors.
11	19.890.004	Aachener Straßenbahn	1989	Meyer	Walter-Josef	Aachen	1. stellv. '
12	19.890.004	Aachener Straßenbahn	1989	Berger	Heiner	Aachen	
13	19.890.004	Aachener Straßenbahn	1989	Derichs	Josef	Baesweiler	
14	19.890.004	Aachener Straßenbahn	1989	Fahl	Burkhard	Aachen	
15	19.890.004	Aachener Straßenbahn	1989	Fücker	Albert	Aachen	
16	19.890.004	Aachener Straßenbahn	1989	Große	Günter	Aachen	
17	19.890.004	Aachener Straßenbahn	1989	Katzy	Dietmar	Aachen	
18	19.890.004	Aachener Straßenbahn	1989	Robert	Simon	Aachen	2. stellv. '

Abb.1: Der digitalisierte Aktienführer.

Oben: Bildigitalisat des Profils der Aachener und Münchener Versicherung AG aus dem Jahrgang 1979.

Unten: Datenexport für die Aufsichtsratsmitglieder der Jahrgänge 1989–1999.

Technisch wurde die Volltexterfassung im Projekt Aktienführer-Datenarchiv I durch ein sogenanntes Double-Keying mit zweimaligem Abtippen und Vergleichen der Ergebnisse vorgenommen, wodurch eine äußerst hohe Text- und Strukturgenauigkeit erzielt werden konnte. Im Projekt Aktienführer-Datenarchiv II werden nun eine automatisierte OCR und Strukturerrfassung erfolgen. Die Ergebnisse aus der ersten Projektphase werden dabei helfen, die automatisierten Methoden so zu trainieren, dass die Erfassungsgüte den wissenschaftlichen Ansprüchen genügt.

Deutscher Reichsanzeiger

Der Deutsche Reichsanzeiger und Preußische Staatsanzeiger ist mit seinen Vorgängerzeitungen sowie allen Beilagen (= i. F. Reichsanzeiger) eine zentrale Quelle für historische und rechtsgeschichtliche Arbeiten über Deutschland im Erscheinungszeitraum (1819 bis 1945). Im Reichsanzeiger finden sich redaktionelle Artikel, amtliche Mitteilungen und Verordnungen und zahlreiche Wirtschaftsdaten wie beispielsweise Firmen- und Handelsregister. Darüber hinaus beinhaltet der Reichsanzeiger auch viele Informationen, die über ein rein akademisches Interesse hinausgehen und für Ahnenforscher relevant sind.

Anfang des 19. Jahrhunderts wurde die Allgemeine Preußische Staats-Zeitung im Zuge der preußischen Reformen nach den napoleonischen Kriegen als amtliches Verkündungsblatt des Königreichs Preußen gegründet. Im Laufe der 127-jährigen Geschichte der Zeitung veränderten sich ihr Geltungsbereich und ihre Bezeichnung.

Den Charakter des Verkündungsblattes erhielt die Zeitschrift vor allem durch die Reduzierung der redaktionellen Teile ab Ende der 1840er Jahre. Dabei besitzt der Reichsanzeiger während der vollständigen Zeit seines Bestehens eine charakteristische Zweiteilung: am Anfang jeder Ausgabe steht ein amtlicher Teil, in welchem offizielle Mitteilungen und Verordnungen von staatlichen Stellen publiziert wurden. Danach folgen in einem nicht-amtlichen Teil

Pressemitteilungen und Artikel zu Begebenheiten im In- und Ausland. Dazu treten ab Mitte 1850er Jahre verschiedene regelmäßig erscheinende Beilagen, wie beispielsweise ein Öffentlicher Anzeiger, das Zentrale-Handelsregister, eine Warenzeichenbeilage sowie eine Börsen-Beilage.

Deutscher Reichs-Anzeiger

und

Königlich Preussischer Staats-Anzeiger.

Das Abonnement beträgt 4 R 50 S
für das Vierteljahr.
Inseratenspreis für den Raum einer Druckzeile 30 S

Alle Post-Anstalten nehmen Gefälligkeit an;
für Berlin außer den Post-Anstalten auch die Expe-
dition: SW. Willehmsstr. Nr. 52.

M. S. Berlin, Sonnabend, den 10. Januar, Abends. 1880.

<p>Se. Majestät der König haben Allergnädigst geruht: dem Ober-Steuer-Inspektor, Steuer-Rath Lange zu Wien. K. K., dem Königl. Kronen-Orden dritter Klasse; dem Kantor und Musikdirektor Seiffert an der Landesmusik- Verein im Kreise Naumburg a. S., den Königl. Kronen- Orden vierter Klasse; sowie dem herrschaftlichen Waldwärter Bernhardt zu Ködinghausen im Kreise Jerteln das All- gemeine Ehrenzeichen zu verleihen.</p> <p style="text-align: center;">Deutsches Reich.</p> <p>Se. Majestät der Kaiser haben Allergnädigst geruht: dem ersten Buchhalter der Negationskasse von Krüger den Charakter als Rechnung-Rath, sowie dem Geheimen Rangier-Sekretär von Normann den Charakter als Rangier- Rath beizulegen.</p>	<table border="1" style="width: 100%; border-collapse: collapse;"> <thead> <tr> <th>Monat</th> <th>Tag</th> <th>Art der Prüfung</th> <th>Ort</th> </tr> </thead> <tbody> <tr> <td>März</td> <td>15. und folg.</td> <td>—</td> <td>—</td> </tr> <tr> <td>—</td> <td>15.—18.</td> <td>—</td> <td>—</td> </tr> <tr> <td>—</td> <td>15.—19.</td> <td>—</td> <td>—</td> </tr> <tr> <td>—</td> <td>15.—19.</td> <td>—</td> <td>—</td> </tr> <tr> <td>—</td> <td>16.—18.</td> <td>—</td> <td>—</td> </tr> <tr> <td>—</td> <td>17.</td> <td>—</td> <td>—</td> </tr> <tr> <td>—</td> <td>18.</td> <td>—</td> <td>—</td> </tr> </tbody> </table>	Monat	Tag	Art der Prüfung	Ort	März	15. und folg.	—	—	—	15.—18.	—	—	—	15.—19.	—	—	—	15.—19.	—	—	—	16.—18.	—	—	—	17.	—	—	—	18.	—	—	<table border="1" style="width: 100%; border-collapse: collapse;"> <thead> <tr> <th>Monat</th> <th>Tag</th> <th>Art der Prüfung</th> <th>Ort</th> </tr> </thead> <tbody> <tr> <td>Sept.</td> <td>4. u. folg. Tage.</td> <td>—</td> <td>—</td> </tr> <tr> <td>—</td> <td>6.—9.</td> <td>—</td> <td>—</td> </tr> <tr> <td>—</td> <td>7.—10.</td> <td>—</td> <td>—</td> </tr> <tr> <td>—</td> <td>10. u. folg. Tage.</td> <td>—</td> <td>—</td> </tr> <tr> <td>—</td> <td>13.—16.</td> <td>—</td> <td>—</td> </tr> <tr> <td>—</td> <td>17.</td> <td>—</td> <td>—</td> </tr> </tbody> </table>	Monat	Tag	Art der Prüfung	Ort	Sept.	4. u. folg. Tage.	—	—	—	6.—9.	—	—	—	7.—10.	—	—	—	10. u. folg. Tage.	—	—	—	13.—16.	—	—	—	17.	—	—
Monat	Tag	Art der Prüfung	Ort																																																											
März	15. und folg.	—	—																																																											
—	15.—18.	—	—																																																											
—	15.—19.	—	—																																																											
—	15.—19.	—	—																																																											
—	16.—18.	—	—																																																											
—	17.	—	—																																																											
—	18.	—	—																																																											
Monat	Tag	Art der Prüfung	Ort																																																											
Sept.	4. u. folg. Tage.	—	—																																																											
—	6.—9.	—	—																																																											
—	7.—10.	—	—																																																											
—	10. u. folg. Tage.	—	—																																																											
—	13.—16.	—	—																																																											
—	17.	—	—																																																											

Abb. 2: Deutscher Reichsanzeiger Ausgabe Nr. 8 von 1880

Die Zeitung wurde mit finanzieller Unterstützung des Zentrums für Insolvenz und Sanierung an der Universität Mannheim e. V. (ZIS) digitalisiert und mit Hilfe der Gerda-Henkel-Stiftung mit Metadaten erschlossen. Aus dieser Quelle sind eine Reihe von Forschungsarbeiten am Lehrstuhl von Ulrich Falk entstanden.⁷ Die Universitätsbibliothek hat noch während der laufenden Forschungsarbeiten die Verantwortung für die Präsentation und Archivierung der Bilddateien übernommen. Ziel ist, die Bilddateien, Volltexte und eine Gliederung nach Ausgaben auf den Servern der Universitätsbibliothek für die wissenschaftliche Beschäftigung mit dieser Quelle langfristig anzubieten.⁸

⁷ Kling 2016.

⁸ <https://digi.bib.uni-mannheim.de/periodika/reichsanzeiger/>.

Eine große Herausforderung ist der Umfang der Daten für den Reichsanzeiger. So wurden die TIFF-Bilddateien für die 127 Jahre mit fast 40.000 Zeitungsausgaben auf neun Festplatten mit insgesamt rund 25 Terabyte geliefert. Da dieses Datenvolumen die vorhandenen Speichermöglichkeiten deutlich überstieg, war zunächst eine verlustfreie Konvertierung in das sparsamere Format JPEG2000 notwendig – ein Prozess, der mehrere Wochen in Anspruch nahm.

Schon früh gab es den Wunsch, die Fülle an Information für Suchtechnologien zu erschließen, indem per OCR aus den Bildern Volltext extrahiert werden sollte. Praktisch alle Artikel sind in Fraktur gesetzt und damit nicht nur für menschliche Leser, sondern insbesondere für maschinelles Lesen nicht gerade einfach zu erfassen. Erfreulicherweise liefert Standardsoftware wie Tesseract eine akzeptable Erkennungsquote und ermöglicht damit eine sinnvolle Suche.

Ancien Droit

Die UB Mannheim besitzt in der Sammlung Desbillons eine Vielzahl von Büchern, die sich mit dem vorrevolutionären französischen Recht (Ancien Droit) befassen. Die Sammlung geht auf den französischen Jesuitengelehrten François-Joseph Terrasse Desbillons (1711–1798) zurück, der ein wichtiger Vertreter der Gegenaufklärung war und sich u. a. als Fabeldichter hervortat. Seine große Leidenschaft galt dem Sammeln von Büchern, und so trug er nach den Kriterien fama und raritas eine Universalbibliothek zusammen, die alle Gebiete des damaligen Wissens umfasste. Mit dem Verbot des Jesuitenordens in Frankreich 1773 verlor er sein Amt als Oberbibliothekar am Ordenskolleg Louis Le Grand und übersiedelte mit seiner Bibliothek an den Hof von Karl Theodor in Mannheim.⁹ Die hohe wissenschaftsgeschichtliche Bedeutung dieser Bibliothek bemisst sich bereits aus ihrer Provenienz, da Desbillons einen Kernbestand der Kollegbibliothek aus Paris mit nach Mannheim nahm und nach dem Verbot des Ordens in den deutschen Territorien ebenfalls die

⁹ Ein antiphilosophisches Experiment im 18. Jahrhundert: F.-J. Terrasse Desbillons 1986, 6–14; Müller 1999, 255f.

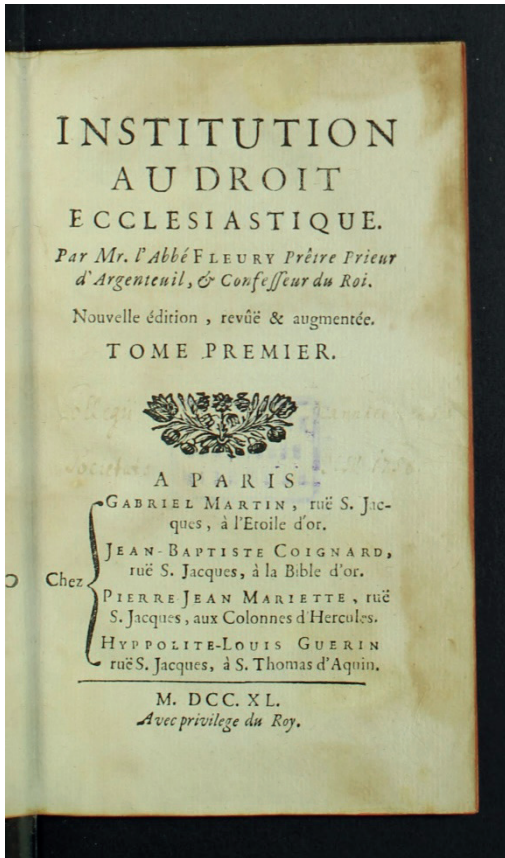


Abb. 3: Institution Au Droit Ecclesiastique

So war er Mitglied zahlreicher Kommissionen, nahm zahlreiche öffentliche Ämter ein und war als Prinzenenerzieher am Hof eingeführt. Sein literarisches Werk zeugt von einer großen thematischen Breite, die Kirchenrecht, Kirchengeschichte, historische, theologische und pädagogische Interessen umfasst.¹¹

Die Bücher der Sammlung Desbillons sind eine hervorragende Grundlage für die Durchführung eines begriffsgeschichtlichen Projekts, da der Bibliothekar und Gelehrte Desbillons die repräsentativen Werke der jeweiligen Disziplinen kannte und für seine Sammlung auswählte. In dem aktuell laufenden DFG-Projekt werden die für die Fragestellung relevanten Themen digitalisiert und bis auf

Bestände der aufgelösten Jesuitenbibliotheken aufkaufte.

Diese Sammlung befindet sich heute im Besitz der UB Mannheim und stellt einen einzigartigen, aus dem 18. Jahrhundert überlieferten Kosmos des zeitgenössischen Wissens dar.

Exemplarisch soll das für das kanonische Recht relevante Werk Institution Au Droit Ecclesiastique vorgestellt werden, das an der UB Mannheim in einer Auflage von 1740 vertreten ist.¹⁰ Der Kleriker Claude Fleury (1640–1723) war eine einflussreiche Persönlichkeit in vielen Bereichen des öffentlichen Lebens.

¹⁰ <http://nbn-resolving.de/urn:nbn:de:bsz:180-digad-12053>.

¹¹ Dictionnaire historique, 333f.

die Kapitelebene inhaltlich beschrieben. Die Bilddateien werden in einem nächsten Schritt im Volltext erschlossen. Der Textkorpus wird mit computerlinguistischen Verfahren analysiert und entsprechende Themen analysiert, die das *ancien droit* getragen haben.

Geschäftsberichte

Geschäftsberichte liefern für ein Geschäftsjahr u. a. in Bilanzen, Gewinn- und Verlustrechnung, Lagebericht und Anhang Informationen über die Tätigkeiten und Ergebnisse eines Unternehmens. Sie sind somit eine interessante und umfassende Quelle für die quantitative und qualitative Forschung in den Wirtschafts- und Sozialwissenschaften. Bibliotheken und Archive mit wirtschaftswissenschaftlichen Schwerpunkten sammelten daher lange Zeit die gedruckten Berichte.¹² So existiert auch an der UB Mannheim mit ca. 50.000 Geschäftsberichten von insgesamt gut 2.000 Firmen eine umfangreiche Sammlung an historischen Geschäftsberichten.

Die ältesten Geschäftsberichte gehen zurück bis ins Jahr 1847. Zwischen 1847 und 1947 lieferten gut 1.000 Firmen regelmäßig Geschäftsberichte, und auch nach der Währungsreform überließen zahlreiche dieser Firmen der UB Mannheim ihre Geschäftsberichte. Mit Aufkommen des Internet und der folgenden Online-Präsentation von Geschäftsberichten durch die Firmen selbst, wurde der Zugang um die Jahrtausendwende immer weniger, bis die Sammeltätigkeit im Jahr 2003 eingestellt wurde. Diese Geschäftsberichte wurden im Wirtschaftsarchiv der Handelshochschule und später der Universität Mannheim untergebracht. In den Anfangsjahren war das Wirtschaftsarchiv im Hauses Liebmann (A 3, 6) und später im Ostflügel des Mannheimer Schlosses untergebracht.

¹² Bsp. ehem. HWWA in Hamburg (heute ZBW – Deutsche Zentralbibliothek für Wirtschaftswissenschaften Kiel/Hamburg), verschiedene Institute der Universität zu Köln, Universität Hohenheim usw.



Abb. 4: Das Wirtschaftsarchiv der Handelshochschule in den 20er Jahren

Die Bestände sind derzeit in einer Excel-Tabelle verzeichnet. Neben Firmennamen, Orten, zum Teil auch Adressen und Signaturen sind auch Verweise auf Vorgänger- und Nachfolgerfirmen erfasst.

Da die gedruckten Bestände nach dem Abbruch der Sammeltätigkeit weiter Lagerfläche beanspruchten, stellte sich die Frage, ob und in welcher Form sie für die Forschung noch relevant sind. Hierzu wurde im Jahr 2015 unter der Federführung der UB Mannheim zusammen mit der ZBW und der USB Köln eine Umfrage unter Wirtschaftswissenschaftlern deutschlandweit durchgeführt. Die Rückmeldung von 147 Wissenschaftlern, die mit Geschäftsberichten forschen, zeigte, dass sie zwar häufiger mit relativ aktuellen Geschäftsberichten arbeiten (ab 1990), aber durchaus auch gesteigertes Interesse an älteren Berichten besteht. Gleichzeitig geben gut ein Drittel der Wissenschaftler an, dass es (eher) schwierig ist, an die benötigten Berichte zu kommen. Dies gilt auch für Geschäfts-

berichte aus Deutschland, insbesondere für kleinere, nicht börsennotierte Unternehmen. Kritisiert wird vor allem auch die Tatsache, dass der Zugang häufig nur über recht teure Datenbanken möglich ist, die jedoch oft nur einen relativ kurzen Publikationszeitraum abdecken (beispielsweise die letzten 10 Jahre). Eine weitere Möglichkeit, nämlich die Geschäftsberichte selbst von den Unternehmen anzufordern, scheidet häufig an der Unterstützung durch die Unternehmen. Inzwischen präsentieren viele Firmen ihre Geschäftsberichte zwar online, jedoch liegt der Fokus dabei nicht auf der mittel- oder gar langfristigen Verfügbarkeit, was aber gerade für die Forschung sehr wichtig ist. Für historische Geschäftsberichte ist daher häufig zwangsweise der Gang in verschiedene Bibliotheken nötig, was einen sehr großen Aufwand für den einzelnen Wissenschaftler bedeutet.

Es ist daher sehr naheliegend, dass Infrastruktureinrichtungen ihre Bestände mit dem umfangreichen Datenmaterial auch digital zugänglich machen, sofern dies urheberrechtlich möglich ist. Die UB Mannheim hat daher mit der Digitalisierung des Bilanzarchivs sowie dessen Erschließung in der Zeitschriftendatenbank (ZDB) begonnen. Geplant ist zudem ein kooperatives Projekt mit der ZBW und der USB Köln zum Aufbau eines digitalen Geschäftsberichte-Archivs, in dem die digitalisierten Bestände der drei Einrichtungen virtuell zusammengeführt werden sollen.

Präsentation in Kitodo

Seit 2014 präsentiert die Universitätsbibliothek ihre digitalisierten Bestände im Internet mit Hilfe der freien Software Kitodo (damals noch unter der Bezeichnung Goobi) und ist so in der Lage, neue Digitalisierungsprojekte ohne zusätzliche Lizenzkosten durchzuführen. Basis ist dabei wie bei der Website der Universitätsbibliothek das Content Management System TYPO3, was die Installation und Wartung deutlich vereinfacht.

Kitodo unterstützt dabei aber nicht nur die Präsentation, sondern auch die gesamten Arbeitsabläufe, die zur Erstellung eines Digitalisates gehören – angefangen vom Scannen bis hin zur Erfassung der Metadaten und der Bereitstellung der Dateien fürs Internet. Ursprünglich eine Entwicklung der Universitäts-

bibliothek Göttingen, wird Kitodo heute unter der Obhut des Vereins Kitodo. Key to digital objects e. V. von Bibliotheken und Firmen gemeinschaftlich gepflegt und ausgebaut. Auch die UB Mannheim engagiert sich im Verein und in der Weiterentwicklung und bringt so eigene Ideen, Korrekturen und Bedürfnisse mit ein.

OCR-Erschließung

Erst mit der Bereitstellung von Text, der durchsucht oder automatisch ausgewertet werden kann, erschließt sich der volle Wert eines digitalisierten Werkes. Die dafür notwendige Technik OCR (Optical Character Recognition) ist daher ein Desiderat für jedes Digitalisierungsprojekt. Dennoch schrecken viele vor dieser noch zu wenig bekannten Technik zurück, denn die damit verbundenen Kosten oder der Aufwand erscheinen schwer kalkulierbar.

Auch die UB Mannheim musste sich das entsprechende Wissen erst erarbeiten und identifizierte drei Softwarepakete, die für OCR geeignet erschienen: die kommerzielle Software ABBYY Finereader sowie die beiden freien Softwareprodukte OCROPY und Tesseract.

ABBYY Finereader wird von der gleichnamigen Firma ABBYY in diversen Varianten für Windows und Linux angeboten und findet sich häufig in wissenschaftlichen Bibliotheken und auch bei OCR-Dienstleistern. Für moderne Schriften (Antiqua) mit neueren Texten erzielt die Software gute Ergebnisse und ist auch recht preisgünstig. Anders sieht es für ältere Texte wie die der Sammlung Desbillons aus. Die deutschen, französischen, spanischen und vor allem lateinischen Texte des Ancien Droit sind zwar in Antiqua gesetzt, verwenden aber einige Spezialitäten wie Ligaturen und insbesondere das lange S. ABBYY unterstützt die aufgeführten historischen Sprachen mit Ausnahme von Latein. Daher muss auf verwandte Sprachen wie Italienisch oder Spanisch mit Unterstützung des langen S ausgewichen werden. Dies führt zu einer geringeren Qualität des erkannten lateinischen Textes. Dazu kommen als ein weiterer Nachteil, dass ABBYY für das OCR historischer Schriften lediglich projektbe-

zogene Lizenzen vergibt, die individuell ausgehandelt werden müssen. Dadurch entstehen für die Bibliotheken nicht kalkulierbare und häufig hohe Lizenzkosten. Schließlich verwendet ABBYY eine zeichenweise Texterkennung – eine Technik, die inzwischen als überholt gilt und durch künstliche neuronale Netze mit zeilenweiser Texterkennung abgelöst wird.

Das von Thomas Breuel an der Universität Kaiserslautern entwickelte OCRopus / OCROPY¹³ implementiert diese Technik, eignet sich gut für Experimente und enthält nützliche Werkzeuge. Allerdings hat dieses Produkt nach wie vor eher den Charakter eines Forschungswerkzeugs und ist daher sehr kompliziert in der Anwendung.

Aus Sicht der UB Mannheim liefert Tesseract¹⁴ (noch in der älteren Version 3) hinreichende Ergebnisse. Diese von Ray Smith ursprünglich bei HP und später bis heute bei Google entwickelte Software wurde an der UB Mannheim als erstes bei der Volltexterschließung des in Frakturschrift gedruckten Reichsanzeigers eingesetzt. Allerdings ergaben Hochrechnungen, dass es mehrere Jahre dauern würde, wenn man den gesamten Reichsanzeiger mit Tesseract auf einem Rechner bearbeiten wollte. Machbar wird diese Aufgabe trotzdem, indem man eben viele Rechner gleichzeitig einsetzt, also die Aufgabe parallelisiert. So konnten mit Hilfe von Rechnern im öffentlichen Bereich der Universitätsbibliothek bereits mehr als 54.000 Seiten mit Tesseract erfasst werden. Ende 2016 erschien eine neue Version von Tesseract, die jetzt auch ein neuronales Netz zur Texterkennung nutzt. Von dieser Version können zukünftig noch besser Ergebnisse erhofft werden – Versuche dazu stehen noch aus.

¹³ <https://github.com/tmbdev/ocropy> .

¹⁴ <https://github.com/tesseract-ocr/tesseract> .

Fazit

Die UB Mannheim hat mit finanzieller Unterstützung der DFG und des Ministeriums für Wissenschaft, Forschung und Kunst Baden-Württemberg in den letzten Jahren begonnen, eine digitale Forschungsbibliothek aufzubauen, die für die Wissenschaftler der Universität Mannheim und die internationale Community gleichermaßen relevant ist. Kitodo und Tesseract sind die zentralen Komponenten der technischen Infrastruktur. Die Inhalte der Sammlungen Aktienführer, Reichsanzeiger und Ancien Droit stehen bereits in digitaler Form zur Verfügung und sind weltweit häufig nachgefragt. Mit OCR generierte Texte und Daten bieten einen Mehrwert, da sie als Grundlage für eine weitere computergestützte Auswertung und Analyse verwendet werden können. Weitere für die Wissenschaft bedeutsame Quellen wie die Geschäftsberichte werden in den nächsten Jahren digitalisiert, so dass die Bedeutung der UB Mannheim als digitale Forschungsbibliothek weiter wachsen wird.

Literatur

Cantoni, Davide. „The Economic Effects of the Protestant Reformation: Testing the Weber Hypothesis in the German Lands.“ *Journal of the European Economic Association*, no. 13 (2015): 561–598.

Forschungsgemeinschaft, Deutsche. „Ausschreibung „Förderung herausragender Forschungsbibliotheken (30.04.2012)““. http://www.dfg.de/download/pdf/foerderung/programme/lis/ausschreibung_forschungsbibliotheken_110914.pdf.

Dictionnaire historique des juristes français: (XIIe–XXe siècle). 1st ed. Paris: Paris : PUF, Presses Univ. de France, 2007.

Ein antiphilosophisches Experiment im 18. Jahrhundert: F.-J. Terrasse Desbillons. Mannheim: Mannheim : Univ., Lehrstuhl Romanistik, 1986.

Kling, Christoph: Deutscher Reichsanzeiger und Preußischer Staatsanzeiger. Einleitung zur Veröffentlichung der Digitalausgabe. Mannheim (2016).

Lehmann-Hasemeyer, Sibylle H., und Streb, Jochen. „The Berlin stock exchange in imperial Germany: A market for new technology?“ *American Economic Review* 106, no.11 (2015): 3558–76.

Michael, Müller. „Die Bibliothèque Desbillons in Mannheim (1764-1789) als Beispiel französisch-jesuitischer Gelehrsamkeit im Umfeld des Mannheimer Hofes des Kurfürsten Karl-Theodor. Ein bibliotheksgeschichtlicher Beitrag zum Karl-Theodor-Jahr 1999/2000.“ *Mannheimer Geschichtsblätter* 6, (1999): S. 255–264.

Schmidt, Hans-Günter. „Kallimachos: Digital Humanities als Auftrag der Universitätsbibliothek Würzburg.“ *ABI-Technik* 36, no. 3 (2016).

Wehry, Matthias. „Die zwei Körper der Forschungsbibliothek.“ *Zeitschrift für Bibliothekswesen und Bibliographie* 60, no. 2 (2013): 70–77.

Wissenschaftsrat. Empfehlungen zu wissenschaftlichen Sammlungen als Forschungsinfrastrukturen, 2011.

Wößmann, Ludger. „iPEHD: The ifo Prussian Economic History Database.“ *Historical Methods: A Journal of Quantitative and Interdisciplinary History* 47, no. 2 (2014): 57–66.