

Digitale Bilder im Archiv

von Peter Worm

Spiegel online meldete am 6. März, dass sich das weltweite Datenvolumen in elektronischer Form in den nächsten 4 Jahren versechsfachen wird: Von 161 Milliarden Gigabyte auf 988 Milliarden Gigabyte. Den Digitalbildern kommt dabei eine nicht unerhebliche Rolle zu: Die Zahl der Digitalkameras wird sich von weltweit knapp 300 Millionen bis 2010 auf über 500 Millionen steigern¹.

Es wundert deshalb kaum, dass fast jedes Archiv inzwischen digitale Bildbestände hat – sei es, dass aus einem Fotografennachlass digitales Originalmaterial ins Archiv gekommen ist, sei es, dass durch das Abfotografieren oder -scannen von Originalen digitale Reproduktionen erzeugt worden sind².

Wie analoge Fotos stellen auch ihre digitalen Verwandten gewisse Anforderungen ans Archiv – in bestandserhalterischer wie in archivfachlicher Sicht. Um

diese Ansprüche verständlich zu machen, wird im Folgenden zunächst erläutert, wie Digitalbilder »funktionieren«. Anschließend wird der Weg des Bilds von der Entstehung über die Lagerung und Sicherung bis hin zur Vorlage bzw. Nutzung am Bildschirm verfolgt. Viele

.....
1 Spiegel ONLINE, 6. März 2007, <http://www.spiegel.de/netzwelt/web/0,1518,469972,00.html>.

2 Die Forderung nach der Digitalisierung und digitalen Bereitstellung von Kulturgut wird sowohl in zunehmendem Maß von Benutzern aber auch von der Politik an die Archive herangetragen. Stellvertretend sei hier auf die im Rahmen des i2010-Programms der Europäischen Union entstandene Empfehlung der Kommission vom 24. August 2006 zur Digitalisierung und Online-Zugänglichkeit kulturellen Materials und dessen digitaler Bewahrung (2006/585/EG). In: Amtsblatt der Europäischen Union, L 236/28 DE, 31.8.2006. Online abrufbar unter http://ec.europa.eu/information_society/activities/digital_libraries/doc/recommendation/recommendation/de.pdf.



Abb. 1: Bitonales Bild

Archiv LWL, Westfälisches Amt für Landes- und Baupflege, Karten. A.07.02.710K –Nr. 115:
Blick ins Münsterland vom Teutoburger Wald, Kreis Tecklenburg

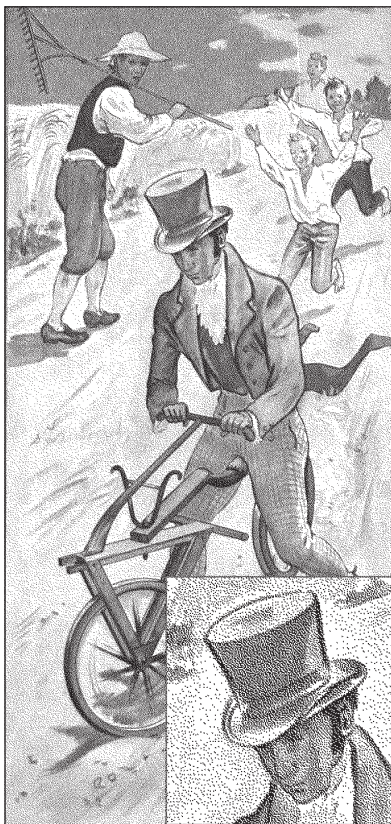


Abb. 2a: Darstellung von Graustufen im Bitonalen Bild mit Detailvergrößerung

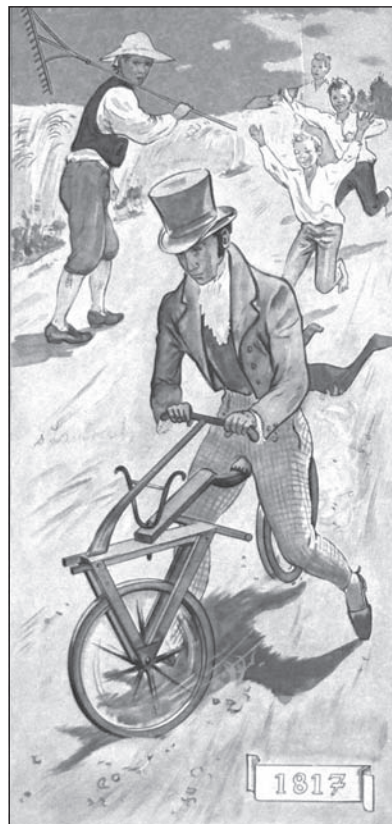


Abb. 2b: Der gleiche Bildausschnitt als Bild mit 256 Graustufen

Archiv LWL, Schulwandbilder A.08.05 – Nr. 534: Die Entwicklung des Fahrrades, 1817,1880, 1949, nach einem Original von E. Grazioli, hier zu 1817






				
300 dpi 300×300 px 90.000 px 116.124 Byte	150 dpi 150×150 px 22.500 px 45.712 Byte	75 dpi 75×75 px 5.625 px 25.400 Byte	30 dpi 30×30 px 900 px 19.396 Byte	15 dpi 15×15 px 225 px 18.480 Byte

Abb. 3: Ein Ausschnitt als Bild mit 256 Graustufen in abnehmender Auflösung

Archiv LWL, Schulwandbilder A.08.05 – Nr. 534 wie Abb. 2

Inhalte können aufgrund des begrenzten Rahmens eines Aufsatzes nur aufgerissen werden.

1.

Eine Bilddatei besteht wie jede andere Datei aus einer Abfolge von Nullen und Einsen, die vom Computer so interpretiert werden können, dass sie zu einer grafischen Ausgabe führen. Für ganz einfache Bilder ordnet man den Bildpunkten, aus denen sie sich zusammensetzen, 0 für weiß und die 1 für schwarz zu und erhält beispielsweise folgende Ergebnisse:

Die von Natur aus im Wesentlichen schwarz-weiße Tuschezeichnung ist gut zu erkennen (Abb. 1), dem Ausschnitt aus eines Schulwandbilds zur Entwicklung des Fahrrads fehlt es aber an den weichen Übergängen und Details (Abb. 2a; Dateigröße TIFF unkompr. 0,165 MB). Grautöne können in diesem einfachsten Bildformat nur durch die Anzahl und Dichte der schwarzen Punkte dargestellt werden; eine Farbdarstellung ist gar nicht möglich. Die Vergrößerung macht den Informationsgehalt des Bildes deutlich. Am ehesten kann die hier verwandte Technik mit der Lithografie verglichen werden.

Möchte man eine lebensnähere Darstellung haben, die fließendere Konturen und Schattierungen zulässt, werden die entstehenden Informationsmengen gleich deutlich größer. Üblich ist es, bei einem Graustufenbild pro Bildpunkt (sog. Pixel) 256 Tonwertstufen darzustellen, dafür werden 8 Bit benötigt. Es entsteht im Vergleich zum bitonalen Bild die rund 8-fache Datenmenge (Abb. 2b; Dateigröße TIFF unkompr. 1,097 MB).

Der Tonwertumfang eines Vollfarbbildes besteht aus 3 Farbkanälen mit je 8 Bit (also insgesamt 24 Bit), das zu gut 16 Millionen möglichen Farbinformationen führt. Hier kommt es noch einmal zu einer Verdreifachung der Datenmenge im Vergleich zum Graustufenbild. Im Vergleich zum bitonalen Bild ist es sogar die rund 20-fache Datenmenge (Dateigröße TIFF unkompr. 3,196 MB).

Der zweite Faktor, der die Dateigröße beeinflusst, ist die sogenannte Auflösung, die Anzahl der Bildpunkte bezogen auf die Bildgröße. Aus dem amerikanischen Maßsystem kommend hat sich auch bei uns die Bezeichnung dpi = dots per inch eingebürgert. Diese Art der Bildgröße wird gern von Buch- und Zeitungsdruckereien vorgegeben, auch für den Drucker auf dem

Schreibtisch im Büro oder zu Hause wird mit dieser Bildauflösung erworben.

Statt der Angabe dieser relativen Auflösung kann auch die absolute Auflösung, also die Anzahl der Pixel insgesamt angegeben werden, um eine Bildgröße zu definieren. Das findet man bei Digitalkameras, die inzwischen mit »10 Megapixeln« – das sind 10 Millionen Pixel – pro Bild daher kommen, aber auch für Bildschirmauflösungen die gleichzeitig das Seitenverhältnis beinhalten. Ein alter Monitor mit 640×480 Pixeln kann also 0,3 Megapixel, ein heute üblicher Monitor mit 1024×768 Pixeln kann 0,8 Megapixel und ein professionelles Gerät kann 1600×1200 Bildpunkte, das knapp 2 Megapixeln entspricht, darstellen.

Unser Auge ist beim Bildschirm beispielsweise nicht in der Lage, eine Auflösung von 72–100 dpi zu erkennen. Die einzelnen Punkte werden in unserem Gehirn also auch bei einer höheren hinterlegten Informationsdichte gleichermaßen zu einem flächigen Gesamteindruck zusammengefügt. Selbst im Papierausdruck unterscheiden sich die ersten zwei Beispielbilder kaum; erst die 75 dpi Auflösung wirkt etwas verwaschener. Bei 30 dpi erahnt man noch das dargestellte Sujet – erst die 15 dpi-Auflösung vermittelt nur noch Kleckse, die unser Gehirn ohne die Vorbilder auf der linken Seite nicht interpretieren könnte.

Als Faustregel kann man sich merken, dass eine Verdopplung der Größe oder der dpi-Zahl eine Ver vierfachung der Dateigröße nach sich zieht. Dass sich das in unseren Beispieldaten nicht ganz so deutlich auswirkt, kommt durch die Fixdaten in der TIFF-Datei, durch die Größe, Farbspektrum und ggf. inhaltliche Informationen, sogenannte Metadaten, festgelegt werden und die als konstante Größe erhalten bleiben. Dieser einführende Teil heißt Kopftext, gebräuchlicher ist der englische Begriff »header«.

2.

Mit diesem Grundwissen im Gepäck kommen wir nun zum Thema der Erstellung von digitalen Bildern. Es gibt leider nicht *die* richtige Auflösung für alle Anwendungsbereiche, sondern sie ist unterschiedlich je nach Ausgabeziel bzw. Verwendungszweck(en). Die folgenden Überlegungen stammen zu großen Teilen aus den Praxisregeln der Deutschen Forschungsgemeinschaft für »kulturelle Überlieferung«, die in diesem Jahr

neu erschienen sind³. Für eine Ausgabe am Computerbildschirm – z. B. als Abbildung im Internet – benötigt man weniger Bildpunkte, sprich eine geringere Auflösung, als für den Druck eines Buchs. Generell gilt, dass ein Bild immer verkleinert oder seine Auflösung reduziert werden kann, dass es umgekehrt jedoch nicht funktioniert. Anders gesagt ist die Reduktion der Bildinformation aus einer hochaufgelösten Aufnahme möglich, die Ergänzung von Bildinformationen aus einer nur wenige Bildpunkte umfassenden Datei ist es nicht. Für die langfristige Nutzung ausgelegte Aufnahmen sollten deshalb i. d. R. in einer angemessenen hohen – sprich in Druckauflösung – abgespeichert werden.

Will man das Original vergrößert abbilden, erhöht sich auch die dafür benötigte Auflösung bei der Aufnahme. Es gilt die zweistufige Formel:

Größe der Ausgabe	=	Vergrößerungs-
Größe des Originals		faktor
Vergrößerungsfaktor		
×	=	Scanauflösung
Bildauflösung		

Neben einer ausreichenden Auflösung sind eine farbtreue und möglichst unverzerrte Aufnahme wichtige Qualitätsfaktoren.

Schließlich muss man sich fragen:

1. Soll man die erstellten Bilder nachträglich digital bearbeiten und erneut abspeichern können und soll das erstellte Digitalisat für möglichst viele Zwecke geeignet sein?
2. Ist ein fixer Zustand erreicht, der keine spätere Nachbesserung erwarten lässt, und nimmt man in Kauf, ggf. für einen geränderten Verwendungszweck erneut zu digitalisieren?

Entscheidet man sich für die erste Variante, kommt nur ein nicht verlustbehaftetes, gut dokumentiertes Langzeitformat für die Speicherung in Frage, das das Bild in einer für den Offset-Druck geeigneten Auflösung vorhält.

Man sollte dann große Datenmengen sicher ablegen können und über die Bereitstellung von Rechnern mit entsprechender Hard- und Softwareausstattung und genügenden Speicherressourcen mit seiner Verwaltung gesprochen haben.

Begnügt man sich mit dem zweiten Lösungsszenario, reichen verlustbehaftete Formate aus, deren Auflösung dem aktuellen Digitalisierungszweck angepasst sein sollten.

Die Variantenbreite soll an den folgenden drei Anwendungsfällen verdeutlicht werden:

1. Beispiel: »Gebrauchsscans« für einen Benutzer

Ein Benutzer bittet um eine Scankopie aus einer Akte. Der verwendete Buchscanner arbeitet am besten mit der größtmöglichen optischen Auflösung. Von der Software her sind viele Geräte in der Lage, Bilddateien mit höheren dpi-Zahlen zu erzeugen. Diese so-

genannte interpolierte »bessere« Auflösung bringt jedoch keine qualitative Verbesserung mehr, da von der eingebauten Hardware bei der Aufnahme nicht mehr Bildpunkte erfasst werden, sondern nur von erfassten Bildpunkten rechnerisch auf die dazwischen liegenden geschlossen wird. Die optische Maximalauflösung wird in der Dokumentation des Scanners angegeben und liegt für die Folio-Doppelseite der Akte beispielsweise bei rund 200 dpi. Davon wird direkt ein Papierausdruck auf DIN A3 erstellt, der dem Benutzer ausgehändigt wird. Für Nachfragen wird die Datei zwei Wochen lang auf der lokalen Festplatte des Rechners, an dem der Scanner angeschlossen ist, gespeichert. Kommen keine Reklamationen wird die Datei gelöscht. Da die Bilddatei nur von temporären Wert ist, wird sie als JPEG mit 90%er Qualität gespeichert. Im Reklamationsfall würde das ein erneutes Ausdrucken und notfalls vorherige kleinere Schärfungs- oder Kontrastverbesserungen erlauben. Der Speicherbedarf wird bei JPEG dadurch stark reduziert, dass Farbpunkte mit ähnlichen Farbwerten zu einer Fläche zusammengefasst werden. Dem Auge fällt dieser Trick zunächst nicht weiter auf. Würde man das Bild nach Veränderungen erneut als JPEG speichern, würden größere Flächen als ähnlich bewertet und zusammengezogen werden. Auf diese Weise entstehen sogenannte Artefakte oder Kästchen, die das Bild letztendlich stark verfremden – nicht unähnlich wie bei einer Aufnahme mit zu geringer Auflösung. Das stört den Archivar in diesem Fall nicht, denn das Bild wird nach den zwei Wochen ohnehin gelöscht.

2. Beispiel: Der Bildinhalt von durch Schimmel angegriffenen Farbnegativen aus einem Fotografen-Nachlass soll durch das Scannen bewahrt werden

Die Negative haben eine Breite von 24 × 36 mm. Die erstellten Digitalisate sollen maximal in DIN A4-Größe in Fotoqualität ausdrückbar sein. So ergibt sich nach der oben genannten ersten Formel ein Vergrößerungsfaktor von rund 10. Setzt man eine ausreichende Auflösung für den Druck von 300 dpi ein, so erhält man eine benötigte Scanauflösung von rund 3000 dpi. Die Durchlichteinheit des eingesetzten Flachbetts scanners würde mit ihren echten 4800 dpi Vergrößerungen bis um den Faktor 16 erlauben, doch würde das die Scandauer und die anfallenden Datenmengen entsprechend vergrößern. Da die so erzeugten Bilder eine Art Ersatzverfilmung der chemisch in Auflösung befindlichen Originale darstellen, kommt nur eine verlustfreie Speicherung in Frage. Zwei Formate bieten sich derzeit an: Das ältere, unkomprimierte TIFF-Format⁴ oder das neuere auf »verlustfrei« eingestellte JPEG2000-.....

³ Deutsche Forschungsgemeinschaft (Hg.), Praxisregeln im Förderprogramm »Kulturelle Überlieferung«. Bonn 2007, bes. S. 8–11. Als Download unter http://www.dfg.de/forschungsfoerderung/formulare/download/12_151.pdf. In weiten Teilen noch gültig ist die Empfehlung der BKK aus dem Jahr 2005: Digitalisierung von archivischem Sammlungsgut. http://www.bundeskonferenz-kommunalarchive.de/empfehlungen/Empfehlung_Digitalisierung.pdf.

⁴ TIFF oder kurz TIF ist die Abkürzung für »Tagged Image File Format«. Die heute gültige Version 6.0 wurde 1992 freigegeben und wird heute vom Marktführer für Grafikanwendungen Adobe betreut. Die Grundstrukturen liegen offen (<http://partners.adobe.com/public/developer/tiff/index.html>), doch ist es jedem Entwickler freigestellt, Ergänzungen zu dem TIFF-Standard zu definieren. TIFF ist also nicht unbedingt gleich TIFF und langfristig sicher lesbar ist ausschließlich eine

Format⁵, das gegenüber TIFF ca. 50 % Speicherplatz spart, aber durch die kompliziertere Kodierung (biorthogonale LeGall-5/3-Wavelettransformation) länger für das Abspeichern benötigt.

Aus diesen Master-Dateien können beispielsweise Plakate für eine Ausstellung zum 100. Geburtstag des Nachlassers und ein Begleitband gestaltet werden. Für die Internetankündigung können 20 Digitalbilder als normale JPEG-Datei mit 300 × 200 Pixel Kantenlänge erzeugt werden.

Werden die Aufnahmen in anderen verlustfreien Formaten abgespeichert, wie es zum Beispiel bei der Migration in aktuellere Formate in vielleicht 20 Jahren notwendig werden wird, führt das nicht zu einer schlechteren Bildqualität. Dauerhaft aufbewahrens-wert sind allein diese hochaufgelösten Dateien, nicht die für konkrete Anwendungszwecke veränderten Arbeitskopien.

3. Beispiel: Im Rahmen der Stadtdokumentation werden mit der Digitalkamera die Straßenzüge eines Viertels durch einen Hobbyfotografen fotografiert.

Der Fotograf benutzt ein gutes Einsteigergerät mit einem Chip, der Aufnahmen mit 5 Megapixel erstellen kann. Die Bilder werden auf der Speicherkarte der Kamera standardmäßig als JPEG mit ganz geringer Komprimierung abgelegt, vom Fotografen auf CD überspielt und dem Stadtarchiv übergeben. Die verhältnismäßig hohe Auflösung erlaubt es, die Fotos in bis zu 21 × 16 cm – also in etwa in DIN A5 – mit 300 dpi auszugeben. Da die Kamera schon die Reduktion auf das JPEG-Format vorgenommen hat, macht eine nachträgliche Überführung in TIFF oder JPEG2000 keinen Sinn – die Qualität würde dadurch nicht verbessert. Erst bei der nächsten Migration in vielleicht 20 Jahren muss darauf geachtet werden, dass die Fotos nicht erneut in einem verlustbehafteten Format abgespeichert werden, da sich in diesem Moment die Artefakte vergrößern würden.

3.

Im Bezug auf die Datensicherung empfiehlt es sich aus technischer Sicht, die Anzahl der verwendeten Dateiformate so gering wie möglich zu halten. Man sollte sich auf ein Speicherformat für die verlustbehaftete Speicherung und eines für die verlustfreie Speicherung konzentrieren. Vermeiden Sie dabei proprietäre oder exotische Formate. Mit JPEG im verlustbehafteten und TIFF bzw. JPEG2000 im verlustfreien Bereich sind sie derzeit auf dem richtigen Weg. Sogenannte RAW-Formate, wie sie durch viele Kamerahersteller angeboten werden, oder die eigenentwickelten Formate einzelner Softwarehersteller sind für langfristige Sicherung nicht geeignet⁶.

Es gibt zur Zeit keine 100 % sicheren Datenspeicher – die Haltbarkeit von selbstgebrannten CDs oder DVDs wird allgemein mit rund 10 Jahren angegeben. Jüngste in der Zeitschrift *Restaurator* veröffentlichte Untersuchungen belegen, dass CD-Rs, also selbst beschriebene CDs, haltbarer sind als DVDs und hier die hellgrünen Rohlinge mit dem Farbstoff Phthalocyanin deutlich besser abschneiden als blaue Rohlinge⁷. Wiederbeschreibbare CDs und DVDs, sogenannte RW oder rewritable CDs, sollte man meiden. Wenn man

diese Träger nutzt, muss man sie unbedingt in dunkler Umgebung in Schutzhüllen bei Raumtemperatur und geringer Luftfeuchte lagern. Stehende Lagerung schützt die Scheiben vorm Verbiegen. Flash-Speicher, wie sie ihn von den Kamera-Speicherkarten und den USB-Sticks her kennen, ist robuster. Seine Zuverlässigkeit nimmt mit der Anzahl der Löschungen und Wiederbeschreibungen ab; eine schleichende Entladung der durch elektrische Aufladung abgelegten Daten führt nach rund 20 Jahren zu Datenverlust. Für die dauerhafte Speicherung eignet sich Flashspeicher so vielleicht etwas besser als CDs, doch ist er auf die Speichergröße bezogen (noch) unverhältnismäßig teurer. Die Speicherung auf Festplatten ist solange sicher, wie die Mechanik reibungslos arbeitet. Leider kündigen Festplatten das Ende ihrer Lauffähigkeit nicht vorher an, sondern beenden ihre Tätigkeit mit einem unangenehmen Geräusch, das keiner vergisst, der es schon einmal gehört hat. Um den Einzelplatzbetrieb sicher zu machen, gibt es sogenannte RAID-Systeme, bei denen die Daten laufend auf eine zweite baugleiche Festplatte im Computer gespiegelt werden. Fällt eine der beiden Speicher aus, sind die Daten auf der anderen in aller Regel noch brauchbar. Sicher ist die Speicherung in der Regel auf einem Netzlaufwerk, das regelmäßig auf einen anderen Server und/oder Bandlaufwerke überspielt wird. Diese Technik findet in vielen Kommunen und Kommunalen Rechenzentren ihren Einsatz. Inwiefern andere Langzeitspeicherformen wie der neu entwickelte Farbmikrofilm eine Alternative in der Langzeitsicherung elektronischer und analoger Bilder bieten können, bleibt abzuwarten⁸.

.....
Datei, die sich ausschließlich am ursprünglichen Aufbau des Formats orientiert.

- 5 JPEG2000 – nicht zu verwechseln mit dem »alten« JPEG-Format – liegt seit 2000 als ISO-Standard 15444 vor. Erweiterungen dieses Formats sind seitdem in Unternormen erfasst worden. Umfangend informiert die englischsprachige Internetseite <http://www.jpeg.org/jpeg2000>, einen guten deutschsprachigen Überblick gewährt <http://de.wikipedia.org/wiki/Jpeg2000>.
- 6 Als Rohdatenformat (RAW-Format) wird ein jeweils modellabhängiges Dateiformat bei Digitalkameras bezeichnet. Die Kamera schreibt die Daten nach der Digitalisierung weitgehend ohne Bearbeitung auf das Speichermedium, so dass eine nachträgliche Bearbeitung und Korrektur besonders gut möglich ist. Zu den Problemen vgl. Spiegel ONLINE vom 20. Juni 2005: Marc Störing, Fotografen fürchten Format-Wirrwarr, <http://www.spiegel.de/netzwelt/technologie/0,1518,360902,00.html>. Inzwischen wurden die Schwierigkeiten auch von Fotografen erkannt, und es wird an einem offenen RAW-Daten-Format gearbeitet. Selbst wenn dieses Format in absehbarer Zeit zur Verfügung steht, schließt sich besonders für die Kunstfotografie die Frage des Originals als dem verbindlichen, vom Fotograf gewollten Endzustand einer Fotografie an. Im RAW-Format werden nur Belichtungszeit und Blende sowie ISO festgelegt, es erlaubt die spätere Änderung aller anderen Parameter wie Weißabgleich, Farbsättigung, Kontrast, Schärfung.
- 7 Joe Iraci, The Relative Stabilities of Optical Disc Formats. In: *Restaurator* 26 (2005), S. 134–150; erst jüngst hat man sich auf ein standardisiertes Verfahren für Langzeittests von optischen Speichersystemen geeinigt: Standard ECMA-379 Test Method for the Estimation of the Archival Lifetetime of Optical Media <http://www.ecma-international.org/publications/standards/Ecma-379.htm>.
- 8 Die Möglichkeit von hochauflösenden Farbmikrofilmen zur Langzeitarchivierung wird derzeit im sog. ARCHE-Projekt untersucht. Es handelt sich um ein Gemeinschaftsunternehmen des Landesarchivs Baden-Württemberg, der Universitätsbibliothek Stuttgart und mehreren Industriepartnern unter der Leitung des Fraunhofer-Instituts für Physikalische Messtechnik IPM in Freiburg mit Unterstützung des Bundesministeriums für Wirtschaft und Arbeit; vgl. eine Projektbeschreibung unter http://www.vdivde-it.de/innonet/projekte/in_pp079_arche.pdf/download (Stand: 19.07.07). Es ist

IPTC-Feldbenennung	IPTC-Beschreibung	Max. Anzahl Zeichen	Archivische Verwendung
Special Instructions	Sonstige Hinweise	256	Archivischer Titel
Object Name	Kurzbezeichnung des Bildes, kann Dateiname sein	64	Signatur
Caption/Abstract	Beschreibung des Bildinhaltes	2000	Originaltitel
Headline	Titel	256	Kurztitel (bei Plakaten: Anlass)
By-line (Author)	Name des Autors/Fotografen	32	Bilder: Fotograf/ Plakate: Grafiker
Credit	Lieferant des Bilds (Agentur)	128	Copyright
Date Created	Aufnahmedatum in der Form CCYYMMDD (z. B. 20070401 für den 1. April 2007)	8	Datierung (wird im Bundesarchiv besonders bestückt)

Das wichtigste Mittel, um Datenverlust zu verhindern, ist eine redundante Datenhaltung: Zu archivierende Bilddaten sollten auf unterschiedlichen Medien räumlich voneinander getrennt aufbewahrt werden. Um die Lesbarkeit dauerhaft sicher zu stellen, müssen die Datenträger regelmäßig kontrolliert und ggf. auf neue Datenträger überspielt werden. Damit für die Prüfung nicht jede Bilddatei aufgerufen werden muss und auch kleinste Veränderungen in der Bitfolge entdeckt werden, die sich nicht unbedingt im optischen Eindruck des Bildes sofort niederschlagen, gibt es automatisierbare Verfahren. Diese nutzen zumeist den sog. Message-Digest Algorithm 5 oder kurz MD5-Hashwert. Mittels einer bei den Bildern abgespeicherten Prüfsumme sind diese Tests auf Unverfälschtheit leicht durchführbar⁹. Sollte absehbar sein, dass eines der verwendeten Bildformate wie TIFF durch die gängige Software nicht mehr unterstützt wird, müssten die in diesem Format vorhandenen Daten möglichst verlustfrei in ein dann gültiges Standardformat überführt werden.

Unachtsamkeit bei der Datenhaltung hat im Gegenzug weitaus schlimmere Folgen als bei der Verwahrung von konventionellem Archivgut: Schlecht gepflegte Digitalaufnahmen sind nicht mehr lesbar und somit weg¹⁰!

Eng mit der physikalischen Speicherung hängt die Wiederauffindbarkeit und Identifizierbarkeit der digitalen Bilder zusammen. Genauso wie bei einer konventionellen Fotografie, die aus einem Album gefallen ist und sich nicht mehr zuordnen lässt, verliert auch das digitale Bild viel von seiner Aussagekraft, wenn man nicht mehr weiß, woher es stammt. Drei Wege können dieses »Verlegen« von Digitalisaten verhindern:

(1) Strukturierte Datenablage

Alle digitalisierte Archivalien in einen Ordner »Scans« abzulegen oder in Ordnern, die das dargestellte Objekt beschreiben, wie »Marktplatz« oder »Portraits« stellt keine tragfähige Lösung dar. Ab einer gewissen Anzahl an Dateien verliert man die Übersicht. Eine Möglichkeit ist deshalb, die digitalisierten Archivalien und Digitalbilder in einer Ordnerstruktur abzulegen, die der Beständeordnung entspricht. Da der Dateiname

nicht eine richtige Erschließung ersetzen kann, kann er stattdessen Rückschlüsse auf die Signatur des dargestellten Objekts ermöglichen.

Wer Bestandssignaturen hat, kann diese als führende Stellen verwenden, gefolgt von der Signatur und ggf. bei mehrseitigen Stücken von der Blatt- oder Seitenzahl. Zur Trennung der einzelnen Stellen verwendet man am einfachsten den Unterstrich. Bei dem im LWL-Archivamt verwendeten Schema würden sich so die folgenden zwei Beispiele ergeben:

Eine Datei mit dem Namen C.Lim.Lim_001_V.tif führt auf das Bild der Vorderseite der Urkunde Nr. 1 von 1322 aus dem Bestand Limburg-Styrum. Ein Bild, das A.01.01.101_408_241_V.tif heißt, zeigt das Papiersiegel und Unterschrift des preußischen Königs Friedrich Wilhelm III. auf dem Statut für eine zu errichtende Hülf-Kasse für die Provinz Westfalen von 1831. Das Stück stammt aus dem Archiv des Landschaftsverbands, aus dem Bestand »Provinziallandtag, Provinzialstände« und von der Vorderseite des Blatts 241 der Akte 408 »Urkunden des 4. Westfälischen Landtages«.

Das Digitalisat wird über die Bezeichnung der Ordnung und den Dateinamen eindeutig in seinen Verzeichnungs-Kontext verwiesen. Sogar der Dateiname allein würde das im Notfall ermöglichen. Diese Variante setzt voraus, dass man eine relativ stabile Beständeordnung hat, da bei eingreifenden Veränderungen alle Dateinamen neu vergeben werden müssten. Das bedeutet wiederum einen Eingriff in die Originaldaten, den man als potentielle Quelle für Verfälschungen und Fehler nur möglichst selten vornehmen sollte. Nach einer Umbenennung müssten auch die Prüfsummen der Dateien neu berechnet werden.

.....
inzwischen möglich, einen solchen Film herzustellen, allerdings liegen die Kosten pro Aufnahme derzeit noch sehr hoch.

⁹ Für die Prüfung ganzer Verzeichnisse oder Datenspeicher gibt es das kostenlose Programm »MD5-Summer«, auch Dateiverwaltungsprogramme wie der Total Commander können auf Knopfdruck MD5-Prüfsummen (hier als »CRC-Quersumme« bezeichnet) bilden.

¹⁰ Wegen dieser Aufwände der Datenpflege, die gerade bei Schutzdigitalisierungen einer »doppelten Bestandserhaltung« (für die analogen Originale und die Digitalisate) gleichkommt, warnt Ulrike Gutzmann, Bestandserhaltung in digitalen Zeiten: Gefahren – Möglichkeiten – Perspektiven. In: Archiv und Wirtschaft 38 (2005), S. 189–197 vor blindem »Drauflosdigitalisieren« und rät in einer Kosten-Nutzen-Abwägung ausschließlich zur Digitalisierung von vielbenutzten bzw. stark gefährdeten Archivalien.

(2) Verzeichnungsprogramme

Eine weitere Möglichkeiten, Bilder zu verwalten und zugänglich zu halten, ist, ihnen bei Ankunft im Archiv eine eindeutige Identifikation zuzuweisen (z. B. eine aufsteigende Nummer, ggf. mit führenden Nullen) und diese in einem Verzeichnungsprogramm zu erfassen¹¹. Die Einordnung in die Tektonik erfolgt dann nur virtuell in der Verzeichnungsdatenbank, wo das digitale Bild ggf. auch in Kleinansichten, sogenannten Thumbnails, hinterlegt wird. Vorteil ist hierbei, dass neben der Erschließung auch die Nutzung des Bildbestands durch Dritte erleichtert wird. Als Nachteil bleibt, dass bei Verlust oder Zerstörung der Datenbank auch die Zuordnung der Bilder schwer fällt. Auf eine dateiinternen Verweis sollte also trotzdem nicht verzichtet werden. Die Datenhaltung der Digitalbilder sollte grundsätzlich nicht innerhalb der Datenbank erfolgen – obwohl das inzwischen in Form sogenannter »Binary Large Objects« kurz BLOBs technisch auch möglich ist, da dadurch das Ausgliedern einzelner Bilddateien und die Datenpflege erschwert werden.

(3) Metadaten

Eine zusätzliche Sicherung der Erschließungsdaten und eine Möglichkeit, Bildinhalte auch außerhalb von Verzeichnungsprogrammen recherchierbar zu machen, ist die Bestückung der Bilder mit Metadaten. Sowohl das TIFF- als auch das JPEG- und das JPEG2000-Format bieten die Möglichkeit an, beschreibende Inhalte in den Kopf der Datei, den sogenannten Header, abzuspeichern. Traditionell gibt es dafür zwei Datenmodelle: Den EXIF-Datensatz für die Abspeicherung der technischen Metadaten und den IPTC-Datensatz für die inhaltlichen und rechtlichen Metadaten. Fast alle gängigen Kamera- und Scannermodelle nutzen die EXIF-Daten, um dort Datum und Uhrzeit der Aufnahme, Belichtungszeit, Blende, Bildgröße, Farbtiefe, Kamerahersteller und -modell festzuhalten. Dass diese Informationen ggf. auch inhaltliche Rückschlüsse erlauben, die für die Erschließung hilfreich sein können, liegt auf der Hand¹².

Der zweite Metadatenatz wurde schon in den Anfängen der digitalen Fotografie durch das International Press and Telecommunications Council, eine internationale Pressevereinigung beschlossen, um Urheber- und Kopierrechte im Bild festzuhalten. Späteren Interessenten sollte dadurch ermöglicht werden, Kontakt zum Fotografen oder der Bildagentur aufzunehmen, um die Veröffentlichungsrechte an einer Aufnahme zu erwerben. Beschreibende Informationen sollten Suchmaschinen erlauben, die Bildbestände nach Personen, Orten oder Ereignissen zu durchforsten. Die Verwandtschaft zur Erschließung im Archiv liegt auf der Hand und wird u. a. vom Bundesarchiv genutzt. Herr Dr. Sander, der dort für diese Thematik zuständig ist, war so freundlich, mir die Bestückung der IPTC-Daten, wie sie in seinem Haus für Bilder, Plakate und Luftbilder praktiziert wird, zur Verfügung zu stellen¹³.

Weitere Felder, die nicht schon mit solchen Standard-Zuweisungen belegt sind, werden u. a. für Informationen zu Bestand, Klassifikation, Altsignatur, Ortsbezeichnung und für den Vermerk genutzt, ob eine Internetveröffentlichung erlaubt ist. Bei Luftbildern werden geografische und flugtechnische Daten, bei

Plakaten ggf. Verlagsdaten ergänzt. Die Bestückung dieser Felder erfolgt aus der normalen Erschließungsmaske des Archivprogramms automatisch, d. h. man muss die Einträge natürlich nicht per Hand in den Kopf der TIFF- oder JPEG-Datei kopieren. Leider unterstützen die bisher auf dem Markt befindlichen Archivprogramme nur selten eine solche Funktionalität, obwohl die technische Realisierung meines Erachtens leicht möglich und auch jetzt schon mit kostenlosen Tools halbautomatisch durchzuführen ist¹⁴. Sollte man keine »sprechenden Dateinamen« vergeben haben, die die Lokalisierung innerhalb der Bestände ermöglichen, sollten entsprechende Angaben unbedingt hier hinterlegt werden. Auch die Funktion eines digitalen Objekts als »Original«, »Sicherungskopie« oder »Nutzungskopie« können hier untergebracht werden. Dokumentieren Sie schließlich ihre Nutzung der IPTC Felder, so dass notfalls auch ihre Nachfolger die Chance haben, ihre Überlegungen nachzuvollziehen.

Aus einem so aufbereiteten Fotobestand kann man bei einem Notfall wie dem Verlust oder der Zerstörung der Datenbank die Erschließungsdaten automatisiert rekonstruieren. Ein Nebeneffekt ist, dass die Einträge im Copyright-Feld die Herkunft der Bilder aus dem eigenen Archiv belegen ohne mit sogenannten »optischen Wasserzeichen« den Eindruck des Archivale zu beeinträchtigen.

Wie bei allem in der Computerwelt gibt es auch im Bereich der Metadatenätze Neuerungen und Verbesserungen. So versucht der Marktführer im Bereich der kommerziellen Bildbearbeitung, Adobe, seit einiger Zeit seinen eigenen Metadatenatz namens XMP (= Extensible Metadata Platform) durchzusetzen¹⁵. In ihm sollen technische und inhaltliche Daten einfließen und er soll durch seine XML-basierten Aufbau flexibler als die bisherigen Lösungen sein. Da IPTC durch Adobe nicht weiter unterstützt wird, ist damit zu rechnen, dass sich XMP auf Dauer durchsetzen und die Nutzung des IPTC-Standards zurückgehen wird. Hier gilt es, die Entwicklung im Auge zu behalten und im eigenen Haus ggf. neue und ältere Software auf Kompatibilität zu untersuchen. Es ist dabei schon jetzt zu meist so, dass es automatisierte Wege gibt die IPTC-

11 In diese Kategorie fällt auch: Sonja Meldau, Erschließung durch eine Bilddatenbank. Die Abteilung Karten und Bilder der Bayerischen Staatsbibliothek. In: Rundbrief Fotografie 13 (2006), S. 15–17.

12 Das EXIF = Exchangeable Image File Format ist ein von der Japan Electronic and Information Technology Industries Association (JEITA) festgelegter Standard. Die neuesten Bemühung zur Standardisierung technischer Metadaten erfolgt in der NISO Z39.87–2006: »Data Dictionary – Technical Metadata for Digital Still Images«, der Text ist als PDF auf den Seiten der National Information Standard Organization (<http://www.niso.org>) zu finden.

13 Ganz herzlich möchte ich an dieser Stelle Herrn Dr. Oliver Sander, Bundesarchiv, Referat B 6, Potsdamer Str. 1, 56075 Koblenz für die freundliche Überlassung der Metadatenbeschreibung danken.

14 Eine solche Metadatenangabe erlaubt beispielsweise das Programm kostenlose Programm Exifer und einige Bildbetrachter. Leider werden die Einträge in den IPTC-Feldern nicht von allen diesen Programmen in gleicher Weise vorgenommen. Testläufe im Vorfeld der Anwendung sind deshalb sehr wichtig. Manchmal können auch von der Pressestelle od. vergleichbaren Einrichtungen eingegebene Metadaten die archivische Erschließung erleichtern. Hier sollte innerhalb einer Verwaltung dringend auf die gegenseitige Austauschbarkeit Rücksicht genommen werden.

15 Adobe stellt die XMP-Spezifikationen sowie ein sog. »Software Development Kit« unter einer Open-Source-Lizenz zur freien Verfügung, <http://www.adobe.com/products/xmp>.

und EXIF-Altdateien in das neuere XMP überführen. Die »Rückwärtskompatibilität«, d. h. die Interpretation von XMP-Daten durch ältere Programme und der Abgleich von IPTC-, EXIF- und XMP-Informationen innerhalb einer Datei ist zumeist nicht vorgesehen¹⁶. Bei der Nutzung von verschiedenen Programmen zur Metadatenbestückung kann es schlimmstenfalls zu unterschiedlichen Angaben in den in einer Datei enthaltenen Metadatensätzen kommen.

4.

Obwohl gerade zu den technischen Details noch Ergänzungen gemacht werden könnten, soll zuletzt auf die Nutzung von Digitalbildern im Archiv eingegangen werden. Die computergestützte Bereitstellungen von Erschließungsinformationen wird heute schon in vielen Archiven praktiziert, doch auch die Bereitstellung von digitalisierten Dokumenten selbst im Lesesaal oder im Internet, sowie der Versand von Digitalisaten an Stelle analoger Kopien wird immer häufiger durchgeführt. Dieser Trend ist zum Beispiel in den Niederlanden schon sehr viel weiter als bei uns fortgeschritten, wird aber auch hier ankommen: der »digitale Lesesaal«¹⁷. Für viele Studenten aber auch für viele Heimat- und Familienforscher gilt bereits heute in Abwandlung des bekannten Rechtsgrundsatzes: »Quod non est in Google non est in mundo« – Was nicht bei Google auftaucht, gibt es auch nicht. Wenn Archive weiterhin als Partner für Bildung und Forschung wahrgenommen werden wollen, geht kein Weg an der breiten Bereitstellung von Erschließungsinformationen und letzten Endes auch von digitalisierten Archivalien vorbei. Zumindest als Appetitmacher auf einen persönlichen Besuch im Archiv sollten wir gerade attraktive Bild-, Plakat- und Kartenbestände für unsere Internetpräsenzen nutzen. Welche engen Grenzen dabei das neue Urhebergesetz setzt, die bei historischen Aufnahmen nicht unbedingt angemessen und im Sinne der interessierten Öffentlichkeit erscheinen, beschreibt Mark Steinert in seinem Beitrag in diesem Heft.

Zumindest mittelfristig wird jedoch ein Großteil der Nutzung – auch von digitalen Archivalien – weiterhin im Benutzerraum des Archivs erfolgen¹⁸. Sowohl für die Recherche als auch für die Präsentation dieser Objekte müssen hier Computerarbeitsplätze zur Verfügung stehen. Die Herausgabe von Benutzungskopien auf einem elektronischen Medium wie einer CD oder DVD und die Durchsicht auf dem vom Benutzer mitgebrachten Laptop ist vielleicht für manche Bestände eine Alternative. Es muss jedoch jedem klar sein, dass man die Benutzer nur schwer davon abhalten kann, von dieser CD/DVD Daten auf die eigene Festplatte zu überspielen und unkontrolliert mitzunehmen. Vielleicht nimmt einem das im Archivaltag manchen Kopierauftrag ab, jedoch sollte man die Benutzer in der gewonnenen Zeit über Nutzungsaufgaben und Kopierrechte aufklären. Um lange Ladezeiten zu verhindern und die Dateien vor Verfälschung und Zerstörung zu schützen, ist es unumgänglich, dem Benutzer nur Kopien in einem Arbeitsformat (z. B. JPEG) anzubieten. Für diese Kopien reicht es aus, wenn diese Datei für die Bildschirmdarstellung mit 72 dpi in Originalgröße mit leichter Komprimierung abgespeichert wird. Will man zur besseren Lesbarkeit eine Vergrößerungsmög-

lichkeit auf doppelte Originalgröße erlauben, so genügen 150 dpi. Die Darstellung kann über ein Benutzungsmodul des Archivprogramms oder über eine externe, für den nichtkommerziellen Gebrauch zumeist kostenlose Ansichtsoftware (sog. Viewer¹⁹) erfolgen.

Bei der Vorlage ganzer Akten oder Fotoalben in digitaler Form, kann das bei konsequenter Namensvergabe ebenfalls in einem solchen Viewer erfolgen. Bewährt hat sich in diesen Fällen als Alternative auch die Präsentation in einer aus den zugehörigen Dokumentenseiten erstellten PDF-Datei.

Jede dieser Formen ist für den Benutzer weit angenehmer als beispielsweise die Nutzung von Mikrofilm oder gar von Mikrofilm. Eine konsequente und verständliche Dateibenennung oder eine entsprechende Verlinkung im Verzeichnungsprogramm ermöglichen die zielgenaue Anwahl des gewünschten Archivals. Für die Archive verkürzt und erleichtert ein digital vorliegendes Archival das Ausheben und auch die Anzahl der Fehlausehebungen von unhandlichen Stücken kann z. B. durch digitale Vorschaubilder von Karten und Plänen reduziert werden. Bei besonders empfindlichen oder vielbenutzten Stücken schont die digitale Ersatzform das Original²⁰.

Ähnlich wie bei Mikroformen geht dem Benutzer der haptische Eindruck des Stücks verloren. Da in vielen Fällen aber zumindest die Farbigkeit der Abbildung erhalten werden kann, ist der Bildschirmindruck oft näher am Original. Gerade älteren Benutzern kann die Vergrößerungsmöglichkeit am Bildschirm sogar eine besondere Hilfe sein.

16 Geplant ist u. a. die Implementierung eines »IPTC Core« in XMP, der alle wichtigen IPTC-Felder im neuen Metaformat abbildet. Da es sich bei XMP nicht um einen Standard, sondern um einen »proprietären Metadaten-Container mit offener Spezifikation handelt«, so Andrea Trinkwalder, Für die Ewigkeit. Metadatenstandards fürs Bildarchiv. In: c't Magazin 16/06, S. 156–158, hier S. 158 sind Änderungen und schwerwiegende Eingriffe durch Adobe nicht auszuschließen.

17 Angelika Menne-Haritz, Digitaler Lesesaal, virtuelle Magazine und Online-Findbücher. Auswirkungen der Digitalisierung auf die archivischen Fachaufgaben. In: Hartmut Weber und Gerald Maier (Hgg.), Digitale Archive und Bibliotheken. Neue Zugangsmöglichkeiten und Nutzungsqualitäten, Stuttgart 2000, S. 25–34; für die niederländische Entwicklung vgl. zuletzt Yvette Hoiink, Die Skizze des virtuellen Lesesaals. In: Archivpflege in Westfalen-Lippe 66 (2007), S. 32–33.

18 Die nordrhein-westfälische Archivgesetzgebung enthält keinen grundsätzlichen Hinderungsgrund, Archivalien im Internet zu veröffentlichen vgl. Michael Grünberger, Archivrechtliche Fragen der Präsentation von Findmitteln und Archivgut im Internet. Archivpflege in Westfalen und Lippe 54 (2001), S. 15–21; zu den Problemen der Internetveröffentlichung von gedruckten und ungedruckten Findmitteln und Regesten- sowie Faksimilewerken vgl. Rainer Polley, Rechtsfragen bei der Präsentation und Benutzung digitaler Publikationen im archivischen Kontext. In: Archivpflege in Westfalen-Lippe 63 (2005), S. 33–39. Da über die rechtlichen Rahmenbedingungen der Veröffentlichung von Findmitteln noch Unsicherheit bestand, hat die Archivreferentenkonferenz des Bundes und der Länder (ARK) ein entsprechendes Rechtsgutachten in Austrag gegeben: »Bereitstellung elektronischer Findmittel in öffentlich zugänglichen Netzen«. Es kann bei Bedarf bei Abteilung 2 Grundsatzfragen und Öffentlichkeitsarbeit des Landesarchivs NRW angefordert werden.

19 Es seien hier nur drei Viewer-Produkte in alphabetischer Folge genannt: FastStone Image Viewer (<http://www.faststone.org>), Irfanview (<http://www.irfanview.de>), XnView (<http://www.xnview.de>).

20 Zur Chance Benutzungssintensität als Planungsinstrument einzusetzen vgl. Peter Worm, Das Staatsarchiv Münster und seine Benutzer (1995–2004). Transferarbeit im Rahmen des Referendariats für den höheren Archivdienst. Münster 2005 <http://archive.nrw.de/dok/>