

Leitfaden zum Umgang mit interaktiven Sprachmodellen wie ChatGPT & Co bei der Erstellung von Abschlussarbeiten FMS und Maturaarbeiten

Stand: Juni 2023

Einleitung

Interaktive Sprachmodelle, die auf künstlicher Intelligenz basieren, werden immer besser und können oft sinnvoll und sehr vielfältig eingesetzt werden. Gleichzeitig bergen sie aber auch Risiken, die es abzuwägen gilt.

Die Entscheidung darüber, wann und wie solche Modelle eingesetzt werden sollen, liegt den Nutzenden. Unstrittig ist aber, dass die Ergebnisse nicht unreflektiert verwendet werden sollten.

Der vorliegende Leitfaden ist eine Richtlinie für den Einsatz von interaktiven Sprachmodellen bei der Erstellung und Entwicklung von Abschlussarbeiten FMS und Maturaarbeiten der Kantonsschule Olten und basiert auf Empfehlungen und Leitprinzipien von Schweizer Universitäten und Fachhochschulen.

Die Entwicklung im Bereich der KI ist sehr dynamisch und so basiert der Leitfaden auf den aktuellen Möglichkeiten dieser Technologie. Weitere Veränderungen sind zu erwarten und können an dieser Stelle nicht prognostiziert werden. Erweiterungen und Anpassungen des Leitfadens sind deshalb möglich.

Was können interaktive Sprachmodelle leisten?

Interaktive Sprachmodelle zielen darauf ab, menschliche Sprache und menschliche Kommunikationsmuster möglichst naturgetreu zu imitieren, um etwa über Chats den Kundenservice zu unterstützen oder andere alltägliche Kommunikationssituationen zu bewältigen. Dementsprechend versuchen sie, Texte zu produzieren, die der menschlichen Schreibleistung gleichen.

Die Qualität der Ergebnisse hängt von vielen Faktoren ab (z.B. Konkretisierungsgrad der Eingabe, thematische Komplexität, usw.), weshalb die Resultate solcher Sprachmodelle stets reflektiert werden sollten.

Welches sind typische KI-gestützte Sprachmodelle?

Derzeit gibt es verschiedene Sprachmodelle, welche auf einer KI-Technologie basieren:

ChatGPT ist das bekannteste Tool. Es wurde vom Entwickler OpenAI erstellt. Da unter anderem Microsoft an diesem Unternehmen betei-

liegt ist, werden Teile der Funktionalität von ChatGPT auch in der Suchmaschine Bing integriert. Zukünftig soll es auch integraler Bestandteil von Office-Produkten werden.

DeepL-Write ist ein KI-basierter Schreibassistent, der ebenfalls helfen kann, Texte zu verbessern: Er kann einen Text auf Grammatik, Rechtschreibung und Stil prüfen. Zudem ist er in der Lage, Synonyme und alternative Formulierungen vorzuschlagen.

Muss ich mit ChatGPT arbeiten?

Nein. Du kannst nicht gezwungen werden, da die Nutzung von ChatGPT aus datenschutzrechtlichen Gründen problematisch ist. Die Technologie von ChatGPT gilt für Datenschützer als Blackbox – solange die Verarbeitung und Verwendung eingegebener Daten unklar ist, sollte auf die Eingabe persönlicher Daten verzichtet werden.

Bitte beachte also, dass deine Fragen und Anweisungen von ChatGPT gespeichert werden. Dadurch gibst du persönliche und sensible Informationen über dich preis, deren weitere Verwendung du nicht kontrollieren kannst.

Natürlich ist es dir aber erlaubt, mit ChatGPT zu arbeiten.

Darf ich meine Texte mit interaktiven Sprachmodellen wie DeepL Write verbessern und optimieren?

Genauso wie dir die Rechtschreibprüfung von Word beim Verfassen deiner Arbeit helfen kann, solltest du auch interaktive Sprachmodelle (z.B: DeepL Write) nutzen, um deinen sprachlichen Ausdruck zu optimieren.

Worauf basieren das «Wissen» oder die Ergebnisse von ChatGPT?

Wir wissen es nicht. Das PT im Namen steht für «pre-trained», was bedeutet, dass ChatGPT sicherlich mit Lehrbüchern, Nachrichten, Geschichte, wissenschaftlichen Arbeiten und weiten Teilen des Webs trainiert wurde, alles aber vor September 2021. Über aktuelles Wissen verfügt ChatGPT nicht.

Weil OpenAI seine Wissensquellen nicht offenlegt, ist ChatGPT selbst nicht in der Lage, wissenschaftliche Arbeit zu produzieren. Stattdessen gibt es auf Anfrage zu jedem Thema einen typischen Überblick über das, was das Netz auch zu bieten hat: das Gute, das Schlechte und das Hässliche.

Sind die Resultate von ChatGPT verlässlich?

Nein. Es gibt unzählige Beispiele, in denen ChatGPT halluziniert, nicht existente Zusammenhänge konstruiert und Unwahrheiten produziert. Dabei wirken die Resultate auf den ersten Blick oft überzeugend und plausibel.

Oft sind die gemachten Angaben aber auch sehr gut.

ChatGPT kann also ein äusserst hilfreiches Instrument sein, z.B. um in Themen einzusteigen, da es wahrscheinlich Aspekte und Bereiche vorschlägt, die man vielleicht nicht bedacht hat, aber es ist immer noch deine Aufgabe, die Ergebnisse zu sortieren und zu überprüfen.

Für jede Art des wissenschaftlichen Schreibens kann ChatGPT daher ein guter Einstieg in den Schreibprozess sein.

Muss ich die von ChatGPT übernommenen Gedanken als Zitate ausweisen?

Nein, und zwar aus zwei Gründen:

- Nur Menschen können urheberrechtlich geschützte Werke schaffen. Deshalb gibt es für Texte, die von ChatGPT erstellt wurden, keinen Urheber.
- Eine Autorenschaft ist immer damit verbunden, Verantwortung für die Arbeit zu übernehmen - und das können Sprachmodelle wie ChatGPT nicht.

Gilt ChatGPT als wissenschaftliche Quelle?

Nein. ChatGPT ist keine wissenschaftlich zitierwürdige Quelle.

ChatGPT entbindet dich also nicht davon, wissenschaftliche Literatur zu suchen, zu studieren und verwendete Angaben mit Quellenverweisen zu versehen.

Können von ChatGPT übernommene Passagen mit Plagiatsfindern identifiziert werden?

Nein. ChatGPT-Antworten sind das Ergebnis eines Algorithmus und daher im Allgemeinen nicht reproduzierbar: ChatGPT wird auf identische Fragen unterschiedliche Ergebnisse liefern. Die Antworten sind nicht einfach Fragmente des Textkorpus, der für das Training verwendet wurde, und daher können die Antwortpassagen nicht mit Plagiatsuchprogrammen gefunden werden. ChatGPT plagiiert nicht.

Wer trägt die Verantwortung für die Ergebnisse von KI-basierten Sprachmodellen?

Die Sprachmodelle sind lediglich Werkzeuge zur Berechnung der Antworten. Sie können nicht für ihre Produkte verantwortlich gemacht werden. Wer mit KI arbeitet, ist letztlich immer selbst dafür verantwortlich, wie er/sie die Resultate nutzt.

Wann verletze ich meine Redlichkeitserklärung?

Die Redlichkeitserklärung ist eindeutig: Deine Arbeiten müssen selbständig verfasst werden.

Wer ganze Textpassagen von ChatGPT übernimmt, hat zumindest Teile der Arbeit nicht selbständig verfasst und verstösst damit gegen die Redlichkeitserklärung.

Kann von diesen Richtlinien abgewichen werden?

Die Vielfalt möglicher Themen für Arbeiten ist sehr gross. Es ist denkbar, dass die betreuende Lehrperson in Einzelfällen von diesen Vorgaben abweichen will. Dies ist grundsätzlich zulässig, muss aber abgesprochen und im Arbeitsvertrag festgehalten werden.

Verfasst von der AG-KI in Zusammenarbeit mit der AG-MA