2012/46 Lifestyle

https://jungle.world/artikel/2012/46/hide-and-seek

Presseverleger wollen ein Stück vom großen Google-Kuchen

Hide and Seek

von **Boris Mayer**

Verlage sollen an den Milliardengewinnen der Suchmaschine Google teilhaben. Geplant ist ein Leistungsschutzrecht für Presseverleger. Dabei gäbe es eine einfache technische Lösung: Wer sich wirklich nicht finden lassen möchte, kann sich im Internet prima vor Google verstecken.

In Frankreich ist es im Streit um eine Suchmaschinenabgabe zu einem regelrechten Machtkampf zwischen Google und den Zeitungsverlagen, ja im Grunde zwischen Google und der französischen Regierung gekommen. Auch hierzulande wollen Verlage mit dem Mittel des Leistungsschutzrechts etwas von Googles Gewinnen abbekommen, weshalb die Regierung dem Bundestag ein entsprechendes Gesetz vorgelegt hat.
Unter den Begriff Leistungsschutzrecht fallen Rechte, die eine Beziehung oder Ähnlichkeit zum Urheberrecht haben, aber wegen einer fehlenden individuellen Gestaltung nicht unter das Urheberrecht fallen. Vom Leistungsschutzrecht geschützt sind zum Beispiel die Produzenten von Tonträgern wie Schallplatten oder CDs sowie Filmproduzenten mit ihren DVDs, aber auch Sendeunternehmen. Grundlage dieser Leistungsschutzrechte ist die Urheberrechtsrichtlinie der EU aus dem Jahr 2001 – im Prinzip die europäische Version des im Internet wohl deutlich bekannteren Digital Millenium Copyright Act, kurz DMCA.
Nicht erfasst sind dabei bisher die Schutzrechte von Presseverlegern, deren Produkte aber unzweifelhaft die gleiche Ähnlichkeit zu urheberrechtlich geschützten Werken haben wie die Produktion einer CD oder DVD. Deshalb forderte im Mai 2009 der Bundesverband Deutscher Zeitungsverleger, gerade im Online-Bereich nicht schlechter als andere Leistungsvermittler gestellt zu werden.

Es ist unbestritten, dass dies ein berechtigtes Anliegen ist. Die ersten Ideen zu einem solchen Gesetz stießen jedoch auf große Kritik, denn viele Blogger befürchteten regelrechte Abmahnwellen durch einschlägige Rechtsanwaltskanzleien. Deshalb geht es in der letzten, im August vom Bundeskabinett verabschiedeten Fassung des Gesetzentwurfs nur noch um die gewerbliche Nutzung durch Suchmaschinenbetreiber und Newsaggregatoren. Ganz ähnliche Debatten und Gesetzesvorhaben gibt es nicht nur in Deutschland, sondern auch in anderen Staaten, zum Beispiel in Frankreich und Italien. Auch außerhalb der EU, etwa in Brasilien, sind entsprechende gesetzliche Regelungen in Planung.
In Frankreich wird vor allem über eine Suchmaschinenabgabe diskutiert. Funktionieren soll sie so: Werden Auszüge aus den Websites von Verlegern angezeigt, soll das die jeweilige

Suchmaschine betreibende Unternehmen aus seinen Werbeeinnahmen einen gewissen Betrag den jeweiligen Verlagen zahlen – wenn sich beide Seiten nicht von sich aus einigen. Google reagierte auf diese Pläne äußerst harsch und kündigte in einem Brief an die französische Regierung an, bei Einführung einer Leistungsschutzabgabe französische Verlage aus den Ergebnissen einfach auszuklammern. Beide Seiten nehmen einander anscheinend sehr ernst, denn inzwischen hat sich sogar Frankreichs Präsident François Hollande mit Googles CEO Eric Schmidt zu einem persönlichen Gespräch in Paris getroffen. Die französische Regierung will bei einer solchen offenen Drohung nicht erpressbar zu wirken.
Google kann sich die Arroganz, die aus der Drohung des Unternehmens gegen die Regierung eines EU-Landes spricht, leisten. Der Suchmaschinenbetreiber ist schließlich derart wichtig, dass eine ganze (wenn auch allgemein verachtete) Branche entstanden ist: SEO-Berater (Search Engine Opimization) beschäftigen sich damit, Websites durch verschiedene Tricks so zu optimieren, dass sie die bestmögliche Platzierung in den Suchergebnissen erreichen. Und so sieht der Streit ums Leistungsschutzrecht für den Beobachter zunächst so aus, als stünden sich die Presseverlage auf der einen Seite und Google auf der anderen Seite gegenüber. Andere Nachrichtenaggregatoren wie die Yahoo News fallen zumindest auf den ersten Blick nicht ins Gewicht.
Das ist nicht von ungefähr so. Google erzielt über die Internetwerbung Milliardengewinne und die Presseverleger möchten an diesen Einnahmen teilhaben. Google, so ihre Argumentation, profitiere im Grunde einseitig von ihrer Arbeit, »ohne eine angemessene Gegenleistung zu erbringen«. So einfach ist die Sache allerdings nicht, wie die Nachfrage nach SEO-Dienstleistungen zeigt: Die Suchmaschine bringt immensen Zulauf auf die Websites der Verlage und Zeitungen, die unter News gelistet werden, Google spricht von »weltweit jeden Monat vier Milliarden Klicks«. (Google hat sich mit seiner Rubrik »News« – sicherlich ungewollt – übrigens als Dokumentarist der nicht mehr an Journalismus, sondern an Content interessierten Arbeitsweise vieler Verlage erwiesen: Sucht man nach einem aktuellen Stichwort, findet man oft 40, 50 und mehr bis hin zur Überschrift gleichlautende Artikel – Agenturmeldungen zu veröffentlichen ist schließlich billiger als journalistische Eigenrecherchen).
Verleger und Suchmaschinenbetreiber beharren auf dem, was sie für ihr Recht halten. Mit der Folge, dass sich nichts bewegt – genau das aber will die französische Regierung unter Präsident Hollande ändern, daher auch die Drohung, eine Abgabe gesetzlich zu regeln, wenn sich die beiden Parteien nicht einigen.

Eine technische Lösung des Problems gibt es jedoch bereits. In der Frühzeit des Internets, als Google noch nicht einmal gegründet worden war, fand die heute so viele Schlagzeilen produzierende Diskussion nämlich schon einmal statt – damals allerdings von der Öffentlichkeit vollkommen unbemerkt.
Im Jahr 1994 war mit »WebCrawler« die erste Volltextsuchmaschine erfunden worden. Um im kompletten Text einer Webpage suchen zu können, musste die Suchmaschine die Seiten komplett einlesen und bei sich so abspeichern, dass Nutzer auf die Eingabe von Stichworten die Seitenvorschläge angezeigt bekommen konnten. Der Begriff »Crawler« oder »Webcrawler« wird noch heute benutzt und steht für ein Programm, das auf einem Server eines Suchmaschinenbetreibers läuft und nichts anderes tut, als Verlinkungen in Websites zu verfolgen, die Zielseite einzulesen, der Datenbank hinzuzufügen und alle darin enthaltenen Verbindungen weiterzuverfolgen.
Zunächst passierte dies nur auf einer einzigen Website und ihren Unterseiten, aber schon bald wurden auch externe Links verfolgt, also solche, die zu anderen Websites führten. Webcrawler

surften also automatisch das ganze mit Links verbundene Internet ab und früher oder später tauchte dadurch die eigene Seite in den Suchergebnissen auf – wenn sich irgendwo in den der Suchmaschine bereits bekannten Seiten ein Link auf sie fand. Die Formulare, mit denen man bis dahin die eigene Website den jeweiligen Suchmaschinen bekannt machen musste, dienten nun nur noch dazu, die Zeit zu verkürzen, bis die jeweilige Seite eingelesen wurde.
Damals stellte sich die Frage, wie man verhindern könne, dass die Suchmaschine bestimmte Seiten einliest – nicht auf Initiative von Website-Betreibern hin, die nicht gefunden werden wollten, sondern auf Wunsch des Unternehmens, das WebCrawler betrieb. Die Lösung dafür hieß »robots.txt« und wurde von Martijn Koster, einem Mitarbeiter bei WebCrawler, entwickelt. Legte man eine solche Datei in das Hauptverzeichnis der eigenen Internetseite ab, so konnte man hier Verbote aussprechen – und das funktioniert auch heute noch. Inzwischen heißt das Konstrukt »Robots Exclusion Standard«, benutzt aber immer noch dieselbe Datei auf den Webservern. In dieser Datei kann man einzelne Crawler – die in der Datei User-Agents heißen – oder auch einfach alle solchen Roboter-Surfer ausschließen. Man kann auch nur einzelne Unterverzeichnisse oder einzelne Seiten so konfigurieren, dass die Crawler sie ignorieren sollen.
Populär wurde das Konzept der robots.txt dann durch die nächste Generation von Suchmaschinen und ihren Crawlern, dem 1994 gestarteten Lycos und dem ein Jahr später gestarteten Altavista, die dann erst Ende der neunziger Jahre in ihrer Popularität von Google überholt wurden.
Zwar müssen sich die Crawler theoretisch nicht an die robots.txt halten – aber gerade Google hält sich daran und erlaubt sogar noch zusätzliche Möglichkeiten. Man kann Google komplett ausschließen (mit User-Agent: googlebot) oder aber auch nur einzelne Dienste von Google – Google-News zum Beispiel mit »User-Agent: googlebot-news«. In der Zeile darunter muss nur noch stehen, welche Unterverzeichnisse man verbieten möchte oder ob es alles sein soll.

Woran liegt es nun, dass die Presseverlage eine solche »robots.txt« eben nicht auf ihren Websites haben, wenn doch Google sich angeblich durch das Verwenden der eigenen Daten nur unfair und einseitig bereichert? Auch der Politik muss die Frage gestellt werden, wieso man gleich zu einer verpflichtenden Abgabe greifen muss, anstatt einfach die Beachtung der robots.txt den Betreibern von Crawlern gesetzlich vorzuschreiben – auch wenn sich damit faktisch nichts ändern würde, denn zumindest Google, Microsoft und Yahoo beachten diese ja bereits.
Eine Modernisierung des Mechanismus wäre im Übrigen problemlos möglich, aber darauf müsste man sich erst international verständigen – das Internet ist schließlich ein globales Netz. Es könnte auch eine Funktion eingebaut werden, die für das Einlesen einer Website durch einen Crawler – und das anschließende Zur-Verfügung-Stellen in Suchmaschinen – einen bestimmten Betrag pro Jahr. Wobei die Crawlerbetreiber dann entscheiden könnten, ob sie das Angebot für diese Website annehmen.

© Jungle World Verlags GmbH