

H. Neuroth, A. Oßwald, R. Scheffel, S. Strathmann, M. Jehn (Hrsg.)

nestor Handbuch

Eine kleine Enzyklopädie
der digitalen Langzeitarchivierung

Version 2.0

Kapitel 18.3
Langzeitarchivierung eines digitalen
Bildarchivs – Projekt zum Auf-
bau eines Langzeitarchivs für
hochaufgelöste digitale Bilddateien
der Staatsgalerie Stuttgart am BSZ

nestor Handbuch: Eine kleine Enzyklopädie der digitalen Langzeitarchivierung
hg. v. H. Neuroth, A. Oßwald, R. Scheffel, S. Strathmann, M. Jehn
im Rahmen des Projektes: nestor – Kompetenznetzwerk Langzeitarchivierung und
Langzeitverfügbarkeit digitaler Ressourcen für Deutschland
nestor – Network of Expertise in Long-Term Storage of Digital Resources
<http://www.langzeitarchivierung.de/>

Kontakt: editors@langzeitarchivierung.de
c/o Niedersächsische Staats- und Universitätsbibliothek Göttingen,
Dr. Heike Neuroth, Forschung und Entwicklung, Papendiek 14, 37073 Göttingen

Die Herausgeber danken Anke Herr (Korrektur), Martina Kerzel (Bildbearbeitung) und
Jörn Tietgen (Layout und Formatierung des Gesamttextes) für ihre unverzichtbare
Unterstützung bei der Fertigstellung des Handbuchs.

Bibliografische Information der Deutschen Nationalbibliothek
Die Deutsche Nationalbibliothek verzeichnet diese Publikation in der Deutschen
Nationalbibliografie; detaillierte bibliografische Daten sind im Internet unter
<http://www.d-nb.de/> abrufbar.

Die Inhalte dieses Buchs stehen auch als Onlineversion
(<http://nestor.sub.uni-goettingen.de/handbuch/>)
sowie über den Göttinger Universitätskatalog (<http://www.sub.uni-goettingen.de>) zur
Verfügung.

Die digitale Version 2.0 steht unter folgender Creative-Commons-Lizenz:
„Attribution-Noncommercial-Share Alike 3.0 Unported“
<http://creativecommons.org/licenses/by-nc-sa/3.0/>



Einfache Nutzungsrechte liegen beim Verlag Werner Hülsbusch, Boizenburg.
© Verlag Werner Hülsbusch, Boizenburg, 2009
www.vwh-verlag.de
In Kooperation mit dem Universitätsverlag Göttingen

Markenerklärung: Die in diesem Werk wiedergegebenen Gebrauchsnamen, Handelsnamen,
Warenzeichen usw. können auch ohne besondere Kennzeichnung geschützte Marken sein und
als solche den gesetzlichen Bestimmungen unterliegen.

Druck und Bindung: Kunsthaus Schwanheide

Printed in Germany – Als Typoskript gedruckt –

ISBN: 978-3-940317-48-3

URL für Kapitel 18.3 „Langzeitarchivierung eines digitalen Bildarchivs - Projekt zum Aufbau
eines Langzeitarchivs für hochaufgelöste digitale Bilddateien des Staatsgalerie Stuttgart am
BSZ“ (Version 2.0): <urn:nbn:de:0008-20090811980>
<http://nbn-resolving.de/urn/resolver.pl?urn=urn:nbn:de:0008-20090811980>



Gewidmet der Erinnerung an Hans Liegmann (†), der als Mitinitiator und früherer Herausgeber des Handbuchs ganz wesentlich an dessen Entstehung beteiligt war.

18.3 Langzeitarchivierung eines digitalen Bildarchivs – Projekt zum Aufbau eines Langzeitarchivs für hochaufgelöste digitale Bilddateien der Staatsgalerie Stuttgart am BSZ

Werner Schweibenz und Stefan Wolf²¹

Der Beitrag beschreibt das Projekt der Staatsgalerie Stuttgart mit dem Bibliotheksservice-Zentrum Baden-Württemberg (BSZ) zum Aufbau eines Langzeitarchivs für hochaufgelöste digitale Bilddateien. Für die Archivierung wird das Langzeitarchiv SWBdepot des BSZ verwendet, die Metadaten für die Langzeitarchivierung werden mit dem Objektdokumentationssystem IMDAS-Pro erzeugt.

Die Situation in der Staatsgalerie Stuttgart

Die international bedeutsamen Sammlungen der Staatsgalerie Stuttgart (SGS)²² reichen vom Tafelaltar bis zur modernen Medienkunst mit praktisch allen Formen z.B. der Malerei, Plastik, Graphik bis hin zu raumgreifenden Installationen. Sie werden fortlaufend erforscht, erschlossen und dokumentiert, gleichzeitig aber auch in verschiedenen Zusammenhängen eingebunden: beispielsweise in den eigenen Ausstellungen, im Leihverkehr zwischen Museen, in der Museumspädagogik, in Publikationen oder auch in Internetauftritten. Das Fotoatelier der SGS fertigt für diese Zwecke laufend eine große Zahl qualitativ hochwertiger Fotos an, dem Fortschritt der Technik folgend heute mit einer hochauflösenden Digitalkamera. Die anfallende Datenmenge wächst kontinuierlich. Für die Zukunft sucht die SGS nach einer kostengünstigen Lösung für Datenspeicherung und –sicherung bei gleichzeitiger Nutzung im Verbund. Dies gilt auch für die aus Sicherheitsgründen notwendige redundante Speicherung an einem zweiten Ort. Als Mitglied im MusIS-Verbund, dem landeseinheitlichen Verfahren für Museumsdokumentation der Staatlichen Museen in Baden-Württemberg wandte sich die SGS an das Bibliotheksservice-Zentrum Baden-Württemberg

21 Die Autoren danken herzlich den Kolleginnen und Kollegen der Staatsgalerie Stuttgart – allen voran Frau Dr. Elke Allgaier – für die kollegiale Unterstützung bei der Erstellung dieses Kapitels zum nestor Handbuch.

22 <http://www.staatsgalerie.de/>

BSZ),²³ das bereits für andere Institutionen ähnliche Dienstleistungen zu Dokumentmanagement und Langzeitarchivierung (LZA) anbietet.

Das Angebot des BSZ im Bereich Langzeitarchivierung

Das BSZ bietet auf dem Gebiet der Langzeitarchivierung eine umfangreiche Dienstleistungspalette an.²⁴ Der Fokus des BSZ richtet sich auf die Verbindung der Objekte mit einer qualitativ hochwertigen Dokumentation und der Schaffung eines Mehrwerts für seine Kunden. Der Gewinn liegt z.B. in der technischen Realisierung spezifischer Geschäftsgänge für verschiedene Häuser auf gemeinsamer Basis, der gemeinsamen Nutzung der gleichen Software oder in der automatisierten Erzeugung und Bereitstellung von Gebrauchsderivaten aus den archivierten Objekten – also auf klassischen Synergieeffekten. Mit diesem Ansatz vereint das BSZ scheinbar Gegensätzliches, indem Produktion, Dokumentation und Nutzung mit der Archivierung wertvoller Daten in einen durchgängigen Arbeits- und Archivierungsprozess gebündelt werden. Für die Kunden wird ein sicheres und kostensparendes Outsourcing großer Datenmengen in Kombination mit einer sicheren Datenhaltung der Originaldateien und einem flexiblen Zugriff auf Derivate geboten, also Mehrwerte, die eine Institution braucht, um sicher, rationell und ökonomisch arbeiten zu können. Für die Bildproduktion der Museen erschien die Durchführung eines Pilotprojekts als notwendig, das vom Ministerium für Wissenschaft, Forschung und Kunst Baden-Württemberg gefördert wird.

Die Leitidee

Die Grundidee des Projektes zum „Aufbau eines Langzeitarchivs im BSZ für hochaufgelöste digitale Bilddateien der Staatsgalerie Stuttgart sowie die Entwicklung eines sicheren und auf Kontinuität basierenden Online-Daten-Transfers der Digitalisate“ (so der Name im Projektantrag) ist, unter strikter Beachtung des Primats einer sicheren und nachhaltigen Archivierung ein praktisches Verfahren zu entwickeln, bei dem möglichst nur ein Computersystem für die Dokumentation der zu archivierenden Bilddateien benötigt wird und bei dem

23 <http://www.bsz-bw.de/>

24 Wolf, Stefan ; Mainberger, Christof ; Schweibenz, Werner: Langzeitarchivierung am Bibliotheksservice-Zentrum Baden-Württemberg : Konzept, Aktivitäten und Perspektiven. – Preprint – Konstanz, BSZ, 2009.

URL: <http://opus.bsz-bw.de/swop/volltexte/2009/465/>

Erscheint voraussichtlich in: Bibliotheksdienst, Heft 43(2009), Heft 3.

die Mitarbeiter in allen beteiligten Arbeitsbereichen im Museum möglichst nur mit den ihnen bereits vertrauten Programmen und Werkzeugen arbeiten. Dieser Weg erhöht die Akzeptanz im Museum und hält den Einarbeitungs- und Schulungsaufwand gering. Gleichzeitig ist dies die Voraussetzung des zweiten Projektziels, das Verfahren nach Abschluss des Pilotprojekts den anderen Museen im MusIS-Verbund zur Verfügung zu stellen. Die Abläufe und Daten integrieren sich nahtlos in das Langzeitarchiv SWBdepot des BSZ. SWBdepot bezeichnet die am BSZ in Betrieb befindliche Speicherinfrastruktur, die nach Bedarf ausgebaut wird und auf der die üblichen Prozesse der Datensicherung wie z.B. Bandsicherung, redundante Speicherung, Konsistenzprüfung abgewickelt werden.

Der Produktionsablauf

Den Auftakt in der Produktion gibt die Bildbestellung, die in der Dokumentation festgehalten wird: nach ihrer Anweisung erstellt das Fotoatelier der SGS die Bildaufnahmen. Seit Januar 2005 arbeitet das Fotoatelier der SGS eingebettet in eine hausintern festgelegte Digitalisierungsstrategie mit einer Digitalkamera: Arbeitsprozesse, Geschäftsgänge und Dokumentation werden durchgängig elektronisch unterstützt und ausgeführt. Die professionellen Qualitätsansprüche an die Bildproduktion löst die digitale Fotografie mittlerweile ein. Je Museumsobjekt entsteht mindestens eine hochauflösende, unkomprimierte Masteraufnahme und ein farbkorrigierter, verlustfreier Submaster gleicher Auflösung im Tagged Image File Format (TIFF)²⁵ von durchschnittlich 50 MB bei 8 Bit Tiefe pro Farbkanal.

Das TIFF-Format bietet die Möglichkeit, IPTC-Metadaten²⁶ im Bild zu erfassen und zu speichern. Dieser Quasistandard der Pressefotografie erlaubt Angaben z.B. zu Bildrechten, Fotograf, Titel, Auflösung und Pixelzahl zu machen. Produkte wie z.B. Adobe Photoshop bieten dafür Erfassungsmasken, die im Fotoatelier von den Fotografen ausgefüllt werden. Teilweise handelt es sich um Angaben, die standardmäßig für jede Aufnahme aus der Staatsgalerie Stuttgart gemacht werden und in den Masken schon vorbelegt sind, teilweise aber auch um individuelle Merkmale, die zur einzelnen Fotografie eingetragen werden.

Sowohl Master als auch Submaster erhalten eindeutige Dateinamen, welche Hinweise auf Bildherkunft, Künstlernamen, Inventarnummer, Aufnahmegegebenheiten und Dateiformat enthalten. Zur Dateinamensgestaltung existiert eine formale und semantische Absprache zwischen SGS und BSZ. Die Konvention

25 http://de.wikipedia.org/wiki/Tagged_Image_File_Format

26 Vgl. <http://de.wikipedia.org/wiki/IPTC-NAA-Standard>

wird strikt eingehalten, da der Dateiname die Ablage im Archiv steuert und mit den IPTC-Daten zur Dokumentation herangezogen wird. Ideal ist es deshalb, wenn vor dem Transfer der Bilddatei das Objekt selbst schon als IMDAS-Pro-Museumsobjekt im Dokumentationssystem der SGS erfasst ist.

Vor dem sFTP-Transfer²⁷ der Dateien an das BSZ wird zur Sicherheit der Integrität die MD5-Prüfsumme²⁸ berechnet. Die Tagesproduktion wird gebündelt und an das BSZ durch einen Cron-Job nachts transferiert. Alle folgenden Arbeitsprozesse werden zur Kontrolle der Datenintegrität mittels Vergleich bzw. Erhebung der MD5-Prüfsumme begleitet.

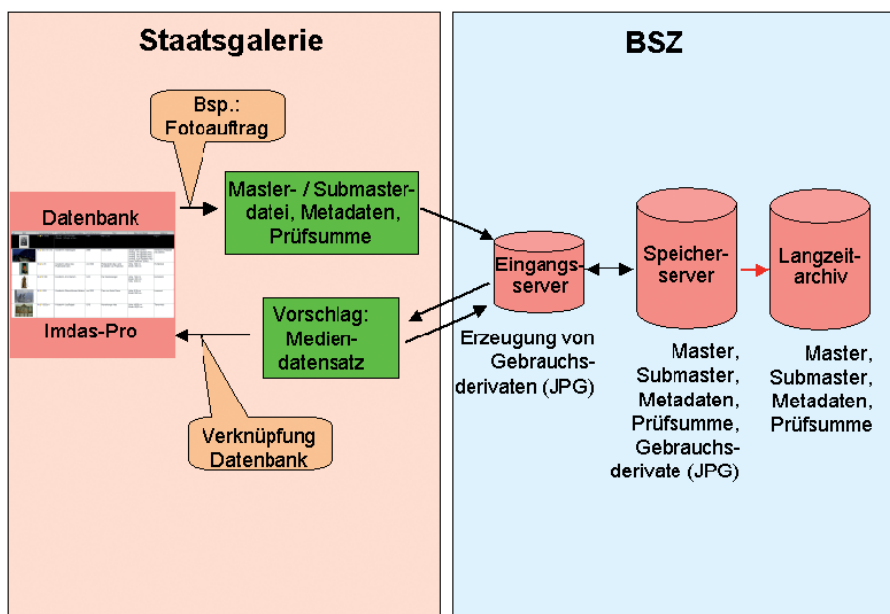


Abbildung 1²⁹

27 http://de.wikipedia.org/wiki/SSH_File_Transfer_Protocol

28 http://de.wikipedia.org/wiki/Message-Digest_Algorithm_5

29 Abbildung 1: Allgaier, Elke (Staatsgalerie Stuttgart): Archivierung von digitalen Bilddaten. Beitrag des nestor-Seminars „Digitale Langzeitarchivierung in Museen und Archiven - Konzepte und Strategien“ Köln, 21.11.2008. In: URL: http://www.langzeitarchivierung.de/downloads/2008-11-21_allgaier.pdf

Aufgabe des BSZ ist die Übernahme, Archivierung und Bereitstellung der Bilddaten. Nach Prüfung der formalen Eingangsvoraussetzung (Einhaltung der Dateinamenskonvention, korrekte Benennung des Dateiformats, gleichzeitige Lieferung von Master, Submaster und Prüfsumme etc.) werden die Bilder vom offenen FTP-Bereich in einen sicheren Arbeitsbereich kopiert. Der Master befindet sich danach in einem geschützten Archivbereich. Er wird nicht weiter verwendet, sondern bildet die Grundlage, wenn später auf die Ursprungsdatei der Bildaufnahme zurückgegriffen werden muss. Der Submaster wird physisch dupliziert – eine Kopie wird mit dem Master im Langzeitarchiv abgelegt, die zweite Kopie bereitgestellt und zur Erzeugung der Gebrauchsderivate im JPG-Format herangezogen. Nötig sind bislang jeweils ein kleines Thumbnail für die Vorschau in IMDAS-Pro und für die Präsentation in BAM, dem gemeinsamen Portal zu Bibliotheken, Archiven und Museen³⁰, sowie eine größere Version zur differenzierten Betrachtung am Bildschirm oder für Restaurierung, Kunstvermittlung und entsprechende Zwecke.

Im gleichen Prozess werden Dateiname und IPTC-Daten gelesen und zusammen verarbeitet. Als Ergebnis entsteht eine Importdatei, die die nötigen Metadaten für eine Vorerfassung der Fotografie als Medienobjekt³¹ in IMDAS-Pro bereitstellt. Sie enthält einen aus dem Dateinamen abgeleiteten Vorschlag zum Künstlernamen und zur Inventarnummer. Wegen der technischen Gegebenheiten in den Zeichensätzen kann die endgültige Ansetzung nicht transportiert werden. Die Importdatei enthält aber auch Angaben zu den technischen Daten der Fotografie und zu den Adressen bzw. Speicherorten der für die Nutzung bereitgestellten Derivate und Submaster, die aus den IPTC-Daten entnommen werden. Die Importdatei wird täglich in IMDAS-Pro eingelesen.

An dieser Stelle wird wieder die Staatsgalerie Stuttgart aktiv: die Dokumentation der Fotografie als IMDAS-Pro-Medienobjekt wird fertig gestellt und mit der Dokumentation des originalen Kunstwerks als IMDAS-Pro-Museumsobjekt verknüpft. Die Vorschläge zu Künstlernamen und Inventarnummer aus der importierten Vorerfassung werden in die endgültige, korrekte Ansetzungsform gebracht. Die Bilder werden geprüft und freigegeben, so dass nun auch im Dokumentationswerkzeug eine Vorausschau auf das Bild verfügbar ist und weitere Derivate sowie der Submaster per Mausklick nach Berechtigung angefordert werden können.

30 <http://www.bam-portal.de/>

31 Das IMDAS-Pro-Medienobjekt ist eine programminterne Dokumentationsklasse, die der Aufnahme von Objekten dient, die mit der Dokumentationsklasse Museumsobjekt in IMDAS-Pro verknüpft werden kann.

Die Verarbeitungsprogramme des BSZ sind modular verkettet, konfigurierbar und parametrisierbar. Weitere Gebrauchsderivate können auf Anforderung erzeugt, weitere Inhalte aus dem IPTC-Header ausgelesen werden. Zur Bildverarbeitung im BSZ ist die OpenSource-Software ImageMagick³² integriert. Gleichzeitig überwacht das Programm auch die Abläufe, so dass Verantwortlichen an BSZ und SGS nach Abschluss des täglichen Jobs die nötigen Rückmeldungen erhalten.

IMDAS-Pro erlaubt den Export von Daten im XML-Format. Sie beinhalten je nach Konfiguration sowohl Daten aus dem IMDAS-Pro-Medienobjekt als auch dem IMDAS-Pro-Museumsobjekt. Für die Langzeitarchivierung notwendig sind Metadaten in austauschfähigen, nicht proprietären Formaten – dafür bietet sich *museumdat*³³ an. Als Arbeitsergebnis der Fachgruppe Dokumentation des Deutschen Museumsbundes stellt es das Standardformat dar, mit dessen Hilfe Beschreibungsdaten aus Museumsbeständen ausgetauscht und gegenseitig nutzbar gemacht werden. Die aus IMDAS-Pro exportierten Daten werden in *museumdat*-Darstellung mit den Mastern und Submastern zu METS-Paketen verbunden und als Submission Information Packages an das Langzeitarchiv übertragen. In diese Pakete werden auch die notwendigen technischen Metadaten der Langzeitarchivierung eingebunden. Die automatische Erhebung dieser Daten geschieht mit der Open-Source-Software JHOVE, die am BSZ seit langem eingeführt ist und z.B. auch in kopal Anwendung findet. Im Archiv stehen diese Pakete den Prozessen der Langzeitarchivierung im engeren Sinn zur Verfügung.

Der Projektstand

Die erste Arbeitssitzung im Projekt lag im April 2008. Heute, im Februar 2009, sind große Teile des Projektes einsatzbereit: die Vorgaben für Dateinamen und IPTC-Daten sind definiert; auf ihrer Basis wurden die Arbeitsrichtlinien in Dokumentation und Fotoatelier der Staatsgalerie erarbeitet, die notwendigen Schulungen fanden statt. Die Verarbeitungsprogramme im BSZ sind fertig gestellt und getestet, die Import-Routinen für IMDAS-Pro stehen bereit. Die notwendigen Anpassungsarbeiten an IMDAS-Pro der Version 4.0 stehen vor dem Abschluss, so dass die Projektpartner damit rechnen, Mitte März die Produktion zu starten. Die Staatsgalerie Stuttgart hat einen großen Teil der seit 2005 hergestellten Fotografien mit mobiler Festplatte an das BSZ übertragen,

32 <http://www.imagemagick.org/script/index.php>

33 XML-Schema und Dokumentation finden sich unter der URL <http://museum.zib.de/museumdat/>

die dort gesichert werden. Mit ihnen wurde ein Massentest der Verarbeitungsprogramme durchgeführt, der zeigte, dass die Programme zuverlässig arbeiten. Gleichzeitig steht mit dem Festplattentransport ein zweiter Lieferweg neben dem sFTP-Transfer zur Verfügung. In Vorbereitung ist die Herstellung und Ablage der Submission Information Packages. Mit Produktionsbeginn wird ein internes Audit zur Einhaltung der BSI-Standards³⁴ und Schutzbedarfsfeststellung durchgeführt.

Die Perspektiven und Zusammenhänge

Nach gemeinsamer Einschätzung der Beteiligten hat das Projekt ein beträchtliches Potential im Hinblick auf Materialien, Anwendung und Ausbreitung. Das im Entstehen befindliche Verfahren entlastet die Staatsgalerie Stuttgart von der laufenden Server- und Softwarewartung für den digitalen Bildbestand und teilt die Verantwortlichkeit für die Langzeitarchivierung der Daten zwischen der SGS als Eigentümer der Daten und dem BSZ als Dienstleister. Neben positiven internen Organisationseffekten werden finanzielle Einsparungen erwartet. Das damit einhergehende fast vollständige Outsourcing des Datenbestandes ist für den Eigentümer der Originale und Originaldaten, die Staatsgalerie Stuttgart, nicht selbstverständlich, aber angesichts der Ergebnisse gewünscht, die Abläufe und Bedingungen zu erleichtern.

Gleichzeitig sind neben den hochauflösenden Fotografien eine Vielzahl weiterer, teilweise bereits historischer Bildbestände aus der Geschichte der Staatsgalerie, aus ihren Werkstätten, zu Veranstaltungen und Ausstellungen vorhanden, die für die Forschung immer relevanter werden. Auch wenn es sich vor allem um Aufnahmen im JPG-Format handelt, steht ihre langfristige Sicherung und Dokumentation an. Schon zur Vermeidung der Doppelerfassung bzw. paralleler Dokumentationswerkzeuge sind sie in die Langzeitarchivierung einzubeziehen.

Daneben entstehen auch Künstlervideos und weitere elektronische Inhalte aus der Museumsdokumentation, aus Ausstellungsunterlagen, Öffentlichkeitsarbeit in Form von Texten, Bildern und Tonaufnahmen, die besondere Anforderungen an die Speichertechnik stellen. Ihre Bearbeitung steht in künftigen Projekten an.

Die in Deutschland maßgeblichen Standards der Langzeitarchivierung werden konsequent beachtet, proprietäre Lösungen vermieden. Die Standardisierung im Museumsbereich ist längst nicht so weit fortgeschritten wie z.B. im

34 http://www.bsi.de/literat/bsi_standard/index.htm

Bibliotheksbereich. Mit museumdat steht jetzt ein Format zur Verfügung, das den Verbindlichkeitsanforderungen an inhaltlich beschreibende Metadaten genügt und sich hoffentlich durchsetzen wird. Seine Benutzung bei der Bildung der zur Langzeitarchivierung üblichen METS-Objekte ist die zukunftsweisende Lösung. Solche Objekte werden geeignet sein, z.B. in eine kopial-Installation übertragen zu werden.

Speicherplatz wird am BSZ bedarfsgerecht bereit gestellt und kann laufend erweitert werden. Damit verbunden sind die üblichen Sicherungsverfahren in einem Rechenzentrum inklusive dislozierter, redundanter Speicherung an einem zweiten Aufbewahrungsort.

Das Ministerium für Wissenschaft, Forschung und Kunst Baden-Württemberg unterstützt das Projekt besonders auch aus einem weiteren Grund: nach Einführung des Verfahrens sollen die Projektergebnisse den anderen Museen des Landes Baden-Württemberg und des MusIS-Verbundes zur Verfügung stehen. Die Begleitung des Projekts durch das Badische Landesmuseum Karlsruhe (BLM) sichert genau diesen Sachverhalt: hausspezifische Lösungen werden vermieden, dafür werden Methoden und Wege gewählt, die sich mit geringem Aufwand verallgemeinern, übertragen und in weiteren Institutionen einführen lassen. Vereinbart ist schon heute, dass nach Inbetriebnahme der Produktion für die Staatsgalerie Stuttgart das BLM seinen Bestand an Fotografien im erarbeiteten Verfahren an das BSZ übertragen wird.

Es bestätigt sich, was in einem Gespräch zur Langzeitarchivierung in Baden-Württemberg festgehalten wurde, an dem Vertreter von Rechenzentren, Bibliotheken, Archiven und Museen beteiligt waren: der Bedarf an Lösungen für die Langzeitarchivierung in Museen besteht. Auch wenn die Digitalisierung der Bestände in den Museen später als z.B. in den Bibliotheken einsetzte und anderen Notwendigkeiten folgt, hat doch die Produktion von Daten, die verlässlich archiviert werden müssen, begonnen und wird einen beträchtlichen Aufschwung nehmen. Die Zusammenarbeit mit einem Dienstleister gewährleistet die gewünschte hohe Datensicherheit.

Literatur

Allgaier, Elke (Staatsgalerie Stuttgart): *Archivierung von digitalen Bilddaten*. Beitrag des nestor-Seminars „Digitale Langzeitarchivierung in Museen und Archiven - Konzepte und Strategien“ Köln, 21.11.2008. In: URL: http://www.langzeitarchivierung.de/downloads/2008-11-21_allgaier.pdf

Wolf, Stefan / Mainberger, Christof / Schweibenz, Werner:
Langzeitarchivierung am Bibliotheksservice-Zentrum Baden-Württemberg : Konzept, Aktivitäten und Perspektiven. – Preprint – Konstanz : BSZ, 2009. URL: <http://opus.bsz-bw.de/swop/volltexte/2009/465/>