



Berner
Fachhochschule

Künstliche Intelligenz: Von der Theorie zur Praxis

Winterkongress 2024


Prof. Dr. Mascha Kurpicz-Briki

Applied Machine Intelligence

Berner Fachhochschule, Biel

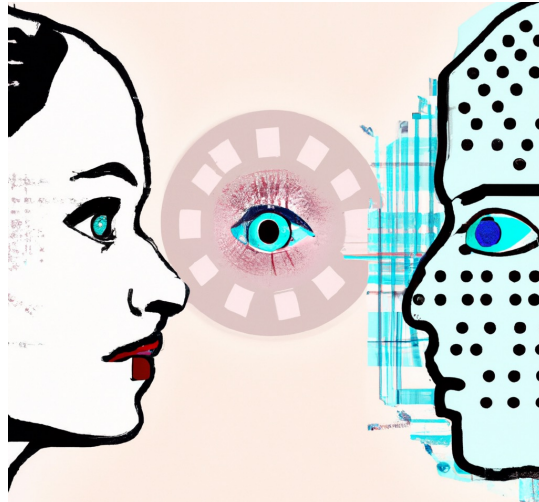
<http://www.bfh.ch/ami>

mascha.kurpicz@bfh.ch

 /maschakurpicz

Generative AI

Viele neue Möglichkeiten durch
schnellen technologischen
Fortschritt



Created by Mascha&DALL-E

Sprachmodelle, Anwendungen
wie ChatGPT, etc.

ChatGPT4

«Generiere ein Bild zu dem Thema:
Wir möchten eine nutzbringende
Anwendung der Technik.»

Created by Mascha&ChatGPT4



Traditionelle Software

Zutaten + spezifische Anleitung =
Resultat



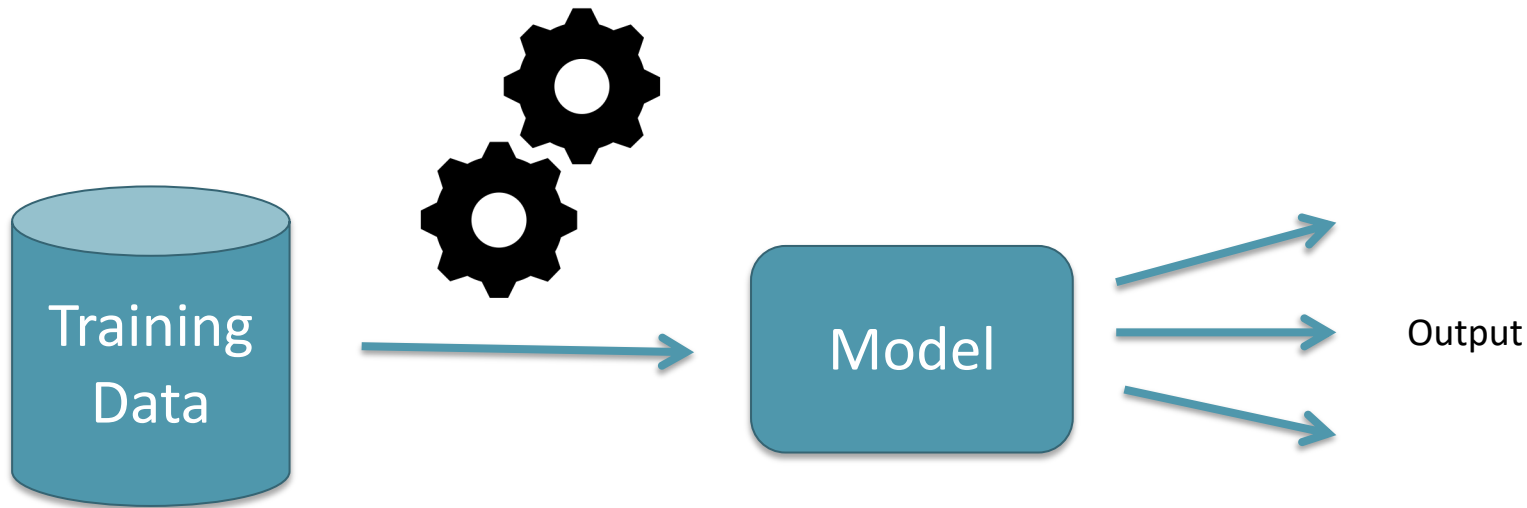
Artificial Intelligence

„machine learning“



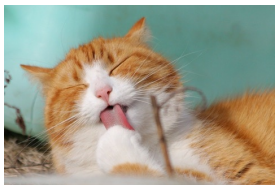
Image Sources: pixabay/stevepb, pixabay/NoName_13

Maschinelles Lernen



Icon Source: Microsoft PowerPoint

Katze



Katze



...

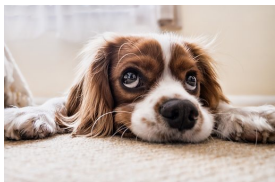


??



Katze

Hund



Hund

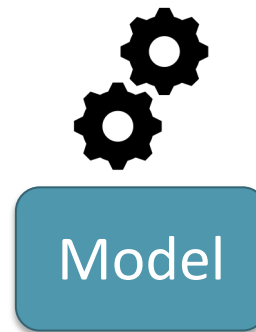
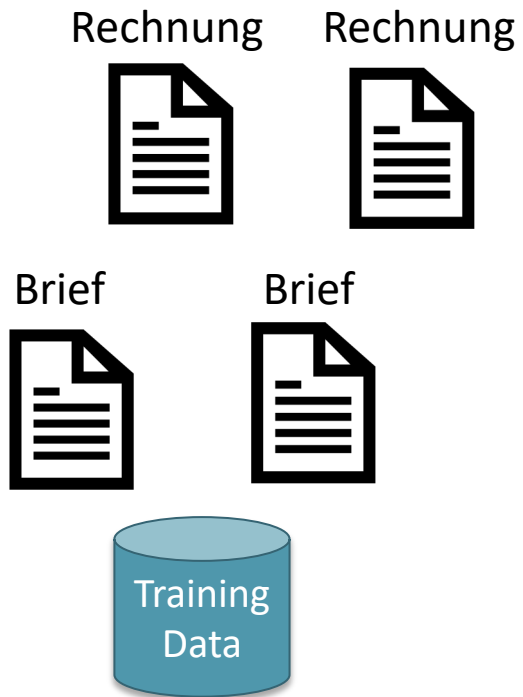
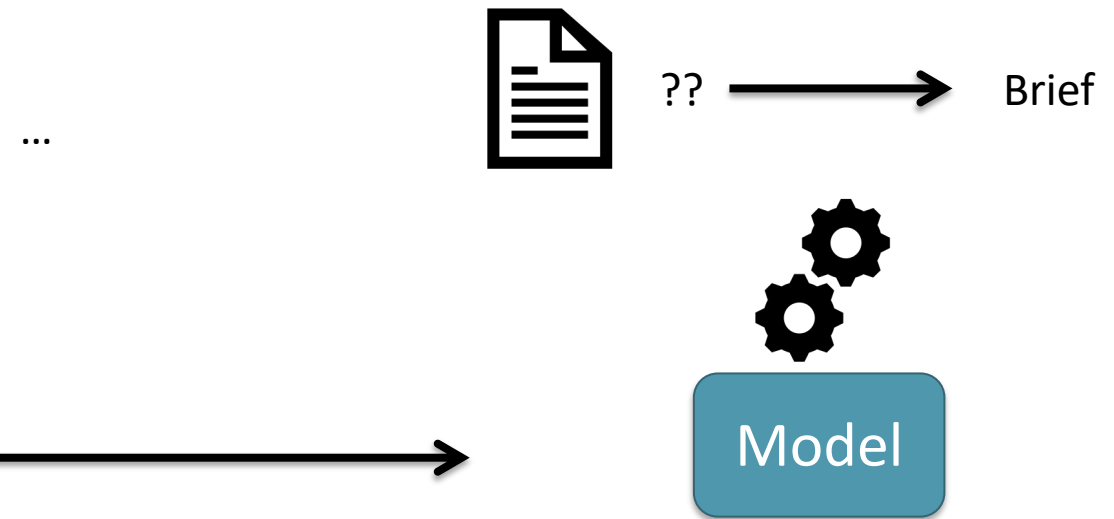


Image Sources: pixabay.com, Icon Source: Microsoft PowerPoint



Anwendungsbeispiel: Dokumente sortieren



Icon Source: Microsoft PowerPoint

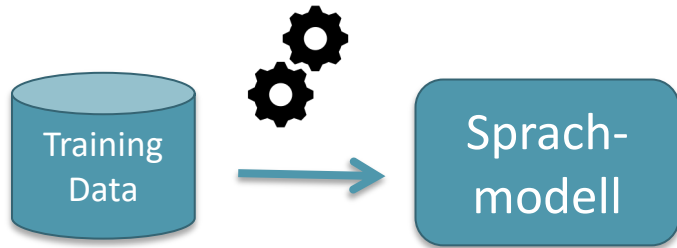
Anwendungsbeispiel: Textgenerierung

Anna ging in den Wald und sah einen ...

97% Baum

60% Ameisenbär

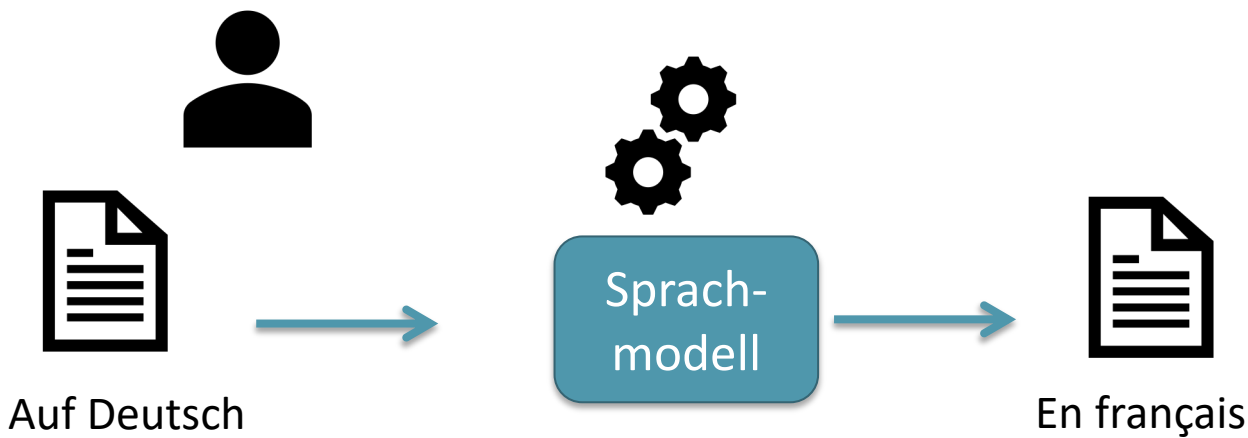
3% lächelte



Welches ist das wahrscheinlichste nächste Wort?

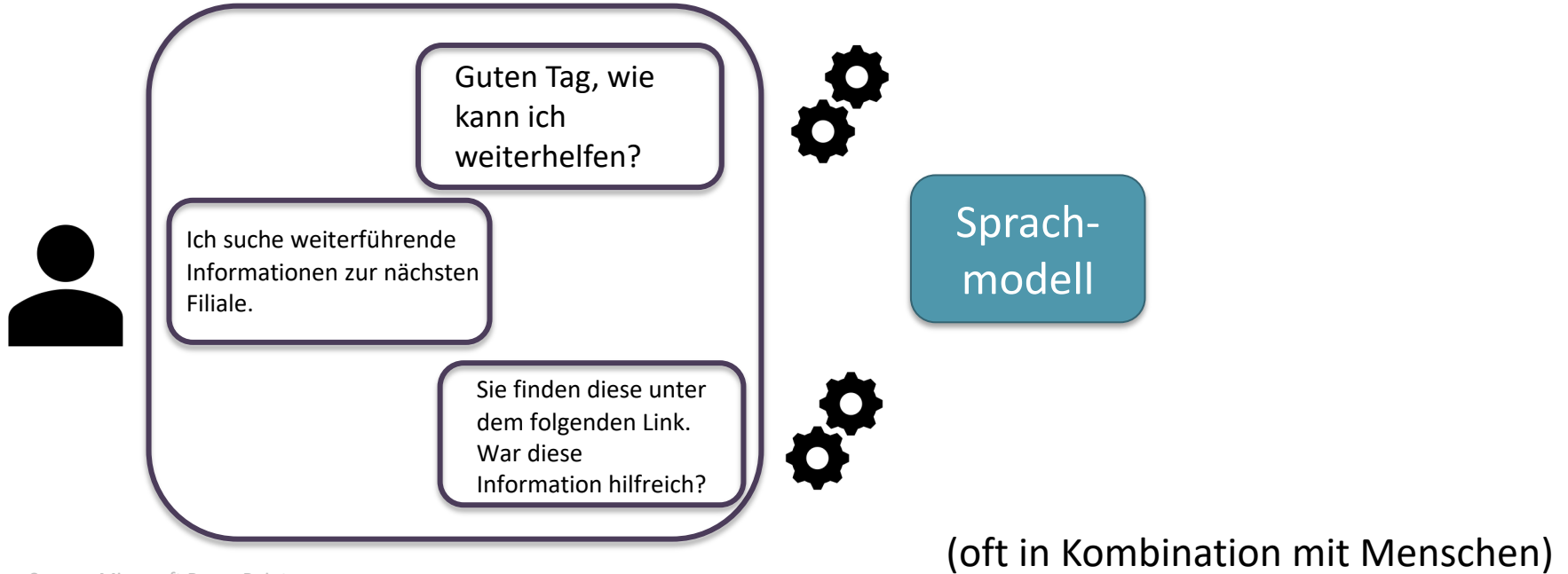
Icon Source: Microsoft PowerPoint

Anwendungsbeispiel: Automatische Übersetzung



Icon Source: Microsoft PowerPoint

Anwendungsbeispiel: Chatbots



Icon Source: Microsoft PowerPoint

Ein Blick hinter die Kulissen von KI zur Textanalyse und -generierung

Word Embeddings

Für automatische Verarbeitung:
Mathematischer Vektor, z.B. 300 Dimensionen



„Katze“

=

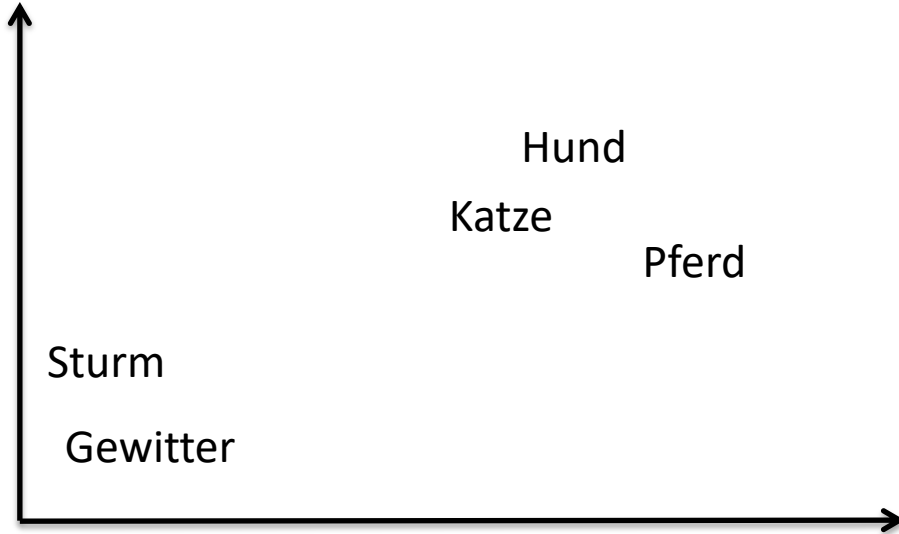
$$\begin{bmatrix} 11.2 \\ 3.4 \\ 4.5 \\ \dots \\ 6.7 \end{bmatrix}$$



Für Menschen: Wort in natürlicher
Sprache, z.B. Deutsch

Icon Source: Microsoft PowerPoint

Word Embeddings



Wörter mit ähnlicher
Bedeutung haben
Vektoren, die näher
beieinander sind

Eigenschaften von Word Embeddings

Diese Differenz zwischen den Vektoren kann genutzt werden:

„Man is to King, as Woman is to X“ X=Queen

weil

$$\vec{\text{Man}} - \vec{\text{Woman}} \approx \vec{\text{King}} - \vec{\text{Queen}}$$

→ Sehr nützlich für vielerlei Anwendungen!

Reference: Bolukbasi, Tolga, et al. "Man is to computer programmer as woman is to homemaker? debiasing word embeddings." *Advances in neural information processing systems*. 2016.

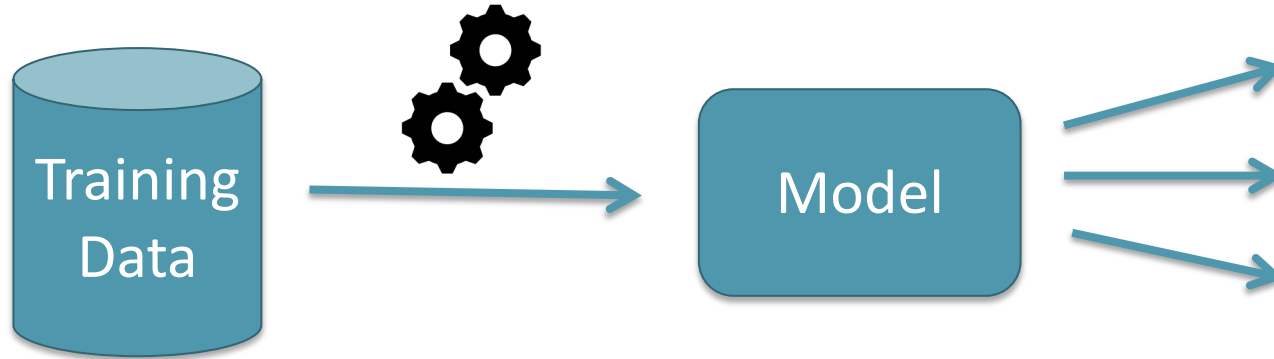
Solche Beziehungen sind nützlich für viele Anwendungen, aber können auch **Stereotypen** enthalten:

$$\begin{array}{c} \longrightarrow \\ \text{father} \end{array} - \begin{array}{c} \longrightarrow \\ \text{mother} \end{array} \approx \begin{array}{c} \longrightarrow \\ \text{doctor} \end{array} - \begin{array}{c} \longrightarrow \\ \text{nurse} \end{array}$$

„Father is to Doctor, as Mother is to Nurse“ ??

Reference: Bolukbasi, Tolga, et al. "Man is to computer programmer as woman is to homemaker? debiasing word embeddings." *Advances in neural information processing systems*. 2016.

Risiken und Limitationen von KI Software: Bias und Stereotypen



Die Wortvektoren und Sprachmodelle
enthalten Stereotypen!

Was bedeutet das für die
Vorhersagen / generierten Texte?

→ Kritisches Hinterfragen durch den/die Benutzer*in erforderlich

Icon Source: Microsoft PowerPoint

Beispiel: maschinelle Übersetzung

Englisch:

The smart student.

The beautiful student.

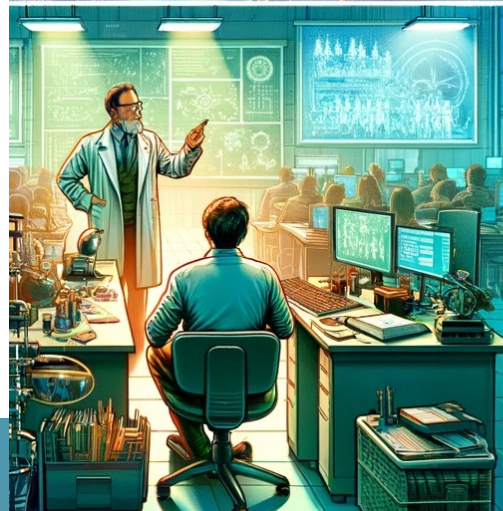
Maschinelle Übersetzung mit gängigen Tools:

Der kluge Student.

Die schöne Studentin.

ChatGPT 4

„Generate 4 pictures of professors in computer science.“



Created by Mascha&ChatGPT4

ChatGPT 4

„Generate 4 pictures of **different** professors in computer science.“



Created by Mascha&ChatGPT4

Risiken und Limitationen von KI Software: Erklärbarkeit

- Die Art und Weise wie KI funktioniert, macht es oft schwer, die Ergebnisse zu erklären
- Bei kommerziellen Anwendungen sind die verwendeten Trainingsdaten oft nicht bekannt
- Das macht es herausfordernd, Risiken zu erkennen

Risiken und Limitationen von KI Software: Fakten & Interpretation

- Generierte Texte können inhaltlich falsch sein
- Interpretation von Bedeutung und kommunikative Absicht in generierten Texten¹
- Sprachmodelle als „Stochastic Parrots“¹




1) Bender, E. M., Gebru, T., McMillan-Major, A., & Shmitchell, S. (2021, March). On the Dangers of Stochastic Parrots: Can Language Models Be Too Big? . In *Proceedings of the 2021 ACM conference on fairness, accountability, and transparency* (pp. 610-623).

Image Source: Pixabay/165106

Risiken und Limitationen von KI Software: Missbrauch

- Texte in bestimmten Stil erstellen: z.B. authentisch aussehende Spam E-Mails, Fake News in verschiedenen Sprachen, ...
- Deepfakes: sehr authentisch wirkende aber gefälschte Bilder und Videos
- Betrugsmaschinen mit Unterstützung von KI
- Cybercrime

Weitere Informationen dazu siehe auch: Bericht Europol 2023, ChatGPT The impact of Large Language Models on Law Enforcement

Risiken: Übersicht

Risiko: fehlende Transparenz und Erklärbarkeit

Risiko: Output ist eine Vorhersage

Risiko: Bias und Stereotypen

Risiko: Menschliche Interpretation

Risiko: Anwendungsfall

Risiko: Datenschutz



Produktwahl

Systemarchitektur

Geeignete Anwendungsfälle

Aufgabenteilung Menschen und KI

Prozesse

Awareness Rising

Risikoabschätzung für Einsatz

...

*Wie soll die digitale Gesellschaft
der Zukunft aussehen?*

«Generiere ein Bild zu dem Thema:
Wir möchten eine nutzbringende
Anwendung der Technik, **mit dem
Menschen im Zentrum.**»

Created by Mascha&ChatGPT4



Augmented Intelligence, statt Artificial Intelligence

- Menschen unterstützen, anstelle sie zu ersetzen
- Viel Potenzial für KI als Tool zur Unterstützung bei repetitiven Arbeiten, damit der Mensch mehr Zeit hat für andere Aufgaben



Image Source: pixabay/fancyrave1

Fazit

- Die KI-Software als Werkzeug sehen und verantwortungsvoll einsetzen
- Sich den Limitationen bewusst sein
 - Ist der Anwendungsfall geeignet?
 - Generierte Inhalte auf Korrektheit prüfen
 - Kritisch hinterfragen (Stereotypen?)

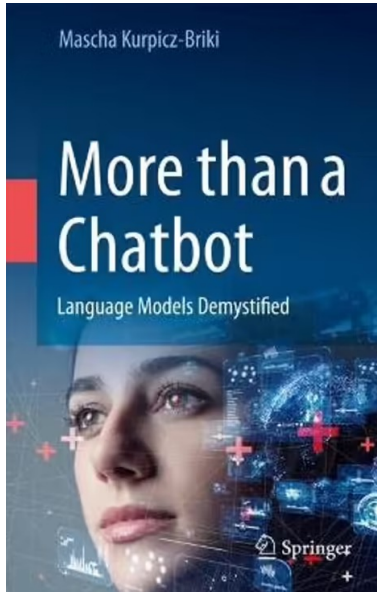


Image Source: pixabay/talhakhali007-5671515



Berner
Fachhochschule

Fragen & Diskussion



Neues Buch (2023):


Was steckt hinter Sprachmodellen und Chatbots, was sind die Limitationen, und wie können wir sie trotzdem verantwortungsvoll nutzen?

Deutsche Ausgabe erscheint 2024!

Prof. Dr. Mascha Kurpicz-Briki

Applied Machine Intelligence
Bern University of Applied Sciences

<http://www.bfh.ch/ami>
mascha.kurpicz@bfh.ch

 /maschakurpicz