DFG-Praxisregeln

„Digitalisierung“

Deutsche Forschungsgemeinschaft
Kennedyallee 40 · 53175 Bonn · Postanschrift: 53170 Bonn
Telefon: + 49 228 885-1 · Telefax: + 49 228 885-2777 · postmaster@dfg.de · www.dfg.de

Inhaltsverzeichnis

Vorbemerkung zu den Praxisregeln „Digitalisierung“ 4

1. Prüfliste für Antragstellende und für die Begutachtung 4

1.1 Allgemeine technische Abläufe/Voraussetzungen 5

1.2 Technische Parameter der digitalen Reproduktion 6

1.3 Metadaten 6

1.4 Volltextgenerierung 8

1.5 Langzeitverfügbarkeit 8

1.6 Organisatorische Fragen 9

1.7 Zitieren, persistente Adressierung 9

1.8 Bereitstellung der Metadaten und Digitalisate für die Öffentlichkeit 9

2. Ziele und Auswahl 11

2.1 Ziele 11

2.2 Auswahl 12

2.3 Dublettenprüfung und Datenabgleich bei der Retrodigitalisierung publizierter Texte 13

3. Digitalisierung 13

3.1 Bereitstellung der Materialien, konservatorische Prüfung 14

3.2 Technische Parameter der digitalen Reproduktion 14

3.2.1 Allgemeine Erläuterungen und Parameter 14

3.2.1.1 Auflösung 15

3.2.1.2 Farbtiefe 16

3.2.1.3 Digitaler Aufnahmeablauf 17

3.2.1.4 Dateiformate 20

3.2.2 Materialspezifische Parameter 22

3.2.2.1 Textwerke 22

3.2.2.2 Grafische Darstellungen 23

3.2.2.3 Fotografien 23

3.2.2.4 Mikroformen 26

3.2.2.5 Dreidimensionale Objekte 26

3.3 Metadaten 30

3.3.1 Erschließung, deskriptive Metadaten 31

3.3.2 Strukturelle Metadaten für digitale Faksimiles 32

3.3.3 Sammlungs- und Bestandsbeschreibung 33

3.3.4 Austausch und Weitergabe der Metadaten 33

3.4 Volltextgenerierung 34

3.4.1 Texterfassung 34

3.4.2 Zeichenkodierung 37

3.4.3 Markup von Volltexten 38

3.4.4 Layout 38

3.5 Langzeitverfügbarkeit 39

4. Organisatorische Fragen – Eigendigitalisierung versus Digitalisierung in Dienstleistung 40

5. Zitieren digitaler Ressourcen, persistente Adressierung 41

6. Bereitstellung der Projektergebnisse für die Öffentlichkeit 42

6.1 Rechte, Lizenzierung und Open Access 42

6.2 Mindestanforderungen für die Bereitstellungssysteme von Digitalisaten 44
6.2.1 Funktionalitätsanforderungen 44
6.2.2 Technische Mindestanforderungen 45

7. Präsentationsstandards für Textwerke, DFG-Viewer 45

7.1 Basisanforderungen und Architektur 46
7.2 Funktionalitätsanforderungen 46
7.3 DFG-Viewer 47

Anhang A: METS/MODS-Profil für die Darstellung im DFG-Viewer und Übermittlung per OAI

Anhang B: METS/TEI-Spezifikation für die Darstellung von digitalisierten Handschriften

Anhang C: LIDO-Kernelemente für die Publikation

Anhang D: Sammlungs- und Bestandsbeschreibung

Vorbemerkung zu den Praxisregeln „Digitalisierung“

Die Deutsche Forschungsgemeinschaft (DFG)[1] fördert im Bereich der Wissenschaftlichen Literaturversorgungs- und Informationssysteme (LIS) Projekte an wissenschaftlichen Einrichtungen, insbesondere Service- und Informationseinrichtungen in Deutschland. Förderziel ist der Aufbau leistungsfähiger Informationssysteme und -infrastrukturen für die Forschung unter überregionalen Gesichtspunkten. Die Ergebnisse der Projekte sollen für die Wissenschaft frei und dauerhaft zugänglich sein.

Ziel der Praxisregeln „Digitalisierung“ im Förderbereich LIS ist es, Antragstellenden die Planung von Digitalisierungsprojekten zu erleichtern und die Begutachtung von Anträgen vergleichbar zu gestalten. Die Praxisregeln wollen keine Hürden aufbauen, sondern durch die Formulierung von Standards einen Beitrag zur Nachhaltigkeit und Zukunftsfähigkeit der unterstützen Projekte leisten.

Die Standards ergänzend finden sich in den Praxisregeln auch weiterführende Ausführungen, beispielsweise zur konservatorischen Prüfung der zur Digitalisierung vorgesehenen Materialien, zum Erheben von Metadaten, zur Herstellung der Digitalisate, zur Indexierung von Bildinhalten, zur Herstellung von Volltexten oder auch zur Perspektive der Langzeitsicherung digitaler Inhalte.

Im Folgenden werden in Kapitel 1 knapp die wichtigsten Anforderungen zusammengefasst. Abweichungen von den dort beschriebenen Regeln sind zwar möglich, müssen aber im Einzelfall begründet werden. In den Kapiteln 2 bis 6 findet sich eine grundsätzliche und weiter ausholende Einführung in die Probleme und Vorgehensweisen, die sich für Projekte stellen, die Objekte aus Bibliotheken, Archiven oder musealen Sammlungen in digitale Form überführen wollen. Diese Kapitel wenden sich bewusst auch an Personen, die derartige Projekte neu planen, also unter Umständen über keine detaillierten Vorkenntnisse verfügen. Kapitel 7 verweist auf für textuelle Objekte erforderliche DFG-Präsentationsstandards und Formate.

1. Prüfliste für Antragstellende und für die Begutachtung

Wie in den folgenden Kapiteln ausgeführt wird, bestehen bei der Planung von Digitalisierungsprojekten zahlreiche Wahlmöglichkeiten. Eine rein mechanische Verpflichtung auf ganz bestimmte Standards würde daher eine unzulässige Einschränkung der zu fördernden Projekte bedeuten und deren dynamische Weiterentwicklung beschränken. Andererseits sollte klar werden, in welchen Bereichen Risiken für die Projektplanung und den Projektverlauf liegen. Die Prüfkriterien sind daher folgendermaßen zu verstehen und anzuwenden:

[1] Die Deutsche Forschungsgemeinschaft (DFG) ist die zentrale Selbstverwaltungseinrichtung der Wissenschaft zur Förderung der Forschung an Hochschulen und öffentlich finanzierten Forschungsinstitutionen in Deutschland. Die DFG dient der Wissenschaft in allen ihren Zweigen durch die finanzielle Unterstützung von Forschungsvorhaben und durch die Förderung der Zusammenarbeit unter den Forschern (http://www.dfg.de). Die DFG unterstützt auch Vorhaben zur Verbesserung der wissenschaftlichen Informations-Infrastrukturen in Deutschland. Die Ergebnisse der geförderten Projekte sollen für die Wissenschaft frei und dauerhaft zugänglich sein (http://www.dfg.de/lis). Zu beachten ist, dass die definierten Trägeraufgaben und -finanzierungen der antragstellenden Einrichtungen durch die Förderung nicht substituiert werden dürfen. Projekte müssen daher in ihrer Profilierung über die regulären Grundaufgaben einer Einrichtung hinausgehen, zeitlich und inhaltlich begrenzt sein sowie herausragende und überregional bedeutende Materialien zum Gegenstand haben. Nicht förderfähig sind damit Vorhaben, die vorrangig der Kulturförderung, Kulturgutvermittlung oder vergleichbaren Zielsetzungen dienen, sowie kommerziell orientierte Projekte.

- Bei allen Anträgen mit Digitalisierungskomponenten werden unabhängig von ihrer inhaltlichen Begutachtung die vorgeschlagenen technischen Konzepte geprüft.
- Darüber hinaus wird geprüft, ob die notwendigen Rechteklärungen bezogen auf das Ausgangsmaterial ebenso erfolgt und nachgewiesen sind, wie die Vorkehrungen, die eine vollumfängliche Nachnutzbarkeit der Projektergebnisse erlauben. (→ 1.8)
- Antragstellende müssen in ihren Anträgen nachvollziehbar belegen, dass die Durchführung des Projekts die im Folgenden aufgelisteten Standards einhält.
- Jede Abweichung von diesen Standards muss ausführlich begründet werden.
- Zur besseren Vergleichbarkeit sollen Anträge (bzw. Arbeitsberichte) Auskünfte über die geschätzten (bzw. tatsächlichen) Kosten für das Scannen inklusive Strukturdatenerfassung pro Image und Titel / Verzeichnungseinheit, Kosten für die Volltexterfassung sowie Kosten für projektspezifische Speicherkapazitäten enthalten.[2]
- Sieht ein Digitalisierungsprojekt vor, die im Folgenden angeführten Standards zu überschreiten, so ist die Notwendigkeit im Einzelnen zu begründen, soweit dadurch erhöhte Kosten entstehen.
- Grundsätzlich müssen Anträge technisch so weit vorbereitet sein, dass eine Gesamtbeurteilung der technischen Anforderungen und Abläufe aus dem Antrag heraus möglich ist. Es besteht also – abgesehen von der Erprobung innovativer Technik – nicht die Möglichkeit, eine erste Projektphase vorzusehen, in der beispielsweise der zeitliche Aufwand für eine Digitalisierungskampagne, die zu wählende Qualität der Digitalisate oder der allgemeine Arbeitsverlauf (Workflow) erst bestimmt wird. Dazu nötige Pilotstudien sind vor der Antragstellung abzuschließen.

1.1 Allgemeine technische Abläufe/Voraussetzungen

Im Antrag ist der geplante Workflow so detailliert zu beschreiben, dass eine Bewertung der folgenden Fragen durch die Gutachterinnen und Gutachter möglich wird:

---

[2] Vgl. dazu das Statistikblatt „Digitalisierung“ unter http://www.dfg.de/lis, das in der Rubrik „Formulare und Merkblätter“ unter „Ergänzende Formulare und Formblätter LIS“ zugänglich ist. Zur Berechnung der realen Kosten gibt es zurzeit noch keine allgemein akzeptierten Standards. Die Praxisregeln empfehlen daher, die Kosten des Scanvorgangs inklusive Strukturdatenerfassung folgendermaßen zu errechnen:

(1) Bei der Vergabe an einen Dienstleister.
  (a) Aus den an den Dienstleister gezahlten Kosten pro Digitalisat.
  (b) Aus den anteiligen Kosten an alle vom Dienstleister verlangten Pauschalen (z. B.: Benennung und Speicherung der Dateien; Strukturdatenvergabe, Transfer der Daten vom Dienstleister zum Projektstandort).
  (c) Aus den anteiligen Personalkosten aller Personen, die ausschließlich oder mit einem kalkulierbaren Anteil die Qualitätskontrolle durchgeführt haben.
(2) Bei der Digitalisierung im Hause.
  (a) Aus den anteiligen Kosten an der neu angeschafften Digitalisierungshardware im engeren Sinne. Diese gilt mit Projektende als abgeschrieben. (Ganze Laufzeit: im Falle eines vier Jahre laufenden Projekts, für das zunächst zwei Jahre beantragt werden, ist also die für die volle vierjährige Laufzeit veranschlagte Zahl von Digitalisaten zu Grunde zu legen.)
  (b) Aus den anteiligen Personalkosten aller Personen, die die Digitalisierungshardware bedienen und Strukturdaten vergeben.
  (c) Aus den anteiligen Personalkosten aller Personen, die ausschließlich oder mit einem kalkulierbaren Anteil die Qualitätskontrolle durchgeführt haben.

In beiden Fällen (1) und (2) sind die Kosten pro Digitalisat zu errechnen. Keine Kosten des Scannens sind alle Kosten, die entstehen:

(a) Durch das Projektmanagement (z. B.: Auswählen und Ausheben der zu digitalisierenden Stücke).
(b) Durch die Erfassung von über die Strukturdaten hinausgehenden Metadaten.
(c) Durch die Langzeitarchivierung.
(d) Indirekte Kosten, wie sie in der Regel im Rahmen betriebswirtschaftlicher Kostenrechnungen erhoben werden.

- Ist die vorgesehene Personalausstattung hinreichend und notwendig? Dazu sind Aussagen über die durchschnittlichen Ressourcen (Arbeitszeit, Speicherkapazität der am Workflow beteiligten Rechner) für jede der dem Mengengerüst zu Grunde liegenden Einheit zu machen.
- Sind die vorgesehenen Bearbeitungszeiten realistisch? Soweit die vorgesehenen Bearbeitungszeiten nicht unmittelbar plausibel sind, sind sie entweder durch die Erfahrungen in vergleichbaren Projekten, durch veröffentlichte und anerkannte Richtwerte oder durch die Ergebnisse selbst vorgenommener Vorabtests zu belegen.

1.2 Technische Parameter der digitalen Reproduktion

Ziel der Digitalisierung ist die möglichst originalgetreue Wiedergabe des Objekts nach Maßgabe der wissenschaftlichen Erfordernisse. Für alle im Rahmen der Projekte zu digitalisierenden Objekte sind die beantragten Qualitäten so zu wählen, dass Darstellungen, die unmittelbar öffentlich zugänglich gemacht werden, ohne menschliche Intervention im Rahmen des Projekts durch Batchprozesse realisiert werden können. Die folgenden Richtlinien für die Mindestanforderungen an die Digitalisierung beziehen sich ausschließlich auf den **digitalen Master** und stellen materialübergreifende Empfehlungen dar. Zu materialspezifischen Empfehlungen, die über diese Mindestanforderungen hinausgehen, vergleiche Kapitel 3.2.2.

- Die Auflösung ist so zu wählen, dass in der Auflösung der Archivkopien die kleinsten relevanten Details klar zu sehen sind, wenn die Datei auf ein Viertel der Ausgangsgröße verkleinert wird. Grundsätzlich wird eine Auflösung von 300 dpi empfohlen. (→ 3.2.1.1)
- Für die Sicherung des digitalen Masters ist eine Farbtiefe von 8 Bit pro Kanal, d. h. 24 Bit ausreichend. (→ 3.2.1.2)
- Die Speicherung der Master erfolgt in Form unkomprimierter Baseline-TIFF-Dateien. Die weitergehenden Optionen von *extended TIFFs* sollten für den digitalen Master nicht genutzt werden. Neben TIFF kann auch TIFF-LZW oder JPEG2000 in seiner verlustfreien Form als Format für den Bildmaster verwendet werden. Für die Speicherung von Mastern im JPEG2000-Format ist allerdings darauf zu achten, dass nur die lizenzfreien Bereiche von JPEG2000 Verwendung finden. Zu den mit JPEG2000 und TIFF-LZW einhergehenden Risiken bezüglich der Langzeitarchivierung vergleiche Kapitel 3.2.1.4.

1.3 Metadaten

Metadaten sind grundsätzlich in einer von der Software unabhängigen und standardkonformen Form bereitzustellen, in aller Regel in einer XML-Codierung.

- Für eine optimale verteilte, auch nachhaltige Nutzbarkeit der deskriptiven Metadaten ist die Erschließung an den einschlägigen Spartenstandards und Referenzmodellen zu orientieren und wo immer möglich mit publizierten Normdaten zu verknüpfen. Für die Erfassung personenbezogener, biografischer und geografischer Information wird die Verwendung der Gemeinsamen Normdatei (GND) der Deutschen Nationalbibliothek erwartet. Darüber hinaus eingesetzte kontrollierte Vokabulare, z. B. Iconclass zur Bilderschließung, müssen national und international anschlussfähig sein. (→ 3.3.1)
- Die Bereitstellung der deskriptiven Metadaten zur weiteren Nutzung gemäß den materialspezifischen Standards ist verpflichtend (→3.3.1 sowie Anhänge A und B):

  - Metadata Encoding and Transmission Standard / Metadata Object Description Schema (METS[3]/MODS) für gedruckte Textwerke und Archivgut
  - Metadata Encoding and Transmission Standard / Text Encoding Initiative (METS/TEI[4]) für Handschriften
  - Lightwight Information Describing Objects (LIDO) für (i.d.R. unikale) bildhafte und dreidimensionale Objekte

  Die bereitgestellten Metadaten müssen gegen das jeweilige XML-Schema valide sein und sind darüber hinaus auf semantische Korrektheit zu überprüfen.
- Die Bereitstellung der deskriptiven Metadaten über eine Open Archives Initiative (OAI[5])-Schnittstelle ist verpflichtend, wahlweise im eigenen System oder über ein geeignetes Zielportal. (→ 3.3.1)
- Die Digitalisate und Metadaten müssen durch persistente Identifier im Nachweissystem (Katalog, Online-Findmittelsystem) verbunden werden. Von antragstellenden Wissenschaftlerinnen und Wissenschaftlern wird erwartet, dass sie die Frage der Metadatenerfassung und Digitalisierung mit einer wissenschaftlichen Infrastruktureinrichtung abstimmen bzw. durch diese durchführen lassen. Der Nachweis digitalisierter Drucke im Zentralen Verzeichnis digitalisierter Drucke (ZVDD)[6] und von digitalisiertem Archivgut im Archivportal-D[7] wird erwartet. Alle Materialien müssen in geeigneten fachlichen oder fachübergreifenden Online-Anwendungen präsentiert und in die Deutsche Digitale Bibliothek (DDB)[8] und mittelbar in die Europeana[9] eingebracht werden. (→ 3.3)
- Es obliegt dem Projektnehmer sicherzustellen, dass die im Projekt hergestellten Digitalisierungseinheiten eindeutig identifiziert und von anderen im selben System vorhandenen Einheiten getrennt durchsucht und aufgerufen werden können. (→ 3.3)
- Die Ablieferung an die genannten Portale sollte gemäß den Standardformaten möglichst über OAI erfolgen. (→ 3.3.1)
- Die Entscheidung über die Erstellung struktureller Metadaten ist immer eine material- und projektspezifische. Falls strukturelle Metadaten Verwendung finden, wird die Konsultation der über die Website des DFG-Viewers[10] zugänglichen Liste von Bezeichnungen erwartet. Empfohlen wird nach dem derzeitigen Stand eine Orientierung an METS oder TEI, wobei der DFG-Viewer bei textuellen Materialien auf jeden Fall unterstützt werden muss. (→ 3.3.2)
- Es wird erwartet, dass Digitalisierungsprojekte mindestens Gegenstand und Umfang der jeweiligen Objektauswahl auf einer Seite im Netz, möglichst auch in Englisch, darstellen. Erwartet wird eine normierte Sammlungs- bzw. Bestandsbeschreibung in XML in einem standardisierten Format, um diese Informationen in Zukunft besser in nationalen oder internationalen Portalen zusammenführen und gezielt recherchieren zu können. Diese Beschreibung kann gemäß dem Dublin Core Collections Application Profile[11] oder in demselben Metadatenstandard erfolgen, in dem auch die Objektbeschrei-

[3] http://www.loc.gov/standards/mets
[4] http://www.tei-c.org
[5] http://www.openarchives.org/
[6] http://www.zvdd.de/
[7] https://www.archivportal-d.de/
[8] https://www.deutsche-digitale-bibliothek.de/
[9] http://www.europeana.eu/portal/de
[10] http://dfg-viewer.de

bungen verfügbar gemacht werden: METS, MODS oder TEI-Header, Encoded Archival Description (EAD(DDB)) sowie LIDO bieten entsprechende Möglichkeiten. (→ 3.3.3 sowie Anhang D)

1.4 Volltextgenerierung

Von allen Anträgen zur Digitalisierung textueller Objekte wird eine Auseinandersetzung mit Möglichkeiten der Volltextbereitstellung erwartet. Für Druckwerke ab Erscheinungsjahr 1850 gilt verpflichtend, dass Volltext hergestellt werden muss und eine bloße Bilddigitalisierung nicht ausreicht. (→ 3.4)

Volltexte umfassen die Zeichen der Vorlage, gegebenenfalls in den Text eingetragene Auszeichnungen bzw. Markup-Daten zur Markierung von Strukturmerkmalen sowie Metadaten, die üblicherweise Teil der gleichen Datei sind. (→ 3.4)

- Die Herstellung von Volltext kann auf zwei Wegen erfolgen, entweder durch Optical Character Recognition (OCR) oder durch Abschreiben bzw. Transkription. (→ 3.4.1)
- Um eine gewisse Einheitlichkeit bei der Beurteilung der Textgenauigkeit zugrunde legen zu können, werden Antragstellende gebeten, diesbezügliche Angaben zur Buchstabengenauigkeit zu machen, d. h. mangelhafte Worttrennungen und Layoutfehler unberücksichtigt zu lassen. Stichproben sollten sich an dem in Kapitel 3.4.1 beschriebenen statistischen Verfahren orientieren. (→ 3.4.1)
- Bezüglich der Zeichencodierung wird empfohlen, die Texte in Unicode abzuspeichern. Zu favorisieren ist UTF-8. (→ 3.4.2)
- Wenn nicht triftige Gründe dagegen sprechen, müssen Volltexte von Drucken und Handschriften nach dem Modell der Text Encoding Initiative (TEI) kodiert bzw. mit Markup versehen werden. (→ 3.4.3)
- In manchen Fällen ist es bei der Präsentation eines Volltextes wichtig, das Layout eines Dokumentes langfristig zu sichern. Die Praxisregeln empfehlen bevorzugt für diese Fälle den Einsatz einer geeigneten Formatierungssprache (z. B. XSLT, XSL-FO, XQuery, CSS), die die Unabhängigkeit von spezieller Software weitgehend sicherstellt. Falls eine Archivierung des Formats mit XML-Techniken aus nachvollziehbaren Gründen nicht möglich ist, können Layoutinformationen zu Textdokumenten auch in PDF nach der ISO-Norm 19005-1 archiviert werden. PDF-Dateien sind aber, wie in Kapitel 3.4.2. dargelegt, kein Ersatz für eine Bereitstellung der mit Markup versehenen Volltexte in XML. (→ 3.4.4)
- Im Sinne des *open access* und der *open source* wird erwartet, dass maschinenlesbare Volltexte bzw. das den Volltexten zu Grunde liegende XML sowie XSLT-Skripte und DTDs oder XML-Schema-Dateien so frei nachnutzbar wie rechtlich möglich angeboten werden. (→ 1.8; 6.1)

1.5 Langzeitverfügbarkeit

Für die Langzeitsicherung und Archivierung digitaler Inhalte gibt es zum jetzigen Zeitpunkt keine verallgemeinerbare Lösung, die für alle Objektgattungen geeignet ist. Im Rahmen der Langzeitsicherung werden Dateien in stabilen, migrationsfähigen Formaten auf einem tech-

nisch wie organisatorisch sicheren Speichersystem gesichert. Die Archivierung digitaler Daten setzt auf einem solchen Speicher auf, beinhaltet jedoch noch weitreichendere technische und organisatorische Festlegungen, die nicht nur eine physische Erhaltung der Daten, sondern Strategien zur Bereitstellung für die Nutzung (Access) einbeziehen. (→ 3.5)

Die Langfristverfügbarkeit der Ergebnisse von Digitalisierungsprojekten ergibt sich einerseits aus der Wahl der Daten- und Metadatenformate. Andererseits ist sicherzustellen, dass die digitalen Daten auch physikalisch verfügbar bleiben. Dabei gilt: Kosten für die projektbezogene Sicherung der Daten werden in DFG-geförderten Digitalisierungsprojekten als Eigenleistung für die Laufzeit des Projekts anerkannt. Eine Förderung dieser Kosten aus DFG-Mitteln kann nicht erfolgen. (→ 3.5)

Es sei darauf hingewiesen, dass Digitalisierungsprojekte aus Sicht der DFG stets Projekte der gesamten Einrichtung sind: Die Unterstützung der das Projekt abwickelnden Fachabteilung durch die IT-Infrastruktur des Hauses wird vorausgesetzt. Dabei wird begrüßt, wenn sich kleinere Einrichtungen der Kompetenz und der Dienstleistung größerer Einrichtungen bedienen. (→ 3.5)

Anträge müssen nachvollziehbare Aussagen zur institutionellen Langzeitsicherung und Archivierung enthalten. Digitalisate sind in redundanter Form zu archivieren. (→ 3.5)

1.6 Organisatorische Fragen

Die Entscheidung darüber, ob ein Vorhaben als *Inhouse*-Projekt konzipiert oder ob auf das Angebot eines Dienstleisters zurückgegriffen wird, ist immer eine projektspezifische Entscheidung, die allein in die Verantwortung der Antragstellerinnen und Antragsteller fällt. Bei der Zusammenarbeit mit Dienstleistern muss die projektnehmende Einrichtung deutlich machen, dass sie ein Projekt kompetent begleiten kann. In Verträgen müssen Leistungen exakt festgelegt werden. Die DFG erwartet, dass ein angemessener Prozentsatz des Rechnungsbetrags als Sicherheit zurückbehalten wird und erst nach einer durchgeführten Qualitätskontrolle zur Auszahlung an das Dienstleistungsunternehmen kommt. (→ 4)

1.7 Zitieren, persistente Adressierung

Digitalisate müssen, um von anderen Objekten oder Datenbanken aus erreichbar zu sein, eindeutig angesprochen werden können. Dazu ist über die übliche Zitierform hinaus, die durch ein Angebot in der Navigationssoftware als klassische Form weitergenutzt werden kann und sollte, die Festlegung und die online zugängliche Dokumentation von Adressierungstechniken erforderlich. Sichergestellt werden muss die Erreichbarkeit und Zitierbarkeit einer Ressource als Ganzer und die Erreichbarkeit und Zitierbarkeit von einzelnen physischen Seiten von diesem Werk. Einrichtungen sollten durch geeignete Mechanismen (Persistent Uniform Resource Locator (PURL), Uniform Resource Name (URN), Digital Object Identifier (DOI), Handle, etc.) die Persistenz der Ressource und der Verknüpfung zu ihr gewährleisten, um zuverlässiges Arbeiten mit den bereitgestellten Quellen in wissenschaftlichen Kontexten zu ermöglichen. Für gedruckte Werke wird die Erzeugung von URNs über die Deutsche Nationalbibliothek empfohlen. (→ 5)

1.8 Bereitstellung der Metadaten und Digitalisate für die Öffentlichkeit

Die Digitalisierung wissenschaftlicher Materialien wird durch die DFG gefördert, um diese Materialien der Forschung in Deutschland und weltweit zugänglich zu machen. Alle Projekte sind

so anzulegen, dass die Ergebnisse der Forschung frühzeitig und nachhaltig zugänglich werden und eine vollumfängliche wissenschaftliche Nachnutzung auch in anderen Forschungskontexten möglich ist.

- Bereits mit Beginn der Projektplanung, spätestens aber zum Zeitpunkt der Antragstellung, muss eine Rechteklärung bezüglich der zu digitalisierenden Objekte erfolgt sein. Insbesondere eventuell vorliegende Urheber-, Persönlichkeits-, Leistungsschutz- und Archivrechte müssen dabei berücksichtigt werden. (→ 6.1)
- Im Sinne des *open access* und der *open source* wird vorausgesetzt, dass alle Ergebnisse – Metadaten, Digitalisate und Volltexte bzw. das den Volltexten zu Grunde liegende XML sowie XSLT-Skripte und DTDs oder XML-Schema-Dateien – so frei nachnutzbar wie rechtlich möglich angeboten werden. Wenn sie bereits Teil der Public Domain sind, sind sie als solche zu markieren und nicht mit Lizenzen zu belegen. Ist dies nicht möglich, sollen sie mit der freiest möglichen Creative Commons Lizenz (CC0, CC BY oder CC BA SA)[12] angeboten und bei geschütztem Material mit standardisierten Rechtehinweisen versehen werden. (→ 6.1).
- Darüber hinaus sollten Images in einer Form bereitgestellt werden, die eine vollumfängliche wissenschaftliche Nachnutzung in anderen Forschungskontexten erlaubt. Dazu sollen entweder hochaufgelöste Derivate im TIFF-Format oder vollaufgelöste JPEG-Images mit einer Kompression von 90, maximal 80 in Kombination mit der vollen Scanauflösung zur Verfügung gestellt werden. (→ 6.1)
- Bei Projekten, in denen nicht nur gemeinfreie Materialien digitalisiert werden und / oder mit kommerziellen Partnern oder Verlagen zusammengearbeitet wird, kann eine verzögerte open access-Publikation (*moving wall*) von bis zu einem Jahr nach Projektende vereinbart werden. (→ 6.1)
- Eine Einschränkung des *open access* kann z. B. datenschutz- urheber- oder archivrechtlich begründet sein, darf aber maximal 5 % der insgesamt zu digitalisierenden Materialien betreffen und setzt ein darzustellendes Konzept zur zeitlichen und technischen Verfügbarmachung voraus.[13]

Digitalisate sollen im Regelfall über unterschiedliche Wege zugänglich sein: (→ 6.2.1)

- über die Website der anbietenden Institution
- über den lokalen und regionalen Bibliothekskatalog / das lokale und regionale Archivportal / die jeweilige materialspezifische Online-Anwendung und die überregionalen Nachweis- und Präsentationssysteme (DDB, Archivportal-D)
- über den lokal implementierten oder an einem anderen Ort betriebenen „DFG-Viewer“, sofern für das angebotene Material zutreffend
- über Internet-Suchmaschinen
- über eine OAI-Schnittstelle

Alle Objekte sind in einer für wissenschaftliche Zwecke hinreichend guten Qualität anzubieten und mit intuitiv zu bedienenden Navigationsmöglichkeiten zu versehen, so dass die Benutzung für die fragliche Zielgruppe und auf an den Hochschulen üblicherweise vorhandenem

[12] Vgl. Lizenzen in der zurzeit gültigen Version: https://creativecommons.org/
[13] Eine Förderung der Bearbeitung nicht im o*pen access* zugänglichen Materials mit DFG-Mitteln ist nicht möglich.

Equipment mühelos möglich ist. Alle jeweils gängigen Browser müssen unterstützt werden, soweit dies der Natur der Sache nach nicht unmöglich ist. (→ 6.2.2;7.2)

Für sämtliche DFG-geförderte Bereitstellungssysteme wird erwartet, dass Nutzerinnen und Nutzer die automatisierte Möglichkeit zum Feedback bezüglich des digitalen Angebots erhalten. Eine entsprechende Möglichkeit ist auf der Projektseite bzw. im digitalen Bereitstellungssystem an zentraler Stelle einzurichten. (→ 6.2.1)

Die DFG erwartet, dass DFG-geförderte Projekte auf ihren im Internet angebotenen Daten eindeutige Herkunftsnachweise sowie gegebenenfalls den Hinweis auf die Förderung durch die DFG anbringen. Dies erfolgt im Fall von Bilddigitalisaten in der Regel durch das Hinzurechnen einer Nachweisleiste in die veröffentlichte Nutzerkopie (zum Beispiel im Grafikformat JPEG). (→ 6.1)

Im Regelfall – wenn die Natur der Digitalisate einen Teil dieses Leistungsspektrums nicht implizit ausschließt – wird erwartet, dass ein Digitalisierungsprojekt Pläne für folgende öffentlich zugängliche Schnittstellen bereitstellt:

- Einen selbständigen Server, der die digitalisierten Objekte mit zu ihrer Nutzung notwendigen Instrumenten bereitstellt.
- Alle Digitalisate sind so anzubieten, dass langfristig zitierfähige URLs möglichst niedriger Granularität entstehen. Die Form, in der zitiert werden soll, ist deutlich anzuzeigen.
- Eine Schnittstelle im technischen Sinn, die den Zugriff auf alle im Rahmen des Projektes erstellten Metadaten ermöglicht. Dieser kann auch über ein zentrales Portal realisiert werden.
- Geeignete Maßnahmen, die dazu führen, dass die Metadaten von Suchmaschinen gefunden werden.

2. Ziele und Auswahl

2.1 Ziele

Digitalisierung gilt heute als ein zentrales Instrument für die geistes- und kulturwissenschaftliche Forschung und die sich entwickelnden *Digital Humanities*. Die direkte Forschung mit den Quellen wurde und wird durch Digitalisierungsmaßnahmen erheblich erleichtert – bei gleichzeitiger Schonung der kostbaren, bisweilen fragilen Originale. Mit der Digitalisierung von Beständen aus Bibliotheken, Archiven und Museen werden aber nicht nur leicht verfügbare Kopien ins Netz gestellt, sondern es entsteht auch eine Infrastruktur, die das Internet zu einem integralen Forschungsraum für eine zunehmend digital ausgerichtete geistes- und kulturwissenschaftliche Forschung macht. Erst die Verbindung mit anderen online verfügbaren Ressourcen ermöglicht es, das Potential des Netzes vollständig auszunutzen. Ziel der Digitalisierung ist also nicht nur das Bereitstellen und Nutzbarmachen, sondern auch und vor allem das Vernetzen der unterschiedlichen Ressourcen zu einer virtuellen Forschungsinfrastruktur.

Es gibt heute einen breiten Stand abgesicherter Kenntnisse zur Durchführung von Digitalisierungsprojekten, die freilich nicht mechanisch angewendet werden können: Was für die Digitalisierung von Handschriften des Mittelalters unverzichtbare konservatorische Sorgfalt ist, kann

für die Bearbeitung von Massenverwaltungsschriftgut des späten 19. Jahrhunderts unnötig zeitraubend und überteuert sein.

Da sich die Mehrzahl der bisher durchgeführten und in Planung befindlichen DFG-geförderten Projekte nach wie vor mit textuellen Materialien befassen, werden im Folgenden die Techniken und Parameter für diesen Bereich besonders gründlich erörtert. Allerdings ist wahrnehmbar, dass auch bildliche Darstellungen beträchtlich an Bedeutung gewinnen. Daher finden sich auch Ausführungen zu bildhaften Materialien, seien es grafische Darstellungen, Fotografien oder Bilddigitalisate von dreidimensionalen Sammlungsobjekten. Die vorliegende Version der PraxisregeIn integriert dabei neuere Entwicklungen in diesen Bereichen, ohne den Anspruch auf Vollständigkeit zu erheben. In DFG-geförderten Pilotprojekten wurden und werden spezifische Techniken und Parameter u. a. für die OCR-gestützte Volltexterzeugung oder Digitalisierung historischer Zeitungen und unikalen Materials (Archivgut, mittelalterliche Handschriften) ermittelt, die nach Abschluss der Pilotphasen Eingang in die Praxisregeln gefunden haben bzw. finden werden.

2.2 Auswahl

Grundsätzlich ist zu beachten, dass die technischen Aspekte der Digitalisierung gut geplant werden können, während der intellektuelle Aufwand bei der Auswahl geeigneter Objekte schwer kalkulierbar ist. Es wird daher immer zu entscheiden sein, ob eine großzügigere Definition des Mengengerüsts für ein Digitalisierungsprojekt insgesamt nicht kostengünstiger und effektiver ist als ein komplexer Bewertungs- und Auswahlprozess. Angeregt wird, auf bereits erbrachte Selektionsleistungen und Vorarbeiten wie Bibliografien oder Fachdatenbanken zurückzugreifen. Als grundsätzliche Kriterien für die Auswahl gelten Forschungsrelevanz und wissenschaftliche Nachfrage.

Die Bildung eines Korpus unter den Gesichtspunkten der Forschungsrelevanz oder der wissenschaftlichen Nachfrage ist nicht immer einfach. Hilfreich für die Plausibilisierung eines Vorhabens ist in diesen Fällen die Kooperation mit einer konkretisierbaren wissenschaftlichen Community oder Institution, die einen nachvollziehbaren Eigenbedarf formulieren kann. Ideal sind Kooperationen, in denen ein wissenschaftliches Vorhaben, z. B. ein philologisches oder rechtshistorisches Forschungs- oder Editionsprojekt, seinerseits ein Internetangebot anstrebt und auf Bibliotheks- oder Archivbestände rekurrieren will, sodass die wechselseitige Vernetzung möglich wird.[14] Alternativ kann man auf einschlägige Fachbibliografien zurückgreifen, die einen Kanon formulieren, oder aber auch eine Mischung aus beidem anstreben. Interessant sind sicher auch Konzepte des *digitization on demand* oder *digitization on use* (Digitalisierung von in Bibliotheken und Archiven genutztem oder stark genutztem Material), wodurch in jedem Fall ein konkretes Nutzungsinteresse sichergestellt ist.

Grundsätzlich sollte bei spezialisierten, eher an einer konkreten Forschungsfrage orientierten Digitalisierungsprojekten gelten, dass sie sich in größere, eher nach formalen Gesichtspunkten vorgehende Projekte integrieren lassen. Umgekehrt sollten fachübergreifende Projekte ihre Angebote für eine vertiefende und spezialisierte Nachnutzung öffnen.

---

[14] In der Regel bauen Forschungsprojekte auf geförderten Infrastrukturen auf. Es ist jedoch ebenfalls möglich, Projekte, die Forschung und Infrastrukturelemente kombinieren, zu beantragen. Alternativ kann man auf einschlägige Fachbibliografien zurückgreifen, die einen Kanon formulieren, oder aber auch eine Mischung aus beidem anstreben.

Die Digitalisierung wissenschaftsrelevanter Materialien wird durch die DFG gefördert, um diese Materialien der Forschung in Deutschland und weltweit zugänglich zu machen. Bereits mit Beginn der Projektplanung, spätestens aber zum Zeitpunkt der Antragstellung, muss die **Rechteklärung bezüglich der zu digitalisierenden Materialien** erfolgt sein und im Antrag nachgewiesen werden. Insbesondere eventuell vorliegende Urheber-, Persönlichkeits- sowie Leistungsschutz- und Archivrechte sind dabei zu berücksichtigen. Die Rechteklärung gestaltet sich häufig aufwendiger als zunächst erwartet. Für die entsprechenden Aufwände ist daher ausreichend Zeit und personeller Einsatz einzuplanen. Stellt sich die Rechteklärung als schwierig dar, wird empfohlen, professionelle Unterstützung, beispielsweise durch das Justitiariat der bestandshaltenden Einrichtung zu suchen. Auch ist schon bei der Auswahl zu beachten, dass die Ergebnisse der durch die DFG geförderten Digitalisierungsprojekte der Forschung im Sinne des *open access* so frei wie möglich zur Verfügung gestellt werden müssen. Wenn eine Markierung als Public Domain rechtlich nicht möglich ist, ist eine freie Creative Commons-Lizenz zu wählen (CC 0, CC BY oder CC BY SA). (→ 6.1).

2.3 Dublettenprüfung und Datenabgleich bei der Retrodigitalisierung publizierter Texte

Zur Vermeidung von Doppeldigitalisierungen ist es sinnvoll, vor Antragstellung zu prüfen, ob die zur Digitalisierung vorgesehenen Objekte in Deutschland oder im Ausland bereits digital vorliegen. Hierfür gelten folgende Anforderungen:

In Anträgen und Berichten werden Aussagen zu bereits vorliegenden oder laufenden nationalen und internationalen Digitalisierungsvorhaben erwartet, soweit sich diese Vorhaben auf das beantragte oder laufende DFG-Projekt und die dort bearbeiteten Bestände beziehen.

Projekte mit größeren Mengengerüsten (z. B. über 1.000 Drucke) sollten in Anträgen und Berichten erläutern, wie sich ihr Vorhaben zu kostenfrei zugänglichen kommerziellen Digitalisierungsangeboten verhält. Ziel sollte es sein, die Zahl der gegebenenfalls auftretenden Doppeldigitalisierungen mit pragmatischem Aufwand möglichst gering zu halten. Bei Drucken der Erscheinungsjahre von 1501 bis 1800 sind das VD16, VD17 und VD18 als Referenzinstrumente für Prüfungen auf Doppeldigitalisierungen heranzuziehen. URN und PURL der Digitalisate sind an diese Verzeichnisse zu melden. Bei der Digitalisierung von Inkunabeln ist der deutsche Inkunabelzensus (ISTC) heranzuziehen.

3. Digitalisierung

Die Digitalisierung umfasst immer die Herstellung digitaler Images und die Erzeugung von Metadaten, im Falle von Textwerken gegebenenfalls zusätzlich auch die Volltexterfassung und die Erzeugung von Strukturdaten und Markup. Wenn im Folgenden von Digitalisierung gesprochen wird, so ist der gesamte Arbeitsgang gemeint (Vorbereitung, Digitalisierung im engeren Sinne, Erzeugung von bibliografischen Metadaten, Strukturdaten, Volltexten sowie Langzeitsicherung/digitale Bestandserhaltung). Auf eine Differenzierung nach Materialarten (u. a. Druckwerke, unikale Dokumente, Objekte) wird in den allgemeinen Kapiteln im Folgenden verzichtet.

3.1 Bereitstellung der Materialien, konservatorische Prüfung

Vorbereitende Tätigkeiten werden bei Digitalisierungsprojekten oft unterschätzt und sollten vor einem Projekt genauestens geprüft werden. Sind die Objekte überhaupt verfügbar? Gibt es möglicherweise konservatorische Bedenken gegenüber der Digitalisierung der Originale? Stehen genügend Mitarbeiterinnen und Mitarbeiter zur Aushebung und Bereitstellung der Objekte zur Verfügung? Ist wissenschaftlich bzw. bibliografisch geschultes Personal zur Hand, das Vollständigkeitskontrollen bzw. Kollationierungen vornimmt, sofern dies aus vorhandenen Katalogeinträgen nicht hervorgeht? Bei Textwerken sollte eine Digitalisierung von unvollständigen oder defekten Drucken nach Möglichkeit vermieden und eine Reproduktion eines vollständigen Exemplars angestrebt werden.

Die konservatorische Prüfung kann sehr viel Zeit in Anspruch nehmen, sollte aber zum Schutz der Objekte keinesfalls unterbleiben. Es wird empfohlen „Checklisten“ für die Prüfung der Digitalisierungsfähigkeit zu verwenden und daraus die Vorgaben für die Digitalisierung zu entwickeln.[15] Stellt die Reproduktion mit den vorhandenen Reproduktionstechniken ein Risiko bzw. eine besondere Belastung für die Originale dar,[16] so sollte sie – sofern möglich – vom bestehenden Mikrofilm[17] erfolgen oder ganz unterbleiben. In jedem Fall sind wertvolle historische Materialien mit der konservatorisch nötigen Sorgfalt zu behandeln, auch wenn dies den Durchsatz bei der Digitalisierung reduziert und einen erhöhten Zeitaufwand erfordert. Die Scantechnik und der Einsatz von Hilfsmitteln, wie Buchkeile oder Fixierhilfen, sind entsprechend den Vorgaben der konservatorischen Prüfung zu wählen.

3.2 Technische Parameter der digitalen Reproduktion

Ziel der Digitalisierung ist die möglichst originalgetreue Wiedergabe des Objektes nach Maßgabe der wissenschaftlichen Erfordernisse. Die anzuwendenden Parameter für die Digitalisierung sind mit Blick auf die Qualität des Bildes, seine Langzeitverfügbarkeit und Interoperabilität zu wählen.

3.2.1 Allgemeine Erläuterungen und Parameter

Bei der Herstellung sind grundsätzlich zwei Formen von Digitalisaten zu berücksichtigen. Zum einen der sogenannte digitale Master, also das Ausgangs- oder Archivformat, und für die Nutzung hergestellte Derivate, meist verkleinerte Kopien in anderen Dateiformaten. Derivate wie z. B. JPEG-Dateien sind in Abhängigkeit von der gewählten Präsentation aus den Mastern zu erzeugen und können nach Bedarf modifiziert werden. Eine Modifikation wird z. B. nötig, wenn sich die bei den Nutzerinnen und Nutzern zu erwartenden Bildschirmauflösungen ändern oder Bildformate zum Einsatz kommen sollen, die optimierte Eigenschaften für die gewünschte Anzeige haben (z. B. stufenfreies Zoomen; fließende Übergänge zwischen Abschnitten großer, gleichzeitig aber detailreicher Objekte).

---

[15] Kriterien einer konservatorischen Prüfung sollten sein: Druckfarbe/Tinte/Tusche gefährdet; Farbauftrag/Grundierung gefährdet; Tintenfraß/Farbfraß; Heftung lose; Bünde oder Überzugmaterial zu steif/unflexibel; Bünde (an-)gebrochen; Überzug im Gelenk (an-)gebrochen; Rückeneinlage zu steif; Einbandrücken beschädigt; Brüche, Risse oder Fehlstellen; Ledernarben im Rücken empfindlich; Rückenvergoldung gefährdet; sehr dicker Pergamentband mit hohlem Rücken; Buchdeckel (an-)gebrochen; Buchdeckel gelöst (vorne/hinten); Schließenriemen zu steif/angebrochen, mikrobieller Befall; Mikroform vorhanden. Darüber hinaus bestehen mitunter auch technische Probleme, die eine Digitalisierung erschweren oder verhindern: Bundsteg zu schmal; Buchblock wellig; Blätter/Lagen sehr steif; Textverlust unvermeidbar; Planlage der Vorlage unmöglich; extremes Format.

[16] Corbach, Almuth: Bestandsschonendes Digitalisieren von schriftlichem Kulturgut. In: Digital und analog. Die beiden Archivwelten. 46. Rheinischer Archivtag. Ratingen 21.-22. Juni 2012. Beiträge.
Hrsg.: LVR-Archivberatungs- und Fortbildungszentrum. Bonn 2013 (Archivhefte 43), S. 90-102;

[17] In Archiven ist die Digitalisierung vom Mikrofilm unabhängig von konservatorischen Erwägungen durchaus Praxis. Zu den materialspezifischen Parametern für eine Digitalisierung vom Mikrofilm vgl. Kapitel 3.2.2.

Der Master ist die Grundlage für alle weiteren Prozesse. Daher sollte man seiner Herstellung besondere Aufmerksamkeit schenken und bei der Langzeitsicherung entsprechende Empfehlungen berücksichtigen.

Die folgenden Richtlinien für die Mindestanforderungen an die Digitalisierung beziehen sich ausschließlich auf den **digitalen Master**.

3.2.1.1 Auflösung

Das gebräuchliche Maß für das Auflösungsvermögen einer Datei ist dpi (*dots per inch*), gleich Bildpunkte pro inch.

Als Untergrenze sollte daher eine Scanauflösung für die Digitalisate gewählt werden, bei der die Details einer Vorlage vollständig in einer gleichgroßen Reproduktion wiedergegeben werden können.[18]

Daher gilt die grundsätzliche Empfehlung von **300 dpi bezogen auf das Vorlagenformat als Zielauflösung** (Vorlagenformat = Ausgabeformat des Digitalisats bei 300 dpi).

300 dpi beziehen sich allerdings nur auf Vorlagen, die für die Betrachtung mit bloßem Auge gedacht sind, wie etwa: Textwerke, Grafiken und fotografische Aufsichtvorlagen. Anders verhält es sich bei Medien, deren vollständige Bildinformation nur in Vergrößerung erfassbar wird. Das sind Miniaturen jeglicher Art und ganz besonders fotografische Durchsichtmedien (z. B. Negative oder Dias).[19]

Weiterhin gibt es Objekte, insbesondere Objekte der Architektur, aber auch überformatige Malereien, die sich auf Grund ihrer Dimensionen den oben beschriebenen Verfahren zur Bestimmung der Auflösung entziehen.

In der bildmäßigen Fotografie für Sammlungsobjekte und Architekturfotografie wird beispielsweise mit einer modernen 50-Megapixel-Digitalkamera eine Auflösung von 300 dpi bei DIN A2 Ausgabegröße erreicht. Diese Auflösung ist ausreichend für vielseitige Verwendungen wie Druckerzeugnisse unterschiedlicher Größe und die Präsentation im Internet. Jedoch können höhere Auflösungen notwendig sein, wenn bildwichtige Details mit diesen Vorgaben nicht dargestellt werden können (z. B. bei Reproduktionen von Gemälden größer als doppelt DIN A0).[20]

---

[18] Ein Ansatz zur Ermittlung der notwendigen Mindestauflösung eines Digitalisats ist das Auflösungsvermögen des menschlichen Auges bei deutlicher Sehweite. Die kürzeste Sehweite, bei der man einen Gegenstand ermüdungsfrei über einen längeren Zeitraum betrachten kann, liegt bei ca. 25 cm. Das menschliche Auge ist in der Lage, zwei Linien noch zu unterscheiden, wenn das Licht auf mindestens zwei nicht benachbarte Sehzellen trifft und mindestens eine Sehzelle dazwischen liegt. Eine Rasterfrequenz von 60 Linien/cm bei 20 cm Betrachtungsabstand kann nicht mehr als getrennte Linien wahrgenommen werden. Daraus resultiert folgende Auflösungsanforderung, die auch in der Druckindustrie gebräuchlich ist:
60 Linien/cm (60-er Raster) benötigen mindestens 120 Pixel zur Darstellung.
120 px/cm x 2,54 cm/inch = 304,8 dpi, rund 300 dpi

[19] Beispiel: Kleinbildnegative mit der Größe von 24 x 36 mm sind für die anschließende Vergrößerung auf Fotopapier hergestellt worden. Sie eignen sich nicht für die Betrachtung mit bloßem Auge. Als fotografische Praxisregel gilt, dass sich Kleinbildnegative (abhängig von Objektiv, Belichtung, Film und Entwickler) ca. 10-fach vergrößern lassen, also auf einen Fotoabzug der Größe von 24 x 36 cm. Um diese Zielgröße von 24 x 26 cm nach obiger Maßgabe im Digitalisat des KB-Negativs zu erreichen, müsste es mit mindestens 3.000 dpi digitalisiert werden.

[20] Bei einem Flächensensor mit 5.700 x 8.600 Pixeln ergibt sich also nach der Beziehung Kamerapixel/Zielauflösung in dpi = Zielgröße in inch folgende Ausgabegröße bei 300 dpi:
8600/300 = 18,7 inch x 2,54 cm/inch = 72 cm
5700/300 = 12,7 inch x 2,54 cm/inch = 48 cm

Hinweise für eine gegebenenfalls sinnvolle Abweichung von der grundsätzlichen Auflösungsempfehlung von 300 dpi bezogen auf das Vorlagenformat werden bei den materialspezifischen Parametern genannt. (→ 3.2.2)

Die erreichbare Auflösung eines digitalen Aufnahmesystems ist dabei nicht nur von der Anzahl der zur Verfügung stehenden Pixel einer Scanzeile oder eines Aufnahmesensors abhängig. Sie ist ein Zusammenspiel aus unterschiedlichen Faktoren. So spielen etwa die technische Qualität des digitalen Aufzeichnungsgeräts (Kamera, Scanner) und die Abbildungsleistung der verwendeten Objektive die entscheidende Rolle.

Ob ein digitales Aufnahmesystem für ein Digitalisierungsvorhaben geeignet ist, lässt sich vorab im Rahmen einer Testdigitalisierung mit genormten Testcharts (ISO-12233-Testchart, USAF-1951-Testchart) ermitteln.

Die effektive Auflösung der Testbilder lässt sich dann mit einer Analysesoftware oder mit entsprechenden, in der Regel beigefügten Auflösungstabellen ermitteln.

3.2.1.2 Farbtiefe

Die Farbtiefe bestimmt die Differenzierung der Helligkeits- und Farbwerte in einem Digitalisat. Da in der digitalen Technik nur diskrete Zustände (ja/nein) möglich sind, können Helligkeits- und Farbunterschiede – im Gegensatz zur analogen Fotografie – nicht kontinuierlich, d. h. mit fließenden Übergängen dargestellt werden.

- Bei einer Farbtiefe von 1 Bit sind zwei unterschiedliche Zustände möglich: Weiß und Schwarz.
- Bei einer Farbtiefe von 8 Bit/Farbkanal sind es $2^8 = 256$ Helligkeitsstufen/Farbkanal von Weiß nach Schwarz.
- Bei einer Farbtiefe von 16/Farbkanal Bit sind es $2^{16} = 65.536$ Helligkeitsstufen/Farbkanal von Weiß nach Schwarz.
- Graustufenbilder mit nur einem Kanal für die Helligkeit haben somit eine Farbtiefe von 8 oder 16 Bit.
- Farbbilder im RGB-Modus (je ein Farbkanal für Rot, Grün und Blau) besitzen somit eine Farbtiefe von 3 x 8 Bit = 24 Bit, bzw. 3 x 16 Bit = 48 Bit.

Der Vorteil der höheren Farbtiefe von 16 Bit pro Kanal liegt also in der größeren Farbdifferenzierung. Dadurch gehen bei einer späteren Bildbearbeitung weniger Tonwerte verloren. Dieser Bearbeitungsspielraum ist wichtig bei bearbeitungsintensiven Digitalisaten, z. B. von SW-Negativen.

Digitale Aufnahmesysteme zeichnen Farben im RGB-Farbmodus auf. Die Analog-/Digitalwandler in hochleistungsfähigen chipbasierten Scanbacks und Zeilenscanbacks können Helligkeitsunterschiede mit 16 Bit differenzieren. Bei den meisten digitalen KB-Kameras und Flachbettscannern sind es 14 Bit.

Digitale Aufnahmesysteme zeichnen grundsätzlich mit einer hohen Farbtiefe von 14 Bit bzw. 16 Bit pro Kanal auf. Um die Informationsdichte der erzeugten Bilddateien zu erhalten, wird im Idealfall eine notwendige Nachbearbeitung der Bilddaten in 16 Bit pro Kanal, d. h. 48 Bit durchgeführt, mindestens jedoch in 8 Bit pro Kanal, d. h. 24 Bit Farbtiefe.

Für die Sicherung des finalen digitalen Master ist eine **Farbtiefe** von **8 Bit pro Kanal**, d. h. **24 Bit** ausreichend, da die heute gängigen Ausgabe- und Anzeigegeräte nur eine Tonwertwiedergabe mit 8-Bit-Differenzierung unterstützen.

3.2.1.3 Digitaler Aufnahmeablauf

Aufnahmetechnik

In der fotografischen digitalen Aufnahmetechnik unterscheidet man zwischen der Zeilenscantechnik, bei der eine trilineare Scanzeile mit fester Breite (je eine Zeile ist für eine Farbe – Rot, Grün, Blau – empfindlich) eine bestimmte Strecke abtastet. Flachbettscanner und Zeilenscanrückteile für Fach- oder Mittelformatkameras arbeiten mit dieser Technik. Hierbei wird für jeden Bildpunkt die Farbinformation physikalisch erzeugt.

Das zweite gebräuchliche Verfahren ist der Flächensensor, bei dem auf einer definierten Fläche eine bestimmte Anzahl von Fotodioden angebracht ist.

Zeilenscanner und multishotfähige Flächensensorkameras finden hauptsächlich Anwendung bei filigranen Motiven wie bildhaften Tiefdruckwerken, Kartenwerken oder Textilien. Zeilenscanner ermöglichen bei Vorlagen bis DIN A0 hohe Auflösungen und extreme Detailgenauigkeit. Für große zweidimensionale Vorlagen ab doppelt DIN A0, die zur Gewährleistung der Lesbarkeit mit höchster Detailwiedergabe aufgezeichnet werden sollen, sind Zeilensensoren in Verbindung mit einem Kamerasystem vorzuziehen.

Mit gebräuchlichen Flachbettscannern lassen sich zweidimensionale Medien bis zu einer Vorlagengröße von DIN A3 digitalisieren. Das Verfahren ist allerdings nicht berührungsfrei, da die Medien mit dem Vorlagenglas des Scanners in Kontakt kommen, sie eignen sich daher nur für bestandserhalterisch unbedenkliche Vorlagen.

Bildrauschen

Um die Details einer Vorlage möglichst vollständig zu reproduzieren, sollten bei der eingesetzten Kameratechnik die herstellerspezifischen Empfehlungen zur Verminderung von Bildrauschen beachtet werden. Grundsätzlich gilt dabei die Faustregel: Je geringer die verwendete ISO-Einstellung ist, desto geringer ist das Rauschen.

Objektive

Weiterhin empfiehlt sich die Verwendung moderner, hochwertiger Objektive, die für das hohe Auflösungspotential digitaler Kameras entwickelt worden sind. Festbrennweiten sind Zoomobjektiven vorzuziehen, da bei ihnen Abbildungsfehler gezielt korrigiert werden und der „Über-Kopf-Einsatz“ auf Dauer die Mechanik beschädigt. Zoomobjektive bieten immer nur einen Qualitätskompromiss über den vorhandenen Brennweitenbereich. Shiftoptiken vermindern Verluste in der Bildqualität, die durch nachträgliches Entzerren stürzender Perspektivlinien entstehen. Ist der verwendete Abbildungsmaßstab unter 1:10, sollten spezielle Makroobjektive verwendet werden, die für diesen Abbildungsbereich konstruiert wurden. Abhängig vom ver-

wendeten Kamerasystem können für Reproduktionen in diesem Maßstabsbereich auch hochwertige Vergrößerungsobjektive eingesetzt werden.[21]

Arbeitsplatz

Zur Vermeidung von Unschärfen durch Verwacklungen während des Digitalisierens sollte darauf geachtet werden, dass bei der Wahl des Standorts für eine Digitalisierungseinheit ein erschütterungsfreier Platz gesucht wird. Holzfußböden z. B. sind ungeeignet, da sie schwingen und somit Vibrationen leicht übertragen. Werden Spiegelreflexkameras benutzt, sollten sie mit Spiegelvorauslösung ausgelöst werden, um Vibrationen des Spiegelschlags auszuschließen. Stative und Reprosäulen sollten so dimensioniert sein, dass sie auch das Gewicht der eingesetzten Kameras problemlos halten können.

Moiré

Bei Objekten mit feinsten, gleichmäßig verteilten Details kann das Oneshot-Verfahren bei Digitalkameras mit Flächensensor zu Farbverfälschungen und Moiré-Effekten führen.[22]

Ein Moiré-Effekt tritt in der digitalen Bilderzeugung auf, wenn in Abhängigkeit von der Pixelgröße der Scanzeile oder des Flächensensors, des Abbildungsmaßstabes und der Linienfrequenz der Vorlage die gleichmäßige Pixelmatrix eines digital erzeugten Bildes mit einer regelmäßigen Linienstruktur der Vorlage interferiert. Das Auftreten dieses Effekts ist schwer vorhersehbar, da die entsprechenden Parameter selten konstant bleiben.

Farbverfälschungen treten bei Oneshot-Flächensensoren unter den gleichen Bedingungen auf, sind aber ein Interpolationsfehler der Kamerasoftware. Bei Oneshot-Kamerasystemen ist ein Pixel jeweils nur für eine der Grundfarben Rot, Grün, Blau empfindlich. Die restlichen Farbinformationen für den jeweiligen Bildpunkt werden aus den Farbinformationen der umliegenden Pixel interpoliert. Sind in Pixelabstand die Farb- und Helligkeitsunterschiede der Vorlage zu extrem, kommt es zu Farbverfälschungen. In solchen Fällen können Zeilenscanner oder multishotfähige Digitalrückteile bessere Ergebnisse liefern, da hier die Farbinformation physikalisch erzeugt wird und nicht rechnerisch. Im Gegenzug erhöht sich jedoch der Digitalisierungsaufwand etwas, da die Scanzeiten länger sind und eine absolut konstante Beleuchtung während des Aufnahmevorgangs unabdingbar ist.

Zur Vermeidung des Moiré-Effekts sollte das notwendige Verfahren und die notwendige Auflösung im Vorfeld eines Digitalisierungsvorhabens an Hand einer Testdigitalisierung ermittelt werden.

Beispielhaft zu nennende Objekte, bei denen Moiré auftreten kann, sind: gedruckte Halbtonvorlagen, Kupferstiche, Textilien.

Beleuchtungstechnik

---

[21] Die Abbildungsleistung eines Objektivs lässt sich auch an den dazugehörigen Modulationstransferfunktions-Diagrammen (MTF-Diagrammen) ablesen. Diese Diagramme werden von diversen Objektivherstellern veröffentlicht. Die Modulationstransferfunktion beschreibt, wie gut ein Objektiv den Kantenkontrast in einer Vorlage auf das fotografische Abbild übertragen kann. Sie bietet somit ein Hinweis auf das Auflösungsvermögen einer Optik. Die technischen Datenblätter der Objektive geben auch Auskunft über den Helligkeits- und Schärfeabfall zum Bildrand. Optische Abbildungsfehler machen sich hier besonders bemerkbar. Zur Reduzierung der Bildfehler sollte ein Objektiv zweimal abgeblendet werden, um hauptsächlich die bessere Abbildungsleistung der Objektivmitte zu nutzen. Wenn keine große Tiefenschärfe benötigt wird, wie bei Reproduktionen planer Vorlagen, sollte mit dieser optimalen Arbeitsblende fotografiert werden.

[22] Das hier behandelte Moiré ist nicht mit dem Phänomen zu verwechseln, das aus der Interferenz von Druckpunkten gerasterter Vorlagen mit den Bildpunkten des Monitors entsteht.

Die Wahl der richtigen Beleuchtungstechnik ist im Vorfeld eines Digitalisierungsvorhabens unter bestandserhalterischen Gesichtspunkten zu prüfen. Zeilenscansysteme benötigen flickerfreies Dauerlicht, da hier über einen bestimmten Zeitraum die Vorlage kontinuierlich abgetastet wird. Lichtleistung und -farbe müssen während des Scanvorgangs konstant sein, um gleichmäßige Ergebnisse zu erzielen. Flächensensoren können auch mit Blitzlicht benutzt werden. Bei Multishotsystemen ist darauf zu achten, dass die Blitzleistung und -farbe während der einzelnen Belichtungen konstant bleiben. Gleiches gilt für Dauerlicht.

Unter konservatorischen Gesichtspunkten lässt sich die Meinung vertreten, dass eine kurzfristige hohe Lichtbelastung weniger schädlich ist als eine hohe Wärmebelastung, zumal bezüglich der Schadwirkung des Lichts die Lichtbilanz zählt, d. h. es ist unerheblich, ob die gleiche Lichtmenge über einen längeren Zeitraum verteilt oder kurzfristig auf das Objekt trifft. Kurzfristige hohe Helligkeitsunterschiede sind jedoch aus arbeitsmedizinischer Sicht für die Mitarbeitenden problematisch.

Bildverarbeitung

Fotografische Aufnahmen werden grundsätzlich im herstellerbedingten Rohdatenformat (RAW-Format) in der maximalen Größe erzeugt, mit einer Farbtiefe von 14 oder 16 Bit pro Farbkanal. RAW-Bilder repräsentieren die originalen Kameradaten. Bildkorrekturen sollten möglichst in RAW-Software vorgenommen werden. Da diese Datei nicht verändert wird, kann sie bei Bedarf verlustfrei neu korrigiert werden. Bearbeitungsvorschriften können überdies als *Preset* auf beliebig viele Bilder angewendet werden, wodurch eine hohe Arbeitseffizienz erreicht wird.

Als Profil für Farbaufnahmen sind entweder Adobe RGB, der von der European Color Initiative (ECI)[23] empfohlene Arbeitsfarbraum ECI-RGB v2 oder der dazu identische Farbraum L-Star RGB[24] zu wählen. CMYK-Farbräume kommen als ausschließliche Druckausgabeprofile nicht in Frage.
Für eine verlässliche Farbabstimmung in der späteren Bildverarbeitung ist es notwendig, Graukeile, Farbkeile oder Farbcharts mit zu produzieren. Sofern man keine Einstellungsänderung vornimmt, ist es ausreichend, den diese pro Bildserie nur einmal repräsentativ für die ganze Serie aufzunehmen.

Bei der Helligkeits- und Tonwertsteuerung ist darauf zu achten, dass das Tonwerthistogramm (für Schwarz und Weiß) seitlich nicht angeschnitten wird, damit der Tonwertumfang erhalten bleibt. Nach heutigem Stand sind Tonwerte von 95 % für Schwarz und 5 % für Weiß differenziert druckbar. Als maximale Helligkeitswerte empfehlen sich:

Schwarz:	RGB 16/16/16, Graustufen 90 %
Weiß:	RGB 232/232/232, Graustufen 10 %

Bei Reproduktionen von zweidimensionalen Vorlagen sind Freisteller so zu gestalten, dass die gesamte Vorlage mit leichtem, umlaufendem Rand abgebildet wird. Nur so ist erkennbar, dass nichts von der Vorlage abgeschnitten wurde. Das Nachschärfen („Unscharfmaskieren“) sollte

[23] European Color Initiative: http://www.eci.org
[24] http://www.colormanagement.org/de/workingspaces.html

für den digitalen Master nur moderat erfolgen. Anwendungsgebunden kann für unterschiedliche Derivate eine zusätzliche Schärfung erfolgen.

Notwendige Bildbearbeitungen in der Bilddatei sind im Idealfall bei 16 Bit Farbtiefe pro Kanal vorzunehmen, mindestens jedoch in 8 Bit pro Kanal, d. h. 24 Bit Farbtiefe. Sind alle Korrekturen vorgenommen, wird die RAW-Datei anschließend in eine TIFF-Datei mit 24 Bit/RGB oder 8 Bit/Graustufen als Master konvertiert. Sind noch weitere Bearbeitungen in der TIFF-Datei vorgesehen, kann die Konvertierung aus der RAW-Datei mit 48 Bit/RGB oder 16 Bit/Graustufen erfolgen. Nach Beendigung aller notwendigen Bearbeitungsschritte wird der finale Master als TIFF mit 24 Bit/RGB oder 8 Bit/Graustufen gesichert und gilt als archivtauglich. Das RAW-Format ist als proprietäres Dateiformat allerdings nicht archivtauglich.

Für einen verlässlichen Abgleich originaler Vorlagen mit der Monitordarstellung des Digitalisats sollten die Vorlagen unter D50-Normlichtumgebung für grafische Arbeitsplätze nach ISO 3664:2000 betrachtet werden. Monitormodelle, die in der grafischen Industrie Verwendung finden, liefern in diesem Zusammenhang die verbindlichsten Darstellungen. Sie bieten über die gesamte Grauachse eine konsistente Farbwiedergabe. Die Monitore sollten mit einem Colorimeter auf die entsprechenden Zielwerte für eine D50-Normlichtumgebung kalibriert werden. Dabei ist zu beachten, dass eine Farbtemperatur von 5.800 K statt 5.000 K am Monitor eingestellt wird, da 5.000 K am Monitor gelber wahrgenommen werden als 5.000 K der D50-Normlichtumgebung. Folgende Eckwerte sind zu empfehlen:[25]

- Luminanz: mindestens 120 cd/m2
- Weißpunkt: 5800 K mit chromatischer Adaption
- Gradation: 1.8 Gamma oder alternativ L*
- Farbräume: ISOcoated v2 und ECI-RGB 1.0/2.0

Digitale Bildnachbearbeitung

Zur Optimierung der Bildqualität digitaler Aufnahmen ist in der Regel eine zusätzliche Nachbearbeitung notwendig. Dabei sollte sich diese auf notwendige Farb- und Tonwertkorrekturen beschränken. Auf jeden Fall zu vermeiden sind Objektdeformationen, das Hinzufügen oder Löschen von Objektteilen, sowie Spezialeffekte, wie z. B. der Einsatz von Verfremdungsfiltern. Zur Steigerung der Bildintegrität können Positionierungshilfen nachträglich aus dem Bild entfernt werden oder Hintergründe bereinigt werden.

3.2.1.4 Dateiformate

Bildmaster von Graustufen oder Farbbildern sollten nach dem derzeitigen Kenntnisstand in „**TIFF uncompressed**“[26] langzeitgesichert werden. Das Format TIFF gibt es schon seit den 1980er Jahren. Es hat sich als einer der wichtigsten De-Facto-Standards etabliert und es ist damit zu rechnen, dass es auch in Zukunft von allen Standardprogrammen unterstützt wird. Allerdings gilt dies nur für so genannte „Baseline-TIFFs“. Die weitgehenden Optionen von Extended TIFFs sollten für den digitalen Master nicht genutzt werden.

---

[25] Vgl. die Empfehlungen des Schweizer Kompetenzzentrums für Medien- und Drucktechnologie: http://www.ugra.ch

[26] http://www.nationalarchives.gov.uk/PRONOM/Format/proFormatSearch.aspx?status=detailReport&id=686; http://www.nationalarchives.gov.uk/PRONOM/Format/proFormatSearch.aspx?status=detailReport&id=1099

Neben TIFF kann auch **TIFF-LZW** oder **JPEG2000**[27] in seiner verlustfreien Form als Format für den Bildmaster verwendet werden. Für die Speicherung von Mastern im JPEG2000-Format ist allerdings darauf zu achten, dass nur die lizenzfreien Bereiche von JPEG2000 Verwendung finden.

In den letzten Jahren gelangten gerade angesichts der zunehmenden Speicherproblematik der ISO-Standard JPEG2000[28] und nach Ablaufen der Lizenzen TIFF-LZW als effiziente Kompressionsformate ins Blickfeld der Gedächtnisinstitutionen. Mit Blick auf die Langzeitarchivierung sollten die Betreibenden von Repositorien die Vor- und Nachteile des Einsatzes beider Formate (TIFF-LZW und JPEG2000) sorgfältig abwägen. Kompressionsformate sind grundsätzlich anfälliger für Bildverluste und über ihren Einsatz sollte nach einer Nutzen-Risiko-Abwägung entschieden werden. Bei der Auswahl eines Formates ist auch seine Verbreitung und Marktdurchdringung zu berücksichtigen. Unabhängig davon, dass international große und einflussreiche Bibliotheken wie die Library of Congress und die British Library auf JPEG2000 setzen, ist seine Verbreitung bisher geringer als die des unkomprimierten TIFF-Formates. Die Lizenzierungssituation ist noch nicht abschließend geklärt, einige Bereiche von JPEG2000 wurden jedoch als frei deklariert.

Auch die unterschiedlichen proprietären RAW-Formate sind nicht zur Sicherung der Bildaufnahme als Master geeignet, zumal sie oftmals nur mit der dazugehörigen RAW-Software angezeigt werden können. Ebenso hat sich das plattformübergreifende, kameraunabhängige Adobe-RAW-Format DNG nicht durchgesetzt und ist somit als Archivformat ungeeignet.

Für die Publikation im Internet empfehlen sich wegen ihrer großen Verbreitung grundsätzlich JPEG und PNG. Jedoch können auch andere Formate oder Techniken zum Einsatz gelangen, wenn diese für die Präsentation z. B. von sehr großen Bilddateien praktikabler sind. Hierbei ist darauf zu achten, dass nur Formate und Techniken eingesetzt werden, die ohne spezielle Technik oder Software von der Wissenschaft genutzt werden können.

Für AV-Medien muss ebenso zwischen Archivformaten und Nutzungsformaten unterschieden werden. Im Bereich Audio hat sich das Waveform Audio File-Format (WAV) in Verbindung mit der Pulse Code Modulation (PCM)[29] als Archivierungsformat etabliert. Als Nutzungsformat hat sich mp3 (MPEG-2 Audiolayer III) durchgesetzt. Für Videoformate kann zurzeit noch keine eindeutige Empfehlung ausgesprochen werden. Jedoch sollte unbedingt darauf geachtet werden, nur lizenzfreie und offen beschriebene Formate zu verwenden. Gerade bei Videodateien liegt das Problem in der enormen Größe, so ist die Anforderung einer unkomprimierten Speicherung bei größeren Projekten kaum umsetzbar.[30]

Digitale 3D-Dokumentation von digitalisierten bzw. digital rekonstruierten 3D-Objekten werden als 3D-Modelle abgelegt. Diese Modelle ermöglichen es, die Form, die Textur und die opti-

---

[27] http://www.jpeg.org/jpeg2000/
[28] ISO/IEC 15444-6:2013
[29] Mindestqualität: 44,1 khz Abtastrate und 16 bit Abtasttiefe
[30] Folgende Formate werden für die Langzeitarchivierung empfohlen, wobei zurzeit noch keine befriedigende Lösung vorliegt: MJPEG2000/MPEG-4 (ISO/IEC 14496-12:2015, ISO/IEC 15444-12:2015), DPX (Standardisiert als SMPTE 268M-2003, v 2.0), MXF/AAF(Standardisiert als SMPTE 377M und als ISO-Standard vorgeschlagen). Die FADGI arbeitet zurzeit an einer Richtlinie zur Archivierung von Filmen: http://www.digitizationguidelines.gov/guidelines/MXF_app_spec.html

schen Materialeigenschaften des Ursprungsobjekts originalgetreu abzubilden, bzw. nicht mehr vorhandene (u. a. zerstörte) Objekte zu dokumentieren und zu visualisieren.

Infrage kommende Datenformate sollten möglichst robust gegenüber Beschädigungen auf den Datenträgern sein, Speicherplatz möglichst effizient nutzen, das Modell logisch strukturiert nachbilden, eine schnelle Datenverarbeitung ermöglichen und entsprechend verbreitet sein.

Als bewährtes Dateiformat im Bereich der 3D-Modelle hat sich OBJ (.obj)[31], als offenes Dateiformat, bewährt. Das Format wird von vielen 3D-Grafikprogrammen unterstützt und ist daher geeignet für die programm- und plattformübergreifende Weitergabe von 3D-Modellen. Optische Materialeigenschaften (z. B. Spiegelung, Transparenz, Glanzlicht usw.) werden in einer separaten Materialdatei (.mtl) definiert, die auch Angaben zu Texturierungen enthalten kann.

Erwähnenswert ist ebenfalls das Format Collada DAE (.dae)[32], ein auf XML basierendes offenes Austauschformat von 3D-Datensätzen. Dabei geht es nicht nur um die Weitergabe von Modellen und Texturen, auch Einstellungen und angewandte Veränderungsschritte sollen so von einem Programm zum anderen übermittelt werden können.

Für die web-basierte Dokumentation und Vermittlung bietet sich das speziell zur Visualisierung von 3D-Modellen innerhalb der WebGL-Technologie entwickelte Format X3D (.x3d)[33] an.

3.2.2 Materialspezifische Parameter

3.2.2.1 Textwerke

Zu den Textwerken zählen in diesem Zusammenhang sowohl gedruckte Werke als auch unikale Dokumente wie Handschriften und Archivgut.

Bei der Altbestandsdigitalisierung ist auf jeden Fall die Imagedigitalisierung vorzunehmen. Auch bei Vorliegen eines maschinenlesbaren Volltexts sollte auf die Imagedigitalisierung bzw. auf die Präsentation des digitalen Faksimiles nicht verzichtet werden, da eine Fülle von Informationen nur über das visuelle Abbild transportiert werden.

Die Lesbarkeit der Texte ist ausschlaggebend für die Wahl der Auflösung und von der im Text verwendeten Schriftgröße abhängig. Die benötigte Scanauflösung richtet sich in diesem Fall weniger nach den Maßen der Vorlage als nach der Lesbarkeit der einzelnen Lettern. Für Textdokumente mit dem kleinsten signifikanten Zeichen ab 1,0 mm ist eine Auflösung von 400 dpi und erst ab 1,5 mm Mindestzeichengröße eine Auflösung von 300 dpi zu empfehlen.[34]

Die Blätter/Seiten der Textwerke werden immer vollständig mit leichtem umlaufendem Rand digitalisiert, um deutlich zu machen, dass nichts von der Vorlage abgeschnitten wurde.

---

[31] http://www.fileformat.info/format/wavefrontobj/egff.htm und https://de.wikipedia.org/wiki/Wavefront_OBJ
[32] https://www.khronos.org/collada/ und https://de.wikipedia.org/wiki/Collada_(Speicherformat)
[33] http://www.web3d.org/x3d/what-x3d und https://de.wikipedia.org/wiki/X3D
[34] Vgl. Federal Agencies Digitization Guidelines Initiative (FADGI): Technical Guidelines for Digitizing Cultural Heritage Materials: Creation of Raster Image Master Files, August 2010: http://www.digitizationguidelines.gov/guidelines/FADGI_Still_Image-Tech_Guidelines_2010-08-24.pdf, S. 59.

3.2.2.2 Grafische Darstellungen

Wie unter 3.2.1.1 beschrieben liegt die empfohlene Mindestauflösung für die Digitalisierung bei 300 dpi unter Berücksichtigung der Ergebnisse einer Testdigitalisierung mit genormter Testtafel.

Häufig wird jedoch bei kleinen Vorlagen eine Auflösung von 300 dpi bezogen auf das Vorlagenformat nicht ausreichen, um die Eigenschaften der Vorlage umfassend erkennbar wiederzugeben. Ist z. B. die künstlerische Technik der Ausführung nicht identifizierbar, sollte unbedingt eine höhere Auflösung gewählt werden. Beispielhaft seien hier erwähnt: Kupferstiche, Briefmarken, Medaillonportraits, Miniaturmalerei.

Die erreichbare Maximalauflösung ist dabei von der eingesetzten Digitalisierungstechnik abhängig. Hierzu ein Rechenbeispieleispiel ohne Berücksichtigung der Objektivqualität: Eine Briefmarke ließe sich mit einer Digitalkamera mit einem Sensor im Kleinbildformat (24 x 36 mm) im Maßstab 1:1 reproduzieren. Bei einem Pixelmaß der Kamera von 3800 Pixeln/cm ergibt sich hochgerechnet eine Auflösung von 4021 dpi.

Umgekehrt kann bei überformatigen Vorlagen (DIN A0 und größer) die Auflösung bezogen auf das Vorlagenformat reduziert werden, wenn die Objekte für einen großen Betrachtungsabstand entworfen wurden. Plakate sind hier als Beispiel zu nennen, sie sind oftmals auf Signalwirkung und Fernsicht ausgelegt. In diesem Fall kann die Auflösung bis auf 150 dpi reduziert werden, da mit wachsendem Betrachtungsabstand auch die Punktgröße wächst, die vom menschlichen Auge nicht mehr unterschieden werden kann. Sind Überformate jedoch detailreich wie beispielsweise Generalkarten oder Kupferstiche sollte mit mindestens 300 dpi digitalisiert werden.

Aufgrund des Mehrwerts für die Benutzung sollte in einem ökonomisch vertretbaren Rahmen die volle Auflösung der eingesetzten Aufnahmetechnik bei den grafischen Darstellungen genutzt werden.

Die Grafiken werden immer vollständig mit leichtem umlaufendem Rand gesichert, um deutlich zu machen, dass nichts von der Vorlage abgeschnitten wurde.

3.2.2.3 Fotografien

Bei Fotografien ist zunächst zwischen Durchsichtmedien, also fotografischen Negativen oder Diapositiven, und Aufsichtmedien, also Fotopositiven z. B. auf Papier, zu unterscheiden. In der Regel handelt es sich bei Letzteren um Kontaktkopien oder Vergrößerungen von Negativen auf Fotopapier.

Durchsichtmedien

Das fotografische Negativ oder Diapositiv ist das Ergebnis einer fotografischen Aufnahme und damit die originale Quelle. Negative sind die Vorlagen für die Erstellung von Reproduktionen zur Verbreitung der Aufnahme als Fotopositiv oder Druck.

Bei Diapositiven lässt sich die Qualität eines Scans visuell mit Hilfe eine Normlichtleuchtkastens mit dem Original vergleichen. Zur Farb- und Tonwertkorrektur sollte die Bilddatei idealerweise in 48 Bit/RGB, mindestens jedoch in 24Bit/RGB vorliegen. Sind keine weiteren Änderungen an der Datei vorgesehen, wird das Endergebnis mit 24 Bit/RGB gesichert. Soll vorrangig der Zustand des Vorlagenmediums zum Zeitpunkt des Scans als digitaler Master doku-

mentiert werden, empfiehlt sich alternativ auch eine Sicherung mit 48 Bit/RGB, besonders wenn die originale Vorlage Alterungserscheinungen aufweist, die für die diversen Derivate bearbeitungsintensiv korrigiert werden sollen.

Anders verhält es sich mit Negativen: Die negative Darstellung ist für die Betrachtung des Motivs ungeeignet. Um das Bildmotiv zu nutzen, muss ein positives Derivat aus dem digitalen Master erzeugt werden. Dieses dient als digitaler Fotoabzug vom Negativ und gibt den Bildinhalt wieder. Die positive Darstellung eines Masters ist aber im Regelfall zu flau und bei Farbnegativen zusätzlich durch deren Farbmaske nicht farbrichtig. Um das Motiv im digitalen Fotoabzug klar erkennbar wiederzugeben, ist daher eine beträchtliche Bildkorrektur nötig. Soll vorrangig der Zustand des Mediums zum Zeitpunkt des Scans als digitaler Master gespeichert werden, empfiehlt sich alternativ auch eine Sicherung mit 48Bit/RGB bzw. 16 Bit/Graustufen. Bei SW-Durchsichtmedien ist die Speicherung in Graustufen ausreichend. Das Speichern von Farbinformationen bei SW-Vorlagen ist nur dann zu vertreten, wenn wichtige Informationen zu dem Medium in Farbe vorliegen, wie Retuschen im Negativ.

Gut ausgearbeitete Negative geben Bilder mit einem Dichteumfang (= Tonwertumfang) von bis zu 12 Blendenstufen wieder, die so auch von der Digitalisierungsvorrichtung reproduziert werden müssen.

Anders als die grafischen Darstellungen wird das fotografische Abbild technisch erzeugt. Die Qualität einer fotografischen Aufnahme ist somit vom Zusammenspiel unterschiedlicher technischer Faktoren abhängig: Aufnahmeformat, Objektiv, Filmtyp, Körnigkeit der Emulsion, Entwicklung, Belichtung und Fokussierung etc.

Um diesen Aspekten in die Bestimmung einer sinnvollen Scanauflösung einfließen zu lassen, bietet sich folgende Alternative zu dem in 3.2.1.1 beschriebenen Verfahren an:

Außer dem Aufnahmeformat sind die Aufnahmebedingungen in der Regel nicht bekannt oder dokumentiert. Das verwendete Filmmaterial lässt sich möglicherweise an Hand einer Kennung, wie Kerben im Planfilm, identifizieren, bei Glasnegativen entfällt dieses Merkmal jedoch.

Um den individuellen Qualitäten eines durchmischten Bestands umfassend gerecht zu werden, kann eine Auflösung von 80 Linien/mm als Ausgangspunkt zur Festlegung der Scangrößen gewählt werden. Das entspricht der Auflösung eines modernen feinkörnigen Diafilms. Dadurch wird gewährleistet, dass alle Details sowohl historischer Aufnahmen als auch Aufnahmen neueren Datums im Digitalisat wiedergegeben werden. Die Auflösungswerte moderner Filme sind größtenteils in den Datenblättern der Hersteller dokumentiert. Eine Scanauflösung von 4000 dpi bildet die Details also ausreichend ab.[35] Auf das Filmmaterial bezogen wäre die theoretisch erzielbare Auflösung für alle Aufnahmeformate konstant.

Für Kleinbildaufnahmen kann eine Scanauflösung von 4000 dpi gewählt werden, da in der Regel die Abbildungsleistung von Kleinbildoptiken besser ist als die von Mittelformat- oder Großbildobjektiven. Für größere Aufnahmeformate kann die Scanauflösung daher verringert werden. Einen entsprechenden Hinweis liefert der maximale Zerstreuungskreisdurchmesser eines Bildpunkts, der aufnahmeformatabhängig vom menschlichen Auge noch als scharf

[35] Um die Filmauflösung in Linien/mm digital darstellen zu können, sind pro Linie mindestens zwei Pixel notwendig. Damit entsprechen 80 Linien/mm 160 Pixel/mm oder 1.600 Pixel/cm. Dieser Wert multipliziert mit 2,54 (1 inch = 2,54 cm) ergibt eine Scanauflösung von 4.064 dpi, gerundet 4.000 dpi.

wahrgenommen wird. Fotografische Praxisregel ist hierbei, dass die Formatdiagonale in mm (entspricht Normalbrennweite) mit 1/1500 multipliziert wird. So ergeben sich folgende maximal zulässige Zerstreuungskreise je Aufnahmeformat:

Kleinbild: 0,03 mm; Mittelformat: 0,05 mm; 9 x 12 cm: 0,1 mm; 18 x 24 cm: 0,2 mm

Daraus abgeleitet ergeben sich folgende Zielgrößen: [36]
Mittelformat: 4.000 dpi x 0,03/0,05 = 2.400 dpi
9 x 12: 4.000 dpi x 0,03/0,1 = 1.200 dpi
18 x 24: 4.000 dpi x 0,03/0,2 = 600 dpi

Die Fotografien werden immer vollständig mit leichtem umlaufendem Rand gesichert, um deutlich zu machen, dass nichts von der Vorlage abgeschnitten wurde.

Aufsichtmedien

Das Fotopositiv ist ein Endprodukt. Es dient anders als das fotografische Negativ nicht als Reproduktionsvorlage oder Ausgangspunkt für eine Vergrößerung. Die Bildqualität ist von der Negativqualität abhängig. Anders als bei Negativfilmen gibt es zum Auflösungsvermögen von Fotopapieren (gemessen in Linienpaaren pro mm, L/mm) keine verlässlichen Angaben. Das Auflösungsvermögen der fotografischen Emulsion im Positiv ist jedoch ausreichend für eine scharfe Bildwiedergabe bei Betrachtung mit bloßem Auge. Als Untergrenze sollte daher eine Scanauflösung für die Digitalisate gewählt werden, die bei Ausgabe für den Druck oder auf Fotopapier die Bildqualität der Vorlage 1:1 wiedergeben kann.

Wie unter 3.2.1.1 beschrieben, liegt die empfohlene Mindestauflösung für die Digitalisierung bei 300 dpi unter Berücksichtigung der Ergebnisse einer Testdigitalisierung mit genormter Testtafel.

Häufig wird jedoch bei kleinen Vorlagen diese Auflösung nicht ausreichen, um die Eigenschaften der Vorlage umfassend erkennbar wiederzugeben. In der ersten Hälfte des 20. Jahrhunderts war es gebräuchliche Praxis auch von 6 x 9 cm Negativen nur Kontaktkopien zu erstellen. [37]

Umgekehrt kann bei überformatigen Vorlagen (DIN A0 und größer) die Auflösung bezogen auf das Vorlagenformat reduziert werden, wenn die Objekte für einen großen Betrachtungsabstand entworfen wurden. Großvergrößerungen sind hier als Beispiel zu nennen, sie sind auf Fernsicht ausgelegt. Bei zu nahem Betrachtungsabstand nimmt man das Filmkorn deutlich wahr. Hier kann die Auflösung bis auf 150 dpi reduziert werden, da mit wachsendem Betrachtungstand auch die Punktgröße wächst, die vom menschlichen Auge nicht mehr unterschieden werden kann.

Soll allerdings bei Mappenwerken mit historischen Fotopositiven die Konfektionierung inklusive Präsentationskarton dokumentiert werden, ist von einer verringerten Auflösung abzusehen,

---

[36] Diese Werte gibt auch die FADGI vor. Vgl. Federal Agencies Digitization Guidelines Initiative (FADGI): Technical Guidelines for Digitizing Cultural Heritage Materials: Creation of Raster Image Master Files, August 2010: http://www.digitizationguidelines.gov/guidelines/FADGI_Still_Image-Tech_Guidelines_2010-08-24.pdf, S. 60.
[37] So empfiehlt die FADGI für Fotoaufsichtmedien im Format 4 x 5 inch eine Auflösung von 800 dpi.Vgl. Federal Agencies Digitization Guidelines Initiative (FADGI): Technical Guidelines for Digitizing Cultural Heritage Materials: Creation of Raster Image Master Files, August 2010: http://www.digitizationguidelines.gov/guidelines/FADGI_Still_Image-Tech_Guidelines_2010-08-24.pdf, S. 62.

damit Bildunterschriften und Aufnahmedetails identifizierbar bleiben, da bei großformatigen Mappenwerken der Präsentationskarton durchaus 50 % oder mehr der Gesamtfläche des Mediums einnehmen kann.

Das Speichern von Farbinformationen bei SW-Vorlagen ist nur dann zu vertreten, wenn wichtige Informationen zu dem Medium in Farbe vorliegen, wie z. B. getonte Abzüge.

Aufgrund des großen Mehrwerts für die Benutzung sollte in einem ökonomisch vertretbaren Rahmen die volle Auflösung der eingesetzten Aufnahmetechnik bei den Fotografien – Durchsicht und Aufsicht – genutzt werden.

Die Fotografien werden immer vollständig mit leichtem umlaufendem Rand gesichert, um zu zeigen, dass nichts von der Vorlage abgeschnitten wurde.

3.2.2.4 Mikroformen

Für Mikroformen (Mikrofilme 16 und 35 mm, Mikrofiche 105x148 mm, positiv und negativ, S/W und Farbe) gelten bezüglich der Parameter die Hinweise, die im vorletzten Abschnitt für Durchsichtmedien gegeben wurden. Hierbei ist aber für die Auflösung abweichend zu berücksichtigen, dass es sich um Sicherungs- oder Schutzreprografien von Originalen handelt. Die Verfilmung kann mit einem Verkleinerungsfaktor zwischen 1:7,5 und 1:96 stattfinden. Die Auflösung bei der Digitalisierung muss sich also soweit möglich am Original orientieren und nicht am Film.

Die Digitalisierung von Mikrofilmen wird häufig als Massendigitalisierung, die mit niedrigen Kosten realisierbar ist, eingesetzt. Dabei lassen sich die angestrebten 300 dpi in Bezug zur Vorlagengröße nur in den seltensten Fällen realisieren, auch wenn die Auflösung des Films dafür theoretisch ausreicht. Der limitierende Faktor sind hierbei die zurzeit marktüblichen Scanner, die in der Lage sind, teilautomatisch in kurzer Zeit ganze Filme zu digitalisieren. Für S/W-Mikrofilme und Mikrofiche können Auflösungen bis 600 dpi bezogen auf den Film erreicht werden. Die Wahl der Auflösung muss sich also am technisch Machbaren der Massenverfahren orientieren. Nur in Ausnahmefällen sollte auf Reproduktionsverfahren zurückgegriffen werden, die Einzelscans mit höherer Auflösung ermöglichen.

Liegen filmische Mikroformen vor, die nach den Standards der Bundessicherungsverfilmung erstellt wurden oder diesen entsprechen (vgl. Kap. 2.3), und dienen diese als Grundlage für eine Digitalisierung, ist zu prüfen, ob die Erstellung und Erhaltung eines Masters sinnvoll ist. Die Mikroverfilmung selbst ist in diesem Fall Langzeitsicherungsmedium und Kopiermaster. Jedoch muss immer geprüft werden, ob die Erstellung von Digitalisaten (Master und Nutzungsversionen) vom Original sinnvoller ist, als die Erzeugung von Benutzungsdigitalisaten vom Mikrofilm.

3.2.2.5 Dreidimensionale Objekte

Bedingt durch den informationstechnologischen Fortschritt der letzten Jahre gewinnt die digitale Erfassung und Vermittlung von 3D-Objekten im Kulturerbebereich laufend an Bedeutung. Die **digitale 3D-Dokumentation** erfasst dabei in erster Linie museale Sammlungsbestände sowie die Architektur und ihre Ausstattung.

Erfolgte diese vor kurzem noch hauptsächlich durch fotografische Abbildung aus verschiedenen Aufnahmestandpunkten, so kann heute auf eine bewährte Auswahl an 3D-Erfassungs-

und Rekonstruktionsmethoden sowie diesen Prozessen beistehende Technologien zurückgegriffen werden.

Im Gegensatz zur Fotografie wird bei der 3D-Dokumentation die gesamte Geometrie eines Objekts, seine Oberflächentextur und nach Möglichkeit seine optischen Materialeigenschaften erfasst, integriert und im digitalen 3D-Modell zusammenzugeführt.

Ein entscheidender Vorteil gegenüber einer Fotografie ist, dass hierbei die Form und die Oberflächen-Licht-Interaktion des Objekts originalgetreu erfasst und wiedergegeben werden können. Das 3D-Modell kann aus beliebigen Perspektiven in der originalen Lichtsituation sowie in beliebigen neuen Lichtsituationen und Umgebungen visualisiert und simuliert werden. Eine *open source* Anwendung zur interaktiven Web-Präsentation von hochaufgelösten 3D-Modellen bietet u. a. “3D Heritage Online Presenter”[38] (3DHOP).

Bei der digitalen 3D-Dokumentation kann zwischen einer **Retrodigitalisierung** infolge der Überführung eines physisch vorhandenen Objekts in ein Digitalisat und einer **Rekonstruktion** eines physisch nicht (mehr) vorhandenen Objekts in Form eines nativ digitalen 3D-Datensatzes unterschieden werden.

Das Ziel der Retrodigitalisierung und der digitalen Rekonstruktion dient vorrangig der Erschließung, Dokumentation, Sicherung und Archivierung sowie der web-basierten Zurverfügungstellung von Kulturerbe für die Forschung und die Öffentlichkeit. Die digitalen 3D-Modelle zeichnen sich durch folgende Vorteile aus:

- praktisch beliebige und parallele Verfügbarkeit (auch in unterschiedlicher Qualität) digitaler 3D-Modelle von Kulturobjekten für Wissenschaftlerinnen und Wissenschaftler
- Ergänzung fehlender (u. a. zerstörter) Bereiche infolge der Rekonstruktion
- Simulation verschiedener hypothetischer Versionen und/oder Varianten eines 3D-Objekts
- Einsatz von digitalen 3D-Modellen im Museumsbetrieb, z. B. zur Ausstellungsplanung, Dokumentation, Beschaffungsplanung, etc.
- virtuelle Präsentation und Ausstellung (in Kombination mit neuen Präsentationstechniken, wie z. B. hybriden Exponaten) als Mittel zum Wissenstransfer und Attraktivitätssteigerung
- Referenz für die Restauration beschädigter Originale bzw. generierungphysischer Repliken auf Basis des digitalen 3D-Modells
- Ersatz für Ausleihe (Vermeidung von Beschädigung, Versicherungskosten, Rechtsunsicherheit zu Eigentum)

Auf der anderen Seite stellen uns digitale 3D-Modelle vor neue Herausforderungen. Bislang haben sich keine Standards im Bereich der softwareunabhängigen Datenformate und der Dokumentation der Digitalisierung/Modellierung und der Resultate durchgesetzt, so dass in diesem Kontext von fehlender Interoperabilität die Rede sein kann. Darüber hinaus verzeichnen wir einen Mangel an adäquater digitaler Forschungs- und Informationsinfrastruktur, die eine Nachhaltigkeit der 3D-Datensätze infrage stellt. Daraus ergibt sich ein großer Handlungsbe-

[38] http://3dhop.net/index.php

darf, der zur Zeit u. a. durch den im November 2014 konstituierten “Rat für Informationsinfrastrukturen”[39] (RfII) definiert wird.

Für die **Retrodigitalisierung** von 3D-Objekten stehen verschiedene Technologien zur Auswahl, die je nach Objektart und Eigenschaften (Größe, Material, Zustand, etc.) zu wählen sind. Gegenwärtig wird in dem Projekt “Colour and Space in Cultural Heritage”[40] (COSCH) versucht einen Online-Ratgeber für die Wahl einer adäquaten Methodik und Technologie je nach Problemstellung zu entwickeln. Zielführend ist jedoch weiterhin der frühe Kontakt mit führenden Technologieentwicklern und Anwendern im Bereich der Massendigitalisierung, wie dem “CultLab3D”[41].

Es gibt verschiedene optische Digitalisierungsverfahren.

Laserscanverfahren werden aufgeteilt in Lauflängenmessung und Triangulierung. Im ersten Fall wird ermittelt, wie lange ein Laserpuls zum Objekt und wieder zurück benötigt, um die Entfernung zu diesem Punkt zu messen. Im zweiten Fall wird mittels einer Kamera aus der Verzerrung der projizierten Laserlinie auf der Oberfläche eines Objekts auf ihren Verlauf in 3D geschlossen.

Streifenlichtverfahren setzen auf die Projektion definierter Muster (meist parallele Linien) auf Objektoberflächen und schließen ebenfalls aus ihrer Verzerrung auf die zugrundeliegende 3D Oberflächengeometrie. Zur Erhöhung der Auflösung können diese Muster über die Oberflächen phasenverschoben werden.

Bei photogrammetrischen Verfahren werden aus verschiedenen Perspektiven Bilder des zu digitalisierenden Objekts angefertigt und nach Merkmalspaaren auf mindestens zwei sich überlagernden Aufnahmen durchsucht, so dass sich für diese mittels Triangulierung die Tiefe des zugehörigen Bildpunktes ergibt. Charakteristische Merkmale können dabei beispielsweise Intensitäts- oder Kontrastunterschiede sein.

Das Prinzip bei Time-of-Flight Kameras entspricht der Lauflängenmessung bei Laserscannern, mit dem Unterschied, dass eine ganze Szene auf einmal beleuchtet, aufgenommen und für jedes Pixel des Kamerasensors eine Tiefe berechnet wird.

Ab bestimmten Objektgrößen, wie etwa größere Statuen oder gar Gebäude, ist eine punktweise Vermessung wie etwa mit Laserverfahren oder Mittels Oberflächenkodierung nicht mehr effizient und photogrammetrische Verfahren zeigen ihre Stärken.

Bei Retrodigitalisierung existiert ein generelles Spannungsverhältnis zwischen Objektgröße und Auflösung, bzw. genauer zwischen dem Messbereich eines Sensors und der Genauigkeit. Je größer das Objekt bzw. je größer die Entfernung, desto gröber werden Details erfasst.

Die Geometrie und Textur werden üblicherweise im selben Arbeitsschritt der Geometrieakquise erfasst, weil eine direkte Korrespondenz zwischen Messpunkten und Punkten des Bildgebers vorliegt und ein Großteil der Akquise-Methoden einen Kamerasensor als integralen Be-

[39] http://www.rfii.de/de/index/
[40] http://cosch.info/
[41] http://cosch.info/

standteil der Erfassung verwenden, um Tiefeninformationen zu extrahieren. Dieser hat pro erfasstem Ausschnitt des Objekts meist bereits die ideale Ausrichtung auf das Objekt, so dass dieselbe Pose des Kamerasensors für die Texturgewinnung genutzt werden kann.

Für die Erfassung der Textur ist eine ausreichende, diffuse und gleichmäßige Beleuchtung erforderlich. Wichtig ist auch ein passender Weißabgleich, da die meist künstliche diffuse Beleuchtung der Objektoberfläche kompensiert werden muss. Eine Farb-Kalibrierung ist hierbei wichtig, da Textur einen wesentlichen Einfluss auf die Realitätsnähe hat.

Zusätzlich zu Geometrie und Textur lassen sich auch optische Materialeigenschaften erfassen. Darunter versteht man beispielsweise Glanz- und Reflexionseigenschaften. Typischerweise geschieht dies in dem man jede Kombination aus Lichteinfall und Perspektive auf ein Objekt aufnimmt, um im Anschluss die für jeden Oberflächenpunkt (Sichtbarkeitsprüfung) relevante Texturen abzuspeichern, so dass bei späterer interaktiver Betrachtung des virtuellen Objektes unter jeglicher Beleuchtung die physikalisch korrekten Texturen appliziert werden und damit die originale Materialanmutung und Licht-Oberflächeninteraktion erreicht wird. In manchen Fällen ist eine komplette Materialakquise eines Objekts nicht möglich, entweder bedingt durch seine Größe, oder weil nur Materialreste erhalten sind. In diesen Fällen bietet sich eine punktuelle Erfassung von Materialeigenschaften und eine anschließende Extrapolation auf Regionen gleichen oder ähnlichen Materials an.

Die **hypothetische Rekonstruktion** nicht vorhandener (u. a. zerstörter) 3D-Objekte basiert auf der Erfassung und Interpretation vom historischen Quellenmaterial. Die computergestützte Rekonstruktion bedient sich der gängigen 3D-Simulations- und Animationssoftware (u. a. Autodesk 3ds Max und Maya, Cinema 4D, SketchUp, Blender, etc.) bei der Erstellung der 3D-Modelle. Als komplementäre Methode zur Retrodigitalisierung bietet uns die Rekonstruktion die Möglichkeit nicht mehr Vorhandenes zu dokumentieren und visualisieren. Bezeichnend ist dabei die Problematik der Darstellung von Wissensunschärfen und Hypothesen. Hinzu gesellen sich dringend geforderte Dokumentationsstandards für die interpretative Modellierung und Visualisierung, die u. a. eine Zuweisung der Quellen und die Abbildung der kreativ-interpretativen Prozesse beinhalten sollen.

“Die digitale Erschließung […] zielt auf die Integration und Konvergenz zwischen digitalisierten und nativ digitalen Daten in einheitlichen, integrierten Arbeitsumgebungen mit dem Ziel dynamischer Wissensintegration”[42], wie der Rat für Informationsinfrastruktur bei der Erläuterung zur Informationsinfrastruktur feststellt. Einen wichtigen Beitrag zur Integration und Konvergenz leisten die strukturierten Daten, welche sich an einschlägigen Metadata-Schemata (z. B. LIDO und CARARE 2.0), kontrollierten Vokabularen (z. B. Getty Vocabularies) und Normdateien (z. B. GND) sowie Referenzontologien (allen voran CIDOC (Documentation International Committee for Documentation) CRM) richten. Die Erweiterung von CIDOC CRM zur besseren Erfassung der Retrodigitalisierung fand innerhalb von CIDOC CRM_{dig} ihren Ausdruck. Die Vorstellung von CIDOC CRM referenzierten Applikationsontologien zur Wissensrepräsentation einer computer-gestützten Rekonstruktion werden gerade in der Fachgemeinschaft vorgestellt und diskutiert.[43]

[42] RfII Berichte No. 1: Begriffsklärungen (2016), S. 13, auf http://www.rfii.de/de/category/dokumente/
[43] http://www.digitale-rekonstruktion.info/

3.3 Metadaten

Die Erzeugung von Metadaten, welche erst die Auffindbarkeit der Objekte gewährleisten und eine kontextualisierende Präsentation ihrer digitalen Images erlauben, ist zentraler Bestandteil der Digitalisierung. Die DFG geht davon aus, dass die der Digitalisierung zu Grunde liegenden analogen Objekte bereits primär in anerkannten digitalen Nachweissystemen erschlossen sind bzw. mit der Digitalisierung einhergehend erschlossen werden. Metadaten, die im Rahmen des Digitalisierungsprojekts entstehen, sind grundsätzlich in einer **von der Software unabhängigen** und **standardkonformen Form** bereitzustellen, in aller Regel in einer **XML-Kodierung**. Dies ist in den Workflow des Projekts so einzubetten, dass ein vollständiger Metadatensatz in software-unabhängiger Form auch dann bereitsteht, wenn das Projekt – aus welchen Gründen auch immer – zu einem vorzeitigen Ende kommt.

Wenn sich die im Rahmen eines geförderten Digitalisierungsprojekts ausgewählten Objekte sachlich für die Einbindung in ein material- oder fachspezifisches Portal eignen, wird erwartet, dass ein Projektantrag entweder erläutert, welche Vorkehrungen projektseitig getroffen werden, um die Anbindung an dieses Portal während und nach der Projektlaufzeit sicherzustellen, oder plausibel macht, warum eine Anbindung aus inhaltlichen Gründen oder aus Gründen des damit verbundenen Aufwands nicht notwendig bzw. sinnvoll ist.

Allgemein werden deskriptive (bibliografische Beschreibung, archivische Erschließung, Beschreibung von unikalen, häufig zudem nicht-textuellen Objekten), strukturelle (Text-, Dokumentstruktur), administrative (z. B. Rechteverwaltung) und technische (z. B. Dateitypen) Metadaten unterschieden. Die folgenden Überlegungen beziehen sich ausschließlich auf deskriptive und strukturelle Metadaten.

Die Verknüpfung zwischen den Metadaten einerseits und den digitalen Images andererseits zu einem Objekt muss dabei immer auf der Ebene der Metadaten gewährleistet sein. Zusätzlich können Metadaten auch in den Header der digitalen Images eingebettet werden, jedoch werden diese von den Software-Produkten unterschiedlich dargestellt und im schlimmsten Fall sogar korrumpiert, so dass die Einbettung in jedem Fall nur eine ergänzende Option ist.

Die Metadatenformate realisieren die Referenzierung der digitalen Images auf unterschiedliche Weise. Mit dem Containerformat **METS** können sowohl deskriptive Metadaten über Mappings in Standardformaten wie z. B., MODS oder TEI als auch strukturelle Metadaten inklusive der Referenzen auf die digitalen Images transportiert werden. Dies bietet sich insbesondere für Volldigitalisate von Textwerken an, die sowohl mit deskriptiven als auch strukturellen Metadaten erschlossen sind. Andere Formate wie **LIDO** enthalten zusätzlich zu den semantischen Elementen zur Objektbeschreibung eigene Elemente zur Referenzierung beliebiger digitaler Ressourcen zum Objekt (neben Images auch Audio- oder Videodaten). Auch hier können beliebig viele Ressourcen mit einem Objekt verknüpft und noch mit eigenen Beschreibungselementen versehen sein. Auch in **EAD(DDB)** ist eine Verknüpfung zu beliebig vielen digitalen Ressourcen einschließlich der Einbindung von METS-Containern möglich. Wichtig ist, dass die digitalen Ressourcen innerhalb der Metadaten über global eindeutige, persistente Adressen, in der Regel URLs, referenziert werden (vgl. persistente Adressierung in Kapitel 5).

Der Nachweis der Digitalisate und Metadaten muss bei Bibliotheksgut entweder durch **Katalogisierung der elektronischen Ausgabe** oder durch Angabe der **PURL der Bilddateien** bzw. **Angabe einer persistenten Verknüpfung im Katalog** (OPAC, Verbundsystem) erreicht

werden. Von Antragstellerinnen und Antragstellern aus universitären Einrichtungen wird erwartet, dass sie die Frage der Katalogisierung mit ihren örtlichen Bibliotheken abstimmen bzw. durch diese durchführen lassen.

Der Nachweis digitalisierter Drucke im **ZVDD** wird erwartet. Die Digitalisate anderer bibliothekarischer Objekte sind in einschlägigen materialspezifischen Portalen nachzuweisen (z. B. Manuscripta Mediaevalia für mittelalterliche Handschriften[44], Kalliope für Nachlässe[45]). Materialien, die nicht in bibliothekarische Nachweisinstrumente eingebracht werden können, sollten in **geeigneten fachlichen oder fachübergreifenden Online-Anwendungen** präsentiert werden. Der Nachweis der Digitalisate und Metadaten muss bei Archivgut im **Archivportal-D** bzw. in der DDB erfolgen. Die Einbringung der Daten in die **DDB** und über die DDB in die **Europeana** wird von allen Projekten erwartet.

Es obliegt dem Projektnehmer sicherzustellen, dass die im Projekt hergestellten Digitalisierungseinheiten eindeutig identifiziert und von anderen im selben System vorhandenen Einheiten getrennt durchsucht und aufgerufen werden können.

Die Ablieferung an diese Portale sollte gemäß den Standardformaten wenn möglich über **OAI** erfolgen (vgl. Kapitel 3.2.1.4).

3.3.1 Erschließung, deskriptive Metadaten

Die Erschließung eines zu digitalisierenden Objekts mit deskriptiven Metadaten erfüllt je nach Material unterschiedliche Funktionen für die Bereitstellung. Zunächst ermöglichen es die deskriptiven Metadaten überhaupt erst, die Inhalte gezielt durch eine Recherche aufzufinden. Das Objekt wird klassifiziert, historisch kontextualisiert und gegebenenfalls in seiner Materialität beschrieben. Im Idealfall bieten die deskriptiven Metadaten Anknüpfungspunkte für unterschiedliche Fragestellungen und Fachdisziplinen. Digitalisierung ohne Nachweis von Metadaten nach den gängigen Community-Standards ist nicht förderfähig.

Sofern in einem Digitalisierungsprojekt über bereits vorliegende deskriptive Metadaten hinaus eine projektspezifische Erschließung und/oder Informationsbereitstellung geplant ist – und dies wird zumindest im Bereich unikaler Objekte der Regelfall sein – sind die wichtigsten Fragestellungen an das Material zu antizipieren und in einem Kernfeldkatalog zu formulieren. Die systematische Bearbeitung der Metadaten auf der Grundlage eines Kernfeldkatalogs ist Voraussetzung für eine optimale Bereitstellung der digitalisierten Bestände gemäß Kap. 6.2.

Für eine optimale verteilte, auch nachhaltige Nutzbarkeit der Metadaten ist die Erschließung an den **einschlägigen Spartenstandards (z. B. RDA[46], ISAD (G)[47])** und **Referenzmodellen** (CIDOC-CRM[48], perspektivisch auch IFLA FRBR/FRBRoo[49]) zu orientieren und wo immer möglich mit **publizierten Normdaten** zu verknüpfen. Ausdrücklich wird für die Erfassung personenbezogener, biografischer Information und für den geografischen Nachweis die Verwendung der **Gemeinsamen Normdatei (GND)** der Deutschen Nationalbibliothek empfohlen.

---

[44] http://www.manuscripta-mediaevalia.de/
[45] http://kalliope.staatsbibliothek-berlin.de/de/index.html.
[46] https://wiki.dnb.de/display/RDAINFO/RDA-Info.
[47] http://www.ica.org/sites/default/files/CBPS_2000_Guidelines_ISAD%28G%29_Second-edition_DE.pdf.
[48] CIDOC Conceptual Reference Model: http://cidoc-crm.org/
[49] IFLA Functional Requirements for Bibliographic Records: http://www.ifla.org/en/publications/functional-requirements-for-bibliographic-records

Darüber hinaus eingesetzte kontrollierte Vokabulare, z. B. **Iconclass**[50] **zur Bilderschließung**, sollten national und international anschlussfähig sein.

Die Bereitstellung der Metadaten zur weiteren Nutzung gemäß den **materialspezifischen Standards** ist verpflichtend: **METS/MODS** für gedruckte Textwerke (s. Anhang A), **METS/TEI** für Handschriften (s. Anhang B), **EAD(DDB)**[51] **in Verbindung mit einem METS/MODS-Mapping** für Archivmaterial (s. Anhang A), **LIDO** für (i.d.R. unikale) bildhafte und dreidimensionale Objekte (s. Anhang C). Die bereitgestellten Metadaten müssen gegen das jeweilige XML Schema valide sein und sind darüber hinaus auf semantische Korrektheit zu überprüfen.

3.3.2 Strukturelle Metadaten für digitale Faksimiles

Wohl zu erwägen ist die Frage der Anwendung von strukturellen Metadaten zur Erschließung von Images, also der Kodierung der strukturellen Elemente eines Dokumentes, wie z. B. Widmung, Vorwort, Kapitel oder Illustration. Die Aufnahme dieser Aspekte nimmt den Gedanken der analytischen Bibliografie auf, die an den Kapitel- und Textstrukturen entlang den Inhalt eines Werkes referiert. In manchen Fällen ist die Erstellung von strukturellen Metadaten von eher nachgeordneter Bedeutung, in anderen Fällen ist die Erstellung eines solchen künstlichen Inhaltsverzeichnisses für die Navigation im Digitalisat unverzichtbar. Beispielsweise ist es keinem Nutzer zuzumuten, in einem 600 Seiten umfassenden digitalen Wörterbuch online zu blättern, um den richtigen Alphabeteinstieg zu finden. Bei manchen Fragestellungen sind strukturelle Metadaten auch für die Recherche interessant. Die Entscheidung über die Erstellung struktureller Metadaten ist daher immer eine **material-** und **projektspezifische**.

Falls strukturelle Metadaten Verwendung finden, wird die Konsultation der über die Website des **DFG-Viewers** zugänglichen **Liste von Bezeichnungen** empfohlen. Sollten darüber hinaus noch andere Bezeichnungen erforderlich sein, sollte man sich im Rahmen konkreter Digitalisierungsprojekte über standardisierte Bezeichnungen verständigen und diese meist fach- oder sachbezogenen Vokabularien über die Projektwebsite und gegebenenfalls auch die Webseite des DFG-Viewers publizieren, um eine umfassende Nachnutzung zu ermöglichen.

Bei der Vergabe von strukturellen Metadaten stellt sich die Frage, ob man sich bei der Erschließung der Dokumente eher am digitalen Faksimile, der physikalischen Seitenfolge, oder an der Text- bzw. Kapitelstruktur des Werks orientiert. Geht man davon aus, dass dem digitalen Faksimile eines alten Drucks oder einer Handschrift die Transkription oder Edition beigegeben werden soll, ist eine Orientierung am **TEI-Kodierungsstandard** zu empfehlen. Bei einer Seitenbeschreibung mit einigen qualifizierenden Merkmalen (z. B. Illustrationen oder Annotationen) ist eher der von der Library of Congress gepflegte **METS** zu empfehlen. Für beide Modelle, die seitenorientierte und die dokumentorientierte, gibt es gute Argumente. Meist lassen sich beide Ansätze auch ineinander überführen. Tendenziell ist eine Struktur, die der Logik des Textes folgt, leistungsfähiger, was spätere Abfragemöglichkeiten und Repräsentation der Vorlage anlangt. Allerdings wird dieser Vorteil durch einen höheren technischen Aufwand bei der Bearbeitung und Präsentation der Dokumente erkauft. Es sei darauf hingewiesen, dass auch eine auf die physikalische Seitenfolge bezogene Kodierung, wie sie in Bibliotheken

[50] http://www.iconclass.nl/
[51] http://www.landesarchiv-bw.de/web/53401.

eher anzutreffen ist, die Verwendung von TEI nicht ausschließt,[52] so dass sich beide Aspekte gegebenenfalls auch sinnvoll verbinden lassen.

Empfohlen wird nach dem derzeitigen Stand bei alten Drucken eine Orientierung an **METS** oder **TEI**. Dessen ungeachtet soll in jedem Fall der **DFG-Viewer** unterstützt werden, der auf METS beruht. Wenn daher TEI für Strukturdaten zum Einsatz kommt, ist im Rahmen des Projekts eine Konversion zu METS notwendig. Da die Basis in beiden Fällen XML ist und die beschriebenen Sachverhalte ähnlich, kann man davon ausgehen, dass diese Konversion keine größere Hürde darstellt.

3.3.3 Sammlungs- und Bestandsbeschreibung

Sammlungs- und Bestandsbeschreibungen kontextualisieren und verorten das Einzelobjekt und ermöglichen dem Nutzer/der Nutzerin so einen Überblick über den Gesamtbestand einer Einrichtung. Grundlage für ein Digitalisierungsprojekt können auch virtuelle Bestände- und Sammlungen (z. B. sachthematisch, materialspezifisch) sein.

Es wird erwartet, dass Digitalisierungsprojekte **mindestens Gegenstand und Umfang der jeweiligen Material- bzw. Objektauswahl auf einer Seite im Netz**, möglichst **auch in Englisch**, darstellen. Sinnvoll dafür ist eine **normierte Beschreibung in XML**, um diese Informationen in Zukunft besser in nationalen oder internationalen Portalen zusammenführen und gezielt recherchieren zu können. Diese Beschreibung kann gemäß dem **Dublin Core Collections Application Profile** (s. Anhang D) oder in demselben Metadatenstandard erfolgen, in dem auch die Objektbeschreibungen verfügbar gemacht werden: **METS**, **MODS** oder **TEI-Header**, **EAD(DDB)** sowie **LIDO** bieten entsprechende Möglichkeiten. Darüber hinaus sollte für eine Sammlung die eindeutige Identifizierung und Beschreibung gemäß ISO 27730, Information and Documentation – International Standard Collection Identifier (ISCI) erwogen werden, die auf der ISIL einer Institution aufbaut.

3.3.4 Austausch und Weitergabe der Metadaten

Für den Aufbau übergreifender Nachweissysteme ist die Schaffung eines **Standards zum Austausch von Metadaten** für Digitalisate von zentraler Bedeutung (vgl. Kapitel 6). Allerdings können sich Standards nur jeweils innerhalb der jeweiligen Community ausbilden und etablieren. Dabei können ein und dieselben Ressourcen durchaus im Horizont ganz verschiedener Fragestellungen erscheinen und entsprechend divergierende Sets von Metadaten erfordern. Ein generalisiertes Verfahren zum Austausch von Metadaten muss daher flexibel unterschiedliche Metadatenformate bzw. communityabhängige Spezifikationen verwalten können. Dieses lässt sich mit dem Protokoll der **OAI** gut erreichen. Die Verwendung von OAI ist mit Blick auf alte Drucke und Handschriften vor allem als technisches Austauschprotokoll sinnvoll. OAI schreibt vor, dass mindestens **Dublin-Core-Daten** geliefert werden müssen; das ist zwar für die deskriptive Beschreibung sowohl alter Drucke und Handschriften[53] als auch nichttextueller Objekte[54] ungenügend, als zusätzliche Information aber von Nutzen. Der OAI-Standard sieht die **parallele Unterstützung weiterer Metadatenformate** explizit vor; eine Verbindung von OAI mit allen XML-basierten Metadatenformaten ist daher grundsätzlich möglich

[52] http://www.tei-c.org/Sample_Manuals/bestpractice.htm

[53] Vgl. Hillmann, Diane I.: Choices: MARC or Dublin Core? In: Anne R. Kenny/Oya Rieger (Hrsg.): Moving Theory into Practice. Digital Imaging for Libraries and Archives. Mountain View: Research Library Group 2000, S. 89f.

[54] Die unzulässig vereinfachende Beschreibung musealer Objekte in Dublin Core und deren Darstellung in Online-Umgebungen war der zentrale Anlass für die Museums-Community zur Entwicklung des Harvestingformats LIDO.

(METS/MODS mit spartenspezifischen Mappings, METS/TEI-HSS, MARCXML, MABxml, EAD(DDB), TEI P5, LIDO etc.).

Die DFG setzt verpflichtend die Bereitstellung von Metadaten für Nutzer/-innen über **OAI** voraus. Eine solche kann auch über ein übergreifendes Portal (z. B: ZVDD, DDB, Archivportal–D) erfolgen. Dabei müssen neben den von OAI vorgeschriebenen **Dublin-Core-Metadaten** materialspezifische Metadaten nach **METS/MODS** für alte Drucke, METS/MODS für Archivgut mit einem für den DFG-Viewer angepassten Mapping, **METS/TEI** für mittelalterliche Handschriften und **LIDO** für bildhafte und dreidimensionale Materialen ausgeliefert werden (vgl. auch Kapitel 7 und Anhang A). Eine Abweichung von dieser Verpflichtung muss im Antrag ausdrücklich begründet werden.

3.4 Volltextgenerierung

Eine vielseitige wissenschaftliche Nachnutzbarkeit, die unter anderem automatisierte Recherchen, quantitative Auswertungen im Rahmen von Text- oder Datamining, semantische Analysen, Mustererkennungen in nicht textuellen Materialien, Anreicherungen, Kontextualisierungen und Weiterverarbeitungen – auch im Rahmen von virtuellen Forschungsumgebungen – ermöglicht, basiert auf der leichten und ungehinderten Nachnutzbarkeit der entsprechenden Daten, entsprechenden Rechtseinräumungen und dem Angebot des digitalen Volltextes. Wo immer möglich und sinnvoll, wird die DFG-Förderung daher im textuellen Bereich auf die Bereitstellung maschinenlesbarer Volltexte abzielen. Von allen Anträgen zur Digitalisierung textueller Materialen wird daher eine **Auseinandersetzung mit Möglichkeiten der Volltextbereitstellung** erwartet. In diesem Zusammenhang wird auf das Koordinierungsprojekt OCR-D[55] verwiesen, das Verfahren zur OCR-gestützten Volltexterfassung prüft und prototypische Workflows entwickelt. Ergebnisse dieses Vorhabens sollen in die nächste Version der Praxisregeln einfließen. Unabhängig von diesen Entwicklungen gilt für **Druckwerke ab Erscheinungsjahr 1850** verpflichtend, dass **Volltext hergestellt werden muss** und eine bloße Bilddigitalisierung nicht ausreicht.

Die Herstellung von Volltext kann auf mehreren Wegen erfolgen: entweder durch **OCR** oder durch **Abschreiben bzw. Transkription**. Die Frage, welches Verfahren gewählt werden soll, ist u. a. abhängig vom Alter und Zustand der Vorlage und der notwendigen Fehlerfreiheit des Textes für die angestrebten Ziele.

3.4.1 Texterfassung

Textgenauigkeiten

Unabhängig davon, ob man einen Text durch OCR oder manuelle Transkription herstellt, ist die Frage zu beantworten, welche Qualität für welche Zwecke benötigt wird und welche Kosten für das jeweilige Ziel angemessen sind. Je nach Projekt werden die Anforderungen an die Textgüte variieren. Ein Editionsprojekt wird allerhöchste Maßstäbe anlegen, während ein Massendigitalisierungsprojekt, das Tausende von Titeln umfasst, auf eine vergleichsweise teure manuelle Transkription verzichten muss. Was dabei noch als ausreichend gilt und welche Kosten für welche Qualität akzeptabel sind, muss in Projekten fallweise, materialabhängig und nach Maßgabe der wissenschaftlichen Anforderungen sorgfältig begründet werden. Gewöhn-

[55] OCR-D: Koordinierungsprojekt zur Weiterentwicklung von Verfahren der Optical Character Recognition (http://www.ocr-d.de/)

lich wird die Qualität von mit OCR bearbeiteten Texten in Prozent angegeben. Dabei herrscht wenig Einigkeit darüber, welche Messkriterien und -verfahren angelegt werden. Genauigkeiten können sich auf die Richtigkeit der Buchstaben, aber auch auf die Richtigkeit der Wörter beziehen. Im ersten Fall ist bei 99 % jeder hundertste Buchstabe falsch, im letzteren jedes hundertste Wort. Ob falsche Layoutinformationen (Marginalie zwar richtig erkannt, aber an der falschen Stelle eingeordnet) oder fehlende Worttrennungen als Fehler gelten, wird von Fall zu Fall anders bewertet.

Um eine gewisse Einheitlichkeit bei der Beurteilung der **Genauigkeit** zugrunde legen zu können, werden Antragstellende gebeten, diesbezügliche Angaben hinsichtlich der **Buchstabengenauigkeit** zu machen, d. h. mangelhafte Worttrennungen und Layoutfehler unberücksichtigt zu lassen. Es ist also zu überprüfen, wie viele Zeichen der Quelle, einschließlich der Interpunktionszeichen, korrekt erkannt wurden. Ideal sind Messungen auf der Basis zuverlässiger Referenztexte, jedoch stehen diese nicht immer zur Verfügung, so dass auf **Stichproben** zurückgegriffen werden muss.

Um transkribierte oder mit OCR erstellte Texte auf ihre Genauigkeit hin zu überprüfen, wird hier unbeschadet späterer neuer Empfehlungen, die z. B. aus den Ergebnissen des OCR-D Projektes kommen, ein **statistisches Verfahren** empfohlen. Ziel ist, mittels einer Stichprobe zu überprüfen, ob eine vom Dienstleister behauptete Erkennungsquote als korrekt eingestuft werden kann, wobei einerseits die Wahrscheinlichkeit für einen Irrtum möglichst gering gehalten werden soll, andererseits aber die Stichprobengröße noch praktikabel ist. Das dazu erforderliche statistische Verfahren ist ein so genanntes **Bernoulli-Experiment**. Da die Berechnung relativ kompliziert ist, wird hier mit fest vorgegebenen Werten gearbeitet, die typische Angaben enthalten. Vorgeschlagen wird eine Stichprobegröße von 500 beliebigen Zeichen; zu empfehlen ist die Benutzung eines Zufallsgenerators, der die Position der Zeichen bestimmt (1. Zeichen: 15. Seite, 24. Zeile, 7. Zeichen. 2. Zeichen: 73. Seite, 3. Zeile, 32. Zeichen usw.). Unter dieser Voraussetzung gilt folgende Tabelle:

Behauptete Erkennungsquote	Mindestzahl der korrekt erkannten Zeichen (Stichprobengröße = 500)
95 %	485
96 %	489
97 %	493
98 %	496
99 %	499
> 99 %	500

In der linken Spalte ist die behauptete Erkennungsquote angegeben, in der rechten die Zahl der in der Stichprobe mindestens korrekt erkannten Zeichen, die vorliegen muss, um überprüfen zu können, ob eine von einem Dienstleister behaupte Erkennungsquote als korrekt eingestuft werden kann. Wenn also ein Dienstleister behauptet, dass ein Text eine Genauigkeit von 96 % hat, müssen in der Stichprobe von 500 Zeichen mindestens 489 Zeichen korrekt erkannt werden, damit bei einer Irrtumswahrscheinlichkeit von 2,5 % die Behauptung des Dienstleisters akzeptiert werden kann. Eine Genauigkeit unter 95 % sollte möglichst nicht vereinbart werden.

Bei einer behaupteten Erkennungsquote von über 99 % müsste die Stichprobegröße erhöht werden. Nachstehend zwei Tabellen, die beispielhaft zeigen, wie hoch die ermittelte Erkennungsquote in Abhängigkeit von einer bestimmten Stichprobengröße sein muss, wenn Texte eine behauptete Erkennungsquote von 99,5 % bzw. 99,7 % haben:

Behauptete Genauigkeit: 99,5 %

Stichprobengröße	Mindestzahl der korrekt erkannten Zeichen
500	500
1000	999
2000	1996
5000	4985
10000	9960

Behauptete Genauigkeit: 99,7 %

Stichprobengröße	Mindestzahl der korrekt erkannten Zeichen
500	500
1000	1000
2000	1998
5000	4995
10000	9990

Natürlich ergeben sich hier pragmatische Grenzen im Aufwand der Stichprobenerhebung. Die Frage, welche Genauigkeit gut, ausreichend oder mangelhaft ist, hängt von der konkreten wissenschaftlichen Fragestellung ab und lässt sich nur aus den konkreten Projektanforderungen her beantworten.

OCR

Die OCR-Technik hat in jüngerer Zeit erheblich Fortschritte gemacht, und wesentliche Verbesserungen konnten auch bei Frakturschriften sowohl der Handpressen- als auch Maschinenpressenzeit erzielt werden. Gleichwohl können die Praxisregeln angesichts der dynamischen Entwicklung und mit Blick auf die zu erwartenden Ergebnisse des OCR-D Projektes zur Frage der OCR-Software und ihrer Anwendbarkeit derzeit noch keine abschließenden Empfehlungen geben.

Der OCR-Prozess selbst kennt verschiedene Stufen, die aufeinander aufbauen. Grob unterscheiden kann man das Preprocessing, in dem ein Image vorbearbeitet wird (cropping, despeckle, deskewing, binarisation), die Optical Layout Recognition (Segmentation bzw. Identifikation von Bild- und Textregionen, Strukturanalyse), Optical Character Recognition (eigentliche Texterkennung) und Postprocessing (Textkorrektur). Insofern hat die Qualität des Ausgangsbilds ganz entscheidenden Einfluss auf den Prozess der Binarisierung und dieser wiederum auf das Erkennungsergebnis. Daher sollten nur solche digitalen Vorlagen einer OCR-Bearbeitung unterzogen werden, bei denen sichergestellt ist, dass die Bildqualität ausreicht. Problematisch sind in diesem Zusammenhang auch intrinsische Phänomene wie Verschmutzungen, Widerdrucksschatten, manuelle Unterstreichungen oder Annotationen u.ä., die sich

nachteilig auf den OCR-Prozess auswirken. Der Binarisierung folgt die Segmentierung, bei der die Software versucht, die jeweiligen Textbereiche auf einem Image zu identifizieren, um die eigentlichen Textbereiche von Illustrationen oder anderen Bildelementen zu trennen. Als besonders problematisch erweisen sich bei diesem Prozess Marginalien in älteren Drucken oder auch Zeitungen mit komplexem Layout. Danach folgt der eigentliche *OCR*-Prozess der Texterkennung, der durch einen nachgeordneten Prozess der Textverbesserung (automatisch mit Hilfe von Wörterbüchern oder manuelle Korrektur) unterstützt werden kann.

Um Daten aus der OCR leicht nachnutzen zu können, wird die Verwendung des Standards ALTO (Analyzed Layout and Text Object) empfohlen, der von der Library of Congress gepflegt wird.[56]

Manuelle Erfassung/double keying

Man unterscheidet bei der Direkterfassung – dem Abschreiben – von Texten zwei Verfahren, das einfache und das *double key*-Verfahren. Bei Letzterem wird ein Text zweimal erfasst und die Abweichungen beider Versionen durch automatischen Textabgleich herausgefiltert. Auf diese Weise sind Erfassungsgenauigkeiten von 99,97 % erreichbar, also ein praktisch fehlerfreier Text. Bei dieser Art Erfassung sollte man sich durch hohe Prozentsätze von Anbietern nicht irreführen lassen. Unterhalb einer Genauigkeit von 99,5 % ist bei manueller Erfassung ein Ergebnis ungenügend.

Sollte die Erfassung durch einen Dienstleister vorgenommen werden, so muss eine brauchbare Textgenauigkeit als Zielvorgabe auch vertraglich fixiert werden. Diese Vorgabe ist an Stichproben des digitalisierten Texts zu überprüfen (s.o.).

Manuelles Erfassen bietet zwar eine hohe Genauigkeit, ist aber auch kostenintensiv, so dass Vor- und Nachteile gegenüber einer OCR Erfassung abgewogen werden müssen. Die eigentliche Texterfassung wird hierbei zumeist im Ausland vorgenommen; der Kontakt mit einer Digitalisierungsfirma sollte jedoch über einen Ansprechpartner in Deutschland erfolgen, da in der Regel eine enge Kooperation und Absprache zu den Erfassungsdetails erforderlich ist.

In einem ersten Schritt muss durch das jeweilige Digitalisierungsprojekt festgelegt werden, welche Eigenschaften der Vorlage mittels eines strukturellen Markups erfasst werden sollen. Hierbei können nur solche Eigenschaften ausgezeichnet werden, die grafisch unterscheidbar sind. Einfache Strukturen können vom Dienstleister automatisch erkannt werden, bei weitergehenden Angaben müssen diese im Bild vor der Übergabe an den Dienstleister entsprechend markiert werden. Das verursacht einen entsprechenden Personalaufwand und muss bei der Kalkulation des Projekts berücksichtigt werden.

Da die meisten Dienstleister den Text nach Zeichenmenge inklusive des Markups berechnen, ist es ratsam, eine zeichenarme Auszeichnungsvariante[57] für diese Zwecke zu verwenden.

3.4.2 Zeichenkodierung

Alle verbreiteten Betriebssysteme unterstützen Unicode. Unicode ist auch das Zeichenkodierungsformat von XML, das die Grundlage für die wichtigsten Strukturdatenauszeichnungssys-

[56] http://www.loc.gov/standards/alto/
[57] Z. B. teitite: http://www.tei-c.org/release/doc/tei-p5-exemplars/html/tei_tite.doc.html

teme darstellt. Daher wird empfohlen, die Texte in **Unicode** abzuspeichern. Zu favorisieren ist das bei europäischen Sprachen sparsamere **UTF-8**. Zeichen, die nicht im Unicode-Standard enthalten sind, können durch Nutzung des *private plane*-Bereichs von Unicode[58] abgebildet und durch entsprechende Grafiken oder Fonts repräsentiert werden. In allen Fällen sind Möglichkeiten der Standardisierung zu prüfen.

Die Kodierung von langem und kleinem s in Frakturschriften oder Ligaturen in Frakturschriften (ch, tz etc.) bzw. die Darstellung von Diphtongen (æ etc.) ist abhängig von fachspezifischen Anforderungen oder editorischen Entscheidungen, die nicht Gegenstand dieser Empfehlungen sein können, aber bei der Erstellung von Volltext im Blick zu behalten und idealerweise zu dokumentieren sind.

3.4.3 Markup von Volltexten

Wenn nicht triftige Gründe dagegen sprechen, müssen **Volltexte von Drucken und Handschriften** nach dem Modell der **TEI kodiert** bzw. **mit Markup versehen** werden. Als transparentes XML-Format ist TEI, sofern sorgfältig dokumentiert, auch für die Langzeitarchivierung die prospektiv beste Wahl. Von PDF/A sollte trotz bestehender ISO-Normen (19005-1:2005 bzw. 19005-2:2011) abgesehen werden, da deren Nachnutzung gerade durch die digitale arbeitenden Geistes- und Kulturwissenschaften mangels struktureller Auszeichnung beschränkt ist. Dessen ungeachtet ist PDF – neben zunehmend auch ePub für mobile Geräte – als derivatives Format für z. B. dynamisch generierte Lesefassungen oder für den Druck aufbereitete Texte gut geeignet und sollte wegen seiner weiten Verbreitung auch zusätzlich von digitalen Bibliotheken angeboten werden (vgl. z. B. die Angebote bei archive.org).

Bei der Kodierung von XML-Strukturen in TEI-Dokumenten muss zunächst entschieden werden, wie und in welchem Umfang man textsortenspezifische Gliederungen wie z. B. Kapitel, Unterkapitel, Jahresbände, Aufsätze etc. berücksichtigt. Hinzu kommen weitere denkbare Strukturmerkmale, z. B. Inhaltsverzeichnisse, Register, Zeilenumbruch, Spaltenumbruch, Seitenumbruch, Kopfzeile/Fußzeile/Kolumnentitel, Seitenzahl, Bilder oder bildähnliche Elemente, Bildunterschriften, Marginalien, Schriftwechsel, z. B. der Wechsel von Fraktur zu Antiqua (etwa für fremdsprachige Zitate), der Wechsel der Schriftgröße, Wechsel der Schriftart (recte, kursiv, fett usw.) u. a. m., Formeln, z. B. mathematische (MathML) oder chemische (CML) Formeln, Fortsetzungsmarkierungen (Kustoden) am Fußende von Seiten (für Anschlussbogen) etc.

Die Wahl des Markups unterliegt in der Regel projektspezifischen Besonderheiten. Um die Austauschbarkeit und Nachnutzung von auf diese Weise mit Markup versehenen Volltexten sicherzustellen, sollten daher die verwendeten XML-Elemente und Attribute im TEI-Header dokumentiert werden. Besondere Beachtung verdienen in diesem Zusammenhang die Bemühung der TEI Community um ein strikt formuliertes Austauschformat: TEI–Simple.[59]

3.4.4 Layout

In manchen Fällen ist es bei der Präsentation eines Volltexts wichtig, das Layout eines Dokumentes langfristig zu sichern. Die Praxisregeln empfehlen für diese Fälle bevorzugt den Einsatz einer **geeigneten Formatierungssprache** (z. B. XSLT, CSS), die die Unabhängigkeit

[58] http://de.wikipedia.org/wiki/Unicode
[59] https://github.com/TEIC/TEI-Simple

von spezieller Software weitgehend sicherstellt. Falls eine Archivierung des Formats mit XML-Techniken aus nachvollziehbaren Gründen nicht möglich ist, können Layoutinformationen zu Textdokumenten auch in PDF nach der ISO-Norm 19005-1 archiviert werden. PDF-Dateien sind aber, wie in 3.4.2. dargelegt, kein Ersatz für eine Bereitstellung des mit Markup versehenen Volltextes in XML.

Das Wesen einer Publikation als XML + Formatierungssprache bringt es mit sich, dass dynamische Anzeigen nach Maßgabe des jeweils vom Nutzer gewünschten Zwecks generiert werden können. Dies sollte bei der Präsentation berücksichtigt und ein möglichst breites Spektrum angeboten werden. Typische Ausgabeformate sind z. B. HTML/XHTML, PDF, ePub oder plain text.

3.5 Langzeitverfügbarkeit

Die **Langzeitsicherung** und **Archivierung** digitaler Inhalte ist auf der Ebene der Bitstream-Archivierung gut untersucht und kann als gelöst angesehen werden; Probleme bestehen allerdings nach wie vor bei der Bewahrung der Ebene der so genannten *information representation* (nach OAIS). Man kann dabei zwischen der Archivierung genuin digitaler Informationen und der Langzeitsicherung von Digitalisaten unterscheiden, die von vorhandenen analogen Objekten erstellt werden. Hier können unterschiedliche Maßstabe an die Archivierung gelegt werden.

Im Rahmen der Langzeitsicherung werden Dateien in stabilen, migrationsfähigen Formaten auf einem technisch wie organisatorisch sicheren Speichersystem gesichert. Die Archivierung digitaler Daten setzt auf einem solchen Speicher auf, beinhaltet jedoch noch weitreichendere technische und organisatorische Festlegungen, die nicht nur eine physische Erhaltung der Daten, sondern auch Strategien zur Bereitstellung für die Nutzung (Access) der Daten, auch im Kontext vorhandener und zukünftiger Informationssysteme. Die Speichersysteme müssen auf Redundanz ausgelegt sein.

Für die Langzeitarchivierung sind besonders technische Informationen und solche über die Veränderungshistorie eines Objekts wichtig.[60] Insbesondere, wenn es um die Veränderungshistorie eines Objekts geht, hat sich PREMIS (Preservation Metadata: Implementation Strategies)[61] als Datenmodell etabliert.

Ob die erzeugten digitalen Inhalte langzeitgesichert oder archiviert werden, hängt von der Strategie der jeweiligen Institution ab, die mit einer Antragstellung erläutert werden muss.[62] Wesentliche Erfolgsfaktoren für eine Langzeitsicherung und Archivierung von digitalen Dokumenten sind

(1) die Schaffung der organisatorischen und wirtschaftlichen Rahmenbedingungen,
(2) die Schaffung der technischen Rahmenbedingungen bzw. die Auswahl einer geeigneten technischen Methode/Strategie.

---

[60] http://www.langzeitarchivierung.de/Subsites/nestor/DE/Standardisierung/Metadaten.html;jsessionid=5239B697C2522475FBECD446877A114F.prod-worker5

[61] http://www.loc.gov/standards/premis/understanding_premis_german.pdf

[62] Eine Institution kann in ihrer Strategie z. B. schlüssig festlegen, dass *born digital*-Objekte archiviert werden und Abbilder von analogen Objekten auf langzeitsicheren Speichern abgelegt werden. Ebenso ist es aus guten Gründen möglich, auch die vom Original erstellten Digitalisate zu archivieren, wenn z. B. die Originale fragil sind.

Für die Archivierung digitaler Daten ist das Open Archival Information System (OAIS) als Referenzmodell anzuwenden.[63] Die „Kriterien für vertrauenswürdige digitale Langzeitarchive“ sind essenziell. Mit ihnen werden u. a. der organisatorische Rahmen, gesetzlichen Rahmenbedingungen, das Qualitätsmanagement sowie die Authentizität für ein vertrauenswürdiges Archiv geschaffen.[64] Die Umsetzung der Langzeitsicherung erfolgt in „Digitalen Magazinen“ oder Reproduktionenverwaltungen, die sich am OAIS-Modell orientieren.

Die Langzeitverfügbarkeit der Ergebnisse von Digitalisierungsprojekten ergibt sich einerseits aus der Wahl der Daten- und Metadatenformate. Überlegungen dazu flossen in die vorigen Kapitel ein. Andererseits ist sicherzustellen, dass die digitalen Daten auch physikalisch verfügbar bleiben. Dabei gilt: **Kosten für die projektbezogene Sicherung der Daten** werden in DFG-geförderten Digitalisierungsprojekten als **Eigenleistung** für die Laufzeit des Projekts anerkannt. Eine **Förderung** dieser Kosten aus DFG-Mitteln kann **nicht** erfolgen.

Die Langzeitsicherungs- bzw. Archivierungsfrage ist ein integraler Bestandteil jedes Digitalisierungsvorhabens. Aufwand und Kosten sollten nicht unterschätzt werden. Nicht nur die Kosten für den Speicherplatz, der je nach Projekt mehrere Terabyte betragen kann, sondern auch der Aufwand für die physische Erhaltung müssen langfristig berücksichtigt werden.

Es sei darauf hingewiesen, dass Digitalisierungsprojekte aus Sicht der DFG stets **Projekte der gesamten Einrichtung** sind: Die Unterstützung der das Projekt abwickelnden Fachabteilung durch die **IT-Infrastruktur des Hauses** wird vorausgesetzt. Dabei wird begrüßt, wenn sich kleinere Einrichtungen der **Kompetenz und der Dienstleistung größerer Einrichtungen** bedienen.

Anträge müssen **nachvollziehbare Aussagen zur institutionellen Langzeitsicherung und Archivierung** enthalten.

4. Organisatorische Fragen – Eigendigitalisierung versus Digitalisierung in Dienstleistung

Die Digitalisierung kann *inhouse* oder in Dienstleistung erfolgen. Im ersten Fall können neben Personalkosten ggf. Gerätekosten beantragt werden, sofern sie projektbedingt und nicht der Grundausstattung der Einrichtung zuzurechnen sind.

Die **Entscheidung** darüber, ob ein Vorhaben als ***Inhouse*-Projekt** konzipiert oder ob auf das **Angebot eines Dienstleisters** zurückgegriffen wird, ist immer eine **projektspezifische Entscheidung** unter Abwägung aller Sach- und Finanzgründe, die allein in die Verantwortung der Antragstellerinnen und Antragsteller fällt. Schließlich ist das Heranziehen eines Dienstleisters vor allem eine Sache des Vertrauens. Anders als bei der Filmdigitalisierung, die für das Objekt unkritisch ist, sollte der Dienstleister im Umgang mit unikalem Material oder bibliothekarischem Altbestand Erfahrungen vorweisen können.

---

[63] Das OAIS- Referenzmodell ist als ISO 14721-Standard verabschiedet: http://public.ccsds.org/publications/archive/650x0m2.pdf

[64] Vgl. Schoger, Astrid/Susanne Dobratz/Reinhard Altenhöner: Kriterienkatalog vertrauenswürdige digitale Langzeitarchive, Frankfurt am Main, 2008. Vgl.: http://nbn-resolving.de/urn:nbn:de:0008-2008021802. Vgl. dazu auch: DIN 31644:2012-04.

Die Arbeit mit einem Dienstleister ist auch dann möglich, wenn die zu digitalisierenden Materialien nicht außer Haus gegeben werden können, da bei größeren Projekten die Dienstleister mit eigenen Geräten und Personal vor Ort arbeiten.

Bei der Formulierung der einschlägigen Verträge ist zu beachten:

- Die **geforderten Leistungsparameter müssen exakt festgelegt werden**, insbesondere die Anforderungen und Formatstandards für die zu liefernden Rohdaten. Sie sind gemäß den Vorgaben der „Praxisregeln Digitalisierung“ projektspezifisch in einem Lastenheft zu formulieren. Die beauftragten Unternehmen sollten zertifizierte Verfahren zur Qualitätssicherung nachweisen können. Der Auftraggeber ist zu einer sorgfältigen Qualitätskontrolle der Lieferungen verpflichtet, bevor Rechnungen vollständig bezahlt werden.
- Die DFG erwartet, **dass ein angemessener Prozentsatz des Rechnungsbetrags als Sicherheit zurückbehalten und erst nach einer durchgeführten Qualitätskontrolle zur Auszahlung an das Dienstleistungsunternehmen** kommt. Gleichzeitig soll vom Dienstleistungsunternehmen die schriftliche verbindliche Zusicherung eingefordert werden, dass kurzfristig und kostenlos Ersatzleistung oder Nachbesserung geleistet wird, wenn dies wegen Nichteinhaltung des Lastenhefts oder wegen sonstiger berechtigter Qualitätsmängel nötig werden sollte.

5. Zitieren digitaler Ressourcen, persistente Adressierung

In den Anfangszeiten der Digitalisierung wurde das Problem der Zitierfähigkeit digitaler Ressourcen meist unterschätzt. Dabei ist es gerade die Zitierbarkeit, die die wissenschaftliche Nutzung einer digitalisierten Quelle auch im Internet erlaubt. Im Unterschied zu vormaligen Sekundärformen wie Mikrofilm oder Papierausdruck ist die Internetressource nicht nur eine Kopie des Originals, die wie das Original behandelt und so zitiert wird, sondern ein selbständiges Objekt in einem sich dynamisch entwickelnden integralen Forschungsraum. Wird ein Digitalisat ins Internet gestellt, muss es, um von anderen Dokumenten oder Datenbanken aus erreichbar zu sein, **eindeutig angesprochen** werden können. Dazu ist über die im Analogen übliche Zitierform hinaus, die durch ein Angebot in der Navigationssoftware als klassische Form weitergenutzt werden kann und sollte, die Festlegung und die online zugängliche Dokumentation von **Adressierungstechniken** nach Maßgabe der in einem Forschungsvorhaben erforderlichen Granularitäten (Buch, Seite, Sammlung, Akte, Objekt, Objekteil, etc.) erforderlich.

Ein positiver Effekt der netzbasierten Zitierform ist, dass das **Referenzieren eindeutig** wird, was man wegen der vielen Fehler in alten Drucken oder auch aufgrund des Mangels an Paginierungen oder Foliierungen etwa in Inkunabeln meist nicht erreichen kann. Daher sollte neben die „inhaltliche“ Zitierform (z. B. p.8, a4, 213r etc.) auch eine „formale“ Zitierform der Bildsequenz treten. So lassen sich auch Images eindeutig zitieren, die nicht zum eigentlichen Werk gehören (Deckel, Spiegel, zusätzliche digitalisierte Wasserzeichen, Teilaufnahmen von Illustrationen etc.).

Bei Volltexten gelten andere Mechanismen, zu denen derzeit noch keine konkreten Empfehlungen ausgesprochen werden können (eindeutige Referenzierungsmöglichkeiten bestünden z. B. mit XPointer, ID-Vergaben oder vergleichbaren Techniken).

Sichergestellt werden muss die Erreichbarkeit und Zitierbarkeit mittels URI bzw. IRI bei Druckwerken und Handschriften sowohl des Werkes als Ganzem als auch die Erreichbarkeit und Zitierbarkeit von einzelnen physischen Seiten aus diesem Werk. Einrichtungen sollten durch geeignete Mechanismen (PURL, URN, DOI, Handle, etc.) die **Persistenz einer Ressource und der Verknüpfung zu ihr** gewährleisten, um zuverlässiges Arbeiten mit den bereitgestellten Quellen in wissenschaftlichen Kontexten zu ermöglichen. Ausdrücklich sei auch auf den Nutzen der eindeutigen Zitierbarkeit mittels URI von Ressourcen im Zuge der Weiterentwicklung des *semnatic web* und *Linked Open Data* (LOD) verwiesen.

Die Erzeugung von **URNs** über die Deutsche Nationalbibliothek mindestens auf Werkebene wird nachdrücklich empfohlen.[65]

6. Bereitstellung der Projektergebnisse für die Öffentlichkeit

6.1 Rechte, Lizenzierung und Open Access

Die Deutsche Forschungsgemeinschaft ist Mitunterzeichnerin der „Berlin Declaration“[66] zu *open access*. Die Ergebnisse der durch sie geförderten Digitalisierungsprojekte stehen der weltweiten Forschung daher grundsätzlich unentgeltlich und im Geiste dieser Erklärung zur Verfügung. Insofern wird grundsätzlich davon ausgegangen, dass Digitalisate in einer Qualität, die die Verfolgung der großen Mehrzahl der üblichen Forschungszwecke erlaubt, kostenfrei über das Internet bereitgestellt werden. Sollte in Projekten davon abgewichen werden, ist dies im Antrag ausdrücklich zu begründen. Eine Einschränkung des *open access* zu den erzeugten Daten kann z. B. datenschutz-, urheber- oder archivrechtlich begründet sein, darf aber maximal 5 % der insgesamt zu digitalisierenden Materialien betreffen.[67]

Die Digitalisierung wissenschaftlicher Materialien wird durch die DFG gefördert, um diese Materialien **der Forschung in Deutschland und weltweit** zugänglich zu machen. Alle Projekte sind also so anzulegen, dass die **Ergebnisse der Forschung frühzeitig und nachhaltig zugänglich** werden. Dies wird in nahezu allen Fällen die Bereitstellung der Digitalisate im Internet einschließen. Sollten Metadaten und / oder Digitalisate aus rechtlichen Gründen nicht frühzeitig nach ihrer Erstellung frei im *open access* verfügbar gemacht werden können, ist sicherzustellen, dass dies zeitnah mit dem Wegfall entgegenstehender rechtlicher Beschränkungen erfolgt.

Während **vor der Digitalisierung** die Rechte an den zu digitalisierenden Materialien zu klären und ggfs. für die Digitalisierung und Zurverfügungstellung einzuholen sind (→ 1.8; 2.2), ist **bei der Digitalisierung** sicherzustellen, dass gegebenenfalls hierbei entstehende Rechte der bewahrenden Institution zufallen. Bei der Herstellung eines digitalen Abbildes des Originals können Rechte entstehen. Grundsätzlich gilt dies nicht bei Massendigitalisierungsverfahren von Flachware.

---

[65] http://www.persistent-identifier.de/

[66] http://oa.mpg.de/openaccess-berlin/berlindeclaration.html

[67] Eine Förderung der Bearbeitung nicht zugänglichen Materials mit DFG-Mitteln ist nicht möglich und ein Konzept zur Zugänglichmachung sowohl hinsichtlich der technischen Lösung als auch der zeitlichen Dimension Voraussetzung.

Bei der **Erstellung von Metadaten** entstehen in der Regel keine Rechte im Hinblick auf die einzelnen Datensätze.[68] Werden existierende Erschließungsinformationen digitalisiert, die über rein formale Angaben hinausgehend auch Beschreibungstexte enthalten und damit rechtlich geschützt sein können, so ist ggfs. eine Rechteeinholung für die Zugänglichmachung und die Nutzungsmöglichkeit der Digitalisate einzuholen.

Im Sinne des *open access* und der *open source* wird vorausgesetzt, dass alle Ergebnisse – Metadaten, Digitalisate und Volltexte bzw. das den Volltexten zu Grunde liegende XML sowie XSLT-Skripte und DTDs oder XML-Schema-Dateien – so frei nachnutzbar wie rechtlich möglich angeboten werden, sei dies durch die Markierung als **Public Domain** oder die Verwendung freier **Creative Commons Lizenzen**[69]. Bei der Digitalisierung gemeinfeien Materials wird die Markierung der Digitalisate als Public Domain erwartet. Bei allen Ergebnissen, bei denen dies nicht möglich ist, weil Rechte bestehen oder bei der Digitalisierung entstanden sind, ist ebenfalls sicherzustellen, dass die Daten im *open access* und so frei nachnutzbar wie möglich zur Verfügung gestellt und lizenziert werden. In Frage kommt ein vollständiger Verzicht auf entstandene Rechte durch die Markierung mit CC0 (z. B. bei Metadaten), aber auch die Verwendung der freien Lizenztypen CC BY oder CC BY SA (→ 1.8). Wird noch geschütztes Material digitalisiert und kann das Digitalisat wegen der noch bestehenden Rechte am Ausgangsmaterial nicht frei lizenziert werden, sollte der Rechtsstatus mit standardisierten Rechtehinweisen versehen werden.[70] Bei Projekten, in denen nicht nur gemeinfreie Materialien digitalisiert werden und / oder mit kommerziellen Partnern oder Verlagen zusammengearbeitet wird, kann eine **verzögerte *open access*-Publikation (*moving wall*) von bis zu einem Jahr nach Projektende** vereinbart werden.

Maßgebliches Ziel der freiest möglichen Verfügbarmachung der Ergebnisse und der vereinheitlichen Rechteangaben ist es, Möglichkeiten zu schaffen, Digitalisate, Metadaten und Volltexte auch in anderen Nutzungskontexten als der unmittelbaren Projektumgebung (z. B. zum Zwecke des *data mining* oder der Datenaggregation) auswerten zu können; dies erfordert das Herunterladen, die Neuindexierung und das Anbieten in eigenen Forschungs- und Präsentationskontexten. Im diesem Kontext gilt für Images aktuell, dass sie in einer Form bereitgestellt werden, die grundsätzlich eine **vollumfängliche wissenschaftliche Nachnutzung** in anderen Forschungskontexten erlaubt (z. B. durch Nutzung von einzelnen Images „in Frames“ nach dem Modell des DFG-Viewers bzw. Googles iFrame-Einbettung oder durch generelle Erteilung einer Erlaubnis, Vorschaubilder auf fremden Servern anzeigen zu dürfen). Dazu sollen bei den Digitalisaten entweder hochaufgelöste Derivate im TIFF-Format oder vollaufgelöste JPEG-Images mit einer Kompression von 90, maximal 80 in Kombination mit der vollen Scanauflösung zur Verfügung gestellt werden. Die Erhebung von Entgelten für Kopien in darüber hinausgehenden Qualitäten, Derivaten oder für die Herstellung anderer Medienformen (CD, Druck, etc.) bleibt davon unberührt.

Die DFG erwartet, dass DFG-geförderte Projekte auf ihren im Internet angebotenen Daten **eindeutige Herkunftsnachweise sowie soweit möglich den Hinweis auf die Förderung durch die DFG** anbringen. Dies erfolgt im Fall von Bilddigitalisaten in der Regel durch das

---

[68] Siehe: Klimpel, Das Eigentum an Metadaten, in Euler / Hagedorn-Saupe / Maier; Handbuch Kulturportale. Online-Angebote aus Kultur und Wissenschaft, Berlin/Boston 2015, S. 57 – 64.
[69] Es ist immer die aktuellste Version der Creative Commons Lizenzen zu verwenden: https://creativecommons.org/.
[70] www.rightsstatements.org

Hinzurechnen einer Nachweisleiste in die veröffentlichte Nutzerkopie (zum Beispiel im Grafikformat JPEG). Im Fall von Volltexten geschieht der Nachweis durch entsprechende Hinweise im Header der Textdatei, ebenso können die Nachweise in die Bildheader eingebettet werden. In jedem Fall aber müssen die Herkunftsangabe und der Hinweis auf die Förderung in den begleitenden Metadaten enthalten sein und im Bereitstellungssystem auch ausgegeben werden.

6.2 Mindestanforderungen für die Bereitstellungssysteme von Digitalisaten

6.2.1 Funktionalitätsanforderungen

Sammlungen/Bestände sind in der Regel über eine Vielzahl von Wegen zugänglich zu machen:

- über die **Website** der anbietenden Institution
- über den lokal implementierten oder an einem anderen Ort betriebenen **„DFG-Viewer“**, sofern für das angebotene Material zutreffend (→ 7.3)
- über eine Suchanfrage an den **lokalen und regionalen Bibliothekskatalog** / das **lokale und regionale Archivportal /** an die **jeweilige materialspezifische Online-Anwendung** und die **überregionalen Nachweis- und Präsentationssysteme** (DDB, Archivportal-D)
- sofern vorhanden über eines der **von der DFG geförderten materialspezifischen Portale**, die einen integrierten Zugriff auf alle im DFG-Programm geförderten digitalen Sammlungen ermöglichen, bzw. ein gemeinsames fachliches Portal z. B. der Fachinformationsdienste
- über **Internet-Suchmaschinen**
- über eine OAI-Schnittstelle.

Über die genannte **OAI-Funktionalität** ist auch der **DFG-Viewer** zu bedienen (sofern das jeweilige Format unterstützt wird). Eine Ablieferung lediglich durch XML-Dateien sollte zugunsten der OAI-Funktionalitäten nicht mehr erfolgen. Neu eingerichtete OAI-Schnittstellen sollten in Frage kommenden Portalen und Fachinformationsdiensten gemeldet werden. Ferner sollten geeignete Maßnahmen ergriffen werden, die dazu führen, dass die Metadaten von Suchmaschinen gefunden werden (z. B. mit dem *sitemap protocol*[71]).

Den Nutzerinnen und Nutzern sollte neben dem gezielten Zugriff auf spezifizierte Dokumente über eine Recherche in Metadaten (*searching*) auch die Möglichkeit geboten werden, in vorab definierten Sammlungen, Sammlungsteilen oder Beständen strukturiert zu navigieren (*browsing*). Bei der Suche ist darauf zu achten, dass im Zweifelsfall einfache, google-orientierte Suchinstrumente einer größeren Benutzergemeinschaft dienen als stark gefelderte Suchmasken, die ein eingehendes Verständnis der Datenstruktur der jeweiligen Sammlung oder des jeweiligen Bestands erfordern. Idealerweise werden beide Aspekte durch die Facettierung der Suche kombiniert, bei der die Nutzerinnen und Nutzer Möglichkeiten an die Hand

[71] http://www.sitemaps.org/

bekommen, selbst größere Treffermengen nach definierten Kriterien (Facetten) zu reduzieren. Darüber hinaus bleibt auch die differenzierte, stark gefelderte Suche als zusätzliches Hilfsmittel für eine hochspezialisierte Fachcommunity ein wünschenswertes Angebot.

Außerdem sind folgende Funktionen zu implementieren:

- Downloadfunktion[72]
- Druckfunktion für die ausgegebene Dokumentansicht[73]
- Aus zentralen DFG-geförderten Informationssystemen (VD16, VD17, VD18, Fachportalen, materialspezifischen Portalen etc. sollte zuerst auf eine Ansicht im Style „DFG-Viewer“ verlinkt werden.
- Für sämtliche DFG-geförderten Bereitstellungssysteme wird erwartet, dass Nutzerinnen und Nutzer die **automatisierte Möglichkeit zum Feedback** bezüglich des digitalen Angebots erhalten. Eine entsprechende Möglichkeit ist auf der Projektseite bzw. im digitalen Bereitstellungssystem an zentraler Stelle einzurichten.

6.2.2 Technische Mindestanforderungen

Server sind so anzulegen, dass sie, soweit anwendbar[74],

- alle Materialien in einer hinreichend guten Qualität anbieten, so dass die Benutzung zu wissenschaftlichen Zwecken auf an den Hochschulen üblicherweise vorhandenem Equipment mühelos möglich ist. Dies wird beispielsweise dadurch erreicht, dass Schrift in einer Größe angeboten wird, die problemlos gelesen werden kann.
- alle Materialien aber auch in einer Qualität anbieten, die über DSL-Anschlüsse ohne unzumutbare Verzögerungen bearbeitbar sind.
- zum wissenschaftlichen Gebrauch den kostenlosen Download kompletter Einheiten in einer einzigen Datei (beispielsweise einzelne gedruckte Werke) ermöglichen.
- alle jeweils gängigen Browser unterstützen, soweit dies der Natur der Sache nach nicht unmöglich ist.[75]

7. Präsentationsstandards für Textwerke, DFG-Viewer

Die oben angesprochenen Grundsätze gelten für alle Arten von Projekten, die digitale Inhalte bereitstellen. Bei der Bereitstellung von Digitalisaten, die den Charakter von digitalen Büchern oder mehrseitigen Dokumenten haben, gelten zusätzlich folgende Mindestanforderungen. Diese betreffen einige Basisanforderungen und die mindestens zu realisierende Funktionalität.

---

[72] Ein Download nach Abschnitten oder einzelnen Seiten ist dann vorzusehen, wenn die Größe der Gesamtdatei nicht mehr handhabbar wäre. Grundsätzlich ist auch eine Bereitstellung in Originalgröße des Bildmasters für die Forschung unentgeltlich (mindestens JPEG 80% Kompression) zu ermöglichen.

[73] Eine Druckfunktion nach Abschnitten oder einzelnen Seiten ist dann vorzusehen, wenn die Größe der Gesamtdatei nicht mehr handhabbar wäre.

[74] Hier wird nachdrücklich das Kriterium der Alltagstauglichkeit gefordert, es geht nicht um die Realisierung abstrakter Wünschbarkeiten. Können Objekte eines Projekts ihrer Natur nach auf Bildschirmen mit einer Auflösung unter 1600 x 1200 nicht sinnvoll dargestellt werden, wird kein Aufwand für Scheinlösungen gefordert; ist ein Objekt unter 3 MB nicht sinnvoll bearbeitbar, verletzt es das Kriterium der DSL-Tauglichkeit des Servers nicht, wenn kein Angebot kleinerer Versionen erfolgt.

[75] Wird ein für eine fortgeschrittene 3D-Anwendung notwendiges Format von einem Browser nicht unterstützt, ist es nicht erforderlich den Aufwand zu betreiben, ein geeignetes Plugin zu entwickeln.

7.1 Basisanforderungen und Architektur

Das Bereitstellungssystem verknüpft die digitalisierten Bild- oder Volltextdateien zu einer Dokumentstruktur, um für die Nutzerinnen und Nutzer eine Navigation innerhalb des Dokumentes zu ermöglichen. Weiterhin stellt es Verbindungen zwischen digitalen Dokumenten bzw. Teilen davon (z. B. Kapitel, Seiten) und Metadaten her, um den Nutzerinnen und Nutzern, ausgehend von einer Recherche in Metadatenbeständen, den Zugriff auf das einzelne Dokument oder auf bestimmte Dokumentteile zu ermöglichen. Schließlich organisiert es digitale Dokumente zu sachlich oder durch Provenienz zusammengehörigen digitalen Sammlungen bzw. Beständen, um den Nutzerinnen und Nutzern die Navigation zwischen Dokumenten und Sammlungen fachgerecht anzubieten. Es stellt entsprechende Benutzeroberflächen für Recherche, Navigation, Zugriff und Abruf von Metadaten, Dokumenten, Sammlungen und Beständen zur Verfügung und unterstützt den weitgehend automatisierten Ex- und Import von standardkonformen Rohdaten. Die Bereitstellungssysteme der einzelnen Informationsinfrastruktureinrichtungen sollten sowohl bei der Navigation in digitalen Sammlungen/Beständen als auch bei der Recherche über Indizes einen institutionenübergreifenden Zugriff ermöglichen. Darüber hinaus ist eine transparente Verknüpfung der Bereitstellungssysteme sowohl mit lokalen Katalog-/Informationssystemen als auch übergreifenden Informationssystemen wünschenswert.

Zur Lösung dieser Aufgaben können unterschiedliche Systemarchitekturen eingesetzt werden. Folgende grundlegende Alternativen sind denkbar:

- Die Metadaten werden zentral in einem Online-Informationssystem (z. B. dem lokalen OPAC bzw. Online-Findmittelsystem oder einem übergreifenden Informationssystem wie einem Bibliotheksverbundkatalog oder dem Archivportal-D) gehalten, die digitalen Dokumentdateien inklusive XML-codierter Strukturdaten werden in einem hierarchisch gegliederten Dateisystem auf einem gesonderten Dokumentenserver für den Online-Zugriff bereitgestellt. Die Struktur der digitalisierten Sammlung bzw. die interne Struktur der digitalisierten Dokumente kann dabei durch die Hierarchie des Dateisystems abgebildet werden.
- Ein Dokumenten-Management-System (DMS) oder Content-Management-System (CMS) kommt zum Einsatz, bei dem sowohl die Metadaten als auch die Digitalisate im Datenbank-System gespeichert sind.

Normalerweise kommt die erste Variante zum Einsatz, die eine verteilte und transparente Informationsinfrastruktur ermöglicht.

7.2 Funktionalitätsanforderungen

Ein wesentliches funktionales Qualitätskriterium ist der Komfort bei der Navigation innerhalb eines aufgefundenen Dokuments. Die folgenden **Navigationsmöglichkeiten** gelten als Basisanforderungen:

- Ansteuern eines beliebigen Images
- Anfang: Springen an den Anfang eines Dokuments
- Ende: Springen an das Ende eines Dokuments
- Vor: Eine Seite vorgehen
- Zurück: Eine Seite zurückgehen

- Volltextrecherche innerhalb der Digitalisate (für Bücher ab 1850 verpflichtend)[76]
- Metadaten-Info: Hier können Nutzerinnen und Nutzer die Informationen aus den im Informationssystem gespeicherten Beschreibungsfeldern zu „ihrem“ digitalen Dokument einsehen.
- Hilfe: Über das Hilfemenü sollte eine detaillierte Beschreibung mit Fallbeispielen zur Navigation und für die Suche in der Digital Library zugänglich sein.

Wenn sachlich möglich und geboten, sind Inhaltsverzeichnisse/Strukturbäume oder funktionale Äquivalente vorzusehen; diese sind grundsätzlich durchsuchbar zu gestalten. Wünschenswert sind Navigationshilfen, z. B. grafische Repräsentationen in einer Kopfzeile, die den Nutzerinnen und Nutzern signalisieren, an welcher Stelle des digitalen Dokuments sie sich gerade befinden. Enthält ein Informationssystem Materialien, die konzeptuell in der Benutzung üblicherweise zu übergeordneten Einheiten zusammengefasst werden (mehrbändige Werke), so sollen diese übergeordneten Einheiten als solche sichtbar werden.

7.3 DFG-Viewer

Um für die wissenschaftliche Benutzung neben den unterschiedlich gestalteten und dezentral verantworteten Webangeboten der jeweiligen Einrichtungen einen einheitlichen Zugriff auf die Daten (Inhalte) aller DFG-geförderter Digitalisate bieten zu können, werden derzeit zwei sich ergänzende Strategien verfolgt:

- Festlegung eines einheitlichen Designprofils für die Visualisierung von in DFG-Förderung erstellten Digitalisaten („DFG-Viewer“)
- Schaffung einer OAI-Schnittstelle mit Unterstützung des METS-Standards: Diese Schnittstelle dient zunächst der einheitlichen Anzeige von Images und ihren Metadaten in DFG-geförderten Projekten und transportiert materialabhängig METS/MODS für Drucke und Archivgut und METS/TEI für Handschriften. Über diese Schnittstelle werden Funktionen des Weiterblätterns, der Anzeige von Metadaten und andere grundsätzliche Funktionen beschrieben und damit einheitliche Anzeige- und Blätterfunktionen geschaffen, die auch innerhalb von zentralen Nachweisportalen einen homogenen Zugriff auf dezentrale Ressourcen (*harvesting*) erlauben.

Zum Zweck der überregionalen Erstpräsentation müssen DFG-geförderte Digitalisierungsvorhaben die genannten Schnittstellen bedienen und/oder den Viewer im DFG-Style an ihrer Einrichtung selbst implementieren. Anhang A und B definieren das METS/MODS- und METS/TEI-Format der Schnittstelle sowie das Viewer-Design.

[76] Nach dem Stand der gegenwärtigen Technik ist eine OCR-Erkennung bei Drucken der Maschinenpressenzeit ab 1850 verpflichtend.

Anhang A:
METS/MODS-Profil für die Darstellung im DFG-Viewer und Übermittlung per OAI

1. DFG-Viewer

Zum Zweck der einheitlichen Präsentation beim Zugriff aus überregionalen Nachweissystemen (z.B. VD16/VD17/VD18, ZVDD, „Virtuelle Fachbibliotheken“, Deutsche Digitale Bibliothek, Archivportal-D) auf lokale digitale Angebote sollten DFG-geförderte Projekte die als DFG-Viewer bekannte Browserdarstellung nutzen sowie die dem Viewer hinterlegten Schnittstellen bedienen. Wissenschaftlerinnen und Wissenschaftlern soll auf diese Weise die Nutzung digitalisierter Inhalte erleichtert werden. Aus dem DFG-Viewer heraus ist die Weiterleitung auf die speziellen lokalen Angebote der jeweiligen Einrichtungen möglich.

Der Viewer[1] wurde im Rahmen der Aktionslinie „Digitalisierung des VD16/VD17“ durch die in der ersten Ausschreibungsrunde geförderten Bibliotheken aufgebaut. Er wird durch die SLUB Dresden in Zusammenarbeit mit weiteren Partnern beständig weiterentwickelt. Gehostet wird die Referenzanwendung des DFG-Viewers derzeit an der SLUB Dresden.

Um den DFG-geförderten Projekten eine größtmögliche Sicherheit bei der Planung ihrer Anträge zu ermöglichen und um die Konformität der Metadaten mit den Anforderungen des DFG-Viewers sicherzustellen, sollten Metadaten aus Projekten gegen eines der von der Website des Viewers abzurufenden Anwendungsprofile valide sein.[2]

Zur Darstellung von Metadaten im Viewer wird METS[3] verwendet. METS dient dabei als ein Rahmenformat (*wrapper*), innerhalb dessen sowohl deskriptive, administrative und strukturelle Metadaten als auch Ressourcen (z. B. Images, Volltexte) verzeichnet werden. Der DFG-Viewer benötigt zum Zwecke der Anzeige von deskriptiven Metadaten (Drucke, Zeitschriften, Zeitungen, Archivgut) nach MODS[4] kodierte Metadaten (vgl. unten, Punkt 2). Für administrative Metadaten (u. a. lokale Anwendung, Homepage, Logo der Einrichtung) wird zur Verlinkung ein spezielles, für den Viewer entwickeltes Format (Namespace *dv*) verwendet.

Eine ausführliche Dokumentation zur Implementierung des METS-Formats befindet sich auf der Homepage der Referenzanwendung des DFG-Viewers.

Diese Vorgaben gelten grundsätzlich für alle digitalisierten Medienarten. Im Einzelfall gibt es jedoch Formatalternativen, die anstelle von METS/MODS verwendet werden können (METS/TEI für mittelalterliche und frühneuzeitliche Handschriften, EAD zur Kodierung von

[1] http://www.dfg-viewer.de/
[2] http://dfg-viewer.de/profil-der-metadaten
[3] http://www.loc.gov/standards/mets/
[4] http://www.loc.gov/standards/mods/

archivischen Findmitteln, LIDO für museale Objekte – dann allerdings ohne Möglichkeit der Anzeige im DFG-Viewer).

2. MODS-DFG-Standard-Set (Bibliotheksbestände)

Der MODS-Standard bietet ein vereinfachtes *subset* zu MARC21, daher sollten automatische Konversionen aus gängigen Katalogen leicht möglich sein. Für die Anzeige im Viewer sind nur wenige Pflichtfelder erforderlich (s. nachstehende Tabelle).

Element /Unterelement	***Wieder-holbar***	***Bemerkungen /Erläuterungen***	***Status***
1 Titelinformation			
<titleInfo>	ja	Titelinformationen; sofern das Werk keinen Titel enthält, muss er fingiert werden. In Fällen von mehrbändigen Werken ohne Stücktitel reicht es aus, den Titel des mehrbändigen Werkes in dem Element `mods:relatedItem` anzugeben. Verpflichtend ist dann aber die Angabe der Bandzählung in dem Element `mods:part` mit den Unterelementen `mods:detail` und `mods:number`.	Pflicht (Ausnahme sind mehrbändige Werke ohne Stücktitel)
<titleInfo> / <title>	nein	Enthält den Haupttitel des Werks.	Pflicht
<titleInfo> / <subTitle>	nein	Enthält den Zusatz zum Haupttitel der Werks.	Pflicht, wenn vh.
2 Person			
<name type="personal" authority="…">	ja	Person, die mit dem Werk in Beziehung steht (z.B. der Autor). Das Attribut @authority enthält den Code für das Regelwerk, nach dem die Person angesetzt wurde; in der Regel „gnd“. Wahlweise kann in @valueURI ein URI für den Personennamen verwendet werden.	Pflicht, wenn vh.
<name> / <namePart type="…">	ja	Enthält Namensbestandteile, deren Art in @type festgelegt ist. Mögliche Werte sind „date“, „family“, „given“, „termsOfAddress“.	
<name> / <displayForm>	nein	Name in der gewünschten Anzeigeform	Empfohlen
<name> / <role>	nein	Rahmenelement zur Funktion der Person	Pflicht, wenn vh.
<name> / <role> / <roleTerm type="code" authority="marcrelator">	nein	Funktion der Person; der Wert des Feldes <roleTerm> ist kodiert (MARC Relator Code)[5].	

[5] Ein Code aus der Liste MARC Value List for Relators and Roles: http://www.loc.gov/marc/sourcecode/relator/relatorlist.html.

Element /Unterelement	*Wieder-holbar*	*Bemerkungen /Erläuterungen*	*Status*
3 Körperschaft			
<name type="corporate" authority="">	ja	Körperschaft, die mit dem Werk in Beziehung steht.	Pflicht, wenn vh.
<name> / <namePart>	ja	s.o.	
<name> / <role>	nein	s.o.	
<name> / <role> / <roleTerm type="code" authority="marcrelator">	nein	s.o.	
4 Erscheinungsvermerk			
<originInfo eventType=“publication“>	ja	Erscheinungsvermerk; der erste <originInfo>-Block ist für Angaben zur Quelle vorgesehen, Angaben zur digitalen Edition können in einem zweiten <originInfo> Block mit dem eventType=“ digitisation“ gemacht werden.	Pflicht, wenn vh.
<originInfo> / <place>	ja	Enthält Elemente zum Erscheinungsort.	Pflicht, wenn vh.
<originInfo> / <place> / <placeTerm type="text">	nein	Enthält den Erscheinungsort; sofern kein Erscheinungsort zu ermitteln ist, „[o.O.]“ angeben.	Pflicht
<originInfo> / <publisher>	ja	Enthält den Verlag/Drucker.	Pflicht, wenn vh.
<originInfo> / <dateIssued keyDate="yes" encoding="w3cdtf">	ja	Enthält das Erscheinungsjahr der Quelle; sofern kein Erscheinungsjahr zu ermitteln ist, „[o.J.]“ angeben.	Pflicht
5 Ausgabevermerk			
<originInfo> / <edition>	ja	Enthält die Ausgabebezeichnung (s.a. oben bezg. Angaben zur elektronischen Ausgabe).	Pflicht, wenn vh.
6 Kollationsvermerk/physische Beschreibung			
<physicalDescription>	nein	Kollationsvermerk	Pflicht
<physicalDescription> / <extent>	ja	Enthält Umfangs-, Format und Illustrationsangaben.	Pflicht, wenn vh.
<physicalDescription> / <digitalOrigin>	nein	Bei digitalisierten Drucken ist der Inhalt des Feldes <digitalOrigin> normalerweise „reformatted digital“.	Pflicht

Element /Unterelement	*Wieder-holbar*	*Bemerkungen /Erläuterungen*	*Status*
7 Übergeordnetes Werk			
<relatedItem type="host"> / <recordInfo> / <recordIdentifier>	nein	<recordIdentifier> ist ein Identifier, der zur Verknüpfung mit hierarchisch übergeordneten Datensätzen dient.	Pflicht, wenn Hierarchie
8 Bandangabe			
<part order=""> / <detail> / <number>	nein	Angaben zu Teilen; Wert des Attributs @order ist ein beliebiger numerischer Wert, der die Reihenfolge der Teile sicherstellt. <number> enthält die Bandangabe.	Pflicht, wenn Hierarchie vh.
9 Sprache			
<language>	ja	Sprachangaben	Pflicht, wenn vh.
<language> / <languageTerm type="code" authority="iso639-2b">	nein	Enthält die Sprache des Werks in gemäß ISO 639-2/B kodierter Form.	
10 Zitierfähige Identifizierung			
<identifier type="…">	ja	Weltweit eindeutige Identifizierung der Ressource (Attribut @type z.B. urn, purl, doi, handle, uri etc.). Angegeben werden muss auch, wenn vorhanden, die GW- und VD-Nummer des Drucks (Attribute @type VD16, VD17, GW).	Pflicht
11 Datenbank-ID			
<recordInfo> / <recordIdentifier>	nein	Datensatz-Identifier zur eindeutigen Identifizierung innerhalb eines Datenbanksystems, z.B. Pica Produktionsnummer	Pflicht
12 Signatur			
<location> <physicalLocation> <shelfLocator>	ja	Standort und Signatur des Originals	Empfohlen

Sollten diese Daten auch für ein Harvesting über OAI in Betracht kommen, ist je nach Materialart und Projektzuschnitt eine Erweiterung dieses Basissets zu erwägen. Für diesen Zweck stehen auf der Seite des DFG-Viewers differenziertere Formatdefinitionen zur Verfügung. Diese enthalten auch weitere Erläuterungen zur Besetzung der jeweiligen Felder.

3. Hinweise für Archive zur Konversion von EAD(DDB) nach MODS

Das Mapping von EAD(DDB) auf entsprechende MODS-Elemente ist grundsätzlich leicht durchführbar. Für die vereinfachte Bearbeitung von archivischen Datensätzen stehen in der archivischen Community auch Tools zur Verfügung, die von interessierten Einrichtungen genutzt werden können.[6]

Für archivische Daten werden folgende Elemente empfohlen; es ist Einrichtungen aber natürlich freigestellt, auch weitere Erschließungsinformationen entsprechend des MODS-Anwendungsprofils anzugeben:

Feldbezeichnung	***EAD(DDB)***	***MODS***	***Status***
Titel	c/did/unittitle	mods:titleInfo, gegebenenfalls mit Unterelementen MODS-Definition: A word, phrase, character, or group of characters, normally appearing in a resource, that names it or the work contained in it. Kommentar: Für das Wurzelstrukturelement in einem METS/MODS-Satz ist die Angabe mindestens eines Titels unter Verwendung des Elements mods:titleInfo mit dem Unterelement mods:title verpflichtend.	Pflicht
Institution	ancestor::archdesc/did/repository/corpname	mods:physicalLocation MODS-Definition: The institution or repository that holds the resource or where it is available.	Pflicht
Signatur	c/did/unitid[not(@type)]	mods:shelfLocator MODS-Definition: Shelfmark or other shelving designation that indicates the location identifier for a copy.	Pflicht

[6] Weitere Informationen unter https://pro.deutsche-digitale-bibliothek.de/dokumente-links-fuer-archive.

Laufzeit	c/did/unitdate	mods:dateCreated MODS-Definition: The date of creation of the resource Kommentar: Datum der Entstehung bzw. der Laufzeit der analogen Vorlage.	Pflicht, wenn v.h.
Kontext[7]	ancestor::c/did/unittitle	mods:relatedItem, gegebenenfalls mit Unterelementen MODS-Definition: Information that identifies other resources related to the one being described. Über mods:relatedItem lassen sich nicht nur hierarchische Beziehungen abbilden. Vielmehr lassen sich mittels des type-Attributs eine Vielzahl von Beziehungen darstellen. Um die Beziehung zwischen Dokumenten oder Dokumentteilen mit MODS herzustellen, müssen eindeutige Identifier verwendet werden.	Empfohlen
ID	c/@id	mods:recordIdentifier MODS-Definition: Contains the system control number assigned by the organization creating, using, or distributing the record.	Pflicht

[7] Dies bezieht sich auf den Bestandskontext der Ressource, also die darüber liegenden Verzeichnungsstufen aus Tektonik und Klassifikation.

4. Beispiele für METS/MODS Datensätze nach dem DFG-Standard

4.1 Für Bibliotheksgut:

```
<?xml version="1.0" encoding="UTF-8"?>
<mets:mets xmlns:mets="http://www.loc.gov/METS/" xmlns:tei="http://www.tei-c.org/ns/1.0"
   xmlns:xlink="http://www.w3.org/1999/xlink" xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
   xmlns:mods="http://www.loc.gov/mods/v3" xmlns:dv="http://dfg-viewer.de/"
   xsi:schemaLocation="http://www.loc.gov/METS/ http://www.loc.gov/mets/mets.xsd">
   <mets:metsHdr>
      <mets:agent ROLE="IPOWNER">
         <mets:name>Herzog August Bibliothek Wolfenbüttel</mets:name>
      </mets:agent>
      <mets:agent ROLE="CREATOR">
         <mets:name>Herzog August Bibliothek Wolfenbüttel</mets:name>
      </mets:agent>
   </mets:metsHdr>
   <mets:dmdSec ID="dmdsec">
      <mets:mdWrap MDTYPE="MODS">
         <mets:xmlData>
            <mods:mods
               xsi:schemaLocation="http://www.loc.gov/mods/v3 http://www.loc.gov/standards/mods/v3/mods-3-
5.xsd">
               <mods:titleInfo>
                  <mods:title>Occulta Philosophia</mods:title>
                  <mods:subTitle>Von den verborgenen Philosophischen Geheimnussen […]</mods:subTitle>
               </mods:titleInfo>
               <mods:identifier type="fingerprint">inl- t?h- esin Woab 3 1613R</mods:identifier>
               <mods:identifier type="vd17">39:122887Y</mods:identifier>
               <mods:originInfo eventType="publication">
                  <mods:dateIssued keyDate="yes" encoding="iso8601">1613</mods:dateIssued>
                  <mods:place>
                     <mods:placeTerm type="text">Franckfurt am Mayn</mods:placeTerm>
                  </mods:place>
                  <mods:publisher>Bringer</mods:publisher>
               </mods:originInfo>
               <mods:name type="personal" valueURI="http://d-nb.info/gnd/119068982">
                  <mods:namePart type="termsOfAddress">Trismegistus</mods:namePart>
                  <mods:namePart type="given">Hermes</mods:namePart>
                  <mods:displayForm>Hermes &lt;Trismgivenegistus&gt;</mods:displayForm>
                  <mods:role>
                     <mods:roleTerm authority="marcrelator" type="code"
valueURI="http://id.loc.gov/vocabulary/relators/oth">oth</mods:roleTerm>
                  </mods:role>
               </mods:name>
               <mods:name type="personal" valueURI="http://d-nb.info/gnd/118507036">
                  <mods:namePart type="termsOfAddress">Valentinus</mods:namePart>
                  <mods:namePart type="given">Basilius</mods:namePart>
                  <mods:displayForm>Basilius &lt;Valentinus&gt;</mods:displayForm>
                  <mods:role>
                     <mods:roleTerm authority="marcrelator" type="code"
valueURI="http://id.loc.gov/vocabulary/relators/oth">oth</mods:roleTerm>
                  </mods:role>
```

```
            </mods:name>
            <mods:name type="personal">
               <mods:namePart type="given">Ludwig Georg</mods:namePart>
               <mods:namePart type="family">Stolberg</mods:namePart>
               <mods:displayForm>Stolberg, Ludwig Georg</mods:displayForm>
               <mods:role>
                  <mods:roleTerm authority="marcrelator" type="code"
valueURI="http://id.loc.gov/vocabulary/relators/dte">dte</mods:roleTerm>
               </mods:role>
            </mods:name>
            <mods:name type="personal">
               <mods:namePart type="given">Johann</mods:namePart>
               <mods:namePart type="family">Bringer</mods:namePart>
               <mods:displayForm>Bringer, Johann</mods:displayForm>
               <mods:role>
                  <mods:roleTerm authority="marcrelator" type="code"
valueURI="http://id.loc.gov/vocabulary/relators/clb">clb</mods:roleTerm>
               </mods:role>
            </mods:name>
            <mods:name type="personal" valueURI="http://d-nb.info/gnd/1037506073">
               <mods:namePart type="given">Johann</mods:namePart>
               <mods:namePart type="family">Bringer</mods:namePart>
               <mods:displayForm>Bringer, Johann</mods:displayForm>
               <mods:role>
                  <mods:roleTerm authority="marcrelator" type="code"
valueURI="http://id.loc.gov/vocabulary/relators/prt">prt</mods:roleTerm>
               </mods:role>
            </mods:name>
            <mods:language>
               <mods:languageTerm type="code" authority="iso639-2b">ger</mods:languageTerm>
            </mods:language>
            <mods:genre authority="aad">Alchemie</mods:genre>
            <mods:location>
               <mods:physicalLocation authority="marcorg">DE-23</mods:physicalLocation>
               <mods:shelfLocator>M: Nd 14 (1)</mods:shelfLocator>
            </mods:location>
            <mods:physicalDescription>
               <mods:extent>[4] Bl., 78 S.</mods:extent>
               <mods:extent>4°</mods:extent>
               <mods:extent>Tbl. r&amp;amp;s, Ill. (Holzschn.)</mods:extent>
            </mods:physicalDescription>
            <mods:recordInfo>
               <mods:recordIdentifier source="DE-23">092275664</mods:recordIdentifier>
               <mods:recordOrigin xml:lang="en">Converted from PICA using a local XSL transformation
script</mods:recordOrigin>
               <mods:recordContentSource authority="marcorg">DE-23</mods:recordContentSource>
            </mods:recordInfo>
         </mods:mods>
      </mets:xmlData>
   </mets:mdWrap>
</mets:dmdSec>
<mets:amdSec ID="amdsec">
   <mets:rightsMD ID="rights">
```

```
      <mets:mdWrap MDTYPE="OTHER" OTHERMDTYPE="DVRIGHTS">
        <mets:xmlData>
          <dv:rights>
            <dv:owner>Herzog August Bibliothek Wolfenbüttel</dv:owner>
            <dv:ownerLogo>http://www.hab.de/images/logo_dfg_viewer.gif</dv:ownerLogo>
            <dv:ownerSiteURL>http://www.hab.de/</dv:ownerSiteURL>
            <dv:ownerContact>mailto:auskunft@hab.de</dv:ownerContact>
          </dv:rights>
        </mets:xmlData>
      </mets:mdWrap>
    </mets:rightsMD>
    <mets:digiprovMD ID="digiprov">
      <mets:mdWrap MDTYPE="OTHER" OTHERMDTYPE="DVLINKS">
        <mets:xmlData>
          <dv:links>
            <dv:reference>http://opac.lbs-braunschweig.gbv.de/DB=2/PPN?PPN=092275664</dv:reference>
            <dv:presentation>http://diglib.hab.de/drucke/nd-14-1s/start.htm</dv:presentation>
          </dv:links>
        </mets:xmlData>
      </mets:mdWrap>
    </mets:digiprovMD>
  </mets:amdSec>
  <mets:fileSec>
    <mets:fileGrp USE="DEFAULT">
      <mets:file ID="file--idp65681520" MIMETYPE="image/jpeg">
        <mets:FLocat xlink:href="http://diglib.hab.de/drucke/nd-14-1s/00001.jpg"
          LOCTYPE="URL"/>
      </mets:file>
      <!--  weitere Verweise auf Images aus Umfangsgründen ausgelassen -->
      <mets:file ID="file--idp65467888" MIMETYPE="image/jpeg">
        <mets:FLocat xlink:href="http://diglib.hab.de/drucke/nd-14-1s/eb04.jpg"
          LOCTYPE="URL"/>
      </mets:file>
    </mets:fileGrp>
    <mets:fileGrp USE="MAX">
      <mets:file ID="file-max-idp65681520" MIMETYPE="image/jpeg">
        <mets:FLocat xlink:href="http://diglib.hab.de/drucke/nd-14-1s/max/00001.jpg"
          LOCTYPE="URL"/>
      </mets:file>
      <!-- weitere Verweise auf Images aus Umfangsgründen ausgelassen -->
      <mets:file ID="file-max-idp65467888" MIMETYPE="image/jpeg">
        <mets:FLocat xlink:href="http://diglib.hab.de/drucke/nd-14-1s/max/eb04.jpg"
          LOCTYPE="URL"/>
      </mets:file>
    </mets:fileGrp>
    <mets:fileGrp USE="MIN">
      <mets:file ID="file-min-idp65681520" MIMETYPE="image/jpeg">
        <mets:FLocat xlink:href="http://diglib.hab.de/drucke/nd-14-1s/min/00001.jpg"
          LOCTYPE="URL"/>
      </mets:file>
      <!--  weitere Verweise auf Images aus Umfangsgründen ausgelassen]-->
      <mets:file ID="file-min-idp65467888" MIMETYPE="image/jpeg">
        <mets:FLocat xlink:href="http://diglib.hab.de/drucke/nd-14-1s/min/eb04.jpg"
```

```
            LOCTYPE="URL"/>
        </mets:file>
      </mets:fileGrp>
      <mets:fileGrp USE="THUMBS">
        <mets:file ID="file-thumbs-idp65681520" MIMETYPE="image/jpeg">
          <mets:FLocat xlink:href="http://diglib.hab.de/drucke/nd-14-1s/thumbs/00001.jpg"
            LOCTYPE="URL"/>
        </mets:file>
        <!--  weitere Verweise auf Images aus Umfangsgründen ausgelassen -->
        <mets:file ID="file-thumbs-idp65467888" MIMETYPE="image/jpeg">
          <mets:FLocat xlink:href="http://diglib.hab.de/drucke/nd-14-1s/thumbs/eb04.jpg"
            LOCTYPE="URL"/>
        </mets:file>
      </mets:fileGrp>
    </mets:fileSec>
    <mets:structMap TYPE="PHYSICAL">
      <mets:div ID="struct-physical" TYPE="physSequence" DMDID="dmdsec" ADMID="amdsec">
        <mets:div ID="struct-physical-idp65681520"
          CONTENTIDS="http://diglib.hab.de/drucke/nd-14-1s/start.htm?image=00001" TYPE="page"
          ORDER="1">
          <mets:fptr FILEID="file--idp65681520"/>
          <mets:fptr FILEID="file-min-idp65681520"/>
          <mets:fptr FILEID="file-max-idp65681520"/>
          <mets:fptr FILEID="file-thumbs-idp65681520"/>
        </mets:div>
        <!-- weitere Verweise auf Strukturen aus Umfangsgründen ausgelassen -->
        <mets:div ID="struct-physical-idp65467888"
          CONTENTIDS="http://diglib.hab.de/drucke/nd-14-1s/start.htm?image=eb04" TYPE="page"
          ORDER="94">
          <mets:fptr FILEID="file--idp65467888"/>
          <mets:fptr FILEID="file-min-idp65467888"/>
          <mets:fptr FILEID="file-max-idp65467888"/>
          <mets:fptr FILEID="file-thumbs-idp65467888"/>
        </mets:div>
      </mets:div>
    </mets:structMap>
    <mets:structMap TYPE="LOGICAL">
      <mets:div ID="struct-logical" TYPE="Monograph" DMDID="dmdsec" ADMID="amdsec"
        CONTENTIDS="http://diglib.hab.de/drucke/nd-14-1s/start.htm">
        <mets:div ID="struct-logical-idp65475312" TYPE="binding">
          <mets:div ID="struct-logical-idp65474288" TYPE="endsheet"/>
        </mets:div>
        <mets:div ID="struct-logical-idp65472496" TYPE="section">
          <mets:div ID="struct-logical-idp65477616" TYPE="title_page"/>
          <mets:div ID="struct-logical-idp65477104" TYPE="dedication"
            LABEL="Dem Wolgebornen Graffen und Herrn/ Herrn Lufwig Georgen/ Graffen zu Stolberg /
Königstein / Roschefort/ Weringeroda und Hohenstein / ...">
            <mets:div ID="struct-logical-idp65478128" TYPE="ornament"/>
          </mets:div>
        </mets:div>
        <mets:div ID="struct-logical-idp65474032" TYPE="section"
          LABEL="Das Erste Theil., Ein Colloquium oder Gesprech zwischen einem alten und jungen
Studenten.">
```

```
            <mets:div ID="struct-logical-idp65483504" TYPE="section"
              LABEL="Adolphus. Mein Freundtlichen Gruß und alle Wolfahrt wünsche ich Euch lieber alter
Senior: ...">
              <mets:div ID="struct-logical-idp65482352" TYPE="ornament"/>
            </mets:div>
            <mets:div ID="struct-logical-idp65494384" TYPE="section"
              LABEL="und Abgöttisch ist / ein ander Werck neben Gottes zusetzen/ ..."/>
            <mets:div ID="struct-logical-idp65498224" TYPE="section"
              LABEL="greiff das ewig Leben/ dazu du auch beruffen bist/ ..."/>
          </mets:div>
          <mets:div ID="struct-logical-idp65499248" TYPE="section"
            LABEL="Das ander Theil Aureliae Occultae Philosophorum.">
            <mets:div ID="struct-logical-idp65498352" TYPE="illustration"/>
            <mets:div ID="struct-logical-idp65505904" TYPE="section"
              LABEL="Schmaragt Tafel Hermetis.">
              <mets:div ID="struct-logical-idp65511152" TYPE="illustration"/>
              <mets:div ID="struct-logical-idp65510640" TYPE="illustration"/>
              <mets:div ID="struct-logical-idp65507952" TYPE="illustration"/>
              <mets:div ID="struct-logical-idp65516400" TYPE="illustration"/>
            </mets:div>
            <mets:div ID="struct-logical-idp65516656" TYPE="section"
              LABEL="Würckung deß Philosophischen Wercks/ erst Figur.">
              <mets:div ID="struct-logical-idp65521520" TYPE="illustration"/>
            </mets:div>
            <mets:div ID="struct-logical-idp65513584" TYPE="section"
              LABEL="Die ander Figur deß Wercks.">
              <mets:div ID="struct-logical-idp65520240" TYPE="illustration"/>
            </mets:div>
           <!-- weitere Strukturmerkmale aus Umfangsgründen ausgelassen-->
          </mets:div>
          <mets:div ID="struct-logical-idp65539824" TYPE="binding">
            <mets:div ID="struct-logical-idp65543536" TYPE="endsheet"/>
          </mets:div>
          <mets:div ID="struct-logical-idp65538672" TYPE="binding">
            <mets:div ID="struct-logical-idp65538928" TYPE="cover"/>
            <mets:div ID="struct-logical-idp65541232" TYPE="paste_down"/>
            <mets:div ID="struct-logical-idp65539312" TYPE="annotation"/>
            <mets:div ID="struct-logical-idp65549296" TYPE="paste_down"/>
            <mets:div ID="struct-logical-idp65549424" TYPE="cover"/>
          </mets:div>
        </mets:div>
      </mets:structMap>
      <mets:structLink>
        <mets:smLink xlink:from="struct-logical" xlink:to="struct-physical"/>
        <mets:smLink xlink:from="struct-logical-idp65475312" xlink:to="struct-physical-idp65681520"/>
        <!-- weitere Verknüpfungslinks aus Umfangsgründen ausgelassen-->
        <mets:smLink xlink:from="struct-logical-idp65549424" xlink:to="struct-physical-idp65467888"
        />
      </mets:structLink>
</mets:mets>
```

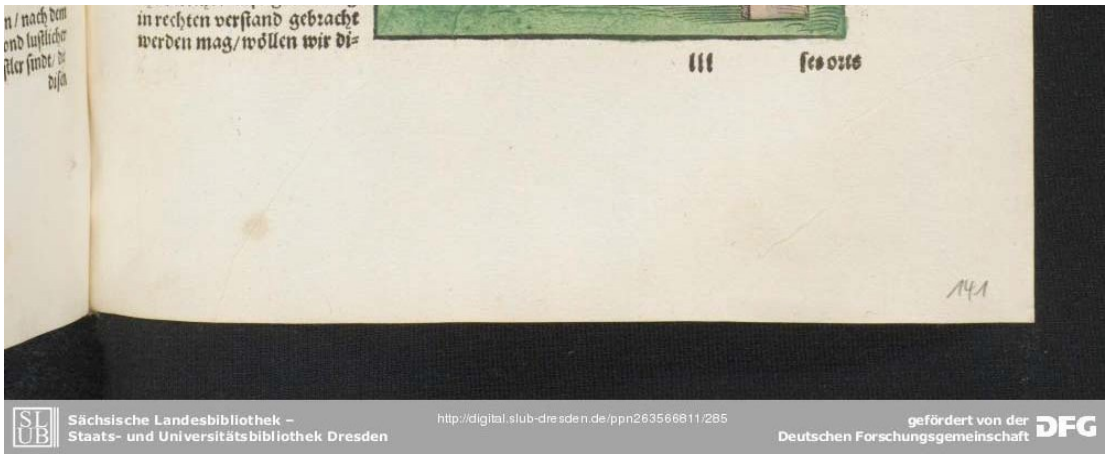
4.2 Für Archivgut:

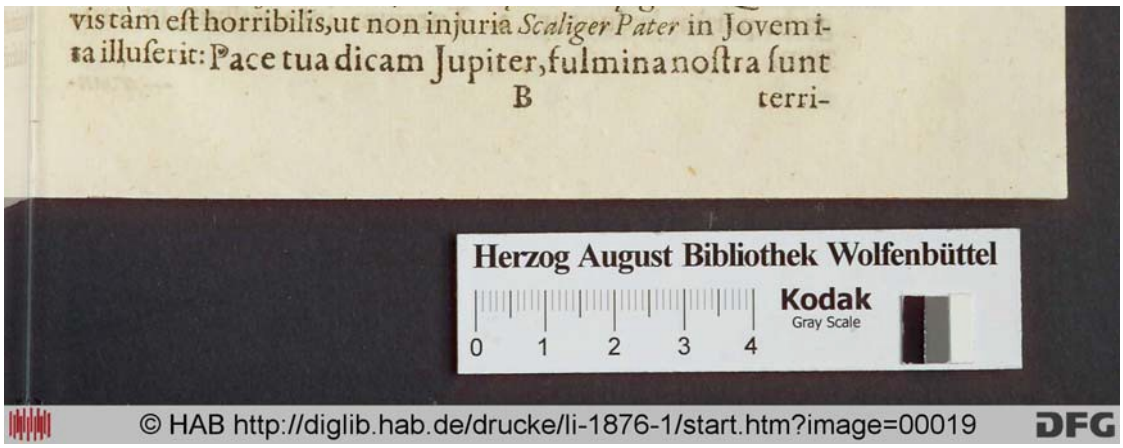
```
<mets:dmdSec ID="DMDLOG_1">
  <mets:mdWrap MDTYPE="MODS">
   <mets:xmlData>
    <mods:mods>
     <mods:recordInfo>
      <mods:recordIdentifier>labw-1-1263080</mods:recordIdentifier>
      <mods:descriptionStandard>DFG-Viewer/Archiv v2.3</mods:descriptionStandard>
     </mods:recordInfo>
     <mods:location>
      <mods:physicalLocation>Landesarchiv Baden-Württemberg, Abt. Hauptstaatsarchiv
Stuttgart</mods:physicalLocation>
      <mods:shelfLocator>H 51 U 589</mods:shelfLocator>
     </mods:location>
     <mods:titleInfo>
      <mods:title>Kaiser Karl publiziert die dreiundzwanzig ersten Kapitel der Goldenen Bulle</mods:title>
     </mods:titleInfo>
     <mods:originInfo eventType="production">
      <mods:dateCreated point="end" qualifier="approximate" encoding="iso8601">1356 Januar
10</mods:dateCreated>
     </mods:originInfo>
     <mods:relatedItem type="host" displayLabel="Gliederung">
      <mods:titleInfo>
       <mods:title>Urkunden</mods:title>
      </mods:titleInfo>
      <mods:relatedItem type="host" displayLabel="Bestand">
       <mods:titleInfo>
        <mods:title>Kaiserselekt</mods:title>
       </mods:titleInfo>
      </mods:relatedItem>
     </mods:relatedItem>
    </mods:mods>
   </mets:xmlData>
  </mets:mdWrap>
 </mets:dmdSec>
 <mets:amdSec ID="amd1">
  <mets:rightsMD ID="rights1">
   <mets:mdWrap MIMETYPE="text/xml" MDTYPE="OTHER" OTHERMDTYPE="DVRIGHTS">
    <mets:xmlData>
     <dv:rights>
      <dv:owner>Landesarchiv Baden-Württemberg, Abt. Hauptstaatsarchiv Stuttgart</dv:owner>
      <dv:ownerSiteURL>http://www.landesarchiv-bw.de/hstas</dv:ownerSiteURL>
      <dv:ownerContact>mailto:landesarchiv@la-bw.de</dv:ownerContact>
      <dv:ownerLogo>https://www.landesarchiv-bw.de/dfgviewer/logo.jpg</dv:ownerLogo>
     </dv:rights>
    </mets:xmlData>
   </mets:mdWrap>
  </mets:rightsMD>
  <mets:digiprovMD ID="digiprov1">
   <mets:mdWrap MIMETYPE="text/xml" MDTYPE="OTHER" OTHERMDTYPE="DVLINKS">
    <mets:xmlData>
     <dv:links>
      <dv:reference>http://www.landesarchiv-bw.de/plink/?f=1-1263080</dv:reference>
      <dv:presentation>http://www.landesarchiv-bw.de/plink/?f=1-1263080-1</dv:presentation>
     </dv:links>
    </mets:xmlData>
   </mets:mdWrap>
  </mets:digiprovMD>
 </mets:amdSec>
```

```
<mets:fileSec>
  <mets:fileGrp USE="DEFAULT">
    <mets:file ID="file-full-image_1-1263080-1" MIMETYPE="image/jpeg">
      <mets:FLocat LOCTYPE="URL"
        xlink:href="http://www.landesarchiv-bw.de/plink/?f=1-1263080-1&amp;ext=456788&amp;format=jpg"/>
    </mets:file>
    <mets:file ID="file-full-image_1-1263080-2" MIMETYPE="image/jpeg">
      <mets:FLocat LOCTYPE="URL"
        xlink:href="http://www.landesarchiv-bw.de/plink/?f=1-1263080-2&amp;ext=456788&amp;format=jpg"/>
    </mets:file>
    <mets:file ID="file-full-image_1-1263080-3" MIMETYPE="image/jpeg">
      <mets:FLocat LOCTYPE="URL"
        xlink:href="http://www.landesarchiv-bw.de/plink/?f=1-1263080-6&amp;ext=456788&amp;format=jpg"/>
    </mets:file>
  </mets:fileGrp>
</mets:fileSec>
<mets:structMap TYPE="PHYSICAL">
  <mets:div ID="phys1" TYPE="physSequence">
    <mets:div ID="image_1-1263080-1" ORDER="1" TYPE="page">
      <mets:fptr FILEID="file-full-image_1-1263080-1"/>
    </mets:div>
    <mets:div ID="image_1-1263080-2" ORDER="2" TYPE="page">
      <mets:fptr FILEID="file-full-image_1-1263080-2"/>
    </mets:div>
    <mets:div ID="image_1-1263080-3" ORDER="3" TYPE="page">
      <mets:fptr FILEID="file-full-image_1-1263080-3"/>
    </mets:div>
  </mets:div>
</mets:structMap>
<mets:structMap TYPE="LOGICAL">
  <mets:div ID="log1" TYPE="act" DMDID="DMDLOG_1" ADMID="amd1"
    LABEL="Kaiser Karl publiziert die dreiundzwanzig ersten Kapitel der Goldenen Bulle" ORDER="1">
    <mets:div ID="Bild1" TYPE="section" LABEL="Schmuckkassette" ORDER="2"/>
    <mets:div ID="Bild2" TYPE="section" LABEL="Einband Vorderseite" ORDER="3"/>
    <mets:div ID="Bild3" TYPE="section" LABEL="fol 01r" ORDER="4"/>
  </mets:div>
</mets:structMap>
<mets:structLink>
  <mets:smLink xlink:from="Bild1" xlink:to="image_1-1263080-1"/>
  <mets:smLink xlink:from="Bild2" xlink:to="image_1-1263080-2"/>
  <mets:smLink xlink:from="Bild3" xlink:to="image_1-1263080-3"/>
</mets:structLink>
</mets:mets>
```

5. Hinzurechnen einer Fußzeile

Die Fußzeile sollte bei den an die Nutzer ausgelieferten bzw. im Internet angebotenen Formaten am unteren Rand des Images hinzurechnet werden. Links sollte sich das Institutionslogo, bei Förderung durch die DFG nach Möglichkeit rechts zusätzlich das Logo der DFG befinden. Die Angabe einer zitierfähigen URL im Mittelbereich (ergänzend zur Übergabe in der Viewer-XML-Datei) wird empfohlen. Schrift und Logos müssen der Auflösung entsprechend skaliert werden. In PDF-Dateien zum Download können – zusätzlich zu einem Vorblatt – ebenfalls Fußzeilen an jedes Bild angefügt werden. Vergleiche die untenstehenden Beispiele:





Anhang B:
METS/TEI-Spezifikation für die Darstellung von digitalisierten Handschriften

1. DFG-Viewer

Ziel der vorliegenden METS/TEI-Spezifikation für digitalisierte mittelalterliche und frühneuzeitliche Handschriften ist die Darstellung der Digitalisate unter der vereinheitlichten Oberfläche des DFG-Viewers. Der ursprünglich für die Präsentation von digitalisierten Druckwerken geschaffene DFG-Viewer wird somit medientypologisch erweitert, ohne dabei die gewohnte Bedienoberfläche aufzugeben. Diese wird lediglich um wenige Funktionen ergänzt, um den spezifischen Anforderungen des neuen Medientyps gerecht zu werden.

Da es sich nicht um eine grundsätzlich neue Implementierung des DFG-Viewers handelt, sondern lediglich um die Unterstützung eines zusätzlichen Anwendungsprofils für METS/TEI, gelten die in Anhang A gemachten Angaben zu Druckwerken analog auch für digitalisierte Handschriften – insbesondere die allgemeinen Informationen zum DFG-Viewer in Abschnitt 1 sowie die Hinweise zum Hinzurechnen einer Fußzeile in Abschnitt 5.

Weitere Informationen zu den vom DFG-Viewer unterstützten Datenformaten und Funktionen sowie zur Einbindung des Webdienstes in das eigene Angebot finden Sie auf der Webseite des DFG-Viewers[1].

2. TEI-DFG-Standard-Set (Handschriftenbestände)

Die folgenden Angaben zur Kodierung einer digitalisierten Handschrift in TEI beschränken sich auf die nötigen Pflichtfelder und stellen nicht das vollständige Spektrum aller Möglichkeiten dar. Eine ausführliche Dokumentation aller Optionen findet sich auf der Webseite des DFG-Viewers.

Grundlagen des Anwendungsprofils sind die TEI-Spezifikation der Text Encoding Initiative[2] sowie die Vorarbeiten des Europeana Regia-Projekts[3] zu denen das DFG-Standard-Set kompatibel ist.

Element / Unterelement	wiederholbar	Bemerkungen / Erläuterungen	Status
1 Angaben zur Identifikation			
Alle in diesem Abschnitt beschriebenen Elemente sind Unterelemente von `teiHeader/fileDesc/sourceDesc/msDesc`			
<head> / <title>	nein	Enthält den Titel der Handschrift nach DFG-Richtlinien zur Handschriftenkatalogisierung.	Pflicht, wenn vh.
<msIdentifier> / <settlement>	nein	Enthält den geographischen Aufbewahrungsort der Handschrift.	Pflicht
<msIdentifier> / <reposi-	nein	Enthält den Namen der die Handschrift auf-	Pflicht

[1] http://dfg-viewer.de/
[2] http://www.tei-c.org/
[3] http://www.europeanaregia.eu/

tory>		bewahrenden Einrichtung.	
<msIdentifier> / <idno>	nein	Enthält einen innerhalb der aufbewahrenden Einrichtung eindeutigen Identifikator, in der Regel die Signatur ggf. ergänzt um eine Angabe der Blätter, über die sich die logische Struktur erstreckt.	Pflicht
<msIdentifier> / <msName>	ja	Enthält einen nicht-kanonischen Titel der Handschrift.	Empfohlen

2 Angaben zur Entstehung

Alle in diesem Abschnitt beschriebenen Elemente sind Unterelemente von `teiHeader/fileDesc/sourceDesc/msDesc/history`.

<origin> / <origPlace>	nein	Enthält den Entstehungsort der Handschrift wie er im DFG-Viewer erscheinen soll.	Empfohlen
<origin> / <origDate>	nein	Enthält Angaben zur Entstehungszeit der Handschrift wie sie im DFG-Viewer angezeigt werden sollen. Exakte numerische Angaben werden in den Attributen *when*, *notBefore*, *notAfter*, *from* und *to* gemacht, wenn möglich.	Empfohlen

3 Angaben zur physischen Beschaffenheit

Alle in diesem Abschnitt beschriebenen Elemente sind Unterelemente von `teiHeader/fileDesc/sourceDesc/msDesc/physDesc/objectDesc`.

<supportDesc material=“…“>	nein	Die Angabe zum Material erfolgt im Attribut `material`. Die Verwendung der Attributwerte ***`paper`***, ***`parchment`***, ***`papyrus`***, sowie ***`mixed`*** wird empfohlen.	Pflicht, wenn vh.
<supportDesc> / <support>	nein	Enthält Angaben zum Beschreibstoff wie sie im DFG-Viewer angezeigt werden sollen.	Pflicht, wenn vh.
<supportDesc> / <extent> / <measure>	nein	Enthält Mengenangaben zum Umfang wie z.B. die Blattzahl, die auch im DFG-Viewer angezeigt wird.	Pflicht, wenn vh.
<supportDesc> / <extent> /<dimensions>	nein	Enthält Angaben zur physischen Abmessung wie etwa das Blattformat, das auch im DFG-Viewer angezeigt wird. Die konkreten Größenangaben erfolgen in den Unterelementen `depth`, `height` sowie `width`.	Pflicht, wenn vh.

4 Angaben zum Inhalt

Alle in diesem Abschnitt beschriebenen Elemente sind Unterelemente von `teiHeader/fileDesc/sourceDesc/msDesc/msContents`.

<summary>	nein	Enthält eine summarische Inhaltsbeschreibung.	Empfohlen
<textLang mainLang="…" otherLangs="…">	nein	Enthält Angaben zu den Sprachen, in denen die Texte der Handschrift verfasst sind. Das Attribut `mainLang` sollte die überwiegend verwendete Sprache benennen, während alle weiteren vorkommenden Sprachen jeweils mit Leerzeichen getrennt im Attribut `otherLangs` kodiert werden sollten. Die Attributwerte müssen Sprachcodes nach den Standards ISO 639-2 oder ISO 639-3 sein.	Empfohlen

5 Administrative Angaben zum Digitalisat

Alle in diesem Abschnitt beschriebenen Elemente sind Unterelemente von `teiHeader/fileDesc`.

<titleStmt> / <title>	nein	Enthält einen Titel für das Digitalisat. In der Regel ist dies die Schlagzeile nach DFG-Richtlinien bestehend aus Signatur der digitalisierten Handschrift und deren Titel. Dieses Feld ist nicht mit dem Titel zu verwechseln und kommt im DFG-Viewer nicht zur Anzeige! Die Verpflichtung ergibt sich aus der allgemeinen TEI-Spezifikation und ist nicht Viewer-spezifisch.	Pflicht
<titleStmt> / <funder>	ja	Enthält einen für die Digitalisierung relevanten Geldgeber, also etwa „Deutsche Forschungsgemeinschaft“. Das kann aber z.B. auch die digitalisierende Institution sein.	Empfohlen
<publicationStmt> / <publisher>	ja	Enthält den Herausgeber des Digitalisats, in aller Regel also die digitalisierende Einrichtung. Mindestens ein Unterelement von `publicationStmt` ist verpflichtend.	Pflicht, wenn vh.
<publicationStmt> / <pubPlace> / <ptr target="…" cRef="…" type="…" />	ja	Enthält einen persistenten Identifier des Digitalisats. Handelt es sich beim Identifier um eine URL, so ist diese im Attribut `target` anzugeben, andernfalls im Attribut `cRef`. Darüber hinaus muss die Art des Identifiers im Attribut `type` mit den Werten ***`urn`*** oder ***`purl`*** angege-	Pflicht, wenn vh.

		ben werden. Mindestens ein Unterelement von `publicationStmt` ist verpflichtend.	
<publicationStmt> / <idno>	ja	Enthält einen Identifier des Datensatzes. Optional kann dessen Typ im Attribut `type` angegeben werden.	Pflicht, wenn vh.

3. Beispiel für einen METS/TEI Datensatz nach dem DFG-Standard

Das folgende Beispiel zeigt den METS/TEI-Datensatz aufgrund seines Umfangs nur ausschnittsweise. Der Auszug stellt einen vollständigen deskriptiven Metadatensatz dar, der alle oben beschriebenen Pflicht- und Optionalfelder exemplarisch enthält. Dieser Auszug befindet sich in einer METS-Datei innerhalb eines `dmdSec`-Abschnitts.

```
<tei:teiHeader>
    <tei:fileDesc>
        <tei:titleStmt>
            <!-- Fingierter Titel für das Digitalisat (obligatorisch)
                 Es handelt sich hierbei nicht um den bibliografischen Titel der
                 Handschrift, sondern um die nach DFG-Richtlinien gebildete Schlagzeile
                 Bestehend aus Signatur und wissenschaftlichem Titel. Der Titel kommt im
                 DFG-Viewer nicht zur Anzeige, ist aber aufgrund der TEI-
                 Spezifikation dennoch verpflichtend. -->
            <tei:title>
                Zeugbuch Kaiser Maximilians I. - BSB Cod.icon. 222
            </tei:title>
            <!-- Förderer der Digitalisierung (optional) -->
            <tei:funder>Deutsche Forschungsgemeinschaft</tei:funder>
        </tei:titleStmt>
        <tei:publicationStmt>
            <!-- Herausgeber des Digitalats (obligatorisch, sofern bekannt) -->
            <tei:publisher>Bayerische Staatsbibliothek</tei:publisher>
            <!-- Persistenter Identifikator für das Digitalisat (obligatorisch) -->
            <tei:pubPlace>
                <tei:ptr
                    target="http://daten.digitale-sammlungen.de/~db/0002/bsb00020956/"
                    type="purl"/>
            </tei:pubPlace>
            <!—Identifikator für den Datensatz (obligatorisch) -->
            <tei:idno type="zend">bsb00020956</tei:idno>
        </tei:publicationStmt>
        <tei:sourceDesc>
            <tei:msDesc>
                <tei:msIdentifier>
                    <!-- Geographischer Aufbewahrungsorts (obligatorisch) -->
                    <tei:settlement>München</tei:settlement>
                    <!-- Name der bewahrenden Institution (obligatorisch) -->
                    <tei:repository>Bayerische Staatsbibliothek</tei:repository>
                    <!-- Signatur o. eindeutiger Identifikator (obligatorisch) -->
                    <tei:idno>Cod.icon. 222</tei:idno>
                    <!-- Nicht-kanonischer Titel (optional, wiederholbar) -->
                    <tei:msName>Zeugbuch Kaiser Maximilians I.</tei:msName>
                </tei:msIdentifier>
                <tei:head>
                    <!-- Handschriftentitel (obligatorisch, sofern vorhanden) -->
                    <tei:title/>
                </tei:head>
                <tei:history>
                    <tei:origin>
                        <!-- Entstehungsort (empfohlen, sofern bekannt) -->
                        <tei:origPlace>Innsbruck</tei:origPlace>
                        <!-- Entstehungszeit (empfohlen, sofern bekannt) -->
                        <tei:origDate notBefore="1500" notAfter="1505">
                            um 1502
                        </tei:origDate>
                    </tei:origin>
                </tei:history>
```

```
                <tei:physDesc>
                    <tei:objectDesc>
                        <!-- Angabe des Materials (obligatorisch, sofern ermittelbar) -->
                        <tei:supportDesc material="mixed">
                            <!-- Angaben zum Beschreibstoff (obligatorisch, sofern
                                 ermittelbar) -->
                            <tei:support>
                                Pergament (2-8, 11-30, 34, 48, 295, 296) und Papier
                            </tei:support>
                            <tei:extent>
                              <!-- Angaben zum Umfang wie z.B. die Blattzahl
                                   (obligatorisch, sofern ermittelbar) -->
                              <tei:measure type="leavesCount">2 + 276 Blätter</tei:measure>
                              <!-- Blattformat (obligatorisch, sofern ermittelbar) -->
                              <tei:dimensions unit="mm" type="leaves">
                                <tei:height>420</tei:height>
                                <tei:width>285</tei:width>
                              </tei:dimensions>
                            </tei:extent>
                        </tei:supportDesc>
                    </tei:objectDesc>
                </tei:physDesc>
                <tei:msContents>
                    <!-- Summarische Beschreibung des Inhalts (empfohlen) -->
                    <tei:summary/>
                    <!-- Textsprache (empfohlen, sofern bekannt) -->
                    <tei:textLang mainLang="ger">Deutsch</tei:textLang>
                </tei:msContents>
            </tei:msDesc>
        </tei:sourceDesc>
    </tei:fileDesc>
</tei:teiHeader>
```

Der übrige Aufbau der METS-Datei folgt analog der Dokumentation für Drucke. Innerhalb des `structMap`-Elements mit dem Attribut `TYPE=LOGICAL` wird in der METS-Datei die logische Struktur der Handschrift kodiert. Jede logische Struktureinheit (z.B. ein Textfragment) kann wiederum auf einen eigenen Abschnitt mit deskriptiven Metadaten verweisen, der ebenfalls dem oben beschriebenen Aufbau folgen muss.

Anhang C:
LIDO-Kernelemente für die Publikation

1. LIDO-Kernelemente für bildhafte und dreidimensionale Objekte

Der LIDO-Standard fokussiert als Harvestingformat auf die Bereitstellung von deskriptiven und administrativen Metadaten für die Recherche und Präsentation von in der Regel unikalen, bildhaften und dreidimensionalen Objekten in Online-Umgebungen.

Es sind nur wenige Pflichtfelder erforderlich. Für unterschiedliche Objektarten kann eine Ausdifferenzierung der allgemeingültigen Beschreibungselemente durch ein type-Attribut erfolgen, jedoch immer unter Beibehaltung der im Standard definierten Semantik eines Elements. Je nach Materialart und Projektzuschnitt ist eine Erweiterung dieses Basissets angezeigt.

Für weitere Erläuterungen zur Besetzung der jeweiligen Felder sei auf die LIDO-Spezifikation sowie weitergehende Informationen auf der LIDO-Webseite unter http://www.lido-schema.org verwiesen.

Vorbemerkung: Zahlreiche LIDO-Elemente (z.B. Objekttyp, Material, Technik) werden mit Datenwerten belegt, die idealerweise als Konzepte („unit of thought“) in kontrollierten Vokabularen (Thesauri, Klassifikationssystem, Wortlisten)[1] definiert sind. Solche Elemente, die im Folgenden in der jeweiligen Erläuterung als *Konzept* markiert werden, bestehen immer aus den Unterelementen *conceptID* zur Angabe des Identifikators im kontrollierten Vokabular und *term* zur Angabe einer Bezeichnung für das Konzept.

Deskriptive Metadaten

Element /Unterelement	***Wieder-holbar***	***Bemerkungen /Erläuterungen***	***Status***
1 LIDO Identifikationsnummer			
<lidoRecordID>	ja	Eindeutige Identifikationsnummer für den LIDO Datensatz, typischerweise zusammengesetzt aus der ISIL-Nummer des Datenlieferanten sowie der lokalen Datensatznummer. (Zu ISIL-Nummern für Museen siehe http://www.museen-in-deutschland.de/)	Pflicht

[1] Dies können sowohl publizierte, zentral gepflegte als auch unpublizierte, lokal gepflegte Vokabulare sein. Im Sinne der Vernetzung ist ersteren i.d.R. der Vorzug zu geben.

Element /Unterelement	*Wieder-holbar*	*Bemerkungen /Erläuterungen*	*Status*
2 Objekttyp			
<objectWorkType>	ja	Objekt- oder Werktyp Konzept	Pflicht
3 Klassifikation			
<classification type="…">	ja	Ein Begriff / Konzept, das ein Objekt in einem größeren Kontext kategorisiert. Das (optionale) type-Attribut qualifiziert die Art der Klassifikation, z.B. Gattung, biologische Systematik o.ä. Konzept	Pflicht, wenn vh.
4 Titel/Objektname			
<titleSet>	ja	Umschlag für Titelinformationen; sofern das Objekt keinen Titel hat, muss er generiert werden, z.B. aus dem Objekttyp.	Pflicht
< titleSet> / <appellationValue>	nur für Sprach-varianten	Titel oder Name, der einem Objekt gegeben wurde.	
5 Aufbewahrung/Standort			
<repositorySet type="current">	ja	Umschlag für die aktuelle Aufbewahrungs- bzw. Standort-Information	Pflicht, wenn vh.
<repositorySet> / <repositoryName> / <legalBodyName> / <appellationValue>	nur für Sprach-varianten	Bezeichnung für die aufbewahrende Institution	
<repositorySet> / <workID>	ja	Eine eindeutige numerische oder alphanumerische Identifikationsnummer, die die aufbewahrende Institution dem Objekt gegeben hat; typischerweise die Inventarnummer	
<repositorySet> / <repositoryLocation> / <namePlaceSet> / <appellationValue>	nur für Sprach-varianten	Bezeichnung des Standorts (insbes. für Architektur Pflichtangabe)	

Element /Unterelement	*Wieder-holbar*	*Bemerkungen /Erläuterungen*	*Status*
6 Objektbeschreibung			
<objectDescriptionSet type="…">	ja	Umschlag für Beschreibungstexte und ihre Quellen. Das (optionale) type-Attribut qualifiziert die Art der Beschreibung, z.B. Objektgeschichte, Katalogtext o.ä.	Empfohlen
< objectDescriptionSet> / <descriptiveNoteValue>	nur für Sprach-varianten	Kurzer, essay-ähnlicher Text, der das Objekt beschreibt.	
7 Maßangaben			
<objectMeasurementsSet>	ja	Umschlag für Maßangaben	Pflicht, wenn vh.
<objectMeasurementsSet> / <displayObjectMeasurements>	nur für Sprach-varianten	Maßangabe in Textform	
<objectMeasurementsSet> / <objectMeasurements> / <measurementsSet>	nein	Maßangabe in strukturierter Form mit den Unterelementen measurementType, measurementUnit, measurementValue	Empfohlen
8 Ereignis			
<eventSet>	ja	Umschlag für Informationen über ein Ereignis im Lebenszyklus des Objekts	Pflicht, wenn vh.
<eventSet> / <event> / <eventType>	nein	Art der Beziehung zwischen Objekt und Ereignis. Je nach Objektart sind typische Ereignisse: Herstellung, Fund, Nutzung. Für empfohlene Ereignistypen sieht http://terminology.lido-schema.org/eventType Konzept	

Element /Unterelement	***Wieder-holbar***	***Bemerkungen /Erläuterungen***	***Status***
Am Ereignis beteiligte identifizierbare Akteure			
<eventSet> / <event> / <eventActor>	ja	Umschlag für Information zur Beteiligung einer Person, Institution, Gruppe, Familie am Ereignis	Pflicht, wenn vh.
<eventSet> / <event> / <eventActor> / <actorInRole> / <actor> / <actorID>	ja	Identifikator für den beteiligten Akteur. Referenziert idealerweise eine Normdatei. Für DFG-Projekte soll die GND verwendet werden.	
<eventSet> / <event> / <eventActor> / <actorInRole> / <actor> / <nameActorSet> / <appellationValue>	nur für Namens-varianten	Name des beteiligten Akteurs	
<eventSet> / <event> / <eventActor> / <actorInRole> / <attributionQualifierActor>	ja	Information zur Zuschreibung eines Objekts zu einem Beteiligten; besonders bei Künstlern relevant	Empfohlen
Kultureller Kontext des Ereignisses			
<culture>	ja	Kultureller Kontext zum Ereignis; vor allem relevant für Herstellung und Nutzung, wenn kein konkretes Individuum als Akteur benannt werden kann Konzept	Pflicht, wenn vh.
Datierung des Ereignisses			
<eventDate>	nein	Datierung des Ereignisses	Pflicht, wenn vh.
<eventDate > / <displayDate>	nur für Sprach-varianten	Datierungsangabe in Textform, in der ggf. Ungenauigkeiten zum Ausdruck gebracht werden können.	
<eventDate > / <earliestDate>	nein	Jahresangabe oder exaktes Datum, das den frühestmöglichen Zeitpunkt angibt, zu dem das Ereignis stattgefunden resp. begonnen hat.	
<eventDate > / <latestDate>	nein	Jahresangabe oder exaktes Datum, das den letztmöglichen Zeitpunkt angibt, zu dem das Ereignis stattgefunden resp. geendet hat.	

Element /Unterelement	***Wieder-holbar***	***Bemerkungen /Erläuterungen***	***Status***
Zeitperiode zum Ereignis			
<periodName>	nur für Angabe frühester / spätester Periode	Zeitperiode zum Ereignis; vor allem relevant im archäologischen und naturhistorischen Kontext Konzept	Pflicht, wenn vh.
Ortsangaben zum Ereignis			
<eventPlace>	ja	Umschlag für Ortsinformation zum Ereignis	Pflicht, wenn vh.
<eventPlace > / <place> / <namePlaceSet> <appellationValue>	nur für Namens-varianten	Name des Ortes	
<eventPlace > / <place> / <gml>	ja	Georeferenzen des Ortes	Empfohlen
Material- und Technikangaben zum Ereignis			
<eventMaterialsTech>	ja	Umschlag für Material- und Technikangaben zum Ereignis; kann für verschiedene Objektteile differenziert werden.	Pflicht, wenn vh.
< eventMaterialsTech > / <materialsTech> / <termMaterialsTech type=“…“>	ja	Ein Material oder eine Technik Das type-Attribut qualifiziert Material oder Technik Konzept	

9 Thema / Bildinhalt

<subjectSet>	ja	Umschlag für Informationen zum Thema oder Bildinhalt eines Objektes Beachte: Unterschiedliche Entitäten (Konzept, Akteur, Ort, Ereignis, Objekt) haben je eigene Unterelemente.	Pflicht, wenn vh.
<subjectSet> / <subject> / <subjectConcept>	ja	Verschlagwortung der Themen oder Bildinhalte eines Objekts Kontrolliert z.B. ICONCLASS Konzept	

Administrative Metadaten

10 Rechte am Objekt

<rightsWorkSet>	ja	Umschlag für Rechte-Information zum Objekt / Werk.	Pflicht, wenn vh.
<rightsWorkSet > / <rightsHolder> / <legalBodyName>/ <appellationValue>	nur für Sprach-varianten	Name des Rechteinhabers.	

11 Datensatz

<recordID>	ja	Eindeutige Identifikationsnummer im (lokalen) System des Datenlieferanten.	Pflicht
<recordType>	nein	Angabe, ob der Datensatz ein Einzelobjekt, eine Sammlung, Serie, Objektgruppe o.a. beschreibt. Konzept.	Pflicht
< recordSource>	ja	Datensatzquelle, i.d.R. die datenliefernde Institution	Pflicht
<recordSource> / <legalBodyName>/ <appellationValue>	nur für Sprach-varianten	Name zur Datensatzquelle.	
<recordRights>	ja	Umschlag für Rechte-Information zum Datensatz.	Pflicht, wenn vh.
<recordRights> / <rightsHolder> / <legalBodyName>/ <appellationValue>	nur für Sprach-varianten	Name des Rechteinhabers.	
<recordInfoSet> / <recordInfoLink>	ja	Link zur Objektbeschreibung.	Pflicht, wenn vh.

12 Reproduktionen

<resourceSet>	ja	Umschlag für Angaben zu (digitalen) Reproduktionen des Objekts.	Pflicht, wenn vh.
<resourceSet> / <resourceRepresentation >	ja	Link(s) zur digitalen Reproduktion des Objekts, , ggf. in verschiedenen Auflösungen und mit technischen Angaben zur Wiedergabe (Audio, Video)	
<resourceSet> / <rightsResource> / <rightsHolder>	ja	Umschlag für Rechte-Information zur digitalen Reproduktion.	
<resourceSet> / <rightsResource> / <rightsHolder> / <legalBodyName>/ <appellationValue>	nur für Sprach-varianten	Name des Rechteinhabers.	

2. Beispiel für einen LIDO Datensatz

```
<?xml version="1.0" encoding="UTF-8"?>
<lido:lido xmlns:lido="http://www.lido-schema.org" xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance" xsi:schemaLocation="http://www.lido-schema.org http://www.lido-schema.org/schema/v1.0/lido-v1.0.xsd">
 <lido:lidoRecID lido:type="local" lido:source="Deutsches Dokumentationszentrum für Kunstgeschichte - Bildarchiv Foto Marburg">DE-Mb112/lido/obj/00154983</lido:lidoRecID>
 <lido:descriptiveMetadata xml:lang="de">
  <lido:objectClassificationWrap>
   <lido:objectWorkTypeWrap>
    <lido:objectWorkType lido:type="Sachbegriff">
     <lido:term>Gemälde</lido:term>
    </lido:objectWorkType>
   </lido:objectWorkTypeWrap>
   <lido:classificationWrap>
    <lido:classification lido:type="Gattung">
     <lido:term>Tafelmalerei</lido:term>
    </lido:classification>
   </lido:classificationWrap>
  </lido:objectClassificationWrap>
  <lido:objectIdentificationWrap>
   <lido:titleWrap>
    <lido:titleSet>
     <lido:appellationValue lido:pref="preferred">La Primavera / Der Frühling</lido:appellationValue>
    </lido:titleSet>
   </lido:titleWrap>
   <lido:repositoryWrap>
    <lido:repositorySet lido:type="current">
     <lido:repositoryName>
      <lido:legalBodyName>
       <lido:appellationValue>Galleria degli Uffizi — Pinacoteca (Florenz)</lido:appellationValue>
      </lido:legalBodyName>
     </lido:repositoryName>
     <lido:workID lido:type="Inventarnummer">8360 (Inv. 1890)</lido:workID>
     <lido:repositoryLocation>
      <lido:namePlaceSet>
       <lido:appellationValue>Florenz</lido:appellationValue>
      </lido:namePlaceSet>
     </lido:repositoryLocation>
    </lido:repositorySet>
   </lido:repositoryWrap>
   <lido:objectMeasurementsWrap>
    <lido:objectMeasurementsSet>
     <lido:displayObjectMeasurements>Höhe x Breite: 203 x 314 cm</lido:displayObjectMeasurements>
     <lido:objectMeasurements>
      <lido:measurementsSet>
       <lido:measurementType>Höhe x Breite</lido:measurementType>
       <lido:measurementUnit>cm</lido:measurementUnit>
       <lido:measurementValue>203 x 314</lido:measurementValue>
      </lido:measurementsSet>
     </lido:objectMeasurements>
    </lido:objectMeasurementsSet>
   </lido:objectMeasurementsWrap>
  </lido:objectIdentificationWrap>
  <lido:eventWrap>
   <lido:eventSet>
    <lido:event>
     <lido:eventType>
```

```
      <lido:conceptID lido:type="URI" lido:source="http://terminology.lido-
schema.org/eventType">http://terminology.lido-schema.org/lido00007</lido:conceptID>
       <lido:term>Herstellung</lido:term>
      </lido:eventType>
      <lido:eventActor>
       <lido:actorInRole>
        <lido:actor lido:type="person">
         <lido:actorID lido:type="local" lido:source="Bildindex">kue 02553338</lido:actorID>
         <lido:actorID lido:type="URL"
lido:source="http://www.bildindex.de">http://www.bildindex.de/dokumente/html/kue02553338</lido:actor
ID>
         <lido:actorID lido:type="URI" lido:source="http://d-nb.info/gnd">http://d-
nb.info/gnd/118514008</lido:actorID>
         <lido:nameActorSet>
          <lido:appellationValue lido:pref="preferred">Botticelli, Sandro</lido:appellationValue>
         </lido:nameActorSet>
         <lido:nameActorSet>
          <lido:appellationValue lido:pref="alternative">Filipepi, Alessandro</lido:appellationValue>
         </lido:nameActorSet>
         <lido:nameActorSet>
          <lido:appellationValue lido:pref="alternative">Filipepi, Sandro</lido:appellationValue>
         </lido:nameActorSet>
         <lido:nationalityActor>
          <lido:term>Italien</lido:term>
         </lido:nationalityActor>
         <lido:vitalDatesActor>
          <lido:earliestDate lido:type="evidenceDate">1445</lido:earliestDate>
          <lido:latestDate lido:type="evidenceDate">1510-05-17</lido:latestDate>
         </lido:vitalDatesActor>
         <lido:genderActor>männlich</lido:genderActor>
        </lido:actor>
        <lido:roleActor>
         <lido:term>Maler</lido:term>
        </lido:roleActor>
       </lido:actorInRole>
      </lido:eventActor>
      <lido:eventDate>
       <lido:displayDate>um 1482</lido:displayDate>
       <lido:date>
        <lido:earliestDate>1472</lido:earliestDate>
        <lido:latestDate>1492</lido:latestDate>
       </lido:date>
      </lido:eventDate>
      <lido:eventMaterialsTech>
       <lido:materialsTech>
        <lido:termMaterialsTech lido:type="Material">
         <lido:term>Tempera</lido:term>
        </lido:termMaterialsTech>
        <lido:termMaterialsTech lido:type="Material">
         <lido:term>Pappelholz</lido:term>
        </lido:termMaterialsTech>
       </lido:materialsTech>
      </lido:eventMaterialsTech>
     </lido:event>
    </lido:eventSet>
    <lido:eventSet>
     <lido:event>
      <lido:eventType>
       <lido:conceptID lido:type="URI" lido:source="http://terminology.lido-
schema.org/eventType">http://terminology.lido-schema.org/lido00034</lido:conceptID>
       <lido:term>Restaurierung</lido:term>
```

```
    </lido:eventType>
    <lido:eventDate>
     <lido:displayDate>1982</lido:displayDate>
     <lido:date/>
    </lido:eventDate>
   </lido:event>
  </lido:eventSet>
 </lido:eventWrap>
 <lido:objectRelationWrap>
  <lido:subjectWrap>
   <lido:subjectSet>
    <lido:subject lido:type="Ikonographie">
     <lido:subjectConcept>
      <lido:conceptID lido:type="URI"
lido:source="http://iconclass.org">http://iconclass.org/23D42</lido:conceptID>
      <lido:term lido:pref="preferred">23 D 42</lido:term>
      <lido:term lido:pref="alternative">Frühling; Ripa: Ver; Primavera</lido:term>
     </lido:subjectConcept>
    </lido:subject>
   </lido:subjectSet>
   <lido:subjectSet>
    <lido:subject lido:type="Ikonographie">
     <lido:subjectConcept>
      <lido:conceptID lido:type="URI"
lido:source="http://iconclass.org">http://iconclass.org/57A9:</lido:conceptID>
      <lido:term lido:pref="preferred">57 A 9</lido:term>
      <lido:term lido:pref="alternative">Menschlichkeit, Höflichkeit; Ripa: Cortesia, Humanità</lido:term>
     </lido:subjectConcept>
    </lido:subject>
   </lido:subjectSet>
   <lido:subjectSet>
    <lido:subject lido:type="Ikonographie">
     <lido:subjectConcept>
      <lido:conceptID lido:type="URI"
lido:source="http://iconclass.org">http://iconclass.org/92C47</lido:conceptID>
      <lido:term lido:pref="preferred">92 C 47</lido:term>
      <lido:term lido:pref="alternative">spezifische Darstellungsformen, allegorische Darstellungsformen;
Venus als Schutzgottheit</lido:term>
     </lido:subjectConcept>
    </lido:subject>
   </lido:subjectSet>
   <lido:subjectSet>
    <lido:subject lido:type="Ikonographie">
     <lido:subjectConcept>
      <lido:conceptID lido:type="URI"
lido:source="http://iconclass.org">http://iconclass.org/92D3</lido:conceptID>
      <lido:term lido:pref="preferred">92 D 3</lido:term>
      <lido:term lido:pref="alternative">Grazien (Chariten), meist als Dreiergespann; Ripa:
Gratie</lido:term>
     </lido:subjectConcept>
    </lido:subject>
   </lido:subjectSet>
   <lido:subjectSet>
    <lido:subject lido:type="Ikonographie">
     <lido:subjectConcept>
      <lido:conceptID lido:type="URI"
lido:source="http://iconclass.org">http://iconclass.org/92E5421</lido:conceptID>
      <lido:term lido:pref="preferred">92 E 54 21</lido:term>
      <lido:term lido:pref="alternative">Zephyrus entführt Flora (oder Chloris), die in der Regel Blumen
streut</lido:term>
     </lido:subjectConcept>
```

```
    </lido:subject>
   </lido:subjectSet>
   <lido:subjectSet>
    <lido:subject lido:type="Ikonographie">
     <lido:subjectConcept>
      <lido:conceptID lido:type="URI"
lido:source="http://iconclass.org">http://iconclass.org/96A2351</lido:conceptID>
      <lido:term lido:pref="preferred">96 A 23 51</lido:term>
      <lido:term lido:pref="alternative">Flora in ihrem Königreich: ein Garten voller Blumen, der ihr von
Zephyrus geschenkt worden ist</lido:term>
     </lido:subjectConcept>
    </lido:subject>
   </lido:subjectSet>
   <lido:subjectSet>
    <lido:subject lido:type="Ikonographie">
     <lido:subjectConcept>
      <lido:conceptID lido:type="URI"
lido:source="http://iconclass.org">http://iconclass.org/92D1521</lido:conceptID>
      <lido:term lido:pref="preferred">92 D 15 21</lido:term>
      <lido:term lido:pref="alternative">Cupido schießt einen Pfeil</lido:term>
     </lido:subjectConcept>
    </lido:subject>
   </lido:subjectSet>
  </lido:subjectWrap>
 </lido:objectRelationWrap>
</lido:descriptiveMetadata>
<lido:administrativeMetadata xml:lang="de">
 <lido:rightsWorkWrap>
  <lido:rightsWorkSet>
   <lido:rightsHolder>
    <lido:legalBodyName>
     <lido:appellationValue>Galleria degli Uffizi — Pinacoteca (Florenz)</lido:appellationValue>
    </lido:legalBodyName>
   </lido:rightsHolder>
  </lido:rightsWorkSet>
 </lido:rightsWorkWrap>
 <lido:recordWrap>
  <lido:recordID lido:type="local">obj/00154983</lido:recordID>
  <lido:recordType>
   <lido:term>Einzelobjekt</lido:term>
  </lido:recordType>
  <lido:recordSource>
   <lido:legalBodyID lido:type="URI" lido:source="ISIL (ISO 15511)">info:isil/DE-
Mb112</lido:legalBodyID>
   <lido:legalBodyName>
    <lido:appellationValue>Deutsches Dokumentationszentrum für Kunstgeschichte - Bildarchiv Foto
Marburg</lido:appellationValue>
   </lido:legalBodyName>
   <lido:legalBodyWeblink>http://www.fotomarburg.de</lido:legalBodyWeblink>
  </lido:recordSource>
  <lido:recordRights>
   <lido:rightsType>
    <lido:conceptID
lido:type="URI">http://creativecommons.org/publicdomain/zero/1.0/</lido:conceptID>
    <lido:term>CC0 1.0 Public Domain Dedication</lido:term>
   </lido:rightsType>
   <lido:rightsHolder>
    <lido:legalBodyID lido:type="URI" lido:source="ISIL (ISO 15511)">info:isil/DE-
Mb112</lido:legalBodyID>
    <lido:legalBodyName>
```

```
      <lido:appellationValue>Deutsches Dokumentationszentrum für Kunstgeschichte - Bildarchiv Foto
Marburg</lido:appellationValue>
     </lido:legalBodyName>
    </lido:rightsHolder>
   </lido:recordRights>
   <lido:recordInfoSet>
    <lido:recordInfoLink
lido:formatResource="html">http://www.bildindex.de/dokumente/html/obj00154983</lido:recordInfoLink>
   </lido:recordInfoSet>
   <lido:recordInfoSet>
    <lido:recordInfoLink lido:formatResource="xml">http://oai.bildindex.de/db/apsisa.dll/oai-
pmh?verb=GetRecord&amp;identifier=DE-
Mb112/lido/obj/00154983&amp;metadataPrefix=lido</lido:recordInfoLink>
   </lido:recordInfoSet>
  </lido:recordWrap>
  <lido:resourceWrap>
   <lido:resourceSet>
    <lido:resourceID lido:type="local">C 654.591</lido:resourceID>
    <lido:resourceRepresentation lido:type="image_thumb">
     <lido:linkResource
lido:formatResource="jpg">http://www.bildindex.de/bilder/fmc654591b.jpg</lido:linkResource>
    </lido:resourceRepresentation>
    <lido:resourceType>
     <lido:term>Negativ</lido:term>
    </lido:resourceType>
    <lido:resourceDateTaken>
     <lido:displayDate>1949</lido:displayDate>
    </lido:resourceDateTaken>
    <lido:rightsResource>
     <lido:rightsType>
      <lido:conceptID lido:type="URI">http://www.europeana.eu/rights/rr-f/</lido:conceptID>
      <lido:term>Rights Reserved - Free Access</lido:term>
     </lido:rightsType>
     <lido:rightsHolder>
      <lido:legalBodyID lido:type="URI" lido:source="ISIL (ISO 15511)">info:isil/DE-
Mb112</lido:legalBodyID>
      <lido:legalBodyName>
       <lido:appellationValue>Deutsches Dokumentationszentrum für Kunstgeschichte - Bildarchiv Foto
Marburg</lido:appellationValue>
      </lido:legalBodyName>
     </lido:rightsHolder>
    </lido:rightsResource>
   </lido:resourceSet>
  </lido:resourceWrap>
 </lido:administrativeMetadata>
</lido:lido>
```

Anhang D:
Sammlungs- und Bestandsbeschreibung

1. Kernelemente für die Sammlungs- und Bestandsbeschreibung

Sammlungs- und Bestandsbeschreibungen können gemäß Dublin Core Collections Application Profile[1] oder in demselben Metadatenstandard erfolgen, in dem auch die Objektbeschreibungen verfügbar gemacht werden: METS, MODS oder TEI-Header, EAD/DDB sowie LIDO bieten entsprechende Möglichkeiten.

In Anlehnung an die im ISO-Standard ISO 27730, Information and Documentation – International Standard Collection Identifier (ISCI) vorgeschriebenen Beschreibungsfelder für eine Sammlung, welche im Dublin Core Collections Application Profile definiert sind, werden folgende Kernelemente für die Sammlungs- und Bestandsbeschreibung empfohlen:

Element	***Wieder-holbar***	***Bemerkungen / Erläuterungen***	***Status***
<dc:identifier>	ja	Sammlungsidentifikator	Pflicht
<dc:type>	ja	Art der beschriebenen Ressource: Sammlung, Bestand	Pflicht
<dc:title>	ja	Name der Sammlung	Pflicht
<dc:language>	ja	Sprache der Sammlungsobjekte	Pflicht, wenn vh.
<dcterms:abstract >	ja	Zusammenfassende Beschreibung der Sammlung.	Pflicht
<cld:isLocatedAt >	ja	Aufbewahrende Institution der Sammlung	Pflicht
<cld:isAccessedVia >	ja	Angebot für den Zugang zu den Sammlungsobjekten	Pflicht
<dcterms:provenance >	ja	Herkunft der Sammlung	Pflicht, wenn vh.
<dc:subject>	ja	Thema der Sammlung	Empfohlen
<dc:creator>	ja	Sammler	Empfohlen
<cld:itemType>	ja	Art der Einzelobjekte der Sammlung	Empfohlen
<cld:itemFormat>	ja	Format der Einzelobjekte der Sammlung	Empfohlen
<dcterms:hasPart>	ja	Teilsammlung	Empfohlen
<dcterms:isPartOf>	ja	Übergeordnete Sammlung	Empfohlen
<cld:associatedCollection >	ja	Andere Sammlung mit Bezug zur beschriebenen Sammlung	Empfohlen

[1] http://dublincore.org/groups/collections/collection-application-profile/

2.1 Erstes Beispiel einer Sammlungsbeschreibung im Dublin Core Collections Application Profile

```
<?xml version="1.0" encoding="UTF-8"?>
<cld:cld xmlns:cld="http://purl.org/cld/terms/" xmlns:dc="http://purl.org/dc/elements/1.1/"
xmlns:dcterms="http://purl.org/dc/terms/" xmlns:marcrel="http://www.loc.gov/loc.terms/relators/"
xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance">
        <dc:identifier>DE-Mb112/Bickell1977</dc:identifier>
        <dc:type>Sammlung</dc:type>
        <dc:title>Bickell, Ludwig</dc:title>
        <dcterms:abstract>Ludwig Bickell (1838-1901) war seit 1892 Bezirkskonservator von
Kurhessen und außerdem Begründer des Marburger Universitätsmuseums. Der Bestand umfaßt seinen
fotografischen Nachlaß, etwa 3.000 Glasplatten (teilweise 30 x 40), überwiegend aus dem 19.
Jahrhundert. Im Archiv der Archive befinden sich zusätzlich 5 Abzüge der Photographischen Anstalt
L.BICKELL, weitere konfektionierte Vintageabzüge (Großformat) von 1893 im
Plattenarchiv.</dcterms:abstract>
        <cld:isLocatedAt>Deutsches Dokumentationszentrum für Kunstgeschichte - Bildarchiv Foto
Marburg</cld:isLocatedAt>
        <cld:isAccessedVia>http://www.bildindex.de/?+qfotograf:"bickell ludwig"</cld:isAccessedVia>
        <dcterms:provenance>Marburg, Verein für hessische Geschichte und Landeskunde,
Zweigverein Marburg e. V.</dcterms:provenance>
        <dcterms:provenance>Der Bestand wurde dem Bildarchiv vom &amp;quot;Verein für hessische
Geschichte und Landeskunde, Zweigverein Marburg e. V.&amp;quot; 1977 als Dauerleihgabe anvertraut.
Hinweis: Der sonstige Nachlaß wird im Hessischen Staatsarchiv Marburg verwaltet. (Staatsarchiv
Marburg, StaM 340, Bickell)</dcterms:provenance>
        <dc:subject>Kunst</dc:subject>
        <dc:subject>Architektur</dc:subject>
        <dc:subject>Hessen</dc:subject>
        <dc:creator>Bickell, Ludwig</dc:creator>
        <cld:itemType>Glasnegativ</cld:itemType>
        <cld:itemType>Abzug</cld:itemType>
        <cld:dateItemsCreated>1869-1900</cld:dateItemsCreated>
        <cld:itemFormat>18 x 24</cld:itemFormat>
        <cld:itemFormat>30 x 40</cld:itemFormat>
</cld:cld>
```

2.2 Zweites Beispiel einer Sammlungsbeschreibung im Dublin Core Collections Application Profile (Projektbeschreibung)

```
<?xml version="1.0" encoding="UTF-8"?>
<cld:cld xmlns:cld="http://purl.org/cld/terms/" xmlns:dc="http://purl.org/dc/elements/1.1/"
xmlns:dcterms="http://purl.org/dc/terms/" xmlns:marcrel="http://www.loc.gov/loc.terms/relators/"
xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance">
        <dc:identifier>http://diglib.hab.de/?link=023</dc:identifier>
        <dc:type>Virtuelle Sammlung</dc:type>
        <dc:title>Helmstedter Drucke Online</dc:title>
        <dcterms:abstract>Ziel des Projektes ist die Image- und Volltextdigitalisierung der in Helmstedt
gedruckten Werke, die sich in der Herzog August Bibliothek befinden. Die Herzog August Bibliothek
besitzt etwa 10.000 Drucke, die bis zur Schließung der Universität Helmstedt 1810 in Helmstedt
gedruckt worden sind. In der ersten zweijährigen Phase des Projektes sollen die auf das 16. und 17. Jh.
entfallenden Drucke als Image präsentiert, mit Strukturdaten zur Navigation versehen und im
Verbundkatalog sowie im VD 16 und VD 17 nachgewiesen werden. Zugleich wird erstmals
experimentell eine größere Menge von 120.000 Seiten per OCR bearbeitet und im Volltext angeboten
werden. In einer zweiten Phase soll auch der aus dem 18 Jh. stammende Anteil digitalisiert, auf
nationalbibliographischem Niveau katalogisiert und zur Nachnutzung einem möglichen VD 18 zur
Verfügung gestellt werden.
Mit der nahezu kompletten Digitalisierung der Produktion eines der wichtigsten norddeutschen
Universitätsdruckorte - Helmstedt findet sich unter den 10 am häufigsten nachgewiesenen Druckorten
im VD 17 - wird nicht nur ein substantieller Beitrag zur Komplettdigitalisierung des deutsche gedruckten
Kulturerbes geleistet, sondern erstmals ein Überblick über die Druckproduktion einer bedeutenden
frühneuzeitlichen Universität geschaffen. Das Projekt flankiert den programmatischen und auf mehrere
Jahre angelegten Forschungsschwerpunkt der HAB zur Erforschung der Universitätsgeschichte
Helmstedts und wird unmittelbar mehreren Forschungsprojekten in diesem Feld zugute
kommen.</dcterms:abstract>
        <cld:isLocatedAt>Herzog August Bibliothek Wolfenbüttel</cld:isLocatedAt>
        <cld:isAccessedVia>http://diglib.hab.de/?link=023</cld:isAccessedVia>
        <dcterms:provenance>ehem. Universitätsbibliothek Helmstedt</dcterms:provenance>
        <cld:itemType>Images</cld:itemType>
                <cld:itemType>Fulltext</cld:itemType>
        <cld:dateItemsCreated>2010-2014</cld:dateItemsCreated>

</cld:cld>
```