

Aktueller Stand der Langzeitarchivierung von Compact Discs in Gedächtnisorganisationen

Bachelorarbeit

Bibliothek und digitale Kommunikation

Fakultät für Informations- und Kommunikationswissenschaften

Technische Hochschule Köln

Vorgelegt von:

Felix Kratzer

am 23.10.2023 bei Prof. Dr. Ursula Arning und Michelle Lindlar

Abstract

Compact Discs sind Datenträger, die in großer Zahl in den Beständen von Gedächtnisorganisationen vorkommen. Weil die Datenträger nur begrenzt haltbar sind, entwickelt sich immer stärker die Notwendigkeit, die Daten von den Trägern zu trennen und gesondert zu archivieren. Die Arbeit gibt einen Überblick darüber, welche Institutionen das Problem auf welche Art und Weise angehen und schlägt Perspektiven für die Zukunft vor. Im theoretischen Teil wird der Begriff „Gedächtnisorganisation“ geklärt, es wird die Bewahrungsproblematik von CDs ausgeführt und es werden Migration und die Erstellung von Disc Images zur Emulation als Archivierungsoptionen erläutert. Im Praxisteil werden mithilfe von Interviews und Literaturarbeit verschiedene Einrichtungen und deren individuelle Strategien vorgestellt. Es zeigt sich deutlich, dass die Langzeitarchivierung optischer Datenträger aktuell noch ein sehr individuelles Thema ist, das Einrichtungen primär aus eigener Motivation angehen und bislang nur wenig Zusammenarbeit stattfindet, um gemeinsame Lösungen zu entwickeln. Abschließend werden Vorschläge formuliert, was unternommen werden kann, um die Langzeitarchivierung in Deutschland zu vereinheitlichen und zu rationalisieren.

Inhalt

Abkürzungsverzeichnis.....	V
Abbildungsverzeichnis	VI
1. Zielsetzung der Arbeit	1
2. Begriffe und Technisches	1
2.1 Gedächtnisorganisationen	2
2.2 Typen von Compact Discs.....	3
2.3 Methoden der Datensicherung	4
2.3.1 Migration	4
2.3.2 Disc Image und Emulation	5
3. Problemaufriss	5
3.1 Rolle von Compact Discs in Gedächtnisorganisationen	7
3.2 Sinkende Anzahl an Abspielgeräten	8
3.3 Begrenzte Haltbarkeit von CDs.....	8
4. Methodik der Datenerhebung	10
4.1 Auswahl von Best-Practice-Beispielen	10
4.2 Auswahlkriterien für Interviewpartner	11
4.3 Fragenkatalog der Interviews.....	12
4.4 Ablauf der Interviews	15
5. Ergebnisse und Auswertung	15
5.1 Relevanz der LZA in den jeweiligen Einrichtungen.....	16
5.2 British Library	17
5.3 Königliche Bibliothek der Niederlande	19
5.4 New York Public Library.....	20
5.5 Deutsche Nationalbibliothek.....	21
5.6 Technische Informationsbibliothek Hannover	22
5.7 Landesarchiv Baden-Württemberg	24
5.8 Zentrum für Kunst und Medien Karlsruhe	25
5.9 Herzog August Bibliothek Wolfenbüttel.....	26
6. Zusammenfassung	26
7. Ausblick	29
7.1 Unterstützung unerfahrener Einrichtungen.....	29
7.2 Datenbank für archivierte Werke.....	30
8. Fazit	31
9. Literaturverzeichnis	32
Anhang 1: Interviewtranskripte	1
Anhang 1.1: DNB-Interview	1

Anhang 1.2: HAB-Interview	10
Anhang 1.3: ZKM-Interview.....	15
Anhang 1.4: LABW-Interview	25
Anhang 1.5: TIB- Interview	37
Anhang 2: E-Mailkorrespondenz.....	55
Anhang 2.1: E-Mail BSB	55
Anhang 2.2: E-Mail Niedersächsische Staats- und Universitätsbibliothek Göttingen.....	56
Anhang 2.3: E-Mail Landesarchiv NRW	57
Anhang 2.4: E-Mail SLUB Dresden.....	58
Anhang 2.5: E-Mail Staatsbibliothek zu Berlin – Preußischer Kulturbesitz	58
Anhang 2.6: E-Mail UB Regensburg.....	59
Anhang 2.7: E-Mails ZB MED	60
Anhang 2.8: E-Mail ZBW	62

Abkürzungsverzeichnis

BL	British Library
BSB	Bayerische Staatsbibliothek München
CD	Compact Disc
CD-DA	Compact Disc Digital Audio, häufiger Audio-CD
CD-R	Compact Disc – Recordable
CD-ROM	Compact Disc – Read Only Memory
CD-RW	Compact Disc – Rewriteable
COPTR	Community Owned digital Preservation Tool Registry
DNB	Deutsche Nationalbibliothek
EMIL	Emulation of Multimedia Objects in Libraries
HAB	Herzog August Bibliothek Wolfenbüttel
IROMLAB	Image and Rip Optical Media Like A Boss
KBNL	Königliche Bibliothek der Niederlande (Koninklijke Bibliotheek)
LABW	Landesarchiv Baden-Württemberg
LZA	Langzeitarchivierung
NYPL	New York Public Library
SLUB	Sächsische Landesbibliothek – Staats- und Universitätsbibliothek Dresden
TIB	Technische Informationsbibliothek Hannover
ZB MED	Zentralbibliothek für Medizin Köln
ZBW	Zentralbibliothek für Wirtschaft Kiel
ZKM	Zentrum für Kunst und Medien Karlsruhe

Abbildungsverzeichnis

Abbildung 1: Absatz von CD-Alben in Deutschland nach Jahr (in Mio. Stck.).....	6
Abbildung 2: Musikstreams gesamt nach Jahr (in Mrd.)	7

1. Zielsetzung der Arbeit

Die vorliegende Arbeit beschäftigt sich mit der Langzeitarchivierung (LZA) von Compact Discs (CDs). Dieses Thema wird immer drängender, weil CDs kein geeignetes Medium sind, um Informationen verlustfrei über mehrere Jahrzehnte hinweg zu bewahren. Bereits 2007 gab das Deutsche Musikarchiv bekannt, dass Ausfälle in einer relevanten Größenordnung festgestellt werden konnten, die die Integrität des Bestandes bedrohen.¹

Zudem soll in einer Momentaufnahme dargestellt werden, wie unterschiedlich deutsche Gedächtnisorganisationen damit umgehen: wer bereits laufende Projekte hat und wie sie angelegt sind, wer sich aktuell mit dem Thema befasst, um Projekte zu starten und wer seine Prioritäten vorerst auf andere Themen legt.

Hierbei erhebt diese Arbeit keinen Anspruch auf Vollständigkeit. Zwar lassen sich anhand der Auswahl der Einrichtungen, die Selbstauskünfte beigesteuert haben oder die zu einem Interview bereit waren, begründete Vermutungen über den Rest der deutschen Bibliotheks- und Archivlandschaft rechtfertigen, dennoch sind die Erkenntnisse nur bedingt repräsentativ.

Darüber hinaus kann der Text als Handreichung verstanden werden: Die Institutionen, die sich noch nicht mit dem Thema befasst haben, erhalten einen Einblick in die Problematik. Diejenigen, die ein entsprechendes Bewusstsein besitzen, allerdings auf sich gestellt überfordert sind, finden Materialien und potenzielle Ansprechpartner. Sonstige Interessierte werden erkennen, dass der Weg noch weit ist.

2. Begriffe und Technisches

Für das bessere Verständnis der Arbeit ist es sinnvoll, vorab Begriffe und Konzepte zu erläutern, die nicht vollständig selbsterklärend sind. Es gibt Erklärungen zum Begriff „Gedächtnisorganisation“, um das Bezugsfeld der Arbeit einzugrenzen, zu den Typen von

¹ vgl. Groth, Susanne, Alexander Uhl (2007): Deutsches Musikarchiv: CD-Zerfall bedroht Kulturerbe. Online unter <https://www.heise.de/news/Deutsches-Musikarchiv-CD-Zerfall-bedroht-Kulturerbe-148962.html> [Abruf am 16.10.2023]

CDs, um die Heterogenität des Untersuchungsgegenstands aufzuzeigen und zu den verschiedenen Arten der Langzeitarchivierung, für die sich Gedächtnisorganisationen anhand ihrer Bedürfnisse entscheiden können. Was bewusst weggelassen wurde, sind tiefgreifende technische Erläuterungen. Zum einen würden sie die Komplexität dieser Arbeit stellenweise unverhältnismäßig erhöhen, ohne das Verständnis des Sachverhalts zu verbessern. Zum anderen wurde die technische Seite der digitalen Kopie und Sicherung von anderen Quellen weitaus umfangreicher und fachkundiger beschrieben. Zu empfehlen ist das entsprechende Dokument von Angela Dappert, British Library (BL).²

2.1 Gedächtnisorganisationen

Der Studiengang „Bibliothek und digitale Kommunikation“ an der TH Köln, dessen Abschluss die vorliegende Arbeit ist, konzentriert sich beinahe ausschließlich auf die namensgebenden Bibliotheken. Nur Bibliotheken zu behandeln wäre allerdings im Kontext dieser Arbeit unzureichend, weil die im folgenden Kapitel dargestellten Problematiken für alle Einrichtungen von Belang sind, die CDs besitzen und erhalten möchten.

Daraus folgt die Notwendigkeit, neben Bibliotheken auch andere Gedächtnisorganisationen zu betrachten. Der Begriff „Gedächtnisorganisation“ oder auch „Gedächtnisinstitution“ ist im Kontext der Langzeitarchivierung sehr gebräuchlich und kaum auf einzelne Urheber zurückzuführen. Je nach Quelle gehören andere Institutionen dazu: Schweibenz und Sieglerschmidt (2010) nennen Bibliotheken, Archive und Museen,³ Hjerpe (2014) zählt außerdem Kulturerbe-Institutionen wie Monumente und Standorte sowie Aquarien und Arboreten, zoologische und botanische Gärten auf.⁴ Maßgeblich ist in jedem Fall, dass Wissen gesammelt, bewahrt und vermittelt wird.⁵

Sinnvoll ist aufgrund der Vergleichbarkeit in der Bestandszusammensetzung eine Beschränkung auf Archive, Museen und wissenschaftliche Bibliotheken. Öffentliche

² Dappert, Angela et al. (2011): Developing a Robust Migration Workflow for Preserving and Curating Hand-held Media. In: Proceedings der Tagung iPRES 2011 - 8th International Conference on Preservation of Digital Objects, 01.-04. November 2011, Singapur, S. 33-43

³ vgl. Schweibenz, Werner; Sieglerschmidt, Jörn (2010): Aktuelle Entwicklungen bei Kultur-Portalen: BAM-Portal, Deutsche Digitale Bibliothek und Europeana. In: Proceedings der Tagung Inventarisierung digital - schmale und breite Wege in den Datenhimmel des Museumsverbandes Baden-Württemberg e.V. am 7. und 8. Mai 2010 im Landesmuseum Württemberg, Stuttgart, S. 1f

⁴ vgl. Hjerpe, Roland (1994): A Framework for the Description of Generalised Documents. In: Proceedings der Tagung ISKO '94 Conference, 21. bis 24. Juni 1994, Kopenhagen, S. 1

⁵ vgl. Byrne, Alex (2014): Memory institutions shaping the past, present and future. In: Proceedings der Tagung IFLA PAC, Geneva, am 13. und 14. August 2014, Genf, S. 4f

Bibliotheken hingegen werden bewusst nicht einbezogen, weil hier selten Archivierung stattfindet, sondern konstant Medien ausgesondert werden, deren Zustand nicht mehr den Erwartungen entspricht oder die nicht nachgefragt werden. Auch wenn der Erhalt der Originaldatenträger – Bücher, Dokumente oder eben CDs – meist Priorität besitzt, etabliert sich seit der Einführung elektronischer Datenverarbeitung auch immer mehr die zusätzliche digitale Speicherung.

2.2 Typen von Compact Discs

Die Compact Disc ist ein geschütztes Format, das eine breite Spanne an Varianten mit unterschiedlichen Funktionen aufweist. Spezifiziert sind diese in den sogenannten „Rainbow Books“, wobei jedes Buch eine eigene Farbe für die jeweilige Variante aufweist. Jede Variante hat eine eigene Haltbarkeitsdauer und wird teils unterschiedlich langzeitarchiviert, worauf in späteren Kapiteln eingegangen wird.

Im Folgenden sind in Auswahl Arten von CDs aufgeführt:

- Red Book: CD-DA (Compact Disc Digital Audio, im Folgenden „Audio-CD“), ursprüngliches CD-Format zum Speichern und Abspielen von Musik und anderen Audiodateien⁶
- Yellow Book: CD-ROM (Compact Disc Read Only Memory), Datenspeicher für Computer⁷
- Orange Book Part II: CD-R (Compact Disc Recordable), einmal beschreibbar durch Endnutzende⁸
- Orange Book Part III: CD-RW (Compact Disc Rewritable), mehrmals beschreibbar durch Endnutzende⁹

⁶ vgl. Van Dijk, Boudewijn; Nijboer, Jaap: The CD Family (1996), S. 31. Online unter <https://www.lscdweb.com/data/downloadables/7/7/the-cd-family-3122-783-0068-1.pdf> [Abruf am 18.10.2023]

⁷ vgl. ebd.

⁸ vgl. Adminius B.V. (o. J.): CD Products. Online unter https://www.lscdweb.com/ordering/cd_products.html [Abruf am 18.10.2023]

⁹ vgl. ebd.

2.3 Methoden der Datensicherung

Grundsätzlich ist es auf zwei Arten möglich, digitale Daten von Datenträgern zu sichern: Migration der Daten sowie die Erstellung emulierbarer Disc Images. Beide Varianten haben jeweils bevorzugte Anwendungsgebiete und Voraussetzungen.

2.3.1 Migration

Bei der Migration stehen die abgespeicherten Daten auf dem Träger im Vordergrund. Der Datenträger selbst sowie das technische Gerüst spielen untergeordnete Rollen. Das Speichermedium wird nur als ein Transfermedium für die Inhalte betrachtet.¹⁰ Nach dem Transfer können die Daten neu organisiert und unterabhängig voneinander weiterverarbeitet werden. Üblich ist, falls nötig, auch eine Überführung in Formate, die besser in die eigene Speicherstruktur passen oder die eher für die Langzeitarchivierung geeignet sind.¹¹ Im Landesarchiv Baden-Württemberg (LABW) beispielweise legen Behörden manchen Papierakten CD-Rs mit ergänzenden elektronischen Daten bei.¹² Die Datenträger werden kopiert, von der entsprechenden Beschriftung wird ein Bild gemacht und alles wird gemeinsam in einem Datensatz abgelegt; im Anschluss wird der Träger entsorgt.¹³ In anderen Einrichtungen ist die Entsorgung optional.

Laut Wikipedia ist ein Bitstrom „eine Bit-Sequenz in zeitlicher Abfolge von unbestimmter Länge“ gegliedert in Bytes und weiter in Blöcke und Datenpakete.¹⁴ Technisch betrachtet braucht jeder Datenspeicher, dessen Daten in Form mehrerer Bitströme abgelegt sind, ein Containerformat, das die Position der Ströme auf dem Träger mithilfe von sogenannten Partitionstabellen verzeichnet. Diese Tabellen stellen zusätzliche Metadaten dar.¹⁵ Bei der Migration werden erwähnte Bitströme einzeln extrahiert, wobei der Container nicht mitkopiert wird – Strukturmetadaten und damit die Authentizität des Datenträgers gehen dadurch verloren.¹⁶ „Authentizität“ bedeutet die Erhaltung der

¹⁰ vgl. Naumann, Kai (2023): LABW-Interview. Zoom-Interview mit Dr. Kai Naumann, Referatsleiter Archivisches Recht im Landesarchiv Baden-Württemberg, durchgeführt von Felix Kratzer am 13.09.2023. Transkription auf S. A25-A36 im Anhang, 05:34-06:09

¹¹ vgl. Dappert (2011), S. 35f.

¹² vgl. Naumann (2023), 1:30-3:09

¹³ vgl. ebd., 28:05-29:27

¹⁴ Wikipedia (2022): Bitstrom. Zuletzt aktualisiert am 22.05.2022. Online unter <https://de.wikipedia.org/wiki/Bitstrom> [Abruf am 19.10.2023]

¹⁵ vgl. Dappert (2011), S. 34

¹⁶ vgl. ebd.

inneren Struktur im Datenträger, also aller Metadaten des Dateisystems, darunter der Bootsektor, die Strukturen und Attribute.¹⁷

2.3.2 Disc Image und Emulation

Legt man jedoch Wert auf die Authentizität des Datenträgers, bietet es sich an, ein Disc Image zu erstellen. Statt in mehreren Bitströmen werden die Daten hier in einem einzigen großen Bitstrom übertragen.¹⁸ Als ideales Ziel gibt die British Library an, eine reversible Migration durchzuführen. Die digitale Entität, die beim Kopieren des Originaldatenträgers erzeugt wird, könnte dann noch genutzt werden, um einen neuen Datenträger mit dem identischen Funktionsumfang des Originals zu erschaffen.¹⁹

Um auf die Kopie zuzugreifen, ist ein PC bzw. ein anderes Abspielgerät mit passendem Betriebssystem nötig, das das entsprechende Disc Image interpretieren kann, weil beispielsweise ein PC mit Windows 10 standardmäßig keine Datenträger für macOS 7.6 abspielen kann. Alternativ kann auch ein Emulator eingesetzt werden, der die Aufgabe übernimmt.²⁰ Ein Beispiel dafür ist „Emulation of Multimedia Objects in Libraries“ (EMiL) der Deutschen Nationalbibliothek, Bayerischen Staatsbibliothek, Albert-Ludwigs-Universität Freiburg und Staatlichen Hochschule für Gestaltung Karlsruhe. EMiL ist laut der DNB nutzungsfreundlich, weist einen hohen Automatisierungsgrad auf und kann dank standardisierten Schnittstellen auch in anderen Institutionen eingesetzt werden.²¹

3. Problemaufriss

Die Compact Disc ist ein veraltetes Format. Auch wenn immer noch neue CDs erscheinen, nimmt der Absatz über die Jahre konsequent ab. Rein digitale Dateien sind wesentlich praktischer. Sie können anders als physische Datenträger in Sekundenbruchteilen

¹⁷ vgl. ebd., S. 35

¹⁸ vgl. ebd., S. 34

¹⁹ vgl. ebd., S. 35

²⁰ vgl. Steinke, Tobias (2023): DNB-Interview. Zoom-Interview mit Tobias Steinke, Deutsche Nationalbibliothek, durchgeführt von Felix Kratzer am 08.09.2023. Transkription auf S. A1-A9 im Anhang, 17:38-18:45

²¹ vgl. Deutsche Nationalbibliothek (2019): EMiL. Bereitstellung von Multimedia-Objekten durch Emulation. Online unter https://www.dnb.de/DE/Professionell/ProjekteKooperationen/Projektarchiv/2016/emulationMultimediaObjekte/emulationMultimediaObjekte_node.html [Abruf am 10.10.2023]

übertragen werden, haben keine von der Anzahl vertriebener Medieneinheiten abhängigen Herstellungskosten und sind jederzeit überall verfügbar. Datenspeicherung im Internet ist einfach, häufig kostenlos und nicht auf fragile Träger angewiesen.

Im Diagramm sind exemplarisch die Absatzzahlen für Musik-CDs in Deutschland visualisiert.

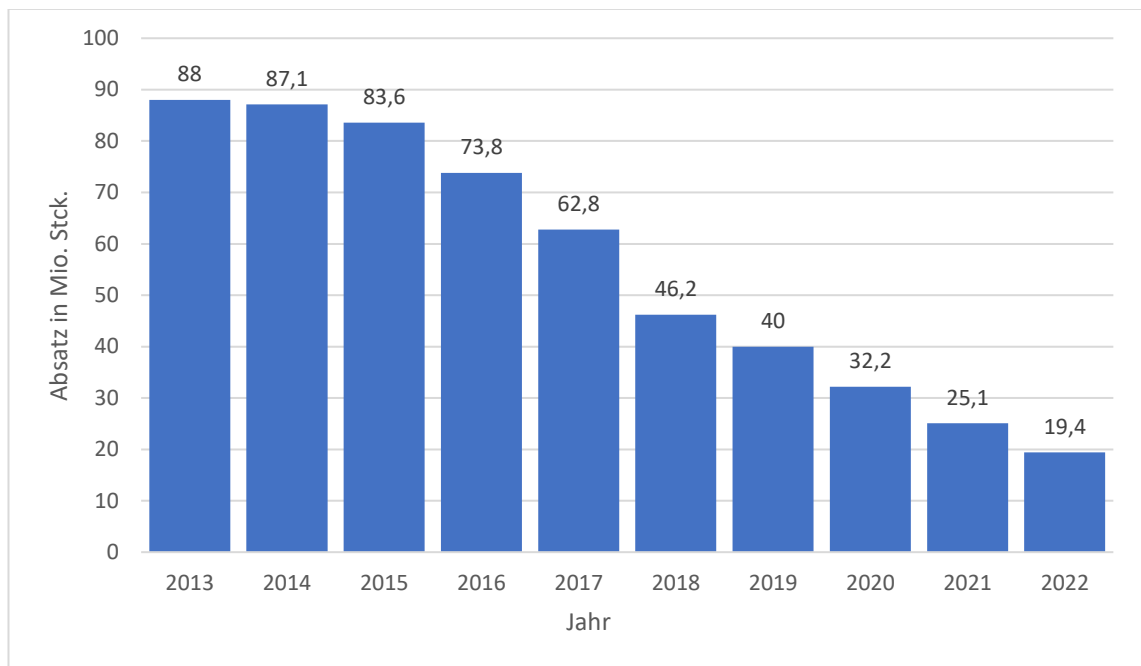


Abbildung 1: Absatz von CD-Alben in Deutschland nach Jahr (in Mio. Stck.)²²

Im Gegenzug nimmt das Musikstreaming weiter zu (siehe folgende Abbildung) und verdrängt physische Absatzformen sowie digitale Downloads²³.

²² vgl. Bundesverband Musikindustrie e.V. (2023): Musikindustrie in Zahlen 2022, S. 16. Online unter <https://www.musikindustrie.de/wie-musik-zur-karriere-werden-kann/markt-bestseller/musikindustrie-in-zahlen-2022>

²³ vgl. ebd., S. 17

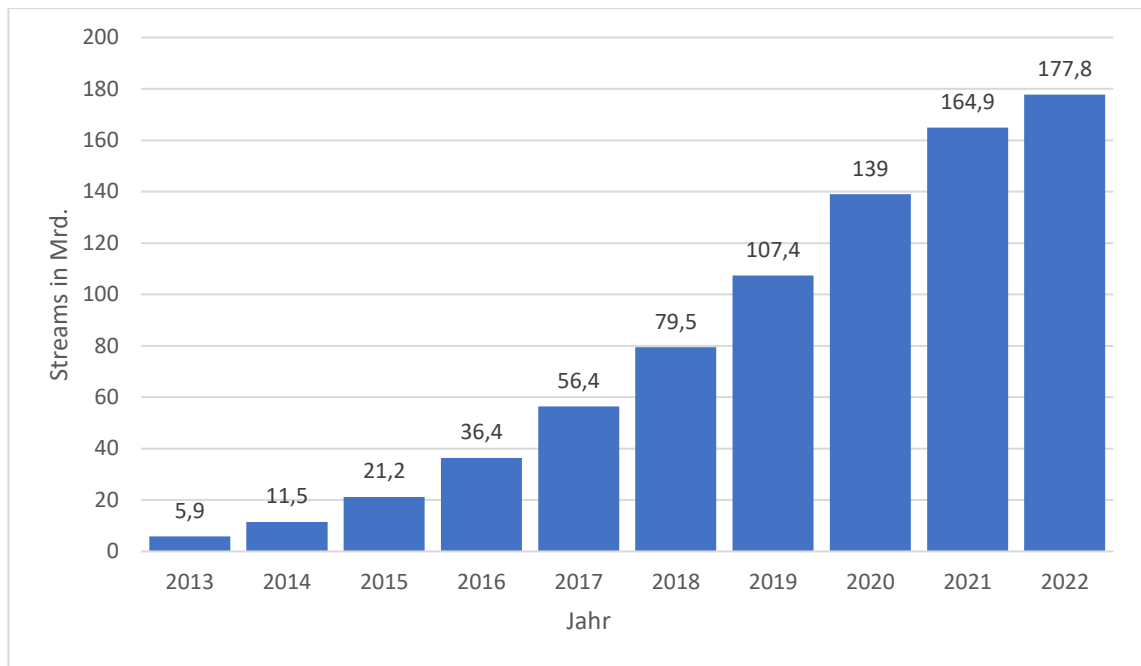


Abbildung 2: Musikstreams gesamt nach Jahr (in Mrd.)²⁴

Die Nachfolgedatenträgerarten DVD und Blu-ray Disc tragen zwar zur technischen Irrelevanz der CD bei, aber auch hier sinken die Kundenzahlen,²⁵ während die Umsätze von Streamingdiensten weiter steigen.²⁶

3.1 Rolle von Compact Discs in Gedächtnisorganisationen

Compact Discs finden sich aus verschiedenen Gründen in den Beständen von Gedächtnisorganisationen wieder. Institutionen mit Pflichtexemplarrecht erhalten alle Publikationen, die in ihrem jeweiligen Sammelgebiet erscheinen und das schließt CDs zu großen Teilen ein.²⁷ Außerdem kommen viele CDs auch als Beilage zu gedruckten Medien ins Haus, die ergänzende Materialien zum Text beisteuern.²⁸ Des Weiteren gibt es CDs, die bewusst angeschafft werden, darunter Videospiele, Nutzsoftware oder Datenbanken.

²⁴ vgl. ebd., S. 18

²⁵ vgl. IVD (2020): Anzahl der Kunden im DVD/Blu-ray-Kaufmarkt in Deutschland in den Jahren 2007 bis 2019. In Statista. Online unter <https://de.statista.com/statistik/daten/studie/168578/umfrage/anzahl-der-kunden-im-dvd-markt-seit-2007/> [Abruf am 06.10.2023]

²⁶ vgl. PwC (2023): Umsätze mit T-VoD und S-VoD in Deutschland in den Jahren 2012 bis 2027*. In Statista. Online unter <https://de.statista.com/statistik/daten/studie/770083/umfrage/umsaetze-im-vod-markt-in-deutschland-nach-segmenten/> [Abruf am 06.10.2023]

²⁷ Die Deutsche Nationalbibliothek sammelt alle CDs, deren Inhalt dem entspricht, was auch in Buchform gesammelt werden würde, Videospiele zählen nicht dazu. Zudem nimmt das Deutsche Musikarchiv, ein Teil der DNB, das Pflichtexemplarrecht unter anderem für Musik-CDs wahr (vgl. Steinke (2023), 1:55-3:19)

²⁸ vgl. Jorzenuk, Saskia (2023): HAB-Interview. Zoom-Interview mit Saskia Jorzenuk, Bereich Neuere Medien der Herzog August Bibliothek, durchgeführt von Felix Kratzer am 11.09.2023. Transkription auf S. A10-A14 im Anhang, 0:13-0:53

Viele dieser Datenbanken existieren nicht in einer lizensierbaren und damit datenträgerfreien Version, sondern nur auf CD.²⁹ Manchmal sind CDs Teil von Nachlässen und übernommenen Sammlungen. Zuletzt erstellen manche Institutionen immer noch selbst CDs, beispielsweise für digitale Editionen.³⁰

3.2 Sinkende Anzahl an Abspielgeräten

Die sinkenden Absatzzahlen und die technische Überalterung führen dazu, dass Abspielgeräte immer weniger nachgefragt und dementsprechend auch schwerer erhältlich werden. Bei Laptops³¹ und Autos³² beispielsweise werden aufgrund sinkender Relevanz und hohem Platzverbrauch teils keine Abspielgeräte mehr ab Werk verbaut. Es ist möglich, dass irgendwann aus Mangel an Geräten nicht mehr auf Daten von aufbewahrten CDs zugegriffen werden kann. Hardwarearchive für alte Computersysteme können helfen, den Zugang zu Werken wieder zu ermöglichen, für die ein Abspielgerät fehlt. Eines wird beispielsweise aktuell in einem Flugzeughangar in Neubiberg eingerichtet.³³ Solche Initiativen sind aber keine dauerhafte und simple Lösung, da alte Technik irgendwann marode wird und ohnehin nur einem begrenzten Nutzerkreis Zugriff auf die überschaubare Anzahl an Geräten gewährt werden kann.

3.3 Begrenzte Haltbarkeit von CDs

Darüber hinaus sind auch die Datenträger selbst nicht sonderlich haltbar, Angaben hierzu sind teils widersprüchlich. Die Herstellerfirma Sony gab 2012 für CDs eine Mindesthaltbarkeit von 20 Jahren mit einer Wahrscheinlichkeit von 95 Prozent unter optimalen Lagerbedingungen an, spätere Statements sprechen darüber hinaus von Haltbarkeiten von über 35 Jahren für die CD-R „Silver“³⁴ und über 100 Jahre für die CD-R in

²⁹ vgl. Ebd., 16:30-16:55

³⁰ vgl. Haardt, Christian (2023): ZKM-Interview. Zoom-Interview mit Christian Haardt, Labor für antiquierte Videosysteme im Zentrum für Kunst und Medien Karlsruhe, durchgeführt von Felix Kratzer am 13.09.2023. Transkription auf S. A15-A24 im Anhang, 15:21-16:02

³¹ vgl. Mills, Matt (2022): Warum hat ein aktueller Laptop kein CD- oder DVD-Laufwerk?. Online unter <https://itigic.com/de/why-doesnt-a-current-laptop-include-a-cd-or-dvd-drive/> [Abruf am 06.10.2023]

³² vgl. Lang, Patrick (2021): Vom Aussterben bedroht. Wo es noch CD-Player gibt und wo nicht. Online unter <https://www.auto-motor-und-sport.de/tech-zukunft/cd-player-autos-modelle-ausstattung/> [Abruf am 06.10.2023]

³³ vgl. Meyer, Ulrich (2010): Ausstellung zeigt wahre Computerschätze. Kennen Sie noch C64, Amiga und Co.?. Online unter https://rp-online.de/digitales/neuheiten/ausstellung-zeigt-wahre-computer-schaetze_aid-12875925 [Abruf am 06.10.2023]

³⁴ Bei „Silver“ handelt es sich um einen Produktnamen, nicht um ein weiteres Rainbow Book.

goldener Farbe, jeweils unter optimalen Bedingungen.³⁵ Prof. Dr.-Ing. Matthias Hemmje von der Fernuniversität Hagen schätzt die Haltbarkeit einer gepressten CD bei optimalen Lagerbedingungen auf 50-80 Jahre, wohingegen selbst gebrannte CDs [CD-R oder CD-RW, Anm. d. Autors] gegebenenfalls schon nach einem Jahr nicht mehr gelesen werden können.³⁶

Folgende Faktoren beeinflussen die Funktionalität:³⁷

1. Lagerbedingungen: Optimale Lagerung findet bei maximal 18°C in staub- und lichtgeschützter Umgebung statt, welche nur bedingt zu gewährleisten ist.
2. Nutzungsintensität: Regelmäßige Nutzung inklusive physischer Belastung kann das Medium schädigen. Schon bei leichtem Verbiegen entstehen Haarrisse im Schutzlack.
3. Materialwahl für Träger, Reflexionsschicht und Aufdruck: der Materialmix kann miteinander reagieren. Gerade die Lacke, die anfangs für die Etiketten verwendet wurden, fressen sich durch die Schichten und beeinträchtigen die Reflexionsfähigkeiten.
4. Entstehungsweise: Von Fabrik aus gepresste Datenträger halten sich länger als selbstgebrannte CDs.

Um die Herstellerversprechen für beschreibbare DVDs überprüfen zu können, führte das Computermagazin c't für einen 2008 publizierten Artikel³⁸ Belastungstests an DVD-Rs von verschiedenen Herstellern durch. Die Ergebnisse legen nahe, dass die Produkte vieler Firmen teils weit hinter den Werbeaussagen zurückbleiben. Nachdem keine CD-Rs mit in den Versuch aufgenommen wurden, lassen sich die Erkenntnisse vermutlich nicht eins zu eins übertragen, dennoch sollte man sich auch hier nicht vollständig auf die versprochenen Laufzeiten verlassen.

³⁵ vgl. Hardwarejournal.de (2013): Wie hoch ist die Lebensdauer von CD- und DVD-Rohlingen?. Online unter <https://www.hardwarejournal.de/cd-lebensdauer-prognose/> [Abruf am 06.10.2023]

³⁶ vgl. Groth (2007)

³⁷ vgl. ebd.

³⁸ vgl. Gieselmann, Hartmut (2008): Silberne Erinnerungen: Archiv-DVDs im Langzeittest. In: c't - Magazin für Computertechnik 16, 2008, S. 116ff

Wie groß der Einfluss optimaler Lagerungsbedingungen auf die Haltbarkeit sein kann, demonstriert die Technische Informationsbibliothek Hannover (TIB). In einem prototypischen Projekt mit 14.000 homogenen CD-Rs oder CD-RWs, die zuvor über 10 Jahre nicht ausleihbar oder nutzbar waren, gab es keinen einzigen Totalausfall und lediglich 19 Discs wiesen Fehler auf.³⁹ Diese Probleme betrafen jedoch nur die äußersten Sektoren der CDs, die am meisten der Umgebung ausgesetzt waren.⁴⁰ Bei stärker nachgefragten Werken sind jedoch größere Ausfälle erwartbar.

4. Methodik der Datenerhebung

In diesem Kapitel folgen die Auswahlprozesse für die Best-Practice-Beispiele und die Interviewpartner sowie der erarbeitete Fragenkatalog und der Ablauf der Interviews.

4.1 Auswahl von Best-Practice-Beispielen

Eines der Produkte eines „SPRUCE“ getauften Projekts von University of Leeds, British Library, Digital Preservation Coalition, London School of Economics und Open Planets Foundation⁴¹ war ein Wiki mit dem Titel „Community Owned digital Preservation Tool Registry“ (COPTR). Im Wiki wurden existierende Listen von Langzeiterhaltungstools in geordneter Form zusammengeführt.⁴² Unter der Kategorie „Workflow: Community Owned Workflows“⁴³ gibt es eine Liste verschiedener größerer Einrichtungen, die Langzeitarchivierung betreiben. Klickt man die einzelnen Einträge an, wird man auf Seiten weitergeleitet, die die jeweiligen internen Workflows aufschlüsseln und erklären. Viele der Institutionen beschäftigen sich nicht mit optischen Datenträgern und daher sind sie für die vorliegende Arbeit nicht von Belang. Von den verbliebenen Einrichtungen waren drei am erfolgversprechendsten: Zur British Library lässt sich viel Hintergrundmaterial

³⁹ vgl. Schwab, Franziska (2023): TIB-Interview. Zoom-Interview mit Franziska Schwab, Preservation Analyst im Team Langzeitarchivierung der TIB, durchgeführt von Felix Kratzer am 17.10.2023. Transkription auf S. A37-A54 im Anhang, 08:59-11:11

⁴⁰ vgl. ebd.

⁴¹ vgl. Pennock, Maureen (2014): HOME. Online unter <http://wiki.opf-labs.org/display/SPR/Home> [Abruf am 18.10.2023]

⁴² vgl. COPTR (2021): About COPTR. Zuletzt aktualisiert am 03.10.2021. Online unter https://coptr.digipres.org/index.php/About_COPTR [Abruf am 18.10.2023]

⁴³ Vgl. COPTR (2021): Workflow:Community Owned Workflows. Zuletzt aktualisiert am 28.04.2021. Online unter https://coptr.digipres.org/index.php/Workflow:Community_Owned_Workflows [Abruf am 18.10.2023]

finden.⁴⁴ Außerdem bauen die Workflows der New York Public Library (NYPL) und der Königlichen Bibliothek der Niederlande (KBNL) darauf auf⁴⁵ und bieten deswegen eine gute Vergleichsbasis. Zusätzlich wurden verschiedene Gedächtnisinstitutionen als Interviewpartner angefragt, um einen Überblick über den Status Quo zu bekommen.

4.2 Auswahlkriterien für Interviewpartner

Ein Fokus auf Deutschland lag nahe, weil er verschiedene Vorteile mit sich bringt:

1. Interviews können auf Deutsch abgehalten werden, was die Verwendung von fachsprachlichen Ausdrücken sowie die gegenseitige Verständigung erleichtert. Die Gefahr für Missverständnisse wird reduziert und gezieltere Rückfragen werden möglich.
2. Es ist weniger Recherche bezüglich der Strukturen in der jeweiligen ausländischen Landschaft von Gedächtnisinstitutionen nötig. Recherche wäre die Voraussetzung dafür, Fragebögen an die Institutionen anzupassen und die Antworten korrekt zu interpretieren.
3. Konzentration auf ein überschaubares Gebiet macht die Ergebnisse repräsentativer. Zwar ist dabei immer noch Vorsicht geboten, aber es lassen sich begründete Vermutungen bezüglich dem Rest Deutschlands anstellen.

Interviews mit ausländischen Institutionen wurden trotzdem nicht kategorisch ausgeschlossen, da sie potenziell ergänzende Informationen beisteuern könnten.

Ein zweiter Fokus lag darauf, bei allen Mitgliedsinstitutionen von nestor e.V. anzufragen, einem deutschen Netzwerk von Kooperationspartnern, die sich mit Langzeitarchivierung befassen.⁴⁶ Hier finden sich nicht nur alle sechs Bibliotheken mit nationaler Bedeutung⁴⁷ wieder, sondern auch das Bundesarchiv, zwei Landesarchive, ein Museum, mehrere Verbände sowie Institute. Es ist zu vermuten, dass in diesem Netzwerk die Konzentration

⁴⁴ z.B. Dappert (2011)

⁴⁵ Rothrock, Michelle et al. (2021): Assessing High-volume Transfers from Optical Media at NYPL. In: Code4Lib Journal 51, 2021-06-14, Heft 51. Online unter: <https://journal.code4lib.org/articles/15908> [Abruf am 19.10.2023]

⁴⁶ vgl. nestor e.V. (2022): Über uns. Online unter https://www.langzeitarchivierung.de/Webs/nestor/DE/nestor/Ueber_uns/ueber_uns_node.html [Abruf am 18.10.2023]

⁴⁷ Die Liste findet sich unter <https://bibliotheksportal.de/informationen/bibliothekslandschaft/bibliotheken-mit-nationaler-bedeutung/> [Abruf am 18.10.2023]

derjenigen Einrichtungen am höchsten ist, die sich mit dem vorliegenden, eher spezifischen Themenbereich befasst haben. Bis auf die Verbünde, die selbst keine eigenen Medien besitzen, wurden alle Mitglieder um Interviews gebeten.

Darüber hinaus gibt es noch zwei weitere Interviewpartner. Das Zentrum für Kunst und Medien (ZKM) in Karlsruhe teilt sich die Bibliothek mit der staatlichen Hochschule für Gestaltung in Karlsruhe, die gemeinsam mit der DNB und BSB an der Entwicklung von EMiL beteiligt war. Zusätzlich interessant wurde sie für die vorliegende Arbeit durch umfangreiche Darstellungen der langzeitarchivarischen Tätigkeiten auf der Homepage.

Die Herzog August Bibliothek in Wolfenbüttel nimmt in der Reihe der übrigen befragten Einrichtungen eine eher untergeordnete Rolle ein. Sie wurde ausgewählt, weil der Autor dort in seinem Praxissemester mit der vom Datenträger getrennten Bereitstellung von CD-Inhalten bereits in Berührung kam.

4.3 Fragenkatalog der Interviews

Der Fragebogen baut auf der Literaturliteratur auf. Er fragt alle Aspekte des Langzeitarchivierungskonzepts für CDs ab, die auch in der British Library eine Rolle spielen, gleichzeitig aber in anderen Einrichtungen abweichen können. Wenn durch die Webpräsenz der jeweiligen Einrichtung schon Fragen vollumfänglich beantwortet wurden, wurde der Fragebogen angepasst. Dasselbe war der Fall, wenn das Studium der Homepage weitere diskussionswürdige Aspekte aufzeigte. Für beide Fälle bietet das Interview mit der DNB ein Beispiel. Vorabrecherchen zeigten, dass zur Emulation die Software EMiL eingesetzt wird, weswegen der Katalog um die Frage „Wie gut funktioniert die Emulation von Disc Images über EMiL?“ ergänzt wurde. Gleichzeitig wurden beispielsweise die ersten drei Fragen durch die Recherchen überflüssig gemacht und gestrichen.

Außerdem wurde im Folgenden der Bearbeitungsprozess der Datenträger mehrmalig als „Digitalisierung“ betitelt. Ursache hierfür ist, dass es für den umfassenden Prozess der Kopierung, Bearbeitung und Bereitstellung optischer Datenträger anders als bei analogen Büchern keinen simplen Begriff gibt, wohingegen „Digitalisierung“ bei Büchern genau diesen Zweck erfüllt. Abgesehen vom folgenden Fragebogen, der sehr früh im Arbeitsprozess entstand, wurde in der restlichen Arbeit bewusst Wert darauf gelegt, besagte Formulierung zu umschreiben.

1. Spielt LZA (z.B. Forschungsdaten, e-Journals/e-Books, Open-Access-Materialien, selbst erstellte Retrodigitalisate) in Ihrer Einrichtung eine relevante Rolle?
2. Besitzt Ihre Einrichtung optische Datenträger (CDs, CD-ROMs etc.)?
3. Werden an Ihrer Einrichtung Daten von ihren optischen Trägern getrennt, beispielsweise, um sie leichter der ganzen Einrichtung zur Verfügung stellen zu können?
4. Spielt die LZA optischer Datenträger eine relevante Rolle?
5. Sind zur LZA optischer Datenträger Projekte abgeschlossen, laufen aktuell oder sind in Planung?
6. Existiert eine einrichtungsbezogene Preservation Policy (interne Richtlinie für den Archivierungsprozess)?
7. Ist LZA ein Thema, das einrichtungsübergreifend angegangen wird, um doppelte Arbeiten zu vermeiden, beispielsweise durch aktive Zusammenarbeit, identische Techniklösungen, ähnliche Workflows oder eine gemeinsame Datenbank?
8. Wie intensiv ist generell der Austausch in dieser Hinsicht?
9. Wie viele CDs wurden bislang digitalisiert? Wie viele fehlen noch? Wird alles, was digitalisiert ist, auch zur Verfügung gestellt, womöglich sogar im Open Access? (Genaue Zahlen sind nicht zwangsweise nötig, die Größenordnung ist ebenfalls ausreichend.)
10. Wie viel Zeit benötigt der Vorgang etwa, ggf. aufgeschlüsselt nach Digitalisierung, Bearbeitung und Bereitstellung?
11. Welche technischen Vorrichtungen (Hardware und Software) werden genutzt, um die Daten zu digitalisieren, aufbereiten und speichern? Wie stark ist der Vorgang automatisierbar bzw. bei Ihnen automatisiert?
12. Wer führt diese Arbeiten durch (z.B. HiWis, feste Angestellte) und in welcher Abteilung ist der Prozess angesiedelt? Benötigt die Arbeit eine Zusatzqualifikation?

13. Welches Datenbanksystem zur Speicherung und zum Abruf wird genutzt? Lassen sich die Daten in das reguläre Bibliothekssystem integrieren oder erfordern sie eine eigene Lösung? Wurde dafür eine eigene Sammlung angelegt?
14. Ist das System im Eigenbau entstanden oder zumindest durch selbst programmierte Elemente ergänzt worden?
15. Kommt ein eigener Server, ein Verbundserver oder ein externer Dienstleister zum Einsatz?
16. War der Ankauf bestimmter Hardware/Software nötig, um die Prozesse zu ermöglichen bzw. zu optimieren (z.B. Disc Robot „Acronova Nimbie“ etc.)?
17. Können die Daten nur lokal oder auch webbasiert abgerufen werden?
18. Bedient sich Ihr System der Migration (Daten werden aus ihrer vorliegenden Struktur entnommen und in ein neues Gerüst konvertiert) oder Emulation (die Abspielplattform imitiert das Betriebssystem, für das der Datenträger entwickelt wurde und gibt die Inhalte in der ursprünglichen Form wieder)?
19. Sind die digitalen Kopien allgemein zugänglich, nur intern verfügbar oder anderweitig zugriffsbeschränkt? Spielen hierbei auch urheberrechtliche Bedenken eine Rolle?
20. Wie werden die durch die Digitalisierung erst entstehenden Urheberrechte gehandhabt, sofern sie eine Rolle spielen?
21. Sind die digitalisierten Inhalte auch maschinell auswertbar?
22. Wie sind die Datensätze angereichert?
23. Gibt es eine Verknüpfung von Digitalisat, Artefakt und Katalogsatz des Artefakts?
24. Sind die Mediennummern für Digitalisat und Artefakt identisch? Deckt sich das Vorgehen hier mit dem bei Digitalisaten von Büchern?
25. Werden Schlagwörter und Normdaten ergänzt?

4.4 Ablauf der Interviews

Jede angefragte Einrichtung erhielt eine E-Mail mit Bitte um ein Interview, dem Fragekatalog aus dem vergangenen Kapitel zur Vorbereitung und jeweils einer Deadline sowohl für die Rückmeldung als auch für ein mögliches Interview.

Die Interviews fanden über Zoom statt. Zu Beginn wurde die Erlaubnis abgefragt, das Interview in Bild und Ton aufzuzeichnen, was die weitere Verwendung erleichtert. Danach wurde die Gelegenheit gegeben, erst einmal frei zur Einrichtung, der dortigen Langzeitarchivierungsstrategie und der Rolle von CDs zu referieren. Im Anschluss wurden die Fragen des Fragebogens behandelt. Nachdem die Interviews qualitativ und nicht quantitativ waren, waren Rückfragen von beiden Seiten erwünscht, um die Situation möglichst genau darzustellen. Auch vertiefende Fragen zu einzelnen Aspekten waren möglich. Zum Ende wurde abgeklärt, ob das Interview veröffentlicht werden darf oder einer Sperrklausel unterliegen muss, um sensible Informationen zu schützen. Zu diesem Zeitpunkt lief die Aufnahme noch, um eine rechtliche Absicherung sicherzustellen.

Nach dem Interview fand die Transkription statt, um das Interview leichter zitierbar zu machen.

5. Ergebnisse und Auswertung

Ursprünglich wäre ein anderer Ansatz für die restliche Arbeit vorgesehen gewesen: Der Plan sah vor, die jeweiligen Archivierungsworkflows verschiedener Gedächtnisinstitutionen in der verfügbaren Literatur zu recherchieren und in Interviews zu erfragen, um sie im Anschluss hier zu beschreiben. Danach hätte man damit beginnen können, sie gegenüberzustellen, Variationen abzuwägen und letztendlich einen Musterworkflow zu erstellen, an dem sich andere Einrichtungen orientieren können. Die Erwartung zu Beginn der Arbeitsphase basierten auf der Annahme, dass die meisten Gedächtnisinstitutionen angesichts des drohenden Daten- und Informationsverlusts damit begonnen hätten, sich in Gruppen zusammenzuschließen, gemeinsam Arbeitsabläufe zu entwickeln und diese jeweils leicht für ihre institutionsspezifischen Bedürfnisse zu modifizieren.

Stattdessen hat sich gezeigt, dass die Realität deutlich anders aussieht: nur wenige Einrichtungen haben überhaupt laufende Projekte, die meisten Einrichtungen arbeiten für sich allein. Zusammenarbeit kommt vereinzelt vor, beispielsweise beim Verbund der drei zentralen Fachbibliotheken,⁴⁸ bei mehreren Archiven⁴⁹ oder bei der zuvor erwähnten Entwicklung der Emulationssoftware EMiL, die 2016 abgeschlossen wurde.⁵⁰ Für dieses Projekt wurden unter anderem von der DNB und der BSB jeweils eine überschaubare Anzahl an Disc Images von CDs erstellt. Nach Projektende hingegen folgten keine Langzeiterhaltungsmaßnahmen, die darauf aufbauen.⁵¹ Wie im vergangenen Kapitel angesprochen, ist von der BSB vorerst keine Änderung des Status Quo zu erwarten. Die DNB plant hingegen ein Folgeprojekt.

Dementsprechend behandelt dieses Kapitel die stattdessen erfolgte Herangehensweise: Die Rückmeldungen der Einrichtungen werden analysiert, sodass sie anderen Einrichtungen eine Orientierung bieten können.

5.1 Relevanz der LZA in den jeweiligen Einrichtungen

Die Menge der Rückmeldungen auf die E-Mailanfragen war zufriedenstellend. Von 18 angefragten Institutionen gaben 13 eine Antwort, die Aufschluss über ihren Status gegeben hat. Fünf davon führten zu Interviews über das Videokonferenztool Zoom.

Zehn der befragten Einrichtungen betreiben aktuell aus verschiedenen Gründen dagegen keine Langzeitarchivierung optischer Datenträger: BSB,⁵² SLUB Dresden⁵³ und UB Regensburg⁵⁴ haben jeweils kein Projekt und es ist vorerst auch keines geplant. HAB⁵⁵ und ZBW⁵⁶ haben ihre Projekte an unterschiedlichen Punkten abgebrochen, bevor eine relevante Menge an Datenträgern bearbeitet war. In der Staatsbibliothek zu Berlin

⁴⁸ vgl. Schwab (2023), 11:54-12:40

⁴⁹ vgl. Naumann (2023), 13:07-13:37

⁵⁰ vgl. Deutsche Nationalbibliothek (2019)

⁵¹ vgl. Steinke (2023), 16:36-17:38

⁵² Schoger, Astrid (2023): Antw: Wtrlt: Interview zu LZA im Rahmen meiner Bachelorarbeit. E-Mail an Felix Kratzer, gesendet am 07.09.2023, Anhang S. A55

⁵³ Hübsch, Gerald (2023): AW: Interview im Rahmen meiner Bachelorarbeit. E-Mail an Felix Kratzer, gesendet am 24.09.2023, Anhang S. A58

⁵⁴ Hübsch, Gerald (2023): AW: Interview im Rahmen meiner Bachelorarbeit. E-Mail an Felix Kratzer, gesendet am 24.09.2023, Anhang S. A58

⁵⁵ vgl. Jorzenuk (2023), 03:24-04:05

⁵⁶ Tunnat, Yvonne (2023): Re: Interview zu LZA im Rahmen meiner Bachelorarbeit. E-Mail an Felix Kratzer, gesendet am 19.09.2023, Anhang S. A62

wurden mögliche Szenarien geprüft, anschließend das Projekt allerdings zurückgestellt.⁵⁷ Bei DNB (siehe Kapitel 5.5), SUB Göttingen⁵⁸ und ZB MED⁵⁹ sind jeweils Projekte kurz- oder mittelfristig in Planung. Die TIB hatte bereits ein Testprojekt und geht nun in die Planung für die Bearbeitung des restlichen Bestands über.⁶⁰

Lediglich drei von dreizehn Einrichtungen haben laufende Projekte: Sowohl das Landesarchiv Baden-Württemberg⁶¹ als auch das Landesarchiv Nordrhein-Westfalen⁶² migrieren laufend alle neu eintreffenden CDs. Überwiegend handelt es sich hierbei um CD-Rs von Behörden.⁶³ Das ZKM setzt auf Disc Images, die unkomprimiert auf Magnetbändern gespeichert werden.⁶⁴

Dementsprechend wird sich das restliche fünfte Kapitel mit den individuellen Langzeitarchivierungsstrategien derjenigen Einrichtungen befassen, die bereits Erfahrungen damit haben, mit ihren Eigenheiten und damit, ob andere Institutionen davon Dinge lernen können. Abseits der Interviews sind hier auch die British Library, die KBNL und die NYPL zu finden, deren Arbeitsweisen aus öffentlich zugänglicher Literatur zu entnehmen ist.

5.2 British Library

Die British Library (BL) ist die Nationalbibliothek des Vereinigten Königreichs und nimmt hierfür auch das Pflichtexemplarrecht wahr. Das führt unter anderem zu einer gewaltigen Menge an optischen Datenträgern, die auf Dauer bewahrt werden müssen. Ihr sogenanntes „Endangered Archives Programme“ umfasste zum Startzeitpunkt etwa 18.000 optische Datenträger, Tendenz steigend.⁶⁵ Um mit derartigen Dimensionen umgehen zu können, ist ein teilautomatisierter Workflow nötig. Hierfür wurden mehrere

⁵⁷ Škarić, Doris (2023): WG: Interview zu LZA im Rahmen meiner Bachelorarbeit. E-Mail an Felix Kratzer, gesendet am 15.09.2023, Anhang S. A58

⁵⁸ Strathmann, Stefan (2023): AW: Interview im Rahmen meiner Bachelorarbeit. E-Mail an Felix Kratzer, gesendet am 22.09.2023, Anhang S. A56

⁵⁹ Markus, Katharina (2023): AW: AW: AW: Anfrage zum Zweitgutachten für meine Bachelorarbeit. E-Mail an Felix Kratzer, gesendet am 18.09.2023, Anhang S. A61

⁶⁰ vgl. Schwab (2023), 04:21-05:53

⁶¹ vgl. Naumann (2023), 15:33-16:18

⁶² Wortmann, Christian (2023): AW: Interview im Rahmen meiner Bachelorarbeit. E-Mail an Felix Kratzer, gesendet am 22.09.2023, Anhang S. A57

⁶³ vgl. Naumann (2023), 01:30-02:21

⁶⁴ vgl. Haardt (2023), 10:29-11:08

⁶⁵ vgl. Dappert (2011), S.34

Disc Robots getestet,⁶⁶ darunter die NSM 7000 Jukebox⁶⁷, der Acronova DupliQ DQ-5610⁶⁸ und der Acronova Nimbie NB11⁶⁹. Ersterer wurde verworfen, weil er besser geeignet war, CDs zu schreiben als zu lesen und sich der Datenträgerwechsel gerade in Hinblick auf den Austausch fehlerhafter Exemplare als schwierig erwies.⁷⁰ Der DupliQ fasste weniger CDs, war schwieriger zu beladen und der CD-Wechsler war unzuverlässiger als beim Nimbie, auf den die Entscheidung fiel.⁷¹ Dieser kann mit bis zu 100 Discs beladen werden kann, die automatisch hintereinander bearbeitet werden. Die Software hingegen wurde in der Version gewählt, wie sie dem Acronova DupliQ beilag, unter anderem, weil die Verarbeitung von Discs im HFS+-Format – einem Dateiformat für Macintosh-Rechner⁷² – weniger problematisch war.

Die gewünschte Vorlageform sind emulierbare Disc Images im ADL-Format (AES31-3⁷³), das als einziges die Ansprüche der BL erfüllt. Die Disc Robots sind mit diesem Format zwar nicht nativ kompatibel, die entstandenen Images können jedoch konvertiert werden.⁷⁴ Anschließend folgt die Erstellung stabilisierter, nutzbarer und mit Metadaten angereicherter Datensätze, in denen das Disc Image mit dem physischen Träger über einen eindeutigen Identifier verknüpft ist.

Mit zwei parallel laufenden Maschinen konnten im Schnitt 1050 CDs pro Monat verarbeitet werden, was etwa 2,2 TB Daten entspricht – der limitierende Faktor hierbei war jedoch nicht die Kapazität und Geschwindigkeit der Hardware, sondern die Notwendigkeit, manuell Metadaten zu erzeugen.⁷⁵ Insgesamt wurden bis 2011 mindestens 16

⁶⁶ vgl. ebd., S. 36f

⁶⁷ vgl. Data Archive Corporation (o. J.): DISC DVD-7000 DVD Library. Online verfügbar über Archive.org unter <https://web.archive.org/web/20110509062816/http://www.dataarchivecorp.com/disc-dvd-7000.htm> [Abruf am 18.10.2023]

⁶⁸ vgl. Acronova Technology, Inc. (o. J.): DupliQ. Online verfügbar über Archive.org unter https://web.archive.org/web/20120118073209/http://www.acronova.com/duplicator_dupliq_usb.htm [Abruf am 18.10.2023]

⁶⁹ vgl. Acronova Technology, Inc. (2012): One's Enough! Nimbie USB. Zuletzt aktualisiert im März 2012. Online verfügbar über Archive.org unter <https://web.archive.org/web/20180424024231/http://www.acronova.com/product/nimbie-usb-auto-loader-publisher-ripper-nb11/2/review.html> [Abruf am 18.10.2023]

⁷⁰ vgl. Dappert (2011), S.36

⁷¹ vgl. ebd., S. 37

⁷² vgl. ebd.

⁷³ Siehe <http://www.edlmax.com/AES31.htm>

⁷⁴ vgl. Dappert (2011), S. 35f.

⁷⁵ vgl. ebd., S. 37

Terabytes an Daten verarbeitet.⁷⁶ Aktuelle Berichte, die Aufschluss über den momentanen Stand geben, sind nicht auffindbar.

5.3 Königliche Bibliothek der Niederlande

Die Koninklijke Bibliotheek van Nederland (KBNL) ist die Nationalbibliothek der Niederlande. Die dortige Discsammlung umfasst in etwa 15.000 bis 20.000 Stück. Die überwiegende Menge stellen CD-ROMs dar. Etwa 25 Prozent sind allerdings Audio-CDs und müssen damit anders bearbeitet werden.⁷⁷ Aufgrund der Vorarbeiten der British Library hat man sich dafür entschieden, einen Acronova Nimbie NB21-DVD⁷⁸ anzuschaffen. Auch die KBNL war mit der mitgelieferten Software nicht zufrieden, weshalb das Workflow-Programm „Image and Rip Optical Media Like A Boss“ (IROMLAB) entwickelt wurde, das verschiedene Softwarekomponenten beinhaltet:

The workflow is built around a number of tried and tested software components. For the disc type identification it uses cd-info, which is part of libcdio, the "GNU Compact Disc Input and Control Library". Extraction of data tracks from CD-ROMs and DVDs is done with IsoBuster. Audio ripping is done with dBpoweramp. Since dBpoweramp only has a graphical interface, we contacted its author, and through a small development contract with the KB he wrote a command-line tool for the core ripping software. This enabled us to integrate dBpoweramp into the workflow as well. Ripped audio files are verified for completeness using either Shntool (WAVE format) or flac (Flac format). ISO images are checked with Isolyzer. Finally, operation of the disc robot is done using the dBpoweramp driver tools.⁷⁹

Nachdem innerhalb der Software die PPN – der Identifikator für Katalogsätze in der KBNL – angegeben wurde, kann die entsprechende Disc der Warteschlange hinzugefügt und der Prozess gestartet werden. Je nach erkanntem Format werden automatisiert andere Programme hinzugezogen, um bedarfsabhängig mit den Daten weiter zu verfahren, was auch die Entscheidung zwischen Disc Image und Migration einschließt. Das lässt einen Teil der Vorsortierung beziehungsweise der Nachbearbeitung wegfallen. Zuletzt werden die kopierten Daten direkt mit Metadaten aus dem Katalog angereichert. Auch

⁷⁶ vgl. ebd., S. 33

⁷⁷ vgl. van der Knijff, Johan (2017): Image and Rip Optical Media Like A Boss. Online unter <https://www.bitsgalore.org/2017/06/19/image-and-rip-optical-media-like-a-boss> [Abruf am 19.10.2023]

⁷⁸ Siehe Acronova Technology, Inc. (2018): Nimbie USB Plus. Zuletzt aktualisiert im Oktober 2018. Online unter <https://disc.acronova.com/product/auto-blu-ray-duplicator-publisher-ripper-nimbie-usb-nb21/9/review.html>

⁷⁹ Ebd.

wenn IROMLAB auf die spezifischen Anforderungen sowie Soft- und Hardwarekombination in der KBNL angepasst ist, wurde es auf GitHub bereitgestellt, damit sich auch andere damit befassen können.⁸⁰

5.4 New York Public Library

Die NYPL, eine der größten⁸¹ Bibliotheken der Welt, nutzt eine modifizierte Version von IROMLAB in Kombination mit einem Aconova Nimble NB21-BR⁸², um die Sicherung umfangreicher Bestände rationalisieren zu können. Ein großer Unterschied zu den beiden Vorbildern BL und KBNL ist, dass Migration über Emulation gestellt wird, unter anderem, weil der authentischen Datenstruktur ein niedrigerer Wert beigemessen wird und ein sehr hoher Anteil der zu bearbeitenden Datenträger ohnehin nur aus CD-Rs und DVD-Rs besteht, die grafische Materialien in Form von JPGs und RAWs enthalten, welche kein Disc Image erfordern. Das Programm IsoBuster als Teil von IROMLAB, der für die Erstellung von Images verantwortlich ist, wurde hierfür durch eine modifizierte Version vom Programm bagit-phyton ersetzt.⁸³

Um genauer bestimmen zu können, ob die transferierten Daten funktional sind, ohne im Anschluss alle einzeln zu testen, wurden nach etwa 3500 Transfers die Logdaten ausgewertet. Ziel war, eindeutige Indikatoren zu finden, die fehlerhafte Übertragungen verzeichnen.⁸⁴ Insgesamt wurden 46 gefunden, bei denen das physische Erscheinungsbild (Dreck, Kratzer etc.), die Daten auf der CD sowie die übertragenen Daten überprüft wurden. Die häufigsten Probleme waren Dateisystemprobleme und Beschädigungen. Für erstere würde dann doch IsoBuster eingesetzt, für zweitere wäre allerdings eine derart zeitaufwändige Wiederherstellung nötig gewesen, dass es sich für die zur Debatte stehende Sammlung nicht gelohnt hätte. Eine zweite Erkenntnis war, dass es hilfreich ist,

⁸⁰ Siehe <https://github.com/KBNLresearch/iromlab>

⁸¹ vgl. Wikipedia (2023): Liste der größten Bibliotheken der Erde. Zuletzt aktualisiert am 08.10.2023. Online unter https://de.wikipedia.org/wiki/Liste_der_gr%C3%B6%C3%9Ften_Bibliotheken_der_Erde [Abruf am 20.10.2023]

⁸² Siehe Acronova Technology, Inc. (2018)

⁸³ Rothrock et al. (2021): Assessing High-volume Transfers from Optical Media at NYPL. In: Code4Lib Journal 51, 2021. Online unter <https://journal.code4lib.org/articles/15908> [Abruf am 20.10.2023]

⁸⁴ Das verwendete Script wurde auf GitHub zur Verfügung gestellt: <https://github.com/chelle-rr/optical-media-check>

problematische Discs aus dem Prozess zu nehmen und im Anschluss gesammelt zu bearbeiten.⁸⁵

Insgesamt brauchte in der Testkollektion jede Disc durchschnittlich ~238 Sekunden Übertragungszeit. Bei 63 Discs wurde nicht auf Anhieb alle Daten übertragen und lediglich drei fielen völlig aus.

5.5 Deutsche Nationalbibliothek

Die DNB ist im deutschen Raum die Bibliothek, die das Pflichtexemplarrecht innehat und ist dementsprechend ebenfalls mit den Herausforderungen konfrontiert, die sehr große Bestände mit sich bringen. Schon 2006 startete ein Pilotprojekt des Deutschen Musikarchivs, das Teil der DNB ist, um erste Erfahrung mit der Migration von Audio-CDs zu sammeln.⁸⁶ Hierfür werden die Daten von den CDs herunterkopiert und in eine einzige große Datei im Broadcast-WAV-Format vereinigt. Beim Zugriff in den Lesesälen werden daraus automatisch wieder einzelne Tracks generiert. Bisher wurden etwa 600.000 Datenträger für diese Art der Bereitstellung umkopiert, allerdings bisher noch nicht dauerhaft im Langzeitarchiv abgelegt. Dieser Schritt ist in Planung.⁸⁷

Auch Daten-CDs werden bislang noch nicht langzeitarchiviert. Das momentane Bereitstellungssystem sieht vor, dass Datenträger, die aus dem Katalog bestellt werden, ausgehoben, verimaged und über ein kommerzielles System der Firma H+H Software GmbH⁸⁸ abgelegt werden, das die Bereitstellung für den Lesesaal übernimmt. Bisher geschieht das noch nicht über einen speziellen Emulator. Stattdessen kann es auch durchaus vorkommen, dass Inhalte nicht angezeigt werden können, weil der verwendete Terminalcomputer nicht kompatibel mit dem Datenformat des Images ist. Außerdem sind die Images ursprünglich nur als temporäre Kopien vorgesehen, auch wenn sie aktuell nicht gelöscht werden.⁸⁹

Die einzigen Disc Images, die bislang zur Langzeitarchivierung vorgesehen sind, wurden bis 2016⁹⁰ händisch im Rahmen der Entwicklung der Emulationssoftware EMiL erstellt.

⁸⁵ vgl. ebd.

⁸⁶ vgl. Groth (2007)

⁸⁷ vgl. Steinke (2023), 05:49-09:24

⁸⁸ siehe <https://www.virtualcd.de/>

⁸⁹ vgl. Steinke (2023), 10:20-12:22

⁹⁰ vgl. Deutsche Nationalbibliothek (2019)

EMiL ist in den Lesesälen der DNB zwar installiert, kann aber momentan nur für diese recht kleine Testmenge an Images verwendet werden. Darüber hinaus ist allerdings geplant, irgendwann das kommerzielle System mit EMiL zu koppeln, sodass also auch mit EMiL alle H+H-Images angezeigt werden können.⁹¹

Im Moment ist ein Projekt in Planung, um neben optischen auch magnetische Datenträger und andere Trägerformen zu sichern. Ein Problem hierbei ist, dass nicht einmal genau festgestellt werden kann, wie viele Datenträger welcher Art im Bestand sind, da in den Katalogsätzen häufig nur in Freitextfeldern der Systemtyp vermerkt ist, aber teils die Information völlig fehlt. Eine simple Katalogabfrage zur genauen Feststellung ist demnach also nicht möglich, weshalb es nötig wird, die aus dem Katalog extrahierbaren Informationen durch Magazinbegehungen zu ergänzen.⁹²

Sobald alle nötigen Informationen zusammengetragen sind, ist es geplant, eine Ausschreibung zu starten und das Projekt an einen externen Dienstleister zu übergeben.⁹³

5.6 Technische Informationsbibliothek Hannover

Die TIB ist neben ZBW und ZB MED eine der drei zentralen Fachbibliotheken in Deutschland. Ihr Fokus liegt auf der umfassenden Sammlung von Literatur und Informationen aus den Bereichen Technik und Naturwissenschaften.⁹⁴ Die genauen Bestände an CDs sind aus ähnlichen Gründen wie bei der DNB nicht genau bestimmbar: Nur einzeln verzeichnete Datenträger lassen sich im Katalog abfragen. Beilagen-CDs sind falls überhaupt in nicht maschinell auswertbaren Freitextfeldern angegeben. Schätzungen aus dem Jahr 2014 kamen auf etwa 40.000 bis 70.000 Datenträger.⁹⁵

Überlegungen, Disc Images von optischen Datenträgern anzufertigen, gab es bereits 2004, als ein Proof-of-Concept zur Imageerstellung erarbeitet wurde, wobei Emulation-as-a-Service der Uni Freiburg als Emulator genutzt wurde. Das System funktionierte zwar, dennoch wurde der Ansatz nicht weiter verfolgt. Grund dafür war das Fehlen von

⁹¹ vgl. Steinke (2023), 16:36-18:45

⁹² vgl. ebd., 11:53-15:56

⁹³ vgl. ebd., 26:17-27:33

⁹⁴ vgl. TIB (o. J.): Die TIB. Mehr als eine Bibliothek. Online unter <https://www.tib.eu/de/die-tib/profil-der-tib> [Abruf am 20.10.2023]

⁹⁵ vgl. Schwab (2023), 20:50-25:15

Datenträgern mit komplexer Software, die Emulation als komplexere Lösung nötig machen würde.⁹⁶

Ein weiteres prototypisches Archivierungsprojekt erfolgte erst im Jahr 2017 und umfasste etwa 14.000 CD-Rs oder CD-RWs, die damals bereits mindestens 10 Jahre alt waren.⁹⁷ Man hat sich zur Bearbeitung für die Migration entschieden, weil es sich um reine Daten-CDs ohne ausführbare Software handelt.⁹⁸ Ausführendes Gerät war auch hier ein Acronova Nimble zusammen mit „Disc to Computer“, einer mitgelieferten Kopiersoftware.⁹⁹ Außerdem wurde folgendes verwendet:

„[F]ür die Fehleranalyse verschiedenste Tools zur Imageerstellung. In der Teststellung haben wir da ausprobiert: das dd-Kommando unter Linux und verschiedene Tools mit grafischer Oberfläche, in dem Fall haben wir geguckt, welche Ergebnisse uns UltraISO bringt, IsoBuster und ImgBurn, aber da befinden wir uns immer noch in der prototypischen Umgebung. Im Fehlerfall war ich dann ein großer Freund der Software dvisaster. Damit kann ich mir nämlich grafisch darstellen lassen, welche Sektoren auf der CD gerade hinüber sind. Und wir haben verschiedene CD-Laufwerke benutzt. Das ist ein spannendes Feld. In Abhängigkeit davon, wie gut die Fehlerkorrektur des Laufwerks ist, bekomme ich unterschiedliche Ergebnisse beim Kopieren.“¹⁰⁰

Nach der Bearbeitung und Fehlerkorrektur werden die Datensätze gemeinsam mit einer Kopie der Metadaten des zugehörigen Katalogsatzes in das Langzeitarchiv eingespeist. Genutzt wird die kommerzielle Archivlösung Rosetta von Ex Libris.¹⁰¹ Die Server sowie die zugehörige Infrastruktur wird von der TIB selbst gehostet, wovon auch die ZBW und die ZB MED profitieren, die mit der TIB in einem Verbund Langzeitarchivierung betreiben.¹⁰²

Abgesehen vom prototypischen Projekt ist auch schon ein Produktivworkflow in der Konzeptionsphase, um auch den restlichen Bestand anzugehen. Die Planung zieht sich allerdings aufgrund von begrenztem Personal.¹⁰³

⁹⁶ vgl. ebd., 1:07:43-1:09:59

⁹⁷ vgl. ebd., 08:59-18:28

⁹⁸ vgl. ebd., 05:53-6:47

⁹⁹ vgl. ebd., 40:32-41:03

¹⁰⁰ ebd., 41:03-42:31

¹⁰¹ vgl. ebd. 57:13-58:16

¹⁰² vgl. ebd., 11:54-13:01

¹⁰³ vgl. ebd., 07:09-7:55

5.7 Landesarchiv Baden-Württemberg

Das Landesarchiv Baden-Württemberg ist eines von 16 Landesarchiven in Deutschland.¹⁰⁴ Als solches sammelt und archiviert es Unterlagen und Akten von Behörden, wobei CDs seit den 90ern primär als Beilage von Akten mit ins Haus kommen.¹⁰⁵ Nachdem diese Zugänge aus ausfallgefährdeten CD-Rs bestehen, ist eine relativ zügige Bearbeitung wichtig, um Datenverluste zu verhindern.¹⁰⁶ Gepresste und damit haltbarere CDs machen einen eher geringen Anteil aus. Sie kommen lediglich dann in den Bestand, wenn Nachlässe übernommen werden.¹⁰⁷ Die Archivierung erfolgte kontinuierlich parallel zum Eingang. Anders als in allen befragten Bibliotheken hat sich im Landesarchiv somit kein übermäßiger Rückstand an unverarbeiteten Datenträgern aufgebaut, der nun in einem Projekt abgearbeitet werden muss.¹⁰⁸ Eingelesen werden die Daten einzeln über USB-Laufwerke.¹⁰⁹

Besonders ist hier, dass das gewählte Bearbeitungsmittel grundsätzlich dasjenige der Migration ist und Emulation überhaupt keine Rolle spielt.¹¹⁰ Die CDs selbst haben keinen eigenen Wert für die Einrichtung, sie dienen rein als Container für die enthaltenen Daten.¹¹¹ Das bedeutet, dass die Erstellung von Disc Images bislang eher uninteressant ist.

Für den Migrationsvorgang an sich gibt es eine Handreichung,¹¹² die alle Materialien und Schritte der Bearbeitung vom Ausheben des Trägers über das Bewerten, Kopieren und Speichern der Daten bis zum letztendlichen Entsorgen der physischen Variante beschreibt. Seit der Vorstellung auf einer Tagung 2019 wurde der Workflow von etwa 200 anderen Archiven im DIMAG-Verbund aufgegriffen und mit Änderungen in den eigenen Arbeitsablauf integriert.¹¹³ Wichtig ist hier allerdings, zu betonen, dass es sich um eine Stichprobe handelt, die sich nicht auf das komplette deutsche Archivsystem übertragen

¹⁰⁴ vgl. Haardt (2023), 36:15-36:29

¹⁰⁵ vgl. ebd., 01:30-02:21

¹⁰⁶ vgl. ebd., 38:39-39:44

¹⁰⁷ vgl. ebd., 04:12-05:00

¹⁰⁸ vgl. ebd., 15:40-16:18

¹⁰⁹ vgl. ebd., 17:10-17:53

¹¹⁰ vgl. ebd., 06:43-07:04

¹¹¹ vgl. ebd., 05:34-06:09

¹¹² Siehe Staatsarchiv Ludwigsburg (2018): Ludwigsburger Handreichung Datenträger. Online unter: <https://www.sg.ch/kultur/staatsarchiv/Spezialthemen-/auds/2019.html> [Abruf am 16.10.2023]

¹¹³ vgl. Haardt (2023), 09:15-13:37

lässt. Denn hier handelt es nicht um eine von höherer Stelle angeordnete Arbeitsweise, sondern es haben sich die einzelnen Archive den Ablauf selbst angeeignet.¹¹⁴

5.8 Zentrum für Kunst und Medien Karlsruhe

Das ZKM ist eine von nur zwei musealen Einrichtungen, die Mitglied von nestor sind.¹¹⁵ Tatsächlich greift der Begriff „Museum“ zu kurz, weil im ZKM versucht wird, alle Medien und Gattungen zu bedienen, neben Malerei und Fotografie zählen beispielsweise auch Video, Musik, Tanz und Performance zum behandelten Spektrum.¹¹⁶ Konsequenz hieraus sind verschiedenste Sammlungen mit Bewahrungsnotwendigkeit, zu denen auch optische Datenträger gehören. Darüber hinaus teilt sich das ZKM die Bibliothek mit der Staatlichen Bibliothek für Gestaltung in Karlsruhe.¹¹⁷

Das ZKM unterscheidet im Bestand zwischen Sammlungen und erworbenen Archiven. In den Sammlungen finden sich etwa 7000 DVDs, davon sind 5000 Kopien der Sammlungswerke und die restlichen 2000 Kaufwerke. Bislang liegen etwa 1100 von 7000 entkörperlicht vor. Im Archiv lagern weitere 1400 DVDs und 200 CDs, wovon etwa 800 bis 900 transferiert sind.¹¹⁸ Bei der Archivierung kommen sowohl Disc Images als auch Migrierung zum Einsatz, das richtet sich nach der Art der Datenträger.¹¹⁹ Es werden keine Disc Robots genutzt, um den Prozess zu beschleunigen, stattdessen wird alles mithilfe von handelsüblichen Laufwerken eingespielt.¹²⁰

Gespeichert werden die Inhalte anschließend unkomprimiert auf Magnetband, was ungewöhnlich ist, weil sich kaum Einrichtungen unkomprimierte Speicherung aufgrund des erhöhten Speicherbedarfs leisten können. Dass das hier möglich ist, ist historisch gewachsen.¹²¹

¹¹⁴ vgl. ebd., 35:12-36:56

¹¹⁵ vgl. nestor e.V. (2023): Mitglieder. Online unter https://www.langzeitarchivierung.de/Webs/nestor/DE/nestor/Mitglieder/mitglieder_node.html [Abruf am 21.10.2023]

¹¹⁶ vgl. ZKM (o. J.): Das ZKM. Online unter <https://zkm.de/de/das-zkm> [Aufruf am 21.10.2023]

¹¹⁷ Siehe <https://zkm.de/de/sammlung-archive/bibliothek>

¹¹⁸ vgl. Haardt (2023), 01:36-03:03

¹¹⁹ vgl. ebd., 14:14-14:55

¹²⁰ vgl. ebd., 16:27-16:43

¹²¹ vgl. ebd., 10:14-11:08

5.9 Herzog August Bibliothek Wolfenbüttel

Die HAB ist eine Forschungsbibliothek, deren Kernkompetenzen in der Bewahrung und Erschließung historischer Bestände liegen.¹²² Der Bestand an optischen Datenträgern ist schwer zu bestimmen: Es sind 450 bis 500 CDs und DVDs mit Einzelsignaturen nachgewiesen, die Mengen an Beilagen-CDs sind aufgrund der fehlenden Erfassung allerdings nicht bekannt.¹²³ Das Projekt zur systematischen Bearbeitung wurde nach Personalwechseln im bibliothekarischen Team und in der EDV schon in der Planungsphase eingestellt.¹²⁴

Dennoch findet in der HAB ähnlich wie in der DNB eine Imageerstellung im Zuge der Bereitstellung statt. Hierfür wird ein Client-Programm der Firma Citrix genutzt: In der HAB wird eine CD eingelesen und ein Image erstellt, das Endnutzer durch Einwahl in den Client abrufen können. Gespeichert wird das Image auf einem Server, der allerdings derart alt und instabil ist, dass schon länger diskutiert wird, ihn abzuschalten.¹²⁵ Anfragen an das Hochschulbibliothekszenrum des Landes Nordrhein-Westfalen um das Jahr 2009 oder 2010 herum, die eine Übernahme des Serverhostings erbat, blieben erfolglos.¹²⁶

6. Zusammenfassung

In dieser Arbeit wurde aufgezeigt, dass sich Compact Discs allmählich zu einem archivarischen Problem entwickeln. Einerseits sind sie technisch überholt, andererseits kann ihre Abspielbarkeit nicht auf Dauer gewährleistet werden, wenn Abspielgeräte seltener werden und sich die Datenträger zersetzen. Es wurde klargestellt, dass bei beschreibbaren Compact Discs wie CD-R und CD-RW ein stärkerer Handlungsbedarf besteht als bei gepressten Trägern wie Audio-CDs und CD-ROMs, dass die tatsächliche Dringlichkeit allerdings stark von der Nutzungsintensität und den Aufbewahrungsbedingungen abhängt.

Die anschließende Datenerhebung umfasst Interviews mit deutschen Gedächtnisorganisationen, von denen die meisten Mitglied von nestor e.V. sind. Der Fragenkatalog der

¹²² vgl. Herzog August Bibliothek (o. J.): Bibliothek. Online unter <https://www.hab.de/bibliothek/> [Aufruf am 23.10.2023]

¹²³ vgl. Jorzenuk (2023), 02:47-03:24

¹²⁴ vgl. ebd., 01:26-02:17

¹²⁵ vgl. ebd., 09:23-12:03

¹²⁶ vgl. ebd., 14:09-14:32

Interviews basiert auf der Literaturarbeit und deckt verschiedene Aspekte der Langzeitarchivierung ab. Zuletzt folgte die Auswertung der Interviews und der Literaturarbeit.

Die British Library muss eine Sammlung mit einem Umfang von 18.000 CDs bewältigen und greift daher auf einen teilautomatisierten Workflow zurück. Verschiedene Disc Robots wurden getestet, von denen der Acronova Nimbie NB11 ausgewählt wurde. Die gewünschte Vorlageform sind emulierbare Disc Images im ADL-Format (AES31-3). Insgesamt verarbeitet die BL etwa 1050 CDs pro Monat.

Die KBNL hat etwa 15.000 bis 20.000 Datenträger, hauptsächlich CD-ROMs. Aufbauend auf den Vorarbeiten der BL wird der Acronova Nimbie NB21-DVD in Kombination mit dem selbstentwickelten Programm IROMLAB verwendet. Dieses Programm entscheidet je nach erkanntem Format eigenständig, ob migriert oder ein Disc Image erstellt wird und reichert den Datensatz direkt mit Metadaten aus dem Katalog an. IROMLAB wurde auf GitHub bereitgestellt.

Die NYPL nutzt eine modifizierte Version von IROMLAB in Kombination mit einem Acronova Nimbie NB21-BR für die Langzeitarchivierung. Die Daten werden migriert, da der zu bearbeitende Bestand überwiegend aus CD-Rs und DVD-Rs besteht. Um die Fehleranalyse stärker automatisieren zu können, werden Logdaten des Auswertungsprozesses hinzugezogen, um fehlerhafte Übertragungen zu identifizieren.

Die DNB hat das Pflichtexemplarrecht und startete 2006 das erste Pilotprojekt zur Migration von Audio-CDs, als im Deutschen Musikarchiv, einem Teil der DNB, erste CD-Ausfälle auftraten. Aktuell wird keine systematische Imageerstellung betrieben, die einzigen Disc Images zur Langzeitarchivierung entstanden im Rahmen der Entwicklung der Emulationssoftware EMiL. Die DNB plant ein Projekt zur Sicherung magnetischer Datenträger.

Die TIB hat bereits 2017 ein prototypisches Migrationsprojekt mit etwa 14.000 CD-Rs oder CD-RWs unterstützt von einem Acronova Nimbie durchgeführt. Die angereicherten Datensätze wurden mithilfe der Software Rosetta auf einem selbst betriebenen Server archiviert, den auch ZBW und ZB MED mitnutzen. Ein Produktivworkflow, um den Rest der 40.000-70.000 Discs zu bearbeiten ist in der Konzeptionsphase.

Das Landesarchiv Baden-Württemberg archiviert CDs als Beilagen von Akten. Nachdem eintreffende Datenträger konstant händisch migriert werden, gibt es keinen übermäßigen Rückstand an unverarbeiteten Datenträgern. Emulation spielt keine Rolle.

Das ZKM besitzt rund 8600 optische DVDs und CDs, von denen etwa 1900 bis 2000 bereits entkörperlicht sind. Je nach Datenträgertyp wird zwischen Migration und Erstellung von Disc Images entschieden, gespeichert werden die Daten ungewöhnlicherweise unkomprimiert auf Magnetbändern. Alle Kopien werden händisch ohne Einsatz von Disc Robots durchgeführt.

Die HAB hatte ein Projekt in Planung, um ihren Bestand von über 450 optischen Datenträgern zu kopieren, das Projekt wurde allerdings schon in der Planungsphase abgebrochen. Dennoch erfolgt eine anfragegesteuerte Erstellung von Disc Images zum Zweck der Bereitstellung über einen Citrix-Client. Es ist momentan nicht sicher, wie lange der verwendete Server mit den vorgehaltenen Images noch betrieben werden kann.

Ursprünglich war geplant, Workflows zu recherchieren und zu vergleichen, um einen Musterworkflow zu erstellen. Allerdings stellte sich heraus, dass nur fünf von 13 Institutionen solche Projekte haben oder hatten, weshalb eine detaillierte Beschreibung der individuellen Vorgehensweisen letztendlich zielführender war. Ohnehin hätte ein einziger Musterworkflow keinen wirklichen Sinn, weil sehr stark von der Bestandsgröße und der Bestandszusammensetzung abhängig ist, welche Lösungen einen Mehrwert bieten. Disc Roboter beispielsweise sind nur für große Institutionen nötig. Die Frage, ob Migration oder Emulation vorzuziehen ist, richtet sich daran aus, ob auf den Trägern statische Dateien oder Softwares enthalten sind. Dennoch hat sich gezeigt, dass die Problematik mindestens in Grundzügen bekannt ist. Mehrere Einrichtungen haben Projekte in der Planung, andere haben sich bewusst dagegen entschieden, weil Schwierigkeiten auftraten oder weil die Priorisierung anderweitig gelagert ist. Abzuwarten ist, ob sich diese Haltung noch ändert, wenn sich irgendwann Ausfälle häufen.

Problematischer ist, dass sich die meisten Institutionen auf eigene Faust mit der Thematik befassen und sich dafür kaum zusammenschließen. Ausnahmen sind der Verbund der drei zentralen Fachbibliotheken, die Kooperation von 200 Archiven im DIMAG-Verbund und die Zusammenarbeit mehrerer Bibliotheken zur Entwicklung von EMiL. Es ist schwierig, von Stichproben Schlüsse auf die Situation in ganz Deutschland zu ziehen,

aber wenn selbst viele Mitglieder eines relevanten Langzeitarchivierungsnetzwerks noch keine Projekte zur Archivierung optischer Datenträger haben, ist es unwahrscheinlich, dass die Quoten außerhalb besser sind. Eine gemeinsame Herangehensweise senkt die Hürden wie beispielsweise Recherche- und Vorbereitungszeit für einzelne Einrichtungen, bis mit der Bearbeitung der eigenen Bestände begonnen werden kann.

7. Ausblick

In diesem Kapitel sollen drei Vorschläge mitgegeben werden, wie man Langzeitarchivierung optischer Datenträger in Deutschland in den kommenden Jahren durch Zusammenarbeit verschiedener Institutionen voranbringen könnte: Unterstützung unerfahrener Einrichtungen durch solche, die sich gut auskennen, Einheitliche Erfassung bereits archivierter Werke und Ausarbeitung eines Kriterienkatalogs zur Priorisierung.

7.1 Unterstützung unerfahrener Einrichtungen

Wie beispielsweise aus 5.9 ersichtlich ist, verfügen nicht alle Gedächtnisorganisationen über die Zeit und das geschulte Personal, um selbst Workflows zur Archivierung zu entwerfen. Die Recherche ist aufwändig, die technische Umsetzung erfordert sehr spezielle IT-Kompetenzen und nicht jede Einrichtung hat ein leistungsfähiges Langzeitarchiv.

Zusammenarbeiten können hier auf verschiedene Arten geschehen:

1. Gemeinsamer Betrieb technischer Infrastruktur wie Servern für digitale Langzeitarchivierung kann Ressourcen bündeln und teure Doppelstrukturen reduzieren, was der Verbund der zentralen Fachbibliotheken zeigt.¹²⁷
2. Aktive Hilfe durch erfahrene Institutionen bei der Einrichtung eigener Workflows, besonders was Hard- und Softwareauswahl sowie deren Abstimmung auf lokale Eigenheiten angeht, weil genannte Schritte viel Recherche benötigen kann.
3. Vollständige Übernahme des Archivierungsprozesses durch Dienstleister ist ebenfalls vorstellbar. Das können sowohl kommerzielle Dienstleister sein, die

¹²⁷ Schwab (2023), 11:54-13:01

keine Gedächtnisorganisationen sind als auch erfahrenere Institutionen. Ein gutes Beispiel hierfür ist die HAB, die ihre Expertise bei der Digitalisierung von Handschriften und alten Drucken regelmäßig anderen Bibliotheken mit Altbeständen zur Verfügung stellt, die nicht die entsprechenden Gegebenheiten vor Ort haben.¹²⁸

4. Es ist möglich, Verbände zum Austausch untereinander und zur Zusammenarbeit entweder neu zu gründen oder die Aktivitäten von bestehenden Zusammenschlüssen zu intensivieren, beispielsweise im nestor-Netzwerk.

7.2 Datenbank für archivierte Werke

Im Moment werden bereits archivierte Werke meist nur im eigenen Katalog nachgewiesen. Auf längere Sicht könnte es sich allerdings lohnen, eine Nationalbibliografie aufzubauen, die verzeichnet, welche Werke bereits langzeitarchiviert vorliegen. Das Vorbild hierbei ist die Sammlung deutscher Drucke,¹²⁹ in der nach und nach eine retrospektive deutsche Nationalbibliografie aufgebaut wird. In einem Äquivalent für digitale Datenträger ginge es bewusst nicht darum, zu verzeichnen, wer die physischen Datenträger besitzt, vielmehr stünde im Vordergrund, alle Disc Images und migrierten Datensätze nachzuweisen. Das hätte zwei Vorteile:

1. Wenn man als Endnutzer nach einem bestimmten Werk sucht, muss man nicht darauf vertrauen, dass der physische Datenträger in der Einrichtung, die ihn nachweist, auch funktioniert. Stattdessen kann man sich direkt zu einer Einrichtung begeben, die über eine archivierte Variante verfügt. Das wird immer relevanter, umso ausfallgefährdeter die CDs aufgrund ihres Alters werden.
2. Gedächtnisorganisationen können besser priorisieren, welche Datenträger zuerst bearbeitet werden müssen: Es ist einfacher zu beurteilen, wie selten bestimmte Werke sind und welche deswegen früher bearbeitet werden müssen, um einen endgültigen Verlust zu verhindern.

¹²⁸ Herzog August Bibliothek (o. J.): Digitalisierung der mittelalterlichen Handschriften der Staats- und Universitätsbibliothek Bremen. In Kooperation mit der Staats- und Universitätsbibliothek Bremen. Online unter <https://www.hab.de/digitalisierung-der-mittelalterlichen-handschriften-der-staats-und-universitaetsbibliothek-bremen/> [Abruf am 23.10.2023]

¹²⁹ Siehe https://www.ag-sdd.de/Webs/agsdd/DE/Home/home_node.html

Ein großer Nachteil gegenüber dem Verzeichnis Deutscher Drucke ist, dass man in einem Verzeichnis optischer Datenträger zwar den Besitznachweis unterbringen kann, nicht aber die Images und migrierten Daten selbst. Das hängt damit zusammen, dass das deutsche Urheberrecht alle bislang entstandenen CDs noch auf mehrere Jahrzehnte so stark schützt, dass ein Abruf weiterhin nur im Lesesaal erlaubt sein wird. Eine einzige Ausnahme ist, dass manche Inhalte im Rahmen einer Dokumentlieferung auf Anfrage bereitgestellt werden dürfen.¹³⁰

8. Fazit

Die Arbeit hat als vorrangiges Ziel, zu dokumentieren, wie verschiedenste Gedächtnisinstitutionen mit der Problematik umgehen, dass ihr CD-Bestand nicht auf Dauer erhalten bleibt, wenn nicht in absehbarer Zeit Erhaltungsmaßnahmen eingeleitet werden. Es konnten sehr unterschiedliche Herangehensweisen aufgezeigt werden, die verdeutlichen, dass zwar einige Abläufe vereinheitlicht werden können und dafür Kooperationen und Verbünde ein gutes Mittel wären, dass aber in der genauen Ausgestaltung für einen allgemeinen Musterworkflow die Voraussetzungen in den einzelnen Einrichtungen zu unterschiedlich sind.

Außerdem wurde dargestellt, dass gerade bei gebrannten CD-Rs die Zeit drängt, Maßnahmen zu ergreifen, davon abgesehen es aber auf die Prioritäten in den jeweiligen Institutionen ankommt, wenn es darum geht, aktiv mit der Sicherung zu beginnen.

Und zuletzt wird deutlich, wie stark der Status Quo im Wandel ist. Manche Einrichtungen stecken bereits in Projekten zur Archivierung, andere Bibliotheken planen gerade aktiv neue Projekte. Es festigt sich ein Bewusstsein für das Thema und die Hoffnung besteht, dass diese Arbeit noch weiter dazu beitragen kann und für unerfahrenere Institutionen zusätzliche Hilfestellung geben kann, tiefer in die Thematik einzusteigen.

¹³⁰ Schwab (2023), 31:17-33:40

9. Literaturverzeichnis

Acronova Technology, Inc. (o. J.): DupliQ. Online verfügbar über Archive.org unter https://web.archive.org/web/20120118073209/http://www.acronova.com/duplicator_dupliq_usb.htm [Abruf am 18.10.2023]

Acronova Technology, Inc. (2012): One's Enough! Nimbie USB. Zuletzt aktualisiert im März 2012. Online verfügbar über Archive.org unter <https://web.archive.org/web/20180424024231/http://www.acronova.com/product/nimbie-usb-auto-loader-publisher-ripper-nb11/2/review.html> [Abruf am 18.10.2023]

Acronova Technology, Inc. (2018): Nimbie USB Plus. Zuletzt aktualisiert im Oktober 2018. Online unter <https://disc.acronova.com/product/auto-blu-ray-duplicator-publisher-ripper-nimbie-usb-nb21/9/review.html> [Abruf am 20.10.2023]

Adminius B.V. (o. J.): CD Products. Online unter https://www.lscdweb.com/ordering/cd_products.html [Abruf am 18.10.2023]

Bundesverband Musikindustrie e.V. (2023): Musikindustrie in Zahlen 2022. Online unter <https://www.musikindustrie.de/wie-musik-zur-karriere-werden-kann/markt-bestseller/musikindustrie-in-zahlen-2022>

Byrne, Alex (2014): Memory institutions shaping the past, present and future. In: Proceedings der Tagung IFLA PAC, Geneva, am 13. und 14. August 2014, Genf

COPTR (2021): About COPTR. Zuletzt aktualisiert am 03.10.2021. Online unter https://coptr.digipres.org/index.php/About_COPTR [Abruf am 18.10.2023]

COPTR (2021): Workflow:Community Owned Workflows. Zuletzt aktualisiert am 28.04.2021. Online unter https://coptr.digipres.org/index.php/Workflow:Community_Owned_Workflows [Abruf am 18.10.2023]

Dappert, Angela et al. (2011): Developing a Robust Migration Workflow for Preserving and Curating Hand-held Media. In: Proceedings der Tagung iPRES 2011 - 8th International Conference on Preservation of Digital Objects, 01.-04. November 2011, Singapur, S. 33-43

Data Archive Corporation (o. J.): DISC DVD-7000 DVD Library. Online verfügbar über Archive.org unter <https://web.archive.org/web/20110509062816/http://www.dataarchivecorp.com/disc-dvd-7000.htm> [Abruf am 18.10.2023]

Deutsche Nationalbibliothek (2019): EMiL. Bereitstellung von Multimedia-Objekten durch Emulation. Online unter https://www.dnb.de/DE/Professionell/ProjekteKooperationen/Projektarchiv/2016/emulationMultimediaObjekte/emulationMultimediaObjekte_node.html [Abruf am 10.10.2023]

Gieselmann, Hartmut (2008): Silberne Erinnerungen: Archiv-DVDs im Langzeittest. In: c't - Magazin für Computertechnik 16, 2008, S. 116ff

Groth, Susanne, Alexander Uhl (2007): Deutsches Musikarchiv: CD-Zerfall bedroht Kulturerbe. Online unter <https://www.heise.de/news/Deutsches-Musikarchiv-CD-Zerfall-bedroht-Kulturerbe-148962.html> [Abruf am 16.10.2023]

Haardt, Christian (2023): ZKM-Interview. Zoom-Interview mit Christian Haardt, Labor für antiquierte Videosysteme im Zentrum für Kunst und Medien Karlsruhe, durchgeführt von Felix Kratzer am 13.09.2023, Anhang S. A15-A24

Hardwarejournal.de (2013): Wie hoch ist die Lebensdauer von CD- und DVD-Rohlingen?. Online unter <https://www.hardwarejournal.de/cd-lebensdauer-prognose/> [Abruf am 06.10.2023]

Herzog August Bibliothek (o. J.): Digitalisierung der mittelalterlichen Handschriften der Staats- und Universitätsbibliothek Bremen. In Kooperation mit der Staats- und Universitätsbibliothek Bremen. Online unter <https://www.hab.de/digitalisierung-der-mittelalterlichen-handschriften-der-staats-und-universitaetsbibliothek-bremen/> [Abruf am 23.10.2023]

Hjerpe, Roland (1994): A Framework for the Description of Generalised Documents. In: Proceedings der Tagung ISKO '94 Conference, 21. bis 24. Juni 1994, Kopenhagen

Hübsch, Gerald (2023): AW: Interview im Rahmen meiner Bachelorarbeit. E-Mail an Felix Kratzer, gesendet am 24.09.2023, Anhang S. A58

IVD (2020): Anzahl der Kunden im DVD/Blu-ray-Kaufmarkt in Deutschland in den Jahren 2007 bis 2019. In Statista. Online unter

<https://de.statista.com/statistik/daten/studie/168578/umfrage/anzahl-der-kunden-im-dvd-markt-seit-2007/> [Abruf am 06.10.2023]

Jorzenuk, Saskia (2023): HAB-Interview. Zoom-Interview mit Saskia Jorzenuk, Bereich Neuere Medien der Herzog August Bibliothek, durchgeführt von Felix Kratzer am 11.09.2023, Anhang S. A10-A14

Lang, Patrick (2021): Vom Aussterben bedroht. Wo es noch CD-Player gibt und wo nicht. Online unter <https://www.auto-motor-und-sport.de/tech-zukunft/cd-player-autos-modelle-ausstattung/> [Abruf am 06.10.2023]

Markus, Katharina (2023): AW: AW: AW: Anfrage zum Zweitgutachten für meine Bachelorarbeit. E-Mail an Felix Kratzer, gesendet am 18.09.2023, Anhang S. A61

Meyer, Ulrich (2010): Ausstellung zeigt wahre Computerschätze. Kennen Sie noch C64, Amiga und Co.?. Online unter https://rp-online.de/digitales/neuheiten/ausstellung-zeigt-wahre-computerschaetze_aid-12875925 [Abruf am 06.10.2023]

Mills, Matt (2022): Warum hat ein aktueller Laptop kein CD- oder DVD-Laufwerk?. Online unter <https://itigic.com/de/why-doesnt-a-current-laptop-include-a-cd-or-dvd-drive/> [Abruf am 06.10.2023]

Naumann, Kai (2023): LABW-Interview. Zoom-Interview mit Dr. Kai Naumann, Referatsleiter Archivisches Recht im Landesarchiv Baden-Württemberg, durchgeführt von Felix Kratzer am 13.09.2023, Anhang S. A25-A36

nestor e.V. (2022): Über uns. Online unter https://www.langzeitarchivierung.de/Webs/nestor/DE/nestor/Ueber_uns/ueber_uns_node.html [Abruf am 18.10.2023]

nestor e.V. (2023): Mitglieder. Online unter https://www.langzeitarchivierung.de/Webs/nestor/DE/nestor/Mitglieder/mitglieder_node.html [Abruf am 21.10.2023]

Noack, Pit (2018): Die Zukunft der Archive: Games nicht nur bewahren, sondern spielbar erhalten. Online unter <https://www.heise.de/news/Die-Zukunft-der-Archive-Games-nicht-nur-bewahren-sondern-spielbar-erhalten-4259331.html?seite=all> [Zugriff am 21.10.2023]

Pennock, Maureen (2014): HOME. Online unter <http://wiki.opf-labs.org/display/SPR/Home> [Abruf am 18.10.2023]

PwC (2023): Umsätze mit T-VoD und S-VoD in Deutschland in den Jahren 2012 bis 2027*. In Statista. Online unter <https://de.statista.com/statistik/daten/studie/770083/umfrage/umsaetze-im-vod-markt-in-deutschland-nach-segmenten/> [Abruf am 06.10.2023]

Schoger, Astrid (2023): Antw: Wtrlt: Interview zu LZA im Rahmen meiner Bachelorarbeit. E-Mail an Felix Kratzer, gesendet am 07.09.2023, Anhang S. A55

Schwab, Franziska (2023): TIB-Interview. Zoom-Interview mit Franziska Schwab, Preservation Analyst im Team Langzeitarchivierung der TIB, durchgeführt von Felix Kratzer am 17.10.2023, Anhang S. A37-A54

Schweibenz, Werner; Sieglerschmidt, Jörn (2010): Aktuelle Entwicklungen bei Kulturportalen: BAM-Portal, Deutsche Digitale Bibliothek und Europeana. In: Proceedings der Tagung Inventarisierung digital - schmale und breite Wege in den Datenhimmel des Museumsverbandes Baden-Württemberg e.V. am 7. und 8. Mai 2010 im Landesmuseum Württemberg, Stuttgart

Škarić, Doris (2023): WG: Interview zu LZA im Rahmen meiner Bachelorarbeit. E-Mail an Felix Kratzer, gesendet am 15.09.2023, Anhang S. A58

Steinke, Tobias (2023): DNB-Interview. Zoom-Interview mit Tobias Steinke, Deutsche Nationalbibliothek, durchgeführt von Felix Kratzer am 08.09.2023, Anhang S. A1-A9

Strathmann, Stefan (2023): AW: Interview im Rahmen meiner Bachelorarbeit. E-Mail an Felix Kratzer, gesendet am 22.09.2023, Anhang S. A56

TIB (o. J.): Die TIB. Mehr als eine Bibliothek. Online unter <https://www.tib.eu/de/die-tib/profil-der-tib> [Abruf am 20.10.2023]

Tunnat, Yvonne (2023): Re: Interview zu LZA im Rahmen meiner Bachelorarbeit. E-Mail an Felix Kratzer, gesendet am 19.09.2023, Anhang S. A62

van der Knijff, Johan (2017): Image and Rip Optical Media Like A Boss. Online unter <https://www.bitsgalore.org/2017/06/19/image-and-rip-optical-media-like-a-boss> [Abruf am 19.10.2023]

Van Dijk, Boudewijn; Nijboer, Jaap: The CD Family (1996), S. 31. Online unter <https://www.lscdweb.com/data/downloadables/7/7/the-cd-family-3122-783-0068-1.pdf> [Abruf am 18.10.2023]

Weig, Heide-Marie (2023): Antw: [EXT] Interview im Rahmen meiner Bachelorarbeit. E-Mail an Felix Kratzer, gesendet am 25.09.2023, Anhang S. A59

Wikipedia (2022): Bitstrom. Zuletzt aktualisiert am 22.05.2022. Online unter <https://de.wikipedia.org/wiki/Bitstrom> [Abruf am 19.10.2023]

Wikipedia (2023): Liste der größten Bibliotheken der Erde. Zuletzt aktualisiert am 08.10.2023. Online unter https://de.wikipedia.org/wiki/Liste_der_gr%C3%B6%C3%9Ften_Bibliotheken_der_Erde [Abruf am 20.10.2023]

Wortmann, Christian (2023): AW: Interview im Rahmen meiner Bachelorarbeit. E-Mail an Felix Kratzer, gesendet am 22.09.2023, Anhang S. A57

ZKM (o. J.): Das ZKM. Online unter <https://zkm.de/de/das-zkm> [Aufruf am 21.10.2023]