

Versuch eines fachbezogenen Vergleichs von Discovery/Recherche-Tools am Beispiel der Physik

Bericht zum Praktikumsprojekt am BSZ

J. ROHRWILD

Version vom 19. August 2016

Zusammenfassung:

Die Websuchmaschine Google Scholar wird mit der Fachdatenbank Scopus und dem Discovery System EDS (über den Freiburger Katalog plus) verglichen. Hierzu werden auf Basis einer kleinen Stichprobe von Artikeln aus dem Bereich der Physik Suchabfragen konstruiert. Die ersten zwanzig Einträge der Treffermenge der einzelnen Recherchertools werden auf ihre Relevanz hin geprüft, um eine grobe Einschätzung der Güte der Treffermengen zu bekommen. Wir finden leichte Unterschiede zwischen den Plattformen, aber keine schneidet deutlich besser oder schlechter ab.

1 Einleitung

Getreu dem Motto „what you can’t find you cannot use“ wurden in den letzten 10 Jahren die Anstrengung seitens der Bibliotheken intensiviert, Zugang zu wissenschaftlichen Aufsätzen nicht nur über stetig steigende Lizenzgebühren zu erwerben, sondern dem Nutzer auch einen einfacheren Weg zur Verfügung zu stellen, diese Artikel schnell zu finden. Der klassische OPAC (Online Public Access Catalogue) wurde (und ist) für eine Suche auf Artekelebene schon wegen fehlender Metadaten¹ nicht ausgelegt. Vor diesem Hintergrund wurden in den letzten zehn Jahren mehrere kommerzielle Resource Discovery Services (RDS) entwickelt, die das Auffinden von elektronischen Aufsätzen erleichtern und direkt oder über Linkresolver Zugriff auf die lizenzierten Volltexte bieten. Das erste System dieser Art war wohl der von OCLC entwickelte WorldCat Local, der 2007 vorgestellt wurde [1]. Nach diesem Startschuss wurden in den folgenden Jahren weitere RDS-Angebote auf den Markt gebracht: Summon (damals Series Solutions, heute ProQuest), EBSCO Discovery Services (EDS) und Primo (ExLibris) um die wichtigsten zu nennen. Mittlerweile verfügen diese RDS jeweils über mehr als eine Milliarde Dokumente („records“). Bei dieser gewaltigen Zahl ist zu bedenken, dass dies nicht allein wissenschaftliche Aufsätze umfassen kann², sondern auch Vorträge, Präsentationen, Auszüge aus Buchkapiteln

¹Im Katalog des SWB finden sich lediglich $\mathcal{O}(300000)$ Aufsätze.

²Die Max-Planck Digital Library schätzt die Zahl der im Jahr veröffentlichten Aufsätze auf ca. 2 Millionen [2]. Für eine Milliarde Artikel müsste das jährliche Publikationsvolumen seit dem Druck der Gutenberg-Bibel annähernd konstant geblieben sein - eine offensichtlich unsinnige Annahme.

usw. beinhaltet. Alternativ zu den kommerziellen Systemen wurden auch frei zugängliche Systeme entwickelt. Das bekannteste ist das 2004 gestartete Google Scholar. Im selben Jahr wurde aber auch beispielsweise die auf Open Access (OA) spezialisierte Bielefeld Academic Search Engine (BASE) für die breite Öffentlichkeit freigeschaltet [3]. BASE verfügt mittlerweile über mehr als 90 Millionen indexierte Dokumente [4]. Ein kommerzieller Open-Access-RDS, oaFindr, wird momentan von 1science aufgebaut [5].

Neben RDS Systemen ermöglichen auch Fachdatenbanken eine Recherche auf Artikelbasis. Im Gegensatz zu einem RDS kann hier nicht gleichzeitig der lokale Katalog mit durchsucht werden. Die Links zu den lizenzierten Volltexten werden auch hier (oft über Linkresolver) direkt bei den Treffern angezeigt. Wie der Name bereits zum Ausdruck bringt, sind Fachdatenbanken in der Regel auf bestimmte wissenschaftliche Disziplinen beschränkt. Teilweise sind diese Datenbanken so wichtig, dass sie in ihrem Fach der de facto Standard sind. Artikel, die nicht in dieser Datenbank gelistet sind, werden vom Fachpublikum praktisch kaum wahrgenommen; Ein Beispiel hierfür ist die freie inSPIRE Datenbank [6] für den Bereich Hochenergie- und Teilchenphysik. Neben Fachdatenbanken, wie z.B. inSPIRE, TEMA, WEMA oder wiso, gibt es auch einige wenige große fachübergreifende Datenbanken; Die bekanntesten sind Web of Science und Scopus.

Die Verfügbarkeit, in Form von Google Scholar, einer großen, frei zugänglichen wissenschaftlichen Suchmaschine zwingt Bibliotheken eine Frage auf: Wie viel Mehrwert bringen kommerziellen Angebote, für die jährlich Lizenzgebühren entrichtet werden müssen? Bei der Beantwortung dieser Frage gibt es grundsätzlich mehrere Stoßrichtungen. Einerseits gibt es den Mehrwert gleichzeitig lokale Kataloge und RDS-Indices durchsuchen zu können. Weiter sind RDS individualisierbar, d.h. die Menge der durchsuchten Datenbanken kann auf die Bedürfnisse und Angebote der Bibliotheken bzw. Forschungseinrichtungen zugeschnitten werden. Wie groß der Effekt eines gut eingestellten Systems sein kann, wurde z.B. in [7] diskutiert. Darüber hinaus stellt sich natürlich die Frage nach der Abdeckung mit Aufsätzen, die die einzelnen RDS bieten. Darunter kann zum Einen das Auffinden eines Artikels auf Basis eines vorhandenen Zitates (und Zugriff auf den Volltext) gemeint sein. Zum Anderen kann es sich auch auf die Treffermenge einer thematischen Recherche über bestimmte Schlagwörter unter Ausnutzung von Fachvokabular beziehen.

Ein Vergleich von Google Scholar mit EDS auf Basis einer Suche nach 500 repräsentativen Artikeln, von denen Titel und Autoren bekannt waren, wurde in [8] durchgeführt. Die Abdeckung und Aktualität von Google Scholar bei Zeitschriftentiteln wurden in [9] im Detail untersucht. Hinzu kommen mehrere Studien, die die Funktionalitäten von freien Suchmaschinen und kommerziellen RDS sowie von Datenbanken untereinander vergleichen, siehe z.B. [10, 11, 12]. Wir werden uns im Folgenden mit der Frage beschäftigen, wie relevant die von den einzelnen Systemen gelieferten Treffermengen sind. Diese Fragestellung wurde bereits für einen Vergleich von verschiedenen Fachdatenbanken mit Google Scholar verwendet [13]. Wir beschränken uns auf thematische Suchen im Bereich der Physik und auf die Datenbank Scopus, sowie Google Scholar und den Freiburger Katalog plus, der an ein EDS-System angebunden ist. Die Untersuchung selbst erfolgte auf Basis einer kleinen Stichprobe, da für Konzeption, Durchführung, Auswertung und Erstellen des Berichtes nur etwas über zwei Wochen zur Verfügung standen.

Diese Arbeit ist wie folgt aufgebaut: Zunächst werden die einzelnen Recherche-Systeme

kurz vorgestellt. Dann beschreiben wir den „experimentellen“ Ablauf der Untersuchung³, insbesondere diskutieren wir den Aufbau der Suchabfragen und die methodischen Probleme, die mit der Untersuchung einhergehen. In Abschnitt 4 werden die Resultate der Untersuchung beschrieben. Nach einem kurzen Vergleich mit anderen Arbeiten zu diesem Thema, fassen wir in Abschnitt 6 die wichtigsten Ergebnisse nochmal zusammen und geben einen knappen Ausblick.

2 Beschreibung der einzelnen Plattformen

Bevor wir uns dem konzeptionellen Aufbau der Untersuchung widmen, stellen wir die drei Recherche-Plattformen kurz vor. Eine detaillierte Beschreibung der Frontends und ein ausführlicherer Abriss der Entstehungsgeschichte kann in z.B. [1, 10, 18] gefunden werden.

Scopus

Scopus ist eine vom Elsevier Verlag betriebene kommerzielle Datenbank mit bibliographischen Daten (inkl. Abstracts) von über 60 Millionen Journal-Artikeln, 7,2 Millionen Konferenzbeiträgen, 27 Millionen Patenten und über 100000 Büchern [14, 15]. Etwa 1/3 der erfassten Veröffentlichungen sind aus der Medizin. Physik, Chemie und Technik machen 29% der Einträge aus. Der Rest stammt aus Sozialwissenschaften, Wirtschafts- und Lebenswissenschaften. Die Datenbank beinhaltet auch bibliometrische Informationen; Zitationen werden auf Ebene von einzelnen Artikeln erfasst. Dies erlaubt es Scopus verschiedene Indikatoren, wie den h-Index [16], oder komplette Zitatanalysen für einzelne Autoren oder Einrichtungen, denen eindeutige Identifikationsnummern zugewiesen werden, zu bestimmen. Insbesondere erlaubt dies auch Suchanfragen nach allen Aufsätzen, die eine vorgegebene Menge von Referenzartikeln zitieren bzw. von diesen zitiert werden. Dies ist hilfreich wenn z.B. nach Veröffentlichungen gesucht wird, die unterschiedliche (experimentelle) Techniken kombinieren.

Die Suche in der Datenbank erfolgt in der Regel über eine modulare Eingabemaske, siehe Abbildung 1. Hierbei können Links- und Rechtstrunkierung sowie Wildcards zum Einsatz kommen. Die Suchanfragen können über Dropdown-Menüs auf spezielle Metadaten-Felder⁴ beschränkt und mit Booleschen Operatoren verknüpft werden. Die Standardeinstellung ist eine Suche mit UND-Verknüpfung in der Feldgruppe „Article Title, Abstract, Keywords“. Da die Treffermengen der einzelnen Suchen gespeichert werden, ist auch ein nachträgliches Vereinigen bzw. Schneiden unterschiedlicher Mengen möglich.

³Die von uns verwendete Methodik ist sehr nahe an der in [13] vorgestellten. Leider hat der Autor diese Publikation erst nach Durchführung der Untersuchung entdeckt. Die Methodik wurde vollkommen unabhängig entworfen.

⁴Felder sind neben Titel, Autor, Abstract auch sog. Schlagworte (Keywords), Affiliationen oder Förderkennzeichen (grant number). Das normierte Vokabular, das die fachliche Zuordnung erleichtert, ist einer der klassischen Vorteile von (gepflegten) Datenbanken.

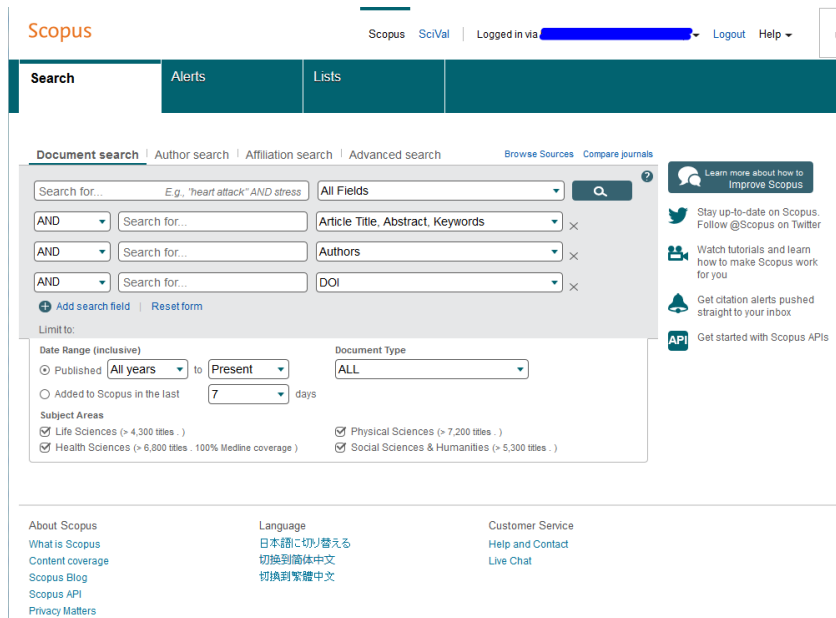


Abbildung 1: Oberfläche von Scopus mit Eingabemaske.

Bei Recherchen ist zu bedenken, dass Scopus zwar auf die Volltexte verlinkt, jedoch nicht im Volltext selbst sucht. Dies wird teilweise durch die Verschlagwortung der einzelnen Artikeln kompensiert, auch wenn manche, vor allem ältere Artikel, (noch) nicht verschlagwortet sind. Neben der Suche über die Eingabemaske gibt es auch eine Experten-suche, die in Textform direkt über Angabe der Feldbezeichnungen ($\mathcal{O}(100)$ Feldcodes) und Boolesche Operatoren durchgeführt wird. Unabhängig davon wie die Suche angestoßen wurde, kann die Treffermenge nach Autor, Zitatzahl, Veröffentlichungsdatum oder Quelle sortiert werden. Wir verwenden als Standardeinstellung eine Sortierung nach Relevanz (relevance ranking).

Google Scholar

Die Suchmaschine Google Scholar wurde 2004 in einer Beta-Version [17] im Web freigeschaltet. Sie stellt den “akademischen” Ableger der Suchmaschine Google dar und ist auf wissenschaftliches Material im weiteren Sinne beschränkt. Darunter fallen neben Artikeln (in Fachzeitschriften und im grünen OA), Konferenzbeiträge, Bücher, Patente aber auch Vorträge und Präsentationen mit wissenschaftlichem Inhalt. Neben der für Websuchmaschinen typischen Suche über einen einzelnen Suchschlitz, siehe Abbildung 2, bietet Google Scholar auch eine erweiterte Suche an. Hier kann beispielsweise die Suche auf die Titel der Artikel, auf eine bestimmte Zeitschrift oder auf Autoren eingeschränkt werden. Die Suchsyntax von Google Scholar erlaubt Boolesche Operatoren. Die UND-Verknüpfung wird standardmäßig zwischen einzelnen Suchtermen ergänzt. Der NICHT-Operator wird durch ein vorangestelltes Minuszeichen (z.B. „-Tanz“, um bei einer Suche nach (Fuß-)Bällen den Opernball aus der Treffermenge zu entfernen). Das logische ODER wird durch ein „OR“ (in Großbuchstaben) dargestellt [19]. Eine Suche

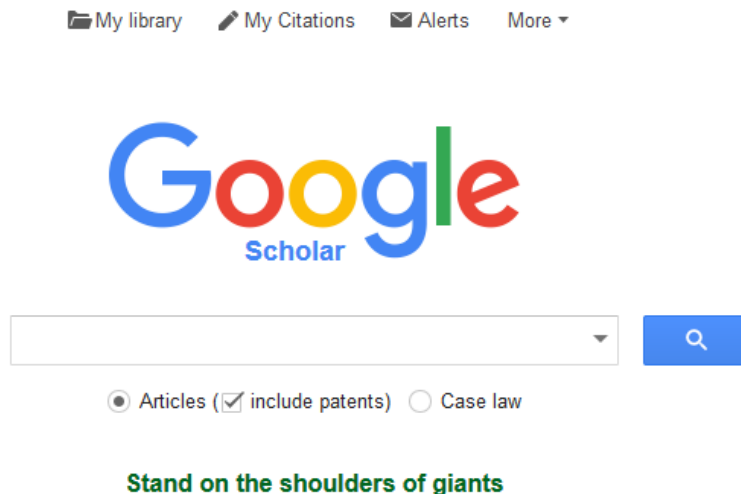


Abbildung 2: *Startseite von Google Scholar.*

in spezifischen Metadatenfeldern kann durch setzen spezieller Tags im Suchausdruck veranlasst werden. Z.B. kann mit "au:Einstein" die Treffermenge auf Artikel mit Einstein als Autor beschränkt werden.

Multiple Treffer für denselben Artikel gruppiert Google Scholar automatisch unter einem einzelnen Eintrag mit der Option, die Einzeltreffer separat anzuzeigen. Die Treffer können nach Datum oder Relevanz sortiert werden; letzteres ist die Standardeinstellung. Wie diese Sortierung erfolgt, ist ähnlich wie bei dem Sortieralgorithmus von Google, nicht bekannt. Bestimmt der Algorithmus, dass die Suche einem bestimmten Aufsatz gilt, so wird nur der entsprechende Treffer angezeigt [20]. Weitere Treffer können aber zugeschaltet werden. Dies soll das schnelle Auffinden von Artikeln durch bloßes Kopieren von Zitatauszügen ermöglichen⁵. Google Scholar stellte eigene bibliometrische Daten zur Verfügung [10] und angezeigte Aufsatz-Treffer werden von Google Scholar mit gefundenen Zitaten versehen. Dies erlaubt es, wie in vielen Aufsatzdatenbanken (Scopus, Web of Science, inSPIRE) üblich, auf einem Artikel aufbauende Veröffentlichungen zu finden.

EDS via Katalog plus

Der Katalog plus ist ein von der Bibliothek der Universität Freiburg entwickeltes Discovery System. Durch eine Anbindung an das EDS erlaubt der Katalog plus gleichzeitig eine Suche im Freiburger Katalog und im EDS. Dabei werden Treffer zu über EDS eingebundenen kostenpflichtigen Datenbanken teilweise nur Nutzern im Universitätsnetz angezeigt. Externe Nutzer erhalten jedoch einen Hinweis darauf, dass der Treffer nicht angezeigt

⁵Bei einem Test lieferte eine Suche nach "Phys. Rev. 105 (1957) 1413" direkt einen Link zur Zeitschriftenhomepage, die den Volltext des original Aufsatzes zum sog. Experiment von Wu [21] enthält, zurück.

Katalog plus

Bücher & mehr Artikel & mehr

» Einfache Suche » Frühere Suchen » Merkliste [0] » Feedback » Hilfe

Suche in den Inhalten zahlreicher Anbieter (teilweise bis auf Volltextebene) und allen lizenzierten EBSCOhost-Datenbanken (EBSCO Discovery Service).

[\[Suchbegriffe entfernen\]](#)

Erweiterte Suche

Suche in allen Feldern:

Personen:

Wörter aus dem Titel:

Schlagwörter:

ISBN:

ISSN:

Erscheinungsjahr:

Veröffentlicht in:

Zusammenfassung:

Suche in allen Feldern

Hier wird in den meisten Datenfeldern gesucht. Das ermöglicht einerseits eine Suche auf breiter Datenbasis, andererseits kann es auch schnell zu großen und ungenauen Treffermengen führen. Je mehr Begriffe Sie eingeben, desto geringer wird die Treffermenge.

Bei Bedarf können einzelne Suchbegriffe mit einem Stern (*) trunkiert werden

Übersicht | Kontakt | Impressum © Copyright 2013 Universität Freiburg A A A

Abbildung 3: Suchmaske von Katalog plus bei der erweiterten Suche.

werden konnte. Bei der Suche über EDS werden von der UB Freiburg vorgegebene Einstellungen berücksichtigt, die auf das Angebot der Universität hin optimiert sind. Andere Einstellungen können zu einer deutlich anderen Trefferzahl und Qualität führen [7]. Die Universität Freiburg verfügt über ein breit aufgestelltes Physikdepartment (Schwerpunkte unter anderem theoretische und experimentelle Hochenergiephysik, kondensierte Materie und Optik). Damit bietet der Katalog plus die Möglichkeit, ein von einer erfahrenen Bibliothek gut eingestelltes EDS einer Universität mit relevantem Fächerspektrum nutzen zu können. Im Hinblick darauf wie wichtig die richtige Einstellung des RDS ist [7], war der Katalog plus die erste Wahl für den Test eines EDS-basierten Systems.

Katalog plus verfügt, wie die anderen Recherche-Oberflächen auch, über eine einfache und eine erweiterte Suche. Die Einstiegsmaske zur erweiterten Suche ist in Abbildung 3 dargestellt. Die wichtigsten Suchfelder werden dabei über die Maske direkt bereitgestellt. Für noch feinere Abfragen kann aus einer großen Zahl von Feldindikatoren ausgewählt werden [22], die in den Suchtext eingebettet werden. Trunkierungen werden dabei unterstützt. In der Sucheingaben können Boolesche Operatoren wie “OR” oder “AND” “AND NOT” analog zu Scopus verwendet werden. Die Treffermenge kann nach Datum und Relevanz sortiert werden. Die Relevanz-Sortierung ist die Standardeinstellung. Treffer, die den gleichen Aufsatz darstellen, aber aus unterschiedlichen Quellen stammen (z.B. BASE, arXiv und Zeitschrift) werden nicht unter einem Eintrag aggregiert, sondern als separate Ergebnisse dargestellt. Zu jeder Suche wird außerdem eine Auswahl an Facetten angeboten. Für registrierte Nutzer können die Suchen gespeichert werden.

Bereich	Titel	Autoren	Referenz
HEP	Gauge theory correlators from noncritical string theory	S. S. Gubser, I. R. Klebanov and A. M. Polyakov	[23]
HEP	The Anti-k(t) jet clustering algorithm	M. Cacciari, G. P. Salam and G. Soyez	[24]
HEP	What is the $\gamma\gamma$ resonance at 750 GeV?	R. Franceschini <i>et al.</i>	[25]
CO	The Inflationary Universe: A Possible Solution to the Horizon and Flatness Problems	A. H. Guth	[29]
CO	Joint Analysis of BICEP2/ <i>KeckArray</i> and <i>Planck</i> Data	P. A. R. Ade <i>et al.</i> [BICEP2 and Planck Collaborations]	[28]
CO	The characterization of the gamma-ray signal from the central Milky Way: A case for annihilating dark matter	T. Daylan, D. P. Finkbeiner <i>et al.</i>	[27]
CO/HEP	Topological Defects and nano-Hz Gravitational Waves in Aligned Axion Models	T. Higaki <i>et al.</i>	[26]
CM	Self-consistent equations including exchange and correlation effects	Kohn, W. and Sham, L.J.	[30]
CM	Two-dimensional gas of massless Dirac fermions in graphene	K.S. Novoselov <i>et al.</i>	[31]
CM	Properties of Cu(In,Ga)Se ₂ solar cells with new record efficiencies up to 21.7%	P. Jackson <i>et al.</i>	[32]
OP	Candela-class high-brightness In-GaN/AlGaN double-heterostructure blue-light-emitting diodes	S. Nakamura, T. Mukai and M. Senoh	[33]
OP	Photonic crystal fibers	P. Russell	[34]
OP	Optical atomic clocks	A.D. Ludlow <i>et al.</i>	[35]

Tabelle 1: *Verwendete Artikel nach Fachgebiet.*

3 Methodik

Zunächst wurden, ohne die Hilfe eines der drei Systeme in Anspruch zu nehmen, Artikel aus den Bereichen Hochenergiephysik (HEP), Astrophysik/Kosmologie (CO), Physik der kondensierten Materie (CM) und der Optik (OP) ausgewählt. Hierfür wurden in jedem Teilgebiet ein älterer, für seine Disziplin fundamentaler Aufsatz, ein weiterer, wichtiger (typischerweise einige Jahre zurückliegender) Artikel und eine beliebige Veröffentlichung, die weniger als ein Jahr alt ist, gesucht. Zusätzlich wurde ein einzelner, hochaktueller Aufsatz (seit einer Woche verfügbar) ausgewählt. In der HEP und der CO wurde die inSPIRE-Datenbank für die Auswahl verwendet; in den anderen beiden Fachbereichen Primo und die Milestone-Liste von Nature. Die Aufsätze sind in Tabelle 1 zusammengefasst.

Danach wurden für jeden Aufsatz vier Suchabfragen konzipiert:

1. Stichwortsuche mit 3-5 Worten aus dem Titel
2. Stichwortsuche wie in (1) zzgl. ein Autorenname
3. Thematische Suche mit vier Gruppen von Schlag- und Stichworten
4. Wie in (3) aber mit grober Einschränkung auf das Erscheinungsjahr des Referenzartikels.

Das bei Suche (4) zusätzlich eingeführte Zeitfenster ist vor allem bei älteren Arbeiten relevant, die sich im Laufe der Zeit zu einer Standardreferenz entwickelt haben. Das

Zeitfenster ist ein Versuch, die in neueren Arbeiten im Zusammenhang mit einer bestimmten Referenz genannten „Buzz-Words“ herauszufiltern. Außerdem ist eine solche Suche denkbar, wenn grundlegende Arbeiten zu einem Thema, beispielsweise zur Einarbeitung in eine Masterarbeit, oder aber nur hochaktuelle, neue Entwicklungen auf einem Forschungsgebiet gesucht werden. Da nahezu alle Publikationen in der Physik auf Englisch verfasst sind, werden nur englische Suchbegriffe verwendet. Wenn möglich werden auch die Treffermengen bei den Suchenabfragen vorab auf englischsprachige Dokumente beschränkt.

Zur besseren Illustration für eine solche thematische Suche betrachten wir ein Beispiel. Ausgehend von dem Artikel „What is the $\gamma\gamma$ resonance at 750 GeV?“ von Franceschini *et al.* [25], einer der drei Aufsätze aus der HEP, wurde eine generelle Fragestellung entworfen: Welche Artikel befass(t)en sich mit Ideen zu dem scheinbaren Signal im am LHC gemessenen Zwei-Photon-Spektrum? Darauf aufbauend wurden vier Gruppen von (englischen) Begriffen konstruiert, die sich für eine thematische Suche eignen (in Großbuchstaben Boolesche Operatoren):

1. (heavy resonance) OR (new particle)
2. bsm OR (beyond sm) OR (beyond standard model)
3. (gamma-gamma) OR diphoton OR (di-photon) OR (two-photon)
4. lhc OR (hadron collider)

Diese vier Bereiche können dann z.B. mit Hilfe einer AND-Verknüpfung kombiniert werden. Diese Suche könnte natürlich noch weiter verfeinert werden. So bieten sich in diesem Fall eine zusätzliche Suche nach „excess OR signal OR bump“ sowie „invariant mass spectrum“ an. Allerdings wurde der Einfachheit halber auf zu detaillierte Suchanfragen verzichtet, um die Gefahr, versehentlich relevante Ergebnisse zu unterdrücken, gering zu halten. Wie genau die einzelnen Abfragen später kombiniert werden, ist von der Syntax der jeweiligen Recherche-Plattform abhängig (siehe Abschnitt 2).

Nachdem die 13 Suchanfragen festgelegt wurden, war der Wunsch diese nicht weiter zu verändern, um nicht versehentlich einzelne Suchmaschinen zu bevorteilen. Allerdings wurde in zwei Fällen eine Korrektur vorgenommen, nachdem festgestellt wurde, dass die Suche sehr viele Treffer aus den Lebenswissenschaften bzw. Computerwissenschaften erbrachte, die auf eine zufällige Überschneidung zweier Fachbegriffe zurückzuführen war. Die Suche wurde deshalb um einen weiteren, fachspezifischen Ausdruck erweitert⁶.

Bei der Stichwortsuche im Titel bzw. in Titel und Autor wurde erfasst, ob der Artikel gefunden wurde, an welcher Stelle im Ranking der Artikel auftritt und wie viel Treffer insgesamt gefunden wurden. Bei Google Scholar und EDS via den Katalog plus wurde die Suche immer in „allen Feldern“ ausgeführt, ohne spezielle Feldbezeichnungen zu benutzen. In Scopus wurde nicht in allen Feldern, sondern explizit im „Authors“ und in den „Article Title, Abstract, Keywords“ Feldern gesucht (siehe Abbildung 1). Der Grund hierfür ist, dass in Scopus die Suche standardmäßig in den Feldern „Article Title, Abstract,

⁶Das Hinzunehmen von „physics“ alleine genügte unglücklicherweise in einem der Fälle nicht.

Keywords“ durchgeführt wird, wenn dies nicht über das Dropdown-Menü geändert wurde. Da ein Nutzer wohl kaum aktiv seine Suchanfrage verschlechtern wird, wurde die Suche hier bewusst nicht auf alle Felder ausgedehnt. Bei den anderen Systemen wurde immer versucht mit einem Suchschlitz zu arbeiten.

Thematische Suchen erfolgten grundsätzlich, unabhängig vom verwendeten System, in allen Feldern. Erfasst wurde hierbei wieder die totale Treffermenge, ob der Titel, der den Anlass zu der Suche gegeben hat, gefunden wurde und wenn ja, an welcher Stelle des Rankings. Weiter wurden die ersten 20, nach der Relevanz sortierten Treffer ausgegeben. Diese wurden daraufhin untersucht, ob sie vor dem Hintergrund der gestellten Frage relevant, marginal oder irrelevant sind. Als relevant wurden, nach Prüfung des hinter dem Link liegenden Abstracts (oder Buchkapitels), solche Treffer befunden, die der Auswertende als hinreichend zutreffend empfand. Marginale Treffer waren solche, die teilweise passend empfunden wurden. Der Rest wurde als irrelevant eingestuft. In diese letzte Klasse wurden auch „tote“ Links eingeordnet.

Um bei dem obigen Suchbeispiel zu bleiben:

- “Indirect probes of the trilinear Higgs coupling: $gg \rightarrow h$ and $h \rightarrow \gamma\gamma$ ” von M. Gorbahn und U. Haisch [36] würde in diesen Zusammenhang (und nur in diesem!) als irrelevant eingestuft
- “Status of a minimal composite Higgs theory” von Z. Fodor *et al.* [38] scheint das eigentliche Suchthema nur am Rand zu behandeln und ist deshalb marginal
- „Diphoton portal to warped gravity” von A. Falkowski und J. Kamenik [37] wäre relevant.

Für jede Suche wurden neben der Einschätzung der ersten 20 Treffer auch die Anzahl der insgesamt zurückgelieferten Treffer festgehalten. Bei Suchen über Google Scholar kann nicht ausgeschlossen werden, dass Informationen über frühere Suchanfragen bei der Bewertung der Treffermenge berücksichtigt werden. Um solche Gedächtniseffekte möglichst gering zu halten, wurde jede Suche mit einer neuen Browser-Session ausgeführt; nach jeder Suche wurde auch die gesamte Browser-Historie und Cookies gelöscht. Bei den Treffern wird bewusst darauf verzichtet zu prüfen, ob tatsächlich Zugriff auf die Volltexte möglich ist. Die Untersuchung befasst sich lediglich mit der (fachlichen) Relevanz der Ergebnisse und nicht mit der Volltextabdeckung. Vereinfacht wird hier angenommen, dass Wissenschaftler entweder direkt aus dem IP-Nummernkreis ihrer Forschungseinrichtung auf die Dokumente zugreifen können oder im Fall der Fälle über eine Fernleihe bzw. über direkte Korrespondenz mit einem der Autoren an eine Postprint-Version kommen können.

Schwächen und Probleme der Methodik

Der oben beschriebene Ablauf hat einige methodische Probleme. Zunächst einmal ist die Anzahl der Stichproben klein. Auch die Auswertung der Treffer beschränkt sich auf eine relativ kleine Menge von nur zwanzig. Damit ist bereits vor Beginn der Untersuchung klar, dass keinerlei statistische Aussagen, z.B. eine Abschätzung der Unsicherheiten, getroffen

	Scopus	Google Scholar	Katalog plus
Top3	8(11)	11(11)	7(10)
Top10	2(0)		2(0)
Top20			1(0)
gefunden	2(1)		2(2)
nicht gefunden	1(1)	2(2)	1(1)*
Median	123	8250	2453

Tabelle 2: *Ranking der gesuchten Artikel in den unterschiedlichen Plattformen. In (ohne) Klammern die Ergebnisse mit (ohne) Autorennamen in der Suchanfrage. Die unterste Zeile gibt den Median der Mächtigkeit der Treffermengen an.*

werden können. Es ist daher wichtig sich vor Augen zu führen, dass die Untersuchung nur einen ersten, qualitativen Eindruck vermitteln kann.

Außerdem sind die thematischen Suchanfragen relativ einfach aufgebaut. Eine detaillierte Literaturrechercheanfrage in einer Fachdatenbank wird typischerweise mehr, dafür aber sehr viel weiter gefasste⁷ Themenbereiche haben und erfordert deshalb eine sorgfältige Planung und einiger Testläufe. Schon im Hinblick auf den zeitlichen Rahmen war ein solcher Suchaufbau nicht möglich. Darüber hinaus verfügt der Auswerter nur in den Teilgebieten Hochenergiephysik und in geringerem Rahmen in Astrophysik/Kosmologie über eigene (mehrjährige) Forschungserfahrung. Die Bereiche Kondensierte Materie und vor allem Optik sind daher fachlich deutlich schwieriger zu beurteilen. Dies machte auch das Konzipieren einer repräsentativen Suchanfrage aufwändiger und erhöht zudem die Wahrscheinlichkeit, einen wichtigen Suchbegriff zu übergehen.

Ein weiteres Problem bei dem oben skizzierten Vorgehen, ist die Tatsache, dass die Entscheidung, ob ein Treffer relevant ist oder nicht letztendlich immer subjektiv ist. Es ist dem Auswerter überlassen, die einzelnen Treffer zuzuordnen. Diese Einschätzung hängt nicht nur vom Fachwissen der Person ab, sondern implizit immer von den Erwartungen wie die Suchergebnisse aussehen sollten. Naiv könnte man erwarten, dass sowohl ein enger gefasstes Bild von der Relevanz der Ergebnisse als auch Schwierigkeiten die Relevanz einzuschätzen sich primär in der absoluten Häufigkeit der Treffer äußern. Bei den relativen Häufigkeiten sollte dies einen geringeren Einfluss haben. Beim Fachwissen kann dieses Argument aber höchstens dann greifen, wenn die unterschiedlichen Recherchertools neben den in die Suche eingebrachten Informationen keine weiteren Informationsquellen, die aus vergangenen, fremden Suchanfragen abgeleitet werden, nutzen. Anderenfalls steht zu befürchten, dass Treffer, die von einem breiten Publikum als interessant oder “einen Klick wert” eingestuft werden, bevorzugt behandelt werden. Dies kann gerade bei fachlich schwierigen Suchen zu erhöhten Fehleinschätzungen bei der Güte der Treffer führen. Hinzu kommt, dass eine vom RDS oder von der Suchmaschine automatisch ausgeführte Query Expansion Einfluß auf Breite und Schärfe der Suche nehmen können.

Schließlich lässt die fachliche Einschränkung keine Schlüsse auf andere Forschungsge-

⁷Alleine die Recherche für ein komplettes Lexikon synonym oder äquivalent gebrauchter Fachtermini erfordert entweder ein zeitintensives Studium von Fachliteratur oder ein vorhandenes tiefes Fachwissen.

biete zu. In der Physik allgemein, vor allem aber in der Astro- und Hochenergiephysik, besteht eine hohe Affinität zu grünem Open Access [39]. Dies sorgt dafür, dass Discovery-Systeme, die aktiv nach Volltexten suchen, diese teilweise schon vor der Einreichung bei einer Zeitschrift in Repositorien oder auf privaten Homepages finden und in den Index einfließen lassen können. Das kann einigen Recherche-Systemen Vorteile bieten, die in anderen Fachbereichen wie den Sozialwissenschaften oder Geisteswissenschaften so nicht gegeben sind.

4 Auswertung

Betrachten wir zunächst die reine Artikelsuche, Suchen (1) und (2) in Abschnitt 3. In Tabelle 2 ist aufgetragen, ob und an welcher Stelle der relevanz-sortierten Treffermenge der Aufsatz gefunden wurde. Die Ergebnisse für eine Suche mit Autor sind in runden Klammern angegeben. Es wird nur unterschieden, ob der Artikel in den ersten drei, zehn oder zwanzig Hits angezeigt wurde. Bei einem der Google Scholar Treffer wurde die in der Einleitung erwähnte Reduktion der angezeigten Hits auf einen Artikel festgestellt. Um ein Gespür für die typische Treffermenge einer Anfrage zu bekommen, wird jeweils der Median der Mächtigkeit der Treffermengen angegeben. Die Veröffentlichung, die von Katalog plus nicht gefunden wurde, war der einzige Fall in dem ein zugriffsgeschützter Treffer (nur aus dem Universitätsnetz sichtbar) an prominenter Stelle auftauchte. Da dieser Treffer auch bei einer Suche nach dem exakten Titel bestehen blieb, liegt die Vermutung nahe, dass es sich hierbei um den gesuchten Artikel handelte.

Berücksichtigt man die typischen Treffermengen, so scheint die Relevanz-Sortierung von Google Scholar das beste Ergebnis zu liefern, allerdings gibt es hier auch zwei Artikel, die gar nicht gefunden wurden: der neuste Artikel, der erst eine Woche vor Durchführung der Untersuchung publiziert wurde und der BICEP2-Artikel, der jedoch mit einer leicht modifizierten Abfrage gefunden wird. Auch muss erwähnt werden, dass der Katalog plus bzw. EDS das einzige Recherche-Tool war, das den neuesten der dreizehn Aufsätze bereits im Index hatte. Bei der Auswertung entstand der Eindruck, dass sowohl Scopus als auch der Katalog plus eine höhere Priorität auf die Aktualität der Aufsätze beim Ranking legen.

Bei der Schlagwortsuche ließe sich im Prinzip zusätzlich zur fachlichen Beurteilung der Treffer, die im Anschluss diskutiert wird, auch die Treffermenge der gefundenen Dokumente vergleichen. Da Scopus als Fachdatenbank für Zeitschriftenartikel naturgemäß deutlich weniger Dokumente umfasst als EDS oder Google Scholar, ist Scopus hierbei außen vor.

Der Vergleich einer bloßen absoluten Anzahl von Treffern ist aber aus zwei Gründen nicht sinnvoll:

- Ohne ein Maß für die Qualität der Treffer kann nicht unterschieden werden, ob mehr tatsächlich besser ist.
- Ein Unterschied zwischen 1000 Treffern und 1 Million ist (ohne eine weitere Verfeinerung der Suche) für den Nutzer im Wesentlichen nicht vorhanden. Schon alleine



Abbildung 4: Relevanzintervalle der dreizehn Suchenabfrage. In cyan (grau) die Ergebnisse von Scopus, in orange (hell grau) Google Scholar und in violett (dunkel grau) Katalog plus. Die Mittelwerte sind mit gestrichelten Balken dargestellt.

1000 Treffer sind nicht mehr „von Hand“ zu inspizieren.

Deshalb verzichten wir im Folgenden auf diesen Vergleich.

Um die Frage nach der thematischen Relevanz der Treffer bei den einzelnen Recherche-Plattformen zu klären, bezeichnen wir zunächst für jede der Suchanfragen die Anzahl der als relevant befundenen Treffer unter den ersten 20 Hits mit N_r . Analog wird N_m und N_i für die Anzahl der marginalen bzw. irrelevanten Artikel verwendet.

Damit kann man die beiden Verhältnisse

$$R_{max} = \frac{N_r + N_m}{N_i + N_m + N_r} \quad \text{und} \quad (1)$$

$$R_{min} = \frac{N_r}{N_i + N_m + N_r} \quad (2)$$

bestimmen. Das Intervall $I_r = (R_{min}, R_{max})$ wird im Folgenden als Relevanz-Intervall bezeichnet werden. Anschaulich entspricht es dem Anteil der Dokumente aus den ersten 20 Treffern, die nach einer groben Bewertung als thematisch gut oder teilweise passend eingestuft wurden.

In Abbildung 4 sind die Intervalle I_r für die dreizehn Suchabfragen auf Basis von Schlagworten und Zeitfenstern (Suche (4) nach der Einteilung aus Abschnitt 3) abgebildet. Die Intervalle für Scopus (cyan/grau), Katalog plus (violett/dunkel grau) und Google Scholar (orange/hell grau) sind den einzelnen Suchabfragen entsprechend gruppiert. Die gestrichelten Relevanzintervalle am unteren Rand der Graphik stellen den naiven arithmetischen Mittelwert für die einzelnen Recherche-Plattformen dar.

Je weiter rechts ein Balken beginnt, desto höher ist der Anteil der voll zutreffenden Suchergebnisse. Bei der Mehrheit der Anfragen gibt es keinen nennenswerten Unterschied zwischen den Plattformen. Allerdings gibt es beim Katalog plus eine Suche, die einen deutlich höheren Anteil von unpassenden Treffern zurücklieferte. Falls man die Mitte der Intervalle als ein Maß für absolute Güte der Suchergebnisse verwenden möchte, so liegen Google Scholar und Scopus mit 64% bzw. 65% knapp vor dem Katalog plus mit 57%. Allerdings erlaubt eine so kleine Stichprobe bei relativ kleinen Unterschieden keine klaren Aussagen. Wäre beispielsweise der oben erwähnten „Ausreißer“ nicht Teil des Samples, schrumpft der Unterschied von Scopus und Katalog plus auf lediglich 4% (weniger als ein zusätzlicher relevanter Treffer unter den ersten 20 Hits). Ein Vergleich mit den Resultaten ohne zeitliche Einschränkung (Suche (3) in Abschnitt 3) zeigt praktisch keine qualitativen Unterschiede und wir verzichten hier auf ein entsprechendes Diagramm.

5 Vergleich mit anderen Arbeiten

Wie wichtig das RDS für die Bibliothekswelt sind, bemerkt man auch an dem breiten Spektrum von Untersuchungen, Berichten und Empfehlungen für die passende Auswahl eines Discovery Systems [1, 40, 41]. Auch Vergleiche von Datenbanken und RDS mit der großen, frei zugänglichen Plattform Google Scholar wurden bereits mehrfach durchgeführt. Viele dieser Arbeiten sind sehr viel umfangreicher als dieser Bericht. So untersucht

[9] im Detail die Abdeckung von Google Scholar an Hand einer Vielzahl (weit über 10000) unterschiedlicher Zeitschriftentitel. Neben der Abdeckung wurde hierbei auch die Verteilung der verwendeten Quellen und Dokumententypen sowie die Aktualität des Indexes beleuchtet. Bei dieser Untersuchung wurde u.a. festgestellt, dass der Google-Index Lücken bei den aktuellsten Titeln aufweist⁸.

Von besonderem Interesse ist auch [8]: ein Vergleich EBSCO Discovery Service und Google Scholar. Dabei wurden 500 Suchen nach Zeitschriftenaufsätzen aus allen Fachgebieten aus den Log-Files der UB Köln zufällig ausgewählt. Diese Anfragen (Known-Item Searches) wurden dann im EDS und in Google Scholar wiederholt und die Treffermenge ausgewertet. Dabei wurde neben dem bloßen Nachweis auch auf die Möglichkeit eines Zugriffs auf den Volltext geachtet. EDS erreichte die besseren Ergebnisse [8]. Unsere Stichprobe von nur 13 Aufsätzen ist natürlich viel zu klein, um mit der sehr viel größeren Stichprobe vergleichbar zu sein. Zwar war EDS das einzige der untersuchten RDS in dem ein Nachweis des neuesten Artikels gefunden werden konnte. Ein einzelner Artikel erlaubt jedoch noch keine Rückschlüsse.

In [13] wurde eine sehr ähnliche Methodik entworfen, wie sie hier zum Einsatz gekommen ist. Die Untersuchung vergleicht verschiedene Fachdatenbanken aus Ingenieur- und Geisteswissenschaften mit Google Scholar. Dabei werden vier thematische Suchen konstruiert und dann in Datenbank und Suchmaschine ausgeführt. Die Ergebnisse wurden in *sehr gut bis gut passende* und *unpassende* Treffer sowie Dubletten eingeteilt. Dabei waren die Suchanfragen scharf genug konzipiert, die meisten Treffermengen vollständig bewerten zu können—ein Vorteil im Vergleich mit unserer Untersuchung. Die Schlußfolgerungen von [13] lassen sich dahingehend zusammenfassen, dass Google eine größere, aber fachlich deutlich unpassendere Treffermenge erzeugt. Insbesondere konnte Google Scholar nicht die zusätzlichen Funktionalitäten von Datenbanken, wie Speichern der Treffermengen, Thesauri oder normiertes Vokabular, die für fortgeschrittene Suchanfragen nötig sind, kompensieren. Die Suchmaschine ist dafür deutlich einsteigerfreundlicher und gut für einfaches, schnelles Suchen geeignet.

6 Zusammenfassung und Diskussion

Ein Vergleich von unterschiedlichen RDS ist ein ausgesprochen komplexes Unterfangen. Die Fragestellung kann aus verschiedensten Blickwinkeln beleuchtet werden. Um nur einige denkbare Fragestellungen zu nennen: Wie leicht erhält der Nutzer Zugriff auf den (lizenzierten) Volltext? Wieviele Dokumente (ohne Dubletten) befinden sich im Fundus/Index? Bietet die Suchumgebung einen einfachen Einstieg, ist Bedienung und Syntax intuitiv? Lassen sich Facetten und Darstellung individualisieren und sind diese übersichtlich? Lassen sich komplexe Suchabfragen konstruieren und liefert die Plattform entsprechende relevante Ergebnisse? Gibt es eine Möglichkeit bibliometrische Daten abzufragen? Hinzukommt, dass jeder einzelne Aspekt als unterschiedlich wichtig, von essentiell bis belanglos, eingestuft werden kann.

⁸ Allerdings ist die Untersuchung mittlerweile knapp zehn Jahre alt und die Suchmaschine wurde in dieser Zeit zweifellos weiterentwickelt.

Diese Untersuchung beschäftigte sich deshalb nur mit einem kleinen Teilaspekt, der Abdeckung mit relevanten Aufsätzen; und auch das nur in einer der kleineren Fachdisziplinen, der Physik. Selbst diese deutlich einfachere Fragestellung kann auf unterschiedliche Weisen angegangen werden. Für diese Arbeit wurde eine kleine Stichprobe von Aufsätzen zur Grundlage von Suchabfragen, deren “beste 20 Treffer” (nach dem Ranking der Plattform) auf ihren Nutzen für einen (hypothetischen) Nutzer bewertet hin wurden. Damit gilt die Hauptstoßrichtung nicht nur dem Umfang des Angebot der Recherche-Plattform, sondern auch wie gut das Relevance-Ranking der einzelnen Anbieter ist. Die Suchanfragen selbst waren von moderater Komplexität und hätten in vielen Fällen noch restriktiver gestaltet werden können. Allen Plattformen wurden möglichst die gleichen Anfragen (unter Würdigung der individuellen Suchsyntax) gestellt.

Keiner der drei untersuchten Anbieter schnitt deutlich schlechter oder besser als die anderen ab. Bei der thematischen Suche hatten Google Scholar und Scopus leichte Vorteile. Der Katalog plus erzielte im Durchschnitt weniger passende Treffer unter den ersten 20 Hits. Der Unterschied lag aber im Mittel nur bei einen bis zwei “guten” Treffern. Wir wollen deshalb hier nur die Schlußfolgerung ziehen, dass das frei zugängliche Produkt Google Scholar den beiden kommerziellen Systemen in Rahmen dieser Untersuchung nicht unterlegen war. Der Katalog plus wies knapp die beste Erfolgsquote bei der Suche nach bekannten Artikeln (Known-Item-Search) auf.

An dieser Stelle soll noch einmal betont werden, dass ein ähnlicher Versuchsaufbau mit Fokus auf eine andere Fachdisziplin sehr wohl zu anderen Ergebnissen führen kann. Es ist unserer Meinung nach nicht ausgeschlossen, dass die Physik ein ungewöhnlich gutes Anwendungsfeld von freien Suchmaschinen ist.

Abschließend soll noch auf denkbare Verbesserungsmöglichkeiten in der Methodik eingegangen werden. Eine deutliche Vergrößerung der Stichprobe, ähnlich der Arbeit [8] scheint der geradlinigste Ansatz zu sein. Eines der Hauptprobleme liegt aber in der subjektiven Einschätzung der “Trefferqualität”. Dies könnte zu einem gewissen Grad dadurch vermieden werden, dass mehrere Personen unabhängig voneinander die selbe Treffermenge begutachten und die Suchanfragen gemeinsam entwerfen. Nach [13] sind gerade präzisen Suchanfragen, die oft eine Auswertung der gesamten Treffermenge gestatten, eine Stärke von kommerziellen Datenbank-Systemen. Eine dahingehende Verbesserung der Abfragen könnte auch die Auswertung erleichtern. Außerdem wäre ein detaillierter Vergleich der Treffer-Relevanz bei unterschiedlich komplexen Abfragen interessant. So könnten drei Komplexitätsstufen (einfach, mittel und hoch) für jedes Suchthema erstellt werden. Damit kann untersucht werden, wie gut die einzelnen Systeme für Suchabfragen von Einsteigern (Suchen zur ersten Seminararbeit), erfahrenen Nutzern (Einarbeitung in die Masterarbeit) und Nutzern mit großem Fach- und Recherchewissen (Erstellen eines Review-Artikels) geeignet sind.

Literatur

- [1] J. Vaughan, "Web scale discovery services," *Library technology reports* **47** 1.
- [2] R. Schimmer, K. K. Geschuh and A. Vogler (2015), "Disrupting the subscription journals' business model for the necessary large-scale transformation to open access ," doi: 10.17617/1.3.
- [3] [https://en.wikipedia.org/wiki/BASE_\(search_engine\)](https://en.wikipedia.org/wiki/BASE_(search_engine)), abgerufen am 17.08.2016.
- [4] <https://www.base-search.net>, abgerufen am 17.08.2016.
- [5] www.1science.com/oafindr.html, abgerufen am 17.08.2016
- [6] <https://inspirehep.net/>, abgerufen am 17.08.2016.
- [7] J. Zhu and J. Kelley, "Collaborating to Reduce Content Gaps in Discovery: What publishers, Discovery Service Providers, and Libraries Can Do to Close the Gaps," *Science & Technology Libraries* **37** (2015) 315 doi: 10.1080/0194262X.2015.1102677.
- [8] H. Kaluza, "Google Scholar versus ESCO Discovery Service: Ein vergleichender Retrieval-Test," *b.i.t.online* **44** MALIS-Praxisprojekte (2013) Projektberichte aus dem berufsbegleitendem Masterstudiengang Bibliotheks- und Informationswissenschaft der FH Köln, Dinges&Frick, Wiesbaden.
- [9] P. Mayr and A.-K. Walter, "Abdeckung und Aktualität des Suchdienstes Google Scholar," *Information - Wissenschaft & Praxis* **57** 3 (2006) 133.
- [10] P. Jasco, "As we may search - Comparison of major feature of the *Web of Science*, *Scopus*, and *Google Scholar* citation-based and citation-enriched databases," *Current Science* **89** 9 (2005) 1537.
- [11] M. Goodsett, "Discovery search tools: a comparative study," *Reference reviews*, Vol. 28, Nr. 8 (2014) 2 doi: 10.1108/RR-12-2013-0312
- [12] S. M. Shafi and R. A. Rather, "Precision and Recall of Five Search Engines for Retrieval of Scholarly Information in the Field of Biotechnology," *Webology* **2** 2 (2005) 12
- [13] F. Handreck and M.W. Mönnich, "Google Scholar als Alternative zu wissenschaftlichen Fachdatenbanken," *B.I.T.online* **4** (2008).
- [14] <https://www.elsevier.com/solutions/scopus/content>, abgerufen am 17.08.2016.
- [15] Scopus Content Coverage Guide, https://www.elsevier.com/__data/assets/pdf_file/0007/69451/scopus_content_coverage_guide.pdf, abgerufen am 17.08.2016.

- [16] J. E. Hirsch, “An index to quantify an individual’s scientific research output,” *Proc. Nat. Acad. Sci.* **46** (2005) 16569 doi: 10.1073/pnas.0507655102 [arXiv:physics/0508025].
- [17] https://en.wikipedia.org/wiki/Google_Scholar, abgerufen am 17.08.2016.
- [18] P. Jasco, “Google Scolar: the pros and cons,” *Online Infomation Review*, **29** (2005) 2
- [19] <https://scholar.google.com/intl/en/scholar/help.html#overview> und https://en.wikipedia.org/wiki/Google_Scholar, abgerufen am 17.08.2016.
- [20] <http://googlescholar.blogspot.de/2016/01/quickly-lookup-references.html>, abgerufen am 17.08.2016.
- [21] C. S. Wu, E. Ambler, R. W. Hayward, D. D. Hoppes and R. P. Hudson, *Phys. Rev.* **105** (1957) 1413. doi: 10.1103/PhysRev.105.1413
- [22] URL:<https://www.ub.uni-freiburg.de/recherche/was-finde-ich-wo/katalog-plus/hilfe/>, abgerufen am 17.08.2016.
- [23] S. S. Gubser, I. R. Klebanov and A. M. Polyakov, “Gauge theory correlators from noncritical string theory,” *Phys. Lett. B* **428** (1998) 105 doi: 10.1016/S0370-2693(98)00377-3 [arXiv:hep-th/9802109].
- [24] M. Cacciari, G. P. Salam and G. Soyez, “The Anti-k(t) jet clustering algorithm,” *JHEP* **0804** (2008) 063 doi: 10.1088/1126-6708/2008/04/063 [arXiv:0802.1189 [hep-ph]].
- [25] R. Franceschini *et al.*, “What is the $\gamma\gamma$ resonance at 750 GeV?,” *JHEP* **1603** (2016) 144 doi: 10.1007/JHEP03(2016)144 [arXiv:1512.04933 [hep-ph]].
- [26] T. Higaki, K. S. Jeong, N. Kitajima, T. Sekiguchi and F. Takahashi, “Topological Defects and nano-Hz Gravitational Waves in Aligned Axion Models,” *JHEP* **1608** (2016) 044 doi: 10.1007/JHEP08(2016)044 [arXiv:1606.05552 [hep-ph]].
- [27] T. Daylan, D. P. Finkbeiner, D. Hooper, T. Linden, S. K. N. Portillo, N. L. Rodd and T. R. Slatyer, “The characterization of the gamma-ray signal from the central Milky Way: A case for annihilating dark matter,” *Phys. Dark Univ.* **12** (2016) 1 doi: 10.1016/j.dark.2015.12.005 [arXiv:1402.6703 [astro-ph.HE]].
- [28] P. A. R. Ade *et al.* [BICEP2 and Planck Collaborations], “Joint Analysis of BICEP2/*KeckArray* and *Planck* Data,” *Phys. Rev. Lett.* **114** (2015) 101301 doi: 10.1103/PhysRevLett.114.101301 [arXiv:1502.00612 [astro-ph.CO]].
- [29] A. H. Guth, “The Inflationary Universe: A Possible Solution to the Horizon and Flatness Problems,” *Phys. Rev. D* **23** (1981) 347. doi: 10.1103/PhysRevD.23.3470.

- [30] W. Kohn and L. J. Sham, “Self-consistent equations including exchange and correlation effects,” *Physical Review*, **140** 4A (1965) A1133. doi: 10.1103/PhysRev.140.A1133.
- [31] K.S. Novoselov et al, “Two-dimensional gas of massless Dirac fermions in graphene,” *Nature* **438**7965 (2005) 197 doi: 10.1038/nature04233.
- [32] P. Jackson et al, “Properties of Cu(In,Ga)Se₂ solar cells with new record efficiencies up to 21.7%,” *Physica Status Solidi - Rapid Research Letters* **9** 1 (2015) 28 doi: 10.1002/pssr.201409520.
- [33] S. Nakamura, T. Mukai and M. Senoh “Candela-class high-brightness InGaN/AlGaN double-heterostructure blue-light-emitting diodes,” *Applied Physics Letters* **64** 13 (1994) 1687 doi: 10.1063/1.111832.
- [34] P. Russell, “Photonic crystal fibers,” *Science* **299** 5605 (2003) 356 doi: 10.1126/science.1079280.
- [35] A.D. Ludlow, M.M. Boyd, J. Ye, E. Peik and P.O. Schmidt “Optical atomic clocks,” *Reviews of Modern Physics* **87** (2015) 2 doi: 10.1103/RevModPhys.87.637.
- [36] M. Gorbahn and U. Haisch, “Indirect probes of the trilinear Higgs coupling: $gg \rightarrow h$ and $h \rightarrow \gamma\gamma$,” arXiv:1607.03773 [hep-ph].
- [37] A. Falkowski and J. F. Kamenik, “Diphoton portal to warped gravity,” *Phys. Rev. D* **94** (2016) no.1, 015008 doi: 10.1103/PhysRevD.94.015008 [arXiv:1603.06980 [hep-ph]].
- [38] Z. Fodor, K. Holland, J. Kuti, S. Mondal, D. Nogradi and C. H. Wong, “Status of a minimal composite Higgs theory,” *PoS LATTICE* **2015** (2016) 219 [arXiv:1605.08750 [hep-lat]].
- [39] URL:<https://open-access.net/DE-EN/open-access-in-individual-disciplines/physics/>, abgerufen am 17.08.2016.
- [40] J. Deodato, “Evaluating Web-Scale Discovery: A Step-by-Step Guide,” *Information Technology and Libraries*, June 2015 doi: 10.6017/ital.v43i2.5745.
- [41] F. W. Chickering and S. Q. Yang, “Evaluation and Comparison of Discovery Tools: An Update,” *Information Technology and Libraries*, June 2016