

Olaf Kopp

MUM UND DIE GOOGLE-SUCHE DER ZUKUNFT



Die neueste Innovation aus dem Hause namens MUM ist ein weiterer großer Schritt in Richtung einer semantischen Suchmaschine, die das Wissen der Welt abbildet und Antworten zu allen Fragen der User nutzerfreundlich bereitstellt. Im gleichen Schritt ist MUM der Versuch seitens Google, verlorene Marktanteile zurückzugewinnen. Der Beitrag von Olaf Kopp gibt Ihnen einen Überblick darüber, wie Google heute und zukünftig funktioniert und was SEOs daraus lernen sollten.

DER AUTOR



Olaf Kopp ist Co-Founder, Chief Business Development Officer (CBDO) und Head of SEO der Aufgang GmbH. Er ist Autor, Podcaster sowie einer der bekanntesten Branchen-Experten für SEO, Online- und Content-Marketing-Strategien entlang der Customer Journey und digitalen Markenaufbau.

INFO

Was ist eine Entität?

Eine Entität ist ein Begriff aus der Philosophie, Semantik und Informatik. Eine Entität beschreibt das Wesen bzw. die Identität eines konkreten oder abstrakten Gegenstands des Seins. Entitäten sind eindeutig identifizierbar und damit einzigartig. Grundsätzlich kann zwischen „Named Entities“, zu Deutsch benannten Entitäten, und Konzepten unterschieden werden. Benannte Entitäten sind Objekte aus der echten Welt, wie z. B. Personen, Orte, Organisationen, Produkte, Events ... Konzepte sind abstrakte Entitäten physikalischer, psychologischer oder sozialer Natur, wie z. B. Entfernung, Quantität, Emotionen, Menschenrechte, Frieden ...

Im Kontext der Informatik stehen benannte Entitäten häufiger im Fokus.

Was bedeutet Semantik?

Semantik – auch als Bedeutungslehre bezeichnet – ist die Theorie oder Wissenschaft von der Bedeutung von Zeichen und Zeichenketten. Zeichen können in diesem Fall Wörter, Phrasen oder Symbole sein. Die Semantik beschäftigt sich typischerweise mit den Beziehungen zwischen den Zeichen und deren Bedeutungen.

Dabei wird unterschieden in folgende Hauptformen der Semantik:

- Semantik von Zeichensystemen (Semiotik)
- Semasiologie und Onomasiologie
- Semantik von Sprachen (Linguistik)

In Bezug auf die semantische Suchmaschinenoptimierung bzw. die semantische Suche und auch das Online-Marketing im Allgemeinen bewegen wir uns in erster Linie im Bereich der Linguistik.

Die Semantik spielt bei der Kategorisierung, Deutung und Auslieferung von Informationen im Internet eine immer größere Rolle. Viele große Online-Plattformen wie Facebook und Google greifen bereits auf semantische Funktionalitäten zurück.

Der Weg zu MUM: Googles Innovationen der letzten zehn Jahre zur semantischen Suchmaschine

Durch Googles große Neueinführungen bezüglich der Suche zieht sich ein roter Faden, der ein klares Ziel erkennen lässt: das vollständige Verständnis aller Inhalte im Netz und eindeutige Interpretation von Suchanfragen. Um die wirkliche Bedeutung von Suchanfragen und Interpretation von Inhalten zu erkennen, muss Google

1. Suchanfragen und deren Suchintention klar identifizieren,
2. Entitäten in Suchanfragen und Inhalten identifizieren und den Index um diese herum organisieren.

Um diese beiden Ziele zu erreichen, benötigt Google leistungsstarke Systeme und Algorithmen, die das komplette Wissen der Welt erschließen und deuten.

Dieses Ziel verfolgt Google seit der Einführung des Knowledge Graphen als semantische Datenbank im Jahr 2012 und der Einführung der neuesten Ranking-Algorithmus-Basis Hummingbird im Jahr 2013 konsequent. Die Erschließung des weltweiten Wissens bedarf eines Index, der als Graph aufgebaut ist. Während Google vor 2012 auf einen Index ähnlich einer tabellarischen Datenbank zugriff, in den Informationen ähnlich einem Verzeichnis gespeichert wurden, kann ein Graphen-Index Beziehungen zwischen Informationen und Entitäten erfassen und abbilden.

Die Entitäten in einer semantischen Datenbank wie dem Knowledge Graphen sind als Knoten erfasst und die Beziehungen werden als Kanten abgebildet. Die Entitäten können mit Labels zu z. B. zu Entitäts-Typen-Klassen, Attributen und Informationen zu in Beziehung zur Entität stehenden Inhalten und digitalen Abbildern wie z. B. Websites, Autorenprofilen oder Profilen in sozialen Netzwerken ergänzt werden.

In Abbildung 2 sehen Sie zur Verdeutlichung ein konkretes Beispiel von Entitäten, die in Beziehung zueinander

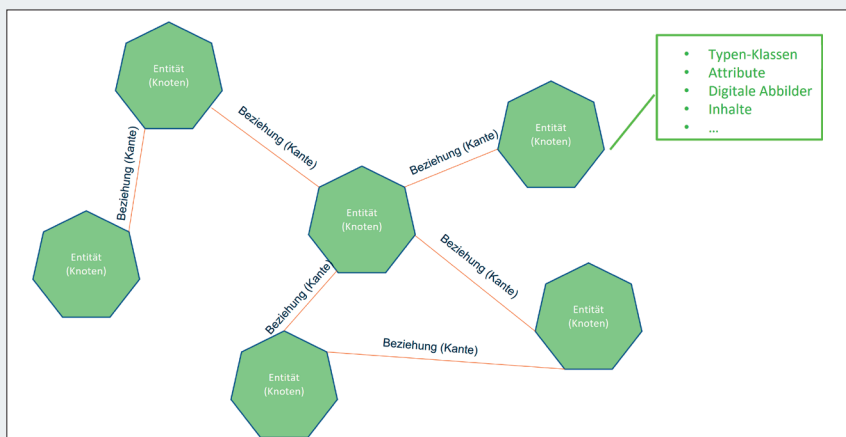


Abb.1: Aufbau eines Graphen-Index (Quelle: Olaf Kopp)

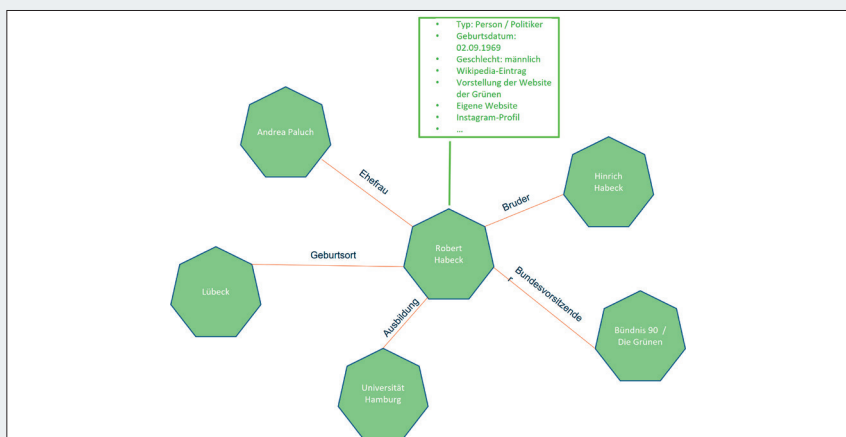


Abb.2: Aufbau eines Graphen-Index in der Praxis (Quelle: Olaf Kopp)



Robert Habeck
Politiker

Robert Habeck ist ein deutscher Politiker und Schriftsteller. Seit dem 27. Januar 2018 ist er gemeinsam mit Annalena Baerbock Bundesvorsitzender der Partei Bündnis 90/Die Grünen. Für die Bundestagswahl 2021 waren Baerbock und er das Spitzenduo der Grünen, wobei Baerbock Kanzlerkandidatin war. Wikipedia

Geboren: 2. September 1969 (Alter 52 Jahre), Lübeck

Ehepartnerin: Andrea Paluch (verh. 1996)

Partei: Bündnis 90/Die Grünen

Ausbildung: Universität Hamburg (2000), MEHR

Filme & Fernsehsendungen: Following Habeck, MEHR

Geschwister: Hinrich Habeck

Bücher Über 25 weitere ansehen

Von hier an anders: Eine poli... 2021

Ruf der Wölfe 2019

Patriotis... Ein linkes Plädoyer 2010

Wer wir sein könnten... 2018

Wird auch oft gesucht Über 10 weitere ansehen

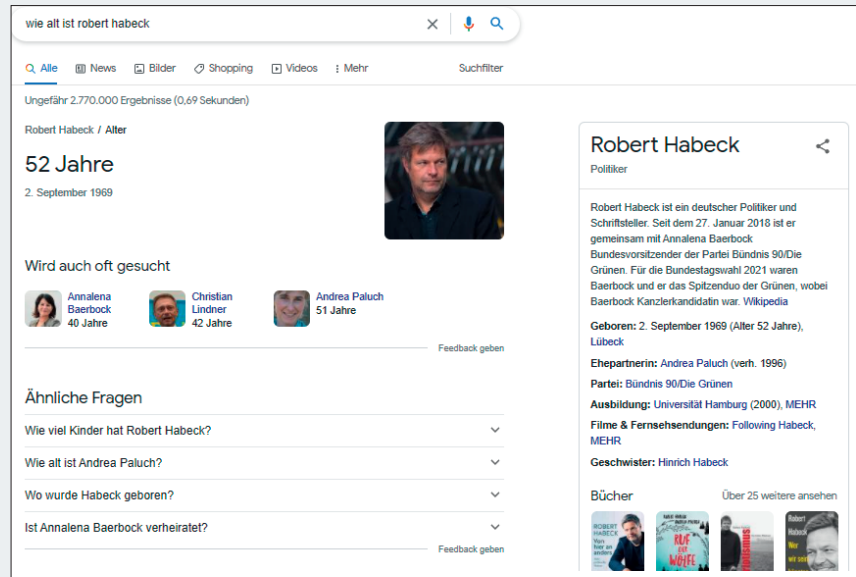
Helene Bubrowski Im Trend

Kevin Kühnert Im Trend

Andrea Paluch

Annalena Baerbock

Abb.3: Knowledge-Panel für Robert Habeck (Quelle: Google)



Wie alt ist Robert Habeck

Ungefähr 2.770.000 Ergebnisse (0,69 Sekunden)

Robert Habeck / Alter

52 Jahre
2. September 1969

Wird auch oft gesucht

Annalena Baerbock 40 Jahre

Christian Lindner 42 Jahre

Andrea Paluch 51 Jahre

Ähnliche Fragen

Wie viel Kinder hat Robert Habeck?

Wie alt ist Andrea Paluch?

Wo wurde Habeck geboren?

Ist Annalena Baerbock verheiratet?

Robert Habeck
Politiker

Robert Habeck ist ein deutscher Politiker und Schriftsteller. Seit dem 27. Januar 2018 ist er gemeinsam mit Annalena Baerbock Bundesvorsitzender der Partei Bündnis 90/Die Grünen. Für die Bundestagswahl 2021 waren Baerbock und er das Spitzenduo der Grünen, wobei Baerbock Kanzlerkandidatin war. Wikipedia

Geboren: 2. September 1969 (Alter 52 Jahre), Lübeck

Ehepartnerin: Andrea Paluch (verh. 1996)

Partei: Bündnis 90/Die Grünen

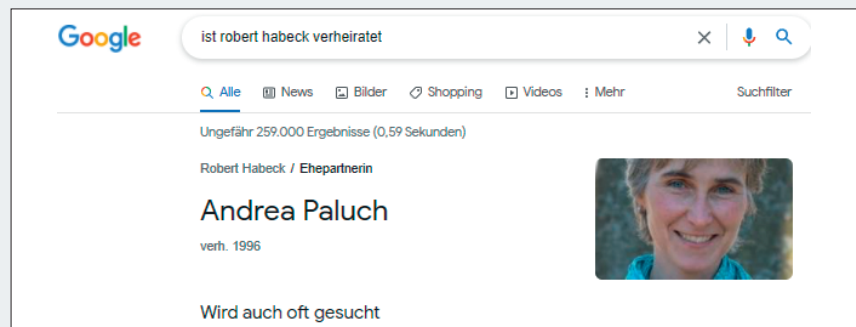
Ausbildung: Universität Hamburg (2000), MEHR

Filme & Fernsehsendungen: Following Habeck, MEHR

Geschwister: Hinrich Habeck

Bücher Über 25 weitere ansehen

Abb.4: Knowledge Card für das Attribut Alter der Entität Robert Habeck (Quelle: Google)



Google

ist Robert Habeck verheiratet

Ungefähr 259.000 Ergebnisse (0,59 Sekunden)

Robert Habeck / Ehepartnerin

Andrea Paluch
verh. 1996

Wird auch oft gesucht

Abb.5: Knowledge Card für die mit Robert Habeck verwandte Entität Andrea Paluch (Quelle: Google)

stehen. Die Haupt-Entität ist Robert Habeck und weitere Entitäten sind seine Ehefrau Andrea Paluch, die Partei die Grünen, sein Bruder Hinrich Habeck, die Stadt Lübeck und die Universität Hamburg.

Die in den SERPs angezeigten Knowledge-Panels sind Entitäten-Boxen, die zum einen die Attribute, den prominenten Entitäts-Typ beinhalten und zum anderen Beziehungen zu weiteren Entitäten abbilden (Abb. 3).

INFO

Was ist eine Ontologie?

Eine Ontologie ist ein Begriff aus der Semantik und beschreibt den Versuch, Dinge des Seins bzw. Entitäten in einen Kontext zu bringen bzw. in Kategorien einzusortieren. Dieser Kontext kann sich z. B. auf thematische Sachverhalte, Beziehungen oder Ereignisse beziehen.

Durch eine Entitäten-basierte bzw. semantische Suchmaschine lassen sich Beziehungen zwischen den Entitäten herstellen, was zu einem deutlich verbesserten Verständnis des Kontexts führt. Die Beziehungen zwischen Entitäten und Attributen sind in den semantischen Datenbanken wie z. B. dem Knowledge Graphen erfasst. Dadurch kann Google auch komplexere Long-Tail-Suchanfragen deutlich besser beantworten, z. B.: „wie alt ist Robert Habeck“, oder: „ist Robert Habeck verheiratet.“

Dadurch lassen sich konkrete Fragen schnell auch via Voice Search beantworten, ohne dass jemand einen Klick in den SERPs machen muss.

Entitäten sind nicht nur eine Aneinanderreihung von Buchstaben, sondern sind Dinge mit einer eindeutig identifizierbaren Bedeutung. Der Suchterm „jaguar“ hat mehrere Bedeutungen. So

kann die Automarke gemeint sein, das Tier oder der Panzer. Die reine Verwendung der Zeichenkette in Suchanfragen und/oder Inhalten reicht nicht aus, um den Kontext zu verstehen. In Kombination mit anderen Attributen bzw. Entitäten wie Rover, Coventry, PS, Auto ... wird die eindeutige Bedeutung klar durch den Kontext, in dem die Entität sich bewegt. Die Bedeutung von Entitäten lässt sich über den Kontext ermitteln, in dem sie in Inhalten und Suchanfragen genannt bzw. benutzt werden.

Durch Vektorraumanalysen lassen sich Suchanfragen und Inhalte in einen thematischen Kontext bringen. Man kann Entitäten bzw. Keywords in einem Vektorraum verorten. Der Abstand der unterschiedlichen Begriffe zueinander gibt Auskunft über den thematischen Kontext, in dem die Entitäten bzw. Keywords genutzt werden. Über diesen

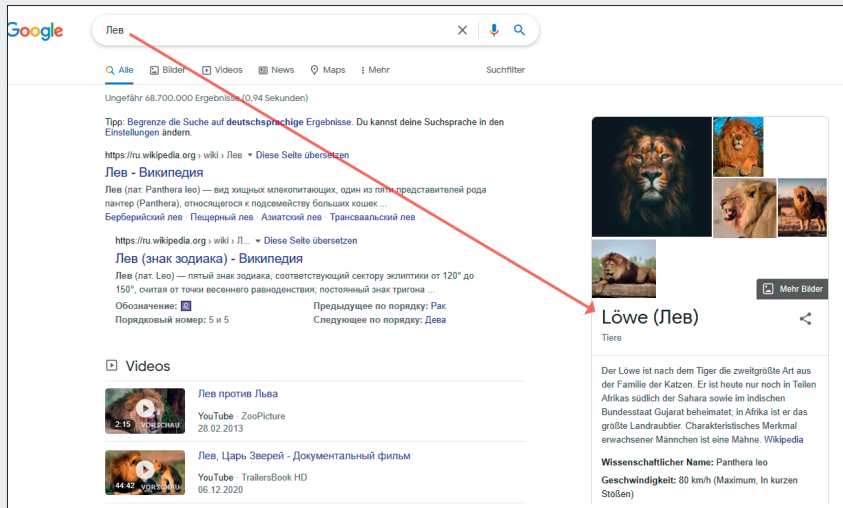


Abb.6: Beispiel eines Knowledge-Panels als Antwort auf eine Entitäten-basierte Suchanfrage in fremder Sprache (Quelle: Google)

Weg lassen sich thematische Ontologien bzw. Kategorien ermitteln, in denen Keywords bzw. Entitäten zu verorten sind.

Diese Methodik hat Google 2015 mit RankBrain offiziell eingeführt, um insbesondere Suchanfragen besser zu verstehen. RankBrain wurde als Innovation für das sogenannte Query Processing, also den Vorgang zur Interpretation von Suchanfragen, vorgestellt. Und mit RankBrain wurde auch das erste Mal seitens Google bestätigt, dass Machine Learning für die Google-Suche zum Einsatz kommt.

Vorteile eines Entitäten-basierten Index und Rankings

Die klassische Keyword-basierte Suche stößt bei Suchmaschinen schnell an ihre Grenzen, da im Endeffekt nur eine Aneinanderreihung von Buchstaben und Termen in Suchanfragen und Dokumenten miteinander abgeglichen wird, aber ein wirkliches Verständnis der Bedeutung nicht vorhanden ist. Google sprach deswegen bei der Einführung des Knowledge Graphen auch von „things, not strings“.

It's why we've been working on an intelligent model—in geek-speak, a “graph”—that understands real-world entities and their relationships to one another: things, not strings.

Dinge haben eine Bedeutung, Zeichenketten nicht. Der große Vorteil an dem Entitäten-basierten Ansatz zeigt sich z. B. bei Sprachen bzw. Internationalisierung. Die Bedeutung der Entität „Löwe“ ist unabhängig von der Sprache gleich. Die Zeichenketten in den unterschiedlichen Sprachen unterscheiden sich aber deutlich voneinander, nämlich Löwe (deutsch), lion (englisch), Лев (russisch), ライオン (japanisch) ... Kennt Google die Entität und deren Bedeutung, lässt sich für jede Sprache der Erde auf den gleichen Entitäts-Index zugreifen. Bisher musste Google für jede Sprache einen eigenen sprachenabhängigen Index pflegen.

Sucht man nach dem russischen Wort für Löwe bei Google, wird ein deutsches Knowledge-Panel ausgegeben.

Das zeigt sehr schön, dass Google unabhängig von der Zeichenkette und der Sprache den Begriff basierend auf der Bedeutung erkennt. Die Suchergebnisse auf der linken Seite werden allerdings noch nach dem Keyword-Term-Match-Prinzip ausgespielt.

Abgrenzung zwischen AI, Machine Learning und Natural Language Processing

Natural Language Processing ist ein Teilbereich von Artificial Intelligence, kurz AI. Im Kontext von Google bzw. Suchmaschinen-Technologien ist es aber

genauer, wenn man von Machine Learning spricht. Eine wirkliche künstliche Intelligenz, die ähnlich einem Menschen fühlt und handelt, gibt es (noch) nicht. In den Medien und von Vertretern der Unternehmen wird gerne von AI oder KI gesprochen, da es spektakulärer klingt als Machine Learning oder Deep Learning. Die heutzutage eingesetzten Technologien basieren aber auf Machine Learning, also einer Vorstufe von AI. Bei Machine Learning werden anhand statistischer Methoden die Algorithmen immer wieder mit Trainingsdaten gefüttert. Diese Trainingsdaten können auch als Erfahrungsdaten verstanden werden, um die Algorithmen schlauer zu machen. Aufgrund dieser Daten können bestehende Modelle validiert oder mit neuen Modell-Klassen ergänzt werden. So entsteht ein immer tieferes „Verständnis“ der Algorithmen zu Themen, Sachverhalten und auch Sprachen.

Die Hauptaufgabe von Algorithmen, die auf Machine Learning beruhen, ist die Klassifizierung von Informationen und das Vorhersehen bestimmter Ereignisse basierend auf dem Erlernten.

Natural Language Understanding ist ein Teilbereich von Natural Language Processing. Über NLU lässt sich über syntaktische und semantische Analysen die Bedeutung eines Satzes und dadurch auch die Bedeutung von ganzen Textpassagen identifizieren. Natural Language Generation ist der zweite Teilbereich von Natural Language Processing, der zur Aufgabe hat, Computer zu einem Dialog zu befähigen, also selbstständig

INFO

Was ist LaMDA?

LaMDA steht für „Language Model for Dialogue Applications“ und ist ein auf einer Transformer-Architektur basierendes KI-Modell. Das Sprachmodell wurde laut Google mit Dialogen trainiert, um natürliche Sprache besser zu verstehen und Antworten auf komplexere Fragen auszugeben.

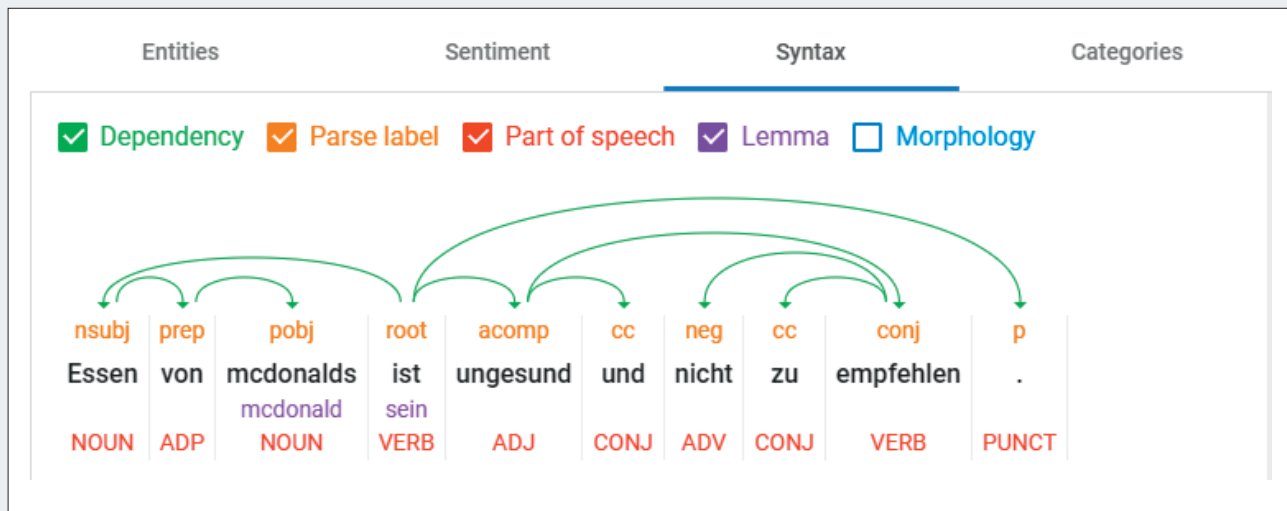


Abb.7: Beispiel einer Syntax-Analyse aus der NLP-API-Demo (Quelle: Google)

zu schreiben und zu sprechen. Zudem können über NLG Zusammenfassungen eines Texts automatisch generiert werden. Hierfür nutzt Google das Sprachmodell LaMDA welches im Zusammenhang mit MUM vorgestellt wurde.

Über das LaMDA-Sprachmodell lassen sich offene Dialoge gestalten, was für folgende Zwecke eingesetzt werden kann:

- » Führung eines Nutzers während einer Google-Suche zu allumfänglichen Informationen zu einem Thema
- » Sprachgesteuerte Dialoge via Google Assistant und Google Home
- » Echtzeit-Übersetzung
- » Chatbots

LaMDA kann über unstrukturierte Daten aus multimedialen frei zugänglichen Internet-Datenquellen trainiert werden. Im Vergleich zu anderen Sprachmodellen wie z. B. GPT-3 sind die Datenquellen nicht manuell geprüft, was die Nutzung für sensible Themen noch schwieriger macht.

Um Machine Learning flächendeckend in voller Breite einzusetzen, bedarf es extrem leistungsstarker Prozessoren und es ist nicht verwunderlich, dass Google seit 2006 bereits im Bereich des Quantencomputings forscht. Ein Quantencomputer ist millionenfach leistungstärker als bisher eingesetzte Rechner-Systeme. Alphabet-Chef Sundar Pichai verkündete in der Eröffnungsrede zur I/O-Konferenz 2021:

„Quantencomputing stellt einen fundamentalen Wandel dar, weil es die Eigenschaften der Quantenmechanik nutzt und uns die beste Chance gibt, die natürliche Welt zu verstehen.“

Funktionalität und Einsatzgebiete von Natural Language Processing in modernen Suchmaschinen

Beim Natural Language Processing als Teilbereich von Machine Learning geht es darum, menschliche Sprache in geschriebener und gesprochener Form besser zu verstehen und unstrukturierte Informationen in maschinenlesbare strukturierte Daten umzuwandeln. Teilaufgaben von NLP sind Übersetzung von Sprachen und die Beantwortung von Fragen. Hier wird schnell klar, wie wichtig diese Technologie für moderne Suchmaschinen wie Google ist.

Generell kann man die Funktionsweise von NLP grob in die folgenden Prozessschritte aufgliedern:

- » Datenbereitstellung
- » Datenvorbereitung
- » Textanalyse
- » Textanreicherung

Die Kernkomponenten von NLP sind Tokenization, zu Deutsch Tokenisierung, Kennzeichnung von Wörtern nach Wortarten (Part of Speech Tagging), Lemmatisierung, Wort-Abhängigkeiten (Dependency Parsing), Parse Labeling, Extraktion benannter Entitäten (Named-

entity Recognition), Saliency-Scoring, Sentiment-Analysen, Kategorisierung, Text-Klassifizierung, Extrahierung von Content-Typen und Identifikation einer impliziten Bedeutung aufgrund der Struktur.

- » **Tokenisierung:** Tokenisierung ist der Vorgang, bei dem ein Satz in verschiedene Begriffe unterteilt wird.
- » **Kennzeichnung von Wörtern nach Wortarten:** Wortartenkennzeichnung klassifiziert Wörter nach Wortarten wie z. B. Subjekt, Objekt, Prädikat, Adjektiv ...
- » **Wortabhängigkeiten:** Wortabhängigkeiten schaffen Beziehungen zwischen den Wörtern basierend auf Grammatikregeln. Dieser Prozess bildet auch „Sprünge“ zwischen Wörtern ab.
- » **Lemmatisierung:** Die Lemmatisierung bestimmt, ob ein Wort verschiedene Formen hat, und normalisiert Abwandlungen zur Grundform. Zum Beispiel ist die Grundform von Tiere Tier oder von verspielt Spiel.
- » **Parsing Labels:** Die Kennzeichnung klassifiziert die Abhängigkeit oder die Art der Beziehung zwischen zwei Wörtern, die über eine Abhängigkeit verbunden sind.
- » **Analyse und Extraktion benannter Entitäten:** Dieser Aspekt sollte uns aus den vorangegangenen Abschnitten bekannt sein. Damit wird versucht, Wörter mit einer „bekannten“

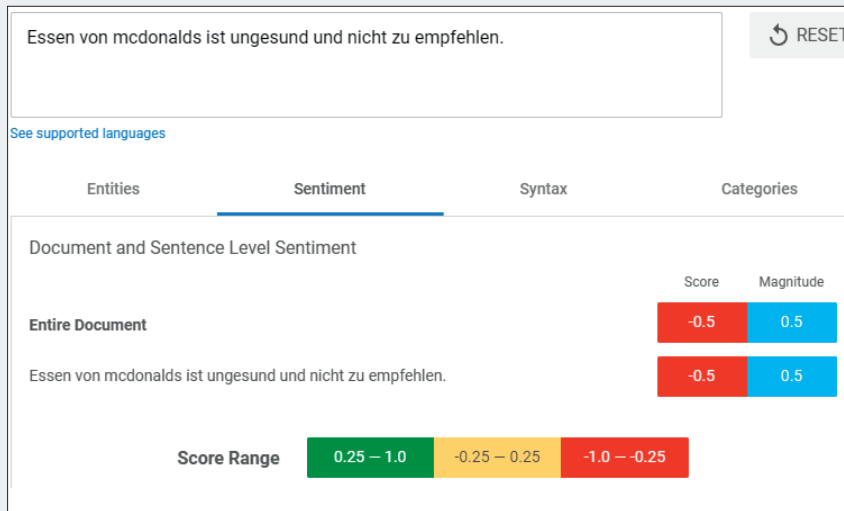


Abb.8: Beispiel einer Syntax-Analyse aus der NLP-API-Demo (Quelle: Google)

Bedeutung zu identifizieren und Klassen von Entitätstypen zuzuordnen. Im Allgemeinen sind benannte Entitäten Menschen, Orte und Dinge (Substantive). Entitäten können auch Produktnamen enthalten. Dies sind im Allgemeinen die Wörter, die ein Knowledge-Panel auslösen. Aber auch Begriffe, die kein eigenes Knowledge-Panel auslösen, können Entitäten sein.

Über Natural Language Processing können aus Suchanfragen, Sätzen und Textabschnitten Entitäten identifiziert sowie die einzelnen Bestandteile in sogenannte Tokens zerlegt und in Beziehung zueinander gesetzt werden. Auch ein grammatikalisches Verständnis kann durch NLP algorithmisch entwickelt werden.

Mit der Einführung von Natural Language Processing ist Google auch in der Lage, mehr als nur die Substantive für die Interpretation von Suchanfragen, Texten und Sprache zu deuten. So sind seit BERT auch Verben, Adverbien, Adjektive für die Ermittlung des Kontexts wichtig. Durch die Identifikation der Beziehungen zwischen den Tokens lassen sich Bezüge herstellen und so können auch Personalpronomen gedeutet werden.

Ein Beispiel:

„Olaf Kopp ist Head of SEO bei Aufgesang. Er beschäftigt sich seit 2005 mit Online-Marketing.“

In der Zeit vor Natural Language Processing konnte Google mit dem Personalpronomen „er“ nichts anfangen, da kein Bezug zur Entität „Olaf Kopp“ hergestellt werden konnte. Für Indexierung und Ranking wurden nur die Begriffe Olaf Kopp, Head of SEO, Aufgesang, 2005 und Online-Marketing berücksichtigt.

Über Natural Language Processing lassen sich nicht nur Entitäten in Suchanfragen und Inhalten identifizieren, sondern auch die Beziehung dieser zueinander.

Dabei wird sowohl die grammatikalische Satzstruktur als auch Bezüge innerhalb von ganzen Absätzen und Texten berücksichtigt. Nomen bzw. Subjekt und Objekt in einem Satz können als poten-

zielle Entitäten identifiziert werden. Über Verben lassen sich Beziehungen zwischen Entitäten herstellen, durch Adjektive ein Sentiment (Stimmung) um eine Entität ermitteln.

In Abbildung 8 sehen Sie ein Beispiel für die Sentimentanalyse über die NLP-API von Google.

Hier analysiert Google die Stimmung bzw. Meinung in dem Satz. Der Sentiment-Score gibt an, ob die Stimmung negativ oder positiv ist. Interessant ist, dass man diese Analyse auch auf Entitäten beziehen kann. Das ist von daher spannend, dass Google Kundenstimmen und Berichterstattungen rund um eine Marke, Produkt, Unternehmen etc. ermitteln kann.

Über Natural Language Processing lassen sich auch konkrete W-Fragen besser beantworten, was für die Bedienung von Voice Search eine deutliche Weiterentwicklung darstellt.

Auch für das von Google 2021 eingeführte Passage-Ranking spielt Natural Language Processing eine zentrale Rolle.

Seit der Einführung von BERT im Jahr 2018 nutzt Google diese Technologie in der Google-Suche. Das 2021 eingeführte Passage-Ranking basiert auf Natural Language Processing, da Google hier einzelne Textpassagen durch die neuen

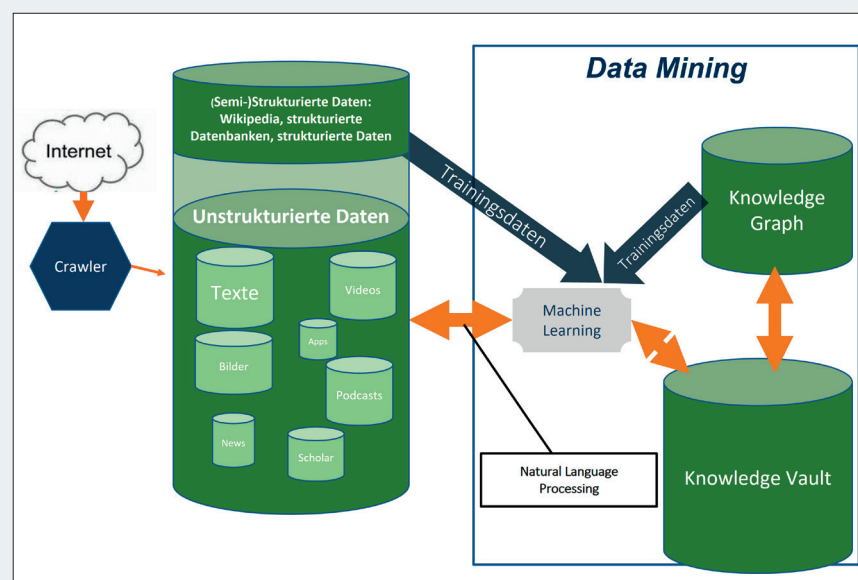


Abb.9: Übersicht Data Mining für den Google Knowledge Graphen (Quelle: Olaf Kopp)

Möglichkeiten besser interpretieren kann.

Googles neueste Innovation namens MUM stellt die nächste Ära für die Google-Suche dar, die maßgeblich durch Natural Language Processing und Natural Language Understanding getrieben wird.

Natural Language Processing zum Aufbau einer semantischen Wissensdatenbank

Während Google bis dato von manuell gepflegten strukturierten und semistrukturierten Informationen bzw. Datenbanken abhängig war, ist es seit BERT möglich, Entitäten und deren Beziehungen aus unstrukturierten Datenquellen zu extrahieren und in einem Graphen-Index zu speichern. Ein Quantensprung in Sachen Data-Mining.

Dafür kann Google die bereits verifizierten Daten aus (semi-)strukturierten Datenbanken wie dem Knowledge Graphen, Wikipedia ... als Trainingsdaten nutzen, um zu lernen, unstrukturierte Informationen zu bestehenden Modellen bzw. Klassen zuzuordnen und neue Muster zu erkennen. Hier spielt Natural Language Processing in Form von BERT und MUM die entscheidende Rolle.

Bereits 2013 hat Google erkannt, dass der Aufbau einer semantischen Datenbank wie des Knowledge Graphen ausschließlich basierend auf strukturierten Daten zu langsam und nicht skalierbar möglich ist, da die große Masse an sogenannten Long-Tail-Entitäten nicht in (semi-)strukturierten Datenbanken erfasst ist. Zur Erfassung dieser Long-Tail-Entitäten bzw. des vollständigen Wissens der Welt stellte Google 2013 den Knowledge Vault vor, der aber seitdem nicht mehr groß Erwähnung fand. Der Ansatz, über eine Technologie das komplette im Internet verfügbare Wissen für eine semantische Datenbank zu nutzen, wird durch Natural Language Processing Realität. Es ist davon auszugehen, dass es neben dem Knowledge Graphen

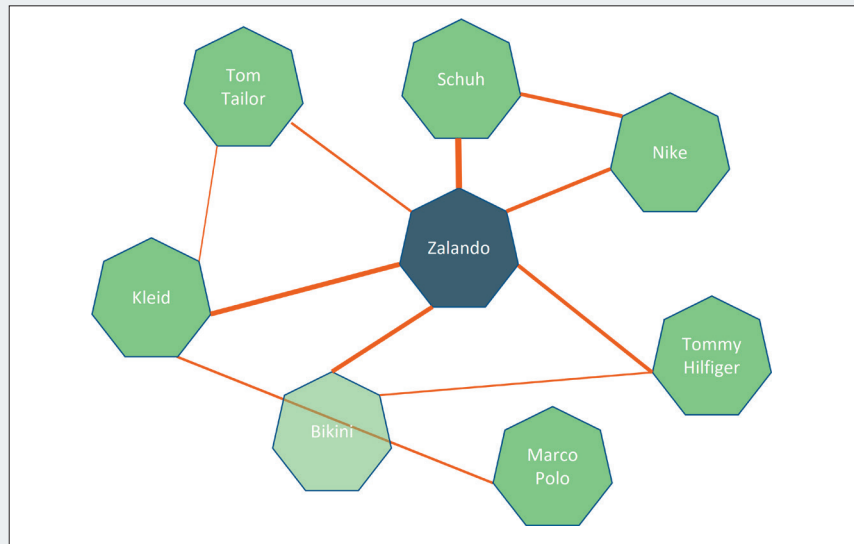


Abb.10: Gewichtete Entitäten-Beziehungen am Beispiel Zalando (Quelle: Olaf Kopp)

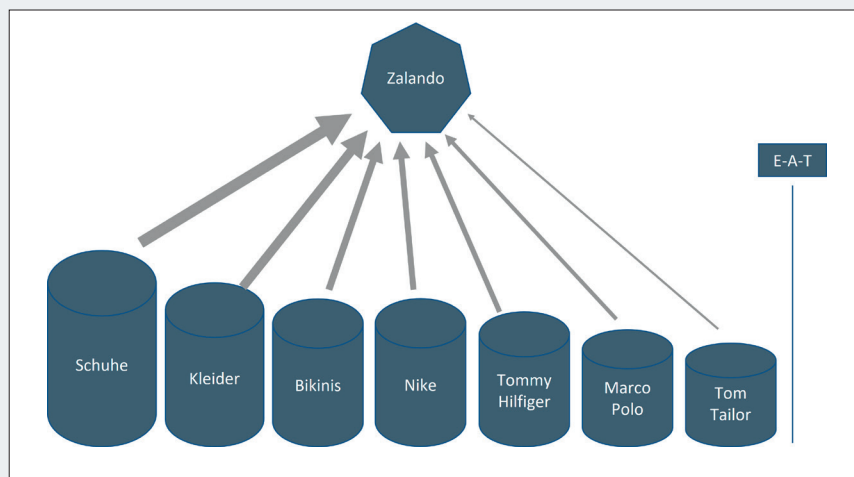


Abb.11: E-A-T-Bewertung nach Themen (Quelle: Olaf Kopp)

eine Art Zwischenspeicher gibt, in dem Google das über Natural Language Processing generierte Wissen erfasst und strukturiert bzw. organisiert. Sobald ein Validitäts-Schwellenwert erreicht ist, werden die Entitäten und Informationen in den Knowledge Graphen überführt. Dieser Zwischenspeicher könnte der Knowledge Vault sein.

Dadurch lässt sich auch erklären, warum der Knowledge Graph gerade in den letzten Jahren sehr schnell gewachsen ist.

"By mid-2016, Google reported that it held 70 billion facts[4] and answered 'roughly one-third' of the 100 billion monthly searches they handled. By May 2020, this had grown to 500 billion facts on 5 billion entities." (Quelle: Wikipedia, <http://einfach.st/wp73>)

Das Zusammenspiel aus Entitäten-basierter Indexierung, Natural Language Processing und E-A-T

E-A-T ist Googles eigenes Konzept, um Publisher und Autoren zu identifizieren, die eine Marke sind. Dadurch möchte Google die Qualität der Suchergebnisse und damit die Nutzererfahrung mit den SERPs verbessern.

Dass E-A-T einen wichtigen Einfluss auf die Rankings hat, ist spätestens seit den Core-Updates der letzten Jahre jedem SEO klar. John Müller sprach auf der Commerce Week 2021 von einem relativ neuen Ranking-Einfluss. E-A-T wird aber bereits seit 2014 in den Quality Rater Guidelines prominent erwähnt.

Es ist sehr wahrscheinlich, dass die manuellen Bewertungen der Such-Eva-

luatoren bzw. Quality Rater seit 2014 als manuell verifizierte von den Machine-Learning-Algorithmen als Lernmaterial für die Erstellung von Modellen zur algorithmischen Bestimmung von E-A-T genutzt werden. Diese Modelle basieren auf erkannten On- und Offpage-Mustern, die für E-A-T signifikant sind.

Seit der Einführung von Hummingbird versucht Google, Entitäten zu identifizieren, zu extrahieren und in Beziehung zu setzen. Die Beziehungen zwischen Unternehmen, Personen und Themen sind für Google wichtig, da sie über diesen Weg kontextuelle Zusammenhänge, die Qualität bzw. Stärke der Beziehung und darüber Autorität und Expertise algorithmisch ermitteln können.

So steht die Entität Zalando in enger Beziehung zu weiteren Entitäten wie Fashion-Brands (z. B. Tom Tailor, Nike, Tommy Hilfiger, Marco Polo ...) und Produktgruppen (Schuhe, Kleider, Bikini ...).

Diese Beziehungen können unterschiedlich stark ausgeprägt sein. Über die Stärke bzw. Ausprägung dieser Beziehungen kann Google eine Bewertung für Expertise und vor allem Autorität erfassen und in ein E-A-T-Konzept einfließen lassen.

Durch die Nutzung der neuen „About this result“-Boxen hat Google eine weitere mögliche Datenquelle, um einen Autor bzw. Publisher hinsichtlich Trusts zu bewerten. Wenn die Nutzer den in den Boxen dargestellten Informationen vertrauen und das Ergebnis aufrufen, ist es ein Vertrauensbeweis in die Quelle. Google könnte diese Nutzesignale dann ebenfalls für E-A-T einbeziehen.

Durch Technologien wie Natural Language Processing lassen sich Entitäten aus unstrukturierten Daten identifizieren, extrahieren und über die Kookkurrenzen mit anderen Entitäten in Beziehung setzen. Je häufiger die Entitäten in Suchanfragen, Texten, Videos und Bildern in Kookkurrenz mit anderen themenrelevanten Entitäten auftreten,

desto mehr Autorität hat die Entität in der jeweiligen thematischen Ontologie.

Was bedeutet MUM und wie funktioniert es?

Im Mai 2021 stellte Google auf der I/O 2021 das Multitask Unified Model, kurz MUM, als 1000-mal leistungsstärkere Weiterentwicklung von BERT vor. Beide Technologien beruhen auf Natural Language Processing. Bei MUM geht es aber um mehr als nur Natural Language Processing. In MUM werden mehrere Technologien zusammengefasst, um die Google-Suche noch semantischer und kontextbasierter zu gestalten mit dem Ziel, die Nutzererfahrung zu verbessern. Durch MUM möchte Google komplexe Suchanfragen beantworten, für die ein normales SERP-Snippet nicht ausreicht.

Die Aufgaben, die von MUM übernommen werden sollen, wurden wie folgt präsentiert:

- » MUM soll ein tiefes Verständnis des weltweiten Wissens ermöglichen.
- » MUM soll durch bis zu 75 Sprachen simultan trainiert werden, um diese zu verstehen. Bei den bisherigen Methoden wurde für jede Sprache ein eigenes Sprachmodell trainiert.
- » MUM soll Informationen neben Text auch von Bildern, Audios und Videos verstehen können.

Durch diese Aussagen seitens Google lässt sich Folgendes ableiten:

- » Google nutzt MUM, um die semantischen Datenbanken wie z. B. dem Knowledge Graphen noch weiter auszubauen und dem Ziel einer voll-

ständigen Wissensdatenbank näher-zukommen.

- » Google möchte mit MUM die internationale Suche noch besser machen, um alle Indizes und Such-Instanzen weltweit auf ein ähnliches qualitatives Niveau zu bringen, damit die Nutzererfahrung in jedem Land und jeder Sprache der Erde auf gleich hohem Niveau zu garantieren und zudem effizienter zu arbeiten, was die eigenen Ressourcen angeht.
- » Die für SEOs wohl spannendste Entwicklung ist, dass MUM auf alle Medienformate zugreifen will, um Informationen zu sammeln (Data-Mining), diese zu verstehen und in den Suchergebnissen so aufzubereiten, dass die Suchintention optimal bedient und die Nutzererfahrung mit der Google-Suche sehr hoch ist.
- » Laut Google stehen neben den Billionen von textbasierten Inhalten nun auch alle Informationen aus allen in den Google-Systemen befindlichen Medienformaten zur Verfügung, um den (Entitäten-)Index zu vergrößern. Dadurch lassen sich noch mehr Informationen zu erfassten Entitäten generieren und zu einer Suchanfrage Informationen aus Text, Videos, Audios und Bildern in den SERPs noch besser zusammengefasst darstellen.

Google veranschaulichte das an dem Beispiel der Suchanfrage:

I've hiked Mt. Adams and now want to hike Mt. Fuji next fall, what should I do differently to prepare.

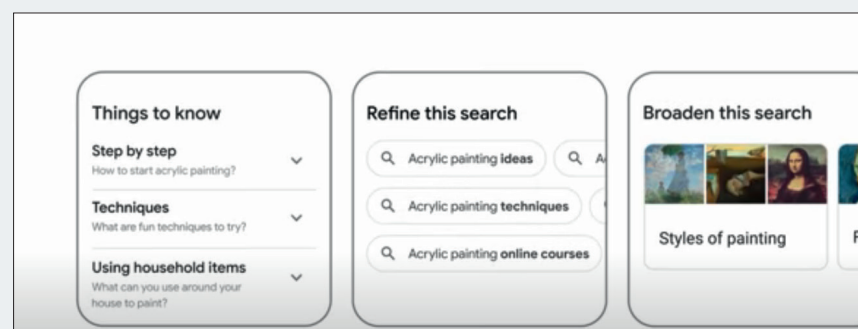


Abb.12: Abbildung aus der Präsentation bei der Google Search On '21 (Quelle: Google)

Diese sehr komplexe Suchanfrage ließe sich durch ein Keyword-Text-Matching nicht beantworten, da über diesen Weg kein Kontext über die Wortbeziehungen möglich wäre. In einer semantischen Suchmaschine, die auf Natural Language Processing setzt, lassen sich die Entitäten Mt. Adams, Mt. Fuji und fall identifizieren. Über die grammatikalische Satzstruktur und die dort genutzten Verben wie hike, prepare kann Google den Kontext in Verbindung mit den Entitäten herstellen und die Bedeutung bzw. die Suchintention wird klarer. Früher konnte Google maximal ermitteln, dass die Begriffe genannt werden, und prüfen, ob es Inhalte im Index gibt, die diese Begriffe enthalten, ohne die Suchanfrage wirklich zu verstehen.

Über Natural Language Processing bzw. MUM ist Google in der Lage zu verstehen, dass der Nutzer Wanderausrüstung für eine Bergwanderung im Herbst sucht. Zudem kann Google daraus ableiten, dass es im Herbst oft regnet und dementsprechende Ausrüstung gefragt ist. Daraufhin kann Google Listen für Wanderausrüstung bzw. Produkte, Informationen zum Wetter in der Zielregion im Vergleich zur Region Mt. Adams angeben und verschiedene Medienformate in den SERPs ausgeben. Und das alles simultan über alle Sprachen.

In der Verbindung mit der verbesserten Bilderkennung von Google in Kombination mit dem besseren Verständnis des Kontexts kann Google konkrete Fragen aus Bild und Text beantworten. So ist es möglich, durch MUM zu erkennen, ob ein auf einem Bild abgebildetes Paar Schuhe geeignet ist, den Mt. Fuji zu besteigen.

In der im September 2021 stattgefundenen Search On 21 gab Google bekannt, dass MUM demnächst für die Google-Suche aktiv sein wird, und gab weitere Erklärungen und Beispiele.

Ein beeindruckendes Beispiel war eine Bild-basierte Suche nach passenden weiteren Kleidungsstücken zu einem Hemd mit extravagantem Muster, das

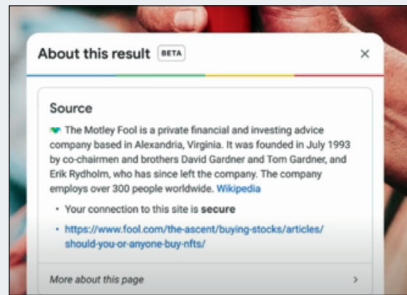


Abb. 13: Abbildung aus der Präsentation bei der Google Search On '21 (Quelle: Google)

sich nur schwer in Worte fassen lässt. Es ist ein weiteres Beispiel, wie Google aus Text und Bild in Kombination einen Kontext und darüber ein Verständnis erlangt, was der Nutzer wirklich sucht. Über die Suchanfrage zu dem Bild des Hemds, via Google Lens aufgenommen, „socken passend zum Muster“, werden in Google-Shopping zu dem Muster passende Socken ausgegeben.

Ein weiteres Beispiel war ein Defekt an einem Fahrrad, der sich nur schwer in einer textbasierten Suche formulieren lässt. Ein Foto des defekten Teils in der Kombination mit einer Frage „wie kann ich das reparieren“ ergibt die Lösung. Daraus lässt sich dann auch wieder die Brücke zu benötigten Werkzeugen oder Teilen schlagen.

Google Lens soll für solche Fälle auch zukünftig auf Desktop PC nutzbar sein, um z. B. Teilbereiche aus Bildern via Screenshot aufzunehmen und darüber z. B. bei Google-Shopping nach einer Lampe, einem Hemd oder anderen Produkten zu suchen.

Auch bei Google-Shopping sind die gelisteten Produkte in einem Shopping-Graphen organisiert. Über die Graphen-Struktur der Datenbank lassen sich Beziehungen zwischen den Produkten nach Attributen, Anbietern ... abbilden. So lassen sich z. B. Anbieter in der Nähe anzeigen, die ein bestimmtes Produkt auf Lager haben.

Neben der Fähigkeit, Suchanfragen richtig zu interpretieren, will Google mit MUM Informationen in den SERPs so aufbereiten, dass tiefere Einblicke möglich sind. Google möchte praktisches Wissen vermitteln, was sonst nur über

erfahrenere Menschen möglich wäre zu erlangen. Dazu stellt Google zukünftig weitere Boxen in den SERPs zur Verfügung.

Über die Box „things to know“ möchte Google dem Nutzer ermöglichen zu spezifizieren, welcher Teilaspekt eines Themas ihn genauer interessiert. Über die Box „refine this search“ bekommt der Nutzer Vorschläge, um die Suche zu verfeinern, was für SEOs ein schöner Fundus sein kann, um die eigenen Inhalte im Long Tail zu optimieren. Über die Box „broaden this search“ lässt sich der Blickwinkel auf ein Thema wieder erweitern. Durch die Boxen wird die Suche für den Nutzer intuitiver.

Zudem soll MUM die SERPs noch bildreicher gestalten, insbesondere bei den Suchanfragen, die eine visuelle Unterstützung benötigen. In der mobilen Ansicht der SERPs sind Bilder in Kombination mit den klassischen Blue Links schon länger zu sehen. Dies soll im Zuge der MUM-Einführung weiter ausgebaut werden.

Auch bei YouTube wird es Erweiterungen zu den Videos geben, um tiefer in das Thema einzusteigen.

Eine weitere große Herausforderung von Google ist neben der Vollständigkeit des Wissens der Welt die Validität von Informationen. Mit E-A-T betreibt Google hierzu bereits seit Jahren eine Qualitäts-offensive, die durch weitere Features in naher Zukunft Unterstützung bekommen soll.

So gibt es im direkten Bezug zum jeweiligen Suchergebnis die Möglichkeit, über eine „About this result“-Box Informationen zur Quelle zu bekommen.

Die Informationen beinhalten sowohl eine Beschreibung des Publishers z. B. aus Wikipedia als auch Informationen dazu, ob die Verbindung zur Website sicher ist. Stichwort: HTTPS. Zusätzlich kann man sich darüber informieren, was der Publisher selbst über sich schreibt, was andere über ihn oder das Thema schreiben.

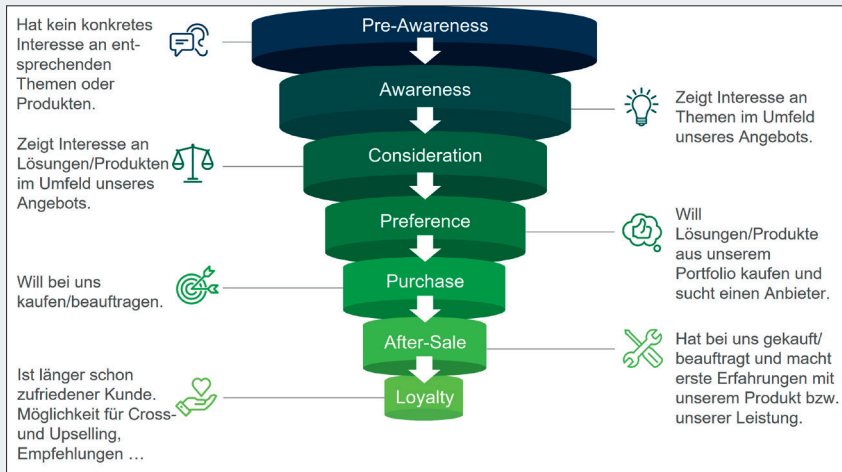


Abb. 14: 7-stufiges Customer-Journey-Modell von Aufgesang (Quelle: Olaf Kopp)

Inwiefern hier MUM zum Einsatz kommt, ist nicht ganz klar. Allerdings ist ziemlich sicher, dass Autorität und Trust der Entität bzw. Quelle auch eine besondere Rolle im Ranking spielen.

MUM als nächster großer Paradigmenwechsel für mehr Performance und Nutzerorientierung

Die Einführung von MUM stellt nach Hummingbird, RankBrain und BERT den nächsten großen Paradigmen-Wechsel für die Google-Suche dar. Während die bisherigen auf Machine Learning basierenden Innovationen auf zahlreiche trainierte Modelle für verschiedene Aufgaben zurückgriffen, ist das Ziel beim MUM, auf nur ein einheitliches Modell für alle relevanten Aufgaben für Indexing, Information-Retrieval und Ranking zugreifen zu müssen.

Das ist aus Performance-Gründen ein deutlicher Fortschritt. Die Google-Suche greift inzwischen auf sehr komplexe Technologien zurück. Der Einsatz von Machine Learning und Natural Language Processing ist nur möglich, wenn die vorhandenen Ressourcen effizient genutzt werden können. Die Reduzierung der parallel laufenden Prozesse ist die Voraussetzung dafür. Die Fokussierung auf die Abfrage von Informationen aus nur einem Datenmodell je Suchanfrage ist ein Effizienz- bzw. Performance-Boost.

In einem Google-Research-Paper (<http://einfach.st/ok11>) mit dem Titel

„Multitask Mixture of Sequential Experts for User Activity Streams“ aus dem Jahr 2020 wird eine Technologie namens MoSE beschrieben, die MUM in vielen Punkten ähnelt.

MoSE kann basierend auf Nutzerdaten sehr performant Klicks und Suchhistorie in einem Datenmodell zusammenfassen. Es funktioniert wie eine Art Marktforschung, beginnend beim Suchmaschinen-Nutzer und nicht beginnend bei den indexierten Informationen wie klassische Suchmaschinen. Die Nutzerintention steht damit im Mittelpunkt und basierend auf dem Modell kann Google vorhersehen, welche Fragen und dementsprechende Antworten ein Nutzer während seiner Recherche benötigen wird.

Daraus lassen sich dann in den SERPs alle notwendigen Informationen zusammenstellen, um den Nutzer lückenlos durch die Customer Journey zu begleiten.

Neues Potenzial für Google-Shopping entlang der Customer Journey

Spannend ist MUM für Google, da sie hier neue Ansatzpunkte für Google-Shopping bekommen. Bei den produktbasierten Suchanfragen hat Google einiges an Boden gegenüber den großen E-Commerce-Marktplätzen wie Amazon, aber auch kleineren Spezial-Anbietern wie z. B. Otto oder Zalando verloren. Nutzer, die sich in der Preference-Phase der Customer Journey befinden, suchen

häufig bei z. B. Amazon direkt nach dem Produkt. Das ist aus wirtschaftlicher Sicht schwierig für Google, da diese Nutzer bzw. kommerziellen Suchanfragen die meisten Klicks auf Anzeigen verursachen. Hier scheint Google nur schwer Nutzer wieder zurückzubekommen, obwohl Google die Shopping-Suche inzwischen wieder für kostenlose Produkt-Listings geöffnet hat.

Die meisten Nutzer:innen benutzen Google für informationsorientierte Suchen in der Awareness-Phase. Aktuell verliert Google aber spätestens in der Preference-Phase an die Wettbewerber.

An den von Google vorgestellten Beispielen wird klar, dass Google Nutzer in frühen Phasen der Customer Journey (Awareness und Consideration) mit wertvollen Informationen versorgen möchte. Durch die neue Gestaltung der SERPs und die Shopping-Suche möchte Google den Nutzer inspirieren, eine Übersicht verschaffen und beim Kauf unterstützen. Also vollständig durch die Customer Journey begleiten, sodass Amazon und Co. auch bei den Produktsuchen leer ausgehen.

Das zeigt, dass Google den direkten Kampf um die Produktsuchen in der Preference-Phase aufgegeben hat und sich auf die eigentlichen Stärken konzentriert – die Organisation und Aufbereitung des Wissens der Welt in eine nutzerfreundliche Form. Hier können die großen E-Commerce-Plattformen nicht mithalten.

Die Zukunft der Google-Suche und was SEOs daraus lernen können

MUM ist das nächste Puzzle-Stück für Google auf dem Weg zu einer reinen semantischen Suchmaschine, die immer besser den Kontext von Suchanfragen sowie Inhalten und damit die Suchintention und die Relevanz von Inhalten und Inhaltspassagen versteht.

Die Entwicklung eines einsetzbaren Quantencomputers ist noch Zukunfts-

musik und so muss sich Google mit effizienteren Technologien wie MUM beschäftigen, um die aktuell fehlende Rechenleistung für Big-Scale Machine Learning zu nutzen. Über diesen Weg kann Google schneller die eigenen Suchsysteme weiterentwickeln, ohne auf die fehlende Performance auf der Hardware-Seite Rücksicht nehmen zu müssen. Man könnte sagen, dass die Software-Entwicklung die Hardware-Entwicklung gerade überholt.

Ein Durchbruch für kommerziell einsetzbare Quantencomputer wird für das Jahr 2029 prognostiziert. Wir können davon ausgehen, dass bis dahin die Google-Suche eine vollständig semantische Suchmaschine ist. Ein Keyword-Text-Matching in der Google-Suche wird dann endgültig der Geschichte angehören.

Zusammengefasst bleibt festzuhalten:

- » Die Google-Suche wird noch nutzerzentrierter.
- » Noch semantischer.
- » Noch nutzerfreundlicher, da immer weniger Klicks notwendig sind, um an Antworten zu kommen.
- » Noch performanter, was in erster Linie den Google-Ressourcen zugutekommt.
- » Die bisher in einzelne Sprachen und Länder aufgeteilte Google-Indizes werden immer mehr verschmelzen.
- » Immer mehr Entitäten können über MUM erkannt und mit Attributen, Beziehungen zu anderen Entitäten versehen und in thematischen Kontext gebracht werden.
- » Entitäts-Typen wie Publishern oder Autoren können immer mehr Inhaltstypen bzw. Medienformate zugeordnet werden, was die thematische Klassifizierung und Bewertung hinsichtlich E-A-T vorantreibt.
- » Die Auszeichnung von Inhalten mit strukturierten Daten wird zukünftig immer unwichtiger, da Google über Natural Language Processing immer weniger strukturierte Daten benötigt.

SEO muss sich verändern!

An dieser Stelle muss man sich auch die Frage stellen, welche Rolle Google zukünftig noch als Traffic-Lieferant spielt und inwiefern SEOs noch einen direkten Einfluss auf die Rankings haben.

Die Einführung von BERT und MUM bringt ähnlich einschneidende Veränderungen in die Branche wie damals Panda und Pinguin. Natural Language Processing treibt die mit Hummingbird und Knowledge Graph fundierte semantische Suche deutlich schneller voran. SEOs müssen mehr in Entitäten und Themen in Bezug auf E-A-T als in Keywords denken.

Für technisches SEO bleibt weiterhin die Aufgabe, das Crawling und die Indexierung der suchrelevanten Inhalte sicherzustellen. Aber Technik macht nicht relevant und schafft auch keine Autorität oder Expertise. Hinsichtlich Trusts (HTTPS) und UX (Page Experience) hat die Technik einige kleine Hebel, um in das Ranking einzugreifen. Diese Hebel sind aber keine Top-Position-Garanten. Technische Aufgabenbereiche wie die Auszeichnung mit strukturierten Daten werden immer mehr obsolet werden, da Google über Natural Language Processing immer weniger strukturierte Informationen für das Verständnis benötigt.

Content und Links bleiben weiterhin die wichtigsten Einflussfaktoren. Zu den Links gesellen sich weitere wichtige Faktoren, die die Autorität untermauern. Kookkurrenzen in Suchanfragen und Inhalten (Texte, Videos, Audios, Bilder) sind wichtige Signale für Vertrauen und Autorität. Google hat durch MUM Zugriff auf deutlich mehr Datenquellen und Informationen. Zudem kann Google über sprachunabhängige Data-Mining alle Informationen der Welt zu Entitäten und Themen sammeln und zusammenführen. Die bisherigen Daten-Silos werden aufgebrochen.

Dadurch kann Google noch besser Fragen beantworten und wirklich tiefes Wissen vermitteln.

Content-Verantwortliche sollten sich

weniger mit der Häufigkeit der Nutzung von Keywords in ihren Inhalten beschäftigen, sondern sich überlegen, aus welchen Perspektiven ein Thema zu behandeln ist. Hier ist die gute alte TF-IDF-Analyse ein immer noch probates Mittel, um wichtige Begriffe zu identifizieren, die das Keyword-Korpus eines Themas beschreiben. Die Recherche von W-Fragen und MAFO ergänzt das Instrumentarium, um herauszufinden, welche Bedürfnisse und Fragen sich deine Nutzer entlang der Customer Journey stellen. Auch hier steht also die Nutzerzentrierung im Fokus. Beschäftige dich mit deinen Zielgruppen!

Content bietet die Antworten auf Fragen. Doch nur Content zu produzieren wird zukünftig nicht mehr ausreichen. Google möchte den Nutzer durch die komplette Customer Journey mit Antworten begleiten, um an den wertvollen produktbezogenen kommerziellen Traffic zu gelangen und diesen in die eigene Shopping-Welt zu überführen. Damit wollen sie Marktanteile zurückgewinnen.

Für Content-Verantwortliche wird es auch aus SEO-Sicht immer wichtiger, Content-Marketing entlang der Customer Journey bereitzustellen, um in der Recherche dem Nutzer so viele Content-Touchpoints wie möglich anzubieten.

Nutzer durchlaufen je nach Wissensstand einen Recherche-Prozess über einen kürzeren oder längeren Zeitraum. Sie sind auf der Suche nach Lösungen. Mit wachsendem Wissen zu einem Thema stellen sich unterschiedliche Herausforderungen und Fragen, die Antworten bedürfen.

Jemand, der frisch in das Thema Suchmaschinenoptimierung einsteigt, wird sich eher die Frage stellen: „Was ist SEO?“. Als Nächstes fragt man sich: „Wie funktioniert SEO?“, um dann zu bemerken, dass das Thema ganz schön komplex ist und man sich eher die Frage stellen sollte: „Wer bietet SEO-Dienstleistungen an?“ Auf dieser Reise sollten Unternehmen die Antworten liefern.

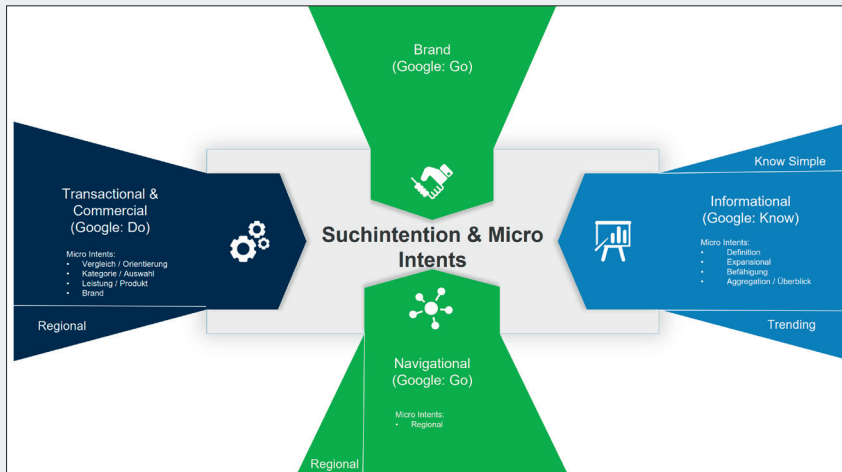


Abb. 15: Übersicht Suchintentionen und Micro-Intents (Quelle: Olaf Kopp)

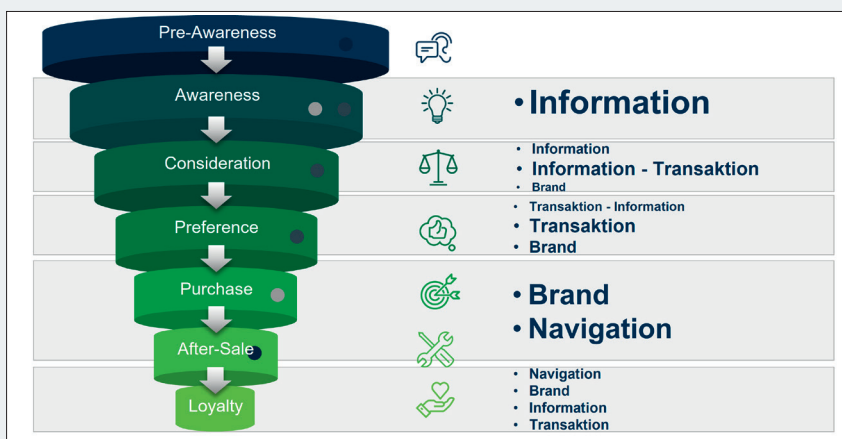


Abb. 16: Suchintentionen in der Customer Journey (Quelle: Olaf Kopp)

Content muss nutzerzentriert gedacht werden und die Bedürfnisse sowie Fragen entlang der Customer Journey antizipieren, genauso wie Google es mit MUM tut. Hier helfen detaillierte SERP-Analysen, um die aktuellen und zukünftigen Suchintentionen zu antizipieren. Dabei reicht eine grobe Klassifizierung nach den klassischen Suchintentionen transactional und informational nicht aus, da daraus keine klaren Handlungsempfehlungen für den Content abgeleitet werden können. Deswegen habe ich die klassischen Suchintentionen in weitere Micro-Intents unterteilt, die mir bei SERP-Analysen immer wieder auffallen.

Informational Micro-Intents

» **Unterhaltung:** Menschen möchten sich die Zeit vertreiben und suchen nach Unterhaltung. Die Unterhaltung kann durch kurzen Snackable Content

wie z. B. Memes oder kurze Videoclips befriedigt werden, die meistens direkt auf der jeweiligen Social-Media-Plattform konsumiert (Social Content) und geteilt werden.

- » **Definition:** Nutzer, die nach grundsätzlichen Antworten suchen, z. B. nach der Frage: Was ist ...? Was bedeutet ...? Wie wichtig ...?, wollen erst einmal verstehen, was etwas bedeutet, um für sich den Kontext eines Themas zu erschließen und/oder zu entscheiden, ob es Sinn macht, sich tiefer mit einem Thema zu beschäftigen. Es sind Einsteiger. Wikipedia-Beiträge sind ein typisches Beispiel für Content, der zu diesem Micro-Intent passt. Diese Suchintention versucht Google teilweise direkt in den SERPs zu beantworten (z. B. Featured Snippets).
- » **Expansional:** Nutzer, die in ihrer Online-Recherche tiefer in ein Thema

einsteigen wollen, benötigen sehr ausführlichen Content, der möglichst viele Perspektiven beleuchtet. Der Content sollte allumfassend ein Thema beschreiben und viele Fragen beantworten. Die sogenannten holistischen Landingpages oder auch Pillar Pages sind hier ein sinnvoller Ansatz, um diese Suchintention zu bedienen. Hier spielen die Tiefe und der Umfang des Contents eine besondere Rolle, aber auch neue, bisher nicht häufig publizierte Perspektiven.

- » **Befähigung:** Nutzer, die sich befähigen wollen, etwas selbst zu tun, benötigen konkrete Anleitungen. Content, der diese Suchintention bedienen soll, sollte Schritt für Schritt die „Wie mache ich ...?“-Fragen beantworten.
- » **Aggregation/Überblick:** Ähnlich wie bei der Expansional-Suchintention geht es dem Nutzer darum, sich einen neutralen Überblick über ein Thema zu verschaffen. Allerdings sollte der Content hier möglichst knapp und übersichtlich gehalten werden, z. B. in Form von Tabellen, thematischen oder Random Listicles.

Transactional / Commercial Micro-Intents

- » **Vergleich/Orientierung:** Der Nutzer befindet sich auf dem Weg zum Kauf einer Dienstleistung eines Produkts oder zeigt zumindest Interesse zu investieren. Um sich einen Überblick zu verschaffen, sucht er nach der besten Lösung. Ranked Listicles, Tests oder Vergleiche machen für diese Suchintention Sinn.
- » **Kategorie/Auswahl:** Bei dieser Suchintention besteht ein konkretes Interesse an Produkten und Dienstleistungen. Der Nutzer weiß im Groben, welche Lösung für ihn die richtige ist, ist sich aber noch nicht sicher, welche Variante einer Leistungs- oder Produktgruppe besser passt. Klassische Shop-Kategorie-Seiten oder Leistungs-Übersichtsseiten sind optimal

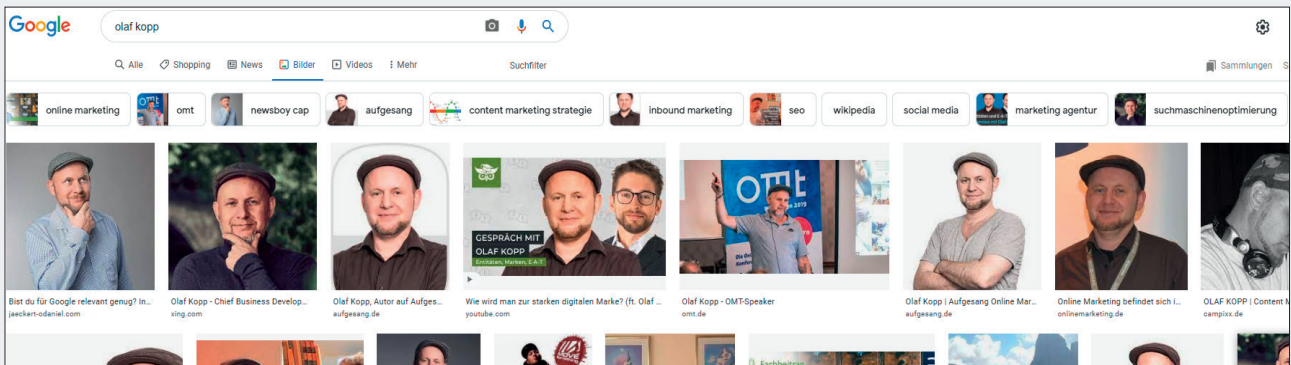


Abb. 17: Bilder-Kategorisierung in der Google-Bilder-Suche (Quelle: Google)

für diese Suchintention. Die Produkte und/oder Leistungen sollten im Fokus des Main Contents (MC) stehen, ggf. begleitet von Informationen, um die Entscheidung für eine Variante zu vereinfachen.

- » **Leistung/Produkt:** Der Nutzer weiß ziemlich genau, was er möchte bzw. welche Lösung die richtige für ihn ist. Er steht kurz vor der Anfrage/Bestellung und möchte sich im Detail über Eigenschaften, Preis, Lieferung, Lieferkosten, Garantien ... informieren. Leistungs- und Produktdetailseiten machen Sinn, um diese Suchintention zu bedienen.
- » **Brand:** Ich habe neben der klassischen Suchintention Brand einen weiteren Micro-Intent Brand ergänzt. Dieser spielt weniger für die Suche eine Rolle. Vielmehr berücksichtigt dieser das Bedürfnis des Nutzers, mehr über die Marke bzw. den Anbieter herauszufinden, um Vertrauen aufzubauen. Typische Content-Arten sind hier Testimonials, Kundenstimmen ...

auch das wichtigste SEO-Tool der Welt, nämlich das erfahrene und geschulte Hirn, zukünftig immer entscheidender.

Ein Nutzer, der nach objektiven Informationen ohne direkte kommerzielle Absicht sucht, befindet sich in der Awareness-Phase. Ein Nutzer, der Mischformen wie Trans/Info- oder Info/Trans-Suchanfragen ausführt, wird einen Schritt weiter sein. Ein Nutzer, der nach Produktgruppen oder konkreten Produkten sucht, wird häufig in der Preference-Phase seiner Journey sein. Es sei denn, es handelt sich um eine völlig neue Produktart oder eine komplexe Dienstleistung, die erst einmal einer grundsätzlichen Erfahrung bedarf.

Einige SEO-Tool-Anbieter geben vor, die Sucherintention je Keyword feststellen zu können, was aus Erfahrung aber oft zu oberflächlich ist und teilweise auch falsche Ergebnisse zur Folge hat. Da die Bestimmung der Such- und Nutzerintention sowohl für Google als auch für SEOs eine der wichtigsten Aufgaben ist, sollte man sich hier nicht auf SEO-Tools verlassen. Eine manuelle SERP-Analyse wird kein SEO-Tool ersetzen können.

Navigational Micro-Intents

- » **Support:** Der Nutzer benötigt Service-Content für die Nutzung eines bestellten Produkts. Benutzungsanleitungen und produktbezogene FAQs machen hier als Content Sinn.

Über die Suchintentionen lassen sich Informationen in eine Reihenfolge entlang der Customer Journey bringen und hier werden sowohl die Google SERPs als

SEO auch für andere Medienformate als Text

SEOs konzentrieren sich in erster Linie auf Text-Content. Durch MUM werden die SERPs deutlich diverser, was Medienformate angeht, da Google immer besser darin wird, neben Text auch Video, Bild und Audio zu verstehen und in einen Kontext zu bringen. Man sieht es

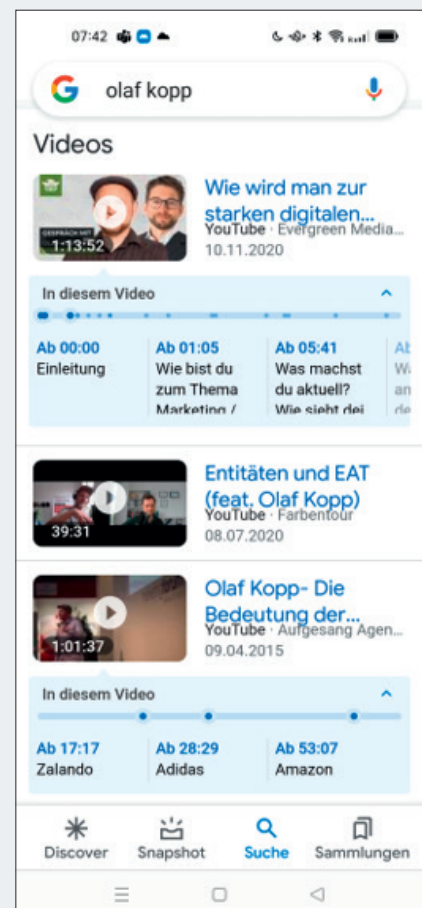


Abb. 18: Automatisierte Entitäten-Erkennung in YouTube-Videos (Quelle: Google)

bereits, wenn man sich z. B. die Klassifizierung von Bildern in der Bilder-Suche ansieht oder die automatisierte Markierung von Stellen in Videos.

Das dritte Video (siehe Abb. 18) wurde nie mit Mark-ups versehen. Google hat die genannten Entitäten selbstständig aus der Ton-Spur des Videos erkannt.

Somit wird Google dem sich verändernden Medien-Konsum hin zu mehr

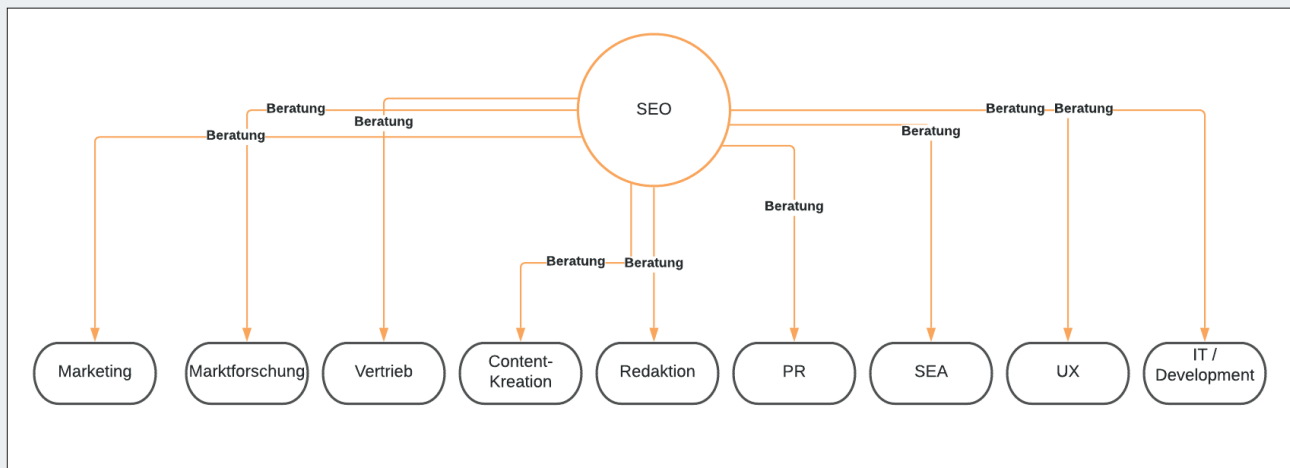


Abb. 19: Die neue Rolle des SEOs, Quelle: Quo Vadis SEO? Welche Rolle spielt SEO in der Zukunft? (Quelle: Olaf Kopp, <http://einfach.st/semd54>)

Audio und Video gerecht und SEOs sollten sich auch mit der Optimierung von Videos, Audios (z. B. Podcasts), Bildern beschäftigen. Und das beginnt nicht erst mit der Optimierung von Text (Titel, Beschreibung ...) um die Medien an sich, sondern schon mit dem Inhalt.

Wie bei einem Text sollte also auch in Videos und Podcasts auf die Verwendung semantisch passender Begriffe geachtet werden.

Schaffung semantischer Offpage-Kontexte

Die Schaffung semantischer Kontexte spielt nicht nur bei der Erstellung von Website-Texten eine wichtige Rolle, sondern auch bei Konkurrenzen der eigenen Entität (Autor und/oder Unternehmen) mit den thementypischen Begriffen außerhalb der eigenen Website, z. B. in Suchanfragen nach der eigenen Brand, Berichterstattung und Nennungen in Drittmedien. Die Einbettung der eigenen Unternehmens- und Personen-Marke in thematisch relevanten Kontext ist mindestens genauso wichtig für das Google-Ranking wie die Schaffung von thematisch relevantem Onpage-Content.

Für die Schaffung dieses semantischen Offpage-Kontexts sind Verantwortliche aus PR, Social-Media-Marketing, Content-Creator, Corporate Influencer und Marketing zuständig.

Die neue Rolle des SEOs

Die semantische Suche stellt SEOs vor neue Herausforderungen. Der direkte Einfluss von SEOs, durch Umsetzung von klassischen SEO-Maßnahmen zum Erfolg zu kommen, sinkt, vor allem in einem in den Suchmaschinen umkämpften Marktumfeld und den Your-Money-Your-Life-Themen wie z. B. Healthcare, Finanzen und E-Commerce.

Hier besteht die neue Aufgabe des SEOs, die Abteilungen bzw. Personen über den möglichen Einfluss ihres Tuns auf die Google-Rankings aufzuklären und sie zu motivieren, das eigene Unternehmen immer wieder in den entsprechenden thematischen Kontext zu setzen.

Zukünftig weniger Besucher aus Google durch MUM?

Durch Innovationen wie MUM und BERT will Google noch mehr Antworten direkt in den SERPs ausspielen, ohne

einen weiteren Klick auf die Quelle des Inhalts notwendig zu machen. Hier besteht die berechtigte Sorge, dass Google weiter den Traffic-Hahn abdrehen und so viele Informationen wie möglich in der eigenen Welt ausspielen wird.

Hier besteht die Gefahr, dass die Interessen von Google und den Content-Publishern auseinandergehen und Google sich dankend der passenden Inhaltspassagen bedient, ohne den Publisher teilhaben zu lassen. Das liegt aber nur in der Hand von Google selbst, wie sie den Ausgleich der Interessen berücksichtigen. Und eins ist klar: Google ist auf aktuellen Content angewiesen, um bestehende und zukünftige Fragen der Nutzer zu beantworten. Und als Technologie-Konzern ist Google in der Lage, Informationen algorithmisch zu indexieren und nutzergerecht aufzubereiten. Sie werden aber vermutlich nie in der Lage sein, gerade tiefgehendes Spezialisten-Wissen selbstständig neu aufzubauen und unabhängig von durch Publisher bereitgestelltem Content abzubilden. Deswegen kann man hier nur vertrauen, dass Google weiterhin guten Content mit Traffic belohnt. ¶

