

Bibliotheca Eugeniana Digital (BED)

Vortrag beim Seminar „Quelle im Fokus: Aktuelle Fragestellungen und Methoden der Kodikologie und der Material Studies“

Simon Mayer Monika Kiegler-Griensteidl Annerose Tartler Florian Windhager

15. Mai 2023



Universität für
Weiterbildung
Krems



ÖAW

ÖSTERREICHISCHE
AKADEMIE DER
WISSENSCHAFTEN

Das Projekt wird aus Mitteln des Go!Digital 3.0-Förderprogramms der Österreichischen Akademie der Wissenschaften finanziert

Ablauf des heutigen Seminars

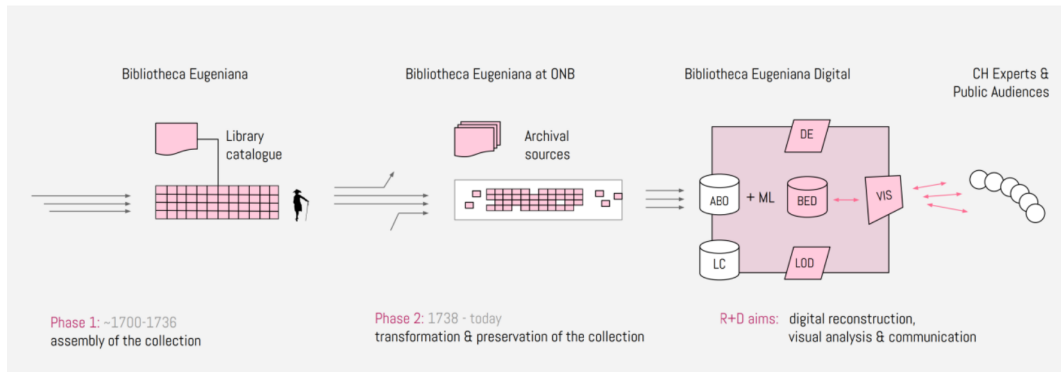
- ▶ Projektvorstellung BED (Simon Mayer und Annerose Tartler, 45 Minuten)
- ▶ Führung durch den Prunksaal (Monika Kiegler-Griensteidl, 30 Minuten)
- ▶ Input zur Sammlungsvisualisierung (Florian Windhager, 30 Minuten)
- ▶ Diskussion von “CollectionVis – Best Practices” (Florian Windhager, 30 Minuten)
plus Q&A (15 Minuten)

- ▶ Prinz Eugen und seine Sammlung
- ▶ Forschungsfragen des Projekts Bibliotheca Eugenia Digital
- ▶ Vorstellung Projektteam
- ▶ Zu erwartende Ergebnisse
- ▶ Vorstellung ausgewählter Methoden

Forschungsfragen des Projekts Bibliotheca Eugenia Digital

- ▶ Fragen zur Geschichte, Ausmaß und Zusammenstellung der Bibliothek Prinz Eugens
 - Wie viele gedruckte Bücher befinden sich in der Bibliotheca Eugenia?
 - Wie können sie identifiziert werden? Wie viele werden noch an der ÖNB gelagert und wo genau?
 - Wie ist die Struktur der Bibliothek und ihrer Wissenskassen?
- ▶ Fragen zur visuellen Analyse und Vermittlung der Sammlung
 - Wie können wir die Zusammenstellung und Provenienz dieser historischen Sammlung Experten vermitteln und visualisieren?
 - Wie können wir die Geschichte der Sammlung einer interessierten Öffentlichkeit vermitteln und visualisieren?

Übersicht Workflow



Vorstellung Projektteam



Bibliotheca Eugeniana Digital - BED



Österreichische
Nationalbibliothek

Universität für
Weiterbildung
Krems



- ▶ Seite online unter labs.onb.ac.at/bed/
- ▶ Kontakt: bed-project@onb.ac.at

Das Projekt „Bibliotheca Eugeniana Digital“ (BED) ist ein von der [Österreichischen Akademie der Wissenschaften](#) im Rahmen des „[go!digital 3.0“-Programms](#) gefördertes Kooperationsprojekt zwischen der [Österreichischen Nationalbibliothek](#) und der [Universität für Weiterbildung Krems](#). Die Laufzeit des Projekts beträgt zwei Jahre, von November 2022 bis November 2024.

Zu erwartende Ergebnisse

- ▶ Digitale Edition des handschriftlichen Katalogs
- ▶ Zusätzliche Metadaten für den modernen Bibliothekskatalog
- ▶ Visualisierungen der Sammlung
- ▶ Vorträge und Publikationen

Vorstellung ausgewählter Methoden

- ▶ Handschriftenerkennung mit Transkribus
- ▶ Bildklassifizierung mit Machine Learning
- ▶ Forschung in Archivmaterialien
- ▶ Sammlungsvisualisierung

Handschriftenerkennung mit Transkribus 1/4

Transkribus Expert Client v1.24.1 (13_01_2023_09:10), Loaded doc: codex_14.376, ID: 1260549, Page 1, file: 14.376_00000009.jpeg [Image Meta Info: (Resolution:1.0, w'h: 4248 * 6044)]

Search current document...

Server Overview Layout Metadata Tools

Document: codex_14.376, ID: 1260549

Collection: BED_codices, ID: 175372

Filename: 14.376_00000009.jpeg

Image URL: <https://files.transkribus.eu/Get?id=AMJEENFTILQGAPLQKOYO/>

Transcript URL: <https://files.transkribus.eu/Get?id=PLBYVTKCZBFLOVJBOHCU/>

Page/Transcript ID: 50958201/114607983

Thumbnail Overview:

Show Document Manager

1: Done 2: Ground Truth 3: Ground Truth 4: Ground Truth

5: Ground Truth 6: Ground Truth 7: Ground Truth 8: Ground Truth

9: Ground Truth 10: Ground Truth 11: Ground Truth 12: Ground Truth

13: Ground Truth 14: Final 15: Done 16: Ground Truth

Create thumbs for local doc Load thumbs

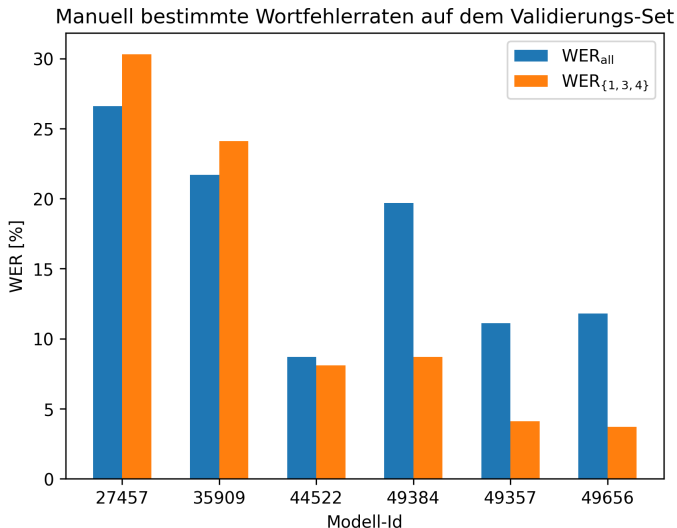
TR
L
BL
W
...
H
V
L
...
...

Catalogus
Librorum Bibliothecae
Serenissimi Principis Eugenij è Sabaudia
Ordine Materialiarum dispositus
& in III. Partes digestus
Cum Indice Alphabetico.
Tomus I.

1-1 Catalogus
1-2 Librorum Bibliothecae
1-3 Serenissimi Principis Eugenij è Sabaudia
1-4 Ordine Materialiarum dispositus
1-5 & in III. Partes digestus
1-6 Cum Indice Alphabetico.
1-7 Tomus I.

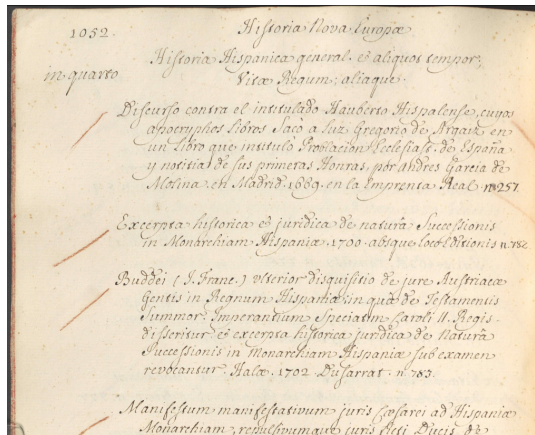
Handschriftenerkennung mit Transkribus 2/4

- ▶ 3 HTR-Modelle wurden trainiert, mit unterschiedlicher Performanz auf dem Validierungs-Set
- ▶ Unterschiede auch in der Performanz auf (un)regelmäßigen Seiten
- ▶ Entscheidung für Modell 49656 aufgrund der geringen Fehlerrate auf regelmäßigen Seiten



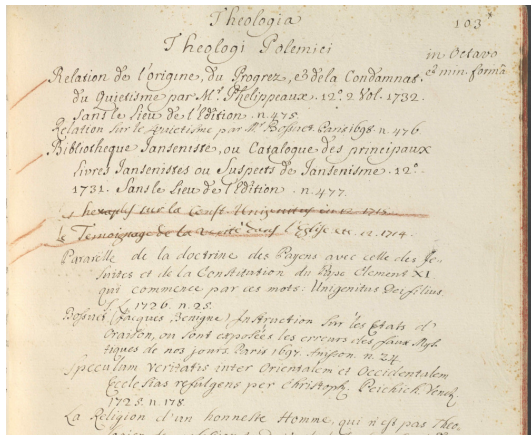
Handschriftenerkennung mit Transkribus 3/4

Regelmäßige Seite



Cod. 14.378, S. 204

Unregelmäßige Seite



Cod. 14.376, S. 119

Verhältnis 77% zu 23% zugunsten der regelmäßigen Seiten (auf 60 Probeseiten)

Umfang der geplanten digitalen Edition

- ▶ Vollständige Transkription (automatisch erstellt und manuell überprüft) der Bände 1–3 (Umfang ca. 1500 Seiten) und Markierung der folgenden Strukturmerkmale auf jeder regulären Katalogseite:
Seitenzahl, Formatsangabe, Wissensklasse, Unterkategorie der Wissensklasse, Katalogeinträge
- ▶ Katalogeinträge sollen (falls möglich) einen Verweis auf den modernen Bibliothekskatalog beinhalten
- ▶ Für die Bände 4 und 5 reduzierte digitale Edition mit automatisch erstellter Transkription geplant
- ▶ Mit Option, Bände 4 und 5 ebenso zu erschließen wie Bände 1–3 falls der Zeitplan es erlaubt

Bildklassifizierung mit Machine Learning 1/4

Verwendung von Machine Learning (ML) für Provenienzerschließung am Beispiel des Wappensupralibros von Prinz Eugen, angewendet auf den Korpus der Austrian Books Online (ABO). Somit können Bucheinbände aus der Privatbibliothek des Prinzen innerhalb des sehr umfangreichen digitalen ABO-Korpus automatisiert gefunden werden. Hier dargestellt: drei verschiedene Varianten des Wappens.



- ▶ Binärer Klassifikator entscheidet, ob das Wappen im Bild vorkommt oder nicht
- ▶ Durch Parallelisierung reduzierter Zeitaufwand
- ▶ Auswertung ermöglicht Verbesserung des Bibliothekskatalogs

Martin Krickl*, Simon Mayer und Emanuel Zangger

Mit Machine Learning auf der Suche nach Provenienzen – ein Use Case der Bildklassifikation an der Österreichischen Nationalbibliothek

<https://doi.org/10.1515/bfp-2021-0090>

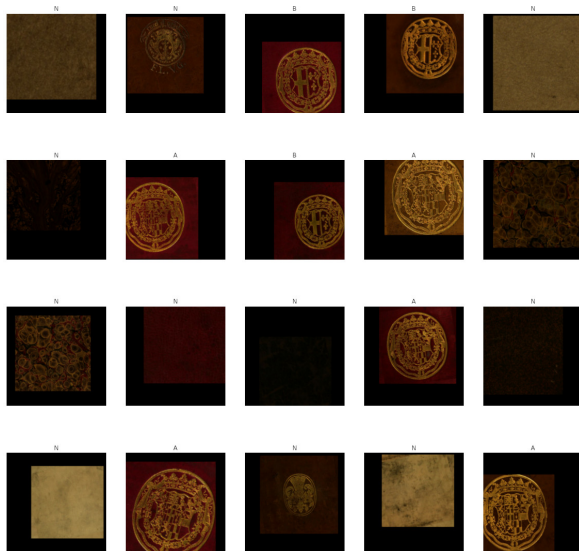
Inhalt

Zusammenfassung: Bibliotheken können in vielen Bereichen künstliche Intelligenz einsetzen. Machine Learning bietet vielversprechende Einsatzmöglichkeiten, unter anderem in der Erschließung von digitalisierten historischen Buchbeständen. Der vorliegende Artikel präsentiert eine Anwendung für die Ermittlung von Provenienzen an zwei Beispielen aus dem Bestand der Österreichischen Nationalbibliothek. Als binäre Bildklassifikatoren wurden Convolutional Neural Networks eingesetzt. Zudem soll gezeigt werden, dass mit einem einfachen Setting bereits gute

1	Künstliche Intelligenz in Bibliotheken	227
1.1	Eine sehr kurze Einführung in Machine Learning	228
1.2	Machine Learning in Bibliotheken	228
1.3	Auf der Suche nach Provenienzen in historischen Druckschriften.	230
2	Anwendungsfall Hermann Hallwich	231
2.1	Fragestellung und Ziel für Einsatz von Machine Learning.	232
2.2	Methode.	232
2.3	Ergebnisse	233

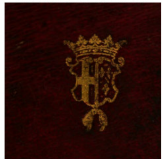
Bildklassifizierung mit Machine Learning 3/4

- ▶ Datenset: 841 Bilder, davon 673 für Training und 168 für Validierung
- ▶ Augmentation der Trainingsdaten: Rotation, Zoom, Translation
- ▶ Je nach Anpassung der Trainings-Parameter zwischen 90% und 97% Genauigkeit auf dem Validierungs-Set
- ▶ Für genaues Absuchen des Hyperparameter-Raums wird die neu erworbene GPU-Infrastruktur der ÖNB verwendet werden

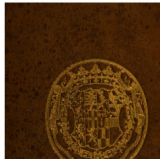


Bildklassifizierung mit Machine Learning 4/4

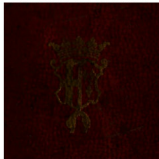
GT: C, P: N, 0.00, 0.00, 0.01, 0.99,



GT: A, P: B, 0.45, 0.54, 0.00, 0.01,



GT: C, P: N, 0.00, 0.00, 0.49, 0.51,



GT: A, P: N, 0.02, 0.00, 0.00, 0.98,



GT: A, P: N, 0.13, 0.00, 0.00, 0.87,



GT: B, P: N, 0.00, 0.04, 0.00, 0.95,



GT: B, P: A, 0.64, 0.36, 0.00, 0.00,



GT: B, P: A, 0.77, 0.23, 0.00, 0.00,



GT: A, P: N, 0.00, 0.00, 0.00, 1.00,

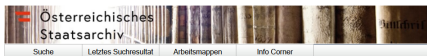


GT: C, P: N, 0.00, 0.00, 0.00, 1.00,



- Evaluation der falschen Vorhersagen auf dem Validierungs-Set
- Entsprechende Anpassung der Trainingsdaten in Folge geplant
- Weitere Anpassungen des Modells geplant (dropout Layer)

Forschung in Archivmaterialien



Startseite | Anmelden | de en fr it nl

- In die Liste wechseln
- In die Bildliste wechseln
- In die Bildübersicht wechseln
- Suche innerhalb der Resultate
- Als PDF anzeigen
- Markierte in Arbeitsmappe legen
- Hilfe

- Legende**
- Lokalisieren im Archivplan
 - Bild vorhanden, klicken zum Anzeigen

Resultatliste

Das Suchresultat enthält online verfügbare Bilder. Wechseln Sie in die [Bildliste](#) um diese anzuzeigen.

235 Datensätze gefunden

<input type="checkbox"/>	Titel/Darln	Zeitraum	Stufe	Signatur	AP	Relevanz
<input type="checkbox"/>	1501 (B) SAVOYEN-CARIGNAN Prinz von Eugen	18.10.1663 - 21.04.1736	Bestand	AT-OeStA/KA NL 1501 (B)		85
<input type="checkbox"/>	Sammlungen des Prinzen Eugen von Savoyen: Nationalbibliothek Wien	15.05.1986 - 31.10.1986	Einzelstück (Aktenstück, Bild, Karte, Urkunde)	AT-OeStA/AVA Plakatesammlung 27.43		46
<input type="checkbox"/>	Wiener Hofkriegsrat (HKR)	1557 - 1848	Bestand	AT-OeStA/KA ZSt HKR		42
<input type="checkbox"/>	Feldzüge des Prinzen Eugen von Savoyen: Beilagen	keine Angabe	Einzelstück (Aktenstück, Bild, Karte, Urkunde)	AT-OeStA/FHKA SJUS KS, S 012		35
<input type="checkbox"/>	Gutachten des Staatsarchivs an die Staatskanzlei über die Nachträge zum Manuskript Friedrich Hellers über die Korrespondenzen des Prinzen Eugen von Savoyen	27.12.1847	Einzelstück (Aktenstück, Bild, Karte, Urkunde)	AT-OeStA/HHSTA StK Interiors Archiv 2-2-101		35
<input type="checkbox"/>	Meriva, Die Stadt und Festung Meriva mit allen von der kaiserlichen Armee unter Kommando des Prinzen Eugen von Savoyen bei Meriva, Virgillara, Borgoforte und Curtatone im Jahre 1702 ausgeführten Verschanzungen	1702	Einzelstück (Aktenstück, Bild, Karte, Urkunde)	AT-OeStA/KA KPS KS H III d, 404		35
<input type="checkbox"/>	Haus-, Hof- und Staatsarchiv wird um Meinungsäußerung zum Manuskript des Rittersmeisters von Heller "Militärische Correspondenz des Prinzen Eugen von Savoyen" ersucht	1848	Akt (Sammelakt, Grundz., Konvolut, Dossier, File)	AT-OeStA/HHSTA SB KA 4/1848		35
<input type="checkbox"/>	Zwei Schreiben von Anna Viktoria von Savoyen, Nichte von Prinz Eugen, an Kaiser Karl VI. mit der Bitte um Übergabe des Erbes nach dem Prinzen Eugen	10.05.1736 - 16.05.1736	Akt (Sammelakt, Grundz., Konvolut, Dossier, File)	AT-OeStA/HHSTA HausA Familienakten 108-2-27		28
<input type="checkbox"/>	Gnadengabe für Prinz Eugen von Savoyen in der Höhe von 400.000 fl.	01.05.1713	Akt (Sammelakt, Grundz., Konvolut, Dossier, File)	AT-OeStA/FHKA AHK HFÖ Akten 1963.1		28

- Ergänzende archivalische Recherchen zum Sammlungsankauf
- Kontaktaufnahme mit Sammlungsspezialist*innen

Sammlungsvisualisierung

Dazu mehr im
Vortrag nach der
Führung ...

