

AR 3319

Christiane Fritze und Martin Krickl

Austrian Books Online – Acht Jahre Digitalisierung des historischen Buchbestandes der Österreichischen Nationalbibliothek mit Google

Zusammenfassung: Dieser Beitrag zur Digitalisierung des historischen Buchbestandes der Österreichischen Nationalbibliothek legt einen Schwerpunkt auf die Herausforderung von Logistik und Metadatenbearbeitung. Ausgehend von einer kritischen Betrachtung zur Nutzbarkeit der erstellten Daten wird der Bogen geschlagen von der aktuellen Präsentation der Digitalisate über den Book-Viewer zu bereits gängigen Forschungsmethoden der digitalen Geisteswissenschaften, die anhand des Projekts Travelogues illustriert werden.

Schlüsselwörter: Massendigitalisierung, Österreichische Nationalbibliothek, Digitale Geisteswissenschaften

Austrian Books Online – Eight Years of Digitization of the Austrian National Library's Historical Book Collection with Google

Abstract: The paper focuses on the challenges of logistics and metadata editing during the digitisation process of the historical book holdings of the Austrian National Library in partnership with Google. Based on a critical examination of the usability of the data created, the current presentation of the digital facsimiles is discussed in the light of already common research methods of the digital humanities, illustrated by the Travelogues project.

Keywords: Mass digitisation, Austrian National Library, digital humanities

1 Einleitung

Im Herbst 2019 feierte Google Books¹ sein 15-jähriges Bestehen mit einem Video² und mit einer eigens für das Jubiläum gestalteten Ausstellung³ in Google Arts and Culture des Google Cultural Institute. Waren es zunächst US-amerikanische Hochschulbibliotheken, die eine Partnerschaft mit Google eingingen, zogen bald erste europäische Bibliotheken nach, wie z.B. die Biblioteca de

¹ <https://books.google.com>.

² https://youtube.com/watch?v=zz_vG9b9dv0.

³ <https://artsandculture.google.com/project/books>.

Catalunya und die Bayerische Staatsbibliothek. Während die Bibliotheken, die in der Frühzeit der *Google Books* Initiative eine Partnerschaft mit Google eingegangen sind, die Digitalisierung der dafür vorgesehenen Bestände inzwischen weitgehend abgeschlossen haben,⁴ beginnen gerade mehrere weitere Bibliotheken Digitalisierungsprojekte im Rahmen von *Google Books*.⁵ Was im ersten Jahrzehnt des zwanzigsten Jahrhunderts noch Neuland und sowohl in Fachkreisen als auch in der breiten Öffentlichkeit nicht unumstritten war,⁶ ist nun gut zehn Jahre später fast eine Selbstverständlichkeit: eine zweistellige Anzahl an Bibliotheken arbeitet mit Google an der Digitalisierung ihrer historischen Buchbestände.

Die Österreichischen Nationalbibliothek führt seit 2011 das Projekt *Austrian Books Online* in einer öffentlich-privaten Partnerschaft mit Google durch. Erste Verhandlungen der beiden Partnerinstitutionen reichen bis in das Jahr 2008 zurück, 2010 wurde die Partnerschaft öffentlich bekannt gegeben.⁷ Nach einer sorgfältig durchgeführten Test- und Planungsphase nahm das Projekt 2011 seinen Vollbetrieb mit zeitweise 70 beteiligten Mitarbeitern seitens der Bibliothek auf. Die Rahmenbedingungen der Kooperation, die Logistik und die Metadatenbearbeitung sowie die Herausforderungen und Möglichkeiten der Qualitätssicherung haben Majewski und Kaiser (2013) ausführlich dargelegt.

Ziel der Partnerschaft war die Digitalisierung in Bild und Volltext des gesamten historischen urheberrechtsfreien Buchbestandes der Österreichischen Nationalbibliothek mit rund 600000 bibliographischen Einheiten. Mit der erfolgreichen Erreichung dieses Ziels ist die Österreichische Nationalbibliothek ihrer im Jahr 2012 entworfenen *Vision 2025*⁸ ein großes Stück näher gekommen. Dabei trug *Austrian Books Online* zu allen fünf Themenfeldern der *Vision 2025* wesentlich bei.

Allerdings ist für die Österreichischen Nationalbibliothek das Kapitel *Google Books* noch lange nicht beendet. Nach der Erreichung der ursprünglichen Projektziele wurde die Digitalisierung mit Google nunmehr in den Regelbetrieb überführt, der seitens der Österreichischen Nationalbibliothek mit Stammpersonal bewältigt wird. Fortan werden jährlich im Rahmen einer mit Google vereinbarten Jahresgrenze (*Moving Wall*) weitere urheberrechtsfreie Werke aus dem Bestand der Österreichischen Nationalbibliothek digitalisiert.

⁴ Siehe Mantels (2019).

⁵ So z.B. die Zentralbibliothek Zürich, die Zentral- und Hochschulbibliothek Luzern und die Universitätsbibliothek Bern siehe Eiholzer (2019) oder auch die Israelische Nationalbibliothek.

⁶ Mit Baileys Bibliografie (Bailey 2011) sei an die intensiv geführte Debatte der Digitalisierung von Kulturgut in öffentlicher Hand durch Google und den dabei geführten Rechtsstreitigkeiten erinnert.

⁷ Österreichische Nationalbibliothek (2010).

⁸ Rachinger (2012).

Zum anderen verfügt die Österreichische Nationalbibliothek nun über einen großen ‘Datenschatz’, den es zu pflegen, zu sichern und für die Nutzer aufzubereiten gilt. Google arbeitet permanent an der Verbesserung der Algorithmen für die Volltexterkennung sowie an der Bildbearbeitung im Post-Processing, um z.B. verzerrt aufgenommene Seiten digital zu glätten. Als *Google Books* Partner kann die Österreichische Nationalbibliothek von den Entwicklungen profitieren und zu mit Google abgestimmten Zeitpunkten neu prozessierte Daten der einmal erzeugten digitalen Faksimiles herunterladen.

2 Logistische Herausforderungen

Nachdem im Jahr 2017 die Digitalisierung der historischen, urheberrechtsfreien Druckwerke aus der Sammlung von Handschriften und Alten Drucken inklusive der im Prunksaal aufgestellten Werke abgeschlossen war, wurde mit der Digitalisierung von urheberrechtsfreien Werken aus anderen Sammlungen sowie dem modernen Magazin, in dem Ausgaben ab dem Publikationsjahr 1851 aufbewahrt werden, begonnen. Die Aufstellung im modernen Kompaktregalmagazin folgt einer anderen Logik als der im historischen Prunksaal, sodass die Logistik der Ausbebevorgänge umgestellt werden musste. Zudem wurde außerdem das Volumen der zu digitalisierenden Titel pro Monat verdoppelt, um die Digitalisierung der ursprünglich geplanten insgesamt 600.000 Bände 2018 abschließen zu können. Als zusätzliche Herausforderung neben der Erhöhung des Durchsatzes stellte sich die Logistik für die Bände aus den Sammlungen dar, die nicht im Hauptgebäude der Österreichischen Nationalbibliothek, sondern an externen Standorten untergebracht sind. Auch wenn die Entfernung mit nur ca. 500–1.000 Metern gering war, mussten diese doch dort ausgehoben und verpackt werden, um zum Haupthaus transportiert zu werden, wo sie wieder für die konservatorische Sicherung, Bearbeitung und Entlehnung ausgepackt wurden, um nach dem Abschluss der Bearbeitung wiederum für den Transport zum Digitalisierungszentrum vorbereitet zu werden.

Ein markanter Einschnitt für den Workflow war der 2018 erfolgte Umstieg vom Bibliothekssystem Aleph zu ALMA inklusive Umstieg vom Codierungsformat MAB2 auf MARC21. Dieser erforderte nicht nur Schulungen und eine grundlegende Adaption des Workflows für die Metadatenbearbeitung, sondern aufgrund der systembedingt verlängerten Bearbeitungszeit pro Band eine Aufstockung von Personal.⁹ Außerdem konnte während der vier Wochen dauernden Datenmigration vom alten in das neue Bibliothekssystem keine Bearbeitung von Katalogdaten vorgenommen werden. Auf all diese

⁹ Vgl. dazu die Beiträge zum Systemumstieg in Mitteilungen der VÖB 72 (2018), insbesondere Christian Beiler u.a.: Erschliessungsarbeit in ALMA - Erfahrungen aus dem OBV vor, während und nach der Aleph-Ablöse, S. 282-306.

Ereignisse, die wiederholt Neuplanungen notwendig machten, hat Google stets sehr kooperativ reagiert.

3 Metadatenbearbeitung – Herausforderungen, Standards, Kompromisse

Neben der Logistik war Inventarisierung und Bearbeitung bibliografischer Metadaten eine fundamentale Säule des Projektes mit relativ hohem Aufwand personeller Ressourcen. Waren zu Beginn des Jahres 2012 noch sechs Vollzeitäquivalente neben zusätzlichem Personal aus der Sammlung von Handschriften und Alten Drucken beschäftigt, so musste auf Grundlage der Anpassungen in Quantität und Frequenz der Lieferungen die Anzahl der mit Metadatenbearbeitung für das Projekt beschäftigten Mitarbeiter erhöht werden. Von Dezember 2012 bis Januar 2013 brachte die Migration der lokalen Datensätze in den Katalog des Österreichischen Bibliothekenverbundes einen ersten systemischen Eingriff in den Workflow mit sich, der eine vorübergehende Auslagerung der Metadatenkontrolle aus dem System bedingte.

Die Inventarisierung der physischen Bände mittels Einkleben spezieller, für historisches Buchgut geeigneter Barcodes, sowie die Überprüfung jedes einzelnen Bandes auf Zustand und Zusammensetzung machte eine Bearbeitung der Bände in Autopsie unabdingbar. Grundlage für die um Autopsie ergänzte Bearbeitung der bibliografischen Metadaten und Bestandsdaten waren die historischen Kataloge der Hofbibliothek. Der seit den 1850er Jahren angelegte und sukzessive erweiterte große alphabetische Zettelkatalog, der mit seiner Fülle an Verweisungen eine bis dato unvergleichliche Erschließungstiefe bot, bildet das Fundament dieser Altdaten. Von 1958 bis 1967 wurden die handschriftlichen Zettel maschinenschriftlich unter Informationsreduktion reproduziert, um damit bibliothekarischen Arbeitskatalog und parallelen Publikums katalog zu speisen.¹⁰ Ende der 1960er bis 1990er Jahre wurde eine Beschlagwortung meist ohne Autopsie ergänzt, anfänglich nach hauseigenem kontrolliertem Vokabular, späterhin in Schlagwortketten nach dem RSWK. Mit dem Wechsel vom Medium Papier auf ein elektronisches Katalogisierungssystem wurde eine Reihe von teils computergestützten, teils in seriellen intellektuellen Arbeitsschritten durchgeführten komplexen Übersetzungs- und Matchingprozessen von Information eröffnet.¹¹ Scannen der Karten, OCR-Lesung, Matching und Abtrennung von Informationseinheiten auf Grundlage von Interpunktionsregeln transformierte die „Zettelflut“ in eine Datenflut.¹² Da historisch gewachsen, erweist sich diese Datenmenge als äußerst heterogen. Zu der auf Grund der Menge nur in Teilschritten annähernd zu erarbeitenden Reduktion von Heterogenität wurde im Projekt *Austrian*

¹⁰ Hoxha et al. (1993).

¹¹ Csákány (2012).

¹² Zu den historischen Zettelkatalogen der Österreichischen Nationalbibliothek Petschar (1999).

Books Online durch Zusammenführung und Angleichung von Aufnahmen, insbesondere dem intellektuellen Einarbeiten von Verweisungen auf Grundlage von Autopsie Erhebliches beigetragen. Die autoptische Arbeit an originalen Ressourcen ermöglichte zudem eine – ebenfalls nie absolut zu erreichende – Ergänzung und Korrektur von Informationen, die in den geschilderten Transformationsprozessen reduziert¹³ oder korrumpiert worden waren und dadurch „diachron unzugänglich“ wurden.¹⁴ Insofern kann neben Datenanreicherung hier auch von Informations-Wiederherstellung gesprochen werden.

Eine Suche nach und ein Abgleich von lokalen Dubletten konnte leider nicht systematisch durchgeführt werden. Ein insbesondere für das 16. und 17. Jahrhundert zeitaufwändiger und erst mit entsprechendem Know-How durchführbarer Abgleich auf Identität oder Varianz von Ausgaben war aus Zeit- und Ressourcengründen nicht zu bewerkstelligen und bleibt damit eine künftige bibliografische Aufgabe zu einzelnen Teilbeständen.

Für die aus der Digitalisierung entstandenen digitalen Ressourcen (Digitalisate) wurden dem Vorbild der Bayerischen Staatsbibliothek folgend bewusst keine eigenen Einträge im Bibliothekskatalog erstellt, sondern als Permalinks in die MARC-Datensätze eingespielt (Kategorie 856).¹⁵ Das Digitalisat wird demgemäß nicht als genuin eigenständige Ressource angesehen, sondern als ein digitales Exemplar derselben Manifestation.

Im Zuge von *Austrian Books Online* wurde auch eine systematische Revision des betroffenen Bestandes möglich, wie sie zuletzt Mitte des 19. Jahrhunderts und bei der Umwandlung der Bestände der Hofbibliothek in öffentliches Kulturgut nach dem ersten Weltkrieg vorgenommen wurde. Viele Bestandsdaten, zumal von Fortlaufenden Ressourcen (Zeitungen und Zeitschriften), wurden autoptisch überprüft und ergänzt.

Im Zuge dieser Revisionsarbeiten wurden aus den Bänden mehrere tausend Objekte geborgen, welche als 'Fundstücke' systematisch gesammelt, versorgt und verzeichnet wurden. Die bibliothekarisch als (Buch-)Einlagen zu bezeichnenden Objekte sind wertvolle Materialien für Biografien von Buchbeständen und die Sammlungsgeschichte,¹⁶ können darüber hinaus aber auch als spezifischer, bisher nur marginal beachteter Objekttypus innerhalb der Kulturgeschichte von Buchpraxen („Historische Anthropologie des Buches“¹⁷) gedeutet werden.¹⁸

¹³ z.B. Stücktitel wie ursprünglich im Zettelkatalog vermerkt.

¹⁴ Schüller-Zwierlein (2014).

¹⁵ Gemäß der zu Beginn des Projektes noch gültigen Regeln für Alphabetische Katalogisierung (RAK) war ein digitales Faksimile eine Sekundärausgabe.

¹⁶ Schlechter (2011).

¹⁷ Ševčenko (2007).

Mit Ende der Digitalisierung und Bearbeitung der historischen Bestände waren folgende Meilensteine erreicht:

- Massendigitalisierung historischer Buchbestände und dadurch erhöhte Sichtbarkeit
- Verbesserung von Metadaten und dadurch erhöhte Auffindbarkeit
- Erstmalige Erfassung von davor über Suchanfragen nicht adressierbaren Texten (Adligate) und damit Erhöhung des im Katalog sicht- und suchbaren Bestandsvolumens
- Re-Inventarisierung der historischen Bestände und damit Erfüllung des gesetzlichen Auftrages zum Ausweis öffentlicher Kulturobjekte
- Umfassende restauratorische Maßnahmen und damit nachhaltige Sicherung von Kulturerbe.¹⁹

Acht Jahre Arbeit mögen Vieles offen gelassen haben, angesichts der Ausgangslage wurde jedoch ein solide Grundlage geschaffen. Wir sind überzeugt, dass die Heterogenität historisch überkommener Daten einerseits, die Notwendigkeit von Interpretationsleistung und Spürsinn in den beschriebenen Prozessen andererseits weiterhin unumgänglich intellektuelle Teamarbeit bedingt.

4 Statistische Auswertung

Insgesamt wurden 594.093 bibliographischen Einheiten digitalisiert. Das bedeutet, dass 190.770.690 Seiten (Bilddigitalisate) im Volltext für die Nutzer durchsuchbar sind. Freilich ist nicht auf jeder dieser Bilddigitalisate Text erkannt worden, da zum einen die Bände vom Bucheinband an gescannt wurden und zum anderen hin und wieder doppelte Scans von Vorsatzblättern angefertigt wurden. Aus der Menge der möglichen statistischen Auswertungen greifen wir uns die Daten zur Sprachverteilung heraus. Grundlage der Auswertungen sind die bibliographischen Metadaten der von Google digitalisierten Titel. Diese Auswertungen wurden auf nicht bereinigten Daten durchgeführt (siehe Abschnitt Metadatenbearbeitung), das heißt, offensichtlich falsche Angaben wurden nicht korrigiert, verschiedene für eine Sprache oder Sprachgruppe vergebene Sprachcodes wurden nicht nachträglich zusammengeführt. Die Grundgesamtheit scheint außerdem verzerrt, da mehrsprachige Titel Mehrfacheinträge in der statistischen Datenbasis erzeugt haben. Das Ergebnis der betrachteten Aspekte ändert sich durch diese Ungenauigkeiten allerdings nicht.

Wie in Abbildung 1 ersichtlich, nimmt die Menge der digitalisierten bibliographischen Einheiten mit steigendem Publikationsjahr zu, knapp 60% stammen aus dem 19. Jahrhundert.

¹⁸ Ein Beitrag von Martin Krickl unter dem Titel "Die Bucheinlage - Annäherungen an ein verdecktes Objekt" wird demnächst erscheinen in der Zeitschrift *Medium Buch - Wolfenbütteler interdisziplinäre Forschungen* 2019.

¹⁹ Wikarski (2017).

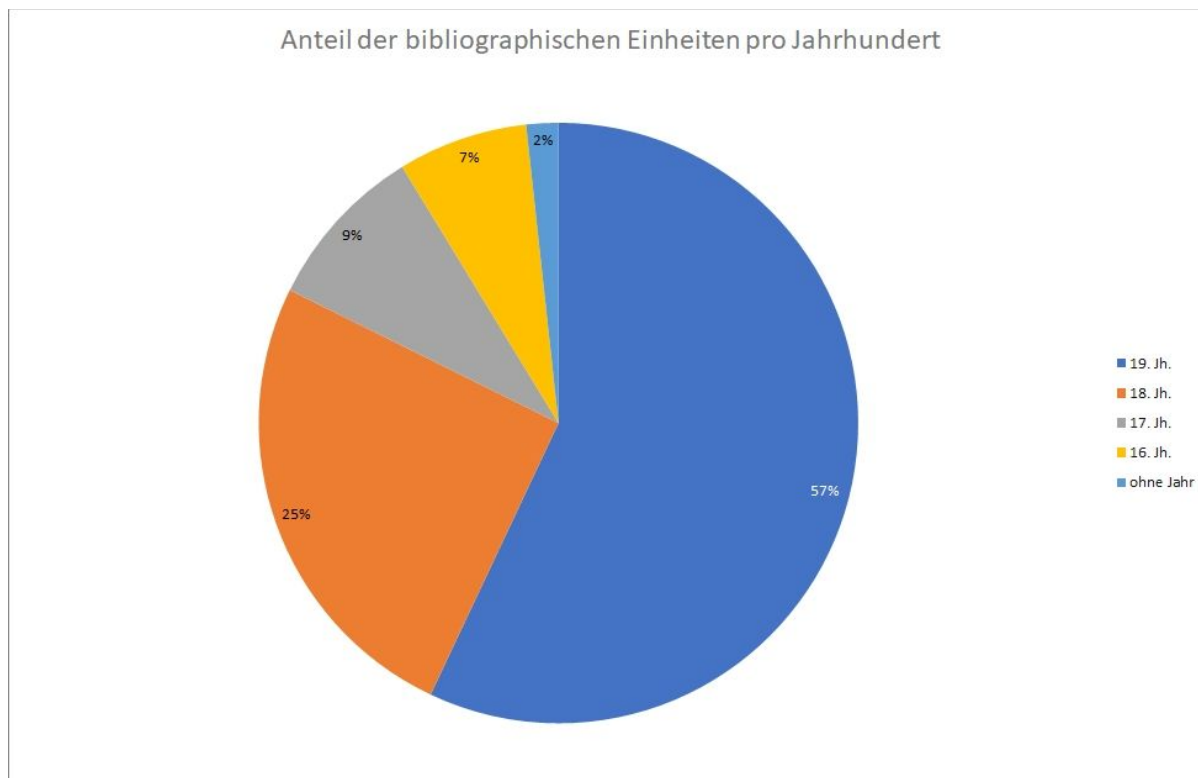


Abb. 1: Verteilung der digitalisierten Titel über die Jahrhunderte

Ebenfalls mit steigendem Publikationsjahr nimmt die Vielfalt der Sprachen zu. In Abbildung 2 zur Verteilung der Sprachen über die Jahrhunderte sind nur die zwölf Sprachen aus der Grundgesamtheit abgebildet, in denen mehr als 1.000 Titel bzw. Ausgaben digitalisiert wurden.

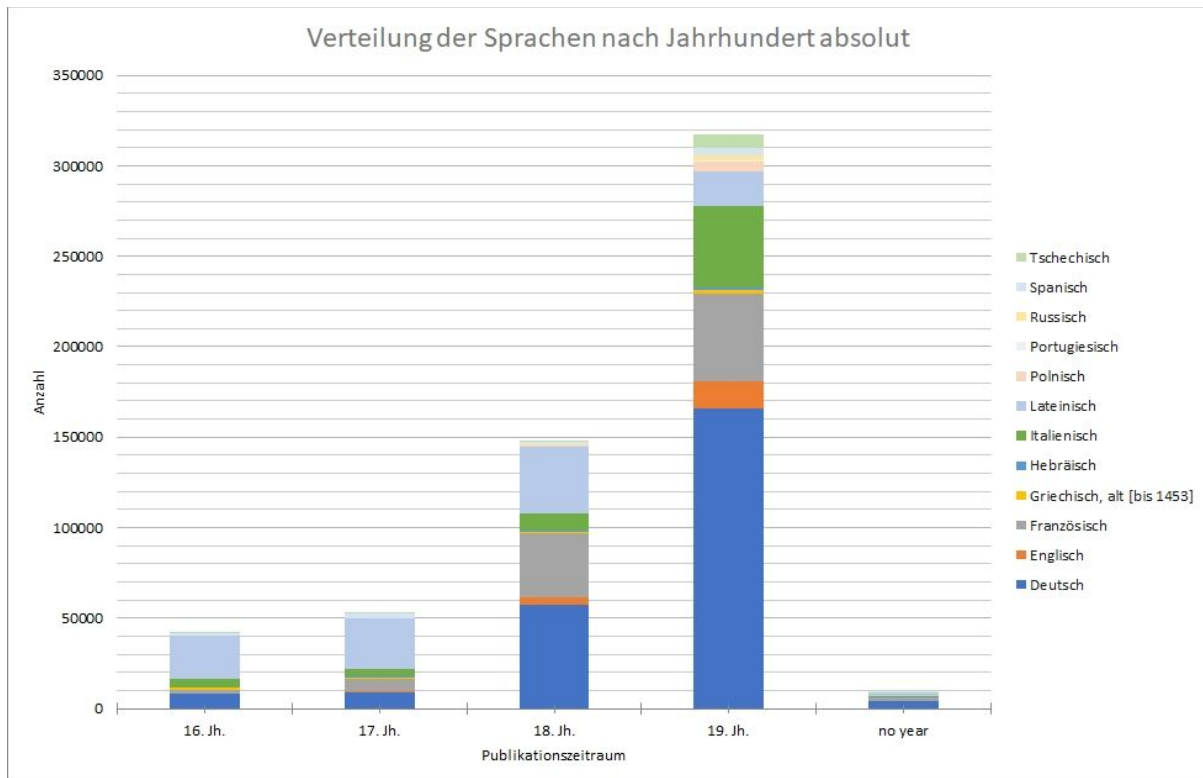


Abb. 2: Verteilung der Titel nach Sprachen über die Jahrhunderte absolut

Wie zu erwarten dominiert im 16. Jahrhundert mit knapp 60% des digitalisierten Bestandes das Lateinische, das jedoch kontinuierlich zurückgeht, sowohl prozentual als auch absolut betrachtet (siehe Abbildung 3). Hingegen nimmt die Menge an gesammelten Titeln, die in deutscher und französischer Sprache publiziert wurden, zunächst zu. Während das Französische im 19. Jahrhundert auf das Niveau des ebenfalls in allen Jahrhunderten stark präsenten Italienisch fällt, steigt die Publikation der deutschsprachigen Titel massiv an. Dass der Anteil spanischer Literatur vom 16. zum 17. Jahrhundert zunimmt und nach 1720 wiederum fällt, lässt sich zum einen aus der Geschichte des Hauses Habsburg und den Folgen des spanischen Erbfolgekrieges erklären, zum anderen aus der 1670/71 erworbenen mehrere tausend Bände umfassenden spanischen Bibliothek des Pedro de Navarra y de la Cueva. Auch der Verlauf der Titel in tschechischer Sprache ist historisch signifikant.

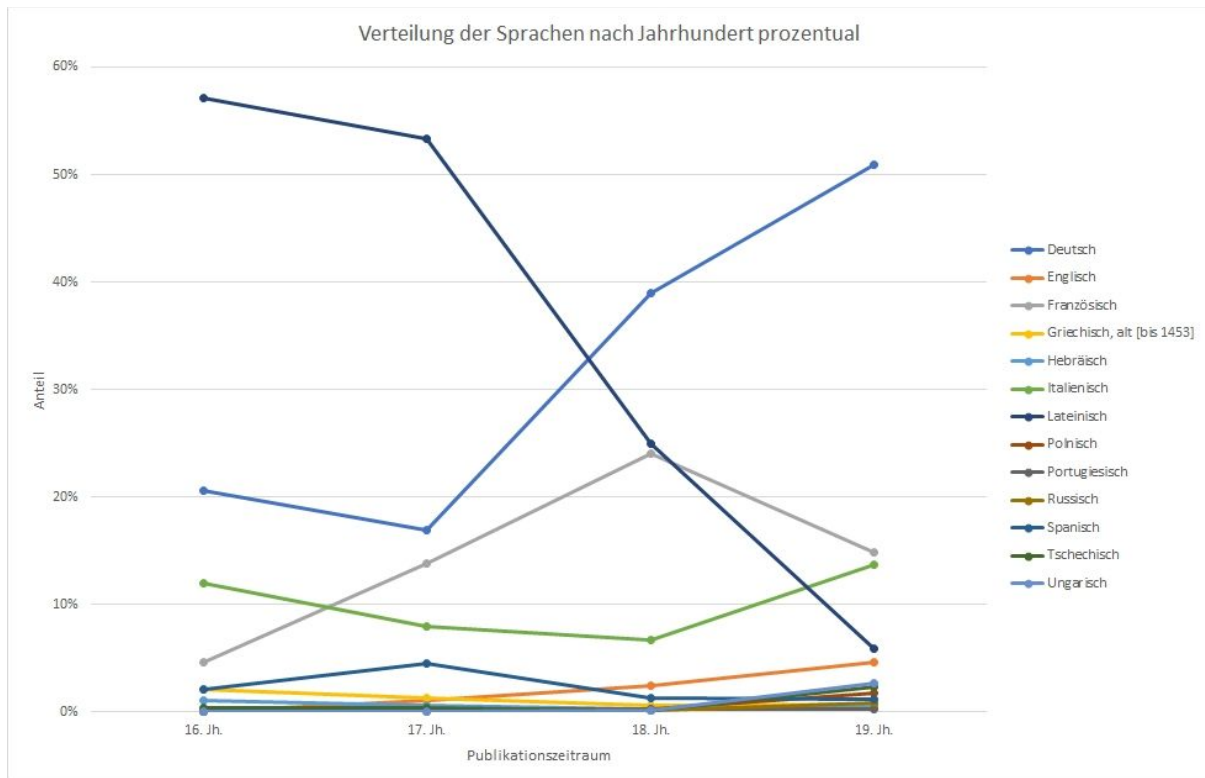


Abb. 3: Prozentuale Verteilung der Sprachen über die Jahrhunderte ohne Darstellung der Titel ohne zu ermittelndes Publikationsjahr

Abbildung 4 veranschaulicht die Verteilung der digitalisierten Publikation nach Sprachen insgesamt. Es ist zu erwarten, dass der Anteil der digitalisierten deutschsprachigen und englischsprachigen Titel mit jedem Jahr, in dem digitalisiert wird, zunehmen wird, wohingegen Latein und Altgriechischen stagnieren und der Gesamtanteil demzufolge sinken wird.

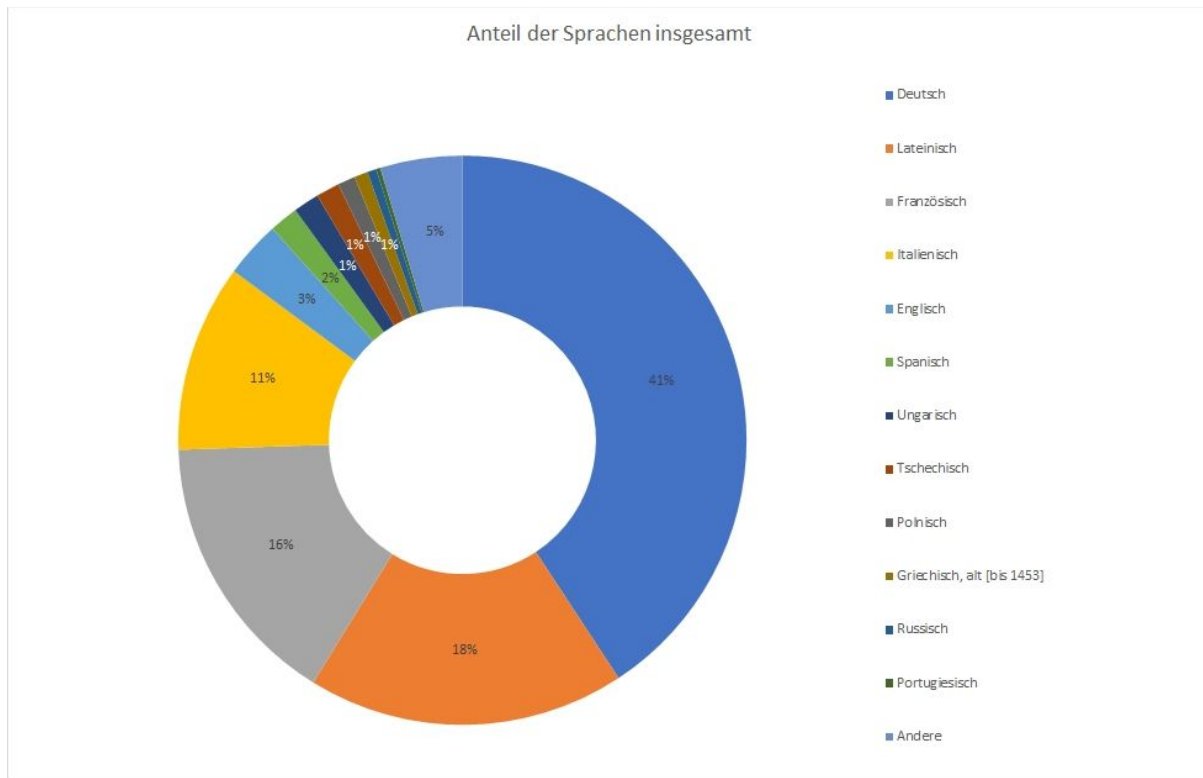


Abb. 4: Verteilung der Sprachen insgesamt.

5 Vom Nutzen der Digitalisate für die Buchforschung

Die im Zuge von *Austrian Books Online* digitalisierten historischen Druckschriften sind nicht zuletzt aufgrund der Extension des Bestandes wertvolles Material für Forschung zu und am Alten Buch. Aus der Community der historisch orientierten Buchwissenschaft, welche mehr am Buch als materiellem Objekt interessiert ist als den darin kommunizierten Inhalten, war anfänglich über die Digitalisate von Google meist wenig Positives zu vernehmen. Die Qualität und somit Vertrauenswürdigkeit der Ergebnisse sei zu gering und letztlich sei es zu wenig transparent, ob die digitale Repräsentation deckungsgleich das originale Objekt abbilde. Störelemente wie Finger von Mitarbeitern an den Digitalisierungsgeräten oder das Abschneiden von Information durch das „Croppen“ von Buchseiten sind Topoi in der Kritik geworden.²⁰ Da bei *Google Books* „kein Interesse an der Phänomenologie des Buchs, sondern nur am [...] Indexierungspotential der Texte“ bestehe,²¹ mögen die Digitalisate zwar immer wieder als unzureichend für Standards der Buchforschung disqualifiziert werden, es erstaunt aber doch, welche vielfältigen Möglichkeiten die Ergebnisse bieten und wie stark sie genutzt werden.²²

²⁰ Rössler (2016).

²¹ Stäcker (2010).

²² Gantert (2019) 246ff.

Die parallele globale Digitalisierung von historischem Buchgut hat eine gewisse Diversifikation im Angebot gebracht. Das in Quantität und Qualität sehr unterschiedliche Angebot bedient dabei unterschiedliche Klientele auf je spezifische Weise zufriedenstellend. Es bedarf stets eines kritischen Verständnisses im Sinne von James E. Dobson²³ und einer differenzierten Sicht auf den Wert von Digitalisaten, der für unterschiedliche Nutzergruppen aus unterschiedlichen Kriterien resultieren mag. Letztlich wird aber – dem wird auch von Seiten der Buchforschung beigeplichtet werden – eine Repräsentation im Digitalisat die Materialität des Originals nicht ersetzen können. Werden die in *Austrian Books Online* erstellten Digitalisate als exemplarische Repräsentation von Manifestationen genutzt, können bibliografische Referenzdatenbanken für Ausgaben wesentlich davon profitieren. So bereichern die Digitalisate seit Ende 2016 von bis dahin intellektuell im VD16 identifizierten Ausgaben wesentlich dieses kollaborative bibliografische Unternehmen.²⁴ Eine Beteiligung am *Masterplan zur Digitalisierung im VD17* – der Bayerischen Nationalbibliothek folgend, die ihre Google-Digitalisate seit Ende 2016 einbringt – wäre wünschenswert und würde die “Schmelze des barocken Eisbergs” voran treiben, zumal für den süddeutschen-österreichischen Raum viele in der Österreichischen Nationalbibliothek unikal überlieferten Drucke noch vereist liegen.²⁵ Auch für das VD18 wären die digitalisierten Bestände der Österreichische Nationalbibliothek, insbesondere der lokalen Unika, eine wesentliche Bereicherung.²⁶

6 Nutzungsszenarien: Status Quo

Von Beginn an hat Google alle digitalisierten Titel unmittelbar nach der Digitalisierung unter dem eigenen Dienst *Google Books* angeboten. 2013 hat die Österreichische Nationalbibliothek die bis dato digitalisierten Titel ebenfalls in ihre Digitale Bibliothek integriert und stellt sie fortan ihren Lesern über einen eigens entwickelten Book-Viewer zur Verfügung.²⁷ Geht *Google Books* in seinem Angebot davon aus, dass eine Suche im Volltextindex ausreicht, um für den Nutzer relevante Literatur auszugeben, setzt eine Bibliothek in erster Linie auf das Angebot einer facettierten Suche in den bibliografischen Metadaten. Eine Suche im Volltextindex über den vollständigen Bestand von *Austrian Books Online* ist ein darüber hinausgehendes zusätzliches Suchangebot. Die Einbindung der Volltextsuche in das User-Interface des Bibliothekskatalogs ermöglicht es den Lesern, auch Titel zu

²³ Dobson (2019).

²⁴ Verzeichnis der im deutschsprachigen Raum erschienen Drucke des 16. Jahrhunderts <https://www.bsb-muenchen.de/sammlungen/historische-drucke/recherche/vd-16/>.

²⁵ Verzeichnis der im deutschsprachigen Raum erschienen Drucke des 17. Jahrhunderts <http://www.vd17.de/files/vd17-masterplan-1.pdf>. Fabian (2010). Zur Entwicklung gegenüber dem VD16 vgl. v.a. Schnellling (2010) 199ff. und Labahn (2013) 14f.

²⁶ Verzeichnis der im deutschsprachigen Raum erschienen Drucke des 18. Jahrhunderts <https://b-u-b.de/digitalisierung-vd18>.

²⁷ Österreichische Nationalbibliothek (2013).

finden, die die gesuchten Wörter oder Phrasen eben nicht in den bibliografischen Metadaten enthalten und somit durch eine Suche in den üblichen Katalogfeldern nicht auffindbar wären. Allerdings wurde der maschinell erzeugte Volltext nicht nachkorrigiert und variiert in der Qualität von Titel zu Titel deutlich. Die Qualität des Volltextes hängt dabei von ganz verschiedenen Faktoren ab. Was zunächst als ergänzende Unterstützung zur klassischen facettierten Katalogrecherche gedacht war, um die Recallquote zu erhöhen, wird zunehmend kritisiert, da die Qualität des Volltextes oft für wissenschaftliche Auswertungen basierend auf dem Volltext nicht ausreicht.

Vom Katalogeintrag des gefundenen Treffers aus gelangt der Nutzer über einen Link zum Book-Viewer, in dem er das digitale Derivat des gesuchten Objekts betrachten kann. Im Hauptbereich des Viewers wird in der Voreinstellung immer ein digitales Faksimile angezeigt. Über ein Icon kann der Nutzer andere Anzeigarten auswählen: Anzeige von zwei digitalen Faksimiles nebeneinander, das die Idee eines aufgeschlagenen Buches imitiert, Vorschau aller digitalen Faksimiles für den schnellen Überblick oder die Ansicht von Bilddigitalisat und Volltext parallel zueinander. Letztgenannte Funktion ist besonders für Leser wichtig, die des Lesens von Frakturschrift unkundig sind, andere mögen sie zur Überprüfung der Qualität des Volltextes nutzen. Eine Lupenfunktion sowie ein Hineinzoomen über einen Slider ergänzen die Möglichkeiten zur Exploration des Bilddigitalisats. Neben einer gängigen Blätternavigation kann auch zu einer bestimmten Seite über die Angabe der Nummer des Digitalisats gesprungen werden. Schließlich findet sich unterhalb des verkürzt wiedergegebenen bibliographischen Hinweises die Suche im Volltext zum angezeigten Objekt. Via rechtem Mausklick auf ein Digitalisat im Hauptbereich öffnet sich ein Dialogfenster, das folgende Möglichkeiten bietet: Drehen der angezeigten Seite um 90° in beide Richtungen, der Download der angezeigten Seite im JPG-Format, Download des gesamten Titels im PDF-Format mit hinterlegtem Volltext sowie eine Kommentarfunktion. Im Zuge der Erarbeitung der strategischen Ziele 2017-2021 zur Erfüllung der Vision 2025 wird der Book-Viewer durch einen auf dem zukunftsweisenden Standard IIF²⁸ basierenden Viewer umgestellt, zu dem bereits Entwicklungsarbeiten durchgeführt werden.

7 Zukünftige Nutzungsszenarien

Lag zu Beginn des großangelegten Digitalisierungsprojekts der Fokus auf dem Angebot an die Bibliotheksbenutzer, direkt aus dem Bibliothekskatalog heraus die digitalen Faksimiles des gesuchten Titels einsehen zu können, haben sich die Anforderungen zehn Jahre später deutlich verschoben,

²⁸ International Image Interoperability Framework <https://iif.io>.

und zwar in Richtung maschineller Auswertung im Rahmen der digitalen Geisteswissenschaften – einer seit wenigen Jahren rasch aufblühenden Disziplin.

Im Zeitalter von Open Science und den digitalen Geisteswissenschaften scheint es nun nicht mehr zeitgemäß, den Lesern ausschließlich über einen vollwertigen Book-Viewer das kulturelle Erbe zugänglich zu machen. Auch die Lieferung der bibliographischen Metadaten an Europeana erweitert nur die Rechercheinstrumente um ein weiteres Einstiegsportal für die Suche. Zunehmend gewünscht ist der barrierefreie Zugang zu den Quelldaten an sich und nicht nur eine Präsentation derer in einer vorgegebenen Applikation, dem Book-Viewer. Forschern reicht es nicht mehr, nur für ihre Fragestellung relevante Titel zu finden und sie unabhängig von Zeit und Ort in Form des *closed reading* nacheinander zu konsumieren, sie wollen auf den Datenquellen im Sinne von *distant reading* selbst Forschung betreiben.²⁹

Über die LibraryLabs³⁰ der Österreichischen Nationalbibliothek werden sämtliche Katalogdaten der Österreichischen Nationalbibliothek als Linked Open Data via Schnittstellen (OAI-PMH und SRU) in den Formaten MARXML, RDF, JSON-LD, RDA und DC angeboten.³¹ Wird dieser Datenbestand der eigenen Fragestellung entsprechend abgefragt, kann die Ergebnisliste in ein Tool zur Erstellung von Datenkollektionen eingespeist werden, um so passgenaue Teilkorpora zu bilden.

Das zu einer bibliographischen Einheit der Digitalisate gehörige Datenset besteht aus verschiedenen Dateien: den Bilddateien für jede einzelne Seite im JPEG2000 Format, dem Volltext, den Textdaten mit den zugehörigen Koordinaten, die der Positionierung jedes einzelnen Wortes auf der Seite entsprechen sowie den technischen Metadaten mit Informationen über den Digitalisierungsprozess. Ist die Qualität des angebotenen Volltextes nicht ausreichend, sollte es in Zukunft möglich sein, dass die Nutzer für klar definierte Subkorpora eigene Texterkennungstools über die digitalen Faksimiles laufen lassen und gegebenenfalls einem Nachkorrekturworkflow unterziehen. Die meisten Fragestellungen werden sich voraussichtlich nur auf spezifische Teile des Datenbestandes erstrecken, andere Analysen können sinnfölig durchaus auf der Gesamtheit des Datenbestandes durchgeführt werden. Für einige Aufgaben ist der vorliegende Volltext notwendige und ausreichende Grundlage, für andere sind zunächst die Bilddigitalisate Voraussetzung. Vorstellbar ist z.B. eine automatische Extraktion aller nichttextuellen Bestandteile pro Seite, die anschließen in Buchschmuck, Tabellen und Grafiken sowie Illustrationen und Abbildungen getrennt werden, um

²⁹ Zu quantitativen Analysemethoden von großen Textkorpora u.a.: Moretti (2007) und Clement (2008) und Jockers (2013). Den Begriff *distant reading* prägte Moretti.

³⁰ Wagner (2018).

³¹ Dokumentation unter <https://labs.onb.ac.at/en/dataset/catalogue>.

mittels Verfahren, die durch Methoden des Machine Learning unterstützt werden, das Set an Illustrationen und Abbildungen zu verschlagworten. Das klingt trivialer als es ist, liegen für derlei historische Bilddaten noch kaum Trainingsdaten vor. Datenpools, die sich aus dem frei verfügbaren Bildangebot des Internets speisen, sind für auf Holzschnitten und Lithographien abgebildete Bildelemente nicht geeignet.

Für Fragestellung mit einem entweder zeitlichen oder thematisch begrenzten Fokus steht zunächst die Erstellung eines Teilkorpus aus dem Gesamtdatenbestand an. Für textbasierte Analysen wie Topic Modelling, Named Entity Recognition, Sentimentanalyse oder Stilometrie ist der angebotene, nicht nachbearbeitete OCR-Volltext qualitativ nicht ausreichend. Daher werden die Nutzer auch hier entweder den angebotenen Volltext für ihre Zwecke nachbearbeiten (s.u. wie etwa im Projekt *Travelogues*) oder anhand der Bilddateien gänzlich neu erstellen. Dass die erwartete Qualität höher sei als die durch *Google Books* erzeugte mag daran liegen, dass durch den gesetzten Fokus das Datenset weniger heterogen in Druck oder Sprache ist und die Erkennungssoftware entsprechend trainiert werden kann. Liegt der Textkorpus einmal vor, beispielsweise von allen Kochbüchern, Enzyklopädien oder Dramen, können diachrone sprachwissenschaftliche Studien zur Entwicklung von Lexik, Grammatik und Stilistik anhand des deutschsprachigen Schriftgutes im Sammelgebiet der ehemaligen Hofbibliothek angelegt werden.

8 Von der Masse zu kuratierten Kollektionen

Die Ergebnisse aus der Massendigitalisierung bedürfen somit intensiver Nachbetreuung, d.h. Investitionen von kuratorischer Arbeitsleistung, zumal – wie Paul Conway betont – “digitization for preservation” noch nicht mit “digital preservation” gleichzusetzen ist.³² Kann erstere in Massendigitalisierungsprojekten effizient abgewickelt werden, so ist letztere ein quasi infiniter kuratorischer Prozess. Über das bloße Angebot der Masse hinaus liegt nun eine wesentliche bibliothekarische Aufgabe darin, Wege durch das Dickicht der Masse zu bahnen, indem etwa Teilbestände in Form von kuratierten Kollektionen gebündelt werden. Diese bedürfen wie Spezialsammlungen intensiver Betreuung, sollten auf einer nachhaltigen Infrastruktur mit persistenten Identifiern langfristig verfügbar gehalten werden und dadurch – vor allem dann, wenn wissenschaftliche Arbeitsleistung investiert wird – klar referenzierbar sein.³³ Kuratierte digitale

³² Conway (2010).

³³ Boté (2019) 259ff.

Kollektionen erhöhen sowohl die Sichtbarkeit der Bestände³⁴ und lassen zudem auf eine „explosion of innovative uses by nontraditional users“ hoffen.³⁵

Kuratorische Prozesse können entweder von Nutzern selbst an den zur Verfügung gestellten Daten je nach Interessenslage vollzogen werden, zu einer bibliothekarischen Kernaufgabe und damit zu einem permanenten Service erhoben werden oder kooperativ zu fruchtbaren Ergebnissen führen. Wollen Nutzer die Datensets der selektierten Kollektionen selbst verarbeiten, ist aufgrund der Komplexität von Datenqualitäten und Datenformaten sowie zur Gewährleistung effizienter Workflows ein unterstützendes Service ratsam, wie es derzeit an der Österreichischen Nationalbibliothek im Rahmen der Library Labs durch die im Rahmen von CLARIAH-AT von der Österreichischen Akademie der Wissenschaften kofinanzierte Initiative „Bring your project!“ angeboten wird.³⁶ Der Bedarf dafür wächst, permanente Personalressourcen sind auf lange Sicht unvermeidbar. In zahlreichen Projekten aus dem Bereich der Digital Humanities werden wertvolle Daten aus oder zu bestehenden Bibliotheksdaten generiert oder diese verbessert, die leider in der Regel nicht wieder in die Datenbestände der Bibliotheken zurückfließen, sondern meist in stand-alone Datenbanken abgekoppelt bleiben. Kooperative Projekte, in welchen der Datenfluss nicht einer Einbahn sondern einem Win-win-Austausch gleicht, sind klar zu unterstützen. Wie laufende Erhebungen zeigen, ist für Kooperationen zwischen wissenschaftlichen Bibliotheken und Forschungsprojekten aus dem Feld der Digital Humanities noch viel Potential ungenutzt,³⁷ das über die servicerende Rolle der Institution Bibliothek in per se interdisziplinär und kooperativ angelegten Digital Humanities-Projekten hinausgeht.

9 Reiseberichte als Case Study

Referenz für ein internationales, interdisziplinäres Kooperationsprojekt aus dem Bereich der Digital Humanities, das auf die in *Austrian Books Online* generierten Daten aufbaut, ist das FWF/DFG-geförderte Projekt *Travelogues: Perceptions of the Other, 1500-1876*.³⁸ Es erhebt den Anspruch als Blaupause für ein kooperatives DH-Projekt mit embedded datalibrarian³⁹ einerseits und für die

³⁴ Tam (2017) und Overholt (2013) 16: „the easier we make it to use our collections, the more we visibly demonstrate their value.“

³⁵ Overholt (2013) 18.

³⁶ <https://labs.onb.ac.at/en/topic/bring-your-project>.

³⁷ Tenopir (2017) und Tenopir (2015).

³⁸ <http://www.travelogues-project.info>. Beteiligte Institutionen sind neben der Österreichischen Nationalbibliothek die Österreichische Akademie der Wissenschaft, das Austrian Institute of Technology, die Universitäten Salzburg und Klagenfurt, sowie das Institut L3S (Hannover).

³⁹ Vgl. Hudson-Vitale (2016), die “embedded options” von “low barrier” bis “resource intensive” skaliert. Das Projekt *Travelogues* kann der Einbettung eines Datalibrarians in Vollzeitstellung demnach als “resource intensive” option bezeichnet werden.

Nachnutzung und Nachbearbeitung von Daten aus dem *Austrian Books Online* Projekt andererseits zu dienen.

Im Projekt werden Bilddigitalisate, Volltexte und Metadaten genutzt, um an deutschsprachigen Reiseberichten der Zeit zwischen 1500 und 1876 mit Einsatz von computergestützten Verfahren Einsichten zum intermediären Genre Reisebericht zu gewinnen. In anderen large scale Projekten werden Erkenntnisse zu Texten, Textsorten oder Genres über längere Zeiträume hinweg in der Regel durch analytischen Zugriff auf Metadaten aus Bibliografien oder Bibliothekskatalogen erarbeitet.⁴⁰ In *Travelogues* werden durch Nutzung von Metadaten und Volltexten über den Einsatz von Methoden des Textminings wie u.a. Named-Entity-Recognition und Topic-Modelling Fragestellungen zu intertextuellen Bezügen und zum Alteritätsdiskurs mit Fokus auf den Orient beantwortet werden. In Kombination von Suchabfragen über die Metadaten im Bibliothekskatalog und traditionellen bibliografischen Methoden wurde ein Korpus von ca. 3.300 intellektuell kategorisierten Reiseberichten erstellt, wovon gut 90 Prozent als Digitalisate mit Volltext verfügbar sind. Zusätzlich wurde mit Verfahren aus dem Machine-Learning (unter Einsatz von Support Vector Machine, multinomial Naive Bayes und Multilayer Perceptron gemäß eines supervised learning) an den Volltexten ein Classifier für Reiseberichte trainiert. Vermittels dieser innovativen Methode konnte das Korpus um weitere 345 davor nicht klassifizierte Reiseberichte aus dem Erhebungszeitraum ergänzt werden.⁴¹

Zureichende Voraussetzung für die Performanz von Classifier und Textmining ist jedoch die Verbesserung der OCR-generierten Volltexte,⁴² der im Projekt durch die zentrale Stellung der OCR-Postcorrection Rechnung getragen wird. Greifen auf Zeichenebene bereits seit geraumer Zeit Verbesserungen, stellt die fehlerhafte Word-Segmentierung weiterhin ein verzerrendes Problem für die computergestützte Auswertung historischer (deutschsprachiger) Druckschriften dar. Auf der Basis eines Neural-Network-Ansatzes werden im Projekt Kombinationen von Modellen eingesetzt, zumal davon ausgegangen wird, dass derzeit kein singuläres Modell die Komplexität des Problems ausreichend zu adressieren vermag.

Dieses Korpus wird als eine kuratierte digitale Kollektion mit sämtlichen Daten, Metadaten, Annotationen und Skripten für Forschungszwecke gemäß der Open Science Maxime frei zugänglich, durchsuchbar und weiter verarbeitbar sein.

⁴⁰ Tuppen (2016) und Belgum (2018) und Lahti (2015)

⁴¹ Der Classifier wurde ausschließlich an einer binären Klassifikation trainiert (travelogue / non travelogue) und gibt als Ergebnis Wahrscheinlichkeitswerte für eine Zuweisung aus, die intellektuell evaluiert wurden. Eine generische Verwertbarkeit des Classifiers wird getestet.

⁴² Vgl. dazu auch die laufenden Arbeiten aus den Projekten OCR-D und NewsEye. Zum Einfluss der OCR auf Topic Modelling etwa Mutuvi et al. (2018).

10 Fazit

Die Mühen haben sich gelohnt. Mit der Digitalisierung des historischen Buchbestandes der Österreichischen Nationalbibliothek ist ein wichtiger Grundstein für die Zukunft mit computergestützten Analysen und Forschungsmethoden der Geisteswissenschaften gelegt. Allerdings gilt es, die vorhandene bibliothekarische Infrastruktur derart auszubauen, dass ein barrierefreier Zugang für analysierenden Zugriff durch die Forschung auf die Datensets gewährleistet ist. Dies schließt kuratierte digitale Kollektionen für den fokussierten Zugriff mit ein. Auch gehört dazu, verifizierte, durch die Forschung verbesserte Daten an die Bibliothek zurückfließen zu lassen, um sie anderen Bibliotheksnutzern zur Nachnutzung zur Verfügung zu stellen.

Literaturverzeichnis

- Bailey, Charles W. Jr. (2011): Google Books Bibliography, 2011. Verfügbar unter <http://www.digital-scholarship.org/gbsb/gbsb.htm>.
- Belgum, Kirsten; Handley, Keith; Bott, Rachel (2018): Mapping Travel Writing: A Digital Humanities Project to Visualise Change in Nineteenth-Century Published Travel Texts. In: *Studies in Travel Writing*, 22 (3), 306-324. Verfügbar unter doi.org/10.1080/13645145.2019.1575765.
- Boté, Juan-José (2019): Dataset Management as a Special Collection. In: *Collection Management*, 44 (2-4), 259-276. Verfügbar unter doi.org/10.1080/01462679.2019.1586613.
- Clement, Tanya; Steger, Sara; Unsworth, John; Uszkalo, Kirsten (2008): How Not To Read a Million books. Verfügbar unter <https://web.archive.org/web/20191101221309/http://www.people.virginia.edu/~jmu2m/hownot2read.html>.
- Conway, Paul (2010): Preservation in the Age of Google: Digitization, Digital Preservation, and Dilemmas. In: *The Library Quarterly* 80 (1), 61-79. Verfügbar unter doi.org/10.1086/648463.
- Csákány, Bianca (2012): Vom Zettelkatalog zum Volltext: über die Entwicklung und Funktion des Kataloges am Beispiel der Österreichischen Nationalbibliothek. Master Thesis im Rahmen des Universitätslehrgangs Library and Informations Studies MSc. Wien.
- Dobson, James E. (2019): *Critical Digital Humanities: the Search for a Methodology*. Urbana u.a.: University of Illinois Press.
- Eiholzer, Stefan (2019): Google Books spannt mit Schweizer Bibliotheken zusammen. SRF (blog), 04.08 2019. Verfügbar unter <https://www.srf.ch/news/schweiz/schnelleres-digitalisieren-google-books-spannt-mit-schweizer-bibliotheken-zusammen>.
- Gantert, Klaus (2019): Handschriften, Inkunabeln, alte Drucke – Informationsressourcen zu historischen Bibliotheksbeständen. Boston u.a.: De Gruyter (Bibliotheks- und Informationspraxis: 60).
- Google Arts and Culture (Hrsg.) (2019): Highlights from Google Books. A handful of stories from the Google Books archive. Exhibition. Verfügbar unter <https://artsandculture.google.com/project/books>.
- Hoxha, Monika; Mersich, Brigitte; Petschar, Hans; Thiel, Johanna (1993): Der Alte Katalog der ÖNB. Vergangenheit – Gegenwart – Zukunft. In: *Altes Buch und Neue Medien*, 370–77. Wien: Vereinigung Österr. Bibliothekarinnen u. Bibliothekare.

- Hudson-Vitale, Cynthia (2016): Embedded Options: A Common Framework. In: *Databrarianship. The Academic Data Librarian in Theory and Practice*, hg. v. Lynda Kellam und Kristi Thompson, 25–34. Chicago: Association of College and Research Libraries.
- Jockers, Matthew L. (2013): *Macroanalysis: Digital Methods & Literary History*. Urbana, Illinois u.a.: University of Illinois Press.
- Kaiser, Max (2011): 600000 Bücher in sechs Jahren. buchreport (blog), 3. Mai 2011, Verfügbar unter <https://web.archive.org/save/https://www.buchreport.de/news/600000-buecher-in-sechs-jahren>.
- Kaiser, Max; Majewski, Stefan (2013): Austrian Books Online: Die Public Private Partnership der Österreichischen Nationalbibliothek mit Google. In: *BIBLIOTHEK – Forschung und Praxis*, 37 (2), 197–208. Verfügbar unter doi.org/10.1515/bfp-2013-0020.
- Labahn, Karsten (2013): Retrodigitalisierung kultureller Überlieferung als kooperatives Vorhaben. Die Digitalisierung von Drucken des 17. Jahrhunderts im Rahmen des VD17-Masterplans. In: *Speicher des Wissens. 444 Jahre Universitätsbibliothek Rostock*, hrsg. von Ulrich Vetter, 14–15. Rostock: Universitätsverlag.
- Lahti, Leo; Ilomäki, Niko; Tolonen, Mikko (2015): A Quantitative Study of History in the English Short-Title Catalogue (ESTC), 1470-1800. In: *LIBER Quarterly*, 25 (2), 87–116.
- Mantels, Ruben (2019): The Book Tower's Long Tail. On the Deal between Google Books and Ghent University Library. The Low Countries (blog), 25. Juni 2019. Verfügbar unter <https://www.the-low-countries.com/article/the-book-towers-long-tail#>.
- Moretti, Franco (2007): *Graphs, Maps, Trees*. London, New York: Verso.
- Mutuvi, Stephen; Doucet, Antoine; Odeo, Moses; Jatowt, Adam (2018): Evaluating the Impact of OCR Errors on Topic Modeling. In: *Maturity and Innovation in Digital Libraries*, hrsg. von Milena Dobрева, Annika Hinze, und Maja Žumer. Cham: Springer, 3-14. Verfügbar unter https://doi.org/10.1007/978-3-030-04257-8_1.
- Österreichische Nationalbibliothek (2013): Österreichische Nationalbibliothek stellt die ersten 100.000 Bücher ins Netz. Verfügbar unter https://web.archive.org/web/20160616193651/http://www.onb.ac.at/services/presse_21764.htm.
- Österreichische Nationalbibliothek (2010). Österreichischen Nationalbibliothek macht ihren gesamten historischen buchbestand online zugänglich. Pressemitteilung. Verfügbar unter <https://webarchiv.onb.ac.at/web/20100825171711/http://www.onb.ac.at/austrianbooksonline/presseinfo.htm>.
- Overholt, John (2013): Five Theses on the Future of Special Collections. In: *RBM: A Journal of Rare Books, Manuscripts, and Cultural Heritage* 14 (1), 15–20. Verfügbar unter doi.org/10.5860/rbm.14.1.391.
- Petschar, Hans; Strouhal, Ernst; Zobernig, Heimo (1999): *Der Zettelkatalog: ein historisches System geistiger Ordnung*. Wien u.a.: Springer.
- Rosenberg, Scott (2017): How Google Book Search Got Lost. Wired Backchannel (blog), 11. April 2017. Verfügbar unter <https://www.wired.com/2017/04/how-google-book-search-got-lost>.
- Rössler, Hole (2016): Google sichtbare Hände. Das Retrodigitalisat als Ware. In: *Zeitschrift für Ideengeschichte*, 10 (2), 115–25.
- Schlechter, Armin (2011): Textträger, archäologisches Objekt und historischer Mosaikstein. Was bleibt vom Alten Buch. In: *Das Ende der Bibliothek? Vom Wert des Analogon*, hg. v. Uwe Jochum und Armin Schlechter, 101–14. Frankfurt am Main: Klostermann.
- Schnelling, Heiner (2010): Die Verzeichnisse der im deutschen Sprachraum erschienenen Drucke des 16., 17. und 18. Jahrhunderts. Kontinuität und Innovation. In: *Schmelze des barocken Eisbergs?*

- Das VD 17 – Bilanz und Ausblick*. Beiträge des Symposiums in der Bayerischen Staatsbibliothek München am 27. und 28. Oktober 2009, hg. v. Claudia Fabian u.a., 199–214. Wiesbaden: Harrassowitz (Bibliothek und Wissenschaft: 43).
- Schüller-Zwierlein, André (2014): Diachrone Unzugänglichkeit: Versuch einer Prozesstypologie. In: *Diachrone Zugänglichkeit als Prozess: kulturelle Überlieferung in systematischer Sicht. Age of Access?* hg. v. Michael Hollmann und André Schüller-Zwierlein, 15–79. Boston: De Gruyter Saur (Grundfragen der Informationsgesellschaft: 4).
- Ševčenko, Nadežda N. (2007): Eine historische Anthropologie des Buches: Bücher in der preußischen Herzogsfamilie zur Zeit der Reformation. Göttingen: Vandenhoeck & Ruprecht.
- Stäcker, Thomas (2010): Digitalisierung buchhistorischer Quellen, Fachportale und buchhistorische Forschung jenseits der Gutenberggalaxie. In: *Buchwissenschaft in Deutschland*, hg. v. Ursula Rautenberg u.a., 711–33. Berlin: De Gruyter.
- Tam, Marcella (2017): Improving Access and “Unhiding” the Special Collections. In: *The Serials Librarian* 73 (2), 179–85. Verfügbar unter doi.org/10.1080/0361526X.2017.1329178.
- Taycher, Leonid (2010): Books of the world, stand up and be counted! all 129,864,880 of you. Inside Google Books Search (blog), 5. August 2010. Verfügbar unter <http://booksearch.blogspot.com/2010/08/books-of-world-stand-up-and-be-counted.html>.
- Tenopir, Carol u.a. (2017): Research Data Services in European Academic Research Libraries. In: *LIBER QUARTERLY*, 27 (1), 23–44. Verfügbar unter doi.org/10.18352/lq.10180.
- Tenopir, Carol u.a. (2015): Research Data Services in Academic Libraries: Data Intensive Roles for the Future? In: *Journal of eScience Librarianship*, 4 (2). Verfügbar unter doi.org/10.7191/jeslib.2015.1085.
- Tuppen, Sandra; Rose, Stephen; Drosopoulou, Loukia (2016): Library Catalogue Records as a Research Resource: Introducing ‘A Big Data History of Music’. In: *Fontes Artis Musicae*, 63 (2), 67–88. Verfügbar unter doi.org/10.1353/fam.2016.0011.
- Wagner, Sophie-Carolin (2018): Die Bibliothek als Laboratorium. Forschungsblog der Österreichischen Nationalbibliothek (blog), 29. November 2018. Verfügbar unter <https://www.onb.ac.at/forschung/forschungsblog/artikel/die-bibliothek-als-laboratorium>.
- Weisl, Annabella (2010): Unlocking our shared cultural heritage: Google partners with the Austrian National Library. Google Europe Blog (blog), 15. Juni 2010. Verfügbar unter https://europe.googleblog.com/2010/06/unlocking-our-shared-cultural-heritage_4958.html.
- Wikarski, Julia (2017): Kulturerbe für die Zukunft sichern. Forschungsblog der Österreichischen Nationalbibliothek (blog), 3. August 2017. Verfügbar unter <https://www.onb.ac.at/forschung/forschungsblog/artikel/kulturerbe-fuer-die-zukunft-sichern>.



Christiane Fritze
Hauptabteilung Digitale Bibliothek

Abt. für Forschung und Entwicklung
Österreichische Nationalbibliothek
Josefsplatz 1
A-1015 Wien
Österreich
christiane.fritze@onb.ac.at



Martin Krickl
Hauptabteilung Digitale Bibliothek
Abt. für Forschung und Entwicklung
Österreichische Nationalbibliothek
Josefsplatz 1
A-1015 Wien
Österreich
martin.krickl@onb.ac.at