

## HANDLUNGSANLEITUNG

## Chatbot Arena

Mit diesem Tool können Fragen, Aufträge oder sonstige Anliegen mit zwei Chatbots gleichzeitig bearbeitet werden. Die Ergebnisse werden immer parallel ausgespielt. Sie können bewerten, welches Ergebnis Ihnen besser gefällt und dafür abstimmen. Nach der Abstimmung enthüllt die Chatbot Arena, welche zwei KI-Systeme gegeneinander angetreten sind. Es gibt darüber hinaus auch eine Wettkampftabelle, in der Sie erkennen können, welches KI-System mit seinem Chatboteinsatz bereits die meisten Punkte in den Duellen gesammelt hat.

Ziel: Chatbots und ihre Arbeitsweise/-ergebnisse miteinander vergleichen

Benötigtes Material: Internetfähiges Endgerät

Dauer: nach Belieben

### Wie nutze ich die Chatbot Arena?

Die Chatbot Arena ist einfach zu bedienen – es gibt auf der Seite auch eine kurze Einführung in englischer Sprache, Sie können mit dem Tool aber in jeder beliebigen Sprache interagieren.

Scrollen Sie auf der Seite nach unten, um zum mittig positionierten Chat-Fenster zu gelangen. Dort können Sie Ihre Frage oder einen Auftrag (Prompt) eingeben. Oberhalb des Chat-Fensters werden nach einem kurzen Moment Bearbeitungszeit parallel zwei verschiedenen Antworten oder Ergebnisse ausgespielt, die auf zwei unterschiedlichen KI-Modellen beruhen, deren Identität Sie zu diesem Zeitpunkt aber nicht erfahren. Das linke Feld wird mit „Model A“ bezeichnet und das rechte Feld mit „Model B“ (rote Pfeile in Abb.1). Dies Antworten oder Ergebnisse können durch Nachfragen oder Verbesserungsvorschläge weiter verändert/ausgeweitet/angepasst werden.

Sind Sie mit Ihren Fragen oder Aufforderungen fertig, vergleichen Sie die Antworten und können diese unterhalb der Antwortfelder bewerten. Sie können wählen zwischen: „A ist besser“, „B ist besser“, „Gleichstand“ oder „beide sind schlecht“ (blauer Pfeil in Abb.1). Nach Ihrer Stimmabgabe wird Ihnen das KI-Modell angezeigt, das jeweils hinter den Antwortfeldern steckt.



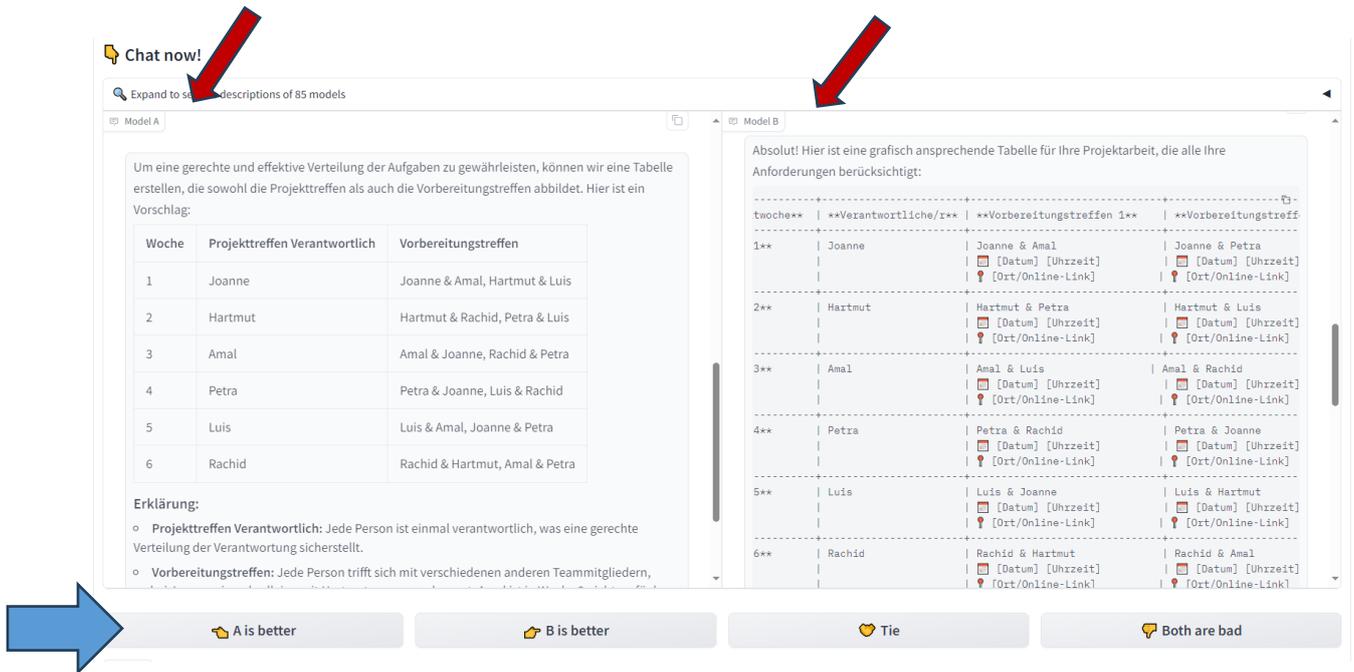


Abbildung 1: Ergebnisfenster mit Abstimmungsbuttons

Oberhalb der Ergebnisfenster können Sie zudem das „Chatbot Arena Leaderboard“ einsehen, also die Wettkampftabelle, welche KI-Systeme die meisten Punkte gemacht haben. Dort finden Sie neben der Rangfolge der Modelle auch einige grafische, statistische Auswertungen und die Informationen, das mit dem neusten Update vom 5. Januar 2025 insgesamt 189 KI-Modelle in den Wettkampf geschickt und bislang über 2,5 Millionen Bewertungsstimmen abgegeben wurden.

### Was kann das Tool noch?

Neben dem Battle-Modus, wo Sie zwei KI-Modelle blind vergleichen, können Sie unter der Registrierkarte „Arena (side-by-side)“ zwei konkrete KI-Modelle auswählen und diese miteinander vergleichen (blauer Pfeil in Abb. 2). Beispielsweise könnten Sie ausprobieren, ob Ihnen die Ergebnisse von ChatGPT (OpenAI) oder von Gemini (Google) besser gefallen, zwei der aktuell (Januar 2025) meist genutzten KI-Modelle für Chatbots.

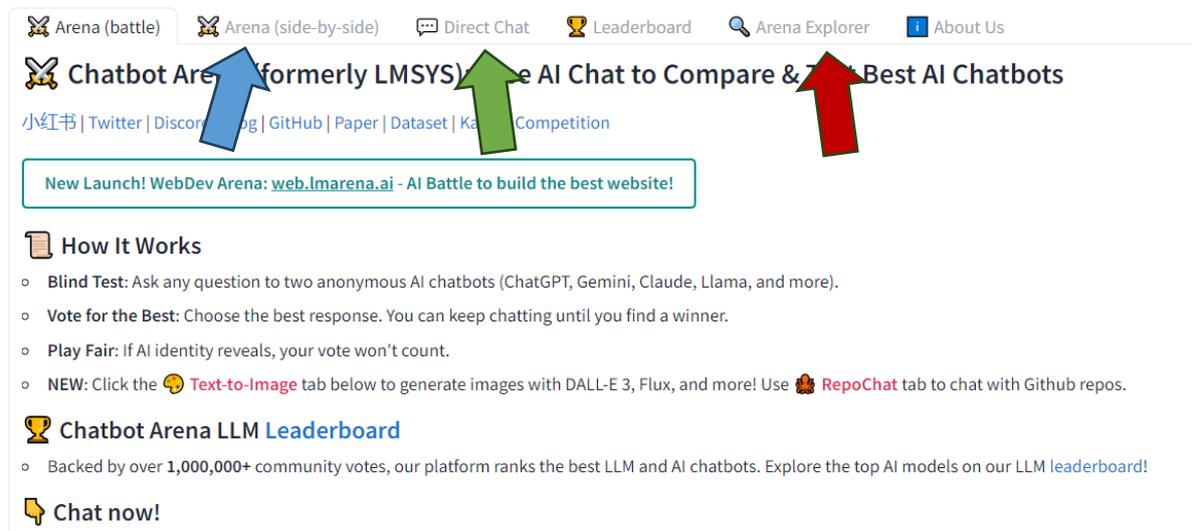


Abbildung 2: Verschiedene Registrierkarten

In der Registrierkarte „Direct Chat“ (grüner Pfeil in Abb.2) können Sie ohne die vergleichende Wettkampfsituation ein konkretes KI-Modell auswählen und mit diesem interagieren. Sie benötigen dafür keinen Zugang und können jedes Modell kostenlos nutzen.

Unter der Registrierkarte „Arena Explorer“ (roter Pfeil in Abb.2) können Sie interaktiv erleben, wie die Community die Chatbot Arena nutzt. Die individuellen Prompts werden hier in allgemeinere Kategorien eingeteilt. Beispielsweise können Sie sehen, dass 13,63% der Nutzenden Prompts eingeben, die dem kreativen Schreiben zugeordnet werden können und 4,45% interessieren sich für geschichtliche Themen. In verschiedenen Unterkategorien werden Beispieldprompts angezeigt, die als Inspiration für eigene Prompts genutzt werden können.

## Warum lohnt der Einsatz in der Erwachsenenbildung?

Die Gegenüberstellung von KI-generierten Chatbot-Ergebnissen ist äußerst interessant. Mittlerweile ist das Wissen um KI-generiertes Wissen zwar so hoch, dass Menschen aus dem Bildungsbereich klar sein sollte, dass dieses Wissen immer zu überprüfen ist, aber die Gefahr halluzinierender KI wird ganz anders deutlich, wenn auf ein und dieselbe Frage unterschiedliche Antworten generiert werden und optisch nebeneinanderstehen. Falschinformationen werden deutlich sichtbar und auch Fehleranfälligkeiten bei Generierungsaufforderungen.



Im Test wurde beispielsweise die Frage gestellt, was es mit dem Feiertag Fronleichnam auf sich hat und woher der Name stammt. Als Erklärung zur Herkunft des Namens wurde durch ein KI-Model eine Herleitung aus dem lateinischen „fructus“ erdacht (Falschinformation!), das im Folgenden mit „Frucht der Leiche“ übersetzt wurde. Diese Aussage in einem ansonsten sehr logisch klingenden und ansprechend formulierten Text fällt nicht so sehr auf, wie der Unterschied zu einer völlig anderen Antwort, die der ersten optisch gegenübergestellt wird.

In einem zweiten Test wurde darum gebeten einen fiktiven Projektplan zu erstellen. Menschen sollten Aufgaben zugeordnet und Zeiten zur Bearbeitung festgelegt werden. Obwohl beide Ausgaben grafisch ansprechende Ergebnisse lieferten, wurde im Vergleich sofort deutlich, dass die eine Variante diverse Vorgaben ignorierte und das Ergebnis auch nach mehrmaliger Umformulierung der ursprünglichen Fragestellung nicht nutzbar wurde.

Die Chatbot Arena eignet sich für Erwachsenenbildner\*innen, sich selbst im Umgang mit KI zu schulen und sensibel für die Art und Weise zu werden, wie Chatbots Antworten und Ergebnissen generieren. Gleichermaßen trainiert sie das Formulieren eigener Prompts, da Sie durch die Ausgabe von jeweils zwei Ergebnissen direkt bemerken, was in Ihrer Eingabe missverständlich oder unvollständig formuliert ist.

Ebenso können Sie mit einer entsprechenden Sicherheitsbelehrung (siehe nachfolgenden Absatz) die Chatbot Arena auch im Kurs einsetzen. Denn auch Ihre Teilnehmenden werden so am eigenen Leib erfahren, wie unterschiedlich Ergebnisse ausgespielt werden, wie schnell die KI bei einem professionellen Eindruck falsche Informationen liefert und welche Fähigkeiten im Umgang mit KI trainiert werden sollten, um diese positiv und produktiv nutzen zu können.

### Was gilt es zu beachten?

Beim Öffnen der Chatbot Arena werden Sie je nach Browser darauf hingewiesen, dass das Seitenangebot eine Art Forschungsvorschau darstellt und nur eingeschränkte Sicherheitsvorkehrungen bereithält. Die Plattform untersagt die Nutzung durch illegale oder schädliche Inhalte oder solche, die Gewalt, Rassismus oder Sex beinhalten.

Die Seitenentwickler\*innen weisen darauf hin, dass keine privaten Daten im Zuge der Chatbot-Nutzung eingegeben werden sollten. Denn die Seite sammelt die



Chatdaten und behält sich vor, diese auch unter einer CC-BY-Lizenz oder einer ähnlichen Lizenz wieder zu veröffentlichen (beispielsweise im Rahmen der Auswertungen und Darstellungen im Arena Explorer).

Dieser Information müssen Sie zustimmen, wenn Sie die Seite nutzen möchten.

Die Chatbot Arena ist eine offene Plattform und steht als Open Source Projekt auch über GitHub zur Verfügung. Interessierte Entwickler\*innen, die das Projekt weiterbringen möchten, sind herzlich eingeladen, sich entsprechend zu engagieren.

[CC BY-SA 3.0](#) DE by **Christina Bliss** für wb-web (16.01.2025)



Dieses Material steht unter der Creative-Commons-Lizenz Namensnennung – Weitergabe unter gleichen Bedingungen 3.0. Um eine Kopie dieser Lizenz zu sehen, besuchen Sie <http://creativecommons.org/licenses/by-sa/3.0/de/>