

Content & SEO: Failure is predictable



Thomas Mindnich

Spätestens seit Content-Marketing auch die Welt der SEOs im Handstreich erobert hat, gibt es keinen Mangel mehr an Ansichten, Anleitungen und Analysen über die perfekte Content-Strategie und „die richtigen Inhalte“. Warum aber erzielen auch viele dieser neuen Inhalte nicht die gewünschten SEO-Erfolge? Ist eine Suchmaschine wie Google nicht längst schlau genug, Webseiten-Inhalte ebenso wie Menschen zu verstehen? Oder muss man sich doch zwischen SEO-Texten für Suchmaschinen und dem besten Inhalt für menschliche Besucher entscheiden?

DER AUTOR



Thomas Mindnich trägt seinen schwarzen Hut mit Stolz und liebt SEO und Algorithmen. Er ist der nerdige Kopf hinter TermLabs.io - einer Tool-Sammlung für die datengetriebene Content-Arbeit.

„Failure is predictable – nur Erfolge sind extrem schwer zu prognostizieren!“

Durch einen kurzen Ausflug in die Statistik wird schnell klar, warum herkömmliche Optimierungsansätze häufig nicht zu den gewünschten positiven Ergebnissen führen.

Statistische Stichprobenverzerrung

Im England zu Zeiten des Zweiten Weltkriegs überlegten sich englische Ingenieure, wie sich die Überlebensrate der eigenen Kampfpiloten steigern ließe. Sie warfen einen genauen Blick auf die Verteilung der Einschusslöcher bei zurückgekehrten Flugzeugen. An den am stärksten betroffenen Stellen wurde anschließend die Panzerung verstärkt.

Allerdings ergab sich hierdurch keine Verbesserung der Überlebensrate. Wäre es nicht sinnvoller gewesen, die Flugzeuge dort stärker zu panzern, wo wenige oder keine Einschusslöcher zu finden waren? Denn dies waren ja offensichtlich die problematischen Stellen, die wahrscheinlich den Absturz auslösten. Auch damals wurde

SURVIVORSHIP BIAS

Dieser Fehlschluss ist bekannt als Survivorship Bias bzw. Verzerrung zugunsten der Überlebenden. Die „Überlebenden“ verzeichnen im Alltag eine größere Sichtbarkeit im Vergleich zu den „Verstorbenen“. Somit kann der Survivorship Bias dazu führen, dass die Chancen auf Erfolg überschätzt werden, da die Misserfolge übersehen werden.

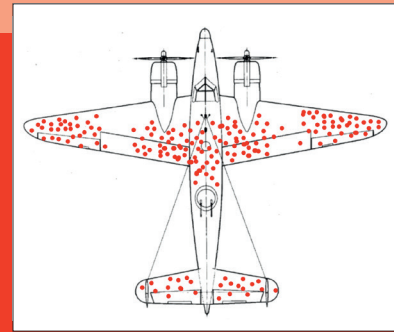
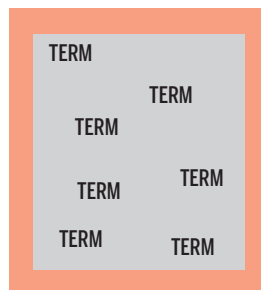
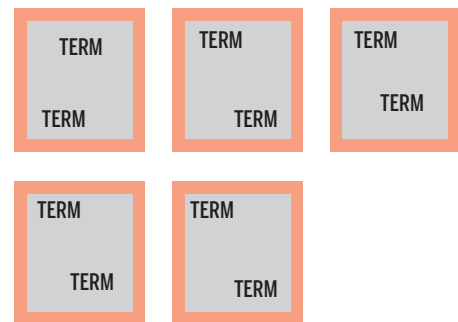


Abbildung 1: Survivorship Bias als Fehlschluss (Quelle: de.wikipedia.org/wiki/Survivorship_Bias; McGeddon/Wikimedia)

Relevanz: TF-IDF (WDF*p*IDF)



Die Relevanz ist **größer**, je häufiger ein Term **im Dokument** vorkommt.



Die Relevanz ist **kleiner**, je häufiger ein Term **in anderen Dokumenten** vorkommt.

Abbildung 2: Relevanz mit TF-IDF einfach erklärt

der falsche Blickwinkel erkannt, vom Mathematiker Abraham Wald.

Ebenso betrachten SEOs gerne die Top 10 oder Top 100 in Google und ziehen Schlussfolgerungen aus entsprechenden Analysen. Hierdurch entstehen einerseits die bekannten Fehlschlüsse von Korrelation auf Kausalität. Andererseits sind die Algorithmen in Zeiten von Machine Learning längst zu komplex, als dass man durch

den Blick auf „nur die Gewinner“ in den Top-Platzierungen valide und gestützte Handlungsanweisungen ableiten könnte.

Viele Basics der Suchmaschinenoptimierung sind eigentlich Hygienefaktoren und erzeugen gar nicht initial selbst positive Ranking-Verbesserungen.

Die „Verlierer“ zeigen aber üblicherweise eindeutige Fehler, die daran schuld sind, dass die gewünschten Rankings nicht erreicht werden können. Diese fehlende Sichtbarkeit ist meist ursächlich auf mindestens eine der folgenden drei Fehlerquellen zurückzuführen. In handwerklich sauber aufgesetzten Tests kann man diese Fehler häufig mit relativ wenig Aufwand signifikant als ursächlich nachweisen.

INFORMATION-RETRIEVAL-THEORIE

Die grundlegenden Funktionsprinzipien moderner Suchmaschinen sind glücklicherweise keine Geheimwissenschaften. Sie basieren auf den Theorien des Information Retrieval (IR), ein Fachgebiet, welches sich mit dem computergestützten Suchen nach Inhalten beschäftigt. Viele Grundlagen und Strategien wurden bereits rund um 1950, also schon lange vor dem Internet, erforscht und beschrieben.

Besonders deutlich kann man auch heute noch viele dieser IR-Verfahren im Kontext der informationellen Suchanfragen beobachten und nachweisen. Eine mehr als interessante Sammlung aktueller Publikationen findet sich direkt bei Google Research unter <https://research.google.com/pubs/InformationRetrievalandtheWeb.html>.

Termgewichtungsanalyse des Dokuments

Welches Buch bekommt man in einer guten Buchhandlung, wenn man nach einem Reiseführer für Rom fragt – einen Reiseführer durch Europa, einen für die schönsten Städte in Italien oder die Spezialausgabe für Rom selbst?

Die höchste Wahrscheinlichkeit auf zufriedene Kunden wird durch die Spezialausgabe für Rom selbst erreicht. Die anderen vom Thema übergeordneten Ausgaben enthalten ggf. auch Reiseinfos zu Rom, doch liegt nur bei der Spezialausgabe für Rom der Fokus klar und eindeutig!

Auch bei Webdokumenten können die Fokussierung und das Themengebiet leicht und zuverlässig über Termgewichtungsanalysen erkannt werden. Dafür werden die Dokumente in die einzelnen Wort-Bestandteile zerlegt (Tokenizing) und diese Terme anschließend mittels Stammformreduktion (Stemming) auf ihren Wortstamm reduziert, damit Konjugationen und Deklinationen bestmöglich zusammengefasst und nicht getrennt erfasst werden.

Jetzt lassen sich die Relevanz-Algorithmen aus der TF-IDF-Klasse (in Deutschland gerne WDF*p*IDF

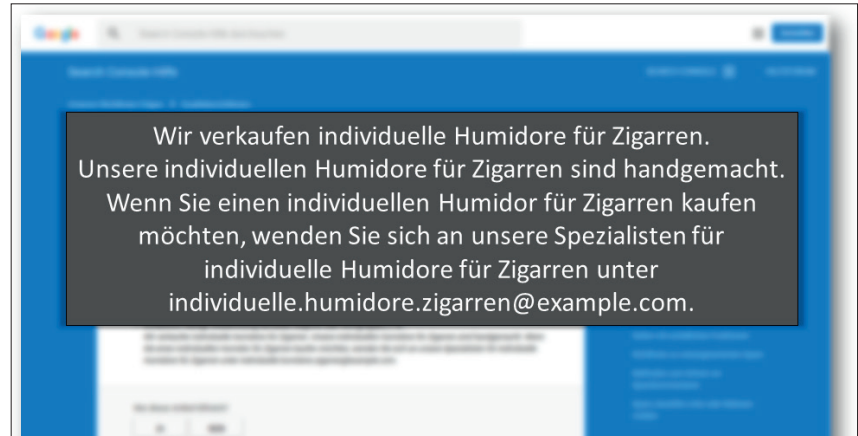


Abbildung 3: Googles Beispiel für Keyword-Stuffing (Quelle: <http://einfach.st/humi>)

genannt) anwenden. Verpasst man diesen Termen (oder genauer Wortstammgruppen) nun eine Ranglisten-sortierung, erhält man auf den ersten Plätzen zuverlässig die prägenden Terme des analysierten Textes oder Dokumentes. Im Vector-Space-Modell aus der Information-Retrieval-Theorie werden diese TF-IDF-Relevanzen z. B. als Values für den Dokumenten-Vektor benötigt.

Die Algorithmen aus der TF-IDF-Klasse sind eigentlich seit einigen Jahrzehnten fest verankert im Information Retrieval, obwohl sich unter SEOs eher fehleranfällige Konzepte wie Keyword-Dichte aufrechterhalten haben.

Auch finden sich diese Algorithmen in den meisten Open-Source-Suchlösungen wie z. B. Lucene, Elasticsearch oder Solr.

Exemplarische Termgewichtungsanalyse der drei Reiseführer:

- » Europa: Spanien, Italien, Mittelmeer, Strandurlaub, ...

„Benutzt ein Schüler andere Wörter als der Rest der Klasse, ist das Aufsatz-Thema verfehlt!“

- » Italien: Toskana, Venedig, Florenz, Rom, ...
- » Rom: Rom, Vatikanstadt, Kolosseum, Petersdom, Tiber, ...

Algorithmisch ist der beste Treffer zur Suchanfrage „Rom“ somit klar der Spezialreiseführer über Rom, weil hier der prägende Term „Rom“ direkt an erster Stelle vorkommt. So lassen sich Web-Dokumente inhaltlich erfolgreich verorten und anschließend Suchanfragen zuordnen.

Wer einen Blick in die Richtlinien von Google wirft und sich das Beispiel (siehe Abbildung 3) zum verbotenen Keyword-Stuffing („überflüssige Keywords“) ansieht, stellt schnell fest, dass das Ziel dieser Taktik eine künstliche Erhöhung der Relevanz-Ausprägung der Ziel-Terme „individuelle Humidore für Zigarren“ ist. Heute im Übrigen längst keine sinnvolle Optimierungstaktik mehr.

GOOGLE QUALITY RATER GUIDELINES

Was sind die Google Quality Rater Guidelines?

Für die manuelle Prüfung der Webseiten-Qualität durch sogenannte „Search Quality Evaluators“ (früher Quality Rater genannt) hat Google einen umfangreichen Richtlinien-Katalog, die „Quality Rater Guidelines“, erstellt. In diesem Leitfaden stehen detaillierte Regeln und Hilfestellungen, wie die qualitative Einschätzung von Webseiten zu erfolgen hat.

Ziel für Google ist die Bewertung der eigenen Suchleistung z. B. nach algorithmischen Updates. Ein direkter Einfluss auf einzelne Rankings existiert hingegen nicht.

Weiterführende Infos finden sich auf <https://www.sem-deutschland.de/die-google-quality-rater-guidelines-was-seos-daraus-lernen-koennen/>.



Abbildung 4: Kookkurrente Terme für „individuelle Humidore für Zigarren“ (Quelle: Google)

Text-Platzierung: Main Content (Hauptinhalt)

Aber auch, in welchen Seitenbereichen die Elemente und Inhalte auf der Webseite vorkommen, ist von Wichtigkeit. Denn genauso wie nicht jeder Bereich, nicht jeder Satz und schon gar nicht jedes Wort subjektiv die gleiche Wertigkeit durch den Leser erhält, bewertet auch eine Suchmaschine die verschiedenen Layout-Positionierungen (Title, Menü, Teaser, Footer, Sidebar oder Hauptinhalt) unterschiedlich.

So finden nicht alle Terme auf der Webseite die gleiche Gewichtung im Dokumenten-Vektor. Hierdurch wird erfolgreich sichergestellt, dass die Webseite vor allem für die Hauptinhalte und nicht für Nebenthemen oder untergeordnete Elemente gefunden wird.

In den Google Quality Rater Guidelines (<https://static.googleusercontent.com/media/www.google.com/de/inside-search/howsearchworks/assets/search-qualityevaluatorguidelines.pdf>) lässt sich ein Blick auf die Einordnung aus Sicht von Google in die drei Bereiche Main Content (Hauptinhalte), Supplementary Content (Zusatzinhalte) und Advertisements (Werbung) werfen.

Eine sinnvolle und korrekte HTML-Tag-Struktur und ein eindeutiges Layout unterstützen sowohl die Suchmaschine bei der algorithmischen Zuordnung der Seitenbereiche als auch den Leser beim Konsumieren der Webseiten.

Treffe ich das SERP-Thema? (Ähnlichkeit)

Woran merkt der Lehrer sofort, dass ein Schulaufsatz nicht vom vorgegebenen Thema handelt? Natürlich wenn keine der in diesem Kontext sinnvollen Wörter enthalten sind!

Auch als SEO kann man sich diese Logik zunutze machen. Und glücklicherweise ist es nicht nötig, aufwendig Reverse Engineering zu betreiben. Die dahinterliegenden Algorithmen der Suchmaschine müssen weder identifiziert noch verstanden werden. Denn über die Suchergebnisse selbst zeigt Google nämlich freizügig, wie der Kontext der entsprechenden Suchanfrage intern eingeschätzt wurde.

Dieser Kontext erschließt sich in einer Analyse der Term-Kookkurrenzen. Mit Term-Kookkurrenz ist das gemeinsame Auftreten zweier Wörter im nahen Kontext bzw. im gleichen Dokument gemeint. Siehe Abbildung 4 am Beispiel der von Google für Keyword-Stuffing genutzten Query „Humidore für Zigarren“.

Enthält der zu überprüfende Inhalt nicht genügend dieser in der Suchergebnis-Analyse festgestellten kookkurrenten Terme, dann sinkt die Wahrscheinlichkeit dramatisch, dass die Suchmaschine das getestete Dokument als validen Treffer für die entsprechende Suchanfrage einschätzt. Ein echter Mensch wird ohne die üblichen und erwarteten Wörter von diesem Inhalt gleichermaßen wenig begeistert sein.

Google Panda-Update

Spätestens seit dem Panda-Update im Jahr 2011 hat Google Verfahren implementiert, die qualitativ schlechte Inhalte datengetrieben erkennen sollen. Abseits der thematischen Ähnlichkeit sind auch grammatikalische oder gar semantische Term-Abhängigkeiten ursächlich für das gemeinsame Term-Auftreten im engen Kontext (Kollokation).

Mithilfe von N-Gram-Algorithmen ist diese Prüfung relativ unaufwendig. Diese finden zuverlässig alleine auf Basis der statistischen Daten aus den N-Grammen heraus, welche Inhalte von der Norm abweichen. Daraus kann dann eine deutlich höhere Wahrscheinlichkeit für mindere Qualität gefolgert werden.

Anmerkung: Ein schönes Beispiel, dass nicht immer eine künstliche Intelligenz (KI) beteiligt ist, wenn für den Betrachter von außen der Eindruck entsteht, dass Google Grammatik oder gar Semantik versteht und nutzt.

Suchabsichten

Vor allem seit dem Hummingbird-Update richtet Google seine Algorithmen und Ranking-Faktoren immer mehr auf die eigentlichen Suchabsichten der Suchenden aus. Viele SEOs nehmen an, dass vor allem Rankbrain Google dabei hilft, den User-Intent (also die Informationsabsichten des Suchenden) besser einzuschätzen.

Wer den User-Intent der Suchanfrage nicht berücksichtigt und bestmöglich befriedigt, hat heute längst keine Chancen mehr auf Top-Rankings. Außerdem macht es selten Sinn, unterschiedliche Intents in einem einzigen Dokument zu vermischen. Diese negativen Auswirkungen lassen sich z. B. bei Webshops beobachten, die plötzlich als informationelle Treffer eingeordnet werden und damit die vorderen Platzierungen bei transaktionalen Suchanfragen verlieren.

„Google Quality Rater Guidelines => „Main Content (MC) is any Part of the page that directly helps the page achieve its purpose. MC is (or should be!) the reason the page exists.“

Schon bei der Keyword-Recherche empfiehlt sich also die Bestimmung der jeweiligen Suchabsicht hinter dem Keyword. Der Blick auf die bisher in den Top 10 rankenden Seiten verrät schnell, wie Google bislang den Intent einschätzt und deswegen entsprechend ausliefert. Wichtig ist mehr die eigene Zielsetzung der Webseite und weniger die spezifische Reichweite des Keywords.

Task Completion Probability

Außerdem sollten Metriken zu Messung der Suchqualität selbst nicht unterschätzt werden. Eine wichtige Metrik ist die sog. „Search Task Completion Probability“, also die Wahrscheinlichkeit, dass durch die Suchergebnisseiten das Informationsbedürfnis eines Suchenden gestillt wurde. Erfassen lassen sich solche Metriken in ausreichender Aussagekraft nur durch die Suchmaschine selbst – vor allem über die Nutzersignale und die Interaktion mit den Suchtreffern.

AUSBLICK

Für eindimensionale Suchanfragen zeigt Google vermehrt Featured Snippets („Hervorgehobene Snippets“) an. Dieses Format erscheint über dem ersten organischen Treffer auf „Position 0“ und soll dem Nutzer eine direkte und prägnante Antwort oder Erklärung auf seine Fragestellung bieten, ohne dass man sich durch die üblicherweise angebotenen Suchergebnisse durchklicken muss.

Ganz neu hat Google mit „Talk to Books“ (<https://books.google.com/talktobooks/>) einen weiteren Schnellzugriff für Suchanfragen in Büchern vorgestellt. Ganz im Gegensatz zu bisherigen Information-Retrieval-Ansätzen wurde hier ein Machine-Learning-Modell zur Beantwortung der Anfragen trainiert. Dieses hat auf Basis von Milliarden Anfrage-Rückgabe-Paaren gelernt, Vorhersagen über gute und passende Antworten zu gestellten Fragen zu geben.

Spätestens bei diesen neuen konversationellen Suchen lässt sich nur schwer (bei Featured Snippets) oder gar nicht mehr

(bei Talk to Books) prognostizieren, wie genau hierfür zu optimieren ist. Aber auch hier gilt: Sind die relevanten Inhalte der Antwort nicht vorhanden, gibt es auch keinen Treffer.

Heißt: Passen die Terme (und damit die Sicht auf die Daten-Ebene) nicht zur üblichen Verwendung (Benchmark) oder den Trainingsdaten, dann sinkt auch hier die Wahrscheinlichkeit für einen Treffer rapide!

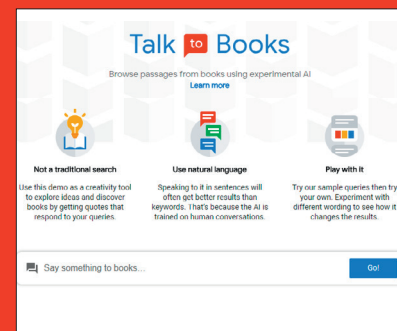


Abbildung 5: Talk to Books (Google)

Möglichst viele der vorhandenen Features des adressierten Keyword-Intents abzudecken, kann vor allem bei informationeller Ausrichtung mit den sog. holistischen Dokumenten eine sinnvolle Optimierungstaktik sein. Rein statistisch gesehen ist das eine Verdrängungsstrategie von bereits existierenden Treffern mit geringer thematischer Breite oder inhaltlicher Tiefe.

Fazit: Für Besucher gemacht, von Google verstanden!

Google betrachtet längst Suchabsichten hinter den Suchanfragen, um die dazu passenden und relevanten Webseiten auszuliefern. Die Zielsetzung ist klar: Größtmögliche Nutzerzufriedenheit erreichen.

Dabei verlässt sich Google nicht auf eine ominöse KI (künstliche Intelligenz), sondern wertet mit statistischen Verfahren (z. B. Machine Learning und Natural Language Processing) Inhalte und Nutzerverhalten aus.

Also müssen Seitenbetreiber dafür Sorge tragen, dass ihre Webseiten aus Datensicht richtig interpretiert werden können und gleichsam das erwünschte Nutzerverhalten auslösen – immer mit einer auf die eigenen Ziele ausgerichteter Optimierungsstrategie.

Es gilt, die angesprochenen drei Fehlerquellen selbst zu vermeiden. Denn nachweislich sind dies auch heute noch grundlegende Auslöser für fehlende oder schlechte Sichtbarkeit, sodass auch durch andere zusätzliche Optimierungstaktiken nur selten das Ruder herumgerissen werden kann. ¶