

Die Ablösung des Lesesaals: Digitalisierung im Schweizer Bundesarchiv mit Hilfe von Kitodo.Production 3.0

Matthias Finck, Stefan Kwasnitza

Einleitung

Das Schweizerische Bundesarchiv (BAR) sichert die Dokumentation staatlichen Handelns und macht sie zugänglich. Das BAR verfolgt dabei die Vision eines durchgängig digitalen Archivs.¹ Unabhängig davon ob Dokumente analog oder digital im Archiv vorliegen: sie lassen sich online konsultieren. Dieser neue Online-Zugang zum BAR wird in zwei Projekten umgesetzt: einem für das kundenorientierte Webportal und einem für die Digitalisierungsinfrastruktur.

Für die Umsetzung des Projekts für den Aufbau der Digitalisierungsinfrastruktur setzt das BAR auf die Open Source Software Kitodo.Production. Kitodo.Production ist das Workflowmanagementmodul der Kitodo-Suite². Es unterstützt den Digitalisierungsprozess von verschiedenen Materialarten wie Akten, Drucken, Periodika, Handschriften, Noten und Musikalien, Einblattmedien und Dokumentennachlässen und kann Workflows hoch individualisiert abbilden.

Im Rahmen eines dreijährigen DFG-Projekts ist die Software Kitodo.Production bis Ende 2019 grundlegend überarbeitet worden³. Das BAR hat diesen Relaunchprozess aktiv begleitet und die Anforderungen an die Software in den Prozess eingebracht. Gleichzeitig wurden umfassend Entwicklungsmittel zur Verfügung gestellt, die auch den Relaunchprozess unterstützen haben.

Mit dieser entwickelten Infrastruktur auf der Basis von Kitodo.Production 3.0 digitalisiert das BAR analoge Dossiers. Mit der Aufschaltung des Webportals kann das BAR einen Teil der Bestellungen digital ausliefern. Der Lesesaal steht weiterhin zur Verfügung. In den folgenden Jahren wird die Kapazität der Digitalisierungsinfrastruktur erhöht, während gleichzeitig

Mit einem neuen Online-Zugang lassen sich in Zukunft alle Angebote des Schweizerischen Bundesarchivs ort- und zeitunabhängig nutzen. Sowohl analog als auch digital im Archiv vorliegende Dossiers können nun online konsultiert werden. Ein Besuch im Lesesaal wird so mittelfristig nicht mehr nötig sein. Damit nähert sich das Bundesarchiv der Vision eines durchgängig digitalen Archivs. Voraussetzung dafür ist die mittels Kitodo.Production 3.0 realisierte archivspezifische Weiterentwicklung der Digitalisierungsinfrastruktur.

das Angebot vor Ort sukzessive reduziert wird. Eine ausführliche Beschreibung dieses Vorhabens, die gleichzeitig Grundlage dieses Beitrags ist, befindet sich im Abschlussband des DFG-Projekts zur Entwicklung von Kitodo.Production 3.0.⁴

Den medienbruchfreien Informationskreislauf als Ziel

Das BAR übernimmt und lagert seit 2009 digitale Geschäftsunterlagen und relationale Datenbanken der Schweizer Behörden. Mit dem Transfer von digitalen Geschäftsverwaltungsunterlagen (GEVER) ins Langzeitarchiv und immer kürzeren Zeitabständen zwischen elektronischer Entstehung und Archivierung wird es zukünftig wichtiger, auch der Verwaltung einen schnelleren Zugriff auf ihre eigenen abgelieferten Unterlagen zu ermöglichen.

Der Zugang zum Archivgut der Eidgenossenschaft geschieht deshalb seit Herbst 2019 über ein neues Portal, das es den berechtigten Nutzenden ermöglicht, orts- und zeitunabhängig Auskunft zu erhalten,

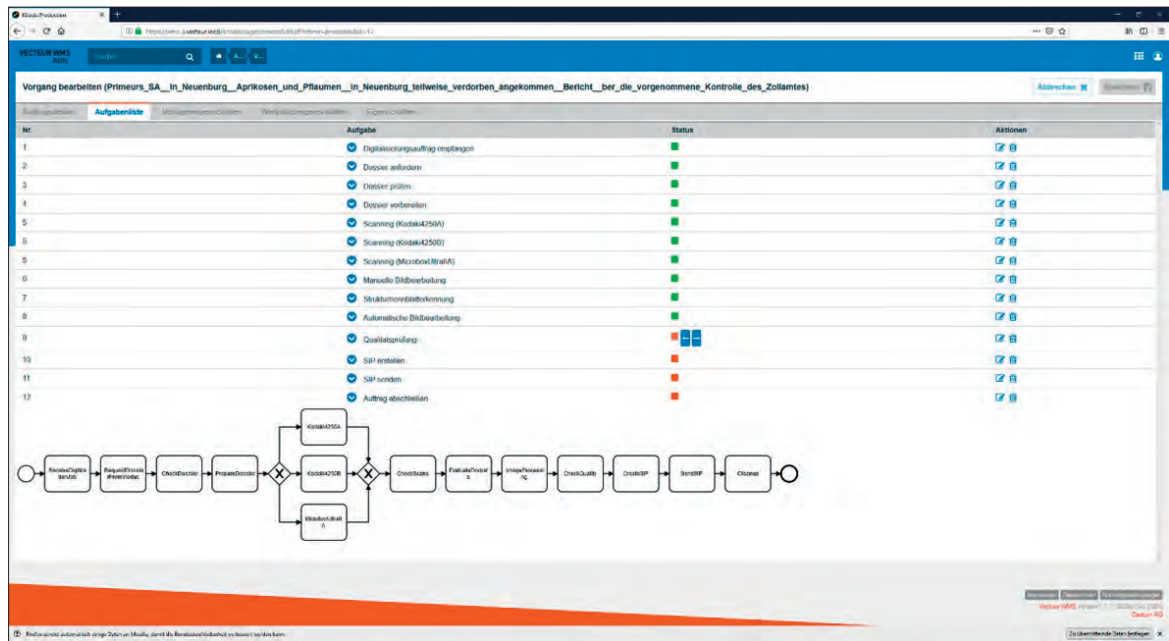
1 Kwasnitza, S. (2020) : Auf dem Weg zum komplett digitalen Archiv. Online-Zugang und Digitalisierung «on-demand» im Schweizerischen Bundesarchiv. In: Politik und Kultur, 3/2020, S. 19.

2 vgl. <https://www.kitodo.org>

3 Finck, M. (2019): Erfolgsfaktoren für konsortiale drittmittelfinanzierte Open Source Projekte – am Beispiel von Kitodo 3.0. In: B.I.T. online, 2/22, b.i.t.verlag, S. 144 – 148.

4 Kwasnitza (2019): DIE ABLÖSUNG DES LESESAALS: DIGITALISIERUNG VON ARCHIVGUT ON-DEMAND – Ein Werkstattbericht aus dem Schweizerischen Bundesarchiv. In Finck, M.; Hermann, E.: NORDBLICK – Forschung für die Wirtschaft – Sonderedition: Abschlussbericht Kooperative Weiterentwicklung der quelloffenen Digitalisierungssoftware Kitodo.Production, 9/2019, Elmshorn.

Abb. 1: Screenshot der Statusübersicht der Aufgaben eines Digitalisierungsauftrags mit Prozessablaufkette in Kitodo. Production 3.0



zu recherchieren, Dokumente einzusehen, Digitalisierungsaufträge zu erteilen und Einsichtsgesuche einzureichen.

Die nun entwickelte Lösung bietet eine Suche in Primär- und Metadaten, die Online-Identifizierung der Kunden im Bestellprozess und die anschließende rechtebasierte Auslieferung der Dokumente per Mausklick aus dem digitalen Archiv. Die Nutzenden können im neuen Webportal je nach Berechtigung unterschiedliche Meta- und Primärdaten suchen. Metadaten beschreiben dabei Bestände, Dossiers oder Dokumente. Primärdaten sind die eigentlichen Unterlagen – unabhängig davon, ob digital entstanden oder digitalisiert. Die Primärdaten sind im digitalen Archiv gespeichert, und mittels OCR wird ein Volltext erzeugt und in den Suchindex geladen.

Von diesen Neuerungen profitiert nicht nur jüngeres, bereits digital produziertes Archivgut. Das BAR digitalisiert Papier-Unterlagen auf Bestellung on-demand und liefert diese ebenfalls über den Online-Zugang aus. In diesem Fall wird im System innerhalb des Bestellvorgangs ein Digitalisierungsauftrag ausgelöst. Bestellen Kundinnen und Kunden ein Papier-Dossier, erhalten sie es rund zwei Wochen später digitalisiert im Online-Zugang.

Die Digitalisierung „on-demand“ ist dabei eine herausfordernde Neuerung für das Archiv. Dabei können Kunden im Bestellvorgang wählen, ob Sie ein analoges Dossier digital oder im Lesesaal einsehen möchten. Wählen die Kunden „digital“, geht ein Auftrag an die interne Digitalisierungsinfrastruktur: das

Digitalisierungsteam des BAR scannt die Unterlagen und speichert sie im digitalen Archiv. Damit stehen die Unterlagen ab diesem Moment allen interessierten Personen langfristig digital zur Verfügung. Eine Herausforderung ist bei der Digitalisierung nicht nur die Menge der bestellten Dossiers, sondern die damit verbundene – für das Archiv neue – Arbeitsweise: Die quasi-industrielle Verarbeitung von Aufträgen. Das Papier-Dossier wird in verschiedenen Schritten auf der internen Digitalisierungsinfrastruktur gescannt und via Ingest-Prozess im digitalen Archiv abgelegt. Die Erzeugung des digitalen Objekts erfolgt neu von der Aushebung des Papier-Dossiers im Magazin bis zum Ingest mit Kitodo.Production⁵.

Das Workflowmanagement-System (WMS), das mit Hilfe von Kitodo.Production implementiert wurde, löst die logistische Lieferung des Papier-Dossiers in die Scanningzone aus, in der zuerst geprüft wird, ob es sich tatsächlich um das bestellte Archivgut handelt. Anschließend wird das Dossier für das Scannen vorbereitet und digitalisiert. Nach der Qualitätsprüfung des Digitalisats wird der erfolgreiche Abschluss per Statusänderung an das Bestellsystem zurückgemeldet und ein validiertes Submission Information Package (SIP) im Digitalen Langzeitarchiv (DIR) gespeichert. Abbildung 1 zeigt exemplarisch den Statusbericht der Aufgaben eines Digitalisierungsauftrags sowie die in Kitodo.Production implementierte Prozesskette.

An diesem Beispiel kann man erkennen, dass die in Kitodo.Production 3.0 neu implementierte Workflow-

⁵ ggl. Meyer, S.; Wendt, K.; Finck, M. (2019). Insight Kitodo 3.0: Die rundum erneuerte Digitalisierungssoftware einfach mal ausprobieren. Vortrag auf dem 7. Bibliothekskongress in Leipzig 2019. <https://opus4.kobv.de/opus4-bib-info/frontdoor/index/index/docId/16497>

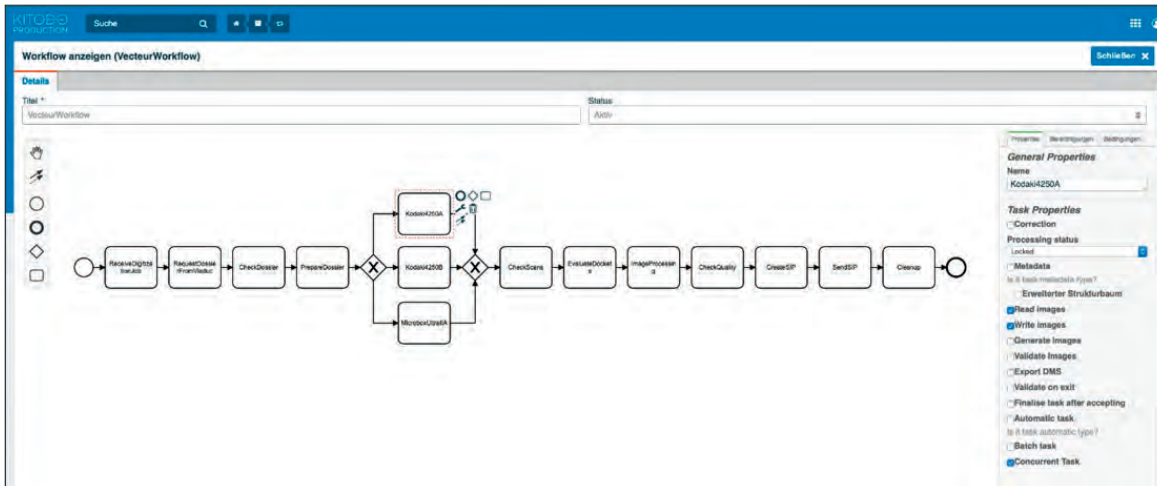


Abb. 2: Screenshot des neuen Workflow-Editors von Kitodo. Production 3.0

Auftragsstapel	Auftragsstapel	Status	Projekt	Laufzeit	Auftragsid	Signatur	Dossier-Titel
<input type="checkbox"/>	Zwetschgen und Pfäuerlein	■	Veclara	27.6.19, 71m	8121	F:9130841969277#1337*	Zwetschgen und Pfäuerlein
<input type="checkbox"/>	Pflauche Pfäuerlein / Zwetschgen	■	Veclara	27.6.19, 96m	8127	F:711041967/04363P*	Pflauche Pfäuerlein, Zwetschgen
<input type="checkbox"/>	Freipassgesuch f. Pfäuerlein, god, inf, ie	■	Veclara	27.6.19, 76m	8126	E:6351C-01#100010414139*	Freipassgesuch für Pfäuerlein gedörft aus Jugoslavien, zum Erhitzen und nachher Export nach nordschweden Ländern
<input type="checkbox"/>	Zollersatz f. getrocknete nicht ausgekost.	■	Veclara	27.6.19, 5m	8129	E:6351F#10001044#899*	Zollersatz für getrocknete nicht ausgekostete Zwetschgen und Pfäuerlein in Packungen von weniger als 50 kg Gewicht
<input type="checkbox"/>	Primeus_SA_in Neuenburg_Aprikosen	■	Veclara	27.6.19, 3m	8130	E:6351C-01#100010414189*	Primeus_SA_in Neuenburg_Aprikosen und Pfäuerlein in Neuenburg belausen verdorben angekommen, Bericht über die vorgenommene Kontrolle des Zolstatus
<input type="checkbox"/>	Aprikosen nach Muster: Tafelverwendung	■	Veclara	27.6.19, 1m	8131	E:6351A#10001029#1321*	Aprikosen nach Muster: Tafelverwendung
<input type="checkbox"/>	Entgemachte Aprikosen_Abfertigung mit	■	Veclara	27.6.19, 59m	8132	E:6351A#10001038#999*	Entgemachte Aprikosen, Abfertigung mit Jahresgutachten
<input type="checkbox"/>	Konservierte Aprikosen	■	Veclara	27.6.19, 51m	8133	E:6351A#10001036#305*	Konservierte Aprikosen
<input type="checkbox"/>	Aprikosen_Allgemeines	■	Veclara	27.6.19, 55m	8134	E:7220A#197570#339*	Aprikosen, Allgemeines
<input type="checkbox"/>	Vfg_Nr_538_in Indische Aprikosen	■	Veclara	27.6.19, 53m	8135	E:7397#1993/362#1148*	Vfg. Nr. 538 Indische Aprikosen

Abb. 3: Liste der laufenden Aufträge in Kitodo

Engine nicht nur komplexe Workflows mit vielen Prozessschritten abbilden kann – hier sind es 12 Workflowschritte vom Bestellvorgang bis zur Ablieferung des Digitalisats – sondern auch parallele Workflowschritte erlaubt. Diese Workflows lassen sich mit Hilfe der neuen Workflow-Engine per Drag’n’Drop gestalten (vgl. Abb. 2).

Archivspezifische Digitalisierung

Durch das hohe Maß an Flexibilität und Anpassbarkeit der neuen Kitodo Version lassen sich nun auch archivspezifische Anforderungen besser abbilden. Im BAR werden freigegebene Aufträge zur Digitalisierung von Dossiers für die Mitarbeitenden im Webportal aufgelistet und nach einer vordefinierten Gewichtung automatisch priorisiert. Die Digitalisierungsinfrastruktur ruft – solange noch Kapazität vorhanden ist – in regelmäßigen Abständen den nächsten Auftrag ab (vgl. Abb. 3 – Auftragsliste). Das Webportal liefert jeweils den am höchsten priorisierten Auftrag inklusive des benötigten Sets an Metadaten. Der Auftrag erscheint nun im WMS und der Status im Webportal wechselt auf „Für Aushebung bereit“.

Nach erfolgter Aushebung des Papier-Dossiers im Magazin prüfen die Mitarbeitenden das bereitgestellte Dossier. Die aus dem Bestellsystem automatisch in Kitodo.Production übermittelten Angaben sind dabei mit der logischen Seitenstruktur des physischen Dossiers sowie den erfassten Metadaten zu vergleichen. Letztlich gilt es, allfällig bestehende Schutzfristen und die davon abhängende Zugänglichkeit zu kontrollieren. Wurde zwar das korrekte Dossier geliefert, aber die verzeichneten Einheiten sind z. B. nicht in der richtigen Reihenfolge erfasst, kann in der logischen Seitenstruktur die Reihenfolge der einzelnen Dokumente im WMS angepasst werden (vgl. Abb. 4).

Nach erfolgreicher Kontrolle wird das Dossier für den Scanningprozess vorbereitet und sein Inhalt in Teilaufträge unterteilt. Dazu wurde in Kitodo.Production im Metadateneditor die Ansicht der Strukturen um die Darstellung eines rein physischen und eines rein logischen Strukturbaums ergänzt (vgl. Abb. 4). Die Substrukturtrennung sowie die Dokumententrennung werden mithilfe entsprechender Strukturtrennblätter vorgenommen, die mit Hilfe der Laufzettelfunktionalität von Kitodo umgesetzt wurden. Die Unterteilung des Dossiers in Teilaufträge

Abb. 4: Meta-
dateneditor in
Kitodo.Production:
Darstellung
der Strukturen,
Auswahl eines
Teilauftrages und
dazugehöriger
Scannertyp

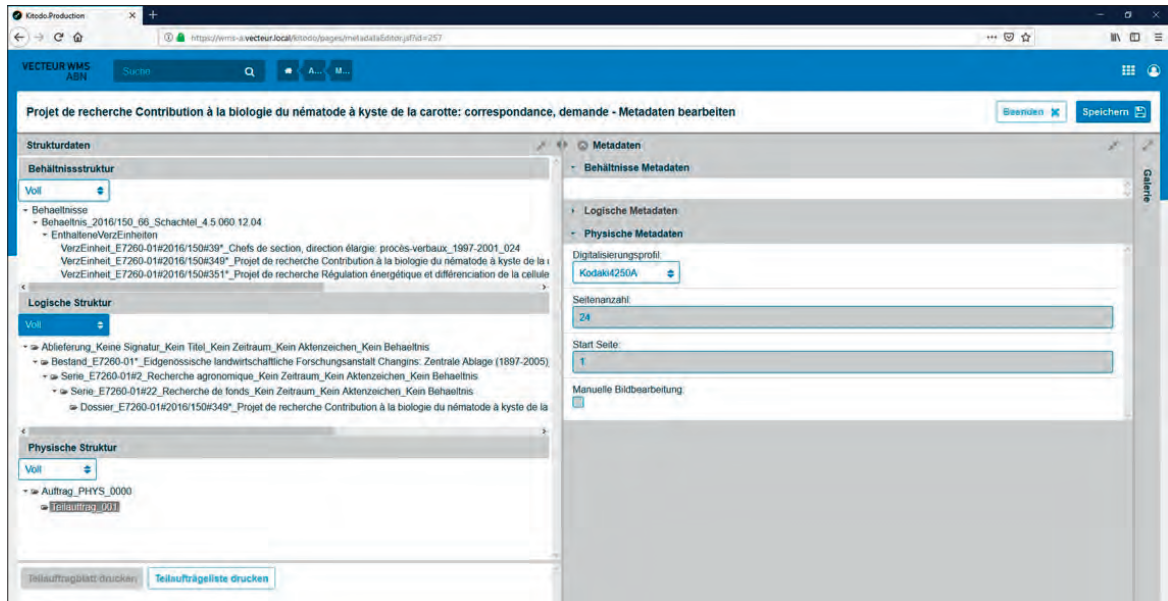
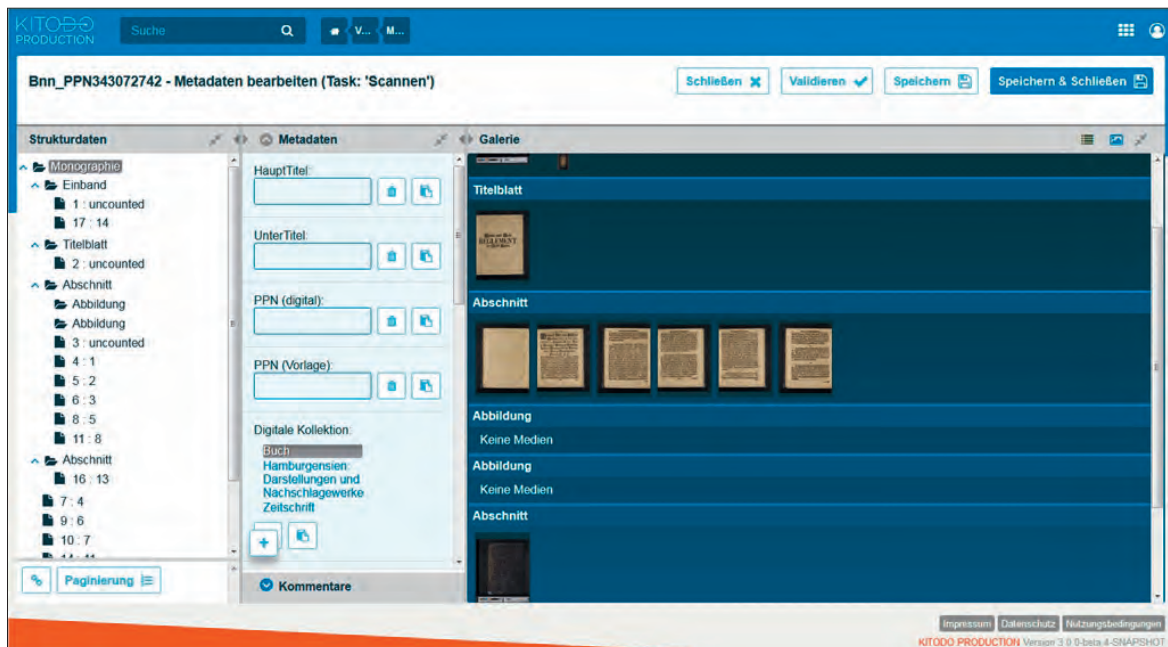


Abb. 5:
Alternative
Konfiguration
des Meta-
dateneditors mit
Galeriefunktion



ermöglicht die anschließende Verteilung eines Scan-Auftrags auf mehrere Scan-Stationen und damit eine bessere Auslastung der Infrastruktur, bestehend aus Einzugs- und Buchscannern. Die auf verschiedenen Digitalisierungsstraßen abgearbeiteten Teilaufträge werden später vom System erkannt und automatisch am Prozessende in ihrer korrekten Abfolge zusammengeführt.

Die vorgängig eingelegten Strukturtrennblätter erzeugen im Digitalisierungsvorgang eine Nachbildung der inneren Struktur eines Dossiers in Form von Strukturelementen und generieren damit die für den Kunden im Archivgutkontext wichtige Orientierung im ausgelieferten Endprodukt.

Diese Umsetzung des Digitalisierungsworkflows am BAR zeigt, wie unterschiedlich intensiv Funktionen

von Kitodo genutzt werden und wie wichtig eine hohe Anpassbarkeit ist. Während das BAR den Metadateneditor hauptsächlich nutzt, um die Dossierstrukturen und die sich daraus ableitenden Aufträge abzubilden und zu pflegen (vgl. Abb. 4), spielt die Qualitätssicherung der Digitalisate und die Erfassung der Metadaten bei wissenschaftlichen Bibliotheken oftmals eine viel zentralere Rolle. Diesen unterschiedlichen Anforderungen kann der neue Metadateneditor gerecht werden, indem er z. B. Galeriefunktionen aktiviert und die Metadatenpflege ins Zentrum stellt (vgl. Abb. 5).

Der Ingest- und Auslieferungsprozess

Am Ende der durch Kitodo.Production gesteuerten Ablaufkette prüft die Qualitätssicherung die Vollständigkeit der Dossier-Elemente und die Qualität der

Scans: Schärfe, Ausschnitt, die Dokumenten- und Substrukturtrennung, die erstellte Baumstruktur sowie die Zuordnung der Scans zu den bereits im Archivinformationssystem verzeichneten Einheiten.

Anschließend wird das Datenpaket mit den PREMIS-Metadaten des Digitalisierungsprozesses ergänzt. Nach erfolgreicher Validierung gemäß den ISO-Standards für die Langzeitarchivierung wird das Datenpaket als SIP in die angeschlossene Speicherlösung übertragen und danach im Kitodo-WMS gelöscht. Die statistische Auswertung der Auftragsdaten bleibt weiterhin möglich. Auch diese Schritte wurden mit in Kitodo integrierten automatisch ablaufenden Skripten realisiert. Die Lieferung an den Kunden erfolgt gleich wie bei den rein digital entstandenen Dokumenten über den Download im Webportal. Beim Abruf der Unterlagen aus dem angeschlossenen digitalen Langzeitarchiv werden dabei technische Transformationsschritte vorgenommen, wie z. B. die Umwandlung von JPG2000 Dateien in PDFs mit hinterlegtem Volltext. Die Digitalisate sind danach als verschlüsselte ZIP-Datei verfügbar, die sich mit separatem Passwort entpacken und im Offline-Viewer auf dem heimischen Computer anzeigen lassen.

Dank der OCR-Erkennung können Nutzende erstmals auch im Volltext der Primärdaten recherchieren. In der Schweiz arbeitet die Bundesverwaltung mehrsprachig. Deshalb werden bei der Eingabe eines Suchbegriffs im Suchschlitz automatisch Synonyme und Übersetzungsvorschläge in den vier Landessprachen generiert. Die vorgeschlagenen Begriffe speisen sich aus der implementierten Terminologie-Datenbank 'Thermdat' der Bundesverwaltung und ergänzen bei Bedarf mehrsprachig die eigene Suche.

Im Online-Zugang zum BAR können Nutzende seit 2019 anonym, als registrierte oder als identifizierte Person recherchieren. Regelsets steuern den Zugang: Je nach Status sind unterschiedliche Funktionen verfügbar. Wer über eine Bewilligung für den Zugriff auf geschützte Unterlagen verfügt, benötigt zu deren Konsultation den Status als identifizierter Nutzer. Hat eine Person keine Berechtigung für einen Zugriff auf geschützte Akten, wird sie automatisch in den Einsichtsgesuchsprozess umgeleitet. Die Suche in öffentlich zugänglichen Metadaten steht dabei sämtlichen Nutzenden offen. Das Webportal und die angeschlossene Kitodo-Digitalisierungsinfrastruktur ermöglichen es Interessierten somit, Archivgut auf der Plattform des Archivs zu recherchieren und, sofern die Rechtslage es erlaubt, diese auszuwerten und weiterzuverarbeiten.

Fazit

Im Gegensatz zu klassischen Archivinformationssystemen wird das Zugangsmodul neu zum Mastersystem für die Verwaltung von Suchindizes, Bestelldaten, Digitalisierungsaufträgen und Nutzeraccounts. Im selben System bewirtschaften die Mitarbeitenden die Prüfung und Freigabe der Bestellungen, steuern den Prozess der Einsichtsgesuche und verwalten die Zugriffsberechtigungen der Nutzenden – ein Ansatz, der zukünftig einen flexibleren Aufbau archivspezifischer IT-Systemlandschaften ermöglichen wird.

Die Vorteile von Kitodo liegen für das BAR im rasch anpassbaren Quellcode, der Interoperabilität mit anderen Systemen durch offene Schnittstellen und der Portierung internationaler Standards. Das BAR adaptiert die bestehende Software in Abstimmung mit der Anwendergemeinschaft erstmals für den gesamten Workflow im archivspezifischen Kontext: eine Entwicklungsleistung, die über den offenen Quellcode auch für andere nutzbar bleibt.

Noch verfügt das BAR nicht über die Kapazität, um alle bestellten analogen Dossiers direkt zu digitalisieren. Bis 2021 soll jedoch die Digitalisierungsinfrastruktur soweit ausgebaut sein, dass ein Besuch im Lesesaal nicht mehr nötig ist. Damit wird ein großer Schritt in Richtung eines durchgängig digitalen Archivs realisiert. **I**



Prof. Dr. Matthias Finck

lehrt an der NORDAKADEMIE Usability Engineering und ist Inhaber der effective WEBWORK GmbH.

NORDAKADEMIE Elmshorn
Köllner Chaussee 11
25337 Elmshorn
matthias.finck@nordakademie.de

effective WEBWORK GmbH
Neuer Wall 18
20354 Hamburg
finck@effective-webwork.de



Stefan Kwasnitza

stellvertretender Direktor
Schweizerisches Bundesarchiv BAR
Archivstrasse 24
3003 Bern
Schweiz
stefan.kwasnitza@bar.admin.ch