



DIE SMX 2024 – DIE LEITKONFERENZ BEIM SEARCH MARKETING

Mario Fischer

Immer im März pilgern die Search Marketingprofis nach München, um sich vor allem auch international über die neuesten Trends zu informieren und sich mit frischem Know-how versorgen zu lassen. Keine Branchenkonferenz hat einen höheren Anteil an internationalen Speakern, die es erlauben, über den deutschsprachigen Tellerrand hinaus zu blicken. Die Teilnehmerinnen und Teilnehmer kamen aus 26 Ländern, was den länderübergreifenden Charakter mehr als verdeutlichte. Doch auch das ebenso wichtige Networking mit alten Bekannten oder auch neuen Kontakten war durch eine grandiose Party am ersten Abend in jedem Fall sichergestellt.

Website Boosting hat sich wie immer an den beiden Tagen für Sie umgeschaut und einige Highlights für Sie herausgepickt.

Nach der Eröffnung der SMX durch die Veranstalterin Sandra Finlay startete Rank Fishkin mit einer Keynote, die den Tod der Marketing Attribution vorhersagte. Er zeigte, dass die sogenannten 3rd Party Cookies durch die zunehmende Blockierung für das Tracken der Customer Journey praktisch nahezu unbrauchbar geworden sind. Mittlerweile werden auch etwa 20 % der Seitenzugriffe von Besuchern für die Erfassung in Google Analytics geblockt. Dazu käme erschwerend hinzu, so Fishkin, dass der Anteil der Klicks in den Suchergebnissen auf etwa ein

Drittel zusammengeschrumpft ist, während der Rest durch direkte Informationen von Google gar keine Klicks auf Websites mehr auslöst. Als wenn das nicht genug wäre, verstecken große Plattformen wie TikTok, WhatsApp, Discord, Slack und viele andere ihren Referrer und werden in Google Analytics als „Direct Traffic“ ausgewiesen. „Dark Social“ nannte Fishkin diese Trafficquellen. Hier findet also schon seit Längerem eine dramatische Verzerrung der Eingangskanäle statt. Viele dieser Plattformen erschweren das „Wegklicken“ von dort, indem sie keine URLs in Postings zulassen.



Abb. 1: Der „Direct“-Eingangskanal ist durch Dark Social für Analysen nahezu unbrauchbar geworden (Quelle: Rand Fishkin).

sen oder Postings mit URLs bewusst in der Reichweite beschränken.

Als Lösung bliebe nur, entweder den großen Werbeplattformen noch mehr Geld hinterherzuwerfen für Verkäufe, die man am Ende wahrscheinlich sowieso bekommen hätte. Oder man investiert in den Aufbau eines Dashboards, das versucht, den Anstieg von Kanälen abzuleiten. Dort misst man dann vor allem Traffic, Conversions und Umsatz. Alles andere solle man getrost vergessen und sowohl bezahlte Werbung als auch die Investition in organisches Ranking einstellen. Den Ausweg brächten schwer zu fassende Metriken wie die Besucherpräsenz, das

Bauchgefühl und „Eitelkeitsmetriken“ (O-Ton). Man solle sich wieder mehr zum Beispiel auf Newsletter und Blogposts zur Kundenakquise konzentrieren und – vereinfacht ausgedrückt – einfach den Mehrwert bieten, den man nirgends anders bekommt.

„Don't let people click out!“, Rand Fishkin über das Mantra Sozialer Netzwerke.

Bastian Grimm klärte über die Möglichkeiten des Einsatzes von künstlicher Intelligenz für Technical SEO auf. Sogenannte LLM (Large Language Models) sind gut, Datenmuster zu erkennen und zu extrahieren, was bei der Entscheidungsfindung sehr nützlich sein kann. Ein ebenso gutes Einsatzgebiet ist die Analyse von Besucherpräferenzen und -verhalten, um personalisierte Empfehlungen geben zu können. Durch das gute Sprachverständnis von LLMs sind sie laut Grimm hervorragend für Chatbots, Übersetzungen, Stimmungsanalysen oder auch Textzusammenfassungen geeignet.

Für alles, was in Trainingsdaten-sätzen nicht vorhanden ist, taugen solche Modelle allerdings nichts. Kontext wirklich zu verstehen oder gar eigenes Wissen zu generieren, ist derzeit nicht möglich. Ethik oder Moral beherrschen



Abb. 2: Rand Fishkin hielt die erste Keynote der SMX.

solche Systeme nicht, auch wenn gut formulierte Antworten das vielleicht ab und zu vermuten lassen. Sie „raten“ eine Antwort aufgrund von Textwahrscheinlichkeiten und verzerren Ergebnisse, wenn ihre Trainingsdaten ebenfalls einseitige Beziehungen enthalten. Was oft als „Halluzinieren“ bezeichnet wird, bedeutet am Ende nichts anderes als: „falsch geraten“.

Eine KI „schreibt“ nichts, sie stellt Worte zufällig nach statistischen Regeln aus ihrem Trainingsset zusammen.

Einer der vielen nützliche Tipps von Grimm zum Einsatz von KI im SEO-Bereich war die Analyse zweier zeitlich auseinanderliegender Website-Crawls mit dem Tool Screaming Frog. Eine Maschine erzeugt daraus blitzschnell Weiterleitungsketten bzw. ermittelt, wohin alte bzw. gelöschte URLs am besten weitergeleitet werden sollten. Das Ergebnis wäre nicht perfekt, so Grimm, aber ca. 90 % aller Vorschläge wären direkt brauchbar.

Besondere Einblicke in das Crawling und die Indexierung bei Google gab Johan von Hülsen. Er schätzte aufgrund



Abb. 3: Bastian Grimm gab nützliche Tipps für KI-Einsatz für SEO.

einiger Quellen, dass Google etwa 400 Milliarden Dokumente im Index halte. Im Hinblick darauf, dass Google jeden Tag mehr als 200 Milliarden Spam-URLs aussortiert, ist das vergleichsweise gar nicht so viel. Da das „normale“ Web pro Jahr hinsichtlich des Inhalts um etwa 10 % wächst, die Größe des Index aber offenbar relativ konstant bleibt, fallen zwangsweise viele Seiten wieder aus dem durchsuchbaren Index heraus. Deutsche Websites stellen nur zwischen 3 und 5 % des gesamten Index (zum Vergleich zum Beispiel Englisch ~19 %, Mandarin ~24 %, Hindu ~8 %, Spanisch ~7 %). Das würde bedeuten, dass der deutsche Index vermutlich zwischen zehn und 20 Milliarden Dokumente enthält. Seine Schlussfolgerung war, dass es bei konstanter Indexgröße, aber wachsendem Volumen verfügbarer (neuer) Seiten immer schwieriger wird, in den Index zu kommen bzw. dort zu bleiben.

„Hast du ein Crawling-Problem, dann ist es oft ‚Google will nicht‘, statt ‚Google kann nicht‘“, Johann von Hülsen.

Wer in der Search Console eine nennenswerte Anzahl von URLs gelistet bekommt, die nicht im Index vertreten sind, dann ist es nach seiner Einschätzung oft so, dass Google die URLs nicht will und nicht, dass diese nicht gecrawled werden könnten. Als einfache Regel meinte von Hülsen, dass man ein Problem hätte, wenn in der Search

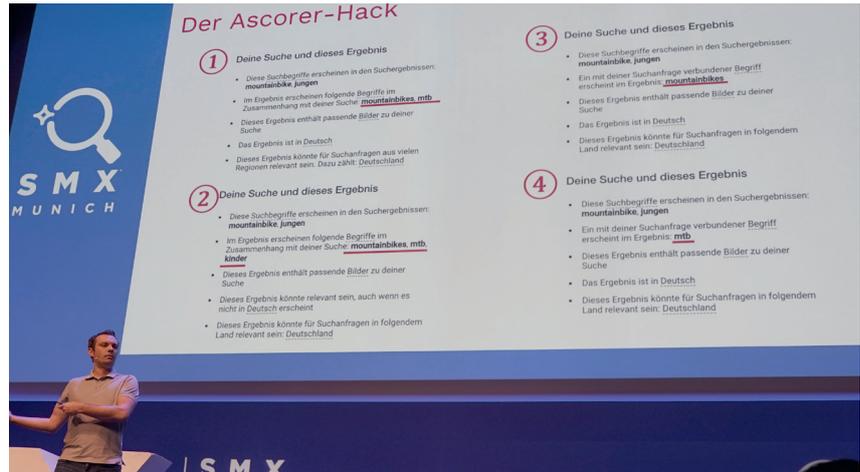


Abb. 4: Johan von Hülsen gab interessante Einblicke in das Thema Ranking und den Dokumentenindex bei Google.

Console die Anzahl der URLs unter „Discovered – currently not indexed“ (Abbildung 4, Ziffer 2) größer ist als die unter „Crawled – currently not indexed“ (Abbildung 4, Ziffer 1). Hier sollte man aktiv werden und sich die URLs genauer ansehen, die Google bewusst links liegen lässt. Typische Lösungen wären, die Content-Qualität zu prüfen und gegebenenfalls zu erhöhen oder für mehr intern eingehende Links zu sorgen. Auch das bewusste Ausschließen von Seiten in der robots.txt, die minderwertigen Content haben, kann die Indexierungsrate für die übrigen Seiten laut von Hülsen ansteigen lassen.

Besonders spannend war eine zufällige Entdeckung bei von Hülsens Analyse der Indexierungsdaten einiger Domains. Er bemerkte, das URLs, die länger als 129 Tage nicht mehr von Google gecrawled wurden, aus dem Index geflogen sind und gleichzeitig die Zahl unter „Crawled – currently not indexed“ (Abbildung 4, Ziffer 1) entsprechend ansteigt. Als er anschließend gezielt weitersuchte, fand er tatsächlich nirgends eine URL, die nach diesem Zeit-

raum (129 Tage) noch im Index war! Wir haben in der Redaktion anschließend ebenfalls nach URLs mit noch früherem Crawl-Datum gesucht und ebenfalls nichts gefunden. Wenn sich diese Vermutung weiter bestätigen lässt, wäre ein längeres Ausbleiben eines „Re-Crawls“ tatsächlich eine Art Ankündigung, dass die URL alsbald aus dem Index verbannt wird – und ein deutliches Signal, rechtzeitig zu reagieren. Das letzte Datum des Crawlens einer URL kann man übrigens recht leicht über die Nutzung/Aktivierung der URL-Inspection API, zum Beispiel über Tools wie Screaming Frog, ermitteln. Wie in Abbildung 6 gezeigt, aktiviert man unter „Crawl-Konfiguration“ bei „Google Search Console“ (Ziffer 1) zunächst den Zugang zur Search Console für die eigene Domain und anschließend unter dem Reiter URL-Inspektion“ (Ziffer 2) das Kästchen „URL-Inspektion aktivieren“ (Ziffer 3). Nach dem Start und Durchlauf des Crawls findet man weiter hinten in den Datenspalten, ob die URL auf Google ist (erste Spalte des Datenausschnitts in Abbildung 6). Bei URLs, die bereits nicht mehr im Index sind, ist die Anzahl der Tage (Ziffer 4) oft größer. Wichtig ist die Dauer für „URL is on Google“ (ab Zeile 4, gelb markiert). Sortiert man diese absteigend, erkennt man sofort Handlungsbedarf. Die mit roter Farbe markierten URLs sind im Beispiel zwischen 70 und 90 Tage schon nicht

Gecrawlt – zurzeit nicht indiziert	1	Google-Systeme	Nicht gestartet	241
Gefunden – zurzeit nicht indiziert	2	Google-Systeme	Bestanden	13

Zeilen pro Seite: 10 | 1 bis 7 von 7

Abb. 5: Search Console: Wenn die Zahl in der zweiten Zeile die der ersten Zeile übersteigt, sollte man reagieren (hier passt alles).

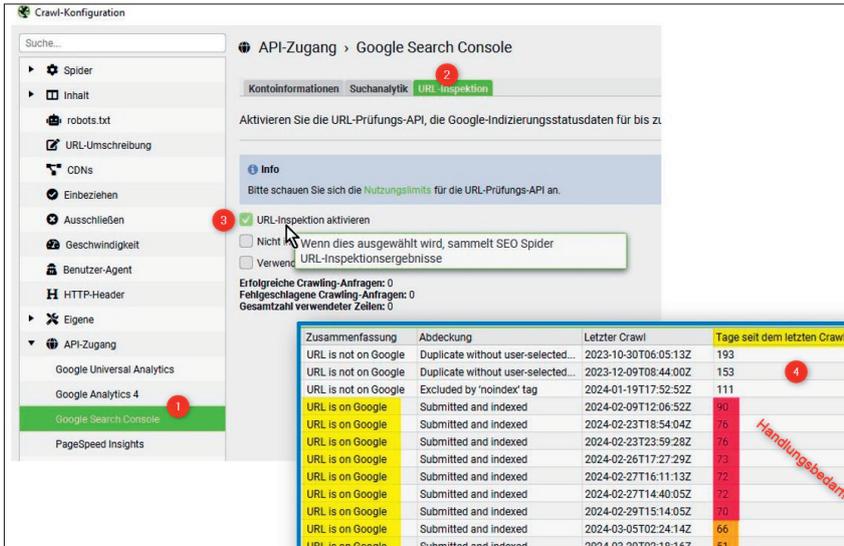


Abb. 6: Tipp: Mit Tools wie dem Screaming Frog die Anzahl der Tage seit dem letzten Crawling ermitteln



Abb. 7: Beatrice Eiring erklärte, wie man Content mit KI-Unterstützung erstellen kann und was dabei zu berücksichtigen ist.

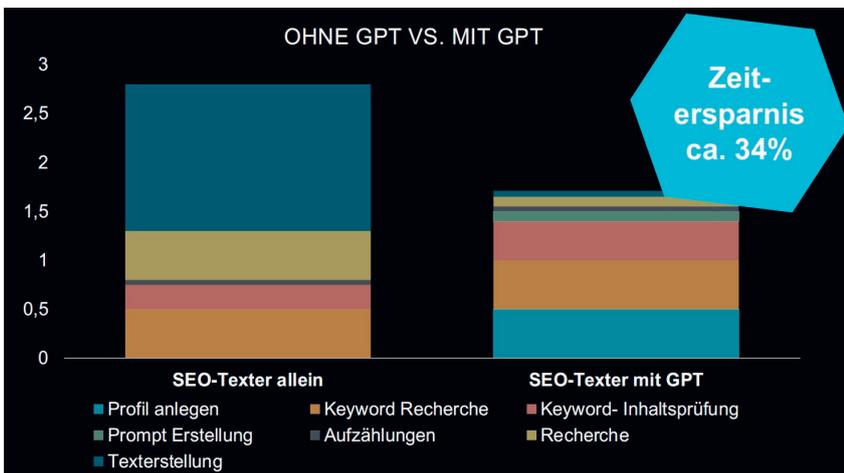


Abb. 8: Etwa ein Drittel Zeit konnte Eiring durch die Unterstützung von ChatGPT einsparen.

mehr aktualisiert wurden und werden nach der oben genannten These wohl in wenigen Wochen aus dem Index fallen.

Wenn Beatrice Eiring etwas zu Content und Sprache erklärt, empfiehlt es sich, gut zuzuhören. Sie hat in Sprachwissenschaften promoviert und verfügt über zehn Jahre praktische Erfahrung in diesem Bereich. Sie hat geprüft, ob komplexe Themen aus dem B2B-Bereich das sprichwörtliche Kryptonit für die KI sind oder man sie dafür bereits heute gut nutzen kann. Spoileralarm: Ohne menschlichen Sachverstand geht es auch hier nicht. ChatGPT 4 neigt bei komplexen Themen dazu, nur zwischen Aufzählungen und Überschriften zu wechseln, und verweigert gerne auch mal die Antwort. Ein Beispiel:

Chat GPT: „Leider kann ich den Text nicht bearbeiten oder formatieren, da das über meine aktuellen Fähigkeiten hinausgeht [...]“

Beatrice Eiring: „Doch, das kannst du!“

Man muss also nur beharrlich bleiben, denn nach dieser Belehrung hat ChatGPT das Gewünschte anstandslos abgeliefert.

„A fool with a tool is still a fool.“

Schlimmer ist allerdings, dass viele Aussagen gut klingen, aber nach der Einschätzung von Eiring einer fachlichen Prüfung nicht standhalten. Auch eine

Missachtung der Prompt-Anweisungen kommt nicht selten vor, hat die Expertin festgestellt. Als Lösung empfahl sie, im Dialogstil in das Sparring mit ChatGPT zu gehen und alles stückweise nach und nach zu prompten. Auch das Anlegen eines eigenen Prompt-Baukastens ist hilfreich, empfahl sie. Ihr Fazit über mehrere Testfälle hinweg war, dass bei Texten mit hoher Qualitätserwartung der Nachbearbeitungsaufwand deutlich ansteigt. Aber auch mit KI-Unterstützung sollte man nie aus dem Auge verlieren, dass man für die Besucher einen spürbaren Mehrwert und auch eine gewisse Einzigartigkeit beim Text erzeugen muss. In Zeiten der steigenden Textfluten mittels KIs wird dies zusammen mit einer starken Nutzerfokussierung explizit sogar noch wichtiger.

Wie immer muss der Versuch scheitern, auch nur annähernd aufzulisten, welche Themen, Tipps und Inspirationen über eine große Fachkonferenz an das Publikum weitergegeben werden. Vielleicht hat Sie unsere bescheidene und kurze Auswahl inspiriert, sich den Termin für die nächste SMX fest im Kalender zu notieren? Sie findet im nächsten Jahr am 18. und 19. März erneut in München im Messezentrum statt. ¶