

Digitale Bestandserhaltung als archivische Herausforderung

Der Erhalt von Archivgut ist sowohl gesellschaftlicher Auftrag als auch archivgesetzliche Pflicht. Insbesondere der Erhalt digitaler Informationen stellt das Landesarchiv Sachsen-Anhalt dabei vor neue Herausforderungen.

Seit jeher zählen die Sicherung und Erhaltung des verwahrten Kulturguts zu den Kernaufgaben von Archiven. Dabei ist es irrelevant, ob es sich bei den Unterlagen um analoges Schriftgut handelt oder um digitale Daten. Elektronische Unterlagen stellen die Archive jedoch vor besondere Herausforderungen.

Gefährdungen

Digitale Archivalien sind wesentlich fragiler als ihre analogen Pendanten. Insbesondere der schnelle technologische Wandel lässt Systeme, Dateiformate und Hardware rasch veralten. Dateien benötigen zwangsläufig ein Speichermedium, wie einen Netzwerkspeicher oder eine Festplatte als Datenträger. Im Laufe der Zeit kommt es schrittweise zu Fehlern beim Auslesen, bis schließlich der Totalausfall eintritt.

In der Folge lassen sich eine Vielzahl von Dateien nach wenigen Jahren nicht mehr öffnen, da es beispielsweise keine Software mehr gibt, die das Dateiformat korrekt interpretiert oder das Dateiformat gänzlich unbekannt ist.

Migration und Emulation

Die Grundlage sämtlicher Bestandserhaltungsmaßnahmen bildet die Erhaltung des Bitstroms (bitstream preservation). Jedes OAIS-konforme Archiv bewahrt seine elektronischen Unterlagen innerhalb einer mehrfach redundanten Speicherinfrastruktur, in der unterschiedliche Speichertechnologien, wie Netzwerk- oder Bandspeicher, verwendet werden. Durch laufende Integritätschecks stellen Archivierungssysteme wie DIMAG sicher, dass ungewollte Veränderungen zeitnah erkannt und behoben werden. Die nun fehlerhafte Repräsentation des digitalen Archivs wird durch die integrale Version eines anderen Speichers ersetzt.

Der Erhalt des Bitstroms allein ist allerdings noch nicht hinreichend für spätere Nutzungsperspektiven, da die Daten noch in einer Softwareumgebung eines (virtuellen) Lesesaals les- und interpretierbar gemacht werden müssen. Derzeitige PC-Arbeitsplätze sind beispielsweise in der Lage, durch installierte Standard-

Reader und Browser PDF-Dateien darzustellen. Von zukünftiger Software kann jedoch nicht erwartet werden, dass diese automatisch erkennen, dass es sich hier um PDF handelt, geschweige denn, dass sie es zur Anzeige bringen kann. Das digitale Archiv muss also neben dem rein technischen Erhalt auch für die korrekte Interpretierbarkeit der Dateien Sorge tragen. Hierfür verfolgen die Archive derzeit im Wesentlichen zwei Strategien: die Migration und die Emulation.

Die Emulationsstrategie basiert auf dem Nachahmen derjenigen Systemumgebung, aus der die elektronischen Unterlagen stammen, umgesetzt in aktueller Software. Für Nutzer*innen bietet sich so neben der Darstellung der Inhalte die Möglichkeit, ein annähernd gleiches Nutzungserlebnis durch die Einbettung in den Kontext zu erfahren. Der Einsatz von Emulatoren lohnt sich beim Verwenden von komplexen und interaktiven Anwendungen, wie Webseiten oder Simulationen.

Ermittlung signifikanter Eigenschaften von Bildformaten

Eigenschaften	Histogramm	EXIF	ExifTool
File Type		TIFF	
File Type Extension		tif	
MIME Type		image/tiff	
Exif Byte Order		Little-endian (Intel, II)	
EXIF			
Subfile Type		Single page of multi-page image	
Image Width		1728	
Image Height		2302	
Bits Per Sample		1	
Compression		T6/Group 4 Fax	
Photometric Interpretation		WhiteZero	
Fill Order		Normal	
Strip Offsets		(Binary data 40 bytes, use -b option)	
Orientation		Horizontal (normal)	
Samples Per Pixel		1	
Rows Per Strip		303	
Strip Byte Counts		(Binary data 39 bytes, use -b option)	
X Resolution		200	
Y Resolution		200	
Planar Configuration		Chunky	
Resolution Unit		inches	
Page Number		0 12	
Software		ImageGear Version: 6.05.020	
Modify Date		2005: 3:04 16:31:52	
Artist		1005 07 AcySoft Inc. All rights reserved	

Bei der Migrationsstrategie werden proprietäre Dateiformate durch regelmäßige Konvertierungen in aktuelle und langzeitstabile Formate umgewandelt. Die Ursprungsdaten sowie alle nachfolgenden Konvertierungen werden dabei versioniert und ebenfalls bewahrt. Als Ergebnis folgt hieraus eine Kette unterschiedlicher Versionen bzw. Repräsentationen zu einer Verzeichnungseinheit, die im Vergleich zueinander sehr heterogen gestaltet sein können. Im Vorhinein sind daher die vorliegenden Dateiformate zu Erhaltungsgruppen zusammenzuführen und für diese sämtliche erhaltenswürdigen bzw. signifikanten Informationen und Eigenschaften zu bestimmen. Unabhängig davon, wie spätere Objektversionen aussehen mögen, hat das digitale Archiv dafür zu sorgen, dass diese signifikanten Eigenschaften stets beibehalten werden, um eine geordnete Überlieferungsbildung der elektronischen Unterlagen zu gewährleisten. Beispielhaft können für Bildformate die Längen- und Breitenangaben, Farbräume und Bittiefen als signifikant deklariert werden.

Momentan gilt die Migration als die Vorzugsvariante für die digitale Bestandserhaltung in Archiven – Emulationen hingegen geraten zunehmend in das Blickfeld bewahrender Einrichtungen. Durch eine vermehrte wissenschaftliche Auseinandersetzung mit dieser Thematik und hieraus entstandene Firmengründungen zeigen sich Möglichkeiten für einen produktiven Einsatz. Perspektivisch ist dabei aus fachlicher Sichtweise der Archive keine Entscheidung für oder gegen eine bestimmte Strategie absehbar, vielmehr wird es wohl auf einen parallelen Einsatz beider Methoden hinauslaufen. Ausgehend von der Fragestellung der Nutzer*innen, dem angedachten Nutzungsszenario sowie den vorliegenden elektronischen Unterlagen muss dies individuell entschieden werden.

DIMAG und DAN

Angesichts der massiven Herausforderungen der Bestandserhaltung elektronischer Unterlagen schloss sich das Landesarchiv Sachsen-Anhalt im Jahr 2016 dem Kooperationsverbund "Digitale Archivierung Nord" (DAN) auch mit der Perspektive an, gemeinsam mit anderen Landesarchiven dieser prioritären fachlichen Aufgabe zu begegnen und sich an den Entwicklungsarbeiten des DIMAG-Verbundes zu beteiligen. Verschiedene Arbeitsgruppen des Verbundes beschäftigen sich direkt oder partiell mit fachlichen Ausarbeitungen zur digitalen Bestandserhaltung. Mit der Konzeption und Entwicklung eines Bestandserhaltungsmoduls (BEM) durch das Staatsarchiv Hamburg, unter Beteiligung der Landesarchive Mecklenburg-Vorpommern und Sachsen-Anhalt, gelang ein wesentlicher Meilenstein hin zu einer produktiven Software-Lösung, die in einer ersten Version seit 2019 vorliegt. Mithilfe des BEM ist es DIMAG-Anwendern möglich, veraltete Dateiformate (Bild- und Office-Formate) zu identifizieren, massenhaft in langzeitstabile Formate, wie PDF-A, zu konvertieren und automatisiert als neue Repräsentation abzulegen.

Ausblick

Mit dem Bedarf für eine vollautomatisierte Migrationslösung im DIMAG-Verbund und der Nachnutzungsmöglichkeit für DIMAG-Partner beginnen in 2020 die Planungen für eine BEM-Version 2.0. Insbesondere die Auswahl und Validierung der signifikanten Eigenschaften verschiedener Archivaliengattungen spielen hierbei eine wichtige Rolle.

Mit einer neugegründeten Dach-AG Bestandserhaltung im DIMAG-Verbund sollen die an unterschiedlichen Stellen anfallenden Ergebnisse thematisch gebündelt werden. Die federführende Verantwortung für die konzeptionellen Vorarbeiten übernimmt dabei das Landesarchiv Sachsen-Anhalt, stellvertretend für DAN.

Björn Steffenhagen

Konfiguration von Migrationsroutinen im Bestandserhaltungsmodul

