

Barbara Bauer M.A.
Erbprinzenstraße 15
76133 Karlsruhe
T +49 721 175-2250
F +49 721 175-2333
presse@blb-karlsruhe.de

Reichenauer Inkunabeln digitalisiert – ein Meilenstein der Texterkennung für Wiegendrucke

Pressemitteilung
Nr. 14 vom
29. April 2024

Mehr als 70.000 Seiten aus 242 Titeln in 123 Bänden: In nur 15 Monaten hat die Badische Landesbibliothek die Inkunabeln der ehemaligen Klosterbibliothek Reichenau digitalisiert und zugleich ihre Texte vollständig durchsuchbar gemacht. Dieses zuvor nicht digital verfügbare Material ist nun über die Digitalen Sammlungen der BLB für eine breite Nutzung zugänglich. Zusätzlich sind drei speziell für Inkunabeln trainierte Transkribus-Modelle zur freien Nachnutzung bereitgestellt.

Seit 15 Jahren digitalisiert die Badische Landesbibliothek Bestände aus dem kulturellen Erbe Südwestdeutschlands. Im Rahmen der Großen Landesausstellung „Welterbe des Mittelalters. 1.300 Jahre Klosterinsel Reichenau“ hat sie nun mit finanzieller Förderung durch die Stiftung Kulturgut Baden-Württemberg ein ehrgeiziges Vorhaben abgeschlossen: die Digitalisierung aller Inkunabeln Reichenauer Provenienz. Der Gesamtdatenbestand an Digitalisaten, Katalogisaten und Strukturdaten steht in den Digitalen Sammlungen der Badischen Landesbibliothek als Open Content zur Verfügung. Zusätzlich erfolgte eine KI-gestützte Volltexterkennung mittels mehrerer speziell trainierter Erkennungsmodelle für Deutsch und Latein sowie für zweisprachige Texte.

Das berühmte, mit zwei UNESCO-Welterbe-Titeln verbundene Kloster auf der Insel Reichenau war eines der einflussreichsten Klöster des frühen und hohen Mittelalters diesseits der Alpen. Der größte Teil seiner äußerst prominenten Bibliothek gelangte mit der Säkularisation 1804 nach Karlsruhe. Neben den stark beforschten frühmittelalterlichen Handschriften sind auch die dazu gehörigen Inkunabeln eine wichtige Forschungsquelle. Sie geben Aufschluss u.a. über kultur- und alltagshistorische Fragestellungen, aufgrund zahlreicher exemplarspezifischer Merkmale aber auch über die Verbindungen und Netzwerke des Klosters.

Mit diesem Projekt treibt die Badische Landesbibliothek die Kulturgutdigitalisierung und insbesondere die Anreicherung von Digitalisaten mit weiterführenden Metadaten aktiv voran. Aufgrund der in ihnen fortgesetzten mittelalterlichen Schreibgewohnheiten mit elaborierten Abbriviaturen und speziellen Layoutvarianten fordern Inkunabeln eine automatische Texterkennung besonders heraus. Das erfolgreich abgeschlossene Projekt setzt nun quantitativ und qualitativ Maßstäbe in der Texterkennung für Inkunabeln und macht es möglich, die Volltexterkennung systematisch auch auf andere digitalisierte Inkunabelbestände auszudehnen.



Herkömmliche OCR-Programme können Inkunabeln bisher nicht zuverlässig erkennen. Das Projekt sollte daher erproben, ob KI-basierte Software, in diesem Fall Transkribus, für die automatisierte Massenerkennung von Inkunabeln trainiert werden kann. Drei nach Sprachen differenzierte Modelle – ein deutsches, ein lateinisches und ein bilinguales deutsch-lateinisches Modell – wurden entwickelt und erreichten Zeichenfehlerraten von ca. 0,5 Prozent – einen Wert, der bis vor kurzem noch als unerreichbar galt.

Als größte Herausforderung dabei erwies sich die Layouterkennung, da Inkunabeln überdurchschnittlich viele problematische Layoutelemente aufweisen. Dazu zählen etwa Klammerglossen in juristischen Kommentaren, gedruckte Marginalien oder handgemalte Initialen. Die Standardeinstellungen von Transkribus führten in diesen Fällen zu großen Textverlusten. Trotzdem konnte eine teilautomatisierte Volltexterkennung erreicht werden.

Die digitalisierten und volltexterkannten Inkunabeln sind ein wertvoller Beitrag zum kulturellen Erbe und stehen pünktlich zur Eröffnung der Großen Landesausstellung sowohl für den lesenden Zugang als auch für weiterführende Forschungsfragen zur Verfügung.

Die angereicherten Digitalisate des Projekts sind unter folgendem Link zu finden:
<https://digital.blb-karlsruhe.de/topic/view/7530707>

Transkriptionsrichtlinien: <https://doi.org/10.57962/regionalia-22875>

Abbildungsmaterial zur Pressemitteilung

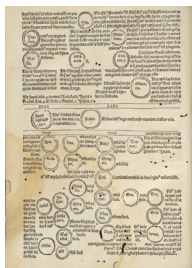


Abb. Nr. 1:
Werner Rolevinck
Fasciculus temporum
Memmingen: Albrecht Kunne, 1482
Badische Landesbibliothek, Lc 23
<https://digital.blb-karlsruhe.de/blrtra/content/pageview/7758560>



Abb. Nr. 2:
Amerigo Vespucci
Von der neu gefunden Region die wol ein welt genempt
mag werden, durch den Cristenlichen künig von Portigal
wunderbarlich erfunden
1505
Badische Landesbibliothek, Dn 401
<https://digital.blb-karlsruhe.de/blrtra/content/pageview/7474184>



Abb. Nr. 3:
Bildschirmaufnahme einer Volltextsuche nach „Achill“ im
Werk „De excidio Troiae historia“ mit Trefferhervorhebung.
<https://digital.blb-karlsruhe.de/blrtra/content/search/7621166?query=Achill>

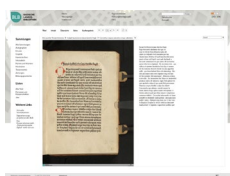


Abb. Nr. 4:
Bildschirmaufnahme einer zweiseitigen Ansicht mit
Digitalisat und Volltext.
<https://digital.blb-karlsruhe.de/blrtra/content/pagetext/7633867>