ParIdigi MasterClass 2023

Grosser Rat Bern

NLP im Justizumfeld

Bern, 7. März 2023

Einführung:

Prof. Dr. Matthias Stürmer, Leiter BFH Institut Public Sector Transformation

Kommentar:

Daniel Brunner, Leiter Informatikdienst Bundesgericht

Eine Veranstaltung von:

Universität Zürich[UZH]

Digital Society Initiative

ParIdigi

Unterstützt durch:

Stiftung Mercator Schweiz

Parldigi Master-Class

Natural Language Processing im Justizumfeld

Prof. Dr. Matthias Stürmer, *Leiter Institut Public Sector Transformation*

Parldigi MasterClass im Grossen Rat Bern
7. März 2023, Staatskanzlei Bern

▸ Institut Public Sector Transformation, Departement Wirtschaft, Berner Fachhochschule

Matthias Stürmer

- Seit 2021 Leiter Institut Public Sector Transformation und Professor an der **Berner Fachhochschule (BFH)**
- Seit 2013 Dozent (Habilitation) an der **Universität Bern** und Leiter Forschungsstelle Digitale Nachhaltigkeit
- 2009 bis 2013 bei **Liip AG** und **EY (Ernst & Young)**
- Doktorat zu Technologie-Management an der **ETH Zürich**
- Studium Wirtschaftsinformatik an der **Universität Bern**

- Präsident Vereine **Digital Impact Network** und **CH Open**
- Gründungsmitglied und Vorstandsmitglied **Opendata.ch**
- Geschäftsleiter Parlamentarische Gruppe Digitale Nachhaltigkeit **Parldigi**
- Mitglied Smart Capital Region **Hauptstadtregion Schweiz**

Prof. Dr. Matthias Stürmer
Leiter Institut Public Sector Transformation

Berner Fachhochschule
Departement Wirtschaft
Brückenstrasse 73
3005 Bern

+41 31 848 41 68
+41 76 368 81 65
matthias.stuermer@bfh.ch
www.bfh.ch/ipst

matthiasstuermer

Institut Public Sector Transformation (IPST)

Unsere Fachgruppen:

1. **Data and Infrastructure**
2. **Digital Democracy**
3. **Public Sector Innovation**
4. **Public Procurement**
5. **Digital Sustainability Lab**

Unsere Aktivitäten:

- ✓ Forschung und Lehre
- ✓ Weiterbildungen
- ✓ Beratung, Studien
- ✓ Technische Realisierung

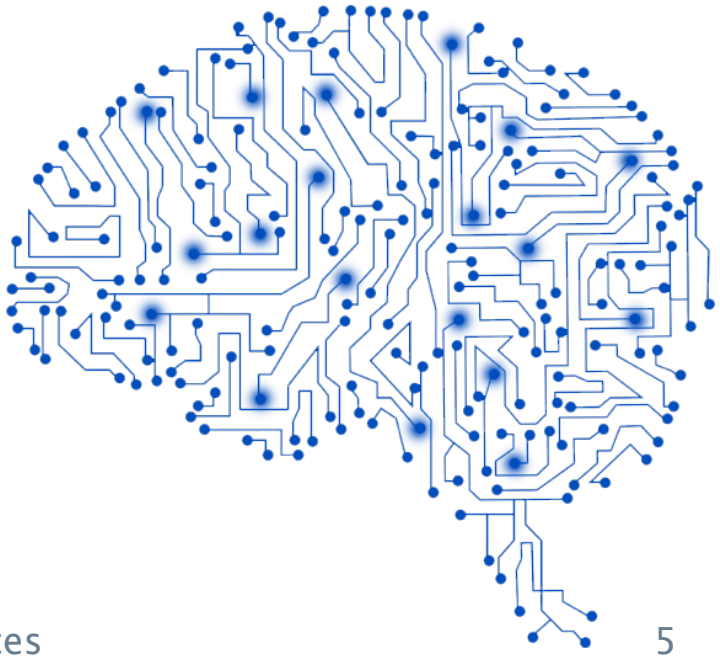
Unsere Themen:

- Demokratische Partizipation
- Innovation und Transformation
- Smart Cities und Smart Regions
- Digitalisierung und Umwelt
- Digitale Nachhaltigkeit und Open Source Software
- Datenökosysteme und Open Government Data
- Datenvisualisierungen und Informationsdesign
- Security und Privacy
- Natural Language Processing (NLP)

https://www.bfh.ch/ipst

Agenda

1. **Einführung Natural Language Processing (NLP)**
2. Transformer-Modelle und GPT-3
3. Forschungsprojekt zu Legal NLP



```
Welcome to
                    EEEEEE  LL      IIII  ZZZZZZZ   AAAAA
                    EE      LL       II       ZZ   AA   AA
                    EEEEE   LL       II     ZZZ    AAAAAAA
                    EE      LL       II    ZZ      AA   AA
                    EEEEEE  LLLLLL  IIII  ZZZZZZZ  AA   AA

Eliza is a mock Rogerian psychotherapist.
The original program was described by Joseph Weizenbaum in 1966.
This implementation by Norbert Landsteiner 2005.


ELIZA: Is something troubling you ?
YOU:   Men are all alike.
ELIZA: What is the connection, do you suppose ?
YOU:   They're always bugging us about something or other.
ELIZA: Can you think of a specific example ?
YOU:   Well, my boyfriend made me come here.
ELIZA: Is it important to you that your boyfriend made you come here ?
YOU:   He says I'm depressed much of the time.
ELIZA: I am sorry to hear that you are depressed.
YOU:   It's true. I am unhappy.
ELIZA: Can you explain what made you unhappy ?
YOU:
```

Verwendung von ChatGPT für Gerichtsentscheid

VICE

MOTHERBOARD
TECH BY VICE

A Judge Just Used ChatGPT to Make a Court Decision

The case is the first time a court has admitted to using the AI text generator's answers in a legal ruling.

By Janus Rose
NEW YORK, US

February 3, 2023, 6:28pm

Rama Judicial
Consejo Superior de la Judicatura
República de Colombia

Radicado No. 13001410500420220045901
Accionante: SALVADOR ESPITIA CHÁVEZ
Accionado: SALUD TOTAL E.P.S
Sentencia No. 032
Fecha de la Providencia: 30-01-2023

estamos clasificados en el grupo VI subgrupo B1 del SISBEN pobreza moderada, circunstancia que acrecienta la necesidad de que se conceda el beneficio de manera inmediata"., no ha sido desvirtuada.

Conforme a lo anterior, las cuotas moderadoras y copagos, para el caso concreto, representan barreras de acceso a los servicios de salud que el Juez Constitucional debe eliminar para que el menor S E C., en las actuales condiciones económicas alegadas por laaccionante.

Por otra parte, atendiendo que la Ley 2213 de 2022 tiene por objeto la incorporación de las TIC en los procesos judiciales., se resolverá extender los argumentos de la decisión adoptada, conforme al uso de inteligencia artificial (IA). Por consiguiente, se procedió a incluir apartes del problema jurídico planteado en la presente sentencia y se ingresó en el aplicativo https://chat.openai.com/chat.

Se hicieron en el aplicativo las siguientes preguntas.

¿**Menor autista esta exonerado de pagar cuotas moderadoras en sus terapias**?

La respuesta del aplicativo IA:

Sí, es correcto. De acuerdo con la normativa en Colombia, los menores con diagnóstico de autismo están exonerados de pagar cuotas moderadoras en sus terapias. Este beneficuento está regulado por la Ley 1753 de 2015 y se aplica a todas las instituciones prestadoras de servicios de salud públicas o privadas que reciban recursos del sistema general de seguridad social en salud.

https://www.vice.com/en/article/k7bdmv/judge-used-chatgpt-to-make-court-decision

Heute gibt ChatGPT vorsichtigere Antwort

Ist ein autistisches Kind in Kolumbien von der Zahlung von Gebühren für die Therapie befreit?

Als KI-Modell verfüge ich über Informationen, die bis September 2021 bekannt waren. Ich kann jedoch sagen, dass ich keine spezifischen Informationen darüber habe, ob autistische Kinder in Kolumbien von der Zahlung von Gebühren für Therapie befreit sind.

Es ist möglich, dass es in einigen Städten oder Regionen spezifische Programme oder Richtlinien gibt, die dies ermöglichen. Es ist jedoch ratsam, sich an lokale Regierungsbehörden oder Autismusorganisationen in Kolumbien zu wenden, um Informationen zu erhalten, ob es spezifische Richtlinien gibt, die die Kosten für die Therapie von autistischen Kindern abdecken.

https://chat.openai.com

Politische Vorstösse zu «ChatGPT»

- Nationalrat sowie Kantone Waadt, Genf, Baselland, Fribourg und Neuchatel:

Geschäft	Titel	Eingereicht von	Art des Geschäftes	Datum Einreichung	Datum Aktualisierung	Status
CH 22.1074	Auswirkungen von ChatGPT und anderen General Purpose Artificial Intelligence Systems	Mäder Jörg, Grünliberale Fraktion	Anfrage	15.12.2022	15.02.2023	Erledigt
VD 23_INT_6	Quelle stratégie face à l'arrivée de ChatGPT et au développement rapide des Intelligences Artificielles ?	Théophile Schenker, VER	Interpellation	17.01.2023	24.01.2023	Überwiesen
GE QUE 1863	ChatGPT et travaux de fin d'études au sein de l'enseignement secondaire II : quelle conciliation ?	Natacha Buffet-Desfayes, PLR	Question urgente écrite	26.01.2023	27.01.2023	Überwiesen
BL 2023/70	Umgang mit ChatGPT an Baselbieter Schulen	Jan Kirchmayr, SP	Postulat	26.01.2023	16.03.2023	Traktandiert
FR 2023-GC-21	Wie geht der Kanton Freiburg mit KI (Künstlicher Intelligenz) um?	Laurent Baeriswyl, Die Mitte	Anfrage	30.01.2023	30.01.2023	Überwiesen
NE 23.157	ChatGPT, attention à ne pas louper le train	Émile Blant, Keine Partei	Interpellation	08.02.2023	08.02.2023	Eingereicht

https://prod.politmonitor.ch

Natural Language Processing (NLP)

- **NLP:** Natural Language Processing
- **Interdisziplinäre Kombination** von Sprachwissenschaft und Informatik → **Computerlinguistik**
- Realisierung von NLP erfolgt über **«Tasks»,** kleinere oder grössere Arbeitsschritte

Einfache NLP Tasks

- Tokenization
- Stemming
- Lemmatization
- Part of Speech Tagging
- Topic Modeling / Text Classification
- etc.

Schwierige NLP Tasks

- Namend Entity Recognition
- Information Retrieval
- Relationship Extraction
- Sentiment Analysis
- Text Similarity
- Coreference Resolution
- etc.

Sehr schwierige NLP Tasks

- Fragen-Antworten
- Automatische Zusammenfassungen
- Text-Vereinfachung
- Text-Generierung
- Maschinelle Übersetzungen
- Chatbots / Dialog-Systeme
- etc.

https://nlpprogress.com

Beispiel «Fragen-Antworten»

GPT | Legal | Question Answering | Summarization | Simplification | Named Entity Recognition | Zero-Shot-Klassifizierung

Gelelctra

Dieses deutschsprachige Modell wurde darauf spezialisiert, in einem Text die Antwort auf eine gestellte Frage zu finden. Es basiert auf XLM-RoBERTa, einem mehrsprachigen Modell mit 270 Millionen Parameter.

Context

Grosser Rat behandelt in den nächsten Tagen über 100 Traktanden
Mehrere Vorstösse und umstrittene Kredite diskutiert der Grosse Rat in Bern in der Frühlingssession.
Nur gerade zwei Gesetzesvorlagen beschäftigen den bernischen Grossen Rat in der Frühlingssession ab kommenden Montag. Die Arbeit geht dem Kantonsparlament aber nicht aus. Über 40 Vorstösse stehen auf der Traktandenliste, dazu mehrere umstrittene Kreditgeschäfte.
In der Wintersession hatte der Grosse Rat eine Solaroffensive eingeleitet. Er überwies eine Reihe von Vorstössen, um den Ausbau der Solarenergie zu beschleunigen. Nun schlägt seine vorberatende Finanzkommission vor, früher als geplant auch Steuervorteile für Photovoltaik- und Solarthermieanlagen einzuführen.
Schon ab 2023 sollen sie von der amtlichen Bewertung ausgenommen werden. Die Kommission möchte so bei der zweiten Lesung der Steuergesetzrevision ein Zeichen setzen: Die Anstrengungen der Bernerinnen und Berner für die Energiewende sollten rasch und spürbar unterstützt werden.
Von der zweiten Neuerung profitieren die Eltern, die ihre Kinder zum Beispiel in einer Kita betreuen lassen. Ebenfalls bereits für das Steuerjahr 2023 soll der Drittbetreuungsabzug auf 16'000 Franken erhöht werden. So soll die Vereinbarkeit von Familien und Arbeit verbessert werden.
Keine grossen Wellen dürfte die zweite Lesung des Lufthygienegesetzes werfen. Hauseigentümer sollen selber entscheiden, wer ihre Öl- oder Gasheizung kontrolliert.

Question

Wie viele Gesetze berät der Grossrat in der Frühlingssession?

Löschen | Absenden

Answer

zwei

Score

0.5884820818901062

Aus den Regionen

Autobahn, Kitas, Solaroffensive

Grosser Rat behandelt in den nächsten Tagen über 100 Traktanden

Mehrere Vorstösse und umstrittene Kredite diskutiert der Grosse Rat in Bern in der Frühlingssession.

Publiziert: 01.03.2023, 09:57

https://digital-sustainability.github.io/nlp-demo

Beispiel «Text-Zusammenfassung»

GPT | Legal | Question Answering | **Summarization** | Simplification | Named Entity Recognition | Zero-Shot-Klassifizierung

T5-Base

Hier wurde ein sogenanntes T5-Sprachmodell auf den Task der Text-Zusammenfassung spezialisiert. Das Modell versteht nur Deutsch und hat 220 Millionen Parameter.

Input

Grosser Rat behandelt in den nächsten Tagen über 100 Traktanden
Mehrere Vorstösse und umstrittene Kredite diskutiert der Grosse Rat in Bern in der Frühlingssession.
Nur gerade zwei Gesetzesvorlagen beschäftigen den bernischen Grossen Rat in der Frühlingssession ab kommenden Montag. Die Arbeit geht dem Kantonsparlament aber nicht aus. Über 40 Vorstösse stehen auf der Traktandenliste, dazu mehrere umstrittene Kreditgeschäfte.
In der Wintersession hatte der Grosse Rat eine Solaroffensive eingeleitet. Er überwies eine Reihe von Vorstössen, um den Ausbau der Solarenergie zu beschleunigen. Nun schlägt seine vorberatende Finanzkommission vor, früher als geplant auch Steuervorteile für Photovoltaik- und Solarthermieanlagen einzuführen.
Schon ab 2023 sollen sie von der amtlichen Bewertung ausgenommen werden. Die Kommission möchte so bei der zweiten Lesung der Steuergesetzrevision ein Zeichen setzen: Die Anstrengungen der Bernerinnen und Berner für die Energiewende sollten rasch und spürbar unterstützt werden.
Von der zweiten Neuerung profitieren die Eltern, die ihre Kinder zum Beispiel in einer Kita betreuen lassen. Ebenfalls bereits für das Steuerjahr 2023 soll der Drittbetreuungsabzug auf 16'000 Franken erhöht werden. So soll die Vereinbarkeit von Familien und Arbeit verbessert werden.
Keine grossen Wellen dürfte die zweite Lesung des Lufthygienegesetzes werfen. Hauseigentümer sollen selber entscheiden, wer ihre Öl- oder Gasheizung kontrolliert.

Summary

Der bernische Grosse Rat diskutiert in der Frühjahrssession über mehrere Vorstösse und umstrittene Kreditgeschäfte. Die Arbeit geht dem Kantonsparlament aber nicht aus.

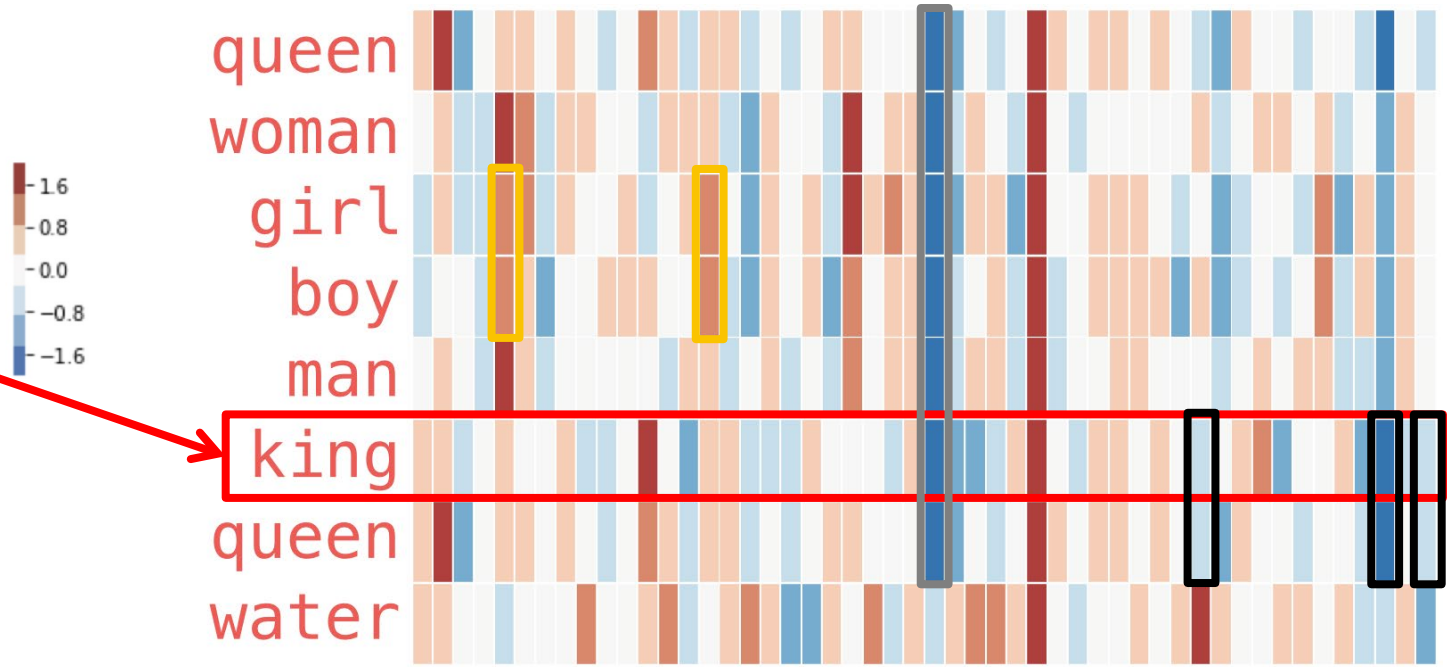
In der Frühlingssession des Grossen Rats in Bern werden über 100 Traktanden behandelt, darunter mehrere umstrittene Kredite und über 40 Vorstösse. Die vorberatende Finanzkommission schlägt vor, Steuervorteile für Solaranlagen früher als geplant einzuführen, sowie den Drittbetreuungsabzug zu erhöhen, um die Vereinbarkeit von Familien und Arbeit zu verbessern. Auch das Lufthygienegesetz steht zur zweiten Lesung an, wobei Hauseigentümer selber entscheiden sollen, wer ihre Öl- oder Gasheizung kontrolliert.

Word Embeddings

50 Vektoren des Worts **«king»** (basierend auf GloVe Vector, trainiert auf Wikipedia):

[0.50451 , 0.68607 , -0.59517 , -0.022801, 0.60046 , -0.13498 , -0.08813 , 0.47377 , -0.61798 , -0.31012 , -0.076666, 1.493 , -0.034189, -0.98173 , 0.68229 , 0.81722 , -0.51874 , -0.31503 , -0.55809 , 0.66421 , 0.1961 , -0.13495 , -0.11476 , -0.30344 , 0.41177 , -2.223 , -1.0756 , -1.0783 , -0.34354 , 0.33505 , 1.9927 , -0.04234 , -0.64319 , 0.71125 , 0.49159 , 0.16754 , 0.34344 , -0.25663 , -0.8523 , 0.1661 , 0.40102 , 1.1685 , -1.0137 , -0.21585 , -0.15155 , 0.78321 , -0.91241 , -1.6106 , -0.64426 , -0.51042]

Vektorisierte Wörter ermöglichen **Vergleiche** und **Mustererkennung:**



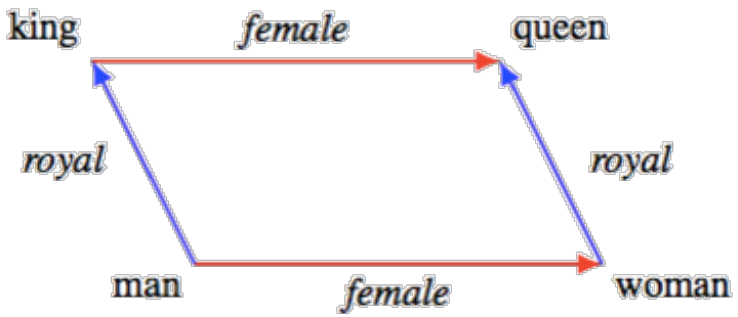
https://jalammar.github.io/illustrated-word2vec/

Berechnungen mit Wort-Vektoren

king – man + woman = ?

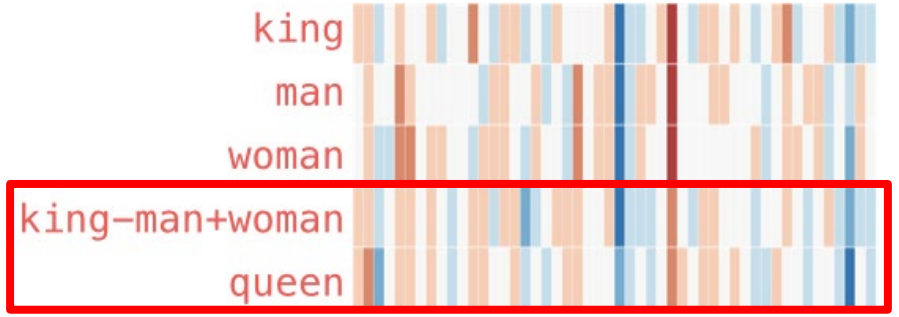
Berechnungen mit Wort-Vektoren

king – man + woman ≈ **queen**



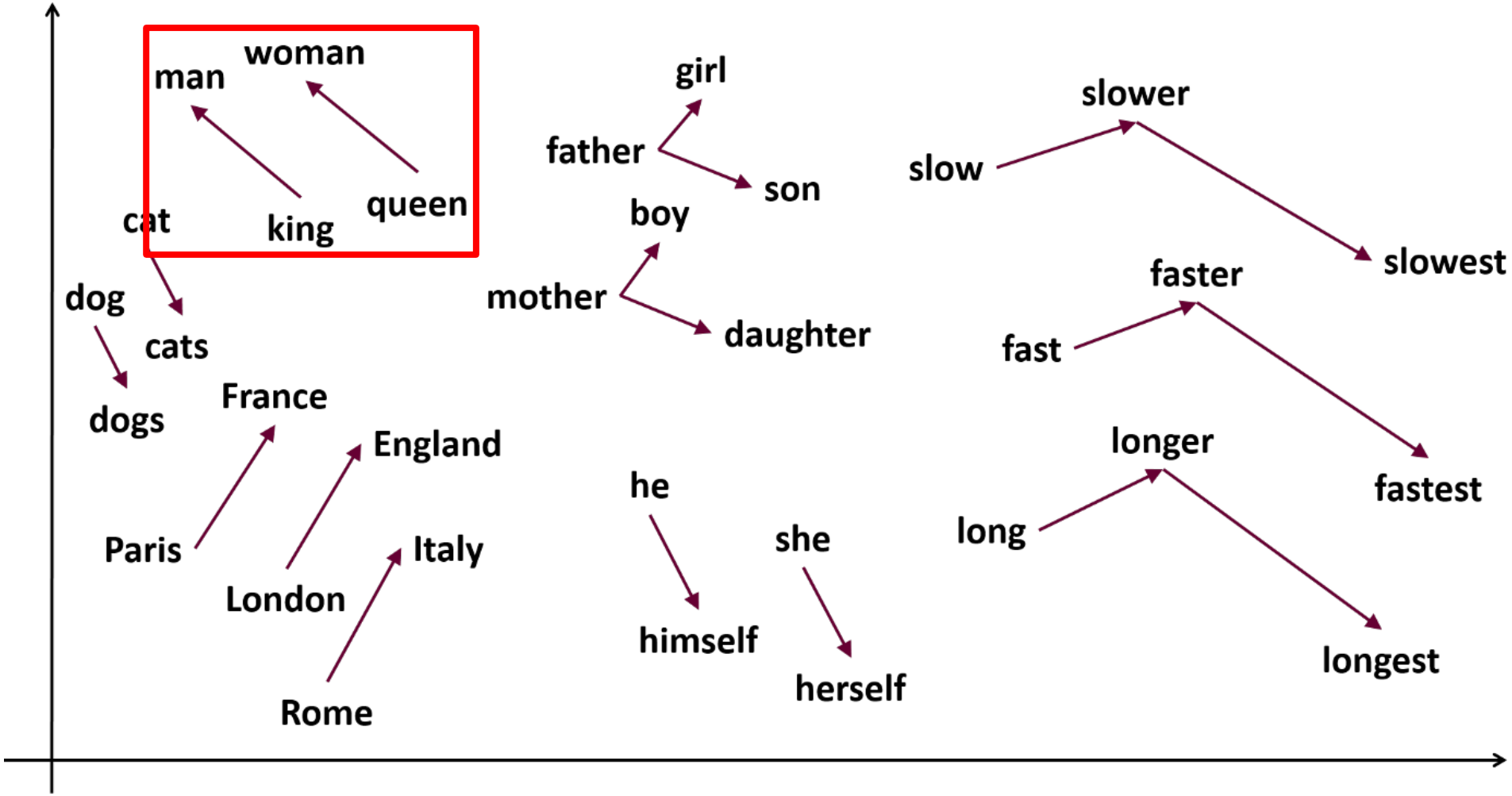

```
model.most_similar(positive=["king","woman"], negative=["man"])
```

```
[('queen', 0.8523603677749634),
 ('throne', 0.7664333581924438),
 ('prince', 0.7592144012451172),
 ('daughter', 0.7473883032798767),
 ('elizabeth', 0.7460219860076904),
 ('princess', 0.7424570322036743),
 ('kingdom', 0.7337411642074585),
 ('monarch', 0.721449077129364),
 ('eldest', 0.7184862494468689),
 ('widow', 0.7099430561065674)]
```



https://kawine.github.io/blog/nlp/2019/06/21/word-analogies.html https://jalammar.github.io/illustrated-word2vec/

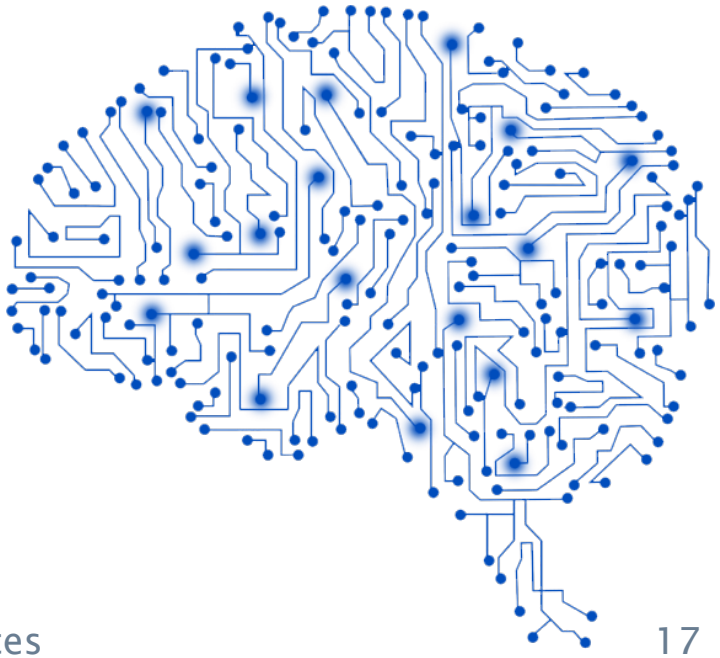
Word2vec (seit 2013)



https://medium.com/analytics-vidhya/implementing-word2vec-in-tensorflow-44f93cf2665f

Agenda

1. Einführung zu Natural Language Processing (NLP)
2. **Transformer-Modelle und GPT-3**
3. Forschungsprojekt zu Legal NLP

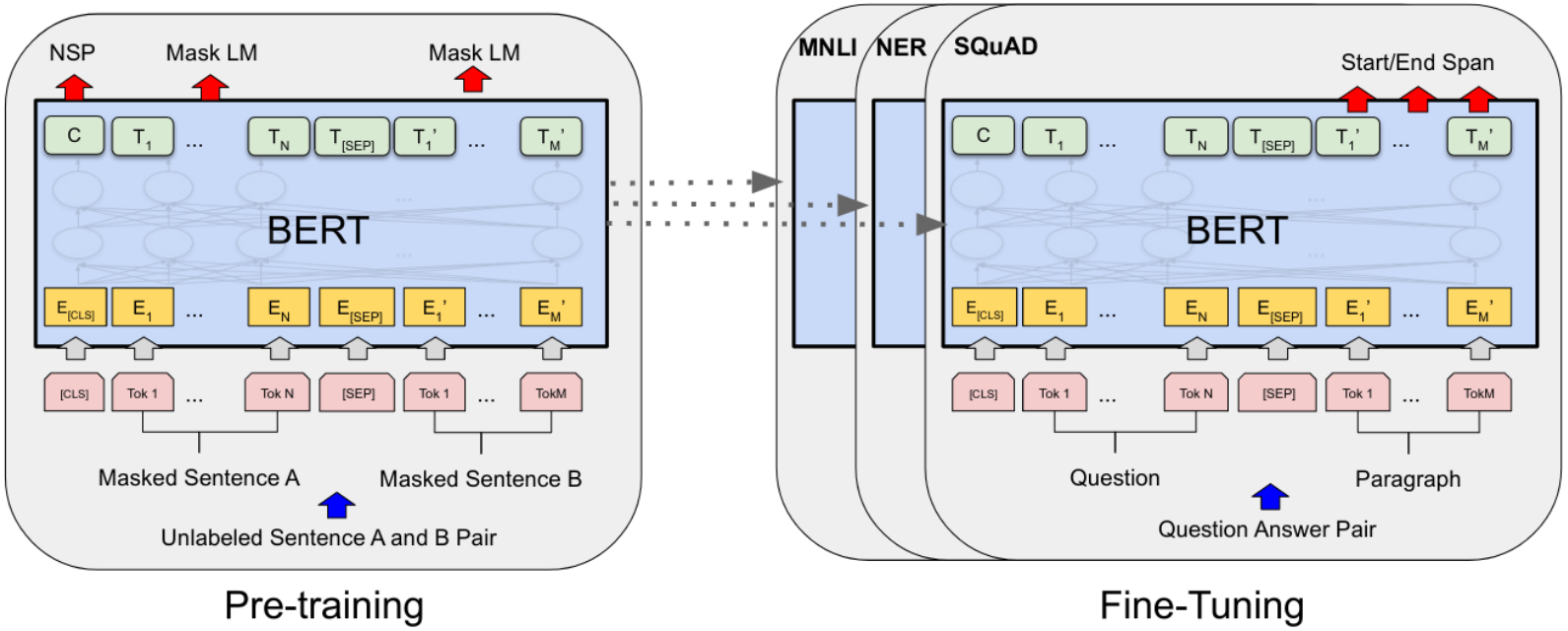


Revolutionäre NLP Technologie BERT (seit 2018)

BERT: Bidirectional **E**ncoder **R**epresentations from **T**ransformers

Schon über 60'000 mal zitiert!

Bert: Pre-training of deep bidirectional transformers for language understanding
J Devlin, MW Chang, K Lee, K Toutanova - arXiv preprint arXiv …, 2018 - arxiv.org
We introduce a new language representation model called BERT, which stands for Bidirectional Encoder Representations from Transformers. Unlike recent language representation models, BERT is designed to pre-train deep bidirectional representations from unlabeled text by jointly conditioning on both left and right context in all layers. As a result, the pre-trained BERT model can be fine-tuned with just one additional output layer to create state-of-the-art models for a wide range of tasks, such as question answering and …
★ Speichern 99 Zitieren Zitiert von: 60775 Ähnliche Artikel Alle 39 Versionen »



BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding

Jacob Devlin Ming-Wei Chang Kenton Lee Kristina Toutanova
Google AI Language
{jacobdevlin,mingweichang,kentonl,kristout}@google.com

arXiv:1810.04805v2 [cs.CL] 24 May 2019

Abstract

We introduce a new language representation model called **BERT**, which stands for **B**idirectional **E**ncoder **R**epresentations from **T**ransformers. Unlike recent language representation models (Peters et al., 2018a; Radford et al., 2018), BERT is designed to pre-train deep bidirectional representations from unlabeled text by jointly conditioning on both left and right context in all layers. As a result, the pre-trained BERT model can be fine-tuned with just one additional output layer to create state-of-the-art models for a wide range of tasks, such as question answering and language inference, without substantial task-specific architecture modifications.

BERT is conceptually simple and empirically powerful. It obtains new state-of-the-art results on eleven natural language processing tasks, including pushing the GLUE score to 80.5% (7.7% point absolute improvement), MultiNLI accuracy to 86.7% (4.6% absolute improvement), SQuAD v1.1 question answering Test F1 to 93.2 (1.5 point absolute improvement) and SQuAD v2.0 Test F1 to 83.1 (5.1 point absolute improvement).

1 Introduction

Language model pre-training has been shown to be effective for improving many natural language processing tasks (Dai and Le, 2015; Peters et al., 2018a; Radford et al., 2018; Howard and Ruder, 2018). These include sentence-level tasks such as natural language inference (Bowman et al., 2015; Williams et al., 2018) and paraphrasing (Dolan and Brockett, 2005), which aim to predict the relationships between sentences by analyzing them holistically, as well as token-level tasks such as named entity recognition and question answering, where models are required to produce fine-grained output at the token level (Tjong Kim Sang and De Meulder, 2003; Rajpurkar et al., 2016).

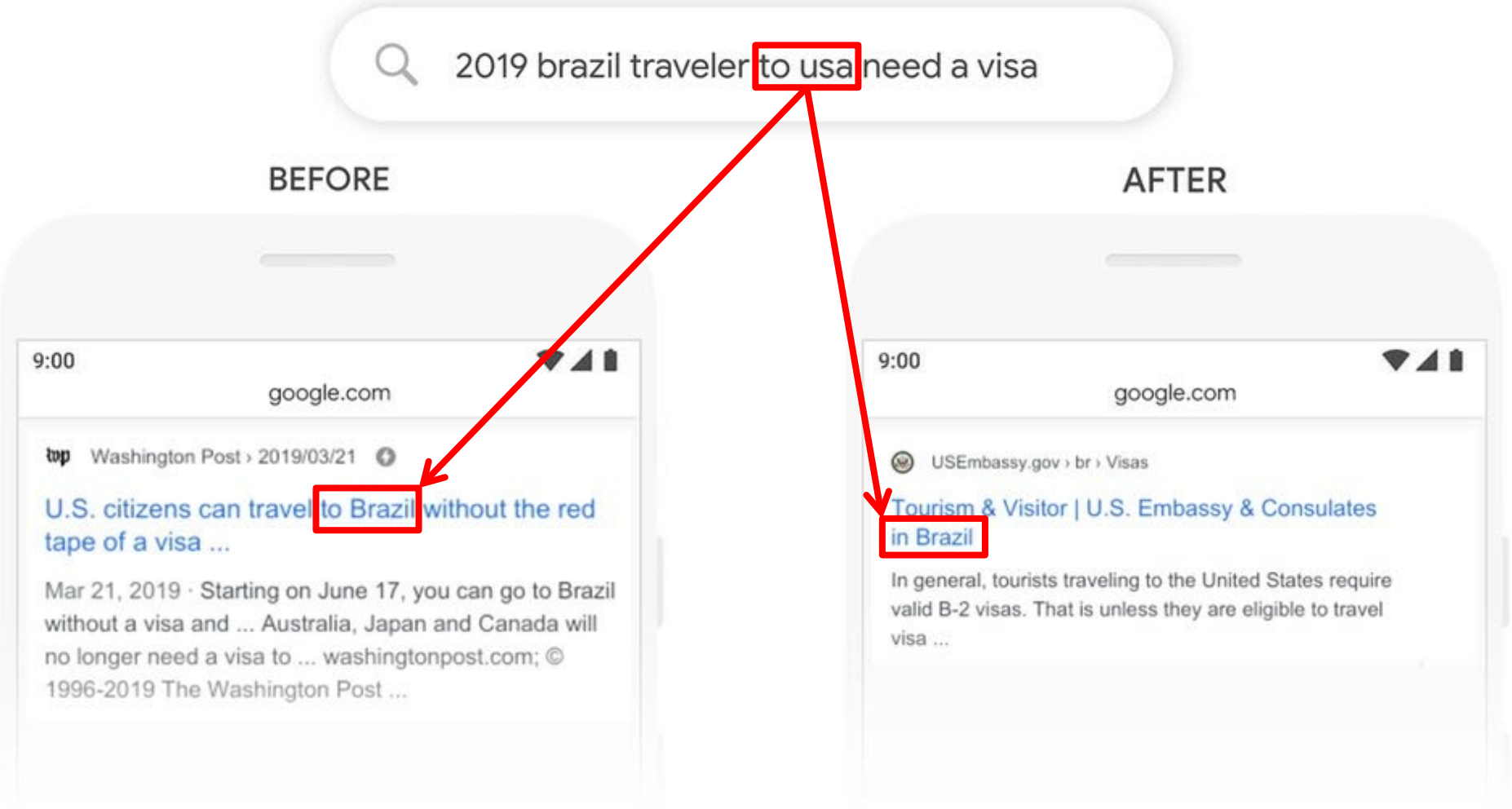
There are two existing strategies for applying pre-trained language representations to downstream tasks: *feature-based* and *fine-tuning*. The feature-based approach, such as ELMo (Peters et al., 2018a), uses task-specific architectures that include the pre-trained representations as additional features. The fine-tuning approach, such as the Generative Pre-trained Transformer (OpenAI GPT) (Radford et al., 2018), introduces minimal task-specific parameters, and is trained on the downstream tasks by simply fine-tuning *all* pre-trained parameters. The two approaches share the same objective function during pre-training, where they use unidirectional language models to learn general language representations.

We argue that current techniques restrict the power of the pre-trained representations, especially for the fine-tuning approaches. The major limitation is that standard language models are unidirectional, and this limits the choice of architectures that can be used during pre-training. For example, in OpenAI GPT, the authors use a left-to-right architecture, where every token can only attend to previous tokens in the self-attention layers of the Transformer (Vaswani et al., 2017). Such restrictions are sub-optimal for sentence-level tasks, and could be very harmful when applying fine-tuning based approaches to token-level tasks such as question answering, where it is crucial to incorporate context from both directions.

In this paper, we improve the fine-tuning based approaches by proposing BERT: **B**idirectional **E**ncoder **R**epresentations from **T**ransformers. BERT alleviates the previously mentioned unidirectionality constraint by using a "masked language model" (MLM) pre-training objective, inspired by the Cloze task (Taylor, 1953). The masked language model randomly masks some of the tokens from the input, and the objective is to predict the original vocabulary id of the masked

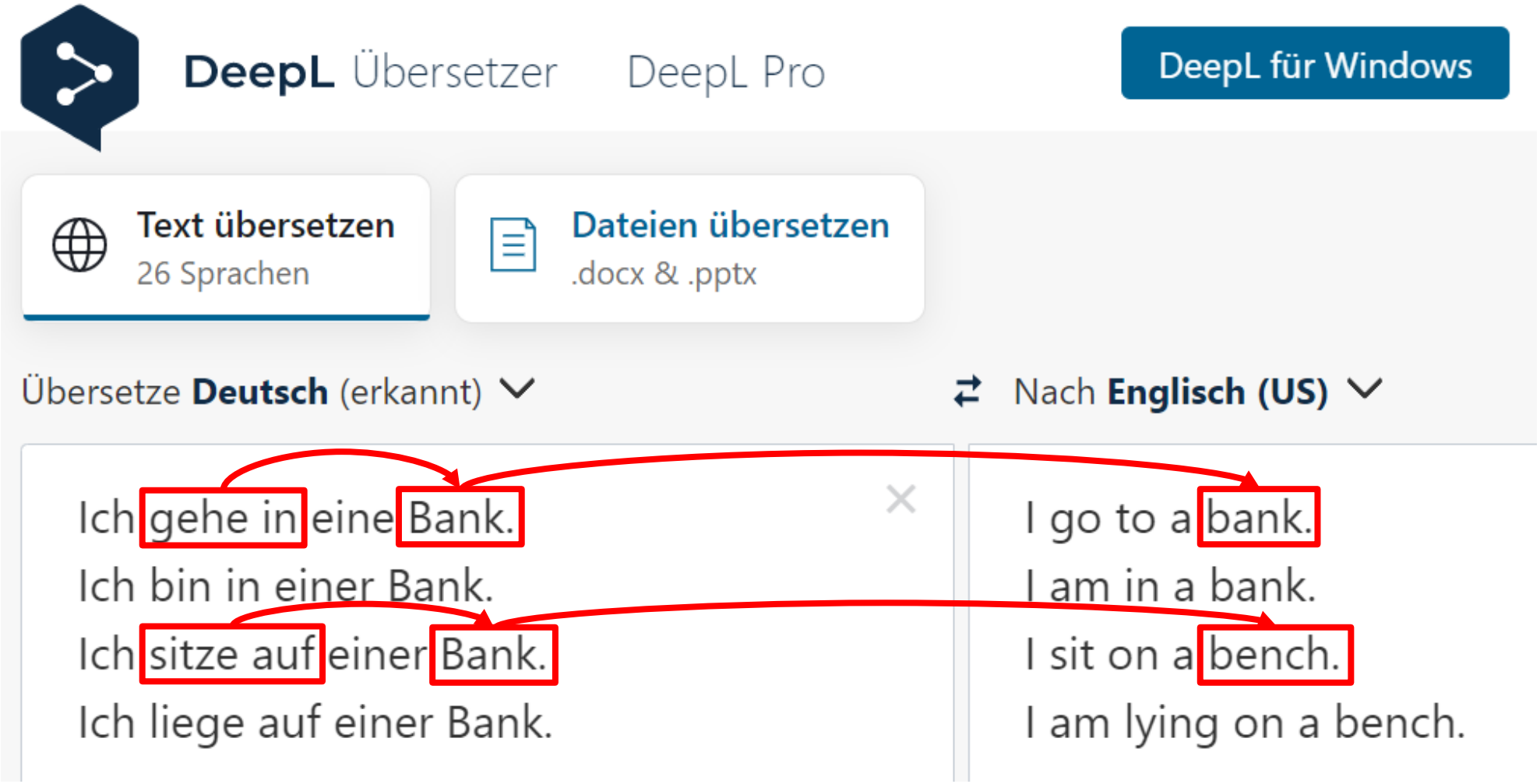
Devlin, J., Chang, M.-W., Lee, K. and Toutanova, K. (2018). BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding, http://arxiv.org/abs/1810.04805 https://ai.googleblog.com/2018/11/open-sourcing-bert-state-of-art-pre.html

Google nutzt BERT für die Internet-Suche seit 2019



https://blog.google/products/search/search-language-understanding-bert

Auch DeepL ‘versteht’ die Bedeutung von Wörtern im Kontext



https://www.deepl.com/translator#de/en/Ich%20gehe%20in%20eine%20Bank.%0AIch%20bin%20in%20einer%20Bank.%0AIch%20sitze%20auf%20einer%20Bank.%0AIch%20liege%20auf%20einer%20Bank

Grundlage von ChatGPT: GPT-3

- **GPT:** Generative Pretrained Transformer
- **GPT-3** wurde 2020 'veröffentlicht', siehe Publikation →
- **Nur als API-Schnittstelle** freigegen, das eigentliche Sprachmodell und die Trainingsdaten sind **proprietär**
- Auf **570 GB Textdaten** trainiert (400 Mrd. Wörter)
- Insgesamt **175 Mrd.** Parameter («Neuronen»)
- GPT-3 kann für **viele NLP Tasks** verwendet werden:
  - Translation
  - Question-Answering
  - Reading Comprehension
  - Text Generation
- Nur wenig oder kein **Fine-Tuning** nötig

Language Models are Few-Shot Learners

Tom B. Brown* Benjamin Mann* Nick Ryder* Melanie Subbiah*

Jared Kaplan† Prafulla Dhariwal Arvind Neelakantan Pranav Shyam Girish Sastry

Amanda Askell Sandhini Agarwal Ariel Herbert-Voss Gretchen Krueger Tom Henighan

Rewon Child Aditya Ramesh Daniel M. Ziegler Jeffrey Wu Clemens Winter

Christopher Hesse Mark Chen Eric Sigler Mateusz Litwin Scott Gray

Benjamin Chess Jack Clark Christopher Berner

Sam McCandlish Alec Radford Ilya Sutskever Dario Amodei

OpenAI

Abstract

Recent work has demonstrated substantial gains on many NLP tasks and benchmarks by pre-training on a large corpus of text followed by fine-tuning on a specific task. While typically task-agnostic in architecture, this method still requires task-specific fine-tuning datasets of thousands or tens of thousands of examples. By contrast, humans can generally perform a new language task from only a few examples or from simple instructions – something which current NLP systems still largely struggle to do. Here we show that scaling up language models greatly improves task-agnostic, few-shot performance, sometimes even reaching competitiveness with prior state-of-the-art fine-tuning approaches. Specifically, we train GPT-3, an autoregressive language model with 175 billion parameters, 10x more than any previous non-sparse language model, and test its performance in the few-shot setting. For all tasks, GPT-3 is applied without any gradient updates or fine-tuning, with tasks and few-shot demonstrations specified purely via text interaction with the model. GPT-3 achieves strong performance on many NLP datasets, including translation, question-answering, and cloze tasks, as well as several tasks that require on-the-fly reasoning or domain adaptation, such as unscrambling words, using a novel word in a sentence, or performing 3-digit arithmetic. At the same time, we also identify some datasets where GPT-3's few-shot learning still struggles, as well as some datasets where GPT-3 faces methodological issues related to training on large web corpora. Finally, we find that GPT-3 can generate samples of news articles which human evaluators have difficulty distinguishing from articles written by humans. We discuss broader societal impacts of this finding and of GPT-3 in general.

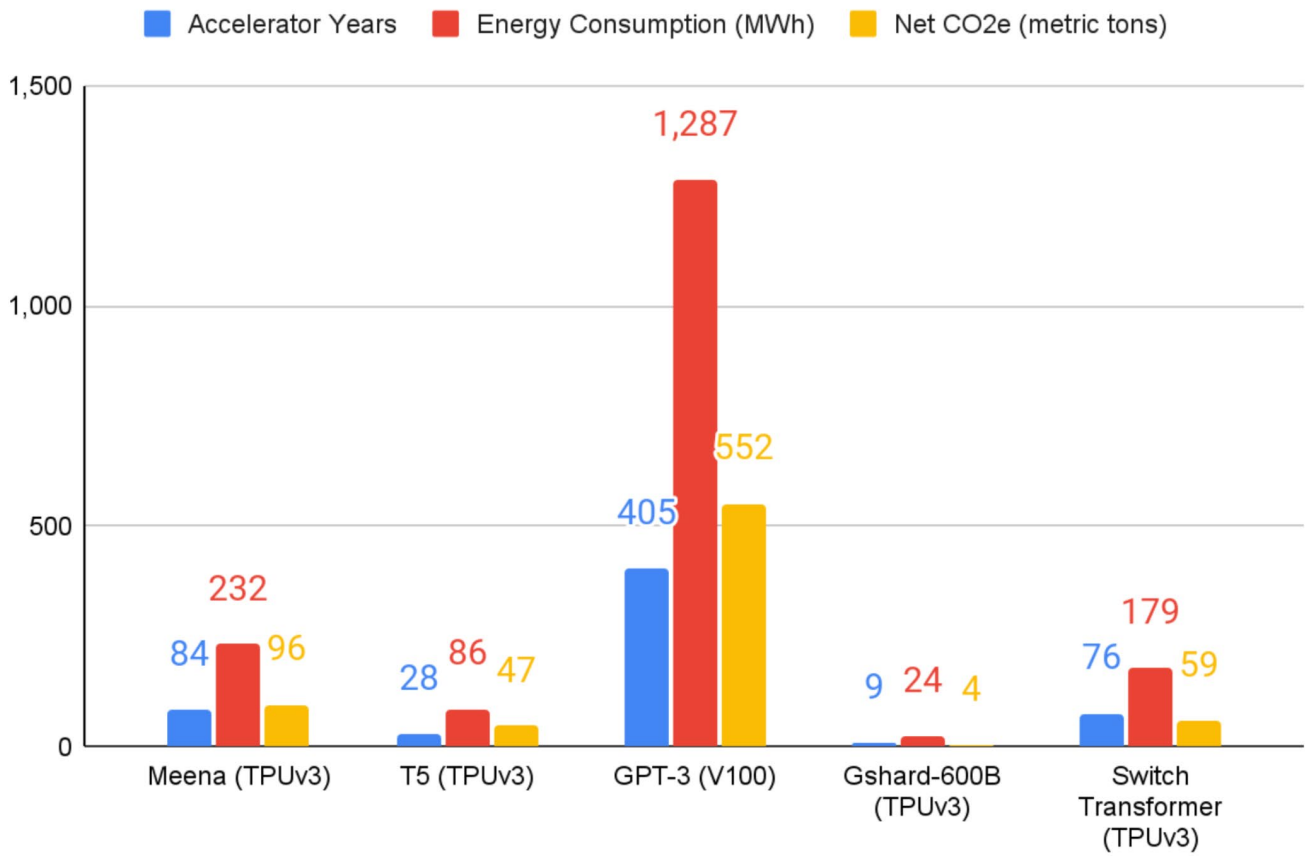
*Equal contribution
†Johns Hopkins University, OpenAI

Author contributions listed at end of paper.

Energieaufwand für Sprachmodell-Training ist immens

Energieaufwand für **GPT-3:**

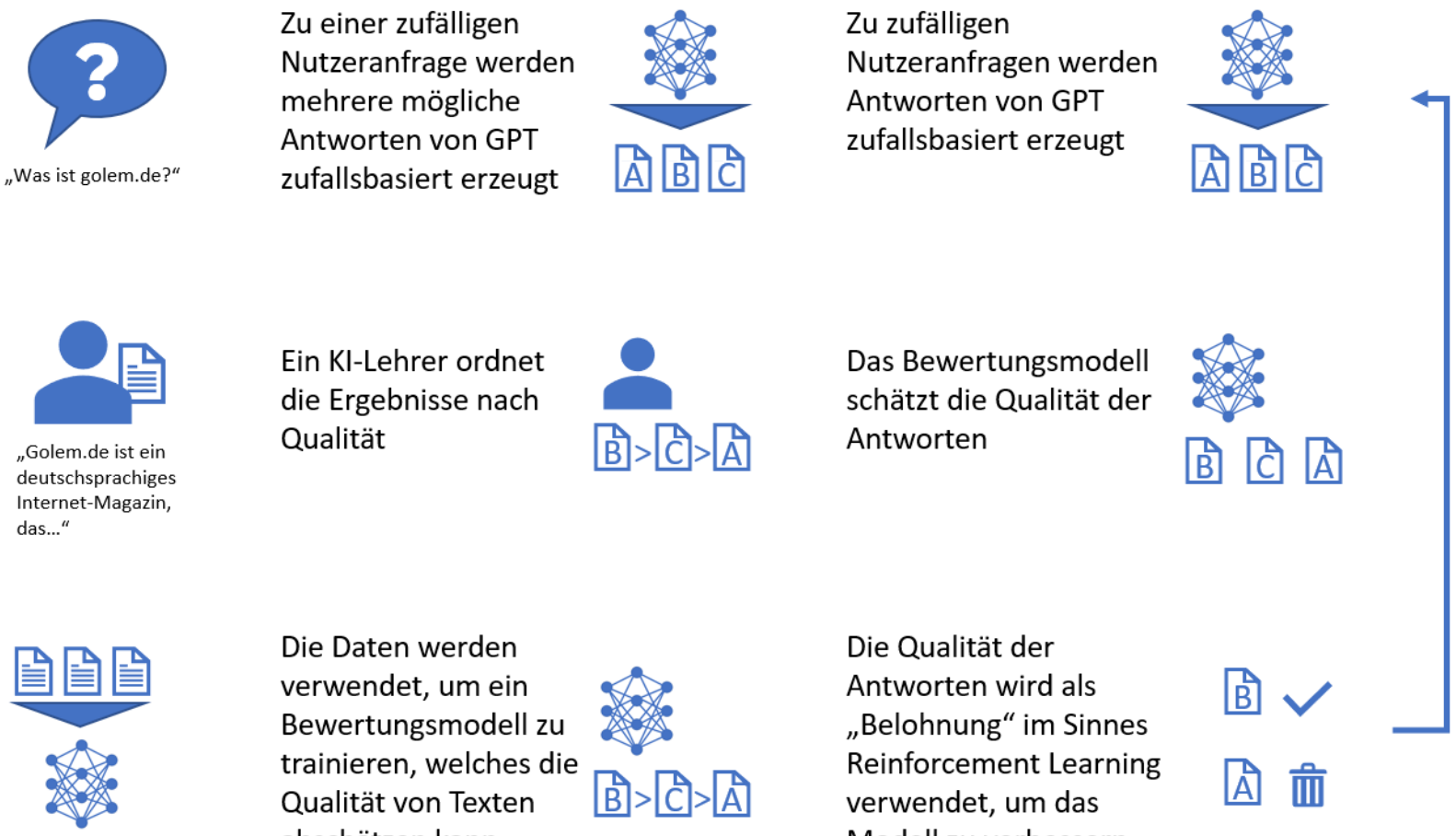
- Dauer von **405 Jahren GPU-Zeit**
- Verbrauch von **1287 MWh** Elektrizität
- Verursachte Footprint von ca. **552 Tonnen CO_2** (tCO2e)



Patterson et al. 2021 “Carbon Emissions and Large Neural Network Training”

Wie funktioniert ChatGPT?

Schritt 1	Schritt 2	Schritt 3
Eine Nutzeranfrage aus der Vergangenheit wird zufällig ausgewählt	Zu einer zufälligen Nutzeranfrage werden mehrere mögliche Antworten von GPT zufallsbasiert erzeugt	Zu zufälligen Nutzeranfragen werden Antworten von GPT zufallsbasiert erzeugt
Ein „KI-Lehrer“ schreibt ein Beispiel für eine gute Antwort	Ein KI-Lehrer ordnet die Ergebnisse nach Qualität	Das Bewertungsmodell schätzt die Qualität der Antworten
Die manuell erzeugten Daten werden zum Modelltraining verwendet	Die Daten werden verwendet, um ein Bewertungsmodell zu trainieren, welches die Qualität von Texten abschätzen kann	Die Qualität der Antworten wird als „Belohnung“ im Sinnes Reinforcement Learning verwendet, um das Modell zu verbessern



https://www.golem.de/news/kuenstliche-intelligenz-so-funktioniert-chatgpt-2302-171644-2.html

Ist ChatGPT wirklich künstliche Intelligenz?

Neue Zürcher Zeitung

INTERVIEW

«Es ist absurd, bei Chat-GPT von künstlicher Intelligenz zu sprechen»

Die Computerlinguistin Emily M. Bender warnte schon vor Jahren davor, Chat-Programme zur Web-Suche zu verwenden. Im Interview erklärt sie, was der derzeitige KI-Hype mit Science-Fiction-Phantasien zu tun hat.

Ruth Fulterer
25.02.2023, 05.30 Uhr

Hören Merken Drucken Teilen

PD

Warum sollten wir versuchen, Computer zu bauen, die denken und fühlen wie wir?

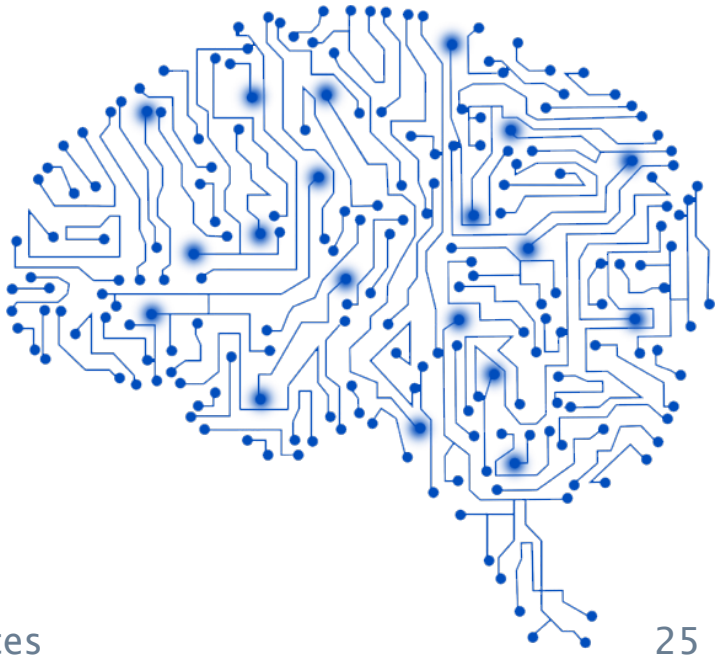
Edwin Koo / Bloomberg

Emily Bender – Professorin an der University of Washington
Emily M. Bender hat im Jahr 2000 an der Universität Stanford in Linguistik promoviert. Nach Forschungsaufenthalten und Erfahrungen im Privatsektor wechselte sie 2003 an die University of Washington. Dort leitet Bender das Forschungslabor für Computerlinguistik. Ausserdem ist sie Vizepräsidentin der Gesellschaft für Computerlinguistik.

https://www.nzz.ch/technologie/es-ist-absurd-bei-chat-gpt-von-kuenstlicher-intelligenz-zu-sprechen-ld.1726924

Agenda

1. Einführung zu Natural Language Processing (NLP)
2. Transformer-Modelle und GPT-3
3. **Forschungsprojekt zu Legal NLP**

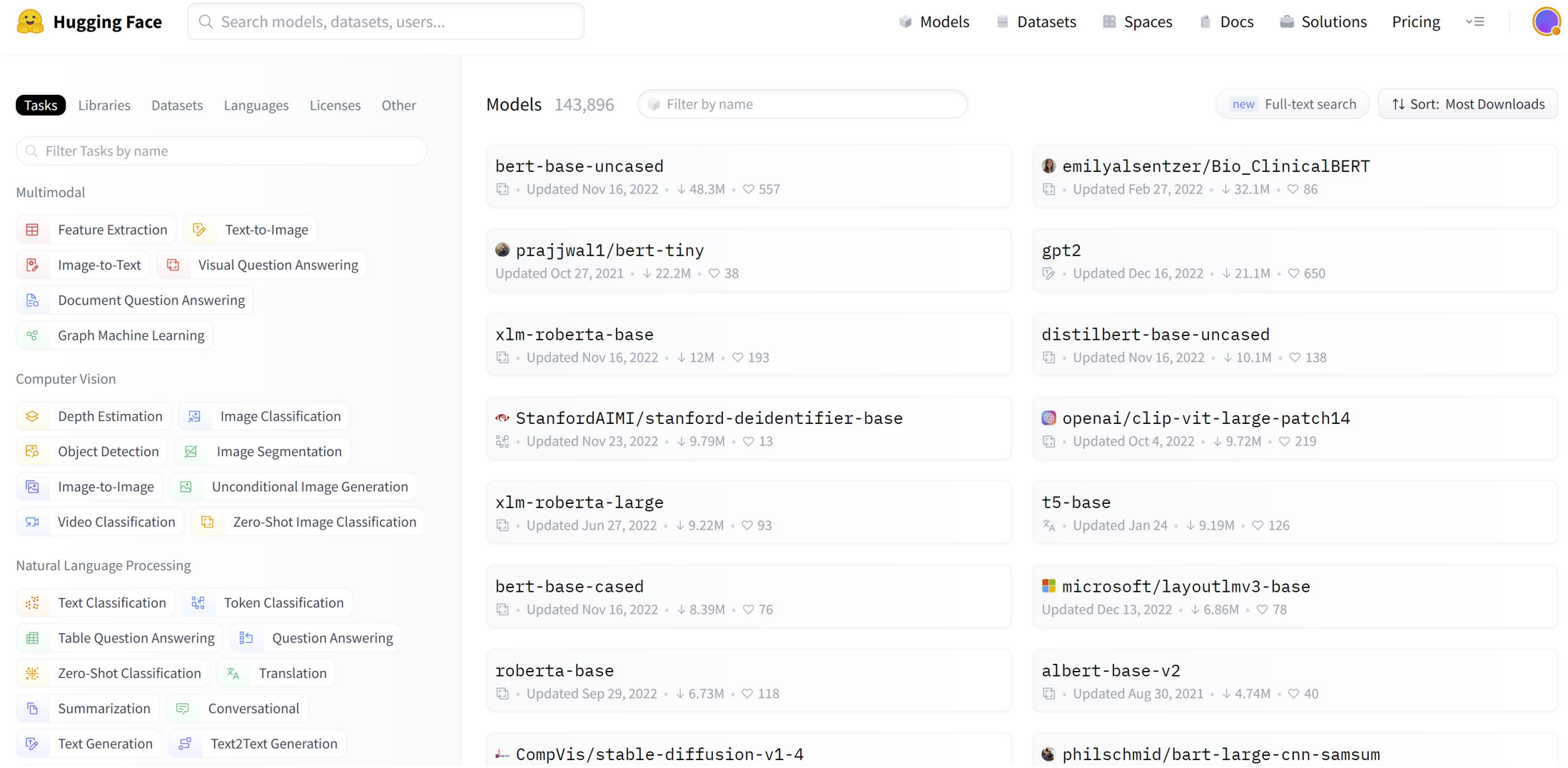


NFP77-Projekt «Open Justice versus Privacy»

- **Transparenz** erfordert Publikation von anonymisierten Gerichtsurteilen
- **Forschungsfrage:** Können diese anonymisierten Entscheide mittels künstlicher Intelligenz de-anonymisiert werden?
- Interdisziplinäres Forschungsprojekt von **Rechtswissenschaften und Informatik** der Universität Bern
- **Dauer:** 2020 bis 2024

https://www.nfp77.ch/portfolio/gerichtsurteile-im-spannungsfeld-zwischen-transparenz-und-privatsphaere/

HuggingFace – frei verfügbare Sprachmodelle



https://huggingface.co/models

Informatik-Doktorand Joel Niklaus

Swiss-Judgment-Prediction: A Multilingual Legal Judgment Prediction Benchmark

Joel Niklaus, Research Center for Digital Sustainability, Institute of Computer Science, University of Bern, Bern, Switzerland, joel.niklaus@inf.unibe.ch
Ilias Chalkidis, Coastal NLP Group, Department of Computer Science, University of Copenhagen, Copenhagen, Denmark, ilias.chalkidis@di.ku.dk
Matthias Stürmer, Research Center for Digital Sustainability, Institute of Computer Science, University of Bern, Bern, Switzerland, matthias.stuermer@inf.unibe.ch

LEXTREME: A Multi-Lingual and Multi-Task Benchmark for the Legal Domain

Joel Niklaus [1,2,6*] Veton Matoshi [2*] Pooja Rani [3] Andrea Galassi [4] Matthias Stürmer [1,2] Ilias Chalkidis [5]
[1]University of Bern [2]Bern University of Applied Sciences [3]University of Zurich [4]University of Bologna [5]University of Copenhagen [6]Stanford University

An Empirical Study on Cross-X Transfer for Legal Judgment Prediction

Joel Niklaus [†*] Matthias Stürmer [†] Ilias Chalkidis [‡◇*]
† Institute of Computer Science, University of Bern, Switzerland
‡ Department of Computer Science, University of Copenhagen, Denmark
◇ Cognitiv+, Athens, Greece

Joel Niklaus
Wissenschaftlicher Mitarbeiter am Institut Public Sector Transformation

Berner Fachhochschule
Departement Wirtschaft
Brückenstrasse 73
3005 Bern

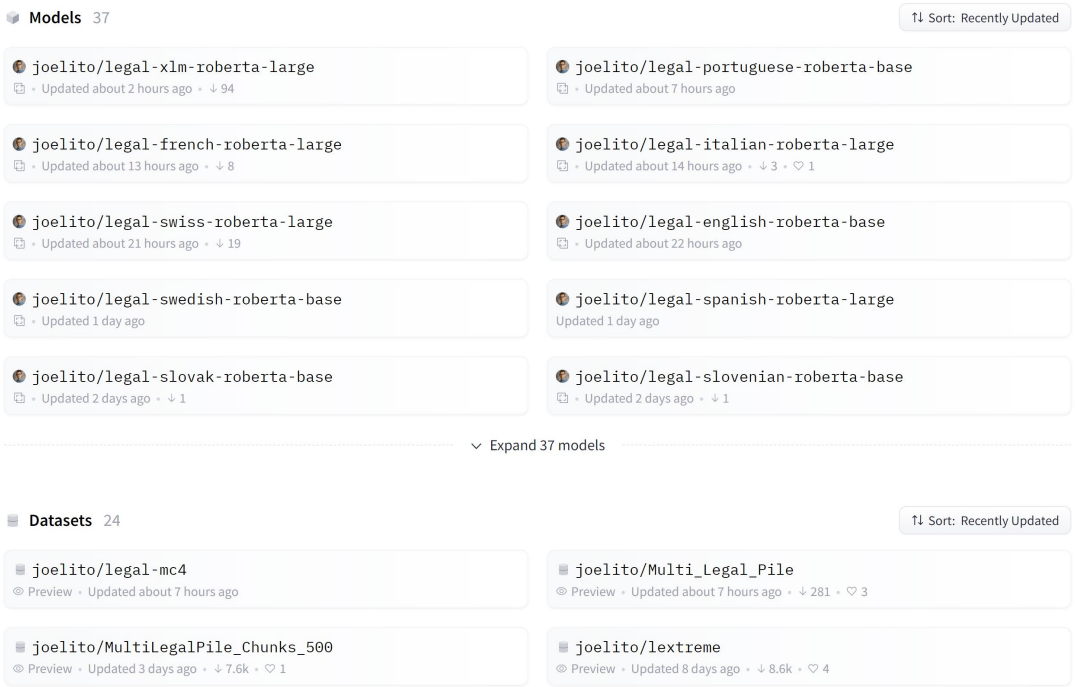
joel.niklaus.2@bfh.ch

https://scholar.google.com/citations?user=qJ8iricAAAAJ https://niklaus.ai

Multi-Lingual Legal Language Model für das Bundesgericht

- **Neues Sprachmodell,** das ausschliesslich auf juristischen Texten trainiert wurde (Pre-Training)
- Weltweit grösste, mehrsprachige **Rechtstextesammlung:**
  - Native Multi Legal Pile (112GB)
  - Eurlex Resources (179GB)
  - Legal MC4 (106GB)
  - Pile of Law (292GB)
- Anwendung beim Bundesgericht für **bessere Anonymisierung der Gerichtsurteile**

Hugging Face

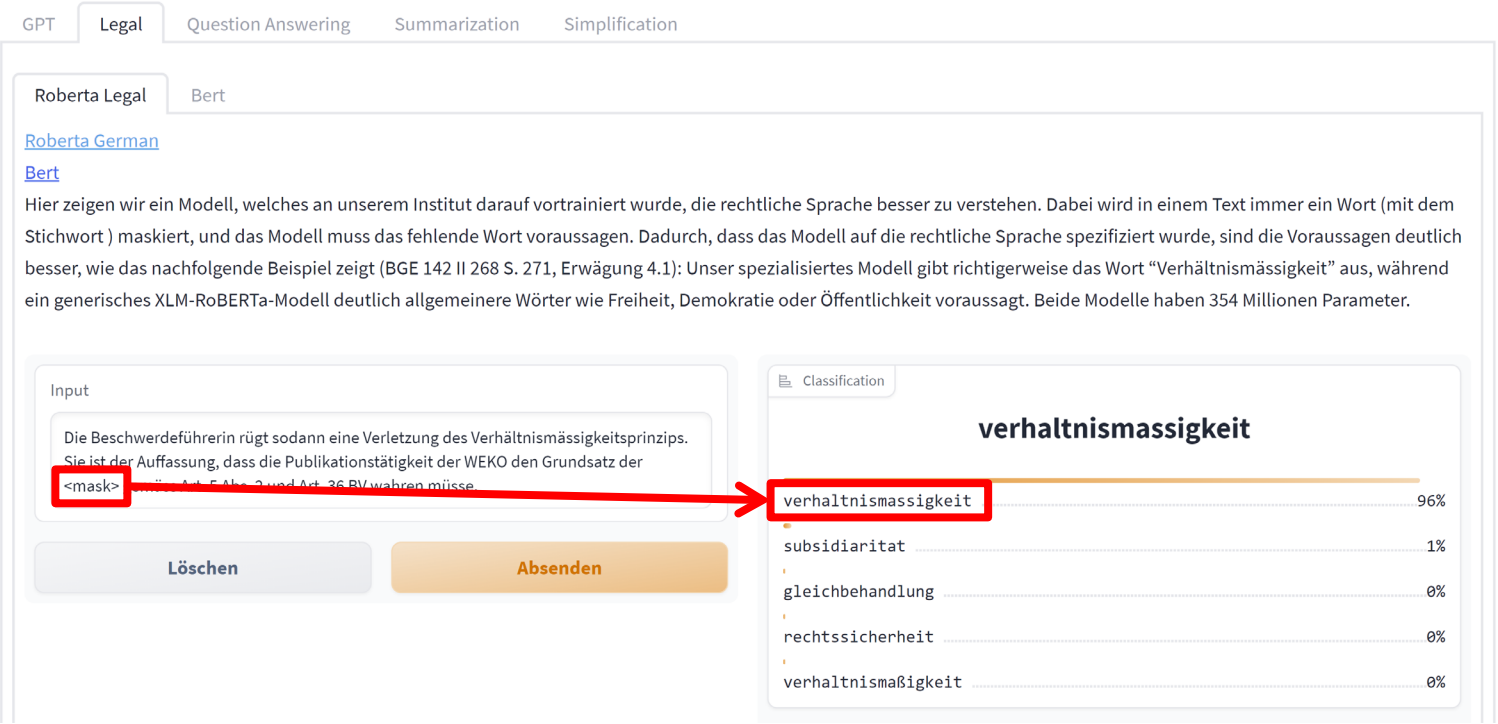


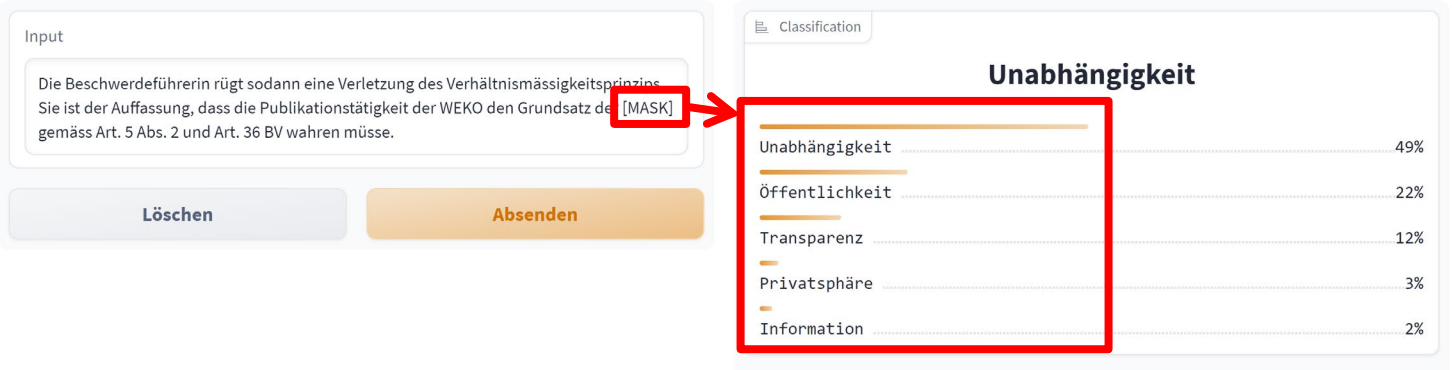
https://www.bfh.ch/de/forschung/forschungsprojekte/2021-432-033-021/ https://huggingface.co/joelito

Demo «Question Answering» in Deutsch

Erste Tests zeigen:

- Unser Sprachmodell **basierend auf Rechtstexten** ergibt präzisere Vorhersagen für Wortmaskierungen
- **Generisches BERT-Modell** kennt die Rechtssprache nicht gut

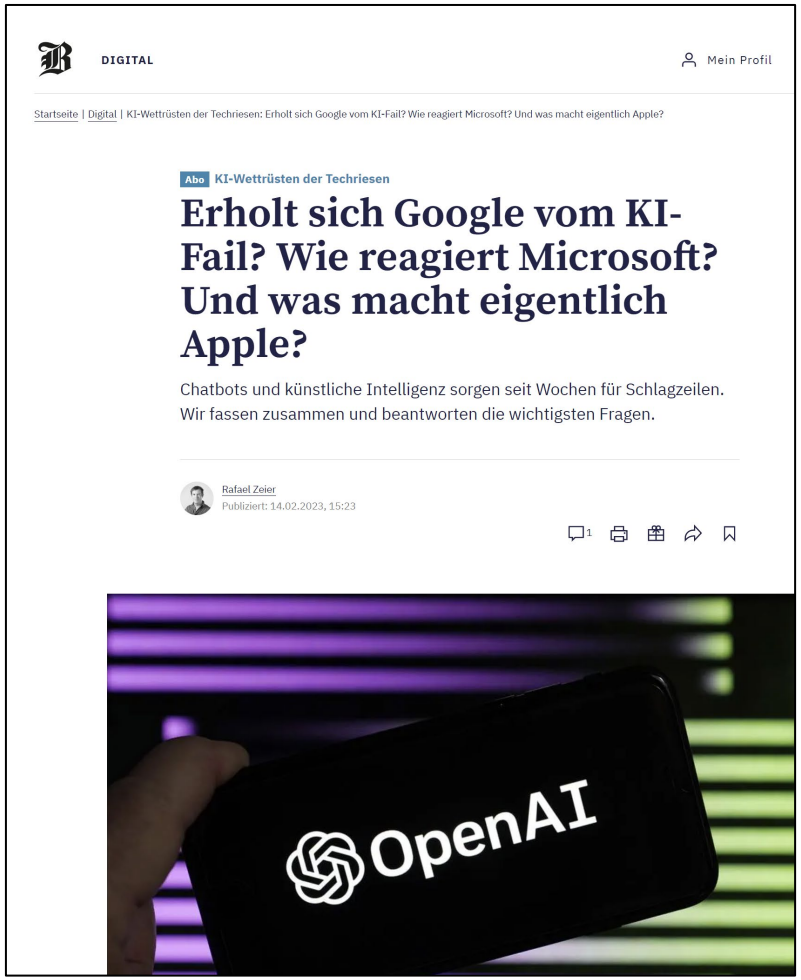




https://digital-sustainability.github.io/nlp-demo

(Zwischen-) Fazit

- Täglich **neue, technologische Entwicklungen** rund um KI, **Big Tech Firmen** investieren immense Summen für (private) Forschung →
- Frage: Wie kann der Staat **Kontrolle und digitale Souveränität** bewahren?
- Meine persönlichen Empfehlungen:
  - **Eigenes KI-Knowhow** auf Führungsebene und bei Mitarbeitenden aufbauen
  - **Vernetzung** mit Hochschulen, zivilgesell-schaftlichen Organisationen etc.
  - **Ausprobieren,** Proof-of-Concepts durchführen, Mitwirken an öffentlichen Hackathons etc.



https://www.derbund.ch/erholt-sich-google-vom-ki-fail-wie-reagiert-microsoft-und-was-macht-eigentlich-apple-422707119546

Berner Fachhochschule

Informationen zum Programm & Ticketverkauf
www.bfh.ch/transform

TRANSFORM 2023: Künstliche Intelligenz im öffentlichen Sektor

Eine Konferenz mit Beiträgen aus Politik, Verwaltung, Wissenschaft und Zivilgesellschaft am Mittwoch, 3. Mai 2023 im Berner Rathaus

Dr. Bertrand Loison
Bundesamt für Statistik

Dr. Paulina Grnarova
DeepJudge AI

Matthias Mazenauer
Kanton Zürich

Prof. Oliver Neumann
Universität Lausanne

Daniel Brunner
Bundesgericht

Dr. Marcel Gygli
Nationalbibliothek

Prof. Nadja Braun Binder
Universität Basel

Min Lin Marti
Nationalrätin SP

Prof. Mascha Kurpicz-Briki
Berner Fachhochschule

Raphael von Thiessen
Kanton Zürich

Syrian Hadad
Kanton Aargau

Dr. des. Angela Müller
AlgorithmWatch

SCHWEIZERISCHES BUNDESGERICHT

KÜNSTLICHE INTELLIGENZ

BERN – 7. MÄRZ 2023

Daniel.Brunner@bger.ch
Leiter Informatikdienst

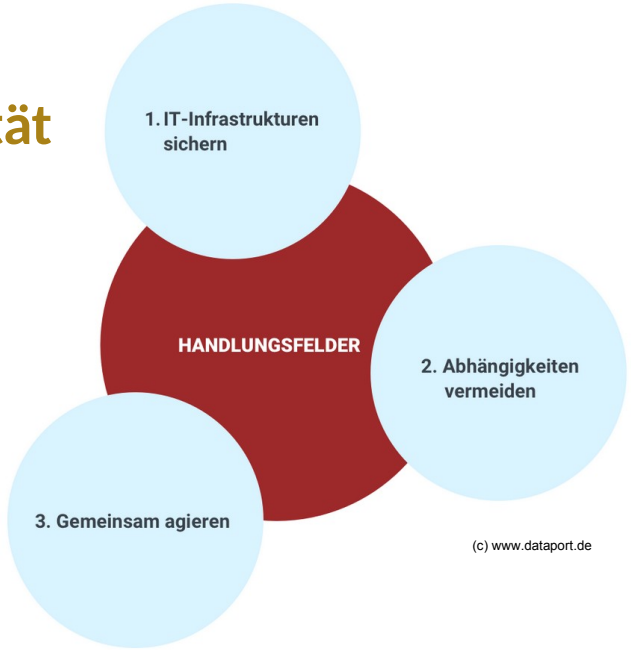
HOHE ANFORDERUNGEN

- ✓ **Das Schweizerische Bundesgericht ist die letzte richterliche Instanz, es gibt daher keine Möglichkeit für Rechts-mittel gegen ihre Entscheide**
  - Es werden rund 8'200 Fälle pro Jahr in 4 Landessprachen behandelt
  - Mehr als 400 Arbeitsplätze, aufgeteilt an 2 Standorten in Lausanne und Luzern
- ✓ **Der Jurist muss die Sicherheit haben, das zu finden, nachdem er sucht, und dies unabhängig von der Sprache**
- ✓ **Beste juristische Informatik CH**

EINE OFFENE STRATEGIE
MIT PRIVATER CLOUD

- ✓ Die Anwendungen und Infrastrukturen basieren auf **offenen Standards** und **Open Source Software.**
  - **Redhat/CentOS/Fedora** und **Gnome 3.x** für die Arbeitsplätze.
  - **Libre Office, Thunderbird** und **Firefox** für die Büroumgebung.
  - **Zimbra** als Server für Nachrichten und Collaboration.
  - **Alfresco** für die elektronische Dokumentenverwaltung und **Postgresql** für die Datenbank.

- Das Bundesgericht ist Herr seiner Infrastruktur und Daten **(Private Cloud)**
- Trotz oder weil der Quellcode **offen ist, sind Anwendungen als sicherer einzustufen**
- Eckpfeiler zur **Digitalen Souveränität**



LibreOffice
The Document Foundation

DER WEG ZUR INNOVATION…

- ✓ Ermitteln von neuen Verwendungszwecken der Technologien oder Anwendungen durch die Benutzer
- ✓ Politik für das Recht auf Fehler !

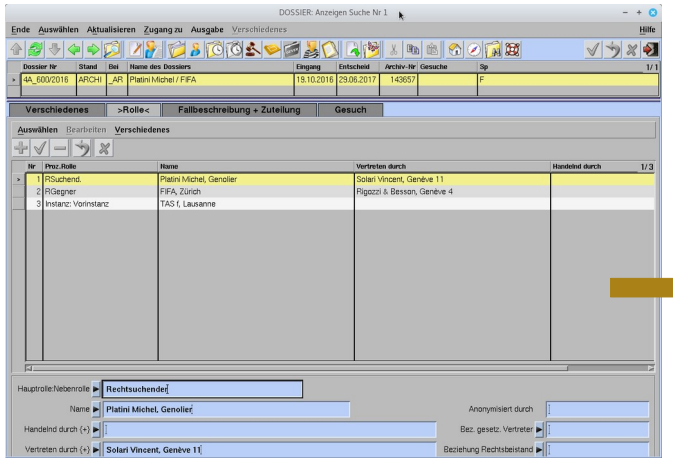
" Innovation entsteht nicht durch ein Pflichtenheft "
-Elmar Mock – co-Erfinder von Swatch

DER WEG ZUR KÜNSTLICHEN INTELLIGENZ…

- ✓ Kompetenzen dank Begleitung einer Master Arbeit zum Thema Künstliche Intelligenz
- ✓ Aufbau interner KI Kompetenzen
- ✓ Zusammenarbeit mit BFH zur Optimierung der KI
- ✓ KOSTBARE DATEN – WIE KANN DANK **KI** DER AM BGER **VORHANDENE DATENSCHATZ** GENUTZT WERDEN?

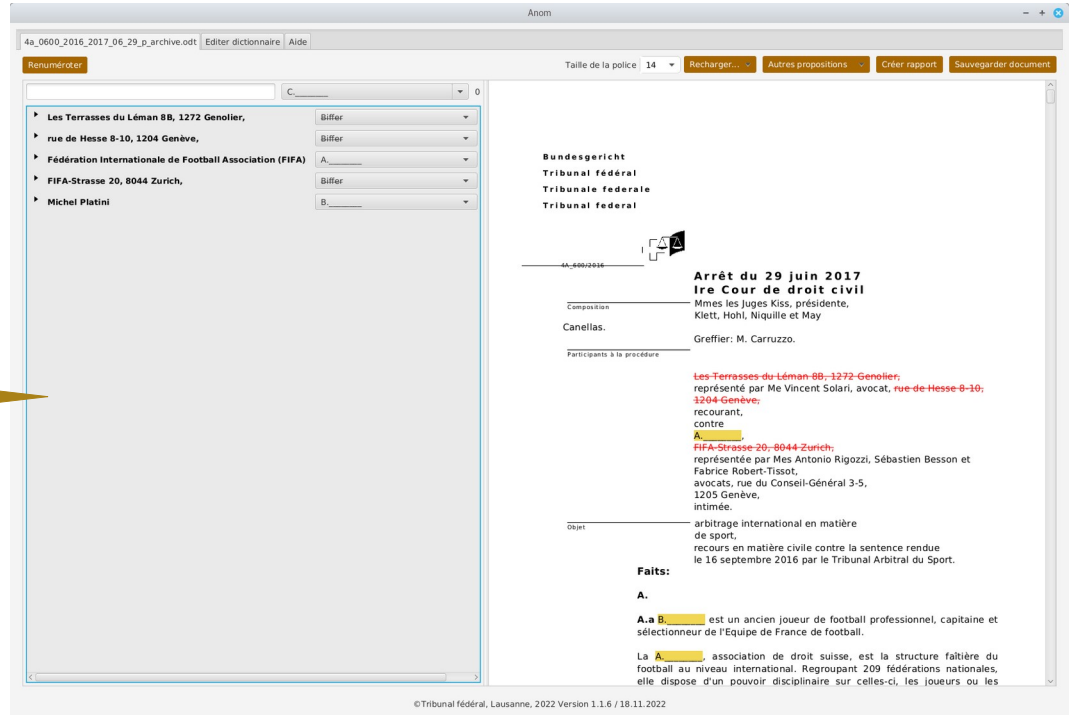
ANONYMISIERUNG DER ENTSCHEIDE
URSPRÜNGLICHE APPLIKATION

- Seit 2013 steht den Benutzern eine Applikation zur Unterstützung bei der Anonymisierung zur Verfügung
- Die Applikation basiert auf einem Regelwerk, das sich unter anderem auf den Parteinamen (Metadaten) bezieht, die in der Geschäftsverwaltung eintragen werden



Regelwerk :
M. Platini
Michel P.
CHF *.*.-
…

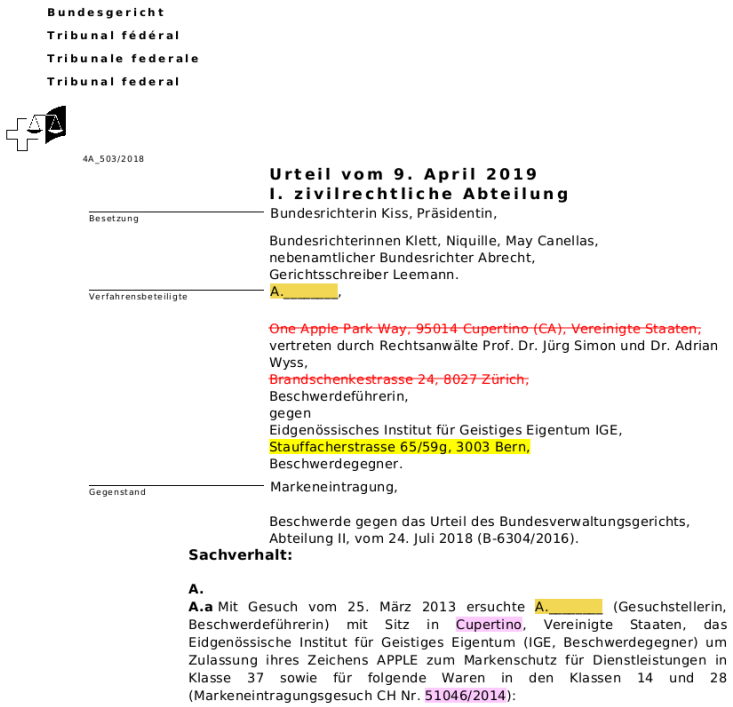
Zeugen ?
Ärzte ?
usw.



ANONYMISIERUNG DER ENTSCHEIDE

PROBLEMATIK

- ✓ **Die Anonymisierung von Gerichtsentscheiden ist eine nicht triviale und zeitaufwändige Aufgabe.**
- ✓ **All diese Zeit könnte für andere Zwecke genutzt werden.**
- ✓ **Das heutige Regelwerk kann die Komplexität der Aufgabe nicht erfassen.**
- ✓ **Viele Benutzer machen alles von Hand!**



Bundesgericht
Tribunal fédéral
Tribunale federale
Tribunal federal

4A_503/2018

Urteil vom 9. April 2019
I. zivilrechtliche Abteilung

Besetzung — Bundesrichterin Kiss, Präsidentin,
Bundesrichterinnen Klett, Niquille, May Canellas,
nebenamtlicher Bundesrichter Abrecht,
Gerichtsschreiber Leemann.

Verfahrensbeteiligte — A._____,
~~One Apple Park Way, 95014 Cupertino (CA), Vereinigte Staaten,~~
vertreten durch Rechtsanwälte Prof. Dr. Jürg Simon und Dr. Adrian Wyss,
~~Brandschenkestrasse 24, 8027 Zürich,~~
Beschwerdeführerin,
gegen
Eidgenössisches Institut für Geistiges Eigentum IGE,
Stauffacherstrasse 65/59g, 3003 Bern,
Beschwerdegegner.

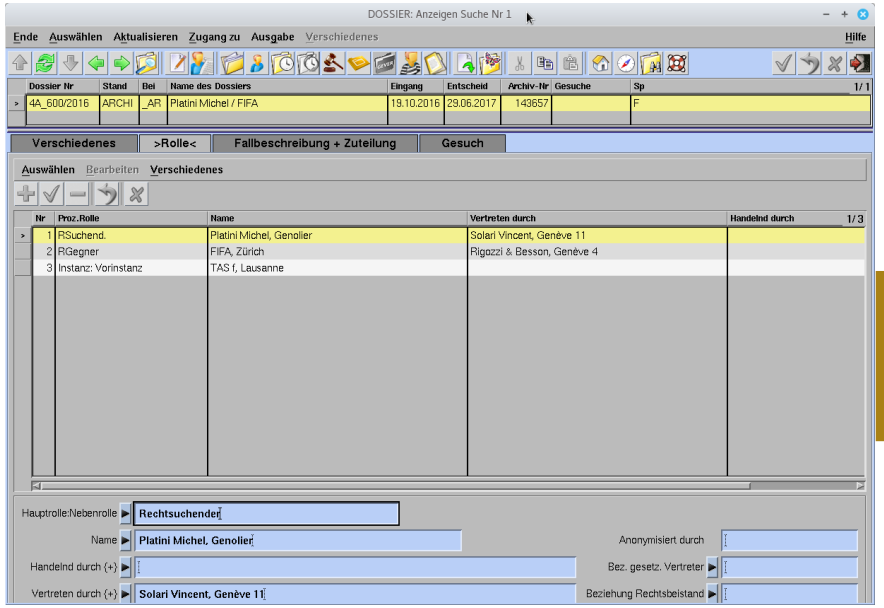
Gegenstand — Markeneintragung,

Beschwerde gegen das Urteil des Bundesverwaltungsgerichts, Abteilung II, vom 24. Juli 2018 (B-6304/2016).

Sachverhalt:

A.
A.a Mit Gesuch vom 25. März 2013 ersuchte A._____ (Gesuchstellerin, Beschwerdeführerin) mit Sitz in Cupertino, Vereinigte Staaten, das Eidgenössische Institut für Geistiges Eigentum (IGE, Beschwerdegegner) um Zulassung ihres Zeichens APPLE zum Markenschutz für Dienstleistungen in Klasse 37 sowie für folgende Waren in den Klassen 14 und 28 (Markeneintragungsgesuch CH Nr. 51046/2014):

ANONYMISIERUNG DER ENTSCHEIDE
KÜNSTLICHER INTELLIGENZ, AUSGANGSLAGE



Die bisher geleistete Arbeit stellt eine Goldmine an Daten dar.

Alles steht in offenen und leicht zu handhabenden Formaten zur Verfügung.

ANONYMISIERUNG DER ENTSCHEIDE
KÜNSTLICHER INTELLIGENZ, AUSGANGSLAGE (2)

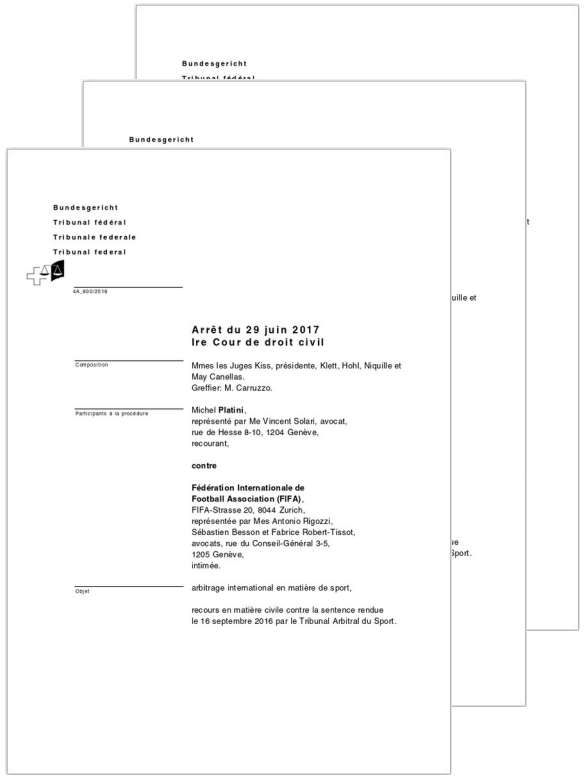
- **Für das Training des KI-Modells (Hirn) wird grosse Rechenleistung benötigt**
- **Aus Datenschutzgründen ist es nicht möglich Rechenleistung bei Google, Amazon usw. einzukaufen**
- **Die Beschaffung von 2 spezialisierten Hochleistungs-Prozessorkarten war notwendig**

amazon

Google AI

Microsoft AI

ANONYMISIERUNG DER ENTSCHEIDE
KÜNSTLICHER INTELLIGENZ, VORGEHEN



Bundesgericht
Tribunal fédéral
Tribunale federale
Tribunal federal

4A_600/2016

Arrêt du 29 juin 2017
Ire Cour de droit civil

Composition

Mmes les Juges Kiss, présidente, Klett, Hohl, Niquille et May Canellas.
Greffier: M. Carruzzo.

Participants à la procédure

Michel **Platini**,
représenté par Me Vincent Solari, avocat,
rue de Hesse 8-10, 1204 Genève,
recourant,

contre

Fédération Internationale de Football Association (FIFA),
FIFA-Strasse 20, 8044 Zurich,
représentée par Mes Antonio Rigozzi,
Sébastien Besson et Fabrice Robert-Tissot,
avocats, rue du Conseil-Général 3-5,
1205 Genève,
intimée.

Objet

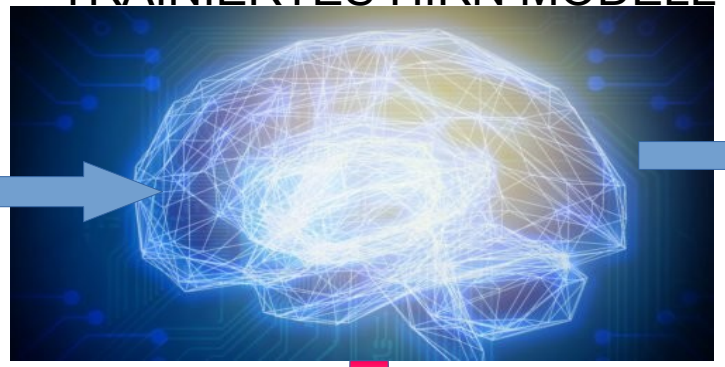
arbitrage international en matière de sport,

recours en matière civile contre la sentence rendue
le 16 septembre 2016 par le Tribunal Arbitral du Sport.

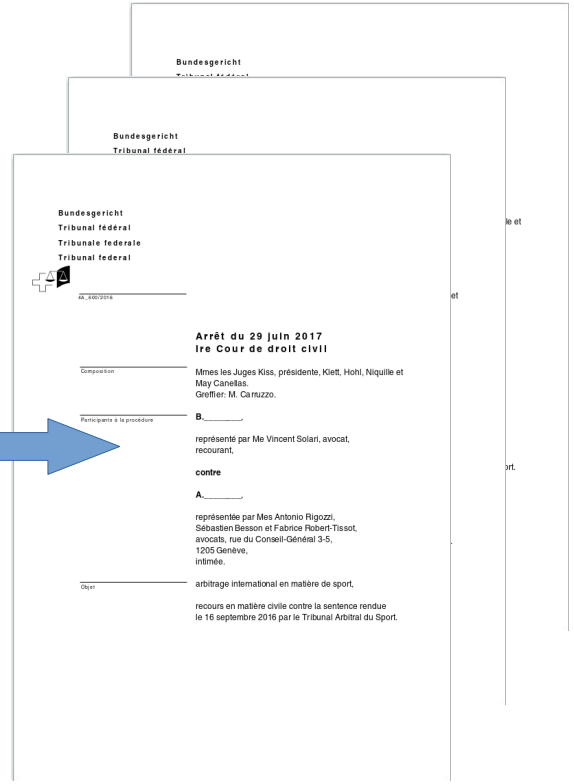
110'000 x
Entscheide
mit Namen

Neuer Entscheid mit Namen

TRAINIERTES HIRN MODELL



Neuer Entscheid Anonymisiert



Bundesgericht
Tribunal fédéral
Tribunale federale
Tribunal federal

4A_600/2016

Arrêt du 29 juin 2017
Ire Cour de droit civil

Composition

Mmes les Juges Kiss, présidente, Klett, Hohl, Niquille et May Canellas.
Greffier: M. Carruzzo.

Participants à la procédure

B._____,

représenté par Me Vincent Solari, avocat,
recourant,

contre

A._____,

représentée par Mes Antonio Rigozzi,
Sébastien Besson et Fabrice Robert-Tissot,
avocats, rue du Conseil-Général 3-5,
1205 Genève,
intimée.

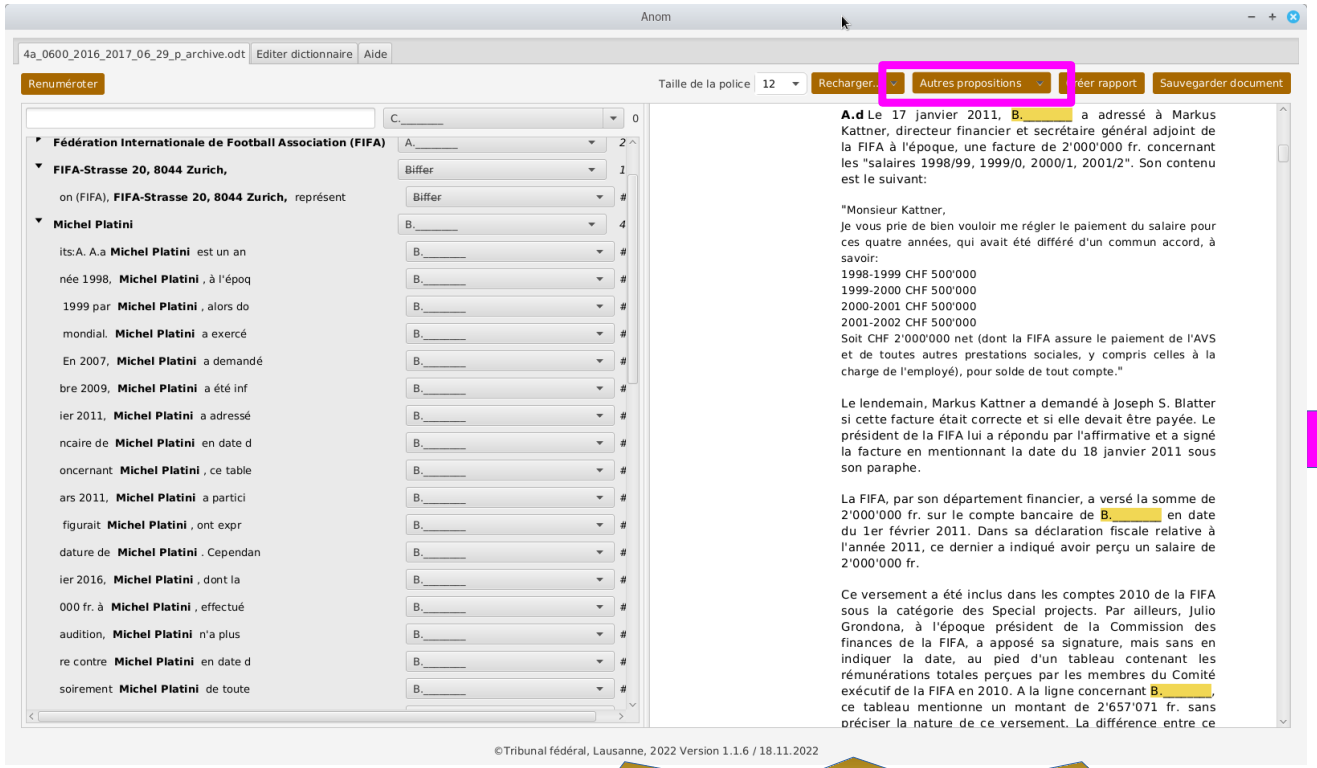
Objet

arbitrage international en matière de sport,

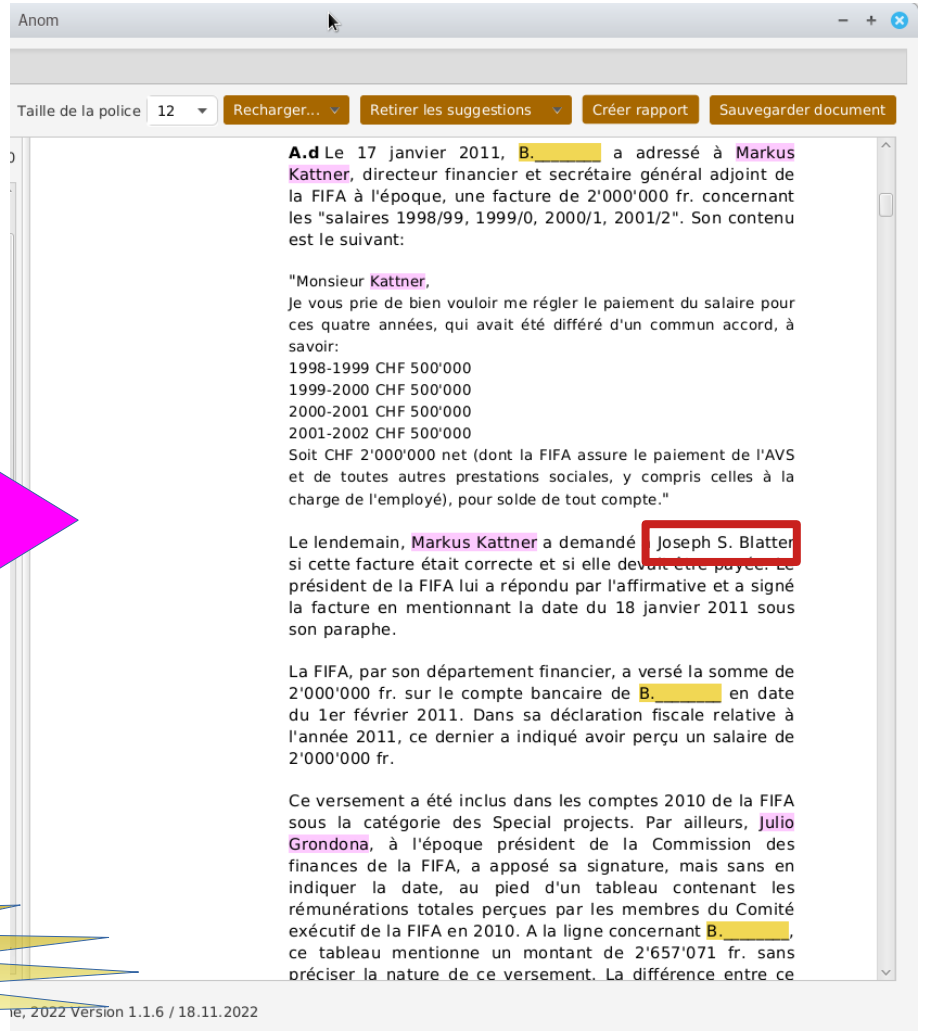
recours en matière civile contre la sentence rendue
le 16 septembre 2016 par le Tribunal Arbitral du Sport.

110'000 x
Entscheide
anonymisiert

ANONYMISIERUNG DER ENTSCHEIDE
KÜNSTLICHER INTELLIGENZ, RESULTAT (2)



Zusätzliche Vorschläge KI



Resultat: Trefferquote -> 90.8%, Genauigkeit -> 91.3 %

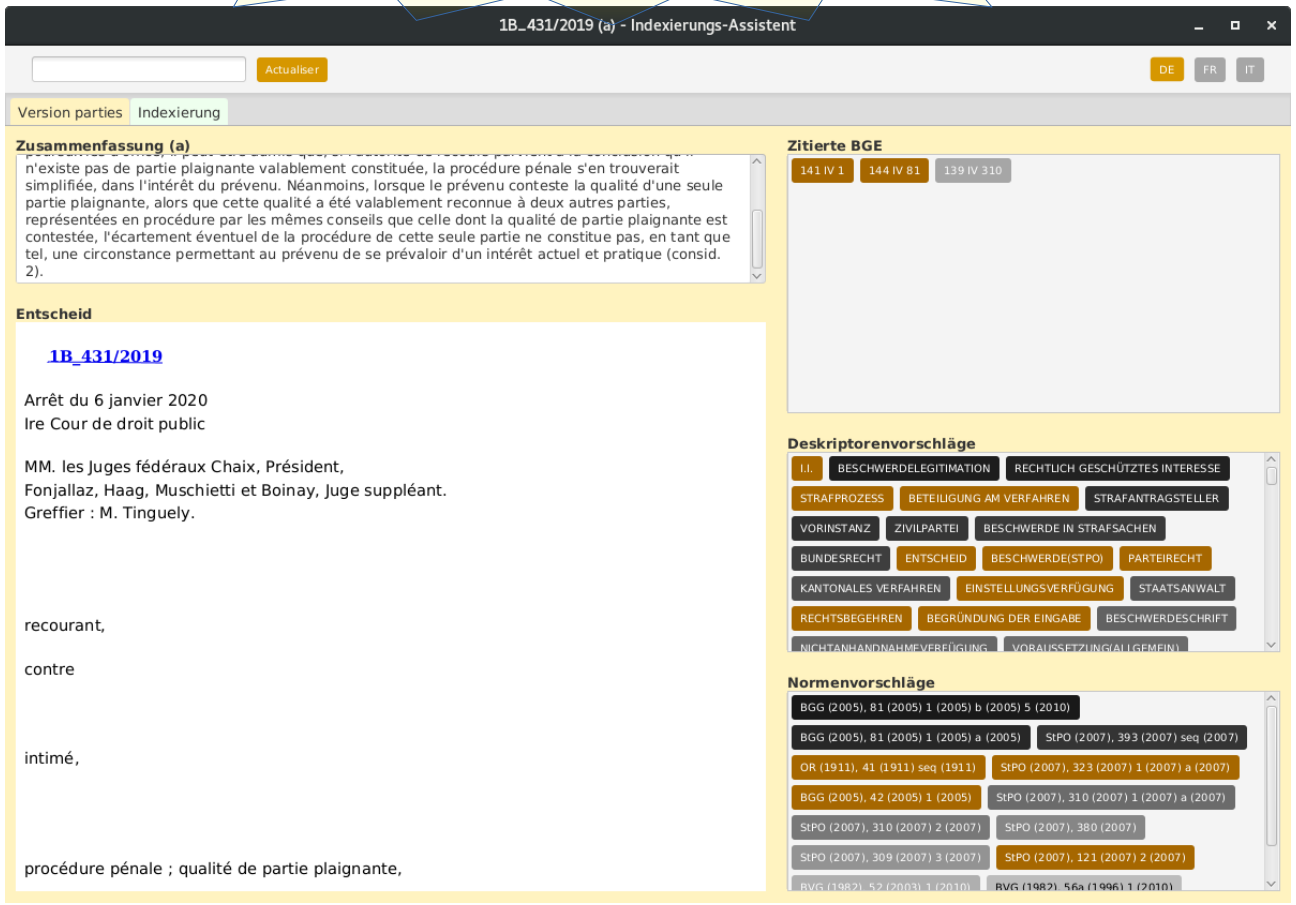
KÜNSTLICHE INTELLIGENZ AM BUNDESGERICHT

AUSBLICK...

Ziel der KI ist den Menschen zu unterstützen, nicht (komplett) zu ersetzen !

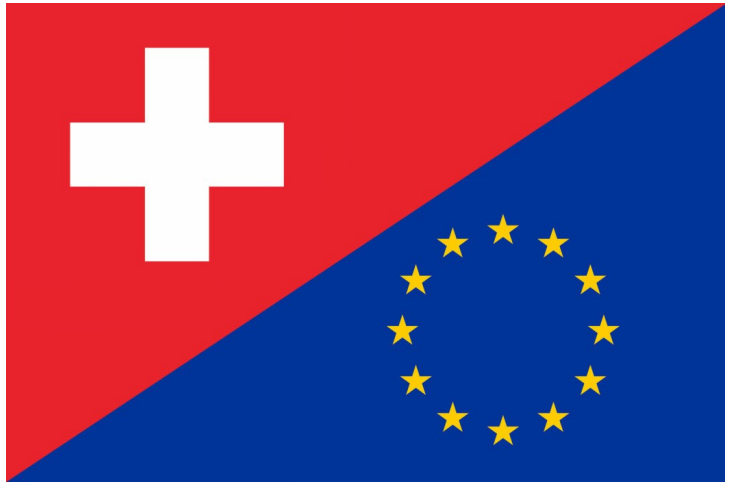
- ✓ **Eine Applikation zur Unterstützung bei der manuellen Indexierung nach Normen und Stichwörter**

- ✓ **Folgende Goldminen stehen dafür zur Verfügung**
  - **Zusammenfassungen**
  - **Manuelle Indexierung von ca. 100'000 Entscheiden**
  - **Indexierungsregeln**
  - **Zitierte Gesetzesartikeln und BGE Referenzen**
  - **History der Suchen in der Entscheiddatensammlung**
  - **Systematische Rechtssammlung**



KÜNSTLICHE INTELLIGENZ IN DER SCHWEIZ

HERAUSFORDERUNGEN..

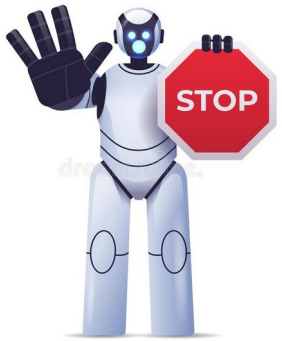


- ✓ **Lokale Kompetenzen**
  - -> KI an den Hochschulen
- ✓ **Rechnenleistung**
  - Lokal versus Global
- ✓ **Datenschutz und Ethik**
  - Abhängigkeiten von USA und Chinesischen Denkweise/Kulturen/Staaten
- ✓ **Datenschatz**
  - Aktualität der Daten
    - Indexierungsregeln, die ändern
    - Bei Vorhersagen von Entscheidungen
      - politische Verteilung
      - Veränderung der Denkweise (z.B. zum Thema Ökologie, Nachhaltigkeit, Energie, usw.)
- ✓ **Qualität und Datenvolumen**
  - z.B. OCR-Erkennung bei eingescannten Texten
  - korrekt anonymiserten Daten beim Benutzen von Cloud-Rechenleistung

KÜNSTLICHE INTELLIGENZ…
DATENSCHATZ versus **DATENSCHUTZ**

KÜNSTLICHE INTELLIGENZ IN DER SCHWEIZ

ETHIK UND KULTUR



KÜNSTLICHE INTELLIGENZ
RISIKEN

FRAGEN ?

Parldigi MasterClass 2023

Grosser Rat Bern

nächste Veranstaltung **am 6. Juni 2023**

Künstliche Intelligenz in der Medizin

Einführung**:**

Dr. Verena Schöning
Klinische Pharmakologie, Universitätsklinik für Allgemeine Innere Medizin, Inselspital

Kommentar:

PD Dr. Markus Christen
Geschäftsführer Digital Society Initiative, UZH

Eine Veranstaltung von:

Universität Zürich[UZH]
Digital Society Initiative

Parldigi

Unterstützt durch:

Stiftung Mercator Schweiz