

Historische Datenbestände für die Nachwelt retten und zugänglich machen

Das Ahnenforschungsportal Ancestry arbeitet bereits seit vielen Jahren mit Archiven, Museen, Sammlungen und Bibliotheken auf der ganzen Welt zusammen, um deren Datenbestände dauerhaft zu bewahren und für die Allgemeinheit nutzbar zu machen. Dabei werden Volkszählungsunterlagen, Passagierlisten, Kirchenbücher, militärhistorische Quellen, Bürger- und Einwohnerverzeichnisse, Gerichts- und Katasterakten, familiengeschichtliche Aufzeichnungen, Zeitschriften und eine Vielzahl unterschiedlicher historisch-biographischer Quellen digitalisiert.

Auch in Deutschland unterstützt ein Team Archive, Bibliotheken, Verlage, Forschungsstellen und genealogische Arbeitsgruppen bei der digitalen Erfassung und Bewahrung ihrer Bestände. Sie alle sind dadurch über das Internet für eine breite Nutzerzahl zugänglich und ihr Inhalt für die Nachwelt gesichert.

Die Digitalisierungsarbeit von Ancestry.de ermöglicht nicht nur eine stetige Erweiterung des Angebots an historischen Datenbeständen für alle Ahnen- und Familienforscher, sondern ist zudem als ein wichtiger Beitrag zur Sicherung und Nutzbarmachung wertvollen Kulturgutes zu sehen. Die historischen Bestände können in digitaler Form uneingeschränkt genutzt und eingesehen werden, ohne dabei beansprucht zu werden oder Schaden zu nehmen.

A) Ancestry.de unterstützt Archive bei der Digitalisierung ihrer Bestände

1. Expertise

Ancestry scannt bereits seit mehr als zehn Jahren historische Unterlagen in verschiedenen Formaten ein und hat bereits mehr als 235 Millionen Scans öffentlich verfügbar gemacht. Die Geschwindigkeit, mit der Bilder in die Datenbank aufgenommen werden, nimmt derzeit weiter zu, so dass davon auszugehen ist, dass pro Jahr mehr als 80 Millionen neue Bilder hinzugefügt werden können.

2. Scannen

2.1 Dokumenten-Scan (Bücher, Zeitungen, Karten, etc.)

Ancestry verfügt über langjährige Erfahrung und modernste Technologien für das Einscannen von Büchern und Dokumenten. Hochwertigste, automatisierte Scanner, die sowohl Einzelblätter als auch gebundene Bücher schonend und zügig bearbeiten können, ermöglichen das Einscannen mehrerer Hunderttausend Seiten pro Monat. In Einzelfällen (z.B. wenn das Dokumentenformat kein automatisiertes Scannen zulässt) erfolgt die digitale Aufnahme von Dokumenten durch ein speziell geschultes Team per Hand.

2.2 Bild-Scan (Mikrofilm, Mikrofiche etc.)

Ancestry scannt Mikrofilme und Mikrofiche mit modernster Ausstattung im JPEG 2000 Standard. Die Digitalisierungskapazität liegt bei mehr als 3,5 Millionen Bildern pro Monat.

3. Scan Optimierungen

Ancestry verwendet einen patentierten Binarisierungsprozess, der die Qualität von Text und Bild entscheidend verbessert. Darüber hinaus ermöglicht die UV-Infrarot-Multispektralscan-Methode das Sichtbarmachen von Schriften, die bei normaler Beleuchtung nicht erkennbar sind.

4. Transkription

Da bei Ancestry schon von Anfang an bei der Bild- und Dokumentenkonversion qualitativ hochwertige Bilder entstehen, lassen sich zusätzlich Systeme zur automatischen Bild- und Texterkennung (OCR bzw. ICR) einsetzen. Hier werden nur die aktuellsten

Softwarelösungen verwendet, welche eine branchenführende Genauigkeit von 98 % erzielen. Mit Multiprozessor-Systemen liegt die Kapazität bei bis zu 2,2 Millionen Seiten pro Monat.

Zusätzlich unterhält das Unternehmen enge Partnerschaften mit den führenden Anbietern von Dateneingabediensten. Dort ist ein fester Stamm von speziell geschulten Mitarbeitern damit beschäftigt, handschriftliche Dokumente exklusiv für Ancestry zu transkribieren. Ein spezielles Prüfungssystem gewährleistet dabei eine Transkriptionsgenauigkeit von 99,1 %.

Mit über 1.500 in der Schriftbearbeitung geschulten Mitarbeitern und routinierten Verfahren können monatlich über 21 Millionen Seiten in Datenbankvorlagen übertragen werden.

5. Qualitätssicherung

Ancestry verfügt über umfangreiche Verfahren zur Qualitätssicherung, in deren Rahmen jeder Schritt des Digitalisierungsprozesses überwacht wird. Dadurch wird beim Endergebnis eine Akkuratheit von über 99 % gewährleistet. Nach der Bearbeitung der Bilder und während des Transkriptionsprozesses wird von einem Prüfungsteam eine vollständige Prüfung aller einzelnen Bilder durchgeführt.

Bevor die Daten über die Webseite der Öffentlichkeit zugänglich gemacht werden, überprüft eine unabhängige Qualitätssicherungsabteilung noch mal das endgültige, integrierte Projekt und gibt dies vor der Veröffentlichung frei. Mit Hilfe einer zufälligen Stichprobe werden die Bilder und der Index von einer Testperson visuell überprüft. Darüber hinaus kontrolliert eine automatische Testsoftware die Index/Bild-Kombination. Nach der Veröffentlichung können Kunden über ein Feedbacksystem für Anmerkungen und Korrekturen etwaige verbleibende Fehler melden.

6. Übergabe der Daten an das Archiv

Nach der Digitalisierung der Dokumente erhalten die Archive die Originale sowie eine Kopie der erstellten digitalen Daten zurück. Ancestry bereitet die Bilder und Indizes für die Zusammenstellung auf externen Datenträgern vor. Die Bilder und Indizes werden in einem Format bereitgestellt, das einfach in ein Dokumentenmanagementsystem importiert werden kann. Alle in einem Projekt genutzten Datenträger werden für die Verwendung in Archiven zertifiziert.

7. Datensicherheit

Datensicherheit ist ein extrem wichtiges Anliegen. Ancestry unterhält und betreibt mehrere hochsichere Datacenter in den USA mit mehr als 2000 Servern. In diesen Datacentern arbeiten nur Mitarbeiter, die nach strengsten Auswahlkriterien ausgesucht wurden.

B) Kooperationen

Ancestry arbeitet eng mit zahlreichen staatlichen und privaten Institutionen zusammen, um seinen Mitgliedern eine stetig wachsende Datenbasis und somit noch mehr Erfolge bei ihrer Recherche nach Vorfahren und Verwandten bieten zu können. In der Regel ergeben die Kooperationen eine beiderseitige Win-Win-Situation, da auch die Institutionen von der Digitalisierung ihrer Datenbestände profitieren. Neben den zahlreichen Kooperationen mit regionalen Archiven, Museen und anderen Institutionen sind es vor allem einige größere Partnerschaften, die eine Zusammenarbeit oft über mehrere Jahre hinweg beinhalten und fortlaufend die Bestände und Sammlungen erweitern. Zu den wichtigsten Partnerprojekten der letzten Jahre zählen folgende Beispiele:

- Deutsche Nationalbibliothek in Frankfurt und Leipzig: Historische Telefonbücher von Berlin, München, Hamburg, Frankfurt und Leipzig
- Staatsarchiv Bremen: Seeleuteregister und Musterungslisten
- Landeshauptarchiv Schwerin: Volkszählungen und Kirchenbuch-Duplikate des 19. Jahrhunderts
- Staatsarchiv Hamburg: Passagierlisten der Auswandererschiffe
- Bayerisches Hauptstaatsarchiv München: Kriegsstammrollen und -ranglisten der Bayerischen Armee im Ersten Weltkrieg

Beispiel: Zusammenarbeit mit dem Bayerischen Hauptstaatsarchiv

Die Zusammenarbeit mit dem Bayerischen Hauptstaatsarchiv / Abteilung Kriegsarchiv in Bezug auf die Digitalisierung der Kriegsstammrollen und -ranglisten der Bayerischen Armee im Ersten Weltkrieg kann als mustergültiges Beispiel für ein Projekt zwischen Ancestry und einer öffentlichen Institution angesehen werden:

Die Kriegsstammrollen und -ranglisten der Bayerischen Armee im Ersten Weltkrieg umfassen ca. 23.000 Bücher mit insgesamt ca. 4 Millionen Namenseinträgen und werden seit dem Ende des Ersten Weltkriegs im Kriegsarchiv München aufbewahrt. Dort können sie lediglich gelagert und auf spezielle Anfragen hin für interessierte Personen zugänglich gemacht werden. Immerhin ist das Interesse von Ahnenforschern an den Soldatenakten mit durchschnittlich mehreren hundert Anfragen im Jahr sehr groß. Eine systematische Recherche in den Beständen nach Namen oder anderen persönlichen Daten war vor der Digitalisierung allerdings nicht möglich, sodass jede Anfrage eine zeitaufwändige Suche vor Ort notwendig machte.

Zum Aufarbeiten, zur Erfassung und zur Indizierung der Bestände fehlten dem Archiv sowohl personelle als auch monetäre Mittel. Mit dem zur Verfügung stehenden Personalbestand hätte eine digitale Erfassung der Bestände Jahrzehnte gedauert. Ancestry wurde daher als Partner begrüßt, da das Unternehmen sowohl die technische und personelle Expertise und Kapazität aufweisen kann und zudem die Ergebnisse der Arbeit über die Webseite einer breiten Öffentlichkeit zugänglich sind.

Die Kooperationsanfrage seitens Ancestry mündete in eine vertragliche Vereinbarung, nach der das Unternehmen über einen bestimmten Zeitraum den Zugang zu den Beständen erhält, diese digital erfasst und indexierbar macht, sowie seinen Mitgliedern online zur Familien- und Ahnenforschung zur Verfügung stellen kann.

Im Gegenzug erhält das Archiv die digitalen Daten und einen Zugang zu Ancestry. Somit können nicht nur Interessierte in aller Welt, sondern auch die Besucher des Archivs die Akten als hochauflösende Scans über die Ancestry-Webseite einsehen und müssen nicht länger die Originale zur Recherche verwenden. Dies spart sowohl den Besuchern als auch den Mitarbeitern des Archivs Zeit und Mühen und schützt zugleich die Bestände vor zu starker Beanspruchung und Abnutzung.

Da die Originalbände der Soldatenakten bislang völlig unaufbereitet waren und selbst eine Microfiche-Version nicht existierte, stellte Ancestry ein spezielles Team zur Verfügung, das die Bestände vor Ort digitalisiert und zur Auswertung und Indexierung weiter leitet. Somit können diese weiterhin im Archiv verbleiben und sind dort für händische Recherche verfügbar.

Die Kooperation zwischen Ancestry und dem Bayerischen Hauptstaatsarchiv startete im März 2008 und erstreckt sich bis zum Abschluss des Projektes voraussichtlich im Herbst 2011.