

Arbeitsstelle für Standardisierung (AFS)
Juli 2021

Expertenteam „RDA-Anwendungsprofil für die verbale Inhaltsererschließung (RAVI)“

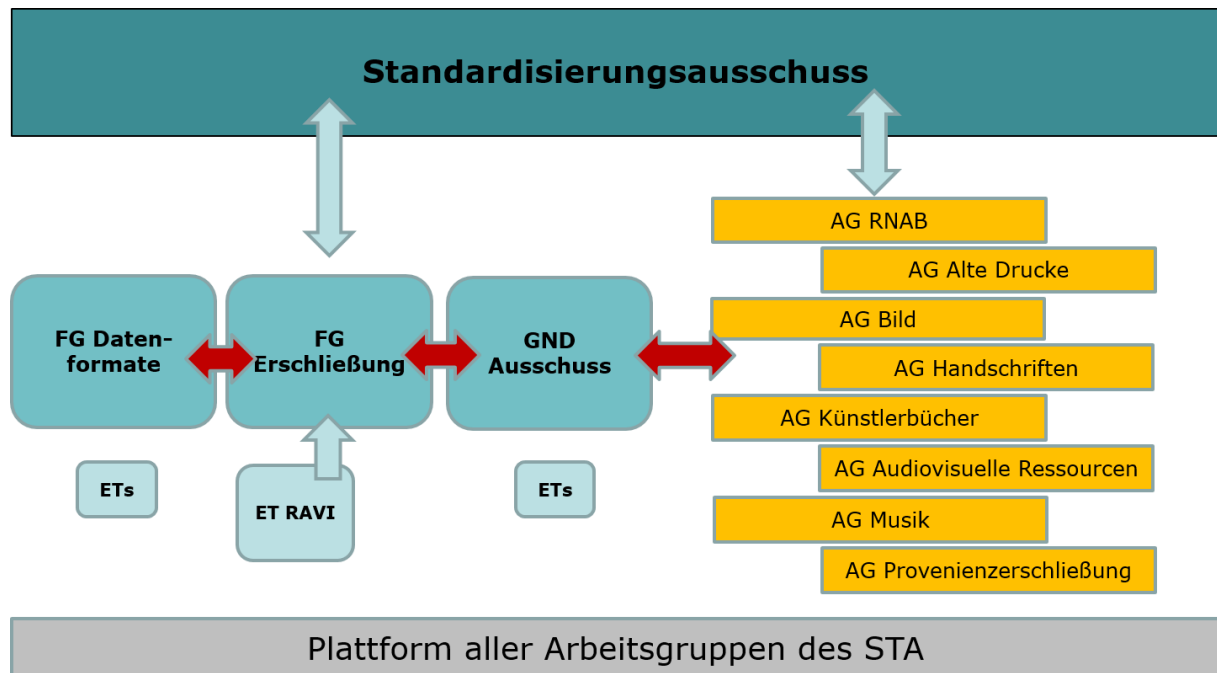
Ansprechpersonen

Zum Inhalt

Dr. Julijana Nadj-Guttandin, Deutsche Nationalbibliothek
Regine Beckmann, Staatsbibliothek zu Berlin

Zur Organisation und zur Einbindung in die Arbeit des Standardisierungsausschusses

Renate Behrens, Arbeitsstelle für Standardisierung der DNB



Inhalt

Berichtszeitraum: Januar 2021 bis Juni 2021	3
Berichtszeitraum: Juli 2020 bis Dezember 2020	4
Berichtszeitraum: Januar 2020 bis Juni 2020	5
Berichtszeitraum: Juli 2019 bis November 2019	6
Berichtszeitraum: Dezember 2018 bis Juni 2019	16
Berichtszeitraum Juli 2018 bis November 2018	17
Berichtszeitraum: Januar 2018 bis Juni 2018	18
Berichtszeitraum: Juni 2017 bis November 2017	20

Berichtszeitraum: Januar 2021 bis Juni 2021

Regelwerksarbeit

Im Berichtszeitraum Dezember 2020 bis Juni 2021 wurde in Unterarbeitsgruppen an zentralen Sachverhalten gearbeitet, deren Neuregelung grundlegend und richtungsgebend für ein anforderungsgerechtes Regelwerk ist: die allgemeinen Grundsätze der Verwendungsregeln (u. a. gleichordnendes / syntaktisches Indexieren, die Erschließungstiefe und die Zerlegungskontrolle), die Abbildung des Zeitaspekts und das neue Konzept von Rollenoperatoren im Indexat am Beispiel der motivischen Darstellung. Des Weiteren wurde ein Template für die zukünftigen Verwendungsregeln entworfen.

Beim kommenden LIS Workshop im Rahmen der virtuellen DSSV-ECDA-Tagung 2021 wird der Fachcommunity der eingeschlagene Weg vorgestellt. Der Vorschlag für die Neuregelung der Motivschlagwörter durch einen Rollenoperator wird im nächsten Treffen der Sonderarbeitsgruppen des Standardisierungsausschusses zur Diskussion gestellt.

Die Bearbeitungsstände wurden in monatlichen Videokonferenzen diskutiert und weiterentwickelt. Besonders deutlich wurde bei allen Themen, dass sehr frühzeitig eine Einbindung in die Katalogisierungsumgebungen und in die verschiedenen Nachweisinstrumente mitbedacht werden muss, damit die Neuregelung überhaupt wirksam werden kann. Hierfür schlägt das ET RAVI vor, die entsprechenden Gremien einzubeziehen.

D-A-CH Dokumentationsplattform

Eine Neuerung mit unmittelbarer Auswirkung auf die laufende Regelwerksarbeit ist die aktuelle Entwicklung der GND-Dokumentation innerhalb der D-A-CH-Dokumentationsplattform. In einer gesonderten Videokonferenz mit dem Leiter der Arbeitsstelle für Standardisierung, Herrn Kett, wurde besprochen, wie sich die GND-Dokumentation und das Regelwerk zur Inhaltserschließung zueinander verhalten können. ET RAVI sieht sich dabei nicht nur für die Verwendungsregeln im bibliothekarischen Bereich zuständig, sondern auch für den übergreifenden Standard bei der Erfassung der Sachschlagwörter, der durch andere Communities anwendungsspezifisch modifiziert werden kann. Ein enger Informationsaustausch ist hier aus RAVI-Sicht wichtig und sollte verstetigt werden, daher ist ein weiterer Termin im 4. Quartal vorgesehen. Das ET RAVI erarbeitet prototypische Textvorschläge für die neue Wikibase-Struktur.

Mitglieder

Hans Schürmann (Swiss Library Service Plattform SLSP) wurde als Nachfolger von Frau Dr. Boxler im RAVI-Team begrüßt.

Weiterhin vakant ist der Platz für ein Mitglied aus dem hebis. Eine Vertretung wäre sehr wünschenswert, um die Impulse und Anforderungen auch dieses Verbundes in die Arbeit des Expertenteams integrieren zu können.

Berichtszeitraum: Juli 2020 bis Dezember 2020

Im Berichtszeitraum Juli – Dezember 2020 hat das ET RAVI in regelmäßigen Videokonferenzen gearbeitet. Einen Schwerpunkt der Tätigkeit bildete die Arbeit an einzelnen Modulen für das neue Regelwerkskonzept. Es wurde beschlossen, sogenannte Templates / Eingabemasken für einzelne Modularten wie z. B. die Erfassung von Normdaten anzulegen. Dadurch wird nicht nur die einheitliche Vorgehensweise bei der Erstellung und Bearbeitung einzelner Module gewährleistet, sondern auch eine zukünftige Versionierung vereinfacht.

Kernelemente werden für möglichst alle Module obligatorisch sein (wie z. B. Thema/Titel, interne Nummerierung, Status, Verweis auf andere verwandte bzw. zu berücksichtigende Module, Verweis auf andere Regelwerke, z. B. RDA, kurze Zusammenfassung des Inhalts, Volltext des Moduls und Beispiele), so dass sich das neue Regelwerk durch eine große inhärente Konsistenz auszeichnet. Es wird darüber hinaus möglich sein, die einzelnen Elemente eines Templates und deren Abfolge individuell anzupassen und ggfs. zu erweitern, wenn besondere Sachverhalte dies erfordern.

Einzelne Module wurden in einer ersten Fassung erarbeitet. Flankierend dazu wurden Fragen nach einer möglichen Strukturierung bzw. Nummerierung sowie terminologische Fragen diskutiert und in einzelnen Wiki-Seiten dokumentiert.

Einen zweiten Schwerpunkt der Arbeit des Expertenteams bildeten Überlegungen zur Festlegung des theoretischen Rahmens für das neue Regelwerk. Zunächst hat eine Unterarbeitsgruppe die dazu bereits bestehenden Vorarbeiten noch einmal gesichtet und gebündelt. Die Diskussion zu diesen konzeptionellen Fragen, insbesondere nach dem theoretischen Gerüst (wie z. B. ISO und DIN-Normen, IFLA LRM) sowie zu Granularität und Spezifität von Erschließung und Erschließungsmethoden wird das Expertenteam in einer zweitägigen Webkonferenz Mitte Dezember führen.

Berichtszeitraum: Januar 2020 bis Juni 2020

Im Berichtszeitraum Januar 2020 – Juli 2020 hat sich das Expertenteam RAVI in regelmäßigen Telefonkonferenzen getroffen. Eine für Ende April geplante Präsenzsitzung musste aufgrund der Corona-Quarantänebestimmungen abgesagt werden. Für die Vertretung des IDS ergab sich ein personeller Wechsel: Seit Januar 2020 ist Frau Dr. Madeleine Boxler Mitglied im ET RAVI.

Die Arbeit an den Modulen für das neue Regelwerk für die verbale Inhaltserschließung bildete den Schwerpunkt der Arbeit des ET RAVI. Dabei haben sich für jedes Modul Bearbeiter*innen gefunden, die für die erste Texterstellung verantwortlich sind. Wie vom Standardisierungsausschuss vorgegeben, werden die einzelnen Module für die Entitäten (Person / Geografikum / Körperschaft / Veranstaltung / Sachschlagwort / Werk) getrennt nach Erfassung sowie Verwendung von Normdaten betrachtet. Neben den Modulen für die Entitäten soll das neue Regelwerk insbesondere eine große Interoperabilität gewährleisten und individuelle Nutzungsbedürfnisse bedienen. Daher werden neben den Erfassungs- und Verwendungsmodulen für die Normdaten weitere Module erarbeitet, die folgende Themen abbilden sollen:

- Verzahnung mit anderen Erschließungsvokabularen
- Verzahnung mit klassifikatorischer Erschließung
- Andocken an maschinelle Erschließungsverfahren

Daneben werden die Bedürfnisse verschiedener (auch nicht-bibliothekarischer) Anwender-Communities berücksichtigt. In einzelnen Modulen werden

- fachbezogene,
- zielgruppenbezogene
- und ressourcenspezifische

Bedürfnisse an inhaltliche Erschließung erarbeitet. Um die Bedürfnisse abzufragen und zu definieren, wird das ET RAVI bei der Bearbeitung dieser Module mit den entsprechenden Stakeholdern, z. B. aus den Sonderarbeitsgruppen des STA, zusammenarbeiten.

Explizit sollen auch Überlegungen zur Darstellung inhaltserschließender Metadaten in Katalogen ins neue Regelwerk mit aufgenommen werden. In einem speziellen Modul werden best practices für Retrieval und Anzeige dargestellt.

Was die konkrete Textarbeit betrifft, so werden die einzelnen Bereiche der RSWK kritisch hinsichtlich der Übertragbarkeit in ein neues, stark modularisiertes Konzept durchgearbeitet. Es wird konsequent auf eine Trennung von Erfassungs- und Verwendungsregeln geachtet und es werden bestehende Defizite identifiziert, die es zu lösen gilt (z. B. der Umgang mit Motiven). Folgende konzeptionelle bzw. strukturelle Fragen werden unsere konkrete Textarbeit begleiten:

- Umgang mit Sprache (Einheitlichkeit der Terminologie, Stil)
- die Einbindung und Verwendung von Beispielen (Umfang, Format, Darstellung, Verständlichkeit)
- die Strukturierung von Verweisungen innerhalb des neuen Regelwerks
- Verweisungen zum neuen RDA-Toolkit
- Überlegungen zum Umgang mit außerhalb von RSWK bestehenden Dokumentationen und Erfassungshilfen zur Inhaltserschließung.

Diese konzeptionellen und strukturellen Fragen werden zu gegebener Zeit mit den Kolleg*innen des Projekts „3R für DACH-Bibliotheken“ rückgekoppelt, damit letztlich eine einheitliche Erschließungsumgebung entstehen kann. Um die Verzahnung der RAVI-Arbeit mit dem geplanten RDA-Erschließungshandbuch zu gewährleisten ist das ET RAVI im Projekt „DACH Dokumentationsplattform“ vertreten.

Berichtszeitraum: Juli 2019 bis November 2019

Im Berichtszeitraum Juli 2019 - Dezember 2019 hat sich das Expertenteam RAVI in monatlichen Telefonkonferenzen mit den drei Themen „Qualität in der Erschließung“, „Eignung der Normdatei und des Regelwerks für intellektuelle und maschinelle Verfahren“ sowie „Sacherschließung am Werknormdatensatz“ beschäftigt. Unterarbeitsgruppen des Expertenteams haben in eigenen Telefonkonferenzen jeweils ein Papier zu den drei Themenkomplexen erarbeitet; sie sind diesem Sachstandsbericht beigelegt. Die Hauptaspekte der Themen sind in *executive summaries* zusammengefasst und folgen diesem einführenden Bericht.

Im Berichtszeitraum haben Mitglieder des ET RAVI die bisherigen Ergebnisse zum Thema „Qualität“ in verschiedenen Kontexten vorgestellt:

- Fortbildungsveranstaltung des VDB-Hessen „Im Fokus: Inhaltserschließung“ am 02. September 2019 in Fulda
- Sitzung der österreichischen Kommission für Sacherschließung am 09. September 2019 in Graz
- Tagung „Netzwerk maschinelle Erschließung“ am 10./11. Oktober 2019 in der DNB in Frankfurt a.M. im Rahmen des Knowledge Cafés „Inhaltserschließung im digitalen Zeitalter für Wissenschaft von heute und morgen“

Die Erarbeitung von messbaren Qualitätskriterien stellt eine große Herausforderung dar. In der Frage der Etablierung von Verfahren zur Messbarkeit der Qualitätsanforderungen schlägt das ET RAVI einen Austausch mit entsprechenden Fachleuten vor. Eventuell wäre dies ein geeignetes Thema für eine nächste Fachtagung „Netzwerk maschinelle Erschließungsverfahren“. Messgrößen für Qualität in der Erschließung können u. a. sein:

- Die codierte Kennzeichnung der Daten hinsichtlich der Erschließungsmethode, des Erschließungsverfahrens und des Erschließungslevels
- Die Auszeichnung nicht erschlossener Materialien
- Die statistische Auswertung der Weiternutzung von Rechercheergebnissen (Download, Ausleihe, iterative Recherche etc.)

Darüber hinaus stellt sich stets die Frage, inwiefern das Gefundene für das jeweilige individuelle Informationsbedürfnis der Benutzerinnen und Benutzer relevant ist. Wünschenswert wäre daher auch ein überregionales Projekt zur Evaluierung der Wirkung der Inhaltserschließungsdaten.

Am 12. November 2019 fand die zweite Präsenzsitzung des ET RAVI in der DNB Frankfurt a.M. statt. Gemäß dem Auftrag des STA entwickelte das Expertenteam erste Überlegungen für ein Konzept für modular aufgebaute Erschließungsregeln für die verbale Inhaltserschließung. Das Konzept soll verschiedene Erschließungslevels berücksichtigen und die Ansprüche, die aus den verschiedenen Qualitätsdimensionen abgeleitet werden, erfüllen. Die Arbeit an diesem Konzept wird den Schwerpunkt der Tätigkeit des ET RAVI im nächsten Jahr bilden. Es ist angedacht, die bisherigen Überlegungen dazu auf dem Bibliothekartag in Hannover im Rahmen eines community-übergreifenden Hands-on Lab zu präsentieren und zu diskutieren. Ein entsprechendes Abstract für den Themenkreis 5 „Content und Digitalisierung“ wird eingereicht.

Sachstandsbericht des ET RAVI - Executive Summaries

Definition allgemeiner verfahrensunabhängiger Qualitätskriterien und ggf. Verfahren zur Qualitätssicherung – Executive Summary

Das Expertenteam „RDA-Anwendungsprofil für die verbale Inhaltsererschließung“ hat sich gemäß des Auftrags des Standardisierungsausschusses¹ im Jahr 2019 in seiner Arbeit auf die „Definition allgemeiner verfahrensunabhängiger Qualitätskriterien und ggf. Verfahren zur Qualitätssicherung“ konzentriert.

Die Qualität inhaltlicher Erschließung resultiert unserem Verständnis nach mindestens aus den folgenden Qualitätsdimensionen:

- Verlässlichkeit und Transparenz (Grunddimensionen)
- Regelbasierte Produktion von Normdaten
- Verwendungsregeln für die Ressourcenbeschreibung
- Transparente Auswertung für Retrieval und Anzeige
- Öffnung der Daten für die Nachnutzung jenseits des klassischen Retrievals

Die Grunddimension Verlässlichkeit versteht sich im Sinne von Konsistenz der Daten, der Regeln und des Retrievals. Transparenz beinhaltet die Offenlegung der Regeln und Standards, die genaue Herkunftskennzeichnung von Daten und das Verdeutlichen von Erschließungslücken und heterogenen Erschließungssituationen (siehe Punkt 3).

Voraussetzung für eine qualitative inhaltliche Erschließung sind regelbasiert erstellte Wissensorganisationssysteme. Deren Anforderungen sind u. A. Eindeutigkeit, Disambiguierung, Gebräuchlichkeit und Aktualität, Maschinenlesbarkeit und Ausrichtung bei der Erstellung an internationalen Standards, z. B. RDA (siehe Punkt 4).

In einem kooperativen Erschließungskontext ist ein gewinnbringender Datenaustausch dann möglich, wenn auf gemeinsame Verwendungsregeln für die inhaltliche Ressourcenbeschreibung zurückgegriffen wird. Diese sind vor dem Hintergrund unterschiedlicher Bedarfe von Fachdomänen, von Institutionen, ihrer Benutzerinnen und Benutzer und in Abhängigkeit vom Dokumententyp zu formulieren. Qualitätsanforderungen an die Verwendungsregeln sind u. A. Verständlichkeit, gut nachvollziehbare Gliederung und Gestaltung, Flexibilisierung durch Einführung von abgestuften und standardisierten Erschließungslevel mit unterschiedlichen Präzisionsgraden sowie die Ausrichtung der Verwendungsregeln an der Eignung der Ergebnisse als Trainingsdaten für maschinelle Verfahren (siehe Punkt 5).

Um die optimale Auffindbarkeit von Ressourcen im Zusammenspiel von Recall und Precision zu gewährleisten, bedarf es der transparenten Auswertung für Retrieval und Anzeige. Anforderung ist eine benutzerfreundliche Visualisierung u. A. durch die Orientierung an den IFLA LRM User tasks, die Homogenisierung der Vielzahl unterschiedlicher inhaltlicher Metadaten sowie Überschaubarkeit, Vollständigkeit und Präzision von Treffermengen (siehe Punkt 6).

Inhaltliche Metadaten und Normdaten sind nicht nur für das klassische Retrieval im bibliografischen Kontext interessant. Das in Metadaten abgelegte vernetzte Wissen kann selbst zum Forschungsobjekt werden oder die Bearbeitung von Forschungsfragen unterstützen (siehe Punkt 7).

¹ Vgl. dazu das [Protokoll der 34. Sitzung des Standardisierungsausschusses am 10. Dezember 2018](#), S. 13.

Eignung der GND und des Regelwerks für maschinelle Verfahren – Executive Summary

Die Themenkomplexe „Eignung der Normdatei GND für maschinelle Verfahren“ und „Eignung des Regelwerks für maschinelle Verfahren“, vor allem im Zusammenspiel mit intellektueller Erschließung, wurden von der Unterarbeitsgruppe getrennt voneinander betrachtet.

Um die Maschinenlesbarkeit aller Informationen eines Normdatensatzes zu erhöhen, sollten alle Elemente des Datensatzes eigene semantische Einstiegspunkte darstellen. Bislang nur in codierter Form vorliegende Informationen am Datensatz sollten als echte Relationen im Sinne des Entity-Relationship-Modells ausgebaut werden. Maschinelle Verfahren bedürfen einer eindeutigen Disambiguierung des kontrollierten Vokabulars. Um dem gerecht zu werden, sollten folgende Maßnahmen umgesetzt werden: Verbesserung und Ausbau der Relationierungen bzw. Hierarchisierung der Datensätze; Verfeinerung der GND-Systematik; stärkere Normierung bei der sprachlichen Gestaltung von Verwendungshinweisen.

Im Hinblick auf die Vergleichbarkeit der Ergebnisse von maschineller und intellektueller Erschließung ist zu diskutieren, welcher Struktur und welcher Logik ein zu erarbeitendes Regelwerk folgen sollte. Neben dieser Vergleichbarkeit sind die Möglichkeiten der Maschinenlesbarkeit des Regelwerks als Regelwerkstext zu prüfen. Unserer Meinung nach wäre es dazu zwingend notwendig, computerlinguistische Kenntnisse und Kompetenzen einzuholen.

Sacherschließung am Werknormdatensatz – Executive Summary

Um die möglichen Effizienzgewinne einer Verlagerung der Sacherschließungsinformationen von der Ebene der Titeldatensätze auf die der Werknormdatensätze realisieren zu können, müssten künftig grundsätzlich und umfassend Werknormdatensätze angelegt und mit den Titeldatensätzen verknüpft werden; auch für die Altdaten müsste dies nachvollzogen werden. Für eine erfolgreiche Implementierung müssten außerdem verschiedene weitere technische und organisatorische Voraussetzungen erfüllt sein. Das Expertenteam RAVI hat diese gesammelt, diskutiert und analysiert. Das Ergebnis ist, dass der Aufwand bei der Umstellung und der künftigen Erschließung in der täglichen Praxis den zu erwartenden Nutzen erheblich übersteigen würde. Auch könnte nicht sichergestellt werden, dass der derzeitige Stand beim Retrieval unter den veränderten Rahmenbedingungen erhalten bliebe. Das Expertenteam empfiehlt deshalb, die Idee zum jetzigen Zeitpunkt nicht weiter zu verfolgen.

Definition allgemeiner verfahrensunabhängiger Qualitätskriterien und ggf. Verfahren zur Qualitätssicherung (ET RAVI)

1. Ausgangslage

In den vergangenen Jahren haben sich die Rahmenbedingungen für die bibliothekarische Kernaufgabe der Inhaltsererschließung deutlich verändert und erweitert. Die zunehmende Menge von Materialien und Publikationsformen und die Heterogenität der Daten birgt zahlreiche immanente Möglichkeiten, aber auch Notwendigkeiten für die Entwicklung und den Einsatz neuer Erschließungsverfahren. Im selben Maße sind die Anforderungen an die inhaltliche Aufbereitung von Ressourcen und die Anwendungsbereiche gewachsen. Regelwerke und Normdateien öffnen sich für Anwendergruppen und Nutzungsszenarien jenseits der Bibliotheken. Die traditionelle intellektuelle Sacherschließung nach den Regeln für die Schlagwortkatalogisierung oder nach einem Klassifikationssystem ist nur noch ein Teil in einem Nebeneinander verschiedener Verfahren und Erschließungssysteme. In erheblichem Umfang sind Bestände nicht inhaltlich erschlossen und entziehen sich der thematischen Auffindbarkeit in der Menge der erschlossenen Bestände. Dementsprechend ist die aktuelle Fachdiskussion über zukunftsfähige Verfahren zur Inhaltsererschließung geprägt von unterschiedlichen Sichtweisen auf die Bedeutung, den Nutzen und die Relevanz von Regelwerken und kontrollierten Vokabularen. In diesem vielschichtigen Kontext spielt der Qualitätsbegriff die zentrale Rolle. Damit unterschiedliche Verfahren zukünftig nicht nur

nebeneinander existieren, sondern unter optimaler Ausnutzung ihrer Stärken zusammenwirken können, bedarf es der Konsolidierung der inhaltlichen Erschließung zugrundeliegenden Qualitätskriterien. Ziel ist die Ablösung bisher hilfswise herangezogener Bewertungskategorien wie z. B. gut, schlecht, nützlich, falsch, präzise, grob durch verlässliche operationalisierbare Definitionen.

2. Auftrag

Auf Basis der Ausweitung des Auftrags an das Expertenteam „RDA-Anwendungsprofil für die verbale Inhaltsererschließung“ (ET RAVI) durch den Standardisierungsausschuss hat sich das Expertenteam im Jahr 2019 in seiner Arbeit auf die „Definition allgemeiner verfahrensunabhängiger Qualitätskriterien und ggf. Verfahren zur Qualitätssicherung“ konzentriert. Gemäß diesem Auftrag hat das Expertenteam die nachstehenden Ergebnisse erarbeitet.

Verfahrensunabhängige Qualität besteht nicht in der Angleichung der Inhalte unterschiedlicher Wissensorganisationssysteme und nicht in der Angleichung der Anwendung, sondern durch die Verbindlichkeit von Standards für die Gestaltung von Wissensorganisationssystemen und für deren Anwendung. Damit wird Interoperabilität der heterogenen Systeme und in heterogenen Umgebungen ermöglicht.

Die Qualität inhaltlicher Erschließung umfasst mindestens die folgenden Qualitätsdimensionen: Sie resultiert aus den übergeordneten Grunddimensionen Transparenz und Verlässlichkeit, aus der regelbasierten Produktion von Normdaten, aus Verwendungsregeln für die Ressourcenbeschreibung, aus der transparenten Auswertung für Retrieval und Anzeige und aus der Öffnung der Daten für die Nachnutzung in anderen Wissenschaftskontexten jenseits des klassischen Retrievals.

3. Grunddimensionen: Verlässlichkeit und Transparenz

Als Grunddimensionen von Qualität gelten generell die Anforderungen Verlässlichkeit und Transparenz. Dabei versteht sich Verlässlichkeit im Sinne von Konsistenz der Daten, der Regeln und des Retrievals. Transparenz beinhaltet die Offenlegung von Regeln und Standards, die genaue Herkunftskennzeichnung von Daten (Erschließungsmethode, Erschließungslevel, Konfidenz) sowie das Verdeutlichen von Erschließungslücken und heterogenen Erschließungssituationen. Durch die optimale Nutzung der Stärken der unterschiedlichen Erschließungssysteme und -methoden wird ein hohes Maß an Benutzerfreundlichkeit erzeugt. Differenzierte Recherchemöglichkeiten verbessern die discoverability unserer Bestände. Weiterhin dienen die beiden Grunddimensionen in zunehmendem Maße auch für die Nachnutzung unserer Daten in anderen Wissenschaftskontexten jenseits des klassischen Retrievals, wie z. B. für die Digital Humanities oder als Grundlage für lernende Verfahren.

4. Qualitätsdimension I: Regelbasierte Produktion von Normdaten

Voraussetzung für die qualitative inhaltliche Erschließung, sei sie intellektuell oder maschinell, sind Normdateien, Thesauri, Klassifikationen und Ontologien. Gegenüber einer reinen Verstichwortung hat die Nutzung solcher Wissensorganisationssysteme Vorteile für die Erzeugung von Qualität, z. B. durch die Herstellung von Kontext. Es bedarf insofern der Regelung bei der Erstellung von Normdaten.

Dabei sind Anforderungen an Qualität:

- Eindeutige Bestimmung des Begriffsumfangs der einzelnen Konzepte
- Disambiguierung im Verhältnis zu anderen Konzepten
- GND-spezifisch: Kritische Überarbeitung der Zerlegungskontrolle, vor allem für solche Normdatensätze, welche mit Verfahren der Named-Entity Recognition exakt zuordnen werden können
- Konsistente terminologische Kontrolle und Anreicherung durch Synonyme
- Relationierung und Hierarchisierung der Datensätze
- Gebräuchlichkeit der Terminologie, die beim Retrieval sowohl den Bedarfen von fachlichen „Laien“ auf der einen Seite bis hin zu den Expertinnen und Experten auf der anderen Seite gerecht wird
- Aktualität und Wissenschaftsnähe
- Ausrichtung an internationalen Standards, wie z. B. der Thesaurus-Norm ISO 25964 oder Resource Description and Access (RDA)
- Maschinenlesbarkeit aller Attribute der einzelnen Normdatensätze und eine geeignete Aufbereitung für das Semantic Web
- Verknüpfung mit anderen Normdaten (Linked Data, semantische Dichte) für die Ermöglichung der logischen Schlussfolgerung durch Mensch und Maschine bzw. zur Navigation in den Daten.

5. Qualitätsdimension II: Verwendungsregeln für die Ressourcenbeschreibung

In einem kooperativen Erschließungskontext sind unterschiedliche Bibliothekstypen – Spezialbibliotheken, Universalbibliotheken, öffentliche Bibliotheken –, Bibliotheksverbünde und ggf. nicht-bibliothekarische Akteure aktiv, die untereinander Daten austauschen. Ein gewinnbringender Datenaustausch ist dann möglich, wenn sich die einzelnen Netzwerkpartner auf gemeinsame Standards für die inhaltliche Ressourcenbeschreibung einigen. Weiterhin werden z. T. konsortiale Softwarelösungen großer Anbieter zur Recherche eingesetzt, die sich für die Optimierung der Suchwerkzeuge an verlässlichen Standards ausrichten.

Bedarfe an die Inhaltserschließung von Fachdomänen, von Institutionen, ihrer Benutzerinnen und Benutzer und in Abhängigkeit vom Dokumententyp können unterschiedlich sein. Verwendungsregeln für die Ressourcenbeschreibung sind vor diesem Hintergrund zu flexibilisieren, um diesen unterschiedlichen Bedürfnissen gerecht zu werden.

Für das Gelingen dieser Flexibilisierung und den optimalen Einsatz der besonderen Stärken der unterschiedlichen Verfahren und Systeme für Retrieval, Anzeige, Facettierung und Nachnutzung der Daten bedarf es im Sinne der Grunddimension Transparenz der maschinenlesbaren und standardisierten Kennzeichnung. Eingesetzte Verfahren, Wissensorganisationssysteme und Erschließungstiefe sind idealerweise aus den einzelnen Datensätzen auslesbar.

Anforderungen an Verwendungsregeln sind:

- Verständlichkeit der Verwendungsregeln
- Gut nachvollziehbare Gliederung und Gestaltung des Richtlinien textes
- Ermöglichung einer flexiblen Handhabung für unterschiedliche Szenarien durch die Einführung von abgestuften und standardisierten Erschließungsleveln mit unterschiedlichen Präzisionsgraden unter Berücksichtigung eines Minimalstandards
- Einsatz international anerkannter und sich weiterentwickelnder Wissensorganisationssysteme (GND, DDC, RVK, BK, ...)
- Präzise inhaltliche Beschreibung einer Ressource in ihrer Gesamtheit
- Ermöglichung der Beschreibung von Teilaspekten (minor terms)

- Ermöglichung einer weniger granularen Beschreibung (broader terms)²
- Maschinenlesbare Kennzeichnung der Erschließungstiefe
- Ausrichtung der Verwendungsregeln an der Eignung der Ergebnisse als Trainingsdaten für maschinelle Verfahren

6. Qualitätsdimension III: Transparente Auswertung für Retrieval und Anzeige

Ziel der Inhaltserschließung ist unter Orientierung an den Bedarfen unterschiedlicher Zielgruppen die optimale Auffindbarkeit von Ressourcen im Zusammenspiel von Recall und Precision. Metadaten der Inhaltserschließung sind dabei die Grundlage für die thematische Suche mit iterativer Herangehensweise. Differenzierte Recherchemöglichkeiten der Suchinstrumente verbessern die discoverability unserer Bestände.

Insofern bedarf es einer benutzerfreundlichen Visualisierung sowie der Offenlegung und der nachvollziehbaren Dokumentation der Erschließungssituation auch im Suchwerkzeug. Informationen zum Einsatz unterschiedlicher Wissensorganisationssysteme, zu Erschließungslücken sowie zu unterschiedlichen Erschließungstiefen sind auch für Nutzerinnen und Nutzer vorzuhalten.

Qualitätsanforderungen für die Gestaltung von Suchwerkzeugen sind:

- Orientierung an den IFLA LRM User tasks
- Sinnvolle Normalisierung und Indexierung, Enrichment
- Auswertung von Cross-Konkordanzen und Relationierungen beim Publishing von Titeldaten zur Homogenisierung der Vielzahl unterschiedlicher inhaltlicher Metadaten in Discovery Systemen
- Optimale Auswertung inhaltlicher Metadaten für das Relevanz-Ranking
- Kontextsensitive Ergebnismengen auf Basis von Benutzer*innen-Profilen bzw. fachlichen Bedarfen
- Überschaubarkeit, Vollständigkeit, Präzision von Treffermengen
- Verständlichkeit der Anzeige inhaltlicher Metadaten
- Angebot von weiterführenden Funktionalitäten auf Basis inhaltlicher Metadaten: Suchausweitung mit Oberbegriffsrelationen oder verwandten Begriffen, Verlinkung in Drittsysteme wie Wikipedia oder andere Referenzwerke
- Anzeige der Relationen als Knowledge Graph oder Themenkarten und Ermöglichung der Navigation an Hand von Normdaten
- Sinnvolle Anordnung und Abfolge von Facettierung, Entwicklung neuer Facetten auf Basis von zeitlichen und geografischen Angaben in Metadaten
- Einfach zu nutzender thematischer Sucheinstieg jenseits des Google-Schlitzes (Browsen, Warenkorb, Karte, Zahlenstrahl)
- Einbindung alternativer inhaltlicher Erschließungselemente (Links zu Klappen- und Volltexten, Inhaltsverzeichnissen etc.)
- Usability, kontextsensitive Hilfe

7. Qualitätsdimension IV: Öffnung der Daten für die Nachnutzung in anderen Wissenschaftskontexten jenseits des klassischen Retrievals

Inhaltliche Metadaten und Normdaten sind nicht nur für das klassische Retrieval im bibliografischen Kontext interessant. Die Vernetzung der GND mit anderen Wissensorganisationssystemen – wie den Library of Congress Subject Headings, VIAF, Wikidata, um nur einige wenige zu nennen – und die Nachnutzung der GND im Bereich von Forschungsprojekten und Digital Humanities nimmt zu. Gleichzeitig kann das in Metadaten abgelegte vernetzte Wissen selbst zum Forschungsobjekt

² Diese kann ggf. auch maschinell aus der präzisen inhaltlichen Erschließung erzeugt werden.

werden oder die Bearbeitung von Forschungsfragen unterstützen. Die im Kontext von Forschungsdaten entstandenen FAIR-Prinzipien³ können als ein Ausgangspunkt für die Definition von Anforderungen dieser Dimension sein. Für die Öffnung der bibliothekarischen Datenwelt und die Bereitstellung von Linked Open Library Data sowie für die Vernetzung in der Linked Open Data Cloud sind die genannten Anforderungen an Qualität insofern auch von Relevanz.

8. Messbarkeit

Nachweisbare Qualität liegt dann vor, wenn der Erfüllungsgrad von Anforderungen objektiv messbar ist. Insofern ist die Frage der Messbarkeit von Qualität der inhaltlichen Erschließung abhängig davon, ob geeignete Metriken entwickelt werden können, um die Umsetzung der obigen Überlegungen zu überprüfen.

Ansatzpunkte für die Messung der Qualität können sein:

- Grad der Abdeckung inhaltlich erschlossener Medien in bibliografischen Metadaten-Pools
- Anteil der Datensätze mit eindeutiger Kennzeichnung von Erschließungsmethode etc.
- Zahl der Relationierungen zwischen Normdatensätzen im selben System, Zahl der Beziehungen zwischen dem eigenen System und anderen Systemen
- Anteil der tatsächlich in der Recherche genutzten Informationen in Normdatensätzen im Vergleich zu den dort enthaltenen (ohne Berücksichtigung von Informationen, die nur für Erschließerinnen und Erschließer von Bedeutung sind)
- Nachvollziehbarkeit des Zustandekommens von Ergebnissen bzw. Treffern nach einer Recherche
- Auswertung von Sucherfolgen im Kontext thematischer Suchanfragen – z. B. Auslösen von Aktionen wie Ausleihe, Aufruf von Volltexten, Download bibliografischer Daten – oder Nachnutzung der angezeigten inhaltlichen Metadaten für weiterführende Suchen
- Im Bereich der maschinellen Erschließung werden klassische Maße wie recall, precision und F-Maß berechnet
- Nachnutzbarkeit und Nachnutzung inhaltlicher Norm- oder Metadaten für die Entwicklung und den Einsatz darauf aufbauender IT-Verfahren in anderen Wissenschaftskontexten (Metadaten als Grundlage für Forschung oder für Digital Humanities)

Eignung der GND und des Regelwerks für maschinelle Verfahren (ET RAVI)

1. Eignung der GND für maschinelle Verfahren

Unsere *Grundannahme* ist, dass zur Erhöhung der Maschinenlesbarkeit der Informationen der Normdatei, alle Elemente des Datensatzes eigene semantische Einstiegspunkte darstellen müssen.

Es können zwei Arten von bislang nur in Form eines Textstrings abgelegten Informationen identifiziert werden: Zum einen sind dies Informationen, die in codierter Form vorliegen wie z. B. die Satzart, der Entitätencode, das Teilbestands- und das Nutzungskennzeichen, den Länder- und den Sprachencode sowie die GND-Systematik.

Bei den codierten Angaben handelt es sich zwar bereits um Identifier (in dem Sinne, dass sie das jeweilige Element eindeutig identifizieren), es gilt allerdings zu überprüfen, ob deren bereits vorhandene linked-data Repräsentationen noch besser für Recherche und Retrieval nutzbar gemacht werden können. Zum anderen liegen Informationen in Freitextform vor wie z. B. die redaktionellen Bemerkungen, die Definition, die biografisch-historischen Angaben, die Benutzungshinweise sowie zu einem gewissen Grad auch die Quellenangaben.

³ Siehe Wikipedia: https://en.wikipedia.org/wiki/FAIR_data

Für die erste Gruppe von unverknüpften Datensatz-Informationen ist es notwendig, die bislang nur in codierter Form vorliegenden Informationen als echte Relationen im Sinne des Entity-Relationship-Modells ggfs. auszubauen.

Darüber hinaus bietet die GND-Systematik großes Potential zur Suchraumerweiterung, denn eine möglichst präzise Einordnung eines Datensatzes in das System der GND-Notationen erlaubt nicht nur eine präzise semantische Kontextualisierung, sondern unterstützt die für die maschinelle Erschließung essentielle Disambiguierung. Die GND-Systematik liegt im RDF-Format vor und müsste in dieser Form am Datensatz abgelegt werden, vgl. <https://d-nb.info/standards/vocab/gnd/gnd-sc.html>). Die Erleichterung der Disambiguierung der Normdatensätze ist sehr wichtig. Um dieser Anforderung gerecht zu werden, könnten unserer Meinung nach unter anderem folgende Maßnahmen umgesetzt werden: Ausbau und Verbesserung der Relationierungen bzw. Hierarchisierung der Datensätze untereinander; Verfeinerung der GND-Systematik; stärkere Normierung bei der sprachlichen Gestaltung von Benutzungshinweisen.

Für die zweite Gruppe von unverknüpften Datensatz-Informationen (Freitext) wird sich die Umsetzung der Maschinenlesbarkeit ungleich schwieriger gestalten, da es sich hier um sehr heterogene Informationen handelt, die zudem oftmals unikalen Charakter haben (so z. B. die Definition einer Entität). Ein Ansatzpunkt wäre z. B., dass Informationen, die noch aus dem GND-Projekt stammen, zukünftig aus dem Normdatensatz entfernt werden (unter der Bedingung, dass sie an anderer Stelle noch einsehbar und eindeutig identifizierbar sind).

Die Benutzungshinweise sind heterogen, z. B. werden sie unterschiedlich angewendet in Formal- und Sacherschließung, es gibt Hinweise auf Verknüpfungsregeln oder auch auf nicht mehr gültige Regeln. Die Benutzungshinweise sollten in ihrer Struktur so verfasst sein, dass sie maschinenlesbar sind. Ein gutes Beispiel dafür ist die Struktur der Hinweisdatensätze, die einer eindeutigen wenn-dann-Logik folgen.

Insgesamt zeigen die hier skizzierten Überlegungen, dass ein Voranbringen der GND in Richtung verbesserter Maschinenlesbarkeit nicht ohne eine Rückkopplung an die im Projekt GND4C entwickelte Öffnung der GND sinnvoll ist. ET RAVI wird sich daher zu geeigneter Zeit mit den Projektverantwortlichen austauschen, um die hier dargelegten praxisorientierten Vorschläge vorzustellen.

2. Eignung des Regelwerks für maschinelle Verfahren

Ein Regelwerk für die verbale Inhaltserschließung bedarf einer fein gegliederten und übersichtlichen Struktur. Es ist zu prüfen, nach welcher Logik die einzelnen Textbausteine strukturiert werden sollten (z. B. mittels wenn-dann-Logik) damit einerseits die Verständlichkeit des Textes sowohl für die intellektuelle Anwendung als auch für maschinelle Verfahren gegeben sein könnte. Andererseits könnte damit eine Vergleichbarkeit der Ergebnisse von maschineller und intellektueller Erschließung im Sinne einer Homogenisierung zu Retrievalzwecken erleichtert werden. Beispiele sollten weiter Bestandteil des Regelwerks sein, müssen aber als solche klar in maschinenlesbarer Form ausgezeichnet sein, damit Maschinen sie zuverlässig identifizieren können.

Darüber hinaus kann möglicherweise die Zerlegungskontrolle, die derzeit ein Grundprinzip des Regelwerks RSWK (Regeln für die Schlagwortkatalogisierung) darstellt, zu einem grundlegenden Problem für maschinelle Erschließungsverfahren werden. Hier gilt es zu überlegen, inwiefern die bisherigen Grundprinzipien ggfs. neu gedacht werden müssen, um für breitere Anwendungsszenarien (Stichwort: GLAM-Institutionen) sinnvoll zu funktionieren. Diese Überlegungen werden den Schwerpunkt der Arbeiten im ET RAVI in der nächsten Zeit bilden.

Zur Klärung der zweiten Frage, nämlich ob das neue Regelwerk selbst in maschinenlesbarer Form gestaltet werden kann, wäre es unserer Meinung nach notwendig, computerlinguistische Kenntnisse und Kompetenzen einzubringen. Dabei geht es nicht darum, cutting-edge-Forschung aus dem Bereich künstliche Intelligenz einzubinden, sondern es sollte geprüft werden, inwieweit

sich durch Softwarelösungen – insbesondere aus dem Bereich data-mining – das Regelwerkskonzept für die verbale Inhaltserschließung technisch anwenden lässt. Wir empfehlen daher, ein eigenes Projekt aufzusetzen um evaluieren zu lassen, inwiefern sich data-mining-Verfahren hinzuziehen lassen.

Sacherschließung am Werknormdatensatz (ET RAVI)

Erläuterung

Das Erfassen von Sacherschließungsinformation in Werknormdatensätzen ist aufgrund des FRBR-Modells naheliegend. Auch verspricht es einen Effizienzgewinn – allerdings nur bei Werken, von denen es mehrere Manifestationen gibt. Eine erfolgreiche Umsetzung ist jedoch nur möglich, wenn eine Reihe von Voraussetzungen erfüllt ist. Im Folgenden werden diese aufgelistet und kommentiert.

1. Die Sacherschließungsinformationen sind im Werknormdatensatz abbildbar

Das GND-Format bietet vom Prinzip her die Möglichkeit, die bisherige Erschließungspraxis in den 5XX-Feldern (Beziehungen zu anderen Entitäten) abzulegen. Durch das Einführen neuer Codes und Unterfelder könnte diese Voraussetzung voraussichtlich erfüllt werden.

2. Für jedes Werk existiert ein Werknormdatensatz und alle Titeldatensätze sind mit dem zugehörigen Werknormdatensatz verknüpft

Dies würde erheblichen Zusatzaufwand bei der Erschließung bedeuten, da derzeit nur bei einem kleinen Teil der Titeldatensätze eine Verknüpfung mit einem Werknormdatensatz durchgeführt wird; bei den RAK-Daten ist die Situation noch schlechter. Auch existieren für die Mehrzahl der Werke keine Normdatensätze. Eine Recherche über heterogene Bestände (mit und ohne Normdatensatzverknüpfung) würde jedoch zu keinen befriedigenden Ergebnissen führen.

3. Die Retrievalsysteme können im Werknormdatensatz abgelegte Sacherschließungsinformationen ebenso gut auswerten wie die bisher in den Titeldatensätzen abgelegte

Bei den derzeit im D-A-CH-Raum benutzten Bibliothekssystemen kann auch mit den in den Normdatensätzen abgelegten Verweisungsformen gesucht werden. Bei einer Umstellung auf Sacherschließung beim Werknormdatensatz wäre es aufgrund der dann komplexeren (zweistufigen) Verknüpfungen bei den Bibliothekssystemen Pica und ALMA derzeit nicht möglich, die Verweisungsformen suchbar zu machen. Es müsste für jedes Bibliothekssystem im D-A-CH-Raum eine Studie durchgeführt werden, unter welchen Bedingungen der jetzige Stand des Sacherschließungs-Retrieval aufrechterhalten werden könnte. In jedem Fall wäre die Umsetzung mit erheblichem technischen Aufwand verbunden, und es bestünde das Risiko, dass gerade die Systeme kleinerer Anbieter die neuen Anforderungen auch längerfristig nicht erfüllen würden.

4. Das Bedürfnis, künftig mehrere unterschiedliche Arten von verbaler Sacherschließung nebeneinander zuzulassen, muss erfüllt werden können

Bisher sind die Sacherschließungsdaten Teil der bibliografischen Beschreibung und gelten nur für den jeweiligen Katalog bzw. Verbundkatalog. Bei einer Umstellung auf Sacherschließung am Werknormsatz würden die Daten hingegen von allen Anwendern der GND im D-A-CH-Raum nachgenutzt. Schon jetzt gibt es unterschiedliche Bedürfnisse und Praktiken bei der verbalen Sacherschließung sowie ein Nebeneinander von intellektueller und maschineller Bearbeitung. Bei einem künftigen, modularisierten Erschließungskonzept wird sich dieser Trend noch verstärken. Deshalb müsste zum einen ein Nebeneinander mehrerer Arten von Erschließung am Werknormdatensatz möglich sein. Bei jeder dort abgelegten Sacherschließungsinformation müssten also auch die jeweils angewendeten Erschließungsprinzipien codiert werden. Die abnehmenden Systeme wiederum müssten mit den entsprechenden Codes umgehen können, um unter mehreren

vorhandenen Erschließungen die jeweils lokal erwünschte herauszuziehen. Auch müssten neue Absprachen getroffen werden, wie mit Sacherschließung umgegangen werden soll, wenn sie nicht der eigenen Präferenz entspricht - z. B. was tun, wenn bestimmte Bibliotheken keine Schlagwortfolgen wollen, am Werknormdatensatz aber Schlagwortfolgen sind? Was tun, wenn andere Bibliotheken eine andere Beschlagwortung wünschen (z. B. engere oder weitere)?

5. Für die Erschließerinnen und Erschließer können Workflows eingerichtet werden, die ebenso schnell und einfach sind wie der bestehende Erfassungsprozess

Bisher müssen die Sacherschließerinnen und Sacherschließer in vielen Fällen bei der Beschlagwortung keine Änderung oder Ergänzung in der GND vornehmen. Künftig müsste jede Sacherschließerin und jeder Sacherschließer schreibenden Zugriff auf die GND haben, da häufig Sacherschließungsinformation am Werknormdatensatz zu verankern wäre. Auch ist ein höherer Zeitaufwand durch den Wechsel der Systeme zu erwarten, sofern nicht spezielle Tools für eine komfortable und effiziente Erfassung unter den neuen Rahmenbedingungen entwickelt werden können.

Empfehlung

Das Expertenteam RAVI kommt angesichts der beschriebenen Anforderungen und den Ergebnissen ihrer Analyse zu dem Schluss, dass die Nachteile einer Sacherschließung am Werknormdatensatz unter den derzeitigen Rahmenbedingungen die möglichen Vorteile überwiegen würden, und empfiehlt deshalb, die Idee zum jetzigen Zeitpunkt nicht weiter zu verfolgen.

Berichtszeitraum: Dezember 2018 bis Juni 2019

Im Berichtszeitraum Dezember 2018 - Juni 2019 hat das Expertenteam RAVI den neuen Auftrag des STA zur Konzeptentwicklung modular aufgebauter Regeln zur inhaltlichen Erschließung und zur Definition allgemeiner verfahrensunabhängiger Qualitätskriterien aufgenommen und in einer ersten Präsenzsitzung, die am 11. Februar 2019 in der DNB stattfand, diskutiert. Zum Thema Qualität (in) der Erschließung wurden zunächst die Dimensionen von Qualität definiert, diese erstrecken sich auf alle Ebenen der Inhaltserschließung:

1. Erfassungsregeln für die Normdatenerstellung
2. Verwendungsregeln für die Ressourcenbeschreibung
3. Regeln und Anforderungen für Retrieval und Anzeige

Weiterhin wurde herausgearbeitet, dass Verlässlichkeit und Transparenz eine Grunddimension von Qualität darstellen. Verlässlichkeit im Sinne von Konsistenz (der Daten, der Regeln, des Retrievals) wird als Grundvoraussetzung für das Erreichen von Qualität erachtet. Transparenz beinhaltet die Veröffentlichung bzw. Offenlegung von Regeln und Standards, die genaue Herkunftskennzeichnung von Daten (Methode, Level, Relevanz) sowie das Verdeutlichen von Erschließungslücken und heterogenen Erschließungssituationen.

Die Ergebnisse des Workshops wurden von den beiden Sprecherinnen des Expertenteams in einem Vortrag auf dem diesjährigen Bibliothekskongress in Leipzig vorgestellt. Der Vortrag erhielt sehr viele positive Rückmeldungen.

Das ET RAVI bearbeitet nun die vom STA als zentral benannten Themen „Qualität“, „Eignung des Regelwerks und der Normdatei für maschinelle Verfahren“ sowie „Sacherschließung beim Normdatensatz“. Dazu wurden Kleingruppen gebildet, welche die Themen in individuellen Telefonkonferenzen diskutieren. Ergebnisse werden auf der nächsten gemeinsamen Telefonkonferenz im Juni präsentiert.

Für den Herbst 2019 ist eine zweite Präsenzsitzung an der DNB geplant, in welcher die Ergebnisse gebündelt und konkrete Arbeitsschritte für die Umsetzung in das künftige Regelwerk abgeleitet werden sollen.

Berichtszeitraum Juli 2018 bis November 2018

Im Berichtszeitraum hat das Expertenteam RAVI entsprechend dem Auftrag des Standardisierungsausschusses aus der Sitzung am 5. Juli 2018 an der Vertiefung der zu klärenden Grundsatzfragen für ein zukünftiges Regelwerk gearbeitet. Dazu wurden die jeweiligen Aspekte der Themenbereiche 2-4 (Eignung der Regelwerke für automatische Verfahren und entsprechende Anforderungen, Abbildung der IFLA LRM user tasks in den RSWK und das Konzept der Inhaltserschließung am Werknormsatz), die im letzten Sachstandsbericht des Expertenteams vom Juni 2018 vorgeschlagen wurden, priorisiert, Kleingruppen gebildet und zwei Telefonkonferenzen durchgeführt.

In der letzten Telefonkonferenz am 26. November 2018 hat sich das Expertenteam auf den mündlich durch die AfS kommunizierten Vorschlag der Initiativgruppe des Standardisierungsausschusses hin vorrangig mit der übergeordneten Frage möglicher Qualitätskriterien für die verbale Inhaltserschließung im Kontext unterschiedlicher Anwendungsprofile und Erschließungsverfahren befasst. Ein erstes Brainstorming wurde begonnen und eine entsprechende Seite im Wiki angelegt. Gemäß dem Vorschlag der Initiativgruppe werden die anderen Aspekte zunächst zurückgestellt, um das gewünschte erste Thesenpapier zu den grundlegenden Qualitätsanforderungen im Februar vorlegen zu können.

Für die Diskussion und effiziente Erarbeitung solch komplexer, konzeptioneller Fragestellungen erweist sich das Format der Telefonkonferenz allerdings als wenig geeignet. Das Expertenteam schlägt daher vor, eine Präsenzsitzung im Februar in der DNB einzuberufen und bittet um die Zustimmung des Standardisierungsausschusses. In Planung ist der 11. oder 12. Februar.

Berichtszeitraum: Januar 2018 bis Juni 2018

Im Berichtszeitraum Januar bis Juni 2018 hat das Expertenteam vier Telefonkonferenzen durchgeführt und die Arbeitsergebnisse in einem eigenen Wiki-Bereich dokumentiert. Gemäß der Vorgabe des Standardisierungsausschusses aus seiner Sitzung am 8. Dezember 2017, bis zu einer konkreten Beauftragung nur grundsätzliche Überlegungen für ein zukünftiges Anwendungsprofil anzustellen, waren die Schwerpunkte unserer Arbeit vor allem folgende Themen und Fragestellungen:

- Wie können unterschiedliche Anwendungsprofile für die inhaltliche Erschließung definiert werden? Basierend auf den drei grundlegenden Qualitätsdimensionen der verbalen Erschließung "Präzision", "Vollständigkeit" und "Themenbildung" wurden Überlegungen angestellt, auf welche Weise modular aufgebaute Erschließungsregeln entwickelt werden können, die je nach Anwendungsbereich unterschiedliche Varianten von inhaltlicher Erschließung ermöglichen und dennoch gut zusammenspielen. Aus unserer Sicht ist hier eine Zusammenschau von intellektueller und maschineller Erschließung von zentraler Bedeutung.
- Inwieweit sind die Normdatei GND und das Regelwerk RSWK in seinem bisherigen Stand für den Einsatz maschineller Verfahren geeignet? Spezifische Probleme der maschinellen Inhaltsererschließung, z. B. die starke Zerlegungskontrolle, wurden benannt und "Baustellen" der RSWK identifiziert, die bereits seit langem einer Lösung harren und die ebenfalls den Einsatz maschineller Verfahren schwierig gestalten (z. B. das Konzept der Motivschlagwörter, die nur teilweise für eine maschinelle Verarbeitung geeignete Abbildung von Zeitaspekten, Splits bei Körperschaften u. a.). Die Bearbeitung dieser Fragen sehen wir als ein wichtiges Aufgabenfeld an.
- Ausgehend von den in IFLA LRM definierten user tasks (find, identify, select, obtain, explore) wurde begonnen, die RSWK einer kritischen Prüfung zu unterziehen, um so noch stärker als bislang die Nutzerfreundlichkeit des Regelwerks in den Blick zu rücken.
- Als zukunftsweisend wurde das Konzept bewertet, Inhaltsererschließung direkt am Werknormsatz zu verankern. Es ist zu erwarten, dass künftig immer mehr Titeldaten mit Werknormsätzen verknüpft werden oder maschinell Werkcluster erstellt werden. Im Expertenteam RAVI können die formattechnischen und regelwerksseitigen Voraussetzungen geprüft und erarbeitet werden.

Die kontinuierliche Weiterbearbeitung dieser grundlegenden Themen als notwendige theoretische Vorarbeit zur erwarteten Beauftragung eines Anwendungsprofils steht in engem Zusammenhang mit der gegenwärtig in der Fachcommunity dynamisch geführten Diskussion um zukunftsfähige Anforderungen an die inhaltliche Erschließung. Auch daran sind die Mitglieder des Expertenteams RAVI intensiv beteiligt, sei es durch Mitarbeit in Gremien und AGs, durch Vorträge und Veröffentlichungen oder in Ausbildung und Lehre. Nicht zuletzt war das Thema „Inhaltsererschließung“ auf dem vergangenen Bibliothekartag durch zahlreiche Veranstaltungen prominent platziert. In der Wahrnehmung des Expertenteams hat sich erfreulicherweise die kontroverse, teils polarisierende Auseinandersetzung insbesondere des letzten Jahres hin zu einer konstruktiven, kooperativen und lösungsorientierten Diskussion entwickelt.

Das Expertenteam RAVI möchte in diesem Kontext den vorliegenden Tätigkeitsbericht an den Standardisierungsausschuss nutzen, um auf die neu gegründete Initiative „Kooperative Erschließung und Informationsversorgung im D-A-CH-Raum am Beispiel der Inhaltsererschließung“ des Standardisierungsausschusses einzugehen.

Die am 08.06.2018 veröffentlichte Stellungnahme der Initiative (IG) des Standardisierungsausschusses wird von dem Expertenteam ausdrücklich begrüßt. Wir halten es für eine gute und wichtige Entwicklung, dass damit Fragen und Herausforderungen der inhaltlichen Erschließung stärker in den Fokus der Standardisierungsarbeit kommen. Die in der Stellungnahme angesprochenen Themen sind auch in der bisherigen Diskussion des Expertenteams RAVI von

zentraler Bedeutung: Engere Zusammenarbeit und Abstimmung zwischen den Bibliotheken - insbesondere bezüglich des Datenaustauschs -, die kooperative Weiterentwicklung maschineller Verfahren sowie das Ausloten des Zusammenspiels zwischen intellektuellen und maschinellen Verfahren mit dem Ziel, die inhaltliche Erschließung insgesamt effizienter zu gestalten. Besonders wichtig sind aus unserer Sicht weiterhin die Frage der Qualitätssicherung intellektueller Erschließung und der Verbesserung der Qualität maschineller Erschließungsverfahren, die Frage nach der Abhängigkeit von Qualitätsniveau und Methode von der Materialart sowie die Frage, ob und wie verschiedene Erschließungsqualitäten und -methoden in ihrer Gesamtheit zu nutzerfreundlichen Rechercheergebnissen führen.

Gerne möchten wir unsere fachliche Kompetenz in die weitere Arbeit an den grundsätzlichen Themen einbringen. Für einen besonders wichtigen Punkt halten wir es dabei, Anwendungsbereiche für verschiedene Erschließungsmethoden (Stichwort: Methodenmix) herauszuarbeiten und Voraussetzungen für deren optimales Zusammenspiel zu schaffen.

Wir hoffen deshalb, dass der Standardisierungsausschuss seinen Auftrag an das Expertenteam in einer Weise aktualisiert und konkretisiert, dass dieser sich nicht nur auf die intellektuelle verbale Erschließung bezieht, sondern eine Gesamtschau unterschiedlicher Erschließungsverfahren, vor allem im Zusammenspiel mit maschineller verbaler Erschließung ermöglicht. Nur so ist unserer Auffassung nach eine sinnvolle Entwicklung eines RDA-Anwendungsprofils für die verbale Inhaltsererschließung möglich. Eine Zusammenarbeit und Abstimmung zwischen IG und EG RAVI würden wir angesichts der zahlreichen Berührungspunkte sehr begrüßen. Sie würde beitragen, Doppelarbeit zu vermeiden und strategisches und operatives Wissen zu bündeln.

Berichtszeitraum: Juni 2017 bis November 2017

Das Expertenteam "RDA-Anwendungsprofil für die verbale Inhaltserschließung" hat am 16. Oktober 2017 seine Arbeit in einer ersten Telefonkonferenz aufgenommen.

Gegenüber dem in seiner Formulierung unveränderten Arbeitsauftrag des STA vom 2. Dezember 2015 hat sich inzwischen die Ausgangslage gravierend verändert. Zum Zeitpunkt des Beschlusses war weder sicher (wenn auch zu vermuten), dass RDA dauerhaft keine detaillierten Regeln zur verbalen Inhaltserschließung enthalten wird, noch war abzusehen, dass es eine derart tiefgreifende Überarbeitung von RDA geben wird. Diese Überarbeitung ist zum gegenwärtigen Zeitpunkt bereits aus dem Zeitplan geraten. Ob dieser gehalten werden kann oder ob es weitere Anpassungen und Korrekturen geben wird, ist ebenfalls noch nicht abzusehen. Die Erarbeitung der Anwendungsprofile ist aber, ebenso wie die ausstehenden Arbeitspakete der Fachgruppe Erschließung, unmittelbar abhängig von der zeitlichen, strukturellen und inhaltlichen Entwicklung des 3R-Projekts (vgl. die Stellungnahme der Fachgruppe Erschließung an den STA).

Eine detaillierte Arbeitsplanung mit belastbarem Terminplan und einer gesicherten Aufwandsabschätzung kann folglich erst vorgelegt werden, wenn der neue RDA-Text vorliegt. Dies wird frühestens im Juni 2018 der Fall sein. Bis dahin hat das Expertenteam Arbeitsschritte festgelegt, welche die notwendige Grundlage für die spätere konkrete Arbeit an den Regeln bilden.

Das Expertenteam geht davon aus, dass sich basierend auf den drei grundlegenden, übergeordneten Qualitätsstandards der verbalen Inhaltserschließung, "Präzision" (d. h. Verwendung von Schlagwörtern auf der präzisen Spezifitätsstufe), "Vollständigkeit" (d. h. Abbildung aller wesentlichen Inhalte einer Ressource) und "Themenbildung" (d. h. Gruppierung von Schlagwörtern zur Repräsentation von Themen) je nach Anwendungsbereich unterschiedliche Anwendungsprofile definieren lassen. Diese Prinzipien gelten dem Anspruch nach gleichermaßen für intellektuelle und maschinelle Verfahren. Die Ausarbeitung dieser Grundidee bildet einen ersten Schwerpunkt im Arbeitsprogramm des Expertenteams. Zugleich werden auch grundsätzliche Überlegungen angestellt sowie einzelne Problembereiche eruiert und bearbeitet.

Nächste Schritte:

- Die Regeln für die Schlagwortkatalogisierung (RSWK) werden daraufhin evaluiert, welche Regeln welchen der drei Grundprinzipien entsprechen, welche Regeln also - je nach verwendetem Anwendungsprofil - weiterverwendet bzw. angepasst werden können oder nicht mehr gelten würden.
- Ebenso werden die RSWK unter dem Aspekt der IFLA LRM User Tasks betrachtet.
- Ein Ideenpool zu den Fragen, welche Anforderungen heute und zukünftig an verbale Inhaltserschließung gestellt werden und mit welchem Möglichkeiten und Regeln eine zeitgemäße thematische Erschließung erreicht werden kann, wird zusammengetragen.
- Spezifische Probleme der maschinellen Inhaltserschließung (z. B. die Problematik der starken Zerlegungskontrolle) werden identifiziert und die Einschätzung der zuständigen Kolleginnen und Kollegen der DNB eingeholt.
- Grundsätzliche Probleme der RSWK, die nach der Neuordnung der Gremien offen geblieben waren und unabhängig von RDA zu lösen sind (Motivschlagwörter, Zeitaspekte u. a.) werden bearbeitet.