

KI als Partner?

Perspektiven auf Kompetenzerweiterung und Arbeitsprozesse in Spezialbibliotheken

Sebastian Nix, Daniel Schlögl

» Bibliothekstagungen ohne das Thema KI? Im Jahr 2025 kaum mehr denkbar, was verdeutlicht, welche explosionsartige Entwicklung seit der gerade einmal gut drei Jahre zurückliegenden Vorstellung von ChatGPT stattgefunden hat. Die systematische Nutzung von Large Language Models (LLMs) für die bibliothekarische Arbeit steht freilich noch am Anfang. Häufig liegt der Schwerpunkt bei fachlichen Tagungen eher auf der Nutzensicht, also z. B. darauf, wie Literaturrecherchen vereinfacht oder neuartige Interaktionen mit Volltexten ermöglicht werden können.

Der bei der 40. ASpB-Tagung am Kölner Tanzarchiv gewählte Fokus war ein anderer: Unter der bewusst mit einem Fragezeichen versehenen Überschrift „KI als Partner?“ begleitete das Thema die Teilnehmenden gleich an zwei Tagen. Die insgesamt vier Vorträge der zweigeteilten Sektion nahmen KI zunächst ganz grundsätzlich in den Blick, um dann aus verschiedenen Perspektiven Nutzungsszenarien für innerbibliothekarische Abläufe vorzustellen. Neben dem „partnerschaftlichen Potenzial“ wurden aber auch die Grenzen von KI-basierten Werkzeugen im Arbeitskontext von Bibliotheksmitarbeitenden beleuchtet.

Zwischen Glaskugel und Statistik: Grundlagen, Potenziale und Risiken von KI

Den Auftakt machte Timucin Cicek mit seinem Vortrag „Mystische Glaskugel oder einfache Mathematik? – Die Magie hinter KI“ (siehe auch <https://nbn-resolving.org/urn:nbn:de:0290-opus4-199943>). Basierend auf seiner eigenen Beschäftigung mit Natural Language Processing und Methoden des maschinellen Lernens im Rahmen eines Kooperationsprojekts mit der Deutschen Nationalbibliothek, erläuterte Cicek zunächst kurz Grundprinzipien und Wesensmerkmale von KI: Auf große Datenmengen werden statistische Verfahren zur Mustererkennung angewendet, so dass im Ergebnis kognitive Fähigkeiten von Menschen imitiert werden können. KI in diesem Sinn hat heute, so Cicek, auf breiter Front Einzug in die Lebensrealität vieler Menschen gehalten: angefangen von Fitnessstrackern und E-Mail-Spamfiltern bis hin zu Empfehlungsdiensten in Onlineshops. Anschließend legte der Referent sein Augenmerk auf Potenziale und Risiken von LLMs. Thematisiert wurden verschiedene Nutzungs-

szenarien: Brainstorming, die Erstellung von Zusammenfassungen oder auch die Unterstützung der Informationsrecherche, Letzteres kurz illustriert am Beispiel der beiden Suchmaschinen „Perplexity“ (für die allgemeine Websuche) und „Consensus“ (für wissenschaftliche Fragestellungen). Potenziale der KI-Anwendung wurden dann kontrastiert mit den dem LLM-Ansatz inhärenten Problemen: beispielsweise „verzerrte“ Abbildungen der Realität im Sinn von Stereotypen und Klischees, die sich unter anderem aus der Nutzung von inhaltlich unausgewogenen Trainingsdaten ergeben können. Ein weiterer Problembereich sind die für das Training und den Betrieb von LLMs notwendigen Ressourcen: Dies gilt sowohl für den immensen Energie- und Wasserverbrauch von Rechenzentren als auch für die intellektuelle Optimierung von LLMs mit der Arbeitskraft von Menschen aus Niedriglohnländern unter teilweise ökonomisch und ethisch fragwürdigen Rahmenbedingungen. Ein dritter Problemkomplex betrifft die kulturpessimistisch konnotierte Befürchtung, dass der Einsatz von LLMs auf lange Sicht das kritische Denken oder die intellektuelle Unabhängigkeit beeinträchtigen könnte.

Angesichts all dessen plädierte Cicek dafür, KI und insbesondere LLM-basierte Anwendungen nicht als Ersatz, sondern zur Erweiterung von Kompetenzen zu nutzen. Unerlässlich dafür: das Wissen um sinnvolle Einsatzszenarien solcher Werkzeuge sowie die Einschätzung, wo sie im Vergleich zu etablierten Anwendungen wie „konventionellen“ Suchmaschinen echten Mehrwert bieten können. Auch sei es stets angezeigt, den Output von LLMs kritisch zu hinterfragen. All dies setzt natürlich voraus, dass die Möglichkeiten und Grenzen LLM-basierter Tools bekannt sind, um ihren reflektierten Einsatz im Sinne einer tatsächlichen Kompetenzerweiterung zu ermöglichen.

Programmieren mit KI: Neue Wege zu IT-Kompetenz und Code Literacy

Um konkrete Beispiele aus dem bibliothekarischen Arbeitsalltag und der akademisch-bibliothekarischen Ausbildung ging es dann in dem Vortrag „Programmieren im Dialog: Mit KI-Chatbots zu mehr KI-Kompetenzen für den Bibliotheksalltag“ (siehe auch <https://nbn-resolving.org/urn:nbn:de:0290-opus4-199959>). Anne Christensen und Matthias Finck, beide gleichermaßen in der akade-

mischen Lehre (HU Berlin bzw. HAW Hamburg) wie in der IT-Dienstleistung (effective WEBWORK GmbH) ausgewiesen, gingen gemeinsam der Frage nach, wie die IT-basierte Lösung typischer Arbeitsaufgaben im Bibliotheksalltag durch den Einsatz KI-gestützter Chatbots unterstützt werden kann.

Unter Verweis auf jüngere Veröffentlichungen zu den fachlichen Erwartungen, die an Bibliothekarinnen und Bibliothekare gerichtet sind, formulierten die Vortragenden zunächst die Prämisse, dass grundlegende IT-Kompetenzen heute für den Arbeitsalltag in Bibliotheken unerlässlich sind, etwa ein Grundverständnis von Software und Skripten, (Meta-)Datenmanagement-Kenntnisse sowie Grundwissen zu Programmierschnittstellen. Davon ausgehend, führte Anne Christensen in das Thema ein, indem sie einen erst während der Tagung mit KI-Unterstützung entwickelten Vergleichsvierier für Tanznotationen vorstellte. Anschließend zeigte sie, welche Erfahrungen sie, ohne selbst Softwareentwicklerin zu sein, bei der Nutzung LLM-basierter Programmierassistenten für die Bewältigung konkreter bibliothekarischer Arbeitsaufgaben sammeln konnte. Die Beispiele umfassten mit der automatisierten Umwandlung/Verbesserung von Katalogdaten, der Entwicklung von Prototypen für digitale Dienste, der Einrichtung eines Webservers und der Optimierung der Suchanfragen in einem Discovery-System ein breites Spektrum. Überzeugend legte Christensen dar, wie die Nutzung LLM-basierter Werkzeuge für solche Aufgaben ihre Handlungsautonomie vergrößert und motivierend gewirkt habe. Trotz zeitaufwändiger Prompt-Erstellung konnten praktisch verwertbare Lösungen auch ohne dezidierten IT-Hintergrund rasch erzielt werden. Zudem habe sich die „Kommunikationsfähigkeit“ mit Software-Entwicklerinnen und -Entwicklern verbessert, woraus sich wiederum Lerneffekte für die eigene IT-Kompetenz ergeben hätten.

Matthias Finck präsentierte dann Erkenntnisse aus dem Einsatz des KI-gestützten Programmierassistenten „GitHub Copilot“ im Rahmen von Lehrveranstaltungen des Studiengangs „Bibliotheks- und Informationswissenschaft“ an der HAW Hamburg. Deutlich erkennbar seien die Studierenden des zweiten und vierten Semesters souveräner bei der Codegenerierung geworden, wobei sie Stärken und Schwächen des genutzten Tools durchaus wahrzunehmen und differenziert einzuschätzen vermochten. Allerdings wirkten sich vorhandene Programmierkenntnisse, die ohne KI-Unterstützung erworben worden waren, positiv auf die Differenziertheit des Urteils aus.

Beide Referierenden zogen ein positives Fazit aus ihren Erfahrungen und sahen, auch unabhängig vom aktuellen „Hype“, im Einsatz von KI-basierten Werkzeugen für bibliotheksbezogene Programmieraufga-

ben eine gute Möglichkeit zur wirksamen Erhöhung der „Code Literacy“.

Alter Bestand, neue Methoden: KI-gestützte Retrokatalogisierung mit „Human in the Loop“

Die Anwendung KI-gestützter Verfahren zur Lösung klassischer bibliothekarischer Problemlagen bildete auch die inhaltliche Klammer des zweiten Teils der KI-Sektion. Der Vortrag von Andreas Lüschoff und Frank Förster mit dem Titel „Alter Bestand neu choreographiert – Zettelkataloge mit Schwung in den OPAC“ (siehe auch <https://nbn-resolving.org/urn:nbn:de:0290-opus4-199920>) griff das Thema Katalogkonversion auf, mit dem sich im Zuge des digitalen Wandels jede Bibliothek auseinandersetzen hatte. Wie andernorts auch, hatten in den 1990er Jahren in der Bundesanstalt für Geowissenschaften und Rohstoffe (BGR) in Hannover Retrokatalogisierungsprojekte stattgefunden. Ein Teil des alten Zettelkatalogs konnte damit erfolgreich transformiert werden; ein nicht unerheblicher „Rest“ von ca. 100.000 Büchern und ebenso vielen Aufsätzen blieb aber von der digitalen Recherchierbarkeit abgekoppelt. Aufgrund der Dreiteilung des Katalogs in die Kategorien „formal“, „sachlich“ und „geographisch“ umfasst dieser 650.000 Katalogkarten. Nicht wenige Bibliotheken haben einer solchen Situation mit dem Aufbau von Image-Katalogen beizukommen versucht, die aber funktional nur unzureichend in die Gesamtumgebung integrierbar waren. Demgegenüber strebt die BGR-Bibliothek langfristig die vollständige Überführung der Titelinformationen aus dem alten Kartenkatalog in ihren OPAC an. Gemäß der Devise „Human in the Loop“ setzt das Projekt auf eine enge Verzahnung automatisierter Aktivitäten, die auf Methoden maschinellen Lernens basieren, mit menschlichen Aktivitäten. Durch die Aufteilung des Projekts in kleine Schritte soll das Gesamtprozedere transparent und steuerbar gehalten werden. Zudem können so auch Mitarbeitende an der Ablaufkontrolle mitwirken, die keine einschlägigen technischen Erfahrungen haben.

Sämtliche Prozessschritte laufen über das eigens entwickelte, modular aufgebaute Software-Tool „ALPACA“. Ganz am Anfang stand das Einscannen sämtlicher Katalogkarten, mithin die Erstellung eines Image-Katalogs. Anschließend werden die Karten anhand ihrer Farben, denen formale Funktionen (Aufsätze, Verweisungen ...) entsprechen, für die Weiterverarbeitung sortiert. Nächste Schritte sind die Differenzierung nach Sprachen sowie die darauf basierende Texterkennung. Wie entscheidend die menschliche Vorstrukturierung, Ergebniskontrolle und darauf ggf. aufbauende Verbesserungsschleifen für die Qualität sind, demonstrierte Andreas Lüschoff plastisch anhand eines OCR-Fehlers, bei dem eine durchge-

zogene Linie als sinnlose Wortfolge interpretiert wurde. Mit Blick auf das Aufwandsvolumen wichtig und nicht trivial sind auch Entscheidungskriterien und Grenzwerte, die für die Eliminierung von Dubletten definiert werden müssen. Damit einher geht die Signatur-Erkennung, die weitere Plausibilitäts- und Vollständigkeitskontrollen, die Gruppierung bzw. Verknüpfung von Aufsätzen und übergeordneten Werken sowie die intellektuelle Vorstrukturierung weiterer maschineller Schritte ermöglicht. Ähnlich komplex gestalten sich die Extraktion von Schlagwörtern und deren Abgleich und Verknüpfung mit dem zugrundeliegenden Thesaurus. Als Ziel des iterativen, sich schrittweise vorantastenden Gesamtablaufs soll am Ende eine möglichst handhabbare Restmenge für die manuelle Prüfung stehen, um Sonderfälle abdecken, verbliebene Datenprobleme bereinigen und schließlich „ein Ende finden“ zu können.

Insgesamt verdeutlicht das noch laufende ALPACA-Projekt das Potenzial und die Grenzen maschinellen Lernens. Es macht damit auch die Nahtstellen erkennbar, an denen weitere KI-gestützte Entwicklungen ansetzen könnten oder wo weiterhin auf die bibliothekarische Expertise gesetzt werden sollte. Man darf gespannt sein, welche Automatisierungsgrade beim Abgleich der bibliografischen Daten sich künftig mittels geeigneter, bibliotheks- oder fachkontextspezifischer LLMs erreichen lassen werden.

Vom Zeitfresser zum Assistenzsystem: KI pragmatisch im Bibliotheksalltag nutzen

Um das Nutzungspotenzial KI-basierter Arbeitstechniken für alltäglichere, auf den ersten Blick weniger komplexe bibliothekarische Aufgaben ging es in dem Vortrag „KI in der Spezialbibliothek: Darüber nachdenken oder einfach machen?“ (siehe auch <https://nbn-resolving.org/urn:nbn:de:0290-opus4-200326>) von Jürgen Warmbrunn (Herder-Institut Marburg) und Ivo Ulrich (Staatsbibliothek zu Berlin). Als lohnende Einsatzfelder für eine KI-Unterstützung identifizierten beide insbesondere solche lästigen Herausforderungen, die sie griffig als „Zeitfresser“, „Routinejobs“, „Stolperfallen“, „Liegenbleiber“ oder „Datenberge“ charakterisierten: Hier sei die Chance besonders groß, mittels maschineller Verarbeitung menschliche Flüchtigkeitsfehler zu vermeiden und Zeitkontingente für andere Aufgaben zu gewinnen. Seit 2024 wurden in der Bibliothek des Marburger Herder-Instituts für bislang zwei Anwendungsfälle LLM-gestützte Assistenzsysteme erfolgreich implementiert. Der „Bibliographic Metadata Structurer“ bereitet heterogen strukturierte Titellisten so auf, dass ein automatisierter Abgleich mit dem Bibliothekskatalog ermöglicht wird. Damit hilft er bei der Ablösung bzw. Vereinfachung der früher langwierigen manuellen Prüfung von Geschenk- oder Tauschangeboten. Der „ToC Structurer“ hingegen dient der Aufbereitung der in den Inhaltsverzeichnissen von Zeitschriften und Sammelbänden enthaltenen Aufsatz-Metadaten, so dass diese direkt in das Bibliothekssystem importiert werden können. Beide Assistenten leisten somit eine automatisierte Analyse heterogener bzw. unstrukturierter bibliografischer Ausgangsdaten und deren Überführung in ein für den jeweiligen Zweck weiterverarbeitbares Format. Ihnen ist ferner gemeinsam, dass sie auf das Sprachmodell GPT-4o mini zurückgreifen und sich dazu des Frameworks LangChain unter Einbeziehung des Tracers

BIS-C PLUS[®] BibliotheksInformationsSystem

BIS-C PLUS ist ein vollintegriertes BibliotheksInformationsSystem für Archive, Bibliotheken und Dokumentationszentren.

Von der Erwerbung über die Katalogisierung bis hin zur Ausleihe deckt BIS-C PLUS sämtliche Arbeitsprozesse ab.

Das modular aufgebaute System ermöglicht eine stufenlose Skalierung vom Einzelpatz bis zum Verbund.



Warum BIS-C PLUS ?

- maßgeschneiderte Bibliothekssystem-Lösung aus Österreich!
- individuell anpassbar
- offene & genormte Schnittstellen
- folgt gängigen Regelwerken, Normen & Standards
- Musiknotation & Notendarstellung
- responsiver & barrierefreier WEB-OPAC
- u. v. m. ...

Ihr PLUS an Individualität !

Gestalten Sie mit uns Ihre Zukunft. Unsere Software ist der Schlüssel zu Ihrem Fortschritt.

Wir unterstützen Sie mit langjähriger Erfahrung, Wissen und Kompetenz.

DABIS

office@dabis.eu
www.dabis.eu



LangSmith bedienen. Im produktiven Einsatz haben sich die LLM-gestützten Tools bereits bewährt; ehemals als „Zeitfresser“ geltende Aufgaben können nun in überschaubaren Minutenkontingenten abgearbeitet werden. Beide Werkzeuge sind grundsätzlich aber auch für weitere, stärker forschungsbezogene Nutzungsszenarien adaptierbar – etwa die dynamische Zusammenstellung von Spezialbibliografien durch Forschende. Ausgehend von den gewonnenen, positiven Erfahrungen hat das Entwicklerduo Warmbrunn/Ulrich weitere Anwendungsmöglichkeiten für die Bibliothek des Herder-Instituts im Visier. Neben einer Weiterentwicklung des automatisierten Metadatenabgleichs zum Zweck der Deduplizierung wird auch an eine KI-Assistenz bei der klassifikatorischen und verbalen Sacherschließung gedacht.

Fazit: KI verändert nicht den Beruf – aber die Anforderungen

Insgesamt zeigten die Vorträge eindrucksvoll, wie groß das Potenzial der KI-Unterstützung bereits jetzt ist – mit entsprechendem Prompting sowohl durch allgemein etablierte LLM-Chatwerkzeuge als auch, entsprechende Entwicklerkompetenz vorausgesetzt, durch Implementierung maßgeschneiderter Lösungen. Als Erfolgsrezept erweist sich jeweils die Zerlegung komplexer Aufgaben in handhabbare Arbeitseinheiten, um den für die Qualität unverzichtbaren „Librarians in the Loop“ die Ergebniskontrolle und Nachjustierung der Folgeschritte zu ermöglichen.

Dass die KI die Bibliothekarinnen und Bibliothekare arbeitslos machen könnte, ist somit keineswegs zu befürchten. Vielmehr werden sich KI-Technologien angesichts der weiterhin zunehmenden Komplexität von bibliothekarischen Aufgaben als unerlässlich erweisen, um mit der Entwicklung Schritt halten zu können. Ebenso dürften sich die Qualifikationen und Aufgaben in den bibliothekarischen Berufen verschieben und in neue Richtungen entwickeln. Die Integration von KI-Kompetenzen in die Ausbildungswege und die Entwicklung berufsbegleitender, niedrigschwelliger Fortbildungsangebote für „Information Professionals“ sind daher zentrale Herausforderungen. **I**



Sebastian Nix

leitet seit 2017 die Abteilung Forschungsdaten & Information am Max-Planck-Institut für Bildungsforschung in Berlin. Er verantwortet die strategische Weiterentwicklung der Informations- und Forschungsdateninfrastruktur unter besonderer Berücksichtigung von Open-Science-Aspekten. Zuvor war er in leitenden Funktionen an der Staatsbibliothek zu Berlin, am Wissenschaftszentrum Berlin für Sozialforschung und am Deutsch-Französischen Institut tätig. Nix studierte Publizistik, Politologie und Romanistik in Mainz sowie Bibliotheks- und Informationswissenschaft in Berlin. Max-Planck-Institut für Bildungsforschung nix@mpib-berlin.mpg.de



Dr. Daniel Schlögl

leitet seit 2008 die Bibliothek des Instituts für Zeitgeschichte München–Berlin (IfZ). Er war zuvor als Arbeitsstellenleiter des Akademienvorhabens „Jahresberichte für deutsche Geschichte“ der Berlin-Brandenburgischen Akademie der Wissenschaften sowie im Arbeitsbereich Digitale Fachinformation und als Bavarica-Referent an der Bayerischen Staatsbibliothek tätig. Schlögl hat Geschichte und Germanistik in München (Promotion 2001) sowie Bibliotheks- und Informationswissenschaft in Berlin studiert. Institut für Zeitgeschichte München–Berlin schloegl@ifz-muenchen.de

Tägliche News auf www.b-i-t-online.de