



SCHWEIZERISCHES BUNDESGERICHT

NLP am BUNDESGERICHT

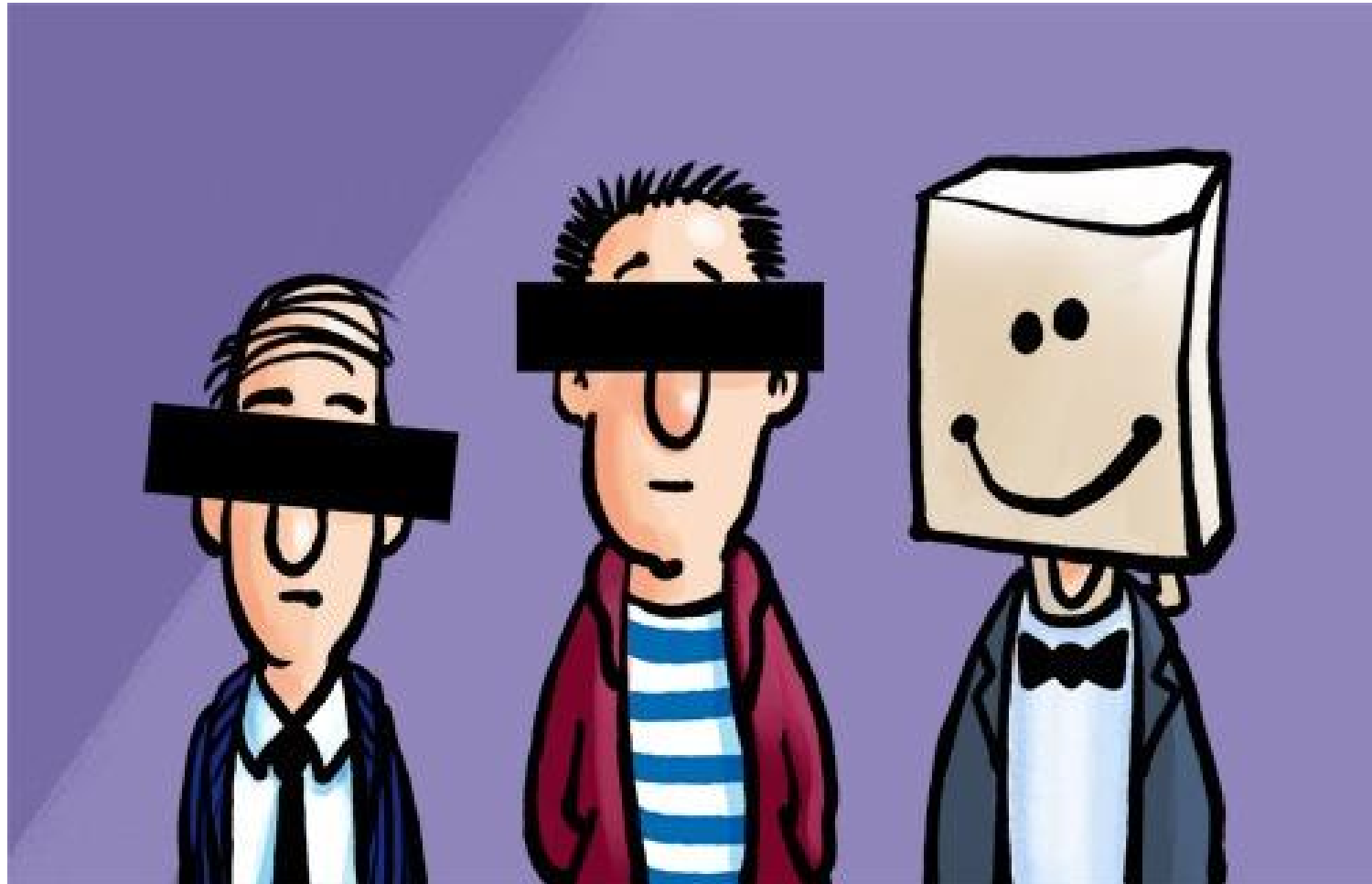
TRANSPARENTE JUSTIZ – 21. JUNI 2024

Daniel.Brunner@bger.ch
Leiter Informatikdienst



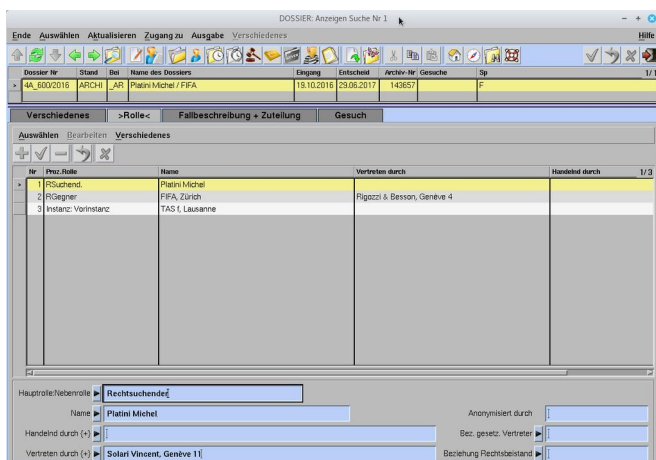
KÜNSTLICHE INTELLIGENZ

ANONYMISIERUNG DER ENTSCHEIDE AM BGer



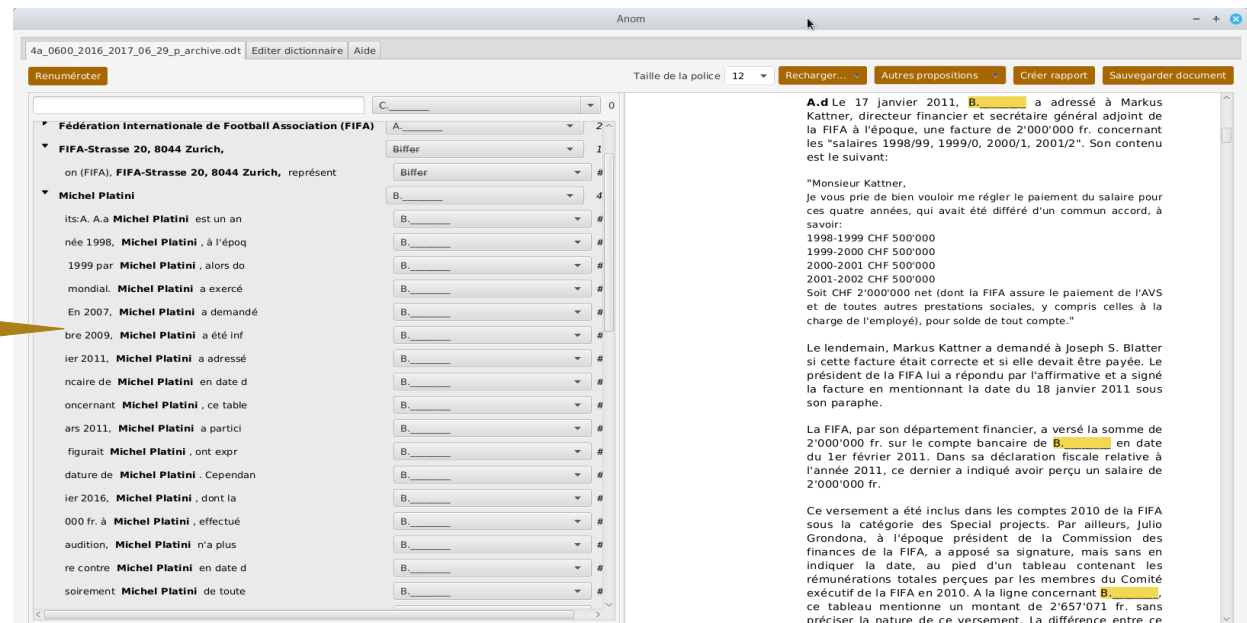
ANONYMISIERUNG DER ENTSCHEIDUNG URSPRÜNGLICHE APPLIKATION

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

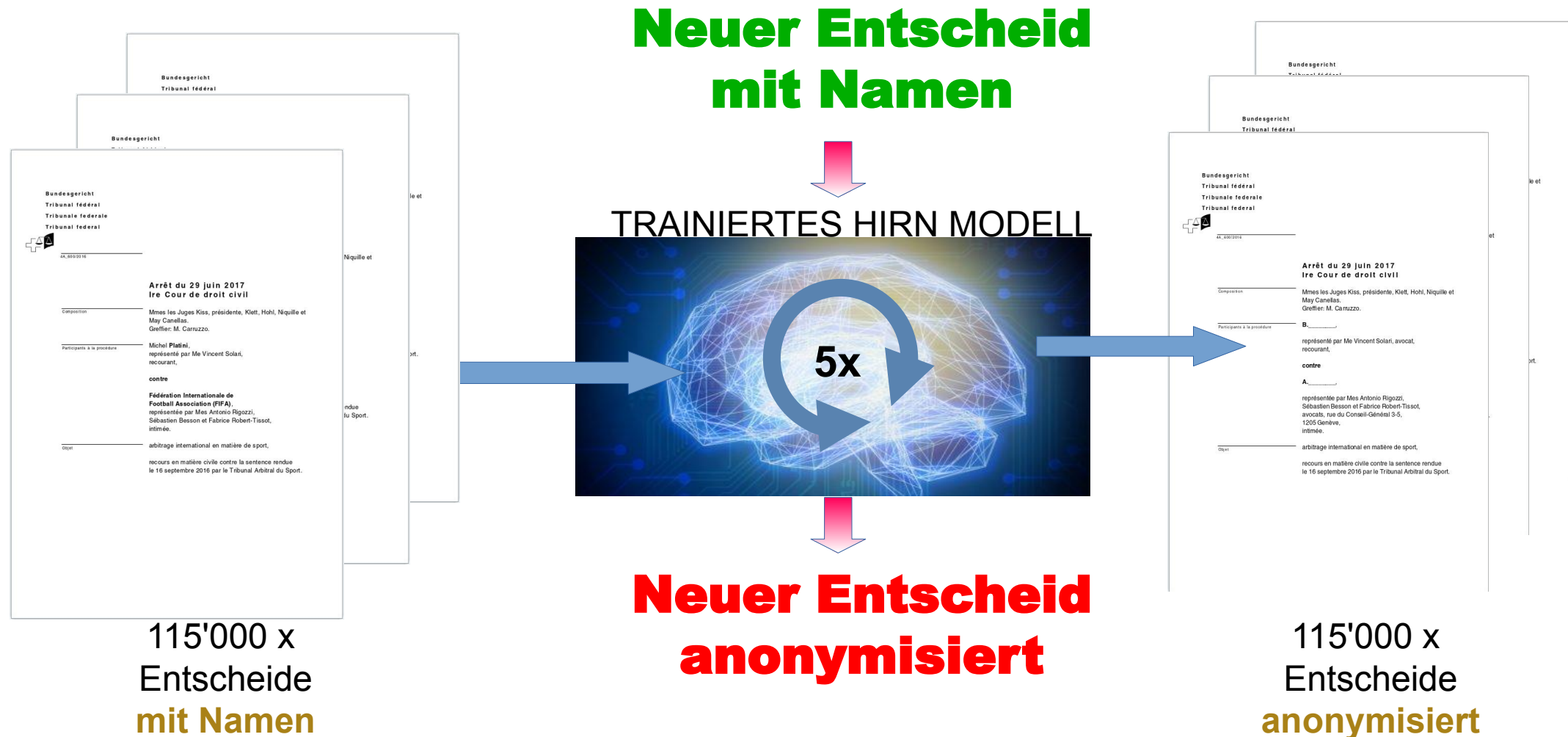


Regelwerk :
M. Platini
Michel P.
CHF *.*-
...

 **Zeugen ?**
Ärzte ?
usw.

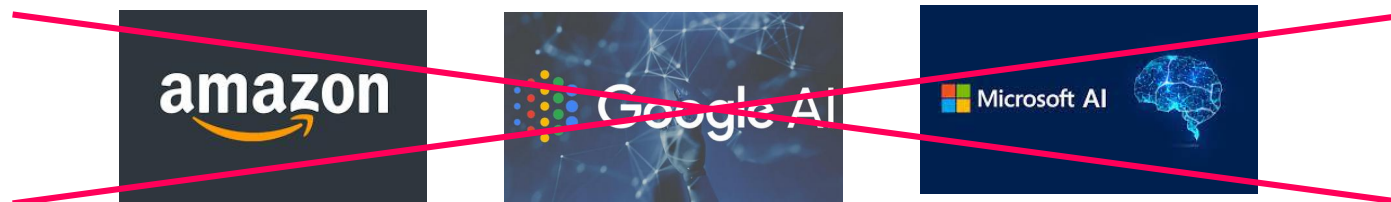


ANONYMISIERUNG DER ENTSCHEIDE KÜNSTLICHER INTELLIGENZ: VORGEHEN



ANONYMISIERUNG DER ENTSCHEIDEN KÜNSTLICHER INTELLIGENZ: VORGEHEN

- 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



ANONYMISIERUNG DER ENTSCHEIDE

KÜNSTLICHE INTELLIGENZ: RESULTATE (1)

The screenshot shows the Anom software interface. On the left, there is a sidebar with a list of names and their associated document IDs. The main area displays a document with several redacted names (e.g., B. _____). A pink box highlights the 'Autres propositions' button in the top toolbar. The document text includes a letter from B. _____ to Markus Kattner, dated January 17, 2011, regarding a salary dispute. The text also mentions Joseph S. Blatter and the FIFA's financial department.

This screenshot shows the same document as the left screenshot, but with a pink arrow pointing to a redacted name 'Joseph S. Blatter' in the text. The text includes a letter from B. _____ to Markus Kattner, dated January 17, 2011, regarding a salary dispute. The text also mentions Joseph S. Blatter and the FIFA's financial department.

Zusätzliche
Vorschläge
KI

Resultate: Trefferquote -> 83.76% Genauigkeit -> 90.72 %



ANONYMISIERUNG DER ENTSCHEIDEN KÜNSTLICHE INTELLIGENZ: RESULTATE (2)

Legal Swiss Roberta

Resultate: Trefferquote -> 92.26% Genauigkeit -> 92.57%

Zusätzliche Vorschläge KI

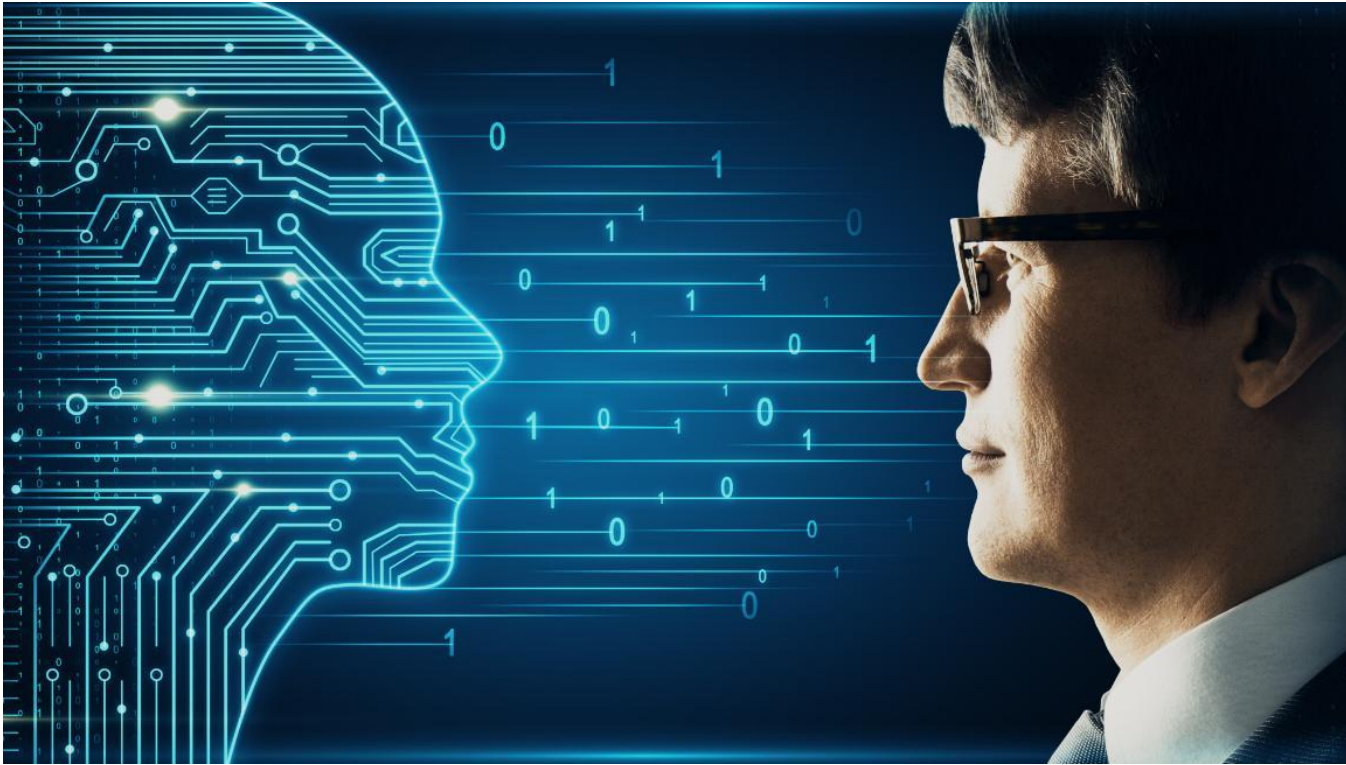
Resultate: Trefferquote -> 83.76% Genauigkeit -> 90.72%

The screenshot displays the Anom software interface for document redaction. The left sidebar shows a document tree with sections like 'Fédération Internationale de Football Association (FIFA)' and 'Michel Platini'. The main text area shows a document snippet with redaction suggestions (e.g., 'B.', 'A.', 'C.') and a 'Legal Swiss Roberta' watermark. The right-hand pane shows the original text with redaction markers. A red jagged shape highlights the top part of the interface, and a yellow jagged shape highlights the bottom part. A pink arrow points from the 'Zusätzliche Vorschläge KI' text to the right-hand pane. A Swiss cross logo is in the bottom left corner.



KÜNSTLICHE INTELLIGENZ

DER MENSCH MUSS LERNEN, MIT KI UMZUGEHEN!



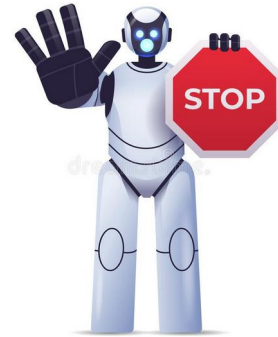
KÜNSTLICHE INTELLIGENZ IN DER SCHWEIZ

ETHIK UND KULTUR



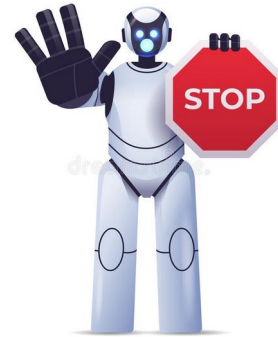
KÜNSTLICHE INTELLIGENZ IN DER SCHWEIZ

ETHIK UND KULTUR



KÜNSTLICHE INTELLIGENZ IN DER SCHWEIZ

ETHIK UND KULTUR



KÜNSTLICHE INTELLIGENZ IN DER SCHWEIZ

ETHIK UND KULTUR: Stereotypen und Vorurteile



KÜNSTLICHE INTELLIGENZ

WELCHE WERTE SOLLEN DIESE SYSTEME WIDERSPIEGELN?

- ✓ Die Daten, mit denen diese Systeme trainiert werden, spiegeln eine Weltanschauung wider
 - Gesellschaftliche Werte unterscheiden sich von Land zu Land und von Bevölkerung zu Bevölkerung
 - Die Werte, die von diesen Modellen gelernt werden, sind mit ihren Schöpfern verbunden
- ✓ Als Informationsquellen werden diese Systeme unsere Sicht der Welt beeinflussen, wenn wir mit ihnen interagieren
- ✓ Wenn wir in der Lage sein wollen, diese Systeme zu studieren, zu überwachen und sie zuverlässig und entsprechend den Erwartungen der Schweizer Gesellschaft einzusetzen, müssen wir sie selbst in der Schweiz erschaffen



freeworldmaps.net



Antoine Bosselut,
EPFL



KÜNSTLICHE INTELLIGENZ AN BUNDESGERICHT

STRATEGIE FÜR DAS SELBSTTRAINING

- ✓ Strategie basierend auf **offenen Sprachmodellen (Open Source)**
- ✓ **Lokale Rechenleistung (private Cloud)**
- ✓ Lernen anhand von **Daten** und den **Kompetenzen von unseren Juristen**
- ✓ **Lokale und spezifische Sprachmodellen** müssen mit **kontrollierten Informationen** trainiert werden
- ✓ Zusammenarbeit mit **Hochschulen** und vor allem
- ✓ **Bewusstsein für Risiken** schaffen und **Bereiche bewerten**, wo die KI unsere Benutzer im Alltag unterstützen kann

ETH und EPFL starten die "Swiss AI"-Initiative



Risiko basierter Ansatz

- Nicht diskriminierend / Ethic
- Sicherheit und Datenschutz
- Nachvollziehbarkeit und Transparenz
- **Am Ende entscheidet immer der Mensch!**



KÜNSTLICHE INTELLIGENZ VERWENDUNGSZWECKE?



Was ist KI ?

KI steht für künstliche Intelligenz. Es handelt sich um ein Bereich der Informatik, der sich mit der Entwicklung von Systemen und Technologien befasst, die in der Lage sind, Aufgaben auszuführen, die normalerweise menschliche Intelligenz erfordern würden. KI zielt darauf ab, Maschinen mit den Fähigkeiten auszustatten, wie Wahrnehmung, Sprachverarbeitung, Lernen, Problemlösung und Entscheidungsfindung.

Es gibt verschiedene Arten von KI, darunter:

1. Schwache KI: Auch bekannt als Narrow AI, bezieht sich auf KI-Systeme, die spezifische Aufgaben oder Probleme lösen können, aber nicht über menschenähnliche Intelligenz verfügen. Beispiele dafür sind Spracherkennungssysteme oder Bilderkennungsalgorithmen.
2. Starke KI: Auch als General AI bezeichnet, bezieht sich auf KI-Systeme, die eine allgemeine Intelligenz besitzen, vergleichbar mit menschlicher Intelligenz. Diese Art von KI ist in der Lage, eine Vielzahl von Aufgaben zu erlernen und auszuführen.

KI wird in verschiedenen Bereichen eingesetzt, wie z.B. Gesundheitswesen, Finanzwesen, Transport, Unterhaltung und vielen anderen. Es hat das Potenzial, komplexe Probleme zu lösen, menschliche Arbeit zu automatisieren und neue Möglichkeiten zu schaffen, um die Welt zu verbessern.



Kannst du es meinem 5-jährigen Sohn erklären ?

Natürlich! Ich werde mein Bestes geben, es deinem 5-jährigen Sohn einfach zu erklären.

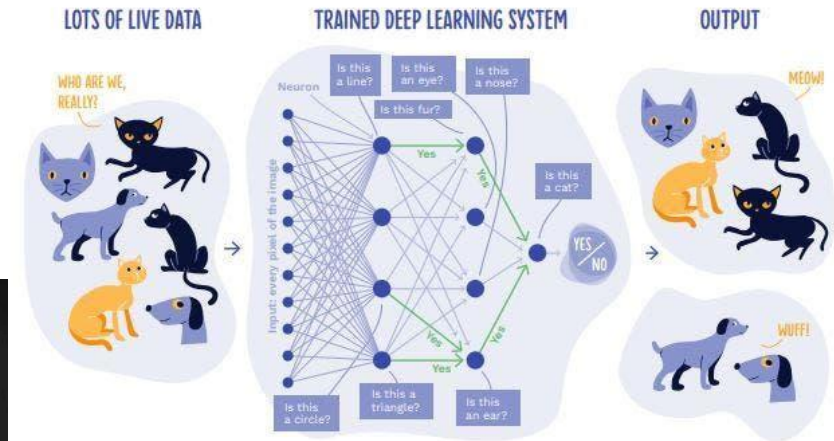
"Künstliche Intelligenz, oder KI, ist etwas, das Wissenschaftler erschaffen haben, um Computern beizubringen, wie Menschen zu denken. Es ist wie ein superkluger Roboter, der uns helfen kann, Dinge zu tun. KI kann uns helfen, Probleme zu lösen oder Fragen zu beantworten, genauso wie du es tust, wenn du lernst und nachdenkst. Manche KIs können auch mit uns sprechen und uns verstehen, so wie du mit deinen Freunden sprichst. Künstliche Intelligenz ist etwas Neues und Spannendes, und viele Menschen auf der ganzen Welt arbeiten daran, sie noch besser zu machen, damit sie uns helfen kann, viele interessante Dinge zu entdecken und zu lernen."

Ich hoffe, das hilft deinem Sohn, das Konzept der künstlichen Intelligenz zu verstehen!



Übersetzen
korrigieren

Kontextualisieren



Kategorisieren



KÜNSTLICHE INTELLIGENZ VERANTWORTUNGSVOLLER EINSATZ

Lernen aus Daten und
Kompetenzen unserer Benutzer

OSS Modelle:
Transparenz
und Datenschutz

- Llama 3
- Mixtral
- Phi 3

Keine
Anonymisierung
des Prompts
notwendig

ChatTF

W Napoleon Bonaparte - WiX +

https://lxchatgaia.bger.a...

Les plus visités es fällt aus - Traductio... OpenJust

+ Neue K...

Biographie von Napole...

Ferien für Putzfrauen und ...

Rechtsbeschwerde gegen ...

Renouvellement du site int...

Kontext verwenden:
Nein

Modell:
GPT 4 Turbo — OpenAI

Anonymisierungsmodell:
spacy

System Prompt:
Sie sind ein Assistent des Schweizer Bundesgerichts. Bitte übernehmen Sie die Rechtschreibung mit 'ss'.

Zusammenfassen

Napoleon Bonaparte, als Kaiser Napoleon I. (französisch Napoléon Bonaparte bzw. Napoléon Ier; * 15. August 1769 in Ajaccio auf Korsika als Napoleone Buonaparte; † 5. Mai 1821 in Longwood House auf St. Helena im Südatlantik), war General der Ersten Republik, Erster Konsul Frankreichs und schließlich Kaiser der Franzosen.

Aus korsischer Familie stammend, stieg Bonaparte während der Französischen Revolution in der Armee auf. Er erwies sich als ein militärisches Talent ersten Ranges. Vor allem die Feldzüge in Italien und in Ägypten machten ihn populär. Dies ermöglichte ihm, durch den Staatsstreich des 18. Brumaire VIII (9. November 1799), zunächst als einer von drei Konsuln, die Macht in Frankreich zu übernehmen. Von 1799 bis 1804 als Erster Konsul der Französischen Republik und anschließend bis 1814 sowie nochmals 1815 als Kaiser der Franzosen stand er einem diktatorischen Regime mit plebiszitären Elementen vor.

Durch verschiedene Reformen – etwa die der Justiz durch den Code civil oder die der Verwaltung – hat Napoleon die staatlichen Strukturen Frankreichs bis in die Gegenwart hinein geprägt und die Schaffung eines modernen Zivilrechts in besetzten europäischen Staaten initiiert. Außenpolitisch errang er, gestützt auf die Armee, zeitweise die Herrschaft über weite Teile Kontinentaleuropas. Er war ab 1805 auch König von Italien und von 1806 bis 1813 Protektor des Rheinbundes. Außerdem setzte er in einigen weiteren Staaten Familienmitglieder und Vertraute als Monarchen ein. Durch die von ihm eingeleitete Auflösung des Heiligen Römischen Reiches 1806 wurde die staatliche Gestaltung Mitteleuropas zu einer zentralen Frage im 19. Jahrhundert. Hatte er anfangs selbst noch den Nationalstaatsgedanken außerhalb Frankreichs verbreitet, erschwerte der Erfolg gerade dieses Gedankens besonders in Spanien, in Deutschland und schließlich auch in Russland die Aufrechterhaltung der napoleonischen Ordnung in Europa.

Napoleon Bonaparte: Französischer General, Erster Konsul, Kaiser, Reformator.

Für den laufenden Monat genutztes Quota: 1.8%

Anonymisieren der Auswahl Deanonymisieren der Auswahl Alles deanonymisieren Löschen

übersetzen auf Französisch :

zusammenfassen in 10 Wörter :

Napoleon Bonaparte, als Kaiser Napoleon I. (französisch Napoléon Bonaparte bzw. Napoléon Ier; * 15. August 1769 in Ajaccio auf Korsika als Napoleone Buonaparte; † 5. Mai 1821 in Longwood House auf St. Helena im Südatlantik) war General der Ersten Republik, Erster Konsul Frankreichs und schließlich Kaiser

Senden

Anonymisierung des
Prompts vor Versand
nur an "kommerzielle"
Modelle

- ChatGPT
- Gemini
- Claude

Deanonymisierung bei
der Rückkehr



KÜNSTLICHE INTELLIGENZ

INITIATIVEN IM GANGE: BGE-ARTIGE RÉSUMÉS !

Text Summarization Service

Uses custom made model to summarize the data

MAIN TEXT

Kürzlich hatte eine Untersuchung die Schweiz erschüttert. Historikerinnen und Historiker der Universität Zürich hatten darin 1002 sexuelle Missbrauchsfälle innerhalb der katholischen Kirche aufgedeckt. Ein Fall entfiel auf das Wallis.

Der Priester wurde wegen pädophiler Verbrechen verurteilt. Vor einigen Tagen erfolgte nun die Reaktion des Papstes. Im Gespräch mit «Le Nouvelliste» sagt Pierre-Yves Maillard (55), Generalvikar der Diözese Sitten: «Man muss daran erinnern, dass nur Rom befugt ist, ein solches Urteil am Ende eines kanonischen Prozesses zu fällen.»

Die zivilrechtlich verjährten Taten wurden in einem kanonischen Prozess geahndet, in dem der Angeklagte zur Rückkehr in den Laienstand, zur Entschädigung der Opfer, zum Rückzug aus der Tätigkeit, die er im «Centre Suisse Immigrés» und ähnlichen Vereinigungen ausübte, sowie zur Übernahme der Kosten des kanonischen Prozesses verurteilt wurde.

Summarize

Get Labels

SUMMARY

Art. 37 Abs. 4 OR; Sexueller Missbrauchs von Kindern durch einen katholischen Geistlichen. Der Papst hat keine Befugnis, anstelle einer Zivilsache eine kirchliche Gerichtsbarkeit einzusetzen.

Text Summarization Service

Uses custom made model to summarize the data

MAIN TEXT

Im Rahmen des Digital Services Act (DSA) nimmt die EU die jeweiligen App Stores von Google und Apple genauer unter die Lupe. Die EU-Kommission fordert von den Tech-Giganten, dass sie Informationen über das Risikomanagement mit der Regierung teilen, wie "Reuters" schreibt. Die Tech-Firmen haben nun bis zum 15. Januar Zeit, dieser Forderung nachzukommen. Anschliessend wird die Kommission basierend auf den eingegangenen Informationen abschätzen, wie hoch systematische Risiken bei den beiden Anbietern ausfallen.

Gemäss des DSA sind unter anderem manipulative Praktiken, welche Nutzer zu einem Kauf drängen sowie explizit auf Kinder ausgerichtete Werbung untersagt. Der Fragenkatalog der EU betrifft auch die Transparenz im Zusammenhang mit Empfehlungssystemen und Online-Werbung, wie es weiter heisst. Bei Verstössen gegen diese Richtlinien drohen den Anbietern Bussen von bis zu sechs Prozent des Jahresumsatzes.

Summarize

Get Labels

SUMMARY

Art. 23 GDPR; Art. 6 ICDRP; Digital Service Act (DSA). Europäischer Datenschutzbehörde hat die Tech-Unternehmen Google und Apple dazu auffordert, ihre Risikomanagementsysteme für ihre App Store in einer Liste einzureihen.

Text Summarization Service

Uses custom made model to summarize the data

MAIN TEXT

Uphold confidentiality and privacy Do not enter any information into a public AI chatbot that is not already in the public domain. Do not enter information which is private or confidential. Any information that you input into a public AI chatbot should be seen as being published to all the world. The current publicly available AI chatbots remember every question that you ask them, as well as any other information you put into them. That information is then available to be used to respond to queries from other users. As a result, anything you type into it could become publicly known. You should disable the chat history in AI chatbots if this option is available. This option is currently available in ChatGPT and Google Bard but not yet in Bing Chat. Be aware that some AI platforms, particularly if used as an App on a smartphone, may request various permissions which give them access to information on your device. In those circumstances you should refuse all such permissions. In the event of unintentional disclosure of confidential or private information you should contact your leadership judge and the Judicial Office.

Summarize

Get Labels

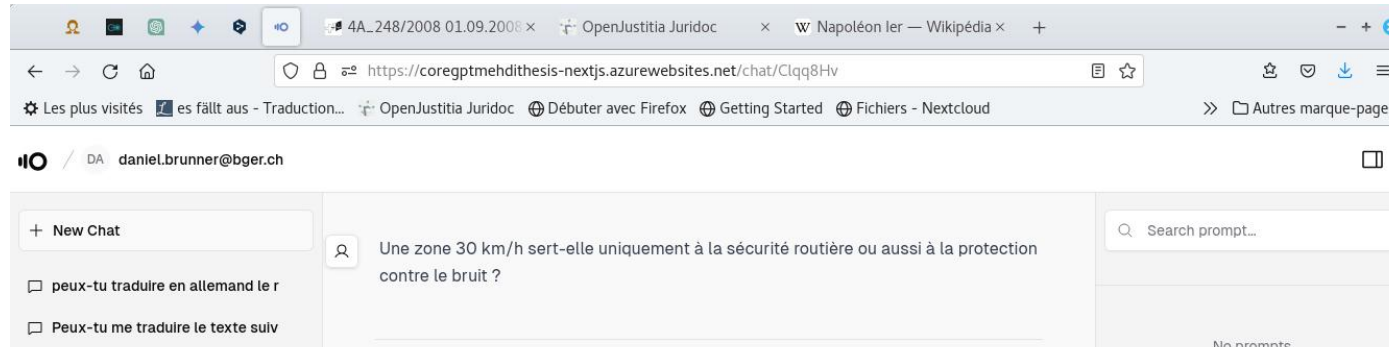
SUMMARY

Artificial Intelligence (AI) Chatbot; Datenschutz und Datensicherheit.

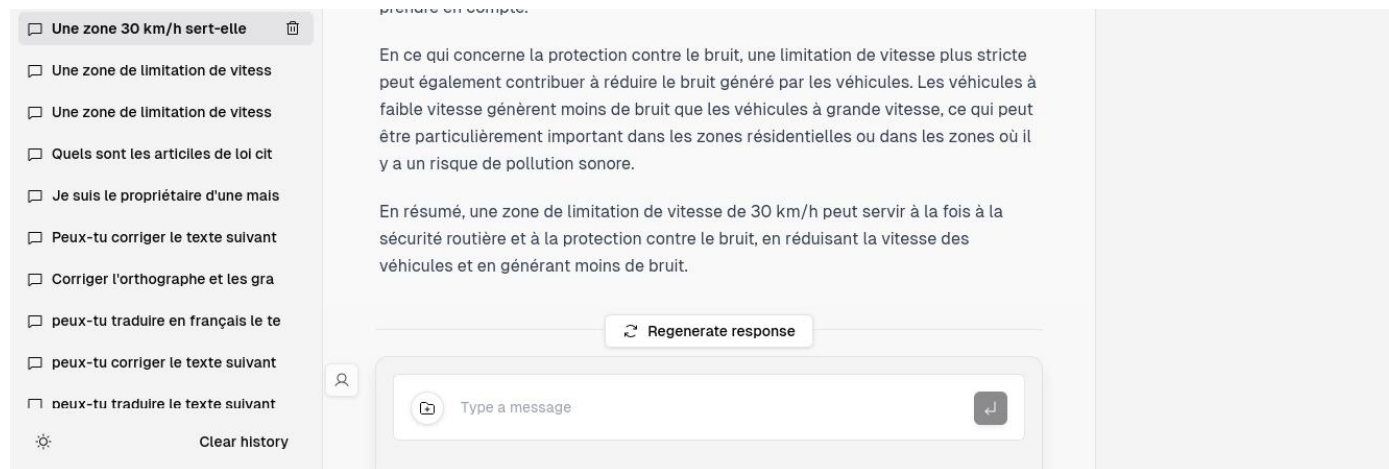


KÜNSTLICHE INTELLIGENZ

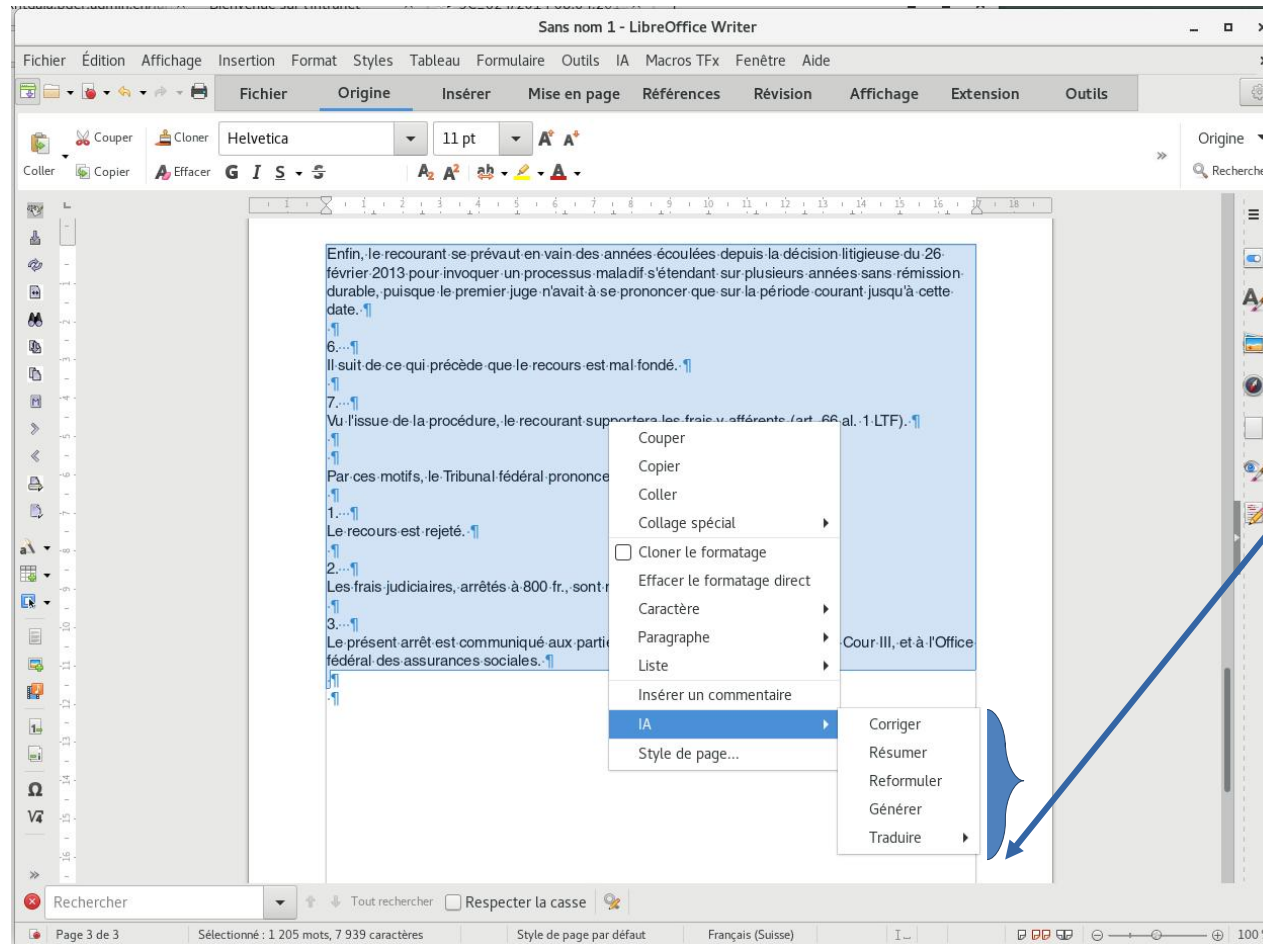
LAUFENDE INITIATIVEN: DISTILLED-CHAT-TF



Die Antworten sind etwas allgemein, wenig präzise, teils auch eher unverbindlich. Das mag auch daran liegen, dass die Fragen eher allgemein gehalten waren. Für einen professionellen Gebrauch wäre wünschbar, dass die Antwort die entsprechenden Quellen (Rechtsgrundlagen, Rechtsprechung etc.) bezeichnet. Bei der Frage zu den Fruchtfolgeflächen (FFF) wird auf französisch der Begriff surfaces d'assolement (SDA) statt surfaces de rotation culturale verwendet. Vielleicht ergäbe das auch eine andere Antwort.



KÜNSTLICHE INTELLIGENZ INITIATIVEN IM GANGE: LIBREOFFICE

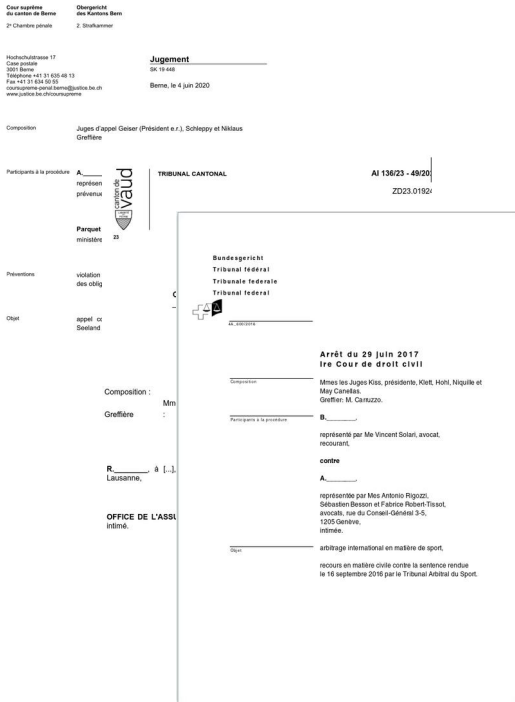


**Intégration eines
OSS-Sprachmodells für das
Ausführen von
Standardaufgaben!**



SPRACHMODELL FÜR DIE SCHWEIZER JUSTIZ

DIE BASIS FÜR EINE VERTRAUENSVOLLERE KI



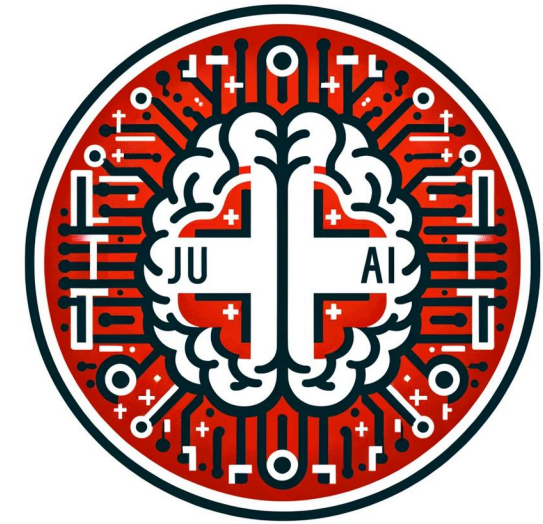
Lehre/Doktrin



Entscheide



SWISS AI Initiative



Swiss Justice Base Model
SJBM



FRAGEN?



Kündigung des Leiters des Superalignment Teams bei OpenAI !

Das Superalignment-Team wurde im vergangenen Juli gegründet und **sollte sich mit den Risiken besonders hoch entwickelter KI-Systeme befassen**

Jan Leike ✓ · 17. Mai 2024
@janleike · Folgen
Antwort an @janleike
Building smarter-than-human machines is an inherently dangerous endeavor.

OpenAI is shouldering an enormous responsibility on behalf of all of humanity.

Jan Leike ✓
@janleike · Folgen

But over the past years, safety culture and processes have taken a backseat to shiny products.

5:57 nachm. · 17. Mai 2024

2.922 Antworten Link kopier.

59 Antworten lesen

