

Feldabgrenzung in einem multidisziplinären Feld am Beispiel der Bildungsforschung

Bibliometrie2014 – 2. Internationale Konferenz und Fachmesse
Regensburg

Referentin: Valeria Aman

Inhalte des Vortrags

- **Feldabgrenzung**
 - Definition
 - klassisches Vorgehen („Makro-Level“)
 - „lex+cite“-Methode
 - am Beispiel der Bildungsforschung
 - Vorteile gegenüber der „Makro-Level“-Abgrenzung

Feldabgrenzung ...

- ...umschreibt die Extraktion von Publikationen aus einer Datenbank, die idealerweise die gesamte relevante Literatur in einem Forschungsfeld repräsentieren
- ...ist ein komplexes Unterfangen ohne allgemeingültige Lösung
- ...wird daher als ein eigenständiges Forschungsfeld betrachtet (Bassecoulard, Lelu & Zitt 2007)

„Makro-Level“-Feldabgrenzung

- Felddefinition erfolgt auf Basis der > 250 Kategorien, denen Zeitschriften in Web of Science (WoS) zugeordnet werden
 - Zeitschriften können genau einer oder bis zu acht dieser sogenannten Subject Categories (SC) zugeordnet werden
 - Fuzziness resultiert aus inkorrektter Zuordnung oder der Mehrfachzuordnung von Zeitschriften
- Klassifikationssystem in WoS bietet eine konsistente Felddefinition und kann für automatisierte, umfangreiche Prozeduren verwendet werden

Grenzen der „Makro-Level“-Feldabgrenzung

- ein Feld ausschließlich anhand von SC abzugrenzen, kann potenziell relevante Artikel ausschließen
 - die in Zeitschriften erschienen sind ohne den gewünschten SC
 - die in multidisziplinären Zeitschriften erschienen sind
- Leydesdorff (2008) merkt an, dass Thomson Reuters die Kategorien auf Basis von Zeitschriften und deren Zitationsmustern vergibt und nicht auf intellektueller Basis
 - in knapp 50% der Fälle korrespondieren die SC nicht mit den Clustern, die aus den Zitationsbeziehungen zwischen den Zeitschriften resultieren

„Makro-Level“-Feldabgrenzung am Beispiel der Bildungsforschung

- relevante ***Subject Categories*** (SC) in **WoS** (von > 250)
 - *Education & Educational Research*
 - *Education, Scientific Disciplines*
 - *Education, Special*
 - *Psychology, Educational*

- relevante ***Descriptions*** in **Scopus** (von insges. 322)
 - *Education*
 - *Developmental and Educational Psychology*

Feldabgrenzung – Schwierigkeiten bei einem multidisziplinären Feld

- wenige Kernzeitschriften
 - Publikationen streuen über diverse Zeitschriften
 - es gibt wenige fachspezifische Terme als in anderen Forschungsfeldern
 - neue Bereiche können entstehen ohne definiertes Vokabular
 - Klassifikationssysteme sind nicht dynamisch genug
- ➔ Abkehr von einer „Makro-Level“-Feldabgrenzung

„Mikro-Level“- Feldabgrenzung

- basiert auf der „**lex+cite**“ Methode, die von Zitt und Bassecoulard (2006) stammt
- geeignete Methode, um komplexe, wachsende und interdisziplinäre Felder abzugrenzen
- „lex+cite“ Methode wird iterativ und unter Mitwirkung von Experten angewandt
- stützt sich auf verschiedene bibliografische Informationen:
 - journalbasierte
 - lexikalische
 - zitationsbasierte

Feldabgrenzung - journalbasiert

- im ersten Schritt empfehlen Zitt und Bassecouard (2006) eine Auswahl an Kernzeitschriften
- besser ist die Definition eines breiteres Sets an Zeitschriften, das im Folgenden verfeinert wird
- hierbei wird auf existierendes Expertenwissen zurückgegriffen

Feldabgrenzung – journalbasiert am Beispiel der Bildungsforschung

- verwendete Journalklassifikationen:
 - **ERIH** (*European Reference Index for the Humanities*)
 - **531** Journals
 - *Pedagogical and Educational Research*
 - **SCImago**
 - **573** Journals mit Stand Juli '13
 - *Social Sciences – Education*
 - **Science-Metrix**
 - **284** Journals
 - *Economic & Social Sciences – Social Sciences – Education*

Feldabgrenzung – journalbasiert am Beispiel der Bildungsforschung

- von den insgesamt 1.388 Zeitschriften sind 1.041 distinkt
- in WoS sind 391 dieser Zeitschriften nachweisbar
- Aggregation aller Publikationen für den Zeitraum 1995-2012 führt zu 228.265 Publikationen
- Experten am DIPF befinden **56%** der Artikel für falsch-positiv

Feldabgrenzung – journalbasiert am Beispiel der Bildungsforschung

- bei dem *Noise* handelt es sich u.a. um
 - Lehrmaterialien (z.B. Werke Shakespeares, Vulkane, etc.)
 - Anleitungen für den Unterricht (z.B. Bedienung eines Bunsenbrenners)
 - Ursache liegt in dem Wort „Education“, das im Englischen vermehrt im Gesundheitswesen gebraucht wird
 - *AIDS education = AIDS-Aufklärung*
- *Noise* kann mit dem lexikalischen Ansatz reduziert werden

Feldabgrenzung – lexikalischer Ansatz

- dient dazu, möglichst alle Aspekte eines Feldes abzubilden, ohne *Noise* zu erhöhen
- WoS bietet kein kontrolliertes Vokabular
 - natürliche Sprache verfügt über Synonyme und Homonyme, sodass eine intellektuelle Prüfung erforderlich bleibt
 - Homonyme sind schwer zu disambiguieren und erhöhen den *Recall*
- Aufbereitung der Suchterme
 - Berücksichtigung von Flexionen, Akronymen, Homonymen, ...

Feldabgrenzung – lexikalischer Ansatz am Beispiel der Bildungsforschung

- in ERIC (*Education Resources Information Center*) sind mehr als 700 relevante Terme unter den folgenden 4 Thesauri zu Bildungsstufen und -prozessen gelistet
 - Educational Levels, Degrees, and Organizations
 - Educational Process: Classroom Perspectives
 - Educational Process: School Perspectives
 - Educational Process: Societal Perspectives
- ➔ im lexikalischen Ansatz ist *Silence* wahrscheinlicher als *Noise*
- weitere Suchterme wurden aus deutschen Publikationen, die Bildungsforschung zum Gegenstand haben, generiert

Feldabgrenzung – zitationsbasierter Ansatz

- dient dazu *Silence* zu reduzieren
 - Nachteil des lexikalischen Ansatzes kann mangelnder *Recall* sein
 - interdisziplinäre Felder können überwiegend alltagssprachliches Vokabular aufweisen, das nicht eindeutig dem Feld zuzuordnen ist
- innerhalb eines Referenzkorpus, der zunächst journalbasiert und dann lexikalisch abgegrenzt ist, wird ein sog. **Seed** definiert
- **Seed** = Kern an Publikationen, die zu 100% das Forschungsfeld zum Gegenstand haben

Feldabgrenzung – zitationsbasierter Ansatz

- Seed wird innerhalb des Referenzkorpus um Publikationen erweitert, die sich idealerweise auf die gleiche intellektuelle Basis stützen
- Publikationen, die über den zitationsbasierten Ansatz erfasst werden, sind:
 - **Cited Items** (vom Seed zitierte Publikationen)
 - **Citing Items** (den Seed zitierende Publikationen)
 - **Siblings** (bibliographisch gekoppelte Items, d.h. Items zitieren dieselben Publikationen wie der Seed)

Feldabgrenzung – Visualisierung

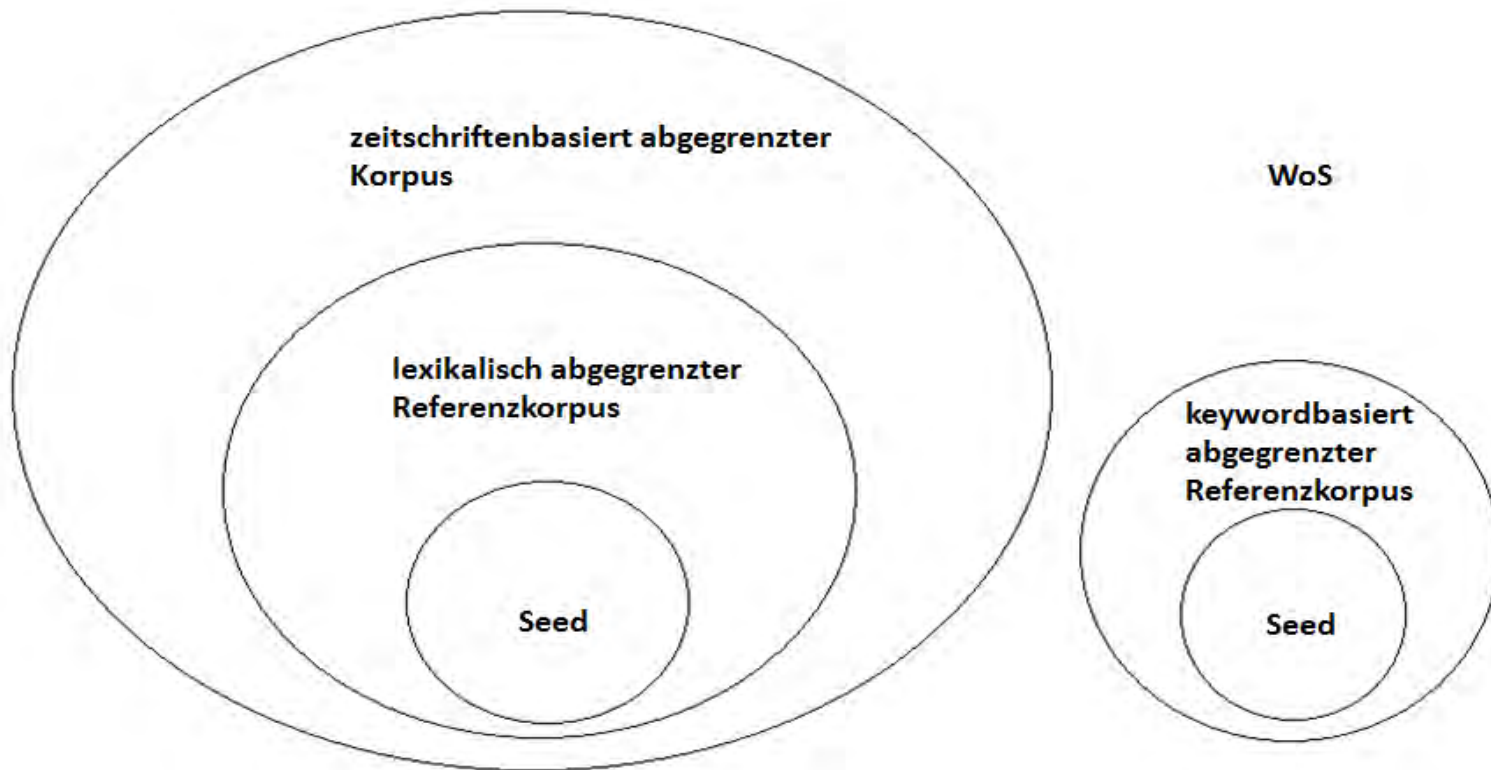


Abbildung 1: Relation der zwei unabhängig voneinander definierten Korpora und deren Seeds in WoS. Quelle: Eigene Darstellung des iFQ.

Feldabgrenzung – zitationsbasierter Ansatz

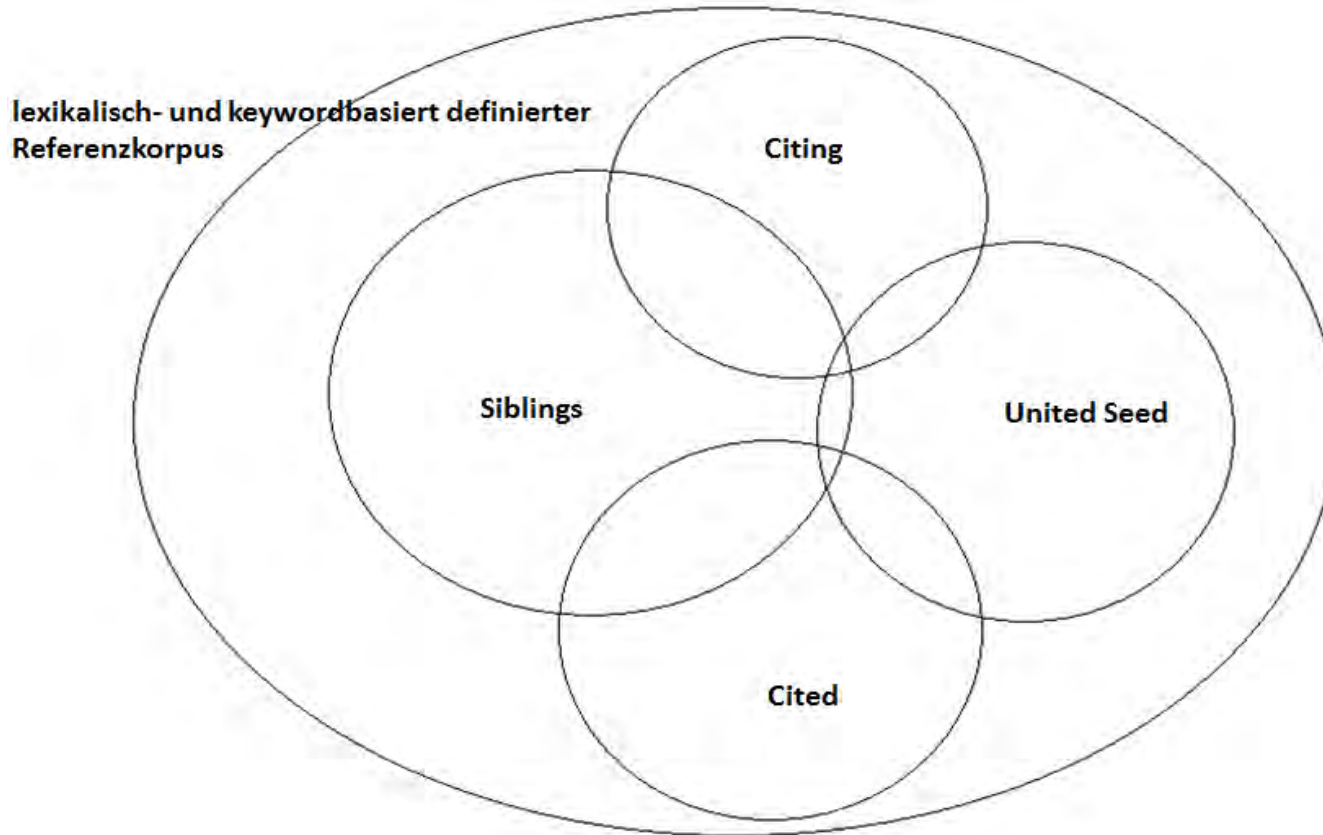


Abbildung 2: Relation der zitationsbasierten Recherche innerhalb des vereinten Korpus und des United Seeds. Quelle: Eigene Darstellung des iFQ.

Feldabgrenzung – Resultat

- insgesamt eine heuristische Suche mit gutem Ergebnis
 - Fokus liegt eher auf *Precision* als auf *Recall*
 - für die angestrebte *Precision Rate* wurden einzelne Schritte iteriert
 - Experten am DIPF haben kontinuierlich Stichproben ausgewertet
- Abgrenzung der Bildungsforschung in WoS führte zu
 - **129.933** Publikationen (1995-2012)
 - *Precision Rate*: **97%**
- Vorteil der hybriden Feldabgrenzung:
 - Überführung auf Scopus

Feldabgrenzung – Scopus

- Abgrenzung der Bildungsforschung in Scopus führte zu
 - **147.160** Publikationen (1996-2012)
 - *Precision Rate: 98%*
- die finalen Korpora sind nicht identisch, da die Feldabgrenzung nicht genuin in Scopus erfolgt ist
- Gründe für Unterschiede zwischen WoS und Scopus
 - verschiedene Journalpolicy
 - Data accuracy (Vollständigkeit, Vergabe von Keywords, ...)
 - verschiedene Zitationszahlen und Zitationsbeziehungen

Verwendete Literatur

- 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: E. Torres-Salinas, H.F. Moed (Eds), 11th International Conference on Scientometrics and informetrics (ISSI 2007), Madrid, Spain, 25-27 June 2007, 74-84.
- 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. (2008) Caveats for the Use of Citation Indicators in Research and Journal Evaluations. *Journal of the American Society for Information Science and Technology*, 59(2): 278-287.
- 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.

Herzlichen Dank für das Interesse!

Fragen, Unklarheiten, Anmerkungen?