

Von word2vec bis irgendwas mit GPT

Entwicklung der Sprachmodelle

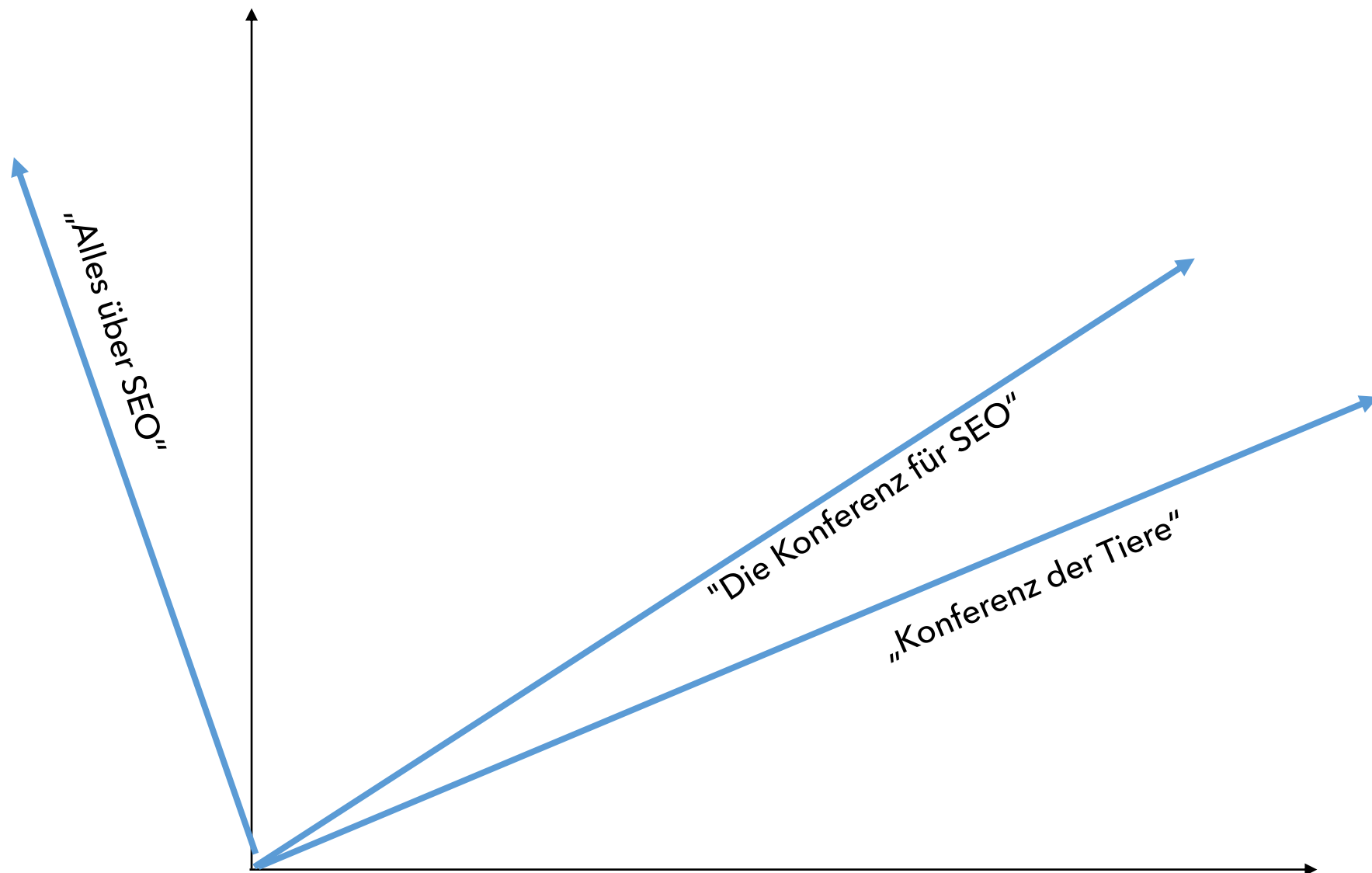
Campixx 2023 – Stefan Fischerländer

Ein Vektor sagt mehr als tausend Wörter

Stefan Fischerländer
SEO Campixx 2017

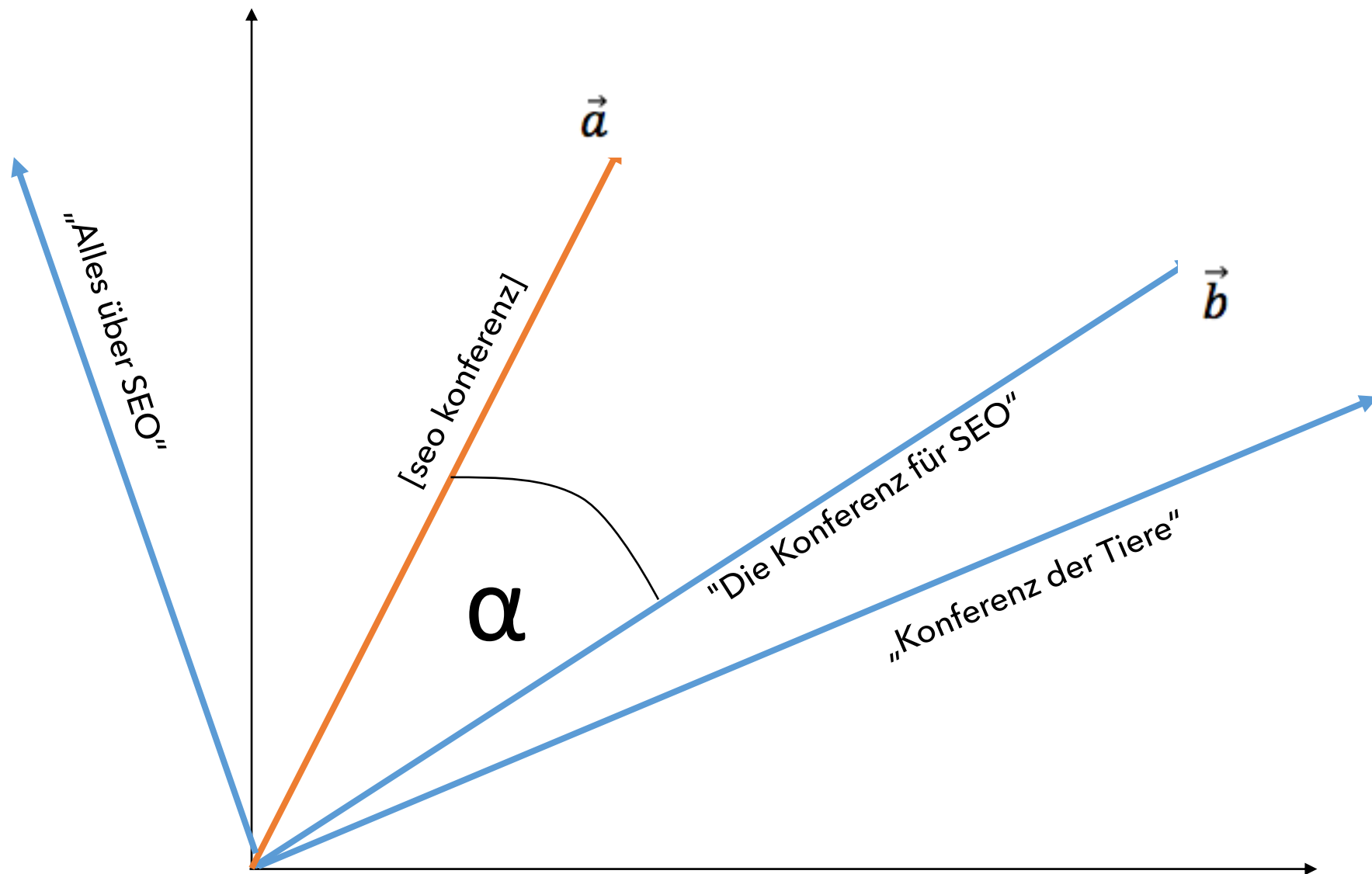
Vector Space Model

Jedes Dokument wird von einem Vektor repräsentiert



Auch die Suchanfrage ist ein Dokument, das durch einen Vektor repräsentiert wird.

Zum Vektor der Suchanfrage $\vec{a}$ wird das Dokument $\vec{b}$ gesucht, das mit $\vec{a}$ den kleinsten Winkel einschließt.



$$\cos(\alpha) = \frac{\vec{a} \cdot \vec{b}}{|\vec{a}| |\vec{b}|}$$

Das Vector Space Model ist die Grundlage der Suchmaschinen

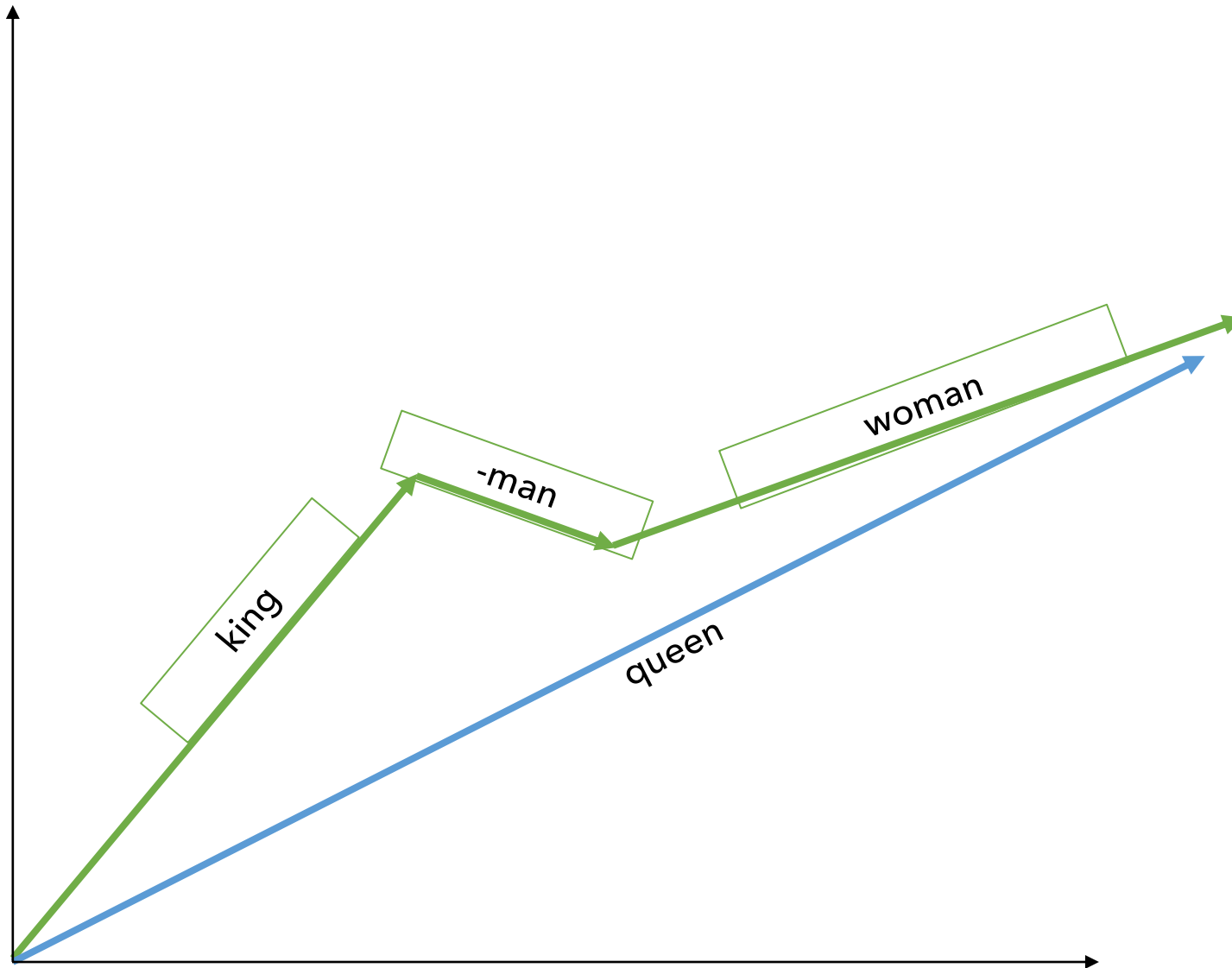
word2vec

Jedes Wort wird von einem Vektor repräsentiert

king – man + woman = queen



king - man + woman = queen



Wörter werden durch Vektoren repräsentiert.

Somit können wir mit Wörtern und ihren Bedeutungen rechnen.

Mein word2vec-Modell (2017)

- Deutsche Wikipedia (ca. 5GB Text)
- Trainingsdauer: ca. 5 Stunden
- Standard-PC

BVB – Dortmund = FCB

Yandex – Russland = Suchmaschine

Canberra – Australien = Hauptstadt

Canberra – Australien + Kanada = Ottawa

Hauptstädte raten

word2vec

7 / 10

Ungarn

Großbritannien

Dänemark

Malta

USA

Kanada

Mali

Ghana

Malawi

Thailand

Was hat sich seit 2017 getan?



A drawing with the letters GPT in front of an impressive sunrise. Comic style.

Ideen erkunden

Kreationen



Inhaltswarnung

Diese Eingabeaufforderung wurde blockiert. Unser System hat diese Eingabeaufforderung markiert, weil sie mit unserer [content policy](#) in Konflikt stehen könnte. Weitere Verstöße gegen die Richtlinien können zur automatischen Sperrung Ihres Zugangs führen.

Wenn Sie der Meinung sind, dass dies ein Fehler ist, melden Sie ihn, um uns bei der Verbesserung zu unterstützen.

Bericht

Zurück

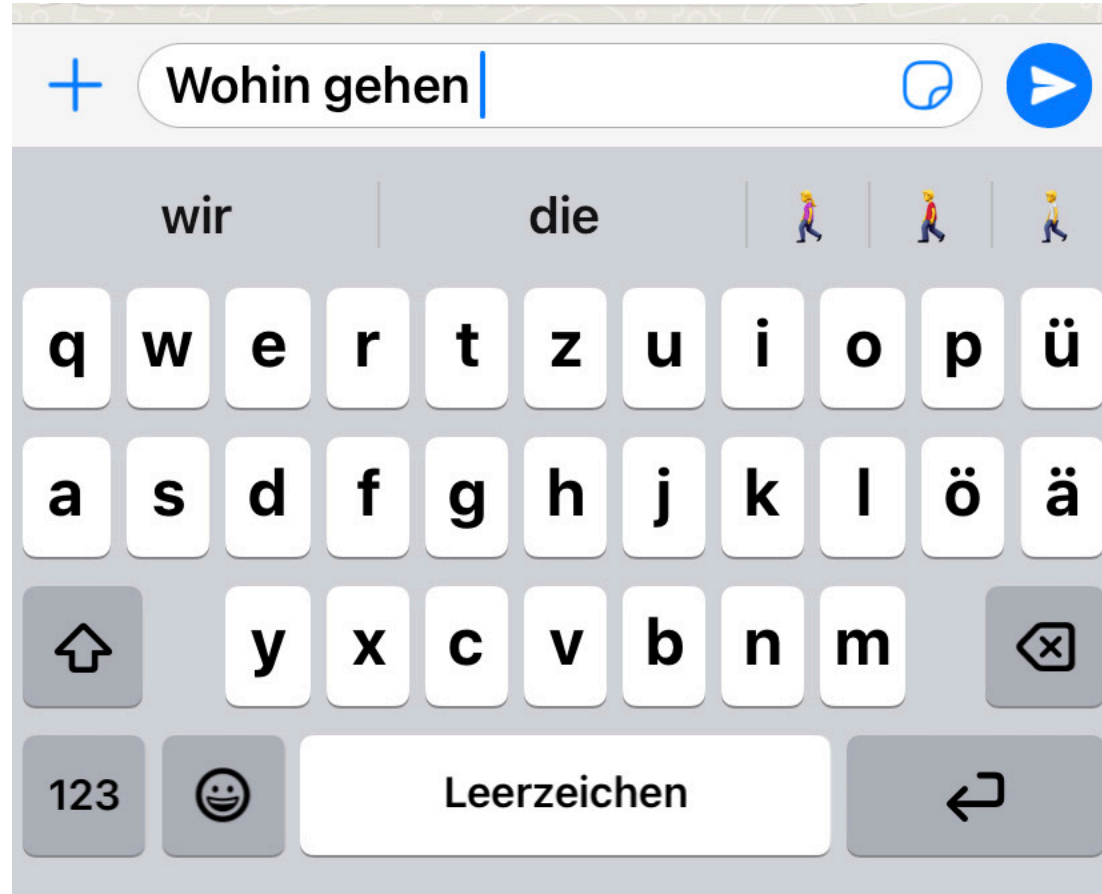
Generative Pre-trained Transformer



Im Großen und ...

... Ganzen

... im Kleinen



Was kann ein Sprachmodell, das
auf 6 KB trainiert wurde?

Wie Ihre Firmenstruktur die SEO-Maßnahmen torpediert

Die Struktur einer Website ist für den Erfolg dieser Website von entscheidender Bedeutung. Das gilt für den Erfolg in Google genau so wie für den Erfolg bei den Nutzern. Beides lässt sich gar nicht mehr voneinander trennen. Sie glauben nun, Sie würden die Struktur Ihrer Website so festlegen, wie es sinnvoll ist und sind vor äußeren Einflüssen, die nichts mit der Website zu tun haben, gefeit?

Dann sollten Sie sich unbedingt mit dem *Gesetz von Conway* auseinander setzen.

Die Wikipedia fasst dieses Gesetz so zusammen:

“ *Organisationen, die Systeme entwerfen, sind auf Entwürfe festgelegt, welche die Kommunikationsstrukturen dieser Organisationen abbilden.*

Wenn Organisationen Systeme bauen, spiegeln sich also die Strukturen der Organisation im System wider.

1 Die Struktur einer Website ist für den Erfolg dieser Website von entscheidender Bedeutung. Das gilt für den Erfolg in
· Google genau so wie für den Erfolg bei den Nutzern. Beides lässt sich gar nicht mehr voneinander trennen. Sie glauben
· nun, Sie würden die Struktur Ihrer Website so festlegen, wie es sinnvoll ist und sind vor äußeren Einflüssen, die
· nichts mit der Website zu tun haben, gefeit?↵

2 ↵

3 Dann sollten Sie sich unbedingt mit dem Gesetz von Conway auseinander setzen.↵

4 ↵

5 Die Wikipedia fasst dieses Gesetz so zusammen: "Organisationen, die Systeme entwerfen, sind auf Entwürfe festgelegt,
· welche die Kommunikationsstrukturen dieser Organisationen abbilden."↵

6 ↵

7 Wenn Organisationen Systeme bauen, spiegeln sich also die Strukturen der Organisation im System wider. Deshalb bauen
· große und komplexe Unternehmen große und komplexe Websites. Jetzt sind Websites, die der Firmenstruktur gleichen, gut
· fürs Unternehmen, weil für jeden Bereich auf der Website klar ist, wer im Unternehmen dafür den Hut aufhat. Doch der
· Nutzer interessiert sich nicht im geringsten für die Struktur der Firma, die hinter der besuchten Website steht. Er
· möchte nur schnell die Informationen finden, die er aktuell benötigt. Die Struktur der Site sollte sich entsprechend
· am Bedürfnis des Nutzers orientieren und nicht an der womöglich lange gewachsenen Firmenstruktur.↵

8 ↵

9 Ich habe vor einigen Jahren ein Unternehmen beraten, das aus der Fusion mehrerer eigenständiger Unternehmen
· hervorgegangen war. Dort wurde sehr genau darauf geachtet, dass alle einstmals eigenständigen Firmen in der
· Außendarstellung gleichberechtigt erscheinen. Ich habe dort dann den „Fehler“ gemacht, nur Produkte aus drei der vier
· Unternehmensteile für die Darstellung auf der Startseite vorzuschlagen; ganz einfach deshalb, weil das Suchvolumen
· des vierten Bereichs so gering war, dass eine derart prominente Platzierung aus SEO-Gesichtspunkten nicht sinnvoll
· erschien. Dieser Vorschlag wurde sofort einkassiert und wir mussten wertvollen Platz auf der Startseite und im
· Seitentitel für unwichtige Begriffe verschwenden. Die Abbildung der Unternehmensstruktur im System „Website“ konnte
· ich auch durch gutes Zureden nicht überwinden. Der SEO-Erfolg war entsprechend geringer als er hätte sein können.↵

10 ↵

11 Der mäßige Erfolg von Windows Vista — böse Zungen sagen es sei ein veritabler Flop gewesen — beruht nach einer
· Microsoft-eigenen Untersuchung darauf, dass die Teams, die Windows Vista gebaut haben, zu komplex waren. Sie waren
· deshalb nach Conway dazu verdammt, ein zu komplexes Produkt zu entwickeln. Und in der Software-Entwicklung heißt das
· meist: Es ist langsam, enthält viele Fehler und richtet sich nicht am Nutzerbedürfnis aus.↵

12 ↵

13 Websites wie auch Apps sind letztlich Systeme, die von Organisationen gebaut werden. Die Strukturen dieser Websites
· und Apps haben entscheidenden Einfluss auf die Usability. In einer idealen Welt, würde die Struktur von Online-Medien

```
python data/sf1/prepare.py input.txt
```

GPT-2-Tokens:

<Die> <Stru> <kt>
<ur> <e> <iner>
<Website> <is> <t>
<f> <ür> <den> <Er>
<fol> <g> <dies> <er>
<Website> <von>
<ents> <che> <id>
<ender> <Bed> <e>
<ut> <ung> <.>

```
python train.py config/train_sf1.py
--eval_iters=100 --log_interval=10 -
-block_size=64 --batch_size=12 --
n_layer=12 --n_head=12 --n_embd=144
--max_iters=20000 --
lr_decay_iters=20000 --dropout=0.0 -
-learning_rate=1e-4 --min_lr=1e-5
```

1.5h warten ...

```
python sample.py --out_dir=out_sf1 -
-device=mps --start="Wenn
Organisationen Systeme "
```

Die Strukturen müssen
ist noch für Sie sich
nichtagenine-
Usability. Das
geringhr keinzerf
Erfolg im
Unternehmen. Doch
häre die Struktur die
Struktur der
Organisationen sind
die häre.nis
aufrieden
Unternehmen.

Was kann ein Sprachmodell, das
auf 1,8 MB trainiert wurde?

Stefans Tiny SEO GPT

Korpus: 1,8 MB englischsprachige
SEO-Tutorials aus dem Web

Training: 3 Stunden auf einem PC mit
10 Jahre alter Grafikkarte

Prompt: „Use these steps to improve
your rankings:“

Write a backlink
campaign

Picking a link

Paid Backlinks

Domain Authority

Find something

Stefans Tiny-SEO-GPT – lyrisch

“

Search engines use the tag.
Search engines use words.

Stefans Tiny-SEO-GPT – hilfreich

“

You should be careful when you pay an
SEO expert

```
(pynanogpt) sf@MBA nanoGPT % python sample.py --out_dir=out-seotutorial --device=mps  
--start="The most important steps for SEO are"  
Overriding: out_dir = out-seotutorial  
Overriding: device = mps  
Overriding: start = The most important steps for SEO are  
number of parameters: 10.23M  
No meta.pkl found, assuming GPT-2 encodings...  
The most important steps for SEO are a web site to keep on your site traffic as possible  
that is also very easy for users that site visitors to set in a given site that  
is the practice of pages of the search engine crawlers are saying that your site may  
crawl, and they need to be updated every web analytics tool. However, we can also  
allow users to tell you what your site will change through the other are relevant. But  
if you can't finish.  
Yahoo! Search Engines can see any other search results, but this's an algorithm and  
is listed in a search engine index, and Yahoo! is not the organic search engine results.  
AdWords is only the actual rank. But Yahoo! has the little better targeted keywords  
in the two search results. The Google are compared to understand, so that's the  
last number of the most users is for the search results will vary in the rankings. I  
sometimes have Google already, and Yahoo! is a highly targeted ranking. Moreover, true  
searchers can also like that page searches must be lower placement, and PPC programs  
are highly targeted.  
Your Business Ranking of the final community are of the most important SEO strategies  
and what companies are connected to your marketing industry. Keywords you're paying  
for marketing and PPC programs that have the beginning to help you're using for the  
keywords that you're taken to your ad copy.
```

Sprachgenerierung

GPT-3 war das erste Modell, mit dem computergenerierte Texte ohne große Nacharbeit schon richtig gut waren.

Sprachverständnis

ChatGPT ist ein GPT-3-Modell, das mit vielen Konversationen feingetuned wurde.

Damit kann es Sprache nicht nur generieren, sondern auch „verstehen“.

Interface

```
model.most_similar(positive=['Canberra', 'Kanada'], negative=['Australien'])
```

```
[(u'Ottawa', 0.7005172371864319),  
(u'Ontario', 0.6972213983535767),  
(u'Montreal', 0.6887485384941101),  
(u'Toronto', 0.6846897006034851),  
(u'Winnipeg', 0.6674027442932129),  
(u'Mississauga', 0.6657336950302124),  
(u'Saskatoon', 0.6404088735580444),  
(u'Halifax', 0.6335928440093994),  
(u'Alberta', 0.6323922276496887),  
(u'Saskatchewan', 0.626349687576294)]
```

ST

What's the capital of Canada? Give me just the answer without any explanations.



Ottawa.



Finetuning

Ein auf allgemeinen Daten trainiertes LLM wird mit spezifischen Daten für eine Nische verfeinert.

Überblick Sprachmodelle

	word2vec	GPT-2	GPT-3
Erschienen:	2013	2019	2020 (2022)
Anmerkung:	Software Open Source	Software Open Source	ChatGPT
Trainingsdaten:	5 GB	40 GB	1.000 GB
Parameter:	n/a	1,5 Mrd.	175 Mrd.
Training:	\$0,20	\$100.000	\$10 Mio.

Hauptstädte raten

word2vec

7 / 10

Ungarn

Großbritannien

Dänemark

Malta

USA

Kanada

Mali

Ghana

Malawi

Thailand

Welcher Begriff gehört nicht dazu?



Stellen Sie Ihre Intelligenz auf die Probe! Jeweils einer der folgende Begriffe passt logisch nicht zu den anderen - welcher ist es?

Wenn mehrere Zusammenhänge möglich sind, sollte die einfachste Erklärung gewählt werden. Es geht nicht um äußerliche Eigenheiten der Worte wie Anfangsbuchstaben, Silbenzahl oder Konsonanten. Lediglich die Begriffe selbst sind von Bedeutung. Die Aufgaben stammen aus offiziellen Intelligenztests.

1

Aus den folgenden Begriffen muss derjenige ausgewählt werden, der logisch nicht zu den anderen passt.

A Biene

B Ameise

C Fliege

D Spinne

NEU IN LEBEN —



Welcher Begriff gehört nicht dazu?

word2vec



ChatGPT



Google Bard



Wie kann das sein?

- Für mein word2vec nur Wikipedia-Korpus benutzt, damit extrem hohe Faktendichte.
- Die Wissensfragen waren der perfekte Use-Case für word2vec.

Sprachmodelle, keine Wissensmodelle

Überblick Sprachmodelle

	word2vec	GPT-2	GPT-3
Erschienen:	2013	2019	2020 (2022)
Anmerkung:	Software Open Source	Software Open Source	ChatGPT
Trainingsdaten:	5 GB	40 GB	1.000 GB
Parameter:	n/a	1,5 Mrd.	175 Mrd.
Training:	\$0,20	\$100.000	\$10 Mio.

Überblick Sprachmodelle

	word2vec	GPT-2	GPT-3	LLaMa
Erschienen:	2013	2019	2020 (2022)	2023
Anmerkung:	Software Open Source	Software Open Source	ChatGPT	Software Open Source
Trainingsdaten:	5 GB	40 GB	1.000 GB	2.000 GB
Parameter:	n/a	1,5 Mrd.	175 Mrd.	65 Mrd.
Training:	\$0,20	\$100.000	\$10 Mio.	\$4 Mio.

Überblick Sprachmodelle

	word2vec	GPT-2	GPT-3	LLaMa	Falcon
Erschienen:	2013	2019	2020 (2022)	2023	2023
Anmerkung:	Software Open Source	Software Open Source	ChatGPT	Software Open Source	Bestes Open Source Model
Trainingsdaten:	5 GB	40 GB	1.000 GB	2.000 GB	1.000 GB
Parameter:	n/a	1,5 Mrd.	175 Mrd.	65 Mrd.	40 Mrd.
Training:	\$0,20	\$100.000	\$10 Mio.	\$4 Mio.	

Überblick Sprachmodelle

	word2vec	GPT-2	GPT-3	LLaMa	Falcon	PaLM2
Erschienen:	2013	2019	2020 (2022)	2023	2023	2023
Anmerkung:	Software Open Source	Software Open Source	ChatGPT	Software Open Source	Bestes Open Source Model	Google Bard
Trainingsdaten:	5 GB	40 GB	1.000 GB	2.000 GB	1.000 GB	5.000 GB
Parameter:	n/a	1,5 Mrd.	175 Mrd.	65 Mrd.	40 Mrd.	340 Mrd.
Training:	\$0,20	\$100.000	\$10 Mio.	\$4 Mio.		\$10 Mio.

Attention Is All You Need

Ashish Vaswani*
Google Brain
avaswani@google.com

Noam Shazeer*
Google Brain
noam@google.com

Niki Parmar*
Google Research
nikip@google.com

Jakob Uszkoreit*
Google Research
usz@google.com

Llion Jones*
Google Research
llion@google.com

Aidan N. Gomez*[†]
University of Toronto
aidan@cs.toronto.edu

Łukasz Kaiser*
Google Brain
lukaszkaizer@google.com

Illia Polosukhin*[‡]
illia.polosukhin@gmail.com

Abstract

The dominant sequence transduction models are based on complex recurrent or convolutional neural networks that include an encoder and a decoder. The best performing models also connect the encoder and decoder through an attention mechanism. We propose a new simple network architecture, the **Transformer**, based solely on attention mechanisms, dispensing with recurrence and convolutions

“

In October 2020, almost every single English-based query was processed by a BERT model.

Quelle: Search Engine Land

SUCHMASCHINE

Google ruft Code Red gegen Bedrohung durch ChatGPT aus

Sundar Pichai hat die Co-Gründer Larry Page und Sergey Brin als Berater eingeladen. Apps wie [ChatGPT](#) seien eine Bedrohung für den Konzern.

Customer Journey



Ich brauche
Urlaub

Reiseziele in
Italien

Ferienhaus
Chianti

Ferienhaus *Villa
Lecchi* buchen

Wie komme ich
zum Ferienhaus?

Italien wäre nett

Urlaub Toskana

Ferienhaus
Chianti mit Hund

Ab wann ist das
Zimmer frei?

Für Suche interessant

ST

Ich möchte Urlaub in Italien machen. Ich mag gutes Essen und guten Wein. Meer und Strand brauche ich nicht unbedingt. Welche Reiseziele kannst du mir empfehlen?



Wenn Sie gutes Essen und guten Wein suchen und Strand nicht unbedingt benötigen, gibt es in Italien viele wunderbare Reiseziele, die Ihren Wünschen entsprechen könnten. Hier sind einige Empfehlungen:



1. Florenz: Die Hauptstadt der Toskana ist bekannt für ihre exquisite toskanische Küche und hervorragende Weine wie Chianti und Brunello di Montalcino. Sie können auch die reiche Kunst- und Kulturszene der Stadt genießen.
2. Bologna: Diese lebendige Stadt in der Emilia-Romagna ist ein Paradies für Feinschmecker. Probieren Sie die berühmte Bolognese-Sauce, Tortellini und Mortadella. Die Region ist auch für ihre hervorragenden Weine wie Lambrusco bekannt.
3. Piemont: Wenn Sie sich für Weine interessieren, sollten Sie das Piemont in Norditalien besuchen. Die Region ist für ihre erlesenen Rotweine wie Barolo und Barbaresco bekannt. Hier können Sie auch köstliche Gerichte wie Trüffel und Agnolotti genießen.

ST

Toskana könnte mir gefallen, aber Florenz ist mir zu groß. Welche kleineren Städte oder ländlichen Gegenden kannst du empfehlen?



Wenn Sie die Toskana erkunden möchten, aber Florenz als zu groß empfinden, gibt es viele charmante kleinere Städte und ländliche Gegenden, die Sie besuchen können. Hier sind einige Empfehlungen:



1. Siena: Siena ist eine wunderschöne mittelalterliche Stadt mit einer gut erhaltenen Altstadt. Sie können die berühmte Piazza del Campo besuchen, den beeindruckenden Dom bewundern und durch die engen Gassen schlendern. Die toskanische Küche und Weine können Sie auch hier genießen.
2. San Gimignano: Diese kleine Stadt ist für ihre markanten Türme bekannt und bietet eine reizvolle mittelalterliche Atmosphäre. Sie können die historische Altstadt erkunden, lokale Weine probieren und das atemberaubende Panorama der umliegenden Landschaft genießen.
3. Volterra: Volterra ist eine charmante Etruskerstadt, die auf einem Hügel liegt. Hier können Sie die gut erhaltene römische und mittelalterliche Architektur bewundern, lokale

ST

Wenn ich in die Gegend um Siena fahren möchte, welche Unterkünfte wären da empfehlenswert? Sollte ich eher nach Hotels schauen oder lieber Ferienwohnungen?



Wenn Sie die Gegend um Siena erkunden möchten, haben Sie verschiedene Optionen für Unterkünfte. Sowohl Hotels als auch Ferienwohnungen können gute Möglichkeiten sein, abhängig von Ihren Vorlieben und Bedürfnissen. Hier sind einige Empfehlungen:



Hotels:

1. Relais La Suvera: Ein luxuriöses Hotel in einem restaurierten Schloss mit antikem Charme und einem atemberaubenden Blick auf die umliegende Landschaft.
2. Hotel Athena: Ein stilvolles Hotel im Herzen von Siena, nur wenige Schritte von der Piazza del Campo entfernt. Es bietet komfortable Zimmer und eine hervorragende Lage für die Erkundung der Stadt.
3. Borgo Scopeto Relais: Ein reizvolles Landhotel in einem renovierten Weingut. Hier können Sie die Ruhe des ländlichen Lebens genießen und gleichzeitig die Nähe zu Siena haben.

ST

Klasse. Weißt du, wo ich die Villa Lecchi buchen kann und was das kosten wird?



Als KI-Modell habe ich keinen direkten Zugriff auf Echtzeitinformationen oder Buchungsplattformen. Allerdings kann ich Ihnen einige allgemeine Hinweise geben, wie Sie eine Buchung für die Villa Lecchi oder ähnliche Unterkünfte vornehmen können.



1. Online-Buchungsplattformen: Websites wie Booking.com, Expedia, Airbnb oder HomeAway bieten oft eine große Auswahl an Ferienvillen und Ferienwohnungen. Sie können nach der Villa Lecchi suchen oder Ihre Präferenzen angeben, um ähnliche Unterkünfte in der Region Siena zu finden.
2. Offizielle Website der Villa Lecchi: Überprüfen Sie, ob die Villa Lecchi eine eigene offizielle Website hat. Auf solchen Websites können Sie oft direkt buchen und erhalten möglicherweise spezielle Angebote oder bessere Tarife.
3. Reiseveranstalter oder Reisebüros: Es kann hilfreich sein, lokale Reiseveranstalter oder Reisebüros zu kontaktieren, die auf die Toskana spezialisiert sind. Sie können Ihnen bei der Suche nach geeigneten Unterkünften und der Buchung behilflich sein.

Customer Journey



Ich brauche
Urlaub

Reiseziele in
Italien

Ferienhaus
Chianti

Ferienhaus *Villa
Lecchi* buchen

Wie komme ich
zum Ferienhaus?

Italien wäre nett

Urlaub Toskana

Ferienhaus
Chianti mit Hund

Ab wann ist das
Zimmer frei?

Für Suche interessant

[↪ Converse](#)[Shopping](#)[Images](#)[Videos](#)[News](#)[Maps](#)[Books](#)[Flights](#)[Finance](#)

Generative AI is experimental



Both Bryce Canyon and Arches National Parks are family-friendly. Although both parks prohibit dogs on unpaved trails, Bryce Canyon has two paved trails that allow dogs.

Bryce Canyon has distinctive features like hoodoos, natural bridges, and waterfalls. At the visitor center's interactive exhibits, children can learn about the geology, wildlife, and people of the area. You can bring your dog on the paved sections of both the Rim Trail and the Shared Use Path, both of which are also stroller-friendly.

At Arches, children may enjoy the rock formations, and some say that Arches has more variety than Bryce Canyon. Pets aren't allowed on any trails, though they are allowed at campgrounds, pullouts, roads, and parking lots.

Both parks offer guided tours, ranger-led programs, picnic areas, and restrooms, and both parks require pets to be leashed.



Kids & Youth -
Bryce Canyon
National Park...



Arches
National Park
vs Bryce...



Visiting Bryce
Canyon
National Park...

[↪ Ask a follow up](#)[↪ How long to spend at Bryce Canyon with kids?](#)[↪ How many days do you need in Arches National Park for kids?](#)

Die Nutzung von LLM in der Suche ist keine primär technologische Aufgabe.

Es ist eine Frage des Produkts.

“

We are a technology company!

Google war noch nie besonders gut darin, gute Produkte jenseits der Suche zu bauen ...

Chatbots als Standard-Interface

- Die Nutzung von Chatbots ist künftig allgegenwärtig
- „Fear of missing out for not using Google“ verschwindet

GPT-Modelle für spezielle Aufgaben

EinsteinGPT

Sprachmodell von Salesforce

Basiert auf OpenAI GPT-3.5

Zugriff auf Kundendaten und Live-Services wie Wetterdaten

BloombergGPT

Sprachmodell für den Finanzbereich

Basiert auf Bloomberg-Meldungen, Unternehmensveröffentlichungen und Nachrichten

Geschätzte Trainingskosten: \$2,7 Mio.

Github Copilot

Sprachmodell für Computersprachen

Basiert auf GPT-3

Trainiert auf öffentlich verfügbarem Sourcecode



MingaGPT

Servus, I bin de MingaGPT und helfe gern bei ois rund um Minga. Wia kann i dia heid?

Wer is da beste Fußball-Verein in Minga?

Des kann i dir leider ned sogn, weil i in de Informationen, die i hab, nix dazu find. Aber wennst a anan Sportverein suchst, dann schau doch mal bei [diesen Sportanbietern](#) oder [hier](#) vorbei. Vielleicht is ja was dabei für di!

Wer is da beste Fußball-Verein in Minga?

Des is a guade Frog! Aber wenn ma danach geht, wer am meistn Titel gholt hod, dann is des der [FC Bayern München](#). Die ham scho so vui Meisterschaften und Pokale gholt, dass ma gar ned mehr zählen ko. Aber natürlich is des a bissl subjektiv und jeder hod a eigene Meinung.

Stell mir a Frog



MingaGPT

[Do geats zu an Blogartikel über mi](#)

Servus und Grüß Gott bei MingaGPT, am Gschprächspartner, der eich bei olle Fragen rund um Minga weiterhelfen mog.

I mecht eich dabei unterstützen, Informationen auf de Seiten [muenchen.de](#) und [stadt.muenchen.de](#) aufzuspürn.

De Daten wer'n an OpenAI weidagegebn, bitte koa persönliche Daten eingeb'n! Merchts eich bitte, dass i nix Offizielles vo der Stadt München bin - i bin nur a Technologiedemo von der Mayflower GmbH.

I hoff allerdings, dass i eich trotzdem a Hand reichen ko.

Dei MingaGPT



Chan-jo Jun 🇺🇦 @Anwalt_Jun · 2h ...

Wir haben zusammen mit [@mayflowergmbh](#) ein solches Tool gebaut. Nötig ist Schnittstelle zu einem LLM und juristische Konfiguration. Wir haben unser "Praxishandbuch Open Source" zum Training reingefüttert. [shop.ruw.de/startseite/498...](#)
5/



Juristische Beraterin Mace

Wann verstoße ich gegen Lizenzbedingungen beim Einsatz von Linux

Ein Verstoß gegen die Lizenzbedingungen beim Einsatz von Linux kann beispielsweise dann vorliegen, wenn ein WLAN-Router, der die Software "netfilter/iptables" und "mtd" enthält, ohne Einhaltung der Lizenzbestimmungen der GPL verkauft wird. Ein solcher Verstoß führt automatisch zum Rechtsrückfall im Sinne einer auflösenden Bedingung nach § 158 Abs. 2 BGB. Es ist jedoch zu beachten, dass nicht jeder Lizenzverstoß automatisch strafbar ist und dass es je nach gegebenem Fall sinnvoll sein kann, eine Klassifizierung der Lizenzen nach deren Risikograd vorzunehmen und einen standardisierten FOSS-Prozess einzuführen.

Stellen Sie eine Frage ...



Eine Technologiedemo von JunIT | Kanzlei für IT- und Wirtschaftsrecht und MAYFLOWER



1



3



29



1.745





Juristische Beraterin Mace

Wann verstoße ich gegen Lizenzbedingungen beim Einsatz von Linux

Ein Verstoß gegen die Lizenzbedingungen beim Einsatz von Linux kann beispielsweise dann vorliegen, wenn ein WLAN-Router, der die Software "netfilter/iptables" und "mtd" enthält, ohne Einhaltung der Lizenzbestimmungen der GPL verkauft wird. Ein solcher Verstoß führt automatisch zum Rechtsrückfall im Sinne einer auflösenden Bedingung nach § 158 Abs. 2 BGB. Es ist jedoch zu beachten, dass nicht jeder Lizenzverstoß automatisch strafbar ist und dass es je nach gegebenem Fall sinnvoll sein kann, eine Klassifizierung der Lizenzen nach deren Risikograd vorzunehmen und einen standardisierten FOSS-Prozess einzuführen.

Stellen Sie eine Frage ...





Bedrohungen für Google



Kernkompetenz unbedeutend

- Crawling und Linkanalyse sind nicht mehr in der Perfektion nötig, die derzeit nur Google kann

2

Mehr Konkurrenz

- Einstiegshürden für Chatbots sind geringer als für eine Suchmaschine
- Open-Source-Software
- Open-Source-Modelle


```
sampling: repeat_last_n = 64, repeat_penalty = 1,100000, presence_penalty = 0,000000
, frequency_penalty = 0,000000, top_k = 40, tfs_z = 1,000000, top_p = 0,950000, typi
cal_p = 1,000000, temp = 0,800000, mirostat = 0, mirostat_lr = 0,100000, mirostat_en
t = 5,000000
generate: n_ctx = 512, n_batch = 512, n_predict = 256, n_keep = 0
```

The most important steps for SEO are still the basic on page factors. If you want to improve your website's ranking, it is crucial that every page of your site has an effective title tag and meta description. Also, make sure that you have a relevant page title for each page. Furthermore, avoid keyword stuffing since this could get you penalized by search engines.

The best way to rank high in Google's SERPs is still creating great content and marketing it. If the right people love your product or service, they will share it on social media. That is how your brand will reach more people and gain more visibility. But don't forget to include keywords so that you appear in searches for those terms. Don't get me wrong. I am not saying that having a blogger outreach campaign isn't important! In fact, it has become an essential part of building a brand. But even the most successful bloggers are going to tell you that backlinks are the single most important factor for ranking high in Google SERPs. Don't be shy in reaching out and asking for backlinks from relevant sites in your industry.

The first thing any marketer needs to do is figure out what topics get searchers to click on

```
llama_print_timings:          load time = 6761,85 ms
llama_print_timings:          sample time = 229,76 ms / 256 runs ( 0,90 ms per
```


3

Neue Geschäftsmodelle

- Bezahlmodelle für Suche sind plötzlich denkbar
- Spezielle Sprachmodelle für einzelne Themenbereiche

4

Weniger Websuchen

- Websuche wird vielfach durch Chatbots ersetzt
- Vor allem die monetär interessanten Suchanfragen gehen verloren

Ausblick



Rethinking Search: Making Domain Experts out of Dilettantes*

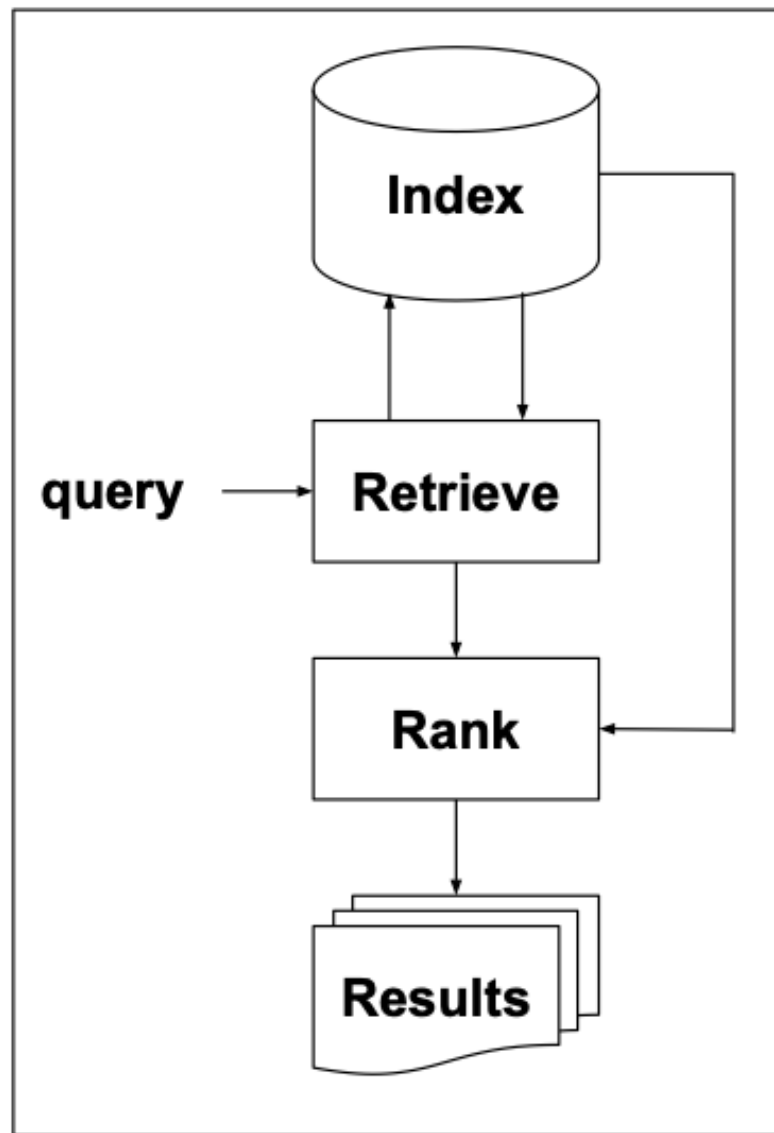
Donald Metzler
Google Research
metzler@google.com

Yi Tay
Google Research
yitay@google.com

Dara Bahri
Google Research
dbahri@google.com

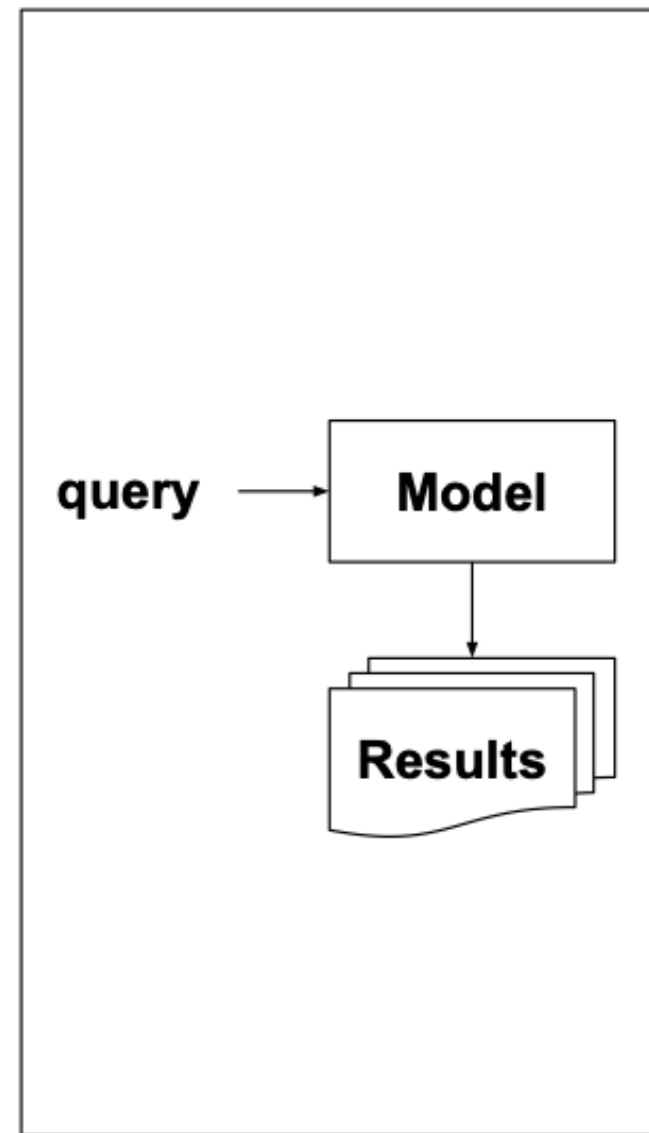
Marc Najork
Google Research
najork@google.com

“The paper makes the case for developing retrieval systems that combine the best elements of document retrieval systems and pre-trained language models



(a) Retrieve-then-rank

SEO



(b) Unified retrieve-and-rank

GAIO



Informational-Traffic nimmt ab

Die Chatbots liefern genau das, was auf Infoseiten mühsam zusammengestellt wurde.

Und der Nutzer kann nachhaken, ohne sich auf einer neuen Seite durch Cookie-Banner, Werbung und Newsletter-Popups kämpfen.



Keine Attributierung

Wie kann die Leistung in Chatbot-Konversationen dem GAIO zugeordnet werden?

Viele weitere Fragen

- Wie wirkt sich die fehlende Attributierung auf die Rolle der SEOs im Unternehmen aus?
- Wie sehr werden die Chats – bzw die ganze hybride Suche – personalisiert sein?
- Wie stellen LLM die Aktualität sicher? Wird es einen Bing-Chat-Dance geben?
- Stimmt meine These, dass sich Chatbots als Standard-Interface etablieren werden?

So wie Kodak die Digitalkamera entwickelt hat, hat Google die Transformer-Technologie entwickelt. Ob Google auch daran scheitert, aus seiner Entwicklung ein tragfähiges Geschäftsmodell zu machen?

Fotopionier beantragt Insolvenz

Kodak ist pleite

19. Januar 2012, 10:19 Uhr | Lesezeit: 1 min

Es war ein langer Überlebenskampf - umsonst: Der US-Kamerakonzern Eastman Kodak ist insolvent. Kodak hatte einst die analoge Fotografie entscheidend geprägt. Mit dem Wechsel zu digitalen Bildern kam das US-Unternehmen jedoch nie klar. In Deutschland können Mitarbeiter jedoch aufatmen. Die deutsche Kodak ist von der Insolvenz nicht betroffen.

Vielen Dank!



Stefan Fischerländer

stefan@fischerlaender.de

@stefanfis