

## Workshop 1:

### **Technische Rahmenbedingungen und Präsentation der Digitalisate**

Leitung: Dr. Thomas Fricke, Udo Herkert (Landesarchiv Baden-Württemberg)

Der Workshop begann mit einem Themenüberblick und dem Hinweis auf drei einschlägige Online-Publikationen, nämlich die *Empfehlung der Bundeskonferenz der Kommunalarchive beim Deutschen Städtetag zur Digitalisierung von archivischem Sammlungsgut* von 2005 ([www.bundeskonferenz-kommunalarchive.de/empfehlungen/Empfehlung\\_Digitalisierung.pdf](http://www.bundeskonferenz-kommunalarchive.de/empfehlungen/Empfehlung_Digitalisierung.pdf)), die von Anna Weymann und anderen erarbeitete *Einführung in die Digitalisierung von gedrucktem Kulturgut. Ein Handbuch für Einsteiger* ([www.iai.spk-berlin.de/fileadmin/dokumentenbibliothek/handbuch/Handbuch\\_Digitalisierung\\_IAI\\_IPK\\_Online\\_druck.pdf](http://www.iai.spk-berlin.de/fileadmin/dokumentenbibliothek/handbuch/Handbuch_Digitalisierung_IAI_IPK_Online_druck.pdf)) und die im Februar 2013 aktualisierten *DFG-Praxisregeln „Digitalisierung“* ([www.dfg.de/formulare/12\\_151/12\\_151\\_de.pdf](http://www.dfg.de/formulare/12_151/12_151_de.pdf)).

Anschließend wurden zunächst die Rahmenbedingungen der Digitalisierung von Archivgut angesprochen, wobei die technischen Rahmenbedingungen im Mittelpunkt standen. Ein Archiv, das seine Archivalien zumindest partiell selbst digitalisieren möchte, benötigt nicht nur einen verdunkelbaren Arbeitsraum mit einem geeigneten Scanner. Notwendig ist auch ein leistungsfähiger PC oder eine Workstation mit einem kalibrierbarem Monitor und eine Software zur Bildbearbeitung. Da beim Digitalisieren von Archivalien schon in kurzer Zeit sehr große Datenmengen entstehen können, sollte auch ein lokaler Massenspeicher (Fileserver, NAS) verfügbar sein, sofern die erzeugten Bilddateien nicht außer Haus, beispielsweise in einem Rechenzentrum, gespeichert werden können. Sehr hilfreich sind außerdem ein lokales Netz mit hoher Transferrate ( $\geq 100$  MBit/s) und eine leistungsstarke Internetanbindung. Gerade für kleinere Archive, die sich nur einen Scanner leisten können, ist die Auswahl des richtigen Scannertyps besonders wichtig. Daher wurden die spezifischen Vor- und Nachteile von Flachbett-, Einzug- und Aufsichtscannern sowie von Scanrobotern, Digitalkameras und Analogkameras mit einem digitalen Rückteil (Scanback) ausführlicher behandelt. Eingegangen wurde auch auf das Einscannen von Dias, Fotonegativen und Mikrofilmen sowie auf die technischen Parameter für die Digitalisierung von Archiv- und Bibliotheksgut. Hier scheinen sich die von der Deutschen Forschungsgemeinschaft vorgegebenen Werte für Auflösung, Farbtiefe und Dateiformate (siehe DFG-Praxisregeln „Digitalisierung“ 02/2013, Abschnitt 2.2.1) als Quasi-Standard zu etablieren. Zum Abschluss dieses ersten Teils des Workshops wurde ein erstes Fazit gezogen: Auch bei den Scannern gibt es

(noch) keine „eierlegende Wollmichsau“, d.h. einen Alleskönner, mit dem alle Arten von Archivgut konservatorisch unbedenklich, einfach, schnell und in guter Qualität digitalisiert werden können. In der Regel sind Aufsichtsscanner – oft auch als Buchscanner bezeichnet – für Archive der beste Kompromiss, doch selbst mit einem besonders vielseitigen (und entsprechend teuren) Aufsichtsscanner lassen sich meist nicht alle Digitalisierungsaufgaben, die in einem Archiv anfallen können, befriedigend durchführen. Daher kann es für kleine wie für große Archive, die eigene Scanner besitzen, nach wie vor sinnvoll sein, bestimmte (Teil-)Bestände von einem Dienstleister digitalisieren zu lassen. Dazu müssen die betreffenden Archivalien nicht unbedingt außer Haus gegeben werden, weil manche Scandienstleister durchaus bereit sind, in Räumlichkeiten des auftraggebenden Archivs zu arbeiten. Da der Wunsch eines Workshop-Teilnehmers nach einer Zertifizierungsstelle für Scandienstleister, die Archiven ihre Dienste anbieten, in absehbarer Zeit wohl nicht realisiert werden wird, bleibt Archivarinnen und Archivaren bei der Suche nach geeigneten Anbietern gegenwärtig nur der altbekannte Weg, im Kollegenkreis entsprechende Erkundigungen einzuziehen.

Der zweite Teil des Workshops befasste sich mit den Arbeiten, die nach dem Scannen der Archivalien erledigt werden müssen, um die erzeugten digitalen Kopien komfortabel nutzen zu können, sei es lokal im archiveigenen Intranet oder weltweit im Internet.

Zunächst stand die weitere Aufbereitung der Digitalbilder im Mittelpunkt der Erörterungen. Bei Digitalisierungsprojekten sollte von jedem analogen Objekt ein für den Druck und die längerfristige Archivierung geeignetes, höherauflösendes Masterdigitalisat erstellt werden und hiervon wiederum eine Internetfassung, die für die Bildschirmdarstellung optimiert ist. Die Masterfassung sollte eine Auflösung 300 dpi bezogen auf die Größe des Originals aufweisen, mit einem ICC-Farbprofil versehen sein und das dargestellte Objekt in allen für die Nutzung wesentlichen Gehalten wiedergeben. Der Master wird in der Regel im TIFF-Format gespeichert. Die Internetfassung unterscheidet sich vom Master durch eine geringere, der Bildschirmdarstellung angepasste Auflösung und die Verwendung eines internettauglichen Dateiformats (JPEG oder PNG mit Komprimierung). Die Massenaufbereitung von Digitalbildern erfolgt am besten mithilfe erprobter Bildbearbeitungsprogramme und -tools wie z.B. Adobe Photoshop, Xnview oder dem Kommandozeilenprogramm Image Magick, das bei sehr großen Mengen am geeignetsten ist.

Ein weiteres Thema innerhalb des Workshops bildete die Verknüpfung der Digitalbilder mit den archivischen Erschließungsdaten. Bei großen Mengen an Bildern erfolgt die Zuordnung seriell bzw. automatisiert durch ein entsprechendes

Programm. Im Landesarchiv Baden-Württemberg wird hierfür das sogenannte bildCMS, eine Eigenentwicklung, verwendet. Damit die einzelnen Digitalbilder seriell archivischen Titelaufnahmen zugeordnet werden können, müssen die Dateinamen so vergeben werden, dass das Programm die zugehörigen Datenbankeinträge automatisch ermitteln kann. Die Zuordnung kann entweder mithilfe der Datenbank-ID oder der Bestellsignatur erfolgen. Nach der Zuordnung besteht der Dateiname eines jeden Digitalisats aus Angaben über den Standort des Digitalisats bzw. die Zugehörigkeit zu einer bestimmten Archivalieneinheit (Titelaufnahme) und einer eindeutigen Kennung. Diese eindeutige Kennung wird auch AID (archival identifier) genannt. Sie wird benötigt, um die persistente Adressierbarkeit des Digitalisats im Internet zu gewährleisten. Jedes Digitalbild wird als Unikat behandelt. Wenn ein anderes Bild an seine Stelle tritt, bekommt es eine neue AID. Die Bilder werden in einer Ordnerstruktur abgelegt, die nach einem einheitlichen Prinzip aufgebaut ist. Die Bezeichnungen der einzelnen Ordner entsprechen der Identifikationsnummer der Findmittel-Datenbank.

Jedes Digitalisat kann zusätzlich zum technischen Identifikator eine textuelle Bezeichnung erhalten, die dann in der Internetpräsentation angezeigt wird. Da die personellen Kapazitäten im archivischen Bereich jedoch nicht dazu ausreichen, jedes Digitalisat entsprechend zu benennen, also z.B. mit einer Seitenzahl zu versehen, die derjenigen im Original entspricht, steht in der Internetanwendung des Landesarchivs Baden-Württemberg eine besondere Funktionalität zur Verfügung, die dem Benutzer die Möglichkeit gibt, sich schnell in einem umfangreichen historischen Dokument zu orientieren. Wenn ein Dokument mehr als drei Seiten umfasst, erscheint eine Auswahlseite mit konfigurierbaren Vorschaubildern. Der Benutzer kann nicht nur die Größe der Vorschaubilder anpassen, sondern auch den Bildbereich festlegen, der jeweils angezeigt wird. Befindet sich die Seitenzahl im Original z.B. am rechten oberen Rand, so kann er die Vorschaubilder so konfigurieren, dass der entsprechende Bereich gut lesbar dargestellt wird. Nun kann er das Dokument bequem durchblättern und schnell die von ihm gesuchte Seite finden.

Im letzten Teil des Workshops wurden die Funktionalitäten des bildCMS im Einzelnen vorgestellt und erläutert. Das Programm bietet, wie erwähnt, die Möglichkeit, Digitalbilder archivischen Titelaufnahmen zuzuordnen. Dies kann aber nicht nur in automatisierter Form aufgrund von Informationen, die in den Dateinamen enthalten sind, geschehen sondern auch halbautomatisch, indem bestimmte Bilder manuell markiert und dann verknüpft werden. Selbstverständlich existiert auch ein Modul, das das Hochladen der Bilder ins Internet ermöglicht. Im Grunde übernimmt das bildCMS die komplette Verwaltung der Digitalisate. Bilder können in Serien oder einzeln verschoben und gelöscht werden. In einem sogenannten Bildexplorer kann

nach Digitalisaten gesucht werden. Desweiteren besteht die Möglichkeit ausgewählte Bilder herunterzuladen, um sie zum Beispiel einem Benutzer auszuhändigen. Und schließlich enthält das bildCMS auch Funktionalitäten, die helfen die Arbeitsabläufe bei der Digitalisierung zu erleichtern, so kann z.B. ein Findbuchauszug in Form einer CSV-Datei erstellt werden, der die Identifikatoren und Signaturen aller Titelaufnahmen eines Bestandes enthält. Dieser Auszug erleichtert die korrekte Benennung der beim Scannen erzeugten Bilddateien, welche wiederum die Voraussetzung für die fehlerlose automatisierte Verknüpfung der digitalen Kopien mit den archivischen Erschließungsdaten bildet.

Der zweite Teil des Workshops verdeutlichte allein schon durch seine Dauer (er nahm rund zwei Drittel der Gesamtzeit ein), dass es bei der digitalen Bereitstellung von Archivgut mit dem Scannen allein nicht getan ist. Wenn die erzeugten digitalen Kopien über das Internet benutzerfreundlich zugänglich gemacht werden sollen, sind nach dem Scannen weitere Arbeitsschritte notwendig, wie die Bildbearbeitung und v.a. das Verknüpfen der Bilddateien mit aussagekräftigen Erschließungsdaten, die der Nutzer maschinell durchsuchen kann. Da digitale Kopien von Archivalien nur selten mittels OCR in maschinenlesbare Textdateien umgewandelt werden können, bleiben die erzeugten Kopien im Grunde digitale Fotos, die erst durch Metadaten, also eine Titelaufnahme oder ein Regest, auffindbar und nutzbar werden. Bei der Planung von Digitalisierungsmaßnahmen muss also auch der Aufwand für jene Arbeitsschritte, die sich an das Scannen anschließen, mit einkalkuliert werden. Dabei ist zu beachten, dass die so genannten Nacharbeiten oft doppelt so viel Arbeitszeit beanspruchen wie das eigentliche Scannen.