

Information Behaviour bei der Erstellung systematischer Reviews

Informationsverhalten von Information Professionals
bei der Durchführung systematischer Übersichtsarbeiten
im Kontext der evidenzbasierten Medizin

Inaugural-Dissertation zur Erlangung der Doktorwürde
der Fakultät für Sprach-, Literatur- und Kulturwissenschaften
der Universität Regensburg

vorgelegt von
Anne-Kathrin Merz
aus
Würzburg

2016

Gutachter (Betreuer): Prof. Dr. Christian Wolff,
Universität Regensburg

Gutachterin: Prof. Dr. Christa Womser-Hacker,
Universität Hildesheim

Danksagung

Zutiefst dankbar bin ich Prof. Dr. Rainer Hammwöhner, der sich 2014 bereit erklärt hatte, mein Promotionsvorhaben zu unterstützen. Er war mir während meiner Forschungstätigkeit ein besonnener und weitsichtiger Mentor und trug in intensiven fachlichen Gesprächen ganz wesentlich zum Gelingen dieser Arbeit bei. Mehr noch prägte er mein wissenschaftliches Denken und fesselte mich mehr und mehr für die Facetten meines Themas und meine Forschungstätigkeit. Ich hätte mich sehr gefreut, ihm an dieser Stelle persönlich für seine Unterstützung zu danken. Leider konnte er die Fertigstellung der Arbeit nicht mehr erleben.

So gilt heute mein Dank vor allem meinem Doktorvater Prof. Dr. Christian Wolff, der angesichts dieses Verlusts kurzfristig die Nachfolgebetreuung übernahm und die Fertigstellung der Arbeit engagiert, konstruktiv und inspirierend unterstützte.

Ebenso möchte ich mich auch bei Frau Prof. Dr. Christa Womser-Hacker an der Universität Hildesheim bedanken, die sich als Zweitgutachterin zur Verfügung stellte und damit garantierte, dass die Arbeit zügig fertiggestellt werden konnte.

Ich danke Ihnen beiden ganz herzlich für Ihre Unterstützung und die vertrauensvolle Zusammenarbeit!

Nie zustande gekommen wäre diese Arbeit ohne Dr. Helge Knüttel, Fachbereichsleiter der Teilbibliothek Medizin an der Universität Regensburg. Er war vor allem in der Anfangsphase des Projekts ein wichtiger Gesprächspartner mit vielen inspirierenden Ideen. Vielen, vielen Dank!

Bedanken möchte ich mich außerdem bei allen Information Professionals, die mich durch die Teilnahme an Umfragen oder durch E-Mail Konversationen aktiv bei der Durchführung meiner Studien unterstützt oder die hierfür nötigen Kontakte hergestellt haben. Sie haben mit dem Einblick in ihre Expertise den wichtigsten Beitrag geleistet.

Unendlich dankbar bin ich meiner Familie, die mich nicht nur darin bestärkte, das Promotionsprojekt in Angriff zu nehmen, sondern durch ihre unermüdliche Geduld und ihren Rückhalt auch ganz wesentlich dazu beitrug, das Projekt zu realisieren. Ganz besonders gilt dieser Dank meiner Mutter, Brigitte Merz, die sich geduldig und kritisch der Korrektur des fertigen Manuskripts annahm.

Context is something you swim in like a fish. You are in it. It is in you.

(Brenda Dervin)

Inhaltsverzeichnis

1	Einführung	1
1.1	Problemstellung	1
1.1.1	Die Übersichtsarbeit als Forschungsmethode	1
1.1.2	Von der evidenzbasierten Medizin zur systematischen Übersichtsarbeit.....	4
1.1.3	Perspektive systematischer Übersichtsarbeiten	6
1.1.4	Anspruch der systematischen Übersichtsarbeit	10
1.1.5	Vergleichbarkeit und Abgrenzung.....	11
1.1.6	Praktische Probleme	15
1.2	Zielsetzung und Methodik.....	21
1.3	Aufbau der Arbeit	24
2	Informationswissenschaftliche Grundlagen	28
2.1	Information und Information Need	28
2.1.1	Was ist Information? Eine Einordnung.....	28
2.1.2	Warum suchen wir Information?.....	32
2.1.3	Vom Information Need zum Informationsverhalten.....	38
2.2	Information Seeking Behaviour	40
2.2.1	Modell von Ellis	41
2.2.2	Modell von Kuhlthau	46
2.2.3	Marchioninis ISP-Modell	48
2.2.4	Modell nach Byström und Järvelin.....	51
2.2.5	Modell von Leckie, Pettigrew und Sylvain.....	54

2.2.6	Modell nach Wilson (1997,1999).....	56
2.2.7	Zusammenfassung.....	57
2.3	Information Retrieval	58
2.3.1	Dokumente und Relevanz	59
2.3.1.1	Dokumente.....	59
2.3.1.2	Relevanz	61
2.3.2	Information Retrieval-Systeme	63
2.3.2.1	Architektur	63
2.3.2.2	Indexierung	64
2.3.2.3	Klassische Information Retrieval-Konzepte	68
2.3.2.4	Query Expansion und Relevance Feedback	73
2.3.2.5	Systemseitige Evaluierung von Retrieval-Systemen.....	74
2.3.2.6	Ausblick.....	77
2.3.3	Boolesches Datenbankmodell	77
2.3.3.1	Grundprinzip des Booleschen Retrieval-Modells.....	78
2.3.3.2	Retrieval in Booleschen Datenbanksystemen	80
2.3.4	Zusammenfassung.....	84
2.4	Suchstrategien auf Booleschen Datenbanken	84
2.4.1	Das Konzept der Suchstrategie	84
2.4.1.1	Suchstrategie als Repräsentation des Information Need.....	85
2.4.1.2	Suchstrategie als Informationsverhalten	86
2.4.2	Das Bates-Modell der Information Search Tactics	87

2.4.3	Operative und konzeptuelle Moves	90
2.4.4	Suchtaktiken in Online-Datenbanken	93
2.4.5	Zusammenfassung.....	96
2.5	Information Retrieval Interaktion im Kontext.....	97
2.5.1	Interaktion: Nutzer und System.....	97
2.5.2	Grundlagenmodell nach Ingwersen und Järvelin	99
2.5.3	Informationsverhalten und Aufgaben.....	102
2.5.4	Kontextdimensionen.....	106
2.6	Zusammenfassung	110
3	Die systematische Übersichtsarbeit	111
3.1	Evidenzquellen: Ergebnisse klinischer Studien	112
3.1.1	Inhaltliche Qualität	112
3.1.2	Verzerrungen in der Forschungs- und Publikationslandschaft	114
3.2	Dokumenten- und Datenverfügbarkeit	118
3.2.1	Bibliographische Fachdatenbanken in der Medizin	118
3.2.2	Volltext,- und Zitationsdatenbanken	120
3.2.3	Zugriff auf Fachdatenbanken	122
3.2.4	Graue Literatur und nicht-bibliographische Datenbanken	126
3.2.5	Klinische und pharmazeutische Studienregister	128
3.2.6	Zusammenfassung.....	130
3.3	Die systematische Übersichtsarbeit als Rechercheaufgabe	130
3.3.1	Formalisierung des Informationsbedarfs.....	130

3.3.1.1	Konzeptualisierung quantitativer Fragestellungen.....	131
3.3.1.2	Konzeptualisierung qualitativer Fragestellungen	134
3.3.1.3	Zusammenfassung	136
3.3.2	Modelle systematischer Literatursuche	137
3.3.2.1	Einordnung.....	137
3.3.2.2	Klassisches Modell der systematischen Literatursuche.....	140
3.3.2.3	Realistische Modelle.....	142
3.3.3	Qualitätsmanagement und Dokumentation	143
3.4	Strukturierte Suchprozesse.....	147
3.4.1	Begriffliche Einordnungen.....	147
3.4.2	Datenbankspezifika	150
3.4.2.1	Thesauri und Metadaten	150
3.4.2.2	Syntaxspezifika	151
3.4.2.3	Funktionalität	154
3.4.3	Methodologische Anforderungen	162
3.4.3.1	Entwicklungsprozess	163
3.4.3.2	Methodologische Suchfilter.....	166
3.4.3.3	Dokumentation.....	169
3.4.3.4	Zusammenfassung	172
3.4.4	Evaluation strukturierter Suchen	173
3.4.4.1	Referenzmenge Gold-Standard	174
3.4.4.2	Referenzmenge Ground Truth	176

3.4.5	Identifikation von Suchtermen	177
3.4.5.1	Intellektuelle Verfahren	177
3.4.5.2	Text Mining-gestützte Verfahren	178
3.4.6	Zusammenfassung	183
3.5	Formale und thematische Suchprozesse	184
3.5.1	Quellenspezifika	184
3.5.1.1	Recherche in klinischen Studienregistern	184
3.5.1.2	Die Rolle von Suchmaschinen	186
3.5.2	Methodologische Anforderung	190
3.5.3	Zusammenfassung	192
3.6	Systematische Übersichtsarbeiten als „Information Interaction in Context“- Problem 193	
3.6.1	Kontext Information und Information Retrieval-System	193
3.6.2	Kontext Arbeits- und Suchaufgaben und Methodenlehre	195
3.6.3	Nutzer und Interaktion als unbekannter Kontext	199
3.7	Forschungsfragen	202
4	Empirische Untersuchung zum Informationsverhalten	205
4.1	Forschungs- und Untersuchungsdesign	205
4.2	Vorstudie	208
4.2.1	Studiendesign	208
4.2.2	Durchführung	210
4.2.3	Methodik der Datenaufbereitung und -analyse	211

4.2.4	Ergebnisauswertung	213
4.2.4.1	Interaktionskontext.....	215
4.2.4.2	Kontextspezifische Interaktion	217
4.2.4.3	Aktivitätsmuster.....	224
4.2.5	Zusammenfassung.....	232
4.3	Konzeption der Hauptstudie	233
4.3.1	Entwicklung des Fragebogens.....	233
4.3.2	Forschungsstrategie Grounded Theory.....	239
4.3.3	Analyseprozess	241
4.4	Ergebnisse der Hauptstudie.....	243
4.4.1	Konstruktion der suchbaren Fragestellung	244
4.4.1.1	Forschungsfrage und Domänenwissen	244
4.4.1.2	Von der Forschungsfrage zum Grundgerüst des Suchprofils.....	246
4.4.2	Suchprofil und Informationsressourcen	249
4.4.2.1	Synonyme.....	250
4.4.2.2	Deskriptoren als datenbankspezifische Suchterme.....	251
4.4.2.3	Freitextterme und Phrasen	252
4.4.2.4	Datenbankspezifischer Suchterm und Suchprofil	253
4.4.3	Bewertung von Synonymen, Kandidatentermen und Suchprofil	256
4.4.3.1	Entscheidung über Synonyme als Kandidatenterme.....	256
4.4.3.2	Problematische Kandidatenterme, problematisches Suchprofil.....	260
4.4.4	Kontrolle des Entwicklungsprozesses	263

4.4.4.1	Einfache Testsuchen.....	263
4.4.4.2	Vergleichende Suchen.....	266
4.4.5	Reformulierung und Präzisierung	270
4.4.5.1	Qualitativ-orientierte Präzisierung durch Suchfilter und Facetten.....	270
4.4.5.2	Quantitativ-orientierte Präzisierung durch Reformulierung.....	273
4.4.6	Nutzung von Informationsressourcen.....	277
4.4.6.1	Ground Truth.....	277
4.4.6.2	Thesauri.....	282
4.4.6.3	Nutzerpräferenzen	284
4.4.7	Verwendung datenbankspezifischer Funktionalitäten und Text Mining-gestützter Verfahren.....	287
4.5	Zusammenfassung	288
5	Generalisierung und Einordnung	290
5.1	Modell zur iterativen Entwicklung verifizierter Suchprofile	290
5.1.1	Vom Synonym zum idealen Suchterm und Suchprofil.....	290
5.1.2	Informationsverhalten von Informationsexperten	294
5.1.2.1	Interaktion mit Information.....	295
5.1.2.2	Informationspotenzial und situativer Informationsbedarf.....	300
5.1.2.3	Iterative Query Reformulation	303
5.1.3	Informationsverhalten im Kontext von Arbeitsumfeld Expertise und Dokumentenverfügbarkeit.....	306
5.1.3.1	Der Akteur in seinem Arbeitsumfeld.....	306
5.1.3.2	Der Akteur als ausgebildeter Informationsexperte	307

5.1.3.3	Der Akteur als Datenbanknutzer	308
5.2	Einordnung.....	309
5.2.1	Gesamtkontext.....	310
5.2.2	Aufgaben und Informationsbedarf.....	312
5.2.2.1	Aufgaben der systematischen Literaturrecherche.....	312
5.2.2.2	Formaler und situativer Informationsbedarf.....	314
5.2.3	Einordnung des Informationsverhaltens	315
5.2.3.1	Prozesse der Informationssuche und -verarbeitung.....	315
5.2.3.2	Domänenspezifische Besonderheiten	317
5.2.3.3	Profession.....	320
5.2.3.4	Datenbankspezifische Kontextvariablen.....	323
5.3	Zusammenfassung	326
6	Das ideale Retrieval-System?.....	329
6.1	Ausgangssituation	329
6.2	Anregungen und weitere Forschungsaspekte	332
7	Fazit.....	339
8	Abbildungsverzeichnis	341
9	Tabellen	345
10	Literatur.....	348
11	Anhang.....	394

1 Einführung

1.1 Problemstellung

In der evidenzbasierten Medizin bilden *systematic reviews* das Fundament für mehr Transparenz in der Gesundheitsversorgung. Sie sind gleichzeitig notwendiges *Werkzeug* zur Bewertung von Therapien und wichtiger Baustein zur Sicherung und Erhöhung von Qualitätsstandards im Gesundheitswesen.

Der englische Begriff Review hat in den vergangenen Jahren auch im deutschen Sprachgebrauch Einzug gehalten und ist in der deutschen Umgangssprache durchaus geläufig. Im Duden sind sowohl die Bedeutung als „kritische Besprechung eines [künstlerischen] Produkts o. Ä.“ als auch „Titel oder Bestandteil des Titels englischer und amerikanischer Zeitschriften“ zu finden (Bibliographisches Institut GmbH Duden Verlag 2016a),- einige Fachzeitschriften verwenden den Begriff in ihrem Zeitschriftentitel. Im Folgenden wird für den Review zum einen die englische, zum anderen der deutsche Begriff „Übersichtsarbeit“ verwendet. Dieser findet sich nicht im Duden, ist jedoch als Fachterminus vor allem für systematische Übersichtsarbeiten gebräuchlich.

1.1.1 Die Übersichtsarbeit als Forschungsmethode

Reviews lassen sich durch zwei Charakteristika definieren (Cooper 1988):

1. Ein Review basiert immer auf Informationen, die bereits in Originalartikeln veröffentlicht wurden. In einem Review werden keine neuen Forschungsergebnisse vorgestellt.
2. Ein Review beschreibt die Ergebnisse aus Originalarbeiten, fasst diese zusammen bewertet, klärt oder integriert sie.

Als solche spielen Übersichtsarbeiten in der Forschung seit vielen Jahrzehnten eine wichtige Rolle.

Strukturell folgt die Übersichtsarbeit einem klassischen Schema aus fünf Phasen:

In der Phase der Problemformulierung wird die Fragestellung des Reviews ausformuliert und präzisiert. Die Literatursuche dient der Identifikation der für die Fragestellung geeigneten Literatur. Diese wird während der Literaturlauswertung hinsichtlich ihrer Relevanz überprüft und anschließend analysiert und interpretiert. Die Übersichtsarbeit endet üblicherweise mit der Präsentation der Ergebnisse, vgl. (Fettke 2006).

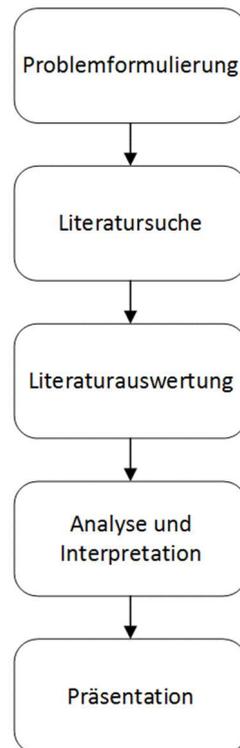


Abbildung 1-1 Struktur und Phasen einer Übersichtsarbeit nach (Fettke 2006)

In den letzten Jahren haben sich in der Forschungslandschaft terminologisch verschiedene Subtypen der Übersichtsarbeit etabliert. Sicherlich prominentester Vertreter der Übersichtsarbeit ist die Literaturübersicht – ein zunächst sehr allgemein gehaltener Begriff, der zu Beginn als Überbegriff für Reviews aller Art verwendet wurde, vgl. (Cooper 1988; Fettke 2006).

Unter einer Literaturübersicht wird im Allgemeinen eine Übersichtsarbeit verstanden, die möglichst breit angelegt ist und die Untersuchung eines aktuellen wissenschaftlichen Forschungsstands zum Ziel hat, dort aber auch Forschungslücken aufdeckt (Rowley und Slack 2004). Detaillierter findet sich eine Charakterisierung bei Cooper: Fokus der Literaturübersicht können beispielsweise wissenschaftliche Methoden und Ergebnisse sein, aber auch Theorien und

Anwendungen. Das Ziel der Literaturübersicht kann sowohl eine Verallgemeinerung, Kritik oder die Identifikation von Forschungslücken sein (Cooper 1988).

Historisch bedingt geht die Verwendung des Terminus auf die Verwendung in zwei großen bibliographischen Datenbanken zurück. Ausgangspunkt war die Definition der Literaturarbeit in der Datenbank ERIC: *“Information analysis and synthesis, focusing on findings and not simply bibliographic citations. Summarizing the substance of the literature and drawing conclusions from it”* (Cooper 1988). Psychological Abstracts als auch Educational Resources Information Center (ERIC) benennen die Identifikation von Dokumenten als *“literature review”* (Cooper 1986) und indexierten in ihren Datenbanken entsprechende wissenschaftliche Veröffentlichungen, die oben genannte Charakteristika erfüllten, mit dem gleichlautenden Deskriptor *„literature review“*.

Bereits in den 1980er Jahren hatte man sich darum bemüht, den Terminus *Literature Review* einheitlich zu definieren und zu klassifizieren (Cooper 1988).

Eine erste Klassifikation wissenschaftlicher Übersichtsarbeiten findet sich bei Harris Cooper in *“Organizing knowledge syntheses. A taxonomy of literature reviews”* aus dem Jahr 1988. Cooper schlug in der Taxonomie vor, Literaturarbeiten anhand der Kriterien Fokus, Ziel, Perspektive, Umfang, Organisation und Zielgruppe festzumachen. Diese Klassifikation wird in der Literatur weiterhin verwendet, um Subtypen von Reviews zu beschreiben (Fettke 2006). Direkte Vergleichskriterien finden sich auch bei Grant und Booth. Sie schlagen eine Klassifikation nach dem Anspruch an die vier zentralen Aufgaben der Übersichtsarbeit vor: Suche, Bewertung, Analyse und Synthese (Grant und Booth 2009).

Ein *Literature Review* kann thematisch sehr breit angelegt sein und von Originalarbeiten bis hin zu aktuellen Forschungsarbeiten jedwede Art von Ergebnissen beinhalten. Die Literaturübersicht setzt ihren Fokus dabei auf die Primärliteratur, erfordert jedoch keine formale und dokumentierte Literatursuche. Umfang und Gründlichkeit der Literatursuche sind ausschließlich vom Umfang der Fragestellung abhängig und werden nicht explizit vorgegeben, siehe auch (Webster und Watson 2002). Daher kann der Literaturüberblick in seinem Suchprozess grundsätzlich thematisch fehlerbehaftet sein oder wesentliche Primärquellen übersehen, der Anspruch der Vollständigkeit ist nicht gegeben. Im Allgemeinen erfolgt während der Analysephase keinerlei Qualitätsbewertung relevanter Literaturstellen. Die Synthese erfolgt häufig zusammenfassend in Berichtform und enthält eine meist

chronologisch oder auch thematisch zusammengefasste Analyse des Forschungsgegenstands. (Grant und Booth 2009)

Nach dieser Abgrenzung unterscheidet sich eine Form ganz entscheidend von anderen Übersichtsarbeiten: die *systematische* Übersichtsarbeit (engl. *systematic review*), im Folgenden auch SR abgekürzt, die direkt auf die Etablierung und Popularität der evidenzbasierten Medizin (im Folgenden auch abgekürzt EBM) zurückzuführen ist.

Die systematische Übersichtsarbeit zeichnet sich durch eine hochstrukturierte Methodik in allen fünf Prozessstufen (vgl. Abbildung 1-1) aus und stellt vor allem in der evidenzbasierten Medizin eine wichtige Forschungs- und Praxisgrundlage dar. Mehr noch legt sie dort einen Grundstein zur Sicherung und Erhöhung von Qualitätsstandards im Gesundheitswesen, vgl. (Jüni et al. 2001; Evans et al. 2013).

Inzwischen ist die systematische Übersichtsarbeit keine exklusive Methode der EBM mehr. Systematische Übersichtsarbeiten werden beispielsweise in den Sozial- und Erziehungswissenschaften wie auch im Software Engineering durchgeführt. Beherrscht wird das Thema jedoch nach wie vor vom Paradigma der EBM, sodass für das Verständnis der Methodik ein Einblick in die Grundlagen der evidenzbasierten Medizin hilfreich ist.

1.1.2 Von der evidenzbasierten Medizin zur systematischen Übersichtsarbeit

Die Idee einer „evidenzbasierten Medizin“ lässt sich in ihre Wurzeln bis ins 18. Jahrhundert zurückverfolgen. Das Deutsche Netzwerk Evidenzbasierte Medizin e.V. nennt den britischen Mediziner William Black als einen der Mitbegründer der Idee, erstmalig findet sich der Begriff „evidenzbasierte Medizin“ bei George Fordyce in seinem Artikel *“An attempt to improve the Evidence of Medicine”*, das 1793 erschien (Deutsches Netzwerk Evidenzbasierte Medizin e.V. 2012; Tröhler 1988).

An dieser Stelle sei auf einen wesentlichen Aspekt hingewiesen:

Während im deutschen Sprachgebrauch „Evidenz“ häufig mit Augenscheinlichkeit assoziiert ist (lat. *evidentia* = Augenfälligkeit, Veranschaulichung, Evidenz (Stowasser et al. 2007)), hat der Begriff im Kontext der evidenzbasierten Medizin eine andere Bedeutung und leitet sich

vom englischen *evidence* (= Aussage, Zeugnis, Beweis, Ergebnis, Unterlage, Beleg) ab. In diesem Sinne ist Evidenz im Deutschen als *empirische* Evidenz zu verstehen.

International geprägt wurde das Konzept erst knapp 200 Jahre später, durch Archie Cochranes Buch *Effectiveness and Efficiency: Random Reflections on Health Services*, das 1972 erschien. Archie Cochrane (1909-1988) war der erste, der die Bedeutung randomisierter Studien im Hinblick auf die Effektivität von Behandlungsmethoden herausstellte (Claridge und Fabian 2005) und sich in seinen Veröffentlichungen vehement für die Förderung und Durchführung systematischer, randomisierter klinischer Studien im Rahmen der medizinischen Versorgungsforschung einsetzte (Deutsches Cochrane Zentrum 2014). Bei diesem Studiendesign werden mindestens zwei Behandlungen an mindestens zwei Patientengruppen unter zufälliger Zuweisung parallel hinsichtlich der Ergebnisse verglichen (Windeler et al. 2008). Archie Cochrane gilt mit seinen Arbeiten heute allgemein als einer der Begründer der modernen evidenzbasierten Medizin und trug nachhaltig zum Einsatz randomisierter klinischer Studien bei.

Sackett und Rosenberg charakterisieren die EBM als Prozess lebenslangen, selbstverantwortlichen Lernens, in dem das Ziel der Patientenversorgung geprägt ist durch Wissensgewinnung und Sammlung klinisch relevanter Informationen über Diagnose, Prognose, Therapie, Entscheidungsfindung und andere Aspekte. In der Praxis bedeutet dies für den Mediziner (Sackett und Rosenberg 1995):

1. jeder Informationsbedarf wird in eine beantwortbare Frage umformuliert
2. im Vordergrund steht die effiziente Suche nach bestmöglicher Evidenz klinischer Beobachtungen, Labortätigkeiten oder der aktuellen Literatur
3. jedes Ergebnis muss kritisch hinsichtlich seiner Validität und seines Nutzens bewertet werden
4. nur Ergebnisse, die dieser Bewertung standhalten, werden in der klinischen Praxis eingesetzt
5. Bewertung der Performanz eigener medizinischer Handlungen

Sackett et al. definieren evidenzbasierte Medizin ein Jahr später als jedes medizinische Handeln, das dem individuellen Patienten eine Behandlung zukommen lässt, die explizit und gewissenhaft ermittelt auf den besten, aktuell zur Verfügung stehenden Daten und Ergebnissen beruht (Sackett et al. 1996). Diese Definition wird heute allgemein anerkannt und findet sich

in kleineren Varianten beispielsweise auch auf den Webseiten des Deutschen Netzwerks Evidenzbasierte Medizin e.V. (Deutsches Netzwerk Evidenzbasierte Medizin e.V. 2011).

Sackett betont wiederholt die Aufgabe im Rahmen des evidenzbasierten Handelns: die (regelmäßige) *systematische* Suche nach evidenzbasierten Ergebnissen in der medizinischen Literatur, sowie deren Bewertung im Hinblick auf die konkrete klinische Fragestellung (Sackett et al. 1996; Sackett und Rosenberg 1995).

An dieser Stelle finden die Bemühungen von Archie Cochrane wie auch der Gruppe um Guyatt und Sackett mit dem Begriff der „evidenzbasierten Medizin“ zusammen: Die randomisierte klinische Studie bildete die Basis für die „neue Disziplin“ der evidenzbasierten Medizin (Daly 2006).

Ausgehend von den Grundsätzen der evidenzbasierten Medizin lässt sich das Ziel einer systematischen Übersichtsarbeit folgendermaßen definieren. Ein SR ist eine nach *systematischen* Methoden durchgeführte Identifikation, Bewertung und Synthese *aller* verfügbaren Evidenz bezüglich einer vorab klar definierten medizinischen Forschungsfrage, der *„well built clinical question“* – laut Richardson et al. der Schlüssel zur EBM (Richardson et al. 1995).

In der evidenzbasierten Medizin bildet die systematische Übersichtsarbeit damit einerseits den Grundstein für mehr Transparenz in der Gesundheitsversorgung, andererseits ist sie als notwendiges *Werkzeug* und als Maßnahme zur Bewertung von Therapien ein wichtiger Baustein zur Sicherung und Erhöhung von Qualitätsstandards im Gesundheitswesen.

1.1.3 Perspektive systematischer Übersichtsarbeiten

Federführend in der Durchführung systematischer Reviews sind häufig unabhängige, nicht-kommerzielle Organisationen, dabei insbesondere die Cochrane Collaboration und ihre Zentren, die sich der Sicherung evidenzbasierter Grundsätze verschrieben haben.

Die Cochrane Collaboration (darunter auch das deutsche Cochrane Zentrum in Freiburg (The Cochrane Collaboration 2015a) als ihr deutscher Repräsentant) hat es sich zum Ziel gemacht, die wissenschaftliche Basis für Entscheidungen im Gesundheitswesen unter Beachtung evidenzbasierter Grundsätze zu verbessern. Sie ist, die Arbeiten Archie Cochranes würdigend,

nach ihm benannt und wurde 1993 als unabhängiges Netzwerk von Medizinern und Naturwissenschaftlern gegründet (The Cochrane Collaboration 2014).

Kern der Arbeiten der Cochrane Collaboration ist die Durchführung systematischer Übersichtsarbeiten (so genannter Cochrane Reviews) und deren Veröffentlichung in der Cochrane Library (John Wiley & Sons, Inc. 2016c). Die Collaboration agiert hierfür weltweit, derzeit sind 52 Review-Gruppen in der Durchführung systematischer Überblicksarbeiten aktiv. Die bei Cochrane angewandten Methoden sind in verschiedenen Cochrane Handbüchern und Manuals beschrieben, z.B. (Higgins und Green 2011; Cochrane Haematological Malignancies Group 2007; The Cochrane Public Health Group 2011; Cochrane Diagnostic Test Accuracy Working Group 2012), werden in Schulungsangeboten und Workshops gelehrt und tragen seither wesentlich zur Verbesserung der Methodik evidenzbasierter Übersichtsarbeiten bei. Mit der Cochrane Library gehört die Cochrane Collaboration zu den größten Anbietern systematischer Übersichtsarbeiten weltweit. Thematisch legt Cochrane seinen Fokus auf:

- interventionelle Reviews mit dem Ziel, Mehrwert und Gefahren spezifischer Behandlungsmethoden in der Medizin zu bewerten
- diagnostische Reviews mit dem Ziel der Bewertung diagnostischer Maßnahmen im Hinblick auf die Erkennung und Diagnose spezifischer Erkrankungen.
- methodologische Reviews zur Bewertung der Durchführung und Berichterstattung systematischer Literaturrecherchen und klinischer Studien

Systematische Methoden sind heute vor allem für die Durchführung interventioneller Reviews beschrieben, die sich der Analyse und Synthese der Ergebnisse randomisierter, kontrollierter klinischer Studien – dem Herzstück des Modells nach Cochrane – widmet, beispielsweise im *Cochrane Handbook for Systematic Reviews of Interventions* (Higgins und Green 2011).

Die Notwendigkeit zur Durchführung systematischer Übersichtsarbeiten wird im Cochrane Logo direkt sichtbar. Es zeigt Daten eines Reviews mit angegliederter Meta-Analyse zum Thema Lungenreife bei drohender Frühgeburtlichkeit. Obwohl aufgrund klinischer Studien bekannt war, dass sich die Gabe von Kortikosteroiden bei drohender Frühgeburt positiv auf das Mortalitätsrisiko der Kinder auswirkt, wurde die Methode in der Praxis nur zögerlich eingesetzt. Mit Durchführung des Reviews konnte die Evidenz des Verfahrens nachgewiesen werden, die Kindersterblichkeit sank mit Etablierung der Methode daraufhin um 30-50%. Die

Linien zeigen immer kleiner werdend die Konfidenzintervalle verschiedener Studien, die letztendlich zur Evidenzbeurteilung des Verfahrens führten (The Cochrane Collaboration 2013).



Abbildung 1-2 Cochrane Logo, mit freundlicher Genehmigung der Cochrane Collaboration

Das australische National Health Medical Research Council (NHMR) führt systematische Übersichtsarbeiten zu Fragestellungen durch, die

- den Effekt einer Intervention beurteilen,
- die Häufigkeiten von Krankheiten beurteilen,
- die Genauigkeit diagnostischer Tests beurteilen,
- ätiologische Daten und Risikofaktoren ermitteln,
- Vorhersagen über einen Gesundheitszustand machen oder
- ökonomische Faktoren von Behandlungsmethoden beurteilen,

um anhand der Ergebnisse nationale Leitlinien im Gesundheitswesen und für die klinische Praxis zu entwickeln und hat seine Methoden ebenfalls in methodischen Leitlinien festgehalten (National Health und Medical Research Council 2000).

Das Centre of Reviews und Dissemination (CRD) ist eine weitere, nicht-kommerzielle Einrichtung, die als Teil des National Institute for Health Research (NIHR) an der University of York in England verankert ist und hierüber auch finanziert wird (Centre for Reviews and Dissemination 2016). Das CRD ist international tätig und fokussiert sich ebenso auf die Nutzenbewertung medizinischer Behandlungen und Therapien sowie die weltweite Verbreitung der Ergebnisse. Das CRD betreibt außerdem eine Reihe von Fachdatenbanken, die beispielsweise auch an die Cochrane Library angeschlossen sind. Aktuell sind über die CRD Datenbanken über 35.000 systematische Übersichtsarbeiten verfügbar (Centre for Reviews and Dissemination 2015). Das CRD bietet darüber hinaus ein umfangreiches Schulungsangebot für die Durchführung

systematischer Überblicksarbeiten an und hat mit *Systematic Reviews: CRD's guidance for undertaking systematic reviews in health care* ein wichtiges Referenzwerk veröffentlicht, das derzeit in der 3. Auflage zum freien Download verfügbar ist (Centre for Reviews and Dissemination 2009).

Neben dem deutschen Cochrane Zentrum führt in Deutschland außerdem das Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen (IQWiG) (Stiftung für Qualität und Wirtschaftlichkeit im Gesundheitswesen 2016b) systematische Überblicksarbeiten durch. Das IQWiG ist ein fachlich unabhängiges, wissenschaftliches Institut mit dem Schwerpunkt der Nutzenbewertung medizinischer Maßnahmen für Patienten, deren Ergebnisse in wissenschaftlichen Berichten, aber auch allgemeinen Gesundheitsinformationen veröffentlicht werden. Aufträge an das Institut erteilen ausschließlich der Gemeinsame Bundesausschuss (G-BA) oder das Bundesministerium für Gesundheit (BMG), des Weiteren werden aber seit über 10 Jahren auch eigeninitiativ Fragestellungen bearbeiten. Finanziert wird das IQWiG aus Beitragszahlungen der gesetzlichen Krankenversicherungen (Stiftung für Qualität und Wirtschaftlichkeit im Gesundheitswesen 2016a). Die Veröffentlichungen des Instituts sind über dessen Website öffentlich zugänglich.

Systematische Übersichtsarbeiten werden inzwischen nicht nur in der Medizin durchgeführt, sondern sind überall da von wesentlichem Interesse, wo Entscheidungen für das Gemeinwohl getroffen werden müssen. Die Campbell Collaboration versteht sich als internationales, wissenschaftliches Netzwerk, das sich systematischen Übersichtsarbeiten allgemeinerer Themen widmet, beispielsweise in der Politik, in den Erziehungswissenschaften, in der Kriminalitätsbekämpfung, in der Sozialfürsorge oder in der Entwicklungshilfe (The Campbell Collaboration 2016). Als der Cochrane Collaboration angegliederte Gesellschaft hat auch sie Richtlinien zur Durchführung systematischer Übersichtsarbeiten veröffentlicht (Kugley et al. 2015).

Das in London ansässige EPPI Centre hat sich auf systematische Übersichtsarbeiten vorwiegend in den Sozialwissenschaften spezialisiert (Social Science Research Unit, UCL Institute of Education 2016). EPPI arbeitet nach seinen eigenen Richtlinien, die in Buchform unter dem Titel *An Introduction to Systematic Reviews* veröffentlicht sind (Gough et al. 2012).

Mit dem Software Engineering hat die systematische Übersichtsarbeit auch in einer anderen Disziplin Einzug gehalten. Die Motivation zur Durchführung systematischer Übersichtsarbeiten

liegt hier weniger in Entscheidungen, sondern vielmehr darin, den aktuellen Stand der Forschung zu überblicken, Lücken aufzudecken und neue Forschungsaspekte zu identifizieren (Kitchenham 2004).

1.1.4 Anspruch der systematischen Übersichtsarbeit

Die systematische Übersichtsarbeit unterscheidet sich von der Literaturübersicht vor allem durch ihren methodologischen Anspruch, der sie durch folgende strikte und unabdingbare Kriterien charakterisiert (Kitchenham 2004; Centre for Reviews and Dissemination 2009; Higgins und Green 2011):

- Basis der systematischen Übersichtsarbeit bilden immer eine spezifische, ausformulierte Forschungsfrage sowie vorab definierte Ein- und Ausschlusskriterien
- Suchprozesse haben den Anspruch, alle verfügbare Evidenz aufzufinden (Kriterium der Vollständigkeit)
- Suchprozesse erfolgen zur Vermeidung systematischer Fehler systematisch (Kriterium der Fehlerrobustheit)
- alle den Bewertungskriterien genügenden Daten werden in die Analyse und Synthese einbezogen
- alle Teilschritte systematischer Übersichtsarbeiten werden reproduzierbar dokumentiert (Kriterium der Reproduzierbarkeit)

Zur Gewährleistung der Kriterien Vollständigkeit, Fehlerrobustheit sowie Reproduzierbarkeit unterliegt die systematische Übersichtsarbeit als einzige Form der Übersichtsarbeit einer strengen Methodologie (griechisch *μεθοδολογία*), einer strengen Lehre oder Theorie wissenschaftlicher Methoden (Bibliographisches Institut GmbH 2016 Stichwort: Methodologie). Federführend sind hierbei Vertreter der evidenzbasierten Medizin, unter anderem die Cochrane Collaboration und ihre Zentren, die es sich zur Aufgabe gemacht haben, Handlungsempfehlungen wie auch Normen festzulegen, nach denen systematische Übersichtsarbeiten durchzuführen sind. Sie definieren dezidiert Aufgabenstrukturen als auch Zielsetzung und Durchführung.

Zu den wichtigsten Werken der Methodenlehre gehören

- *Cochrane Handbook for Systematic Reviews of Interventions* (Higgins und Green 2011)
- *Systematic Reviews. CRD's guidance for undertaking reviews in healthcare* (Centre for Reviews and Dissemination 2009)
- *How to review the evidence: systematic identification and review of the scientific literature* (National Health und Medical Research Council 2000)
- *Searching for studies: A guide to Information Retrieval for Campbell Systematic Reviews* (Kugley et al. 2015)

An der Entwicklung dieser Handlungsempfehlungen und Normen sind häufig interdisziplinäre Teams aus Epidemiologen und Informationsspezialisten beteiligt. Innerhalb der Cochrane Collaboration ist für die Entwicklung von Methoden und die Qualitätssicherung von Retrievalprozessen die seit 2004 aktive Cochrane Information Retrieval Methods Group (Cochrane IRMG) zuständig (The Cochrane Collaboration 2015b).

Mitarbeiter einzelner Institutionen unterliegen in ihrer Tätigkeit den Empfehlungen und Grundsätzen ihrer Institutionen. Dies gilt insbesondere für Cochrane Reviews sowie Reviews der Campbell Collaboration. 2012 wurde von der Cochrane Collaboration mit den *Methodological Expectations of Cochrane Intervention Reviews (MECIR)* ein Positionspapier veröffentlicht, in dem dezidiert festgelegt ist, welche der in Handbüchern vorgeschlagenen Methoden bei der Durchführung von Cochrane Reviews empfohlen oder erforderlich sind und wie diese Maßnahmen protokollarisch festgehalten werden müssen (Chandler et al. 2013).

Die Methodenlehre, die in Kapitel 3 der Arbeit eingehend erläutert wird, umfasst *alle* Bereiche der systematischen Übersichtsarbeit. Systematische Suchprozesse – der Fokus der vorliegenden Arbeit – sind geprägt von Methoden zur Entwicklung höchst sensitiver, gleichzeitig aber auch höchst spezifischer Suchanfragen. Diese stellen gleichzeitig sicher, dass im Gesamtprozess keine systematischen Fehler passieren, welche die Ergebnisse verfälschen könnten.

1.1.5 Vergleichbarkeit und Abgrenzung

In ihrem Anspruch an Reproduzierbarkeit, Vollständigkeit sowie eine strukturierte, fehlerrobuste Methodik in *allen* Teilschritten lässt sich die systematische Übersichtsarbeit mit anderen Formen wie beispielsweise der Literaturübersicht vergleichen und auch abgrenzen.

Critical Review

Ein kritischer Review dient typischerweise der Rezension, legt also den Fokus auf exzessive Recherche mit späterer kritischer Bewertung *einzelner* Beiträge aus der Primärliteratur. Grant und Booth definieren den kritischen Review allgemein als Methode, die typischerweise Hypothesen erhärten oder Modelle bilden sollen. Kritische Reviews finden also keine Antworten, sondern dienen der Modellbildung. Der Suchprozess erfolgt daher charakteristischerweise umfassend, um möglichst viele Informationen zu einem Thema zusammenzutragen, er kann damit ähnlich fehlerrobust wie der einer systematischen Übersichtsarbeit angesehen werden, eine spätere kritische Qualitätsbewertung findet jedoch im Allgemeinen nicht statt oder erstreckt sich auf kleine Teile der aufgefundenen Literatur. Die Synthese folge beim Critical Review keiner bestimmten Systematik und zielt mit der Analyse vor allem auf die Identifikation von Konzepten oder neuen Modellen ab. (Grant und Booth 2009)

Critical Reviews lassen sich also hinsichtlich des Suchprozesses grundsätzlich mit systematischen Überblicksarbeiten vergleichen. Mit dem Ziel, Theorien zu erhärten oder Modelle zu bilden, muss allerdings nicht zwingend eine umfassende Qualitätsbewertung stattfinden.

Scoping Review

Der Scoping Review dient als Übersichtsarbeit der Festlegung eines Untersuchungsrahmens, beispielsweise in Vorbereitung geplanter Forschungsarbeiten. Scoping Reviews werden beispielsweise durchgeführt, um die Notwendigkeit systematischer Übersichtsarbeit beurteilen zu können, werden also gerne im Vorfeld einer systematischen Übersichtsarbeit eingesetzt. Die Suche im Scoping Review orientiert sich daher typischerweise an verfügbaren Ressourcen und mündet ohne qualitative Bewertung in einer mehr oder weniger umfangreichen Zusammenfassung mit dem Fokus auf den Umfang verfügbarer Literatur. Auch hier können sich aufgrund eines nicht notwendig systematischen Vorgehens systematische Fehler einschleichen. Da der Scoping Review typischerweise einer allgemeinen Übersicht dient, ist eine Qualitätsbewertung hier nicht nötig. (Grant und Booth 2009)

Im Kontext der evidenzbasierten Medizin kann das von Arksey und O'Malley 2005 in einer Fallstudie definierte Framework als Leitfaden angesehen werden (Arksey und O'Malley 2005).

Vor allem als Vorbereitung auf systematische Übersichtsarbeiten wird der Scoping Review in der Fachliteratur seit Jahren eingehend diskutiert (vgl. beispielsweise (Levac et al. 2010; Kastner et al. 2012; Shemilt et al. 2013)), - es wird auch hier versucht, mehr Systematik in den Gesamtprozess zu bringen, um nicht nur im Systematic Review selbst, sondern bereits im Vorfeld möglichst fehlerfrei zu arbeiten. Brien et al. stellen die Unterschiede gegenüber (Brien et al. 2010):

Tabelle 1-1 Gegenüberstellung systematischer Übersichtsarbeit und Scoping Review nach (Brien et al. 2010)

Systematic Review	Scoping Review
Fokus auf einer klar beschriebenen Forschungsfrage	Forschungsfragen sind häufig breit angelegt
Ein- und Ausschlusskriterien sind vorab genau festgelegt	Ein –und Ausschlusskriterien können auch nachträglich festgelegt oder verändert werden
Qualitätskriterien werden üblicherweise vorab definiert	Qualitätskriterien haben keine Priorität
Es erfolgt eine detaillierte Datenauswertung	Eine Datenauswertung kann optional erfolgen
Enthält häufig eine quantitative Analyse	Die Analyse erfolgt vorzugweise qualitativ, typischerweise nicht quantitativ
Formale Qualitätsbewertung, die letztendlich zu einer Beantwortung der Forschungsfrage führt	Wird eingesetzt, um Merkmale und Lücken im aktuellen Forschungsstand aufzudecken

Neben weniger systematischen Verfahren gibt es hinsichtlich Zielsetzung und Methodik der systematischen Übersichtsarbeit ähnliche Verfahren. Auch sie werden in der medizinischen Forschung häufig eingesetzt und sind dementsprechend in der Fachliteratur eingehend beschrieben und diskutiert.

Meta-Analyse

Die Meta-Analyse unterscheidet sich von der systematischen Überblicksarbeit im Wesentlichen durch ihre Zielsetzung und durch damit verbundene Technik.

Das Ziel einer Meta-Analyse ist die statistische Auswertung und Beurteilung wissenschaftlicher

Ergebnisse aus quantitativen Studien, vgl. (Egger et al. 1997; Egger und Smith 1997). Sie zielt damit vor allem auf die Bewertung von Effekten ab. Daher unterscheidet sich vom SR nicht im Anspruch an den Suchprozess, sondern im Wesentlichen durch die späteren Analysemethoden sowie die Art der Darstellung im Bericht. Da die Meta-Analyse mit statistischen Verfahren arbeitet, kommen der Art und vor allem der Qualität der zur Verfügung stehenden Daten eine wesentliche Rolle zu (Grant und Booth 2009).

Egger et al. schlagen vor, den Begriff Meta-Analyse zu verwenden, wenn statistische Methoden verwendet werden, während eine systematische Übersichtsarbeit jede andere Form der Übersichtsarbeit sei, der klar definierte Methoden und Kriterien zugrunde liegen (Egger und Smith 1997). Chalmers et al. fassen Systematic Review und Meta-Analyse unter dem Begriff *Research Analysis* zusammen und setzen beides in einen gemeinsamen Kontext (Chalmers et al. 2002, S. 15):

“Systematic Review: The application of strategies that limit bias in the assembly, critical appraisal, and synthesis of all relevant studies on a specific topic. Meta-analysis may be, but is not necessarily, used as part of this process.

Meta-Analysis: The statistical synthesis of the data from separate but similar, i.e. comparable studies, leading to a quantitative summary of the pooled results”

Meta-analytische Verfahren können in der medizinischen Forschung also Bestandteil einer systematischen Übersichtsarbeit sein, um durch statistische Analysemethoden quantitative Ergebnisse zur Synthese beizutragen.

Zusammenfassend sind systematische Übersichtsarbeit und Meta-Analyse die einzigen beiden Methoden, die methodologisch klar vorgegeben und damit in der Fachliteratur auch eingehend beschrieben sind.

Rapid Review

Kritisch beurteilt wird im Kontext systematischer Übersichtsarbeiten der so genannte Rapid Review, der ebenfalls bei Grant und Booth eingeordnet wird. Historisch bedingt lässt sich die

Bezeichnung und Methodik von Rapid Reviews auf die Notwendigkeit evidenzbasierter Entscheidungen zurückführen, die aber unter zeitlich kritischen Bedingungen erfolgen. Rapid Reviews zeichnen sich durch eine höchst umfangreiche und wenig methodologische Anwendung von Suchtaktiken aus, darunter auch Methoden des systematischen Reviews, Informationen zu finden und kritisch zu bewerten, vgl. (Harker und Kleijnen 2012). Insgesamt ist der Umfang dieser Prozesse aber ressourcenabhängig, deutlich weniger gut dokumentiert und damit hinsichtlich ihrer Validität und ihrer Eignung vor allem im medizinischen Kontext nicht empirisch belegbar, siehe hierzu (Khangura et al. 2012). Heute sind Rapid Reviews definiert als Überblicksarbeiten zur Bewertung bereits bekannter Erkenntnisse (oder Behandlungsmethoden), bei der Methoden systematischer Überblicksarbeiten zum Einsatz kommen. (Grant und Booth 2009)

1.1.6 Praktische Probleme

Das strenge Protokoll macht die systematische Übersichtsarbeit in der Praxis nicht nur zu einem personal-, sondern auch zu einem zeit- und kostenintensiven Prozess:

Die in der Primärliteratur veröffentlichten Ergebnisse klinischer Studien sind nach wie vor die Hauptquelle der systematischen Übersichtsarbeit. Die per definitionem umfangreiche und erschöpfende Suche erstreckt sich daher vorwiegend auf bibliographische Fachdatenbanken, darunter auch die bekannteste medizinische Fachdatenbank MEDLINE, die unter anderem über die Metadatenbank Pubmed (National Center for Biotechnology Information (NCBI) 2016) recherchierbar ist.

Alle für die Recherche verfügbaren Fachdatenbanken basieren auch heute noch auf dem Booleschen Retrieval-Modell, das in Kapitel 2 eingehend erläutert wird. Dieses aus informationswissenschaftlicher Sicht heute durch nutzerfreundlichere, alternative Retrieval-Modelle ersetzte Konzept erfordert vom Nutzer eine exakte Formalisierung seines Informationsbedarfs sowie die Kenntnis datenbankspezifischer Syntax, vgl. (Belkin und Croft 1987). Als „*exact match*“ System liefert es den Nutzern ausschließlich exakt zur Fragestellung passenden Dokumente zurück, - in der Praxis, so die Kritik, gäbe es eine derart genau messbare Relevanz nicht (Hjørland 2010).

Genau auf diesem Kritikpunkt, der sich in der Tatsache eines häufig nicht klar formalisierbaren Informationsbedarfs von Nutzern begründet, fußt das Grundprinzip systematischer Über-

sichtsarbeiten und wird ihr hier zum Vorteil: Die initiale Definition einer klinischen Fragestellung entspricht in der Methodologie der Grundlage für eine (idealerweise weitestgehend vollständigen) Formalisierung und späteren vollständigen Reproduzierbarkeit Boolescher Suchanfragen. Gleichzeitig ermöglicht das Konzept des Booleschen Retrievals die permanente Kontrolle aller Suchvorgänge, vgl. (Hjørland 2015).

Wohl aufgrund der anspruchsvollen Retrieval-Techniken gehören Informationsspezialisten und speziell geschulte Bibliothekare zur typischen – wenn auch nicht exklusiven – Nutzergruppe bibliographischer Fachdatenbanken. Für die Literatursuche als integraler Bestandteil der systematischen Übersichtsarbeit werden daher meist Informationsspezialisten, beispielsweise Bibliothekare, hinzugezogen (McGowan und Sampson 2005).

Resultierend daraus kann die systematische Literaturrecherche als geführter Prozess definiert werden, in dem ein Nutzer mit Informationsbedarf (der Forschungsfrage) einen Experten zur Hilfestellung konsultiert – in diesem Fall zur Durchführung systematischer Übersichtsarbeiten. Der Nutzer ist Domänenexperte, beispielsweise ein forschender Mediziner mit Expertenwissen in seiner fachlichen Domäne. Der Experte hingegen ist üblicherweise ein hochtrainierter und erfahrener Bibliothekar, der von einer anerkannten Ausbildungsstätte akkreditiert wurde, vgl. (Hjørland 2015) und entsprechend über Rechercheexpertise verfügt.

Häufig werden zur Recherche vor allem im medizinischen Kontext Bibliothekare eingesetzt, die auch medizinisches Wissen mitbringen oder speziell geschult wurden. Der professionelle Suchende ist jedoch grundsätzlich weniger Domänenexperte, sondern bringt eher technische Expertise mit, vgl. (McGowan und Sampson 2005). Die systematische Übersichtsarbeit kann daher als interdisziplinäre Aufgabe verstanden werden, an der nicht nur eine Person, sondern ein Team arbeiten. Cochrane verfügt derzeit über 52 multidisziplinäre Review Teams weltweit (The Cochrane Collaboration 2016b), in denen Mediziner, Epidemiologen (beide im Folgenden Domänenexperten genannt), Biometriker, Informationsvermittler oder Bibliothekare (im Folgenden Informationsspezialisten oder Informationsexperten) zusammenarbeiten. Anmerkung: Es handelt sich beim Begriff Informationsvermittler nicht um einen in Deutschland anerkannten Ausbildungsberuf. Als eigentliche typische Berufsbilder kommen beispielsweise studierte Informationswissenschaftler oder diplomierte Informationswirte in Frage, vgl. hierzu (Bachmann 2000).

Professionelle Suche kann auch durch einen professionellen beruflichen Kontext definiert und damit vom Amateur abgegrenzt werden. Dies entspricht im englischsprachigen der Definition des Begriffs *profession* (Oxford University Press 2016e Stichwort: professional). Professionelle Suche definiert sich als Auftragsrecherche, die im Gegensatz zu Suchen aus persönlichem Interesse in Form einer Arbeitsanweisung auf professioneller, bezahlter Basis durchgeführt wird. Sie ist häufig charakterisiert durch einen klaren Suchauftrag sowie ein Budget und eventuell auch zeitliche Rahmenbedingungen (Tait 2014).

Aus diesen Rahmenbedingungen lassen sich wesentliche Charakteristika in der professionellen Suche ausmachen:

Zum einen ist die professionelle Suche durch potenzielle Missverständnisse in der Fragestellung oder mangelndes Fachwissen des Rechercheexperten immer fehleranfällig. Zum anderen ist sie als Auftragswerk weniger flexibel als eine Amateursuche und als Gesamtprozess wesentlich umfangreicher und zeitintensiver. Üblicherweise arbeiten professionell Suchende auf Basis von Protokollen und Best Practices (Tait 2014). Das Informationsverhalten in der professionellen Suche kann demnach ganz klar im Kontext der Auftragsrecherche und ihren zugrundeliegenden Aufgaben betrachtet werden. Vergleichbar mit Suchprozessen aus der Patentrecherche, die in verschiedenen Aufgabenbereichen auch immer dem Anspruch der Vollständigkeit unterliegen (vgl. (Bonino et al. 2010; Alberts et al. 2011)) oder auch der Bibliometrie (siehe (Ball und Tunger 2007)), unterscheidet sich der SR in einem Punkt ganz wesentlich: In der systematischen Literaturrecherche ist das Protokoll durch die Natur der systematischen Übersichtsarbeit strikt vorgegeben, Best Practices sind nicht nur überliefert oder gelernt, sondern in Handbüchern festgelegt oder wie im Fall der Cochrane Collaboration in Form von Normen wie dem MECIR Statement festgelegt.

In den vergangenen Jahren wird die systematische Literaturrecherche im Rahmen der Übersichtsarbeit durch den Einsatz von Informationsspezialisten nicht mehr durch die Notwendigkeit komplexer Retrieval-Techniken oder die Verfügbarkeit von Information, sondern vor allem aufgrund des exponentiellen Wachstums an verfügbarer medizinischer Primärliteratur erschwert. In der Praxis führen professionell durchgeführte Suchprozesse in bibliographischen Datenbanken häufig zu immensen Treffermengen, die in der Analysephase hinsichtlich ihrer Relevanz überprüft werden müssen, siehe (Lefebvre et al. 2011).

Abbildung 1-3 zeigt beispielhaft den jährlichen Zuwachs neu indexierter Dokumente nur der Metadatenbank Pubmed, in den Jahren 1998 bis 2015 (Corlan 2004), der gemessen ab 1955 im Jahr 2015 zu insgesamt knapp 25 Millionen indexierten Pubmed Dokumenten führt (Stand 10.04.2016). Hinweis: Über die vorgenommene Auswertung waren die Daten für das Jahr 2015 noch nicht vollständig abrufbar, der Trend ab 2014 ist in der Graphik demnach unvollständig.

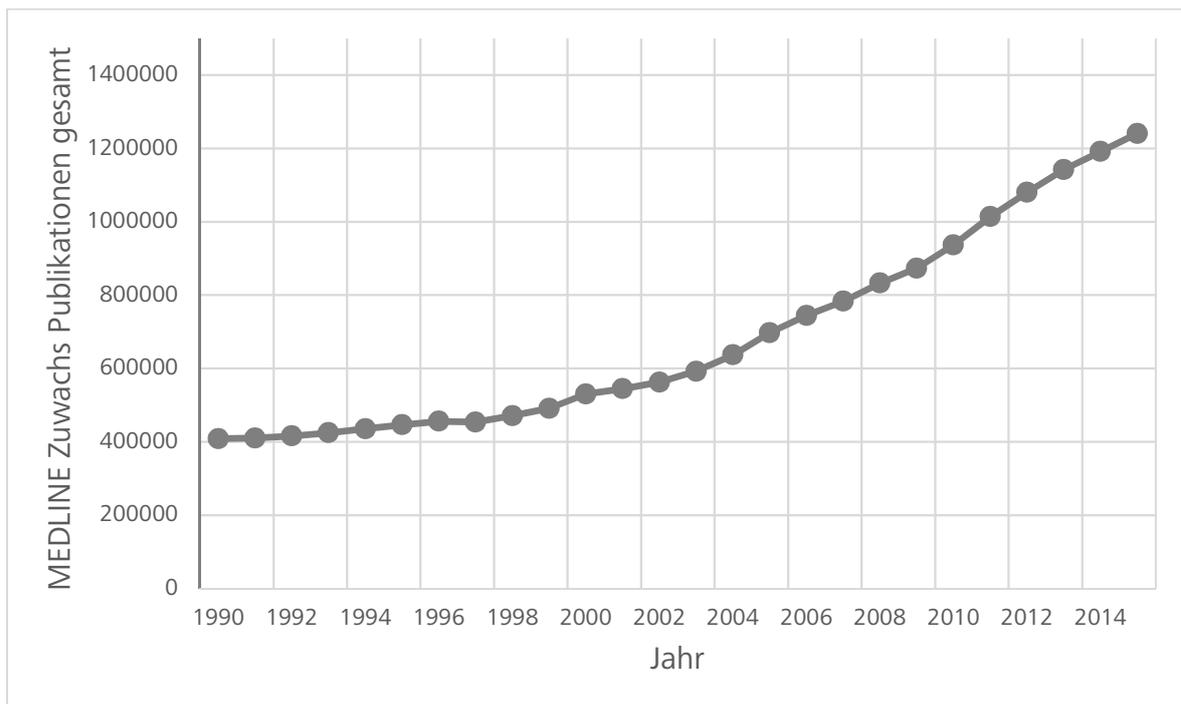


Abbildung 1-3 Jährlicher Zuwachs indexierter Dokumente in Pubmed/MEDLINE

Gleichzeitig steigt auch die Zahl der in MEDLINE verfügbaren systematischen Übersichtsarbeiten, sodass seit 1955 bis 2015 über 2.200.000 als systematische Übersichtsarbeiten indexierte Dokumente zur Recherche verfügbar sind (Quelle siehe oben, Stand 20.04.2016). Auch in Abbildung 1-4 waren über die vorgenommene Auswertung die Daten für das Jahr 2015 noch nicht vollständig abrufbar, sodass der Trend ab 2014 in der Graphik demnach ebenfalls unvollständig dargestellt wird.

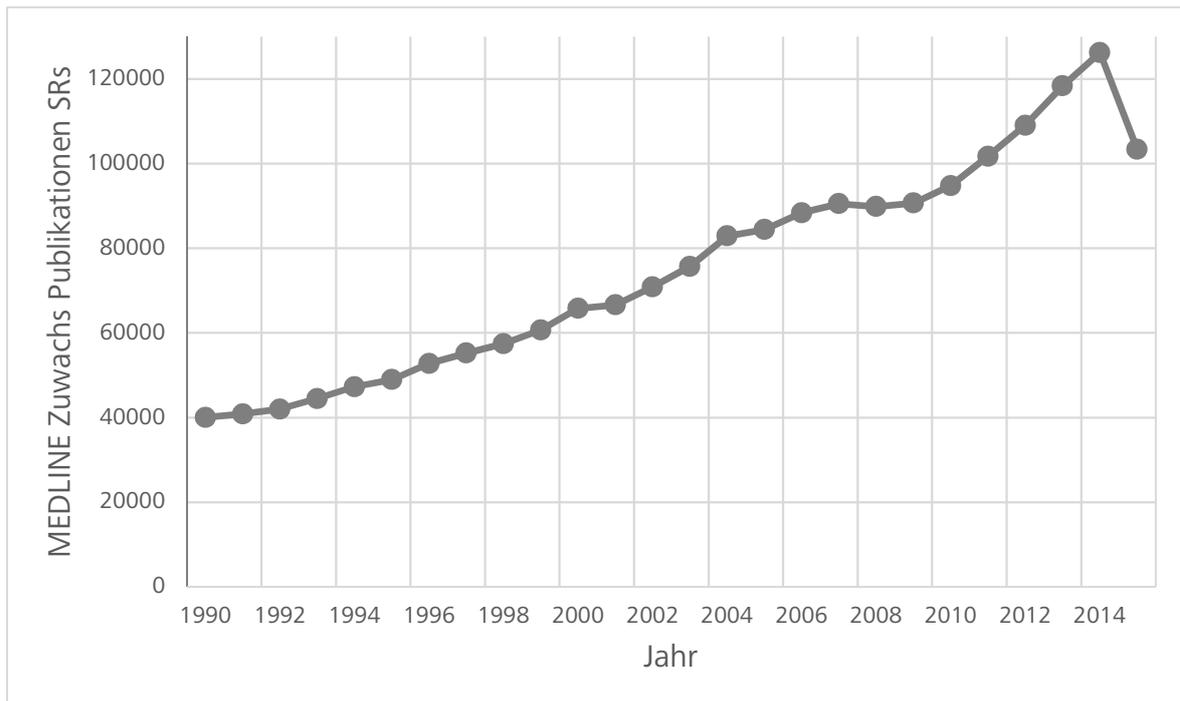


Abbildung 1-4 Jährlicher Zuwachs von als SRs indexierter Dokumente in Pubmed/MEDLINE

Gemessen an den beispielsweise von Google im Jahr 2015 indexierten 60 Trillionen Seiten, was auch in der Websuche zum Problem der Auffindbarkeit führt (Wright 2016), klingt dies verschwindend gering, vor dem Hintergrund evidenzbasierter Ergebnisse allerdings können sowohl das Übersehen wesentlicher Information als auch der Anspruch eines möglichst vollständigen Retrievals für die Gesundheitsforschung fatale Folgen haben. In der Patentrecherche gilt dies analog natürlich auch, da durch Fehler vor allem wirtschaftliche Verluste drohen.

Allen und Olkin ermittelten bereits 1999 für die methodisch vergleichbare Meta-Analyse durchschnittliche Tätigkeiten von 588 Stunden für Suche und Retrieval-Prozesse, 144 Stunden für statistische Analyse, 206 Stunden für Reporting Tätigkeiten und 201 Stunden für administrative Aufgaben (Allen und Olkin 1999) – bei einer Arbeitszeit von 8 Stunden pro Tag dauert die Suche damit im Schnitt über 73 Arbeitstage. *Are systematic reviews up-to-date at the time of publication?* fragen Fachleute (Beller et al. 2013). Sie verweisen dabei nicht nur auf die Vielzahl der zu durchsuchenden Quellen, die zur Sicherung vollständiger Retrieval-Prozesse erforderlich sind, sondern auch auf die Zeit, die zwischen letztem Suchprozess und Publikation liegt (im Mittel 8 Monate). Die Dauer in der Entwicklung von Reviews führt dazu, dass Ergebnisse häufig nicht rechtzeitig in der klinischen Praxis einsetzbar sind (Tsafnat et al.

2013). Eine ebenfalls häufig von Rechercheexperten durchgeführte Patentrecherche benötigt im Mittel 12 Stunden (Joho et al. 2010).

Erschwert wird die Suche auch durch systematische Verzerrungen in der Forschungs- und Literaturlandschaft, darunter fehlerhafte Studiendesigns wie auch Publikationsauswahlverfahren. Wie in Kapitel 3 erläutert, resultieren diese Verzerrungen beispielsweise in einer Überrepräsentation „positiver“ Studienergebnisse in der Primärliteratur, in einer überproportionalen Indexierung englischsprachiger Veröffentlichungen und anderen für Entscheidungen in der EBM kritischen Faktoren. Der Prozess der Literaturrecherche, als auch der der Analyse, sind von diesen Verzerrungen besonders betroffen (Felson 1992; Egger und Smith 1998; Song et al. 2000).

Die Suche selbst ist, wie in Kapitel 3 beschrieben, daher per definitionem geprägt von der Forderung nach einer fehlerrobusten und systematischen Vorgehensweise zur Vermeidung systematischer Fehler (Biases) - im deutschen Sprachgebrauch eine „durch falsche Untersuchungsmethoden (z. B. Suggestivfragen) verursachte Verzerrung des Ergebnisses einer Repräsentativerhebung“ (Bibliographisches Institut GmbH Duden Verlag 2016c). So kommt der Auswahl geeigneter Suchquellen bei der Durchführung systematischer Überblicksarbeiten eine wesentliche Bedeutung zu, die sich damit nicht mehr nur auf bibliographische Fachdatenbanken beschränkt, sondern auch auf so genannte klinische Studienregister und Graue Literatur ausgedehnt werden muss.

Seit Jahren werden Methoden entwickelt und erprobt, mit deren Hilfe die Durchführung systematischer Übersichtsarbeiten, insbesondere aber die Entwicklung robuster und vollständiger Suchstrategien erleichtert und beschleunigt werden kann. Hierbei spielen Text Mining-Technologien eine große Rolle, vgl. beispielsweise (Cohen et al. 2010; Thomas et al. 2011; Shemilt et al. 2013). So entstanden im Laufe der vergangenen Jahre eine Reihe an Softwarelösungen zur Automatisierung von Teilprozessen und damit ein Flickenteppich an Tools.

Neben der Tatsache, dass teilweise der Nutzen dieser Tools und Automatismen empirisch nicht belegt ist, stellt sich die grundsätzliche Frage, welche der einzelnen Teilschritte überhaupt automatisiert werden können und wie der Workload von Nutzern unter den restriktiven methodologischen Rahmenbedingungen reduziert werden kann (Tsafnat et al. 2014).

Die unter strikter Methodologie durchgeführte, systematische Übersichtsarbeit als strukturierte, professionelle Suche in riesigen bibliographischen Datenbeständen mit der Gefahr systematische Verzerrungen in Durchführung und Ergebnisinterpretation ist die Motivation für die vorliegende Arbeit.

1.2 Zielsetzung und Methodik

Warum Menschen Informationen suchen, wird informationswissenschaftlich seit Jahren eingehend beforscht. Die Forschungsergebnisse münden in unserem heutigen Verständnis des Informationsbedarfs, einem Konzept, dem in Kapitel 2 ein eigener Abschnitt gewidmet ist.

Information Retrieval-Systeme wie bibliographische Fachdatenbanken, aber auch Suchmaschinen dienen der Rückgewinnung von Information und sind damit wichtiges Werkzeug für die Suche im Rahmen der systematischen Übersichtsarbeit, vgl. hierzu im Folgenden Abschnitt 2.3.2 der vorliegenden Arbeit. Sie sollen den Informationsbedarf als Formalisierung einer vordefinierten, klinischen Fragestellung decken und damit zu Antworten und Entscheidungen in der evidenzbasierten Medizin beitragen.

Ausgehend von den frühen Arbeiten von Wilson (Wilson 1981) herrscht informationswissenschaftlich spätestens mit *The Turn* von Ingwersen und Järvelin (Ingwersen und Järvelin 2005) der Konsens, dass Konzeption und Evaluation von Information Retrieval-Systemen immer im Hinblick auf den Nutzer zu erfolgen haben und Informationsverhalten (engl. *information behaviour*) als Interaktion mit und Nutzung von Information im Kontext verschiedener Rahmenbedingungen untersucht werden sollte.

Dieser Ansatz ist das Instrumentarium der vorliegenden Arbeit:

Wie in Kapitel 2 Abschnitt 2.2 und 2.5 dargestellt, stehen informationswissenschaftlich eine Vielzahl theoretischer Modelle zur Verfügung, mit deren Hilfe das Informationsverhalten von Nutzern klassifiziert und anhand kontextueller Rahmenparameter beschrieben werden kann. Professionelle Informationsbeschaffung im Spannungsfeld von Domänenwissen und Rechercheexpertise unter derart kritischen Rahmenbedingungen wie sie die systematische Übersichtsarbeit prägen, finden dabei bisher keine Beachtung.

Vor dem Hintergrund der beschriebenen Probleme in der Durchführung systematischer Übersichtsarbeit soll daher in der vorliegenden Arbeit ein Modell zur Beschreibung aufgabenspezifischen Nutzerverhaltens im Kontext verschiedener Rahmenbedingungen entwickelt werden. Ziel ist es, zu beurteilen, welche Teilschritte der systematischen Literaturrecherche als Teilaufgabe der systematischen Übersichtsarbeit automatisiert werden können, um den ressourcenintensiven Prozess zu entschlacken und Gestaltungsempfehlungen für die im SR genutzten Informationsressourcen auszusprechen.

Vielversprechend scheinen grundsätzlich Ansätze, die professionelles Informationsverhalten im Kontext von Aufgaben beschreiben, vgl. hierzu (Leckie et al. 1996; Vakkari 2001; Byström und Hansen 2005), auch wenn diese empirisch aus dem Informationsverhalten von Domänenexperten gewonnen wurden. Ausgehend von *working tasks*, den übergeordneten Aufgaben, ergeben sich durch einen Informationsbedarf so genannte *information seeking tasks*, die wiederum in konkreten Suchaufgaben, den *searching tasks* münden. Jede Aufgabe ist dabei ein mehrstufiger Prozess, der je nach Komplexität in Teilaufgaben zerlegt werden kann. Teilaufgaben wiederum können als Teilschritte oder eine Abfolge von Aktivitäten und Aktionen interpretiert werden. Im Allgemeinen haben Aufgaben aber auch ein klar definiertes Ziel, ein Ergebnis sowie einen Start- und Endpunkt. Diese Sichtweise ermöglicht die Beschreibung von Aufgaben sowohl als statische Konstrukte durch verschiedene charakteristische Attribute (*task description*) als auch in Form von Prozessabläufen (*task processing*). Prozessabläufe informationsintensiver Aufgaben sind, wie wir heute aus der informationswissenschaftlichen Forschung wissen, als Informationsverhalten durch diverse kontextuelle Rahmenparameter beeinflusst (Byström und Hansen 2005; Ingwersen und Järvelin 2005) und können so entsprechend modelliert werden. Dieses Rahmenwerk kann zur Konzeption und Bewertung automatisierter Verfahren in der systematischen Übersichtsarbeit herangezogen werden.

Für die systematische Übersichtsarbeit sind aus der Methodologie per definitionem eine Vielzahl an *working tasks* benannt und in methodologischen Handbüchern teils als *search tasks*, den Suchaufgaben, konkretisiert. Durch die Methodenlehre und Forschung stehen ausreichend Daten zur Verfügung, die ein differenziertes Bild auf die im Kontext der systematischen Übersichtsarbeit relevanten Rahmenbedingungen liefern und diese statisch charakterisieren, darunter sowohl Evidenzquellen – jegliche Information über Ergebnisse (idealerweise) randomisierter klinischer Studien –, als auch verfügbare Informationsressourcen.

Ein umfassendes aufgabenzentriertes Modell, das zum Verständnis des gesamten informationsintensiven Prozesses beitragen kann, existiert bisher nicht, lediglich rudimentäre Ansätze zur Definition einzelner Arbeitsaufgaben sind bekannt, vgl. hierzu (Tsafnat et al. 2014; Booth et al. 2012). Die systematische Übersichtsarbeit wird bisher nicht als kontextuell geprägtes Phänomen wahrgenommen, das Informationsverhalten involvierter Domänenexperten und Informationsspezialisten wurde bisher ebenso überhaupt nicht betrachtet.

Aufgrund der Tatsache, dass das Thema aus diesem Blickwinkel bisher nicht beforscht wurde, ist ein eher explorativer Forschungsansatz notwendig, der nicht darauf abzielen kann, ein vollständiges Bild zu zeichnen. Wie sich im Verlauf der Arbeit zeigen wird, kann dieser jedoch die bisher vorwiegend system- und methodologisch orientierte Forschungslandschaft durch eine neue Sichtweise zu ergänzen.

Methodisch-empirisch muss sich die vorliegende Arbeit an den Rahmenbedingungen der systematischen Übersichtsarbeit orientieren: Wie in Kapitel 4 eingehend dargelegt, sind Informationsspezialisten als Testgruppe schwer erreichbar. Sie arbeiten international, was die Verfügbarkeit aus ökonomischen Gründen erschwert und sind beruflich in komplexe Arbeitsgebiete eingebunden, was ihre Bereitwilligkeit zur Teilnahme an zeitintensiven Studien deutlich reduziert. Zur Entwicklung dieses Modells wird daher eine mehrstufige, empirische Untersuchung aus einer eingehenden Literaturrecherche, Beobachtungsstudien sowie Fragebögen durchgeführt.

Die umfassende Literaturrecherche dient der Analyse und informationswissenschaftlichen Einordnung der aus Methodologie und Forschung bekannten Rahmenparameter sowie der Entwicklung eines rudimentären Grundgerüsts aus

1. methodologisch vorgeschriebenen Working Tasks und Search Tasks
2. einer Charakterisierung von für die systematische Übersichtsarbeit relevanten Dokumenten und Daten
3. sowie einer Grundstruktur und Charakterisierung verfügbarer und tatsächlich notwendiger Informationsquellen zum Auffinden evidenzbasiert ermittelter Ergebnisse.

Die Literaturrecherche ermöglicht außerdem die Identifikation und Benennung typischer Problemfelder, die einleitend bereits benannt wurden.

Die Beobachtungsstudie in Form einer Vorstudie soll das Grundgerüst für Konzeption der im Folgenden durchgeführten Fragebogenstudie liefern. Durch sie sollten vor allem

1. Muster in Suchprozessen identifiziert und
2. bisher ungeklärte Problemfelder abseits methodologisch benannter Aufgaben aus Nutzersicht ermittelt werden.

Die Fragebogenstudie mit Teilnehmern aus dem professionellen Umfeld als Hauptstudie der vorliegenden Arbeit dient der konkreten Vervollständigung des bis dahin ermittelten Modells zum professionellen Informationsverhalten um folgende Aspekte:

1. Benennung weiterer im Rahmen systematischer Übersichtsarbeit auftretenden Informationsbedarfe und daraus resultierender *Searching Tasks*
2. Wahrnehmung der in der systematischen Übersichtsarbeit notwendigen und in der Praxis zusätzlich anfallenden Suchaufgaben
3. Einsatz von Suchstrategien und Suchtaktiken zur Bewältigung anstehender Aufgaben
4. Persönliche, normative und kontextuelle Rahmenbedingungen des individuellen Informationsverhaltens

Die so gewählte Untersuchungsmethodik soll die in der Methodologie der systematischen Übersichtsarbeit vorwiegend bibliothekarische Sichtweise mit informationswissenschaftlichen Theorien vereinen und damit eine solide Basis schaffen, die es den Entscheidern und Machern in der Methodologie systematischer Reviews ermöglicht, die Entwicklung nutzerfreundlicher Systeme und Werkzeuge mitzutragen, die im Kontext der EBM die Durchführung systematischer Reviews erleichtern und ggf. sogar vollständig automatisieren können.

1.3 Aufbau der Arbeit

Die vorliegende Arbeit besteht abseits des aktuellen, einleitenden Kapitels aus fünf Themenblöcken und gliedert sich wie folgt:

Kapitel 2 **Informationswissenschaftliche Grundlagen** beschreibt die Basis für die informationswissenschaftliche Theorie der vorliegenden Arbeit.

Zunächst werden mit Information und Informationsbedarf grundlegende Themen aufgegriffen, die das Fundament des informationswissenschaftlichen Grundverständnisses von Informationsverhalten bilden.

Einige der in den 1980er bis 1990er Jahren theoretisch und empirisch entwickelten Modelle zum *Information Behaviour* sind sowohl für das Verständnis der gesamten Forschungslandschaft als auch für die vorliegende Arbeit von Bedeutung. Sie werden chronologisch vorgestellt und bieten dem Leser damit gleichzeitig einen historischen Abriss über die Forschungsaktivitäten nutzerzentrierten Information Retrievals.

Der darauffolgende Abschnitt widmet sich als Gegenpol der eher systemorientierten Seite des klassischen Information Retrievals und beleuchtet sowohl Dokumente und Relevanzbegriff, als auch einzelne Systemarchitekturen und theoretische Modelle der für die systematische Übersichtsarbeit relevanten Informationsquellen, darunter insbesondere Boolesches Retrieval. Wie bereits eingangs erwähnt, gehört das Boolesche Retrieval-Modell heute eher zu den antiquierten Technologien des Information Retrieval, die inzwischen aus informationswissenschaftlicher Sicht kaum mehr beforscht werden. Historisch bedingt finden sich jedoch in den Anfängen bibliographischer Datenbanken eine Reihe von Veröffentlichungen, die für das Verständnis sowohl der Methodologie systematischer Reviews als auch des späteren Nutzerverhaltens sehr hilfreich sind. Ihre Darstellung in der vorliegenden Arbeit dient gleichzeitig dem Verständnis des in der Informationswissenschaft und den Bibliothekswissenschaften sehr unterschiedlich verwendeten Begriffs der *Suchstrategie*.

Das Kapitel endet mit dem für die Arbeit zentralen Forschungsschwerpunkt: Informationsverhalten und Information Retrieval im Kontext, dem heute gängigen Ansatz, der versucht, beide Seiten – Informationsverhalten und klassisches Information Retrieval – durch die kontextuelle Betrachtung aller beteiligten Entitäten und deren Beziehung zueinander zu vereinen.

Kapitel 3 beschreibt **die systematische Übersichtsarbeit** als informationswissenschaftlichen Forschungsgegenstand und beleuchtet ihn entsprechend aus dem Kontext von „Evidenzquellen“ – verstanden als alle Information, die valide Ergebnisse klinischer Studien vorhält, Informationsressourcen und Methodologie.

Zunächst beleuchtet das Kapitel die Rahmenbedingungen.

Zum einen sind dies die Evidenzquellen, charakterisierbar durch ihre Qualität und Verfügbarkeit, beides geprägt von unterschiedlichsten Biases schon während der Durchführung klinischer Studien bis hin zur Veröffentlichung oder Nicht-Veröffentlichung von Ergebnissen. Da das Grundverständnis von Biases wesentlich die Methodologie systematischer Suchprozesse

bestimmt, werden typische, in der Fachliteratur beschriebene systematische Fehlerquellen eingehend beschrieben. Der Anspruch systematischer Literaturrecherchen macht deutlich, welche unterschiedlichen Quellen für eine erschöpfende Recherchetätigkeit genutzt werden müssen, um unverzerrte Ergebnisse zu erhalten.

Exemplarisch können in der vorliegenden Arbeit im Anschluss daran nur die wichtigsten Informationsquellen vorgestellt werden, darunter neben bibliographischen Fachdatenbanken, Volltextdatenbanken, Grauer Literatur auch klinische Studienregister, eine den systematischen Verzerrungen geschuldete Besonderheit medizinischer Berichterstattung und Recherche.

Abschnitt 3.3 beschreibt die systematische Übersichtsarbeit als Rechercheaufgabe, in der sich neben strukturierten Suchprozessen auch formale und thematische, weniger strukturierte Suchaufgaben und Suchtaktiken identifizieren lassen.

Da die Beschreibung der mit der Suche assoziierten Aufgabenstruktur untrennbar mit der Methodenlehre systematischer Übersichtsarbeiten verbunden ist, die ihre Vollständigkeit, Reproduzierbarkeit und Fehlerrobustheit garantieren sollen, skizziert dieser Abschnitt die Methodenlehre strukturierter wie auch weniger strukturierter Suchprozesse.

Das Kapitel endet mit der Einbettung der systematischen Übersichtsarbeit in ein informationswissenschaftliches Rahmenmodell, das Herausforderungen und Probleme in Retrievalprozessen der systematischen Übersichtsarbeit herausstellt und damit zu den Forschungsfragen und dem empirischen Teil der Arbeit überleitet.

Kapitel 4 **Empirische Untersuchung zum Informationsverhalten** stellt die Ergebnisse der mehrstufigen Untersuchung vor, die im Rahmen der vorliegenden Dissertationsarbeit durchgeführt wurde. Es dient der Vervollständigung des in Kapitel 3 entwickelten Rahmenmodells und der Entwicklung eines Modells zum Teilaspekten des Informationsverhaltens in der professionell durchgeführten systematischen Übersichtsarbeit.

Einleitend wird zunächst das Forschungs- und Untersuchungsdesign eingehend beschrieben, kritisch beleuchtet und begründet.

Forschungsergebnisse der verschiedenen Untersuchungsdesigns werden separat vorgestellt, wobei einleitend jeweils Studiendurchführung und Datenauswertung beschrieben werden.

Das Kapitel endet mit einer Synthese aller in der Gesamtuntersuchung gewonnenen Daten und bettet dieses in verschiedene informationswissenschaftliche Modelle ein:

Grundlage bildet die granulare Aufgabenstruktur systematischer Übersichtsarbeiten, wie sie von Nutzern wahrgenommen und in der Praxis zu beobachten ist, die in verschiedenen Prozessen Informationsverhalten mündet. Wie diese Prozesse durchlaufen werden, kann anhand

der Studienergebnisse an verschiedenen kontextuellen Rahmenbedingungen festgemacht werden.

Kapitel 5 **Generalisierung und Einordnung** erörtert die in Kapitel 4 vorgestellten Ergebnisse im Kontext informationswissenschaftlicher Forschung als auch der Methodologie systematischer Übersichtsarbeiten anhand eines Modells der „iterativen Entwicklung verifizierter Suchprofile.

Kapitel 6 **Das ideale Retrieval-System?** hinterfragt, inwiefern die vorliegenden Ergebnisse zur Reduktion zeit- und ressourcenintensiver systematischer Übersichtsarbeiten beitragen können. Wie eingehend beschrieben, betrifft dies sowohl die Automatisierung einzelner Aufgaben als auch die Empfehlung für die Gestaltung moderner Information Retrieval-Systeme, die weiterhin den hohen Ansprüchen der systematischen Übersichtsarbeit Genüge tun können, gleichzeitig aber dem Informationsbedarf und dem Informationsverhalten der Nutzer entgegenkommen. Das Kapitel schließt mit einem Ausblick auf weitere, an diese Arbeit anschließende Forschungsthemen, die in konkreten Maßnahmen und Technologien münden könnten.

2 Informationswissenschaftliche Grundlagen

Kapitel 2 dient der Einführung in die theoretischen Grundlagen der Fragestellung. Es stellt daher zunächst die historische Entwicklung vor, die zu unserem heutigen Verständnis der Konzepte Information, Informationsbedarf (*information need*) und Informationsverhalten (*Information Behaviour*) führt. Ausgehend davon werden verschiedene Modelle vorgestellt, mit deren Hilfe das Informationsverhalten von Nutzern erklärt und beschrieben werden kann.

Information Retrieval-Systeme dienen der Gewinnung von Information. Die Interaktion mit diesen Systemen ist als Reaktion auf einen vom Benutzer erkannten Informationsbedarf mit dem Ziel verbunden, Wissenslücken zu schließen. So kommt der Gestaltung von Information Retrieval-Systemen, die diese Aufgabe möglichst nutzergerecht unterstützen, eine zentrale Rolle zu. Das Kapitel führt in die technischen Grundlagen der für die Arbeit wichtigsten Konzepte des Information Retrieval ein und beleuchtet dabei sowohl die technische Systemseite als auch den Nutzer. Es schließt mit einem Ausblick auf aktuelle Forschungsansätze zur Betrachtung der „Information Retrieval Interaktion im Kontext“, mit deren Hilfe Informationsverhalten erklärt und moderne, nutzerfreundliche Information Retrieval-Systeme konzipiert als auch evaluiert werden können. Dabei stehen entsprechend der vorliegenden Arbeit informationsintensive Aufgaben professioneller Suchprozesse im Vordergrund, *Information Retrieval Interaction in Context* bildet auch die Basis für die vorliegende wissenschaftliche Arbeit.

2.1 Information und Information Need

Wollen wir Informationsverhalten beobachten und beschreiben, ist es zunächst kurz nötig, zu definieren, was Information bedeutet und wie ein „Informationsverhalten“ daraus abgeleitet werden kann.

2.1.1 Was ist Information? Eine Einordnung

Versuche der Abgrenzung und Definition des Begriffs Information nahmen in der Informationswissenschaft zu Beginn großen Raum ein, allein schon, um damit die Informationswissenschaft als Wissenschaft zu rechtfertigen und den Forschungsgegenstand zu klären (Belkin

1978). Sie münden jedoch letztendlich in der unter anderem von Belkin formulierten Ansicht, Information lasse sich nicht definieren, sondern müsse als Konzept verstanden werden (Belkin 1978):

"... we are not concerned with definitions of information, but rather with concepts of information. The distinction is that a definition presumably says what the phenomenon defined is, whereas a concept is a way of looking at, or interpreting, the phenomenon. Although there have been various reasons proposed for the significance of this distinction, the crucial one for this argument is that by accepting the idea of a concept one becomes free to look for a useful concept, rather than a universally true definition, of information."

Warum dies sinnvoll ist, lässt sich schon an Etymologie und Sprachvielfalt des Begriffs erkennen. Information lässt sich auf das lateinische *informare* (formen, gestalten, bilden; bilden, unterrichten; unterweisen, hinweisen; darstellen, schildern; sich denken, sich vorstellen) (Stowasser et al. 2007) zurückführen und wird schon dort in vielerlei verschiedenen Facetten verwendet.

Ähnlich vielfältig gestaltet sich die Verwendung des Begriffs in unserem heutigen Sprachgebrauch. Im Englischen lässt sich die Verwendung des Begriffs *information* in verschiedenen Bedeutungsvarianten bis ins 14. Jahrhundert zurückverfolgen und wird häufig im Kontext der Wissensvermittlung verwendet, dabei aber auch Dingen zugeschrieben, die aufgrund ihrer Natur „informativ“ sind. Das Oxford English Dictionary liefert uns einen guten Hinweis auf die sprachliche Vielfalt des Begriffs: *"Knowledge communicated concerning some particular fact, subject, or event; that of which one is apprised or told; intelligence, news"*, *"The action or fact of imparting the knowledge of a fact or occurrence; communication of news; notification."*, aber auch *"Contrasted with data: that which is obtained by the processing of data"* (Oxford University Press 2016c). Im Duden findet sich zur Information die Definitionen „Unterrichtung über eine bestimmte Sache“, „[auf Anfrage erteilte] über alles Wissenswerte in Kenntnis setzende, offizielle, detaillierte Mitteilung über jemanden, etwas“ als auch „Gehalt einer Nachricht, die aus Zeichen eines Codes zusammengesetzt ist“ (Bibliographisches Institut

GmbH Duden Verlag 2016i). Im Alltag müssen wir zwischen dieser erkennbaren Unterscheidung zwischen Information als Prozess und Information als Nachricht im Allgemeinen nicht differenzieren: wir wissen, was gemeint ist (vgl. auch (Case 2012), S. 38).

Ausgehend von unserem alltäglichen Sprachgebrauch lässt sich Information nach Buckland unter vier Aspekten betrachten, die Information sowohl statisch als auch dynamisch klassifizieren und noch einmal die Facetten des Begriffs verdeutlichen, unter dem Information aus informationswissenschaftlichen Sicht betrachtet werden sollte:

Tabelle 2-1 Information als Entität und Prozess nach Buckland (Buckland 1991), zur Darstellung siehe (Capurro 2015)

	Physikalisch nicht fassbar	Physikalisch fassbar
Entität	Information als Wissen Wissen	Information als Sache Daten, Dokumente, aufgezeichnetes Wissen
Prozess	Information als Prozess „informiert werden“	Informationsverarbeitung Datenverarbeitung, Dokumentenverarbeitung

Informationswissenschaft ist eine noch recht junge Disziplin, die 1958 mit Gründung des Institute of Information Scientists (IIS) in Großbritannien das Licht der Welt erblickte. Mit dem Begriff der Informationswissenschaft wollte man Wissenschaftler, die mit Information arbeiten von Laborwissenschaftlern abgrenzen – Informationswissenschaft versteht sich also interdisziplinär. Historisch bedingt fokussiert sich die Informationswissenschaft zunächst auf Information im Sinne von Daten sowie deren Verarbeitung; Dinge und Prozesse also, die physikalisch fassbar sind (Ingwersen 1992).

Das heutige informationswissenschaftliche Verständnis des Begriffs Information leitet sich ab aus den Bemühungen, den informationswissenschaftlichen Forschungsgegenstand im interdisziplinären Kontext zu definieren, als auch eine Abgrenzung zu anderen Disziplinen herzustellen, siehe Abbildung 2-1.

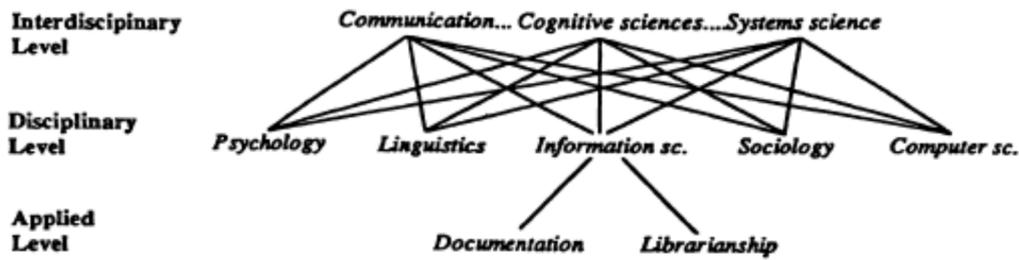


Fig. 1. Information science viewed as one of several sciences on information.

Abbildung 2-1 Informationswissenschaft im Kontext wissenschaftlicher Disziplinen aus (Ingwersen 1992, S. 103)

Belkins kognitiver Ansatz führt zu einem neuen Verständnis der Informationswissenschaft und damit zu einer erweiterten Sichtweise auf den Begriff der Information im Rahmen eines Kommunikationsprozesses.

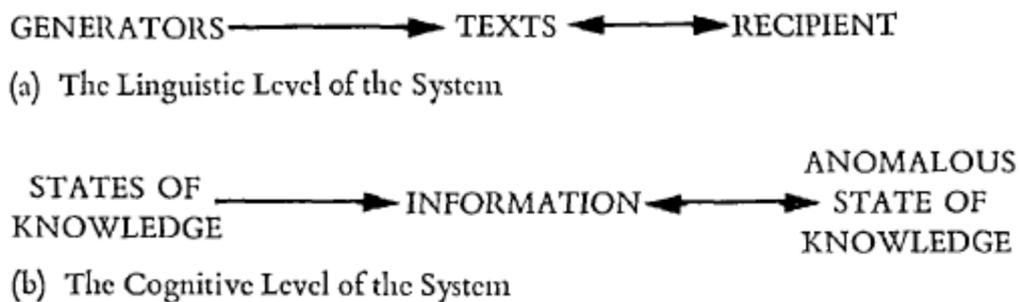


Abbildung 2-2 Belkins kognitives Modell der Information aus (Belkin 1978, S. 81)

Informationswissenschaft betrifft Information nach diesem Konzept mindestens (Belkin 1978):

- alle Information in menschlichen, kognitiven Kommunikationssystemen
- die Beziehung zwischen Menschen und Herstellern von Information (*generator*)
- das Verhältnis von Information und Nutzer
- Charakteristika der gewünschten Information
- Effektivität von Information, aber auch Effektivität von Informationsvermittlung (*information transfer*)

Belkin interpretiert Information damit multidimensional:

Der Erzeuger eines Dokuments macht dieses über ein Kommunikationssystem verfügbar und

definiert damit einen Wissensstatus zu einem spezifischen Thema. Information entspricht zu diesem Zeitpunkt der Repräsentation als Daten. Die so verfügbaren Informationen folgen einer gewissen Struktur, können aber in eine „kommunizierbare“ Struktur gebracht werden. Diese kommunizierbare Struktur ist sowohl abhängig von der Wissensrepräsentation im System als auch vom Wissensstatus potenzieller Nutzer. Die Anforderung von Information durch den Nutzer erfolgt auf Basis eines „anormalen“ Wissensstatus – ein Konzept, das im Folgenden detaillierter erläutert wird. Zum Abruf der Informationen ist eine Formalisierung der „Suchanfrage“ in eine kommunizierbare Form nötig, sie wird je nach Wissensstatus als Information oder Wissen wahrgenommen und kann zur Kompensation des anomalen Wissensstatus genutzt werden. (Belkin 1978)

Informationswissenschaftlich fallen alle diese Aspekte in das Forschungsgebiet der Informationswissenschaft. Untrennbar ist Information dabei beim Empfänger der Information mit Wissen verbunden – „Information ist Wissen in Aktion“ (Kuhlen 1995), sodass für die folgende Arbeit aus Gründen der Vereinfachung Wissen, Information, aber auch Daten teils synonym verwendet werden.

2.1.2 Warum suchen wir Information?

Belkins Definition führt zu einer weiteren Frage: Was bedeutet es, wenn wir davon sprechen, dass Menschen Informationen benötigen? Warum suchen Menschen Informationen?

Die Suche nach Information wird häufig ausgelöst durch den Wunsch nach oder den Bedarf an Information. In der Fachterminologie wird dieser Informationsbedarf englischsprachig als *information need* bezeichnet. Der folgende historische Abriss beleuchtet verschiedene Sichtweisen dieses Konzepts, das vor allem zu Beginn der Informationswissenschaft die erklärende Komponente des Umgangs von Menschen mit Information war und die Forschung jahrelang beherrschen sollte.

Mit seinen Grundlagenartikeln *The process of asking questions* und *Question negotiation and Information seeking in libraries* prägte Robert Taylor in den 1960er Jahren mit Konzept des *Information Need* als die Suche nach Antworten bei der Suche von Fachinformationen in Fachdatenbanken (Taylor 1962, 1968).

Zu dieser Zeit wurden Fachdatenbanken ausschließlich von Experten bedient. Verfügbare Datensätze wurden in langen Zeiträumen aktualisiert, lagen teilweise noch in Form von Magnetbändern vor (vgl. Abbildung 2-3) oder waren online ausschließlich über Telekommunikationsnetze verfügbar. Eine Literaturrecherche in Fachdatenbanken war daher immer eine Auftragsrecherche, die dazu führte, dass die eigentliche Forschungsfrage in einem vorbereitenden Gespräch zwischen Anfragendem und einem Informationsspezialisten zunächst genau definiert werden musste.

Damals war es üblich, sich mit seiner Fragestellung in einer Bibliothek an den Informationdesk zu wenden – Taylor nennt dies den *Mittelsmann*. Dieser Mittelsmann handelte mit dem Auftraggeber schrittweise zunächst den Informationsbedarf und schließlich die Formulierung der Suchanfrage aus. Die Aufgabe des Mittelsmanns war es demnach, den zu Beginn eventuell noch nicht formal fassbaren Informationsbedarf des Nutzers im Gespräch auszuarbeiten und schließlich in Form einer „suchbaren“ Fragestellung zu formalisieren. Der Mittelsmann erfüllte eine wichtige Aufgabe: er war der Schlüssel zur Information. Mehr noch: „*For most people, I am the information system*“ zitiert Taylor einen seiner Interviewpartner (Taylor 1968).



Abbildung 2-3 Eingabe von Datensätzen für die Speicherung auf Magnetbändern, Images from the History of Medicine (NLM), <https://ihm.nlm.nih.gov/images/A17414> (09.02.2016)

Wie dieser Prozess typischerweise abließ, wird bei Taylor 1968 anhand eines Stufenmodells eingehend beschrieben (Taylor 1968):

- Q1 bezeichnet die Stufe, in der ein Informationsbedarf erkannt wurde, aber noch nicht artikuliert wird oder artikuliert werden kann (*visceral need*)
- Q2 ist die Ebene, auf der ein Informationsbedarf geistig ausgearbeitet wird (*conscious need*)
- Q3 definiert das Level, in dem der Informationsbedarf formalisiert werden kann (*formalized need*)
- Q4 als oberste Ebene entspricht dem Informationsbedarf, der bereits als suchbare Fragestellung vorliegt (*compromised need*).

Die grundsätzliche Lesart des Taylor Modells ist linear, die Phasen Q1 bis Q4 werden also Schritt für Schritt durchlaufen:



Abbildung 2-4 Typologie des Informationsbedarfs nach Taylor (Taylor 1968)

Um diesen Prozess erfolgreich zu durchlaufen, müssen, so Taylor, Nutzer verschiedene „Filter“ passieren. Taylor legt mit diesen so genannten Filtern die Rahmenbedingungen fest, denen der Q1-Q4 Prozess unterliegt (Taylor 1968):

- die Definition des Themas
- Motivation und Intention des Nutzers
- persönliche Charakteristika des Nutzers
- Verhältnis der Datenbankstrukturen zur Nutzeranforderung
- abgelehnte und akzeptierte Antworten auf den Informationsbedarf

Taylors Modell trägt mit seiner allgemeinen Sichtweise wesentlich zum Grundverständnis des Konzepts bei und ist für die vorliegende Arbeit seines Fokus auf die professionelle Suche in Fachdatenbanken und der dabei nötigen Formalisierung des Informationsbedarfs auch heute noch interessant. Insbesondere aber legt das Modell den Grundstein für weitere Modelle und Konzepte wie dem kontextuellen Informationsverhalten, das im Folgenden eingehend betrachtet wird.

In seinem Modell des *anomalous state of knowledge* (ASK) greift Nikolas Belkin 1980 grundsätzliche Prinzipien des Taylor Modells als theoretische Grundlage für die Konzeption von Information Retrieval auf. Ein Informationsbedarf lässt sich nach der ASK-Theorie, erstmals definiert in (Belkin 1980), mit dem Erkennen eines anomalen Wissensstands erklären. Dieser ist am ehesten mit der Stufe Q1 des Taylor Modells zu assoziieren, definiert einen problematischen Aspekt. Die Suche nach Information dient der Reduktion von „Unsicherheit“ und kann entweder zu einem befriedigenden Wissensstand führen oder neue Unsicherheit aufzeigen. Das Auftreten eines *anomalous state of knowledge* sieht Belkin in situativen Unterschieden der mentalen Modelle auf System – und auf Nutzerseite und daraus resultierenden unterschiedlichen Wissensstatus, der vom Nutzer ausgeglichen werden muss. (Belkin 1980)

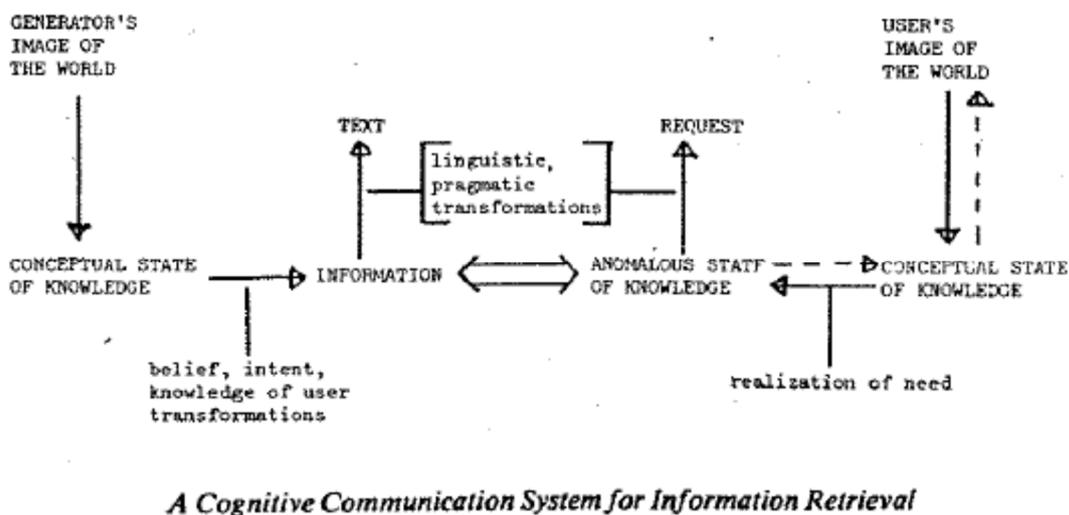


Abbildung 2-5 Belkins Modell des kognitiven Kommunikationssystems aus (Belkin 1980, S. 135)

Historisch entstammt das Konzept des *information needs* Nutzerbeobachtungen in Bibliotheken und findet mit dem ASK Modell in den 1980er Jahren direkt seine Anwendung in der Konzeption neuer Information Retrieval-Systeme, wie beispielsweise elektronischer Fachdatenbanken, vgl. hierzu beispielsweise (Belkin et al. 1982).

Einen vergleichbaren Ansatz verfolgt die *Sensemaking*-Theorie von Brenda Dervin. Ähnlich eines *anomalous state of knowledge* postuliert das Sensemaking-Modell *gaps* – Lücken oder Abgründe – die situativ auftreten und einen fortlaufenden Prozess zunächst blockieren. Um

zum gewünschten Ergebnis zu kommen, müssen diese Gaps überbrückt werden, idealerweise durch geeignete Information. (Dervin und Nilan 1986)

Die ausschließliche Fokussierung auf den *information need* und andere Konzepte, die einen Bedarf an Information erklären sollten, wurde bald kritisiert und als unzureichend beurteilt: Ein Informationsbedarf sei als Spezifikum einzelner Individuen schwer fassbar, zudem mündeten Versuche einer Definition des Information Need in der Praxis häufig in der Beobachtung eines Nutzerverhaltens (Case 2012, S. 87).

Dies wird beispielsweise bei Marcia Bates Berry Picking-Modell deutlich. Nach dem Berry Picking-Modell (siehe Abbildung 2-6) beginnen Nutzer die Suche häufig mit einem bekannten relevanten Dokument und durchkämmer bei der Informationssuche eine Vielzahl an Informationsquellen. Insbesondere definiert Bates den Suchprozess als dynamisch und geht nicht mehr von einem statischen, sondern vielmehr einem sich dynamisch verändernden Informationsbedarf aus.

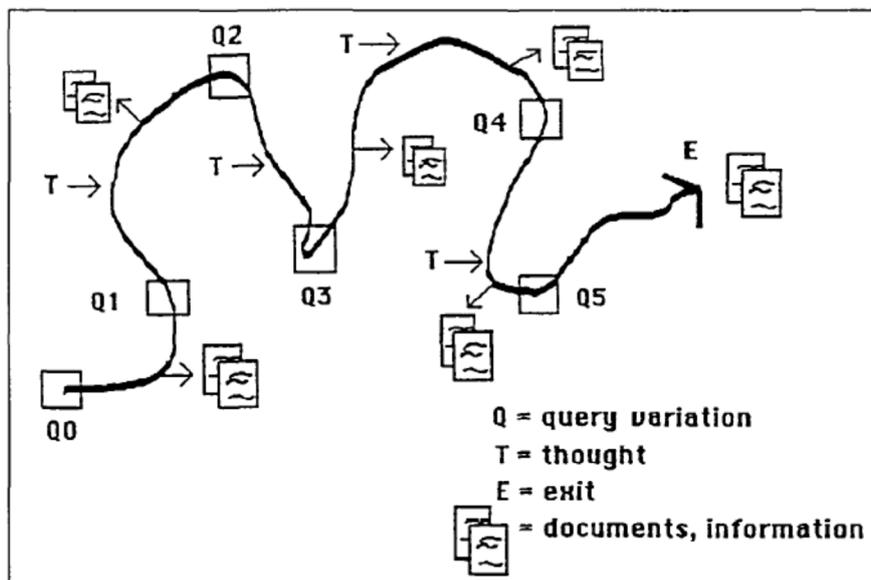


Abbildung 2-6 Bates Berry Picking Modell, aus (Bates 1989, S. 410)

Nach ihrem Modell führt jede neu aufgefundene Information zu neuen Ideen und möglichen Richtungen und demzufolge einer neuen Sichtweise auch auf den Informationsbedarf. So wird die Suchanfrage während des Suchprozesses verändert, um den sich gegebenenfalls auch

ändernden Informationsbedarf möglichst gut zu decken, Suchterme verändern daher ebenfalls sich kontinuierlich. Gleichzeitig kann der Nutzer mit jeder Suche wieder auf neue Information zurückgreifen. Die Suchanfrage wird also nicht in einem Schritt befriedigt, sondern ist ein Prozess aus vielen einzelnen Schritten, die iterativ den Suchprozess ausmachen. Suchanfragen führen damit nicht mehr zu einer einzelnen Treffermenge, sondern zu einem iterativen Prozess der Informationssammlung. Einzelne Teilschritte benennt Bates als „*berry picking*“ (Bates 1989).

Berry Picking hat viele Facetten (Bates 1989):

- beim **Footnote Chasing** folgt der Nutzer Fußnoten und Referenzlisten aus Büchern und Artikeln, führt also eine Rückwärts Zitationsuche durch
- **Citation Searching** bezeichnet eine Vorwärts Zitationsuche in einem Zitationsindex, der Nutzer sucht also nach Artikeln, die einen dem Nutzer bekannten Artikel zitieren
- beim **Journal Run** durchsucht der Nutzer Inhaltsverzeichnisse einzelner Ausgaben
- **Area Scanning** bezeichnet eine Suche innerhalb eines thematischen Fachgebiets
- **Subject Searching** ist eine dem Area Scanning verwandte Suchstrategie, die jedoch Thesauri nutzt
- **Author Searching** bezeichnet die Suche nach Veröffentlichungen bekannter Autoren.

Das Ende der 1980er Jahre von Bates postulierte Berry Picking-Modell (Bates 1989) sowie die daraufhin folgende Veröffentlichung *Where should the person stop and the information search interface start?* (Bates 1990) fordern die Entwicklung von Interfaces, die die Bedienung von Information Retrieval-Systemen bedienerfreundlicher machen sollten, indem sie dem natürlichen Suchverhalten von Benutzern entgegenkommen (Bates 1989):

“This model, it is argued, is much closer to the real behavior of information searchers than the traditional model of Information Retrieval is, and, consequently, will guide our thinking better in the design of effective interfaces.”

Aus informationswissenschaftlicher Sicht stellte man sich also zunehmend der Frage nach Informationsverhalten, d.h. der kontextspezifischen Interaktion mit Information, die kontextspezifische Nutzung von Information und letztendlich kontextspezifischer Strategien zur Informationsnutzung und Wissensgewinnung. Man folgte damit einem nutzerzentrierten im Vergleich zum bisher vorherrschenden systemorientierten Ansatz. Die Frage war also: Wie suchen und nutzen Menschen Information?

2.1.3 Vom Information Need zum Informationsverhalten

T.D. Wilsons Grundlagenmodell schafft ein Rahmenmodell des Untersuchungsgegenstands Informationsverhalten (Wilson 1981). Es definiert Informationsverhalten (eng. *information behaviour*) als Reaktion auf das Erkennen eines Informationsbedarfs und beleuchtet unterschiedliche Facetten dieses Phänomens (Abbildung 2-7).

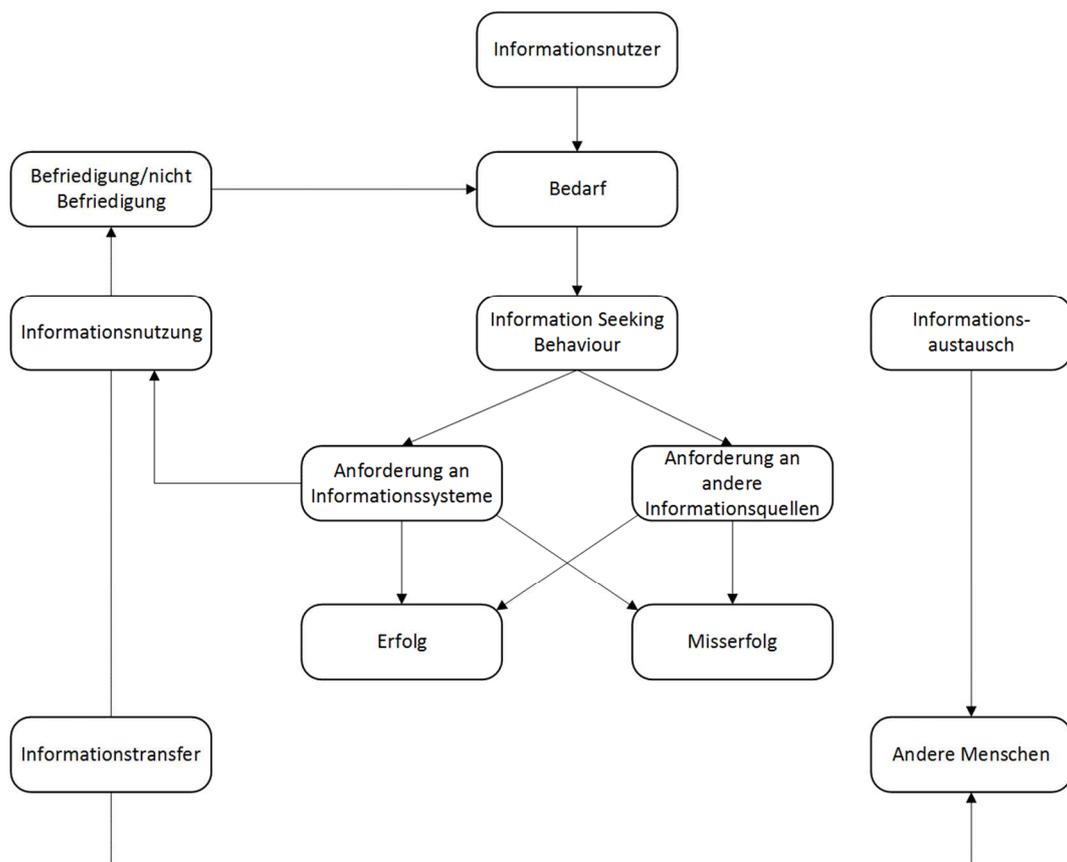


Abbildung 2-7 Modell zum Informationsverhalten nach Wilson (Wilson 1981, S. 4)

Das aus einem Bedarf resultierende Verhalten (*information seeking behaviour*) kann verschiedene Ausprägungen annehmen, beispielsweise in Form von Anfragen an Informationssysteme (darunter Bibliotheken ebenso wie Online-Systeme), aber auch in Form weniger formaler Informationsbeschaffungsprozesse (Anfragen bei Immobilienmaklern oder Autohändlern). Auch menschliche Kommunikation erklärt Wilson als - in diesem Fall bidirektionale - Informationsbeschaffungsmaßnahme (in Form eines Informationsaustauschs). Jede Informationsbeschaffung mündet in der Nutzung von Information, kann aber, so das Modell, auch scheitern und führt davon ausgehend gegebenenfalls zu einem iterativen Prozess.

Ein Merkmal des Wilson Modells ist die Herauslösung des Konzepts *information need* und eine Differenzierung psychologischer, kognitiver und affektiver Bedürfnisse, die alle kontextabhängig zu betrachten und dafür verantwortlich sind, dass sowohl individuelle Parameter des Nutzers als auch Umgebungsvariablen das Nutzerverhalten beeinflussen und erklären können (Wilson 1981).

Wie auch der Begriff Information ist auch der Begriff des Informationsverhaltens facettenreich, insbesondere aber lassen sich verschiedene Ebenen identifizieren, die als Klassifikationschema für das Konzept Informationsverhalten dienen. Die im Folgenden genannten Begriffe werden aufgrund ihrer Verwendung in der internationalen Fachliteratur englischsprachig verwendet.

Information behaviour umfasst alles menschliche Verhalten im Umgang mit Informationsquellen und Informationskanälen, aktive Nutzung oder passive Rezeption. Informationsverhalten beschreibt demnach sowohl beispielsweise *face to face* Kommunikation ebenso wie Radiohören oder Fernsehen. Aus Informationsverhalten lässt sich keine Intention ableiten, auf so gewonnene Information aktiv zu reagieren.

Information seeking behaviour ist Suchverhalten, das zielgerichtet einen spezifischen Bedarf decken soll. Dies umfasst beispielsweise die Nutzung einer Zeitung, einer Bibliothek, einer Datenbank oder ganz allgemein des Internets.

Information searching behaviour definiert ein Mikro-Level des Verhaltens und beschreibt die Interaktion des Suchenden mit Informationssystemen aller Art. Dazu gehören zum einen rein physikalische Interaktionen wie der Mausclick, beispielsweise bei Nutzung einer Software, eines Systems oder das Folgen eines Links im Internet. Zudem umfasst *information searching*

behaviour aber auch intellektuelle Tätigkeiten wie die Formulierung einer Booleschen Suchanfrage in Fachdatenbanken oder Relevanzbeurteilungen aufgefundener Information.

Information use behaviour umfasst alles Verhalten, das nötig ist, um aufgefundene Information im Kontext aktuellen Wissens bewerten oder vergleichen zu können. (Wilson 2000)

In den vergangenen Jahrzehnten wurden Dutzende konkreter Modelle zur Beschreibung von Informationsverhalten veröffentlicht. Modelle legen ihren Fokus im Gegensatz zu Theorien meist auf spezifische Aspekte, sie sind die vereinfachte Repräsentation eines Phänomens oder Sachverhalts, repräsentieren Variablen und deren Assoziationen und werden häufig grafisch dargestellt. Zu den wesentlichen Unterscheidungskriterien einzelner Modelle gehören ihre Struktur und Darstellungsweise, ihre Evidenz aber auch ihr Abstraktions- oder Detailgrad (Ingwersen und Järvelin 2005, S. 15).

Die im Folgenden vorgestellten Modelle unterscheiden sich zudem unter anderem sowohl hinsichtlich der beobachteten Nutzergruppe als auch hinsichtlich der Qualität und des Umfangs der empirischen Methodik.

2.2 Information Seeking Behaviour

Vor allem auf Ebene des *information seeking behaviour* wurden eine Vielzahl an Modellen vorgestellt. Sie fokussieren sich nicht ausschließlich auf Suchprozesse im Sinne des *information searching behaviour*, sondern betrachten Nutzung von Information und Interaktion mit Information aus einem allgemeinen Blickwinkel. Information wird im *information seeking* Modell sowohl in seiner Form als Ding, strukturelle Wissensrepräsentation als auch Prozess wahrgenommen und entsprechend im Kontext unterschiedlichster Informationskanäle untersucht, vgl. hierzu 2.1.1. *Information seeking behaviour* Modelle können demnach als Grundlage auch für die Entwicklung von Information Retrieval-Systemen verstanden werden, bedienen aufgrund ihrer allgemeinen, übergeordneten Perspektive hierbei jedoch ausschließlich Designprinzipien und nicht die Interaktionsebene. (Wilson 1999)

Die im Folgenden weitestgehend in chronologischer Reihenfolge vorgestellten Modelle sind entweder für das Verständnis des Konzepts *information seeking behaviour* von grundlegender Bedeutung oder haben thematisch direkten Bezug zur vorliegenden Arbeit, indem sie sich empirisch im Kontext der professionellen Suche verorten lassen. Sie alle zeigen verschiedene

Sichtweisen auf das Konzept des Informationsverhaltens und reichen von statischen, sehr allgemeinen Modellen bis hin zu Aktivitätsmodellen mit konkreten Szenarien.

2.2.1 Modell von Ellis

Ellis entwickelte in zu Beginn der 1990er Jahren ein Verhaltensmodell der professionellen Suche von Domänenexperten in Information Retrieval-Systemen, das in verschiedene Disziplinen, beispielsweise in den Naturwissenschaften, bei Ingenieuren und industriell tätigen Wissenschaftlern (Ellis 1989), entwickelt, getestet und später anhand weiterer Studien inhaltlich und konzeptionell ergänzt wurde, vgl. (Ellis 1993; Ellis et al. 1993; Ellis und Haugan 1997). Ziel des empirisch abgeleiteten Modells war es, das Informationsverhalten der Probanden aus ihrer Wahrnehmung zu beschreiben (Ellis 1989).

Gegenüber vielen anderen Modellen verzichtet Ellis auf die typische graphische Darstellung und beschreibt verschiedene Charakteristika und Aktivitätsmuster während der Suche wie folgt. Die folgenden Definitionen und Erläuterungen sind, sofern nicht anders angegeben der Originalarbeit von Ellis (Ellis 1989) entnommen und erläutert:

Starting: die initiale Suche nach Informationen, um potenziell relevante Informationsquellen zu identifizieren

Starting subsumiert in Ellis Modell alle Muster der Informationssuche, die dem Ziel dienen, sich in thematisch in neue Themengebiete einzuarbeiten. Individuelle Faktoren wie Kenntnisstand und Erfahrung bedingen Starting-Prozesse, beispielsweise die Nutzung von Schlüsselreferenzen, Übersichtsarbeiten oder Sekundärquellen. Die Schlüsselreferenz als Dokument erfüllt in Starting Aktivitäten daher zwei Rollen: Informationsquelle zum schnellen, thematischen Einstieg, also auch Quelle für weitere Dokumente und damit dem Einstieg in Chaining-Aktivitäten. Übersichtsarbeiten verschaffen Nutzern zusätzlich einen umfassenden Überblick. (Ellis 1989)

Chaining: Rückwärts- oder Vorwärtssuche in Zitationen (Zitationsuche) und anderen Informationsquellen, die in irgendeiner Form miteinander verknüpft sind

In der Rückwärtssuche folgt der Nutzer bibliographischen Referenzen aus Originalarbeiten, eine Methode, die im Kontext der Sozialwissenschaften zu den wichtigsten Methoden der

Informationsgewinnung aus bekannten Dokumenten gilt und in der Literatur vielfach beschrieben wird (Line 1971; Skelton 1973). In der Vorwärtssuche werden aus vorliegenden relevanten Artikeln oder Zitationen über spezialisierten Fachdatenbanken wie Web of Science (Thomson Reuters 2016c) (so genannte Zitationsdatenbanken) alle neueren Veröffentlichungen ermittelt, die einen Referenzartikel zitieren. Die Idee zur Zitationssuche stammt bereits aus den 1960er Jahren (Garfield 1964), wurde durch ihre Implementierung in Fachdatenbanken und Datenbanksystemen aber erst später populär. Dieses Konzept findet sich unter anderem in Bates Berry Picking Strategien, siehe Abschnitt. 2.1.2.

Browsing: semi-direktionale Suche in Quellen potenziellen Interesses, beispielsweise das Überfliegen von Inhaltsverzeichnissen oder Überschriften

Das Konzept Browsing ist etymologisch auf das englische Wort *browse* zurückzuführen, das historisch das Abgrasen eines Geländes durch Weidetiere beschreibt (Oxford University Press 2016e, Stichwort: browsing). Informationswissenschaftlich werden unter Browsing heute vorwiegend informelle oder unvorhergesehene Verhaltensmuster im Umgang mit Information zusammengefasst (Case 2012, S. 100).

Browsing kann beschrieben werden als (Bates 2007):

"[...] the activity of engaging in a series of glimpses, each of which may or may not lead to closer examination of a (physical or represented) object, which examination may or may not lead to (physical and/or conceptual) acquisition of the object."

eine aufeinanderfolgende Serie flüchtiger Blicke, die nicht notwendigerweise zu einer näheren Untersuchung eines physikalischen repräsentierten Objekts führen, aber im Folgenden zum physikalischen oder konzeptuellen Erwerb dieses Objekts führen können.

Im Sinne des Ellis Modells wird Browsing als semi-gerichtetes, semi-strukturiertes Verhalten definiert und beispielsweise in Form eines „Durchsehens“ von Inhaltsverzeichnissen oder Bücherregalen beobachtet (Ellis 1989).

Differentiating: Zugriff auf evtl. auch gefilterte Informationen mit dem Ziel des Vergleichs oder der Bewertung

Das multidimensionale Konzept Differentiating umfasst im Wesentlichen Vergleichs- und Bewertungsmaßnahmen durch den Suchenden mit dem Ziel inhaltlicher oder qualitativer Validierung von Information. (Ellis 1989).

Monitoring: über neueste Entwicklungen eines Forschungsgebiets durch Beobachtung von Informationsquellen oder Informationsitems „auf dem Laufenden bleiben“

Die von Ellis identifizierten Monitoring Aktivitäten sind vielschichtig: Sie umfassen nicht nur formaler Informationsbeschaffungsprozesse wie die regelmäßige Nutzung von Informationsquellen wie Zeitschriften oder Bücher, sondern auch Informationsaustausch im Sinne informeller Tätigkeiten, beispielsweise Kollegenkontakt oder Netzwerke. (Ellis 1989).

Extracting: systematische Überprüfung von Informationsquellen mit dem Ziel, dort relevantes Material zu extrahieren

Verifying: Kontrolle und Validierung von Information

Ending: ergänzende Suche am Ende des Projekts, beispielsweise eine finale Literatursuche

Ellis lässt sowohl Beziehungen als auch Übergänge zwischen einzelnen Aktivitäten offen und betont, dass diese von den jeweiligen Umständen abhängen (Ellis 1989):

“the detailed interrelation or interaction of the features in any individual information seeking pattern will depend on the unique circumstances of the information seeking activities of the person concerned at that particular point in time. The relationships between the features of the model can, therefore, only be indicated in the most abstract and general terms unless there is reference to a particular information seeking pattern”

Erklärungsversuche, wie die von Ellis identifizierten Muster in eine logische und zeitliche Abfolge zu bringen wären, folgten später: Starting, Chaining und Browsing können als Aktivitäten interpretiert werden, die vorwiegend in die frühe Phase der Informationssuche fallen, vgl. (Wilson 1999; Vakkari 2003), Extracting und Verifying beenden den Suchprozess.

Das Ellis Modell wurde später entsprechend aufbereitet und von Wilson grafisch dargestellt, indem er mögliche Aktivitätsketten modelliert (Wilson 1999):

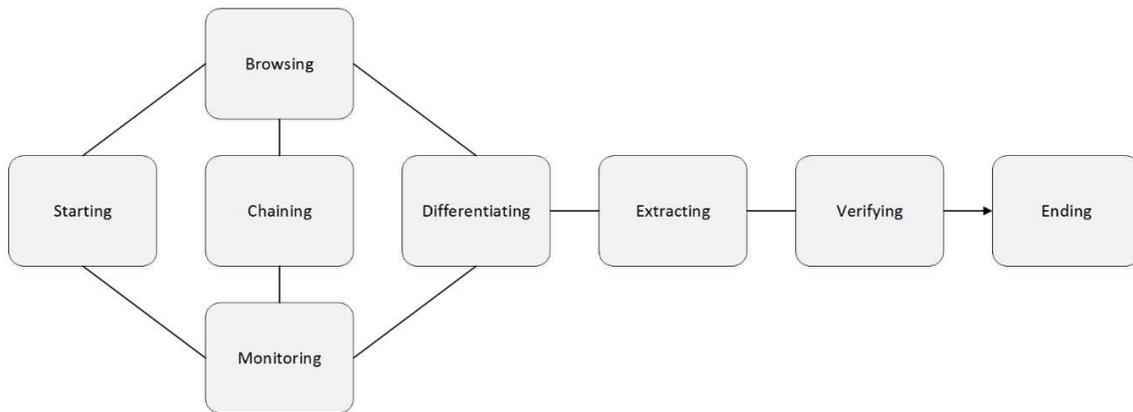


Abbildung 2-8 Das Ellis Modell in der Darstellung nach Wilson (Wilson 1999)

Die Stärken des Ellis Modells liegen in seiner Empirie, der vielfachen praktischen Anwendung und seiner Anwendbarkeit in vielen experimentellen Settings. Die Schwäche des Modells liegt in seiner statischen und sehr allgemeinen Beschreibung sowie den fehlenden kausalen Faktoren und Triggern für einzelne Aktivitäten (vgl. hierzu (Ingwersen und Järvelin 2005; Case 2012)).

Zusammenfassend liegt mit dem Ellis Modell also ein empirisch entwickeltes Modell vor, das mehrfach in verschiedenen Disziplinen validiert und erweitert werden konnte, sodass wir daraus heute ein grundsätzliches Bild auf Aspekte des *information seeking behaviour* von Domänenexperten bei der Nutzung von Information Retrieval-Systemen für Forschungsprojekte ableiten können.

Das Modell von Meho und Tibbo (2003) ist eine der Forschungsarbeiten, welche die Vorarbeiten von Ellis auf Basis einer umfassenden Studie mit einem interdisziplinären Kollektiv an Testpersonen empirisch bestätigen, es kann damit als Weiterentwicklung gelten (Meho und Tibbo 2003). Die von Ellis identifizierten Suchaktivitäten werden im Meho und Tibbo-Modell differenziert in den Kontext der Suche nach Information (*searching*) sowie der Verarbeitung von Information (*processing*) gestellt, können jedoch hier und da verortet werden, siehe Abbildung 2-9. Die Suche ist definiert als Prozess, in dem relevantes und potenziell relevantes Material identifiziert wird und umfasst sowohl verschiedene Methoden der Informationsbeschaffung als auch den Umgang mit verschiedenen Informationsquellen.

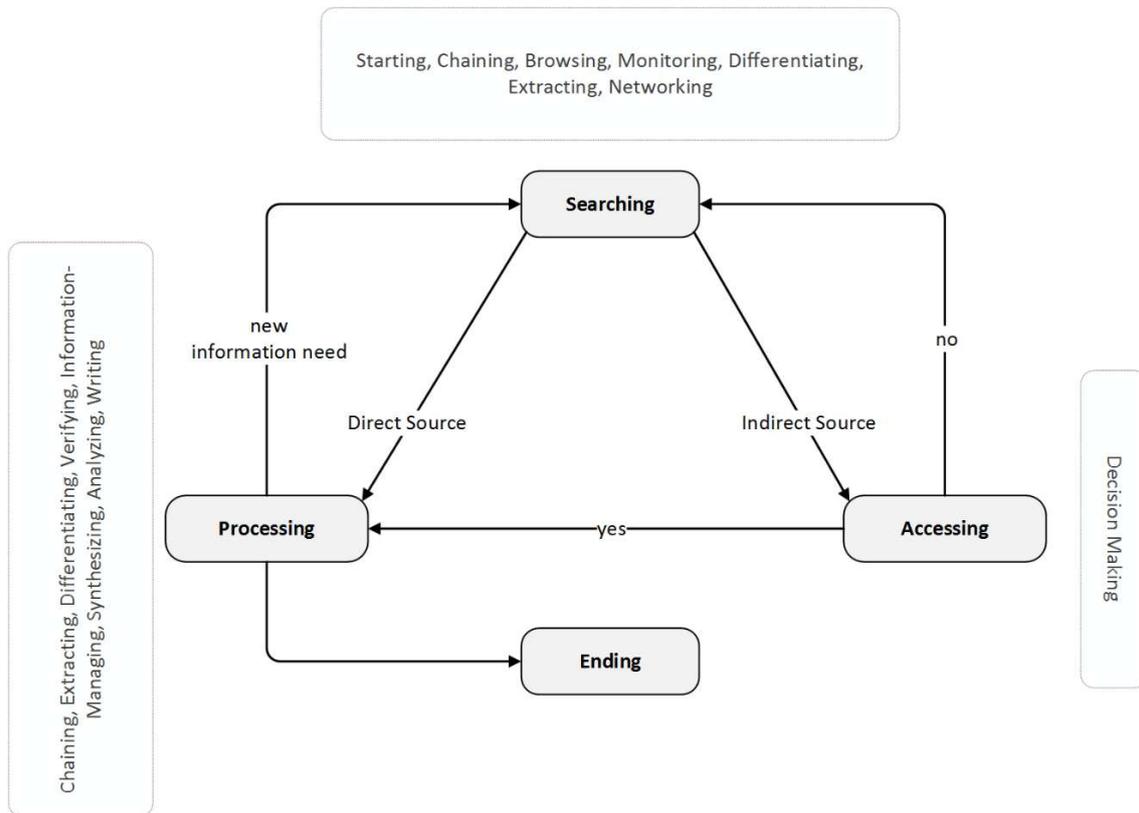


Abbildung 2-9 Modell nach Meho und Tibbo (Meho und Tibbo 2003)

Das Meho und Tibbo-Modell umfasst ergänzend zum Ellis-Modell weitere Ausprägungen der Suche und Verarbeitung von Information, siehe für die folgenden Definitionen die Originalarbeit von (Meho und Tibbo 2003):

Networking beschreibt kommunikative Aspekte der Informationsbeschaffung, beispielsweise den Kontakt zu Kollegen und kann konkret auf E-Mail-Kontakte und Internet Netzwerke ausgedehnt werden. Vor allem die Verfügbarkeit und Aktualität machen diese Form der informellen Informationsbeschaffung für Wissenschaftler attraktiv.

Information Managing umfasst alle Prozesse rund um das Management von Information, beispielsweise die Speicherung von Referenzen in Literaturverwaltungsprogrammen.

Während der Verarbeitungsphase werden Informationen analysiert und synthetisiert, sie führt letztendlich zu einem finalen Ergebnis, der Nutzung aufgefundener Information (*ending*).

Die Beschaffungsphase (*accessing*) definiert als Brücke zwischen Suche und Verarbeitung im Falle der Nutzung indirekter Informationsquellen wie bibliographischer Fachdatenbanken den notwendigen und nicht immer erfolgreichen Zwischenschritt der Informationsbeschaffung. Meho und Tibbo begründen die Notwendigkeit dieses Schritts in der Vielfalt verfügbarer Informationsquellen. Die für die Suchphase charakteristischen Aktivitäten Starting, Chaining, Browsing, Monitoring, Extracting sowie Networking nutzen zur Informationssuche häufig indirekte Quellen in Form bibliographischer Datenbanken, persönlicher Kontakte, Empfehlungen von Kollegen oder Zitationssuchen (Meho und Tibbo 2003).

Das Meho Tibbo-Modell stellt also ergänzend zum Ellis-Modell Informationsverhalten als Prozess dar. In diesem können Suche, Beschaffung und Verarbeitung von Information durch die Verfügbarkeit von Informationsquellen sowie neu getriggertem Informationsbedarf (als charakteristisches Merkmal des Prozesses) iterativ zu notwendigen Entscheidungen führen.

2.2.2 Modell von Kuhlthau

Das Konzept der Unsicherheit des ASK Modells (siehe Abschnitt 2.1.2) findet sich in den Untersuchungen von Carol Kuhlthau (Kuhlthau 1988), die letztendlich in einem fundierten und für die Informationswissenschaft wegweisenden Modell des Informationsverhaltens (Kuhlthau 1991, 1993a, 1993b) mit einer kognitiven Sichtweise auf das *information seeking behaviour* münden.

Das Kuhlthau Modell wurde empirisch in Untersuchungen mit studentischen Testpersonen entwickelt und scheint daher im Kontext der vorliegenden Arbeit weniger relevant. Als Grundlagenmodell prägte es jedoch die Forschung um das Information Seeking Behaviour nachhaltig und ist daher für das Verständnis der weiteren Arbeit hilfreich, sodass es hier der Vollständigkeit halber genannt werden soll.

Kuhlthau beschreibt den *information search process*, ein anerkannt universelles und allgemeingültiges Modell zum Informationsverhalten, vgl. (Case 2012, S. 145; Ingwersen und Järvelin 2005, S. 65), indem einzelne Stufen des Prozesses mit kognitiven Aspekten, Aktionen und möglichen Aufgaben assoziiert werden.

	Initiation	Selection	Exploration	Formulation	Collection	Presentation
Feelings	Uncertainty	Optimism	Confusion Frustration Doubt	Clarity	Sense of Direction/ Confidence	Satisfaction or Disappointment
Thoughts	vague			clearer	Increased interest	focused
Actions	Seeking background information		Seeking relevant information		Seeking pertinent information	
Appropriate Tasks	Recognize	Identify, investigate	Identify, investigate	Formulate	Gather	Complete

Abbildung 2-10 Kuhlthaus Information Search Process-Modell nach (Kuhlthau 1991)

Die Stadien des Information Search Process beschreibt Kuhlthau wie folgt (Kuhlthau 1991):

Initiation ist das Stadium, in dem Nutzern der Informationsbedarf bewusst und ein Problem erkannt werden. Es ist assoziiert mit Unsicherheit, ein Konzept, das sich auch im ASK Modell findet.

In der **Selection** Phase des Prozesses werden allgemeine Informationen identifiziert, ausgewählt und untersucht.

Exploration beschreibt die Phase des Suchprozesses, in der nach allgemeiner, relevanter Information gesucht wird, sie kann sowohl in Verwirrung, Frustration als auch Zweifel münden und lässt sich eindeutig von der nächsten Stufe des Prozesses abgrenzen.

Formulation benennt das Stadium der Suche, in dem die Fragestellung fokussiert formuliert und präzisiert wird, sodass der Informationsbedarf letztendlich konkrete Strukturen annimmt.

Collection umfasst die daraufhin folgenden Such- und Sammelprozesse, die sich eindeutig auf das fokussierte Problem beziehen.

Presentation beschließt die Suche und beendet sie mit der Nutzung aufgefundener Information.

Das Kuhlthau Modell wurde zunächst im studentischen Umfeld entwickelt, später aber in weiteren Arbeiten auf anderen Domänen übertragen (Kuhlthau 1997, 1999) und als Grundlage für weitere *information seeking*-Modelle verwendet, beispielsweise (Vakkari 2001): Vakkaris Theorie „übersetzt“ die bei Kuhlthau definierten Stadien des Informationsprozesses in Pre-Focus, Formulation und Post-Focus, bei denen im Gegensatz zu Kuhlthaus Modell nicht die Nützlichkeit (*usefulness*), sondern die Relevanz von Information das Entscheidungskriterium ist, siehe hierzu auch Abschnitt 2.5.3.

2.2.3 Marchioninis ISP-Modell

Marchioninis achtstufiges ISP-Modell (Marchionini 1995) liefert eine theoretische Beschreibung des Informationsverhaltens als *information seeking process* (ISP) und modelliert Informationsverhalten wie die bisherigen Modelle ebenfalls als Prozess. Das ISP-Modell beschreibt jedoch nicht nur einzelne Stadien der Suche, sondern auch konditionale Übergänge zwischen diesen Stadien.

Die folgende Erläuterungen des ISP-Modells sind sofern nicht anders angegeben der Originalquelle (Marchionini 1995), S. 49ff. entnommen.

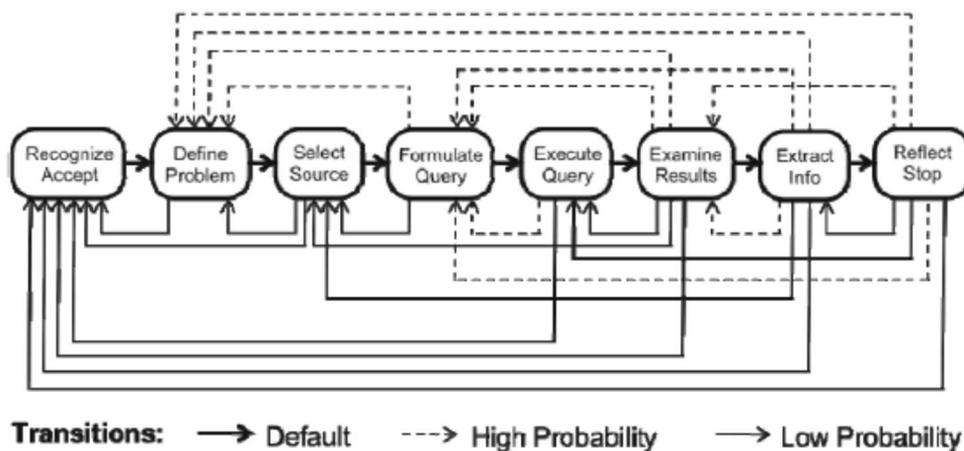


Abbildung 2-11 Information Seeking Process Marchionini aus (Marchionini 1995, S. 49)

Mit dem Erkennen und Akzeptieren eines Informationsbedarfs (**Recognize Accept**) greift das Modell Konzepte wie Taylors *visceral need* oder Belkins ASK (siehe 2.1.2) auf. Das Erkennen eines Informationsbedarfs führt nicht notwendigerweise direkt in einen Suchprozess, sondern

kann, beispielsweise aufgrund von Kosten- Nutzen-Entscheidungen, prinzipiell auch unterdrückt werden. Erst dessen Akzeptanz definiert nach dem ISP-Modell den Zweck und die Instanziierung von Suchprozessen und ist sowohl von thematisch konzeptionellem Wissen, Wissen um verfügbare Informationssysteme als auch von infrastrukturellen Rahmenbedingungen beeinflusst.

Das Stadium der Problemdefinition (**Define Problem**) führt als kognitive Aufgabe letztendlich zur konkreten Formulierung einer Suchaufgabe sowie zur Entwicklung von Suchstrategien und kann direkt mit Taylors *conscious need* assoziiert werden. Da es aus anderen Phasen des ISP getriggert werden kann, ist es im Modell über den gesamten Suchprozess präsent. Erklärt wird das Verständnis um die Suchaufgabe und damit Problemdefinition vorwiegend über Domänenwissen, - gleichzeitig ist dies auch eine erklärende Variable für das Scheitern von Suchprozessen.

Die Auswahl von Suchquellen (**Select Source**) ist wesentlich beeinflusst von persönlicher Erfahrung, Domänenwissen, als auch infrastrukturellen Rahmenbedingungen – Domänenexperten sind häufig spontan in der Lage, gezielt zu beurteilen, wo die erforderlichen Informationen aufzufinden sind, während der naive Suchende häufig auf Standardquellen zurückgreift. In der professionellen Suche führen neue Themengebiete dagegen zur Notwendigkeit, aufgrund von Hintergrundinformationen über potenzielle Informationsquellen die Auswahl geeigneter Ressourcen einzugrenzen.

Mit der Formulierung der Suchanfrage (**Formulate Query**) als Übergang zum *compromised need* und schließlich *formalized need*, der Formalisierung des Informationsbedarfs, erfolgt die erste Interaktion mit Informationsquellen, sie ist demnach abhängig von Systemfunktionalitäten und geprägt von Routine im Umgang mit Quellenspezifika.

Die Ausführung der Suche (**Execute Search**) beschreibt entgegen der Formulierung von Suchanfragen physikalische Aktionen zum Zugriff auf Information. Unser Verständnis um diese Aktivitäten im Rahmen der Informationsbeschaffung ist vielfältig und hat in den letzten Jahren sicher die größte Veränderung erfahren. Heute drücken wir bei der Nutzung von Suchmaschinen und Datenbanken nach Eingabe von Suchbegriffen schlicht die Return-Taste, um Treffermengen zu erhalten, folgen in einem Hypertextsystem, einem Link oder öffnen eine E-Mail.

Im Alltag interpretieren wird dies heute sicher als prototypische Form der Informationsgewinnung. Eine Ausführung der Suche kann aber auch durch das Entnehmen eines Buches aus einem Bücherregal oder dem Griff zum Telefonhörer erfolgen.

Die Untersuchung von Treffermengen (**Examine Results**) ist die direkte Konsequenz aus der Ausführung der Suche und der Antwort eines Informationssystems, ohne die eine Beurteilung des Erfolgs der Suche nicht möglich ist. Die Untersuchung von Treffermengen im Sinne einer Relevanzbeurteilung im Hinblick auf den aktuellen Suchprozess kann aus Perspektive des Suchenden als Entscheidung für die nächsten Schritte im ISP Prozess interpretiert werden (Marchionini 1995):

„From an information seeker's perspective, relevance may be considered as a decision on what action to take next in the information-seeking process.“

Das in der Informationswissenschaft wichtige Konzept der Relevanz, detailliert erläutert in Abschnitt 2.3.1.2, wird hier im Sinne einer situativen Relevanz (Wilson 1973) verstanden: Situativ relevante Information dient der Beantwortung von situativ zu betrachtenden Fragen und ist daher sowohl abhängig von spezifischen Informationsproblemen als auch veränderlich. Als Entscheidungsparameter führt Relevanz im ISP-Prozess zu einer Vielzahl möglicher Teilprozesse. Neben dem Abbruch des Suchprozesses durch erfolgreichen Abschluss der Suche sind Entscheidungen wie eine eingehende Untersuchung von Informationen ebenso denkbar wie ein Wechsel der Informationsquelle, die Reformulierung von Suchanfragen oder eine Neudefinition des Informationsproblems an sich. Das ISP-Modell zeigt daher Übergänge zu allen Stufen im Gesamtprozess.

Das ISP-Modell verknüpft den Prozess der Relevanzbeurteilung direkt mit dem Exzerpieren von Information (**Extract Information**). Relevante Information kann dabei partiell oder vollständig einem Informationsbedarf genügen und führt im ISP-Modell zu alternativen Handlungsoptionen: Eine positive Relevanzbeurteilung kann direkt in der Auswahl und Nutzung von Information münden, jedoch auch in temporärer Speicherung und Wiedervorlage von Dokumenten. In diesem Fall können iterative Relevanzbeurteilungen anhand inzwischen neu gewonnener Dokumente situativ zu vollkommen anderen Entscheidungen führen.

Die Auswahl von Information erfolgt durch Modi wie Lesen, Scannen, Klassifizieren oder auch

Speicherung von Information. Neue Information wird durch Integration in das domänenspezifische Wissen in Form rein physikalischer Aktivitäten wie Speichern, Kopieren und Einfügen nutzbar gemacht. Dieser Prozess ist ein gutes Beispiel für situative Relevanz: Neue Information kann in einem früheren Stadium des iterativen ISP Prozesses relevanter eingestuft werden als zu einem Zeitpunkt, zu dem bereits mehr Information zur Verfügung steht – die Relevanz von Information ist immer auch im Kontext des bereits verfügbaren Wissens zu interpretieren.

Der ISP Prozess endet mit der Reflektierungsphase (**Reflect Stop**), die zu iterativen Loops wie auch zur Beendigung des Prozesses führen kann.

Das ISP-Modell ist auch im Kontext der professionellen Informationsbeschaffung beschrieben. Beispielsweise findet es sich in adaptierter Form in der Patentrecherche wieder (Jürgens et al. 2014) und identifiziert dort spezifische Übergänge zwischen einzelnen Aktivitäten unter den in der Patentrecherche üblichen Aufgaben sowie geltender Rahmenbedingungen.

2.2.4 Modell nach Byström und Järvelin

Das Modell von Byström und Järvelin (Byström und Järvelin 1995) beschreibt Informationsverhalten als Problemlösungsstrategien unterschiedlich komplexer Suchaufgaben.

Die Abhängigkeit von Informationsverhalten, Informationssuchprozessen und Arbeitsaufgabe ist in der Fachliteratur seit Jahrzehnten akzeptiert, vgl. hierzu (Wersig 1973; Belkin et al. 1982): Aufgaben erfordern verfügbare Information, um erfolgreich erledigt werden zu können. Aus dieser Notwendigkeit kann sich demnach ein Informationsbedarf ergeben, der ein spezifisches, von verschiedenen Faktoren abhängiges Informationsverhalten auslöst (Byström und Järvelin 1995).

Das Modell geht von einer fünfstufigen Skala zur a-priori Klassifikation von Aufgaben aus. Bei wenig komplexen (einfachen) Aufgaben ist es möglich sowohl das Ergebnis, die erforderlichen Informationsquellen als auch den Prozess selbst vorab zu beschreiben. Bei hochkomplexen Aufgaben sind weder Ergebnis, noch die erforderlichen Informationsquellen oder der letztendliche Prozessablauf von vornherein vorherzusagen (Byström und Järvelin 1995), siehe Abbildung 2-12 und die folgende Beschreibung.

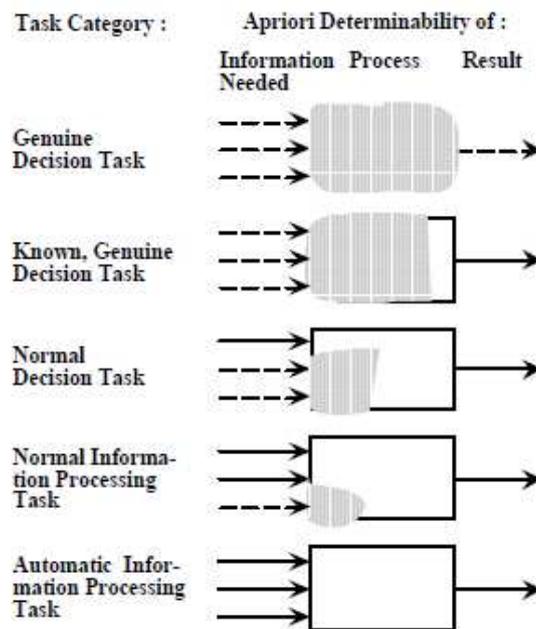


Abbildung 2-12 Klassifizierung der Komplexität von Aufgaben, aus (Byström und Järvelin 1995, S. 6)

Automatische Informationsverarbeitungsaufgaben sind soweit vorhersehbar, dass sie „im Prinzip automatisiert werden könnten“.

Normale Informationsverarbeitungsaufgaben lassen sich im Allgemeinen vorsehen, allerdings mit gelegentlichen Randbedingungen, die von Fall zu Fall abgeglichen werden können, sodass weder der Gesamtprozess und Informationsbedarf immer vorhersehbar sind.

Normale Entscheidungsaufgaben weisen zwar deutliche Strukturen auf, sind jedoch sehr individuell.

Bekannte, authentische Entscheidungsaufgaben sind solche, deren Ergebnis zwar vorab bekannt ist, die aber bisher keine klaren Abläufe kennen. Dies macht diesen Aufgabentyp sowohl in seiner Prozessierung als auch im Informationsbedarf unvorhersehbar.

Authentische Entscheidungsaufgaben sind definiert als unerwartete Aufgaben, die entsprechend unstrukturiert sind. Daher sind weder Ergebnis, noch Prozess, noch Informationsbedarf absehbar.

Das Modell klassifiziert zudem drei das aufgabenspezifische Informationsverhalten beeinflussende Informationstypen (Byström und Järvelin 1995):

Problemspezifische Information beschreibt Strukturen, Eigenschaften und Anforderungen, die ein Problem von Natur aus an den Menschen stellt. Beim Brückenbau gehören beispielsweise Standort und Einsatzort zu problemspezifischen Informationen.

Domänenspezifische Information umfasst Fakten, Konzepte, Gesetze und Theorien im Rahmen der Problemdomäne. In oben genanntem Beispiel gehört beispielsweise die Kenntnis über Brückenkonstruktionen in diese Kategorie. Domänenspezifische Information findet sich typischerweise in Fachliteratur.

Problemlösungsinformationen sind Informationen, die direkt mit der Problemlösung assoziiert sind, beispielsweise die Beurteilung verschiedener denkbarer Brückenkonstruktionen für den Stand- und Einsatzort. Diese Informationen sind häufig über Experten verfügbar.

Das Informationsverhalten beginnt nach dem Byström und Järvelin-Modell mit der individuellen Wahrnehmung der übertragenen Aufgabe, weiteres Verhalten ist assoziiert mit persönlichen Faktoren sowie situativen Rahmenbedingungen, die sowohl die Art der aufzusuchenden Informationen beeinflussen, als auch die Art und Weise, wie diese Informationen aufgesucht werden. Suchprozesse münden als Feedback-Loop in einer Evaluation der Informationssuche, die entweder scheitern kann, erledigt ist oder in weiterem Informationsbedarf endet (siehe Abbildung 2-13).

Aus der Studie an 14 Staatsbediensteten (Byström und Järvelin 1995) kann abgeleitet werden, dass

- die Komplexität von Aufgaben positiv mit der Komplexität des Informationsbedarfs korreliert,
- die Komplexität von Aufgaben positiv mit dem Bedarf an domänenspezifischer Information als auch Information zu Problemlösungsstrategien korreliert,
- die Komplexität von Aufgaben positiv mit der Nutzung allgemeiner Informationsquellen und negativ mit der Nutzung spezifischer Informationsquellen korreliert,
- die Erfolgsrate von Suchstrategien sinkt und

- die Anzahl benötigter Informationsquellen steigt.

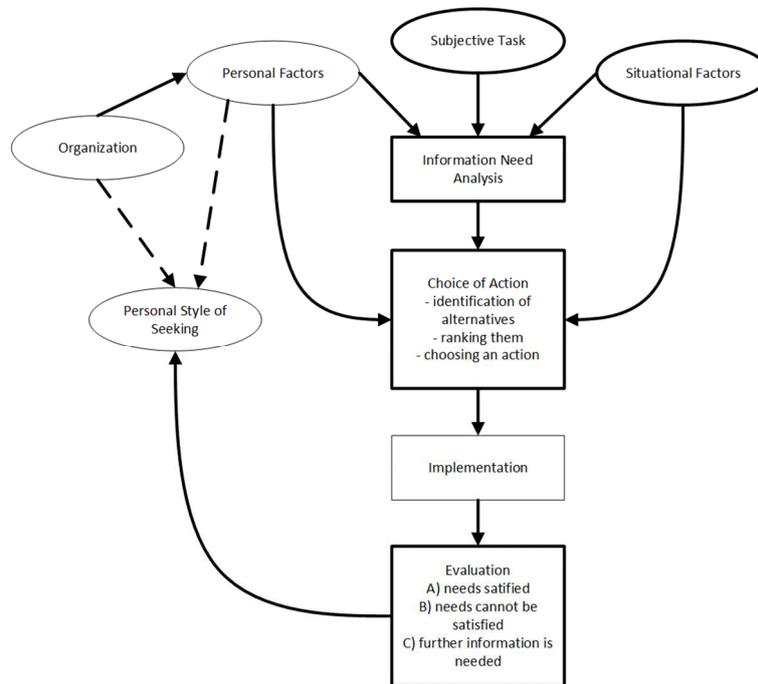


Abbildung 2-13 Byström Järvelin Modell nach (Byström und Järvelin 1995)

Das Byström Järvelin Modell ist das erste, das Informationsverhalten im Kontext von Aufgaben beleuchtet, ein Aspekt, der in Abschnitt 2.5.3 noch konkretisiert wird.

2.2.5 Modell von Leckie, Pettigrew und Sylvain

Das Modell von Leckie et al. (Leckie et al. 1996) ist neben dem Byström Järvelin Modell eines der wenigen zum *information seeking behaviour*, das in empirischen Untersuchungen in einem professionellen Kontext entwickelt wurde, wobei sich die Profession hier auf Domänenwissen bezieht, darunter auch das von Medizinern. Die graphische Darstellung zeigt Abbildung 2-14.

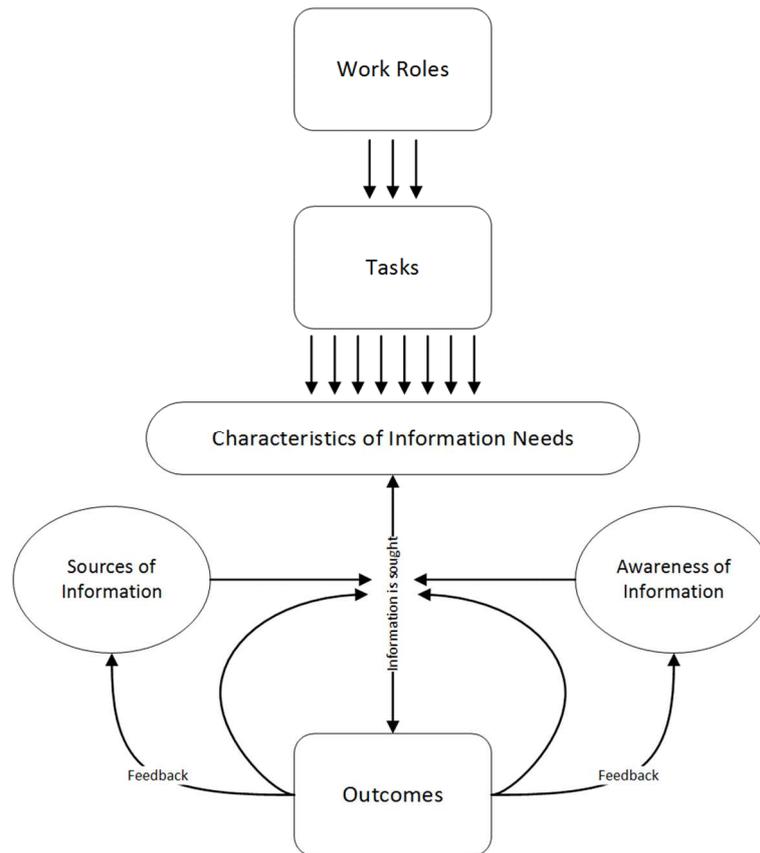


Abbildung 2-14 Modell nach Leckie et al. nach (Leckie et al. 1996)

Das Modell leitet *information need* aus Aufgaben ab, die sich direkt aus dem Betätigungsfeld des Nutzers ergeben; entsprechend ist der Informationsbedarf charakterisiert durch eine Vielzahl von Faktoren:

- individuelle demographische Faktoren (Alter, Berufsverfahrung, Arbeitsumfeld)
- situative Faktoren des Informationsbedarfs (intern oder extern initiiert)
- Häufigkeit des Informationsbedarfs (neu, wiederkehrend, Routinetätigkeiten)
- die Vorhersagbarkeit eines Informationsbedarfs (planbar, unerwartet)
- die Relevanz eines Informationsbedarfs (Grad der Dringlichkeit)
- die Komplexität eines Informationsbedarfs.

Erklärende Variable des Leckie-Modells sind daher Verfügbarkeit und Routine im Umgang mit Informationsquellen, verfügbare Suchstrategien, die Vertrauenswürdigkeit, Zeit- und Kosten sowie Verfügbarkeit von Informationsquellen.

Leckie et al. betonen unter anderem den Stellenwert der persönlichen Routine und Erfahrung (Leckie et al. 1996):

“it is not surprising that studies have shown that professionals tend to rely on their own personal knowledge and experience first when faced with a work-related decision or problem”

Das Leckie-Modell ist, wie von den Autoren richtig zusammengefasst, deshalb interessant, weil es die Komplexität des professionellen Informationsverhaltens aufzeigt und die Notwendigkeit deutlich macht, Informationsverhalten nicht nur quellenabhängig zu studieren, sondern weitere Faktoren zu berücksichtigen, siehe (Leckie et al. 1996).

2.2.6 Modell nach Wilson (1997, 1999)

In einem weiteren Modell (Wilson 1997, 1999) konkretisiert Wilson sein initiales Modell des Informationsverhaltens (vgl. 2.1.3), graphisch in Abbildung 2-15 dargestellt.

Ausgehend vom Ursprungsmodell wird *information seeking behaviour* über einen Informationsbedarf getriggert, nun aber in den Kontext verschiedener bedingender Variablen gesetzt (*intervening variables*) und als aktives oder auch passives Verhalten dargestellt, welches mit der Verarbeitung und Nutzung von Information in einem Feedback-Loop – und damit der Entscheidung, ob und wie der Informationsbedarf gedeckt werden konnte – mündet. Die Entscheidung für konkretes Informationsverhalten begründet das Wilson-Modell mit verschiedenen Theorien und stellt sie in den Kontext unterschiedlicher Einflussfaktoren wie psychologische, demographische, soziokultureller Eigenschaften oder Umgebungsvariablen (vgl. (Wilson 1999)). Wilsons Modell ist nicht empirisch belegt, sondern beschreibt das Konzept *des information seeking behaviour* als interdisziplinären Forschungsgegenstand (vgl. (Wilson 1997)).

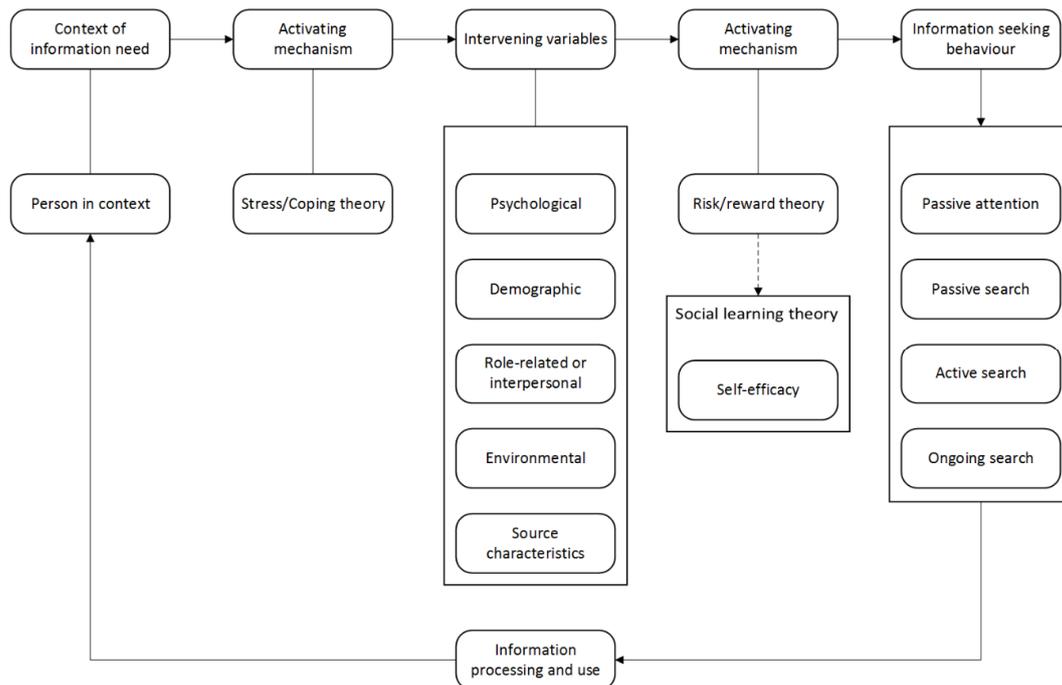


Abbildung 2-15 Wilsons Modell des Information Behaviour (Wilson 1999)

Wie Wilson selbst betont, handelt es sich um die Beschreibung eines „*macro-behaviour*“ (Wilson 1999), das Raum für Erweiterungen, Hypothesen und weitere Forschung lässt. Als postuliertes Modell kann es, wie sein Vorgänger, im Wesentlichen die Aufgabe haben, einen Forschungsgegenstand grob zu definieren.

2.2.7 Zusammenfassung

Wilson's erstes Modell schafft die wissenschaftlichen Rahmenbedingungen zur Erforschung des Informationsverhaltens, das in den folgenden Jahren von der Forschergemeinde aufgegriffen und in verschiedenen Facetten beleuchtet wird.

Das Modell von Ellis zeichnet ein allgemeingültiges und insbesondere empirisch hergeleitetes Bild der Charakteristika des Informationsverhaltens und identifiziert Verhaltensmuster im Umgang mit Information, ohne diese näher zu spezifizieren. Diese statische Sichtweise auf das Informationsverhalten macht das Modell universell, zeigt gleichzeitig aber auch offene Forschungsfragen auf, die zunächst insbesondere Übergänge zwischen einzelnen Verhaltensmustern betreffen. Dies wird in den Modellen von Kuhlthau als auch Marchionini weitergeführt,

in denen Informationsverhalten als Prozess verstanden wird. Das Kuhlthau-Modell hingegen beschreibt den Prozess anhand kognitiver Faktoren, spezifiziert durch eine Gefühlsebene, eine Gedankenebene und Aktionen, ohne dabei auf Kontextvariablen einzugehen –Ansatzpunkt ihres Modells ist das Konzept der Unsicherheit, das sich bereits in den Arbeiten von Dervin (vgl. Abschnitt 2.1.2) findet. Das Marchionini-Modell versteht den ISP als von Rahmenbedingungen abhängigen Problemlösungsprozess im Kontext von Informationsbedarf, deklarativem und prozeduralem Wissen.

Byström und Järvelin sowie Leckie et al. beleuchten *information seeking behaviour* als Resultat eines *information need* erstmals im Kontext von Aufgaben in einem professionellen Arbeitsumfeld, insbesondere Leckie und ihre Kollegen machen die Komplexität des Forschungsgegenstands deutlich.

Die hier vorgestellten Modelle beleuchten die umfassende Sichtweise auf das Phänomen des Information Behaviour, das durch das zweite Wilson Modell in seinen Grundsätzen theoretisch erklärt ist. Keines der Modelle spannt den Bogen zwischen Informationsverhalten und Information Retrieval, vgl. (Ingwersen und Järvelin 2005, S. 105) – dem Prozess der Informationsrückgewinnung mittels spezieller Information Retrieval-Systeme.

2.3 Information Retrieval

Information Retrieval als Disziplin meint die Rückgewinnung von Information: „*The action of recovering stored information, esp. information stored in a computer*“ (Oxford University Press 2016f) und beleuchtet im Wesentlichen die Systemsicht auf den Informationsbeschaffungsprozess.

Die Grundidee des Information Retrieval geht zurück auf Vannevar Bushs visionäres Essay *As we may think* aus dem Jahr 1945. Bush fordert darin die Flexibilität und Zugänglichkeit von Daten für die Wissenschaft (Bush 1945):

“A record if it is to be useful to science, must be continuously extended, it must be stored, and above all it must be consulted.”

und macht deutlich, dass automatische Indexierung anders funktioniert als der menschliche Verstand. Es sei also nötig, Systeme zu entwickeln, die diese Kluft durch technische Lösungen

zumindest in Teilen überbrücken können, um Menschen den Zugang zu gespeicherter Information zu erleichtern. Dies führt direkt zum Ziel des Information Retrieval-Systems und gleichzeitig seiner größten Herausforderung (Baeza-Yates und Ribeiro-Neto 2011, S. 4):

“The primary goal of an IR system is to retrieve all the documents that are relevant to a user query while retrieving as few non-relevant documents as possible.”

- das primäre Ziel eines Information Retrieval-Systems ist es, alle in Bezug auf eine Nutzeranfrage relevanten Dokumente, dabei aber möglichst wenig nicht-relevante Dokumente zu liefern. Die Konzepte *Dokument* und *Relevanz* sind dabei von zentraler Bedeutung, sie werden daher im Folgenden kurz charakterisiert.

2.3.1 Dokumente und Relevanz

2.3.1.1 Dokumente

Der Begriff Dokument ist aus dem Lateinischen abgeleitet (lat. *documentum* = Beispiel; Beweis; Schriftstück sowie *docere* = lehren, unterrichten (Stowasser et al. 2007, Stichwort: *docere*)), im englischen Sprachgebrauch ist es sowohl als „Instruktion“, als auch etwas Gegenständliches beschrieben, das Beweise oder Information zu einem spezifischen Thema enthält. Im Deutschen wie auch im Englischen („*A collection of data in digital form that is considered a single item and typically has a unique filename by which it can be stored, retrieved, or transmitted*“ (Oxford University Press 2016a)) wird Dokument im Zusammenhang mit Urkunden als mit „strukturierten, als Einheit erstellte und gespeicherte Menge von Daten; [Text]datei“ (Bibliographisches Institut GmbH Duden Verlag 2016j) verwendet.

Die Diskussion, ob ein Objekt ein Dokument sein kann, führte zu folgender Definition: Informationswissenschaftlich sind Objekte dann Dokumente, wenn sie (Buckland 1997)

- gegenständlich sind (was auch digitale Dokumente umfasst)
- intentional sind, eine Bedeutung transportieren
- entwickelt oder verfasst wurden
- und als Dokumente wahrgenommen werden.

Anderswo werden Dokumente als das Ergebnis menschlicher Dokumentation (Lund 2010) definiert oder als Wissensartefakte, welche die kulturelle Umgebung widerspiegeln, in der sie entstanden sind (Smiraglia 2008), vgl. auch (Stock et al. 2013, S. 63). Dokumente umfassen demnach nicht nur Texte, sondern auch nicht-textuelles Material, in digitaler Form beispielsweise Bilder, Audio, Video oder auch Software.

Für die vorliegende Arbeit sind vorwiegend Textdokumente aller Art von Interesse. Textdokumente sind heute in verschiedenen Medien zu finden und durchlaufen unterschiedliche Publikationsprozesse, vgl. (Stock et al. 2013, S. 68):

Formal publizierter Text hat einen formalen Publikationsprozess durchlaufen, was für Presseartikel ebenso gilt wie für Bücher, wissenschaftliche Primärliteratur oder beispielsweise Patentschriften. Formal publizierte Textdokumente sind häufig verifiziert, beispielsweise durch Review-Prozesse vor Veröffentlichung wissenschaftlicher Originalartikel, vgl. (Stock et al. 2013, S. 69).

Weniger formal läuft die Publikation bei Grauer Literatur ab. Das Konzept „Graue Literatur“ (engl. *grey (gray) literature*) ist historisch bis in die 1970er Jahre zurückzuverfolgen. Die heute gängige Definition wurde auf der Grey Literature Conference, Luxembourg, 1997 festgelegt: “[Grey literature is] that which is produced on all levels of government, academics, business and industry in print and electronic formats, but which is not controlled by commercial publishers.” (Schöpfel 2010). Unter Graue Literatur fällt also in der Bibliothekswissenschaft jede Publikation, die nicht über Verlage vertrieben wird und daher kontrolliert werden kann, sondern beispielsweise über Vereine oder nicht-profitorientierte Organisationen zugänglich gemacht wird, häufig auch im Internet. Graue Literatur umfasst auch wissenschaftliche Veröffentlichungen, die aus verschiedenen Gründen nicht formell publiziert werden konnten.

Gänzlich informell publizierte Texte haben im Gegensatz zu formal publizierten Dokumenten häufig überhaupt keinen Verifikationsprozess durchlaufen, beispielsweise Webseiten, aber auch Tweets oder Facebook Nachrichten, siehe (Stock et al. 2013, S. 71).

In großen Datenbeständen werden Dokumente häufig nach standardisiertem Verfahren mit Metadaten assoziiert, um sie zu beschreiben und zueinander in Beziehung zu setzen. Metadaten sind als „Daten über Daten“ zu verstehen. Ein gängiges Metadatenschema ist der Dublin Core, der verschiedene Felder vorsieht, Dokumente zu beschreiben (Dublin Core Metadata Initiative (DCMI) 2016). Typische Metadatensätze wie der Dublin Core enthalten neben tech-

nischen Angaben unter anderem Angaben zum Autor, dem Publikationsdatum, der Publikationsquelle, aber auch zum Inhalt. Man kann also zwischen deskriptiven und semantischen Metadaten unterscheiden, vgl. (Baeza-Yates und Ribeiro-Neto 2011, S. 205).

Die entscheidende Rolle bei der semantisch-inhaltlichen Erschließung und Klassifizierung großer Dokumentenbestände für bibliographische Datenbanken spielen Thesauri, die an anderer Stelle noch eingehend erläutert werden. Sie erfüllen bei der Datenbankentwicklung, -pflege und auch beim Retrieval die Aufgabe eines kontrollierten Vokabulars zur Unterstützung von Indexierungs- und Suchprozessen. Das Wort Thesaurus hat sowohl griechische (altgriechisch *θησαυρός thesaurós*) als auch lateinische (lat. *thesaurus*) Ursprünge und bedeutet Schatz, Vorrat, Schatzkammer oder Vorratskammer, metaphorisch auch Fundgrube (Stowasser et al. 2007). Im Information Retrieval bezeichnet ein Thesaurus einen Vorrat bzw. eine Schatzkammer an Wörtern oder auch Phrasen. Die Einträge des Thesaurus sind dabei durch Beziehungen verknüpft, um wiederum Ähnlichkeiten oder hierarchische beziehungsweise assoziative Verbindungen zwischen einzelnen Einträgen zum Zwecke eines verbesserten Retrievals nutzen zu können (National Information Standards Organization (U.S.) 2005). Der Thesaurus als kontrolliertes Vokabular einer Datenbank ist ein wichtiges Werkzeug, um ein „inhaltlich bereits vollständig erschlossenes Informationsangebot zielgenau nach fachlichen Gesichtspunkten zu recherchieren“ (Mayr 2006).

In bibliographischen Datenbanken, also solche, die keine Dokumente, sondern lediglich Referenzen auf diese Dokumente vorhalten, ist Information Retrieval direkt assoziiert mit der Informationssuche nach Metadaten – den bibliographischen Referenzen und damit sowohl abhängig von Datenbank- als auch Thesaurusstrukturen.

2.3.1.2 Relevanz

Ausgehend von einer Suche aufgrund eines Informationsbedarfs sind für den Nutzer nur die Dokumente von Interesse, die in Bezug auf eine initiale Fragestellung *relevant* sind – Information Retrieval-Systeme dienen dem Auffinden dieser Dokumente. So gilt das Konzept Relevanz als wesentliches Element des Information Retrieval und wurde vor allem in den 1950 bis 1970er Jahren, aber auch später eingehend beforscht und hinterfragt (Mizzaro 1997).

Nutzer mit spezifischem Informationsbedarf haben eine konkrete Erwartungshaltung hinsichtlich der in Bezug auf ihren Informationsbedarf relevanten Dokumente. Auch in unserem

Sprachgebrauch haben wir eine sehr intuitive Vorstellung von Relevanz: „Bedeutsamkeit, Wichtigkeit in einem bestimmten Zusammenhang“ beschreibt der Duden den Begriff (Bibliographisches Institut GmbH Duden Verlag 2016f). Relevanz hat jedoch im Information Retrieval auch eine andere Seite: die Systemsicht. Aus Systemsicht ist relevant, was *objektiv* zur Befriedigung eines Informationsbedarfs beitragen kann, die nutzerzentrierte Sicht liefert einen *subjektiven* Blickwinkel. Die systemspezifische Sicht auf Relevanz ist vor allem für die Entwicklung von Information Retrieval-Systemen von Bedeutung.

Unabdingbares Attribut der Relevanz ist die *Relation* – Relevanz kann ausschließlich in Beziehung zu etwas gewertet werden, vgl. für die folgende Ausführung (Saracevic 1996):

Systemspezifische oder algorithmische Relevanz beschreibt die systemspezifische Eigenschaft eines Information Retrieval-Systems, auf eine Suchanfrage genau die Treffer zurückzuliefern, die in Bezug auf diese Suchanfrage systemspezifisch und demnach algorithmisch implementiert als relevant gelten. Sie beschreibt eine Relation zwischen einer Query und einer Menge an Dokumenten. Algorithmische Relevanz kann anhand von Vergleichen empirisch ermittelt werden.

Thematische Relevanz beschreibt die Beziehung zwischen einem Thema, wie es in einer Suchanfrage formuliert ist und dem Thema, wie es im verfügbaren Dokumentenbestand abgedeckt und repräsentiert ist.

Subjektive Relevanz (auch Pertinenz oder kognitive Relevanz) beschreibt die Beziehung zwischen dem aktuellen Wissensstand eines Nutzers, seinem Informationsbedarf und den auf eine Suchanfrage zurückgelieferten oder nicht zurückgelieferten Dokumenten.

Situative Relevanz ist die Relation zwischen der aktuellen Situation, der Aufgabe oder dem Problem und den auf eine Suchanfrage zurückgelieferten oder nicht zurückgelieferten Dokumenten. Situative Relevanz ist dynamisch und kontextspezifisch. Ein Dokument, das zu einem Zeitpunkt noch relevant war, kann zu einem späteren Zeitpunkt, zu dem bereits mehr Information vorliegt, an Relevanz verlieren, vgl. hierzu auch (Wilson 1973).

Affektive Relevanz ist die Beziehung zwischen Intentionen, Zielen und Motivation des Nutzers und den auf eine Suchanfrage zurückgelieferten oder nicht zurückgelieferten Dokumenten.

So bewegt sich der Begriff der Relevanz im Spannungsfeld zwischen Nutzer und System und zielt direkt auf die Interaktion von Nutzern mit Information Retrieval-Systemen ab.

2.3.2 Information Retrieval-Systeme

Die Geschichte der Entwicklung klassischer Information Retrieval-Modelle und Systeme reicht bis in die 1960er Jahre und beginnt mit Dialog, dem ersten kommerziellen Datenbank Online Service, über den die staatlichen Datenbanken ERIC (Institute of Education Sciences (IES) 2016) und NTIS (National Technical Information Service 2016) recherchiert werden konnten. Mit MEDLARS folgte ein Online-Service zur Recherche in der medizinischen Fachdatenbank MEDLINE, vgl. (Stock et al. 2013, S. 97).

2.3.2.1 Architektur

Das Grundprinzip des klassischen Information Retrieval-Systems kann anhand des „Urmodells“ (Abbildung 2-16) veranschaulicht werden.

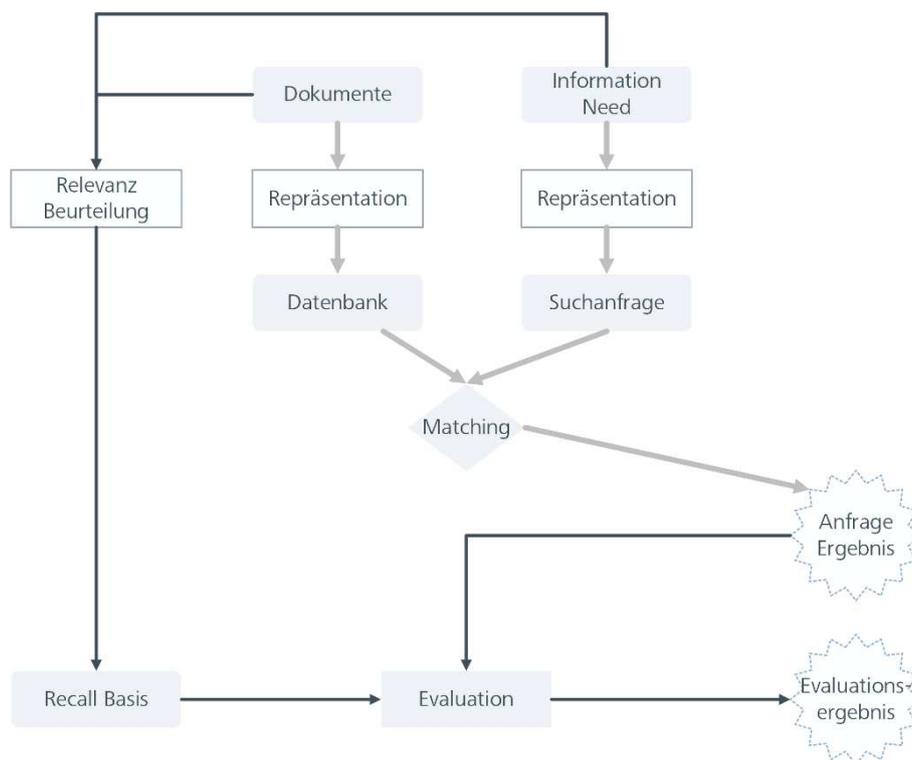


Abbildung 2-16 Grundlagenmodell des Information Retrieval nach (Ingwersen und Järvelin 2005, S. 115)

Retrieval-Systeme bestehen immer aus den zwei zentralen „Komponenten“: Der Dokumentenrepräsentation sowie der Repräsentation eines Informationsbedarfs.

Die Ebene der Dokumentenrepräsentation umfasst dabei die Dokumentinhalte, Dokumentenstruktur als auch Metadaten, wobei Inhalte sowohl durch manuelle als auch automatisierte Indexierungsprozesse erschlossen.

Die Repräsentation der Suchanfrage meint die Formulierung syntaktisch korrekter Suchanfragen aus Suchtermen und je nach Retrieval-Technologie auch Operatoren.

Der Matching-Algorithmus bringt beide Komponenten zusammen und liefert ein Anfrageergebnis. Wie dies technisch umgesetzt wird, ist vom Retrieval-Verfahren abhängig. (Ingwersen und Järvelin 2005, S. 115)

Ergänzend sei hinzugefügt, dass das klassische Suchmaschinen-Modell anders aussieht:

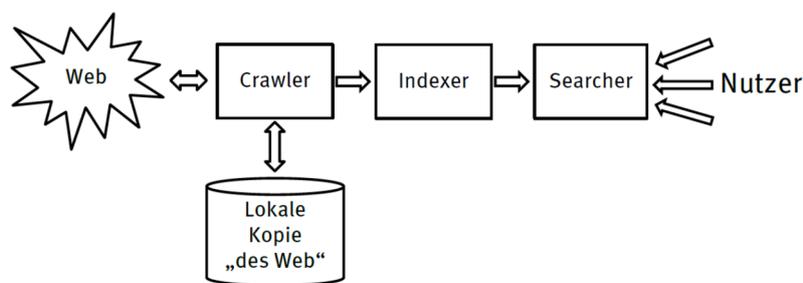


Abbildung 2-17 Aufbau einer Web-Suchmaschine aus (Lewandowski 2013, S. 498)

Die Datenbasis von Suchmaschinen bilden die Inhalte des Webs, die entsprechend verteilt sind, sich dynamisch ändern und durch die Suchmaschine erst aufgefunden werden müssen. Aufgabe von Crawlern ist es, anhand von Hyperlinkstrukturen neue Dokumente zu identifizieren und bestehende auf ihre Aktualität zu prüfen. Über den Indexer werden gecrawlte Dokumente so aufbereitet, dass sie für die Suche gut verarbeitet werden können. Die Anfrage von Nutzern erfolgt durch die Abfrage des Indexers. (Lewandowski 2013, S. 498)

2.3.2.2 Indexierung

Während Websites für die Inhaltserschließung und Indexierung von Websites gecrawlt werden, um schnell und effizient so viele Websites und Verlinkungsstrukturen wie möglich zu

erschließen, erfolgt die Inhaltserschließung zur Dokumentenrepräsentation im klassischen Information Retrieval-System in der Regel über Thesauri, die Werkzeuge zur Indexierung.

Thesauri

Terme oder Deskriptoren bezeichnen im Thesaurus die Indexierungskomponenten und stehen für einzelne fachlich-inhaltliche Konzepte. Ein Deskriptor kann dabei sowohl ein einzelnes Wort als auch eine Phrase sein, also eine Verknüpfung mehrerer Einzelworte. Üblicherweise wird in einem Thesaurus die Pluralform eines Wortes verwendet, sofern dieses im Sprachgebrauch vorkommt. Ähnlichkeitsbeziehungen können als typische Synonyme (Bibliographisches Institut GmbH 2016, Stichwort: Synonym) einzelner Deskriptoren interpretiert werden, während hierarchische Beziehungen der Klassifikation einzelner Terme in Kategorien und Unterkategorien dienen – so sind Thesauri häufig polyhierarchisch konzipiert. Parallel dazu können Terme über assoziative Verbindungen in einen bestimmten Kontext gestellt werden. Thesauri sind üblicherweise proprietäre Konstrukte, die vom Datenbankanbieter selbst entwickelt werden. So verfügen verschiedene bibliographische Fachdatenbanken über eigene Thesauri. (Baeza-Yates und Ribeiro-Neto 2011, S. 229).

Der für die medizinische Fachrecherche sicherlich wichtigste Thesaurus sind die Medical Subject Headings (MeSH), ein polyhierarchisch strukturierter Thesaurus.

Die Medical Subject Headings bilden das Herzstück der Indexierung aller bei der National Library of Medicine (NLM) vorgehaltenen Informationen und Publikationen, unter anderem für die bibliographische Fachdatenbank MEDLINE: So werden MeSH Terme nicht nur zur Indexierung von Artikeln in Datenbanken, sondern auch weiterer Produkte wie beispielsweise Bücher, die über die NLM vertrieben werden, eingesetzt. Der Thesaurus dient damit der Verschlagwortung jeder bibliographischen Referenz und wird daher auch in Suchanfragen eingesetzt, um Datenbankeinträge zu verschiedenen Topics zu finden. (U.S. National Library of Medicine 2016c)

Die Anfänge der Medical Subject Headings in seiner heutigen Form liegen in den 60er Jahren (U.S. National Library of Medicine 2011), sie basieren heute noch auf dem Index des Katalogsystems Index Medicus (National Library of Medicine (U.S.). History of Medicine Division 2014).

MeSH ist nach hierarchischem Prinzip in aktuell 12 Ebenen aufgebaut und verfügt im Jahr

2016 über mehr als 27.000 so genannte Deskriptoren, den so genannten *Main Headings* und über 200.000 Terme, den so genannten *Entry Terms* (U.S. National Library of Medicine 2013). Die *Main Headings* entsprechen den Hauptschlagwörtern und beschreiben als kontrolliertes Vokabular den Inhalt der in den Datenbanken vorgehaltenen Einträge. Eine Menge so genannter Qualifier (*Subject Headings* oder *Subheadings*) beschreibt die möglichen kontextuellen Konzepte hinter den Deskriptoren. Für den Deskriptor „*Ascorbic Acid*“ sind beispielsweise die Qualifier „*pharmacology*“ (Pharmakologie) und „*diagnostic use*“ (diagnostischer Nutzen/Einsatz) hinterlegt.

Unterstützt werden diese Deskriptoren durch Synonyme, den so genannten *Entry Terms*. Für den MeSH Deskriptor „*Ascorbic Acid*“ (Ascorbinsäure) existieren beispielsweise 10 weitere Terme, die in einer Ähnlichkeitsverbindung stehen, unter anderem auch „*Vitamin C*“. Weiterhin wird der Indexterm mit Hilfe assoziativer Verbindungen sowohl im Kontext „*pharmacology*“ (Pharmakologie) als auch „*diagnostic use*“ (diagnostischer Nutzen/Einsatz) definiert. Gleichzeitig reiht sich „*Ascorbic Acid*“ in der Hierarchie unterhalb der Konzepte „*Acids, Acyclic*“ und hier wieder letztendlich unterhalb „*Organic Chemicals*“ ein, einem Indexterm auf oberster Ebene, siehe (U.S. National Library of Medicine 2015b).

MeSH Heading	Ascorbic Acid
Tree Number	D02.241.081.844.107
Tree Number	D02.241.511.902.107
Tree Number	D09.811.100
Annotation	/defic = ASCORBIC ACID DEFICIENCY or SCURVY
Scope Note	A six carbon compound related to glucose. It is found naturally in citrus fruits and many vegetables. Ascorbic acid is an essential nutrient in human diets, and necessary to maintain connective tissue and bone. Its biologically active form, vitamin C, functions as a reducing agent and coenzyme in several metabolic pathways. Vitamin C is considered an antioxidant.
Entry Term	Ascorbic Acid, Monosodium Salt
Entry Term	Ferrous Ascorbate
Entry Term	Hybrin
Entry Term	L-Ascorbic Acid
Entry Term	Magnesium Ascorbate
Entry Term	Magnesium Ascorbicum
Entry Term	Magnesium di-L-Ascorbate
Entry Term	Magnorbin
Entry Term	Sodium Ascorbate
Entry Term	Vitamin C
Allowable Qualifiers	AA AD AE AG AI AN BI BL CF CH CL CS CT EC GE HI IM IP ME PD PH PK PO RE SD SE ST TO TU UR
Pharm. Action	Antioxidants
Pharm. Action	Vitamins

Abbildung 2-18 Sog. Concept View des MeSH Deskriptors "Ascorbic Acid" (National Center for Biotechnology Information 2016), Stand 16.06.2016

Die Struktur des MeSH folgt auf Ebene der Deskriptoren sowohl einer alphabetischen als auch

hierarchischen Struktur, welche sich konzeptuell vom Allgemeinen zum Speziellen vorarbeitet. So ist beispielsweise „*pharmacology*“ als MeSH Deskriptor in zweiter Hierarchieebene unter den übergeordneten Deskriptoren „*Natural Science Disciplines*“ wie auch „*Health Occupations*“ zu finden.

Indexierungsverfahren

Die Indexierung von Dokumenten kann sowohl intellektuell, semi-automatisiert als auch vollständig automatisiert erfolgen.

Die intellektuelle Indexierung hat das Ziel, den Dokumenteninhalte mittels passender Indexterme möglichst gut wiederzugeben und damit zu repräsentieren. Die Stärke der intellektuellen Indexierung zeigt sich darin, dass inhaltlich ähnliche oder inhaltlich gleiche Dokumente konsistent indexiert werden können – eine Stärke gegenüber semi-automatisierten und voll-automatisierten Verfahren. (Lepsky 2013)

MEDLINE und andere medizinische Fachdatenbanken beispielsweise bedienen sich der intellektuellen Indexierung. Vor allem hier gibt es eine Reihe an Aspekten, die den Indexierungsprozess beeinflussen, siehe (Stock et al. 2013, S. 760):

- auf Dokumentenebene geht es bei der Indexierung dabei, die zur Charakterisierung des Dokuments zentralen, inhaltlichen Kernkonzepte zu identifizieren
- auf Ebene des Verschlagwortenden ist das für die Indexierung nötige Domänenwissen ausschlaggebend
- auf Ebene des Datenbanknutzers ist die Frage wichtig, ob Dokumente im Kontext von Suchanfragen mit einer spezifischen Indexierung relevant sind und ob die dort identifizierten Kernkonzepte von Nutzern erkannt und verwendet werden
- auf Datenbankebene muss sichergestellt werden, welche bereits verfügbaren Dokumente ein Konzept vielleicht besser beschreiben und ob ein spezifisches Dokument einen neuen Beitrag dazu leisten kann
- auf Ebene der Wissensrepräsentation muss geklärt werden, ob ein Verschlagwortungskonzept im Kontext der verwendeten Methoden sinnvoll ist oder nicht

- auf Ebene des Domänenwissens muss geklärt werden, ob die identifizierten Kernkonzepte eines Dokuments im Kontext einer wissenschaftlichen Debatte wichtig sind und ob es historisch oder semantisch orientiert indexiert werden sollte.

Intellektuelle Indexierungsverfahren können semi-automatisiert unterstützt werden, um zum einen den hohen Workload manueller Verfahren zu reduzieren und gleichzeitig eine hohe Indexierungsqualität zu gewährleisten. Dies ist bei vollautomatisierten Verfahren in der Regel nicht möglich, weil sie rein über linguistische und statistische Verfahren durchgeführt werden. (Lepsky 2013)

Die Indexierungsproblematik wird in Kapitel 3 wiederholt aufgegriffen und am Beispiel einiger weiterer medizinischer Fachdatenbanken vertieft.

2.3.2.3 Klassische Information Retrieval-Konzepte

Klassische Information Retrieval-Systeme können im Hinblick auf ihren Matching-Algorithmus klassifiziert werden. Man unterscheidet zwischen *exact match* und *partial match* Technologien, eine Klassifikation, die auf Belkin und Croft zurückgeht, vgl. (Belkin und Croft 1987; Ingwersen und Järvelin 2005, S. 117) und heute eine Vielzahl technologischer Ansätze umfasst. Erste Systeme wie Dialog oder MEDLARS basierten auf dem Booleschen Retrieval-Modell, einem *exact match* Modell.

Exact Match: Boolesches Retrieval-Modell

Das Boolesche Retrieval-Modell folgt als *exact match* Modell dem Prinzip „ganz oder gar nicht“: Dokumente sind – auf Systemseite - entweder relevant oder nicht relevant, je nachdem, ob die in der Suchanfrage verwendeten Suchbegriffe in einem Dokument vorkommen oder nicht. Boolesches Retrieval basiert auf der von George Boole entwickelten Booleschen Algebra, einer algebraischen Struktur aus logischen Operatoren AND, OR und NOT beziehungsweise der mengentheoretischen Konzepte Durchschnitt, Vereinigung und Komplementärmenge, siehe detaillierter Abschnitt 2.3.3.1. Die Nutzung der Booleschen Algebra ermöglicht eine binäre Sichtweise auf die Suche: Entweder die Suche nach A UND B führt zum Erfolg, weil der Datenbestand Informationen vorhält, die sowohl in Menge A als auch Menge B enthalten sind, oder nicht. (Belkin und Croft 1987)

Boolesches Retrieval galt bereits früh als relativ starre Technik, was insbesondere auf das *exact match* Prinzip zurückzuführen ist. Die Nachteile von Booleschen Retrieval-Systemen wurden beispielsweise von Belkin und Croft genannt (Belkin und Croft 1987):

- viele relevante Dokumente werden nicht gefunden, weil ihre Repräsentation im Datenbestand die Anfrage nur teilweise erfüllt und sie dadurch nicht aufgefunden werden können
- es findet kein Ranking der zurückgelieferten Treffermenge statt
- Boolesches Retrieval kennt keine Möglichkeit, bestimmte Begriffe innerhalb der Anfrage unterschiedlich zu gewichten, auch Gewichtungen innerhalb von Texten wird nicht berücksichtigt
- die Formulierung der Anfragen ist kompliziert
- die Repräsentation der Anfrage und die Repräsentation des Dokuments müssen im gleichen Vokabular vorliegen

Aus diesen Gründen ist die Boolesche Suche daher in modernen Retrieval-Systemen zwar teilweise noch implementiert, beispielsweise in der Web-Suche, sie wird in der Praxis von nicht-professionellen Nutzern üblicherweise kaum verwendet. Informationsexperten hingegen bevorzugen das Boolesche Retrieval gerade aufgrund seiner Präzision und der Reproduzierbarkeit von Suchanfragen. Das Boolesche Retrieval-Modell findet sich daher heute vorwiegend in der Enterprise-Suche und spielt für die Web-Suche und private Suchprozesse kaum eine Rolle. Historisch betrachtet basieren jedoch die meisten Fachdatenbanken und Hosts, den Hauptquellen für Primärliteratur, auch heute noch auf dem Booleschen Datenbankmodell. (Manning et al. 2008, S. 15)

Best Match-Modelle

Mit dem *best match*-Prinzip versuchte man, Systeme zu schaffen, die anders als das Boolesche Modell. Für die folgende Zusammenstellung siehe (Ingwersen und Järvelin 2005, S. 119). Die grundlegende Idee war, ein System zu schaffen, das

- eine flexible und abgestufte, aber keine harte Unterscheidung zwischen zurückgelieferten und nicht zurückgelieferten Dokumenten macht,

- die Treffer innerhalb der Ergebnisliste entsprechend ihrer Relevanz rankt, indem die relevantesten Treffer zuerst gelistet werden,
- die volle Kontrolle über die Größe der Treffermenge garantiert,
- eine Gewichtung von Suchtermen zulässt und
- einfacher bedienbare Benutzerschnittstellen mitbringt, über die Suchanfragen in natürlicher Sprache gestellt werden können.

Diesen Ansatz verfolgt unter anderem das **Vektorraum-Modell**, einem der prominentesten Vertreter der *best match*-Modelle. Das in den 1960er und 1970er Jahren von Salton basiert auf der Annahme, dass jedes Dokument durch eine Menge zugewiesener Schlüsselbegriffe (*keywords*) und Indexterme als n-dimensionaler Vektor repräsentiert wird. Dies ermöglicht sowohl für Dokumente D_i als auch Suchanfragen Q_i eine vektorisierte Darstellung der Form

$$D_i = (t_{i1}, t_{i2}, \dots, t_{in})$$

$$Q_i = (q_{i1}, q_{i2}, \dots, q_{in})$$

wobei n die Anzahl der im System verfügbaren Indexterme ist und t_{ik} respektive q_{ik} die Werte des k -ten Terms für das Dokument D_i beziehungsweise Q_i . Im Dokumentenraum lassen sich über diese Vektoren Ähnlichkeitsbeziehungen sowohl zwischen einzelnen Dokumenten als auch zwischen Dokumenten und Suchanfragen berechnen. (Salton 1988)

Das Auffinden relevanter Dokumente erfolgt nun durch die Suche nach den Dokumenten, deren Vektor möglichst ähnlich dem der Suchanfrage ist. Dabei bestimmen sowohl die Lage des Vektors im Vektorraum als auch Winkel zwischen einzelnen Vektoren das Ähnlichkeitsmaß von Suchanfrage und Dokument, vgl. (Stock et al. 2013, S. 291). Dokumente werden also auch dann zurückgeliefert, wenn sie mit der Suchanfrage nur in Teilen übereinstimmen (*partial match*).

Grundsätzlich bestand der Fortschritt des Vektorraum-Modells für den Laien also genau darin, dass durch reine Freitext-Suchanfragen für den Anwender die vergleichsweise komplexe Nutzung Boolescher Operatoren entfällt, weil die Notwendigkeit, Suchterme mittels Operatoren in Verbindung zu setzen, entfiel. Positiv hervorzuheben sind daher Eigenschaften, die *exact match* Modelle nicht mitbringen: Durch Term-Gewichtung kann das Retrieval-Ergebnis deut-

lich verbessert werden, partielle Übereinstimmung ermöglicht zudem das nach Relevanz geordnete Retrieval zurückgelieferter Dokumente.

Im Gegensatz zum klassischen Booleschen Modell erlaubt das Vektorraum-Modell eine nach Relevanz geordnete Rückgewinnung von Treffermengen – in Form einer nach Relevanz geordneten Liste. Entscheidend beim Erfolg dieser *partial match* Systeme sind also vor allem die Güte des Ranking-Algorithmus, insbesondere natürlich bei großen Dokumentenbeständen.

Dies unterscheidet das Vektorraum-Modell ganz klar vom Booleschen Retrieval-Modell und legte mit seiner Ranking-Methode gleichzeitig konzeptionell den Grundstein für die Entwicklung von Suchmaschinen, die vom Nutzer keine exakten Suchanfragen erwarten und mit ihren riesigen Datenbeständen daher auf Ranking-Algorithmen angewiesen sind.

Auch **probabilistisches Retrieval** versucht, die Probleme des *exact match* Boolescher Systeme zu überbrücken. Probabilistisches Retrieval beruht auf statistischen Methoden und der Frage, wie wahrscheinlich es ist, dass einzelne Dokumente im Hinblick auf eine spezifische Suchanfrage relevant sind. Probabilistisches Retrieval rankt daher Dokumente anhand der auf Systemseite berechneten Wahrscheinlichkeit ihrer Relevanz. (Robertson 1977)

Das „probabilistische Ranking-Prinzip“ verdeutlicht die Idee hinter diesem Retrieval-Modell: Probabilistische Modelle versuchen, für ein Dokument D_i und eine vorgegebene Suchanfrage Q die Wahrscheinlichkeit abzuschätzen, mit der Nutzer dieses Dokument als relevant einstufen. Hierbei gilt die Annahme, dass die Wahrscheinlichkeit für Relevanz ausschließlich von der Suchanfrage und dem Dokument abhängt und damit systemabhängig als auch nutzerunabhängig ist. Das Modell geht weiterhin von einer „idealen Antwortmenge“ R aus, in der die Wahrscheinlichkeit am höchsten ist. Für Dokumente aus dieser Antwortmenge R kann vorhergesagt werden, dass sie relevant sind, für Dokumente außerhalb, dass sie nicht relevant sind. (Robertson 1977)

Mathematisch lässt sich der Sachverhalt mittels des Bayes-Theorems folgendermaßen darstellen (Stock et al. 2013, S. 301):

$$P(D|Q) = \frac{P(Q|D) \cdot P(D)}{P(Q)}$$

$P(D|Q)$ bezeichnet die Wahrscheinlichkeit der Relevanz eines Dokuments D unter der Bedingung (Suchanfrage) Q . $P(D|Q)$ ist – vereinfacht gesagt – nicht die einzige Unbekannte dieser Gleichung. In der Praxis ist $P(Q|D)$ nicht bekannt und muss daher mittels *Relevance Feedback*-Loops (siehe hierzu auch Abschnitt 2.3.2.4 abgeschätzt werden – entweder auf System- oder auf Nutzerseite, siehe ebenfalls (Stock et al. 2013 am angegebenen Ort).

Weiterentwicklungen, wie das von Salton et al. 1983 vorgeschlagene *Extended Boolean Information Retrieval* mischen die Vorteile exakter Matching-Algorithmen mit dem Prinzip des Rankings. Dokumente sind in diesem Modell dann relevant oder relevanter, wenn sie mit wenigstens einem Teil der Suchanfrageterme übereinstimmen und werden entsprechend in einer gerankten Treffermenge ausgegeben (Salton et al. 1983).

Zum Thema des *Relevance Ranking* im Kontext Boolescher Datenbanken vgl. (Davis 1995), der die Technik folgendermaßen begründet:

“The problem is this: perfectly good Boolean logic can lead to enormous retrieval sets given the sizes of our research libraries and specialized databases. Ranking output with a well established but still little-known technique called weighted-term searching provides a general and straightforward solution to the problem.”

In den 1960er bis 1980er Jahre wurden eine Vielzahl an Forschungsfeldern in Information Retrieval bearbeitet, darunter nicht nur Systeme zur Verbesserung des *partial match*-Prinzips, sondern auch Methoden zur Verbesserung der Retrieval-Qualität oder Benutzerschnittstellen. So wurden zahlreiche weitere Retrieval-Modelle entwickelt, bestehende wurden verbessert und ausgebaut.

Vektorraum-Modell und probabilistisches System wurden beispielsweise ergänzt um netzwerkbasierte Systeme, Hypermedia-Systeme, Sprachmodell und andere. Eine Übersicht findet sich beispielsweise bei (Baeza-Yates und Ribeiro-Neto 2011), (Ingwersen und Järvelin 2005, Kapitel 4) oder (Stock et al. 2013).

Wie bereits angedeutet, hat sich die Informationslandschaft in den vergangenen Jahren vor allem durch das Internet noch einmal drastisch verändert und zog in der Konzeption ergänzender Information Retrieval-Technologien und dem Erkennen neuer Forschungsaspekte eine umfangreiche Weiter- und Neuentwicklung nach sich. Das Web nutzt ausschließlich Ranking-

Verfahren und macht sich dabei seine grundlegenden Konzepte wie Verlinkung auch für Ranking-Algorithmen zunutze, vgl. (Manning et al. 2008, S. 421).

Das Boolesche Retrieval-Modell als *exact match* Modell selbst findet sich heute vorwiegend in der Enterprise-Suche und spielt abseits davon heute kaum eine Rolle. Für die weitere Arbeit sind im Wesentlichen Systeme nach dem Booleschen Retrieval-Modell interessant: Historisch betrachtet basieren die meisten Fachdatenbanken und Hosts, den Hauptquellen für Primärliteratur, auch heute noch auf dem Booleschen Datenbankmodell, vgl. (Baeza-Yates und Ribeiro-Neto 2011, S. 700).

2.3.2.4 Query Expansion und Relevance Feedback

Ein Problem aller Retrieval-Systeme ist das sprachlicher Synonyme und des „idealen“ Suchbegriffs: ein und dasselbe Konzept kann prinzipiell mittels einer Vielzahl an Synonymen ausgedrückt werden – eine Tatsache, die den Nutzer vor allem im Booleschen Retrieval vor Probleme stellt, aber für gute Retrieval-Qualität essentiell ist.

Eine typische Lösung dieses Problems ist die der Modifikation von Suchanfragen, im Einzelnen *Query Expansion* oder *Query Reformulation*, der Erweiterung oder der gänzlichen Reformulierung. Die Methoden hierfür sind vielfältig und sollen hier nur kurz skizziert werden. Im einfachsten Fall erfolgt die Reformulierung von Suchanfragen manuell durch den Nutzer, beispielsweise im Booleschen Retrieval (Manning et al. 2008, S. 177). Idealerweise jedoch wird dieser Prozess durch das System unterstützt.

Systemgestützte Methoden können zunächst grob in global und lokal eingeteilt werden:

Globale Methoden umfassen Techniken, die unabhängig von einer gegebenen Query und unabhängig von gegebenen Treffermengen sind. Hierzu gehören beispielsweise Methoden zum Einsatz von Thesauri oder Korrekturen an Suchanfragen durch Tippfehler oder falscher Buchstabierung. Ein typisches Beispiel für eine solche globale Methode ist das bei Pubmed eingesetzte *automatic term mapping*, das in Abschnitt 3.4.2.3 beschrieben wird. Hierbei werden vom Nutzer gestellte Suchanfragen mit dem Thesaurus abgeglichen, die eigentlich abgesendete Suchanfrage nutzt dann nicht nur die Suchterme des Nutzers sondern auch korrespondierende Deskriptoren. (Manning et al. 2008, S. 177)

Die typische „*find similar*“ Funktion, also die Ähnlichkeitssuche findet sich heute nicht nur im Web, sondern auch auf Booleschen Datenbanken wie Pubmed, siehe hierzu im Folgenden Abschnitt 3.4.5.1. Untersuchungen zur Analyse von Suchanfrage-Reformulierungen und der Implementierung unterstützender Verfahren nehmen in der Forschung großen Raum ein und werden auch im medizinischen Kontext durchgeführt, vgl. (Spink et al. 2001; Herskovic et al. 2007; Lu et al. 2009; Wacholder 2011; Mosa und Yoo 2013).

Lokale Methoden justieren Suchanfragen anhand der korrespondierenden Treffermenge. erfordert mittels des so genannten *Relevance Feedback* eine direkte oder indirekte Rückmeldung durch den Nutzer (Manning et al. 2008, S. 177).

Die Idee hinter *Relevance Feedback* ist, den Nutzer in den Retrieval Prozess einzubinden, um das Retrieval-Ergebnis zu verbessern, indem der Nutzer Feedback zum initialen Suchergebnis gibt. Den typischen Ablauf beschreiben Manning et al. folgendermaßen (Manning et al. 2008, S. 178):

- Der Nutzer sendet eine meist kurze und einfache Suchanfrage
- Das System liefert eine erste Treffermenge auf Basis dieser Suchanfrage
- Der Nutzer markiert die für ihn relevanten wie auch irrelevanten Treffer
- Das System berechnet anhand dieses Feedbacks eine dem Informationsbedarf angemessenere Treffermenge
- Das System liefert die modifizierte Treffermenge zurück

Obwohl *Relevance Feedback* in der Praxis zu signifikant besseren Retrieval-Ergebnissen führen kann, ist es bei Nutzern tendenziell eher unpopulär: Sie geben schlicht ungern eine Rückmeldung, vgl. (Fuhr 2013).

Wie sich im Folgenden zeigen wird, ist die Reformulierung von Suchanfragen vor allem in der professionellen Recherche häufig noch Aufgabe des Nutzers, auch wenn Fachdatenbanken inzwischen (vor allem globale) Methoden der *Query Reformulation* implementieren.

2.3.2.5 Systemseitige Evaluierung von Retrieval-Systemen

Aus Systemsicht eignen sich zur Evaluierung von Information Retrieval-Systemen zwei Größen, die sich direkt aus Anspruch und Ziel des Systems ableiten lassen: Nutzer sollen möglichst alle

für ihre Suchanfrage relevanten Dokumente finden und dabei mit einer möglichst geringen Menge nicht-relevanter Dokumente konfrontiert werden.

Die Qualität von Suchergebnissen wird daher anhand der Größen Recall und Precision gemessen. Hierfür sind drei Parameter von Bedeutung, vgl. (Stock et al. 2013, S. 114):

- Anzahl zurückgelieferter relevanter Dokumente r_p
- Anzahl nicht-relevanter Dokumente in der Treffermenge f_p (auch Ballast oder *noise*)
- Anzahl relevanter Dokumente der Gesamtmenge, die nicht zurückgeliefert wurden r_n (*lost documents*)

Der Recall ist definiert als der Quotient aus der Anzahl korrekt zurückgelieferter relevanter Dokumente und der Summe aller relevanten Dokumente:

$$\frac{r_p}{r_p + r_n}$$

Die Precision ist definiert als der Quotient aus der Anzahl korrekt zurückgelieferten relevanten Dokumenten und der Summe aller zurückgelieferten Dokumente:

$$\frac{r_p}{r_p + f_p}$$

Die beiden Größen Precision und Recall sind demnach rein systemspezifische Größen, um die Performanz von Retrieval-Systemen messbar zu machen. Im Gegensatz zur Precision, die exakt berechnet werden kann, ist der Recall in der Praxis jedoch ein theoretisches Konstrukt, da der Wert für r_n nicht direkt ermittelt werden kann. Aufgrund der schweren Messbarkeit des Recalls insbesondere in großen Datenbeständen ist es möglich, auf den *relativen* Recall auszuweichen, der definiert ist als der Quotient aus der Anzahl zurückgelieferter relevanter Dokumente und der von Nutzern erwarteten Anzahl an zurückgelieferten relevanten Dokumenten, vgl. (Baeza-Yates und Ribeiro-Neto 2011, S. 145). Vor allem für die Informationsgewinnung im Internet ist der Recall als Größe daher irrelevant: Es ist hier schlicht nicht möglich, ihn zu bestimmen.

Wie sieht die Evaluation in der Praxis aus?

Standardmäßig ist zur Evaluation eines Information Retrieval-Systems mittels der Größen Precision und Recall eine Testmenge erforderlich, die folgendes erfüllt – eine Anforderung, die auf das so genannte Cranfield-Paradigma zurückgeht, vgl. (Manning et al. 2008, S. 152; Womser-Hacker 2013):

- sie enthält eine Dokumentsammlung
- sie verfügt über eine Testmenge vorgegebener Fragestellungen, die wiederum als Suchanfragen formulierbar sind
- sie verfügt über eine Menge an Relevanzbewertungen für jedes Suchanfrage-Dokumenten-Paar und klassifiziert diese als relevant beziehungsweise nicht-relevant

Unter anderem zur Forschungsunterstützung des Information Retrieval und zur Evaluierung von Information Retrieval-Systemen wird seit 1992 die TREC Conference abgehalten (National Institute of Standards and Technology (NIST) 2016), für weitere Testsuiten und Verfahren siehe (Manning et al. 2008, S. 153). Der Nutzen dieser TREC-Konferenzen liegt in der gemeinsamen Verwendung von Test-Kollektionen und damit der direkten Vergleichbarkeit der Ergebnisse einzelner Arbeitsgruppen, insbesondere aber in der Bildung einer Information Retrieval-Community (Sanderson 2010).

Die Evaluierung von Systemen erfolgt hierbei auf Testdatenbanken mit entsprechend konstruierten Dokumentenbeständen, die Recall-Beurteilung erfolgt auf Basis vorgegebener Queries und Relevanzbeurteilungen wie oben beschrieben.

Der jährlich stattfindende TREC Workshop besteht dabei aus einer Reihe sogenannter *tracks*, den Forschungsschwerpunkten, in denen wiederum spezifische Aufgaben definiert werden. Ein *track* erfüllt dabei mehrere Ziele:

Zum einen geht es um ein Problemverständnis allgemein, zum anderen aber auch um die Entwicklung nötiger Infrastrukturen, bestehend unter anderem aus Test Kollektionen und Evaluationskriterien. Tracks demonstrieren zudem die Robustheit bestehender Technologien im Hinblick auf verschiedene Task (TREC Conference 2002).

Die klassische TREC-Aufgabe ist das so genannte Ad Hoc-Retrieval, bei dem die Aufgabe darin besteht, eine statische Menge von Dokumenten anhand verschiedener, vordefinierter Themen zu durchsuchen, vgl. (Womser-Hacker 2013).

Erst in den vergangenen Jahren ist eine Abkehr von der systemzentrierten Bewertung von Information-Retrieval Systemen zu beobachten, die sich mit der Frage beschäftigt, wie zufrieden Nutzer mit der Retrieval-Qualität einzelner Systeme sind. Eine Methode zur Einbeziehung des Nutzers sind Maße wie *cumulative gain*, beschrieben bei (Järvelin und Kekäläinen 2002), bei dem beurteilt wird, wie viel Ertrag der Nutzer aus der ermittelten Treffermenge ziehen kann. Varianten dieser Vorgehensweise berücksichtigen hierbei zudem den Aufwand des Nutzers bei der Abarbeitung der Treffermenge, vgl. (Womser-Hacker 2013).

Neueste Modelle, die unter anderem in der professionellen Suche zur nutzerzentrierten Bewertung von Retrieval-Systemen verwendet werden, nutzen beispielsweise Use-Cases (Hansen et al. 2014).

2.3.2.6 Ausblick

Aus systemseitiger Sicht bleiben – so Ingwersen und Järvelin – wichtige Fragen offen, unter anderem: Wie kann der Nutzen von Information Retrieval-Systemen in bestimmten Nutzungskontexten bewertet und verbessert werden? (Ingwersen und Järvelin 2005, S. 188)

Dies führte bereits in den 1980er Jahren zu einer eher nutzer-zentrierten und kognitiven Sichtweise auf die Entwicklung und Beforschung von Information Retrieval-Systemen.

Nachdem im folgenden Abschnitt das für die Arbeit wesentliche Boolesche Datenbankmodell eingehend beschrieben wird, stellt Abschnitt 2.4 die in diesem Kontext relevanten Bemühungen vor, darunter insbesondere die für die Suche in Online-Datenbanken grundlegenden Arbeiten von Marcia Bates und Raya Fidel.

2.3.3 Boolesches Datenbankmodell

Wie bereits angedeutet, basieren die meisten Fachdatenbanken auch heute noch auf dem Booleschen Retrieval-Modell. Der vorliegende Abschnitt beleuchtet die Grundelemente und die daraus resultierenden Methoden zur Informationssuche.

2.3.3.1 Grundprinzip des Booleschen Retrieval-Modells

Boolesches Retrieval baut auf der nach George Boole benannten Booleschen Algebra auf. Es beruht auf der Tatsache, dass zu jedem Indexterm in einem Information Retrieval-System die eindeutige Aussage getroffen werden kann, ob er einem Dokument zugewiesen ist (1 = wahr) oder nicht (0 = falsch).

Tabelle 2-2 Boolesche Operatoren

Boolescher Operator	Beschreibung
AND	A AND B findet alle Dokumente, auf die A und gleichzeitig auch B zutrifft
OR	A OR B findet alle Dokumente, auf die A oder B zutrifft
NOT	NOT A schließt Dokumente aus, auf die A zutrifft
XOR	A XOR B schießt Dokumente ein, auf die entweder ausschließlich A oder ausschließlich B zutreffen.

Die Menge zurückgelieferter Dokumente folgt entsprechend des Booleschen Retrieval je nach Suchanfrage einem „ganz oder gar nicht“ Prinzip (*exact match*). Ein Dokument ist je nach Suchanfrage relevant oder nicht relevant, daher werden jeweils nur die Dokumente zurückgeliefert, die eine gesuchte Kombination von Wörtern enthalten.

Für ein Vokabular $V = \{A, B, C, D\}$ sind folgende Dokumentenrepräsentationen denkbar:

Ein Dokument D_1 enthält A, B und C, jedoch nicht D: $D_1 = \{1, 1, 1, 0\}$

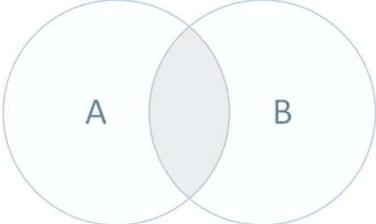
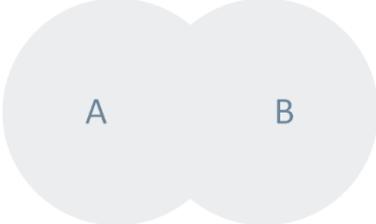
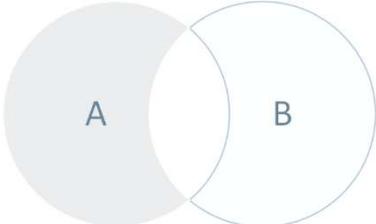
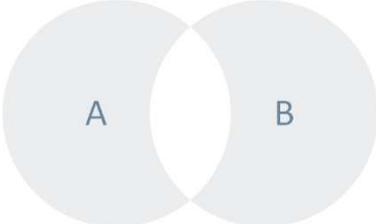
Ein Dokument D_2 enthält A und B, jedoch nicht C und D: $D_2 = \{1, 1, 0, 0\}$

Ein Dokument D_3 enthält B, C und D jedoch nicht A: $D_3 = \{0, 1, 1, 1\}$

Eine Suchanfrage (A OR B) AND C würde demnach die Dokumente D_1 und D_3 , D_2 hingegen würde durch das Suchraster fallen. (B OR C) NOT D würde D_1 und D_2 zurückliefern.

Mengenalgebraisch stellt sich die Verwendung der Booleschen Operatoren wie folgt dar:

Tabelle 2-3 Mengenalgebraische Sicht auf die Booleschen Operatoren

Boolescher Operator	Darstellung
AND	
OR	
NOT	
XOR	

Die Verwendung des NOT-Operators in großen Datenbeständen ist in der Praxis kritisch zu bewerten, da durch unbedachten Ausschluss einzelner Konzepte (alle Dokumente die ein Konzept nicht enthalten) die Gesamttreffermenge ungeahnt anwachsen kann (Baeza-Yates und Ribeiro-Neto 2011, S. 259).

Mathematisch gesehen definiert eine Boolesche Algebra eine Menge aus dem Nullelement 0 und dem Eins-Element 1 mit den beiden zweistelligen Verknüpfungen UND (in der Schreib-

weise von Peano \wedge) als Multiplikation und ODER (\vee) als Addition sowie der einstelligen Verknüpfung NOT (\neg) (Monk 2002).

Auf dieser Algebra gelten daher unter anderem das Kommutativgesetz, das Assoziativgesetz sowie das Distributivgesetz, aber auch weitere mathematische Gesetzmäßigkeiten, wie in folgender Tabelle zusammengestellt.

Tabelle 2-4 Gesetze im Umgang mit Booleschen Operatoren nach (Monk 2002)

Gesetz	Beispiel
Kommutativgesetz	$A \text{ OR } B = B \text{ OR } A$ $A \text{ AND } B = B \text{ AND } A$
Assoziativgesetz	$(A \text{ OR } B) \text{ OR } C = A \text{ OR } (B \text{ OR } C)$ $(A \text{ AND } B) \text{ AND } C = A \text{ AND } (B \text{ AND } C)$
Distributivgesetz	$A \text{ AND } (B \text{ OR } C) = (A \text{ AND } B) \text{ OR } (A \text{ AND } C)$ $A \text{ OR } (B \text{ AND } C) = (A \text{ OR } B) \text{ AND } (A \text{ OR } C)$
Idempotenzgesetz	$A \text{ AND } A = A$ $A \text{ OR } A = A$
De Morgan'sche Gesetze	$\text{NOT } (A \text{ AND } B) = (\text{NOT } A) \text{ OR } (\text{NOT } B)$ $\text{NOT } (A \text{ OR } B) = (\text{NOT } A) \text{ AND } (\text{NOT } B)$
Komplementärgesetze	$A \text{ AND } (\text{NOT } A) = 0$ $A \text{ OR } (\text{NOT } A) = 1$
Absorptionsgesetze	$A \text{ OR } (A \text{ AND } B) = A$ $A \text{ AND } (A \text{ OR } B) = A$

2.3.3.2 Retrieval in Booleschen Datenbanksystemen

Die Suche in einem Booleschen Retrieval-System erfolgt anhand einer Suchanfrage, aufgebaut aus einzelnen Termen oder Phrasen in Verbindung mit Booleschen Operatoren, im Folgenden auch synonym als Query bezeichnet.

Terme und Phrasen

Aus sprachwissenschaftlicher Sicht kann ein Term als „nicht wahrheitsfähiger, wohlgeformter Bestandteil einer Aussage“ interpretiert werden (Bußmann et al. 2008, Stichwort: Term). In der Definition von Stock (Stock 2007, S. 321) findet sich des Weiteren neben den Definitionen unbearbeitete, flektierte Wortform, Grundform und Wortstamm auch der Hinweis auf die Sonderform „*named entities*“. So können also neben „Wörtern“ auch trunkierte Wortformen als Terme interpretiert werden. Dies gilt insbesondere für Suchterme, also Terme, die in einer datenbankspezifischen Query eingesetzt werden.

Phrasen sind in einer Sprache als Ausdrücke zu verstehen, die aus mehreren Einzelworten gebildet sind. Besonders wichtig ist die Phrase im Englischen (da es dort keine zusammengesetzten Substantive gibt) oder aber bei vielen medizinischen Fachbegriffen ("*Diabetes mellitus*"), Firmennamen etc. (Stock 2007, S. 248). Phrasen spielen daher im Kontext der englischsprachigen als auch der medizinischen Literaturrecherche eine wichtige Rolle.

Bei Suchanfragen auf bibliographischen Datenbanken basierend auf Booleschem Retrieval muss zwischen Freitexttermen und kontrolliertem Vokabular, den Deskriptoren, unterschieden werden. Dabei ist das kontrollierte Vokabular nicht begrenzt auf Einzelwörter, sondern enthält ebenfalls Phrasen, vgl. hierzu Abschnitt 2.3.2.2.

Weitere Suchterme werden unter dem Begriff Freitextterme zusammengefasst. Freitextterme sind sowohl Terme, die im Verschlagwortungskatalog als Synonyme zu Deskriptoren genannt werden (wie beispielsweise Entry Terms im MeSH-Thesaurus, siehe auch hierzu Abschnitt 2.3.2.2) als auch Textwörter oder Textphrasen, die direkt mit der im Datensatz referenzierten Originalarbeit assoziiert sind, weil sie beispielsweise im Titel, im Abstract oder als Keyword eines bibliographischen Dokuments vorkommen und damit im Metadatensatz der bibliographischen Referenz zu finden sind. Zu den Freitexttermen zählen aber auch sprachliche Synonyme, die dem Nutzer zwar bekannt sind, von ihm aber weder im Thesaurus noch in irgendeiner bibliographischen Referenz aktiv als Textwort identifiziert wurden – in diesem Fall greift der Nutzer also nicht auf einen bekannten oder wenigstens für ihn reproduzierbar vorhandenen Term des Gesamtsystems zu, sondern nutzt „auf gut Glück“ eigenes Vokabular.

Für einen flexiblen Umgang mit Freitexttermen bieten viele Datenbanken Trunkierungsmöglichkeiten in ihren Retrievalsprachen an. Verallgemeinert kann hier von Suchmustern gesprochen werden, englisch *pattern matching*.

Kernkonzept hinter diesen Mustern bilden reguläre Ausdrücke. Unter einem solchen versteht man eine Zeichenkette (*string*), in dem normale, beispielsweise alphanumerische Zeichen mit so genannten Metazeichen kombiniert sind. Diese wiederum geben Auskunft über weitere Konzepte, beispielsweise die Menge, Position oder die Art von Zeichen, vgl. (Stubblebine 2008, S. 11). Mittels Mustersuche können in Datenbanken Texte aufgefunden werden, die den durch reguläre Ausdrücke definierten Regeln genügen. Reguläre Ausdrücke werden in Programmiersprachen allerdings nicht einheitlich verwendet und dies gilt auch für Datenbanken.

Der einfachste Fall eines Musters sind Wildcards, meist durch ein * gekennzeichnet. Das * ersetzt dabei jedes beliebige Zeichen. Wildcards finden nicht nur in Datenbanken Anwendung, sondern beispielsweise auch bei Suchmaschinen oder auf Betriebssystemen zur Suche nach Dateien. Sie dienen dazu, Wortfragmente zu finden, so liefert beispielsweise „männ*“ sowohl Männer als auch männlich.

Operatoren

Während die Booleschen Operatoren in ihrer Verwendung international vereinheitlicht sind, handelt es sich bei den ebenfalls häufig verfügbaren PROXIMITY- oder Wortabstandsoperatoren um eher proprietäre Konstruktionen mit zwar identischer Logik jedoch häufig individueller Syntax.

Proximity kann als „Auflockerung“ der Phrasensuche verstanden werden (Baeza-Yates und Ribeiro-Neto 2011, S. 258). Hierbei wird neben einer beliebigen Anzahl einzelner Terme oder Phrasen beispielsweise der längste erlaubte Wortabstand zwischen diesen Termen angegeben. Mit Hilfe von PROXIMITY-Operatoren ist es möglich, einzelne Suchterme oder Phrasen im Titel, Abstract oder in der Metabeschreibung eines Datenbankeintrags zu spezifizieren.

A ADJ5 B findet beispielsweise A im Abstand von 5 Wörtern zu B, wobei datenbankspezifische Stopwörter ignoriert werden (Stock et al. 2013, S. 248). Aus Retrieval Sicht werden auch Operatoren, die einen Suchterm innerhalb eines bestimmten Datenbankfelds, beispielsweise

dem Titel oder Abstract einer bibliographischen Referenz, lokalisieren, als PROXIMITY-Operatoren bezeichnet (Baeza-Yates und Ribeiro-Neto 2011, S. 704).

Information Retrieval in einer Datenbank entspricht also einer Suchanfrage mittels eines datenbankspezifisch ausformulierten Suchprofils bestehend aus Termen, Phrasen und Operatoren. Abschnitt 3.4.2 beleuchtete Spezifika der Suchanfrage auf medizinischen Fachdatenbanken.

Benutzeroberflächen

Typischerweise erfolgt die Suchanfrage heute häufig über ein einfaches Texteingabefeld in Kombination mit formular- oder menübasierten Webinterfaces, was gleichzeitig die Kenntnis der dahinterliegenden Retrievalsprache minimiert. Mit dieser Funktionalität wird vor allem Domänenexperten die Suche in Booleschen Datenbanken erleichtert (vgl. hierzu (Stock et al. 2013), S. 253f.).

Viele Datenbanken und Hosts bieten einen parallelen Profimodus, in dem die Suchanfrage wie auch die Anzeige der Treffermenge auf Kommandoebene erfolgt. Dies entspricht der Nutzung wie sie in den Anfängen online-verfügbarer Datenbanken üblich und notwendig war, vgl. (Markey und Cochrane 1978). In diesem Fall ist die Kenntnis der Retrievalsprache und ihrer Details weiterhin unumgänglich.

Üblicherweise ermöglichen Fachdatenbanken außerdem den direkten Zugriff auf ihre Thesauri und damit das Browsen durch kontrolliertes Vokabular, vgl. (Baeza-Yates und Ribeiro-Neto 2011, S. 705).

Vor allem die Anbieter großer bibliographischer Fachdatenbanken befinden sich heute in direkter Konkurrenz mit dem Internet. Sie sind daher gezwungen, sich bei der Gestaltung und Funktionalität ihrer Produkte an den Wünschen ihrer zahlungswilligen Kunden zu orientieren, die meist in Wirtschaft und Wissenschaft zu finden sind. Viele Boolesche Datenbanken werden daher um Funktionen wie die Entfernung von Duplikaten, komplexe Ranking-Algorithmen oder Sortierfunktionen erweitert, vgl. (Baeza-Yates und Ribeiro-Neto 2011, S. 705). Dieser Aspekt wird im Verlauf der Arbeit bei genauerer Betrachtung medizinischer Fachdatenbanken an der ein oder anderen Stelle deutlich.

2.3.4 Zusammenfassung

Der vorliegende Abschnitt diente dem Einblick in das Grundprinzip des Information Retrieval und stellte zunächst die drei technischen Hauptrichtungen an Information Retrieval-Systemen kurz vor. Für die weitere Arbeit sind ausschließlich Systeme nach dem Booleschen Retrieval-Modell relevant, da die meisten Fachdatenbanken und Hosts auch heute noch auf dem Booleschen Datenbankmodell basieren. Dass sich daran bis heute kaum etwas geändert hat, kann neben gewachsenen Strukturen vor allem auf die erforderliche Rechenleistung in großen Datenbanken zurückgeführt werden, die im Booleschen Retrieval noch vergleichsweise gering ausfällt (Baeza-Yates und Ribeiro-Neto 2011, S. 703).

Wer also heute noch in bibliographischen Fachdatenbanken recherchiert, ist mit dem Problem konfrontiert, seine Suchanfrage sowohl syntaktisch einwandfrei als auch inhaltlich präzise zu formulieren.

2.4 Suchstrategien auf Booleschen Datenbanken

Das Konzept des Booleschen Retrieval mündet vor allem bei Themengebieten, in denen das Auffinden *aller* themenrelevanter Dokumente von wesentlicher Bedeutung ist, beispielsweise der Patentsuche oder aber auch der evidenzbasierten Medizin, in der Notwendigkeit, sich Suchtechniken und Suchstrategien anzueignen, die sicherstellen, dass diesem Anspruch so gut wie möglich Genüge getan werden kann.

2.4.1 Das Konzept der Suchstrategie

Die Forschungsaktivitäten um das Verständnis und die Beschreibung von Suchtechniken und Suchstrategien im Kontext bibliographischer Datenbanken ist aus informationswissenschaftlicher Sicht eng verknüpft mit Marcia Bates (Bates 1979a, 1979b, 1989, 1990). Ihre Arbeiten sind grundlegend für das Verständnis wie auch die Beschreibung von Nutzeraktivitäten in klassischen Online Information Retrieval-Systemen und münden im heutigen Verständnis des Begriffs der *Suchstrategie*.

Der Terminus Suchstrategie wird heute informationswissenschaftlich vor allem im Kontext von Informationsverhalten verwendet. Das war nicht immer so: historisch bedingt findet sich der Begriff bereits zur Beschreibung der Formulierung von Suchanfragen auf Booleschen Datenbanken. Er wird, wie sich im Verlauf der Arbeit noch zeigen wird, noch heute in diesem Sinne fachterminologisch auch in der systematischen Übersichtsarbeit verwendet und soll daher historisch eingeordnet und abgegrenzt werden.

2.4.1.1 Suchstrategie als Repräsentation des Information Need

Eine der ersten expliziten Nennungen des Begriffs der Suchstrategie stammt aus dem Jahr 1973. Ein gutes Verständnis für den Informationsbedarf des Auftraggebers im Sinne des Taylorschen Stufenmodells, vorgestellt in Abschnitt 2.1.2, führe letztendlich zu besseren „Suchstrategien“ (Jewell und Brandhorst 1973), also einer möglichst exakten formalen Repräsentation des Informationsbedarfs in Form einer Booleschen Query – damals Standardtechnologie für datenbankgesteuerte Fachrecherchen.

Mit der Etablierung erster Online-Datenbanken standen dem Nutzer, in diesem Fall in erster Linie Rechercheexperte, plötzlich neue Möglichkeiten zur Verfügung: *„Innumerable alternatives are now available to the searcher, and a great variety of search methods can be employed....“* schreiben Oldroyd und Citroen 1977 (Oldroyd und Citroen 1977).

Aufgrund der zunehmenden Onlineverfügbarkeit von Datenbanken begann man bei ERIC in den späten 1970er Jahren mit Schulungsmaßnahmen, die vor allem darauf abzielten, den Nutzern eine qualitative und quantitative Bewertung ihrer Booleschen Suchanfragen anhand einer „perfekten Suchanfrage“ zu ermöglichen. Hierfür hatten Forscher an der Universität in Berkeley für 29 verschiedene Fragestellungen aus dem Index der ERIC Datenbank zunächst ideale Antwortmengen aus dem Datenbestand ermittelt. Außerdem hatten sie für jede Fragestellung verschiedene Suchanfragevariationen entwickelt, die entweder zu einem hohen Recall der Antwortmenge, zu einer hohen Precision der Query in Bezug auf die Antwortmenge oder zu möglichst geringen Kosten (Anm.: wohl mit ausgewogenem Verhältnis von Precision und Recall) führten. Die Zusammenführung aller Antwortmengen wurde als „perfekte Antwortmenge“ einer Fragestellung definiert, die Suchanfragevariationen als *Suchstrategien* bezeichnet (Markey und Cochrane 1978).

So bekommt die Suchstrategie eine weitere Qualität. Sie ist nicht mehr nur Formalisierung einer wissenschaftlichen Forschungsfrage in Form einer Booleschen Datenbank-Query (im Sinne einer Strategie, sich der Fragestellung zu nähern), sondern eine Query, die in Bezug auf eine bekannte Referenzmenge hinsichtlich vorgegebener Qualitätsmaße optimiert wurde.

2.4.1.2 Suchstrategie als Informationsverhalten

Etymologisch lässt sich der Begriff Strategie ganz klar im militärischen Kontext verorten, englischsprachig ist er vor allem im 19. Jahrhundert verwurzelt („*The art of a commander-in-chief; the art of projecting and directing the larger military movements and operations of a campaign*“). Später wird der Begriff “[...] *a plan for successful action based on the rationality and interdependence of the moves [...]*“ (Oxford University Press 2016g) zunehmend verallgemeinert.

Bates definierte den damals fachterminologisch weit verbreiteten Begriff der Suchstrategie aus informationswissenschaftlicher Sicht neu und führte zusätzlich den Begriff der Suchtaktik ein. Sie leitete hierfür begrifflich Strategien wie auch Taktiken aus dem oben genannten militärischen Kontext ab, wo Strategien mit der Gesamtplanung assoziiert sind, Taktiken hingegen mit kleinen Manövern zur Unterstützung dieser Strategien (Bates 1979b). Die von Bates definierten Begriffe beziehen sich dabei historisch ausschließlich auf klassische Suchaufgaben in online verfügbaren Booleschen Datenbanken.

Bates definiert drei Level an Suchaktivitäten, die letztendlich die Suchstrategie bestimmen:

Tabelle 2-5 Klassifikation von Suchaktivitäten nach Bates (Bates 1990, 1979b)

Suchaktivitäten	Beschreibung
Move	eine einzelne identifizierbare Aktion im Rahmen der Suche
Taktik	eine Abfolge einzelner Aktionen (= Moves), welche die Suche voranbringen
Stratagem	eine komplexe, Informationsquellen-spezifische Kombination oder Abfolge an Taktiken oder Moves
Strategie	Plan der gesamten Suche und Suchstrategie, die wiederum aus Moves, Taktiken und Stratagemen besteht,

	als Forschungsgegenstand auch die Untersuchung von Theorien, Prinzipien und Praxis in der Anwendung von Suchtaktiken, Stratagemen und Strategien.
--	---

Terminologisch sind die Begriffe Taktik, Stratagem und Strategie in der Literatur nicht immer klar voneinander abzugrenzen. So benennt selbst Bates einige Konzepte in ihren Veröffentlichungen mal als Taktik, mal als Stratagem (Bates 1989, 1990), anderswo wird Stratagem auch mit Strategie gleichgesetzt (Xie 2000). Wichtiger als die Begrifflichkeiten scheint die hierarchische Einordnung und Beziehung dieser Konzepte sowie das Verständnis von Aktivitäten und Aktionen, die in bestimmter Reihenfolge und in bestimmtem Kontext dazu eingesetzt werden können, Taktiken umzusetzen, die im Rahmen einer vorab geplanten Strategie zur Erledigung von Aufgaben führen sollen.

2.4.2 Das Bates-Modell der Information Search Tactics

Mit Hilfe ihrer Klassifikation definiert Bates 29 verschiedene – nicht empirisch untersuchte – Taktiken zur Suche in Online-Datenbanken und liefert damit ein sehr granulares, theoretisches Modell zur Beschreibung von Aktivitäten im Rahmen klassischer Tätigkeiten bei der Suche auf Booleschen Datenbanken. Bates beleuchtet mit ihrem Modell nicht nur die Query an sich, sondern auch den Umgang mit Datenbankstrukturen und vorbestehenden Trefferlisten. Auf eine Übersetzung der Benennungen wird verzichtet, da diese in der Fachliteratur meist wie von Bates benannt referenziert werden.

Monitoring-Taktiken sind solche, die der Kontrolle, Bewertung und Validierung von Queries und Suchprozessen dienen

Tabelle 2-6 Monitoring-Taktiken nach Bates (Bates 1979b)

Taktik	Beschreibung
M1. CHECK	Kontextuelle Überprüfung und Vergleich einer aktuellen mit der initialen Suchanfrage
M2. WEIGHT	Durchführung regelmäßiger Kosten,- Nutzenkontrolle bei Strategieänderungen der Suchanfrage

M3. PATTERN	Analyse, Veränderung und Aktualisierung von Suchmustern, beispielsweise zur Erhöhung der Effizienz
M4. CORRECT	Kontrolle von Suchanfragen im Hinblick auf syntaktische Fehler oder Schreibfehler
M5. RECORD	Dokumentation durchgeführter und bisher nicht durchgeführter Suchpfade

File Structure-Taktiken beinhalten alle Techniken, die sich die Datenstrukturen einzelner Datenbanken zu Nutze machen, um eine Suche voranzutreiben.

Tabelle 2-7 Datenbankspezifische Taktiken nach Bates (Bates 1979b)

Taktik	Beschreibung
F1. BIBBLE	Aufsuchen einer bereits vorbestehenden Bibliographie passend zu einem aktuellen Informationsbedarf
F2. SELECT	Aufbrechen komplexer Suchanfragen in Teilprobleme zu deren Einzelbetrachtung
F3. SURVEY	Überprüfung aktuell möglicher möglicher Entscheidungswege zu jedem Zeitpunkt des Suchprozesses
F4. CUT	Auswahl einer Suchoption, welche eine vorgegebene Domäne optimal disjunkt zerlegt, in dem sie den größten Teil der Domäne ausschließt
F5. STRETCH	Nutzung einer Ressource in anderer Form als intendiert
F6. SCAFFOLD	Nutzung indirekter Hilfspfade zum Auffinden gesuchter Information
F7. CLEAVE	Binäre Suche nach einem bestimmten Objekt in geordneten Strukturen

Taktiken zur Suchformulierung unterstützen den Aufbau oder die Neustrukturierung von Datenbank Queries.

Tabelle 2-8 Taktiken zur Suchformulierung nach Bates (Bates 1979b)

Taktik	Beschreibung
S1. SPECIFY	Suche mit spezifischen Termen
S2. EXHAUST	Ausbau der Treffermenge durch Hinzunahme weiterer möglicher Suchterme zu einer bestehenden Query
S3. REDUCE	Reduktion der Treffermenge durch Herausnahme einzelner Suchterme aus einer bestehenden Query
S4. PARALLEL	Breitere Suche durch Hinzunahme von Synonymen zu einer Query
S5. PINPOINT	Präzision der Suchanfrage durch Minimierung von Termen und Synonymen, Beschränkung auf kontrolliertes Vokabular
S6. BLOCK	Expliziter Ausschluss einzelner Terme

Term-Taktiken unterstützen die Auswahl und Formulierung spezifischer Terme in einer Query.

Tabelle 2-9 Term-Taktiken nach Bates (Bates 1979b)

Taktik	Beschreibung
T1. SUPER	Verwendung eines Terms, der hierarchisch breiter angelegt ist
T2. SUB	Verwendung eines Terms, der hierarchisch spezifischer angelegt ist
T3. RELATE	Verwendung eines verwandten Terms in gleicher Hierarchieebene
T4. NEIGHBOR	Verwendung ähnlicher Terme, beispielsweise alphabetische Nähe
T5. TRACE	Nutzung bereits verfügbarer Information zum Auffinden weiterer Terme
T6. VARY	Variation oder Ersatz einzelner Suchterme – nicht näher spezifiziert
T7. FIX	Verwendung alternativer Präfixe, Suffixe oder Infixe
T8. REARRANGE	Variation der Reihenfolge von Wörtern in Phrasen

T9. CONTRARY	Suche mit Termen entgegengesetzter Bedeutung
T10. RESPELL	Verwendung anderer Schreibweisen für einzelne Terme
T11. RESPACE	Verwendung von Abstandsoperatoren

Die von Bates ermittelten Taktiken bilden ein statisches und theoretisches Modell zur Beschreibung möglicher Aktivitäten im Rahmen Boolescher Suchen auf bibliographischen Datenbanken. Praktisch wurden einige dieser Taktiken in den Arbeiten von Fidel experimentell im Suchverhalten professioneller Informationsspezialisten mit vorwiegend medizinischem Arbeitsumfeld beobachtet. Sie werden dort als Moves im Sinne kleiner Aktionen bezeichnet, die taktisch eingesetzt werden, um die Ergebnisse einer Suche zu verändern (Fidel 1985, 1991c).

2.4.3 Operative und konzeptuelle Moves

Die empirisch belegten Beobachtungsstudien von Fidel bilden eine wichtige Grundlage zum Verständnis konkreter Aktionen bei der Formulierung von Queries in bibliographischen Online-Datenbanken, sie wurden zwischen 1984 und 1991 veröffentlicht (Fidel 1984, 1991a, 1991b, 1991c).

Laut Fidel lassen sich Moves aus drei Problemkategorien ableiten und entsprechend verorten, die aus empirischen Untersuchungen mit Datenbanknutzern identifiziert werden konnten (Fidel 1991c):

- vom Nutzer als zu klein beurteilte Treffermengen, welche zu Maßnahmen führen, um diese zu vergrößern
- von Nutzer als zu groß beurteilte Treffermengen, welche zu Maßnahmen führen, um diese zu verkleinern
- vom Nutzer als „off-target“ beurteilte Treffermengen, die zu Maßnahmen führen, mit denen ein ausgewogenes Verhältnis zwischen Sensitivität und Spezifität wiederhergestellt werden soll

Moves zur Expansion von Treffermengen dienen dabei in erster Linie der Verbesserung der Sensitivität der Suchanfrage, d.h. einer Vergrößerung der Treffermenge. Moves zur Reduktion

von Treffermengen dienen jedoch nicht ausschließlich zur Verbesserung der Spezifität einer Suchanfrage, sondern häufig einfach zur Verkleinerung der Treffermenge (Fidel 1991c).

Anmerkung: Fidel benennt in ihrem Artikel die hier genannte Sensitivität mit Recall, die Spezifität mit Precision. Im Folgenden soll Spezifität für nutzerinitiierte Änderungen an Suchanfragen verwendet werden, die eine Erhöhung des Recalls zur Folge oder zumindest zum Ziel haben, analog wird Spezifität für nutzerinitiierte Änderungen an Suchanfragen verwendet, die eine Erhöhung der Precision zur Folge oder zum Ziel haben. Diese Definition hat zum Ziel, die in Abschnitt 2.3.2.5 definierten Evaluierungsparameter Precision und Recall ausschließlich systemseitig zu verstehen.

Gleichzeitig klassifiziert Fidel operative und konzeptionelle Moves.

Die operationale Suche ist charakterisiert durch Suchaktivitäten, die einzig auf Funktionalitäten des Retrieval-Systems aufbauen und die konzeptuelle Bedeutung von Treffermengen nicht modifizieren, während konzeptuelle Suchprozesse durchaus auch die Bedeutung von Treffermengen verändern können. Ein operationaler Move klassifiziert beispielsweise einen Thesaurus Deskriptor als „Main Heading“ oder variiert eine Suche durch Synonyme oder Schreibweisen. Operationale Moves verengen, verbreitern oder spezifizieren eine Suche also und vergrößern beziehungsweise reduzieren damit die Treffermenge. Konzeptionelle Moves hingegen verwenden beispielsweise verwandte oder hierarchisch über- und untergeordnete Deskriptoren oder Freitextterme und können damit das Suchergebnis auch semantisch verändern. (Fidel 1984)

Somit ergibt sich eine Klassifizierung von Aktivitäten zur Modifikation von Datenbank Queries im Kontext verschiedener Problemkategorien wie in Tabelle 2-10 und Tabelle 2-11 dargestellt.

Tabelle 2-10 Operative Moves zur Reduktion von Treffermengen nach (Fidel 1991)

Move	Beschreibung
Reduktion der Treffermenge	
Weight I	Klassifikation eines Deskriptors als Hauptdeskriptor
Weight II	Schnittmenge von Freitexttermen mit einem breit angelegten Deskriptor

Weight III	Einschränkung von Freitexttermen auf ein bestimmtes Datenbankfeld, z.B. Titel
Weight IV	Nutzung von PROXIMITY-Operatoren
Weight V	Einschränkung auf bestimmte Dokumenttypen
Negate	Nutzung des NOT-Operators
Limit I	Einschränkung einer Treffermenge auf bestimmte Sprachen
Limit II	Einschränkung einer Treffermenge auf ein bestimmtes Publikationsdatum
Limit III	Einschränkung auf bestimmte Dokumentenmenge, Journal oder Datenbank
Limit IV	Einschränkung auf Dokumentenbestände mit bestimmtem Titel
Cut	Übermittlung einer beliebigen Teilmenge einer Treffermenge

Tabelle 2-11 Operative Moves zur Expansion von Treffermengen nach (Fidel 1991)

Move	Beschreibung
Expansion der Treffermenge	
Add I	Hinzunahme von Synonymen und alternativen Schreibweisen
Add II	Hinzunahme von Deskriptoren und Freitexttermen
Add III	Hinzunahme von Termen, die in aufgefundenen relevanten Dokumenten ermittelt wurden
Add IV	Hinzunahme von Deskriptoren mit höheren Treffermengen
Add V	Wechsel zu einer anderen Datenbank
Include	Hinzunahme von in der Thesaurus Hierarchie darunterliegenden Deskriptoren
Cancel	Aufheben eventuell bereits gemachter Restriktionen
Refine	Ermittlung besserer Suchterme

Fidel konkretisiert mit ihrem Modell die Vorarbeiten von Bates. Sie unterscheidet dezidiert zwischen Manövern für die strukturierte Datenbanksuche mittels kontrolliertem Vokabular

sowie der Freitextsuche. In Summe stehen durch die Arbeiten von Bates und Fidel konkrete Methoden zur Verfügung, Prozesse zur Formulierung von Suchanfragen auf Booleschen Datenbanken nicht nur zu beschreiben, sondern auch in den Kontext systematischer Übersichtsarbeiten zu setzen.

2.4.4 Suchtaktiken in Online-Datenbanken

Nicht empirisch, sondern aus der Praxis sind die von Pauline Cochrane und Karen Markey vorgestellten Suchtaktiken für Online-Datenbanken zu bewerten, die im ERIC Online User Manual veröffentlicht wurden. Das Manual ist auch Quelle für die Beschreibung der folgenden Manöver (Markey und Cochrane 1978).

Building Blocks

Die so genannte Building Blocks Strategie lässt sich direkt aus dem Konzept des Booleschen Retrieval ableiten.

Die Aufgabe in der Entwicklung einer Suchstrategie wird bei ERIC definiert „*Analyzing the search topic into facets or concept groups. Planning approaches to search strategy for combining concepts of the topic.*“ (Markey und Cochrane 1978). Das Konzept dieser Topics oder auch Building Blocks bezeichnen die Autoren als „dominant“ – es ist also ein gängiger Ansatz der Suche auf Fachdatenbanken.

Bei der Building Blocks Strategie wird die Frage in Building Blocks, also Themenblöcke (engl. auch *facets*), zerlegt. Für jeden dieser Themenblöcke werden Textwörter, Synonyme, Phrasen und kontrolliertes Vokabular aus den Thesauri der zu durchsuchenden Datenbanken ermittelt.

A₁, A₂, A₃, ... Textwörter, Synonyme, Phrasen und kontrolliertes Vokabular zu Block A

B₁, B₂, B₃, ... Textwörter, Synonyme, Phrasen und kontrolliertes Vokabular zu Block B

C₁, C₂, C₃, ... Textwörter, Synonyme, Phrasen und kontrolliertes Vokabular zu Block C

Diese müssen über den Booleschen Operator OR miteinander verknüpft werden, da sie als alternative Suchterme fungieren.

$(A_1 \text{ OR } A_2 \text{ OR } \dots A_N)$

$(B_1 \text{ OR } B_2 \text{ OR } \dots B_N)$

$(C_1 \text{ OR } C_2 \text{ OR } \dots C_N)$

Die Suche in den Datenbanken erfolgt pro individuelm Block, führt also zu einer Reihe an Queries, über die anschließend mit dem AND Operator geschnitten wurden. Bei der Durchführung resultiert dies natürlich zunächst zu sehr großen Treffermengen, die mit jeder Hinzunahme eines weiteren Blocks und den Operator AND verkleinert werden.

Klassischerweise folgt das fertiggestellte Suchprofil also vereinfacht der Formel

$(A_1 \text{ OR } A_2 \text{ OR } \dots A_N) \text{ AND } (B_1 \text{ OR } B_2 \text{ OR } \dots B_N) \text{ AND } (C_1 \text{ OR } C_2 \text{ OR } \dots C_N)$.

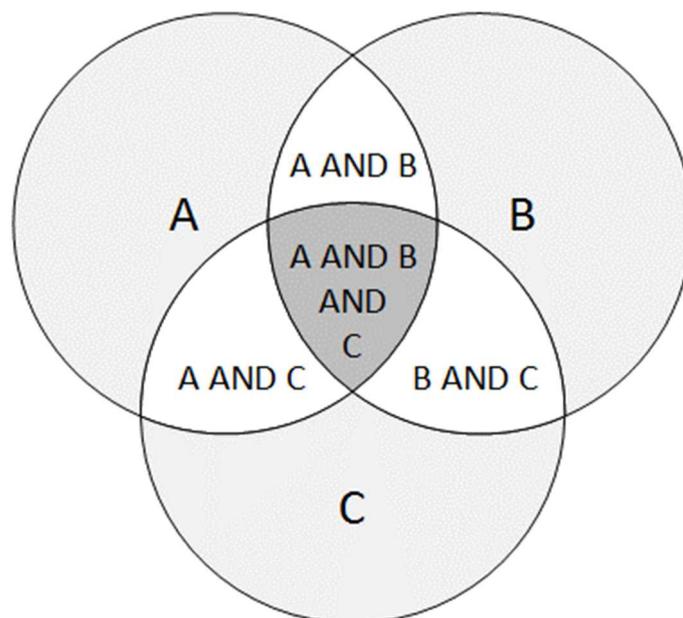


Abbildung 2-19 Building Blocks

Die Building Block-Strategie ist unter Rechercheexperten nach wie vor eine verbreitete Strategie zur Formulierung von Suchanfragen, in der Praxis zeigt sich jedoch, dass Endnutzer wie

beispielsweise Domänenexperten, die natürlich heute auch Zugang zu Fachdatenbanken haben, weiterhin großen Klärungsbedarf in der Nutzung auch dieser Strategie haben, siehe hierzu (Hertzberg und Rudner 1999).

Building Blocks haben eine große Ähnlichkeit mit der späteren Booleschen Suche, erfordern aber Arbeitsgedächtnis und – ein Faktor, der heute keine Bedeutung mehr hat, aber damals auch Kostengründen durchaus Bedeutung hatte – auch längere Online Zeit für die Formulierung der Suchanfrage, vgl. insgesamt (Markey und Cochrane 1978).

Citation Pearl Growing

Der Ansatz Citation Pearl Growing wird bei Markey und Cochrane als interaktive Methode der Suche beschrieben. Die Metapher des Perlenzüchtens findet sich in der Vorgehensweise beim Pearl Growing wieder: das Ziel ist es, über eine spezifische Suchanfrage zunächst mindestens eine relevante bibliographische Referenz aufzufinden. Aus dieser Referenz werden alle für die Suchanfrage relevanten Indexterme und Freitextterme notiert und in eine neue Suchanfrage eingearbeitet. Ab da wiederholt sich der Prozess iterativ so lange, bis keine neuen relevanten Suchergebnisse mehr aufgefunden werden können. Citation Pearl Growing ist durch die Entwicklung dynamischer Suchanfragen charakterisiert. (Markey und Cochrane 1978)

Successive Fractions

Beim Ansatz Successive Fractions, auch „*divide and conquer*“ oder „*file partitioning*“ (Markey und Cochrane 1978) genannt, fokussiert sich der Suchende zunächst auf den thematisch wesentlichen Building Block. Im Folgenden werden iterative alle weiteren Building Blocks mittels des AND Operators ergänzt und sofort evaluiert:

$(A_1 \text{ OR } A_2 \text{ OR } \dots \text{ OR } A_N)$ - wesentlicher Building Block

$(B_1 \text{ OR } B_2 \text{ OR } \dots \text{ OR } B_N)$ - weiterer Building Block

$(C_1 \text{ OR } C_2 \text{ OR } \dots \text{ OR } C_N)$ - weiterer Building Block

Auf die Suchanfrage $(A_1 \text{ OR } A_2 \text{ OR } \dots A_N) \text{ AND } (B_1 \text{ OR } B_2 \text{ OR } \dots B_N)$ folgt demnach die direkte Bewertung, erst dann wird die Suchanfrage zu $(A_1 \text{ OR } A_2 \text{ OR } \dots A_N) \text{ AND } (B_1 \text{ OR } B_2 \text{ OR } \dots B_N) \text{ AND } (C_1 \text{ OR } C_2 \text{ OR } \dots C_N)$ ergänzt.

Mit dem Successive Fractions-Ansatz arbeiten sich die Suchenden iterativ durch sich mit jedem Schritt reduzierende Treffermengen. Die Suche kann jederzeit terminiert werden, gleichzeitig ist die Kontrolle über den Verlust relevanter Artikel wie auch die Größe der Treffermenge besser als beim Building Block-Ansatz (Markey und Cochrane 1978).

Most Specific Facet First

Bei der Most Specific Facet First Strategie beginnt die Suche einer aus mehreren Building Blocks bestehenden Suchanfrage mit dem spezifischsten Building Block. Markey und Cochrane beschreiben diesen Ansatz als geeignet für Fragestellungen, bei denen gerade die spezifischen Aspekte durch unzureichende Indexierung über andere Strategien eventuell übersehen werden könnten (Markey und Cochrane 1978).

Lowest Postings Facet First

Beim Lowest Postings Facet First Ansatz gehen Nutzer von dem Building Block aus, für den sie die geringste Treffermenge erwarten. Ist diese Treffermenge überschaubar oder wie erwartet klein, so gibt es strategisch keinen Grund, durch die Hinzunahme weiterer Building Blocks die Treffermenge von Natur aus weiter zu verkleinern (Markey und Cochrane 1978).

Die von Markey und Cochrane beschriebenen Strategien sind grundlegend. Im weiteren Verlauf der Arbeit zeigt sich, dass sie trotz ihrer fehlenden Empirie in der professionellen Literaturrecherche im Rahmen systematischer Übersichtsarbeit auch heute noch eine wichtige Rolle spielen, siehe beispielsweise (Ramer 2005; Schlosser et al. 2006; Booth 2008).

2.4.5 Zusammenfassung

Suchstrategien und Taktiken sind insbesondere bei der Bedienung von auf Booleschem Prinzip basierenden Information Retrieval-Systemen für den Nutzer von enormem Wert. Sie dienen der effizienten Suche von Fachinformationen und garantieren ihm bestmögliche Kontrolle des

Suchprozesses in klassischen Online-Datenbanken. Suchtaktiken speziell auf Booleschen Datenbanken bis hin zu fragmentarischen Moves wurden vor allem in den 1980er und 1990er Jahren aus informationswissenschaftlicher Sicht intensiv untersucht und liefern auch heute noch eine Grundlage zur Beschreibung des Informationsverhaltens in klassischen Suchprozessen, dem *information searching behaviour*.

2.5 Information Retrieval Interaktion im Kontext

Die bisherigen Abschnitte beleuchteten die Themen Informationsbedarf und Informationsverhalten. Sie ordneten zudem Boolesche Retrieval-Systeme im Kontext systemspezifischer Implementierungen ein und gaben einen Überblick über die Charakteristika dieser Systeme und dem Informationsverhalten, das in diesem Kontext zu effizienten Retrieval-Strategien führt. Bereits in Abschnitt 2.2 wurde deutlich, dass Informationsverhalten nicht isoliert betrachtet werden kann. Dies wurde 2004 und 2005 von Järvelin und Ingwersen mit dem Konzept der Interaktion mit Information Retrieval im Kontext konkretisiert (Järvelin und Ingwersen 2004; Ingwersen und Järvelin 2005).

Das Ziel ist die Entwicklung von Systemen, welche die Nutzer beim Erreichen ihrer Suchaufgaben unterstützen sollen – Information Retrieval als Interaktion mit Systemen, betrachtet im Kontext verschiedenster Rahmenbedingungen.

2.5.1 Interaktion: Nutzer und System

Interaktion beschreibt in unserem Sprachgebrauch „aufeinander bezogenes Handeln zweier oder mehrerer Personen; Wechselbeziehung zwischen Handlungspartnern“ (Bibliographisches Institut GmbH Duden Verlag 2016k), im englischen Allgemeiner „*reciprocal action; action or influence of persons or things on each other*“ (Oxford University Press 2016d), in diesem Fall also beispielsweise auch die Interaktion zwischen dem Nutzer und einem Information Retrieval-System mittels geeigneter Bedienoberflächen.

Aufgabe der Bedienoberfläche (*Interface*) von Information Retrieval-Systemen ist es, die Nutzer dabei zu unterstützen, ihren Informationsbedarf zu formulieren, ihn in der passenden Anfragesprache zu formalisieren, Antwortmengen zu analysieren und zu verstehen sowie den

Fortschritt des gesamten Suchprozesses zu unterstützen. Zentrale Elemente der Nutzeroberfläche sind dabei zunächst ein Texteingabefeld zur Formulierung der Suchanfrage und ein Button zum Absenden dieser Suchanfrage. Trefferlisten werden meist als vertikale Listen implementiert. (Hearst 2009, Kap. 1.1)

Dieses sehr einfach gehaltene Konstrukt kann beispielsweise durch folgende Aspekte erklärt werden: Suchprozesse sind häufig Teil einer umfassenderen Aufgabe, sodass Nutzer nicht nur komplexe Bedienoberflächen gestört werden sollten. Gleichzeitig gehören Suchprozesse zu den kognitiv komplexen Aufgaben, was eine Fokussierung auf diese Aufgabe erfordert. Nicht zuletzt müssen Bedienoberflächen wie auch die Anzeige von Treffermengen vom Suchenden auch intellektuell verstanden werden können. (Hearst 2009 Kap. 1.1)

Vor allem vor dem Hintergrund Boolescher Retrieval-Systeme muss man sich bewusst machen, dass die Entwicklung von User-Interfaces vor allem durch die weltweite Verfügbarkeit des Web und Suchmaschinen einen radikalen Wandel vollzogen hat, vgl. (Hearst 2009 Kap. 1.2): Bereits in Abschnitt 2.4.1 wurde deutlich, dass Information Retrieval vor dieser Zeit lange eine Aufgabe einer kleinen, sehr spezialisierten Nutzergruppe war und sich Suchprozesse vor allem auf Bibliotheken, Gesetzestexte und ähnliche Inhalte reduzierten. Suchanfragen waren entsprechend komplex, die Recherche war ausgebildeten Informationsexperten vorbehalten.

Die Analyse und Beschreibung von Bedienoberflächen umfasst systemseitig verschiedene Dimensionen, vgl. hierzu (Ingwersen und Järvelin 2005, S. 135).

Eine Dimension betrifft die für die Interaktion erforderlichen „**Objekte**“.

Hierzu gehören wie bereits genannt ein Texteingabefeld, ein Suchbutton als auch die Darstellung von Treffermengen, die beispielsweise listenweise oder aber auch graphisch denkbar ist, vgl. hierzu (Nowell et al. 1996). Zusätzlich fallen darunter die Darstellung des gesamten Dokumentenbestands, die Darstellung seiner thematischen Organisation, beispielsweise durch Thesauri oder eines einfachen Index, aber auch die Darstellung einzelner Dokumente. (Ingwersen und Järvelin 2005, S. 135)

Eine weitere Dimension betrifft die dem Nutzer verfügbare **Funktionalität**, ein Aspekt, der aus zwei Seiten betrachtet werden kann. Zum einen betrifft Funktionalität die Möglichkeiten

zur Interaktion des Nutzers mit dem System, zum anderen betrifft sie aber auch Funktionalitäten, die einzelne Stufen des Retrieval-Prozesses unterstützen (Ingwersen und Järvelin 2005, S. 136).

Bates (Bates 1990) definiert verschiedene Stufen der Nutzer-System-Interaktion:

- Level 0 überlässt alle Suchaktivitäten dem Nutzer
- Level 1 interagiert mit dem User auf Anforderung, zeigt beispielsweise mögliche Aktivitäten an
- Level 2 führt Aktivitäten auf Nutzeranforderung hin aus
- Level 3 zeichnet Suchprozesse auf und spricht optional Empfehlungen aus
- Level 4 führt Aktivitäten automatisch durch, wobei der Nutzer darüber optional informiert werden kann.

Nach der Idee von Bates unterstützen die Bedienoberflächen nutzerfreundlicher Information Retrieval-Systeme daher verschiedene Suchtechniken, sie kritisiert Systemarchitekturen, welche dem Nutzer durch Vollautomatisierung einzelner Prozesse die Kontrolle über den Suchprozess nehmen. Gleichzeitig jedoch macht sie deutlich, dass beispielsweise Boolesche Systeme wie bibliographische Datenbanken bis dato lediglich Level 1 Interaktionsmodi bedienen, statt beispielsweise in Level 1 oder Level 2 auch Suchstrategien wie Berry Picking abzudecken (Bates 1990).

Die letzte Dimension ist die **Human-Computer-Interaction**, die benutzerfreundliche Gestaltung von Interaktionsmöglichkeiten. Da dies nicht Fokus der vorliegenden Arbeit ist, sei hier auf (Hearst 2009, Kap 1.1) sowie die Arbeiten von Shneiderman et al. (Shneiderman et al. 1997) verwiesen.

2.5.2 Grundlagenmodell nach Ingwersen und Järvelin

Mit dem Ebenen-Framework *Information seeking and retrieval design and evaluation*, siehe (Järvelin und Ingwersen 2004; Ingwersen und Järvelin 2005) wird 2004 ein neuer theoretischer Rahmen vorgestellt. Dieses Framework erweitert bisher bestehende Modelle und Theorien (wie beispielsweise Abschnitte 2.2.1- 2.2.5) um den Aspekt Kontext und definiert damit neue Forschungsaspekte zur Gestaltung nutzerfreundlicher Information Retrieval-Systeme.

Information Retrieval-Forschung und *information behaviour*-Forschung werden ausgehend von diesem Modell als *information retrieval interaction* unter einem Dach vereint, siehe Abbildung 2-20.

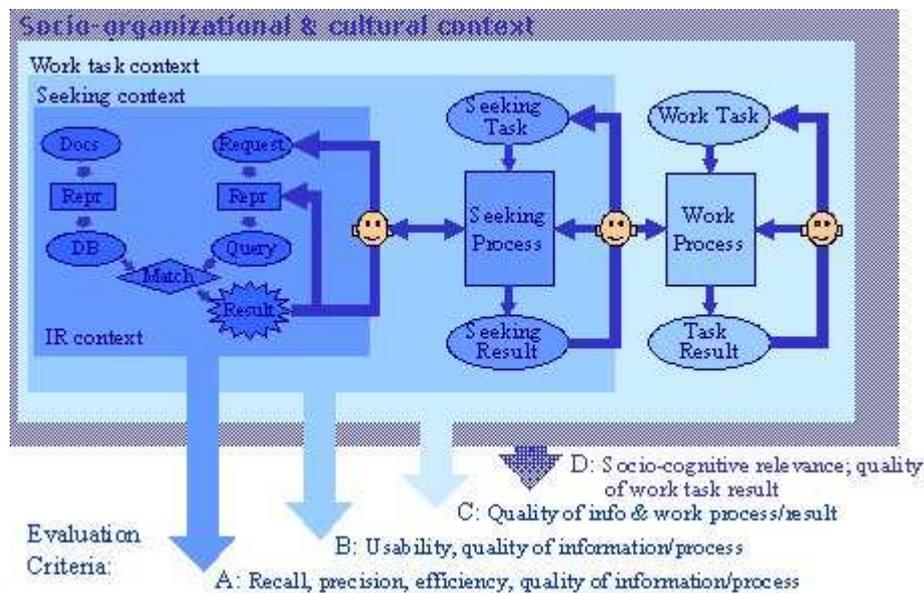


Abbildung 2-20 Das Ingwersen und Järvelin Framework, aus (Järvelin und Ingwersen 2004)

Die innerste Ebene, der so genannte *IR context* betrifft das Design von Information Retrieval-Systemen an sich: Es umfasst auf Nutzerseite konkrete Suchprozesse, auf Systemseite Dokumentrepräsentation, Suchanfrage und Retrieval-Matching. Klassische Evaluationsmaße sind daher, sofern diese Ebene aus dem Kontext betrachtet wird, Precision und Recall.

Information Retrieval wird als Teil eines *information seeking behaviour* in einem *information seeking context* verstanden. Dieser Kontext umfasst verfügbare Informationsquellen und Informationssysteme. Design- und Bewertungskriterien sind demnach Faktoren wie Usability oder die Qualität von Suchprozessen und deren Bewertung durch den Akteur, den Nutzer. Information Seeking wiederum resultiert aus Arbeitsaufgaben und ist daher im *work context* zu betrachten. Über allem stehen soziokulturelle und organisatorische Faktoren als *socio-organizational & cultural context*. Wesentliches Element des Modells jedoch ist seine Verschachtelung: Kontextfaktoren äußerer Ebenen bestimmen immer auch Informationsverhalten der weiter innen gelegenen Ebenen – Informationsverhalten kann daher nie isoliert betrachtet werden. (Järvelin und Ingwersen 2004)

Erklärt werden kann das neue Forschungsparadigma, das sich heute in der Forschung an verschiedenen Stellen etabliert hat (vgl. hierzu beispielhaft die Gesamtwerke (Ingwersen und Järvelin 2005; Russell-Rose und Tate 2013)) anhand des Information Retrieval-Prozessmodells von Ingwersen und seiner Komponenten – dem Gesamtkontext, dargestellt in Abbildung 2-21.

Information Retrieval-Prozesse werden initiiert von **Arbeitsaufgaben**. Diese führen beim Nutzer zu einer kognitiven Wahrnehmung des vorliegenden Problems, der zu erreichenden Ziele sowie einer individuellen Einschätzung des vorherrschenden Informationsbedarfs, der wie in Abschnitt 2.1.2 dargestellt, auch durch Unsicherheit gekennzeichnet werden kann. Dies mündet letztendlich in einem „Informationsverhalten“ (vgl. hierzu Abschnitt 2.2) zur Abarbeitung von Aufgaben und damit auch in einer Interaktion mit Informationssystemen. Die **Interaktion** erfolgt mit einem System, das aufgrund der Implementierung von Information Retrieval-Techniken, Datenbankstrukturen als auch Indexierungsverfahren bestimmten Regeln folgt, vgl. hierzu auch Abschnitt 2.3.2. In diesem Information Retrieval-System werden **Informationsobjekte** vorgehalten, deren Eigenschaften wiederum nicht nur das Information Retrieval-System technisch begründen und erklären, sondern auch das Nutzerverhalten beeinflussen, vgl. hierzu (Ingwersen 1996).

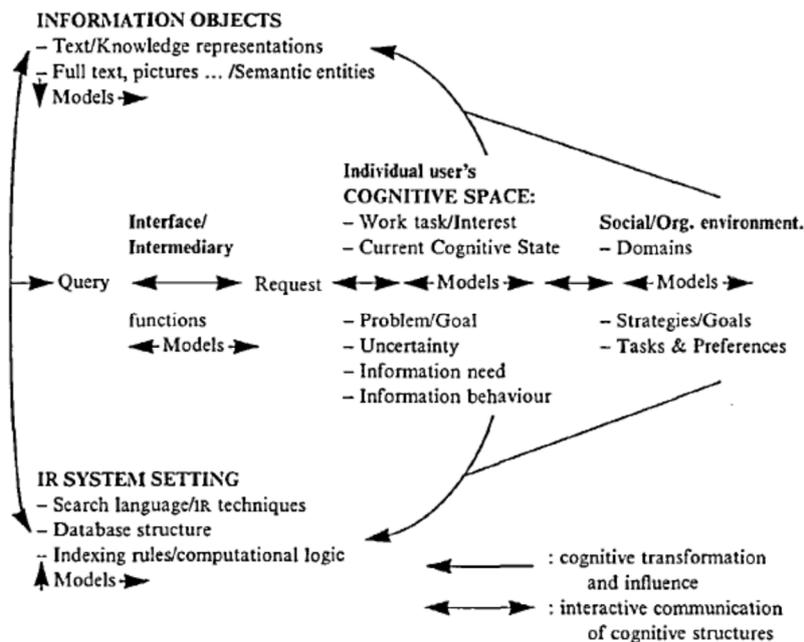


Abbildung 2-21 Modell des Information Retrieval Prozesses aus (Ingwersen 1996, S. 9)

Mittelpunkt beider Modelle ist die Arbeitsaufgabe als Initiator eines Informationsverhaltens, was zur Überlegung führt, dass Informationsverhalten beispielsweise anhand von Aufgaben charakterisiert werden kann, vgl. neben Abschnitt 2.2.4 auch (Byström 1999, 2002; Li und Belkin 2008). Da die systematische Übersichtsarbeit als aufgabeninitiiert verstanden werden muss, ist dieser Aspekt für die vorliegende Arbeit von besonderem Interesse.

2.5.3 Informationsverhalten und Aufgaben

Eine Aufgabe (engl. *task*) wird allgemein verstanden als „etwas, was jemandem zu tun aufgegeben ist; Auftrag, Obliegenheit“ (Bibliographisches Institut GmbH Duden Verlag 2016b), im Englischen *“Any piece of work that has to be done“* (Oxford University Press 2016i) („ein Stück Arbeit, das erledigt werden muss“) und kann damit allgemein hin als zielgerichtete Aktivität verstanden werden.

Die Abhängigkeit von Informationsverhalten, Informationssuchprozessen und Arbeitsaufgabe ist in der Fachliteratur seit Jahrzehnten akzeptiert (vgl. Abschnitt 2.2.4): Aufgaben erfordern verfügbare Information, um sie erfolgreich erledigen zu können. Aus dieser Notwendigkeit kann sich demnach ein Informationsbedarf ergeben, der wiederum ein spezifisches, von verschiedenen Faktoren abhängiges Informationsverhalten auslöst.

In der Informationssuche ist mit dem Begriff Aufgabe meist ein mehrstufiger Prozess gemeint, der je nach Komplexität auch in Teilaufgaben zerlegt werden kann und dem Erreichen eines Ziels dient (Hansen 1999). Teilaufgaben wiederum können als Teilschritte oder eine Abfolge von Aktivitäten und Aktionen interpretiert werden. Im Allgemeinen haben Aufgaben ein klar definiertes Ziel, ein Ergebnis sowie einen Start- und Endpunkt. Diese Sichtweise ermöglicht die Analyse und Beschreibung von Aufgaben sowohl als statische Konstrukte durch verschiedene charakteristische Attribute, beispielsweise deren Komplexität (vgl. Abschnitt 2.2.4), als auch in Form von Prozessabläufen (vgl. Abschnitt 2.2.3). Hierbei können verschiedene Aspekte im Fokus stehen, beispielsweise die Reihenfolge der Abarbeitung von Aufgaben oder einzelne Schritte und Entscheidungen (Vakkari 2003).

In der Fachliteratur wird in einem mindestens zweistufigen Ebenen Modell zwischen sog. Arbeitsaufgaben (*work tasks*) und Suchaufgaben (*search tasks*) differenziert (Hansen 1999), später dann zwischen *seeking* und *searching tasks* unterschieden (vgl. u.a. (Byström 1999; Vakkari 2003; Byström und Hansen 2005)). Wie im Folgenden erkennbar ist, geht diese Sichtweise konform mit der Definition des *information behaviour*, *information seeking behaviour* und *information searching behaviour* Modells von Wilson (Wilson 1999).

Die Forschungsergebnisse und Ansätze zum Informationsverhalten im Kontext von Aufgaben sind wesentlich geprägt durch die Arbeiten von Byström und Hansen (Hansen 1999; Byström 1999; Byström und Hansen 2005) sowie Vakkari (Vakkari 2001, 2003).

Work Tasks

Die Arbeitsaufgabe kann als übergeordnete Aufgabenstellung verstanden werden, die nicht notwendigerweise klar definiert sein muss. Work Tasks werden in der Fachliteratur in unterschiedlichen Facetten klassifiziert, beispielsweise anhand ihrer Komplexität (vgl. 2.2.4), ihres Ziels, ihrer Struktur oder ihres Prozesses (vgl. Abschnitte 2.2.1-2.2.3).

Arbeitsaufgaben sind im Allgemeinen stark kontextuell geprägt und führen zu einem Nutzerverhalten, das sich extrem an Umgebungsvariablen wie dem Arbeiten anhand von Normen oder Regelwerken, aber auch an (in Unternehmen) verfügbaren Ressourcen orientiert. Gleichzeitig spielen aber auch situative Aspekte in die Erledigung von Aufgaben: Neben individuellen Faktoren wie dem Wissenstand, Erfahrung oder persönlicher Ambition, sind dies Ressourcen-bezogene Faktoren wie Zeit und Quellen-bezogene Faktoren wie die zur Verfügung stehenden technischen Möglichkeiten des Einzelnen. (Hansen 1999; Byström und Hansen 2005). Dieser Zusammenhang ist auch in Abbildung 2-20 sichtbar.

Seeking Tasks

Aus Arbeitsaufgaben, den *working tasks*, ergeben sich Aufgaben im Umgang mit Information (*information seeking tasks*), beispielsweise durch das Erkennen eines Informationsbedarfs während der Erledigung der übergeordneten Arbeitsaufgabe, vgl. hierzu die Abschnitte 2.2.4 und 2.2.5. Diese Teilaufgaben werden *immer* über einen Informationsbedarf getriggert, beispielsweise die Notwendigkeit, einen Fachbegriff in einem Lexikon nachschlagen zu müssen

oder verfügbare Informationen miteinander zu vergleichen. *Information seeking tasks* sind aber auch gekennzeichnet durch die zwangsläufige Entscheidung, ob einem erkannten Informationsbedarf notwendigerweise nachgegangen werden muss oder ob auf diesen Schritt verzichtet werden kann. Als Teil der Arbeitsaufgabe unterliegen auch die *information seeking tasks* Umgebungsvariablen wie oben genannt. (Byström und Hansen 2005)

Information seeking tasks sind wesentliche Bestandteile meist komplexer und informationsintensiver Arbeitsaufgaben. In der Literatur werden sie – wie auch *work tasks* in unterschiedlichen Bezeichnungen als meist dreistufige Prozesse definiert (Byström und Hansen 2005), beispielhaft wird dies an folgenden Forschungsergebnissen deutlich:

Tabelle 2-12 Working Tasks und Information Seeking Tasks: Prozessmodelle

Referenz	Stufe 1	Stufe 2	Stufe 3
Marchionini (Marchionini 1995)	Understanding	Plan and Execution	Evaluation and Use
(Vakkari 2001)	Pre-Focus	Formulation	Post-Focus
Byström/Hansen (Byström und Hansen 2005)	Task Construction	Task Performance	Task Completion

In Stufe 1, als Konstruktionsphase, Verständnis oder auch Pre-Focus bezeichnet, wird anhand eines zu entwickelnden Grundverständnisses für Rahmenbedingungen, Zielsetzungen und Aufgabenstellung die Abarbeitung der Aufgabe geplant. Wie umfangreich die Planungsphase in der Praxis ausfällt, ist abhängig von der individuellen Wahrnehmung der Aufgabe, beispielsweise wie schwierig und komplex diese empfunden wird (Byström und Hansen 2005).

Stufe 2 beschreibt die Phase, in der Aufgaben durch Aktivitäten und Aktionen abgearbeitet werden, was sich vor allem im Modell von Marchionini an den Bezeichnungen „*plan and execution*“ widerspiegelt. Hier fokussiert sich die Informationssuche auf ein sehr konkretes Ziel mit ebenso konkreten Handlungen. Im Gegensatz zur Konstruktionsphase lässt sich das Nutzerverhalten der Stufe 2 sehr gut beobachten (Byström und Hansen 2005).

Mit Stufe 3 ist die Aufgabe abgeschlossen (was allerdings nicht mit einem erfolgreichen Abschluss gleichzusetzen ist). Das Ergebnis kann anschließend bewertet werden, alternativ kann

die über die Suche ermittelte Informationen für die Erledigung von Arbeitsaufgaben genutzt werden (Byström und Hansen 2005).

Information seeking tasks und das daraus resultierenden Informationsverhalten werden thematisch in den in Abschnitt 2.2 vorgestellten Modellen zum Informationsverhalten behandelt. In das Prozessmodell des *information seeking tasks* der Tabelle 2-12 lassen sich direkt Elemente der in Abschnitt 2.2 vorgestellten Prozessmodelle von Kuhlthau und Marchionini verorten, vgl. hierzu auch (Hansen 1999; Byström und Hansen 2005).

Vakkari (Vakkari 2001) stellt aufgabeninitiiertes Informationsverhalten direkt dem im Abschnitt 2.2.2 vorgestellten kognitiven Modell von Kuhlthau gegenüber. Er identifiziert während des Gesamtprozesses Strategien und Taktiken im Informationsverhalten, die beispielsweise bereits von Bates und Fidel (siehe Abschnitte 2.4.2 und 2.4.3) beschrieben sind. Im Gegensatz zum Kuhlthau-Modell ist im Vakkari-Modell nicht der Nutzen eines Dokuments das ausschlaggebende Relevanzkriterium, sondern es sind vorab definierte Relevanz-Kriterien und die Precision wesentlich. (Vakkari 2001).

Die Modelle von Byström und Leckie (Abschnitte 2.2.4 und 2.2.5) widmen sich konkret den kontextuellen Parametern aufgabenbezogenen Informationsverhaltens.

Searching Tasks

Information seeking tasks wiederum führen zu einer Reihe von Teilaufgaben, zu denen auch die Suchaufgaben (*information searching tasks*) gehören. Die Suchaufgabe bezieht sich auf eine einzelne, spezifische Quelle, beispielsweise eine bibliographische Datenbank und ist daher direkt im Information Retrieval Kontext zu betrachten. *Information seeking tasks* können definitionsgemäß aus einzelnen sowie einer Reihe an Suchaufgaben bestehen, die Suchaufgabe entspricht dabei dem eigentlichen Suchen im Sinne der Eingabe einer Suchanfrage. (Byström und Hansen 2005)

Entsprechend gehören Suchaufgaben kontextuell zu den einzigen Aufgaben, die im Ingwersen und Järvelin-Modell auf Information Retrieval-Ebene anzusiedeln sind, siehe Abbildung 2-20. Sie sind insbesondere in den Modellen zum Informationsverhalten auf Online-Datenbanken eingehend beschrieben, beispielsweise bei Bates oder Fidel (Bates 1979a, 1979b,

1989; Fidel 1991a, 1991b, 1991c) und den Arbeiten von Markey und Cochrane (Markey und Cochrane 1978), siehe Abschnitt 2.4.

2.5.4 Kontextdimensionen

Bildungssprachlich steht „Kontext“ (lat. *contextere* = zusammenweben) (Stowasser et al. 2007) sowohl im Deutschen als auch Englischen für „Zusammenhang“ (Bibliographisches Institut GmbH 2016, Stichwort: Kontext), sprachwissenschaftlich wird Kontext auch als Menge der Elemente einer Kommunikationssituation definiert, die das Verständnis einer Äußerung bestimmen (Bußmann et al. 2008, Stichwort: Kontext). Im Verständnis des Informationsverhaltens findet sich „Kontext“ zunächst als Grundidee bei Wilson (vgl. Abschnitte 2.1.3 und 2.2.6).

Der Begriff Kontext stand informationswissenschaftlich immer wieder im Fokus der Information Behaviour-Forschung (u.a. (Dervin 1997; Talja et al. 1999; Johnson 2003; Courtright 2007)) und führte zu unterschiedlichen Facetten, vor allem:

- Kontext als Menge beteiligter Entitäten (Akteure, Strukturen und deren Eigenschaften) (Talja et al. 1999)
- Kontext als situative Interpretation des menschlichen Akteurs (Wilson 1981, 1997, 1999)

Im Information Searching- und Retrieval-Modell lässt sich die Definition des Kontexts im Sinne der Menge beteiligter Entitäten direkt an den Elementen des Information Retrieval Prozesses (Abbildung 2-21) festmachen, die als „Spielbälle“ des Prozesses fungieren. Sie haben eigene Strukturen und Charakteristika, stehen aber gleichzeitig miteinander in Verbindung und definieren so einen systemischen Kontext. Gleichzeitig steht der Mensch mit seinem Informationsbedarf als Akteur im Zentrum (Järvelin und Ingwersen 2004; Ingwersen und Järvelin 2005).

Arbeitsaufgabe und Suchaufgabe

Die Kontextdimension Arbeitsaufgabe umfasst die *working tasks* wie auch *search tasks*. Wie bereits beschrieben sind Arbeitsaufgaben stark geprägt von soziokulturellen Faktoren und unterliegen domänenspezifischen Anforderungen. Die Durchführung von Aufgaben ist

sowohl abhängig von Unternehmenskulturen, gleichzeitig aber auch von Unternehmensstrukturen sowie organisationspezifischen Ressourcen und damit von Möglichkeiten, wie und ob Nutzer (in dem Fall Rechercheexperten) mit Systemen wie auch mit Menschen interagieren können. Anhand vorgegebener Arbeitsaufgaben können aus einem organisatorisch-unternehmerischen Kontext Suchaufgaben vordefiniert sein oder sich aus der Arbeitsaufgabe ergeben. In der professionell durchgeführten Suche sind Arbeitsaufgaben grundsätzlich zugewiesen, Suchaufgaben ergeben sich aus der Notwendigkeit von Arbeitsaufgaben, sodass es sich hier ausschließlich um extern motivierte Tätigkeiten handelt, vgl. (Tait 2014).

Ingwersen und Järvelin betonen die Notwendigkeit, folgende Charakteristika von Aufgaben in die Analyse einzubeziehen, einige davon wurden im Verlauf der Arbeit bereits angesprochen (Ingwersen und Järvelin 2005, S. 324):

- Struktur von Aufgaben
- Strategien und Praktiken im Umgang mit Aufgaben
- Komplexität von Aufgaben
- Abhängigkeiten zwischen einzelnen Aufgaben
- Anforderung zur Erledigung von Aufgaben
- Kontext von Aufgaben und deren Domäne

Vor allem *working tasks* und ihre Wahrnehmung durch den Akteur, so die Kritik des Ingwersen und Järvelin-Modells fanden in der traditionellen Information Retrieval-Forschung gar kein, in der Forschung um das Informationsverhalten nur unzureichende Beachtung (Ingwersen und Järvelin 2005, S. 314).

Akteur

Arbeitsaufgaben und Suchprozesse unterliegen individuellen, subjektiven Faktoren, beispielsweise deklarativem (*know what*) oder prozeduralem Wissen (*know how*). Deklaratives Wissen umfasst dabei Domänenexpertise als auch Wissen um Informationsquellen und ihre Strukturen. Prozedurales Wissen wird verstanden als das technische Wissen, beispielsweise die Bedienung einer Datenbank oder das Wissen um die Syntax von Datenbankabfragen, auf Domänenseite zusätzlich auch klassischen Problemlösungsstrategien, vgl. (Byström und Hansen 2005)).

Vorgegebene Arbeits- und Suchaufgaben führen daher beim Akteur zu subjektiv wahrgenommenen Arbeits- und Suchaufgaben wie auch zum Bewusstsein eines Informationsbedarfs: Typisch für komplexe Suchaufgaben ist ein häufig auftretender Informationsbedarf durch Lücken im deklarativen oder prozeduralen Wissen, siehe hierzu Abschnitt 2.1.2. Die Erkennung eines Informationsbedarfs wiederum führt klassischerweise nicht nur zur Deckung dieses Informationsbedarfs, sondern auch zu Entscheidungssituationen, in denen ermittelt wird, ob die Schließung von Informationslücken für die Durchführung von Aufgaben relevant ist (Byström und Hansen 2005). Daraus resultiert, dass sowohl individuelle Faktoren wie Kenntnisstand und Erfahrung von Nutzern als auch Systemparameter einzelner Quellen, deren Verfügbarkeit und technische Beschaffenheit die Durchführung von Aufgaben das Informationsverhalten wesentlich beeinflussen. Letztendlich spielt auch die Wahrnehmung von Arbeits- oder Suchaufgaben durch die Akteure eine Rolle. Sie resultiert aus der Summe oben genannter Parameter und führt dazu, dass Aufgaben als einfach oder komplex empfunden werden und entsprechend mehr oder weniger strukturiert durchgeführt werden können, siehe hierzu (Byström und Järvelin 1995) bzw. Abschnitt 2.2.4.

Zusammengefasst nennen Ingwersen und Järvelin folgende Variable, die in der Forschung Berücksichtigung finden sollten (Ingwersen und Järvelin 2005, S. 332):

- das Domänenwissen von Akteuren
- die Rechercheexpertise von Akteuren
- die Erfahrung von Akteuren im Umgang mit den zu bewältigenden Arbeits- und Suchaufgaben
- die Wahrnehmung des Akteurs bezüglich des aktuellen Stadiums in der Erledigung von Aufgaben
- die Wahrnehmung des Akteurs bezüglich seines Arbeitsumfelds
- problematische Situationen
- Motivation und Emotionen des Akteurs

Dokumente, Systeme und Interaktion

Gleichzeitig bildet die Menge systemspezifischer Elemente einen Kontext, der insbesondere für klassische Suchaufgaben wesentlich ist (*IR Kontext*). Hierzu gehören die Menge verfügba-

rer Informationen, das Retrieval-Modell sowie Funktionalität von Suchmaschinen und Datenbanken und im Speziellen deren Interfaces sowie die Möglichkeit zur Interaktion, vgl. (Ingwersen und Järvelin 2005, S. 340).

Zur vollständigen Charakterisierung von Abhängigkeiten schlagen Ingwersen und Järvelin auf Dokumentenebene unter anderem folgende Untersuchungsgegenstände vor (Ingwersen und Järvelin 2005, S. 340):

- Struktur von Dokumenten
- Publikationstypen
- Dokumenttypen
- Art der vorgehaltenen Information
- Kommunikative Funktionen von Dokumenten
- zeitliche Aspekte von Dokumenten
- Sprache, Stil und Layout von Dokumenten
- Hyperlink-Strukturen,

und auf Interaktionsebene zusätzlich (Ingwersen und Järvelin 2005, S. 351):

- die Dauer einer Interaktion
- alle Entitäten, die an einer Interaktion beteiligt sind
- die Art der Interaktion
- Interaktionsstrategien und -taktiken
- Zweck menschlicher Interaktion und Mensch-Maschine-Interaktion
- Interaktionsmodi
- Aufwand.

Die hier genannten Parameter erheben keinen Anspruch auf Vollständigkeit, insbesondere dürfen sie lediglich als Klassifikatoren gelten. Sie müssen in ihren Nuancen problemspezifisch im Detail betrachtet und hinsichtlich ihrer Relevanz beurteilt werden.

Dies ist Ziel der vorliegenden Arbeit:

Informationswissenschaftlich führen sowohl die isolierte Betrachtung von Information Retrieval-Systemen als auch die isolierte Betrachtung von Informationsverhalten zu einer sehr restriktiven Sichtweise auf ein komplexes Gesamtsystem, vgl. (Ingwersen und Järvelin 2005, S. 318). Information Retrieval-Interaktion im Kontext betrachtet hingegen Suchaufgaben und

Arbeitsaufgaben sowie das daraus resultierende Informationsverhalten im Kontext oben genannter Dimensionen und führt damit zu einem möglichst vollständigen Bild des Akteurs, seiner Bedürfnisse und Aktivitäten (siehe Abbildung 2-20).

2.6 Zusammenfassung

Das vorliegende Kapitel diene dem Überblick über die informationswissenschaftliche Forschung im Bereich des Information Retrieval und dem daraus resultierenden Konzept des *information behaviour*. Wie deutlich wurde, führt die Entwicklung von Information Retrieval-Systemen nach einer zunächst rein technischen Betrachtung in den 1980er Jahren zu einer Fokussierung auf menschliches Informationsverhalten und in der Folge zur Betrachtung von menschlicher Interaktion mit Information Retrieval-Systemen. Die heute in der Forschung anerkannte Betrachtung dieser Interaktion im Kontext aller beteiligten Entitäten (vom Nutzer bis zum System) bildet die theoretische Grundlage für die vorliegende Arbeit.

Die Sichtweise auf Informationsverhalten als aufgabeninitiiertes Phänomen, wie von Ingwersen, Järvelin und anderen propagiert, ist in der informationswissenschaftlichen Forschung vor allem für die professionelle Informationsrecherche von besonderem Interesse: Informationsintensive Aufgaben wie Auftragsrecherchen bedingen eine hochfrequente Nutzung von Information, Suchaufgaben sind dabei – wie sich im Folgenden zeigt – Routineaufgaben.

3 Die systematische Übersichtsarbeit

Das folgende Kapitel dient der Einordnung der systematischen Überblicksarbeit mit Fokus auf interventionellen Themen in der EBM (vgl. Abschnitt 1.1.3) in den Kontext des *Information Retrieval Interaction in Context*-Modells wie in Abschnitt 2.5 vorgestellt.

Zunächst beleuchtet das Kapitel Informationen und Informationsressourcen:

Evidenzquellen – jegliche Form von Information, die evidenzbasierte Ergebnisse aus klinischen Studien (allgemeiner: empirisch ermittelten Studienergebnissen) – vorhalten, bilden einen wichtigen Baustein der systematischen Übersichtsarbeit. Sie werden zunächst hinsichtlich ihrer Qualität und Verfügbarkeit charakterisiert.

Es folgt eine Übersicht über die für systematische Literaturrecherchen nutzbaren fachspezifischen Informationssysteme und die dort verfügbaren Retrieval-Methoden. Ausgehend davon werden Suchprozesse der systematischen Übersichtsarbeit eingehend methodologisch wie auch aus verschiedenen kontextuellen Blickwinkeln erläutert. Dabei stehen entsprechend des Fokus der Arbeit Reviews in der Medizin im Mittelpunkt.

Ausgehend von den Grundsätzen der evidenzbasierten Medizin lässt sich das Ziel einer systematischen Übersichtsarbeit als die nach *systematischen* Methoden durchgeführte Identifikation, Bewertung und Synthese *aller* verfügbaren Evidenz bezüglich einer vorab klar definierten medizinischen Forschungsfrage, der „*well built clinical question*“ definieren (vgl. Abschnitt 1.1.2).

Diese Aufgabenstruktur ist in der EBM untrennbar verbunden mit der Methodenlehre zur Durchführung systematischer Übersichtsarbeiten, welche ihre Vollständigkeit, Reproduzierbarkeit und Fehlerrobustheit garantieren soll. Natürlich übrigens ist man sich auch in der Methodenlehre systematischer Übersichtsarbeiten durchaus bewusst, dass der Anspruch eines „*total recall*“ lediglich einen fiktiven, aus der EBM ableitbaren Anspruch erfüllt, der in der Praxis nie realisierbar sein kann, weil schon das Universum verfügbarer potenziell relevanter Information nicht bekannt ist, vgl. (Gough et al. 2012, S. 112), per definitionem ist dieser Anspruch in der EBM jedoch einer der Grundpfeiler der reinen Methodenlehre.

Das Kapitel führt abschließend in ein informationswissenschaftliches Rahmenwerk. Dieses ist an das Ingwersen und Järvelin-Modell angelehnt und stellt Kontextparameter, Herausforderungen sowie Probleme in Retrievalprozessen der systematischen Übersichtsarbeit heraus. Es dient der Überleitung zu den Forschungsfragen und dem empirischen Teil der Arbeit.

3.1 Evidenzquellen: Ergebnisse klinischer Studien

3.1.1 Inhaltliche Qualität

Wie bereits einleitend erläutert, versteht die EBM den Begriff der Evidenz im Sinne einer empirischen Evidenz, die sich direkt in der Studien- oder Publikationsqualität widerspiegelt:

Wegweisendes Element systematischer Übersichtsarbeiten in der Medizin ist die Frage nach der empirischen Beweiskraft klinischer Studienergebnisse, die sich sowohl an ihrer methodologischen als auch klinischen Validität festmachen lässt. Die randomisierte klinische Studie bildet die Basis für die „neue Disziplin“ der evidenzbasierten Medizin (Daly 2006), sie steht für einen methodisch und klinisch validen Forschungsansatz und steht methodisch an höchster Stufe der Evidenz.

Stufen der Evidenz wurden 2001 erstmalig von Haynes in der so genannten „4S Struktur“ definiert (Haynes 2001) und später von Dicenso et al. in ein 6-stufiges Modell einer Evidenzpyramide überführt (Dicenso et al. 2009), siehe Abbildung 3-1.

Klinische Entscheidungssysteme definieren den höchsten Grad der Evidenz. Sie erlauben den Abgleich detaillierter Patientendaten mit einer Wissensmanagementbasis und tragen nachhaltig zu erfolgreichen Entscheidungsprozessen bei, sind allerdings aufgrund ihrer praktisch schwierigen Implementierung und insgesamt kostenintensiven Anschaffung im Alltag nur selten verfügbar (Dicenso et al. 2009).

In der Pyramide werden sie gefolgt von medizinischen Leitlinien wie beispielsweise die AWMF Leitlinien. Diese, so die AWMF (Arbeitsgemeinschaft der Wissenschaftlichen Medizinischen Fachgesellschaften e.V. (AWMF) e.V. 2016), sind als „Hilfen für Ärzte zur Entscheidungsfindung in spezifischen Situationen“ zu verstehen, sie „beruhen auf aktuellen wissenschaftlichen Erkenntnissen und in der Praxis bewährten Verfahren und sorgen für mehr Sicherheit in der

Medizin, sollen aber auch ökonomische Aspekte berücksichtigen“. Leitlinien sind „für Ärzte rechtlich nicht bindend“. (Arbeitsgemeinschaft der Wissenschaftlichen Medizinischen Fachgesellschaften e.V. (AWMF) e.V. 2016)

Die nächste Evidenzstufe erreichen vergleichende Übersichten zu systematischen Reviews und systematische Reviews selbst, zusammengefasst beispielsweise in speziellen Datenbanken wie der Cochrane Library (John Wiley & Sons, Inc. 2016c).

Den geringsten Grad an Evidenz haben sowohl Originalartikel als auch Übersichtsarbeiten zu Studien (d.h. solche ohne Anspruch auf Systematik und Vollständigkeit).

Die Suche nach Evidenz erfolgt nach der Evidenzpyramide idealerweise entsprechend ihrer Verfügbarkeit von oben nach unten (Dicenso et al. 2009). Für die systematische Übersichtsarbeit sind demnach systematische Übersichtsarbeiten die umfangreichste, qualitativ hochwertigste und daher offensichtlichste Informationsquelle (Lefebvre et al. 2011).

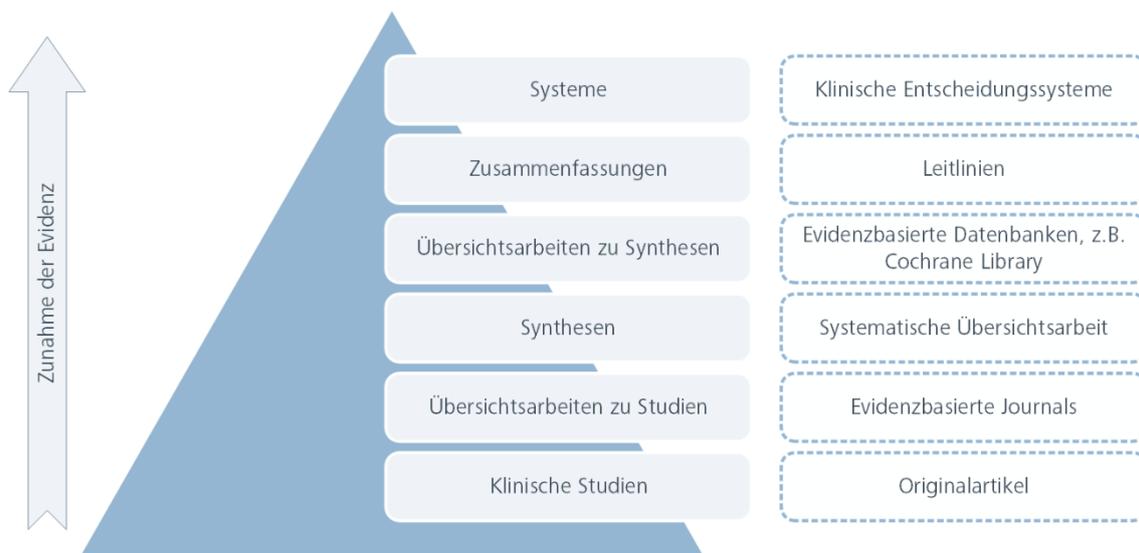


Abbildung 3-1 Die 6S Evidenzpyramide nach (Dicenso et al. 2009)

Die hierarchische Struktur der Evidenzpyramide lässt sich analog auf das Studiendesign anwenden (Concato et al. 2000):

So sind grundsätzlich alle randomisierten kontrollierten klinischen Studien (engl. *randomized controlled trials*, kurz RCT) hinsichtlich ihrer Evidenz höher zu werten als solche ohne Randomisierung. Beides ist hinsichtlich der Evidenz wiederum stärker als die methodisch weniger

strengen Kohorten,- oder Fall-Kontroll-Studien zu beurteilen. Case-Reports oder Expertenmeinungen als subjektivste Konzepte sind in der Evidenzpyramide ganz unten anzusiedeln.

3.1.2 Verzerrungen in der Forschungs- und Publikationslandschaft

Die primären Quellen für systematische Literaturrecherchen sind nach der Evidenzpyramide die Ergebnisse klinischer Studien, veröffentlicht in Form von Originalartikeln, der so genannten Primärliteratur, definitionsgemäß die „Gesamtheit der literarischen, philosophischen o. ä. Texte, die selbst Gegenstand einer wissenschaftlichen Untersuchung sind“ (Bibliographisches Institut GmbH Duden Verlag 2016e).

Primärliteratur als „eigentlicher literarischer Text“ (Wilpert 1989) lässt sich von Sekundärliteratur, d.h. „wissenschaftlichen und kritischen Werken“ über die Primärliteratur, abgrenzen (Wilpert 1989). Im Englischen spricht man – auch in den Bibliothekswissenschaften - mit *primary source* nicht von Literatur sondern Quelle, vgl. (Oxford University Press 2016e, Stichwort: source).

Die systematische Übersichtsarbeit selbst zählt daher definitionsgemäß nicht zur Primärliteratur, nutzt diese aber (Cochrane Collaboration 2015):

“A systematic review is a high-level overview of primary research on a particular research question that tries to identify, select, synthesize and appraise all high quality research evidence relevant to that question in order to answer it.”

Der Zugriff auf die in Fachzeitschriften veröffentlichte Primärliteratur erfolgt in der Regel über Fachdatenbanken. Fachdatenbanken enthalten häufig eine Vielzahl an Publikations- und Studientypen, sodass hier Ergebnisse aus RCTs in Form von Originalarbeiten oder Referenzen auf diese Originalarbeiten ebenso aufzufinden sind wie systematische Übersichtsarbeiten oder Referenzen auf diese. Datenbanken können in folgende Kategorien eingeteilt werden, eine detaillierte Betrachtung folgt in Abschnitt 3.2:

- *bibliographischen* Datenbanken (synonym auch Referenzdatenbank oder Literaturdatenbank) halten Referenzen auf Originalarbeiten vor und verweisen häufig auf Ressourcen zum Zugriff auf zugehörige Volltextdokumente
- reine *Volltextdatenbanken* stellen den direkten Zugriff auf Volltexte der Originalarbeiten bereit
- *Zitationsdatenbanken* enthalten Zitationen und Informationen darüber bereit, welche Veröffentlichung von anderen Veröffentlichungen zitiert wird

Erschwert wird die Recherche nach Primärliteratur und damit den Originalergebnissen unter anderem durch systematische Fehler (*biases*).

Unter einem Bias versteht man im deutschen Sprachgebrauch eine „durch falsche Untersuchungsmethoden (z.B. Suggestivfragen) verursachte Verzerrung des Ergebnisses einer Repräsentativerhebung“ (Bibliographisches Institut GmbH Duden Verlag 2016c).

Ein nicht entdeckter Bias kann aufgrund seiner systematischen Verzerrung letztendlich zu einer Reihe für die evidenzbasierte Medizin systematischer Fehler und damit beispielsweise zur Überbewertung positiver Studienergebnisse führen, vor allem sind hier der Prozess der Literaturrecherche als auch der der Analyse betroffen (Felson 1992; Egger und Smith 1998; Song et al. 2000).

Eine Übersicht über die im Kontext der systematischen Übersichtsarbeit bekannten Biases findet sich in Tabelle 3-1.

Tabelle 3-1 Biases im Zusammenhang mit systematischen Literaturrecherchen nach (Booth et al. 2012, S. 185) und (Egger und Smith 1998)

Bias Typ	Auslöser
Datenbankbias	Verzerrte Indexierung von Studien in einer Fachdatenbank
Finanzierungsbias	Veröffentlichung favorisierter Ergebnisse bei Industrie-finanzierten Studien
Graue Literatur Bias	Divergenz von Ergebnissen in grauer Literatur und Fachjournalen
Sprachbias	Wahrscheinlichkeit, dass positive Ergebnisse eher englischsprachig veröffentlicht werden
Bias bei Mehrfachpublikation	Mehrfachpublikationen auf Basis einer einzigen Studie

Publikationsbias	Selektive Übermittlung und Annahme von Studien in Journals
Zitationsbias	Inhaltliche Fokussierung zitierender Veröffentlichungen und Korrelation von Studienergebnissen
Bias durch Studienqualität	Assoziation von Studien geringer Studienqualität und positiven Ergebnissen
Ergebnisbedingter Bias	Reduktion der Veröffentlichung auf signifikante Ergebnisse

Eine problematische, systematische Verzerrung von Ergebnissen kann vor allem durch den so genannten *Publikationsbias* zustande kommen.

Der Begriff Publikationsbias wird in der Literatur unterschiedlich verwendet, die folgende Definition geht auf Dickersin zurück, welche Studienergebnisse in statistisch-signifikant (positiv) und nicht signifikant (negativ) unterteilt (Dickersin 1990). Nach dieser Vorstellung entsteht der Publikationsbias, wenn Verlage Einreichungen von Studienergebnissen mit statistisch signifikanten (i.S. von positiven) Ergebnissen tendenziell öfter oder eher akzeptieren und veröffentlichen als Arbeiten mit nicht-signifikanten Ergebnissen (Song et al. 2000). In diesem Fall sind Studien, die ein diagnostisches Verfahren oder eine Therapie erfolgversprechend bewerten gegenüber solchen Studien, die keinen Vorteil des Verfahrens oder der Therapie sehen, in der Literatur überrepräsentiert, was zu einer irreführenden Darstellung von Publikationen in Fachzeitschriften und damit letztendlich zu einer systematischen Verzerrung in der Gesamtdarstellung von Studienergebnissen führt.

Der Publikationsbias wurde das erste Mal in Sozialwissenschaften berichtet, ist aber heute auch in der Medizin dokumentiert (Felson 1992). Für medizinische Fachzeitschriften wurde die tatsächliche Existenz des Publikationsbias beispielsweise in einer retrospektiven Studie von Easterbrook et al. nachgewiesen (Easterbrook et al. 1991).

Wie dieser Bias zustande kommt, ist nicht abschließend geklärt. Beispielsweise könnten Prüfärzte, Peer Review und andere verantwortlich sein, eventuell seien es die Autoren selbst, die Studien mit nicht-signifikanten Studienergebnissen erst gar nicht veröffentlichen würden (Song et al. 2000). Dubben und Beck-Bornholdt vermuten zusätzlich die Vorselektion negativer Ergebnisse bei pharmazeutisch finanzierten Studienergebnissen – in diesem Fall ein sogenannter ergebnisbedingter oder auch ein Finanzierungsbias (Dubben und Beck-Bornholdt 2004).

Ein Zitationsbias kann beispielsweise aus übermäßigem Forschungskontakt zu einzelnen Kollegen oder einer überproportionalen Sammlung der Literatur aus einzelnen Referenzlisten resultieren. Egger et al. (Egger und Smith 1998) nennen hierfür ein konkretes Beispiel: Bei der Suche zum Thema Cholesterinsenkung, so fanden die Autoren, waren Veröffentlichungen mit medikamentöser Unterstützung in der Literatur überrepräsentiert. Die Beschränkung auf Referenzlisten in Primär- und Sekundärliteratur führte in diesem Fall dazu, dass existierende Veröffentlichungen mit differenzierterer Betrachtung, die zur objektiven Beantwortung der klinischen Fragestellung wichtig gewesen wären, übersehen wurden.

Der Datenbankbias ist ähnlich definiert, er kommt üblicherweise zustande, wenn sich die Originalarbeiten, die in einer Datenbank indexiert sind, von den Ergebnissen nicht publizierter Studien systematisch unterscheiden (Felson 1992; Song et al. 2000).

Tendieren Wissenschaftler dazu, positive Ergebnisse vorwiegend in großen, englischsprachigen Journals zu veröffentlichen, negative jedoch in lokalen, anderssprachigen Journals, führt dies zu einem Sprach- als auch einem Datenbankbias. Vermieden werden sollte auch die Berücksichtigung von Mehrfachveröffentlichungen eines Autors oder einer Forschergruppe zu ein und derselben Fragestellung: Mehrfach in die spätere Analyse eingehende Daten verfälschen spätere Ergebnisse. (Egger und Smith 1998)

Ergebnisbedingte Biases und Publikationsbias führten in der klinischen Forschung zu konkreten Gegenmaßnahmen wie beispielsweise der Registrierungspflicht für klinische Studien. Seit 2004 dürfen nur noch Ergebnisse der klinischen Studien in Fachzeitschriften veröffentlicht werden, die vorab registriert wurden (Angelis 2005).

Die Motivation zur Einführung klinischer Studienregister fassen Antes et al. zusammen: Wer klinische Studien durchführe, habe eine Verpflichtung, deren Ergebnisse zu veröffentlichen. Jede unterlassene Veröffentlichung könne zu einer verzerrten Wahrnehmung führen, beispielsweise eine „Überbetonung“ eines Therapienutzens. Dem könne unter anderem mit der Veröffentlichung klinischer Studien in Studienregistern entgegengewirkt werden (Antes et al. 2009). Antes et al. betonen die Relevanz klinischer Studienregister für die Gesundheitsversorgung – und damit natürlich auch ihre Relevanz im Rahmen systematischer Übersichtsarbeiten: Vielfach werden wissenschaftliche Arbeiten, gerade in der Medizin, überhaupt nicht veröffentlicht. Die Rate nicht publizierter medizinischer Ergebnisse nennen Antes et al. „völlig inakzeptabel“ (Antes et al. 2009).

Klinische Studienregister, spezielle Datenbanken, in denen alle relevanten Informationen über abgeschlossene und laufende klinische Studien vorgehalten werden können, sind daher heute eine wichtige weitere Quelle für die systematische Übersichtsarbeit.

3.2 Dokumenten- und Datenverfügbarkeit

3.2.1 Bibliographische Fachdatenbanken in der Medizin

Die in Fachkreisen bekannteste und eine der weltweit größten bibliographischen Datenbanken für medizinische Primärliteratur ist MEDLINE (*Medical Literature Analysis and Retrieval-System Online*) der U.S. National Library of Medicine (NLM). MEDLINE umfasst neben den Fachgebieten aus der Humanmedizin und Zahnmedizin auch Nachweise veterinärmedizinischer Veröffentlichungen sowie solche des öffentlichen Gesundheitswesens aus aktuell über 5600 verschiedenen, internationalen Fachzeitschriften (U.S. National Library of Medicine 2015d) (Stand 02.02.2016).

Historisch basiert MEDLINE auf dem von der NLM herausgegebenen *Index Medicus*. Dieser ging auf das Engagement des amerikanischen Mediziners Dr. John Shaw Billings (1838-1913) zurück, der sich in den 1880er Jahren sukzessive darum bemühte, die seit Beginn des 19. Jahrhunderts durch das Militär gesammelten Veröffentlichungen zur chirurgischen Versorgung von Soldaten systematisch zu katalogisieren und in eine nationale Bibliothek zu überführen. Dieser gedruckt verfügbare Index wurde von der NLM bis 2004 herausgegeben (National Library of Medicine (U.S.). History of Medicine Division 2014). MEDLINE indexiert heute alle im Index Medicus verfügbaren Datensätze ab 1966.

Der Einsatz von Computern im bibliographischen Information Retrieval geht auf die 1950- und 1960er Jahre zurück, siehe (Baeza-Yates und Ribeiro-Neto 2011, S. 699), auch die Ursprünge der Datenbank MEDLINE fallen in die 1960er Jahre:

1964 wurde mit MEDLARS (Medical Literature Analysis and Retrieval-System) der erste breit angelegte, computerbasierte und öffentlich zugängliche Suchservice ins Leben gerufen, damals noch auf Magnetbändern, die von teilnehmenden Institutionen abgerufen werden konnten. Im Oktober 1971 folgte mit MEDLINE (*MEDLARS Online*) der erste Online-Service zur

Suche in NLM Datenbanken, die über Bandlaufwerke auch Universitäten und anderen Einrichtungen zugänglich gemacht wurden. MEDLINE ist damit die erste große Datenbank, die über Telekommunikationsnetzwerke verteilt wurde. Da spezielle Schulungen für die Erlernung von Suchanfrage Syntax und den Einsatz von Operatoren nötig waren, wurde MEDLINE (im Auftrag von Wissenschaftlern) zu dieser Zeit fast ausschließlich von Bibliothekaren und Informationsspezialisten bedient, siehe auch Abschnitt 2.1.2.

Seit 1997 kann auf MEDLINE über das von der NLM vertriebene Zugangssystem Pubmed online kostenfrei zugegriffen werden (vgl. (National Center for Biotechnology Information (NCBI) 2016)). Pubmed fungiert heute als wichtigste Schnittstelle zum Zugriff auf die Datenbank MEDLINE. Aktuell sind über Pubmed mehr als 26 Millionen Einträge aus MEDLINE und anderen NLM-Quellen verfügbar (Stand 28.06.2016). Die Aktualisierung des Pubmed Datenbestands erfolgt laufend an den Wochentagen Dienstag bis Samstag (U.S. National Library of Medicine 2016a), die Datenbank ist also als tagesaktuell anzusehen. Aktualität und Status einzelner Einträge des Pubmed Datenbestands sind jedoch höchst unterschiedlich: Pubmed enthält tagesaktuell auch Datensätze, die zwar bereits publiziert, aber noch nicht oder nicht vollständig in MEDLINE indexiert wurden. Daneben gehören zu Pubmed Einträge aus dem Datenbestand OLDMEDLINE, dem Datenbestand von vor 1966 (U.S. National Library of Medicine 2014a), der insgesamt nicht vollständig indexiert ist und in Zukunft nicht mehr indexiert wird.

Auch der kommerzielle Anbieter Ovid Technologies, ein Teil der niederländischen Verlagsgruppe Wolters Kluwer, bietet über den Host Ovid Zugang zu MEDLINE an. Das Angebot von Ovid MEDLINE umfasst neben dem aktuellen MEDLINE Datenbestand auch Zugriff auf nicht-indexierte oder derzeit in Bearbeitung befindliche MEDLINE Einträge sowie OLDMEDLINE, dem Datenbestand zwischen 1948 und 1965. Die Ovid MEDLINE Datenbanken werden je nach Typ täglich, wöchentlich oder vierteljährlich aktualisiert. (Ovid Technologies 2016)

Eine weitere große und für die medizinische Literaturrecherche wichtige bibliographische Datenbank ist EMBASE (*Excerpta Medica Database*) des Herstellers Reed Elsevier (RELX Group), eine kommerziell vertriebene Datenbank (Elsevier B.V. 2016). EMBASE hält Nachweise internationaler Literatur aus der Humanmedizin mit besonderem Schwerpunkt auf Pharmakologie, Toxikologie und angrenzender Gebiete, hat sich dabei jedoch auf europäische Veröffentlichungen spezialisiert. Aktuell sind über EMBASE 31,5 Millionen Datensätze aus über 8400

Fachzeitschriften verfügbar, davon auch Veröffentlichungen bis ins Jahr 1947 (Stand 28.06.2016). Inhaltlich baut der Datenbestand auf den ebenfalls bei Reed Elsevier vertriebenen Excerpta Medical Full Set Series (Elsevier 2016a) auf, einer Serie aus Journals, die bibliographische Referenzen und Abstracts aus 4000 Fachzeitschriften aus 70 Ländern beinhaltet. Auch EMBASE kann über den Host der Ovid Technologies recherchiert werden.

Der Datenbestand von EMBASE und MEDLINE hat größere Schnittmengen und bietet seit 2010 alle Fachzeitschriften an, die in MEDLINE *indexiert* sind. Zusätzlich können aber über EMBASE Beiträge aus weiteren 2700 nicht in MEDLINE der Pubmed indexierten Journals und 300.000 zusätzlich indexierte Konferenzbeiträge abgerufen werden.

Für Literaturrecherchen im Bereich der Psychologie und angrenzender medizinischer Fachgebiete bietet die American Psychological Association (APA) mit PsycINFO eine weitere große bibliographische Datenbank an, die aktuell 4 Millionen bibliographische Datensätze umfasst und über 2500 Fachjournals indexiert (American Psychological Association 2016). Wie auch EMBASE indexiert PsycINFO neben Originalartikeln zusätzlich Konferenzabstracts und andere Publikationstypen.

Weitere bibliographische Fachdatenbanken haben sich auf Originalarbeiten regionaler Forschung spezialisiert, beispielsweise LILACS (Literature in the Health Sciences in Latin America and the Caribbean) auf die lateinamerikanische Fachliteratur (Latin American and Caribbean Center on Health Sciences Information 2016), der African Index Medicus auf afrikanische Fachliteratur (WHO Regional Office for Africa 2016).

Wichtig ist anzumerken, dass sich die hier genannten Zahlenangaben und Informationen natürlich jährlich ändern. Insbesondere müssen aktuelle Daten immer direkt über den Anbieter recherchiert werden, was vor allem bei kommerziellen Hosts nicht immer einfach ist.

3.2.2 Volltext,- und Zitationsdatenbanken

Volltextdatenbanken sind neben bibliographischen Datenbanken ein weiteres wichtiges Werkzeug der Literatursuche. Im Gegensatz zu bibliographischen Datenbanken stellen sie Volltexte bereit, sodass die Beschaffung der Originalarbeiten – frei verfügbar oder mit Lizenzgebühren – direkt über den Datenbankanbieter erfolgen kann.

Die Cochrane Library (John Wiley & Sons, Inc. 2016c) gilt weltweit als eine der hochwertigsten Quellen zur evidenzbasierten Medizin. Sie beinhaltet 650.000 Einträge bzw. Reviews von Fachzeitschriften, Tagungsbeiträgen und anderen Quellen zu aktuellen medizinischen Behandlungen sowie zu kontrollierten klinischen Studien und ist damit erste Anlaufstelle für das Auffinden systematischer Übersichtsarbeiten sowie Ergebnissen randomisierter klinischer Studien. Die Cochrane Library wächst jährlich um ca. 30.000 Einträge und wird vierteljährlich aktualisiert. Die Datenbanken der Cochrane Library werden von der Cochrane Library und dem Anbieter Wiley hergestellt und sind über die Websites von Wiley verfügbar.

Die Volltextdatenbank „Cochrane Database of Systematic Reviews“ (CDSR), abrufbar über die Cochrane Library, zählt weltweit zu den führenden Quellen für systematische Übersichtsarbeiten in der Medizin. Die CDSR beinhaltet nicht nur so genannte Cochrane Reviews (systematische Übersichtsarbeiten, die nach den methodologischen Vorgaben der Cochrane Collaboration durchgeführt wurden), sondern auch Leitartikel. Mit einem *Impact Factor* von 6,03 liegt die zugehörige Fachzeitschrift *Cochrane Database of Systematic Reviews* derzeit unter den 10 meistzitierten medizinischen Fachzeitschriften.

Das Cochrane Central Register of Controlled Trials (CENTRAL) ist als Literaturdatenbank auf randomisierte klinische Studien spezialisiert. Sie wird ebenfalls von der Cochrane Collaboration herausgegeben und ist über die Cochrane Library (John Wiley & Sons, Inc. 2016c) aufrufbar. Über CENTRAL lassen sich nicht nur Verweise aus MEDLINE und anderen bibliographischen Datenbanken recherchieren, sondern auch Referenzen auf in der Primärliteratur nicht veröffentlichte Studien aus anderen Quellen.

Damit sind die CDSR und Cochrane Central essentielle Quellen für die Erstellung systematischer Übersichtsarbeiten, die dort vorgehaltenen Dokumente sind von immenser Bedeutung für die systematische Übersicht (vgl. Abschnitt 3.1.1).

Titel und Abstracts beider Datenbanken sind über das Internet frei durchsuchbar, der Zugriff auf Volltexte unterliegt verschiedenen Lizenzierungsmodellen (John Wiley & Sons, Inc. 2016a).

Zitationsdatenbanken wiederum sind spezialisierte Datenbanken, die ausschließlich Zitationen enthalten – Angaben darüber also, welche Publikationen von anderen zitiert werden.

Die sicherlich bekannteste Zitationsdatenbank ist der Social Sciences Citation Index des Anbieters Thomson Reuters mit Zitationsangaben aus über 3000 internationalen Journals in 50 Disziplinen (Thomson Reuters 2016b). Ergänzt wird das kommerzielle Angebot von Thomson

Reuters um die Zitationsdatenbank Web of Science (ISI Web of Knowledge), die derzeit 59 Millionen Dateneinträge enthält und im Datenbestand Originalartikel bis zum Ende des 19. Jahrhunderts berücksichtigt (Thomson Reuters 2016c).

Vom Anbieter Elsevier steht mit Scopus eine weitere Abstract- und Zitationsdatenbank für die Recherche zur Verfügung (Elsevier 2016b).

3.2.3 Zugriff auf Fachdatenbanken

In den vergangenen Jahrzehnten ist der Zugang zu Fachdatenbanken für Wissenschaftler und Rechercheexperten durch die Verfügbarkeit im Internet deutlich einfacher geworden: Sowohl bibliographische Datenbanken als auch Volltextdatenbanken und Zitationsdatenbanken sind heute im Internet weltweit zugänglich, allerdings je nach Hersteller häufig kostenpflichtig. Die Suche in Fachdatenbanken erfolgt dabei allerdings nicht ausschließlich direkt über den Hersteller und Datenbankanbieter, sondern auch über so genannte Hosts, einen Datenbankvertrieb.

Hosts stellen über eine proprietäre Software für den parallelen Zugriff auf eine Vielzahl meist thematisch verwandter Fachdatenbanken bereit und bestimmen damit zum einen Recherche-funktionalität, zum anderen aber auch die Lizenzmodelle für einzelne Datenbanken in Form von Geschäftsmodellen mit dem Endnutzer. Die verteilten Rollen von Datenbankanbietern und Hosts finden sich beispielsweise bei Baeza-Yates und Ribeiro-Neto:

Tabelle 3-2 Rollen von Datenbank Anbietern und Hosts nach (Baeza-Yates und Ribeiro-Neto 2011, S. 702)

Datenbank Anbieter	Host
Entwicklung und Definition der Datenbank Struktur	Entwicklung einer Retrieval Software für eine Vielzahl an Datenbanken
Sammlung von Primärliteratur, standardisierter bibliographischer Informationen bis hin zu Abstracts	Lizenzierung des Zugriffs auf Datenbanken
Entwicklung von Thesauri und Indexierungsverfahren mit kontrolliertem Vokabular	Mounting von Datenbanken unter Einsatz invertierter Indizes

Standardisierte bibliographische Informationen	Standardisierte Strukturen für Datensätze
Datei Updates in regelmäßigen Abständen	Update der angebotenen Datenbanken in regelmäßigem Turnus
	Dokumentation für Endnutzer
Geschäftsmodelle basierend auf Updates bei Hosts	Geschäftsmodelle mit dem Endnutzer
	Zusätzliche Service Dienstleistungen für Endnutzer durch Schulungsangebote

Der Vorteil bei der Nutzung eines Datenbankvertriebs besteht zweifelsohne in der Zusammenführung des Zugriffs auf verschiedene thematisch verwandte Datenbanken unter einem Dach, wobei die Syntax einzelner Anbieter für alle dort verfügbaren Datenbanken klassischerweise meist einheitlich gestaltet ist und über ein Interface, häufig per Single-Sign-On erreichbar ist (Baeza-Yates und Ribeiro-Neto 2011, S. 701).

Dies klingt zunächst einfach und scheint auch für Endnutzer ohne professionelle Kenntnisse machbar. Doch stehen diesen vermeintlichen Vorteilen auch Nachteile gegenüber:

Kosten

Bis auf wenige Ausnahmen sind Datenbankvertriebe kommerzieller Art, was häufig mit der Finanzierung der nach wie vor kostspieligen Implementierung zu tun hat. Damit ist die Nutzung dieser Vertriebe kostenpflichtig: Die Hosts lizenzieren Datenbanken von den Anbietern selbst und bieten dem Endnutzer verschiedene Abonnement-Modelle (Baeza-Yates und Ribeiro-Neto 2011, S. 702). Anfallende Kosten betreffen dabei in bibliographischen Datenbanken häufig nicht nur die Beschaffung von Volltexten, sondern bereits die Anforderung von Ergebnislisten pro zurückgeliefertem Treffer.

Die Informationen zu diesen Treffern ist je nach Host und Suchabonnement, für den Laien vergleichsweise rudimentär, vgl. (DIMDI: Deutsches Institut für Medizinische Dokumentation und Information 2014): Lediglich Meta-Informationen wie Autorenschaft, Titel des Artikels, Journal, Erscheinungsjahr werden angegeben. Der Nutzer erhält also zunächst lediglich Infor-

mationen über die Existenz eines Originalartikels sowie dessen Beschaffungsort beim Fachverlag. Werden vom Nutzer unbedacht und wenig zielgerichtet Suchanfragen angefordert, kann dies für den weniger versierten Anwender schnell zu immensen Kosten und einer unvorteilhaften Kosten-Nutzen-Rechnung führen.

Information Retrieval und Retrieval-Sprachen

Insgesamt sind die Services in Entwicklung und Wartung teuer, auf Anbieter- als auch auf Host Seite. Vor allem Hosts sind heute vorwiegend in der Hand großer Verlage, sodass die Trennung zwischen Datenbankanbieter und Datenbank Vertrieb vielfach verschmilzt. Beide finden sich im Verlauf der Jahre in direkter Konkurrenz mit dem Internet. Insbesondere Hosts sind daher gezwungen, ihre Interfaces und Services an die Bedürfnisse der Nutzer anzupassen (Baeza-Yates und Ribeiro-Neto 2011), S. 706.

Zu den weltweit größten Datenbankvertrieben zählen der Host Dialog des amerikanischen Unternehmens Dialog, ein Tochterunternehmen der ProQuest, ebenfalls ein Datenbankhost. Auch Pubmed kann als Host interpretiert werden: Die NLM hat mit Pubmed ein Interface geschaffen, über das neben MEDLINE auf viele weitere bibliographische Datenbanken und zum Teil auch Volltextdatenbanken zugegriffen werden kann.

Pubmed bietet als Dienst der U.S. National Library of Medicine umfangreiche Möglichkeiten der Suche sowohl im gesamten MEDLINE Datenbestand wie auch in anderen NLM Datenbeständen, beispielsweise in der Volltext-Datenbank Pubmed Central und ist im Gegensatz zu vielen anderen Zugriffssystemen im Internet frei zugänglich. Mit seiner Eröffnung 1997 gehört es mit zu den jüngsten Produkten der NLM rund um die Datenbank MEDLINE und ist heute für viele Forscher in der Medizin – sicher auch aufgrund der öffentlichen Verfügbarkeit – die erste Anlaufstelle bei der Suche nach medizinischer Fachliteratur. Der Zugriff auf verlinkte Volltexte ist im akademischen Bereich abhängig von der Lizenzierung an den jeweiligen Universitäten oder Einrichtungen.

Bundesweit fungiert das Deutsche Institut für Medizinische Dokumentation und Information DIMDI (DIMDI: Deutsches Institut für Medizinische Dokumentation und Information 2014), eine nachgeordnete Behörde des Bundesministeriums für Gesundheit, derzeit noch unter anderem als Host für die Datenbankrecherche in der Medizin. Das DIMDI bietet derzeit Zugriff auf 30 Fachdatenbanken und damit über 140 Millionen Dokumente, darunter auch MEDLINE,

EMBASE sowie Cochrane CENTRAL. Im Februar 2016 hat das DIMDI angekündigt, seine Hosting Services ab 2017 einzustellen (DIMDI: Deutsches Institut für Medizinische Dokumentation und Information 2016a).

Werden mehrere Hosts oder Fachdatenbanken für eine Recherche eingesetzt, so erfordert dies vom Nutzer die Kenntnis mehrerer meist proprietärer, Anbieter-spezifischer Syntaxen, d.h. Retrievalsprachen. Hosts untereinander verfügen also über keine vereinheitlichte Retrievalsprache, keine vereinheitlichte Retrieval-Software und ebenso über keinen gemeinsamen Dokumentenstandard. Vor allem der Konkurrenzdruck ist ein Grund dafür, dass Datenbankanbieter oder Hosts ihren Nutzern eine Vielzahl weiterer, auf den Kenntnisstand des Nutzers angepasster Funktionalitäten anbieten, um die Suche zu vereinfachen komfortabler zu machen. Aufgrund der teils langsamen und eingeschränkten systemseitigen Verarbeitung vor allem benutzerfreundlicher Bedien-Interfaces für die Zielgruppe der Domänenexperten werden daher parallel Abfragemodi mit Kommandozeilen-Bedienung zur Verfügung gestellt, die von professionellen Nutzern eingesetzt werden können. (Baeza-Yates und Ribeiro-Neto 2011, S. 699)

Die Recherche bei DIMDI erfolgt beispielsweise in verschiedenen Lizenzmodellen und bietet zudem Suchoberflächen, die dem Kenntnisstand einzelner Nutzergruppen angepasst sind:

Einsteigern ohne Vertrag bietet DIMDI die Suche über eine menügesteuerte Suchoberfläche (DIMDI SmartSearch) Zugriff auf einen Teilbestand der Datenbanken, exemplarisch dargestellt in Abbildung 3-2. Die Suche wie auch die Ausgabe der Dokumententitel sind dabei kostenfrei, Volltexte können kostenpflichtig direkt über das DIMDI oder Bibliotheken beschafft werden. Darüber hinaus bietet das DIMDI Lizenzmodelle für professionelle Recherchen in allen vom DIMDI lizenzierten Datenbanken in einer professionellen, kommandozeilenbasierten Umgebung (die so genannte DIMDI Classic Search).

0 Datenbankdokumente
0 Volltexte
0,00 EUR
zum Warenkorb >

Suche
Suchergebnis
Dokumentaushabe
Merkliste (0)

Suchformulierung: (AU=REVIEW ? OR (REVIEW#)) AND SYSTEMAT###

Trefferzahlen in den einzelnen Datenbanken ▲

Kürzel	Name	Info	Trefferzahl	Ergebnis
ME10	MEDLINE	Info	77208	anzeigen
BA10	BIOSIS Previews	Info	37640	anzeigen
EA08	EMBASE Alert	Info	2927	anzeigen
EM10	EMBASE	Info	126365	anzeigen
GA03	German Medical Science - Journals	Info	301	anzeigen
GM03	German Medical Science - Meetings	Info	272	anzeigen
IS10	SciSearch	Info	73923	anzeigen

Gefundene Dokumente: 318636

Suchergebnis sortieren

Titel 1-10 von 318636 aller ausgewählten Datenbanken

Nr.	Dokumenttitel / Datenbank	€	Aktionen
<input type="checkbox"/>	1. Effectiveness of Stretch Interventions for Children With Neuromuscular Disabilities: Evidence-Based Recommendations. MEDLINE (ME10)	0,00	↔
<input type="checkbox"/>	2. Prolapsus rectal révélant une tumeur : intérêt de l'échographie abdominale, à propos de 3 cas. - [Rectal prolapse revealing a tumor: The role of abdominal ultrasound]. MEDLINE (ME10)	0,00	↔
<input type="checkbox"/>	3. Systematic reviews of observational studies: evaluating evidence quality. MEDLINE (ME10)	0,00	↔
<input type="checkbox"/>	4. Nonbismuth concomitant quadruple therapy for Helicobacter pylori eradication in Chinese regions: A meta-analysis of randomized controlled trials. MEDLINE (ME10)	0,00	↔
<input type="checkbox"/>	5. Return to Sport After Tibial Shaft Fractures: A Systematic Review. MEDLINE (ME10)	0,00	↔
<input type="checkbox"/>	6. Guidelines for the diagnosis, prevention and management of osteoporosis. MEDLINE (ME10)	0,00	↔

Abbildung 3-2 DIMDI Suche (28.06.2016)

3.2.4 Graue Literatur und nicht-bibliographische Datenbanken

Im Rahmen medizinischer systematischer Übersichtsarbeiten besteht in der Einbeziehung Grauer Literatur (vgl. Abschnitt 2.3.1.1) ergänzend zur Suche in Studienregistern die große Chance zur Vermeidung von Biases, somit erfüllt der Zugang zu Grauer Literatur einen Qualitätsanspruch (Hopewell et al. 2007b).

Der Zugriff auf Graue Literatur erweist sich im Vergleich zur der auf Originalarbeiten vergleichsweise mühsam und umfasst methodisch neben der Suche in speziellen Datenbanken zu grauer Literatur üblicherweise auch Suchprozesse wie den Kontakt mit Kollegen (Cook et al. 2001).

Seit 1992 bemüht sich die Initiative Grey Literature Network Service um die Belange der grauen Literatur, darunter auch die Veröffentlichung, Distribution und Zugänglichkeit grauer Literatur. Das Netzwerk stellt auf seinen Seiten einen umfangreichen, themenspezifischen Index zum Zugriff auf Graue Literatur bereit (INIST-CNRS - Institut de l'Information Scientifique et Technique 2016a).

Eine wichtige Informationsquelle für Graue Literatur ist die bibliographische Datenbank Open Grey (*System für Information on Grey Literatur in Europe*). Sie ermöglicht den freien Zugriff auf derzeit 700.000 bibliographische Referenzen auf in Europa produzierter grauer Literatur zu den Themen Wissenschaft, Technologie, biomedizinische Forschung, Wirtschaft und Sozialwissenschaften, darunter wissenschaftliche Reportings, Promotionen, Konferenzbeiträge und andere Formen Grauer Literatur (INIST-CNRS - Institut de l'Information Scientifique et Technique 2016b).

Mit dem Ziel, auch Veröffentlichungen zu identifizieren, die in bibliographischen Datenbanken nicht oder nur unzureichend indexiert wurden (vgl. (Lefebvre et al. 2011)), bilden Journals und Datenbanken, die sich auf Tagungsbände und Konferenzbeiträge spezialisiert haben, eine weitere Informationsressource in der systematischen Übersichtsarbeit.

Die Nutzung dieser Quellen führt direkt zur Notwendigkeit einer weniger strukturierten Form der Recherche: der Handsuche. Über sie werden manuell einzelne Journals und Tagungsbände manuell durchsucht (Hopewell et al. 2002).

Das „Handsearching Projekt“ in Freiburg (Das Deutsche Cochrane Zentrum 2014) führt in verschiedenen Projektverbänden seit 1995 eine systematische Handsuche in medizinischen Zeitschriften bis 1948 durch, erfasst die dort ermittelten randomisierten oder kontrollierten klinischen Studien im Register *Cochrane Central Register of Controlled Clinical Trials* (CENTRAL) und macht die Ergebnisse damit für die Wissenschaft leichter verfügbar (John Wiley & Sons, Inc. 2016b).

Die Handsuche in Grauer Literatur sowie die Berücksichtigung dort veröffentlichter Ergebnisse in Meta-Analysen oder systematische Übersichtsarbeiten wurden in den Anfängen kritisch beurteilt, da die Qualität nicht veröffentlichter Studien zunächst als zu niedrig eingeschätzt wurde oder sie methodisch schwerer zu beurteilen waren als dies bei Veröffentlichungen in

Journals mit Review-Prozessen der Fall ist. In der Praxis kann dies jedoch durch eine vorgeschaltete Definition von Qualitätskriterien reguliert werden (Hopewell et al. 1996; Conn et al. 2003b; Hopewell et al. 2007b), sodass diese Methode inzwischen vollständig etabliert ist.

3.2.5 Klinische und pharmazeutische Studienregister

Eine weitere wichtige Quelle für medizinische Evidenz sind klinische und pharmazeutische Studienregister, spezielle Datenbanken, in denen alle relevanten Informationen über abgeschlossene und laufende klinische Studien vorgehalten werden können.

In Deutschland erfolgt die Registrierung klinischer Studien im europäischen Register EudraCT (European Union Drug Regulating Authorities Clinical Trials), das seit 2004 durch die EU ins Leben gerufen wurde und seit 2011 für die Öffentlichkeit zugänglich ist (European Medicines Agency 2016). Das Register umfasst in Europa durchgeführte pädiatrische sowie interventionelle Studien und kann als behördliches Studienregister verstanden werden. Zu jeder Studie sind Angaben zu deren Durchführung, Finanzierung, Intervention sowie zum derzeitigen Status abrufbar. Für abgeschlossene Studien sind Abschlussberichte verfügbar.

Das Deutsche Register für klinische Studien (DRKS) als nationales Register ist als gemeinsames Projekt des Instituts für Bibliometrie und Medizinische Informatik und des Deutschen Cochrane Zentrums am Universitätsklinikum Freiburg verortet (Deutsches Register Klinischer Studien, Institut für Medizinische Biometrie und Statistik 2016) und wird seit 2007 vom Bundesministerium für Bildung und Forschung (BMBF) gefördert. Das DRKS richtet sich wie viele Studienregister nicht nur an Fachkreise, sondern auch an die interessierte Öffentlichkeit sowie Wissenschaft, Forschung und Industrie. Das DRKS basiert auf einer langjährigen Zusammenarbeit mit der WHO und angegliederten Gremien und baut daher technisch auf den Vorgaben des WHO Meta-Registers auf – damit übrigens wurde auch die Anerkennung des Registers durch die WHO garantiert (Dreier et al. 2009).

Die WHO ist mit ihrer Plattform ICTRP federführend bei den internationalen Bemühungen um die Registrierung *aller* international durchgeführten klinischen Studien (World Health Organization 2016). Die Vorteile einer gemeinsamen Organisation sind offensichtlich und werden

bei Dreier et al. (Dreier et al. 2009) genannt: Die Bündelung solcher Prozesse ermöglicht internationale Kollaboration und kann eine international einheitliche Qualität der erfassten Daten weitest möglich garantieren.

Eines der größten klinischen Studienregister weltweit ist neben dem ICTRP das Register ClinicalTrials.gov, das vom U.S. National Institutes of Health (NIH) und der National Library of Medicine (NLM) herausgegeben wird (U.S. National Institutes of Health (NIH) 2016). Das Studienregister (*registry*) wurde 1997 initiiert, die Datenbank (*result database*) ist seit 2008 öffentlich zugänglich. Über das Studienregister wie auch die Datenbanken können alle Informationen zu in den USA durchgeführten klinischen Studien sowie internationaler Studien abgerufen werden – die Registrierung von Studien ist dort international möglich. Die Datenbank bietet dezidiert Informationen über den Gesamtverlauf klinischer Studien sowie zugehörige Ergebnisse und statistische Analysen. (U.S. National Institutes of Health (NIH) 2016)

Zu den klassischen Erfassungsparametern klinischer Studienregister gehören beispielsweise Studientitel, kurze Beschreibungen oder Informationen zum Studiendesign, eventuell auch im Rahmen der Studie entstandene Veröffentlichungen. Trotzdem existiert für die Erfassung von Studien in klinischen Studienregistern bisher kein einheitlicher Standard, vgl. (Hausner und Kaiser 2010). Klinische Studienregister werden im Gegensatz zu bibliographischen Fachdatenbanken häufig von nicht-kommerziellen Anbietern bereitgestellt, siehe hierzu die Websites der Anbieter. (U.S. National Institutes of Health (NIH) 2016; World Health Organization 2016; Deutsches Register Klinischer Studien, Institut für Medizinische Biometrie und Statistik 2016).

Pharmazeutische Studienregister

Neben klinischen Studienregistern stellen pharmazeutische Studienregister eine „unverzichtbare Quelle bei der Suche nach nicht publizierten Daten“ dar (Hausner und Kaiser 2010). Bundesweit kann über die Website des Verbands Forschender Arzneimittelhersteller e.V. auf eine Liste verfügbarer pharmazeutischer Studienregister zugegriffen werden (Verband Forschender Arzneimittelhersteller e.V. 2016), in den USA stellen die „Pharmaceutical Research and Manufacturers of America“ (PhRMA) ein äquivalentes Produkt zur Verfügung (Pharmaceutical Research and Manufacturers of America 2016).

3.2.6 Zusammenfassung

Für die Identifikation weitestgehend aller verfügbaren Originalquellen zu einer gegebenen klinischen Fragestellung stehen zunächst klassische Fachdatenbanken mit Verweisen auf Originalliteratur oder Volltextdatenbanken zur Verfügung. Verzerrungen in der Literaturlandschaft machen es jedoch nötig, die Suche auf weitere Informationsressourcen auszudehnen. Die für die Recherche verfügbaren und erforderlichen Quellen erstrecken sich daher auf proprietäre, teils kommerzielle und mächtige Datenbanken mit Vollfunktionalität bis hin zu weniger bis kaum strukturierten Informationsquellen wie klinische Studienregister und Graue Literatur.

Klinische und pharmazeutische Studienregister sind eine Spezialität systematischer Übersichtsarbeiten in der Medizin, als Recherchequelle sind sie insbesondere zur Sicherung der Vollständigkeit systematischer Übersichtsarbeiten von Bedeutung.

3.3 Die systematische Übersichtsarbeit als Rechercheaufgabe

Wie bereits einleitend verdeutlicht, handelt es sich bei der systematischen Übersichtsarbeit um eine klassische Rechercheaufgabe im professionellen Kontext, die sich im Unterschied zu anderen Formen der Übersichtsarbeit durch eine strukturierte, klar vorgegebene Vorgehensweise auszeichnet.

3.3.1 Formalisierung des Informationsbedarfs

Klar definierte Fragestellungen garantieren die Anwendung der durch systematische Übersichtsarbeit ermittelten Ergebnisse auf spezifische klinische Probleme und bilden den Schlüssel zur evidenzbasierten Medizin (Richardson et al. 1995). Im Handbuch der Cochrane Collaboration heißt es: Einer systematischen Übersichtsarbeit liegen immer sowohl eine klar strukturierte Fragestellung als auch klar definierte Bewertungskriterien zugrunde (Lefebvre et al. 2011).

Die Formalisierung dieser Fragestellung gibt für strukturierte als auch unstrukturierte Suchprozesse den Handlungsrahmen vor. In diesem vorbereitenden Schritt zeigt sich, dass die Vorgehensweise der systematischen Übersichtsarbeit abhängig ist von Forschungsgegenstand beziehungsweise zugrundeliegenden Studiendesigns zu identifizierender Studien, die grob in qualitativ und quantitativ unterteilt werden können. Die vorliegende Arbeit beschränkt sich auf die in der Methodenlehre umfangreich beschriebenen quantitativen Fragestellungen als klassisches Element der EBM, der Vollständigkeit halber werden aber auch Unterschiede zu qualitativen Fragestellungen kurz erläutert.

3.3.1.1 Konzeptualisierung quantitativer Fragestellungen

In der evidenzbasierten Medizin erfolgt die Formalisierung in der Regel über Hilfsschemata. Eine gut strukturierte und klar definierte klinische Fragestellung zu interventionellen Themen folgt einem Muster, das sich von dem bereits 2002 vorgestellten so genannten PICO-Schema ableitet. Das Akronym PICO steht dabei für *Population/Problem, Intervention/Exposure, Comparison* und *Outcome* – eine Strukturierung wie sie klassischerweise in der Epidemiologie auftritt. Dieses Schema wurde von Guyatt et al. der epidemiologischen Forschergruppe an der McMasters Universität im Kontext evidenzbasierter Medizin als Hilfsmittel zur Formulierung evidenzbasierter klinischer Forschungsfragen beschrieben (Guyatt und Rennie 2002):

„Foreground questions—targeted questions that provide the evidentiary basis for specific clinical decisions—are best structured using the framework of patient, intervention or exposure, a possible comparison intervention, and outcomes of interest: the PICO format“

Heute stellt das PICO ein wichtiges Werkzeug bei der Durchführung systematischer Reviews zu interventionellen Fragestellungen dar: Das PICO-Schema erleichtert nicht nur dem Domänenexperten eine strukturierte Formulierung der Forschungsfrage, sondern hilft auch dem Informationsspezialisten bei der Übersetzung des Forschungsgegenstands in eine „suchbare“ Fragestellung.

Tabelle 3-3 PICO-Schema nach (Banfield 2016)

Begriff	Erläuterung und Hilfestellung
Patient, Population, Problem	Beschreibung des Patientenkollektivs
Intervention, Prognostic Factor	Beschreibung der Behandlung sowie prognostischer Faktoren
Control/Comparison	Beschreibung der zu vergleichenden Behandlungsmöglichkeiten
Outcome	Methoden zur Beurteilung von Effekten und Behandlungserfolgen

Ausformuliert folgt eine nach PICO-Schema entwickelte Fragestellung typischerweise folgendem Aufbau:

Führt die Behandlung (*Intervention*) bei einem Patientenkollektiv mit den Charakteristika (*Patient/Population*) zu einem (...) Ergebnis (*Outcome*) im Vergleich zu einer Kontrollgruppe/einer Kontrollbehandlung (*Control*)?

Derivate des PICO-Schemas helfen, auch nicht-interventionelle Studientypen über ein Hilfsschema zu erfassen.

Die Variante PICOC ergänzt das klassische PICO-Schema um das Konzept Kontext (engl. *context*), es wurde 2005 von Petticrew und Roberts vorgeschlagen, um die zu beobachtende Intervention im Kontext des zugrundeliegenden Studiendesigns beschreiben zu können (Petticrew und Roberts 2006). Eine alternative Version ist das PICOT Schema, das zusätzlich einen Zeitfaktor, beispielsweise die Dauer einer Behandlung (*time*), berücksichtigt (Davies 2011).

Eine Reihe weiterer Hilfsschemata nimmt sich sehr spezieller oder vom PICO deutlich abweichender Fragestellungen an. Schlosser et al. erweitern PICO zu PESICO (*Person – Environment – Stakeholders – Intervention – Comparison – Outcome*) als nützliches Schema für Fragen zum Thema Rehabilitation (Schlosser und O'Neil-Pirozzi 2006), Kloda und Bartlett nennen in ihrer Zusammenstellung beispielsweise PECODR (*Patient or Problem or Population, Exposure, Comparison, Outcome, Duration, Results*) (Kloda und Bartlett 2013).

Für die Durchführung systematischer Übersichtsarbeiten zu diagnostischen Testverfahren schlägt das Cochrane Handbuch vor, in die Fragestellung ein Konzept zum diagnostischen Testverfahren und eines zur Beschreibung der Zielbedingungen aufzunehmen (Cochrane Diagnostic Test Accuracy Working Group 2012).

Das Konzept der Hilfsschemata ist grundsätzlich auf andere Themengebiete abseits der Medizin übertragbar, beispielsweise im Software Engineering (Kitchenham 2004):

Tabelle 3-4 Übertragbarkeit des PICO-Schemas auf das Software Engineering nach (Kitchenham 2004)

Begriff	Erläuterung und Hilfestellung
Population	Rolle verschiedener Akteure im Software Engineering Kenntnisstand eines Nutzers oder Programmierers Anwendungsgebiet einer Software
Intervention	typische Probleme, die das Software Engineering adressiert, darunter die Verwendung spezieller Technologien, Systemtests oder Kostenkalkulation
Outcome	Zielvorgaben im Prozess des Software Engineering, beispielsweise Reliabilität, Reduktion von Produktionskosten oder Verringerung von Produktionszeiten

Das PICO-Schema und seine Derivate sind trotzdem prototypisch für die aus der EBM abgeleiteten Methodenlehre der systematischen Übersichtsarbeit: In den medizinischen Fachgebieten zielt die Suche typischerweise auf die Ergebnisse klinischer Studien, idealerweise im randomisierten klinischen Studiendesign, ab. Die Forschungsfrage kann daher hochstrukturiert formalisiert werden, quantitative Ergebnisse identifizierter Studien können dabei außerdem direkt in Meta-Analysen einfließen.

Ein Problem in der Formalisierung des Informationsbedarfs betrifft die Komplexität zugrundeliegender klinischer Fragestellungen, dabei insbesondere zwei Komponenten des PICO-Schemas (siehe Abschnitt 3.3.3): zum einen das *Outcome*, zum anderen die *Intervention*.

Eine klinische, interventionelle Fragestellung wird unter anderem dann als komplex bezeichnet, wenn (Craig et al. 2008):

- sie eine Vielzahl interagierender Komponenten zur Interaktion in Versuchs- und Kontrollgruppe enthält
- eine Vielzahl von Verhaltensweisen bei Ärzten oder Patienten zur Verabreichung oder bei Erhalt einer Intervention möglich macht
- sich die Intervention an eine Vielzahl von Zielgruppen richtet
- oder wenn Interventionen zu einer Vielzahl möglicher Ergebnisse (*Outcomes*) führen können.

Die Lösung für komplexe *Outcomes* ist methodologisch verhältnismäßig einfach. Seitens Cochrane wird empfohlen, klinische Fragestellungen generell nicht zu sehr auf das Konzept *Outcome* zu fokussieren, was einen sehr pragmatischen Grund hat: Die Folgen klinischer Behandlungen werden weitaus seltener in den Titeln und Abstracts genannt als Einzelheiten zum Patientenkollektiv oder zur Behandlungsmethode. Sie sind daher in bibliographischen Fachdatenbanken nicht gesichert recherchierbar. Typischerweise kann das Konzept *Outcome* daher in der Praxis auch entfallen. (Lefebvre et al. 2011)

Nach wie vor problematisch gestaltet sich aufgrund ihrer Homogenität die Formalisierung klinische Fragestellungen mit komplexen Interventionen als notwendiges Element des PICO-Schemas und die Entwicklung von Suchanfragen interventioneller Studien. Komplexe Interventionen finden sich unter anderem in der Verhaltensforschung, siehe hierzu beispielhaft den Systematic Review von Loveman et al. (Loveman et al. 2015) mit dem Thema *Parent-only interventions for childhood overweight or obesity in children aged 5 to 11 years*, aber auch in den Sozialwissenschaften, vgl. (Petticrew 2011). Sie sind demnach für viele Institutionen, die sich der Durchführung systematischer Übersichtsarbeiten verschrieben haben, grundsätzlich eine Herausforderung. Dieser Aspekt nimmt entsprechend Raum in der methodologischen Forschung der systematischen Übersichtsarbeit ein, vgl. beispielsweise (Pawson et al. 2005; Petticrew 2011; Petticrew et al. 2013a; Petticrew et al. 2013b; Squires et al. 2013) und wird im Folgenden eingehend erläutert.

3.3.1.2 Konzeptualisierung qualitativer Fragestellungen

Komplexe Interventionen finden sich insbesondere in den Sozialwissenschaften und angegliederten Fachgebieten. In der Praxis führt hier schon die Vorbereitung systematischer Übersichtsarbeiten häufig zu iterativen Prozessen und demnach auch abweichenden Methoden der Suche. Aus Gründen der Vollständigkeit soll daher kurz dargestellt werden, warum die

Formalisierung mit Hilfsschemata vor allem im Kontext qualitativer Forschung nicht problemlos funktioniert.

Sozialwissenschaftlich sind potenzielle Forschungsfragen nicht wie in der Medizin taxonomisch erfasst und werden unter anderem mit folgendem Ziel durchgeführt (Gough et al. 2012, S. 58):

- Beschreibung oder Analyse von Forschungsgebieten
- Zusammenfassung der Ergebnisse aus Forschungsgebieten
- Identifikation von Wissen zu Forschungsgebieten
- Anwendung von Forschungsergebnissen in verschiedenem Kontext.

Der Umfang möglicher Forschungsfragen zeigt sich entsprechend variabler und führt letztendlich zu anderen Forschungsaktivitäten und auch anderen Forschungsprodukten, darunter Beschreibungen, Messungen, Vergleiche, Beziehungen oder Bewertungen (Gough et al. 2009). Dies macht die Formalisierung des Informationsbedarfs in der Schärfe interventioneller Reviews unmöglich, vgl. (Gough et al. 2012, S. 58).

Üblicherweise wird daher versucht, in iterativen Prozessen Forschungsfrage, Ziel der Arbeit und Forschungstheorien über ein konzeptuelles Framework zu vereinen und in Einzelschritten zu spezifizieren, was zu einer Vielzahl thematischer Hilfsschemata führt, die analog zu PICO durch Akronyme benannt sind, wie oben bereits beschrieben.

PROGRESS steht beispielsweise für Wohnort (*place of residence*), Ethnie oder Kultur (*race, ethnicity, culture, language*), Beruf (*occupation*), Geschlecht (*gender*), Religion (*religion*), Erziehung (*education*), sozioökonomischen Status (*socioeconomic status*) sowie soziales Kapital (*social capital*), es kann genutzt werden, um sozial bedingte Unterschiede im Gesundheitsstatus zu identifizieren (Evans und Brown 2003).

Da die Verwendung des PICO-Schemas in Treffermengen typischerweise zu einer Überrepräsentation quantitativer Studien führt, wie von Cooke et al. ausgeführt, berücksichtigen alternative Suchschemata zur qualitativen wie beispielsweise SPIDER als Derivat des PICO-Schemas auch das Forschungsdesign (Cooke et al. 2012).

Tabelle 3-5 Vergleich PICO und SPIDER Schema (Cooke et al. 2012)

PICO	SPIDER
P = Patient, Population, Problem	S = Sample
I = Intervention, Prognostic Factor	PI = Phenomenon of Interest
C = Control/Comparison	D = Design
O = Outcome	E = Evaluation
	R = Research Type

Eine umfangreiche Übersicht derzeit bekannter Hilfsschemata wie auch Probleme in der Durchführung qualitativer systematischer Reviews findet sich beispielsweise in der umfangreichen systematischen Übersichtsarbeit von Booth et al. (Booth 2016).

3.3.1.3 Zusammenfassung

Ausgehend von den Grundsätzen der evidenzbasierten Medizin lässt sich das – wie bereits einleitend deutlich gemacht – fiktive Ziel einer systematischen Übersichtsarbeit, eine nach *systematischen* Methoden durchgeführte Identifikation, Bewertung und Synthese *aller* verfügbaren Evidenz bezüglich einer vorab strukturierten medizinischen Forschungsfrage, der „*well built clinical question*“ definieren (Richardson et al. 1995). Die systematische Übersichtsarbeit definiert die Formalisierung des Informationsbedarfs daher als notwendigen Einstiegspunkt in die systematische Übersichtsarbeit.

Dieser Prozess ist für interventionelle, qualitative Studien klar definiert und folgt üblicherweise dem in der evidenzbasierten Medizin etablierten PICO-Schema. Derivate des PICO-Schemas erleichtern die Formulierung klar formulierter, strukturierter Fragestellungen für andere Themengebiete. Vor allem für komplexe Interventionen und qualitative Studiendesigns erweist sich die Formalisierung des Informationsbedarfs als Problem. Dies beeinflusst – wie im Folgenden eingehend dargestellt – die Vorgehensweise bei der Identifikation relevanter Dokumente und Studienergebnisse.

3.3.2 Modelle systematischer Literatursuche

Methodologisch unterliegt die Suche dem Anspruch,

- verfügbare Evidenz so umfassend wie möglich aufzufinden (Kriterium der Vollständigkeit)
- systematisch zu erfolgen (Kriterium der Fehlerrobustheit)
- reproduzierbar dokumentiert zu werden (Kriterium der Reproduzierbarkeit).

Sie deckt demnach so viele der verfügbaren Informationsquellen ab wie möglich und erfolgt dort systematisch und mit dem Ziel der Vollständigkeit nach Methoden, die ihre Fehlerrobustheit als auch Reproduzierbarkeit garantieren.

3.3.2.1 Einordnung

In der Fachliteratur zur systematischen Übersichtsarbeit sind Methoden der Suche wiederholt in unterschiedlichen Detailgraden benannt worden.

Tsafnat et al. identifizieren ausgehend vom Cochrane Handbuch und interner Guidelines 15 verschiedene Aufgabenbereiche und betten diesen in den Kontext der vier wesentlichen Elemente der systematischen Übersichtsarbeit – Vorbereitung, Retrieval, Bewertung und Synthese – ein (Tsafnat et al. 2014).

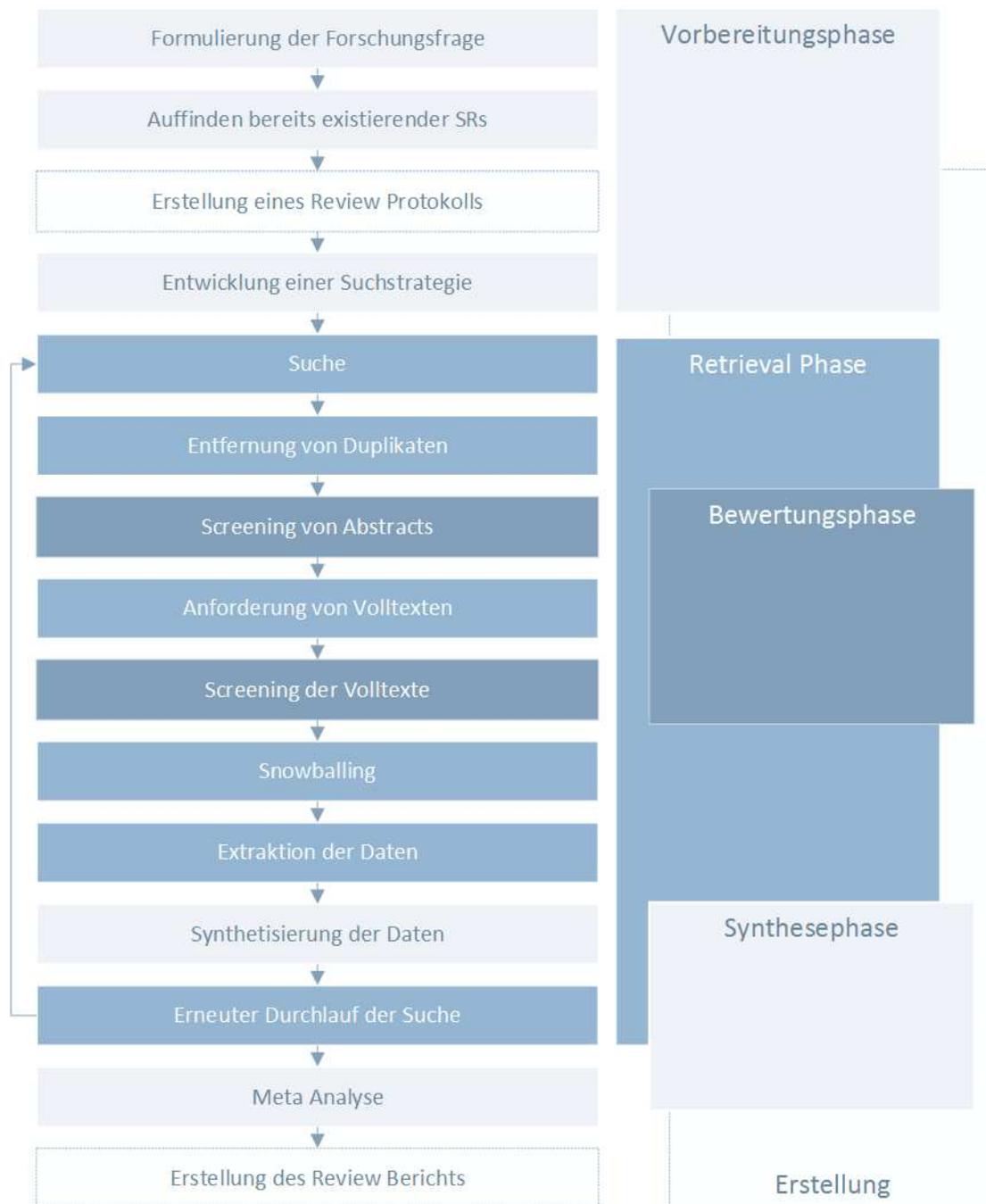


Abbildung 3-3 Schritte zur Durchführung systematischer Übersichtsarbeiten nach (Tsafnat et al. 2014)

Der Retrieval-Prozess umfasst nach Tsafnat et al. unter anderem

- die (systematische) Suche (in verschiedenen Informationsquellen)
- die Entfernung von Duplikaten

- die Anforderung von Volltexten zu allen relevanten, bibliographischen Fundstellen
- weitere Suchprozesse wie bibliographische Suche und Zitationssuche
- ein erneuter Durchlauf der Suche

Er lässt sich als iterativer Prozess wie in Abbildung 3-3 dargestellt als wesentliches Element der systematischen Übersichtsarbeit verorten und wird immer wieder teilweise von der Analyse und Synthese ermittelter Ergebnisse überlagert. Methodisch zentral ist außerdem die *Wiederholung* systematischer Suchprozesse, die durch die hohe Frequenz medizinischer Neuveröffentlichungen wie auch die Dauer systematischer Übersichtsarbeiten notwendig ist, um den Vollständigkeitsanspruch der systematischen Übersichtsarbeit final garantieren zu können (Tsafnat et al. 2014).

Eine andere, weniger prozessorientierte Zusammenstellung von Aufgaben und potenziellen Methoden der systematischen Literatursuche findet sich bei Booth et al.:

Tabelle 3-6 Iterative Aufgaben im Suchprozess der systematischen Literaturrecherche, aus (Booth et al. 2012, S. 72)

Festlegung des Untersuchungsrahmens
Identifikation geeigneter Informationsquellen (Datenbanken, Kataloge)
Identifikation thematischer Konzepte und Suchterme
Entwicklung und Dokumentation einer Suchstrategie
Durchführung der Suche
Suche in allen Datenbanken unter Berücksichtigung bisher identifizierter Suchterme, Freitextsuche sowie Thesauri, Einsatz von Datenbankabfragesprachen
Suche in Grauer Literatur
Einsatz und Bewertung methodologischer Suchfilter
Dokumentation von Änderungen in der Suchstrategie
Bibliographische Suche
Suche in Referenzlisten und Bibliographien aller bereits ermittelten Studien
Identifikation von Schlüsselreferenzen und Durchführung einer Suche über Zitationen

Handsuche in relevanten Journals
Verifikation
Überprüfung hinsichtlich Studien, die von der bisherigen Suchstrategie nicht erfasst wurden
Anpassung der Suchstrategie
Kommunikation mit Domänenexperten zur Überprüfung des Gesamtergebnisses
Dokumentation
Dokumentation eingesetzter Suchquellen, Retrievalergebnisse, Suchstrategien und weiterer Referenzen

Abbildung 3-3 sowie Tabelle 3-6 zeigen zum einen klassische und hochstrukturierte Retrieval-Prozesse wie die Suche in bibliographischen Datenbanken und die formale Suche in klinischen Studienregistern, daneben aber auch die Anwendung typischer Suchtaktiken wie bibliographische Suche und Zitationssuche als formale und thematische Suchprozesse (vgl. hierzu die Berry Picking-Techniken nach Bates in Abschnitt 2.1.2).

Grundsätzlich sind nicht alle der in Abbildung 3-3 sowie Tabelle 3-6 zusammengestellten Schritte erforderlich und werden in der Praxis daher nicht umfassend durchgeführt (vgl. (Booth et al. 2012, S. 72; Tsafnat et al. 2014)), trotzdem werden sie in den meisten methodologischen Handbüchern empfohlen, siehe hierzu im Folgenden die Abschnitte 3.4 und 3.5.

3.3.2.2 Klassisches Modell der systematischen Literatursuche

Wegweisend für die Methodenlehre ist das klassische Modell der Suche, wie es beispielsweise im Handbuch der Cochrane Collaboration (Higgins und Green 2011), hier speziell (Lefebvre et al. 2011), aber auch bei der Campbell Collaboration beschrieben ist.

Cochrane Reviews unterliegen mit dem MECIR Statement einem strengen Protokoll und schreiben folgende Aufgaben bei der Suche nach Studien verpflichtend vor (Chandler et al. 2013):

- Suche in Schlüsseldatenbanken

- Suche in klinischen Studienregistern
- bibliographische Suchen
- Strukturierung von Suchstrategien
- Entwicklung von Suchstrategien für bibliographische Datenbanken
- Dokumentation des Suchprozesses
- Wiederholung von Suchprozessen auf bibliographischen Datenbanken,

und definieren folgende Methoden der Suche als „höchst erwünscht“:

- Suche in spezialisierten Datenbanken
- Suche in Grauer Literatur
- Suche in verfügbaren, weiteren systematischen Übersichtsarbeiten
- Einsatz von Suchfiltern
- Einbeziehung der durch die Wiederholung des Suchprozesses identifizierten Dokumente.

Der Fokus des Suchprozesses liegt auf klassischem Retrieval in Fachdatenbanken, sie sind der einfachste und effektivste Weg, eine initiale Menge relevanter Studienergebnisse zu identifizieren, und führen im Gesamtergebnis üblicherweise zu großen Treffermengen (Lefebvre et al. 2011), die dann sowohl einen Deduplikations,- als auch einen Screening-Prozess durchlaufen müssen, siehe auch Abbildung 3-3.

Als Minimaldatensatz zu durchsuchender Fachdatenbanken nennen einzelne Handbücher (vgl. Abschnitt 1.1.4) neben dem Cochrane Central Register of Controlled Trials (CENTRAL) die beiden bibliographischen Datenbanken MEDLINE und EMBASE sowie eine Reihe nationaler und regionaler Datenbanken. Themenspezifisch ist es erforderlich, weitere Fachdatenbanken angrenzender Fachgebiete wie beispielsweise psychologische Fachdatenbanken zu durchsuchen. Fachspezifisch nennt Campbell analog die Datenbanken ERIC und PsycINFO (Kugley et al. 2015). Entscheidungen über themenspezifische Datenbanken überlassen die Herausgeber der Guidelines den Institutionen und Durchführenden selbst und verweisen dabei auf deren Verfügbarkeit und das für den Review verfügbare Budget (Lefebvre et al. 2011; Kugley et al. 2015).

Die Suche in klinischen Studienregistern als medizinspezifische Domäne sollte laut Cochrane Handbuch (Lefebvre et al. 2011) mindestens in nationalen und internationalen Studienregistern erfolgen, kann aber auf themenspezifische Register und Studienregister pharmazeutischer Konzerne ausgeweitet werden (vgl. Abschnitt 3.2.5).

Zusätzliche Methoden wie die Suche in Grauer Literatur dienen der abschließenden Absicherung der Vollständigkeit strukturierter Suchprozesse.

Ergänzend werden daher Suchen in Journals empfohlen, darunter sowohl elektronisch verfügbare Volltextdatenbanken als auch Conference und Proceedings Datenbanken wie BIOSIS, ein Angebot des Verlags Thomson Reuters, siehe (Thomson Reuters 2016a), vgl. auch Abschnitt 3.2.4. Die Suche in Grauer Literatur (vgl. Abschnitt 3.2.4) umfasst zusätzlich die Suche in Dissertationsdatenbanken, darunter ProQuest Dissertations & Theses Global (ProQuest 2016) und andere.

Zur Identifikation evidenzbasierter Studienergebnisse gehören weitere Methoden: Unter anderem sollten die bibliographischen Einträge verfügbarer Reviews zu gleichen oder ähnlichen Themen geprüft und Wissenschaftler, non-profit Organisationen, ggf. auch Hersteller kontaktiert werden. Ähnliche Empfehlungen finden sich beispielsweise im Leitfaden des Center for Review and Dissemination (Centre for Reviews and Dissemination 2009) oder Leitlinien des National Health and Medical Research Council (National Health and Medical Research Council 2000).

Ein wichtiges Instrument zur Sicherung der Reproduzierbarkeit von Such- und Retrieval-Prozessen im Rahmen systematischer Übersichtsarbeiten ist in der Methodenlehre die so genannte *Suchstrategie*. Dieser iterative Prozess wird, wie auch das Konzept des methodologischen Suchfilters, in Abschnitt 3.4 eingehend beschrieben.

3.3.2.3 Realistische Modelle

So genannte *Realist Reviews* versuchen, die Methodik der systematischen Übersichtsarbeit an die Rahmenbedingungen anzupassen, die sich abseits der klassischen evidenzbasierten Medizin aus der Komplexität von Fragestellungen ergeben und häufig mit qualitativen Fragestellungen assoziiert sind. Zu den Problemen gehören neben der erschwerten Formalisierung des Informationsbedarfs vorwiegend die Verteilung von Originalliteratur zu Studienergebnissen

auf unterschiedliche Informationsressourcen, aber auch die Diversität der Fachterminologie sowie die teils unzureichende thematische Indexierung in Fachdatenbanken (Papaioannou et al. 2010; Pawson et al. 2004).

Reviews zu komplexen Interventionen werden zur Entwicklung von Suchanfragen häufig an eine sogenannte *Scoping Search* gekoppelt, die der schnellen Identifikation von Kernkonzepten als auch der wesentlichen Informationsquellen zur Recherche eines Forschungsgebiets dienen (Pawson et al. 2005):

„Because it deals with the inner workings of interventions, realist review is much more likely to make use of grey literature rather than relying solely on articles in academic journals.“

Scoping Searches haben das Ziel, schnell Schlüsselkonzepte eines Forschungsgebiets, wesentliche Informationsressourcen und verfügbare, evidente Forschungsergebnisse zu identifizieren. *Scoping Searches* sind eigene Projekte und werden vor allem bei komplexen oder bisher nicht eingehend beforschten Themengebieten eingesetzt (Mays et al. 2001). Sie unterliegen weder dem Anspruch der Vollständigkeit, noch der Fehlerrobustheit oder Reproduzierbarkeit (Arksey und O'Malley 2005).

In der systematischen Übersichtsarbeit zu sozialwissenschaftlichen Themen ist eine Mischung verschiedener Suchtaktiken zur Ermittlung von Primärliteratur der Ausarbeitung erschöpfender Suchanfragen überlegen und auch notwendig, um im Ergebnis zu einer umfassenden Literatursuche zu gelangen, darunter vor allem Zitationssuchen und Pearl Growing-Methoden, wie auch der direkte Kontakt zu Forschern. Ein Abbruchkriterium für nicht-strukturierte Suchen in Datenbanken kann in diesem Fall eine über weniger formale Suchmodi erreichte Sättigung des aktuell ermittelten Datenbestands sein. (Papaioannou et al. 2010)

3.3.3 Qualitätsmanagement und Dokumentation

Die Dokumentation des gesamten Suchprozesses hat im klassischen Modell (siehe Abschnitt 3.3.2.2) methodologisch in einem Detailgrad erfolgen, der die korrekte Zusammenfassung in der systematischen Übersichtsarbeit ermöglicht, sie umfasst (Higgins und Green 2011; Kugley et al. 2015):

- Angabe aller durchsuchten Datenbanken, darunter bibliographische Datenbanken, Zitationsdatenbanken und klinische Studienregister
- Datum der letzten Suche und Angabe des gesamten Retrievalprozesses
- Angabe aller durchsuchten Quellen zu Grauer Literatur
- Angabe aller persönlichen Kontakte
- Angabe durchsuchter Journals und Tagungsbeiträge
- Angabe sonstiger durchsuchter Quellen, darunter auch Internetsuchen und bibliographische Suchen

Die Angabe der Queries strukturierter Suchprozesse sind für alle durchsuchten Datenbanken Teil des Protokolls, für Cochrane Reviews ist deren Angabe nach dem MECIR Statement verpflichtend (Chandler et al. 2013).

Im so genannten PRISMA (*Preferred Reporting Items for Systematic Reviews and Meta-Analyses*)-Statement werden die durch Suchprozesse identifizierten Veröffentlichungen und Daten mit zugrundeliegenden Studien zusammengeführt, sodass zum Ende der systematischen Übersichtsarbeit transparent dargelegt werden kann, wie viele klinische Studien in das Ergebnis des Reviews eingeflossen sind und in wie zugehörige Veröffentlichungen identifiziert wurden, siehe Abbildung 3-4.

Nach dem PRISMA-Schema werden zunächst die Anzahl aller durch strukturierte Datenbanksuchen identifizierter Dokumente als auch die Anzahl von „in anderen Quellen“ identifizierter Dokumente aufgelistet. Diesem Prozess folgen üblicherweise eine Deduplizierung als auch ein Screening zum Abgleich aufgefundener Dokumente mit den vorab definierten Ein- und Ausschlusskriterien. Das PRISMA-Schema endet mit der Angabe der Volltexte, die zur Bewertung in die systematische Übersichtsarbeit beschafft wurden und der mit diesen Veröffentlichungen assoziierten Studien, die gegebenenfalls auch in eine Meta-Analyse eingehen. (Liberati et al. 2009; Moher et al. 2009)

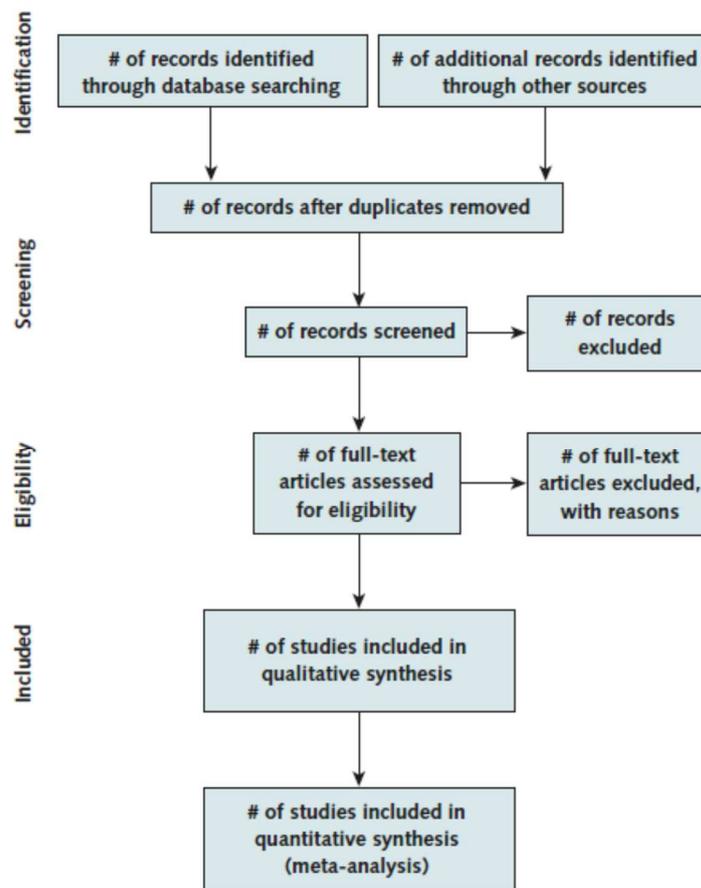


Abbildung 3-4 PRISMA Statement aus (Liberati et al. 2009, S. 68)

Abbildung 3-5 zeigt die Anwendung des PRISMA-Schemas zur Dokumentation der Such- und Bewertungsprozesse, die in der Praxis charakteristischerweise häufig sehr umfangreich ausfallen. In diesem Fall stehen 13769 eindeutigen Dokumenten nach Deduplizierung 38 Publikationen nach dem Screening-Prozess gegenüber, dies entspricht 0,28 Prozent. Gleichzeitig zeigt das PRISMA-Schema anhand der Zahlen auch das Ungleichgewicht strukturierter Prozesse auf Datenbanken und der Suche über andere Informationsquellen oder -kanäle.

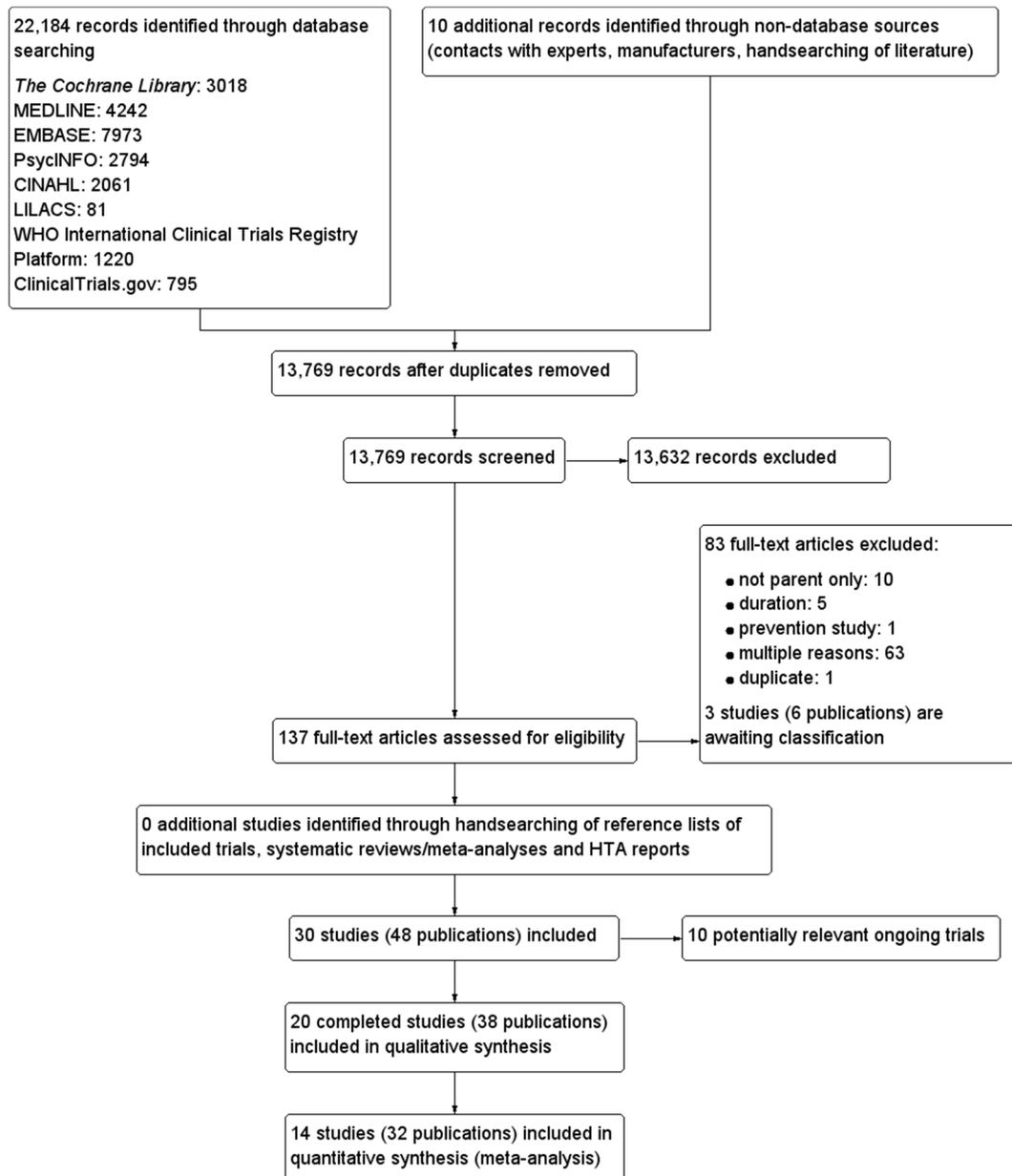


Abbildung 3-5 Dokumentation von Suchprozessen nach PRISMA aus (Loveman et al. 2015, S. 12)

Modelle wie die des Realist Review sind entgegen des klassischen Modells nicht vollständig reproduzier- und dokumentierbar (Pawson et al. 2004).

Die Dokumentation strukturierter Suchprozesse wie auch das PRISMA-Modell sind daher exklusiv auf die klassische systematische Übersichtsarbeit anwendbar und eng verknüpft mit der zugrundeliegenden klinischen Fragestellung.

3.4 Strukturierte Suchprozesse

Die Suche in bibliographischen Fachdatenbanken mit strukturiertem Datenbestand ist trotz der Vielfalt verfügbarer Ressourcen das Herzstück der systematischen Übersichtsarbeit und wird in der Literatur entsprechend umfangreich beschrieben, vgl. u.a. (Lefebvre et al. 2011; Centre for Reviews and Dissemination 2015; Gough et al. 2012). Begrifflich wird sie im Folgenden als strukturierte Suche von formalen und thematischen Suchprozessen (siehe Abschnitt 3.5) abgegrenzt.

3.4.1 Begriffliche Einordnungen

Kern der bibliographischen Suche ist die Entwicklung von *Suchstrategien*. Da der Terminus in der Fachliteratur häufig zu finden ist, aber anders verwendet wird als in der Informationswissenschaft üblich, ist eine neuerliche begriffliche Einordnung nötig.

Im PICO-Schema (siehe Abschnitt 3.3.1.1) findet sich klar sichtbar die Suchstrategie der Building Blocks (vgl. Abschnitt 2.4.1) wieder.

Suchstrategien lassen sich – sicherlich der historischen Entwicklung geschuldet – entsprechend der terminologischen Verwendung in verfügbaren Normenwerken (vgl. (Higgins und Green 2011; Centre for Reviews and Dissemination 2009; Kugley et al. 2015)) als strukturierte, datenbankspezifische Suchanfragen und damit als *Struktur* oder auch ein *Produkt* (vgl. Abschnitt 2.4.1) verstehen:

P_1, P_2, P_3, \dots Textwörter, Synonyme, Phrasen und kontrolliertes Vokabular zu Block Population:
(P_1 OR P_2 OR ... P_N)

I_1, I_2, I_3, \dots Textwörter, Synonyme, Phrasen und kontrolliertes Vokabular zu Block Intervention:
(I_1 OR I_2 OR ... I_N)

C_1, C_2, C_3, \dots Textwörter, Synonyme, Phrasen und kontrolliertes Vokabular zu Block Control: (C_1 OR C_2 OR ... C_N)

O_1, O_2, O_3, \dots Textwörter, Synonyme, Phrasen und kontrolliertes Vokabular zu Block Outcome:
(O_1 OR O_2 OR ... O_N)

Klassischerweise folgt die fertiggestellte Suchanfrage also vereinfacht der Formel

$$(P_1 \text{ OR } P_2 \text{ OR} \dots P_N) \text{ AND } (I_1 \text{ OR } I_2 \text{ OR } \dots I_N) \text{ AND } (C_1 \text{ OR } C_2 \text{ OR } \dots C_N) \text{ AND } (O_1 \text{ OR } O_2 \text{ OR } \dots O_N)$$

oder entsprechenden PICO-Varianten.

Begrifflich ist der so verstandene Terminus *Suchstrategie* in der Methodenlehre systematischer Übersichtsarbeiten weit verbreitet (Lefebvre et al. 2011; Kugley et al. 2015).

Abseits bibliographischer Fachdatenbanken zeigt sich die Menge an Informationsressourcen und recherchierbaren Dokumenttypen zunehmend inhomogen, die Suche nach Evidenz wird zwar entsprechend der Evidenzpyramide von oben nach unten und damit absteigend ihres Evidenzgrads durchgeführt. Sie beschränkt sich heute notwendigerweise nicht mehr nur auf Fachdatenbanken, sondern muss auf weitere Informationsressourcen wie klinische Studienregister oder Graue Literatur ausgedehnt werden. Aus dem Anspruch der systematischen Übersichtsarbeit heraus muss die Suchstrategie jedoch für alle diese Quellen dokumentiert werden (Chandler et al. 2013), darunter auch solche, in denen klassisches Retrieval nicht möglich ist oder Suchprozesse nicht wie bei Booleschen Fachdatenbanken dem Alles-oder-Nichts Prinzip folgen und damit Treffermengen nicht klar reproduzierbar sind – beispielsweise bei Suchmaschinen.

Daher finden sich in der Fachliteratur freiere Beschreibungen zum Begriff *Suchstrategie*, die diesem Sachverhalt gerecht werden, vgl.(Egger und Smith 2001) und (Gough et al. 2012, S. 155)).: In einer Suchstrategie werden alle zu durchsuchenden Datenbanken gelistet, die gesamte Menge an Termen zusammengestellt und über Boolesche Operatoren definiert, wie diese Terme zu ressourcenspezifischen Suchanfragen kombiniert werden.

Festzuhalten ist, dass in der Methodenlehre der Begriff *Suchstrategie* in der Fachliteratur zur evidenzbasierten Medizin immer dann im Sinne einer Query verstanden werden sollte, wenn Untersuchungen zum Auffinden der Evidenz in Fachdatenbanken und der Optimierung dieser Datenbankabfragen angestellt werden.

Für die folgende Arbeit gilt weiterhin die informationswissenschaftlich etablierte Definition im Sinne von Bates (siehe Abschnitt 2.4.1). Dies entspricht unserer gängigen Auffassung des Begriffs einer Strategie, die im Deutschen definiert ist als *„genauer Plan des eigenen Vorgehens, der dazu dient, ein militärisches, politisches, psychologisches, wirtschaftliches o. ä. Ziel zu erreichen, und in dem man diejenigen Faktoren, die in die eigene Aktion hineinspielen könnten, von vornherein einzukalkulieren versucht“* (Bibliographisches Institut GmbH Duden Verlag 2016g).

In diesem Sinne ist die Suchstrategie als Planung und Dokumentation des gesamten Suchprozesses zu verstehen und beinhaltet nicht nur die Formulierung von Datenbankabfragen, sondern auch die Überlegung, welche Informationsquellen für den gesamten Suchprozess herangezogen werden. Als ein systematischen Übersichtsarbeit übergeordneter Plan berücksichtigen Suchstrategien (Gough et al. 2012, S. 116):

- das Ziel der Suche sowie die Sicherung bestmöglicher Methoden
- die Informationsquellen, die für die Suche herangezogen werden
- Vor- und Nachteile der Suche in einzelnen Dokumentenbeständen
- Verfügbarkeit von Ressourcen
- Kosten- und Nutzenrechnung der Auswahl zu durchsuchender Informationsquellen

Einen wesentlichen Bestandteil von Suchstrategien systematischer Übersichtsarbeiten bilden Suchprofile. Suchprofile definieren alle für die Suche auf Fachdatenbanken zu verwendenden Freitextterme und Terme aus kontrolliertem Vokabular anhand der durch die Fragestellung vorgegebenen Konzepte und bilden damit das Grundgerüst für datenbankspezifische Queries (siehe hierzu Tabelle 3-12). In der finalen Dokumentation des gesamten Suchprozesses werden Suchprofile im Protokoll mit einer Liste aller durchsuchten Fachdatenbanken zusammengeführt und gelistet.

3.4.2 Datenbankspezifika

Die Suche in bibliographischen Datenbanken ist grundsätzlich im Kontext von Datenbankspezifika zu sehen, darunter zum einen datenbankspezifische Thesauri und Metadaten, Syntaxspezifika, aber auch die Funktionalitäten, die einzelne Anbieter für die Recherche zur Verfügung stellen.

3.4.2.1 Thesauri und Metadaten

MEDLINE und EMBASE gelten als die Schlüsseldatenbanken für die bibliographische Suche (Lefebvre et al. 2011). Sie verfügen wie alle großen medizinischen Fachdatenbanken über Thesauri zur Verschlagwortung der dort vorgehaltenen Dokumente.

Die Medical Subject Headings MeSH (siehe Abschnitt 2.3.2.2) spielen als Thesaurus eines nicht-kommerziellen Anbieters eine große Rolle in der Welt der medizinischen Fachdatenbanken und sind damit in großen Fachdatenbanken teils implementiert, teils historisch übernommen. Sie sind das Verschlagwortungsinstrument für alle Produkte der National Library of Medicine, unter anderem MEDLINE, Pubmed und Pubmed Central. Als Unterstützung der medizinischen Fachrecherche werden die Medical Subject Headings auch außerhalb der NLM eingesetzt. Die Cochrane Library verschlagwortet in ihren bibliographischen Datenbanken mit MeSH, das DIMDI veröffentlicht eine jährlich aktualisierte, deutschsprachige Version der Medical Subject Headings auf ihren Websites (DIMDI: Deutsches Institut für Medizinische Dokumentation und Information 2016b).

Die Datenbank EMBASE verfügt mit EMTREE über einen eigenen Thesaurus (Embase: Indexing with Life Science Thesaurus Emtree | Elsevier 2015), der ebenfalls nach hierarchischem Prinzip aufgebaut ist. Derzeit enthält EMTREE über 70.000 so genannter „*preferred terms*“, also Deskriptoren, davon 31.000 mit pharmazeutisch-chemischem Schwerpunkt. Diese werden im EMTREE Thesaurus als *Major Terms* bezeichnen. Mit diesen sind über 290.000 Synonyme (*Minor Terms*) assoziiert. EMTREE verfügt zudem über spezifische Terme für allgemeine und medizinische Geräte sowie mehrere tausend Terme für medizinischen Verfahren. Ähnlich dem MeSH-Thesaurus fungieren spezielle Terme als *Subheadings*, mit deren Hilfe Datensätze konzeptuell beschrieben werden können. EMTREE kennt derzeit unter anderem 64 mit Medikamenten assoziierte und 14 mit Krankheitsbildern assoziierte Subheadings.

MeSH und EMTREE unterscheiden sich trotz ihres konzeptionell ähnlichen Aufbaus in einigen Aspekten wesentlich (Elsevier R&D Solutions 2015):

- EMTREE nutzt im Gegensatz zu MeSH natürlichsprachige Terminologien (*myeloid leukemia vs. Leukemia, myeloid*)
- EMTREE ist mit ca. 300.000 Einträgen größer als MeSH mit ca. 200.000
- EMTREE verfügt über mehr Schlagworte zur Indexierung von Medizinprodukten und entsprechend seines Schwerpunkts auch zu Pharmazeutika
- MeSH nutzt so genannte „Scope Notes“, um die Vergabe von Deskriptoren zu erläutern
- EMTREE wird in Gegensatz zu MeSH dreimal im Jahr statt lediglich einmal jährlich aktualisiert
- MeSH verfügt über eine umfangreiche Menge an Deskriptoren zu den Themen Pflege, Veterinärmedizin und Zahnheilkunde.

3.4.2.2 Syntaxspezifika

Wie viele andere Fachdatenbanken basieren auch MEDLINE und EMBASE auf dem klassischen Konzept des Booleschen Retrieval (vgl. Abschnitt 2.3.3). Eine klassische Suchanfrage für drei Konzepte A, B, C folgt demnach beispielhaft dem Muster:

$(A_1 \text{ OR } A_2 \text{ OR } \dots A_N) \text{ AND } (B_1 \text{ OR } B_2 \text{ OR } \dots B_N) \text{ AND } (C_1 \text{ OR } C_2 \text{ OR } \dots C_N)$

mit Suchtermen A_i, B_i, C_i in Form von Deskriptoren als auch Freitexttermen.

Für die Suche gelten dabei wichtige, datenbankspezifische Unterschiede sowohl bei der Verwendung von Deskriptoren als auch Freitexttermen und Phrasen.

Verwendung von Schlagwörtern

Die Suche mit Deskriptoren folgt der datenbankspezifischen Syntax. Hierbei ist, verdeutlicht am Beispiel der Suche MEDLINE via Ovid und MEDLINE via Pubmed einiges zu beachten (Motschall 2014):

Bei der Suche mit Pubmed wird der Hierarchiebaum des Thesaurus automatisch nach unten mit durchsucht, während bei der Nutzung von OVID die Suche im Hierarchiebaum untergeordneter Schlagwörter über das Schlüsselwort *exp* explizit angefordert werden muss. Umgekehrt erfordert die Restriktion auf die gewünschte Hierarchieebene bei Pubmed die Angabe *noexp*.

Gleichzeitig variiert auch die Schreibweise: Pubmed erfordert bei Mehrwort-Deskriptoren sowohl Anführungszeichen als auch die zusätzliche Angabe [MeSH] oder auch [mh], bei OVID ist der abschließende Slash / ausreichend, alternativ .sh..

Tabelle 3-7 Schlagwortsuche MEDLINE OVID, MEDLINE Pubmed (getestet am 25.02.2016)

MEDLINE OVID	MEDLINE Pubmed
Myocardial Infarction/ äquivalent Myocardial Infarction.sh. (Treffer: 148502)	"myocardial infarction"[MeSH: noexp] äquivalent "myocardial infarction"[mh:noexp] (Treffer: 148168)
exp Myocardial Infarction/ (Treffer: 152917)	"myocardial infarction"[MeSH] äquivalent "myocardial infarction"[mh] (Treffer: 152576)

Freitextterme und Phrasen

Die Suche mit Freitexttermen erfolgt über spezifische Datenbankfelder einzelner bibliographischer Dokumente und kann über die zusätzliche Angabe dieser Suchfelder als so genannte *Qualifier* auch auf das Vorkommen in diesen Feldern eingeschränkt werden. Zum beispielhaften Verständnis des Retrievals ist ein kurzer Einblick in die Datenstrukturen hilfreich.

Die wesentlichen Felder zur Klassifizierung von Einträgen finden sich direkt auf den Dokumentationsseiten der NLM (U.S. National Library of Medicine 2014b), in der folgenden Tabelle findet sich ein grober Auszug daraus, sofern mit Suchfeldern identisch – diese wiederum finden sich im Pubmed Hilfesystem (National Center for Biotechnology Information (US) 2014). Es ist zu beachten, dass sich Datenbankfelder im Laufe der Jahre geändert haben, diese Historie ist in der folgenden Tabelle jedoch nicht wiedergegeben.

Tabelle 3-8 Auszug aus den wichtigsten Datenbankfeldern der Datenbank MEDLINE, entsprechend (U.S. National Library of Medicine 2014b) und (National Center for Biotechnology Information (US) 2014), alphabetische Reihenfolge

Datenbank Feld (Abkürzung)	Beschreibung
Abstract (AB)	Englischsprachige Abstracts – und nur diese -werden direkt aus den publizierten Originalartikeln übernommen, sofern verfügbar. Anderenfalls bleibt das Feld leer. Dieses Feld ist beschränkt auf 400 Zeichen.
Author (AU)	Autoren des Originalartikels im Format NACHNAME Initiale1 [Intiale2] Suffix
Journal Title Abbreviation (TA)	Standardabkürzung des Journal Titels, z.B. JAMA (<i>Journal of the American Medical Association</i>)
MeSH Terms (MH)	MeSH Term mit Identifikationsmöglichkeit von Haupt- und Untertermen, z.B. MH - Cardiovascular Diseases/etiology/*mortality
MeSH Subheadings (SH)	Spezifikation eines kontrollierten Terms als Subheading
Publication Date (DP)	Publikationsdatum des Originalartikels im Format yyyy/mm/dd oder auch in Bereichen wie „last X days“
PMID (PMID)	Pubmed Identifikation-ID (PMID), 1-8-stellige eindeutige Nummer zur Identifikation des Datensatzes.
Publication Type (PT)	Publikationstyp
Title (TI)	Englischsprachiger, kompletter Titel des Originalartikels. Titel nicht-englischsprachiger Originalartikel werden in [] angegeben. Bereits aufgenommene, aber noch nicht übersetzte Titel sind ebenfalls in [] angegeben.

Will man Datenbankfelder in die Freitextsuche einbeziehen, so unterscheiden sich die Suchanfragen von MEDLINE OVID und MEDLINE Pubmed (Motschall 2014):

Tabelle 3-9 Freitextsuche MEDLINE Ovid und MEDLINE Pubmed (getestet am 25.02.2016)

MEDLINE OVID	MEDLINE Pubmed
myocardial infarction.ti. (Treffer: 59874)	"myocardial infarction"[ti] (Treffer: 63847)
myocardial infarction.ab. (Treffer: 111826)	"myocardial infarction"[ab] (Treffer: 204760)
myocardial infarction.mp. (Treffer: 193769)	"myocardial infarction"[tw] (Treffer: 204707)

Die Angaben .mp. (Ovid) und [tw] (Pubmed) lokalisieren dabei als sogenannte *Qualifier* den Suchterm in einer Reihe von Feldern.

Neben der einfachen Wortsuche mit Einzelworttermen kann bei Pubmed eine Phrasensuche durchgeführt werden, bei der zwei oder mehrere zusammengehörige Wörter in doppelten Anführungsstrichen verknüpft werden, z.B. "*myocardial infarction*". Pattern-basierte Suche ist über Pubmed nur rudimentär möglich, lediglich die Trunkierung durch * im Sinne eines Wildcard Operators wird von Pubmed unterstützt, die Endtrunkierung ist im Gegensatz zu Ovid auf maximal 600 Variationen begrenzt. Dies bedeutet, dass dem Nutzer nur 600 von gegebenenfalls auch deutlich mehr verfügbaren Variationen angezeigt werden (U.S. National Library of Medicine 2015c).

Via Ovid kann MEDLINE mit PROXIMITY-Operatoren durchsucht werden, dies ist über Pubmed nicht möglich.

3.4.2.3 Funktionalität

Das Retrieval in fachspezifischen Informationssystemen ist, wie bereits in Kapitel 2 beschrieben, stark geprägt durch die zugrundeliegenden Information Retrieval-Technologien, in allen bibliographischen Fachdatenbanken, also durch Boolesches Retrieval.

Da die bedeutenden medizinischen Fachdatenbanken nicht nur über den Anbieter, sondern auch verschiedene Hosts recherchierbar sind, sind die Recherchemöglichkeiten als auch das Retrieval zusätzlich abhängig von Benutzeroberflächen und Funktionalitäten, die einzelne Anbieter oder Vertriebe ihren Nutzern zur Verfügung stellen. Was diese Rahmenbedingungen in der Praxis für die Recherche in Fachdatenbanken bedeuten, wird im Folgenden am Beispiel

MEDLINE im Vergleich zwischen dem Zugriff über Ovid und Pubmed dargestellt. Wie bereits bemerkt, kann MEDLINE sowohl über Pubmed als auch den Anbieter Ovid recherchiert werden, gleichzeitig können Teile der Datenbank MEDLINE auch über EMBASE durchsucht werden. Ovid bietet dabei unter anderem ebenso Zugriff auf die Datenbank EMBASE an (siehe Abschnitt 3.2.1).

Die Interaktion mit dem Retrieval-System erfolgt, wie in Abschnitt 2.5.1 erläutert, typischerweise auf mehreren Ebenen, sie umfasst

- Eingabe und Modifikation von Suchanfragen
- Analyse von Treffermengen
- Kontrolle des Suchverlaufs

Bibliographische Fachdatenbanken unterstützen dies mit verschiedenen Funktionen.

Eingabe und Modifikation von Suchanfragen

Sowohl Pubmed als auch Ovid bieten ihren Nutzern zwei verschiedene Suchmodi: eine einfache Suche und einen professionellen Suchmodus – ein Wechsel zwischen beidem ist jederzeit möglich. Die folgende Abbildung zeigt den einfachen Sucheingabemodus aus Texteingabefeld und Absende-Button.



Abbildung 3-6 Pubmed einfache Suche (25.02.2016)

Die einfache Suche wird bei Pubmed unterstützt durch Termvorschläge aus dem kontrollierten Vokabular wie in folgender Abbildung dargestellt, bei Ovid erfolgt dies auf explizite Anforderung hin.

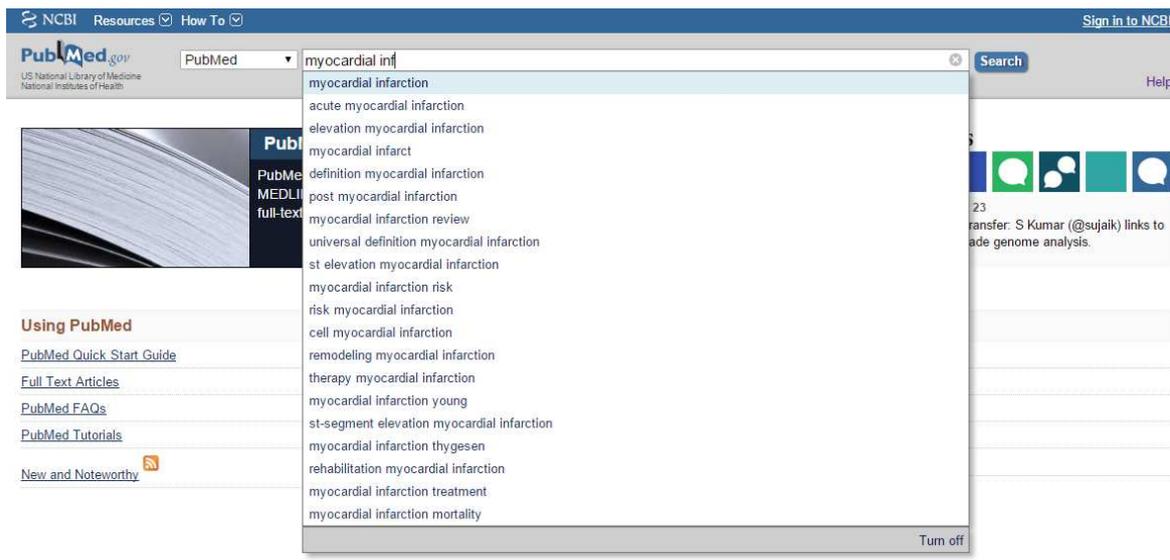


Abbildung 3-7 Pubmed Term-Vorschläge im einfachen Suchmodus (25.02.2016)

Die Termvorschläge (*query suggestions*) werden dabei bei Pubmed direkt aus den Query-Logs der täglich über 2 Millionen abgesetzten Suchanfragen ermittelt, man will dabei in erster Linie den Domänenexperten entgegenkommen, die üblicherweise ihre Suchanfragen häufig modifizieren, bevor sie das Dokument finden, das für ihren Informationsbedarf relevant ist. (Lu et al. 2009):

“It is common for PubMed users to repeatedly modify their queries (search terms) before retrieving documents relevant to their information needs. To assist users in reformulating their queries, we report the implementation and usage analysis of a new component in PubMed called Related Queries, which automatically produces query suggestions in response to the original user’s input.”

Zusätzlich geht es darum, die in den Pubmed Query Logs ermittelten Fehlerquellen wie Tippfehler, Fehler in der Schreibweise, Anfragen mit leeren Treffermengen sowie die Verwendung fremdsprachlicher, englischsprachig nicht gültiger Symbole zu korrigieren (Lu et al. 2009).

Pubmed als auch Ovid bieten neben der einfachen Suche für die Expertensuche auch einen Profi-Suchmodus an, dessen Funktionsweise in Abschnitt 3.4.2 näher erläutert wird:

Use the builder below to create your search

[Edit](#) [Clear](#)

Builder

All Fields [Show index list](#)

AND All Fields [Show index list](#)

[Search](#) or [Add to history](#)

History [Download history](#) [Clear history](#)

Search	Add to builder	Query	Items found	Time
#1	Add	Search myocardial infarction Sort by: PublicationDate	210470	05:29:36

Abbildung 3-8 Pubmed angewandte Suche (25.02.2016)

Das Absenden von Suchanfragen führt vom einfachen als auch vom professionellen Suchmodus aus zur webbasierten Anzeige einer mit der Query korrespondierenden Treffermenge. Die Anzeige der Treffer erfolgt Seiten-basiert, wie wir es auch von Suchmaschinen kennen. Ähnlich wie bei der klassischen Websuche kann die angezeigte Treffermenge dabei je nach Spezifikation der Suchanfrage sehr umfangreich werden. Hierbei ist zu beachten, dass Retrieval in der Booleschen Suche zu einer zwar grundsätzlich sortierten (bei Pubmed ist dies das Publikationsdatum), aber nicht-gerankten Treffermenge führt. Sie kann vom Nutzer mit weiteren Werkzeugen nach weiteren Kriterien sortiert, eingeschränkt und damit manipuliert werden.

Ein typisches Ergebnis einer Suchanfrage zeigt Abbildung 3-9. Die Ergebnisse werden hier seitenweise ausgegeben, per Default sind das bei Pubmed 20 Ergebnisse pro Seite, bei Ovid 10 Ergebnisse.

Die systematische Übersichtsarbeit

The screenshot shows the PubMed search interface. At the top, there's a search bar with 'myocardial infarction' entered. Below the search bar, there are options for 'Create RSS', 'Create alert', and 'Advanced'. The search results are displayed in a list format, with the first five results visible. Each result includes a title, authors, journal name, date, and PMID. The results are sorted by 'Publication Date'. On the left side, there are various filters and options, such as 'Article types', 'Text availability', 'PubMed Commons', 'Publication dates', 'Species', and 'Clear all'. On the right side, there are additional filters and options, such as 'Filters: Manage Filters', 'New feature', 'Results by year', 'Find related data', 'Search details', and 'Recent Activity'.

NCBI Resources How To Sign in to NCBI
PubMed.gov PubMed myocardial infarction Search Help
US National Library of Medicine National Institutes of Health
Create RSS Create alert Advanced

Article types
Clinical Trial
Review
Customize ...

Text availability
Abstract
Free full text
Full text

PubMed Commons
Reader comments
Trending articles

Publication dates
5 years
10 years
Custom range...

Species
Humans
Other Animals

Clear all
Show additional filters

Summary 20 per page Sort by Publication Date Send to: Filters: Manage Filters

Search results
Items: 1 to 20 of 210470 << First < Prev Page 1 of 10524 Next > Last >>

1. [Graphene quantum dots FRET based sensor for early detection of heart attack in human.](#)
Bhatnagar D, Kumar V, Kumar A, Kaur I.
Biosens Bioelectron. 2016 May 15;79:495-9. doi: 10.1016/j.bios.2015.12.083 [2]. Epub 2015 Dec 25.
PMID: 26748366
[Similar articles](#)

2. [China Patient-centered Evaluative Assessment of Cardiac Events Prospective Study of Acute Myocardial Infarction: Study Design.](#)
Li J, Dreyer RP, Li X, Du X, Downing NS, Li L, Zhang HB, Feng F, Guan WC, Xu X, Li SX, Lin ZQ, Masoudi FA, Spertus JA, Krumholz HM, Jiang LX; China PEACE Collaborative Group.
Chin Med J (Engl). 2016 5th Jan;129(1):72-80. doi: 10.4103/0366-6999.172596 [2].
PMID: 26712436 Free Article
[Similar articles](#)

3. [Discharge Communication in Patients Presenting to the Emergency Department With Chest Pain: Defining the Ideal Content.](#)
Ackermann S, Heierle A, Bingisser MB, Hertwig R, Padiyath R, Nickel CH, Langewitz W, Bingisser R.
Health Commun. 2016 May;31(5):557-65. doi: 10.1080/10410236.2014.979115 [2]. Epub 2015 Oct 26.
PMID: 26503453
[Similar articles](#)

4. [Glucose and Inflammatory Cells Decrease Adiponectin in Epicardial Adipose Tissue Cells: Paracrine Consequences on Vascular Endothelium.](#)
Fernández-Trasancos Á, Guerola-Segura R, Paradela-Dobarro B, Álvarez E, García-Acuña JM, Fernández ÁL, González-Juanatey JR, Eiras S.
J Cell Physiol. 2016 May;231(5):1015-23. doi: 10.1002/jcp.25189 [2]. Epub 2015 Sep 29.
PMID: 26406271
[Similar articles](#)

5. [No evidence for a J-shaped curve in treated hypertensive patients with increased cardiovascular risk: The VALUE trial.](#)
Kjeldsen SE, Berge E, Bangalore S, Messerli FH, Mancina G, Holzhauser B, Hua TA, Zappe D, Zanchetti A, Weber MA, Julius S.
Blood Press. 2016 Apr;25(2):83-92. doi: 10.1109/08037051.2015.1106750 [2]. Epub 2015 Oct 29.

New feature
Try the new Display Settings option - Sort by Relevance

Results by year
Download CSV

Find related data
Database: Select
Find items

Search details
"myocardial infarction"[MeSH Terms] OR ("myocardial"[All Fields] AND "infarction"[All Fields]) OR "myocardial infarction"[All Fields]
Search See more...

Recent Activity
Turn Off Clear
Q myocardial infarction (210470) PubMed
Q "myocardial infarction"[all] (204760) PubMed
Q "myocardial infarction"[tw] (204707) PubMed
Q "myocardial infarction"[All Fields] (204760) PubMed
Q "myocardial infarction"[ti] (63847) PubMed

Abbildung 3-9 Pubmed Suche mit angezeigten Treffern (25.02.2016)

Ein Werkzeug zur Einschränkung und Organisation der Treffermenge auf Pubmed bildet beispielsweise die Facettensuche (siehe Abbildung 3-9 linke Menüleiste), zur Verwendung des Terminus vgl. (The PubMed Health Team 2013).

Die Facettensuche ermöglicht es Nutzern, sich explorativ durch den (in diesem Fall gesuchten) Datenbestand zu klicken und damit die Treffermenge einzuschränken. Die typischen Filterattribute einer Facettennavigation ergeben sich dabei direkt aus der Struktur des Datenbestands, sie folgen entgegen einer typischen Facettenklassifikation keiner hierarchischen Ordnung. Die Facettensuche auf der Datenbank Pubmed erlaubt die Einschränkung nach Artikeltyp, nach Verfügbarkeit eines Volltexts, nach Publikationsdatum oder nach Spezies – Fragestellungen in der Medizin zielen häufig auf eine Einschränkung auf klinische Studien am Menschen oder im Tierversuch ab, vgl. hierzu auch „*faceted nomenclature*“ bei Stock. (Stock et al. 2013, S. 713)

Die einfache Pubmed Suche funktioniert für den Laien auch ohne die Verwendung Boolescher Operatoren als reine Freitext-Suche.

Wichtig zum Verständnis der angezeigten Treffermenge bei diesen Pubmed Suchanfragen mit verschiedenen, nicht durch Boolesche Operatoren verknüpften Suchtermen ist deren Pubmed-interne „Übersetzung“. Dieses so genannte *automatic term mapping* (ATM) ist dargestellt in Tabelle 3-10. Aus informationswissenschaftlicher Sicht ist dieses Verfahren mit einer klassischen *Query Expansion* gleichzusetzen, vgl. (Manning et al. 2008, S. 191), siehe hierzu auch Abschnitt 2.3.2.

In der Praxis nutzt Pubmed ohne konkrete Anforderung interne Suchfelder, insbesondere Deskriptoren und verknüpft diese mit den Suchtermen des Nutzers zu einer erweiterten Booleschen Query. Hierfür wird für Textwortterme zunächst ein Abgleich mit dem Thesaurus versucht. Schlägt dies fehl, so werden Terme und Phrasen mit Journalnamen abgeglichen, ebenso mit Autoren und Auftraggebern (U.S. National Library of Medicine 2015a).

Tabelle 3-10 Suchanfrage bei Pubmed: Nutzeranfrage und Pubmed-interne Verarbeitung mit ATM (durchgeführt am 16.06.2016)

Suchanfrage Nutzer/Pubmed-Intern
myocardial infection
("myocardium"[MeSH Terms] OR "myocardium"[All Fields] OR "myocardial"[All Fields]) AND ("infection"[MeSH Terms] OR "infection"[All Fields])

Die Anzeige der Treffer erfolgt standardmäßig rückwärtschronologisch nach Erstellungsdatum (des Datensatzes – nicht der Originalpublikation), neu ist 2013 eine Sortierung nach Relevanz hinzugekommen – ein Konzept, das mit dem Booleschen Retrieval zunächst nicht verträglich scheint und aus dem *extended Boolean Retrieval* bekannt ist, siehe Abschnitt 2.3.2.3.

Über den konkreten Algorithmus des Pubmed Rankings ist nichts Näheres bekannt, die Autorin äußert sich lediglich folgendermaßen (Canese 2015):

"The relevance sort order for search results is based on an algorithm that analyzes each PubMed citation that includes the search terms. For each search query, "weight" is calculated for citations depending on how many

search terms are found and in which fields they are found. In addition, recently-published articles are given a somewhat higher weight for sorting [...]”

Letztendlich wird mit Ranking-Funktionalitäten der Screeningprozess ermittelter Treffermengen für den Laien zwar deutlich erleichtert, der professionelle Nutzer jedoch greift bei Verwendung des Relevance Ranking allerdings auf Funktionen zurück, die den Suchprozess aus methodologischer Sicht intransparenter machen.

Ein für die systematische Literatursuche wichtige Funktion ist die Verarbeitung von Treffermengen zur weiteren Nutzung in Literaturverwaltungsprogrammen und zur späteren Auswertung: Treffermengen beliebiger Suchanfragen können üblicherweise in verschiedenen Formaten zur Übernahme in Literaturverwaltungsprogramme exportiert werden. Bei Pubmed erfolgt dies bequem über einen Klick, Ovid begrenzt die Anzahl gleichzeitig speicherbarer Ergebnisse und macht daher die schrittweise Speicherung über eine spezielle Syntax notwendig (Motschall 2010).

Manipulation und Kontrolle der Suchhistorie

Der so genannte „*PubMed Advanced Search Builder*“ wie auch die angewandte Suche über OVID ermöglichen eine angewandte Suche über ein benutzerfreundliches Webinterface, beispielsweise durch die Zusammenstellung einzelner bereits getätigter Suchanfragen aus der Suchhistorie mittels Boolescher Operatoren.

Über den so genannten Builder (bei OVID „Suchverlauf“) kann der Nutzer (übrigens auch ohne Kenntnis spezifischer Datenbankfeld-Benennungen) zusätzlich konkretisieren, in welchen Feldern einzelne Suchterme oder Phrasen vorkommen dürfen und anschließend einzelne Suchanfragen mittels Boolescher Operatoren zusammenfügen. Dabei können neben neuen Suchphrasen auch bereits getätigte Suchanfragen aus der Suchhistorie beliebig ausgeführt und komplett in neue Queries übernommen werden (siehe Abbildung 3-10 und Abbildung 3-11).

Use the builder below to create your search

[Edit](#) [Clear](#)

Builder

All Fields [Show index list](#)

AND All Fields [Show index list](#)

or [Add to history](#)

History [Download history](#) [Clear history](#)

Search	Add to builder	Query	Items found	Time
#6	Add	Search (#1 OR #5)	22168	07:52:10
#5	Add	Search ("Acupuncture"[Mesh]) OR "Acupuncture Therapy"[Mesh]	18053	07:51:46
#1	Add	Search ("acupuncture") OR "acupuncture therapy"	21113	07:49:28

Abbildung 3-10 Pubmed Advanced Search (31.03.2015)

Ovid® Wolters Kluwer

Mein Account Mein PayPerView Support & Training Hilfe Abmelden

Suche Zeitschriften A-Z Liste Ihrer Bücher bei Ovid Meine Multimedia-Ressourcen Mein Arbeitsbereich

Suchverlauf (7 Suchen) (schließen) Gespeicherte Suchen

<input type="checkbox"/>	#	Suchen	Ergebnisse	Suchtyp	Aktionen
<input type="checkbox"/>	7	3 or 6	19779	Erweitert	Anzeigen Mehr >
<input type="checkbox"/>	6	4 or 5	18551	Erweitert	Anzeigen Mehr >
<input type="checkbox"/>	5	Acupuncture Therapy.mp. [mp=title, abstract, original title, name of substance word, subject heading word, keyword heading word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier]	12309	Erweitert	Anzeigen Löschen Mehr >
<input type="checkbox"/>	4	Acupuncture.mp. [mp=title, abstract, original title, name of substance word, subject heading word, keyword heading word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier]	18551	Erweitert	Anzeigen Löschen Mehr >
<input type="checkbox"/>	3	1 or 2	18067	Erweitert	Anzeigen Mehr >
<input type="checkbox"/>	2	exp Acupuncture/	1259	Erweitert	Anzeigen Mehr >
<input type="checkbox"/>	1	exp Acupuncture Therapy/	17291	Erweitert	Anzeigen Mehr >

| Auswahl verknüpfen mit:

Freitextsuche | Referenz suchen | Suchwerkzeuge | Feldsuche | Expertensuche | Mehrfeldsuche

1 Ressource ausgewählt | [Ausblenden](#) | [Ändern](#)

Ovid MEDLINE(R) 1946 to March Week 4 2015

Stichwort oder Phrase eingeben („" oder „S" bei Trunkierung)

Stichwort Autor Artikeltitlel Zeitschriftenname

Eingrenzen (öffnen) Einschließlich Multimedia Begriff einem Schlagwort zuordnen

Abbildung 3-11 Angewandte Suche MEDLINE via OVID (31.03.2015) über Lizenz der Universitätsbibliothek Regensburg

Zusammenfassung

Mit den großen bibliographischen Datenbanken stehen Werkzeuge zur Verfügung, die den gesamten Suchprozess komfortabel unterstützen. Sie spielen vor allem bei der Suche nach Primärliteratur eine wichtige Rolle.

Retrieval-Funktionalität und Benutzeroberflächen einzelner Datenbankanbieter und Hosts unterscheiden sich in der Praxis jedoch erheblich und führen damit zu teils vollkommen unterschiedlichen Ergebnissen (Katchamart et al. 2011; Boeker et al. 2012). Nutzer, die bei der Recherche nach Fachliteratur auf möglichst vollständige Ergebnisse angewiesen sind und dabei auf verschiedenen Systemen suchen müssen, sind notwendigerweise gezwungen, sich einzelne Funktionen anzueignen und Unterschiede in der Funktionsweise zu kennen.

Eine Sonderrolle spielt die bibliographische Volltextdatenbank Cochrane Library. Auch sie ist ein mächtiges Werkzeug zur Literatursuche. Der Anbieter Wiley ist in sehr engem Kontakt mit den Cochrane Review Gruppen und daher auch bereit, Nutzerschnittstellen in Kooperation mit den Nutzern gezielt auf deren Bedürfnisse anzupassen (Foxlee und Tovey 2013). Dieser Trend zur engen Zusammenarbeit von Anbieter und Kunden wurde bereits in Abschnitt 2.3.3.2 angesprochen. Die Cochrane Library macht offenbar insofern eine Ausnahme, dass hier die Anpassungen in erster Linie für die professionelle Nutzergruppe vorgenommen werden.

3.4.3 Methodologische Anforderungen

Die systematische Literaturrecherche führt insbesondere in bibliographischen Fachdatenbanken zu charakteristisch hochstrukturierten, möglichst sensitiven Suchprozessen, versucht aber trotzdem die größtmögliche Balance zwischen Sensitivität und Relevanz zu finden, um die Menge nicht-relevanter Dokumente gering zu halten (Lefebvre et al. 2011):

“Searches for systematic reviews aim to be as extensive as possible in order to ensure that as many as possible of the necessary and relevant studies are included in the review. It is, however, necessary to strike a balance between striving for comprehensiveness and maintaining relevance when developing a search strategy. Increasing the comprehensiveness (or

sensitivity) of a search will reduce its precision and will retrieve more non-relevant articles.“

Das, was den systemseitigen Kern eines Retrieval-Systems ausmacht, nämlich bei größtmöglichem Recall die Precision ebenso möglichst zu optimieren, wird in der systematischen Literaturrecherche als Charakteristikum der Suchanfrage definiert und letztendlich als Aufgabe dem Nutzer überlassen: die Sensitivität von Suchanfragen hat auch beim Ziel der Balance immer Vorrang. Dies zeigt, wie sehr der Nutzer sowohl bei der Entwicklung von Queries als auch im späteren Screening-Prozess gefordert ist.

3.4.3.1 Entwicklungsprozess

Methodologisch sind die Anforderungen an klassische Retrievalprozesse in Booleschen Datenbanken in den verschiedenen Handbüchern, insbesondere bei Cochrane klar definiert und charakterisiert durch die Anforderung einer möglichst sensitiven Suche.

Die Entwicklung strukturierter Suchprofile mündet in strukturierten, datenbankspezifischen Queries der Form, siehe Abschnitt 3.3.1.1:

$(P_1 \text{ OR } P_2 \text{ OR} \dots P_N) \text{ AND } (I_1 \text{ OR } I_2 \text{ OR } \dots I_N) \text{ AND } (C_1 \text{ OR } C_2 \text{ OR } \dots C_N) \text{ AND } (O_1 \text{ OR } O_2 \text{ OR } \dots O_N).$

Der Entwicklungsprozess umfasst demnach alle Methoden, die Building Blocks der formalisierten Fragestellung durch die Hinzunahme weiterer Suchterme $T_{1,\dots}, T_k, \dots$ aus kontrolliertem Vokabular als auch Freitexttermen auszubauen.

Den Einsatz sowohl an Deskriptoren als auch Textwörtern begründet das Cochrane Handbuch mit der Notwendigkeit der sensitiven Suche im Kontext von Indexierungspraktiken bibliographischer Fachdatenbanken (Lefebvre et al. 2011):

“[...] indexers are not always experts in the subject areas or methodological aspects of the articles that they are indexing. In addition, the available indexing terms might not correspond to the terms the searcher wishes to use.“

Die bereits in Abschnitt 2.3.2.2 angesprochene problematischen Aspekte der Indexierung wie Domänenwissen und Eignung kontrollierten Vokabulars müssen in der systematischen Übersichtsarbeit also vom Nutzer bewusst verinnerlicht sein, Suchanfragen müssen so konzipiert werden, dass an dieser Stelle jede erdenkliche Fehlerquelle überbrückt werden kann.

Die datenbankspezifische Syntax unter Verwendung von Wildcards und PROXIMITY-Operatoren resultiert ebenso wie die eingesetzte Menge an synonymen Suchtermen aus der Vorgabe einer möglichst sensitiven Suche:

Für Freitextterme werden aus methodologischer Sicht sowohl unterschiedliche Schreibweisen, Synonyme und Akronyme berücksichtigt, als auch Trunkierungen und PROXIMITY-Operatoren eingesetzt. Die Suche mit in der Thesaurus Hierarchie innenstehenden Deskriptoren sollten entsprechend je nach Datenbank über Nutzung der *Explode*-Funktion auf darunterliegende Baumstrukturen erweitert werden (siehe Abschnitt 3.4.2.2).

Eine durchdachte Suchstrategie für die Suche in bibliographischen Datenbanken findet ein ausgewogenes Verhältnis zwischen der geforderten Vollständigkeit der aufzufindenden Information sowie einer Präzisierung der Suchanfrage zur Reduktion des Workloads bei der Analyse von Daten. Die Entwicklung dieser sensitiven Suchprofile wird als strukturierter iterativer Prozess beschrieben (Lefebvre et al. 2011):

“Developing a search strategy is an iterative process in which the terms that are used are modified, based on what has already been retrieved.”

Iterativ beginnt die Entwicklung laut Methodenlehre mit einer rudimentären, datenbankspezifischen Query, die alle Building Blocks vollständig abdeckt. Die mit diesem aktuellen Stand des Suchprofils durchgeführten Testsuchen dienen anhand der zurückgelieferten Treffermenge der Beurteilung der Vollständigkeit des Suchprofils, aber auch der Identifikation weiterer Suchterme und letztendlich der Entscheidung darüber, ob

- in relevanten Dokumenten identifizierte, bisher unbekannte Suchterme in das Suchprofil eingearbeitet werden müssen
- das Suchprofil aufgrund des Verlusts bekannter relevanter Dokumente modifiziert werden muss
- das Suchprofil aufgrund der Menge nicht-relevanter Terme modifiziert werden muss

- oder die Entwicklung des Suchprofils abgeschlossen werden kann.

Dies entspricht dem klassischen Citation Pearl Growing-Ansatz, der bereits in Abschnitt 2.4.1.2 beschrieben wurde. Iterativ reduziert sich der Aufwand des Suchprozesses mit der Zeit, da mit jedem Durchlauf weniger neue relevante Treffer hinzukommen. In der Praxis ist dies das Abbruchkriterium. Suchprofile sind vollständig, wenn sich der Aufwand weiterer Iterationen nicht mehr lohnt. Die restriktive Methodenlehre der Cochrane Collaboration weist jedoch darauf hin, dass es durchaus möglich sei, 120 Treffer einer Datenbanksuche pro Stunde und damit ca. 1000 am Tag hinsichtlich oben genannter Entscheidungen zu screenen, ein Hinweis auf exzessives Durchlaufen dieses iterativen Prozesses. (Lefebvre et al. 2011)

Die Entwicklung von Suchprofilen unterscheidet zwischen neuen Reviews und Updates von Reviews. Bei Updates werden bisherige Queries für alle in der Vorgängerversion durchsuchten Datenquellen auf neue Indexterme oder die Verwendung neuer Fachterminologien in der Forschung überprüft. Parallel kann die Suche auf weitere Datenbanken ausgedehnt werden. vgl. hierzu (Lefebvre et al. 2011; Kugley et al. 2015).

Strukturell beinhaltet eine Suchstrategie immer die Kernkonzepte der Review-Fragestellung (siehe Abschnitt 3.3.1). Ausgehend von dieser Struktur liefert eine Query alle die Dokumente zurück, die zu jedem Konzept mindestens einen der in der Query enthaltenen Suchterme beinhaltet. Daher werden Suchterme eines Konzepts entsprechend der Struktur der Suchanfrage mit dem Booleschen Operator OR verknüpft, die Konzepte wiederum mit dem Operator AND.

Die Verwendung des NOT Operators wird wie auch anderswo in der Fachliteratur weitestgehend als kritisch betrachtet, um die Gefahr des Verlusts relevanter Dokumente möglichst gering zu halten. Werden Studien zu Frauen gesucht, würde beispielsweise ‚NOT male‘ alle Studien ausschließen, die Frauen und Männer gleichermaßen betreffen (vgl. (Lefebvre et al. 2011; Kugley et al. 2015).

Zur Vermeidung von Biases, insbesondere des Publikations- und Sprachenbias (siehe Abschnitt 3.1.2) weist das Cochrane Handbuch explizit darauf hin, keine Eingrenzung auf Sprachen vorzunehmen. Auch eine Eingrenzung auf Publikationszeiträume ist nur dann gerechtfertigt, wenn Behandlungen oder andere Aspekte der Fragestellung vor einem gewissen Datum gesichert nicht verfügbar waren. Auch eine Eingrenzung auf spezifische Publikationsformate sollte nicht Teil einer Suchanfrage für systematische Übersichtsarbeiten sein (Lefebvre et al.

2011). Für das Auffinden von Publikationen mit entsprechend der Evidenzpyramide qualitativ hochwertiger Studiendesigns wie der randomisierten kontrollierten Studie schlägt die Methodenlehre die Verwendung so genannter methodologischer Suchfilter vor, ihrem Einsatz kommen in der medizinischen Literaturrecherche besondere Bedeutung zu.

3.4.3.2 Methodologische Suchfilter

Methodologische Suchfilter werden eingesetzt, um spezielle Studiendesigns oder auch Publikationstypen zu identifizieren, beispielsweise die Ergebnisse randomisierter klinischer Studien oder systematische Übersichtsarbeiten. Jenkins definiert methodologische Suchfilter als „*a search strategy combining search terms relating to the research methodology*“ (Jenkins 2004), also als eine spezielle, ausgearbeitete Form des Suchprofils für einzelne Datenbanken und spezifische Fragestellungen. Cochrane spezifiziert den Zweck von Suchfiltern weiter mit „*search strategies that are designed to retrieve specific types of records, such as those of a particular methodological design.*“ (Lefebvre et al. 2011).

Die Idee in der Verwendung eines Suchfilters ist folgende:

Nachdem bereits ein thematisch sauber ausgearbeitetes Suchprofil für einzelne Datenbanken ausgearbeitet wurde, können zusätzliche Suchfilter eingesetzt werden, um spezielle Studientypen leichter zu identifizieren bzw. die Treffermenge weiter zu spezifizieren, ohne dabei eklatante Verluste bisher identifizierter relevanter Dokumente in Kauf nehmen zu müssen. Man kann sich den Suchfilter also wie ein feines Sieb vorstellen, das zusätzlich über die thematisch identifizierte Treffermenge gelegt wird.

Hierfür wird ein Suchprofil der Form

$$(P_1 \text{ OR } P_2 \text{ OR } \dots P_N) \text{ AND } (I_1 \text{ OR } I_2 \text{ OR } \dots I_N) \text{ AND } (C_1 \text{ OR } C_2 \text{ OR } \dots C_N) \text{ AND } (O_1 \text{ OR } O_2 \text{ OR } \dots O_N)$$

mit Hilfe eines Suchfilters

$$(M_1 \text{ OR } M_2 \text{ OR } \dots M_N)$$

weiter spezifiziert:

$(P_1 \text{ OR } P_2 \text{ OR } \dots P_N) \text{ AND } (I_1 \text{ OR } I_2 \text{ OR } \dots I_N) \text{ AND } (C_1 \text{ OR } C_2 \text{ OR } \dots C_N) \text{ AND } (O_1 \text{ OR } O_2 \text{ OR } \dots O_N) \text{ AND } (M_1 \text{ OR } M_2 \text{ OR } \dots M_N)$

Suchfilter sind wie Suchprofile immer datenbankabhängig: Sie enthalten nicht nur datenbank-spezifische Syntax, sondern aufgrund von Thesauri auch spezifisches, kontrolliertes Vokabular.

Die NLM hat als Anbieter der großen Datenbanken MEDLINE und Pubmed in Kooperation mit der McMaster University eine Reihe methodologischer Suchfilter (engl. *hedges*) speziell für MEDLINE Suchen veröffentlicht, die in regelmäßigen Abständen überarbeitet werden, beispielsweise für klinische Fragestellungen, für Fragestellungen aus der Genetik als auch zum Auffinden systematischer Überblicksarbeiten (National Library of Medicine (U.S.). Bibliographic Services Division; McMaster University Health Information Research Unit 2013). Basis für die Validierung und Qualitätsbewertung dieser von Haynes et al. definierten Suchfilter ist ein von der McMaster University per Handsuche initiiertes und gepflegter Datenbestand (Haynes 2004, 2005), der heute als McMaster Plus (Health Information Research Unit 2016) bekannt ist.

Auch Cochrane bietet in seinem Handbuch verschiedene Suchfilter an, die ihren Schwerpunkt entweder auf Recall oder Precision legen und für die beiden großen Datenbanken MEDLINE und EMBASE ausgelegt wurden: *The Cochrane Highly Sensitive Search Strategies for identifying randomized trials in MEDLINE* ist in mehreren Varianten wie auch alternativ für die Datenbank EMBASE verfügbar (Lefebvre et al. 2011, Kap. 6.4.11):

- *Cochrane Highly Sensitive Search Strategy for identifying randomized trials in MEDLINE: sensitivity-maximizing version (2008 revision); PubMed format* (Fokus auf Sensitivität)
- *Cochrane Highly Sensitive Search Strategy for identifying randomized trials in MEDLINE: sensitivity- and precision-maximizing version (2008 revision); PubMed format* (Optimierung auf ausgewogenes Verhältnis von Sensitivität und Spezifität)
- *Cochrane Highly Sensitive Search Strategy for identifying randomized trials in MEDLINE: sensitivity-maximizing version (2008 revision); Ovid format*
- *Cochrane Highly Sensitive Search Strategy for identifying randomized trials in MEDLINE: sensitivity- and precision-maximizing version (2008 revision); Ovid format*

Dass es für den MEDLINE-Filter sowohl eine Pubmed als auch eine OVID-Variante gibt, ist wichtig. Zum einen unterscheidet sich die Syntax der Anbieter (siehe Abschnitt 3.4.2.2), zum anderen aber ist bekannt, dass sich die Anbieter mit Zugriff auf die Datenbank MEDLINE bei gleicher Semantik in den Treffern als auch in ihrer Sensitivität unterscheiden (Boeker et al. 2012; Katchamart et al. 2011). Es ist also nicht unerheblich, ob man MEDLINE über Pubmed durchsucht oder einen anderen Datenbankhost zur Suche heranzieht.

Tabelle 3-11 Suchfilter: Sensitivity and Precision Maximizing Version aus (Lefebvre et al. 2011)

<p>Cochrane Highly Sensitive Search Strategy for identifying randomized trials in MEDLINE: sensitivity- and precision-maximizing version (2008 revision); PubMed format</p>
<pre>#1 randomized controlled trial [pt] #2 controlled clinical trial [pt] #3 randomized [tiab] #4 placebo [tiab] #5 clinical trials as topic [mesh: noexp] #6 randomly [tiab] #7 trial [ti] #8 #1 OR #2 OR #3 OR #4 OR #5 OR #6 OR #7 #9 animals [mh] NOT humans [mh] #10 #8 NOT #9</pre>
<p>Cochrane Highly Sensitive Search Strategy for identifying randomized trials in MEDLINE: sensitivity- and precision-maximizing version (2008 revision); Ovid format</p>
<pre>1 randomized controlled trial.pt. 2 controlled clinical trial.pt. 3 randomized.ab. 4 placebo.ab 5 clinical trials as topic.sh. 6 randomly.ab. 7 trial.ti. 8 1 or 2 or 3 or 4 or 5 or 6 or 7 9 exp animals/ not humans.sh. 10 8 not 9</pre>

Über die Website der ISSG sind thematisch geordnet eine Reihe teils evaluierter Suchfilter abrufbar (InterTASC Information Specialists' Sub-Group (ISSG) 2015), ebenso wie beispielsweise über die Seiten der McMaster University (McMaster University Health Information Research Unit 2013), der laut einer Umfrage und Interviews (Beale et al. 2014) neben Cochrane wichtigsten Quelle für das Auffinden geeigneter Suchfilter.

Die Verwendung von Suchfiltern ist in der Fachwelt inzwischen gängige Praxis und für Cochrane Reviews nach dem MECIR Statement mindestens erwünscht (Chandler et al. 2013). Bei der Auswahl spielen neben der Spezialisierung auf Recall und Precision vor allem die Aktualität von Filtern als auch der Name des Anbieters eine große Rolle, auf der anderen Seite scheint deren Einsatz als auch Nutzen in der Praxis allerdings nach wie vor nicht ausreichend dokumentiert (Beale et al. 2014).

Das Potential in der Nutzung von Filtern zeigt sich in der Praxis von zwei Seiten: Studiendesigns als auch Indexierung einzelner Datenbanken. Thesauri wurden erst in den 1990er Jahren um kontrolliertes Vokabular zur Kennzeichnung randomisierter klinischer Studien ergänzt (Lefebvre et al. 2013). Bis dahin bibliographisch erfasste Artikel zu randomisierten klinischen Studien sind allerdings auch heute nicht entsprechend indexiert und daher über reine Thesaurus-Suchen nicht identifizierbar.

Nicht nur RCTs sind für die systematische Übersichtsarbeit von Interesse: Aus Sicht der evidenzbasierten Medizin wäre es wünschenswert, Filter auch für nicht-randomisierten Studientypen zur Verfügung zu stellen, vgl. hierzu (Bekhuis und Demner-Fushman 2012), Ansätze hierzu liefert beispielsweise die Arbeit von Wilczynski et al. (Wilczynski et al. 2005).

Um nicht nur möglichst sensitive, sondern auch möglichst objektive Suchfilter zu generieren, wurden für deren Entwicklung in den vergangenen Jahren zusätzlich Checklisten ausgearbeitet, welche die Objektivität des Entwicklungsprozesses garantieren und die Anfälligkeit für systematische Fehler reduzieren sollen, beispielsweise die ISSG Search Filter Appraisal Checklist (Glanville et al. 2008), die über die ISSG Website abrufbar ist (The InterTASC Information Specialists' Sub-Group 2016). Ergänzend dazu bietet Abschnitt 3.4.4.1 einen Überblick über die Evaluierung von Suchfiltern.

3.4.3.3 Dokumentation

Wesentlicher Teil des Qualitätsmanagements ist die Dokumentation von Suchprofilen, die zur Reproduzierbarkeit strukturierter Suchprozesse beiträgt.

Schon zur Fehlervermeidung wird empfohlen, die Syntax über *Copy and Paste* direkt so zu übernehmen, wie sie in den Datenbanken durchgeführt wurden, siehe beispielhaft Tabelle 3-12. Die Daten wurden dabei einer Übersichtsarbeit entnommen, die aus dem zum Zeitpunkt der Verfassung dieses Abschnitts neuesten Ausgabe der CDSR unter der Vorgabe möglichst

gut dokumentierter und möglichst kurzer Suchprofile ausgewählt worden war (Schlager et al. 2016).

Tabelle 3-12 Gegenüberstellung Queries aus (Schlager et al. 2016)

Cochrane Central Suchanfrage
<pre>#1 MeSH descriptor Psoriasis explode all trees #2 (psoria*) #3 (#1 OR #2) #4 MeSH descriptor Scalp explode all trees #5 (scalp*) #6 (#4 OR #5) #7 (#3 AND #6)</pre>
MEDLINE OVID Suchanfrage
<pre>1. exp Psoriasis/ or psoria\$.mp. 2. exp Scalp/ 3. scalp\$.mp. 4. 2 or 3 5. randomized controlled trial.pt. 6. controlled clinical trial.pt. 7. randomized.ab. 8. placebo.ab. 9. clinical trials as topic.sh. 10. randomly.ab. 11. trial.ti. 12. 5 or 6 or 7 or 8 or 9 or 10 or 11 13. exp animals/ not humans.sh. 14. 12 not 13 15. 1 and 4 and 14 [Lines 5-14: Cochrane Highly Sensitive Search Strategy for identifying randomized trials in MEDLINE: sensitivity- and precision-maximizing version (2008 revision)]</pre>
EMBASE OVID Suchanfrage
<pre>1. exp psoriasis/ 2. psoria\$.ti,ab. 3. 1 or 2 4. exp scalp/ 5. scalp\$.ti,ab. 6. 4 or 5 7. crossover procedure.sh. 8. double-blind procedure.sh. 9. single-blind procedure.sh. 10. (crossover\$ or cross over\$).tw.</pre>

```
11. placebo$.tw.  
12. (doubl$ adj blind$).tw.  
13. allocat$.tw.  
14. trial.ti.  
15. randomized controlled trial.sh.  
16. random$.tw.  
17. or/7-16  
18. exp animal/ or exp invertebrate/ or animal experiment/ or  
animal model/ or animal tissue/ or animal cell/ or nonhuman/  
19. human/ or normal human/  
20. 18 and 19  
21. 18 not 20  
22. 17 not 21  
23. 3 and 6 and 22
```

Cochrane empfiehlt die Entwicklung von Queries in separaten Textdokumenten. Hierbei werden einzelne Suchterme zeilenweise erfasst, dann zeilenweise als einzelne Queries in datenbankspezifischer Syntax im Datenbank Interface eingegeben und im nächsten Schritt über das Datenbank Interface mittels Boolescher Operatoren verknüpft (vgl. (Lefebvre et al. 2011)), siehe beispielhaft Abbildung 3-12.

#	Suchen	Ergebnisse	Suchtyp	Aktionen
1	exp Psoriasis/ or psoria\$.mp.	38521	Erweitert	Anzeigen Mehr >>
2	exp Scalp/	11354	Erweitert	Anzeigen Mehr >>
3	scalp\$.mp.	31453	Erweitert	Anzeigen Mehr >>
4	2 or 3	31453	Erweitert	Anzeigen Mehr >>
5	randomized controlled trial.pt.	406624	Erweitert	Anzeigen Mehr >>
6	controlled clinical trial.pt.	90068	Erweitert	Anzeigen Mehr >>
7	randomized.ab.	303867	Erweitert	Anzeigen Mehr >>
8	placebo.ab.	155146	Erweitert	Anzeigen Mehr >>
9	clinical trials as topic.sh.	174956	Erweitert	Anzeigen Mehr >>
10	randomly.ab.	215136	Erweitert	Anzeigen Mehr >>
11	trial.ti.	131937	Erweitert	Anzeigen Mehr >>
12	5 or 6 or 7 or 8 or 9 or 10 or 11	928730	Erweitert	Anzeigen Mehr >>
13	exp animals/ not humans.sh.	4189142	Erweitert	Anzeigen Mehr >>
14	12 not 13	850021	Erweitert	Anzeigen Mehr >>
15	1 and 4 and 14	116	Erweitert	Anzeigen Mehr >>

Abbildung 3-12 Exemplarische Durchführung der OVID Suche nach (Schlager et al. 2016), 29.02.2016

3.4.3.4 Zusammenfassung

Die Entwicklung von Suchstrategien – als Sammlung datenbankspezifischer Suchprofile – fokussiert sich auf Suchprozesse in bibliographischen Datenbanken mit Booleschem Retrieval und umfasst in der Praxis aufgrund der Forderung nach Vollständigkeit und Fehlerrobustheit Suchprofile für eine Vielzahl verfügbarer Primärliteraturre Ressourcen.

Suchprofile sind charakterisiert als möglichst sensitive, datenbankspezifische Formalisierung der dem Review zugrundeliegender klinischer Fragestellung, bestehend aus kontrolliertem Vokabular und Freitexttermen.

Als Zusammenstellung dieser hochstrukturierten Suchprofile garantieren Suchstrategien die Anforderung an die Reproduzierbarkeit systematischer Übersichtsarbeiten.

Methodologisch unterliegt die Entwicklung von Suchprofilen strengen Regeln, die insbesondere der regelmäßigen Bewertung der Vollständigkeit geschuldet sind und daher in einem iterativen Prozess münden, bei dem in regelmäßigen Abständen sowohl das Retrieval bekannter Schlüsselreferenzen als auch Sensitivität und Relevanz des Suchprofils überprüft werden müssen.

Die Entwicklung von Suchstrategien mit Suchprofilen in Form einer vollständigen Repräsentation der klinischen Fragestellung ist dann abgeschlossen, wenn Testsuchen keine weiteren relevanten Treffer liefern und durch neuerliche Änderungen keine relevanten Dokumente verloren gehen.

3.4.4 Evaluation strukturierter Suchen

Die objektive qualitative als auch quantitative Bewertung von Suchprofilen im Hinblick auf ihre Fehlerrobustheit macht einen wesentlichen Beitrag in der Methodenlehre der Durchführung systematischer Übersichtsarbeiten aus, wie sich beispielsweise am Einsatz methodologischer Suchfilter zeigt. Warum dies so ist, lässt sich mit den Grundzügen der EBM erklären: Optimierte Suchanfragen sollten möglichst objektiv ermittelt werden, um systematische Verzerrungen zu vermeiden (Jenkins 2004), sie sollten außerdem später möglichst leistungsstark sein, um die Sensitivität und Spezifität der späteren Treffer zu maximieren (Lefebvre et al. 2011).

Dabei führen aus methodologischer Sicht jede Ergänzung und Modifikation am Suchprofil sowie die Beurteilung von Testsuchen notwendigerweise zur Entscheidung, ob die Entwicklung eines Suchprofils abgeschlossen kann oder ob sie aufgrund neuer relevanter Dokumente weitergeführt werden muss (siehe Abschnitt 3.4.3.1).

Aus dem Objektivitätsanspruch heraus sind in der Praxis die „klassischen Retrievalmaße“ wie Precision und Recall, gemessen an „idealen Treffermengen“ die Richtschnur für die Optimierung dokumentierter und dokumentierbarer Suchvorgänge während einer systematischen Literaturrecherche.

Anmerkung: In Abschnitt 2.4.3 wurde festgelegt, dass für die vorliegende Arbeit die Termini Precision und Recall ausschließlich im Kontext systemseitiger Evaluierung verwendet werden

würden, Sensitivität und Spezifität hingegen für nutzerinitiierte Maßnahmen. In der Fachliteratur werden die beiden Begriffe Precision und Recall hingegen teils anders definiert als in der vorliegenden Arbeit und auch informationswissenschaftlich verstanden (Lefebvre et al. 2011):

“Sensitivity is defined as the number of relevant reports identified divided by the total number of relevant reports in existence. Precision is defined as the number of relevant reports identified divided by the total number of reports identified.”

nämlich als Synonym für die klassischen Retrievalmaße Precision und Recall.

Da es in der Praxis nicht möglich ist, den Recall zu messen, erfolgt die Evaluierung von Queries ähnlich der eines Information Retrieval-Systems (siehe Abschnitt 2.3.2) und misst die Performanz einer Suchstrategie anhand einer bekannten Referenzmenge. Die Ermittlung von Referenzmengen und Methoden zur Berechnung des Recalls bestimmt dabei die Güte und Validität von Evaluierungsverfahren (Booth 2010). Die Methodenlehre greift hier als auf klassische, systemseitige Evaluierungsmethoden zurück, siehe hierzu 2.3.2.5.

Zur der Evaluierung von Suchstrategien mittels eines „Gold-Standard“ wird der Recall einer vorliegenden Suchstrategie (im Sinne einer Suchanfrage) gegen die Mächtigkeit einer bekannten, objektiv generierten Referenzmenge getestet, der Quotient aus diesen beiden Größen bestimmt die Güte der Suchstrategie. Er findet sich klassischerweise bei der Entwicklung von Suchfiltern.

3.4.4.1 Referenzmenge Gold-Standard

Die Anfänge methodologischer Suchfilter, bei Jenkins als „erste Generation an Suchfiltern“ bezeichnet, lagen in Erfahrungswerten von Informationsexperten, die ersten Suchfilter sind also das Ergebnis subjektiver Expertenmeinung und basierten häufig auf Wissen um Datenbankstrukturen, zugrundeliegende Thesauri und fachlichen Inhalten. So wurden Suchfilter zu Beginn selten gegen eine bekannte Referenzmenge getestet (Jenkins 2004). In der weiteren Entwicklung, der „zweiten Generation Suchfilter“, begann man, die Ergebnisse eines Filters gegenüber einem so genannten „Gold-Standard“ zu evaluieren, wie oben beschrieben (Jenkins 2004).

Mit der neuesten Generation methodologischer Suchfilter wurde der Objektivitätsanspruch höher: Nach wie vor werden die Filter gegen einen „Gold-Standard“ getestet, gleichzeitig wird aber versucht, bereits bei der Entwicklung des Suchfilters objektivere Verfahren einzusetzen (Jenkins 2004). Die in einem Suchfilter verwendeten Terme werden hierfür beispielsweise durch Häufigkeitsanalysen und statistische Verfahren wie logistischer Regression gewonnen (Jenkins 2004), eine große Initiative wurde bereits an anderer Stelle genannt: McMaster PLUS (Health Information Research Unit 2016). Die Bemühungen um die Entwicklung von Gold-Standards für die Evaluation von Suchfiltern ist demnach vergleichbar mit dem TREC Setting, das in Abschnitt 2.3.2.5 beschrieben wurde.

Den Terminus „Gold-Standard“ liest man in der medizinischen Fachliteratur seit den 1950er Jahren häufig. Da er immer wieder in verschiedenen Bedeutungen verwendet wird, wird über die Verwendung des Terms in der Literatur allerdings seit Jahren diskutiert. Wiederholt hat die unreflektierte Verwendung gerade in der Medizin bereits zu Kritik geführt, sodass seit langem gefordert wird, auf die Verwendung des Terminus „Gold-Standard“ in jeglichem Kontext zu verzichten (Claassen 2005; Duggan 1992). Die eigentliche Bedeutung geht auf einen Artikel aus dem Jahr 1979 zurück und definiert den „Gold-Standard“ als neues diagnostisches oder therapeutisches Verfahren, das im Vergleich mit anderen Methoden als *das* aktuell gültige Standardverfahren gehandelt wird (Claassen 2005). Dies muss nicht immer zwingend das beste Verfahren sein, es entspricht jedoch einer Benchmark, die unter dokumentierten Bedingungen ermittelt und gesetzt wurde. So gesehen ist der Begriff Gold-Standard prinzipiell auch anwendbar, wenn direkte Maße fehlen, anhand derer die Performanz des Standards objektiv evaluiert werden könnte (Versi 1992).

Aus Gründen der Differenzierung wird diese Referenzmenge im Folgenden als Gold-Standard bezeichnet, sofern sie automatisiert und methodologisch so objektiv wie möglich oder nach entsprechenden Qualitätsstandards ermittelt wurde. Wurde eine solche Referenzmenge intellektuell entwickelt oder durch Handsuchen aufgebaut, wird sie Ground Truth benannt (im Oxford English Dictionary findet sich *groundtruth* als „*A fundamental truth. Also: the real or underlying facts; information that has been checked or facts that have been collected at source.*“ (Oxford University Press 2016b)).

3.4.4.2 Referenzmenge Ground Truth

Suchprofile können ähnlich wie in Abschnitt 3.4.4.1 für den Gold-Standard beschrieben gegenüber dieser Ground Truth getestet werden.

Die Referenzmenge Ground Truth für eine systematische Literaturrecherche beinhaltet nach Cochrane Handbuch idealerweise mindestens bekannte, bereits durchgeführte systematische Übersichtsarbeiten zu einer gegebenen klinischen Fragestellung (Lefebvre et al. 2011). Diese Übersichtsarbeiten gelten in der Fachliteratur als weitestgehend fehlerfreie Evidenzquelle (siehe Abschnitt 3.1.1) und stellen damit die optimale Referenzmenge dar. Für die Entwicklung von Suchprofilen sind thematisch passende systematische Übersichtsarbeiten sowohl als Quelle für potenziell relevante Studien interessant, als auch als Quelle für Suchterme aus bibliographischen Angaben. Mehr noch enthalten sie aufgrund der Dokumentationspflicht der gesamten Suchstrategien häufig auch komplette Suchprofile. Hauptquelle für systematische Übersichtsarbeiten in der Medizin sind die Cochrane Library sowie DARE, die Suche in diesen Datenbanken ist daher für medizinische Übersichtsarbeiten verpflichtend (Chandler et al. 2013).

Nicht immer allerdings sind systematische Überblicksarbeiten für einzelne Fragestellungen überhaupt verfügbar. In diesem Fall muss die Ground Truth durch eine reine Handsuche in Fachdatenbanken und anderen Quellen erfolgen. Dies unterscheidet sie sowohl im Umfang als auch ihrer Generierung vom Gold-Standard. Ground Truth-Mengen sind daher in der Praxis als willkürlicher und weniger fehlerrobust anzusehen, vgl. (Hausner et al. 2012).

Im iterativen Prozess der Entwicklung von Suchprofilen wird mit Ergänzungen oder Modifikationen des Suchprofils typischerweise auch die Ground Truth vergrößert. Bei der Bewertung des optimalen Abbruchzeitpunkts muss also das vollständige Retrieval der Ground Truth gesichert sein. Der Entwicklungsprozess des Suchprofils wird jedoch idealerweise erst dann abgebrochen, wenn weitere Änderungen keine oder nur wenige zusätzliche, relevante Dokumente liefern. Dieses hart formulierte Abbruchkriterium ist in der Praxis jedoch abhängig von verfügbaren Ressourcen (Lefebvre et al. 2011; Kugley et al. 2015).

3.4.5 Identifikation von Suchtermen

Suchprofile wachsen durch die Hinzunahme von Suchtermen für einzelne Building Blocks, sie werden im Suchprofil im einzelnen Building Block über den OR Operator verknüpft.

Wie jedoch ermitteln Nutzer Suchterme? Welche Aussage macht die Methodenlehre zu dieser Frage?

3.4.5.1 Intellektuelle Verfahren

Werden Reviews aktualisiert, so liegt aus der Vorgängerversion bereits eine voll ausgearbeitete Suchstrategie vor. Die Hauptaufgabe besteht dann in der kritischen Überprüfung und gegebenenfalls Aktualisierung vorbestehender Suchprofile. Praktisch bedeutet dies eine Wiederholung aller strukturierten Suchprozesse unter Kontrolle aktualisierter Indexierungsterme wie auch Änderungen in der Datenbanksyntax (Lefebvre et al. 2011).

Bei der Entwicklung neuer Reviews schlägt die Methodenlehre eine Reihe an Suchtaktiken zur Identifikation von Suchtermen vor, die im Folgenden beschrieben werden.

Die mit der klinischen Fragestellung assoziierten Deskriptoren können beispielsweise über die Suche nach bibliographischen Referenzen bekannter, relevanter Schlüsseldokumente identifiziert werden, vgl. (Lefebvre et al. 2011). Dieses auch in der Fachliteratur als Citation Pearl Growing bekannte Prinzip (Booth 2008) findet sich bereits in den von ERIC vorgestellten Suchtaktiken auf Online-Datenbanken (siehe Abschnitt 2.4.1.1).

Weitere Möglichkeiten sind die Nutzung datenbankspezifischer Grundfunktionalitäten wie beispielsweise „*similar articles*“ (PubMed) beziehungsweise „*find similar*“ (OVID) (Lefebvre et al. 2011). Zusätzliches kontrolliertes Vokabular sollte direkt über die Thesauri einzelner Fachdatenbanken identifiziert werden (Kugley et al. 2015). Hierfür stehen den Nutzern mit dem MeSH Browser (vgl. Abschnitt 3.4.2.1) oder bei EMBASE „Begriff einem Schlagwort zuordnen“ (vgl. Abschnitt 3.4.2.3) entsprechende Datenbank Funktionen zur Verfügung.

Die Ground Truth in ihrer Gesamtheit erfüllt bei der Erstellung und Bewertung von Suchstrategien zwei Funktionen: Sie ist einerseits ein Pool relevanter Dokumente zur Bewertung der

Qualität einer Suchstrategie (Booth 2010), zum anderen ist sie aber auch ein Pool für Freitextterme, Synonyme und (über die bibliographischen Einträge) kontrolliertes Vokabular sowie im Fall der systematischen Übersichtsarbeit außerdem für fertige Suchprofile (Lefebvre et al. 2011).

Die Literatur gibt prinzipiell weitere Empfehlungen, die Methodenlehre jedoch nicht. Nicht empirisch belegt ist beispielsweise die Nutzung klassischer Suchtaktiken zur Identifikation von Suchtermen auf Basis der Ground Truth (Booth 2008), beispielsweise bibliographische Suchen (Rückwärtssuchen), Zitationssuchen (Vorwärtssuche), Journal Runs und andere Suchtaktiken, die bei Bates unter dem Sammelbegriff Berry Picking subsummiert werden (vgl. Abschnitt 2.1.2). Genutzt werden diese Taktiken in der Praxis jedoch durchaus, insbesondere in der qualitativen Forschung (Grant 2004).

Zusammenfassend finden sich für die Identifikation von Suchtermen (darunter Deskriptoren ebenso wie Freitextterme und Phrasen) die klassischen Elemente der Suche in Online-Datenbanken wieder. Weitere Vorgaben macht die Methodenlehre nicht. Für den Nutzer bleibt dabei vor allem die Frage nach der Objektivität in der Gewinnung potenziell relevanter Suchterme als auch der Vollständigkeit von Suchprofilen im Hinblick auf die Menge aufgefundener Synonyme offen. Im folgenden Abschnitt werden beispielhaft einige Text Mining-gestützte Verfahren vorgestellt, die in der Literatur zur Unterstützung dieses Prozesses beschrieben sind.

3.4.5.2 Text Mining-gestützte Verfahren

Der rasante Zuwachs an medizinischer Primärliteratur, die Verzerrungen in der medizinischen Forschung und Literatur, die dadurch ebenso zunehmenden zeitlichen und personellen Ressourcen führten zu dem Wunsch und der Forderung nach objektiveren Methoden in der Entwicklung von Suchprofilen und einer verbesserten Reproduzierbarkeit (Lefebvre et al. 2013). So haben sich in den vergangenen Jahren Methoden etabliert, einzelne Teilaufgaben der systematischen Übersichtsarbeit unter anderem mittels Text Mining-Verfahren zu automatisieren.

Text Mining ist eine Technologie, die in den vergangenen Jahren zunehmend an Bedeutung gewonnen hat. Sie dient der Unterstützung der Informationsgewinnung in großen textuellen Datenbeständen, beispielsweise Datenbanken, aber auch Suchmaschinen. Im Gegensatz zu

Information Retrieval-Technologien ist das Ziel des Text Mining nicht die Informationsgewinnung, sondern die Erkennung von Mustern in Dokumentenbeständen. Text Mining umfasst Methoden, die es Nutzern erleichtern sollen, anhand von in Datenbeständen identifizierten Mustern die für sie relevanten Informationen aufzufinden, beispielsweise durch Dokumentenklassifikation, Clustering-Verfahren oder die Extraktion von Information aus Text, siehe beispielsweise (Aggarwal und Zhai 2012) und kann daher als unterstützende Maßnahme zum Information Retrieval verstanden werden.

Dieser Ansatz ist auch in der Medizin nicht neu: Vor allem im Bereich der biomedizinischen Forschung werden seit Jahren Text Mining-Verfahren eingesetzt, um Rechercheprozesse zu unterstützen, wie beispielsweise die aktuelle Übersicht von (Fleuren und Alkema 2015) zeigt. Diese fokussieren sich jedoch häufig auf die Identifikation von Proteinen oder Genen, sie sind daher für die systematische Übersichtsarbeit nur teilweise interessant.

Auf Nutzerseite werden Text Mining-gestützte Verfahren während der Entwicklung systematischer Übersichtsarbeiten prinzipiell positiv bewertet vgl. (Paynter et al. 2016). In der Entwicklung von Suchprofilen werden vor allem die Beschleunigung des Gesamtvorgangs, die Qualität und die Reproduzierbarkeit von Ergebnissen hervorgehoben (Balan et al. 2014). Ananiadou et al. (Ananiadou et al. 2009) betten verfügbare Technologien in den Ablauf der systematischen Übersichtsarbeit ein:

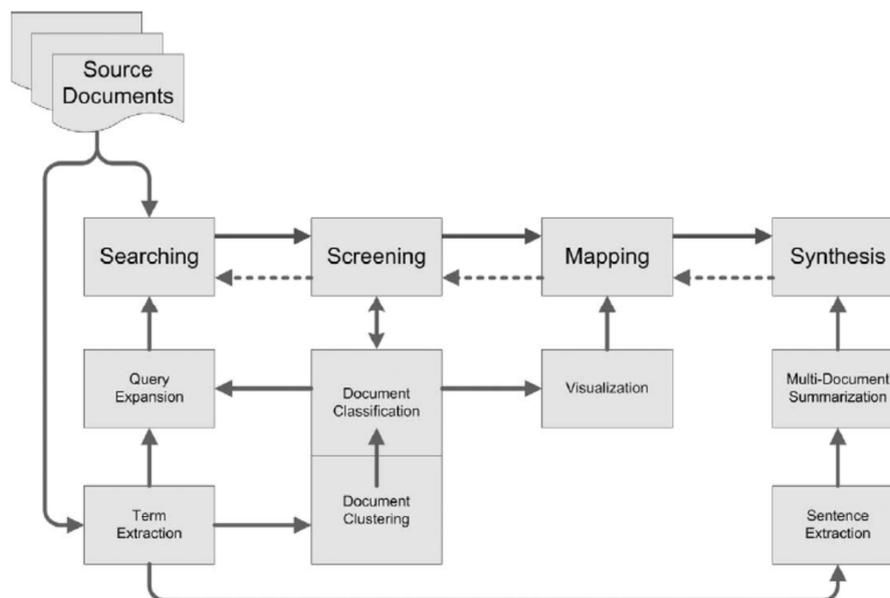


Abbildung 3-13 Text Mining-Verfahren im Workflow systematischer Übersichtsarbeiten aus (Ananiadou et al. 2009, S. 512)

Werkzeuge

Insbesondere beim EPPI Centre, aber auch beim CRD werden seit Jahren Text Mining-Verfahren in der Entwicklung systematischer Übersichtsarbeiten eingesetzt und hinsichtlich ihrer Eignung evaluiert.

Die in der bibliographischen Fachliteratur in den vergangenen Jahren vorgestellten und teils etablierten Text Mining-gestützten Lösungsansätze zur Ermittlung von Suchtermen und dem Aufbau erschöpfender Suchprofile basieren dabei meist auf dem so genannten Bag-of-Words Prinzip, u.a. bei (Frantzi et al. 2000; Ananiadou et al. 2009; Hausner et al. 2012; Slater 2014; Hausner et al. 2015). Unter dem Bag-of-Words Prinzip versteht man eine vereinfachte Repräsentation von Texten als Menge von Wörtern oder Phrasen, ohne Berücksichtigung von deren Grammatik oder deren semantischer Bedeutung. Stopwörter, also Begriffe, die in einer Sprache sehr häufig vorkommen und keine tragende Bedeutung haben, werden hierzu vorab aus dem Text gestrichen, sodass für die weitere Verarbeitung eine sinntragende Version von Text zur Verfügung steht. (Aggarwal und Zhai 2012)

Beispielhaft sollen zwei Werkzeuge vorgestellt werden, die nach diesem Prinzip funktionieren:

TerMine ist ein Online verfügbares Tool zur Ermittlung von Multiwort-Termen aus Texten, es kann neben Plain Text auch Textdokumente, PDF und HTML Dateien verarbeiten. Die Motivation für die Entwicklung von TerMine liegt in der biomedizinischen Forschung begründet, in der die Zahl an Fachtermini und Neologismen ständig steigt. Neben der Extraktion von Termen erkennt TerMine auch Akronyme. Algorithmisch wurde TerMine auf die Verarbeitung großer Datenmengen ausgelegt und kann so eine Menge von über einer Million MEDLINE Abstracts mit geschätzten 2 GB Datenvolumen in ca. 10 Minuten sprachlich verarbeiten und dabei bis zu knapp 10 Millionen Terme extrahieren. (Okazaki 2015)

PubReMiner gehört mit zu den populärsten Werkzeugen der Literatursuche in medizinischen bibliographischen Datenbanken. Das Tool versteht sich als Frontend zur Datenbank Pubmed, über das Suchanfragen anhand von Text-Mining Analysen sukzessive verfeinert und vom Nutzer analysiert werden können. (Koster 2016)

The screenshot shows the PubMed PubReMiner interface. At the top, it displays the search query "Psoriasis[mh] AND Scalp[mh]" and indicates that the query resulted in 105 references. Below the query, there are buttons for "Goto PubMed with query" and "Create CV output". A section for "Manual adjustment" includes a dropdown for "Abstract limit: 1000" and a "Search with Manual Adjustment" button. Below this, there is a table of search results with columns for OR Year, OR, Author, Count, OR, Word, Mesh, and OR, publication type. The table lists various authors and their associated counts for different terms and MeSH categories.

# OR Year	# OR	Author	# Count	OR Word	Mesh	# OR	publication type
2	2015	GANSLANDT C	105	471	PSORIASIS	105	/ pathology
7	2014	KUMKS	105	397	SCALP*	102	Epidemiology
10	2013	PARK HK	100	109	HUMAN*	100	Scalp
8	2012	SHRUMS	58	165	PATHOLOGY	100	Humans
8	2011	VAN DE KERKHOFF JC	55	210	PATIENT*	82	/ therapeutic use
4	2010	SHILU J	55	136	THERAPY*	51	/ drug therapy
4	2009	ARMSTRONG AW	51	86	DISEASE*	48	/ administration & dosage
4	2008	BRUNFALCO Q	49	65	ADULT*	46	Male
5	2007	CANEVY SW	46	48	MALE*	45	Female
1	2006	COLONLE	46	168	TREAT*	43	Adult
1	2005	ENEVOLD A	45	114	DRUG*	39	Scalp/pathology
2	2004	FELDMAN SR	45	47	FEMALE*	39	/ adverse effects
2	2003	GOTTSCHALK RW	39	78	AGE*	33	Middle Aged
2	2002	HAW CR	37	89	THERAPEUTIC	32	/ diagnosis
2	2001	KRAGALLE K	36	91	SKIN	28	Adolescent
1	2000	LUGER TA	35	91	TOPIC*	28	/ metabolism
3	1998	LYNDE C	34	49	CLINIC*	27	Psoriasis/drug therapy
2	1997	MCCORMACK FL	33	33	MIDDLE	8	- Psoriasis/drug therapy/pathology
2	1995	PAFF K	32	81	AGENT*	1	- Psoriasis/drug therapy/systemology/therapy
1	1994	FOULIN Y	32	80	EFFECT*	1	- Psoriasis/drug therapy/genetics/immunology
3	1993	RUNNE U	32	72	PSORIATIC*	1	- Psoriasis/drug therapy/hygiene/pathology/therapy
1	1992	SAMARASEKERA EJ	30	54	INFLAMMATOR*	1	- Psoriasis/drug therapy/economics
						26	/ complications

Abbildung 3-14 PubReMiner, Screenshot vom 09.03.2016

Die Eingabe von Suchanfragen erfolgt in PubReMiner analog zu Pubmed und mündet in einer Ausgabe in Form von Häufigkeitstabellen mit den Elementen Publikationsjahr, Journal, Autor, Freitextterme, MeSH Term und weitere. Nicht in der Query enthaltene Suchterme oder Entitäten können interaktiv in die vorbestehende Suchanfrage übernommen und neu analysiert oder direkt für die Suche auf Pubmed übernommen werden. (Slater 2014)

Beispiel für den Einsatz von Text Mining-Technologien in der Praxis

Einen verpflichtenden, objektivierten Ansatz zur Entwicklung von Suchanfragen praktiziert das IQWiG und greift damit den Entwicklungsprozess von Suchfiltern auf – hier Unternehmensstandard (Hausner et al. 2012). Es ist in der Praxis das einzige Beispiel für empirisch entwickelte und evaluierte, standardisierte Text Mining-Verfahren.

Der Entwicklungsprozess stellt sich graphisch folgendermaßen dar:

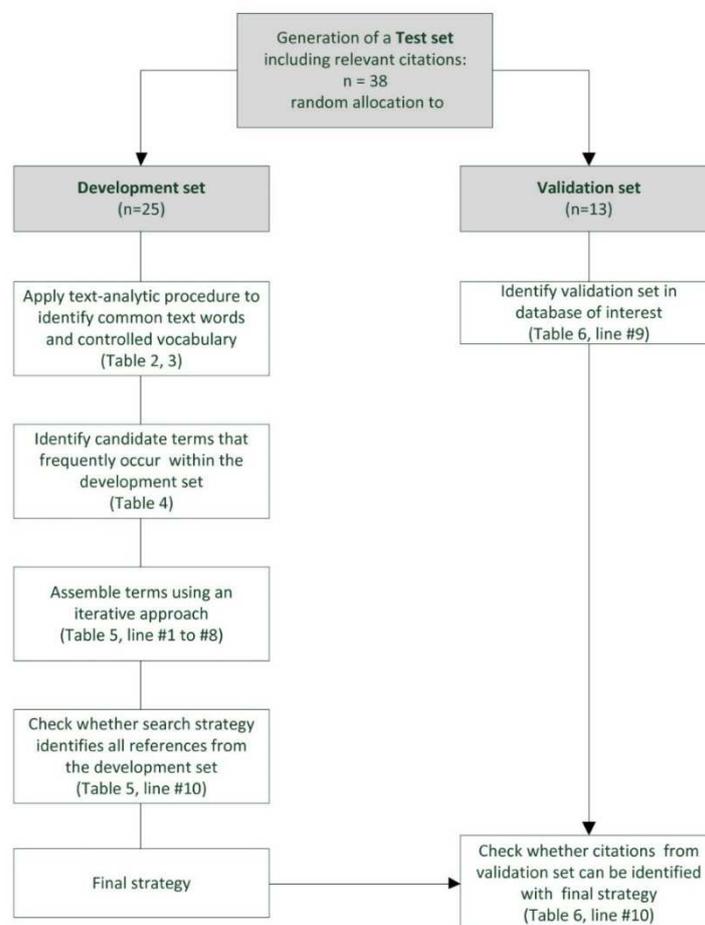


Abbildung 3-15 Objektivierter Ansatz zur Entwicklung von Suchprofilen aus (Hausner et al. 2012, S. 6)

Bei diesem Verfahren werden zur Identifikation von Suchtermen und zur Entwicklung von Suchprofilen alle im Rahmen einer Vorrecherche identifizierten bibliographischen Referenzen herangezogen und randomisiert in ein Entwicklungs- und ein Testset aufgeteilt.

Mit Hilfe von Text Mining-Verfahren und statistischer Software wird die Term-Häufigkeit potenzieller Deskriptoren und Freitextterme der Referenzen aus der Entwicklungsmenge ermittelt. Terme, die in mindestens 20% der Dokumente vorkommen, werden als potenzielle Suchterme extrahiert.

Eine zusätzlich zufallsgeneriert ausgewählte Menge bibliographischer Referenzen aus der Ziel-datenbank dient der Kontrolle bisher ermittelter Freitextterme. In das spätere Suchprofil werden ausschließlich Terme aufgenommen, die in der Datenbank eine geringe Sensitivität, für die Fragestellung jedoch eine hohe Sensitivität aufweisen. Ein ähnliches Verfahren wird auf identifiziertes kontrolliertes Vokabular angewandt. Ausgehend davon wird das Suchprofil in einem intellektuellen Prozess zunächst durch Prüfung des vollständigen Retrievals der Entwicklungsmenge adjustiert. Das ausgearbeitete Suchprofil wird gegen das Testset, im Sinne eines Gold-Standards, getestet. (Hausner et al. 2012)

3.4.6 Zusammenfassung

Strukturierte Suchprozesse sind das Herzstück systematischer Übersichtsarbeiten, sie zeichnen sich insbesondere durch ihre Reproduzierbarkeit aus, die durch die Dokumentation der Suchprofile im Anhang systematischer Übersichtsarbeiten gewährleistet ist.

Die Entwicklung von Suchprofilen erfordert in der Praxis sowohl die Kenntnis als auch die Berücksichtigung datenbankspezifischer Eigenheiten wie Syntaxspezifika, das Wissen um Funktionalitäten einzelner Systeme sowie unterstützende Verfahren zur Kontrolle strukturierter Suchprozesse.

Methodologisch unterliegen Suchprofile vor allem der Forderung nach möglichst vollständigen Retrieval-Ergebnissen. Zur Qualitätsüberprüfung können sie gegen eine im Idealfall objektiv automatisiert oder intellektuell generierte Testmengen hinsichtlich Recall und Precision bewertet werden.

Die Identifikation von Suchtermen aus kontrolliertem Vokabular und Freitexttermen mit dem parallel verlaufenden Ausbau einer Ground Truth als Menge bekannter, relevanter Dokumente wird als iterativer Prozess beschrieben, der lediglich für Deskriptoren als datenbankspezifische Indexierungskomponenten reproduzierbar ist, für Freitextterme jedoch weniger empirische und reproduzierbare Methoden erfordert.

Mit der Automatisierung und Objektivierung in der Entwicklung von Suchprofilen werden Methoden vorgestellt, mit deren Hilfe auch der Prozess der Identifikation von Suchtermen transparenter und reproduzierbarer werden soll.

3.5 Formale und thematische Suchprozesse

Der Nutzung bibliographischer Fachdatenbanken mit reproduzierbaren Suchanfragen als klassisches Retrieval stehen in der Methodenlehre weitere Suchprozesse gegenüber, die sich in ihrer Vorgehensweise deutlich von strukturierten Prozessen unterscheiden. Obwohl sie je nach Methodenlehre verpflichtend vorgeschrieben sind, wurden sie dort bisher wenig spezifiziert. Aus diesem Grund wurden Nutzen und Evidenz weniger strukturierter Suchmethoden in den vergangenen Jahren immer wieder auch zum kritischen Forschungsgegenstand. Das folgende Kapitel gehört thematisch nicht zum Schwerpunkt der Arbeit, unterstützt aber das Verständnis systematischer Übersichtsarbeiten und vervollständigt entsprechend das Gesamtbild.

3.5.1 Quellenspezifika

3.5.1.1 Recherche in klinischen Studienregistern

Wie bei Fachdatenbanken ist die zugrundeliegende technische Struktur von Studienregistern nicht einheitlich, die Abfragemöglichkeiten sind im Allgemeinen im Vergleich zu denen großer Datenbankanbieter einfacher gehalten, auch ihre Funktionalität ist im Vergleich mit kommerziellen Datenbanken typischerweise geringer (Gough et al. 2012). Insgesamt scheint der Aufbau klinischer Studienregister aus technischer und konzeptioneller Sicht in der wissenschaftlichen Literatur vergleichsweise wenig dokumentiert. Die beiden derzeit größten klinischen Studienregister clinicaltrials.gov das vom U.S. National Institutes of Health (NIH) herausgegeben wird, sowie ICTRP, die International Clinical Trials Registry Plattform der WHO, unterscheiden sich technisch, sie verwenden sowohl eine andere Suchtechnologie als auch unterschiedliche Suchoptionen (Glanville et al. 2014).

Allgemein gliedern sich die Informationen zu einzelnen Studien in administrative Angaben, in wissenschaftliche Aspekte sowie Verlaufsdaten. Bei vielen Datenbankfeldern handelt sich um

Freitextfelder, auch die Übermittlung zusätzlicher Dokumente ist möglich. Die dezidierte Angabe von Informationen zum Studienzweck, zu Studienteilnehmern und anderen Merkmalen ermöglicht es, im Rahmen systematischer Literaturrecherchen gezielt nach den Konzepten der klinischen Fragestellung zu suchen, die im Vorfeld der Literaturrecherche per definitionem detailliert festgelegt werden müssen, vgl. u.a. (Higgins und Green 2011; Gough et al. 2012; Booth et al. 2012). Darüber hinaus bietet die Volltextsuche in eventuell verfügbaren, zusätzlichen Publikationen einen direkten Zugriff auf die wissenschaftlichen Ergebnisse.

Wie bibliographische Datenbanken basieren die großen klinischen Studienregister ClinicalTrials.gov und das ICTRP der WHO technisch auf Booleschem Retrieval und bieten sowohl einfache als auch eine angewandte Bedienmodi an. Beide Register implementieren zumindest für bestimmte Felder einen Thesaurus, clinicaltrials.gov die Medical Subject Headings, das ICTRP das UMLS (Unified Medical Language System) der NLM, eine auf MeSH basierende Ontologie (U.S. National Library of Medicine 2016b). Die genannten Studienregister bieten dem Nutzer neben Booleschen Retrieval im Rahmen einer Volltextsuche ein webbasiertes Formular zur Spezifikation bestimmter Studieneckdaten. Die aktuelle Literatur liefert zu Funktionalitäten und User-Interfaces klinischer Studienregister nur rudimentär wissenschaftliche Veröffentlichungen –Gough et al. bestätigen diese Einschätzung (Gough et al. 2012, S. 125).

Für den Nutzer stehen für die Datenbank ClinicalTrials.gov drei Mechanismen der Suche zur Verfügung: Eine einfache Suche, eine so genannte fokussierte Suche zur Eingrenzung von Suchparametern auf bestimmte Kriterien sowie ein Browsing-Modus, beispielsweise nach Krankheitsbildern. Zusätzlich bietet die Suchoberfläche eine Auto-Complete Funktion mit integrierten Synonymen sowie einen Spell-Checker zur rechtschriftlichen Überprüfung an (Tse et al. 2002). Weitere Autoren berichten 2007 und später 2012 über Änderungen an der Nutzeroberfläche der webbasierten Datenbank ClinicalTrials.gov: So verfügt die Datenbank seit 2007 über eine Facettensuche nach Publikationsdatum und Studientyp sowie verschiedene Ansichten zur Organisation einzelner Suchanfragen Dies wurde 2012 weiter ausgebaut. (Williams und Tse 2007; Huston et al. 2014).

Aufgrund der höchst unterschiedlichen konzeptionellen Struktur, uneinheitlicher Indexierungsmechanismen als auch unterschiedlicher, teils vergleichsweise rudimentärer Benutzeroberflächen (vgl. auch Abschnitt 3.5.1.1) kann für die Suche in Studienregistern kein Standardvorgehen geschildert werden.

Empfehlungen zur Suche in klinischen Studienregistern sind in der Fachliteratur vergleichsweise wenig beschrieben: Generell sind Suchanfragen in klinischen Studienregistern weniger strukturiert, Hilfsschemata wie PICO kommen nicht zu Einsatz (Hausner und Kaiser 2010). Methodisch werden die großen klinischen Studienregister ClinicalTrials.gov und ICTRP idealerweise nur mit einem Building Block durchsucht, hierbei werden zur Sicherung der Sensitivität ausreichend Synonyme verwendet (Glanville et al. 2014).

3.5.1.2 Die Rolle von Suchmaschinen

Wichtig scheint es, im Kontext vor allem der Grauen Literatur, die Rolle der Suchmaschinen wie Google und insbesondere des Services Google Scholar (Google 2016) aus Sicht der Methodologie systematischer Übersichtsarbeiten zu beleuchten.

Mit der Veröffentlichung des Dienstes Google Scholar bieten sich Wissenschaftlern vollkommen neue Möglichkeiten der Fachrecherche, was auch in medizinischen Fachkreisen schnell diskutiert wurde. Scholar bietet Zugriff auf eine Vielfalt verschiedener Publikationstypen wie beispielsweise Primärliteratur, Hochschulschriften, Bücher, Abstracts aus verschiedenen Quellen wie Verlagen, Fachgesellschaften, Online-Repositories, Universitäten und anderen Websites (Google 2013).

Google Scholar ist damit nicht nur inhaltlich breiter gestreut als beispielsweise medizinische Fachdatenbanken (Gehanno et al. 2013), sondern ermöglicht gleichzeitig den Zugriff auf eine Vielzahl unterschiedlicher Quellentypen, die in anderen professionellen Suchkontexten wie der domänenspezifischen Suche in den Rechtswissenschaften durchaus eine Rolle spielen, vgl. auch (Mielke und Wolff 2008). Das Retrieval in Google Scholar erfolgt entgegen der klassischen bibliographischen Fachdatenbank in Form von Freitext-Anfragen und erfordert keine Booleschen-Operatoren. Die Ergebnisse werden wie bei Suchmaschinen üblich als gerankte Liste ausgegeben.

Einige Vor- und Nachteile der Nutzung von Google Scholar für die Wissenschaft werden 2005 unter anderem bei Noruzi (Noruzi 2005) genannt, nicht alle sind heute noch so bedeutungsvoll wie vor einigen Jahren (beispielsweise die Verfügbarkeit über das Internet):

Tabelle 3-13 Vor- und Nachteile des Services Google Scholar, nach (Noruzi 2005)

Vorteile	Nachteile
Internationales Angebot an wissenschaftlichen Inhalten	Klar erkennbarer Sprach-Bias
Möglichkeit einer breit angelegten, umfangreichen und interdisziplinären Suche	Inkonsistenz im Zitationsstil
Objektive Auswahl von Journals und somit geringe Gefahr von Publikationsbias	Unvollständige Autorennamen bei Mehrautorenschaft
Vielfältige Auswahl an Publikationsformen	Eingeschränkte Quellen und fehlende Journals
Fehlerrobustheit bei Suchanfragen	Keine angewandten Suchmöglichkeiten und keine Indexierung
Flexible Abfragesprachen	
Freie Verfügbarkeit über das Internet	

So deutlich die Vorteile sichtbar werden, ist ein Punkt für die professionelle, methodologisch gesteuerte Recherche im Kontext der evidenzbasierten Medizin kritisch: Über den Ranking-Algorithmus von Google Scholar ist wenig bekannt. Untersuchungen von Beel und Gipp lassen unter anderem folgende Rückschlüsse zu, die allerdings heute schon nicht mehr aktuell sein müssen (Beel und Gipp 2009):

- der Ranking-Algorithmus beruht insbesondere auf der Anzahl der Zitationen eines Artikels: Google Scholar spielt seine Stärken daher dann aus, wenn häufig zitierte Standardliteratur oder aber die neuesten Trends recherchiert werden
- Titelwörter werden für das Ranking höher gewichtet als Terme im Textkorpus
- Google Scholar berücksichtigt keine Synonyme zu dem vom Nutzer verwendeten Suchbegriffen
- das Ranking berücksichtigt keine Term-Häufigkeiten
- neuere Artikel werden älteren Artikeln gegenüber bevorzugt
- die Rankingalgorithmen in der Schlagwortsuche, der Suche nach „*related articles*“ und der Zitationssuche unterscheiden sich
- es erfolgt eine hohe Gewichtung von Autorennamen und Journal Namen

Ein wesentlicher Kritikpunkt an Google Scholar, den aus Sicht der EBM wurde bereits 2005 benannt (Henderson 2005):

Der Suchalgorithmus, der Treffer unter anderem auf Basis von Zitationen rankt (Google

2013), könnte zu systematischen Fehlern durch Überrepräsentation einzelner, vielzitiert veröffentlichten führen.

Von Beginn an wurde dringend davon abgeraten, Google Scholar als einzige Quelle für medizinische Fachinformationen heranzuziehen (Giustini 2005), insbesondere nicht für systematische Literaturrecherchen (Wang und Howard 2012). Unbestritten ist hingegen, dass Google und Google Scholar zumindest in der nicht-professionellen Recherche (gemeint ist hier auch die Recherche durch Domänenexperten) schnell zur Hauptquelle für medizinische Informationen wurden (Gehanno et al. 2013). Welchen Stellenwert Google Scholar in der Fachrecherche für systematische Überblicksarbeiten haben sollte, wird seither in der Fachwelt teils hitzig diskutiert (Gehanno et al. 2013; Giustini et al. 2013; Boeker et al. 2013).

Methodologisch wiesen Gehanno et al. 2013 im Test gegen einen vorab identifizierten „Gold-Standard“ aus systematischen Überblicksarbeiten zwar einen 100% Recall bei Google Scholar-Suchen (Gehanno et al. 2013), d.h. es war mittels Suchen in Google Scholar möglich, alle im „Gold-Standard“ gesammelten Zitationen, darunter auch Graue Literatur, vollständig zu identifizieren, was mittels strukturierter Suchen nicht gelang. Die Autoren folgerten daraus (Gehanno et al. 2013)

“the coverage of GS [Anmerkung: GS = Google Scholar] is much higher than previously thought for high quality studies. GS is highly sensitive, easy to search and could be the first choice for systematic reviews [...]. It could even be used alone. It just requires some improvement in the advanced search features to improve its precision and to become the leading bibliographic database in medicine.”

Boeker et al. zeigen methodologisch, wie die für bibliographische Fachdatenbanken ausgearbeiteten Suchprofile in der Praxis aufgrund technischer Restriktionen des Google Scholar-Interfaces reduziert und angepasst werden müssen, um sie auf Google Scholar überhaupt anwenden zu können und konnten die Schlussfolgerungen („some improvement“) von Gehanno et al. damit widerlegen (Boeker et al. 2013). Die Autoren kritisierten dabei insbesondere die mangelnde Precision von Google Suchanfragen, vgl. auch (Walters 2008). Große Treffermengen, aber eine schwankende und zum Teil geringe Precision waren auch das Ergebnis anderer Studien (Anders und Evans 2010; Bramer et al. 2016).

Im Vergleich zu Interfaces medizinischer und anderer Fachdatenbanken unterliegt die Suche über Google Scholar, wie bereits deutlich wurde, in der Praxis einigen Restriktionen, vgl. (Boecker et al. 2013; Bramer et al. 2013; Bramer et al. 2016):

- Gegenüber Fachdatenbanken dürfen Suchanfragen bei Google Scholar sowohl in der einfachen als auch angewandten Suche 256 Zeichen nicht überschreiten
- die Trefferanzeige ist immer auf 1000 Einträge limitiert, ein Export ist ausschließlich in Schritten von 20 Items möglich
- Google Scholar verfügt über kein kontrolliertes Vokabular.
- Google Scholar verfügt nur eingeschränkt über Möglichkeiten der Trunkierung und Maskierung
- Boolesche Operatoren sind nur in einer Ebene verschachtelbar einsetzbar
- die Reihenfolge der Suchterme beeinflusst das Ranking der Treffer
- komplexe Boolesche Queries sind aufgrund des begrenzten Nutzerinterfaces nicht oder nur auf Basis komplexer Vorarbeiten möglich
- die Aktualität des Datenbestands unterliegt langen Intervallen
- Google Scholar verfügt über keinerlei weiterführende Referenzen auf ggf. fehlende Volltexte
- die Kombination von Suchfeldern ist nicht möglich
- ein Export von Treffermengen ist nicht möglich

Tabelle 3-13 führt die Fehlerrobustheit von Suchanfragen auf Google Scholar als großen Vorteil auf. Dies scheint für die Durchführung systematischer Übersichtsarbeiten irrelevant. Es wurde nämlich deutlich, dass Google Scholar trotz seines Ranking-Retrieval-Verfahrens, das in Suchanfragen ohne Operatoren auskommt, in der systematischen Übersichtsarbeit genutzt wird wie eine bibliographische Datenbank. Es erfolgt also keine Freitext-Suchanfrage, sondern eine auf die restriktiven Booleschen Funktionalitäten von Google Scholar hin angepasste strukturierte Suchanfrage. Insbesondere versuchen die Nutzer, ihre für Fachdatenbanken entwickelten Suchprofile für die Nutzung in Google Scholar zu adaptieren – und scheitern hier aufgrund der von Booleschen Datenbanken abweichenden, den Nutzern weitestgehend unbekanntem Funktionsweise von Google Scholar in der strikten Anwendung der Methodenlehre.

Die ideale Einsatzmöglichkeit der Suchmaschine formulieren Gray et al.: *“Google Scholar may be useful for initial and supplemental information gathering, but lacks a deeper reliability*

than other existing services currently provide scholars." (Gray et al. 2012):

Google Scholar ist zwar nützlich für eine initiale oder unterstützende Informationssuche, Hauptkritikpunkt in der Nutzung von Google Scholar ist aber die mangelnde Reliabilität von Suchprozessen. Im Kontext der medizinischen Forschung wie auch medizinischer Übersichtsarbeiten wird der Einsatz der wissenschaftlichen Suchmaschine Google Scholar also trotz besserer Performanz weiterhin kritisch beurteilt.

3.5.2 Methodologische Anforderung

Die Suche in klinischen Studienregistern ist eine Spezialität medizinischer Disziplinen, sie wird daher auch nur in der medizinischen Methodologie beschrieben. Die Notwendigkeit der Suche dient der Fehlerrobustheit sowie der Absicherung der Vollständigkeit durch die potenzielle Identifikation noch laufender oder nicht publizierter Studien. Sie ist ausschließlich für Cochrane Reviews zwingend vorgeschrieben und wird auch beim IQWiG durchgeführt (Chandler et al. 2013; Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen (IQWiG) 2015).

Die Handsuche ist ein zusätzliches Element in der systematischen Übersichtsarbeit zur Vermeidung von Biases, beispielsweise aufgrund fehlerhafter Indexierung oder unvollständiger, teils englischsprachig fokussierter Erfassung in Fachdatenbanken (Blümle und Antes 2008). Auch die gewachsene Funktionalität von Datenbanken ist ein Grund für die Notwendigkeit von Handsuchen: Die in der systematischen Übersichtsarbeit bei MEDLINE verfügbaren wichtigen Publikationstypen „*randomized controlled trial*“ und „*controlled clinical trial*“ beispielsweise wurden erst in den 1990er Jahren in den Medical Subject Headings eingeführt (Hopewell et al. 2007a). Vor diesem Datum publizierte Ergebnisse randomisierter Studien sind ohne passenden Suchfilter daher schwer auffindbar. Die Handsuche umfasst zum einen Journals, die nicht in gängigen bibliographischen Datenbanken gelistet sind, zum anderen aber auch Graue Literatur, beispielsweise also Abstract Datenbanken oder Dissertationsdatenbanken, siehe Abschnitt 3.3.2).

Methodologisch ist der Begriff der Grauen Literatur ebenfalls eng verknüpft mit dem der Handsuche, in jedem Fall aber mit „unkonventionelleren“ und weniger strukturierten Such- und Beschaffungsmethoden als bei der Suche in bibliographischen Datenbanken.

Ob und wann die Suche in Grauer Literatur als Suchstrategie in Erwägung gezogen werden sollte, wird in der Literatur vielfach an der initialen Fragestellung ausgemacht. Während eine systematische Literaturrecherche für alle Fragestellungen, die sich klar und prägnant in ein PICO-Schema übertragen lassen, exzellente Ergebnisse für Forschung und klinischen Alltag liefern kann, wird dies mit zunehmender Komplexität der Konzepte „Intervention“ und „Outcome“ schwieriger (Pawson et al. 2005; Benzies et al. 2006).

So findet sich beispielsweise folgende Entscheidungshilfe: Je mehr der folgenden Fragen mit Ja beantwortet werden, desto mehr spricht dies für die Durchführung einer Suche in der grauen Literatur (Benzies et al. 2006):

- Handelt es sich um eine komplexe Intervention?
- Ist die Intervention kontextabhängig?
- Handelt es sich um einen komplexen *Outcome*?
- Wie gut ist der *Outcome* messbar?
- Wie hoch ist der Grad der Evidenz?
- Wie hoch ist die Qualität der Evidenz?

Conn et al. stellen verschiedene Methoden und Informationsquellen, die in der systematischen Übersichtsarbeit im Allgemeinen genutzt werden in den Kontext der Suche nach Grauer Literatur (Conn et al. 2003b):

- Überprüfung von Datenbanken
- bibliographische Suchen
- Zitationssuchen
- Einbeziehung wissenschaftlicher Fach-Communities
- Handsuchen in Journals
- Kontakt mit Kollegen
- Tagungsbände und Abstracts
- Internetsuche

Einige der dort vorgeschlagenen Methoden werden bei Conn et al. anhand der Maße Recall und Precision miteinander verglichen:

Die bibliographische Suche sollte nicht nur die eigentlichen Literaturangaben, sondern auch die im Kontext von Primärliteratur geführten Diskussionen beinhalten. Sie führt daher typischerweise zu einer hohen Sensitivität und insbesondere bei großen Datenbanken zu mäßiger

Spezifität.

Die Zitationssuche liefert sehr ähnliche Artikel bei je nach Quelle sehr unterschiedlicher Qualität von Sensitivität und Spezifität.

Die Handsuche in Journals wird hinsichtlich der Precision als sehr positiv gegenüber klassischen Suchprozessen beurteilt.

Der Kontakt zu Kollegen scheint bei der Suche nach Grauer Literatur ein wesentlicher und erfolgversprechender Ansatz. (Conn et al. 2003a)

Besonders problematisch erweist sich die Beschaffung nicht veröffentlichter Daten industriegesponserter Studien. In einer Umfrage zeigte sich, dass die Beschaffung nicht-publizierter pharmazeutischer Studienergebnisse häufig über direkten Kontakt mit den Auftraggebern oder Forschern erfolgt und der Erfolg damit stark von der Kooperationswilligkeit beispielsweise von Pharmaunternehmen abhängig ist. Der Kontakt zu Autoren scheiterte häufig daran, dass diese für die Daten nicht verantwortlich waren, aber auch daran, dass Autoren nicht bereit waren die Daten und Ergebnisse weiterzugeben. Dies sollte laut Umfrage zur direkten Kontaktaufnahme mit den Auftraggebern, beispielsweise Pharmakonzernen, führen. Hier allerdings hatten viele Studienteilnehmer bereits negative Erfahrungen gemacht: Häufig werden grundsätzlich nur Zusammenfassungen oder Protokolle öffentlich gemacht, ebenso häufig werden die Ergebnisse nur bestimmten Fachkreisen übermittelt oder Pharmakonzerne wollen genaue Auskunft darüber, wie die Daten verwendet werden (Wolfe et al. 2013)

3.5.3 Zusammenfassung

Charakteristika von Informationsquellen wie klinischen Studienregistern, Fachdatenbanken für Graue Literatur als auch Suchmaschinen führen in der formalen und thematischen Suche zu Suchprozessen, die weniger umfangreich und weniger strukturiert sind als Suchanfragen in klassischen Retrieval-Prozessen.

Unbestritten spielt auch in der systematischen Übersichtsarbeit vor allem Google Scholar für die Suche nach Grauer Literatur eine entscheidende Rolle. Trotzdem führen technische Implementierung als auch die restriktiv implementierten Syntaxspezifika mit Boolescher Funktionalität bei Google Scholar zu deutlicher Kritik.

Dies zeigt, dass die klassische Form der systematischen Übersichtsarbeit bisher ganz klar mit dem Konzept Boolescher Datenbanken verwurzelt ist und in der – teils notwendigen, teils geforderten – Nutzung weiterer Informationsquellen an ihre methodischen Grenzen stößt.

3.6 Systematische Übersichtsarbeiten als „Information Interaction in Context“- Problem

Die systematische Übersichtsarbeit unterscheidet sich in ihrer Struktur aus Suche, Analyse, kritischer Bewertung und Synthese wissenschaftlicher Forschungsergebnisse zu einer vorab klar definierten Forschungsfrage nicht wesentlich von anderen Formen der Literaturrecherche. Ihr Anspruch, der sich aus den Anforderungen der Vollständigkeit, Fehlerrobustheit und Reproduzierbarkeit in allen diesen Teilschritten ergibt, macht sie zur strengsten Form der Übersichtsarbeit, die durch eine umfangreiche Methodologie untermauert wird.

Aus informationswissenschaftlicher Sicht stellt sich die systematische Übersichtsarbeit als methodologisch höchst restriktiver, zeit- und ressourcenintensiver Prozess mit klarer Zielvorgabe dar. Suchaufgaben unterliegen technischen Restriktionen verfügbarer Informationsquellen und sind geprägt sind von Verzerrungen in der medizinischen Forschung und Literatur.

Der folgende Abschnitt stellt die aus der Theorie und Methodenlehre bekannten Charakteristika einzelner Entitäten der systematischen Übersichtsarbeit wie in den vorhergehenden Abschnitten dargestellt in einem rudimentären, dem Modell nach Ingwersen und Järvelin gelehnten *Information Interaction in Context* Modell zusammen. Dies führt zu den noch offenen Forschungsfragen. Die folgenden Darstellungen lehnen sich demnach ebenso an die Arbeiten von Ingwersen und Järvelin an, siehe hierzu Abbildung 2-21.

3.6.1 Kontext Information und Information Retrieval-System

Das Paradigma der EBM wie auch die Methodenlehre der systematischen Übersichtsarbeit geben ein differenziertes Bild über den Kontext verfügbarer Informationen wie auch Informationsressourcen, siehe Abschnitt 3.2.

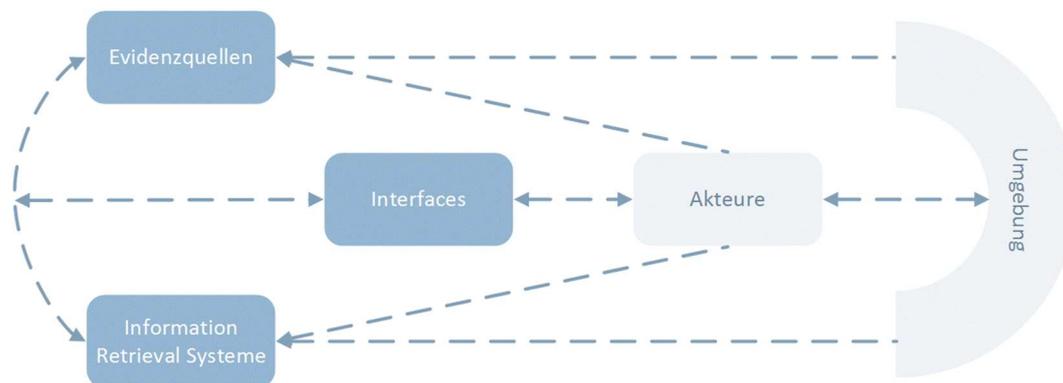


Abbildung 3-16 Kontext Dokument und Information Retrieval-System, Darstellung nach (Ingwersen 1996)

Ergebnisse klinischer Studien sind in der systematischen Übersichtsarbeit zur EBM wertvolle Evidenzquellen, die randomisierte klinische Studie stellt laut Evidenzpyramide die ideale, weil valideste Evidenzquelle dar (siehe Abschnitt 3.1.1). Aufgrund systematischer Verzerrungen in der Literaturlandschaft liegen diese Ergebnisse nicht nur in Form von Primärliteratur vor, sondern sind auch als Datensätze in klinischen Studienregistern oder über den Kontakt zu Forschergruppen und Pharmaunternehmen verfügbar (siehe Abschnitt 3.1.2). Die Menge verfügbarer Information ist daher per se inhomogen und reicht von strukturierten bibliographischen Referenzen über Volltexttexte bis hin zu inhomogenen Datenbankstrukturen oder unstrukturierten Textdokumenten, beispielsweise in Grauer Literatur oder auch als Beigabe in klinischen Studienregistern. Wichtige Informationsquelle der systematischen Übersichtsarbeit ist die systematische Übersichtsarbeit selbst. Sie beinhaltet nicht nur evidenzbasierte Ergebnisse, sondern ist auch Hilfsmittel zur Identifikation weiterer evidenzbasierter Ergebnisse. Suchprozesse erstrecken sich nicht nur auf bibliographische Datenbanken als primäre Informationsressource, sondern verteilen sich auf Systeme mit unterschiedlichen Datenstrukturen und unterschiedlichen Dokumenten,- und Datenrepräsentationen.

Da die Ergebnisse klinischer Studien meist in Form von Originalarbeiten und systematischen Übersichtsarbeiten vorliegen, sind bibliographische Fachdatenbanken die primäre Ressource für die systematische Literaturrecherche. Diese ist damit wiederum gekoppelt an Retrieval auf *exact match* basierenden Systemen mit Boolescher Retrieval-Technologie. Sie tragen wesentlich zur Reproduzierbarkeit von Suchprozessen und damit zur Einhaltung methodologischer Ansprüche bei.

Die verfügbaren Systeme (siehe unter anderem Abschnitte 3.2.1 und 3.2.2) sind als Hochleistungsdatenbanken mit mächtigen Thesauri und an spezifische Nutzergruppen angepasster Funktionalität und Benutzeroberfläche charakterisierbar (siehe 3.4.2). In Zusammenarbeit mit federführenden Institutionen wurden einzelne bibliographische Fachdatenbanken angepasst, um dem methodologischen Anspruch systematischer Übersichtsarbeiten gerecht werden zu können, so beispielsweise die Einführung des Thesaurus-Terms „*systematic review as topic*“ in den MeSH-Thesaurus (vgl. (Lefebvre et al. 2013)).

Auch bei klinischen Studienregistern oder den in der Methodologie vorgeschlagenen, teils verpflichtenden Datenquellen zu Grauer Literatur überwiegen Systeme mit Boolescher Retrieval-Technik (siehe Abschnitte 3.2.5 und 3.5). Wie sehr die Nutzung von Datenquellen auf die möglichst vollständige Reproduzierbarkeit von Suchprozessen fokussiert ist, zeigt die Tatsache, dass die Suche über Suchmaschinen methodologisch fast schon verpönt ist, auch wenn sie in der Praxis ähnlich vollständige Treffermengen liefern kann (siehe Abschnitt 3.5.1.2).

Zusammenfassend ist der Dokumenten- und Information Retrieval-Kontext stark methodologisch geprägt, Systeme werden zwar angepasst, berücksichtigen aber ausschließlich die für die Qualitätssicherung nötigen Maßnahmen, jedoch kaum den Nutzer dieser Systeme. Mehr noch projiziert die Methodenlehre die Anforderungen, die aus informationswissenschaftlicher Sicht an ein Retrieval-System gestellt werden, auf den Nutzer: Er muss alle Probleme, die während Indexierung und Retrieval passieren könnten, durch ausgefeilte Suchanfragen kompensieren können, um den Vollständigkeitsanspruch der systematischen Übersichtsarbeit gewährleisten zu können. Für die systematische Übersichtsarbeit macht dies den Bibliothekar oder Informationsspezialisten als Mittelsmann und Exekutive strukturierter Suchprozesse unentbehrlich.

3.6.2 Kontext Arbeits- und Suchaufgaben und Methodenlehre

Arbeitsaufgaben und die daraus resultierenden klassischen Suchaufgaben sind, wie bereits an anderer Stelle deutlich gemacht, untrennbar mit der Methodenlehre systematischer Übersichtsarbeiten verbunden.

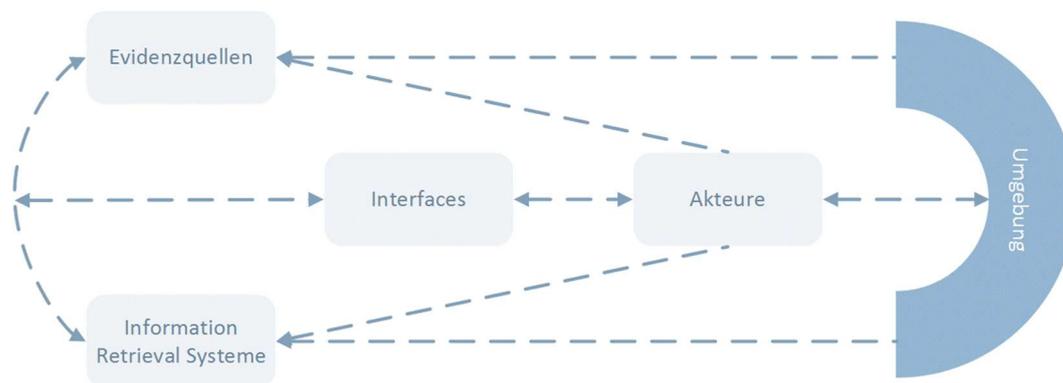


Abbildung 3-17 Kontext Arbeits- und Suchaufgaben, Darstellung nach (Ingwersen 1996)

Diese Methodologie fußt auf den grundlegenden Arbeiten von Institutionen, die sich streng den Grundsätzen evidenzbasierter Medizin verschrieben haben. Diese umfasst nicht nur den Suchprozess an sich, sondern auch die Vorbereitung der Suche, die Auswahl anhand der vordefinierten Forschungsfrage geeigneten Studien sowie Veröffentlichungen und deren kritischer Bewertung, die Synthese, Berichterstattung und Protokollierung der systematischen Übersichtsarbeit, auf die hier teils nicht oder nur kurz eingegangen werden konnte.

Suchprozesse schließen aufgrund systematischer Verzerrungen in der Literaturlandschaft (*biases*) nicht nur das Retrieval auf bibliographische Fachdatenbanken ein, sondern auch die formale, thematische Suche in klinischen Studienregistern, die Suche in Grauer Literatur sowie klassische Suchtaktiken, wie informationswissenschaftlich bereits umfangreich beschrieben.

Die Entwicklung und spätere Dokumentation so genannter *Suchstrategien* für das Retrieval in bibliographischen Fachdatenbanken, im informationswissenschaftlichen Sinn Suchanfragen, sind unerlässliche Aufgaben und werden methodologisch ausführlich beschrieben. Methodologisch beginnt dieser Prozess üblicherweise mit dem Aufbau einer Ground Truth, die sowohl als Referenzmenge als auch Quelle für Deskriptoren und Freitextterme fungiert.

Die Entwicklung von Suchprofilen führt aufgrund der theoretischen Anforderungen systematischer Übersichtsarbeiten zu zeitintensiven Suchprozessen, die in regelmäßigen Abständen von Entscheidungen hinsichtlich Vollständigkeit und Relevanz aufgefundener Dokumente und Daten, aber auch der Fehlerrobustheit des gesamten Suchprozesses geprägt sind.

Die der strukturierten Suche gegenüberstehenden formalen und thematischen Suchprozesse sind methodologisch zwar teils vorgeschrieben, aber deutlich seltener Gegenstand der Forschung und damit in der reinen Methodenlehre weitestgehend unbeachtet. In der Literaturlandschaft zur systematischen Übersichtsarbeit finden sich vorwiegend quantitativ nach dem Vorbild interventioneller Reviews durchgeführte Übersichtsarbeiten, die in ihrer Forschungsfrage auf die Einhaltung von Methodologien oder der Verbesserung von Retrieval-Ergebnissen hin abzielen (vgl. (Hopewell et al. 2007b; Horsley et al. 2011)). Vor allem in systematischen Übersichtsarbeiten zu komplexen Interventionen oder in den Sozialwissenschaften spielen sie aufgrund terminologischer Restriktionen und der Verteilung von Information auf viele Informationsressourcen jedoch eine wesentliche Rolle.

Ausgehend vom *work task*, einer „nach *systematischen* Methoden durchgeführten Identifikation *aller* verfügbaren Evidenz bezüglich einer vorab strukturierten medizinischen Forschungsfrage“, lassen sich im Suchprozess interventioneller Reviews auch Suchaufgaben festmachen:

Suche nach existierenden Reviews, Entwicklung der Ground Truth

Die Entwicklung der Ground Truth (siehe Abschnitt 3.4.4.2) dient in erster Linie der Generierung einer validierten Referenzmenge, über welche die Vollständigkeit strukturierter Suchprofile evaluiert werden kann. Methodologisch wird die Ground Truth auch als Hilfsmittel zur Identifikation von Suchtermen beschrieben.

In der Praxis werden Ground Truth Mengen wie auch „Gold-Standards“ häufig durch Handsuchen generiert. Ob und wie diese Ground Truth objektiv und valide ermittelt werden kann und in der Praxis ermittelt wird, ist bisher jedoch weitestgehend unklar. Dieses Dilemma kann in der Praxis zu fehlerbehafteten Ausgangssituationen führen und damit die Verzerrung des Gesamtergebnisses ungewollt bereits an dieser Stelle initiieren.

Identifikation von Suchtermen

Die Identifikation von Suchtermen in Form kontrollierter Vokabularien als auch Freitextterme ist zentrale Aufgabe bei der Entwicklung vollständiger und reproduzierbarer Suchprofile.

Als Struktur sind Suchprofile eingehend methodologisch beschrieben und hinsichtlich ihrer Performanz über Referenzmengen und klassische Retrievalmaße auch messbar. Wie Suchterme in der Praxis ermittelt werden, ist bisher ebenso unklar wie ein potenzieller und zu vermutender Informationsbedarf, der sich während der Identifikation geeigneter Deskriptoren und Freitextterme ergibt. Einzige Vorarbeiten hierzu liefern mit Ausnahmen die Arbeiten von Booth sowie Arksey und O'Malley (Arksey und O'Malley 2005; Booth 2008).

Strukturierte Suche nach Evidenzquellen

Die strukturierte Suche nach Evidenzquellen zielt auf eine möglichst vollständige Identifikation thematisch relevanter Primärliteratur ab. Sie ist als iterativer Prozess direktes Ergebnis der Entwicklung datenbankspezifischer Suchprofile und fokussiert sich auf Retrievalprozesse auf Booleschen Datenbanken.

Strukturierte Suchprozesse sind in ihrem Umfang und Anspruch methodologisch klar vorgegeben und durch Queries und Speicherungsfunktionalitäten einzelner Datenbanken sowohl messbar als auch reproduzierbar.

Seit Jahren werden, wie einführend bereits dargestellt, Methoden entwickelt und erprobt, mit deren Hilfe die Durchführung strukturierter Suchen im Rahmen systematischer Übersichtsarbeiten erleichtert und beschleunigt werden kann. Hierbei spielen Text Mining-Technologien eine große Rolle. So entstanden im Laufe der vergangenen Jahre eine Reihe an Softwarelösungen zur Automatisierung von Teilprozessen und damit ein Flickenteppich an Tools, deren methodologischer Nutzen allerdings ausschließlich auf Performance Ebene betrachtet wird. Neben der Tatsache, dass teilweise der Nutzen dieser Tools und Automatismen empirisch nicht belegt ist, stellt sich die grundsätzliche Frage, welche der einzelnen Teilschritte überhaupt automatisiert werden können und wie der Workload von Nutzern reduziert werden kann, vgl. (Tsafnat et al. 2014).

Formale und thematische Suche nach Evidenzquellen

Formale und thematische Suchen sichern die Vollständigkeit und Fehlerrobustheit strukturierter Suchen. Neben der Identifikation von Dokumenten aus Grauer Literatur umfasst die formale und thematische Suche auch reine Daten.

Die ausgehend von Definition und in der systematischen Übersichtsarbeit verortbaren Suchaufgaben zielen im Wesentlichen auf die Durchführung und Evaluierung strukturierter Suchprozesse ab.

Während die Suche in bibliographischen Fachdatenbanken als klassischer *search task* (siehe Abschnitt 2.5.3) verstanden werden muss, sind die Entwicklung einer Ground Truth als auch die Identifikation von Suchtermen zur Erarbeitung vollständiger Suchprofile typische *seeking tasks* mit teils unklarem Informationsbedarf. In der Methodologie und Fachliteratur sind mögliche Suchtaktiken zwar beschrieben (siehe Abschnitte 3.3.2 und 3.4.5) jedoch nicht empirisch ermittelt. Bei der formalen und thematischen Suche ist die Formalisierung des Informationsbedarfs aufgrund meist technischer Restriktionen erschwert.

Die in der systematischen Übersichtsarbeit identifizierbaren und teils auch in der weiterführenden Fachliteratur beschriebenen Suchaufgaben geben den Rahmen für umfassende und – so zu vermuten – teils komplexe Retrievalprozesse vor. Methodologisch sind diese insbesondere im Hinblick auf die Performanz von Suchprozessen gut beforscht. Bisher existiert jedoch kein Modell, das die in der systematischen Literaturrecherche anfallenden *information seeking tasks* und *information searching tasks* vollständig identifiziert, klassifiziert und im nächsten Schritt hinsichtlich ihres Informationsverhaltens im Kontext analysiert.

3.6.3 Nutzer und Interaktion als unbekannter Kontext

Während die Entitäten Dokumente, Informationsquellen als auch die Methodologie Umgebungsvariablen der systematischen Übersichtsarbeit darstellen, die fundiert charakterisiert werden können, gibt es aus Sicht der *Information Retrieval Interaction* zwei zentrale Komponenten, die von der Forschung bisher nicht eingehend berücksichtigt wurden: die Nutzer, deren Interaktion mit Retrieval-Systemen und der im Kontext aller Umgebungsvariablen resultierende kognitive Workload.

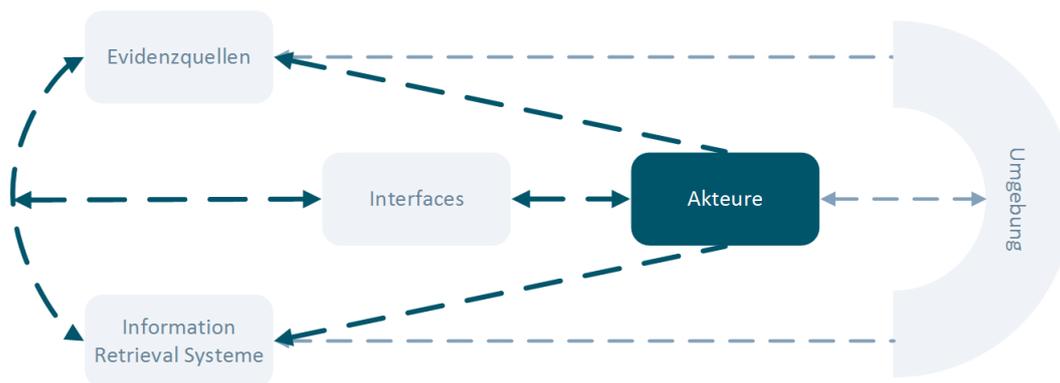


Abbildung 3-18 Unbekannter Kontext: Nutzer und Interaktion, Darstellung nach (Ingwersen 1996)

Nutzer sind in diesem Fall vorwiegend Informationsspezialisten, auch wenn es Usus ist, systematische Übersichtsarbeiten in einem interdisziplinären Team aus Domänenspezialisten und Informationsexperten durchzuführen (siehe Abschnitt 1.1.4). Die Rolle des Informationsspezialisten für die Recherche im Rahmen anspruchsvoller Reviews wurde bereits früh herausgestellt (Schell und Rathe 1992; Mead und Richards 1995). Die Skills, die Informationsspezialisten für die systematische Übersichtsarbeit unentbehrlich machen, betreffen vor allem die Rechercheexpertise und das Wissen um technische Anforderungen und Restriktionen einzelner Fachdatenbanken (McGowan und Sampson 2005). Sie zielen drauf ab, die methodologisch hohen Anforderungen an systematische Übersichtsarbeiten garantieren zu können. Heute sind Informationsspezialisten fester Bestandteil des Review Teams, vgl. u.a. (Harris 2005; McKibbon 2006). In der Praxis sind die Rollen, die dem Informationsspezialisten in diesem Team zukommen, heute vielfältig. Er fungiert häufig als Projektleiter, Manager, als Spezialist für die Literatursuche und das Literaturmanagement, ist aber auch für die Literaturbeschaffung und die Bewertung von Zitationen und Dokumenten zuständig. Darüber hinaus sind Informationsexperten auch in die Synthese und Berichterstattung involviert (vgl. (Beverley et al. 2003)).

Der Vollständigkeit halber soll erwähnt werden, dass nicht nur bei der Durchführung systematischer Literaturrecherchen Informationsspezialisten und Bibliothekare in medizinische Recherchetätigkeiten eingebunden werden: Abgegrenzt werden muss die Tätigkeit im Rahmen der systematischen Übersichtsarbeit vor allem gegenüber „Clinical Medical Librarian“ (CML) Programmen wie sie in englischsprachigen Ländern bereits in den 1970er Jahren etabliert wurden (Cimpl 1985). In CML Programmen ist ein Bibliothekar oder Informationsspezialist Teil

eines klinisch tätigen Teams und unterstützt den Kliniker *at the point of care* bei der Informationsbeschaffung zu medizinisch Fragestellungen aus dem klinischen Alltag heraus.

Ziel des CML Programms war es, medizinische Teams schnell mit Informationen zu versorgen, auf der anderen Seite aber auch, das Informationsverhalten von Klinikern zu schulen, während gleichzeitig schlicht erkannt wurde, dass der Bedarf an nutzer-zentrierten Ansätzen der Informationssuche im klinischen Alltag gegeben war – durch die Einstellung so genannter *clinical librarians* (Cimpl 1985). In Deutschland sind derartige Programme bis heute nicht etabliert.

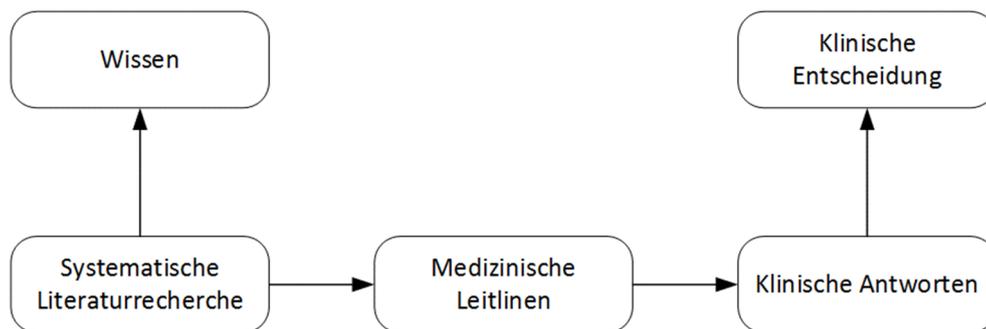


Abbildung 3-19 Systematische Literaturrecherchen: Wissen und klinische Entscheidungen (Abbildung nach (Booth 2008, S. 583))

CML Programme erfordern daher eine *spezifische* und *relevante* Suche, die aber schon aufgrund zeitlicher Restriktionen keinen Anspruch auf Vollständigkeit haben kann; so gesehen sollten sie einem Suchprotokoll folgen, um wenigstens die Wahrscheinlichkeit zu erhöhen, alle potenziell wichtigen Ressourcen identifiziert zu haben (Booth 2008). Rahmenbedingungen wie diese, so findet sich nicht-empirisch belegt bei Booth, beeinflussen auch die Auswahl adäquater Suchstile und Suchtaktiken. (Booth 2008)

Die Differenzierung von Literaturrecherchen zum Zwecke des Wissenserwerbs (und in der Folge zur Erstellung medizinischer Leitlinien oder ähnlicher Regelwerke) und auf der anderen Seite zur klinischen Entscheidung verdeutlicht gleichzeitig noch einmal, in welchem Rahmen systematische Übersichtsarbeiten durchgeführt werden. Zentral ist in beiden Fällen die Rolle des Informationsspezialisten als „Zugang zur Information“. Nach wie vor ist er also der *Mittelsmann* (siehe Abschnitt 2.1) zwischen dem antragstellenden Domänenexperten mit Informationsbedarf und einem in den vergangenen Jahrzehnten exponentiell wachsenden Dokumenten- und Informationsbestands.

3.7 Forschungsfragen

Das Modell von Ingwersen (Ingwersen 1996), sowie Ingwersen und Järvelin (Ingwersen und Järvelin 2005) setzt Nutzerverhalten in den Kontext verschiedener Ebenen, der vorliegende Absatz dient dazu, die in Kapitel 3 charakterisierten Parameter und Rahmenbedingungen in dieses Modell einzuarbeiten. Hinsichtlich der im Folgenden beschriebenen empirischen Untersuchungen zum Informationsverhalten stellt sich die systematische Übersichtsarbeit als „Information Interaction in Context“- Problem dabei bisher unvollständig dar:

Wesentlich geprägt ist die systematische Übersichtsarbeit durch das Paradigma der evidenzbasierten Medizin und der darauf fußenden Methodenlehre. Die Methodenlehre, die nicht nur der Qualitätssicherung systematischer Übersichtsarbeiten, sondern auch der Ausbildung von Bibliothekaren und Rechercheexperten dient, scheint wesentlicher Faktor für das Informationsverhalten während der systematischen Literaturrecherche, sie gibt insbesondere dem Informationsspezialisten als Akteur den normativen Rahmen seines Handelns vor. Besondere Beachtung sollte ihr daher als Normenwerk und institutionelle Verpflichtung zukommen. Wie methodologische Vorgaben in der Praxis tatsächlich umgesetzt werden und wie sie das Informationsverhalten beeinflussen, ist bisher jedoch unklar und soll empirisch untersucht werden.

Verfügbare Information und Informationsressourcen bilden eine sowohl systematischen Verzerrungen in der Literatur- und Forschungslandschaft als auch der zunehmenden Menge an Erstpublikationen geschuldete, sehr komplexe und inhomogene Arbeitsumgebung. Der Informationsexperte als Auftragnehmer der systematischen Literaturrecherche hat sich dabei diesen Rahmenbedingungen zu stellen. Wie er dies tut, ist bisher lediglich aus methodologischer Sicht bekannt. Ziel der vorliegenden Arbeit sind weitere Einblicke auch in diesen Aspekt.

Bestehende Modelle zum Information Behaviour können teilweise Auskunft darüber geben, welche weiteren kontextuellen Rahmenparameter *Information Behaviour* in der professionellen, systematischen Literaturrecherche wesentlich beeinflussen könnten.

Das Modell von Byström und Järvelin (siehe Abschnitt 2.2.4) versteht Informationsverhalten aus aufgabeninitiiert und stellt dieses in den Kontext der Komplexität zu bewältigenden Aufgaben. In der systematischen Literaturrecherche sind Aufgabenstrukturen zwar in verschiedenen Detailgraden benannt (vgl. Abbildung 3-3 oder Tabelle 4-4), diese Darstellungen sind jedoch wie vieles stark normativ geprägt und geben keinerlei Hinweis auf den tatsächlichen

Informationsbedarf und die daraus resultierenden *information seeking tasks*. Grundlegende Aufgabe der empirischen Untersuchung ist daher die Identifikation und Klassifikation von Aufgabenstrukturen und Informationsverhalten.

Ausgehend vom Modell nach Leckie et al. (siehe Abschnitt 2.2.5) wird das Informationsverhalten von Domänenexperten wesentlich von demographischen Faktoren des Nutzers beeinflusst. In der systematischen Übersichtsarbeit wurde zwar die Rolle des Informationsexperten wiederholt definiert und charakterisiert, über den Akteur hinter dieser Rolle ist jedoch bisher wenig bekannt. Die systematische Literatursuche im Rahmen von Reviews ist – wie einleitend bereits beschrieben – in der Praxis eine interdisziplinäre Aufgabe von Domänenexperten als auch Bibliothekaren und Informationsspezialisten. Informationswissenschaftlich findet das Informationsverhalten von Informationsexperten gegenüber dem von Domänenexperten bisher kaum Beachtung. Die vorliegende Arbeit versucht, diese Lücke zu schließen.

Dem gegenüber stehen informationswissenschaftliche Modelle, die Prozesse der Informationsgewinnung beschreiben. Es ist anzunehmen, dass sich diese Prozesse auch in der Durchführung systematischer Übersichtsarbeiten finden: das Modell von Ellis (siehe Abschnitt 2.2.1) wurde empirisch in verschiedenen domänenspezifischen Kontexten bestätigt, das ISP-Modell von Marchionini auf die mit dem SR teilweise vergleichbare Patentrecherche übertragen (siehe Abschnitt 2.2.3).

Modelle wie diese können allerdings lediglich als Anhaltspunkt für die Beschreibung des Informationsverhaltens von Rechercheexperten dienen, letztendlich sind sie im Kontext der evidenzbasierten Medizin zu entwickeln.

1. Die empirische Untersuchung fokussiert sich auf die Entwicklung komplexer Suchprofile für strukturierte Suchprozesse wie in Abschnitt 3.4 beschrieben, einem Teilaspekt der systematischen Übersichtsarbeit, in dem der Informationsexperte besonders gefordert ist. Insbesondere die von den Nutzern angewandten Strategien zur Ermittlung von Suchtermen wurden in der Methodologie bisher nur unzureichend beleuchtet. Ein verbesserter Einblick in die dort angewandten Methoden und Muster könnte wesentlich zur Fehlervermeidung in der Vorbereitung von Reviews beitragen, aber auch Antworten auf die Frage geben, welche Strategien im Sinne der Zeitersparnis zu einer optimalen Kosten-Nutzen-Rechnung beitragen.
2. Gleichzeitig stellt sich die Frage, ob und welche Empfehlungen für die Gestaltung von Fachdatenbanken ausgesprochen werden können, die den gegebenen methodologischen Anforderungen strukturierten Suchprozesse nicht nur Genüge tun, sondern den

Nutzern von Informationssystemen bei der Suche nach Ergebnissen evidenzbasierter Medizin unter den methodologisch strengen Rahmenbedingungen entgegenkommen.

4 Empirische Untersuchung zum Informationsverhalten

Das folgende Kapitel stellt die empirischen Untersuchungen zum Informationsverhalten während der Durchführung systematischer Übersichtsarbeiten vor, welche im Rahmen der vorliegenden Dissertation zwischen März 2014 und November 2015 in eigenständiger Arbeit durchgeführt wurden. Die Ergebnisse dieser Untersuchungen dienen der Beantwortung der einleitend formulierten und in Abschnitt 3.7 konkretisierten Forschungsfragen und münden in einer Ergänzung des *Information Behaviour and Retrieval in Context*-Modells der systematischen Übersichtsarbeit.

Das Kapitel beschreibt zunächst allgemein Forschungsdesign, Datenerhebung und Datenanalyse und Teilergebnisse einzelner empirischer Untersuchungen, endet mit der Präsentation der Ergebnisauswertung und mündet in der Ergänzung des oben genannten Modells zum Verständnis des kontextuellen Informationsverhaltens in der professionell durchgeführten systematischen Übersichtsarbeit.

4.1 Forschungs- und Untersuchungsdesign

Im Vorfeld der Untersuchung musste die Frage nach einem geeigneten Forschungs- und Untersuchungsdesign geklärt werden. Forschungsdesign meint damit allgemein die „methodische Vorgehensweise einer Studie“ (Döring und Bortz 2016, S. 182), sie entscheidet unter anderem über die Aussagekraft wissenschaftlicher Ergebnisse, ist aber auch abhängig von forschungsökonomischen Gesichtspunkten, vgl. (Döring und Bortz 2016, S. 182).

Aufgrund der beschriebenen offenen Forschungsfragen mit dem Ziel einer Theoriebildung lag ein qualitatives Forschungsdesign nahe, das nicht nur eine interpretative Auswertung der Daten sondern auch flexible Datenerhebungsmethoden ermöglicht (Döring und Bortz 2016, S. 184).

Das Informationsverhalten von Informationsexperten ist bisher in der Forschung wenig beleuchtet. Die evidenzbasierte Medizin macht hier insoweit eine Ausnahme, dass wir anhand

der strengen Methodenlehre wenigstens erahnen können, wie Experten tatsächlich suchen. Mit Normen wie MECIR oder Handbüchern wie dem Cochrane Handbuch oder den Richtlinien des CRD stehen zwar ausreichend normative Daten zur Verfügung, subjektive Einschätzungen sind jedoch kaum zu finden. Daten, mit deren Hilfe die drei wesentlichen Fragen:

- Wie gehen die Experten in der Praxis tatsächlich vor?
- Wie gehen Experten mit Entscheidungssituationen um?
- Wie werden normative Regelwerke in der Praxis umgesetzt?

beantwortet werden konnten, mussten daher zunächst generiert werden.

Die methodische Erarbeitung dieser Daten für die vorliegende Arbeit gestaltete sich aus organisatorischen Gründen nicht unproblematisch und unterlag einigen forschungsökonomischen Randbedingungen:

Untersuchungsgruppe der vorliegenden Untersuchung mussten Experten in der Durchführung systematischer Übersichtsarbeiten sein. Obwohl die systematische Übersichtsarbeit überwiegend in einem interdisziplinären Team aus Domänen- und Rechercheexperten durchgeführt wird, liegen doch die Rechercheexperten als die für die Suche verantwortliche Nutzergruppe im Fokus des Interesses.

Verfügbare Experten sind weltweit verteilt und damit für den Forscher zunächst nicht leicht erreichbar. Vor allem im deutschsprachigen Raum gibt es abseits des in Freiburg ansässigen deutschen Cochrane Zentrum inzwischen nur noch wenige anerkannte Experten, die sich auf die Beschaffung medizinischer Fachinformationen spezialisiert haben: Auf den Websites des DIMDI werden im November 2015 lediglich 13 Anlaufstellen genannt, von denen die meisten im universitären Kontext angesiedelt sind (DIMDI: Deutsches Institut für Medizinische Dokumentation und Information 2016c). Von den 52 Cochrane Review Groups sind 2015 zwei in Deutschland angesiedelt, weitere Review Groups gibt es im deutschsprachigen Raum nicht. Ebenso führt das Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen (IQWiG) systematische Übersichtsarbeiten durch. Diese verhältnismäßig kleine potenzielle Zielgruppe ist für die Forschung theoretisch problemlos persönlich erreichbar. International kann der Kontakt zu Rechercheexperten in persönlichem Kontakt, auf Kongressen oder über Mailinglisten hergestellt werden.

Aufgrund der eingeschränkten Zielgruppe musste im Vorfeld auch ausgelotet werden, wie kooperativ sich die Rechercheexperten bundesweit und international zeigen und wie groß

ihre Bereitschaft zur Unterstützung dieses Forschungsprojekts sein würde.

Für die Datenerhebung quantitativer Studien stehen grundsätzlich verschiedene Methoden zur Verfügung, darunter Beobachtungsstudien, Interviews und Fragebögen. Abseits methodologischer Qualitätsaspekte dieser Datenerhebungsmethoden, die im Folgenden noch eingehend betrachtet werden, stellte sich zu Beginn die Frage, welche Methoden von den Nutzern akzeptiert werden.

Für einen Rechercheexperten in Regensburg waren Beobachtungsstudien grundsätzlich denkbar. Die Überlegung, umfangreiche Beobachtungsstudien mit verschiedenen Informationsexperten durchzuführen, erwies sich in der Praxis jedoch als nicht realisierbar: die Durchführung systematischer Übersichtsarbeiten ist keine Arbeit von Stunden, sondern Tagen und Wochen. In der Praxis hätten also Beobachtungsstudien über mehrere Sessions hinweg erfolgen müssen und wären mit viel Organisationsaufwand für die Protokollführerin als auch die Probanden verbunden gewesen. In persönlichen Gesprächen mit Experten auf dem Cochrane Workshop in Freiburg im März 2015 musste die Idee daher aus Machbarkeitsgründen verworfen werden. Ebenso verworfen werden musste die Idee einer Diary-Studie oder sonstiger elektronisch gestützter Beobachtungsverfahren, wie sich auf oben genanntem Workshop zeigte: Gerne wären Experten bereit, sich für Gespräche zur Verfügung zu stellen, die zusätzliche Arbeitsbelastung durch das Führen von Tagebüchern wurde jedoch ebenso als Ablehnungskriterium genannt wie Verschwiegenheitserklärungen von Experten im Rahmen ihrer Recherchetätigkeit. Eine Kooperationsanfrage über die oben erwähnte Liste beim DIMDI führte bei kommerziell tätigen Informationsbrokern zu deutlicher Ablehnung (hierzu liegt der Verfasserin der Arbeit eine Mail vom 19.09.2015 vor), sodass diese komplett aus der potenziellen Zielgruppe entfielen. Die Datenerhebungsmethoden mussten sich demnach auf den akademischen Bereich beschränken und gleichzeitig den Workload der Profis möglichst geringhalten.

Unter den gegebenen Bedingungen wurde nach dem Prinzip der methodischen Triangulation (vgl. (Flick 2007)) ein hybrides Forschungsdesign qualitativer Methoden gewählt, das aus einer Vorstudie in Form einer Beobachtungsstudie sowie einer Hauptstudie in Form eines Fragebogens bestand.

Während quantitative Forschungsdesigns typischerweise anhand der klassischen Evaluationskriterien Objektivität, Reliabilität und Validität bewertet werden können, stellt sich die Situation in der qualitativen Forschung ungleich komplexer dar, vgl. (Steinke 2007) und bedarf einiger Erläuterungen.

Die folgenden Abschnitte gliedern sich daher jeweils in Konzeption und Durchführung der Studie, theoretischer Hintergrund sowie Datenauswertung und die daraus ermittelten Ergebnisse.

4.2 Vorstudie

Zwischen März 2014 und Dezember 2014 wurden an der Universität Regensburg informelle Interviews und Anwendungsbeobachtungen in Kooperation mit der Universitätsbibliothek Regensburg durchgeführt. Diese dienten dem Ziel, einen Überblick über die Tätigkeiten des Rechercheexperten sowie die Interaktion mit verfügbaren Informationsressourcen zu gewinnen.

4.2.1 Studiendesign

Beobachtungsprotokolle ermöglichen per se semi-strukturierte oder vollstrukturierte Beobachtungen. Dies unterstützte das vorgegebene Ziel, jeden einzelnen Schritt des Gesamtprozesses schnell dokumentieren zu können, um die Reaktivität der Beobachtungssituation möglichst gering zu halten, vgl. (Döring und Bortz 2016, S. 330).

Ausgehend davon erfolgte die Dokumentation der Beobachtungsprotokolle mittels eines papierbasierten Formulars (siehe Anhang 1 und Anhang 2), mit folgenden Erfassungsparametern, die sich sehr eng an der Methodenlehre systematischer Übersichtsarbeiten und dem daraus ableitbaren Tätigkeitsprofil orientierten und im Folgenden erläutert werden:

Tabelle 4-1 Erfassungsparameter der Beobachtungsstudie

Parameter	Beschreibung
Modus	Modus beobachteter Aktivitäten
Kontext	Kontext, Einsatz von Informationsressourcen
Move	Suchtaktiken und Taktiken zur Ideenfindung
DB Feature	Spezifikation genutzter Benutzeroberflächen in Fachdatenbanken
Ertrag	Ergebnis bzw. Ertrag der Aktivität
Bewertung	Beurteilung der Aktivität durch den Nutzer
Dauer/Uhrzeit	Dauer bzw. Uhrzeit dokumentierter Aktivitäten

Bemerkung	Freitextbemerkung
------------------	-------------------

Modus sollte die Dokumentation eines Aufgabenschwerpunkts möglich machen und orientierte sich an den in der Methodenlehre benannten Aufgaben in der Entwicklung von Suchprofilen.

Bereits in vorbereitenden Gesprächen war aufgefallen, dass die von Bates definierten Taktiken zur Suche in Online-Datenbanken (Bates 1979b) und die von Bates definierten *idea tactics* (Bates 1979a) (siehe hierzu Abschnitt 2.4.2.) in der Interaktion mit Booleschen Datenbanken sowie in der Strategiefindung zur systematischen Literaturrecherche ihre Gültigkeit zu haben scheinen. Daraus entwickelte sich die Idee, diese für eine schnelle Dokumentation *typischer Aktivitäten* (*moves*) im Kontext von Suchprofil und in der Interaktion mit Booleschen Fachdatenbanken zu nutzen.

Mit Vorgabe dieser Taktiken war gleichzeitig der **Kontext** an Informationsressourcen und Entitäten vorgegeben (siehe Tabelle 2-6 bis Tabelle 2-9), der um Gegebenheiten systematischer Literaturrecherchen in der Medizin angepasst wurde und damit die Dokumentation der Nutzung dieser Informationsressourcen ermöglichte.

Ergänzt wurde das Formular um ein Feld **Ertrag** sowie Angaben zur subjektiven **Erfolgsbewertung**.

Der Konzeption des formularbasierten Dokumentationsprozesses war gleichzeitig die Wahl der späteren Analysemethoden geschuldet. Die aus den Beobachtungsdaten extrahierbaren und analysierbaren Daten waren zum einen quantitativer, durch die Freitextbemerkungen jedoch vor allem qualitativer Natur. Aufgrund der Forschungsfragen bot sich eine Fokussierung auf qualitative Forschungsmethoden, in dem Fall der qualitativen Inhaltsanalyse, an.

Da Strukturierungsdimensionen und Kategorien durch den Formularaufbau des Beobachtungsbogens teilweise vorgegeben waren, eignete sich eine qualitative Inhaltsanalyse mit deduktiver Kategorienbildung.

Hierbei werden „schon vorher festgelegte, theoretisch begründete Auswertungsaspekte an das Material“ (Mayring 2000) herangetragen und methodisch abgesichert zu Textstellen zugeordnet (siehe ebenfalls oben angegebene Zitation). Die Aufgabe der qualitativen deduktiven Inhaltsanalyse besteht also im Wesentlichen darin, genau zu definieren, unter welchen Regeln den Textstellen des Materials so genannten Kategorien zugeordnet werden können, siehe Abbildung 4-1.

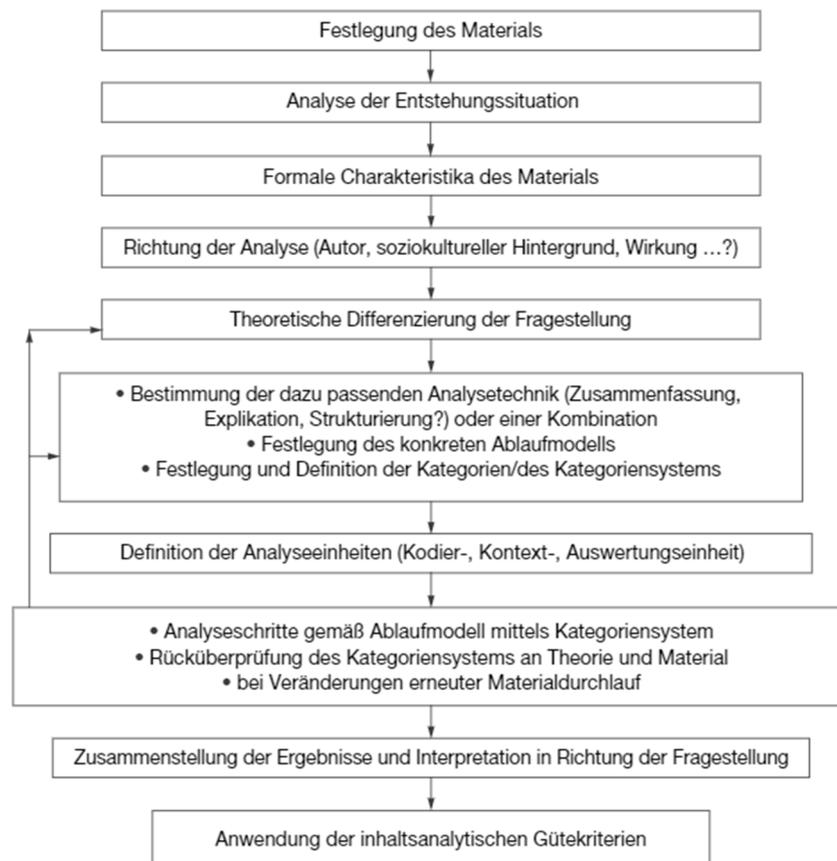


Abbildung 4-1 Qualitative Inhaltsanalyse nach Mayring, aus (Mayring 2015, S. 62)

Die qualitative Untersuchung zielte also darauf ab, mittels der aus Freitextanmerkungen und Annotationen gewonnenen zusätzlichen Daten die sogenannten Strukturierungsdimensionen und Kategorien anhand eines Kodier-Leitfadens iterativ zu überarbeiten. Im iterativen Kodierungsprozess konnten entsprechend folgenden Kodier-Leitfadens insbesondere Aktivitäten und Kontexte identifiziert werden.

4.2.2 Durchführung

Die Durchführung der Anwendungsbeobachtung während der Erstellung eines Suchprofils für eine systematische Übersichtsarbeit erfolgte im November und Dezember 2014 in 6 Einzelsessions, die zwischen 45 Minuten und 3,5 Stunden dauerten. Bei jeder Sitzung war neben dem Informationsspezialisten die Protokollführerin anwesend, bei drei Sitzungen zudem die Auftraggeberin der systematischen Literaturrecherche. Rechercheexperte und Auftraggeberin

hatten sich im Vorfeld der Beobachtungsstudie bereits ausgetauscht. Damit war die klinische Forschungsfrage eingehend fixiert. Zudem lagen eine Reihe relevanter Dokumente als Ground Truth vor, darunter auch eine systematische Übersichtsarbeit zu einem ähnlichen Thema, welche die Auftraggeberin zu Beginn selbst recherchiert hatte.

Pro Sitzung wurden zwischen 15 und 86 einzelne Protokollseiten angefertigt, die zur Vermeidung von Erinnerungsfehlern (vgl. (Döring und Bortz 2016, S. 331)) nach den Sitzungen umgehend eingescannt, auf Dokumentationsfehler hin überprüft und annotiert wurden. Die Annotation war der Tatsache geschuldet, dass Freitextbemerkungen handschriftlich erfolgt waren und teilweise zudem Rückschluss auf zeitlich bedingt nicht dokumentierte Erfassungsp Parameter gaben, gleichzeitig dokumentierten sie aber auch Korrekturen aufgrund offensichtlicher Dokumentationsfehler. Diese Vorgehensweise ist bereits Teil der Datenaufbereitung, siehe (Döring und Bortz 2016, S. 583)

4.2.3 Methodik der Datenaufbereitung und -analyse

Diese annotierten Dokumente wurden in einem iterativen Verfahren inhaltlich auf das Wesentliche reduziert, anschließend sprachlich vereinheitlicht, in eine Microsoft Excel-Liste übertragen und dabei in weitere Einzelschritte zerlegt, wenn

- sich Freitext Anmerkungen in mehrere, bereits bekannte Einzelschritte zerlegen ließen
- iterative Aufgaben dokumentiert worden waren.

Nähere Details zu den Methoden der Datenaufbereitung finden sich bei (Döring und Bortz 2016, S. 583).

Anhand dieser Datenaufbereitung war es möglich, jeden Teilschritt einem dokumentierten Schritt und einer Session zuzuweisen und gleichzeitig eine inhaltliche Analyse von Teilabläufen wie auch des Gesamtablaufs durchzuführen.

Für die Datenauswertung lagen zum Ende der Datenaufbereitung qualitative Daten vor, die sich sowohl aus den Vorgaben des Formulars („Was tut die Testperson?“) als auch der Annotationen zusammensetzten. Gleichzeitig bot die Liste sowohl eine Sicht auf die Quantität von

Einzelschritten („Wie häufig tut die Testperson etwas?“) als auch Prozesse („In welcher Reihenfolge werden Aufgaben erledigt?“).

Ziel der Analyse war es, aus den vorliegenden Daten ein strukturiertes, lesbares Rahmenmodell zu entwickeln, aus dem in Summe ablesbar war:

- was die Versuchsperson gerade tut
- welche Entitäten (Dokumente, Informationsressourcen) genutzt werden
- welches Ziel die Testperson mit einzelnen Teilschritten oder Manövern verfolgt.

Die Datenaufbereitung ermöglichte darüber hinaus auch eine Mustererkennung sowie eine quantitative Inhaltsanalyse mit dem Ziel

- Muster in der Abfolge einzelner Teilschritte zu identifizieren und
- Tätigkeitsschwerpunkte der einzelnen Beobachtungssessions zu ermitteln.

Zur weiteren Bearbeitung wurden die Daten in die auf qualitative Datenanalyse und *Mixed Methods* (einer Methode, für Forschungszwecke sowohl qualitativ als auch quantitative Daten auszuwerten, vgl. (Döring und Bortz 2016, S. 72)) spezialisierte Software MaxQDA Version 11 übertragen, Annotationen der eingescannten Dokumente sofern notwendig als ergänzende Einträge angelegt.

Hierüber erfolgte zunächst die Kodierung aller durch das Formular und die Datenaufbereitung gewonnenen Daten anhand der durch den Erfassungsbogen vorgegebenen Strukturierungsdimensionen Aktivität, Kontext und soweit wie möglich auch Ziele und Problemfelder (vgl. Tabelle 4-1), sodass letztendlich jedem dokumentierten Schritt der Beobachtungsstudie jeweils ein Code mindestens einer dieser Strukturierungsdimensionen zugeordnet werden konnte.

Der Formularaufbau erwies sich zur Dokumentation zwar als sehr hilfreich und garantierte die schnelle Dokumentierbarkeit, in der Praxis aber als unzureichend, weil unvollständig: Häufig waren Einzelschritte nicht identifizierbar, nicht eindeutig zuzuweisen oder Einzelschritte wurden in schneller Abfolge iterativ durchgeführt.

Viele der Einzelschritte konnten hinsichtlich der vordefinierten Erfassungsparameter, vor allem dem Modus sowie der Uhrzeit und Dauer, daher nur unvollständig dokumentiert werden, während Freitext-Bemerkungen während der Studie zunehmend an Bedeutung gewannen.

Daher enthielten viele Schritte Freitext-Anmerkungen, einzelne Sessions mit der Auftraggeberin des systematischen Reviews wurden rein als Freitext dokumentiert. Die Anmerkungen umfassten dabei sowohl Aussagen der Testperson als auch Memos der Protokollführerin. Wenn iterative Schritte bereits mehrfach ausgeführt worden waren, wurde die Dokumentation dieser Iterationen bei Wiederholung auf die wesentlichen, abweichenden Aspekte reduziert. Die über die Beobachtungsstudie ermittelten Prozesse waren also anhand der Daten nicht direkt sichtbar, sondern mussten durch eine eingehende Analyse identifiziert werden.

Zur Absicherung der Vollständigkeit und Validität der Kodierung wurden die Ergebnisse engmaschig mit den Originaldaten verglichen und mit dem Interviewpartner diskutiert, vgl. (Steinke 2007) zu Details dieser Vorgehensweise.

Im Folgeschritt wurden alte und neue Strukturierungsdimensionen iterativ verknüpft, um beobachtete Tätigkeiten im Interaktionskontext zu identifizieren und damit ein Grundgerüst anfallender Aufgaben zu generieren.

Über visuelle Tools der Software MaxQDA war es darüber hinaus möglich, Prozesse hinter diesen Aufgaben sichtbar zu machen und daraus typische Handlungsabfolgen zu identifizieren.

4.2.4 Ergebnisauswertung

Der folgende Abschnitt enthält und ergänzt Ergebnisse, die in einer Poster-Präsentation im Rahmen des 23. Cochrane Colloquium vom 03.-07.10.2015 in Wien vorgestellt wurden (Merz und Knüttel 2015).

Bei der Darstellung der Ergebnisauswertung wird wiederholt auf das Transkript der Beobachtungsstudie referenziert, einzelne Schritte sind über die Kodierung `SessionNummer.TransskriptNummer.FormularNummer` eindeutig zuzuordnen, Code 1.16.12 beispielsweise gehört zu Session 1, wurde dokumentiert in Zeile 16 und referenziert auf Formularblatt 12. Für die Sessions 3 und 5, die nicht über das Formularblatt dokumentiert wurden, liegen ausschließlich Transskriptnummern, jedoch keine Formularnummer vor.

Die Daten zur Beobachtungsstudie finden sich auf einer der Arbeit beiliegenden CD.

Zum Verständnis der Ergebnisse ist ein Überblick über den Ablauf der Beobachtungsstudie hilfreich, die in 6 Sessions mit inhaltlich sehr unterschiedlichem Schwerpunkt durchgeführt wurde:

Für Session 1 lagen dem Informationsexperten eine Reihe relevanter Referenzen, teils als Zitation, teils als Volltext vor. Deren Analyse sowie der Aufbau des Suchprofils durch die Sammlung geeigneter Suchterme prägten das Tätigkeitsprofil von Beobachtungseinheit 1.

In Session 2 wurden diese Tätigkeiten weitergeführt, der Schwerpunkt lag dabei allerdings auf dem Ausbau des bisherigen Suchprofils.

In Session 3 war neben dem Informationsexperten auch die Auftraggeberin anwesend, um finale Suchprozesse zu definieren, darunter beispielsweise Umfang und Art zu durchsuchender Informationsquellen und Datenbanken.

Während der Beobachtung in Session 4, bei der ebenfalls die Auftraggeberin anwesend war, wurde die Syntax des bisherigen Suchprofils auf die gewählten Datenbanken angepasst und als Probedurchlauf durchgeführt. Das Ergebnis von final ungefähr 466.000 Treffern machte es notwendig, sowohl die initiale Fragestellung als auch das Suchprofil zu überdenken. Zur besseren Planbarkeit weiterer Schritte war es Aufgabe der Auftraggeberin, eine aus der Treffermenge gewählte Prüfmengung von ca. 10.000 Referenzen manuell zu sichten und auf relevante Treffer hin zu analysieren.

Session 5 ist geprägt von Optimierungsbemühungen zum Erreichen einer „handhabbaren“ und validen Treffermenge.

Die abschließende Session 6 dokumentiert weitere Optimierungsbemühungen sowie die finale Suche der systematischen Übersichtsarbeit.

Die Entwicklung datenbankspezifischer Suchprofile kristallisierte sich wie zu erwarten als die dominanteste Tätigkeit und Aufgabe der gesamten Beobachtungsstudie heraus. Die Suche in Grauer Literatur als auch in Studienregistern wurde während Session 3 an die Auftraggeberin delegiert, sodass der Informationsexperte ausschließlich zur Durchführung systematischer Suchprozesse zuständig war. Weitere in der Methodenlehre genannten formalen und thematischen Suchprozesse wurden nicht beobachtet. Die folgenden Ergebnisse beschreiben daher ausschließlich strukturierte Suchprozesse (vgl. Abschnitt 3.4).

4.2.4.1 Interaktionskontext

Im Hinblick auf die im Folgenden geplante Hauptstudie als auch das daraus ableitbare Interaktionsmodell wurden aus der Beobachtungsstudie heraus verschiedene in der Praxis genutzte Interaktionskontexte klassifiziert.

Sie bilden den Interaktionsradius des Nutzers bei der Entwicklung strukturierter Suchen und legen die für den Nutzer verfügbaren Interaktionsformen fest und werden wie folgt definiert:

Tabelle 4-2 Interaktionskontexte in der Entwicklung und Durchführung strukturierter und systematischer Literatursuchen

Interaktionskontext	Beispiele
Informationsressourcen	Bibliographische Datenbanken Thesauri Hosts Suchmaschinen
Informationseinheiten	Zitation Bibliographisches Dokument Volltext Thesauruseinträge Treffermengen
Suchprofil und Queries	Kontrolliertes Vokabular Freitextterme Phrasen Artefakte
Kontakte	Kontakt zum Domänenexperten

Unter dem Interaktionskontext **Informationsressourcen** werden entsprechend der bisherigen Terminologie der Arbeit alle Informationsquellen zusammengefasst, die zum Auffinden von Informationen im Sinne von Dokumenten genutzt werden, darunter bibliographische Datenbanken, korrespondierende Thesauri, Hosts sowie Suchmaschinen.

Art und Umfang genutzter bibliographischer Fachdatenbanken müssen immer im Kontext der Verfügbarkeit gesehen werden und können daher nicht verallgemeinert werden. Für die Literaturrecherche an der Universität Regensburg beispielsweise stehen Domänenexperten eine

Vielzahl an Datenbanken zur Verfügung, darunter auch frei verfügbare wie Pubmed. Die Fachrecherche über den Host DIMDI ist in Regensburg ausschließlich über die Universitätsbibliothek lizenziert.

Aufgrund der Spezialisierung des Informationsspezialisten auf die Entwicklung von Suchprofilen für bibliographische Datenbanken wurden über die Beobachtungsstudie nur ein Teil verfügbarer und in der Fachliteratur benannter Informationsressourcen (siehe Abschnitt 3.2) genutzt, die direkt mit strukturierten Suchprozessen assoziiert sind.

Informationseinheiten sind im Folgenden definiert als jegliche Art der Informationen, die zur Ermittlung von Suchtermen genutzt werden können. Der Interaktionskontext Informationseinheit ist vielfältig und umfasst neben bibliographischen Angaben (Zitationen), bibliographischen Dokumenten auf Fachdatenbanken auch Volltexte und Thesauruseinträge als Spezifika datenbankangelegelter Thesauri sowie Treffermengen durchgeführter Datenbanksuchen.

Informationseinheiten sind demnach gekennzeichnet durch unterschiedlichste Strukturen und im Hinblick auf Suchterme auch unterschiedliche Erträge, siehe Tabelle 4-3.

Tabelle 4-3 Struktur und direkter Ertrag von Informationseinheiten

Informationseinheit	Struktur	Direkter Ertrag
Bibliographisches Dokument	Titel (Autoren) Journal Abstract Keywords Indexierung	Kontrolliertes Vokabular Freitextterme
Zitation	Titel Autor Journal	Freitextterme
Volltext	Titel Autoren Journal Abstract Textkorpus Bibliographie	Freitextterme

Thesauruseinträge	Beschreibung	Kontrolliertes Vokabular
	Hierarchie	Subheadings
	Subheadings	Freitextterme

Ein Teil der über die Beobachtungsstudie identifizierten Informationseinheiten bildet die *Ground Truth*, siehe hierzu Abschnitt 3.4.4. Die Kodierung „Ground Truth“ wurde während der Analyse des Datenmaterials angewandt, wenn sich Interaktionen direkt in der Ground Truth, also einer vorab verfügbaren oder noch zu entwickelnden Referenzmenge relevanter Dokumente, abspielten. Diese Referenzmenge bestand aus einer Liste an Zitationen, für die vereinzelt auch Volltexte in Form von PDF Dokumenten vorlagen, vgl. hierzu insb. Session 1.

Suchprofil und Query bilden den dritten Interaktionskontext. Das Suchprofil kann verstanden werden als strukturiertes Dokument (siehe 1.5.3a, 1.6.4) aus kontrolliertem Vokabular, Freitexttermen, Phrasen und Artefakten als datenbankspezifisch angepasste, meist trunkierte Suchterme oder komplexe Phrasen.

Kontrolliertes Vokabular liegt sowohl in Form von Einzeltermen als auch Mehrworttermen vor, ist jedoch datenbankspezifisch als solches erkennbar, siehe hierzu Abschnitt 3.4.2.2.

Freitextterme als Einzelwortterme wie auch Phrasen als Multiwort-Terme können mittels datenbankspezifischer Syntax trunkiert und mittels Proximity-Operatoren syntaktisch zu Artefakten verknüpft werden, anderswo auch in Summe „*atomic search arguments*“ benannt (Stock et al. 2013, S. 242).

Kontakte bilden den letzten Interaktionskontext. Sie finden in erster Linie mit dem Domänenexperten, meist dem Auftraggeber statt.

4.2.4.2 Kontextspezifische Interaktion

Die im methodologischen Kontext der systematischen Übersichtsarbeit definierte Aufgabe „systematische Entwicklung strukturierter Suchprofile“ (in Handbüchern benannt als „Entwicklung von Suchstrategien“, vgl. hierzu Abschnitt 3.4.1) lässt sich in der Methodenlehre (siehe Abschnitt 3.4) per definitionem als Abfolge folgender Teilaufgaben interpretieren und ist dort als iterativer Prozess verschiedener *working tasks* beschrieben:

- Aufbau einer suchbaren Fragestellung in Form von Building Blocks

- Aufbau eines umfassend formulierten, in Building Blocks strukturierten Suchprofils bestehend aus Deskriptoren, Freitexttermen und Phrasen
- Übersetzung von Suchprofilen in datenbankspezifische Queries
- Kontrolle der Vollständigkeit von Suchprofilen

Ziel der Analyse war es, ausgehend von den in der Methodenlehre bekannten *working tasks* Aufgabenstrukturen aus *searching-* und *seeking tasks* sowie daraus resultierende Tätigkeiten und Taktiken abzuleiten und in verschiedenen Interaktionskontexten zu beschreiben. Grundlage hierfür boten zum einen die im formularbasierten Protokoll bereits vorgegebenen datenbankspezifischen Moves (siehe hierzu Abschnitt 2.4.2), zum anderen aber auch Informationsverhalten, das im informationswissenschaftlichen Kontext bereits beschrieben wurde, siehe hierzu die Modelle in den Abschnitten 2.2.1-2.2.6. So referenziert die folgende Darstellung terminologisch wiederholt auf verschiedene informationswissenschaftlichen Modelle, darunter insbesondere die Modelle von Ellis und Marchionini.

Verarbeitung von Informationseinheiten

Die exzessive Sammlung geeigneter Deskriptoren und Freitextterme bestimmt vor allem die Anfangsphase in der Entwicklung von Suchprofilen, siehe Session 1 und Session 2. Sie ist charakterisiert durch die gezielte Überprüfung verfügbarer Informationseinheiten mit dem Ziel, dort für die Suchanfrage geeignete Suchterme zu identifizieren.

Extracting wird bei Ellis beschrieben als „systematische Überprüfung von Informationsquellen mit dem Ziel, dort relevantes Material zu extrahieren“, siehe Abschnitt 2.2.1. Extracting ist mit dem Ziel der Identifikation neuer Suchterme bei der Entwicklung von Suchprofilen die vorherrschende Aktivität, sie nutzt sowohl Dokumente der Ground Truth als auch einzelne Thesauruseinträge, bibliographische Dokumente, Volltexte, Zitationen sowie Treffermengen aus Ad Hoc-Suchprozessen.

Die Beobachtungsstudie zeigte, dass die Sammlung von Suchtermen in Dokumenten der Ground Truth im Wesentlichen zwei Ziele verfolgt:

1. Den Aufbau eines umfassenden, möglichst vollständigen Suchprofils aus kontrolliertem Vokabular und Freitexttermen (siehe 1.5.3a ff.)

2. Die Sicherung des Retrievals vorab bekannter relevanter Referenzen im Sinne einer Ground Truth durch Aufnahme mindestens eines Suchterms pro Building Block (siehe 1.17.13 ff.)

Das Retrieval eines Dokuments ist durch die Aufnahme mindestens eines Suchterms pro Building Block gesichert (dies kann theoretisch bereits durch die Identifikation von Suchtermen in sehr spezifischer Dokumententitel erfolgen). Der Aufbau eines umfassenden, möglichst vollständigen Suchprofils aus kontrolliertem Vokabular und Freitexttermen erfordert jedoch gegebenenfalls die Analyse korrespondierender bibliographischer Dokumente in den zu durchsuchenden Fachdatenbanken wie auch zugehörige Volltexte.

Bei der Analyse bibliographischer Dokumente werden Strukturen einzelner Informationseinheiten genutzt, sie erfolgte entsprechend systematisch. Bei der Analyse von *Volltexten* aus der Ground Truth werden gezielt Titel, Abstract, Textkorpus als auch Zitationen der Bibliographie analysiert, in Thesauri sowohl die Entry Terms als auch Beschreibungen (*scope notes*) zu einzelnen Thesauruseinträgen. Häufig ist Extracting daher eine Tätigkeit, die der Analyse verfügbarer, meist komplexer Informationseinheiten (**Analyzing**) oder der Bewertung des Retrievals einzelner Dokumente durch das aktuell verfügbare Suchprofil folgt.

Davon abgegrenzt werden muss die eher explorative Untersuchung von Informationseinheiten (**Exploring**), das zum einen mit Lernprozessen assoziiert ist, beispielsweise zum Verständnis der Verwendung spezifischer Fachtermini (vgl. 1.121.58), zum anderen mit dem Wunsch, weitere, potenziell interessante Suchterme aufzufinden, die bisher bei anderen Analyseprozessen nicht identifiziert werden konnten (vgl. 1.161.76 ff). Vor allem Session 1 und 2 sind von explorativen Untersuchungsprozessen geprägt.

Die Identifikation potenziell relevanter Suchterme mit dem Ziel des Ausbaus bestehender Suchprofile erfolgt häufig weniger strukturiert in einer Form von **Screening** und führt zur Extraktion potenziell relevanter Suchterme aufgrund von „als relevant erkennen“ (u.a. 1.19.14 ff), bei unspezifischeren Terminologien in Titel oder Abstract aber auch zur Notwendigkeit von Verifikationsprozessen (u.a. 1.38.20).

Creating Artefacts ist eine Standardaktivität während der Entwicklung von Suchprofilen, die häufig im Anschluss an Extracting oder Screeningprozessen folgt, das in erster Linie auf die

Identifikation relevanter Wortstämme (*Lexeme* = „Einheit des Wortschatzes, die die begriffliche Bedeutung trägt“ (Bibliographisches Institut GmbH Duden Verlag 2016d)) abzielt. Mit der Modifikation von Lexemen zu suchtauglichen, meist trunkierten, also bereits im Hinblick auf sensitive Suchanfragen ausgerichteten Suchtermen, werden datenbankspezifische Suchterme (Artefakte) geschaffen. Die über die Datenanalyse so kodierten Aktivitäten subsummieren Tätigkeiten, die auf möglichst sensitive Suchprofile hin ausgerichtet sind und umfassen die unter anderem bei Bates und Fidel beschriebenen Moves, siehe Abschnitt 2.4.2 und 2.4.3.

Synthesizing beschreibt die Eingliederung neu identifizierter Suchterme und Artefakte in bestehende Suchprofile, meist in direkter Kombination mit der Erstellung suchtauglicher Artefakte, u.a. 1.17.13 – 1.21.14. Da die Eingliederung in Suchprofile nicht datenbankspezifischer Natur ist, soll sie begrifflich von „formulate query“, der Formulierung einer Suchanfrage, aus dem ISP-Modell (siehe Abschnitt 2.2.3) abgegrenzt werden.

Die Extracting – Creating Artefacts – Synthesizing Kette ist eine der dominantesten Verhaltensmuster, die während der Beobachtungsstudie identifiziert werden konnte.

Identifikationsprozesse

Identifikationsprozesse dienen im Wesentlichen der Ermittlung weiterer Informationseinheiten mit dem Ziel, aus diesen weitere Suchterme zu extrahieren und damit die Sättigung des Suchprofils voranzutreiben.

Searching bezeichnet bei Ellis ein mögliches, finales Element der Literatursuche (siehe Abschnitt 2.2.1) und kann beispielsweise als Anwendung einer in vorhergehenden Prozessen entwickelten Query interpretiert werden, vgl. hierzu auch „formulate query“ und „execute query“ im Modell nach Marchionini (siehe Abschnitt 2.2.3). In der Entwicklung von Suchprofilen spielt Searching bei der systematischen und finalen Anwendung komplexer Suchprofile als datenbankspezifische Queries im Sinne von Ellis eine zentrale Rolle (vgl. hierzu Beobachtungsstudie 4.12.9 – 4.42.29 ff), ist jedoch auch wichtiges Hilfsmittel zur Evaluierung der Sensitivität und Spezifität von Suchtermen, Phrasen und ausgearbeiteten Queries (vgl. hierzu Beobachtungsstudie 6.18.5 ff).

Gleichzeitig dienen Ad Hoc-Suchen auch dem Ziel der Identifikation weiterer relevanter Dokumente ermittelter Treffermengen, aus deren Titeln oder bibliographischen Referenzen im

Folgenden neue Freitextterme, Synonyme und Deskriptoren gewonnen werden können. In diesem Fall münden Ad Hoc-Suchen in Analyse- oder Screening-Prozessen wie oben beschrieben.

Specifying ist eine Aufgabe, die vor allem im Kontext der Ground Truth zu beobachten war, wenn zu Zitationen mit „unspezifischem Titel“ die Volltexte beschafft oder zugehörige bibliographische Dokumente identifiziert werden mussten (u.a. 1.136.69, 1.171.78). Sie entspricht der klassischen „*known item search*“, die informationswissenschaftlich definiert ist als Suche nach einem Item (Dokument) mit bekanntem Titel oder Autor, vgl. (Lee et al. 2006). Known Item Searches wurden sowohl im Kontext von Suchmaschinen wie Google Scholar als auch bibliographischen Fachdatenbanken beobachtet, sind aber prinzipiell auf alle problemlos verfügbaren Informationsquellen ausdehnbar. Die Verwendung des Terminus Specifying erfolgt bewusst als Abgrenzung zum Terminus Searching, vgl. auch (Belkin et al. 1995).

Unter **Chaining** versteht Ellis (siehe Abschnitt 2.2.1) sowohl die Rückwärts- als auch die Vorwärtssuche. Sie ist im Rahmen systematischer Literatursuchen eine klassische Aktivität im Kontext bibliographischer Angaben der Ground Truth und mündet beispielsweise in der Aktivität Specifying zur Beschaffung assoziierter Volltexte.

Weitere Aktivitäten können als Bewertungsprozesse zusammengefasst werden.

Bewertungsprozesse

Der Aufbau des Suchprofils erfolgt zunächst als strukturierte Liste in Form einer umfangreichen Sammlung an Kandidatentermen, im Sinne potenziell für das vollständige Suchprofil relevanter Suchterme, die in regelmäßigen hinsichtlich ihres Ertrags als auch der Relevanz erzeugter Treffermengen bewertet werden, siehe u.a. 1.59.25.

Differentiating ist eine vergleichende Aktivität, meist mit dem Ziel der semantischen oder qualitativen und vergleichenden Bewertung von Informationseinheiten, Freitexttermen und Deskriptoren (2.11.9, 2.13.11), siehe hierzu auch Ellis in Abschnitt 2.2.1. Bei Ellis insbesondere als vergleichende Bewertungsmaßnahme von Informationseinheiten beschrieben, dient der Vergleich beim Aufbau strukturierter Suchprofile der qualitativen Bewertung von Suchtermen, beispielsweise um die Verwendung und Popularität von Fachterminologien in der Fachliteratur

(z.B. 1.121.58) und damit die Tauglichkeit von Suchtermen im Hinblick auf vollständige Suchprofile zu bewerten. Im Gegensatz dazu dienen quantitative Vergleiche der Bewertung des Retrievals einzelner Dokumente, was abgrenzend dazu als **Comparing** definiert wird.

Abzugrenzen ist Differentiating im Kontext der vorliegenden Untersuchung von **Verifying**, einer Aktivität, die ebenfalls bei Ellis (Abschnitt 2.2.1) beschrieben ist und die Beurteilung von Zuverlässigkeit und Richtigkeit von Information meint. Bei der Entwicklung von Suchprofilen dient die Verifikation weniger der Information einzelner Dokumente als vielmehr der Verifizierung in der Indexierung bibliographischer Referenzen oder der Bedeutung bisher unbekannter Fachterminologien.

Mit **Evaluating** stehen qualitativen Bewertungsprozessen quantitative Evaluierungsprozesse gegenüber. Sie spielen in der Entwicklung von Suchprofilen aufgrund der Anforderung möglichst sensitiver Suchprofile eine wesentliche Rolle und werden anhand von Testsuchen realisiert, um die Mächtigkeit von Treffermengen beurteilen zu können. Gleichzeitig dienen Evaluierungsmaßnahmen auch der Kontrolle modifizierter Artefakte oder der gesamten Query durch vergleichende Testsuchen (u.a. 2.26.20, 2.30.22). Während Evaluierungsmaßnahmen zu Beginn meist (der Bestätigung) einer Sensitivitätssteigerung dienen, werden sie gegen Ende des Gesamtprozesses auch für eine Präzisierung der Suchanfrage verwendet und münden direkt in einer Änderung von Suchtermen, beispielsweise durch eine Veränderung an PROXIMITY-Operatoren oder Entfernung einer Trunkierung (5.11ff sowie 6.15ff).

Die iterative Kontrolle von Suchtermen durch Validierungs- und Evaluierungsprozesse wird insbesondere in Session 1 der Beobachtungsstudie deutlich, in der alle bis dato in das Suchprofil aufgenommenen Terme als „Kandidatenterme“ bezeichnet werden (1.29.25).

Recording Prozesse dienen dem Monitoring und der Dokumentierbarkeit des gesamten Entwicklungsprozesses, sie betreffen vor allem Suchprofil und Queries. Unterstützt werden diese Prozesse häufig durch datenbankspezifische Funktionalitäten, siehe hierzu auch Abschnitt 3.4.2.3.

Tabelle 4-4 fasst die Ergebnisse zusammen:

Tabelle 4-4 Tätigkeiten während der Entwicklung von Suchprofilen im Kontext verschiedener Informationskontexte (siehe hierzu Tabelle 4-2)

Aktivität	Kontext	Beispiel
Identifikationsprozesse		
Searching	Informationsressourcen	Ad Hoc-Suche mit einzelnen Suchtermen und Phrasen als auch Suche mit ausgearbeiteten Queries
Specifying	Informationsressourcen	Known Item-Suche, u.a. existierende systematische Übersichtsarbeiten, bibliographische Dokumente, Volltexte
Chaining	Informationseinheiten	Bibliographische Suche (Rückwärtssuche) Zitationssuche (Vorwärtssuche)
Verarbeitung von Informationseinheiten		
Analyzing	Informationseinheiten	Strukturelle Analyse im Kontext von Zitationen, Volltexten, Literaturangaben, Treffermengen, bibliographischen Dokumenten, Thesauruseinträgen
Extracting	Informationseinheiten	Gezielte Sammlung von Kandidatentermen aus Zitationen, Volltexten, Literaturangaben, Treffermengen, bibliographischen Dokumenten, Thesauruseinträgen
Exploring	Informationseinheiten	Explorative Analyse im Kontext von Zitationen, Volltexten, Literaturangaben, Treffermengen, bibliographischen Dokumenten, Thesauruseinträgen
Entwicklung und Modifikation von Suchprofil und Query		
Creating Artefacts	Suchprofil/Query	Entwicklung von Suchtermen und Suchphrasen durch Trunkierung, Maskierung und Einsatz von PROXIMITY-Operatoren, datenbankspezifisch
Synthetizing	Suchprofil/Query	Hinzunahme neuer Terme zu einer vorbestehenden Query, Neustrukturierung von Suchprofil und Query
Kontrolle des Suchprozesses		

Comparing	Informationseinheiten Suchprofil/Query	Vergleichende qualitative Bewertung im Kontext von Thesauri, bibliographischen Dokumenten, Volltexten, Zitationen sowie Suchprofil sowie vergleichende, quantitative Bewertung von Suchtermen und Suchprofilen durch Vergleich von Treffermengen
Validating	Informationseinheiten	Qualitative Bewertung, Thesauri, bibliographische Dokumente, Volltexte, Zitationen, Treffermengen aus Ad Hoc Suchprozessen
Comparing	Informationseinheiten	Vergleichende qualitative Bewertung, beispielsweise zur Verwendung kontrollierten Vokabulars
Evaluating	Informationsressourcen	Sensitivitätsbewertung von Deskriptoren, Freitexttermen, Synonymen, Phrasen und Artefakten als auch Teilen einer Query
Monitoring des Gesamtprozesses		
Recording	Informationseinheiten Suchprofil/Query	Volltexte, Suchprofile, Queries

4.2.4.3 Aktivitätsmuster

Wie bereits einleitend beschrieben, waren die während der Beobachtungsstudie dokumentierten Sessions inhaltlich unterschiedlich geprägt. Während Session 1 und 2 überwog das Ziel, ein Suchprofil aus einer möglichst umfangreichen Menge an Deskriptoren und Freitexttermen zu entwickeln.

Eine wichtige Rolle für die Entwicklung von Suchprofilen spielt – soweit vorhanden – die Ground Truth als Menge bekannter, relevanter Dokumente. Ihre Bearbeitung erfolgt iterativ verschlauft und wurde in der Beobachtungsstudie immer wieder als gezielte Strategie zur Sicherung des Retrievals der Ground Truth sowie der Expansion des Suchprofils identifiziert. Sie dominierte das Informationsverhalten der Beobachtungsstudie, sodass sich die folgenden Ergebnisse auf diesen Aspekt beschränken.

Retrieval-gesteuerte Ground Truth-Loops

Der Ground Truth-Loop auf dem Building Block Level beschreibt das strukturierte Vorgehen einer systematischen Abarbeitung von Dokumenten und Zitationen der Ground Truth mit dem primären Ziel, das Retrieval eines Dokuments mittels der aktuell ausgearbeiteten Query zu sichern (vgl. u.a. 1.40.21ff).

Der Ground Truth-Loop auf dem Level Informationseinheiten beschreibt die wiederum systematische Abarbeitung der gesamten Ground Truth.

Beides kann als iterative Interaktionskette in verschiedenen wie in Abschnitt 4.2.4.1 beschriebenen Interaktionskontexten beschrieben werden und ist in Abbildung 4-2 graphisch dargestellt. Die in der folgenden Abbildung verwendeten Abkürzungen beziehen sich dabei auf die in Tabelle 4-2 definierten Interaktionskontexte.

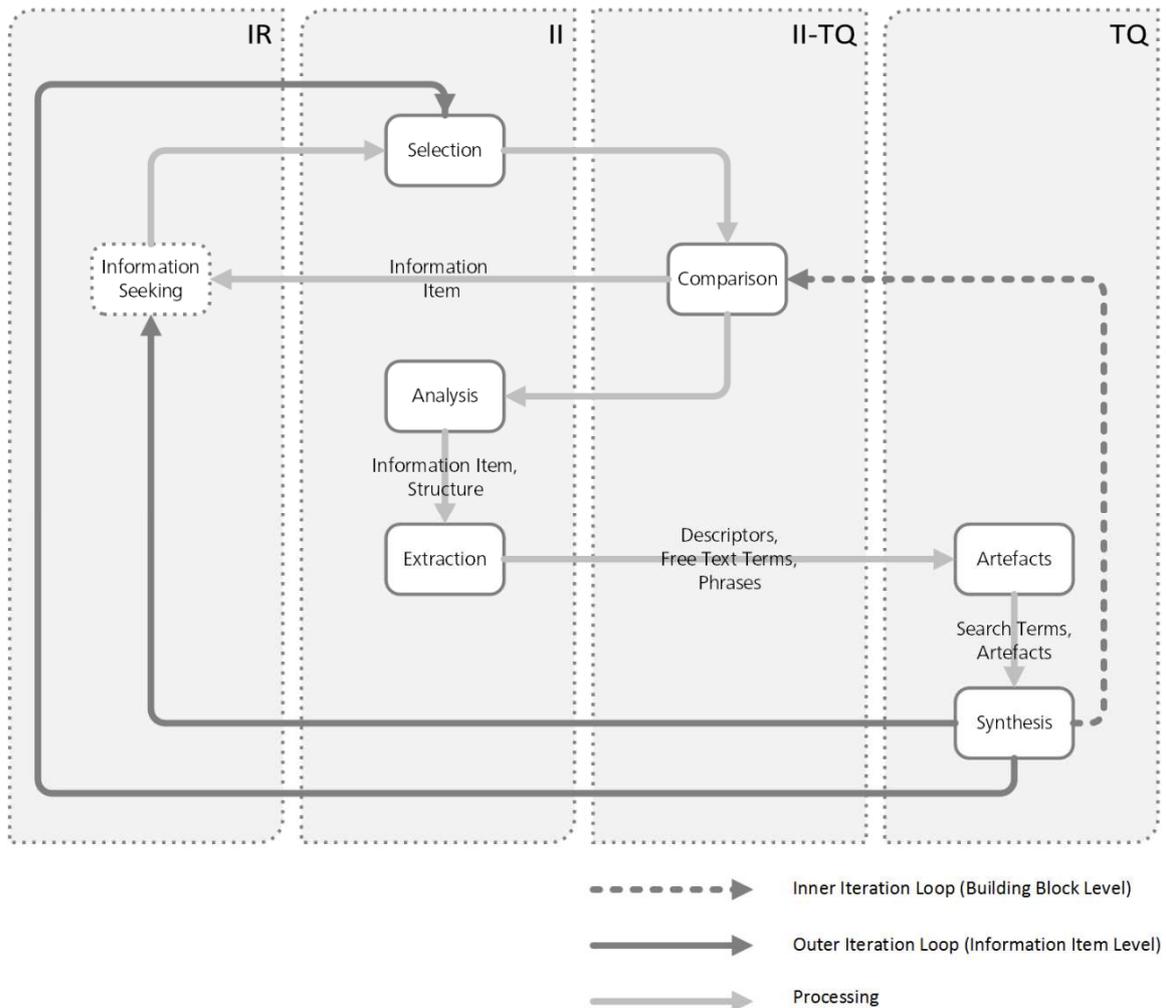


Abbildung 4-2 Retrieval-gesteuerte Ground Truth Loops

Die Grundidee der in Abbildung 4-2 als *Inner Iteration Loop (Building Block Level)* benannten Schleife ist folgendermaßen und wurde an anderer Stelle (Abschnitt 3.4.4.2) bereits eingehend beschrieben: Das Retrieval eines Dokuments gilt als vollständig gesichert, wenn jeder Building Block durch das Vorkommen bereits gesammelter Suchterme im Titel, Abstract oder dem kontrollierten Vokabular eines Dokuments bzw. seiner datenbankspezifischen, bibliographischen Referenz repräsentiert ist.

Verfügt das Suchprofil bereits über Suchterme, so erfolgt zunächst der Abgleich von allen zum Dokument verfügbaren Informationen und dem Suchprofil im Hinblick auf vollständige Sicherung des Retrievals, in dem zu jedem Building Block Konzept für jeden als relevant beurteilten Suchterm aus Titel oder Abstract überprüft wird, ob dieser bereits im Suchprofil enthalten ist

oder nicht. Liegen nicht ausreichend Informationen zu einem Dokument vor, um das Retrieval garantieren zu können, beispielweise lediglich Informationen in Form einer Zitation mit einem bezüglich der Fragestellung eher unspezifischen Titel, müssen bibliographische Dokumente oder Volltexte zunächst beschafft werden. Grundsätzlich werden alle über die Analyse als potenziell relevant erkannten Suchterme als Kandidatenterme in das Suchprofil aufgenommen.

Mit der Abarbeitung eines Dokuments der Ground Truth folgen weitere Dokumente, bis die Ground Truth vollständig gesichert ist.

Sensitivitäts-gesteuerte Ground Truth-Loops

Ist das Retrieval durch das aktuelle Suchprofil gesichert, so können Dokumente der Ground Truth auf weitere optionale Suchterme und Synonyme zu beliebigen Konzepten der Fragestellung hin untersucht werden. Die Untersuchung erfolgt hier explorativer (siehe u.a. 1.167.76, 1.190.84). Da sie teilweise auch dem Verständnis neu identifizierter, potenziell relevanter, weil sensitivitätssteigernder Suchterme dient, münden Analyseprozesse häufig in Vergleichsprozessen mit anderen Dokumenten oder der Validierung von Suchtermen in anderen Quellen.

Die Analyse bibliographischer Angaben zu Volltexten ist nicht nur eine Methode zur Identifikation neuer Freitextterme, sondern gleichzeitig eine Quelle für weitere potenziell relevante Dokumente, beispielsweise in Form der Nutzung datenbankspezifischer Funktionalitäten wie „*similar articles*“.

Neu identifizierte und sofort als relevant eingestufte Dokumente werden in die Ground Truth aufgenommen und analog aller bisher analysierten Dokumente verarbeitet.

Die graphische Darstellung findet sich in Abbildung 4-3:

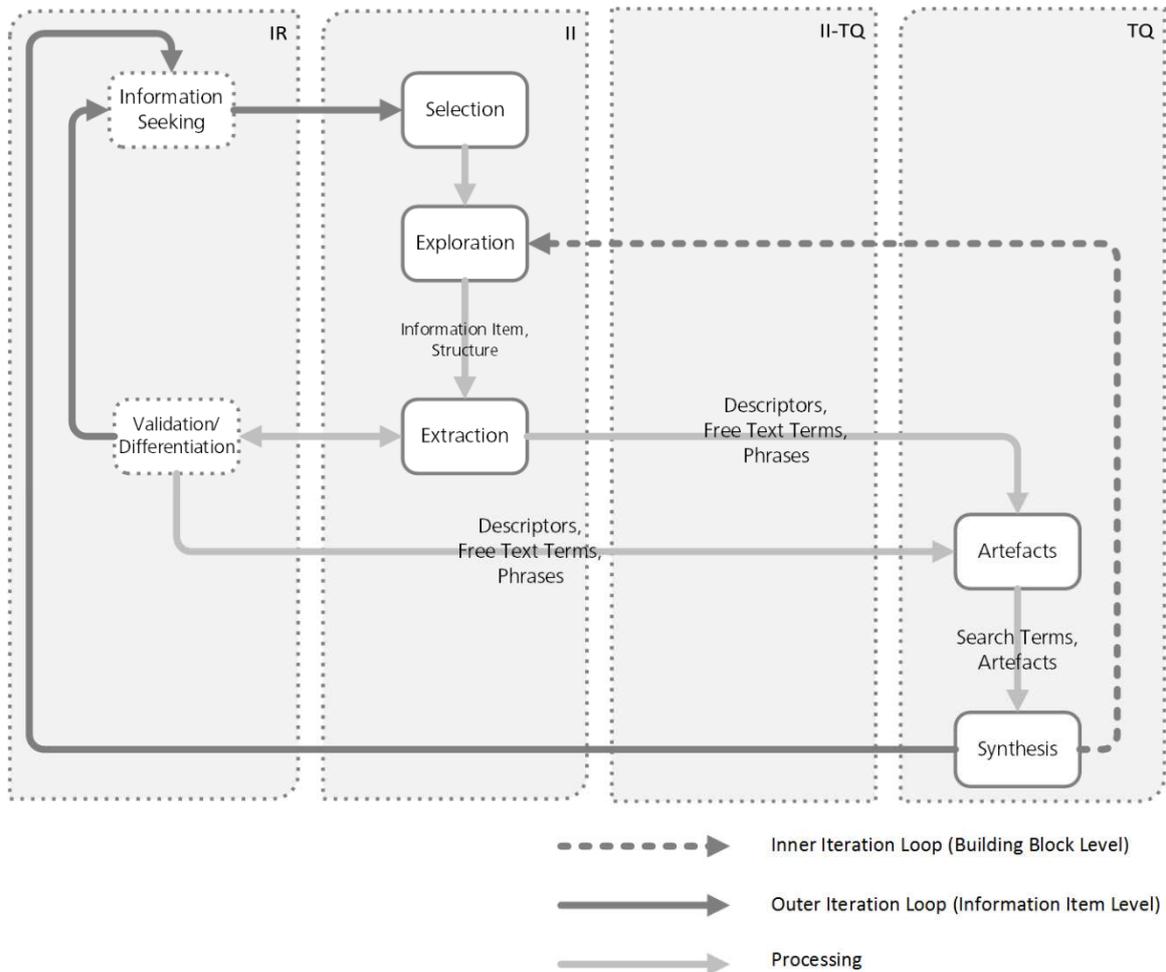


Abbildung 4-3 Sensitivitäts-gesteuerte Ground Truth Loops

In beiden Ground Truth Loop-Modellen zeigt sich die methodologische Grundidee in der Entwicklung von Suchprofilen: *“Developing a search strategy is an iterative process in which the terms that are used are modified, based on what has already been retrieved.”* (Lefebvre et al. 2011), vgl. hierzu auch Abschnitt 3.4.3.1. Ground Truth-Loops sind vorwiegend in Session 1 und Session 2 als iterative Prozesse zu beobachten und dominieren hier das Informationsverhalten.

Creating Artefacts konnte, wie bereits in Abschnitt 4.2.4.2 beschrieben, aufgrund des formal vorgegebenen Protokollierungsprozesses in weitere Teilaufgaben zerlegt werden, wobei nahezu alle bei Bates und Fidel beschriebenen Modifikationen von Suchtermen und Suchprofil angewandt wurden.

Mit dem Ziel der Entwicklung umfassender, erschöpfender Suchprofile fokussieren sich die Taktiken auf solche, die einzelne Suchterme und damit das gesamte Suchprofil erweitern, siehe Tabelle 4-5 sowie Tabelle 4-6.

Es sei dabei auf Folgendes hingewiesen: In dieser aus der Veröffentlichung von Bates entnommenen tabellarischen Übersicht wird der Begriff Query verwendet, der in der vorliegenden Arbeit exklusiv auf datenbankspezifische Suchabfragen hin abzielt. Die vorliegende Übersicht verwendet die „Bates-Terminologien“ in allgemeiner Weise und meint damit sowohl Suchterme im Suchprofil (als strukturierte Auflistung von Suchtermen) als auch datenbankspezifische Suchterme in Suchanfragen.

Tabelle 4-5 Ergebnisse der Beobachtungsstudie: Taktiken zur Suchformulierung nach (Bates 1979b)

Taktik	Beschreibung
S1. SPECIFY	Suche mit spezifischen Termen
S2. EXHAUST	Ausbau der Treffermenge durch Hinzunahme weiterer möglicher Suchterme zu einer bestehenden Query
S4. PARALLEL	Breitere Suche durch Hinzunahme von Synonymen zu einer Query

Tabelle 4-6 Ergebnisse der Beobachtungsstudie: Taktiken zur Modifikation von Termen nach (Bates 1979b)

Taktik	Beschreibung
T1. SUPER	Verwendung eines Terms, der hierarchisch breiter angelegt ist
T3. RELATE	Verwendung eines verwandten Terms in gleicher Hierarchieebene
T5. TRACE	Nutzung bereits verfügbarer Information zum Auffinden weiterer Terme
T10. RESPELL	Verwendung anderer Schreibweisen für einzelne Terme
T11. RESPACE	Verwendung von Abstandsoperatoren

Ground Truth Loops: Rahmenbedingungen

Verarbeitungsprozesse der Ground Truth werden in der Beobachtungsstudie durch verschiedene Rahmenbedingungen beeinflusst und damit gegebenenfalls auch unterbrochen oder ergänzt:



Abbildung 4-4 Bedingende Faktoren der Ground Truth-Loops

Die Nutzung systematischer Übersichtsarbeiten als qualitativ hochwertiger **Studientyp** mit hohem Ertragspotenzial mündet häufig in einer *umfassenden* Untersuchung zugehöriger Volltextdokumente. Als Ziele lassen sich sowohl die Identifikation weiterer Synonyme festmachen, als auch die Analyse von Literaturverzeichnisse zur Identifikation weiterer relevanter Dokumente. Diese dienen wiederum als Quelle für Suchterme und vergrößern unter Umständen gleichzeitig die Ground Truth (siehe beispielhaft Beobachtungsstudie 1.9.7, 1.11.9, 1.161.76 ff).

Sensitivität und Noise von Suchprofilen bestimmen sowohl die Anfangsphase der Entwicklung umfassender, erschöpfender Suchprofile als auch deren Optimierung. Während in der Anfangsphase ein hoher Ertrag vorrangig ist und zu häufigen sensitivitätsgetriebenen

Ground Truth-Loops führt, bestimmt in der Optimierungsphase die „Handhabbarkeit von Treffermengen“ die Vorgehensweise (siehe Sessions 5 und 6).

Das Ziel eines vollständigen **Retrievals** von Ground Truth Dokumenten war in der Beobachtungsstudie als zentraler Motor des Gesamtvorgehens erkennbar und führt zu exzessiven Retrieval-gesteuerten Ground Truth-Loops.

Das Durchlaufen verschiedener Ground Truth-Schleifen ist abhängig von der **Verfügbarkeit** von Informationseinheiten sowie der Struktur der Ground Truth. Obwohl letztendlich alle in der Ground Truth verfügbaren relevanten Dokumente hinsichtlich ihres Retrievals geprüft wurden (siehe 1.129.66), erfolgte die Ermittlung von Suchtermen nach Aufwand: Zu Beginn des Prozesses wurden hierfür zunächst lediglich die Titel relevanter Dokumente im Hinblick auf Suchterme untersucht, bei Unklarheiten und zur finalen Vervollständigung des Suchprofils auch verfügbare Volltexte (siehe u.a. 1.10.8 ff, 1.38.20 ff). Die Beschaffung bisher nicht vorliegender Volltexte oder die Suche korrespondierender bibliographischer Dokumente erfolgte in erster Linie bei Zitationen mit thematisch wenig spezifischen Titeln.

Die Analyse relevanter Informationseinheiten orientiert sich gezielt an der **Struktur** verfügbarer Information und fokussierte sich neben kontrolliertem Vokabular vorwiegend auf den Titel oder den Abstract bibliographischer Angaben oder Volltexte. Durchbrochen wird der Prozess lediglich bei Informationseinheiten der Ground Truth, deren Relevanz aufgrund verfügbarer Daten, beispielsweise aufgrund „unspezifischer“ Terminologien in Titel oder Abstract, nicht eindeutig beurteilt werden kann.

Die **Spezifität und Diversität in der Terminologie** als ein von der initialen Fragestellung abhängiger Faktor (siehe hierzu auch Abschnitt 3.3.1) ist in der Methodenlehre als kritischer Faktor insbesondere systematischer Übersichtsarbeiten zu interventionellen Themen bekannt und wurde auch in der Beobachtungsstudie deutlich. Unklarheiten führen zu häufigen Validierungs- und Vergleichsprozessen, um sowohl die Spezifität von Suchtermen zu beurteilen als auch die durch Suchterme verursachten Noise zu bewerten. In der Beobachtungsstudie führten sie des Weiteren häufig zur Planung von Austauschgesprächen mit der Domänenexpertin und zu einem Gefühl der Unsicherheit hinsichtlich der Verwendung einzelner Fachterminologien in der Praxis und damit auch deren Notwendigkeit für das zu entwickelnde Suchprofil (siehe beispielhaft 1.60.26, 1.60.27, 2.15.13, 5.10).

Spannungsfelder

Das gesamte Informationsverhalten der Beobachtungsstudie war geprägt durch das Spannungsfeld von Domänenwissen und intellektueller Anforderung an den Retrieval-Prozess. Verfügbares Domänenwissen wurde vor allem von der Auftraggeberin eingebracht, der Informationsexperte war sich seiner Lücken in der Verwendung von Fachtermini, als auch in der Beurteilung der aktuellen Forschungslandschaft durchaus bewusst und erkannte in allen Sessions, die ohne die Domänenexpertin durchgeführt wurde, die Notwendigkeit, sich mit ihr gelegentlich noch einmal abzusprechen.

Das Spannungsfeld zwischen Domänenwissen und Rechercheexpertise wurde vor allem bei der Strategiefindung zum Management intellektuell nicht-handhabbarer Treffermengen (initial 4.48.21) in den Sessions, bei denen die Auftraggeberin wieder anwesend war, deutlich: Der Proband orientierte sich strikt an der Vorgabe der Methodenlehre, die eine umfassende und erschöpfende Suche vorschreibt. Daher wurden Suchprofil als auch die Wahl der zu durchsuchenden bibliographischen Fachdatenbanken im Absprache mit der Auftraggeberin entsprechend breit angelegt und führten zu einer nicht-handhabbaren Menge an Treffern. Letzteres resultierte auf Wunsch der Auftraggeberin in der Notwendigkeit, das Suchprofil entsprechend zu reduzieren. Dieser gezielt angestoßene Prozess mit dem Ziel einer deutlichen Reduzierung der Treffermenge wird in den Beobachtungsprotokollen der Sessions 5 und 6 strukturiert beschrieben. Er führte zu intensiven Diskussionen über Methodenlehre und wurde letztendlich durch die Auftraggeberin gesteuert.

4.2.5 Zusammenfassung

Mit den Ergebnissen der Vorstudie lag ein subjektives, teils methodologisch vorgegebenes Rahmenmodell zum Informationsverhalten bei der Erstellung von Suchprofilen zur Durchführung systematischer Übersichtsarbeiten vor, welches durch die folgende Hauptstudie untermauert und ergänzt werden sollte. Insbesondere galt es mit der Hauptstudie, kontextuelle Rahmenbedingungen in Entscheidungssituationen kritisch zu hinterfragen.

4.3 Konzeption der Hauptstudie

Für die Hauptstudie wurde, wie bereits in Abschnitt 4.1 beschrieben, ein Fragebogen aus Freitextfragen entwickelt. Da bundesweit schätzungsweise nicht mehr als 10 Personen als Interviewpartner zur Verfügung gestanden hätten, sollte dieser international verteilt werden und die Durchführung realer Interviews mit ausgewählten Interviewpartnern ersetzen.

Methodisch wären klassische Interviews zu bevorzugen gewesen: Interviews ermöglichen es, wie Fragebögen auch, subjektives Erleben zu erfragen, das in Beobachtungsstudien nicht erfasst werden kann. Gegenüber Fragebögen ist die Methode des Interviews im Allgemeinen mit geringerer Hemmschwelle verbunden und kann in einem Rahmen stattfinden, in dem Personen direkt und persönlich angesprochen werden. Gegenüber Fragebögen hat das Interview den Vorteil, in kürzerer Zeit wesentlich mehr Informationen zu liefern und auch komplexere Sachverhalte ansprechen zu können (Döring und Bortz 2016, S. 356).

So sind in einem Fragebogen umfangreiche und komplexe Antworten tendenziell weniger häufig zu erwarten, gleichzeitig sollte man sich gegenüber Interviews im Fragebogen auf wenige und sehr prägnante Befragungsinhalte beschränken, vgl. (Döring und Bortz 2016, S. 398).

Interviews waren aufgrund der über alle Kontinente verteilten Experten im Live-Setting, als auch aufgrund einiger kritischer Stimmen hinsichtlich der Nutzung elektronischer Medien wie Skype digital nicht realisierbar, siehe hierzu auch Abschnitt 4.1. Damit war die Fragebogenmethode die einzig durchführbare und gleichzeitig aber effizientere Forschungsgrundlage (Döring und Bortz 2016), ebenda.

4.3.1 Entwicklung des Fragebogens

In der Beobachtungsstudie zeigte sich ein Trend, der auch in der Methodenlehre und Fachliteratur deutlich sichtbar ist: Der Fokus der Tätigkeiten von Informationsexperten liegt auf der Arbeit mit bibliographischen Fachdatenbanken, siehe u.a. Abschnitt 3.6.3.

Dies untermauerte die Entscheidung, den Fragebogen konzeptuell gezielt auf Informationsverhalten während der Entwicklung datenbankspezifischer Suchprofile für Boolesche Datenbanken hin auszurichten.

Inhaltlich zielte der Fragebogen auf das individuelle Informationsverhalten von Experten ab, legte aber den Fokus auf den Umgang mit bekannten, kritischen Entscheidungssituationen, die bereits in der Literatur beschrieben sind. Hierzu gehören zum einen die Ergebnisse der Studien von Raya Fidel (Fidel 1991a, 1991b, 1991c), die sehr dezidiert Auskunft über Lösungsstrategien für die von Fidel identifizierten und auch in der systematischen Übersichtsarbeit relevanten Problemkategorien liefern (siehe hierzu auch Abschnitt 2.4.3):

- Vom Nutzer als zu klein beurteilte Treffermengen, welche zu Maßnahmen führen, um diese zu vergrößern
- Von Nutzer als zu groß beurteilte Treffermengen, welche zu Maßnahmen führen, um diese zu verkleinern
- Vom Nutzer als „*off-target*“ beurteilte Treffermengen, die zu Maßnahmen führen, mit denen ein ausgewogenes Verhältnis zwischen Sensitivität und Spezifität wiederhergestellt werden soll

Diese waren durch die Beobachtungsstudie als auch durch Einzelgespräche als kritische Aspekte weitestgehend bestätigt und benannt worden. Gleichzeitig lag mit den Ergebnissen der Beobachtungsstudie ein erstes Prozessmodell vor, das durch die Fragebogenaktion bewertet werden sollte.

Für die Konzeption des Fragebogens galt es, die Formulierung suggestiver Fragen zu vermeiden, um normativ geprägtes Informationsverhalten vom deskriptiv beschriebenen Informationsverhalten abgrenzen zu können. Aus diesem Grund wurde ein teilstrukturierter Fragebogen entworfen. Inhaltliche Fragen wurden offen in Form von Freitextfragen konzipiert, zusätzliche demographische Informationen dienen der Analyse des Gesamtkollektivs der Befragten und wurden geschlossen formuliert.

Vorherrschende Fragen waren:

Welche Quellen werden zur Identifikation von Suchtermen herangezogen und wie werden diese Quellen genutzt?

Kann die Vorgehensweise bei der Entwicklung von Suchprofilen allgemeingültig dargestellt werden oder gibt es erkennbare Muster in der Vorgehensweise?

Welche Entscheidungen müssen während der Entwicklung von Suchprofilen getroffen werden?

Diese Fragen wurden ergänzt mit weiteren Items, die sich vorwiegend auf kritische Entscheidungssituationen bezogen:

Welche Kriterien sind für die Aufnahme von Suchtermen in Suchprofile und in finale Queries ausschlaggebend?

Wie erfolgt der Umgang mit eher „unspezifischen“ (d.h. allgemein gehaltenen) Synonymen?

Wie wird großen Treffermengen entgegengewirkt?

Welche Rolle spielen Testsuchen?

Um die deskriptiven Ergebnisse später besser im Kontext bestehender Standards und Normen, wie beispielsweise dem Cochrane Handbook, interpretieren zu können, war die Frage nach deren Rolle in der Praxis von wesentlicher Bedeutung: Mitglieder von Cochrane Review Groups müssen sich bei der Durchführung systematischer Übersichtsarbeiten aus Gründen der Qualitätssicherung beispielsweise verpflichtend an ihre Standards halten. Normen und Handbücher stehen Experten jedoch auch anderswo für Schulungen und Methodenlehre zur Verfügung. Folgende Fragen sollten daher beantwortet werden:

Welche Standards und Handbücher zur Durchführung systematischer Übersichtsarbeiten werden tatsächlich genutzt?

Gibt es Nutzer, die in ihrer Arbeit auf „*mixed methods*“ zurückgreifen, sich also verschiedener Methodenlehren bedienen?

Bereits während der Vorbereitung der vorliegenden Arbeit zeigten Literaturrecherchen und persönliche Gespräche mit Experten, wie unterschiedlich einzelne Fachtermini interdisziplinär verwendet werden. Es sollte daher auch Aufgabe sein, fachspezifische bibliothekswissenschaftliche und informationswissenschaftliche Terminologien zu identifizieren und auf eine gemeinsame Basis zu bringen. Da der Fragebogen für die Interviewpartner möglichst ohne Rückfrage verständlich sein musste, um die Rücklaufquote nicht zu gefährden, war wichtig, bei der Formulierung der Fragen die eher bibliothekswissenschaftlich angesiedelte Terminologie zu verwenden. Aus diesem Grund wurde die Entwicklung von Suchprofilen durchgängig als „Entwicklung von Suchstrategien“ bezeichnet, auch wenn dies aus informationswissenschaftlicher Sicht kritisch zu beurteilen ist (siehe Abschnitte 2.4.1 und 3.4.1).

Viele der online verfügbaren Umfragewerkzeuge lassen nur mit erhöhten Lizenzkosten eine Zwischenspeicherung zu, um eine spätere Weiterbearbeitung zu ermöglichen. Dies wäre aufgrund des Konzepts der Freitextfragen natürlich wünschenswert gewesen. Leider ließ das Budget, das für die vorliegende Arbeit zur Verfügung stand, dies nicht zu.

Ohne diese Zwischenspeicherung wäre aufgrund dieses Konzepts mit einem Akzeptanzproblem im Online-Verfahren zu rechnen gewesen. Daher wurde der Fragebogen zur leichteren und flexiblen Bearbeitung und Möglichkeit der Überarbeitung als Word-Dokument versendet und mit Antworten überarbeitet wieder zurückgefordert.

Im Verlauf der Untersuchung wurde ein initialer deutschsprachiger Fragebogen zunächst an 5 ausgewählten, deutschsprachigen Experten getestet. Diese 5 persönlich kontaktierten Experten lieferten die Fragebögen zwischen Ende Mai und Ende Juli 2015 per Mail zurück. Die Ergebnisse der deutschsprachigen Fragebögen wurden anschließend für eine englischsprachige Überarbeitung genutzt.

Nach der ersten Analyse des Vorab-Fragebogens wurden einige Fragen (4-12) für den finalen Fragebogen aufgrund unklarer Antwortmuster gestrichen. Die in der Literatur weit verbreiteten Fachtermini in den Fragen 5-11 wurden teilweise offenbar nicht richtig interpretiert (Mail Teilnehmerin 5), zusätzliche wurden auf die genannten Fragen teils wenig differenzierte Antworten gegeben. Um den Fragebogen nicht zu überfrachten, wurden diese schließlich gestrichen. Beide Versionen des Fragebogens finden sich codiert in deutscher und englischer Sprache im Anhang 3.

Die englischsprachige Version des Fragebogens wurde im Oktober 2015 an die Mailing-Liste der Cochrane Information Retrieval and Methods Group (siehe <https://lists.cochrane.org/mailman/listinfo/irmg>) mit zum 04.12.2015 über 320 Mitgliedern sowie zwei weitere ausgewählte Experten versandt, die sich in persönlichen Gesprächen auf dem Cochrane Colloquium vom 03.-07.10.2015 in Wien zu einer Teilnahme bereit erklärt hatten. In diesem Turnus liefen 11 beantwortete Fragebögen zurück, davon zwei aus dem deutschsprachigen Raum. Insgesamt lagen damit 15 Antworten vor. Mit 10 Rückläufern aus der Mailing-Liste entspricht dies einer Rücklaufquote von ca. 3%. Über die Rücklaufquote echter Rechercheexperten können keine Angaben gemacht werden, da nicht bekannt ist, wie viele Abonnenten der Mailing-Liste tatsächlich Literaturrecherchen durchführen und damit zu den direkten Adressaten der Anfrage zu zählen sind.

Alle Teilnehmer wurden im Vorfeld darauf hingewiesen, dass ihre Angaben vertraulich verwendet und in der Arbeit anonymisiert werden würden.

Im direkten Kontakt per E-Mail erläuterten einige der Teilnehmer ergänzend zu ihrem rückgesendeten Antwortbogen aus persönlicher Sicht ihre Arbeit oder standen bei Bedarf für weitere Rückfragen zur Verfügung. Auch diese Daten wurden, ggf. in Form von Memos und Freitextanmerkungen in die Analyse einbezogen.

Die folgende Tabelle gibt einen Überblick über die Interviewteilnehmer, darunter auch Berufserfahrung (BE), Geschlecht und Sprache:

Tabelle 4-7 Übersicht der Codings der Interviewpartner

Code	Sprache	BE (Jahre)	Geschlecht	Anmerkung
1	DE	4	w	zusätzliche E-Mail Konversation liegt vor
2	DE	27	w	zusätzliche E-Mail Anmerkungen liegen vor
3	DE	23	w	
4	DE	15	w	
5	DE	10	w	
6	EN	6	m	zusätzliche E-Mail Konversation liegt vor
7	EN	9	w	
8	EN	15	w	
9	EN	12	w	
10	EN	10	w	
11	EN	20	w	
12	EN	8	w	
13	DE	8	w	normativ geprägte Aussagen
14	EN	8	w	
15	DE	11	m	zusätzliche E-Mail Anmerkungen liegen vor

16	EN	10	w	
----	----	----	---	--

Eine Teilnehmerin des deutschsprachigen Fragebogens gab an, keine systematischen Übersichtsarbeiten durchzuführen. Aufgrund ihrer langjährigen Berufserfahrung wurden ihre Angaben dennoch in die Analyse mit einbezogen. Zwei Teilnehmer führen als klinische Bibliothekare vorwiegend Rapid Reviews durch. Beide haben Erfahrung in der Durchführung systematischer Übersichtsarbeiten, wiesen sowohl im Vorfeld als auch im weiteren E-Mail-Kontakt und bei der Beantwortung einzelner Fragen auf Unterschiede und Gemeinsamkeiten zwischen systematischen Übersichtsarbeiten und Rapid Reviews hin. So konnten auch sie in die Analyse einbezogen werden. Die Aussagen einer Teilnehmerin waren stark normativ geprägt, hier konnten leider nicht alle Antworten in die Analyse einbezogen werden. 7 Befragte waren oder sind aktive Mitglied einer Cochrane Review Group, 7 Befragte haben eine Aus- oder Weiterbildung im Bibliothekswesen, drei Befragte nennen als Ausbildung Informationswissenschaft oder eine vergleichbare Ausbildung, vier Befragte haben zusätzlich oder exklusiv ein naturwissenschaftliches Studium absolviert, eine Befragte hat eine Pflegeausbildung.

Tabelle 4-8 stellt die Ergebnisse der im Fragebogen abgefragten demographischen Daten zusammen.

Tabelle 4-8 Demographische Daten Fragebogenstudie

Parameter	Mittelwert	Median
Berufserfahrung (Jahre)	12,25	10
SRs pro Jahr	34,1	10
Regelmäßig genutzte bibl. Datenbanken	5,1	5
Regelmäßig genutzte Studienregister	4,4	2,5

Methodologische Standards werden von allen Beteiligten als Grundlage für die Durchführung systematischer Übersichtsarbeiten genannt, die Details listet Tabelle 4-9 auf. Dass hierbei vor allem Cochrane genannt wird, mag auf die Verteilung der Teilnehmer zurückzuführen sein, könnte aber auch daran liegen, dass das Cochrane Handbuch der wohl unter Informationsexperten populärste Standard ist.

Tabelle 4-9 Nutzung methodologischer Standards

Parameter	Nennungen
Cochrane Handbook	11
Cochrane in Kombination mit MECIR	3
CRD	3
Campbell	1
Sonstige, eigene	7

Ein Teilnehmer (Code 6) bat explizit darum, bei näheren Angaben mit Namen genannt zu werden. Neben den von ihm gemachten Angaben im Fragebogen lag ein subjektiv gehaltenes Internetdokument vor, das sehr dezidiert seine individuelle Vorgehensweise bei der Erstellung von Suchprofilen erläutert und daher als zusätzliche, nicht anonymisierte Quelle hinzugezogen wurde (Bramer und Jonge 2015). Dieses Dokument war die einzige über Literaturrecherchen ermittelbare themenrelevante Quelle mit subjektiven Beschreibungen zur Entwicklung strukturierter Suchanfragen und ergänzt damit die Daten der Fragebogenstudie.

4.3.2 Forschungsstrategie Grounded Theory

Gegenüber der Beobachtungsstudie als Vorstudie waren die Fragebögen breiter und allgemeiner angelegt, sie dienten neben der Bestätigung bereits bekannter Phänomene in erster Linie der Theoriegewinnung, wobei einzelne Ergebnisse der Beobachtungsstudie selbstverständlich als bereits vorliegende Information in die Analyse einfließen sollten und auch mussten.

Methodologisch eignet sich für diese Form der Fragestellung die von Glaser und Strauss 1967 vorgestellte Grounded Theory (Glaser und Strauss 1967), deren Idee es ist, anhand qualitativer Analysen soziale Phänomene sowohl zu erklären als auch zu beschreiben. Die Grounded Theory wird als „Forschungsstrategie“ bezeichnet, auf deren Basis eine – wie der Name bereits sagt – gegenstandsverankerte (*grounded*) und erklärende Theorie abgeleitet werden kann, vgl. (Döring und Bortz 2016, S. 545). Die Daten zur qualitativen Analyse werden dabei weitestgehend frei gelassen, es ist demnach gewollt, alles verfügbare Datenmaterial zu verwenden und in die Analyse mit einzubeziehen. Corbin und Strauss weisen darauf hin, dass Quellen

wie Dokumente, Videoaufzeichnungen, Zeitungen, Briefe und Bücher ebenso kodiert werden können wie die Aufzeichnungen zu Beobachtungsstudien oder Interviews (Corbin und Strauss 1990).

Die vorherrschende Analysetechnik der Grounded Theory ist die Methode des permanenten Vergleichs in einem iterativen und parallelen Prozess des Datensammelns, Kodierens und Analysierens verfügbarer Daten, vgl. (Döring und Bortz 2016, S. 603).

Die Kodierung der Daten erfolgt zunächst über offene Codes oder Merkmale, um damit Unterschiede, Gemeinsamkeiten und Handlungsmuster im Datenmaterial herauszuarbeiten. Auf diese Weise entstehen zunächst häufig lange Codelisten, die iterativ gebündelt und abstrahiert werden können. Während des darauffolgenden axialen Kodierens werden in einem „Kodier-Paradigma“ aus Bedingungen, Kontext, Strategien und Konsequenzen Relationen zwischen einzelnen Kategorien herausgearbeitet. Sind die so ermittelten Kategorien bekannt, kann die Kodierung auch selektiver und zielgerichteter erfolgen (Corbin und Strauss 1990). Begleitet wird der gesamte Kodier-Prozess durch das Schreiben so genannter Memos, schriftlicher Notizen, in denen Zwischenergebnisse und neue Ideen festgehalten werden, vgl. (Döring und Bortz 2016, S. 546).

Ein für das Verständnis wichtiger Unterschied zwischen der Forschungsmethode qualitative Inhaltsanalyse und Grounded Theory kann an der Entwicklung des Kategoriensystems und dem Kategorienbegriff an sich festgemacht werden. Während, wie bereits beschrieben, die qualitative Inhaltsanalyse darauf angelegt ist, im Fall der induktiven Methode bereits bekannte Merkmale an das Datenmaterial heranzutragen, oder im Fall der deduktiven Methode durch das Datenmaterial abzusichern, bleiben die Kategorien in der Grounded Theory bis zum Ende des Forschungsprozesses für die weitere Überarbeitung offen, vgl. (Muckel 2007).

Der Einsatz der Grounded Theory in der Information Behaviour-Forschung ist nicht neu, vgl. beispielsweise die Studien von (Ellis 1993; Lee 2010; Vassilakaki und Johnson 2015). Für die vorliegende Arbeit eignet sie sich zur Analyse des Datenmaterials der Hauptstudie insbesondere, weil die aus der Vorstudie nicht theoriefrei gewonnenen Daten mit zusätzlichem Material valide ergänzt werden können.

4.3.3 Analyseprozess

Auch für die Analyse des Fragebogens wurde die Software MaxQDA Version 11 verwendet, die neben Kodier-Werkzeugen auch die Annotation des Datenmaterials durch Memos erlaubt.

Die Antworten der einzelnen Fragebögen wurden vorab für jede einzelne Frage tabellarisch erfasst, nach Frage und Interviewpartner codiert und dann pro Frage in die Software MaxQDA Version 11 übertragen. In der Analysesoftware lagen demnach für jede Frage einzelne Dokumente pro Interviewpartner vor, deren Benennung der Codierung der Fragen und Teilnehmer folgt, beispielsweise „GENERAL 3“ für die Antwort der Teilnehmerin 3 auf die mit GENERAL codierte Frage 2 des Fragebogens. Die Übertragung passierte dabei nicht ad hoc, sondern iterativ. So war es während der Einarbeitungsphase möglich, sich zunächst auf eine Frage des Fragebogens zu konzentrieren. Schrittweise wurden aus der tabellarischen Liste weitere MaxQDA Dokumente zu weiteren Fragen ergänzt und neu kodiert.

Zur besseren Handhabbarkeit während der Analyse wurden alle in MaxQDA importierten Textfragmente sprachlich abstrahiert und vereinheitlicht, englischsprachige Antworten wurden hierzu übersetzt. Während der Analyse lagen also vereinfachte Texte vor, die einfach kodiert werden konnten, parallel dazu bot die Excel-Tabelle bei eventuellen Unklarheiten und zur Zitation eine übersichtliche und leicht zu durchsuchende Quelle für die zugehörigen Originaltextstellen.

Ein Ziel der Analyse war es, aus der klar definierten Arbeitsaufgabe „Entwicklung eines Suchprofils“ mit den beispielsweise im Cochrane Handbuch dezidiert definierten Teilaufgaben „Strukturierung einer suchbaren Fragestellung“, „Identifikation von Deskriptoren“ und „Identifikation von Freitexttermen“ zunächst Teilaufgaben zu identifizieren, zu charakterisieren und im Kontext der methodologischen Anforderungen der systematischen Literaturrecherche zu interpretieren. In einem nächsten Schritt sollten diese Informationen genutzt werden, um die ebenso zu identifizierenden Lösungsstrategien kontextuell zu beschreiben und zu klassifizieren.

In der Einarbeitungsphase der Analyse lag der Fokus zunächst auf der Verwendung sehr offensichtlicher Codes wie beispielsweise Deskriptor, Suchterm, Datenbank oder bibliographi-

sches Dokument, um analog zur Analyse der Beobachtungsstudie zu ermitteln, welche verschiedenen Entitäten bei der Entwicklung von Suchprofilen eine Rolle spielen. Dies führte zu einer Vielzahl einzelner Codes, die teilweise sehr gut klassifiziert werden konnten, aber zunächst nur rudimentär zum Verständnis des Gesamtprozesses beitrugen. Es war zu diesem Zeitpunkt also qualitativ, vor allem aber quantitativ gut erfassbar, welche Objekte existieren und wie diese in Zusammenhang stehen oder wie sie genutzt werden. Eine Aussage über deren Relevanz für den Gesamtprozess und den Einfluss äußerer Rahmenbedingungen war jedoch kaum zu treffen.

Vor allem sollten aus den Daten Aktivitäten und Prozesse sowie Entscheidungswege der Beobachtungsstudie belegt, widerlegt und neu identifiziert werden. Einzelne Antworten wie auch gesamte Interviews ließen zwar deutlich Prozesse erkennen, mit semi-automatisierten Kodierungsverfahren waren jedoch die zugrundeliegenden Kontextparameter nicht sauber abzubilden. Hierzu war es nötig, die vorliegenden Interviews wiederum als individuelles Ganzes zu betrachten. In einem iterativen Verfahren wurden daher in einem nächsten Schritt konsequent für alle thematisch klar voneinander abgrenzbaren Elemente und Textfragmente der Interviews separate Memos erstellt, miteinander verglichen und mit Originalzitate verknüpft.

Schritt für Schritt konnten so neue Aspekte der eigentlichen Fragestellung identifiziert werden. Die Analyse und Beschreibung des gesamten Prozessablaufs wurden daher von zwei neuen Fragestellungen überlagert, die als Kernkategorien im Sinne der Grounded Theory mehr und mehr in den Fokus der Untersuchungen gerieten und gleichzeitig zum Schlüssel für die Frage nach den zu identifizierenden Suchaufgaben wurden:

- Wie gewinnen die Experten Vertrauen in die Qualität der von ihnen entwickelten Suche?
- Was macht die Qualität eines Suchterms oder einer Query aus?

Ein von Beginn an bekannter Aspekt, der während der vorliegenden Arbeit zudem geklärt werden sollte, war die interdisziplinär unterschiedliche Verwendung von Fachtermini. Diese Problematik findet sich auch in den Interviews wieder und wurde durch das Vorliegen deutsch- und englischsprachiger Interviews zur besonderen Herausforderung und Chance. Wann immer Analyseergebnisse es möglich und erforderlich machten, einzelne unklar verwendete Fachtermini im interdisziplinären und sprachlichen Kontext näher zu erläutern und

final im Kontext der vorliegenden Arbeit zu definieren, passiert dies an der entsprechenden Stelle.

4.4 Ergebnisse der Hauptstudie

Wie auch in der Beobachtungsstudie ersichtlich, wird der Entwicklungsprozess von Suchprofilen auch von den Interviewpartnern durchgängig als nicht linearer, iterativer, mehrfach verschlauerter Prozess geschildert, von einem Interviewpartner auch genau so benannt (GENERAL 11).

In Frage 2 des Fragebogens wurden die Teilnehmer gebeten, ihre typischer Vorgehensweise bei der Erstellung von Suchprofilen zu beschreiben (*Can you describe your typical approach building your search strategies in some simple steps? Do you have a typical order to work on this task?*). Die Angaben zu dieser Frage waren wie die Darstellungen der gesamten Interviews unterschiedlich ausführlich. In Summe führte die sehr offen formulierte Frage dazu, dass nicht nur verschiedene Aspekte des Gesamtprozesses, sondern auch wesentliche Aufgaben identifiziert werden konnten, die beim Aufbau von Suchprofilen eine Rolle spielen. Die Frage kristallisierte sich daher schnell als Einstiegsstelle in die Analyse heraus, da sie die verschiedenen Aspekte und Teilprozesse der Entwicklung von Suchprofilen beleuchtet und in Kombination mit anderen Fragenkomplexen ihre erklärende Komponente erhält. Sie wird daher häufig Einstiegspunkt der folgenden Ergebnisse sein.

Die Analyse der Fragebogenstudie bestätigte im Wesentlichen die Ergebnisse der Beobachtungsstudie und zeigte, dass sich die typische Vorgehensweise unabhängig von einzelnen Testpersonen grob in folgende Phasen und Aufgabenfelder aufteilen lässt- dies bestätigt die Struktur der systematischen Übersichtsarbeit und ist daher nicht überraschend:

- Konstruktion der suchbaren Fragestellung
- Iterativer Prozess der Entwicklung von Suchprofilen inkl. Vollständigkeitskontrolle und Relevanzüberprüfung
- Optimierung von Suchprofilen im Hinblick auf Sensitivität und Noise

Die folgenden Abschnitte beleuchten diese einzelnen Phasen und entwickeln eine erweiterte Struktur zugehöriger *information seeking tasks* und *information searching tasks* (siehe Abschnitt 2.5.3). Anschließend werden jeweils Prozesse und Taktiken beschrieben, die zur Erledigung dieser Aufgaben führen und kontextuelle Rahmenparameter identifiziert, welche die taktischen Überlegungen und das Informationsverhalten beeinflussen. In den Ergebnissen wird – falls nötig – auf die Ergebnisse der Beobachtungsstudie Bezug genommen. Gleichzeitig werden Ergebnisse im Kontext der Methodenlehre beleuchtet, die bei allen Teilnehmern bekannt ist und zur Methodik der Durchführung auch herangezogen wird.

Zitate der Fragebogenteilnehmer sind, sofern nicht ins Deutsche übersetzt, durch eine spezielle Formatierung („Zitat“) gekennzeichnet und können über die Angabe des Fragencodes und der Nummer des Teilnehmers eindeutig zugewiesen werden.

Die Daten zur Hauptstudie finden sich auf einer der Arbeit beiliegenden CD.

4.4.1 Konstruktion der suchbaren Fragestellung

Obwohl nicht explizit erfragt, beschrieben die meisten Interviewteilnehmer in ihrer Antwort auf Frage 2 eingehend ihren Einarbeitungsprozess. Dies zeigt zum einen den Stellenwert der thematischen Einarbeitung, aber auch das Wissen um die eigene Expertise (die Recherche) gegenüber der fachlichen Expertise des Domänenexperten.

4.4.1.1 Forschungsfrage und Domänenwissen

In der Vorbereitungsphase scheint die Zusammenarbeit mit Auftraggebern als Domänenexperten von entscheidender Bedeutung. Im Vorfeld werden daher beim Auftraggeber häufig Zusatzinformationen zur Forschungsfrage angefordert, um so den Grundstein für die Entwicklung von Suchprofilen zu legen.

Ob und welche Informationen angefordert werden, hat zum Teil mit persönlichen Erfahrungswerten zu tun, hängt aber auch von institutionellen Vorgaben ab. So berichtet beispielsweise eine Interviewpartnerin, in ihrer Projektgruppe, einer Cochrane Review Group, würde keine systematische Übersichtsarbeit begonnen, für die der Auftraggeber nicht mindestens zwei randomisierte klinische Studien benennen könne (GENERAL 1). Zu den häufig angeforderten

Materialien zählen neben (englischsprachigen) Schlüsselbegriffen und Synonymen (GENERAL 3, GENERAL 5) vor allem bekannte relevante Schlüsselreferenzen (GENERAL 3, GENERAL 5, GENERAL 10), also Veröffentlichungen, die zu diesem Thema bereits bekannt sind. In welcher Form diese relevanten Schlüsselreferenzen vorliegen scheint unterschiedlich, die Interviewpartner verwenden hier ein breites Spektrum an Begrifflichkeiten wie „relevante Beispielzitate“ (GENERAL 3), „bekannte relevante Artikel“ (GENERAL 5, auch GENERAL 10), „eligible papers“ (GENERAL 7) oder „key relevant papers“ (GENERAL 9). Es kann theoretisch vorkommen, dass Schlüsselreferenzen digital oder ausgedruckt als Volltexte geliefert werden, ebenso kann eine Liste von Schlüsselreferenzen in Form eines bibliographischen Formats wie RIS, BibTex vorliegen. Wichtig scheint an dieser Stelle ausschließlich das Wissen um die *inhaltliche* Relevanz der Schlüsselreferenz als Element der Ground Truth.

Nur in einem Interview werden „Datenbanken und Suchstrategien, in / nach denen der Nutzer bereits selbst recherchiert hat“ sowie „Suchstrategien von Vorgängerrecherchen“ (GENERAL 3) erwähnt.

Die Lieferung von Zusatzinformationen ist keine freiwillige Leistung des Auftraggebers, sondern durchaus eine ihm übertragene Aufgabe. Dies machen Formulierungen wie „der Auftraggeber soll [...]“ (GENERAL 5) oder „lasse ich mir [...] zuschicken“ (GENERAL 3) deutlich.

Immer wieder wird klar, dass sich die Suchenden durchaus der Notwendigkeit eines inhaltlichen Verständnisses und der Notwendigkeit einer engen Zusammenarbeit mit dem Domänenexperten bewusst sind (GENERAL 1, GENERAL 2, GENERAL 4, GENERAL 10, GENERAL 14, GENERAL 15), was von einer Teilnehmerin besonders plastisch dargestellt wird (GENERAL 4):

„Nachfragen und evtl. Treffen mit Auftraggeber. Sich die Fragestellung genau und ausführlich erklären lassen [...] In welche Richtung soll gesucht werden, wo liegt der Schwerpunkt der Fragestellung, was will man damit zeigen, verbessern... [...] Liegen evtl. tatsächlich mehrere Fragestellungen vor. Kann man die Fragestellung enger/weiter fassen? Hat schon jemand eine ähnliche Fragestellung bearbeitet? [...] Oft wird so die genaue Frage die vorliegt erst herausgearbeitet, Blöcke gebildet.“

Wie der Kontakt mit dem Domänenexperten stattfindet, ist unterschiedlich. Eine Teilnehmerin berichtet von initialem E-Mail-Kontakt (GENERAL 3), häufig werden aber auch Treffen mit dem Auftraggeber genannt (GENERAL 1, GENERAL 3, GENERAL 5, GENERAL 10). Das Treffen mit dem Auftraggeber ist eine Möglichkeit zur Klärung offener Fragen, wie unter anderem folgende Interviewpartnerin deutlich macht: „Ggf. nehme ich Kontakt mit den Autoren auf,

um weitere offene Fragen zu klären“ (GENERAL 1). Anmerkung: bei Auftraggebern handelt es sich häufig um die späteren Autoren der zu veröffentlichten Reviews.

Der Ansatz von Bramer, den gesamten Rechercheprozess grundsätzlich gemeinsam mit dem Auftraggeber durchzuführen, scheint eine Ausnahme (GENERAL 6) und ist sicher ausschließlich durch den zeitoptimierten Ansatz (Bramer und Jonge 2015), auf den später noch im Detail eingegangen wird, überhaupt möglich. Eine andere Interviewpartnerin bietet diese Option zumindest an: „Teilweise treffe ich mich mit dem Nutzer und wir erarbeiten die Suchstrategie gemeinsam.“ (GENERAL 3).

Ein meist interdisziplinäres Reviewteam, das beispielsweise den Informationsspezialisten in Cochrane Review Groups zur Verfügung steht, kann den Auftraggeber als Domänenexperten und Ansprechpartner ersetzen (GENERAL 1, GENERAL 14). Für die Einarbeitung stehen hier andere Möglichkeiten zur Verfügung, beispielsweise Vorabsuchen (*Scoping Searches*), deren Ergebnisse im Team besprochen werden (GENERAL 14).

Im Gegenzug dazu wird die Rolle von Auftraggebern oder Domänenexperten von den beiden klinisch tätigen Informationsspezialisten an keiner Stelle des schriftlichen Interviews erwähnt (GENERAL 11, GENERAL 12), er steht offenbar aufgrund vollkommen anderer Rahmenbedingungen für weitere Rückfragen zumindest nicht in diesem Umfang zur Verfügung.

4.4.1.2 Von der Forschungsfrage zum Grundgerüst des Suchprofils

Die Überführung der Forschungsfrage in eine suchbare Fragestellung wird von vielen Befragten ganz klar als Einstieg in die Entwicklung von Suchprofilen genannt. Hauptaufgabe ist dabei die Identifikation „potenzieller Suchkonzepte“ (GENERAL 9), „Elemente“ (GENERAL 6, GENERAL 10) oder auch „Achsen“ und „potenzieller Dimensionen“ (GENERAL 11). Die Verwendung des ein oder anderen Terminus scheint dabei eher willkürlich und variabel (siehe z.B. „elements/concepts“ (GENERAL 12)). Es ist davon auszugehen, dass die jeweils verwendeten Begriffe innerhalb der Fachwelt trotzdem für Kollegen unmissverständlich sind.

Der Terminus Suchprofil wird von den Interviewpartnern übrigens nicht genutzt, ausnahmslos alle Interviewpartner verwenden hierfür den Begriff der Suchstrategie oder englisch *search strategy* wie in der entsprechenden Fachliteratur üblich (siehe Abschnitt 3.4.1). Dies mag zum einen der bewusst gewählten Verwendung des Begriffs in der Formulierung von Frage 2 geschuldet sein, zeigt aber auch sehr deutlich, dass Suchprofile (und damit letztendlich Queries)

die *individuelle* Strategie widerspiegeln, mit der sich Rechercheexperten in Kooperation mit dem Domänenexperten einer Forschungsfrage nähern.

Das Auffinden geeigneter Konzepte stellt für den Suchenden durchaus eine Herausforderung dar. Häufig werden hierfür Hilfsschemata wie das bekannte PICO-Schema (siehe Abschnitt 3.3.1) herangezogen, das von fünf der Befragten genannt wird, aber aufgrund der Nennung in Normen und Handbüchern und seiner engen Verknüpfung mit dem Paradigma der evidenzbasierten Medizin allen Befragten geläufig sein muss. Dass die strukturelle Gliederung der Forschungsfrage in Schlüsselbegriffe den Patienten betreffend, die Intervention betreffend sowie Ein- und Ausschlusskriterien betreffend vor allem bei interventionellen Studien immer im Vordergrund steht, wird auch an anderer Stelle deutlich. So beschreibt eine Interviewpartnerin, wie sie sich während der thematischen Einarbeitung vor allem auf „die Beschreibung der Population sowie Intervention und deren Ein- und Ausschlusskriterien“ (GENERAL 1) fokussiert und Fragestellungen bezüglich komplexer Interventionen als schwierig zu suchen einstuft (PROBLEMS 1).

An anderer Stelle werden „weit gefasste Fragen“, die nur schwer in Blöcke gefasst werden können, als problematisch angesehen, da Schemata wie PICO nicht anwendbar sind (PROBLEMS 4).

Inhaltlich erfolgt die Klärung geeigneter Konzepte grundsätzlich interdisziplinär mit dem Auftraggeber oder einem Domänenexperten im Reviewteam (vgl. GENERAL 1, GENERAL 4).

Mit der Identifikation dieser Konzepte wird die Forschungsfrage nicht nur thematisch erfasst, sondern auch syntaktisch gegliedert, wie folgende Passage deutlich macht (GENERAL 16):

„I start by parsing the question into searchable concepts“

So entsteht ein grobes Strukturmodell der späteren Query, in dem die thematischen Konzepte der klinischen Fragestellung als so genannte Building Blocks syntaktisch verknüpft sind (GENERAL 15):

„An erster Stelle steht das Erstellen von Konzepten (Building Blocks) als Basis für die weiteren strukturierten Suchen.“

Die thematischen Konzepte der Forschungsfrage haben also in den Building Blocks, die später im Rahmen der Datenbanksuche mit AND-Operatoren verknüpft werden, ihr direktes Äquivalent, die suchbare Fragestellung entspricht damit dem rudimentären Grundgerüst des späteren Suchprofils.

Die Auswahl geeigneter Konzepte unterliegt aber nicht nur der initialen Forschungsfrage, sondern durchaus auch ökonomischen Gesichtspunkten:

„Choose which 2-4 concepts to use (this may involve a little testing to see which concepts are most simple to search)” (GENERAL 9),

“decide which elements (usually 2) to search” (GENERAL 10)

Eine Suchstrategie kann demnach auch dadurch geprägt sein, dass auf einzelne Konzepte der initialen Fragestellung aus strategischen Gründen verzichtet wird, beispielsweise, weil diese „nicht einfach“ zu suchen sind. Gleichzeitig führt die Identifikation geeigneter (weil einfach zu suchender) Building Blocks zur Notwendigkeit von Testsuchen, um beurteilen zu können, welche dieser Blöcke geeignet sind.

Suchprofile werden nicht parallel für verschiedenen Datenbanken entwickelt, sondern zunächst für eine Datenbank erarbeitet und nach Fertigstellung für weitere Datenbanken übersetzt, genannt beispielsweise bei GENERAL 1, GENERAL 5, GENERAL 6, GENERAL 10). Ein Interviewpartner gibt an, auf einem Host zu recherchieren und daher eine einheitliche Syntax und Retrievalsprache für eine Vielzahl dort verfügbarer Fachdatenbanken nutzen zu können (GENERAL 15). Da auch auf dem Host datenbankspezifische Felder und weitere Sonderfälle Berücksichtigung finden, fokussiert sich auch dieser Interviewpartner zunächst auf eine der zu durchsuchenden Datenbanken (GENERAL 15):

„Ich recherchiere bei einem Host (DIMDI), der die simultane Suche in zahlreichen Datenbanken mit einer einheitlichen und mächtigen Syntax/Retrievalsprache erlaubt (datenbankspezifische Felder, Thesauri und Schlagwörter können genutzt werden und diese berücksichtige ich). Dies umfasst die wichtigsten und umfangreichsten bibliographischen Datenbanken [...]. Ich beginne mit der Suchstrategie für MEDLINE [...] und ergänze dann um die Spezifika [...] der weiteren, ausgewählten Datenbanken.“

Fast einheitlich wird MEDLINE als die Datenbank genannt, für die das initiale Suchprofil erstellt wird. Lediglich in den Ausführungen von Bramer ist EMBASE die Datenbank der Wahl. Es

scheint pragmatische, vielleicht auch ökonomische Gründe zu haben, die auch genannt werden und zur Empfehlung „EMBASE first“ führen. Bramer nutzt Microsoft Word und dort programmierbare Makros zur semi-automatisierten Übersetzung von Suchprofilen zwischen einzelnen Datenbanken und schreibt (Bramer und Jonge 2015):

“The embase thesaurus is more complex, therefore it is easier to translate Emtree to MeSH than vice versa. Recommendation: If you have embase, start in embase, otherwise start in Medline”

Wie fokussiert die Experten in ihrem Tätigkeitsbereich auf die Suche in bibliographischen Datenbanken sind, zeigt sich darin, dass kein Interviewpartner bei Beantwortung der sehr allgemein gehaltenen Frage 2 auf die Suche in Grauer Literatur eingeht, sondern sich ausschließlich auf den Kontext bibliographischer Datenbanken beschränkt. Wie bereits in den Ergebnissen der Beobachtungsstudie deutlich gemacht, sind thematische und formale Suchprozesse häufig Aufgabe des Auftraggebers (vgl. beispielsweise TACTICS_BS 1, GENERAL 15).

Suchprofile sind zu Beginn der Entwicklung rudimentäre Konstrukte aus Building Blocks und werden sukzessive ausgebaut und weiterentwickelt. Bereits in der Beobachtungsstudie lag dies im Fokus, für die Fragebögen wurde dies konzeptuell um folgende Fragen erweitert:

Welche Aufgaben müssen durchgeführt werden, um ein qualitativ hochwertiges Suchprofil zu generieren und fertigzustellen?

Wie werden diese Aufgaben durchgeführt?

Welche Rahmenbedingungen beeinflussen die Wahl entsprechender Suchtaktiken?

Wie sehen Entscheidungswege aus?

4.4.2 Suchprofil und Informationsressourcen

Erste Aufgabe bei der Entwicklung von Suchprofilen ist die Identifikation geeigneter Suchterme zum Aufbau einer möglichst vollständigen Query in einzelnen bibliographischen Datenbanken.

Der folgende Abschnitt widmet sich dem Zusammenhang von verfügbarer Informationsressource und Suchprofil sowie der Wahrnehmung spezifischer Charakteristika einzelner Informationseinheiten durch den Informationsspezialisten.

Per definitionem kann diese Aufgabe der Identifikation von Suchtermen zunächst in zwei komponentenabhängige Teilaufgaben zerlegt werden: Das Auffinden und Ergänzen von Deskriptoren sowie das Auffinden und Ergänzen von Freitexttermen und Phrasen.

Die Suche mit geeigneten Deskriptoren dient dem Auffinden aller in der Datenbank entsprechend indexierten bibliographischen Referenzen. Um auch nicht-indexierte oder unzureichend indexierte Referenzen auffinden zu können und damit den Anforderungen der systematischen Übersichtsarbeit gerecht zu werden, ist eine Suche mit Freitexttermen und gegebenenfalls auch Phrasen unerlässlich (beschrieben bei GENERAL 2 und GENERAL 11). Dies spiegelt die klassische Methodenlehre wider.

4.4.2.1 Synonyme

Die zentralen Komponenten sind die zu den durch Schlüsselbegriffe repräsentierten Building Blocks passenden *Synonyme*, ein Begriff, dessen Verwendung zunächst geklärt werden muss.

Von den Befragten wird der Begriff Synonym sehr häufig, jedoch in unterschiedlicher sprachlicher Schärfe verwendet, beispielsweise lässt die Verwendung des Terminus Synonymen für Begriffe, die vorbereitend vom Auftraggeber genannt werden (GENERAL 5), darauf schließen, dass unter Synonymen im Wesentlichen Begriffe mit gleicher oder ähnlicher Bedeutung des Keywords subsummiert sind. Andere Interviewpartner verwenden den Begriff im Kontext der für Deskriptoren genannten Querverweise, im MeSH-Thesaurus *Entry Terms* genannt, bei EM-TREE *synonyms*: „Handelt es sich um ein Synonym, welches im Thesaurus nicht verzeichnet ist [...].“ (UNSPEC_SYN 2) oder „Verweise auf verwandte Deskriptoren sowie Synonymbezeichnungen“ (SELECTION_DESC 3).

Im Deutschen definiert der Terminus Synonym Begriffe, die „mit einem anderen Wort oder einer Reihe von Wörtern von gleicher oder ähnlicher Bedeutung [sind], sodass beide in einem bestimmten Zusammenhang austauschbar sind; sinnverwandt“ (Bibliographisches Institut GmbH Duden Verlag 2016h). Das Oxford English Dictionary definiert das Synonym als „*strictly, a word having the same sense as another (in the same language); but more usually, either or any of two or more words (in the same language) having the same general sense, [...]*“ (Oxford University Press 2016h).

Im vorgestellten Modell fungiert das Synonym als allgemein gehaltenere Überbegriff für Deskriptoren als auch Freitextterme und Phrasen, darunter auch der oben genannte Spezialfall

der in Thesauri gelisteten *Entry Terms*. Das Synonym steht damit für jeden Begriff, der die Kernkonzepte der suchbaren Fragestellung sinngemäß optimal repräsentiert.

4.4.2.2 Deskriptoren als datenbankspezifische Suchterme

Als Quelle für Deskriptoren fungieren sowohl die die Indexierungsangaben aus Dokumenten bibliographischer Datenbanken (genannt in ORDER 1-3, GENERAL 7, ORDER 9, ORDER 14) als auch die Thesauri bibliographischer Datenbanken selbst (genannt in GENERAL 2-7).

Der Umgang mit Thesauri und Deskriptoren ist in der Praxis eng verknüpft mit den von den Suchenden vermuteten oder bekannten Indexierungscharakteristika einzelner Fachdatenbanken. Hierbei spielen mehrere Faktoren eine Rolle: Die Regelmäßigkeit der Indexierung sowie die Verwendung verfügbarer Deskriptoren im Kontext der thematischen Abdeckung einzelner Datenbanken. Genannt werden beispielsweise Indexierungszeitpunkt und Indexierungsturnus (GENERAL 2) einzelner Datenbanken wie in Kapitel 3 beschrieben. Sind aktuelle Einträge noch nicht indexiert, so sind sie über eine exklusive Suche mit Deskriptoren nicht auffindbar. Gleichzeitig sind ältere Einträge je nach Datenbank eventuell gar nicht indexiert.

Auch der Verwendungszeitraum und die Aktualisierung von Thesauri, beispielsweise durch die Ergänzung um neue Deskriptoren, wird als ein für die Recherche relevanter Aspekt genannt. So beschreibt eine Befragte, wie wichtig Testsuchen sind, um bei Aktualisierung des Thesaurus die Verwendung neu ergänzter Deskriptoren zu überprüfen und diese in den Kontext ähnlicher, älterer Deskriptoren zu setzen (TEST 8):

”Checking on the output of new-ish terms in a subject thesaurus, [...] Similar concepts may have appeared before, but the buzzword has only appeared recently”

Andere Interviewpartner weisen auf die Breite der Indexierung (PRECISION 14, PRECISION 15) hin, vorwiegend für die Datenbank EMBASE. Hier wird, wie vereinzelt von Interviewpartnern angegeben, teils „sehr breit indexiert“ (PRECISION 15) oder Deskriptoren werden überbeansprucht (PRECISION 14).

Ein Teilnehmer macht mit der kritischen Bemerkung

”database personnel are known to miss important aspects of articles when assigning indexing terms (particularly at Medline)”

auf die Indexierungsqualität manuell indexierter Datenbanken aufmerksam und ist sich bewusst, dass einzelne Themen über Datenbanken hinweg nicht immer gleich gut verschlagwortet sind (GENERAL 11).

Aufgrund ihrer jahrelangen persönlichen Erfahrung, der fachlichen Schulung zu einzelnen Datenbanken und anderen Weiterbildungsmöglichkeiten, sowie einem teils regen Austausch mit Kolleginnen und Kollegen ist davon auszugehen, dass die genannten Datenbankspezifika allen Interviewpartnern bekannt sind. Sie können daher auch ohne explizite Nennung als teils kritische Kontextparameter verstanden werden, derer sich die Experten bewusst sind.

Eine Befragte weist darauf hin, wie wichtig es ist, Deskriptoren nicht nur zu identifizieren, sondern auch deren Verwendung zu verstehen (GENERAL 8):

”Always check what some subject terms from the thesaurus from the database do (in terms of output) – to check on indexer interpretation of scope notes for subject descriptors.“

Sie überprüft die Suchergebnisse (“in terms of output“) einzelner Deskriptoren, um die Verwendung von Deskriptoren zu analysieren. „*Scope Notes*“ ist eine Ovid-spezifische Bezeichnung für die kurzen Erläuterungstexte zu Deskriptoren des MeSH-Thesaurus, die Befragte nutzt diese Erläuterungstexte also zum Verständnis der Indexierung.

4.4.2.3 Freitextterme und Phrasen

Freitextterme wie auch Phrasen umfassen zu Beginn der Entwicklung von Suchprofilen lediglich alle die gewählten Building Blocks repräsentierenden, offensichtlichen Keywords sowie die vom Domänenexperten gelieferten Fachtermini und Synonyme.

Die offensichtlichste Quelle für synonyme Freitextterme und Phrasen bilden bibliographische Dokumente (genannt in ORDER 1-3, GENERAL 7, ORDER 9, ORDER 14), darüber hinaus Volltexte sowie systematische Übersichtsarbeiten. Die umfassende Suche nach Freitexttermen und Phrasen wird mit den bereits beschriebenen Indexierungscharakteristika einzelner Fachdatenbanken als auch mit der Notwendigkeit, ein breit angelegtes Suchprofil zu generieren, begründet. Freitextterme sind zusätzlich in Beschreibungstexten kontrollierten Vokabulars verfügbar, siehe hierzu auch Abschnitt 3.4.3.

Volltexte, auf die in bibliographischen Datenbanken meist zumindest ein Querverweis verfügbar ist (siehe Abschnitt 3.1), haben für die Identifikation von Synonymen unterschiedlichen Stellenwert. Wie eine Befragte deutlich macht, stünden sie in den „seltensten Fällen [...] für das Retrieval zur Verfügung“ (IDENT_FULL 1), ein Hinweis darauf, dass sich die Suche auf bibliographische Datenbanken beschränkt und Suchbegriffe aus dem Corpus von Volltexten für das Retrieval zugehöriger bibliographischer Einträge daher nicht relevant sind.

Andere Befragte nutzen Volltexte, sofern sie ihnen zur Verfügung stehen (GENERAL 3, GENERAL 5, IDENT_FULL 2), siehe hierzu auch die Ergebnisse der Vorstudie Abschnitt 4.2.4.1.

Die Beschaffung von Volltexten wird dabei durchaus als kritische Aufgabe wahrgenommen, wie eine Interviewpartnerin deutlich macht: „Aber uns stehen längst nicht alle Volltexte sofort zur Verfügung.“ (IDENT_FULL 2).

Wo die Probleme liegen und wie bei der Beschaffung von Volltexten vorgegangen wird, ist aus den Daten nicht ermittelbar. Es ist aber zu vermuten, dass Probleme mit der Beschaffung weniger mit der Lokalisierung von Volltexten als vielmehr der Lizenzierung von Journals an den Universitäten und Einrichtungen zu tun haben. In der den Interviews vorgelagerten Beobachtungsstudie wurde in solchen Situationen häufig Google Scholar genutzt, vgl. Abschnitt 4.2.4.1.

4.4.2.4 Datenbankspezifischer Suchterm und Suchprofil

Bei der Sammlung von Synonymen handelt es sich zunächst um „Kandidaten“ (GENERAL 1, SELECTION_TERM 15), Terme also, die als Vorschlag für das spätere finale Suchprofil zu interpretieren sind. Auch dies war in der Beobachtungsstudie bereits deutlich geworden.

Diese Sammlung erfolgt beispielsweise in einem Microsoft Word Dokument (GENERAL 1, (Bramer und Jonge 2015)) oder einem Texteditor (Beobachtungsstudie). Die *Sammlung* von Suchtermen in Form eines Suchprofils ist damit klar abzugrenzen von der *Suche* mit Suchtermen (der Query), wie auch einzelne Formulierungen deutlich machen: es handelt sich hier vielmehr um eine Zusammenstellung oder „Abfolge“ (GENERAL 15) von Termen, die „notiert“ werden (GENERAL 1, GENERAL 3).

Das Suchprofil besteht in der Praxis also nicht aus einer Sammlung aus mittels Booleschen Operatoren verknüpften Suchtermen, sondern stellt ein mittels der Building Blocks strukturiertes Dokument an Suchtermen dar. Erst durch Anwendung dieser Suchterme in einzelnen

Queries und die logische Verknüpfung der Treffermengen und Building Blocks direkt auf bibliographischen Datenbanken wird daraus die Query für die gesamte suchbare Fragestellung. Dies spricht deutlich für die ausschließliche Verwendung professioneller Benutzeroberflächen, die die Verkettung mehrerer, in der Suchhistorie verfügbarer Suchanfragen ermöglicht, siehe hierzu Abschnitt 3.4.2.3.

Kandidatenterme werden von den Experten nicht nur als inhaltlich verwandter Begriff, sondern insbesondere bei Freitexttermen und Phrasen als Lexeme wahrgenommen und auch so behandelt: Bereits in der Beobachtungsstudie zeigte sich, dass Synonyme, nachdem sie im Text erkannt wurden, sofort als über den Wortstamm definierte Freitextterme in einer üblichen datenbankspezifischen Syntax trunkiert oder maskiert notiert werden (siehe Abschnitt 4.2.4.1). Es war demnach keine klare Trennung der Aufgaben „Identifikation eines Synonyms“ und einer im Suchprofil verwendeten „datenbankspezifischen Anpassung“ zu erkennen, der Prozess wirkte eher automatisiert. Wann immer ein Synonym identifiziert wurde, folgte im nächsten Schritt eine prototypische, an die Datenbanksyntax angepasste Erfassung dieses Synonyms im Suchprofil. Die Identifikation des (fiktiv gewählten) Begriffs *searching* führt beispielsweise ohne weiteres Nachdenken zum Notieren des Suchterms *search**, die Identifikation des Begriffs *behaviour* analog zum Suchterm *behavi?or*, der amerikanisches als auch britisches Englisch gleichermaßen abdeckt.

Dies beschreibt auch Bramer, der mit seiner Vorgehensweise diesen Prozess semi-automatisiert, indem er in einem Word Dokument eine Reihe vorgefertigter datenbankspezifischer Syntaxbausteine für Trunkierung und Maskierung vorsieht, die über Copy & Paste wiederverwendet werden können (Bramer und Jonge 2015). Äußerungen anderer Interviewteilnehmer lassen darauf schließen, dass auch synonyme Phrasen als Mehrwortterme sofort datenbankspezifisch notiert oder mittels PROXIMITY-Operatoren verknüpft werden (GENERAL 1). Mehr noch haben die Experten ihre individuelle und standardisierte Methode im Umgang mit der Breite von PROXIMITY-Operatoren – meist schwankt diese zwischen 2 und 5 Wörtern Abstand (ADJ_TEST 4, ADJ_TEST 15).

Eine sehr stringente Vorgehensweise findet sich beispielsweise bei Bramer, der auch bei PROXIMITY-Operatoren den (an dieser Stelle nach EMBASE Syntax) vorgefertigten Baustein (Bramer und Jonge 2015)

“(O NEAR/3 O)”

nutzt und an anderer Stelle betont, gerne große Abstände zu nutzen, um von Beginn an eine hohe Sensitivität garantieren zu können (ADJ_TEST 6).

Bei der Identifikation der Phrasen *information searching behaviour* sowie *information seeking behaviour* würde demnach beispielsweise der Suchterm *information ADJ3 behavio?r* beziehungsweise *information NEAR2 behavio?r* notiert werden. In den Interviews wird dieser Schritt vom Synonym zum datenbankspezifischen Freitextterm nur von einem Interviewpartner ausdrücklich erwähnt, der Interviewpartner nutzt häufig PROXIMITY-Operatoren und weist darauf hin, dass auch sprachliche Besonderheiten wie amerikanische und britisches Englisch oder Eigenheiten nicht-muttersprachlicher Autoren berücksichtigt werden müssen (GENERAL 11):

”Keyword searches are often supplemented with proximity searches [...]. Authors whose native language is not English, or who publish in non-US journals, may use phrasings different from the majority of writers, [...] a proximity search will capture many of them. I also make sure to use wildcards as necessary to pick up plurals, British spellings, and other variants”

Der Vorgang des *Notierens* ist also mehr als nur die Identifikation und das Zusammenschreiben von Kandidatentermen, es umfasst im Umgang mit Freitexttermen und Phrasen gleichzeitig die routinierte Formulierung an die Datenbanksyntax angepasster Suchterme, im Umgang mit Deskriptoren die entsprechende Kennzeichnung als kontrolliertes Vokabular.

Abbildung 4-5 fasst die Begrifflichkeiten abschließend graphisch zusammen:

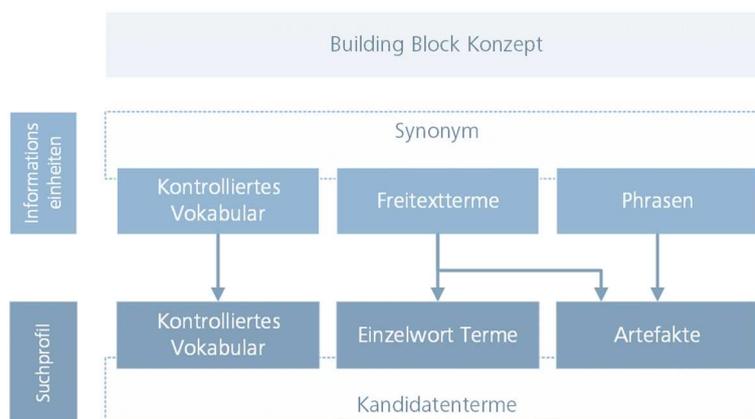


Abbildung 4-5 Vom Synonym zum Kandidatenterm

4.4.3 Bewertung von Synonymen, Kandidatentermen und Suchprofil

Aus der Definition systematischer Übersichtsarbeiten heraus ist die ideale Suchanfrage auf bibliographischen Fachdatenbanken so sensitiv wie nötig und so spezifisch wie möglich anzulegen, vgl. Abschnitt 3.4.3.

Bisher ist nicht bekannt, wie Informationsexperten mit dieser Anforderung, die häufig zu hohen Treffermengen und einem umfangreichen Screening und Analyseprozess führt, in der Praxis umgehen. Der folgende Abschnitt soll zunächst Aufschluss über mögliche Bewertungsschemata für Suchterme und Suchprofil geben.

Mehrere Fragen zielten in diese Richtung:

1. Welche Kriterien sind Ihnen bei der Aufnahme eines Terms aus einem relevanten Dokument in Ihre Suchstrategie besonders wichtig? (*Which are the crucial factors and conditions to add new synonyms into your search query?*)
2. *What makes you feel confident about synonyms?*
3. Wie erfolgt typischerweise der Umgang mit einem eher „unspezifischen“ Synonym, das in einem relevanten Volltext gefunden wird? Welche Entscheidungen sind hierfür ausschlaggebend? (*How do you typically handle with a rather "non-specific" synonym term? Which factors are particularly crucial when adding or dropping such "non-specific" terms to your search query?*)

Dies führte während der Analyse zur Beantwortung weiterer Fragen

- Wie gewinnen die Experten Vertrauen in die Qualität der von ihnen entwickelten Suche?
- Was macht die Qualität eines Suchterms oder einer Query aus?

4.4.3.1 Entscheidung über Synonyme als Kandidatenterme

Mit Abschluss der Vorbereitungsphase steht üblicherweise eine Reihe an Informationen zur Verfügung, die für die Sammlung von Kandidatentermen genutzt werden können:

- Die strukturierte suchbare Fragestellung aus Building Blocks als durch Schlüsselbegriffe repräsentierte thematische Konzepte
- bekannte Schlüsselreferenzen
- alle vom Domänenexperten benannten Synonyme und Fachtermini
- Auswahl zu durchsuchender bibliographischer Datenbanken und korrespondierende Thesauri

Welche Faktoren entscheiden nun darüber, ob Synonyme als Kandidatenterme aufgenommen werden, was macht einen „guten Suchterm“ aus?

Entscheidung über kontrolliertes Vokabular, Freitextterme oder Phrasen

In den Interviews wird deutlich, dass die Präferenz von kontrolliertem Vokabular gegenüber Freitexttermen und Phrasen deutlich überwiegt, die Sammlung von Suchtermen in strukturierten Suchprozessen ist demnach stark datenbankspezifisch geprägt und in der Indexierung bibliographischer Dokumente begründet.

Die im Folgenden zitierte Befragte bezieht sich mit dem Begriff Dokument auf bibliographische Dokumente medizinischer Fachdatenbanken:

„Zunächst Übernahme von Deskriptoren. Danach schauen, ob im Dokument weitere relevante Suchworte auftauchen.

Wenn kein passender Deskriptor vorhanden [...] haben die Suche nach Synonymen und Umschreibungen, Phrasen und Ausdrücken einen höheren Stellenwert als wenn eine sachliche Suche schon gut durch einen Deskriptor abgedeckt ist.“ (SELECTION_TERM 5)

Dies wird von einigen Teilnehmern ausdrücklich erwähnt, findet sich aber auch in der Formulierung des allgemeinen Vorgehens, unter anderem:

„[...] suche ich zunächst nach Deskriptoren“ (GENERAL 2)

„zuerst Deskriptorensuche“ (GENERAL 3)

Synonyme als eindeutige Repräsentanten von Building Blocks

Deskriptoren als auch Freitextterme sind als Kandidatenterme dann „ideal“, wenn sie sich 1:1 mit den Schlüsselbegriffen einzelner Building Blocks decken. Als verdeutlichendes Beispiel sei hier noch einmal ein bereits an anderer Stelle zitierter Review mit dem Titel *Parent-only interventions for childhood overweight or obesity in children aged 5 to 11 years* (Loveman et al. 2015) genannt. Die Schlüsselbegriffe „*overweight*“ und „*obesity*“ (Übergewicht und Fettleibigkeit) sind als kontrolliertes Vokabular Teil des MeSH-Thesaurus und daher „ideale“ Suchterme (SELECTION_TERM 5, GENERAL 11, SELECTION_TERM 15) und damit auch idealer Einstiegspunkt (GENERAL 11):

„I examine the indexing (MESH, EMTREE) to ascertain whether the database does or does not have a precise term for that dimension, and if it does, start with that.“

Synonyme als Notwendigkeit für das Retrieval relevanter Dokumente

Zwei Gründe beherrschen Entscheidungssituationen über die Aufnahme von Synonymen als Kandidatenterme:

Die Sicherung des Retrievals eines als relevant bekannten Dokuments sowie der Wunsch nach einem umfassenden, erschöpfenden Suchprofil, was von den Befragten eindeutig formuliert wird (SELECTION_TERM 1, SELECTION_TERM 3, SELECTION_TERM 5- 7, SELECTION_TERM 15). Das erschöpfende Suchprofil ist per definitionem dann erreicht, wenn mit der Hinzunahme weiterer Suchterme keine neuen relevanten Treffer mehr identifiziert werden können, siehe auch Abschnitt 3.4.3. Umgekehrt sind alle Suchterme, die weitere, bisher nicht bekannte relevante Dokumente liefern oder solche, ohne die vorab bekannte, relevante Treffer verloren gingen, notwendig zum Erreichen eines erschöpfenden Suchprofils. Die folgenden Zitate aus Fragebögen untermauern dies.

„[weiteres Synonym, Retrieval des Dokuments, Vorkommen in bestimmten Feld...] Die beiden erstgenannten sind sehr wichtig.“ (SELECTION_TERM 1)

„weitere Synonyme für Freitextsuche für die eigene Suchstrategie finden“ (SELECTION_TERM 3)

“Reason to add a term is that it possibly results in higher recall.” (SELECTION_TERM 6) – hier wird deutlich, dass es sich tatsächlich um einen Kandidatenterm handelt, der später hinsichtlich dieses Kriteriums weiter überprüft werden muss.

„Crucial factor is if I need the term to capture one of the papers in the validation set” (SELECTION_TERM 7)

“New synonyms would be definitely be used if they [...] retrieve further relevant results that the previous search did not find.” (SELECTION_TERM 9)

„Beim SR liegt der Fokus oft auf der Maximierung des Recalls.“ (SELECTION_TERM 15)

Dabei scheint auch relevant, in welchen Strukturelementen (vorwiegend bibliographischer Dokumente) die Kandidatenterme identifiziert werden, wobei die Aussagen hier eher indifferent sind:

“New synonyms would be definitely be used if they appear in the title [...]” (SELECTION_TERM 9)

„Dabei gibt es eine ungefähre Rangfolge: Häufig verwendete Terme, die möglichst auch oft im Titel vorkommen.[...]“ (SELECTION_TERM 15)

„Nach brauchbaren Suchworten wird zunächst in den Abstracts gesucht, auch in den Titeln [...] Wichtige Suchworte kommen nicht unbedingt gleich im Titel vor.“ (SELECTION_TERM 5)

„Das Vorkommen in einem bestimmten Feld evtl. auch, aber nachgeordnet.“ (SELECTION_TERM 1)

Zusammenfassend unterliegt die Entscheidung über die Aufnahme eines Synonyms in die Liste der Kandidatenterme ausgehend von den vorliegenden Ergebnissen durchaus einigen Regeln: Grundsätzlich ist kontrolliertes Vokabular, soweit verfügbar, bei der strukturierten Suche der Verwendung von Freitexttermen und Phrasen vorzuziehen. Der ideale Deskriptor findet sich in der direkten Entsprechung zu mit Building Blocks korrespondierenden Schlüsselbegriffen als 1:1 Match.

In Summe sind Entscheidungssituationen zur Aufnahme geeigneter Synonyme als Kandida-

tenterme stark von der methodologischen Anforderung nach umfassenden und erschöpfenden Suchprofilen geprägt und folgen demnach dem Prinzip „Collect all“. Hierbei spielt das Retrieval bekannter relevanter Dokumente eine tragende Rolle.

4.4.3.2 Problematische Kandidatenterme, problematisches Suchprofil

Das Wissen um das Vertrauen in Suchterme und deren Nutzen im Hinblick auf Ertrag und Performanz des aktuellen Suchprofils ist nicht grundsätzlich gewährleistet:

“finding out about synonyms can be difficult – depends on background literature to the research question, personal knowledge within review team (or any advisory group of the review)[...]” (TERM_CONFID 14)

„[...] bin ich auch oft gar nicht sicher, ob ein bestimmtes Synonym notwendig im Rahmen einer Suchstrategie ist“ (SELECTION_TERM 15)

Dieses notwendige Vertrauen in Suchterme und Suchprofil hat sowohl eine qualitative Ebene, die zur Notwendigkeit führt, Suchterme zu validieren, als auch eine quantitative Ebene, die dazu führt, dass Suchterme hinsichtlich ihres Ertrags und ihrer Precision evaluiert werden müssen.

Die Aufgabe der Validierung von Synonymen dient nicht nur der Überprüfung, ob „ein Term inhaltlich in die richtige Richtung geht“ ((TEST_REASONS 15), wie es von einem Interviewpartner beschrieben wird (vgl. auch IDENT_FULL 12, PROBLEMS 14), sondern auch dem eigenen „Lernen über das Thema und seinen Kontext“ (ORDER 15). Eine Notwendigkeit zur Validierung von Synonymen ergibt sich meist aus einem spontanen, vorab nicht erkennbaren Informationsbedarf, der durch die Entdeckung „neuer“, d.h. bisher unbekannter Synonyme getriggert wird (z.B. ORDER 15, TEST-REASONS 14), und gehört demnach nicht zu den Routineaufgaben. Mehr noch führt dies häufig zur Unterbrechung von Routinetätigkeiten, weil „[...] regelmäßig neu entdeckten Termen und ihrer Bedeutung oder Verwendung [...]“ (ORDER 15) nachgegangen werden muss.

Nicht immer müssen Synonyme validiert werden, sodass folgende Klassifizierung möglich ist:

- **Exakte Übereinstimmung/exact match:** Synonyme, die sofort als ideale Kandidatenterme erkannt werden oder ein exact match mit einem Building Block

bilden („However, it is simply just a recognition of that word as a synonym for the concept you are looking for. E.g. when you see the word ‘diet variety’ you realize that you should probably include it in your food concept of ‘food repertoire’” (TERM_CONFID 16))

- **Partielle Übereinstimmung/partial match:** Synonyme, die als potenzielle interessante und gleichzeitig kritische, wahrscheinlich zu allgemein gehaltene Kandidatenterme erkannt werden („Ich versuche [...] initial Kandidatenterme zu finden, die nach Testsuchen sowie nach Absprache mit den Domänenexperten möglicherweise auch wieder [...] gestrichen werden [...].“ (SELECTION_TERM 15))
- **Unklare Übereinstimmung/unclear match:** sowie Synonyme, die erst verifiziert werden müssen, weil sie dem Suchenden inhaltlich als auch im Hinblick auf ihren Ertragswert im Suchprofil unklar sind

Die Frage *What makes you feel confident about synonyms?* zeigte, dass Synonyme nach bestimmten Kriterien auch einen „Vertrauensvorschuss“ genießen können, weil sie entweder durch eine entsprechende Fundstelle validiert sind (beispielsweise Nennung durch den Auftraggeber) oder durch die Notwendigkeit des Retrievals bekannter oder neu hinzugekommener relevanter Dokumente für eine vollständige Suchstrategie erforderlich sind oder scheinen.

„I would add synonyms based on experience of use in previous reviews [...]“ (TERM_CONFID 8)

“I would add synonyms [...] and by comparing outputs [...] whether we are missing a set of possible studies that may only be retrieved by adding a new synonym.” (TERM_CONFID 8)

“Use in previous reviews. Occurrence in specific field. Recommendations from stakeholders.” (SELECTION_TERM 12)

Nach Auskunft einiger Befragter ist die Verwendung eines Suchterms in thematisch verwandten Reviews kein hinreichendes Kriterium für die Aufnahme als Kandidatenterm, insbesondere sind Kopien vorbestehender Suchprofile keine Strategie in der Entwicklung von Suchprofilen, beispielsweise:

“[...] never use search strategies from other reviews. I hardly reuse earlier search strategies, as in my opinion every search is new. Only if a researcher returns and says: I want

exactly that, but now with one difference I reuse part of an old strategy” (SELECTION_TERM 6)

„[...] Suchstrategien früherer SRs waren mir meist nur Inspirationsquellen oder als Ausgangs- oder Ergänzungspunkt hilfreich. [...] Außerdem waren Suchstrategien oft nicht reproduzierbar dokumentiert. [...] konnte ich die Ratio hinter Suchstrategien nicht nachvollziehen bzw. diese erschien mir fragwürdig und deshalb ungeeignet.“ (SELECTION_TERM 15)

Gleichzeitig sind Kandidatenterme besonders geeignet, wenn Sie möglichst wenig Ballast mitbringen, also so sensitiv wie möglich angelegt sind, aber Treffermengen mit geringer Noise (synonym auch Grundrauschen) zurückliefern, die definiert werden kann als $1 - precision$, siehe hierzu auch Abschnitt 2.3.2.5.

“Feel confident if the synonym makes sense and doesn’t introduce too much noise” (TERM_CONFID 7)

“I feel confident adding a synonym when it brings in new relevant articles without bringing in a lot of useless articles.” (TERM_CONFID 16).

Als problematisch zeigen sich im Umkehrschluss Kandidatenterme, die wenig oder einen schlecht beurteilbaren, neuen Ertrag an weiteren relevanten Dokumenten liefern (Delta-Sensitivität) sowie ein erhöhtes Grundrauschen mitbringen:

„Unspezifische Synonyme werden möglichst weggelassen, um die Precision nicht übermäßig zu verschlechtern.“ (UNSPEC_SYN 1)

„...schaue mir an, durch welche Begriffe/Schlagwörter irrelevante Treffer gefunden werden. [...]“ (GENERAL 1)

„Nachdem klar ist, dass durch die Suche zuviel Ballast mit an Bord kommt, [...]“ (UNSPEC_SYN 5)

“if they appear in the title and they add a manageable number of new relevant records then I feel more confidence [...]” (UNSPEC_9)

Da Kandidatenterme aufgrund der Konzeption von Suchanfragen als Building Blocks ausschließlich im gesamten Suchprofil getestet werden können, betrifft jede quantitative Bewertung nicht nur den Term selbst, sondern das gesamte Suchprofil und kann daher als Evaluierung für den gesamten, datenbankspezifischen Suchprozess gewertet werden.

Vor allem Beurteilung der Vollständigkeit von Suchprofilen wird unter den Befragten als komplexe Aufgabe und problematischer Aspekt wahrgenommen, was einige der Interviewpartner auf die Frage „Was sind Ihrer Meinung nach die problematischsten Aspekte beim Auffinden von Deskriptoren und Freitexttermen und deren „passender“ Eingliederung ins Suchprofil?“ (*“What are the most problematic or resource-intensive aspects during search strategy development?”*) folgendermaßen formulieren:

„Am problematischsten ist die fehlende theoretische oder empirische Grundlage für die Entscheidung, einen Suchterm mit in die Suche aufzunehmen oder nicht“ (PROBLEMS 15)

„Man kann nie sicher sein, alle fuer die Suche relevanten Ausdruecke gefunden und verwendet zu haben.“ (PROBLEMS 5)

„Vielfalt der sprachlichen Ausdrucksmöglichkeiten, besonders bei psycho-sozialen Fragestellungen“ (PROBLEMS 4)

4.4.4 Kontrolle des Entwicklungsprozesses

Wichtiges Hilfsmittel in der Kosten -, Nutzenkontrolle als auch Optimierung von Suchtermen und Suchanfragen sind Testsuchen, es kann zwischen einfachen Testsuchen und vergleichenden Testsuchen unterschieden werden. Im Folgenden werden Methoden beschrieben, die zum Teil sehr individuell zu sein scheinen, andere wiederum dürften aufgrund der Häufigkeit ihrer Nennungen zum üblichen Repertoire gehören.

4.4.4.1 Einfache Testsuchen

Einfache Testsuchen geben den aktuellen Stand des derzeitigen Suchprofils wieder. Sie werden daher meist durch datenbankspezifische Suchanfragen mit dem vollständigen, derzeit verfügbaren Stand der Query durchgeführt.

Einfache Testsuchen haben mehrere Ziele und wurden in den Interviews von einigen Teilnehmern genannt, die Zitate stellen aufgrund der Datenlage jedoch lediglich prototypische Beschreibungen dar:

Ausbau der Sammlung an Kandidatentermen

mittels der Identifikation weiterer Kandidatenterme durch Screening der Treffermenge. Hierbei muss nicht zwingend eine direkte Analyse bibliographischer Referenzen erfolgen, siehe (GENERAL 6):

“Read the abstracts of relevant articles and scan them for extra relevant words. Add those relevant words in the title abstract part of the element [...] Scan thesaurus terms of relevant articles for missed relevant terms. Add those to the original search string“

Bramer et al. beschreiben hierfür eine sehr strukturierte Vorgehensweise, die allerdings nicht als prototypisch bezeichnet werden kann:

Für jeden Building Block wird dabei iterativ die Query unter Ausschluss einzelner Freitextterme durchgeführt, beispielsweise (Bramer und Jonge 2015)

*“For each element, but one at a time: Instead of OR between the thesaurus terms type NOT:
(‘kinesiotherapy’/exp NOT (kinesiotherap* OR kinesitherap* OR ((exercis*) NEAR/3 (technique* OR treat* OR therap*))) :ab,ti) AND (‘hip osteoarthritis’/exp OR (coxarth* OR ((cox OR hip*) NEAR/3 (arthrit* OR arthros* OR osteoarth*))) :ab,ti)“*

Die aus der Testsuche resultierenden relevanten Dokumente der Treffermenge werden hinsichtlich zusätzlicher Freitextterme untersucht, die dann als Kandidatenterme aufgenommen werden. Diese Variante separiert Treffer, die ausschließlich über Synonyme in Titel und Abstract gefunden wurden und zielt damit auf Artikel mit ähnlicher Indexierung ab, um daraus zusätzliche fachterminologische Freitextterme zu identifizieren.

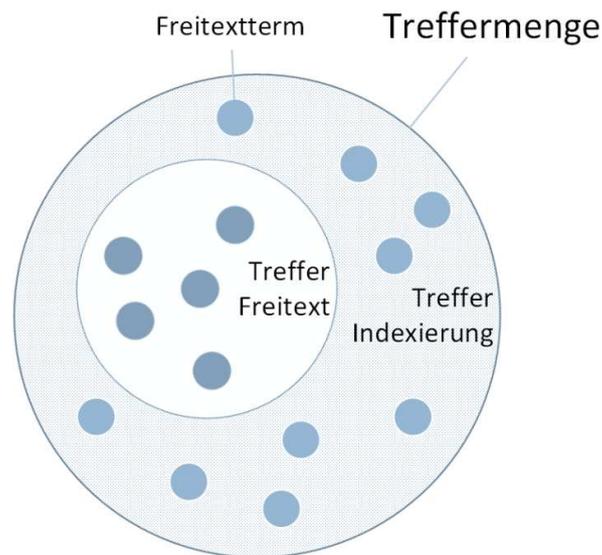


Abbildung 4-6 Thesaurus NOT Free Text Manöver nach (Bramer und Jonge 2015)

Eine alternative Vorgehensweise wird für den Ausschluss einzelner Freitextterme und Artefakte beschrieben, die Analyse relevanter bibliographischer Referenzen zielt in diesem Fall ausschließlich auf das Auffinden zusätzlicher Deskriptoren ab.

Die umgekehrte Variante zielt auf Dokumente ähnlichen Inhalts mit unterschiedlicher Indexierung ab und macht es möglich, die in Abschnitt 4.4.2.2 angesprochene Indexierungsproblematik zu überwinden.

Ausbau der Ground Truth

mittels der Identifikation weiterer hoch relevanter Dokumente zur späteren Analyse korrespondierender bibliographischer Dokumente, (GENERAL 1):

„Ich nutze „Similar articles“ (früher: „Related articles“) in PubMed, um schnell weitere RCTs zu identifizieren. Falls ich fündig werde, analysiere ich diese ebenso.“

Kontrolle der Sensitivität des Suchprofils

mittels der Evaluation der Treffermenge, beispielsweise (GENERAL 1)

„Ich probiere diese Suchstrategie aus (in Ovid), um zu sehen wie viele Treffer von den einzelnen Suchbegriffen erzeugt werden.“

Kontrolle der Noise des Suchprofils

mittels Screening von Treffermengen zur Relevanzanalyse von Titeln und Analyse bibliographischer Referenzen (Qualitätsüberprüfung von Kandidatentermen), siehe (GENERAL 1)

„Ich screene die ersten 10-20 Referenzen der Gesamttreffermenge und schaue mir an, durch welche Begriffe/Schlagwörter irrelevante Treffer gefunden werden.“

Überprüfung von Relevanz oder Semantik

mittels Analyse bibliographischer Referenzen im Hinblick auf Relevanz einzelner Suchterme, beispielsweise aufgrund der Vielfalt in der Verwendung von Fachterminologien, wie im Folgenden beschrieben (SELECTION_TERM 14):

„Awareness that different authors may express words in a specific way, and use different terminology.“

Ein Aspekt ist allen gemeinsam: Sofern der Fokus von Nutzern auf dem gezielten Ausbau der Sensitivität von Suchprofilen liegt, ist an eine Testsuche gleichzeitig die Sammlung weiterer Kandidatenterme gekoppelt, d.h. die Aufnahme neuer Synonyme erfolgt *immer*, mündet dann aber notwendigerweise in vergleichenden Testsuchen, um die neu generierte Suchanfrage zu bewerten.

4.4.4.2 Vergleichende Suchen

Vergleichende Testsuchen dienen in erster Linie der Kontrolle und Bewertung bei Veränderung einzelner Kandidatenterme oder Optimierung von Suchtermen wie auch Building Blocks.

Sie können, wie bereits in der Beobachtungsstudie deutlich wurde, iterativ durchgeführt werden:

“Run a couple of versions of the strategy with varying levels of precision” (BALANCE 7)

Eine Form der vergleichenden Testsuche wird über NOT-Querys durchgeführt, sie dient der Relevanzbeurteilung einzelner Kandidatenterme:

„Vergleich von Treffermengen; arbeiten mit „NOT“, um zu analysieren, welche Referenzen nur anhand eines bestimmten Thesaurus-Begriffs gefunden werden.“ (SELECTION_DESC 1)

„über eine NOT-Kombination ermittle ich dann die zusätzlichen Treffer der breiteren Suche, die der Nutzer ebenfalls sichten kann.

#3 #1 NOT #2 91003

#2 "Kidney Diseases"[Majr]

#1 "Kidney Diseases"[Mesh] 422136“ (GENERAL 3)

Diese Methode ist vor allem bei der Verwendung von Major Terms effektiv und nutzt Indexierungsspezifika einzelner Datenbanken: Bei den mit Major Terms verschlagworteten Dokumenten handelt es sich um eine echte Teilmenge der mit MeSH-Termen verschlagworteten Dokumente, wie die am 08.04.2016 durchgeführte Testsuche mit obigem Beispiel zeigt und die folgende Graphik verdeutlicht. Sie ist aus technischer Sicht eine Variante der oben genannten Manöver zur Identifikation weiterer Freitextterme und kontrollierten Vokabulars, verfolgt jedoch ein anderes Ziel, sie dient der Beurteilung der Mächtigkeit von Treffermengen und kann auch zur Relevanzbeurteilung herangezogen werden, siehe (SELECTION_DESC 1).

Tabelle 4-10 Beispiel: Testsuche über „NOT Manöver“, Durchführung am 08.04.2016

Nr.	Query	Treffer
#6	(#4 NOT #5)	92450
#5	"Kidney Diseases"[Majr]	348152
#4	"Kidney Diseases"[Mesh]	440602

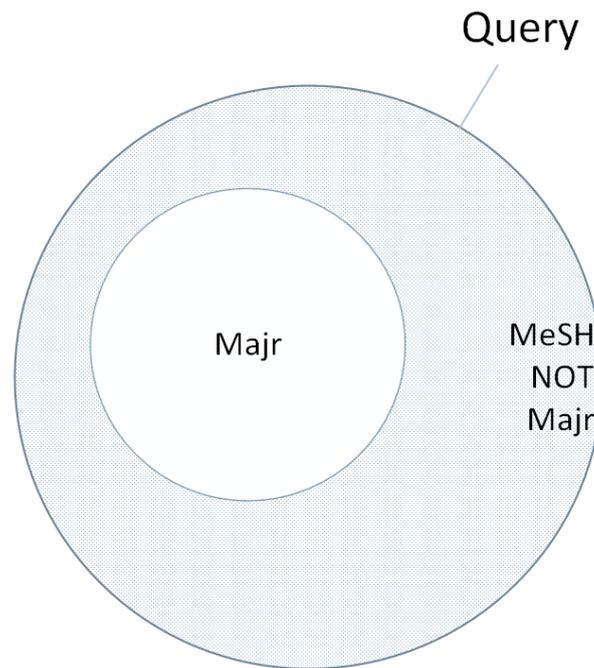


Abbildung 4-7 Graphische Darstellung des NOT Manövers

Kontrolle der Sensitivität von Kandidatermen/Suchtermen/Building Blocks

durch Bewertung der Treffermenge im Hinblick auf Erhöhung der Sensitivität und auf änderungsbedingt fehlende relevante Dokumente

“Then run a search using NOT to see what will be missed if we choose the more precise search.” (BALANCE 7)

“[...] by comparing outputs from different database searches one can get an idea whether we are missing a set of possible studies that may only be retrieved by adding a new synonym” (TERM_CONFID 8)

Kontrolle des Grundrauschens von Kandidatermen/Suchtermen/Building Blocks

mittels Screening von Treffermengen als auch Analyse bibliographischer Referenzen zur Qualitätsüberprüfung von Kandidatermen

Die Kontrolle des Grundrauschens wird vor allem dann nötig, wenn Phrasen und Artefakte verwendet werden. Sie sind nach Auskunft der Befragten häufig aufgrund der Komplexität und der sprachlich unterschiedlichen Verwendung von Phrasen erforderlich (TEST_REASONS 1-3, TEST_REASONS 9, TEST_REASONS 15), beispielsweise bei (TEST_REASONS 1):

„Komplexe Phrasen („Luteinizing hormone relasing hormone analogue“);
wenn die Reihenfolge der Wörter äußerst unterschiedlich vorkommt (weight changes,
changes in weight, changes in body weight, changed body weight).“

Auch in vergleichenden Testsuchen ist der Ausbau der Sammlung an Kandidatentermen häufig Teilaufgabe, d.h. die Aufnahme zufällig entdeckter neuer Synonyme erfolgt immer, mündet dann aber notwendigerweise in weiteren vergleichenden Testsuchen.

Vergleichende Testsuchen zur Konzeption suchbarer Fragestellungen werden von den Befragten auch im Kontext von Building Blocks beschrieben, ein Ansatz, der aus den Anfangszeiten online verfügbarer Datenbanken als *Successive Fractions* bekannt ist, siehe Abschnitt 2.4.4, beispielsweise (GENERAL 11):

“here A and B are clearly defined elements and C is not clearly defined enough.
1 A and B and C (missing relevant hits because C is not well indexed)
2 A and B and (X or Y) (where X and Y might be outcome measures associated with C)
3 A and B and (filter for RCT or other high value study type)
4 1 or 2 or 3
Another approach which will work if the topic is narrow enough would be [(A and B) or
(A and C) or (B and C)]”

Vergleichende Suchen können, wie auch aus der Beobachtungsstudie deutlich wird, entweder direkt auf die Sammlung einzelner Kandidatenterme folgen oder am Ende des gesamten Prozesses stehen, bei dem die Präzisierung datenbankspezifischer Queries im Vordergrund steht. Bei der Reformulierung von Suchtermen und Queries ist die vergleichende Testsuche zwingend erforderlich, siehe hierzu auch den folgenden Abschnitt 4.4.5.2.

Im Ergebnis wurde deutlich, dass der im Fragebogen gewählte Begriff **Testsuche** nicht korrekt gewählt war: Testsuchen werden während der systematischen Übersichtsarbeit von vielen Befragten regelmäßig und iterativ eingesetzt, um Kontrolle über den Entwicklungsprozess zu behalten. Testsuchen führen bei Aufnahme weiterer Kandidatenterme auch bei (aus Vergleichssuchen resultierenden) Änderungen an Suchtermen grundsätzlich zu einer Reformulierung von Suchanfragen.

4.4.5 Reformulierung und Präzisierung

Die beiden Fragen

- Welche Maßnahmen zur Erhöhung der Precision halten Sie im Rahmen systematischer Übersichtsarbeiten für geeignet/ungeeignet? (*Which actions/steps are suitable to increase precision from your point of view?*) und
- Welche Maßnahmen zur Balance von Precision und Recall halten Sie bei systematischen Übersichtsarbeiten für geeignet? (*Which actions/steps are suitable to get the optimal balance between precision and recall?*)

sollten Aufschlüsse darüber geben, welche Maßnahmen in der Praxis verwendet werden und geeignet scheinen.

In der Beobachtungsstudie hatte sich die Frage nach der Handhabbarkeit großer Treffermengen durch die exzessive Sammlung an Suchtermen als problematisch erwiesen und war vor allem durch Intervention der Auftraggeberin wahrscheinlich eher untypisch verlaufen.

4.4.5.1 Qualitativ-orientierte Präzisierung durch Suchfilter und Facetten

Der Einsatz methodologischer Suchfilter ist vor allem bei Informationsspezialisten aus Cochrane Review Groups Usus und dient zur Einschränkung von Treffermengen auf die entsprechend der Evidenzpyramide methodologisch hochwertigen randomisierten klinischen Studien (RCT), siehe Abschnitt 3.1.1. Als Methode wird sie von 8 Befragten als ideale Methode zur Präzisierung von Suchanfragen genannt (Befragte 1, 2, 3, 5, 6, 11, 14, 16), beispielsweise (FILTER_SPEC 1):

„Ich nutze die im Cochrane Handbook publizierten RCT-Filter sowie die von der Gruppe um Haynes RB entwickelten Filter (McMaster University, Hamilton, Kanada), da sie mir als methodisch robust bekannt sind.“

Die inhaltliche Begründung für den Einsatz methodologischer Filter spricht hier klar für die Umsetzung eines methodologischen Qualitätsmerkmals sowie die Anforderung, wirklich keinen relevanten Treffer zu übersehen. Wenn nötig, werden methodologische Suchfilter dabei während des Optimierungsprozesses auch ausgetauscht (siehe BALANCE 1).

„RCT-Filter fast immer bei Therapiestudien. Systematic Review Filter (Montori modifiziert)” (FILTER_SPEC 4)

“Frequently, Cochrane, HIRU, SIGN” (FILTER_SPEC 7)

“RCT, and the pretest-posttest filters (latter a bit unreliable)” (FILTER_SPEC 8)

“I regularly use the Cochrane RCT filter in RCT systematic reviews. I use adaptations and in-house developed methodological search filters for cost-effectiveness and qualitative study searches. I use clinical queries in Ovid – but only for scoping or targeted searches, not for systematic review searches.” (FILTER_SPEC 9)

„Sehr oft; standardmäßig wird bei uns häufig nach RCTs und SR gesucht, daher benutzen wir diese sehr oft. Es werden aber auch andere Studienfilter eingesetzt (z. B. Prognose, Diagnose, Screening, NON RCT, etc.). Es müssen extern validiert[e] Filter sein; die in ausreichender Weise empirisch erstellt wurden (siehe CRD Bewertung von Studienfilter: <https://sites.google.com/a/york.ac.uk/issg-search-filters-resource/>).“ (FILTER 13)

“I use the Glanville RCT filter for RCTs and a modified version of the Haynes review filter to find SRs. [...] a modified version of the SIGN guidelines filters for observational studies and I have one I created for case studies.” (FILTER_SPEC 16)

Wie aus den Aussagen der Befragungen ersichtlich, werden vor allem methodologisch hochwertige, evaluierte Filter und deren Modifikationen genutzt. Dies deckt sich mit den Ergebnissen einer Studie von Beale et al. Sie stellten fest, dass für die Auswahl eines Filter neben individuellem Kenntnisstand über den Einsatz dieser Filter, deren Herkunft, Validität als auch die Möglichkeit flexibler Anpassungen ausschlaggebend sind (Beale et al. 2014).

Warum werden methodologische Filter nicht eingesetzt? Darüber kann die Fragebogenstudie keine finale Auskunft geben. Ein Teilnehmer spricht jedoch konkret das Vertrauen in Suchfilter an:

“Very sparcely (even the once I created myself I hardly use). Only the HSSS from cochrane I use more frequently. I never take a filter for granted always add some of my own words to make it more sensitive. Often combining study type filters with worde describing an outcome of interest.“ (FILTER 6),

konkretisiert in (Bramer und Jonge 2015):

“Advise: never trust filters, even when they are validated. They can always be improved (see the Cochrane HSSS which does not contain randomization)”

Ein anderer Befragter begründet den Nichteinsatz durch mangelnde diesbezügliche Anforderungen durch Auftraggeber (FILTER 15).

Methodologische Suchfilter sind demnach weiterhin ein Spezifikum stark methodologisch geprägter systematischer Übersichtsarbeiten. Sie werden vorwiegend von institutionell in der EBM verankerten Informationsspezialisten verwendet, während anderswo auch eigene, weniger gut validierte Filtermethoden angewandt werden. In diesem Fall lassen die Befragten jedoch ausschließlich eine Einschränkung auf den Studientyp und das Studiendesign (beispielsweise Studien an Tieren oder Patienten) zu: „ggf. Eingrenzen der Treffermenge der Freitextsuche über Filter wenn Filter hier nicht einsetzbar sind (human, male / female, Altersgruppen) grenze ich über entsprechende Freitextbegriffe ein“ (GENERAL 3).

Die Eignung von Filtern [Anmerkung: es ist auch bei Befragter 3 davon auszugehen, dass es sich um methodologische Filter handelt] ist abhängig von suchbaren Fragestellungen. Sind sie nicht verfügbar, so erfordert dies anderweitige Maßnahmen zur Qualitätssicherung und Optimierung, beispielsweise die Konkretisierung von Altersgruppen.

Dies ist insbesondere deshalb bemerkenswert, weil nicht-institutionell in der EBM verankerte Befragte den Einsatz datenbankspezifischer Filtermechanismen, klassischerweise die Facetten-
suche (siehe Abschnitt 3.4.2.3), durchaus in Erwägung ziehen, um Suchanfragen mit erhöhtem Grundrauschen zu präzisieren und damit Treffermengen handhabbarer zu machen, beispielsweise (FILTER 3):

- Bei großen Treffermengen und nach Vorgabe des Nutzers nutze ich häufig die Filter
- Publikationstyp zusätzlich bzw. für aktuelle Publikationen auch über Freitextsuche eingrenzen
- Publikationszeitraum
- Sprache der Artikel
- human für aktuelle Publikationen über Freitextsuche eingrenzen
- Geschlecht für aktuelle Publikationen über Freitextsuche eingrenzen
- Altersgruppen für aktuelle Publikationen über Freitextsuche eingrenzen

und bei (GENERAL 2):

„Je nach Anfrage werden die Treffermengen durch das Setzen von Filtern eingeschränkt, z.B. Publikationsart, Altersgruppen, Veröffentlichungszeitraum, Originalsprache.“

Facettenfilter werden hier ebenso wie methodologische Filter als Maßnahmen zur Qualitätssicherung von Queries beschrieben, sie sind aus methodologischer Sicht jedoch nur auf Basis fachlicher Begründung einsetzbar. Die Beschränkung auf das Publikationsdatum beispielsweise ist nach Einschätzung einiger Befragter und konform mit der Methodenlehre lediglich dann zulässig, wenn es sich um die Aktualisierung eines Reviews handelt oder der Forschungsgegenstand vor einem bekannten Datum nicht bekannt war (vgl. BALANCE 1 sowie (Bramer und Jonge 2015)). Die Auswahl geeigneter Filtermethoden liegt zusätzlich in einem Spannungsfeld zwischen Methodenlehre sowie Anforderungen des Auftraggebers (siehe FILTER 3 sowie FILTER 15).

4.4.5.2 Quantitativ-orientierte Präzisierung durch Reformulierung

Die bereits mehrfach angesprochenen Arbeiten von Bates und Fidel, siehe Abschnitte 2.4.2 und 2.4.3 benennen verschiedene Methoden zur Reduzierung großer Treffermengen. Im Folgenden werden sie Präzisierungsmethoden im Sinne einer optimalen Balance zwischen vollständigem Recall (gemessen an der Ground Truth) und Precision genannt, da die Manöver zur Reduktion des Grundrauschens immer davon geprägt sind, den Vollständigkeitsanspruch in der systematischen Literaturrecherche zu garantieren. Die ist, wie im Folgenden dargestellt, auch der Hauptgrund für kritische Beurteilung einzelner Methoden, beispielsweise die Verwendung des NOT Operators. Der Begriff Optimierung, wie von einer Befragten (7) genannt, wird in der Fachwelt abgrenzend auch für den exzessiven Ausbau von Suchprofilen verwendet, vgl. (Bramer und Jonge 2015).

Vor allem unspezifische Synonyme resultieren in einfachen Testsuchen häufig in hohem Grundrauschen. Für den Umgang mit unspezifischen, weil zu allgemein gehaltenen Synonymen gilt häufig eine Vermeidungsstrategie (u.a. UNSPEC_SYN 1, UNSPEC_SYN 5, UNSPEC_SYN 6, UNSPEC_SYN 10), vgl. hierzu auch Tabelle 4-11.

Da das Retrieval bekannter Schlüsselreferenzen und die Forderung nach einem möglichst sensiblen Suchprofil immer Vorrang gegenüber einer Spezifizierung des Suchprofils haben (die ja die Gefahr des Verlusts relevanter Dokumente in sich birgt) (u.a. UNSPEC_SYN_CRITERIA 6,

UNSPEC_SYN_CRITERIA 8, UNSPEC_SYN_CRITERIA 11), sind aufgrund von Testsuchen gegebenenfalls Präziserungsmaßnahmen nötig, um das Grundrauschen der Treffermenge möglichst gering halten zu können.

Zur rein quantitativ orientierten Präzisierung von Suchanfragen kommen in der Praxis sehr spezifische Methoden zum Einsatz, die in Tabelle 4-11 zusammengefasst sind und im Folgenden erläutert werden. Sie decken sich weitestgehend mit den bei Bramer und de Jonge beschriebenen Methoden, siehe (Bramer und Jonge 2015) bzw. (PRECISION 6).

Tabelle 4-11 Methoden zur Spezifizierung von Suchanfragen, siehe auch (Bramer und Jonge 2015)

Methode
Entfernung von Freitexttermen
Austausch von Suchtermen (unspezifisch, zu spezifisch)
Einsatz von Filter nicht näher spezifiziert
Einsatz methodologischer Suchfilterfilter
Verwendung NOT Operator
Spezifizierung von Deskriptoren
Spezifizierung von Freitexttermen
Optimierung von Building Blocks, auch einzeln
Reduktion von Informationsquellen

Spezifizierung von Freitexttermen und Phrasen

Neben dem Einsatz methodologischer Filter und Facettenfilter dominiert in den Beschreibungen der Befragten vor allem die Methode der Spezifizierung von Freitexttermen, die beispielsweise aus der Notwendigkeit resultieren kann, mangels Alternativen allgemeingehaltene Begriffe oder Phrasen in das Suchprofil aufnehmen zu müssen.

Hierfür werden verschiedene Möglichkeiten genannt, unter anderem die Bildung von Phrasen, die Entfernung von Trunkierungen (gleichzusetzen mit der Vergrößerung von Wortstämmen) und die Verknüpfung unspezifischer mit spezifischen Termen mittels Proximity-Operatoren sowie mit AND (UNSPEC_SYN 1, UNSPEC_SYN 6, UNSPEC_SYN 14, UNSPEC_SYN 15):

„Häufig kann man unspezifische Synonyme [...] in eine spezifische Wortkombination umwandeln. [...] funktioniert mit Adjacency-Operatoren [...] besser, da das unspezifische Synonym dann in der Nähe eines anderen Terms auftauchen muss.“ (UNSPEC_SYN 1)

“If it is not very specific – perhaps it can be ANDed [...] with another word.” (UNSPEC_SYN 14)

Die Modifikation von Phrasen und Artefakten ist der am häufigsten genannte Grund für vergleichende Testsuchen.

Die Beschränkung von Suchtermen auf das Vorkommen in vordefinierten Feldern bibliographischer Datenbanken (meist sind dies Titel oder Abstract) (siehe hierzu Abschnitt 3.4.2.2) ist ebenfalls eine Methode, Suchanfragen zu spezifizieren und wird hierfür auch verwendet (BALANCE 4, PRECISION 5, PRECISION 6, PRECISION 14, PRECISION 16).

Sie wird jedoch nicht durchgängig als ideale Methode beschrieben: „Nur ungern wende ich eine Einschränkung von Termen auf das Vorkommen im Titel oder eine Gewichtung von Schlagwörtern an.“ (UNSPEC_SYN_CRITERIA 15), vgl. auch BALANCE 4: „Nicht geeignet: Suche im Titel für höhere Precision. Gefahr des Verlustes zu groß“.

Die Verwendung des NOT Operators wird vorwiegend kritisch beurteilt, er sollte „nur in Ausnahmefällen“ (BALANCE 1) und mit Vorsicht (PRECISION 14) eingesetzt werden und zur Evaluierung der Vollständigkeit immer an eine abschließende, vergleichende Testsuche gekoppelt werden (BALANCE 1, BALANCE 11, PRECISION 14).

Spezifizierung von kontrolliertem Vokabular

Methodologisch wird beim Einsatz von kontrolliertem Vokabular in Suchanfragen die Nutzung der Explode Funktion empfohlen, siehe hierzu Abschnitt 3.4.2.1. Eine Empfehlung zur Reduktion von Grundrauschen lautet: „*Maybe narrower terms from the thesaurus cause noise. See if using a no explode function solves this problem*“ (Bramer und Jonge 2015).

Ein Spezifikum bibliographischer Suchen auf medizinischen Fachdatenbanken ist der Einsatz so genannter *floating subheadings* (OVID Terminologie) oder auch *free floating subheadings* (Pubmed Terminologie).

Subheadings wurden in den Abschnitten 2.3.2.2 sowie 3.4.2.1 als so genannte Qualifier beschrieben, mit deren Hilfe kontrolliertes Vokabular kontextuell eingeschränkt werden kann, beispielsweise „`thromboembolism/prevention and control [mh]`“, abgekürzt auch „`thromboembolism/pc [mh]`“, vgl. (National Center for Biotechnology Information (US) 2016). Sie können in der Suche auch als einzelne Suchterme eingesetzt werden, beispielsweise: „`prevention and control[sh]`“, führen in der Praxis aber häufig zu erhöhtem Grundrauschen.

Bei der Nutzung so genannter *floating subheadings* werden die Qualifier nicht direkt mit dem kontrollierten Vokabular gekoppelt, sondern über die Spezifikation als *subheadings* mittels Operatoren mit kontrollierten Vokabular verbunden, z.B. `breast neoplasms[mh] AND trends[sh]` als Alternative zu `breast neoplasms/trends[mh]`.

Beschrieben wird diese Methode lediglich bei einem Befragten (Nr. 6, vgl. auch (Bramer und Jonge 2015)) in Kombination mit AND- und PROXIMITY-Operatoren:

“Floating subheadings can retrieve irrelevant results. This can be reduced by combining thesaurus terms with subheadings. When this method is used, it is wise to also use the next method, combining the free text words of these elements to phrases or proximity instead of AND.”

Spezifizierung von Building Blocks

Die Präzisierung von Suchanfragen kann alternativ auch durch die Modifikation oder die Hinzunahme von Building Blocks erfolgen (vgl. BALANCE 6, PRECISION 6, PRECISION 8, GENERAL 11, UNSPEC_SYN 11 (Bramer und Jonge 2015)). Diese Methode birgt allerdings die Gefahr, relevante Treffer zu verlieren und wird daher als zeitintensive Methode beschrieben (Bramer und Jonge 2015), siehe hierzu auch die vergleichenden Testsuchen in Abschnitt 4.4.4.2.

„Ich hatte schon wiederholt den Fall, dass die Fragestellung konkretisiert oder eingeschränkt werden musste. Nach meinem Eindruck lag dies aber nicht primär an einer grundsätzlich ungeeigneten Suchstrategie, sondern daran, dass die Domänenexperten die Mengen und Bandbreite der Literatur im Themenfeld unterschätzt hatten.“ (UNSPEC_SYN_CRITERIA 15)

4.4.6 Nutzung von Informationsressourcen

1:1 mit Schlüsselbegriffen einzelner Building Blocks korrespondierende Deskriptoren stellen per se den idealen Kandidatenterm strukturierter Suchprozesse dar.

Auch die vom Domänenexperten genannten Synonyme und Fachtermini als für die Fragestellung sehr charakteristische Kandidatenterme fließen als Synonyme direkt in die Liste an Kandidatentermen ein (siehe auch Abschnitt 4.4.1.1). Darüber hinaus ließen sich über Beobachtungsstudie und Interviews ressourcenspezifische Strategien in der Entwicklung von Suchprofilen identifizieren.

4.4.6.1 Ground Truth

Der Einstieg über die Ground Truth ist die vorherrschende Taktik zur Identifikation von Suchtermen. Sie wird insbesondere von den Experten, die bereits in der Vorbereitungsphase die Beschaffung oder Bereitstellung von Schlüsselreferenzen durch den Auftraggeber als notwendige, vorbereitende Maßnahme erwähnen, als sehr strukturierter Prozess beschrieben (beispielsweise (ORDER 15):

„Habe ich bspw. initial einige relevante Treffer, Literaturstellen oder Volltexte, gehe ich diese jeweils komplett auf der Suche nach Termen durch, mit einem Schwerpunkt auf Titel und Abstract.“

Das Prinzip des Retrieval-gesteuerten Ground Truth-Loop wurde in Abschnitt 4.2.4.3 graphisch dargestellt und beschrieben. In der Methodenlehre der systematischen Übersichtsarbeit wird sie vor allem zur Qualitätsbewertung von Queries genutzt, siehe hierzu Abschnitt 3.4.4.2.

Ground Truth-Loops

Bei Fokussierung auf die Ground Truth werden alle verfügbaren Schlüsselreferenzen systematisch ausgeschöpft. In diesem Prozess lassen sich zwei Ziele festmachen, welche die Ergebnisse der Beobachtungsstudie bestätigen:

- Sicherung des Retrievals aller Ground Truth Dokumente durch Aufnahme passender Synonyme für jeden einzelnen Building Block aus Titel, Abstract und Indexierung der bibliographischen Referenzen
- Ausbau der Menge an Kandidaterme durch weitere Synonyme aus Titel, Abstract und Indexierung

Da die Ground Truth als die ideale Referenzmenge beschrieben wird, werden Suchterme, Deskriptoren wie Freitextterme, teilweise ausschließlich deshalb als Kandidaterme aufgenommen, um im finalen Suchprofil das Retrieval dieser Dokumente zu garantieren – man möchte mit dem eigenen Suchprofil gleichzeitig die ideale Referenzmenge abdecken (SELECTION_TERM 1, SELECTION_TERM 7, SELECTION_TERM 9) und hat damit ein Bewertungskriterium für Suchprofile geschaffen. Um das Retrieval einzelner Dokumente zu sichern, ist es technisch unerlässlich, für jeden der Building Blocks, die später über den Booleschen AND Operator verknüpft werden, mindestens einen Freitextterm oder Deskriptor in das Suchprofil zu übernehmen, sofern nicht schon dort enthalten.

Was sich bereits in der Beobachtungsstudie gut prozessorientiert darstellen ließ, wird mit den Ergebnissen der Fragebogenstudie nun untermauert.

Zielt die Analyse von Dokumenten auf das Retrieval hin – was bei den Befragten wie bekannt häufig als expliziter Grund für die Aufnahme einzelner Kandidaterme genannt wurde – werden dabei alle verfügbaren bibliographischen Referenzen gezielt sprachlich-inhaltlich „analysiert“ (genannt in GENERAL 1, GENERAL 3, TYP5, GENERAL 7, GENERAL 14, ORDER 15), pro Building Block im Hinblick auf Deskriptoren aus der Indexierung sowie Freitextterme aus Titel und Abstract geprüft (GENERAL 3, GENERAL 5), welche als Kandidaterme notiert werden (GENERAL 1, GENERAL 14). Die sprachlich-inhaltliche Analyse kann als *Abgleich* von Dokumentenprofil und Suchprofil interpretiert werden (SELECTION_TERM 1, GENERAL 5, ORDER 9). Besonders deutlich wird der iterative Loop in der Beobachtungsstudie, in der die Versuchsperson systematisch das Dokument im Hinblick auf jeden einzelnen Building Block „abhakte“, ggf. Suchterme im Suchprofil notierte und nach Abschluss zum nächsten Dokument der Ground Truth wechselte.

Ist das Dokumentenprofil durch das aktuelle Suchprofil vollständig repräsentiert, weil zu allen Building Blocks bereits Suchterme im Suchprofil enthalten sind, ist der weitere Abgleich opti-

onal, bei dem aus Ground Truth Dokumenten zusätzlich weitere potenziell passende Kandidenterme erfasst, die sich in Titel, Abstract oder der Indexierung bibliographischer Referenzen identifizieren lassen (GENERAL 1, ORDER 9). Wie sich in der Beobachtungsstudie (Session 1) zeigte, vermischen diese beiden Vorgänge in der Praxis wahrscheinlich: Wird ein Dokument hinsichtlich eines spezifischen Keywords abgeglichen, so wird nicht nur ein exakt passendes Synonym in das Suchprofil aufgenommen, sondern gegebenenfalls auch mehrere, exakt oder partiell übereinstimmende Synonyme.

Relevante bibliographische Referenzen werden zunächst in ihrer Funktion als Quelle für Suchterme beschrieben und scheinen wichtige Anforderung für die Entwicklung von Suchprofilen. Relevante Artikel der Ground Truth (und damit ihre bibliographischen Referenzen) stehen im Kontext der Fragestellung allerdings nicht isoliert, sondern haben als ideale Repräsentanten inhaltliche und konzeptionelle Gemeinsamkeiten. Die thematische Einarbeitung kann daher über die Erkennung gemeinsamer Konzepte bibliographischer Referenzen erfolgen (GENERAL 7).

Struktur genutzter Informationseinheiten

Frage 3 des Interviews zielte darauf ab, zu erfassen, ob während der Analyse bibliographischer Referenzen oder Volltextdokumente die Lokalisierung von Synonymen im Titel, Abstract, dem Volltext Corpus oder in der Verschlagwortung wahrgenommen und besonders berücksichtigt wird und welche Dokumentstrukturen während der Verarbeitung von besonderem Interesse sind: *Do you prefer a specific order when processing relevant citations or documents? (eg. collect descriptors first, collect terms in the order of building blocks, scan TI- AB- TW – bibliography)* – die genutzten Abkürzungen TI, AB, TW waren der Ovid-typischen Syntax der MEDLINE Suche geschuldet und dienen in Anlehnung an Qualifier lediglich der Abkürzung für die Begriffe Titel, Abstract und Textkörper- sofern es sich um Volltexte handelt).

Bibliographische Angaben werden dabei zwar als strukturierte Dokumente wahrgenommen, aber nicht durchgängig in ihrer strukturierten Form genutzt (“I tend to take the paper as a whole – look at subject headings and any keywords at the same time” (ORDER 16)).

Die bibliographische Referenz ist in der Praxis in erster Linie eine wichtige Quelle für Deskriptoren (GENERAL 1, GENERAL 2, GENERAL 3, GENERAL 7, GENERAL 8).

Viele Interviewpartner berichten durchgängig von der gezielten Identifikation von Synonymen

im Titel und im Abstract bibliographischer Angaben (GENERAL 1, GENERAL 3, GENERAL 4, GENERAL 8, GENERAL 9), eine Fokussierung auf Terme aus Titel oder alternativ Abstract ist nicht eindeutig erkennbar, mehr noch strittig. Eine Reihe der Experten bevorzugt bei der Sammlung Freitextterme aus dem Titel, andere betonen, es sei gerade nicht so, dass die wirklich relevanten Synonyme bereits im Titel vorkäme (SELECTION_TERM 5), siehe hierzu auch Abschnitt 4.4.3.1.

Die Struktur verfügbarer Informationseinheiten spielt, so lässt sich zusammenfassend feststellen, nur dann eine Rolle, wenn es sich um bibliographische Referenzen handelt. Für die Boolesche Suche sind technisch bedingt ausschließlich Suchterme von Relevanz, die in bibliographischen Dokumenten in Titel, Abstract oder als kontrolliertes Vokabular vorkommen. Innerhalb der Ground Truth als Pool relevanter Zitationen ist die Analyse zugehöriger Volltexte daher zweitrangig, siehe auch Abschnitt 4.4.3.1. Über eine individuelle Bevorzugung von Termen aus Titel oder Abstract lassen weder Beobachtungsstudie noch Fragebogenstudie eindeutige Aussagen zu, siehe hierzu ebenso Abschnitt 4.4.3.1.

Klassische Interaktionsstrategien im Kontext Ground Truth

Dokumente der Ground Truth sind nicht nur Quelle für Suchterme, sondern auch Quelle für weitere relevante Dokumente. Sie werden dahingehend auch genutzt, um die Ground Truth als Referenzmenge auszubauen und gleichzeitig weitere Synonyme zu identifizieren.

Aus den Ergebnissen der Fragebogenstudie lassen sich dahingehend klassische Pearl Growing-Strategien identifizieren (Interviewpartner 1,3-5), wie sie als Interaktionsstrategie bereits von Arksey und O'Malley beschrieben wurden, siehe Abschnitt 2.4.1.2.

„Ich nutze „Similar articles“ (früher: „Related articles“) in PubMed, um schnell weitere RCTs zu identifizieren. Falls ich fündig werde, analysiere ich diese ebenso [...]“ (GENERAL 1)

Für den Informationsexperten ist dies „wichtig, um schnell weitere relevante Referenzen zu finden (und deren Schlagworte + Freitextbegriffe zu analysieren)“ (TACTICS_CPG 1), zum „Auffinden von Dokumenten und ggf. Freitexttermen“ (TACTICS_CPG 3) und vor allem „sehr wichtig für den Einstieg in die Suche, s.o. Auffinden von Suchbegriffen, Deskriptoren und Kontrolle der Suchergebnisse und Dokumentation der Suche“ (TACTICS_CPG 5).

Die bei Bates beschriebenen Berry Picking-Strategien (siehe Abschnitt 2.1.2), darunter insbesondere bibliographische Suchen und Zitationssuchen sind häufig nicht Aufgabe des Informationsspezialisten, sondern werden vielmehr an den Auftraggeber delegiert:

„erledigen unsere Autoren, nicht ich“ (TACTICS_CS 1)

„Zusätzlich können Suchen mit mehr manuellen Schritten sinnvoll sein, wobei ggf. direkt eine Relevanzbeurteilung durch die Domänenexperten enthalten ist. Beispiele sind die Durchsicht von Literaturverzeichnissen relevanter Artikel [...], das Sichten zitierender Arbeiten von mehr als nur einzelnen Artikeln, das Sichten von Related Citations/Similar Articles in PubMed [...]. Dies kann ich oft nicht selbst leisten, sondern weise die Domänenexperten auf diese Möglichkeiten hin [...].“ (GENERAL 15)

Stellenwert der systematischen Übersichtsarbeit

Besonderen Stellenwert haben in der systematischen Übersichtsarbeit aus methodologischer Sicht vorbestehende systematische Übersichtsarbeiten oder solche zu ähnlichen Themen. Sie fungieren als Evidenzquelle (siehe Abschnitt 3.1.1) und sind wertvolle Quelle sowohl für Suchterme als auch ganze Suchprofile, sowie weitere bibliographische Referenzen (dies sind die in den Review eingeschlossenen Primärveröffentlichungen zu randomisierten klinischen Studien).

Diese Vielfalt an weiterführender Information ist auch der am häufigsten und teils auch ausschließlich genannte Grund für deren Beschaffung als Volltexte (IDENT_FULL 14, IDENT_FULL 15, IDENT_FULL 16). An dieser Stelle muss hinzugefügt werden, dass die Beschaffung systematischer Übersichtsarbeiten vor allem in der Cochrane Library für lizensierende Institutionen problemlos möglich ist, da die Inhalte nicht nur als PDF, sondern auch im HTML Format als Volltext abrufbar sind.

Sowohl im Cochrane Handbuch (Higgins und Green 2011), dem MECIR Statement (Chandler et al. 2013) und anderen Handreichungen wird die Nutzung bereits bestehender systematischer Übersichtsarbeiten mindestens empfohlen. Als qualitativ hochwertige wissenschaftliche Arbeit ist die systematische Übersichtsarbeit nicht nur Quelle für den Ausbau einer Ground Truth, sondern liefert mit den zumindest für Cochrane Reviews verpflichtend veröffentlichten Suchstrategien gleichzeitig eine Vorlage für die neu zu entwickelnden Suchprofile.

Obwohl von den Interviewpartnern 7 in einer Cochrane Review Group tätig sind oder tätig waren, wird die Anforderung systematischer Übersichtsarbeiten nicht explizit als erforderliche Aufgabe in der Vorbereitung des Reviews genannt. Wie bereits an anderer Stelle erwähnt, werden mit Fachartikeln zu randomisierten kontrollierten Studien zwar beispielsweise methodisch hochwertige Forschungsergebnisse angefordert (GENERAL 1), meist aber ist lediglich von relevanten Schlüsselreferenzen die Rede, die dann in Summe die Ground Truth bilden. Nur eine Interviewpartnerin weist darauf hin, dass sie zunächst in einer breit angelegten Vorabsuche systematische Übersichtsarbeiten zu lokalisieren versucht (GENERAL 12). Es bleibt offen, ob die Anforderungen des MECIR Statements so verinnerlicht sind, dass den Interviewpartnern die Erwähnung der Suche nach systematischen Übersichtsarbeiten nicht erforderlich scheint.

Systematische Übersichtsarbeiten werden wie andere Dokumente, allerdings meist über vorliegende Volltexte zur Identifikation von Kandidatermen genutzt (IDENT_FULL 15, IDENT_FULL 16), insbesondere werden Suchterme, die bereits in anderen Übersichtsarbeiten genutzt werden, als Kandidaterme aufgenommen (SELECTION_TERM 12), ebenso solche, die von den Informationsspezialisten bereits in anderen, thematisch verwandten Reviews verwendet werden (TERM_CONFID 8). Die Übernahme kompletter Suchprofile wird allerdings von keinem Teilnehmer der Studie erwähnt, mehr noch lehnt dies ein Befragter strikt ab (SELECTION_TERM 6):

„[...] never use search strategies from other reviews. I hardly reuse earlier search strategies, as in my opinion every search is new. Only if a researcher returns and says: I want exactly that, but now with one difference I reuse part of an old strategy“

Es bleibt anhand der Ergebnisse weitestgehend unklar, ob bei der Analyse systematischer Übersichtsarbeiten mit dem Ziel der Identifikation von Synonymen der Gesamttext, die dort veröffentlichte Suchstrategie oder auch die Durchsicht der in den Review eingeschlossenen Studien beziehungsweise deren Titel im Vordergrund stehen.

4.4.6.2 Thesauri

Neben der Ground Truth sind Thesauri zu durchsuchender bibliographischer Fachdatenbanken ein weiterer wichtiger Einstiegspunkt in die Entwicklung von Suchstrategien.

Die Struktur der für medizinische systematische Literaturrecherchen wichtigsten Thesauri MeSH und Emtree wurde in den Abschnitten 2.3.2.2 als auch 3.4.2.1 eingehend beschrieben.

Thesauri einzelner Datenbanken, im Wesentlichen MeSH als Thesaurus der am häufigsten durchsuchten Datenbank MEDLINE, sind für die Experten in der Praxis nicht nur Quelle für Deskriptoren, sondern auch für Thesaurus spezifische Synonyme, wie beispielsweise die MeSH *Entry Terms* (siehe Abschnitt 3.4.2.1) (GENERAL 3, GENERAL 5, GENERAL 6, UNSPEC_SYN 2) oder sonstige Freitextterme, die beispielsweise in der Beschreibung des kontrollierten Vokabulars verwendet werden.

Der Einstieg in den Thesaurus und der Zugriff auf passende Deskriptoren erfolgen in der Praxis entweder direkt über verfügbare Webinterfaces (GENERAL 2, GENERAL 3, GENERAL 11) oder aber über eine einfache Keyword Suche und Analyse des „*automatic term mapping*“ der Datenbank Pubmed (GENERAL 4, GENERAL 5). Nicht in den Interviews genannt, aber ebenso verfügbar ist eine analoge Funktionalität für die Datenbank OVID. Es ist zu vermuten, dass die Befragten zum Einstieg in den Thesaurus häufig die einfache Keyword-Suche nutzen und dabei auf die sowohl in Pubmed als auch Ovid verfügbare Funktionalität zurückgreifen, Freitextterme mit zugehörigen Deskriptoren zu mappen (siehe Abschnitt 3.4.2.3).

Fokussieren sich Interviewpartner zunächst auf die Nutzung von Thesauri, werden die Building Blocks des Suchprofils im ersten Schritt gezielt ausschließlich mit geeigneten Deskriptoren aufgefüllt und danach um Freitextterme zum kontrollierten Vokabular ergänzt (erwähnt beispielsweise bei GENERAL 2, GENERAL 9, GENERAL 11). Die Freitextterme entsprechen dabei entweder 1:1 dem kontrollierten Vokabular (als Einzelwortterm oder Phrase) oder werden dem Beschreibungstext beziehungsweise korrespondierenden *Entry Terms* entnommen (GENERAL 3, SELECTION_DESC 5).

Anschließend werden die so erstellten Suchprofile auf Basis durchgeführter Testsuchen (siehe Abschnitt 4.4.4.1) erweitert.

Ein idealer Einstiegspunkt über den Thesaurus scheint für Experten gegeben, wenn sich Schlüsselbegriffe einzelner Building Blocks 1:1 mit einem im Datenbank-Thesaurus verfügbaren Deskriptor decken:

„I examine the indexing (MESH, EMTREE) to ascertain whether the database does or does not have a precise term for that dimension, and if it does, start with that.” (GENERAL 11).

“It depends on the search. If it is a specific area that could have discrete controlled terms that are clear then I would start with controlled terms. Where it is an area that is less well defined I might start with free-text. [...]” (ORDER 14)

Auch beim Einstieg über die Indexierung spielt die Ground Truth - soweit verfügbar - eine Rolle, in diesem Fall allerdings vorwiegend als Quelle für die Terme, für die vermutet wird, es gäbe hierfür geeignete Deskriptoren:

„Einstieg: Suchbegriffe im Titel -> Verschlagwortung ansehen zur Orientierung, welche zutreffend sind -> mit Schlagwörtern weitersuchen, mit Textwörtern ergänzen.“ (GENERAL 4).

4.4.6.3 Nutzerpräferenzen

„Collect Sense“: Spezifitätsorientierte Nutzertypen

Wenige Nutzertypen (insbesondere Nr. 8 und 11, auch 5) beschreiben eine eher spezifitätsorientierte Vorgehensweise, die sich zentral an der Indexierung bibliographischer Dokumente in den einzelnen Fachdatenbanken orientiert und hierfür vor allem Thesaurus-Funktionalitäten nutzt, was sich bei einer Befragten sehr deutlich zeigt:

“I do not follow the librarian preferred procedure of constructing a massive search strategy at the outset. I try to keep it simple and work in a spiral fashion so that I know at the beginning which terms are likely to be problematic in trying to produce an output that is relevant, meaningful – and manageable.” (GENERAL 8)

Die Gründe hierfür liegen nicht in äußeren methodologischen Anforderungen: “I work in health services research, which makes construction of search strategies a little more challenging than for clinical research.” – Systematische Übersichtsarbeiten mit dem Ziel wissenschaftlicher Beurteilungen im Gesundheitswesen machen nach Angaben der Befragten die Entwicklung von Suchprofilen zu einer etwas anspruchsvolleren Aufgabe als in der klinischen Forschung (GENERAL 8).

Das Vorgehen ist geprägt von einem erhöhten Informationsbedarf hinsichtlich der Verwendung kontrollierten Vokabulars sowie der Indexierungsqualität, siehe hierzu auch Abschnitt 4.4.2.2:

“Always check what some subject terms from the thesaurus from the database do [...] – to check on indexer interpretation of scope notes [...]“ (GENERAL 8)

und damit von einer Vielzahl iterativer Testsuchen zum Zweck der Validierung (“to check whether the articles really are being indexed the way I expect them to be” (IDENT_BIB 11)).

In diesem Fall rückt die Präzisierung von Suchanfragen während der Entwicklung von Suchprofilen als Aufgabe eher in den Hintergrund (PRECISION 8), wahrscheinlich da die Treffermengen aufgrund einer weniger stark wachsenden Menge vorab bereits validierter Suchterme proportional weniger schnell steigt.

Nutzer, die diese Vorgehensweise beschreiben, arbeiten häufig als klinische Bibliothekare oder fühlen sich institutionell keiner Methodenlehre verpflichtet. Gleichzeitig berichten sie im Hinblick auf ihre eigene Expertise im Gegensatz zum Großteil der Befragten ausschließlich von persönlicher und langjähriger Erfahrung (Teilnehmer 8 und 11).

„Collect all you can get“: Ground Truth für sensitivitätsorientierte Nutzer

Die vorwiegende Zahl der Informationsspezialisten bevorzugt zum Einstieg in den Entwicklungsprozess eine Menge relevanter Beispielreferenzen, die so genannte Ground Truth wie in Abschnitt 4.4.6.1 beschrieben. Ihre Beschaffung ist, wie in Abschnitt 4.4.1.1 dargestellt, häufig Aufgabe des Auftraggebers.

Die „Collect all you can get“ Strategie nutzt mit der Ground Truth und Fokussierung auf bibliographische Referenzen vorrangig Informationsressourcen mit hohem Ertrag (kontrolliertes Vokabular, Freitextterme, Phrasen, bibliographische Literaturangaben sowie „similar articles“). Sie ist damit geprägt durch ein höchst effizientes, gleichzeitig aber auch strukturiertes Vorgehen (beispielsweise die „Ground Truth-Loop“).

Die so beschriebene Strategie ist ausgehend von der in den Abschnitten 3.4.3ff. dargestellten Methodenlehre stark methodologisch geprägt, sie wird vorwiegend von Befragten beschrieben, die institutionell in entsprechenden Einrichtungen verankert sind und entsprechenden Richtlinien unterliegen oder solchen, die sich bei ihrer Vorgehensweise auf Schulungsangebote entsprechender Institutionen berufen (PERSONAL_STYLE 7, PERSONAL_STYLE 9, PERSONAL_STYLE 15, PERSONAL_STYLE 16). Sie folgt entsprechend methodologischer Anforderungen dem Ziel eines umfassenden und erschöpfenden Suchprofils und führt im Allgemeinen zu großzügig angelegten Sammlungen an Kandidatentermen (siehe u.a. die Erwähnung bei den Teilnehmern 6, 11, 15) wird aber auch durch den Wunsch möglichst objektiv generierter Suchprofile begründet: "Using synonyms helps make search less subjective (slightly). I would rather use lots of synonyms provided this doesn't generate too much irrelevant additional items." (SELECTION_TERM 14)

In der Praxis führt diese Vorgehensweise häufig zu Problemen bei der Beurteilung der Vollständigkeit von Suchprofilen (u.a. PROBLEMS 5, PROBLEMS 15), dem Umgang mit der Terminologie komplexer Interventionen (u.a. PROBLEMS 1, PROBLEMS 3, PROBLEMS 4), PROBLEMS 14), mit großen Treffermengen sowie übermäßigem Grundrauschen (PROBLEMS 15, PROBLEMS 16).

Diese Probleme resultieren beim Nutzer häufig im Gefühl der Unsicherheit auf der einen Seite:

„aber das muss man ja auch erst einmal abschätzen können“ (BALANCE 1)

und zu einer erhöhten intellektuellen Belastung, beispielsweise im Screening-Prozess:

„Das Problem [...] ist zumeist die geringe Spezifität, die für den angestrebten Recall oft Trefferzahlen liefert, die für die intellektuelle Trefferelektion zu groß sind.“ (UNSPEC_SYN_CRITERIA 15).

Während Screening-Prozesse systematischer Übersichtsarbeiten im Umfeld der Cochrane Collaboration und anderer Institutionen standardisiert ablaufen und dabei in größeren, interdisziplinären Teams arbeiten können, erfolgt die Relevanzbeurteilung von Treffermengen strukturierter Suchen anderenfalls häufig durch den einzelnen Auftraggeber. Der Umgang mit großen Treffermengen ist daher geprägt von der „Erwartungshaltung“ des Auftraggebers. Die Aufgabe des Informationsspezialisten scheint dabei je nach Wunsch des Auftraggebers mit

Generierung einer „handhabbaren“ Treffermenge beendet, was sich sehr deutlich bereits in der Beobachtungsstudie zeigte.

4.4.7 Verwendung datenbankspezifischer Funktionalitäten und Text Mining-gestützter Verfahren

Keine der im Fragebogen formulierten Fragen zielte direkt auf die Verwendung datenbankspezifischer Funktionalitäten (siehe hierzu auch Abschnitt 3.4.2.3) oder verfügbarer Text Mining Werkzeuge (siehe Abschnitt 3.4.2.3) ab. Da die Datenanalyse hierzu aber Einblicke zulässt, die für das Gesamtbild hilfreich sind, werden diese im Folgenden kurz zusammengefasst.

Es ist davon auszugehen, dass nicht nur der Durchlauf finaler Suchprozesse, sondern auch die Durchführung von Testsuchen weitestgehend im professionellen Datenbankmodus durchgeführt und darüber auch dokumentiert werden (hierfür spricht beispielsweise GENERAL 3), vgl. hierzu Abbildung 3-10 und Abbildung 3-11.

Die Funktion „*automatic term mapping*“ (siehe Abschnitt 3.4.2.3) kann für Pubmed Suchen zur Identifikation von mit Schlüsselbegriffen direkt korrespondierenden Deskriptoren genutzt werden:

„einfach Suchbegriffe eingeben und Umwandl[u]ng per Automatic Term Mapping ansehen, um zu schauen, ob brauchbare Thesaurusbegriffe vorgeschlagen werden“ (GENERAL 4)

„Automatic Term Mapping‘ um zu sehen, wie Anfrage nach relevanten Begriffen umgesetzt wird“ (GENERAL 5)

Bei Zugriff auf MEDLINE über den Host OVID ist die Funktion „Begriff einem Schlagwort zuordnen“ in der angewandten Suche direkt verfügbar, siehe Abschnitt 4.4.6.2.

In der Auswahl der hier befragten Informationsspezialisten spielen Text Mining-Funktionalitäten bisher eine eher geringe Rolle. Vorwiegend wird hierbei das bereits an anderer Stelle vorgestellte Tool PubReminer (siehe Abschnitt 3.4.5.2) genannt. Während dies in der Methodenlehre vorwiegend als unterstützende Maßnahme zur Generierung möglichst objektiver Such-

anfragen beschrieben ist (siehe ebenfalls Abschnitt 3.4.5.2), wird es bei Informationsspezialisten auch eingesetzt, um sich in bisher unbekannte Themen einzuarbeiten und sich inspirieren zu lassen:

“I may use PubReminer text analysis for subjects I am not familiar with to help pick and check terms.” (GENERAL 9)

„evtl. in PubMed Reminer eingeben und die Listen anschauen: Häufigkeit MeSH, Stichworte anschauen“ (GENERAL 5)

Eine Befragte äußert sich dezidiert dazu, dass Empfehlungen aus Text Mining-Analysen kein Entscheidungskriterium für die Aufnahme von Suchtermen seien: “Text mining analysis (PubReminer) recommendation would not determine whether the synonym is added or not, but would inform the decision.” (SELECTION_TERM 9).

Insgesamt wird die Existenz solcher Tools bei den Nutzern zwar durchaus positiv und nützlich beurteilt (PROBLEMS 6, SELECTION_TERM 15), aus den Daten lassen sich jedoch weder eine frequente Nutzung noch eine Zielsetzung und durchgängige Nutzungsform ableiten. Explizit wünschen sich Nutzer Datenbank-übergreifend verfügbare Funktionen, wie beispielsweise das Highlighten von in Suchanfragen verwendeter Freitextterme in Treffermengen (verfügbar in OVID) (PRECISION 1) oder Werkzeuge wie PubReminer (PROBLEMS 5).

4.5 Zusammenfassung

Ziel der empirischen Untersuchungen war

- eine aus den *working tasks* der systematischen Literaturrecherche ableitbaren *seeking* und *searching tasks* in Anlehnung an die in Abschnitt 2.5.3 beschriebene Klassifikation
- eine kontextuelle Einordnung dieser Aufgaben sowie
- eine Klassifikation kontextabhängiger Informationsverhalten und Lösungsstrategien.

Mit der als Vorstudie konzipierten Beobachtungsstudie konnten wesentliche Aufgaben während der Entwicklung strukturierter Suchprofile im Rahmen systematischer Übersichtsarbeiten identifiziert und bestätigt werden:

Die Analyse der Beobachtungsstudie bestätigte die Aufgabenstruktur der Methodenlehre systematischer Übersichtsarbeit und ermöglichte zusätzlich die Identifikation spezifischer *search tasks* und *seeking tasks*, die mit typischen, für strukturierte Suchprozesse verfügbaren Interaktionskontexten assoziiert werden konnten.

Vorherrschende Taktik für die Identifikation von Suchtermen war die systematische Analyse relevanter Dokumente einer Ground Truth und die Bewertung von deren Rückgewinnung bei Änderungen am Suchprofil. Dies ist ein Muster, das in der Methodenlehre beschrieben wird und darüber hinaus auch in der Qualitätssicherung strukturierter Suchprozesse Verwendung findet, vgl. Abschnitt 3.4.4.2.

Da es sich bei der Beobachtungsstudie thematisch um einen Review zu einer komplexen Intervention handelte, wurden Probleme hinsichtlich der Strukturierung von Building Blocks ebenso sichtbar wie die der Diversität von Fachterminologien, die ebenfalls in der Methodenlehre eingehend beschrieben sind, siehe Abschnitt 3.4.3.1.

Die Beobachtungsstudie zeigt zusammenfassend ein mögliches, wenn auch sicherlich nicht prototypisches Bild der Entwicklung und Durchführung systematischer Übersichtsarbeiten und unterliegt daher der deutlich sichtbaren Kritik mangelnder Objektivität: Die Untersuchung wurde an nur einem Probanden durchgeführt, was keinerlei Verallgemeinerung der Ergebnisse zulässt und zudem von nur einer Protokollführerin erfasst, was trotz objektiverer Analyseverfahren die Gefahr systematischer Verzerrungen nicht ausschließen lässt.

Eine angegliederte Hauptstudie hinterfragte die bisher vorliegenden Ergebnisse kritisch und ergänzte sie um weitere Faktoren, die über die Beobachtungsstudie nicht identifiziert werden konnten. So konnten Aufgabenstrukturen, Entscheidungssituationen und potenzielle Lösungsstrategien konkretisiert und bedingende Rahmenparameter festgemacht werden. Dabei muss allerdings kritisch angemerkt werden, dass aufgrund der Datenlage auch hier einzelne Aspekte nur prototypisch dargestellt werden konnten, beispielsweise Gründe für die Durchführung von Testsuchen wie auch die Methodik einzelner Testsuchen.

5 Generalisierung und Einordnung

Das aktuelle Kapitel präsentiert ein aus den Ergebnissen der Studie entwickeltes, theoretisches Modell des professionellen Informationsverhaltens in der Entwicklung von strukturierter Suchanfragen für die systematische Suche in bibliographischen Fachdatenbanken und ordnet es anhand der in Kapitel 2 dargestellten Modelle auf kritische Weise informationswissenschaftlich ein.

Es beschränkt sich mit der Entwicklung von Suchstrategien für bibliographische Fachdatenbanken auf eine zentrale Aufgabe der systematischen Übersichtsarbeit.

5.1 Modell zur iterativen Entwicklung verifizierter Suchprofile

Ausgangspunkt der Entwicklung einer Suchstrategie für die strukturierte Suche in bibliographischen Fachdatenbanken ist der Auftrag der Durchführung einer systematischen Literaturrecherche. Basis einer Suchstrategie bildet die „suchbare Fragestellung“ als rudimentäres Konstrukt einer späteren Booleschen Suchanfrage.

5.1.1 Vom Synonym zum idealen Suchterm und Suchprofil

Aus den Ergebnissen der empirischen Untersuchungen lässt sich ableiten, dass das Vertrauen in Suchterme und Suchprofil erst dann gegeben ist, wenn jedes aufgefundene Synonym die Phase vom Kandidatenterm zum verifizierten Suchterm durchlaufen hat, also sowohl inhaltlich validiert als auch im Suchprofil hinsichtlich seines Ertrags evaluiert wurde.

Im Ergebnis lassen sich daraus folgende *Evolutionsschritte* für Suchterme und Suchprofil definieren, die zum einen den Anspruch der evidenzbasierten Medizin an die methodologisch streng definierte systematische Übersichtsarbeit verdeutlichen, zum anderen den Motor für das Informationsverhalten bilden:

Tabelle 5-1 Zustände von Suchtermen

Zustand	Charakterisierung
Kandidatenterm	Deskriptoren, Einzelwort Terme, Artefakte, die als subjektiv relevant beurteilt wurden
Evaluiertes Kandidatenterm	Kandidatenterm, der für ein vollständiges Suchprofil notwendig ist
Validiertes Kandidatenterm	Kandidatenterm, der für die Suche als situativ relevant beurteilt wurde
Verifizierter Suchterm	Kandidatenterm, der sowohl evaluiert als auch validiert wurde
Idealer Suchterm	Idealer Suchterm mit möglichst hoher Spezifität und geringem Grundrauschen

Dabei gilt: Kandidatenterme sind im Suchprofil unter anderem dann **notwendig**,

- wenn sie Keywords der Building Blocks eindeutig repräsentieren
- wenn das Suchprofil durch sie weitere relevante Treffer liefert
- wenn das Suchprofil nur durch sie das Retrieval bekannter Schlüsselreferenzen sichert.

Kandidatenterme können vorab **validiert** sein, wenn sie

- vom Domänenexperten genannt wurden
- in vorhergehenden Versionen systematischer Reviews verwendet wurden
- in ähnliche Reviews verwendet wurden
- dem Informationsexperten als thematisch relevant bekannt sind
- Konzepte einzelner Building Blocks exakt abbilden
- benötigt werden, um das Retrieval von Elementen aus der Ground Truth zu sichern
- konsistent für die Indexierung bekannter relevanter Artikel aus der Ground Truth genutzt werden.

Suchprofile sind per definitionem dann und nur dann als vollständig anzusehen, wenn die Hinzunahme weiterer Suchterme keine neuen, relevanten Treffer mehr liefert.

Ein weiterer Aspekt ist die Spezifität von Suchprofilen. Suchprofile können dann als **ideal** definiert werden, wenn Suchterme datenbanksyntaktisch so optimiert wurden, dass das Suchprofil möglichst wenig Grundrauschen produziert, ohne dabei das Retrieval bekannter, relevanter Dokumente zu gefährden. Der ideale Suchterm kann also definiert werden als ein Term, der die Sensitivität des Suchprofils erhöht, gleichzeitig aber ein möglichst geringes Grundrauschen verursacht.

Quantitative Bewertungsprozesse nehmen in der Beobachtungsstudie großen Raum ein und werden von allen Interviewpartnern der Fragebogenstudie in verschiedenem Kontext genannt. In Kombination mit der Frage nach dem Umgang mit unspezifischen, im Sinne allgemein gehaltener Synonyme, können folgende Bewertungsregeln für die quantitative Beurteilung von Kandidatentermen abgeleitet werden:

- Kandidatenterme mit hohem Grundrauschen werden als kritisch beurteilt
- Kandidatenterme mit nur geringem zusätzlichem Ertrag werden als kritisch beurteilt
- Kandidatenterme werden positiv bewertet, wenn sie die Sensitivität bezüglich der klinischen Fragestellung erhöhen
- Kandidatenterme werden als ideal bewertet, wenn sie nicht nur die Sensitivität erhöhen, sondern auch geringes Grundrauschen mitbringen

Validierte, aber schlecht evaluierte Freitextterme oder Phrasen können durch eine PROXIMITY-Kombination mit gut evaluierten Suchtermen oder idealen Suchtermen zu neuen, bereits validierten Kandidatentermen modifiziert werden, müssen dann aber zunächst den Evaluierungsprozess wieder durchlaufen. Schlecht evaluierte Kandidatenterme können durch die Verwendung von Qualifiern auf das Vorkommen in bestimmten Feldern eingeschränkt oder alternativ aus dem Suchprofil gestrichen werden.

Aufgrund der bereits wiederholt dargestellten Struktur der suchbaren Fragestellung und ihrer Repräsentation in Form von Building Blocks wird deutlich, dass Kandidatenterme nicht isoliert evaluiert werden, sondern dies immer im Kontext des aktuellen Suchprofils erfolgen muss. Durch die Boolesche Verknüpfung von Suchtermen einzelner Building Blocks mit dem OR-

Operator sowie die thematische Verknüpfung der Building Blocks entsprechend des PICO-Schemas mit dem AND-Operator können Änderungen am Suchterm selbst, Änderungen am Building Block durch Entfernen oder Hinzunahme von Suchtermen sowie Änderungen an den Building Blocks zu extremen Qualitätsschwankungen hinsichtlich Sensitivität und Spezifität führen.

Daher ist jede Änderung am Suchprofil mit der Kontrolle des Retrievals bekannter, relevanter Dokumente verbunden.

Der in der Methodenlehre häufig „Entwicklung einer Suchstrategie“ benannte Aufgabenkomplex lässt sich auf *working task* Ebene anhand der Ergebnisse aus Vor- und Hauptstudie zunächst in verschiedene Teilaufgaben zerlegen:

- Der Ausbau von Suchprofilen durch das Auffinden einer möglichst umfassenden Menge an Deskriptoren und Freitexttermen als auch Phrasen für jeden Building Block zur Sicherung der Vollständigkeit von Suchprofilen
- Vollständigkeitskontrollen in Form von Kosten- und Nutzenkontrolle einzelner Suchterme im Suchprofil zur Überprüfung zusätzlichen Ertrags und unnötigen Grundrauschens unter Beibehaltung des Retrievals bekannter Schlüsselreferenzen
- Die Optimierung von Kandidatentermen als datenbankspezifische Suchterme und gesamtem Suchprofil

Insbesondere die zur Bewertung von Suchprofil und finaler Suchanfrage nötigen Validierungsprozesse führen bei der Entwicklung von Suchanfragen zu einem zusätzlichen, situativen Informationsbedarf, sodass das Informationsverhalten als multidimensionales Problem betrachtet werden muss.

Die zur Vertrauensgewinnung in Synonyme, spätere Suchterme und das Suchprofil nötigen Arbeitsschritte führen zu zusätzlichen Aufgaben sowie charakteristischen Taktiken und Strategien.

5.1.2 Informationsverhalten von Informationsexperten

Aus der Methodenlehre definiert sich die Entwicklung erschöpfender Suchanfragen (**Exhaustive Query**) für strukturierte Suchprozesse als iterativer Prozess, in dem durch die Nutzung bereits identifizierter Dokumente so lange weitere Suchterme in Form kontrollierten Vokabulars, Freitexttermen und Phrasen in die Query aufgenommen werden, bis diese im Testlauf keine weiteren relevanten Treffer mehr zu Tage fördert.

Abbildung 5-1 stellt das aus den Ergebnissen der Studie resultierende *information seeking* und *information searching behaviour* während der Entwicklung erschöpfender, verifizierter Suchanfragen graphisch dar. Aus Gründen der Internationalisierung aller bekannten informationswissenschaftlichen Modelle und dem persönlichen Wunsch nach kontextueller Einordnung sind auch in folgendem Modell alle Terminologien englischsprachig gehalten. Das Modell orientiert sich in seiner Darstellung am Grundlagenmodell von Ingwersen und Järvelin (siehe Abbildung 2-20), wie dies auch von vergleichbaren Darstellungen in anderem Kontext bekannt ist, vgl. beispielsweise (Elsweiler et al. 2011). Es zeigt verschiedene Facetten des aus den empirischen Untersuchungen abgeleiteten Informationsverhaltens im Kontext der auch bei Ingwersen und Järvelin beschriebenen kontextuellen Rahmenparameter.

Processing beschreibt die Interaktion mit verfügbarer Information und wird eingehend in Abschnitt 5.1.2.1 erläutert. **Informationspotenzial und situativer Informationsbedarf** führen zu Routinetätigkeiten (*routine*) oder außerplanmäßigem Informationsverhalten (*on demand*), siehe Abschnitt 5.1.2.2. Wichtigste Routineaufgabe ist die iterative **Neuformulierung des Suchprofils**, das in verschiedenen Bewertungszyklen zum idealen Suchprofil führen soll und im Detail in Abschnitt 5.1.2.3 verallgemeinert dargestellt wird. Entsprechend des Ingwersen und Järvelin-Modells ist jedes professionelle Informationsverhalten in verschiedenen, sich wiederum bedingenden Kontexten zu sehen, die im genannten Grundlagenmodell bereits in Teilen definiert sind (siehe Abschnitt 3.6) und in der vorliegenden Arbeit konkretisiert werden müssen. *Cultural Context* definiert den Kontext Arbeitsumfeld und wird in Abschnitt 5.1.3.1 als Einflussfaktor beschrieben, *Information Seeking and Processing Context* beleuchtet den Akteur als ausgebildeten Informationsspezialisten (siehe Abschnitt 5.1.3.2). Schließlich bildet der *IR Context* einen Kontext, in dem der Akteur als Datenbanknutzer zu betrachten ist, was in seinen Facetten in Abschnitt 5.1.3.3 zusammenfassend beschrieben wird.

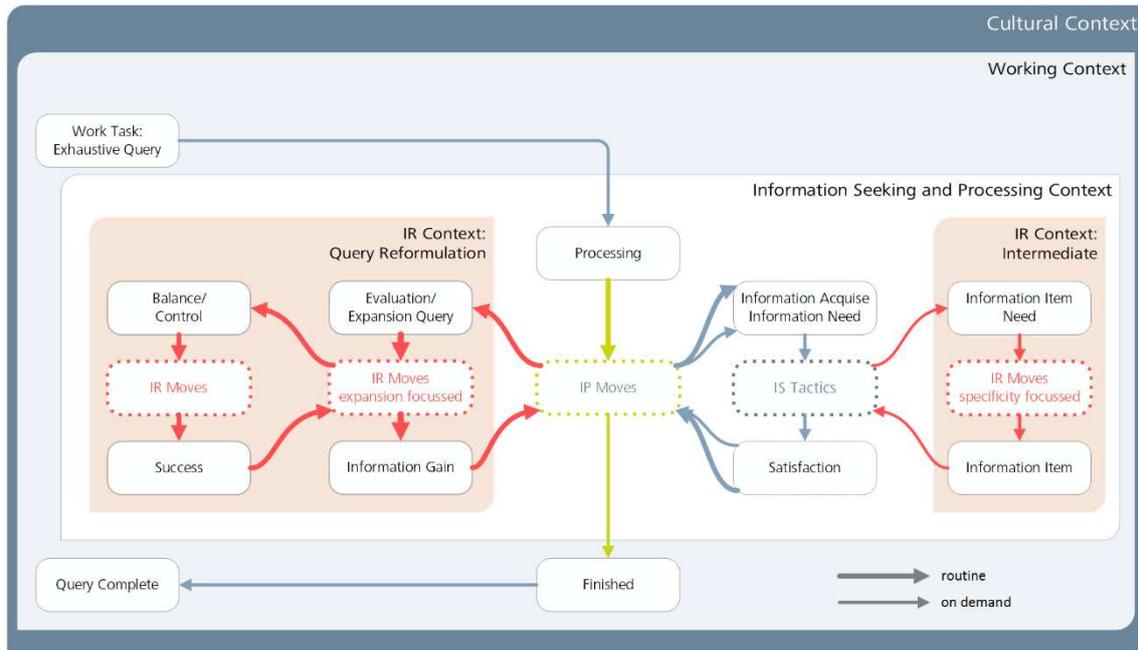


Abbildung 5-1 Aufgabenbezogenes Information Behaviour in der Entwicklung erschöpfender Suchfragen

5.1.2.1 Interaktion mit Information

Bei der Entwicklung von Suchprofilen steht die Verarbeitung verfügbarer Informationseinheiten (**Processing**) mit dem Ziel der Identifikation von Kandidatermen im Mittelpunkt von Aufgabenstruktur und Informationsverhalten, siehe Abbildung 5-1. Einige davon sind übliche Routineaufgaben, andere wiederum können sich aus einem Informationsbedarf ergeben und werden daher situativ durchgeführt.

Interaktionsmodi

Aus den empirischen Untersuchungen konnte eine Reihe klassischer Verarbeitungstätigkeiten (**IP Moves**), die bereits in Abschnitt 4.2.4.2 beschrieben wurden und kurzgefasst wie folgt definiert werden:

Analyzing definiert die strukturierte Analyse verfügbarer, bereits als relevant bekannter Informationseinheiten mit dem konkreten Ziel der Gewinnung weiterer Suchterme, meist mit dem Ziel des Dokumenten-Retrieval als auch die strukturierte Analyse von Thesauruseinträgen mit dem Ziel der Gewinnung von Kandidatermen

Exploring definiert weniger strukturierte Untersuchungsprozesse, die häufig mit Lern- und Verständnisprozessen im Kontext der klinischen Fragestellung assoziiert sind und nicht notwendigerweise in der Gewinnung neuer Suchterme enden

Screening meint die explorative Untersuchung von Information, meist Treffermengen durchgeführter Testsuchen und dient sowohl der inhaltlichen Kontrolle als auch der Gewinnung neuer Suchterme. Beim Screening werden Trefferlisten wie ausgegeben auf neue Kandidaterme hin überprüft.

Comparing beschreibt den Vergleich von Dokument und aktuellem Stand des Suchprofils im Hinblick auf die vollständige Sicherung des Dokumenten-Retrievals, eine darauffolgende Analyse ist dabei abhängig vom Ergebnis dieses Vergleichs

Differentiating definiert den Vergleich zweier Informationseinheiten mit dem Ziel der Verifikation von Synonymen

Extracting beschreibt die Gewinnung weiterer Kandidaterme als Ergebnis verschiedener Verarbeitungsprozesse

Entsprechend methodologischer Richtlinien stehen Umfang und Vollständigkeit der Sammlung von Kandidatermen im Vordergrund der Sammelphase, die Akteure priorisieren eine großzügige thematische Abdeckung gegenüber der Schärfe von Kandidatermen und Building Blocks.

Bei der Interaktion mit Informationseinheiten gilt in der Aufbauphase nach ökonomischen Gesichtspunkten vorwiegend das **Maximalprinzip**, das sich an spezifischen Eigenschaften verfügbarer und genutzter Informationseinheiten festmachen lässt und im Folgenden erläutert wird:

Charakteristika von Informationseinheiten

Informationseinheiten sind definiert als Elemente, die als Quelle für Synonyme, zu deren inhaltlicher Verifikation, als auch zum Auffinden weiterführender Informationseinheiten genutzt werden können.

Informationseinheiten stehen für sich, sind jedoch Teil eines größeren Ganzen, beispielsweise einer bibliographischen Datenbank oder eines Thesaurus. Informationseinheiten können auch durch Queries erzeugt werden, beispielsweise durch Testsuchen mit Suchprofilen. Sie sind charakterisierbar und unterscheidbar durch spezifische Eigenschaften wie in Tabelle 5-2 zusammengestellt.

Bibliographische Dokumente liefern mit Freitexttermen, Deskriptoren, Treffermengen und einem potenziell verfügbaren Zugriff auf den korrespondierenden Volltext den größten direkten Ertrag. Spezifische, das Informationsverhalten bestimmende Charakteristika betreffen die Studienqualität, insbesondere bei älteren oder sehr aktuellen Artikeln zudem deren Indexierung in den Datenbanken, und auch die Frage, ob das Retrieval eines Dokuments bereits gesichert ist oder nicht. Ein Spezifikum der bibliographischen Referenz ist ihre Struktur aus Titel, Abstract, Autorenangaben, Journalangaben sowie der Indexierung.

Thesauruseinträge liefern dem Nutzer in erster Linie Deskriptoren, durch ihre Beschreibungstexte aber auch Freitextterme. Das wichtigste Spezifikum eines Deskriptors ist dessen hierarchische Position im gesamten Thesaurus.

Treffermengen sind das Resultat von Testsuchen. Aus ihnen lassen sich im Screening-Prozess Freitextterme aus den Titel der Treffermenge extrahieren, gleichzeitig bieten Treffermengen aber auch den direkten Zugriff auf bibliographische Dokumente. Hinsichtlich der Evaluation von Kandidatentermen und Suchprofil sind sowohl der für den Nutzer sofort ersichtliche Umfang der Menge als auch das abzuschätzende Grundrauschen spezifische Größen.

Volltexte sind für die systematische Übersichtsarbeit nur dann von wirklicher Bedeutung, wenn es sich um systematische Übersichtsarbeiten selbst handelt. Im Fall der systematischen Übersichtsarbeit reicht Ertrag von Volltexten über Freitextterme über Zitationen hinaus bis zu vollständigen Suchprofilen. Theoretisch stehen in Freitexten auch zusätzliche Meta-Angaben zur Verfügung, die beispielsweise von Journals selbst vergeben wurden. Diese wurden in der Beobachtungsstudie jedoch nicht kommuniziert, sollen hier aber der Vollständigkeit halber genannt werden.

Zu den Spezifika von Volltexten gehört neben der zugrundeliegenden Studienqualität auch die Struktur des Dokuments.

Tabelle 5-2 Charakteristika von Informationseinheiten

Bibliographische Dokumente	
Direkter Ertrag	Freitextterme Deskriptoren Treffermengen (beispielsweise „similar articles“) Volltexte
Spezifika	Qualität (Evidenzpyramide) Indexierung ja/nein Retrieval gesichert ja/nein
Struktur	Titel Abstract Metadaten Indexierung
Thesauri/Deskriptoren	
Direkter Ertrag	Deskriptoren Subheadings Freitextterme (auch Entry Terms) Treffermengen
Spezifika	Hierarchische Position
Treffermengen	
Direkter Ertrag	Freitextterme Bibliographische Dokumente
Spezifika	Größe Grundrauschen
Volltexte	
Direkter Ertrag	Freitextterme Deskriptoren Suchanfragen (Systematic Review) Zitationen
Spezifika	Qualität (Evidenzpyramide)
Struktur	Titel Abstract

	Text Corpus Meta-Angaben
--	-----------------------------

Strategien in der Nutzung von Informationseinheiten

Vor allem in der Anfangsphase der Entwicklung von Suchprofilen werden qualitativ hochwertige Informationseinheiten bevorzugt, die zum einen hohen Ertrag an Suchtermen und hohes zusätzliches Informationspotenzial erwarten lassen, gleichzeitig aber zu wenig zusätzlichem Informationsbedarf führen.

In der Nutzung von Informationseinheiten sind dies Dokumente, die mit in der Evidenzpyramide (siehe Abschnitt 3.1.1) weit oben angesiedelten Studiendesigns assoziiert sind, beispielsweise randomisierte, klinische Studien oder systematische Übersichtsarbeiten; alternativ kontrolliertes Vokabular, das exakt mit den Konzepten einzelner Building Blocks übereinstimmt.

Die Verarbeitung von Informationseinheiten ist immer an eine Relevanzbeurteilung gekoppelt (zum Relevanzbegriff siehe Abschnitt 2.3.1.2).

Werden Dokumente als situativ relevant eingestuft, so führt dies zu Maßnahmen, welche die Sicherung seines Retrievals gewährleisten. Suchterme sind dabei entsprechend des Evolutionsmodells bereits validiert.

Gleichzeitig werden – sofern verfügbar – datenbankspezifische oder dokumentspezifische Interaktionsmöglichkeiten iterativ ausgeschöpft und analysiert. Insbesondere bibliographische Dokumente und Volltexte systematischer Übersichtsarbeiten bieten dem Nutzer diese direkten Interaktionsmöglichkeiten mit zusätzlichem Informationspotenzial:

In der Praxis führt dies unter anderem zu den in Abschnitt 4.2.4.3 definierten und in Abschnitt 4.4.6 konkretisierten Ground Truth-Loops oder zur Nutzung datenbankspezifischer Funktionalitäten wie „*similar articles*“, aber auch der Facettensuche, siehe Abschnitt 3.4.2.3. Die Wahl geeigneter Informationsressourcen ist damit nicht nur ökonomischen Gesichtspunkten geschuldet, sondern lässt sich direkt in der Methodenlehre der systematischen Übersichtsarbeit verorten, in der die Suche nach vorbestehenden systematischen Übersichtsarbeiten oder Arbeiten auf Basis randomisierter klinischer Studien methodisch empfohlen ist.

Wird während einer situativen Relevanzbeurteilung zusätzlicher Informationsbedarf erkannt, so muss diesem nachgegangen werden. Verarbeitungsprozesse sind abhängig von der Wahl

und Verfügbarkeit geeigneter Informationseinheiten per se nicht immer erfolgreich und sind dabei unterschiedlich komplex. Gründe für zusätzlichen Informationsbedarf sind insbesondere

- die Notwendigkeit des Dokumenten-Retrievals objektiv relevanter Dokumente, deren Relevanz jedoch aufgrund vorliegender Informationen für den Rechercheexperten nicht einschätzbar ist,
- unklare Indexierung von Dokumenten,
- Indexierungspraktiken und inkonsistente Verwendung kontrollierten Vokabulars,
- Relevanzbeurteilung der für die initiale Fragestellung als potenziell interessant beurteilten Suchterme.

5.1.2.2 Informationspotenzial und situativer Informationsbedarf

Verarbeitungsprozesse führen häufig zum Erkennen eines zusätzlichen Informationspotenzials (**Information Acquire**) oder zum Erkennen eines situativen Informationsbedarfs (**Information Need**), siehe Abbildung 5-1. Ersteres resultiert aus Vollständigkeitsanforderung, ist vor allem im Hinblick auf ein umfassendes Suchprofil gewollt und führt gezielt zum Ausbau der Ground Truth, letzteres stört Verarbeitungsprozesse und damit auch typische Abläufe.

Informationspotenzial

Ein durch die Nutzung von Informationseinheiten erkennbares zusätzliches Informationspotenzial wird von den Akteuren insbesondere bei qualitativ hochwertigen und als relevant eingestuften Informationseinheiten wahrgenommen, wenn nicht gar gezielt provoziert, und kann sowohl auf die Identifikation weiterer Informationseinheiten als auch die Identifikation weiterer Kandidaterme abzielen.

Zusätzliches Informationspotenzial wird entweder durch datenbankspezifische Funktionalitäten oder durch Anwendung typischer Information *search tactics* nutzbar gemacht und strategisch auch eingeplant. Dabei ist die Strategie „Best Representation“ vorherrschend:

Best Representation

Best Representation-Suchen dienen dem Ausbau der Ground Truth. Best Match-Dokumente sind dabei nicht nur thematisch verwandt, sondern nutzen auch ein ähnliches, meist in der

Evidenzpyramide oben angesiedeltes Studiendesign. Für die Best Match-Suche werden datenbankspezifische Funktionalitäten wie die Facettensuche oder „*similar articles*“ genutzt (siehe Abschnitt 3.4.5.1), die in großen bibliographischen Fachdatenbanken üblicherweise verfügbar sind. Berry Picking-Techniken, wie bei Bates beschrieben (siehe hierzu Abschnitt 2.1.2), sind Interaktionsmodi, die aus Zeitgründen häufig nicht durchgeführt oder an den Domänenexperten delegiert werden.

Situativer Informationsbedarf

Situativer Informationsbedarf ergibt sich häufig aus der zunächst subjektiv nicht beurteilbaren Relevanz einzelner Informationseinheiten oder Synonyme, er führt üblicherweise zu Validierungsprozessen. Diese sind beispielsweise dann nötig, wenn die Quelle von Synonymen keine subjektiv erkennbare Aussage über deren Notwendigkeit oder Validität zulässt, unter anderem Zitationen mit unspezifischem Titel, unspezifischen Abstract und unklarer Indexierung. Situativer Informationsbedarf kann auch durch das Erkennen spezifischer wie auch unspezifischer Synonyme entstehen.

Bezugnehmend auf das Byström und Järvelin-Modell, dargestellt in Abschnitt 2.2.4, ergibt sich situativer Informationsbedarf vorwiegend aus Mangel an domänenspezifischer Information oder domänenspezifischem Wissen. In der vorliegenden Arbeit wurde auch das Wissen um potenziell fehlerbehaftete oder unscharfe Indexierungspraktiken einzelner Datenbanken als Motor für situativen Informationsbedarf identifiziert.

Taktiken zur Lösung eines situativen Informationsbedarfs lassen sich klassifizieren in:

Good Luck

Die Suche „auf gut Glück“ ergibt sich aus einem meist initialen Informationsbedarf aufgrund „Lernen zum Thema“ in der Anfangsphase der Entwicklung von Suchprofilen.

Sie wird typischerweise durch Ad Hoc-Suchen mit Freitexttermen realisiert und dient insbesondere der Identifikation von thematisch mit den Konzepten einzelner Building Blocks korrespondierenden kontrollierten Vokabulars.

Für die Good Luck-Suche werden typischerweise datenbankspezifische Funktionalitäten wie „*automatic term mapping*“ (Pubmed) oder Text Mining-gestützte Softwarelösungen genutzt, welche die Erschließung von Thesauri auf Basis einfacher Freitextsuchen ermöglichen.

Refining

Informationsbedarf, der sich aus der Notwendigkeit des Retrievals relevanter Zitationen mit unspezifischem Titel ergibt, mündet in typischen „*known item searches*“, dem bereits in Abschnitt 4.2.4.2 beschriebenen **Specifying**. *Known item* Suchen können sowohl die Identifikation bibliographischer Dokumente als auch die Identifikation von Volltexten zum Ziel haben. Die grundsätzliche Verfügbarkeit bibliographischer Dokumente ist in der *strukturierten* Suche dabei durch die Nutzung und institutionelle Lizenzierung großer bibliographischer Datenbanken weitestgehend gesichert. Zur Beschaffung von Volltexten werden dabei jedoch auch andere Informationsressourcen wie Suchmaschinen oder Google Scholar genutzt.

Object of Comparison

Der Umfang nötiger Validierungsmaßnahmen nimmt zu, wenn Zitationen neben einem unspezifischen Titel auch einen unspezifischen Abstract führen oder unklar indexiert sind, sodass auch bibliographische Dokumente keine direkte Extraktion validierter Suchterme zulassen.

Ein „Object of Comparison“ dient im Validierungsprozess als Vergleichsobjekt.

Diese Form der Suche mündet direkt in weiteren Verarbeitungsprozessen (**Comparing**). Vergleichbar ist „Object of Comparison“ mit „Good Luck“ Suchen. Die Suche erfolgt hier allerdings weniger explorativ, sondern nutzt auch gezielte Thesaurus-Suchen.

Die für den Akteur problematische Situation ergibt sich daraus, dass situativer Informationsbedarf in der Entwicklung erschöpfender Suchprofile aufgrund der Vollständigkeitsforderung immer gedeckt werden muss. Dies betrifft vor allem Refining und Object of Comparison Suchen. Akteure tendieren daher mit Hilfe der bereits angesprochenen Strategie der Nutzung vorab bereits validierter Informationseinheiten dazu, einen situativen Informationsbedarf weitestgehend zu vermeiden, indem sie auf Informationseinheiten zurückgreifen, die möglichst valide Kandidatenterme liefern können.

Gelingt dies nicht, sind die Taktiken zur Überwindung dieses Informationsbedarfs entsprechend wenig stringent. Sie umfassen neben der explorativen Suche in verschiedensten, sowohl online als auch analog verfügbaren Datenbeständen vor allem die Kommunikation mit dem Domänenexperten. Dieser ist vor allem in der Vorbereitungsphase der Suche wichtiger Ansprechpartner für die Klärung von Fragen und wird auch mit der Beschaffung situativ relevanter Information beauftragt. Eine weitere Methode zur Deckung situativen Informationsbedarfs besteht daher in der Kommunikation mit Domänenexperten. Entgegen der streng me-

thodologisch vorgegebenen Schritte der datenbankspezifischen Suche scheint bei der Deckung situativen Informationsbedarfs „alles erlaubt“, so lange dieser Bedarf gedeckt werden kann.

Situativer Informationsbedarf kann während des Gesamtprozesses prinzipiell immer auftreten, wird aber häufig während der Anfangsphase der Entwicklung von Suchanfragen beobachtet. Im Verlauf der Entwicklung von Suchprofilen sind Suchprozesse, die sich aus situativem Informationsbedarf ergeben, nicht dokumentiert.

5.1.2.3 Iterative Query Reformulation

Dem gegenüber steht die iterative Reformulierung von Suchanfragen (**Query Reformulation**), die vorwiegend mit der Bewertung von Suchprofilen assoziiert und daher ausschließlich durch stark datenbank-orientiertes Informationsverhalten charakterisiert ist, siehe Abbildung 5-1.

Eine Reformulierung der Suchanfrage wird grundsätzlich immer durch die Aufnahme neuer Kandidatenterme in das Suchprofil getriggert und führt grundsätzlich zu Testsuchen im Retrieval-System.

Unterschieden werden muss zwischen Query Reformulation Schritten, die der Vollständigkeitsbeurteilung und damit der **Expansion** der Suchanfrage dienen, und solchen, die der Kontrolle sowie der Balance zwischen Sensitivität und Spezifität der Suchanfrage dienen (**Balance/Control**).

Evaluation von Suchtermen und Expansion von Suchprofilen

Die Expansion der Suchanfrage durch die Aufnahme weiterer Suchterme dient der gezielten Erweiterung des Suchprofils, sie erfolgt ausschließlich durch Testsuchen. Sie führt demnach zu zwei Entscheidungssituationen:

- Liefert ein Kandidatenterm weitere relevante Treffer?
- Wie groß ist das Grundrauschen einzelner Kandidatenterme?

Im Gegensatz zum situativen Informationspotenzial werden bei der Expansion von Suchprofilen keine datenbankspezifischen Funktionalitäten wie „*similar articles*“ oder die Facettensuche genutzt, vielmehr erfolgt der Ausbau der Query durch eine gezielte iterative Anwendung von Query und *Query Reformulation* im professionellen Bedienmodus, gegebenenfalls auch im Vergleich zweier Queries. Die aufgrund der Entscheidungssituation notwendige Relevanzbeurteilung ist daher weniger subjektiv, sondern eher situativ zu werden. Eine aus der Durchführung von Queries resultierende positive Relevanzbeurteilung klassifiziert neue Kandidaterme als validierte Kandidaterme und führt durch das Screening von Treffermengen immer zur Verarbeitung neuer, relevanter Informationseinheiten.

Ein typisches Beispiel hierfür ist das iterative Citation Pearl Growing Prinzip, das bereits aus den Anfangszeiten online verfügbarer bibliographischer Fachdatenbanken bekannt ist (siehe hierzu Abschnitt 2.4), bei dem iterativ bibliographische Angaben relevanter Treffer aus Testsuchen genutzt werden, um weiteres kontrolliertes Vokabular zu identifizieren. Eine Ausarbeitung der Citation Pearl Growing Strategie sind das **Thesaurus NOT Free Text** beziehungsweise das **Free Text NOT Thesaurus** Manöver, siehe hierzu Abschnitt 4.4.4.1.

Query Reformulation in der Expansionsphase von Suchprofilen ist geprägt durch datenbanksyntaktische Formulierung, die eine möglichst sensitive Suche zulassen, vor allem Trunkierung und Maskierung. Das iterative, datenbankspezifische Testen von Queries und *Query Reformulation* gehört für den Informationsexperten zu den Routineaufgaben, kann jedoch zu situativem Informationsbedarf führen, beispielsweise wenn das Screening von Treffermengen keine eindeutige Aussage über den zusätzlichen Ertrag und damit der situativen Relevanz eines Kandidaterms zulässt.

Die Durchführung von Testsuchen ist ganz klar in einem datenbankspezifischen Kontext zu sehen und erfordert vom Akteur nicht nur Kenntnis der Datenbanksyntax, sondern auch das Wissen um Benutzerinterfaces und Datenbankfunktionalität. So werden verfügbare Datenbankfunktionen im Allgemeinen auch ausgereizt. Mehr noch ist der Akteur abhängig von der Verfügbarkeit als auch Funktionstüchtigkeit dieser Features, darunter beispielsweise das Highlighten verwendeter Suchbegriffe in Treffermengen. Interessanterweise verlassen sich die Akteure trotz ihrer sonst sehr strukturierten und dokumentierbaren Vorgehensweise beim Screening von Treffermengen auf wenig transparente Datenbankfunktionalitäten wie die Standard-sortierung von Treffermengen nach dem Modus „neueste Veröffentlichungen zuerst“ oder

Sortiermechanismen beziehungsweise das *Relevance Ranking* bei Pubmed und reduzieren den Screeningprozess auf eine für sie überschaubare Menge.

Sensitivität und Spezifität

Beurteilen Informationsexperten das Grundrauschen von Kandidatentermen als zu hoch, erfolgen in der Regel Testsuchen zur Optimierung von Sensitivität und Spezifität, die wiederum von einer Entscheidungssituation abhängen:

Führt die Query Reformulation durch Spezifizierung zum Verlust als relevant bekannter Dokumente?

Da die Gefahr des Verlusts einzelner relevanter Dokumente bei der Präzisierung von Suchanfragen die größte Gefahr der systematischen Literaturrecherche darstellt, werden Manöver zur Spezifizierung von Suchanfragen von den Akteuren als eher kritisch beurteilt und sind immer mit entsprechenden Kontrollen verbunden. So nimmt der Informationsspezialist lieber einen erhöhten Workload während der Analysephase der Gesamttreffermenge in Kauf, als einzelne Dokumente zu verlieren.

Bevorzugt werden zur Verbesserung der Balance zwischen Sensitivität und Spezifität iterative Änderungen an Building Blocks vorgenommen, um mit Taktiken wie „Most specific facet first“ (siehe Abschnitt 2.4.4) zunächst handhabbare, in jedem Fall aber wesentliche Aspekte der initialen Fragestellung vollständig abzudecken. Auch diese sind bereits in den frühen Jahren online verfügbarer Systeme beschrieben worden.

Maßnahmen zur Spezifizierung von Suchanfragen werden insgesamt indifferent beurteilt. Aus technischer Sicht wird – vor allem im Hinblick auf nicht prüfbare Verluste – häufig generell davon abgeraten, die Operatoren AND und NOT einzusetzen, gegenüber der Lokalisierung einzelner Suchterme in speziellen Datenbankfeldern durch Qualifier (siehe Abschnitt 3.4.2.2), scheint es keine klaren Präferenzen zu geben.

Ein domänenspezifisches Charakteristikum balance-orientierter Suchprofile ist die Verwendung methodologischer Suchfilter, siehe hierzu Abschnitt 3.4.3.2.

5.1.3 Informationsverhalten im Kontext von Arbeitsumfeld Expertise und Dokumentenverfügbarkeit

Das professionelle Informationsverhalten bei der Entwicklung verifizierter Suchprofile ist nicht isoliert, sondern immer im Kontext verschiedener Umgebungsvariablen zu sehen. Wie in Abbildung 5-1, angelehnt an das Modell von Ingwersen und Järvelin, dargestellt, finden sich hier verschiedene Kontextebenen.

5.1.3.1 Der Akteur in seinem Arbeitsumfeld

Die systematische Literaturrecherche und der daraus resultierende Arbeitsauftrag „Entwicklung einer Suchstrategie“ sind dem Informationsexperten zugewiesene Aufgaben, aus der sich klassischerweise eine Reihe zwingend durchzuführender Suchaufgaben ergeben.

Die Notwendigkeit zur Durchführung systematischer Übersichtsarbeiten ergibt sich per definitionem aus der evidenzbasierten Medizin, wie in Kapitel 1 eingehend beschrieben: Jedes medizinische Handeln beschreibt, das dem individuellen Patienten eine Behandlung zukommen lässt, die gewissenhaft ermittelt auf den besten, aktuell zur Verfügung stehenden Daten und Ergebnissen beruht (Sackett et al. 1996).

Genau aus dieser Definition lässt sich die wesentliche Aufgabe im Rahmen des evidenzbasierten Handelns ableiten: Die (regelmäßige) *systematische* Suche nach evidenzbasierten Ergebnissen in der medizinischen Literatur, sowie deren Bewertung im Hinblick auf die konkrete klinische Fragestellung (Sackett et al. 1996; Sackett und Rosenberg 1995). Mehr noch untermauert sie die Notwendigkeit einer systematischen Vorgehensweise bei jeder Recherche- und Bewertungstätigkeit.

Die in der Methodenlehre vorgegebenen Richtlinien und Normen zur Durchführung der systematischen Literatursuche sind daher die bestimmenden Kontextparameter in der Durchführung systematischer Übersichtsarbeiten, siehe hierzu Abschnitt 3.4.3. Diese Best Practices und Normen bestimmen wesentlich den Gesamtprozess der Entwicklung von Suchprofilen.

Ihnen unterliegen vor allem die Informationsspezialisten, die institutionell entsprechend verankert sind und sich von Berufs wegen ihrer Einhaltung verschrieben haben (beispielsweise dem MECIR Statement, siehe u.a. Abschnitt 3.3.3).

Da sie bei nicht entsprechend institutionell verankerten Informationsspezialisten als Handreichung für eine methodologisch einwandfreie Vorgehensweise dienen, bestimmen sie auch dort als ein dem Gesamtprozess übergeordnetes Prinzip in der Regel immer das Informationsverhalten.

Ausgehend davon werden alle in Handbüchern beschriebenen Aufgaben im Kontext der **Query Reformulation** (siehe Abbildung 5-1) meist durchgängig und routiniert durchgeführt.

Gestört wird die Anwendung der Methodenlehre im nicht institutionellen, meist akademischen Umfeld durch Faktoren, die sich direkt aus den Charakteristika der systematischen Literaturrecherche als Auftragsrecherche ergeben. Während Informationsexperten, die für EBM-assozierte Institutionen arbeiten, direkt den dort verankerten Grundsätzen unterliegen, trifft dies für den einzelnen Informationsspezialisten, beispielsweise im universitären Umfeld auf den direkten Auftraggeber zu: Obwohl Informationsspezialisten gerne methodologisch sauber arbeiten, unterliegen sie im letztendlichen Verfahren doch den Wünschen und Anforderungen des Auftraggebers. Da dieser im Allgemeinen weder über die nötige Rechercheexpertise verfügt, noch den Umfang existierender sowie durch Retrievalprozesse zurückgelieferter Informationen abschätzen kann und damit die auf ihn selbst zukommenden Analyseprozesse unterschätzt, führt dies in der Praxis auch zu einem Informationsverhalten, welches der Experte methodologisch kritisch bewertet oder aus eigener Erfahrung eher ablehnt.

5.1.3.2 Der Akteur als ausgebildeter Informationsexperte

Informationsspezialisten tendieren generell dazu, überwiegend die während ihrer Ausbildung gelernten oder von Kollegen überlieferten Handgriffe und Techniken anzuwenden und sind in ihrem Informationsverhalten, insbesondere in klassischen *Search Tasks* durch eine konservative, höchst strukturierte Vorgehensweise charakterisierbar. Wie sehr sich die Ausbildung von Informationsspezialisten sowohl an Funktionalität und Grundprinzip Boolescher Datenbanken als auch der Methodenlehre systematischer Übersichtsarbeiten orientiert, wurde bereits in den Abschnitten 3.4.2 sowie 3.4.3 deutlich gemacht. So beinhaltet dies sowohl die effiziente Durchführung klassischer Suchprozesse in Form komplex formulierter Queries als auch die Anwendung geeigneter *Information Retrieval Tactics*, wie sie beispielsweise bereits bei Markey und Cochrane und später bei Booth beschrieben werden. Beides gilt insbesondere für Experten aus den Bibliothekswissenschaften.

Überlagert wird dies von Faktoren wie der Ausbildung, der Berufserfahrung, persönlicher Routine im Umgang mit genutzten Informationsressourcen, dem Wissen um die Verfügbarkeit von Informationsressourcen sowie um die technische Implementierung von Information Retrieval-Systemen.

Vor allem Berufserfahrung und die Anzahl bereits durchgeführter systematischer Übersichtsarbeiten führt dazu, dass Akteure die von ihnen genutzten Datenbanken und ihre Thesauri kennen und entsprechend routiniert nutzen.

Die im Verlauf der Untersuchung aufgestellte Hypothese, dass Akteure mit informationswissenschaftlichem Hintergrund sowie Akteure, die kein EBM-spezifisches Schulungsprogramm durchlaufen haben, eine grundsätzlich andere Herangehensweise an die systematische Übersichtsarbeit haben, konnte aufgrund der Datenlage nicht seriös belegt werden. Hier wären weitere Untersuchungen mit einer größeren Teilnehmergruppe nötig.

Die vorliegenden Daten lassen vermuten, dass diese Faktoren zu einem von der Methodenlehre grundsätzlich abweichenden Informationsverhalten führen, indem nicht die Anforderung größtmöglicher Sensitivität, sondern von Beginn an die Anforderung nach bestmöglicher situativer Relevanz im Vordergrund stehen. Dies führt zwar im Gegensatz zur methodologisch geprägten Vorgehensweise zu vermehrt thematischen und subjektiven Relevanzbeurteilungen, reduziert aber andererseits die Notwendigkeit von Kontrollsuchen. Gleichzeitig erhöht die individuelle Expertise sowohl das Vertrauen in die eigene Praxis als auch in die Validität der eigenen Ergebnisse.

5.1.3.3 Der Akteur als Datenbanknutzer

Informationsspezialisten sind vorwiegend Datenbanknutzer. Suchprozesse, die sich nicht direkt auf bibliographischen Datenbanken (anzunehmend sei ergänzt: klinischen Studienregistern) erledigen lassen, werden häufig an den Domänenexperten delegiert. Als Datenbanknutzer verfügen Informationsspezialisten von Berufs wegen meist über Expertise in der Recherche in mehreren bibliographischen Fachdatenbanken, darunter mindestens der Cochrane Library, Ovid und EMBASE.

Leider, so bestätigen es die Interviews, haben nicht alle Informationsexperten gleichermaßen Zugriff auf verfügbare und für eine umfassende Recherche notwendige Fachdatenbanken. Während institutionell verankerte Experten teilweise über ausreichende Infrastrukturen verfügen, scheint es vor allem im universitären Umfeld zu Lizenzierungsproblemen zu kommen, welche die Recherchemöglichkeiten deutlich einschränken.

Das Informationsverhalten der Informationsspezialisten ist stark datenbankspezifisch geprägt. Die Nutzer kennen sowohl Funktionalitäten als auch Syntaxspezifika einzelner Datenbanken, mehr noch kennen sie Stärken und Schwächen nicht nur der Benutzerinterfaces, sondern auch datenbankspezifischer Indexierungspraktiken und daraus resultierender Probleme, vgl. Abschnitt 3.4.2.

Für strukturierte Suchprozesse werden, wie sich aus den Daten schließen lässt, vorwiegend professionelle Suchmodi genutzt, die Ergebnisse der Untersuchungen lassen auf eine routinierte Nutzung dieser Funktionen schließen. Funktionen, die in der Praxis für Domänenexperten konzipiert sind, darunter die Facettensuche, die Nutzung der Funktion „*similar articles*“ oder das Pubmed spezifische „*automatic term mapping*“ werden insbesondere zur Deckung situativen Informationsbedarfs oder in der initialen Lernphase genutzt.

Probleme in der Nutzung von Datenbanken ergeben sich sowohl aus proprietären Softwarelösungen einzelner Anbieter (und der daraus resultierenden Notwendigkeit, sich neue Strategien aneignen zu müssen) als auch durch Unterschiede in den Benutzeroberflächen und dem Funktionsumfang einzelner Datenbanken. Es ist anzunehmen, dass sich dies noch deutlicher in der Nutzung klinischer Studienregister (siehe Abschnitt 3.2.5) zeigen würde, die nicht Teil der empirischen Untersuchung war.

5.2 Einordnung

In Abschnitt 5.1 wurde ein aus den Ergebnissen der empirischen Untersuchung gewonnenes Modell zur iterativen Entwicklung verifizierter Suchprofile entwickelt und in den Kontext der Rahmenbedingungen systematischer Übersichtsarbeiten in der Medizin gesetzt. Die vorliegende Arbeit fokussierte sich dabei ausschließlich auf systematische Übersichtsarbeiten zu *interventionellen* Themengebieten, die Basis der evidenzbasierten Medizin (siehe hierzu Kapitel 1).

Der folgende Abschnitt dient der Diskussion des Modells und seiner Einordnung in die Landschaft informationswissenschaftlicher Modelle und Theorien zum Informationsverhalten. Dabei werden die in Abschnitt 2.5.4 formulierten Kontextvariablen aufgegriffen. Auf deren systematische und vollständige Betrachtung wird dabei jedoch zu Gunsten der wichtigsten Ergebnisse verzichtet. Zunächst aber soll die Untersuchung kritisch in den Gesamtkontext der professionellen Suche eingeordnet werden. Dabei ist insbesondere die Frage zu beantworten,

ob ein spezialisiertes Modell für das Suchverhalten von Informationsexperten in der systematischen Übersicht notwendig und sinnvoll ist.

5.2.1 Gesamtkontext

Ein Grund für ein solches Modell liegt in der erst in den vergangenen Jahren zunehmenden Fokussierung auf die professionelle Suche:

Die Entwicklung strukturierter Suchanfragen im Rahmen systematischer Übersichtsarbeiten in der Medizin ist schon allein durch ihre Definition als Auftragsrecherche ganz klar in die Riege der professionellen Suchen einzuordnen, anderswo als *Enterprise Search* bezeichnet (siehe Abschnitt 2.3.2). Der Begriff der Profession spielt für die informationswissenschaftliche Einordnung eine wichtige Rolle:

Das Informationsverhalten von Rechercheexperten ist im Gegensatz zu dem der Domänenexperten (vgl. (Ellis 1989; Marchionini 1995; Byström und Järvelin 1995; Leckie et al. 1996; Meho und Tibbo 2003)) informationswissenschaftlich generell bisher kaum beforscht, vgl. (Vassilakaki et al. 2014). Dem gegenüber stehen Modelle wie beispielsweise Kuhlthau (Kuhlthau 1991), das sich mit kognitiven Aspekten der Informationssuche beschäftigt, die für die Auftragsrecherche prinzipiell kaum Relevanz haben, allein schon, weil sie aus einem anderen experimentellen Kontext heraus entwickelt wurden.

Neuere Entwicklungen widmen sich dem Informationsverhalten in der webbasierten Enterprise-Suche, beispielsweise (Russell-Rose et al. 2011; Russell-Rose und Tate 2013), ein Kontext, der für die vorliegende Arbeit nur von geringem Interesse ist.

Die Frage nach der Einordnung professionellen Informationsverhaltens insbesondere im Kontext von Fachdatenbanken ist also bisher nicht beantwortet.

Ein zweiter wesentlicher Grund liegt darin, dass die systematische Übersichtsarbeit trotz ihrer Einordnung in die professionellen Suchen bisher kaum Gegenstand der Forschung ist. Hier jedoch sollte sie allein aufgrund ihrer teils sehr speziellen Anforderungen auch ihren Platz finden:

Vergleichbare Anforderungen wie in der systematischen Literaturrecherche finden sich beispielsweise bei Patentrecherchen, die mit dem Ziel der Vollständigkeit des Retrieval-Ergebnisses und entsprechender Qualitätsrichtlinien ähnlich restriktiven Rahmenbedingungen unterliegen und zudem ebenso wie die systematische Literaturrecherche häufig in Booleschen Fachdatenbanken mit ähnlich großem Datenbestand durchgeführt werden, vgl. (Jürgens et al.

2014). Die systematische Übersichtsarbeit hat in ihrem Anspruch und ihren Rahmenbedingungen generell also durchaus Parallelen mit der Patentrecherche.

Es gibt jedoch auch Unterschiede. Die systematische Übersichtsarbeit verlangt von den Akteuren, in dem Fall Auftragnehmern, eine dezidierte Dokumentation und Reproduzierbarkeit von Suchprozessen, die in der Veröffentlichung komplexer Suchanfragen mündet. Die Dokumentation dieser sogenannten Suchstrategie als Auflistung aller Queries wie in Fachdatenbanken durchgeführt, ist ein Spezifikum der systematischen Literaturrecherche. Ihre Vorgehensweise ist zudem im Kontext nicht-kommerzieller Institutionen methodologisch streng vorgegeben, während in der Patentrecherche lediglich methodologische Empfehlungen verfügbar sind (vgl. hierzu (United Nations Development Programme 2012; Gujarat Technological University 2013).

Auch in der Aufgabenstruktur sind Patentrecherche und systematische Literaturrecherche nur bedingt vergleichbar. Die Fachliteratur benennt in der Patentrecherche verschiedene Aufgabentypen, vgl. (Tseng und Wu 2008; Joho et al. 2010).

Die in der Patentrecherche typische Aufgabe, einen Überblick über verfügbare Patente zu erhalten, typischerweise genannt State of the Art (SOA) oder Freedom to Operate (FTO), vgl. (Joho et al. 2010) erfüllen einen ähnlichen Auftrag, weitere aus der Patentrecherche bekannten Aufgabenkomplexe lassen sich in dieser Form in der systematischen Übersichtsarbeit nicht identifizieren. Gleichzeitig sind in der Patentrecherche auch Aufgaben beschrieben, die lediglich die Identifikation eines einzigen Dokuments zum Ziel haben (Alberts et al. 2011) – ein dem Auftrag der systematischen Literaturrecherche diametral gegensätzliches Konzept. Die in der systematischen Literatursuche zentrale Aufgabe fokussiert sich *ausschließlich* auf die Identifikation einer in diesem Fall möglichst umfangreichen Menge an Information zu einer spezifischen Fragestellung. Die Methodenlehre benennt dies ungeachtet der Unmöglichkeit dieses Auftrags strikt „Identifikation aller Information“.

Ein ganz wesentlicher Punkt wird jedoch in der Fachliteratur zur professionellen Suche überhaupt nicht beachtet: Der Informationsexperte als Akteur der systematischen Übersichtsarbeit.

Professionelle Suche in der Medizin meint in der Fachliteratur derzeit offenbar ausschließlich domänenspezifische Suchprozesse, beschäftigt sich also mit dem Suchverhalten von Medizern. Eine Differenzierung zwischen professioneller Suche von Ärzten auf der Suche nach Fachliteratur auf der einen Seite (die domänenspezifische Suche) und der systematischen Li-

teraturrecherche als professionelle Suche auf der anderen Seite, ist in der informationswissenschaftlichen Literatur bisher nicht vorgenommen worden. Ziele und Anspruch domänenspezifischer Suchen werden im Vergleich zur systematischen Literaturrecherche nicht ausreichend definiert, siehe beispielsweise (Vassilakaki et al. 2014; Lupu et al. 2014). Auch wenn Tait die professionelle Suche als Auftragsrecherche definiert (Tait 2014), so beleuchtet die Forschung doch häufig ausschließlich den Domänenexperten.

Da die professionelle Suche durchgeführt von Domänenexperten sowohl in ihrem Umfang als auch in ihrer Systematik nicht vergleichbar mit der durch Informationsexperten durchgeführten systematischen Übersichtsarbeit ist (Herskovic et al. 2007; Mosa und Yoo 2013), entsprechen die in der Fachwelt durchaus diskutierten Aspekte des Suchverhaltens medizinischer Domänenexperten in keiner Weise methodologischen Anforderungen des Informationsexperten, siehe unter anderem (Coumou und Meijman 2006; Flynn und McGuinness 2011; Lykke et al. 2012; Korjonen-Close 2015).

Die systematische Übersichtsarbeit ist daher durchaus als eigener Forschungsgegenstand zu interpretieren, auch wenn es vor allem hinsichtlich der Anforderungen an strukturierte Suchen in Booleschen Datenbanken Gemeinsamkeiten beispielsweise mit der Patentrecherche gibt und daher wahrscheinlich Synergien genutzt werden können.

5.2.2 Aufgaben und Informationsbedarf

Charakteristisch für Auftragsrecherchen sind sowohl klar strukturierte Aufgaben als auch ein – meist wenigstens formal – vorliegender Informationsbedarf im Sinne einer Rechercheaufgabe, vgl. (Tait 2014). Dies findet sich auch in der systematischen Übersichtsarbeit.

5.2.2.1 Aufgaben der systematischen Literaturrecherche

In der systematischen Übersichtsarbeit sind insbesondere Strukturen von *working tasks* sowohl über die Definition der EBM als auch die Methodenlehre vorgeben und in der Literatur auch weiter durch *search tasks* konkretisiert, siehe beispielsweise (Tsafnat et al. 2014). Ein differenziertes Bild, das wie informationswissenschaftlich üblich hierarchisch zwischen *working tasks*, *information seeking tasks* und *information searching tasks* unterscheidet, existiert bisher jedoch nicht.

Die vorliegende Arbeit ergänzt dieses Bild:

Neben den in der Methodenlehre bereits vielfach benannten Teilaufgaben des initialen *working tasks* der systematischen Literaturrecherche (als Teilaufgabe der systematischen Übersichtsarbeit), vgl. hierzu Abbildung 3-3, konnten weitere Aufgaben auf der Ebene der *information seeking tasks* identifiziert und eingeordnet werden. Hierzu zählen insbesondere Aufgaben, die sich aus situativem Informationsbedarf ergeben, darunter Verifikationsprozesse zur Beurteilung potenziell relevanter Suchterme.

Darüber hinaus ermöglicht das Modell der iterativen Entwicklung verifizierter Suchprofile ein differenziertes Bild auf bekannte Aufgabenstrukturen auch im Kontext der *information searching tasks*, der Retrieval-Ebene, unter anderem durch die Definition verschiedener Prozesse der Durchführung von Testsuchen.

Entsprechend der Klassifikation von Byström und Hansen (siehe Abschnitt 2.2.4) ergibt sich mit den vorliegenden Ergebnissen ein granulares Bild der Aufgabenstruktur: Die Methodenlehre fokussiert sich weitestgehend auf Aufgabentypen, die sich auf *working task* und *information searching task* Ebene bewegen. Strategien und Praktiken im Umgang mit *working tasks* und *searching tasks*, hier also den klassischen Retrieval-Aufgaben, werden in der systematischen Übersichtsarbeit durch die Methodenlehre streng vorgegeben. Dies ist eine Besonderheit, die sich anderswo in dieser Form nicht findet. Sie gehören vorwiegend zu den bei Byström und Hansen so benannten automatisierten oder normalen Informationsverarbeitungs- und Entscheidungsaufgaben.

Authentische Aufgaben, deren Strukturen, Abläufe und Erfolgsaussichten nicht bekannt sind (als Beispiel können hier wiederum Aufgaben resultierend aus situativem Informationsbedarf genannt werden) sind ein wesentlicher Beitrag zum Verständnis des Informationsverhaltens.

Gänzlich neu ist die Definition des Modells „vom Synonym zum idealen Suchterm und Suchprofil“ (Abschnitt 5.1.1), auf dessen Grundlage Aufgaben und Prozesse der Entwicklung umfassender Suchprofile in Evaluierungsprozessen für Suchterme und Suchprofile einbettet werden können und damit iterativ und quellenspezifisch vom Kandidatenterm zum idealen Suchterm und fertigen Suchprofil führen.

5.2.2.2 Formaler und situativer Informationsbedarf

Ein deutlicher Unterschied des vorliegenden Modells zu den in den Abschnitten 2.2.1 bis 2.2.6 vorgestellten Modellen des Informationsverhaltens zeigt sich im Umgang mit Informationsbedarf, ein Konzept, das in der systematischen Literaturrecherche noch einmal differenziert betrachtet werden muss:

Mit der meist interdisziplinär durchgeführten Formulierung einer suchbaren Fragestellung liegt aus informationswissenschaftlicher Sicht ein formalisierter Informationsbedarf vor, der im Sinne des fundamentalen Taylor Modells (Taylor 1968) final als *compromised need* bezeichnet werden kann und bis zum Ende der Einarbeitungsphase zumindest Teile des Q1-Q4 Stufenmodells durchlaufen hat, vgl. hierzu Abschnitt 2.1.2. In der Praxis ist natürlich nicht klar, wie präzise dieser Informationsbedarf tatsächlich formuliert werden konnte und in welchem Stadium sich die Formulierung des Informationsbedarfs überhaupt befindet. Aufgabe des Informationsexperten ist es daher, diesen Informationsbedarf sukzessive zu konkretisieren. Das Ergebnis ist das wiederholt benannte verifizierte Suchprofil als Produkt.

Der **formale Informationsbedarf** des Informationsspezialisten bezieht sich während der Entwicklung von Suchprofilen primär nicht auf vollständige Inhalte oder Teilinhalte von Dokumenten, sondern auf die ideale *Repräsentation* eines Dokuments durch Synonyme im Hinblick auf diesen definierten Informationsbedarf, die „*well built clinical question*“, vgl. Abschnitt 1.1.2.

Inhalte, die diesen Informationsbedarf decken, werden erst nach verifizierter Finalisierung vollständiger Suchanfragen überhaupt analysiert und dann gezielt verwertet (vgl. hierzu das einführende Kapitel 1).

Mit Annahme des Auftrags zur Recherche ergibt sich im Verlauf gegebenenfalls ein, meist mit mangelndem Domänenwissen assoziierter **situativer Informationsbedarf**, der wiederum mit Theorien wie *sensemaking* (Dervin und Nilan 1986) verglichen werden kann und zur Notwendigkeit führt, die daraus resultierenden *gaps* zu überwinden.

Insgesamt ist es daher zulässig, nicht von einem initial durch eine vordefinierte klinische Fragestellung vorgegebenen Informationsbedarf zu sprechen, sondern diesen als dynamisch zu definieren.

Das ISP-Modell geht davon aus, dass ein erkannter Informationsbedarf auch verdrängt werden darf und Informationsverhalten erst akzeptiert werden muss, um Informationsverhalten

zu aktivieren (Marchionini 1995). In der professionellen Literaturrecherche ist dies nicht gültig: Die Methodenlehre legt mit seinem Vollständigkeitskriterium fest, dass jeder Informationsbedarf akzeptiert werden muss und im Idealfall auch voll befriedigt werden muss. Es gibt daher keine Klassifikation von als wichtig oder weniger wichtig wahrgenommenen Aufgaben und demzufolge auch keine dringenden oder weniger dringenden Aufgaben wie anderswo beschrieben, vgl. hierzu beispielsweise (Li und Belkin 2008).

Insbesondere situativer Informationsbedarf kann in der Praxis zu kognitiv problematischen und umfangreichen Suchprozessen führen und die systematische Vorgehensweise stören. Gleichzeitig spielen bei der systematischen Literaturrecherche im Gegensatz zum Rapid Review-Modell, wie es beispielsweise in Clinical Librarian-Tätigkeitsfeldern (siehe Abschnitt 3.6.3), zeitliche Rahmenbedingungen keine Rolle. Umfangreiche und möglichst vollständige Retrievalprozesse führen daher in der Praxis weiterhin schlicht zu ressourcenintensiven Screening-Prozessen.

5.2.3 Einordnung des Informationsverhaltens

Mittelpunkt der Studie war die Identifikation kontextuellen Informationsverhaltens in der Entwicklung umfassender Suchprofile als ein Teilaspekt systematischer Übersichtsarbeiten. Die Ergebnisse bestätigen dabei bereits bekannte, das Informationsverhalten beeinflussende Faktoren, zeigen aber insbesondere domänenspezifische Besonderheiten.

5.2.3.1 Prozesse der Informationssuche und -verarbeitung

Die in der informationswissenschaftlichen Fachliteratur beschriebenen typischen Aktivitäten wie beispielsweise in den Modellen von Ellis oder Marchionini benannt, werden in der jüngeren Fachliteratur immer wieder verwendet und adaptiert, um Prozesse des Informationsverhaltens dazustellen, vgl. beispielsweise (Russell-Rose et al. 2011; Russell-Rose und Tate 2013; Jürgens et al. 2014).

Kaum verwunderlich ist, dass diese teils sehr allgemein gehaltenen Termini auch für die Darstellung der hier vorliegenden Studienergebnisse mit kleineren deskriptiven Erläuterungen mehr oder weniger problemlos für die Beschreibung des Informationsverhaltens herangezogen werden konnten.

Es sollte an dieser Stelle ergänzt werden, dass die Erläuterungen klassischer Aktivitäten wie beispielsweise „*searching*“ in der Fachliteratur nicht immer einheitlich verwendet werden,

siehe beispielhaft (Ellis 1989; Belkin und Cool 2002). Aus diesem Grund müssen Begrifflichkeiten von Untersuchung zu Untersuchung neu hinterfragt, eingeordnet und definiert werden. Auf eine gänzliche Neudefinition unter anderer Terminologie kann jedoch dann verzichtet werden, wenn die Begrifflichkeiten im englischen oder deutschen Sprachraum konzeptuell verbreitet sind.

Information Seeking

Das Modell der iterativen Entwicklung verifizierter Suchprofile bewegt sich durch den spezifischen Informationsbedarf gegenüber bestehenden informationswissenschaftlichen Modellen jedoch auf einer Mikro-Ebene: Informationsbedarf meint in erster Linie Suchterme sowie weiterführende Informationseinheiten, er ist nicht beschränkt auf die Inhalte verfügbarer Information. Das Endprodukt dieser Teilaufgabe der systematischen Literaturrecherche ist ganz klar definiert als vollständiges Suchprofil, also eine weitestgehend vollständige Repräsentation des formalen Informationsbedarfs.

Aus diesem Grund werden die initial bei Ellis definierten und in der vorliegenden Arbeit referenzierten Aktivitäten im vorliegenden Modell um aufgabenspezifische Elemente ergänzt, die sich zwischen Verarbeitungsprozessen und klassischen Suchprozessen bewegen, darunter beispielsweise die Generierung suchbarer, datenbankspezifischer Artefakte.

Gleichzeitig finden sich insbesondere im Fall situativ auftretenden Informationsbedarfs, der aus dem Spannungsfeld aus Rechercheexpertise und mangelndem Domänenwissen resultiert, auch Elemente des Modells nach Ellis wie auch Meho und Tibbo (Meho und Tibbo 2003), dargestellt in Abschnitt 2.2.1 wieder, beispielsweise die Notwendigkeit einer Beschaffung von Dokumenten und deren spätere Analyse zum inhaltlichen Verständnis der klinischen Fragestellung.

Information Retrieval-Ebene

Der Mikro-Ebene ist eine andere Abweichung geschuldet: Kontrollprozesse in der Entwicklung verifizierter Suchprofile sind immer datenbankspezifisch und in den meisten Fällen sensitivitätsorientiert.

Für die Beobachtungsstudie wurden auf Information Retrieval-Ebene zusätzlich gezielt Beschreibungsmuster zur Suche auf Online-Datenbanken genutzt.

Zu den üblichen Techniken gehören insbesondere die Forschungsergebnisse von Marcia Bates (Bates 1979b), als auch die von Raya Fidel (Fidel 1991a, 1991b, 1991c), die in den Abschnitten 2.4.2 und 2.4.3 eingehend dargestellt wurden. Hierin werden Reaktionen auf drei große Entscheidungssituationen kategorisiert, die sich in der systematischen Literatursuche weniger als Problemsituationen, sondern mehr noch als klar definierte Arbeitsaufgaben finden.

Die Nutzung typischer Boolescher Datenbankfunktionalitäten wie der Einsatz von Trunkierung, Wildcards und Qualifiern ist in der Praxis noch ausgeprägter als beispielsweise in der Patentrecherche, vgl. (Azzopardi et al. 2010), dies gilt insbesondere für Methoden der *Query Expansion*:

Aufgrund methodologischer Anforderungen an die systematische Übersichtsarbeit stehen Information Retrieval Methoden, die zur Vergrößerung der Treffermengen beitragen, im Vordergrund. Dies bedeutet nicht, dass *Query Reformulation*-Prozesse mit dem Ziel der Verkleinerung von Treffermengen nicht präsent sind: Die Spezifizierung von Suchanfragen steht methodologisch dem Wunsch nach erschöpfenden Suchprofilen klar entgegen und kann daher nur dann durchgeführt werden, wenn das Dokumenten-Retrieval einer bekannten Ground Truth überprüft wurde und gesichert ist.

Die Notwendigkeit für die Nutzung von Information Retrieval Techniken zum Umgang mit „*off target*“ Treffermengen (also solche, die thematisch unpassend sind), wie bei Fidel beschrieben (Fidel 1991c), wird in der systematischen Übersichtsarbeit durch Strategien wie die Nutzung einer Ground Truth als ideale, weil valide Quelle für Suchterme als auch ideale Testmenge weitestgehend reduziert.

Hier zeigt sich ein Spezifikum der systematischen Literaturrecherche: Während beispielsweise in der Patentrecherche vorwiegend die Suchergebnisse aus Testsuchen verwendet werden, um daraus weiteres Suchvokabular abzuleiten (vgl. (Tseng und Wu 2008)) kann die systematische Literaturrecherche schon per definitionem auf die Methodik einer idealen Referenzmenge, der Ground Truth, zurückgreifen – auf der Suche nach einer möglichst umfangreichen Menge an Artikeln zu einer spezifischen Fragestellung ist davon auszugehen, dass zu einem Thema bereits Veröffentlichungen verfügbar sind.

5.2.3.2 Domänenspezifische Besonderheiten

In der Erklärung des Nutzerverhaltens finden sich einige domänenspezifische Aspekte. Aufgrund der in der Forschung bisher fehlenden Untersuchungen zum Informationsverhalten in

der professionellen medizinischen Literatursuche, ergänzt dieses Ergebnis die Forschungslandschaft. Gleichzeitig bestätigt es die Ergebnisse der Untersuchungen von Leckie et al. (Leckie et al. 1996), siehe Abschnitt 2.2.5: Das Informationsverhalten muss im Kontext des Betätigungsfelds von Akteuren betrachtet werden und wird hier von einer Vielzahl von Faktoren beeinflusst.

Qualität und Ertrag von Information

Anders als bisherige Modelle stellt das hier vorgestellte Modell Aktivitäten und daraus resultierende Strategien in den Kontext verfügbarer Information und deren inhaltlich-methodischer Qualität: Die Verarbeitung verfügbarer Information hängt nicht nur von der situativen Zielsetzung des Akteurs, sondern auch von der methodologischen Qualität zugrundeliegender Studien sowie dem erwarteten Ertrag eines Dokuments ab, siehe Abschnitt 5.1.2.1.

Mit dieser Abhängigkeit lässt sich wahrscheinlich auch begründen, dass sich keine allgemeingültigen Prozessmodelle zur Identifikation von Synonymen und weiterführenden Informationseinheiten festmachen ließen, wie dies beispielsweise im allgemein gehaltenen ISP-Modell von Marchionini (Marchionini 1995) der Fall ist. In Verarbeitungsprozessen zeigten sich mit den Ground Truth-Loops und anderem vielmehr ressourcenspezifische Muster, eine Darstellung, die sich auch an anderer Stelle findet, vgl. (Russell-Rose und Tate 2013, Kap. 2). Sie zeigen in ihrer Gesamtheit einen umfangreichen iterativen und vor allem facettenreichen Prozess aus unterschiedlichsten Aktivitäten im Kontext verschiedener Informationseinheiten.

Die Abhängigkeit von verfügbarer Information und daraus resultierende Informationsverarbeitung ist domänenspezifisch:

Während die Verwendung „idealer Referenzmengen“ eine Strategie ist, die bereits zu Beginn der Online-Datenbanken verwendet wurde, um Queries zu bewerten (vgl. (Markey und Cochrane 1978)) und Testmengen auch übliches Mittel zur systemseitigen Evaluierung von Information Retrieval-Systemen sind (siehe Abschnitt 2.3.2.5), gewinnt die Ground Truth vor allem durch die Berücksichtigung von den Originalartikeln zugrundeliegenden Studienqualitäten wie beispielsweise der randomisierten klinischen Studie eine neue Dimension. Die in Abschnitt 3.1.1 eingehend erläuterte Evidenzpyramide ist daher als wichtiger Motor des Informationsverhaltens zu interpretieren.

Methodenlehre

Bereits die Existenz einer Methodenlehre kann als Domänenspezifikum gesehen werden, findet sie sich doch in keiner anderen Form der professionellen Literaturrecherche in dieser strikten, institutionell vorgegebenen Strenge.

Die Ergebnisse bestätigten daher, dass auch individuelles Informationsverhalten in der Durchführung systematischer Literaturrecherchen vor allem in der evidenzbasierten Medizin *immer* im Kontext der Methodenlehre zu sehen ist. Sie gibt für einen Großteil der Nutzer dezidierte Aufgaben vor (vgl. (Chandler et al. 2013)), für andere schafft sie durch Handreichungen und Methodologie (siehe Abschnitt 3.4.3) einen Rahmen der Sicherheit, den hohen Ansprüchen genügen zu können.

Darüber hinaus ist deren Anwendung auch die Grundvoraussetzung dafür, die aus systematischen Übersichtsarbeiten resultierenden Fachartikel möglichst hochkarätig, beispielsweise in der Cochrane Library überhaupt veröffentlichen zu können.

Es ist anzunehmen, dass insbesondere Initiativen wie das MECIR Statement (Chandler et al. 2013) diese Entwicklung in Zukunft noch verstärken.

Kritisch sei an dieser Stelle angemerkt, dass die Methodenlehre von Natur aus keinerlei Hinweise zum Umgang mit situativem Informationsbedarf geben kann. An dieser Stelle zeigen sich dann auch keinerlei methodenspezifische Verhaltensmuster, entsprechende Lösungsstrategien sind weniger stringent.

In der Bestätigung geeigneter *Information Retrieval Tactics*, wie sie beispielsweise bereits bei Markey und Cochrane (Markey und Cochrane 1978), siehe Abschnitt 2.4.1.1, wird deutlich, dass Informationsspezialisten generell dazu tendieren, überwiegend die während ihrer Ausbildung gelernten oder von Kollegen überlieferten Handgriffe und Techniken anzuwenden. Sie sind in ihrem Informationsverhalten, insbesondere in klassischen *Search Tasks* durch eine eher konservative Haltung charakterisierbar

Sprachliche Diversität

Die Untersuchung bestätigt die in der Fachliteratur bereits bekannte Tatsache, dass die Durchführung systematischer Literaturrecherchen von der Wahl der Fragestellung und dem Variantenreichtum sprachlicher Konzepte abhängt. Ein gutes Beispiel hierfür sind komplexe Interventionen, beschrieben in Abschnitt 3.3.2.2.

Die Wahrscheinlichkeit, eine systematische Übersichtsarbeit „reibunglos“ durchführen zu können, steigt also, wenn es sich um eine klassische interventionelle Fragestellung handelt, deren Konzepte eindeutig durch Schlüsselbegriffe und Synonyme repräsentierbar sind. Umgekehrt führen sprachlich komplexe Konzepte häufig zu Informationsverhalten, das entweder durch situativen Informationsbedarf gestört wird oder aber durch die Fülle an Fachterminologien in einem erhöhten kognitiven Workload resultiert.

Probleme wie der Sprachenbias führen des Weiteren dazu, dass Suchprozesse nicht auf englischsprachige Dokumente und sprachlich spezialisierte Informationsressourcen beschränkt bleiben dürfen. Auch dies erschwert das Auffinden geeigneter Suchterme zusätzlich. Dies allerdings ist kein domänenspezifisches Charakteristikum der systematischen Übersichtsarbeit, sondern findet sich auch anderswo in der professionellen Suche wieder, siehe u.a. (Bonino et al. 2010; Cornacchia et al. 2013).

5.2.3.3 Profession

Bereits das Modell von Leckie et al. macht deutlich, dass Informationsverhalten durch individuelle demographische Faktoren bestimmt ist, darunter Alter, Berufserfahrung und Arbeitsumfeld (Leckie et al. 1996). Leider lässt die vorliegende Studie abseits des Faktors Methodenlehre keine validen Schlüsse zu, auch wenn zu vermuten ist, dass Alter und Berufserfahrung durchaus Faktoren sind, die Tätigkeiten zur Routine werden lassen und insgesamt zu anderen Suchprozessen führen.

Nicht von der Hand zu weisen ist in jedem Fall der Faktor der Rechercheprofession gegenüber der Domänenexpertise. Bereits Byström und Järvelin klassifizieren die Notwendigkeit um das Wissen problemspezifischer Information im Gegensatz zu Domänenwissen und Problemlösungsinformation als wesentliche Einflussfaktoren für Informationsverhalten, siehe Abschnitt 2.2.4.

Das Spannungsfeld zwischen dem für die Literaturrecherche erforderlichen Domänenwissen, beispielsweise zur Beurteilung verwendeter Fachterminologien, sowie der Rechercheexpertise ist auch ein wesentliches Problem im Informationsverhalten der systematischen Literaturrecherche. Datenbankbasierte Suchprozesse sind, der Aufgabenteilung in der systematischen Übersichtsarbeit entsprechend, Aufgabe des Informationsexperten.

Die meisten dieser Akteure sind professionell geschult und durchlaufen häufig methodologisch geführte Workshops, in denen die seit Jahren etablierten und dort aufgrund persönlicher Rechercheexpertise von Dozenten ausgebauten Inhalte vermittelt werden, vgl. beispielsweise die Schulungsmaterialien von Motschall (Motschall 2010, 2014). Mehr noch ist anzunehmen, dass für vor allem institutionell verankerte Informationsexperten eine Schulung anhand entsprechender Handbücher eine notwendige Qualifikation darstellt.

Im Cochrane Handbuch heißt es zur Methodenlehre der systematischen Übersichtsarbeit, die in Abschnitt 3.4.3 dargestellt wurde, (Lefebvre et al. 2011):

“Searches for systematic reviews aim to be as extensive as possible if[...]. It is, however, necessary to strike a balance between striving for comprehensiveness and maintaining relevance when developing a search strategy. Increasing the comprehensiveness (or sensitivity) of a search will reduce its precision and will retrieve more non-relevant articles.”

wie auch:

“[...] indexers are not always experts in the subject areas or methodological aspects of the articles that they are indexing. In addition, the available indexing terms might not correspond to the terms the searcher wishes to use.”

- beides Aspekte, deren sich Nutzer durchaus bewusst sind, die aber auf der anderen Seite auch die Rolle des Akteurs beleuchten:

Der Akteur ist nicht nur Informationsexperte, der den Auftrag hat, „das Beste“ bezüglich einer vordefinierten Fragestellung aus einem System herauszuholen. Er ist der Spezialist für die eingesetzten Informationsressourcen und er ist auch der, der sich aller potenziellen Fehlerquellen, die beispielsweise durch technische und menschliche Gegebenheiten eines Indexierungsprozesses auftreten können, bewusst sein muss. So ist er damit beauftragt, Wege zu finden,

diese Probleme zu überbrücken. Dies ist aus informationswissenschaftlicher Seite eigentlich eher Aufgabe eines Retrieval-Systems und nicht des Nutzers. Aus Sicht dieser Nutzer jedoch wird – so zumindest die Ergebnisse der Studie – dieses Problem nicht hinterfragt.

Die während des Entwicklungsprozesses vorwiegend genutzten professionellen Benutzeroberflächen Boolescher Datenbanken unterstützen die Entwicklung und Dokumentation komplexer Suchanfragen als iterative Reformulierung der Suchanfrage, jedoch kaum die Überbrückung von darüber hinaus gehendem und in der Methodenlehre nicht berücksichtigtem domänenspezifischem, gegebenenfalls auch situativem Informationsbedarf.

In der Praxis ist dieses Problem bekannt und wird durch Workarounds entsprechend umgangen, in erster Linie durch Kollaboration:

Auch wenn die Literatursuche selbst von Informationsexperten durchgeführt wird, ist der Gesamtprozess als interdisziplinäre und kollaborative Aufgabe zu interpretieren, bei der den verfügbaren Domänenexperten insbesondere in der Planungsphase konkrete Aufgaben zugewiesen werden.

Eine noch engere Zusammenarbeit von Domänenexperten und Rechercheexperten kann in der Praxis zu qualitativ besseren Suchanfragen führen. Vor allem die Komplexität iterativer und filigraner Prozesse macht es in der Zukunft erforderlich, kollaborative Arbeitsprozesse zwischen Domänenexperten und Rechercheuren besser zu unterstützen (Swinkels et al. 2006). Zu ähnlichen Ergebnissen kamen auch entsprechende Untersuchungen in der Patentrecherche, vgl. (Hansen und Järvelin 2005).

Das Ziel der Forschung war es in den vergangenen Jahren, Information Retrieval-Systeme für Recherche-Laien nutzbarer zu machen, was an eine verbesserte Retrieval-Funktionalität mit diversen, nutzerfreundlichen Möglichkeiten zur Formulierung von Suchanfragen und Methoden beziehungsweise Interfaces zur explorativen Erschließung des Dokumentenbestands gekoppelt ist, beispielsweise die für Nutzer praktikable Formulierung von Freitext-Suchanfragen, Ranking-Methoden oder auch Facetten-Suchen. Sie unterstützt damit vor allem Rechercheprozesse von Domänenexperten. Sichtbar wird dies übrigens auch in der Gestaltung nutzerfreundlicher User-Interfaces auf bibliographischen Fachdatenbanken. Diese verfügen neben dem professionellen Suchmodus auch über einfache Suchmöglichkeiten, bei denen Suchprozesse mittels Facettensuche oder Ranking-Verfahren unterstützt werden, siehe Abschnitt 3.4.2.3.

In der professionellen Suche, insbesondere in der systematischen Literaturrecherche bleiben Probleme bestehen, da die Systeme zwar auf der einen Seite die professionelle Recherche unterstützen, die für den Domänenexperten gedachten Funktionen auf der anderen Seite aufgrund des Qualitätsanspruchs für den Informationsexperten nicht problemlos nutzbar sind, beziehungsweise die Reproduktion von Suchprozessen deutlich erschweren.

5.2.3.4 Datenbankspezifische Kontextvariablen

Die Nutzung bibliographischer Datenbanken und Boolescher Retrieval Techniken ist in der systematischen Literaturrecherche dominant. So überrascht es nicht, dass das Modell auf Information Retrieval Ebene in der Informationssuche vorwiegend Verhalten bestätigte, das bereits in den Anfangszeiten online verfügbarer Information Retrieval-Systeme bekannt war, zum anderen überrascht es nicht, dass sich Akteure meist sehr stringent an die über ihre Arbeitgeber vorgegebenen Richtlinien halten.

Die Untersuchungen bestätigten eine informationswissenschaftlich bekannte Tatsache: Die Nutzer stellen aufgrund der für sie idealen Dokumentierbarkeit und Kontrolle strukturierter Suchen die starre Technik Boolescher Datenbanken nicht in Frage (vgl. (Tait 2013)).

Sie reagieren auf die zunehmende Datenflut mit der Perfektionierung ihrer eigenen Technik (Bramer und Jonge 2015). Zusätzlich werden Methoden objektiven Peer Reviewings von Suchprofilen vorgeschlagen, vgl. hierzu (Sampson et al. 2008; Sampson et al. 2009), bei denen mehrere Rechercheexperten Suchprofile unabhängig voneinander methodologisch bewerten.

Wie sehr sich die Ausbildung von Informationsspezialisten sowohl an Funktionalität und Grundprinzip Boolescher Datenbanken als auch der Methodenlehre systematischer Übersichtsarbeiten orientiert, wurde bereits in den Abschnitten 3.4.2 sowie 3.4.3 deutlich gemacht. Die Ergebnisse bestätigen diese Abhängigkeit, mehr noch zeigen sie die Symbiose zwischen dem Konzept „systematische Übersichtsarbeit“, ihrer Methodenlehre und der Technologie Boolescher Datenbanken. Vor allem die fachlichen Diskussionen über die Nutzung von Google Scholar zeigen (siehe Abschnitt 3.5.1.2) wie sehr sich Methodik und Methodenlehre der systematischen Literaturrecherche und das Konstrukt des Booleschen Retrieval bedingen. Grundsätzlich wird die Qualität von Google Scholar positiv eingeschätzt. Zur Sicherung der Methodik jedoch wird auch eine wissenschaftliche Suchmaschine nach Methodenlehre der erschöpfenden Suchstrategie (im Sinne einer idealen Repräsentation des formalen

Informationsbedarfs) mit klassischen Booleschen Datenbankabfragen strukturiert und systematisch durchsucht.

Im Hinblick auf die systematische Übersichtsarbeit lassen sich im Kontext des Information Retrieval mehrere ganz grundlegende Probleme festmachen, die in der Literatur zum Teil auch beschrieben sind. Sie sind in erster Linie der Anforderung nach Vollständigkeit und der Methodenlehre geschuldet.

Vollständigkeitsbeurteilungen

Der Umgang vor allem mit der methodologisch vorgeschriebenen Vollständigkeitsbeurteilung von *Queries* als ideale und möglichst objektiv ermittelte Repräsentation der Forschungsfrage durch eine meist umfangreiche Sammlung an Suchtermen ist für die meisten Nutzer der problematischste Aspekt der systematischen Literaturrecherche. Dieser wird anhand der Untersuchungen vorwiegend im Kontext terminologisch breiter, komplexer Interventionen oder Outcomes beschrieben, dürfte in der Praxis aber auch auftreten, wenn es mit Hilfe multilingual ausgerichteter Suchanfragen um die Überwindung von Sprachbiases (siehe Abschnitt 3.1.2) geht. Neben der semantischen Beurteilung von Kandidatentermen geht es in der Praxis darum, möglichst alle in der Fachliteratur verwendeten Termini aufzufinden.

Erschwert wird dies durch für den Akteur nicht immer reproduzierbare Indexierungsmechanismen einzelner Fachdatenbanken sowie eine deutliche Diversität in der Qualität und den Indexierungspraktiken verfügbarer Thesauri, vgl. Abschnitt 3.4.2.1.

In der methodologischen Fachliteratur sind Strategien zur Bewältigung datenbankspezifischer Unterschiede bei der Entwicklung komplexer Suchanfragen daher seit Jahren Forschungsthema, vgl. hierzu exemplarisch (Boynton et al. 1998; Allison et al. 1999; Agoritsas et al. 2012; Boeker et al. 2012).

Die Notwendigkeit datenbankspezifischer Suchstrategien zeigt sich natürlich insbesondere in der Verwendung von Suchfiltern (vgl. hierzu Abschnitt 3.4.3.2), die in den letzten Jahren zu einer Vielzahl verfügbarer Produkte geführt haben (Haynes 2004, 2005; Haynes et al. 2005; Lefebvre et al. 2011).

Abschätzung des Recalls

Ein weiterer Punkt betrifft die zur Vollständigkeitsbeurteilung von Treffermengen notwendige Abschätzung des Recalls durch die Abarbeitung von Treffermengen. Wir wissen, dass Nutzer den tatsächlichen Recall in der Praxis meist überschätzen (Bache 2011). Für eine tatsächliche Abschätzung sind Akteure entweder gezwungen, riesige Treffermengen zu screenen oder sich auf proprietäre Ranking Algorithmen einzelner Anbieter zu verlassen. Auch dies ist ein Thema, das beispielsweise aus der Patentrecherche bekannt ist, vgl. (Tait 2013).

In der systematischen Übersichtsarbeit sind zur Lösung dieses Problems in der Literatur verschiedene Ansätze beschrieben:

Die *Capture-Mark-Recapture* Technik, im Deutschen auch Rückfangmethode, findet sich ursprünglich in der Biologie. Hierbei wird eine Stichprobe einer zu beobachtenden Population gefangen, markiert und danach freigelassen. Anhand der Anzahl markierter Tiere einer neuen Stichprobe aus der Gesamtheit kann auf den Anteil der zu beobachtenden Population an der Gesamtpopulation geschlossen werden. Das Prinzip wurde mehrfach in der systematischen Literaturrecherche angewandt und dort auch empirisch als mögliches Stop-Kriterium nachgewiesen, vgl. (Spoor et al. 1996; Bennett et al. 2004; Kastner et al. 2009). Kritisiert wird vielfach allerdings, dass diese und andere Methoden bisher nicht evidenzgesichert sind, vgl. (Booth 2010).

In Summe ergibt sich für den Informationsspezialisten oder Bibliothekar die Schwierigkeit im Umgang mit Booleschen Fachdatenbanken häufig nicht aus der direkten Interaktion mit dem System – dieser ist erlernt und Routine: Es ist, wie Hjørland betont, eine Aufgabe, die Stärken und Schwächen von Systemen zu kennen. Probleme ergeben sich daher vielmehr aus der Interaktion mit der „Welt verfügbaren Wissens“ (Hjørland 2015) und deren Repräsentation in verfügbaren Datenbeständen.

Die in bibliographischen Fachdatenbanken vorliegenden Informationen sind zwar manuell indexiert und damit prinzipiell auffindbar. Struktur und Wissensrepräsentation der für die Beantwortung der klinischen Fragestellung relevanten Inhalte sind jedoch im Text selbst, beispielsweise den Abstracts, eben nicht so formalisiert repräsentiert wie es formalisierte Suchanfragen vorgeben. Letztendlich bewegt sich die Informationssuche nach der Methodenlehre der EBM damit in einem konzeptionellen Bruch, der in der systematischen Übersichtsarbeit

bisher nur durch massiv ressourcenintensive, intellektuelle Screening-Prozesse überbrückt werden kann.

5.3 Zusammenfassung

Das vorliegende Kapitel verallgemeinerte die Ergebnisse der empirischen Studie in einem allgemeinen Modell zur iterativen Entwicklung verifizierter Suchterme und stellte dies in einen informationswissenschaftlichen Kontext. Insbesondere sollte dabei Information Retrieval wie von Ingwersen und Järvelin verstanden, als Interaktion und im Kontext verschiedener Rahmenbedingungen interpretiert werden.

Die vorliegende Arbeit schließt an dieser Stelle eine Lücke: Bisher wurde das Informationsverhalten von Rechercheexperten in der Durchführung systematischer Übersichtsarbeiten weder in der der Bibliothekswissenschaften noch in der Informationswissenschaft eingehend empirisch untersucht.

Ergänzend zu den in der Methodenlehre bekannten Aufgabenstrukturen (siehe Abschnitt 3.3.2) unterstreicht das Modell den iterativen Charakter der systematischen Literaturrecherche, beleuchtet aber auch bisher nicht formulierte Aufgabenstrukturen, ressourcenspezifische Suchtaktiken, die Fokussierung auf datenbankspezifische Suchprozesse sowie das Spannungsfeld zwischen Domänenexpertise und erforderlicher Rechercheexpertise.

Aus den Untersuchungen wird deutlich, dass die systematische Übersichtsarbeit nicht nur methodologisch eng mit den per definitionem vorgegebenen Anforderungen und den in der Fachliteratur verfügbaren, daraus resultierenden Best Practices verknüpft ist, sondern im Kontext verfügbarer Informationsressourcen und deren Anbieter gesehen werden muss. Als professionelle Tätigkeit, die in dieser Form fast ausschließlich von Informationsexperten durchgeführt wird, zeigt die systematische Literaturrecherche hier einige charakteristische Gemeinsamkeiten und Probleme, die auch in der professionellen Patentrecherche bekannt sind und sich sicherlich am Vollständigkeitsanspruch festmachen lassen.

Natürlich muss die vorliegende Arbeit auch kritisch betrachtet werden.

Die für die empirische Untersuchung verfügbaren Daten mögen ein prototypisches Bild abgeben, das jedoch auf Basis weiterer Forschung mit weitaus mehr Daten empirisch belastbarer und noch granularer gestaltet werden könnte und müsste. Gleichzeitig ist auch die Auswahl der Studienteilnehmer kritisch zu hinterfragen: durch die Teilnahme am Cochrane Colloquium wurden einige Teilnehmer persönlich rekrutiert, die Kontaktaufnahme innerhalb der Mailing-Liste erfolgte ebenfalls über einen persönlichen Kontakt. Auch wenn die IMRG Mailing-Liste nicht ausschließlich Mitglieder der Cochrane Collaboration beinhaltet, wäre es wünschenswert gewesen, weitere Institutionen, die sich der Durchführung systematischer Übersichtsarbeiten in der EBM verschrieben haben, ähnlich repräsentiert zu haben.

Eine interessante, weil sehr authentische Quelle böten hierfür die Mailinglisten der IMRG Gruppe der Cochrane Collaboration sowie die Expert-Searching Mailingliste (siehe http://pss.mlanet.org/mailman/listinfo/expertsearching_pss.mlanet.org). Leider waren die dort verfügbaren Archive für eine Auswertung nicht verfügbar, da wiederholte Anfragen seit Dezember 2015 nicht beantwortet wurden. Für weitere Untersuchungen wäre es wichtig, die Betreiber beider Listen noch einmal anzufragen und die Wichtigkeit der dort verfügbaren Informationen für weitere Forschungen zu betonen. Alternativ wäre es gegebenenfalls auch möglich, die Beteiligten einzelner Threads exklusiv um ihre Einwilligung zu bitten.

Das Modell beschränkt sich des Weiteren auf die Entwicklung strukturierter Suchprozesse in bibliographischen Datenbanken. Dies ist – den systematischen Verzerrungen in der medizinischen Literaturlandschaft geschuldet – lediglich ein Teilaspekt der für die systematische Übersichtsarbeit notwendigen Teilschritte, siehe Abschnitt 3.2. Für die Gewinnung eines vollständigen Bildes sind in jedem Fall ergänzende Untersuchungen zum Thema der Informationsgewinnung aus klinischen Studienregistern als auch die Nutzung weiterführender Informationsressourcen wie beispielsweise Google Scholar unumgänglich. Parallel dazu müssten abweichende Besonderheiten für die Durchführung nicht-interventioneller systematischer Übersichtsarbeiten gesondert herausgearbeitet werden.

Dies alles war nicht Teil der vorliegenden Arbeit und könnte in weiteren Forschungsprojekten zur Vervollständigung des Gesamtbilds beitragen.

Ein weiterer Aspekt, den die vorliegende Arbeit nicht abdeckt, ist der der zugrundeliegenden Fragestellung, darunter zum einen nicht-interventionelle Reviews, zum anderen aber auch die in der Forschung zunehmend interessanten qualitativen Fragestellungen, vgl. hierzu Abschnitt

3.3.1.2. Da hier zum einen qualitative Analysemethoden zum Einsatz kommen und auch strukturierte Suchprozesse häufig nicht anwendbar sind, ist davon auszugehen, dass sich hier ein vollkommen anderes Informationsverhalten zeigt als in der quantitativen Forschung. Auch dies wäre ein Thema, das für ein abschließendes Gesamtbild untersucht werden sollte.

Das abschließende Kapitel widmet sich ausgehend von der aktuellen Forschungslage den einleitend formulierten Forschungsfragen. Es spannt damit den Bogen zurück zur Ausgangslage und gibt Anregungen für innovative und nutzerfreundliche Retrieval-Umgebungen für die systematische Literaturrecherche.

6 Das ideale Retrieval-System?

In Kapitel 1 wurde angedeutet, dass von Informationsspezialisten seit Jahren Methoden entwickelt und erprobt werden, mit deren Hilfe die Durchführung systematischer Übersichtsarbeiten erleichtert und beschleunigt werden kann.

Bilden diese gemeinsam mit den Ergebnissen der vorliegenden Arbeit eine solide theoretische Basis für die Entwicklung eines „idealen“ Retrieval-Systems?

6.1 Ausgangssituation

Bereits in Abschnitt 3.4.5.2 wurden Text Mining-gestützte Ansätze in der Entwicklung von Suchprofilen vorgestellt. Diese spielen in der Methodenlehre eine zunehmend große Rolle, einen Überblick bieten unter anderem (Ananiadou et al. 2009; Cohen et al. 2010; Thomas et al. 2011; Shemilt et al. 2013). Grundsätzlich folgt die Methodenlehre hier einer Tendenz, die sich auch im ähnlichen Patentretrieval findet, vgl. hierzu unter anderem (Tseng et al. 2007). So entstand im Laufe der vergangenen Jahre eine Reihe an Softwarelösungen zur Automatisierung von Teilprozessen, die Text Mining-gestützte Verfahren wie Automatic Term Recognition (ATR), Document Clustering, Document Classification und Document Summarization implementieren, um Teilprozesse zu unterstützen (Thomas et al. 2011). Mit der Website SR Toolbox steht eine für den Endnutzer übersichtliche Zusammenstellung aktueller Entwicklungen zur Verfügung (Marshall 2016).

Aktuell nutzbare Tools reizen jedoch nur einen Bruchteil verfügbarer Technologien überhaupt aus: Die in der bibliographischen Fachliteratur in den vergangenen Jahren vorgestellten und teils etablierten Text Mining-gestützten Lösungsansätze zur Ermittlung von Suchtermen und dem Aufbau erschöpfender Suchprofile basieren meist auf dem so genannten Bag-of-Words Prinzip, u.a. bei (Frantzi et al. 2000; Ananiadou et al. 2009; Hausner et al. 2012; Slater 2014; Hausner et al. 2015), beispielhaft wurden hierfür in Abschnitt 3.4.5.2 zwei Werkzeuge vorgestellt. Die entsprechenden Text Mining Ansätze wie *Automatic Term Recognition* (ATR) geben zwar Aufschluss über potenziell interessante Suchbegriffe, jedoch keinerlei Auskunft über deren semantische Beziehung, wie folgende Einschätzung der ATR Software TerMine verdeutlicht (Thomas et al. 2011):

"[...] it expands the review in favour of the literature that uses the same language as the documents that have already been found. This method will not identify cross-disciplinary research very well. [...]. This method on its own will not assist the reviewer in identifying literatures that use different words to describe the same concepts"

So erfüllen sie zwar die in der Methodenlehre vielfach geforderte Objektivität in der Gewinnung von Suchtermen, können jedoch das Problem der Semantik und Sprachenvielfalt nicht lösen und werden aus technischer Sicht heute auch nicht mehr als adäquate Lösung angesehen, sondern durch semantische Methoden abgelöst, vgl. (Jiang 2012; Crain et al. 2012).

Erst in den vergangenen zwei bis drei Jahren werden auch für die Text Mining-gestützte Durchführung systematischer Übersichtsarbeiten vermehrt Topic Modeling-Ansätze hinsichtlich ihrer Eignung und Validität untersucht, die im Übrigen im Hinblick auf die Sensitivität von Suchanfragen als ein wichtiges Qualitätskriterium in der Durchführung systematischer Übersichtsarbeiten gegenüber Bag-of-Words Modellen deutlich besser beurteilt werden, vgl. (Jonnalagadda und Petitti 2013; Miwa et al. 2014; Mo et al. 2015). Es gibt jedoch in der Literatur keinen Hinweis darauf, dass diese Entwicklungen bereits evidenzbasiert evaluiert wurden, was wiederum ein entscheidendes Kriterium nicht nur für die Akzeptanz dieser Verfahren ist, sondern als Qualitätskriterium gleichzeitig deren gefahrlosen Einsatz in der systematischen Übersichtsarbeit rechtfertigt.

Es sei an dieser Stelle auf ein generelles Problem hingewiesen: Häufig werden methodisch neue Verfahren an eigenen Testmengen, beispielsweise einer mittels systematischer Übersichtsarbeiten ermittelten Ground Truth, evaluiert und mit herkömmlichen Methoden verglichen, vgl. beispielsweise (Hausner et al. 2012; Hausner et al. 2015). Keines der in der Methodenlehre getesteten Verfahren jedoch verwendet bisher größere und vor allem thematisch allgemein gehaltene Testmengen, häufig finden sich auch Fehler in der Bewertungskonzeption, vgl. (Olorisade et al. 2016).

Die als besonders zeitaufwändig bekannte Aufgabe des Screenings großer Treffermengen (nicht nur im finalen Analyseprozess, sondern bereits während der Entwicklung von Suchprofilen) bildet neben der Entwicklung von Suchprofilen den nächsten Problemkomplex der systematischen Übersichtsarbeit. Sie dient während der Suche der Entscheidung für das weitere Vorgehen, beispielsweise dem Abbruch von Suchprozessen, während des finalen Screening-

Prozesses der Entscheidung über die Aufnahme von Studien in den finalen Bewertungsprozess. Die Reduktion des Workload im Screening-Prozess als klassisches Klassifizierungsproblem nimmt in der entsprechenden Forschungslandschaft großen Raum ein, vgl. beispielhaft (Bekhuis und Demner-Fushman 2010, 2012; Wallace et al. 2010; Shemilt et al. 2013; Miwa et al. 2014). Die Idee scheint zu sein, in der Praxis weiterhin auf möglichst erschöpfende Suchprofile zu setzen, dabei die Spezifität vollkommen außer Acht zu lassen und den Workload der daraus resultierenden, noch umfangreicheren Screening-Prozesse zu automatisieren. Der Vorteil dieser Vorgehensweise liegt klar erkennbar darin, dass Screening-Prozesse durch die Möglichkeit eines Exports auch riesiger Treffermengen außerhalb der Datenbanken stattfinden können und damit eine Unabhängigkeit vom Datenbankanbieter gewährleistet ist.

Federführende Institutionen scheinen sich derzeit weiterhin auf die Verbesserung intellektueller Prozesse zu verlagern. In den vergangenen Jahren etabliert sich vor allem im Screening-Prozess mit *Crowd-Sourcing* eine neue Entwicklung. Hierbei werden typische Aufgaben an eine beliebige Auswahl von Nutzern herangetragen, die damit als Teil einer Community aktiv an einem Projekt mitwirken, vgl. hierzu (Ranard et al. 2014). Mit dem EMBASE-Project wurden 2014 erste konkrete Versuche vorgestellt, Web-User in die Identifikation randomisierter klinischer Studien mit einzubeziehen, vgl. (Dooley et al. 2014; Noel-Storr et al. 2015a, 2015b). Dieses Projekt ist seit 2016 auch online verfügbar (The Cochrane Collaboration 2016a). Crowd-Sourcing Modelle wurden in der Medizin bereits an verschiedenen Stellen erfolgreich eingesetzt, darunter in Szenarien wie Problemlösung, Datenverarbeitung, Kontroll- und Überwachungsprozesse sowie bei Umfragen. Die positiven Erwartungen an den Nutzen, Erfolg und die Qualität wurden in einem 2014 veröffentlichten Review grundsätzlich bestätigt (Ranard et al. 2014), anderswo wurden diesbezüglich vor allem in der Praxis einige wesentliche Hürden identifiziert (Sims et al. 2014). Derzeit jedoch fehlen auch an dieser Stelle evidenzbasierte Ergebnisse zur Eignung dieser Methoden.

Leider sind moderne Information Retrieval – und Text Mining-Technologien in der Praxis bisher also nicht flächendeckend etabliert: Auch, wenn insbesondere Text Mining-Verfahren seit Jahren als unterstützende Maßnahme gefordert werden und ihr Einsatz entsprechend beforscht wird, haben sich lediglich einige wenige Institutionen der Entwicklung innovativer Text Mining-gestützter Verfahren verschrieben und setzen diese auch ein, beispielsweise das IQWiG, vgl. (Hausner et al. 2012), oder das EPPI Centre, vgl. (Thomas et al. 2011; O'Mara-Eves et al. 2013; O'Mara-Eves et al. 2015).

Hervorzuheben ist die Tatsache, dass es sich bei den verfügbaren Lösungen nicht um native Implementierungen einzelner Retrieval-Systeme handelt oder auch um Meta-Systeme, die Zugriff auf verschiedene Fachressourcen bieten. Unter den angebotenen Lösungen finden sich ausschließlich Tools, die frei verfügbar, teils aber auch kommerziell genutzt werden können, um die Suche in einzelnen bibliographischen Datenbeständen unterstützen, vgl. (Paynter et al. 2016). Es bleibt unklar, ob die Verfügbarkeit dieser Werkzeuge oder der Zweifel an deren Einsatz dafür ausschlaggebend sind. Eine Tendenz zur Skepsis gegenüber nicht eindeutig reproduzierbaren Daten sowie nicht evidenzbasiert hergeleiteten Methoden ist aus den Ergebnissen der Hauptstudie mindestens zu vermuten.

Text Mining-Verfahren zur Unterstützung systematischer Reviews sind also nicht neu, was fehlt, ist, sie als festen Bestandteil im Gesamtprozess zu etablieren oder sie für die Masse der Nutzer verfügbar zu machen und damit auch ihre Akzeptanz zu steigern. Dabei sollten Probleme, die durch Text Mining-Verfahren gelöst werden könnten, nicht mehr isoliert, sondern in seiner Gesamtheit gesehen werden, sie sollten zudem den Informationsspezialisten als Akteur sowie externe Rahmenbedingungen besser berücksichtigen.

6.2 Anregungen und weitere Forschungsaspekte

Wie könnte sie unter diesen Gegebenheiten nun aussehen, die ideale Retrieval-Umgebung für strukturierte Suchprozesse systematischer Übersichtsarbeiten in der Medizin?

Eine Frage, die sicherlich zunächst diskutiert werden muss, ist:

Wie ist der Einsatz Boolesche Retrieval-Systeme, die aus als aus informationswissenschaftlicher Sicht längst flächendeckend durch alternative Systemkonfigurationen abgelöst wurden, im Kontext systematischer Übersichtsarbeiten einzuschätzen?

Angesichts der Tatsache, dass wie in Kapitel 2 dargestellt, Boolesches Retrieval aus informationswissenschaftlicher Sicht überholt scheint, weil mit *partial match* Systemen und Ranking-Algorithmen andere, nutzerfreundliche und leistungsstärkere Systemkonfigurationen zur Verfügung stehen, ist man geneigt, die grundsätzliche Empfehlung für eine grundlegende Neukonzeption bibliographischer Fachdatenbanken wie auch Expert Search-Systeme auszusprechen. Suchprozesse bewegen sich heute in einer Welt „Google-ähnlicher“ Systeme, sodass

sich auch die Frage stellt, ob der Thesaurus, als Entwicklung klassischen Information Retrievals in dieser modernen Welt noch einen Platz hat (Hjørland 2016). Wie angesprochen, stehen mit *partial match* Technologien und Ranking-Algorithmen moderne Systemkonfigurationen zur Verfügung, deren Einsatz im Kontext der systematischen Übersichtsarbeit bisher nicht eingehend hinterfragt wurde, in Bibliothekssystemen wie OPACs jedoch Einzug hält, vgl. beispielsweise (Behnert und Lewandowski 2015).

In der systematischen Übersichtsarbeit rütteln *partial match* Technologien jedoch zunächst an den Grundfesten der Methodenlehre und führend derzeit zu einem durch die Methodenlehre unüberwindbaren Problem: Treffermengen sind weniger gut hinsichtlich ihrer Vollständigkeit beurteilbar (vgl. hierzu die Problematik in der Nutzung der Suchmaschine Google Scholar in Abschnitt 3.5.1.2.) Diese Einschätzung findet sich in direktem Kontext auch bei Tait (Tait 2013):

“The real point I want to make [...] is that it is often said patent searchers and legal searchers prefer Boolean to ranked retrieval: I believe what they really want is some of the specific properties they obtain from Boolean Retrieval (like reproducibility), and this should be born in mind by researchers and systems vendors.”

Tait bezieht sich hier darauf, dass Nutzer an Booleschen Retrieval-Systemen spezifische Eigenschaften schätzen, darunter vor allem die Reproduzierbarkeit ihrer Suchanfragen und die Abschätzbarkeit von Treffermengen und Recall.

Auch zeigt sich die von Tait ebenso angesprochene Diskrepanz zwischen dem Wunsch nach einer Formalisierung von Suchanfragen durch den Rechercheexperten auf der einen Seite und der Fokussierung der Information Retrieval-Forschung auf Ranking-basierte Technologien auf der anderen Seite (Tait 2013) in der evidenzbasierten Medizin ganz deutlich:

Das Prinzip der Building Blocks und PICO-Schema, das sich 1:1 mit dem Grundprinzip des Booleschen Retrievals deckt (vgl. Abschnitte 2.4.4 und 3.4.3.1), ist zentrales Element der evidenzbasierten Medizin und daher Grundvoraussetzung für deren Gelingen. Eine Abkehr von oder der Formalisierung des Informationsbedarfs in Form von Building Blocks würde ein Umdenken in der Konzeption und Durchführung systematischer Reviews erforderlich machen. Die Formalisierung des Informationsbedarfs analog zum Grundprinzip Boolescher Systeme kann daher kaum in Frage gestellt werden, ohne gleichzeitig auch Säulen der Methodenlehre

systematischer Übersichtsarbeiten zu hinterfragen – zumindest sofern diese nicht den Qualitätsstandard der EBM betreffen, der aufgrund seiner eminent wichtigen Rolle für die Qualitätssicherung im Gesundheitswesen selbstverständlich nicht in Frage gestellt werden darf.

Denkbar sind beispielsweise Systeme, in denen es grundsätzlich möglich ist, mittels einer nach Booleschem Prinzip aufgebauten strukturierten Fragestellung in den Datenbestand einzusteigen, damit auf Basis von Building Blocks, fachspezifisch also dem PICO-Schema und seiner Derivate (vgl. Abschnitt 3.3.1), den aus Nutzersicht korrekt formalisierten, wenn auch noch nicht vollständigen Informationsbedarf vorzugeben und gegebenenfalls mittels einer Ground Truth zu unterfüttern. Ansätze für Information Retrieval-Systeme, die eine solche Experten-suche unterstützen, finden sich beispielsweise im ezDL-System (Beckers et al. 2014), das dem Nutzer Suchmodi vorschlägt, wie sie beispielsweise bei Bates oder Fidel beschrieben sind. Im Kontext biomedizinischer Übersichtsarbeiten wurden zudem Experimente im erweiterten Booleschen Retrieval-Modell positiv evaluiert, indem unter anderem die Zahl relevanter Treffer durch Einsatz von Termgewichtung gegenüber dem klassischen Modell deutlich erhöht werden konnte (Pohl et al. 2010), ohne dabei das grundsätzliche Suchverhalten zu stören.

Wie bereits deutlich wurde, ist es aufgrund der drastisch steigenden Zahl an Erstveröffentlichungen jedoch unabdingbar, das Wissen hinter dem Dokumentenbestand sichtbar zu machen. Dieser Denkansatz deckt sich mit der Einschätzung anderer Autoren (Petcu und Dragusin 2013; Hjørland 2015), die Herausforderung könnte sich nach (Hjørland 2015) folgendermaßen formulieren lassen:

"[...] the Boolean model should be recognized and not replaced with less powerful models. In the Boolean model, a great range of strategies are available to increase "recall" and "precision" [...] To utilize such devices in optimal ways, the user has to know about the databases, search facilities, documents, genres, languages, paradigms, and so on, in which he or she is searching. This should be part of what is often termed information literacy."

Im Mittelpunkt der Forschung sollte laut Hjørland demnach nicht die Neuentwicklung mächtiger Retrieval-Systeme stehen, sondern die Förderung von Bildung und das Wissen um Datenbanken, Strukturen und Recherchemöglichkeiten. Nicht außer Acht gelassen sollte dabei

allerdings werden, dass der Autor bei einem Paradigmenwechsel in der professionellen medizinischen Fachrecherche hin zu auch von Domänenexperten unter den gegebenen Voraussetzungen erfolgreich bedienbaren Systemen auch die Daseinsberechtigung von Informationsexperten in Frage stellt (Hjørland 2015):

“If information searching can be done by everybody with best-match systems and with good results, the role of professional searchers is outplayed.”

Die zentrale, bereits zu Beginn der Arbeit formulierte Frage, die direkt auf Anregungen und neue Denk- und Forschungsaspekte abzielt, ist:

Welche der bekannten Aufgaben während der systematischen Literaturrecherche könnten automatisiert oder wenigstens teilautomatisiert werden?

Sie mündet direkt in eine weitere Frage:

Decken aktuelle technologische Entwicklungen den wirklichen Bedarf der Nutzer und welche Vor- und Nachteile ergäben sich aus dem Einsatz alternativer Retrieval-Technologien?

Die Fülle an derzeit verfügbaren Text Mining-Verfahren, darunter insbesondere semantisch gestützte Methoden, gibt grundsätzlich einen Hinweis darauf, dass (abgesehen von möglichen Performanz-Problemen in riesigen Datenbeständen, was nicht Thema der vorliegenden Arbeit sein kann) prinzipiell fast alle während der Entwicklung von Suchprofilen anfallenden Aufgaben auf Basis von Text Mining mindestens semi-automatisiert unterstützt oder vereinfacht werden könnten, darunter insbesondere für den Nutzer kritische Aufgaben

- die Auswahl und Bewertung geeigneter Kandidaterme
- die bewertende und vergleichende Evaluation von Suchprofilen
- Relevanz- und Vollständigkeitsbeurteilungen von Suchprofil und Testsuchen

sowie im Verlauf auch Screening- und Analyseprozesse. An dieser Stelle sei auf die Bemühungen zur Text Mining-gestützten Patentrecherche wie auch der Biomedizin hingewiesen (Abbas et al. 2014; Fleuren und Alkema 2015).

Das vor allem für die Qualitätssicherung im Gesundheitswesen drängendste Problem der systematischen Übersichtsarbeit ist, wie bereits im einleitenden Kapitel deutlich gemacht, nicht

der Mangel, sondern die Flut an Information, die einhergeht mit einer sprachlichen Diversität an Fachterminologien, einer Fülle an teils methodisch nicht akzeptablen Studiendesigns sowie einer der Publikationspolitik geschuldeten systematischen Verzerrung der Literaturlandschaft. Erschwert wird dies zusätzlich durch die von Datenbankanbieter zu Datenbankanbieter unterschiedliche Indexierung von Datenbeständen sowie Unterschiede in thematischen Schwerpunkten einzelner Thesauri und Ontologien. Da sich Informationsverhalten in der professionellen Literatursuche häufig an der Qualität verfügbarer Information (insbesondere also dem zugrundeliegenden Studiendesign) festmachen lässt, gilt es, zudem Methoden zu schaffen, die abseits einer manuellen Indexierung Publikationsformen wie systematische Übersichtsarbeiten oder Studientypen wie die randomisierte klinische Studie leichter identifizierbar machen.

Mehr noch müssen wir uns zur Lösung dieses zentralen Problems andere Aspekte bewusstmachen, die dringender sind, als lediglich einzelne Aspekte der systematischen Übersichtsarbeit zu automatisieren. Wie die Ergebnisse der vorliegenden Arbeit zeigen, wird vor diesem Hintergrund das Spannungsfeld zwischen Domänenwissen und notwendiger Rechercheexpertise zum kritischen Punkt.

Das ideale Retrieval-System macht daher das in großen Datenbeständen zugrundeliegende Wissen in erster Linie transparent und sichtbar und hilft dem Rechercheexperten bei der Überbrückung situativen Informationsbedarfs sowie der Schließung von Lücken im domänenspezifischen Wissen. Die Formalisierung eines Informationsbedarfs in Form einer Booleschen Query mag für das Konzept der systematischen Literaturrecherche per definitionem tauglich sein, deckt jedoch nicht den Informationsbedarf des Nutzers.

Es kann daher nicht allein darum gehen, Text Mining-Technologien einzusetzen, um das in Dokumentenbeständen verwendete Vokabular zu extrahieren, sondern darum, implizit verfügbares Wissen sichtbar zu machen und so den situativen Informationsbedarf besser decken zu können.

Interessante Anknüpfungspunkte liefern abseits Boolescher Technologien aktuelle Entwicklungen im Bereich der Suchmaschinen, welche direkt auf das Spannungsfeld zwischen formalisiertem Informationsbedarf und Dokumentenrepräsentation abzielen und bei Wright beschrieben sind (Wright 2016):

Bisher lag der Forschungsfokus wie üblicherweise im Information Retrieval darauf, eine ge-

bene Suchanfrage mit einer Sammlung von Dokumenten abzugleichen, eine Menge relevanter Dokumente zu berechnen und diese dem Nutzer listenartig zu präsentieren. Dies ist nichts anderes als die ursprüngliche Idee des Information Retrieval, wie sie bereits in den 1970er Jahren formuliert wurde und auch heute noch von allen großen Suchmaschinen so praktiziert wird. Es hat sich bisher also an der Grundidee nicht viel geändert, sie fokussierte sich in den vergangenen Jahren lediglich mehr und mehr auf Probleme der semantischen Analyse, beispielsweise der Identifikation von Synonymen, dem Umgang mit Tippfehlern oder orthographischen Problemen und sonstigen sprachlichen Unklarheiten.

Mehr und mehr zentriert sich die Entwicklung derzeit auf die sprachlich-semantische Analyse der Suchanfrage mit der Idee, durch die Verfügbarkeit großer Mengen an Suchanfragen und Suchtermen tiefere Einblicke in den wahren Informationsbedarf von Nutzern zu gewinnen. Wright beschreibt die Entwicklung von User Interfaces, die auf Basis spezieller oder eher allgemein gehaltener Suchterme die Darstellung von Treffermengen anpassen (Wright 2016).

Auch für große bibliographische Datenbanken liegen Query Logs vor. Sie zu analysieren, ist nicht neu und führte zu einer Vielzahl entsprechender Publikationen zu Analyse des Nutzerverhaltens, beispielsweise (Herskovic et al. 2007; Mosa und Yoo 2013). Sie aber für die Analyse des Informationsbedarfs zu nutzen, wäre ein vielversprechender Ansatz, der die Kluft zwischen Domänenwissen auf der einen und Rechercheexpertise auf der anderen Seite überwinden helfen könnte.

Der Vorteil für die professionelle Literaturrecherche, dabei insbesondere die systematische Übersichtsarbeit in der Medizin, läge sicher in dem notwendigen, tieferen Verständnis für aktuelle fachliche Entwicklungen und sprachliche Diversität. Insbesondere ist anzunehmen, dass mittels dieser Ansätze auch die Probleme mit komplexen Interventionen und komplexen Outcomes leichter überwunden werden könnten.

Der bereits angesprochene konzeptionelle Bruch zwischen der formalisierten Suchanfrage einerseits und der Wissensrepräsentation in verfügbaren Dokumenten andererseits, kommt hier besonders zum Tragen. Auch eine noch so ausgefeilte Boolesche Suchanfrage kann den eigentlichen Informationsbedarf nie hundertprozentig transportieren. Ein reiner Abgleich von Suchbegriffen mit spezifischen Textworten bibliographischer Dokumente und ihrer (gegebenenfalls falschen oder unzureichenden) manuellen Verschlagwortung ist für den Anspruch der evidenzbasierten Medizin vor allem in zunehmend wachsenden Datenbeständen und vor dem Hintergrund diverser systematischer Verzerrungen langfristig nicht zu verantworten.

Aufgrund der Ergebnisse der vorliegenden Arbeit müssen unabhängig von technologischen Implementierungen allerdings zwei nicht zu unterschätzende Kriterien berücksichtigt werden: Die Akzeptanz durch die Nutzer und die Gewährleistung aller Anforderungen, die per definitionem an die systematische Übersichtsarbeit gekoppelt sind.

Dies kann nach aktueller Einschätzung nur gewährleistet werden, wenn

- die Nutzung der Funktionen keinen weiteren, erhöhten Lernaufwand für die Informationsspezialisten bedeutet
- Akteure ihre bisherige Arbeitsweise bei Routineaufgaben so weit wie möglich erhalten können
- Implementierte Text Mining-Verfahren für den Akteur transparent sind
- Änderungen durch Nutzung innovativer Funktionen jederzeit vergleichbar und reversibel sind
- die Nutzung dieser Funktionen dokumentierbar und reproduzierbar ist
- der Nutzer bei Bedarf die vollständige Kontrolle über den Einsatz verfügbarer Funktionen und die daraus resultierenden Ergebnisse hat

Zusätzlich gilt für alle Innovationen auch, die Akzeptanz durch EBM-nahe Institutionen zu sichern. Dies scheint insbesondere gewährleistet, wenn der Implementierung neuer Verfahren eine umfangreiche, vergleichende Evaluation im Hinblick auf die Vollständigkeit ermittelter Ergebnisse vorausgegangen ist, sodass diese als evidenzbasiert gelten können.

Vor allem aber gilt es, Methoden zu entwickeln, die zur Beschleunigung und Verbesserung von Informationsgewinnungsprozessen in der evidenzbasierten Medizin beitragen und damit die Qualitätssicherung im Gesundheitswesen weiterhin gewährleisten können. Dabei sind trotz strikter Methodenlehre der EBM im Wesentlichen die bereits angesprochenen „Google-ähnlichen“ Retrieval-Verfahren als große Chance zu sehen: Sie könnten helfen, die in der Medizin aufgrund höchst unterschiedlicher Studienqualität als auch kommerziell-gesteuerter Publikationspolitik drastisch spürbaren systematischen Verzerrungen in der Forschungs- und Literaturlandschaft transparenter zu machen.

7 Fazit

Die systematische Übersichtsarbeit in der Medizin ist bekannt als zeit- und ressourcenintensiver Prozess der professionellen Informationsbeschaffung mit spezifischen Anforderungen, die insbesondere strengen methodologischen Rahmenbedingungen wie der Vollständigkeitsbeurteilung und Reproduzierbarkeit von Suchprozessen unterliegt. Es ist nur konsequent, ein Umdenken verfügbarer Systeme zu fordern, um den Workload der Akteure – ausschließlich geschulte Informationsspezialisten und Rechercheexperten – zu reduzieren.

Information behaviour mit dem Ziel der Entwicklung nutzerzentrierter Information Retrieval-Systeme und ihrer Bewertung ist seit Jahrzehnten als informationswissenschaftlicher Forschungsgegenstand etabliert. Erst seit einigen Jahren gerät dabei auch das Informationsverhalten in der professionellen, meist auftragsinitiierten Suche in den Fokus des Forschungsinteresses, darunter insbesondere die Patentrecherche.

Die medizinische Übersichtsarbeit wurde dabei in der Literatur weitestgehend vernachlässigt, was dazu führte, dass über das Informationsverhalten in der systematischen Übersichtsarbeit bisher wenig bekannt war. Methodenlehre und Forschung ist seit Jahrzehnten klassischerweise von Bibliothekswissenschaftlern dominiert, was zum einen zu einer extrem Information Retrieval-lastigen Forschung und daraus resultierend auch zu unterschiedlichen Terminologien führte, während sich die Informationswissenschaft in den vergangenen Jahrzehnten neuen Information Retrieval-Technologien widmete und sich von der traditionellen Systemsicht des Information Retrieval verabschiedete.

Mit der hier vorgestellten Forschungsarbeit sollte eine Lücke geschlossen werden, um beide Wissenschaften wieder näher zusammenzubringen.

Mit der vorliegenden Arbeit wird daher erstmals der Akteur der systematischen Übersichtsarbeit sichtbar.

Das im Ergebnis entwickelte Rahmenmodell zeigt einen wie zu erwarten stark methodologisch geprägten, teils wenig nutzerfreundlichen Prozess unter kontinuierlich erschwerten Recherchebedingungen. Das Modell ergänzt die ebenso methodologisch geprägte Forschung um den informationswissenschaftlich vielversprechenden Ansatz des Informationsverhaltens im Kontext unterschiedlicher Rahmenbedingungen und bietet damit eine neue Sichtweise für die Konzeption und Bewertung von Retrieval-Systemen.

Da, wie abschließend beleuchtet, moderne Information Retrieval-Ansätze neue Optionen für die Durchführung systematischer Übersichtsarbeiten bieten können, scheint ein Aspekt für die erfolgreiche Weiterentwicklung und Implementierung ressourcenschonender und nutzerfreundlicher Systeme unumgänglich:

Sowohl Datenbankanbieter als auch federführende Institutionen müssen erkennen, dass ein Loslassen alter Traditionen und die Zusammenarbeit mit der Informationswissenschaft neue Wege ermöglichen könnte.

Umgekehrt ist auch die Informationswissenschaft gefordert, denn die professionelle Informationsrecherche ist ein Thema, das hier in den vergangenen Jahrzehnten eher stiefmütterlich behandelt wurde und erst in den letzten Jahren verstärkt in den Fokus gerät.

So kann die Arbeit dazu beitragen, die systematische Übersichtsarbeit als Forschungsaspekt der professionellen Suche in Erinnerung zu bringen, neue Forschungsaspekte zu etablieren und mit durch das Wissen um die Akteure, ihren Informationsbedarf und ihr Informationsverhalten die aktive innovative Weiterentwicklung in der Suche nach Primärliteratur und aktuellen Forschungsergebnissen zu fördern.

Es muss in Zukunft die Aufgabe sein, sich noch mehr der Frage nach den Nutzerbedürfnissen und ihrer Hintergründe zu widmen und gleichzeitig das Thema der idealen und für die Medizin so ungeheuer wertvollen Wissensrepräsentation im Spannungsfeld von Forschung, Publikationspolitik und verfügbarer Information weiter zu beforschen.

Die technologischen Möglichkeiten für Innovation sind gegeben.

8 Abbildungsverzeichnis

Abbildung 1-1 Struktur und Phasen einer Übersichtsarbeit nach (Fettke 2006)	2
Abbildung 1-2 Cochrane Logo, mit freundlicher Genehmigung der Cochrane Collaboration	8
Abbildung 1-3 Jährlicher Zuwachs indexierter Dokumente in Pubmed/MEDLINE.....	18
Abbildung 1-4 Jährlicher Zuwachs von als SRs indexierter Dokumente in Pubmed/MEDLINE	19
Abbildung 2-1 Informationswissenschaft im Kontext wissenschaftlicher Disziplinen aus (Ingwersen 1992, S. 103).....	31
Abbildung 2-2 Belkins kognitives Modell der Information aus (Belkin 1978, S. 81).....	31
Abbildung 2-3 Eingabe von Datensätzen für die Speicherung auf Magnetbändern, Images from the History of Medicine (NLM), https://ihm.nlm.nih.gov/images/A17414 (09.02.2016)	33
Abbildung 2-4 Typologie des Informationsbedarfs nach Taylor (Taylor 1968)	34
Abbildung 2-5 Belkins Modell des kognitiven Kommunikationssystems aus (Belkin 1980, S. 135)	35
Abbildung 2-6 Bates Berry Picking Modell, aus (Bates 1989, S. 410).....	36
Abbildung 2-7 Modell zum Informationsverhalten nach Wilson (Wilson 1981, S. 4).....	38
Abbildung 2-8 Das Ellis Modell in der Darstellung nach Wilson (Wilson 1999)	44
Abbildung 2-9 Modell nach Meho und Tibbo (Meho und Tibbo 2003)	45
Abbildung 2-10 Kuhlthaus Information Search Process-Modell nach (Kuhlthau 1991).....	47
Abbildung 2-11 Information Seeking Process Marchionini aus (Marchionini 1995, S. 49)..	48
Abbildung 2-12 Klassifizierung der Komplexität von Aufgaben, aus (Byström und Järvelin 1995, S. 6)	52

Abbildung 2-13 Byström Järvelin Modell nach (Byström und Järvelin 1995)	54
Abbildung 2-14 Modell nach Leckie et al. nach (Leckie et al. 1996).....	55
Abbildung 2-15 Wilsons Modell des Information Behaviour (Wilson 1999).....	57
Abbildung 2-16 Grundlagenmodell des Information Retrieval nach (Ingwersen und Järvelin 2005, S. 115).....	63
Abbildung 2-17 Aufbau einer Web-Suchmaschine aus (Lewandowski 2013, S. 498).....	64
Abbildung 2-18 Sog. Concept View des MeSH Deskriptors "Ascorbic Acid" (National Center for Biotechnology Information 2016), Stand 16.06.2016	66
Abbildung 2-19 Building Blocks	94
Abbildung 2-20 Das Ingwersen und Järvelin Framework, aus (Järvelin und Ingwersen 2004)	100
Abbildung 2-21 Modell des Information Retrieval Prozesses aus (Ingwersen 1996, S. 9)..	101
Abbildung 3-1 Die 6S Evidenzpyramide nach (Dicenso et al. 2009).....	113
Abbildung 3-2 DIMDI Suche (28.06.2016)	126
Abbildung 3-3 Schritte zur Durchführung systematischer Übersichtsarbeiten nach (Tsafnat et al. 2014).....	138
Abbildung 3-4 PRISMA Statement aus (Liberati et al. 2009, S. 68).....	145
Abbildung 3-5 Dokumentation von Suchprozessen nach PRISMA aus (Loveman et al. 2015, S. 12)	146
Abbildung 3-6 Pubmed einfache Suche (25.02.2016).....	155
Abbildung 3-7 Pubmed Term-Vorschläge im einfachen Suchmodus (25.02.2016).....	156
Abbildung 3-8 Pubmed angewandte Suche (25.02.2016)	157
Abbildung 3-9 Pubmed Suche mit angezeigten Treffern (25.02.2016).....	158

Abbildung 3-10 Pubmed Advanced Search (31.03.2015)	161
Abbildung 3-11 Angewandte Suche MEDLINE via OVID (31.03.2015) über Lizenz der Universitätsbibliothek Regensburg	161
Abbildung 3-12 Exemplarische Durchführung der OVID Suche nach (Schlager et al. 2016), 29.02.2016	172
Abbildung 3-13 Text Mining-Verfahren im Workflow systematischer Übersichtsarbeiten aus (Ananiadou et al. 2009, S. 512)	180
Abbildung 3-14 PubReMiner, Screenshot vom 09.03.2016	181
Abbildung 3-15 Objektivierter Ansatz zur Entwicklung von Suchprofilen aus (Hausner et al. 2012, S. 6)	182
Abbildung 3-16 Kontext Dokument und Information Retrieval-System, Darstellung nach (Ingwersen 1996)	194
Abbildung 3-17 Kontext Arbeits- und Suchaufgaben, Darstellung nach (Ingwersen 1996)	196
Abbildung 3-18 Unbekannter Kontext: Nutzer und Interaktion, Darstellung nach (Ingwersen 1996)	200
Abbildung 3-19 Systematische Literaturrecherchen: Wissen und klinische Entscheidungen (Abbildung nach (Booth 2008, S. 583))	201
Abbildung 4-1 Qualitative Inhaltsanalyse nach Mayring, aus (Mayring 2015, S. 62)	210
Abbildung 4-2 Retrieval-gesteuerte Ground Truth Loops	226
Abbildung 4-3 Sensitivitäts-gesteuerte Ground Truth Loops	228
Abbildung 4-4 Bedingende Faktoren der Ground Truth-Loops.....	230
Abbildung 4-5 Vom Synonym zum Kandidatenterm	255
Abbildung 4-6 Thesaurus NOT Free Text Manöver nach (Bramer und Jonge 2015)	265
Abbildung 4-7 Graphische Darstellung des NOT Manövers.....	268

Abbildung 5-1 Aufgabenbezogenes Information Behaviour in der Entwicklung erschöpfender Suchanfragen 295

9 Tabellen

Tabelle 1-1 Gegenüberstellung systematischer Übersichtsarbeit und Scoping Review nach (Brien et al. 2010).....	13
Tabelle 2-1 Information als Entität und Prozess nach Buckland (Buckland 1991), zur Darstellung siehe (Capurro 2015).....	30
Tabelle 2-2 Boolesche Operatoren	78
Tabelle 2-3 Mengenalgebraische Sicht auf die Booleschen Operatoren.....	79
Tabelle 2-4 Gesetze im Umgang mit Booleschen Operatoren nach (Monk 2002)	80
Tabelle 2-5 Klassifikation von Suchaktivitäten nach Bates (Bates 1990, 1979b).....	86
Tabelle 2-6 Monitoring-Taktiken nach Bates (Bates 1979b)	87
Tabelle 2-7 Datenbankspezifische Taktiken nach Bates (Bates 1979b)	88
Tabelle 2-8 Taktiken zur Suchformulierung nach Bates (Bates 1979b).....	89
Tabelle 2-9 Term-Taktiken nach Bates (Bates 1979b).....	89
Tabelle 2-10 Operative Moves zur Reduktion von Treffermengen nach (Fidel 1991).....	91
Tabelle 2-11 Operative Moves zur Expansion von Treffermengen nach (Fidel 1991).....	92
Tabelle 2-12 Working Tasks und Information Seeking Tasks: Prozessmodelle	104
Tabelle 3-1 Biases im Zusammenhang mit systematischen Literaturrecherchen nach (Booth et al. 2012, S. 185) und (Egger und Smith 1998).....	115
Tabelle 3-2 Rollen von Datenbank Anbietern und Hosts nach (Baeza-Yates und Ribeiro-Neto 2011, S. 702).....	122
Tabelle 3-3 PICO-Schema nach (Banfield 2016)	132
Tabelle 3-4 Übertragbarkeit des PICO-Schemas auf das Software Engineering nach (Kitchenham 2004).....	133

Tabelle 3-5 Vergleich PICO und SPIDER Schema (Cooke et al. 2012).....	136
Tabelle 3-6 Iterative Aufgaben im Suchprozess der systematischen Literaturrecherche, aus (Booth et al. 2012, S. 72).....	139
Tabelle 3-7 Schlagwortsuche MEDLINE OVID, MEDLINE Pubmed (getestet am 25.02.20116)	152
Tabelle 3-8 Auszug aus den wichtigsten Datenbankfeldern der Datenbank MEDLINE, entsprechend (U.S. National Library of Medicine 2014b) und (National Center for Biotechnology Information (US) 2014), alphabetische Reihenfolge	153
Tabelle 3-9 Freitextsuche MEDLINE Ovid und MEDLINE Pubmed (getestet am 25.02.2016)	154
Tabelle 3-10 Suchanfrage bei Pubmed: Nutzeranfrage und Pubmed-interne Verarbeitung mit ATM (durchgeführt am 16.06.2016)	159
Tabelle 3-11 Suchfilter: Sensitivity and Precision Maximizing Version aus (Lefebvre et al. 2011)	168
Tabelle 3-12 Gegenüberstellung Queries aus (Schlager et al. 2016)	170
Tabelle 3-13 Vor- und Nachteile des Services Google Scholar, nach (Noruzi 2005).....	187
Tabelle 4-1 Erfassungsparameter der Beobachtungsstudie	208
Tabelle 4-2 Interaktionskontexte in der Entwicklung und Durchführung strukturierter und systematischer Literatursuchen	215
Tabelle 4-3 Struktur und direkter Ertrag von Informationseinheiten	216
Tabelle 4-4 Tätigkeiten während der Entwicklung von Suchprofilen im Kontext verschiedener Informationskontexte (siehe hierzu Tabelle 4-2)	223
Tabelle 4-5 Ergebnisse der Beobachtungsstudie: Taktiken zur Suchformulierung nach (Bates 1979b)	229
Tabelle 4-6 Ergebnisse der Beobachtungsstudie: Taktiken zur Modifikation von Termen nach (Bates 1979b).....	229

Tabelle 4-7 Übersicht der Codings der Interviewpartner	237
Tabelle 4-8 Demographische Daten Fragebogenstudie.....	238
Tabelle 4-9 Nutzung methodologischer Standards	239
Tabelle 4-10 Beispiel: Testsuche über „NOT Manöver“, Durchführung am 08.04.2016	267
Tabelle 4-11 Methoden zur Spezifizierung von Suchanfragen, siehe auch (Bramer und Jonge 2015)	274
Tabelle 5-1 Zustände von Suchtermen	291
Tabelle 5-2 Charakteristika von Informationseinheiten	298

10 Literatur

Abbas, A.; Zhang, L.; Khan, S. U. (2014): A literature review on the state-of-the-art in patent analysis. In: *World Patent Information* 37, S. 3–13. DOI: 10.1016/j.wpi.2013.12.006.

Aggarwal, C. C.; Zhai, C. (2012): An Introduction to Text Mining. In: C. C. Aggarwal und C. Zhai (Hg.): *Mining Text Data*. Boston, MA: Springer US, S. 1–10.

Agoritsas, T.; Merglen, A.; Courvoisier, D. S.; Combesure, C.; Garin, N.; Perrier, A.; Perneger, T. V. (2012): Sensitivity and predictive value of 15 PubMed search strategies to answer clinical questions rated against full systematic reviews. In: *J Med Internet Res* 14 (3), S. e85. DOI: 10.2196/jmir.2021.

Alberts, D.; Yang, C. B.; Fobare-DePonio, D.; Koubek, K.; Robins, S.; Rodgers, M. et al. (2011): Introduction to Patent Searching. In: M. Lupu, K. Mayer, J. Tait und A. J. Trippe (Hg.): *Current Challenges in Patent Information Retrieval*, Bd. 29. Berlin, Heidelberg: Springer-Verlag Berlin Heidelberg (The Information Retrieval Series, 29), S. 3–43.

Allen, I. E.; Olkin, I. (1999): Creative Nonfiction. In: *JAMA* 282 (7), S. 634–635. Online verfügbar unter <http://jama.jamanetwork.com/data/Journals/JAMA/4676/JBK0818.pdf>.

Allison, J. J.; Kiefe, C. I.; Weissman, N. W.; Carter, J.; Centor, R. M. (1999): The art and science of searching MEDLINE to answer clinical questions. Finding the right number of articles. In: *Int J Technol Assess Health Care* 15 (2), S. 281–296.

American Psychological Association (Hg.) (2016): PsycINFO. Online verfügbar unter <http://www.apa.org/pubs/databases/psycinfo/>, zuletzt geprüft am 21.06.2016.

Ananiadou, S.; Rea, B.; Okazaki, N.; Procter, R.; Thomas, J. (2009): Supporting Systematic Reviews Using Text Mining. In: *Social Science Computer Review* 27 (4), S. 509–523. DOI: 10.1177/0894439309332293.

Anders, M. E.; Evans, D. P. (2010): Comparison of PubMed and Google Scholar literature searches. In: *Respir Care* 55 (5), S. 578–583.

Angelis, C. D. de (2005): Is This Clinical Trial Fully Registered? In: *Ann Intern Med* 143 (2), S. 146–148. DOI: 10.7326/0003-4819-143-2-200507190-00016.

Antes, G.; Dreier, G.; Hasselblatt, H.; Blümle, A.; Schumacher, M. (2009): Register für klinische Studien. In: *Bundesgesundheitsbl.* 52 (4), S. 459–462. DOI: 10.1007/s00103-009-0832-6.

Arbeitsgemeinschaft der Wissenschaftlichen Medizinischen Fachgesellschaften e.V. (AWMF) e.V. (Hg.) (2016): AWMF: Leitlinien. Online verfügbar unter <http://www.awmf.org/leitlinien.html>, zuletzt geprüft am 21.06.2016.

Arksey, H.; O'Malley, L. (2005): Scoping studies: towards a methodological framework. In: *International Journal of Social Research Methodology* 8 (1), S. 19–32. DOI: 10.1080/1364557032000119616.

Azzopardi, L.; Vanderbauwhede, W.; Joho, H. (2010): Search system requirements of patent analysts. In: F. Crestani und S. Marchand-Maillet (Hg.): Proceedings of the 33rd international ACM SIGIR conference on Research and development in information retrieval. SIGIR '10 The 33rd International ACM SIGIR conference on research and development in Information Retrieval. Geneva, Switzerland, 19.-23.07.2010. New York, NY: ACM, S. 775–776.

Bache, M. (2011): Measuring and Improving Access to the Corpus. In: M. Lupu, K. Mayer, J. Tait und A. J. Trippe (Hg.): Current Challenges in Patent Information Retrieval. Berlin, Heidelberg: Springer-Verlag Berlin Heidelberg (The Information Retrieval Series, 29), S. 147–165. Online verfügbar unter http://link.springer.com/chapter/10.1007/978-3-642-19231-9_7, zuletzt geprüft am 28.04.2016.

Bachmann, J. (2000): Der Information-Broker. Informationen suchen, sichten, präsentieren. München: Addison-Wesley (Die Integrata-Qualifizierung).

Baeza-Yates, R.; Ribeiro-Neto, B. (2011): Modern information retrieval. The concepts and technology behind search. Second edition. New York: Addison Wesley.

Balan, P. F.; Gerits, A.; Vanduffel, W. (2014): A practical application of text mining to literature on cognitive rehabilitation and enhancement through neurostimulation. In: *Frontiers in Systems Neuroscience* 8. DOI: 10.3389/fnsys.2014.00182.

Ball, R.; Tunger, D. (2007): *Bibliometrische Analysen - Daten, Fakten und Methoden*. Grundwissen Bibliometrie für Wissenschaftler, Wissenschaftsmanager, Forschungseinrichtungen und Hochschulen. Unveränd. Nachdr. Jülich: Forschungszentrum (Schriften des Forschungszentrums Jülich : Reihe Bibliothek, 12).

Banfield, L. (2016): *Resources for Evidence-Based Practice: Forming Questions*. McMaster University Health Sciences Library Health Sciences Library Home. Online verfügbar unter <http://hsl.mcmaster.libguides.com/c.php?g=306765&p=2044787>, zuletzt aktualisiert am 18.02.2016, zuletzt geprüft am 14.03.2016.

Bates, M. J. (1979a): Idea Tactics. In: *J. Am. Soc. Inf. Sci.* 30 (5), S. 280–289. DOI: 10.1002/asi.4630300507.

Bates, M. J. (1979b): Information search tactics. In: *J. Am. Soc. Inf. Sci.* 30 (4), S. 205–214. DOI: 10.1002/asi.4630300406.

Bates, M. J. (1989): The design of browsing and berrypicking techniques for the online search interface. In: *Online Information Review* 13 (5), S. 407–424. DOI: 10.1108/eb024320.

Bates, M. J. (1990): Where should the person stop and the information search interface start? In: *Information Processing & Management* 26 (5), S. 575–591. DOI: 10.1016/0306-4573(90)90103-9.

Bates, M. J. (2007): What is browsing—really? A model drawing from behavioural science research. In: *Information Research* 12 (4), S. 330. Online verfügbar unter <http://InformationR.net/ir/12-4/paper330.html>, zuletzt geprüft am 11.02.2016.

Beale, S.; Duffy, S.; Glanville, J. M.; Lefebvre, C.; Wright, D.; McCool, R. et al. (2014): Choosing and using methodological search filters: searchers' views. In: *Health Info Libr J* 31 (2), S. 133–147. DOI: 10.1111/hir.12062.

Beckers, T.; Dungs, S.; Fuhr, N.; Jordan, M.; Kontokotsios, G.; Kriewel, S. et al. (2014): ezDL: An Interactive IR Framework, Search Tool, and Evaluation System. In: G. Paltoglou, F. Loizides und P. Hansen (Hg.): *Professional Search in the Modern World*. Cham: Springer International Publishing (Lecture Notes in Computer Science), S. 118–146.

- Beel, J.; Gipp, B. (2009): Google Scholar's Ranking Algorithm: An Introductory Overview. In: B. Larsen und J. Leta (Hg.): Proceedings of ISSI 2009 - the 12th International Conference of the International Society for Scientometrics and Informetrics, 12th International Conference on Scientometrics and Informetrics: ISSI 2009. Rio de Janeiro, 16.-19.09.2009, S. 230–241. Online verfügbar unter http://www.sciplore.org/publications/2009-Google_Scholar%27s_Ranking_Algorithm_-_An_Introductory_Overview_-_preprint.pdf, zuletzt geprüft am 16.06.2016.
- Behnert, C.; Lewandowski, D. (2015): Ranking Search Results in Library Information Systems — Considering Ranking Approaches Adapted From Web Search Engines. In: *The Journal of Academic Librarianship* 41 (6), S. 725–735. DOI: 10.1016/j.acalib.2015.07.010.
- Bekhuis, T.; Demner-Fushman, D. (2010): Towards automating the initial screening phase of a systematic review. In: *Stud Health Technol Inform* 160 (Pt 1), S. 146–150.
- Bekhuis, T.; Demner-Fushman, D. (2012): Screening nonrandomized studies for medical systematic reviews: A comparative study of classifiers. In: *Artificial intelligence in medicine* 55 (3), S. 197–207. DOI: 10.1016/j.artmed.2012.05.002.
- Belkin, N. J. (1978): Information Concepts for Information Science. In: *Journal of Documentation* 34 (1), S. 55–85. DOI: 10.1108/eb026653.
- Belkin, N. J.; Cool, C. (2002): Classification of interactions with information. In: Martha Kellogg Smith (Hg.): Emerging frameworks and methods. Proceedings of the Fourth International Conference on Conceptions of Library and Information Science (COLIS4). Fourth International Conference on Conceptions of Library and Information Science (CoLIS4). Seattle, Washington, 21 - 25 July 2002, S. 1–15.
- Belkin, N. J.; Cool, C.; Stein, A.; Thiel, U. (1995): Cases, Scripts, and Information-Seeking Strategies: On the Design of Interactive Information Retrieval Systems. In: *Expert Systems with Applications* 9 (3), S. 379–395. DOI: 10.1016/0957-4174(95)00011-W.
- Belkin, N. J.; Croft, W. B. (1987): Retrieval techniques. In: Annual review of information science and technology, vol. 22. Elsevier Science Inc, S. 109–145.

Belkin, N. J.; Oddy, R. N.; Brooks, H. M. (1982): Ask for Information Retrieval: Part I. Backgrounds and Theory. In: *Journal of Documentation* 38 (2), S. 61–71. DOI: 10.1108/eb026722.

Beller, E. M.; Chen, J. K.-H.; Wang, U. L.-H.; Glasziou, P. (2013): Are systematic reviews up-to-date at the time of publication? In: *Sys Rev* 2 (1), S. 36. DOI: 10.1186/2046-4053-2-36.

Bennett, D. A.; Latham, N. K.; Stretton, C.; Anderson, C. S. (2004): Capture-recapture is a potentially useful method for assessing publication bias. In: *Journal of Clinical Epidemiology* 57 (4), S. 349–357. DOI: 10.1016/j.jclinepi.2003.09.015.

Benzies, K. M.; Premji, S.; Hayden, K. Alix; Serrett, K. (2006): State-of-the-evidence reviews: advantages and challenges of including grey literature. In: *Worldviews Evid Based Nurs* 3 (2), S. 55–61. DOI: 10.1111/j.1741-6787.2006.00051.x.

Beverley, C. A.; Booth, A.; Bath, P. A. (2003): The role of the information specialist in the systematic review process: a health information case study. In: *Health Information & Libraries J* 20 (2), S. 65–74. DOI: 10.1046/j.1471-1842.2003.00411.x.

Bibliographisches Institut GmbH (Hg.) (2016): Duden | Startseite. Online verfügbar unter <http://www.duden.de/>, zuletzt aktualisiert am 21.06.2016, zuletzt geprüft am 21.06.2016.

Bibliographisches Institut GmbH Duden Verlag (Hg.) (2016a): Duden. Online verfügbar unter http://www.duden.de/rechtschreibung/Review_Kritik_Besprechung, zuletzt geprüft am 21.06.2016.

Bibliographisches Institut GmbH Duden Verlag (2016b): Duden | Auf-ga-be | Rechtschreibung, Bedeutung, Definition, Synonyme. Online verfügbar unter <http://www.duden.de/rechtschreibung/Aufgabe>, zuletzt aktualisiert am 20.06.2016, zuletzt geprüft am 21.06.2016.

Bibliographisches Institut GmbH Duden Verlag (Hg.) (2016c): Duden | Bi-as | Rechtschreibung, Bedeutung, Definition, Herkunft. Online verfügbar unter <http://www.duden.de/rechtschreibung/Bias>, zuletzt aktualisiert am 21.06.2016, zuletzt geprüft am 21.06.2016.

Bibliographisches Institut GmbH Duden Verlag (Hg.) (2016d): Duden | Le-xem | Rechtschreibung, Bedeutung, Definition, Herkunft. Online verfügbar unter <http://www.duden.de/rechtschreibung/Lexem>, zuletzt aktualisiert am 21.06.2016, zuletzt geprüft am 21.06.2016.

Bibliographisches Institut GmbH Duden Verlag (2016e): Duden | Pri-mär-li-te-ra-tur | Rechtschreibung, Bedeutung, Definition. Online verfügbar unter <http://www.duden.de/rechtschreibung/Primaerliteratur>, zuletzt aktualisiert am 21.06.2016, zuletzt geprüft am 21.06.2016.

Bibliographisches Institut GmbH Duden Verlag (Hg.) (2016f): Duden | Re-le-vanz | Rechtschreibung, Bedeutung, Definition, Synonyme, Herkunft. Online verfügbar unter <http://www.duden.de/rechtschreibung/Relevanz>, zuletzt aktualisiert am 21.06.2016, zuletzt geprüft am 21.06.2016.

Bibliographisches Institut GmbH Duden Verlag (Hg.) (2016g): Duden | Stra-te-gie | Rechtschreibung, Bedeutung, Definition, Synonyme, Herkunft. Online verfügbar unter <http://www.duden.de/rechtschreibung/Strategie>, zuletzt aktualisiert am 21.06.2016, zuletzt geprüft am 21.06.2016.

Bibliographisches Institut GmbH Duden Verlag (Hg.) (2016h): Duden | sy-n-o-nym | Rechtschreibung, Bedeutung, Definition, Synonyme, Herkunft. Online verfügbar unter <http://www.duden.de/rechtschreibung/synonym>, zuletzt aktualisiert am 21.06.2016, zuletzt geprüft am 21.06.2016.

Bibliographisches Institut GmbH Duden Verlag (Hg.) (2016i): Duden | In-for-ma-ti-on | Rechtschreibung, Bedeutung, Definition, Synonyme, Herkunft. Online verfügbar unter <http://www.duden.de/rechtschreibung/Information>, zuletzt aktualisiert am 20.06.2016, zuletzt geprüft am 21.06.2016.

Bibliographisches Institut GmbH Duden Verlag (Hg.) (2016j): Duden | Do-ku-ment | Rechtschreibung, Bedeutung, Definition, Synonyme, Herkunft. Online verfügbar unter <http://www.duden.de/rechtschreibung/Dokument>, zuletzt aktualisiert am 21.06.2016, zuletzt geprüft am 21.06.2016.

Bibliographisches Institut GmbH Duden Verlag (Hg.) (2016k): Duden | In-ter-ak-ti-on | Rechtschreibung, Bedeutung, Definition, Synonyme. Online verfügbar unter <http://www.duden.de/rechtschreibung/Interaktion>, zuletzt aktualisiert am 28.06.2016, zuletzt geprüft am 28.06.2016.

Blümle, A.; Antes, G. (2008): Handsuche nach randomisierten kontrollierten Studien in deutschen medizinischen Zeitschriften. In: *Deutsche medizinische Wochenschrift (1946)* 133 (6), S. 230–234. DOI: 10.1055/s-2008-1017501.

Boeker, M.; Vach, W.; Moshontz, H. (2012): Semantically equivalent PubMed and Ovid-MEDLINE queries: different retrieval results because of database subset inclusion. In: *Journal of Clinical Epidemiology* 65 (8), S. 915–916. DOI: 10.1016/j.jclinepi.2012.01.015.

Boeker, M.; Vach, W.; Motschall, E. (2013): Google Scholar as replacement for systematic literature searches: good relative recall and precision are not enough. In: *BMC Med Res Methodol* 13, S. 131. DOI: 10.1186/1471-2288-13-131.

Bonino, D.; Ciaramella, A.; Corno, F. (2010): Review of the state-of-the-art in patent information and forthcoming evolutions in intelligent patent informatics. In: *World Patent Information* 32 (1), S. 30–38. DOI: 10.1016/j.wpi.2009.05.008.

Booth, A. (2008): Unpacking your literature search toolbox: on search styles and tactics. In: *Health Info Libr J* 25 (4), S. 313–317. DOI: 10.1111/j.1471-1842.2008.00825.x.

Booth, A. (2010): How much searching is enough? Comprehensive versus optimal retrieval for technology assessments. In: *Int J Technol Assess Health Care* 26 (4), S. 431–435. DOI: 10.1017/S0266462310000966.

Booth, A. (2016): Searching for qualitative research for inclusion in systematic reviews: a structured methodological review. In: *Systematic reviews* 5 (1), S. 1. DOI: 10.1186/s13643-016-0249-x.

Booth, A.; Papaioannou, D.; Sutton, A. (2012): *Systematic approaches to a successful literature review*. Los Angeles [etc.]: Sage.

Boynton, J.; Glanville, J.; McDaid, D.; Lefebvre, C. (1998): Identifying systematic reviews in MEDLINE: developing an objective approach to search strategy design. In: *Journal of Information Science* 24 (3), S. 137–154. DOI: 10.1177/016555159802400301.

Bramer, W. M.; Giustini, D.; Kramer, B.; Anderson, P. F. (2013): The comparative recall of Google Scholar versus PubMed in identical searches for biomedical systematic reviews: a review of searches used in systematic reviews. In: *Syst Rev* 2, S. 115. DOI: 10.1186/2046-4053-2-115.

Bramer, W. M.; Giustini, D.; Kramer, B. M. R. (2016): Comparing the coverage, recall, and precision of searches for 120 systematic reviews in Embase, MEDLINE, and Google Scholar: a prospective study. In: *Systematic reviews* 5 (1), S. 1. DOI: 10.1186/s13643-016-0215-7.

Bramer, W. M.; Jonge, G. B. de (2015): Improving efficiency and confidence in systematic literature searching. Workshop at EAHIL+ICAHIS+ICLC Conference, Edinburgh, UK 11.-12.06.2015. https://www.researchgate.net/publication/278300171_Improving_efficiency_and_confidence_in_systematic_literature_searching_-_eahil2015_researchminded. Rotterdam, Niederlande. Internet (Word Dokument), zuletzt geprüft am 21.06.2016.

Brien, S. E.; Lorenzetti, D. L.; Lewis, S.; Kennedy, J.; Ghali, W. A. (2010): Overview of a formal scoping review on health system report cards. In: *Implement Sci* 5, S. 2. DOI: 10.1186/1748-5908-5-2.

Buckland, M. K. (1991): Information and information systems. New York: Greenwood Press (New directions in information management, no. 25).

Buckland, M. K. (1997): What is a "document"? In: *Journal of the American Society for Information Science* 48 (9), S. 804–809. DOI: 10.1002/(SICI)1097-4571(199709)48:9<804::AID-ASI5>3.0.CO;2-V.

Bush, V. (1945): As We May Think. Online verfügbar unter <http://www.theatlantic.com/magazine/archive/1945/07/as-we-may-think/303881/>, zuletzt geprüft am 16.02.2016.

Bußmann, H.; Gerstner-Link, C.; Lauffer, H. (Hg.) (2008): Lexikon der Sprachwissenschaft. Mit 14 Tabellen. 4., durchges. und bibliogr. erg. Aufl. Stuttgart: Kröner.

Byström, K. (1999): Task complexity, information types and information sources. Examination of relationships. Doctoral Thesis. University of Tampere, Finland, Tampere, Finland. Dept. of Information Studies.

Byström, K. (2002): Information and information sources in tasks of varying complexity. In: *J. Am. Soc. Inf. Sci.* 53 (7), S. 581–591. DOI: 10.1002/asi.10064.

Byström, K.; Hansen, P. (2005): Conceptual framework for tasks in information studies. In: *J. Am. Soc. Inf. Sci.* 56 (10), S. 1050–1061. DOI: 10.1002/asi.20197.

Byström, K.; Järvelin, K. (1995): Task complexity affects information seeking and use. In: *Information Processing & Management* 31 (2), S. 191–213. DOI: 10.1016/0306-4573(95)80035-R.

Canese, K. (2015): PubMed Relevance Sort. U.S. National Library of Medicine. Online verfügbar unter https://www.nlm.nih.gov/pubs/techbull/so13/so13_pm_relevance.html, zuletzt aktualisiert am 01.10.2015, zuletzt geprüft am 29.04.2016.

Capurro, R. (2015): Vorlesung Informationsbegriff. Online verfügbar unter <http://www.capurro.de/infovorl-index.htm>, zuletzt aktualisiert am 08.08.2015, zuletzt geprüft am 09.02.2016.

Case, D. O. (2012): Looking for Information. A survey of research on information seeking, needs and behavior. 3rd ed. Bingley, UK: Emerald Group Pub. (Library and information science).

Centre for Reviews and Dissemination (2009): Systematic Reviews. CRD's guidance for undertaking reviews in healthcare. 3. ed. York: Centre for Reviews and Dissemination, University of York (Systematic reviews).

Centre for Reviews and Dissemination (Hg.) (2015): NIHR Centre for Reviews and Dissemination - CRD Database. University of York. Online verfügbar unter <http://www.crd.york.ac.uk/CRDWeb/AboutPage.asp>, zuletzt aktualisiert am 03.06.2015, zuletzt geprüft am 21.06.2016.

Centre for Reviews and Dissemination (Hg.) (2016): Centre for Reviews and Dissemination - Home page. University of York. Online verfügbar unter <http://www.york.ac.uk/inst/crd/index.htm>, zuletzt geprüft am 20.06.2016.

Chandler, J.; Churchill, R.; Higgins, J. P.T.; Lasserson, T.; Tovey, D. (2013): Methodological Expectations of Cochrane Intervention Reviews (MECIR): Methodological standards for the conduct of new Cochrane Intervention Reviews. Cochrane Collaboration. Online verfügbar unter http://editorial-unit.cochrane.org/sites/editorial-unit.cochrane.org/files/uploads/MECIR_conduct_standards%202.3%2002122013_0.pdf, zuletzt geprüft am 02.03.2016.

Cimpl, K. (1985): Clinical medical librarianship: a review of the literature. In: *Bull Med Libr Assoc* 73 (1), S. 21–28.

Claassen, J. A. H. R. (2005): The gold standard: not a golden standard. In: *BMJ* 330 (7500), S. 1121. DOI: 10.1136/bmj.330.7500.1121.

Claridge, J.A.; Fabian, T.C. (2005): History and Development of Evidence-based Medicine. In: *World J. Surg.* 29 (5), S. 547–553. DOI: 10.1007/s00268-005-7910-1.

Cochrane Collaboration (2015): Evidence-based health care and systematic reviews | Cochrane Community (beta). Online verfügbar unter <http://community.cochrane.org/about-us/evidence-based-health-care>, zuletzt aktualisiert am 2015, zuletzt geprüft am 24.02.2016.

Cochrane Diagnostic Test Accuracy Working Group (Hg.) (2012): Handbook for Diagnostic Test Accuracy Reviews. Cochrane Collaboration. Online verfügbar unter <http://srda.cochrane.org/handbook-dta-reviews>, zuletzt aktualisiert am 23.07.2015, zuletzt geprüft am 20.06.2016.

Cochrane Haematological Malignancies Group (Hg.) (2007): How to develop a search strategy – a step by step approach. Cochrane Collaboration. Online verfügbar unter <http://chmg.cochrane.org/sites/chmg.cochrane.org/files/uploads/How%20to%20develop%20a%20search%20strategy-support-manual.pdf>, zuletzt geprüft am 30.04.2014.

Cohen, A. M.; Adams, C. E.; Davis, J. M.; Yu, C.; Yu, P. S.; Meng, W. et al. (2010): Evidence-based Medicine, the Essential Role of Systematic Reviews, and the Need for Automated Text Mining Tools. In: T. Veinot (Hg.): Proceedings of the 1st ACM International Health Informatics Symposium. 1882992: ACM, S. 376–380. Online verfügbar unter <http://doi.acm.org/10.1145/1882992.1883046>.

Concato, J.; Shah, N.; Horwitz, R. I. (2000): Randomized, controlled trials, observational studies, and the hierarchy of research designs. In: *N. Engl. J. Med.* 342 (25), S. 1887–1892. DOI: 10.1056/NEJM200006223422507.

Conn, V. S.; Isaramalai, S.; Rath, S.; Jantarakupt, P.; Wadhawan, R.; Dash, Y. (2003a): Beyond MEDLINE for Literature Searches. In: *J Nursing Scholarship* 35 (2), S. 177–182. DOI: 10.1111/j.1547-5069.2003.00177.x.

Conn, V. S.; Valentine, J. C.; Cooper, H. M.; Rantz, M. J. (2003b): Grey literature in meta-analyses. In: *Nursing research* 52 (4), S. 256–261.

Cook, A. M.; Finlay, I. G.; Edwards, A. G.K.; Hood, K.; Higginson, I. J.; Goodwin, D. M. et al. (2001): Efficiency of Searching the Grey Literature in Palliative Care. In: *Journal of Pain and Symptom Management* 22 (3), S. 797–801. DOI: 10.1016/S0885-3924(01)00315-3.

Cooke, A.; Smith, D.; Booth, A. (2012): Beyond PICO: the SPIDER tool for qualitative evidence synthesis. In: *Qual Health Res* 22 (10), S. 1435–1443. DOI: 10.1177/1049732312452938.

Cooper, H. M. (1986): Literature-Searching Strategies of Integrative Research Reviewers A First Survey. In: *Science Communication* 8 (2), S. 372–383. DOI: 10.1177/107554708600800217.

Cooper, H. M. (1988): Organizing knowledge syntheses. A taxonomy of literature reviews. In: *Knowledge in Society* 1 (1), S. 104–126. DOI: 10.1007/BF03177550.

Corbin, J. M.; Strauss, A. L. (1990): Grounded theory research: Procedures, canons, and evaluative criteria. *Qualitative Sociology*. In: *Qual Sociol* 13 (1), S. 3–21. DOI: 10.1007/BF00988593.

Corlan, A. Dan (2004): Medline trend: automated yearly statistics of PubMed results for any query. Online verfügbar unter <http://dan.corlan.net/medline-trend.html>, zuletzt aktualisiert am 17.05.2016, zuletzt geprüft am 21.06.2016.

Cornacchia, R.; Kamps, J.; Alink, W.; Vries, A. P. de (2013): Searching political data by strategy. In: M. Lupu, M. Salampasis, N. Fuhr, A. Hanbury, B. Larsen und H. Strindberg (Hg.): *Proceedings of the Integrating IR Technologies for Professional Search Workshop, Moscow, Russia (March 24, 2013)*. IRPS 2013 Integrating IR Technologies for Professional Search. Moscow, Russia, 24.03.2013, S. 88–91.

Coumou, H. C.H.; Meijman, F. J. (2006): How do primary care physicians seek answers to clinical questions? A literature review. In: *J Med Libr Assoc* 94 (1), S. 55–60.

Courtright, C. (2007): Context in information behavior research. In: *Annual review of information science and technology* 41 (1), S. 273–306. DOI: 10.1002/aris.2007.1440410113.

Craig, P.; Dieppe, P.; Macintyre, S.; Michie, S.; Nazareth, I.; Petticrew, M. (2008): Developing and evaluating complex interventions: the new Medical Research Council guidance. In: *BMJ* 337, S. a1655. DOI: 10.1136/bmj.a1655.

Crain, S. P.; Zhou, K.; Yang, S.-H.; Zha, H. (2012): Dimensionality Reduction and Topic Modeling: From Latent Semantic Indexing to Latent Dirichlet Allocation and Beyond. In: C. C. Aggarwal und C. Zhai (Hg.): Mining Text Data. Boston, MA: Springer US, S. 129–161.

Daly, J. (2006): A Short History of Evidence-Based Medicine. In: Victor M. Montori (Hg.): Evidence-Based Endocrinology. Totowa, NJ: Humana Press (Contemporary endocrinology), S. 11–24.

Das Deutsche Cochrane Zentrum (Hg.) (2014): Handsearching-Projekt in Freiburg. Online verfügbar unter <http://www.cochrane.de/handsearching-projekt-freiburg>, zuletzt aktualisiert am 22.01.2014, zuletzt geprüft am 24.02.2014.

Davies, K. (2011): Formulating the Evidence Based Practice Question: A Review of the Frameworks. In: *Evidence Based Library and Information Practice* 6 (2), S. 75–80. Online verfügbar unter <http://ejournals.library.ualberta.ca/index.php/EBLIP/article/view/9741>, zuletzt geprüft am 13.11.2014.

Davis, C. H. (1995): Beyond Boole: The next logical step. In: *Bulletin of the American Society for Information* (21), S. 17–20. Online verfügbar unter <https://www.asis.org/Bulletin/Jun-95/cdavis.html>, zuletzt geprüft am 28.04.2016.

Dervin, B. (1997): Given a context by any other name: methodological tools for taming the unruly beast. In: P. Vakkari, B. Dervin, Savolainen und Reijo (Hg.): ISIC '96 Proceedings of an international conference on Information seeking in context. Information seeking in context. Tampere, Finland. Taylor Graham Publishing, S. 13–38. Online verfügbar unter <http://dl.acm.org/citation.cfm?id=267191>, zuletzt geprüft am 23.02.2016.

Dervin, B.; Nilan, M. (1986): Information needs and uses. In: *Annual review of information science and technology* 21, S. 3–33.

Deutsches Cochrane Zentrum (2014): Das Deutsche Cochrane Zentrum. Online verfügbar unter <http://www.cochrane.de>, zuletzt aktualisiert am 04.08.2014, zuletzt geprüft am 04.08.2014.

Deutsches Netzwerk Evidenzbasierte Medizin e.V. (Hg.) (2011): Definitionen — Deutsches Netzwerk Evidenzbasierte Medizin e.V. Online verfügbar unter <http://www.ebm-netzwerk.de/was-ist-ebm/grundbegriffe/definitionen/>, zuletzt aktualisiert am 30.12.2011, zuletzt geprüft am 21.06.2016.

Deutsches Netzwerk Evidenzbasierte Medizin e.V. (Hg.) (2012): Geschichte der EbM — Deutsches Netzwerk Evidenzbasierte Medizin e.V. Online verfügbar unter <http://www.ebm-netzwerk.de/was-ist-ebm/geschichte>, zuletzt aktualisiert am 29.05.2012, zuletzt geprüft am 21.06.2016.

Deutsches Register Klinischer Studien, Institut für Medizinische Biometrie und Statistik (Hg.) (2016): DRKS - Deutsches Register Klinischer Studien. Universitätsklinikum Freiburg. Online verfügbar unter https://drks-neu.uniklinik-freiburg.de/drks_web/, zuletzt geprüft am 12.11.2014.

Dicenso, A.; Bayley, L.; Haynes, R. B. (2009): Accessing pre-appraised evidence: fine-tuning the 5S model into a 6S model. In: *Evid Based Nurs* 12 (4), S. 99–101. DOI: 10.1136/ebn.12.4.99-b.

Dickersin, K. (1990): The Existence of Publication Bias and Risk Factors for Its Occurrence. In: *JAMA* 263 (10), S. 1385–1389. DOI: 10.1001/jama.1990.03440100097014.

DIMDI: Deutsches Institut für Medizinische Dokumentation und Information (2014): DIMDI - Deutsches Institut für Medizinische Dokumentation und Information. Online verfügbar unter <http://www.dimdi.de/static/de/index.html>, zuletzt aktualisiert am 04.05.16, zuletzt geprüft am 21.06.2016.

DIMDI: Deutsches Institut für Medizinische Dokumentation und Information (2016a): Datenbankangebot ab 2017: Fakten statt Literatur. Online verfügbar unter http://www.dimdi.de/static/de/amg/aktuelles/news_0395.html_319159480.html, zuletzt aktualisiert am 01.02.2016, zuletzt geprüft am 21.06.2016.

DIMDI: Deutsches Institut für Medizinische Dokumentation und Information (2016b): DIMDI - MeSH - Medical Subject Headings. Online verfügbar unter https://www.dimdi.de/static/de/klassi/mesh_umls/mesh/index.htm, zuletzt aktualisiert am 10.02.2016, zuletzt geprüft am 21.06.2016.

DIMDI: Deutsches Institut für Medizinische Dokumentation und Information (2016c): DIMDI - Informationsvermittlungsstellen. Online verfügbar unter <https://www.dimdi.de/static/de/db/service/ivs/index.htm>, zuletzt aktualisiert am 27.01.2016, zuletzt geprüft am 21.06.2016.

Dooley, G.; Noel-Storr, A.; Glanville, J. (2014): The Embase project: the first year. Poster. In: The Cochrane Collaboration (Hg.): Evidence-Informed Public Health: Opportunities and Challenges. Abstracts of the 22nd Cochrane Colloquium. 22nd Cochrane Colloquium. Evidence-Informed Public Health: Opportunities and Challenges. Hyderabad, India, 21.-26.09.2014.

Döring, N.; Bortz, J. (2016): Forschungsmethoden und Evaluation in den Sozial- und Humanwissenschaften. 5. vollst. überarb., aktualisierte und erw. Aufl. Berlin: Springer (Springer-Lehrbuch).

Dreier, G.; Hasselblatt, H.; Antes, G.; Schumacher, M. (2009): Das Deutsche Register Klinischer Studien: Begründung, technische und inhaltliche Aspekte, internationale Einbindung. In: *Bundesgesundheitsbl.* 52 (4), S. 463–468. DOI: 10.1007/s00103-009-0833-5.

Dubben, H.-H.; Beck-Bornholdt, H.-P. (2004): Unausgewogene Berichterstattung in der medizinischen Wissenschaft - publication bias -. Institut für Allgemeinmedizin des Universitätsklinikums Hamburg-Eppendorf. Hamburg. Online verfügbar unter https://www.researchgate.net/publication/228390859_Unausgewogene_Berichterstattung_in_der_medizinischen_Wissenschaft-publication_bias, zuletzt geprüft am 04.08.2014.

Dublin Core Metadata Initiative (DCMI) (2016): DCMI Specifications. Online verfügbar unter <http://dublincore.org/specifications/>, zuletzt geprüft am 17.02.2016.

Duggan, P. F. (1992): Time to abolish "gold standard". In: *BMJ* 304 (6841), S. 1568–1569. DOI: 10.1136/bmj.304.6841.1568-b.

Easterbrook, P.J.; Gopalan, R.; Berlin, J. A.; Matthews, D.R. (1991): Publication bias in clinical research. In: *The Lancet* 337 (8746), S. 867–872. DOI: 10.1016/0140-6736(91)90201-Y.

Egger, M.; Smith, G. D. (1997): Meta-analysis. Potentials and promise. In: *BMJ* 315 (7119), S. 1371–1374. DOI: 10.1136/bmj.315.7119.1371.

Egger, M.; Smith, G. D. (1998): meta-analysis bias in location and selection of studies. In: *BMJ* 316 (7124), S. 61–66. DOI: 10.1136/bmj.316.7124.61.

Egger, M.; Smith, G. D. (2001): Principles of and Procedures for Systematic Reviews. In: Matthias Egger, George Davey Smith und D. G. Altman (Hg.): *Systematic Reviews in Health Care*. London, UK: BMJ Publishing Group, S. 23–42.

Egger, M.; Smith, G. D.; Phillips, A. N. (1997): Meta-analysis. Principles and procedures. In: *BMJ* 315 (7121), S. 1533–1537. DOI: 10.1136/bmj.315.7121.1533.

Ellis, D. (1989): A behavioural approach to information retrieval system design. In: *Journal of Documentation* 45 (3), S. 171–212. DOI: 10.1108/eb026843.

Ellis, D. (1993): Modeling the information-seeking patterns of academic researchers: A grounded theory approach. In: *The Library Quarterly* 63 (4), S. 469–486.

Ellis, D.; Cox, D.; Hall, K. (1993): A comparison of the information seeking patterns of researchers in the physical and social sciences. In: *Journal of Documentation* 49 (4), S. 356–369. DOI: 10.1108/eb026919.

Ellis, D.; Haugan, M. (1997): Modelling the information seeking patterns of engineers and research scientists in an industrial environment. In: *Journal of Documentation* 53 (4), S. 384–403. DOI: 10.1108/EUM0000000007204.

Elsevier (Hg.) (2016a): Excerpta Medica Full Set Series | FS00-8549 | Elsevier. Online verfügbar unter <https://www.elsevier.com/journals/excerpta-medica-full-set-series/fs00-8549>, zuletzt aktualisiert am 15.06.2016, zuletzt geprüft am 21.06.2016.

Elsevier (Hg.) (2016b): Scopus | The largest database of peer-reviewed literature | Elsevier. Online verfügbar unter <https://www.elsevier.com/solutions/scopus>, zuletzt aktualisiert am 17.06.2016, zuletzt geprüft am 21.06.2016.

Elsevier B.V. (Hg.) (2016): Embase: Biomedical Database | Elsevier. Online verfügbar unter <http://www.elsevier.com/online-tools/embase>, zuletzt geprüft am 20.06.2016.

Elsevier R&D Solutions (2015): A Comparison of Emtree® and MeSH®. Online verfügbar unter http://supportcontent.elsevier.com/Support%20Hub/Embase/Files%208%20Attachments/4685-Embase_White%20Paper_Comparison%20of%20Emtree%20and%20MeSH_July%202015.pdf, zuletzt geprüft am 25.02.2016.

Elsweiler, D.; Wilson M.L.; Kirkegaard Lunn, B. (2011): Understanding Casual-Leisure Information Behaviour. In: A. Spink und J. Heinström (Hg.): *New directions in information behaviour*. 1. Aufl. Bingley, U.K.: Emerald Group Publishing Limited, S. 211–241.

Embase: Indexing with Life Science Thesaurus Emtree | Elsevier (2015). Online verfügbar unter <http://www.elsevier.com/online-tools/embase/about/emtree>, zuletzt aktualisiert am 18.03.2015, zuletzt geprüft am 19.03.2015.

European Medicines Agency (Hg.) (2016): EU Clinical Trials Register - Update. Online verfügbar unter <https://www.clinicaltrialsregister.eu/>, zuletzt aktualisiert am 31.05.2016, zuletzt geprüft am 21.06.2016.

Evans, I.; Thornton, H.; Chalmers, I.; Glasziou, P. (2013): *Wo ist der Beweis? Plädoyer für eine evidenzbasierte Medizin*. Unter Mitarbeit von G. Antes. Bern: Verlag Hans Huber. Online verfügbar unter http://de.testingtreatments.org/wp-content/uploads/2013/07/wo_ist_der_beweis_volltext.pdf.

Evans, T.; Brown, H. (2003): Road traffic crashes. Operationalizing equity in the context of health sector reform. In: *Injury Control and Safety Promotion* 10 (1-2), S. 11–12. DOI: 10.1076/icsp.10.1.11.14117.

Felson, D. T. (1992): Bias in meta-analytic research. In: *Journal of Clinical Epidemiology* 45 (8), S. 885–892. DOI: 10.1016/0895-4356(92)90072-U.

Fettke, P. (2006): State-of-the-Art des State-of-the-Art. In: *Wirtsch. Inform.* 48 (4), S. 257–266. DOI: 10.1007/s11576-006-0057-3.

Fidel, R. (1984): Online searching styles. A case-study-based model of searching behavior. In: *J. Am. Soc. Inf. Sci.* 35 (4), S. 211–221. DOI: 10.1002/asi.4630350404.

Fidel, R. (1985): Moves in online searching. In: *Online Review* 9 (1), S. 61–74. DOI: 10.1108/eb024176.

Fidel, R. (1991a): Searchers' selection of search keys: I. The selection routine. In: *Journal of the American Society for Information Science* 42 (7), S. 490–500. DOI: 10.1002/(SICI)1097-4571(199108)42:7<490::AID-ASI4>3.0.CO;2-V.

Fidel, R. (1991b): Searchers' selection of search keys: II. Controlled vocabulary or free-text searching. In: *Journal of the American Society for Information Science* 42 (7), S. 501–514. DOI: 10.1002/(SICI)1097-4571(199108)42:7<501::AID-ASI5>3.0.CO;2-V.

Fidel, R. (1991c): Searchers' selection of search keys: III. Searching styles. In: *Journal of the American Society for Information Science* 42 (7), S. 515–527. DOI: 10.1002/(SICI)1097-4571(199108)42:7<515::AID-ASI6>3.0.CO;2-F.

Fleuren, W. W.M.; Alkema, W. (2015): Application of text mining in the biomedical domain. In: *Text mining of biomedical literature* 74, S. 97–106. DOI: 10.1016/j.jymeth.2015.01.015.

Flick, U. (2007): Zur Qualität qualitativer Forschung — Diskurse und Ansätze. In: U. Kuckartz, H. Grunenberg und T. Dresing (Hg.): *Qualitative Datenanalyse: computergestützt. Methodische Hintergründe und Beispiele aus der Forschungspraxis*. 2., überarb. und erw. Aufl. Wiesbaden: VS, Verl. für Sozialwissenschaften (SpringerLink : Bücher), S. 188–209.

Flynn, M. G.; McGuinness, C. (2011): Hospital clinicians' information behaviour and attitudes towards the 'Clinical Informationist': an Irish survey. In: *Health Info Libr J* 28 (1), S. 23–32. DOI: 10.1111/j.1471-1842.2010.00917.x.

Foxlee, R.; Tovey, D. (2013): A New Search Interface for the Cochrane Library. In: *Cochrane Database Syst Rev* (2). DOI: 10.1002/14651858.ED000052.

Frantzi, K.; Ananiadou, S.; Mima, H. (2000): Automatic recognition of multi-word terms: the C-value/NC-value method. In: *International Journal on Digital Libraries* 3 (2), S. 115–130. DOI: 10.1007/s007999900023.

Fuhr, N. (2013): Modelle im Information Retrieval. In: Rainer Kuhlen, Wolfgang Semar und Dietmar Strauch (Hg.): *Grundlagen der praktischen Information und Dokumentation. Handbuch zur Einführung in die Informationswissenschaft und -praxis*. 6., völlig neu gefasste Ausg. Berlin: De Gruyter, S. 322–335.

Garfield, E. (1964): Citation indexing: A natural science literature retrieval system for the social sciences. In: *The American Behavioral Scientist (pre-1986)* 7 (10), S. 59.

Gehanno, J.-F.; Rollin, L.; Darmoni, S. (2013): Is the coverage of Google Scholar enough to be used alone for systematic reviews. In: *BMC Med Inform Decis Mak* 13, S. 7. DOI: 10.1186/1472-6947-13-7.

Giustini, D. (2005): How Google is changing medicine. In: *BMJ* 331 (7531), S. 1487–1488. DOI: 10.1136/bmj.331.7531.1487.

Giustini, D.; Kamel, B.; Maged, N. (2013): Google Scholar is not enough to be used alone for systematic reviews. In: *OJPHI* 5 (2). DOI: 10.5210/ojphi.v5i2.4623.

Glanville, J. M.; Bayliss, S.; Booth, A.; Dundar, Y.; Fernandes, H.; Fleeman, N. David et al. (2008): So many filters, so little time: the development of a search filter appraisal checklist. In: *J Med Libr Assoc* 96 (4), S. 356–361. DOI: 10.3163/1536-5050.96.4.011.

Glanville, J. M.; Duffy, S.; McCool, R.; Varley, D. (2014): Searching ClinicalTrials.gov and the International Clinical Trials Registry Platform to inform systematic reviews: what are the optimal search approaches?*. In: *J Med Libr Assoc* 102 (3), S. 177–183. DOI: 10.3163/1536-5050.102.3.007.

Glaser, B. G.; Strauss, A. L. (1967): The discovery of grounded theory. Strategies for qualitative research. Chicago: Aldine Pub. Co. (Observations).

Google (Hg.) (2013): About Google Scholar. Online verfügbar unter <http://scholar.google.com/intl/en/scholar/about.html>, zuletzt aktualisiert am 22.08.2013, zuletzt geprüft am 09.03.2016.

Google (Hg.) (2016): Google Scholar. Online verfügbar unter <http://scholar.google.de/>, zuletzt geprüft am 08.11.2014.

Gough, D.; Oliver, S.; Newman, M.; Bird, K. (2009): Transparency in planning, warranting and interpreting research.

Gough, D.; Oliver, S.; Thomas, J. (2012): An introduction to systematic reviews. London, Thousand Oaks, Calif.: Sage.

Grant, M. J. (2004): How does your searching grow? A survey of search preferences and the use of optimal search strategies in the identification of qualitative research. In: *Health Info Libr J* 21 (1), S. 21–32. DOI: 10.1111/j.1471-1842.2004.00483.x.

Grant, M. J.; Booth, A. (2009): A typology of reviews: an analysis of 14 review types and associated methodologies. In: *Health Info Libr J* 26 (2), S. 91–108. DOI: 10.1111/j.1471-1842.2009.00848.x.

Gray, J. E.; Hamilton, M.C.; Hauser, A.; Janz, M. M.; Peters, J.P.; Taggart, F. (2012): Scholarish: Google Scholar and its Value to the Sciences. In: *Issues in Science and Technology Librarianship*. Online verfügbar unter DOI:10.5062/F4MK69T9, zuletzt geprüft am 04.03.2015.

Gujarat Technological University (2013): Patent Search Methodology. Online verfügbar unter http://www.gtu.ac.in/circulars/13Aug/Module%20Patent%20Search%20Methodology__3rdAug2013.pdf, zuletzt geprüft am 29.04.2016.

Guyatt, G.; Rennie, D. (2002): Users' guides to the medical literature. A manual for evidence-based clinical practice / the Evidence-Based Medicine Working Group. Chicago, IL: AMA Press.

Hansen, P. (1999): User Interface Design for IR Interaction: A Task-Oriented Approach. Research Paper. In: T. Aparac, T. Saracevic, P. Ingwersen und P. Vakkari (Hg.): *Digital Libraries: Interdisciplinary Concepts, Challenges and Opportunities*. CoLIS3 Proceedings. Digital Libraries: Interdisciplinary Concepts, Challenges and Opportunities. Dubrovnik, Croatia, 23-26 May 1999: Benja Publishing, Lokve, Croatia. Online verfügbar unter https://www.researchgate.net/publication/221507626_User_Interface_Design_for_IR_Interaction_A_Task-Oriented_Approach, zuletzt geprüft am 21.02.2016.

Hansen, P.; Järvelin, A.; Eriksson, G.; Karlgren, J. (2014): A Use Case Framework for Information Access Evaluation. In: G. Paltoglou, F. Loizides und P. Hansen (Hg.): *Professional Search in the Modern World*. Cham: Springer International Publishing (Lecture Notes in Computer Science), S. 6–22.

Hansen, P.; Järvelin, K. (2005): Collaborative Information Retrieval in an information-intensive domain. In: *Information Processing & Management* 41 (5), S. 1101–1119. DOI: 10.1016/j.ipm.2004.04.016.

Harker, J.; Kleijnen, J. (2012): What is a rapid review? A methodological exploration of rapid reviews in Health Technology Assessments. In: *Int J Evid Based Healthc* 10 (4), S. 397–410. DOI: 10.1111/j.1744-1609.2012.00290.x.

- Harris, M. R. (2005): The librarian's roles in the systematic review process: a case study. In: *Journal of the Medical Library Association : JMLA* 93 (1), S. 81–87.
- Hausner, E.; Guddat, C.; Hermanns, T.; Lampert, U.; Waffenschmidt, S. (2015): Development of search strategies for systematic reviews: validation showed the noninferiority of the objective approach. In: *J Clin Epidemiol* 68 (2), S. 191–199. DOI: 10.1016/j.jclinepi.2014.09.016.
- Hausner, E.; Kaiser, T. (2010): Suche in Studienregistern: Ablauf und Dokumentation. In: *Verzerrte Welt: Auswirkungen von Publikationsbias auf die Patientenversorgung* 104 (4), S. 292–297. DOI: 10.1016/j.zefq.2010.03.032.
- Hausner, E.; Waffenschmidt, S.; Kaiser, T.; Simon, M. (2012): Routine development of objectively derived search strategies. In: *Sys Rev* 1 (1). DOI: 10.1186/2046-4053-1-19.
- Haynes, R. B. (2001): Of studies, syntheses, synopses, and systems. The "4S" evolution of services for finding current best evidence. In: *Evidence-Based Medicine* 6 (2), S. 36–38. DOI: 10.1136/ebm.6.2.36.
- Haynes, R. B. (2004): Optimal search strategies for retrieving scientifically strong studies of diagnosis from Medline: analytical survey. In: *BMJ* 328 (7447), S. 1040. DOI: 10.1136/bmj.38068.557998.EE.
- Haynes, R. B. (2005): Optimal search strategies for retrieving scientifically strong studies of treatment from Medline: analytical survey. In: *BMJ* 330 (7501), S. 1179. DOI: 10.1136/bmj.38446.498542.8F.
- Haynes, R. B.; McKibbon, K. A.; Wilczynski, N. L.; Walter, S. D.; Werre, S. R. (2005): Optimal search strategies for retrieving scientifically strong studies of treatment from Medline: analytical survey. In: *BMJ* 330.
- Health Information Research Unit (Hg.) (2016): Health Information Research Unit - HIRU ~ McMaster PLUS. McMaster University. Online verfügbar unter http://hiru.mcmaster.ca/hiru/HIRU_McMaster_PLUS_projects.aspx, zuletzt aktualisiert am 09.02.2016, zuletzt geprüft am 21.06.2016.

Hearst, M. (2009): Search user interfaces. Online Ausgabe. Cambridge, New York: Cambridge University Press. Online verfügbar unter <http://searchuserinterfaces.com/book/>, zuletzt geprüft am 28.06.2016.

Henderson, J. (2005): Google Scholar: A source for clinicians? In: *CMAJ* 172 (12), S. 1549–1550. DOI: 10.1503/cmaj.050404.

Herskovic, J. R.; Tanaka, L. Y.; Hersh, W.; Bernstam, E. V. (2007): A day in the life of PubMed: analysis of a typical day's query log. In: *Journal of the American Medical Informatics Association : JAMIA* 14 (2), S. 212–220. DOI: 10.1197/jamia.M2191.

Hertzberg, S.; Rudner, L. (1999): Quality of Researchers' Searches of the ERIC Database. In: *education policy analysis archives* 7 (0), S. 25. DOI: 10.14507/epaa.v7n25.1999.

Higgins, J. P.T.; Green, S. (Hg.) (2011): Cochrane Handbook for Systematic Reviews of Interventions. Version 5.1.0 [updated March 2011]. [S.l.]: Cochrane Collaboration. Online verfügbar unter <http://handbook.cochrane.org/>.

Hjørland, B. (2010): The foundation of the concept of relevance. In: *Journal of the American Society for Information Science and Technology* 61 (2), S. 217–237. DOI: 10.1002/asi.21261.

Hjørland, B. (2015): Classical databases and knowledge organization: A case for boolean retrieval and human decision-making during searches. In: *Journal of the Association for Information Science and Technology* 66 (8), S. 1559–1575. DOI: 10.1002/asi.23250.

Hjørland, B. (2016): Does the traditional thesaurus have a place in modern information retrieval. In: *Knowledge Organization* 43 (3).

Hopewell, S.; Clarke, M.; Lefebvre, C.; Scherer, R. (2007a): Handsearching versus electronic searching to identify reports of randomized trials. In: *Cochrane Database Syst Rev* (2), S. MR000001. DOI: 10.1002/14651858.MR000001.pub2.

Hopewell, S.; Clarke, M.; Lusher, A.; Lefebvre, C.; Westby, M. (2002): A comparison of handsearching versus MEDLINE searching to identify reports of randomized controlled trials. In: *Stat Med* 21 (11), S. 1625–1634. DOI: 10.1002/sim.1191.

Hopewell, S.; McDonald, S.; Clarke, M.; Egger, M. (2007b): Grey literature in meta-analyses of randomized trials of health care interventions. In: *Cochrane Database Syst Rev* (2), S. MR000010. DOI: 10.1002/14651858.MR000010.pub3.

Hopewell, S.; McDonald, S.; Clarke, M.; Egger, M.; Hopewell, S. (1996): Grey literature in meta-analyses of randomized trials of health care interventions. In: *Cochrane Database of Systematic Reviews*. Chichester, UK: John Wiley & Sons, Ltd.

Horsley, T.; Dingwall, O.; Sampson, M. (2011): Checking reference lists to find additional studies for systematic reviews. In: *Cochrane Database Syst Rev* (8), S. MR000026. DOI: 10.1002/14651858.MR000026.pub2.

Huston, M.; Williams, R. J.; Bergeris, A.; Fun, J.; Tse, T. (2014): New Style and New Content for ClinicalTrials.gov. Hg. v. U.S. National Library of Medicine. Lister Hill Center for Biomedical Communications (NLM Tech Bull., 2012 Jul-Aug;(387):e5). Online verfügbar unter http://www.nlm.nih.gov/pubs/techbull/ja12/ja12_clinicaltrials.html, zuletzt aktualisiert am 14.02.2014, zuletzt geprüft am 12.11.2014.

Ingwersen, P. (1992): Information and Information Science in Context. In: *Libri* 42 (2), S. 99–135. DOI: 10.1515/libr.1992.42.2.99.

Ingwersen, P. (1996): Cognitive Perspectives of Information Retrieval Interaction: Elements of a cognitive IR theory. In: *Journal of Documentation* 52 (1), S. 3–50. DOI: 10.1108/eb026960.

INIST-CNRS - Institut de l'Information Scientifique et Technique (Hg.) (2016a): Grey Literature - GreySource, A Selection of Web-based Resources in Grey Literature. Online verfügbar unter <http://www.greynet.org/greysourceindex.html>, zuletzt aktualisiert am 19.06.2016, zuletzt geprüft am 21.06.2016.

INIST-CNRS - Institut de l'Information Scientifique et Technique (Hg.) (2016b): OpenGrey. Online verfügbar unter <http://www.opengrey.eu/>, zuletzt geprüft am 21.06.2016.

Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen (IQWiG) (2015): IQWiG: Allgemeine Methoden 4.2. Köln, zuletzt geprüft am 29.02.2016.

Institute of Education Sciences (IES) (Hg.) (2016): ERIC - Education Resources Information Center. U.S. Department of Education. Online verfügbar unter <https://eric.ed.gov/>, zuletzt geprüft am 21.06.2016.

InterTASC Information Specialists' Sub-Group (ISSG) (2015): ISSG Search Filters Resource. Online verfügbar unter <https://sites.google.com/a/york.ac.uk/issg-search-filters-resource/>, zuletzt aktualisiert am 12.02.2015, zuletzt geprüft am 19.02.2015.

Järvelin, K.; Ingwersen, P. (2004): Information seeking research needs extension towards tasks and technology. In: *Information Research* 10 (1).

Järvelin, K.; Kekäläinen, J. (2002): Cumulated gain-based evaluation of IR techniques. In: *ACM Trans. Inf. Syst.* 20 (4), S. 422–446. DOI: 10.1145/582415.582418.

Jenkins, M. (2004): Evaluation of methodological search filters--a review. In: *Health Info Libr J* 21 (3), S. 148–163. DOI: 10.1111/j.1471-1842.2004.00511.x.

Jewell, S.; Brandhorst, W.T. (1973): Search Strategy Tutorial; Searcher's Kit. Online verfügbar unter <http://files.eric.ed.gov/fulltext/ED082763.pdf>.

Jiang, J. (2012): Information Extraction from Text. In: C. C. Aggarwal und C. Zhai (Hg.): *Mining Text Data*. Boston, MA: Springer US, S. 11–41.

John Wiley & Sons, Inc. (Hg.) (2016a): Access Options for Cochrane Library. Online verfügbar unter <http://www.cochranelibrary.com/help/access-options-for-cochrane-library.html>, zuletzt geprüft am 21.06.2016.

John Wiley & Sons, Inc. (Hg.) (2016b): CENTRAL. Online verfügbar unter <http://www.cochranelibrary.com/about/central-landing-page.html>, zuletzt geprüft am 21.06.2016.

John Wiley & Sons, Inc. (Hg.) (2016c): The Cochrane Library. Online verfügbar unter <http://www.cochranelibrary.com/>, zuletzt geprüft am 20.06.2016.

Johnson, J.D. (2003): On contexts of information seeking. In: *Information Processing & Management* 39 (5), S. 735–760. DOI: 10.1016/S0306-4573(02)00030-4.

Joho, H.; Azzopardi, L. A.; Vanderbauwhede, W. (2010): A survey of patent users. an analysis of tasks, behavior, search functionality and system requirements. In: N. J. Belkin und D.

Kelly (Hg.): Proceedings of the third symposium on Information interaction in context. IliX 2010 Information Interaction in Context Symposium. New Brunswick, NJ, USA, August 18 - 21, 2010. ACM, S. 13–24. Online verfügbar unter http://delivery.acm.org/10.1145/1850000/1840789/p13-joho.pdf?ip=132.199.181.149&id=1840789&acc=ACTIVE%20SERVICE&key=24A4B5F6D28C022A%2E4D4702B0C3E38B35%2E4D4702B0C3E38B35%2E4D4702B0C3E38B35&CFID=584839567&CFTOKEN=47213005&__acm__=1457946599_7de53f9e69c771177fc6649151c51b8f, zuletzt geprüft am 14.03.2016.

Jonnalagadda, S.; Petitti, D. (2013): A new iterative method to reduce workload in systematic review process. In: *IJCBDD* 6 (1/2), S. 5. DOI: 10.1504/IJCBDD.2013.052198.

Jüni, P.; Altman, D. G.; Egger, M. (2001): Systematic reviews in health care. Assessing the quality of controlled clinical trials. In: *BMJ* 323 (7303), S. 42–46. DOI: 10.1136/bmj.323.7303.42.

Jürgens, J. J.; Womser-Hacker, C.; Mandl, T. (2014): Modeling the interactive patent retrieval process: an adaptation of Marchionini's information seeking model. In: D. Elswiler und B. Ludwig (Hg.): IliX '14: Proceedings of the 5th Information Interaction in Context Symposium. IliX '14: 5th Information Interaction in Context Symposium. Regensburg, 26.-30. August 2014. New York, NY, USA: ACM, S. 247–250. Online verfügbar unter http://delivery.acm.org/10.1145/2640000/2637034/p247-jurgens.pdf?ip=132.199.40.169&id=2637034&acc=ACTIVE%20SERVICE&key=24A4B5F6D28C022A%2E4D4702B0C3E38B35%2E4D4702B0C3E38B35%2E4D4702B0C3E38B35&CFID=770046509&CFTOKEN=11113321&__acm__=1460535493_48cb15fe1deb9ef882cc875b9e1e1bc8.

Kastner, M.; Straus, S. E.; McKibbin, K. A.; Goldsmith, C. H. (2009): The capture-mark-recapture technique can be used as a stopping rule when searching in systematic reviews. In: *J Clin Epidemiol* 62 (2), S. 149–157. DOI: 10.1016/j.jclinepi.2008.06.001.

Kastner, M.; Tricco, A. C.; Soobiah, C.; Lillie, E.; Perrier, L.; Horsley, T. et al. (2012): What is the most appropriate knowledge synthesis method to conduct a review? Protocol for a scoping review. In: *BMC Med Res Methodol* 12, S. 114. DOI: 10.1186/1471-2288-12-114.

Katchamart, W.; Faulkner, A.; Feldman, B.; Tomlinson, G.; Bombardier, C. (2011): PubMed had a higher sensitivity than Ovid-MEDLINE in the search for systematic reviews. In: *Journal of Clinical Epidemiology* 64 (7), S. 805–807. DOI: 10.1016/j.jclinepi.2010.06.004.

Khangura, S.; Konnyu, K.; Cushman, R.; Grimshaw, J.; Moher, D. (2012): Evidence summaries: the evolution of a rapid review approach. In: *Systematic reviews* 1 (1), S. 1. DOI: 10.1186/2046-4053-1-10.

Kitchenham, B. (2004): Procedures for Performing Systematic Reviews. Keele University Technical Report TR/SE-0401.

Kloda, L.; Bartlett, J. (2013): Formulating answerable questions: Question negotiation in evidence-based practice. In: *Journal of the Canadian Health Libraries Association* 34, S. 55–60.

Korjonen-Close, H. (2015): The information needs and behaviour of clinical researchers: a user-needs analysis. In: *Health information and libraries journal* 22, S. 96–106. Online verfügbar unter <http://onlinelibrary.wiley.com/store/10.1111/j.1471-1842.2005.00562.x/asset/j.1471-1842.2005.00562.x.pdf?v=1&t=i5c9u089&s=6283de5de7581115d28bec3b1f7309cdc285d943>, zuletzt geprüft am 25.01.2015.

Koster, J. (2016): PubMed PubReMiner: a tool for PubMed query building and literature mining. Online verfügbar unter <http://hgserver2.amc.nl/cgi-bin/miner/miner2.cgi>, zuletzt geprüft am 21.06.2016.

Kugley, S.; Wade, A.; Thomas, J.; Mahood, Q.; Jørgensen, A.-M. K.; Hammerstrøm, K.; Sathe, N. (2015): Searching for studies. A guide to information retrieval for Campbell Systematic Reviews. Hg. v. The Campbell Collaboration. Online verfügbar unter [10.4073/csrs.2010.1](http://www.campbellcollaboration.org/10.4073/csrs.2010.1), zuletzt aktualisiert am 30.10.2015.

Kuhlen, R. (1995): Informationsmarkt. Chancen und Risiken der Kommerzialisierung von Wissen. Konstanz: UVK, Univ.-Verl. (Schriften zur Informationswissenschaft, 15).

Kuhlthau, C. C. (1988): Developing a Model of the Library Search Process: Cognitive and Affective Aspects. In: *RQ* 28 (2), S. 232–242.

Kuhlthau, C. C. (1991): Inside the search process: Information seeking from the user's perspective. In: *JASIS* 42 (5), S. 361–371.

Kuhlthau, C. C. (1993a): An uncertainty principle for information seeking. In: *Journal of Documentation* 49 (4), S. 339–355.

Kuhlthau, C. C. (1993b): Seeking meaning: A process approach to library and information services: Ablex Norwood, NJ.

Kuhlthau, C. C. (1997): The influence of uncertainty on the information seeking behavior of a securities analyst: Taylor Graham Publishing.

Kuhlthau, C. C. (1999): The role of experience in the information search process of an early career information worker: Perceptions of uncertainty, complexity, construction, and sources. In: *Journal of the American Society for Information Science* 50 (5), S. 399–412. DOI: 10.1002/(SICI)1097-4571(1999)50:5<399::AID-ASIS>3.0.CO;2-L.

Latin American and Caribbean Center on Health Sciences Information (Hg.) (2016): LILACS. Online verfügbar unter <http://lilacs.bvsalud.org/en/>, zuletzt geprüft am 21.06.2016.

Leckie, G. J.; Pettigrew, K. E.; Sylvain, C. (1996): Modeling the information seeking of professionals. A general model derived from research on engineers, health care professionals, and lawyers. In: *The library quarterly : a journal of investigation in library and information studies*.

Lee, J. H.; Renear, A.; Smith, L. C. (2006): Known-Item Search: Variations on a Concept. In: *Proceedings of the ASIST Annual Meeting* 43.

Lee, S.-S. (2010): Library Service Users' Information Seeking Behavior Model Based on the Grounded Theory Methodology. In: *Journal of the Korean Society for Library and Information Science* 44 (1), S. 5–28. DOI: 10.4275/KSLIS.2010.44.1.005.

Lefebvre, C.; Glanville, J. M.; Wieland, L. S.; Coles, B.; Weightman, A. L. (2013): Methodological developments in searching for studies for systematic reviews: past, present and future? In: *Sys Rev* 2 (1), S. 78.

Lefebvre, C.; Manheimer, E.; Glanville, J. M. (2011): Searching for Studies. In: J. P.T. Higgins und S. Green (Hg.): *Cochrane Handbook for Systematic Reviews of Interventions*. Version 5.1.0 [updated March 2011]. [S.l.]: Cochrane Collaboration.

Lepsky, K. (2013): Automatische Indexierung. In: Rainer Kuhlen, Wolfgang Semar und Dietmar Strauch (Hg.): Grundlagen der praktischen Information und Dokumentation. Handbuch zur Einführung in die Informationswissenschaft und -praxis. 6., völlig neu gefasste Ausg. Berlin: De Gruyter, S. 272–285.

Levac, D.; Colquhoun, H.; O'Brien, K. K. (2010): Scoping studies: advancing the methodology. In: *Implement Sci* 5, S. 69. DOI: 10.1186/1748-5908-5-69.

Lewandowski, D. (2013): Suchmaschinen. In: Rainer Kuhlen, Wolfgang Semar und Dietmar Strauch (Hg.): Grundlagen der praktischen Information und Dokumentation. Handbuch zur Einführung in die Informationswissenschaft und -praxis. 6., völlig neu gefasste Ausg. Berlin: De Gruyter, S. 494–508.

Li, Y.; Belkin, N. J. (2008): A faceted approach to conceptualizing tasks in information seeking. In: *Information Processing & Management* 44 (6), S. 1822–1837. DOI: 10.1016/j.ipm.2008.07.005.

Liberati, A.; Altman, D. G.; Tetzlaff, J.; Mulrow, C.; Gøtzsche, P. C.; Ioannidis, J. P.A. et al. (2009): The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate healthcare interventions: explanation and elaboration. In: *BMJ* 339, S. b2700.

Line, M. B. (1971): The information uses and needs of social scientists: an overview of INFROSS. In: *Aslib proceedings* 23 (8), S. 412–434. Online verfügbar unter <http://www.emeraldinsight.com/doi/abs/10.1108/eb050298>.

Loveman, E.; Al-Khudairy, L.; Johnson, R. E.; Robertson, W.; Colquitt, J. L.; Mead, E. L. et al. (2015): Parent-only interventions for childhood overweight or obesity in children aged 5 to 11 years. In: *Cochrane Database Syst Rev*. DOI: 10.1002/14651858.CD012008.

Lu, Z.; Wilbur, W. J.; McEntyre, JR; Iskhakov, A.; Szilagyi, L. (2009): Finding Query Suggestions for PubMed. In: *AMIA Annual Symposium Proceedings 2009*, S. 396–400.

Lund, N. Windfeld (2010): Document, text and medium: concepts, theories and disciplines. In: *Journal of Documentation* 66 (5), S. 734–749. DOI: 10.1108/00220411011066817.

Lupu, M.; Salampasis, M.; Hanbury, A. (2014): Domain Specific Search. In: G. Paltoglou, F. Loizides und P. Hansen (Hg.): Professional Search in the Modern World. Cham: Springer International Publishing (Lecture Notes in Computer Science), S. 96–117.

Lykke, M.; Price, S.; Delcambre, L. (2012): How doctors search: A study of query behaviour and the impact on search results. In: *Information Processing & Management* 48 (6), S. 1151–1170. DOI: 10.1016/j.ipm.2012.02.006.

Manning, C. D.; Raghavan, P.; Schütze, H. (2008): Introduction to information retrieval. Reprinted. Cambridge: Cambridge Univ. Press. Online verfügbar unter <http://nlp.stanford.edu/IR-book/>, zuletzt geprüft am 16.06.2016.

Marchionini, G. (1995): Information seeking in electronic environments. Cambridge: Cambridge Univ. Press (Cambridge series on human-computer interaction, 9). Online verfügbar unter <http://lccn.loc.gov/94044629>.

Markey, K.; Cochrane, P. A. (1978): ONTAP: Online Training and Practice Manual for ERIC Data Base Searchers. 2nd Edition. Online verfügbar unter <http://files.eric.ed.gov/full-text/ED212296.pdf>.

Marshall, C. (2016): Systematic Review Toolbox. York Health Economics Consortium, University of York. Online verfügbar unter <http://systematicreviewtools.com/contact.php>, zuletzt geprüft am 24.06.2016.

Mayr, P. (2006): Thesauri, Klassifikationen & Co - die Renaissance der kontrollierten Vokabulare? In: Petra Hauke und K. Umlauf (Hg.): Vom Wandel der Wissensorganisation im Informationszeitalter. Festschrift für Walther Umstätter zum 65. Geburtstag. Berlin: Institut für Bibliotheks- und Informationswissenschaft (Vom Wandel der Wissensorganisation im Informationszeitalter), S. 151–170.

Mayring, P. (2000): Qualitative Content Analysis. In: *Forum Qualitative Sozialforschung / Forum: Qualitative Social Research* 1 (2). Online verfügbar unter <http://www.qualitative-research.net/index.php/fqs/article/download/1089/2384>, zuletzt geprüft am 21.06.2016.

Mayring, P. (2015): Qualitative Inhaltsanalyse. Grundlagen und Techniken. 12., überarb. Aufl. Weinheim: Beltz (Beltz Pädagogik).

Mays, N.; Roberts, E.; Popay, J. (2001): Synthesising research evidence. In: *Studying the organisation and delivery of health services: Research methods*, S. 188–220.

McGowan, J.; Sampson, M. (2005): Systematic reviews need systematic searchers. In: *Journal of the Medical Library Association : JMLA* 93 (1), S. 74–80.

McKibbin, K. A. (2006): Systematic Reviews and Librarians. In: *Library Trends* 55 (1), S. 202–215. Online verfügbar unter <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.178.7202&rep=rep1&type=pdf>, zuletzt geprüft am 04.02.2015.

McMaster University Health Information Research Unit (Hg.) (2013): Health Information Research Unit - HIRU ~ Hedges Project. McMaster University. Online verfügbar unter http://hiru.mcmaster.ca/hiru/HIRU_Hedges_home.aspx, zuletzt aktualisiert am 30.04.2013, zuletzt geprüft am 02.03.2015.

Mead, T. L.; Richards, D. T. (1995): Librarian participation in meta-analysis projects. In: *Bulletin of the Medical Library Association* 83 (4), S. 461–464.

Meho, L. I.; Tibbo, H. R. (2003): Modeling the information-seeking behavior of social scientists: Ellis's study revisited. In: *J. Am. Soc. Inf. Sci.* 54 (6), S. 570–587. DOI: 10.1002/asi.10244.

Merz, A.-K.; Knüttel, H. (2015): Search and Retrieval Strategies in the Systematic Literature Search. Terminologies and Taxonomies in Information and Library Sciences. In: The Cochrane Collaboration (Hg.): 23rd Cochrane Colloquium. Filtering the information overload for better decisions. 23rd Cochrane Colloquium: Filtering the information overload for better decisions. Wien, 03.-07.10.2015. The Cochrane Collaboration, S. 272–273. Online verfügbar unter <http://epub.uni-regensburg.de/32556/>.

Mielke, B.; Wolff, C. (2008): Wissenschaftsportale und virtuelle Fachbibliotheken in der juristischen Recherchepraxis. In: E. Schweighofer, A. Geist, G. Heindl und C. Szücs (Hg.): Komplexitätsgrenzen der Rechtsinformatik. Tagungsband des 11. Internationalen Rechtsinformatik Symposions. 11. Internationales Rechtsinformatik Symposium IRIS 2008. Salzburg, 21.-23. Februar 2008. Stuttgart: Richard Boorberg Verlag (Internationales Rechtsinformatik Symposium IRIS 2008), S. 256–268.

- Miwa, M.; Thomas, J.; O'Mara-Eves, A.; Ananiadou, S. (2014): Reducing systematic review workload through certainty-based screening. In: *Journal of Biomedical Informatics* 51, S. 242–253. DOI: 10.1016/j.jbi.2014.06.005.
- Mizzaro, S. (1997): Relevance: The whole history. In: *J. Am. Soc. Inf. Sci.* 48 (9), S. 810–832. DOI: 10.1002/(SICI)1097-4571(199709)48:9<810::AID-ASL6>3.0.CO;2-U.
- Mo, Y.; Kontonatsios, G.; Ananiadou, S. (2015): Supporting systematic reviews using LDA-based document representations. In: *Systematic reviews* 4 (1), S. 1. DOI: 10.1186/s13643-015-0117-0.
- Moher, D.; Liberati, A.; Tetzlaff, J.; Altman, D. G.; The PRISMA Group (2009): Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. In: *PLoS Med* 6 (7), S. e1000097. DOI: 10.1371/journal.pmed.1000097.
- Monk, J. Donald (2002): *The Mathematics of Boolean Algebra*. Center for the Study of Language and Information (CSLI), Stanford University. Online verfügbar unter <http://plato.stanford.edu/entries/boolalg-math/>, zuletzt geprüft am 18.02.2016.
- Mosa, A. S. M.; Yoo, I. (2013): A study on PubMed search tag usage pattern: association rule mining of a full-day PubMed query log. In: *BMC medical informatics and decision making* 13, S. 8. DOI: 10.1186/1472-6947-13-8.
- Motschall, E. (2010): *Vor/Nachteile bei PubMed und Ovid*. Hg. v. Department für Medizinische Biometrie und Medizinische Informatik. Universität Freiburg. Online verfügbar unter <https://portal.uni-freiburg.de/imbi/dateien/bibliothek/pdfs/vor-nachteile-pubmed-ovid.pdf>, zuletzt geprüft am 25.02.2016.
- Motschall, E. (2014): *Vergleich der Syntax von Pubmed und Ovid*. Hg. v. Department für Medizinische Biometrie und Medizinische Informatik. Universität Freiburg. Online verfügbar unter <https://portal.uni-freiburg.de/imbi/dateien/bibliothek/pdfs/syntax-ovid-pubmed>, zuletzt geprüft am 25.02.2016.
- Muckel, P. (2007): Die Entwicklung von Kategorien mit der Methode der Grounded Theory. In: *Historical Social Research/Historische Sozialforschung. Supplement*, S. 211–231. Online verfügbar unter <http://nbn-resolving.de/urn:nbn:de:0168-ssoar-288620>.

National Center for Biotechnology Information (Hg.) (2016): Home -MeSH -NCBI. U.S. National Library of Medicine. Online verfügbar unter <http://www.ncbi.nlm.nih.gov/mesh>, zuletzt geprüft am 21.06.2016.

National Center for Biotechnology Information (US) (2014): PubMed Help. Bethesda (MD): National Center for Biotechnology Information (US); 2005-. PubMed Help. [Updated 2014 Sep 18]. Online verfügbar unter <http://www.ncbi.nlm.nih.gov/books/NBK3827/#pubmed-help.FAQs>, zuletzt aktualisiert am 19.10.2014, zuletzt geprüft am 19.10.2014.

National Center for Biotechnology Information (US) (2016): [Table, MeSH Subheadings]. Online verfügbar unter http://www.ncbi.nlm.nih.gov/books/NBK3827/table/pubmed-help.T.mesh_subheadings/?report=objectonly, zuletzt aktualisiert am 14.02.2016, zuletzt geprüft am 21.06.2016.

National Center for Biotechnology Information (NCBI) (Hg.) (2016): Home - PubMed - NCBI. National Institutes of Health (NIH). Online verfügbar unter <http://www.ncbi.nlm.nih.gov/pubmed>, zuletzt geprüft am 21.06.2016.

National Health; Medical Research Council (2000): How to review the evidence: systematic identification and review of the scientific literature. Canberra: National Health and Medical Research Council (Handbook series on preparing clinical practice guidelines). Online verfügbar unter <http://www.nhmrc.gov.au/guidelines/publications/cp65>, zuletzt geprüft am 08.02.2013.

National Information Standards Organization (U.S.) (2005): Guidelines for the construction, format, and management of monolingual thesauri. An American national standard. Bethesda, Md: NISO Press (National information standards series, ANSI/NISO Z39.19-2005).

National Institute of Standards and Technology (NIST) (Hg.) (2016): Text REtrieval Conference (TREC) Home Page. Online verfügbar unter <http://trec.nist.gov/>, zuletzt aktualisiert am 31.05.2016, zuletzt geprüft am 21.06.2016.

National Library of Medicine (U.S.). Bibliographic Services Division (Hg.): PubMed Clinical Queries. NLM. Online verfügbar unter <https://www.ncbi.nlm.nih.gov/pubmed/clinical>, zuletzt geprüft am 02.03.2015.

National Library of Medicine (U.S.). History of Medicine Division (2014): The Story of NLM Historical Collections. Unter Mitarbeit von National Library of Medicine (U.S.). History of Medicine Division. U.S. National Library of Medicine. Online verfügbar unter <http://www.nlm.nih.gov/hmd/about/collectionhistory.html>, zuletzt aktualisiert am 19.06.2014, zuletzt geprüft am 10.11.2014.

National Technical Information Service (Hg.) (2016): NTIS.gov - NTIS Database. Online verfügbar unter <http://www.ntis.gov/products/ntis-database/>, zuletzt geprüft am 21.06.2016.

Noel-Storr, A.; Dooley, G.; Glanville, J.; Foxlee R. (2015a): The Embase project 1: crowd characteristics and behaviour after 18 months. Poster. In: The Cochrane Collaboration (Hg.): 23rd Cochrane Colloquium. Filtering the information overload for better decisions. 23rd Cochrane Colloquium: Filtering the information overload for better decisions. Wien, 03.-07.10.2015. The Cochrane Collaboration.

Noel-Storr, A.; Dooley, G.; Glanville, J.; Foxlee R. (2015b): The Embase project 2: crowdsourcing citation screening. In: The Cochrane Collaboration (Hg.): 23rd Cochrane Colloquium. Filtering the information overload for better decisions. 23rd Cochrane Colloquium: Filtering the information overload for better decisions. Wien, 03.-07.10.2015. The Cochrane Collaboration.

Noruzi, A. (2005): Google Scholar: The New Generation of Citation Indexes. In: *Libri* 55 (4). DOI: 10.1515/LIBR.2005.170.

Nowell, L. Terry; France, R. K.; Hix, D.; Heath, L. S.; Fox, E. A. (1996): Visualizing search results. In: H.-P. Frei, D. Harman, P. Schaübie und R. Wilkinson (Hg.): SIGIR '96 Proceedings of the 19th annual international ACM SIGIR conference on Research and development in information retrieval. The 19th annual international ACM SIGIR conference. Zurich, Switzerland, August 18 - 22, 1996. New York, NY, USA: ACM Press, S. 67-75.

O'Mara-Eves, A.; Thomas, J.; McNaught, J.; Miwa, M.; Ananiadou, S. (2015): Using text mining for study identification in systematic reviews: a systematic review of current approaches. In: *Systematic reviews* 4 (1), S. 1. DOI: 10.1186/2046-4053-4-5.

Okazaki, N. (2015): Acromine Demonstration. Hg. v. The National Centre for Text Mining NaCTeM. University of Manchester. Online verfügbar unter <http://www.nactem.ac.uk/software/termine/>, zuletzt aktualisiert am 16.07.2015, zuletzt geprüft am 21.06.2016.

Oldroyd, B. K.; Citroen, C. L. (1977): Study of strategies used in on-line searching. In: *Online Review* 1 (4), S. 295–310. DOI: 10.1108/eb023957.

Olorisade, B. K.; Quincey, E. de; Brereton, P.; Andras, P. (2016): A critical analysis of studies that address the use of text mining for citation screening in systematic reviews. In: S. Beecham (Hg.): EASE '16: Proceedings of the 20th International Conference on Evaluation and Assessment in Software Engineering. EASE '16 20th International Conference on Evaluation and Assessment in Software Engineering. Limerick, Ireland, June 01 - 03, 2016. New York, NY, USA: ACM.

O'Mara-Eves, A.; Brunton, G.; McDaid, D.; Kavanagh, J.; Oliver, S.; Thomas, J. (2013): Techniques for identifying cross-disciplinary and 'hard-to-detect' evidence for systematic review. In: *Res. Syn. Meth.* 5, S. 50–59. DOI: 10.1002/jrsm.1094.

Ovid Technologies (2016): Wolters Kluwer | Ovid - Home. Hg. v. Ovid Technologies. Wolters Kluwer. Online verfügbar unter <http://www.ovid.com/site/index.jsp>, zuletzt geprüft am 20.06.2016.

Oxford University Press (Hg.) (2016a): document, n. : Oxford English Dictionary. Online verfügbar unter <http://www.oed.com/view/Entry/56328>, zuletzt geprüft am 21.06.2016.

Oxford University Press (Hg.) (2016b): ground truth, n. : Oxford English Dictionary. Online verfügbar unter <http://www.oed.com/view/Entry/249130>, zuletzt geprüft am 21.06.2016.

Oxford University Press (Hg.) (2016c): information, n. : Oxford English Dictionary. Online verfügbar unter <http://www.oed.com/view/Entry/95568>, zuletzt geprüft am 21.06.2016.

Oxford University Press (Hg.) (2016d): interaction, n. : Oxford English Dictionary. Online verfügbar unter <http://www.oed.com/view/Entry/97519>, zuletzt geprüft am 21.06.2016.

Oxford University Press (Hg.) (2016e): Oxford English Dictionary. Online verfügbar unter <http://www.oed.com/>, zuletzt geprüft am 21.06.2016.

Oxford University Press (Hg.) (2016f): retrieval, n. : Oxford English Dictionary. Online verfügbar unter <http://www.oed.com/view/Entry/164463>, zuletzt geprüft am 21.06.2016.

Oxford University Press (Hg.) (2016g): strategy, n. : Oxford English Dictionary. Online verfügbar unter <http://www.oed.com/view/Entry/191319>, zuletzt geprüft am 21.06.2016.

Oxford University Press (Hg.) (2016h): synonym, n. : Oxford English Dictionary. Online verfügbar unter <http://www.oed.com/view/Entry/196522>, zuletzt geprüft am 21.06.2016.

Oxford University Press (Hg.) (2016i): task, n. : Oxford English Dictionary. Online verfügbar unter <http://www.oed.com/view/Entry/198017>, zuletzt geprüft am 21.06.2016.

Papaioannou, D.; Sutton, A.; Carroll, C.; Booth, A.; Wong, R. (2010): Literature searching for social science systematic reviews: consideration of a range of search techniques. In: *Health Info Libr J* 27 (2), S. 114–122. DOI: 10.1111/j.1471-1842.2009.00863.x.

Pawson, R.; Greenhalgh, T.; Harvey, G.; Walshe, K. (2004): Realist synthesis: an introduction. ESRC Research Methods Programme. London (ESRC Working Paper Series).

Pawson, R.; Greenhalgh, T.; Harvey, G.; Walshe, K. (2005): Realist review--a new method of systematic review designed for complex policy interventions. In: *Journal of health services research & policy* 10 Suppl 1, S. 21–34. DOI: 10.1258/1355819054308530.

Paynter, R.; Bañez, L. L.; Berliner, E.; Erinoff, E.; Lege-Matsuura, J.; Potter, S.; Uhl, S. (2016): EPC Methods: An Exploration of the Use of Text-Mining Software in Systematic Reviews. Rockville, MD (AHRQ Publications) (16-EHC023-EF). Online verfügbar unter <https://effective-healthcare.ahrq.gov/ehc/products/625/2214/text-mining-report-160419.pdf>, zuletzt geprüft am 09.06.2016.

Petcu, P.; Dragusin, R. (2013): Considerations for the Development of Task-Based Search Engines. In: M. Lupu, M. Salampanis, N. Fuhr, A. Hanbury, B. Larsen und H. Strindberg (Hg.): Proceedings of the Integrating IR Technologies for Professional Search Workshop, Moscow, Russia (March 24, 2013). IRPS 2013 Integrating IR Technologies for Professional Search. Moscow, Russia, 24.03.2013. Online verfügbar unter http://ceur-ws.org/Vol-968/irps_14.pdf, zuletzt geprüft am 28.04.2016.

Petticrew, M. (2011): When are complex interventions 'complex'? When are simple interventions 'simple'? In: *The European Journal of Public Health* 21 (4), S. 397–398. DOI: 10.1093/eurpub/ckr084.

Petticrew, M.; Anderson, L. M.; Elder, R.; Grimshaw, J.; Hopkins, D.; Hahn, R. et al. (2013a): Complex interventions and their implications for systematic reviews: a pragmatic approach. In: *J Clin Epidemiol* 66 (11), S. 1209–1214. DOI: 10.1016/j.jclinepi.2013.06.004.

Petticrew, M.; Rehfuess, E.; Noyes, J.; Higgins, J. P.T.; Mayhew, A.; Pantoja, T. et al. (2013b): Synthesizing evidence on complex interventions: how meta-analytical, qualitative, and mixed-method approaches can contribute. In: *J Clin Epidemiol* 66 (11), S. 1230–1243. DOI: 10.1016/j.jclinepi.2013.06.005.

Petticrew, M.; Roberts, H. (2006): *Systematic Reviews in the Social Sciences*. Oxford, UK: Blackwell Publishing Ltd.

Pharmaceutical Research and Manufacturers of America (Hg.) (2016): PhRMA | Pharmaceutical Research and Manufacturers of America. Online verfügbar unter <http://www.phrma.org/>, zuletzt aktualisiert am 21.06.2016, zuletzt geprüft am 21.06.2016.

Pohl, S.; Zobel, J.; Moffat, A. (2010): Extended Boolean retrieval for systematic biomedical reviews. In: Bernard Mans und Mark Reynolds (Hg.): *ACSC '10: Proceedings of the Thirty-Third Australasian Conferenc on Computer Science - Volume 102. Thirty-Third Australasian Computer Science Conference (ACSC 2010)*. Brisbane, Australia, January 2010. Darlinghurst, Australia, Australia: Australian Computer Society, Inc, S. 117–126.

ProQuest (Hg.) (2016): *ProQuest Dissertations & Theses Global - ProQuest Dissertations & Theses Global*. Online verfügbar unter <http://www.proquest.com/products-services/pqdtglobal.html>, zuletzt geprüft am 21.06.2016.

Ramer, S. L. (2005): Site-ation pearl growing: methods and librarianship history and theory. In: *J Med Libr Assoc* 93 (3), S. 397–400.

Ranard, B. L.; Ha, Y. P.; Meisel, Z. F.; Asch, D. A.; Hill, S. S.; Becker, L. B. et al. (2014): Crowdsourcing—Harnessing the Masses to Advance Health and Medicine, a Systematic Review. In: *J GEN INTERN MED* 29 (1), S. 187–203. DOI: 10.1007/s11606-013-2536-8.

Richardson, W. S.; Wilson, M. C.; Nishikawa, J.; Hayward, R. S. (1995): The well-built clinical question: a key to evidence-based decisions. In: *ACP J. Club* 123 (3), S. 3.

Robertson, S. E. (1977): The probability ranking principle in IR. In: *Journal of Documentation* 33 (4), S. 294–304. DOI: 10.1108/eb026647.

Rowley, J.; Slack, F. (2004): Conducting a literature review. In: *Management Research News* 27 (6), S. 31–39. DOI: 10.1108/01409170410784185.

Russell-Rose, T.; Lamantia, J.; Burrell, M. (2011): A Taxonomy of Enterprise Search and Discovery. In: *Proceedings of the HCIR 2011*. Online verfügbar unter <https://docs.google.com/a/kent.edu/viewer?a=v&pid=sites&srcid=ZGVmYXVsd-GRvbWFpbnoY2lyd29ya3Nob3B8Z3g6NzdmYjc3OWY2ZjQ2Zjg4MQ>, zuletzt geprüft am 29.06.2016.

Russell-Rose, T.; Tate, T. (2013): *Designing the search experience. The information architecture of discovery*. Amsterdam: Morgan Kaufmann.

Sackett, D. L.; Rosenberg, W. M. (1995): On the need for evidence-based medicine. In: *Journal of public health medicine* 17 (3), S. 330–334.

Sackett, D. L.; Rosenberg, W. M. C.; Gray, J. A. M.; Haynes, R. B.; Richardson, W. S. (1996): Evidence based medicine: what it is and what it isn't. In: *BMJ* 312 (7023), S. 71–72. DOI: 10.1136/bmj.312.7023.71.

Salton, G. (1988): On the relationship between theoretical retrieval models. Online verfügbar unter <https://uhdspace.uhasselt.be/dspace/bitstream/1942/845/1/salton263.pdf>.

Salton, G.; Fox, E. A.; Wu, H. (1983): Extended Boolean information retrieval. In: *Commun. ACM* 26 (11), S. 1022–1036. DOI: 10.1145/182.358466.

Sampson, M.; McGowan, J.; Cogo, E.; Grimshaw, J.; Moher, D.; Lefebvre, C. (2009): An evidence-based practice guideline for the peer review of electronic search strategies. In: *Journal of Clinical Epidemiology* 62 (9), S. 944–952. DOI: 10.1016/j.jclinepi.2008.10.012.

Sampson, M.; McGowan, J.; Lefebvre, C.; Moher, D.; Grimshaw, J. (2008): *PRESS: Peer Review of Electronic Search Strategies - Publication* | CADTH. Ottawa: Canadian Agency for Drugs and Technologies in Health. Online verfügbar unter <http://www.cadth.ca/publication/781>, zuletzt geprüft am 08.02.2015.

Sanderson, M. (2010): Test Collection Based Evaluation of Information Retrieval Systems. In: *FNT in Information Retrieval* 4 (4), S. 247–375. DOI: 10.1561/1500000009.

Saracevic, T. (1996): Relevance reconsidered '96. In: P. Ingwersen und N. O. Pors (Hg.): *CoLIS 2. Second International Conference on Conceptions of Library and Information Science: Integration in Perspective*. CoLIS 2. Second International Conference on Conceptions of Library and Information Science: Integration in Perspective. Kopenhagen, Dänemark, October

13-16, 1996. Copenhagen: The Royal School of Librarianship., S. 201–218, zuletzt geprüft am 17.02.2016.

Schell, C. L.; Rathe, R. J. (1992): Meta-analysis: a tool for medical and scientific discoveries. In: *Bulletin of the Medical Library Association* 80 (3), S. 219–222.

Schlager, J. Gabriel; Rosumeck, S.; Werner, R. N.; Jacobs, A.; Schmitt, J.; Schlager, C.; Nast, A. (2016): Topical treatments for scalp psoriasis. In: *Cochrane Database Syst Rev*. DOI: 10.1002/14651858.CD009687.pub2.

Schlosser, R. W.; O'Neil-Pirozzi, T. M. (2006): Problem Formulation in Evidence-based Practice and Systematic Reviews. In: *Contemporary Issues in Communication Science and Disorders (CICSD)* 33. Online verfügbar unter <http://www.asha.org/uploadedFiles/asha/publications/cicsd/2006SProblemFormulationinEBP.pdf>, zuletzt geprüft am 04.02.2015.

Schlosser, R. W.; Wendt, O.; Bhavnani, S.; Nail-Chiwetalu, B. (2006): Use of information-seeking strategies for developing systematic reviews and engaging in evidence-based practice: the application of traditional and comprehensive Pearl Growing. A review. In: *Int J Lang Commun Disord* 41 (5), S. 567–582. DOI: 10.1080/13682820600742190.

Schöpfel, J. (2010): Towards a Prague Definition of Grey Literature. Twelfth International Conference on Grey Literature: Transparency in Grey Literature. Grey Tech Approaches to High Tech Issues. Prague, 6-7 December 2010, Dec 2010, Czech Republic. In: D. J. Farace und J. Frantzen (Hg.): GL12 Conference Proceedings: Transparency in Grey Literature, Grey Tech Approaches to High Tech Issues. Twelfth International Conference on Grey Literature : Transparency in Grey Literature, Grey Tech Approaches to High Tech Issues. Prague, CZ, 06.-07.12.2010. International Conference on Grey Literature. Amsterdam (GL conference series, no. 12), S. 11–26. Online verfügbar unter http://archivesic.ccsd.cnrs.fr/sic_00581570.

Shemilt, I.; Simon, A.; Hollands, G. J.; Marteau, T. M.; Ogilvie, D.; O'Mara-Eves, A. et al. (2013): Pinpointing needles in giant haystacks: use of text mining to reduce impractical screening workload in extremely large scoping reviews. In: *Res. Syn. Meth.*, S. n/a–n/a. DOI: 10.1002/jrsm.1093.

Shneiderman, B.; Byrd, G.; Croft, W. B. (1997): Clarifying Search: A User-Interface Framework for Text Searches. In: *D-Lib Magazine*. Online verfügbar unter

- <http://www.dlib.org/dlib/january97/retrieval/01shneiderman.html>, zuletzt geprüft am 28.06.2016.
- Sims, M. H.; Bigham, J.; Kautz, H.; Halterman, M. W. (2014): Crowdsourcing medical expertise in near real time. In: *Journal of hospital medicine* 9 (7), S. 451–456. DOI: 10.1002/jhm.2204.
- Skelton, B. (1973): Scientists and Social Scientists as Information Users: a comparison of results of science user studies with the investigation into information requirements of the social sciences. In: *Journal of Librarianship and Information Science* 5 (2), S. 138–156. DOI: 10.1177/096100067300500205.
- Slater, L. (2014): PubMed PubReMiner. In: *Journal of the Canadian Health Libraries Association/Journal de l'Association des bibliothèques de la santé du Canada* 33 (2), S. 106–107.
- Smiraglia, R. P. (2008): Rethinking What We Catalog. Documents as Cultural Artifacts. In: *Cataloging & Classification Quarterly* 45 (3), S. 25–37. DOI: 10.1300/J104v45n03_04.
- Social Science Research Unit, UCL Institute of Education (Hg.) (2016): EPPI-Centre Home. University College London. Online verfügbar unter <http://eppi.ioe.ac.uk/cms/>, zuletzt geprüft am 21.06.2016.
- Song, F.; Eastwood, A.; Gilbody, S.; Duley, L.; Sutton, A. (2000): Publication and related biases: a review. In: *Health Technology Assessment* 4 (10), S. 115. DOI: 10.3310/hta4100.
- Spink, A.; Wolfram, D.; Jansen, M. B. J.; Saracevic, T. (2001): Searching the web: The public and their queries. In: *Journal of the American Society for Information Science and Technology* 52 (3), S. 226–234. DOI: 10.1002/1097-4571(2000)9999:9999<:AID-ASI1591>3.0.CO;2-R.
- Spoor, P.; Airey, M.; Bennett, C.; Greensill, J.; Williams, R. (1996): Use of the capture-recapture technique to evaluate the completeness of systematic literature searches. In: *BMJ* 313 (7053), S. 342–343. DOI: 10.1136/bmj.313.7053.342.
- Squires, J.; Valentine, J. C.; Grimshaw, J. M. (2013): Systematic reviews of complex interventions: framing the review question. In: *J Clin Epidemiol* 66 (11), S. 1215–1222. DOI: 10.1016/j.jclinepi.2013.05.013.

Steinke, I. (2007): Qualitätssicherung in der qualitativen Forschung. In: U. Kuckartz, H. Grunenberg und T. Dresing (Hg.): Qualitative Datenanalyse: computergestützt. Methodische Hintergründe und Beispiele aus der Forschungspraxis. 2., überarb. und erw. Aufl. Wiesbaden: VS, Verl. für Sozialwissenschaften (SpringerLink : Bücher), S. 176–187.

Stiftung für Qualität und Wirtschaftlichkeit im Gesundheitswesen (Hg.) (2016a): Aufgaben und Ziele. Online verfügbar unter <https://www.iqwig.de/de/ueber-uns/aufgaben-und-ziele.2946.html>, zuletzt geprüft am 21.06.2016.

Stiftung für Qualität und Wirtschaftlichkeit im Gesundheitswesen (Hg.) (2016b): iqwig.de - Startseite. Online verfügbar unter <https://www.iqwig.de/>, zuletzt geprüft am 20.06.2016.

Stock, W. G. (2007): Information Retrieval. Informationen suchen und finden. München: Oldenbourg (Einführung in die Informationswissenschaft, 1). Online verfügbar unter [http://www.regensburger-katalog.de/InfoGuideClient.ubrsis/start.do?Login=i-gubr&Query=540="978-3-486-58172-0"](http://www.regensburger-katalog.de/InfoGuideClient.ubrsis/start.do?Login=i-gubr&Query=540=).

Stock, W. G.; Stock, M.; Becker, P. (2013): Handbook of information science. Online verfügbar unter <http://search.ebscohost.com/login.aspx?direct=true&scope=site&db=e000xat&AN=641726>.

Stowasser, J. M.; Petschenig, M.; Skutsch, F.; Pichl, R.; Christ, A.; Lošek, F. (Hg.) (2007): Stowasser. Lateinisch-deutsches Schulwörterbuch. [Nachdr.]. Düsseldorf, Wien: Oldenbourg; öbvhpt.

Stubblebine, T. (2008): Reguläre Ausdrücke. Kurz & gut ; [für Perl, Ruby, PHP, C, Python, Java & .Net]. 2. Aufl. Beijing u.a.: O'Reilly (O'Reillys Taschenbibliothek).

Swinkels, A.; Briddon, J.; Hall, J. (2006): Two physiotherapists, one librarian and a systematic literature review: collaboration in action. In: *Health Info Libr J* 23 (4), S. 248–256. DOI: 10.1111/j.1471-1842.2006.00689.x.

Tait, J. (2013): Issues and Non-Issues in Professional Search. In: M. Lupu, M. Salamasis, N. Fuhr, A. Hanbury, B. Larsen und H. Strindberg (Hg.): Proceedings of the Integrating IR Technologies for Professional Search Workshop, Moscow, Russia (March 24, 2013). IRPS 2013 Integrating IR Technologies for Professional Search. Moscow, Russia, 24.03.2013. Online verfügbar unter http://ceur-ws.org/Vol-968/irps_3.pdf, zuletzt geprüft am 28.04.2016.

- Tait, J. I. (2014): An Introduction to Professional Search. In: G. Paltoglou, F. Loizides und P. Hansen (Hg.): Professional Search in the Modern World, Bd. 8830. Cham: Springer International Publishing (Lecture Notes in Computer Science), S. 1–5.
- Talja, S.; Keso, H.; Pietiläinen, T. (1999): The production of 'context' in information seeking research: a metatheoretical view. In: *Information Processing & Management* 35 (6), S. 751–763. DOI: 10.1016/S0306-4573(99)00024-2.
- Taylor, R. S. (1962): The process of asking questions. In: *American documentation* 13 (4), S. 391–396.
- Taylor, R. S. (1968): Question-negotiation and information seeking in libraries. In: *College & research libraries* 29 (3), S. 178–194.
- The Campbell Collaboration (Hg.) (2016): About Us : The Campbell Collaboration. Online verfügbar unter http://www.campbellcollaboration.org/about_us/index.php, zuletzt geprüft am 21.06.2016.
- The Cochrane Collaboration (Hg.) (2013): The story of the Cochrane logo | ARCHIVE. Online verfügbar unter <http://community-archive.cochrane.org/features/story-cochrane-logo>, zuletzt aktualisiert am 06.01.2013, zuletzt geprüft am 21.06.2016.
- The Cochrane Collaboration (Hg.) (2014): Home - The Cochrane Library. Online verfügbar unter <http://www.thecochranelibrary.com/view/0/index.html>, zuletzt aktualisiert am 04.08.2014, zuletzt geprüft am 04.08.2014.
- The Cochrane Collaboration (Hg.) (2015a): Cochrane Deutschland. Online verfügbar unter <http://www.cochrane.de/>, zuletzt aktualisiert am 15.09.2015, zuletzt geprüft am 21.06.2016.
- The Cochrane Collaboration (Hg.) (2015b): Welcome | Cochrane Methods Information Retrieval. Online verfügbar unter <http://methods.cochrane.org/irmg/>, zuletzt geprüft am 21.06.2016.
- The Cochrane Collaboration (Hg.) (2016a): Cochrane crowd. Online verfügbar unter <http://crowd.cochrane.org/index.html>, zuletzt aktualisiert am 13.06.2016, zuletzt geprüft am 23.06.2016.

The Cochrane Collaboration (Hg.) (2016b): Review Groups | Cochrane. Online verfügbar unter <http://www.cochrane.org/contact/review-groups>, zuletzt aktualisiert am 20.06.2016, zuletzt geprüft am 21.06.2016.

The Cochrane Public Health Group (Hg.) (2011): Guide for developing a Cochrane protocol. Online verfügbar unter http://ph.cochrane.org/sites/ph.cochrane.org/files/uploads/Guide%20for%20PH%20protocol_Nov%202011_final%20for%20website.pdf, zuletzt aktualisiert am 24.11.2011, zuletzt geprüft am 21.06.2016.

The InterTASC Information Specialists' Sub-Group (Hg.) (2016): ISSG search filter checklist proforma - ISSG Search Filters Resource. Online verfügbar unter <https://sites.google.com/a/york.ac.uk/issg-search-filters-resource/issg-search-filter-checklist-proforma>, zuletzt aktualisiert am 20.06.2016, zuletzt geprüft am 21.06.2016.

The PubMed Health Team (Hg.) (2013): Refine your search results with PubMed Health's new faceted search - Blog - National Library of Medicine - PubMed Health. Online verfügbar unter <http://www.ncbi.nlm.nih.gov/pubmedhealth/blog/2013/07/refine-search-results-PubMedHealth-faceted-search/>, zuletzt aktualisiert am 11.07.2013, zuletzt geprüft am 10.04.2016.

Thomas, J.; McNaught, J.; Ananiadou, S. (2011): Applications of text mining within systematic reviews. In: *Res. Syn. Meth.* 2 (1), S. 1–14. DOI: 10.1002/jrsm.27.

Thomson Reuters (Hg.) (2016a): BIOSIS Previews. Online verfügbar unter <http://thomsonreuters.com/en/products-services/scholarly-scientific-research/scholarly-search-and-discovery/biosis-previews.html>, zuletzt aktualisiert am 20.06.2016, zuletzt geprüft am 21.06.2016.

Thomson Reuters (Hg.) (2016b): Social Sciences Citation Index. Online verfügbar unter <http://thomsonreuters.com/en/products-services/scholarly-scientific-research/scholarly-search-and-discovery/social-sciences-citation-index.html>, zuletzt aktualisiert am 20.06.2016, zuletzt geprüft am 21.06.2016.

Thomson Reuters (Hg.) (2016c): Web of Science [v.5.22] - All Databases Home. Online verfügbar unter http://apps.webofknowledge.com/UA_GeneralSearch_input.do?product=UA&search_mode=GeneralSearch&SID=Y2N449s9ZQXQcHKgrjU&preferencesSaved=, zuletzt geprüft am 21.06.2016.

TREC Conference (2002): TREC Tracks. Hg. v. NIST National Institute of Standards and Technology. Online verfügbar unter <http://trec.nist.gov/tracks.html>, zuletzt aktualisiert am 10.04.2014, zuletzt geprüft am 05.05.2016.

Tröhler, U. (1988): To Improve the Evidence of Medicine": Arithmetic Observation in Clinical Medicine in the Eighteenth and Early Nineteenth Centuries. In: *History and Philosophy of the Life Sciences* 10, S. 31–40.

Tsafnat, G.; Dunn, A.; Glasziou, P.; Coiera, E. (2013): The automation of systematic reviews. In: *BMJ* 346 (jan10 1), S. f139. DOI: 10.1136/bmj.f139.

Tsafnat, G.; Glasziou, P.; Choong, M. K.; Dunn, A.; Galgani, F.; Coiera, E. (2014): Systematic review automation technologies. In: *Syst Rev* 3, S. 74. DOI: 10.1186/2046-4053-3-74.

Tse, T.; Johnson, S.; Ripple, A. (2002): Update: Expanded ClinicalTrials.gov Search Capabilities. NLM Technical Bulletin. Jul-Aug 2002. Hg. v. U.S. National Library of Medicine. Lister Hill National Center for Biomedical Communications (NLM Tech Bull., 2002 Jul-Aug;(327):e5). Online verfügbar unter http://www.nlm.nih.gov/pubs/tech-bull/ja02/ja02_ctgov.html, zuletzt aktualisiert am 18.04.2013, zuletzt geprüft am 12.11.2014.

Tseng, Y.-H.; Lin, C.-H.; Lin, Y.-I. (2007): Text mining techniques for patent analysis. In: *Patent Processing* 43 (5), S. 1216–1247. DOI: 10.1016/j.ipm.2006.11.011.

Tseng, Y.-H.; Wu, Y.-J. (2008): A study of search tactics for patentability search. a case study on patent engineers. In: John Tait (Hg.): Proceedings of the 1st ACM workshop on Patent information retrieval. CIKM '08 Conference on Information and Knowledge Management. Napa Valley, California, USA, October 26 - 30, 2008. New York, NY: ACM, S. 33–36. Online verfügbar unter <http://dl.acm.org/citation.cfm?id=1458581>.

U.S. National Institutes of Health (NIH) (2016): ClinicalTrials.gov. Hg. v. U.S. National Institutes of Health (NIH). Online verfügbar unter <https://clinicaltrials.gov/>, zuletzt aktualisiert am n.a., zuletzt geprüft am 24.02.2016.

U.S. National Library of Medicine (2011): NLM 175th Anniversary | Our Milestones. Hg. v. U.S. National Library of Medicine. Online verfügbar unter <http://apps.nlm.nih.gov/175/milestones.cfm>, zuletzt aktualisiert am 16.10.2014, zuletzt geprüft am 16.10.2014.

U.S. National Library of Medicine (2013): Fact Sheet Medical Subject Headings (MeSH®). Online verfügbar unter <http://www.nlm.nih.gov/pubs/factsheets/mesh.html>, zuletzt aktualisiert am 09.12.2013, zuletzt geprüft am 16.10.2014.

U.S. National Library of Medicine (2014a): OLDMEDLINE Data. Online verfügbar unter http://www.nlm.nih.gov/databases/databases_oldmedline.html, zuletzt aktualisiert am 03.03.2014, zuletzt geprüft am 31.03.2015.

U.S. National Library of Medicine (2014b): MEDLINE®/PubMed® Data Element (Field) Descriptions. Online verfügbar unter <http://www.nlm.nih.gov/bsd/mms/medlineelements.html>, zuletzt aktualisiert am 25.08.2014, zuletzt geprüft am 08.11.2014.

U.S. National Library of Medicine (2015a): Automatic Term Mapping. Pubmed Tutorial. Online verfügbar unter https://www.nlm.nih.gov/bsd/disted/pubmedtutorial/020_040.html, zuletzt aktualisiert am 27.05.2015, zuletzt geprüft am 21.06.2016.

U.S. National Library of Medicine (2015b): MeSH Browser - 2016. Online verfügbar unter <https://www.nlm.nih.gov/mesh/MBrowser.html>, zuletzt aktualisiert am 09.11.2015, zuletzt geprüft am 21.06.2016.

U.S. National Library of Medicine (2015c): Truncation. Pubmed Tutorial. Online verfügbar unter https://www.nlm.nih.gov/bsd/disted/pubmedtutorial/020_460.html, zuletzt aktualisiert am 25.06.2015, zuletzt geprüft am 21.06.2016.

U.S. National Library of Medicine (2015d): Fact Sheet MEDLINE®. Online verfügbar unter <http://www.nlm.nih.gov/pubs/factsheets/medline.html>, zuletzt aktualisiert am 12.02.2015, zuletzt geprüft am 02.02.2016.

U.S. National Library of Medicine (2016a): Fact Sheet PubMed®: MEDLINE® Retrieval on the World Wide Web. Online verfügbar unter <https://www.nlm.nih.gov/pubs/factsheets/pubmed.html>, zuletzt aktualisiert am 06.05.2016, zuletzt geprüft am 21.06.2016.

U.S. National Library of Medicine (Hg.) (2016b): Metathesaurus. Online verfügbar unter https://www.nlm.nih.gov/research/umls/knowledge_sources/metathesaurus/, zuletzt aktualisiert am 12.04.2016, zuletzt geprüft am 21.06.2016.

- U.S. National Library of Medicine (2016c): National Library of Medicine - National Institutes of Health. Online verfügbar unter <https://www.nlm.nih.gov/>, zuletzt aktualisiert am 09.06.2016, zuletzt geprüft am 21.06.2016.
- United Nations Development Programme (2012): Patent Methodology. Online verfügbar unter <http://www.undp.org/content/dam/undp/library/hivaid/English/Patent%20Information%20and%20Transparency.pdf>, zuletzt geprüft am 29.04.2016.
- Vakkari, P. (2001): A theory of the task-based information retrieval process: a summary and generalisation of a longitudinal study. In: *Journal of Documentation* 57 (1), S. 44–60. DOI: 10.1108/EUM0000000007075.
- Vakkari, P. (2003): Task-based information searching. In: *Ann. Rev. Info. Sci. Tech.* 37 (1), S. 413–464. DOI: 10.1002/aris.1440370110.
- Vassilakaki, E.; Garoufallou, E.; Johnson, F.; Hartley, R. J. (2014): Users' Information Search Behavior in a Professional Search Environment. In: G. Paltoglou, F. Loizides und P. Hansen (Hg.): *Professional Search in the Modern World*, Bd. 8830. Cham: Springer International Publishing (Lecture Notes in Computer Science), S. 23–44.
- Vassilakaki, E.; Johnson, F. (2015): The use of grounded theory in identifying the user experience during search. In: *Library & Information Science Research* 37 (1), S. 77–87. DOI: 10.1016/j.lisr.2014.06.006.
- Verband Forschender Arzneimittelhersteller e.V. (Hg.) (2016): Register nicht-interventioneller Studien. Online verfügbar unter <http://www.vfa.de/de/anzneimittel-forschung/datenbanken-zu-anzneimitteln/nisdb>, zuletzt geprüft am 21.06.2016.
- Versi, E. (1992): "Gold standard" is an appropriate term. In: *BMJ* 305 (6846), S. 187.
- Wacholder, N. (2011): Interactive query formulation. In: *Ann. Rev. Info. Sci. Tech.* 45 (1), S. 157–196. DOI: 10.1002/aris.2011.1440450111.
- Wallace, B. C.; Trikalinos, T. A.; Lau, J.; Brodley, C. E.; Schmid, C. H. (2010): Semi-automated screening of biomedical citations for systematic reviews. In: *BMC Bioinformatics* 11, S. 55. DOI: 10.1186/1471-2105-11-55.

Walters, W. H. (2008): Google Scholar Search Performance: Comparative Recall and Precision. In: *portal: Libraries and the Academy* 9 (1), S. 5–24. DOI: 10.1353/pla.0.0034.

Wang, Y.; Howard, P. (2012): Google Scholar Usage: An Academic Library's Experience. In: *Journal of Web Librarianship* 6 (2), S. 94–108. DOI: 10.1080/19322909.2012.672067.

Webster, J.; Watson, R. T. (2002): Analyzing the Past to Prepare for the Future: Writing a Literature Review. In: *MIS Quarterly* 26 (1), S. 13–26. Online verfügbar unter <http://www.jstor.org/stable/4132319>, zuletzt geprüft am 09.03.2015.

Wersig, G. (1973): Informationssoziologie. Hinweise zu einem informationswissenschaftlichen Teilbereich. Frankfurt /M.: Athenäum-Verl.

WHO Regional Office for Africa (2016): African Index Medicus - AIM. Online verfügbar unter <http://indexmedicus.afro.who.int/>, zuletzt aktualisiert am 20.06.2016, zuletzt geprüft am 21.06.2016.

Wilczynski, N. L.; Morgan, D.; Haynes, R. Brian (2005): An overview of the design and methods for retrieving high-quality studies for clinical care. In: *BMC medical informatics and decision making* 5 (1), S. 1. DOI: 10.1186/1472-6947-5-20.

Williams, R. J.; Tse, T. (2007): New Look for ClinicalTrials.gov. Hg. v. U.S. National Library of Medicine. Lister Hill National Center for Biomedical Communications (NLM Tech Bull., 2007 Sep-Oct; (358):e2). Online verfügbar unter http://www.nlm.nih.gov/pubs/tech-bull/so07/so07_clinical_trials.html, zuletzt aktualisiert am 16.11.2013, zuletzt geprüft am 12.11.2014.

Wilpert, G. von (1989): Sachwörterbuch der Literatur. 7., verb. und erw. Aufl. Stuttgart: A. Kröner (Kröners-Taschenausgabe, Bd. 231).

Wilson, P. (1973): Situational relevance. In: *Information Storage and Retrieval* 9 (8), S. 457–471. DOI: 10.1016/0020-0271(73)90096-X.

Wilson, T. D. (1981): On user studies and information needs. In: *Journal of Documentation* 37 (1), S. 3–15. DOI: 10.1108/eb026702.

Wilson, T. D. (1997): Information behaviour: An interdisciplinary perspective. In: *Information Processing & Management* 33 (4), S. 551–572. DOI: 10.1016/S0306-4573(97)00028-9.

Wilson, T. D. (1999): Models in information behaviour research. In: *Journal of Documentation* 55 (3), S. 249–270. DOI: 10.1108/EUM0000000007145.

Wilson, T. D. (2000): Human Information Behaviour. In: *Informing Science* 3 (2).

Windeler, J.; Antes, G.; Behrens, J.; Donner-Banzhoff, N.; Lelgemann, M. (2008): Randomisierte kontrollierte Studien: Kritische Evaluation ist ein Wesensmerkmal ärztlichen Handelns. In: *Dtsch Arztebl International* 105 (11), S. 565. Online verfügbar unter <http://www.aerzteblatt.de/int/article.asp?id=59361>.

Wolfe, N.; Gøtzsche, P. C.; Bero, L. (2013): Strategies for obtaining unpublished drug trial data: a qualitative interview study. In: *Systematic reviews* 2, S. 31. DOI: 10.1186/2046-4053-2-31.

Womser-Hacker, C. (2013): Evaluierung im Information Retrieval. In: Rainer Kuhlen, Wolfgang Semar und Dietmar Strauch (Hg.): *Grundlagen der praktischen Information und Dokumentation. Handbuch zur Einführung in die Informationswissenschaft und -praxis*. 6., völlig neu gefasste Ausg. Berlin: De Gruyter, S. 396–410.

World Health Organization (Hg.) (2016): International Clinical Trials Registry Platform (ICTRP). Online verfügbar unter <http://www.who.int/ictrp/en/>, zuletzt geprüft am 21.06.2016.

Wright, A. (2016): Reimagining search. In: *Commun. ACM* 59 (6), S. 17–19. DOI: 10.1145/2911971.

Xie, H. (2000): Shifts of interactive intentions and information-seeking strategies in interactive information retrieval. In: *J. Am. Soc. Inf. Sci.* 51 (9), S. 841–857. DOI: 10.1002/(SICI)1097-4571(2000)51:9<841::AID-ASI70>3.0.CO;2-0.

11 Anhang

Anhang 1 Erfassungsformular 1 Beobachtungsstudie

Modus	<input type="checkbox"/> Kontrolle von Suchanfragen	<input type="checkbox"/> Struktur des Suchprofils	<input type="checkbox"/> Ideenfindung
	<input type="checkbox"/> Informationsstruktur und Suchanfrage	<input type="checkbox"/> Einsatz von Suchtermen	
Kontext	<input type="checkbox"/> Suchmaschine	<input type="checkbox"/> Studienregister	<input type="checkbox"/> Metainfo
	<input type="checkbox"/> DB oder Host	<input type="checkbox"/> Graue Literatur	<input type="checkbox"/> Suchphrase
	<input type="checkbox"/> Thesaurus/Lex.	<input type="checkbox"/> Ergebnismenge	<input type="checkbox"/> Termpool
	<input type="checkbox"/> Sek. Literatur	<input type="checkbox"/> Dokument	
Move	<input type="checkbox"/> Check	<input type="checkbox"/> Pattern	<input type="checkbox"/> Record
	<input type="checkbox"/> Weigh	<input type="checkbox"/> Correct	
	<input type="checkbox"/> Bibble	<input type="checkbox"/> Cut	<input type="checkbox"/> Cleave
	<input type="checkbox"/> Select	<input type="checkbox"/> Stretch	
	<input type="checkbox"/> Survey	<input type="checkbox"/> Scaffold	
	<input type="checkbox"/> Specify	<input type="checkbox"/> Reduce	<input type="checkbox"/> Pinpoint
	<input type="checkbox"/> Exhaust	<input type="checkbox"/> Parallel	<input type="checkbox"/> Block
	<input type="checkbox"/> Super	<input type="checkbox"/> Trace	<input type="checkbox"/> Contrary
	<input type="checkbox"/> Sub	<input type="checkbox"/> Vary	<input type="checkbox"/> Respell
	<input type="checkbox"/> Relate	<input type="checkbox"/> Fix	<input type="checkbox"/> Respace
	<input type="checkbox"/> Neighbor	<input type="checkbox"/> Rearrange	
DB Feature	<input type="checkbox"/> Einfache Suche	<input type="checkbox"/> Angewandte Suche	
	<input type="checkbox"/> Facettensuche	<input type="checkbox"/> bestehende Suchfilter	
Ertrag/Ergebnis	<input type="checkbox"/> Datenbank	<input type="checkbox"/> Studienregister	<input type="checkbox"/> Dokument
	<input type="checkbox"/> Thesaurus/Lex	<input type="checkbox"/> Graue Literatur	<input type="checkbox"/> Suchphrase
	<input type="checkbox"/> Sek. Literatur	<input type="checkbox"/> Ergebnismenge	<input type="checkbox"/> Term(pool)
Bewertung	<input type="checkbox"/> positiv	<input type="checkbox"/> negativ	<input type="checkbox"/> unklar
Dauer/Uhrzeit			
Bemerkung			

Anhang 2 Erfassungsf formular 2 Beobachtungsstudie

Modus	<input type="checkbox"/> Ideenfindung		
Kontext	<input type="checkbox"/> Suchmaschine	<input type="checkbox"/> Studienregister	<input type="checkbox"/> Metainfo
	<input type="checkbox"/> DB oder Host	<input type="checkbox"/> Graue Literatur	<input type="checkbox"/> Suchphrase
	<input type="checkbox"/> Thesaurus/Lex.	<input type="checkbox"/> Ergebnismenge	<input type="checkbox"/> Termpool
	<input type="checkbox"/> Sek. Literatur	<input type="checkbox"/> Dokument	
Move	<input type="checkbox"/> Think	<input type="checkbox"/> Meditate	<input type="checkbox"/> Rescue
	<input type="checkbox"/> Brainstorm	<input type="checkbox"/> Consult	<input type="checkbox"/> Wander
	<input type="checkbox"/> Change	<input type="checkbox"/> Stop	<input type="checkbox"/> Jolt
	<input type="checkbox"/> Catch	<input type="checkbox"/> Breach	<input type="checkbox"/> Focus
	<input type="checkbox"/> Break	<input type="checkbox"/> Reframe	<input type="checkbox"/> Dilate
	<input type="checkbox"/> Skip	<input type="checkbox"/> Notice	
Ertrag/Ergebnis	<input type="checkbox"/> Datenbank	<input type="checkbox"/> Studienregister	<input type="checkbox"/> Dokument
	<input type="checkbox"/> Thesaurus/Lex	<input type="checkbox"/> Graue Literatur	<input type="checkbox"/> Suchphrase
	<input type="checkbox"/> Sek. Literatur	<input type="checkbox"/> Ergebnismenge	<input type="checkbox"/> Term(pool)
Bewertung	<input type="checkbox"/> positiv	<input type="checkbox"/> negativ	<input type="checkbox"/> unklar
Dauer/Uhrzeit			
Bemerkung			

Anhang 3 Fragen und Codes des Freitext Fragebogens der Hauptstudie

Sprache	Frage	Code
DE	Nach welchen methodologischen Standards führen Sie systematische Übersichtsarbeiten üblicherweise durch?	STANDARD
EN	According to what methodological standards do you usually conduct systematic reviews?	
DE	Können Sie Ihre typische allgemeine Vorgehensweise beim Aufbau von Suchstrategien in einzelnen kurzen Schritten stichpunktartig beschreiben?	GENERAL
EN	Can you describe your typical approach building your search strategies in some simple steps?	
DE	Haben Sie eine typische Reihenfolge? (z.B. Deskriptoren zuerst etc.)	ORDER
EN	Do you prefer a specific order when processing relevant citations or documents?	
DE	Wo sehen Sie die Stärken der folgenden Suchtaktiken? (Auffinden von Dokumenten, von Freitextermen, Deskriptoren, Kontrolle der Suchergebnisse etc.)	TACTICS
DE	Citation Pearl Growing	TACTICS_CPG
DE	Related Article Funktion	TACTICS_RAF
DE	Journal Handsuche	TACTICS_JH
DE	Bibliographische Suche	TACTICS_BS
DE	Zitationssuche	TACTICS_CS
DE	Suche nach Autoren/Websites/Forschergruppen	TACTICS_AU
DE	Themensuche in Grauer Literatur	TACTICS_GL
DE	Welche Tools und welche Entscheidungen helfen Ihnen, zu breite oder zu eng gefasste Deskriptoren in die Suchstrategie aufzunehmen?	SELECTION_DESC
DE	Welche Kriterien sind Ihnen bei der Aufnahme eines Terms aus einem relevanten Dokument in Ihre Suchstrategie besonders wichtig?	SELECTION_TERM
EN	Which are the crucial factors and conditions to add new synonyms into your search query?	
EN	What makes you feel confident about synonyms?	TERM_CONFID

DE	Suchen Sie für relevante Volltextdokumente <u>immer</u> bibliographische Einträge in den einschlägigen Datenbanken, beispielsweise um zugehörige Deskriptoren zu identifizieren?	IDENT_BIB
EN	What might be your reasons to identify corresponding bibliographic entries for relevant full text documents you have at hand?	
DE	Suchen Sie für interessante bibliographische Einträge sicherheitshalber auch immer den Volltext, beispielsweise um weitere Synonyme zu identifizieren?	IDENT_FULL
EN	What might be your reasons to access corresponding full-text documents for a relevant or interesting citation?	
DE	Wie erfolgt <u>typischerweise</u> der Umgang mit einem eher „unspezifischen“ Synonym, das in einem relevanten Volltext gefunden wird?	UNSPEC_SYN
EN	How do you typically handle with a rather "non-specific" synonym term?	
DE	Welche Entscheidungen sind hierfür ausschlaggebend?	UNSPEC_SYN_CRITERIA
EN	Which factors are particularly crucial when adding or dropping such "non-specific" terms to your search query?	
DE	Unter welchen Umständen ist es Ihrer Meinung nach <u>dringend</u> erforderlich, Testsuchen mit einzelnen Termen oder Phrasen durchzuführen?	TEST_REASONS
EN	Under what circumstances do you think is urgently necessary to conduct test searches with terms or phrases?	
DE	Führen Sie bei der Verwendung von ADJ Operatoren <u>grundsätzlich</u> Vergleichssuchen durch, um den Wortabstand hinsichtlich Precision/Recall zu optimieren?	ADJ_TEST
EN	Under which circumstances do you conduct test searches before using phrases with ADJ operators to optimize precision and recall?	
DE	Welche Maßnahmen zur Erhöhung der <u>Precision</u> halten Sie im Rahmen systematischer Übersichtsarbeiten für geeignet/ungeeignet?	PRECISION
EN	Which actions/steps are suitable to increase precision from your point of view?	

DE	Welche Maßnahmen zur Balance von Precision und Recall halten Sie bei systematischen Übersichtsarbeiten für geeignet?	BALANCE
EN	Which actions/steps are suitable to get the <u>optimal balance</u> between precision and recall?	
DE	Führen Sie <u>immer</u> Testsuchen durch, bevor Sie Terme oder Phrasen erweitern (z.B. ADJ oder Trunkierung) oder reduzieren (z.B. Lokalisation in bestimmten Feldern TI., AB...)?	ADJ_MOD
EN	Under which circumstances do you conduct test searches before expanding (ADJ or truncation) or reducing (LOC Operators TI-AB...) terms or phrases in a search query?	
DE	Wie häufig nutzen Sie methodologische Suchfilter und wenn ja, welche?	FILTER
EN	How often do you use methodological search filters? Which ones?	
DE	Nutzen Sie neben methodologischen Suchfiltern auch andere gespeicherte Snippets?	FILTER_SPEC
DE	Was sind Ihrer Meinung nach die problematischsten Aspekte beim Auffinden von Deskriptoren und Freitexttermen und deren „passender“ Eingliederung ins Suchprofil?	PROBLEMS
EN	What are the most problematic or resource-intensive aspects during search strategy development?	
EN	Do you have any suggestions or approaches to solve these problems?	PROB_SOLVING
EN	How much is your personal searching style affected by external constraints?	EXT_CONSTRAINTS
EN	How did you acquire your profession and personal searching style?	PERSONAL_STYLE
DE	Wie viele Jahre Berufserfahrung haben Sie in der Durchführung systematischer Übersichtsarbeiten?	EXPERIENCE_YEARS
EN	How many years of professional experience in conducting systematic reviews do you have?	
DE	Welche einschlägige, professionelle Berufsausbildung haben Sie?	PROFESSION
EN	Which professional education do you have?	
DE	Haben Sie selbst ein medizinisches Studium absolviert?	MED

DE	Recherchieren Sie vorwiegend in einem bestimmten medizinischen Fachgebiet? Wenn ja, welches?	SPECIALTY
EN	Do you mainly conduct systematic reviews in a specific medical specialty? If yes, which one?	
EN	Are you a member of a Cochrane Review Group? If yes, which one?	REVGROUP
EN	How many systematic reviews do you conduct annually?	SR_YEAR
EN	How many different clinical trial registers do you use regularly?	CTR_REG
EN	How many different bibliographic databases do you use regularly?	BIB_REG