

Frankfurter Allgemeine

Künstliche Intelligenz für Journalisten

Patrick Bernau, 29. September 2023



Agenda

Workshop:

1. Was ist Künstliche Intelligenz?
2. ChatGPT und Bing
3. Andere Werkzeuge



Was ist Künstliche Intelligenz?



Teil hier bis 16:20

Der Stammbaum der KI

- Symbolische KI
- Subsymbolische KI / Machine Learning
 - Generative KI
 - ❖ Bildgeneratoren (Midjourney, Dall-E, Stable Diffusion, Adobe Firefly)
 - ❖ Large Language Models
 - Luminous (Aleph Alpha)
 - LaMDA, PaLM etc. (Google)
 - Bard
 - LLaMA (Facebook)
 - GPT (OpenAI)
 - ChatGPT
 - Sydney (Bing)
 - Perplexity
 - ...
 - ❖ etc.



Neuronale Netzwerke

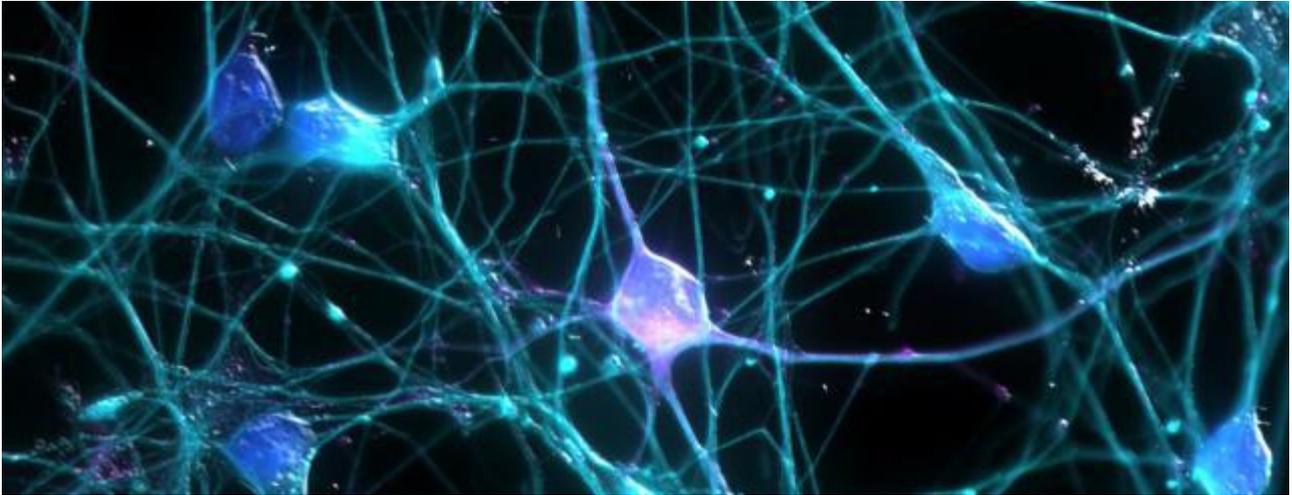


Foto: UCI Research / CC BY-NC



Vielleicht deshalb so ähnlich wie Menschen: Ihr Verhalten ist schwer zu erklären, sie sind total kreativ, haben aber Schwierigkeiten mit formeller Logik, und manchmal erzählen sie voller Überzeugung großen Mist.

Je komplexer Modelle werden, desto menschenähnlicher werden sie

So funktioniert das dann

„Die Hartz-Reformen von Bundeskanzler ...“

Gerhard	40%	↓ Temperatur
Schröder	30%	
Helmut	7%	
Kohl	3%	
Schmidt	2%	
...		



Andere Parameter:

- top-k: Beschränke auf erste x Einträge / top-p: Beschränke auf entsprechende Wahrscheinlichkeit
 - frequency penalty, presence penalty
- **Frequency_penalty:** This parameter is used to discourage the model from repeating the same words or phrases too frequently within the generated text. It is a value that is added to the log-probability of a token each time it occurs in the generated text. A higher frequency_penalty value will result in the model being more conservative in its use of repeated tokens.
 - **Presence_penalty:** This parameter is used to encourage the model to include a diverse range of tokens in the generated text. It is a value that is subtracted from the log-probability of a token each time it is generated. A higher presence_penalty value will result in the model being more likely to generate tokens that have not yet been included in the generated text.

LLMs wurden trainiert an ...

- Patentanmeldungen
- Wikipedia
- Nachrichtenseiten
- Wissenschaftliche Publikationen

Bemerkenswert wenig:

- Nutzerforen (außer Bard)
- Bücher

RANK	DOMAIN	CATEGORY	PERCENT OF ALL TOKENS
1	parents.google.com	Law & Government	0.40%
2	wikipedia.org	News & Media	0.35%
3	scribd.com	News & Media	0.07%
4	nytimes.com	News & Media	0.06%
5	pubmed.ncbi.nlm.nih.gov	Science & Health	0.06%
6	latimes.com	News & Media	0.05%
7	theguardian.com	News & Media	0.05%
8	forbes.com	News & Media	0.05%
9	forbes.com	News & Media	0.04%
10	parents.com	Law & Government	0.04%
11	washingtonpost.com	News & Media	0.03%
12	couriers.org	Info & Education	0.03%
13	foia.com	Business & Industrial	0.03%
14	pubmed.org	Science & Health	0.02%
15	researchgate.net	Technology	0.02%

The Post believes it is important to present the complete contents of the data fed into AI models, which promise to govern many aspects of modern life. Some websites in this data set contain highly offensive language and we have attempted to mask these words. Disputable content may remain.

Note: Some websites were unable to be categorized and, in many cases, are no longer accessible.



Washington Post / Common Crawl,

Google Bard: inkl. Foren

Schwächen von LLMs

- Zuverlässigkeit („Halluzinieren“)
- Erklärbarkeit

Für die Praxis:

- Fakten unzuverlässig
- phrasenhaft (orientiert sich an viel geschriebenem)
- inhaltliche Schlagseiten (Vorurteile, politische Positionierung, z.T. bewusste Blockaden)
- Formulierung ausgewogen bis langweilig
- Urheberrecht offen
- Daten- und Quellenschutz



OpenAI's ChatGPT and GPT-4 were the most [left-leaning](#) and libertarian. Google's BERT models were more socially conservative than OpenAI's models, and Meta's LLaMA was the most right-leaning and authoritarian

Die kleine Warnung

Vorsicht mit dem Schutz von Daten:

- Urheberrecht an den Beiträgen
- Daten von Informanten
- Daten des Unternehmens



Bis 16:20

ChatGPT und Bing

- ... unterscheiden sich wenig
- ... ChatGPT kennt Plugins, Bing surft besser in der kostenlosen Version
- ... am besten nutzen, als würde man fix einen Kollegen fragen
- ... um die Kreativität zu steigern

Wir arbeiten im Verlageinsatz auch mit Aleph Alpha



Teil hier bis 16:40

LLMs in der Recherche

- Große Ergebnislisten durchsuchen ("Wo wohnt Boris Becker?"), lässt sich leicht verifizieren
- Ideen für Ansprechpartner
- Ideen für vergessene Aspekte
- Ideen für Interviewfragen
- Ideen für Serienbeiträge
- Später: schnelles Scannen langer Berichte

... und wofür haben Sie es schon benutzt?



LLMs beim Schreiben

- Metaphern suchen
- Überleitungen vorschlagen
- Bessere Formulierungen suchen (DeepL Write!)
- Überschriftenvorschläge

... und wofür haben Sie es schon benutzt?



Der richtige Prompt

- detailliert, Überlegungen vorwegnehmen
- Experten oder andere Personen aufrufen
(Was würde Franz Beckenbauer / Hansi Flick / Schiedsrichter-Chef Lutz Michael Fröhlich sagen)
- Bei kreativen Aufgaben: Antworten unsachlich einschränken
(„Alle Wörter müssen mit S beginnen“)
- Bei Denkaufgaben: „Denke Schritt für Schritt“

... und was hat bei Ihnen gut funktioniert?



Bis 17:00

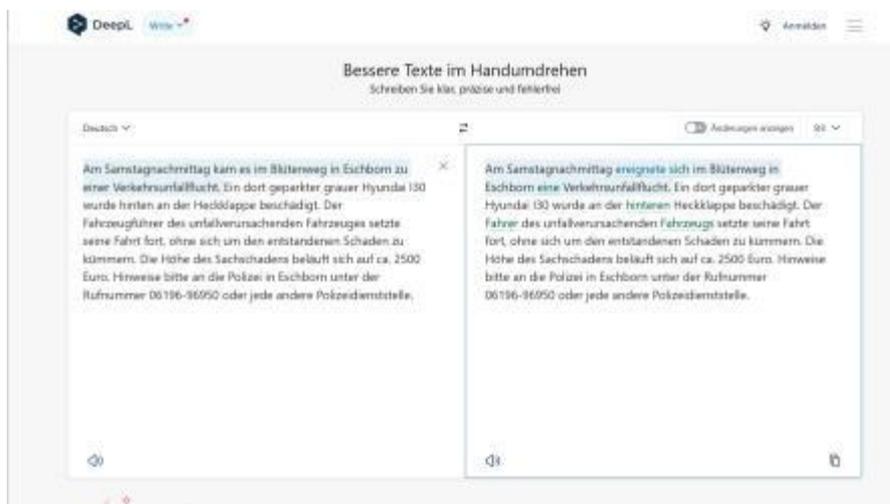
Mehr Werkzeuge

- Sprache
- Transkription
- Dokumente durchsuchen
- Recherchesammlung
- Podcast-Bearbeitung
- Videogenerierung
- Jenni



Teil hier bis 17:00

DeepL Write



Künstliche Intelligenz für Journalisten

16 Reporter:innen-Forum Schweiz — 29. September 2023



Transkription



- Besser: Trint, evtl. Amberscript
- Datenschutz-freundlicher: Whisper

PDFs zusammenfassen: PDF.ai

The screenshot displays the PDF.ai web interface. On the left, a table of contents is visible under the heading 'Inhaltsverzeichnis'. The table lists various sections and their corresponding page numbers. On the right, a summary of an article is shown, titled 'Wie hat sich die Beschäftigung seit der Einführung des Mindestlohns entwickelt?'. The summary discusses the employment development since the introduction of the minimum wage in 2015, noting that while overall employment in 2022 was similar to 2015, the pandemic led to a higher employment rate. It also mentions that the minimum wage had little impact on the reduction of social security obligations, but led to negative effects on the number of low-skilled employees. The summary is attributed to p. 15.

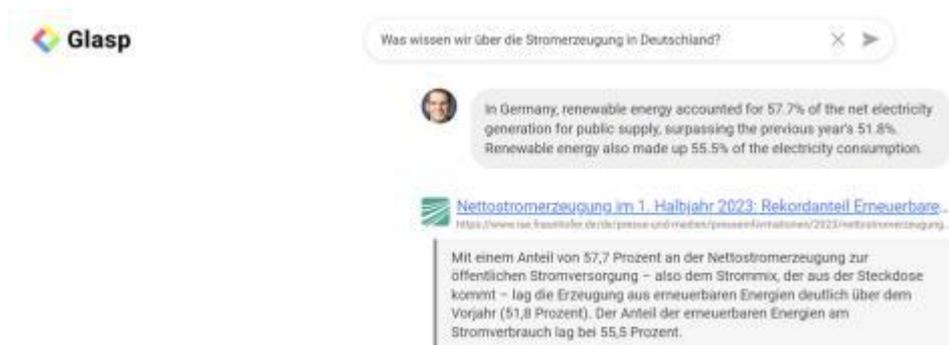
Section	Page
Inhaltsverzeichnis	3
Wort	4
Die Essenz	7
Zusammenfassung	8
1. Ausgangslage und Rahmenbedingungen	22
1.1 Juristische Aufregung der Mindestlohnminister	23
1.2 Wirtschaftlich-ökonomische Hintergründe	24
1.3 Die unterschiedliche Impact in der Beschäftigung	25
1.4 Mindestlohn im internationalen Kontext	26
1.5 Auswirkungen von individuellen Faktoren der Mindestlohnminister	27
1.6 Zusammenfassung	28
1.7 Bestmögliche Handlungsstrategie von Experten der Jahre 2015 bis 2022	29
2. Mindestlohn und Zukunft der Arbeitnehmerinnen und Arbeitnehmer	30
2.1 Mindestlohn	31
2.2 Mindestlohn und soziale Sicherung	32
2.3 Mindestlohn und soziale Sicherung	33
3. Mindestlohn und Beschäftigung	34

Künstliche Intelligenz für Journalisten

18 Reporter:innen-Forum Schweiz — 29. September 2023



Recherchesammlung: glasp.co



Glasp

Was wissen wir über die Stromerzeugung in Deutschland? ✕ ▶

 In Germany, renewable energy accounted for 57.7% of the net electricity generation for public supply, surpassing the previous year's 51.8%. Renewable energy also made up 55.5% of the electricity consumption.

 [Nettostromerzeugung im 1. Halbjahr 2023: Rekordanteil Erneuerbare](https://www.tat.fraunhofer.de/foer/promo-und-medien/pressenotizen/2023/nettostromerzeugung)
https://www.tat.fraunhofer.de/foer/promo-und-medien/pressenotizen/2023/nettostromerzeugung

Mit einem Anteil von 57,7 Prozent an der Nettostromerzeugung zur öffentlichen Stromversorgung – also dem Strommix, der aus der Steckdose kommt – lag die Erzeugung aus erneuerbaren Energien deutlich über dem Vorjahr (51,8 Prozent). Der Anteil der erneuerbaren Energien am Stromverbrauch lag bei 55,5 Prozent.

Podcasts bearbeiten

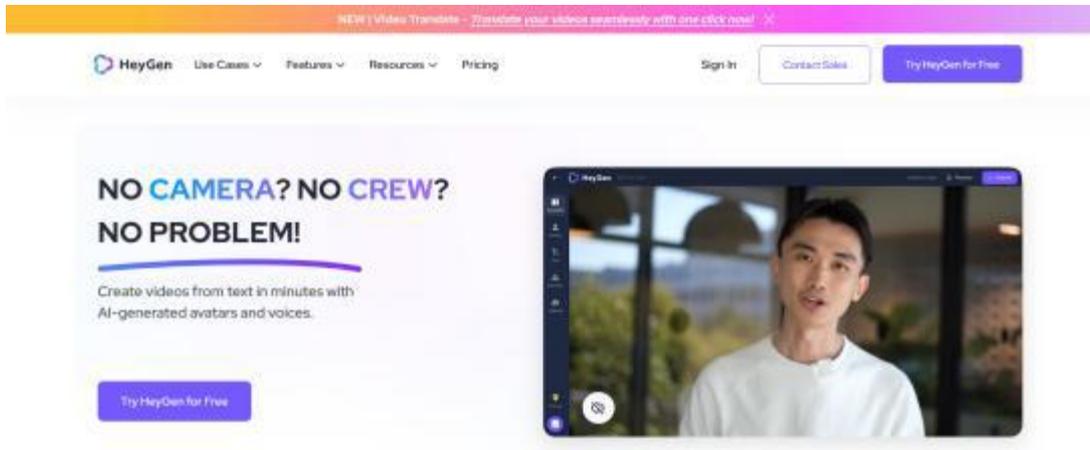


Künstliche Intelligenz für Journalisten

20 Reporter:innen-Forum Schweiz — 29. September 2023



Videos generieren: HeyGen



21 Künstliche Intelligenz für Journalisten
Reporter:innen-Forum Schweiz — 29. September 2023



Beta-Version: Übersetzt und lässt aussprechen, mit Lippenbewegung

Jenni



Viel Spaß damit

... und vielen Dank für das Interesse!

