

Nutzen und Grenzen von Suchmaschinen

Für den immensen Erfolg von Suchmaschinen im digitalen Zeitalter gibt es mehrere Gründe. Zum Einen ist dies auf die denkbar einfache Bedienung zurückzuführen. Auf der Startseite mit der Einfach-Suche steht ein Suchfeld zur Verfügung, über das nach Eingabe eines Suchbegriffs der gesamte Index durchsucht werden kann. Zum Anderen sorgen minimale Antwortzeiten und eine umfassende Trefferauflistung für einen (subjektiven) Rechercheerfolg. Es ist wichtig zu wissen, dass der Suchmaschinen-Tycoon Google über Jahre hinweg einen Anteil von über 90 Prozent auf dem deutschen Suchmaschinenmarkt hat – in Frankreich oder Spanien ist er noch höher – und damit faktisch vorgibt, was wir im Netz überhaupt finden können. Suchergebnisse des Konkurrenten Yahoo werden im Übrigen seit August 2011 von Microsoft (Bing) zur Verfügung gestellt. Bei wissenschaftlichen Informationsrecherchen mit Hilfe von Suchmaschinen ist jedoch Vorsicht geboten:

- Schätzungen und Untersuchungen zufolge ist der Teil des Internets, wie themenspezifische Datenbanken oder Bibliothekskataloge, der nicht mit Hilfe von Suchmaschinen recherchierbar ist (englisch ‚invisible web‘), 40- bis 500-mal größer als der sichtbare Teil des Web (englisch ‚visible web‘, Bergmann, 2001).
- Art, Umfang, Struktur und Qualität der Datenmenge im Internet sind den Nutzerinnen und Nutzern ebenso weitgehend unbekannt wie die linktopologischen Rankingverfahren, die für die vermeintliche Sortierung der Trefferliste nach Relevanz verantwortlich sind. Durch dieses Verfahren und das in letzter Zeit oftmals vorkommende Index Spamming, bei dem sogenannte Suchmaschinenoptimierer eine Webseite so verändern, dass diese in der Trefferliste eine vordere Platzierung erzielen, wird den Internetnutzer/innen nicht selten eine Relevanz der gefundenen Dokumente vorgetäuscht, die einer näheren Prüfung in Bezug auf Inhalt und Authentizität nicht immer standhalten kann. Ergebnisse, die nicht auf der ersten Seite der Trefferliste stehen, werden meistens ausgesprochen selten gesichtet.
- Mangelhafte Trunkierungsmöglichkeiten (Suche nach Wortbestandteilen) sowie die Nichtberücksichtigung von Synonymen können die Suche im Internet erschweren.

Nicht unerwähnt bleiben sollte ein in den Informationswissenschaften bekanntes Phänomen (‚Serendipity‘, Hull et al., 2008), das Suchenden bei Google oder Yahoo sicherlich schon widerfahren ist: Im Zuge einer Recherche stößt man unbeabsichtigt auf ein Informationsjuwel, einem äußerst interessanten Treffer, der sich als glücklicher Zufall und überraschende Entdeckung zugleich von etwas ursprünglich nicht Gesuchtem erweist.



In der Praxis

Im Sommer 2013 standen die Enthüllungen des IT-Experten Edward Snowden wochenlang im Fokus der politischen Berichterstattung. Der ehemalige Mitarbeiter des US-Geheimdienstes NSA hatte die Internetüberwachung der NSA öffentlich gemacht. Das Programm mit dem Codenamen PRISM diente demnach dem massenhaften Sammeln und Speichern von Telefon- und Email-Verbindungsdaten mit dem Ziel der Auswertung. Betroffen sind hier Google, Facebook & Co., die offensichtlich mit dem Geheimdienst kooperierten. Manche halten die Empörung hierüber für überzogen, der Suchmaschinenverein (SUMA) hingegen spricht von ‚Überwachungsmaschinerie‘ und ‚Internet-GAU‘. Die EU arbeitet an einem neuen europäischen Datenschutzabkommen, wonach US-Konzerne wie Google, Microsoft, Apple oder Facebook Strafen bis zu zwei Prozent des weltweiten Jahresumsatzes drohen, sollten sie künftig Daten von EU-Nutzerinnen und -Nutzern an die US-Behörden weiterreichen (Kölnische Rundschau, 2013). Suchende im Netz haben schon jetzt Alternativen: Statt Google empfiehlt sich die Suche mit Startpage (startpage.com) – eine anonyme Google-Suche über Proxy-Server. Den Betreibern zufolge werden keine IP-Adressen, Suchanfragen oder Tracking-Cookies gespeichert. Gleiches gilt für die Metasuchmaschine Ixquick (vom selben Betreiber) sowie das deutsche Flaggschiff unter den Metasuchmaschinen Metager. Metager arbeitet zurzeit an einer objektiven Verbesserung seines Treffer-Rankings – ein Grund mehr, dieser Metasuchmaschine den klassischen Anbietern wie Google und Yahoo/ Bing mit ihren nicht gerade einfach nachvollziehbaren Ranking-Algorithmen den Vorzug zu geben.

Nutzen und Grenzen von wissenschaftlichen Suchdiensten

Ende 2004 brachte Google seinen wissenschaftlichen Suchdienst Google Scholar in der deutschen Version auf den Markt. Erklärtes Ziel ist die Unterstützung der Scientific Community beim Auffinden wissenschaftlicher Arbeiten. Die oben beschriebenen Mängel sollen unter anderem dadurch kompensiert werden, indem a priori nur wissenschaftlich relevante Inhalte indiziert werden. Google Scholar versucht, die in einem Fachbeitrag zitierte Fachliteratur zu erkennen und als solche suchbar zu machen. Die Ergebnisse werden gemäß dem Page-Ranking von Google und der Zitationshäufigkeit aufgelistet. Google Scholar durchsucht zahlreiche wissenschaftliche Server, wobei auch Volltexte kostenpflichtiger Dokumente kommerzieller Anbieter durchsucht werden. Wie hoch der Anteil der durch Google Scholar erfasste Teil wissenschaftlicher Publikationen im Netz ist, aus welchen Fachgebieten sie stammen und welcher Aktualisierungszyklus zugrunde liegt, kann nicht genau verifiziert werden. In einer Untersuchung an der Uni Karlsruhe (Mönnich & Handreck,

2008) wurden die Ergebnisse von Literaturrecherchen in Fachdatenbanken zu vier Themengebieten aus dem Fächerangebot einer deutschen Universität mit den Ergebnissen von Google Scholar verglichen und unter dem Aspekt der Relevanz bewertet. Die Wissenschaftler/innen ziehen das Fazit, „dass trotz der erheblichen inhaltlichen Defizite anzunehmen ist, dass der Nutzerkreis von Google Scholar weiter zunehmen wird“. Für den Einstieg in eine Thematik oder eine ergänzende Nachrecherche ist Google Scholar in jedem Fall nützlich, auch wenn die bei Fachdatenbanken selbstverständliche Transparenz bei der Quellenauswertung und deren Qualität der bibliographischen Daten weitgehend fehlt. In früheren Untersuchungen (Mayr & Walter 2007; 2008) bemängeln die Verfasser/innen, dass Open-Access-Journals der Untersuchung zufolge bei Google Scholar unterrepräsentiert, manche Ergebnisse nicht sehr aktuell seien und nach wie vor das ‚alte Manko unklarer Quellen‘ bestehe.

Revision #2

Created 28 February 2025 21:12:22 by Bernd Grabner

Updated 13 February 2026 14:18:58 by Github Admin