

Quelleneditionen auf CD-ROM?*

Als die Arbeitsgemeinschaft außeruniversitärer historischer Forschungseinrichtungen (AHF) bei mir anfragte, ob ich bereit sei, für ihr Jahrbuch einen kleinen Beitrag über historische „Quelleneditionen auf CD“ zu schreiben, wozu sich später die Bitte gesellte, in der Mitgliederversammlung schon vorab darüber zu berichten, habe ich mir Bedenkzeit erbeten. Ich wollte vor einer Zusage erst erkunden, was an wissenschaftlichen historischen CD-Quelleneditionen überhaupt „auf dem Markt“ ist. Der einfachste Weg, dies zu tun, schien mir, sich an die Sondersammelgebietsbibliotheken für Geschichte zu wenden, die von der DFG dabei unterstützt werden, auch spezielle und hochspezielle Fachliteratur anzukaufen, wobei - wie ich aus langjähriger Mitgliedschaft im Bibliotheksausschuß der DFG wußte - ein spezieller Topf für den Erwerb von CDs zur Verfügung steht. Sondersammelbibliothek für Geschichte ist die Bayerische Staatsbibliothek, für bestimmte Bereiche auch die Staats- und Universitätsbibliothek Göttingen. Eine abendliche mehrstündige Katalogrecherche im Internet und eine fernmündliche Nachfrage, ob dies wirklich alles sei, ergaben einen eindeutigen Befund: Es lohnt nicht. Was ich im Katalog an CD-Editionen fand, war nicht viel, kannte ich bereits, und mehr gab es offensichtlich nicht; wie man mir versicherte, wurde alles Wichtige gekauft.¹

Als ich dieses Ergebnis der AHF-Geschäftsstelle mitteilte, war man nur mäßig erschüttert. Man schlug mir vor, der Frage nachzugehen, warum dies so sei; ausgehend von diesem Problem allgemeinere Fragen der digitalen Edition von historischen Quellen zu erörtern und dabei die Erfahrungen des DFG- Förderprogramms der „Retrospektiven Digitalisierung“ mit einzubringen; denn daß ich den DFG-Gutachterausschüssen für „Elektronische Publikationen“ und „Retrospektive Digitalisierung“ angehörte, mehrere Jahre auch vorsah, war der Grund gewesen, mich um eine entsprechende Stellungnahme zu bitten.² Und in der Tat: Sowenig das DFG-Digitalisierungsprogramm mit der CD als Träger digitaler Daten steht und fällt, sowenig steht und fällt mit ihr die digitale Edition historischer Quellen; denn die Frage ihrer Chancen und Probleme läßt sich ebensowenig auf die Frage des Trägers, wie die Frage des Trägers auf die CD reduzieren; wenn man das Anfang/Mitte der 90er Jahre noch anders sah, so zeigt das nur einmal mehr, wie schnelllebig dieses Metier ist, wie sich manche Prognosen nicht erfüllten, durch die Entwicklung überholt wurden, revidiert werden mußten, wie einige Probleme sich rasch von selbst erledigten, andere blieben, manche erst mit der Zeit erkennbar wurden. Das alles

* Vortrag, gehalten auf der Mitgliederversammlung der AHF am 4. März 2002.

¹ Vgl. die Alphabetische Liste der Datenbanken / CD-ROM der Bayerischen Staatsbibliothek (<http://www.bsb-muenchen.de/benuetzu/cdabc.htm>); zwar ist sie im Ausdruck über 100 Seiten lang, aber selektiert man nach Fachgebieten (hier: Geschichte), Gegenstand (Quelleneditionen) und Träger (CD) ist das Ergebnis so, wie es eben beschrieben wurde. Es findet sich bestätigt in der Göttinger Systematischen Übersicht über ihren CD-ROM-Bestand (http://www.sub.uni-goettingen.de/ebene_2/2_cdfach.htm). Wer es zumindest an einem Ausschnitt genauer wissen möchte, sei verwiesen auf: *Markus Sehmeyer*, CD-ROMs und Internet in der spätantiken und mittelalterlichen Geschichtsforschung, in: *Historische Zeitschrift* Bd. 274, Heft 2 (April 2002), S. 367-386. Für Unterstützung bei den folgenden Ausführungen sage ich den Damen und Herrn des Digitalisierungszentrums der Bayerischen Staatsbibliothek München, der Gruppe Wissenschaftliche Literaturversorgungs- und Informationssystem in der Geschäftsstelle der DFG Bonn sowie Herrn Professor Dr. Gerhard Schneider vom Rechenzentrum der Universität Freiburg herzlichen Dank.

² Dazu auch meinen Beitrag zum 4. Gemeinsamen Kolloquium der Digitalisierungszentren der BSB München und der SUB Göttingen am 28./29. November 2000 in München (*Helmut Altrichter*, *Retrodigitalisierung in Deutschland. Versuch einer Zwischenbilanz*), mit den anderen dort gehaltenen Referaten nachzulesen unter <http://www.bsb-muenchen.de/mdz/forum/altrichter/index.htm>

läßt sich *auch* an der CD zeigen: Ich meine 1. die Vorteile, die digitale Editionen gegenüber gedruckten nun einmal haben; 2. die unterschiedlichen Funktionen, die die Digitalisierung von Texten in den einzelnen historisch arbeitenden Wissenschaften spielen kann (was mir selbst auch erst so recht während der Arbeit in den genannten DFG-Gremien einsichtig wurde); und 3. welche Folgen sich daraus sowie aus der technischen Entwicklung für die Frage des Datenträgers ergeben. Weil das so ist, sind wir beim ursprünglichen Thema („Quelleneditionen auf CD“) geblieben und haben dahinter nur ein Fragezeichen gesetzt. Ich will zu jedem dieser Punkte ein paar Bemerkungen machen:

1. Zu den Vorteilen der elektronischen Edition gegenüber den Printmedien

CD-ROM steht bekanntlich für Compact Disc - Read Only Memory (nur lesbarer Speicher, was sie von gewöhnlichen Audio-CDs unterscheiden soll). Ein Beispiel dafür wäre - aus meiner eigenen kleinen Sammlung - das „Archiv der Gegenwart“, früher „Keesings Archiv“, eine weltumspannende Dokumentation zu Ereignissen der Politik, die seit 1931 auf der Basis der Auswertung einer Vielzahl von Nachrichtenbulletins, Pressemitteilungen, Tageszeitungen, Nachrichtenmagazinen und offiziellen Dokumentationen erarbeitet wird und Monat für Monat in Druckform erscheint.³ Zu Jahrgängen zusammengebunden ergeben sie jeweils einen stattlichen Band; der Jahrgang für 1989 etwa hat über 1000 großformative Seiten. Die Bände für die Jahre von 1931 bis 2000 füllen ein kleines Regal. Ihre Informationen (mithin von über 50.000 Seiten) passen unschwer alle auf eine CD.

Die Seiten sind dabei nicht nur als Bilder („Images“) abgespeichert, sondern Buchstabe für Buchstabe, Wort für Wort (als „Volltext“, wie man auch sagt), was es möglich macht, über die Bandgrenzen und das jährliche Register hinaus nach Personen und Ereignissen, aber auch nach Orten, Daten und Begriffen zu suchen, die das gedruckte Register *nicht* enthält. Neben der Platzersparnis ein deutlicher „Mehrwert“ der digitalen gegenüber der gedruckten Form.

Möglich wird dies durch einen enormen Speicherplatz. Hatten die erst 8 Zoll, dann 5 ¼ Zoll großen flexiblen Disketten („Floppys“) zunächst einen Speicherplatz von weniger als 100 (später 360) Kilobytes, so passen auf eine (im Umfang sogar kleinere) CD-ROM 650 MB (also 6,5 Millionen mal so viele Informationen). Auch wenn man sich ins Gedächtnis ruft, daß bei einem Ende der 80er Jahre gekauften Computer eine eingebaute Festplatte von 20 bis 40 MB (mit einem Fassungsvermögen von 10.000 bis 20.000 Schreibmaschinenseiten schon als völlig ausreichend erschien, wird deutlich, welchen Qualitätssprung die 15- bzw. 30fache Speicherfläche der CD-ROM bedeutete.

Wer sich weniger mit zeitgeschichtlichen Ereignissen als mit spätantiken/frühmittelalterlichen Texten beschäftigt, wird das Beispiel des „Migne“ noch überzeugender finden Die Patrologia Latina, die über 200 Bände umfassende Textsammlung lateinischer Kirchenväter aus der Zeit zwischen 200 und 1216 n.Chr., die (herausgegeben von Jacques-Paul Migne) in den 40er und 50er Jahren des 19. Jahrhunderts erstmals erschien⁴, wurde in den 1990er Jahren vom britischen Verleger Chadwyck-Healey auf CD-ROM herausgebracht, äußerst komfortabel zu benutzen, mit den oben skizzierten Möglichkeiten der Volltext-Recherche über die Bände und Jahrhunderte hinweg. Bibliotheken und Institute, die das ihren Benutzern möglich machen wollten, wurden allerdings heftig zur Kasse gebeten (sofern ich es aus der BSB richtig weiß, kostete die Einzelplatznutzung 25.000 Pfund). Da die Nachfrage dennoch, wie man hörte, rege war, die Vorteile der digitalen Nutzungsform auf der Hand lagen, der CD im Unterschied zur Diskette eine relativ hohe Lebenserwartung vorausgesagt wurde, optimistische Schätzungen von einer Haltbarkeit von 40, ja 100 Jahren sprachen, Firmen wie Chadwyck-Healey bereits weitere große Projekte planten (im Bereich der englischen Literatur, aber auch für Luther

³ Archiv der Gegenwart. Die weltweite Dokumentation für Politik und Wirtschaft, gegr. von Heinrich v. Siegler. Königswinter 1931 ff.; auch in englischer, französischer und niederländischer Ausgabe; 1945-1955 als Keesing's Archiv der Gegenwart (Bonn).

⁴ Patrologia cursus completus: sive bibliotheca universalis [...] omnium ss. patrum, doctorum scriptorumque ecclesiasticorum qui ab aevo apostolico ad usque Innocentii III tempora floruerunt, accurante J.-P. Migne, Parisiis 1844-1855; 1862-1865, 221 Bde.

und Goethe, Kafka und Brecht), lag es nahe, in der CD das Medium der Zukunft zu sehen, das bei vielbändigen Reihenwerken (Bibliographien, Lexika, Text-, Bild- und Kartensammlungen) das Buch in absehbarer Zeit ersetzen würde. Ja, es schien an der Zeit, generell nach der Rolle der elektronischen Medien in der Bibliothek der Zukunft zu fragen.

2. Die Rolle der elektronischen Medien in der „wissenschaftlichen Bibliothek der Zukunft“ und die Förderpolitik der Deutschen Forschungsgemeinschaft

Da inzwischen in den Bibliotheken als den klassischen Einrichtungen der Informationsversorgung über den in Zukunft zu erwartenden quantitativen und qualitativen Bedarf an elektronischen Serviceangeboten erhebliche Unsicherheiten bestanden, setzte der Bibliotheksausschuß der DFG auf seiner Frühjahrstagung 1994 zum Themenkomplex „Elektronische Publikationen“ eine Arbeitsgruppe ein, die im Frühsommer 1995 eine entsprechende Denkschrift vorlegte.⁵ Teilweise parallel dazu hatte eine Arbeitsgruppe des Bibliotheksausschusses und der Kommission für Rechenanlagen der DFG den Handlungsbedarf im Bereich „Neuer Informations-Infrastrukturen für Forschung und Lehre“ sondiert und dazu „Empfehlungen“ gegeben.⁶ Beider Überlegungen griff eine weitere Arbeitsgruppe auf, deren Aufgabe es war, das „Memorandum zur überregionalen Literaturversorgung“ fortzuschreiben und die ihren Entwurf der Neufassung 1997 vorlegte.⁷ Aufeinander aufbauend und sich ergänzend entwickelten die genannten Gremien das Leitbild einer „Verteilten digitalen Fachbibliothek“. Diese „verteilte“, „digitale“, „virtuelle Fachbibliothek“ sollte eine fachliche Einheit bilden, ohne daß dafür ihre Bestandteile in Wirklichkeit an einem Ort versammelt sein müssen. Dabei sollte die jeweilige Sondersammelgebietsbibliothek Leitfunktion übernehmen und ein entsprechendes Fachinformationssystem aufbauen, d.h. Göttingen für die Reine Mathematik und die Geowissenschaften, die BSB München für Geschichte und Osteuropa, Saarbrücken für die Psychologie, Erlangen für die Philosophie, Frankfurt am Main für die Germanistik usw. usf. Zugang zu diesem Fachinformationssystem von außerhalb sollte die Homepage der Leitbibliothek schaffen. Dieses Fachinformationssystem sollte in einem gemeinsamen elektronischen Katalog die Print- und elektronischen Medien verzeichnen, erlauben, in den entsprechenden Fachbibliographien zu recherchieren und diese Recherchen nach Möglichkeit mit einer Bestellfunktion, einem Dokumentenlieferdienst zu verbinden; es sollte Einblick geben, was auf den Servern anderer wichtiger Facheinrichtungen liegt und dem Benutzer ermöglichen, per Mausclick dorthin zu gelangen.

Zu dieser „Virtuellen Fachbibliothek“ sollte zugleich ein „Virtueller Lesesaal“ gehören. Darin sollten nicht nur die bereits erwähnten Fachbibliographien stehen; zu einem Lesesaal gehören schließlich auch für die Wissenschaft einschlägige Nachschlagewerke, Lexika, Fachzeitschriften. Sie lagen in aller Regel nicht in elektronischer, sondern in Druckform vor. Um das zu ändern, schuf die DFG ein eigenes (seit 1997 bestehendes) Förderprogramm zur „Retrospektiven Digitalisierung von Bibliotheksbeständen“.

Zu einem Gutteil war diese DFG-Entscheidung Reaktion: Sie schuf mit dem neuen Förderprogramm ein Pendant zu ähnlichen, teilweise noch sehr viel ehrgeizigeren Projekten in England, Frankreich und den USA. So hatten die Leiter der 15 größten US-amerikanischen Bibliotheken und Archive sowie der „Commission on Preservation and Access“ am 1. Mai 1995 eine Vereinbarung zur Zusammenarbeit beim Aufbau einer „National Digital Library Federation“ unterzeichnet. Die Library of Congress hatte \$ 15 Millionen an staatlicher Unterstützung und Spenden für die Digitalisierung ihrer historischen Dokumente zur Geschichte der USA erhalten. Bis zum Jahr 2000 sollten, so die Ankündigung, in einer „National Digital Library“ 5 Millionen „items on American History“ gespeichert und über das Internet online abrufbar sein oder auf CD-ROM vertrieben werden. Im europäischen Bereich wollte die Bibliothèque Nationale de France bis 1996 etwa 100.000

⁵ Elektronische Publikationen im Literatur- und Informationsangebot Wissenschaftlicher Bibliotheken. Positionspapier des Bibliotheksausschusses, Bonn (Juni) 1995; in elektronischer Form (wie auch die folgend genannten Dokumente) unter: <http://www.dfg.de/foerder/biblio/publikationen.html>

⁶ Neue Informations-Infrastrukturen für Forschung und Lehre. Empfehlungen des Bibliotheksausschusses und der Kommission für Rechenanlagen, Bonn (Dezember) 1995.

⁷ Weiterentwicklung der überregionalen Literaturversorgung. Memorandum, Bonn (Februar) 1998.

Bücher und Mikroformen, bis zur Jahrtausendwende 300.000 Bände digitalisieren. Vergleichbare Projekte waren in Großbritannien angelaufen. Darauf wollte und mußte die DFG reagieren.

Zur Vorbereitung des eigenen Retrodigitalisierungsprogramms hatte der Bibliotheksausschuß der DFG zwei Arbeitsgruppen eingesetzt, die Empfehlungen (erstens) zur inhaltlichen Auswahl der Bibliotheksmaterialien und (zweitens, worauf ich hier nicht näher eingehen will) der technischen Durchführung der retrospektiven Digitalisierung ausarbeiten sollten. Sie trafen sich - unter Göttinger Federführung - jeweils zu drei Sitzungen und legten Anfang 1997 die vorläufigen Ergebnisse ihrer Beratungen vor.⁸

Entscheidende Kriterien für die inhaltliche Auswahl sollten

- der wissenschaftliche Wert,
- das aktuelle Interesse der Forschung,
- die bisherige Nutzung der Bestände,
- der Grad der Verbesserung der Benutzbarkeit durch die Digitalisierung
- und die Einbindung in internationale Digitalisierungsmaßnahmen und Forschungsvorhaben sein.

Das alles machte die enge Zusammenarbeit von Bibliothekaren und Fachwissenschaftlern zur *conditio sine qua non*. Bei der Auswahl konnte auch der Aspekt des Schutzes singulärer und bestandswürdiger Materialien sowie die Verbesserung des Zugangs zu schwer zugänglichen Materialien eine Rolle spielen.

Bei den Textgattungen sah man, dem Leitbild des Virtuellen Lesesaals folgend, vorrangig die Digitalisierung von:

- Enzyklopädien und Nachschlagewerken,
- Biographischen Lexika,
- Bibliographien, Katalogen und Verzeichnissen,
- Sprachwörterbüchern von historischem Wert
- sowie Kultur-, Literatur- und Fachzeitschriften vor (wobei das amerikanische JSTORE-Unternehmen das Vorbild bildete).

Schon kritischer sah man - und durchaus zurecht - die Digitalisierung von:

- Monographischer Literatur und
- Handbüchern.

In modellhaften Versuchen sollte ferner die Einbeziehung von

- historischen Quellen und literarischen Texten,
- von archivalischen Quellen
- und von Bildquellen

getestet werden.

Über eine Sache hatte sich der für die Betreuung des Förderprogramms eingesetzte Gutachterausschuß von Anfang an nicht zu beklagen: über einen Mangel an Anträgen. Schon bei der ersten Sitzung im Januar 1997 waren es 14; bis zum Herbst 2000 waren fast 80 bei der Geschäftsstelle der DFG eingegangen (wovon 56 bewilligt, 21 abgelehnt wurden); inzwischen sind noch etliche, gewichtige Anträge hinzugekommen. Zu den Antragstellern gehörten keineswegs nur Bibliotheken, sondern auch Akademien, Max-Planck- und Hochschul-

⁸ Retrospektive Digitalisierung von Bibliotheksbeständen für eine Verteilte Digitale Forschungsbibliothek, Bericht der Arbeitsgruppe Inhalt, Berlin 1998; Retrospektive Digitalisierung von Bibliotheksbeständen. Bericht der Arbeitsgruppe Technik, Berlin 1998.

institute, Archive. Ähnlich breit war das Fächerspektrum, es reichte von der Ägyptologie, den Altertumswissenschaften und der Altorientalistik; über die Mediävistik, die Neuere und Neueste Geschichte; die Kunstgeschichte, Literatur- und Musikwissenschaft; die Pädagogik, Theologie und Rechtswissenschaft; bis zur Geschichte der Naturwissenschaften und der Mathematik, was zugleich heißt, daß die geisteswissenschaftlichen, historisch arbeitenden Disziplinen bei weitem überwogen.

Was herauskam, erinnert in manchem, ja vielem an das Leitbild des „Virtuellen Lesesaals“ und geht doch zugleich weit darüber hinaus. Selbstverständlich sind unter den Textgattungen jene stark vertreten, denen schon die Arbeitsgruppe „Inhalt“ hohe Priorität einräumte: Enzyklopädien, Nachschlagewerke, Sammlungen und Wörterbücher. Dazu gehören

- das Große Vollständige Universal-Lexicon aller Wissenschaften und Künste von Johann Heinrich Zedler (68-bändig, die wichtigste deutsche Enzyklopädie des 18. Jahrhunderts, und in der digitalisierten Form, wie ich in der Bayerischen Staatsbibliothek hörte, ein „Renner“);
- die Digitalisierung und Kumulierung der Register der 19-bändigen Neuen Deutschen Biographie und der 56-bändigen Allgemeinen Deutschen Biographie;
- die „Regesta imperii“ (jenes Grundlagenwerk, dessen erster Band 1831 erschien und das sich zum Ziel gesetzt hat, alle urkundlichen und historiographischen Quellen der römisch-deutschen Kaiser/Könige von den Karolingern bis zu Maximilian I. sowie der Päpste des frühen und hohen Mittelalters zu verzeichnen, bis jetzt auf über 22.000 Seiten);
- die Sammlungen preußischer Rechtsquellen, des Corpus Constitutionum Marchicarum (CCM) und des Novum Corpus Constitutionum Marchicarum (NCC), mit Quellentexten vom Ende des 13. bis zum Anfang des 19. Jahrhunderts;
- das Deutsche Rechtswörterbuch, das seit 1932 erscheint (mit seinen heute über 80.000 Wortartikeln);
- die digitale Aufbereitung und elektronische Publikation von Georg Wenkers „Sprachatlas des Deutschen Reiches“, des ersten und bis heute umfangreichsten Sprachatlas überhaupt, aufbauend auf Erhebungen (zwischen 1876 und 1887) an 40.000 Orten und mit dem Versuch, die dialektologischen Ergebnisse unter anderem auf 1.647 handgezeichneten Karten in 22 Farben festzuhalten⁹;
- das Grimmsche Wörterbuch (das 1854-1971 in 32 Teilbänden und einem Registerband erschien);
- das Grammatisch-kritische Wörterbuch der hochdeutschen Mundart von Johann Christoph Adelung (das maßgebliche Wörterbuch des 18. Jahrhunderts);
- die mittelhochdeutschen Wörterbücher von Lexer und Benecke/Müller/ Zarncke (die Liste ließe sich fortsetzen).

Es gibt eine Reihe weiterer Materialien, die man ohne weiteres zu den „Lesesaalbeständen“ rechnen könnte:

- wie die Digitalisierung eines großen, in die Tausende gehenden historischen Kartenbestandes;
- die Digitalisierung von Städte- und Landschaftsansichten in den Buchillustrationen des 15., 16. und 17. Jahrhunderts;
- die Digitalisierung ausgewählter Emblembücher aus der frühen Neuzeit;
- die Digitalisierung eines kolonialen Bildarchivs und unter anderem auch
- eines Nationalen Bildarchivs zur Kunst und Architektur, mit 1,3 Millionen Fotografien zur Kunst und Architektur in Deutschland, die in den vergangenen 20 Jahren nach und nach aus Denkmalämtern, Museen und Bibliotheken in Marburg zusammengeführt, redaktionell geordnet, auf Mikrofiche verfilmt und dann den Eigentümern zurückgegeben wurden.

Zum „virtuellen Lesesaal“ würden auch die digitalisierten Bibliographien, die Kunst-, Literatur- und kulturhistorischen Zeitschriften gehören (die Jahrbücher für deutsche Geschichte, die Schriften der Königlich Preußi-

⁹ Vgl. <http://www.uni-marburg.de/das>

schen Akademie zwischen 1700 und 1900, führende Rezensionenorgane und Literaturzeitschriften des 18. und 19. Jahrhunderts, die Pädagogischen Zeitschriften aus dem Zeitraum von 1760 bis 1870, jüdische Periodika aus dem deutschsprachigen Raum zwischen 1837 und 1938, die Zeitungen und Zeitschriften des deutschsprachigen Exils 1933-1945, die Digitalisate aus dem Jahrbuch über die Fortschritte der Mathematik usw.); erst recht ein neues digitales historisches Rezensionenorgan wie die „Sehepunkte“ in München. Auch das Gegenstück zum amerikanischen JSTORE-Projekt, die systematische retrospektive Digitalisierung und Bereitstellung von Fachzeitschriften im engeren Sinne, ist (unter Göttinger Federführung und dem Akronym „Digi-Zeit“) inzwischen angelaufen.¹⁰ Daß von den genannten Beständen manche durch die Digitalisierung nicht nur einen Mehrwert, sondern eine neue Qualität gewinnen, sei nur am Rande angemerkt.

Doch nicht alles läßt sich, so oder so, in das Bild des „Virtuellen Lesesaals“ pressen. Ich erwähne aus einer Vielzahl nur sechs, die alle auch mit unserem engeren Thema, den „Quelleneditionen“, zu tun haben:

1. die Digitalisierung Tausender von Papyri (in Gießen, Heidelberg und Köln), wobei das Unternehmen in einem internationalen Rahmen erfolgt (zumindest nach den Vorstellungen der DFG erfolgen soll), der in einer gemeinsamen Datenbank (APIS, Advanced Papyrological Information System) wieder zusammenzuführen versucht, was einst zusammengehörte und durch die Papyrikartelle, die Universitätsinstitute im 19. / beginnenden 20. Jahrhundert mit „Material“ versorgten, auseinandergerissen wurde¹¹;
2. die Digitalisierung der Turfanhandschriften (es handelt sich dabei um einen Teil der 40.000 Fragmente, die zu Beginn dieses Jahrhunderts in der Oase Turfan und in benachbarten Gebieten gefunden wurden und in der Regel der 2. Hälfte des 1. Jahrhunderts nach Chr. entstammen); sie befinden sich im Besitz der Berlin-Brandenburgischen Akademie der Wissenschaften, an ihrer Entzifferung wird noch lange Zeit gearbeitet werden und der Sinn ihrer digitalen Veröffentlichung besteht darin, sie auf diese Weise - unter Schonung der fragilen Originale - einem „breiteren“ Kreis von Experten zugänglich zu machen¹²;
3. die Digitalisierung des gesamten mittelalterlichen Handschriftenbestand der Diözesan- und Dombibliothek zu Köln (d.h. von ca. 400 Codices mit ca. 130.000 Seiten), deren Bereitstellung für die Forschung dazu beitragen soll, der Kanonisierung des Textkorpus der mittelalterlichen Geschichte gegenzusteuern, wobei die Manuskripte in ein Erschließungssystem eingebettet werden, das unter Rückgriff auf bereits vorliegende Dokumentationen die Objekte schon unmittelbar nach ihrer Digitalisierung bereitstellt, sie andererseits aber durch Einbindung von Sekundärliteratur in einen größeren Kontext bringt¹³;
4. die Digitalisierung von rund 500 relativ seltenen Drucken des 17. Jahrhunderts zur Festkultur des Barock, gesammelt in der Herzog August Bibliothek Wolfenbüttel, die damit einer an der Kulturgeschichte interessierten, interdisziplinären Frühneuzeitforschung als Quellensammlung zur Verfügung stehen werden¹⁴ (was sich gut zusammenfügt mit dem Aufbau des Servers „Frühe Neuzeit“¹⁵ durch die Bayerische Staatsbibliothek und das Historicum München, was von der DFG mitgefördert wurde);
5. die Digitalisierung der wohl umfangreichsten und geschlossensten Sammlung von Flugblättern, Anschlägen, Aufrufen, Plakaten, Liedern, Karikaturen und Protokollen zur Revolution von 1848 im deutschsprachigen Raum, die zuvor sicherheitsverfilmt und optisch so geglättet wurden, daß eine kostengünstige Übertragung per Scanning möglich war. Sie sollten (als Projekt der Stadt- und Universitätsbibliothek Frankfurt) einer an kul-

¹⁰ Vgl. die Aufstellung mit einer Kurzbeschreibung der Projekte unter <http://www.sub.uni-goettingen.de/de/vdf-dvdf-liste.shtml>

¹¹ Vgl. als Beispiel http://www.uni-giessen.de/ub/ub_allg/papyri.html#pap_digit

¹² Vgl. http://www.bbaw.de/vh/turfan/dta/dta_main_index.html

¹³ Vgl. http://www.dbi-berlin.de/projekte/d_lib/einzproj/retrodig/p24.htm

¹⁴ Vgl. <http://www.hab.de/forschung/de/vdf/index.htm>

¹⁵ Vgl. <http://www.sfn.uni-muenchen.de>

tur- und mentalitätsgeschichtlichen Fragestellung interessierten Revolutionsforschung als eher ungewöhnliche Quellengattung zur Verfügung gestellt werden¹⁶;

6. die digitale Erschließung, Speicherung, Archivierung und Bereitstellung der Stenographischen Berichte des Deutschen Reichstags von 1867-1895, ein vielbändiges Werk, dessen Register (General- und Sprechregister) im Volltext erfaßt und mit den (im Imageformat digitalisierten) 107.695 Seiten der 165 Einzelbände verlinkt wurden, ein Unternehmen, das allein durch seine Größe die digitale Aufbereitung vor erhebliche Probleme stellte, andererseits aber richtungweisend für die Behandlung ähnlicher Projekte (man denke nur an Gesetzes-sammlungen) sein könnte¹⁷, vor allem dann, wenn die Frakturschrift keine befriedigende OCR-Ergebnisse liefert und der Aufwand einer nochmaligen Abschrift nicht finanzierbar erscheint.

Ich habe im Herbst 2000, nach 4 Jahren Laufzeit des DFG-Förderprogramms, die Geschäftsstelle gebeten, es auszählen zu lassen: Es wurden Bewilligungen für 9.568.468 Seiten ausgesprochen, damit sie mit Unterstützung des DFG-Förderprogramms retrospektiv digitalisiert, eingescannt und mit dem Inhaltsverzeichnis verlinkt, mit OCR-Software behandelt, in Südostasien neu abgeschrieben werden können; unter diesen Seiten sind weit über 1 Million bildliche Darstellungen und Tausende von Karten. Seither sind - überschlagsweise - 3,6 Millionen hinzugekommen, so daß von einer Gesamtzahl von rd. 13 Millionen Seiten auszugehen ist. Darunter auch, wie wir gesehen haben, eine nicht unerhebliche Anzahl historischer Quelleneditionen, wobei die „Edierung“ unterschiedliche „Tiefe“ aufweist und unterschiedliche, disziplinspezifische Ziele verfolgt. Ich will und kann diesen Punkt hier nicht vertiefen.

Die DFG ist selbstverständlich nicht die einzige Institution, die Digitalisierungen anregt und fördert. Wer sich in den elektronischen Katalog einer großer Staats- oder Universitätsbibliothek¹⁸ einwählt und etwa unter der Rubrik „Volltextressourcen“ ihren Linksammlungen folgt, stößt auf eine Vielzahl von Initiativen. Man sehe sich, davon ausgehend, nur einmal das „Internet Ancient History Sourcebook“¹⁹, das „Internet Medieval Sourcebook“²⁰ und das „Internet Modern History Sourcebook“²¹ der New Yorker Fordham University an; oder die Seite des Historical Text Archive²²; oder „EuroDocs: Primary Historical Documents From Western Europe“²³; oder - mehr literaturgeschichtlich angelegt - das Oxford Text Archive²⁴; oder das „Electronic Text Center at the University of Virginia“²⁵; oder die Seiten von „Projekt Gutenberg - DE“²⁶; „Project Wittenberg“²⁷; „Project Runeberg“²⁸ usw. usf. Wer bei der Suchmaschine „google“ nach „Digitale[n] Archive[n]“ sucht, bekommt nach 0,39 Sekunden das Ergebnis von „ungefähr 33.900“ Treffern; wer „American Heritage“ eingibt (den Namen des „Virtuellen Archiv Projektes“ der Universitäten Berkeley, Stanford, Duke und Virgi-

¹⁶ Vgl. <http://dbib.uni-frankfurt.de/1848/index.html>

¹⁷ Vgl. <http://mdz.bib-bvb.de/digbib/reichstag>

¹⁸ Vgl. unter http://www.ub.fu-berlin.de/literatursuche/literatur_im_web/volltext/sammlungen.html oder unter http://www.ub.fu-berlin.de/internetquellen/fachinformation/geschichte/digital_archives.html im elektronischen Angebot der Universitätsbibliothek der Freien Universität Berlin.

¹⁹ Vgl. <http://www.fordham.edu/halsall/ancient/asbook.html>

²⁰ Vgl. <http://www.fordham.edu/halsall/sbook.html>

²¹ Vgl. <http://www.fordham.edu/halsall/mod/modsbook.html>

²² Vgl. <http://historicaltextarchive.com>

²³ Vgl. <http://library.byu.edu/~rdh/eurodocs>

²⁴ Vgl. <http://ota.ahds.ac.uk>

²⁵ Vgl. <http://etext.lib.virginia.edu>

²⁶ Vgl. <http://www.gutenberg2000.de>, wo man allerdings erfährt, daß der Sponsor AOL das Projekt zur Zeit „aus technischen Gründen“ vom Netz genommen hat.

²⁷ Vgl. <http://www.iclnet.org/pub/resources/text/wittenberg/wittenberg-home.html>

²⁸ Vgl. <http://www.lysator.liu.se/runeberg>

nia), verbucht nach 0,08 Sekunden „ungefähr 326.000“ Treffer; und wer mehr über „Making of America“, das Projekt der Universitäten von Michigan und Cornell zur amerikanischen Sozialgeschichte wissen möchte, in das sich dann auch die „Digital Library Federation“ und die Library of Congress mit ihrem „American Memory“-Programm einklinkten (ein „gateway“ zu Primärquellen der amerikanischen Kultur und Geschichte, der heute Zugang zu „more than 7 million digital items from more than 100 historical collections“ bietet²⁹), bekommt, wenn er „google“ danach fragt, nach wenigen Sekunden über 2,8 Millionen Treffer angezeigt, was ihn vermutlich dazu bringt, auf weitere Informationen freiwillig zu verzichten.

Aber widerspreche ich mir damit nicht selbst, meiner anfänglichen Behauptung, daß das Thema nicht lohne? Das führt mich zu meinem 3. Punkt:

3. CD-ROM als Datenträger

Nein, die Anfangsbehauptung ist damit nicht widerlegt, denn kaum eine der genannten Volltexteditionen erschien primär oder auch auf CD. Dafür lassen sich mehrere Gründe nennen. Eher von untergeordneter Bedeutung ist wohl, daß

1. die Haltbarkeit der CDs mittlerweile differenzierter gesehen wird. Sie hängt mit von der Beschichtung (Silber, Gold oder Platin) ab, was sie in unterschiedlicher Weise von chemischen Zerfallsprozessen abhängig macht. So gibt man einer silbernen „selbstgebrannten“ CD-ROM bei ungünstigen Bedingungen nur 10-15, einer goldenen 20-25 Jahre, einer platinenen 30-40 Jahre. Wichtiger scheint jedoch die technische Entwicklung, denn
2. das Nachfolgermodell, die DVD (die Digital Versatile Disc), mit höherer Spur- und Pitdichte steht bereits in den Startlöchern. Sie weist mit 4,2 bis 8,4 GB eine zehnfach höhere Speicherkapazität auf, Kapazitäten bis zu 17 GB scheinen möglich. Zwar sind auch die Rohlinge (mit 10 Euro) noch wesentlich teurer, aber mit wachsender Stückzahl ist auch hier eine Preissenkung zu erwarten. Daß die CD-ROMs bei der Bereitstellung und Vorhaltung der Daten eine geringere Rolle spielen, als früher vielfach erwartet, hängt
3. auch mit der Entwicklung von Alternativen, der Magnetbänder und vor allem der Festplatten ab. Festplatten von 160 GB sind heute schon für 300 Euro zu haben. Auf sie passen 250 CD-ROMS. Damit entfallen nicht nur die mechanischen Probleme der Abspielung (über CD-Roboter). Auch der Zugriff ist bei der Platte 10-15 mal schneller als bei der CD-ROM. Der Hauptgrund aber dürfte sein, daß
4. das Netz sich sehr viel rasanter entwickelte, als das ursprünglich vorauszusehen war, so daß auch größere Mengen von Daten über das Netz übermittelt, verschickt und recherchiert werden können, ohne zusätzlich lokal vorgehalten zu werden, daß über das Netz selbst Verfahren kommerzieller Datennutzung abgewickelt werden können, ohne eine CD-ROM „erwerben“ zu müssen.

Das alles löst zwar nicht das „alte“ Problem der Datenlangzeitarchivierung, aber wenn ich recht sehe, wird es mittlerweile weniger „prinzipiell“, eher locker, sportlich, optimistisch gesehen: Bei der Mobilität, Schnelligkeit und Schnelllebigkeit des Objekts sei nicht gleichzeitig Dauerhaftigkeit zu erwarten. So müßten die Datenmengen eben alle 10 Jahre auf neue Standards und Träger überspielt werden. Das sei mit der erreichten und weiter zu erwartenden Normierung - vorausgesetzt, sie werde eingehalten - möglich und durchaus bezahlbar. Wir müssen das wohl glauben - zumal, was wäre die Alternative?

Mit der rasanten Entwicklung des Netzes und seinen anarchischen Strukturen hängt auch die neue Unübersichtlichkeit zusammen, von der die obigen Ausführungen zu meinen Netzrecherchen mittels marktüblichen Suchmaschinen bereits eine Vorstellung gaben. Umso wichtiger wäre der rasche Auf- und Ausbau der Fachinformationssysteme, zu der - aus meiner Sicht - jeweils nicht nur die Aktualisierung der Links, sondern auch die Ordnung und Bewertung der Informationsmassen, ihre Sichtung, Filterung und Hierarchisierung gehören muß. Doch das ist - so oder so - ein weites Feld.

²⁹ Vgl. <http://memory.loc.gov>