

sind nach den bisherigen Erfahrungen in nur ca. 40 Prozent aller Fälle neu anzusetzen. Die restlichen sind entweder schon als Stammsatz vorhanden, kommen in verschiedenen Sätzen mehrmals vor oder sind falsche Eintragungen, die gelöscht werden können. Auch zum Anlegen eines neuen Körperschafts-Stammsatzes muss grundsätzlich das Buch vom Fach geholt werden.

Vom September 2001 bis Mitte März 2003 wurden auf diese Weise insgesamt ca. 23.000 Datensätze vollständig bearbeitet. Übrig sind noch ca. 63.000 Mailboxen, deren Bearbeitung in ungefähr anderthalb Jahren abgeschlossen sein dürfte. Ein großes

Problem stellt der von der Firma nicht mehr vollständig bearbeitete Systematische Katalog dar. Hier sind nicht nur die Rohdaten, sondern eigentlich alle Aufnahmen unter Kontrolle der jeweiligen Katalogkarten zu überarbeiten. Es wurden dafür erst kürzlich Sondermittel beantragt – ob diese gewährt werden, ist aber noch ungewiss.

Erschwert wird die Arbeit im Moment auch noch durch die im Hause stattfindenden Bauarbeiten. Ein Großteil des Bestandes wurde im März in die Galleria Nazionale d'Arte Moderna ausgelagert. Die Bearbeitung dieser Fälle muss daher bis zur Aufnahme eines regulären Betriebes dort verschoben werden.

RETROKONVERSION DER ALPHABETISCHEN KATALOGE DES KUNSTHISTORISCHEN INSTITUTS IN FLORENZ

Ruth Werner – (Kunsthistorisches Institut in Florenz, Max-Planck-Institut, Bibliothek, Florenz)

Das Kunsthistorische Institut in Florenz, 1897 gegründet, verfügte über die längste Zeit seines Bestehens nicht über einen nach heutigen Kriterien konzipierten, vollständigen alphabetischen Katalog. Der gesamte Bibliotheksbestand war in Freihandaufstellung zugänglich, die Aufstellung so einleuchtend gestaltet, dass die Suche am Standort in den meisten Fällen ausreichende Ergebnisse brachte. Noch der Vorgänger des jetzigen Bibliotheksleiters war mit guten Gründen davon überzeugt, dass 90 % der Suchen auf Künstler oder topographische Einheiten gerichtet seien, und dass die Sachgruppen-Aufstellung diesen Suchanfragen vollauf genüge.

Erfassung und Erschließung der Florentiner Bestände bis 1996

Das wichtigste Instrument zur Erschließung der Bestände war der Standort-Katalog. Es gab jedoch von Anfang an auch einen alphabetischen Katalog, der handschriftlich, mit Tinte und Feder, bis zum Anfang der achtziger Jahre des zwanzigsten Jahrhunderts fortgeführt wurde. Die Erfassungsregeln wurden nicht schriftlich festgelegt und haben sich im Laufe von über achtzig Jahren Katalogisierungszeit beträchtlich geändert. Das Anwachsen des Bestandes (heute mehr als 250.000 Bände), Besonderheiten in der Aufstellung (drei Formate, die sich

nicht in der Signatur ausdrücken und an drei verschiedenen Orten aufgestellt sind) sowie neue, interdisziplinäre Fragestellungen in der Kunstgeschichte führten aber seit den sechziger Jahren dazu, dass man, weiterhin mündlich, detailliertere Hausregeln festlegte, die mechanische Wortfolge und Staatsbürgerprinzip bei der Namensansetzung (nicht wie in der Bibliotheca Hertziana, als Hommage an das Gastland Italien, „VonGoethe“), aber keine Eintragungen unter Serien oder Körperschaften vorsahen.

Einstieg für Bestandskataloge von Museen und Ausstellungskataloge war der Ort der Einrichtung, dann ggf. ein ungefähres Alphabet der Museumsnamen für die Bestandskataloge und eine chronologische Folge der Ausstellungen. Bei für die Institutzwecke wichtigen Orten wie etwa Rom, Florenz, London gibt es weitere Unterteilungen für Ausstellungen moderner Kunst (moderner Künstler ist im KHI, wer nach 1850 geboren ist) und Auktionskataloge. Vermieden wurde nach Möglichkeit eine Haupteintragung unter dem Sachtitel (unter dem man die Ausstellungskataloge also nicht findet), was auch zur Ausbildung einer Gruppe „Atti“ (Kongressakten) mit einer eigenen Sachgruppen-Unterteilung führte. Die Kongressakten fand man also auch nicht unter dem Titel, sondern unter dem Datum des Kongresses – gelegentlich

auch unter dem Namen des ersten Herausgebers. Nebeneintragen gibt es wenige, dafür nicht selten mehrere Aufnahmen für das gleiche Werk, etwa bei Wanderausstellungen, die aber nicht aufeinander verweisen.

Die Anlage des Katalogs erinnert in mancher Hinsicht an den berühmten Katalog des British Museums, in dem ebenfalls mit Sachgruppen, Formbegriffen und chronologischen Unterteilungen in einem alphabetischen Katalog gearbeitet wurde. In der Handhabung ist der nunmehr mehr als hundert Jahre alte Holzkastenkatalog, in dem die etwa DIN A 5 großen Blätter mit je einer Stange gesichert sind, außerordentlich unpraktisch. Man braucht Automechaniker-Werkzeug, um ein Blatt für Ergänzungen herauszuholen, dementsprechend wurden Bandnachtragungen nicht gerne gemacht – die Benutzer sollten, ganz im Sinne einer „finding list“, die Signatur notieren und dann bitte am Regal nachsehen, was im Bestand war.

Im Jahre 1983 wechselte man zu einem Katalog mit Kärtchen im internationalen Format, der dank der Speicherschreibmaschine mit mehr Nebeneintragen bestückt werden konnte, behielt jedoch im Wesentlichen die für den ersten Katalog beschlossenen Sonderregeln bei. Die Bandnachtragungen wären infolge einer neuen Schraubkonstruktion an den Kästen einfacher gewesen, blieben aber unbeliebt. Die inzwischen erfolgte Einstellung von bibliothekarisch geschultem Personal sorgte dafür, dass viele formale Regeln der International Standard Bibliographic Description (ISBD) angewendet wurden, wie etwa Deskriptionszeichen und die genormte Abfolge der bibliographischen Beschreibung. Nicht erfasst wurde aber zum Beispiel weiterhin die ISBN. Die Katalogkärtchen haben alle ein handschriftliches Element: die mit Bleistift geschriebene Signatur. In einer Kunstbibliothek liegt es nahe, dass der eine oder andere Bearbeiter einen Sinn für Kalligraphie hat – oft ist die Signatur mit den unterschiedlichsten Schnörkeln verziert, was die Umsetzung nicht sonderlich befördert hat.

Vorüberlegungen bei der Retrokonversion in Florenz

Die Max-Planck-Gesellschaft hatte als erste der damals drei Trägereinrichtungen des Kunstverbundes erkannt, wie wichtig der Nachweis des Gesamtbestandes im 1996 mit den Neuerwerbungen begonnenen gemeinsamen Kubikat-Online-Katalog sei, so dass die Bibliotheca Hertziana als erste der drei Partnerbibliotheken die Retrokonversion europaweit ausschreiben und eine holländische Firma damit beauftragen konnte. Im Zuge der Verbundabsprachen wurde beschlossen, dass auch das damals vom Bundesministerium für Bildung, Wissenschaft, Forschung und Technologie getragene Kunsthistorische Institut mit derselben

Firma arbeiten sollte, was die gleichmäßige Qualität der eingegebenen Daten garantieren, den externen Dublettenabgleich zwischen Rom und Florenz zur Vermeidung von Doppeleingaben in unsere Datenbank ermöglichen und die Entwicklungskosten für den Datenimport reduzieren sollte. In Florenz beschloss man, angesichts beschränkter Finanzen, zunächst nur den „leichter“ umzusetzenden maschinenschriftlichen Alphabetischen Katalog für die Erwerbungsjahre 1983-1996 zu konvertieren.

Ein Grund für die langsamere Vorgehensweise in Florenz war aber auch, erst einmal Erfahrungen mit einer Firma sammeln zu wollen, die bislang vor allem mit Retrokonversions-Projekten auf nationalbibliothekarischer Ebene hervorgetreten war, auf der die zugrunde liegenden bibliographischen Daten sehr gut und stark standardisiert sind. In Spezialbibliotheken sind die Kataloge wegen der Freihand-systematik aber oft sehr viel weniger präzise, was allgemeine bibliographische Anforderungen angeht, und viel genauer in Wissenszusammenhängen, nicht zuletzt dank der leicht resignierten Einsicht, dass ein nicht bibliothekarisch geschulter Katalogbenutzer mit manchen von den RAK vorgeschriebenen Informationen nicht allzu viel anfangen kann. Änderungen von Körperschaftsnamen mussten beispielsweise in den Zettelkatalogen weder vom Bibliothekar noch vom Benutzer nachvollzogen werden, da man die Sache durch den Einstieg unter dem Ort entschärft hatte, der sich bei den ortsfesten Kunsteinrichtungen so schnell nicht ändert. Das erleichterte die Suche im Katalog und sparte Katalogisierungszeit, hatte aber für die Konvertierung zur Folge, dass man die Aufnahmen auch im Kontext der Einordnung sehen muss, um möglichst alle Informationen in einen Online-Katalog zu übertragen. Auch gab es Unterschiede im Katalogisierungsniveau: die Altdaten der Bibliotheca Hertziana, in der immer bibliothekarisch ausgebildetes Fachpersonal katalogisiert hatte, hatten ein für spezialbibliothekarische Verhältnisse sehr hohes Niveau. Von der Qualität der Florentiner Altdaten war man weniger überzeugt, weil in Florenz wegen chronischer Personalknappheit schon immer mit flugs angelernten Aushilfskräften katalogisiert worden war.

Einscannen des Katalogs, Datenerfassung und notwendige Nacharbeiten

Der AK 1983-1996 wurde während der Bibliotheksschließung im August 2000 in Holland gescannt. Wir haben es damals leider versäumt, eine Kopie des Scans für das KHI zu verabreden, so dass sich erst im Zuge der Datenlieferungen herausstellte, dass mehr als 5.000 von insgesamt ca. 45.000 Titeln die Mailboxkategorie „Signatur unleserlich“ trugen. Die Firma hatte alle Kärtchen mit einer einzigen Einstellung gescannt, mit diesem Resultat. Die

Daten-Erfassung erfolgte in Heimarbeit durch holländische Hausfrauen, in einem zweiten Durchgang sollte eine bibliothekarische Kontrolle stattfinden, die vor allem auch zwischen Rom und Florenz dublette Titel ermitteln und mittels „Ansigeln“ der Lokaldaten nur eine Aufnahme verwenden sollte. Rom und Florenz haben in Zusammenarbeit mit der Firma je ein Pflichtenheft erstellt. Das Florentiner basiert auf dem römischen, ist aber nicht identisch damit. Wir haben beispielsweise für Florenz eine abweichende Behandlung von Sonderdrucken vereinbart, die die Recherche der Fundstelle erlaubt, waren aber durch viele in Zusammenarbeit von Bibliotheca Hertziana, Thomas Berger (unserem Allegro-Koordinator) und der Retrokonversions-Firma getroffene Vereinbarungen von vornherein festgelegt.

Die Einsparung der eigenen Entwicklungsarbeit hatte große Vorteile für das KHI, aber uns schmerzt beispielsweise die Entscheidung, der Firma nicht die gesamte Datenbank, sondern nur eine Auswahl der vorhandenen Stammsätze zur Verfügung gestellt zu haben. Dies hat Nacharbeiten zur Folge, deren Ende nicht abzusehen ist. Die Firmenmitarbeiter waren stets freundlich und höflich, legten aber auch auf ausdrückliche Nachfragen nie die verwendeten Methoden klar, und konnten kurz vor dem Insolvenzverfahren, das die Firmenaktivitäten schließlich beendete, nicht auch noch mit Dublettenabgleichen behelligt werden. Vorsichtige Schätzungen waren von 35-40 % Bestandsüberschneidung zwischen Rom und Florenz ausgegangen, tatsächlich angesigelt wurde schließlich nur in weniger als 10 % der Aufnahmen. Sofort nach der Dateneinspielung im März/April 2002 wurde in etwa drei Monaten in Florenz ein erster Nachbearbeitungslauf für wenige Felder, insbesondere die Signatur, mit drei angelernten Kräften durchgeführt.

Die Nachbearbeitung erfolgt anhand des Original-Katalogs, um zu sehen, ob Aufnahmen fehlten – was zum Glück nur in geringem Umfang der Fall war. Hingegen stellte sich heraus, dass mehr als hundert Signaturen an die falschen Aufnahmen angesigelt waren, offenbar durch einen nicht genügend differenzierten automatisierten Dublettenabgleich. Sachtitel und Erscheinungsjahr genügen in Kunstbibliotheken nicht zur Identifizierung eines Werkes, vor allem nicht für Publikationen in Jubiläumsjahren von Künstlern. Bei jeglichem Problem (Dubletten, an die falsche Aufnahme angesigelt, Strukturprobleme wie mehrbändiges Werk versus Serienkatalogisat) werden nur die Identnummer des Satzes und ein Kürzel für die Art des Problems für eine weitere, spätere Nachbearbeitung notiert – auf Zetteln. Diese zweite Nachbearbeitung erfolgt seit Herbst 2002 durch eine besonders tüchtige studentische Hilfskraft, die hoffentlich einmal die Bibliothekslaufbahn einschlagen wird, und läuft

halbtags noch mindestens bis zum Jahresende 2003. In Absprache mit den Bearbeiterinnen in Rom werden auch die Mailboxen für fehlende Verknüpfungen und Nachbesserungen aller Art bearbeitet. Thomas Berger verdanken wir nicht nur die Entwicklung von Tests und des Datenimports, sondern auch diese Mailboxen, die beim Einspielen ins Allegro-System generiert wurden, und auch diverse Mechanismen für eine bequemere Nachbearbeitung. Per Verbundabsprache ist festgelegt, dass, wer eine Aufnahme bearbeitet, sie zu vervollständigen hat (also etwa auch die Münchner Kollegen anlässlich einer antiquarischen Erwerbung) – die gleiche Arbeitsteilung, die wir auch bei Neuaufnahmen anwenden.

Neuorientierung durch die neue Trägerschaft

Seit dem 1.1.2002 ist auch das KHI ein Institut der Max-Planck-Gesellschaft, und man konnte sich – finanziell und in der Projektarbeit mehr gefördert – der Konversion der restlichen Altdaten für Monographien vor Erwerbungs-jahr 1983 zuwenden: 96 Katalogkästen à ca. 1.500 Aufnahmen, insgesamt etwa 145.000, von denen Zeitschriftentitel, Mehrfachaufnahmen und Verweisungen abzuziehen sind, andererseits aber auch oft mehrere Auflagen oder Ausgaben eines Werkes auf einer Karte stehen.

Angesichts der Erfahrungen mit der ersten Retrokonversion mussten neue Wege eingeschlagen werden. Der Katalog wurde auch dieses Mal gescannt, weil die Bearbeitung des Scans und der Datenbank auf *einem* Monitor ergonomischer ist, mindestens dann, wenn es sich um 21-Zoll-Flachbildschirme handelt, auf denen man mehrere Bilder und Datenbanken zugleich geöffnet haben kann, ohne dass sie sich überlagern. Diesmal haben wir die Rechte an den Bildern und eine Kopie davon im Hausnetz zur Verfügung gestellt. Die Ordner entsprechen den Katalogkästen, auf die so nicht nur die Retro-Katalogisierer zugreifen können. Die besonderen Schwierigkeiten des alten alphabetischen Katalogs – kaum Normierung (noch nicht einmal die Signatur steht immer an der gleichen Stelle), Minimal-Angaben im Sinne einer Short Title List, unterschiedliche Erfassung durch unterschiedliche Bearbeiter im Lauf der Jahre – ließen voraussehen, dass zahllose Rückfragen anfallen würden, die in vielen Fällen nur im Hause selbst entweder von den Mitarbeitern oder anhand des Buches selbst zu beantworten wären. Bei einer hausinternen Erfassung stellen sich keine Import-Probleme, keine Probleme mit der zwischenzeitlich erfolgten Aktualisierung der ja lebenden Datenbank (Firma hat an Aufnahme 1 angesigelt, die zum Zeitpunkt der Einspielung aber nicht mehr existiert, da aus hoffentlich guten Gründen auf Aufnahme 2 umgeleitet).

Aber eine wirklich hausinterne Erfassung durch ein Team von mehreren Halbtagskräften (mehr als 4–5 Stunden kann man sich auf diese Art von Arbeit nicht konzentrieren) unter Leitung einer der beiden Diplom-Bibliothekarinnen des KHI, die von anderen Aufgaben nicht befreit werden kann, wäre schon aus organisatorischer Sicht ein Fiasko geworden. Eine Florentiner Firma, gut eingeführt auf dem lokalen Markt durch verschiedene Retrokon-Projekte etwa in der Bibliothek der Uffizien und im Gabinetto Vieusseux und außerdem nationaler Preisträger für die beste bibliothekarische Scan-Indexierungs-Software, hatte Interesse an dem Projekt. Unsere Vorgaben waren: Direkterfassung in die lokale Florentiner Datenbank, und zwar in der DOS-Version von Allegro, die weniger sprachabhängig ist als die Windows-Version a99; eine Teamleiterin für die Einarbeitung und Arbeitsorganisation des möglicherweise während der Projektzeit wechselnden Personals, für die Qualitätskontrollen und Zusammenarbeit mit der Bibliothek. Außer einem neuen Pflichtenheft entstand auch eine ausführliche Anleitung zum Arbeiten mit Allegro, alles auf Italienisch.

Die Einrichtung der Arbeitsplätze, Schulungen, Erarbeiten von Pflichtenheft und Anleitungen erfolgten im Oktober/November 2002. Sechs Titelaufnehmer arbeiten in zwei Schichten à 4 bis 5 Stunden zwischen 8.30 und 18.00 Uhr. Pro Mitarbeiter und Stunde müssen im Schnitt sechs Aufnahmen gemacht werden, entweder durch Neuanlegen oder durch Bearbeiten einer schon vorhandenen Aufnahme, zum Beispiel aus den römischen Retrokon-Aufnahmen. Für den Datenbankadministrator ist zu bedenken: fest angestellte Bibliotheksmitarbeiter sind in der Regel nicht untröstlich, wenn man ihnen mitteilen muss, dass die Bibliotheksdatenbank einige Stunden nicht zur Verfügung stehen kann, eine andere Sache ist es, Ausfallzeiten einer Fremdfirma plausibel zu machen. Nicht dass Allegro-Datenbanken ständig ausfielen, im Gegenteil, aber einmal im Monat ist Wartungs-Nachmittag, und gelegentlich gibt es Sonderaktionen.

Das KHI bezahlt pro Aufnahme, der vereinbarte Preis ist wegen der umfangreichen Recherchen (auch für die Stammsätze) der gleiche für Neuanlegen oder Bearbeiten vorhandener Aufnahmen. Die Firmenmitarbeiter konnten wählen, ob sie pro Aufnahme oder Stunde bezahlt werden wollten, und haben sich alle für Bezahlung pro Stunde entschieden. Zunächst waren wir von 8 Aufnahmen pro Stunde ausgegangen, mussten aber nach den ersten Erfahrungen auf 6 zurückgehen. Problematisch ist, dass es im Moment eher schon zu viele Aufnahmen für ein Werk in der Datenbank gibt, da die ersten Retro-Maßnahmen ja noch längst nicht durchgehend nachbearbeitet sind. Täglich werden von den Katalogisierern etwa 150 Aufnah-

men bearbeitet, monatlich etwa 2.500. Mit dem Abschluss der Arbeiten rechnen wir im Jahr 2005. Die Mitarbeiter sind Italiener, alle mit bibliothekarischer Erfahrung und/oder Ausbildung, teils schon in anderen Projekten der Firma bewährt, teils über die Online-Jobbörse der Associazione Italiana Biblioteche angeworben. Die Teamleiterin ist Kunsthistorikerin mit einer zusätzlichen bibliothekarischen Ausbildung und guten Deutschkenntnissen.

Auf einem guten Weg

Nach den turbulenten ersten drei Monaten Oktober bis Dezember 2002 haben sich jetzt Kontrollverfahren und Kommunikation per Mailboxen an der Aufnahme, per Mail, über Gruppenverzeichnisse im Hausnetz eingespielt. Die Teamleiterin kontrolliert mindestens 10 % der bearbeiteten oder neu erstellten Aufnahmen, die Aufnahmen der ersten drei Monate werden vollständig überprüft. Ein zurzeit noch großer Anteil der Kontrollen findet in direkter Zusammenarbeit der Teamleiterin mit der Bibliotheksbeauftragten statt – ein Verfahren, das sich als zeitsparender und weniger zäh als das Verfassen von Beanstandungsmeldungen erwiesen hat, und das gleichzeitig die Einarbeitung von Erläuterungen und Beispielen in die Arbeitsanleitungen erlaubt. Nach unseren Erfahrungen kann so die Tagesproduktion von sechs Titelaufnehmern an einem Arbeitstag kontrolliert werden. Das Pflichtenheft steht kurz vor der Abnahme und ist im April 2003 von uns mit demjenigen für die im Zentralinstitut für Kunstgeschichte in München beginnende Retrokonversion abgeglichen worden.

Der Katalog erweist sich in vielen Fällen als noch schwieriger zu interpretieren, als man ohnehin gedacht hatte: Das raffinierte hauseigene System der Umfangsangabe ist so raffiniert, dass es offenbar auch damals im Original-Katalog nicht von allen Bearbeitern richtig interpretiert worden ist. Sonderdrucke sind oft ohne präzise Angabe oder ganz ohne Fundstelle aufgenommen. Dazu die Besonderheiten alter Drucke und alter Körperschaften, von Konvoluten, die Florenz in anderer Zusammenstellung als Rom besitzt, oder solchen von Sonderdrucken aus Zeitschriften und Büchern, die nicht nur Verknüpfungen mit der Fundstelle, sondern auch untereinander bräuchten, in zwei Kategorien, die aber zusammen in einer Aufnahme nicht vorkommen dürfen. Alles in allem sehen wir aber gute Fortschritte und sind trotz aller Schwierigkeiten sehr davon überzeugt, dass sich die erheblichen Investitionen von Zeit und Geld lohnen.

Uns bleibt, die Konversion des Aufsatzkatalogs vor 1996 zu organisieren (Kubikat enthält bibliographisch selbstständige und unselbstständige Einheiten). Der Aufsatzkatalog der Bibliotheca Hertziana ist schon in der gemeinsamen Datenbank, muss aber

noch umfangreich nachbearbeitet werden. Viele Stammsatz-Verknüpfungen mit Zeitschriftentiteln fehlen – die Kategorie ist vorhanden, aber nur mit Text besetzt, so dass keine Bestandsangaben für Florenz oder München angezeigt werden. Es wird

zwischen den Bibliotheksleitern darüber verhandelt, ob diese Nachbearbeitung und die Ergänzung der Florentiner Sacherschließungsdaten gemeinsam in einer Maßnahme in Angriff genommen werden können.

WER SUCHET, DER FINDET ... ABER NICHT IMMER DAS, WAS ER SUCHT

ZUR BEDEUTUNG VON METADATENSTANDARDS

Dirk Witthaut – (Sales & Consulting Manager, ADLIB Information Systems, Nürnberg)

Datenbanken im Internet

Museen, Bibliotheken und Archive produzieren laufend Unmengen an Informationen. Objektdokumentationen werden erstellt, wissenschaftliche Beiträge in Katalogen, Zeitschriften etc. publiziert. Zunehmend wird dazu auch die Möglichkeit der Online-Publikation im Internet genutzt. Ursprünglich schlichte Internetseiten mit rudimentären Informationen zu den Museen entwickeln sich zu Wissensportalen, die in einigen Fällen auch den Zugang zu Objektdatenbanken ermöglichen. Eine Flut von digitalen Informationen wartet in den endlosen Weiten des Internets nur darauf, von Wissenschaftlern und interessierten Laien zur Kenntnis genommen zu werden. Doch die Frage ist, wie der Wissensdurstige auf schnellstem Wege zur Quelle der Weisheit gelangt.

Suchmaschinen

Der Erfolg von Internetsuchmaschinen wie Google hat auch in der Welt der Universitäten und Museen zu hohen Erwartungen an die Suche nach wissenschaftlichen Informationen im Internet geführt. Es gilt fast schon als selbstverständlich, mit den relativ einfach zu handhabenden Suchmaschinen schnell zu umfassenden Suchergebnissen zu kommen. Es mag zwar sein, dass jemand auf seine Anfrage hin Tausende von Treffern erhält, doch von einem gezielten Retrieval oder Vollständigkeit der Ergebnisse kann nicht einmal ansatzweise die Rede sein. Internet-Suchmaschinen sind ohne Frage ein höchst nützliches Werkzeug, doch ein großer Teil der im Prinzip im Internet vorhandenen Informationsquellen bleibt ihnen verschlossen. Das gilt insbesondere für den Inhalt von Datenbanken.

Die Ergebnisse einer Suchmaschine basieren auf der Indexierung von HTML- und seit kurzem auch PDF-Dokumenten durch so genannte Roboter oder Spinnen. Bedauerlicherweise hat der Prozess der Indexierung an der Schwelle zu Datenbanken

ein Ende. Jeder Hersteller von Datenbanken hat seine eigenen Methoden, um Datenbankinformationen im Internet zugänglich zu machen. Solange eine Suchmaschine die spezifische Suchsyntax einer Datenbank nicht kennt, kann sie den Inhalt der Datenbank nicht indexieren und folglich auch nicht in eine Suche mit einbeziehen. Nicht umsonst werden Datenbanken auch als die schwarzen Löcher des Internets bezeichnet. Wichtige Informationen zu Sammlungsbeständen von Museen und Bibliotheken können nur abgefragt werden, wenn man die betreffenden Internetseiten aufruft und eine Suchanfrage startet. Vorausgesetzt, die Datenbank des betreffenden Museums ist überhaupt online zugänglich. Für den Forscher, der zum Beispiel nach Objekten für eine Ausstellung sucht, die auf mehrere Museen verteilt sind, bedeutet dies, dass er kostbare Zeit damit vertut, Datenbanken zu suchen, in ihnen zu recherchieren oder schriftliche Anfragen an die einzelnen Museen zu verschicken.

Open Archives Initiative

Einen wichtigen Ansatz zur Lösung dieses Problems hat die 1999 ins Leben gerufene Open Archives Initiative¹ (OAI) entwickelt. Ziel von OAI ist es, den breiten Zugang zu wissenschaftlichen Informationen im Internet zu ermöglichen. Herbert van der Sompel und Carl Lagoze von der Cornell University haben ein Protokoll entwickelt, mit dem es auf der Basis von HTTP und XML möglich ist, auch den Inhalt einer Datenbank mit einer Internetsuchmaschine zu erfassen. Der Schlüssel zu einer effektiven Suche in Datenbanken sind die so genannten Metadaten.

OAI hat ihren Ursprung in der Welt des E-Print. Gerade im Bereich der wissenschaftlichen Publikationen besteht schon seit längerem das Interesse an Alternativen zur traditionellen Form der gedruckten Publikation. Das Internet bietet sich geradezu