

Universität Innsbruck:

ABBYY Fine Reader Scripting Edition als Universal-Tool

Die Universität Innsbruck unterhält eine eigene Abteilung für Digitalisierung und elektronische Archivierung. Die Abteilung unter Leitung Dr. Günter Mühlbergers ist auch als Dienstleister aktiv und hat bereits zahlreiche Projekte für externe Organisationen realisiert. Für die Implementierung von OCR-Prozessen setzen die Spezialisten der Universität Innsbruck auf ABBYY-Technologie. Die Zusammenarbeit besteht seit vier Jahren und begann durch eine Partnerschaft im EU Projekt METADATA ENGINE. Das Projekt wurde von der Universität Innsbruck geleitet, im Rahmen des Projekts wurde auch die Basisentwicklung von FineReader XIX durchgeführt. (Link: <http://meta-e.aib.uni-linz.ac.at/>).

Zurzeit setzt die Abteilung für Digitalisierung und elektronische Archivierung ABBYY FineReader Scripting Edition in den Versionen 6.0 und 7.0 ein, deren Funktionen sie in den unterschiedlichsten Bereichen nutzt. Die SDK Version 7.1 wird derzeit gerade getestet. Drei Projekte sind in diesem Zusammenhang besonders hervorzuheben:

1. Die Digitalisierung des kompletten Karteikartenkatalogs der Universitätsbibliothek mit 1,1 Millionen Karteikarten im Format A6.
2. Die Bereitstellung alter Werke in Frakturschrift als eBook mit der Möglichkeit zur Volltextsuche.
3. Die Digitalisierung eines großen Zeitungsausschnittsarchivs mit rund einer Million Artikel. Das digitale Archiv wird sowohl im Volltext als auch mittels Bildsuche erschlossen.

1. Digitalisierung des Katalogs

Mit der Digitalisierung des Zettelkatalogs, der von 1500 bis 1988 reicht und mehr als 1,1 Millionen Karteikarten umfasst, hatte sich die Abteilung für Digitalisierung und elektronische Archivierung etwas Besonderes vorgenommen. Für die Online-Nutzung sollte der Katalog die Funktionsweise der traditionellen Karteikästen möglichst exakt nachbilden, d.h. der gesamte Inhalt des Katalogs sollte zwar per Volltextsuche zu erschließen sein, aber die ursprüngliche Struktur sollte erhalten bleiben: Anwender sollten wie gewohnt im Katalog blättern und zudem die Ergebnisse der Digitalisierung am Original kontrollieren bzw. ergänzen können.

Aufgrund begrenzter Budgets kamen nur zwei Vorgehensweisen in Frage: Die automatische Digitalisierung per OCR-Software oder manuelles Abtippen in einem Billiglohnland. Zwei Argumente gaben letztlich den Ausschlag für den Einsatz der ABBYY OCR-Software: Die hohe Präzision und Zuverlässigkeit des Systems und die Kosten. „Wir gehen davon aus, dass wir durch den Einsatz der ABBYY-Software Budget in der Größenordnung von 1:10 gegenüber einer Auslagerung in ein Billiglohnland gespart haben“, schätzt Mühlberger und ergänzt: „Wir haben das Projekt im März 2004 abgeschlossen und die Ergebnisse geben uns Recht. Wir haben die Suchresultate in unserem neuen Online-Katalog in Stichproben untersucht und festgestellt, dass lediglich 15% der Karteikarten nicht auf Anhieb zu finden sind. Mit etwas Spielerei bei der Formulierung der Suchanfrage und durch die Möglichkeit, im Katalog wie in einem herkömmlichen Karteikartenkatalog zu blättern, erzielt man aber sehr gute Ergebnisse.“

Der Online-Katalog ist unter www2.uibk.ac.at/ub/dea allgemein zugänglich und liefert die OCR-Ergebnisse Seite an Seite neben einer gescannten Karteikarte: (siehe Abbildung 1.0)

Zur Zeit verzeichnet der Katalog täglich ca. 1500 Abfragen, was eine deutliche Verbesserung gegenüber den früheren Werten bedeutet. Die Universität und ihre Studierenden können somit einen weitaus umfangreicheren Gebrauch der vorhandenen Ressourcen machen. Die Altbestände wurden vor dem Versinken in den Tiefschlaf gerettet. Auch in anderer Hinsicht profitiert die Universität Innsbruck von dem Projekt: Aufgrund des großen Erfolges bietet die Abteilung für Digitalisierung und elektronische Archivierung den Aufbau eines elektronischen Archivs externen Anwendern als Dienstleistung an. Derzeit wird mit Interessenten aus Österreich und Deutschland über weitere Projekte verhandelt.

ABBYY
FineReader
Scripting Edition **OCR 7.0**



Über Universität Innsbruck Abteilung für Digitalisierung und elektronische Archivierung – DEA

Die Universität Innsbruck unterhält eine eigene Abteilung für Digitalisierung und elektronische Archivierung. Die Abteilung unter Leitung Dr. Günter Mühlbergers ist auch als Dienstleister aktiv und hat bereits zahlreiche Projekte für externe Organisationen realisiert. Die DEA bietet die Digitalisierung von Büchern, Zeitschriften, Karteikarten, Manuskripten und anderen Materialien. So wurde bereits jedes Buch der Universitätsbibliothek Innsbruck – sofern es nicht mehr regulär im Buchhandel vertrieben wird – als digitales e-Book zur Verfügung gestellt. Darüber hinaus bietet die DEA die komplette Konvertierung von Zettelkatalogen: Scannen, Texterkennung, Indexierung mittels Oracle Datenbank und Hosting des Katalogs auf den Rechnern des Zentralen Informatik Instituts der Universität Innsbruck.

Kontakt

Universität Innsbruck
Digitalisierungsprojekte
Dr. Günter Mühlberger und Dr. Kurt Habitzel
Innrain 52
A-6020 Innsbruck
Tel +43/+512/507/9053
Fax +43/+512/5072607

