

Das ist doch alles im Netz!

- Angebot und Nutzen von digitalisierten Altbestandsquellen im Internet

Thomas Stäcker (HAB Wolfenbüttel)
Vortrag Düsseldorf, Bibliothekartag, 15.3.2005

Abstract

Inkunabeln und alte Drucke der Zeit bis 1850 werden in zunehmendem Maße auch über Internet als digitale Ressource zur Verfügung gestellt. Das Angebot reicht von Einzeldigitalisierungen bis zur Bereitstellung umfangreicher Sammlungen, von Imagedigitalisaten bis zu Volltexten. Vorhandene Ressourcen sollen gesichtet und auf Ihre Zugänglichkeit und Nutzbarkeit hin bewertet sowie Perspektiven für kooperative Entwicklungsschritte und Digitalisierungskampagnen skizziert werden.

Mit dem Wachstum digitaler Sammlungen von Handschriften, Inkunabeln und Drucken stellt sich zunehmend die Frage nach deren Zugänglichkeit und Nutzbarkeit. Dabei haben sich in der Beurteilung dieser Aspekte in den letzten Jahren die Ansprüche gewandelt. Während in der Frühzeit der Digitalisierung noch das *factum brutum* der Bereitstellung im Netz und bis vor kurzem die effiziente Befriedigung von Bedürfnissen einer definierten Nutzungsklientel als ausreichend erachtet wurde, verlangt man heute von digitalen Angeboten, dass sie als wieder verwertbare Bausteine im Kontext weltweiter Ressourcen konzipiert werden.¹ Das bedeutet, digitale Sammlungen müssen "für ihre Wiederverwendbarkeit, Interoperabilität, Persistenz und Verifizierbarkeit qualifiziert" sein² (ib. S.136).

Aus der Sicht des wissenschaftlichen Nutzers gesprochen, der von digitalisierten Altbestandsquellen Gebrauch machen will, kann man diese Anforderungen auf drei Fragen bringen. Wie leicht werden Digitalisate gefunden? Wie ist deren Nutzbarkeit und wie kann ich sie für meine wissenschaftliche Arbeit weiterverwenden?

Die Frage nach der Findbarkeit von von Bibliotheken angebotenen Ressourcen ist besonders heikel. Beanspruchen doch –klassischerweise - die wissenschaftlichen Bibliotheken wie kein anderer öffentlicher Institutionenverband das Feld der Informationsvermittlung und der qualitativen Medienbereitstellung für sich. Ob sich dieser Anspruch im Bereich von digitalen Angeboten zum alten Buch einlösen lässt, soll Gegenstand dieses Artikels sein. Zu prüfen sind die Angebote von Handschriften, Inkunabeln und alten Drucken. Die Nutzbarkeit und Möglichkeit der wissenschaftlichen Weiterverwendung wird bei den einzelnen Angeboten erläutert. Eine detaillierte Analyse wird nicht unternommen, sondern der *ad hoc* Zugriff und unmittelbare Eindruck bewertet, was auch der Nutzersicht am ehesten entsprechen dürfte.³

¹ Vgl. die 3. Stufe der IMLS (Institute of Museum and Library Science) nach Manfred Thaller [Gesamtredaktion]: Retrospektive Digitalisierung von Bibliotheksbeständen. Evaluierungsbericht über einen Förderschwerpunkt der DFG Retroeval, Köln 2005, S. 136 (wird in Kürze im Netz über die Seite der DFG (www.dfg.de) veröffentlicht).

² Ebd. S. 136

³ Zu qualitativen Anforderungen an Digitalisate im Bereich altes Buch vgl. Stäcker, Thomas: Erschließungsformen Alter Drucke im Internet, in: Margit Rützel-Banz (Hrsg.): 91. Deutscher Bibliothekartag in Bielefeld 2001. Bibliotheken- Portale zum globalen Wissen. Frankfurt a.M. 2001 (ZfBB Sonderheft, 81), S. 134-148.

Natürlich können hier nur Stichproben gemacht werden. Die beschriebenen Fälle sind aber soweit exemplarisch, dass sie leicht auf andere Angebote übertragbar sind. Neben dem Blick auf die großen Anbieter, galt es auch die Nebenstrassen abzuschreiten, da es hier überraschend viel zu entdecken gibt. Als Kriterien der Handhabbarkeit bzw. Nutzbarkeit der hier unternommenen Kurzbewertung sind Navigation im Buch, Qualität der Bilder, Zitierfähigkeit bzw. Persistenz und Kontextinformationen.

Die Ergebnisse der Bewertung werden mit der DFG geförderten Studie des Instituts für Historisch-Kulturwissenschaftliche Informationsverarbeitung (HKI)⁴ verglichen und insgesamt Vorschläge für den weiteren Ausbau digitaler Bibliotheken gemacht.

Wissenschaftliche Nutzer verfolgen in der Regel zwei Suchstrategien. Zum einen die formale, zum anderen die thematische Suche. Im ersten Fall wird ein konkreter Titel oder eine bereits vorliegende Information gesucht. Der Nutzer hat definierte Schlüsselwörter oder Titelinformationen, nach denen er suchen kann. Im zweiten Fall sucht der Nutzer allgemein Literatur oder Informationen zu einem Gegenstand, besitzt nur eine vage Vorstellung, von dem, was ihn erwartet und findet ggf. Interessantes, was er gar nicht gesucht hat.⁵ Naturgemäß ist der zweite Bereich nur sehr schwer abzubilden und folglich qualitativ nicht leicht zu bewerten, da es so gut wie nicht vorhersehbar ist, welchen Suchweg der Nutzer einschlägt und welches Angebot das für ihn wichtige war.

Die hier leitende Fragestellung nach der Qualität eines Angebotes beschränkt sich auf die formale Suche, geht also davon aus, dass der Nutzer nach bestimmten Schlüsselwörtern und Materialien sucht.⁶ Ausgang wird von einem vorhandenen Digitalisat genommen. Das ist ein Wissen, das der Nutzer normalerweise nicht hat. Man muss daher damit rechnen, dass er die Suche schnell abbricht, wenn er unmittelbar keine Ergebnisse oder weiterführenden Hinweise findet. Die Überprüfung erfolgt in zwei Schritten, zunächst wird in den bibliothekarischen Hilfsmitteln und Katalogen recherchiert, sodann eine Suche in Google gestartet. Die meisten Nutzer werden es vermutlich umgekehrt machen.

Ich werde insgesamt 9 Beispiele vorstellen, darunter eine Handschrift, eine Inkunabel und sieben alte Drucken. Nachlassmaterialien und Autographen werden aus Zeitgründen hier nicht behandelt.

1. Beispiel (Handschrift) Suche nach den Incipit-Stichworten *Parabulae Salom(o)n(i)s* aus Beda Venerabilis super Parabolas Salomonis [Angebot: CEEC]⁷

Das zentrale deutsche Portal DFG geförderter Projekte für den Nachweis mittelalterlicher Handschriften ist *Manuscripta mediaevalia* bzw. die Handschriftendatenbank.⁸ Die Suche nach dem *Incipit* führt zu etlichen Treffern, jedoch zu keinem Digitalisat, auch innerhalb der Rubrik digitalisierte Handschriften⁹ wird der Nutzer nicht fündig. Die Kölner Daten sind nicht eingespielt.

⁴ <http://www.hki.uni-koeln.de>

⁵ Zu diesem Komplex vgl. Wegmann, Nikolaus: Bücherlabyrinth. Suchen und Finden im alexandrinischen Zeitalter. Böhlau, Köln u.a. 2000.

⁶ Auch in diesem Bereich gibt es natürlich Unschärfen. Da sind zum einen nicht normalisierte oder normierte Schreibweisen, zum anderen im Internetkontext das Problem mehrsprachiger Angebote.

⁷ <http://www.ceec.uni-koeln.de/ceec-cgi/kleioc/0010/exec/pagedmed/%22kn28%2d0105%5f001%2ejpg%22/segment/%22body%22> [2005-03-13]

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

⁹ <http://www.manuscripta-mediaevalia.de/hs/hs-online.htm>

Die Google-Suche führt indessen zum Erfolg, zumindest wenn man versiert genug ist, *parabola* nicht mit der klass. Endung –ae, sondern der mittelalterlichen –e zu suchen.

Das Angebot von CEEC ist gut strukturiert, die Navigation und Suchmöglichkeiten sind intuitiv verständlich und leicht nutzbar. Die präsentierten Metadaten und weiterführenden Informationen (Handapparat) als Kontextinformationen sind vorbildlich. Die Bildqualität ist hervorragend, die höchste Auflösung lässt sicher auch unter kodikologischen Gesichtspunkten kaum Fragen offen. Unklar ist die Zitierfähigkeit. Welche URL kann stabil zitiert werden, wenn die Adresse z.B. in eine Publikation übernommen werden soll? Die Bilder selbst sind als einzelne erreichbar, so dass der wissenschaftliche Kollegen in Übersee per Link auf eine bestimmte Seite gelangen kann.

2. Beispiel (Inkunabel): Aegidius Romanus, Theoremata de h[o]stia co[n]secrata, 1490 [Angebot: JALB Emden, Slg. Hardenberg]¹⁰

Der Titel ist als Digitalisat über den Verbund nachgewiesen (was nicht unbedingt zu erwarten ist, da Inkunabeln als besondere Materialien sich normalerweise nicht in OPACs finden bzw. vollständig verzeichnet werden können); das Link über den GBV funktioniert leider nicht. Wechselt man von dort in den OPAC (bequem über Link möglich) funktioniert der Link und man gelangt zum Inhaltsverzeichnis bzw. Digitalisat. Durch den Nachweis im GBV ist das Digitalisat auch über den KVK zugänglich. Im Portal INKA (Inkunabelportal deutscher Bibliotheken), wo auch zahlreiche Digitalisate nachgewiesen sind, ist es nicht zu finden, da die Emden Bestände nicht eingespielt wurden. Auch über Google ist das Digitalisat nicht greifbar bzw. fördert tote Links zutage. Ob Inkunabelforscher im KVK suchen werden, ist ungewiss.

Ein Zitieren einzelner Seiten ist möglich, zur Persistenz wird keine Aussage gemacht. Die Navigation ist gut handhabbar und übersichtlich, die Digitalisate von sehr guter Qualität. Kontextinformationen zur Sammlung Hardenberg werden über die Startseite der Bibliothek geboten.

3. Beispiel: Mader, Joachim Johann: Antiquitates Brvnsvicenses..., Helmestadi: Müller, 1678¹¹ [Angebot: ULB Münster, Bibliothek Nünning]

Ein zentrales Portal für den Nachweis digitalisierter Drucke fehlt bis heute in Deutschland. Auf das DFG geförderte Projekt "Portal Digitalisierter Drucke" (PDD), das diesem Mangel abhelfen soll, wird noch einzugehen sein.¹²

Im KVK wird der Titel nachgewiesen. Das wäre theoretisch ein guter Zugangsweg zur digitalen Edition. Unterstellt, der Nutzer klickt nicht auf die in anderen Verbänden

¹⁰ <http://www.jalb.de:8087/servlet/digbib-jalb?template=view.html&id=5186&startpage=5&endpage=186&image-path=http%3A%2F%2Fwww%2Ejalb%2Ede%2Fimage%2F38&image-subpath=38&pagenumber=5&zoom-factor=null&imageset-id=38> [2005-03-13]

¹¹ http://miami.uni-muenster.de/servlets/DocumentServlet?XSL.H_main=BROWSE&XSL.H_searchId=1110099447118&id=287 [2005-03-13]

¹² Theoretisch käme hierfür auch EROMM in Frage, das aber aus verschiedenen Gründen nicht systematisch genutzt wird. Zu EROMM vgl. unlängst Schwartz, Werner. "EROMM und alte Drucke". In: ZfBB 51 (2004) H.4. S.236-238.

nachgewiesenen Exemplare, sondern auf das Angebot des HBZ, dem die ULB Münster zugehört, wird er zum Verbundkatalog geführt, der ihm aber an dieser Stelle nicht weiterhilft. Auch ein Stöbern in der Digitalen Bibliothek NRW führt zu keinem Ergebnis. Wer allerdings zufällig im Dokumenten- und Publikationsserver der Humboldt Universität zu Berlin in der OAI-Suche recherchiert¹³, wird den Titel finden¹⁴. Immerhin zeigt sich hier ein sehr innovatives Konzept. Das Münsteraner System Miami tritt als OAI-Dataprovider auf und ist damit bestens für die Zukunft gerüstet. Dem heutigen Nutzer bringt dies allerdings noch wenig. Er wird sich eher über den Google-Treffer auf einer philologischen Seite in England freuen, die digitalisierte Titel gelistet hat und einen Link zu Mader bietet.

Die Darstellung ist übersichtlich, die Navigation intuitiv und unkompliziert; der Text der Drucke ist deutlich lesbar, Annotate oder andere exemplarspezifische Informationen sind hingegen wegen der bitonalen Digitalisierung leider nicht oder nur schlecht zu erkennen. Die Sammlung selbst wird im Kontext des Angebotes ausführlich beschrieben, der Nutzer weiß, was er hier bekommt. Zur persistenten Zitierfähigkeit werden unmittelbar keine Angaben gemacht, es gibt jedoch über die Hauptseite der digitalen Bibliothek Münster allgemeine Hinweise zum Zitieren von Internetquellen. Einzelne Seiten können per URL referenziert werden.

4. Beispiel: Luthers Ermanunge zum Friede, 1525. [Lutherhalle Wittenberg; Angebot erstellt durch HKI].¹⁵

Die Verbünde liefern zum digitalisierten Titel keine Informationen. Leider führt auch Google zu keinem Treffer. Man gelangt zum Digitalisat nur über die Datenbank des Instituts für Kulturhistorische Informationsverarbeitung bzw. via Verlinkung der Lutherhalle Wittenberg. Das doch recht umfangreiche und attraktive Angebot Wittenbergs ist also für den normalen Netznutzer bedauerlicherweise praktisch unsichtbar.

Die Navigation ist unproblematisch und leichtgängig, die Qualität der Digitalisate in den verschiedenen Auflösungen gut bis hervorragend. Negativ ist, dass sich keinerlei erläuternde Angaben zum Angebot finden. So kann man aus den insbesondere am HKI vorfindlichen, sehr spärlichen Informationen abnehmen, dass es wohl Drucke und Handschriften aus der Reformationszeit sind und dass es sich um 90.000 Digitalisate handelt. Welche und wie viele Titel es sind, ist nicht oder nur im ungefähren über die Indices zu ermitteln. Zur Zitierfähigkeit werden keine Angaben gemacht.

5. Beispiel Maercker, Georg: Nachlassbehandlung, Berlin 1894 [MPI Rechtsgeschichte Frankfurt]¹⁶

Das Digitalisat ist nicht im Verbund nachgewiesen, wird jedoch auf Anhieb in Google gefunden.

¹³ http://edoc.hu-berlin.de/e_suche/oai.php

¹⁴ Bedauerlich ist, dass der Berliner OAI Server nicht auf der zentralen OAI-Seite in <http://www.openarchives.org/service/listproviders.html> aufgeführt wird.

¹⁵ <http://luther.hki.uni-koeln.de/luther-cgi/kleioc/0010KILuther/exec/druckseite/%22Ag+4%7B-%7D+206+r%7B%7DTDI00040.jpg%22> [2005-03-13]

¹⁶ http://dlib-pr.mpiet.mpg.de/mfer-cgi/kleioc/0010MFER/exec/bigpage/%22184701_00000001%22 [2005-03-13]

Die Datenbank ist vom HKI entwickelt, für die Navigation gelten daher im Prinzip die obigen Angaben. Angaben zur Zitierweise oder Persistenz werden nicht gemacht. Die Kontextinformationen sind gut. Hervorzuheben ist, dass trotz der immensen Größe der Sammlung für alle Titel Strukturinformationen geboten werden. Die Bilder liegen als bitonale Scans vor, was aber angesichts der Materialien (19 Jh.) unproblematisch ist, insofern für dieses Zeitsegment exemplarspezifische und druckanalytische Fragestellungen eine eher untergeordnete Rolle spielen.

6. Beispiel: Nas, Johannes: Schaw Leser diese Bilder an/ Und dich darbey selbstem gemahn/ Waß massn Ecclesia militans, Heu clamet parturiens, Einblattdruck 1588 [Angebot: BSB München, DFG Projekt „Einblattdrucke der frühen Neuzeit“]¹⁷

Interessanterweise nicht im Sisis OPAC zu finden, obwohl eine Sisis-Nummer angegeben wird; im BibliotheksVerbund Bayern wird nur ein Exemplar für Augsburg nachgewiesen; ergo auch kein Weg zum Digitalisat der BSB; über Google stößt man über das im Projekt angebotene PDF-Dokument „Auswahlbibliographie zur Geschichte des Einblattdrucks in der frühen Neuzeit“ auf einen Hinweis, aus dem der Nutzer schließen kann, dass das Dokument digitalisiert wurde. Mit etwas Geduld kann er sich von dort zum Projekt durchsuchen. Einen direkten Link erhält er über Google nicht.

Der hier verwendete Zoom-Server der BSB ist leider ohne die Installation des Plugins nicht besonders komfortabel, im gezeigten Falle ist es schwierig, einen bestimmten Ausschnitt einzustellen. Nach Installation des Plugin ist die Nutzung jedoch gut, allerdings könnte der Gesamtausschnitt größer sein; leider ist die Nutzung des Plugins an den Internetexplorer gekoppelt. Für Leute wie mich, die die Verwendung des Internetexplorer mit ungeschütztem Netzverkehr gleichsetzen, ist das natürlich ein Problem. Die Qualität der Bilder ist gut und dürfte den allermeisten wissenschaftlichen Fragestellungen genügen. Kontextinformationen sind vorhanden, jedoch gelangt man dorthin nicht direkt vom Digitalisat. Angaben zur Zitierweise oder Persistenz des Angebots fehlen.

7. Beispiel Georg Fabricius, Poematum libri, Basel 1567 [Angebot: Mannheim/Heidelberg, Camena]¹⁸

In den Verbänden ist der Titel nicht zu finden, jedoch auf Anhieb in Google. Da das Camena-Projekt auch Volltexte bietet, kann man auch danach recherchieren. Sucht man z.B. vom selben Autor etwas aus den *carmina nuptialia* über Google, z.B. die Phrase *Phares, sponsae vero Rachelis* so findet man das anstandslos. Problematisch wird es allerdings, wenn man mit unspezifischeren Begriffen ohne Verfasserangabe sucht, z.B. *Ille uxorem rapuit* (rund 400 Treffer). Die Volltextsuche in Camena ist leider etwas gewöhnungsbedürftig und wegen mangelnder Übersichtlichkeit der Ergebnismengen nicht optimal.

Das Angebot kann vom Umfang des zum jeweiligen Druck Gebotenen her als mustergültig gelten, wenigstens in den Fällen, in denen alle konzipierten Aspekte realisiert wurden, das sind digitale Images, ASCII-Volltext und getaggtter, d.h. strukturierter Volltext in XML. Das läßt in puncto Metadaten kaum Wünsche offen. Sammlungs- bzw. Kontextinformationen sind

¹⁷ http://zoom.bib-bvb.de/StyleServer/calcrn?cat=einbl&item=/300000026_0_r.sid&wid=750&hei=500&style=bsb/einbl.xsl&plugin=false [2005-03-13]

¹⁸ <http://www.uni-mannheim.de/mateo/camena/fabr2/te01.html> [2005-03-13]

vorhanden bzw. teilweise etwa im Bereich *thesaurus eruditionis*¹⁹ vorbildlich. Die Images sind von guter bis sehr guter Qualität und als einzelne referenzierbar. Aussagen zur Persistenz oder Zitierweise des Angebots fehlen jedoch. Bei Camena verkompliziert sich das Zitierproblem dadurch, dass man als Nutzer möglicherweise auch den Volltext zitieren möchte. Dafür fehlen nach meinem Wissen jedoch insgesamt brauchbare Lösungen. Möglicherweise werden es in Zukunft Adressierungstechniken mit XPath erlauben, gezielt einzelne Passagen in XML-Dateien anzusteuern.

8 Beispiel: Joseph Addison, Remarks on several parts of Italy, [et]c. 1705 [GDZ Göttingen, Sammlung Reiseliteratur]²⁰

Das Digitalisat wird über den KVK gefunden, da im GBV verzeichnet. Die Situation in Google ist unübersichtlich, da eine Recherche nur mit Titelstichwörtern zu über 26.000 Treffern führt. Zusammen mit dem Erscheinungsjahr und dem Stichwort "digital" findet man den Titel verlinkt auf einer Seite namens "Digital Book Index".²¹

Die Navigation ist übersichtlich und leicht zu handhaben. Das Ansteuern einzelner Images ist unproblematisch. Es wird eine PURL (Persistent URL) angegeben, so dass verlässliches Zitieren möglich ist. Die Bildqualität reicht für das Lesen aus. Die Bitonalität der Bilder erlaubt jedoch keine weitergehende Analyse gerade der älteren Drucke. Da der Schwerpunkt der Göttinger digitalen Sammlungen eher Mitte des 18. bis 19. Jh. liegt, ist dieser Mangel möglicherweise nicht so schwerwiegend, da in dieser Zeit die druckanalytischen bzw. exemplarbezogenen Fragestellungen gegenüber der Textbereitstellung zurücktreten. Die gebotenen Kontextinformationen sind leider etwas versteckt. Von der Hauptseite startet man direkt vom Link "Itineraria" in die Datenbank. Den Artikel zur Sammlung von Eck findet man erst nach etwas Suchen unter *Projects*.

9. Beispiel: De Bry, Johann Theodor, Proscenium vitae sive Emblematum Saecularium, 1627 [HAB Wolfenbüttel, Wolfenbütteler Digitale Bibliothek- Emblematica Online]²²

Das Angebot ist im Verbund und so KVK nachgewiesen und wird, dank eingestellter Titellisten, auch über Google gefunden. Einzelne Seiten sind persistent adressierbar, das Zitierverfahren wird erläutert. Problematisch ist das Auffinden ggf. vorhandener Strukturinformation, diese werden, obwohl wir sie auch als flache Informationen ins Netz gestellt haben, derzeit zumeist nur über lokale Datenbanken gefunden. Damit Google etwas indexiert, ist es nötig diese Informationen zu verlinken (vgl. <http://diglib.hab.de/drucke/gm-4f-411/intro.htm>). Von den 300 ins Netz gestellten Dateien sind bisher nur 3 über Google erreichbar. Man wird also um weitere Verlinkungen nicht herumkommen, wenn auch Strukturinformationen gefunden werden sollen. Kontextinformationen erhält man über das Link "Digitale Bibliothek".

Da ich die Navigation selbst eingerichtet habe, ich es jetzt müßig, sie gut zu finden. Kritik und Verbesserungsvorschläge nehme ich gern entgegen. Die Qualität der überwiegenden Anzahl der Digitalisate halte ich für sehr gut. Allerdings gibt es auch noch unschöne Leichen an der HAB.²³ Diese Scans vom Microfiche, in einer Art *good will* –Aktion mit ABM- Kräften

¹⁹ <http://www.uni-mannheim.de/mateo/camenahtdocs/camenaref.html> [2005-03-13]

²⁰ <http://dz-srv1.sub.uni-goettingen.de/sub/digbib/loader?ht=VIEW&did=D164743&p=2> [2005-03-13]

²¹ http://www.digitalbookindex.com/_search/search010litenglish18thca.asp [2005-03-13]

²² <http://diglib.hab.de/drucke/xb-6550/start.htm> [2005-03-13]

²³ <http://diglib.hab.de/drucke/227-1-quod/start.htm?image=00000006> [2005-03-13]

hergestellt, sind in etwa 5% der Fälle praktisch unleserlich. Da wir aber eine grundsätzliche Erklärung zur Persistenz und Zitiersicherheit auf Seitenebene abgegeben haben, müssen die Seiten jetzt Schritt für Schritt ersetzt und können nicht gelöscht werden, - so peinlich es ist. Hier zeigt sich, dass auch *good will* Aktionen organisatorisch gut überlegt sein wollen, möchte man sich späteren Ärger ersparen.

Resümee

Aus den hier kurz vorgestellten digitalen Angeboten lässt sich folgendes Bild gewinnen. Als sehr problematisch muss die bibliothekarische Seite der Erreichbarkeit der Digitalisate angesehen werden. In Verbundkatalogen findet sich nur das Wenigste, selbst bei Bibliotheken, bei denen man davon ausgehen könnte, dass sie ihre digitalisierten Titel über Verbünde nachweisen, ist dies nicht selbstverständlich. Des Weiteren sind nur vereinzelt die allgemeinen Forderungen nach OAI-Schnittstellen auch bei Altbestandsdigitalisierung umgesetzt worden. - - mir ist nur der Münsteraner Fall bekannt, und in diesem Fall vermutlich nur, weil der Altbestand gleichsam "zufällig" miterfasst wurde. Eine übergreifende Erschließung von Strukturdaten oder Volltexten jenseits von Google gibt es bisher nicht.

Als einzige Konstante kristallisiert sich heraus, dass die meisten Titel, Strukturen und Volltexte über Google gefunden werden. Dafür gelten aber einige Einschränkungen. Die Suchbegriffe müssen spezifisch genug sein, um zu einer sinnvollen Eingrenzung der Treffermenge zu kommen. Komplexere Suchwünsche können nur bedingt befriedigt werden. Nicht alle Informationen werden indexiert, die Kontrolle über die Indexierung liegt nicht beim Indexer. So vorderhand nützlich und im Moment auch unverzichtbar Google & Co sein mögen, so unbefriedigend stellt sich die Situation für die Bibliotheken und mittelbar auch für die Nutzer dar, die sich letztlich nicht sicher sein können, ob das Gesuchte nicht doch irgendwo vorhanden ist. Vor dem Hintergrund dieses Befundes nimmt es nicht wunder, dass in der von Manfred Thaller betreuten sog. Retroeval-Studie ein erhebliches Informationsdefizit bei den Benutzern festgestellt wurde. Viele waren überrascht, wie viel schon digitalisiert ist, die wenigsten wussten, wo die für sie relevanten Altbestandsquellen zu finden waren.²⁴

Wenn auch die Situation in Sachen Bekanntheitsgrad und Findbarkeit düster ist, so lässt sich ein anderer hier festgestellter Trend durch die Retroeval-Studie positiv bestätigen. Diejenigen Nutzer, die auf ein digitales Angebot stießen, zeigten sich ob der Angebote angenehm überrascht und waren zum überwiegenden Teil mit der Qualität der Images, den Navigationsmöglichkeiten, der Handhabbarkeit zufrieden und hatten auch mit der Orientierung im Angebot kaum Probleme.²⁵

Ein sicher noch weiter zu untersuchender Aspekt ist die Frage der gelieferten Hintergrundinformationen und Hilfsmittel. Leider bieten nicht alle Projekte auch Texte zu den jeweiligen digitalen Sammlungen und machen es dem Benutzer unnötig schwer, die Art und Zusammensetzung der vorgefundenen digitalen Sammlung einzuschätzen. Neben der Titelbeschreibung sollte bei jeder Digitalisierungskampagne als Mindeststandard ein kurzer Text über das, was man inhaltlich und quantitativ erwarten kann, dazu gehören. Hier mögen sich in Zukunft die Aktivitäten zur Verbesserung der *collection description* positiv auswirken; schön wäre eine Art Handbuch für digitalisierte historische Buchbestände.

²⁴ Vgl. Retroeval

²⁵ Vgl. Retroeval-Studie, S. 47 und 2.3 Umfrage II (qualitative Studie), S. 77.ff.

Die Zitierfähigkeit der Quellen ist leider in den meisten Fällen ungewiss. Nur zwei der genannten Angebote bieten dem Nutzer beim Digitalisat explizite Hinweise, dass das, was er vorfindet, auch persistent mit sog. PURLs zitiert werden kann. Kein Angebot verwendet, zumindest nach den vorgefundenen Metadaten, URNs oder DOIs für den Altbestand. Hier besteht erheblicher Verbesserungsbedarf.

Perspektiven

Alle bisherigen Evaluierungen und Untersuchungen – und davon gibt es nicht viele²⁶ – deuten darauf hin, dass es ein vordringliches Ziel sein muss, die disparat vorliegenden Projekte zusammenzuführen, sonst bleibt Suchen und Finden eine Sache des Zufalls. Sinnvollerweise geschieht dies in materialspezifischen Portalen, die den traditionellen formalen Erschließungskonzepten Rechnung tragen. Die DFG fördert derzeit drei solcher zentraler Portale für den Nachweis von Altbestand. Das sind die Handschriftendatenbank *Manuscripta mediaevalia* insbesondere für mittelalterliche Handschriften, *Kalliope*²⁷ für Autographen und Nachlassmaterialien und das *Portal Digitalisierter Drucke (PDD)*. Während die ersteren beiden auf eine schon längere Entwicklung zurückblicken und noch nicht mit der klaren Perspektive, auch als zentrale Portale für den Zugriff auf Digitalisate zu fungieren, konzipiert wurden, startet letzteres am 1.4.05 als Gemeinschaftsprojekt der Verbundzentrale Göttingen, des Hochschulbibliothekszentrums in Köln und der Sammlung Deutscher Drucke und setzt sich explizit zum Ziel, alle einschlägigen digitalen Materialien insbesondere aus DFG Projekten zusammenzufassen. Dabei sollen sowohl Titel, Struktur- und Volltextdaten berücksichtigt werden. Ein weiteres Portal war immer wieder im Gespräch, ist aber bis heute nicht realisiert worden, nämlich ein zentrales Inkunabelportal. Möglicherweise wirkt hier einfach die Kraft des Faktischen, da INKA auf der Basis von deutschen Census bzw. ISTC-Daten diesen zentrale Zugriff auf viele digitalisierte Exemplare bereits bietet (die Wolfenbütteler Digitalisate aus dem vdlb Projekt²⁸ sind dort schon nachgewiesen, die aus Köln werden folgen). Ggf. wird auch der ISTC in London selbst – sofern die Webversion einmal in Betrieb geht – ebenfalls diese Funktion übernehmen.

Um den Datenaustausch und die Datenlieferung an die Portale sicherzustellen, ist die Festlegung von Formaten auf der Basis von XML unverzichtbar. Hier gibt es eine Reihe von Entwicklungen und es zeichnen sich auch konsolidierte Formate ab. Im Bereich Handschriften ist es das MASTER-Format, das im Zuge einer Überarbeitung durch die *Text Encoding Initiative* in Kürze in der Version P5 der TEI aufgehen wird. Damit wäre angesichts der Bedeutung der TEI auch ein internationaler Standard de facto festgeschrieben. Derzeit führen München und Wolfenbüttel in Verbindung mit der Handschriftendatenbank ein Teilprojekt durch, das die Adaption der DFG-Richtlinien zur Handschriftenbeschreibung auf das Master/TEI-P5 Format zum Ziel hat.²⁹ Mittels eines XML Schemas wird sodann die Validität von Beschreibungen ma. Handschriften überprüfbar sein, mit der Entwicklung von XSLT-Skripten können Masterdaten in lokalen Kontexten nachgenutzt, PDF-Versionen für Kataloge generiert und vor allem Daten über eine definierte Schnittstelle an die Handschriftendatenbank für den zentralen Nachweis geliefert werden. Ein entsprechendes Format gibt es auch schon für Kalliope. Es setzt auf der EAD (Encoded Archival Description)-DTD auf und bezieht die RNA-Richtlinien ein. Für den Bereich PDD sollen

²⁶ Vgl. Retroeval, Einleitung

²⁷ <http://kalliope.staatsbibliothek-berlin.de>

²⁸ <http://inkunabeln.ub.uni-koeln.de>

²⁹ <http://www.hab.de/forschung/projekte/master.htm>

solche Formate im Rahmen des Projektes noch entwickelt werden. Am aussichtsreichsten scheint derzeit METS³⁰ für alle Kategorien von Metadaten.

Zum Transport aller dieser Daten – auch dies ist mittlerweile eine feste Größe – wird OAI empfohlen. Im PDD Projekt werden alle SDD Bibliotheken, die noch keine OAI Schnittstelle haben, eine solche nachrüsten. OAI, dies sei hier noch einmal betont, fungiert dabei als Transportprotokoll. Dass zu OAI auch Mindestmetadaten nach DC gehören, spielt für die wenigsten Applikationen eine Rolle. DC lässt sich als ein reduziertes Subset aus vorhandenen Metadaten, die insgesamt mehr Informationen als DC bieten, relativ leicht automatisch extrahieren. Eine Primärerfassung von Daten in DC ist in Altbestandskontexten jedoch nicht zu empfehlen.³¹ Weitere, nicht über OAI zu beziehende Daten ließen sich auch per Harvester nach Google-Vorbild einsammeln.

In der idealen digitalen Welt böten so die genannten Portale einen zentralen Zugriff auf alle Digitalisate von Quellenmaterial. Definierte Schnittstellen und Formate sorgten für einen reibungslosen Verkehr zu diesen Portalen, die Zweigleisigkeit von ggf. spezialisiertem lokalem und überregionalem allgemeinem Angebot wäre gewährleistet. Wichtig ist, die Portale vor allem in der Wissenschaft zu bewerben, um dem in der Retroeval-Studie festgestellten eklatanten Informationsdefizit abzuhelpfen. Mit anderen Worten, die Portale müssten mindestens so bekannt und effizient sein wie der KVK. So bestünden gute oder doch zumindest bessere Aussichten, dass alles was im Netz ist – und das ist schon sehr viel - , auch gefunden wird und die Basis gelegt wird für große, national abgestimmte Digitalisierungskampagnen. Diese, das sei unterstrichen, machen nämlich nur Sinn, wenn auch Nachweis und Bekanntheit mit dem Wachstum der digitalen Sammlungen Schritt halten. Über die Auswahl des künftig zu Digitalisierenden und die zu verfolgende Strategie kann man natürlich füglich streiten. *Tot professores, tot opiniones*: von der Forschung wird es nach meinem Eindruck hierüber entweder keine oder, was dasselbe ist, heterogene oder –doxe Auskunft geben. Man könnte hier versuchen, wenn dies möglich ist, über Statistiken der Lesesaalnutzung bestimmte Bände herauszufiltern, die besonders intensiv genutzt werden oder aber bestimmte Gattungsgruppen, hinter denen man eine Forschungscommunity weiß, digitalisieren. Ob Kampagnen a la Google auch im deutschsprachigen Bereich durchführbar und sinnvoll sind, müsste noch geprüft werden. Grundsätzlich ist das natürlich eine schöne Idee. Bei der Umsetzung beneide ich die "betroffenen" Einrichtungen allerdings nicht. Ein wichtiger Schritt für den weiteren Aufbau, auch dies eine Empfehlung aus der Retroeval-Studie, dürfte sein, dass nicht an wenigen zentralen Stellen Digitalisierungsressourcen aufgebaut werden, sondern die "Digitalisierung von bibliothekarischen Ressourcen oder von Objekten des kulturellen Erbes mittelfristig zum Standardleistungsspektrum aller oder zumindest der meisten wissenschaftlichen Bibliotheken zählen".³²

³⁰ <http://www.loc.gov/standards/mets>

³¹ Vgl Diane L. Hillmann, in: "Kenny, Anne R./Rieger Oya Y. (Hrsgg.): Moving Theory into Practice. Digital Imaging for Libraries and Archives, Mountain View: Research Library Group, 2000, S. 89f

³² Retroeval, S.12