

Archivierung von digitalisierten Archivalien in einem Rechenzentrum

Von GABRIELE STÜBER

Das Thema der Archivierung digitalisierter Archivalien berührt einen Bereich, der im Kontext einer Sicherung von Unterlagen aus einer neuen medialen Dimension zunehmende Wichtigkeit erlangt. Der kurze Beitrag stellt am Beispiel des Zentralarchivs der Evangelischen Kirche der Pfalz dar, wie diese Aufgabe angegangen und umgesetzt wird. Im Anschluss werden einige grundsätzliche Fragen formuliert, die die Archive gleich welcher Sparte aus Sicht der Verfasserin begleiten und zur Stellungnahme herausfordern.

Kurzporträt des Zentralarchivs der Evangelischen Kirche der Pfalz, Speyer

Das Zentralarchiv der Evangelischen Kirche der Pfalz ist zuständig für alle in der Landeskirche erzeugten Unterlagen inklusive die der Diakonie. Im Archiv arbeiten sieben hauptamtliche Kräfte, davon drei Fachkräfte, vier ehrenamtliche Kräfte sowie wechselnde Aushilfskräfte. Wir beschäftigen Praktikantinnen und Praktikanten aus Schule und Studium sowie im Rahmen von beruflichen Wiedereingliederungsmaßnahmen und wechselnde Projektkräfte.

Derzeit umfasst der Umfang aller aufbewahrten Unterlagen etwa 5 000 laufende Meter in drei Magazinegebäuden. In dem hier zu behandelnden Kontext sind folgende Unterlagen wichtig: 15 000 Bildquellen und 20 000 Baupläne. Aufgrund seiner regionalen Zuständigkeit und seines Bestandsprofils von etwa 800 Beständen ist das landeskirchliche Zentralarchiv mit einem Staatsarchiv vergleichbar, allerdings einem sehr kleinen. Ungeachtet dessen fallen hier alle Aufgaben an, die in jedem Archiv bewältigt werden müssen.

Erzeugung von Datenmengen durch Digitalisierung

In den vergangenen 15 Jahren wurden nach und nach Bildquellen (Fotos und Objekte der Sammlung Volksfrömmigkeit) digitalisiert, dies geschah in Eigenregie. 45 Filme aus der Ostasienmission wurden durch eine Firma digitalisiert, desgleichen die Glasdias der Mission. Alle Digitalisate wurden auf der landeskirchlichen Serverfarm gespeichert, hier kam es bald zu Engpässen, und die EDV-Abteilung wies das Archiv auf die Problematik der wachsenden Datenmengen hin.

Eine weitere Verschärfung der Lage und ein gewaltiger quantitativer Sprung wurden durch das Projekt einer Digitalisierung der Kirchenbücher erzeugt. Die Digitalisierung von ca. 6 000 Kirchenbüchern aus dem 16. bis 20. Jahrhundert mit durchschnittlich 300 Seiten war in der

Serverlandschaft nicht mehr zu realisieren. Diese Digitalisierung erfolgt seit 2011 mit einem professionellen Anbieter auf zwei Wegen.

Aus den bereits vorhandenen Sicherungsfilmern werden für Publikationszwecke und Ausdrücke im Rahmen von Beglaubigungen Dateien mit einer Auflösung von 300 dpi im JPEG-Format mit 80-prozentigem Qualitätserhalt erzeugt, für die Anwendung im Lesesaal und die Einspielung in das geplante Kirchenbuchportal der Evangelischen Kirche in Deutschland werden diese JPEG-Bilddateien mit 30-prozentigem Qualitätserhalt sowie mit einem hinterlegten Wasserzeichen gefertigt.¹

Bei mäßiger Verfilmungsqualität werden die Originalkirchenbücher nochmals mit 24 Bit Farbtiefe digitalisiert. Die Digitalisate der Sicherungsfilme sind – dem Standard der Sicherungsverfilmung entsprechend – in Schwarzweiß gehalten. Bei der Digitalisierung der Kirchenbücher werden keine DFG-Standards angewendet.

Sicherung der Digitalisate des Zentralarchivs im Rechenzentrum

Über die landeskirchliche Verwaltung wurde im März 2011 ein Hosting-Vertrag mit dem Kirchlichen Rechenzentrum in Eggenstein-Leopoldshafen bei Karlsruhe geschlossen. Dieser Vertrag hat zunächst eine Laufzeit von vier Jahren. Das Rechenzentrum stellt einen Fileserver mit 10 TB für Kirchenbücher, Fotos und Filme zur Verfügung. Der Zugang für das Archiv erfolgt über eine FTP-Verbindung mit Benutzungsrechten für derzeit vier Personen, was als ausreichend erachtet wird.

Die Daten werden mittels eines externen Datenträgers (Wechselfestplatte) importiert.

Eine Vollsicherung erfolgt im Rechenzentrum nach jedem größeren Datenzuwachs. Außerdem wird an jedem Werktag eine nächtliche Sicherung vorgenommen. Die folgende Abbildung veranschaulicht die Konfiguration.

¹ Vgl. hierzu *Gabriele Stüber*: Ein neuer Zugang zu Kirchenbüchern: www.kirchenbuchportal.de. In: *Blätter für pfälzische Kirchengeschichte und religiöse Volkskunde* 77 (2010), S. 217–223.

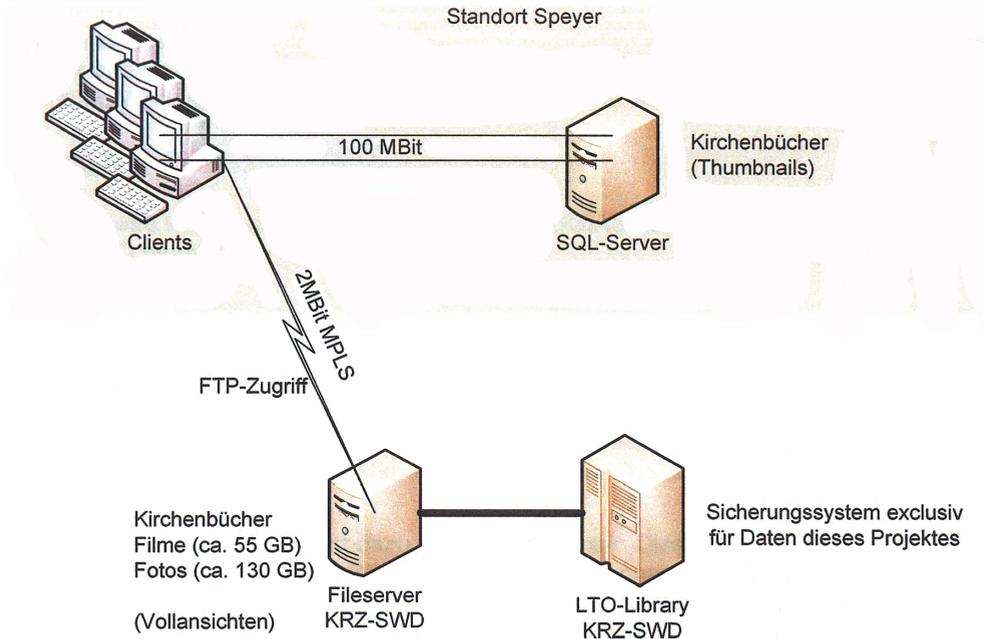


Abb. 1: Konfiguration der Datensicherung des Zentralarchivs der Evangelischen Kirche der Pfalz im Rechenzentrum Eggenstein-Leopoldshafen bei Karlsruhe, erstellt vom Rechenzentrum.

Durch die FTP-Verbindung werden die einzelnen Bilddateien hoch- bzw. heruntergeladen, wobei von einer Häufigkeit von bis zu 10 Bildern pro Tag in der Anfangsphase ausgegangen wird. Diese geringe Zahl ist erstaunlich, doch sei darauf hingewiesen, dass nur im Falle einer Publikation oder zu Beglaubigungszwecken die hochauflösenden Formate benötigt werden. Als Storage-Einheit ist exklusiv für das Archiv eine LTO-Library vorhanden.

Die JPEG-Dateien in geringer Auflösung bleiben als Arbeitsbilder auf den Servern der Landeskirche. Im Tagesgeschäft wird fast ausschließlich mit diesen gearbeitet.

Service und Support bestehen rund um die Uhr an 365 Tagen im Jahr. Die Reaktionszeit bei Störungen beträgt je nach Fehlerklasse zwei bis vier Stunden.

Die Kosten setzen sich zusammen aus

- Einmalkosten für Installation und Konfiguration von 2.135 € brutto

Dazu gehören die Bereitstellung der notwendigen Hardwareinfrastruktur, des erforderlichen Betriebssystemes sowie die Einrichtung der Sicherungsjobs für die Daten und Wartung des Fileservers.

- Laufende Kosten pro Monat: 878 € brutto

Aus diesen Rahmendaten ergeben sich an Gesamtkosten für das erste Betriebsjahr (März – Dezember 2011) 10.900 € brutto, für das zweite Betriebsjahr (Januar – Dezember 2012) 10.540 €. Im Anschluss an das zweite Jahr erfolgt eine Kostenprüfung und nochmalige Verhandlung mit dem Rechenzentrum.

Probleme der Datenzugänglichkeit

Via FTP ist grundsätzlich keine Voranzeige von Images möglich. Man muss also die gewünschte Seite bei einer Anforderung für eine Publikation genau bestimmen können, was aber mittels der Seitenkennung keine Schwierigkeit darstellt.

Im laufenden Betrieb sind bisher keine gravierenden Probleme aufgetreten. Der häufigere Rückgriff auf die Digitalisate in Druckqualität wird zeigen, ob es zu Engpässen kommt.

Bei den Fotos haben wir das FTP-Problem umgangen, indem wir zwar die Fotodaten im Rechenzentrum sichern, aber die Großdateien auf dem Archivserver der Landeskirche vorhalten. Da das Archiv als landeskirchliche Fotoagentur sehr häufig in Anspruch genommen wird, wäre ein anderes Verfahren mit zu hohem Zeitaufwand verbunden. Der weitere Fortschritt der Technik wird zeigen, ob diese Zugangsprobleme bei den großformatigen Dateien nur ein vorübergehendes Phänomen sind.

Ausblick und Fragen

Es wurde bereits darauf hingewiesen, dass der Digitalisierung im Zentralarchiv keine DFG-Normen zugrunde liegen. Die Umsetzung der DFG-Anforderungen erzeugt ein großes Datenvolumen, dessen Unterhaltskosten so manches Archivbudget sprengen, jedenfalls ist das im Zentralarchiv der Fall.

Ich betone diese Praxis noch einmal unter Hinweis auf die Tatsache, dass wir eines der vielen kleineren Archive repräsentieren. Aus der Sicht von Staatsarchiven mag sich die Erfüllung der DFG-Norm als unproblematisch erweisen, zumal die Drittmittelförderung der DFG an die Voraussetzung der Normeinhaltung gebunden ist. Die Kirchenarchive zahlen die Digitalisierung in der Regel aus eigenen Mitteln und befinden daher auch über die zugrunde liegende Norm.

Schreibt man diese Entwicklung fort und bezieht alle Archivsparten in die Betrachtung ein, so erhalten wir unter Umständen zwei (und mehr) Klassen von Archiven in der digitalen Welt. Das hängt vor allem mit der Verschärfung der finanziellen Situation im Kulturbereich zusammen, die sich bei den Archivträgern schon jetzt sehr unterschiedlich darstellt.

Aus der vorfindlichen Konstellation erwachsen folgende Leitfragen:

1. Welche Sicherungsanforderungen (ich vermeide bewusst den Begriff Archivierungsanforderungen) sind an Digitalisate zu stellen, die aus papierbasiertem Archivgut erzeugt wurden? Es ist sicher eine andere Messlatte anzulegen als bei digital born documents, aber welche? Sicherungsanforderungen in diesem Bereich sind zweifellos erforderlich, da die Digitalisierung Kosten erzeugt.

2. Aus der Sicht kleinerer Archive stellt sich die fast schon existentielle Frage, wie sie mit der digitalen Herausforderung umgehen und angesichts ihrer sinkenden Personal- und Sachmitteleinsatzung agieren bzw. reagieren sollen. Aber vielleicht ist das nicht nur eine Frage an die kleineren Archive... Diese digitale Herausforderung umfasst drei Bereiche:

- die Archivierung von Unterlagen aus digitalen Systemen (das Thema dieses Arbeitskreises seit 1997 und eine neue Standardaufgabe von Archiven);
- die Erzeugung digitaler Unterlagen aus herkömmlichem Archivgut und deren Präsentation im Netz, weil der Druck der Öffentlichkeit erheblich ist und ein schneller und niedrigschwelliger Zugang zu Archivgut eine Standardforderung geworden ist. Ein Archiv, das nicht im Netz präsent ist, läuft Gefahr, mit seinen Beständen nicht mehr wahrgenommen zu werden;
- die Sicherung der digitalisierten Archivalien.

Die Kosten für die Erzeugung, Archivierung und Sicherung von digitalen Unterlagen sind eine neue Kategorie für die Budgets aller Archive. Angesichts schrumpfender öffentlicher Haushalte und bleibender Kostenstellen im Bereich der traditionellen Kernaufgaben der Archive sind alle Kostenarten bereits mittelfristig nicht mehr zu schultern. Es könnte zu einer Verschiebung der Finanzmittel zu Lasten der traditionellen Kernaufgaben kommen.

Im Rahmen einer prinzipiell offenen Kosten-Leistungs-Rechnung drängt sich eine Frage auf, die nicht nur Rechnungshöfe stellen könnten: In welchem Rahmen können Archive ihre Leistungen zukünftig noch gebührenfrei anbieten? Müssen sie nicht Segmente ihres Services mit Gebühren belegen, um zumindest eine Teilrefinanzierung zu erzielen? Die Investitionen der Deutschen Forschungsgemeinschaft im Bemühen um ein *public brain* etwa im Kontext der Deutschen Digitalen Bibliothek lassen digitale Datenmengen zurück, die auf Dauer gesichert sein wollen – wer trägt die Folgekosten?

Angesichts dieser Herausforderungen bleibt aber auch die Hoffnung, dass mit einer leistungsfähigeren Technik unter Umständen auch die Kosten dafür sinken werden. Überdies könnte sich die Bereitschaft der Archive aller Sparten, Aufgaben gemeinsam in einem Verbund von Einrichtungen zu bewältigen, als eine zukunftsfähige Vernetzung erweisen.²

² Die Ausgabe der Zeitschrift *Archivar*, 65. Jg. Heft 1/Februar 2012 belegt diverse Verbundlösungen.