

Einflüsse auf die Ergebnisse kommerzieller Suchmaschinen

Modellbildung und empirische Ergebnisse

Dirk Lewandowski, Sebastian Schultheiß, Sebastian Sünkler

Hochschule für Angewandte Wissenschaften Hamburg, Deutschland

{[dirk.lewandowski](mailto:dirk.lewandowski@haw-hamburg.de), [sebastian.schultheiss](mailto:sebastian.schultheiss@haw-hamburg.de), [sebastian.suenkler](mailto:sebastian.suenkler@haw-hamburg.de)}

@haw-hamburg.de

Abstract

In diesem Beitrag stellen wir anhand eines Modells die von uns in den letzten Jahren durchgeführten empirischen Studien zu den Einflüssen auf die Suchergebnisse kommerzieller Suchmaschinen vor. Wir behandeln die Einflüsse durch die externen Akteursverbände Suchmaschinenoptimierung, Suchmaschinenwerbung und Nutzende und zeigen die Wechselwirkungen zwischen diesen auf. Die Zusammenfassung der berichteten Studien zeigt die bisher gewonnenen Erkenntnisse und stellt diese in den Kontext. Weiterhin werden wichtige weitere Forschungsfragen identifiziert. Wir verdeutlichen, an welchen Stellen besonderes Potenzial für die informationswissenschaftliche Forschung besteht. Schließlich geben wir eine Übersicht der von uns veröffentlichten Datensätze und Software, welche für alle interessierten Forschenden frei zur Verfügung stehen.

Schlagworte: Suchmaschinen, Ergebnispräsentation, Suchmaschinenwerbung, Suchmaschinenoptimierung, Einflussnahme auf Suchergebnisse

1 Einleitung

Kommerzielle Suchmaschinen wie Google sind aus dem Alltag der Informationssuchenden schon lange nicht mehr wegzudenken. Ihre Bedeutung für den Wissenserwerb in der Gesellschaft ist kaum zu überschätzen. Es ist daher verwunderlich, dass sich die Forschung mit der Einflussnahme auf die Er-

gebnisse kommerzieller Suchmaschinen und die daraus resultierenden Folgen für die Nutzenden vergleichsweise wenig auseinandergesetzt hat. Vergleicht man beispielsweise die Menge der veröffentlichten Arbeiten zu Suchmaschinen mit denen zu Social Media, zeigt sich ein großer Unterschied. Für das hohe Interesse an Social Media mag es unterschiedliche Gründe geben; einer liegt sicher darin, dass zumindest von Twitter leicht Daten zu bekommen sind, auf denen Analysen aufgebaut werden können. In der Konsequenz zeigt sich auch, dass ein großer Anteil der empirischen Studien zu Social Media Daten von Twitter verwendet. Einen Zugang zu Daten von kommerziellen Suchmaschinen gibt es in vergleichbarer Weise nicht; auch dies mag die relative Zurückhaltung der Forschenden in diesem Bereich erklären. In diesem Beitrag werden wir anhand der in den letzten Jahren in unserer Forschungsgruppe durchgeführten Arbeiten beschreiben, was wir für einen gangbaren Weg der Suchmaschinenforschung halten und wie wir Daten auf unterschiedliche Weise erheben. Dabei wollen wir aber nicht nur unsere Ansätze und Ergebnisse vorstellen, sondern interessierten Forschenden auch einen Weg aufzeigen, wie sie sich selbst mit kommerziellen Suchmaschinen beschäftigen können. Wir werden uns hier sowohl auf die Methoden als auch auf offene Forschungsfragen beziehen.

Natürlich soll nicht der Eindruck erweckt werden, dass es keine Forschung zu kommerziellen Suchmaschinen gebe. Im Gegenteil hat sich in den letzten ca. 20 Jahren ein robuster Korpus an Arbeiten zu diesen Suchsystemen entwickelt. In diesen Arbeiten werden allerdings häufig zentrale Eigenschaften der kommerziellen Suchmaschinen nicht – oder nicht ausreichend – berücksichtigt. Dies sind vor allem die Eigeninteressen der Suchmaschinenbetreiber, die aus Sicht der Nutzenden unklare Trennung von Suchergebnissen und Werbung sowie die externe Einflussnahme auf die Suchergebnisse durch Suchmaschinenoptimierung (SEO). Gerade für die verstärkt auftauchenden Studien zu gesellschaftlichen Einflüssen von Suchmaschinen(ergebnissen) beispielsweise auf Nutzende, die sich zu politischen Wahlen informieren möchten (u.a. Robertson et al., 2018; Urman et al., 2021), sind diese Punkte jedoch von entscheidender Bedeutung. Forschungen, die stärker auf die Weiterentwicklung von Information-Retrieval-Systemen ausgerichtet sind, verwenden in Evaluationsstudien oft kommerzielle Suchmaschinen beispielhaft, um daraus Erkenntnisse für die Weiterentwicklung anderer Systeme abzuleiten (z.B. Bais & Choi, 2022; Dragusin et al., 2013). Während dieses Vorgehen zu validen Ergebnissen führen kann, fällt auf, dass Suchergebnisseiten oft auf die zentrale Liste der organischen Ergebnisse reduziert wird, was der Kom-

plexität der Ergebnisdarstellungen in den Systemen, die u. a. Ergebnisse aus vertikalen Suchmaschinen, Werbetreffer und Faktenergebnisse beinhalten, nicht gerecht wird.

Wir haben uns in den letzten Jahren intensiv mit Einflüssen auf die Suchergebnisseiten von kommerziellen Suchmaschinen beschäftigt. Dies geschieht vor allem in zwei von der Deutschen Forschungsgemeinschaft geförderten Projekten, SEO-Effekt und RAT (*Result Assessment Tool*). Aufgrund der überragenden Marktstellung von Google haben wir uns vor allem auf diese Suchmaschine konzentriert, wo sinnvoll und möglich aber auch andere Suchmaschinen berücksichtigt. In diesem Beitrag wollen wir unsere in den letzten Jahren publizierten Ergebnisse zusammenfassen und in einem Modell verdeutlichen. Das Modell liefert einen Beitrag für das Verständnis der Zusammenhänge und kann weitere Forschung in dem beschriebenen Themenfeld anleiten. Ein besonderes Anliegen ist uns dabei, auch die genuin informationswissenschaftliche Herangehensweise an das Thema zu verdeutlichen: Den Beitrag der Informationswissenschaft sehen wir in der Fokussierung auf die informationssuchenden Personen und einer ganzheitlichen Betrachtung von Informationsprozessen (bzw. der ganzen „information chain“; Bawden & Robinson, 2022, S. 3). Methodisch liegt unser Ansatz in einer Verbindung von Methoden der Informatik und der Sozialwissenschaften.

2 Theoretischer Hintergrund und Methodik

Die Zusammenstellung von Suchergebnisseiten in kommerziellen Suchmaschinen lässt sich als eine Verhandlung zwischen Akteursverbänden ansehen (Röhle, 2010, S. 80ff.). Die Suchmaschinenbetreiber sind nicht allein für die Zusammenstellung der Ergebnisse verantwortlich, sondern weitere Akteursverbände nehmen Einfluss auf das, was den Nutzenden auf den Suchergebnisseiten gezeigt wird. Angelehnt an Röhle (2010, S. 80) handelt es sich um folgende Akteursverbände:

1. Suchmaschinenbetreiber (bei Röhle mit „Google“ benannt)
2. Inhaltenanbieter
3. Suchmaschinenoptimierung
4. Suchmaschinenwerbung (bei Röhle nicht explizit genannt)
5. Nutzende.

Im Vergleich zu Röhle wurde die Liste um den Akteursverbund Suchmaschinenwerbung erweitert und der Akteursverbund „Google“ im Sinne eines erweiterten Erkenntnisinteresses in „Suchmaschinenbetreiber“ umbenannt. Da das Augenmerk unserer aktuellen Arbeiten auf den externen Einflüssen auf die Suchergebnisse liegt, wird der Akteursverbund der Suchmaschinenbetreiber im Folgenden weitgehend ausgeklammert. Es soll jedoch zumindest erwähnt werden, dass Suchmaschinenbetreiber über die bevorzugte Anzeige von Ergebnissen aus eigenen Angeboten (bspw. Google Shopping oder YouTube in Google-Suchergebnissen) Nutzende innerhalb ihrer Angebote halten und die Suchen damit besser monetarisieren können. Das bekannteste Beispiel sind hier sicher Googles Shopping-Ergebnisse, für deren unfair bevorzugte Anzeige Google von der Europäischen Kommission zu einer Kartellstrafe von 2,42 Milliarden Euro verurteilt wurde (European Commission, 2017).

Aus dem theoretischen Ansatz ergeben sich Konsequenzen für das methodische Vorgehen. Um ein umfassendes Bild von den Einflüssen auf die Ergebnisse kommerzieller Suchmaschinen zu erlangen, müssen zum einen die Nutzenden in ihren Kenntnissen, ihren Erwartungen und ihrem Verhalten untersucht werden, zum anderen die von den Suchmaschinen ausgegebenen Ergebnisse selbst. Betrachtungen, die sich nur entweder auf den einen oder den anderen Bereich beziehen, reichen nicht aus, um ein umfassendes Verständnis der Einflüsse zu erlangen: Entweder fehlt die Quantifizierung der Inhalte bzw. Quellen der Suchergebnisse oder aber ein Verständnis dessen, wie Nutzende die präsentierten Ergebnisse wahrnehmen und mit ihnen umgehen. Dies bedeutet nicht, dass jede Studie immer beide Bereiche abdecken muss. Das Argument hier ist vielmehr, dass Forschungsprogramme, die sich mit dem Einfluss einer oder mehrerer Akteursverbände beschäftigen, unvollständig bleiben müssen, sofern sie nicht beide Bereiche berücksichtigen.

Um die Akteursverbände angemessen zu berücksichtigen, haben wir uns in unserem Forschungsprogramm für eine Kombination von informatischen und sozialwissenschaftlichen Methoden entschieden. Anknüpfend an die Diskussion der Frage, ob die Informationswissenschaft eigene Methoden verwendet oder „nur“ Methoden aus anderen Disziplinen importiert, sehen wir gerade die Methodenkombination in der Zusammenführung mit der ganzheitlichen Betrachtung der Informationsprozesse (Bawden & Robinson, 2022, S. 3) und der Fokussierung auf die informationssuchende Person als einen genuin informationswissenschaftlichen Ansatz.

Methoden der Informatik verwenden wir bei der automatischen Erfassung von Suchergebnissen mittels Scraping sowie der automatischen Analyse von Suchergebnissen. Hier wird schon deutlich, dass ein Teil unserer Forschungsarbeit aus der Entwicklung von spezialisierter Forschungssoftware besteht, die diese Zwecke erfüllt. Unser Ansatz hierbei ist, Software zu schaffen, die sich durch ihren modularen Aufbau für eine Vielzahl möglicher Studien, die sich mit Suchergebnissen aus unterschiedlichen Suchsystemen beschäftigen, eignet. Mit dem *Result Assessment Tool (RAT)*¹ entwickeln wir ein System, das diese Anforderungen erfüllt. Der Software-Code liegt ‚open source‘ vor, weiterhin unterstützen wir interessierte Forschende dabei, mithilfe der Software ihre eigenen Studien durchzuführen.

Aus dem Bereich der Sozialwissenschaften verwenden wir unterschiedliche Methoden aus dem Spektrum der empirischen Sozialforschung. Dies ist für die Informationswissenschaft erst einmal keine Besonderheit, allerdings haben wir uns bei verschiedenen Fragestellungen dafür entschieden, groß angelegte Befragungen, die durch die Bearbeitung von Aufgaben ergänzt werden, durchzuführen. Dabei ist natürlich nicht die Größe des Samples entscheidend, sondern seine Repräsentativität (wir halten uns hier an die Vorgaben der Arbeitsgemeinschaft Onlineforschung, AGOF²). Solche Studien erlauben valide Aussagen über die Haltung zu und die Nutzung von Suchsystemen. Es liegt auf der Hand, dass die Studien kaum ohne externe Unterstützung durch ein Marktforschungsinstitut durchgeführt werden können, da insbesondere die Gewährleistung der Repräsentativität der Stichprobe mit erheblichem Aufwand verbunden ist. Da dieser Aufwand neben personellen auch finanzielle Ressourcen erfordert, sind repräsentative Studien in der Informationswissenschaft selten. Wir haben in dieser Hinsicht sehr gute Erfahrungen gemacht und können interessierten Forschenden nur empfehlen, eine externe Durchführung von Befragungsstudien in Auftrag zu geben – wo dies finanziell möglich ist –, um den Evidenzwert ihrer Ergebnisse zu verbessern. Dies sorgt nicht nur fachintern für eine Verbesserung der Ergebnisqualität, sondern macht informationswissenschaftliche Forschung auch für andere Disziplinen attraktiver. Neben Befragungen führen wir auch Laborstudien durch, die, wie in der Informationswissenschaft üblich, durch technische Erfassungsmethoden wie Eyetracking ergänzt werden.

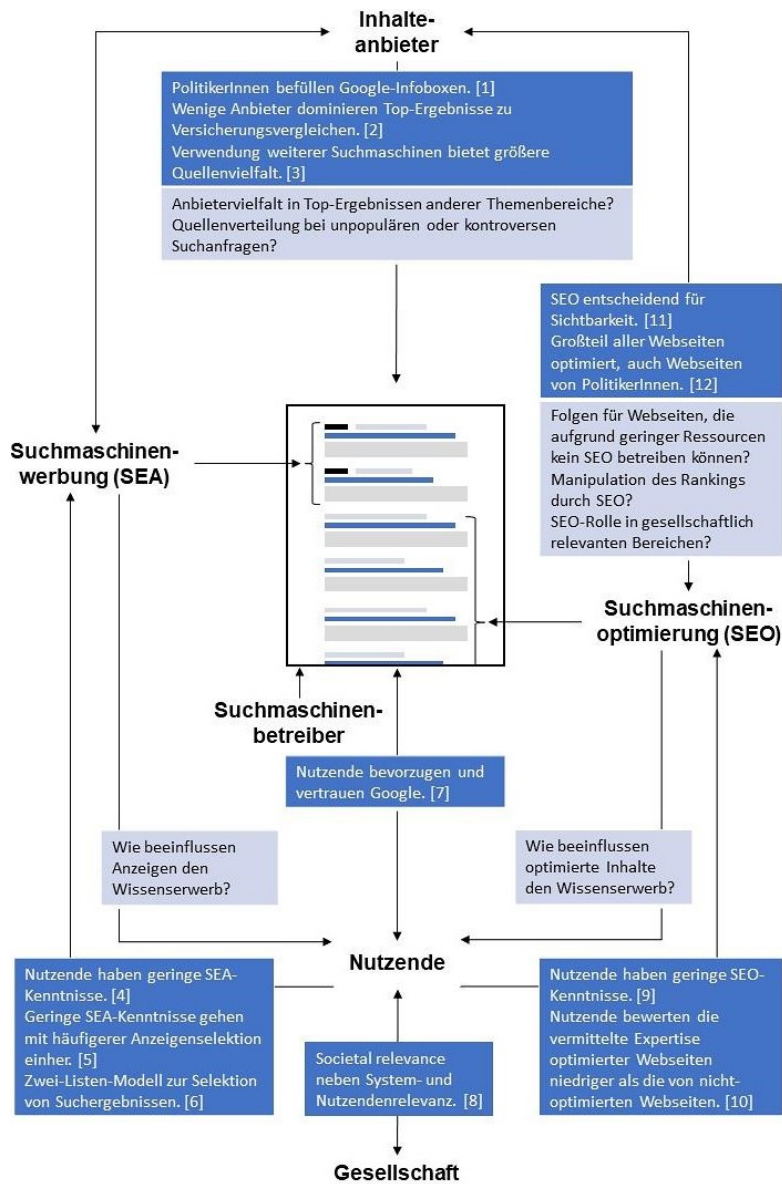
1 <https://searchstudies.org/de/research/rat/>

2 <https://www.agof.de/>

3 Beiträge zum Einfluss der Akteursverbände auf die Suchergebnisse

In diesem Kapitel berichten wir die zentralen Ergebnisse aus den in den letzten Jahren von uns durchgeführten Studien und ordnen diese in den Kontext, also Studien zur Einflussnahme auf Suchergebnisse und daraus resultierenden Folgen, ein. Um aussagekräftige Ergebnisse zu erzielen, verfügen unsere Studien einerseits über sinnvolle und wo möglich repräsentative Stichproben. Andererseits sind die Studien im Hinblick auf die nur bei den Suchmaschinenbetreibern vorliegenden Daten in Verbindung mit deren Forschungsinteressen ausgerichtet (Lewandowski et al., 2020). In diesem Sinne für die informationswissenschaftliche Forschung relevante Fragestellungen liegen außerhalb des Bereichs, der von den Suchmaschinenbetreibern anhand der ihnen vorliegenden Daten auf einer valideren empirischen Basis bearbeitet werden kann. Weiterhin relevant sind Studien, an denen die Suchmaschinenbetreiber selbst kein Interesse haben, weil die erwarteten Ergebnisse mit den Unternehmenszielen konfliktieren.

Berücksichtigt man die im Kapitel 2 genannten Akteursverbände und ihre Einflüsse auf die Suchergebnisse, ergeben sich Beziehungen, die in Abbildung 1 dargestellt sind. Dabei entstehen nicht nur gerichtete Einflüsse, sondern auch Wechselwirkungen, wenn sich zwei Akteursverbände gegenseitig beeinflussen. Neben den Beziehungen zwischen den Akteursverbänden werden im Schaubild auch die zentralen Ergebnisse aus unseren eigenen Arbeiten sowie zentrale Fragen für künftige Forschung benannt. Der Akteursverband der Suchmaschinenbetreiber bleibt in der weiteren Betrachtung weitgehend ausgeklammert, da unser Interesse hier auf der externen Beeinflussung der Suchergebnisse liegt.



bisherige Ergebnisse
offene Fragen

[1]: Hinz et al. (2020) [5]: Schultheiß & Lewandowski (2021b) [9]: Lewandowski & Schultheiß (2022)
 [2]: Lewandowski & Sünkler (2019) [6]: Lewandowski (2022) [10]: Schultheiß et al. (2022)
 [3]: Yagci et al. (2022) [7]: Schultheiß & Lewandowski (2021a) [11]: Schultheiß & Lewandowski (2020)
 [4]: Lewandowski et al. (2017, 2018); Lewandowski & Schultheiß (2022) [8]: Haider & Sundin (2019); Sundin et al. (2022) [12]: Hinz, Sünkler, Lewandowski (2023)

Abb. 1 Akteursverbände und ihre Einflüsse auf die Suchergebnisse

3.1 Inhalteanbieter

Zunächst einmal können die Inhalteanbieter selbst die Suchergebnisse beeinflussen: Mit jedem Dokument, das erstellt wird und potenziell in die Suchergebnisse eingehen kann, wird ein Einfluss auf potenzielle Ergebnislisten genommen. Dabei sind die Grenzen zwischen einfacher Texterstellung und Maßnahmen der Suchmaschinenoptimierung fließend: Schon allein der Versuch, gut verständliche Texte zu erstellen, kann implizit Maßnahmen der Suchmaschinenoptimierung verwenden. Auf der Seite der Analyse der Suchergebnisse ergibt sich damit das Problem, dass SEO-Maßnahmen, sofern sie sich direkt auf die erstellten Texte beziehen, nicht eindeutig identifizierbar sind. Insofern fällt eine eindeutige Einordnung empirischer Arbeiten entweder in den Bereich Inhalteanbieter oder Suchmaschinenoptimierung schwer.

Eindeutig dem Bereich der Inhalteanbieter zuzuordnen sind Maßnahmen, die in einem beschränkten Bereich unmittelbar auf die Inhalte der Suchergebnisseite einwirken. Im Rahmen der Bundestagswahl 2017 gab Google allen Kandidierenden die Möglichkeit, jeweils bis zu drei Statements zu ihren Grundsätzen und Prioritäten zu formulieren und diese direkt in Infoboxen auf den Google-Suchergebnisseiten einzubinden. Eine Analyse der Suchergebnisseiten von allen Kandidierenden (Hinz et al., 2020) zeigte, dass nur 478 der 4.828 Kandidierenden die Info-Boxen für ihre Selbstdarstellung verwendeten. Der Anteil der Nutzung der Info-Boxen war unter den Kandidierenden der etablierten Parteien am höchsten. Wer stärker etabliert und professionalisiert war, bemühte sich stärker um Sichtbarkeit im Netz.

In qualitativen Interviews mit Inhalteanbietern, Suchmaschinenoptimierern und Online-JournalistInnen (Schultheiß & Lewandowski, 2020) wurde übereinstimmend geäußert, dass Inhalteanbieter ohne die Hilfe von Suchmaschinenoptimierung nicht die gewünschte Sichtbarkeit in Suchmaschinen erreichen können. SEO wird als Standardmethode für die Generierung von Sichtbarkeit betrachtet; die Maßnahmen beginnen bereits vor der Erstellung der Inhalte. Dies wird durch Interviews mit OnlinejournalistInnen in anderen Ländern bestätigt (Giomelakis et al., 2019). Zusammengefasst verdeutlicht dies die Verschränkung dieser Akteursverbände in ihrem Interesse, Sichtbarkeit für bestimmte Inhalte zu schaffen.

Forschungslücken in Bezug auf die Inhalteanbieter sehen wir vor allem im Hinblick auf die Anbieterverteilung innerhalb der (Top-)Suchergebnisse (s. Abschn. 3.6). Beispielhafte Fragen sind hier, welche Anbietervielfalt sich in den Ergebnissen für unterschiedliche Themen finden lässt und wie sich

diese auf die Nutzenden auswirkt, sowie die Auswertung von Quellenverteilungen bei kontroversen Themen. Hierzu gibt es schon einige Veröffentlichungen (z.B. Ballatore, 2015; Gezici, 2021), allerdings sehen wir angesichts der Bedeutung des Themas hier noch Bedarf für weitere Arbeiten.

3.2 Suchmaschinenoptimierung

Suchmaschinenoptimierung ist für die Sichtbarmachung von Inhalten in Suchmaschinen von ausschlaggebender Bedeutung. Dies wird auch von den Suchmaschinenoptimierern selbst so gesehen, wie eine von uns durchgeführte Expertenbefragung zeigt (Schultheiß & Lewandowski, 2020). Interessant dabei ist, dass die Suchmaschinenoptimierer ihre Arbeit durchweg als positiv bewerten: Sie äußerten übereinstimmend, dass die Tätigkeit der Suchmaschinenoptimierung einen positiven Beitrag für die Ergebnisqualität der Suchmaschine und zum Internet generell leiste. Trotz des hohen Stellenwerts, dem SEO beigemessen wurde, nahmen die ExpertInnen an, dass SEO den meisten SuchmaschinennutzerInnen nicht bekannt ist (zur Perspektive der Nutzenden s.u.).

Wir haben den Einsatz von Techniken der Suchmaschinenoptimierung in unterschiedlichen Bereichen untersucht. Dazu haben wir Software entwickelt, die anhand eines umfangreichen Indikatorensets die Wahrscheinlichkeit, dass ein bestimmtes Dokument optimiert wurde, ermittelt (Lewandowski et al., 2021). Wir haben mit dem System unterschiedliche Themenbereiche (besonders populäre Suchanfragen, Corona, rechtsradikale Suchanfragen) mit insgesamt $n = 256.853$ Ergebnissen getestet. Themenübergreifend hat sich gezeigt, dass die von Google angezeigten Suchergebnisse zu einem großen Teil optimiert sind. Mit sinkender Trefferposition sinkt der Anteil der optimierten Ergebnisse zwar, allerdings sind auch bei Treffern nach Position 100 noch mehr als 80% optimiert. Dies unterstreicht den Einfluss, den optimierte Ergebnisse auf das haben, was Nutzende in Suchmaschinen zu sehen bekommen.

Im Kontext der Studien zum Thema Suchmaschinenoptimierung fanden wir Hinweise auf einen Zusammenhang zwischen SEO-Maßnahmen und kommerziellen Motivationen der Inhalteanbieter (Schultheiß et al., 2022). Für gesundheitsbezogene Suchergebnisse konnten wir zeigen, dass SEO-Maßnahmen vermehrt von kommerziell motivierten Anbietern wie Pharmaunternehmen und weniger häufig von nicht-kommerziellen Anbietern, beispielsweise Behörden, durchgeführt werden. Gleichzeitig verfügen optimierte

Webseiten über einen durchschnittlich höheren Rang innerhalb der Suchergebnisse. Vor diesem Hintergrund stellten wir die Frage, ob Nutzende zwischen optimierten und nicht-optimierten Webseiten Qualitätsunterschiede wahrnehmen. In einer Nutzerstudie ließen wir $n = 61$ Teilnehmende die Qualität optimierter und nicht-optimierter gesundheitsbezogener Webseiten anhand diverser Kriterien wie Vertrauenswürdigkeit, Objektivität und Expertise bewerten. Diese Kriterien gehören zu jenen, die Nutzende bei der Bewertung der Qualität von online verfügbaren Gesundheitsinformationen am häufigsten verwenden. Die Teilnehmenden bewerteten die vermittelte Expertise der nicht-optimierten Webseiten als signifikant höher als die der optimierten Seiten. Begründet wurde dies durch ein kompetenteres bzw. seriöseres Erscheinungsbild der BetreiberInnen und AutorInnen nicht-optimierter Webseiten.

Im Bereich der politischen Kommunikation haben wir eine Studie durchgeführt, die im Rahmen der Bundestagswahl 2021 untersuchte, ob und in welchem Maße SEO von Kandidierenden für den Bundestag eingesetzt wurde (Hinz et al., 2023). Insgesamt hatten $n = 1.372$ der $n = 6.211$ Kandidierenden eine eigene persönliche Website als Kommunikationsmittel für den Wahlkampf. Auf 93% der persönlichen Seiten konnte der Einsatz von SEO ermittelt werden. SEO ist also auch in der politischen Kommunikation im Web von großer Bedeutung. Es zeigte sich, dass, je professionalisierter ein Kandidat ist, desto höher die Wahrscheinlichkeit der SEO-Nutzung ist. Ebenso zeigten sich Zusammenhänge zwischen der Zugehörigkeit der Partei zum Parlament und der Nutzung von SEO.

Die Ergebnisse zum Einfluss der Suchmaschinenoptimierung zeigen, dass Methoden der Suchmaschinenoptimierung themenübergreifend eingesetzt werden und sich auch auf gesellschaftlich sowie persönlich relevante Inhalte beziehen. Die Forschung zum Einfluss von SEO steht allerdings noch am Anfang und wir sehen hier viele offene Forschungsfragen. Eine grundlegende, für die Informationswissenschaft besonders relevante Frage ist die nach der Manipulation oder zumindest Veränderung der Rankings durch SEO, da dadurch direkt die Relevanz der Suchergebnisse beeinflusst wird. Durch die Einflussnahme verschiebt sich die Frage nach der Relevanz von einer allein auf den Anbieter von Suchsystemen gerichteten hin zu einer Aushandlung der Relevanz von Suchergebnissen. Auch die Diskussion um nicht-kommerzielle bzw. alternative Suchmaschinen würde erheblich profitieren, wenn sie die Frage der Relevanz unter den gegebenen Bedingungen stärker berücksichtigen würde. Eine weitere interessante Frage ist, wie Anbieter mit fehlenden oder nur geringen Ressourcen für SEO in Suchmaschinen repräsentiert

sind. Auf der einen Seite kann hier SEO einfach als ein Wettbewerb gesehen werden, bei dem die stärkere Partei gewinnt; auf der anderen Seite ist es gerade bei den gesellschaftlich sowie persönlich relevanten Inhalten wichtig, dass Suchmaschinen dazu beitragen, mit den Suchergebnissen den gesellschaftlichen Diskurs zu befördern, anstatt ihn anhand der stärksten Anbieter auszurichten.

3.3 Suchmaschinenwerbung

Der Ausgangspunkt unserer Betrachtungen der Suchmaschinenwerbung ist, dass es sich bei der Werbung um einen Typ von Suchergebnissen handelt. Die Werbung wird ebenso wie die organischen Ergebnisse als Antwort auf eine Suchanfrage ausgegeben und bezieht sich damit direkt auf die Informationsbedürfnisse der Nutzenden. Da die Werbetreibenden in dieser Werbeform per Klick bezahlen (im Gegensatz zu anderen Werbeformen, bei denen nach der Zahl der Einblendungen abgerechnet wird) und Suchmaschinen sich weit überwiegend aus dieser Werbung finanzieren, besteht für Suchmaschinenbetreiber der Anreiz, die Grenzen zwischen organischen Suchergebnissen und Werbung zu vermischen. Dieses Problem begleitet die Geschichte der Suchmaschinen seit den Anfängen der Suchwortvermarktung; eine für die Nutzenden eindeutige und verständliche Trennung wird immer wieder angemahnt (Marvin, 2019; Sullivan, 2013).

Aus der Berücksichtigung der Suchmaschinenwerbung als Ergebnistyp haben wir ein Modell entwickelt, das je nach Kenntnis und Bewertung der Suchmaschinenwerbung unterschiedliche Verhalten berücksichtigt (Lewandowski, 2022). Das Modell erlaubt eine Prognose von Klicks und kann als Basis für Nutzerstudien zum Einfluss der Anzeigen dienen.

Für die künftige Forschung im Bereich Suchmaschinenwerbung bieten sich zum einen Studien, die sich mit der Gestaltung und dem Effekt der Werbung auf Klick-Entscheidungen beschäftigen, an. Am bedeutendsten dürfte zum anderen aber die Frage sein, wie der Wissenserwerb von informationssuchenden Personen in Abhängigkeit ihrer Kenntnis bzw. Berücksichtigung der Werbung beeinflusst wird.

3.4 Nutzende

In einer Multimethodenstudie (Befragung, taskbasierte Onlinestudie, Online-Experiment) mit einer repräsentativen Stichprobe von $n = 1.000$ Teilnehmenden

den konnte ein umfassendes Bild der (Un-)Kenntnis der SuchmaschinennutzerInnen hinsichtlich der Mischung von Werbung und „neutralen“ Ergebnissen auf den Suchergebnisseiten gezeichnet werden (Lewandowski et al., 2017, 2018). Im Befragungsteil der Studie wurde festgestellt, dass 58% der deutschen InternetnutzerInnen entweder die Möglichkeit der prominenteren Platzierung (mittels Anzeige) nicht kennen oder nicht wissen, dass Anzeigen von organischen Ergebnissen unterscheidbar sind. Im zweiten Teil hatten die Befragten in vier Aufgaben je nach Aufgabenstellung organische oder bezahlte Ergebnisse auf fünf Screenshots von *search engine result pages* (SERPs) zu markieren. Lediglich 1% der Teilnehmenden gelang es, alle Bereiche vollständig und ohne Fehler (z.B. Anzeige fälschlicherweise als organisches Ergebnis bewertet) zu kennzeichnen. Weitere 10% machten keine fehlerhaften Markierungen, wenngleich sie dabei nicht sämtliche Ergebnisse identifiziert haben. In einer weiteren repräsentativen Befragung (Lewandowski & Schultheiß, 2022) wurden die Wissensfragen zur Suchmaschinenwerbung wie auch die Markieraufgaben erneut integriert. Die Wissensfragen kamen zu denselben Ergebnissen: Wieder lag der Anteil der Nutzenden, die die Platzierungsmöglichkeit mittels Anzeige nicht kennen oder keine Unterscheidbarkeit zwischen Anzeigen und organischen Ergebnissen annehmen, bei 58%. Die Markieraufgaben ergaben, dass etwa jede zweite Anzeige korrekt als solche identifiziert wurde, wobei eine andere Methodik sowie andere SERP-Screenshots als Stimuli zum Einsatz kamen als in der vorherigen Studie (Lewandowski et al., 2017, 2018), weshalb die Ergebnisse der Markieraufgaben beider Studien nur bedingt vergleichbar sind. Deutsche InternetnutzerInnen sind somit zum Großteil nicht in der Lage, verlässlich zwischen Anzeigen und organischen Ergebnissen zu unterscheiden. Der Kenntnisstand zu den Anzeigen beeinflusst auch das Auswahlverhalten auf den Suchergebnisseiten: Versuchspersonen, die die Anzeigen nicht von den organischen Ergebnissen unterscheiden können, wählten etwa doppelt so häufig ein als Anzeige gekennzeichnete Ergebnis aus als Teilnehmende, die Anzeigen von organischen Ergebnissen unterscheiden konnten. Dies bestätigt die Annahme, dass Nutzende, die über eine geringe Anzeigenkenntnis verfügen, Anzeigen und Suchergebnisse als eine einzige Liste betrachten („Zwei-Listen-Modell“, s. Lewandowski, 2022). Anschließend an diese Studie haben wir in einem Eye-Tracking-Experiment inklusive Fragebogenelementen und $n = 100$ Teilnehmenden untersucht, welchen Einfluss der Kenntnisstand über Suchmaschinenwerbung auf das Rechercheverhalten ausübt (Schultheiß & Lewandowski, 2021b). Da Anzeigen insbesondere am Smartphone große Teile des

Bildschirms einnehmen, fand die Studie neben dem Desktop-PC auch am Smartphone statt. Die Ergebnisse zeigen einen Einfluss des Endgeräts auf das Blickverhalten. So schenkten die Versuchspersonen den Anzeigen insbesondere am Smartphone eine hohe visuelle Aufmerksamkeit. Darüber hinaus wurde ein Zusammenhang zwischen Anzeigenkenntnis und Auswahlverhalten festgestellt. Versuchspersonen, die über geringe Anzeigenkenntnisse verfügten, wählten signifikant häufiger eine Anzeige aus als Versuchspersonen mit hohen Anzeigenkenntnissen.

In einer für die deutsche Internet-Bevölkerung repräsentativen Online-Umfrage mit $n = 2.012$ befragten Personen (Lewandowski & Schultheiß, 2022) zeigte sich, dass nur eine Minderheit der Nutzenden die Einflussmöglichkeit durch Suchmaschinenoptimierung (SEO) kennt. Ein noch geringerer Anteil kennt „SEO“ als Begriff, kann SEO-Maßnahmen benennen oder auf Screenshots von Suchergebnisseiten (SERPs) SEO mit organischen Ergebnissen assoziieren. Die befragten InternetnutzerInnen nahmen sowohl positive als auch negative Folgen an, die aus SEO resultieren könnten. Unter anderem wurde SEO einerseits zugesprochen, die Ergebnisqualität zu erhöhen, andererseits jedoch befürchtet, dass kleinere Webseitenbetreiber aus den Ergebnissen verdrängt werden könnten, insofern diese über keine ausreichenden Ressourcen für SEO-Maßnahmen verfügen.

In einer auf die Frage nach dem Vertrauen in Suchmaschinen ausgerichteten Auswertung des gleichen Datensatzes (Schultheiß & Lewandowski, 2021a) zeigte sich, dass Nutzende, die bevorzugt Google nutzen, einerseits über geringere SEO-Kenntnisse verfügten und andererseits stärker darauf vertrauten, dass Google eine faire und unvoreingenommene Informationsquelle ist und ihnen korrekte und vertrauenswürdige Inhalte liefert. Dahingegen konnte bei Personen, die bevorzugt andere Suchmaschinen als Google nutzen, Gegenteiliges beobachtet werden – nämlich eine im Schnitt höhere SEO-Kenntnis bei niedrigerem Vertrauen in Google.

Im Bereich der Nutzerstudien, die den Kontext der beeinflussten Suchergebnisse in kommerziellen Suchmaschinen berücksichtigen, ergeben sich zahllose Möglichkeiten. Wir sehen diesen Bereich als besonders fruchtbar für die informationswissenschaftliche Forschung, da sich hier die in der Informationswissenschaft vorhandenen Modelle und Methoden leicht übertragen lassen. Während es sich bei den von uns durchgeführten Studien um quantitative Analysen handelt, die sich auf große und repräsentative Samples beziehen, sehen wir auch erhebliches Potenzial für qualitative Studien.

3.5 Einfluss Nutzende/Gesellschaft

Die Frage der Relevanz der Suchergebnisse wird in der Informationswissenschaft traditionell als System- oder als Nutzerrelevanz betrachtet (Saracevic, 2016). Im Hinblick auf die durch die Interessen der unterschiedlichen Akteursverbände – die, wie wir gezeigt haben, die von den Nutzenden wahrgenommenen Suchergebnisse erheblich beeinflussen – stellt sich allerdings die Frage nach einer dritten Form von Relevanz, nämlich der gesellschaftlichen Relevanz (*societal relevance*; beispielsweise bei COVID-19-bezogenen Suchanfragen; Haider & Sundin, 2019; Sundin et al., 2022). Die Einflüsse, die Suchmaschinen vermittelt über die Nutzenden auf die Gesellschaft haben, lassen sich erst durch die Betrachtung dieser drei Bereiche der Relevanz angemessen erfassen.

3.6 Verteilung der Angebote innerhalb der (Top-)Suchergebnisse

Neben der Untersuchung des Einflusses der einzelnen Akteursverbände haben wir auch Studien durchgeführt, in denen die Verteilung der (Top-)Suchergebnisse nach Quellen untersucht wurde. Die ermittelten Verteilungen lassen sich also nicht auf den Einfluss bestimmter Akteursverbände zurückführen, zeigen jedoch ein Bild der Quellen, die Suchenden bevorzugt angezeigt werden, wenn sie Suchmaschinen benutzen.

In einer Studie wurden die Unterschiede zwischen den Ergebnissen von Google im Vergleich mit anderen Suchmaschinen untersucht (Yagci et al., 2022). Dabei wurden die Top-10-Ergebnisse von Google, Bing, DuckDuckGo und MetaGer anhand von $n = 3.537$ Suchanfragen aus Google Trends aus Deutschland und den USA verglichen. Dabei zeigte sich, dass Google mehr verschiedene Domains in den Top-Ergebnissen anzeigt als seine Konkurrenten. Wikipedia und Nachrichten-Websites sind insgesamt die am häufigsten innerhalb der Top-Ergebnisse vertretenen Quellen. Auch wenn einige Top-Quellen die Suchergebnisse dominieren, ist die Verteilung der Domains über alle Suchmaschinen hinweg recht ähnlich: Es dominieren vor allem Wikipedia und bekannte Nachrichtenquellen. Die Überschneidung zwischen Google und Bing liegt immer unter 32%, während MetaGer eine größere Überschneidung mit Bing aufweist als DuckDuckGo mit Bing, nämlich bis zu 78%. Diese Studie zeigt, dass die Verwendung einer weiteren Suchmaschine, insbesondere zusätzlich zu Google, eine größere Vielfalt an Quellen bietet und die Nutzenden dazu bringen kann, neue Perspektiven zu finden.

Während diese Studie durch die Verwendung eines Suchanfragensets auf der Basis von Google Trends Aussagen über die Quellenverteilung bei der Eingabe von besonders populären Suchanfragen treffen kann, haben wir uns in weiteren Untersuchungen mit spezifischen Themen beschäftigt. In einer Studie im Bereich Verbraucherschutz (Lewandowski & Sünkler, 2019) haben wir anhand von $n = 121$ verschiedenen Suchanfragen zum Thema Versicherungsvergleich ermittelt, dass nur wenige Anbieter die Top-Suchergebnisse mit ihren Angeboten dominieren. Sie erscheinen häufig auf der ersten Suchergebnisseite, auch unter der Verwendung verschiedener Domains. Es wird deutlich, dass kleinere Anbieter für Versicherungsvergleiche damit weniger Chancen haben, von Nutzenden wahrgenommen zu werden. Dies betrifft insbesondere auch Non-Profit-Angebote wie die Stiftung Warentest, die bei ihren Versicherungsvergleichen deutlich höhere Standards ansetzt als die dominierenden kommerziellen Anbieter (Stiftung Warentest, 2022).

Wenn Anfragen, die auf rechtsradikale Inhalte hindeuten, gestellt werden, dann zeigt Google vor allem Ergebnisse aus Mainstream-Quellen (bspw. Nachrichtenportale, Wikipedia). Fügt man jedoch Ortsnamen hinzu, um die Suchanfragen zu lokalisieren, ändert sich das Bild (Norocel & Lewandowski, 2023). Je nach Suchanfragen-Level (nach „Rechtsradikalität“) ändert sich das Quellenspektrum. Je ‚radikaler‘ die Suchanfrage, desto häufiger finden sich Ergebnisse von offen rechtradikalen Websites. Diese Studie weist darauf hin, dass bei Themen, bei denen die Top-Suchergebnisse vermeintlich ‚sauber‘ sind, sich das Bild radikal wandeln kann, wenn Nutzende spezifischer suchen. Insbesondere anknüpfend an die (konzeptionelle) Forschung zu Data Voids (Golebiewski & Boyd, 2018) ergeben sich hier Anknüpfungspunkte für weitere empirische Arbeiten.

Forschung zur Verteilung von Quellen innerhalb von (Top-)Ergebnissen hat bereits eine gewisse Tradition; es fällt allerdings auf, dass gerade im Zuge der Diskussion um den gesellschaftlichen Einfluss von Suchmaschinen bzw. Suchergebnissen das Interesse an solchen Forschungen neu erwacht ist und in den letzten Jahren einige Studien publiziert wurden (u. a. Agrawal, 2016). Hier sehen wir auch die Chance, das Thema *Search Engine Bias*, das in vielen Punkten unter einer schlechten empirischen Basis leidet, auf ein solides empirisches Fundament zu stellen.

4 Diskussion

In der Beschreibung der von uns durchgeführten Studien und dem Ausblick auf besonders interessante Forschungsfragen in den jeweiligen Bereichen wurde schon deutlich, dass in der Erforschung der Einflüsse auf die Ergebnisse kommerzieller Suchmaschinen ein erhebliches Potenzial für die informationswissenschaftliche Forschung liegt.

Für einen Kern des Fachs, nämlich die Fokussierung auf den Begriff der Relevanz, ändern sich die Voraussetzungen, wenn davon ausgegangen werden muss, dass Suchergebnisseiten sowohl durch direkt kommerzielle Ergebnisse (in Form von Werbetreffern) als auch durch hochgradig optimierte organische Ergebnisse sowie die Eigeninteressen der Suchmaschinenbetreiber beeinflusst sind. Wenn die Informationswissenschaft weiterhin Relevanz als ihren zentralen Begriff ansieht, dann muss dieser im Kontext von Anbietern von Informationssystemen, die einerseits Eigeninteressen verfolgen und deren Ergebnisse andererseits durch externe Akteure beeinflusst wird, neu betrachtet werden. Im Hinblick auf die aktuelle Herangehensweise in Informationswissenschaft und Information Retrieval kann man allerdings von einer naiven Betrachtung sprechen, die interessanterweise gerade eher der Betrachtungsweise der Nutzenden als der der hochprofessionalisierten Akteursverbände der Suchmaschinenbetreiber, der Suchmaschinenoptimierer und der Suchwortvermarkter entspricht.

Die Informationswissenschaft kann in dem skizzierten Forschungsbereich mit ihrer besonderen Expertise Vorreiter für andere Fächer sein, die sich ebenfalls mit diesem Themenfeld auseinandersetzen bzw. gerade auseinanderzusetzen beginnen. Wir halten es hier strategisch allerdings für sinnvoll, dass sich die in der Informationswissenschaft Forschenden auch außerhalb des Fachs engagieren und informationswissenschaftliche Forschung dort bekannt machen, wo an sie in anderen Fachcommunitys angeknüpft werden kann. Natürlich handelt es sich hier um einen Spagat, da sowohl die eigene Community als auch verwandte Communitys bedient werden müssen.

Eine Schwachstelle aller bisherigen Studien, die Einflüsse auf die Suchergebnisse auf der Basis einer Auswahl von Suchanfragen messen, ist die Auswahl der Suchanfragen. Die verwendeten Suchanfragensets werden in aller Regel intellektuell zusammengestellt, um ein vorab bestimmtes Thema abzudecken. Allerdings stellt sich die Frage, ob diese Suchanfragen das Thema tatsächlich vollständig abdecken. Weiterhin berücksichtigen bisherige

Studien nicht die tatsächlichen Häufigkeiten, mit denen Suchanfragen gestellt werden. Dadurch kann es zu Verzerrungen in den Ergebnissen kommen.

Die Frage nach dem Einfluss von SEO dürfte sich in Zukunft noch stärker stellen. Zum einen ist eine weitere Ausweitung von SEO auf informative Inhalte zu beobachten. Schon seit Längerem werden nicht nur Inhalte zu Produkten und Dienstleistungen optimiert, sondern bspw. auch journalistische Inhalte. Ein besonderes Augenmerk sollte hier allerdings auf die interessen-geleitete Kommunikation gelegt werden. Insbesondere im Zusammenhang mit dem hohen Vertrauen, das Nutzende in Suchmaschinen – insbesondere Google – und ihre Suchergebnisse setzen, besteht hier die Gefahr einer starken und unbemerkten Einflussnahme.

Weiterhin ist zu beobachten, dass sich SEO nicht nur auf die allgemeinen Suchmaschinen beschränkt, sondern auf alle Plattformen ausgeweitet wird. So werden schon seit Längerem die sog. „Academic SEO“ und „Amazon-SEO“ diskutiert. Eine Ausweitung auf weitere Bereiche ist als wahrscheinlich anzusehen; letztlich wird jede Plattform, von der sich Inhalteanbieter einen nennenswerten Zustrom von Nutzenden erwarten, von SEO betroffen sein. Interessant ist hier, wie sich die Plattformbetreiber zu solchen Maßnahmen positionieren werden und wie die Gesellschaft mit Plattformen umgehen wird, deren Ergebnisse nicht nur durch die Relevanzentscheidungen der Systemersteller, sondern auch extern beeinflusst werden.

5 Ressourcen

Mit diesem Beitrag möchten wir nicht nur die Methodik und Ergebnisse unserer Forschung beschreiben und einordnen, sondern auch weitere Forschung in diesem Themenfeld stimulieren. Dazu haben wir zum einen auf aus unserer Sicht interessante Fragen für die weitere Forschung hingewiesen. Zum anderen möchten wir auch auf die zahlreichen von uns bereitgestellten Ressourcen hinweisen, die interessierte Forschende nachnutzen können. Es handelt sich dabei um die Forschungsdaten unserer Studien, aber auch um den Software-Code sowie Software, die wir als Service anbieten. Tabelle 1 gibt einen Überblick zu allen Forschungsdaten, die im Rahmen unserer Studien entstanden sind. In der Tabelle werden die Inhalte in den Datensätzen sowie die Verfügbarkeit aufgezeigt.

Tab. 1: Forschungsdaten und Software

Datensatz / Software	Beschreibung	Verfügbarkeit
repräsentative Online-Befragung	Online-Umfrage aus dem Jahr 2020 mit einer repräsentativen Stichprobe der deutschen Online-Bevölkerung im Alter von 16 bis 69 Jahren durch ($n = 2.012$). Die Online-Befragung umfasste zwölf suchmaschinenbezogene Abschnitte. In der Befragung wurden weit mehr Items abgefragt als in unseren Aufsätzen berichtet wurden; außerdem ist der Datensatz sehr detailliert hinsichtlich der Eigenschaften der Teilnehmenden.	https://doi.org/10.17605/OSF.IO/PG82E Schultheiß & Lewandowski, 2022
SEO auf Suchergebnissen bei Google	Ermittlung der Wahrscheinlichkeit von SEO auf insgesamt $n = 256.853$ Ergebnissen mit Suchanfragen und SEO-Indikatoren	https://doi.org/10.17605/OSF.IO/RVX54 Lewandowski et al., 2021
Unterschiede in den Suchergebnissen bei verschiedenen Suchmaschinen	Top-10-Ergebnisse von Google, Bing, DuckDuckGo und MetaGer mit $n = 3.537$ Suchanfragen von Google Trends Deutschland und USA	https://doi.org/10.17605/OSF.IO/NT3WV Yagci et al., 2022
Googles Suchergebnisse für deutsche Suchanfragen zu Versicherungsvergleichsangeboten	$n = 121$ relevante Suchanfragen zum Thema Versicherungsvergleich mit den dazugehörigen Suchergebnissen sowie Merkmalen der Anbieter	https://zenodo.org/record/2572691#.YtAm-3ZByUk Lewandowski & Sünkler, 2019
SEO im Bundestagswahlkampf 2021	Rohdaten und statistische Auswertungen zur Ermittlung von SEO auf persönlichen Webseiten der Kandidierenden für die Bundestagswahl 2021	https://doi.org/10.17605/OSF.IO/2WS5C Hinz, Sünkler, & Lewandowski, 2023
SEO-Effekt Demo Tool	Demoanwendung, die die Wahrscheinlichkeit von SEO auf einer Webseite überprüft	https://searchstudies.org/research/seo-effekt/demo
Software Repository	öffentlich zugänglicher Software-Code von allen in der Forschung entstandenen Tools	https://github.com/searchstudies

Am bedeutendsten für Forschende, die sich für die Durchführung von Studien auf der Basis von Suchergebnissen interessieren, dürfte allerdings das von uns entwickelte *Result Assessment Tool* (RAT) sein. Mit dieser Software lassen sich sehr flexibel Studien erstellen und durchführen. Eine erste Version der Software liegt nicht nur ‚open source‘ vor, sondern wir bie-

ten die Software auch als Service an, der von interessierten Forschenden genutzt werden kann. Weiterhin können wir Hilfe bei der Durchführung eigener Studien anbieten, sodass auch technisch weniger versierte Personen von den Funktionen der Software profitieren können.

Danksagung

Die meisten der in diesem Text behandelten Studien wurden im Rahmen der von der Deutschen Forschungsgemeinschaft (DFG) geförderten Projekte „Relevance Assessment Tool (RAT): Entwicklung und Erweiterung einer Software zur Durchführung von Studien, die auf Ergebnissen aus Suchsystemen beruhen“ (Förderkennzeichen 460676551) und „Der Effekt der Suchmaschinenoptimierung auf die Suchergebnisse von Web-Suchmaschinen: Modellentwicklung, empirische Überprüfung und Triangulation mit NutzerInnen- und ExpertInneneinschätzungen (SEO-Effekt)“ (Förderkennzeichen 417552432) durchgeführt. Im Rahmen der Studien wurden große Mengen von Daten intellektuell klassifiziert; für ihre Mitarbeit möchten wir den in den Projekten beschäftigten Hilfskräften (in alphabetischer Reihenfolge) danken: Lena Haberzettl, Michael Haring, Marieke Hartmann, Nina Niesche und Ngan Pham. Die Befragung der deutschen Internet-Bevölkerung zu ihren Kenntnissen und Einstellungen zur Suchmaschinenoptimierung wurde durch die Fittkau & Maaß Consulting GmbH durchgeführt. Wir möchten uns für die kompetente und reibungslose Beratung und Durchführung bedanken.

Referenzen

- Agrawal, R. (2016). Overlap in the Web Search Results of Google and Bing. *Journal of Web Science*, 2(1), 17–30. <https://doi.org/10.1561/106.00000005>
- Bais, S., & Choi, B. (2022). Developing a Search Engine for Social Networks. *The 8th International Conference on Frontiers of Educational Technologies (ICFET)*, Yokohama (S. 176–186). ACM. <https://doi.org/10.1145/3545862.3545891>
- Ballatore, A. (2015) Google chemtrails: a methodology to analyze topic representation in search engine results. *First Monday* 20(7). <https://doi.org/10.5210/fm.v20i7.5597>
- Bawden, D., & Robinson, L. (2022). *Introduction to Information Science* (2nd edition). Facet Publishing.

- Dragusin, R., Petcu, P., Lioma, C., Larsen, B., Jørgensen, H. L., Cox, I. J., Hansen, L. K., Ingwersen, P., & Winther, O. (2013). FindZebra: A search engine for rare diseases. *International Journal of Medical Informatics*, 82(6), 528–538. <https://doi.org/10.1016/j.ijmedinf.2013.01.005>
- European Commission (2017). Antitrust: Commission fines Google €2.42 billion for abusing dominance as search engine by giving illegal advantage to own comparison shopping service – Factsheet. http://europa.eu/rapid/press-release_MEMO-17-1785_en.htm
- Gezici, G. (2021). Biased or Not? The Story of Two Search Engines. *arXiv*. <https://arxiv.org/abs/2112.12802>
- Giromelakis, D., Karypidou, C., & Veglis, A. (2019). SEO inside Newsrooms: Reports from the Field. *Future Internet*, 11(12), 261. <https://doi.org/10.3390/fi11120261>
- Golebiewski, M., & Boyd, D. (2018). Data Voids: Where Missing Data Can Easily Be Exploited. *Data & Society*, (May 2018). https://datasociety.net/wp-content/uploads/2018/05/Data_Society_Data_Voids_Final_3.pdf
- Haider, J., & Sundin, O. (2019). *Invisible Search and Online Search Engines*. Routledge.
- Hinz, K., Sünkler, S., & Lewandowski, D. (2020). Selbstdarstellung und Positionierung von Kandidatinnen und Kandidaten zur Bundestagswahl 2017 in Google-Infoboxen. *Medien & Kommunikationswissenschaft*, 68(1–2), 94–112. <https://doi.org/10.5771/1615-634X-2020-1-2-94>
- Hinz, K., Sünkler, S., & Lewandowski, D. (2023). SEO im Wahlkampf: Welche Kandidierende durch Suchmaschinenoptimierung ihre Sichtbarkeit zu erhöhen versuchen. In K.-R. Korte (Hrsg.), *Die Bundestagswahl 2021: Analysen aus Sicht der Wahl-, Parteien-, Regierungs- und Kommunikationsforschung*. VS Springer. https://doi.org/10.1007/978-3-658-35758-0_19-1
- Lewandowski, D. (2022). A two-list selection model of search engine results [Manuskript zur Veröffentlichung eingereicht].
- Lewandowski, D., Kerkmann, F., Rümmele, S., & Sünkler, S. (2018). An empirical investigation on search engine ad disclosure. *Journal of the Association for Information Science and Technology*, 69(3), 420–437. <https://doi.org/10.1002/asi.23963>
- Lewandowski, D., & Schultheiß, S. (2022). Public awareness and attitudes towards search engine optimization. *Behaviour & Information Technology* (20 S.). <https://doi.org/10.1080/0144929X.2022.2056507>
- Lewandowski, D., & Sünkler, S. (2019). What does Google recommend when you want to compare insurance offerings? *Aslib Journal of Information Management*, 71(3), 310–324. <https://doi.org/10.1108/AJIM-07-2018-0172>

- Lewandowski, D., Sünkler, S., & Kerkmann, F. (2017). Are Ads on Google Search Engine Results Pages Labeled Clearly Enough? The Influence of Knowledge on Search Ads on Users' Selection Behaviour. In M. Gäde, V. Trkulja, & V. Petras (Hrsg.), *Everything Changes, Everything Stays the Same? Understanding Information Spaces. Proceedings of the 15th International Symposium of Information Science (ISI 2017)* (S. 62–74). Hülsbusch.
- Lewandowski, D., Sünkler, S., & Schultheiß, S. (2020). Studies on Search: Designing Meaningful IIR Studies on Commercial Search Engines. *Datenbank-Spektrum*, 20(1), 5–15. <https://doi.org/10.1007/s13222-020-00331-1>
- Lewandowski, D., Sünkler, S., & Yagci, N. (2021). The influence of search engine optimization on Google's results. *13th ACM Web Science Conference 2021* (S. 12–20). <https://doi.org/10.1145/3447535.3462479>
- Marvin, G. (2019). *Updated: A visual history of Google ad labeling in search results*. Search Engine Land. <https://searchengineland.com/search-ad-labeling-history-google-bing-254332>
- Norocel, O. C., & Lewandowski, D. (2023). Google, data voids, and the dynamics of the politics of exclusion. *Big Data & Society*, 10(1), 205395172211490. <https://doi.org/10.1177/20539517221149099>
- Robertson, R. E., Jiang, S., Joseph, K., Friedland, L., Lazer, D., & Wilson, C. (2018). Auditing Partisan Audience Bias within Google Search. *Proceedings of the ACM on Human-Computer Interaction*, 2 (CSCW), Art. No. 148. <https://doi.org/10.1145/3274417>
- Röhle, T. (2010). *Der Google-Komplex: Über Macht im Zeitalter des Internets*. Transcript.
- Saracevic, T. (2016). *The Notion of Relevance in Information Science: Everybody knows what relevance is. But, what is it really?* Springer. <https://doi.org/10.2200/S00723ED1V01Y201607ICR050>
- Schultheiß, S., Häußler, H., & Lewandowski, D. (2022). Does Search Engine Optimization come along with high-quality content? *ACM SIGIR Conference on Human Information Interaction and Retrieval* (S. 123–134). <https://doi.org/10.1145/3498366.3505811>
- Schultheiß, S., & Lewandowski, D. (2020). “Outside the industry, nobody knows what we do” SEO as seen by search engine optimizers and content providers. *Journal of Documentation*, 77(2), 542–557. <https://doi.org/10.1108/JD-07-2020-0127>
- Schultheiß, S., & Lewandowski, D. (2021a). Misplaced trust? The relationship between trust, ability to identify commercially influenced results and search engine preference. *Journal of Information Science*, 016555152110141. <https://doi.org/10.1177/01655515211014157>

- Schultheiß, S., & Lewandowski, D. (2021b). How users' knowledge of advertisements influences their viewing and selection behavior in search engines. *Journal of the Association for Information Science and Technology*, 72(3), 285–301. <https://doi.org/10.1002/asi.24410>
- Schultheiß, S., & Lewandowski, D. (2022). Data set of a representative online survey on search engines with a focus on search engine optimization (SEO): A cross-sectional study [version 2; peer review: 2 approved]. *F1000Research*, 11(376). <https://doi.org/10.12688/f1000research.109662.2>
- Stiftung Warentest. (2022). Wie die Stiftung Warentest arbeitet. <https://www.test.de/unternehmen/testablauf-5017344-0/>
- Sullivan, D. (2013). FTC Updates Search Engine Ad Disclosure Guidelines After “Decline In Compliance.” Search Engine Land. <http://searchengineland.com/ftc-search-engine-disclosure-164722>
- Sundin, O., Lewandowski, D., & Haider, J. (2022). Whose relevance? Web search engines as multisided relevance machines. *Journal of the Association for Information Science and Technology*, 73(5), 637–642. <https://doi.org/10.1002/asi.24570>
- Urman, A., Makhortykh, M., & Ulloa, R. (2021). The Matter of Chance: Auditing Web Search Results Related to the 2020 U.S. Presidential Primary Elections Across Six Search Engines. *Social Science Computer Review*, 40(5), 1323–1339. 089443932110068. <https://doi.org/10.1177/08944393211006863>
- Yagci, N., Sünkler, S., Häußler, H., & Lewandowski, D. (2022). A Comparison of Source Distribution and Result Overlap in Web Search Engines. *Proceedings of the 85th Annual Meeting of the Association of Information Science and Technology, Pittsburgh, PA*. <https://doi.org/10.1002/pr2.758>

In: A. Imeri, K. Scheibe, F. Zimmer (Hrsg.): *Informationswissenschaft im Wandel. Wissenschaftliche Tagung 2022 (IWWT22)*. Düsseldorfer Konferenz der Informationswissenschaft, 6.–7. Oktober 2022, Haus der Universität Düsseldorf. Glückstadt: Verlag Werner Hülsbusch, S. 123–144. DOI: <https://doi.org/10.5281/zenodo.7457708>.