

Digitalisierungsprojekte, Datenbankportale und Literaturmanagement in den Geisteswissenschaften Anforderungsprofile – Softwarelösungen – Realisation

Kay Heiligenhaus / Till Schicketanz

Ich könnte mir vorstellen, daß die größte Wirkung von Computern nicht darin bestehen wird, Raketen zum Mars zu steuern, Buchhaltungsirrtümer auszuschließen oder Herztransplantationen so genau zu überwachen, daß nichts mehr schiefgehen kann. Nein, die größte Wirkung des Computers wird darin bestehen, daß er einen neuen Typus Mensch schaffen kann – einen Fragensteller.

John Sculley

Datenbanksysteme werden in der Praxis gemeinhin als ›Hilfsmittel‹ angesehen, vordefinierte Ziele, deren Umsetzung auf eine komplexe Masse heterogener Informationen rekuriert, effizienter zu erreichen. Ob dies die Bedienung eines Mobiltelefons oder jenen Raumflug zum Mars betrifft – stets oszilliert die Absicht zwischen dem Bedürfnis, dem Individuum das Leben angenehmer zu gestalten, und dem Wunsch, der Menschheit vermeintliche Quantensprünge zu ermöglichen. Wenige Jahrzehnte nach Inbetriebnahme der ersten Großrechneranlagen haben Datenbanken unseren industriestaatlichen Alltag bis ins Privatleben hinein längst vollständig infiltriert. Das bedeutet im Umkehrschluss freilich, dass jeder, der ›konkurrenzfähig‹ werden oder bleiben möchte, sich dieses Werkzeugs bedienen muss, um ›Wettbewerbsvorteile‹ zu behaupten. Dies gilt für alle Bereiche, in denen Datenbanken sinnvoll Anwendung finden können, insbesondere für die Wissenschaften. Die Frage nach dem »Sinn und Nutzen von Datenbanken in den Geisteswissenschaften« stellt sich also nur rhetorisch – verschwiegen werden soll jedoch nicht, dass sich der Geisteswissenschaftler mit einem klaren Ja zu dieser Technologie, neben etlichen Vorteilen, sogleich auch eine ganze Reihe von Herausforderungen einhandelt.

Zieht man allgemeine Betrachtungen zurate, was eine Datenbank diesseits aller Technik darstellt, verwundert es zunächst, warum ihr Nutzen in den Geisteswissenschaften erst vergleichsweise spät erkannt wurde. Als »data structure used to store organized information«¹ bzw. »collection of data, or information, that is specially organized for rapid search and retrieval«² scheinen Datenbanken wie

¹ University of Mary Washington Style Guide: Appendix VI. Computer Terms and E-mail Etiquette,
(http://www.umw.edu/policies/style_guide/computer_terms_email_etiq/default.php)
(05.07.2006)

² Encyclopædia Britannica Online, Artikel »Database«,

geschaffen dafür, das ureigene Terrain des Geisteswissenschaftlers zu beackern, nämlich systematisch die Bestandteile unseres Weltwissens zu katalogisieren, zueinander in Beziehung zu setzen, Gesetzmäßigkeiten aufzudecken und somit letztlich neues Wissen zu generieren. In diesem Kontext jedoch – weniger in der hartnäckigen Technophobie der Disziplin – findet sich zugleich eine der Hauptursachen, warum Datenbanken im Arbeitsalltag der Geisteswissenschaften ihr Schattendasein noch nicht gänzlich überwinden konnten: Das, was die Welt im Innersten zusammenhält, die Relationen zwischen den geistigen ›Elementarteilchen‹ wie auch deren Verhältnis zum Ganzen ist im höchsten Maß ambig und bedarf stets der erneuten Deutung. Unser Weltwissen als immaterieller Untersuchungsgegenstand der Geisteswissenschaften setzt sich nicht zusammen wie eine Formelsammlung oder ein Telefonbuch. Das Ganze ist mehr als die Summe seiner Einzelteile. Datenbanken in den Geisteswissenschaften liefern keine *Antworten*, wie sie dies millionenfach in alltäglichen Anwendungsbereichen leisten – sie helfen uns letzten Endes nur, unsere *Fragen* präziser zu formulieren. Jede gegenteilige Hoffnung läuft zwangsläufig auf das bekannte Szenario hinaus:

»Er ... good morning, O Deep Thought,« said Loonquawl nervously, »do you have ... er, that is ...«

»An answer for you?« interrupted Deep Thought majestically. »Yes, I have.« [...]

»To Everything? To the great Question of Life, the Universe and Everything? [...] Tell us!« [...]

»Forty-two,« said Deep Thought, with infinite majesty and calm. ...

»Forty-two!« yelled Loonquawl. »Is that all you've got to show for seven and a half million years' work?«

»I checked it very thoroughly,« said the computer, »and that quite definitely is the answer. I think the problem, to be quite honest with you, is that you've never actually known what the question is.«³

Zu verlockend ist die Aussicht, eine Datenbank könne auch nur eine der ›großen‹ Fragen beantworten, die uns tagtäglich umtreiben. Aber betrachten wir einmal vorab, was Datenbanken in diesem Bereich heute schon leisten, welches Potential mittelfristig in ihnen steckt und stellen uns die Frage, ob der konsequente, flächendeckende Einsatz der verfügbaren Technologien nicht schon mehr ist, als wir erhoffen können.

(<http://www.britannica.com/eb/article-9029424>) (05.07.2006)

³ Douglas Adams: *The Hitch Hiker's Guide to the Galaxy*. London: Pan Books 1979, S. 181.

Datenbanken *speichern* Informationen. Dies nicht etwa in Analogie zu Keilschrifttafeln, Papyrusrollen, Pergamenten, Quipus oder Büchern – sie repräsentieren vielmehr die folgerichtige Weiterentwicklung dieser klassischen Datenspeicher der Menschheit. Zugleich werden die Informationen *gesichert* – im Gegensatz zu »database« hebt die deutsche Bezeichnung dies hervor –, wodurch Datenbanken herkömmlichen analogen Speichermedien deutlich überlegen sind, denn gerade ihre Immaterialität erweist sich theoretisch als unschlagbare Stärke: Was physisch nicht existent, aber per Mausklick jederzeit duplizierbar ist, kann auch keinen Schaden nehmen – vorausgesetzt, alle nötigen Vorkehrungen wurden getroffen.⁴ Ihre ›Unkörperlichkeit‹ verleiht Datenbanken darüber hinaus einen weiteren, unschätzbaren Vorteil: Sie speichern und sichern Material und Informationen in einer Erschließungs- und Präsentationsform, die in analoger Form gar nicht gegeben ist. Insbesondere retrospektive Digitalisierungsprojekte, die historische Quellencorpora über das Internet zugänglich machen, illustrieren diesen wichtigen Aspekt des datenbankgestützten Speicherns, Sicherns und Kreierens von Materialsammlungen.

Als Beispiel mag das Internetarchiv jüdischer Periodika »Compact Memory«⁵ dienen, das auf seinen Webseiten einen strukturierten Zugriff auf die digitalen Abzüge von über 100 deutsch-jüdischen Periodika des 18., 19. und 20. Jahrhunderts anbietet. Insgesamt steht ein etwa 750.000 Originalgraphiken umfassendes Zeitschriftencorpus zur Verfügung, das in unterschiedlicher Erschließungstiefe als Quellenreservoir alle relevanten Lebensbereiche des deutschsprachigen Judentums der Neuzeit dokumentiert. Die Bezeichnung ›Archiv‹ unterstreicht das besondere Merkmal des Angebots: »Compact Memory« stellt nicht nur eine nachhaltige Speicher- und Sicherungsplattform des digitalen Materials dar – das Archiv ist darüber hinaus vollkommen *virtuell*; es bündelt Quellen und kreiert eine Sammlung, die in dieser Vollständigkeit an keinem Ort der Welt realiter existiert. Gesammelt, erschlossen und präsentiert werden Daten und Digitalisate vieler hundert Zeitschriftenjahrgänge, die bislang über Dutzende in- und ausländische Bibliotheken und Archive verstreut waren, nicht mehr in den

⁴ Vgl. zum Komplex der ›Nachhaltigkeit‹ z. B. die Arbeit des Kompetenznetzwerkes Langzeitarchivierung *nestor* (<http://www.langzeitarchivierung.de>) (05.07.2006).

⁵ Das von 2000 bis 2006 von der Deutschen Forschungsgemeinschaft geförderte Digitalisierungsprojekt (www.compactmemory.de) wurde in Kooperation des Lehr- und Forschungsgebiets Deutsch-jüdische Literaturgeschichte an der RWTH Aachen, der Universitätsbibliothek Frankfurt/Main sowie der Bibliothek Germania Judaica e.V., Köln realisiert. Die technische Unterstützung erfolgte von Seiten der semantics Kommunikationsmanagement GmbH. Vgl. Till Schicketanz / Kay Heiligenhaus: »Inseln im Meer des Beliebigen«. Architektur und Implementierung eines Internetportals Deutsch-jüdische Periodika. In: Jahrbuch für Computerphilologie 5 (2003), S. 67–96 (<http://computerphilologie.uni-muenchen.de/jg03/schicketanz-heiligenhaus.html>) (25.05.2006).

Leihverkehr gelangen, nun aber als einmalig komplettierter Bestand zentral abgerufen und recherchiert werden können.

»Compact Memory« verdeutlicht eine weitere zentrale Funktion: Datenbanken *verteilen* Informationen. Vor allem unzureichend überliefertes Material, das bislang eingeschränkt vor Ort zugänglich war, dem Gros der Interessierten verschlossen blieb bzw. dessen Distribution für Bibliotheken und Archive eine enorme Mehrbelastung darstellte, kann global verfügbar gemacht werden. Datenbanken wie jene von »Compact Memory« dienen der allgemeinen Verbesserung der Informations- und Literaturversorgung und stellen dabei nicht nur für Forschung und Lehre einen qualitativ wichtigen Zugewinn dar – in Zeiten leerer Haushaltskassen bieten sie auch ernstzunehmende ökonomische Alternativen und Vorteile. Einerseits tragen solche Digitalisierungsprojekte im Informationszeitalter stets zu einem positiven, öffentlichkeitswirksamen Image der verantwortlichen Institutionen bei. Andererseits können die eingesetzten Technologien kostensparend in vergleichbaren Projekten nachgenutzt werden, ohne unüberschaubare Mehraufwände zu verursachen. Das Schwesterprojekt von »Compact Memory«, das an der Universitätsbibliothek Frankfurt am Main angesiedelte Portal »Jiddische Drucke« illustriert diesen Zusammenhang.⁶

Auf analoger technischer Basis wie »Compact Memory« realisiert, präsentiert »Jiddische Drucke« ein ca. 130.000 Einzelseiten bzw. rund 1.000 äußerst wertvolle Monographien umfassendes Corpus jiddischer und deutsch-jüdischer Alltags- und Gebrauchsliteratur vom 16. bis zum 19. Jahrhundert. Aufgrund ihrer homogenen Datenmodelle wären beide Portale problemlos unter dem gemeinsamen Dach einer ›Virtuellen Bibliothek des Judentums‹ zu vereinigen, wobei die Datenbanken durchaus an verteilten Standorten vorgehalten bzw. über Schnittstellen weitere Datenbanken integriert werden könnten. Hierzu sind selbstverständlich einheitliche, technische Standards erforderlich, an denen es in den Geisteswissenschaften aufgrund allzu vieler ›Insellösungen‹ leider immer noch mangelt. Aus der Inkompatibilität des überwiegenden Teils geisteswissenschaftlicher Datenbankangebote, die uns heute noch wie Jäger und Sammler durch das World Wide Web streifen lässt und oft genug mit Zufallsfunden nach Hause schickt, resultiert eine unumstößliche Lehre: Datenbanken verteilen Informationen nicht allein an Endnutzer aus Fleisch und Blut – sie distribuieren Daten ebenso an andere Datenbanken, Verbände von Datenbanken oder übergeordnete Portale. Die Entwicklung einer Datenbank hat folglich dem Prinzip der *Ubiquität* zu folgen, was einerseits bedeutet, dass die Datenbestände global

⁶ Zum DFG-Projekt *Jiddische Drucke* (www.literatur-des-judentums.de) der Universitätsbibliothek Frankfurt/Main vgl. Rahel Heuberger: Die Bestände der Judaica-Sammlung auf dem Weg ins Internet. Zwei Digitalisierungsprojekte an der Stadt- und Universitätsbibliothek Frankfurt a. M. In: *Tribüne* 39 (2000), H. 154 (<http://publikationen.ub.uni-frankfurt.de/volltexte/2005/1313/html/tribuene.html>) (05.07.2006); dies., Jiddische Bücher im Internet. In: *Jiddistik Mitteilungen*, Nr 30, November 2003, S. 15–23.

verfügbar sein müssen, in gleicher Qualität und Quantität wie am Erfassungsort, unabhängig von individuell genutzten Hard- oder Softwareplattformen. Andererseits ist technisch zu gewährleisten, dass auch die Datenerfassung an verteilten Standorten durchgeführt werden kann und die Datenbasis migrationsfähig ist, um sie verlustfrei in Fremdsysteme zu überspielen bzw. Fremddaten ins eigene System zu importieren.

Diese Interoperabilität der Systeme auf der Basis gemeinsamer Standards ist unverzichtbar, wenn eine Datenbank alle potentiellen Nutzer erreichen und auf ebenso konsequente wie effiziente Weise mit Informationen versorgen soll. ›In-sellösungen‹, die ausschließlich in ihrer spezifischen Umgebung arbeiten, außerhalb ihres Frameworks aber nicht oder nur unzureichend adressierbar sind, riskieren, früher oder später der Bedeutungslosigkeit anheim zu fallen, da sie den Nutzer mangels Schnittstellen immer wieder zum Wechsel seines virtuellen Arbeitsplatzes nötigen. Dergestalt isolierte Systeme verspielen – unabhängig von ihren konkreten Inhalten – eine wesentliche Stärke der neueren Technologien: Datenbanken *vervielfältigen* Informationen. Dies bedeutet, dass das erschlossene Material nicht exklusiv auf einer Plattform bereitgestellt wird, an die der Nutzer fortan gebunden bleibt, sondern gleichsam ›multipliziert‹ über verschiedene Zugänge abgerufen werden kann. Die Vorteile, welche dieses wichtige Anforderungskriterium nach sich zieht, zeigt ein Blick auf Großprojekte wie die Online-Version der »Bibliographie der deutschen Sprach- und Literaturwissenschaft« (BDSL-Online).⁷

BDSL-Online, die digitale Ausgabe des in der internationalen Germanistik renommierten ›Eppelsheimer-Köttelwesch‹, bietet zunächst die Nutzungs- und Funktionsbreite vergleichbarer, bibliographischer Angebote, die in den letzten Jahren zurecht ihre neue Heimat im World Wide Web gefunden haben: Diverse Eingabemasken ermöglichen die Recherche nach gesuchten Informationen, wobei aufgrund der jahrgangsübergreifend frei kombinierbaren Kategorien (Autor, Titel, Publikationsform, Schlagwort, behandelte Werke usw.) ein deutlicher Mehrwert gegenüber der Druckfassung erzielt wird. Trefferlisten fassen die Suchergebnisse zusammen, deren Detailansicht jeweils alle erforderlichen Titelangaben enthält bzw. auf übergeordnete Werke, Rezensionen, weitere Veröffentlichungen eines Autors oder auf schlagwort-identische Publikationen verweist. Suchrecherchen einer Sitzung werden gespeichert, um später nochmals, ggf. in variiert Form durchgeführt zu werden. Darüber hinaus können Treffer in eine Auswahlliste übernommen und im Anschluss als Textdatei oder

⁷ Die an der Universitätsbibliothek Frankfurt a. M. unter Leitung von Dr. Wilhelm R. Schmidt herausgegebene Online-Version der *Bibliographie der deutschen Sprach- und Literaturwissenschaft* (www.bdsl-online.de) umfasst die Berichtsjahrgänge der gedruckten Fassung des sog. ›Eppelsheimer-Köttelwesch‹ ab 1985. Frei zugänglich sind die mit Förderung der DFG retrodigitalisierten Jahrgänge bis einschließlich 1995; lizenzpflichtige Updates aktualisieren den Datenbestand jeweils in vierteljährlichem Rhythmus.

im weit verbreiteten Austauschformat Dublin Core auf dem eigenen Rechner abgespeichert werden. Zwei weitere Funktionen der BDSL-Online, die das Prinzip der Interoperabilität verdeutlichen, resultieren aus einem wichtigen Anforderungsprofil, das berechtigterweise an internetgestützte Datenbankangebote gestellt wird.

Mehrere Systeme sind interoperabel, wenn sie miteinander ›kommunizieren‹, d. h. Informationen austauschen und somit den Nutzer bei der Literatursuche und -beschaffung so effizient wie nachhaltig unterstützen. Im Fall der bibliographischen Recherche, die in den Geisteswissenschaften zum täglichen Geschäft gehört, sucht der Nutzer in der Regel nach Publikationen, die ihm als Basis seiner eigenen Arbeit dienen. Der Recherche folgt die Beschaffung der gewünschten Quellen, der stets noch eine Verfügbarkeitsprüfung vorausgeht. Machte dies zu analogen Zeiten noch die Konsultation des lokalen Bibliotheks- bzw. der überregionalen Standortkataloge erforderlich, können solche Suchanfragen heute verlust- und verzugsfrei von einem System an ein anderes weitergeleitet werden. Die BDSL-Online integriert zu diesem Zweck sog. ›Linkingservices‹, die dem Nutzer in seiner gewohnten Arbeitsumgebung zum einen per OpenURL den unmittelbaren, standortabhängigen Zugang auf seinen lokalen Online Public Access Catalogue (OPAC) ermöglichen, der Auskunft über die momentane Verfügbarkeit des gesuchten Titels gibt. Zum anderen gewährleistet dieselbe Schnittstelle die parallele Recherche nach lizenzierten und freien elektronischen Volltexten der gesuchten Dokumente im Web bzw. adressiert verschiedene Internet-Buchhandlungen, über die sich das Werk käuflich erwerben lässt.

Über diese konkreten Standortinformationen und Beschaffungsfunktionen hinaus stellen die Linkingservices sicher, dass der Nutzer nicht permanent zwischen unterschiedlichen Internetangeboten wechseln und Rechercheanfragen unter Umständen mehrfach wiederholen muss. Die Vorstellung, Literaturrecherche und -beschaffung redundanzfrei in einer Arbeitsumgebung zu bündeln, mündet in der Idee, ein Maximum von Datenbanken in einem Interface zusammenzufassen, über das sich alle für das eigene Fachgebiet und Forschungsinteresse einschlägigen Angebote individuell konfigurieren und ansteuern lassen. Das Online-Modul »Visual Library .NET«, das zu diesem Zweck entwickelt wurde, setzt an diesem Punkt an und vereinigt als Rechercheportal insgesamt ca. 400 in- und ausländische Allgemein- und Fachdatenbanken, die über eine einheitliche Benutzeroberfläche adressiert werden können.⁸ Informa-

⁸ Vgl. www.visuallibrary.net – Neben allgemein ausgerichteten Datenbanken, Katalogen von Internetbuchhandlungen, Lexika und Suchmaschinen sowie den Katalogen nationaler und internationaler Bibliotheksverbände, Universitätsbibliotheken, FH-Bibliotheken und Spezialbibliotheken sind freie und lizenzpflichtige Angebote zu folgenden Fach- und Forschungsschwerpunkten in der *Visual Library* recherchierbar: Geschichte; Länder & Kulturen; Literatur & Sprache; Medien & Kommunikation; Musik, Kunst & Theater; Naturwissenschaft & Technik; Medizin & Pharmazie; Pädagogik & Psychologie;

tions- und Literaturrecherchen, Verfügbarkeitsprüfungen sowie die eigentliche Textbeschaffung werden nicht mehr in verschiedenen Angeboten, sondern zentral in einer Arbeitsumgebung durchgeführt. Anbindungen an Referenzorgane wie die BDSL-Online oder an Kataloge wie jenen des Gemeinsamen Bibliotheksverbundes (GBV) mit allein mehr als 760 teilnehmenden Bibliotheken bzw. über 45 Millionen erfassten Titeln gewährleisten dabei eine gleichermaßen homogen aufbereitete wie maximale Trefferausbeute. – Auf die »Visual Library« wird weiter unten, im Zusammenhang mit dem Einsatz von Literaturmanagementsystemen in den Geisteswissenschaften, zurückzukommen sein. Ein weiteres, wichtiges Kriterium, das Datenbanken erfüllen müssen und die Aufbereitung bzw. Präsentation von Informationen betrifft, soll vorab kurz skizziert werden.

* * *

Mit dem Einzug von Datenbanken in die Geisteswissenschaften bedurfte es einige Jahre, bis Kompromisse gefunden wurden, welche ergonomische Nutzererwartungen hinsichtlich der Datenwiedergabe erfüllen konnten. Eher die Regel als die Ausnahme stellen dagegen oft heute noch Datenbankangebote dar, die in dieser Hinsicht am Endnutzer »vorbeientwickelt« wurden: Wohlbestückt mit hochinteressanten Materialien für das jeweilige Fachgebiet, liefern Rechercheläufe dennoch immer nur periphere Informationen, die den Nutzer über kurz oder lang verzweifeln und die Datenbank irgendwann zum Datenfriedhof werden lassen. Oftmals erhält man auch nur punktuelle Informationen über den Gesamtumfang der erschlossenen Daten und erfasst folglich gar nicht, welche Schätze möglicherweise in der Datenbank ruhen. Unabhängig von zentralen Fragen, die z. B. die Nachhaltigkeit des Datenformats, die Architektur des Systems oder Aspekte der Langzeitarchivierung thematisieren, werden konkrete Nutzererwartungen von den Anbietern nicht selten stiefmütterlich behandelt. Ein Entwicklungsbereich, der diese Problematik besonders deutlich illustriert, betrifft den Gesamtkomplex der Datenpräsentation.⁹

Die heute noch durchaus gängige und oft genug einzige Form, einen gezielten Zugriff auf den Datenbestand eines Systems zu erhalten, bildet eine Eingabemaske zur Suche nach singulären Datensätzen bzw. -teilen: In Interfaces zur »einfachen« oder »erweiterten Suche« erfasst der Nutzer Personen- oder Autornamen, wahlweise Titelangaben oder Schlagworte zur gewünschten Publikation bzw. sonstige die Recherche eingrenzende Angaben. Stimmt nur einer der Pa-

Philosophie & Theologie; Wirtschaft, Politik & Recht.

⁹ Vgl. Alexander Czmiel / Martin Iordanidis / Pia Janczak / Susanne Kurz / Gesamtedaktion: Manfred Thaller: Retrospektive Digitalisierung von Bibliotheksbeständen.

Evaluierungsbericht über einen Förderschwerpunkt der DFG. Universität zu Köln, Januar 2005

www.dfg.de/forschungsfoerderung/wissenschaftliche_infrastruktur/lis/download/retro_digitalisierung_eval_050406.pdf (05.07.2006).

parameter nicht mit den hinterlegten Daten überein, liefert das System keine Treffer – umgekehrt formuliert: Der Nutzer muss bereits über ein unter Umständen beträchtliches Wissen darüber verfügen, was er sucht, um überhaupt etwas zu finden.

So unverzichtbar elaborierte Eingabemasken zum gezielten ›Searching‹ sind – das zukunftsreichere Potential moderner Technologien besteht darin, in Datenbanken enthaltene Informationen zu *visualisieren*, d. h. das intuitive ›Browsing‹ nach relevanten Materialien zu unterstützen. Nicht mehr nur die gezielte Suche nach bibliographischen Angaben, die oft einem Stochern im Heuhaufen gleichkommt, fördert eine Treffermenge zutage – der Nutzer hat ergänzend die Möglichkeit, den Datenbestand in toto systematisch und strukturiert auf konkrete Fragestellungen hin zu durchforsten, d. h. ohne exakte Eingaben zu einschlägigen Ergebnissen zu gelangen. Für die »Bibliography of Linguistic Literature / Bibliographie Linguistischer Literatur« (BLLDB), des Schwesterprojekts der BDSL-Online, bedeutete dies eine Klassifikationsstruktur von Fachtermini zu etablieren, entlang denen sich der Nutzer zu den Publikationen seines Interessenschwerpunkts navigieren kann. Die hierarchisch organisierte Visualisierung von Datenbeständen zur Ergänzung herkömmlicher Suchfunktionen stellt ein globales Prinzip dar, das für Datenbanken jedweden Typs Geltung beansprucht. Analog zur gestaffelten Schlagwortnavigation in bibliographischen Angeboten wie BLLDB oder BDSL-Online sollte daher auch in retrospektiven Digitalisierungsprojekten, die schwerpunktmäßig Originalgraphiken historischer Quellen präsentieren, auf eine strukturierte Visualisierung des abrufbaren Materials gesetzt werden.

Die skalierbaren Seitengraphiken der jüdischen Periodika sind in »Compact Memory« gemäß ihrer ›natürlichen‹ Ordnung zugänglich, wobei über den Navigationsbaum im linken Frame zunächst die gewünschte Zeitschrift, dann der Jahrgang und schließlich das gesuchte Heft sowie ein Verzeichnis der darin enthaltenen Einzelbeiträge aufgerufen wird. Datenbankportale sollten den Nutzern grundsätzlich solche ergonomisch sinnvollen, visualisierten und intuitiven Zugriffe auf ihre Datenbestände ermöglichen: Im Fall von »Compact Memory« bedeutet dies, dass auf das internalisierte Wissen der Nutzer rekurriert wird, welches sie sich im Umgang mit periodisch erscheinender Literatur in ›realen‹ Bibliotheken erworben haben. Diese Prämisse gilt allerdings nicht nur für die nutzerseitige Datenausgabe, sondern – wie im Folgenden kurz skizziert werden soll – gleichermaßen für die Datenerfassung.

* * *

Die Entwicklung von Datenbanken, die Katalogisierung und Erschließung eines Datenbestands, dessen Überführung in ein Internetportal sowie die nachhaltige Serverbetreuung – kurz: alle Arbeitsphasen von der Planung über die Realisation bis zum Unterhalt eines solchen Projekts – verschlingen hohe Sach- und Perso-

nalkosten und fordern in der Regel ein üppig bemessenes Zeitbudget. Im Gegenzug sind meistens nur Großprojekte, die sich an ein breites und daher in den Geisteswissenschaften vergleichsweise seltenes Publikum wenden, in der Lage, die entstandenen Aufwände durch angemessene Lizenzmodelle abzufedern, falls das Vorhaben nicht ohnehin von der öffentlichen Hand finanziert wird oder dem Open Access-Modell folgt. Neben dem bereits erwähnten technischen Kriterium der *Ubiquität* ist also jede Institution, die ein solches Projekt erwägt, gut beraten, Planung und Realisation auch unter die Prämissen der *Quantität*, *Simultaneität* und *Simplizität* zu stellen.

Datenbank- und Digitalisierungsprojekte machen aufgrund der hohen, langfristigen Investitionen ökonomisch nur Sinn, wenn eine ›Masse‹ von Daten für eine ›Masse‹ von Nutzern aufbereitet wird. Vorhaben, die lediglich eine überschaubare Klientel bedienen bzw. eine schmale, womöglich sehr spezialwissenschaftliche, niederfrequent genutzte Datenbasis anbieten, rechnen sich nicht, so bedeutsam das Ansinnen auch erscheinen mag. Dies bedeutet nun freilich nicht, dass Projekte, die nicht wenigstens einige Hunderttausend Datensätze bzw. Digitalisate ein paar Tausend potentiellen Nutzern zur Verfügung stellen, bereits in der Planungsphase verworfen werden müssen. Vielmehr sollten in quantitativer Hinsicht ›kleine‹ Projekte möglichst jede technische Eigenentwicklung vermeiden und konsequent und kompromißlos auf die kostengünstigere Nachnutzung bestehender Softwaresysteme setzen – selbst wenn dies heißt, nicht alle Projektziele verwirklichen zu können.

Aber nicht allein aus Kostengründen sollte der Versuchung widerstanden werden, Softwareentwicklungen in Eigenregie vorzunehmen: Die Ergebnisse des vergangenen Jahrzehnts, in dem überproportional viele Digitalisierungsprojekte in den Geisteswissenschaften angegangen wurden,¹⁰ zeigen, dass die Auseinandersetzung mit moderner Softwaretechnologie, insbesondere im Rahmen von Großprojekten, zum einen nicht unbedingt zu den Stärken des Fachs zählt, zum anderen illustriert die Flut der entstandenen ›Insellösungen‹ hinreichend die geringe Nachhaltigkeit sowie die ubiquitäre Beschränkung vieler Angebote. Insbesondere im technischen Bereich, darauf wurde wiederholt hingewiesen, fehlt es an unerlässlichen Synergieeffekten,¹¹ die vor allem die teilweise gravierend langen Verzugszeiten vermeiden helfen, bis lauffähige Online-Datenbanken ins Netz gelangen – sofern diese Zielmarke überhaupt erreicht wird. Ein zentrales Kriterium wurde in der langwährenden Diskussion über ein-

¹⁰Vgl. die Übersicht der von der DFG-geförderten Digitalisierungsprojekte (http://www.dfg.de/forschungsfoerderung/wissenschaftliche_infrastruktur/lis/download/projektuebersicht_orte.pdf) (05.07.2006).

¹¹Vgl. z. B. in Reaktion auf den »Evaluierungsbericht« (wie Anm. 9) Monika Wicki: Digitalisierte historische Kinderbücher aus Beständen der Universitätsbibliotheken Oldenburg und Braunschweig (Retrospektive Digitalisierung) [Rezension]. In: H-Soz-u-Kult (<http://hsozkult.geschichte.hu-berlin.de/rezensionen/id=90&type=rezwww>) (05.07.2006).

zuhaltende technische Anforderungen und Richtlinien weitgehend an den Rand der Aufmerksamkeit gedrängt: Simultanität.

Der Umstand, daß in Digitalisierungsprojekten teilweise Jahre zwischen der Erfassung der ersten Digitalisate, Metadaten und Volltexte bis zu ihrer Bereitstellung im Internet vergehen, stellt sich in ökonomischer und repräsentativer Hinsicht um so negativer dar, als heutzutage für diese Verzugszeiten technisch betrachtet kein Anlass mehr besteht. Verfügbare, professionelle Softwaresysteme gewährleisten mittlerweile, daß der Produktionsprozeß zeitlich auf ein Minimum reduziert wird, ohne irgendwelche verbindlichen, technischen Normen preisgeben zu müssen. Nach dem Prinzip der Simultanität erfolgt die datenbankgestützte Einspeisung der bibliographischen Metadaten weitgehend parallel zur strukturierten Zuordnung und Aufbereitung der graphischen Inhalte, die bei Bedarf in der gleichen Arbeitsumgebung per Optical Character Recognition (OCR) in digitalen Volltext umgewandelt werden. Zu jedem Zeitpunkt eines Arbeitsganges, der von einer Korrektur einzelner Datensätze bis hin zur Digitalisierung einer umfangreichen Vorlage variieren kann, sollte das System in der Lage sein, die gewonnenen Produktionsergebnisse direkt, autark und ohne größere Aufwände auf den Webseiten des Projekts zu aktualisieren. Dabei folgt ein solches System auch dem Gebot der *Simplizität*: Dies bedeutet etwa, Synergieeffekte zu vergleichbar gelagerten Projekten zu nutzen, grundsätzlich also das Rad nicht neu erfinden zu wollen oder im konkreten Fall auf aufwändige Eigenentwicklungen zu verzichten, wenn etablierte, leistungsstarke Software, z. B. OCR-Programme, Bildbearbeitungs-Tools oder Graphikkonverter, modular in die Benutzeroberfläche implementiert werden können. Das Prinzip des »keep it simple« kommt aber auch in der gesamten Benutzerführung und Produktionslinie zum Tragen, deren ergonomisches Design jegliche Redundanzen vermeidet und deren Verständnis keine zeitintensiv zu erwerbenden Vorkenntnisse verlangen darf.

Entsprechend dieses Anwendungsprofils wurde mit der »Image Library« eine homogene, modular aufgebaute Arbeitsplattform entwickelt, die die strukturierte Erfassung, Indizierung und Volltexterkennung, ferner die Bearbeitung und Bereitstellung beliebiger grafischer und textueller Materialien im Internet unterstützt. Die Software erlaubt es, große Mengen von Grafiken übersichtlich und schnell auf einen lokalen Dokumentenserver zu überspielen und dort systematisch mit den zugehörigen Metadaten der Publikation zu verknüpfen – je nach Vorlage mehrere Tausend Images pro Stunde. Ein Vorschauenfenster zeigt dem Bearbeiter den Inhalt des Quellverzeichnisses an, aus dem Images per Drag-and-Drop ins Zielverzeichnis kopiert werden, wobei die Digitalisate im integrierten Graphikbetrachter einzeln oder in Form von Thumbnails aufgerufen und in einem Arbeitsgang von Schattierungen oder Verschmutzungen gereinigt werden. Ein Navigationsbaum, der das Zielverzeichnis des Datenbankservers abbildet, gibt die erfassten, annotierbaren Metadaten wieder, die der Bearbeiter parallel in

der »Image Library« katalogisiert oder vorzugsweise über eine Schnittstelle automatisch importiert hat. Im Anschluss an diese Schritte können die Graphiken mittels eines OCR-Moduls in Volltexte umgewandelt und redaktionell bearbeitet werden, um das erschlossene Material dann per Mausklick in Form von dynamisch generierten Webseiten im Internetportal zur Verfügung zu stellen.

* * *

An die Ausführungen zu den Technologien, die in institutionellen Großprojekten sehr erfolgreich eingesetzt werden, knüpft sich die Frage an, welche Aufgaben Datenbanken für Geisteswissenschaftler im privaten Anwendungsbereich übernehmen können? Die Fortschritte, die im Großen verzeichnet werden, strahlen allmählich auf diesen Entwicklungssektor aus und bieten dem Privatanwender in Form von Desktop-Lösungen zunehmend mehr Möglichkeiten, sein wichtigstes Gut zu schützen – Zeit. Darüber hinaus wird er in die Lage versetzt, sich den Grundlagen seiner Fachdisziplin wie auch der Erschließung neuer Forschungserträge umfassender, strukturierter, ökonomischer und flächendeckender zu widmen, als dies ausschließlich mit Hilfe traditioneller Medien realisierbar war. Dabei ist der Stellenwert, den die »Neuen Medien« und ihre Entwicklungswerkzeuge im vergangenen Jahrzehnt gewonnen haben, zweifellos dem Quantensprung gleichzustellen, den in den Geisteswissenschaften zuletzt Gutenbergs Erfindung des Buchdrucks auslöste.

Die »Elektronische Datenverarbeitung« galt in den Geisteswissenschaften lange als ureigenstes Ressort technisch interessierter Forscher und gelangte ohne Mithilfe versierter Informatiker selten zum Einsatz. Mit dem Personal Computer begann der Rückzug der Schreibmaschine vom Arbeitsplatz des Geisteswissenschaftlers, aber die ersten Entwicklungsgenerationen der Textverarbeitungsprogramme revolutionierten die althergebrachten Arbeits- und Publikationsweisen noch nicht im Kern. Dies geschieht erst, seitdem leistungsstarke, erschwingliche Literaturmanagementsysteme eingeführt wurden, welche – in Kombination mit den Informationsangeboten im Internet – grundsätzlich neue Wege der fachwissenschaftlichen Produktions- und Publikationsmethoden eröffnen. Im Produktionsstadium basieren diese modernen Technologien dennoch weiterhin auf den klassischen *rhetorices partes* und berücksichtigen damit die internalisierten, humanistischen Arbeitsmethoden, um deren konkrete Umsetzung mit Hilfe von datenbankgestützten Anwendungen einschneidend und umfassend zu optimieren.

Hier setzt das Literaturmanagementsystem »Visual Composer .NET« an, das auf derselben technischen Plattform aufsetzt wie dessen auf Digitalisierungsprojekte spezialisiertes Derivat, die »Image Library«:¹² Die Datenbank dieser Literaturverwaltung enthält mit bibliographischen Metadaten, ihnen zuge-

¹²Zu näheren Informationen vgl. <http://www.visualcomposer.net>.

wiesenen, systematisch archivierten und indexierten Exzerpten oder Volltexten, ferner den Referenzen zu sachverwandten oder übergeordneten Quellen sowie den Publikationen, Projekten, Konzepten und Gliederungen des Nutzers eine umfassende, auf die individuellen Forschungsinteressen zugeschnittene Materialbasis. Sie dient im übertragenen Sinn jener ›Auffindung des Stoffes‹ (*inventio*), die den Ausgangspunkt eines aktuellen Projekts bildet, um mit der Zeit kontinuierlich anzuwachsen – u. a. durch neues, fachbezogenes Material, das mit Hilfe des integrierten Online-Moduls »Visual Library« systematisch aus dem Internet gefiltert und in die Dankbank importiert werden kann.

Die ›Gliederung des Stoffes‹ (*dispositio*), ließe sich sagen, geschieht in der Projektverwaltung des »Visual Composers«: Der Nutzer legt zunächst ein neues ›Projekt‹ an – dies kann jede Form einer geplanten Publikation, eines Forschungsvorhabens oder eines Digitalisierungsprojekts sein. Alle relevanten Daten und Materialien, die zur Realisation erforderlich sind – Metadaten, eigene Exzerpte, zugehörige Dokumente, Graphiken oder sonstige Referenzen – werden per Drag & Drop systematisch einer frei konfigurierbaren Gliederungsstruktur zugewiesen, welche den argumentativen, organisatorischen oder logistischen Aufbau des Projekts widerspiegelt. Hier wie schon bei der vorangegangenen Datenerfassung spielt das Prinzip der Visualisierung eine maßgebende Rolle: Die Datenbank präsentiert sich nicht als ›Black Box‹, die dem Nutzer an der Arbeitsoberfläche immer nur einen beschränkten Ausschnitt seiner Datenbasis abbildet, sondern ihm einen visualisierten Zugriff auf alle Bereiche seines Corpus bietet, auf das er zudem wechselnde Sichtweisen erzeugen kann.

Bildet ein Text das Produktionsziel, wie dies bei Geisteswissenschaftlern die Regel ist, steht dem Nutzer im »Visual Composer« ein Editor zur Verfügung, der alle Kernfunktionen herkömmlicher Textverarbeitungsprogramme bietet. Darin bearbeitet er sein gegliedertes Konzept nach textuellen Kriterien und bereitet es unter sprachlichen Gesichtspunkten für die eigentliche Publikation vor (*elocutio*). Dieser Produktionsschritt widmet sich vorrangig der rhetorischen Durchgestaltung und kommunikativen Intention des Textes, verlangt allerdings aufgrund des wissenschaftlichen Anspruchs stets die Beachtung zahlreicher formaler Elemente und Kriterien. Typographische und bibliographische Vorgaben wechselnder Stylesheets, wie z. B. variierende Richtlinien hinsichtlich der Zitation von Quellen, müssen in diesem Arbeitsgang ebenso berücksichtigt werden wie technische Anforderungen, etwa die zunehmend geforderte Bereitstellung der Dokumente in spezifischen Formatsprachen oder deren Indexierung entsprechend moderner Auszeichnungssysteme. In Literaturmanagementprogrammen wie dem »Visual Composer« entfallen diese redundanten Tätigkeiten weitgehend: Spezifische Gliederungsstrukturen oder homogenes Datenmaterial, z. B. Literaturangaben oder Fußnoten, werden automatisch erstellt, mitgeführt und aktualisiert, wobei vorgegebene oder individuell konfigurierbare Stylesheets

die Einhaltung typographischer bzw. formatsprachlicher Vorgaben gewährleisten.

Eine gänzlich neue Dimension gewinnt der rhetorische Akt der *memoria* durch den Einsatz von Datenbanken in den Geisteswissenschaften: Hatte sich der klassische Orator seine Rede für den freien, öffentlichen Vortrag einzuprägen, »memoriert« die Datenbank nicht nur eine konkrete, situationsabhängige Ergebnisstufe, sondern auch sämtliche direkten und indirekten Vorstufen sowie alle über- und untergeordneten Elemente des Produktionsprozesses des Textes. Die Bestandteile, aus denen das Arbeitsergebnis letztlich resultiert, werden im »Visual Composer« getrennt von einander in nicht-proprietärer Form systematisch archiviert, um entweder bei Bedarf separat aktualisiert zu werden oder in neue Projekte einzugehen.

* * *

Das Einsatzgebiet von Datenbanken, dies sollten die Anmerkungen zum »Visual Composer .NET« ebenfalls verdeutlichen, kann sich in den Geisteswissenschaften nicht auf die Verwaltung homogener, hochstrukturierter Materialbestände, z. B. bibliographische Metadaten, beschränken. Die spezifische Stärke datenbankgestützter Literaturmanagementsysteme, die mittlerweile auch »kleine« Desktoplösungen für den Privatanwender auszeichnen, liegt vielmehr darin, effizient und ergebnisbezogen Materialien und Informationen zu synthetisieren, deren formale Charakteristika ebenso wie ihre konkrete Nutzenanwendung z. T. hochgradig variabel sind. Dieses Potential – kombiniert mit der multiplen Einsatztauglichkeit dieser Systeme, welche nicht zuletzt aus ihrer modularen Architektur und ihrem ergonomischen Design resultiert – gewährleistet, dass solche Anwendungen praktisch sämtliche Aufgabenschwerpunkte ihrer Nutzer abdecken bzw. begleitend unterstützen. Auf der Grundlage einer erweiterbaren Datenbasis kann der Geisteswissenschaftler in einer einheitlichen, webaktiven Softwareumgebung alle Arbeiten durchführen, die sich im Rahmen von Studium, Lehre und Forschung ergeben – Datenverwaltung bzw. -archivierung, Informationsrecherche, Verfügbarkeitsprüfungen, Literaturbeschaffung, Textauswertung und -verarbeitung, Veranstaltungs- bzw. Projektorganisation, Publikationsmanagement usw. Der Fokus all dieser Tätigkeiten und Anstrengungen, die den Tagesablauf des Geisteswissenschaftlers bestimmen, richtet sich in der Regel letztlich darauf, erstellte Lehrmaterialien oder gewonnene Forschungsergebnisse zu veröffentlichen. Das Spektrum möglicher Publikationstypen ist dabei grade in den »Humanities« denkbar breit und reicht von der Rezension oder Miszelle über den Essay, die Dissertation oder Habilitation bis zu großangelegten Projekten wie umfangreiche Einzelstudien, kommentierte Quelleneditionen, fachwissenschaftliche Nachschlagewerke und datenbankgestützte Webportale. Diesen Veröffentlichungsformen gesellt sich in den Geisteswissenschaften der nicht unbedeutende Bereich der sog. Grauen Literatur zur Seite: in der Lehre z.B. Seminar- und Examensarbeiten, aber auch Arbeitsblätter, Lektürelisten, Reader oder

sonstige Studienvorlagen, dagegen in der Forschung beispielsweise Tagungsunterlagen, Institutsschriften oder Quellenmaterialien und Dokumentationen, wie sie auf vielfältige Weise im Rahmen von Forschungsprojekten gesammelt werden. Publikationswürdige Akten, Nachlässe, unveröffentlichte Forschungsbeiträge oder alte, rare Druckwerke, die aus Kosten- bzw. Absatzgründen nicht neu aufgelegt werden, ergänzen diese fast endlose Liste.

Zweifellos ist nicht grundsätzlich alles, was geschrieben, aufgezeichnet oder überliefert wurde, wert, auch veröffentlicht zu werden. Eine verquere Subventionspolitik hat in den letzten Jahren indessen in den Geisteswissenschaften eine Publikationslandschaft geschaffen, die an Absurdität kaum zu überbieten ist: Die öffentliche Hand bezuschusst zunächst Verlage mit hohen Beträgen, Forschungsergebnisse der steuerzahlenden Allgemeinheit zugänglich zu machen. Liegen diese Druckwerke dann vor, ›erwirbt‹ sie die Öffentlichkeit zu übersteuerten Preisen tatsächlich ein *zweites* Mal, um sie in steuerfinanzierte Bibliotheken einstellen zu können. Überspitzt ließe sich formulieren: Veröffentlicht wird an geisteswissenschaftlicher Fachliteratur möglicherweise nur noch, was einen Druckkostenzuschuss erhält, wobei dessen Bewilligung nicht notwendig aus der Güteklasse der Publikation resultieren muss. Die ›Todesspirale‹ der erwerbspolitischen Zwangsentscheidungen der Hochschulbibliotheken, wie sie z. B. in den Massenkündigungen von Zeitschriftenabonnements zutage tritt, stellt nur eine der schwerwiegenden Folgen dieser Abwärtsentwicklung dar.

Aus technischer und ökonomischer Sicht sind Datenbanken ein wirkungsmächtiges Werkzeug, dieser bildungs- und forschungspolitischen Dauermisere erfolgreich Paroli zu bieten und qualifizierte, nachhaltige Auswege aus der Krise zu weisen. Datenbanken *publizieren* Informationen – im World Wide Web sehr viel aktueller, umfassender, funktionaler und preiswerter als jedes Druckwerk, das zudem immer mit Verzugszeiten und medial bedingten Abstrichen zu kalkulieren hat, sich im Umfang des Gebotenen beschränken muss und kaum zu rechtfertigende Kosten verursacht. Der besondere Vorzug, dass webbasierte Publikationen sehr viel mehr Öffentlichkeit erzeugen und Leser erreichen als kostspielige, standortgebundene Fachliteratur in Kleinstauflagen, muss nicht eigens hervorgehoben werden.

Hat der Geisteswissenschaftler seine bibliophilen Bedenken nun erfolgreich fallengelassen und die rationalen Argumente, die für eine (parallele) Webveröffentlichung sprechen, akzeptiert, stellt sich ihm die Frage, wo er in den Weiten des Internet seine Texte plazieren soll. Auch wenn die Publikationsmöglichkeiten zahllos erscheinen mögen – im Wesentlichen reduzieren sie sich auf vier Alternativen: der Veröffentlichung auf einem institutionellen, ggf. fachbezogenen Dokumentenserver, der Bereitstellung seiner Texte auf einer privaten Homepage, der Einspeisung seines Materials in ›offene Wissensportale‹ wie »Wikipedia« oder der Publikation auf einer synergetischen Plattform, welche die individuellen Vorteile dieser drei erstgenannten Möglichkeiten ver-

eint. Mit »Bibliopedia« wurde der prototypische Versuch unternommen, auf der Basis der *MediaWiki*-Software die Stärken der Einzelalternativen zu bündeln und eine fachspezifische Publikationsplattform aufzusetzen, welche die inhaltlichen, technischen und ergonomischen Erwartungen des Geisteswissenschaftlers bezüglich einer Veröffentlichung im Netz erfüllen soll.

Die Entscheidung, »Bibliopedia« technisch in unmittelbarer Nähe zur »Wikipedia« anzusiedeln, fiel einerseits aufgrund des praxisbewährten Softwareprofils der »MediaWiki«, ihres Funktionsumfangs sowie ihres ergonomischen Designs, das selbst Laien in kürzester Frist in die Lage versetzt, Texte jedweder Art online zu publizieren. Ausschlaggebend war andererseits, dass eine fachorientierte, interaktive, allgemein zugängliche Publikationsplattform nach dem Vorbild der »Wikipedia« grundsätzlich dem Prinzip der allseitigen Offenheit folgen muss, um die Diskursfreiheit der Geisteswissenschaften zu sichern bzw. die Qualität der eingebrachten Einzelbeiträge im Autorenkollektiv zu sichern und womöglich zu steigern. Daraus resultiert ein weiteres Kriterium, das bei der Systemwahl eine zentrale Rolle spielte: Bei der Entwicklung der »MediaWiki« als eines verteilten Redaktionssystems stand, neben der ›Philosophie der Offenheit‹, die Indexierung der Inhalte bei Suchmaschinen und bestmögliche Platzierung in den Ergebnislisten im Vordergrund – der Google-Faktor. In dem Maß nämlich, in dem ein Text Öffentlichkeit erzeugt und mittels automatischer Suchtechniken global leichter verfügbar ist, wird er als Plagiat unattraktiv. Es reicht aus, eine entsprechende Textstelle zu ›googeln‹.

Aufgrund ihrer enzyklopädischen Ausrichtung und ihres kollektiven Redaktionssprinzips eignet sich »Wikipedia« selbst allerdings nur ansatzweise als fachspezifische Publikationsplattform – ja, will dies auch gar nicht sein: Um den diesbezüglichen Erwartungen und Anforderungen eines Geisteswissenschaftlers zu genügen, wurde in »Bibliopedia« eine Reihe von Standardisierungen implementiert, die ihr Profil als Fachportal schärfen, das gesamte Spektrum der disziplinen- genen Publikationstypen zulässt bzw. mittels einer ergonomisch optimierten Benutzerführung inhaltliche und funktionale Mehrwerte generiert. Dies impliziert zunächst, daß die Artikelkonventionen in »Bibliopedia« weitaus liberaler gehandhabt werden, als es in »Wikipedia« der Fall ist: In dieser Hinsicht werden also nicht nur objektive ›Lexikonartikel‹ sanktioniert, sondern auch Beiträge autorisiert, die stilistisch und inhaltlich sowohl das konkrete Erkenntnisinteresse ihres Verfassers widerspiegeln als auch in der äußeren Form den spezifischen Erfordernissen des Mitgeteilten entsprechen dürfen und sollen. Vordefinierte Einsprungspunkte – im Portal ›Literaturwissenschaft‹ etwa *Autoren*, *Werke*, *Epochen* und *Themen* – erleichtern einerseits die Platzierung des eigenen Texts gemäß den einschlägigen Kategorien der jeweiligen Fachdisziplin. Nach dem Prinzip der Visualisierung unterstützt dieses Feature andererseits die Navigation im Gesamtdatenbestand bzw. dessen systematische, kontextbezogene Anreicherung mit detailspezifischen Materialien. Die Integration verschiedener Module, z. B. zur

Verschlagwortung der eingebrachten Texte oder normierten Erfassung bibliographischer Titelangaben, erleichtern dabei vorrangig die Dateneingabe, tragen darüber hinaus aber wesentlich dazu bei, dass die statisch-enzyklopädische Modellierung der »Wikipedia« gleichermaßen flexibel wie homogen den besonderen Interessen geisteswissenschaftlicher Nutzer angepasst wird. Nicht an letzter Stelle fällt in diesen Bereich die Alternative, dass Publikationen wie in der »Wikipedia« sowohl im Kollektiv redigiert als auch gemäß geltenden Urheberrechts ganz oder teilweise gesperrt werden können.

* * *

Ziel dieses ›Werkstattberichts‹ war es, aus entwicklungstechnischer Sicht an einigen konkreten Beispielen das Einsatzgebiet zu umreißen, in dem Datenbanken heute in den Geisteswissenschaften genutzt werden. Neben Fragen, die der praktischen Realisation von Datenbankprojekten und den ihnen zugrunde liegenden Anforderungsprofilen galten, sollte gezeigt werden, dass sich der Anwendungsbereich dieser Technologien zunehmend ausweitet und in Zukunft noch stärker ausdehnen wird: Großvorhaben, z. B. Digitalisierungsprojekte, sind ohne Datenbankunterstützung nicht denkbar, wobei festzuhalten bleibt, dass sich deren technologische Basis – verbunden mit allem zusätzlichen Aufwand und Nutzen – in wachsendem Maß spezialisiert, also professionelle Kompetenzen erfordert. Ausgehend von EDV-gestützten Katalogsystemen ›erobern‹ Datenbanken in den letzten Jahren ferner Publikationstypen, die bislang mangels Alternativen dem Buchdruck vorbehalten waren: Fachbibliographien oder Referenzorgane stehen diesbezüglich aufgrund ihrer hochstrukturierten und homogenen Daten an erster Stelle, zumal der ›elektronische Zugriff‹ und moderne Technologien einen erheblichen Mehrwert offerieren, den analoge Medien auch langfristig nicht werden bieten können. Eine zentrale Rolle übernimmt in diesem Kontext selbstverständlich das Internet, das potentiell jede Datenbank der Welt standortunabhängig verfügbar macht, um als Medium zugleich – »Wikipedia« und ihre Derivate wurden genannt – Datenbanken als fachspezifische Diskussions- und Publikationsplattformen vorzuhalten. Von der wachsenden Popularität und der stetig steigenden Entwickler- und Benutzerfreundlichkeit profitieren schließlich nicht zuletzt die Endanwender: So werden Datenbanken im Großen nicht nur für Benutzerkreise anwendungstauglich, denen diese Technologien bislang mangels tieferer Kenntnisse verschlossen bleiben mussten. Der Trend geht zudem in die Richtung, daß datenbankgestützte Literatur- und Dokumentenmanagementsysteme immer mehr Individualanwender finden und die Abkehr von traditionellen Arbeitsprozessen in den Geisteswissenschaften wohl nicht mehr lange auf sich warten lassen wird. Die Frage nach dem »Sinn und Nutzen von Datenbanken in den Geisteswissenschaften« stellt sich also vielleicht doch nicht nur rhetorisch: Der Aufbau von Datenbanken sowie der Umgang mit diesen ›Wissensspeichern‹ prägen und schulen eine neue Art des Suchens, Findens und Verarbeitens von Antworten und Erkenntnissen. Datenbanken in diesem Sinn wären folglich nicht nur ›For-

schungswerkzeuge, sondern in gleichem Maß genuine ›Forschungsgegenstände‹ der Geisteswissenschaften im 21. Jahrhundert.

Literaturverzeichnis

Adams, Douglas: *The Hitch Hiker's Guide to the Galaxy*. London: Pan Books 1979.

Heuberger, Rahel: Die Bestände der Judaica-Sammlung auf dem Weg ins Internet. Zwei Digitalisierungsprojekte an der Stadt- und Universitätsbibliothek Frankfurt a. M. In: *Tribüne* 39 (2000).

Schicketanz, Till und Heiligenhaus, Kay: »Inseln im Meer des Beliebigen«. Architektur und Implementierung eines Internetportals Deutsch-jüdische Periodika. In: *Jahrbuch für Computerphilologie* 5 (2003).

Thaller, Manfred: Retrospektive Digitalisierung von Bibliotheksbeständen. Evaluierungsbericht über einen Förderschwerpunkt der DFG. Universität zu Köln, Januar 2005.