



Bundesministerium für Bildung und Forschung



Wie entstehen Daten?

HPI Hasso Plattner Institut
Digital Engineering · Universität Potsdam



Bundesministerium für Bildung und Forschung

Supporttickets im Kundenservice

Suchanfragen in Google

Röntgenbilder

Daten entstehen i.d.R. über Zeit in Form von NutzerInnen in Form eines Prozesses.

Langlebigkeit vs. Kurzlebigkeit

Johannes Hötter, Christian Warmuth

openHPI

Wie entstehen Daten?







Bundesministerium

Langlebigkeit vs. Kurzlebigkeit

Johannes Hötter, Christian Warmuth

openHPI

2020 2040

Student	hat studiert
Christian	Data
Warmuth	Engineering
Johannes	Data
Hötter	Engineering

Student	hat studiert
Christian	Data
Warmuth	Engineering
Johannes	Data
Hötter	Engineering
Jon Doe	IT Systems Engineering

wächst mit der Zeit ...

Wie entstehen Modelle?











Langlebigkeit vs. Kurzlebigkeit

Johannes Hötter, Christian Warmuth

openHPI

Datengrundlage

Student	hat studiert
Christian	Data
Warmuth	Engineering
Johannes	Data
Hötter	Engineering

Modell



Wie entstehen Modelle?







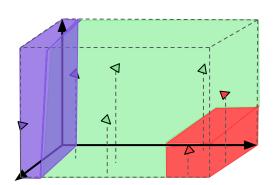


Langlebigkeit vs. Kurzlebigkeit

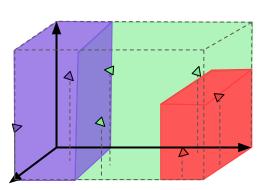
Johannes Hötter, Christian Warmuth

openHPI

Modell A



Modell B



Sind (i.d.R.) austauschbar! z.B. von proprietär auf open-source

Recap

Ist nicht jedes Machine Learning-Modell datenzentriert?



Trends:

- Komplexe Algorithmen sind in open-source Bibliotheken bereits umgesetzt
- Anbieter wie OpenAI oder Hugging Face (open-source) stellen vortrainierte
 Modelle bereit, welche per Finetuning einfach angepasst werden können





Konsequenzen:

- Für die meisten Anwendungen kein Mehrwert, eigene Algorithmik zu implementieren
- Fokus lieber auf der Datensammlung und -aufbereitung
- Vorteil: Daten sind langlebig und modellunabhängig

Langlebigkeit vs. Kurzlebigkeit

Johannes Hötter, Christian Warmuth

openHPI

Recap Verschiedene Wege der Umsetzung generativer KI







Prompt Engineering

Kontext mitliefern

In-Context Learning

Fine Tuning

—

Alle diese Optionen sind datenzentriert!

Langlebigkeit vs. Kurzlebigkeit

Johannes Hötter, Christian Warmuth

openHPI





Bundesministerium für Bildung und Forschung

