Digitalisierung in der Bayerischen Staatsbibliothek in München

V.L.N.R.: WOLFGANG MARIA WEBER; BSB; STAATS- UND UNIVERSITÄTSBIBLIOTHEK, DRESDEN; BSB

BIBLIOTHEKEN

Babylonischer Bau

Die Deutsche Digitale Bibliothek will Millionen Bücher, Filme, Bilder und Tonaufnahmen im Internet zugänglich machen. Über 30 000 Bibliotheken, Museen und Archive sollen ihr digitalisiertes Kulturgut beisteuern. Kann ein derart ehrgeiziger Plan überhaupt gelingen?

Wenn alles gutgeht, schafft dieser Leser bis zu 1216 Seiten in der Stunde. Leise zischend liest er Buch um Buch weg. Ab und zu macht er „pffft“.

Ein Roboter neuester Bauart ist hier im Einsatz. Jedes Buch, das man ihm aufgeschlagen hinlegt, liest er automatisch ein: Ein schmaler Keil senkt sich hinab in den Falz, saugt dort links und rechts je ein Blatt an und zieht die Beute senkrecht empor. Dabei wird sie abfotografiert. Mit einem sachten Luftstoß – pffft – blättert der Roboter sodann um.

So geht es tagein, tagaus im Münchner Digitalisierungszentrum der Bayerischen Staatsbibliothek. 45 000 Werke sind bereits elektronisch erfasst – vom „Nibelungenlied“ auf Pergament bis hin zur Originalpartitur aus der Hand Gustav Mahlers.

Die Schätze aus der Frühzeit der Buchkultur werden freilich meist noch von Hand abgelichtet. Der Roboter kapituliert angesichts brüchiger Folianten, die mit ihren Ledereinbänden oder Holzdeckeln nicht selten einen halben Zentner wiegen. Zu den schönsten Stücken zählt ein prachtvolles Familienbuch der Kaufleute Fugger aus dem 16. Jahrhundert.

Vom Reichtum der Münchner Datenspeicher soll künftig ein neues Portal im Internet profitieren: die Deutsche Digitale Bibliothek (DDB). Geplant ist eine Online-Zentrale für Millionen Bücher, Zeitschriften, Fotos und Filme. Aus dem ganzen Land, heißt es, werden Bibliotheken, Museen und Archive ihre digitalisierten Kulturgüter einspeisen.

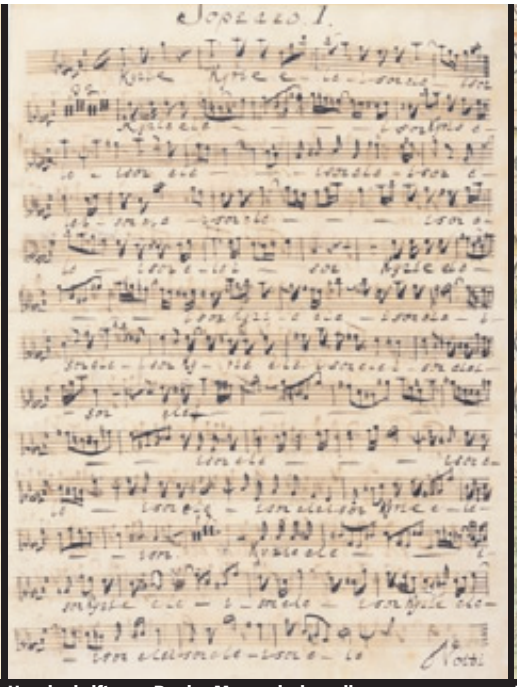
Es wird allerdings noch eine Weile dauern. Erst 2011 dürfte eine erste Probeversion ans Netz gehen – „und auch die nur für einen begrenzten Nutzerkreis“, sagt Ute Schwens, Direktorin der Deutschen Nationalbibliothek in Frankfurt, die den Aufbau der DDB koordiniert.

Umso ehrgeiziger ist die langfristige Vision. Von einem „Jahrhundertprojekt“ schwärmte Kulturstaatsminister Bernd Neumann (CDU). Eine virtuelle Wunderkammer verheißen die Initiatoren, geeignet sowohl für Laien als auch für Forscher, die gezielt nach Quellen und wissenschaftlichen Dokumenten fahnden.

Wer „Beethoven“ eingibt, findet dann nicht nur Bücher über den Komponisten, sondern auch handschriftliche Notenblätter, Musikbeispiele zum Anhören oder vielleicht eine Filmaufnahme des „Fidelio“.

Anfang Dezember gab das Bundeskabinett grünes Licht. Ziel ist es, die DDB der-

Familienbuch der Fugger
Bayerische Staatsbibliothek



Handschrift von Bachs Messe in h-moll
Sächsische Landesbibliothek Dresden

Historische Landkarte Bayerns (Ausschnitt)
Bayerische Staatsbibliothek

Zahlreiche Bibliotheken und Museen sind seit Jahren damit beschäftigt, ihre Sammlungen zu digitalisieren (siehe Beispiele). Diese Vorarbeiten bilden den Grundstock der geplanten Deutschen Digitalen Bibliothek. Sie ist gedacht als Portal, das im Internet die digitalisierten Kulturgüter der angeschlossenen Häuser unter einem gemeinsamen Dach präsentiert. Die Exponate selbst bleiben auf den Computern der Institute, die zumeist ihre eigenen Internetangebote damit betreiben. Über die Website der DDB sind alle Werke erstmals zentral aufzurufen; dafür werden sie zuvor einheitlich verschlagwortet.

einst einzugliedern in das europäische Portal Europeana, das ein Jahr zuvor mit vergleichbaren Ambitionen gestartet ist.

Die europäische Emsigkeit ist vor allem dem Suchmaschinenriesen Google geschuldet, der sich weltweit bereits mehr als zehn Millionen Bücher digital einverleibt hat. Vom drohenden Kulturmonopol eines Privatkonzerns war da die Rede.

Europeana und DDB versprechen nun, strikt die Urheberrechte zu achten, die Google im Eifer des Einlesens bislang eher widerstrebend zur Kenntnis nimmt. Jean-Noël Jeanneney, ehemals Präsident der Französischen Nationalbibliothek, sprach von einem „antikapitalistischen Gegenmodell zum Ermächtigungsakt Googles“.

Zugleich ist die Digitalisierung auch eine Maßnahme gegen die Hinfälligkeit des Mediums Buch. Als im Jahr 2004 die Weimarer Anna Amalia Bibliothek brannte, wurden 50 000 teils unersetzliche Bände zerstört. Digitale Sicherheitskopien könnten solche Verluste künftig begrenzen.

Doch während die Fachleute in Deutschland nun mit den Vorarbeiten beginnen, zeichnet sich bereits ab, dass das Unternehmen vor furchterregenden Herausforderungen steht. Höchst ambitioniert erscheinen allein schon die technischen Ziele: Das Fraunhofer-Institut IAIS in Sankt Augustin bei Bonn, das für die Computertechnik der DDB zuständig ist, arbeitet an Programmen, die in Filmen die Personen erkennen, Sprachaufnahmen in durchsuchbaren Text verwandeln und Dokumente automatisch verschlagworten sollen.

Vollends verwegen erscheint die angestrebte Reichweite des neuen Portals: Über 30 000 Museen, Archive und wissenschaftliche Sammlungen in ganz Deutschland sollen sich anschließen. Bisher können die DDB-Macher froh sein, wenn auch nur hundert Häuser mitmachen. Selbst renommierte Institutionen wie die Hamburger Kunsthalle oder das Frankfurter Städel-Museum sind vorerst nicht darunter.

Rolf Griebel, der Generaldirektor der Bayerischen Staatsbibliothek, der das Projekt eigentlich „für gut und überfällig“ hält, warnt deshalb vor allzu ehrgeizigen Plänen: „Ich habe erhebliche Bedenken, ob man die DDB anständig und in vernünftiger Zeit mit Inhalt befüllen kann.“

Für das Einlesen eines Buchs aus dem 16. oder 17. Jahrhundert rechnet Griebel, je nach Aufwand, mit Kosten von 70 bis 140 Euro. Zeitgenössische Titel sind billiger, doch geht es hier um gewaltige Mengen. Der Deutsche Bibliotheksverband schlägt vor, zunächst in zehn Jahren eine Auswahl von etwa fünfeinhalb Millionen Bänden zu digitalisieren. Das Unterfangen würde mindestens 165 Millionen Euro kosten. Doch woher sollte ein solche Summe kommen?

Neidvoll blicken die Deutschen nach Frankreich, wo Staatspräsident Nicolas Sarkozy kürzlich versprach, 750 Millionen Euro für die Digitalisierung der französischen Nationalkultur aufzubringen.

Das deutsche Angebot dagegen dürfte auf Jahre hin vor allem aus Lücken bestehen. Wird aber der Nutzer damit zufrieden sein, dass er im Glücksfall schöne Funde macht, ansonsten aber immer wieder ins Leere läuft? Sollte man nicht lieber weniger machen, das aber richtig gut? „Für den Anfang wäre sicher eine Beschränkung auf ausgewählte Schwerpunkte und Themen sinnvoll“, meint Bibliotheksleiter Griebel.

Damit aber wollen sich die DDB-Planer nicht begnügen: Alle Kultursparten, alle Wissenschaften, alle Arten von Dokumenten sollen erfasst werden, und das am besten von allen Museen und Bibliotheken.

Und die Suchtechnik soll um einiges raffinierter ausfallen als das einfache Aufstöbern von Suchwörtern, wie Google es anbietet. Die Bestände der DDB, so der Plan, lassen sich gezielt nach vielerlei Kriterien erschließen – etwa Ort, Zeit oder Sachgebiet. Das geht aber nur, wenn die Objekte auch detailliert beschrieben sind.

Dafür steuert das vom Bund geförderte Theseus-Programm hilfreiche Grundlagentechnik bei. Seit 2007 arbeiten die Forscher dort an Verfahren, wie sich Bilder, Filme, Tonaufnahmen und Bücher maschinell erfassen lassen. Wenn der Computer ansatzweise versteht, worum es dabei geht, kann er schon mal etliche Datenfelder automatisch ausfüllen – unabdingbar

bei den Massen von Dokumenten, mit denen die DDB fertig werden muss.

Erste Fortschritte melden die Forscher etwa beim Erkennen von Objekten in Filmen und Fotografien. „Gesichter sind noch schwierig, aber mit Bäumen, Autos oder Gebäuden geht das schon ganz gut“, sagt Thomas Niessen, Leiter des Theseus-Programms. Auch verwandelt der Computer mit einigem Erfolg gesprochene Rede in durchsuchbaren Text – außerdem versucht er, die erwähnten Personen, Orte und Ereignisse herauszufiltern.

Zugleich laufen bereits die Debatten ums große Ganze: Wie kann es gelingen, die DDB für Laien und Forscher zugleich attraktiv zu machen? Und was gehört zu einem idealen Portal alles dazu?

Noch darf geträumt werden. Reinhard Altenhöner von der Deutschen Nationalbibliothek etwa hält es für denkbar, dass die Nutzer selbst Beiträge einstellen.

JENS MEYER / AP

Anna Amalia Bibliothek in Weimar*: *Hinfälligkeit des Mediums Buch*

„Wenn ein Stadtarchiv Material über die Geschichte einer Straße anbietet“, sagt er, „könnten die Bewohner das mit ihren eigenen Geschichten und Fotos anreichern.“

Museen wiederum dürfen in den Suchergebnissen Links zu passenden Ausstellungen platzieren, die bei ihnen gerade laufen. Eine kleine Demonstration am Bildschirm zeigt, wie das funktionieren könnte. „Und hier“, sagt Altenhöner und klickt auf einen weiteren Link, „hier können Sie auch gleich das Bahnticket kaufen.“

Solch ausgetüftelte Extras sucht man bei Google vergebens. Der Suchmaschinenkonzern macht lieber Projekte, die in einem Satz zu erklären sind. Im Fall der Digitalisierung lautet das Ziel schlicht: alle Bücher der Welt, bedienungsfreundlich präsentiert. Andernfalls nützt die beste Erschließungstechnik wenig.

Wohin es führt, wenn dieser Grundsatz missachtet wird, zeigt das Dachportal Europeana. Die Sammlung soll nun, nach längerer Stagnation, bis Mitte des Jahres anwachsen auf zehn Millionen Kulturgüter. „Damit sind wir weltweit führend“, sagt der Berliner Informationswissenschaftler Stefan Gradmann, der dem Leitungsgremium der Europeana angehört.

Ein Teil der Exponate ist probeweise über eine intelligente Suche erschließbar: Wer etwa „Paris“ eintippt, dem liefert die Europeana auch den Montmartre und die Tuilerien; gesondert erscheinen Quellen, die sich mit dem Königssohn Paris aus der griechischen Mythologie befassen. Dessen verhängnisvolle Tat, der „Raub der Helena“, ist der Suchmaschine ebenfalls bekannt: Sie findet auch Dokumente, in denen das Suchwort gar nicht vorkommt. Doch macht es wenig Vergnügen, in der Europeana zu stöbern. Die Funde erscheinen nur mit briefmarkengroßen Vorschaubildchen. Und wer sich durchklickt, um einen Überblick zu gewinnen, wird von jedem Bild zum jeweiligen Institut weitergeschaltet. Bald irrt der Suchende ratlos herum auf einem Dutzend verschiedener Websites von Museen und Bibliotheken – und geht am Ende verloren irgendwo zwischen der „Vlaamse Kunstcollectie“ und der „Wielkopolska Biblioteka Cyfrowa“.

Ist es da nicht besser, die Exponate allesamt in den vertrauten Rahmen der Europeana einzubinden? „So hätten wir es gern“, sagt Gradmann. „Aber da machen die Museen nicht mit.“ Die nämlich bestehen darauf, sich selbst mit ihren Schätzen zu präsentieren.

Sollte die DDB ebenfalls der Eitelkeit der beteiligten Institute nachgeben, droht ein babylonischer Bau mit 30 000 angeschlossenen Häusern. Wer hätte die Ausdauer, sich mit deren 30 000 höchsteigenen Web-Angeboten herumzuschlagen?

Auch das Versprechen, das Urheberrecht vorbildlich zu achten, wirft Probleme auf. Unbeschwert einlesen darf die DDB nur Werke von Autoren, die mindestens 70 Jahre tot sind. Für jüngere Dokumente, deren Urheber nicht erreichbar sind, wird eine Regelung mit den zuständigen Verwertungsgesellschaften erarbeitet.

DDB-Koordinatorin Schwens will aber auch die Geistesproduktion der Gegenwart einbinden. „Es wäre schade“, sagt sie, „wenn etwa die aktuelle Wissenschaft über die DDB nicht auffindbar wäre.“ Verhandlungen mit den Verlagen sind schon im Gange. Ideal wäre, meint Schwens, ein „One Stop Shop“, wo der Nutzer das Werk, das ihn interessiert, auch gleich elektronisch ausleihen oder kaufen kann.

Für die Verkaufsgeschäfte würde sich der Internetladen Libreka anbieten, den der Börsenverein des Deutschen Buchhandels betreibt. Libreka steht allerdings in zweifelhaftem Ruf: Viele elektronische Bücher sind dort mit einem umständlichen Kopierschutz versehen. Die Verlage wollen es so. Und ihre gefragtesten Bücher geben sie aus Angst vor Raubkopierern erst gar nicht ins Sortiment.

So droht sich die Digitalisierung zwischen zwei Mühlsteinen aufzureiben: Für das Ablichten der alten Werke gibt es zu wenig Geld, und den Zugang zu den neuen, die zumeist schon digital vorlägen, verweigern aus Ängstlichkeit die Verlage.

Sollte also doch besser Google die Sache übernehmen? Mitte des Jahres will die Firma in den Handel mit elektronischen Büchern einsteigen. Gleich eine halbe Million Titel sind für das Projekt „Google Editions“ vorgesehen. Von jedem Verkauf gehen 63 Prozent an den jeweiligen Verlag, den Rest behält Google ein.

Die Bayerische Staatsbibliothek jedenfalls hat bislang gute Erfahrungen mit Google gemacht. Seit 2007 digitalisiert der Suchmaschinenkonzern den Münchner Kulturhütern kostenlos die gemeinfreien Bücher aus dem 18. und 19. Jahrhundert – immerhin rund eine Million Bände. In harten Verhandlungen sicherte sich die Bibliothek das Recht auf eine eigene Kopie eines jeden Buchs; sie kann damit machen, was sie will. So bleibt der öffentliche Zugang zu den Schätzen gewahrt.

Und das Einlesen läuft wie am Schnürchen. Jede Woche verlassen derzeit rund 5000 Bände die Hallen der Staatsbibliothek. Ein Lastwagen schafft sie an einen streng geheimen Ort in Bayern, wo die Scanner von Google unentwegt am Werk sind. Bei dem Tempo wird in kaum vier Jahren schon alles erledigt sein.

Manfred Dworschak

* Nach der Wiedereröffnung des bei einem Brand am 2. September 2004 zerstörten Rokokosaals.