

Mit Discovery Tools unterwegs zu besseren Recherchemöglichkeiten?

Mit welchem Aufwand und zu welchem Preis lassen sich Discovery Tools in Bibliothekscloudmanagementsystemen implementieren

Stephan Holländer

Die Suchportale von Bibliotheken haben im Laufe der letzten 20 Jahre eine immer stärkere Reduktion auf möglichst einfache Suchmöglichkeiten im Stile des Google Frontend erfahren. Das kommt zwar den Gewohnheiten der Kundinnen und Kunden der Bibliotheken entgegen, schränkt aber auch die Möglichkeiten einer erfolgreichen Recherche in Bezug auf Präzision und Vollständigkeit der Suchergebnisse ein. Abhilfe sollen explorative Suchwerkzeuge wie beispielsweise Discovery Tools bringen. Damit diese nutzergerecht umgesetzt werden können, braucht es eine hohe Qualität der zu durchsuchenden Daten und standardisierte Funktionalitäten der Software für die Recherche, um den Erwartungen der Nutzenden zu entsprechen. Anstelle eines Discovery Tools „out of the box“ empfiehlt sich eine Ausrichtung und Individualisierung zusätzlicher Recherchertools für konkrete Anwendungszwecke und für die Bedürfnisse der verschiedenen Nutzergruppen. Um einem bei Nutzerinnen und Nutzern festgestellten Bedürfnis nach Integration des zunehmenden Anteils elektronischer Ressourcen im World Wide Web in die Suchergebnisse nachzukommen, sollten bibliografische Informationshinweise in einer „One-Stop-Shop“-Strategie angezeigt werden können. Am Anfang der Entwicklung begannen Bibliotheken mit dem Eigenbau eigener Module wie etwa dem Link Resolver von Herbert van de Sompel oder mit dem Rückgriff auf bereits im E-Commerce-Bereich vorhandene Lösungsansätze mit deren Discovery Modulen, die für die Ansprüche der Bibliotheken angepasst wurden¹. Bald entwickelten kommerzielle Anbieter jedoch bibliotheksspezifische Lösungen wie beispielsweise Primo von Ex Libris oder VuFind, das auf einer Open-Source-Software basierte. Im zweiten Jahrzehnt dieses Jahrhunderts kamen dann neu konzipierte Bibliothekssysteme mit Cloud-Anbindung auf den Markt. Man versprach sich davon in erster Linie niedrigere Anschaffungskosten, da bei diesen Bibliothek-

systemen die Hardwarekosten als Investitionskosten wegfallen. Für die Nutzenden wurde eine weiterentwickelte Discovery-Lösung angeboten, die weiteren Recherche-Komfort versprach. Entsprechend hohe Erwartungen bei Nutzerinnen und Nutzern auf das künftige Bibliothekssystem mit seinen noch zu entdeckenden Suchwerkzeugen wurden damit geweckt.

Bei wissenschaftlichen und öffentlichen Bibliotheken und Bibliotheksverbänden begann die Ablösung der Client-Server-Systeme durch Cloud-Systeme. Dabei zeigen sich deutliche Unterschiede zwischen wissenschaftlichen und öffentlichen Bibliotheken bei der Anschaffung eines Cloudsystems, was an den finanziellen Möglichkeiten der Bibliotheken liegt². Die Anschaffung eines modernen Cloudsystems ist mit einer erheblichen Investitionssumme für eine einzelne Bibliothek verbunden, was gemäß den Kenntnissen des Schreibenden zumindest in zwei Fällen zu einem nationalen Zusammenschluss von Bibliotheken geführt hat.³

Nach einem meist komplexen Migrationsprozess war dann das neue Bibliotheksmanagementsystem in der Cloud betriebs- und einsatzbereit. Die Nutzerinnen und Nutzer hatten nun Gelegenheit, die neuen Suchmöglichkeiten zu entdecken. Ihre hauptsächliche Kritik richtete sich vor allem gegen das Discovery-Tool. Einhelliger Tenor der Kritik war: Wer etwas sucht, findet nichts mehr. Das gesuchte Medium wurde im Bibliotheksmanagementsystem in der Cloud nicht angezeigt, obwohl Nutzende wie beispielsweise ein Anglistikprofessor einer hiesigen Universität wussten, dass das gesuchte Buch nur wenige Meter von ihrem Schreibtisch entfernt im Regal stand.

Nach wenigen Monaten fanden solche Kritiken den Weg in die Gespräche der Hochschulangehörigen. Eine Wochenzeitung veröffentlichte einen Artikel, der diese und andere Kritikpunkte aufnahm und einer breiteren Öffentlichkeit bekannt machte⁴.

1 Antelman Kristin, Lyneman Emily, Pace Andrew K.: Toward a 21st century catalog in: *Information Technology and Libraries*, 2006, vol. 25, n. 3, S.128-139

2 Linda Schünhoff, Benjamin Flämig, *Discovery Lösung für Öffentliche und Unternehmens-Bibliotheken*, Wiesbaden, 2018, S. 152

3 Siehe b.i.t.online Heft 6/2022, Seite 506 und b.i.t.online.Heft 3/2022, Seite 312

4 WOZ Nr. 13/2021 vom 01.04.2021.

Was war geschehen? Beim kritisierten Discovery Tool handelt es sich um ein neukonzipiertes Modul⁵, das mit künstlicher Intelligenz arbeitet. Mit ausgeklügelten Algorithmen soll die fehlende Recherchekompetenz der Nutzenden kompensiert und die Recherchegewohnheiten der Nutzenden beim Suchen analysiert werden mit dem Ziel, zusätzlich zu Informationen aus dem Medienbestand der beteiligten Mitgliedsbibliotheken auch zutreffende Informationen aus dem World Wide Web und externen lizenzierten Informationsangeboten als Links im Suchergebnis anzuzeigen. Zum Auffinden dieser externen Informationen greift das Discovery Tool auf eine interne Datenbank zurück, welche Links zu Informationsangeboten enthält, die vom Produzenten des Bibliothekssystems eingekauft und in die Datenbank eingespeist wurden. Die Software braucht aber eine gewisse Lernphase, um sich die Recherchegewohnheiten der Nutzenden zu merken. So mancher Treffer ließ die Nutzenden zunächst frohlo-

Damit breitete sich Enttäuschung unter den Nutzerinnen und Nutzern aus. Der Applaus der Nutzenden hielt sich in Grenzen, wie in Fachpublikationen vermerkt wurde⁶. Die befragten Bibliothekarinnen und Bibliothekare konnten nicht weiterhelfen. Das liegt an der Tatsache, dass viele von ihnen mit einem anderen Frontend als die Nutzenden arbeiten. Erste Aktualisierungen bei Discovery Tools brachten leichte Verbesserungen im ersten Halbjahr des Betriebes, aber das Resultat befriedigte die Nutzenden noch immer nicht.

Die Gründe dafür sind vielschichtig, wie die ersten Berichte in der Fachpresse gezeigt haben. Es zeigte sich, dass die neuen Bibliothekssysteme in der Cloud mit ihrer an die Internet-Suchmaschinen angelehnten Technologie andere Recherchestrategien nötig machen als sie für die Recherche in vorherigen Client-Server-Bibliothekssystemen mit Boole'schen Operatoren üblich und für die Nutzenden gewohnt war. Dies liegt an der Konzeption eines Discovery Tools, das auf eine interne Datenbank zurückgreift. Das Ranking der Resultate nach der Suche wird gemäß dem Rechercheverhalten der jeweiligen Nutzenden und der Ausprägung der gewählten Suchbegriffe oder auch nach Sicherheitsaspekten des Zugangs zu externen Quellen beeinflusst angezeigt⁷. Von den Nutzenden werden die generierten Rankings des Bibliothekssystems als zu wenig transparent und vielfach auch als ungenau empfunden. Die Nutzung der bisherigen Normdaten liegt beim neuen Bibliothekssystem im Vergleich zu den gemachten Erfahrungen der Nutzenden mit den bisherigen Client-Serversystemen zurück. Als ein praktisches Beispiel sei angeführt, dass die einzelnen Bände einer Serie nicht mehr einzeln am Bildschirm angezeigt werden können.

Eine erste Voraussetzung für eine erfolgreiche Recherche ist eine hohe formale und inhaltliche Erschließungsqualität der dem System zugrundeliegenden Normdateien⁸. Gerade bei der Zusammenführung ursprünglich getrennter Katalogisate in einer gemeinsamen Cloud wird eine konsequente vorherige Aufarbeitung der Katalogisate nach einer einheitlichen Katalogisierungsnorm verlangt. Dies klingt banal, ist aber gerade bei mehrsprachigen und eingekauften Katalogisaten nicht immer selbstverständlich, wie Beispiele aus der Praxis zeigen.

Ein Beispiel aus der Schweiz zeigt, dass man in einem solchen Fall mit Katalogisaten unterschiedlicher Herkunft auch anders vorgehen kann. Vor der Inbetriebnahme der kooperativen Speicherbibliothek Büron, die von mehreren Bibliotheken gemeinsam betrieben wird, wurden die

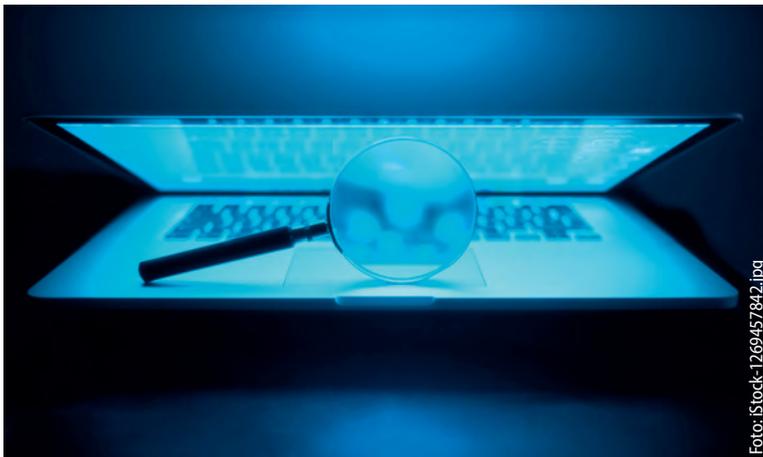


Foto: iStock-1269457842.jpg

Neue Suchtechnologien finden oft auch ein kritisches Echo bei den Nutzenden

cken, dass man gerade eine vielversprechende Ressource gefunden hätte. Leider erschien dann beim Anklicken des angezeigten Links oft die Meldung, dass die Ressource nicht durch die Mitgliedsbibliothek lizenziert sei und daher das Dokument nicht angezeigt werden könne. Hier handelt es sich um ein Phänomen, dass amerikanische Bibliotheken bereits kennen und auf ihre Art gelöst haben. Auf ihren Webseiten werden lange Listen mit den Ressourcen veröffentlicht, welche von der Mitgliedsbibliothek lizenziert sind und welche nicht. Darauf und auf die daraus resultierenden Probleme für Bibliotheken wird im Weiteren noch näher eingegangen

⁵ Auch OCLC bietet mit WMS ein Discovery System an.

⁶ Martin Blenke, Claudia Bodem, Mission possible? Vortrag gehalten am Bibliothekartag Bremen 2014. Foliensatz siehe <https://de.slideshare.net/blenkle/mission-possible-36228584>, zuletzt aufgerufen am 18.01.2023.

⁷ Dazu auch Linda Schünhoff, Benjamin Flämig, Discovery Lösung für Öffentliche und Unternehmens-Bibliotheken, Wiesbaden, 2018, S.137.

⁸ Darauf weist Ex Libris auch selbst hin https://knowledge.exlibrisgroup.com/Alma/Knowledge_Articles/Article_from_A-Z_Database_is_not_in_Primo_VE, zuletzt aufgerufen am 18.01.2023.

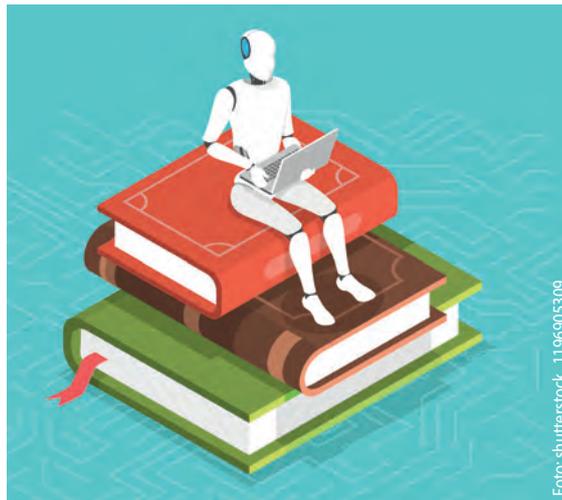
Katalogisate auf Konformität und Uniformität überprüft und harmonisiert. Das war mit einem ziemlichen Arbeits- und Kostenaufwand verbunden, aber das Resultat hat die Investition mehr als gerechtfertigt. Der Aufwand hängt aber auch vom Datenvolumen und den dem Projekt zur Verfügung stehenden Geldmitteln ab.

Neue Kompetenzen wie beispielsweise Grundlagen der Datenwissenschaft sind sowohl bei den Nutzenden wie auch bei den Bibliothekarinnen und Bibliothekaren vonnöten, um das Potential der Cloudsysteme mit deren Suchtechnologie voll nutzen zu können.

Dabei soll nicht verschwiegen werden, dass Suchmaschinen wie zum Beispiel Google mit dem Hummingbird-Algorithmus9 und dem Knowledge Graph sowie Baidu mit dem RankDex nicht nur über einen patentierten Algorithmus verfügen, den sie selbst entwickelt haben, sondern auch über eine wesentlich längere Erfahrung im Umgang mit dieser Technologie verfügen als die beiden größten Anbieter von Bibliothekssystemen.

Da Bibliotheken nicht alle Quellen lizenziert haben, die als Links in der internen Discovery-Datenbank hinterlegt sind, haben nordamerikanische Bibliotheken Listen mit den von ihnen lizenzierten Quellen aufgeschaltet. Damit weisen sie gleichzeitig auch darauf hin, welche Informationsquellen sie ihren Nutzerinnen und Nutzern empfehlen⁹. Dies ist eine nutzerfreundliche Hilfe, die auch europäische Bibliotheken übernehmen sollten, die Discovery Tools verwenden.

Beim Einsatz von Discovery Tools sehen sich Bibliotheken mit einem weiteren Phänomen konfrontiert, die es mit den bisherigen Bibliothekssystemen bisher noch nicht gegeben hat. Discovery Tools greifen auf interne Datenbanken mit bei verschiedenen Firmen erworbenen Links von verschiedenen Anbietern wie beispielsweise Verlagen zu. Das haben einige Verlagsverantwortlichen erkannt. Damit bietet sich für sie eine gute Marketingchance, das eigene Informations- und Verlagsangebot bewerben zu lassen, ohne dass der Verlag selbst als Anbieter in Erscheinung tritt. Stellvertretend unter vielen soll dazu auf ein Stelleninserat von Oxford University Press hingewiesen werden, das Anfang des Jahres mit der blumigen Berufsbezeichnung Discovery Manager erschienen ist¹⁰. Ein Schelm, wer dabei Böses denkt, aber Kollegen in angelsächsischen Ländern haben sich bereits in verschiedenen Artikeln in der Fachpresse mit dem sogenannten Tracking der Nutzerinnen und Nutzer bei der Recherche auseinandergesetzt. Bekannt ist die Vorgehensweise des Trackings auf



Helfen uns künftig Chatbots wie ChatGPT bei den Recherchen?

kommerziellen Informationsplattformen wie dem Web of Science, Scopus oder Google Scholar, wo sie aufgrund der verwendeten Technologie schon länger praktiziert wird. Neu für Mitarbeitende in Bibliotheken und für die Nutzenden ist aber jetzt, dass dies nun auch auf Bibliothekssystemen mit Discovery Tools stattfindet.

Mit dem zunehmenden Anteil an digitalen Medien im Bestand der Hochschulbibliotheken wandeln sich die verwendeten Technologien in den Bibliothekssystemen mit Discovery Tools von einem assistierenden Recherchemittel zum Nachweis und Auffinden von wissenschaftlicher Literatur zur technischen Möglichkeit des Überwachens des Recherche- und Informationsverhaltens der Bibliotheksnutzerinnen und -nutzer, wie eine amerikanische Publikation nachweist¹¹. Diese Vorgehensweise gibt es aber schon länger, wenn man an die Verbreitung von E-Readern und den Bezug von E-Books von einschlägigen Anbieterplattformen denkt, aber auch Informationsplattformen wissenschaftlicher Verlage haben diese Technologien übernommen. Nun gehören aber immer mehr Informationsangebote wissenschaftlicher Verlage zum integrierten und lizenzierten Informationsangebot wissenschaftlicher Bibliotheken, die den Nutzenden durch Discovery Tools angezeigt werden. Geänderte Authentifizierungsverfahren beim Zugang zu diesen Plattformen der Verlage haben eine größere Aufmerksamkeit gegenüber diesen Praktiken bei betroffenen Wissenschaftlern hervorgerufen. Eine Reaktion des Ausschusses für Wissenschaftliche Bibliotheken und Informationssysteme (AWBI) der Deut-

9 Als Beispiel seien hier die entsprechenden Webseiten der kanadischen Seneca Bibliotheken zitiert: [Home - Database Content Not Covered by Library Search - LibGuides at Seneca Libraries \(senecacollege.ca\)](https://www.senecacollege.ca/libguides/), aufgerufen am 30.08.2022

10 [Discovery Manager, Oxford University Press – The North Carolina Central University School of Library and Information Sciences \(nccuslis.org\)](https://www.oxfordup.com/press-releases/2023/01/18/discovery-manager), zuletzt aufgerufen am 18.01.2023.

11 Siehe dazu den ausführlichen Beitrag von Cody: Hanson, User Tracking on Academic Publisher Platforms, <<https://www.codyh.com/writing/tracking.html>>, zuletzt aufgerufen am 18.01.2023.

schen Forschungsgemeinschaft (DFG) und Empfehlungen ließen nicht auf sich warten¹². Diese Reaktionen beschränken sich nicht nur auf Europa, auch die American Library Association hat zu dieser Problematik unmissverständlich Stellung genommen¹³.

Gerade die Corona-Pandemie und als Folge davon die Einschränkungen bei der Benutzung von Bibliotheken vor Ort haben zu einer erhöhten Nutzung digitaler Medien geführt, die durch das vorübergehende Freischalten von Informationsplattformen von Hosts und wissenschaftlichen Verlagen noch beflügelt wurde. Dass Menschen beim Surfen eine Vielzahl von Daten im World Wide Web hinterlassen, ist seit längerem bekannt. Dass aber Bibliothekssysteme mit Discovery Tools nun technologisch dazu in der Lage sind, das Nutzerverhalten abzuschöpfen, stellt einen Paradigmenwechsel dar. Das sogenannte Third-Party-Tracking durch von Bibliotheken lizenzierte Informationsangebote mittels Discovery Tools in Bibliotheksmanagementsystemen¹⁴ in der Cloud ist vielen Bibliotheksnutzerinnen und -nutzern, die diese Systeme als Recherchemittel bei ihrer Arbeit brauchen, nicht bewusst. Die Trackingpraktiken, wie sie in mehreren Fachbeiträgen dargestellt wurden, sind nicht völlig von der Hand zu weisen¹⁵. Dies zeigt die sehr zurückhaltende kurze Stellungnahme von Springer Nature¹⁶ zu erst vor kurzem veröffentlichten Fachbeiträgen¹⁷. Die Frage ist, ob beim Datentracking allenfalls von einem Verstoß gegen Datenschutzbestimmungen und dem Recht auf informationelle Selbstbestimmung gesprochen werden muss, wie dies der hessische Datenschutzbeauftragte Alexander Rosnagel in einem Interview¹⁸ tat, in dem er auch die Frage einer allfälligen Mithaftung durch Bibliotheken gleich mit beantwortet hat.

Lassen sich die standardisierten Discovery Tools der Hersteller in ihrer Treffsicherheit verbessern? Ein Projekt des Schweizerischen Wirtschaftsarchivs in Basel hat mit seinem Recherchetool mit Daten aus Alma bewiesen, dass man die Nutzerzufriedenheit steigern kann, wenn das Modul im Rechercheportal²⁰ den nur für das Wirtschafts-

archiv relevanten und kleineren Datenbestand aus dem gesamten Datenbestand von Swiscovery nutzt und den Nutzerinnen und Nutzern für die Recherche zur Verfügung stellt. Des Weiteren können die Nutzenden selbst bestimmen, welche Felder bei der Recherche bevorzugt abgefragt werden sollen sowie eine Voreinstellung der Filter bei der Facettenrecherche im Rechercheportal ermöglicht werden. Für das Beschlagwortmodul werden dann die selbsttrainierten Daten durch das Schweizerische Wirtschaftsarchiv verwendet. Im November letzten Jahres wurde die Webseite mit diesen Suchmöglichkeiten unter Einsatz der Software Elastic Search¹⁹ freigeschaltet. Einige Recherchen durch den Verfasser haben ein spürbar besseres Resultat bei den angezeigten Dokumenten gezeigt. Voraussetzung ist allerdings die Unterstützung der Archivare und Bibliothekare durch entsprechend geschulte Fachkräfte in der IT-Abteilung und im Archiv, um die anvisierten Verbesserungen zu erreichen. Ein weiterer Erfolgsfaktor ist eine umfassende Information der Nutzenden, die mit Anwendungsfällen (used cases) ergänzt wird. In modernen Bibliotheksmanagementsystemen in der Cloud wird auch der interbibliothekarische Leihverkehr mittels spezifischer Softwaremodule automatisiert und rationalisiert werden können. Ex Libris und OCLC bieten dazu die Softwaremodule Rapido und Tipasa an, deren Anschaffung bei mehreren europäischen Bibliotheken in der Evaluation oder vor der Anschaffung stehen. Sieht man sich dazu die diesbezügliche Werbung der beiden großen Bibliothekssystemanbieter für diese beiden Module an, so fällt auf, dass speziell darauf hingewiesen wird, dass mit diesen Modulen sichergestellt werden kann, dass das Urheberrecht und die Lizenzbedingungen von Fachpublikationen eingehalten werden können. Das heißt, dass Nutzungsdaten mit den Modulen gesammelt werden, die möglicherweise nicht nur bei den beteiligten Bibliotheken bleiben, da beide Module auf Servern von Ex Libris und OCLC laufen, auf denen auch die Nutzerkonten gespeichert sind. Was mit diesen Nutzungsdaten geschieht, ist für die Bibliotheken und Verbände, die diese Module

12 DFG: Datentracking in der Wissenschaft. Aggregation und Verwendung bzw. Verkauf von Nutzungsdaten durch Wissenschaftsverlage. Bonn 2021. Online: <https://www.dfg.de/download/pdf/foerderung/programme/lis/datentracking_papier_de.pdf> und Empfehlungen zu Methoden zur Kontrolle des Zugriffs auf wissenschaftliche Informationsressourcen. Ein gemeinsames Papier von Deutscher Bibliotheksverband e.V. (dbv) und Schwerpunktinitiative Digitale Information der Allianz der deutschen Wissenschaftsorganisationen, 27.11.2019, Online: <https://dbv-cs.e-fork.net/sites/default/files/2020-11/2019_11_27_dbv_Stellungnahme_Empfehlungen%20zu%20Methoden%20zur%20Kontrolle%20des%20Zugriffs%20auf%20wissenschaftliche%20Informationsressourcen.pdf>, beides zuletzt aufgerufen am 18.01.2023.

13 Siehe <https://www.arl.org/category/our-priorities/data-analytics/>, zuletzt aufgerufen am 18.01.2023

14 Siehe dazu Renke Siems, Das Lesen der andern, o-bib, Bd. 9 Nr. 1 (2022), zuletzt aufgerufen am 18.01.2023

15 Sehr ausführlich dazu ein Interview mit Felix Reda, Daten und Wissenschaft schützen in: <https://www.zbw-mediatalk.eu/de/2022/01/tracking-in-der-wissenschaft-so-koennen-bibliotheken-daten-und-wissenschaftsfreiheit-schuetzen>, zuletzt aufgerufen am 18.01.2023

16 Springer Nature Group, Kurz-Statement in ABI Technik 2022, 42 (2), S.108

17 Peter Reuter, Konstanze Söllner, Datentracking in der Wissenschaft – ein Sonderfall? und Michael Freiberg Third-Party-Tracking bei Wiley und Springer, beides erschienen in ABI Technik 2022, 42 (2), S.94-95 und S.96-104.

18 Third-Party-Tracking - ein Problem aus Sicht des Datenschutzes? Interview mit Prof. Alexander Rosnagel in: ABI Technik 2022, 42 (2), S.105-107.

19 Elasticsearch ist eine kostenlose und offene verteilte Suchmaschine und Analytics-Engine für Daten aller Art – von Textdaten über numerische Daten und Geodaten bis hin zu strukturierten und unstrukturierten Daten. Elasticsearch basiert auf Apache Lucene und wurde 2010 von Elasticsearch N.V. (jetzt unter dem Namen „Elastic“ firmierend) eingeführt, siehe <https://www.elastic.co/de/elasticsearch/>, zuletzt aufgerufen am 18.01.2023.

anschaffen, nicht transparent. Bleibt zu hoffen, dass die rechtlichen Bestimmungen der DSGVO und auch die Regelungen der künftigen E-Discovery-Verordnung der EU eingehalten werden kann. Rein technisch betrachtet wäre es möglich allfällige Lizenzvertragsverletzungen zurückzuverfolgen oder Auswertungen über den Gebrauch und das Volumen des interbibliothekarischen Leihverkehrs zu beobachten und auszuwerten, ohne dass die Bibliotheken und die Bibliothekskundinnen und -kunden etwas davon merken. Man bewegt sich hier in einer rechtlichen Grauzone, die der Klärung bedarf. Eine minimale datenschutzrechtliche Mindestvoraussetzung, um datenschutzrechtliche Vorgaben einzuhalten, wäre eine vorgängige Einverständniserklärung der Nutzenden, aber es scheint, dass dafür weitere Anpassungen an diesen Softwaremodulen nötig sind, da nicht nur die datenschutzrechtliche Regelung, sondern auch die Struktur der Verbände anders sind als in den Vereinigten Staaten.

Im Dezember 2022 richtete die Fach- und Allgemeinpresse ihre Aufmerksamkeit auf einen neuen Chatbot als Prototyp, der die Suchtechnologie mit künstlicher Intelligenz unterstützen soll²⁰. Die Suchanfragen können dabei in natürlicher Sprache formuliert werden. ChatGPT²¹ ist ein auf den ersten Blick beeindruckender Chatbot, der nicht nur die Fachwelt erstaunt, sondern das Potential von künstlicher Intelligenz (KI) auch einem breiten Publikum zeigt. Die Texte, die der Bot am Bildschirm anzeigt, wirken oft gut geschrieben. Tatsächlich wurde ChatGPT mit unzähligen Datensätzen trainiert, die im World Wide Web vorhanden sind. Hauptsächlich wurde der Bot mit Reddit-Daten aus dem Jahr 2021 trainiert. Die Verwendung von Trainingsdaten ist ein wichtiger Bestandteil des Lernprozesses von ChatGPT und hilft, dass der Prototyp in der Lage ist, ausführliche Dialoge zu führen. Auf dieser Grundlage entscheiden sich die Algorithmen der Software für die wahrscheinlichste Antwort, wenn eine Frage gestellt wird. Jeder Text ist also im Wesentlichen eine mathematische Entscheidung. Das Vorgehen ist ähnlich wie beim Smartphone. Man fängt gerade an, etwas ins Textfeld zu tippen und die Software macht bereits Vorschläge für die nächsten Wörter. Es gibt auch eine Anwendung für Bilder und Photos²². Erste Reaktionen von Nutzern im Netz wollen herausgefunden haben, dass KI-Anwendungen auch ihre Bilder verwenden, um die Algorithmen zu trainieren. Prisma Labs als Herstellerin der Lensa-App dementiert die Vorwürfe, die App kopiere den persönlichen Stil der

Künstler. Sie erklärt, dass die künstliche Intelligenz beim Lernen einfach sehr ins Detail gehe, um möglichst echt aussehende Gemälde zu produzieren. Technologie-Unternehmen setzen bereits KI als Technologie ein. Forschungseinrichtungen in Europa, China und den USA arbeiten bereits an neuen Methoden, um künstliche Intelligenz noch besser in den menschlichen Alltag zu integrieren. Klar ist aber auch, dass dieses Jahr auch deshalb ein wichtiges Jahr für KI werden könnte, weil OpenAI mit ChatGPT-4 den Nachfolger herausbringen will und Microsoft 42% des Aktienpaketes erwerben will. GPT soll die hauseigene Suchmaschine Bing revolutionieren und die Lücke zu Google verkleinern.

Es ist noch zu früh, um zu sagen, ob sich diese Technologien auch bei den Applikationen für Bibliotheken durchsetzen werden. Sollte sich jedoch im laufenden Jahr zeigen, dass dies der Fall sein sollte, würde dies die bisherigen Bibliothekssysteme als veraltet aussehen lassen und den Rhythmus für die Erneuerung dieser Systeme weiter verkürzen.

Nicht zuletzt sei darauf hingewiesen, dass je nach gewähltem Geschäftsmodell des Cloud-Computings²³ die Lizenzierung weiterer Softwaremodule auch immer eine größere Ressourcennutzung der Serverstruktur und des Speicherplatzes bei wachsendem Datenumfang zur Folge hat. Die Kosten für den Betrieb des Bibliothekssystems steigen damit im gleichen Umfang. Anfangs sind die Betriebskosten noch geringer als bei herkömmlichen Client-Server-Systemen, nehmen jedoch mit einem umfangreicheren Dienstleistungsangebot und dem zunehmenden Bestand an Katalog- und Nutzerdaten auch zu. Die anfänglich gemachten Betriebskostenannahmen werden damit obsolet und müssen aktualisiert werden. ■



Stephan Holländer

Lehrbeauftragter, Basel
stephan@stephan-hollaender.ch

20 ChatGPT ist ein Prototyp für einen auf künstlicher Intelligenz basierenden Chatbot. Er wurde von *OpenAI* entwickelt und im November 2022 veröffentlicht. Um das Modell zu verbessern, wird die künstliche Intelligenz (KI) von ChatGPT ständig durch menschliches Feedback verbessert.

21 <https://openai.com/blog/chatgpt/>, zuletzt aufgerufen am 18.01.2023.

22 <https://prisma-ai.com/lensa>, zuletzt aufgerufen am 18.01.2023.

23 Joachim Hackmann, Cloud Computing im Kostencheck, in Computerwoche, Link: <https://www.computerwoche.de/a/cloud-computing-im-kostencheck,1235100>, zuletzt aufgerufen am 18.01.2023 und Florian Weigmann, die vier Kostentreiber für die Cloud, in Cloud Computing Insider: <https://www.cloudcomputing-insider.de/die-vier-groessten-kostentreiber-fuer-die-cloud-a-1111139/>, zuletzt aufgerufen am 18.01.2023.