

Karen Kristina Köhn

Maschinelle Sacherschließungsverfahren bei medizinischen Publikationen

Erfahrungen an der DNB

Inhaltsverzeichnis

- Maschinelle Inhaltserschließung an der DNB
- Begleitstudie zur Ausweitung maschineller Verfahren auf Printmedien am Beispiel der Sachgruppe 610
- Qualität und Qualitätsmanagement
- Fazit



Maschinelle Inhaltsererschließung an der DNB

- Maschinelle Klassifikation
 - DDC-Sachgruppen
 - DDC-Kurznotationen
- Maschinelle Beschlagwortung
 - GND-Schlagwörter
- Maschineller Abgleich
 - paralleler Netz- und Printpublikationen



Meilensteine der maschinellen Erschließung

- 2012 Maschinelle Sachgruppenvergabe bei Netzpublikationen
- 2014 Maschinelle Beschlagwortung bei Netzpublikationen
- 2015 Medizinische Kurznotationen bei Netzpublikationen
- 2017 Ausweitung maschineller Verfahren auf Printmedien

Ausweitung auf Printmedien

- Begleitstudie in der Sachgruppe 610 Medizin, Gesundheit
- Warum diese Sachgruppe?
 - aussagekräftige Titel
 - alle maschinellen Verfahren etabliert
 - zahlenmäßig großer Zugang

Maschinelle Erschließung von Printmedien

- Monografien
- Reihe B und H, die über ein Inhaltsverzeichnis verfügen
- Alle Sachgruppen mit Ausnahme der Belletristik
- Quelltext: Titel und gescanntes Inhaltsverzeichnis

Quelltext bei
Netzpublikationen:
Titel und elektronische
Volltexte

Maschinelle Klassifikation von Printmedien

- Deutsch- und englischsprachige Publikationen
- DDC-Sachgruppe
 - intellektuelle Prüfung bei Printmedien
- DDC-Kurznotation
 - nur für die Sachgruppe 610 (Medizin, Gesundheit)

Bei Netzpublikationen:
610 Medizin, Gesundheit (2015)
004 Informatik (2018)
300 Sozialwissenschaften (2018)
540 Chemie (2018)
weitere in Arbeit

Maschinelle Beschlagwortung von Printmedien

- Deutschsprachige Publikationen
 - Indexat aus GND-Schlagwörtern
- Eigene Konfiguration
 - TableOfContents (S_TOC*)
- Englischsprachige Publikation
 - Indexat aus GND- und LCSH- Schlagwörtern in Arbeit

Drei Konfigurationen bei Netzpublikationen:
S_BOD* BookOnDemand-Publikationen
S_ART* Artikel bzw. kurze Texte
S_WA* Hochschulschriften und alle anderen Objektgruppen

Qualität und Qualitätsmanagement

DDC-Sachgruppen (Begleitstudie)

	Reihe B	Reihe H	Gesamt
Testmenge	221 Titel	1429 Titel	1650 Titel
SG 610 richtig	51,1 %	89,6 %	84,5 %
SG 610 falsch vergeben	19,9 %	6,0 %	7,9 %
SG 610 nicht erkannt	29,0 %	4,4 %	7,6 %

Qualität und Qualitätsmanagement

DDC-Kurznotationen (Begleitstudie)

- [140 Kurznotationen](#)
- in ca. 68% ($\pm 8\%$) richtig vergeben
- abhängig von der vergebenen DDC-Sachgruppe
- abhängig vom Literaturlaufkommen einzelner Kurznotationen
- unabhängig von der Reihenzugehörigkeit

Qualität und Qualitätsmanagement

Indexat aus GND-Schlagwörtern (Begleitstudie)

- Stichprobenartiges Qualitätsmanagement
 - Gesamtbewertung des maschinellen Indexates
 - Einzelbewertung maschinell vergebener Schlagwörter
 - Ergänzung fehlender Aspekte
 - Optional intellektuelle Schlagwortfolge

Beispiele

Titel: Motivierende Gesprächsführung in der Parodontologie



RSWK

- Arzt
- Patient
- Gesprächsführung
- Risikofaktor
- Parodontitis
- Therapie



maschinell

- Gesprächsführung
- Parodontitis
- Parodontologie

Titel: Optimierung und Vergleich von Behandlungsplänen zur Schwerionen- und Protonentherapie für Patienten mit lokal fortgeschrittenem Pankreaskarzinom



RSWK

- Bauchspeicheldrüsenkrebs
- Schwerionentherapie
- Protonentherapie
- Bestrahlungsplan



maschinell

- Bauchspeicheldrüsenkrebs
- Bestrahlungsplan
- Protonentherapie
- Therapieplan

Titel: Diabetes im Alter - besondere Berücksichtigung von Schulung, Therapie und Autonomie des älteren Menschen mit Diabetes mellitus



RSWK

- Alter
- Diabetes mellitus
- Sekundärkrankheit
- Patientenschulung



maschinell

- Alter
- Dieb
- Diabetes mellitus
- Therapie
- Einschließung

Wovon hängt die Qualität ab?

- Vokabular aus Titel und Inhaltsverzeichnis

1. Einleitung	8
2. Grundlagen	10
2.1. Parodontitis.....	10
2.1.1. Ätiologie der Parodontitis.....	10
2.1.2. Klassifizierung und Behandlung der Parodontitis.....	12
2.1.3. Prävention der Parodontitis.....	13
2.2. Motivierende Gesprächsführung.....	14
2.2.1. MI in der Anwendung.....	14
2.2.2. Training in MI.....	16
2.2.3. Motivierende Gesprächsführung in der Zahnmedizin.....	17
2.2.4. Bewertung der motivierenden Gesprächsführung.....	20
2.3. Selbstwirksamkeitserwartung	22
2.3.1. Selbstwirksamkeitserwartung im Gesundheitswesen.....	24
Material und Methode (Teil 1: systematische Literaturübersicht).....	26
2.4. Literatursuche	26
2.5. Ergebnisse der systematischen Literaturübersicht	31
2.6. Ergebnisse der eingeschlossenen Studien.....	33
2.7. Biasrisiko der eingeschlossenen Studien	38
2.7.1. Parodontaltherapeutische Bias und MI Bias.....	38
2.7.2. MI-Bias.....	44
2.8. Charakteristiken der eingeschlossenen Studien.....	48
3. Material und Methode (Teil 2: Datenerhebung)	51
3.1. Probanden	51
3.2. Einschlusskriterien.....	51
3.3. Ausschlusskriterien.....	51

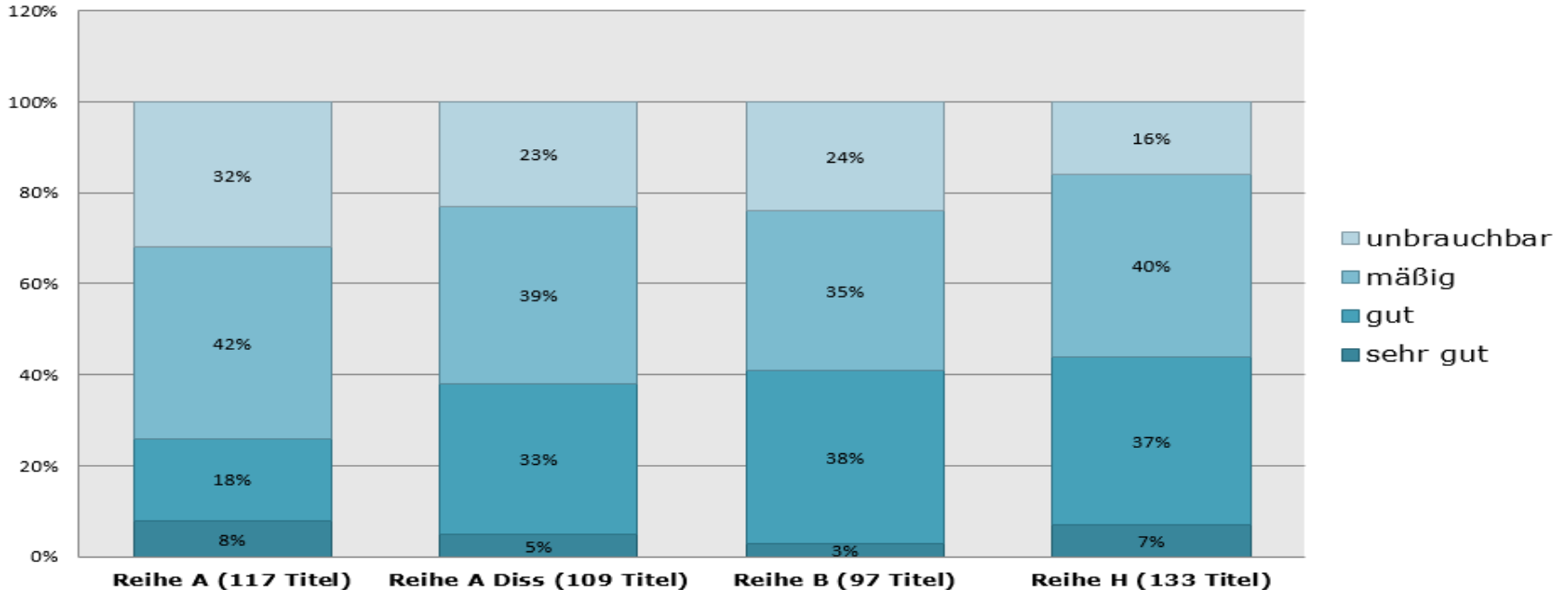
INHALTSVERZEICHNIS

Zusammenfassung	1
Einleitung	2
Publikation	5
Abstract.....	5
Introduction	7
Discussion.....	11
Future perspective	13
Conclusions	13
Appendices.....	14
References.....	20
Liste wissenschaftlicher Publikationen	24
Danksagung.....	25
Aktueller Lebenslauf	26

Wovon hängt die Qualität ab?

- Vokabular aus Titel und Inhaltsverzeichnis
- Scan und OCR (technische Qualität)
- Konfiguration der Averbis-Software
- Wörterbuchpflege und GND-Aktualität im Wörterbuch
- GND-Pflege
 - Neuansetzungen
 - Disambiguierung
 - Synonymie-Verweisungen

Gesamtbewertung Indexat Sachgruppe 610



Fazit

- Die maschinelle Beschlagwortung bietet zusätzlich zu Titelstichwörtern weitere Sucheinstiege über GND-Schlagwörter.
- Große Mengen können schnell mit GND-Schlagwörtern angereichert werden.
- Verstehen und Interpretieren von Sprache kann die Maschine (noch) nicht.

Vielen Dank für Ihr Interesse!

Haben Sie Fragen?

Karen Kristina Köhn

Deutsche Nationalbibliothek

Inhaltserschließung

Telefon: +49 69 1525-1537

k.koehn@dnb.de