



Bibliometrische Feldabgrenzung in Web of Science und Scopus: Ein Methodenvergleich anhand des interdisziplinären Feldes Bildungsforschung

Valeria Aman

iFQ – Institut für Forschungsinformation und Qualitätssicherung

Abstract

Den meisten bibliometrischen Analysen geht eine Abgrenzung des zu untersuchenden Feldes voran. Dabei handelt es sich um eine komplexe Aufgabe für die keine generelle Lösung existiert. Daher wird Feldabgrenzung („field delineation“) als ein eigenständiges Forschungsfeld der Bibliometrie erachtet. Zumeist geschieht die Feldabgrenzung anhand des in der zugrundeliegenden Datenbank implementierten Klassifikationssystems. So sind die in *Web of Science* indexierten Zeitschriften Themenkategorien zugewiesen. Die als „Makro-Level“ bezeichnete Felddefinition geschieht auf Basis der Zeitschriften, die unter der zutreffenden Kategorie klassifiziert sind. Diese Form der Abgrenzung kommt bei interdisziplinären Feldern wie der Bildungsforschung an ihre Grenzen. Der Aufsatz zeigt zunächst die Probleme der Makro-Level-Abgrenzung in *WoS* und *Scopus* anhand der Bildungsforschung auf. Als Alternative wird die von Zitt & Bassecouard (2006) vorgeschlagene „lex + cite“-Methode angewandt, die sich sowohl lexikalischer als auch zitationsbasierter Merkmale von Publikationen bedient. Die Ergebnisse zeigen, dass die hybride „lex + cite“-Methode in *WoS* als auch in *Scopus* zu vergleichbaren abgegrenzten Korpora führt, die in der Sprache des *Information Retrievals* ein zufriedenstellendes Verhältnis aus *Recall* und *Precision* aufweisen.

1. Einleitung

Bei der bibliometrischen Analyse eines Forschungsfeldes besteht der erste Schritt in der adäquaten Abgrenzung der Untersuchungsmenge. Es soll sich nicht nur um das richtige Feld handeln, sondern auch eine genügende Anzahl an Publikationen bereitstehen. Da es sich hierbei um eine komplexe Aufgabe handelt, für die es keine generelle Lösung gibt, wird Feldabgrenzung (engl. „field delineation“) als ein eigenständiges Forschungsfeld der Bibliometrie erachtet (Bassecoulard, Lelu & Zitt, 2007).

Bibliometrische Feldabgrenzung ist eine spezielle Anwendung des *Information Retrievals* (IR), wobei gefundene Publikationen nach ihrem *Recall* („Trefferquote“) und ihrer *Precision Rate* („Genauigkeit“) evaluiert werden (Zitt & Bassecoulard, 2006). Mit *Silence* wird die Menge der vermuteten relevanten Publikationen bezeichnet, die nicht gefunden werden konnten, während *Noise* die Menge der falsch-positiven Treffer bezeichnet. Ziel der bibliometrischen Feldabgrenzung ist somit die Extraktion von Publikationen aus einer Datenbank, die idealerweise alle relevanten Publikationen eines Forschungsfeldes repräsentieren.

In der Bibliometrie ist es üblich Forschungsfelder auf dem „Makro-Level“ abzugrenzen. Damit wird das Vorgehen bezeichnet, alle Publikationen einer Zeitschrift, die einem Feld oder mehreren Feldern zugeordnet ist, für die Definition des Feldes heranzuziehen. Die Zuordnung von Zeitschriften zu Themenfeldern erfolgt durch die Datenbankbetreiber auf Basis des implementierten Klassifikationssystems. Laut Laurens, Zitt & Bassecoulard (2010, S.647) führt eine Makro-Level-Abgrenzung zu keinem zufriedenstellenden Verhältnis von *Noise* und *Precision*, sobald es sich um breite Felder oder interdisziplinäre Felder handelt. Die Bildungsforschung ist ein interdisziplinäres Feld, das nicht nur direkten Bezug zu den Erziehungswissenschaften hat, sondern auch zur Psychologie und den Sozialwissenschaften. Die Bildungsforschung ist jedoch ein noch viel breiteres Feld, da im Grunde jede Disziplin zur Bildungsforschung beitragen kann (Bildungsrat, 1974). Sie ist ferner gekennzeichnet durch Heterogenität, die sich in der Vielfalt der Publikationsmedien niederschlägt, der Diversität der Methoden und Forschungsthemen, der Ausrichtung auf den

nationalen Raum und der unmittelbaren Anwendung und Umsetzung von Forschungsergebnissen (siehe Hicks, 2004; Fry et al. 2009). In einem interdisziplinären Feld wie der Bildungsforschung kommt die traditionelle Makro-Level-Abgrenzung an ihre Grenzen. Zitt & Bassecoulard (2006) empfehlen daher für die Abgrenzung heterogener Felder die „lex + cite“-Methode, die sie anhand des Feldes *Nano-sciences* in *Web of Science* (*WoS*) durchführten. Diese hybride Methode bedient sich lexikalischer als auch zitationsbasierter Merkmale von Publikationen und führt im Ergebnis zu einem zufriedenstellenden Verhältnis von *Precision* und *Recall*. In dem vorliegenden Aufsatz gilt es die traditionelle Makro-Level-Feldabgrenzung in *Web of Science* als auch *Scopus* anhand des Feldes Bildungsforschung aufzuzeigen, samt den daraus resultierenden Problemen. Dem gegenübergestellt wird die „lex + cite“-Methode in *WoS* als auch in *Scopus*, die sich zur Abgrenzung der Bildungsforschung eignet.

2. Datenbasis *Web of Science* und *Scopus*

Die verwendeten Daten stammen aus den Bibliometrie-Datenbanken des *Kompetenzzentrums Bibliometrie für die deutsche Wissenschaft*¹, dessen Geschäftsstelle am *iFQ* angesiedelt ist. Zum einen verfügt das *Kompetenzzentrum Bibliometrie* (KB) über *Thomson Reuters'* Publikations- und Zitationsdaten des *Web of Science* (*WoS*). *WoS* ist eine Agglomeration aus verschiedenen Indexen. Angefangen mit dem *Science Citation Index* (SCI) in den frühen 1960ern, sind ab 1973 der *Social Sciences Citation Index* (SSCI) und 1978 der *Arts & Humanities Citation Index* (A&HCI) hinzugekommen. Da für die Aufnahme von Zeitschriften Kriterien wie zeitnahe Publikation, Aktualität der Inhalte und deren Rezeption in der Wissenschaftscommunity von Bedeutung sind, steht *WoS* stellvertretend für Qualität. *WoS* weist momentan mehr als 12.000 internationale Zeitschriften nach und bezeichnet sich selbst als selektiv bei der Aufnahme neuer Zeitschriften.² Sobald die Aufnahmekriterien von bereits indextierten Zeitschriften nicht länger erfüllt werden können, werden Zeitschriften aus der Datenbank herausgenommen.

Zum anderen verfügt das KB über die von *Elsevier* bereitgestellten *Scopus*-Daten. Die Zitations- und Abstractdatenbank *Scopus* wird seit 2004 von *Elsevier* betrieben und wurde

¹ <http://www.bibliometrie.info/>

² *The Thomson Reuters Journal Selection Process*. <http://wokinfo.com/essays/journal-selection-process/>

rückwirkend bis zum Jahr 1996 mit Publikations- und Zitationsdaten angereichert. In *Scopus* sind mehr als 20.000 peer-referierte Zeitschriften indexiert.³ Des Weiteren pflegt das KB eine Dublettendatenbank, die es ermöglicht, einer Publikation in *WoS* eine identische Publikation aus *Scopus* zuzuordnen. Dies erfolgt anhand eines Abgleichs bibliografischer Daten, wie Erstautor, Zeitschriftentitel, ISSN, Publikationsjahr, Seitenangaben oder DOI. Da die Zahl der in den Datenbanken indexierten Zeitschriften über die Jahre nicht konstant ist, ist es notwendig bei Ausweisung von Zeitschriftenzahlen Jahresangaben zu machen. So beschränken sich alle nachfolgenden Analysen auf den Publikationszeitraum 2009-2013.

3. Standardvorgehen der Feldabgrenzung

Bis zum heutigen Tag gibt es keine internationale Standardklassifikation, die die bibliometrische Forschungsarbeit unterstützt. Das von *Thomson Reuters* entwickelte Klassifikationssystem ist jedoch in der bibliometrischen Community akzeptiert und wird für die Felddefinition bibliometrischer Indikatoren wie z.B. der feldnormalisierten Zitationsrate (siehe

Waltman et al., 2011) verwendet. Das Klassifikationssystem in *WoS* besteht aus 255 sogenannten „Subject Categories“, denen Zeitschriften zugeordnet werden. In Abhängigkeit von dem Spektrum einer Zeitschrift, kann diese einer einzigen oder bis zu acht verschiedenen dieser *Subject Categories* (SC) angehören. Zusätzlich zu den SC wird jede Zeitschrift einem sogenannten „broad field“ in der *Essential Science Indicators* Datenbank zugeordnet.⁴ Anders als bei der Zuweisung von SC, sind Zeitschriften genau einem „broad field“ zugeordnet (Leydesdorff & Rafols, 2009, S.349). *Scopus* verfügt ebenfalls über ein eigenes Klassifikationssystem, bestehend aus 334 sogenannten „Descriptions“. 27 dieser *Descriptions* fungieren als Oberkategorien zur Bezeichnung von aggregierten Forschungsfeldern. Beim Standardvorgehen zur Abgrenzung eines Feldes werden die dem abzugrenzenden Feld entsprechenden Kategorien in *WoS* bzw. *Scopus* selektiert. Für die Bildungsforschung lassen sich die in Tabelle 1 aufgeführten Kategorien als relevant einstufen. Die Tabelle listet ferner die Zeitschriftenzahl, die unter der jeweiligen Kategorie in den Publikationsjahren 2009-2013 indexiert ist.

Tabelle 1: Übersicht über Themenkategorien in *WoS* bzw. *Scopus*, die Bildungsforschung zum Gegenstand haben, samt der Zahl indexierter Zeitschriften im Zeitraum 2009-2013.

<i>WoS</i>		<i>Scopus</i>	
Subject Category	n	Description	n
Education & Educational Research	265	Education	948
Education, Scientific Disciplines	40		
Education, Special	39		

Aus der Tabelle wird sichtbar, dass es in *WoS* drei relevante SC gibt, von denen eine explizit die Bildungsforschung in der Bezeichnung trägt. Obwohl das Klassifikationssystem in *Scopus*, angesichts der Zahl unterscheidbarer Kategorien, granularer als in *WoS* ist, gibt es lediglich die Kategorie „Education“. Hierunter sind im 2009-2013-Zeitraum 948 Zeitschriften nachweisbar. In *WoS* beträgt die Gesamtzahl potenziell relevanter Zeitschriften 305.⁵ Im nächsten

Schritt werden die Publikationen dieser Zeitschriften für den Publikationszeitraum 2009-2013 aggregiert. In *WoS* führt das zu einer Gesamtzahl von 83.353 Publikationen, während es in *Scopus* mit 171.458 mehr als doppelt so viele sind. Ein Abgleich der Publikationen in der Dublettendatenbank zeigt, dass zu 63.588 Publikationen in *WoS* eine identische Publikationen in *Scopus* identifizierbar ist. Dies entspricht 76,3% der *WoS*-Publikationen, jedoch nur

³ *Scopus: Content Overview*. <http://www.elsevier.com/online-tools/scopus/content-overview>

⁴ Die *Essential Science Indicators* dienen als Statistik für Regierung, Universitäten oder Forschungsauftraggeber. <http://ipscience-help.thomsonreuters.com/incitesLiveESI/ESIGroup/overviewESI/esiJournalsList.html>

⁵ Aufgrund der Mehrfachvergabe von SC handelt es sich um 305 Zeitschriften und nicht die Summe der drei SC in Tabelle 1.



37,1% der Scopus-Publikationen. WoS-Publikationen, zu denen keine Dublette in Scopus gefunden werden konnte, sind entweder nicht in Scopus indexiert oder mit einer anderen Kategorie als „Education“ versehen. Wie stark sich die auf dem Makro-Level angewandte Feldab-

grenzung in WoS und Scopus voneinander unterscheidet, wird in der nachfolgenden Tabelle sichtbar. Hierin sind die in dem jeweiligen Korpus erfassten Zeitschriften aufgeführt, absteigend nach der Zahl der darin veröffentlichten Publikationen.

Tabelle 2: Übersicht der Zeitschriften in WoS und Scopus, samt Publikationszahl und Prozentsatz an dem abgegrenzten Korpus zur Bildungsforschung. Die Daten basieren auf der Makro-Level-Abgrenzung in dem Publikationszeitraum 2009-2013.

WoS			Scopus		
Zeitschrift	n	%	Zeitschrift	n	%
Academic Medicine	2.097	2,5	Laryngoscope	3.483	2,0
Journal of Chemical Education	1.942	2,3	Advanced Science Letters	3.400	2,0
Medical Teacher	1.646	2,0	Academic Medicine	2.206	1,3
Journal of Intellectual Disability Research	1.492	1,8	Journal of Chemical Education	1.995	1,2
Research in Developmental Disabilities	1.442	1,7	Medical Teacher	1.708	1,0
Medical Education	1.258	1,5	American Journal of Pharmaceutical Education	1.400	0,8
Computers & Education	1.253	1,5	International Journal of Learning	1.358	0,8
American Journal of Pharmaceutical Education	1.003	1,2	Medical Education	1.290	0,8
Nurse Education Today	941	1,1	Children and Youth Services Review	1.283	0,7
American Journal of Physics	910	1,1	Computers & Education	1.254	0,7
European Journal of Physics	868	1,0	Chemical Research in Chinese Universities	1.105	0,6
Phi Delta Kappan	857	1,0	Journal of Extension	999	0,6
British Journal of Educational Technology	834	1,0	Nurse Education Today	973	0,6
Educational Leadership	815	1,0	Journal of Dental Education	868	0,5
International Journal of Engineering Education	742	0,9	Phi Delta Kappan	812	0,5
	18.100	21,7		24.134	14,1

Fett hervorgehoben sind diejenigen Zeitschriften in Tabelle 2, die sowohl in WoS als auch in Scopus zu den Zeitschriften mit der höchsten Publikationszahl gehören. In WoS wird der Korpus dominiert von der Zeitschrift *Academic Medicine*, während in Scopus dieser Zeitschrift zwei andere vorangestellt sind. Die Prozentwerte zeigen, dass die aufgeführten Zeitschriften in WoS zu 21,7% des gesamten abgegrenzten Korpus beitragen, während es in Scopus nur 14,1% sind. Dies ist primär darauf zurückzuführen, dass die Publikationen in Scopus über mehr als dreimal so viele Zeitschriften wie in WoS streuen. Auffällig ist, dass viele der Zeitschriften dem Gesundheitswesen und der Medizin angehören. Dies liegt zum einen

darin, dass medizinische Zeitschriften nahezu vollständig in den Datenbanken indexiert sind. Zum anderen sind Zeitschriften der Medizin „publikationsstark“, d. h. sie weisen eine durchschnittlich höhere Zahl an Publikationen pro Jahrgang auf, als Zeitschriften anderer Disziplinen. Insgesamt zeigt die Tabelle 2, dass das Verständnis darüber, welche Zeitschriften zur Bildungsforschung zu zählen sind, in den Datenbanken WoS und Scopus divergiert.

Die Klassifikation von Zeitschriften in WoS als auch Scopus erfolgt anders als vermutet nicht allein auf intellektueller Basis. SC in WoS werden von Mitarbeitern Thomson Reuters' auf Basis mehrerer Kriterien vergeben, so z.B. an-

hand des Zeitschriftentitels und der Zitationsmuster (Leydesdorff & Rafols, 2009; Pudovkin & Garfield, 2002). Die Klassifikationsschemata müssen daher als das Ergebnis eines agglomerierenden Prozesses betrachtet werden, bei dem Ähnlichkeiten zwischen zwei Einheiten anhand von Gleichheit oder Ungleichheit ausfindig gemacht werden. Dies erfolgt durch die Analyse von Zeitschriftennetzwerken, worin Zeitschriften die Knoten darstellen und Kanten durch aggregierte Zitierungen auf Zeitschriftenartikel repräsentiert werden. Die Analyse von Zitationsbeziehungen unter Zeitschriften ermöglicht es Cluster abzubilden, wobei zwei Zeitschriften mit gleichem Muster der eingehenden Zitierungen als gleich erachtet werden (Leydesdorff, 1987). Trotz dieses angewandten referenzbasierten Verfahrens zur Klassifikation von Zeitschriften fanden Boyack et al. (2005), dass knapp die Hälfte der SC in WoS nicht mit den Clustern korrespondiert, die aus Zitationsbeziehungen zwischen Zeitschriften generiert werden. Dies führt dazu, dass Zeitschriften zu mehreren Kategorien zugeordnet werden oder zur Kategorie „Multidisciplinary“. Das gewährleistet zwar, dass das gesamte thematische Spektrum einer Zeitschrift erfasst wird, es entsteht jedoch das Problem, dass diese Zeitschriften bei einer SC-basierten Makro-Level Feldabgrenzung unberücksichtigt bleiben. Wenn Zeitschriften zusehends als „multidisziplinär“ klassifiziert werden, können Klassifikationsschemata kein vollständiges Bild der Forschungsleistung in einem zu untersuchenden Feld widerspiegeln, insbesondere wenn relevante hochzitierte Publikationen in diesen multidisziplinären Zeitschriften erscheinen (Bornmann, et al., 2008). Ein Klassifikationssystem sollte weitgehend stabil bleiben und zugleich regelmäßig aktualisiert werden, um Inkonsistenzen zu vermeiden. Wissenschaft ist dynamisch, es entstehen fortwährend neue Forschungsfelder und die Terminologie ist ständig im Wachstum begriffen. Es sollte vermieden werden, dass neugegründete Zeitschriften in ein ex ante Klassifikationssystem gepresst werden, obwohl die Neudefinition einer Kategorie vonnöten wäre. Die SC in WoS werden zwar auf jährlicher Basis aktualisiert, können aber nicht im Nachhinein geändert werden (Leydesdorff & Rafols, 2009). Daher wird eine Zeitschrift, die ehemals unter Psychologie klassifiziert war, sich jedoch nun vermehrt der Bildungsforschung widmet, zwar in den aktuellen Jahrgängen unter *Education & Educational Research* klassifiziert sein, die früheren Jahrgänge bleiben jedoch der Psychologie zugeordnet. Dies scheint plausibel vor dem Hintergrund, dass die Zeitschriftenklassifikation in WoS ursprünglich zum Zwecke des *Retrievals* geschaffen wurde

(Glänzel & Schubert, 2003). Durch die Mehrfachvergabe von SC bzw. *Descriptions* wird der Zweck des *Retrievals* übererfüllt, wenn Zeitschriften unter einer Vielzahl verschiedener Kategorien auffindbar sind.

Unschärfe bei der Makro-Level-Abgrenzung entsteht somit nicht nur durch inkorrekte Zuordnung von Zeitschriften oder eine Mehrfachzuordnung, sondern auch aus dem zu untersuchenden Forschungsfeld. Bildungsforschung ist ein multidisziplinäres Feld mit einer geringen Anzahl an Kernzeitschriften, stattdessen einer breiten Streuung von Publikationen über diverse Zeitschriften. Goodyear et al. (2009) untersuchten, ob sich ein Kern an Zeitschriften in der Bildungsforschung bestimmen lässt. Sie schlussfolgerten, dass aufgrund hoher Diversität der Zeitschriften und Inhalte, es keinen Konsens über „core journals“ geben kann (ebd., S.702). Eine Feldabgrenzung auf Makro-Level ist somit bei breiten Forschungsfeldern als auch bei hochspezialisierten Forschungsfeldern problematisch. Es werden per se Artikel solcher Zeitschriften ausgeschlossen, die nicht explizit mit einer SC versehen ist, die mit dem abzugrenzenden Feld assoziiert wird.

4. Hybride Feldabgrenzung nach der „lex + cite“-Methode

Die Makro-Level-Abgrenzung des Feldes Bildungsforschung in WoS und Scopus zeigt, dass die abgegrenzten Korpora sich quantitativ als auch bezüglich der erfassten Zeitschriften stark unterscheiden. Daher gilt es im Folgenden die Abgrenzung der Bildungsforschung in Anlehnung an die hybride „lex + cite“-Methode durchzuführen, wie sie von Zitt & Bassecoulard (2006) am Beispiel der Nanowissenschaften empfohlen wird. Laurens, Zitt & Bassecoulard (2010) verwendeten diese Methode, um das Feld der Genomik abzugrenzen. Die „lex + cite“-Methode kann folgendermaßen skizziert werden: Zu Anfang schlagen Zitt & Bassecoulard (2006) die Selektion von Kernzeitschriften vor. Im nächsten Schritt werden lexikalische Suchstrings formuliert, um einen Kern („Seed“) an Dokumenten zu extrahieren, der eine hohe *Precision* aufweist. Im letzten Schritt wird dieser Kern an relevanten Publikationen zitationsbasiert ausgeweitet. Mit dem Ziel eine möglichst hohe *Precision Rate* zu erreichen, können die Schritte in der lexikalischen und zitationsbasierten Suche beliebig oft iteriert werden.

4.1 Zeitschriftenbasierte Abgrenzung

Da es in den Datenbanken an Transparenz bei der Zuordnung von Kategorien zu Zeitschriften mangelt, sollten für die Bestimmung von

Kernzeitschriften in einem Feld Experten herangezogen werden. Alternativ kann man sich mit bereits bestehenden Klassifikationssystemen behelfen, die von Experten geschaffen wurden, die über entsprechendes Hintergrundwissen verfügen. Eine wichtige Quelle für die Bildungsforschung ist der *European Reference Index for the Humanities and Social Sciences (ERIH-PLUS)*, bis 2014 unter *ERIH* geläufig.⁶ Seit 2008 veröffentlicht *ERIH* Zeitschriftenlisten, die nach thematischen Kategorien der Geisteswissenschaften durchsuchbar sind. In 2014 sind die *Social Sciences* hinzugekommen, was

durch den Zusatz „*PLUS*“ kenntlich gemacht wurde. Unter der Kategorie „*Pedagogical & Educational Research*“ sind mit Stand 2013 insgesamt 531 Zeitschriften gelistet. Die ISSN dieser Zeitschriften wurde für einen Abgleich in *WoS* bzw. *Scopus* extrahiert und aufbereitet. In *WoS* lassen sich 246 (46,3%) dieser Zeitschriften auffinden, während es in *Scopus* 375 (70,6%) sind. Die nachfolgende Tabelle gibt Auskunft darüber, unter welchen Kategorien die *ERIH*-Zeitschriften in *WoS* bzw. *Scopus* klassifiziert sind.

Tabelle 3: Übersicht über die 10 häufigsten Kategorien, der in *WoS* bzw. *Scopus* zwischen 2009 und 2013 nachweisbaren *ERIH*-Zeitschriften der Kategorie *Pedagogical & Educational Research*.

Subject Category in WoS	n	%	Description in Scopus	n	%
Education & Educational Research	165	68,8	Education	326	88,8
Psychology, Educational	31	12,9	Psychology (all)	121	33,0
Education, Special	25	10,4	Social Sciences (all)	118	32,2
Rehabilitation	18	7,5	Arts and Humanities (all)	101	27,5
Linguistics	18	7,5	Developmental and Educational Psychology	64	17,4
Education, Scientific Disciplines	17	7,1	Health Professions (all)	32	8,7
Language & Linguistics Theory	14	5,8	Linguistics and Language	26	7,1
Social Sciences, Interdisciplinary	10	4,2	Language and Linguistics	23	6,3
Psychology, Developmental	6	2,5	Sociology and Political Science	20	5,4
Health Care Sciences & Services	6	2,5	Computer Science Applications	16	4,4

Tabelle 3 zeigt, dass die Mehrheit der *ERIH*-Zeitschriften in *WoS* unter *Education & Educational Research* klassifiziert ist. In *Scopus* sind fast 90% der Zeitschriften unter *Education* auffindbar. Sowohl in *WoS* als auch in *Scopus* folgt an zweiter Stelle die Psychologie, die jedoch in *WoS* mit dem Zusatz „*Educational*“ versehen ist. Insgesamt wird deutlich, dass Kategorien, die stellvertretend für die Linguistik, das Gesundheitswesen und die Sozialwissenschaften stehen, für die Klassifizierung der *ERIH*-Zeitschriften fungieren. Die *ERIH*-Zeitschriften streuen in *WoS* über insgesamt 59 *SC*, während es in *Scopus* 125 *Descriptions* sind, die für die Klassifikation der *ERIH*-Zeitschriften gebraucht wurden. Die Publikationen dieser Zeitschriften wurden für den Zeitraum 2009-2013 aggregiert. In *WoS* führt das zu 67.850 Publikationen, in *Scopus*, in Einklang mit der höheren Zahl indexierter *ERIH*-Zeitschriften, zu 76.726 Publikationen. Eine erste Durchsicht der Artikeltitel zeigt, dass es sich bei der Mehrheit der Publikationen um Un-

terrichtsinhalte oder Anleitungen für den Unterricht handelt. Beispiele hierfür sind „Plato“, „ph Basics“, „Platinum“ oder „The Stranger“. Ferner gibt es Artikeltitel, die keinerlei Auskunft über den Inhalt der Publikation geben, wie z.B. „Leaving“, „Progress“ oder „Reaction“. Das Ziel der Feldabgrenzung ist es jedoch genuine Forschungsartikel zur Bildungsforschung zu identifizieren und nicht Inhalte des Unterrichts. Daher gilt es im nächsten Schritt den *Noise* in den *ERIH*-Publikationen zu reduzieren.

4.2 Lexikalische Abgrenzung

Der lexikalische Ansatz ermöglicht eine präzise Extraktion relevanter Literatur, indem a priori definierte Terme mit bibliografischen Daten von Publikationen abgeglichen werden. Laut Zitt & Bassecouard (2006) besteht beim lexikalischen Ansatz eine größere Wahrscheinlichkeit für *Silence* als für *Noise*. Das Ziel der lexikalischen Abgrenzung besteht jedoch darin möglichst alle Aspekte der Bildungsforschung zu er-

⁶ <https://dbh.nsd.uib.no/publiseringkanaler/erihplus/>



fassen, ohne den *Noise* zu erhöhen. Das *Retrieval* mit lexikalischen Termen ist ein mächtiges Werkzeug, solange die richtigen Terme verwendet werden. Daher sollte die Zusammenstellung brauchbarer Terme in Zusammenarbeit mit Experten erfolgen. Alternativ kann man sich jedoch bereits bestehenden Expertenwissens bedienen. So bietet das *Education Resources Information Center (ERIC)* einen Thesaurus zu insgesamt 41 Kategorien, angefangen bei *Agriculture and Natural Resources* hin zu *Tests and Scales*. Für die Charakterisierung der Bildungsforschung eignen sich die nachfolgenden vier Kategorien:

- Educational Levels, Degrees, and Organizations
- Educational Process: Classroom Perspectives
- Educational Process: School Perspectives
- Educational Process: Societal Perspectives

Insgesamt befinden sich unter diesen vier Kategorien zu Bildungsstufen und –prozessen 781 Terme. Diese wurden für die lexikalische Suche aufbereitet, indem Zusätze in Klammern entfernt wurden, Flexionen und Singular-/Pluralformen berücksichtigt wurden und die amerikanische und britische Schreibweise vereinheitlicht wurde. So können durch den Term „behav“ Publikationen gefunden werden, die entweder „behaviour“ (Br.) oder „behavior“ (Am.) zum Gegenstand haben. Gegenüber der Groß- und Kleinschreibung ist die Suche nicht sensitiv, sodass mit dem Suchterm „behav“ auch der Artikelanfang „Behaviour...“ gefunden werden kann.

Die *ERIC*-Terme können einerseits mit dem Artikeltitel einer Publikation abgeglichen werden. Da dieser jedoch nicht immer Auskunft über den Inhalt gibt, kann zusätzlich in den Keywords einer Publikation gesucht werden.⁷ Nicht jede der *ERIH*-Publikationen ist mit Keywords versehen. In *WoS* sind es 74,0%, während es in *Scopus* 66,8% der *ERIH*-Publikationen sind. Daher bilden Abstracts eine weitere relevante Quelle für den Abgleich valider Terme. In *WoS* beläuft sich der Anteil der *ERIH*-Publikationen mit Abstract auf 74,2%, während es in *Scopus* mit 85,6% deutlich mehr sind. Dass es Publikationen ohne Abstract oder Keywords gibt, liegt u.a. daran, dass es neben Forschungsartikeln andere Publikationsarten wie z.B. „Letters“, „Corrections“ oder „Notes“ gibt. Um jeglichen Bias zu vermeiden, der aus

der Art der Publikation oder der Qualität der Datenbanken resultiert, ist die Suche sowohl auf Basis des Artikeltitels, der Keywords als auch des Abstracts erfolgt. Die *Precision Rate* der *ERIH*-Publikationen wurde gesteigert, indem vorausgesetzt wurde, dass eine valide Publikation über mindestens einen *ERIC*-Term verfügen muss, unabhängig davon, ob dieser im Artikeltitel, in den Keywords oder im Abstract vorkommt. Dies resultiert in *WoS* in einer Menge von 52.733 Publikationen und damit 77,7% der Ausgangsgröße. In *Scopus* verfügen 83,6% (64.106) der Publikationen über mindestens einen *ERIC*-Term. Durch die lexikalische Abgrenzung wurden rigide *ERIH*-Publikationen eliminiert, die weder im Titel, noch in den Keywords oder im Abstract über einen *ERIC*-Term verfügen und somit keinen Bezug zur Bildungsforschung erkennen lassen. Einige der *ERIC*-Terme mögen zwar zu allgemein erscheinen, um den *Noise* zu reduzieren, jedoch wird die lexikalische Abgrenzung innerhalb der Menge der *ERIH*-Publikationen durchgeführt, sodass für eine Kontextualisierung gesorgt ist. Eine Suche nach „school“ in der gesamten Datenbank würde auch zu Treffern führen, die nicht die Lehranstalt, sondern den Fischschwarm zum Gegenstand haben. Durch die vorangehende zeitschriftenbasierte Abgrenzung wird jedoch selbst ein allgemeines Wort wie „learning“ mit Lehranstalt und lebenslangem Lernen assoziiert, anstatt mit „Maschinelle Lernen“ - einem Terminus der Informationstechnologie.

Für den nachfolgenden Schritt der zitationsbasierten Ausweitung ist es notwendig innerhalb der lexikalisch abgegrenzten Publikationen einen Kern an Publikationen zu definieren, der idealerweise eine *Precision Rate* von 100% aufweist (Zitt & Bassecouard, 2006). Die Wahrscheinlichkeit, dass eine Publikation Bildungsforschung zum Gegenstand hat, steigt mit der Anzahl der vorkommenden *ERIC*-Terme. Daher wurde vorausgesetzt, dass eine *Seed*-Publikation über mindestens drei verschiedene *ERIC*-Terme verfügen muss, unabhängig davon, ob diese im Titel, als Keyword oder im Abstract vorkommen. Dies resultiert in 27.633 *Seed*-Publikationen innerhalb der *ERIH*-Publikationen in *WoS*. In *Scopus* sind es 35.551 Publikationen, die über drei oder mehr *ERIC*-Terme verfügen. Diese *Seed*-Publikationen dienen als valide Basis für den Expansionsprozess, der im nächsten Kapitel beschrieben wird. Die nachfolgende Abbildung illustriert die bisherigen Schritte der Abgrenzung.

⁷ Unter den *ERIH*-Publikationen in *WoS* gibt es 499, die im Artikeltitel mit „Untitled“ belegt sind.



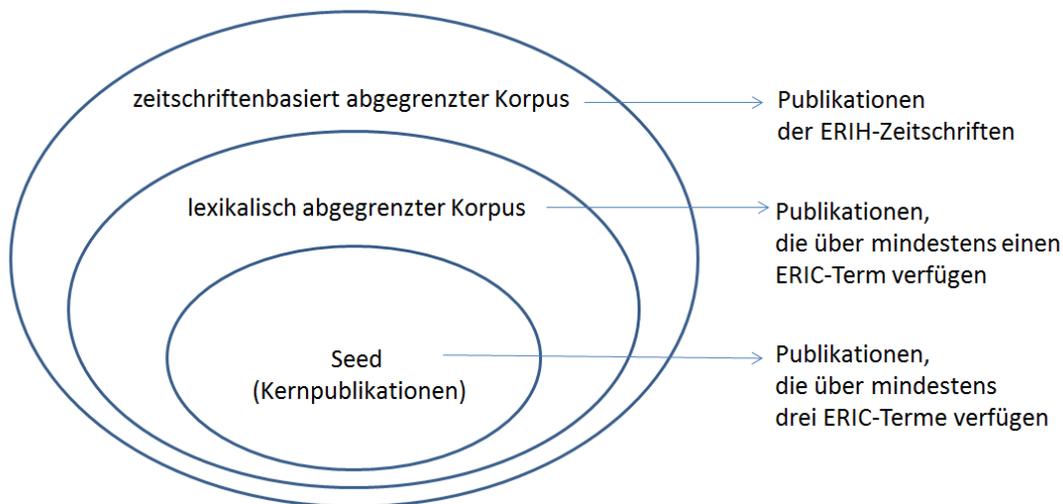


Abbildung 1: Übersicht über Relationen der zeitschriftenbasierten und lexikalischen Abgrenzung und der Definition eines Seeds.

Die lexikalische Abgrenzung erhöht die Menge der relevanten Publikationen ohne den *Noise* zu erhöhen. In diesem Stadium der Feldabgrenzung ist ein geringer *Noise*-Wert wichtig, da es erst in der zitationsbasierten Abgrenzung darum geht *Silence* zu reduzieren (Zitt & Bassecouard, 2006). Der Nachteil der lexikalischen Abgrenzung kann jedoch im geringen *Recall* bestehen, insbesondere in einem interdisziplinären Feld mit unspezifischem Vokabular. Bildungsforschung ist terminologisch schwer zu erfassen, da es sich überwiegend der Allgemeinsprache bedient. Anders als in den Naturwissenschaften, mangelt es an spezifischen Termen, die eindeutig der Bildungsforschung zuzuschreiben sind. Weder in *WoS* noch in *Scopus* liegt kontrolliertes Vokabular vor und natürliche Sprache ist gekennzeichnet durch Synonyme und Homonyme. Laut Zitt & Bassecouard (2006, S.1518) liegt die primäre Ursache für „Noise“ in Homonymen, die schwer zu disambiguieren sind. Die Suche nach Homonymen erhöht zwar den *Recall*, damit verbunden aber auch den Aufwand diese zu disambiguieren. So kann das Wort *Education* bereits „Noise“-Publikationen erzeugen, da es sich nicht allein mit Bildung übersetzen lässt, sondern ebenso Erziehung und Aufklärung bezeichnet. Ein Aufsatz zu „AIDS education“ hätte somit das Gesundheitswesen zum Thema und weniger die Bildungsforschung. Ferner sind der natürlichen Sprache keine Grenzen gesetzt. Eine lexikalische Suche einzig im Artikeltitel nach dem potenziell relevanten Aufsatz „*Txtng: The gr8 db8*“, würde mit a priori definierten

Suchtermen nicht gelingen, sodass diese Publikation zu der Menge der „*Silence*“-Publikationen zu zählen wäre. Gleichermäßen können sich neue Aspekte der Bildungsforschung herausbilden, einhergehend mit neuem Vokabular, dem mit a priori definierten Termen nicht Rechnung getragen werden kann. Diese und andere Limitierungen kann die im nächsten Kapitel beschriebene zitationsbasierte Abgrenzung abwenden.

4.3 Zitationsbasierte Abgrenzung

Während der lexikalische Ansatz dazu dient Publikationen mit einer hohen *Precision Rate* zu finden (notfalls zu Lasten des *Recalls*), ist das Ziel der zitationsbasierten Abgrenzung die Seed-Publikationen um solche Publikationen zu ergänzen, die idealerweise die gleiche Wissensbasis teilen (Zitt & Bassecouard, 2006, S.1521). In der zitationsbasierten Abgrenzung werden Publikationen hinzugenommen, die vom Seed zitiert werden, d.h. die Referenzen der Seed-Publikationen (ebd.). Ferner wird der Seed um zitierende und um bibliografisch gekoppelte Publikationen ergänzt. Bibliografische Kopplung ist ein Maß für thematische Ähnlichkeit zweier Publikationen, das anhand der Zahl gleicher Referenzen ermittelt wird. Wenn eine Publikation die gleiche Referenz aufweist wie eine Seed-Publikation, so sind diese zwei Publikationen bibliografisch gekoppelt. Je mehr gemeinsame Referenzen zwei Publikationen aufweisen, desto ähnlicher sind sie sich bezüglich ihres thematischen Spektrums (Persson, 1994, S.31). Durch die Ausweitung des Seeds mittels

dieser drei zitationsbasierten Beziehungen umfasst der finale Korpus in *WoS* 37.239 Publikationen. In *Scopus* führt die zitationsbasierte Abgrenzung zu insgesamt 47.985 Publikationen.

Die Beziehungen und Dimensionen der zitationsbasierten Abgrenzung sind in der nachfolgenden Abbildung visualisiert.

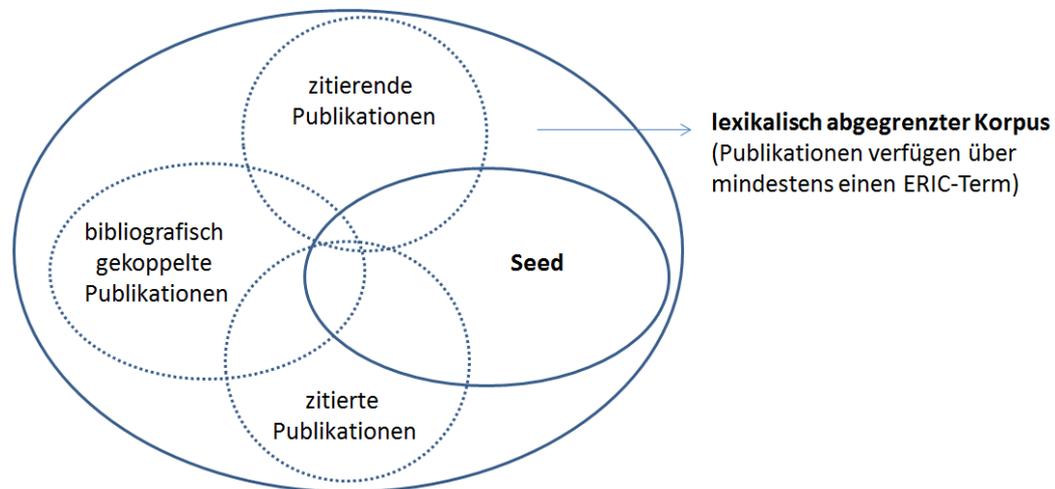


Abbildung 2: Übersicht über die Relationen der zitationsbasierten Abgrenzung.

Der Prozess der zitationsbasierten Abgrenzung ist notwendig, um alle relevanten Artikel zu identifizieren, die mit a priori definierten Termen nicht gefunden werden konnten. Diese zitationsbasierten Publikationen können zum Seed hinzugefügt werden, da sie innerhalb des Kontexts der Bildungsforschung sind. Durch den hohen Grad der Konnektivität unter Publikationen, die den Seed zitieren oder von diesem zitiert werden, ist der Prozess der Ausweitung jedoch begrenzt.

Der Vorteil der zitationsbasierten Abgrenzung liegt darin, dass es die Grenzen kontrollierter Suchterme kompensiert. Ein weiterer Vorteil der zitationsbasierten Abgrenzung ist, dass diese automatisch prozessiert werden kann, während der lexikalische Ansatz ohne Experten, die das relevante Vokabular kennen, schwer zu implementieren ist (Zitt & Bassecouard 2006, S.1519). Die Feldabgrenzung nach der „lex + cite“-Methode hat ihren Fokus auf der *Precision* als auf dem *Recall* (ebd.). Es ist nahezu unmöglich das gesamte Feld der Bildungsforschung zu erfassen. Da es in der Literatur zugleich keine Angabe über die Forschungsleistung in der Bildungsforschung gibt,

die sich in Zeitschriftenaufsätzen niederschlägt, fehlt es einer Orientierungsgröße für den zu erwartbaren *Recalls*. Daher sollte das abgegrenzte Publikationsset in *WoS* als auch in *Scopus* als Ausschnitt angesehen werden. Bei diesem Ausschnitt handelt es sich jedoch mit hoher Wahrscheinlichkeit um Publikationen, die die Bildungsforschung zum Gegenstand haben.

Dies wird u.a. dadurch gewährleistet, dass nicht gesamte Zeitschrifteninhalte zur Abgrenzung des Feldes herangezogen wurden, sondern nur die darin befindlichen relevanten Publikationen. Die nachfolgende Tabelle zeigt das Ergebnis der Feldabgrenzung in *WoS* und *Scopus* in Hinblick auf die erfassten Zeitschriften, samt Publikationszahl. In der Tabelle sind diejenigen Zeitschriften in *WoS* und *Scopus* gelistet, die zu den meisten Publikationen in dem final abgegrenzten Korpus beitragen. Zusätzlich ist der Prozentsatz, der durch die Feldabgrenzung erfassten Publikationen an der Gesamtzahl der in der jeweiligen Zeitschrift im Zeitraum 2009-2013 veröffentlichten Publikationen, angegeben.



Tabelle 4: Übersicht über die in den finalen Korpora dominierenden Zeitschriften, samt des Prozentsatzes der als relevant identifizierten Publikationen an der Gesamtzahl der in den Zeitschriften im Zeitraum 2009-2013 veröffentlichten Publikationen.

<i>WoS</i>			<i>Scopus</i>		
Zeitschrift	n	%	Zeitschrift	n	%
Computers & Education	1200	95,8	Computers & Education	1190	94,9
Medical Teacher	1078	65,5	Medical Teacher	1047	61,3
Computers in Human Behavior	676	54,5	Journal of Chemical Education	939	47,1
Medical Education	675	53,7	Computers in Human Behavior	652	52,3
Journal of Chemical Education	650	33,5	Teaching and Teacher Education	641	92,6
Teaching and Teacher Education	647	92,3	Medical Education	617	47,8
Research in Developmental Disabilities	582	40,4	Research in Developmental Disabilities	561	38,5
International Journal of Engineering Education	541	72,9	International Journal of Engineering Education	553	74,2
International Journal of Science Education	538	92,8	International Journal of Science Education	521	93,4
Economics of Education Review	453	89,3	Higher Education	465	94,7
Higher Education	446	87,5	British Journal of Educational Technology	463	77,8
Teachers College Record	446	93,5	Teachers College Record	445	96,9
Kuram ve uygulamada eğitim bilimleri	437	79,6	Economics of Education Review	441	86,5
Learning and Individual Differences	435	74,1	Learning and Individual Differences	429	73,3
Educational Technology & Society	409	78,7	Kuram ve uygulamada eğitim bilimleri	406	79,8
	9.213			9.370	

Fett hervorgehoben sind diejenigen Zeitschriften, die in den abgegrenzten Korpora in *WoS* als auch in *Scopus* erfasst wurden. Aus der Tabelle geht hervor, dass die Zeitschriften *Computers & Education* und *Medical Teacher* in *WoS* als auch in *Scopus* die meisten Publikationen zum jeweiligen finalen Korpus beitragen. Während jeweils ca. 95% der Publikationen aus *Computers & Education* durch die Feldabgrenzung erfasst wurden, sind es aus der Zeitschrift *Medical Research* lediglich zwei Drittel aller Publikationen. Den geringsten Prozentsatz erfasster Publikationen weist die Zeitschrift *Research in Developmental Disabilities* auf. Etwa 40% der Publikationen dieser Zeitschrift wurden mit der „lex + cite“-Methode erfasst und daher

nur solche, die einen Bezug zur Bildungsforschung haben.⁸ Die erfassten Zeitschriften in *WoS* und *Scopus* sind überwiegend gleich, auch wenn es Schwankungen in den Publikationszahlen gibt. Diese können aus der Qualität und Vollständigkeit bibliografischer Daten resultieren, wie z.B. dem Fehlen von Abstracts oder Keywords. Eine weitere Ursache liegt darin, dass nur solche Zitationen erfasst werden können, die in der Datenbank als zitierende Dokumente vorhanden sind. In die Gesamtmenge der Publikationen einer Zeitschrift in Tabelle 4 fließen alle Dokumentarten ein. Bei der lexikalischen Abgrenzung werden jedoch Dokumentarten wie *Editorials*, *Letters* oder *Corrections* mit einer geringeren Wahrscheinlichkeit gefunden, da sie selten über Abstract oder Keywords

⁸ Laut Elsevier ist der Scope von *Research in Developmental Disabilities* behavior analysis, pharmacotherapy, traditional assessment, behavioral assessment, speech training and occupational therapy. <http://www.journals.elsevier.com/research-in-developmental-disabilities/>

verfügen. Auch in der zitationsbasierten Abgrenzung werden diese Publikationsarten mit geringerer Wahrscheinlichkeit identifiziert, da diese seltener zitiert werden als Forschungsartikel. Die „lex + cite“-Methode vernachlässigt somit Dokumentarten wie *Letters*, *Editorials* oder *Notes* und trägt dadurch zur Steigerung der *Precision Rate* bei, da Forschungsergebnisse primär in Artikeln erscheinen. Ein Abgleich der Korpora in der Dublettendatenbank zeigt, dass zu 34.169 Publikationen in *WoS* eine identische Publikation in *Scopus* gefunden werden kann. Dies entspricht 91,8% der *WoS*-Publikationen. Umgekehrt entspricht die Menge der identischen Publikationen 77,6% der *Scopus*-Publikationen. Diese recht hohen Prozentwerte zeigen, dass die Anwendung der „lex + cite“-Methode in beiden Datenbanken zu einem zufriedenstellenden Ergebnis führt. Die Qualität der Feldabgrenzung hängt nicht nur von der verwendeten Methode ab, sondern auch von der Abdeckung der zugrundeliegenden Datenbank und der Genauigkeit und Vollständigkeit der Daten. Die in den finalen Korpora vorliegenden Publikationen haben nicht nur das Feld Bildungsforschung zum Gegenstand, sondern erlauben auch eine Vergleichbarkeit des Feldes unter den Datenbanken *WoS* und *Scopus*.

5 Zusammenfassung

Feldabgrenzung ist ein wesentlicher Bestandteil bibliometrischer Studien und wird nicht umsonst als ein eigenständiges Forschungsfeld gesehen. Die Qualität der Feldabgrenzung ist essenziell für die Analyse von Publikationen. Laut Glänzel & Schubert (2003) ist die Klassifizierung von wissenschaftlicher Literatur in Themenkategorien eine der Grundvoraussetzungen für eine valide bibliometrische Analyse. Die wachsende Interdisziplinarität der Forschung erschwert zusehends die Grenzziehung zwischen Disziplinen. Wissenschaftliche Disziplinen sind ebenso wie Zeitschriften im Wandel begriffen, weshalb Bibliometriker schon von Anfang an mit dem Problem der Feldabgrenzung konfrontiert sind (Leydesdorff, 2004). Laut Pudovkin & Garfield (2002) haben selbst die gebräuchlichsten Zeitschriftenklassifikationen ihre Mängel und sind nur Annäherungen und vage Abbilder von Forschungsfeldern. Referenzbasierte Klassifikationssysteme sind zwar leicht zu implementieren, unterliegen jedoch Inkonsistenzen und Fehlinterpretationen aufgrund von Unterschieden in der Publikations- und Zitationsaktivität innerhalb von Sub-Disziplinen (Glänzel & Schubert, 2003). Autoren begrenzen sich selten auf eine einzige Sub-Disziplin und bedienen stattdessen ein breites Spektrum an Sub-Disziplinen mit verschiedener Gewichtung.

Nichtsdestotrotz können Zeitschriftensets auf aggregierter Ebene als Indikatoren für die intellektuelle Organisation der Wissenschaften erachtet werden (Leydesdorff, 1987). Die Klassifikationssysteme von Zeitschriften können auf jährlicher Basis relativ konstant reproduziert werden und als Indikatoren des wissenschaftlichen Wandels genutzt werden (Leydesdorff, 2002). Es besteht jedoch weiterhin das Problem Zeitschriftensets konsistent über längere Zeiträume voneinander abzugrenzen (ebd.). Eine Feldabgrenzung auf Zeitschriftenbasis wird keine so genauen Ergebnisse erzielen, wie eine auf Artikelebene, da Zeitschriften zwar einer Kategorie zugeordnet sein können, aber dennoch ein breites Themenspektrum abdecken. Daher wird eine Feldabgrenzung auf dem Makro-Level mit hoher Wahrscheinlichkeit Artikel inkludieren, die einen schwachen Bezug zum abzugrenzenden Feld haben und zugleich relevante Artikel verfehlen (Gómez et al., 1996, S.224). Um interdisziplinäre Felder adäquat abgrenzen zu können, erdachten Zitt & Bassecoulard (2006) eine Methode, die sich der Kombination aus Schlagwortsuche und Zitationsanalyse bedient. Die hier vorgestellte Methode zur Abgrenzung der Bildungsforschung beruhte im Wesentlichen auf dieser Top-Down-Methode. Dazu wurden zunächst potenziell relevante Zeitschriften zurate gezogen. Durch die Verwendung von validen Termen konnte die *Precision* gesteigert werden und der *Noise* reduziert werden. Der zitationsbasierte Ansatz wurde wiederum genutzt, um den *Recall* zu erhöhen und die *Silence* zu senken.

Sich von der Makro-Level-Abgrenzung abzuwenden und stattdessen eine eigene Feldabgrenzung vorzunehmen, mit dem Ziel einer hohen Genauigkeit, geschieht oftmals zu Lasten der Vergleichbarkeit mit anderen Studien. Die hier vorgestellte Methode führt jedoch unter Verwendung externer Ressourcen wie der *ERIH*-Zeitschriften und der *ERIC*-Terme in *WoS* als auch *Scopus* zu vergleichbaren Korpora. Die Zahl und Diversität von Klassifikationssystemen und Datenbanken führen bei vergleichenden internationalen Studien zu Schwierigkeiten, sobald verschiedene Datenquellen und die darin implementierten Klassifikationsschemata gebraucht werden (Gómez et al., 1996). Wichtig ist jedoch eine systematische Herangehensweise, um reproduzierbare Ergebnisse zu garantieren.



6 Referenzen

- Bassecouard, E., Lelu, A., & Zitt, M. (2007). A modular sequence of retrieval procedures to delineate a scientific field : from vocabulary to citations and back. In: Proceedings ISSI 2007 (Madrid). 74-84.
- Bornmann, L., Mutz, R., Neuhaus, C. & Daniel, H. (2008). Citation counts for research evaluation: standards of good practice for analyzing bibliometric data and presenting and interpreting results. *Ethics in Science and Environmental Politics*, 8, 93-102.
- Boyack, K. W., Klavans, R. & Börner, K. (2005). Mapping the backbone of science. *Scientometrics*, 64, 351-374.
- Deutscher Bildungsrat (1974). Aspekte für die Planung der Bildungsforschung (Empfehlungen der Bildungskommission). Bonn: Dt. Bildungsrat.
- Fry, J. et al. (2009). Communicating knowledge: how and why UK researchers publish and disseminate their findings. A Research Information Network report, September 2009. London: The Research Information Network. 47 S.
- Glänzel, W. & Schubert, A. (2003). A new classification scheme of science fields and subfields designed for scientometric evaluation purposes. *Scientometrics*, 56 (3), 357-367.
- Gómez, I., Bordons, M., Fernández, M. T. & Méndez, A. (1996). Coping with the problem of subject classification diversity. *Scientometrics*, 35 (2), 223-235.
- Goodyear et al. (2009). The Intellectual Foundations of Education: Core Journals and Their Impacts on Scholarship and Practice. *Educational Researcher*, 38 (9), 700-706.
- Hicks, D. (2004). The Four Literatures Of Social Science. In H.F. Moed, W. Glänzel & U. Schmoch (Ed.). *Handbook of Quantitative Social Science and Technology Research*. Kluwer Academic, S.473-496.
- Laurens, P., Zitt, M. & Bassecouard, E. (2010). Delineation of the genomics field by hybrid citation-lexical methods: interaction with experts and validation process. *Scientometrics* (2010) 82:647–662. DOI 10.1007/s11192-010-0177-9.
- Leydesdorff, L. (1987). Various methods for the mapping of science. *Scientometrics*, 11, 291-320.
- Leydesdorff, L. (2002). Dynamic and evolutionary updates of classificatory schemes in scientific journal structures. *Journal of the American Society for Information Science and Technology*, 53 (12), 987-994.
- Leydesdorff, L. (2004). Clusters and maps of science journals based on bi-connected graphs in the Journal Citation Reports. *Journal of Documentation*, 60(4), 371-427.
- Leydesdorff, L. & Rafols, I. (2009). A global map of science based on the ISI subject categories. *Journal of the American Society for Information Science and Technology*, 60(2), 348-362.
- Persson, O. (1994). The Intellectual Base and Research Fronts of JASIS 1986-1990. *Journal of the American Society for Information Science*, 45(1), 31-38.
- Pudovkin, A. I. & Garfield, E. (2002). Algorithmic procedure for finding semantically related journals. *Journal of the American Society for Information Science and Technology*, 53 (13), 1113-1119.
- Rafols, I. & Leydesdorff, L. (2009). Content-based and algorithmic classifications of journals: Perspectives on the dynamics of scientific communication and indexer effects. *Journal of the American Society for Information Science and Technology*, 60(9), 1823-1835.
- Waltman, L. et al. (2011). Towards a new crown indicator: Some theoretical considerations. *Journal of Informetrics* 5 (1), 37-47.
- Zitt, M. & Bassecouard, E. (2006). Delineating complex scientific fields by an hybrid lexical-citation method: An application to nanosciences. *Information Processing and Management* 42, 1513–1531.