

VII Wissenschaftliche Suchmaschinen

Für Bereiche der wissenschaftlichen Literatur- und Informationsrecherche haben sich einige spezialisierte Suchmaschinen etabliert. Vor allem zwei Arten sind zu unterscheiden: in einen Fall handelt es sich um konventionelle Suchmaschinenteknologie, der Suchraum wird jedoch so eingeschränkt, dass nur relevante Daten indexiert werden. Im anderen Fall werden neue, semantische Technologien eingesetzt, um bessere Suchergebnisse zu erzielen. Auch eine Kombination aus beiden Methoden ist möglich. Bei einem Teil dieser Suchmaschinen basiert die Kommunikation zwischen dem Datenprovider und der Suchmaschine auf dem *Protocol for Metadata Harvesting* der *Open Archives Initiative* (OAI-PMH), das als einfacher, offener Standard von vielen Informationssystemen im wissenschaftlichen Bereich unterstützt wird. Der eingeschränkte Suchraum von wissenschaftlichen Suchmaschinen kann unterschiedlich definiert werden, z. B. bezogen auf Server einer bestimmten Art (z. B. Open-Access-Repositorien), auf Institutionen einer bestimmten Art (z. B. Hochschulserver) oder auf Dokumente eines bestimmten Typs (z. B. *Google Patent Search* für Patente). Er kann darüber hinaus auch fachbezogen sein, indem nur Server und Quellen dieses Fachgebietes ausgewertet werden. Beispiele dafür sind *CiteSeerX* für Informatik oder *BizNar* für die Wirtschaftswissenschaften.

Bei den im Folgenden vorgestellten Suchmaschinen, *Google Scholar*, *BASE*, *Semantic Scholar*, *Science Research* und *WorldWideScience.org* liegt der Fokus auf der Suche nach interdisziplinärer wissenschaftlicher Literatur. Neben der Beschränkung des Suchraums auf die Anbieter wissenschaftlicher Informationen bieten Suchdienste vielfach noch weitere, spezielle Vorteile; so ermöglicht *Google Scholar* eine Zitationsanalyse der indexierten Inhalte und verbindet kostenpflichtige Angebote mit dem Bestand von lizenzierenden Bibliotheken, *Semantic Scholar* erlaubt neben der Zitationsanalyse auch die semantische Analyse von Suchanfragen und Indexdaten, *Science Research* bietet Bibliotheken die Möglichkeit, individuell definierte Datenbestände in den Suchraum zu integrieren und *WorldWideScience.org* arbeitet mit multilingualen Suchzugriffen und mehrsprachigen Thesauri.

Titel	Google Scholar
Urheber	Google
Typ	Wissenschaftliche Suchmaschine mit Zitationsanalyse
Inhalt	Die Suchmaschine indexiert kostenfreie und kostenpflichtige digitale Dokumente einschließlich der Literaturangaben in diesen Dokumenten (Zitate). Überwiegend handelt es sich bei den indexierten Quellen um Repositorien von Wissenschaftsorganisationen und wissenschaft-

liche Verlage. Das Quellenverzeichnis ist nicht einsehbar; auch die Gesamtzahl der indexierten Dokumente wird nicht angegeben. Bei dem Material handelt es sich überwiegend um Zeitschriftenaufsätze, ebenfalls enthalten sind jedoch auch Bücher, Abschlussarbeiten, Präsentationen, Poster und andere Formen wissenschaftlicher Veröffentlichungen.

Beschr. Die Besonderheit dieser Suchmaschine liegt neben der Volltextindexierung in der Auswertung der Zitate. Die Erweiterte Suche bietet die Unterscheidung zwischen der Suche in den Volltexten bzw. den Metadaten, eine Autorensuche, eine Suche nach Zeitschriften und Quellen sowie eine zeitliche Einschränkung. Für die Facettierung der Trefferlisten stehen weitere Filter zur Verfügung. Auf Open-Access-Publikationen wird direkt verlinkt, Closed-Access-Publikationen sind ggf. über die Linking-Services von Bibliotheken zu erreichen. Treffer lassen sich markieren, bibliographische Daten können in verschiedenen Formaten und Styles exportiert werden. – Die Zitationsanalyse bietet die Möglichkeit der Anzeige aller Zitate eines Titels sowie seine Zitationen in späteren Arbeiten; ausgehend von diesen Daten werden auch weitere, thematisch vergleichbare Dokumente angezeigt. – Existieren von einer Publikation verschiedene Versionen, so können alle angezeigt werden, unterschiedliche Zugangsmöglichkeiten können so genutzt werden. – Vielfältig sind die Möglichkeiten von Google Scholar für angemeldete Nutzer. So lassen sich persönliche Einstellungen vornehmen, u. a. in Bezug auf die Anzahl und den Umfang der angezeigten Treffer, die Sprachen, die bei der Suche berücksichtigt werden sollen, das bevorzugte Format für den Datenexport und – sicherlich die wichtigste Funktion – die „Bibliothekslinks“. Über diese Funktion können Nutzer Bibliotheken definieren, in denen sie angemeldet sind. Auf diese Weise können sie nicht nur das Material im Volltext nutzen, das open access angeboten wird, sondern auch die dort lizenzierten Dokumente. Auch Alerting-Dienste zu einzelnen Suchanfragen können eingerichtet werden. Weitere personalisierte Dienste sind „Meine Bibliothek“ (dort lassen sich interessante Treffer dauerhaft abspeichern) und „Mein Profil“ für Autoren; hier findet sich ein Überblick der eigenen Publikationen mit Zitationsbeziehungen, Metriken und Interessensgebieten. Aufgrund des Profils werden u. a. interessante neue Publikationen und mögliche Koautoren vorgeschlagen. Für die Nutzung weiterer Dienste und die Vernetzung kann das eigene Profil auch veröffentlicht werden.

Bezieh. Für Chrome stehen verschiedene Dienste von Google Scholar auch als Browser-Plugin zur Verfügung. – Bei *BASE* wird von den Ergebnissen zu Google Scholar verlinkt. – *Google Books* und Google Scholar sind inzwischen eng verknüpft, durch die Zusammenarbeit mit Verlagen sind vielfach die gleichen Bücher auch in Google Scholar enthalten.

[ZITATION] **Library and information science** text for Africa

LO Aina - 2004 - Third World Information Services

☆ ⓘ Zitiert von: 743 Ähnliche Artikel In RefWorks importieren ⓘ ⓘ

Information literacy at school level: A comparative study between the Netherlands and South Africa

AK Boekhorst, JJ Britz - ... journal of libraries and information science, 2004 - journals.co.za

... **Information Poverty and Social Justice**, South African Journal of **Library and Information Science**, 67 (2): 63-69, Dervin, B., & Nilan, M. 1986 ... South African Journal of **Library and Information Science**, 60, (1):39-46 Webster, F. 2002. Theories of the **Information Society** ...

☆ ⓘ Zitiert von: 57 Ähnliche Artikel Alle 7 Versionen In RefWorks importieren SUB Göttingen Link (SFX) Linkingservices TIB ⓘ ⓘ

[PDF] journals.co.za

SFX@LMU

[DOC] **Relevance reconsidered**

T Saracevic - ... of library and information science (CoLIS ..., 1996 - tefkos.cominfo.rutgers.edu

The paper is a critical review of the progress in thinking about the nature of relevance in **information science**. To a lesser extent, studies dealing with manifestations of relevance are reviewed as well. Four frameworks about nature of relevance emerged over time: systems ...

☆ ⓘ Zitiert von: 734 Ähnliche Artikel Alle 7 Versionen In RefWorks importieren HTML-Version ⓘ ⓘ

[DOC] rutgers.edu

Abb. 122: Trefferanzeige in Google Scholar mit direkten Links zu weiteren Angeboten und Informationen (unter dem Titel) sowie Links zum jeweiligen Volltext (rechts)

Titel **BASE – Bielefeld Academic Search Engine**

Urheber Universitätsbibliothek Bielefeld

Typ Suchmaschine für Dokumente auf akademischen Servern

Inhalt Indexiert werden Ressourcen aller Fachgebiete; berücksichtigt werden rund 10 000 Quellen, überwiegend institutionelle Repositorien (meist von Hochschulen) und Angebote von wissenschaftlichen Verlagen. Das Quellenverzeichnis ist einsehbar. Insgesamt werden rund 320 Millionen Dokumente in die Suche einbezogen. Mehr als ein Drittel der indexierten Dokumente sind Open-Access-Publikationen.

Beschr. Die Suchmöglichkeiten in BASE sind überaus vielfältig. Dies wird besonders deutlich in der Erweiterten Suche. Dort werden nicht nur zahlreiche Suchkriterien angeboten (u. a. Autor, Titel, Schlagwörter, Verlag, Erscheinungsjahr, DOI und ORCID-ID), zur Verfügung steht auch eine Volltextsuche, allerdings werden hierbei vielfach nur die Metadaten und die Abstracts berücksichtigt, da nur ein kleinerer Teil der Dokumente vollständig indexiert ist. Open-Access-Dokumente können bevorzugt angezeigt werden, auch die gewünschte Nutzungs-

art (Lizenzierung) kann ausgewählt werden. Bei den Dokumentarten wird unter anderem unterschieden zwischen Büchern, Artikeln aus Zeitschriften und Zeitungen, Konferenzbeiträgen, Reports und Rezensionen, Dissertationen und anderen Abschlussarbeiten, Patenten, Noten, Karten, Forschungsdaten, Softwareprodukten, Bildern und Videos. Verschiedene Wortformen können in die Suche einbezogen werden. Über die Aktivierung der mehrsprachigen Suche lassen sich auch Übersetzungen der Suchbegriffe in 25 weiteren Sprachen finden. Ebenfalls eingebunden sind mit dem *EuroVoc Thesaurus* und dem *Agrovoc Thesaurus* zwei mehrsprachige Thesauri. Die Trefferlisten können auf bestimmte Länder, Erscheinungsjahre oder auch Lizenzierungen begrenzt werden. Weitere Recherchemöglichkeiten bietet das Browsing. Hier lassen sich die Dokumente über die Dewey-Dezimalklassifikation (DDC) auswählen (allerdings ist nur ein Teil der Dokumente entsprechend klassifiziert), über die Dokumentenart, die Lizenzierung sowie über die Möglichkeiten des Zugangs. – In der Ergebnisliste werden formale und inhaltliche Filter angeboten. Bei jeder Kategorie ist die Anzahl der Ergebnisse aufgeführt. Besonders aufschlussreich ist diese Angabe für die Herkunft der Treffer (Quellen). Datensätze lassen sich exportieren, Suchanfragen können über verschiedene Feeds abonniert werden; angemeldete Nutzer können Suchergebnisse und -anfragen auch dauerhaft abspeichern. Gegebenenfalls vorhandene weitere Versionen (Preprints, Sekundärveröffentlichungen etc. und Angebote von anderen Plattformen) werden ebenfalls angezeigt.

Bezieh. Von einem Metadatensatz aus kann zur Anzeige des Treffers in *Google Scholar* gewechselt werden, wo auch die Zitationsanalyse zur Verfügung steht. – BASE wurde im bibliothekarischen Umfeld entwickelt, sein Index ist in viele andere Informationsressourcen einbezogen, insbesondere in Kataloge und Fachinformationsressourcen, daneben in Meta-Suchportale wie dem *Karlsruher Virtuellen Katalog*; auch über den weltweit verbreiteten *EBSCO Discovery Service* können die Daten von BASE direkt im Bibliothekskatalog genutzt werden.

Titel **Semantic Scholar**

Urheber Allen Institute for Artificial Intelligence (AI2)

Typ Wissenschaftliche Suchmaschine basierend auf Verfahren der künstlichen Intelligenz und der semantischen Analyse; Berücksichtigung von Zitationszusammenhängen

- Inhalt** Die Suchmaschine greift auf einen Suchraum von mehr als 200 Millionen Forschungspublikationen zu; berücksichtigt werden sowohl Open-Access- als auch Closed-Access-Publikationen.
- Beschr.** Um die Qualität der Suchergebnisse zu erhöhen, nutzt Semantic Scholar vielfältige Methoden der Sprachverarbeitung und des maschinellen Lernens, um in Dokumenten und Suchbegriffen relevante Beziehungen und Entitäten (Sachbegriffe, Personen, Institutionen, Orte und Daten) zu erkennen und diese für die Suche zu nutzen. Auch Abstracts, Graphiken, Zitate etc. werden automatisch erkannt und separat angezeigt. – Jeder einzelne Titeldatensatz bietet ein sehr umfangreiches Funktionsangebot: Auf die Volltexte wird verlinkt, Treffer lassen sich exportieren, in Sozialen Medien teilen und zitieren, angezeigt werden die Zitationen, die Zitate und relevante Treffer (zitationsbasierter Recommender-Dienst). Existieren zu einem Dokument auch ergänzende Forschungsdaten, Daten aus klinischen Studien, Präsentationen, Videos, wird auch auf diese verlinkt. – Umfangreiche Möglichkeiten bestehen für angemeldete Nutzer (u. a. Alerting-Dienste, Anlegen eigener Research Feeds). – Besonders interessant an Semantic Scholar ist die Funktion *TLDR* (Extreme Paper Summaries). *TLDR* steht für den Ausdruck „too long; didn’t read“ und bezeichnet ein auf künstlicher Intelligenz basiertes Verfahren, die Grundaussage eines umfangreichen Textes in einem knappen, natürlichsprachlichen Satz auszudrücken. Diese Funktion nutzt Semantic Scholar und zeigt bereits in der Kurztitelanzeige damit die Grundaussage der angezeigten Publikationen an. – Vielfältig sind auch die Möglichkeiten, die wichtigsten und einflussreichsten Arbeiten oben in der Trefferliste zu ranken (Sort by Relevance, Sort by Citation Count, Sort by Most Influential Paper, Sort by Recency).
- Bezieh.** Semantic Scholar greift nicht nur bei den indextierten Dokumenten (Publishing Partners), sondern auch bei den Indizes und Linkingdiensten (Data Partners) auf Angebote anderer Institutionen zurück. Besonders wichtig in diesem Zusammenhang sind die Dienste von *unpaywall* und *Microsoft*. Bis Ende 2021 betrieb Microsoft selbst die wissenschaftliche Suchmaschine *Microsoft Academic* mit einem Suchraum von 260 Millionen Dokumenten, die sich vor allem durch die inhaltliche Analyse der Suchbegriffe auszeichnete. Durch die semantische Inferenz wurden so die eingegebenen Suchbegriffe erkannten Entitäten zugewiesen. – Neben dem Projekt Semantic Scholar arbeitet das Allen Institute for Artificial Intelligence in verschiede-

nen Projekten, u. a. in den Bereichen Natural language processing (NLP), maschinellm Lesen und Verstehen und der visuellen Wissensextraktion; der Grundgedanke besteht jeweils darin, breitere und offene Anwendungen für künstliche Intelligenz zu schaffen. – Gemeinsam mit weiteren Projektpartnern wurde das Projekt CORD-19 (COVID-19 Open Research Dataset) entwickelt, das einen separaten Suchraum für Publikationen und Forschungsdaten zu COVID-19 bietet.

LSTM: A Search Space Odyssey
 Klaus Greff, R. Srivastava, J. Koutník, Bas R. Steunebrink, J. Schmidhuber · Computer Science, Medicine ·
 IEEE Transactions on Neural Networks and Learning... · 13 March 2015

TLDR We present the first large-scale analysis of eight **LSTM** variants on three representative tasks: speech recognition, handwriting recognition, and polyphonic music modeling. [Expand](#)

👍 2.553 PDF 📄 · 🔍 View on IEEE 📌 Save 🔔 Alert 🗨️ Cite 📈 Research Feed

Abb. 123: Semantic Scholar, Kurztitelanzeige mit TLDR-Funktion (maschinelle inhaltliche Zusammenfassung)

 SEMANTIC SCHOLAR Klaus Greff

DOI: 10.1109/tnnls.2016.2582924 · Corpus ID: 3356463

LSTM: A Search Space Odyssey
 Klaus Greff, R. Srivastava, +2 authors · J. Schmidhuber · Published 2017 · Computer Science, Medicine ·
 IEEE Transactions on Neural Networks and Learning Systems

Several variants of the long short-term memory (LSTM) architecture for recurrent neural networks have been proposed since its inception in 1995. In recent years, these networks have become the state-of-the-art models for a variety of machine learning problems. This has led to a renewed interest in understanding the role and utility of various computational components of typical LSTM variants. In this paper, we present the first large-scale analysis of eight LSTM variants on three representative... [Expand](#)

Share This Paper    

2.517 Citations
 Highly Influential Citations 218
 Background Citations 1.017
 Methods Citations 571
 Results Citations 46

Abstract Supplemental Code Figures and Topics Paper Mentions 2.517 Citations 64 References Related Papers

Abb. 124: Semantic Scholar, Trefferanzeige mit Zitationsanalyse

Titel	Science Research
Urheber	Deep Web Technologies
Typ	Föderative Suchplattform (Federated Search) unter Berücksichtigung verschiedener Datenbestände
Inhalt	Die frei zugängliche Suchmaschine Science Research bietet eine integrierte Suche über alle oder separat auswählbare Suchbereiche mit einem Schwerpunkt im Bereich der Naturwissenschaften. Individuali-

sierte Anwendungen einzelner Bibliotheken können diese Zusammenstellung um weitere, kostenpflichtige Suchräume erweitern.

Beschr. Die föderierte Suche bietet die Möglichkeit, verschiedene Datenquellen in eine Suchanfrage einzubeziehen, wobei die jeweiligen Suchergebnisse zusammengeführt, bewertet und dem Nutzer in einer integrierten Trefferanzeige mit einem gemeinsamen Ranking präsentiert werden (Explorit Everywhere). Die Möglichkeit, auf diese Weise verschiedene, definierte Suchräume anzusprechen, bietet Bibliotheken die Möglichkeit, die Suchoberfläche an ihre Bedürfnisse anzupassen und bei der Suche neben öffentlich zugänglichen Datenbeständen auch eigene und lokal lizenzierte Datenbestände zugänglich zu machen, auf die allgemeine und wissenschaftliche Suchmaschinen nicht zugreifen können (Deep Wep). Dies kann dazu führen, dass hochwertige, lizenzierte Inhalte von Bibliotheksnutzern stärker berücksichtigt werden. Die Erweiterte Suche von Science Research bietet vielfältige Recherchemöglichkeiten, zeitliche Einschränkungen und eine Eingrenzung der Suche auf frei zugängliche Publikationen. Trefferlisten lassen sich facetrieren, wobei die Clusterung auch graphisch dargestellt werden kann; Alerting-Dienste können eingerichtet werden.

Bezieh. Die Technik der föderierten Suche findet Anwendung in einer Vielzahl von individualisierten bibliothekarischen und fachspezifischen Lösungen (u. a. *BizNar* im Bereich der Wirtschaftswissenschaften oder *Mednar* im Bereich der Medizin). Eine weitere Anwendung bildet *WorldWideScience.org*.

Titel **WorldWideScience.org**

Urheber WorldWideScience Alliance, verantwortlich: U.S. Department of Energy's Office of Scientific and Technical Information

Typ Wissenschaftliche Suchmaschine mit Übersetzungstool; internationales Rechercheportal

Inhalt In die Suche einbezogen werden in diesem Fall nicht Tausende dezentrale Repositorien, Verlage und andere Quellen, sondern 105 große nationale und internationale Anbieter und Aggregatoren für wissenschaftliche Publikationen, aber auch für Multimedia-Objekte, Forschungsdaten und Software. Auf diese Weise können über 275 Millionen Objekte zugänglich gemacht werden.

Beschr. Das Rechercheportal bietet eine integrierte multilinguale Suche nach wissenschaftlicher Literatur auf weltweit verteilten Repositorien. Die Besonderheit liegt in einer multilingualen Recherche mit Hilfe des *Microsoft Translator*. Für die Eingabe von Suchbegriffen stehen zehn

Sprachen zur Verfügung. So kann z. B. in deutscher Sprache und mit lateinischer Schrift nicht nur in englischen und französischen, sondern auch in chinesischen, russischen und arabischen Datenbanken gesucht werden. Die originalsprachlichen Titeldaten der Trefferlisten können in andere Sprachen übertragen werden. Auf diese Weise soll die Forschung auch über Sprachgrenzen hinweg gefördert werden. – Die Erweiterte Suche bietet vielfältige Recherchezugriffe und Filtermöglichkeiten. Eine thematische Clusterung der Treffer lässt sich mit der Anzeige der jeweiligen Trefferzahlen auch visuell darstellen.

Bezieh. Deutscher Partner der WorldWideScience.org ist die TIB Hannover. – Auf derselben technischen Grundlage (*Deep Web Technologies*) basiert die Suchmaschine *Science Research*.

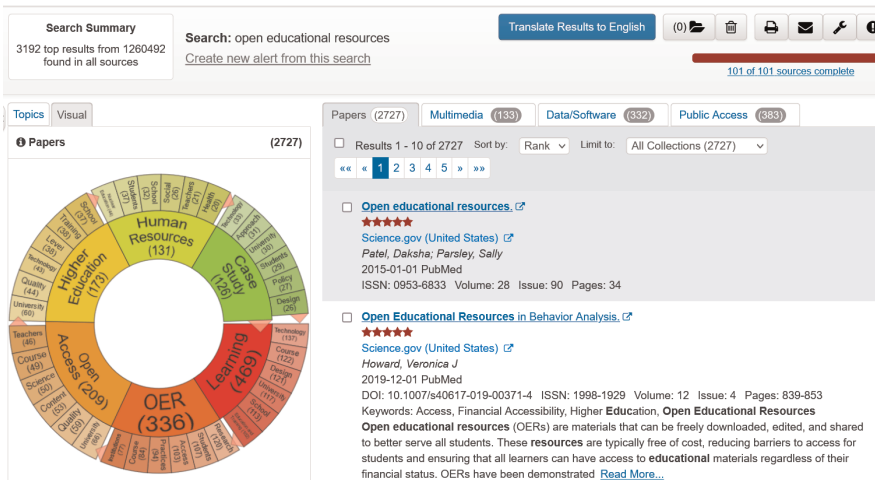


Abb. 125: Trefferanzeige in WorldWideScience.org mit thematischer Clusterung

