

## Sauver les CD

### Resultate der Bestandesaufnahme

#### Inhalt

1	Ausgangslage .....	1
2	Ablauf der Bestandesaufnahme .....	1
3	Ergebnisse .....	2
4	Folgerungen .....	3

#### 1 Ausgangslage

Das Projekt "Sauver les CD" basiert auf der Prämisse, dass in vielen Archiven bereits seit einigen Jahren CD-R, DVD-R und andere ephemere (digitale und analoge) Datenträger aufbewahrt werden, und dass wegen ihres Alters und ihrer archivischen Bedeutung Handlungsbedarf besteht, die darauf gespeicherten Daten vor dem Medienzerfall zu retten. Bevor geeignete Strategien zu dieser Rettung erarbeitet werden, galt es, die Prämisse durch eine Bestandesaufnahme in Mitgliedsarchiven der KOST zu erhärten, allenfalls zu widerlegen oder zu korrigieren.

Am Projekt beteiligen sich neun Mitgliedsarchive der KOST, nämlich die Staatsarchive von AG, BS, BE, LU, SZ, UR, VD und VS sowie das Landesarchiv FL. Damit ist die Hälfte der KOST-Mitglieder vertreten; die Untersuchung hat mithin eine hinreichend breite Basis.

#### 2 Ablauf der Bestandesaufnahme

Die Bestandesaufnahme wurde in jedem der beteiligten Archive einzeln in Zusammenarbeit mit der KOST durchgeführt. Sie umfasste zwei Schritte. Der erste bestand aus einer Mengenanalyse, das heisst im Wesentlichen aus der Erstellung einer Liste aller vorhandenen ephemeren Datenträger. In gewissen Archiven sind digitale oder auch im weiteren Sinn audiovisuelle Datenträger gesondert verzeichnet, was einen einfachen und vollständigen Überblick über diese Bestände ermöglicht. Andere Archive mussten einen solchen Überblick durch geeignete Suchabfragen im Archivinformationssystem erstellen, wobei die Vollständigkeit nicht hundertprozentig garantiert ist.

Der zweite Schritt der Bestandesaufnahme zielte darauf ab, via eine Drill-Down-Analyse inhaltliche und technische Angaben zu den gefundenen digitalen Unterlagen zu machen. Dies wurde teilweise durch eine Grobanalyse auf der Basis der Verzeichnungsinformationen und des Wissens der Archivmitarbeitenden erreicht. In den meisten Fällen schloss sich eine detailliertere Analyse der Datentypen, Erzeugungsdaten, Datenformate und anderer Eigenschaften auf einer zufälligen Stichprobe der Datenträger an. Dabei leistete ein CD-Analysetool der KOST gute Dienste.

Obwohl das Projekt "Sauver les CD" ausdrücklich jede Art von ephemeren Datenträgern in den Blick nahm, also auch analoge, konzentrierte sich die Drill-Down-Analyse auf die digitalen Datenträger. Analoge Datenträger wie VHS-Kassetten, Filmrollen, Audiokassetten, Schallplatten etc. stellen in ihrer Mehrzahl keine akuten

konservatorischen Probleme. Dafür ist der Zugang zu diesem Archivgut oft schwierig. Mittelfristig ist eine Digitalisierung in Betracht zu ziehen (dazu mehr unten). Die analogen Datenträger wurden in der Bestandesaufnahme nicht detailliert quantifiziert und werden in den folgenden statistischen Analysen nicht bzw. nur zu Vergleichszwecken einbezogen.

### 3 Ergebnisse

#### 3.1 Anzahl und Art der Datenträger

Das erste Resultat der Mengenanalyse ist, dass in den beteiligten Archiven tatsächlich relativ bedeutende Bestände ephemerer digitaler Datenträger aufbewahrt werden, gesamthaft 6219 Stück. Die Spannweite ist gross: Sie reicht von 32 Stück in Luzern (wo im Rahmen des Projekts weitere 150 CDs als nicht archivwürdig kassiert werden konnten) bis zu 2339 Stück im Fürstentum Liechtenstein (wo vor allem die Filmsammlung und das Rheinberger-Musikarchiv ins Gewicht fallen); der Durchschnittswert ist 691, der Median 223. 60% aller verzeichneten Datenträger sind CDs (CD-R, CD-ROM und Audio-CD), 16% gehören zu den verschiedenen DVD-Formaten. Die restlichen 24% sind weitere digitale Datenträger wie MiniDisc (568 Exemplare im LLA), DAT-Bänder (Bestände von 700 Stück im StABS und von rund 100 Stück im StAUR), Disketten, ZIP-Disks, digiBeta, Beta SP und DV Cam (insgesamt 82 Exemplare in 4 Archiven)<sup>1</sup>.

#### 3.2 Corpora

93% aller erfassten digitalen Datenträger gehören jeweils einem grösseren, einheitlichen Corpus an. Beispiele solcher Corpora sind die Filmsammlung im FL (800 DVDs), die In-House-Digitalisate von Bildern und Fotos in VD (145 CDs), die (sämtlich von der gleichen Firma ausgeführten) Bilddigitalisate in SZ (130 CDs) oder die Tonaufnahmen der Ratsdebatten in UR (360 CDs). Diese Corpora wurden oft mit einheitlichen Methoden erstellt, die entsprechenden Dateien liegen in einem oder wenigen gut bekannten Formaten vor. Datenträger, die aus einzelnen Ablieferungen von Privatarchiven oder von staatlichen Stellen stammen, machen nur 7 Prozent aus<sup>2</sup>.

#### 3.3 Originale und Kopien

Für die Beurteilung des weiteren Vorgehens war entscheidend, wie viele der auf ephemeren Datenträgern gespeicherten Daten Originale sind. Unter einem Original verstehen wir in diesem Zusammenhang eine digitale Datei, von welcher kein analoges Pendant existiert und die damit der Träger der Überlieferung ist. Das können sowohl digital erzeugte Dateien sein, sogenannte *originär digitale Dateien (born-digital)*, als auch Digitalisate, deren (analoge) Vorlagen nicht mehr existieren oder dem Archiv nicht zugänglich sind.

---

<sup>1</sup> Die Tatsache, dass das Staatsarchiv BS und das Landesarchiv FL jeweils für über einen Drittel der im ganzen Projekt erfassten digitalen Datenträger verantwortlich sind, verzerrt diese prozentualen Analysen. Deshalb geben wir auch die über die neun beteiligten Archive normalisierten Zahlen: Ein durchschnittliches Archiv hat 71% CDs, 15% DVDs und 13% andere digitale Datenträger.

<sup>2</sup> Die normalisierten Zahlen: Ein durchschnittliches Archiv hat 75% Datenträger aus Corpora und 25% Einzelstücke.

Als Kopien, d.h. als nicht überlieferungstragend, können fast die Hälfte der untersuchten Datenträger bezeichnet werden, nämlich rund 46% (normalisiert: 44%). Die übrigen 54% sind Originale im obigen Sinn. Dabei handelt es sich nur zu einem geringen Teil um Archivalien der staatlichen Überlieferung, nämlich um rund 9% (normalisiert: 19%) aller betrachteten Datenträger. Es sind dies praktisch ausschliesslich die Audioaufnahmen der Ratsdebatten in BE und UR; dazu kommen einige wenige Ablieferungen (38 Datenträger) staatlicher Stellen in anderen Kantonen. Der grosse Rest der Originale kann als Sammlungsgut angesprochen werden, nämlich rund 45% der Gesamtzahl (normalisiert 37%).

Entscheidend ist hier die Erkenntnis, dass praktisch sämtliche im Projekt betrachteten Daten als archivwürdig eingeschätzt werden müssen. Dies gilt ohne nähere Begründung für die Originale sowohl staatlicher als auch privater Provenienz. Es gilt jedoch auch für einen Grossteil der Kopien. Diese sind zwar reproduzierbar, häufig aber nur mit grossem Aufwand, sind sie doch zumeist im Rahmen mehr oder weniger umfangreicher Digitalisierungsprojekte entstanden. Als Zugangskopien sind sie auch für den Archivbetrieb von Bedeutung. Einige wenige Ausnahme bestätigen die allgemeine Regel der Archivwürdigkeit: So wurden im StALU 150 CDs mit Aufnahmen vom Autobahnneubau zu Gunsten eines kleinen Sample kassiert. In VD lässt die Stichprobe darauf schliessen, dass viele der Datenträger aus Privatarchiven nur die digitale Version von Office-Dokumenten beinhalten, die im Bestand ebenfalls enthalten sind und das archivische Original darstellen. Diese Datenträger könnten folglich problemlos kassiert werden.

### **3.4 Dateiformate**

In Sachen Dateiformate wurde die relative Geschlossenheit eines grossen Teils der untersuchten Daten bereits erwähnt. Entsprechend machen Bildformate (TIFF, JPEG), Audioformate (WAV, MP3) sowie Videoformate (MPEG2) den grössten Teil dieser Daten aus. In Beständen aus Privatarchiven und vereinzelt Ablieferungen staatlicher Stellen finden sich jedoch in sehr kleiner Zahl auch weitere Formate: Officeformate (DOC, XLS und andere) sowie spezielle Applikationsformate.

## **4 Folgerungen**

### **4.1 Keine Überlieferungslücke...**

Die Prämisse des Projekts "Sauer les CD" war, dass in den Staatsarchiven wegen der bereits archivierten ephemeren Datenträgern eine Überlieferungslücke droht. Diese Befürchtung kann relativiert werden. Zwar sind in den beteiligten Archiven tatsächlich grosse Bestände an CDs, DVDs etc. vorhanden. Die meisten davon (die erwähnten 54% originär digitalen Daten sowie offenbar ein Grossteil der "Kopien") sind auch wirklich überlieferungswürdig. Der grösste Teil dieser Datenträger ist jedoch nicht älter als 4 Jahre; die ältesten gefundenen digitalen CDs stammen von 2001. Nicht mehr lesbare Datenträger waren absolute Einzelfälle in den Stichproben (einige 5¼-Zoll-Disketten und eine fehlerhafte CDs in VD).

### **4.2 ...aber Handlungsbedarf**

Zwei Gefahren wurden jedoch identifiziert, die in ihrer Verbindung zu raschem Handeln Anlass geben: In gewissen Archiven existiert noch keine Sicherheitsstrategie für digitale Datenträger; die CDs und DVDs sind in einigen Fällen der einzige Träger der

Überlieferung, und es gibt keine Sicherungskopie. Zudem gibt das Alter der ältesten vorhandenen CDs (7 Jahre) auf Grund der bekannt schlechten Stabilität gebrannter optischen Disks Grund zur Sorge. Deshalb ist eine Sicherungsstrategie für die ephemeren digitalen Datenträger schnell zu entwickeln.

### 4.3 Strategie für das weitere Vorgehen

Die KOST hat a priori zwei verschieden umfassende und intensive Strategien ins Auge gefasst. Für kleine Bestände war das regelmässige Umkopieren auf neue Datenträger als ideal angesehen worden. Für grössere Bestände oder ein archivisch saubereres Vorgehen wurden die Inbetriebnahme einer archivtauglichen Speicherinfrastruktur und das Umkopieren der Datenträger darauf vorgesehen.

Sämtliche am Projekt beteiligten Archive haben sich in der Bestandesaufnahme für die Anschaffung einer archivtauglichen Speicherinfrastruktur und damit gegen das Umkopier-Konzept ausgesprochen. Das hat zur Folge, dass dieses Konzept nicht weiter verfolgt wird. Stattdessen wird der zweite Teil des Projekts *Sauver les CD* mit dem Projekt *arcun* zusammengeführt (Archival Repository for Collaborative Use). Im Rahmen dieses Projekts erarbeitet die KOST Vorschläge für eine von den Staatsarchiven gemeinsam genutzte Speicherinfrastruktur. Diese soll danach für die Speicherung der Daten auf ephemeren Datenträgern verwendet werden.

### 4.4 Analoge audiovisuelle Datenträger

Die Bestandesaufnahme identifizierte zusätzlich zu den digitalen Datenträgern auch teils grosse Bestände analoger ephemerer Datenträger. Dabei handelt es sich praktisch durchgehend um audiovisuelle Datenträger.

Auf diese analogen Datenträger ging das Projekt, wie oben bereits angemerkt, nicht im Detail ein. Gleichwohl können ein paar summarische Aussagen gemacht werden:

- Die Vielfalt analoger ephemerer Datenträger ist bemerkenswert. Audiodaten sind auf verschiedenen Tonbandspulen, Kompaktkassetten und Schallplatten gespeichert, Filme sowohl auf Schmalfilm als auch auf Videokassetten, jeweils in diversen Formaten.
- Aussagen zur Lesbarkeit dieser Datenträger sind schwierig zu machen. Die *Bitstream Preservation* als solche scheint nicht das Hauptproblem zu sein, sondern vielmehr die Benutzung. Die Staatsarchive haben in der Regel keinerlei Abspielgeräte für diese audiovisuellen Träger; im Bedarfsfall müssen Mediatheken oder andere spezialisierte Institutionen einspringen.
- Zur Reduktion der Anzahl Trägerformate im Archiv, zum erleichterten Zugang, aber auch zur Vermeidung des Lesbarkeitsverlusts ist in den betroffenen Archiven mittelfristig wohl die Digitalisierung dieser audiovisuellen Datenträger in Betracht zu ziehen. Die Bestände sind in der Regel relativ übersichtlich, so dass eine durchgehende Digitalisierung finanziell und organisatorisch machbar sein sollte. Dazu müssen jedoch zwei Vorbedingungen erfüllt sein: Es muss eine archivtaugliche Speicherinfrastruktur zur Verfügung stehen, und die Parameter der Digitalisierung müssen im Hinblick auf die Langzeiterhaltung sorgfältig erarbeitet werden. Die KOST plant, zu dieser Fragestellung nächstes Jahr einen eigenen Beitrag zu leisten.