

Deutsche
Forschungsgemeinschaft

**Wissenschaftliche
Literaturversorgungs- und
Informationssysteme (LIS):**

**DFG-Praxisregeln
„Digitalisierung“**

Stand: April 2009

DFG

DFG-Praxisregeln „Digitalisierung“ zu den Förderprogrammen der Wissenschaftlichen Literaturversorgungs- und Informationssysteme.
Erarbeitet durch den Unterausschuss für Kulturelle Überlieferung am 12. und 13. Februar 2009,
beschlossen durch den Ausschuss für Wissenschaftliche Bibliotheken und Informationssysteme am 20.03.2009.

© April 2009
Deutsche Forschungsgemeinschaft
Geschäftsstelle: Kennedyallee 40,
D-53175 Bonn
Postanschrift: D-53170 Bonn
Telefon: ++49/(0)228/885-2358
Telefax: ++49/(0)228/885-2272
E-Mail: lis@dfg.de
Internet: <http://www.dfg.de/lis>

Inhalt

Vorbemerkung	3
1. Ziele und Auswahl	4
1.1 Ziele	4
1.2 Auswahl.....	4
1.3 Dublettenprüfung und Datenabgleich bei der Imagedigitalisierung	5
2. Digitalisierung von Druckwerken und unikalene Dokumenten	6
2.1 Bereitstellung der Materialien, konservatorische Prüfung	6
2.2 Imagedigitalisierung	6
2.2.1 Digitalisierungsparameter	6
2.2.2 Dateiformate	8
2.3 Volltextgenerierung	9
2.3.1 Zeichenkodierung	10
2.3.2 Auszeichnung von Strukturdaten	10
2.3.3 Layout.....	10
2.3.4 Texterfassung	11
2.4 Langzeitsicherung	12
2.5 Digitalisierung inhouse oder mit Dienstleistern?.....	14
2.6 Metadaten	15
2.6.1 Sammlungs- und Bestandsbeschreibung, Katalogisierung, Deskriptive Metadaten	15
2.6.2 Strukturelle Metadaten zur Imagedigitalisierung	16
2.7 Austausch und Weitergabe von Metadaten, Bekanntmachung.....	17
3. Zitieren digitaler Drucke und Handschriften, persistente Adressierung	19
4. Bereitstellung der Digitalisate für die Öffentlichkeit	20
4.1 Open Access.....	20
4.2 Mindestanforderungen für die Bereitstellungssysteme digitaler Bibliotheken.....	20
4.2.1 Basisanforderungen und Architektur.....	20
4.2.2 Funktionalitätsanforderungen	21
4.2.3 Technische Mindestanforderungen.....	22
5. Präsentationsstandards („DFG-Viewer“) und Formate (METS/MODS).....	23
6. Prüfliste für Antragsteller und Gutachter	23
6.1 Allgemeine technische Abläufe / Voraussetzungen	25
6.2 Datenqualität und -formate.....	25
6.3 Langzeitsicherung	26
6.4 Zusammenarbeit mit Dienstleistern	27
6.5 Metadaten	27
6.6 Austausch und Weitergabe	27
6.7 Zitieren, persistente Adressierung.....	27
6.8 Bereitstellung der Digitalisate, öffentlich zugängliche Schnittstellen	27
7. Regeln für die Durchführung von Digitalisierungsprojekten.....	29
 Anhänge	
Anhang A: METS/MODS Profil für die Darstellung im DFG Viewer und Übermittlung per OAI	30

Vorbemerkung zu den Praxisregeln „Digitalisierung“

Die Deutsche Forschungsgemeinschaft (DFG) unterstützt im Förderbereich Wissenschaftliche Literaturversorgungs- und Informationssysteme (LIS) Projekte in Deutschland, die dem Aufbau leistungsfähiger, vernetzter und überregional agierender Informationssysteme für alle Fachgebiete der deutschen Forschung dienen. Die Ergebnisse der Projekte sollen für die Wissenschaft frei und dauerhaft zugänglich sein.¹

Ziel der Praxisregeln „Digitalisierung“ im Förderbereich der Wissenschaftlichen Literaturversorgungs- und Informationssysteme ist es, Antragstellern 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 unterstützten Projekte leisten.

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

Im Folgenden wird in den Abschnitten 1 bis 4 eine grundsätzliche und weiter ausholende Einführung in die Probleme und Vorgehensweisen gegeben, die sich für Projekte stellen, die Druckwerke und unikale Dokumente in digitale Form überführen wollen, also für die Mehrzahl der derzeit unternommenen Projekte. Diese Abschnitte wenden sich bewusst auch an Personen, die derartige Projekte neu planen, also unter Umständen über keine detaillierten Vorkenntnisse verfügen. Abschnitt 5 verweist auf erforderliche DFG-Präsentationsstandards und Formate. In Abschnitt 6 sind knapp noch einmal die wichtigsten Anforderungen zusammengefasst. Abweichungen von den dort beschriebenen Regeln sind zwar möglich, müssen aber im Einzelfall begründet werden. Abschnitt 6 nennt auch andere Formen digitaler Ressourcen (insb. AV-Medien), mit denen derzeit nur wenige praktische Erfahrungen vorliegen und die daher in den vorangegangenen Abschnitten nicht behandelt wurden. In Abschnitt 7 werden abschließend wichtige Verfahrens-Regeln für die Durchführung eines DFG-Vorhabens benannt.

Auch wenn dies im Folgenden öfter wiederholt werden wird, ist ein Prinzip so wichtig, dass es explizit vorangestellt werden soll: Die wissenschaftlich motivierte Digitalisierung der kulturellen Überlieferung wird nicht als neue Technik, sondern als Normalfall verstanden. In der Praxis bedeutet dies für die Konzeption von Projekten, dass es zwar weiterhin wichtig ist, Digitalisate zu erzeugen, deren Qualität für die Forschung ohne jeden Widerspruch ausreicht, dass aber gleichzeitig auch darauf zu achten ist, dass die gewählten Vorgehensweisen effektiv und kostenbewusst genug sind, um systematisch auf große Mengengerüste angewendet werden zu können.

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.

1. Ziele und Auswahl

1.1 Ziele

Digitalisierung ist ein wichtiges Instrument für die Erschließung wissenschaftlicher Information, sowohl der gedruckten als auch der handschriftlichen, geworden. Ehemals schwer zugängliche oder konservatorisch heikle Materialien lassen sich bequem von zu Hause aus oder in der Bibliothek bzw. im Archiv am PC nutzen. Die direkte Forschung mit den Quellen hat sich dadurch erheblich erleichtert – bei gleichzeitiger Schonung der kostbaren, bisweilen fragilen Originale. Mit der Digitalisierung von Altbeständen aus Bibliotheken und Beständen aus Archiven werden aber nicht nur leicht verfügbare Kopien ins Netz gestellt, es entsteht darüber hinaus eine Infrastruktur, die das Internet zu einem integralen Forschungsraum auch für die geistes- und kulturwissenschaftliche Forschung macht. Erst die Verbindung mit anderen online verfügbaren Ressourcen, mit Katalogen, Lexika, Bibliographien, Editionen, Sekundärliteratur u.a. nutzt das volle Potential des Netzes aus. Ziel ist also nicht nur das Bereitstellen, sondern auch und vor allem das Vernetzen.

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. Jh. unnötig zeitraubend und übersteuert sein. Diese Empfehlungen gehen von der Vorstellung aus, dass die Digitalisierung für eine wissenschaftliche Informationseinrichtung des 21. Jahrhunderts eine normale Dienstleistung, kein „Elitemerkmal“, darstellt. Der digitale Zugang wird in naher Zukunft die Regel und nicht die Ausnahme bilden. Die Digitalisierung hervorragender Einzelstücke oder bedeutender Sammlungen steht hierbei in keinem Gegensatz zu großflächigen Digitalisierungsmaßnahmen.

Da sich die Mehrzahl der bisher durchgeführten und in Planung befindlichen Projekte mit handschriftlichen Beständen und Alten Drucken beschäftigt haben, werden im Folgenden die Techniken und Parameter in diesen Bereichen besonders gründlich erörtert. Dies sollte jedoch nicht zum Trugschluss verleiten, dass die Digitalisierung moderner Materials nicht möglich sei oder von der wissenschaftlichen Community nicht erwartet werde. Anforderungen, die für den Bereich des älteren Material gelten, können selbstverständlich in vollem Umfange auf die wesentlich umfangreichere spätere Überlieferung übertragen werden. In mittelfristiger Perspektive sollen diese Praxisregeln auch weitergehende Empfehlungen zum Umgang mit Fotos, Filmen, Rundfunk- und 3D-Aufnahmen etc. enthalten. Bereits jetzt finden sich in der Zusammenfassung unter Punkt 6 erste Hinweise zu den genannten Medienarten.

1.2 Auswahl

Grundsätzlich ist zu beachten, dass die technischen Aspekte der Digitalisierung gut geplant werden können, während der intellektuelle Aufwand bei der Auswahl geeigneter Stücke schwer kalkulierbar ist. Es wird daher immer zu entscheiden sein, ob eine großzügigere Definition des Mengengerüstes für ein Digitalisierungsprojekt insgesamt nicht kostengünstiger und effektiver ist als ein komplexer Bewertungs- und Auswahlprozess. Angeregt wird, ausdrücklich auf bereits erbrachte Selektionsleistungen und Vorarbeiten wie Bibliographien 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 rekurren will, so dass die wechselseitige Vernetzung möglich wird. Hierzu gibt es seitens der DFG geeignete Förderangebote.² Alternativ kann man auf einschlägige Fachbibliographien 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. Zu weiteren Gesichtspunkten der Auswahl wird auf die bestehenden Best-Practice-Handbücher verwiesen.³

Größere Digitalisierungsprojekte sollten sich in ein Gesamtprogramm einfügen und abgestimmt sein. Die DFG-Förderung im Bereich der Massendigitalisierung hat zur Voraussetzung, dass bereits hochwertige Metadaten zu den Druckwerken vorliegen (beispielweise bei der Digitalisierung anhand der Verzeichnisse VD16 / VD17 oder bei der Digitalisierung von Materialien aus den Sondersammelgebieten).

Bei Überlegungen zur Auswahl können sich organisatorische und fachliche Prioritäten widersprechen. Während es z.B. organisatorisch sinnvoll sein kann, mit mehreren Partnern nach Zeitsegmenten und/oder regionalen Prinzipien vorzugehen, interessiert sich die Forschung in der Regel nur für den fachlichen Zugang (unabhängig davon, ob ein Buch im 16. oder 17. Jh. oder in den Niederlanden oder Deutschland gedruckt wurde). Ziel einer größeren Kampagne muss es sein, beide Aspekte zu versöhnen. Man kann indes davon ausgehen, dass, je größer ein Digitalisierungsprojekt ist, desto mehr organisatorische Überlegungen an Gewicht gewinnen werden. Umgekehrt lassen sich, sofern die derzeit im Aufbau befindlichen zentralen Nachweisinstrumente und Schnittstellen realisiert sind, durchaus auch Projekte konzipieren, die größere Mengen fachbezogene Literatur bzw. themenspezifische Quellen digitalisieren.

1.3 Dublettenprüfung und Datenabgleich bei der Imagedigitalisierung

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

- 1.3.1 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.
- 1.3.2 Projekte mit größeren Mengengerüsten (z.B. über 1.000 Drucke) sollten in Anträgen und Berichten erläutern, wie sich ihr Vorhaben zu kostenfrei zugänglichen kommerziellen Digitalisierungsprojekten (Google, Microsoft, Yahoo etc.) verhält. Erwartet wird auch eine Aussage zu den durch Google an der Bayerischen Staatsbibliothek digitalisierten Beständen.⁴ Ziel sollte sein, die Zahl der ggf. auftretenden Doppeldigitalisierungen mit pragmatischen Aufwand möglichst gering zu halten.
- 1.3.3 Bei Vorhaben, die eine Digitalisierung von Titeln mit Erscheinungsjahr vor 1601 oder die Digitalisierung nicht gemeinfreier Materialien zum Ziel haben⁵, entfallen die Punkte 1.3.1 und 1.3.2. Bei Materialien der Erscheinungsjahre von 1501 bis 1700 sind das VD16 und das VD 17 als Referenzinstrument für Prüfungen auf Doppeldigitalisierungen heranzuziehen. URN und URL der Digitalisate sind an diese Verzeichnisse zu melden. Bei der Digitalisierung von Inkunabeln ist der deutsche Inkunabelzensus (ISTC) heranzuziehen.

2 Vgl. <http://www.dfg.de/lis>

3 Zu weiteren Gesichtspunkten der Auswahl vgl. u.a. das Minerva Good Practice Handbook: http://www.minervaeurope.org/structure/workinggroups/goodpract/document/bestpracticehandbook1_2.pdf.

4 Vgl. die BSB-Website (http://www.bsb-muenchen.de/Informationen_fuer_Antragsstel.1844.0.html).

5 Vgl. das Merkblatt 12.15.

2. Digitalisierung von Druckwerken und unikalene Dokumenten

Die Digitalisierung umfasst die Herstellung digitaler Images und gegebenenfalls zusätzlich die Volltextfassung sowie die Erzeugung von Struktur- und Metadaten. Wenn im Folgenden von Digitalisierung gesprochen wird, so ist der gesamte Arbeitsgang gemeint (Vorbereitung, Digitalisierung im engeren Sinne, Katalogisierung/Erschließung bzw. Erzeugung von Metadaten sowie Langzeitsicherung/digitale Bestandserhaltung).

2.1 Bereitstellung der Materialien, konservatorische Prüfung

Vorbereitende Tätigkeiten werden bei Digitalisierungsprojekten oft unterschätzt und sollten vor einem Projekt genauestens geprüft werden. Sind die Materialien überhaupt verfügbar? Gibt es möglicherweise konservatorische Bedenken gegen die Digitalisierung der Originale? Soll daher eher vom vorhandenen Mikrofilm digitalisiert werden? Stehen genug Kräfte zur Verfügung, um die Bücher auszuheben? Ist wissenschaftlich bzw. bibliographisch geschultes Personal zur Hand, das Vollständigkeitskontrollen bzw. Kollationierungen vornimmt, sofern dies aus vorhandenen Katalogeinträgen nicht hervorgeht? Eine Digitalisierung von unvollständigen oder defekten Drucken sollte 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 Bücher keinesfalls unterbleiben. Stellt die Reproduktion ein Risiko bzw. eine besondere Belastung für die Originale dar,⁶ so sollte sie sofern möglich vom bestehenden Mikrofilm erfolgen oder ganz unterbleiben. Zu denken wäre auch an Verfahren, bei denen mehrere Institutionen kooperieren und untereinander vereinbaren, Titel, die in einer Institution, z.B. wegen zu engem Falz, nicht aufgenommen werden können, in einer anderen digitalisiert werden. Auf jeden Fall sind wertvolle historische Drucke mit der konservatorisch nötigen Sorgfalt zu behandeln, auch wenn dies den Scandurchsatz reduziert und einen erhöhten Zeitaufwand fordert.

2.2 Imagedigitalisierung

Bei der Altbestandsdigitalisierung ist auf jeden Fall die Imagedigitalisierung vorzunehmen. Auch bei Vorliegen eines maschinenlesbaren Volltextes sollte auf die Imagedigitalisierung bzw. auf die Präsentation des digitalen Faksimiles nicht verzichtet werden, da eine Fülle von Informationen nur über das visuelle Abbild des Druckes oder der Handschrift transportiert werden. Sofern nicht aus konservatorischen oder anderweitigen Gründen eine Digitalisierung vom Film angezeigt ist, empfiehlt sich bei älteren gedruckten Materialien bis ca. 1750 und grundsätzlich bei Handschriften eine Reproduktion in Farbe vom Original. Digitale Kameras und Scanner nehmen heute standardmäßig in Farbe auf, die Kosten für Speichermaterialien sinken bezogen auf das MB exponentiell. Auch die erwarteten Schwierigkeiten beim Farbmanagement haben sich bei weitem nicht als so gravierend herausgestellt, als seinerzeit befürchtet wurde. Probleme bei den Aufnahmen lassen sich durch beigelegte Targets oder standardisierte Kontrollmechanismen bei der Bilderzeugung kompensieren. Die *Speicherung* von 48 Bit Bildern (die von der *Digitalisierung* mit 48 Bit sorgfältig zu unterscheiden ist) ist derzeit nur in Ausnahmefällen sinnvoll. Diese Bildqualität bietet ausschließlich dann Vorteile, wenn äußerst hochwertige Scans erstellt werden sollen, was nur bei wenigen Materialgruppen vorkommen wird.

2.2.1 Digitalisierungsparameter

Ziel der Digitalisierung ist die möglichst originalgetreue Wiedergabe des Druckes oder der Handschrift nach Maßgabe der dafür geltenden wissenschaftlichen Erfordernisse. Die anzuwendenden Parameter für die Digitalisierung sind mit Blick auf die Qualität des Bildes, seine

6 Petersen, Dag-Ernst: Die Mikroform: Chance und Gefahr für das Buch. In: IADA Preprints 1999. 9. IADA Kongress, 16. - 21. August 1999 in Kopenhagen, S. 181 – 183.

Langzeitverfügbarkeit und Interoperabilität zu wählen. Neben den folgenden Regeln können zur vergleichenden Orientierung auch die Empfehlungen der Digital Library Federation (DLF)⁷ oder anderer einschlägiger Institutionen herangezogen werden.

Bei der Herstellung sind grundsätzlich zwei Formen von Digitalisaten zu berücksichtigen. Zum einen der sogenannte *digital master*, also das Roh- oder Archivformat, und für die Nutzung hergestellte *Derivative*, meist verkleinerte Kopien in anderen Dateiformaten. Die nachstehenden Parameter beziehen sich nur auf den Master. Derivate wie z.B. JPEG oder GIF Dateien sind in Abhängigkeit von der gewählten Präsentation aus den Mastern zu erzeugen und können, sofern die Master ausreichend Qualität bieten, nach Belieben modifiziert werden. Das wird z.B. nötig, wenn sich die bei den Benutzern zu erwartenden Bildschirmauflösungen ändern oder Bildformate zum Einsatz kommen sollen, die optimierte Eigenschaften für die gewünschte Anzeige haben (stufenfreies Zoomen; fließende Übergänge zwischen Abschnitten großer, gleichzeitig aber detailreicher Objekte wie zum Beispiel Landkarten oder mittelalterliche Urkunden).

Der Master ist die Grundlage für alle weiteren Prozesse. Daher sollte man seiner Herstellung und Archivierung besondere Aufmerksamkeit schenken. Jedoch gilt für den Master, dass man nicht einfach mechanisch die zu einem Zeitpunkt gerade höchste realisierbare technische Qualität heranzieht. In älteren Dokumenten wird oft explizit oder implizit die Auffassung vertreten, dass der Akt der Digitalisierung so aufwändig sei, dass stets die „höchste technisch machbare“ Qualität angestrebt werden sollte, damit die Digitalisierung „nie“ wieder wiederholt werden müsste. Diese Aussage ist nicht haltbar. Eine nüchterne, wenn auch großzügige Abwägung, welche Formen der Nutzung zu erwarten sind, sollte also ebenso selbstverständlich sein wie die Sorge um langfristig ausreichende Qualität.

2.2.1.1 Auflösung und Bildqualität

Für Graustufen oder Farbaufnahmen wird eine Mindestauflösung von 300 dpi bezogen auf das Format der Originalvorlage als Richtwert empfohlen. Bei Handschriften oder Kartenwerken mit sehr feinen Strichen oder Beschriftungen können auch bis zu 400 dpi erforderlich sein. Wenn bitonale Scans hergestellt werden sollen, sind 600 dpi erforderlich.

Eine darüber hinaus gehende Auflösung macht meist wenig Sinn, weil alle wichtigen Informationseinheiten visuell erkannt werden können. Anders mag der Fall liegen, wenn man sich Sonderfragen wie der Untersuchung von Papierstrukturen widmet, die eine beträchtliche Vergrößerung erfordern. Das ist aber nicht Gegenstand dieser Empfehlungen.

Die Auflösung ist indes nur als einer von mehreren Aspekten zu betrachten. Gute Bildqualität ist auch von der Art der eingesetzten Technik abhängig. Das entstandene Bild sollte daher auf seine Farbechtheit und Originaltreue hin sorgfältig untersucht werden. Empfohlen wird, zu diesem Zweck die Bildschirme zu kalibrieren⁸ und ein definiertes Umgebungslicht zu schaffen, um Abbildungen auf dem Bildschirm objektiv beurteilen zu können.

Beim Einsatz von Scannern ist der Abstand zum Objekt definiert. Die Auflösung des Scanners ist bis zu einer vorgegebenen Maximalgröße des aufgenommenen Objekts immer gleich (z.B. 300 dpi bis zu DIN A3). Bei Digitalkameras ist die Auflösung abhängig von der Entfernung zum aufgenommenen Objekt. Um sicherzugehen, dass 300 dpi erreicht werden, muss einmalig berechnet werden, wie groß das Objekt für eine bestimmte Kamera maximal sein darf. Wenn eine Kamera z.B. eine Matrix von 4000 x 3000 Pixel (12 Mio. Pixel) aufweist, dürfen nur Objekte bis

7 Vgl. hierzu die Empfehlungen der Digital Library Federation: Benchmark for Faithful Digital Reproductions of Monographs and Serials. URL: <http://www.diglib.org/standards/bmarkfin.htm>, PURL: <http://purl.oclc.org/DLF/benchrepro0212>

8 Üblicherweise kommen hier spezielle Kalibrierungsinstrumente zum Einsatz, bei denen ein Messgerät mit einem Saugnapf auf den Bildschirm gepresst wird und anhand einer vordefinierten Farbskala die Farbechtheit des Monitors prüft. Die mitgelieferte Software erstellt dann ein Profil, das dafür sorgt, dass abweichende Farbwerte korrigiert werden.

zu einer Größe von 33,9 x 25,4 cm aufgenommen werden. Dies wird gemäß folgender Formel berechnet:

Anzahl Pixel (Dots) der Kamera
 ----- = max. Größe des Objekts in Inch (1 Inch = 2,54 cm)
 Auflösung in Dots per Inch (dpi)

Bei 300 dpi:

Längsseite: $4000/300 = 13,3$ inch = 33,9 cm.

Kurzseite: $3000/300 = 10$ inch = 25,4 cm

Bei der Digitalisierung vom Film muss bei der gewünschten Zielgröße 300 dpi in Bezug auf das Original der Verkleinerungsfaktor des Mikrofilms oder -fiches berücksichtigt werden. Hat man z.B. eine Vorlage mit 24 x 36 cm auf einem klassischen Kleinbild-Dia der Größe 24 x 36 mm aufgenommen, muss, wenn wiederum vom Film gescannt wird, bei der Auflösung der Faktor 10 in Bezug auf das Original hinzugerechnet werden. Mit anderen Worten: das Kleinbild-Dia müsste mit 3000 dpi gescannt werden. Voraussetzung für eine genaue Bestimmung der Auflösung ist, dass man die Größe des Originalobjektes kennt oder zumindest eine begründete Schätzung (z.B. Folioband nicht größer als 40 cm) zugrunde legt, um eine Unterschreitung der Mindestauflösung von 300 dpi in Bezug auf die Originalgröße zu vermeiden.

2.2.1.2 Farbtiefe

Beim bitonalen Scannen (s/w) wird mit einer Farbtiefe von 1 Stufe (1 Bit) per Pixel gearbeitet. Dabei nimmt ein Pixel entweder die Form 1 (=schwarz) oder 0 (=weiß) an. Graustufenbilder werden mit 256 Stufen pro Pixel digitalisiert. Farbbilder sind nach den drei Farbkanälen rot, grün und blau (RGB-Farben), eine Kombination aus 3 x 256 Stufen. Technisch gesehen ist also eine Farbaufnahme eine dreifache Graustufenaufnahme, indem die Farbwerte pro Kanal anteilig in einen Pixel additiv eingemischt werden (z.B. 35 Rot + 233 Grün + 186 Blau). So ergeben sich 16,7 Mio. Farben ($256 \times 256 \times 256$). Eine 256-stufige Differenzierung benötigt im Rechner 8 Bit bzw. 1 Byte (jedes Bit nimmt den Zustand 0 oder 1 an). Eine Farbaufnahme erfolgt somit in der Regel mit 24 Bit Farbtiefe (3 x 8 Bit / 3 Byte). Manche Kamera- oder Scannerhersteller bieten sogar noch mehr Farbtiefe bis zu 48 Bit an, was an sich unnötig ist, da für die Darstellung am Computer-Bildschirm eine Farbtiefe von 24 Bit völlig ausreicht. Dennoch mag dies sinnvoll sein, wenn man ein gescanntes Bild nachbearbeiten möchte, da durch Tonwertkorrekturen Farbwerte verloren gehen können. Gerade die *Aufnahme* bei 48 Bit bietet oft Vorteile und ein ausgewogeneres Bild, da sich menschliche und kameratechnische Spektralwahrnehmung unterscheiden.

Eine *Archivierung* von 48-Bit-Bildern ist jedoch selten zu rechtfertigen (vgl. 2.2). Zu bedenken ist auch, dass sich der Speicherbedarf für ein 48-Bit-Bild verdoppelt, was sich bei sehr großen Archiven durchaus zum Problem auswachsen kann. Man sollte daher eine Farbtiefe, die über 24 Bit hinausgeht, auf diejenigen Materialien beschränken, bei denen es – aus wissenschaftlichen Gründen – auf eine möglichst exakte Farbwiedergabe oder auf spezifische Farbräume ankommt bzw. umfangreiche Nachbearbeitungsschritte anfallen können.

2.2.2 Dateiformate

Bildmaster von Graustufen- oder Farbbildern sollten nach dem derzeitigen Kenntnisstand in „TIFF uncompressed“ archiviert werden. Für bitonale Bilder kann TIFF mit Group 4 Kompression verwendet werden. TIFF gibt es schon seit den 80er Jahren und hat sich als einer der wichtigsten Standards etabliert. Es ist damit zu rechnen, dass es auch in Zukunft von allen Standardprogrammen unterstützt wird. Allerdings gilt dies nur für so genannte „Baseline-TIFFs“. Hinter dieser Bezeichnung verbirgt sich folgender Sachverhalt: Der TIFF-Standard insgesamt ist ausgesprochen reich und ermöglicht auch die Speicherung von Bildern mit exotischen Eigen-

schaften – beispielsweise aufgeteilt auf so genannte „Tiles“, die es ermöglichen, aus einem Gesamtbild einzelne Teilbilder unabhängig voneinander zu laden, was etwa für den Bereich hoch auflösender kartographischer Aufnahmen sehr nützlich ist. Ein sehr reicher Standard ist aber nur sehr mühsam vollständig zu implementieren: Deshalb wird innerhalb des Standards zwischen einem relativ kleinen Nukleus von Bildeigenschaften unterschieden, die eine jede Implementation des Standards unterstützen *muss*, wenn sie von sich behauptet, TIFF zu unterstützen, und zahlreichen darüber hinausgehenden Erweiterungen, die Implementationen verwenden *können*, die von anderer TIFF unterstützender Software aber ggf. ignoriert wird, ohne dass dies den Anspruch, dass dieses andere Softwareprodukt „TIFF unterstützt“, aufheben würde.

Bei Speicherplatzproblemen könnte theoretisch auch mit einem verlustfreien Kompressionsverfahren (z.B. mit LZW komprimiertes TIFF) archiviert werden. Grundsätzlich stellt die Verwendung eines Kompressionsformates ein Risiko dar, weil schon durch geringe Beschädigung der Datei oder Kippen einzelner Bits das gesamte Bild beeinträchtigt wird. Solche Beschädigungen können durch Defekte der Speicherträger oder auch beim Kopieren der Dateien auftreten. Dies tritt besonders häufig beim JPEG-Standard auf. Bekannt sind Phänomene, wo z.B. nur noch die Hälfte des Bildes sichtbar ist. Solcherart beschädigte Bilder sind meist irreparabel. Das Tückische ist, dass das Bild auf dem einen Medium in Ordnung sein kann, nach einem etwa durch eine Migration auf ein neues Speichermedium erfolgten Kopiervorgang aber beschädigt. JPEG ist insofern als Archivformat ungeeignet. Wenn aus Kostengründen oder anderweitigen Überlegungen heraus auf Kompression nicht verzichtet werden kann, sollte im Zweifelsfall auf die Dateiformate PNG oder TIFF-LZW zurückgegriffen werden.

Eine neuere Entwicklung ist JPEG2000⁹, das nicht nur einen gegenüber JPEG effizienteren Kompressionsalgorithmus bietet, sondern auch weniger störungsanfällig ist. Verlustfreies Speichern ist im Unterschied zum JPEG-Format möglich. Weitere Vorteile sind progressiver Bildaufbau (je mehr vom Bild geladen ist, desto mehr Details zeigen sich) oder die Möglichkeit, Metadaten aufzunehmen. Als Schlüsselvorteil könnte sich die Eigenschaft von JPEG2000 erweisen, aus einem großen Bild verschiedene Auflösungen und auch Ausschnitte zu generieren, was die Verwaltung eines Bildarchivs vereinfachen würde, indem sich die Vorhaltung diverser Auflösungen erübrigt. Ob und inwiefern JPEG2000 geeignet ist, TIFF als Master zu ersetzen, kann zum derzeitigen Zeitpunkt noch nicht entschieden werden. Weder der Verbreitungsgrad noch die Softwaresituation, die wichtige Indikatoren für den Einsatz eines Master-Formates sind, lassen im Moment eine eindeutige Empfehlung zum Einsatz des Formates als Archivformat zu.

Für die Publikation im Internet empfiehlt sich wegen seiner großen Verbreitung JPEG und PNG. GIF kommt wegen seiner begrenzten Farbpalette nur bei bitonalen oder Graustufenscans in Frage. Letztlich ist dies aber freizustellen, denn diese oder andere durch die gängigen Browser lesbaren Formate können leicht nachträglich aus dem Master generiert werden, so dass eine Entscheidung ggf. revidiert bzw. auf veränderte Rahmenbedingungen reagiert werden kann.

2.3 Volltextgenerierung

Volltexte können durchsucht, quantitativ ausgewertet und weiterverarbeitet werden. Sie können in größere Textsammlungen integriert, für spezifische Fragestellungen weiter erschlossen oder für neue Lesegeräte aufbereitet werden. Volltexte umfassen die Zeichen der Vorlage, in den Text eingetragene Auszeichnungen bzw. Markup-Daten zur Markierung von Strukturmerkmalen sowie Metadaten, die üblicherweise Teil der gleichen Datei sind.

Die Herstellung von Volltext kann auf zwei Wegen erfolgen: entweder durch OCR oder durch Abschreiben. 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. Wichtig ist es, am Anfang des Projekts festzulegen, zu welchem Zweck die Voll-

9 <http://www.jpeg.org/jpeg2000/>

textgenerierung durchgeführt werden soll. Dieses Ziel sollte auch während des gesamten Digitalisierungsprozesses nicht aus den Augen verloren werden und immer wieder in Betracht gezogen werden, insbesondere bei den Entscheidungen über die Frage, ob und wie bestimmte Textmerkmale erfaßt werden sollen.

Vor der eigentlichen Erfassung, sei es durch Abschreiben oder durch OCR, ist es häufig nützlich, neben dem unveränderten Masterscan noch eine Kopie des Bilddigitalisats anzufertigen, die entweder durch die Scansoftware oder durch eigenständige Programme vorverarbeitet wird. Insbesondere Entzerrungen, Schärfung, Kontrastveränderungen und Zeilenbegradigungen können so vorgenommen werden und so zu einer besseren Texterkennung beitragen.

2.3.1 Zeichenkodierung

Alle verbreiteten Betriebssysteme unterstützen Unicode, und Unicode ist auch das Zeichenkodierungsformat von XML, das die Grundlage für die wichtigsten Strukturdatenauszeichnungssysteme darstellt. Daher wird empfohlen die Texte in Unicode abzuspeichern. Zu favorisieren ist UTF-8 oder UTF-16.

In der Regel enthalten Vorlagen aber auch Zeichen, die nicht Teil des Unicode-Standards sind. Es gibt mehrere Möglichkeiten solche Zeichen zu kodieren (Private plane bei Unicode, Entities, Grafiken).

Bei der Kodierung ist auch zu bedenken, ob je nach fachspezifischem Bedürfnis zum Beispiel Unterschiede zwischen dem langen und dem kleinen s in Frakturschriften oder Ligaturen in Frakturschriften (ch, tz etc.) erhalten bleiben oder aufgelöst werden sollen.

2.3.2 Auszeichnung von Strukturdaten bei Druckwerken (vgl. auch 2.6.2)

Bei der Auszeichnung von Strukturdaten muss zunächst entscheiden werden, wie und in welchem Umfang man textsortenspezifische Gliederungen wie z.B. Kapitel, Unterkapitel, Jahressbände, Aufsätze etc. berücksichtigt. Hinzukommen weitere denkbare Strukturmerkmale, zum Beispiel Inhaltsverzeichnisse, Register, Zeilenumbruch¹⁰, 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.

All diese Beispiele sind *nicht* als verpflichtende Vorgaben gedacht, sondern als Aspekte, die bei einer Volltextdigitalisierung durchdacht werden sollten. Die getroffenen Entscheidungen sollten auf jeden Fall dokumentiert werden.

2.3.3 Layout

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 XML-Auszeichnungssprache (z.B. XSLT, XSL:FO), die die Unabhängigkeit von spezieller Software weitgehend sicherstellt. Falls eine Archivierung des Formats mit XML-Techniken aus nachvollziehbaren Gründen nicht möglich ist, können Textdokumente auch nach der ISO-Norm 19005-1, die ein langfristiges sicheres Subset von PDF spezifiziert, archiviert werden. PDF-Dateien sind aber in jedem Fall kein Ersatz für eine Bereitstellung der Strukturdaten im XML-Format, da dies in PDF nicht möglich ist.

¹⁰ Auf folgendes Problem ist aufmerksam zu machen: Erhält man die Trennzeichen, dann kann nach dem Wort in den meisten Suchmaschinen nicht gesucht werden. Löscht man das Trennzeichen, dann geht im Vergleich zur Vorlage Originaltext verloren.

2.3.4 Texterfassung

OCR

Bisher galt, dass die OCR-Bearbeitung erst bei jüngeren Antiquaschriften ab dem 19. Jahrhundert sowie bei Frakturschriften ab der zweiten Hälfte des 19. Jahrhunderts zu akzeptablen Ergebnissen führt. Da sich der Markt anbieterseitig dynamisch weiterentwickelt und neue Produkte auf den Markt kommen, können die Praxisregeln zur Frage der OCR-Software und ihrer Anwendbarkeit zum jetzigen Zeitpunkt keine weiteren Empfehlungen geben.

Insbesondere bei großen Textbeständen, die in einem gleichmäßigen Druck vorliegen, kann möglicherweise durch entsprechendes Training der OCR-Software ein brauchbares Resultat erreicht werden. Bislang beherrscht noch keine OCR-Software den eigenständigen Wechsel von Fraktur- zu Antiquatypen. Sind die verschiedenen Typen in jeweils eigene Blöcke segmentiert, kann per Hand umgeschaltet werden. Liegt eine Mischung verschiedener Schriften im fortlaufenden Text vor, wird nur die voreingestellte Schriftart zuverlässig erkannt. Sind in der Vorlage neben Text auch Formeln u.a.m. enthalten, so können diese mittels des Einsatzes von OCR-Software derzeit noch nicht in ausreichendem Umfang erkannt werden.

Anders liegt der Fall, wenn man mit einer gleichsam „schmutzigen“ Version (bekannt auch als schmutziges OCR) arbeiten möchte. In diesen Fällen dient der Volltext nur zur Positivsuche, nicht aber als verlässliche Textbasis. Positivsuche bedeutet, dass nur positive Treffer erzielt werden. Ein negatives Ergebnis bedeutet umgekehrt nicht, dass nicht doch ein Treffer vorhanden ist. Ebenso wenig kann man sicher sein, dass mit den gefundenen Treffern alle theoretisch möglichen Treffer gefunden wurden. Ein solches Verfahren kann als Überbrückung bis zur Herstellung besserer Texte gute Dienste leisten. Es sei aber ausdrücklich vor der Gefahr gewarnt, dass Nutzer bei mangelnder Dokumentation zu falschen Ergebnissen geführt werden. Auf die Suchbedingungen sollte an deutlich sichtbarer Stelle hingewiesen werden. Vorteilhaft ist auch, den Volltext – selbst bei mäßiger Qualität – anzuzeigen, damit sich der Nutzer selbst ein Bild von der Qualität seiner Recherchebasis machen kann.

Manuelle Erfassung / double keying

Man unterscheidet bei der Direkterfassung (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 99,997% 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 im Grunde wertlos. Anders ausgedrückt, bei 99% wäre jeder 100ste Buchstabe(!) falsch, pro Zeile also ca. ein Fehler.

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 Textes zu überprüfen. Wird die vorgegebene Erfassungsgenauigkeit unterschritten, so sollten die vertraglich festgelegte Reaktionen in Kraft treten, z.B. Minderung oder Aussetzung der Zahlung

Manuelles Erfassen ist zwar ebenfalls fehleranfällig, allerdings können mit der doppelten Eingabe der Texte und einem anschließenden Vergleichslauf zur Fehlersuche die besten Textqualitäten erreicht werden. Dies Verfahren ist zur Zeit jedoch auch das kostenintensivste. Die eigentliche Texterfassung wird hierbei zumeist im Ausland vorgenommen; der Kontakt mit einer Digitalisierungsfirma sollte jedoch über einen Ansprechpartner in Deutschland geschehen, da in der Regel eine enge Kooperation und Absprache zu den Erfassungsdetails von Nöten 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 für diese Zwecke zu verwenden.

2.4 Langzeitsicherung

Für die Langzeitsicherung oder Archivierung digitaler Inhalte gibt es zum jetzigen Zeitpunkt keine verallgemeinerbare Lösung, die für alle Objekt- und Materialgattungen geeignet ist. Wesentliche Erfolgsfaktoren für eine Langzeitarchivierung 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.

Die langfristige Sicherheit der Ergebnisse von Digitalisierungsprojekten ergibt sich einerseits aus der Wahl der Daten- und Metadatenformate. Überlegungen dazu flossen in die vorigen Abschnitte ein. Andererseits ist sicherzustellen, dass die digitalen Daten auch physikalisch verfügbar bleiben. Dabei gilt: Die DFG kann Kosten für die Langzeitsicherung nur übernehmen, soweit sie sich auf die Erstellung entsprechender langzeitsicherer Kopien der Daten beziehen. Es wird erwartet, dass ein Antrag erläutert, auf Grund welcher institutioneller Vorkehrungen die während des Projekts angefallenen Daten langfristig sichergestellt werden. Dies schließt sowohl die Feststellung ein, dass die Mittel für die langfristige Weiterführung der für das digitale Angebot notwendigen Softwareplattform im Rahmen eines Gesamtkonzepts bereitstehen als auch die Feststellung, dass es ein explizites Konzept zur langfristigen Sicherung der Daten gibt.

Der Archivierungsfrage wird oft zu wenig Aufmerksamkeit geschenkt sowie Kosten und Aufwand unterschätzt. Je mehr Auflösung eine Kamera oder ein Scanner bietet, umso mehr Speicherbedarf verursachen die erzeugten Bilder. Der *digital master* von Farbbildern kann 20 bis 80 oder sogar mehr MB umfassen. Hier wachsen unter Umständen schnell Bildarchive von mehreren Terabyte heran, die abgelegt werden müssen.

Derzeit spielen bei der digitalen Archivierung vor allem vier Trägermaterialien eine Rolle:

- (1) optische Wechseldatenträger wie CD/DVD,
- (2) Bandlaufwerke (streamer),
- (3) Festplatte,
- (4) Mikrofilm.

(Zu 1) Die Archivierung auf optischen Wechseldatenträgern wie CD-Rs oder DVD-Rs ist bei größeren Mengengerüsten nicht sinnvoll und nicht förderungsfähig. Darüber hinaus sind CDs/DVDs aufgrund des raschen Technologiewandels und der Materialbeschaffenheit kein geeignetes Speichermedium für die Langzeitsicherung. Zur Archivierung auf CD oder DVD werden vor allem diejenigen greifen, die mit dem Aufbau eines digitalen Archivs beginnen oder nur eine begrenzte Anzahl von Digitalisaten herstellen wollen.¹¹ Bei größeren Datenmengen kommt man jedoch schnell an Kapazitätsgrenzen. Handelsübliche CDs fassen 700 MB, DVDs 4,7 GB, mit neuen Techniken kommt man auf rund 8 GB. Um nur ein Terabyte abzulegen (das entspricht etwa 180 Büchern¹²), bedürfte es 1429 CDs, 213 DVDs zu 4,7 GB oder 125 DVD zu 8 GB. Eine

11 Iraci, Joe: Die relative Haltbarkeit verschiedener optischer Speicherplatten. CD, DVD. In: Restaurator 26,2 (2005) S. 134-150.

12 Realistischer Schätzwert für ein TB: im Mittel 25 MB pro Image, 220 Seiten pro Buch, also 5,5 GB pro Buch, was 182 Büchern für ein TB entspricht.

Archivierung auf CD wird daher angesichts dieser Zahlen schnell unhandlich. Aber auch DVDs stellen bei großen Datenmengen keine wirklich befriedigende Lösung dar. Die Scheiben müssen gebrannt, beschriftet und archiviert werden, Vorgänge, die u.U. sehr viel Zeit beanspruchen. Direktzugriffe sind, wenn man keine Roboter oder CD-Server einsetzt, umständlich und ortsabhängig. Bei der Verwendung von CDs/DVDs ist in jedem Fall eine redundante (mindestens doppelte) Herstellung sinnvoll.

(Zu 2) Die Archivierung auf Bändern bietet eine etwas bequemere Variante der Massenarchivierung, hat aber den Nachteil, dass Bänder relativ langsam sind. Wenn häufige Zugriffe auf die *digital master* benötigt werden, sind Bänder keine gute Lösung. Sie müssen darüber hinaus regelmäßig bewegt werden, um das Zusammenkleben zu verhindern. Wenn ein Bandarchivsystem genutzt werden soll, ist Universitätsbibliotheken – oder Bibliotheken, die organisatorisch an Universitäten angebunden sind, – dringend empfohlen, bei der Entscheidung über die Strategie für die Langzeitarchivierung den Kontakt zum Rechenzentrum der Universität zu suchen. Moderne Bandarchivierungssysteme, „Robots“, die stets auf mehrere 100 Terabyte angelegt sind, garantieren, dass von Bändern mehrere Kopien angelegt werden und dass die Bandkassetten in den nötigen Mindestabständen bewegt werden.

(Zu 3) Besonders zu empfehlen im Sinne einer Migrationsstrategie ist die redundante Datenerhaltung auf Festplatten (z.B. RAID 5) in Form sogenannter Network Attached Storage (NAS) Systeme oder Storage Attached Network (SAN) im Rechenzentrumsbetrieb. Zur Sicherheit muss eine Bandsicherung oder alternativ eine weitere Kopie auf Festplatte erfolgen (Bandsicherung bzw. Kopie dürfen nicht am selben Ort aufbewahrt werden). Dieses Modell erlaubt den schnellen und unkomplizierten Zugriff auf die Daten und erleichtert ggf. nötige Migrationen. Das Archivierungsmodell „Festplatte“ geht davon aus, dass Daten permanent aktuell sind, dass es also kein Szenario des Weg- oder Beiseitelegens gibt, bei dem nicht Hardware oder Softwareprozesse die Integrität sicherstellen. Insofern geht dieses Modell einen anderen Weg der Langzeitsicherung, mit dem man als solchem noch keine Erfahrungen hat, wie überhaupt die Langzeitarchivierung digitaler Medien ein bisher noch nicht abschließend diskutierter Bereich ist. Es sei hier nur auf die einschlägigen nationalen und internationalen Aktivitäten verwiesen, z.B. Nestor¹³ oder Portico.¹⁴

(Zu 4) Eine weitere Möglichkeit der Langzeitsicherung besteht in der Ausbelichtung der Digitalisate auf Mikrofilm im Rahmen einer Konversionsstrategie. Mikroformen bieten im Vergleich zu sonstigem analogen Fotomaterial bei optimaler Lagerung hinsichtlich ihrer Haltbarkeit die besten Perspektiven. Zu einem späteren Zeitpunkt können die Filme bei Bedarf mit entsprechenden Filmscannern wieder in ein digitales Format überführt werden.¹⁵ In der Zwischenzeit gibt es auch Lösungen für die Ausbelichtung von hochwertigen Farbdigitalisaten auf Farbmikrofilm.¹⁶ Zu diesen neu entwickelten Techniken kann zum gegenwärtigen Zeitpunkt keine abschließende Bewertung vorgenommen 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 können sich kleinere Einrichtungen sehr wohl der Kompetenz und der Dienstleistung großer Einrichtungen bedienen.

Den im Rahmen der wissenschaftlichen Informationsinfrastrukturen geförderten Bibliotheksprojekten wird – falls nicht ohnehin gesetzlich geregelt – die Ablieferung eines komplettes Datensets an die Deutsche Nationalbibliothek empfohlen.¹⁷

13 <http://www.langzeitarchivierung.de/>

14 <http://www.portico.org/>

15 Solche Filmscanner gibt es bereits seit längerer Zeit für S/W-Mikrofilme.

16 Vgl. zum Prinzip der Ausbelichtung auf langzeitstabilen Farbmikrofilm z.B. das Projekt ARCHE (<http://www.landesarchiv-bw.de/arche>). Für 35 mm-Farbmikrofilme gibt es allerdings zum gegenwärtigen Zeitpunkt noch keine befriedigende Lösung für eine rationelle, kostengünstige Filmdigitalisierung.

17 <http://www.d-nb.de/netzpub>

2.5 Digitalisierung inhouse oder mit Dienstleistern?

Die Digitalisierung kann inhouse oder durch Beauftragung eines Dienstleisters erfolgen. Im ersten Fall können Geräte, aber auch projektbedingte Personalkosten, bei der DFG beantragt werden. Gerade bei größeren Mengen ist es manchmal aus Sach- und Finanzgründen günstiger, die Arbeiten im eigenen Haus mit eigenen Geräten und eigenem Personal durchzuführen. Ein anderes Mal ist die Beauftragung eines Dienstleisters wieder die bessere und kostengünstigere Wahl. Schließlich kann es in Einzelfällen sinnvoll sein, Personal- oder Sachmittel zu beantragen, die dann aber an einem anderen Standort, beispielsweise einer großen Bibliothek mit einschlägiger Digitalisierungskompetenz, zum Einsatz kommen.

Die Entscheidung, ob ein Vorhaben als Inhouse-Projekt konzipiert oder ob auf das Angebot eines Dienstleisters zurückgegriffen wird, ist daher immer eine projektspezifische Entscheidung, die allein in die Verantwortung der 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 Altbestand Erfahrungen vorweisen können. Da in diesem Bereich dieselben Kriterien und Handlungsrichtlinien gelten wie beim Film, kann hier auf eine weitere Erörterung verzichtet werden.

Die Prüfung der von einem Dienstleister gelieferten Qualität ist meist nicht einfach. Bei der Filmdigitalisierung ist oft nicht klar festzustellen, ob eine mangelhafte Digitalisierung auf einen schlechten Film oder ungenügende Scantechnik des Dienstleisters zurückzuführen ist. Um eine angemessene Bewertung der gelieferten Qualität zu erleichtern, sollte man Wert darauf legen, dass bei einer Digitalisierung vom Original möglichst Farbtafeln und ein Zentimetermaß beigelegt werden, die es erlauben, Farbtreue und Auflösung zuverlässig zu beurteilen.

Nachdrücklich sei dabei darauf hingewiesen, dass Dienstleister für die Durchführung einzelner Arbeitsschritte zwar unbestreitbar sehr kostengünstige und effektive Ergebnisse erbringen können, dass dies die Auftrag gebende Bibliothek aber nicht davon entlasten kann, allgemeines Knowhow über Digitalisierungsprojekte bereitzuhalten (für die langfristige Aufrechterhaltung des digitalen Angebotes einerseits, für die Abwicklung einer effektiven Zusammenarbeit mit dem beauftragten Dienstleister andererseits).

Bisherige Erfahrungen haben gezeigt, dass die Vergabe von Teilaufgaben eines Digitalisierungsprojektes an externe Dienstleister sehr effektiv sein *kann*, aber keineswegs immer ist. Grundsätzlich ist festzuhalten, dass auch der konsequente Einsatz von Dienstleistern eine Bibliothek oder ein Archiv des 21. Jahrhunderts nicht der Aufgabe enthebt, Kompetenzen im Bereich der Digitalisierung im Hause bereitzuhalten – allein schon deshalb, weil nur dann kompetente Vertragsverhandlungen mit Dienstleistern möglich sind. Schließlich sei nochmals hervorgehoben, dass die DFG davon ausgeht, dass die Digitalisate langfristig innerhalb der Infrastruktur der Häuser gepflegt werden. In der Antrag stellenden Institution selbst müssen daher ausreichende Personal- und Sachmittel für die Projektleitung und die Projektkontrolle sowie die Auswahl und Vorbereitung der Materialien zur Digitalisierung bereitstehen. Des Weiteren werden Arbeiten zur spezifischen Erschließung und Aufbereitung der digitalisierten Materialien nahezu immer in den antragstellenden Bibliotheken, Archiven und wissenschaftlichen Einrichtungen selbst durchgeführt.

Liegen diese Bedingungen vor, wird die Beauftragung externer Dienstleister häufig vorteilhaft sein. Dies 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. Auch nachgelagerte Produktionsstufen, wie die Bereitstellung der Daten im Netz, Vertrieb und Vermarktung von Offline-Produkten, spezialisierte Druckzentrums-Dienstleistungen oder die Migration zur Langzeitarchivierung, können durch Fremdfirmen manchmal wirtschaftlicher erfolgen als durch die Bibliotheken und wissenschaftlichen Einrichtungen selbst.

Bei der Formulierung der einschlägigen Verträge ist zu beachten:

- (1) Die geforderten Leistungsparameter müssen exakt festgelegt werden, insbesondere die Anforderungen und Formatstandards für die zu liefernden Rohdaten. 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.
- (2) 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 abgefordert werden, dass kurzfristig und kostenlos Ersatzleistung oder Nachbesserung geleistet wird, wenn dies wegen Nichteinhaltung der „Praxisregeln Digitalisierung“ oder wegen sonstiger berechtigter Qualitätsmängel nötig werden sollte.
- (3) Die Deutsche Forschungsgemeinschaft geht bei einer Bewilligung davon aus, dass die Digitalisierung und die spätere Benutzung unter Beachtung der Vorschriften des Urheberrechts erfolgen und dass nötigenfalls das Einverständnis der Berechtigten eingeholt wird. Es muss sichergestellt sein, dass der Eigentümer der Originaldigitalisate keine Rechte an die beauftragten Firmen abtritt.

2.6 Metadaten

Sofern Daten außerhalb vorhandener Bibliotheksverbünde oder zentraler Portale separat erfasst werden, sind Metadaten in einer von Software unabhängigen Form bereitzustellen. Dies ist in den Workflow des Projekts so einzubetten, dass ein vollständiger Metadatensatz in software-unabhängiger Form – was in aller Regel auf eine XML-Kodierung hinauslaufen wird – auch dann bereit steht, wenn das Projekt aus welchen Gründen auch immer zu einem vorzeitigen Ende kommt. Eine Projektplanung, in der die Bereitstellung herstellerunabhängiger Metadaten erst in einer späten Projektphase oder nach Projektende eingeplant wird, ist in hohem Maße problematisch. Beim Einsatz proprietärer Softwaresysteme ist also von Anfang an darauf zu achten, dass diese die Ausgabe der Daten in einem herstellerunabhängigen Format ermöglichen.

Wenn sich die im Rahmen eines geförderten Digitalisierungsprojektes ergebenden Materialien sachlich für die Einbindung in ein von der DFG gefördertes Portal und/oder eine Virtuelle Fachbibliothek eignen, wird erwartet, dass ein Projektantrag entweder erläutert, welche Vorkehrungen projektseitig getroffen werden, um die Anbindung an dieses Portal während der Projektlaufzeit sicherzustellen, oder plausibel macht, warum eine Anbindung aus inhaltlichen Gründen oder aus Gründen des damit verbundenen Aufwandes nicht notwendig bzw. sinnvoll ist.

Allgemein werden deskriptive (meist bibliographische Beschreibung oder archivarische Erschließung), strukturelle (Text-, Dokumentstruktur), administrative (z.B. Rechteverwaltung) und technische (z.B. Dateitypen) Metadaten unterschieden. Die folgenden Überlegungen beziehen sich nur auf deskriptive und strukturelle Metadaten.

2.6.1 Sammlungs- und Bestandsbeschreibung, Katalogisierung, Deskriptive Metadaten

2.6.1.1 Sammlungs- und Bestandsbeschreibung, Projektinformationen

Schon in den klassischen Formen des bibliothekarischen und archivarischen Angebots spielte die Sammlungs- oder Bestandsbeschreibung eine wichtige Rolle, um dem Nutzer einen Überblick über die Art und Zusammensetzung der in einer Altbestandsbibliothek oder einem Archiv vorhandenen Bestände zu ermöglichen. Es wird erwartet, dass Digitalisierungsprojekte Gegenstand und Umfang des jeweiligen Projektes auf einer Seite im Netz möglichst auch in Englisch darstellen. Auf die Förderung durch die DFG ist hinzuweisen. Erwünscht ist darüber hinaus eine

normierte Beschreibung in XML, um diese Informationen in Zukunft besser in nationalen oder internationalen Portalen zusammenführen und gezielt recherchieren zu können. Hinzuweisen ist hier auf das Dublin Core Collection Description Application Profile (DC CD), das Elemente des Collection Description Format (CLD) aufnimmt und mit eigenen Elementen anreichert.¹⁸ Für den archivischen Bereich ist es sinnvoll, sich an den internationalen Erschließungsstandards ISAD (G)¹⁹ und EAD (Encoded Archival Description)²⁰ zu orientieren.

2.6.1.2 Katalogisierung und Erschließung

Digitale Reproduktionen von alten Drucken oder Archivmaterialien müssen mindestens deskriptive – bibliographische – Metadaten beigegeben werden. Digitalisierung ohne Nachweis zumindest der bibliographischen Metadaten bzw. der archivischen Titelaufnahmen nach gängigen bibliothekarischen und archivischen Standards ist nicht sinnvoll! Der Nachweis kann bei Bibliotheksgut entweder durch Katalogisierung der elektronischen Ausgabe oder durch Angabe der URL der Bilddateien bzw. Angabe einer persistenten Verknüpfung im Katalog (OPAC, Verbundsystem) erreicht werden. Von Antragstellern aus universitären Einrichtungen wird erwartet, dass sie die Frage der Katalogisierung mit ihren örtlichen Bibliotheken abstimmen bzw. durch diese durchführen lassen. Die Digitalisate müssen in einschlägigen überregionalen Portalen nachgewiesen werden. Mit Nachdruck empfohlen wird die Bereitstellung einer OAI-Schnittstelle, die neben Dublin Core METS/MODS ausliefert (s. Anhang A), so dass sichergestellt ist, dass einschlägige Portale die Daten harvesten können.

Digitale Reproduktionen von mittelalterlichen Handschriften sollten in der Handschriftendatenbank (*manuscripta mediaevalia*) nachgewiesen werden. Bei separater Erfassung oder Archivierung der Metadaten von mittelalterlichen Handschriften sollte das Handschriftenformat der TEI P5 verwendet werden.²¹

2.6.2 Strukturelle Metadaten zur Imagedigitalisierung (vgl. auch 2.3.2)

Wohl zu erwägen ist die Frage der Anwendung von strukturellen Metadaten, 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 Bibliographie 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 starken digitalen Wörterbuch online zu blättern, um den richtigen Alphabeteinstieg zu finden. Die Entscheidung über die Erstellung struktureller Metadaten ist daher immer eine objektspezifische.

Falls strukturelle Metadaten Verwendung finden, wird die Konsultation der über die Website des sog. DFG-Viewers²² 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 sachbezogene Vokabularien auch über die Projektwebsite publizieren, um eine Nachnutzung zu ermöglichen.

18 <http://www.ukoln.ac.uk/metadata/rslp/schema/>

19 Beispiele für Online-Beständeübersichten finden sich z.B. im Internet-Angebot des Landesarchivs Baden-Württemberg (<http://www.landesarchiv-bw.de>).

20 Zur Anwendung von EAD für die Erschließung im deutschen Archivwesen siehe das abgeschlossene Projekt „daofind“ des Bundesarchivs (www.daofind.de).

21 Eine Bearbeitung nach Maßgabe der DFG Richtlinien zur Handschriftenkatalogisierung steht an der HAB Wolfenbüttel zur Verfügung (vgl. <http://www.manuscripta-mediaevalia.de/hs/kataloge/HSKRICH.htm>). Bei neueren Handschriften oder Nachlassmaterialien ist EAD (<http://www.loc.gov/ead/>) zu empfehlen.

22 <http://www.dfg-viewer.de>

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 Werkes orientiert. Geht man davon aus, dass dem digitalen Faksimile eines alten Druckes oder einer Handschrift die Transkription oder Edition beigegeben werden soll, ist eine Orientierung am TEI-Kodierungsstandard²³ zu empfehlen. Bei einer Seitenbeschreibung mit einigen qualifizierenden Merkmalen (z.B. Illustrationen oder Annotationen) ließe sich der von der Library of Congress gepflegte *Metadata Encoding and Transmission Standard (METS)*²⁴ nutzen. 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 eher anzutreffen ist, die Verwendung von TEI nicht ausschließt,²⁵ so dass sich beide Aspekte ggf. 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 Projektes 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.

2.7 Austausch und Weitergabe von Metadaten, Bekanntmachung

Für den Aufbau einer verteilten digitalen Bibliothek ist die Schaffung eines übergreifenden Standards zum Austausch von Daten und Metadaten von zentraler Bedeutung (vgl. Kap. 5). Allerdings können sich Standards, soweit sie sich unterhalb der Ebene der klassischen bibliothekarischen Formalerschließung bewegen, 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 neben den bibliothekarischen Metadaten auch unterschiedliche Metadatenformate bzw. *community*-abhängige Spezifikationen verwalten können. Dieses lässt sich mit dem Protokoll der Open Archive Initiative (OAI)²⁶ 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 alter Drucke und Handschriften ungenügend,²⁷ 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 möglich (MARCXML, MABxml, EAD, METS/MODS, TEI P5 etc.).

Die DFG empfiehlt ausdrücklich die Bereitstellung von Metadaten über OAI. Dabei sollten neben den vorgeschriebenen Dublin-Core-Metadaten zusätzlich Metadaten gemäß des METS-Standards (bei Drucken METS/MODS) übermittelt werden (vgl. auch Kap. 5). TEI, EAD oder Dokumente anderer domänenspezifischer Standards sollten nach Möglichkeit ebenfalls innerhalb von METS, das in diesem Fall als sog. *wrapper* fungiert, übermittelt werden. Über die beschriebene OAI Funktionalität kann der sog. DFG-Viewer bedient werden. Neu eingerichtete OAI Schnittstellen sollten in Frage kommenden Portalen und Fachbibliotheken gemeldet wer-

23 <http://www.tei-c.org>

24 <http://www.loc.gov/standards/mets>

25 http://www.tei-c.org/Sample_Manuals/bestpractice.htm

26 <http://www.openarchives.org/>

27 Vgl. Diane L. Hillmann, in: Kenny, Anne R./Rieger Oya (Hrsg.): *Moving Theory into Practice. Digital Imaging for Libraries and Archives*, Mountain View: Research Library Group, 2000, S. 89f.

den. Ferner sollten geeignete Maßnahmen ergriffen werden, die dazu führen, dass die Metadaten von Suchmaschinen gefunden werden (z.B. mit dem Sitemap Protokoll²⁸).

Die DFG erwartet, dass Projektnehmer gezielte Anstrengungen unternehmen, um einen hohen Bekanntheitsgrad und eine hohe Nutzung der erstellten Ressourcen zu garantieren. Dies sollte mindestens durch die Einbindung der Ressourcen in vorhandene oder entstehenden materialspezifischen Portale und Kataloge bzw. in die fachspezifischen Virtuellen Bibliotheken geschehen.

28 <http://www.sitemaps.org/>

3. Zitieren digitaler Drucke und Handschriften, persistente Adressierung

In den Anfangszeiten der Digitalisierung hat man 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 selbstständiges Objekt in einem sich dynamisch entwickelnden integralen Forschungsraum. Während also die klassische Fotokopie als Kopie zu vernachlässigen ist, bedarf die digitale Kopie als Kopie eigener Zitationsregeln, die im Übrigen nur für das Internet gelten. Versionen auf CD oder anderen nicht vernetzen Speichermedien kann man wie Filme oder Papierkopien behandeln und zitieren. Wird aber eine Kopie ins Internet gestellt, muss sie, um von anderen Dokumenten 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.

Ein positiver Effekt der netzbasierten Zitierform, die sich in der Regel an der physischen Bildfolge orientieren wird, ist, dass das Referenzieren eindeutig wird, was man wegen der vielen Fehler in alten Drucken oder auch Mangel an Paginierungen oder Foliiierungen etwa in Inkunabeln meist nicht behaupten 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 eigentlich Corpus gehören (Deckel, Spiegel, zusätzliche digitalisierte Wasserzeichen, Teilaufnahmen von Illustrationen etc.). Einzige Voraussetzung dafür ist, dass sich ein bestimmtes Bild in einer alphanumerisch aufsteigenden Sequenz eindeutig verorten lässt (z.B. 00001, 00002, 00002a, 00003 etc; hier wurde ein Bild 0002a eingefügt). Zitierziel ist hierbei ein Image. Bei Volltexten gelten andere Mechanismen, zu denen derzeit noch keine konkreten Empfehlungen ausgesprochen werden können (eindeutige Referenzierungsmöglichkeiten bestünden z.B. mit XPATH oder vergleichbaren Techniken).

Im Grundsatz sollte immer die größtmögliche Granularität angestrebt werden. Vor allem zwei Funktionalitäten sind bei der Netzpräsentation zu nennen. Zum einen die Erreichbarkeit des Werkes als Ganzem, zum anderen die Erreichbarkeit einzelner Seiten oder Doppelseiten aus diesem Werk. Die Struktur eines „Zitats“ könnte wie folgt aussehen (fingiert):

<http://digitalebibliothek.ubique.de?titelid=234&image=0002>

Sichergestellt werden muss mindestens die Erreichbarkeit und Zitierbarkeit eines Werkes als Ganzem, künftig auch die Erreichbarkeit und Zitierbarkeit von einzelnen physischen Seiten von 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.

Die Erzeugung von URNs über die Deutsche Nationalbibliothek auf Werkebene wird empfohlen.²⁹

²⁹ Vgl. <http://www.persistent-identifizier.de/>

4. Bereitstellung der Digitalisate für die Öffentlichkeit

4.1 Open Access

Die Digitalisierung von wissenschaftlichen 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.

Die Deutsche Forschungsgemeinschaft ist Mitunterzeichnerin der „Berlin Declaration“ zum Open Access. Die Ergebnisse der durch sie geförderten Digitalisierungsprojekte stehen der weltweiten Forschung daher unentgeltlich und im Geiste dieser Erklärung zur Verfügung.³⁰ 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. Die digitale Ressource sollte in einer Form bereitgestellt werden, die eine wissenschaftliche Nutzung in anderen Forschungskontexten erlaubt (z.B. durch Angebot eines Images ohne Navigationskontext). Die Erhebung von Gebühren oder Entgelten für Kopien in darüber hinausgehenden Qualitäten, Derivaten oder für die Herstellung anderer Medienformen (CD, Druck, etc.) bleibt davon unberührt. Die Herkunft der Digitalisate sollte dabei auch in späteren Verwertungsumgebungen immer eindeutig erkennbar sein.

Die DFG erwartet, dass DFG-geförderte Projekte auf ihren im Internet angebotenen Daten eindeutige Herkunftsnachweise sowie ggf. den Hinweis auf die Förderung durch die DFG anbringen. Dies geschieht im Fall von Bilddigitalisaten in der Regel durch das Hinzurechnen einer Nachweisleiste in die veröffentlichte Nutzerkopie (zum Beispiel im Grafikformat JPEG), im Fall von Volltexten durch entsprechende Hinweise im Header der Textdatei. In Anhang A werden die technischen Spezifikationen und Formate ausführlich beschrieben. Diese gelten für alle DFG-geförderten Digitalisierungsprojekte und alle Materialgattungen.

Bei Projekten, in denen nicht nur gemeinfreie Materialien digitalisiert werden und daher mit kommerziellen Partnern oder Verlagen zusammengearbeitet wird, kann eine verzögerte Publikation („moving wall“) von bis zu einem Jahr nach Projektende vereinbart werden.

4.2 Mindestanforderungen für die Bereitstellungssysteme digitaler Dokumente

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 Dokumenten haben, gelten konkret folgende Mindestanforderungen. Diese betreffen einige Basisanforderungen und einen Katalog der mindestens zu realisierenden Funktionalität.

4.2.1 Basisanforderungen und Architektur

Das Bereitstellungssystem verknüpft die digitalisierten Bild- oder Volltextdateien zu einer Dokumentstruktur, um für den Benutzer 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 dem Benutzer, 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 zusammengehörige digitale Sammlungen bzw. Bestände, um dem Benutzer die Navigation zwischen Dokumenten und Sammlungen analog zu einer Freihandbibliothek in sachlicher Aufstellung zu bieten. Es stellt entsprechende Benutzeroberflächen für Recherche, Navi-

³⁰ <http://oa.mpg.de/openaccess-berlin/berlindeclaration.html>

gation, 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 Bibliotheken und Archive sollten sowohl bei der Navigation in digitalen Sammlungen/Beständen als auch bei der Recherche über Indizes und Register einen institutionenübergreifenden Zugriff ermöglichen. Darüber hinaus ist eine transparente Verknüpfung der Bereitstellungssysteme mit lokalen Katalogsystemen und Verbunddatenbanken wünschenswert.

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

- (1) Die Metadaten werden zentral in einem Katalogsystem (z. B. dem lokalen OPAC oder einem Bibliotheksverbundkatalog) gehalten, die digitalen Dokumentdateien (inkl. elektronischem Inhaltsverzeichnis und Register) 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.
- (2) Ein Dokumenten-Management-System (DMS) oder Content-Management-System (CMS) kommt zum Einsatz.

4.2.2 Funktionalitätsanforderungen

Unabhängig von der gewählten Architektur müssen mindestens folgende Funktionalitäten bereitgestellt werden:

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

- über die Website der anbietenden Institution
- über eine OAI Schnittstelle
- über den lokal implementierten oder an einem anderen Ort betriebenen sog. „DFG-Viewer“ (vgl. Kap. 5)
- über eine Suchanfrage an den lokalen und regionalen Bibliothekskatalog / an das lokale Online-Findmittelsystem
- über ein gemeinsames Portal der „Virtuellen Fachbibliotheken“ bzw. eines der von der DFG geförderten, materialspezifischen Portale, die einen integrierten Zugriff auf alle im DFG-Programm geförderten digitalen Sammlungen ermöglichen
- über Internet-Suchmaschinen.

Dem Benutzer 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 Bestandes erfordern.

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 (für Bücher ab 1850)³¹
- Metadaten-Info: Hier kann der Benutzer die Informationen aus den im DMS gespeicherten Beschreibungsfeldern zu „seinem“ 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 dem Benutzer signalisieren, an welcher Stelle des digitalen Dokuments er sich gerade befindet. Enthält ein Server Materialien, die konzeptuell vom Benutzer üblicherweise zu übergeordneten Einheiten zusammengefasst werden (mehrbändige Werke), so sollen diese übergeordneten Einheiten als solche sichtbar werden.

Darüber hinaus sind folgende Funktionen zu implementieren:

- Downloadfunktion³²
- Druckfunktion als PDF-Datei³³
- Aus zentralen DFG-geförderten Informationssystemen (VD16, VD17, Fachportalen etc.) sollte zuerst auf eine Ansicht im Style „DFG-Viewer“ verlinkt werden.

4.2.3 Technische Mindestanforderungen

Server sind so anzulegen, dass sie, soweit anwendbar³⁴:

- (1) 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.
- (2) Alle Materialien aber auch in einer Qualität anbieten, die über DSL Anschlüsse ohne unzumutbare Verzögerungen bearbeitbar sind.
- (3) Zum wissenschaftlichen Gebrauch den kostenlosen Download kompletter Einheiten in einer einzigen Datei (bspw. einzelne gedruckte Werke) ermöglichen.
- (4) Alle jeweils gängigen Browser unterstützen, soweit dies der Natur der Sache nach nicht unmöglich ist.³⁵

31 Nach dem Stand der gegenwärtigen Technik sollte eine OCR-Erkennung bei Drucken der Maschinenpressenzeit ab 1850 stets in Erwägung gezogen werden.

32 Ein Download nach Abschnitten oder einzelnen Seiten ist dann vorzusehen, wenn die Größe der Gesamtdatei nicht mehr handhabbar wäre.

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

34 Hier wird nachdrücklich das Kriterium der Alltagstauglichkeit gefordert, es geht nicht um die Realisierung abstrakter Wünschbarkeiten. Können Objekte eines Projekt ihrer Natur nach auf Bildschirmen mit einer Auflösung unter 1600 mal 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.

35 Wird ein für eine fortgeschrittene 3D-Anwendung notwendiges Format von einem Browser nicht unterstützt, braucht auch kein Aufwand betrieben zu werden, ein geeignetes Plugin zu entwickeln.

5. Präsentationsstandards („DFG-Viewer“) und Formate (METS/MODS)

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

- (1) Festlegung eines einheitlichen Designprofils für die Visualisierung von in DFG-Förderung erstellten Digitalisaten (sog. „DFG-Viewer“).
- (2) Schaffung einer definierten technischen Schnittstelle auf der Basis des METS Standards: Diese Schnittstelle dient zunächst der einheitlichen Anzeige von Images und ihren Metadaten in DFG-geförderten Projekten. Ziel ist es, über eine XML-Schnittstelle im METS/MODS-Format (bzw. künftig METS/TEI-P5 für Handschriften und METS/EAD für Archivmaterialien), die das Weiterblättern, Anzeige von Metadaten und andere grundsätzliche Funktionen beschreibt, einheitliche Anzeige- und Blätterfunktionen zu schaffen, die auch innerhalb von zentralen Nachweisportalen einen homogenen Zugriff auf dezentrale Ressourcen erlauben. Diese Schnittstelle ist mit dem METS/MODS AP zur Ablieferung von bibliographischen und Strukturdaten kompatibel und kann bei Bedarf durch Anreicherung zu einer vollwertigen Schnittstelle zur Datenlieferung, z.B. per OAI, ausgebaut werden. Der DFG-Viewer vermag auch Daten per OAI zu lesen.

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

6. Prüfliste für Antragsteller und Gutachter

Wie aus den bisherigen Abschnitten klar wird, bestehen bei der Planung von Digitalisierungsprojekten, selbst im Bereich der Digitalisierung von Bibliotheksgut im engeren Sinne, zahlreiche Wahlmöglichkeiten. Dies gilt um so mehr für Projekte, bei denen es um die Bereitstellung von Materialien geht, mit denen erst wenige Erfahrungen bestehen – man denke an 3D-Rekonstruktionen von Gebäuden, Globen etc. Zudem entwickeln sich die technischen Gegebenheiten ständig weiter.

Eine rein mechanische Verpflichtung auf ganz bestimmte Standards würde daher eine unzulässige Einschränkung der zu fördernden Projekte bedeuten und deren dynamische Weiterentwicklung unzulässig einschränken. Andererseits ist klar, wo die Gefahrenpunkte liegen. Die folgende Prüfliste – in die auch Empfehlungen aus anderen Bereichen als der Buchdigitalisierung eingehen – ist daher folgendermaßen zu verstehen.

- (1) Bei allen Anträgen mit Digitalisierungskomponenten werden unabhängig von ihrer inhaltlichen Begutachtung die vorgeschlagenen technischen Konzepte geprüft.
- (2) Antragsteller müssen in ihren Anträgen in der Regel nachvollziehbar belegen, dass die Durchführung des Projekts die im Folgenden aufgelisteten Standards einhält.
- (3) Jede Abweichung von diesen Standards muss ausführlich begründet werden.

- (4) Sieht ein Digitalisierungsprojekt vor, die im Folgenden angeführten Standards zu überschreiten, so ist, soweit dadurch erhöhte Kosten entstehen, deren Notwendigkeit im Einzelnen zu begründen.
- (5) 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.
- (6) Bei der Digitalisierung von Material der Kulturellen Überlieferung findet eine insgesamt gut verstandene Technik Verwendung. Es ist daher davon auszugehen, dass die dabei auftretenden Kosten laufend sinken. Bei Beantragung und Begutachtung ist daher darauf zu achten, dass spätere Projekte die in den Abschlussberichten früherer Projekte nachgewiesenen Kosten mindestens nicht überschreiten.
- (7) Zur besseren Vergleichbarkeit sollen Anträge (bzw. Arbeitsberichte) Auskünfte über die geschätzten (bzw. tatsächlichen) Kosten für das Scannen (die Rohdigitalisierung) pro Image enthalten.³⁶

36 Grundsätzlich besteht Interesse an den realen Kosten. Zu deren Berechnung gibt es jedoch zur Zeit keine allgemein akzeptierten Standards. Die Praxisregeln empfehlen daher, die Kosten des Scannens (Rohdigitalisierung) 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; DVDs für den 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.
 - (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 (Rohdigitalisierung) sind alle Kosten, die entstehen:

- (a) Durch das Projektmanagement (z.B.: Auswählen und Ausheben der zu digitalisierenden Stücke).
- (b) Durch die Erfassung wie auch immer gearteter Metadaten (Ausn.: Benennung und Speicherung der Dateien).
- (c) Durch die Langzeitarchivierung.
- (d) Indirekte Kosten, wie sie in der Regel im Rahmen betriebswirtschaftlicher Kostenrechnungen erhoben werden.

PRÜFLISTE für Antragsteller und Gutachter

6.1 Allgemeine technische Abläufe / Voraussetzungen

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

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

6.2 Datenqualität und -formate

- (1) Für alle im Rahmen der Projekte zu digitalisierenden Materialien 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.
- (2) Für die langfristige Sicherung (die Archivkopien) sind folgende Richtwerte einzuhalten (bzw. Abweichungen³⁷ von ihnen zu begründen):

Bilder:

Die Speicherung der Master erfolgt in der Regel in Form unkomprimierter Baseline-TIFF oder PNG-Dateien. Für monochrome Dateien ist die Verwendung von TIFF mit Group 4 Kompression zu empfehlen (→ 2.2.2).

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. Liegt kein ungewöhnliches Material vor, wird angenommen, dass eine Auflösung – bezogen auf die Größe der Originalvorlage – von 300 dpi bei Farb- und Grauwertbildern, 600 dpi bei monochromen Bildern dies sicherstellt.

Eine Digitalisierung in Form monochromer Bilder ist zu wählen, wenn eindeutig plausibel ist, dass ein Text frei von Bildern oder Grauabstufungen ist. Farbbilder sind als 24 Bit Bilder, Grauwertbilder als 8 Bitbilder zu speichern (→ 2.2.1.1. und 2.2.1.2).

Bei Unter- oder Überschreitung dieser Richtwerte ist der Begründung der Abweichung eine Serie von Testscans beizulegen, die die Möglichkeit / Notwendigkeit der Abweichung belegen.

37 Bei Abweichungen von den – bewusst sehr restriktiv gehandhabten – Dateiformatempfehlungen, wird dringend empfohlen die unter <http://www.digitalpreservation.gov/formats> angegebenen Präferenzen zu berücksichtigen. Die folgende Auflistung gibt, soweit anwendbar, eine Untermenge der eben zitierten Empfehlungen wieder. Darin wird Bezug genommen auch auf die ebenfalls berücksichtigte, auf praktischen Untersuchungen basierende Empfehlung des Florida Center for Library Automation: www.fcla.edu/digitalArchive/pdfs/recFormats.pdf

Audio:

Waveform Audio File Format (WAVE) mit linearem PCM Bitstream (essentiell: unkomprimiert) oder Audio Interchange File Format (AIFF) mit linearem PCM Bitstream.

Qualitätsempfehlungen können auf Grund der relativ geringen vorliegenden Erfahrungen noch nicht ausgesprochen werden.

Bewegtbilder (Video):

MPEG-1 und MPEG-2, mit einem der Profile: Simple, Main, oder 4:2:2. Für die langfristige Bearbeitbarkeit von Videosequenzen, bei denen die Fähigkeit zum präzisen Schnitt langfristig erhalten bleiben sollte, ist Motion JPEG zu berücksichtigen.

Qualitätsempfehlungen können auf Grund der relativ geringen vorliegenden Erfahrungen noch nicht ausgesprochen werden.

3D-Daten:

Nach aktueller Lage der internationalen Normierungsbestrebungen und der verfügbaren Werkzeuge sind X3D und COLLADA die empfohlene Wahl für die langfristig sichere Verwendung von 3D-Modellen. Von der Verwendung von VRML wird aus Gründen der zahlreichen Varianten und unterschiedlich ausgerichteten Viewer und Verarbeitungswerkzeuge abgeraten.

Datenbankinhalte:

Soweit Datenbanken als Zugriffsinstrumente auf Metadaten als Bestandteil von Digitalisierungsprojekten betrieben werden, sind die oben beschriebenen Anforderungen zu erfüllen. Wenn in einzelnen Projekten Datenbanken als Projektbestandteile anfallen, die über den Umfang von Metadaten hinausgehen, ist zu beachten:

Der Workflow ist so zu gestalten, dass neben der Speicherung der Daten im jeweiligen Datenbanksystem auch eine von der Software unabhängige Variante der Datenbankinhalte erzeugt wird und auch bei einem abrupten Projektende verfügbar bleibt (→ 2.6). Dazu eignen sich stets XML-Datenbankauszüge auf der Basis einer dokumentierten DTD, im Falle von SQL-fähigen Datenbanken auch die zur Erstellung der Datenbank geeigneten SQL-DDL-Statements (wobei auf die Verwendung von SQL-Dialekten möglichst verzichtet werden sollte).

Texte

Die Herstellung von Volltext kann auf zwei Wegen erfolgen: entweder durch OCR oder durch Abschreiben. Es wird empfohlen die Texte in Unicode abzuspeichern. Zu favorisieren ist UTF-8 oder UTF-16.

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 XML-Auszeichnungssprache (z.B. XSLT, XSL:FO), die die Unabhängigkeit von spezieller Software weitgehend sicherstellt.

Die Auswahl der wichtigsten Strukturdaten, die erfasst werden sollen, muß inhaltlich begründet werden.

Die Wahl des jeweiligen Texterfassungsverfahren sollte abhängig von den Anforderungen an die Textqualität begründet werden (→ 2.3).

6.3 Langzeitsicherung

Der Antrag muss nachvollziehbare Aussagen zur institutionellen Langzeitarchivierung enthalten. Digitalisate sind in redundanter Form zu archivieren. Die Abgabe eines kompletten Datensets an die Deutsche Nationalbibliothek wird – falls nicht ohnehin gesetzlich geregelt – ergänzend empfohlen (→ 2.4).

6.4 Zusammenarbeit mit Dienstleistern

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 Rechnungsbetrages als Sicherheit zurückgehalten und erst nach Durchführung der Qualitätskontrolle ausbezahlt wird. Der Projektnehmer hat darauf zu achten, dass das Urheberrecht oder andere Schutzrechte nicht missachtet werden (→ 2.5).

6.5 Metadaten

Jedes Digitalisat muss zumindest auf Titelebene nach einschlägigen bibliothekarischen oder archivischen Standards katalogisiert und in einem zentralen Nachweissystem (bibliothekarischer Verbund, zentrales Portal, Virtuelle Fachbibliothek etc.) nachgewiesen sein. Für Archive gelten analoge Regeln. Sofern Daten außerhalb vorhandener Bibliotheksverbünde oder zentraler Portale erfasst werden, wird die Archivierung von Zwischenergebnissen in XML erwartet (→ 2.6).

Zur Text- und Seitenbeschreibung und Beschreibung von Strukturelementen buchartiger Dokumente wird der Einsatz eines einheitlichen Strukturdatenformats empfohlen.³⁸ Für Archivalien sollte EAD, für mittelalterliche Handschriften TEI-P5 verwendet werden. In begründeten Fällen kann von dieser Empfehlung abgewichen werden (→ 2.6.1).

6.6 Austausch und Weitergabe

Digitalisate müssen über zentrale Portale nachgewiesen werden können. Mit Nachdruck empfohlen wird die Bereitstellung einer OAI-Schnittstelle, die neben Dublin Core METS/MODS ausliefert (s. Anhang A), so dass sichergestellt ist, dass einschlägige Portale die Daten harvesten können (→ 2.7).

Digitalisierungsprojekte müssen das in Anhang A angegebene METS/MODS-Format unterstützen sowie aus überregionalen Kontexten heraus primär auf eine Webpräsentation des Styles DFG-Viewer verlinken (→ 5).

6.7 Zitieren, persistente Adressierung

Sichergestellt werden soll mindestens die Erreichbarkeit und Zitierbarkeit eines Werkes als Ganzem, künftig auch die Erreichbarkeit und Zitierbarkeit von einzelnen physikalischen Seiten von 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 (→ 3).

6.8 Bereitstellung der Digitalisate, öffentlich zugängliche Schnittstellen

Digitalisate und Arbeitsergebnisse müssen kostenfrei über Internet angeboten werden. Eine „moving wall“ von maximal 1 Jahr kann vereinbart werden (→ 4.1).

Digitalisate sollen im Regelfall über unterschiedliche Wege zugänglich sein:

- (1) über die Website der anbietenden Bibliothek/des anbietenden Archivs
- (2) über eine OAI-Schnittstelle

³⁸ Vgl. auch <http://dfg-viewer.de/profil-der-strukturdaten>

- (3) über eine Suchanfrage an den lokalen und regionalen Bibliothekskatalog bzw. archivisches Online-Findmittel-System
- (4) über ein gemeinsames Portal der „Virtuellen Fachbibliotheken“ bzw. eines der von der DFG geförderten, materialspezifischen Portale, die einen integrierten Zugriff auf alle im DFG-Programm geförderten digitalen Sammlungen ermöglichen (→ 4.1.2).
- (5) über Internet-Suchmaschinen.
- (6) aus überregionalen Systemen heraus über eine Präsentation im Style „DFG-Viewer“ (→ 5).

Alle Materialien 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 Equipment mühelos möglich ist. Alle jeweils gängigen Browser müssen unterstützt werden, soweit dies der Natur der Sache nach nicht unmöglich ist (→ 4.1.2 und 4.1.3).

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:

- (1) Einen selbständigen Server, der das digitalisierte Material mit zu seiner Nutzung notwendigen Instrumenten bereitstellt.
- (2) Alle Digitalisate sind so anzubieten, dass langfristig zitierfähige URLs möglichst niedriger Granularität entstehen. Die Form, wie zitiert werden soll, ist deutlich anzuzeigen (→ 3).
- (3) Eine Schnittstelle im technischen Sinn, die einem für den jeweiligen Materialbereich von der DFG geförderten Portal den Zugriff auf alle im Rahmen des Projektes erstellten Metadaten ermöglicht, insbesondere eine OAI Schnittstelle, die neben DC METS/MODS (METS/TEI-P5, METS/EAD) Metadaten an geeignete Harvester liefert (→ 2.7).
- (4) Geeignete Maßnahmen, die dazu führen, dass die Metadaten von Suchmaschinen gefunden werden (→ 2.7).

7. Regeln für die Durchführung von Digitalisierungsprojekten

- (1) Zwischen der Förderzusage der DFG und dem tatsächlichen Arbeitsbeginn vergeht üblicherweise eine gewisse Zeit, während der die Hardware gekauft, Verträge mit den Dienstleistern geschlossen und die Kräfte des Projekts rekrutiert werden. Die in den Projektanträgen vorgesehenen Arbeitspläne gehen dagegen üblicherweise davon aus, dass das Projektteam vollständig und arbeitsfähig ist. Um den Verlauf der Projekte besser begleiten und Fehlentwicklungen besser gegensteuern zu können, wird daher erwartet, dass die Projektnehmer die beschriebenen vorbereitenden Tätigkeiten in einer Vorbereitungsphase durchführen, während derer wenig oder keine Personalmittel ausgegeben werden. Sind diese Vorbereitungen abgeschlossen, ist der DFG der tatsächliche Projektbeginn schriftlich mitzuteilen. Alle weiteren Fristen zählen ab diesem Zeitpunkt. Zwischen der Förderzusage und dem tatsächlichen Projektbeginn darf in der Regel nicht mehr als ein Jahr verstreichen.
- (2) Alle Projekte sind verpflichtet, zu einem Zeitpunkt, da eine Gegensteuerung gegen Fehlentwicklungen noch möglich ist, ein funktionsfähiges Modell des Internetangebotes zu präsentieren (in der Regel nach Ablauf des ersten Jahres der Förderung). Dieses Modell muss den Nachweis erbringen, dass die oben angeführten Mindestansprüche eingehalten werden. Alle Komponenten des technischen Angebotes müssen zu diesem Zeitpunkt prinzipiell funktionsfähig sein. Entstehen Rückfragen oder Zweifel, ist das Internetangebot ggf. vor Ort zu demonstrieren.
- (3) Erst nachdem der Server in seinen Grundfunktionalitäten abgenommen ist und monierte Mängel abgestellt sind, sind Verlängerungsanträge möglich.
- (4) Als Bestandteil aller Zwischenberichte ist anzugeben:
 - (a) Ein wie großer Bestandteil des Mengengerüsts tatsächlich digitalisiert wurde.
 - (b) Ein wie großer Bestandteil des Mengengerüsts im Internet bereitsteht.
 - (c) Welche aussagekräftigen Zugriffszahlen sich aus den Logdateien des Servers ablesen lassen.

Wesentliche Unterschreitungen der im Projektantrag beschriebenen Arbeitsgeschwindigkeit sind häufig ein Zeichen ungenügender Projektorganisation und stellen keine erfolgversprechende Begründung für Verlängerungsanträge dar. Sind sie signifikant, ist auch die Förderung bereits ursprünglich angekündigter Projektphasen kritisch zu überprüfen.

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/VD 17, ZVDD, Virtuelle Fachbibliotheken) 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 Viewer³⁹ wurde im Rahmen der Aktionslinie „Digitalisierung des VD 16 / VD 17“ durch die in der ersten Ausschreibungsrunde geförderten Bibliotheken aufgebaut. Er wird durch diese 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 das von der Website des Viewers abzurufende XML Schema valide sein.⁴⁰

Zur Darstellung von Metadaten im Viewer wird METS⁴¹ 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 bibliographischer Metadaten (nur Drucke) nach MODS⁴² 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 Formates befindet sich auf der Homepage der Referenzanwendung des DFG-Viewers.

Diese Vorgaben gelten zunächst nur für den Bereich der gedruckten Überlieferung (METS/MODS). Die Nutzung des Viewers zur Präsentation anderer Materialien (z.B. handschriftlicher Überlieferung) ist in Vorbereitung und wird sich an den TEI-P5 (METS/TEI-P5) und EAD (METS/EAD) orientieren.

2. MODS DFG-Standard-Set (Printbestände)

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).

39 <http://www.dfg-viewer.de/>

40 <http://dfg-viewer.de/profil-der-metadaten>

41 <http://www.loc.gov/standards/mets/>

42 <http://www.loc.gov/standards/mods/>

Element/Unterelement	Wiederholbar	Bemerkungen / Erläuterungen	Status
----------------------	--------------	-----------------------------	--------

1 Titelinformation

<titleInfo>	ja	Titelinformationen; sofern das Werk keinen Titel enthält, muss er fingiert werden.	Pflicht
<titleInfo> / <title>	nein	Enthält den Haupttitel des Werkes.	Pflicht
<titleInfo> / <subTitle>	nein	Enthält den Zusatz zum Haupttitel der Werkes	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 „pnd“.	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) ⁴³	

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>	ja	Erscheinungsvermerk; der erste <originInfo>-Block ist für Angaben zur Quelle vorgesehen, Angaben zur digitalen Edition können in einem zweiten <originInfo> Block 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

43 Ein Code aus der Liste *MARC Value List for Relators and Roles*:
<http://www.loc.gov/marc/sourcecode/relator/relatorlist.html>.

Element/Unterelement	Wiederholbar	Bemerkungen / Erläuterungen	Status
----------------------	--------------	-----------------------------	--------

5 Ausgabevermerk

<originInfo> / <edition>	ja	Enthält die Ausgabebezeichnung	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

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 type="host" 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 Werkes 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 Druckes (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. Beispiel für einen METS/MODS Datensatz nach dem DFG-Standard

```

<mets:mets xmlns:mets="http://www.loc.gov/METS/"
  xmlns:xlink="http://www.w3.org/1999/xlink"
  xmlns:mods="http://www.loc.gov/mods/v3"
  xmlns:dv="http://dfg-viewer.de/">
  <mets:dmdSec ID="dmd_586140484">
    <mets:mdWrap MDTYPE="MODS">
      <mets:xmlData>
        <mods:mods>
          <mods:recordInfo>
            <mods:recordIdentifier source="WDB_OPAC"
              >oai:diglib.hab.de:ppn_586140484</mods:recordIdentifier>
          </mods:recordInfo>
          <mods:location>
            <mods:physicalLocation>Herzog August Bibliothek Wolfenbüttel</mods:physicalLocation>
            <mods:shelfLocator>M: Gn 4° 1572:1</mods:shelfLocator>
          </mods:location>
          <mods:identifier type="purl"
            >http://diglib.hab.de/drucke/gn-4f-1572-1b/start.htm</mods:identifier>
          <mods:identifier type="urn"
            >urn:nbn:de:gbv:23-drucke/gn-4f-1572-1b3</mods:identifier>
          <mods:language>
            <mods:languageTerm type="code" authority="iso639-2b"
              >lat</mods:languageTerm>
          </mods:language>
          <mods:name type="personal">
            <mods:namePart type="family">Leibniz</mods:namePart>
            <mods:namePart type="given">Gottfried Wilhelm</mods:namePart>
            <mods:displayForm>Leibniz, Gottfried Wilhelm</mods:displayForm>
            <mods:namePart type="date">1646-1716</mods:namePart>
            <mods:role>
              <mods:roleTerm type="code" authority="marcrelator">ctb</mods:roleTerm>
            </mods:role>
          </mods:name>
          <mods:originInfo>
            <mods:place>
              <mods:placeTerm type="text">Hanoverae</mods:placeTerm>
            </mods:place>
            <mods:publisher>Foerster</mods:publisher>
            <mods:dateIssued keyDate="yes" encoding="w3cdtf"
              >1707</mods:dateIssued>
          </mods:originInfo>
          <mods:originInfo>
            <mods:place>
              <mods:placeTerm type="text"
                >Wolfenbüttel</mods:placeTerm>
            </mods:place>
            <mods:publisher>Herzog August Bibliothek</mods:publisher>
            <mods:dateIssued encoding="w3cdtf">2008</mods:dateIssued>
            <mods:edition>[Electronic ed.]</mods:edition>
          </mods:originInfo>
          <mods:physicalDescription>
            <mods:extent>2°</mods:extent>
            <mods:extent>[22] Bl., 1004 S., [1] Bl</mods:extent>
            <mods:digitalOrigin>reformatted digital</mods:digitalOrigin>
          </mods:physicalDescription>
          <mods:titleInfo>
            <mods:title>Scriptores Rerum Brnsvicensium Illustrationi
              Inservientis, Antiqui Omnes Et Religionis Reformatione
              Priores ...</mods:title>
          </mods:titleInfo>
          <mods:relatedItem type="host">
            <mods:recordInfo>

```

```

        <mods:recordIdentifier source="WDB_OPAC"
          >oai:diglib.hab.de:pnp_58614045X</mods:recordIdentifier>
      </mods:recordInfo>
    </mods:relatedItem>
    <mods:part type="host" order="1">
      <mods:detail>
        <mods:number>[1]</mods:number>
      </mods:detail>
    </mods:part>
  </mods:mods>
</mets:xmlData>
</mets:mdWrap>
</mets:dmdSec>
<mets:amdSec ID="amd_586140484">
  <mets:rightsMD ID="amd_dvrights_586140484">
    <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>auskunft@hab.de/</dv:ownerContact>
        </dv:rights>
      </mets:xmlData>
    </mets:mdWrap>
  </mets:rightsMD>
  <mets:digiprovMD ID="amd_dvlinks_586140484">
    <mets:mdWrap MDTYPE="OTHER" OTHERMDTYPE="DVLINKS">
      <mets:xmlData>
        <dv:links>
          <dv:reference>
            http://sunny.biblio.etc.tu-
            bs.de:8080/DB=2/SET=2/TTL=2/CMD?ACT=SRCHA&IKT=1016&SRT=YO
            P&TRM=url+diglib.hab.de%5C%2Fdrucke%5C%2Fgn-4f-1572-
            1b%5C%2Fstart.htm
          </dv:reference>
          <dv:presentation>http://diglib.hab.de/drucke/gn-4f-1572-1b/start.htm</dv:presentation>
        </dv:links>
      </mets:xmlData>
    </mets:mdWrap>
  </mets:digiprovMD>
</mets:amdSec>
<mets:fileSec>
  <mets:fileGrp USE="DEFAULT">
    <mets:file ID="drucke_gn-4f-1572-1b_00001" MIMETYPE="image/jpeg">
      <mets:FLocat
        xlink:href="http://diglib.hab.de/drucke/gn-4f-1572-1b/00001.jpg"
        LOCTYPE="URL"/>
    </mets:file>
    <!-- Kürzung um 00001.jpg bis 01057.jpg DEFAULT Auflösung -->
    <mets:file ID="drucke_gn-4f-1572-1b_01058" MIMETYPE="image/jpeg">
      <mets:FLocat
        xlink:href="http://diglib.hab.de/drucke/gn-4f-1572-1b/01058.jpg"
        LOCTYPE="URL"/>
    </mets:file>
  </mets:fileGrp>
  <mets:fileGrp USE="MIN">
    <mets:file ID="min_drucke_gn-4f-1572-1b_00001" MIMETYPE="image/jpeg">
      <mets:FLocat
        xlink:href="http://diglib.hab.de/drucke/gn-4f-1572-1b/min/00001.jpg"
        LOCTYPE="URL"/>
    </mets:file>
    <!-- Kürzung um 00002.jpg bis 01057.jpg Minimale Auflösung -->
  </mets:fileGrp>

```

```

<mets:file ID="min_drucke_gn-4f-1572-1b_01058" MIMETYPE="image/jpeg">
  <mets:FLocat
    xlink:href="http://diglib.hab.de/drucke/gn-4f-1572-1b/min/01058.jpg"
    LOCTYPE="URL"/>
</mets:file>
</mets:fileGrp>
<mets:fileGrp USE="MAX">
  <mets:file ID="max_drucke_gn-4f-1572-1b_00001" MIMETYPE="image/jpeg">
    <mets:FLocat
      xlink:href="http://diglib.hab.de/drucke/gn-4f-1572-1b/max/00001.jpg"
      LOCTYPE="URL"/>
  </mets:file>
  <!-- Kürzung um 00002.jpg bis 01057.jpg Maximale Auflösung -->
  <mets:file ID="max_drucke_gn-4f-1572-1b_01058" MIMETYPE="image/jpeg">
    <mets:FLocat
      xlink:href="http://diglib.hab.de/drucke/gn-4f-1572-1b/max/01058.jpg"
      LOCTYPE="URL"/>
  </mets:file>
</mets:fileGrp>
<mets:fileGrp USE="THUMBS">
  <mets:file ID="thumbs_drucke_gn-4f-1572-1b_00001" MIMETYPE="image/jpeg">
    <mets:FLocat
      xlink:href="http://diglib.hab.de/drucke/gn-4f-1572-1b/thumbs/00001.jpg"
      LOCTYPE="URL"/>
  </mets:file>
  <!-- Kürzung um 00002.jpg bis 01057.jpg Thumbnails -->
  <mets:file ID="thumbs_drucke_gn-4f-1572-1b_01058" MIMETYPE="image/jpeg">
    <mets:FLocat
      xlink:href="http://diglib.hab.de/drucke/gn-4f-1572-1b/thumbs/01058.jpg"
      LOCTYPE="URL"/>
  </mets:file>
</mets:fileGrp>
</mets:fileSec>
<mets:structMap TYPE="LOGICAL">
  <mets:div ID="logMD_586140484" TYPE="Monograph" DMDID="dmd_586140484"
    ADMID="amd_586140484"/>
</mets:structMap>
<mets:structMap TYPE="PHYSICAL">
  <mets:div ID="physMD_586140484" TYPE="physSequence">
    <mets:div ID="physMD_586140484_1" TYPE="page" ORDER="1">
      <mets:fptr FILEID="drucke_gn-4f-1572-1b_00001"/>
      <mets:fptr FILEID="min_drucke_gn-4f-1572-1b_00001"/>
      <mets:fptr FILEID="max_drucke_gn-4f-1572-1b_00001"/>
      <mets:fptr FILEID="thumbs_drucke_gn-4f-1572-1b_00001"/>
    </mets:div>
    <!-- Kürzung um Zuordnungen von 00002 bis 01057 -->
    <mets:div ID="physMD_586140484_1058" TYPE="page" ORDER="1058">
      <mets:fptr FILEID="drucke_gn-4f-1572-1b_01058"/>
      <mets:fptr FILEID="min_drucke_gn-4f-1572-1b_01058"/>
      <mets:fptr FILEID="max_drucke_gn-4f-1572-1b_01058"/>
      <mets:fptr FILEID="thumbs_drucke_gn-4f-1572-1b_01058"/>
    </mets:div>
  </mets:div>
</mets:structMap>
<mets:structLink>
  <mets:smLink xlink:from="logMD_586140484" xlink:to="physMD_586140484" />
</mets:structLink>
</mets:mets>

```

4. 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 Bibliothekslogo, 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:

