Findmittel-Digitalisierung im Staatsarchiv Basel-Stadt

Die Digitalisierung der Repertorien

Ein Projektbericht

von

Lambert Kansy

Übersicht

Vorstellung des Projekts zur Findmittel-Digitalisierung 2003-2005 aus der Perspektive des Staatsarchivs Basel-Stadt

1. Kontext
2. Lösungsansatz
3. Verlauf
4. Ergebnis
5. Ausblick

Kontext

Teil eines Gesamtprojekts zur Informatisierung

Zielsetzung

Es werden „...bestehende Archivverzeichnisse ... In eine digitale Form konvertiert und und online zugänglich gemacht..“

Angestrebte Ergebnisse:

- Orts- und zeitunabhängige Recherche ermöglichen
- Einheitliches Recherchesystem
- Basis für elektronisches Bestellwesen
- Freisetzen von Ressourcen im Lesesaal

Lösungsansatz

Integration der bestehenden Findmittel in die Archivdatenbank scopeArchiv

Erste Schritte:

- Auswertung von Vorarbeiten 2000-2001
- Evaluation der Findmittel
- Prüfung der technischen Umsetzung

Vorarbeiten 2000-2001

Pilotdigitalisierung

30 Findmittel digitalisiert

Mithilfe Scanning, OCR-Bearbeitung

Entwicklung Findmittel-Assistent

Ergebnis unbefriedigend:

- Scannen der Findmittel
- Qualität OCR-Bearbeitung und Strukturierung
- Datenqualität der Findmittel

Konzept Findmittel-Digitalisierung

Zielsetzungen, Portfolio, Abläufe, Methoden

Umsetzungsvorschlag

Orientiert eher an technisch innovativen Verfahren und Erschliessungsbedürfnisse als an Benutzungsperspektive

Auswahlkriterien & Portfolio

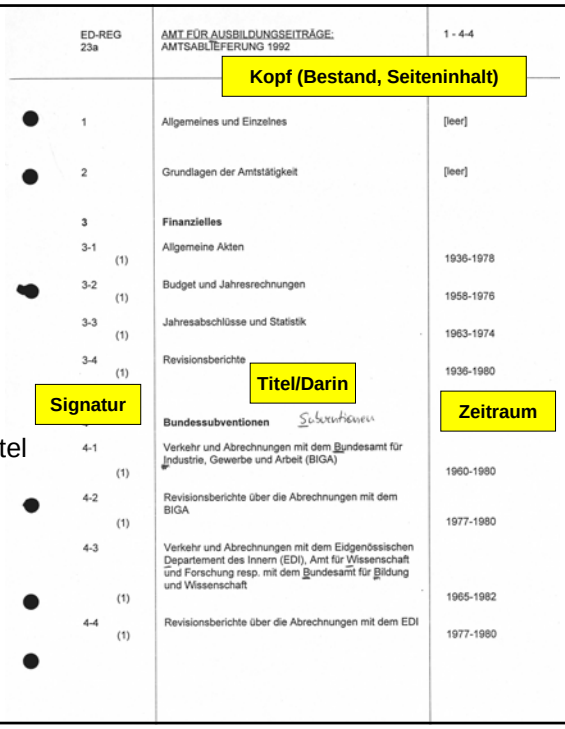
- Nutzungshäufigkeit
- Zugänglichkeit
- Bestellprozesse
- Relevanz für Überlieferungsbildung
- Umfang (Menge, Zeitraum)

- Repertorium A
  Älteres Hauptarchiv und Ältere Nebenarchive (Pertinenzbestände), ca. 4000 Seiten
- Repertorien B und C
  Neueres Hauptarchiv (B) und Neuere Nebenarchive (C) (Provenienzbestände), ca. 800 Seiten
- Repertorium D
  Privatarchive, ca. 4000 Seiten

Findmittel (1)

Findmittelstruktur

- Format A4
- Einzelseiten
- 3 Spalten
- Maschinenschrift
- Einseitig beschriftet
- Pro Bestand ein Findmittel
- Einheitliche Struktur pro Findmittel



ED-REG 23a	AMT FÜR AUSBILDUNGSBEITRÄGE: AMTSABLIEFERUNG 1992	1 - 4-4
1	Allgemeines und Einzelnes	[leer]
2	Grundlagen der Amtstätigkeit	[leer]
3	**Finanzielles**	
3-1 (1)	Allgemeine Akten	1936-1978
3-2 (1)	Budget und Jahresrechnungen	1958-1976
3-3 (1)	Jahresabschlüsse und Statistik	1963-1974
3-4 (1)	Revisionsberichte	1936-1980
	Bundessubventionen Subventionen	
4-1 (1)	Verkehr und Abrechnungen mit dem Bundesamt für Industrie, Gewerbe und Arbeit (BIGA)	1960-1980
4-2 (1)	Revisionsberichte über die Abrechnungen mit dem BIGA	1977-1980
4-3 (1)	Verkehr und Abrechnungen mit dem Eidgenössischen Departement des Innern (EDI), Amt für Wissenschaft und Forschung resp. mit dem Bundesamt für Bildung und Wissenschaft	1965-1982
4-4 (1)	Revisionsberichte über die Abrechnungen mit dem EDI	1977-1980

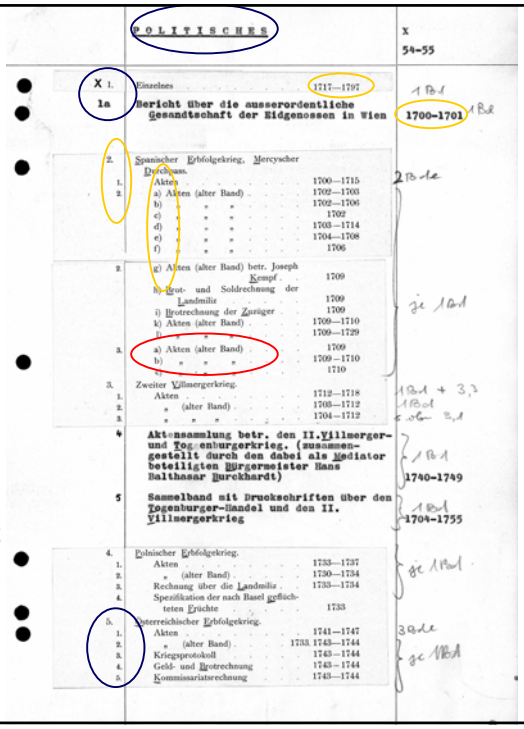
Findmittel (2)

Innerhalb eines Findmittels:

- **Springende Spalten**
- **Platzhalter/fehlende Titel**
- **verteilte Signatur**

Findmittel übergreifend:

- **uneinheitliche Signaturen**
- **heterogene Struktur/Layout**


POLITISCHES

X
54-55

X 1. Einzelnes . . . . . 1717—1797 1 Bd
1a Bericht über die ausserordentliche Gesandtschaft der Eidgenossen in Wien 1700-1701 1 Bd

2. Spanischer Erbfolgekrieg, Mercyscher Durchzug.
1. Akten . . . . . 1700—1715 2 Bde
2. a) Akten (alter Band) . . . . . 1702—1703
b) " " " . . . . . 1702—1706
c) " " " 1702
d) " " " . . . . . 1703—1714
e) " " " . . . . . 1704—1708
f) " " " 1706
2 g) Akten (alter Band) betr. Joseph Kempf . 1709
h) Brot- und Soldrechnung der Landmiliz . . 1709
i) Brotrechnung der Zuzüger . . 1709
k) Akten (alter Band) . . . . 1709—1710
l) " " " . . . . 1709—1729 je 1 Bd
3. a) Akten (alter Band) 1709
b) " " " . . . . 1709—1710
c) " " " 1710

3. Zweiter Villmergerkrieg.
1. Akten . . . . . 1712—1718 1 Bd + 3,3
2. " (alter Band) . . . . . 1703—1712 1 Bd
3. " " " . . . . . 1704—1712 siehe 3,1

4 Aktensammlung betr. den II. Villmerger- und Toggenburgerkrieg. (zusammengestellt durch den dabei als Mediator beteiligten Bürgermeister Hans Balthasar Burckhardt) 1740-1749 1 Bd

5 Sammelband mit Druckschriften über den Toggenburger-Handel und den II. Villmergerkrieg 1704-1755 1 Bd

4. Polnischer Erbfolgekrieg.
1. Akten . . . . . 1733—1737 je 1 Bd
2. " (alter Band) . . . . . 1730—1734
3. Rechnung über die Landmiliz . . 1733—1734
4. Spezifikation der nach Basel geflüchteten Früchte 1733

5. Österreichischer Erbfolgekrieg.
1. Akten . . . . . 1741—1747 3 Bde
2. " (alter Band) . . . . . 1733. 1743—1744 je 1 Bd
3. Kriegsprotokoll . . . . . 1743—1744
4. Geld- und Brotrechnung . . . . 1743—1744
5. Kommissariatsrechnung . . . . 1743—1744

Methodenwahl

Testlauf manuelle Erfassung

- Pertinenzbestände
- durch Fachpersonal

Ergebnis

1 Seite (importfertig): 0.6 Std.

Evaluation automatisierte Digitalisierung (docWORKS)

- Erste Kontakte Frühjahr 2003
- Sommer 2003 Kontakt mit CCS: Testseiten
- September 2003: DFG-Workshop präsentiert docWORKS
- Herbst 2003: Kontakt mit Hauptstaatsarchiv Düsseldorf

Ergebnis:

- **Einsatz für StABS-Findmittel grundsätzlich möglich**
- **Anpassung docWORKS notwendig & möglich**

Ergebnis der Modellrechnung

- Manuelle Erfassung

  Zeitbedarf: 5400 Std. = 3. Jahre (0.6 h/S.)

  Kosten: ca. Fr. 200‘000.-
- Automatisierte Digitalisierung (docWORKS)

  Zeitbedarf: 2700 Std. = 1 1/2 Jahre (0.3 h/S.)

  Kosten: ca. Fr. 170‘000.-

Annahmen für inhouse Digitalisierung:

- 2 Projektmitarbeitende (Fachpersonal)
- 9000 Seiten
- Manuelle Erfassung: 0.6 Std./Seite
- Bearbeitung mit docWORKS: 0.3 Std./Seite

Gründe für docWORKS:

- **deutliche kürzere Laufzeit**
- **potentielle Nachnutzung möglich**
- **Kosten vorteilhafter**

Variantenwahl (Kriterien und Gewichtung)

Gewichtung	Kriterium	Projekt-MA	Dienstleister
20%	Kosten (im Rahmen des Budgets?)	0	20
15%	Kostenrisiko	0	15
15%	Einhaltung Zeitplan (Abschluss bis Ende 2005?)	7.5	7.5
10%	Aufwand StABS Administration	0	10
10%	Aufwand StABS Qualitätssicherung	5	5
10%	Know how-Erwerb StABS (Nachnutzung)	10	5
10%	Einarbeitungszeit	0	10
10%	Entwertung getätigter Investitionen (docWorks-Beschaffung)	10	5
Total		**32.5**	**77.5**

Punktwerte: 0-5-10

Entscheid zugunsten Vergabe an Dienstleister

Vergabekriterien

- Zwingende Kriterien :
  - Kostendach
  - Fristeinhaltung
- Weitere Kriterien (inkl. Gewichtung):
  - 30% Kosten
  - 10% Fristen
  - 30% Know how i. S. Konversion
  - 10% Know how-Transfer (Nachnutzung)
  - 10% Referenzen
  - 10% Räumliche Zusammenarbeit (inhouse/externe Durchführung)

Digitalisierung durch ArchivInForm

Dauer: 18. April 2004 bis 25. Februar 2005

Ablauf:

1. Scannen der Findmittel (extern
2. Findmittelanalyse, Erfassungs
3. Produktionsvorbereitung docW
4. Bearbeitung der Findmittel (26-34 Wochen)
5. Konvertierung in Zielformat (1 Woche)
6. Import nach scopeArchiv, inkl. Nachbearbeitung (4 Wochen)

Ressourcen StABS:

Paul Müller (Informatik; Gesamtprojektleitung)

Lambert Kansy (Erschliessung/Archivinformatik

docWORKS: Ernüchterung

Etappe 1

- Findmittelanalyse (Struktur, Layout)
- Erfassungsregeln erarbeiten

Etappe 2

- Einrichtung docWORKs (Regelwerk)

Komplexere Strukturen und Erfassungsregeln als erwartet.

Einrichtung docWORKs schwieriger als erwartet.

• Anpassung des Regelwerks

•Fehler bei Verarbeitung

Anpassungen der Software notwendig

Manuelle Erfassung

Anpassung docWORKS:

- In Mai-Juni 2004 Definition der Änderungen, Kontakt mit Hersteller (CCS).
- Keine verlässliche Zusage hinsichtlich Realisierung.

Manuelle Erfassung als Alternativverfahren geprüft und beschlossen.

Manuelle Erfassung erfolgt durch Partner von ArchivInForm.

Ergebnisse im selben Format wie docWORKS-Resultate geliefert.

„Produktion“

Phase 2:

- Erstellung generelle Erfassungsregeln: Juni 2004
- Beantwortung von Rückfragen zu einzelnen Findmitteln: Juni - November 2004

Phase 3:

- Juni 2004: erste Testergebnisse
- September/Oktober 2004: 3000 Seiten importfertig geliefert
- Bis Mitte Februar 2005: Bearbeitung aller Seiten und Lieferung als Importdateien

doxTransform & Import scopeArchiv

- Keine „blosse“ Transformation von einer xml-Struktur in die andere.
- Nachbearbeitungsschritte integriert:
  - Zusammenziehen von Signaturen
  - Erstellung der Hierarchie
  - Konvertierung der Datumsformate
- Import nach scopeArchiv mittels scopeÜbernahme-Assistent
  - Archivseitige Nachkontrolle (OCR-Qualität, Strukturierung, Vollständigkeit)
  - Definition von Feldzuordnungen Importdatei-Datenbank
  - Setzen von Standardwerten (Benutzungsbedingungen, Status, Darstellung/Masken)

Nachbearbeitung und Qualität

Anforderungen erfüllt
OCR-Fehlerquote geringer als erwartet
Spezifische Fehlerarten pro Erfassungsmethode

Qualität - Anforderungen

- 0% Fehler in Signaturen
- 1% Fehler bei Titel, Darin, Zeitraum

Aufwand im Rahmen der Erwartungen
Grossteil der archivseitigen Arbeit fiel hier an.
z. T. verlagert in Phase Produktion

Nachbearbeitung

- vor Import: Standardwerte für Verzeichungseinheiten setzen
- nach Import: Strukturierung, Zeitraumkumulation

Ergebnis

- Repertorien A-D (9000 Seiten) liegen digitalisiert vor.
- Volumen der scopeArchiv-Datenbank verdoppelt.
- Digitalisierung innerhalb des Zeitplans und zu vereinbarten Kosten (ca. Fr. 110‘000.-)
- Import bzw. Nachbearbeitung aufwendiger als geplant
- Ansatz Digitalisierung 1:1
  - Führt zu raschen Resultaten
  - Schafft Erklärungsbedarf für Benutzung

Neue Recherchequalität erreicht:

Fast Gesamter Aktenbestand inkl. Pertinenzbestände in scopeArchiv vorhanden und via Query recherchierbar

Fazit

- Realisierung durch Vergabe erst spät evaluiert.
  - ➔ Erhöhter Aufwand für Entscheid
- Erwartungen an docWORKS wurden enttäuscht.
  - ➔ Einsatz für komplexe und heterogene Findmittel - auch maschinenschriftliche - nur sehr bedingt ausserhalb des Findmittel-Korpus des DFG-Projekts möglich.
- Digitalisierung durch erfahrenen Dienstleister war sinnvoll
  - ➔ Kosten- und Zeitrahmen konnten eingehalten werden
  - ➔ Aufwand im StABS blieb überschaubar (total 6 Monate)
  - ➔ Schwierigkeiten mit docWORKS konnte begegnet werden

Ausblick

Evaluation der Digitalisierung weiterer Findmittel 2005-2006

- Alter Planarchiv-Katalog
  ca. 14‘000 Karteikarten, maschinenschriftlich
  Grund: bes. Benutzungsreglen, Strukturierung
- Urkunden-Regesten
  ca. 25‘000 Karteikarten, v. a. handschriftlich
  Grund: Handschriftliches Findmittel
- Findmittel der Bildersammlung
  Heterogene Strukturen und Formate, z. T. handschriftlich
  Grund: bes. Benutzungsregeln; Teil der Bilderdigitalisierung, Strukturen

Vielen Dank für Ihre Aufmerksamkeit

Kontakt:
Lambert Kansy
Staatsarchiv Basel-Stadt
Erschliessung und Archivinformatik
Email: lambert.kansy@bs.ch
Tel. +41 61 267 86 98
Fax +41 61 267 65 71