ZBW *Publikationsarchiv*

Publikationen von Beschäftigten der ZBW – Leibniz-Informationszentrum Wirtschaft
Publications by ZBW – Leibniz Information Centre for Economics staff members

Seeliger, Frank et al.

Article — Published Version

Zum erfolgversprechenden Einsatz von KI in Bibliotheken – Diskussionsstand eines White Papers in progress – Teil 2

b.i.t.online

Suggested Citation: Seeliger, Frank et al. (2021) : Zum erfolgversprechenden Einsatz von KI in Bibliotheken – Diskussionsstand eines White Papers in progress – Teil 2, b.i.t.online, ISSN 2193-4193, b.i.t.verlag gmbh, Wiesbaden, Vol. 24, Iss. 3, pp. 290-299, https://www.b-i-t-online.de/heft/2021-03-fachbeitrag-seeliger

This Version is available at:
http://hdl.handle.net/11108/490

Kontakt/Contact
ZBW – Leibniz-Informationszentrum Wirtschaft/Leibniz Information Centre for Economics
Düsternbrooker Weg 120
24105 Kiel (Germany)
E-Mail: info@zbw.eu
https://www.zbw.eu/de/ueber-uns/profil-der-zbw/veroeffentlichungen-zbw

Standard-Nutzungsbedingungen:
Dieses Dokument darf zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden. Sie dürfen dieses Dokument nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, aufführen, vertreiben oder anderweitig nutzen. Sofern für das Dokument eine Open-Content-Lizenz verwendet wurde, so gelten abweichend von diesen Nutzungsbedingungen die in der Lizenz gewährten Nutzungsrechte.

Terms of use:
This document may be saved and copied for your personal and scholarly purposes. You are not to copy it for public or commercial purposes, to exhibit the document in public, to perform, distribute or otherwise use the document in public. If the document is made available under a Creative Commons Licence you may exercise further usage rights as specified in the licence.

ZBW Leibniz-Informationszentrum Wirtschaft
Leibniz Information Centre for Economics

Mitglied der Leibniz-Gemeinschaft

«WORK IN PROGRESS»

Zum erfolgversprechenden Einsatz von KI in Bibliotheken

Diskussionsstand eines White Papers in progress – Teil 2

Frank Seeliger, Frank Puppe, Ralph Ewerth, Thorsten Koch, Anna Kasprzik, Jan Frederik Maas, Christoph Poley, Elisabeth Mödden, Andreas Degkwitz, Elke Greifeneder

Ist der Suchschlitz ein Vintage Look&Feel des letzten Jahrhunderts? Oder– die Frage sei erlaubt, – unter welchen Bedingungen es Bibliotheken schaffen könnten, dass Nutzerinnen und Nutzer die Erstsuche nach Informationen über sie laufen lassen? Das sind beispielhafte Fragen, die man sich angesichts gewonnener Erfahrungen aus bisherigen Innovationen stellen kann, aber ebenfalls mit Blick auf das, was Neues an Technologien sich auf dem Markt befindet und zunehmend viele Lebensbereiche durchdringt. Das Forschungsgebiet der Künstlichen Intelligenz (KI) ist in aller Munde und fordert von Informationseinrichtungen eine Antwort darauf, ob der Einsatz von KI-Technologien als weitere Facette der Automatisierung und Digitalisierung von Workflows und Service in Bibliotheken sinnvoll und hilfreich, und ob Engagement auf dieser Linie notwendig ist. Auch wenn sich Bibliotheken in ihrem steten Bemühen um Innovation mit ihrer Belegschaft und Leitungen noch differenziert bis reserviert zu KI bekennen, ist ein weiterer Verlauf der Entwicklungen gemäß dem chinesischen Sprichwort: „Was du nicht aufhalten kannst, solltest du begrüßen", absehbar. Was heißt das aber für Bibliotheken, den Wandel zu gestalten bis hin zu dem Moment, wenn ein Dienst über die Pilot-, Prototyp- und damit Projektphase hinaus als fester Bestandteil des Serviceangebots fest etabliert und verstetigt werden soll? Welche Ressourcen und Kompetenzen sind z.B. in der Konsequenz von der Einrichtung selbst vorzuhalten? Weitere Fragen betreffen die Bewältigung der Masse an Daten und ihre Offenheit für die Anwendung von KI-Algorithmen.

Warum sollten Bibliotheken Alternativen zu den Technologie-Giganten anbieten?

Der Einsatz von Künstlicher Intelligenz ist nicht an den Ort der Organisation gebunden, die diesen Einsatz benötigt, sondern kann überall dort durchgeführt werden, wo die Daten, Speicher- und Rechenkapazität in ausreichendem Maß zur Verfügung stehen. Darüber hinaus kann sie umso effektiver und fehlerärmer eingesetzt werden, je mehr Daten für das Training der ihr zugrundeliegenden Algorithmen zur Verfügung stehen („Coverage"). Somit können Organisationen, die über große Datenmengen und viel Rechenleistung verfügen, KI-Technologien besonders gut zum Einsatz bringen und haben in Konkurrenzsituationen starke Vorteile.

Große Player wie Google, Amazon, Facebook etc. oder auch kleinere, aber nicht weniger durchschlagende Newcomer wie DeepL (mit einer Belegschaft unter einhundert Mitarbeiterinnen und Mitarbeitern) bieten daher massentaugliche Services und Hilfen im Alltag. Und sie geben in der technologischen Entwicklung durch ihre enormen Ressourcen und Erfindergeist den Takt an. Sollten sich Bibliotheken angesichts der Datenübermacht und dem technologischen Vorsprung diesem Wettbewerb stellen, komplementär und alternativ, oder ausschließlich diese Plattformen nachnutzen? Haben Bibliotheken überhaupt eine Chance und wo bieten sie den feinen Unterschied? Aus Sicht der Diskutanten des Arbeitskreises sprechen mehrere Gründe für ein wohlüberlegtes und zielorientiertes Engagement auf dem weiten Feld der KI.

Langfristigkeit und Offenheit

Obwohl Bibliotheken und Informationseinrichtungen nicht über die Möglichkeiten und die enormen Ressourcen von Konzernen wie Google oder Facebook verfügen, haben sie auch organisatorische Vorteile. Bibliotheken müssen keine Shareholder kurzfristig zufriedenstellen, sondern können es sich erlauben, Services im Sinne ihrer Nutzenden auf die lange Sicht hin auszubauen und anzubieten – das Einverständnis der Geldgeber natürlich vorausgesetzt. Gerade Google hat in den vergangenen Jahren verschiedene Produkte eingestellt, weil diese nicht mehr den erwünschten wirtschaftlichen Nutzen erbrachten. Bibliotheken und Informationseinrichtungen können dagegen die verlässliche Verfügbarkeit von technischen Services garantieren. Natürlich müssen auch Bibliotheken ressourcensparend agieren, können aber im Rahmen ihres Auftrags zum Beispiel Spezialdienste für die Spitzenforschung (etwa im Rahmen eines Forschungsinformationsdienstes) anbieten, auch wenn die Nutzerzahl eines solchen Services sehr gering ist. Da Bibliotheken und Informationseinrichtungen in deutlich weniger ausgeprägten Konkurrenzverhältnissen zueinander stehen als privatwirtschaftliche

Organisationen, können Bibliotheken ihre Daten zum Training von KI-Systemen in vielen Fällen frei zur Verfügung stellen und mit anderen Einrichtungen austauschen. So können kooperativ aggregierte Datenbestände entstehen, die zwar nicht die Größenordnung der Big-Data-Bestände von Google et al. besitzen, die aber zum Training von KI-Verfahren in konkreten Fällen ausreichen können. Aus demselben Grund können Bibliotheken auch die verwendeten Algorithmen, den Sourcecode und die trainierten Systeme zur Verfügung stellen, austauschen und kooperativ weiterentwickeln. Die konsequente Verfolgung von Offenheit (Open Source, Open Access, Creative Commons bzw. CC-Lizenzen) und das Anwenden der FAIR-Prinzipien (**F**indable, **A**ccessible, **I**nteroperable, and **R**e-usable) ist auch an dieser Stelle nicht nur Selbstzweck, sondern Grundlage von synergetischen Kooperationen.

Transparenz und ethische Aspekte

Von der öffentlichen Hand getragene Einrichtungen müssen aufgrund ihrer öffentlichen Finanzierung eine hohe ethische Messlatte an ihre Tätigkeiten anlegen. Dies kann nicht nur eine Herausforderung, sondern auch ein großer Vorteil sein: Nutzende können sich darauf verlassen, dass die von ihnen erhobenen Daten z.B. DSGVO-konform eingesetzt werden, während bei vielen scheinbar kostenfreien Services im Internet der „Nutzer das Produkt ist", man also den kostenfreien Messaging-Dienst mit den persönlichen Daten „bezahlt". Während bei kommerziellen Produkten oft wenig Interesse daran besteht, die Funktionsweise eines KI-Systems offenzulegen, können und sollten sich Informationseinrichtungen darum bemühen, dies zu tun – zum Beispiel durch:

1. das Dokumentieren und Offenlegen des KI-Einsatzes für interessierte Nutzende, inklusive der Darlegung der verwendeten Verfahren und Datenquellen,
2. das konsequente Suchen, Offenlegen und Beheben von „Vorurteilen" (bias) in den Trainingsdaten und
3. durch das Einsetzen von Methoden aus dem Forschungszweig der „explainable AI", also Möglichkeiten, den Output von KI-Verfahren zumindest in Teilen nachvollziehbar zu machen.

Durch den ethischen und transparenten Einsatz von KI kann das Vertrauen der Nutzenden in die Methoden und in die sie anwendenden Einrichtungen gestärkt werden.

Lokale und kontextspezifische Dienste

Wie aus dem Gesagten ersichtlich, ist es oft wenig sinnvoll, das direkte Konkurrenzverhältnis zu den KI-Systemen der großen IT-Unternehmen zu suchen. Die Daten-, Personal- und Technologieübermacht der Unternehmen ist schlichtweg zu groß. Allerdings arbeiten KI-Verfahren dann am besten, wenn man spezifische und wohl definierte Probleme damit lösen möchte und über ausreichend Trainingsdaten des Problemkontextes verfügt. Dies eröffnet Möglichkeiten des Einsatzes von KI-Technologien in Bibliotheken in Bereichen, in denen

1. bibliotheksspezifische Probleme gelöst werden können,
2. bibliotheksspezifische Daten vorliegen oder
3. in denen im Idealfall beides zutrifft.

So scheint es sinnvoll, den KI-Einsatz in Bezug auf die Räumlichkeiten einer Bibliothek zu prüfen. Der „Seatfinder" der KIT-Bibliothek[1] ist so ein Beispiel für einen gelungenen Einsatz von lernenden Algorithmen[2]. Darüber hinaus gibt es Daten in Bibliotheken, die möglicherweise noch nicht vollständig maschinenlesbar und -interpretierbar erfasst wurden, die aber die Services der Bibliothek maßgeblich verbessern können. Beispiele hierfür sind Erwerbungsdaten zur Optimierung des Bestandsaufbaus, das maschinelle Training von Relevance-Ranking-Algorithmen in Discovery-Systemen durch Nutzungsdaten, etc. Auch und vielleicht gerade bei solchen eher unscheinbar wirkenden Anwendungsfällen kann der Einsatz von Künstlicher Intelligenz einen großen Nutzen bringen.

Fallstricke und uneinlösbare Annahmen oder Trittbrettfahren ohne konkretes Ziel

Ein flexibles Anpassungsprocedere und IT-Handwerk, Reverse Engineering, Projektarbeit, Nachnutzen von Open Source, Frameworks etc. kennzeichnen kreative Bibliotheksarbeit im IT-Bereich. Der Markt an Möglichkeiten, der jedem über das freie Web auch von den großen Playern zur Verfügung gestellt wird, scheint für Data Scientists oder Data Engineers ins schier Unendliche zu gehen. Aber alle Anwendungs- und Adaptionsarbeiten und das Nutzen freier Tools haben ihre Grenzen. Freie Software und Frameworks wie z.B. TensorFlow von Google können als Werkzeug für eigene Daten verwendet werden, diese Werkzeuge für eigene Daten einzusetzen. Diese Möglichkeit sollte aber nicht darüber hinwegtäuschen, dass eine Anpassung von Tools für hauseigene Problemstellungen nicht ohne teils erheblichen Aufwand einhergeht. Worauf auch

1 https://www.bibliothek.kit.edu/freie-lernplaetze.php
2 https://seatfinder.bibliothek.kit.edu/

Thorsten Koch hinweist: *There ain't no such thing as a free lunch*.[3]

Die Entwicklungskosten für oben genannte Produkte wie AlphaGo, das wirklich nur Go spielen kann, wurden im mehrstelligen Millionenbereich geschätzt. Für die Ausfinanzierung entsprechender Forschungsabteilungen mit über 700 Mitarbeitern des in London ansässigen Unternehmens DeepMind muss Google jährlich ca. 400 Mio. Pfund aufbringen. Grenzen der Machbarkeit bei öffentlichen Einrichtungen wie Bibliotheken sind damit schnell aufgezeigt. Aber Hoffnung, mit freien Tools auch ggf. in einer Garage gut durchstarten zu können, schafft z.B. der Blick an den Rhein

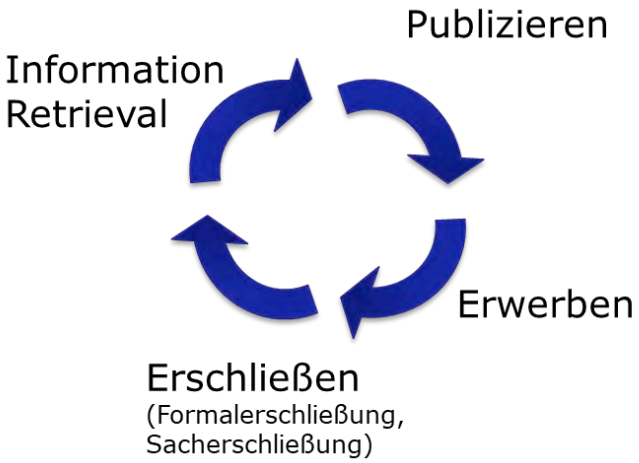

Abbildung 1: (Verarbeitungs-)Kreislauf einer Monographie

von Köln, wo DeepL mit zwanzig Mitarbeiterinnen und Mitarbeitern startete und einen sehr überzeugenden Onlinedienst mit der maschinellen Übersetzung schuf!

Ein zweiter wichtiger Aspekt ist die Leistungsbeschreibung dessen, was man sich als Produkt oder Assistenzsystem für welche konkrete Aufgabe wünscht. Das ist kein rein KI-spezifisches Problem. Eine gute Anforderungsanalyse ist elementar für jede Software-Entwicklung. Passiert das nicht, können die Konsequenzen haarig auch für KI-Systeme sein. Aus dem Science-Fiction-Roman ‚Per Anhalter durch die Galaxis' wissen wir, die "Answer to the Ultimate Question of Life, The Universe, and Everything," ist die Zahl 42. Sie ergibt keinen Sinn, denn "I think the problem, to be quite honest with you, is that you've never actually known what the question is."

Mit Thorsten Koch: Einen „Wunderalgorithmus" auf eine Menge Daten, über die man nichts weiß, zu werfen und zu hoffen, die Antwort auf die Frage zu bekommen, die man nicht kennt, ist nicht das Anliegen des vorliegenden Beitrags.

Das mag in der Theorie richtig sein, taugt aber nicht in der Praxis, oder doch?[4]

Neben den oben genannten Konferenzen mit Bezug zu KI-Ansätzen sind in Bibliotheken mehrere Projekte, welche die KI-Technologie in den Mittelpunkt stellten, umgesetzt worden. Sind ihre Konzepte und Ergebnisse so erfolgversprechend, dass man sich an ihnen ein strategisches Beispiel nehmen kann?

Die ersten Beispiele der Anwendung kreisen um die Kernaufgabe von Informationseinrichtungen und dabei ist der Effekt des KI-Recyclings interessant.

Hierbei ist prinzipiell festzustellen, dass in diesem Kontext sich existierende KI-Verfahren für verschiedenartige Szenarien verwenden lassen, in denen automatisiert neue Daten (aus einem diskreten Vokabular) erzeugt werden sollen. Als Ausgangspunkt der Betrachtung dient ein (Verarbeitungs-)Kreislauf einer Monografie (Abbildung 1). Hieran lässt sich zeigen, dass eine Mehrfachnutzung von KI-Verfahren möglich ist. So findet exemplarisch das Toolkit Annif der Finnischen Nationalbibliothek nicht nur beim Erschließen seine Anwendung. Es ist ebenfalls in Werkzeuge integriert, die publizierenden Autoren helfen, Stichwörter aus einem normierten Vokabular für eine Annotation auszuwählen.[5] Auch das Information Retrieval in Discovery-Systemen bedient sich gerne der KI-Techniken. Zunächst können direkt vor oder bei der Indexierung in einer Suchmaschine Metadaten durch KI-Methoden angereichert werden, indem weitere Metadaten wie Titel, Abstracts oder auch Volltexte verarbeitet werden. Weiterhin lassen sich Suchanfragen um assoziierte Terme oder Identifikatoren erweitern, um die Vollständigkeit der Suchergebnisse zum Beispiel unabhängig von ihrer zugrundeliegenden Sprache zu verbessern. Die hier eingesetzten Verfahren unterscheiden sich oftmals nicht oder nur kaum. Ähnliche Ideen finden sich auch an unterschiedlichen Stellen im Bereich der Erwerbung wieder, wo KI-Verfahren Fachreferentinnen und Fachreferenten bei der Auswahl von Medien assistieren, die zum Erwerbungsprofil einer bestimmten Bibliothek passen.

Die **DNB** beschäftigte sich als eine der ersten Bibliotheken mit der Nachnutzung von KI-Verfahren und kann auf zehn Jahre Erfahrungen mit dem produktiven Einsatz von Software-Lösungen für die automatisierte

3 siehe https://www.th-wildau.de/files/Bibliothek/Bilder/InnoCamp_2020/2020-11-05-Koch-InnoCampus-TANSTAAFL.pdf

4 in Anlehnung an Immanuel Kants Abhandlung „Über den Gemeinspruch: Das mag in der Theorie richtig sein, taugt aber nicht für die Praxis." von 1793

5 siehe Anwendungsbeispiele siehe https://www.annif.org

[Seite enthielt Werbung]

Erschließung von Medienwerken zurückblicken.[6] Eine maschinelle DDC-Sachgruppenvergabe,[7] die Vergabe von DDC-Kurznotationen und von GND-Schlagwörtern sind vor allem für unkörperliche Medienwerke etablierte Workflows. Verfahren zur Qualitätskontrolle der maschinell generierten Erschließungsdaten sind eingerichtet.

Dabei haben die Kolleginnen und Kollegen der DNB in der Vergangenheit differenzierte Erfahrungen sammeln können. Ein Projekt zur halbautomatischen Formalerschließung von gedruckten Dissertationen basierte auf maschinellen Methoden zur Gewinnung von Metadaten aus den Titelblättern. Allerdings wurden die Erwartungen an die Qualität der Daten und die Zeitersparnis durch den Einsatz eines solchen Verfahrens nicht erfüllt.

Aktuell arbeitet die DNB daran, ein neues Erschließungssystem aufzusetzen, das die vorhandenen produktiven Werkzeuge ablösen soll. Das neue System besitzt einen modularen Aufbau und erlaubt den Austausch sowie die Kombination einzelner Komponenten und Verfahren. Die Kernfunktionalitäten für die Vergabe von DDC-Sachgruppen, DDC-Kurznotationen und GND-Schlagwörtern werden mit dem Toolkit Annif der Finnischen Nationalbibliothek umgesetzt. Weitere Bestandteile des neuen Erschließungssystems sind KI-gestützte Assistenztools zur Textextraktion (OCR) und Textaufbereitung bei digitalen Objekten, Spracherkennung sowie Tools zur Einbindung in die vorhandenen Erschließungsworkflows der DNB. Zu den Erfahrungen und Perspektiven mit Annif veranstaltete die DNB Ende 2020 einen Online-Workshop[8] für das Netzwerk Maschinelle Verfahren in der Erschließung.[9]

Weitere Beispiele und ihre Referenzen, ohne dem Anspruch auf Vollständigkeit gerecht zu werden, deuten die Vielfalt an Möglichkeiten mit dem Set an KI-Technologien an:

Die **ZBW** – Leibniz-Informationszentrum Wirtschaft in Hamburg und Kiel befasst sich schon einige Jahre aktiv mit der eigenständigen Entwicklung von KI-basierten Lösungen für die Automatisierung der Inhaltserschließung.[10] Das Ziel dabei ist die Verschlagwortung von Publikationen mit dem ‚Standardthesaurus Wirtschaft' (STW). Der bisherige Ansatz besteht im Wesentlichen darin, für Publikationen auf der Basis von Textmaterial aus den entsprechenden Metadatensätzen (Titel, Autoren-Keywords) über eine intelligente Kombination verschiedener Machine-Learning-Algorithmen STW-Deskriptoren vorzuschlagen. Über die Algorithmenschicht wird eine weitere Schicht zur Qualitätskontrolle gelegt, die sich ebenfalls teilautomatisieren lässt. Das Verfahren wird kontinuierlich weiterentwickelt und mittlerweile werden auch Deep-Learning-Ansätze erprobt. Die ZBW führt also ihre eigene angewandte Forschung zur Inhaltserschließung durch und setzt gleichzeitig die so erarbeiteten Lösungen im eigenen Haus nahtlos in produktive Dienste um und bettet sie dafür in eine geeignete Softwarearchitektur ein. Umgebende Maßnahmen sind eine kommunikative und operative Verschränkung mit der intellektuellen Inhaltserschließung im Haus und die Zusammenarbeit mit internen und externen Stakeholdern insbesondere auch im Hinblick auf eine stetige Verbesserung der Inputdaten. Die ZBW kommuniziert und kooperiert national und international mit anderen Institutionen zum Thema automatisierte Inhaltserschließung, so beteiligt sie sich zum Beispiel aktiv an der Weiterentwicklung des Open-Source-Toolkits Annif der finnischen Nationalbibliothek und berät entsprechend auch Institutionen wie die DNB zum Einsatz von Annif und zum Aufbau einer umgebenden Softwarearchitektur dafür.

Die Technische Informationsbibliothek (**TIB**), das Leibniz-Informationszentrum für Technik und Naturwissenschaften hat im Jahr 2011 gemeinsam mit dem Hasso-Plattner-Institut begonnen das TIB AV-Portal (https://av.tib.eu/) zu entwickeln. Hier werden wissenschaftliche AV-Medien englisch- und deutschsprachig u.a. nach der GND automatisiert mit Methoden der Sprach- und Bilderkennung erschlossen.[11]

Seit dem Jahr 2015 hat die TIB mit assoziierten Professuren ihre eigenen Forschungskapazitäten deutlich ausgebaut. So gibt es mittlerweile mehrere Forschungsgruppen (z.B. Data Science and Digital Libraries (Prof. Sören Auer), Knowledge Infrastructures (Dr. Stocker), Learning and Skill Analytics (Dr. Kishmihok), Scientific Data Management (Prof. Maria-Esther Vidal), Visual Analytics (Prof. Ralph Ewerth)), die sich intensiv mit semantischen Technologien und maschinellen Lernverfahren beschäftigen.

6 siehe als Beispiel für viele andere Vorträge https://www.th-wildau.de/files/Bibliothek/Bilder/InnoCamp_2020/20201105_innocampDNB2020final-1.pdf

7 siehe https://www.dnb.de/DE/Professionell/DDC-Deutsch/DDCinDNB/ddcindnb_node.html

8 siehe dazu https://wiki.dnb.de/display/FNMVE/Erfahrungen+und+Perspektiven+mit+dem+Toolkit+Annif

9 siehe dazu https://wiki.dnb.de/display/FNMVE/Netzwerk+maschinelle+Verfahren+in+der+Erschliessung

10 siehe dazu https://www.zbw.eu/de/ueber-uns/arbeitsschwerpunkte/automatisierung-der-erschliessung und lfd. aktualisierte Publikationen und Vorträge von Anna Kasprzik

11 siehe Strobel, Sven; Plank, Margret: Semantische Suche nach wissenschaftlichen Videos – Automatische Verschlagwortung durch Named Entity Recognition. In: Zeitschrift für Bibliothekswesen und Bibliographie, 2014, Heft 4, S. 254-8

Hervorzuheben ist die unter anderem von der TIB vorangetriebene Initiative zum Open Research Knowledge Graph (ORKG, https://orkg.org) mit dem Ziel, die bisher in Artikeln durchgeführte wissenschaftliche Kommunikation durch die Verwendung von Wissensgraphen völlig neu zu organisieren. Dies soll u.a. dadurch, dass Beziehungen zwischen Entitäten und anderen Artikeln explizit modelliert werden, und so direkt für Suchanfragen dienlich sind. So können etwa die Related-Work-Suche oder der Vergleich von Fachartikeln einfacher und systematisch realisiert werden. Die Forschungsgruppe Visual Analytics erforscht z.B. für das AV-Portal neue Verfahren zur Inhaltserkennung in Videos mittels Deep-Learning-Verfahren, beschäftigt sich aber auch mit der Informationsextraktion aus wissenschaftlichen Texten, Diagrammen und Bildern. Ein aus der Forschungsgruppe Visual Analytics entstandenes und in den TIB Labs (https://labs.tib.eu/) dokumentiertes Projekt beschäftigt sich spielerisch mit der automatisierten Erkennung des Aufnahmeorts von Fotos (siehe https://labs.tib.eu/geoestimation/). Darüber können auch eigene Fotos getestet werden. Weiterhin beschäftigt sich die Forschungsgruppe im Forschungfeld „Search as Learning“ damit, wie der physische Lernraum Bibliothek in die Onlinewelt übertragen werden kann. Konkret geht es darum, wie Suchmaschinen für das informelle Lernen der Nutzenden optimiert werden können, z.B. indem Lernabsichten oder Wissenszuwachs automatisch geschätzt werden. Die Ergebnisse auch zu den drei Vorträgen über KI auf der #vbib20 können sich sehen und hören lassen, siehe http://av.tib.eu.

Mit dem kommerziellen Produkt YEWNO als semantischer bzw. Konzept-Suchmaschine geht die Bayerische Staatsbibliothek (**BSB**) seit 2017 neue Wege.[12] Im Index hinterlegt sind mehrere einhundert Millionen Publikationen, viele Open Access, in vorwiegend englischer Sprache, somit Metadaten und die dazugehörigen Volltexte. Automatisch werden Konzepte (insgesamt 500 Mio.) aus den Textkorpora gezogen, ohne dabei Ontologien zu verwenden, sondern nur auf Basis der KI-Methoden wie des Maschinellen Lernens. Die beständige Aktualisierung der Konzepte und ihre logische Verknüpfung erlaubt damit auch Zusammenhänge post festum zu erkennen. Als Beispiel erhielt am Ende des Ersten Weltkriegs stattgefunden habende Kerenskj-Offensive erst viel später in der Geschichtsschreibung diesen Titel. Ein anderes Beispiel sind die Azteken, eine Wortschöpfung des 18. Jahrhunderts. Trotzdem würde man Texte mit den Konzepten verknüpft finden, auch wenn die Begrifflichkeiten selbst im Text nicht erscheinen. Die graphische Visualisierung zeigt die Verknüpfungen an. Die Erfahrungen der Nutzerinnen und Nutzer werden als sehr unterschiedlich wiedergegeben, von nützlichen Zufallsfunden bis hin zu einer gewissen Umständlichkeit und Unvertrautheit im Umgang mit dem graphischen Frontend. Parallel zu YEWNO werden alle weiteren Suchwerkzeuge weiterhin den Kundinnen und Kunden angeboten.

An der Universität **Würzburg** wurden Erfahrungen mit der Erschließung alter Drucke gesammelt. Ein Team um Prof. Puppe und dessen Lehrstuhl für Künstliche Intelligenz und Wissenssysteme arbeitet zusammen mit dem neu gegründeten Zentrum für Philologie und Digitalität Kallimachos in verschiedenen Projekten an der Weiterentwicklung von Layout- und Texterkennung (OLR und OCR) zur Transkription alter Drucke, aber auch alter musikalischer Handschriften und gescannter Dokumente mit Tabellen.[13] Dabei werden Techniken des Deep Learning mit semantischer Analyse kombiniert. Ein weiterer Schwerpunkt liegt in der inhaltlichen Erschließung von Romanen, Märchen usw. im Hinblick auf die Erkennung von Personen und deren Relationen in Figurennetzwerken mittels Techniken der maschinellen Verarbeitung natürlicher Sprache (NLP).

Das BMBF-Projekt „Qurator: Curation Technologies“ ist angetreten, KI-Verfahren für die digitale Kuratierung in branchenübergreifenden Kontexten zu entwickeln. Die Staatsbibliothek zu Berlin - Preußischer Kulturbesitz (**SBB-PK**) ist im Projekt für den Bereich des digitalisierten Kulturerbes zuständig.[14] Indem aktuelle Verfahren aus der Forschung aufgegriffen und für die Anforderungen der Digitalisierung historischer Dokumente adaptiert werden, entwickelt die SBB-PK einen Workflow für die Strukturerkennung, Texterkennung, Entitätenerkennung und -verlinkung von Digitalisaten. Für ein Fazit ist es noch zu früh.

Zum Wissenschaftsjahr 2019 mit Motto „KI” legte das **Goethe-Institut** das Programm „Generation A=Algorithmus” auf, u.a. mit aufgezeichneten Online-Tutorials als Couch Lessons zu KI.[15] Mit dem Einsatz Künstlicher Intelligenz wird sich, so der Ausgangs-

12 siehe https://www.bsb-muenchen.de/article/kuenstliche-intelligenz-bayerische-staatsbibliothek-testet-zukunftsweisendes-semantisches-recherchewerkzeug-1668/ und nachfolgende Vorträge und Berichte, u.a.https://www.th-wildau.de/files/Bibliothek/Bilder/InnoCamp_2020/Innocamp_2020_Yewno.pdf

13 siehe https://av.tib.eu/media/47150?hl=vbib20+puppe

14 siehe https://blog.sbb.berlin/qurator-digitale-kuratierung/

15 siehe https://www.goethe.de/prj/one/de/gea.html

punkt des Programms, unsere Gesellschaft von Grund auf verändern. Das Programm „Generation A=Algorithmus“ setzt sich mit genau diesen Veränderungen und ihren Auswirkungen auf eine noch in den Kinderschuhen steckende Generation A auseinander. Die Formate umfassen u.a. eine europaweite Umfrage unter Jugendlichen, ein „Robots in Residence“-Programm, die Etablierung einer „A(I)lliance“, wöchentliche „Couch Lessons“ sowie ganz praxisorientierte „Data Detox“ Fortbildungen für Bibliothekarinnen und Bibliothekare.

Die zwei Leben der Chatbots

Ein Gigant der Mathematik und ein Vordenker sowohl des Computers als auch der Künstlichen Intelligenz ist Alan Turing. Vielen begegnet er indirekt z.B. beim sogenannten CAPTCHA-Test, ausformuliert: *completely automated public Turing test to tell computers and humans apart*, oder deutsch „vollautomatischer öffentlicher Turing-Test zur Unterscheidung von Computern und Menschen“. Man sichert sich mit in Kacheln unterteilten anklickbaren Fotos, die etwas Bestimmtes enthalten, und weiteren interaktiven Methoden dabei ab, dass das Gegenüber am Bildschirm eine Person und keine Maschine ist. Von Turing stammt ebenfalls ein hoher formaler Anspruch in der Mensch-Maschinen-Interaktion: Der sogenannte Turing-Test formuliert umgangssprachlich den Anspruch, dass eine Maschine dann als künstlich intelligent einzustufen ist, wenn eine Person unbemerkt mit einem Computer kommuniziert und die eloquente Person die antwortende Gerätschaft im Hintergrund nicht bemerkt. Dass es KI-geführte Dialogsysteme geben kann, bewies der oben bereits erwähnte Weizenbaum 1966 mit ELIZA. Dies führte kurz nach der Jahrtausendwende bereits dazu, dass verschiedene Bibliotheken sich für **Chatbots** engagierten. Es wurden Dialogsysteme integriert, die z.B. außerhalb der Öffnungs- und Servicezeiten vollautomatisiert Kundenanfragen zielführend beantworten konnten (Ausleihfristen, Öffnungszeiten, Fernleihe etc.).[16] Der Aufwand für die Pflege der knowledge base war ein Grund, der der ersten Generation von Chatbots ein baldiges Ende bescherte. Für die Chatbots wurde schon vor fünf Jahren ein Revival festgestellt.[17]

Anlass dafür sind neue Methoden, die mehrdeutige Sprache formal und technisch genauer abbilden zu können. Zum Beispiel können mathematische Vektoren Worte in einem mehrdimensionalen Vektorraum abbilden. Man spricht von den *word embeddings*, die sich effizient auf riesigen unannotierten Daten trainieren. Word2Vec stellt einen solchen Lernalgorithmus als Beispiel dar. Die Vektorarithmetik wird auf Sätze, Konzepte, aber auch auf Sätze angewendet. Aktuell arbeiten zwei Einrichtungen an dieser Form der Renaissance von Chatbots, die ZBW in Hamburg/Kiel mit Low Fidelity[18] und die TH Wildau[19].

Eine besondere Herausforderung in der Virtualisierung des genannten Informationsdienstes, hier beispielgebend gleichfalls für andere Ansätze von KI, stellt die Verfügbarkeit von Daten dar. Es sind keine typischen Dialoge zwischen bibliothekarischer Fachauskunft und dem Kunden digital verfügbar, womit man Systeme trainieren kann. Ein Weg, an solche Daten zu gelangen, könnte das automatisierte Verarbeiten von Online-Tutorials aus Bibliotheken sein, wie sie gerade in der Zeit der Pandemie zahlreich entstanden sind. Weniger zielführend wäre als Gegenbeispiel das Publikationskorpus in bibliothekarischen Fachzeitschriften. Interessant ist dieses Sprachmaterial zudem für die Weiterverwendung, wenn z.B. NLP-Systeme gleichfalls auf das Zuhören (Natural Language Understanding (NLU)) und Artikulieren als Leistungen eines Assistenzsystems trainiert werden sollen. Der klassische Anwendungsfall stellt hier der Roboter dar, wie er gerade mit Produkten wie Pepper und Nao Einzug hält in Informationseinrichtungen.

Das Konzept von Chatbots, NLP und NLU einzusetzen, ist – um kurz den Blick über den Tellerrand zu wagen – viel umfassender, als ausschließlich auf Dialogsysteme zu setzen. Methoden der NLP und NLU kann ebenso beim Kategorisieren von Texten helfen, als auch selbständig auf einer breiten Datenbasis von z.B. wissenschaftlichen Publikationen kreativ neue Antworten zu finden. Hiervon künden Beiträge wie in The Guardian,[20] modelliert aus der KI-Software GPT-3, oder der IBM Project Debator[21].

Zahlreiche weitere Beispiele ließen sich mit Bibliotheksbezug finden, die häufig der Herausforderung nachgehen, aus vielen Publikationen das Gewünschte zu filtern oder auf Beziehungen untereinander hinzuweisen. Neben dem zuvor genannten Open Re-

16 siehe zu Chatbots in Hamburg oder Köln in „BuB : Forum Bibliothek und Information“ Heft 1 von 2007 (S. 20) und Heft 3 von 2012 (S. 229-32)

17 siehe https://www.zbw-mediatalk.eu/de/2016/05/revival-der-chatbots-revolution-des-kundenkontakts-disruption-des-software-marktes/

18 siehe https://www.th-wildau.de/files/Bibliothek/Bilder/InnoCamp_2020/ChatBot_Planungen_im_Kontext_von_EconBiz_InnoCamp.pdf

19 siehe chatbot.th-wildau.de

20 siehe vom September 2020 „Are you scared yet, human?“ unter https://www.theguardian.com/commentisfree/2020/sep/08/robot-wrote-this-article-gpt-3

21 siehe zur Frage der Subventionierung von Vorschulen www.research.ibm.com/artificial-intelligence/project-debater/

search Knowledge Graph (ORKG) kann hier z.B. die AI-gestützte Suchmaschine Semantic Scholar (**semanticscholar.org**) mit inhaltlichen Schwerpunkten in Information und Biomedizin genannt werden. Die seit 2015 bereitgestellte Suchmaschine stellt eine gute Ergänzung gegenüber anderen Plattformen wie PubMed und Google Scholar in der Welt der wissenschaftlichen Publikationen dar.

Im Fazit: KI für Bibliotheken kann eine Automatisierungs- und Erschließungshilfe sein, um von Routinen zu entlasten und eine neue Qualität an Daten inkl. ihres Findens zu erreichen. KI muss als probates Mittel der Automatisierung und Digitalisierung mit berücksichtigt werden und kann darüber hinaus Ausgangspunkt für neue Services sein.

Voraussetzungen zur erfolgreichen Implementierung von KI-Systemen

Überzeugen die *best-practice*-Beispiele mit den gewonnenen Erfahrungen, ihren Konzepten und ihren Visionen, ist ein wichtiger Impuls dafür gegeben, sich in der eigenen Einrichtung mit der gewählten KI-Technologie und ihrem verheißungsvollen Potential zu beschäftigen. Der lange Weg zu einem ersten konkreten Projekt und Ergebnis beginnt damit, sich über Voraussetzungen, Rahmenbedingungen und Grundsatzfragen zu verständigen, wenn diese neue technologische Ausrichtung, die auch ein Kurswechsel sein kann, z.B. ebenfalls vom Bibliotheksteam, Stakeholdern und weiteren wichtigen Protagonisten langfristig mitgetragen werden soll.

Ein weiterer wichtiger Aspekt betrifft den Kompetenzaufbau zu KI-Technologie in den Informationseinrichtungen selbst, wie er z.B. im Rahmen der Datenvisualisierung über das Format der *library carpentries* erbracht wird. Es sollte der Anspruch bestehen, Prozesse bis zu einem gewissen Grad zu verstehen, ähnlich wie Ranking-Abfolgen von Discovery-Tools erklärbar und nachvollziehbar bleiben sollten.[22] Der folgerichtige nächste Schritt nach Durchführung erster KI-Projekte ist zu entscheiden, ob das Ergebnis eine Daueraufgabe bibliothekarischen Handelns im Sinne eines neuen Dienstes werden soll. Soll dies der Fall sein, muss das Projekt nach Auslaufen mit allen notwendigen Ressourcen für einen Dauerbetrieb geführt werden. Wie am bayerischen Anwendungsbeispiel der semantischen Suchmaschine YEWNO zu sehen war, ist eine Grundvoraussetzung für das erfolgreiche Trainieren von KI-Systemen eine hohe Anzahl von offenen und maschinentauglichen Daten. Komplett verfügbare Volltexte, mindestens die sie beschreibenden Daten ist eine *conditio sine qua non*, damit KI-Werkzeuge verwertbare Resultate liefern.

Für die offenen Daten bis hin zu Open Access bei Volltexten muss, ähnlich bei Forschungsdaten, das FAIR-Prinzip gelten (**F**indable, **A**ccessible, **I**nteroperable, and **R**e-usable).

Diese Voraussetzung gilt in allen weiteren Kontexten, z.B. den Chatbots. Würden die verschiedenen Wissensbasen von solchen dialogorientierten Frage-Antwort-Maschinen aus dem Bibliothekskontext allgemein verfügbar sein, würden Systeme vor Ort besser auf die unterschiedlichen und manchmal unerwartbaren Anfragen reagieren können.

Neben der Offenheit von Daten bildet die Qualität von Metadaten und Strukturen eine wichtige Basis für KI-Anwendungen. Weitere Randbedingungen sind über eine ausreichende KI-Infrastruktur (GPU statt CPU) gegeben, verbindlichen Kooperationen für entsprechende Leistungen u.a. zur stetigen Verbesserung des erreichten Standes.

Ein Bündel an Grundsatzfragen

Bibliotheken ist spartenübergreifend gemein, für eine gewisse Form und Qualität von veröffentlichten Informationen als Wissensspeicher zu stehen und für den Zugang zu ihnen zu sorgen. Beides – Information und den Bibliothekskunden – zusammenzubringen mit dem Anspruch „wie es gesucht wird, so gefunden“, gibt Bibliotheken ihre Aufgabe unabhängig jeglicher „Vermittlungs“-Technologie vor. Die besonderen Anforderungen entstehen aus Randbedingungen wie der der Diversität an Nutzerinnen und Nutzern. Es gibt nicht eine Nutzerin oder den Nutzer einer Spezial-, Schul-, Stadt-, Hochschul- oder Forschungsbibliothek. Dementsprechend vielfältig müssen Strategien der medialen Verknüpfung mit ihren Interessenten ausfallen. Lösungen von der Stange (*ready-made* oder *off-the-shelf*) werden diesen funktionalen Anforderungen nicht gerecht. Der Markenkern von Informationseinrichtungen liegt neben dem Besitz von Medien in der Herstellung des Zugangs über geeignete Nachweisinstrumente und Recherchetools zum gewünschten Werk. Kann hierbei KI Abhilfe schaffen oder sogar Neues? Schreitet man zu neuen Ufern, muss die Frage beantwortet werden, was man bislang erreicht hat, z.B. beim jeweiligen Um-

22 Es ist von außen nicht immer ersichtlich, ob bibliothekarische Technologien wie z.B. Werkzeuge der Sacherschließung (Digitale Assistenten), Roboter oder Discovery-Tools bereits über Ansätze der sog. schwachen KI verfügen. So kann ein Discoverysystem ein intellektuell optimiertes Relevance Ranking besitzen oder eines, das beispielsweise über „Learning-to-Rank“-Verfahren (https://en.wikipedia.org/wiki/Learning_to_rank) verbessert wurde. Auch unter den Autor/-innen dieses Beitrags gab es die abweichenden Ansichten, solche Systeme aufgrund ihrer Funktionsweise als KI-Systeme zu bezeichnen oder alternativ das Label KI/Nicht-KI ausschließlich von den intern verwendeten Verfahren abhängig zu machen.

stieg von Katalogkästen über webOPAC's bis hin zu suchmaschinenzentrierten Discovery-Lösungen. Was sind z.B. unsere Erfahrungen aus bisherigen Nutzungen von Nachweisinstrumenten, Recherchetools und Wissenspräsentationen?

Wie misst man die Qualität von Retrievalergebnissen? Wenn, wie häufig kolportiert, angeblich bis zu 90% der Suchanfragen an bibliothekseigene Darstellungen über Google und Scholar, Web of Science oder Scopus, oder verlagseigene Portale in das System kommen, wo liegen die Gründe? Oder sollten Bibliotheken bessere Services anbieten? Ein ambivalentes Merkmal der großen Suchmaschinen liegt in ihren leistungsfähigen Ranking-Algorithmen, die oft nicht transparent, aber trotzdem einen großen Einfluss haben – wer schaut sich schon den hundertsten Treffer an? Nach welchen Kriterien sollten Bibliotheken ihre Ergebnisse sortieren? Hilfreich sind auch nutzergesteuerte Filtermechanismen der Treffermenge – aber welche Filter wünschen sich die Nutzerinnen und Nutzer? Wie gut eignen sich dafür die routinemäßig erhobenen Metadaten bei der Sacherschließung? Sollen nur Ergebnisse der eigenen Bibliothek oder auch Treffer aus anderen Quellen im World Wide Web angezeigt werden, eventuell beschränkt auf Seiten, die frei zugänglich sind? Wie gut kennen Bibliotheken ihre Nutzerinnen und Nutzer – ist es sinnvoll Suchanfragen nach unterschiedlichen Zielgruppen unterschiedlich zu beantworten? Bieten Ratings der Nutzerinnen und Nutzer für Informationsangebote ähnlich wie in Online-Shops einen Mehrwert für andere Nutzerinnen und Nutzer bei der Suche?

Ist der Trend im Nutzungsverhalten bei der Informationssuche zu den großen Suchmaschinen vielleicht sogar umkehrbar? Möchte man vermeiden, dass zukünftig nur Angebote der TechGiganten zum Befriedigen der Informationsbedürfnisse genutzt werden, sollte das Schaffen von Mehrwerten durch den Einsatz neuer Technologien eine zentrale Strategie sein. Dabei stößt man aber schnell auf potentielle Konflikte: Viele frei verfügbare Tools sind erfolgreich dadurch, dass sie Nutzer- und Nutzungsdaten (*digital fingerprinting*, Statistiken) in das Ranking mit einfließen lassen. Daher muss die Frage gestellt werden: Benötigen wir neben einem nationalen Hosting, Verbund etc. evtl. auch eine **nationale Strategie der Auswertung des Informationsverhaltens** der Bibliotheksnutzerinnen und -nutzern mit den von uns betriebenen Tools an zentraler Stelle? Ist ein solcher Ansatz unter Einhaltung der DSGVO überhaupt möglich? (ggf. ELSI-Merkmale (Ethical, Legal and Social Implications))

Es liegt zugegebenermaßen eine Ungewissheit darin vor, will man sich konkret die Zukunft der Bibliothek über zehn, zwanzig Jahre hinaus vorstellen. Entwicklungen wie die des Buchdrucks oder das Aufkommen des Internets stehen für eine Zäsur, auch in der inhaltlichen Ausrichtung von Bibliotheken. Neben neuen Technologien verändert sich das Portfolio von Informationseinrichtungen ebenfalls in ihren Beständen. Man denke im wissenschaftlichen Kontext an das zu begrüßende Voranschreiten von Open-Access-Publikationen oder des Vorliegens digitaler Publikationen. Fallen Lizenzierungen und Archivrechte bei digitalen Kollektionen weg, wie werden dann Nachweis- und Recherchewerkzeuge konzipiert sein?

Mit Blick auf das unhinterfragte Muss bibliothekarischer Wertschöpfung durch Erzeugung von bibliographischen, die Publikationen beschreibenden Daten zur besseren Auffindbarkeit soll an dieser Stelle kurz auf einen Diskurs, wie er z.B. durch den Leiter der TIB Hannover, Sören Auer, immer wieder angestoßen wird, hingewiesen werden. Welchen **Wert** werden künftig **Metadaten** haben?[23] Sie sind seit Bibliotheksgedenken eine Grundvoraussetzung für das Organisieren und Auffinden von Informationen. Mit dem Vorzeigebeispiel von Google als Suchmaschine, die nicht auf die Nutzung intellektuell erstellter Metadaten angewiesen ist, ist die Frage zu stellen, welchen künftigen Beitrag sie dazu in der digitalen Welt leisten werden. Informationstechnik, das ist das Unbequeme, rückt Selbstverständlichkeiten in neues Licht.

Ggf. hat man Aussagen von Forschenden dieser Art bereits vernommen: Selbst im Bereich der Kernkompetenz rezipiert man nicht mehr selbst aktuelle Beiträge, sondern „lässt" lesen. Umfassender wird diese Art der indirekten Rezeption bei inter- und transdisziplinären Ansätzen und der dazugehörigen Literatur. Man kann nicht mehr alles zu einem Thema lesen, man lässt vielleicht KI für sich lesen und unter einer bestimmten Fragestellung das Relevante filtern. Wird aus Resignation vor der Menge eine Vision für ein neues Assistenzsystem?

Ausblick mit offenem Schluss

Die genannten Fallbeispiele, das Wissen um Möglichkeiten von KI, das Erahnen einer neuen Selbstverfasstheit von Bibliotheken lassen konkrete Ziele eigenen Handelns auf greifbar-operativer Ebene erkennen. Vorzuschlagende Ziele bibliothekarischen Handelns mittels KI-Assistenz können sehr vielfältig sein. Hier

23 siehe Sören Auer im Interview mit b.i.t.-online (2020) auf die Frage: Was halten Sie für überbewertet? „Metadaten. Diese waren über Jahrhunderte eine absolute Grundvoraussetzung für das Organisieren und Auffinden von Informationen – in der digitalen Welt hingegen kann oft auf Metadaten verzichtet werden, da Suchmaschinen ... Google ist hier das Paradebeispiel," [online unter https://www.b-i-t-online.de/heft/2020-02-letzte_seite.pdf]

eine beispielhafte Sammlung als Ergebnis eines Brainstormings:

- Entlastung der Mitarbeiterinnen und Mitarbeiter von intellektuellen Routinearbeiten
- KI wandelt das bibliothekarische Handeln dahingehend, dass MitarbeiterInnen und Mitarbeiter in den Bibliotheken beim Einsatz solcher Verfahren die Rolle der fachlichen Expertinnen und Experten wahrnehmen, um die Qualität der eingesetzten KI-Verfahren zu beurteilen und zu verbessern.
- Erzeugen von Metadaten höchster Qualität der Erschließung (Validierung, Goldstandard)
- Zielgenaue Informationsversorgung, z.B. durch künstlich-intelligente Antizipation von Informationsinteressen
- Verknüpfung von Wissen und Materialien, interdisziplinär und crossmedial (*knowledge graph*)
- Anbieten von Informationssystemen mit
  - natürlichsprachlicher Interaktion, Question-Answering-Dialogen
  - konzept- und textbasiertem Retrieval (nicht mehr nur Einzelworte, ganze Passagen)
  - einer durch KI verbesserten Anzeige von *related work*
  - einer verbesserten Trefferdarstellung durch Mehrwerte wie die Präsentation von intelligent ausgewählten Textschnipseln etc.
- Plagiatskonzept invers als Empfehlungsdienst mit Hinweis auf Zitationslücken
- Lesezeit bei E-books generieren und Ausleihdauer dementsprechend deklarieren
- Adaptierung von Geschäftsprozessen von Menschen- hin zu Maschinenlesbarkeit und auf der Zielgeraden wieder menschenlesbar! (neue hierarchisierte Strukturierung, Datenmodellierung, entitätenzentriert)
- finden von modularen KI-basierten Lösungen auf Basis der Kooperation (Bibliothek, Anbieter, Forschung)

Will man sich diesen Zielen zuwenden, müssen ebenfalls greifbar-konkrete Aufgaben angegangen werden:

- Aufbau einer Hauskompetenz (ML, NLP, NLU, Deep Learning etc.)
- KI-Infrastruktur (GPU, fachkompetente man-power etc.)
- bibliothekarische Prozesse konsequenter maschinenlesbar denken d.h.
  a) auf Grundlage des RDF-Modells,
  b) Linked-Data-Prinzipien
  c) entitäten- und nicht mehr dokumentenbasierte Erschließung (und auch keine Trennung von Inhalts- und Formalerschließung mehr)
  d) Einbeziehung von Struktur- und Relationsbeziehungen wie aus Ontologien
- KI-Algorithmen müssen interne und externe Strukturen bei der Datenaufbereitung berücksichtigen (ggf. nicht in der Anzeige, nur bei der Verarbeitung)
- Use Cases erstellen

Fazit

Der Beitrag listet viele Informationen auf und gibt zunächst wenig Anregungen für das Tagesgeschäft. Dies mag ungewollt an den Sponti-Spruch erinnern, gemäß: *Nutze die Chance, die Du nicht hast!* Wie hoch hängen die Früchte für eine KI-Perspektive der Informationseinrichtungen, wo liegt der Weg aus der vielleicht gefühlten Aussichtslosigkeit? Der Blick zu anderen Entwicklungen wie der kompletten Automatisierung von Fabriken mit Logistikkreisläufen bei Amazon, der Übermacht von Google und weiteren TechGiganten mag manchmal defätistisch stimmen, soll es aber gerade nicht. Aus Sicht der Autorinnen und Autoren ist bei diesem Technologiethema sehr stark das strategische Management gefordert, damit konkrete Lösungen mit Relevanz für das Kerngeschäft in greifbare Nähe rücken. Vermutlich auf nahezu jeder Arbeitsebene im Prozessablauf oder Organigramm einer Bibliothek ließen sich auch auf der Basis von KI-Technologien Automatisierungspotentiale mit vertretbaren Aufwänden erfassen, wenn man denn um die Potentiale von KI weiß. Weitere Pilotierungen von KI-Assistenz- bis hin zu Dialogsystemen werden Wegweiser sein, die wenige große Einrichtungen zu erbringen in der Lage sind. Für eine spartenübergreifende Marktdurchdringung von KI an allen Einrichtungen bedarf es aus unserer Sicht strategischer Papiere wie des White Papers, um den Entscheidern mit ihrer Prokura und Handlungsspielräumen eine mögliche Schwerpunktsetzung nahezulegen bis hin dazu, im Rahmen der Allokationsethik die Relevanz von KI für Informationseinrichtungen aufzuzeigen. Entsprechende Ressourcenallokationen als *conditio sine qua non* sind Impulsgeber, das Potential von KI auszuschöpfen und eine Roadmap zu entwerfen. Hierfür, den angestoßenen Transformationsprozess über Open Access hinaus auch in diesem Technologieumfeld zu denken, dafür wirbt der vorliegende Beitrag zum aufgezeigten Diskussionsstand.

Indice completo saltat scriptor pede laeto

Für die Autorinnen und Autoren

Dr. Frank Seeliger
Leiter der Hochschulbibliothek TH Wildau
fseeliger@th-wildau.de
http://www.th-wildau.de/bibliothek.html