

Christian Lüthi, Daniel Tschirren

Massendigitalisierung in Schweizer Universitätsbibliotheken

Vier Bibliotheken in Bern, Luzern, Basel und Zürich digitalisieren mehrere 100 000 Bücher in Zusammenarbeit mit Google

Die Universitätsbibliotheken Bern und Basel, die Zentral- und Hochschulbibliothek Luzern sowie die Zentralbibliothek Zürich digitalisieren seit 2019 in Zusammenarbeit mit Google einen großen Teil ihres historischen Buchbestandes. Die Universitätsbibliothek Bern schloss ihr Projekt mit 84 000 digitalisierten Titeln im September 2022 ab, die Zentral- und Hochschulbibliothek Luzern folgte mit über 70 000 Titeln im November 2022. Trotz Unterbrüchen während des Pandemie-Lockdowns konnten die beiden Bibliotheken ihre Projekte erfolgreich durchführen und bieten ihren Kundinnen und Kunden nun einen einfacheren Zugang zu ihrem Bestand aus dem 18. und 19. Jahrhundert. 2023 starten Basel und Zürich mit ihren Anschlussprojekten.

Im Juli 2018 kontaktierte ein Google-Mitarbeiter die Universitätsbibliothek Bern (UB Bern) und fragte, ob diese an einer Zusammenarbeit im Google-Books-Programm interessiert wäre. Google betreibt seit 2007 in München ein Scan-Center, wo die historischen Bestände der Bayerischen Staatsbibliothek und der Österreichischen Nationalbibliothek gescannt wurden.¹ Diese beiden großen Projekte standen vor dem Abschluss, weshalb Google weitere Partner in Europa suchte. Die UB Bern war sehr interessiert an einer Kooperation. Der Ablauf eines solchen Projektes war seit der Zusammenarbeit von Google mit der Bibliothèque cantonale et universitaire in Lausanne in den Jahren 2007 bis 2009 bekannt.² Die UB Bern empfahl Google, auch die ZHB Luzern, die UB Basel, die ZB Zürich und die ETH-Bibliothek Zürich für eine Zusammenarbeit anzufragen. Mit Ausnahme der ETH-Bibliothek, die den größten Teil ihrer historischen Sammlung schon auf e-rara.ch digitalisiert hatte, waren alle Bibliotheken an der Kooperation interessiert.

Um überhaupt mit Google ins Geschäft zu kommen, musste zuerst ein Non-Disclosure-Agreement unterzeichnet werden. Damit verpflichteten sich die Bibliotheken im September 2018, bis zum Abschluss des Vertrages für das Projekt die Verhandlungen und Gespräche höchst vertraulich zu behandeln, und nur die absolut notwendigsten Mitarbeitenden in die Pläne einzuweihen.

Bevor im Juli 2019 der Zusammenarbeitsvertrag mit Google unterzeichnet werden konnte, waren umfangreiche Abklärungen notwendig. So mussten die Rektorate der Universitäten und weitere Stakeholder – so zum Beispiel die Korporation Luzern als Besitzerin der ehemaligen Bürgerbibliothek und die

Bürgergemeinde Bern – dem Vorhaben zustimmen. Mit dem Vertrag verpflichtete sich jede der vier Bibliotheken, mindestens 60 000 Bände Monografien aus dem Publikationszeitraum 1700 bis 1900 für das Scanning durch Google zur Verfügung zu stellen. Google übernahm sämtliche Kosten für die Transporte und das Scanning. Den Aufwand für die Bereitstellung und das Zurückstellen der Bücher vor Ort mussten die Bibliotheken bezahlen. Ende Juni 2019 informierten die Bibliotheksdirektionen ihre Mitarbeitenden und Mitte Juli die Medien über die Kooperation.³

Von der »Candidate List« zur operativen Umsetzung

Bis 2019 hatte Google schon rund sieben Millionen Dokumente mit Erscheinungsjahr vor 1900 für seine Plattform Google Books digitalisiert. Deshalb musste zuerst geklärt werden, welche Publikationen aus den Schweizer Beständen noch nicht bei Google Books zugänglich sind.

Mit dem Vertrag verpflichtete sich jede der vier Bibliotheken, mindestens 60 000 Bände Monografien aus dem Publikationszeitraum 1700 bis 1900 für das Scanning durch Google zur Verfügung zu stellen.

Die UB Basel exportierte dazu sämtliche Titel mit Erscheinungsjahr 1700 bis 1900 der beteiligten Schweizer Bibliotheken aus der gemeinsamen Katalogdatenbank Swissbib, filterte Titel heraus, die bereits auf der Plattform e-rara digitalisiert verfügbar waren, und schickte den Datensatz an Google. Dort glichen Datenspezialisten die Liste mit bereits gescannten Titeln auf Google Books ab und lieferten Anfang 2019 die sogenannte Candidate List zurück. Diese umfasste 896 000 Titel, ohne Doppel Exemplare 558 000 Titel. Den Partnerbibliotheken von Google stand es frei, Spezialbestände auszusondern und nicht ins Scan-Center zu liefern. Bei der UB Bern war das zum Beispiel der Spezialbestand Sammlung Lauterburg, der in einem Projekt ab 2023 auf e-rara digitalisiert werden soll. Zudem befanden sich viele Broschüren mit wenigen Seiten auf der Candidate Liste (Var-Signaturen). Diese wurden auch von der Liste gestrichen.

Auch die ZHB Luzern schied bestimmte Bestände von vornherein aus, welche nicht in das Projekt Google Books aufgenommen wurden. Dies betrifft den gesamten Tresorbestand der Sondersammlung, der unter anderem Alte Drucke des 15. bis 17. Jahrhunderts umfasst, sowie analog zur UB Bern Konvolute, die in der Mehrheit aus zusammengebundenen Kleindrucken bestehen.

Zur Vorbereitung war es einer Gruppe von knapp 20 Mitarbeitenden aus allen vier beteiligten Bibliotheken möglich, im Oktober 2019 exklusiv das Scan-Center von Google in München zu besuchen, um so einen Eindruck von der Arbeitsweise beim Scanning zu erhalten. In München fand anschließend ein Treffen mit der Projektleitung des Google Projektes der Bayerischen Staatsbibliothek statt. Dort erhielten die Teilnehmenden wertvolle Informationen und Tipps für die Projektdurchführung.

Anschließend entschieden sich die vier Bibliotheken, die Workflow-Datenbank der Firma ImageWare, die diese für das Projekt der Bayerischen Staatsbibliothek entwickelt hatte, in angepasster Form zu übernehmen. Damit ließen sich die Auswahl der Bücher und der Lieferprozess effizient steuern und abwickeln.⁴

5 000 Bände pro Monat digitalisiert

Google stellt für ihre Bibliothekspartner diverse verschriftlichte Informationen für das operative Geschäft zur Verfügung (Onboarding Guide etc.). Darin enthalten sind formale Kriterien wie der benötigte Öffnungswinkel für den Scanprozess, der Zustand der Einbände oder die minimalen und maximalen Dimensionen eines Buches. Diese formalen Kriterien bildeten eine erste Grundlage zur Sichtung und Sperrung weiterer Signaturgruppen in den historischen Beständen.

Nebst den formalen Kriterien konnten Bücher auch aufgrund ihres konservatorischen Zustands vom Scanprozess ausgenommen werden. Hierfür begann die konservatorische Vorprüfung sämtlicher infrage kommender Titel ab Februar 2020 und dauerte bis Juni 2022. Nebst der Vorprüfung fand nach erfolgter Lieferung auch eine stichprobenweise Nachkontrolle der Bücher auf Schäden statt.

Bei diesem Massendigitalisierungsprojekt mussten monatlich 5 000 Bände ins Scan-Center von Google geliefert werden, um die erforderliche Menge Bücher innerhalb eines überschaubaren Zeitraums digitalisieren zu können. Um dieses Liefervolumen bewältigen zu können, musste zusätzliches Personal angestellt werden. Neben einer Projektleiterin waren weitere Personen im Büchermagazin tätig, um die Bücher in den Regalen zu holen, zu verbuchen und auf Bücherwagen bereitzustellen. Nach dem Rücktransport kümmerten sie sich auch um die Rückbuchung und Rückstellung vor Ort.

Für die ZHB Luzern wurde die Kooperative Speicherbibliothek Schweiz in Büron als Abholort definiert, da in diesem Außenmagazin die überwiegende Mehrheit der Alten Drucke des 18. und 19. Jahrhunderts aufbewahrt werden. Die Speicherbibliothek stellte mit ihrem Stammpersonal ohne zusätzliche Ressourcen in 14 monatlichen Lieferungen insgesamt über



Gepackter Bücherwagen, vorbereitet für die Digitalisierung durch Google. Foto: Sara Jenni

70 000 Exemplare bereit. Die Prozesse der Kommissionierung und Rückstellung der Exemplare verliefen während der gesamten Projektlaufzeit weitestgehend fehlerfrei. Lediglich in einer Lieferung wurden 47 Titel eruiert, die nicht auf dem Lieferverzeichnis erschienen und deren Metadaten nachgeliefert werden mussten.

Nebst den formalen Kriterien konnten Bücher auch aufgrund ihres konservatorischen Zustands vom Scanprozess ausgenommen werden.

Für den Transport stellte Google Transportwagen zur Verfügung. Eine Lieferung bestand jeweils aus 20 dieser Wagen. Damit die Bibliothek nicht in Zeitnot geriet, waren Reservewagen eingeplant, die bereits im Voraus für die übernächste Lieferung bereitgehalten werden konnten.

Der Transport der Lieferungen wurde durch eine Firma aus Deutschland abgewickelt. Da sämtliche Lieferungen zweimal den Zoll passieren mussten, war es Aufgabe der Bibliotheken, die benötigten Zollpapiere vorzubereiten. Pro Lieferung wurde ein sogenanntes ATA Carnet inklusive detailliertem Warenverzeichnis und benötigter Vollmacht für die Transportfirma bereitgestellt.



Mehrere Mitarbeitende waren in den Büchermagazinen tätig, um die Bücher aus den Regalen zu holen, zu verbuchen und auf Bücherwagen bereitzustellen. Foto: Sara Jenni

Im März 2020 fand das eigentliche Kickoff-Meeting mit Google statt. Angesichts des Beginns der Corona-Pandemie konnten die Google Mitarbeitenden nicht aus den USA und aus München nach Bern reisen, weshalb das Meeting in Bern mit einer Online-Verbindung ins Ausland stattfand.

Der geplante Start der Buchlieferungen ab April 2020 musste angesichts des Pandemie-Lockdowns verschoben werden. Für die Magazin-Mitarbeitenden mussten andere Arbeiten gesucht werden.

Nach dem pandemiebedingt verzögerten Start erfolgte im August 2020 eine Testlieferung von rund 40 Büchern aus Bern an Google. Diese diente dazu, die Prozesse und Schnittstellen zwischen Google und der UB Bern zu überprüfen. Nach erfolgreicher Testlieferung konnte im September 2020 eine erste vollständige Menge von 20 Bücherwagen an Google geschickt und digitalisiert werden. Es folgten noch zwei weitere Lieferungen im Oktober und November 2020, bevor die operativen Arbeiten wegen des zweiten Corona-Lockdowns eingestellt werden mussten.

Die Unterbrechung dauerte sechs Monate bis Juni 2021. Das Magazinpersonal musste erneut vorübergehend mit anderen Aufgaben beschäftigt werden. Danach konnten die

operativen Arbeiten wieder aufgenommen werden. Im September 2022 gelangte die letzte Lieferung zurück nach Bern und bildete somit den Abschluss der operativen Arbeiten.

Auch die ZHB Luzern musste den Projektstart verschieben, die erste Lieferung verließ die Kooperative Speicherbibliothek im August 2021. Die dortigen Mitarbeitenden konnten während des Projektunterbruchs ihre angestammten Tätigkeiten weiterführen.

Im Sommer 2022 programmierte die IT-Abteilung der UB Bern eine Schnittstelle, um die Links zu Google Books in Alma / Swisscovery einspielen zu können. Der Import der Daten in Alma war im November abgeschlossen.

Ergebnisse: Rund 155 000 Titel aus der Schweiz neu auf Google Books

Google digitalisierte 84 362 Titel aus dem Bestand der Universitätsbibliothek Bern sowie 70 436 Titel aus der Zentral- und Hochschulbibliothek Luzern. Der ganze Prozess war sehr gut organisiert und verlief ohne Pannen oder Verzögerungen ab. Bei weniger als einem Prozent des Bestandes traten durch die Digitalisierung kleinere Schäden am Einband, beim Signaturschild oder als kleinere Risse von Seiten auf, die einer konservatorischen Nachbearbeitung bedürfen. Google dokumentiert die Schäden minutiös. Diese waren aber nicht gravierender als dies bei Ausleihen in den Lesesaal entstehen kann.

Zu einem späteren Zeitpunkt sollen die Daten der Digital Library Copy bei Google bestellt werden, um sie in einem Langzeitarchiv zu speichern. Die UB Basel wird die Koordination dieses nachgelagerten Projektes in Zusammenarbeit mit den anderen drei UBs übernehmen. Gemäß Vertrag mit Google, können diese Daten auch auf bibliothekseigenen Plattformen öffentlich zugänglich gemacht oder anderweitig (nicht kommerziell) nachgenutzt werden.

Google digitalisierte 84 362 Titel aus dem Bestand der Universitätsbibliothek Bern sowie 70 436 Titel aus der Zentral- und Hochschulbibliothek Luzern.

Mit den Digitalisaten und den von Google extrahierten Volltexten und Strukturinformationen existiert ein wertvoller Datenbestand für die Zukunft. Dieser bietet Bibliotheken die Möglichkeit, neue innovative Angebote zur Nachnutzung und für die Forschung zu erarbeiten. Dabei müssen einerseits rechtliche Aspekte berücksichtigt, aber auch die Kosten im Auge behalten werden. So gibt es Bibliotheken, wie die BCU Lausanne, die ihre Daten nur über Google Books zugänglich machen, während andere, wie die Bayerische Staatsbibliothek, diese interoperabel über das International Image Interoperability Framework (IIIF, gesprochen »TripleAiEff«) mit einer eigenen dezentral nutzbaren Volltextsuche zur Verfügung stellen. Was hier die aktuell sinnvollste Lösung ist, soll noch im Rahmen einer Arbeitsgruppe erörtert werden. Die zuvor erwähnte Langzeitarchivierung der

Daten ist jedoch zwingend für den Fall, dass Google die gescannten Bücher einmal nicht mehr frei zugänglich halten sollte.

- 1 Wilhelm Hilpert: 10 Jahre Partnerschaft mit Google. Auswirkungen und Spuren an der Bayerischen Staatsbibliothek. In: Klaus Ceynowa, Martin Hermann (Hrsg.): Bibliotheken, Innovation aus Tradition. Berlin etc. 2015. doi/10.1515/9783110310511/html. Max Kaiser, Stefan Majewski: Austrian Books Online: Die Public Private Partnership der Österreichischen Nationalbibliothek mit Google. In: Bibliothek, Forschung und Praxis, 37(2013), S. 197–208. Christiane Fritze, Martin Krickl: Austrian Books Online – Acht Jahre Digitalisierung des Historischen Buchbestandes der Österreichischen Nationalbibliothek mit Google. In: Bibliothek, Forschung und Praxis, 44(2020), S. 89–99.
- 2 Hubert Villard: L'opération «Googlos» à la Bibliothèque cantonale et universitaire de Lausanne. In: Réseau patrimoine, no. 13(2012), S. 89–94. Peter Haber: Bücher ins Netz. Das Projekt «Google Books» fasst nun auch in der Schweiz Fuss. In: Neue Zürcher Zeitung, 9./10.6.2007, S. 50.
- 3 Der Bund, 16.7.2019, S. 21; Stefan Eiholzer: Google Books spannt mit Schweizer Bibliotheken zusammen, Radio SRF, 4.8.2019, www.srf.ch/news/schweiz/schnelleres-digitalisieren-google-books-spannt-mit-schweizer-bibliotheken-zusammen
- 4 Martin Baumgartner et al.: Zur Workflowsteuerung der Massendigitalisierung. Der Weg der Bücher und der Digitalisate. In: B.I.T. online, 11(2008), Nr. 3, S. 267–271. Martin Baumgartner, Wilhelm Hilpert: Halbzeit – ohne Pause. Stand und Erkenntnisse der industriellen Massendigitalisierung an der Bayerischen Staatsbibliothek. In: B.I.T. online, 14(2011), Nr. 2, S. 133–138.



Christian Lüthi ist Vizedirektor der Universitätsbibliothek Bern und dort unter anderem verantwortlich für die Retrodigitalisierung. Er koordinierte zudem die Zusammenarbeit der Schweizer Bibliotheken mit Google Books.

Daniel Tschirren ist stellvertretender Direktor der Zentral- und Hochschulbibliothek Luzern und leitet in Personalunion den Standort Sempacherstrasse. Er ist unter anderem verantwortlich für das Retrodigitalisierungsprojekt im Rahmen von Google Books, das die ZHB Luzern in Zusammenarbeit mit der Kooperativen Speicherbibliothek Schweiz in Büron durchführte.



ANZEIGE

BIS-C

Archiv- und Bibliotheks- InformationsSystem

- Katalogisierung
- Erwerbung
- Ausleihe
- OPAC Web GUI

Informations-
und Wissensmanagement
für Bibliotheken, Archive und
Dokumentationszentren

Offene Schnittstellen,
freie Generierbarkeit,
Standards & Normen

DABIS - Gesellschaft für Datenbank-Informationssysteme

ist ein unabhängiges Software-Haus
und seit Neugründung im Jahr 1989
mehr als 30 Jahre führender Software-Anbieter
in der deutsch- und mehrsprachigen
Bibliotheks-, Archiv- und Dokumentationswelt

DABIS GmbH

Heiligenstädterstrasse 213, A-1190 Wien

Telefon: +43 (1) 318 9777 15

Mail: support@dabis.eu

www.dabis.eu

Ihr Partner für Archiv-, Bibliotheks- und Dokumentationsysteme