

Colibri – Corpus Libri et Liberi

Digitalisierung von Kinder-
und Jugendliteratur des 19. Jahrhunderts –
ein DFG-gefördertes Projekt



Ausgaben von „Brauseköpfchen“
von Hedwig Prohl von 1899 und 1900

Im Rahmen des Projekts ‚Colibri: Digitalisierung von Kinder- und Jugendliteratur des 19. Jahrhunderts‘ bauen die Staatsbibliothek zu Berlin – Preußischer Kulturbesitz, die Universitätsbibliothek der Technischen Universität Braunschweig und die Universitätsbibliothek Bielefeld sowie die Internationale Jugendbibliothek München eine digitale Sammlung deutschsprachiger historischer Kinder- und Jugendliteratur auf. Das Projekt wird von der Deutschen Forschungsgemeinschaft (DFG) gefördert und startete im Juni 2021 mit einem virtuellen Kick-Off-Meeting.

Der Projektrahmen

Die vier oben genannten Einrichtungen und Projektpartner verfügen auf dem Gebiet der historischen Kinder- und Jugendliteratur deutschlandweit über sehr umfangreiche und wertvolle Buchbestände. Für das ‚Colibri‘-Projekt sollen aus diesen Kollektionen 15.000 bibliographische Einheiten digitalisiert werden. Der Erscheinungszeitraum der Einheiten liegt zwischen 1801 und 1914. Damit knüpft das Projekt

einerseits an die von der DFG geförderte Digitalisierungsstrategie für Drucke des 18. Jahrhunderts im Rahmen des ‚Verzeichnisses der im deutschen Sprachraum erschienenen Drucke des 18. Jahrhunderts (VD 18)‘ an; andererseits schließt es eine Lücke zu den im Projekt ‚European Collections 1914 – 1918‘ digitalisierten Kinder- und Jugendbüchern.

Sie werden in einem Online-Portal zusammenfließen, das die digitalisierten historischen Kinder- und Jugendbücher präsentiert und damit sichtbar und uneingeschränkt frei nachnutzbar macht.

Die Ergebnisse des Colibri-Projekts leisten damit einen wichtigen Beitrag zur Bildung eines Korpus zur Erforschung deutschsprachiger historischer Kinder- und Jugendliteratur des sogenannten ‚Langen 19. Jahrhunderts‘¹ mittels Methoden der ‚Digital Humanities‘. Sie werden in einem Online-Portal zusammenfließen, das die digitalisierten historischen Kinder- und Jugendbücher präsentiert und damit sichtbar und uneingeschränkt frei nachnutzbar macht. Die Digitalisate der historischen Kinder- und Jugendbücher werden mit Meta- und Strukturdaten versehen, auch sollen durch Optical-Character-Recognition-Verfahren (OCR-Verfahren) Volltexte erstellt werden, um umfassende Recherchemittel für die wissenschaftliche Analyse des Korpus bereitzustellen. Auf diese Weise sollen Forschende und

interessierte Nutzerinnen und Nutzer aus aller Welt die Möglichkeit haben, Zugang zu umfangreichen kinder- und jugendliterarischen Datenmengen zu erhalten und diese auszuwerten.

Das Projektkonsortium wird von einem neunköpfigen Beirat unterstützt, der den antragstellenden Einrichtungen in bibliothekarischer, literatur- und buchwissenschaftlicher, sprachtechnologischer und informationstechnischer Hinsicht beratend zur Seite steht. Darüber hinaus wird das gesamte Projekt mit seinem Fokus auf die Literatur- und Informationsversorgung der Wissenschaft von Literaturwissenschaftlerinnen und -wissenschaftlern der Universitäten Bielefeld und Leipzig begleitet.

Voraussetzungen und praktische Umsetzung des Colibri-Projekts

Ein solches Kooperationsprojekt von vier Bibliotheken bedarf guter Planung und Vorbereitung. Seit März 2020 tauschen sich die Kooperationspartnerinnen regelmäßig in Videokonferenzen untereinander, aber auch mit Mitgliedern des Beirats aus. Die Projektpartnerinnen legten Arbeitsschritte sowie Regeln für die Erfassung und Bearbeitung der 15.000 zu digitalisierenden Kinder- und Jugendbücher fest.



Auswahl historischer Bücher für das Colibri-Projekt



Bilderbuch-ABC von ca. 1880

Allen Arbeitsschritten des Projekts liegen die ‚Praxisregeln Digitalisierung‘ der DFG zugrunde². Die Produktion und Speicherung der kinder- und jugendliterarischen Digitalisate erfolgen in den Digitalisierungs-Workflow-Systemen Kitodo bzw. Goobi. Beide Plattformen basieren auf den XML-basierten LOC-Metadaten-Standards METS und MODS, implementieren weitgehend identische Strukturdatenregelsätze und sind beide mit dem ‚DFG-Viewer‘³ kompatibel, in dem die digitalisierten Kinder- und Jugendbücher angezeigt werden.

In das Colibri-Projekt fließen nur Buchbestände ein, die bereits nach bibliothekarischen Standards erfasst worden sind, sodass die zu digitalisierenden Bestände der vier Bibliotheken eingegrenzt, gesichtet und verglichen werden können. Damit wurde sichergestellt, dass Katalogdaten vorhanden sind, die in den jeweiligen Verbundsystemen nachgenutzt werden können.

Zu Beginn des Projekts führen die Bibliotheken in Berlin, Bielefeld, Braunschweig und München in Abstimmung zunächst eine Dublettenkontrolle der für das Projekt ausgewählten Bestände durch. Mit der Überprüfung soll sichergestellt werden, dass identische Exemplare aus den verschiedenen Einrichtungen nicht doppelt digitalisiert werden. Die ausgewählten Bestände werden weiterhin darauf geprüft, inwieweit bereits im Web frei zugängliche und qualitativ hochwertige Digitalisate vorhanden sind, sodass die entsprechenden Kinder- und Jugendbücher für das Colibri-Projekt nicht ein zweites Mal gescannt werden müssen.

Darüber hinaus wird überprüft, in welchem Umfang mit urheberrechtlich geschütztem Material bzw. Material, dessen urheberrechtliche Situation nicht eindeutig zu klären ist, gerechnet werden muss, da dieses Material für das geplante Projekt ausgeschlossen werden sollte.

An diese Vorabkontrollen schließt sich dann die Vorbereitung der Bestände an, die in mehreren Lieferungen an einen Scan-Dienstleister gehen und dort nach und nach digitalisiert werden sollen. Nach dem Ausheben der Bücher aus den jeweiligen Lagern und Magazinen müssen die Bücher auf ihren konservatorischen Zustand geprüft werden, d. h. daraufhin, ob der Zustand eine Digitalisierung überhaupt zulässt. Lose Blätter oder Buchlagen sowie ein eingeschränkter Öffnungswinkel der Bücher erschweren oder verhindern gar die Eignung.

Die voraussichtlich 5.000 Bände, die die Internationale Jugendbibliothek aus ihren großen Beständen zu dem Projekt beisteuern wird, sollen nicht nur auf dem Projekt-Portal, sondern auch in der Virtuellen Bibliothek Bayern gehostet und präsentiert werden.

Nach dem Scannen der Bücher bei den Dienstleistern werden jeweils eine Qualitätskontrolle und möglicherweise Reklamationen fehlerhafter Scans durch die Bibliotheken erfolgen. Als weitere Arbeitsschritte schließen sich die Meta- und Strukturdatenerfassung der Digitalisate sowie das Implementieren einer automatischen Texterkennung auf Basis von optischer Zeichenerkennung an.

Es folgen die Präsentation der digitalen Ausgaben mittels des DFG-Viewers sowie die Langzeitsicherung der Daten.

Zur ersten Präsentation der Digitalisate, Metadaten und Volltexte soll in den letzten Abschnitten des Projekts ein Online-Portal angelegt werden. In diesem Projekt-Portal werden die Daten aller beteiligten vier Bibliotheken zusammengeführt. Bereits jetzt verfügt das Projekt über eine eigene Website (www.colibri-portal.eu). Die voraussichtlich 5.000 Bände, die die Internationale Jugendbibliothek aus ihren großen Beständen beisteuern wird, sollen nicht nur auf dem Projekt-Portal, sondern auch in der Virtuellen Bibliothek Bayern gehostet und präsentiert werden. Ein Workshop, der die Ergebnisse der Forschenden-Community vorstellt, wird zum Abschluss des Projekts stattfinden.

Die teilnehmenden Bibliotheken freuen sich, im Zuge des Colibri-Projekts erstmals ein umfangreiches spezifisch kinder- und jugendliterarisches Text-Korpus aus dem ‚Langen 19. Jahrhundert‘ bereitstellen zu können, das im Rahmen der Digital-Humanities-Entwicklung in der deutschsprachigen Forschungslandschaft als Quelle für die historische Kinder- und Jugendliteraturforschung sowie verwandte Forschungsbereiche dienen kann.

Von Katja Wiebe und Jutta Reusch

Katja Wiebe ist die Koordinatorin des Colibri-Projekts für die Internationale Jugendbibliothek. Jutta Reusch ist Leiterin der Bibliothekarischen Dienste in der Internationalen Jugendbibliothek.



¹ Der Erscheinungszeitraum der Einheiten liegt zwischen 1801 und 1914, also in dem Zeitraum des sogenannten ‚Langen 19. Jahrhunderts‘.

² www.dfg.de/formulare/12_151/12_151_de.pdf

³ <http://dfg-viewer.de>