

## Die Qualität von Suchmaschinen auf dem Prüfstand – eine vergleichende Inhaltsanalyse der Trefferlisten von Google, Bing, Ask, Lycos und Ixquick

„Gabriel gegen Google“ – so könnte man den am 16. Mai in der FAZ erschienenen Artikel des amtierenden Wirtschaftsministers überspitzt zusammenfassen. Darin warnt dieser nicht nur vor Google als „Datenkrake“, sondern kündigt auch politische Schritte gegen den Internetriesen an: So werde das Bundeskartellamt prüfen, ob Google seine marktbeherrschende Stellung missbrauche, um Wettbewerber „systematisch“ zu verdrängen. Wichtig sei das „Gebot der Nichtdiskriminierung von alternativen Anbietern, die Platzhirsche innovativ herausfordern“ (Gabriel 2014).

Dass Google auf dem deutschen Suchmaschinenmarkt eine Monopolstellung innehat, lässt sich nicht bestreiten. Bereits die Tatsache, dass das Wort „googeln“ mittlerweile ganz selbstverständlich in den deutschen Sprachgebrauch übergegangen ist und sogar in den Duden aufgenommen wurde, spiegelt die hohe gesellschaftliche Bedeutung des US-Konzerns hierzulande wider. Diese zeigt sich auch in den Nutzerstatistiken: Über 90 Prozent aller Suchanfragen in Deutschland<sup>1</sup> entfallen auf Google (Statista 2014a,b,d; SEO United 2014). Googles Erfolg zeigt sich auch am immer weiter steigenden jährlichen Umsatz und Gewinn: 2013 lag der weltweite Umsatz bei fast 60 Milliarden US\$, der Nettogewinn bei fast 13 Milliarden US\$ (Statista 2014c,e).

Eine maßgebliche Ursache dieser enormen Marktkonzentration ist das Verhalten der Nutzer: Die meisten greifen nur auf diese eine Suchmaschine zurück, weil sie sie schlichtweg für die beste halten (Anonymisiert 2014). Dadurch verleihen sie einem gewinnorientierten Unternehmen, das im Unterschied zu den traditionellen Massenmedien nicht an eine öffentliche Aufgabe rückgebunden ist, einen enormen gesamtgesellschaftlichen Einfluss, womit eine große gesellschaftliche Verantwortung für das Unternehmen einhergeht. In der Folge stellt sich – eng verbunden mit den medienökonomischen Befürchtungen einer Wettbewerbsverzerrung – die medienpolitische Frage, wie es um die Qualität des Angebots von Google bestellt ist: Wird Google durch die Erfüllung bestimmter Qualitätsmaßstäbe seiner gesellschaftlichen Verantwortung gerecht? Und ist Googles Angebot tatsächlich qualitativ hochwertiger als das der Mitbewerber?

Dies untersucht die vorliegende Untersuchung mittels eines inhaltsanalytischen Vergleichs der Trefferlisten von fünf Suchmaschinen (Google, Bing, Ixquick, Lycos und Ask<sup>2</sup>) hinsichtlich vier klassischen kommunikationswissenschaftlichen Kriterien für die Qualität von Medieninhalten im Internet: Vielfalt, Aktualität, Relevanz und Nutzerfreundlichkeit. Die Studie leistet damit auch einen Beitrag zu der Frage, wie die Qualität von Online-Inhalten gemessen werden kann, worüber derzeit noch kein Konsens besteht (Neuberger 2012: 40).

Der Begriff „Qualität“ beschreibt an sich wertneutral die Eigenschaften eines Objekts. Oft ist er aber positiv konnotiert, weil damit die Erfüllung bestimmter (hoher) Standards assoziiert ist (Hohlfeld 2003: 205). Rosengren, Carlsson und Tagerud (1991: 22) bezeichnen die Qualität aus kommunikationswissenschaftlicher Perspektive als „one or more characteristics satisfying certain standards backed up by more or less central values and norms“. Es handelt sich also um zugeschriebene Eigenschaften, basierend auf einem Werthorizont, der eng mit der öffentlichen Aufgabe traditioneller Massenmedien verbunden ist. Diese gesellschaftliche Dimension von Medienqualität wurde in der bisherigen Forschung zur Qualität von Suchmaschinen vernachlässigt. Entsprechende Studien stammen vor allem aus den Informationswissenschaften, die sich auf rein technische Qualitätsmerkmale beschränken (z.B. Vaughan 2004; Lewandowski & Hochstötter 2007). In einer der ersten kommunikationswissenschaftlichen Studien unter-

<sup>1</sup> Weltweit ist Googles Vormachtstellung – bedingt durch die Vormachtstellung der Suchmaschine Baidu in China – zwar nicht ganz so stark, aber immer noch deutlich: Im März 2014 suchten ca. 70 Prozent der Onliner weltweit hauptsächlich mit Google (Statista 2014a).

<sup>2</sup> Suchmaschinen wie Yahoo oder T-Online wurden nicht untersucht, da diese nur die Ergebnisse von Google oder Bing anzeigen.

sucht Beiler (2013) Trefferlisten mehrerer Suchmaschinen inhaltsanalytisch, beschränkt sich dabei aber auf die Suche nach Nachrichteninhalten, was nicht der Nutzungsrealität der meisten User entspricht (Anonymisiert 2014).

Um diese Nutzungsrealität besser abzubilden, wurden für die vorliegende Studie sechs Suchanfragen bzw. Themen ausgewählt, die langfristige gesellschaftliche Probleme bzw. dauerhaft relevante Themen widerspiegeln und deren Ergebnisse durch den Zeitpunkt der Abfrage – beispielsweise durch starke Präsenz aktueller Themen in den traditionellen Massenmedien – nicht stark verzerrt werden dürften: Stress bei Schulkindern, Mindestlohn, neuer Rundfunkbeitrag, Organspende, Homo-Ehe und Videoüberwachung in Deutschland. Zum einen wurde mittels simplen Ein- bzw. Zweiwort-Suchanfragen nach diesen Themen gesucht. Zum anderen wurden zu vier der sechs Themen komplexere Suchanfragen formuliert, um den Einfluss von Suchoperatoren bzw. spezifischeren Suchanfragen aus mehr als zwei Suchbegriffen auf das Suchergebnis zu prüfen. Somit ergeben sich zehn verschiedene Suchanfragen und entsprechende Trefferlisten pro Suchmaschine (n = 50 Trefferlisten). Orientiert am Verhalten der Nutzer, die in aller Regel nur die erste Ergebnisseite von Trefferlisten ansehen (Anonymisiert 2014), beschränkt sich die Analyse auf ebendiese, die jeweils meist zehn einzelne Treffer umfassen (n = 511 Treffer). Die Untersuchung wurde im Juni (Pretests) und Juli (Hauptuntersuchung) 2013 durchgeführt. Um größtmögliche Vergleichbarkeit zwischen den Suchmaschinen herzustellen, wurden die Suchanfragen am 9. Juli 2013 in allen Suchmaschinen zeitgleich ausgeführt und die Trefferlisten lokal gespeichert (nach vorheriger Löschung der Cookies, um keine personalisierten Ergebnisse zu erhalten). Geprüft wird auf Ebene (1) der Trefferlisten und (2) der einzelnen Treffer, wie gut die vier Angebote kommunikationswissenschaftliche Qualitätskriterien erfüllen, die für Suchmaschinen besonders relevant sind:

*Vielfalt* meint die Sicherstellung einer möglichst großen Breite an Themen und Meinungen (Maurer 2005: 93). Dieses Kriterium könnte durch die Monopolstellung von Google besonders stark in Gefahr geraten. Der durch die Marktkonzentration bedingte Mangel an externer Vielfalt könnte aber durch interne Vielfalt in den Trefferlisten zumindest teilweise ausgeglichen werden. Da es sich bei den Suchanfragen nicht um große und vielfältige Themenkomplexe handelt, sollte eine Suchmaschine viele verschiedene Arten von Seiten anbieten – also z.B. sowohl Online-Ableger von Massenmedien als auch Portale, Blogbeiträge und Internetlexika. Die darauf basierende Gattungsvielfalt der Trefferlisten wird mittels des Maßes der standardisierten Entropie operationalisiert.

*Aktualität* beschreibt den zeitlichen Abstand zwischen dem Entstehungszeitpunkt einer Information und dem Zeitpunkt ihrer Publikation bzw. Abrufbarkeit (Gräfe 2005: 29). Weil aktuelle Informationen im Internet eine sehr große Rolle spielen, sollte dieses Kriterium auch für Suchmaschinen gelten. Bestimmt wird die Aktualität der einzelnen Treffer anhand ihrer Datumsangabe.

*Relevanz* bezieht sich darauf, ob eine Information für den Nutzer aus dessen subjektiver Sicht einen Mehrwert besitzt (Beiler 2013: 114ff.; Gräfe 2005: 28) – für viele Nutzer eines der entscheidendsten Kriterien, um die Leistung und Qualität von Suchmaschinen zu bewerten (Anonymisiert 2014). Aufgrund der Subjektivität dieses Kriteriums ist es auf einer allgemeinen Ebene schwer zu messen. Daher wurde in der Analyse mit Suchszenarien gearbeitet, welche die Informationsbedürfnisse eines fiktiven Durchschnittsnutzers simulierten. Als Relevanzindikatoren dienen in technischer Hinsicht das Vorkommen der Suchbegriffe im Titel oder Teasertext des Treffers sowie der logische Zusammenhang zwischen den Suchwörtern, in inhaltlicher Hinsicht die Erfüllung von Nutzererwartungen gemäß der Szenarien.

*Nutzerfreundlichkeit* als onlinespezifisches Qualitätskriterium steht in Verbindung mit der formalen Professionalität, die u.a. eine Orientierungsfunktion für den Nutzer fordert (Fahr 2001: 40). Besonders nutzerfreundlich ist eine Suchmaschine, wenn der Nutzer möglichst schnell das findet, wonach er ge-

sucht hat. Hilfreich sind hierbei z.B. fett gedruckte Suchwörter (Anonymisiert 2014), Suchvorschläge und ein möglichst geringer Werbeanteil.

Für jedes dieser vier Qualitätskriterien wird ein Index errechnet. Die vier Indexwerte wiederum bilden die Grundlage für einen übergreifenden Qualitätsindex mit einem Wertebereich von 0 (= sehr geringe Qualität) bis 1 (= sehr hohe Qualität). Die Indexbildung ermöglicht, verschiedene Suchmaschinen hinsichtlich der einzelnen Qualitätskriterien und ihrer Qualität insgesamt zu vergleichen, und lässt Aussagen darüber zu, wie nah die einzelnen Suchmaschinen einer normativen Idealvorstellung von Qualität kommen. Diese anschauliche Darstellung des abstrakten Begriffs „Suchmaschinenqualität“ stellt eine methodische Innovation dar.

Die Ergebnisse zeigen zum Teil recht deutliche Qualitätsunterschiede zwischen den einzelnen Suchmaschinen (Tabelle 1): Hinsichtlich der *Aktualität* enthält Google (abgesehen von drei Treffern bei Bing) als einzige Suchmaschine Treffer, die vom selben Tag oder aus derselben Woche stammen wie die Suchanfrage. Bei der *Vielfalt* schneidet Google etwas schlechter ab als die Wettbewerber, die Indexwerte bewegen sich aber bei allen Suchmaschinen auf einem hohen Niveau, weshalb bei keiner von geringer Vielfalt gesprochen werden kann. Die *Relevanz* der Google-Trefferlisten ist sowohl in technischer als auch inhaltlicher Hinsicht sehr hoch, doch Bing und Lycos erreichen vergleichbare Indexwerte. Ein Vorsprung von Google zeigt sich aber bei der *Nutzerfreundlichkeit*, bei der lediglich Ixquick eine vergleichbare Qualität bei den fettgedruckten Suchbegriffen erreicht.

Fasst man die vier Indexwerte zum Gesamtindex zusammen, ergibt sich für Google eine signifikant höhere Qualität als für die vier anderen Suchmaschinen. Das bedeutet, dass die Nutzer auch aus kommunikationswissenschaftlicher Sicht das beste Angebot präferieren, was zumindest teilweise erklärt, wie es zu der Vormachtstellung von Google auf dem Suchmaschinenmarkt kommt. Um mögliche Zusammenhänge zwischen der Marktkonzentration und dem inhaltlichen Angebot von Google beurteilen zu können, kann die vorliegende Untersuchung aber nur ein erster Schritt sein. Weiterführende Inhaltsanalysen dazu sollten beispielsweise die Vielfalt der verlinkten Ergebnisse untersuchen, um beurteilen zu können, ob die Dominanz eines bestimmten Anbieters auch zu einer Verengung der Perspektiven im öffentlichen Diskurs führen kann.

## Literatur

- Autoren (2014). Sammelbandbeitrag.
- Beiler, M. (2013). Nachrichtensuche im Internet. Inhaltsanalyse zur journalistischen Qualität von Nachrichtensuchmaschinen. Konstanz: UVK.
- Fahr, A. (2001). Katastrophale Nachrichten? Eine Analyse der Qualität von Fernsehnachrichten. München: Reinhard Fischer.
- Gabriel, S. (16. Mai 2014). Unsere politischen Konsequenzen aus der Google-Debatte. In: faz.net. Verfügbar unter: [http://www.faz.net/aktuell/feuilleton/debatten/die-digital-debatte/sigmar-gabriel-konsequenzen-der-google-debatte-12941865.html?printPagedArticle=true#pageIndex\\_2](http://www.faz.net/aktuell/feuilleton/debatten/die-digital-debatte/sigmar-gabriel-konsequenzen-der-google-debatte-12941865.html?printPagedArticle=true#pageIndex_2) [17.07.2014].
- Gräfe, G. (2005). Informationsqualität bei Transaktionen im Internet. Eine informationsökonomische Analyse der Bereitstellung und Verwendung von Informationen im Internet (Business-to-Business-Marketing). Wiesbaden: Deutscher Universitäts-Verlag.
- Hohlfeld, R. (2003). Objektivierung des Qualitätsbegriffs: Ansätze zur Bewertung von Fernsehqualität. In H.-J. Bucher & K.-D. Altmeppen (Hrsg.), Qualität im Journalismus: Grundlagen – Dimensionen – Praxismodelle (S. 203-221). Wiesbaden: Westdeutscher Verlag.
- Lewandowski, D. & Hochstötter, N. (2007). Qualitätsmessung bei Suchmaschinen – system- und nutzerbezogene Evaluationsmaße. Informatik Spektrum, 30 (3), 159-169.
- Maurer, T. (2005). Fernsehnachrichten und Nachrichtenqualität: Eine Längsschnittstudie zur Nachrichtenentwicklung in Deutschland. München: Reinhard Fischer.

- Neuberger, C. (2012). Journalismus im Internet aus Nutzersicht. Ergebnisse einer Onlinebefragung. *Media Perspektiven*, 1, 40-55
- Rosengren, K.E., Carlsson, M. & Tagerud, Y. (1991). Quality in Programming: Views from the North. *Studies of Broadcasting*, 27, 21-80.
- SEO United (2014). Suchmaschinenverteilung in Deutschland. Verfügbar unter: <http://www.seo-united.de/suchmaschinen.html> [17.07.2014].
- Statista (2014a). Marktanteile der Suchmaschinen weltweit und in Deutschland im März 2014. Verfügbar unter: <http://de.statista.com/statistik/daten/studie/222849/umfrage/marktanteile-der-suchmaschinen-weltweit/> [17.07.2014].
- Statista (2014b). Marktanteile von Suchmaschinen bei der Desktop-Suche und bei der mobilen Suche in Deutschland im Mai 2014. Verfügbar unter <http://de.statista.com/statistik/daten/studie/301012/umfrage/marktanteile-der-suchmaschinen-und-marktanteile-mobile-suche/> [17.07.2014].
- Statista (2014c). Nettogewinn von Google in den Jahren von 2001 bis 2013 (in Millionen US-Dollar). Verfügbar unter: <http://de.statista.com/statistik/daten/studie/76456/umfrage/nettogewinn-von-google-seit-2001/> [17.07.2014].
- Statista (2014d). Suchmaschinenverteilung in Deutschland im Juni 2014. Verfügbar unter: <http://de.statista.com/statistik/daten/studie/167841/umfrage/marktanteile-ausgewaehlter-suchmaschinen-in-deutschland/> [28.07.2014].
- Statista (2014e). Umsatz von Google weltweit in den Jahren von 2002 bis 2013 (in Milliarden US-Dollar). Verfügbar unter: <http://de.statista.com/statistik/daten/studie/74364/umfrage/umsatz-von-google-seit-2002/> [17.07.2014].
- Vaughan, L. (2004). New Measurements for Search Engine Evaluation Proposed and Tested. *Information Processing & Management* 40, 677-691.

## Anhang

**Tabelle 1: Qualität von Google, Bing, Ixquick, Ask und Lycos – Indexwerte**

Mittelwerte	Suchmaschine					Gesamt
	Google (A)	Bing (B)	Ixquick (C)	Ask (D)	Lycos (E)	
<b>Index Aktualität</b>	.25 <sup>B,C,d,E</sup> (n=107)	.05 (n=104)	.01 (n=100)	.14 <sup>b,C,E</sup> (n=100)	.0025 (n=100)	.09 (n=511)
<b>Index Vielfalt</b>	.71 (n=107)	.75 (n=104)	.83 (n=100)	.90 (n=100)	.79 (n=100)	.79 (n=511)
Inhaltliche Relevanz	.91 <sup>C,D</sup> (n=107)	.85 (n=100)	.75 (n=95)	.73 (n=98)	.86 <sup>d</sup> (n=97)	.82 (n=497)
Anteil an vorkommenden Suchbegriffen	.95 <sup>D</sup> (n=107)	.96 <sup>D</sup> (n=104)	.93 (n=100)	.85 (n=100)	.96 <sup>D</sup> (n=100)	.93 (n=511)
Anteil an zusammenhängenden Suchbegriffen	.48 (n=107)	.46 (n=104)	.38 (n=100)	.36 (n=100)	.51 (n=100)	.44 (n=511)
<b>Index Relevanz</b>	.78 <sup>c,D</sup> (n=107)	.75 <sup>D</sup> (s=.19) (n=100)	.69 (s=.22) (n=95)	.64 (s=.23) (n=98)	.78 <sup>c,D</sup> (s=.20) (n=97)	.73 (s=.21) (n=497)
Anteil der relevanten an den gesamten Suchvorschlägen	.54 <sup>B,C,D,E</sup> (n=107)	.30 <sup>C,E</sup> (n=104)	.00* (n=100)	.32 <sup>C,E</sup> (n=100)	.00* (n=100)	.24 (n=511)
Anteil an fettgedruckten Suchbegriffen	.81 (n=107)	.72 <sup>c</sup> (n=104)	.83 (n=100)	.00* (n=100)	.00* (n=100)	.48 (n=511)
Index Werbung	.92 <sup>B,C,D,E</sup> (n=107)	.62 <sup>C,D</sup> (n=104)	.48 <sup>D</sup> (n=100)	.27 (n=100)	.75 <sup>b,C,D</sup> (n=100)	.61 (n=511)
<b>Index Nutzerfreundlichkeit</b>	.76 <sup>B,C,D,E</sup> (n=107)	.55 <sup>C,D,E</sup> (n=104)	.44 <sup>D,E</sup> (n=100)	.20 (n=100)	.25 <sup>D</sup> (n=100)	.44 (n=511)
<b>Qualitätsindex gesamt</b>	.62 <sup>B,C,D,E</sup> (n=107)	.53 <sup>c,D,E</sup> (n=100)	.49 <sup>E</sup> (n=95)	.47 (n=98)	.46 (n=97)	.52 (n=497)

<sup>Aa</sup> Die Buchstaben hinter den Werten zeigen an, zu welchem Wert der betreffende Wert sich signifikant unterscheidet. Bei Kleinbuchstaben handelt es sich um signifikante Unterschiede auf einem Niveau von  $p < 0,05$ ; Großbuchstaben zeigen einen hochsignifikanten Unterschied ( $p < 0,01$ ) an. Steht hinter einem Wert beispielsweise ein „a“, so besagt das, dass der betreffende Wert sich von dem Wert in Spalte (A) signifikant unterscheidet, und zwar mit einer Irrtumswahrscheinlichkeit von  $< 5\%$ .

\* Die Suchmaschine bietet die entsprechende Funktion nicht an.