

Harald Fischer, Claudia Schorcht

# Die Digitalisierung der Exilmonografien in der Deutschen Nationalbibliothek

## Die Perspektive des Dienstleisters

### Die Herausforderung

24.000 Bände,  
4 Millionen  
Seiten,  
8 Monate

24.000 Bände, 4 Millionen Seiten, 8 Monate. Das waren die drei Zahlen, die den Rahmen des Projekts in der Leistungsbeschreibung der Ausschreibung absteckten. Dazu die Anforderung, ein Monitoringsystem bereitzustellen, das der Deutschen Nationalbibliothek (DNB) schnellen Zugang zu Digitalisaten, OCR-Ergebnissen und Metadaten sowie Auswertungen zum Bearbeitungsstand und auch die Möglichkeit direkter Kommunikation mit dem Dienstleister bei Fragen und Reklamationen bieten sollte.

Bewerbung  
nicht selbst-  
verständlich

Die Entscheidung, ein Angebot für diese Ausschreibung abzugeben, war nicht selbstverständlich. Ein so großes Volumen war weder mit den vorhandenen Geräten noch mit den verfügbaren festen Mitarbeiterinnen und Mitarbeitern in der vorgegebenen Zeit zu bewältigen. Das geforderte Monitoringsystem allerdings, ein zentraler Punkt in der Leistungsbeschreibung, war als Eigenprogrammierung aus einem anderen Projekt schon weitgehend verfügbar, die Anpassung an die spezifischen Projektvorgaben kurzfristig möglich.

Ist das zu  
schaffen?

Die entscheidende Frage lautete also: Ist die Digitalisierung inklusive sämtlicher zu erbringender Vorleistungen wie Beschaffung und Aufbau der Geräte, Anpassung des Monitoringsystems, Einrichtung des Workflows und der Schnittstellen zwischen DNB und Produktionssystem und die Einstellung und Einarbeitung neuer Mitarbeiterinnen und Mitarbeiter im vorgegebenen Zeitraum überhaupt zu schaffen?

Heute, im Rückblick auf das Projekt, können wir diese Frage entspannter betrachten als bei Projektbeginn und sie uneingeschränkt bejahen. Tatsächlich wurden in der vorgegebenen Zeit nicht nur die als Gesamtvolumen geschätzten 4 Millionen Seiten,

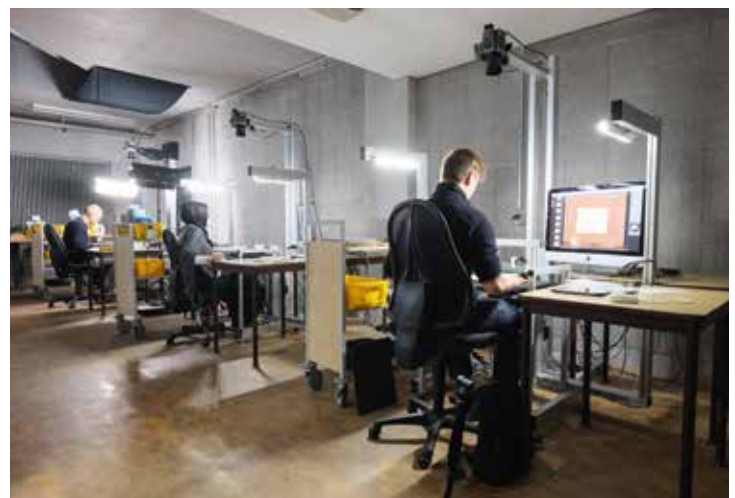
sondern mit rund 4,7 Millionen Seiten sogar knapp 20 Prozent zusätzlich bearbeitet.

Damit bestätigte sich die Vorabkalkulation der erforderlichen Digitalisierungskapazität: Unter Berücksichtigung eines Puffers für Nachbearbeitungen und für den Fall, dass das gesamte Seitenvolumen die Prognosen übertreffen sollte, war das täglich zu bearbeitende Seitenvolumen festgesetzt worden. Verfügbar waren rund 200 Arbeitstage inklusive Samstagen. Um auf der sicheren Seite zu sein, wurde ein durchschnittlich zu bearbeitendes Tagesvolumen von 24.000 Seiten, also insgesamt 4,8 Millionen Seiten, vorgesehen. Nach den vorliegenden Erfahrungswerten konnte dieses Volumen mit fünf Produktionslinien, die im Schichtbetrieb durchschnittlich 12 Stunden pro Tag laufen würden, erreicht werden.

Diese Schätzung erwies sich als äußerst vorsichtig und wurde regelmäßig übertroffen. Der zusätzliche Einsatz einer sechsten Produktionslinie im August, der geringere Tagesleistungen während der Anfangszeit des Projekts ausgleichen sollte, führte in diesem Monat regelmäßig zu Tagesleistungen von weit über 40.000 Seiten.

Ja, und  
sogar noch  
etwas mehr

Kalkuliertes  
Tagesvolumen  
regelmäßig  
übertroffen



Blick in die Digitalisierungswerkstatt: Aufnahmearbeitsplätze  
Foto: Deutsche Nationalbibliothek, Stephan Jockel

## Digitalisierungstechnik

Bei der Wahl der Aufnahmegeräte waren folgende Kriterien ausschlaggebend:

Kriterien  
Aufnahme-  
geräte

- Hohe Scangeschwindigkeit: Die Geräte durften auf keinen Fall langsamer als die Person sein, die daran arbeitet. Das heißt: Je nach Material und Mitarbeiter mussten Spitzenwerte von bis zu 400 Aufnahmen oder 800 Seiten pro Stunde möglich sein.
- Hohe Verfügbarkeit: Stand- und Reparaturzeiten länger als 24 Stunden waren angesichts des engen Zeitplans nicht akzeptabel; Reparaturen mussten möglichst von uns selbst ausgeführt werden können; einige Ersatzteile mussten vorrätig sein.
- Robuster Dauereinsatz: Täglich bis zu 14 Stunden und bis zu 4.500 Aufnahmen mussten alle Komponenten der Ausstattung mitmachen.
- Durchgehend höchste Scanqualität mit konstanter Schärfe und Farbverbindlichkeit. Kalibrieren, Nachjustieren etc. sollten nicht häufig erforderlich sein und mussten mit geringem Zeitaufwand durchgeführt werden können.
- Schonung der Digitalisierungsvorlagen und flexible Einsatzmöglichkeit der Geräte bei unterschiedlichen Formaten und Ausstattungen; auch Sonder- und Überformate innerhalb einer Bindeeinheit mussten ohne Gerätewechsel bearbeitet werden können.
- Leichte Bedienbarkeit: Kurze Einarbeitungszeiten und schnelles Erreichen der Mengenvorgaben bei hoher Qualität von Anfang an.

Um den beschriebenen Anforderungen gerecht zu werden, kombinierten wir die besten Digitalkameras der Welt, Hasselblad-Kameras, mit Buchtischen aus eigener Entwicklung und leistungsstarken Rechnern von Apple.

Die optischen Qualitäten von Hasselblad-Kameras, also Auflösung und Schärfe, liegen jenseits aller Zweifel. Sie schaffen etwa 1,5 Bilder pro Sekunde und haben nur ein kleines mechanisches Bauteil, die Verschlussblende. Der vom Hersteller garantierte Lebenszyklus des Verschlusses liegt bei 100.000 Aufnahmen, tatsächlich haben die im Projekt eingesetzten Kameras bis auf eine Ausnahme mit dem ersten Verschluss jeweils rund 400.000 Aufnahmen hinter sich gebracht. Als Aufnahmesoftware kam Hasselblads »Phocus« zum Einsatz.

Hasselblad

Digitalisiert wurde mit 300 dpi Auflösung und 24 bit Farbtiefe, als letzte Aufnahme wurde bei jedem Buch eine Farbkontrollkarte aufgenommen. Bei den Buchtischen, digitalisiert wurde fast durchgängig unter Glas, wurde auf die Schonung der Vorlagen beim Digitalisierungsvorgang geachtet, bei gleichzeitig optimaler Lagerung mit Blick auf die Qualität der Digitalisate. Für überformatige Bände stand ein größerer Buchtisch zur Verfügung, »Problembände«, die nicht 180 Grad geöffnet werden sollten, wurden an einem mit zwei Kameras ausgestatteten 120-Grad-Buchtisch bearbeitet.

## Monitoringsystem und Workflowmanagement

Neben der Aufnahmetechnik waren das Monitoringsystem und das Workflowmanagement zentrale Komponenten bei der Durchführung des Projekts. Über das Monitoringsystem wurde der Bibliothek Zugriff auf Digitalisate, Metadaten und OCR-Ergebnisse nach den Vorgaben der Leistungsbeschreibung gewährt und der jeweilige Bearbeitungsstatus jedes Werks dokumentiert. Über das Workflowmanagement wurde der gesamte Workflow von der Aufnahme bis zur Ablieferung der Daten an die Bibliothek organisiert und weitgehend automatisiert durchgeführt.



Blick in die Digitalisierungswerkstatt: Eingangsbereich mit neu angelieferten Büchern auf dem Bücherwagen und digitalisierten Werken im Regal, Kontrollarbeitsplätze  
Foto: Deutsche Nationalbibliothek, Stephan Jockel

Vorgabe der DNB war, dass die meterweise aufgehobenen und bereitgestellten Bücher jeweils geschlossen in der korrekten Signaturfolge zurückgegeben werden sollten. Da absehbar war, dass es immer wieder einzelne Bände geben würde, deren Bearbeitung (inklusive interne und externe Qualitätskontrolle) länger als die vorgesehenen zwei Wochen Durchlaufzeit dauern würde, musste davon ausgegangen werden, dass bis zu 3.000 Bände gleichzeitig im Digitalisierungsraum aufzubewahren sein würden. Die schnelle und einfache Identifikation und Auffindbarkeit jedes Bandes während aller Bearbeitungsschritte war deshalb von zentraler Bedeutung. Vor diesem Hintergrund wurde die Aufteilung der angelieferten Bände in Boxen mit jeweils zehn bis 20 Bänden und demgemäß eine fortlaufende, aber kleinteilige interne Kennzeichnung der Bücher vorgesehen.

Im Einzelnen setzte sich der Bearbeitungsablauf der Digitalisierung aus folgenden Schritten zusammen:

Bearbeitungsablauf

- Übernahme der auf Bücherwagen angelieferten Bände.
- Einstellen jedes Bands in das Monitoringsystem durch Einlesen des Akz-Barcodes. Dabei erfolgten automatisch: SRU-Abfrage des Katalogs zur Übernahme der IDN des digitalen Objekts und des später für die OCR benötigten Sprachcodes in das System. Vergabe der internen zweiteiligen Kennung und Ausdruck eines Laufzettels mit Akz, IDN und interner Kennung.
- Prüfung jedes Bands auf Erhaltungszustand, möglichen Öffnungswinkel und Besonderheiten wie Faltafeln etc.; Aussortieren von Dubletten (Erkennung durch das Monitoringsystem) und Bänden, bei denen eine Nachbearbeitung der Katalogdaten oder eine Rücksprache mit der DNB erforderlich war. Bände, die eine besondere Handhabung bei der Digitalisierung erforderten, wurden bereits hier ausgesondert, um sie gezielt an bestimmte Mitarbeiter/Aufnahmegерäte weitergeben zu können.
- Bereitstellung der Bände in Boxen zur Verteilung an die einzelnen Aufnahmegерäte.
- Durchführung der Digitalisierung (dabei Benennung der Digitalisate mit der von der DNB vergebenen IDN).
- Einstellen der digitalisierten Bände in fortlaufender Boxenzählung in die Buchregale des Produktionsraums.

- Export der Rohdaten und Speicherung der TIFF-Dateien auf dem zentralen Speicher, Verteilung der Digitalisate an die Kontrollarbeitsplätze zur Qualitätskontrolle, OCR-Vorbereitung und Erstellung der Strukturdateien.
- Abholung von Digitalisaten und Strukturdaten von den Kontrollarbeitsplätzen, Bereitstellung der Digitalisate für die OCR,<sup>1)</sup> Erstellung von Thumbnails und Preview-Dateien, Zusammenführung von Katalog-, Struktur-, projektspezifischen und Bildmetadaten nach dem METS-Standard.
- Im Anschluss an die OCR-Bereitstellung der kompletten Daten für die Qualitätssicherung der DNB.
- Nach Freigabe der Daten durch die DNB beziehungsweise nach Ablauf der Prüfungsfrist Erstellung des tar-Transferpakets (Bilddaten, Metadaten und OCR-Ergebnisse [TXT, ALTO] mit MD5-Prüfsumme) und Weiterleitung der Daten an einen externen Speicherdienstleister;<sup>2)</sup> im späteren Projektverlauf gleichzeitiger Upload in den Hotfolder der DNB. Von der DNB reklamierte Werke wurden ebenso wie Fragen zur Bearbeitung über das Monitoringsystem an die Kontrollarbeitsplätze zurückgegeben.
- Nach Freigabe der digitalisierten Werke wurden die fertig bearbeiteten Bände aus dem Monitoringsystem ausgetragen und in der ursprünglichen Reihenfolge auf Bücherwagen zur Rückgabe an die DNB bereitgestellt.

## Der Faktor Mensch

Die Digitalisierung der Exilmonografien in nur wenigen Monaten war zweifellos nur mit einer für diesen Kontext optimalen technischen Ausstattung möglich. Dennoch sollte hier abschließend auch ein Blick auf den Beitrag der beteiligten Personen geworfen werden.

Im Umgang zwischen den Projektleitungen von DNB und Dienstleister wurde konsequent auf kurze, effiziente Kommunikations- und Entscheidungswege gesetzt. Mehrmals wöchentlich kam die Projektleiterin der DNB in die Digitalisierungswerkstatt, um sich über den aktuellen Projektstand zu informieren, Fragen zu klären und Lösungswege für Probleme zu besprechen, die insbesondere in der Anfangszeit aus der Praxis heraus auftraten und

Kurze Kommunikations- und Entscheidungswege

von beiden Seiten so nicht vorhersehbar waren. Vortermi­nierte längere Besprechungen konnten so auf ein Mindestmaß reduziert werden. Zur Effizienz der Kommunikation trug auch wesentlich bei, dass Zuständigkeiten auf beiden Seiten klar definiert waren und je nach Sachverhalt die betroffenen Mitarbeiterinnen und Mitarbeiter ohne weitere Umwege miteinander kommunizieren und anstehende Fragen klären konnten.

Effizientes  
Arbeiten durch  
hochwertige  
Technik

Im Verhältnis Technik/Mensch erwies es sich als richtig, dass beim Digitalisierungsgerät auf eine hochwertige und flexible Technik gesetzt worden war, die weder besonders schnell und effizient arbeitende Mitarbeiterinnen und Mitarbeiter ausbremste, noch an aufwändigen speziellen Digitalisierungsanforderungen scheiterte. Auch bewährte es sich, dass mit dem eigenen Monitoringsystem und Workflowmanagement auf ein schlankes, auf die konkreten Projekterfordernisse zugeschnittenes System gesetzt worden war, das eine rasche Einarbeitung erlaubte und auch im laufenden Projekt noch weiter optimiert werden konnte.

Fazit: Eine hochwertige und flexible technische Ausstattung als Werkzeug, engagierte und kompe-



Mitarbeiterin bei der Digitalisierung  
Foto: Deutsche Nationalbibliothek, Stephan Jockel

tente Mitarbeiterinnen und Mitarbeiter und deren Einsatz nach ihren individuellen Stärken, die von Anfang an konstruktive und zuverlässige Kommunikation und Zusammenarbeit zwischen Auftraggeber und Dienstleister als Rahmenbedingung – der Faktor Mensch hat Gewicht.

Fazit

## Anmerkungen

- 1 Angesichts des zu bearbeitenden Volumens und engen Zeitplans war von Anfang an vorgesehen worden, die OCR nicht selbst vorzunehmen, sondern dafür ein anderes Unternehmen über einen Subauftrag in das Projekt einzubinden. Beauftragt wurde die Firma Herrmann und Kraemer, Garmisch-Partenkirchen, die in direkter Anbindung an unseren Workflow die OCR in Frankfurt am Main durchführte.
- 2 Auch bei der Speicherdienstleistung war von Anfang an die Zusammenarbeit mit einem anderen Unternehmen geplant. Beauftragt wurde das Unternehmen cw it services, Dreieich.