

Fachbeitrag

Michael Franke-Maier

Anforderungen an die Qualität der Inhalterschließung im Spannungsfeld von intellektuell und automatisch erzeugten Metadaten

<http://doi.org/10.1515/abitech-2018-4005>

Zusammenfassung: Spätestens seit dem Deutschen Bibliothekartag 2018 hat sich die Diskussion zu den automatischen Verfahren der Inhalterschließung der Deutschen Nationalbibliothek von einer politisch geführten Diskussion in eine Qualitätsdiskussion verwandelt. Der folgende Beitrag beschäftigt sich mit Fragen der Qualität von Inhalterschließung in digitalen Zeiten, wo heterogene Erzeugnisse unterschiedlicher Verfahren aufeinandertreffen und versucht, wichtige Anforderungen an Qualität zu definieren.

Schlüsselwörter: Inhalterschließung, Automatisierung, Qualität

Requirements concerning the quality of subject indexing with special regard to intellectually and automatically generated metadata

Abstract: Since the 2018 German Librarians' Convention, if not earlier, the discussion concerning automatic methods of subject indexing by Deutsche Nationalbibliothek has shifted from a political discussion to a discussion about quality. This paper deals with the quality of subject indexing in digital times, when different methods produce heterogeneous metadata; it seeks to define important quality requirements.

Keywords: Subject Indexing, Automatization, Quality

1 Qualitätsoffensive – und nun?

Als 1994 die Expertengruppe Online-Kataloge der Kommission des Deutschen Bibliotheksinstituts für Erschließung und Katalogmanagement in ihrem Band zu Sacherschließung in Online-Katalogen schrieb, dass die „Sacherschließung [...] ein Thema [sei], das für Bibliotheken im Online-Zeitalter erheblich an Bedeutung gewinnt“,¹ war das, historisch eingeordnet, die folgerichtige Antwort auf die Einführung von Online-Katalogen, ja sogar wie im Fall des Gemeinsamen Bibliotheksverbunds bereits die Antwort auf die Pica-Verbundkatalogisierung mit 2,3 Millionen Nachweisen.²

Heute, 25 Jahre später, gilt dieser Satz nach wie vor, stellt man ihn in den Kontext von Mega-Indices wie K10plus-zentral mit 202 Millionen bibliografischen Nachweisen. Gleichzeitig scheint es, dass das Verhältnis des Wissenswachstums und der Ergebnismengen in Suchwerkzeugen aus den Fugen geraten ist – die Beziehung zwischen den klassischen Messgrößen der Inhalterschließung, zwischen Recall und Precision, stimmt nicht mehr.

Welche Qualitätsanforderungen für die Inhalterschließung können definiert werden, damit Suchwerkzeuge weiterhin Recherche als strategische Erkundung, wie in den Threshold Concepts definiert, ermöglichen?³

¹ Kommission des Deutschen Bibliotheksinstituts für Erschließung und Katalogmanagement, Expertengruppe Online-Kataloge. *Sacherschließung in Online-Katalogen*. Berlin: Dt. Bibliotheksinstitut, 1994, 5.

² Vgl. dazu Diedrichs, Reiner. „Vorwort.“ In *Jahresbericht 2017*. Hrsg. von der Verbundzentrale des Gemeinsamen Bibliotheksverbundes. Göttingen: VERLAG 2018, 3. https://www.gbv.de/Verbundzentrale/Publikationen/PDF/Jahresbericht_2017_web.pdf (30.08.2018).

³ Vgl. dazu Franke, Fabian. „Das Framework for Information Literacy. Neue Impulse für die Förderung von Informationskompetenz in Deutschland?!“ *o-bib. Das offene Bibliotheksjournal* 4, 4 (2017): 26. <https://doi.org/10.5282/o-bib/2017H422-29> (30.08.2018).

Vor allem dann, wenn automatisch erstellte inhaltliche Metadaten auf intellektuelle treffen?

Nach der Definition der ISO-Norm für Qualitätsmanagement ist Qualität der „Grad, in dem ein Satz inhärenter Merkmale eines Objekts Anforderungen erfüllt“⁴ – sie ist also dynamisch und abhängig von den Anforderungen, die wir definieren. Zur Sicherung muss ihre Erfüllung messbar gemacht werden. Dies ist im Hinblick auf die Inhaltserschließung sicher kein einfaches Unterfangen. Hierfür spricht sich dennoch eine Initiativgruppe des Standardisierungsausschusses in ihrer Stellungnahme zur Entwicklung der Inhaltserschließung im D-A-CH-Raum aus, indem sie „effektive, überprüfbare Standards für die Qualitätssicherung und Transparenz“⁵ für maschinelle Verfahren der Inhaltserschließung fordert. Diese Stellungnahme beendet gleichzeitig eine politische Diskussion⁶ um den Einsatz von automatischen Verfahren durch die Deutsche Nationalbibliothek und eröffnet Spielräume für eine positive Grundhaltung bezogen auf Qualität von intellektueller und maschineller Inhaltserschließung auf hohem Niveau auf allen Ebenen der Bibliotheksorganisation.

2 Kartierung von Qualitätsdimensionen

Um Qualität zu erreichen, reicht es nicht, sich auf einen Bereich des bibliothekarischen Arbeitens zu beschränken. Erst ein Blick auf alle Produkte und Prozesse der Inhaltserschließung lässt Interdependenzen erkennen, und es entsteht ein Bild unterschiedlicher voneinander abhängiger Qualitätsdimensionen, die im Folgenden – ohne Anspruch auf Vollständigkeit – beschrieben werden.

⁴ Wikipedia. *Qualität*. <https://de.wikipedia.org/w/index.php?title=Qualität&oldid=177258864> (30.08.2018).

⁵ Initiativgruppe des Standardisierungsausschusses. *Stellungnahme zur Entwicklung der Inhaltserschließung im D-A-CH-Raum*. <http://www.dnb.de/SharedDocs/Downloads/DE/DNB/standardisierung/protokolle/stellungnahmeEDachRaum.html> (30.08.2018).

⁶ Vgl. dazu Franke-Maier, Michael. *Qualitätsanforderungen zwischen intellektueller und algorithmisierter Inhaltserschließung*. Vortrag, 22. Verbundkonferenz des Gemeinsamen Bibliotheksverbundes. Kiel: 29.08.2018. Folie 3. https://verbundkonferenz.gbv.de/wp-content/uploads/2018/09/Franke-Maier_Qualität_MFM_final.pdf (17.09.2018).

2.1 Qualität der Konzepte und Konzept-Korpora

Grundlegend für alle Vorgänge der Inhaltserschließung, sei sie intellektuell oder automatisch, sind Thesauri, Ontologien und Normdateien – allgemeiner: Konzept-Korpora mit ihren Entitäten. Für den deutschsprachigen Raum sind hier beispielhaft die Gemeinsame Normdatei (GND) und die Regensburger Verbundklassifikation (RVK) zu erwähnen. Neben die klassischen inhaltlichen Anforderungen wie Eindeutigkeit, Informationsgehalt, Gebräuchlichkeit⁷ oder Aktualität im Sinn von Wissenschaftsnähe⁸ treten technische Anforderungen der Maschinenlesbarkeit.

Weiterhin sind Hierarchisierung und Relationierung untereinander, ggf. mit anderen Normdaten, zu nennen. Die GND hat diesbezüglich seit der Einführung des Entity-Relationship-Models eine große Qualitätshürde genommen – wären da nicht die aufzuarbeitenden Altdaten. Die RVK ist seit Kurzem durch die Einführung von Identifiern und einer CC0-Lizenz positiv aufgefallen und hat sich zu einer gut nutzbaren, wenn auch in Teilen inkonsistenten Klassifikation entwickelt. Eine der wichtigsten und sicherlich schwer zu beantwortenden Qualitätsfragen, die sich im Spannungsfeld von intellektueller und automatischer Inhaltserschließung stellt, ist jene nach der Eignung der Konzepte bzw. der Konzept-Korpora zur präzisen Inhaltserschließung durch Mensch und Maschine.

2.2 Qualität der Regelwerke

Zur Anwendung der Korpora bedarf es Regeln. Diese Regeln ihrerseits wirken sich auf die Konzeptfassung aus, z. B. durch die Zerlegungskontrolle wie sie aus den Regeln der Schlagwortkatalogisierung (RSWK) bekannt ist.⁹ Dort, wo das Regelwerk zerlegt und intellektuell mit der Kombination mehrerer Schlagwörter agiert, erkennt die Maschine den zusammengesetzten Begriff im Volltext als engstes Schlagwort. Die Erzeugung von Präzision wirft insofern folgende Fragen auf: Eignen sich unsere

⁷ Vgl. dazu Mödden, Elisabeth, Helga Karg. *Maschinelle Indexierung der Deutschen Nationalbibliothek*. Vortrag, 107. Deutscher Bibliothekartag. Berlin: 2018. Folie 16. <https://nbn-resolving.org/urn:nbn:de:0290-opus4-157942> (11.09.2018).

⁸ Vgl. dazu Universitätsbibliothek Regensburg. *RVK – Regensburger Verbundklassifikation. For the good neighbourhood of books...* Regensburg, 10.2011. https://rvk.uni-regensburg.de/images/stories/Con2011/Beitraege/flyer%20rvk_2011.pdf (11.09.2018).

⁹ Arbeitsstelle für Standardisierung. *Regeln für die Schlagwortkatalogisierung*. 4., vollständig überarb. Aufl. Leipzig, Frankfurt am Main: Deutsche Nationalbibliothek, 2017. § 8,5, 53 und § 304,3, 121.

Regelwerke für die Welt digitaler Volltexte? Und können wir sie für die Erzeugung kongruenter Erschließungsdaten programmieren?

Weitere Anforderungen an Regelwerke sind die Minimierung der Effekte der Erschließung durch unterschiedliche Personen bzw. zu unterschiedlichen Zeitpunkten (Inter- bzw. Intra-Encoder-Reliability) oder eine leichte Anwendbarkeit mit geringem Schulungsaufwand.

2.3 Qualität der digitalen Werkzeuge

Die Entwicklung assistierender Systeme ist nicht zuletzt deswegen erfolgt, weil gängige Bibliotheksmanagementsysteme nicht in der Lage sind, den Workflow der intellektuellen Inhaltserschließung optimal und effizient zu gestalten. Ob assistierende Systeme, wie z. B. das Unterstützungstool für Fachreferatsarbeit der UB Mannheim¹⁰ oder der Digitale Assistent von Eurospider,¹¹ den Workflow besser abbilden, sei dahingestellt. In Zeiten von Wischen, Zoomen und Tippen auf mobilen Endgeräten scheinen mausbasierte Click-Oberflächen in ihrer Usability zumindest strittig.

Auch wenn auf Managementebene eine Peer-Review-Funktionalität schwer durchzusetzen wäre, also die Prüfung von intellektuell erstellten Indexaten durch Dritte, ist sie im föderalen System der Verbundkataloge im Zusammenhang mit Korrekturingriffen bei Fremddatenübernahmen bereits Realität, auch ohne Kennzeichnung der vorgenommenen Qualitätssicherung. Ein digitaler Assistent als verbundübergreifendes soziales Netzwerk könnte gerade bei inhaltlich schwierig zu erfassenden Publikationen eine Lösung sein, da er den Austausch bibliothekarischer Fach-Communities stärken könnte.

2.4 Qualität der Algorithmen

Bei automatischen Erschließungsverfahren greifen zunächst die klassischen Messgrößen Recall, Precision und F1-Score sowie Konfidenzwerte. Recall und Precision sind dabei abhängig von Methode und Datenpool: Bei einem schwer zu erschließenden Datenpool kann ein niedriger Precision-Wert bereits gut sein, wohingegen bei

einem Verfahren, das nur grob erschließt, eine hohe Precision wenig aussagt.

Bei Werk-Cluster-Verfahren wie Culturegraph¹² ist die exakte Gruppierung gleicher Titelsätze von hohem Gewicht. Qualität entsteht durch die Aufbereitung der inhaltlichen Metadaten vor dem Verteilen an das gesamte Werk-Cluster.

Bei assistierenden Systemen stellt sich die Frage, ob sie tatsächlich Mehrwert erzeugen oder nur Substitut für einen nicht funktionierenden Datenaustausch sind. Mehrwert wird dort erzeugt, wo ein Assistenzsystem der erschließenden Person bei der Beschlagwortung gänzlich uner-schlossener Publikationen behilflich ist und das nicht nur mittels einer Ähnlichkeits-Algorithmik. Wichtig scheint doch vielmehr – wenn Vorschläge zur Erschließung auf Basis von ähnlichen Titeln erfolgen –, dass der inhaltliche Unterschied¹³ der zu erschließenden Publikation klar herausgestellt wird. Nur so kann präzise erschlossen werden.

Letztlich ist es eine Frage der Effektivität: Können wir mit solchen Systemen das Richtige tun? Verbleibt die Algorithmik in Händen gewinnorientierter Firmen, wird dies schwer zu entscheiden sein. Ein Blick in aktuelle Forderungskataloge zu „Gütekriterien für algorithmische Prozesse“ lohnt ebenfalls für bibliothekarische Anwendungen.¹⁴

2.5 Qualität des Titeldatensatzes

Die Qualität des Titeldatensatzes ist abhängig von Datenformaten wie MAB, MARC oder RDF und deren Möglichkeiten bzw. Limitationen. Zentral für eine spätere Steuerung in Discovery Systemen sind die Erfassung der Provenienz von Datenelementen und der verwendeten Konzept-Korpora, von Identifiern sowie ggf. von Rollenoperatoren.

Zum Schutz vor Datenmanipulationen kann eine Zertifizierung bzw. die Angabe einer qualitätssichernden Maßnahme dienen. Bestenfalls ist der Datensatz historisierbar

¹⁰ Universitätsbibliothek Mannheim. *Unterstützungstool für Fachreferatsarbeit*. [http://data.bib.uni-mannheim.de/malibu/isbn/suche.html?isbn=\(11.09.2018\)](http://data.bib.uni-mannheim.de/malibu/isbn/suche.html?isbn=(11.09.2018)).

¹¹ Vgl. dazu Beckmann, Regine, Imma Hinrichs. *Alles unter einer Haube: Die nächste Generation des Digitalen Assistenten DA-3*. Vortrag, 107. Deutscher Bibliothekartag, Berlin: 2018. <https://nbn-resolving.org/urn:nbn:de:0290-opus4-36355> (11.09.2018).

¹² Vgl. dazu Deutsche Nationalbibliothek. *culturegraph.org – Plattform für Wissensvernetzung*. <http://www.culturegraph.org/> (17.09.2018).

¹³ Vgl. dazu die Definition zu „property“. International Organization for Standardization. *ISO 5127:2017(en). Information and documentation – Foundation and vocabulary*. 3.1.1.03. <https://www.iso.org/obp/ui/#iso:std:iso:5127:ed-2:v1:en> (17.09.2018).

¹⁴ Vgl. dazu Rohde, Noëlle. *Gütekriterien für algorithmische Prozesse. Eine Stärken- und Schwächenanalyse ausgewählter Forderungskataloge. Arbeitspapier*. Gütersloh: Bertelsmann Stiftung, 2018. <https://doi.org/10.11586/2018027> (17.09.2018).

und frühere Zustände in einer Bearbeitungshistorie sind einsehbar.¹⁵

Weiterhin kann man Vollständigkeit fordern: Ist eine Publikation verbal, klassifikatorisch und automatisch erschlossen? Gibt es ein Abstract und ein maschinenlesbares, sauberes Inhaltsverzeichnis?

2.6 Qualität des Titeldatensatzes im Kontext anderer Datensätze

Titelsätze stehen im Kontext zu anderen Titelsätzen und zu Normdaten. Klassische Qualitätsmaße sind Konsistenz und Kongruenz: „Gleiche Themen werden gleich erschlossen“¹⁶ fordern z. B. die Regeln für die Schlagwortkatalogisierung (RSWK). Unerschlossene Dokumente fallen aus dem Raster und wirken sich, ebenso wie nicht deckungsgleich erschlossene, negativ auf den Recall aus.

Ferner führte die Einführung von Resource Description and Access (RDA), den Functional Requirements for Bibliographic Records (FRBR) und deren Weiterentwicklung zum IFLA Library Reference Model¹⁷ zu einer maßgeblichen, formalen Qualitätssteigerung bei der Verwendung bibliografischer Daten durch eine vernetzte Datenumgebung.

Letztendlich braucht es einen Grundstock an hochqualitativen, validierten Titeldaten, ggf. fachspezifisch, an dem sich die Ergebnisse der Inhaltserschließung automatischer Verfahren messen können und der als Trainingspool für lernende Verfahren fungiert. Da Wissenschaft sich ändert, bleibt die Generierung eines solchen Datensets eine Daueraufgabe.

2.7 Qualität des Datenaustauschs

Eine der wichtigsten Anforderungen beim Datenaustausch ist der Erhalt von Qualität: Der Anspruch ist, dass die Importqualität mindestens der Ursprungsqualität und die Exportqualität der internen Qualität entspricht. Im föderalen Verbundsystem bleiben Feldkonkordanzen und Form-

konverter für die bibliothekarische Allgemeinheit eine Black Box, und es ist viel zu oft eine Qualitätsminderung zu attestieren.

2.8 Qualität der Suchwerkzeuge

Die letzte Dimension betrifft unsere Suchwerkzeuge: Gemeinhin bieten wir Funktionalitäten wie Themensuche und -facette und Verlinkungen zur Weitersuche an, aber von der Ermöglichung der Recherche als strategische Erkundung und der Überwindung des Google-Schlitz-Dogmas hin zu einem thematischen Browsen bzw. einem Knowledge Discovery sind wir weit entfernt.

Durch die Masse der heterogen erschlossenen Nachweise in Discovery-Systemen und den Wechsel vom „exact match“ hin zum „best match“ liefert das Relevanz-Ranking auch irrelevanten Bias, und die kontextabhängige Auswertung erfolgreicher Suchanfragen hin zu einem lernenden Relevanz-Ranking wird durch Datenschutzbestimmungen erschwert.

3 Beschreibungsgrad und Erschließungsaufwand – Fazit

Wenn über Qualität gesprochen wird, stellt sich schnell die Frage nach dem Preis bzw. Aufwand. Mit wachsendem Beschreibungsgrad wird die Qualitätserzeugung aufwändiger, dies gilt gleichermaßen für intellektuelle als auch für automatische Verfahren. Das Narrativ der teuren intellektuellen Inhaltserschließung hält sich hartnäckig, obwohl zu den tatsächlichen Kosten keine veröffentlichten Zahlen vorliegen. Dabei wird bei der Argumentation vergessen, dass auch eine gute automatische Inhaltserschließung ihren Preis haben wird: Auch dazu erfährt man wenig, da Ausgaben, z. B. für das Verfahren der DNB, nicht transparent sind.

Der maßgebliche Grund für die Legitimation unterschiedlicher, sich ergänzender Erschließungsmethoden jedoch ist die Publikationsflut. Hierbei stoßen in einem Suchwerkzeug heterogene Beschreibungsgrade aufeinander und erzeugen einen inkongruenten Datenpool. Verschiedene Erschließungsmethoden können gleichwertig nebeneinanderstehen, solange es ein Bekenntnis zu hoher Qualität gibt und die Wege der Datenerzeugung gekennzeichnet werden, um sie für ein optimales Suchergebnis auswertbar zu machen!

¹⁵ Vgl. dazu z. B. Bruce, Thomas R., Diane I. Hillmann. „The Continuum of Metadata Quality: Defining, Expressing, Exploiting.“ In *Metadata in Practice*. Hrsg. von Diane I. Hillmann, Elaine L. Westbrooks, American Library Association. 238–256. Chicago: American Library Association, 2004.

¹⁶ Arbeitsstelle für Standardisierung 2017: § 6.4, 49.

¹⁷ Riva, Pat, Patrick Le Boeuf, Maja Žumer, IFLA. *IFLA Library Reference Model. A Conceptual Model for Bibliographic Information*. Den Haag: 2017. https://www.ifla.org/files/assets/cataloguing/frbr-lrm/ifla-lrm-august-2017_rev201712.pdf (17.09.2018).

Anmerkung: Dieser Tagungsbeitrag fasst die vom Autor als Impulse vorgetragene Ideen beim Workshop der FAG „Erschließung und Informationsvermittlung“ des GBV am 29. August 2018 in Kiel zusammen. Der Workshop fand im Rahmen der 22. Verbundkonferenz des GBV statt.

Autoreninformationen



Michael Franke-Maier
Universitätsbibliothek der Freien
Universität Berlin
Garystraße 39
14195 Berlin
Deutschland
franke@ub.fu-berlin.de
orcid.org/0000-0003-4263-2754