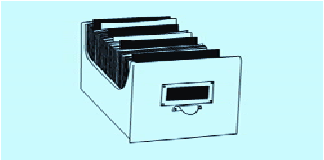
Karteikarten im Dt. Bundesarchiv - Erfahrungen mit der Digitalisierung



Vielen Dank, daß Sie mir noch einmal zuhören möchten. In diesem Vortrag geht es um ein Projekt, in dem ausschließlich Karteikarten digitalisiert wurden.

Die Vorgabe/Voraussetzungen

Die Stiftung Archiv der Parteien und Massenorganisationen der DDR im deutschen Bundesarchiv bewahrt Archivgut aus den zentralen Leitungsebenen der Parteien, Gewerkschaften und Massenorganisationen der DDR. Um den Zugang zu diesem Material zu verbessern, sollten die Findmittel digitalisiert werden. Dabei waren folgende Punkte problematisch:

1. der große Umfang der Findmittel macht es unmöglich, diese durch die Archivare parallel zum täglichen Archivbetrieb zu digitalisieren. Deswegen bot sich ein Outsourcing an.

2. auf Grund der Überlieferungsgeschichte verfügen die einzelnen Bestände über sehr heterogene Findmittel erstens in Bezug auf Art (Karteien unterschiedlicher Formate sowie Findbücher) und zweitens in Bezug auf die Qualität (unsystematische Titelbildung, handschriftliche Ergänzungen).
POWERPOINT Beispielkarten

3. die Findmittel sollten nur möglichst kurz aus der Benutzung genommen werden.

Weitere Rahmenbedingungen waren eine Fehlerfreiheit von 99,98% für die allgemeine Texterfassung und die Signaturen.

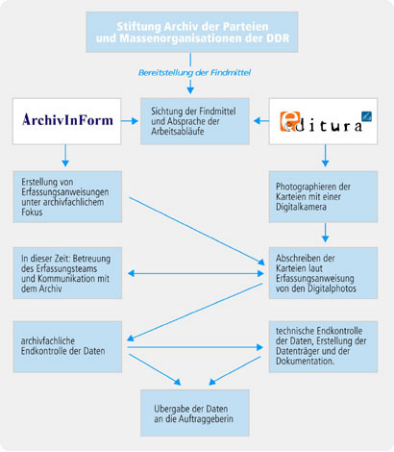
Das erste Projekt umfasst die Verarbeitung von ca. 150.000 Datensätzen, die sich auf etwa 143 Beständen bzw. ca. 4.500 lfm. Archivgut beziehen. Inzwischen arbeiten wir in einem Rahmenvertrag für alle Außenstellen des Bundesarchivs, hier werden sowohl Karteien als auch Findbücher verarbeitet. Wieviel Meter Archivgut nun dahinterstecken weiß ich nicht.

Als Zielformat wurden Daten für MidosaXML gewünscht, weil dieses System zur Findbuchredaktion und -erstellung optimal die Bedürfnisse des Bundesarchivs erfüllt. Inzwischen arbeiten wir auch mit dem zweiten im Bundesarchiv vorhandenen Datenbankprogramm Basys. Dieses Programm hat noch kein fertiges Übernahmetool und somit ist die Lieferung von Daten recht schwierig.

Fazit: Diese Digitalisierung ist technisch und archivfachlich sehr anspruchsvoll.

Ich möchte Ihnen im Folgenden zeigen, wie wir unter Berücksichtigung dieser schwierigen Rahmenbedingungen ein sehr gutes Ergebnis erzielen konnten.

Unser Lösungsweg



Die Zusammenarbeit sieht im Schema hochkompliziert aus. Aber jede Firma bringt ihr spezielles Fachwissen ein, das in verschiedenen Schritten einer Digitalisierung notwendig ist:
Die Firma editura bietet technisches KnowHow, was xml-Strukturen und Hard- und Software angeht, Erfahrungen mit der Umsetzung von umfangreichen literarischen Textsammlungen sowie preiswertes und zahlreiches Personal in Rumänien. Wir von ArchivInForm bringen archivfachliche und analytische Kompetenz sowie Erfahrungen mit Digitalisierungen von Findmitteln ein.

Zur wirtschaftlichen Realisierung dieses Projektes haben wir erfolgreich folgenden Weg beschritten:

1. Photographieren der Karteikarten mit einer Digitalkamera bei mehrmaligem Prüfen der Vollständigkeit

2. parallel dazu: Sichtung der Findmittel und Erstellung von Erfassungsanweisungen durch ausgebildete Archivare

3. Abschreiben der Findkarteien von den Digitalphotos laut Erfassungsanweisung bei einem Editura-Tochterunternehmen in Kronstadt/Rumänien

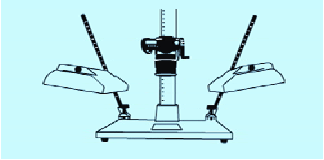
4. parallel dazu: Betreuung der archivfachlichen Fragen des Erfassungsteams durch die ArchivInForm GmbH und Kommunikation mit dem Auftraggeber

5. archivfachliche Endkontrolle in Deutschland durch die Firma ArchivInForm GmbH

6. technische Endkontrolle der Daten, Erstellung der Datenträger und der Dokumentation

7. Übergabe der Daten an das Bundesarchiv

1 Vorbereitung der Karteien und Digitalphotographie



Im Gegensatz zum Projekt in Basel, wurden die Vorlagen von uns als Auftragnehmer gescannt. Um die Karteien nur kurz aus der Benutzung nehmen zu müssen, erfolgte die Imageerstellung vor Ort in den Räumen des Bundesarchives. Bei Karteien handelt es sich meistens um Unikate, von denen keine Kopien vorliegen, die herausgegeben werden können.
In einem ersten Schritt wurden die Karteien gestempelt und in Hunderter-Blöcke eingeteilt, die wir Batches nennen. Imageerstellung und Weiterverarbeitung erfolgte dann anhand dieser Blöcke, um handliche Einheiten für die Qualitätssicherung zu haben.

Die Karten wurden dann mit einer 6-Megapixel-Digitalkamera photographiert, die in einem Reprostativ mit Beleuchtung montiert war. Damit konnten die Karten in ihrem Originalzustand digitalisiert werden. Formatanpassungen oder eine Entmetallisierung sind bei diesem Vorgehen nicht notwendig, im Gegensatz zu den störanfälligen Einzugscannern.

Parallel zum Photographieren wurde ein Protokoll geführt, in dem Besonderheiten vermerkt waren, z.B. zusammen geklammerte Karten, Rückseiten, Gliederungspunkte oder besondere Hinweise für die Erfassung. An Hand dieses Protokolles wurde vor dem Überspielen geprüft, ob die Karten eines Batches vollständig erfasst wurden.

War ein Batch fertig photographiert und stimmte die Anzahl der Photos in der Kamera mit den Werten auf der Protokollkarte überein, wurde die Speicherkarte aus der Kamera entnommen und die Daten wurden auf ein Notebook überspielt. Auf dem Notebook wurden die Bilddaten noch einmal stichprobenhaft auf Vollständigkeit und Lesbarkeit geprüft.

Durch dieses Verfahren werden ca. 5000 Karten pro Tag digitalisiert. Dies ist ein Wert, den wir mit ruhigem Gewissen verantworten können, denn die Methode ist effizient, aber auch gründlich genug.

2 Erfassungsanweisung

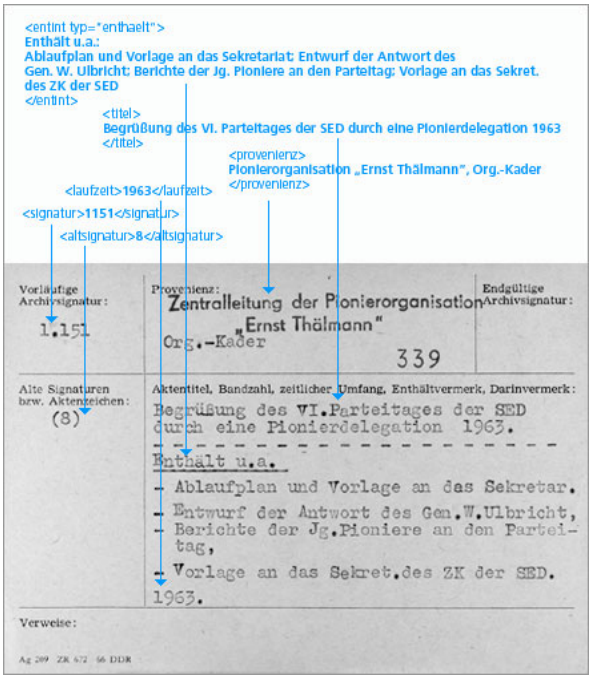


Abb: Beispiel für Feldzuweisungen

Vor der Erfassung werden alle Wünsche des auftraggebenden Archivs bezüglich Erfassung, Vereinheitlichung, z.B. von Laufzeiten, und Nichterfassung von Titelbestandteilen aufgenommen. Es werden zunächst die Allgemeinen Erfassungsanweisungen erstellt, die sich auf alle Bestände beziehen. Hier sollen so viele Fragen wie möglich schon vorher festgelegt werden, um später im Internet ein einheitliches Bild der Findmittel bieten zu können. Die Beteiligung mehrerer Außenstellen sowie der Zentralabteilung in Koblenz machte die Kommunikation jedoch schwerfällig. So hat die Ausarbeitung der allgemeinen Erfassungsanweisungen sehr viel Zeit in Anspruch genommen. Dieser Zeitaufwand lohnt sich jedoch, da die speziellen Anweisungen, die für einzelne Bestände gelten, dann schneller formuliert sind. Unsere Erfahrung mit mehreren Digitalisierungsprojekten hilft uns dabei, die Archive eingehend durch Lösungsvorschläge zu beraten.

Im zweiten Schritt und mit Hilfe dieses Vorwissens werden die Vorlagen beständeweise von unseren Archivaren analysiert. In einer erneuten Kommunikation mit dem Bundesarchiv werden bestandsspezifische Einzelfragen geklärt, die sich aus der Analyse ergeben haben.

Der Vorteil dieser Voranalyse ist die Entlastung der Archivare vor Ort, das Archiv bereitet die Findmittel also nicht oder nur in geringem Maße selbst vor. Da die Karteien schon vor der Erfassung auf Besonderheiten einzelner Bestände genau geprüft und entsprechende Anweisungen formuliert werden, wird der weitere Ablauf nicht durch häufige Nachfragen gestört.

Der Unterschied zu dem Projekt des Staatsarchivs Basel liegt also auf der Hand. Die Vorlagen sind nicht von Archivaren erstellt worden, sondern von Registratoren oder anderen Mitarbeitern der DDR-Verwaltungen. In den Wendezeiten wurden diese zum

Teil durch Archivare überarbeitet, und sind deswegen äußerst heterogen. Es waren also umfassendere Analysen, Kontrollen und Korrekturen als für Basel nötig.

Die Vielzahl an beteiligten Personen erfordert ein Höchstmaß an Organisation und umfangreiche Protokollpflichten. Um die Kommunikation so einfach und direkt wie möglich zu halten, haben wir eine Online-Platform eingerichtet: Wir stellen all unsere Erfassungsanweisungen sowie Übersichten über den Arbeitsstand und aktuelle Mitteilungen in ein Firmen-Wikipedia, auf das sowohl alle Beteiligten der Firmen ArchivInForm und editura, als auch das Bundesarchiv zugriff haben.

3 Datenerfassung

Die Erfassung der Daten erfolgte an Hand der Digitalphotos im Einzelfall aber auch anhand von Kopien in Papierform in Kronstadt/Rumänien. Pro Bestand und unterteilt nach Batches kam je ein Team von erfahrenen Erfassern, Korrektoren und Technikern zum Einsatz. Grundlage dafür waren die allgemeinen und bestandsspezifischen Erfassungsanweisungen.

Es wird zur Erfassung eine selbst entwickelte Software bentutz, die bereits während der Eingabe die Daten nach formal-technischen Kriterien automatisch prüft. Die intellektuelle und sprachliche Prüfung wurde durch Korrektoren geleistet, denen wegen der archivfachlichen Anforderungen und wegen der geforderten Genauigkeit eine große Bedeutung zukam. Es arbeiten in Rumänien ausschließlich deutsche Muttersprachler oder Mitarbeiter, die z.B. als Lehrer Deutsch studiert haben.

Betreuung der Erfassung

Während der Erfassung auftretende Fragen wurden an die ArchivInForm GmbH gestellt, deren Archivare in ca. 95% der Fälle selbst entscheiden konnten. Ca. 5% der Fragen mussten mit dem Bundesarchiv geklärt werden. Damit wird deutlich, welche enorme Entlastung den Archivaren durch eine Zusammenarbeit mit externen Fachkräften entsteht.

Die Archivare von ArchivInForm betreuen die Retrokonversion von Anfang an und kennen die allgemeinen und speziellen Anforderungen an das Projekt. Sie prüfen die Einhaltung der Vorgaben zu Struktur und Inhalt sowie die Vollständigkeit der Erfassung. Einzelne von den Erfassern nicht lösbare Probleme werden nachgearbeitet und Handschriften nacherfasst. Werden Mängel oder Ungenauigkeiten erkannt, gehen die Dateien zur Nachbearbeitung zurück an die Erfassung.

Endkontrolle

Bei der Endkontrolle der Daten bedürfen Handschriften der besonderen Fürsorge. Auch wenn in Rumänien die deutsche Sprache Voraussetzung für eine Mitarbeit ist, ist die Grammatik oftmals wegen Zerschreibungen von Wort-Endungen schwierig. Für solche Karteikarten werden in Rumänien Datensätze mit den entsprechenden Feldern und Texten angelegt und als »Handschrift« markiert. Dadurch sind sie für die deutschen Korrektoren der Firma ArchivInForm GmbH schnell auffindbar. 100% der Handschriften werden in Deutschland überprüft und korrigiert.

Qualitätssicherung

Die Qualitätssicherung erfolgt in mehreren Stufen. Wir unterscheiden zwischen technischer und inhaltlicher Qualität.

Die technische Qualitätsprüfung umfasst die Vollständigkeit der Karten anhand der Stempelnummern, Batches und Protokolle. Ebenso werden die Daten während der Erfassung auf formale Kriterien hin geprüft.

Die Fehlerfreiheit von 99,98 % erreichen wir durch die doppelte Korrektur in Rumänien und die Endkontrolle in Deutschland. Die Endkontrolle bezieht sich vorrangig auf die Hierarchie der Gliederungspunkte und Datensätze sowie die 100%-ige Fehlerfreiheit der Signaturen und der Feldzuweisungen.

Letzte Fragen werden mit dem Bundesarchiv geklärt und danach die Daten zur Abgabe freigegeben.

Durch die Aufteilung der Qualitätssicherung in einen internen (Editura GmbH) und einen externen Bereich (ArchivInForm GmbH) garantieren wir die hohe Fehlerfreiheit sowie eine Reduzierung des Prüfaufwands durch die Auftraggeber.

Datenübergabe

Dem Auftraggeber wurden die XML-Daten im Format von MidosaXML bzw. Basys sowie die Images von den Findkarteien übergeben. Diese wurden in das lokale Netz des Archives eingespielt. Damit stehen sie sofort den Mitarbeitern zur internen Benutzung zur Verfügung.

Zusammenfassung und Perspektiven

Der gesamte Auftrag von über 150.000 Datensätzen konnte innerhalb von nur 8 Monaten realisiert und dem Bundesarchiv zur Nutzung übergeben werden. Durch die mehrschichtige Kontrolle der Ergebnisse konnte eine hohe Fehlerfreiheit realisiert werden. Die Zufriedenheit des Bundesarchivs sowohl mit der unkomplizierten und direkten Ansprache im laufenden Projekt sowie mit dem Ergebnis gibt uns bei dem von uns gewählten Vorgehen Recht.

Nichtsdestotrotz ist diese Projektbeschreibung nur eine weitere der möglichen Vorgehensweisen, die an unterschiedlichste Vorgaben und Wünsche der Archive angepasst werden kann. Denn: Jedes Archiv hat eine andere Geschichte und somit unterschiedlichste Bestände und Findmittel. Aber: wir schaffen das!