

Till Schicketanz/Kay Heiligenhaus

»Inseln im Meer des Beliebigen«.  
Architektur und Implementierung eines Internetportals  
*Deutsch-jüdische Periodika*

*Hans Otto Horch zum 60. Geburtstag*

Abstract

Historical newspapers and journals are invaluable sources for the study of the past. Especially, Jewish periodicals provide a detailed impression of the cultural development of Jewry and shed a quite different light on German history. Yet, it is tremendously difficult to obtain and to investigate these sources: Besides »normal« losses (paper destruction et cetera), a great stock of Jewish newspapers was destroyed under the Nazi regime or are scattered all over the world – a fact, that forces a researcher to expensive stays in different libraries or archives and compels librarians to an extra burden of work. Within the funding program *Scientific Library Services and Information Systems* of the Deutsche Forschungsgemeinschaft the Aachen Chair of German-Jewish Literary History, the Cologne library Germania Judaica and the Town and University Library Frankfurt/Main carry out a digitization project, which already provides more than twenty representative periodicals, containing about 300.000 images in the WWW.

The essay emphasizes the implications and the broad range of problems as well as their solutions in the process of digitizing periodicals. A discussion of the project's technological goals demonstrates the single strategies and procedures to publish large amounts of images, text and bibliographical data in an economic and structured way.

1.

Das Informationszeitalter begrüßt den Philologen, so will es scheinen, jeden Tag mit einer Überraschung – zumindest, was die Unterstützung seiner Arbeit durch Informations- und Kommunikationstechnologien betrifft. Die Pionierleistungen der Computerphilologie – an deutschen Universitäten etwa erste, seit den späten 1960er Jahren mit Großrechnern und Lochkarten erstellte Textanalysen, Indices und Editionen – lie-

ferten die Ausgangsbasis einer technologischen Entwicklung, deren Dimension erst allmählich deutlich wird. Personal Computer, preiswerte Speicher- und Distributionsmedien, zunehmend nutzeradäquate Software, schließlich die ubiquitäre Verfügbarkeit des Internets schufen die Rahmenbedingungen eines heute kaum mehr überschaubaren Arsenal an technologischen Verfahren und Applikationen zur Informationserfassung und -distribution.

Dabei erlebt der Philologe den Anbruch dieser neuen Ära am denkbar bequemsten Ort – an seinem individuellen Computerarbeitsplatz: In zunehmendem Maß konkurrieren virtuelle Kataloge mit Bibliotheken vor Ort. Dokumentenlieferdienste versenden beliebige Texte per Mail. Exzerpte und Bibliographien verwaltet das persönliche Dokumentenmanagementsystem, und dank automatischer Texterkennungssoftware können auch umfangreichste Corpora nach Schlüsselbegriffen durchsucht werden. Last not least bietet das World Wide Web eine exponentiell wachsende Informationsbasis, die den Gang in »reale« Bibliotheken irgendwann vollends zu erübrigen scheint.

Kaum ein Bereich veranschaulicht die Technisierung und Globalisierung der Philologie so schlagend wie die weltweiten Initiativen zum Aufbau virtueller Bibliotheken. Die Perspektive ist in der Tat atemberaubend: Allein von der Deutschen Forschungsgemeinschaft wurden in den letzten Jahren circa 90 philologisch-bibliothekarische Digitalisierungsprojekte ins Leben gerufen, mit deren Hilfe langfristig die wissenschaftliche Literaturversorgung verbessert und zugleich der stetig ansteigende, aber immer schwerer zu finanzierende Arbeitsaufwand der Bibliotheken verringert werden kann. Ein Schwerpunkt der geförderten Vorhaben liegt dabei auf der »retrospektiven« Digitalisierung, die vorrangig ältere, urheberrechtsfreie Bibliotheksbestände aufbereitet und im Internet zur Verfügung stellt.<sup>1</sup> Das Spektrum ist denkbar breit und reicht von Turfan-

<sup>1</sup> Im Rahmen des »Förderprogramms Retrospektive Digitalisierung von Bibliotheksbeständen« <[http://www.dfg.de/forschungsfoerderung/wissenschaftliche\\_infrastruktur/lis/informationen\\_antragsteller/verteilte\\_digitale\\_forschungsbibliothek/retrospekt\\_digitalisierung.html](http://www.dfg.de/forschungsfoerderung/wissenschaftliche_infrastruktur/lis/informationen_antragsteller/verteilte_digitale_forschungsbibliothek/retrospekt_digitalisierung.html)> (22.1.2004). – Vgl. Jürgen Bunzel: Die Verteilte Digitale Forschungsbibliothek als Infrastrukturförderung der Deutschen Forschungsgemeinschaft. In: Hartmut Weber/Gerald Maier (Hg.): Digitale Archive und Bibliotheken. Neue Zugangsmöglichkeiten und Nutzungsqualitäten. Stuttgart: Kohlhammer 2000 (Werkhefte der Staatlichen Archivverwaltung Baden-Württemberg, Serie A: Landesarchivdirektion, 15), S. 67-82; Ewald Brahm: Digital Library Initiatives of the Deutsche Forschungsgemeinschaft. In: D-Lib Magazine 7 (Mai 2001), Nr. 7. <<http://www.dlib.org/dlib/may01/brahms/05brahms.html>> (22.1.2004); Sigrun Eckelmann: Förderschwerpunkte der DFG im Bereich digitaler Bibliotheken. Vortrag anlässlich der Sun Summit Bibliotheken und Museen, 25.9.2002, Frankfurt a. M. <<http://www.sun.de/Downloads/Praesentationen/2002/Summit-Bibliotheken/pdf/eckelmann.pdf>> (22.1.2004).

handschriften und tibetanischen Archivbeständen über Braille-Musik-Matrizen, Tonaufnahmen semitischer Sprachen und neulateinische Dichtungen bis zum Grimm'schen Wörterbuch. Ein Dutzend Vorhaben widmet sich Zeitschriften und Jahrbüchern beziehungsweise periodisch erschienenem Schrifttum.<sup>2</sup> Alle Projekte gehen im Konzept der *Verteilten Digitalen Forschungsbibliothek* auf und leisten heute bereits unschätzbare Dienste, wo es gilt, weltweit verstreute oder entlegene Bestände virtuell zusammenzuführen und den mit ihnen befassten Disziplinen einen ungehinderten Zugriff auf das Material zu bieten.

Vergleicht man die deutsche Digitalisierungslandschaft mit Angeboten aus den USA,<sup>3</sup> ist freilich festzustellen, dass der Aufbau digitaler Fachbibliotheken hierzulande gerade erst das Anfangsstadium hinter sich gelassen hat. Dies betrifft *nicht* die Auswahl teilweise recht exotischer Textcorpora oder das Faktum, dass der Bestand von fachübergreifend einschlägigen und vielgenutzten Quellen nur langsam wächst. Vielmehr sei, summiert die DFG die Entwicklung der letzten Jahre, die »Verteilte digitale Forschungsbibliothek [...] zunächst unstrukturiert gewachsen« und setzt als »Gegenmaßnahme« derzeit auf die »Einrichtung eines Portals »Sammlung digitaler Drucke«, mit dem Ziel, retrodigitalisierte Dokumente leichter auffindbar zu machen.«<sup>4</sup>

Erkennbare Defizite bestehen [...] gegenwärtig noch bei der Einbindung digitaler Angebote in vorhandene Informationssysteme, insbesondere einer über die jeweilige Besitzbibliothek hinausgehenden Erschließung der digitalisierten Bestände und einer aktiven Bekanntmachung der verfügbaren Materialien.<sup>5</sup>

Der »Eindruck einer Vielfalt unterschiedlicher Ressourcen« gipfelt in der »Unübersichtlichkeit des vorhandenen Angebots«, das »auf der Ebene

<sup>2</sup> Vgl. die »Projektübersicht für das Programm »Retrospektive Digitalisierung von Bibliotheksbeständen«: <[http://www.dfg.de/forschungsfoerderung/wissenschaftliche\\_infrastruktur/lis/gefoerderte\\_projekte/download/programm\\_retrospektive\\_digitalisierung\\_von\\_bibliotheksbestaenden.pdf](http://www.dfg.de/forschungsfoerderung/wissenschaftliche_infrastruktur/lis/gefoerderte_projekte/download/programm_retrospektive_digitalisierung_von_bibliotheksbestaenden.pdf)> (22.1.2004). – Links zu den Einzelprojekten bieten die Digitalisierungszentren München und Göttingen: <<http://www.bsb-muenchen.de/mdz/proj2.htm>> und <<http://gdz.sub.uni-goettingen.de/de/index.html>> (22.1.2004).

<sup>3</sup> Vgl. für Digitalisierungsprojekte z.B. in den USA die »Digital Initiatives Database« der *Association of Research Libraries*: <<http://www.arl.org/did/>> (22.1.2004).

<sup>4</sup> Sigrun Eckelmann: Förderschwerpunkte der DFG im Bereich digitaler Bibliotheken, S. 9 und 14. (Fußnote 1).

<sup>5</sup> Die Erschließung und Bereitstellung digitalisierter Drucke. Vorschläge des Unterausschusses für Kulturelle Überlieferung. Durch den Bibliotheksausschuss der Deutschen Forschungsgemeinschaft im Oktober 2002 verabschiedet, S. 6 <[http://www.dfg.de/forschungsfoerderung/wissenschaftliche\\_infrastruktur/lis/aktuelles/download/konzept\\_digitale\\_drucke.pdf](http://www.dfg.de/forschungsfoerderung/wissenschaftliche_infrastruktur/lis/aktuelles/download/konzept_digitale_drucke.pdf)> (22.1.2004).

der überlokalen Informationssysteme dann völlig undurchdringbar« wird, »wenn [...] eine unsystematische und weitgehend vom Zufall abhängige Auswahl der jeweils nachgewiesenen Digitalisierungsaktivitäten hinzutritt«.<sup>6</sup> Unter praktischem Gesichtspunkt hat es darüber hinaus oft noch den Anschein, dass viele Projekte technologisch bedingte Inkonsistenzen aufweisen beziehungsweise in summa keiner homogenen Architektur folgen. Ebenso scheint bislang nur wenig Einigkeit über allgemeine Produktionskriterien und -verfahren zu bestehen, die eine ökonomische, ressourcensparende Verarbeitung umfangreicher Corpora gewährleisten.<sup>7</sup> Die unter funktionalen und ergonomischen Aspekten sehr unterschiedlich gestalteten Websites deutscher Digitalisierungsprojekte lassen den Nutzer erahnen, welche Schwierigkeiten die Ermittlung und Umsetzung produktions- und designtechnischer Standards derzeit noch bereiten.<sup>8</sup>

## 2.

Ungeachtet dieser Problematik, auf die im Folgenden näher eingegangen werden soll, hat sich das Konzept der internetgestützten Informationsversorgung als integraler Bestandteil und richtungsweisender Imperativ deutscher Bibliotheks- und Bildungspolitik etabliert – allen kulturpessimistischen Unkenrufen zum Trotz, eine »zur Cyberscience hochgerüstete Wissenschaft« würde letztlich »nichts weniger als ihre Wissenschaftlichkeit ein[büßen]«.<sup>9</sup> Gehörte es vor knapp zehn Jahren fast noch zum guten Ton, die »Technisierung« der Geisteswissenschaft rundheraus abzu-

<sup>6</sup> A.a.O.

<sup>7</sup> Vgl. ebd., S. 8f. sowie Marianne Dörr: Planung und Durchführung von Digitalisierungsprojekten. In: Hartmut Weber/Gerald Maier (Hg.): Digitale Archive und Bibliotheken, S. 103-112, bes. S. 103. (Fußnote 1).

<sup>8</sup> Die Digitalisierungszentren Göttingen und München (siehe Fußnote 2) sollen in dieser Hinsicht Abhilfe schaffen. Das Kompetenzzentrum an der Universität Trier widmet sich speziell Fragen elektronischer Erschließungs- und Publikationsverfahren in den Geisteswissenschaften. Siehe folgende Adresse: <<http://www.kompetenzzentrum.uni-trier.de/index.html>> (22.1.2004). – Einen »Katalog« technischer Richtlinien hat die Landesarchivdirektion Baden-Württemberg im Zuge zweier Digitalisierungsprojekte auf der Basis entsprechender DFG-Empfehlungen (siehe Fußnote 20) zusammengestellt: »Digitalisierung von Archiv- und Bibliotheksgut« <<http://www.lad-bw.de/digpro/index.htm>> sowie »Workflow und Werkzeuge zur digitalen Bereitstellung größerer Mengen von Archivgut« <<http://www.lad-bw.de/workflow/index.htm>> (22.1.2004).

<sup>9</sup> Uwe Jochum/Gerhard Wagner: Cyberscience oder vom Nutzen und Nachteil der neuen Informationstechnologie für die Wissenschaft. In: Zeitschrift für Bibliothekswesen und Bibliographie 43 (1996), H. 6 (November/Dezember), S. 579-593. <[http://www.klostermann.de/verlegen/jochu\\_02.htm](http://www.klostermann.de/verlegen/jochu_02.htm)> (22.1.2004).

lehnen, ist der »Außenseiterstatus« der die Zeichen der Zeit verkennenden »Verächter und Verweigerer« längst evident geworden.<sup>10</sup> Kein Wissenschaftler, der die Effizienz von Personal Computer und World Wide Web erkannt hat, will fortan auf diese Arbeitshilfen verzichten. Nur die letzten standhaften Verteidiger einer überkommenen Buchdruckromantik leugnen heute noch das Faktum, dass auch Gutenbergs revolutionärer Quantensprung in erster Linie eine *technologische* Leistung darstellte, deren Ergebnis – zunächst als Teufelswerk verschrien – innerhalb kurzer Zeit fest in den frühneuzeitlichen Wissenschaftsalltag integriert wurde, diesen gar allererst konstituiert hat.<sup>11</sup>

Die Grundsatzdebatte kann angesichts jüngerer bildungs- und forschungspolitischer Entscheidungen, die sich frühen DFG-Empfehlungen anschließen und auf den nachhaltigen Ausbau der digitalen Informationsversorgung drängen, als beendet gelten.<sup>12</sup> In Anbetracht der voll im Gange befindlichen »Umgestaltung der wissenschaftlichen Informationslandschaft«,<sup>13</sup> die – wie Eli M. Noam 1995 prophezeite – die klassischen Strukturen der universitären Wissensvermittlung tief greifend wandeln wird,<sup>14</sup> gewinnt die oft als »Technophobie« getadelte Skepsis der Geisteswissenschaftler allerdings neuerlichen Auftrieb. Der Ursprung dieser pessimistisch-misstrauischen Haltung ist genauer zu lokalisieren: Als »Angst vor dem Unbekannten« oder »mangelnde Informationskompetenz bei den Nutzern«<sup>15</sup> ist diese aus Unsicherheit und Unkenntnis resultierende

<sup>10</sup> Rainer Baasner: Digitalisierung – Geisteswissenschaften – Medienwechsel? Hypertext als fachgerechte Publikationsform. In: Jahrbuch für Computerphilologie 1 (1999). <<http://computerphilologie.uni-muenchen.de/jahrbuch/jb1/baasner.html>> (22.1.2004).

<sup>11</sup> Vgl. hierzu ausführlich Michael Giesecke: Der Buchdruck in der frühen Neuzeit. Eine historische Fallstudie über die Durchsetzung neuer Informations- und Kommunikationstechnologien. Mit einem Nachwort zur Taschenbuchausgabe 1998. Frankfurt a. M.: Suhrkamp 1998 (Suhrkamp Taschenbuch Wissenschaft, 1357).

<sup>12</sup> Vgl. das DFG-Memorandum zur »Weiterentwicklung der überregionalen Literaturversorgung« <[http://www.dfg.de/aktuelles\\_presse/reden\\_stellungnahmen/download/memo.pdf](http://www.dfg.de/aktuelles_presse/reden_stellungnahmen/download/memo.pdf)> (22.1.2004) und die »Empfehlungen zur digitalen Informationsversorgung durch Hochschulbibliotheken« des Wissenschaftsrats vom 13. Juli 2001 <<http://www.wissenschaftsrat.de/texte/4935-01.pdf>> (22.1.2004) sowie das Positionspapier »Information vernetzen – Wissen vernetzen« des Bundesministeriums für Bildung und Forschung vom September 2002 <[http://www.bmbf.de/pub/information\\_vernetzen-wissen\\_aktivieren.pdf](http://www.bmbf.de/pub/information_vernetzen-wissen_aktivieren.pdf)> (22.1.2004).

<sup>13</sup> Sigrun Eckelmann: Förderschwerpunkte der DFG im Bereich digitaler Bibliotheken, S. 24. (Fußnote 1).

<sup>14</sup> Vgl. Eli M. Noam: Electronics and the Dim Future of the University. In: Science 270 (Oktober 1995), S. 247-249. <<http://www.uta.fi/FAST/JH/noam.html>> (22.1.2004).

<sup>15</sup> BMBF-Studie zur Zukunft der wissenschaftlichen Information. Pressemitteilung des Bundesministeriums für Bildung und Forschung vom 16.9.2002 anlässlich der Vorstel-

Befangenheit gegenüber neuen Technologien ein nicht zu unterschätzender Faktor, der schon im Vorfeld über Wohl und Wehe jeder Digitalisierungsinitiative entscheidet.<sup>16</sup>

Doch wer ist der »Nutzer«? Im Zeitalter des »analogen« Informationsaustausches, in der Ära der klassischen Bibliothek, lag die Antwort auf der Hand: der »Leser«, im universitären Bereich also vor allem Wissenschaftler und Studierende. In der Epoche der hybriden Bibliotheken und Archive,<sup>17</sup> die neben Schrifttum in Buch- oder Microform alle erdenkbaren Arten und Formate von digitalisierten Materialien bereitstellen müssen, wird der Bibliothekar beziehungsweise Archivar zunächst selbst zum Nutzer – an erster Stelle zum »User« technischer Apparaturen, Computeranwendungen sowie Format- und Auszeichnungssprachen zur Erfassung, Indizierung und Verbreitung von Informationen. Zwar soll nicht unterstellt werden, dass das Bibliotheks- und Archivpersonal bislang einen Bogen um den Computer gemacht hätte. Die Anforderungen, die heute an den Bibliothekar und Archivar gestellt werden, unterscheiden sich jedoch immens von Fertigkeiten, wie sie seit den 1970er Jahren hinsichtlich der EDV-gestützten Katalogisierung und Datendistribution verlangt wurden. Nicht allein der neue Medientypus, vor allem die technische Peripherie zwingt zur grundlegenden Neuorientierung:

Nicht nur die Andersartigkeit der Medien, auch die gleichzeitig veränderte Welt der Informationstechnik, die schnellen Netze, die hohe Speicherdichte und die Diversifikation der Informationsmärkte stellen unsere Vorstellungen radikal in Frage.<sup>18</sup>

Amerikanische Digitalisierungsspezialisten insistieren daher zurecht darauf, dass jeder Initiative die selbstkritische Bewertung der eigenen technischen Kompetenz vorausgehen muss:

The impulse to embrace things digital is strong, but too often infrastructure – costs, personnel, systems, and preservation – gets insufficient thought and delivery falls short of the promise. Information professionals can little afford to make mistakes in initiating and maintaining digital programs. They must assess care-

---

lung des Positionspapiers »Information vernetzen – Wissen vernetzen«. <<http://www.bmbf.de/presse01/720.html>> (22.1.2004). Vgl. Fußnote 12.

<sup>16</sup> Vgl. Hermann Leskien: Der Einfluss digitaler Medien auf die bibliothekarischen Tätigkeiten. In: Hartmut Weber/Gerald Maier (Hg.): *Digitale Archive und Bibliotheken*, S. 51-63. Hier S. 63. (Fußnote 1).

<sup>17</sup> Zur »Hybriden Bibliothek« vgl. Chris Rusbridge: *Towards the Hybrid Library*. In: *D-Lib Magazine*, Juli/August 1998. <<http://www.dlib.org/dlib/july98/rusbridge/07rusbridge.html>> (22.1.2004).

<sup>18</sup> Hermann Leskien: *Der Einfluss digitaler Medien auf die bibliothekarischen Tätigkeiten*, S. 51. (Fußnote 16).

fully the pros and cons of technology choices in a cultural context. The best way to ensure good decisions is to become a knowledgeable consumer of the technology.<sup>19</sup>

Der vernetzte Bibliothekar muss über Erfahrungen im Projektmanagement verfügen, vor allem aber technisches Wissen und Urteilsvermögen mitbringen. Der Erwerb dieser Fertigkeiten bleibt allerdings mangels längerfristiger, flächendeckender Erfahrungen im Digitalisierungsbereich heute noch meist seiner Eigeninitiative überlassen. Entsprechende Ausbildungsmaßnahmen, die normative, technische Infrastrukturen und Standards voraussetzten, stellen bislang ebenso eine Ausnahme dar wie allgemein gültige Übereinkünfte, die die praktische Durchführung von Digitalisierungsprojekten zu regeln hätten. Technisch-inhaltliche Empfehlungen, vereinzelte Projektdokumentationen oder Sammelwerke sowie neue Diskussionsforen zeigen,<sup>20</sup> dass auf zahlreiche drängende Fragen noch keine verbindlichen Antworten gefunden wurden. Abhilfe könnte eine repräsentative, didaktisch eingängige Bündelung des erreichten Kenntnisstandes nach Vorbild amerikanischer Standardwerke schaffen,<sup>21</sup> die den betroffenen Bibliothekaren, Archivaren und Wissenschaft-

<sup>19</sup> Anne R. Kenney/Oya Y. Rieger (Hg.): *Moving Theory into Practice. Digital Imaging for Libraries and Archives*. Mountain View, CA: Research Libraries Group 2000, Introduction, S. 3.

<sup>20</sup> Vgl. die »Empfehlungen zur inhaltlichen Auswahl von Bibliotheksmaterialien für die retrospektive Digitalisierung« der Facharbeitsgruppe »Inhalt« und den Bericht der Facharbeitsgruppe »Technik« im Förderbereich »Verteilte Digitale Forschungsbibliothek« <[http://www.sub.uni-goettingen.de/ebene\\_2/vdf/empfehl.pdf](http://www.sub.uni-goettingen.de/ebene_2/vdf/empfehl.pdf)> bzw. <[http://www.sub.uni-goettingen.de/ebene\\_2/vdf/endfas.pdf](http://www.sub.uni-goettingen.de/ebene_2/vdf/endfas.pdf)> (22.1.2004). – Einen vorbildlichen, in technischen Fragen freilich nicht mehr ganz aktuellen Werkstattbericht lieferten Stefan Aumann/Hans-Heinrich Ebeling/Hans-Reinhard Fricke/Manfred Thaller: *Innovative Forschung in Duderstadt. Das digitale Archiv*. Begleitheft zur Ausstellung in der Sparkasse Duderstadt, 5.-16. Mai 1997. Mit einer Beständeübersicht. Duderstadt: Mekte 1997. Vgl. folgende Adresse: <<http://www.archive.geschichte.mpg.de/duderstadt/dud.htm>> (22.1.2004). – Einen fundierten Einstieg in den Gesamtkomplex bietet der Sammelband *Digitale Archive und Bibliotheken* von Hartmut Weber/Gerald Maier (siehe Fußnote 1). – Vgl. ferner das »Digital Library Forum« der Fraunhofer Gesellschaft zur Förderung der angewandten Forschung e.V. <<http://www.dl-forum.de>> (22.1.2004).

<sup>21</sup> Anne R. Kenney/Stephen Chapman: *Digital Imaging for Libraries and Archives*. Ithaca: Cornell University Library 1996; Anne R. Kenney/Oya Y. Rieger (Hg.): *Moving Theory into Practice* (siehe Fußnote 19); weitere Literaturhinweise, ebd., S. 9. – Eine ausgezeichnete Einführung bietet das »Digital Imaging Tutorial« des Library and Research Departments der Cornell University: <[http://www.library.cornell.edu/preservation/tutorial/tutorial\\_English.pdf](http://www.library.cornell.edu/preservation/tutorial/tutorial_English.pdf)> (22.1.2004).

lern als richtungsweisender Ratgeber zu dienen hätte.<sup>22</sup> Der Gewinn einer zunehmenden ›Technosensibilisierung‹ liegt auf der Hand:

[...] digital projects are usually experimental and permit a rare and precious let's-try-it-and-see attitude. This early period of technological innovation turns you and your staff into valuable assets for your community. You have an expertise that has been dearly bought by your institution, and its investment in you and the conversion projects you manage means that you have its attention.<sup>23</sup>

Die technologischen und praktischen Widerstände wiegen desto schwerer, als die Forderung nach dem Auf- und Ausbau digitaler Archive und Bibliotheken aus Gründen der zeitgemäßen Informationserhaltung, -verwaltung und -versorgung immer lauter erhoben wird. Zwar ist zu erwarten, dass das »learning-by-doing« viele Defizite allmählich kompensiert, institutionelle Schwierigkeiten überwunden werden und sich im Zuge nationaler und internationaler Kooperationen – so genannte »Collaboratories« – einheitliche Standards etablieren.<sup>24</sup> Gegenwärtig sind die Planer jedes Digitalisierungsprojekts aber bereits im Vorfeld ihrer Anstrengungen mit einer Reihe weitreichender Fragestellungen konfrontiert, die im technischen und distributiven Bereich spezielle Design- und Produktionsstrategien nach sich ziehen.

### 3.

Die Crux jedes Projekts ist selbstverständlich die Finanzierung eines solchen Vorhabens, wobei die Aussichten seit Start des DFG-Förderprogramms zur ›Retrospektiven Digitalisierung von Bibliotheksbeständen‹ weniger trübe als in anderen Wissenschaftsbereichen sind.<sup>25</sup> Dennoch bleibt festzuhalten: Digitalisierungsprojekte sind *sehr* kostenintensiv, wobei oft übersehen wird, dass gerade der zur Realisierung der Pläne unverzichtbare technische Unter- und Überbau erhebliche finanzielle Inve-

<sup>22</sup> Vgl. Marianne Dörr: Planung und Durchführung von Digitalisierungsprojekten, S. 103. (Fußnote 7).

<sup>23</sup> Abby Smith: Real-Life Choices. In: Anne R. Kenney/Oya Y. Rieger (Hg.): Moving Theory into Practice, S. 2-3. Hier S. 3. (Fußnote 19). – Gleichlautend erklärt die DFG den »[k]onzentrierte[n] Aufbau des entsprechenden spezialisierten technischen und organisatorischen Wissens« zu einem der Hauptziele ihrer jüngsten Förderinitiativen. Vgl. Sigrun Eckelmann: Förderschwerpunkte der DFG im Bereich digitaler Bibliotheken, S. 24. (Fußnote 1).

<sup>24</sup> Vgl. Elmar Mittler: Collaboratories – auf dem Weg zu neuen Formen der technisch unterstützten Kooperation. In: Hartmut Weber/Gerald Maier (Hg.): Digitale Archive und Bibliotheken, S. 95-100. (Fußnote 1).

<sup>25</sup> Vgl. Fußnote 1.



stitutionen erfordert. Dies betrifft zum einen die Herstellung der so genannten Digitalisate, zum anderen die Finanzierung leistungsstarker Hardware. Weitaus kostenintensiver ist jedoch der Investitionsaufwand im Softwarebereich: Selbst wenn ein geeignetes Dokumentenmanagementsystem gefunden wurde, mit dem alle Arbeitsschritte von der Erfassung bis zur Präsentation der Quellen durchgeführt werden können, verursachen Software-Anpassungen stets weiteren Finanzierungsbedarf. Inwiefern die gewählte Software die anstehenden Aufgaben allerdings tatsächlich zu bewältigen vermag, erfordert wiederum technische Urteilsfähigkeit. Spitzentechnologie – hier sollten keine Zugeständnisse gemacht werden – ist teuer und erfordert professionelle Kenntnisse. Die Verwendung veralteter Systeme oder die Entscheidung, den Auf- oder Ausbau der technischen Infrastruktur »Hobbyprogrammierern« zu überlassen, heißt, im sensibelsten Bereich des Projekts eine Zeitbombe zu platzieren.

Die Auswahl des Corpus, das digitalisiert werden soll, hat sich einerseits technischen Design- und Produktionsmöglichkeiten zu beugen: Da der finanzielle Aufwand erheblich ist, wird jeder Projektträger früh Anschlussvorhaben ins Auge fassen und nur in Verfahren investieren, von denen Folgeprojekte profitieren können. Andererseits müssen verfahrenstechnische Strategiekonzepte die ökonomischen und inhaltlichen Auswahlkriterien berücksichtigen, wobei es grundsätzlich zwischen »guten« und »schlechten« Entscheidungsparametern zu unterscheiden gilt:<sup>26</sup> Ob die Digitalisierung zur Sicherung gefährdeter Bibliotheksbestände beiträgt und langfristig die herkömmliche Microverfilmung ablöst, spielt in diesem Zusammenhang eine nachgeordnete Rolle:

Digitalisierung macht die Aufbewahrung der Originale nicht überflüssig; sie löst auch nicht die konservatorischen Probleme der Bibliotheken. Aber sie erleichtert die Benutzung historischer Texte in entscheidender Weise und ist geeignet, neue Fragestellungen anzuregen.<sup>27</sup>

Entscheidend ist vielmehr, dass Projekte dieser Art unter dem Primat stehen, »durch den Einsatz digitaler Technik die wissenschaftliche Literaturversorgung zu verbessern«; sie gewährleisten den »Direktzugriff auf für die Forschung und Lehre wichtige Bestände« sowie den »Mehrfachzugriff auf vielgenutzte Literatur« und haben »die digitale Bereitstellung schwer zugänglicher Bestände« wie auch »die erweiterte Nutzung bisher nur wenig bekannter Materialien« zu sichern.<sup>28</sup>

<sup>26</sup> Vgl. z.B. den Überblick von Abby Smith: Real-Life Choices. (siehe Fußnote 19).

<sup>27</sup> Die Erschließung und Bereitstellung digitalisierter Drucke, S. 3. (Fußnote 5).

<sup>28</sup> Empfehlungen zur inhaltlichen Auswahl von Bibliotheksmaterialien für die retrospektive Digitalisierung, S. 2. (Fußnote 20).

Die praktischen Konsequenzen, die diese DFG-Prämissen für die Corpus-Auswahl nach sich ziehen, liegen auf der Hand und fordern spezielle design- und produktionstechnische Vorentscheidungen. Folgende Faktoren sind hier von ausschlaggebender Bedeutung:

- *Quantität* – Die Stärke von Hard- und Softwaresystemen besteht darin, große Datenmengen erfassen, verwalten und aufbereiten zu können. Auf Grund des finanziellen und organisatorischen Aufwands sowie der anzustrebenden, möglichst hohen Nutzerfrequenz sollte bei der Auswahl zu digitalisierender Materialien bewusst auf »Masse« gesetzt und Corpora bereitgestellt werden, die wenigstens einige zehn- oder hunderttausend Seiten umfassen. Die technische Infrastruktur muss die problemlose Erfassung und Wiedergabe großer, dabei in der Regel heterogener Datenmengen garantieren, wobei stets das oft eingeschränkte, technische Know-how der Bearbeiter zu berücksichtigen ist.
- *Simultaneität* – Es ist keine übertriebene Schätzung, dass mit Beginn der Förderung eines Digitalisierungsprojekts bis zur Präsentation erster Ergebnisse einige Jahre vergehen – ein Umstand, der weder die Projektträger noch die Nutzer befriedigt. Der technische Produktionsweg sollte daher zwei kategoriale Bedingungen erfüllen: Erstens sind die einzelnen Produktionsschritte zu modularisieren, so dass sukzessiv autarke Ergebnisse zur Verfügung gestellt werden können. Zweitens darf zwischen der primären Erfassung der Daten und ihrer Bereitstellung im Internet möglichst kein Zeitverlust entstehen. Konkret bedeutet dies einerseits, dass ein aus diversen Gründen *jederzeit* möglicher Projektabbruch der Trägerinstitution keinen allzu großen Schaden zufügt. Andererseits ist technisch zu gewährleisten, dass das Ergebnis jeder Aktion der Bearbeiter *sofort* dem Endnutzer bereitgestellt wird.
- *Ubiquität* – Selbstverständlich müssen die digitalisierten Bestände global verfügbar sein, in gleicher Qualität und Quantität wie am Erfassungsort, unabhängig von individuell genutzten Hard- oder Softwareplattformen. Im Netzwerk der *verteilten* digitalen Bibliothek, die überregionale oder internationale Kooperationen anstrebt, ist ferner technisch zu gewährleisten, dass die Datenerfassung ohne größeren Aufwand von jedem beliebigen Standort aus durchgeführt werden kann. Dies impliziert, dass die Datenbasis migrationsfähig zu sein hat: Sie muss – nicht nur aus Gründen der Langzeitsicherung – in andere standardisierte Formate und Speichermedien überführbar sein; alle

»eigenen« Daten müssen, zum Beispiel zum Zweck der Verbundrecherche, verlustfrei in »fremde« Systeme überspielt werden können wie umgekehrt die Integration von Fremddaten ins lokale System gesichert sein muss.

4.

Der Aufbau einer effizienten Produktionslinie, die große, heterogene Datenmengen ohne Zeitverzug zur ubiquitären Nutzung im Internet aufbereitet, bildet den technologischen Fokus des DFG-Kooperationsprojekts »Retrospektive Digitalisierung jüdischer Periodika im deutschsprachigen Raum« ([www.compactmemory.de](http://www.compactmemory.de)), das seit Frühjahr 2000 vom Aachener Lehr- und Forschungsgebiet Deutsch-jüdische Literaturgeschichte, dem Sondersammelgebiet Judentum der Frankfurter Stadt- und Universitätsbibliothek sowie der Kölner Bibliothek Germania Judaica durchgeführt wird. Im Verlauf von sechs Jahren soll der Großteil der seit Ende des 18. Jahrhunderts im deutschen Sprachraum erschienenen jüdischen Zeitschriften, Zeitungen und Jahrbücher erschlossen und bereitgestellt werden. Das Vorhaben schließt eine gravierende Lücke, die bislang die Arbeit der Jüdischen Studien maßgeblich erschwerte: Einerseits bilden die rund 5.000 jüdischen Periodika, die seit dem 17. Jahrhundert weltweit erschienen, ein gewaltiges, gar nicht zu überschätzendes Quellenreservoir der jüdischen Geschichte und Kultur.<sup>29</sup> Über drei Jahrhunderte versuchten jüdische Periodika alle wissenschaftlichen, beruflichen, literarischen, pädagogisch-didaktischen beziehungsweise geistigen Bedürfnisse ihrer Leser zu befriedigen, wodurch sie zu einem kulturhistorisch einmaligen »Archiv« wurden, das sämtliche religiösen, politischen, sozialen und kulturellen Richtungen innerhalb des Judentums dokumentiert.<sup>30</sup> Andererseits sind die erhaltenen Bestände, vor

<sup>29</sup> Vgl. G[eorg] H[erlit]z/M[endel] P[robst]: Presse, jüdische. In: Georg Herlitz/Bruno Kirschner (Hg.): Jüdisches Lexikon. Ein enzyklopädisches Handbuch des jüdischen Wissens in vier Bänden. Berlin: Jüdischer Verlag 1927-1930, Bd. IV/1, Sp. 1102-1110 <[http://www.compactmemory.de/project/doku02\\_link.html](http://www.compactmemory.de/project/doku02_link.html)> (22.1.2004) und David Flinker/Shalom Rosenfeld/Mordechai Tsanim (Hg.): The Jewish Press That Was. Accounts, Evaluations and Memories of Jewish Papers in Pre-Holocaust Europe. Jerusalem: Jerusalem Post Press 1980.

<sup>30</sup> Zum historisch-wissenschaftlichen Stellenwert jüdischer Periodika vgl. Hans Otto Horch/Till Schicketanz: »Ein getreues Abbild des jüdischen Lebens«. *Compact Memory* – Ein DFG-Projekt zur retrospektiven Digitalisierung jüdischer Periodika im deutschsprachigen Raum. In: Menora. Jahrbuch für jüdische Geschichte 12 (2001), S. 387-405; dies.: *Compact Memory* – Ein Projekt zur retrospektiven Digitalisierung jüdischer Periodika im deutschsprachigen Raum. In: Michael Nagel (Hg.): Zwischen

allem infolge der Verluste im Zweiten Weltkrieg und der systematischen Zerstörungen der Nazis, in alle Himmelsrichtungen zerstreut und vollständige Jahrgänge nur an wenigen Bibliotheken erhalten. Wegen ihres schlechten Erhaltungszustandes gelangt das Gros der Originale längst nicht mehr in den Leihverkehr, so dass interessierte Forscher und Laien zu häufigen und kostspieligen Bibliotheksreisen gezwungen sind, was angesichts der hohen Nutzungsfrequenz jüdischer Periodika auch für das Bibliothekspersonal einen erheblichen zusätzlichen Arbeits- und Zeitaufwand bedeutet.

Die erste Forderung bestand folglich darin, die gemäß ihrer historischen Bedeutung, heutigen Nutzungsfrequenz und technischen Tauglichkeit ausgewählten Periodika dem Nutzer per Internet am individuellen Arbeitsplatz zur Verfügung zu stellen. Die Präsentation des insgesamt rund 700.000 Seiten umfassenden Corpus sollte einen zugleich ökonomischen, ergonomisch sinnvollen und intuitiven Zugriff gestatten – Kriterien mithin, die in Anbetracht der enormen Menge zu digitalisierender Analogdaten selbstverständlich auch für die bibliothekarische Erfassung der Daten gelten.<sup>31</sup> Zudem war zu garantieren, dass dem User erste Ergebnisse ohne Zeitverzug in Form eines strukturierten Archivs navigier- beziehungsweise skalierbarer Grafiksammlungen bereitgestellt werden können.

---

Selbstbehauptung und Verfolgung. Deutsch-jüdische Zeitungen und Zeitschriften von der Aufklärung bis zum Nationalsozialismus. Hildesheim u.a.: Olms 2002 (Haskala, 25), S. 351-359.

<sup>31</sup> Die Auswahl umfasst u.a. folgende Periodika: *Allgemeine Zeitung des Judenthums* (1837-1922), *Altneuland* (1904-1906), *Berliner Vereinsbote* (1895-1901), *C.V.-Zeitung* (1922-1938), *Der Israelit* (1860-1938), *Der Jude* (1832-1833), *Der Jude* (1916/17-1924), *Der Morgen* (1925/26-1938), *Der Orient* (1840-1851), *Die Freistadt* (1913/14), *Die Welt* (1897-1914), *Esra* (1919/1920), *Im deutschen Reich* (1895-1922), *Israelitische Rundschau* (1901-1902), *Jahrbuch für jüdische Geschichte und Literatur* (1898-1931, 1936-1938), *Jeschurun* (1854/55-1869/70; 1883-1888; 1914-1930), *Jüdische Rundschau* (1902-1938), *Menorah* (1923-1932), *Mitteilungen des Gesamtarchivs der Deutschen Juden* (1908/09-1914/15, 1926), *Monatsschrift für Geschichte und Wissenschaft des Judentums* (1851/52-1887; 1893-1939), *Neue jüdische Monatshefte* (1916/17-1919/20), *Ost und West* (1901-1923), *Palästina* (1902-1938), *Sulamith* (1806/08-1845/48), *Wissenschaftliche Zeitschrift für jüdische Theologie* (1835-1847), *Zeitschrift für Demographie und Statistik der Juden* (1905-1923; 1924-1931), *Zeitschrift für die Geschichte der Juden in Deutschland* (1887-1892; 1929/30-1937). – Bis 2006 sollen insgesamt 120 Periodika zugänglich gemacht werden.

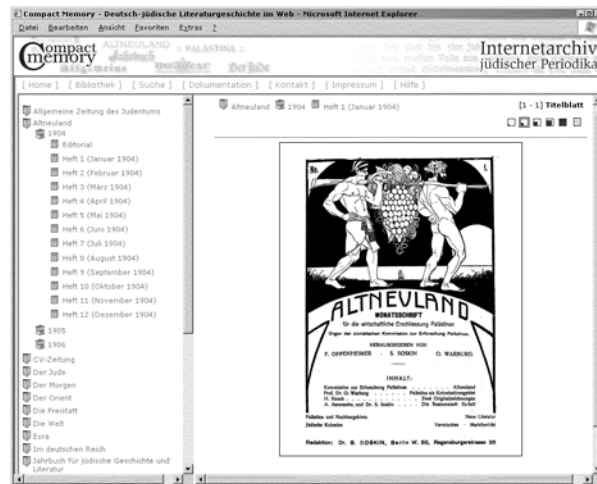


Abb. 1: Auswahl und Anzeige eines Zeitschriftenheftes

Wie in einer »realen« Bibliothek wählt der Besucher zunächst das Periodikum über einen Navigationsbaum aus, um über den gewünschten Jahrgang zur gesuchten Nummer zu gelangen (siehe Abb. 1). Die Anzeige der Images, in denen man wie im papierenen Original »blättern« kann, erfolgt in den gängigen Grafikformaten; separate Optionen dienen der Thumbnailansicht, der Vergrößerung beziehungsweise Verkleinerung sowie dem Druck oder Download der Images.

Die Bereitstellung der Grafiken erfolgt mit dem spezifisch auf den Bedarf von Digitalisierungsprojekten zugeschnittenen Produkt *Visual Library* der Firma *semantics*. Diese Softwareplattform ermöglicht mit den Modulen *Library Manager* und *Library Scout* die strukturierte Erfassung, Indizierung, Volltexterkennung, Bearbeitung und Bereitstellung beliebiger grafischer und textueller Materialien im Internet.<sup>32</sup> Dabei setzt der *Library Manager* als zentrales Arbeitswerkzeug das Digitalisierungsteam in den Stand, große Mengen von Grafiken übersichtlich und schnell auf einen lokalen Datenbankserver zu überspielen (circa 1.000 Images pro Stunde). Ein Vorschaufenster zeigt wahlweise den Inhalt des Quellverzeichnisses an, aus dem Images per Drag-and-Drop ins Zielverzeichnis kopiert werden. Im integrierten Grafikbetrachter werden die Images einzeln oder in Form von Thumbnails aufgerufen und in einem Arbeitsgang von Schattierungen oder Verschmutzungen gereinigt. Über einen Navigationsbaum, der die serverinterne Zielverzeichnisstruktur

<sup>32</sup> Nähere Angaben zum Funktionsumfang der *Visual Library* unter <[http://www.semantics.de/produkte/visual\\_library/](http://www.semantics.de/produkte/visual_library/)> (22.1.2004).

abbildet, legt der Bearbeiter neue, annotierbare Zeitschriftentitel, Jahrgänge oder Hefte an, wobei jedes Image zudem typisierbar ist (Titelblatt, Inhaltsverzeichnis, Artikel und so weiter). Mit diesen Arbeitsschritten stehen die erschlossenen Materialien unmittelbar unter Verwendung des *Library Scouts* in Form dynamisch generierter Webseiten zu Recherchezwecken zur Verfügung.

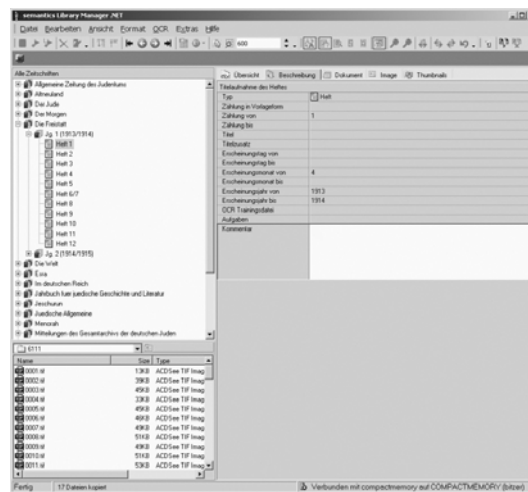


Abb. 2: Library Manager – Einspeisung der Images ins lokale Verzeichnissystem



Abb. 3: Library Manager – Imageansicht

Ein Vorteil dieses ersten Produktionsschritts besteht darin, große Mengen digitalen Quellenmaterials ohne Umschweife im Internet zur Verfügung zu stellen. Die Datenerfassung folgt ergonomischen und ökonomischen Prinzipien und setzt keine besonderen technischen Kenntnisse voraus. Darüber hinaus ist diese Phase der Produktion vollkommen autark: Mit geringem Aufwand lassen sich auch weltweit verstreute Materialien in Form virtueller Gesamtbestände zentralisieren, auf die fortan global zugegriffen werden kann. Der Umstand, dass in diesem Stadium keine digitalen Volltexte angeboten werden und das Archiv erst oberflächlich strukturiert ist – im Fall von Periodika gemäß ihrer »natürlichen« Hierarchie (Titel, Jahrgang, Einzelheft) –, erweist sich nur auf den ersten Blick als Nachteil: Waren zuvor aufwändige, oft erfolglose Bibliotheksreisen, Archivaufenthalte oder Bestellvorgänge nötig, wird der Nutzer die bloße Verfügbarkeit bislang schwer zugänglicher Corpora – und sei die Erschließungstiefe vorläufig noch so gering – als ungemeine Arbeitserleichterung begrüßen. Bibliotheken und Archive wiederum werden es zu schätzen wissen, knappe Personalressourcen schonen und die bedrohten Originale schützen zu können.

Diese einfache Bereitstellungsform, die als erster, selbstständiger Produktionsschritt angestrebt werden sollte, mag in manchen Fällen bereits vollkommen genügen – sie stellt jedenfalls ein vergleichsweise einfach, günstig und schnell zu erzielendes Arbeitsergebnis dar, das Nutzer und Anbieter gleichermaßen entlastet. Die geringe Komplexität dieses Verfahrens, das sich leicht auf andere Publikationstypen oder Überlieferungsformen applizieren lässt, mag ferner ein Argument für Institutionen darstellen, die bislang keine Erfahrungen im digitalen Bereich gesammelt haben, dieses Segment jedoch aus Gründen der »Selbsterhaltung« rasch besetzen sollen:

Es steht den Geisteswissenschaften nicht mehr frei, sich aus den wandlungsinintensiven Bedingungen der Kommunikation herauszuhalten. Und wenn dies [...] im Stillen oder lautstark gefordert wird, dann um den Preis der Selbstmarginalisierung. [...] Was jetzt im Internet als Wissensbestand und Geltungsanspruch nicht angemessen markiert wird, kann mittelfristig bereits von der Weltkarte der geläufigen Kenntnisse verschwunden sein. Es gerät, wenn es für eine computergestützte Benutzung nicht in mediengerechter Form zur Verfügung steht, an den Rand jenes Feldes, das als Raum des allgemein Wissenswertes betrachtet werden kann.<sup>33</sup>

---

<sup>33</sup> Rainer Baasner: Digitalisierung – Geisteswissenschaften – Medienwechsel? (Fußnote 10).

Sofern das Corpus eine tiefere Erschließungsebene erfordert oder die entsprechende Nachfrage besteht, sollten sich Digitalisierungsvorhaben nicht darauf beschränken, Quellen als »Loseblattsammlungen« anzubieten. Eine wichtige Vorgabe, die die Attraktivität eines Digitalisierungsprojekts sichert, besteht bekanntlich darin, dass die Effizienz des Zugriffs auf das Textcorpus die hergebrachten Möglichkeiten des Buchs oder der Microform übersteigen sollte. Dies betrifft vor allem die Recherchemöglichkeiten.

## 5.

Keinesfalls wollen die Nutzer bei jeder neuen Fragestellung immer wieder eine Unzahl von Grafiken nach den gesuchten Materialien durchsuchen – eine mühsame und zeitraubende Prozedur, die schon die Arbeit mit Microfilmen oder papierenen Vorlagen erschwerte. Die Minimalerwartung der Benutzer besteht selbstverständlich darin, die den analogen »Originalen« entsprechenden digitalen »Kopien« einsehen zu können. Darüber hinaus will der User direkt auf zugehörige bibliographische Kerndaten zugreifen sowie in den digitalen Volltexten recherchieren. Als separater Produktionsschritt, der von der Einspeisung der Grafiken ebenso wie von der Erfassung der Volltexte getrennt werden sollte, mag die Erschließung der bibliographischen Daten relativ unaufwändig und unproblematisch sein – allerdings nur, sofern es sich um Monographien handelt: Hinter einem eindrucksvollen Archiv von abertausend Seiten verbergen sich oft nur einige Hundert Titelaufnahmen, die eventuell längst erfasst wurden oder kurzfristig katalogisiert werden können.<sup>34</sup> Die normkonforme Katalogisierung unselbstständig erschienener Literatur erfordert hingegen einen weitaus höheren Arbeitsaufwand, den angesichts chronischer Ressourcenverknappung kaum noch eine Bibliothek aufzubringen im Stande ist. Möglicherweise wird man in Zukunft wenigstens die Titel des laufend in Fachzeitschriften, Jahrbüchern und Sammelwerken erscheinenden Schrifttums digital erfassen können – unter

---

<sup>34</sup> Vgl. das »Schwesterprojekt« von *Compact Memory*, das digitale Archiv *Jiddische Drucke* der Stadt- und Universitätsbibliothek Frankfurt a. M. unter der folgenden Adresse: <<http://www.literatur-des-judentums.de>> (22.1.2004) sowie die Dokumentation von Rachel Heuberger: Die Bestände der Judaica-Sammlung auf dem Weg ins Internet. Zwei Digitalisierungsprojekte an der Stadt- und Universitätsbibliothek Frankfurt a. M. In: *Tribüne* 39/154 (2000); vgl. folgende Adresse <<http://www.stub.uni-frankfurt.de/publikationen/tribuene.htm>> (22.1.2004).



dem Kriterium der »Realität des Leistbaren«<sup>35</sup> kann diese Aufgabe *retrospektiv* wohl auch langfristig nur im Ausnahmefall erbracht werden. Das statistische Verhältnis zwischen selbst- und unselbstständigen Publikationen verdeutlicht das Problem: Repräsentiert die Titelaufnahme einer Monographie ungefähr 200 bis 300 Seiten, umfasst – gemäß den Erfahrungen von *Compact Memory* – ein Beitrag aus einem historischen Periodikum durchschnittlich kaum vier bis fünf Seiten Text. Ein Zeitschriftencorpus von circa 500.000 Seiten würde demnach die Erfassung von mehr als 100.000 Artikeln erforderlich machen, während im Fall von Monographien nur rund 2.000 Einträge anfielen.

Stellt die Erfassung unselbstständigen Schrifttums, die einem Standard wie zum Beispiel den »Regeln für die alphabetische Katalogisierung in wissenschaftlichen Bibliotheken« (RAK-WB) beziehungsweise den bislang nur als Entwurf vorliegenden »Regeln für die alphabetische Katalogisierung unselbstständiger Werke« (RAK-UW) folgen sollte, nicht eine herkulische Leistung dar? Ist es unter ökonomischen Gesichtspunkten überhaupt zu rechtfertigen, dass sich ein Bibliotheksteam über Jahre dieser Aufgabe widmet – wohl wissend, dass das Ergebnis nur einen Tropfen auf dem heißen Stein ausmacht? Lange vor Anbruch des digitalen Informationszeitalters wurde die Forderung laut, dass verstärkt auch unselbstständiges Schrifttum katalogisiert werden müsse. In den vergangenen Jahrzehnten übernahmen teilweise Fachbibliographien diese Aufgabe. Mitte der 1990er Jahre folgten entsprechende Internetangebote, die sich aus nahe liegenden Gründen zumeist auf die laufend neuerscheinenden, hauptsächlich naturwissenschaftlich-technischen Fachzeitschriften konzentrieren.<sup>36</sup> Die *retrospektive* Katalogisierung historischer Bestände ist hingegen sicher nicht grundlos immer wieder aufgeschoben oder nur im Einzelfall angegangen worden.

In diesem Zusammenhang lautet die zentrale Frage vor allem, ob sich die Mühe in Zeiten der zunehmend effizienter arbeitenden Texterkennungsprogramme überhaupt lohnt: Ohne Zutun eines Bibliothekars könnte ein umfangreiches Corpus von Grafiken automatisch texterkannt und in digitalen, das heißt recherchierbaren Volltext umgewandelt werden. Im Ergebnis differenzierte das System zwar nicht zwischen distinkten, bibliographischen Einheiten wie »Autor«, »Titel« und so weiter, was

<sup>35</sup> Marianne Dörr: Planung und Durchführung von Digitalisierungsprojekten, S. 106. (Fußnote 7).

<sup>36</sup> Vgl. Initiativen einzelner Bibliotheken bzw. Bibliotheksverbünde vor allem das Kooperationsprojekt JADE, das die Recherche nach ca. 24.000.000 Aufsätzen aus rund 42.000 Fachzeitschriften ermöglicht, die über den kostenpflichtigen Dokumentlieferdienst JASON bestellt werden können.

eine – wiederum aufwändige – Nachindizierung der Texte voraussetzte. Der Nutzer wäre aber dennoch in der Lage, nach bestimmten Zeichenfolgen zu recherchieren – eben auch solchen, die zum Beispiel einen Autornamen oder den Titel eines gesuchten Beitrags repräsentieren.

Die Entscheidung, bibliographische Kerndaten manuell zu katalogisieren, bleibt unter wirtschaftlichen beziehungsweise technischen Gesichtspunkten stets anfechtbar. Man kann darüber spekulieren, ob künftig neue Technologien die klassische Form der Katalogisierung obsolet machen werden. Von dieser Entwicklung, die keinesfalls eine Zukunftsvision darstellt, einmal abgesehen, sind die neuen, hochinformativen Möglichkeiten jedoch an einem traditionellen Kriterium zu messen: Demzufolge basiert der »Wert« eines digitalen Archivs nicht ausschließlich auf der Zweckmäßigkeit, die ein solches Angebot für einen bestimmten Nutzerkreis besitzt, wie im Übrigen ja auch die Bedeutung einer traditionellen Bibliothek keinesfalls nur in der Literaturversorgung besteht. Gerade die »Erschließung alter und wertvoller Bestände«, die »von gesamtstaatlicher oder überregionaler Bedeutung« sind,<sup>37</sup> gilt zurecht als maßgebendes Förderkriterium, sofern dadurch der eigentliche Mehrwert jedes Einzelarchivs innerhalb der Verteilten Digitalen Forschungsbibliothek konstituiert wird. Den aus Einzelprojekten resultierenden Datenbanken ist folglich ein bleibender kulturhistorischer Stellenwert zu Eigen, der nicht unterschätzt werden kann: Wie ein Buch, das unkatalogisiert in eine Bibliothek eingestellt wurde, für den Nutzer schlichtweg nicht existiert, stellt erst die distinkte Titelaufnahme eines Zeitschriftenbeitrags die initiale Materialisierungsstufe seines potentiellen Informationsgehalts dar. Die *Summe* aller Titelaufnahmen bildet die Voraussetzung der optimalen *Informationsvermittlung*; der *Zweck* der Katalogisierung besteht indes ebenso in der reinen *Informationserhaltung*. – Das Zukunftsszenario mag erschrecken, unrealistisch ist es keineswegs: Wenn die Originale eines Tages zu Staub zerfallen sind und einige Jahrzehnte oder Jahrhunderte später die archivierten Microfilme ausgebleicht sein werden, geben beizeiten konvertierte Datenbanken wenigstens darüber Auskunft, welche Informationen der Menschheit verloren gingen.

In den weltweit aktiven Jüdischen Studien muss die Titeldatenbank von *Compact Memory* zudem das Kriterium erfüllen, auch vom Grafikbeziehungsweise Volltextangebot unabhängige Recherchen zu ermöglichen. Als digitales »Quellenverzeichnis« schließt die Datenbank die große Lücke zwischen biographischen Nachschlagewerken, Fachbibliographien

---

<sup>37</sup> Sigrun Eckelmann: Förderschwerpunkte der DFG im Bereich digitaler Bibliotheken, S. 7. (Fußnote 1).

sowie einschlägigen Lexika und Enzyklopädien, wobei der Vorteil darin besteht, dass die Daten online verfügbar sind – nötigenfalls auch auf anderen Plattformen, in fremden Informationsverbünden oder auch als Referenzorgan in Buchform. Dies setzt voraus, dass die bibliothekarischen Kerndaten standardisiert und vollständig katalogisiert werden, wobei die Dateneingabe – um die Masse halbwegs zu bewältigen – überregional erfolgen sollte und keine Redundanzen aufweisen darf.

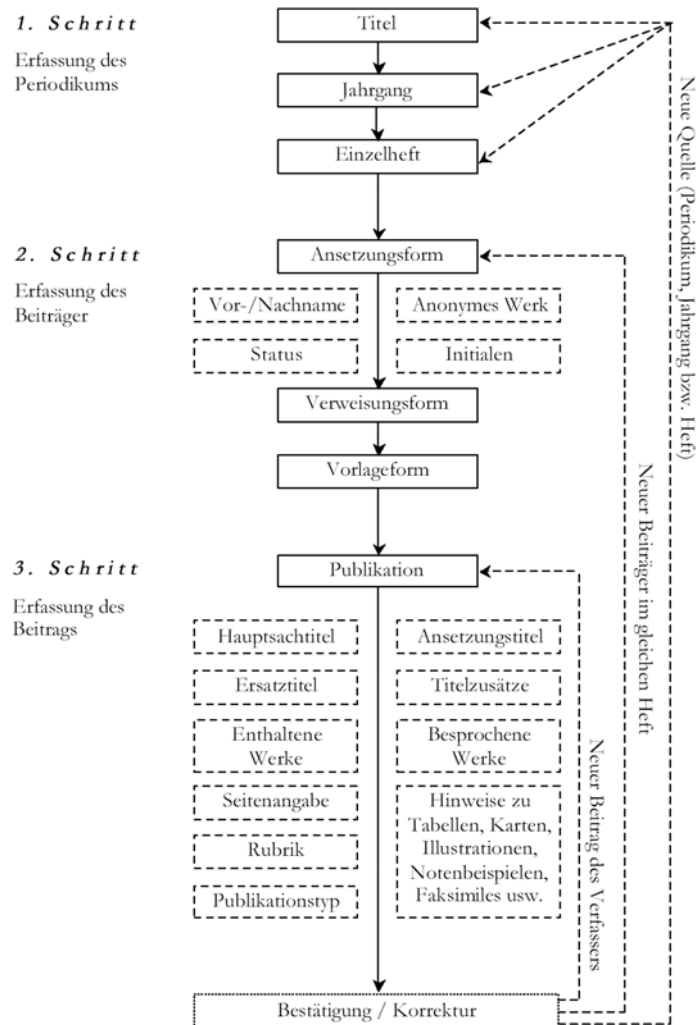


Abb. 4: Produktionsschema zur Erfassung bibliographischer Kerndaten

Die Erfassung der bibliographischen Kerndaten sieht laut Schema (siehe Abb. 4) drei aufeinander folgende Produktionsstufen vor, die mittels eines im Projekt entwickelten, webbasierten Eingabeinterfaces durchlaufen werden: Der Aufnahme des Periodikums folgt zunächst die Erfassung beziehungsweise Auswahl des Jahrganges und des Einzelheftes, dessen Beiträge katalogisiert werden sollen (siehe Abb. 5). Dem Eintrag dieser gestaffelten Quellenvermerke, die nur einmal vorgenommen werden müssen, um an allen Arbeitsstandorten abrufbar zu sein, schließt sich als zweiter Schritt die RAK-konforme Aufnahme der an der Publikation beteiligten Personen oder Körperschaften an.

Abb. 5: Erfassung eines Zeitschriftenheftes

Die technische Realisation dieses Schritts stellte eine der komplexesten Aufgaben dar (siehe Abb. 6): Neben anonymen, nur mit Initialen versehenen oder von Körperschaften autorisierten Beiträgen müssen vor allem die im osteuropäischen Bereich häufig variierenden Namensschreibweisen berücksichtigt werden, ohne redundante Mehrfacheinträge für dieselbe Person zu erzeugen, welche die Homogenität des Datenmodells beeinträchtigen. In verschärfter Form tritt diese Problematik bei Pseudonymen zu Tage, die als solche vom Bibliothekspersonal oft nicht oder nur zufällig zu erkennen sind. Führt die »Ansetzungsform« in der Regel den Geburtsnamen auf, geben die zugeordneten »Verweisungsformen« die Pseudonyme oder wechselnden Namensschreibungen des Beiträgers wieder, wobei man die Angaben jederzeit separat erweitern oder Zuordnungen revidieren kann.

Der Bearbeiter muss ferner zur Aufnahme eines neuen Beiträgers alle bereits erfassten Ansetzungsformen sowie die zugehörigen Verweisungsformen durchsuchen können. Im Gegenzug galt es sicherzustellen, dass auch der Nutzer bei der Autorenrecherche eine vollständige Ergebnis-

menge der zugewiesenen Publikationen erhält – also auch diejenigen Artikel, die der Autor unter Pseudonym oder variierender Schreibweise seines Namens veröffentlichte. Da zudem zwei oder mehr Autoren beziehungsweise Körperschaften für einen Beitrag verantwortlich zeichnen können, muss die Personenerfassung beliebig oft wiederholbar sein, wobei es sich als sinnvoll erwies, den eventuell unterschiedlichen »Status« der Urheber zu verzeichnen (Verfasser, Illustrator, Übersetzer und so weiter). Der Datenbankserver, auf den die Clients zugreifen, verhindert auch bei diesem Arbeitsschritt redundante Doppel- oder Mehrfacherfassungen, da sämtliche Aufnahmen oder Änderungen unmittelbar allen beteiligten Standorten zur Verfügung stehen.

http://www.compactmemory.rwth-aachen.de/cm/erfassung/personen.asp?Zeitschrift\_ID=18;Jahrgang\_1...

Datei Bearbeiten Ansicht Favoriten Extras ?

Digitalisierung jüdischer Periodika Bibliothekarische Erfassung in Frankfurt und Köln

Erfasste Daten

Zeitschrift Der Jude 1 (1916/1917) 1 (4. 1916)  
Ansetzungsformen Keine

Ansetzungsformen

Nachname Buber  
Vorname Martin  
Status Verfasser  
☐ Initialen ☐ Anonymes Werk

Erfasste Ansetzungsformen

Keine Ansetzungsformen aufgenommen.

OK Abbrechen

Fertig Internet

Abb. 6: Erfassung eines Zeitschriftenbeitrags

Die Titelaufnahme schließt als dritter und letzter Produktionsschritt die Erfassung ab (siehe Abb. 7). Neben den Angaben zum Hauptsachtitel, zum Ansetzungstitel, einem eventuellen Ersachtitel oder diversen Titelnzusätzen beziehungsweise obligatorischen Hinweisen zur Fundstelle (Rubrik, Seitenangaben) erschien es praktisch, die betreffenden Beiträge wenigstens ansatzweise zu verschlagworten: So werden im Fall von Rezensionen die besprochenen Werke in Kurzform verzeichnet; ebenso zentral für wissenschaftliche Recherchen, zum Beispiel für Fragen der Kanonbildung, ist die Möglichkeit, in Sammelbeiträgen enthaltene Werke, zum Beispiel Gedichte verschiedener Verfasser in einer Zeitschriftenanthologie, erfassen zu können. Die Treffermenge lässt sich darüber hinaus mit Hilfe des Publikationstyps (Beitrag, Rezension, Nachricht, Illustration und so weiter) oder der bereits zugewiesenen Rubrik (Leitarti-

kel, Gemeindenachrichten, Feuilleton und so weiter) eingrenzen. Angaben zu Tabellen, Karten, Abbildungen, Notenbeispielen und so weiter liefern weitere Hinweise.

Abb. 7: Titelaufnahme

Bei der Entwicklung von Arbeitsoberflächen dieser Form und Funktionalität sind drei Überlegungen von entscheidender Bedeutung: (1) In Anbetracht der großen Datenmengen müssen während der Katalogisierung jeglicher Zeitverlust vermieden und redundante Tätigkeiten neutralisiert werden. (2) Die Arbeiten dürfen an die Clients keine speziellen Hard- oder Softwareanforderungen stellen beziehungsweise den bibliothekarischen Nutzern keine tieferen technischen Kenntnisse abverlangen und müssen simultan von verteilten, das heißt letztlich beliebigen Standorten aus durchgeführt werden können. (3) Die erfassten bibliothekarischen Kerndaten sind dem Nutzer unmittelbar nach jeder einzelnen Titelaufnahme in strukturierter Form im Internet zur Verfügung zu stellen.

Wie angedeutet, wurde aus ökonomischen Gründen besonders das Prinzip der »Rekursivität« berücksichtigt: Da das Bibliothekspersonal die Katalogisierung – von Nachträgen oder Korrekturen abgesehen – gemäß der Druckreihenfolge der Einzelbeiträge durchführt, »merkt« sich das System das zuletzt aufgerufene Einzelheft, den übergeordneten Jahrgang sowie das zugehörige Periodikum. Der erste Produktionsschritt – die Erfassung des Periodikums, Jahrganges und Einzelheftes – ist gestaffelt, wodurch der Arbeitsaufwand minimiert wird: Die rekursiven Quellenan-

gaben sind erst zu aktualisieren, wenn der Bearbeiter in ein neues Heft, einen neuen Jahrgang oder ein neues Periodikum wechselt. Die zentrale Verwaltung aller personen- und körperschaftsbezogenen Angaben, die nach einmaliger Erfassung an allen Standorten in Form von Auswahllisten zur Verfügung stehen, optimiert die Ergonomie und Effizienz des Produktionsverfahrens. Zugleich wird auf diesem Weg die einheitliche Datenerfassung und die Homogenität des Datenbestands gesichert – ein Faktor, dem vor allem in einem Projekt mit verteilten Standorten immense Bedeutung zukommt. Insgesamt konnten auf diese Weise innerhalb von knapp drei Jahren circa 60.000 Einzelbeiträge beziehungsweise rund 6.000 Personen- und Körperschaftsangaben katalogisiert und zur Recherche freigegeben werden (Stand: Januar 2004).

Ein webbasiertes Eingabeinterface mag im Vergleich zu einer fest am individuellen Arbeitsplatz installierten Softwarelösung einige Nachteile aufweisen. So sind zum Beispiel der technische Funktionsumfang und die ergonomischen Möglichkeiten eines lokalen Erfassungstools weniger limitiert, als dies bei einer Eingabeplattform der Fall ist, die über einen Internet-Browser angesteuert wird. Unter Umständen spricht jedoch gerade diese Alternative für ein Web-Interface: Die Datenbank kann nicht nur an jedem Ort der Welt genutzt werden – dies leistet auch jede moderne Erfassungssoftware –, die beteiligten Personen und Institutionen können vor allem ohne jedwede Anpassung ihres lokalen Systems arbeiten. Ein Wechsel des Rechnertyps, den eine Software eventuell voraussetzt, ist ebenso unnötig wie etwaige Umstellungen oder Aktualisierungen der individuellen Betriebssysteme. Die retrospektive Katalogisierung von Bibliotheks- und Archivbeständen kann mit Hilfe internetbasierter Eingabeinterfaces quasi voraussetzungslos von verteilten Standorten aufgenommen werden, wobei auch der anfallende Entwicklungs- und laufende Wartungsaufwand vergleichsweise gering wäre und keine Lizenzgebühren anfielen. Der Preis, der im Zuge einer solchen Entscheidung zu entrichten ist, besteht im Verzicht, *alle* in einem Projekt anfallenden Aufgaben auf *einer* integralen Plattform zu lösen. Grafiken und Volltexte mit einem Tool, bibliographische Kerndaten hingegen mittels einer Internetmaske in die Datenbank einzuspeisen, bedeutet, dass im Ergebnis kongruente Arbeitsabläufe separiert werden. Die später erforderliche Synthese der verschiedenen Datenmengen ist unter verfahrenstechnischen Aspekten selten ohne Reibungsverluste zu bewerkstelligen, wobei der vielleicht nur geringfügig höhere Arbeitsaufwand weniger schwer wiegt als die Preisgabe der technischen Homogenität und Effizienz.

Die Diskussion der Vor- und Nachteile, die eine konkrete technologische Fragestellung nach sich zieht, verdeutlicht einen entscheidenden Punkt: Digitalisierungsinitiativen stecken ein Terrain ab, auf dem gegenwärtig noch intensiv »experimentiert« werden muss, um effiziente und verbindliche Produktionskonzepte zu entwickeln.<sup>38</sup> Digitalisierungsprojekte entwerfen, erproben und evaluieren Design- beziehungsweise Produktionsstrategien, um ihre Erfahrungen in einem langfristigen, nachhaltigen Digitalisierungsprogramm aufgehen zu lassen, welches den sukzessiven Aufbau »digitale[r] themenorientierte[r] Informationsnetze«<sup>39</sup> forciert und somit das Rückgrat des geplanten DFG-Portals »Sammlung digitalisierter Drucke« darstellen könnte. Dieses Experimentierfeld wurde in *Compact Memory* bewusst abgeschritten – die Entscheidung indessen, welche technische Alternative gewählt wird, basiert im Kern auf der Kompetenz und Bereitschaft der beteiligten Institutionen, neue, prototypische Technologien in bestehende Systeme zu integrieren, um dadurch den Aufbau benutzerorientierter Informations-Infrastrukturen voranzutreiben.

Welche Produktionsvariante im skizzierten Fall letzten Endes bevorzugt wird, hängt einerseits von den ins Auge gefassten Projektzielen, von den verfügbaren Ressourcen und nicht zuletzt von der technologischen Kompetenz des Mitarbeiterstabs ab. Andererseits verdeutlichen die Ausführungen zur bibliothekarischen Datenerfassung, dass die gewählte Lösung dem Gebot der Ökonomie zu folgen hat und die Datenbasis migrationstauglich sein muss, um weltweit in Form überregionaler, internationaler Gateways einen standardisierten Zugang zu ermöglichen.<sup>40</sup>

---

<sup>38</sup> So auch das Fazit von Marianne Dörr: Planung und Durchführung von Digitalisierungsprojekten, S. 110 (Fußnote 7): »Digitalisierungsprojekte sind eine neue Aufgabe der Bibliotheken und Archive, vor allem handelt es sich – und das wird oft unterschätzt – um eine sehr komplexe Aufgabe. Damit sind Fehler unvermeidbar und vermutlich muss eine Reihe von Fehler gemacht werden, bis auch die deutschen Bibliotheken und Archive praktikable Checklisten für Digitalisierungsprojekte aufstellen können.«

<sup>39</sup> Sigrun Eckelmann: Förderschwerpunkte der DFG im Bereich digitaler Bibliotheken, S. 19. (Fußnote 1).

<sup>40</sup> Vgl. im Zusammenhang der retrospektiven Zeitschriftendigitalisierung das Projekt *Dieper* (Digitised European Periodicals) der Niedersächsischen Staats- und Universitätsbibliothek Göttingen, in dessen Rahmen ein standardisierter Zugriff auf digitale Zeitschriftenbestände innerhalb und außerhalb Europas ermöglicht werden soll. <<http://gdz.sub.uni-goettingen.de/dieper/>> (22.1.2004).



6.

Im Allgemeinen interessiert es den Nutzer nicht, auf welche Weise Daten in eine Datenbank gelangen – entscheidend ist für ihn, wie das erfasste Material aufbereitet und im Internet zur Verfügung gestellt wird. Dabei erweist sich der Grad, in dem das Corpus in formaler und inhaltlicher Hinsicht erschlossen wurde, als ebenso zentraler Faktor wie die ergonomische und taktile Funktionalität der Zugriffsmöglichkeiten. Die Attraktivität eines digitalen Archivs steigt folglich in dem Maß, wie es den individuellen Arbeitsgewohnheiten seiner Nutzer entgegenkommt und traditionelle Wege der Informations- und Literaturbeschaffung erleichtert.

Von der Volltextsuche abgesehen, sind es im Fall der bibliographischen Recherche in einem digitalen Zeitschriftencorpus im Wesentlichen drei typische Suchstrategien, welche die Anbieter berücksichtigen müssen: In der Regel will der Nutzer über gängige Suchoptionen gezielt und ohne Verzug bestimmte Materialien aufrufen, deren Quellenangaben ganz oder teilweise bekannt sind (Simple Search). Will man Suchergebnisse einschränken oder liegen nur rudimentäre Hinweise vor, müssen mittels kombinierter Suchmöglichkeiten hierarchisch organisierte Trefferlisten generiert werden können (Advanced Search). Zuletzt darf nicht ignoriert werden, dass viele Nutzer im Bestand »stöbern« möchten: Wie der Besucher einer realen Bibliothek mal zu diesem, mal zu jenem Band greift, klickt der Nutzer eines digitalen Archivs mal diesen, mal jenen Link an, um sich von Zufallsfunden überraschen zu lassen oder in bekannten Kontexten gezielt zu lesen.

Es empfiehlt sich daher grundsätzlich, dass ein digitales Archiv seinen *gesamten* Datenbestand in strukturierter Form visualisiert – vor allem um den Nutzern einen Überblick über den Umfang, die Vollständigkeit und die Erschließungstiefe des Textcorpus zu vermitteln. Endloslisten, die bibliographische Daten nach singulären Kriterien aufführen, erweisen sich als unpraktisch und unübersichtlich. Dagegen wird der Einstieg zweifellos erleichtert, wenn das Vorwissen der Nutzer bezüglich des Umgangs mit bestimmten Textcorpora berücksichtigt wird und die Präsentation des digitalen Bestands der analogen »Urform« des Mediums folgt: Aus Erfahrung »weiß« der Nutzer, dass er in einer Bibliothek zunächst ein Periodikum auswählt und dann zu einem bestimmten Jahrgang greift. Er hat »gelernt«, dass ein Jahrgang eventuell ein Inhaltsverzeichnis enthält, auf jeden Fall aber eine beliebige, prinzipiell chronologisch geordnete Anzahl Einzelhefte umfasst; erst in den Heften erwartet der Nutzer, auf »Text« in Form einzelner Artikel zu stoßen.

Der Medienwechsel muss nicht notwendigerweise eine Umstellung internalisierter Gewohnheiten beziehungsweise praxiserprobter Strukturierungsformen bedeuten: Ein digitales Archiv sollte dieses erworbene Vorwissen vielmehr kreativ umsetzen und dem Nutzer – neben diversen Suchfunktionen – stets auch den intuitiven, sozusagen »plastischen« Zugriff auf das Textcorpus ermöglichen. Dementsprechend wurde in *Compact Memory* eine Präsentationslösung angestrebt, die die Recherche, Navigation und Orientierung im Corpus erleichtert, indem alle Daten mit Hilfe des *Library Scouts* bis in die Einzelhefte hinein visualisiert werden (siehe Abb. 8).

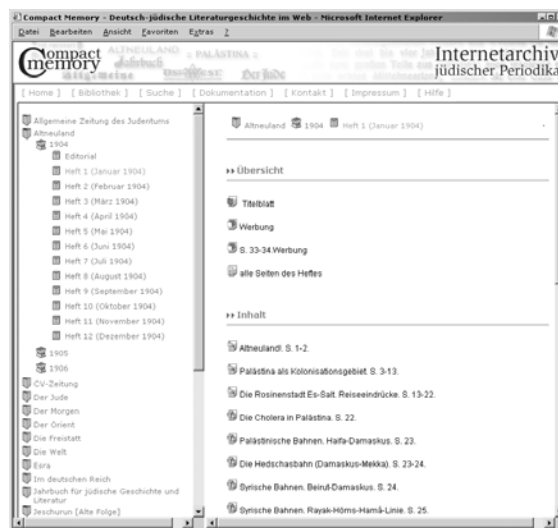


Abb. 8: Anzeige des Inhalts eines Einzelheftes

Gegenüber der Titelrecherche bildet die Volltextsuche einen logisch und arbeitsteilig weitgehend eigenständigen Aufgabenbereich, der im Anschluss an die Präsentation der Digitalisate und die Erfassung bibliographischer Kerndaten die dritte, separate Produktionsstufe darstellt. Die Problematik ist allgemein bekannt: Nur in seltenen Fällen lässt sich vom Titel eines Beitrags auf dessen Inhalt oder historischen Stellenwert schließen; bedeutende Beiträge tragen ausdruckschwache Überschriften oder können von mehr oder minder unbekannten Verfassern stammen. Fördert die Suche nach Titelschlagworten also stets ein Konglomerat relevanter und peripherer Angaben zu Tage, vermag erst die Volltextrecherche die Spreu vom Weizen zu trennen, indem die Textinhalte nach bestimmten Zeichenfolgen durchsucht und statistisch aufbereitete Trefferlisten generiert werden.

Die Umwandlung der Grafiken in Volltexte bedeutet einen erheblichen Mehraufwand, auch wenn dieser Arbeitsschritt mit Hilfe automatisierter OCR-Software durchgeführt wird. Zwar sind Texterkennungsprogramme heute deutlich leistungsstärker, leichter zu handhaben und preiswerter als vor einigen Jahren. In einem für die *retrospektive* Digitalisierung zentralen Punkt stoßen aber auch beste OCR-Programme an ihre Grenzen: Kann Schrift in Antiqua, gute Vorlagenqualität vorausgesetzt, in der Regel mit Trefferquoten von annähernd 100 Prozent erkannt werden, erfordert die bis in die 1920er Jahre im deutschen Sprachraum weit verbreitete Frakturschrift bislang einen beträchtlichen manuellen Trainingsaufwand, um halbwegs zufrieden stellende Ergebnisse zu erzielen. Die fortschreitende OCR-Entwicklung wird hier über kurz oder lang Abhilfe schaffen, vor allem sobald die Softwareindustrie dieses Marktsegment entdecken sollte. Bis dahin wird jedes Digitalisierungsprojekt individuell zu entscheiden haben, ob der zu erbringende Aufwand in einem vertretbaren Verhältnis zum Nutzen steht, wodurch letztlich immer auch die CorpUSAuswahl betroffen ist.

So zentral wie die Frage nach der erzielbaren *Qualität* der Volltexte ist das Problem, wie man die enorme *Quantität* an Text auf ökonomische Weise bewältigt. Allein aus Gründen der Ergonomie und Übersichtlichkeit können mehrere zehn- oder hunderttausend Seiten nicht einfach einem separat arbeitenden OCR-Programm zugeführt und im Anschluss manuell in eine Datenbank überführt werden. Die Einspeisung der Grafiken in das hierarchisch strukturierte Verzeichnissystem und die Zuweisung der seitenweise erzeugten Volltexte sind vielmehr als logisch parallel laufende Produktionsstufen zu organisieren. Zur Vermeidung von Reibungsverlusten sollte es demnach möglich sein, die Erfassung der Grafiken und die Erzeugung beziehungsweise Verknüpfung der zugehörigen Volltexte nicht von getrennt arbeitenden Programmen, sondern auf einer integralen Plattform durchzuführen.

Der in *Compact Memory* eingesetzte *Library Manager* wurde zu diesem Zweck um ein OCR-Modul erweitert,<sup>41</sup> das einen auszuwählenden Bestand von Grafiken – komplette Periodika, bestimmte Jahrgänge, einzelne Hefte oder Seiten – in Stapelverarbeitung in digitalen Text überführt. Die resultierenden Textdateien werden in der Datenbank automatisch den entsprechenden Grafiken zugeordnet und stehen unmittelbar

---

<sup>41</sup> Genutzt wird derzeit die *FineReader 6.0 Engine* des ABBYY Software House. Vgl. die URL: <[http://www.abbyy.com/developer\\_toolkits.asp?param=2395](http://www.abbyy.com/developer_toolkits.asp?param=2395)> (22.1.2004). – Die Ende 2003 releaste Betaversion 7.0 soll standardmäßig auch Frakturschrift erkennen können. Bis Redaktionsschluß konnten keine Tests durchgeführt werden.

der Volltextrecherche zur Verfügung. Der *Library Manager* bietet ferner alle zentralen Funktionen des OCR-Programms, darunter die Möglichkeit, besondere Zeichensätze wie Fraktur zu trainieren. Mit Hilfe des integrierten Editors können die vorliegenden Texte darüber hinaus nach Bedarf redigiert sowie TEI-konform im XML-Format ausgezeichnet werden.<sup>42</sup>

## 7.

Der vorliegende Werkstattbericht versuchte anhand konkreter Erfahrungen und Fragestellungen, die im Zuge der digitalen Zeitschriftenreformation auftreten, ein Produktionskonzept zu skizzieren, dessen Schwerpunkt in der verzugsfreien, halbautomatisierten und standardisierten Massendigitalisierung historischer Drucke besteht. Vor allem der souveräne Umgang mit heute verfügbaren Technologien zur Massendigitalisierung bildet die Voraussetzung, eine Zukunftsvision zu verwirklichen, die auch verantwortliche Institutionen nicht mehr hinter vorgehaltener Hand diskutieren:

Der insgesamt – auch international – erreichte Stand der retrospektiven Digitalisierung lässt jedoch heute die Vision realistisch erscheinen, dass in einer oder zwei Generationen *die gesamten historischen Buchbestände des Landes*, ergänzt durch entsprechende Digitalisate handschriftlicher, bzw. nichtschriftlicher Teile des kulturellen Erbes über eine einheitliche Oberfläche vom Schreibtisch jedes und jeder Interessierten direkt und ohne nennenswerte Zeitverzögerung zugänglich sein könnten. Auch eine konservative Hochrechnung technischer Entwicklungen lässt erwarten, dass ein derartiges Ziel in einigen Jahrzehnten erreicht werden kann.<sup>43</sup>

Ausländische Großinitiativen,<sup>44</sup> aber auch viele prototypische Lösungen deutscher Einzelprojekte demonstrieren, dass dieser gleichsam enzyklo-

<sup>42</sup> Zu TEI, dem standardisierten Auszeichnungssystem der *Text Encoding Initiative* auf Basis der Extensible Markup Language (XML) vgl. vor allem die TEI-Homepage <<http://www.tei-c.org/>> (22.1.2004). Eine Einführung bietet Fotis Jannidis: TEI in der Praxis. In: editio 11 (1997), S. 152-177. <<http://computerphilologie.uni-muenchen.de/praxis/teipraxis.html>> (22.1.2004).

<sup>43</sup> Die Erschließung und Bereitstellung digitalisierter Drucke, S. 18. (Fußnote 5); unsere Hervorhebung, T.Sch./K.H.

<sup>44</sup> Vgl. z.B. das *Gallica*-Projekt der Bibliothèque National (<<http://gallica.bnf.fr/>> (22.1.2004)), das *Cervantes*-Projekt spanischer und lateinamerikanischer Bibliotheken (<<http://cervantesvirtual.com/proyectoEN/BIMICESA.shtml>> (22.1.2004)), die Kordic Digital Newspaper Library *Tiden* (<<http://tiden.kb.se>> (22.1.2004)) oder das

pädische Auftrag heute bereits zu bewältigen ist. Die Hürden, die im Verbund deutscher Bibliotheken, Archive und Universitäten noch genommen werden müssen, sind weniger technischen als organisatorischen Ursprungs. Drei Aspekte stehen dabei im Vordergrund: (1) Schaffung eines zentralen Internet-Portals, das einen fachübergreifenden Zugriff auf vorhandene digitale Drucke ermöglicht; (2) Synchronisation und Evaluation laufender beziehungsweise zukünftiger Digitalisierungsprojekte auf Basis eines zu erstellenden, verbindlichen »Kriterienkatalogs«; (3) Planung und Prioritätensetzung der Digitalisierung historischer Bibliotheks- und Archivbestände im Rahmen eines nationalen Gesamtkonzepts.<sup>45</sup>

In Erweiterung der so genannten Sondersammelgebiete stellt das Konzept der Virtuellen Fachbibliothek zweifellos die wichtigste Alternative dar, potentielle Nutzer mittels »qualitätsgesicherte[r] Erschließungs- und Zugangssysteme« über einschlägige Ressourcen wie digitale Sammlungen zu informieren.<sup>46</sup> Angesichts der zunehmenden Spezialisierung der Einzelwissenschaften ist vor allem die Virtuelle Fachbibliothek das probate Mittel, »dem Benutzer einen einigermaßen umfassenden Nachweis der in Deutschland verfügbaren digitalen Bestände geben und ihm den Zugriff darauf ermöglichen« zu können.<sup>47</sup> Ebenso unerlässlich wie die einheitliche Anwendung technischer und methodischer Standards, welche die nachhaltige Nutzung solcher Gesamtsysteme garantieren, erfordert der Aufbau Virtueller Fachbibliotheken die Synthese von bibliothekarischer, fachwissenschaftlicher und technologischer Kompetenz. Allein die aktive Zusammenarbeit zwischen Informationsanbietern, Nutzern und Systemkonstrukteuren kann ein Angebot schaffen, das die Interessen aller Beteiligten zu berücksichtigen vermag.

Dem weltweit aktiven Forschungszweig »Jüdische Studien« würde eine nationale wie internationale Initiativen synchronisierende Fachbibliothek unschätzbare Dienste leisten. Eine solche Virtuelle Forschungsbibliothek wäre in der Lage, sämtliche Zeugnisse jüdischer Tradition in sich zu vereinen – historische, literarische und wissenschaftliche Primärtexte, musikalische oder grafische Quellen sowie Nachlässe jedweder Provenienz. Darüber hinaus könnten Enzyklopädien, Nachschlagewerke, Bibliographien, Kataloge und Verzeichnisse, aber auch Wörterbücher und Periodika zur Verfügung gestellt werden. Die hierzu erforderlichen, interna-

---

*American Heritage Project* der Library of Congress (<<http://sunsite.berkeley.edu/amher/>> (22.1.2004).

<sup>45</sup> Die Erschließung und Bereitstellung digitalisierter Drucke, S. 2, 5 u.ö. (Fußnote 5).

<sup>46</sup> Ebd., S. 6. Vgl. die Homepage der *Virtuellen Fachbibliothek* <<http://www.virtuelle-fachbibliothek.de/>> (22.1.2004).

<sup>47</sup> Die Erschließung und Bereitstellung digitalisierter Drucke, S. 8. (Fußnote 5).

tional kooperierenden Initiativen besäßen in der Forschungsbibliothek aber nicht nur einen virtuellen »Lesesaal«. Sie bildeten eine globale Plattform, die einerseits der Erfassung und Verbreitung von benötigten Textcorpora dient. Andererseits entstünde ein Forum zur Präsentation von Forschungserträgen und -initiativen, das weitaus aktueller als herkömmliche Printmedien sein könnte.