

CLOUDERA

Open Source Sprachmodell in Cloudera Data Platform Machine Learning

Nutzen und Einsatzmöglichkeiten in der Verwaltung
September 2023

Friedrich Rothkoetter
frothkoetter@cloudera.com
Jürgen Bienzeisler
bienzeisler@cloudera.com

Worüber wir sprechen werden

Neue Formen Generativer KI sicher und regulationskonform nutzen im Behördenumfeld

Sicherer Einsatz von ChatGPT im Behördenumfeld

- Welche Hilfsmittel stellt generative AI zur Unterstützung bei der Automatisierung von Verwaltungsvorgängen zur Verfügung?
- Welchen Nutzen können Behörden aus dieser Technologie erwarten?
- Was gibt es beim datenschutzkonformen Einsatz dieser Technologie zu bedenken?
- Wie können ChatGBT oder Large Language Modules (LLMs) bereits heute datenschutzkonform und in der Praxis bei Behörden eingesetzt werden?
- Praxisbeispiele für die Herangehensweise an ein Projekt mit Cloudera

Was ist ein vor-trainiertes Sprachmodell (LLM) ?

Large Language Model

Ein vor-trainiertes **Sprachmodell (LLM)** ist eine **spezielle Form künstlicher Intelligenz (KI)**, die auf Basis großer Textmengen trainiert wurde, um **vorhandene Inhalte zu verstehen** und **neue Inhalte zu generieren**.

Das Sprachmodell besitzt die Fähigkeit, auf eine **Vielzahl von Themen und Fragen zu antworten** und sich **wie ein menschlicher Gesprächspartner** zu verhalten und **komplexe Konversationen führen**.

Dank seines maschinellen Lernens und seiner **neuronalen Netze** ist es in der Lage, **auf natürliche Weise zu kommunizieren** und **komplexe Konversationen zu führen**.

Welche Fähigkeiten besitzt das vor-trainierte Sprachmodell (LLM) ?

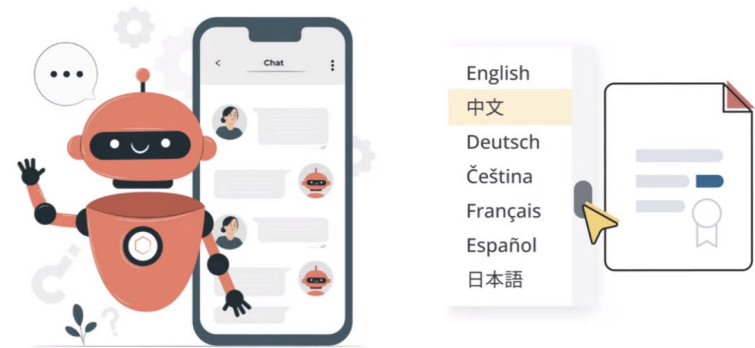
eine nicht vollständige Aufzählung

- **mehr Sprachen** als ein Mensch
 - Übersetzungen erstellen
 - Englisch, Französisch, Spanisch, Italienisch, Portugiesisch, Deutsch, Niederländisch, Russisch, Koreanisch und Japanisch sowie „weitere Sprachen“
- **Textinhalte** zusammenfassen
 - Briefingdokumente oder Protokolle
- **Mails oder Textentwürfe** schreiben
- Inhalte, Dokumentationen oder **Vorschriften finden**
- **Software** schreiben
- **Bilder** oder Musik **generieren**

Sprachmodelle (LLMs) haben das größte Automatisierungspotential

Fachkräftemangel und zu erwartende zunehmende Arbeitsbelastung in der Verwaltung

- Sprachmodelle (LLMs) verwenden **natürliche Sprache als Interaktions-Medium**
 - **Keine Programmier- oder Abfragesprachen** erforderlich
 - **Keine spezifischen Trainings** oder Einführungen erforderlich
- Sprachmodelle (LLMs) sind **universell einsetzbar**
 - **Keine Limitierung** hinsichtlich der Anwendungsfälle



Welchen Nutzen können Behörden von dieser Technologie erwarten?

Fachkräftemangel und zu erwartende Arbeitsbelastung in der Verwaltung

Jeder Behördenmitarbeiter - und falls gewünscht auch jeder Bürger - können mit dem **Instrument der natürlichen Sprache** auf Daten und Informationen der Verwaltung Zugriff erhalten **ohne spezifische Kenntnisse oder Abfragesprachen** zu beherrschen oder **Trainings** zu absolvieren.

Damit können **Verwaltungsvorgänge** **erfolgreich, durchgängig und rasch digitalisiert** werden.

Bestehende Fachanwendungen der **Verwaltung** können **mit Sprachmodellen als Schnittstelle** ausgestattet werden.

Mögliche Anwendungen des Sprachmodells im Behörden Alltag

Bürgerservice & Onlinezugangsgesetz (OZG)

- Assistenzsystem für Bürgerservice
 - bestehende digitale Angebote der Verwaltung finden und nutzen
- Onlinezugangsgesetz (OZG)
 - Initiativen zügig realisieren
- Angebote im Bürgerservice zügig erweitern
 - Automatisierungspotentiale des Sprachmodells erproben

Mögliche Anwendungen des Sprachmodells im Behörden Alltag

Assistenzsysteme für die Verwaltungsmitarbeiter

- **Assistenzsystem für Verwaltungsmitarbeiter** (interne Verwaltungs-Prozesse)
 - **Unterstützung der Sachbearbeitung**
 - **Feststellen des Bearbeitungsstandes** eines Vorgangs
 - **Anwendbarkeit von Richtlinien** und Vorgaben **prüfen**
 - **Informationen zu Anfragen** suchen und **finden**
 - **Erstellen von Entwürfen** für Protokolle, Entscheidungsvorlagen oder Briefings
 - **Elektronische Kommunikation** im Entwurf erstellen lassen
 - **Übersetzung von Dokumenten** im Entwurf
- **Vorgaben für Barrierefreiheit** in der digitalen Verwaltung umsetzen
- Daten **Monitoring- und Reporting zur Vorhersage** und Frühwarnung
- **Programmier Code** automatisiert erstellen - Effizienz der IT steigern !

Fähigkeiten des Sprachmodells – Allgemein

Beispiel

Erläutere den
Zulassungsvorgang
eines KfZ in Essen
In **einfacher Sprache**

Welche **Regelungen**
Existieren zum
Erwerb der deutschen
Staatsbürgerschaft
In **russischer Sprache**

Fähigkeit

Erinnert sich an den
Zusammenhang des
zuvor Gesagten in der
Konversation

Übersetzungen aus der
**Originalsprache in
andere Sprachen**



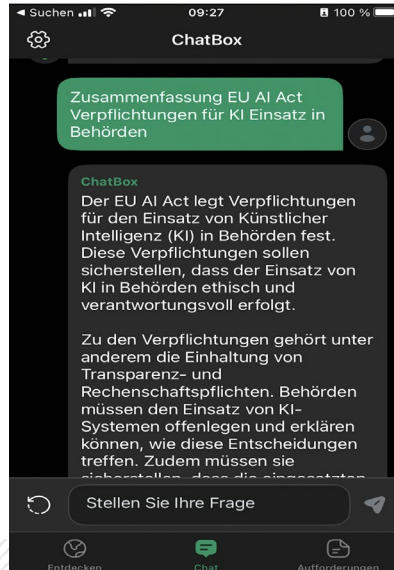
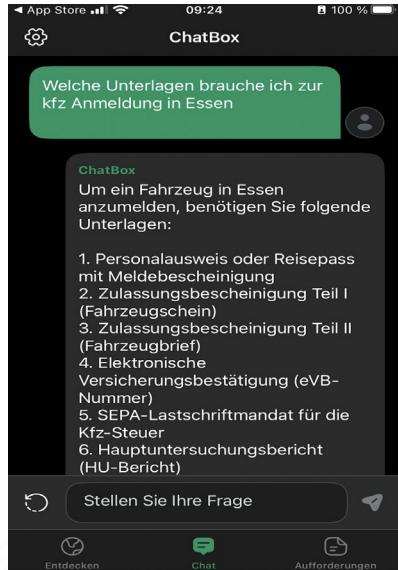
Hallo **Herr Meyer**,
Wie kann ich Ihnen bei
der Zulassung Ihres
Fahrzeugs helfen?



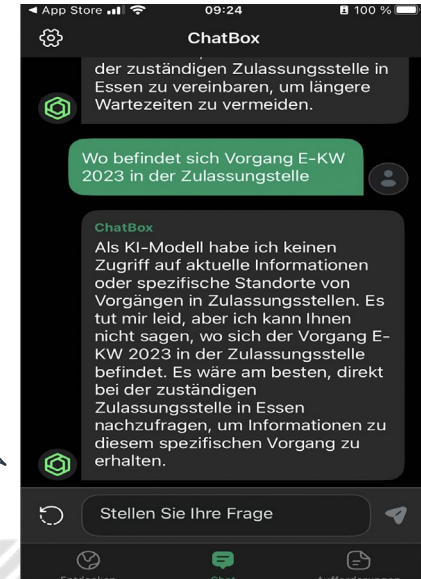
Beschränkungen des Sprachmodells - fachlich und inhaltlich

am Beispiel von ChatGBT

- Sprachmodelle erschließen die Inhalte des WWW
 - denn darauf wurden sie trainiert
- Sprachmodelle haben **keine Kenntnis von behörden-internen Daten oder Prozessen**



Als KI-Modell habe ich **keinen Zugriff auf aktuelle Informationen** oder **spezifische Standorte von Zulassungsstellen.**



SaaS Sprachmodelle

sind in der öffentlichen Verwaltung nicht nutzbar

Gründe:

- Cybersecurity Risiko
- Data Privacy Risiko
- Ethische Risiken
- Verwaltungs Richtlinien
- Falsche Informationen auch wenn sie plausibel erscheinen

Darum legt das “Office of the National Counterintelligence Executive” U.S. fest, dass keine sensitive information in ChatGBT prompts verarbeitet werden darf.

Auch Samsung hat nach einer Exposition interner Daten in ChatGBT reguliert, dass keine persönlichen Informationen oder Firmendaten in ChatGBT prompts verarbeitet werden dürfen.

Auch die Erstellung von Program Code wurde stark reguliert.

Cloudera löst diese Problematik mit der Cloudera Data Platform

Es entstehen keine Kosten für die Nutzung des LLM und der Betrieb ist vorkonfiguriert in der Plattform

- Das Sprachmodell wird in der **geschützten Umgebung Ihres Service Providers** oder Rechenzentrums betrieben
- Es besteht **kein Risiko, Behördendaten zu kompromittieren** oder nach außen zu geben
- Das Sprachmodell ist ebenso **leistungsfähig** wie Chat GPT-3
- Governance- und **regulative-Anforderungen sind automatisiert**
- Das Sprachmodell ist in der Lage, den **Fach-Kontext** der Behörde zu verarbeiten, um **sachlich richtige Antworten** zu generieren

Cloudera Data Platform enthält ein vor-trainiertes Sprachmodell

Vorkonfiguriert und integriert in die Plattform ist es **betriebsbereit für den Projekteinsatz**

Der **Code des Sprachmodells ist zu 100 % Open Source**, sodass jeder Cloudera Kunde ihn unmittelbar nutzen kann und alle Cloudera-Kunden ihn **mit wenigen Klicks in ihrem CML-Arbeitsbereich bereitstellen** können.

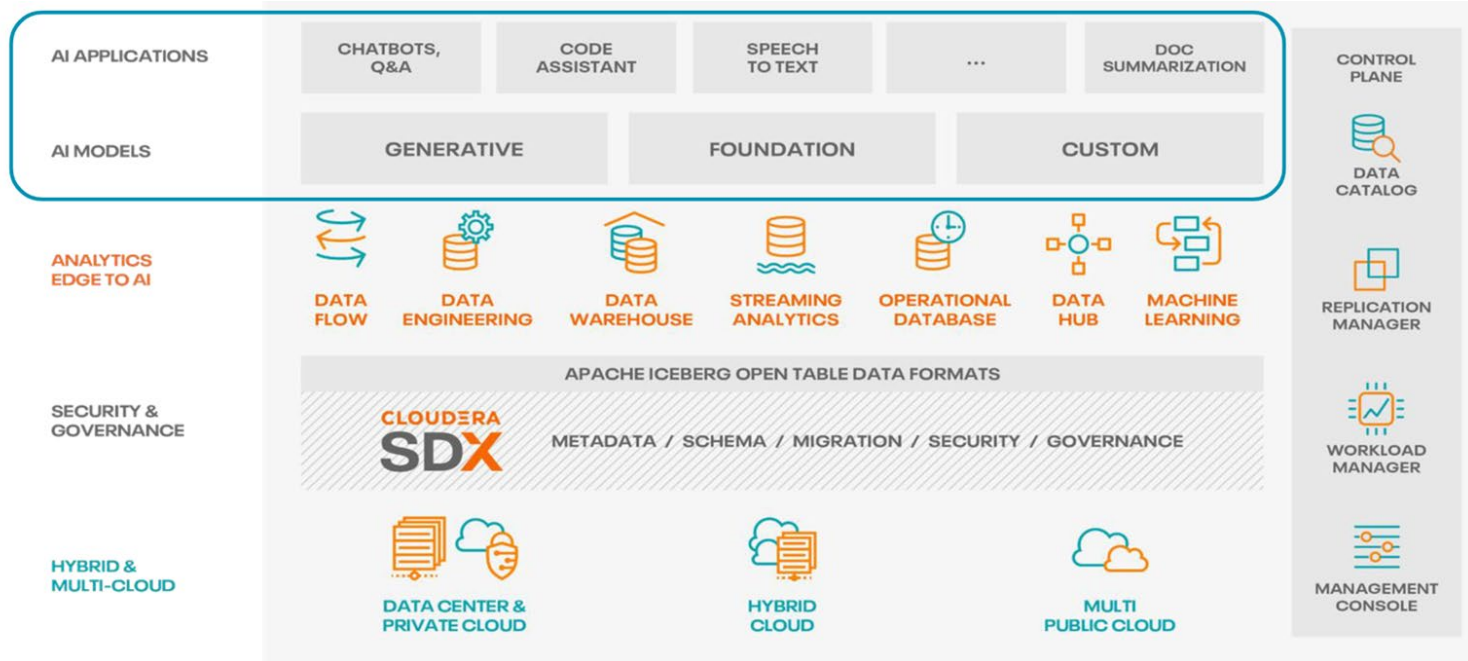
Das vor-trainiertes Sprachmodell stammt aus dem Ecosystem at Hugging Face:

<https://huggingface.co/models>

<https://huggingface.co/blog/os-llms>

HuggingFace veröffentlicht fortlaufend neue Sprachmodelle (LLMs), die von verschiedenen Forschungsteams und Organisationen entwickelt wurden und auf modernsten Techniken und Architekturen basieren und das kollektive Fachwissen der breiteren KI-Community nutzen.

Das Sprachmodell ist Teil einer integrierten Plattform für AI Entwicklung



schnellere und effektivere Datenverwaltung und Datenanalyse für Daten, mit optimaler Leistung, Skalierbarkeit und durchgängige Sicherheit



Das Open Source Sprachmodell in der Cloudera Data Platform



LLM Chatbot Augmented with Enterprise Data

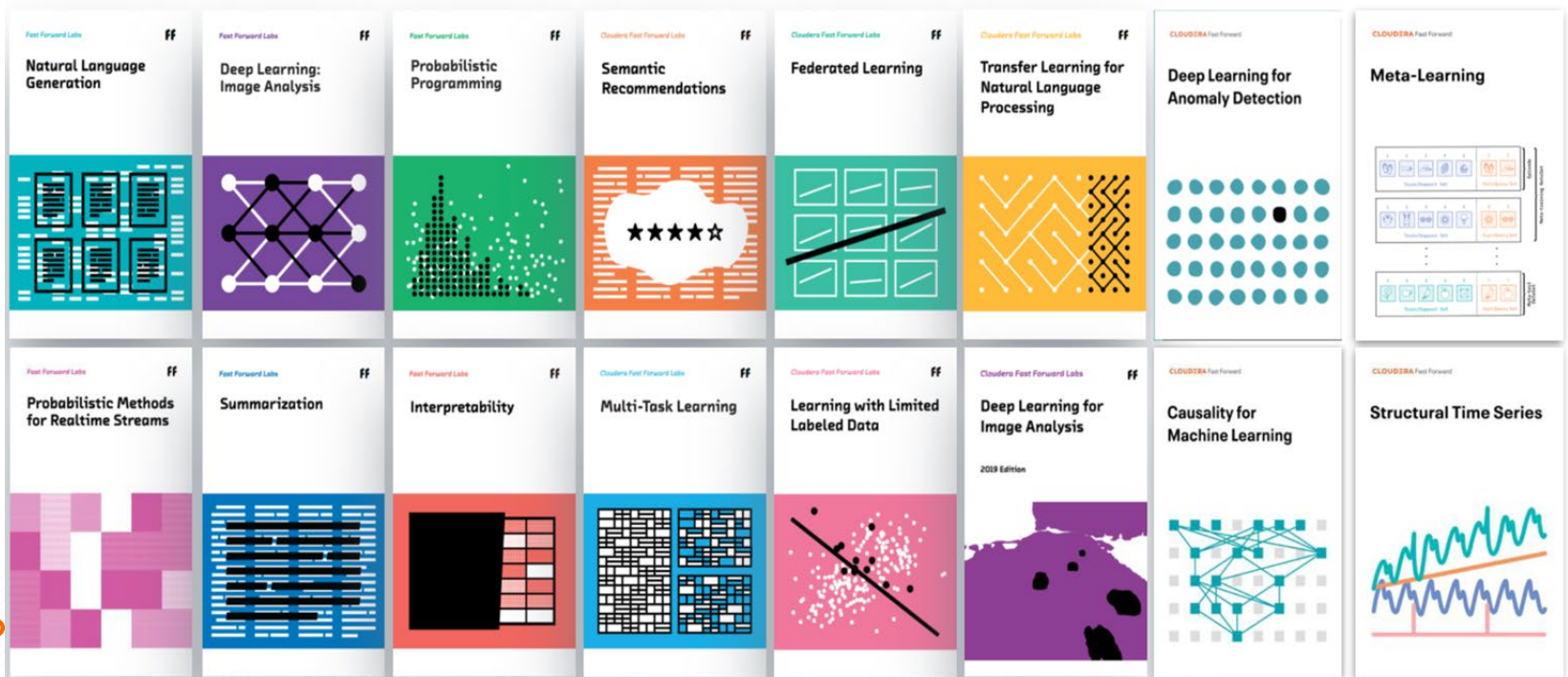
CHATBOT

LLM

- Open Source Sprachmodell (LLM)
- Ohne externe Aufrufe an OpenAI oder einen anderen externer SaaS-KI-Dienst
- Generieren sachlich richtiger Antworten
- Schnelle Umsetzung in kurzer Zeit
- Moderne und Entwickler-freundliche Machine Learning Umgebung
- Anpassbare und flexible Data Integration

Cloudera Data Platform: Vielzahl von Modellen verfügbar

Starten Sie Ihre KI Projekte in der Cloudera Data Plattform mit **Modellen**, die von **Cloudera** entwickelt und gepflegt werden. Dies führt zu **Qualität und Geschwindigkeit** in der Entwicklung von KI Projekten



Keine sachlich richtigen Antworten ohne Fachinformation der Behörde

Aus dem Vortrag von Stefan wissen wir, dass **Sprachmodelle** mit der Information aus dem **Internet vor-trainiert** sind und daher **keine Informationen aus deutschen Behörden und Ihren Fachverfahren** kennen können.

Ohne diese fachspezifischen Informationen ist das Sprachmodell jedoch unfähig, **Behörden-spezifische Fragestellungen zu beantworten oder zu automatisieren.**

Fachinformationen aus Behörden Verfahren hinzufügen

Die Wissensdatenbank

Das **Sprachmodell** wird mit **behörden-spezifischen Kontext zu Fachverfahren** versorgt, um die **Antwortqualität des Sprachmodells** durch die Entwicklung eines **aufgabenspezifischen Verständnisses** signifikant zu verbessern.

Erst dadurch ist das Sprachmodell in der Lage, **spezifische Fachanfragen zu verstehen, bessere Antworten zu geben und die Nuancen der kundenspezifischen Sprache** geschickt zu handhaben.

Auf diese Weise können auch Antworten in der gewünschten Sprachebene erreicht werden (akademische, einfache Sprache oder Verwaltungs-Jargon).

Einige Begriffsdefinitionen zur Erläuterung der Technologie

Wissensdatenbank

“**Wissensdatenbank**” oder “knowledge base” ist der **Speicherort für die Dokumente, Prozessbeschreibungen oder fachspezifischen Informationen**, die mit dem Sprachmodell verarbeitet werden sollen, um sachlich korrekte Antworten zu liefern

Datenquelle dafür könnte z.B. Ihr **Sharepoint Server** sein, auf dem Sie behördenspezifische Informationen, Datenverzeichnisse, Verfahrensbeschreibungen, Regelwerke oder Prozessabläufe hinterlegt haben

Einige Begriffsdefinitionen zur Erläuterung der Technologie

Vektor-DB

Damit das Sprachmodell **aus der Fülle der Informationen** und Daten in der **“Wissendatenbank”** zur Beantwortung einer Fragestellung **schnell die relevanteste Information findet**, wird eine **Technik** verwendet, die sich **Vektor-DB** nennt.

Im Grunde handelt es sich dabei um einen ähnlichen Vorgang, wie wir in relationalen Datenstrukturen mit Indices schnell Datensätze finden. Allerdings ist die Technologie hier eine völlig andere und sie arbeitet mit Vektoren und sogenannten **“embeddings”**.

Wie wird das Sprachmodell konfiguriert, betrieben und verwendet?

Erläuterung unserer Technologie und Projekt Tätigkeiten

1. Initiale Befüllung der Wissensdatenbank mit Ihren Dokumenten und Daten
 - **Datenquelle** dafür könnte z.B. Ihr **Sharepoint Server** sein, auf dem Sie behördenspezifische Informationen, Datenverzeichnisse, Verfahrensbeschreibungen, Regelwerke oder Prozessabläufe hinterlegt haben
2. Automatisches Befüllen der Vektor-DB, um relevante Informationen, die zur Anfrage an das Sprachmodell passen, schnell und zuverlässig zu finden
 - Dies ist ein automatischer Prozess, der von der Cloudera Datenplattform durchgeführt wird
3. Befüllung der Wissensdatenbank mit Updates und Ergänzungen

Automatische Befüllung der Vector DB mit embeddings (=Vektoren)

Als Quellen für die “Wissensdatenbank” kommen in Frage:

Behördenspezifische Informationen und Datenverzeichnisse, Verfahrensbeschreibungen, Regelwerke oder Prozessabläufe.

Datenquelle könnte z.B. Dokumente und Daten in Sharepoint sein.

Diese Dokumente werden automatisiert in eine Vector Datenbank importiert.

Zweck ist es, schnell und zuverlässig die relevante Information in der “Wissensdatenbank” zu identifizieren und diese als Kontext der Anfrage für die sachlich richtige Beantwortung zu verwenden.

```
Populate Vector DB with documents embeddings Success  
By Alex Song — Session — 1 vCPU / 4 GiB Memory — 13 days ago See job details  
  
Session Logs  
  
-----  
/  \| /  _/ \| | / / / / / / / / _/  
/ / \| / / / / \| | / / / / / \ \  
/_/ / _/ / / / / / / / / / / / / / {Lite}  
  
Welcome to use Milvus!  
  
Version: v2.2.8-lite  
Process: 153  
Started: 2023-08-18 16:51:37  
Config: /home/cds/milvus-data/configs/milvus.yaml  
Logs: /home/cds/milvus-data/logs  
  
Ctrl+C to exit ...  
v2.2.8-lite  
Milvus database is up and collection is created  
Generating embeddings for: ml-product-overview-2.txt  
Generating embeddings for: ml-product-overview-4.txt  
Generating embeddings for: ml-product-overview.txt  
Generating embeddings for: ml-product-overview-3.txt  
Generating embeddings for: ml-product-overview-5.txt  
Generating embeddings for: ml-architecture-overview-spark-on-kubernetes.txt  
Generating embeddings for: ml-architecture-overview1.txt  
Generating embeddings for: ml-architecture-overview-cml-1.txt  
Generating embeddings for: ml-architecture-overview-cml-2.txt  
Generating embeddings for: ml-architecture-overview-runtimes.txt  
Generating embeddings for: ml-architecture-overview-provisioning.txt  
Generating embeddings for: iceberg-snippet.txt  
Generating embeddings for: ozone-snippet.txt  
Total number of inserted embeddings is 13.  
Finished loading Knowledge Base embeddings into Milvus
```

Zum Einsatz des vor-trainierten Sprachmodells in 6 Schritten

1. Auswahl des Prototypen LLM aus dem Katalog in Cloudera Machine Learning Bibliothek
2. Konfigurieren des Projektes in der Plattform
3. Aus der Menüoberfläche Download starten, um das Sprachmodell und seine Abhängigkeiten zu laden
4. "Wissensdatenbank" befüllen
5. Vektor Datenbank befüllen durch einen automatisierten Prozess in der Plattform
6. Starten der Sprachmodell Anwendung

**Wir zeigen Ihnen jetzt,
wie dies in der Praxis geht ...
Achtung: tech stuff 10 min**

Aus Cloudera Machine Learning Bibliothek das Modell auswählen

LLM Chatbot Augmented with Enterprise Data auswählen

The screenshot shows the Cloudera Machine Learning interface. On the left is a dark sidebar with navigation options: Projects, Sessions, Experiments, Model Deployments, Model Registry, Jobs, Applications, User Settings, AMPs, Runtime Catalog (highlighted with a red circle), Site Administration, and Learning Hub. The main area is titled 'Applied ML Prototypes' and features a search bar, a 'Source' dropdown, and a 'Tags' dropdown. A red circle highlights the 'LLM Chatbot Augmented with Enterprise Data' prototype, which includes a 'CHATBOT' and 'LLM' tag. Other visible prototypes include 'Churn Modeling with scikit-learn', 'Deep Learning for Image Analysis', 'Deep Learning for Anomaly Detection', 'Structural Time Series', 'Analyzing News Headlines with SpaCy', 'Deep Learning for Question Answering', and 'Explaining Models with LIM SHAP'. The interface also displays various data visualizations and model performance metrics.

Modell Beschreibung lesen und Projekt konfigurieren

Verwenden Sie “Configure Project”

The screenshot shows the Cloudera Machine Learning interface. On the left is a navigation sidebar with options like Projects, Sessions, Experiments, Model Deployments, Model Registry, Jobs, Applications, User Settings, AMPs, Runtime Catalog, Site Administration, and Learning Hub. The main area displays 'Applied ML Prototypes' with a search bar and filters. Two project cards are visible: 'LLM Chatbot Augmented with Enterprise Data' and 'Structural Time Series'. A modal dialog is open for the first project, containing the following text:

LLM Chatbot Augmented with Enterprise Data

IMPORTANT: Please read the following before proceeding. By configuring and launching this AMP, you will cause h2oai/h2ogpt-oig-oasst1-512-6.9b, which is a third party large language model (LLM), to be downloaded and installed into your environment from the third party's website. Please see <https://huggingface.co/h2oai/h2ogpt-oig-oasst1-512-6.9b> for more information about the LLM, including the applicable license terms. If you do not wish to download and install h2oai/h2ogpt-oig-oasst1-512-6.9b, click "cancel" below. By clicking "Configure Project" below, you acknowledge the foregoing statement and agree that Cloudera is not responsible or liable in any way for h2oai/h2ogpt-oig-oasst1-512-6.9b. Author: Cloudera Inc.

This AMP builds a Retrieval Augmented Generation (RAG) Question-Answer Large Language Model (LLM) Bot application which demonstrates how context from local documents can be used with pre-trained LLM models to perform context retrieval and chat response generation with factual responses.

Configuration tags: CHATBOT, LLM, HUGGINGFACE, GENERATIVE AI, RAG, VECTOR DB, MILVUS, TRANSFORMERS, NLP.

Buttons: [View on Github](#), [Cancel](#), [Configure Project](#)

LLM Chatbot Augmented with Enterprise Data Projekt konfigurieren

Werte für die Variablen der Konfiguration auswählen - Damit definieren Sie die Ressourcen

The screenshot displays the 'Configure Project' page in Cloudera Machine Learning. The page title is 'Configure Project: LLM Chatbot Augmented with Enterprise Data - charu'. The AMP Name is 'LLM Chatbot Augmented with Enterprise Data (v1)'. The description states: 'This AMP demonstrates how to use an open source pre-trained instruction-following LLM (Large Language Model) to build a ChatBot-like web application. The responses of the LLM are enhanced by giving it context from an internal knowledge base. This context is retrieved by using an open source Vector Database to do semantic search.'

Environment Variables
This prototype does not define any environment variables.

Runtime
Editor

The 'Runtime' section contains four dropdown menus, which are circled in red in the image: 'Workbench', 'Python 3.9', 'Nvidia GPU', and '2022.11'. Below these is a toggle for 'Enable Spark' (disabled) and a dropdown menu showing 'Spark 3.1.1 - CDP 7.2.11 - CDE 1.13 - H0...'. The 'Runtime Image' field contains the value 'docker.repository.cloudera.com/cloudera/cdsw/ml-runtime-workbench-python3.9-cuda:2022.11.2-b2'. A blue box below the runtime image states 'No Runtime Addon is required for this AMP.'

Setup Steps
 Execute AMP setup steps

At the bottom right, there are 'Cancel' and 'Launch Project' buttons.

Download des Sprachmodells und Jobstart zum Befüllen der Vector-DB

3 Schritte mit vorgefertigten jobs in der Plattform - nacheinander auszuführen

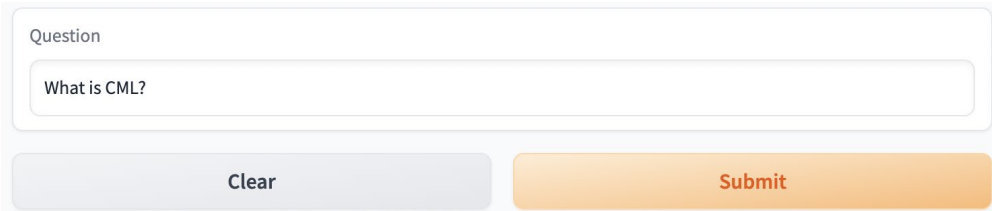
The screenshot displays the Cloudera Machine Learning (CML) interface. On the left is a dark sidebar with navigation options: All Projects, Overview, Sessions, Data, Experiments, Model Deployments, Jobs, Applications, Files, and Collaborators. The main content area shows the 'AMP Status' for a project named 'charu / LLM Chatbot Augmented with Enterprise Data - charu'. The page title is 'AMP Setup Steps'. Below the title, it states 'AMP Name: LLM Chatbot Augmented with Enterprise Data (v1)' and provides a brief description of the AMP's purpose. A progress indicator shows 'Completed 0 of 6 steps'. The setup steps are listed as follows:

- Step 1:** Install Dependencies. Status: started 5/16/2023 2:49 PM. Includes a 'View details' link.
- Step 2:** Create job to download pre-trained models. Status: not yet started. Description: Create job to download open source pre-trained models required by the LLM Chatbot application. All models are downloaded to a local directory.
- Step 3:** Run job to download pre-trained models. Status: not yet started.
- Step 4:** Create job to populate Vector Database with document embeddings. Status: not yet started. Description: Create job to launch Milvus Vector Database locally and insert embeddings for documents. Embeddings are generated by the locally running embeddings model.

Red ovals are drawn around the descriptions of Step 2, Step 3, and Step 4, highlighting the jobs to be executed.

Starten der Anwendung mit Frontend für Benutzer

Demonstration der Wirkungsweise der Wissensdatenbank



Question

What is CML?

Clear Submit

Antwort ohne Cloudera Fachinformation in der Wissensdatenbank

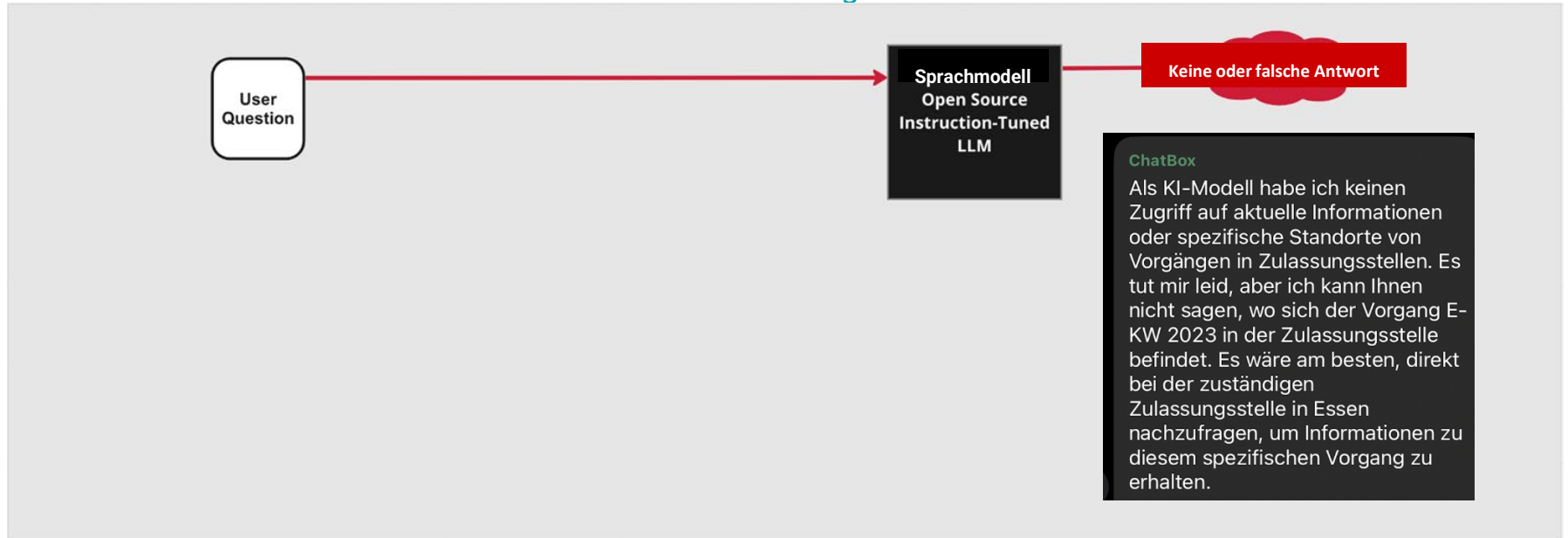
CML stands for Chronic Myeloid Leukemia, a type of blood cancer that affects white blood cells called neutrophils. It's caused by a mutation in the gene that controls how cells divide and differentiate. The mutation causes the cells to become cancerous and accumulate in the bone marrow, where they can cause symptoms such as fatigue, weight loss, and anemia.

Antwort mit Cloudera Fachinformation in der

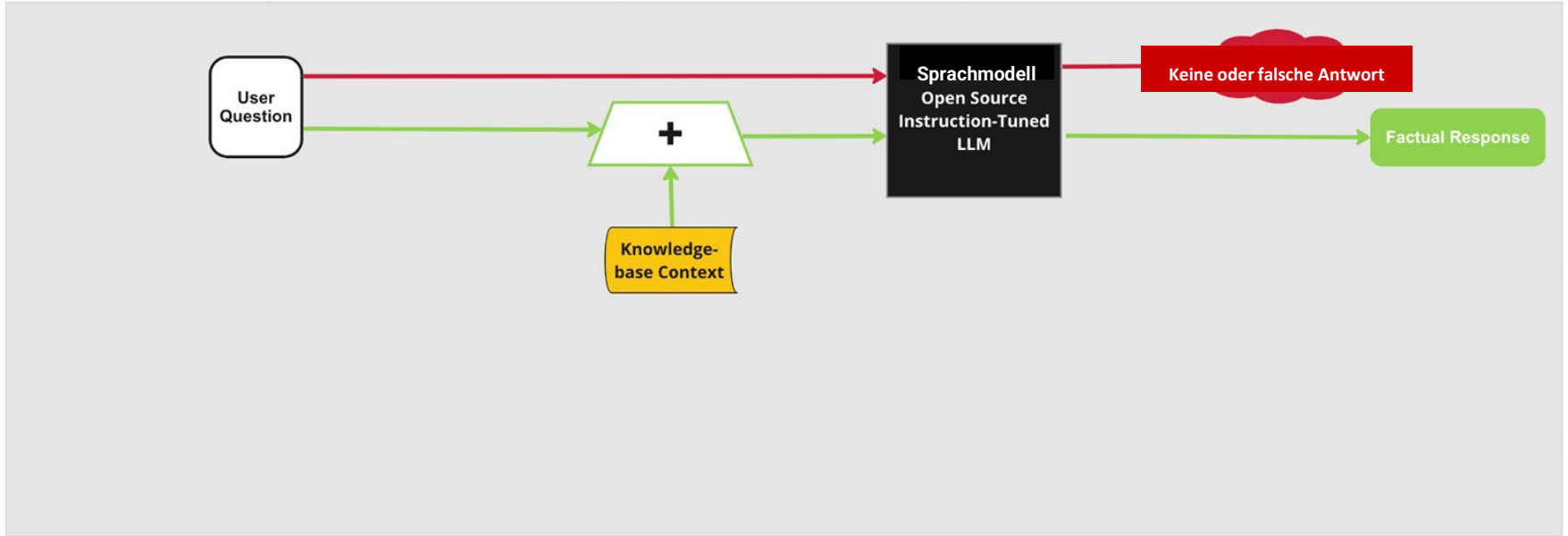
Cloudera Machine Learning (CML) is a three-tier architecture that provides a convenient way to manage and scale your analytical pipeline. It includes a web interface, a cluster of nodes, and a set of Docker containers that contain the code and data needed to run your analysis.

Die sachlich richtige Antwort und der Inhalt der Wissensdatenbank

Das Sprachmodell benötigt relevante Information in der Wissensdatenbank, um sachlich richtige Antworten zu geben

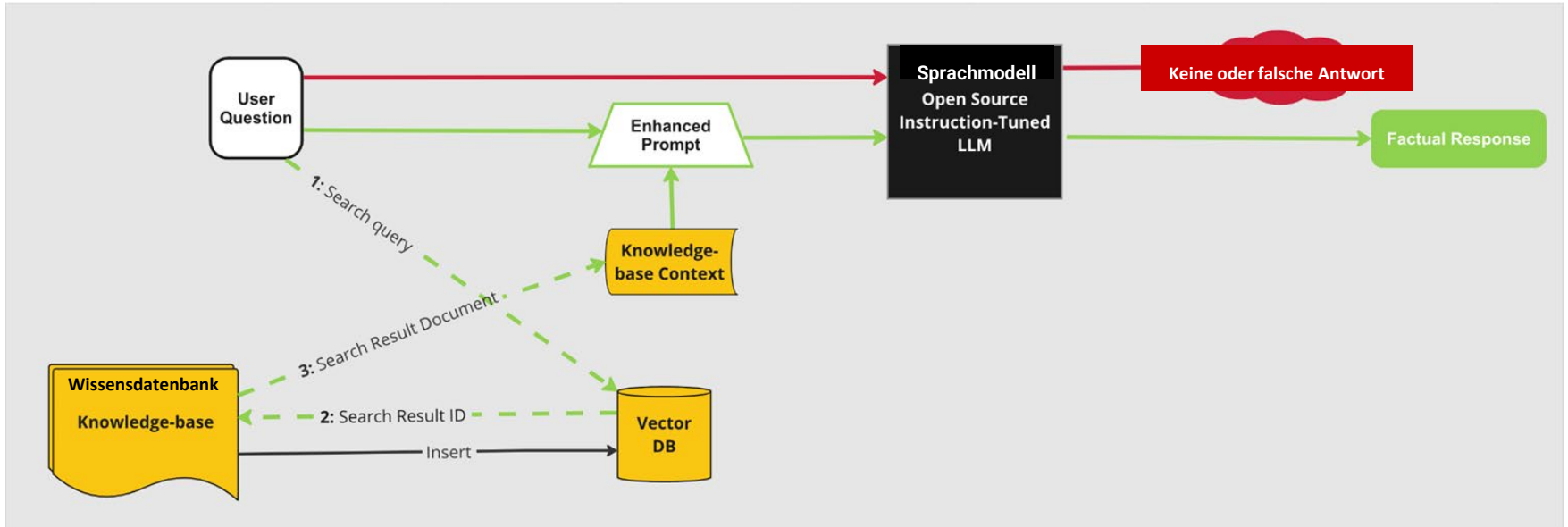


Der Anfrage wird der fachliche Kontext - Fachinformation mitgegeben aus der Wissensdatenbank stammend



Anfrage Kontext in Vector DB gefunden, relevanter Inhalt aus Wissensdatenbank wird verarbeitet

Fachlich korrekte Antwort für User Anfrage wird generiert



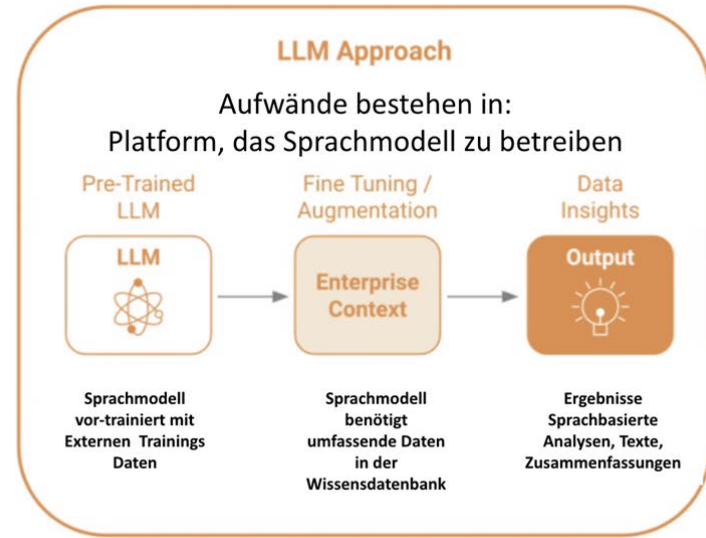
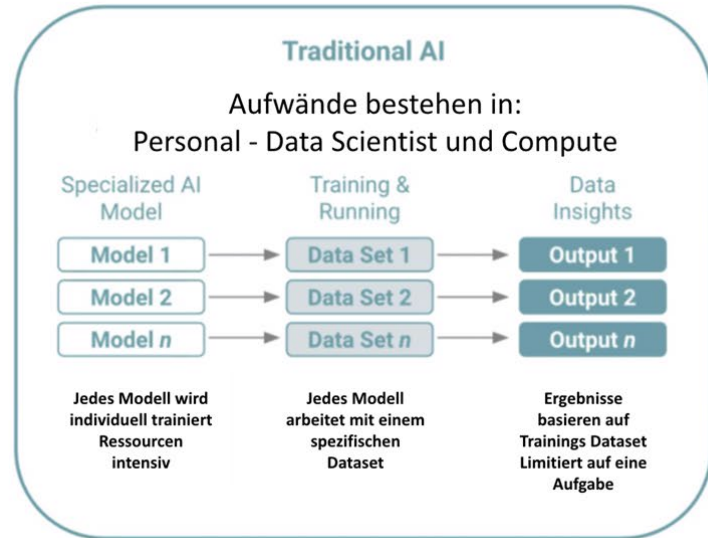
Anforderungen an die Datenarchitektur bei Nutzung von Sprachmodellen

Wissensdatenbank wird größer werden bei Nutzung des Sprachmodells

- Die **Menge der Daten in Ihrer “Wissensdatenbank”** wird **steigen**, um das vor-trainierte Sprachmodell **umfassend und aktuell mit Behörden-Kontext** zu versorgen
- Soll das Sprachmodell für **tagesaktuelle Antworten** sorgen, erfordert dies **tagesaktuelle Daten** in der “Wissensdatenbank”

Datenarchitektur bei Nutzung von Sprachmodellen

Vergleich Ressourcen, Datenmengen für ML Modelle und Sprachmodell



Wo könnten Sie dieses Sprachmodell betreiben?

Dienstleister der öffentlichen Verwaltung haben **Cloudera Data Platform** im Service Portfolio

- ITZBund
- Dataport
 - Data[port]ai
- Ihr eigener Dienstleister? Ihr eigenes RZ ?
- Public Cloud Provider (AWS, Azure)?

CLOUDERA DATA PLATFORM ÜBERSICHT

hybrid & shared data platform for advanced analytics, AI und ML



Data Fabric



Data Lakehouse



Data Mesh



AWS, Azure,
Google Cloud



On-Premises



Data Engineering



Data Flow



Data Warehouse



Operational Database



Machine Learning



Open data fabrics,
Data lakehouses
with data
anywhere at scale

Multi-cloud &
on-premises
data management
and analytics

“Write once,
run anywhere”
cloud-native data
analytics portability

Unified security &
governance with
open cloud-native
storage formats

Konformitätserklärung

Ich, der Referent Jürgen Bienzeisler, bestätige bei der Erstellung meiner Präsentation die rechtlichen/urheberrechtlichen Vorgaben für die Nutzung von Werken Dritter beachtet zu haben. Ich versichere, dass für den Fall der Nutzung fremder Werke in meiner Präsentation die erforderlichen Lizenzen / Genehmigungen eingeholt wurden.

Ich bin Urheber der Präsentation und zu erreichen unter: Jürgen Bienzeisler, Cloudera GmbH, Birketweg 31, 80639 Munich,

jbienzeisler@cloudera.com

+49 170 963 1887

Kontakt

Danke

