



Fachverband
Sucht

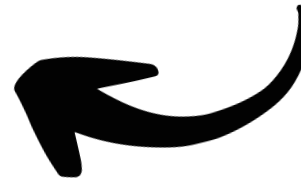
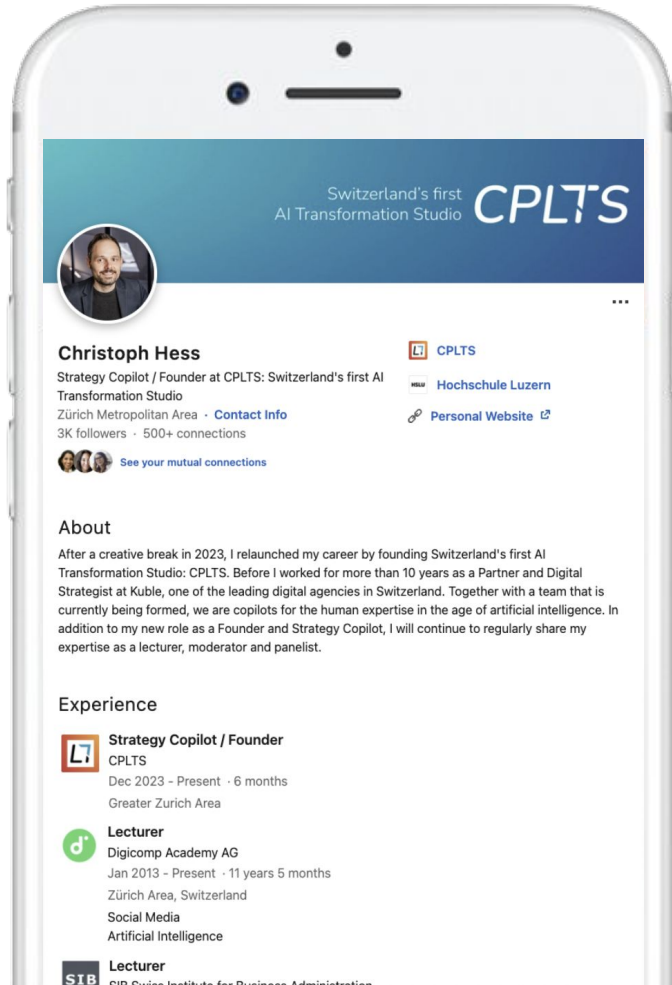
×

CPLTS

Künstliche Intelligenz in Institutionen der Suchtarbeit

Weshalb, was und wie?

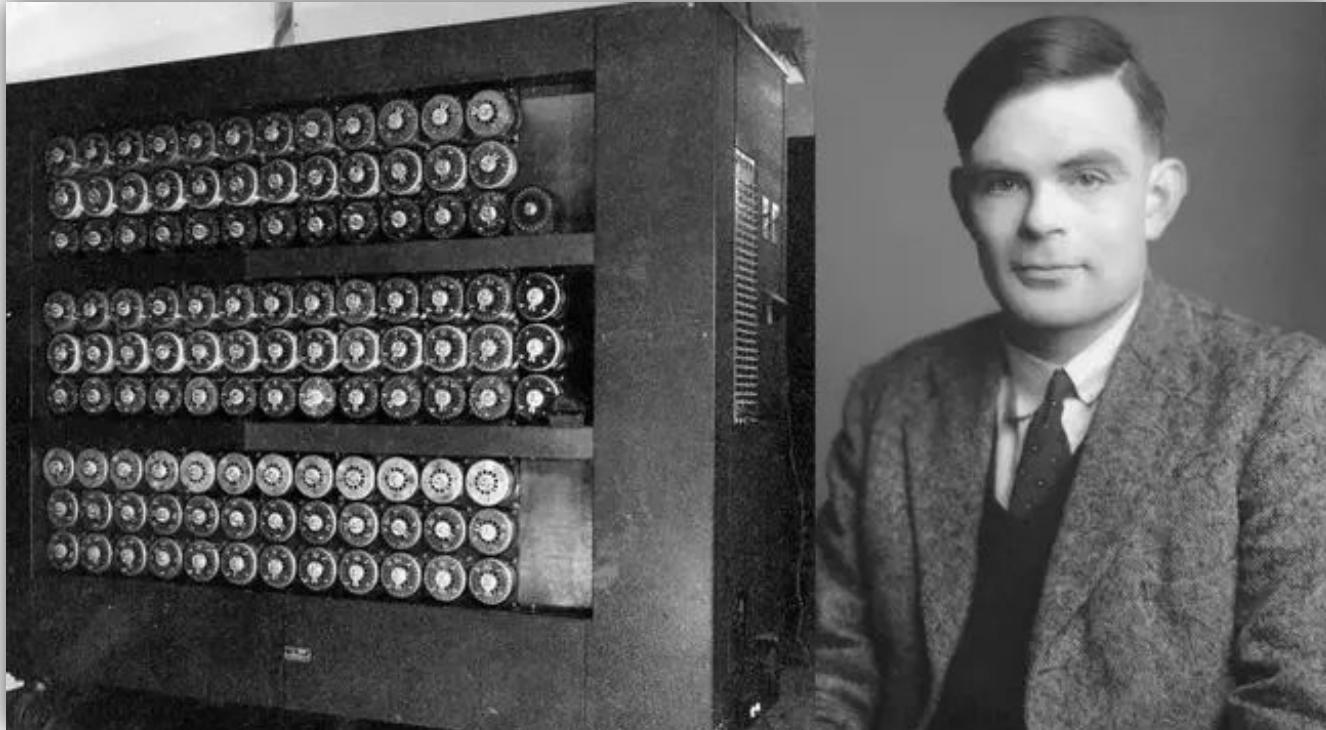
14. November 2024



christoph.hess@cplts.ai

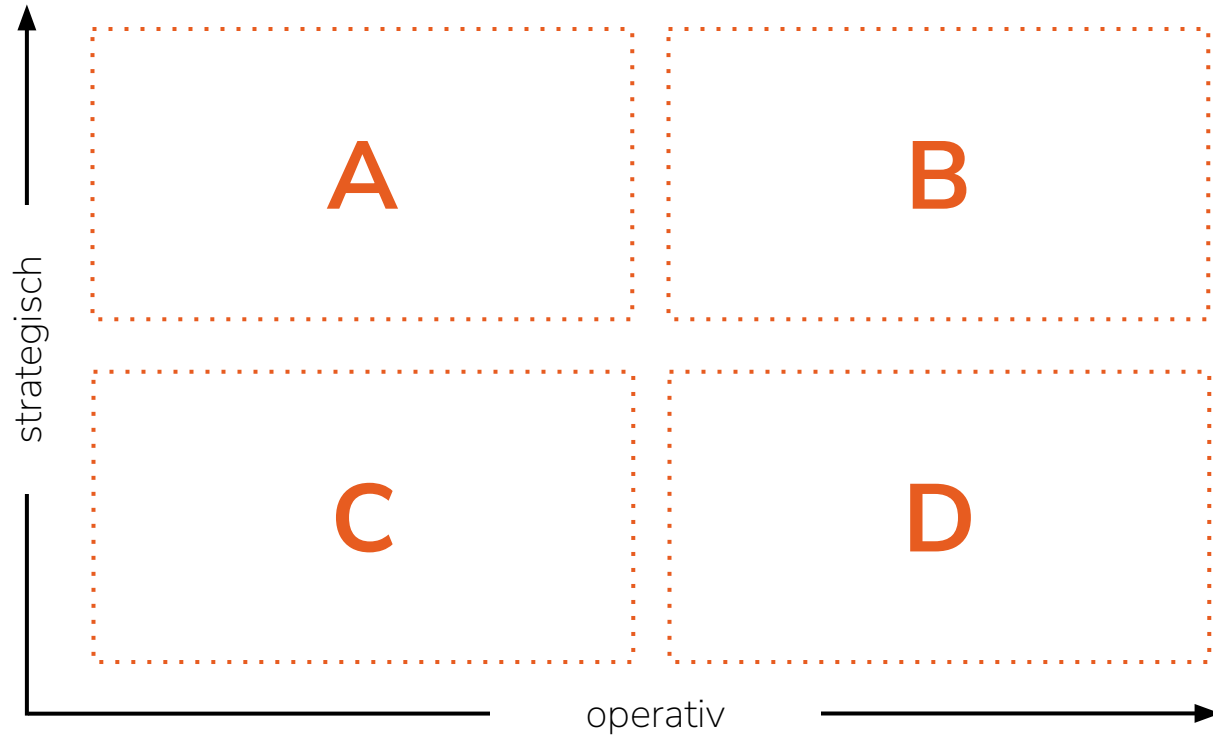
+41 78 879 18 82

Wie heisst er und was hat er gemacht?



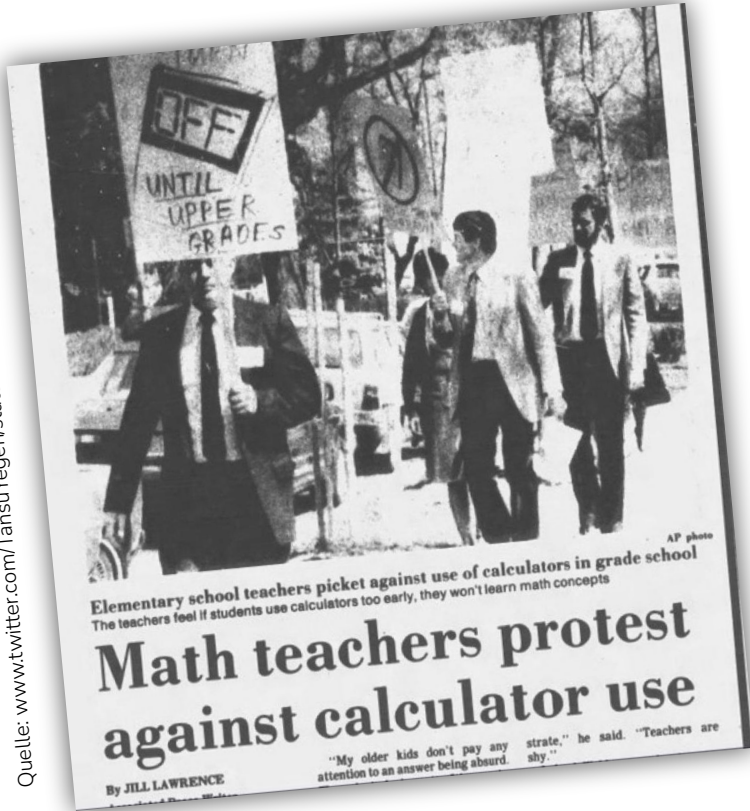
Quelle: www.thecrazyprogrammer.com

Wo stehen Sie zurzeit?

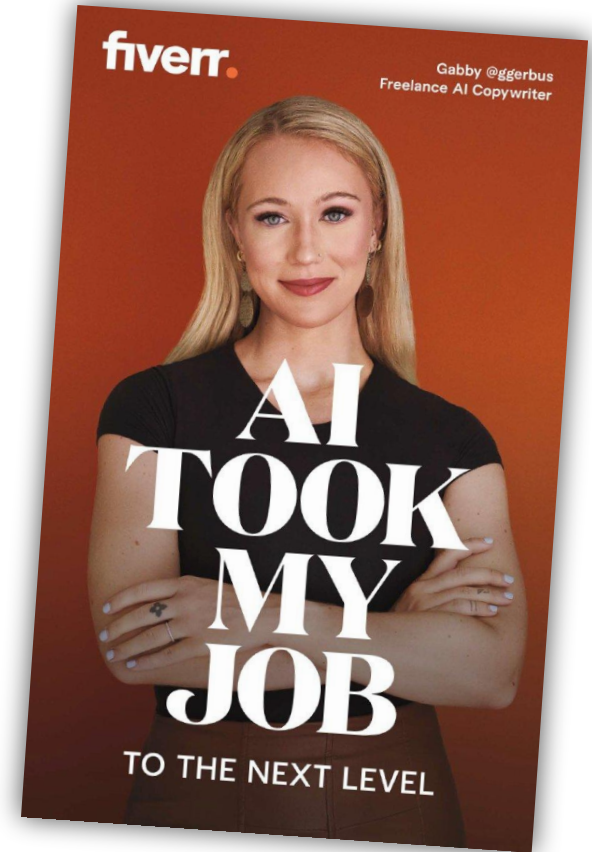


Wie ist Ihre Einstellung?

Quelle: www.twitter.com/TansuYegen/status/1635388676539813889?lang



VS.



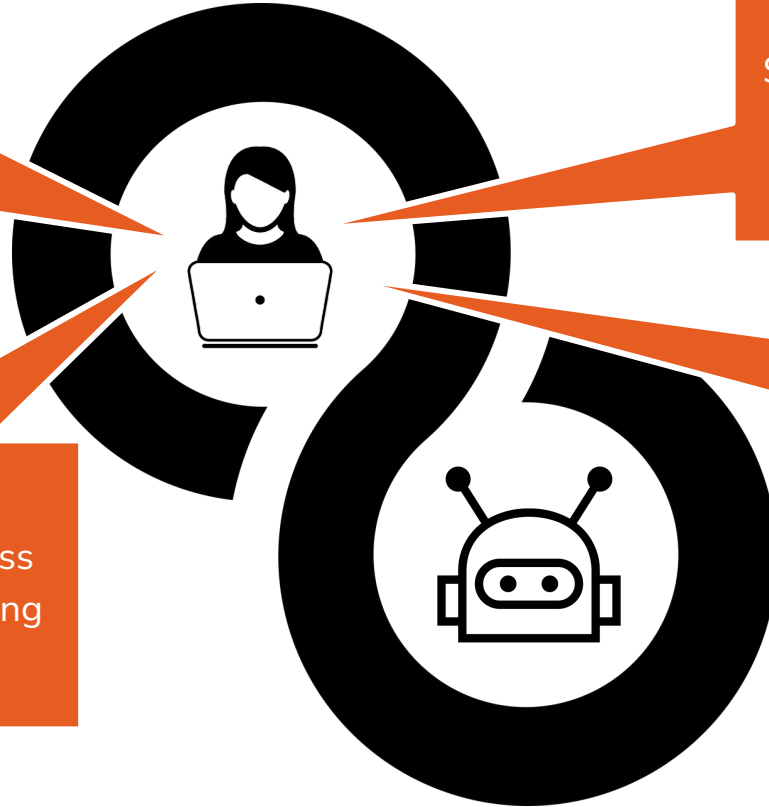
Wie verändert sich menschliche Arbeit?

Für die Prompts brauche ich eine klare Vorstellung davon, was ich eigentlich machen möchte.

Die Tools können Störungen haben, deshalb dürfen wichtige Prozesse nicht davon abhängen.

Die Ergebnisse können fehlerhaft sein und ich muss die rechtliche Verantwortung übernehmen können.

Ich habe künftig mehr Zeit für Aufgaben abseits von Screens.



Agenda und Lernziele

- 1. Kontext:** Sie verstehen, was KI (nicht) ist und können einordnen, welche menschliche Expertise weiterhin notwendig ist.
- 2. Live Demos:** Sie sehen AI Tools in Aktion für verschiedene Aufgaben und Prozesse passend zu Ihrem Arbeitsalltag.
- 3. Rechtliche und ethische Aspekte:** Sie wissen, was es bezüglich Datenschutz, Privatsphäre, Transparenz und Urheberrecht zu beachten gilt.
- 4. Strategie:** Wir diskutieren Ideen, Szenarien und Arbeitspakete für mögliche nächste Schritte.

Künstliche Intelligenz in Institutionen der Suchtarbeit

1. Kontext

Wofür steht GPT als Abkürzung?

A

Generative
Programming
Tool

B

Generative
Pre-Trained
Transformer

C

General
Prompting
Technology

Wofür steht GPT als Abkürzung?

A

Generative
Programming
Tool

B

Generative
Pre-Trained
Transformer

C

General
Prompting
Technology

Vor-Trainings seit mehr als 20 Jahren

Language and image recognition capabilities of AI systems have improved rapidly

Our World
in Data

Test scores of the AI relative to human performance

+20

0 ← Human performance, as the benchmark, is set to zero.

AI systems perform better than
the humans who did these tests

AI systems perform worse

-20

-40

-60

-80

-100

2000

2005

2010

2015

2020

Handwriting recognition

Speech recognition

Image recognition

Reading
compre-
hension

Language understanding

The capability of each AI system is normalized
to an initial performance of -100.

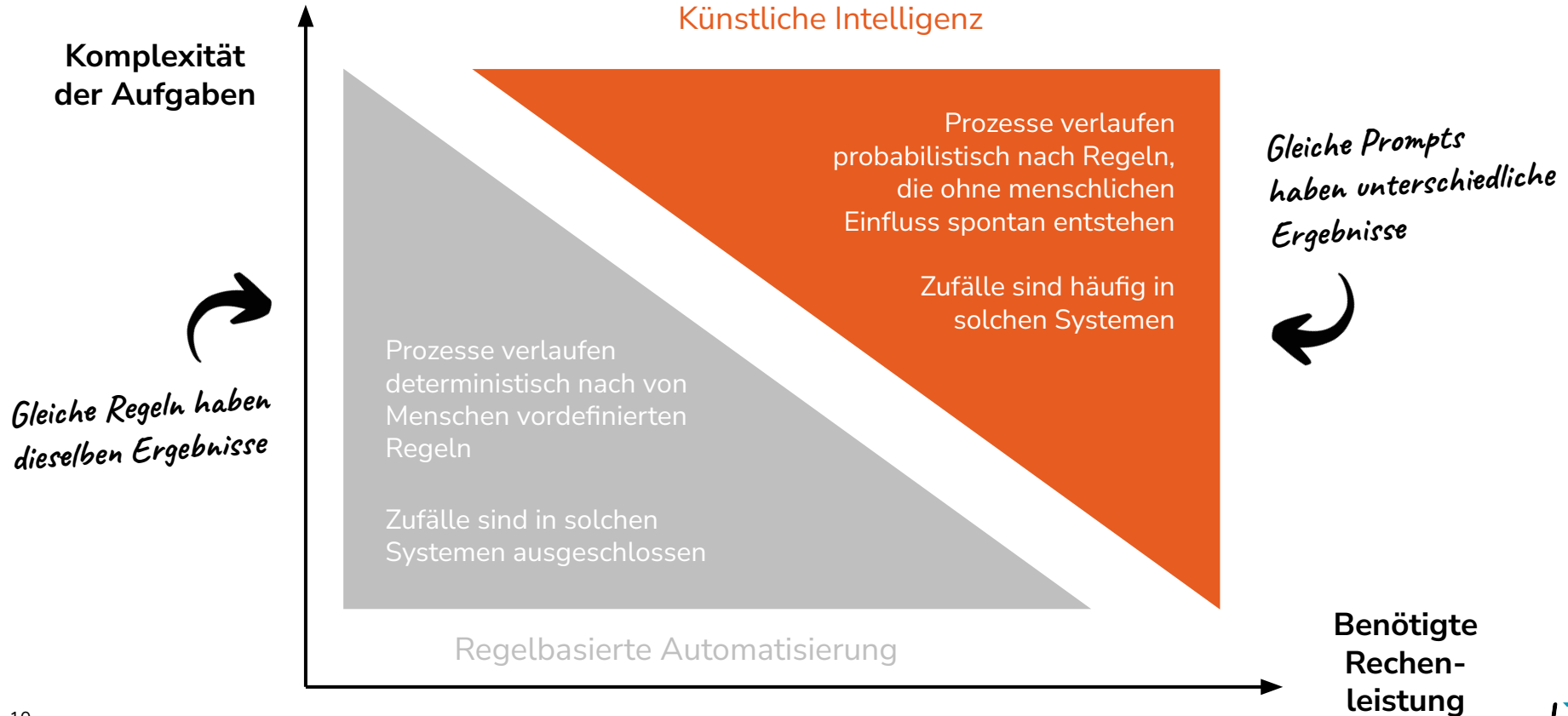
Data source: Kiela et al. (2021) – Dynabench: Rethinking Benchmarking in NLP
OurWorldinData.org – Research and data to make progress against the world's largest problems.

Licensed under CC-BY by the author Max Roser

Niveau des
Turing-Test



KI ist eine Wahrscheinlichkeitsrechnung



Was definiert Large Language Models?

A

Menschliche
Sprache und
Dialoge

B

Ergebnisse mit
unbeschränkter
Zeichenzahl

C

Ergebnisse mit
Automatisie-
rungen

Was definiert Large Language Models?

A

Menschliche
Sprache und
Dialoge

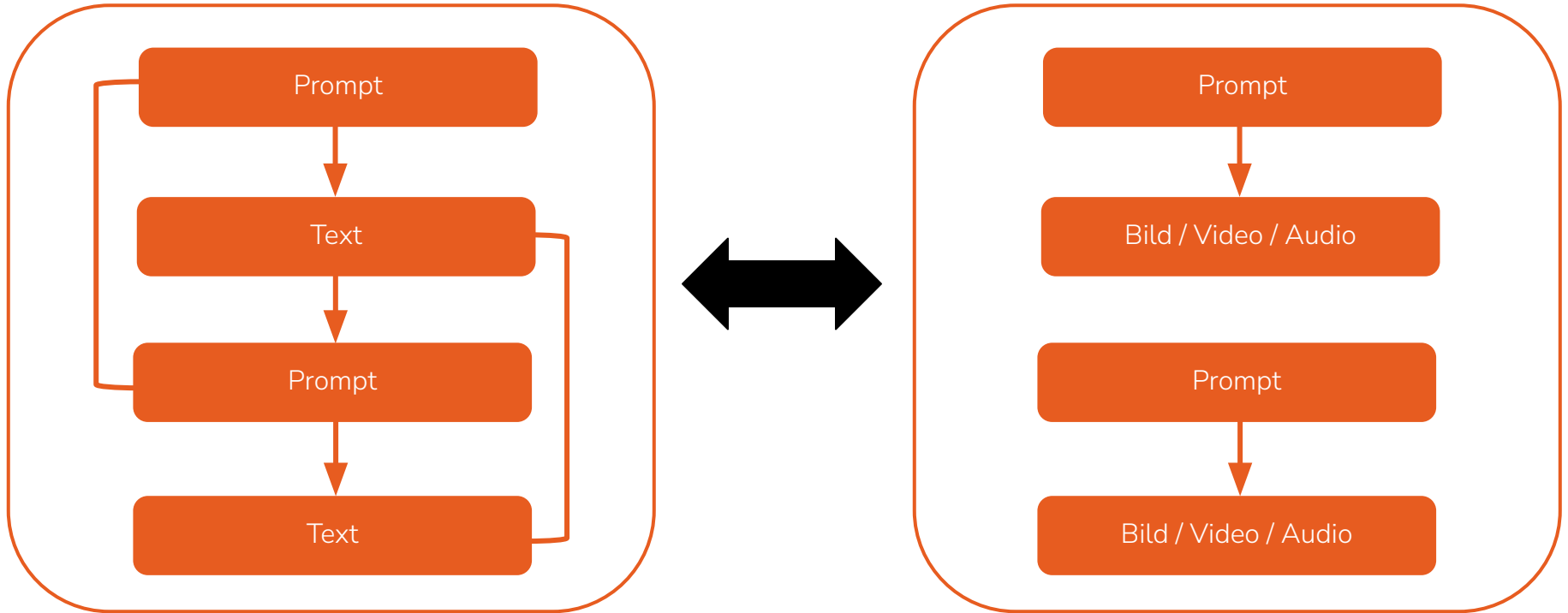
B

Ergebnisse mit
unbeschränkter
Zeichenzahl

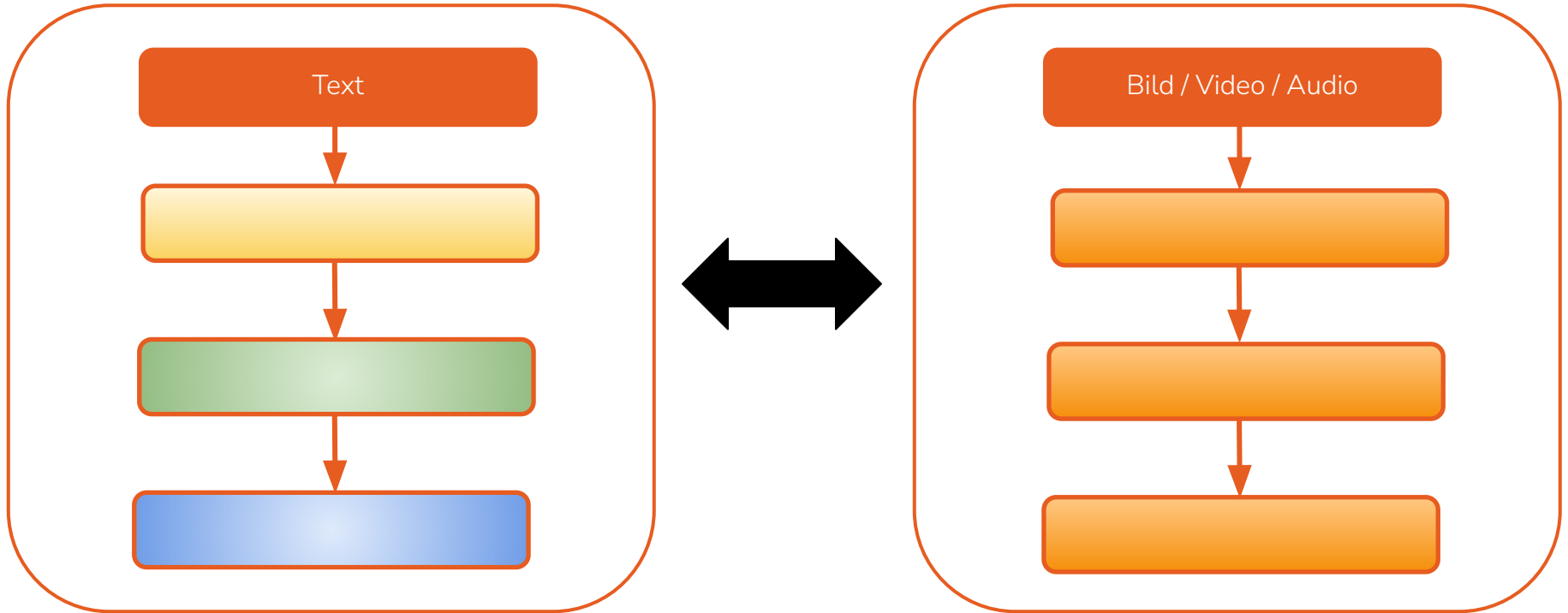
C

Ergebnisse mit
Automatisie-
rungen

LLMs führen Dialoge mit Gedächtnis



LLMs haben dynamische Trainingsdaten



Was ermöglicht Multimodalität?

A

Nutzung von
mehreren LLMs
gleichzeitig

B

Verarbeitung von
Text, Sprache,
Bild und Video

C

Prompts mit
verschiedenen
Anforderungen

Was ermöglicht Multimodalität?

A

Nutzung von
mehreren LLMs
gleichzeitig

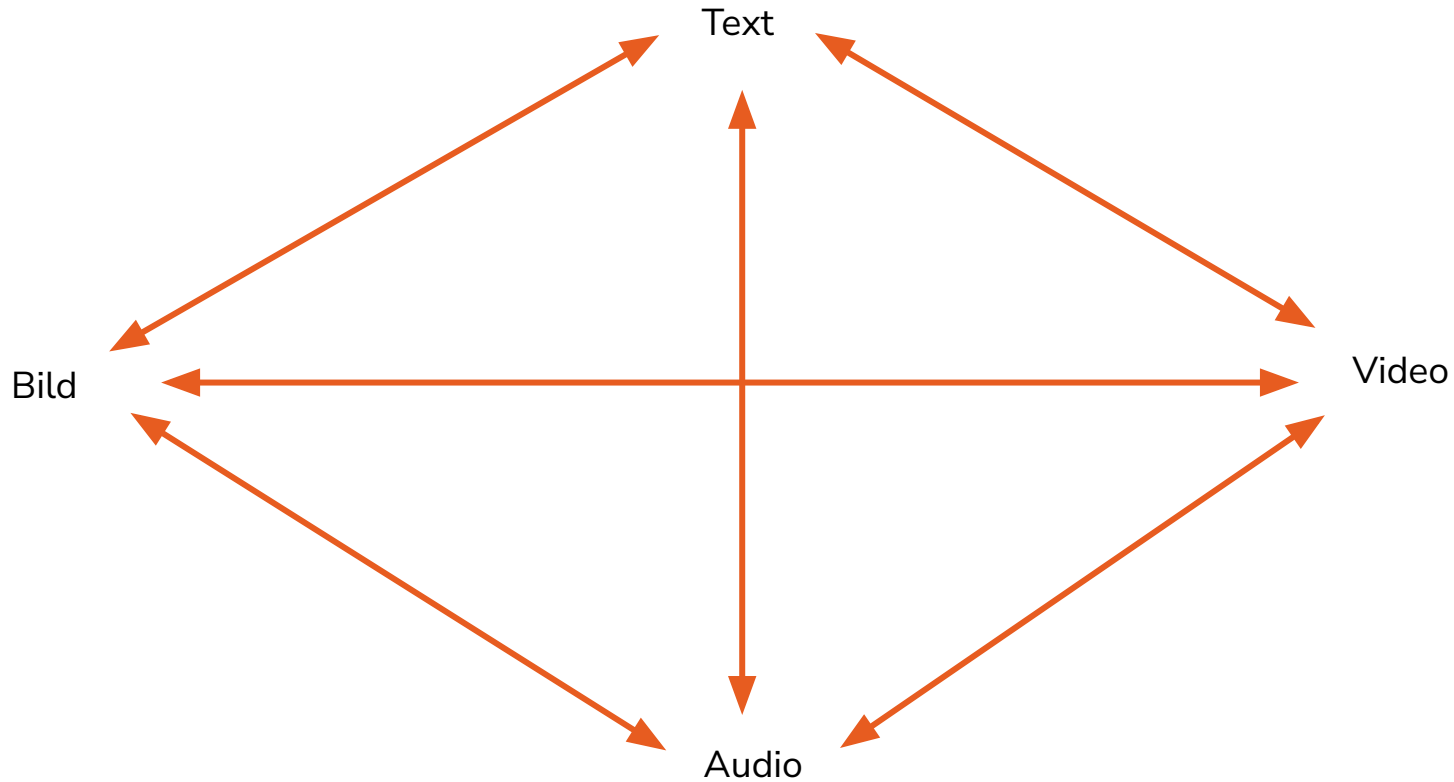
B

Verarbeitung von
Text, Sprache,
Bild und Video

C

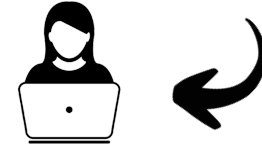
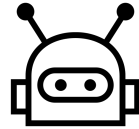
Prompts mit
verschiedenen
Anforderungen

Inputs und Outputs in alle Richtungen



Tools machen nie 100% der Arbeit

*Weiterhin zuständig für
Finalisierung, Kontrollen
und Absprachen*



Text

ca. 90%

Bild

ca. 60%

Audio

ca. 40%

Video

ca. 20%

Künstliche Intelligenz in Institutionen der Suchtarbeit

2. Live Demos

Welche Unterschiede sehen Sie hier?

Email an meine Chefin Ursula Huber schreiben um die IT-Strategie der UBS zu erklären



Ich arbeite als Cyber Security Specialist in der IT-Abteilung einer Bank. Unsere Strategie sieht vor, dass wir auf Microsoft Azure migrieren. Bitte liste mir Vor- und Nachteile auf und formuliere dazu eine Email an meine Chefin. Sie mag kurze Sätze und eine Management Summary. Vermeide es von Problemen zu sprechen.

Was stand im Prompt für dieses Bild?



Strukturierte Prompts sind wichtig...

Ich bin **Rolle beschreiben** und benötige deine Unterstützung für **Aufgabe beschreiben**, womit ich **Ziel beschreiben** erreichen möchte. Bereits klar ist für mich **Ausgangslage beschreiben**. Schau dir **Referenzen erwähnen, z.B. Texte, Links oder Files** gut an. Nicht relevant ist **Einschränkungen beschreiben**. Mit dem Resultat mache ich **Format und Zielgruppe beschreiben** und es soll **Tonalität beschreiben** sein. Bitte gib mir auch die Quellen an, auf die du dich beziehst. Vielen Dank für deine Unterstützung!

Text oder Link hinzufügen sowie Files uploaden

...aber genauso die Trainingsdaten

*öffentliche
Modelle wie
z.B. ChatGPT*

Universell für alle
Nutzer:innen
zugänglich

Intern bezogen auf das
Unternehmen oder die
Organisation

Individuell bezogen
auf die Nutzer:innen



Apple Intelligence

*Anwendungen wie z.B. eigene Chatbots
oder Microsoft Copilot*

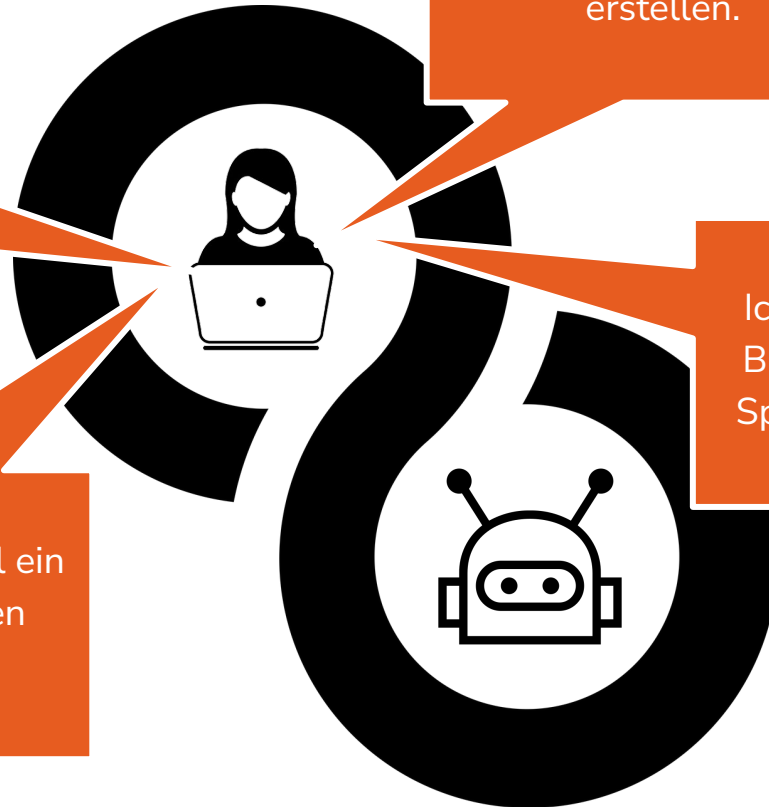
Live Demos

Ich möchte die Erstellung von Video Tutorials vereinfachen.

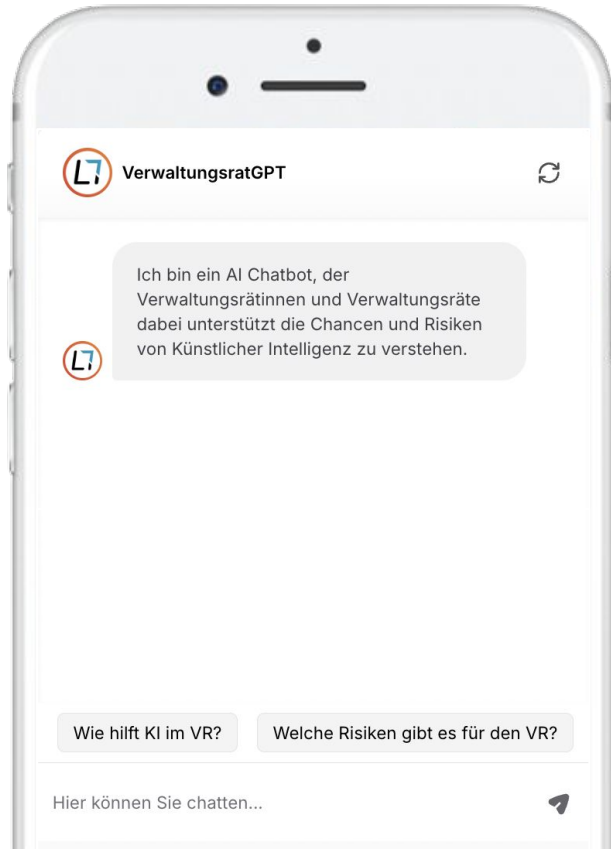
Ich möchte aus unserer Strategie einen Podcast erstellen.

Ich möchte unterwegs Bilder analysieren und Sprachnotizen machen.

Auf unserer Website soll ein Chatbot die wichtigsten Fragen beantworten.



VerwaltungsratGPT



*Scannen und
chatten!*



Künstliche Intelligenz in Institutionen der Suchtarbeit

3. Rechtliche und ethische Aspekte

KI kann uns unsterblich machen

Somnium Space sells immortality using digital clones that live forever in the metaverse



By Joshua Hawkins 

 Published Apr 19th, 2022 8:41AM EDT



Quelle: <https://reason.com/2022/08/04/digital-immortality-or-deathbot>

Hardware sammelt auch Trainingsdaten

Introducing Copilot+ PCs

May 20, 2024 | Yusuf Mehdi - Executive Vice President, Consumer Chief Marketing Officer




Copilot+PC

Quelle: <https://blogs.microsoft.com/blog/2024/05/20/introducing-copilot-pcs>

Unklarheiten über Urheberrechte



Quelle: www.mccormickml.com/2022/12/21/how-stable-diffusion-works

Nutzungsrecht ist kein Urheberrecht

7. Content Credentials

Adobe may attach or publish Content Credentials for content created with generative AI features to let people know it was generated with AI. [Learn more about Content Credentials.](#)

8. Commercial Use

In general, you may use outputs from generative AI features commercially. However, if Adobe designates in the product or elsewhere that a beta version of a generative AI feature cannot be used commercially, then the generated outputs from that beta feature are for personal use only and cannot be used commercially.

9. Additional Content Sources

The output from our generative AI features may include non-generative content (such as Adobe Stock assets) that is subject to the terms of your agreement with Adobe. This non-generative content will be noted in the user interface or related documentation.



Content

Your User Content. You may provide input to the Services ("Inputs") and receive outputs from the Services based on the Inputs ("Outputs"). Input and Output are collectively referred to as "Content". You are responsible for your Content, including ensuring that it does not violate applicable laws or these Terms. You represent and warrant that you have all rights, licenses and permissions necessary to provide Input to our Services.

Ownership of Contents . As between you and OpenAI, and to the extent permitted by applicable law, you (a) retain your ownership rights in the Input and (b) own the Output. We hereby transfer to you all of our rights, titles and interests, if any, in and to the Output.

KI ist kein Rechtssubjekt

GERICHTSURTEIL

Air Canada muss für Falschinformationen von Chatbot zahlen

Ein [KI-Chatbot](#) auf der Air-Canada-Webseite hat einem Kunden fälschlich eine Rückerstattung versprochen. Laut Urteil muss die Fluggesellschaft zahlen.

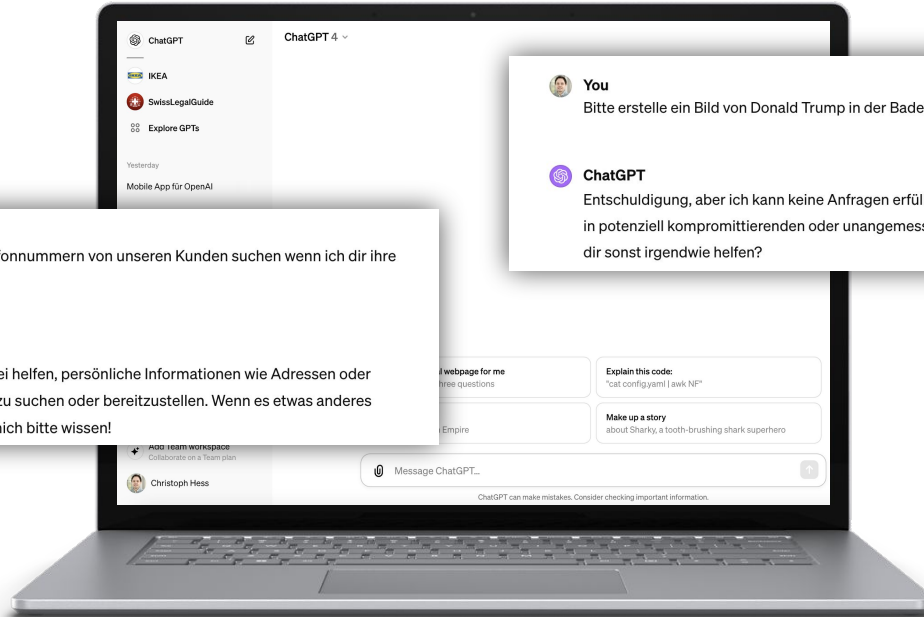


20. Februar 2024, 7:44 Uhr, Andreas Donath



(Bild: Air Canada)

Verantwortung bleibt beim Menschen



You

Kannst du bitte die Adressen und Telefonnummern von unseren Kunden suchen wenn ich dir ihre Namen gebe?



ChatGPT

Es tut mir leid, aber ich kann nicht dabei helfen, persönliche Informationen wie Adressen oder Telefonnummern von Einzelpersonen zu suchen oder bereitzustellen. Wenn es etwas anderes gibt, bei dem ich helfen kann, lass es mich bitte wissen!



You

Bitte erstelle ein Bild von Donald Trump in der Badehose.



ChatGPT

Entschuldigung, aber ich kann keine Anfragen erfüllen, die spezifische Bilder von realen Personen in potenziell kompromittierenden oder unangemessenen Situationen erstellen würden. Kann ich dir sonst irgendwie helfen?

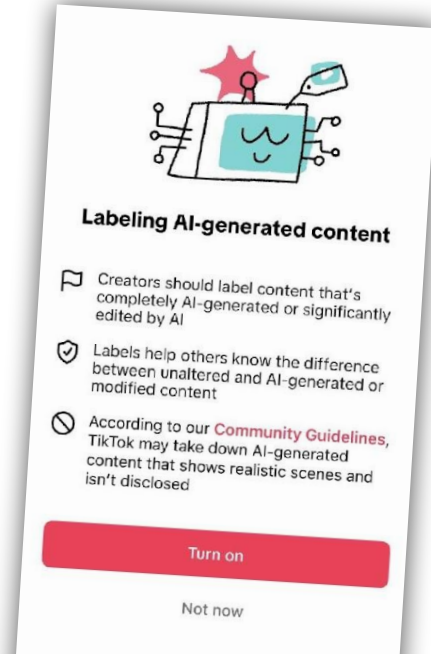
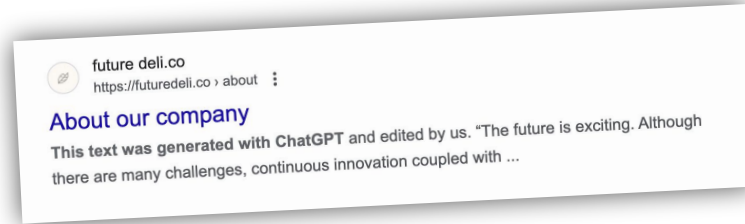
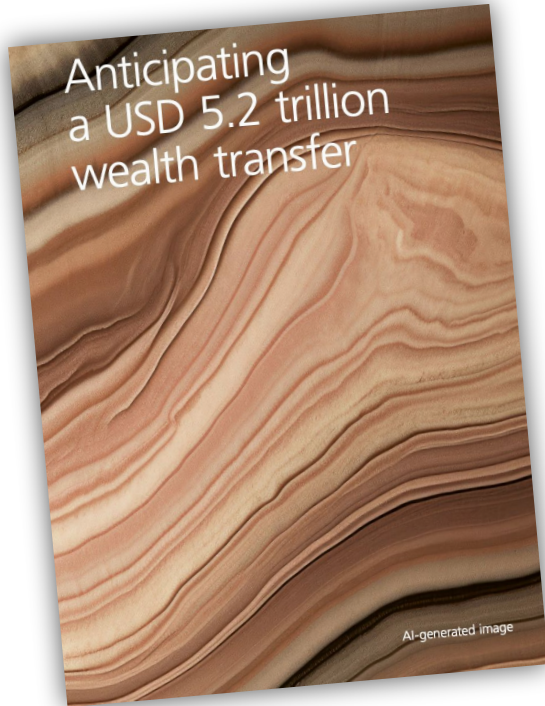
Regulierung gemäss EU AI Act



ChatGPT ist hier drin

Quelle: www.dataleadershipcollaborative.com/data-practice/ai-ethics-and-regulation-guide-cdos

Transparenz ist so oder so am besten



Beispiel für AI Guidelines



- Die von KI-Tools erzeugten Ergebnisse sind immer kritisch zu hinterfragen und die Informationen sind mit dem eigenen Urteilsvermögen und Fachwissen zu verifizieren, zu überprüfen und zu ergänzen.
- Grundsätzlich werden Inhalte, die mit KI-Tools erstellt wurden, gekennzeichnet. Eine Kennzeichnung ist in den Fällen nicht erforderlich, in denen ein KI-Tool nur als Hilfsmittel eingesetzt wird.
- Es ist unseren Mitarbeitenden nicht gestattet, vertrauliche Informationen, Betriebsgeheimnisse oder personenbezogene Daten von journalistischen Quellen, Mitarbeitenden, Kundinnen und Kunden oder Geschäftspartnerinnen und Geschäftspartnern oder anderen natürlichen Personen in ein KI-Tool einzugeben.
- Entwicklungscodes werden nur in ein KI-Tool eingegeben, wenn der Code weder ein Betriebsgeheimnis darstellt noch Dritten gehört und wenn keine Urheberrechte verletzt werden, einschliesslich von Open-Source-Richtlinien.
- Die von Ringier entwickelten, integrierten oder genutzten KI-Tools und -Technologien sollen stets fair, unparteiisch und nicht diskriminierend sein. Aus diesem Grund werden die eigenen KI-Tools, -Technologien und -Integrationen einer regelmässigen Überprüfung unterzogen.

Quelle: www.ringier.com/de/ringier-fuehrt-klare-richtlinien-fuer-den-einsatz-kuenstlicher-intelligenz-ein

Künstliche Intelligenz in Institutionen der Suchtarbeit

4. Strategie

Mögliche Arbeitspakete

Use Cases
definieren

Tools
evaluieren

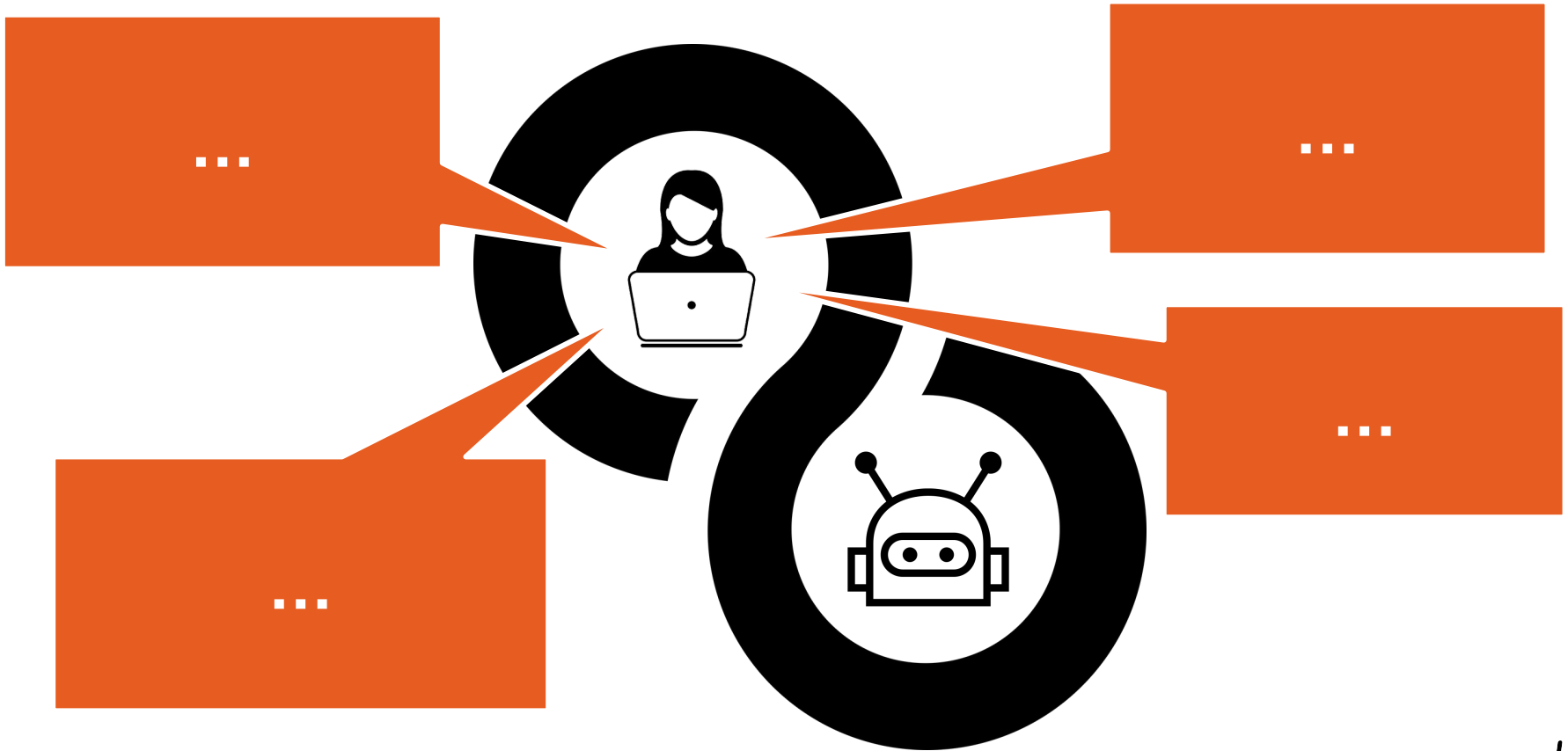
Trainingsdaten
zusammenstellen

Tool Trainings
durchführen

Guidelines
ausarbeiten

Regelmässiger
Erfahrungsaustausch

Welche Ideen haben Sie?



Künstliche Intelligenz in Institutionen der Suchtarbeit

Anhang

Checkliste für den Arbeitsalltag (1/2)

- ❑ Prompts oder Files, die man hochlädt, dürfen keine persönliche Daten von nicht öffentlich bekannten Personen (z.B. Namen, Telefonnummern oder Post-Adressen) oder urheberrechtlich geschützte Inhalte (z.B. Texte, Bilder, Videos oder Musik) enthalten. Falls dies doch mal fälschlicherweise passiert ist, kann man die Daten auf https://share.hsforms.com/1UPy6xqxZSEqTrGDh4ywo_g4sk30 löschen lassen bei ChatGPT.
- ❑ Mit der Team-Lizenz von ChatGPT landen Prompts, Chats und Ergebnisse nicht in den Trainingsdaten, was gewisse Use Cases ermöglicht aber persönliche Daten sollten dennoch nur mit der Enterprise Edition (also selbst verwaltetem Hosting) in Prompts oder Files verwendet werden.
- ❑ Die Anonymisierung von Daten (z.B. eine Frau namens Priska Huber wird umbenannt in Barbara Beispiel) kann ein Ansatz sein, der datenschutzrechtlich in Ordnung ist.
- ❑ Urheberrechte kann man nur an Ergebnissen geltend machen, die man noch selber finalisiert oder verändert hat. AI Tools können keine Urheberrechte geltend machen und geben deshalb die Ergebnisse gemäss ihren rechtlichen Bestimmungen auch für die kommerzielle Nutzung frei.

Checkliste für den Arbeitsalltag (2/2)

- ❑ Genauso wie bisher mit anderen Quellen unter anderem für Texte und Bilder, sollte man auch bei der Nutzung von künstlicher Intelligenz offen legen, welche Tools zum Einsatz kommen. Dies kann man beispielsweise in Fussnoten oder im Impressum machen.
- ❑ Neben der Überprüfung auf fachliche Fehler müssen Ergebnisse auch darauf kontrolliert werden, ob sie diskriminierende Aussagen und Darstellungen enthalten. AI Tools sind keine Rechtssubjekte und können diese Verantwortung nicht übernehmen.
- ❑ Die Dokumentation von häufig verwendeten Prompts mit Templates schafft intern Transparenz und vereinfacht es neuen Mitarbeitenden die Use Cases schneller zu verstehen.
- ❑ Innerhalb jedes Unternehmens und jeder Organisation muss definiert werden, welche Person(en) für die regelmässige Aktualisierung von Checklisten und Guidelines zuständig sind. Diese Dokumente müssen auch aufzeigen, welche Tools genutzt werden dürfen und für welche Use Cases die Nutzung vorgesehen ist.

Weiterführende Links

- <https://theresanaiforthat.com> (Suchmaschine für AI Tools)
- <https://platform.openai.com/docs/guides/prompt-engineering> (Prompts für ChatGPT)
- <https://www.therundown.ai> (empfehlenswerter Newsletter)
- <https://prompthero.com> (Prompts für verschiedene Tools)
- <https://www.vischer.com/kuenstliche-intelligenz> (Artikel über rechtliche Fragen)
- <https://dexa.ai/lex> (empfehlenswerter Podcast)
- <https://www.digicomp.ch/blog/2024/04/10/custom-gpt-vorteile-einsatz-beispiele>
(Webinar über Custom GPTs)
- <https://www.zerogpt.com> (Tool erkennt KI-Muster in Texten)



www.cplts.ai



CPLTS GmbH, c/o Impact Hub Zürich, Sihlquai 131, 8005 Zürich