

Wie Google funktioniert und was Google in der Websuche sonst noch so macht

	Beschreibung	Praxisbedeutung
Das Web durchsuchen	Google durchsucht das Netz kontinuierlich und speichert die gefundenen Seiten als Volltext auf eigenen Rechnern ab. Viele Seiten findet Google aber nicht – aus technischen, juristischen und anderen Gründen (siehe „Was Google alles nicht findet“).	Journalisten müssen wissen, dass Google nach wie vor nicht das ganze Internet abbildet und dass für sie wichtige Informationen auch im Deep Web liegen können.
Im Suchschlitz eingegebene Wörter werden innerhalb der Google-„Datenbank“ gesucht	Google sucht die eingegebenen Wörter <u>nicht live</u> im Internet, sondern lediglich in der auf den eigenen Rechnern gespeicherten Kopie des Internets – im sog. Index oder auch cache	Über das Cache lassen sich oft gerade veränderte oder aus dem Netz genommene Informationen noch finden/belegen.
Ergebnisse sortieren	Google sortiert die gefundenen Treffer nach einem Algorithmus. Die Hauptkriterien hierfür sind bekannt: Anzahl der Links, Page Rank der Site und der verweisenden Sites, Position und Häufigkeit des Suchbegriffs.	Das Ranking entspricht nicht ganz journalistischen Kriterien; insbesondere durch Suchmaschinenoptimierung.
Geld verdienen	Google verdient den Großteil seines Gewinns mit Anzeigen, die passend zu den Suchbegriffen angezeigt und pro Klick bezahlt werden	Anzeigeninhalte werden unabhängig von ihrer Relevanz angezeigt
Sortierung variiert mit Standort und gewählter Google-Site	Google wertet u.a. die IP-Adresse aus und passt die Suchergebnisse dem Standort an. Das verändert vor allem die Sortierung, aus juristischen Gründen aber mitunter auch die Anzahl der Treffer.	Durch das Löschen der Cookies, notfalls auch das Verwenden von Proxy-Diensten kann die Suche weitgehend neutralisiert werden.
Automatisch trunkieren	Google trunkiert automatisch – die Eingabe [journalistischen] liefert auch die Ergebnisse „Journalismus“ und „Journalist“	Wenn das nicht gewünscht ist, kann das mit dem Operator F oder „“ unterbunden werden.
Reihenfolge der Suchbegriffe berücksichtigen	Bei 2 oder 3 Suchbegriffen ändert Google die Sortierung der Ergebnisse. Einprägsames Beispiel: [Paris Hilton] vs. [Hilton Paris]	Reihenfolge variieren kann gute Treffer nach oben befördern
Ignoriert Groß- und Kleinschreibung	Google ignoriert Groß- und Kleinschreibung – lediglich die Operatoren OR und AROUND müssen groß geschrieben werden	



Operator	Bedeutung	Anwendung (mit Beispiel)
[„Suchbegriff“] (funktioniert auch mit [-„Suchbegriff“])	Suchbegriffe müssen in der vorgegebenen Reihenfolge stehen, keine anderen Wörter dazwischen	[„netzwerk recherche“] findet Texte, in denen „recherche“ unmittelbar nach „netzwerk“ steht; nicht aber die Wortfolge „netzwerk fördert Recherche“ gut für die Suche nach Zitaten, Eigennamen (Personen, Institutionen, Titel)
[Suchbegriff1 OR Suchbegriff2]	entweder der Begriff <u>vor</u> oder der Begriff nach dem Operator müssen in dem gesuchten Text vorkommen	[Benzin OR Diesel OR Autogas OR Gas OR Erdgas OR Super] findet Texte in denen einer dieser Begriffe vorkommt (OR muss groß geschrieben werden.)
[-Suchbegriff] (kein Leerzeichen!)	schließt den Suchbegriff aus	[“Marcus Lindemann“ –Bundesliga –Premiere –Sky] schließt Treffer zu meinem Namensvetter, dem Bundesliga-Kommentator, aus <u>Achtung:</u> Damit werden auch Treffer ausgeschlossen, in denen etwas über mich steht, aber einer der drei ausgeschlossenen Begriffe auftaucht!
[site: domain.tld] (funktioniert auch in der Variante [- site:])	findet Treffer ausschließlich von der genannten Domain	[site:bundestag.de] findet ausschließlich Treffer von dieser Domain: <u>Achtung:</u> [site:www.bundestag.de] liefert andere Ergebnisse! Tipp: Hinter der Toplevel-Domain lassen sich auch Unterverzeichnisse gezielt durchsuchen; ebenso kann man subdomains angeben. Beispiele: [site:mil], [site:heute.zdf.de], [site:bundestag.de/abgeordnete18]
[filetype: xxx]	findet ausschließlich Treffer des genannten Dateityps	z.B.: [filetype:pdf] findet ausschließlich pdf-Dateien funktioniert auch für doc, xls, ppt, rtf, ps, dwf, kml, kmz, swf Undokumentiert:txt,mp3, mov u.v.m. (unvollständig) – bei diesen Dateitypen werden aber nicht die Inhalte, sondern nur Dateinamen und Beschreibung durchsucht
[intitle: Suchbegriff], Analog: [inurl: Suchbegriff]	Suchbegriff muss im Titel der Seite vorkommen	Alle Treffer aus dem Print-Archiv des Spiegel enthalten die Wörter „Der Spiegel“; [intitle: „Der Spiegel“] – zusammen mit [site:spiegel.de} eine kluge Archivsuche.
[allinbody: Suchbegriff] ähnlich: allintitle: , allinurl:	alle nachfolgenden Begriffe müssen im entsprechenden Seitenteil vorkommen	[allintitle: Marcus Lindemann] findet Internetseiten, in denen der Name im Titel der Seite steht, das sind bei Namen, u.a. Profile in sozialen Netzwerken, Lebensläufe etc.
[inanchor: Suchbegriff]	durchsucht Linkbeschreibungen	



Operator	Bedeutung	Anwendung (mit Beispiel)
Nicht mit anderer Syntax kombinierbare Operatoren		
related: domain.tld (nicht bei Bing)	findet Sites, die der angegebenen ähnlich sind (d.h. in der gleichen Kategorie im Google-Verzeichnis liegen)	[Related:volkswagen.de] liefert eine sehr umfangreiche Liste von Automarken (sowie einige Seiten, die mit VW zu tun haben wie die Autostadt oder die VW-Bank).
link: domain.tld (nicht bei Bing)	Findet Sites, die auf die genannte Site verlinken	[link:martinlutherking.org] etwa findet Seiten, die auf diese Domain aus ganz unterschiedlichen Gründen verweisen: a) als Beispiel für eine Site, hinter der etwas anderes steckt als es den Anschein hat und b) Seiten, die das gut finden
Nicht in der erweiterten Suche dokumentierte Syntax: (außer IP-Suche nicht bei Bing)		
„Suchbegriff1 * Suchbegriff2“	Der Asterix (*) innerhalb einer Phrasensuche erlaubt, dass an der Stelle mehrere Wörter stehen dürfen (und mindestens ein Zeichen stehen muss).	[„Henryk * Broder“] findet „Henryk M. Broder“, „Henryk Mittelname Broder“ und auch „Henryk Modest Broder“ Achtung: Texte, in denen ausschließlich „Henryk Broder“ vorkommt, werden mit dieser Asterix-Suche nicht gefunden!
~Suchbegriff	Findet auch Synonyme zum Suchbegriff (funktioniert jetzt auch mit dt. Begriffen gut!)	[~jobs] findet auch „Arbeit“, „Stelle“, „karriere“ und „Stellenangebot“
+Suchbegriff	unterdrückt das ungewollte Trunkieren und die Berücksichtigung von anderen Schreibweisen	[+Photographie] findet nur Seiten, die das Wort „Photographie“ in dieser Schreibweise enthalten, während [Photographie] auch „Fotografie“ und „fotografisch“ findet.
cache:domain.tld	ruft das Cache der angegebenen Domain auf	cache:recherche-info.de führt direkt (ohne den Klick auf „im cache“) zu der im Google-Index gespeicherten Kopie der angegebenen Site
Zahl1..Zahl2 (numrange)	sucht neben den beiden angegebenen Zahlen auch alle ganzen Zahlen dazwischen	[10000..17000 Berlin] sucht nach allen Postleitzahlen in dem angegebenen Zahlenbereich
Suchbegriff1 AROUND (n) Suchbegriff2	findet 2 Suchbegriffe in einem vorgegebenen Maximalabstand (n-1= Anzahl der maximalen Wörter zwischen den beiden Suchbegriffen)	[Henryk] AROUND(2) Broder] findet sowohl „Henryk Broder“ als auch „Henryk M. Broder“
ip: ipadresse (nur bei Bing!)	Findet alle Seiten, die auf dem angegebenen Server liegen	[ip:82.165.68.224] findet alle Seiten auf dem angegebenen Server



Funktion	Bedeutung	Anwendungsmöglichkeiten
Die erweiterte Suche bietet zunächst Funktionen, für die es auch Operatoren gibt („“, -, OR, Dateiformat, Position, Domains/Site) – siehe Tabelle 2		
Sprache	Filtert nach Sprache	In der Praxis in der Regel ohne Bedeutung, da über die Suchbegriffe die Sprache in der Regel schon gewählt ist. Interessant nur in Kombination mit anderen Operatoren: etwa arabische Seiten auf .de-Domains oder mp3 in Katalanisch
Region	Filtert nach Regionen, in der Regel nach Ländern	auch hier entsteht der Nutzen erst aus der Kombination mit anderen Funktionen oder aber man recherchiert wie ein Thema in einem bestimmten Land behandelt wird. Sucht man z.B. [Assange] in Australien ist eine Unterstützer-Site weit oben.
Datum	Filtert nach Datum funktioniert in der NEWS Suche tadellos	Funktioniert mittlerweile ganz gut. Google scheint das Datum der ersten Indexierung einer Seite zu speichern. In keinem Fall lässt die Datumssuche einen Rückschluss auf die Aktualität der Informationen zu! <u>Achtung:</u> unter Suchoptionen auf der Ergebnisseite erhalten Sie die Option „ Zeitraum festlegen “! Nutzen: Bei Themen, die häufiger in der Presse sind, lassen sich Zeiträume, die nicht zum gewählten Aspekt gehören, rausfiltern (z.B.: BP vor „Deep Water Horizon“) <u>Zusatznutzen:</u> Sobald ein Zeitraum festgelegt ist, können Sie nach Datum statt nach Relevanz sortieren!
Nutzungsrechte	Filtert nach Inhalten, die unter creative commons Lizenzen stehen	Recherche in verschiedenen Abstufungen möglich. Anfrage beim Urheber bleibt in der Regel aber erforderlich, etwa, um zu klären, ob dieser die geplante Nutzung als „kommerziell“ bewertet.
Erweiterte Suche in anderen Google Services		
Bilder	Gesichter/Fotos/Clipart/Strichzeichnung	Filtert nach Bildertypen (weitere Filter: nach Farben)
Kamerasymbol in der Bildersuche	Bilder mit Bildern suchen	Sucht ohne sucht Begriffe anhand von Bildern, die man per URL eines Bildes oder hochgeladenem Foto vorgibt.
News	Archiv	Nach der Eingabe von Suchbegriffen lassen sich Zeiträume Tag genau definieren
source:	News und Scholar	Mit dem Operator source lassen sich gezielt Quellen auswählen.
author:	Scholar	Suche nach den Verfassern wissenschaftlicher Texte
Video		Umfangreicher als youtube! u.a. Suche nach Länge und Auflösung

