



Erschließung strukturierter Massendaten aus Datenbanken

Was ist nötig, um solche Daten interpretierbar und benutzbar zu machen?

Sigrid Schieber

Inhalt

I. Hintergrund

II. Annahmen

III. Gruppen von Metainformationen

IV. Differenzierung in klassische Erschließung und Binnenerschließung

V. offene Fragen

I. Hintergrund

- Digitales Archiv Hessen: Seit 2010 Übernahme von Daten aus 17 Fachverfahren
 - Archivierung von Datenexporten, nicht der Anwendung.
 - Übernahme von definierten Teilmengen der Daten nach Feinbewertung des jeweiligen Fachverfahrens
- Dabei war klar: Metainformationen sind wichtig ...
- ... und Hauptsache, die nötigen Metainformationen werden überhaupt gesichert, im Zweifelsfall lieber mehr davon als weniger.
- 2017: Phase der Reflektion und Systematisierung

II. Annahmen

- Datenbanken enthalten häufig nicht explizit alle Metainformationen, die zum Verständnis der in ihnen verwalteten Informationen nötig sind.
- Bei der Erzeugung eines Datenexports können nicht immer alle vorhandenen Metainformationen mit übernommen werden.
- Durch die Bewertung und die Erzeugung eines Datenexportes wird ein neues Informationsobjekt geschaffen, das nicht selbsterklärend ist.

Beispiel- Fachverfahren (dramatisiert)

Werteliste 1 | 2 | 3 | 4
(Klartext zu den
Werten ist in der
Anwendung fest
codiert.)

Geburtsname wird
als Bemerkungsfeld
zweckentfremdet
(Information bei der
Schnittstellenum-
setzung)

Datenexport

Nachname; Geschwisterfolge; Geburtsname; DATUM

Müller; 2; ; 0

Maier; 1; Mittwochs ist nur die Großmutter zu erreichen; -1

Schmidt; 3; ; 1

Werteliste 0 | 1 | -1
(fehlerhafte Checkbox:
0 = „nein“
1 und -1 = „ja“)

Warum fehlt Person X?
(Bewertungsentscheidung: Es
wurden nur Daten öffentlicher
Schulen übernommen)

III. Gruppe 1: Informationen über das ursprüngliche Fachverfahren

- Was

- Zu welchem Zweck und von wem wurde es wie genutzt?
- Welche Informationen wurden darin gespeichert?

- Wo

- Dokumentationsmaterial zum Fachverfahren
(Anwenderhandbücher, Screenshots der Oberfläche, Datenbankmodelle ...)
- Aktenvermerke aus der Bewertungsphase

III. Gruppe 2: Informationen über die Bewertungsentscheidung

- Was
 - Welche Entscheidung wurde getroffen?
 - Warum?
- Wo
 - Aktenvermerk(e) zur Bewertungsentscheidung

III. Gruppe 3: Informationen über die Erstellung und die Struktur des Datenbankexports

■ Was

- Welche Felder aus welchen Tabellen wurden übernommen?
- Wurden dabei Daten aufbereitet? Wie?
- Welche Spezifikationen galten für die übernommenen Felder in der Herkunftsanwendung?
- Wie werden Datenschnitte voneinander abgegrenzt?
- Besonderheiten?

■ Wo

- Aktenvermerk + Feldtabelle zum Datenexport

IV. Was tun mit all diesen Informationen?

- Welche davon brauchen die späteren Benutzer/innen?
- Wo und wie sollen die Informationen zugänglich gemacht werden?
- Lösungsansatz aktuell: Differenzierung in
 - klassische Erschließung (im AFIS, nur Informationen, die für die Recherche sinnvoll sind)
 - Binnenerschließung (Wo? Was? Wie?)

IV. Klassische Erschließung

Arcinsys Navigator Suche | Merkliste | Beantragen | Nutzen | Übernehmen | Managen | Aufbewahren

Hessen

- Archive in Hessen
 - Hessisches Hauptstaatsarchiv
 - 651 Nassauischer Zentralstudienfonds (nach 1945)
 - Gliederung
 - Allgemeine Verwaltung
 - Finanzverwaltung
 - Domänenverwaltung
 - Mittelverwendung
 - Förderung
 - Stipendien**
 - Darlehen
 - Zuschüsse

HHStAW Bestand 651 Nr. 717 Neue Repräsentation Exportieren

Bearbeiten | Löschen | Verschieben | Drucken | Merken / Bestellen | Verlinken | Ver

Beschreibung - Repräsentationen - Eigenschaften

Beschreibung

Identifikation

<i>Titel</i>	Datenbank zur Förderung der Stipendiaten des Nassauischen Zentralstudienfonds
<i>Laufzeit</i>	1983-2013

Informationen / Notizen

<i>Zusatzinformationen</i>	Digitales Objekt Zur Aufbereitung und Beschreibung der Datenbank siehe HHStAW Abt. 651 Nr. 50000
<i>Interne Notizen</i>	Digitales Objekt in DIMAG: aID

Repräsentationen

Aktion	Typ	Bezeichnung	Zugang	Aufbewahrungsort	Info
Detailseite	Digitales Archivobjekt	R_1 Übernahmeformat (mdb)		Server DIMAG	
Detailseite	Digitales Archivobjekt	R_2 Archivformat (csv)		Server DIMAG	

IV. Binnenerschließung: „Aufbereitungsbericht“

- menschenlesbar, „Aufsatz“
- relativ unstrukturiert
- wird aktuell als Archivalie angelegt und kann von Nutzer/innen zusätzlich zum Datenexport bestellt werden
- kann durch vorarchivisches Dokumentationsmaterial zur Fachanwendung (z.B. Handbücher) ergänzt werden

IV. Binnenerschließung: „Aufbereitungsbericht“

- Enthält idealtypisch
 - Angaben zum Zweck der ursprünglichen Fachanwendung („Inhalt und Überlieferungsgeschichte“) = Gruppe 1
 - Angaben zur Bewertungsentscheidung (was, warum) = Gruppe 2
 - Angaben zum Übernahmeprozess (wer, wann, wie) ?
 - Nutzungshinweise ?

Hessisches Landesarchiv
Digitales Archiv



GZ 9.2.1.004-(0010)

Archivsignatur: HHStAW Abt. 651, Nr. 717

Nassauischer Zentralstudienfonds

Datenbank zur Förderung der Stipendiaten des Nassauischen Zentralstudienfonds
(1983-2013)

Aufbereitungsbericht

Inhalt und Überlieferungsgeschichte:

Der Nassauische Zentralstudienfonds (NZF) wurde am 29. März 1817 als öffentlich-rechtliche Stiftung mit eigener Rechtspersönlichkeit durch das Nassauische Schuledikt des Herzogs Wilhelm von Nassau gegründet¹. Das Stiftungsvermögen wird u.a. für die Vergabe von Stipendien verwendet. Antragsberechtigt sind ausschließlich Bewerberinnen und Bewerber, die im Gebiet des ehemaligen Herzogtums Nassau geboren sind. Sie sind „als ordentlich Studierende (Vollzeitstudium) an einer staatlichen oder staatlich anerkannten Universität, Hochschule, Fach- oder Kunsthochschule im Inland, einem Mitgliedsstaat der EU, des EWR oder der Schweiz immatrikuliert und haben zu Beginn des erstmals geförderten Studiums das 25. bzw. bei Masterstudiengängen das 35. Lebensjahr noch nicht vollendet. Die zum Studium berechtigende Prüfung wurde mindestens mit der Note 2,7 abgeschlossen. Gefördert werden können nur „nassauische Landeskinder“ mit begrenzten finanziellen Möglichkeiten.“²

Die Access-Datenbank zur Verwaltung der Daten der Stipendiaten wurde vom Nassauischen Zentralstudienfonds in den Jahren 1983 bis 2013 geführt. Sie bietet Grunddaten zu 2519 Personen und dazugehörige Kontaktdaten, Bankverbindungen, Angaben zu Studium und Einkommen sowie

IV. Binnenerschließung: „Feldtabelle“ (Gruppe 3)

- Beschreibt die Struktur des Datenexportes
 - technische Vorgaben (z.B. CSV mit spezieller CSV-Spezifikation)
 - listet tabellarisch folgende Angaben auf:
 - Feldname (im Export),
 - Name im Herkunftssystem (Feld und Tabelle),
 - inhaltliche Beschreibung
 - Feldtyp,
 - Feldlänge,
 - Codierung,
 - in der Datenbank vorhanden seit,
 - Pflichtfeld,
 - Besonderheiten
 - ...

Feldname Export	Fachliche Beschreibung	Tabellen- /Feldname LUSD	Feldtyp	Feldlänge	Codierung	Tabellenfeld in LUSD existiert seit	Besonderheiten
VerkehrsspracheFamilie	Sprache, die in der Familie des Schülers vorwiegend gesprochen wird	T021_022_Sprache Bezeichnung	Text	40			Kein Eintrag bedeutet i.d.R. deutsch, wurde v.a. befüllt, wenn die Verkehrssprache nicht deutsch war.
ZuzugsDatumBRD	Datum des Zuzugs nach Deutschland	T400_036_Zuzugsd atumBRD	Datum	10	dd.mm.yyyy		
IstVerstorben	Gibt an, ob ein Schüler verstorben ist	T400_049_IstVersto rben	Text	4	Ja Nein		Aussagekraft zweifelhaft . wurde oft zweckentfremdet, um versehentlich doppelt angelegte Datensätze zu Schüler/innen für die Sachbearbeiter/innen „unsichtbar“ zu machen. (Es gab keine Löschmöglichkeit für Datensätze.)
Konfession	Konfession des Schülers	T102_022_Katalog Wert	Text	100			
SchuelerStatus	Anzeige des Status des Schülers: Aktiv oder Ehemalig (wobei nur die ehemaligen Schüler ausgegeben werden, die in dem Halbjahr ehemalig wurden und das hessische Schulsystem verlassen haben)	T144_022_Katalog Wert	Text	100	Aktiv Ehemalig		
AbgangsDatum	Datum, an dem der Schüler von der berichtenden Schule						

V. Offene Fragen

- Lässt sich die Binnenerschließung von strukturierten Daten stärker standardisieren? Was wären Vorteile einer Standardisierung?
- Ist Maschinenlesbarkeit – z.B. für die Feldtabelle – von Vorteil?
- Wo und wie sollte die Binnenerschließung zukünftig abgelegt werden? AFIS, Teil der Metadaten des DIP, referenzierte Archivalie, ganz anders? Oder ist das eigentlich auch egal, Hauptsache die Informationen finden sich irgendwo?
- Wie weit gehen wir bei der Dokumentation? Was ist noch effizient, wo sind die Grenzen erreicht?

Kontakt

Sigrid Schieber

Sigrid.Schieber@hhstaw.hessen.de