Zeitschrift: Arbido

Herausgeber: Verein Schweizerischer Archivarinnen und Archivare; Bibliothek Information Schweiz

Band: 18 (2003)

Heft: 3

Artikel: Sicherung und Bewertung von Datenbanken im Schweizerischen Bundesarchiv

Autor: Zürcher Thrier, Thomas

DOI: https://doi.org/10.5169/seals-769890

Nutzungsbedingungen

Die ETH-Bibliothek ist die Anbieterin der digitalisierten Zeitschriften auf E-Periodica. Sie besitzt keine Urheberrechte an den Zeitschriften und ist nicht verantwortlich für deren Inhalte. Die Rechte liegen in der Regel bei den Herausgebern beziehungsweise den externen Rechteinhabern. Das Veröffentlichen von Bildern in Print- und Online-Publikationen sowie auf Social Media-Kanälen oder Webseiten ist nur mit vorheriger Genehmigung der Rechteinhaber erlaubt. Mehr erfahren

Conditions d'utilisation

L'ETH Library est le fournisseur des revues numérisées. Elle ne détient aucun droit d'auteur sur les revues et n'est pas responsable de leur contenu. En règle générale, les droits sont détenus par les éditeurs ou les détenteurs de droits externes. La reproduction d'images dans des publications imprimées ou en ligne ainsi que sur des canaux de médias sociaux ou des sites web n'est autorisée qu'avec l'accord préalable des détenteurs des droits. En savoir plus

Terms of use

The ETH Library is the provider of the digitised journals. It does not own any copyrights to the journals and is not responsible for their content. The rights usually lie with the publishers or the external rights holders. Publishing images in print and online publications, as well as on social media channels or websites, is only permitted with the prior consent of the rights holders. Find out more

Download PDF: 14.07.2025

ETH-Bibliothek Zürich, E-Periodica, https://www.e-periodica.ch

Sicherung und Bewertung von Datenbanken im Schweizerischen Bundesarchiv

Thomas Zürcher Thrier
Schweizerisches Bundesarchiv
Wissenschaftlicher Mitarbeiter der Fachstelle ARELDA

Datenbanken in der Bundesverwaltung

Wann ist Herr X zum Oberleutnant befördert worden? Wie viel verdiente der Beamte Y 1991? Hatte Z tatsächlich ein Gesuch um Einsicht in seine Staatsschutzfiche gestellt?

Datenbanken sind in der Bundesverwaltung allgegenwärtig. Sie bilden Arbeitshilfen (Adressverwaltung etc.) und werden für statistische Zwecke gebraucht, sie werden aber auch immer öfter zur Kontrolle von Geschäftsvorgängen verwendet. Datenbanken bilden die Grundlage von Dokumentenverwaltungssystemen und sie stehen hinter interaktiven Webanwendungen. Die oben genannten und zahllose weitere ähnliche Fragen lassen sich anhand der Datenbanken beantworten, die im Magazin des Schweizerischen Bundesarchivs lagern.

Das Projektteam ARELDA erarbeitet gegenwärtig Metadatenmodell und Datenkatalog zur vorarchivischen Beschreibung von Datenbanken. Die Beschreibung soll garantieren, dass die in den Datenbanken enthaltenen Informationen auch nach einem halben Jahrhundert nicht nur verfügbar, sondern auch verstehbar sind. Dazu müssen nicht nur die Daten selbst übernommen, sondern auch möglichst vollständige Angaben zur jeweiligen Datenbank (Metadaten) erfasst werden. Dies ist nur möglich, wenn das Archiv schon zu Beginn des Lebenszyklus der Datenbanken auf die Einhaltung der archivischen Datenqualität drängt. Nur was von Anfang an definiert und während dem Betrieb erfasst wird, lässt sich später als Metadaten übernehmen.

Die Systeme müssen so konzipiert werden, dass Geschäftsabläufe auch später rekonstruierbar sind. So sollten die Daten bei der Aktualisierung beispielsweise nicht überschrieben werden, sondern es sollte ein neuer Datensatz angelegt werden (Historisierung). Auch muss das System mittels so genannter Log-Files dokumentieren, wer wann welche Daten verändert hat. Auch die Schnittstellen zum Archiv müssen im Voraus definiert werden, um die Übernahme von Daten und Metadaten mit vertretbarem Aufwand zu ermöglichen. Demgemäss sind die eidgenössischen Behörden dazu angehalten, das Bundesarchiv bereits bei der Planung neuer digitaler Informationssysteme beizuziehen[1].

Bewertung von Datenbanken: eine Typologie von Informationssystemen

Schon seit geraumer Zeit wird im Bundesarchiv das Prinzip der prospektiven Bewertung sukzessive umgesetzt. Danach werden die Geschäftsunterlagen nicht erst bewertet, wenn sie zur Ablieferung angeboten werden, sondern es werden bereits die Registraturplanpositionen im Hinblick auf die spätere Aktenproduktion bewertet. Das Prinzip der prospektiven Bewertung ist besonders im Hinblick auf die digitalen Geschäftsunterlagen unabdingbar. Der grosse organisatorische und finanzielle Aufwand, der mit der Sicherung und dem Archivieren von Datenbanken verbunden ist, lässt sich nur dann rechtfertigen, wenn möglichst schon zum Zeitpunkt der Konzeption einer Datenbank feststeht, ob sie später einmal archivwürdige Daten enthalten wird oder nicht.

Datenbanken bestehen aus zwei Teilen: aus einer Datensammlung zu einem bestimmten Thema und aus einem Applikationsteil. Unter Daten werden in diesem Zusammenhang strukturierte und für die Verarbeitbarkeit in Computern codierte Informationen verstanden. Der Applikationsteil, das so genannte Datenbankmanagementsystem, erleichtert die Datenerfassung und ermöglicht einen effizienten Zugang nach unterschiedlichen Abfragekriterien.

Unterteilt werden Datenbanken oft nach technischen Kriterien (relationale Datenbanken, objektorientierte Datenbanken etc.) oder nach dem Inhalt: z.B. Personendatenbank, geografische Informationssysteme etc. Um Datenbanken aus archivischer Sicht zu bewerten, ist es aber empfehlenswert, sie als Informationssysteme zu verstehen, die zu einem bestimmten Zweck entwickelt wurden. Demgemäss haben die Fachstelle Bewertung des Bundesarchivs und ARELDA gemeinsam eine Typologie entwickelt, die auf folgenden Kriterien basiert:

- *Geschäftsrelevanz:* Hat das System primär den Zweck, Geschäftsvorgänge zu belegen oder dient es in erster Linie als Hilfssammlung?
- *Originalität und Herkunft der Daten:* Handelt es sich um Daten, die ausschliesslich vom Amt selbst oder in dessen Auftrag erfasst wurden, oder übernimmt das System vorwiegend Fremddaten?
- *Dokumentenbezug:* Werden im System ganze Dokumente verwaltet (Textdateien, Bilder, Web-Seiten) oder enthält es vorwiegend homogen strukturierte Datenfelder (Codes in Form von Zahlen, einzelnen Wörtern oder Abkürzungen)?

Aus diesen drei Hauptkriterien lassen sich eine Reihe von Datenbanktypen ableiten:

RMS (Records Management Systeme oder Geschäftsverwaltungssysteme): Sie dienen zur Geschäftskontrolle. Ein ins System integrierter Registraturplan und eine Workflow-Komponente ermöglichen die lückenlose Dokumentation zu Geschäftsabläufen. In der Bundesverwaltung sind Systeme im Einsatz, die lediglich die Metadaten der Geschäftsdossiers verwalten. Einzelne Ämter führen aber auch RMS mit vollständigen Dokumentmanagementfunktionen. Geschäftsverwaltungssysteme sind wegen ihres hohen Evidenzwertes archivwürdig[2].

DMS (Dokument Management Systeme): Auch Dokumentmanagementsysteme ohne RMS-Funktion sind in der Bundesverwaltung im Einsatz. Sie erlauben keine explizite Dokumentation der Geschäftsab-

[1] Weisungen über die Anbietepflicht und die Ablieferung von Unterlagen an das Schweizerische Bundesarchiv vom 28. September 1999, Art. 4, Abs. 3.

[2] Zur Unterscheidung von Evidenz- und Informationswert siehe: Schellenberg, Theodore R.: Die Bewertung modernen Verwaltungsschriftguts. Übers. und hrsg. von Angelika Menne-Haritz. Marburg: Archivschule, 1990 (= Veröffentlichungen der Archivschule Marburg; Nr. 17).

läufe. Sie haben nur dann einen hohen Evidenzwert, wenn die Dokumente gemäss dem Registraturplan abgelegt wurden. Sonst ist die Archivwürdigkeit dieser Systeme fraglich. Sie können allenfalls als Findmittel betrachtet werden, die ein möglichst effizientes Auffinden von bestimmten Dokumenten erlauben (Beschlagwortung).

Erhebungs- und Messdatenbanken: Zu diesem Typ gehören die ersten Datenbanken, die dem Bundesarchiv zur Archivierung angeboten wurden. Sie dienen zur Erfassung von mess- oder zählbaren Vorgängen, sei es zur späteren Publikation, sei es für eine spätere interne Auswertung. Ihr Evidenzwert ist in der Regel gering, dagegen ist ihr Informationswert gerade wegen den verschiedenen Zugängen, welche Datenbanken im Gegensatz zu gedruckten Statistiken erlauben, für die sozialhistorische Forschung hoch[3]. Hier entscheidet die Frage nach der Originalität der Daten über ihre Archivwürdigkeit. Für diesen Datenbanktyp hat das Schweizerische Bundesarchiv einen entsprechenden departementsübergreifenden Grundsatzentscheid gefällt[4].

Hilfssammlungen: Zahllos sind die kleinen und mittleren Datenbanken mit arbeitserleichternder Funktion (Inventare, Adressverzeichnisse). Sie sind meist nicht archivwürdig. Eine Ausnahme bilden solche Datenbanken, die Dossiernummern enthalten und damit als Findmittel dienen können. Unter Umständen kommt solchen Datenbanken sogar grosser Evidenzwert zu, weil sie die Existenz von Dossiers nachweisen, die in der Zwischenzeit «verschwunden» sind.

Datenbanken und ihre Dokumentation

Damit sie auch in ferner Zukunft noch verstehbar sind, müssen zu archivierende Datenbanken sorgfältig dokumentiert werden. Neben den auch für Papierakten üblichen Angaben zur Provenienz oder zum Umfang der Ablieferung sind für Datenbanken spezifische Metadaten notwendig. Vor allem vollständige Angaben zur Datenstruktur und die Bedeutung der verwendeten Codes und Abkürzungen müssen minutiös festgehalten werden. Auch ist darauf zu achten, dass die Felddefinitionen und die Struktur der Datenbank im Lauf der Zeit verändert werden. Unterschiedliche Ablieferungen zu derselben Datenbank können daher nicht mit Sicherheit zusammengeführt werden.

[3] Bütikofer, Niklaus: Archivierung von statistischen Daten. In: Studien und Quellen – Zeitschrift des Schweizerischen Bundesarchivs 1990–1991, 16/17: S.263–293.

[4] Einzusehen ab März 2003 auf *http://www.bundesarchiv.ch/*.

ARELDA: Varianten der Erschliessungsvorschläge

Variante	Teilbestand / Dossier		Dossier / Subdossier	Dokument(e)
1	**Datenbank**			*Tabellen*
2	**Datenbank**		*Menge(n) von Datensätzen*	*Datensatz*
3	**Datenbank**	*Menge(n) von Datensätzen*	*Datensatz*	*Datenfeld*

Tab. 1: Eingliederung von Datenbanken in die übrige Bestandestektonik des Bundesarchivs gemäss Vorschlag ARELDA:
Datensatz = Menge der dasselbe Geschäft (z.B. eine Person) betreffenden Zeilen aus einer oder mehreren Tabellen
Datenfeld = Einzelfeld aus einer Tabelle (z.B. Einzelbrief)
fett: Erschliessung in erster Priorität, *kursiv:* Erschliessung in zweiter Priorität

Diese Grundsätze entsprechen allgemeiner Praxis. Das britische National Digital Archive of Datasets (NDAD) dokumentiert beispielsweise seine Datenbanken nach einer für Datenbanken entwickelten Adaptation von ISAD(G)[5]. Die Beschreibungen sind online zugänglich und gegen Bezahlung können die Daten heruntergeladen werden[6]. Auch das Data-Archiv der Universität von Essex, das über 5000 Datenbanken im Bereich Sozial- und Geisteswissenschaften beherbergt, dokumentiert die Datenbanken beispielhaft[7]. Als Grundlage dient hier ein von der Data Description Initiative (DDI) speziell für politik- und sozialwissenschaftlich relevante Daten entwickelter Beschreibungsstandard, der in XML codiert ist. Die dazugehörige DTD (Document Typ Definition) basiert auf dem Dublin Core und enthält über 200 Beschreibungselemente[8].

Diese bisherige Beschreibungspraxis für Datenbanken hat zwei wesentliche Nachteile. Erstens handelt es sich um Archive, die ausschliesslich Datenbanken beherbergen. Demgemäss werden die Datenbanken alle auf derselben Beschreibungsstufe angeordnet und nicht in andere Archivbestände integriert. Zweitens sind die bisherigen Ansätze primär auf Erhebungsdatenbanken ausgerichtet. So enthalten die Dokumentationen detaillierte Angaben zu den Erhebungen, die zur Datensammlung geführt haben (Instruktionen für die Datenerhebung, einzelne Erhebungsbogen etc.). Für die Erschliessung von Geschäftsverwaltungssystemen wäre hingegen die Benutzungsgeschichte von weit grösserer Bedeutung. Für die Rekonstruktion von Geschäftsabläufen ist es beispielsweise aufschlussreich zu wissen, welcher Angestellte welche Daten wie eingeben oder konsultieren konnte.

[5] Shepherd, Elizabeth und Smith, Charlotte: The application of ISAD(G) to the description of archival datasets. In: Journal of the Society of Archivists 2000, 21, (1): S.55–86.

[6] *http://ndad.ulcc.ac.uk/* [kons. 5.2.2003].

[7] Musgrave, Simon und Winstanley, Bridget: Archiving Statistical Data: The Data Archive at the University of Essex. In: RLG DigiNews 1998, 2, (6), *http://www.rlg.org/preserv/diginews/diginews2-6.html – essex* [kons. 5.2.2003].

[8] *http://www.icpsr.umich.edu/DDI/CODEBOOK/index.html* [kons. 5.2.2003].

Mehrstufige Erfassung von Datenbanken nach Geschäftsrelevanz

Ein Teil der vorarchivisch erfassten Metadaten wird bei der Übernahme der Datenbank vom archivischen Findsystem übernommen. Der primäre intellektuelle Zugang zum Archivgut ist nur über die Rekonstruktion des Entstehungszusammenhangs möglich, die Bestandestektonik bietet den Schlüssel dazu. Betrachtet man Datenbanken nicht primär unter technischem Aspekt, sondern unter ihrem Verwendungszweck, so sind sie in diese Tektonik adäquat einzugliedern. Eine Datenbank als mehrgliedriges Gebilde sollte auch mehrstufig beschrieben werden. Die Erschliessungsvorschläge von ARELDA sehen die in der obigen Tabelle angezeigten Varianten vor.

Die erste Variante ist ausschliesslich für einfache Erhebungs- und Messdatenbanken vorgesehen, die nur aus einer oder ganz wenigen Tabellen bestehen. Moderne relationale Datenbanken bestehen aus zahllosen Einzeltabellen, die hauptsächlich Referenzwerte auf andere Tabellen enthalten. Diese sind rein technische Gebilde ohne Geschäftsrelevanz.

Die zweite Variante sieht eine Gruppierung von Datensätzen, die dasselbe Geschäft (zum Beispiel dieselbe Person) betreffen, vor.

Die dritte Variante ist für Datenbanken mit Dokumenten geeignet. Die einzelnen Datenfelder entsprechen den im System gespeicherten Dokumenten. Sie werden mit (automatisch generierten) Angaben wie Dokumententitel, Autor, Empfänger etc. erfasst.

Ob die Datenbank als Teilbestand oder Dossier angesetzt wird, hängt von ihrer Stellung beim Aktenbildner ab. Falls eine Datenbank ein Sachgebiet abdeckt, das in die Kompetenz des ganzen Amtes fällt, wird die Beschreibung der Datenbank möglichst hoch angesetzt. In der Regel erhält die Datenbank in diesem Fall keine Position im

Registraturplan. Sie wird gewissermassen als eigene «Ablage» mit eigenem Ordnungssystem betrachtet und müsste folglich im Archiv als Teilbestand angesetzt werden. Falls die Datenbank hingegen nur einem Teilgebiet des Aktenbildners zuordenbar ist (zum Beispiel dem Personalwesen), wird die Beschreibung auf Dossierebene angesetzt. Idealerweise erhält die Datenbank in diesem Fall auch eine entsprechende Position im Registraturplan.

Zukunftsmusik

Auch wenn möglichst viele Metadaten automatisch übernommen werden: die koordinierende und ergänzende Hand des Archivars oder der Archivarin bleibt unersetzlich. Neben der Bestandeskontrolle und der Vereinheitlichung von Rechtschreibung, Namensgebungen und Abkürzungen muss das Archiv zusätzliche, sachorientierte Findmittel entwickeln. Sie sollen jenen BenutzerInnen, die mit den Eigenheiten der Verwaltung nicht vertraut sind oder die nicht die Zeit haben, sich in umfangreiche Dokumentationen einzuarbeiten, einen bestandesübergreifenden Zugang zum Archivgut ermöglichen. Auch solche Deskriptoren können aus Datenbanken extrahiert werden. Hier bieten sich vor allem die Findmittel-Personendatenbanken an, die neben Personennamen Dossiernummern enthalten. Die Extraktion solcher Indices bildet keinen grossen Aufwand, diese aber so zu vereinheitlichen, dass sie auch eine bestandesübergreifende Recherche ermöglichen, hingegen schon. Die BenutzerInnen müssen jedenfalls sicher sein, dass sie restlos alle Daten zu ihrem Peter Muster erhalten und dass die Einträge zu unterschiedlichen Personen mit demselben Namen unterscheidbar sind.

Vorläufig finden die meisten Konsultationen des Archivgutes im Lesesaal des Bundesarchivs statt. Es ist vorgesehen, mit der Zeit ausgewählte Bestände auch über ein Internetportal zugänglich zu machen[9]. Erschliessung als «Versuch, archivierte Unterlagen und die in ihnen enthaltenen Informationen für eine vorhergesehene wie auch für eine unvorhergesehene Nutzung zugänglich zu machen»[10], muss auch diese Benutzungsvariante vorwegnehmen oder zumindest nicht verbauen. Gerade die sozialwissenschaftlich relevanten Erhebungsdaten könnten für ein breites internationales Publikum von Interesse sein. Die Webbesucherin der Zukunft wird aber nicht einen virtuellen Archivlesesaal nach dem anderen abklappern, um zu den gesuchten Informationen zu gelangen. Sie wird ein möglichst zentrales Portal konsultieren, das fähig ist, Informationen aus unterschiedlichsten Archiven simultan zu konsultieren. Ein solches Portal unterhält beispielsweise das EU-Network European Social Science Tools and Resources (NESSTAR), das den DDI-Standard anwendet[11]. Eine ähnliche Funktion könnte die Encoded Archival Description (EAD) für geschäftsrelevante Daten übernehmen[12]. Archive tun jedenfalls gut daran, solche standardisierte Metadatensets zu berücksichtigen. ■

contact:
E-Mail:
thomas.zuercher@bar.admin.ch

[9] Chiquet, Simone et al.: Zugang öffnen: Die Vermittlung im Schweizerischen Bundesarchiv. In: Studien und Quellen – Zeitschrift des Schweizerischen Bundesarchivs 2002, 28: S.351–376.

[10] Förster, Bärbel und Kellerhals, Andreas: Das Erschliessungskonzept des Schweizerischen Bundesarchivs: Vom Findmittel zum Findsystem. In: Studien und Quellen – Zeitschrift des Schweizerischen Bundesarchivs 1997, 23: S.335–353, S.335.

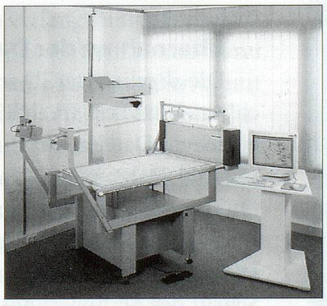
[11] *http://www.nesstar.org/* [kons. 5.2.2003].

[12] *http://lcweb.loc.gov/ead/* [kons. 5.2.2003].

Anzeigen

SMA 6650:
Scanner für grosse Aufgaben

Der SMA 6650 kann alles was der Scanfox kann: Scannen, Kopieren und Archivieren.
Genauso bedienerfreundlich und genauso unkompliziert.
Seine wahre Größe spielt er bei Vorlagen bis DIN A1 (SMA 1 6650) oder bis DIN A 0 (SMA 0 6650) aus.
Wie beim ScanFox wird die Vorlage normal im Sitzen eingelegt und mit einem Fußschalter die Anpress-Automatik ausgeführt.
Durch den Höhenausgleich der Buchwippe können sogar Bücher bis 20 cm Dicke verarbeitet werden.
Drei Leseköpfe je 5000 Pixel sorgen für höchste Scanqualität.
Der Lesekopf scannt in Farbe, Graustufen oder in Schwarzweiss.
Optional erhalten Sie einen Aufsatz durch den das Gerät mit wenigen Handgriffen in einen Durchlauf-Scanner umfunktioniert werden kann.



Bitte fordern Sie weitere Informationen an.

TECNOCOR ACC AG
ARCHIVING COMPETENCE CENTER
CH-6030 Ebikon, Luzernerstrasse 28 | Telefon +41-(0)41-440 74 22, Telefax +41-(0)41-440 85 84 | E-mail: info@tecnocor.ch, http://www.tecnocor.ch

Wir drei Mikrofilmer treiben Ihre Dokumente in die Enge.

BÜRGER SPITAL BASEL MIKROGRAFIE
Holbeinstrasse 58
4051 Basel
mikrografie@buespi.ch
Tel. 061 276 98 80
Fax 061 276 98 81

dreischiibe
Rosengartenstr. 3
9006 St.Gallen
info@dreischiibe.ch
Tel. 071 243 58 41
Fax 071 243 58 90

STIFTUNG BEHINDERTENBETRIEBE URI
Mikrografie
Rüttistrasse 57
6467 Schattdorf
behindertenbetriebe.uri@sburi.ch
Tel. 041 874 15 15
Fax 041 874 15 99