

Michael Franke-Maier, Anna Kasprzik,
Andreas Ledl, Hans Schürmann
Editorial

Es ist naheliegend, einen Band zur Qualität in der Inhaltserschließung damit zu beginnen, den Begriff *Qualität* als solchen zu reflektieren. Wenn wir Qualität gemäß der *ISO-Norm 9000:2015-1* definieren als „Grad, in dem ein Satz inhärenter Merkmale eines Objekts Anforderungen erfüllt“¹, was sind dann die Qualitätsanforderungen an die Inhaltserschließung? Auf diese Frage gibt es zweifellos keine eindeutige und allgemeingültige Antwort – zumindest war das das Ergebnis der Diskussion unter den Herausgeber:innen dieses Bandes. Dementsprechend beleuchten wir die Frage im Folgenden aus vier unterschiedlichen Perspektiven: zunächst aus einer eher *idealistischen*, im Berufsethos verankerten, dann aus einer rein *fachlichen* und schließlich aus einer *praktischen* Perspektive. Die vierte, die *deskriptive* bzw. *historische*, geht als eigener Aufsatz von Andreas Ledl in den Band ein.

1 Qualität beginnt im Kopf

Eine erste, *idealistische* Annäherung an eine Antwort kann aus der charakteristischen Funktion von Bibliotheken hergeleitet werden – nämlich aus der Kernaufgabe, Ressourcen für Bildung, Lehre, Forschung und zur Unterhaltung zugänglich zu machen. Nicht zuletzt wird im *IFLA Code of Ethics* dieses Verfügbarmachen als zentrales Anliegen bezeichnet: „The core mission of librarians and other information workers is to ensure access to information [...]“² Diese Mission kann sogar so weit gefasst werden, dass sie auch den *access to misinformation* bzw. *disinformation* miteinschließt.³ Das ist aus Sicht der

1 Vgl. dazu z. B. Zollondz, Hans-Dieter u. a.: *Lexikon Qualitätsmanagement*. Berlin, Boston: De Gruyter Oldenbourg 2016, S. 857–858.

2 International Federation of Library Associations: *IFLA Code of Ethics for Librarians and other Information Workers*, full version. <https://www.ifla.org/publications/node/11092#access-to-information> (16.11.2020).

3 Zu den feinen Unterschieden bzgl. der beiden Begriffe sowie weiterer verwandter Begriffe siehe z. B.: Turčilo, Lejla und Mladen Obrenović: *Fehlinformationen, Desinformationen, Malinformationen. Ursachen, Entwicklungen und ihr Einfluss auf die Demokratie*. Demokratie im Focus #3. Heinrich Böll Stiftung, 2020, S. 8–9. <https://www.boell.de/de/2020/08/25/fehlinformationen-desinformationen-malinformationen> (1.12.2020). Zur Diskussion um die

Informationstheorie von Shannon und Weaver aus dem Jahr 1948/1949 folgerichtig, da nach ihr der Begriff *Information* zunächst keine Wertigkeit besitzt: „In fact, two messages, one of which is heavily loaded with meaning and the other of which is pure nonsense, can be exactly equivalent, from the present viewpoint, as regards information.“⁴

Vor dem Hintergrund der demokratischen und gesellschaftspolitischen Funktion⁵ von Bibliotheken wird schnell klar, dass solche Information mindestens zu kuratieren bzw. zu kontextualisieren ist. Damit bewegen wir uns im Kontext der Bedeutungszuschreibung, des Verstehens, der Bewertung und Einordnung von Information, nicht zuletzt also der Organisation von Wissen, der *knowledge organization* – einer unbestrittenen Kernaufgabe von Bibliotheken.

Wissensorganisation, sowohl auf der Ebene der Begriffsbildung im Normdatenbereich⁶ als auch bei der Inhaltsanalyse von Dokumenten, ist mit dem im bibliothekarischen Berufsethos impliziten Neutralitätsgebot behaftet, welches sich z. B. auch in den *Regeln für die Schlagwortkatalogisierung*⁷ ausdrückt, letztlich aber von der Perspektive und der Sozialisation der erschließenden Person, im Speziellen ihrer politischen, fachlichen und bibliothekarischen, abhängig bleibt.⁸

Bereitstellung z. B. rechtsextremistischer Literatur in Bibliotheken, die Jörg Sundermeyer, Verleger und Gründer des Verbrecher-Verlages, während des 107. Deutschen Bibliothekartages in Berlin anstieß und die anschließend in diversen Bibliothekszeitschriften und z. B. auf der Digitalkonferenz #vBIB20 weitergeführt wurde, vgl.: Sundermeier, Jörg: Rechte Verlage und ihre Produkte. Sollten Bücher aus rechten Verlagen im Bestand geführt werden? In: BuB – Forum Bibliothek und Information (2018) Jg. 70 H. 6. S. 331–333. <https://b-u-b.de/wp-content/uploads/2018-06.pdf> (16.11.2020).

⁴ Shannon, Claude E. und Warren Weaver: *The mathematical theory of communication*. Urbana: The University of Illinois Press, 1964, S. 8. https://pure.mpg.de/rest/items/item_2383164/component/file_2383163/content (16.11.2020).

⁵ Vgl. dazu z. B. die Positionen des Deutschen Bibliotheksverbands DBV zu Bibliotheken und Demokratie: <https://www.bibliotheksverband.de/dbv/themen/bibliotheken-und-demokratie.html> (16.11.2020).

⁶ Vgl. dazu z. B. den Vortrag von Guido Bee auf der GNDCon 2018 zu *gefährlichen Wörtern* in der GND. https://wiki.dnb.de/download/attachments/145591706/GNDCon_Folien_D2_Bee.pdf (9.12.2020).

⁷ Vgl. dazu RSWK, § 4,2: „Der Standpunkt oder eine Weltanschauung des Verfassers wird im Allgemeinen nicht berücksichtigt.“

⁸ Nicht zuletzt deswegen gibt es Initiativen wie das Cataloging Ethics Steering Committee, <https://sites.google.com/view/cataloging-ethics/home> (14.12.2020). Im zweiten Entwurf des Cataloging Code of Ethics heißt es bspw.: „We recognize that cataloguing is not a neutral act and endorse critical cataloguing as an approach to our shared work with the goal of making metadata more inclusive and accessible.“ <https://docs.google.com/document/d/1tnTetySv6mgCa4RfScaffykQdtViZZ8M5nwgowZboQk/edit> (14.12.2020).

Die Idee, dass Algorithmen eine Art technische Neutralität oder Objektivität besitzen könnten, die die des Menschen übertrifft, wurde in neuester Zeit immer wieder relativiert.⁹ Ein aktuelles Beispiel dafür, das beim Verfassen dieses Editorials gerade in den sozialen Medien Aufmerksamkeit hervorrief, war der Fall der Top-Forscherin Özlem TÜreci, die zusammen mit Uğur Şahin für das Unternehmen BioNTech einen Impfstoff gegen COVID-19 entwickelte. Die Google-Suche kategorisierte TÜreci lediglich als Şahins Ehefrau, obwohl sie ebenso maßgeblich an der Erforschung des Impfstoffes beteiligt war.¹⁰ Algorithmen können offensichtlich in Daten und Texten wiederkehrende Haltungen und deren Relationen perpetuieren, ohne dass sie dabei eine Möglichkeit haben, diese zu reflektieren und ggf. zu korrigieren.

Angesichts dieser Komplikationen, die gerade bei lernenden Verfahren dauerhaft als Verzerrung der Bewertungsbasis der Algorithmen ins Gewicht fallen können,¹¹ ist es zumindest beim aktuellen Stand der Wissenschaft in vielen Fällen notwendig, dass im Nachgang eine intellektuelle Kontrolle und Anreicherung der Ergebnisse mit Kontextinformationen erfolgt. Und selbst bei Fortschreiten und Optimierung der Algorithmen stellt sich die Frage, ob die besten Ergebnisse nicht doch durch ein Zusammenwirken von Mensch und Maschine erzielt werden können.

In den letzten Jahrzehnten sind auch im bibliothekarischen Kontext immer größer werdende bibliografische Mega-Indices, Knowledge Bases und Zitationsdatenbanken sowie (digitale und digitalisierte) Sammlungen von Forschungs- und Kulturdaten entstanden. Mit der Bereitstellung dieser Informationsquellen geht eine große Verantwortung bezüglich ihrer Zugänglichkeit, Nachnutzbarkeit und Interpretationsmöglichkeiten einher. Im Zusammenhang mit Inhaltserschließung liest man das Wort Verantwortung indessen eher selten, obwohl mit ihr der Zugang zu Information nachhaltig festgelegt wird – sie dient letztlich der bestmöglichen Versorgung mit relevanten Inhalten in Bezug auf Bildungsziele, Lehrinhalte, Forschungsfragen oder Unterhaltungsvorlieben. Stattdessen geht es doch meist um die Kostenersparnis oder um eine Effizienzsteigerung, die durch eine Automatisierung herbeigesehnt wird, weniger aber um eine

⁹ Die Autorin Cathy O’Neil weist in ihrem Buch *Weapons of Math Destruction* sogar auf potenzielle Gefahren der vermuteten Objektivität von Algorithmen hin – dies aber nur nebenbei.

¹⁰ Rainer, Anton: Ungenaue Algorithmen. Google stuft Impfforscherin zur „Ehefrau“ herab. Spiegel Netzwelt vom 10.11.2020, 16:29. <https://www.spiegel.de/netzwelt/web/biontech-impfstoff-wie-google-eine-impf-forscherin-zur-ehefrau-macht-a-e8ace050-8329-4706-9543-8a4e3ff6cd34> (16. November 2020).

¹¹ Vgl. dazu z. B. Knight, Will: The Dark Secret at the Heart of AI. In: MIT Technology Review, 11. April 2017. <https://www.technologyreview.com/2017/04/11/5113/the-dark-secret-at-the-heart-of-ai/> (14.12.2020).

Orientierung an Qualitätskriterien über die Basisfunktionen der Literatursuche hinaus. Letztlich kann in diesem Zusammenhang wieder der *IFLA Code of Ethics* bemüht werden – in welchem der bibliothekarische Beruf als nach den „highest standards of service quality“¹² strebend charakterisiert wird –, um den Appell an die strategische Managementebene zu formulieren, dass zumindest die Relevanz der Inhaltserschließung anerkannt und entsprechend die Orientierung an einem geeigneten Qualitätsstandard handlungsleitend für die Ausrichtung unserer Bibliothekslandschaft wird. Kurz gesagt: „Qualität beginnt im Kopf.“¹³

Idealerweise drückt sich diese Anerkennung mindestens darin aus, dass auf Personal- und Sachebene ausreichende Mittel für eine an überprüfbaren Qualitätsstandards orientierte Inhaltserschließung bereitgestellt werden – das gilt zunächst einmal sowohl für eine intellektuelle Inhaltserschließung durch eine kooperativ agierende Community als auch für einen halb- oder vollautomatischen Einsatz von Algorithmen. Die Automatisierung der Inhaltserschließung durch den Einsatz moderner semantischer Technologien und Technologien der Künstlichen Intelligenz muss allerdings noch viel weitreichender als bisher als transformative Daueraufgabe anerkannt und etabliert werden, um in Forschungsprojekten entwickelte Prototypen konsequent und dauerhaft in die produktiven Erschließungsworkflows von (wissenschaftlichen) Bibliotheken integrieren zu können.¹⁴ Erst dann sind die Kapazitäten gegeben, um praxistaugliche Lösungen für eine intelligente(re) Verzahnung von intellektuellen Wissensorganisationsmethoden, automatisierter Inhaltserschließung und Retrievalalgorithmen zu finden und umzusetzen.

2 Garbage in, garbage out

Die zweite Herangehensweise zur Beantwortung der eingangs gestellten Frage nach geeigneten Qualitätskriterien für die Inhaltserschließung ist eine *fachliche* aus Sicht der Informationswissenschaften, der Informatik und anderen angrenzenden Disziplinen. Die informationswissenschaftliche Perspektive orientiert

12 IFLA Code of Ethics for Librarians and other Information Workers. <https://www.ifla.org/publications/node/11092#neutrality> (16.11.2020).

13 Dieses Motto war 2018 die Headline des Weiterbildungsprogramms der Deutschen Gesellschaft für Qualität.

14 Vgl. dazu Kasprzik, Anna u. a.: Automatisierte Sacherschließung als Produktivverfahren für wissenschaftliche Bibliotheken – Herausforderungen und Lösungsansätze. Konferenzfolien, #vBIB20, 28.5.2020. Folie 14. <https://nbn-resolving.org/urn:nbn:de:0290-opus4-173944> (18.11.2020). Mitschnitt: <https://doi.org/10.5446/36437>.

sich im Wesentlichen daran, wie gut Information *nach messbaren Kriterien* aufbereitet und bereitgestellt werden kann, um den Nutzer:innen alle Aktivitäten rund um eine erfolgreiche Literaturrecherche zu ermöglichen wie in den IFLA LRM *User Tasks*¹⁵ beschrieben: *find, identify, select, obtain, explore*.

Inbesondere durch den fortschreitenden Einsatz von maschinellen Methoden zur Digitalisierung und zur Automatisierung der Informationsbereitstellung bieten sich stochastische Größen aus dem Machine Learning und aus dem Information Retrieval als messbare Qualitätskriterien für die Informationsaufbereitung und -bereitstellung an. So geben zum Beispiel die Metriken *precision, recall* und der F1-Wert an, wie viele erwünschte Dokumente im Zuge einer Recherche gefunden werden, wie viele Dokumente unnützer Beifang sind und in welchem Verhältnis die beiden Größen stehen. Statistische Konfidenzwerte treffen eine Aussage dazu, wie sicher ein durch ein trainiertes Modell geliefertes Ergebnis (etwa ein Schlagwort zur inhaltlichen Beschreibung einer Ressource) zutrifft.

Inwiefern solche Größen den Anspruch, den Bibliotheken insgesamt an die Qualität ihrer Tätigkeiten und insbesondere ihrer Erschließung haben, adäquat abbilden, ist eine ganz andere Frage – und dieser vorgelagert ist eine noch tiefgreifendere Frage, die das Potential hat, das berufliche Aufgabenprofil von an Bibliotheken Tätigen weitreichend zu verändern: Die Frage, inwieweit bibliothekarische Inhaltsererschließung angesichts moderner Mustererkennungs- und Retrievalalgorithmen umgestaltet und neu gedacht werden muss, um gewinnbringend mit modernen Technologien verzahnt werden zu können.

Eine Weiterentwicklung der bibliothekarischen Inhaltsererschließung erfordert in jedem Fall zunächst einmal eine umfassende und kritische Aufklärung auf allen Ebenen der Bibliotheksorganisation, was moderne Technologien leisten können und was nicht, auf welchen Prinzipien sie aufbauen und welche Voraussetzungen für ihren Einsatz gegeben sein müssen.

Es lässt sich jedenfalls festhalten, dass ein System, das eine zielgerichtete und erschöpfende Suche auf hohem Niveau ermöglicht, von einer hohen Beschreibungsdichte und einer möglichst weitreichenden Kontextualisierung durch eine hochgradige Vernetzung mit weiteren Informationsquellen und zwischen für die Beschreibung relevanten Entitäten profitiert. Letzteres kann unter anderem durch eine Unterfütterung mit möglichst dicht vernetzten Wissensorganisationssystemen erreicht werden.¹⁶

¹⁵ IFLA Library Reference Model (LRM). <https://www.ifla.org/publications/node/11412> (2.12.2020).

¹⁶ Vgl. dazu auch Freyberg, Linda: Density of Knowledge Organization Systems. In: Knowledge Organization for Digital Humanities. Proceedings of Wissensorganisation 2017, S. 25–30. http://dx.doi.org/10.17169/FUDOCs_document_000000028863.2.

Und auch auf der Datenebene zeichnet sich immer deutlicher eine Wechselwirkung zwischen maschinellen Verfahren und einer hochqualitativen (für feingranuläre fachliche Qualitätsaspekte weiterhin auch intellektuellen) Datenaufbereitung ab: Maschinelle Verfahren (sowohl solche aus der Künstlichen Intelligenz als auch semantische Technologien) profitieren von formattechnisch sauberen, maschinenlesbaren und vernetzten Daten im Sinne der FAIR-Prinzipien *findable, accessible, interoperable, reusable*. Es gilt *garbage in, garbage out* – nicht umsonst verbringen *Data Scientists* ca. 60 % ihrer Zeit damit, Daten zu säubern und zu strukturieren.¹⁷ Sowohl in Bezug auf Maschinenlesbarkeit als auch in Bezug auf eine generell höhere Datenqualität, Konsistenz und Abdeckung erwartet Bibliotheken folglich in den nächsten Jahren noch eine beträchtliche Anstrengung.¹⁸

3 Ergebnisoptimierung durch Erwartungsreduktion?

Die dritte Herangehensweise zur Beantwortung der Frage nach geeigneten Qualitätskriterien für die Inhaltserschließung ist eine *praktische* aus Sicht der Bibliotheken.¹⁹ Dabei werden aus der Nutzendenperspektive die unterschiedlichsten Ansprüche und Forderungen an qualitativ hochstehende Inhaltserschließung gestellt. Es macht einen Unterschied, ob die Nutzung der Daten durch Menschen oder durch datenverarbeitende Software erfolgen soll. Die Regelwerke für die Erschließung, Ergebnisse aus Nutzungsumfragen und Tracing von Nutzungsverhalten liefern hier durchaus widersprüchliche Anforderungen,

17 CloudFlower: Data Science Report 2016, S. 6. https://visit.figure-eight.com/rs/416-ZBE-142/images/CrowdFlower_DataScienceReport_2016.pdf (9.12.2020). Vgl. zu diesem Thema auch eine Aussage von Karen Foster in einem Blogbeitrag zur nationalen Datenstrategie des JISC: *Being able to gather and analyse data is one thing, but making sure that data is ‚clean‘ and meaningful is quite another*. <https://www.jisc.ac.uk/blog/the-national-data-strategy-a-golden-opportunity-to-level-up-the-use-of-data-and-innovation-08-dec-2020> (9.12.2020).

18 Vgl. dazu auch Kasprzik, Anna: Putting Research-based Machine Learning Solutions for Subject Indexing into Practice, Abschnitt 2 Challenges and Potentials. In: *Qurator 2020. Proceedings of the Conference on Digital Curation Technologies*. Hrsg. von Paschke u. a. Berlin 2020, S. 3 ff. http://ceur-ws.org/Vol-2535/paper_1.pdf (9.12.2020).

19 „Ergebnisoptimierung durch Erwartungsreduktion?“ ist der Claim eines Berliner Street-Art-Kunstwerks, s. Franke-Maier, Michael: Qualitätsanforderungen zwischen intellektueller und algorithmisierter Inhaltserschließung. 22. Verbundkonferenz des Gemeinsamen Bibliotheksverbundes, Kiel am 29. August 2018, Konferenzfolien, Folie 6. https://verbundkonferenz.gbv.de/wp-content/uploads/2018/09/Franke-Maier_Qualität_MFM_final.pdf (14.12.2020).

obwohl alle Zugänge dasselbe Ziel einer guten Informationsversorgung verfolgen. Weder das Bibliothekspersonal noch die Algorithmen können jedoch widersprüchlichen Anforderungen gerecht werden. Insofern ist es entscheidend, dass man sich bei der Formulierung von Qualitätskriterien auf den Bezugsrahmen verständigt. Letztlich ist die Festlegung, was Qualität in der Inhaltserschließung sichert, von der Entscheidung abhängig, wie die Bibliotheken den Zugang zu Information gewährleisten wollen.²⁰ Je größer die Metadatenmenge, desto bedeutender werden maschinelle Verfahren. Je vernetzter die Daten sind, desto standardisierter sollten die Codierungen sein. Je kleiner die zu erschließende Sammlung ist, desto lokaler kann die Erschließung sein. Die Liste ließe sich beliebig erweitern, wobei schnell klar wird, dass sich hier Widersprüche ergeben. Leidtragend ist das Bibliothekspersonal, das zum einen die Differenz zwischen dem Rechercheangebot und der Nutzungserwartung ausgleichen muss und zum anderen die Hoheit über die Metadatenpflege zunehmend verliert. Insofern wundert es nicht, dass das Bibliothekspersonal zunehmend in die Beratung und Betreuung von Kund:innen verschoben und die Datenpflege der Information and Communication Technology (ICT) zugeschrieben wird. Datenpflege reduziert sich dann auf Stichprobenanalyse und Erarbeitung von Massenkorrekturen systematisch auftretender Fehler. Aber auch dieser Bereich hat mit intellektueller Aufarbeitung nicht viel zu tun. Vielmehr geht es darum, die Indexierungsfehler nicht ausufern zu lassen. Intellektuell wird dann nur noch in Einzelfällen eingegriffen. Das kann funktionieren, wird aber in dem Moment problematisch, wenn dieselben Metadatensätze in unterschiedlichen Kontexten verwendet werden. Dann bedarf es in den Normdatensätzen Verwendungscodes und in den bibliografischen Datensätzen Provenienzkennzeichnungen, um sicherzustellen, dass Retrieval auch spezifisch nutzbar ist. Gerade kleine Bibliotheken mit spezifischem Bestand sehen sich sonst gezwungen, ihre Daten aus den Datennetzen wieder in lokale Datenbanken zu separieren.²¹

In dieser Spannung zwischen semantischer Metadatenbildung und technischer Datenverarbeitung muss ein Weg der Verständigung gefunden werden, wer mit welchen Vorzügen zu einem optimalen Qualitätslevel für die Informationsbereitstellung in einem multidimensionalen Umfeld beitragen kann. Das wiederum geht nur über die Formulierung eines Bezugsrahmens, an dem sich alle Beteiligten orientieren können.

20 Vgl. dazu auch Schürmann, Hans: Sacherschließung nach RDA. In: 027.7 Zeitschrift für Bibliothekskultur (2015) Bd. 3 H. 2. S. 74–80. <https://doi.org/10.12685/027.7-3-2-64>.

21 Vgl. dazu auch Schürmann, Hans: Subject indexing in a digital age. Presented at the General Assembly of the European Theological Libraries (47th General Assembly of BETH), Nürnberg 2018. <https://doi.org/10.5281/zenodo.1420492>.

4 Qualität ist multidimensional

Spätestens seit 2017, mit der Veröffentlichung des angepassten Erschließungskonzepts der Deutschen Nationalbibliothek (DNB) unter dem Titel *Grundzüge und erste Schritte der künftigen inhaltlichen Erschließung von Publikationen in der Deutschen Nationalbibliothek*²² und der anschließenden Diskussion z. B. während des Bibliothekartags 2018, ist das Interesse zu Fragen der Qualität der Inhaltserschließung im DACH-Raum stark gestiegen. Der vorliegende Band führt diese Diskussion insofern fort, da er aus ihr erwachsen ist und als strukturierende Grundlage für den Call for Papers und für sein Narrativ der (leicht veränderten) Kartierung von Qualitätsdimensionen folgt, wie sie auf der 22. Verbundkonferenz des Gemeinsamen Bibliotheksverbundes in Kiel am 29. August 2018 vorgestellt wurde.²³

Zur Vervollständigung der eingenommenen Perspektiven des Editorials und zur historischen Anbindung an die bibliothekarische Qualitätsdiskussion generell beginnt der Band zunächst mit einem Beitrag von Andreas Ledl zur *Qualität in der Inhaltserschließung – Ein Überblick aus 50 Jahren (1970–2020)*. Dieser zeichnet die großen Linien des Qualitätsdiskurses im deutschsprachigen Raum seit den 1970er Jahren chronologisch nach. Dabei kristallisiert sich heraus, dass die Frage nach der Güte der Sacherschließung schon immer unterschiedlich beantwortet wurde und mit der Weiterentwicklung des Bibliothekswesens kontinuierlich neue Perspektiven hinzutraten. Ob Menschen oder Maschinen, Bibliothekar:innen, Fachcommunitys oder Lai:innen alle oder nur ausgewählte Bestände kooperativ oder isoliert, intellektuell oder automatisiert, verbal oder klassifikatorisch, gleichordnend oder syntaktisch, mit Schlagwörtern oder Tags inhaltlich erschließen sollen – stets wurden auch qualitative Argumente angeführt, um bestimmte Positionen zu vertreten, zu verwerfen oder zwischen ihnen zu vermitteln.

Im Folgebeitrag *Fit for Purpose – Standardisierung von inhaltserschließenden Informationen durch Richtlinien für Metadaten* fokussiert sich Joachim Laczny

22 Deutsche Nationalbibliothek: *Grundzüge und erste Schritte der künftigen inhaltlichen Erschließung von Publikationen in der Deutschen Nationalbibliothek*, Stand Mai 2017. <https://www.dnb.de/SharedDocs/Downloads/DE/Professionell/Erschliessen/konzeptWeiterentwicklungInhaltserschliessung.pdf> (17.12.2020).

23 Siehe Franke-Maier, Michael: *Qualitätsanforderungen zwischen intellektueller und algorithmisierter Inhaltserschließung*. Folie 9–12. https://verbundkonferenz.gbv.de/wp-content/uploads/2018/09/Franke-Maier_Qualität_MFM_final.pdf (14.12.2020); und Franke-Maier, Michael: *Anforderungen an die Qualität der Inhaltserschließung im Spannungsfeld von intellektuell und automatisch erzeugten Metadaten*. In: *ABI Technik* (2018), Bd. 38 H. 4. S. 327–331. <https://doi.org/10.1515/abitech-2018-4005>.

auf die Frage, inwieweit sich die Ausarbeitung, Veröffentlichung und Anwendung von Metadaten-Richtlinien auf den Qualitätsanspruch an die Inhaltserschließung auswirken. Er gibt dabei einen tiefgründigen Einblick in die wissenschaftliche Diskussion zu Qualitätskriterien der letzten zehn bis zwanzig Jahre und kommt zu der Schlussfolgerung, dass sich solche Richtlinien positiv auf „nutzungsorientierte Metadatenangebote“ von Bibliotheken auswirke – vor allem dann, wenn die Betreuung in der Kompetenz des Bibliothekspersonals liege, da dann „zeitgemäße Datendienstleistungen nutzungsorientiert und auf bibliothekarisch angemessenem Qualitätsniveau“ angeboten werden könnten.

Was insbesondere Bibliotheken mit zentralem Auftrag wie die Deutsche Nationalbibliothek (DNB) tun können, um die Qualitätsansprüche an ihre Dienstleistungen fortlaufend zu bedienen, und mit welchen Maßnahmen die DNB der Herausforderung begegnet ist, trotz der teilweise kontroversen Diskussion der Ergebnisse in der Anfangsphase seit 2010 zunehmend den Einsatz von maschinellen Methoden zu etablieren, schildern Ulrike Junger und Frank Scholze in ihrem Beitrag *Neue Wege und Qualitäten – Die Inhaltserschließungspolitik der Deutschen Nationalbibliothek*. Weiterhin geben sie einen Einblick in die aktuellen und zukünftigen Entwicklungs- und Handlungsfelder der DNB in diesem Bereich.

Auch wenn in ihrem Beitrag nicht auf das im November 2020 verabschiedete WikiLibrary-Manifest²⁴ Bezug genommen wird, das die Kooperation zwischen der Deutschen Nationalbibliothek und Wikimedia festigt und gleichzeitig weitere Bibliotheken zur Teilhabe auffordert, werden sich künftig sicherlich zahlreiche Schnittpunkte zur Erschließungspolitik ergeben. Insbesondere ein Austausch zwischen DNB und Wikimedia in Bezug auf Möglichkeiten, Qualitätsprobleme in sehr großen Normdateien zu erkennen und zu beheben, wird sich lohnen, wie es die beiden folgenden Beiträge in diesem Band zeigen. Lydia Pintscher, Peter Bourgonje, Julián Moreno Schneider, Malte Ostendorff und Georg Rehm beschreiben in ihrem Aufsatz *Wissensbasen für die automatische Erschließung und ihre Qualität am Beispiel von Wikidata* die Aspekte von Datenqualität, die Arten von Qualitätsproblemen sowie vorhandene und zukünftige Werkzeuge und Prozesse zum Finden und Beheben dieser Probleme, was aufgrund der Offenheit von Wikidata sowie der großen Zahl an Editierenden eine besondere Herausforderung darstellt. Weiterhin gibt der Aufsatz einen praktischen Einblick in die Nachnutzung von Wikidata durch das Deutsche Forschungszentrum für Künstliche Intelligenz.

²⁴ Vgl. https://www.dnb.de/DE/Professionell/ProjekteKooperationen/_content/gndWikimedia_teaaser.html und <https://www.wikimedia.de/projects/wikilibrary-manifest/> (11.1.2021).

Ein weniger offenes, allerdings zentrales Instrument für die verbale Inhaltserschließung von Bibliotheken im DACH-Raum ist die Gemeinsame Normdatei (GND), die kooperativ geführt und von der DNB koordiniert wird. Die Qualität eines derart massiv nachgenutzten Wissensorganisationssystems ist natürlich essenziell für die Qualität der daraus resultierenden Erschließung durch Mensch und Maschine. Im Spannungsfeld zwischen Universalgültigkeit und flexibler Einsatzfähigkeit ist es eine anspruchsvolle Aufgabe, die formale und inhaltliche Qualität kontinuierlich zu überwachen und sicherzustellen. Esther Scheven gibt in ihrem Beitrag *Qualitätssicherung in der GND* eine Übersicht über die Entstehung der GND und über die Regelwerke, mit denen sie interagiert, und beleuchtet eine Reihe von Aspekten und Voraussetzungen, sowie fachliche und organisatorische Maßnahmen für die Qualitätssicherung der GND. In diesem Zusammenhang geht sie auch darauf ein, wie diese Maßnahmen angepasst und erweitert werden müssen, wenn die GND nun sukzessive zur Anwendung für weitere Gruppen außerhalb des Bibliothekswesens geöffnet wird.

Als fünfter Beitrag folgt unter dem Titel *Qualitätskriterien und Qualitätssicherung in der inhaltlichen Erschließung* ein bisher unveröffentlichtes und für diesen Band überarbeitetes Thesenpapier des Expertenteams RDA-Anwendungsprofil für die verbale Inhaltserschließung (ET RAVI). Es wurde im Auftrag des Standardisierungsausschusses erarbeitet und diesem Ende 2019 vorgelegt. Es schließt an die beiden vorangegangenen Beiträge insofern an, als es sich auch mit der Qualitätsdimension der „[r]egelbasierte[n] Produktion von Normdaten“ beschäftigt. Gleichzeitig geht es darüber hinaus und nennt neben den zwei „Grunddimensionen: Verlässlichkeit und Transparenz“ drei weitere Dimensionen, die im Fortgang dieses Bandes immer wieder eine Rolle spielen werden, die „Verwendungsregeln für die Ressourcenbeschreibung“, die „[t]ransparente Auswertung für Retrieval und Anzeige“ sowie die „Öffnung der Daten für die Nachnutzung in anderen Wissenschaftskontexten jenseits des klassischen Retrievals“. Neben vielen weiteren Elementen der Qualitätserzeugung wird im Thesenpapier das Augenmerk auch auf die „Verknüpfung mit anderen Normdaten [...], um logische Schlussfolgerungen durch Mensch und Maschine sowie die Navigation in den Daten zu ermöglichen“, sowie auf die „Auswertung von Cross-Konkordanzen und Relationierungen zur Homogenisierung der Vielzahl unterschiedlicher inhaltlicher Metadaten“ gelegt.

Genau darum geht es im Folgebeitrag *coli-conc – Eine Infrastruktur zur Nutzung und Erstellung von Konkordanzen* von Uma Balakrishnan, Stefan Peters und Jakob Voß, der sich mit der Interoperabilität zwischen Wissensorganisati-

onssystemen (KOS²⁵) beschäftigt. Als einen der Hauptgründe zur Entwicklung der Infrastruktur für einen vereinfachten Zugriff auf KOS und die computergestützte Erstellung von Mappings zwischen ihnen nennen die Autor:innen die Verbesserung der Qualität von Konkordanzen. Die coli-conc-Datenbank,²⁶ der Hauptbestandteil der Infrastruktur, beinhaltet über 370 000 Mappings zwischen unterschiedlichen KOS in einem einheitlichen Format und steht zur freien Nutzung zur Verfügung. Die Infrastruktur basiert auf Modulen zum einheitlichen Zugriff auf KOS, Konkordanzen und Konkordanzbewertungen. Alle Komponenten sind über einheitliche und offene APIs miteinander verbunden, sodass eine verteilte Infrastruktur für jede Funktion entsteht, indem jede Komponente einzeln erweiterbar und nachnutzbar ist.²⁷ Die Module sind in einer Webanwendung zur effektiven Bearbeitung von Konkordanzen zusammengeführt: in dem Mapping-Tool Cocoda²⁸ sind mehr als 170 KOS eingebunden. Die Infrastruktur bietet die Möglichkeit, den Aufbau eines umfassenden KOS-Management-Tools mit Hilfe eines Konzept-Hubs zu realisieren.

Solche Tools können also enorme Fortschritte für die Titel- und die Normdatenpflege darstellen. Dass es jedoch oft für die Qualität von Werkzeugen und Algorithmen nicht eine einzige Kennzahl gibt, sondern verschiedene Dimensionen berücksichtigt werden müssen und entsprechend viele Metriken denkbar sind, die jeweils unterschiedliche Aspekte und Perspektiven abbilden, zeigen Clemens Neudecker, Mike Gerber, Konstantin Baierer, Julián Moreno Schneider, Karolina Zaczynska und Georg Rehm exemplarisch für die optische Texterkennung in ihrem Beitrag *Methoden und Metriken zur Messung von OCR-Qualität für die Kuratierung von Daten und Metadaten*. Sie kommen zu dem Schluss, dass eine hochqualitative *Optical Character Recognition* (OCR) eine wertvolle Zuarbeit für die Erschließung darstellen kann. Um echte Mehrwerte sowohl für die formale Metadatenanreicherung als auch für eine tiefere Inhaltserschließung zu erhalten, reicht laut den Autor:innen die reine OCR allerdings nicht aus, sondern muss mit Layoutanalysen und sprachwissenschaftlichen Verfahren verschränkt werden, die es erlauben, Form und Inhalt zueinander in Beziehung zu setzen.

Zwei Beiträge des Bandes beschäftigen sich explizit mit formatabhängigen Aspekten der Qualität bzw. mit der formalen Datenqualität. Jakob Voß legt dabei seinen Fokus auf die *Datenqualität als Grundlage qualitativer*

25 Die Abkürzung KOS für Wissensorganisationssysteme kommt von der englischen Sprachvariante Knowledge Organization Systems.

26 <https://coli-conc.gbv.de> (21.1.2021).

27 <https://coli-conc.gbv.de/publications/#software> (21.1.2021).

28 <https://coli-conc.gbv.de/cocoda/> (21.1.2021).

Inhaltserschließung und definiert vier Anforderungen an die formale Datenqualität: Sie erfordere Regeln, Dokumentation, Validierung und Praxisbezug. Weiterhin sei „[d]ie Form von Daten [...] keine rein technische Nebensächlichkeit“, da „eine inhaltlich hervorragende Erschließung unbrauchbar [werde], wenn die Erschließungsdaten aufgrund inkompatibler Datenformate nicht verwendet werden können“. Dem eher theoretischen Ansatz von Voß schließt sich eine empirische Analyse von Péter Király und Rudolf Ungváry an. Sie vergleichen in ihrem Aufsatz *Bemerkungen zu der Qualitätsbewertung von MARC-21-Datensätzen* mehrere Gesamtkataloge von National- und Universitätsbibliotheken aus Ungarn, Deutschland und anderen europäischen Ländern daraufhin, ob und in welchem Maße die Möglichkeiten des Datenformates MARC 21 ausgeschöpft werden. Aus Sicht des Datenformates diskutieren sie Möglichkeiten, wie anhand der effektiven Codierungen Aussagen über die Qualität der Inhaltserschließung gemacht werden können. Dabei stellt sich heraus, dass das Potential des Formates zur Kontrolle der Feldbelegungen und zur Behebung von Inkonsistenzen bei weitem nicht ausgeschöpft ist.

Sina Menzel, Hannes Schnaitter, Josefine Zinck, Vivien Petras, Clemens Neudecker, Kai Labusch, Elena Leitner und Georg Rehm nähern sich in ihrer Studie *Named Entity Linking mit Wikidata und GND – Das Potenzial handkuratierter und strukturierter Datenquellen für die semantische Anreicherung von Volltexten* dem Thema Qualität nicht über Bibliotheksdatenformate, sondern über die Frage nach der Eignung verschiedener Normdateien für die automatisierte Erkennung von Entitäten in historischen Texten. Bei der Studie wird dabei auf den Volltextkorpus der digitalisierten Zeitungsseiten der Staatsbibliothek zu Berlin – Preußischer Kulturbesitz (SBB) mit der stolzen Zahl von 2 078 127 Zeitungsseiten zurückgegriffen, um die beiden großen Wissensbasen Wikidata und GND hinsichtlich ihres Potenzials für die Inhaltserschließung zu vergleichen, und zwar in Bezug auf etablierte Verfahren für Named Entity Recognition und Named Entity Linking.

Fabian Steeg und Adrian Pohl betonen in ihrem Beitrag *Ein Protokoll für den Datenabgleich im Web am Beispiel von OpenRefine und der Gemeinsamen Normdatei (GND)*, dass Normdaten als externe Datenquellen eine wichtige Rolle für die Qualität der Inhaltserschließung spielen. Sie stellen am Beispiel von *lobid-gnd*²⁹ die Möglichkeiten einer Reconciliation-API anhand ihrer Verwendung in OpenRefine zum Abgleich mit der GND dar. Anschließend wird als Ausblick über die Arbeit der Entity Reconciliation Community Group des World Wide Web Consortiums (W3C) berichtet, in der ein Protokoll zur Standardisierung dieser Funktionalität entwickelt wird. Der Beitrag behandelt so die Quali-

²⁹ <https://lobid.org/gnd> (21.1.2021).

tät des Datenaustauschs (speziell die Beschreibung und Verbesserung des Protokolls selbst im Rahmen einer Standardisierung), die Qualität von Datensätzen im Kontext anderer Datensätze (speziell die Einheitlichkeit der Verknüpfung und Ermöglichung von Datenanreicherung durch Nutzung von Normdaten statt der Verwendung von Freitext), sowie die Qualität von Werkzeugen und Algorithmen (speziell von Werkzeugen, die das Protokoll nutzen, konkret OpenRefine, sowie durch das Protokoll ermöglichte Arbeitsabläufe).

Gleich drei Beiträge beschäftigen sich mit der Qualität der Inhaltserschließung im Zusammenhang mit der Recherche und beleuchten diverse Aspekte von Online-Katalogen bzw. Discovery-Systemen. Heidrun Wiesenmüller betrachtet in ihrem Beitrag *Verbale Erschließung in Katalogen und Discovery-Systemen – Überlegungen zur Qualität* ausgehend von den Qualitätsdimensionen des Thesenpapiers des ET RAVI zwanzig Recherchewerkzeuge mit Blick auf die drei Funktionen von Inhaltserschließung: Zugang erhalten, orientieren und explorieren. Ihre Feststellungen kulminieren in dem zentralen Satz, dass „Bibliotheken und andere Informationseinrichtungen [...] darüber hinaus auch deshalb von Inhaltserschließung [profitieren], weil diese den Wert des Bestands steigert und ihn sozusagen ‚veredelt‘“. Dies gelte allerdings nur, wenn sich der Mehrwert dieser Inhaltserschließung in der Zielanwendung, also in den genutzten Recherchewerkzeugen, auch voll entfalten könne. Hier sieht Heidrun Wiesenmüller noch starke Defizite. Insofern belegt ihr Aufsatz den Bedarf für eine der technischen Ebene vorgelagerte strategische Ebene zur Gestaltung der Abstimmungsprozesse zwischen Spezialist:innen der Inhaltserschließung und der IT-Entwicklung. Genau hier setzt der Beitrag *Inhaltserschließung für Discovery-Systeme gestalten* von Jan Frederik Maas an: Nach einer Beschreibung der technologischen Grundlagen von Discovery-Systemen skizziert er einen Ansatz zur Umsetzung dieser Prozesse anhand von Personas und Scenario-Based Design. Dieser soll sowohl dazu dienen, auf beiden Seiten ein Bewusstsein für die Notwendigkeit von Synergieeffekten zu schaffen, als auch dazu, durch das Definieren von übergeordneten Zielen gemeinsame Leitlinien für eine Verschränkung der beiden bisher noch getrennt gedachten Geschäftsprozesse der Erschließung und des Retrievals festzulegen.

Inwiefern sich ein Blick durch die Brille des Retrievals tatsächlich als Maß für die Qualität von Inhaltserschließung interpretieren lässt, erörtern Christian Wartena und Koraljka Golub in ihrem Beitrag *Evaluierung von Verschlagwortung im Kontext des Information Retrievals*. Sie fassen zunächst zusammen, welche Zwecke Verschlagwortung erfüllen soll und kann, und beleuchten dann verschiedene Aspekte, die in die Bewertung eines Suchergebnisses einfließen können, etwa das Suchszenario, die verwendeten Thesauri oder Ontologien, oder die eingesetzten Rankingalgorithmen. Sie ziehen das Fazit, dass das Retrieval

aufgrund der vielen Einflussfaktoren als Qualitätsmaß für die Inhalterschließung nicht unproblematisch ist, dass es sich aber trotzdem lohnt, weitere Untersuchungen und Tests in diese Richtungen durchzuführen, um mehr über die Interaktion von Verschlagwortung und Retrievalsystemen herauszufinden.

Cyrus Beck legt mit seinem Aufsatz eine beobachtende Mikrostudie vor, in der er *Die Qualität der Fremddatenanreicherung FRED* anhand der Regeln für die Schlagwortkatalogisierung und der Grundaspekte von Erschließungsqualität Vollständigkeit, Richtigkeit, Präzision und Themenbildung untersucht. Er kommt zu dem Ergebnis, dass „[d]ie Anforderungen des Regelwerks [...] für die ausgewählten Fächer immerhin zu einem guten Teil erfüllt“ und „nur verhältnismäßig wenige falsche oder nicht präzise Schlagwörter importiert“ werden. Weiterhin gibt er einen Einblick in Fragen bzgl. der Nutzung von FRED in der neuen Umgebung der *Swiss Library Service Platform (SLSP)*.

Während FRED die Effizienz der Inhalterschließung steigert und damit die Quantität der Inhalterschließung erhöht, fragt der Beitrag *Quantität als Qualität – Was die Verbände zur Verbesserung der Inhalterschließung beitragen können* von Rita Albrecht, Barbara Block, Mathias Kratzer und Peter Thiessen explizit danach, inwiefern die Quantität der Inhalterschließung selbst als Qualitätskriterium gelten kann. Dabei stehen Bibliotheken und Bibliotheksverbände mit ihren Tools für den Datenaustausch im Zentrum der Analyse. Gefordert wird eine solide Basis für technische Infrastrukturen zur Sicherung der Qualität in der Datenübernahme, die letzten Endes nur im Dialog zwischen Entwickler:innen und Gremien erfolgreich geschaffen werden kann.

Der Band wird durch den Beitrag *Hybride Künstliche Intelligenz in der automatisierten Inhalterschließung* von Harald Sack abgerundet, in dem er zunächst kurz die Entwicklung der Künstlichen Intelligenz in den letzten acht Jahrzehnten und die Unterscheidung zwischen symbolischen (logikbasierten) und subsymbolischen (machine-learning-basierten) Ansätzen der Wissensrepräsentation umreißt. Anschließend beleuchtet er anhand von exemplarischen Problemstellungen, was ausgewählte State-of-the-Art-Methoden aus der symbolischen und subsymbolischen Künstlichen Intelligenz sowohl für die Inhalterschließung und angrenzende Gebiete als auch für die Suche und das Retrieval leisten könnten. Für die Erschließung greift er die verbale und klassifikatorische Inhalterschließung, das Transkribieren historischer Dokumente und die Pflege von Wissensorganisationssystemen und Wissensbasen heraus, für das Retrieval die semantische Suche, die Visualisierung von Suchergebnissen, die explorative Suche und Empfehlungssysteme. Er hält fest, dass kombinierte symbolische und subsymbolische Techniken der Wissensrepräsentation zwar einen qualitativen Quantensprung zu aktuell eingesetzten Technologien darstellen, dass jedoch die Erfüllung der diversen Erwartungshaltungen an Informationssysteme

nach wie vor eine Herausforderung ist und bleibt. Harald Sack prognostiziert, dass es wohl noch eine Weile dauern wird, bis automatisierte Verfahren die intellektuelle Erschließung abgelöst haben werden und dass der Weg dorthin eher über eine intelligente Kombination beider Herangehensweisen führen wird. Er mahnt jedoch auch an, dass die Geschwindigkeit dieser Transformation nicht allein von der technologischen Entwicklung, sondern vor allem auch vom notwendigen Commitment und von der allgemeinen Akzeptanz in der Zielcommunity (in dem Fall also dem Bibliotheks- und Archivwesen) gegenüber solchen innovativen Verfahren abhängt.

Die Beiträge dieses Bandes zur Qualität der Inhaltserschließung zeigen, dass sich Inhaltserschließung in einem multidimensionalen Raum bewegt. Neben die regelbasierte intellektuelle Verschlagwortung treten weitere Werkzeuge und Technologien, so dass die Praktiken zur Erreichung des Ziels – eine optimale Unterstützung der Recherche – einem Wandel unterliegen, der einer genauen Abstimmung bedarf, um hohe Ansprüche an Qualität halten zu können. Der Einsatz verschiedener Automatisierungsverfahren birgt ein großes Innovationspotenzial für die Inhaltserschließung, vielleicht mehr noch als für andere Aufgabenbereiche von Bibliotheken – insbesondere, seit Verfahren aus der Künstlichen Intelligenz ein Entwicklungsstadium erreicht haben, in dem sie realistische Optionen für die Praxis darstellen, wie in einigen Beiträgen beschrieben. Die Vielfalt der Ansätze und Perspektiven in diesem Band belegt ein breites Spektrum von Herangehensweisen an dieses komplexe und spannende Thema, das sicher auch in den kommenden Jahrzehnten eine dynamische Entwicklung durchlaufen wird. Insofern hoffen wir, dass dieser Band der Inspiration dient und darüber hinaus auch Lust macht, sich (weiterhin) mit Inhaltserschließung zu beschäftigen. Stay tuned!

5 Editorische Notiz und Danksagung

Es war uns ein Anliegen, diesen Band weitgehend geschlechtergerecht bzw. genderneutral zu gestalten, und wir haben uns auf die Verwendung neutraler Formulierungen oder auf den Einsatz des Gender-Doppelpunktes geeinigt. Letzterer wurde aufgrund der besseren Lesbarkeit für Screen-Reader gewählt.

Unser Dank gebührt allen am Band beteiligten Autor:innen für die hervorragende Zusammenarbeit, der Freien Universität Berlin für die Finanzierung der Open-Access-Kosten aus Mitteln ihres Open-Access-Publikationsfonds, weiterhin Sean Nowak für die Übersetzung zweier englischer Beiträge in die deutsche Sprache, Claudia Heyer und Elise Wintz vom Verlag De Gruyter Saur für die

Organisation des Korrektorats sowie die vielseitige Unterstützung und nicht zuletzt Klaus Gantert und Ulrike Junger für die Aufnahme des Bandes in die Reihe Bibliotheks- und Informationspraxis (BIPRA).

6 Literaturverzeichnis

- Bee, Guido: Gefährliche Wörter. Die GND und die problematischen Begriffe. GNDCon 2018, Deep Dive Session „Sachbegriffe. Kontinuität und Wandel – Herausforderungen für das Sachschlagwort in der GND“, 3. Dezember 2018. Konferenzfolien. https://wiki.dnb.de/download/attachments/145591706/GNDCon_Folien_D2_Bee.pdf (29.1.2021).
- Cataloging Ethics Steering Committee: Draft Cataloguing Code of Ethics, revised Sept. 2020. <https://docs.google.com/document/d/1tnTetySv6mgCa4RfScaffykQdtViZZ8M5nwgowZboQk/edit> (29.1.2021).
- CloudFlower: Data Science Report, 2016. https://visit.figure-eight.com/rs/416-ZBE-142/images/CrowdFlower_DataScienceReport_2016.pdf (29.1.2021).
- Deutscher Bibliotheksverband e. V.: Bibliotheken und Demokratie. <https://www.bibliotheksverband.de/dbv/themen/bibliotheken-und-demokratie.html> (29.1.2021).
- Deutsche Gesellschaft für Qualität: Qualität beginnt im Kopf. Das Weiterbildungsprogramm 2018 für alle Themen rund um Managementsysteme und Prozesse.
- Deutsche Nationalbibliothek: Grundzüge und erste Schritte der künftigen inhaltlichen Erschließung von Publikationen in der Deutschen Nationalbibliothek, Stand Mai 2017. <https://www.dnb.de/SharedDocs/Downloads/DE/Professionell/Erschliessen/konzeptWeiterentwicklungInhaltserschliessung.pdf> (29.1.2021). Deutsche Nationalbibliothek: Normdaten der Gemeinsamen Normdatei und Wikimedia. https://www.dnb.de/DE/Professionell/ProjekteKooperationen/_content/gndWikimedia_tteaser.html (29.1.2021).
- Foster, Karen: The National Data Strategy – a golden opportunity to level up the use of data and innovation for education and research. In: JISC Blog, 8.12.2020. <https://www.jisc.ac.uk/blog/the-national-data-strategy-a-golden-opportunity-to-level-up-the-use-of-data-and-innovation-08-dec-2020> (29.1.2021).
- Franke-Maier, Michael: Anforderungen an die Qualität der Inhaltserschließung im Spannungsfeld von intellektuell und automatisch erzeugten Metadaten. In: ABI Technik (2018) Bd. 38 H. 4. S. 327–331. <https://doi.org/10.1515/abitech-2018-4005>.
- Franke-Maier, Michael: Qualitätsanforderungen zwischen intellektueller und algorithmisierter Inhaltserschließung. 22. Verbundkonferenz des Gemeinsamen Bibliotheksverbundes, Kiel am 29. August 2018, Konferenzfolien. https://verbundkonferenz.gbv.de/wp-content/uploads/2018/09/Franke-Maier_Qualität_MFM_final.pdf (29.1.2021).
- Freyberg, Linda: Density of Knowledge Organization Systems. In: Knowledge Organization for Digital Humanities. Proceedings of Wissensorganisation 2017: 15. Tagung der Deutschen Sektion der Internationalen Gesellschaft für Wissensorganisation (ISKO) (WissOrg'17). Hrsg. v. Christian Wartena, Michael Franke-Maier und Ernesto De Luca. German Chapter of the ISKO / Freie Universität Berlin 2017, S. 25–30. https://doi.org/10.17169/FUDocs_document_00000028863.2

- International Federation of Library Associations: IFLA Code of Ethics for Librarians and other Information Workers, full version. <https://www.ifla.org/publications/node/11092> (29.1.2021).
- International Federation of Library Associations: IFLA Library Reference Model (LRM). <https://www.ifla.org/publications/node/11412> (29.1.2021).
- Kasprzik, Anna, Moritz Fürneisen und Timo Borst: Automatisierte Sacherschließung als Produktivverfahren für wissenschaftliche Bibliotheken – Herausforderungen und Lösungsansätze. #vBIB20, Session „KI als Aufgabe für Bibliotheken in Forschung, Lehre und Anwendung, oder: Zwischen Hype, Wirklichkeit und Durststrecke“, 28. Mai 2020. Konferenzfolien. <https://nbn-resolving.org/urn:nbn:de:0290-opus4-173944> (29.1.2021).
- Kasprzik, Anna, Moritz Fürneisen und Timo Borst: Automatisierte Sacherschließung als Produktivverfahren für wissenschaftliche Bibliotheken – Herausforderungen und Lösungsansätze. #vBIB20, Session „KI als Aufgabe für Bibliotheken in Forschung, Lehre und Anwendung, oder: Zwischen Hype, Wirklichkeit und Durststrecke“, 28. Mai 2020. Mitschnitt: <https://doi.org/10.5446/36437>.
- Kasprzik, Anna: Putting Research-based Machine Learning Solutions for Subject Indexing into Practice. In: Qurator 2020. Proceedings of the Conference on Digital Curation Technologies. Berlin, Germany, January 20th to 21st, 2020. Hrsg. v. Adrian Paschke, Clemens Neudecker, Georg Rehm, Jamal Al Qundus und Lydia Pintscher. Berlin 2020. http://ceur-ws.org/Vol-2535/paper_1.pdf (29.1.2021).
- Knight, Will: The Dark Secret at the Heart of AI. In: MIT Technology Review, 11. April 2017. <https://www.technologyreview.com/2017/04/11/5113/the-dark-secret-at-the-heart-of-ai/> (29.1.2021).
- O’Neil, Cathy: Angriff der Algorithmen. Wie sie Wahlen manipulieren, Berufschancen zerstören und unsere Gesundheit gefährden. München: Carl Hanser Verlag 2016.
- Rainer, Anton: Ungenaue Algorithmen. Google stufte Impfforscherin zur „Ehefrau“ herab. Spiegel Netzwelt vom 10.11.2020, 16:29. <https://www.spiegel.de/netzwelt/web/biontech-impfstoff-wie-google-eine-impf-forscherin-zur-ehefrau-macht-a-e8ace050-8329-4706-9543-8a4e3ff6cd34> (29.1.2021).
- Scheven, Esther und Julijana Nadj-Guttandin (Bearb.): Regeln für die Schlagwortkatalogisierung: RSWK. 4., vollständig überarb. Aufl. Erarb. von der Expertengruppe Sacherschließung im Auftrag des Standardisierungsausschusses. Leipzig, Frankfurt a. M.: Deutsche Nationalbibliothek 2017. <https://nbn-resolving.org/urn:nbn:de:101-2017011305> (29.1.2021).
- Schürmann, Hans: Sacherschließung nach RDA. In: 027.7 Zeitschrift für Bibliothekskultur (2015) Bd. 3 H. 2. S. 74–80. <https://doi.org/10.12685/027.7-3-2-64>.
- Schürmann, Hans: Subject indexing in a digital age. Presented at the General Assembly of the European Theological Libraries (47th General Assembly of BETH), Nürnberg 2018. <https://doi.org/10.5281/zenodo.1420492>.
- Shannon, Claude E. und Warren Weaver: The mathematical theory of communication. Urbana: The University of Illinois Press, 1964. https://pure.mpg.de/rest/items/item_2383164/component/file_2383163/content (29.1.2021).
- Sundermeier, Jörg: Rechte Verlage und ihre Produkte. Sollten Bücher aus rechten Verlagen im Bestand geführt werden? In: BuB – Forum Bibliothek und Information (2018) Jg. 70 H. 6. S. 331–333. <https://b-u-b.de/wp-content/uploads/2018-06.pdf> (29.1.2021).
- Turčilo, Lejla und Mladen Obrenović: Fehlinformationen, Desinformationen, Malinformationen. Ursachen, Entwicklungen und ihr Einfluss auf die Demokratie. Demokratie im Focus #3.

Heinrich Böll Stiftung, 2020. <https://www.boell.de/de/2020/08/25/fehlinformationen-de-sinformationen-malinformationen> (29.1.2021).

Wikimedia Deutschland: Das WikiLibrary Manifest. <https://www.wikimedia.de/projects/wikibrary-manifest/> (29.1.2021).

Zollondz, Hans-Dieter, Ketting, Michael und Raimund Pfundtner (Hrsg.): Lexikon Qualitätsmanagement. Handbuch des Modernen Managements auf Basis des Qualitätsmanagements, 2. Aufl. Berlin, Boston: De Gruyter Oldenbourg 2016. <https://doi.org/10.1515/9783486845204>.