

# Datenbanken im Archiv für Zeitgeschichte der ETHZ

Autor(en): **Ehrsam, Thomas**

Objektyp: **Article**

Zeitschrift: **Geschichte und Informatik = Histoire et informatique**

Band (Jahr): **10 (1999)**

PDF erstellt am: **16.07.2024**

Persistenter Link: <https://doi.org/10.5169/seals-8092>

## **Nutzungsbedingungen**

Die ETH-Bibliothek ist Anbieterin der digitalisierten Zeitschriften. Sie besitzt keine Urheberrechte an den Inhalten der Zeitschriften. Die Rechte liegen in der Regel bei den Herausgebern.

Die auf der Plattform e-periodica veröffentlichten Dokumente stehen für nicht-kommerzielle Zwecke in Lehre und Forschung sowie für die private Nutzung frei zur Verfügung. Einzelne Dateien oder Ausdrucke aus diesem Angebot können zusammen mit diesen Nutzungsbedingungen und den korrekten Herkunftsbezeichnungen weitergegeben werden.

Das Veröffentlichen von Bildern in Print- und Online-Publikationen ist nur mit vorheriger Genehmigung der Rechteinhaber erlaubt. Die systematische Speicherung von Teilen des elektronischen Angebots auf anderen Servern bedarf ebenfalls des schriftlichen Einverständnisses der Rechteinhaber.

## **Haftungsausschluss**

Alle Angaben erfolgen ohne Gewähr für Vollständigkeit oder Richtigkeit. Es wird keine Haftung übernommen für Schäden durch die Verwendung von Informationen aus diesem Online-Angebot oder durch das Fehlen von Informationen. Dies gilt auch für Inhalte Dritter, die über dieses Angebot zugänglich sind.

# Datenbanken im Archiv für Zeitgeschichte der ETHZ

---

Thomas Ehram, Archiv für Zeitgeschichte

Das Archiv für Zeitgeschichte der ETH Zürich (AfZ) sichert und dokumentiert für die Öffentlichkeit private Quellen von gesamtschweizerischer Bedeutung zur Zeitgeschichte: Privatnachsätze und Archive von privaten Organisationen, Mikrofilme aus ausländischen Archiven zur Schweiz, Fotos, Tonaufzeichnungen und Zeitungsdokumentationen. Im Laufe seiner etwa dreissigjährigen Geschichte (nach privaten Anfängen in den sechziger Jahren wurde es 1974 von der ETH übernommen und dem Institut für Geschichte angegliedert) haben sich klare Sammlungsschwerpunkte herausgebildet: Jüdische Zeitgeschichte mit den Themen Emigration, Judenverfolgung, Antisemitismus, Flüchtlingspolitik, Juden in der Schweiz; ferner Wirtschaft und Zeitgeschichte, verstanden als Wirtschafts- und Aussenwirtschaftspolitik, und neuerdings der Kalte Krieg. Seit kurzem liegt ein Handbuch vor, das erstmals einen Überblick über alle Bestände im AfZ ermöglicht.<sup>1</sup>

Seit 1987 wird im AfZ die elektronische Datenverarbeitung eingesetzt. Am Anfang stand die zentrale Datenbank und (natürlich) ein Office-Paket (WordPerfect). Seither kamen andere Informatikaufgaben hinzu, von denen die wichtigsten wenigstens genannt seien: der Auftritt im Internet, die Digitalisierung der historischen Zeitungsdokumentation der Wirtschaftsförderung mit gleichzeitiger Mikroverfilmung zur Langzeitarchivierung – zugleich ein Pilotprojekt für weitere Digitalisierungen –, die Sicherung von Tonaufzeichnungen unserer Kolloquien mit Zeitzeugen auf CD's, schliesslich der Einsatz des Desktop-Publishings für eigene Publikationen. Im Folgenden werde ich mich aber auf die Darstellung der Datenbanken beschränken, die das Herzstück der Archiv-Informatik bilden. Von der Akquisition über die Erschliessung bis zur Recherche und der Findbucherstellung: alles geschieht in und mit der Datenbank. Das war nicht immer so. Als wir 1987 die erste, vor kurzem abgelöste zentrale Datenbank einführten, war ihr Einsatz zunächst auf den Nachweis der Akten-Dossiers beschränkt: sie ersetzte den zentralen Zettelkatalog.

Es gab zu diesem Zeitpunkt noch kaum Standardsoftware für Archive; wir waren deshalb zu einer Eigenentwicklung gezwungen, für die wir als

---

<sup>1</sup> Klaus, Urner; Däniker, Marie-Claire; Ehram, Thomas; Hoerschelmann, Claudia: *Das Archiv für Zeitgeschichte und seine Bestände*. Zürich 1999.

Datenbankprogramm DataPerfect der damaligen Firma WordPerfect wählten. Es handelte sich um eine relationale DOS-Datenbank, die bei grosser Leistungsfähigkeit durch einfache Handhabung und extrem kompakte Datenspeicherung auffiel. Als Knacknuss für die Entwicklung erwies sich schon damals die Klassifikation, die in einem Archiv mit vielen kleinen und nur relativ wenig grossen Beständen, für die jeweils eine eigene hierarchische Ordnung erarbeitet werden muss, von besonderer Bedeutung ist. Da man 1987 noch gar nicht daran denken konnte, eine Klassifikationshierarchie und einen zu einer bestimmten Klassifikation gehörenden Datensatz nebeneinander so darzustellen, dass ein Überblick über die Klassifikation ermöglicht und Klassifikationen und Dossiers nach Bedarf verschoben werden können, haben wir darauf verzichtet, direkt in der Datenbank zu erschliessen. Statt dessen haben wir die Bestände bis vor kurzem in WordPerfect erfasst, dessen Gliederungsfunktion es erlaubt hat, Klassifikationen leicht hin- und herzuschieben. Ein eigenes Programm hat dann die Aufteilung des WP-Files in verschiedene Dateien und Felder besorgt, die anschliessend in DataPerfect importiert werden konnten. So leicht damit die Erfassung war, so hat sich doch auch der Nachteil dieses Vorgehens gezeigt: da Bestände erst fertig erschlossen importiert werden können, vergeht gerade bei grossen Nachlässen oder Organisationsarchiven oft zu viel Zeit, bis sie abgefragt werden können. Zudem bedingt dieses Vorgehen eine sehr bescheidene Strukturierung der Daten.

Die Ablösung der alten durch eine neue Datenbank hatte aber noch handfestere Gründe: eine DOS-Datenbank, die zudem nicht mehr weiterentwickelt wird, hat in einer Zeit der grafischen Benutzeroberflächen keine Zukunft und die Volltextabfrage war bei etwas komplexeren Abfragen entschieden zu langsam. Dem sollte mit dem neuen System abgeholfen werden.

### **Die zentrale Datenbank: DACHS-A**

Im Herbst 1997 schlossen wir mit den Wissenschaftshistorischen Sammlungen der ETH Bibliothek eine Vereinbarung zur gemeinsamen Entwicklung oder Beschaffung eines relationalen und skalierbaren Datenbanksystems, das mit einem probabilistischen Information Retrieval System der aus der ETH hervorgegangenen Firma Eurospider gekoppelt werden sollte, da die schnelle und vor allem effiziente Volltextsuche nicht die Stärke von relationalen Datenbanken ist.

Zunächst zur Datenbank: Sie soll bewusst nicht die Dokumente selbst, sondern deren Beschreibungen, also Metatexte, und zwar in aller Regel

Beschreibungen von Dokumentengruppen (Dossiers) enthalten. Es sind nicht nur technische und finanzielle Gründe, die uns am Konzept «Findmittel» festhalten lassen, allerdings an dem eines beständeübergreifenden elektronischen Findmittels (das ausgedruckte Findbuch eines Bestands ist damit keineswegs obsolet geworden!). Damit ist nun gar nichts gegen die Digitalisierung von Dokumenten und deren Veröffentlichung im Internet gesagt (natürlich im Rahmen dessen, was der Datenschutz zulässt), aber die Digitalisierung wird die Datenbank der Metatexte nicht ersetzen können – auch dann nicht, wenn die schiere Menge der Dokumente technisch keine Rolle mehr spielen sollte. Sie kommt, wo sie mehr bieten soll als eine bloße Illustration oder eine Sammlung von aufbereiteten Leckerbissen für ein breites Publikum, vor allem für klar definierte Dokumententypen in Frage: für Dokumente, deren Kontext durch ihre Bezeichnung weitgehend gegeben ist und nicht erst gebildet werden muss (z.B. Protokolle und Geschäftsberichte). Es ist Aufgabe der intellektuellen Erschliessung, diesen Kontext dort, wo er fehlt, durch Bildung von Klassifikationen und Gruppierung von Dokumenten zu rekonstruieren. Die Klarheit einer Struktur und die Kürze und Prägnanz der Dossier-Beschreibung wird jeder zu schätzen wissen, der sich zunächst orientieren möchte, um dann gezielt zu den einzelnen Dokumenten vorzustossen.

Da die eingeführten kommerziellen Lösungen kaum die Bedürfnisse von Privatarchive mit umfangreichen Nachlassbeständen berücksichtigen und insbesondere mit mehrstufigen Klassifikationen nur unbefriedigend umgehen, wählten wir eine neu auf den Markt gekommene Archivsoftware, die sich für Nachlassarchive ebenso eignet wie für Ablieferungsarchive: DACHS-A 3.0 der Berliner Firma DISOS GmbH, einer IBM-Tochter. Im Rahmen einer Vereinbarung mit der Herstellerfirma konnte sich unsere Arbeitsgruppe, bestehend aus Rudolf Mumenthaler (Leiter Sonder-sammlungen der ETH-Bibliothek), Erich Scherer (IT-Dienste der ETH-Bibliothek) und Thomas Ehram (IT-Verantwortlicher des AfZ), an der Weiterentwicklung von DACHS-A beteiligen.

Das System arbeitet in einer Client-Server-Struktur und basiert auf einer Oracle-Datenbank. Die Navigation erfolgt über eine sehr benutzerfreundliche, am Windows-Explorer angelehnte, grafische Oberfläche. Damit ist DACHS-A insbesondere in der Lage, mehrstufige Klassifikationen elegant abzubilden und zu verwalten. Unsere konzeptionelle Arbeit bei der Weiterentwicklung betraf hauptsächlich die bis anhin noch fehlende Stammdatenverwaltung, die Erfassung von Mikrofilmen, die weitere Ausgestaltung der Klassifikation und die Abklärung der Kompatibilität mit dem Standard ISAD(G). Besonders das Stammregister war eine spezifische Forderung



unsererseits. Wir richteten uns bei unserer Lösung nach dem Standard US-MARC, wodurch wir eine Übereinstimmung mit dem in Aleph500 zur Anwendung kommenden «Alphabetischen Formalregister» der ETH-Bibliothek erreichten.

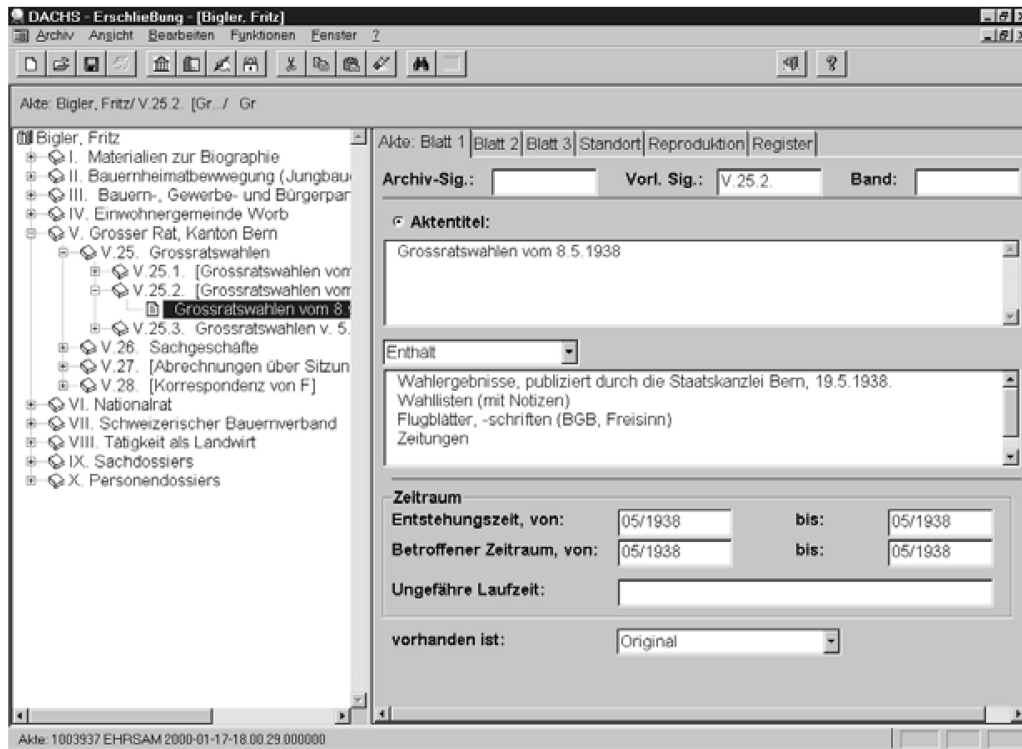


Abb. 1: Die Maske für die Aktenerschliessung in DACHS-A: links die Klassifikation, rechts ein Aktendossier (sichtbar ist das erste von fünf Erfassungsblättern).

Die nun vorliegende Version 3.1 umfasst ein Erfassungs-, ein Benutzungs-, ein Auswertungs-, ein Administrations- und ein Import-Exportmodul. DACHS-A kann damit eingesetzt werden:

- für die Verzeichnung von Akten, Dokumenten, Bildern etc., wobei für die verschiedenen Dokumenttypen je eigene Masken zur Verfügung stehen (bis hin zur Möglichkeit, Urkunden mitsamt Volltext, Übersetzung, Dorsualvermerk und was der schönen Dinge mehr sind, zu erfassen),
- für die Erstellung von bestandsbezogenen Findbüchern und die Etikettierung der Bestände,
- für die Abfrage und
- für die Verwaltung von Akzessionen und Beständen.

Für eine spätere Version sind Module für den Benutzerdienst und die Ausleihe, bzw. die Lagerbewirtschaftung vorgesehen, zudem ist ein Web-Client in Arbeit. Die Abfrage kann beständeübergreifend nach den ver-

schiedensten Kriterien erfolgen, mit dem Oracle Context-Server können zudem Volltextabfragen mit Booleschen Operatoren durchgeführt werden.

DACHS-A steht im AfZ seit Anfang Dezember im produktiven Einsatz. Seit diesem Zeitpunkt werden neue Bestände direkt in der Datenbank ohne Umweg über eine Textverarbeitung erfasst. Die Migration der alten Daten aus DataPerfect in die neue Datenbank ist momentan in Arbeit und sollte bis auf die Mikrofilmbestände Anfang Februar abgeschlossen sein.

### **Information Retrieval: Eurospider**

Der Volltextabfrage wurde von Anfang an ein hoher Stellenwert gegeben, da sich die Beschlagwortung (auf die nicht verzichtet werden soll) auf Namen, Orte und Körperschaften wird beschränken müssen – der Aufbau eines Thesaurus für Sachbegriffe würde die Kapazitäten der beteiligten Institutionen sprengen. Die herkömmliche (nun auch von Oracle angebotene) boolesche Abfrage hat sich insofern als beschränkt erwiesen, als man mit ihr allzu häufig viel zu viele oder gar keine Treffer erhält, wobei zentrale und unwichtige Datensätze wahllos beieinander stehen. Ein probabilistisches Information Retrieval System wie Eurospider normiert bei der Indexierung sowohl Dokument wie auch Abfragetext (aus «Häuser» wird «Haus») und bewertet die gefundenen Dokumente nach ihrer voraussichtlichen, mittels Wahrscheinlichkeitsrechnung ermittelten Relevanz für die Abfrage. Diese bestimmt auch die Reihenfolge, in der die gefundenen Akten dargestellt werden. Dadurch erreicht ein solches System eine weit höhere und genauere Trefferquote als die herkömmliche Volltextsuche: es findet mehr relevante Dokumente und weniger «Lärm». Es genügt nun allerdings nicht, die in der Datenbank gespeicherten Texte von einem Information Retrieval System indexieren und abfragen zu lassen. Die Verbindung von Datenbank und Information Retrieval erreicht erst dann ihr Ziel, wenn sie für den Benutzer zur Einheit wird, d. h. wenn die jeweiligen Recherchemöglichkeiten kombiniert werden können: Volltext- und Schlagwortabfrage (letztere für Namen, Orte und Körperschaften) sollen gekoppelt und bei Bedarf mit Laufzeitangaben verbunden werden können. A propos Laufzeiten: hier unterscheiden wir zwischen den Entstehungszeiten und den betroffenen Zeiträumen, da diese gerade bei privaten Akten mit vielen Erinnerungsschriften oft nicht übereinstimmen. Aber auch eine Fragestellung der Art: «Wie wurde 1995 die Geschichte der Schweiz im Zweiten Weltkrieg rezipiert?» setzt diese Trennung voraus.

Da jeder Internet-Browser Frontend von Eurospider sein kann, bedeutet der Einsatz dieser Software auch, dass die Datenbank ohne weiteren Auf-

wand auf dem Netz angeboten werden kann – wobei selbstverständlich nur die freigegebenen Daten von Eurospider erfasst und abgesucht werden. Die Implementierung von Eurospider ist für die erste Hälfte 2000 geplant.

### **Andere Datenbanken**

Der Entscheid für eine zentrale Datenbank kann bei den vielfältigen informationstechnischen Aufgaben eines Archivs heute nicht mehr bedeuten, dass daneben keine anderen Datenbanksysteme verwendet werden. So schön es wäre, mit einem System auszukommen, so unsinnig wäre es auch, das zentrale System für alle Einzelbedürfnisse fit machen zu wollen. Das würde nicht nur bedeuten, mit Kanonen auf Spatzen zu schiessen, es würde auch heissen, projektbezogene Aufgaben mit den langfristigen zu vermengen oder die zentrale Anwendung für die Bedürfnisse einer kleinen und spezialisierten Benutzergruppe mit Features zu belasten, die im Rahmen einer Archiv-Datenbank wenig Sinn machen. Ein Beispiel für den letztgenannten Fall ist die Bilddatenbank.

Obwohl sich Bilder im DACHS-A auch erfassen lassen (die nächste Version wird voraussichtlich auch binäre Dateien aufnehmen können), streben wir, wiederum in Zusammenarbeit mit den Wissenschaftshistorischen Sammlungen der ETH-Bibliothek und einigen anderen Institutionen auf dem Platz Zürich, den Einsatz einer eigenen Bilddatenbank (Image-Finder) an, da Bilder einen anderen Interessentenkreis ansprechen als Texte: Grafiker, Ausstellungsmacher, Verlage, Zeitungen, elektronische Medien etc. Da sich hier auch die Möglichkeit der Vermarktung und damit von Einnahmen öffnet – in Zeiten der zunehmenden Eigenfinanzierung hochwillkommen –, lohnt sich ein spezieller Aufwand. Die Bilddatenbank soll den Bedürfnissen und Gewohnheiten dieser Benutzer u. a. dadurch gerecht werden, dass die in Frage kommenden Bilder auf einen virtuellen Leuchttisch geschoben und dort arrangiert, beurteilt und schliesslich ausgewählt und direkt elektronisch bestellt werden können. Damit sich die Benutzer nicht an jede Institution einzeln wenden müssen, um nach Bildern zu einem Thema zu fahnden, soll in der ETH-Bibliothek eine gemeinsame Bilddatenbank aufgebaut werden, in die die beteiligten Institutionen ihre Bilder jeweils einspeisen. Auch Archive mit relativ kleinen Bildbeständen sollen damit eine Chance erhalten, von potentiellen Kunden wahrgenommen zu werden.

FileMaker Pro - [FI-Flüchtlinge.fp3]

Datei Bearbeiten Modus Auswahl Format Script Fenster Hilfe

assung

**FLÜCHTLINGE** Erfassung

Namenliste

**Personalien**

Signatur VSJF Doeblin, Hugo ID\_Nr 3196

Name Doeblin Signatur bilden Gibt es Doppel?

Vorname Hugo Emil Titel

Mädchenname Geschl. m Zivilstand geschieden

Geburtsdatum 29. 10.1876  genau  nur Monat u. Jahr  nur Jahr  ungeprüft

Konfession jüdisch / israelit. / mos.

Ehepartner

Geburtsort Stettin

Staatsbürgerschaft I Staatenlos Bemerk. Nach 1950 wieder deutscher Pass  
II Deutsches Reich (zur Staatsbürgerschaft)

Nat. Identität

Gestorben am 5. 11.1960  genau  nur Monat u. Jahr  nur Jahr  ungeprüft

in Zürich

Beruf früher Schauspieler

in der Schweiz

Einreise Ort Kanton

am 1. 4.1938  genau  nur Monat u. Jahr  nur Jahr  ungeprüft

von Wien / Oesterreich

Grund

Ausreise  ausgereist  ausgewiesen  Verbleib unbekannt

am  genau  nur Monat u. Jahr  nur Jahr  ungeprüft

von nach

Blättern

Abb. 2: Die Eingabemaske der Flüchtlingsdatenbank (VSJF) auf Filemaker (Ausschnitt)

Für projektbezogene Datenbanken haben wir als Nachfolgeprodukt von DataPerfect Filemaker gewählt. Zur Erschliessung der ca. 14'000 Flüchtlingsdossiers des Verbands Schweizerischer Jüdischer Fürsorgen (VSJF) wurde eine eigene Forschungsdatenbank mit über 100 Feldern aufgebaut, die nicht nur das Auffinden der einzelnen Dossiers, sondern v. a. auch Forschungen zur schweizerischen Flüchtlingspolitik, zur Emigration und Migration, zu verschiedenen Flüchtlingsbewegungen etc. unterstützen und damit den immensen und in Europa einmaligen Bestand für qualitative und quantitative Fragestellungen zugänglich machen soll. Neben ausführlichen Personalien mit Angabe der Verwandten enthält die VSJF-Datenbank

Informationen zu Ein- und Ausreise, zum Fluchtweg, zu Lageraufenthalten im Ausland und in der Schweiz, zum rechtlichen Status und schliesslich Hinweise auf besonders reich dokumentierte Materialien. Dank dieser Hinweise kann man sich beispielsweise die Dossiers anzeigen lassen, in denen sich aussagekräftige Unterlagen zu Stichworten wie Unterstützung, Unterbringung, Schulung, Aus- und Weiterreise, Vermögenslage oder -verluste, Wiedergutmachung u.a. finden – bei der unabsehbaren Menge der Akten zweifellos eine wichtige Hilfe für Forscher und Forscherinnen.