



2006/35 Thema

<https://jungle.world/artikel/2006/35/suchen-und-finden-lassen>

Suchen und finden lassen

Von **oona leganovic**

Es gibt kein objektives Pageranking, aber ohne ein solches geht es auch nicht. Die Gewichtung von Inhalten ist kein rein mathematisches Problem. von oona leganovic

Wer sucht, der findet, heißt ein altes Sprichwort. Das Wort »Suchmaschine« legt nahe, dass man suchen lässt – und damit auch das Finden aus der Hand gibt. Jedenfalls stimmt, dass für das Finden die Methode des Suchens entscheidend ist. Die Funktion von Internet-Suchmaschinen besteht darin, die über das World Wide Web verstreuten Informationen möglichst unkompliziert zugänglich zu machen. Zu diesem Zweck werden Webseiten katalogisiert, ihre Inhalte indiziert, sie werden gewichtet, und teilweise werden Kopien der Seiten zwecks besseren Zugriffs zwischengelagert. Bei Google nennt sich die Methode des Gewichtens »Pageranking«.

Der Pagerank wird durch einen patentierten Algorithmus festgestellt und beschreibt die Wahrscheinlichkeit, mit der ein zufällig durch die Gegend klickender Internet-Nutzer auf einer bestimmten Seite landen wird. Dabei wird die Linkstruktur des WWW als ausschlaggebender Faktor benutzt: Ein Link auf eine Seite funktioniert als eine »Stimme« bzw. Zustimmung für diese Seite, so dass stark verlinkte Seiten einen hohen Rang zugewiesen bekommen. Ein Nutzer landet eben eher dort, wo viele Wege hinführen. Allerdings zählt nicht jeder Link gleich: Links von Seiten, die ihrerseits einen hohen Rang haben, werden stärker gewertet. Der Algorithmus wurde von Larry Page (deshalb »Page«ranking), einem der beiden Gründer von Google, erfunden und bildet die Basis für den Erfolg des Unternehmens.

Von diesem Pageranking zu unterscheiden ist die Relevanz einer Seite in Bezug auf einen bestimmten Suchbegriff, die wiederum durch Textanalyse ermittelt wird. Hier geht es darum, wie oft und an welchen Stellen der Suchbegriff auf der Seite auftaucht und mit welchen Begriffen als sichtbarem Linktext diese Seite verlinkt wird: Wenn sich auf Seite A ein Link zu Seite B findet, dessen Text zum Beispiel »Schwulmulbel« lautet, wird nicht nur der Pagerank von Seite B erhöht, sondern auch ihre Relevanz für den Suchbegriff »Schwulmulbel«. Das Suchergebnis, das dem Nutzer dann geliefert wird, kommt also durch eine Kombination dieser Kriterien zustande.

Je erfolgreicher das Unternehmen Google wird, desto wichtiger wird seine Bewertung einer Seite für deren Wahrnehmung. So kann eine Seite für viele Nutzer praktisch unsichtbar sein, wenn sie beim Googeln nicht angezeigt wird oder in der Trefferliste auch nur weit unten erscheint. Die enorme Bedeutung einer Suchmaschine wie Google beruht auf dem gegenseitigen Bedürfnis von denen, die suchen, und denen, die gefunden werden wollen.

Wer im Internet publiziert, will auch schnell auffindbar sein. Kein Wunder also, dass diejenigen, die gefunden werden wollen, sich allerlei einfallen lassen, um im Ranking höher zu rutschen, sei es nun durch Linkkauf, Kommentar- und Forenspam oder, wie Anfang dieses Jahres BMW, durch schlüsselwortgespickte Köderseiten, die möglicherweise nichts mit denen zu tun haben, die den Nutzern präsentiert werden. Die Firma Google muss schon allein zur Qualitätssicherung gegen diese Methoden vorgehen, und deshalb ist es gar nicht möglich, sich auf den Algorithmus zu verlassen. Kein Wunder ist es auch, dass die Mechanismen der Bewertung, soweit bekannt, sich für allerlei Tricks ausnutzen lassen, bekannt unter anderem als »Google-Bombing«.

Ein Beispiel ist die bewusst inszenierte Verdrängung der antisemitischen Seite »jewwatch« von Platz Eins für die Suchanfrage »Jew« im Jahre 2004, und der Umstand, dass das Suchwort »miserable failure« als ersten Treffer eine Seite des Internetauftritts des Weißen Hauses liefert. Google scheint allerdings keine Anstrengungen zu unternehmen, diese Art der Manipulation einzudämmen – einerseits, weil die Wirkungen meist nicht von Dauer sind und die »manipulierten« Treffer häufig von Berichten über eben diese Aktion verdrängt werden, andererseits, weil diese Art der Manipulation eben auch eine »reale« Wertung darstellt.

Aber da das Unternehmen Google inzwischen über die relative Sichtbarkeit oder Unsichtbarkeit einer Seite entscheidet, wird es immer mehr als Zensurinstanz wahrgenommen. Scientology etwa wollte eine kritische Seite aus den Suchergebnissen entfernen lassen, die Modemarke Vuitton verklagte Google in Frankreich erfolgreich, weil die Seiten von Produktfälschern angezeigt wurden. Staatliche Zensur wie die rechtsradikaler Seiten in verschiedenen europäischen Ländern oder ganzer Themenfelder wie im Iran und in China will Google nicht umgehen, sondern unterstützt sie.

Ein grundlegendes Problem ist dabei, dass Google mit der Zustimmung zu zensierenden Maßnahmen sich selbst im großen Maßstab einer wichtigen Verteidigungslinie gegen Zensurbemühungen beraubt hat: Das Unternehmen kann nun nicht mehr behaupten, eine grundlegende Zensur sei nicht möglich, da das Such- und Bewertungsverfahren vollständig automatisiert sei.

Die reale Machtkonzentration reizt zu Verschwörungstheorien. Die Vorwürfe reichen von der Annahme, dass Google heimlich die Suchergebnisse manipuliere, bis zu der Beschreibung von Google als »Krake«, die die totale Überwachung und Kontrolle organisiere. Dabei muss mal der menschliche Eingriff, mal das automatisierte Verfahren als Verdachtsmoment herhalten. Und einige deutsche Rechtsradikale, für die Google schon allein deswegen unter Generalverdacht steht, weil einer der Gründer Jude ist, fühlen sich gar persönlich verfolgt (wie z.B. der Bund für echte Demokratie, obwohl dessen Seiten über google.de problemlos zu finden sind).

Man kann Google alles Mögliche vorwerfen, sei es der bedenkliche Erwerb und Umgang mit den Daten der Nutzer oder der angeblich »unfaire« Suchalgorithmus – man ist aber nicht gezwungen, Google zu benutzen: Neben anderen Suchmaschinen wie Yahoo oder MSN und Metasuchmaschinen wie MetaCrawler, die es einem ermöglichen, die Suchergebnisse verschiedener Suchmaschinen zu vergleichen, gibt es außerdem auch die auf Google basierende Seite scroogle.org, die einem die Ergebnisse von Google ohne die einschlägigen Cookies liefert. »Cookies« sind jene kleinen Textdateien, mit deren Hilfe Google seine Nutzer identifiziert und ihre Daten sammelt.

Es ist überdies anzunehmen, dass Google als Suchmaschine den Zenit ihrer Bedeutung erreicht hat und in Zukunft andere Methoden der Informationsorganisation im Netz zunehmen werden, vor allem solche, die sich eben doch wieder auf die Bewertung durch Menschen statt Maschinen stützen. Nur eben durch sehr viele Menschen, die dafür auch gar nicht bezahlt werden, wie beim »social bookmarking« etwa des Dienstes del.icio.us, die das Mitteilungsbedürfnis der Nutzer dazu einsetzt, Inhalte zu klassifizieren. Womit sich wiederum ganz neue Missbrauchsmöglichkeiten und Datenschutzprobleme auftun. Auch hier, wie im Falle sowohl von Google als auch von staatlichen Maßnahmen, besteht ein wesentliches Problem darin, dass sich die meisten Nutzer nicht für Datenschutz interessieren. Auch deshalb, weil sich so wenige Menschen bewusst machen, wem sie eigentlich wann für wie lange welche Daten zur Verfügung stellen wollen, werden immer mehr Daten gesammelt und gespeichert.