

Abspeichern statt Abheften – den digitalen Schuhkarton sortieren

Leitfaden für den Einstieg in die Archivierung digitaler Daten

Annika Mühling, Pia Piontkowitz, Gabriele Pütz, Lukas Wähler

Durch den digitalen Wandel wird täglich eine Vielzahl persönlicher digitaler Objekte produziert. Einer Archivierung für den privaten Gebrauch stehen nicht nur die unübersichtliche Masse, sondern auch fehlende Kenntnisse im Bereich der Dateiformate, Speichermedien und Metadaten im Wege. In einem Projekt an der TH Köln unter Leitung von Prof. Achim Obwald erarbeiteten Studierende der Bibliothekswissenschaft Leitfäden für die persönliche digitale Langzeitarchivierung und organisierten einen Workshop in der Stadtbibliothek Köln um diesen Problem entgegenzutreten. Dieses Projekt wurde mit dem TIP-Award 2017 ausgezeichnet.

Due to the digital transition a multitude of personal digital objects is created daily. Archiving for the private use is made complicated not only by the sheer mass of objects, but also due to a lack of knowledge in the fields of file formats, storage media and metadata. In a project of the Technical University of Cologne under the direction of Prof. Achim Obwald, students of library science created guidelines for personal digital long-term archiving and organized a workshop in the public library of Cologne. This project received the TIP-Award 2017.

Digitale Langzeitarchivierung und Bibliotheken werden bisher fast ausschließlich mit der Digitalisierung und Sicherung von schätzenswerten Beständen in Zusammenhang gebracht. Insbesondere wissenschaftliche Bibliotheken mit großen historischen Beständen arbeiten seit Jahren an der Digitalisierung zwecks Erhaltung und verbessertem Zugang zu ihren Beständen. Doch durch den seit Jahren anhaltenden digitalen Wandel der Gesellschaft entstehen täglich unzählige persönliche Objekte, die nicht mehr physisch sondern in digitaler Form vorliegen. Fotos, Briefe, Rechnungen und Sprachnachrichten finden sich verstreut auf Computern, Handys, Tablets oder Servern. Während vor wenigen Jahren diese Erinnerungsstücke im Schuhkarton oder Fotoalbum Platz fanden und so an die nächste Generation weitergegeben wurden, ist dies heute nicht mehr möglich. Neben der schier endlosen Menge an Objekten, bei der es schwer fällt den Überblick zu behalten, sind die dezentrale Verteilung auf mehreren Geräten und die nicht garantierte Kompatibilität der Dateiformate Hindernisse, die der Erinnerungskultur im Wege stehen. Während sich das Bewusstsein für die Datensicherung des heimischen Computers und Smartphones auf externen Speichermedien immer weiter

durchsetzt, erfüllen die dafür eingesetzten Methoden nicht die Standards der Langzeitarchivierung. Die Erstellung von „Backups“ geht meistens nicht über eine redundante Speicherung hinaus. Die weitverbreitete Annahme, dass bereits das Digitalisieren und mehrfache Speichern einer Archivierung entspricht, wird nicht nur durch kommerzielle Anbieter suggeriert¹, sondern auch durch Institutionen wie den WDR unterstützt.² Das Spiegeln und Mehrfachspeichern der Daten auf einem anderen Speichermedium sorgt nur für eine Sicherung der Dateien, nicht aber deren Inhalte. Wichtige Elemente wie die passende Auswahl des Datenformats, welche die Langzeitverfügbarkeit und spätere Nutzung garantieren, werden bisher nur im professionellen Umfeld angewandt. Es droht weiterhin der Verlust durch inkompatible Datei- und Speicherformate. Auch Information über die Daten gehen verloren. Während in der Vergangenheit auf der Rückseite eines Papierfotos häufig Datum, Ort und abgebildete Personen vermerkt wurden, macht sich kaum jemand die Mühe, seine privaten Urlaubsbilder ausreichend mit Metadaten anzureichern. Ohne ein systematisches Vorgehen bei der Sicherung der persönlichen digitalen Objekte droht der Totalverlust der digitalen Erinnerungsstücke. Bibliotheken können mit geringem Aufwand diesem Szenario entgegenwirken und so stärker als Ort des digitalen Wandels ins Bewusstsein der Nutzer gelangen.

Forschungsstand

Abseits der digitalen Langzeitarchivierung im professionellen Umfeld existieren in Deutschland bisher nur wenige Arbeiten die sich mit der privaten Langzeitarchivierung befassen. In den USA und Großbritannien liegt der Forschungsvorsprung in diesem Bereich hingegen bereits bei etwa 15 Jahren. Auf wissenschaftlicher Ebene beschäftigte sich beispielsweise Catherine C. Marshall mit dem Problem. Sie formulierte in ihrer vielbeachteten Arbeit „Rethinking Personal Di-

¹ Amazon: Datenarchivierung. URL: <https://aws.amazon.com/de/archive/> [14. Juni.2017].

² WDR: WDR Digit. URL: <http://digit.wdr.de> [14. Juni.2017].

gital Archiving“³ vier zentrale Fragen, die der Nutzer sich im Rahmen der erfolgreichen Archivierung eigener Daten stellen muss. Die wesentlichen Entscheidungen umfassen die Selektion (Was soll bewahrt werden?), die Wahl des Speicherorts und -mediums (Wo und worauf soll gespeichert werden?), die Formatwahl (Welches Datenformat soll benutzt werden?) und die Wiederauffindbarkeit (Welche Metadaten und Ordnerstruktur soll verwendet werden?). Ein weiterer bedeutender Akteur in diesem Bereich ist die Library of Congress, welche nach unterschiedlichen Medienarten sortierte Leitfäden für die Langzeitarchivierung persönlicher digitaler Daten veröffentlichte.⁴ Diese in sehr einfacher Sprache verfassten Anweisungen richten sich an den Endnutzer und verschaffen einen Überblick über die Schritte, die für eine Archivierung notwendig sind.

Rahmenbedingungen

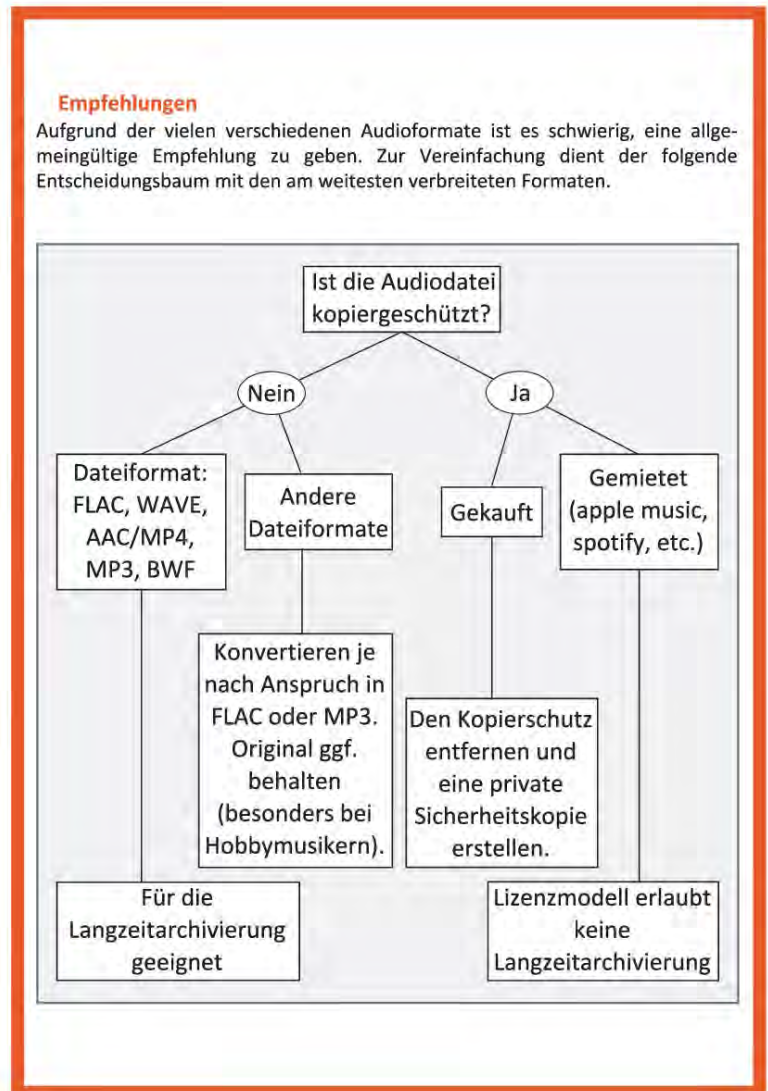
Im Rahmen eines Wahlpflichtfaches des Bachelorstudiengangs Bibliothekswissenschaft an der TH Köln haben sich im Sommersemester 2016 und Wintersemester 2016/2017 unter Leitung von Prof. Achim Oßwald 14 Studierende des Themas angenommen. Das Ziel des Projekts war es, einen Lösungsansatz für das im vorigen beschriebene Problem zu finden, der es Laien ermöglicht, eine Vielzahl unterschiedlicher Kompetenzen im Bereich der Sicherung und Überlieferung (sprich der Archivierung) ihrer privaten Daten anzuwenden. Die Konzeption des Ansatzes stützte sich im Wesentlichen auf die Empfehlungen zur persönlichen Archivierung der Library of Congress und der Arbeit von Marshall. Das darin empfohlene schrittweise Vorgehen wurde für das erstellte Angebot adaptiert und in Form einer praktischen Beratung und schriftlicher Leitfäden umgesetzt.

Umsetzung

Der erste Schritt ist es demnach, dass der Nutzer sich einen Überblick über den eigenen Datenbestand verschafft, um danach fundiert entscheiden zu können, wie er bei der Archivierung dieser Daten vorgehen möchte. Dazu sind verschiedene Entscheidungen notwendig, die für alle unterschiedlichen Medientypen, aus denen private Sammlungen bestehen, im gleichen Maße Gültigkeit besitzen. Diese Entscheidungen beginnen mit dem Bereich der Auswahl – denn dabei spielen gleich zwei entscheidende Faktoren eine

3 Marshall, Catherine C.: Rethinking Personal Digital Archiving Part 2. Implications for Services, Applications, and Institutions, in: D-Lib Magazine 14 (2008) 3/4. URL: <http://www.dlib.org/dlib/march08/marshall/03marshall-pt2.html> [14. Juni.2017].

4 Library of congress: Personal Archiving. Preserving your digital memories. URL: <http://www.digitalpreservation.gov/personalarchiving/> [14. Juni.2017].



Entscheidungsbaum Audio

Rolle. Zum einen ist es wichtig, dass der Nutzer seine gesamten Daten zentral zusammenführt. Das häufig anzutreffende Phänomen der verteilten Speicherung an unterschiedlichen Orten bedeutet eine potentielle Gefahr für den Erhalt und die Langzeitverfügbarkeit der Daten, weshalb es wichtig ist diese vor der Archivierung an einen zentralen Ort zu überführen. Zum anderen sollte aus diesem Konvolut dann eine systematische Auswahl von nach persönlichen Kriterien relevanter Daten erfolgen, die später archiviert werden sollen. Viele Dateien liegen in unterschiedlichen Versionen redundant vor und es ist nur bedingt sinnvoll diese alle zu archivieren. Vielmehr bietet es sich an, von Fotos beispielsweise das Original in der höchsten Auflösung zu archivieren, um bei Bedarf daraus weitere Derivate zu generieren. Nach diesem Auswahlprozess ist es in weiteren Schritten notwendig zu entscheiden, in welchem Format die ausgewählten Dateien zur Archivierung gespeichert werden sollen und welche Speichermedien dabei Verwendung finden. Da



*Pia Piontkowitz
und Lukas
Wähler bei der
Preisübergabe
des TIP Award*

es auch für diese Entscheidungen keine allgemeingültigen Lösungen gibt und es maßgebend ist, dass Nutzer diese Auswahl in Bezug auf ihre persönlichen Bedürfnisse treffen, war ein wesentlicher Teil des Projektes die Erstellung von Handlungsempfehlungen in Form verschiedener Leitfäden (je einen für die unterschiedlichen Datentypen Bild, Text, Audio und Video, sowie einen separaten Leitfaden mit allgemeingültigen Hinweisen), die es den Nutzern ermöglichen sollten anhand von Empfehlungen einen eigenen auf sie zugeschnittenen Archivierungsplan zu entwickeln.

Der dem Projekt zugrundeliegende Kurs begann zunächst mit einer allgemeinen Einführung in die Thematik des Personal Digital Archiving. Wo finden sich die Unterschiede, sofern es sie denn gibt, zwischen dem Personal Digital Archiving und der bibliothekarischen Langzeitarchivierung? Dieser Frage stellten sich die Studierenden zu Beginn des Kurses selbst gegenüber und betrachteten kritisch ihr eigenes Speicherverhalten. Eine erste Brainstorming-Session zeigte auf, dass in der Archivierung persönlicher Dokumente sowohl Wert auf offiziell relevante Dokumente als auch auf private Erinnerungen gelegt wurde. Bei der Frage, wo und wie sicher all diese Daten gespeichert seien herrschte aber betretenes Schweigen unter den Informationsspezialisten.

Der Auftrag stand also: Der gesamte Prozess der eigenen Archivierung von Daten, beginnend bei der Sammlung der zu speichernden Objekte bis hin zur Speicherung auf einem geeigneten Datenträger sollte verständlich und übersichtlich gemacht werden. Leitfäden für die Archivierung sollten erstellt sowie für jeden Nutzer nachvollziehbar ausgestaltet werden. Hierzu teilte sich der Kurs zunächst in Gruppen, die sich an drei der von Catherine Marshall gestellten Fragen anlehnten: Die Datenspeicherung, das Datenformat und die Metadaten. In den einzelnen Gruppen

wurden die Fragen tiefergehend betrachtet und erste Antwortversuche entworfen. Die Gruppe der Datenspeicherung beschäftigte sich mit sämtlichen Ideen rund um Speicherorte und -medien. Welche Medien sind veraltet, welche bewähren sich immer wieder und welche beweisen sich gerade? Was ist bei der Cloud-Speicherung zu beachten und wie lange ist ein Speichermedium haltbar? Das Team Datenformat betrachtete sämtliche vorgeschlagene archivierwürdige Daten und teilte diese nach Medientyp ein. Die daraus entstandenen Unterkategorien von Text, Bild, Audio, und Video wurden dann individuell auf ihre verschiedenen Formate geprüft. Was sind die derzeit gängigen Formate und ist absehbar, inwiefern diese sich auf dem Markt halten werden? Unter welchen Umständen kann, sollte oder muss eine Datei konvertiert werden? Zuletzt forschte die Gruppe rund um Metadaten, inwiefern diese auf dem Heim-PC zu speichern sind und welche Empfehlungen die Studierenden als Bibliothekare aussprechen würden. Gibt es Programme, die die Organisation der privaten Metadaten erleichtern? Sind die vom System gegebenen Möglichkeiten zur Veränderung der Dateieigenschaften genügend oder empfiehlt sich der Einsatz externer Programme und Lösungen?

Die Ergebnisse trugen die Gruppen zum Abschluss der ersten Hälfte des Kurses in einem Dokument zusammen. Bei der Betrachtung dieses Endobjekts fiel dann aber auf, dass die angestrebten Leitfäden weit verfehlt wurden. Das Endergebnis war eine wissenschaftliche Arbeit, die äußerst unübersichtlich war – ein Laie in dieser Thematik konnte sich hierin sicher nicht zurechtfinden. Um die Übersicht für den Nutzer zu gewährleisten teilten sich der Kurs erneut auf, diesmal in Gruppen zu den als relevant ermittelten Medientypen Text, Bild, Audio und Video. Diese Gruppen griffen die bisherigen Ergebnisse auf und betrachteten zusätzlich für die einzelnen Typen relevante Probleme, die bei der Archivierung auftreten können. Die gesammelten Resultate wurden dann jeweils in einem internen ausführlichen Leitfaden und einem Handout im Flyer-Format zusammengetragen. In diesen Handouts wurden Themen wie Besonderheiten des Dateityps, mögliche Softwares, vorherrschende Formate, Metadaten und die Empfehlungen der studentischen Gruppen zur Sprache gebracht. Ein Ampelsystem zu den für den jeweiligen Medientyp zur Verfügung stehenden Formaten sollte den Nutzern einen schnellen Überblick über geeignete Formate geben, welche nach Kriterien wie Verbreitung, Komprimierung und Herstellerabhängigkeit eingestuft wurden. Der ebenfalls in den Handouts zu findende „Entscheidungsbaum“ fungiert als eine Kurzanleitung, die es

Nutzern ermöglicht auch ohne genaue Betrachtung des gesamten Handouts die Archivierung erfolgreich nach ihren individuellen Anforderungen durchführen zu können. Zusätzlich zu den vier Medientypen wurden ein Leitfaden und ein Handout mit allgemeinen Informationen erstellt, die auf Dateien aller Typen zutreffen. Dazu zählen u.a. Speichertypen, Ordnerstrukturen und die möglichen Dateibenennungen. Die Handouts wurden bewusst in einer einfachen und verständlichen Sprache verfasst, sodass auch Personen mit wenigen IT-Kenntnissen den Inhalt des Textes gut erfassen können. Auf Fachvokabular und Fachtermini wurde weitestgehend verzichtet.

Workshop

Die Möglichkeit der praktischen Umsetzung ergab sich in der Zusammenarbeit mit der Stadtbibliothek Köln, die bereit war einen von den Studierenden konzipierten Workshop im Rahmen ihrer Veranstaltungsreihe „Digitale Werkstatt“ anzubieten.⁵ Die Veranstaltung begann mit einem circa fünfzehn minütigen Vortrag in dem es um die grundsätzliche Problematik der Aufbewahrung und Sicherung von digitalen Daten ging. Allen Teilnehmern wurde daran anschließend das Handout mit den allgemeinen Informationen an die Hand gegeben, welches die Informationen des Vortrags in verständlicher Sprache zusammenfasste. Es folgte zuerst eine allgemeine Fragerunde bevor es individuelle Beratungen ausgerichtet nach den einzelnen Medientypen gab. Durch die vorhergehende Aufteilung in Gruppen waren die Studentinnen und Studenten jetzt Spezialisten für Text-, Bild-, Audio- oder Videodateien und somit in der Lage die Teilnehmer je nach ihren individuellen Bedürfnissen und Wünschen zu einzelnen Medientypen zu beraten.

Reflektion

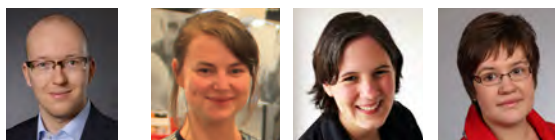
Das Thema Personal Digital Archiving ist ein aktuelles Thema, das besonders den Bereich der Öffentlichen Bibliotheken mit der Vermittlung von Informationskompetenz im Rahmen von Nutzerschulungen betrifft. Auch für Wissenschaftliche Bibliotheken ist es ein relevantes Thema, da gerade im Bereich des wissenschaftlichen Arbeitens das Wiederauffinden von einmal erstellten oder recherchierten Inhalten von zentraler Bedeutung ist. Da die häufigste Ursache für Datenverlust nicht etwa die gefürchteten Hardwarefunktionsstörungen, sondern die Unfähigkeit eine vorhandene Datei wiederzufinden ist, könnten auch wissenschaftliche Bibliotheken entsprechende Schulungen z.B. für

Studenten anbieten, um schon früh in der Karriere zukünftiger Wissenschaftler eine fundierte Basis für eine strukturierte Datenarchivierung zu legen.⁶

Bei vielen Nutzern herrscht zudem eine große Unsicherheit im Umgang mit den eigenen Daten. Dies zeichnete sich auch in der hohen Nachfrage nach dem angebotenen Workshop ab. Bibliotheken hätten hier also die Möglichkeit sich durch die Abdeckung eines aktuellen Angebotsfeldes neu zu positionieren. Für uns war die Konzeption und Durchführung des Workshops eine gute Chance sich so auf die praktische Seite des Berufes vorzubereiten und in unterschiedlichen Teams verschiedene Aspekte eines komplexen Projektes zu bearbeiten. Das große Interesse der Nutzer und die Anfrage nach einer Wiederholung des Angebots stellten dabei einen besonderen Erfolg für uns dar und zeigten deutlich, dass tatsächlich Interesse und Bedarf an diesem aktuellen Thema bestehen.

Danksagung

An dieser Stelle möchten wir uns noch einmal besonders bei unserem Dozenten, Herrn Prof. Dr. Achim Oßwald, für seine Unterstützung und seinen Input während des gesamten Projekts bedanken. Dank auch an die Stadtbibliothek Köln, die es uns ermöglicht hat, den von uns konzipierten Workshop in ihrem Haus durchzuführen und so in der Praxis zu testen. Zusätzlich bedanken wir uns bei unseren Kommilitonen aus dem Seminar BWT 24 für ihren Beitrag zur Erstellung und Durchführung des Projektes. |



Lukas Wähler

Bibliothekswissenschaft [B.A.]
lukaswaehler@posteo.de

Pia Piontkowitz

Bibliothekswissenschaft [B.A.]
pia.piontkowitz@web.de

Annika Mühling

Bibliothekswissenschaft [B.A.]
Kunstgeschichte [M.A.]
annika.muehling@gmx.de

Gabriele Pütz

Bibliothekswissenschaft [B.A.]
Leitung der Bibliothek des Historischen Instituts
der RWTH Aachen
Bibliothek@histinst.rwth-aachen.de

⁵ Stadtbibliothek Köln: Workshops. URL: <http://www.stadt-koeln.de/leben-in-koeln/stadtbibliothek/bibliotheken-archiv/workshops> [14. Juni 2017].

⁶ John, Jeremy Leighton/ Rowlands, Ian/ Williams, Pete/ Dean, Katrina: Digital Lives. Personal digital archives for the 21st century – an initial synthesis. Beta version 0.2, 2010, S. 23. URL: <http://britishlibrary.typepad.co.uk/files/digital-lives-synthesis02-1.pdf> [14. Juni 2017]