

Abstract

In großen Digitalisierungsprojekten, in deren Rahmen viele verschiedene Archivbestände Gegenstand der Digitalisierung sind, können die Anforderungen sehr vielfältig sein. Sowohl bei einer Inhouse- als auch bei einer externen Digitalisierung muss daher sichergestellt werden, dass die Qualität der gelieferten Daten den in der Leistungsbeschreibung formulierten Grundsätzen entspricht. Bei einem Mengengerüst von mehreren zehntausend Verzeichnungseinheiten ist darüber hinaus die Dokumentation der Fehler, das Fehlermanagement gegenüber einem externen Dienstleister sowie die Deklaration der „gültigen“ Version der neu entstandenen Repräsentation eine große Herausforderung. Die Qualitätssicherung ist somit aus quantitativer und organisatorischer Sicht ein nicht zu unterschätzender Posten in einem Digitalisierungsprojekt.

ArchivInForm hat für die Begleitung von Digitalisierungsprojekten als externer Dienstleister eine Software entwickeln lassen, die den gesamten Prozess der Qualitätssicherung im Pre-Ingest abbildet, bevor der Auftraggeber die digitalisierten Archivalien in sein Digitales Archiv übernimmt. Diese wird im folgenden Beitrag näher vorgestellt.

Ein Pre-Ingest-Tool für die Qualitätssicherung in Digitalisierungsprojekten? Christoph Jobs (ArchivInForm GmbH)

Einführung

An die Retrokonversion archivischer Findmittel als Grundlage für die Onlinepräsentation von Archivgut schließt sich nun die großflächige Digitalisierung der Bestände selbst an.¹ Auch wenn die Archive seit einiger Zeit oft genutzte Bestände digitalisieren, ist die Herausforderung die Menge der in den Archiven verwahrten und noch nicht digitalisierten Bestände fachgerecht zu digitalisieren, groß.² Neben der Diskussion über den Aufbau Digitaler Archive und der Etablierung nationaler und internationaler Portale, waren zunächst vor allem die Zielstellung der Digitalisierung und die Kriterienfindung für die Priorisierung zu digitalisierender Bestände diskursbestimmend. Um Digitalisierungsprojekte für archivische Belange zu standardisieren, erarbeitete das von der Deutschen Forschungsgemeinschaft (DFG) geförderte Forschungsprojekt *Digitalisierung archivalischer Quellen* Kriterien für die fachgerechte Archivgutdigitalisierung und erprobte Beispielworkflows für die Planung und Durchführung von Digitalisierungsprojekten. Hieraus entstand eine Förderlinie der DFG.³

Die Digitalisierungsvorhaben können dabei verschiedene Ziele verfolgen. Einerseits werden die Originale nach einer Digitalisierung geschont, da sie nicht mehr für die Benutzung ausgehoben werden müssen. Andererseits erlaubt die Digitalisierung eine orts- und zeitunabhängige Benutzung und eine damit einhergehende Verbesserung der Sichtbarkeit der Archive für potenzielle Benutzerinnen und Benutzer.⁴ Die Bestände können darüber hinaus durch automatische Textanalysen und die Anreicherung der Daten mit Normdaten tiefer inhaltlich erschlossen werden, wodurch in Zukunft ganz neue Nutzungsszenarien entstehen werden.

Für die Erreichung der Digitalisierungsziele, müssen die digitalen Reproduktionen, die in Digitalisierungsprojekten entstehen, der ausgeschriebenen Leistung (*Lastenheft*) und somit den Ansprüchen des *Digitalen Archivs* und seiner *vorgesehenen Zielgruppe* entsprechen. Die in einem Digitalisierungsprojekt zu absolvierenden Tätigkeiten gehen dabei weit über das reine Scannen hinaus und reichen von der Projektplanung und -steuerung bis zur Archivierung und Onlinestellung. Ein Teil dieses Prozesses ist die Qualitätssicherung, die in der Fachliteratur in sehr lineare Gesamtprozesse eingebettet und häufig als in sich geschlossener Teilprozess dargestellt wird.⁵ Der

1 Vgl. Konferenz der Leiterinnen und Leiter der Archivverwaltungen des Bundes und der Länder (KLA) (2017a): Wirtschaftliche Digitalisierung in Archiven – Empfehlungen der Konferenz der Leiterinnen und Leiter der Archivverwaltungen des Bundes und der Länder (KLA), S. 1. Online verfügbar unter: https://www.bundesarchiv.de/imperia/md/content/bundesarchiv_de/fachinformation/ark/2017-01-04_kla_fta_wirtschaftliche_digitalisierung_final.pdf, zuletzt geprüft am 24.05.2018.

2 Vgl. Glauert, Mario (2013): Dimensionen der Digitalisierung. Kosten, Kapazitäten und Konsequenzen. In: Claudia Kauertz (Hrsg.): Digital und analog. Die beiden Archivwelten : Beiträge / 46. Rheinischer Archivtag, 21. - 22. Juni 2012 in Ratingen, Bonn: Habelt, S. 51ff.; Vgl. auch Stumpf, Marcus (2014): Digitalisierungsstrategien in Deutschland – Versuch einer Bestandsaufnahme. In: Archivpflege in Westfalen-Lippe (80), S. 34.

3 Vgl. Beck, Emily (2015): Pilotprojekt "Digitalisierung von archivalischen Quellen" - Gefördert durch die Deutsche Forschungsgemeinschaft (DFG). Online verfügbar unter: <https://www.archivschule.de/DE/forschung/forschungsprojekte/digitalisierung/dfg-projekt-archivgutdigitalisierung.html>, zuletzt geändert am 13.08.2015, zuletzt geprüft am 24.05.2018; Schleiter, Klaus (2018): Digitalisierung archivalischer Quellen – Ausschreibung Digitalisierung archivalischer Quellen der DFG. Online verfügbar unter: <https://www.archivschule.de/DE/forschung/digitalisierung-archivalischer-quellen/>, zuletzt geändert am 19.01.2018, zuletzt geprüft am 24.05.2018.

4 KLA (2017a): S. 2.; Vgl. auch Grau, Bernhard (2016): Digitalisierung von Archivgut – Was und warum. In: Archive in Bayern 9, S. 134.

5 Vgl. Glauert (2013): S. 49; Vgl. auch Stockert, Harald (2015): Musterworkflow für die Planung und Durchführung einer Inhouse Digitalisierung und einer Digitalisierung durch einen externen Dienstleister. Online verfügbar unter: https://www.archivschule.de/uploads/Forschung/Digitalisierung/Handreichungen/Idealtypischer_Digitalisierungsworkflow_fuer_Inhouse- und externe_Digitalisierung.pdf, zuletzt geprüft am 24.05.2018; Vgl. auch Konferenz der Leiterinnen und Leiter der Archivverwaltungen des Bundes und der Länder (KLA) (2017b): Handreichung zur Digitalisierung von Archivgut, S. 7. Online verfügbar unter: <http://www.bundeskonferenz->

vorliegende Aufsatz reflektiert sie hinsichtlich ihrer Bedeutung für ein nachhaltiges Digitalisierungsvorhaben, formalisiert die Prozesse und überträgt sie auf die Qualitätssicherungssoftware Panoptes.

Bedeutung von Qualitätssicherung für Digitalisierungsprojekte

Um langfristige Auswirkungen von Fehlern aus dem Digitalisierungsprozess zu vermeiden, ist sicherzustellen, dass die Digitalisierung von Anfang an den im Lastenheft festgelegten Qualitätsmaßstäben entspricht. Gerade im Hinblick auf die begrenzten personellen und finanziellen Ressourcen der Archive ist die zeitnahe und projektbegleitende Qualitätssicherung die beste Möglichkeit sie strukturiert und effizient zu absolvieren. Qualitätsmängel, die bei der späteren Benutzung festgestellt werden, haben einen Einfluss auf die Vertrauenswürdigkeit und Transparenz des Digitalen Archivs.⁶ Darüber hinaus sind retrospektive Nacharbeiten ungleich aufwändiger als eine zeitnahe Prüfung, auch wenn diese selbst im Aufwand nicht zu unterschätzen ist.⁷

Ein Teil des Aufwands für die Qualitätssicherung kann durch Softwareunterstützung reduziert werden. Dies betrifft sowohl die Prüfung der technischen Parameter, als auch die Unterstützung des intellektuellen Prüfprozesses (Ziehung von Stichproben, Ausgabe von Fehlerprotokollen, standardisierte Fehlereingabe und –verknüpfung). Der Aufwand und die Komplexität der intellektuellen Sichtprüfung hängen aber letztlich immer von der Beschaffenheit der Bestände ab. So sind auch die Ergebnisse der Qualitätssicherung bei homogenen Beständen besser auf den nicht geprüften Teil des digitalisierten Bestands übertrag- und verallgemeinerbar.⁸ Bei heterogenen Beständen sind die Fehler sehr individuell, sodass eine hohe Prüfquote anzustreben ist.⁹ Die Stichproben sind daher flexibel daran anzupassen und müssen von Bestand zu Bestand variieren können.

Einordnung der Qualitätssicherung in den Kontext der Digitalen Archivierung

Die Qualitätssicherung kann als Schnittstelle zwischen dem Produzenten und dem Digitalen Archiv verstanden werden. Dabei ist der Produzent in diesem Fall kein Informationsproduzent bzw. Provenienz einer intellektuellen Entität, sondern der Produzent einer Repräsentation (Digitale Reproduktion des Originals). Die Qualitätssicherung ist dafür verantwortlich, dass die Grundsätze des Digitalen Archivs für die Übernahme von Informationen eingehalten werden. Entsprechend der DIN-Norm 31644 sind daher v. a. die Kriterien 14, 17 und 18, die die Aufnahmeschnittstelle betreffen für die Qualitätssicherung relevant.

Das *Kriterium 14: Integrität Aufnahmeschnittstelle* stellt die Erfordernis, dass die Aufnahmeschnittstelle es „ermöglicht den Produzenten und der Administration des digitalen Langzeitarchivs, die Integrität der Repräsentationen zu überprüfen“.¹⁰ Hieraus ergibt sich die Anforderung an das Durchführen von Integritätschecks. Das *Kriterium 17: Authentizität: Aufnahme* erfordert, dass „das digitale Langzeitarchiv Verfahren ein[setzt], die die Beurteilung der Authentizität

kommunalarchive.de/empfehlungen/Handreichung_Digitalisierung_von_Archivgut_endfassung.pdf, zuletzt geprüft am 24.05.2018.

⁶ Vgl. Conway, Paul (2011): Archival quality and long-term preservation: a research framework for validating the usefulness of digital surrogates. In: Archival Science 11 (3-4), S. 297.

⁷ Vgl. KLA (2017b): S. 7.

⁸ Vgl. Stadtarchiv Mannheim – Institut für Stadtgeschichte (o. J.): Abschlussbericht des Produktivpiloten Digitalisierung von archivalischen Quellen, S. 37. Online verfügbar unter:

https://www.archivschule.de/uploads/Forschung/Digitalisierung/Ergebnisse/ISG_StadtA_MA_Abschlussbericht_DFG_Archivgutdigitalisierung.pdf, zuletzt geprüft am 24.05.2018.

⁹ Der Abschlussbericht zum Projekt Digitalisierung von archivalischen Quellen der Generaldirektion der Staatlichen Archive Bayerns empfiehlt diesbezüglich eine lückenlose Prüfung. Vgl. Generaldirektion der Staatlichen Archive Bayerns (o. J.):

Abschlussbericht des Produktivpiloten Digitalisierung von archivalischen Quellen, S. 8. Online verfügbar unter:

https://www.archivschule.de/uploads/Forschung/Digitalisierung/Ergebnisse/GDA_Abschlussbericht_DFG-Archivgutdigitalisierung.PDF, zuletzt geprüft am 24.05.2018; Vgl. auch Vgl. auch KLA (2017b): S. 7.

¹⁰ Vgl. Keitel, Christian; Schoger, Astrid (2013): Vertrauenswürdige digitale Langzeitarchivierung nach DIN 31644. Berlin: Beuth (Beuth-Kommentar), S. 57.

der Repräsentationen bei der Aufnahme sowie die Beurteilung und Sicherung der Authentizität der Transferpakete ermöglichen“.¹¹ Es ist zu prüfen, ob das „Informationsobjekt inhaltlich genau das ist, was es zu sein vorgibt“.¹² Darüber hinaus verlangt *Kriterium 18: Authentizität: Erhaltungsmaßnahmen*, dass „Das digitale Langzeitarchiv Verfahren ein[setzt], die bei der Durchführung der Langzeiterhaltungsmaßnahmen die Authentizität der Objekte sicherstellen bzw. den Grad an Authentizität dokumentieren“.¹³

Dies impliziert vor allem auch Verantwortlichkeiten, die die *inhaltliche* Prüfung der vom Produzenten zu übernehmenden Repräsentationen erfordert und eine große Auswirkung auf die Vertrauenswürdigkeit des Digitalen Archivs haben. So gilt, sofern man das Archiv bzw. die Archivierung ganzheitlich versteht,¹⁴ Kriterium 18 auch für die Transformation analogen Archivguts in eine digitale Repräsentationsform, da es sich streng genommen auch um eine Erhaltungsmaßnahme handelt, bei der der Erhalt der signifikanten Eigenschaften gefährdet ist.¹⁵ Daher ist der Prozess der Qualitätssicherung im Detail zu dokumentieren. Das bezieht sich sowohl auf die technischen Prüfungen, als auch auf archivfachlich geschultes Personal durchgeführte intellektuelle Prüfungen, sodass die Prüfverfahren während der späteren Nutzung der digitalisierten Archivalien noch nachvollziehbar sind.

Formalisierung von Qualitätssicherungsprozessen

Erfahrungen mit Digitalisierungsprojekten führten bei ArchivInForm zu der Erkenntnis, dass es für das Absolvieren der Qualitätssicherung unerlässlich ist, die neu entstehenden Repräsentationen kontrolliert zu verarbeiten. Die bisherige Begleitung von Digitalisierungsprojekten erfolgte mit Hilfe von Excel-Dateien. Die Digitalisate wurden in Dateiodnern geliefert, ohne für den Qualitätssicherungsprozess eine echte Verknüpfung/Kontrolle über die Digitalisierungsdaten und Verzeichnungsinformationen herzustellen.

Das führte dazu, dass die Prüfung nur schwer formalisierbar war und auch die Prüfung nach technischen Gesichtspunkten, obwohl Softwaretools eingesetzt wurden,¹⁶ durch diverse Medienbrüche doch sehr subjektiv, fehleranfällig und schwer nachvollziehbar erschien (Anmutungsqualität). Die Idee zu Panoptes entstand vor diesem Hintergrund mit dem Ziel alle Prozesse bei der Qualitätssicherung abzudecken. Von der Übergabe der analogen Archivalien über Rücknahme, technische Prüfung, Stichprobenziehung und Sichtprüfung bis zur Übergabe der korrekten qualitätsgeprüften Versionen an den Auftraggeber kann sie alle Prozessschritte abdecken und die Kommunikation mit allen Projektbeteiligten erleichtern.¹⁷

Mit dieser Zielstellung wurde ein Prozess für die Formalisierung der Qualitätssicherung entwickelt, welcher immer einen Prozessstrang für den Lebenszyklus des analogen Archivale im Projekt abbildet und einen, den die digitale Repräsentation durchläuft.

11 Keitel, Schoger (2013): S. 59.

12 Ebd.

13 Ebd. S. 60.

14 Im Sinne der Verantwortlichkeit für das in Obhut befindliche analoge und digitale Archivgut.

15 Der Definition von signifikanten Eigenschaften liegt in diesem Fall aber eine Einbeziehung inhaltlicher Maßstäbe zugrunde.

16 Wie z.B. DPF-Manager (<http://dpfmanager.org/>, zuletzt geprüft am 24.05.2018); EXIFTool (<http://owl.phy.queensu.ca/~phil/exiftool/>, zuletzt geprüft am 24.05.2018).

17 Der Fokus liegt zunächst auf der formalen, inhaltlichen und strukturellen Prüfung und der Abbildung eines medienbruchfreien Prozesses. Die Anwendung automatisierter Prüfverfahren hinsichtlich der Aufnahmequalität (bspw. mittels Universal Test Target (UTT)) war in bisherigen Projekten nicht vorgesehen. Überlegungen zum Qualitätsmanagement mit UTT gibt es bspw. von Röschner, Matthias (2016): Alles unter Kontrolle? Möglichkeiten und Grenzen des Qualitätsmanagements bei der Zusammenarbeit mit externen Scandienstleistern. In: Archive in Bayern 9, S. 213-220.

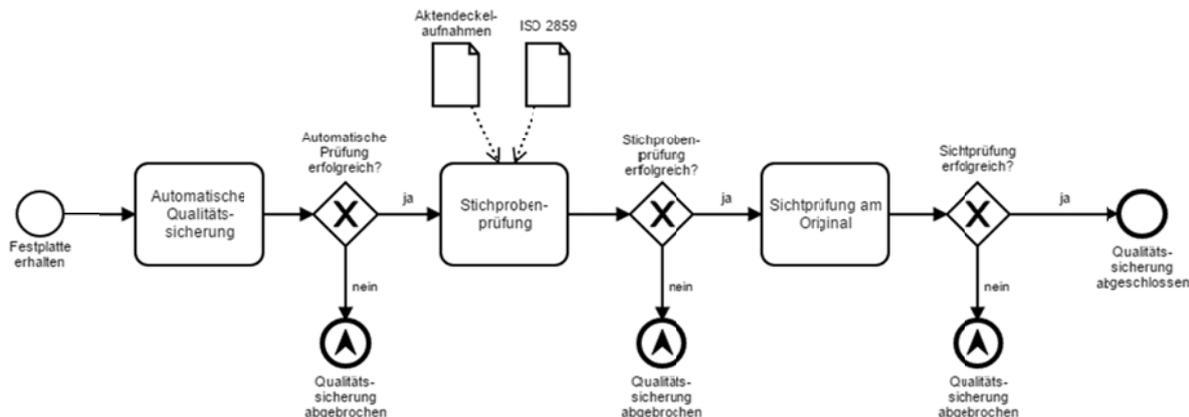


Abbildung 1: Prozess der Qualitätssicherung (stark vereinfachter Prozess)¹⁸

Geprüft wird vom Allgemeinen zum Speziellen. Die Aufgabe *Automatische Qualitätssicherung* besteht dabei aus sechs Teilprozessen. Dazu gehören folgende Prüfungen:

- Abgleich der Prüfsummen,
- Abgleich der Benennung mit der erwarteten Benennung,
- Import der vorgegebenen Masterdigitalisate und Derivate,
- Mengenabgleich und –übereinstimmung der Benennung der versch. Repräsentationsformen (z.B. Lieferung im TIFF-Format, JPEG-Format und PDF/A-Format¹⁹),
- Validierung der Formate gemäß Format-Spezifikation²⁰,
- Validierung von Parametern (z.B. über die Metadaten)²¹

Durch diesen Schritt wird die *rational bedingte Sachqualität*²² geprüft, die objektiv messbar ist und strukturell bedingte Verarbeitungsfehler vollständig identifizieren kann. Alle weiteren Schritte bestehen aus Prüfungen hinsichtlich der *Anmutungsqualität*²³, die subjektiver Natur ist und vor allem individuelle Fehler identifizieren kann. Hierfür sind in der Probephase eines Projekts Prüfkriterien festzulegen und zu bestimmen, wie die intellektuellen Prüfungen zu dokumentieren sind.²⁴

18 Für ein detaillierteres Modell siehe: Jobs, Christoph (2018): Ein Pre-Ingest-Tool für die Qualitätssicherung in Digitalisierungsprojekten? Ein Praxisbericht, Vortragsfolien, S. 12. Online verfügbar unter: https://www.staatsarchiv.sg.ch/home/auds/22/jcr_content/Par/downloadlist_1638086843/DownloadListPar/download_1552214939.ocFile/03_Ingest_Jobs.pdf, zuletzt geprüft am 24.05.2018.

19 An dieser Stelle nicht die einzelnen Dateien, sondern die Anzahl der in die PDF/A-Datei eingebetteten Dateien.

20 Durch Einbindung verschiedener Formatvalidatoren, z.B. JHOVE, Callas pdfaPilot.

21 Der im PREFORMA-Projekt entwickelte DPF Manager unterscheidet hierbei bspw. zwischen *Implementation Checker* und *Policy Checker*, vgl. Ohne Autor (o. J.): DPF Manager User Manual, S. 13ff. Online verfügbar unter: <http://dpfmanager.org/Downloads/User%20Manual.pdf>, zuletzt geprüft am 24.05.2018.

22 Für die Definition der zwei Qualitätstypen vgl. Markgraf, Daniel (2018): Qualität. In: Gabler Wirtschaftslexikon. Online verfügbar unter: <https://wirtschaftslexikon.gabler.de/definition/qualitaet-45908/version-269195>, zuletzt geändert am: 16.02.2018, zuletzt geprüft am 24.05.2018.

23 Vgl. ebd.

24 Bisher sind bzgl. der Formalisierung der Qualitätssicherung vor allem die Werkzeuge für Digitalisierungsprojekte im Landesarchiv Nordrhein-Westfalen zu erwähnen, die u. a. aus verschiedenen Prozessmodellen, einem Projektablaufplan und einer Tabelle mit standardisierten Fehlerbeschreibungen bestehen. Vgl. Kistenich, Johannes (2013): Werkzeuge für Digitalisierungsprojekte im Landesarchiv Nordrhein-Westfalen, Standardworkflow, technische Vorgaben, Qualitätssicherung, Projektplanung, In: Archivar 66, Heft 2, S. 244; Für weitere Fehlermodelle etc. vgl. auch Holzapfl, Julian; Makowski, Stephan (2014): Qualitätssicherung in archivischen Digitalisierungsprojekten. Erfahrungen und Diskussionsanregungen auf dem Weg zur Professionalisierung. Vortrag auf der Konferenz Made digital - Wege und Ziele am 26.11.2014 in Koblenz. Online verfügbar unter: https://www.bundesarchiv.de/imperia/md/content/bundesarchiv_de/fachinformation/informationstechnologie/digitalisier_tesarchivgut/holzapflmakowski_qualitaetssicherung.pdf, zuletzt geprüft am 24.05.2018; Conway (2011): S. 301; Conway, Paul (2013): Preserving Imperfection: Assessing the incidence of Digital Imaging Error in HathiTrust. In: Preservation, Digital Technology & Culture 42(1), S. 20ff.

Unter den intellektuellen Prüfungen lassen sich folgende Aufgaben zusammenfassen:

- Digitalisatstichprobe 1: Aktendeckelprüfung (Kuckucksakte) unter der Fragestellung: *Wurden die Images des Archivals abgelegt, die laut Ordner- und Dateibenennung vorliegen sollten?*
- Digitalisatstichprobe 2: Eine Stichprobe zur Prüfung der Aufnahmequalität unabhängig vom Kontext (Ziehung der von Zufallssamples nach ISO-Norm 2859) unter der Fragestellung: *Ist die Herangehensweise bei der Aufnahme generell zufriedenstellend?*
- Eine Stichprobe für die Sichtprüfung am Original unter der Fragestellung: *Sind die Archivalien vollständig digitalisiert worden?*

Die Pre-Ingest-Software **Panoptes** - Das Datenmodell

Auf dem Kern der Software HT.dIVAS^{®25} wurde das Datenmodell an die Bedarfe der Qualitätssicherung für Digitalisierungsprojekte angepasst. Grundsätzlich lässt sich bei den verwendeten Entitäten zwischen Projektteknik, Archivteknik und Ablauforganisation unterscheiden. Die Projektteknik ermöglicht die Berücksichtigung und Unterscheidung von Einstellungen für die Automatische Prüfung (z.B. Einstellung der Validatoren). Die Archivteknik (Bestand, Archivalie, Version) ermöglicht die einheitliche Verarbeitung eines Bestands (z.B. Festlegung einheitlicher Stichprobengrößen je Bestand) und im Anschluss an die Digitalisierung und Qualitätssicherung einen einheitlichen Export der gültigen Versionen als Submission Information Packages (SIP).



Abbildung 2: Navigation durch das Hauptmenü

Das Archivalie ist Dreh- und Angelpunkt des Systems. Alle Prozessinformationen/-ereignisse und Prüfergebnisse werden an den Archivaliendatensatz gekoppelt. Da davon auszugehen ist, dass digitalisierte Repräsentationen Fehler aufweisen können, muss es möglich sein, dass mehrere Versionen nebeneinander existieren, miteinander verglichen und ungültige gegen gültige Versionen ausgetauscht (bzw. ungültige gelöscht) werden können.

Die Reiter für die Ablauforganisation (Nachbesserung, Übergabeprotokoll, Lieferschein, Festplatte und Stichprobe, Import) ermöglichen die Dokumentation der zwischen Auftraggeber, Digitalisierungsdienstleister und Qualitätssicherung ausgetauschten Informationen. Das Übergabeprotokoll enthält ausschließlich Informationen darüber welche Archivalien an den Digitalisierungsdienstleister übergeben wurden, der Lieferschein indes hat mehrere Funktionen. Einerseits bildet er ab, welche analogen und digitalen Archivalien der Digitalisierungsdienstleister an die Qualitätssicherung übergibt, andererseits ermöglicht er eine Verknüpfung mit der Importquelle (z.B. Festplatten-ID der gelieferten Festplatte). Nach dem Import der Daten ermöglicht der Lieferschein einen Abgleich der erwarteten mit den tatsächlich gelieferten Daten (Vollständigkeitsprüfung).

25 Vgl. H & T Greenline GmbH (o. J.): Revisionssichere, vertrauenswürdige Langzeitspeicherung mit HT.dIVAS[®]. Online verfügbar unter: <https://htgreenline.de/de-de/produkte/htdivas.aspx>, zuletzt geprüft am 24.05.2018.

Details Lieferschein

Lieferscheinnummer	24																																																	
Lieferdatum	11.05.2018																																																	
Kartonanzahl																																																		
Sichprobengröße(je Prozent)																																																		
Auftragnehmer																																																		
Lieferscheinnummer (Auftragnehmer)	20180511																																																	
Übergabeprotokolle																																																		
Lieferpositionen	<table border="1"><thead><tr><th>Signatur</th><th>Bestand</th><th>Charge</th><th>Analog Content</th><th>Digital Content</th><th>Master Digitalisate</th><th>Festplattenkennung</th></tr></thead><tbody><tr><td>DE2043_ZA001_M09101</td><td>DE2043_ZA001</td><td>1(LOS 1)</td><td></td><td></td><td>491</td><td>20180511</td></tr><tr><td>DE2043_ZA001_M09102</td><td>DE2043_ZA001</td><td>1(LOS 1)</td><td></td><td></td><td>357</td><td>20180511</td></tr><tr><td>DE2043_ZA001_M09103</td><td>DE2043_ZA001</td><td>1(LOS 1)</td><td></td><td></td><td>429</td><td>20180511</td></tr><tr><td>DE2043_ZA001_M09104</td><td>DE2043_ZA001</td><td>1(LOS 1)</td><td></td><td></td><td>647</td><td>20180511</td></tr><tr><td>DE2043_ZA001_M09105</td><td>DE2043_ZA001</td><td>1(LOS 1)</td><td></td><td></td><td>651</td><td>20180511</td></tr><tr><td>DE2043_ZA001_M09106</td><td>DE2043_ZA001</td><td>1(LOS 1)</td><td></td><td></td><td>518</td><td>20180511</td></tr></tbody></table>	Signatur	Bestand	Charge	Analog Content	Digital Content	Master Digitalisate	Festplattenkennung	DE2043_ZA001_M09101	DE2043_ZA001	1(LOS 1)			491	20180511	DE2043_ZA001_M09102	DE2043_ZA001	1(LOS 1)			357	20180511	DE2043_ZA001_M09103	DE2043_ZA001	1(LOS 1)			429	20180511	DE2043_ZA001_M09104	DE2043_ZA001	1(LOS 1)			647	20180511	DE2043_ZA001_M09105	DE2043_ZA001	1(LOS 1)			651	20180511	DE2043_ZA001_M09106	DE2043_ZA001	1(LOS 1)			518	20180511
Signatur	Bestand	Charge	Analog Content	Digital Content	Master Digitalisate	Festplattenkennung																																												
DE2043_ZA001_M09101	DE2043_ZA001	1(LOS 1)			491	20180511																																												
DE2043_ZA001_M09102	DE2043_ZA001	1(LOS 1)			357	20180511																																												
DE2043_ZA001_M09103	DE2043_ZA001	1(LOS 1)			429	20180511																																												
DE2043_ZA001_M09104	DE2043_ZA001	1(LOS 1)			647	20180511																																												
DE2043_ZA001_M09105	DE2043_ZA001	1(LOS 1)			651	20180511																																												
DE2043_ZA001_M09106	DE2043_ZA001	1(LOS 1)			518	20180511																																												

Abbildung 3: Ansicht Lieferschein

Über den Lieferschein können für die darin enthaltenen Lieferpositionen Stichproben gezogen werden. Diese werden bei der Entität Stichprobe in Listenform dokumentiert und dienen als Grundlage für die Aushebung der Stichproben für die Sichtprüfung am Original.

Workflow einer Datenlieferung in Panoptes

Die im Projekt zu digitalisierenden Archivalien werden über eine Importliste im System hinterlegt.²⁶ Werden vom Digitalisierungsdienstleister Daten (z.B. auf einer Festplatte) geliefert, werden diese ins System importiert. Die Digitalisate werden als Version (roter Rahmen in Abbildung 4) automatisch mit dem Datensatz des Archivale verknüpft.

Details Archivalie

Charge	1(LOS 1)														
Bestand	DE2043_ZA001														
Signatur	DE2043_ZA001_M09102														
Titel	Taufen														
Laufzeit	1899-1901														
Priorität															
Status	Importiert														
Sichtprüfung	Abgeschlossen														
Systemvalidationsergebnis	Archivalie Validation ungültig														
Nutzvalidationsergebnis	Archivalie Validation ungültig														
Versionen	<table border="1"><thead><tr><th>Festplattenkennung</th><th>Zeitpunkt Datenimport</th><th>Master Digitalisate</th><th>Systemvalidationsergebnis</th><th>Nutzvalidationsergebnis</th></tr></thead><tbody><tr><td>20180511</td><td>16.05.2018 10:05:22.863</td><td>357</td><td>Archivalienversion Validation ungültig</td><td>Archivalienversion Validation ungültig</td></tr></tbody></table>	Festplattenkennung	Zeitpunkt Datenimport	Master Digitalisate	Systemvalidationsergebnis	Nutzvalidationsergebnis	20180511	16.05.2018 10:05:22.863	357	Archivalienversion Validation ungültig	Archivalienversion Validation ungültig				
Festplattenkennung	Zeitpunkt Datenimport	Master Digitalisate	Systemvalidationsergebnis	Nutzvalidationsergebnis											
20180511	16.05.2018 10:05:22.863	357	Archivalienversion Validation ungültig	Archivalienversion Validation ungültig											
Auftragspositionen	<table border="1"><thead><tr><th>Nummer</th><th>Archivalienausgabe</th><th>Betrifft Nachlieferung</th><th>Priorität</th><th>Signatur</th><th>Bestand</th><th>Charge</th></tr></thead><tbody><tr><td>2c6263aa-971a-48f1-a895-648551748851</td><td></td><td></td><td></td><td>DE2043_ZA001_M0...</td><td>DE2043_ZA001</td><td>1(LOS 1)</td></tr></tbody></table>	Nummer	Archivalienausgabe	Betrifft Nachlieferung	Priorität	Signatur	Bestand	Charge	2c6263aa-971a-48f1-a895-648551748851				DE2043_ZA001_M0...	DE2043_ZA001	1(LOS 1)
Nummer	Archivalienausgabe	Betrifft Nachlieferung	Priorität	Signatur	Bestand	Charge									
2c6263aa-971a-48f1-a895-648551748851				DE2043_ZA001_M0...	DE2043_ZA001	1(LOS 1)									

Abbildung 4: Details eines Archivalies

Anhand bei Projektbeginn definierter Validatoreinstellungen werden sie automatisch qualitätsgesichert. Die Ergebnisse werden auf mehreren Ebenen (Archivalie, Version, Digitalisat) gespeichert (Systemvalidationsergebnis). Hinzu kommen die Ergebnisse der intellektuellen Prüfungen (Digitalisatstichproben und Archivalienstichprobe), welche als Nutzvalidationsergebnisse zusammengefasst werden.

²⁶ Im Moment entsteht eine Version von Panoptes, die die Informationen über zu prüfende Archivalien aus METS-Dateien extrahiert und darüber den Import steuern kann.

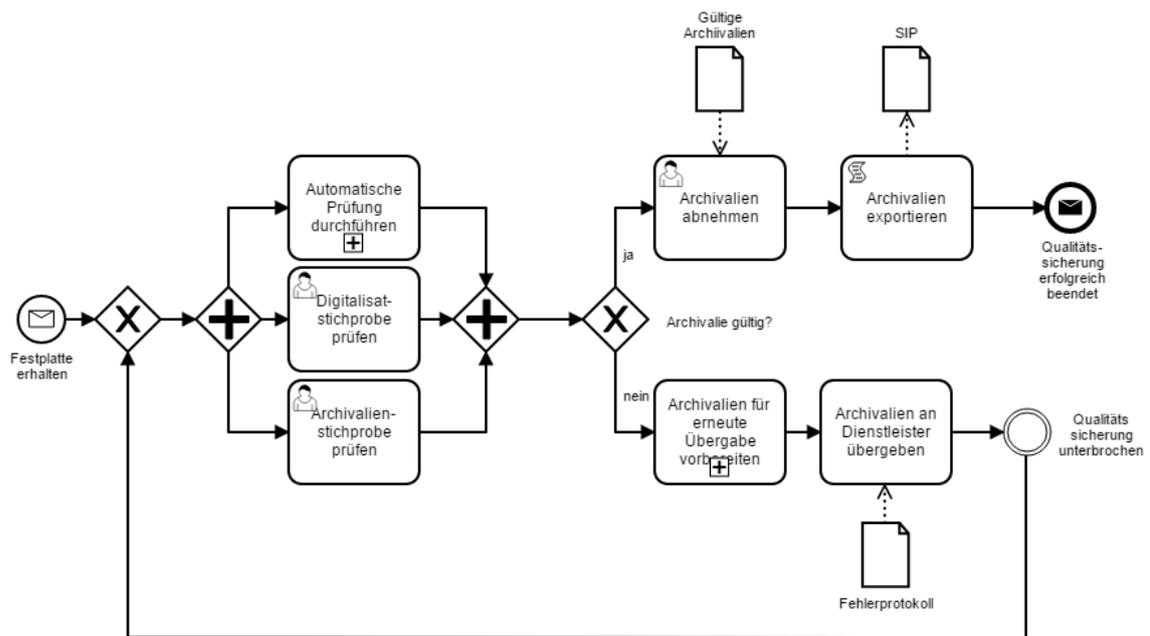


Abbildung 5: Qualitätssicherungsworkflow (einschließlich Export der SIP)

Nach Abschluss dieser Prüfungen entscheidet sich, welche Archivalien im System abgenommen werden können und welche Archivalien durch den Digitalisierungsdienstleister nachgebessert werden müssen. Eine Archivale kann nach einem Qualitätssicherungsprozess nur aus dem System exportiert werden, wenn die technische und intellektuelle Prüfung – sofern das Archivale für eine intellektuelle Prüfung ausgewählt wurde - erfolgreich war.

Die Prüfungsumgebung

In Panoptes ist ein Viewer integriert, der es ermöglicht die verschiedenen Prüfschritte in einer Umgebung zu absolvieren. Daher können die Aktendeckelstichprobe, die Digitalisatstichprobe und die Stichprobe für die Sichtprüfung am Original im selben Viewer absolviert werden.

Record Version Thumbnail Browser Page

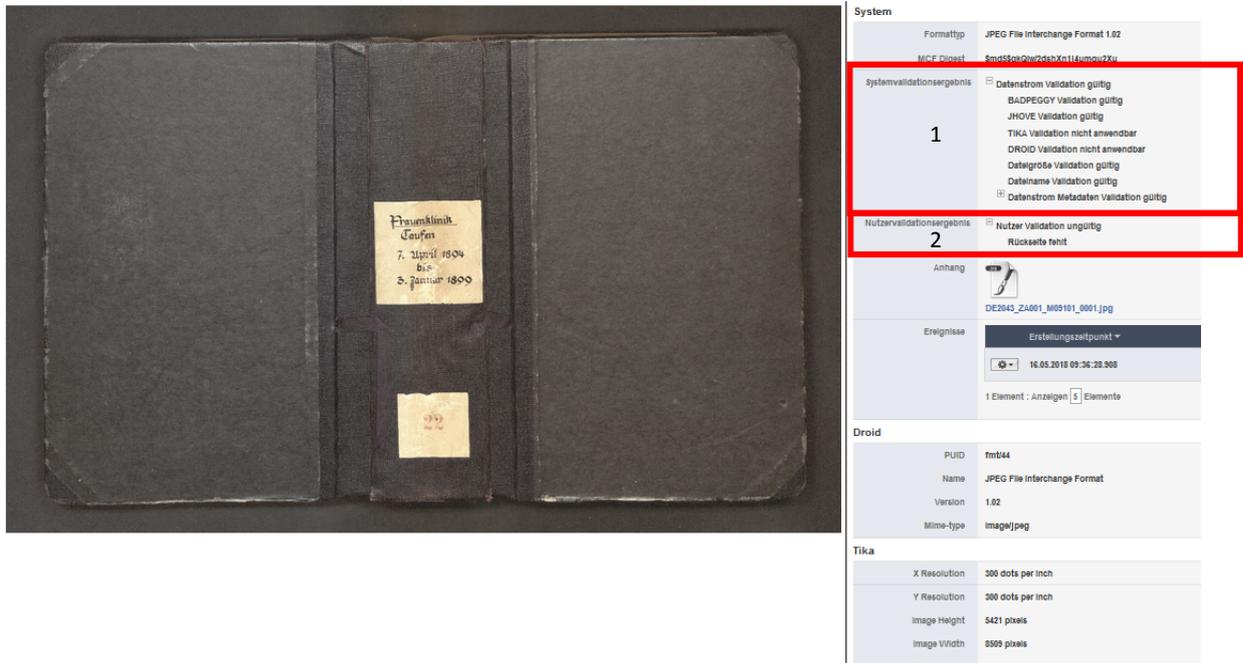
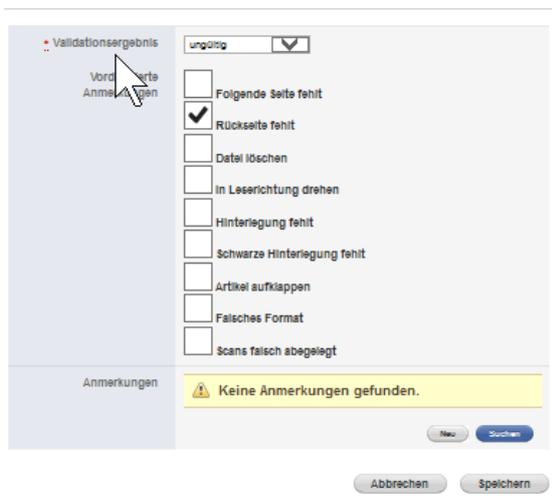


Abbildung 6: Viewer in Panoptes mit Übersicht über Metadaten und Validationsergebnisse

Die Digitalisate sind über die verschiedenen Stichproben erreichbar und navigierbar, durch sie kann aber auch unabhängig davon, ob sie in einer Stichprobe sind, navigiert werden. Auf der linken Seite des Viewers befindet sich das Digitalisat.²⁷ Auf der rechten Seite befinden sich die Metadaten zum Digitalisat, die System- [1] und Nutzervalidationsergebnisse [2] und falls nötig nennenswerte/besondere Ereignisse aus dem Importprozess.²⁸



Die gefundenen Fehler werden über eine standardisierte Checkboxliste eingetragen und direkt mit dem Digitalisat verknüpft. Die Auswahl der vordefinierten Fehlertypen wurde von der Reihenfolge her nach der Häufigkeit des Auftretens in aktuellen Projekten gewählt. Von der Norm abweichende Fehler können im Freitextfeld eingetragen werden. Da die möglichen Fehlertypen zu vielfältig sind, können aus administrativer Sicht nicht alle Fehler vordefiniert werden. Zu den nicht vordefinierten Fehlern gehören z.B. Reihenfolgefehler (Beispiel Ordne Scan x und Scan y hinter Scan z ein).

Abbildung 7: Standardisierte Fehlerliste der Prüfungsumgebung

²⁷ Die Default-Einstellung ist die Anzeige der Masterdatei in verkleinerter Form (Thumbnail-Ansicht).

²⁸ Hier werden bspw. nicht auslesbare oder fehlerhaft ausgelesene Prüfsummen aufgeführt. Diese sind aber zusätzlich über ein zusammenfassendes Prüfprotokoll (Excel-Datei) ausgebenbar.

Fehlerbeschreibungen, die mit den Digitalisaten verknüpft sind, können nach Abschluss der Prüfung zusammengefasst werden. Die Zusammenfassung der Fehler erfolgt unterhalb der jeweiligen Archivalienversion.



Abbildung 8: Fehlerzusammenfassung für eine Archivalienversion

Darüber hinaus können sie in einem Fehlerprotokoll (z. B. Excel-Tabelle) ausgegeben werden. Das Fehlerprotokoll ermöglicht das Ausgeben der Systemvalidationsergebnisse, der Nutzervalidationsergebnisse und der Ereignisse (z. B. fehlgeschlagene Importe aufgrund von Dateibenennungsfehlern oder Lieferung in einem falschen Dateiformat).

Fazit und Ausblick

In Panoptes wird ein (Pre-)Ingestprozess abgebildet. Damit wird die Qualitätssicherung für Digitalisierungsprojekte formalisiert und automatisiert. Die Strukturierung der automatischen Prüfung, die Festlegung konkreter Regeln für die Datenimporte und die Vordefinition von Fehlertypen für die Sichtprüfung ermöglichen eine einheitliche und effiziente Bearbeitung durch mehrere Mitarbeiter und erleichtern die Kommunikation mit dem Digitalisierungsdienstleister. Durch die Stichprobenmodule wird eine objektive Ziehung von Zufallsstichproben ermöglicht.

In Zukunft können die abgebildeten Prüfprozesse um weitere Prüfungen erweitert werden. In den Prüfprozess könnten Erkennungsverfahren für die Aufnahmequalität eingegliedert werden.²⁹ Ein mögliches Szenario bietet auch die Weiterverarbeitung bzw. manuelle Umstrukturierung von Transferpaketen. Diese könnten vereinzelt werden und als granularere SIP an den Auftraggeber übergeben werden (z.B. bei Rollfilmdigitalisierung oder Strukturierung von Akten und Vorgängen).

Letztlich könnte auch die Generierung von Nutzungsderivaten (z.B. im PDF- oder JPEG-Format) oder OCR-Volltexten vom Produktionsprozess beim Digitalisierungsdienstleister hinter den Abschluss der Qualitätssicherung verlagert werden. Der Konvertierungsprozess würde hierdurch rationalisiert und gleichzeitig Fehlerquellen vermieden, da erst für gültige Archivalienversionen weitere Repräsentationen entstünden.

29 z. B. <https://www.image-engineering.de/products/software/376-iq-analyzer>, zuletzt geprüft am 24.05.2018.

Literaturverzeichnis

Beck, Emily (2015): Pilotprojekt "Digitalisierung von archivalischen Quellen" - Gefördert durch die Deutsche Forschungsgemeinschaft (DFG). Online verfügbar unter:

<https://www.archivschule.de/DE/forschung/forschungsprojekte/digitalisierung/dfg-projekt-archivgutdigitalisierung.html>, zuletzt geändert am 13.08.2015, zuletzt geprüft am 24.05.2018.

Conway, Paul (2011): Archival quality and long-term preservation: a research framework for validating the usefulness of digital surrogates. In: *Archival Science* 11 (3-4), S. 293-309.

Conway, Paul (2013): Preserving Imperfection: Assessing the incidence of Digital Imaging Error in HathiTrust. In: *Preservation, Digital Technology & Culture* 42(1), S. 17-30.

Conway, Paul (2015): Digital Transformations and the archival nature of surrogates. In: *Archival Science* 15 (1), S. 51-69.

Fresa, Antonella et al. (2015): Digital curation and quality standards for memory institutions: PREFORMA research project. In: *Archival Science* 15 (2), S. 191-216.

Generaldirektion der Staatlichen Archive Bayerns (o. J.): Abschlussbericht des Produktivpiloten Digitalisierung von archivalischen Quellen. Online verfügbar unter:

https://www.archivschule.de/uploads/Forschung/Digitalisierung/Ergebnisse/GDA_Abschlussbericht_DFG-Archivgutdigitalisierung.PDF, zuletzt geprüft am 24.05.2018.

Glauert, Mario (2013): Dimensionen der Digitalisierung. Kosten, Kapazitäten und Konsequenzen. In: Claudia Kauertz (Hrsg.): *Digital und analog. Die beiden Archivwelten : Beiträge / 46*. Rheinischer Archivtag, 21. - 22. Juni 2012 in Ratingen, Bonn: Habelt, S. 48-59.

Grau, Bernhard (2016): Digitalisierung von Archivgut – Was und warum. In: *Archive in Bayern* 9, S. 127-151.

Holzapfl, Julian; Makowski, Stephan (2014): Qualitätssicherung in archivischen Digitalisierungsprojekten. Erfahrungen und Diskussionsanregungen auf dem Weg zur Professionalisierung. Vortrag auf der Konferenz *Made digital - Wege und Ziele* am 26.11.2014 in Koblenz. Online verfügbar unter:

https://www.bundesarchiv.de/imperia/md/content/bundesarchiv_de/fachinformation/informations-technologie/digitalisiertesarchivgut/holzapflmakowski_qualitaetssicherung.pdf, zuletzt geprüft am 24.05.2018.

Jobs, Christoph (2018): Ein Pre-Ingest-Tool für die Qualitätssicherung in Digitalisierungsprojekten? Ein Praxisbericht, Vortragsfolien. Online verfügbar unter:

https://www.staatsarchiv.sg.ch/home/auds/22/_jcr_content/Par/downloadlist_1638086843/DownloadListPar/download_1552214939.ocFile/03_Ingest_Jobs.pdf, zuletzt geprüft am 24.05.2018.

Keitel, Christian; Schoger, Astrid (2013): Vertrauenswürdige digitale Langzeitarchivierung nach DIN 31644. Berlin: Beuth (Beuth-Kommentar).

Kistenich, Johannes (2013): Werkzeuge für Digitalisierungsprojekte im Landesarchiv Nordrhein Westfalen, Standardworkflow, technische Vorgaben, Qualitätssicherung, Projektplanung, In: *Archivar* 66 (2013), Heft 2, S. 228-244.

Konferenz der Leiterinnen und Leiter der Archivverwaltungen des Bundes und der Länder (KLA) (2017a): *Wirtschaftliche Digitalisierung in Archiven – Empfehlungen der Konferenz der Leiterinnen*

und Leiter der Archivverwaltungen des Bundes und der Länder (KLA). Online verfügbar unter: https://www.bundesarchiv.de/imperia/md/content/bundesarchiv_de/fachinformation/ark/2017-01-04_kla_fta_wirtschaftliche_digitalisierung_final_.pdf, zuletzt geprüft am 24.05.2018.

Konferenz der Leiterinnen und Leiter der Archivverwaltungen des Bundes und der Länder (KLA) (2017b): Handreichung zur Digitalisierung von Archivgut - Beschluss der Bundeskonferenz Kommunalarchive in Wolfsburg am 26.09.2017. Online verfügbar unter: http://www.bundeskonferenz-kommunalarchive.de/empfehlungen/Handreichung_Digitalisierung_von_Archivgut_endfassung.pdf, zuletzt geprüft am 24.05.2018.

Markgraf, Daniel (2018): Qualität. In: Gabler Wirtschaftslexikon. Online verfügbar unter: <https://wirtschaftslexikon.gabler.de/definition/qualitaet-45908/version-269195>, zuletzt geändert am: 16.02.2018, zuletzt geprüft am 24.05.2018.

Ohne Verfasser (o. J.): DPF Manager User Manual. Online verfügbar unter: <http://dpfmanager.org/Downloads/User%20Manual.pdf>, zuletzt geprüft am 24.05.2018.

Röschner, Matthias (2016): Alles unter Kontrolle? Möglichkeiten und Grenzen des Qualitätsmanagements bei der Zusammenarbeit mit externen Scandienstleistern. In: Archive in Bayern 9, S. 213-220.

Schleiter, Klaus (2018): Digitalisierung archivalischer Quellen – Ausschreibung Digitalisierung archivalischer Quellen der DFG. Online verfügbar unter: <https://www.archivschule.de/DE/forschung/digitalisierung-archivalischer-quellen/>, zuletzt geändert am 19.01.2018, zuletzt geprüft am 24.05.2018.

Stadtarchiv Mannheim – Institut für Stadtgeschichte (o. J.): Abschlussbericht des Produktivpiloten Digitalisierung von archivalischen Quellen. Online verfügbar unter: https://www.archivschule.de/uploads/Forschung/Digitalisierung/Ergebnisse/ISG_StadtA_MA_Abschlussbericht_DFG_Archivgutdigitalisierung.pdf, zuletzt geprüft am 24.05.2018.

Stockert, Harald (2015): Musterworkflow für die Planung und Durchführung einer Inhouse Digitalisierung und einer Digitalisierung durch einen externen Dienstleister. Verfügbar unter: https://www.archivschule.de/uploads/Forschung/Digitalisierung/Handreichungen/Idealtypischer_Digitalisierungsworkflow_fuer_Inhouse- und_externer_Digitalisierung.pdf, zuletzt geprüft am 24.05.2018.

Stumpf, Marcus (2014): Digitalisierungsstrategien in Deutschland – Versuch einer Bestandsaufnahme. In: Archivpflege in Westfalen-Lippe (80), S. 27-35.