



Use-Case

Einfacher Coding-Assistent am eigenen Laptop

Einbettung

Die KI wird als einfach zugängliches Lernwerkzeug im Lehr-Setting zum Programmieren eingesetzt. Sie ist bewusst niederschwellig gestaltet und kann auf jedem Gerät lokal genutzt werden. Im Mittelpunkt stehen:

- erste Berührungspunkte mit KI im Lernprozess
- selbstständiges Erkunden und Ausprobieren

Lernziele

Nach Abschluss des Lernsettings können die Lernenden ...

- erklären, wie die KI zur Unterstützung beim Programmieren eingesetzt werden kann,
- die KI gezielt zur Klärung von Verständnisfragen zu Code und Konzepten anwenden,
- analysieren, welche Antworten der KI hilfreich sind und wo deren Grenzen liegen,
- unterschiedliche Lösungsansätze beim Programmieren (eigene vs. KI-unterstützte) vergleichen,
- begründen, wann der Einsatz lokaler KI beim Programmieren sinnvoll ist und wann nicht.

```
./Llama-3.2-1B-Instruct-06_K.llamafile
... werden, in der 5 Bestseller-Bücher mit ihrem Veröffentlichungsjahr gespeichert
... erden. Beim Programmstart sollen alle Bücher ausgegeben werden. freeing cont
> ./Llama-3.2-1B-Instruct-06_K.llamafile

LLAMAFILE

software: llamafile 0.9.2
model: Llama-3.2-1B-Instruct-06_K.gguf
compute: Apple Metal GPU
server: http://127.0.0.1:8080/

A chat between a curious human and an artificial intelligence assistant. The assi
stant gives helpful, detailed, and polite answers to the human's questions.
>>> Schreibe ein Python-Programm. Es soll eine Variable als Dictionary angelegt
werden, in der 5 Bestseller-Bücher mit ihrem Veröffentlichungsjahr gespeichert w
erden. Beim Programmstart sollen alle Bücher ausgegeben werden.
Hier ist ein Python-Programm, das die gewünschte Funktion implementiert:

'''python
class Bestseller:
    def __init__(self):
        self.bücher = {}
```

Lokale Kommandozeile

Warum Lokale KI?

Lokale KI unterstützt Lernende in einem geschützten und datensicheren Lernraum bei ersten Berührungspunkten mit Künstlicher Intelligenz. Sie ermöglicht einen einfachen, plattformunabhängigen Zugang und fördert einen selbstständigen sowie reflektierten Umgang mit KI als Lernhilfe. Dabei bleibt die Kontrolle über Daten und Nutzung vollständig bei den Lernenden und Lehrenden.

DSGVO

Durch den lokalen Betrieb werden keine personenbezogenen Daten an Dritte übertragen, wodurch die volle Kontrolle über Datenverarbeitung gewährleistet ist.

Compliance

Die lokale KI ermöglicht die Einhaltung institutioneller und rechtlicher Vorgaben durch eine transparente Systemarchitektur und nachvollziehbare Prozesse.

Datensicherheit

Sensible Daten werden ausschließlich lokal verarbeitet, was das Risiko von Datenlecks und externem Zugriff deutlich reduziert.

Urheberrecht

Proprietärer Code und Lernmaterialien verbleiben vollständig im lokalen System, wodurch eine klare Kontrolle über Trainings- und Nutzungsdaten erhalten bleibt.



Laptop
Endgerät

Umsetzung mit Lokaler KI

oer.one/ki-lokal/alf

Hardware Voraussetzung

- Technisches Know-How: mittel
- Betriebssystem: Windows, macOS oder Linux
- RAM: mindestens 8 GB
- Grafikkarte: keine Voraussetzung
- Speicherplatz: je nach Modell, ab 2 GB
- Internet: nach dem Download nicht notwendig

Anleitung

Schritt 1: Webseite öffnen

Rufe die Projektseite Llamafile auf (z. B. über [Mozilla.ai](#) oder [GitHub](#)) und wähle ein passendes Llamafile-Modell aus der Liste der Beispielmodelle aus.

Schritt 2: Llamafile herunterladen

Lade die gewünschte Datei herunter. Es handelt sich um eine einzelne Datei, die bereits das Modell und die Chat-Oberfläche enthält.

Schritt 3: Datei ausführbar machen

- Windows: Datei in *.exe umbenennen
- macOS / Linux: Datei ausführbar machen (chmod +x dateiname)

Schritt 4: KI nutzen

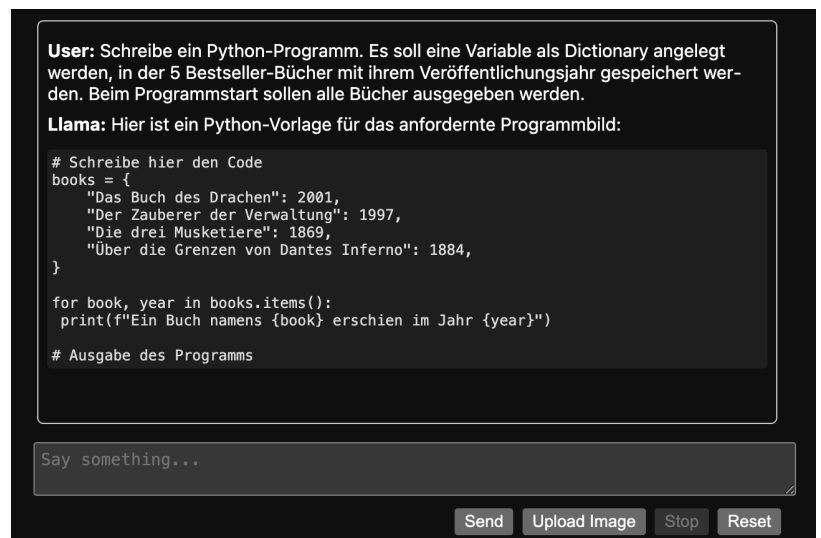
Llamafile in der Kommandozeile starten.

Die KI läuft nun vollständig lokal auf dem eigenen Gerät. Fragen können direkt in der Kommandozeile oder im Browser gestellt werden. Es ist keine Anmeldung oder Cloudzugang erforderlich.

Model	Size	License	Llamafile	other quants
LLaMA 3.2 1B Instruct	1.11 GB	LLaMA 3.2	Llama-3.2-1B-Instruct.Q6_K.llamafile	See HF repo
LLaMA 3.2 3B Instruct	2.62 GB	LLaMA 3.2	Llama-3.2-3B-Instruct.Q6_K.llamafile	See HF repo
LLaMA 3.1 8B Instruct	5.23 GB	LLaMA 3.1	Llama-3.1-8B-Instruct.Q4_K_M.llamafile	See HF repo
Gemma 3 1B Instruct	1.32 GB	Gemma 3	gemma-3-1b-it.Q6_K.llamafile	See HF repo
Gemma 3 4B Instruct	3.50 GB	Gemma 3	gemma-3-4b-it.Q6_K.llamafile	See HF repo
Gemma 3 12B Instruct	7.61 GB	Gemma 3	gemma-3-12b-it.Q4_K_M.llamafile	See HF repo

Liste an Llamafiles

Quelle: https://mozilla-ai.github.io/llamafile/example_llamafiles



Lokale Chat-Oberfläche im Browser



CC BY 4.0 International
Benedikt Brünner (TU Graz), Andreas Zitek (BOKU University), Martin Ebner (TU Graz)

