



knopflogik
GmbH

Live-Demo

Ein KI-Assistent für Kitodo.Production

Dr.-Ing. Thomas Low

13.11.2025

Wer bin ich?

Dr.-Ing. Thomas Low

- Doktorand an der Otto-von-Guericke Universität Magdeburg bis 2017
- Mitgründer des Startups LegalHorizon AG
- Doktorabschluss zum Thema ~Explorative Suche~ im Januar 2022
- Gründung der knopflogik GmbH aus Eigenkapital im Mai 2021

Kontakt

Telefon: +49-391-40594560

Mail: thomas.low@knopflogik.de

Web: <https://knopflogik.de/>

Worum geht es?

Ein Experiment aus Eigeninitiative mit

- Open-Source Software,
- frei verfügbaren KI-Technologien, und
- bezahlbarer Hardware

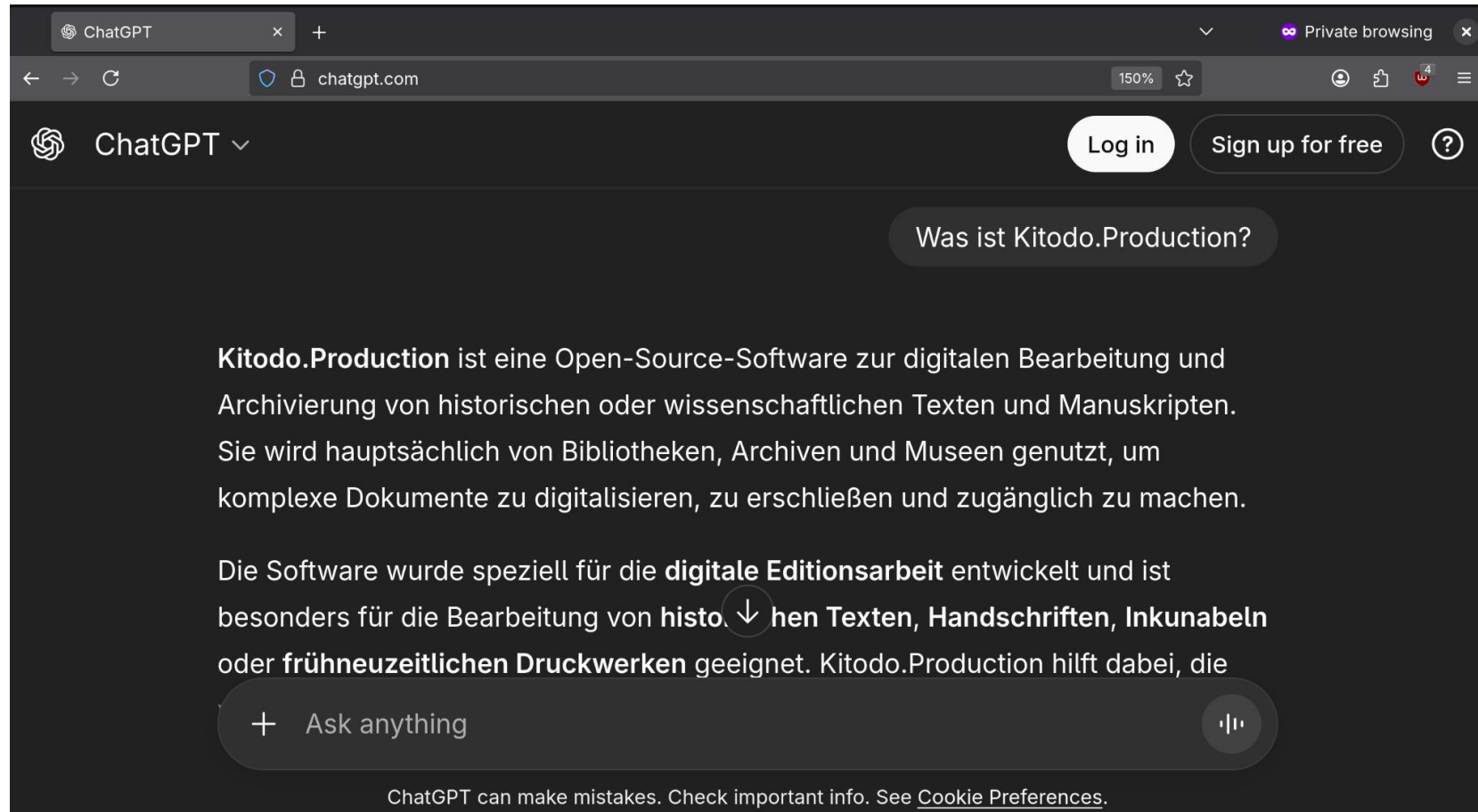
im Szenario von

- Kitodo.Production

zur Beantwortung der Fragen

- Was ist aktuell mit KI-Assistenten umsetzbar?
- Welche Probleme gibt es?

KI und Sprachmodelle



Von Sprachmodellen zu KI-Assistenten

Ein Sprachmodell, zusätzlich dafür trainiert, um Befehlsaufrufe zu erzeugen:

1. Anfrage: *Wie ist das aktuelle Wetter in Magdeburg?*
2. Sprachmodell: *{ command: "weather", location: "Magdeburg" }*
3. Wetter-API: *{ temperature: "10°C", condition: "cloudy" }*
4. Sprachmodell: *Magdeburg ist aktuell bewölkt bei 10°C.*

Wozu ein KI-Assistent?

Im Szenario von Kitodo.Production

- Unterstützung beim Onboarding neuer Nutzer
- Bessere Barrierefreiheit durch Sprachinteraktion
- Automatisierung von Workflows ohne Programmierkenntnisse

Von Sprachmodellen zu KI-Assistenten

Kommerzielle Sprachmodelle

OpenAI ChatGPT
Google Gemini
Microsoft Copilot
Anthropic Claude
Perplexity AI

...

Frei verfügbare Sprachmodelle

Gemma (von Google)
Llama (von Meta)
GPT-oss (von OpenAI)
DeepSeek
Qwen (von Alibaba)

...

Von Sprachmodellen zu KI-Assistenten

Es gibt Sprachmodelle in verschiedenen Größen:

Anzahl an Parametern (in Milliarden)

0.6b → 1.7b → 8b → 30b → 235b

Hardwareanforderungen

Mobil → Office-PC → PC mit Grafikkarte → KI-Server → Rechenzentrum

Einschränkungen zur Live-Demo

Nur einzelne Funktionen sind eingebunden:

- zur Vorgangsliste und Aufgabenliste navigieren
- Aufgaben nach Aufgabenschritt filtern
- Eine Aufgabe bearbeiten
- Die Gesamtanzahl an Aufgaben ausgeben

Der Assistent (bisher)

- hat keinen Zugriff auf die Bedienungsanleitung, GitHub Tickets, o.Ä.
- "sieht" die Benutzeroberfläche nicht,
- kann die Maus nicht bedienen,
- kann keinen Text eingeben.

Live-Demo (als Video)

The screenshot displays the KITODO PRODUCTION web application. At the top, there is a search bar with the text 'Suche' and a magnifying glass icon. Below the search bar, the page title 'Vorgänge' is visible, along with a 'Treffer: 81282' indicator and a 'Filter' input field. The main content area is a table with columns: 'ID', 'Vorgangstitel', 'Status', 'Projekt', 'Laufzeit', and 'Aktionen'. The table lists several tasks, each with a status bar (green, yellow, or red) and a duration. A chat overlay is present on the left side of the table, showing a conversation with a user named 'Halle'. The chat messages include: 'Hallo! Wie kann ich Ihnen heute helfen?', 'Wie kann ich mich anmelden?', and 'Um sich anzumelden, geben Sie bitte Ihre Benutzername und Passwort in die entsprechenden Felder ein und klicken Sie auf „Anmelden“. Falls Sie Ihr Passwort vergessen haben oder Zugangsprobleme haben, wenden Sie sich bitte an die Institution, die dieses Kitodo.Production-Exemplar hosts.' Below the chat, there is a button 'zur Vorgangsliste navigieren' and a message: 'Sie wurden erfolgreich zur Seite „Alle Vorgänge“ navigiert. Hier können Sie alle vorhandenen Vorgänge im System anzeigen lassen. Möchten Sie weitere Informationen oder Aktionen in Bezug auf bestimmte Vorgänge durchführen?'. At the bottom of the chat, there is a text input field with the placeholder 'eine Frage oder ein Befehl' and a 'Senden' button. In the bottom right corner, there is a small video feed showing a person's face. The text 'KITODO.PRODUCTION Version N/A' is visible at the bottom right of the interface.

ID	Vorgangstitel	Status	Projekt	Laufzeit	Aktionen
81299	Towaensu_1920706593	Green	Example Project	221d; 1h	[Icons]
81287	DeepLefo3_1911775154	Red	Example Project	329d; 4h	[Icons]
81285	Process of Another Project	Yellow	Another Project	383d; 18h	[Icons]
81281	DeepLeinW_1905076460	Red	Extended Structure	388d; 23h	[Icons]
81280	InteFrImA_189089026X	Red	Example Project	514d; 23h	[Icons]
81270	Unbekannt	Yellow	Example Project	1269d; 5h	[Icons]
		Red	Example Project	1269d; 5h	[Icons]
		Red	Example Project	1269d; 5h	[Icons]
		Red	Example Project	1269d; 5h	[Icons]
		Red	Example Project	1269d; 5h	[Icons]
		Red	Example Project	1269d; 6h	[Icons]
		Red	Example Project	1269d; 6h	[Icons]
		Red	Example Project	1269d; 6h	[Icons]
		Red	Example Project	1270d; 4h	[Icons]
		Red	Example Project	1270d; 4h	[Icons]
		Yellow	Extended Structure	1275d; 6h	[Icons]
		Red	Example Project	1295d; 3h	[Icons]

<https://knopflogik.de/video/eafdd31d>

Wie ist der KI-Assistent umgesetzt?

Eigener Prompt

You are a helpful assistant for the tool Kitodo.Production.

Kitodo.Production is a workflow management software for [...]

In case a user asks about login credentials, recommend that he needs to contact the institution that is hosting this instance of Kitodo.Production for more information. [...]

Always answer in German.

Wie ist der KI-Assistent umgesetzt?

Befehle

1. Das Sprachmodell erhält eine Beschreibung und das Datenformat

Navigate to the edit own task page for a specific row in the list of all tasks.
{ command: "edit_own_task_page", row: 1 }

2. Befehle werden im Browser per JavaScript ausgeführt
3. Die Ausgabe eines Befehls instruiert das Sprachmodell über (Mis-)Erfolg

"You need to navigate to the all tasks page before editing a task."

Was geht, was geht nicht?

Funktioniert gut

- Einfache Sprachbefehle
- Vorlesen von angezeigten Informationen
- Beantworten von einfachen Fragen

Schwierig ist

- Ableiten von Aktionen auf Basis eines übergeordneten Ziels
- Ohne Begleitung lange Befehlsketten ausführen
- Ungewöhnliche Begriffe per Sprache eingeben

Eingesetzte Technologien

Open-Source-Software

- Sprachmodell: [qwen3:8b](#) (Apache 2.0) via [ollama](#) (MIT)
- Text-to-Speech: [sherpa-onnx](#) (Apache 2.0)
- Speech-to-Text: [whisper](#) (MIT), Variante "small"

Eingesetzte Hardware

- Office-Workstation mit Grafikkarte für Sprachmodell / Speech-to-Text
(8-Core CPU, 64 GB RAM, RTX 3080 TI VRAM 12 GB)
- Server ohne Grafikkarte für Text-to-Speech
(12-Core CPU, 32 GB RAM)

Ein KI-Assistent für Kitodo.Production

Wie geht es weiter?

- Forschungsprojekt
- Masterarbeit
- Abwarten auf bessere, allgemeine Browser-Assistenten
 - z.B. [BrowserOS](#), ChatGPT Atlas, Perplexity Comet

Diskussion

Vielen Dank für Ihre Aufmerksamkeit!

Kontakt

Telefon: +49-391-40594560

Mail: thomas.low@knopflogik.de

Web: <https://knopflogik.de/>