

Google über alles? Wissenschaftssuchmaschinen als alternativer Weg zu relevanten Fachpublikationen in der Landwirtschaft

Hubert Schlieber¹

Abstract - Freier und kostenloser Zugang zum Wissen der Welt für alle? Erhalten wir diesen im Internetzeitalter nicht ohnehin unkompliziert und umfassend, wenn wir eine Suchmaschine wie *Google* einsetzen? Doch die Anzahl der indextierten Webseiten des Marktführers *Google* wächst exponentiell an und erreicht bereits Zahlen in Milliardenhöhe. Damit steigt aber auch der darin enthaltene *Datenmüll*, der es immer schwieriger macht, wissenschaftliche und in diesem Falle landwirtschaftliche Fachinformation schnell, relevant und zuverlässig aufzufinden. Wer als Student oder Lehrperson kostenfreien Zugang zu lizenzpflichtigen Zeitschriftenvolltext- und Faktendatenbanken an der Universitätsbibliothek hat, der hat den Vorteil, gezielt Fachrecherchen nach bekannten, bewährten und bewerteten Informationsquellen durchführen zu können. Was machen aber Forscher außerhalb der Hochschul-Community ohne Zugriff auf universitäre Onlineresourcen und mit geringen finanziellen Mitteln? Ein gangbarer Weg zur Abhilfe besteht in speziell entwickelten *Wissenschaftssuchmaschinen*, die im Gegensatz zu *Google* mit ausführlichen Filter- und Sortiermechanismen ausgestattet sind und überdies Inhalte des *Deep Web* sowie kostenlose Publikationen der *Open-Access-Bewegung* aufspüren können. Anhand einiger zentraler landwirtschaftlicher Themen (Biolandbau, Klimawandel, etc.) werden in der vorliegenden Arbeit die derzeit bedeutendsten Wissenschaftssuchmaschinen (*BASE*, *Google Scholar*, *Scirus*, etc.) auf ihre Nützlichkeit und Verlässlichkeit für die wissenschaftliche Fachrecherche - und damit auch auf ihre Einsatzfähigkeit als Alternative zu *Google* - untersucht.

AUSGANGSSITUATION

Ausgangspunkt der im Frühjahr 2011 vorgenommenen Untersuchung war *einerseits* die zusehends abnehmende Brauchbarkeit einer minimalistischen Universalsuchmaschine wie *Google* für die gezielte wissenschaftliche Recherche nach landwirtschaftlicher Fachinformation (Zeitschriftenartikel, Forschungsreports, Agrarstatistiken, etc.) durch exponentiell anwachsenden Datenmüll (Müller, 2011), *andererseits* die kostenlose Bereitstellung von hochwertigen Datenbanken und Informationsnetzen (*Web of Knowledge*, *Scopus*, *SpringerLink*, etc. für die Nutzung von e-journals und e-books) *ausschließlich* für Benutzer von Groß- bzw. Universitätsbibliotheken.

Was macht nun aber Suchmaschinen (SM) wie *Google*, *Yahoo*, *Bing* etc. so ungeeignet für wissenschaftliches Arbeiten? Der Grund besteht darin, dass die Indices von Suchmaschinen großteils auf unstrukturiertem, unsegmentiertem und damit nicht näher zu spezifizierendem Inhalt aufsetzen, wobei neben qualitätvollen Texten jede Menge Datenschnitt das Suchergebnis verwässert. Die ungeprüften, unrevidierten und vielfach unvollständigen (Strzolka, 2008; Lewandowski, 2007a) Suchergebnisse bzw. Texte lassen selten einen Aufschluss über Qualität, Verlässlichkeit und Erstellungsdatum des Inhalts zu, auch die Autorenschaft lässt sich nicht immer zweifelsfrei feststellen (Lewandowski, 2007b) - ein wesentliches Kriterium zur Einschätzung der Artikelqualität. Es fehlen in den allermeisten Fällen also die *Metadaten* (Hermann, 2010) zu einer Publikation. Überdies sind die Möglichkeiten, die Suche vorab mit vielfältig wählbaren Filterkriterien einzuschränken bzw. das Suchergebnis in brauchbarer Weise zu sortieren, bei den eingangs erwähnten Universalsuchmaschinen nicht oder nur sehr ansatzweise vorhanden.

Was machen nun Wissenschaftssuchmaschinen (Wiss-SM) besser? Der einprägsamste Unterschied zu Universalsuchmaschinen besteht darin, dass man hierbei vom Universalitätsanspruch abrückt und eine bewusste Auswahl von wissenschaftlich relevanten und qualitätvollen Quellen trifft. Die bereitgestellten Daten *sollen* Metadaten (z.B.: Autor, Erstellungsdatum, Keywords, Sprache, etc.) in einem genormten Datenformat aufweisen. Vielfach konzentrieren sich Wissenschaftssuchmaschinen auf bestimmte Themen, wie etwa *AgEcon*² auf agrarökonomische Literatur.

Wissenschaftlich relevante Inhalte (Dissertationen, Forschungsberichte, Fachartikel etc.) befinden sich immer öfter auf Dokumentenservern (Repositories³) im Hochschulbereich, auf persönlichen Webseiten von Forschern oder als Artikel in einer elektronischen Zeitschrift (e-journal). Vielfach kann auf diese mit Metadaten angereicherten Inhalte via *Open Access* kostenlos zugegriffen werden. Das zentrale Merkmal von *Open Access* (OA) ist, dass öffentlich gemachte und mit öffentlichen Mitteln produzierte wissenschaftliche Literatur für jeden

¹ H. Schlieber arbeitet an der Bundesanstalt für Agrarwirtschaft und leitet das Referat Bibliothek und Dokumentation, Wien, Österreich (hubert.schlieber@awi.bmlfuw.gv.at).

² <http://ageconsearch.umn.edu/>

³ Repository: Dokumentenserver, ursprüngl. Quelle, Speicher, Lagerort, Verwahrungsort

kostenfrei und öffentlich im Internet zugänglich gemacht wird, damit Interessierte die Volltexte lesen, herunterladen, kopieren, verteilen, drucken, in ihnen suchen, auf sie verweisen und sie auch sonst auf jede denkbare legale Weise benutzen können⁴.

Des Weiteren ist anzumerken, dass hochwertige Fachinformation wie sie etwa im Hochschul- und Forschungsbereich zu finden ist, häufig im Bereich des *Deep Web* oder *Invisible Web* abgelegt und daher für Universalsuchmaschinen nicht erreichbar ist (Lewandowski, 2007b).

STRUKTUR- UND LEISTUNGSVERGLEICH AUSGEWÄHLTER WISSENSCHAFTSSUCHMASCHINEN

Zur besseren Vergleichbarkeit wurden vom Autor wie schon bei einer Vorstudie (Schlieber, 2008) bei jeder Wiss-SM die gleichen, landwirtschaftlich brisanten Themen sowohl in deutsch als auch in englisch abgefragt: Biologische Landwirtschaft, LW und Klimawandel, GAP-Reform, Risikomanagement, Nachhaltigkeit, Gentechnik, Agrarförderung, Nachwachsende Rohstoffe (Nawaro), Bioenergie, Biomasse. Die Themenvorschläge wurden von MitarbeiterInnen der Bundesanstalt für Agrarwirtschaft eingebracht.

Es lassen sich drei Typen von Wiss-SM unterscheiden:

1. SM speziell für Open-Access Inhalte⁵
2. SM für OA und kommerzielle Angebote vermischt⁶
3. Repositories mit Suchoberflächen⁷

Gruppe 1: BASE, MS Academic Search, Scientific Commons, Cite Seer, Open-J-Gate, OpenDOAR, ROAR, DOAJ, OAISTER, SpringerLink OA

Gruppe 2: Scirus, Google Scholar

Gruppe 3: ScienceGate, AgEcon, Organic eprints, EDOC

Abgesehen von einer möglichst hohen Trefferanzahl pro Abfrage nach den oben angeführten Themen, was für sich allein genommen zu wenig aussagekräftig ist, stand das Vorhandensein folgender Merkmale der Wiss-SM im Fokus der Untersuchung:

- ergebniseinschränkende (=weniger, aber präzisere Treffer): Suche nach Metadaten (Autor, Jahr, Keyword,...), Filter und Sortieroptionen *vor* und *nach* Suche, logisches UND
- ergebniserweiternde (=mehr, aber weitläufigere Treffer): Volltextsuche, Stemming⁸, themengeführte Suche, Trunkierung, logisches ODER
- Sonstige: Quellenangabe, Anzahl aller vorhandenen Medien, RSS-Abo, Suchhistorie nachvollziehbar, Export von Suchergebnissen, Trefferaktualität

Die Suchergebnisse sowie die allenfalls vorhandenen Merkmale wurden mit MS Excel in eine *Matrix* eingetragen, die mehrfache Auswertungen erlaubt.

ERGEBNISSE UND AUSBLICK

Was die Trefferanzahl insgesamt anbelangt, hatte SCIRUS deutlich die Oberhand (> 6 Mio. Titel), gefolgt von GOOGLE SCHOLAR und BASE, wobei letztere Volltextsuche erlauben, was der Präzision des Ergebnisses nicht immer zuträglich ist. Obwohl jede Wiss-SM etwas andere thematische Schwerpunkte setzt, zeichnete sich doch ein Trend ab: Titel zu Nawaros wurden praktisch überall nachgewiesen, gefolgt von Nachhaltigkeit, Biolandbau und Klimawandel – ein Ergebnis, das dem realen Geschehen in Umwelt und Landwirtschaft Rechnung trägt.

Betrachtet man die themenunabhängige Leistung bzw. die erweiterten Recherchemöglichkeiten so dominiert BASE vor OAISTER und SCIRUS, wobei die Bool'schen Operationen UND/ODER sowie die Suche in Metadatenfeldern bei *allen* Wiss-SM aufzufinden waren.

Wenn auch die Untersuchung hier nur stark verkürzt dargestellt dargelegt werden konnte, hoffe ich doch, neugierig auf die ständig in Weiterentwicklung begriffenen Wiss-SM gemacht und auf eine bedenkenswerte Alternative zu Google hingewiesen zu haben. Überdies war es dem Autor ein Anliegen, auf die wachsende Anzahl von qualitativen und bewerteten Fachpublikationen hinzuweisen, die im Zuge der Open-Access-Bewegung zusehends für die (Agrar-) Forschung bereitgestellt werden – ein Aspekt, den man in Zeiten allgemeiner Budgetkürzungen in privaten und öffentlichen Forschungseinrichtungen nicht zu gering einschätzen sollte.

LITERATUR

Hermann, L. (2019). Möglichkeiten und Grenzen von Suchmaschinen bei der wissenschaftlichen Recherche im Internet : Magisterarbeit. - Berlin : Humboldt - Univ. Berlin. Online: <http://edoc.hu-berlin.de/master/hermann-lars-2010-01-28/PDF/hermann.pdf>

Lewandowski, D. (2007a). Deep Web - Suchen wir was man finden könnte? Online: http://www.durchdenken.de/lewandowski/doc/Euroforum_Lewandowski.pdf

Lewandowski, D. (2007b). Google, Deep Web und Fachdatenbanken. Online: <http://www.durchdenken.de/lewandowski/doc/swp2007.pdf>

Müller, D. (2011). Das Web versinkt im Datenmüll. Blogbeitrag.- <http://tinyurl.com/6dz9992>

Schlieber, H. (2008). Der freie (kostenlose) Zugang zu Publikationen aus Wissenschaft und Bildung im Internet über Open Access (OA) : Schwerpunkt Landwirtschaft ; August 2008. - Wien : AWI http://eprints.rclis.org/14781/1/Open_Access.pdf

Strzalka, R. (2008). Das Internet als Weltbibliothek: Suchmaschinen und ihre Bedeutung für den Wissenserwerb. - Berlin : Simon Verl. f. Bibliothekswissen.

⁴ http://open-access.net/at_de/allgemeines/was_bedeutet_open_access/

⁵ (bieten keinen eigenen Inhalt an, suchen über OA-Inhalte und/oder lizenzierte Inhalte)

⁶ S. Fussnote 5

⁷ bieten eigenen Inhalt an, meist Hochschul-Dokumentenserver

⁸ (Grundformenreduktion, Normalformenreduktion)