

Sinn und Nutzen von Datenbanken in den Geisteswissenschaften

Veranstaltet von Prof. Dr. Theo Girshausen, Ingo Jonas M.A.,
Institut für Theaterwissenschaft (ITW) der Universität Leipzig
Leipzig, 24. bis 25. März 2006

Der „Sinn und Nutzen von Datenbanken in den Geisteswissenschaften“ liegt in ihrem Gebrauch. Dass hinter dieser simplen Einsicht durchaus kein trivialer Gedanke steht, konnten Ende März ca. 50 Teilnehmer in einem Leipziger Workshop ganz konkret erfahren. Anders als es der ultimative Titel vielleicht vermuten lässt, ging es den Veranstaltern weniger um prinzipielle Sinnfragen, als vielmehr um den ganz praktischen, problemorientierten Erfahrungsaustausch.

Theo Girshausen und *Ingo Jonas* (Institut für Theaterwissenschaft der Universität Leipzig) hatten über ein Call for Paper bundesweit Betreiber von geisteswissenschaftlich orientierten Fachdatenbanken eingeladen, sich über ihre Projekte auszutauschen und die mitunter sehr speziellen Probleme zwischen idealem Anspruch und technischer Realisation zu diskutieren. Neben der anwendungsbedingten Vielfalt der Projekte standen vor allem Fragen der Projekt-Organisation, zu verwendeter Software und zu dem Metadaten- und Content-Management im Vordergrund. Schnell wurde klar, dass unter den Teilnehmern ein enormer Informationsbedarf zur Entwicklung und Betreuung von Datenbanken besteht, die längst nicht mehr nur Hilfsmittel, sondern unverzichtbare Produkte geisteswissenschaftlicher Arbeit sind. Auch die DFG hat in den letzten Jahren die Vorteile von Datenbanken für die Erschließung von Quellen und die geisteswissenschaftliche Forschung insgesamt erkannt und diese Projekte entsprechend gefördert.

Günther Heeg (ITW) eröffnete den Workshop und wies darauf hin, dass Datenbanken als komplexe Gebilde in der Regel nicht mehr durch Einzelpersonen betreut werden können, sondern die Zusammenarbeit verschiedenster Wissenschaftler und Fachdisziplinen erfordern, andererseits eine interdisziplinäre und internationale Zusammenarbeit ermöglichen. 19 verschiedene Datenbank-Projekte wurden dem Plenum von den Entwicklern in jeweils halbstündigen Präsentationen vorgestellt und zur Diskussion gestellt.

Referate

Ingo Jonas, Initiator des Workshops, eröffnete die Runde mit einer Datenbank zu den Stücken und der Aufführungspraxis von August Wilhelm Iffland (1759-1814), Autor, Schauspieler und Direktor, welche Teil seines Dissertationsprojektes ist. Aufgrund des enormen Korpus bürgerlicher Familiendramen sei es mit Hilfe der relationalen Datenbank möglich, die Stücktexte hinsichtlich gesellschaftlicher Konventionen (Beruf, Stand, Verwandtschaft) zu untersuchen. Mit der Betrachtung von Theaterzetteln, Briefen und Theaterperiodika werden zusätzliche Informationen aufgenommen, die die Aufführungspraxis Iffland'scher Stücke untersuchbar machen. Dabei können Aussagen über geographische Verteilungen und gesellschaftliche Relevanz in Form von Stückbewertungen und Anzahl der Aufführungen bestimmter Stücke getroffen werden. Zugleich seien auch Aussagen über eine Zeit und die gesellschaftliche Situation möglich, die so bisher noch kaum erforscht sei.

Der Informatiker *Holger Gast* stellte das Dissertationsprojekt von Almut Mehling (Ur- und Frühgeschichte) vor. Beide haben an der Universität Tübingen eine relationale Datenbank zur detaillierten Analyse und Erfassung von Bildinhalten von Lebensbildern entwickelt. Ziel des Projektes sei ein fundierter Nachweis

zwischen Lebensbild und seinem Entstehungszeitraum gewesen. Gerade der Umgang mit „weichen Inhalten“ machte hier eine klare, eindeutige und entsprechend diversifizierte Definition der Beschreibungskategorien schon für die Dateneingabe unerlässlich, so dass man sich hier für die Atomarisierungsthese der Faktenaufnahme entschied. Die wichtigsten Ergebnisse der Studie, die durch die statistische Aufbereitung des Materials möglich waren, werden im Juli 2006 von Almut Mehling als Buch publiziert. Die Datenbank selbst sei allerdings nicht öffentlich zugänglich. Gast machte deutlich, wie die Arbeit mit der Datenbank für Bildbeschreibung über die ursprünglich gefassten Thesen hinausführte. Mit der fortschreitenden Auswertung des Materials seien viele neue Fragen aufgetaucht, die sich anfangs nicht gestellt hätten.

Diese Einsicht nahm bereits Äußerungen und Erfahrungen anderer Referenten vorweg, die die Relevanz von Datenbanken für die Forschung bekräftigten, immer vorausgesetzt, qualifizierte Fragen leiten die Arbeit am Material. Meist ist aber schon mit der Erschließung und Digitalisierung großer, historischer Quellenkonvolute wie z.B. Handschriften, Tageszeitungen etc. eine neue Ausgangslage für eine intensivere Forschung gegeben.

Über die Möglichkeiten und Grenzen einer projektübergreifenden Datenbank in einem interdisziplinären Forschungsverbund berichtete *Stefanie Rüther*. Sie koordiniert seit mehreren Jahren die Zusammenarbeit verschiedener Fachdisziplinen innerhalb des SFB 496 „Symbolische Kommunikation und gesellschaftliche Wertesysteme vom Mittelalter bis zur französischen Revolution“ an der Universität Münster. Ziel der Datenbank ist es, die Informationen so aufzubereiten, dass sie für interdisziplinäre Fragestellungen nutzbar gemacht werden können. Zum Beispiel für den diachronen Vergleich einzelner Zeichen, Bilder, Rituale und Werte in ihren unterschiedlichen Verwendungszusammenhängen. Zentral sei hier die Frage nach Wandel sowie Interdependenz symbolischer Kommunikation und gesellschaftlicher Konflikt- bzw. Friedensrituale im Spätmittelalter.

Als problematisch habe sich der Umgang mit verschiedenen Quellengattungen (Bildern, Gerichtsakten, historiographischen und fiktionalen Texten) als auch der verwendeten fachbedingten Termini erwiesen, auch die Suche nach einer technisch optimalen Vernetzung der sieben Teilprojektdatenbanken. Als willkommener Nebeneffekt habe sich allerdings auch die datenbankbedingte Zusammenarbeit und Verständigung der einzelnen Fachbereiche verbessert. (<http://www.uni-muenster.de/SFB496/>)

Till Schicketanz und *Kay Heiligenhaus* brachten als Inhaber ihrer Firma „semantics Kommunikationsmanagement GmbH“ (Aachen) die Perspektive der Datenbank-Entwickler ein. Sie entwickeln und betreuen ganz unterschiedliche Datenbankprojekte mit geisteswissenschaftlichem Hintergrund (Literaturmanagement, Digitalisierungsprojekte, Datenbankportale, etc.). Kay Heiligenhaus stellte drei der in enger Kooperation mit Hochschulinstitutionen entwickelten Internetprojekte vor: Die Online-Version der „Bibliographie der deutschen Sprach- und Literaturwissenschaft“ (<http://www.bdsl-online.de>) und das Zeitschriftenportal „Compact Memory“ (<http://www.compactmemory.de>). Dieses „Wissenschaftsportal für Jüdische Studien“ enthält unter anderem ein Internetarchiv jüdischer Periodika. Weit über einhundert digitalisierte deutsch-jüdische Zeitschriften aus den Jahren 1806-1938 sind hier weltweit verfügbar. Schließlich wurde auch die virtuelle Sammlung „Jiddische Drucke“ der UB Frankfurt (<http://www.Literatur-des-Judentums.de>) vorgestellt.

Heiligenhaus wies nachdrücklich darauf hin, dass auch in den Geisteswissenschaften die Herausforderung wachse, die Recherche und Erfassung bzw. Verwaltung und Analyse von digitalem Content in die eigene Arbeit zu integrieren. Forscher, Lehrende und Studierende aller geistes- und kulturwissenschaftlichen Disziplinen seien vor die Aufgabe gestellt, nicht nur rezipierend, sondern aktiv produzierend am datenbank- und internetgestützten Auf- und Ausbau digitaler Angebote mitzuwirken.

Heiligenhaus stellte unter anderem das Online-Recherchemodul „Visual Library“ vor (<http://www.visuallibrary.net>). Diese von „semantics“ entwickelte Suchmaschine führt die Suche in Bibliothekskatalogen, Buchhandlungen, Fachdatenbanken, Enzyklopädien und Archiven in einer einzigen

Suchmaske zusammen und erleichtert so das Auffinden spezieller Bücher, Aufsätze, etc. (<http://www.semantics.de>). Während seiner Reise durch die verschiedensten bibliographischen Online-Datenbanken wurde schnell deutlich, dass es zwar ein enormes Angebot für die Metasuche und Literatur-Recherche im Netz gibt, zu viele Wissenschaftler aber offenbar kaum in der Lage sind, all diese Möglichkeiten auch intensiv zu nutzen, weil sie immer noch zu wenig bekannt sind und beherrscht werden.

Karin Orchard vom Sprengel Museum Hannover stellte das dort angesiedelte Kurt Schwitters-Archiv vor. Es enthält eine umfassende Dokumentation von Schwitters Werk, ein Werksverzeichnis, dass ständig ergänzt wird. Kurt Schwitters war eine der herausragenden Künstlerpersönlichkeiten der zwanziger Jahre. Das Archiv ist chronologisch geordnet und umfasst derzeit etwa 3000 Werke. Von ca. 2000 Werken existieren Fotografien oder Ektachrome. Die Arbeit des Archivs bestehe im wesentlichen darin, die bereits vorhandenen Angaben zu überprüfen und zu ergänzen, neue Arbeiten weltweit ausfindig zu machen und in das Archiv aufzunehmen. Das Archiv ist Anlaufstelle für wissenschaftliche Recherchen und Anfragen zu Kurt Schwitters.

Grundlage der Forschungsarbeit zu Schwitters Werk ist eine umfangreiche Online-Datenbank (<http://www.kurt-schwitters.org>), auf der die Abbildung, die technischen Daten eines Werkes sowie dessen Provenienz, Ausstellungs- und Publikationsgeschichte dokumentiert werden. Die Recherche über einzelne Werke wird durch eine starke Vernetzung der einzelnen Informationen sehr erleichtert.

Dana Pflugmacher von der Niederlandistik der Universität Leipzig steht erst am Beginn einer Online-Bibliographie der deutschen Übersetzung niederländischer Literatur vom Mittelalter bis 1830. Das von der DFG geförderte Projekt entsteht in Zusammenarbeit mit der FU Berlin, die den Zeitraum ab 1550 bearbeitet. Die Datenbank soll ein Instrumentarium für zukünftige Forschung über die niederländisch-deutschen Literaturbeziehungen sein und die vorhandene Bibliographie ab 1830 von Herbert van Uffelen um die früheren Perioden ergänzen. Geplant sei nach Abschluss der Arbeit sowohl eine Veröffentlichung in Druck als auch über das Internet durch eine Online-Datenbank, die weitere Aktualisierungen ermöglicht. Die Bibliografie werde selbständige und unselbstständige, integrale und fragmentarische Übersetzungen aufnehmen, wobei der Begriff der Übersetzung im Rahmen des Projekts auch den der Bearbeitung einschließe. Erfasst werden sollen alle Textarten, die Gegenstand der Literaturwissenschaft seien, so dass epische, dramatische und poetische, aber auch geschichtliche und religiöse Texte sowie Werke aus dem großen Bereich der Artes-Literatur Berücksichtigung fänden. Ziel des Projektes ist eine möglichst vollständige Erfassung aller bis dato angefertigten deutschsprachigen Übersetzungen und Bearbeitungen der niederländischsprachigen Literatur aus dem Zeitraum vom Mittelalter bis 1830.

Iacopo Cividini vom Institut für Musikwissenschaft der Universität München referierte über das Bayerische Musiker-Lexikon Online (<http://www.bmlo.lmu.de>), welches einmal die Biographien aller Musiker enthalten soll, die jemals in Bayern und der italienischen Peripherie aufgetreten sind. Es soll dem Nutzer neben den Biographien von Komponisten und Dirigenten, Sängern und Instrumentalisten, Musikverlegern, Instrumentenbauern und Musikforschern aller Epochen auch entsprechende Video- und Ton-Beispiele, Bilder etc. bieten. Das Portal ist ein gemeinsames Pilotprojekt des Musikwissenschaftlichen Instituts der Universität München und der Bayerischen Staatsbibliothek, welche mit dem Sondersammelgebiet Musik einen beachtlichen Quellen-Umfang einbringt. Ziel sei ein digitales Nachschlagewerk zur bayerischen Musikgeschichte, das dann u.a. in verschiedenen Internet-Fachportalen der Bayerischen Staatsbibliothek nutzbar sein soll. Biographische Musikforschung wolle man hier mit der landeskundlichen Literatur zusammenführen. Dazu würden bestehende Lexika retrodigitalisiert sowie Spezialliteratur und weitere Bibliotheks- und Archivbestände erschlossen und vernetzt. Das BMLO soll die vielfältige Recherchefunktionalität einer Datenbank mit der Informationsfülle eines traditionellen Lexikons und der medialen Bandbreite des Internets verbinden, so Cividini. Schon jetzt deute sich an, dass die multimediale Ausrichtung der Seite die Entwickler vor das Rechtproblem stellen wird.

Viele Entwickler beklagten im Lauf des Workshops grundsätzlich das Rechtproblem für die Speicherung von Bildern etc. als ein enormes Hindernis für eine effektivere Ausstattung und Nutzung der Datenbanken innerhalb der Forschung.

Stefan Bordag, Diplom-Informatiker der Abteilung für automatische Sprachverarbeitung an der Universität Leipzig, sprach über die „Extraktion und Visualisierung von komplexen Zusammenhängen in Texten“. Er hat sich auf die Entwicklung von Algorithmen für die Filterung und Analyse von Volltext spezialisiert. Diese Algorithmen kommen innerhalb einer Datenbank als hoch effektive Werkzeuge zum Einsatz, können Textsorten weitgehend automatisch erkennen und für die Weiterverarbeitung innerhalb einer Datenbank Informationen aus dem Text extrahieren. Schnell wurde deutlich, dass der Einsatz solcher Werkzeuge, auf die jeweiligen Anforderungen richtig angepasst, enorme Kosten und Arbeitsaufwand sparen kann oder bestimmte Datenbank-Funktionen erst ermöglicht. Bordag sprach über die automatische Analyse von Sprachen, Texten und Sätzen und informierte über verschiedene Ansätze innerhalb der Algorithmen-Forschung. Auf der Basis eines enormen Text-Korpus europäischer Zeitungen und Webseiten, die permanent gesammelt werden, experimentiere die Arbeitsgruppe mit Algorithmen, deren Effektivität und Fehlerverhalten sich so leicht testen lässt. Statt einzelne Fehler zu beseitigen, werde der Algorithmus so angepasst, dass ganze Fehlertypen vermieden werden können. Ein Arbeitsschwerpunkt Bordags sei derzeit die Entwicklung von Algorithmen, die „Sprache besser verstehen“. Die Texte bzw. Sätze würden analysiert nach Schlagworten (Frequenz), Wortklassen, Wortähnlichkeiten, Morphologie, Mehrdeutigkeiten, Häufigkeiten, Signifikanz etc. Durch Vergleich mit statistischen Daten aus dem Textkorpus ließen sich über die Abweichungen Rückschlüsse auf Textart, Thema etc. ziehen oder typische Gebrauchsprofile von Wörtern erstellen. (<http://www.wortschatz.uni-leipzig.de/asv/>)

Hans-Jörg Bibiko vom Leipziger Max Planck-Institut für Evolutionäre Anthropologie informierte über die verblüffenden Möglichkeiten der Visualisierung von Daten mit SVG (Scalable Vector Graphics). Diese auch für Neulinge leicht anwendbare Graphik-Software wurde auch für die Visualisierung der Forschungsergebnisse in dem ehrgeizigen Sprachatlas des MPI verwendet (Worldatlas of Language Structures), für den 2600 Sprachen ausgewertet und nach typologischen Strukturen sowie globaler Verteilung erforscht wurden. Bibiko zeigte, wie sich auch die zeitliche, evolutionäre Entwicklung von Daten mit Hilfe dieser Software in graphische ‚Karten‘ umsetzen lässt, da die Datenreihen mit einem beliebigen Hintergrund und anderen Datenreihen sehr flexibel kombinierbar seien. Prinzipiell können mit Hilfe dieser Software Zusammenhänge, Korrelationen und Signifikanzen für alle erdenklichen Präsentationszwecke auf der Basis ganz beliebiger Datenreihen visualisiert werden. Bibiko plädierte eindringlich für den verstärkten Umgang und Einsatz entsprechender Programme innerhalb der Forschung und Lehre. (<http://www.eva.mpg.de/lingua/files/wals.html>)

Klaus Gerlach stellte die Online-Datenbank „Berliner Klassik“ (<http://www.berliner-klassik.de>) vor. An der Berlin-Brandenburgischen Akademie der Wissenschaften angesiedelt, rekonstruiert diese sukzessiv zu erweiternde, bereits beachtliche Datenbank Informationen über das Berlin zur Wende vom 18. zum 19. Jahrhundert. Dass es neben der Weimarer Klassik auch eine Berliner gab, beide nicht getrennt voneinander zu verstehen sind, diese Einsicht wollen die Entwickler durch umfangreiche Informationen über Personen, rund 150 Gesellschaften, Vereine und Clubs sowie Theater und entsprechende Literatur belegen. Eine komplexe Großstadtkultur habe das gesellschaftliche Leben im Berlin zwischen 1786-1815 geprägt. Die heuristische These ist, dass es in der deutschen Kulturblüte um 1800 *zwei* gleichwertige, wenn auch der Erscheinungsform nach sehr unterschiedliche kulturelle Zentren gab, nämlich Weimar *und* Berlin. Zum nationalen Mythos sei aber nur Weimar geworden. Während Weimar eher Ort künstlerischer Produktion gewesen sei und dort eine Stadtkultur mit entsprechendem Publikum gefehlt habe, sei das bürgerliche Berlin Ort intensiver Rezeption gewesen. Gerlach verwies auf den ausgeprägten Goethe- und Schillerkult in Berlin. Durch Erforschung des intensiven Wechselverhältnisses zwischen den Städten sollen tradierte Gegensätze aufgehoben und neue Zusammenhänge sichtbar werden.

In einer der vier Teildatenbanken werden Informationen zu den momentan 500 Personen und deren literarischen Werken gesammelt, die für Berlin kulturprägend waren. Eine Theaterdatenbank erfasst das Repertoire des Berliner „Königlichen Nationaltheaters“ und dessen Rezeption, die Geselligkeitsdatenbank Informationen zu den Gesellschaften und Vereinen, die in Berlin zwischen 1786-1815 existierten. Ein virtueller Stadtplan von Berlin sei in Arbeit.

Dass Datenbanken auch in Form von komplexen Hypertext-Strukturen angelegt und genutzt werden können, wie in den Geisteswissenschaften zweckmäßig, beweist der Literaturwissenschaftler *Benno Wagner* von der Universität Siegen mit seinem „Virtuellen Kafka-Bureau“ (<http://www.benno-wagner.net/kdp.htm>).

Die Kombination von Online-Datenbank und Hypertext soll dazu dienen, Kafkas Schreibbewegung und seine Lektüren sichtbar zu machen. Gerade ein Autor wie Kafka eignet sich besonders für dieses ambitionierte Projekt. Er selbst hat sich, beruflich bedingt, mit neuen Verwaltungs- und Kommunikationstechniken auseinandergesetzt. Wagner wies darauf hin, dass Kafka seine systematischen Textstrukturen und Verweise geradezu vorsätzlich an solche Formen anlehne. Die Konsequenz daraus müsse sein, die herkömmliche germanistisch-philologische Rezeption, die lediglich neue Lesarten entwickle, zu überwinden. So will Wagner und sein Team mit der Kafka-Seite zugleich ein neues „digitales Arbeitsumfeld für Literatur-, Medien- und Kulturwissenschaftler“ entwickeln.

Die erste Projektphase soll zunächst das Feld der – intermedial verstandenen – Kafkaschen Lektüren in Form einer relationalen Datenbank erschließen. Darauf aufbauend wollen die Forscher die abfragbaren Datensätze dieser „Virtuellen Mediathek Kafkas“ in der Logik eines dynamischen Hypertextes für die Nutzer navigierbar machen. Dessen Verknüpfung orientiere sich an dem Ensemble literarischer Chiffren, als das Kafkas Texte hier verstanden werden. Ziel des Projektes ist nicht weniger als die „elektronische Remedialisierung des Kafkaschen Schreibraums“.

Mit der neuen Seite soll eine Rezeption möglich sein, die zunächst eher einer Begehung des Textes gleiche und lediglich auf der Folie linearen Lesens stattfindet. Kafkas Originaltext diene als „Matrix“, an den ein „Echotext“ angelagert werde. Das heißt Originaltext und mögliche Konnotationen, Bezüge und Verweise auf offensichtlich verarbeitetes Material werden für eine Reise durch das Kafka-Universum angeboten, ohne dass bestimmte Interpretationen privilegiert werden sollen. Der Vorteil eines Hypertextes ist hier, dass die verschiedenen Kontexte und Referenztexturen in eine dynamische und wechselseitige Beziehung zueinander treten können. Was dabei entsteht, wäre Kunst und Wissenschaft in einem – ein kulturelles Archiv der Postmoderne.

Der Vortrag bildete insofern eine Ausnahme, da er sich intensiv mit den Visualisierungsmöglichkeiten von Datenbankergebnissen, hier eben in einem Hypertext, beschäftigte. Die Seite werde voraussichtlich erst im letzten Jahresviertel online gehen.

Den zweiten Tag des Workshops eröffnete *Kerstin Albrecht* von der FH Anhalt Dessau, die sich in ihrem Vortrag der Digitalisierung und Inventarisierung von Bilddateien widmete. Sie stellte das „Digitale Designarchiv“ (<http://www.dda-dessau.de>) vor, welches Designobjekte aller Art erfassen und als Abbildung inklusive einiger Zusatzinformationen zugänglich machen will. Wie bei vielen Datenbanken stellt sich hier das Problem der Erstellung eines geeigneten fachspezifischen Thesaurus. Da es die Designgeschichte immer mit Alltagsobjekten zu tun hat, kann hier nicht auf herkömmliche Verschlagwortungen zurückgegriffen werden. Die von der Referentin gewählte Lösung ist eine sogenannte Polyhierarchie. Dabei werden semantische Netze, ein Ansatz aus der Lexikologie, erstellt. Über Knotenpunkte als Oberbegriffe können so verschiedenen Unterbegriffe (auch mehrfach) zugeordnet werden.

Eine Schwierigkeit vieler Datenbanken beschrieb auch Albrecht für das dda. Die Vollständigkeit der Datenbank ist nicht erreichbar, so dass sie immer für Erweiterungen offen bleiben muss.

Ein Beispiel, wie gründlich die entsprechenden Vorüberlegungen für die Erstellung einer Datenbank sein müssen, lieferte *Regina Teske* von der BstU Berlin mit der „Datenbank über politische Gegnerschaft und unbekannten Widerstand in der DDR“ (<http://www.bstu.bund.de>). In Auswertung der Karteien des MfS, das bizarrerweise eine unschätzbare Vorarbeit leistete, wird in Berlin eine Datenbank über eine DDR-Datenbank erstellt, die anhand von Straftatbeständen politische Gegnerschaft dokumentieren und nachweisen will. Ob aber anhand des Verstoßes gegen Gesetze der DDR automatisch auf politische Gegnerschaft geschlossen werden kann, bleibt fragwürdig, ebenso, inwieweit durch dieses Vorgehen die Vorannahmen des MfS lediglich übernommen und so sanktioniert werden. Ein Teilnehmer schlug vor, im Titel der Datenbank „Widerstand“ durch „Verfolgung“ zu ersetzen. Neu war die Erfahrung, dass auch mit Datenbanken schwierige politische und gesellschaftliche Probleme angesprochen sein können.

Das Konzept eines E-Museums verfolgt die Stadt Düsseldorf. *Margret Schild* vom Theatermuseum Düsseldorf zeigte wie mit Hilfe eines einheitlichen Museums-Management-Systems ein umfassendes digitales Kunst- und Kulturarchiv einer Stadt entstehen soll. Die Datenbank „d:kult“ dient vor allem der internen Verwaltung und weniger dem geneigten Düsseldorferinteressenten. Die Düsseldorfer Museen erhoffen sich durch diese Archivierung sowohl eine Kostenersparnis, als auch neue Erkenntnisse durch erstmals mögliche Analysen über den Gesamtbestand.

(http://www.duesseldorf.de/presse/pld/d2004/d2004_09/d2004_09_30/p16234.shtml)

Dem Veranstaltungsort des Workshops durchaus angemessen thematisierte *Margret Schild* den theater-spezifischen Bereich der Datenbank. So sei es schwierig, eine Kunst zu inventarisieren, die weitgehend ohne Artefakte operiert. Die Inszenierungen können dabei nur als fiktive Objekte generiert werden, an die die weiteren Informationen (wie Schauspieler, Regisseur, Kostüm etc.) anzulagern wären.

Ein Projekt ganz anderer Art stellte die Historikerin *Katrinette Bodarwé* von der Universität Göttingen vor. Die Datenbank „FrKIDB“ (<http://www.frauenkloester.de>) versucht, alle mittelalterlichen Frauenklöster Europas zu verzeichnen, um so einen konstitutiven Beitrag zur Frauengeschichtsforschung zu leisten, dienen die Klöster doch für Frauen als singuläre Orte des Lernens im Mittelalter.

Bei der Verwendung von Datenbanken in den Geschichtswissenschaften ergeben sich spezielle Probleme. Dass der Status von Quellen aus dem Mittelalter häufig unsicher ist oder falsche Quellen mit langer Tradition ebenso mit registriert werden müssen, stellt dabei noch die geringste Schwierigkeit dar. Viel größer ist das Problem dagegen bei der Entscheidung für einen spezifischen Umgang mit Quellen und deren Selektion. *Bodarwé* betonte, dass mit jeder Erweiterung der Datenbank gleichzeitig eine Einschränkung an anderer Stelle einhergehe. Durch die Verwendung des Instrumentes Datenbanken ändert sich demnach auch die Struktur und Evidenz des Wissens über das Forschungsobjekt. Eine Datenbank zu erstellen kann demnach niemals heißen, das Quellenmaterial dem Shredder zu übergeben. Die Datenbank ist Werkzeug, nicht aber Ersatz für Quellen.

Dass es durch die neuen Medien auch in der Kunstgeschichte zu differenten Methoden kommen muss, bewies *Klaus Weber* aus Mainz. An der dortigen Uni wird versucht, eine fachübergreifende und interdisziplinäre Multimediatatenbank zu installieren. Bisher sind ca. 20.000 Bilder in die Datenbank aufgenommen. Diese verhältnismäßig geringe Zahl ist laut Weber hauptsächlich Resultat uni-interner Verwaltungsprobleme und vor allem bisher ungeklärter Bild- und Urheberrechte. Ohne eine einheitliche und praktikable Gesetzeslage sei man mit Bilddatenbanken am Ende, bevor man richtig anfangen könne, so Weber.

Eine bereits vollständig funktionierende und agierende Datenbank präsentierte *Daniel Burckhardt*. „Clio-Online“, angegliedert an das große H-Net und gefördert durch die DFG, sammelt Rezensionen zu Veröffentlichungen aus den Geschichtswissenschaften. Der Bestand liegt momentan bei ca. 30.000 Rezensionen zu 25.000 Publikationen. Laut Burckhardt bietet das Online-Angebot von Rezensionen enorme Vorteile und hat auf dem Markt der Besprechungen für eine kleine Revolution gesorgt. Durch Datenbanken wie Clio-Online ist der Abstand zwischen Publikations- und der Rezensionen-Zeitpunkt auf ein Minimum geschrumpft.

Zudem hat sich die Zugänglichkeit deutlich verbessert. Musste man früher, in den 1990er Jahren, noch tage-lang in Zeitschriften blättern, ist es nun möglich, per Suchmaske innerhalb von Sekunden die gewünschte Rezension zu finden.

Besonders innovativ ist die teilautomatisierte Einspeisung von Rezensionen in die hier vorgestellte moderierte Datenbank. Ein bestimmtes Tool durchforstet das Internet anhand der verbindlichen Kategorie der ISBN-Nummer nach neuen Publikationen/ Rezensionen und pflegt sie selbstständig in die Datenbank ein. Diese Methode führt einerseits zu großen Einsparungen, andererseits zu einer Ausweitung auch auf das globale Wissenschaftsfeld. (http://www.clio-online.de/site/lang__en/40208174/default.aspx)

Auch in den neuen EU-Ländern wird bereits intensiv an der Erstellung von Datenbanken gearbeitet. So stellte *Agnieszka Seidel-Grzesinska* aus Wroclaw/ Polen das Projekt „MIDAS“ vor. Hier sollen architektonische Monumente aller Art, vor allem jedoch Denkmäler katalogisiert werden. Das größte Problem dabei sei der internationale Anspruch und die damit verbundene Multilingualität, die laut Seidel-Grzesinska für eine kleine Sprache wie Polnisch unumgänglich sei. Vor allem die Übersetzung der Thesauri und das Integrieren der Sonderzeichen bestimmter Sprachen bereite dabei Schwierigkeiten. Die Referentin stellte die These auf, der Aufbau einer Sprache bestimme zu großen Teilen die Struktur der Datenbank. (<http://www.oss.wroc.pl/biuletyn/ebib06/midas.html>)

Den Abschlussvortrag hielt *Kristina Lowis*, die als Vertreterin der „Bibliographie de l’Histoire de l’Art Paris“ (BHA) referierte. Sie stellte das wohl größte Projekt auf dieser Tagung vor, welches zudem über die längste Tradition verfügt. Seit bereits mehr als 100 Jahren wird die internationale, vor allem westliche, Forschungsliteratur zur abendländischen Kunst katalogisiert. Mit der Digitalisierung dieser Bibliographie ist es nun möglich, auf über 330.000 Dokumente in 48 Sprachen zuzugreifen und einen jährlich Zugang von ca. 24.000 Lemmata zu bewältigen.

Das BHA versteht sich selbst als „Pfadfinder und Wege-Bahner“. Die Mitarbeiter aus Frankreich und den USA (Getty Research Institute) erstellen zu jeder Publikation kurze Resümees. Allein aufgrund des Umfangs der Datenbank ist hier in jahrzehntelanger Arbeit ein unvergleichliches wissenschaftliches Archiv entstanden. (<http://www.inist.fr/PRODUITS/bha.php>)

Ergebnisse

Web-basierte geisteswissenschaftliche Datenbanken können nicht nur sehr komplexe Quellen-Konvolute erschließen und verwalten, sie können auch Unterschiede, Wandel und die Dauer von Phänomenen repräsentieren. Die ökonomischen und erkenntnissteigernden Vorteile dieser Form der Datenorganisation und -präsentation sind offensichtlich. Aber wie kann die Datenbank einem größerem Kreis von potentiell interessierten Forschern zugänglich gemacht werden? Noch immer bleibt der Nutzer bei der Suche nach speziellen Informationen im Netz auf seine Intuition und Instinkte angewiesen. Deshalb sollten die vielen Angebote verstärkt in themenbezogenen Portalen gebündelt und das Material sinnvoll vernetzt werden.

Im Verlauf des Workshops kam die Frage auf, ob nicht durch die Formalisierung von Daten für die Erfassung innerhalb entsprechender Kategorien der Datenbank die Gefahr besteht, dass unbemerkt anthropologische Konstanten konstruiert oder Ergebnisse präfiguriert würden. Das führte zu der Einsicht, dass die Prämissen der Kategorienbildung für die Erfassung der Daten eventuell auch für die Nutzer transparent gemacht werden sollten. Nicht zuletzt um der Illusion einer voraussetzungslosen Objektivität zu begegnen. Denn die spezifische Eigenlogik einer Datenbank erfordert je nach Thema und Funktion fast immer bestimmte Zurichtungen des Materials und die Eingrenzung auf eine bedingte Geltung und Signifikanz der Ergebnisse.

Wiederholt wurde neben der Frage der Pflege auch das Problem der Archivierung von Datenbanken nach Förderungsende berührt. Pflegekosten sind auch hier ein enormer Faktor, zumal wenn es sich um Online-Projekte handelt. Es wurde deutlich, dass in diesen Fragen noch erheblicher Informationsbedarf besteht.

Da zwischen der ersten Planungssitzung bis zur vollen Nutzungsreife der Datenbank meist mehrere Jahre liegen, sei eine sorgfältige, d.h. vor allem realistische Planung der Ressourcen angeraten, so der Tenor. Schon früh die richtigen Fragestellungen und Begrifflichkeiten zu entwickeln, kann viel Geld und Zeit sparen, da eventuell auftretende Probleme und Nutzungserweiterungen so in die Planung einbezogen und die Programmierung entsprechend flexibel angelegt werden kann.

Die Entwickler sollten sich zunächst über Aufgaben, Umfang, Verfahren, Masken und Möglichkeiten der Online-Nutzung klar werden und die Nutzung an einigen Beispiel-Datensätzen durchspielen, ehe sie sich mit der Wahl einer adäquaten Software schicksalhaft an einen bestimmten Anbieter binden. Schon hier ist die Beratung durch einen erfahrenen Datenbank-Spezialisten sinnvoll.

Der erhebliche Kostenaufwand für eine Datenbank wirft die Frage auf, inwieweit nicht schon die Programmierung bereits beendeter bzw. anderer Projekte für weitere genutzt werden könnte. Diese Formen der Mehrfachverwertung sind offensichtlich noch unüblich, da einerseits das Interesse ganz automatisch auf die Neuentwicklung fixiert ist und es zudem keine entsprechenden Informations-Strukturen gibt, die dies organisieren könnten. Gerade in Rücksicht auf die angespannte Förderpraxis wären so auch Low-Budget-Projekte möglich, auch wenn die Anpassung der Software zunächst zusätzliche Ressourcen bindet.

Kay Heiligenhaus brachte als Inhaber und Till Schicketanz als wissenschaftlicher Projektleiter der Firma semantics Kommunikationsmanagement GmbH (Aachen) die Perspektive der Datenbank-Entwickler ein.

Ein Tagungsband mit den Beiträgen wird voraussichtlich im dritten Jahresviertel 2006 erscheinen.

René Damm, Sebastian Göschel

Kontakt:

rene.damm@theaterforschung.de

sebastian.goeschel@theaterforschung.de

Anmerkung

Wir möchten an dieser Stelle darüber informieren, dass unser Professor und Lehrer Prof. Dr. Theo Girshausen am 25. März nach kurzer schwerer Krankheit gestorben ist. Die Konferenz hatte er mitkonzipiert, konnte sie aber nicht mehr eröffnen. Wir trauern um einen herausragenden Wissenschaftler und Lehrer, der uns, seine Studenten, fachlich wie menschlich außerordentlich geprägt hat. Wir wollen ihm gedenken und versuchen, seinem Verständnis von Wissenschaft, Offenheit und Dialog in unserer Arbeit gerecht zu werden.

Copyright

Arbeitsgemeinschaft historischer Forschungseinrichtungen in der Bundesrepublik Deutschland e.V., 2006.

Kein Teil dieser Publikation darf ohne ausdrückliche schriftliche Genehmigung der AHF in irgendeiner Form reproduziert oder unter Verwendung elektronischer Systeme verarbeitet, vervielfältigt oder verbreitet werden.

AHF, Schellingstraße 9, 80799 München

Telefon: 089 – 13 47 29, Fax: 089 – 13 47 39

E-Mail: info@ahf-muenchen.de, Website: <http://www.ahf-muenchen.de>

Empfohlene Zitierweise / recommended citation style:

AHF-Information. 2006, Nr.062

URL: <http://www.ahf-muenchen.de/Tagungsberichte/Berichte/pdf/2006/062-06.pdf>