

Textgenerierende KI zur Erleichterung der Arbeit von Lehrpersonen? Ein Vorschlag am Beispiel der Aufgabendifferenzierung mithilfe von ChatGPT

Tobias Bahr, Mario Manzocco, Dennis Schuster und Albrecht Wacker

Zusammenfassung

Textgenerierende KI bietet Lehrpersonen neue Möglichkeiten in der Unterrichtsplanung, um Materialien adaptiv auf die Voraussetzungen der Schüler*innen anzupassen. Der Artikel zeigt, wie Aufgaben mithilfe von ChatGPT differenziert werden können.¹ Am konkreten Beispiel „Beschreibe die Gattung Fabel“ wird eine Differenzierung anhand der allgemeindidaktischen Kriterien der Bloom'schen Taxonomie und der sprachlogischen Komplexität vorgenommen. Hierzu werden zwei Verfahren vorgestellt: Einerseits die direkte Eingabe in ChatGPT, andererseits Eingaben in eine erstellte Webseite. Möglichkeiten und Grenzen des Einsatzes von ChatGPT zur Aufgabendifferenzierung werden abschließend diskutiert.

Schlüsselbegriffe: ● *Schulpädagogik* ● *Medienpädagogik* ● *Large Language Model / LLM* ● *textgenerierende KI* ● *Individualisierung* ● *Heterogenität* ● *Aufgabendifferenzierung* ● *ChatGPT* ● *Unterrichtsplanung*

Einleitung

Spätestens mit der Veröffentlichung von *ChatGPT* im November 2022 wurden große Sprachmodelle (Large Language Models, LLMs) als ein Teilbereich der künstlichen Intelligenz (KI) in der breiten Öffentlichkeit diskutiert (vgl. Knaus et al. 2023; Krenn 2022; Marx et al. 2023; Rizvi/Waite/Sentance 2023). *ChatGPT*, entwickelt von *OpenAI* (US-amerikanische Firma), nutzt als Chatbot das große Sprachmodell *GPT* (kurz für Generative Pretrained Transformer) und greift auf dessen große Datenbank zurück (vgl. Gao et al. 2020). Durch diese Datenbank

¹ Dieser Artikel ist im Rahmen der Innovationsinitiative „Online-Tool zur bildungsplanorientierten Aufgabendifferenzierung für alle Schulfächer“ entstanden, die von der Professional School of Education Stuttgart-Ludwigsburg (PSE) im Jahr 2024 gefördert wird.

ist es für den Chatbot möglich, relativ lange Texte zu generieren. Dadurch wurde *ChatGPT* auch hinsichtlich des Einsatzes im schulischen Kontext diskutiert (vgl. Knaus et al. 2023; Krenn 2022; Marx et al. 2023; Rizvi/Waite/Sentance 2023). So wurden in verschiedenen Beiträgen Potenziale von Chatbots für die Schule hervorgehoben (vgl. Knaus et al. 2023). Diese Potenziale umfassen unter anderem das Generieren von Texten (z. B. für den Unterricht oder für Hausaufgaben), das Bearbeiten oder Adaptieren erstellter Texte (z. B. Korrektur der Grammatik) und das Diskutieren von Themen (z. B. im sokratischen Gespräch, vgl. Opper 2023; Knaus et al. 2023). Mit diesen Potentialen stieg auch die Anzahl an Anleitungen (Artikel und Videos) für Lehrpersonen, *ChatGPT* oder andere auf LLMs basierende Chatbots für die Unterrichtsvorbereitung zu nutzen (vgl. Knaus et al. 2023). Jedoch sind die Ergebnisse von *ChatGPT* öfters sachlich falsch, zu kurz oder es ist eine Verfeinerung der Eingabe notwendig (vgl. de Winter 2023; Knaus et al. 2023). Mit dieser Ausgangssituation stellte sich den Autoren die Frage, inwiefern *ChatGPT* gewinnbringend für die schulische Praxis eingesetzt werden kann.

Im aktuellen Diskurs werden Chatbots auch im Hinblick auf das (fachliche) Lernen der Schüler*innen thematisiert, unter anderem für die Unterrichtsfächer Biologie (vgl. Thyssen 2023), Deutsch (vgl. Birk 2023), Physik (vgl. Rühling 2023) und Technik (vgl. Nepper/Ruch 2023). In manchen Bundesländern wurde zusätzlich das Thema KI in der Sekundarstufe I und II durch von Hochschulen begleitete Projekte in den Informatikunterricht integriert (vgl. z. B. Gräßl 2023; Jatzlau et al. 2019; Lindner et al. 2023). Noch verhalten hingegen werden bislang die Möglichkeiten thematisiert, die Chatbots möglicherweise für die Arbeit der Lehrpersonen beinhalten. In der Literatur finden sich diesbezüglich erste Hinweise darauf, dass KI insgesamt ein Potenzial zum Umgang mit Heterogenität beinhaltet und Chatbots wie *ChatGPT* zur Differenzierung von Aufgaben in der Unterrichtsplanung eingesetzt werden können (vgl. z. B. Fleischmann 2023; Knaus et al. 2023; Nuxoll 2023). Der Beitrag skizziert am Beispiel der allgemeindidaktischen Aufgabendifferenzierung eine Möglichkeit, wie LLMs die Arbeit von Lehrpersonen unterstützen und möglicherweise zu deren zeitlicher Entlastung beitragen können: Im ersten Teil (1) wird die Notwendigkeit der Aufgabendifferenzierung skizziert und die (2) allgemeindidaktische Aufgabenanalyse in ihren Kriterien dargestellt. Anhand (3) der konkreten Aufgabenstellung „Beschreibe die Fabel“ wird die Idee den Leser*innen illustriert. Zwei Möglichkeiten zur praktischen Umsetzung werden vorgestellt: Einerseits (4) eine direkte Eingabe in *ChatGPT*, andererseits (5) die Verwendung einer hierzu angelegten Webseite. Eine (6) Zusammenfassung und kritische Würdigung, die auch auf die Grenzen verweist, rundet den Artikel ab.

1. Notwendigkeit der Aufgabendifferenzierung in der Schule

In der Nachkriegszeit wurde das dreigliedrige Schulsystem der Weimarer Republik weitgehend restauriert und in den 1960er-Jahren dreigliedrig ausgebaut. Hier versuchte man, der leistungsbezogenen Heterogenität der Schüler*innen mit äußerer Differenzierung der Schularten (Hauptschule, Realschule, Gymnasium) zu begegnen. Fragen der inneren Differenzierung wurden durch die neuen Gesamtschulen, die in den 1970er-Jahren in neun Bundesländern eingeführt wurden (vgl. Wacker 2023), angestoßen und zunehmend diskutiert. Diese Diskussion wurde nochmals intensiviert durch die seit etwa 2007 vielfach eingeführten neuen Sekundarschularten, „die zwei oder drei Bildungsgänge zusammenfassen“ (Ridderbusch 2022,

S. 151) und damit die Frage der Binnendifferenzierung akzentuiert stellen. An den meisten Sekundarschularten erweiterten diese Strukturreformen die Leistungsheterogenität der Schülerschaft und machen (bis heute) darauf bezogene adaptive Konzepte erforderlich.

In diesem Beitrag betrachten wir die innere Differenzierung über die Variation von Aufgaben als eine Möglichkeit zum Umgang mit den leistungsheterogenen Voraussetzungen von Schüler*innen. Sie umfasst die Differenzierung des Unterrichtsmaterials und der Lernaufgaben (und auch daran anknüpfend der Leistungsaufgaben und der Leistungsbeurteilung) entsprechend der Voraussetzungen der Schüler*innen). Damit steigen die Anforderungen an die Lehrpersonen in der vorausgehenden Unterrichtsplanung. Als Herausforderungen sind u. a. die Aufgabendifferenzierung auf verschiedenen Niveaustufen – häufig auch für Schüler*innen mit sonderpädagogischem Förderbedarf im Rahmen ihrer inklusiven Beschulung an Regelschulen – sowie die damit verbundenen zeitlichen Ressourcen zu nennen.

Wir greifen in unserem Beitrag auf den gemeinsamen Bildungsplan für die Sekundarstufe in Baden-Württemberg zurück, in dem eine Graduierung über das grundlegende Niveau (G-Niveau), das mittlere Niveau (M-Niveau) und das erweiterte Niveau (E-Niveau) konstitutiv verankert wurde (vgl. Pant 2016, S. 13). Die Niveaustufen des Bildungsplans unterscheiden sich (zumeist) hinsichtlich der Durchdringungstiefe und des Abstraktionsgrads (vgl. Pant 2016, S. 13), enthalten aber kaum Anhaltspunkte für eine gelingende pädagogisch-didaktische Aufgabendifferenzierung, wie diese von den Lehrpersonen im Hinblick auf die spezifischen Voraussetzungen der Schüler*innen eigens zu leisten ist. Sie erfordert die Planung konkreter Aufgaben, die adaptiv an die jeweiligen Voraussetzungen der Lernenden anschließen. Hier können LLMs vor allem in ihrer Funktion eingesetzt werden, Lehrmaterialien umfänglicher zu individualisieren, um über die gezielte „Förderung *aller* Schülerinnen“ (vgl. Knaus et al. 2023, S. 20, Herv. durch Autor) letztlich zur Bildungsgerechtigkeit beizutragen (vgl. Kreye 2023, S. 41).

Bevor wir die Verwendungsmöglichkeiten von Chatbots (hier *ChatGPT*) an einem Beispiel illustrieren, ist es erforderlich, die allgemeindidaktischen Kriterien der Aufgabenanalyse, die wir hierfür in Betracht ziehen, zu skizzieren.

1. Allgemeindidaktische Kriterien der Aufgabenanalyse und -differenzierung

In diesem Beitrag beschränken wir uns auf allgemeindidaktische Kriterien zur Aufgabendifferenzierung, also auf Kriterien, die über alle Fächer hinweg Gültigkeit beanspruchen können, und greifen hierfür auf einen Vorschlag von Maier und seiner Arbeitsgruppe zurück (vgl. Maier et al. 2010). Diese Wahl kann zum einen pragmatisch mit dem Ziel legitimiert werden, allen Leser*innen, unter denen (angehende) Lehrpersonen den Hauptteil ausmachen dürften, ungeachtet ihrer zu unterrichtenden Fächern eine Anschauung zu geben. Selbstredend sind diese allgemeindidaktischen Kriterien in einem weiteren Schritt um fachdidaktisch begründete Merkmale aus den einzelnen Fächern und Fachbereichen zu erweitern und zu ergänzen. Der überfachliche Fokus begründet sich zum anderen inhaltlich im Sachverhalt, dass ein wesentlicher Bereich der allgemeindidaktischen Aufgabenanalyse in der Bestimmung des kognitiven Potenzials von Aufgaben liegt und sie deshalb für die (meisten) Fächer herangezogen werden kann.

Maier und seine Arbeitsgruppe benennen im erwähnten Übersichtsbeitrag zur Analyse des

kognitiven Potenzials folgende Merkmale, die unter Rückgriff auf allgemeindidaktische Lernzieltaxonomien und Befunde der Kognitionspsychologie gewonnen wurden: (1) Die Art des in Aufgaben geforderten Wissens (z. B. Faktenwissen oder prozedurales Wissen), (2) die in Aufgaben verlangten kognitiven Prozesse (z. B. Reproduktion oder Transfer), (3) die Anzahl der in Aufgaben benannten oder verlangten Wissenseinheiten, (4) die Offenheit der Aufgabenstellung (z. B. eindeutiger Arbeitsauftrag, der mehrere Lösungen zulässt), (5) der Lebensweltbezug (Aufgaben ohne und mit realem Lebensweltbezug), (6) die sprachlogische Komplexität der Aufgabenstellung (z. B. einfach oder hoch) und (7) Repräsentationsformen des Wissens, die wir hier jedoch vernachlässigen wollen (vgl. Maier et al. 2010). Diese Kriterien zur Analyse bieten sich zur Differenzierung an, bedürfen aber der Auswahl und Reduktion. Für unser Beispiel greifen wir lediglich auf zwei allgemeindidaktische Differenzierungskriterien zurück:

- Die Bloom'sche Taxonomie (vgl. Bloom 1956) in ihrer revidierten Fassung durch Anderson und Krathwohl (2001) und
- die sprachlogische Komplexität (vgl. Jordan et al. 2006; Maier et al. 2010).

Die von Bloom erstmals eingeführte Taxonomie, später durch Anderson und Krathwohl (2001) revidiert, beschreibt die Domäne kognitiver Fähigkeiten in sechs hierarchischen Stufen (Erinnern, Verstehen, Anwenden, Analysieren, Bewerten, Erschaffen). Sie bildet damit eine Grundlage, um Aufgaben aus einer allgemeindidaktischen Perspektive hinsichtlich ihrer kognitiven Anforderungen zu differenzieren. Auch wenn an dieser Stufung und insbesondere an den drei hierarchisch übergeordneten Stufen von Bloom vielfach Kritik geübt wurde (auf die wir hier nicht näher eingehen können), verwenden wir dieses Modell aufgrund seiner Bedeutung. Es bildet ebenfalls die erkennbare Grundlage, mit denen die erwähnten Niveaustufen des Bildungsplans in ihrem Abstraktionsgrad und ihrer Durchdringungstiefe differenziert sind (vgl. Pant 2016, S. 13). Die sprachlogische Komplexität ziehen wir ergänzend heran.

2. Die konkrete Aufgabenstellung „Beschreibe die Gattung Fabel“

Ziel ist es, zu einer gegebenen Aufgabe differenzierte Aufgaben mit Hilfe von *ChatGPT* zu erstellen. Für das Anwendungsbeispiel der Aufgabendifferenzierung greifen wir im Beispiel primär auf die Software ChatGPT-4 zurück (vgl. OpenAI 2024a). Die Verwendung dieser (kostenpflichtigen) Software führt unseres Erachtens nach zu besseren Resultaten im Vergleich zum kostenfreien Programm ChatGPT-3.5 Modell (vgl. OpenAI 2024b).

Wir wählen hierzu exemplarisch das Thema „Fabel“ aus dem Fach Deutsch aus, das in den Lehrplänen aller Bundesländer als Gegenstand verankert ist, und greifen in unserem Beispiel exemplarisch auf den erwähnten gemeinsamen Bildungsplan für die Sekundarstufe in Baden-Württemberg zurück, in dem eine Graduierung über drei Niveaustufen konstitutiv verankert wurde (vgl. Pant 2016, S. 13). Aus den oben angeführten Punkten zur Unterteilung der Niveaustufen stellt das G-Niveau den Ausgangspunkt für unser Fallbeispiel dar; auf die Ausgestaltung der weiteren Niveaustufen im Bildungsplan gehen wir nicht ein.

Im gemeinsamen Bildungsplan für die Sekundarstufe in Baden-Württemberg für die Klassenstufen 5 und 6 ist als inhaltsbezogene Kompetenz im G-Niveau vermerkt: „Die Schülerinnen und Schüler können unter Verwendung von einigen Gattungsmerkmalen einfache Gedichte und epische Kleinformen (Märchen, Fabel) beschreiben“ (vgl. Ministerium für Kultus,

Jugend und Sport Baden-Württemberg 2016, S. 19). Diese Kompetenz erfordert von den Schüler*innen ein Wissen um die konstitutiven Merkmale und Verlaufsformen der Gattung Fabel, das die Grundlage bildet, um später Texte werten und einordnen zu können (vgl. Ministerium für Kultus, Jugend und Sport Baden-Württemberg 2016, S. 19). Für den formulierten Arbeitsauftrag gilt es nun, differenzierte Aufgaben über allgemeindidaktische Kriterien der Aufgabenanalyse zu gewinnen. Hierzu stellen wir sowohl die Möglichkeit vor, Eingaben in der Software selbst vorzunehmen (vgl. Abschnitt 3) als auch diese über eine Website einzugeben, die wir entwickelt haben (vgl. Abschnitt 4).

3. Aufgabendifferenzierung über Prompts

Der nachfolgende Abschnitt beschreibt, wie mit ChatGPT Eingaben getätigt werden können und illustriert dies an einem Beispiel.

Eingaben in ChatGPT

Mit direkten Eingaben in ChatGPT sollen nun Aufgaben zur Beispiel-Aufgabe „Beschreibe die Gattung Fabel“ differenziert werden. Dabei ist zu beachten, dass immer die gleiche Konversation (bzw. Chatinstanz) genutzt werden sollte und bereits getätigte Eingaben nicht gelöscht werden, da ChatGPT die vorherigen Eingaben (Differenzierungsparameter) speichert. Zuerst muss der Sachverhalt in eine für ChatGPT verständliche Form gebracht werden. Die Eingaben in den Chat, die hierzu erforderlich sind, werden Prompts genannt. Die Nutzer*innen können über diese Prompts mit ChatGPT kommunizieren und sie werden im Chatfenster am unteren Bildschirmrand getätigt (vgl. Abbildung 1).

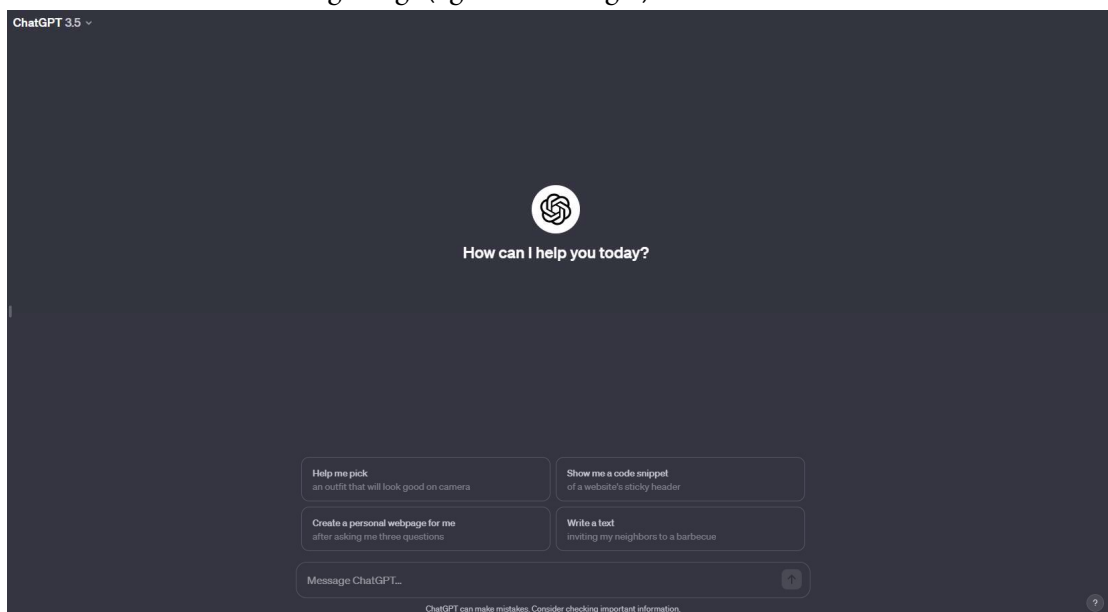


Abbildung 1: Screenshot von ChatGPT

Um eine differenzierte Aufgabenstellung zu erhalten, bietet sich der nachfolgende Prompt an:
„Vereinfache die nachfolgende Aufgabe so, dass die Bloom'sche Taxonomiestufe danach "Erinnern" entspricht.

Beschreibe die Gattung Fabel.“

ChatGPT gibt als Ergebnis bzw. Output dann zum Beispiel „Was ist eine Fabel?“ aus. Mit diesem Beispiel sollte demonstriert werden, dass in kürzester Zeit einfache Aufgaben generiert werden können. Jedoch wird in diesem Output (der Ausgabe von ChatGPT) kein Operator verwendet, was bereits einen Hinweis darauf gibt, dass die Formulierung der Prompts in ChatGPT differenzierter erfolgen muss. ChatGPT verwendet eben nur die Kriterien zur Differenzierung, die von der nutzenden Person mitgeteilt werden.

Zur Eingabe „Verändere die nachfolgende Aufgabe so, dass die Bloom’sche Taxonomiestufe danach ‚Analysieren‘ entspricht und generiere einen Hinweis. Beschreibe die Gattung Fabel“, erhalten Nutzer*innen zum Beispiel das Ergebnis: „Analysiere die Struktur und die typischen Merkmale einer Fabel. Gebe dabei einen Hinweis darauf, wie die moralische Lehre in einer Fabel präsentiert wird.“

Die beiden Beispiele verdeutlichen, dass gute Prompts verschiedene Kriterien beinhalten sollten, um hilfreiche Resultate zu erzielen (vgl. OpenAI 2024c). Zu diesen gehören unter anderen:

- präzise Formulierungen der Prompts,
- die Angabe von Kontexten (in unserem Beispiel die Kriterien, nach denen differenziert werden soll),
- die Angabe von personalen Rollen der Nutzenden (in unserem Beispiel für eine erfahrene Lehrperson der Sekundarstufe I); hier könnte jedoch auch eingegeben werden, dass die Antwort wie von einem Historiker, einem Zeitzeugen, einem Journalisten oder einem Drehbuchautor gegeben werden soll (vgl. Kreye 2023, S. 41),
- die Verfeinerung von Prompts in iterativen Schleifen unter Angabe von Feedback, ob ein (bisheriges) Ergebnis gut war oder nicht (vgl. OpenAI 2024c).

Daher wird es von uns als sinnvoll erachtet, mit Prompts, die diese Kriterien erfüllen, zunächst in Schritt 1 ChatGPT zu instruieren. Für die Aufgabendifferenzierung wird nachfolgend in weiteren Schritten beschrieben, welche Prompts in welcher Reihenfolge bei ChatGPT eingegeben werden müssen.

Schritt 1: Initialisierung

Für Schritt 1 wird der Prompt aus dem Anhang (siehe Prompt im Anhang dieses Beitrags) kopiert und in das Chatfenster eingefügt. Diesen Prompt fügt die benutzende Person in den Chat per „Copy und Paste“ ein und bestätigt. Anschließend antwortet ChatGPT mit „Bitte geben sie mir eine Aufgabe, die ich analysieren soll.“

Schritt 2: Eingabe und Analyse der Aufgabe

Nun wird die zu differenzierende Aufgabe eingegeben und analysiert. In unserem Fallbeispiel ist es die Aufgabe „Beschreibe die Gattung Fabel“. Dazu wird folgender Prompt genutzt:

Bitte analysieren Sie die folgende Schulaufgabe basierend auf den vorgegebenen Metriken:

Beschreibe die Gattung Fabel

Für die Differenzierung anderer Schulaufgaben können Anwender*innen in den Prompt an der dunkelgrauen Stelle eigene Aufgaben eingeben.

ChatGPT analysiert die Aufgabe „Beschreibe die Gattung Fabel“ nach der Beschreibung von Schritt 1. Zunächst analysiert ChatGPT die eingegebene Aufgabe und gibt dies als Ergebnis aus. Die Ausgabe von ChatGPT beinhaltet eine Ziffer für die Lernzielstufe, eine Ziffer für die sprachlogische Komplexität sowie einen wahren oder falschen Wert für den Hinweis und das Beispiel.

In diesem Beispiel sagt ChatGPT, dass die aktuelle Lernzielstufe der eingegebenen Aufgabe der Ziffer 2 (Verstehen) und die sprachlogische Komplexität der Ziffer 3 (mittleres Niveau) entsprechen (Abbildung 2).

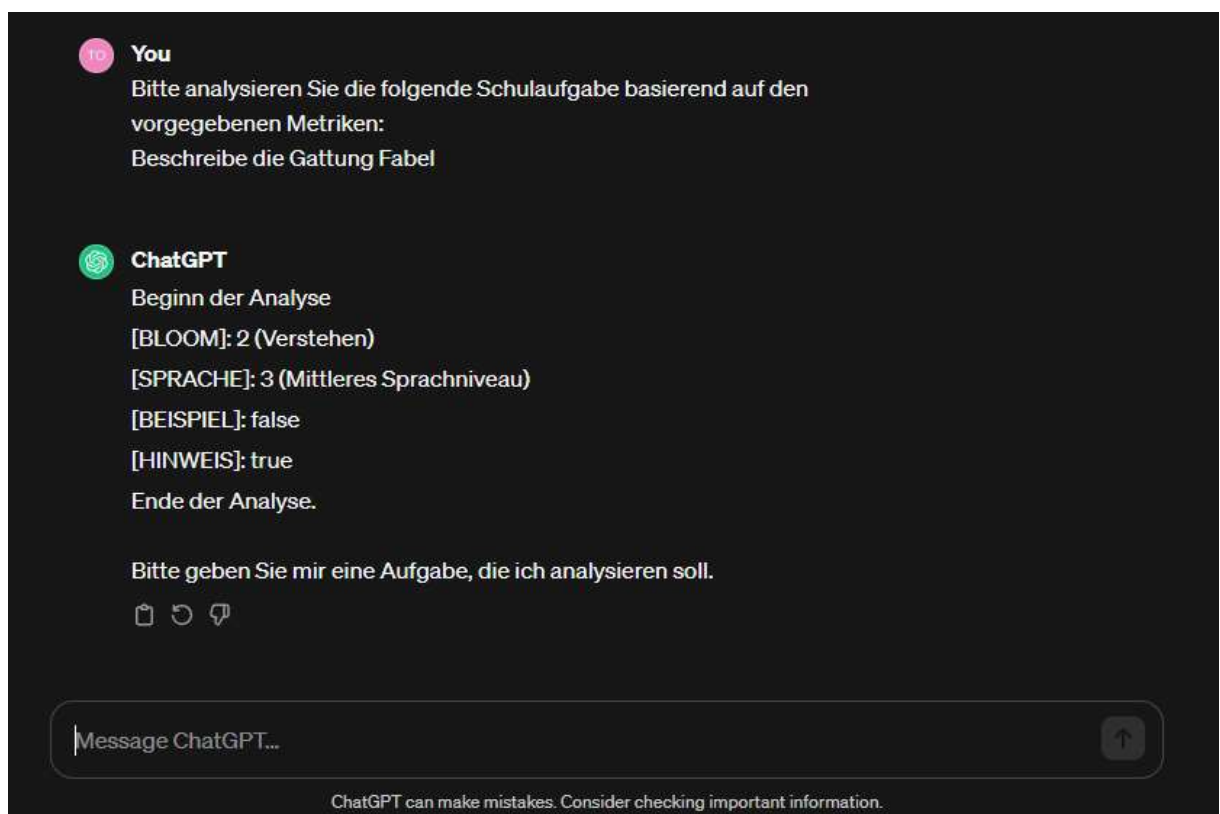


Abbildung 2: Screenshot von ChatGPT zur Ausgabe von Schritt 2

Diese Aufgabe soll nun im nächsten Schritt differenziert werden.

Schritt 3: Differenzierte Aufgabe erstellen

Zunächst werden in ChatGPT das Lernziel des Bildungsplans und die Differenzierungskriterien hinterlegt, nach denen die differenzierte, „neue“ Aufgabe erstellt werden soll. Hierfür wird der nachfolgende Prompt kopiert und ausgeführt.

Bitte erstelle eine angepasste Version der zuvor analysierten Schulaufgabe. Diese Anpassung sollte den folgenden Kriterien entsprechen, wobei das ursprüngliche Lernziel beibehalten wird:

[BLOOM] : (Wert zwischen 1 und 6 einfügen)

[SPRACHE] : (Wert zwischen 1 und 6 einfügen)

[BEISPIEL] : (true/false)

[HINWEIS] : (true/false)

Beachten Sie, dass das Lernziel und die Kernintention der Aufgabe unverändert bleiben sollen und das Sprachniveau für die Klassen 5 bis 10 geeignet ist.

Das Lernziel für die Aufgabe ist: (Lernziel einfügen)

Genau wie in Schritt 2 sind die dunkelgrauen Felder für die Benutzer*innen zur Anpassung gedacht. Für das Anwendungsbeispiel soll nun die Lernzielstufe auf 1 (Erinnern) und die sprachlogische Komplexität auf 2 (einfaches Niveau) verändert werden. Die Angaben zum Beispiel und dem Hinweis bleiben auf falsch, sodass nichts erstellt wird. Außerdem muss am Ende des Prompts das Lernziel aus dem Bildungsplan eingefügt werden: „Die Schülerinnen und Schüler können unter Verwendung von einigen Gattungsmerkmalen einfache Gedichte und epische Kleinformen (Märchen, Fabel) beschreiben“ (vgl. Ministerium für Kultus, Jugend und Sport Baden-Württemberg 2016, S. 19). Der/die Anwender*in erhält nachfolgende Ausgabe (Abbildung 3).

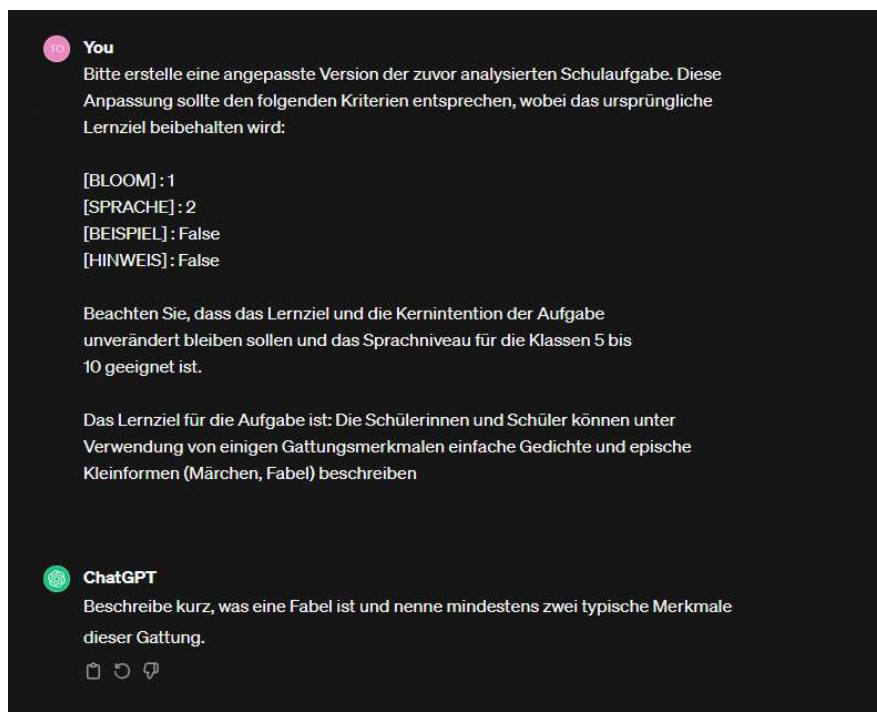


Abbildung 3: Screenshot von ChatGPT zur Ausgabe von Schritt 3

Schritt 4: Weitere differenzierte Aufgaben erstellen

In unserem Beispiel, das auf den Bildungsplan in Baden-Württemberg zielt, ist es wie o. a. möglich, differenzierte Aufgaben für das grundlegende (G), das mittlere (M) und das erweiterte (E) Niveau des gemeinsamen Bildungsplans für die Sekundarstufe I zu erstellen und darüber hinaus noch weitere Aufgaben zu erhalten, die auf die Voraussetzungen der Schüler*innen zugeschnitten sind und eine Individualisierung des Lernens befördern. Dafür muss der nachfolgende Prompt mit den jeweiligen Informationen ausgeführt werden:

Ich benötige differenzierte Aufgaben für die (Klassenstufe einfügen) te Klasse. Das Fach ist: (Fach einfügen), das Thema ist: (Thema einfügen) und das Lernziel ist: (Lernziel einfügen). Bitte erstelle Aufgaben für die Schwierigkeitsstufen G-Niveau, M-Niveau und E-Niveau entsprechend den angegebenen Kriterien.

Hier sind, analog zur Vorgehensweise in Schritt 2 und 3, in den dunkelgrauen Kästchen die für die benutzende Person relevanten Informationen einzufügen. Die Klassenstufe, das Fach, das Thema und das Lernziel des Bildungsplans sind hierfür an die Stellen des dunkelgrauen Kästchens kopiert worden. Basierend auf den vorliegenden allgemeindidaktischen Differenzierungskriterien im Chat kann ChatGPT nun neue Aufgaben auf den drei Niveaustufen, und sofern gewünscht, auch darüber hinaus erstellen.

4. Webseite zur Aufgabendifferenzierung

Um den oben erläuterten Prozess der ChatGPT-Nutzung zu vereinfachen, wurde von unserer Arbeitsgruppe eine geschützte Webseite erstellt.² Abbildung 4 zeigt einen Screenshot der Webseite mit dem oben aufgeführten Fallbeispiel.

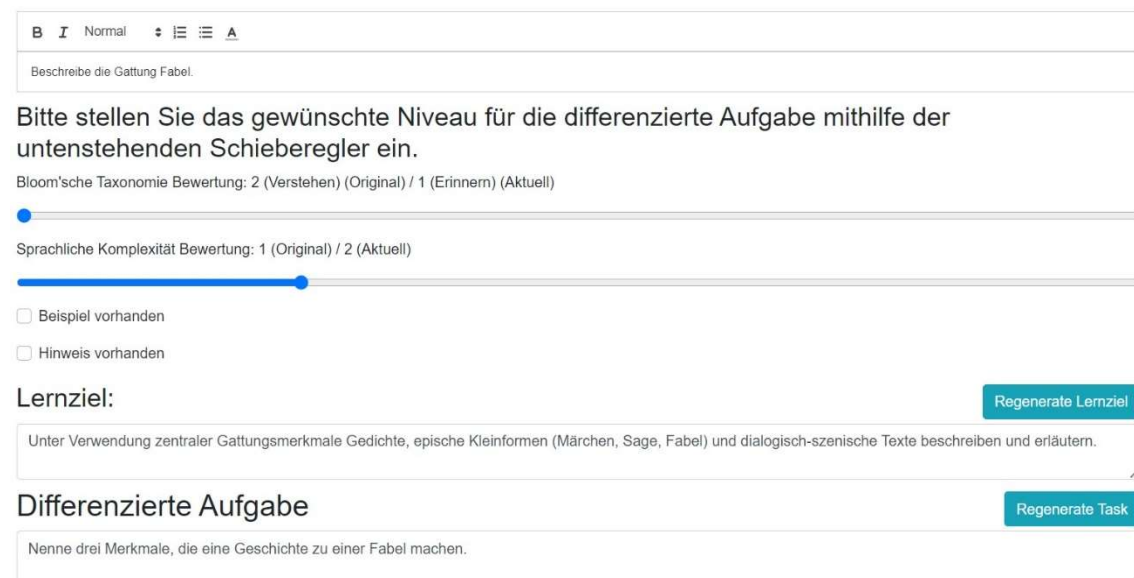


Abbildung 4: Anwendungsbeispiel zur Aufgabendifferenzierung (eigene Darstellung)

² Zugang zur Webseite können Interessierte über eine Anfrage an die Autoren des Artikels erhalten.

Die Webseite greift im Hintergrund auf ChatGPT zu und bietet die Möglichkeit, die in Ziffern ausgedrückten Abstufungen (siehe Beispiel oben) mittels eines Schiebereglers einzugeben.

Wie funktioniert nun die Differenzierung mithilfe unserer Webseite? Im Vergleich zur direkten Nutzung von ChatGPT entfällt der dort beschriebene Schritt 1, da eine Initialisierung automatisch von der Webseite vorgenommen wird. Analog zu Schritt 2 ist die nutzende Person aufgefordert, die zu differenzierende Aufgabe „Beschreibe die Gattung Fabel“ einzugeben. Dies geschieht im oberen Textfenster (Abbildung 4).

Im nächsten Schritt steht unter dem Textfenster (Abbildung 4) die Aufforderung „Bitte stellen Sie das gewünschte Niveau für die differenzierte Aufgabe mithilfe der untenstehenden Schieberegler ein.“. Die Schieberegler stellen die Stufen der Bloom'schen Taxonomie und die sprachlogische Komplexität dar. Diese wurden wie im Anwendungsbeispiel oben (vgl. Abschnitt 3) auf 1 (Erinnern) und 2 (einfaches Niveau) gestellt (vgl. Abbildung 4, blaue Schieberegler).

Auf der Webseite wird der benutzenden Person die aktuelle Stufe über dem Schieberegler angezeigt. So steht dort „1 (Erinnern) (Aktuell)“ und „2 (Aktuell)“ (vgl. Abbildung 4).

Wie in Schritt 3 des Anwendungsbeispiels muss die benutzende Person das Lernziel aus dem Bildungsplan kopieren und in das Textfeld unter „Lernziel:“ einfügen (vgl. Abbildung 4). Zudem besteht optional die Möglichkeit, von ChatGPT ein Lernziel generieren zu lassen.

Durch Klicken auf den türkisfarbenen Button „Regenerate Task“ in der unteren rechten Bildschirmcke wird eine neue differenzierte Aufgabe erstellt. Diese lautet „Nenne drei Merkmale, die eine Geschichte zu einer Fabel machen.“ und erscheint im untersten Textfeld nach der Überschrift „Differenzierte Aufgabe“ (rote Markierung, vgl. Abbildung 5).

B I Normal + ≡ ≡ A

Beschreibe die Gattung Fabel.

Bitte stellen Sie das gewünschte Niveau für die differenzierte Aufgabe mithilfe der untenstehenden Schieberegler ein.

Bloom'sche Taxonomie Bewertung: 2 (Verstehen) (Original) / 1 (Erinnern) (Aktuell)

Sprachliche Komplexität Bewertung: 1 (Original) / 2 (Aktuell)

Beispiel vorhanden

Hinweis vorhanden

Lernziel: Regenerate Lernziel

Unter Verwendung zentraler Gattungsmerkmale Gedichte, epische Kleinformen (Märchen, Sage, Fabel) und dialogisch-szenische Texte beschreiben und erläutern.

Differenzierte Aufgabe Regenerate Task

Nenne drei Merkmale, die eine Geschichte zu einer Fabel machen.

Abbildung 5: Ergebnis der Differenzierung (eigene Darstellung)

The screenshot shows a digital interface for task differentiation. At the top, there is a text input field containing the prompt "Beschreibe die Gattung Fabel." Below this, there are two horizontal sliders. The first slider is labeled "Bitte stellen Sie das gewünschte Niveau für die differenzierte Aufgabe mithilfe der untenstehenden Schieberegler ein." and "Bloom'sche Taxonomie Bewertung: 2 (Verstehen) (Original) / 3 (Anwenden) (Aktuell)". The second slider is labeled "Sprachliche Komplexität Bewertung: 1 (Original) / 2 (Aktuell)". Below the sliders are two checkboxes: "Beispiel vorhanden" and "Hinweis vorhanden", both of which are unchecked. There are two buttons: "Regenerate Lernziel" and "Regenerate Task". The "Lernziel:" section contains the text: "Unter Verwendung zentraler Gattungsmerkmale Gedichte, epische Kleinformen (Märchen, Sage, Fabel) und dialogisch-szenische Texte beschreiben und erläutern." The "Differenzierte Aufgabe" section contains the text: "Anhand eines Fabeltextes deiner Wahl, erläutere wie typische Merkmale einer Fabel in diesem Text zur Anwendung kommen."

Abbildung 6: Zweites Anwendungsbeispiel zur Aufgabendifferenzierung (eigene Darstellung)

Im Hinblick auf die Zielgruppe von Schüler*innen der Klassenstufe 6, die wir hier im Beispiel im Blick haben, erscheint es uns angemessener, diesen eine Fabel vorzugeben anstatt wie in der erstellten Aufgabe einen Fabeltext „deiner Wahl“ auszusuchen.

Unabhängig davon, welche Version (kostenpflichtig oder nicht) von ChatGPT oder anderen KI-Chatbots benutzt wird, sollten die Ergebnisse unserer Meinung nach unbedingt auf ihre Plausibilität, beispielsweise ihre Passung auf die Lernziele und die Voraussetzungen der Schüler*innen, überprüft werden (und ggf. geändert werden). Dies zeigt sich in der Literatur auch für andere Anwendungsbereiche (vgl. de Winter 2023; Knaus et al. 2023).

5. Zusammenfassung und Ausblick

Lehrpersonen stehen u. a. vor dem Hintergrund zahlreicher Schulreformen, welche insbesondere in den Sekundarstufen die Leistungsheterogenität der Lerngruppen erhöhten, vor der Herausforderung, sowohl Lernaufgaben für den Unterricht als auch Aufgaben für die Leistungsbeurteilung zu differenzieren. Im Beitrag wird hierzu eine Möglichkeit vorgestellt, differenzierte Aufgaben einerseits über Prompts, die direkt in KI-Chatbots (hier am Beispiel von ChatGPT) eingegeben werden können, andererseits über eine Webseite, die auf einen KI-Chatbot zugreift, zu entwickeln. Um beide Möglichkeiten zu illustrieren, wurde ein Beispiel aus dem Fach Deutsch des Gemeinsamen Bildungsplans für die Sekundarstufe in Baden-Württemberg herangezogen und dieses nach (lediglich) zwei allgemeindidaktischen Kriterien differenziert. Das gewählte Beispiel zur Gattung Fabel kann auf andere Fächer und Bundesländer übertragen werden.

Im Ergebnis verweisen beide Verfahren auf Variationsmöglichkeiten: Wenn Lehrpersonen im Anwendungsbeispiel „Beschreibe die Gattung Fabel“ andere Werte bei den Differenzierungskriterien eingeben („3“ Anwenden, „2“ einfache sprachliche Komplexität), wird eine veränderte Aufgabe ausgegeben (Abbildung 6). Dies erfordert einen prüfenden Blick der Lehrpersonen vor dem Hintergrund ihrer fachwissenschaftlichen und fachdidaktischen Profession sowie im Hinblick auf die Voraussetzungen der Schüler*innen.

In weitergehenden Forschungen ist geplant, einen systematischen Satz an Aufgaben eines Unterrichtsfaches mit variierenden Kriterien zu differenzieren. Anschließend sollen die Aufgaben von Expert*innen aus der Bildungspraxis anhand verschiedener Kriterien bewertet werden. Diese könnten unter anderem die schulpraktische und curriculare Relevanz, Verständlichkeit der Aufgabenstellung für die Zielgruppe und weitere Kriterien sein. Somit wäre es möglich, grundlegende Aussagen zur Qualität der differenzierten Aufgaben mit der Webseite bezüglich eines Faches zu tätigen.

In unserem illustrierenden Beitrag ziehen wir allgemeindidaktische Kriterien als bedeutsam heran, weil sie über den Fächern liegende Merkmale zur Aufgabendifferenzierung darstellen. Selbstredend bedarf der Vorschlag darüber hinaus in einem weiteren Schritt des Einbezugs von fachdidaktischen Kriterien, wie sie in letzter Zeit zunehmend in den Fachdidaktiken für die einzelnen Fächer und Fachbereiche theoretisch und empirisch herausgearbeitet worden sind (vgl. z. B. Reinhardt/Rehm/Wilhelm 2021, u.v.a.m.). Um die Webseite hier weiter zu entwickeln, sind Workshops mit verschiedenen Stakeholder*innen aus allen Phasen der Lehrerbildung angedacht.

Als zwingend notwendig zeigt sich, zur Differenzierung den jeweils gültigen Lehrplan grundzulegen. Mit ihm sind die (politischen) Leitplanken hinsichtlich der fachlichen und überfachlichen Ziele vorgegeben und durch seine Verwendung sichergestellt, dass der KI-Chatbot Vorschläge auf der Grundlage dieses Fundaments unterbreitet (vgl. Kreye 2023, S. 41).

Insgesamt bleibt festzuhalten, dass der Einsatz von KI-Chatbots in der Unterrichtsplanung helfen könnte, differenzierte und an die Voraussetzungen der Schüler*innen ansetzende Aufgaben zu erstellen. Ob damit auch eine wünschenswerte Zeitersparnis für die Lehrpersonen einhergeht und ob die Aufgaben adaptiver, anspruchsvoller und kreativer werden, bleibt der weiteren Forschung im Projekt (siehe Fußnote) vorbehalten. Neben den oben skizzierten Forschungsdesideraten planen die Autoren, die Webseite nach den Workshops von Lehrpersonen im Schulnetzwerk digital.macht.schule einsetzen zu lassen, um in Form von qualitativen Rückmeldungen zur Optimierung und zum Thema Zeitersparnis belastbare Aussagen zur Nutzung von KI-Chatbots zur Zeitersparnis zu generieren.

Literatur

- Anderson, Lorin Willard/Krathwohl, David Reading (2001): *A Taxonomy for Learning, Teaching and Assessing: A Revision of Bloom's Taxonomy of Educational Objectives*. New York: Longman.
- Birk, Christian (2023): ChatGPT im Deutschunterricht: Kann künstliche Intelligenz den Schreibunterricht bereichern? In: *Schulmagazin*, 5-10, 2023 (5+6), S. 34–39.
- de Winter, Joost Cornelis (2023): Can ChatGPT Pass High School Exams on English Language Comprehension?. *International Journal of Artificial Intelligence in Education*. doi.org/10.1007/s40593-023-00372-z.
- Fleischmann, Daniel (2023): ChatGPT ist ein tolles Erklärinstrument. *Transfer. Berufsbildung in Forschung und Praxis*.
- Gao, Leo/Biderman, Stella/Black, Sid/Golding, Laurence/Hoppe, Travis/Foster, Charles/Phang, Jason/He, Horace/Thite, Anish/Nabeshima, Noa/Presser, Shawn/Leahy, Connor (2020): *The pile: An 800gb dataset of diverse text for language modeling*. arXiv preprint [arXiv:2101.00027](https://arxiv.org/abs/2101.00027).
- Gräßl, Cornelia (2023): Unravel – Bilderkennung mit neuronalen Netzen. In: *Informatische Bildung in Schulen*, 1(1), S. 1–11. doi.org/10.18420/ibis-01-01-09.

- Jatzlau, Sven/Michaeli, Tilman/Seegerer, Stefan/Romeike, Ralf (2019): It's not magic after all – machine learning in snap! using reinforcement learning. In: 2019 IEEE blocks and beyond workshop (B&B), pp. 37–41. IEEE.
- Jordan, Alexander/Ross, Nathalie/Krauss, Stefan/Baumert, Jürgen/Blum, Werner/Neubrand, Michael/Löwen, Katrin/Brunner, Martin/Kunter, Mareike (2006): Klassifikationsschema für Mathematikaufgaben: Dokumentation der Aufgabenkategorisierung im COACTIV-Projekt. Berlin: MPI.
- Knaus, Thomas et al (2023): KI in Schule und Hochschule – Impulse aus bildungstheoretischer Perspektive. In: Ludwigsburger Beiträge zur Medienpädagogik – LBzM, 23, S. 1–42. doi.org/10.21240/lbzm/23/19.
- Krenn, Daniela (2022): Was Chat GPT kann und vor allem: was (noch) nicht. falter.at/mail/20221214/was-chat-gpt-kann-und-vor-allem-was-noch-nicht.
- Kreye, Adrian (2023): Mein Lehrer, Herr KI. Süddeutsche Zeitung vom 30.09.2023, S. 41.
- Lindner, Annabel/Müller-Unterwiesing, Michaela/Berges, Marc/Rösch, Mathias (2023): KI in der Kiste: Unterrichtsaktivitäten aus einem Lernlabor zum Thema Künstliche Intelligenz. Informatische Bildung in Schulen, 1 (1), S. 1–11. 10.18420/ibis-01-01-03.
- Maier, Uwe/Kleinknecht, Marc/Metz, Kerstin/Bohl, Thorsten (2010): Ein allgemeindidaktisches Kategoriensystem zur Analyse des kognitiven Potenzials von Aufgaben. Beiträge zur Lehrerbildung, 28, S. 84–96. 10.25656/01:13734.
- Marx, Erik/Leonhardt, Thimo/Bergner, Nadine (2023): Secondary school students' mental models and attitudes regarding artificial intelligence-A scoping review. In: Computers and Education: Artificial Intelligence, 5, 100169. doi.org/10.1016/j.caeai.2023.100169.
- Ministerium für Kultus, Jugend und Sport Baden-Württemberg (2016): Gemeinsamer Bildungsplan für die Sekundarstufe I. Fach Deutsch. www.bildungsplaene-bw.de/site/bildungsplan/get/documents/lbw/export-pdf/depot-pdf/ALLG/BP2016BW_ALLG_SEK1_D.pdf.
- Nepper, Hannes Helmut/Ruch, Armin (2023): ChatGPT. Implikationen für den Technikunterricht. In: technik-education (tedu). Fachzeitschrift für Unterrichtspraxis und Unterrichtsforschung im allgemeinbildenden Technikunterricht, 3 (1), S. 3–10. URN: urn:nbn:de:0111-pedocs-268112 - DOI: 10.25656/01:26811.
- Nuxoll, Florian (2023): ChatGPT: Create a catchy headline: Wie sich chatbots und KI tools auf den Englischunterricht auswirken. In: Der fremdsprachliche Unterricht Englisch, 2023 (182), S. 46– 48.
- OpenAI (2024a): Prompt engineering. platform.openai.com/docs/guides/prompt-engineering.
- OpenAI (2024b): GPT-4. openai.com/research/gpt-4.
- OpenAI (2024c): GPT-3.5. openai.com/research/gpt-3-5.
- Opper, Katharina (2023): Im Sokratischen Gespräch mit KI. Hochschulforum Digitalisierung. hochschulforumdigitalisierung.de/im-sokratischen-gespraech-mit-ki.
- Pant, Hans Anand (2016): Einführung in den Bildungsplan 2016. In: Ministerium für Kultus, Jugend und Sport Baden-Württemberg (Hrsg.): Bildungsplan 2016. Lehrerbegleitheft, Villingen-Schwenningen: Necker-Verlag, S. 4–15.
- Reinhardt, Volker/Rehm, Markus/Wilhelm, Markus (2021): Wirksamer Fachunterricht. Hohengehren: Schneider.
- Ridderbusch, Jens (2022): Schulstrukturreformen im Kontext der demografischen Entwicklung. Zur Ausbreitung integrierter Sekundarschulformen in Schleswig-Holstein und in weiteren Bundesländern. In: Die Deutsche Schule, 114 (2), S. 150–165. doi.org/10.31244/dds.2022.02.03.
- Rizvi, Saman/Waite, Jane/Sentance, Sue (2023): Artificial Intelligence teaching and learning in K-12 from 2019 to 2022: A systematic literature review. In: Computers and Education: Artificial Intelligence, 100145. doi.org/10.1016/j.caeai.2023.100145.

- Rühling, Ute (2023): Die KI schreibt mein Versuchsprotokoll... geht das? ChatGPT im naturwissenschaftlichen Unterricht. In: Unterricht Physik, 2023 (195+196), S. 83–87.
- Thyssen, Christoph (2023): ChatGPT–KI in der Schule: Wie ist die Lage? Orientierungsgedanken für den Biologieunterricht. In: Unterricht Biologie, 2023 (483), S. 9.
- Wacker, Albrecht (2023): Schulstruktur und Zweigliedrigkeit: Umbau des Bildungssystems. In: Bohl, Thorsten/Budde, Jürgen/Rieger-Ladich, Markus (Hrsg.): Umgang mit Heterogenität in Schule und Unterricht. Bad Heilbrunn: Klinkhardt, S. 197–212.

Informationen zu den Autoren



Tobias Bahr promoviert aktuell am Institut für Erziehungswissenschaft der Universität Stuttgart, Abteilung Berufspädagogik mit Schwerpunkt Technikdidaktik im Bereich der Informatikdidaktik zum Profilmfach Informatik, Mathematik, Physik (IMP). Forschungsschwerpunkte: Kompetenzforschung im Bereich Informatik, Genderforschung im Bereich der MINT-Bildung.

Tobias.bahr@ife.uni-stuttgart.de



Mario Manzocco studiert Lehramt Sekundarstufe I (BA) an der Pädagogischen Hochschule Ludwigsburg mit den Fächern Informatik und Englisch.

mario.manzocco@stud.ph-ludwigsburg.de



Dennis Schuster studiert Lehramt Sekundarstufe I (BA) an der Pädagogischen Hochschule Ludwigsburg mit den Fächern Informatik und Geschichte.

dennis.schuster@stud.ph-ludwigsburg.de



Dr. Albrecht Wacker ist Professor für Schulpädagogik der Sekundarstufe I an der Pädagogischen Hochschule Ludwigsburg. Forschungsschwerpunkte: Unterrichts- und Schulentwicklung, Fragen der Schulstruktur und Problematiken gegliederter Schulsysteme in der Sekundarstufe.

albrecht.wacker@ph-ludwigsburg.de

Zitationshinweis:

Bahr, Tobias/Manzocco, Mario/Schuster, Dennis/Wacker, Albrecht (2024): Textgenerierende KI zur Erleichterung der Arbeit von Lehrpersonen? Ein Vorschlag am Beispiel der Aufgabendifferenzierung mithilfe von ChatGPT. In: *Ludwigsburger Beiträge zur Medienpädagogik – LBzM*, 24, S. 1–16. doi.org/10.21240/lbzm/24/07.

Anhang

Prompt für Schritt 1:

Du schlüpfst in die Rolle eines erfahrenen Lehrers für die Sekundarstufe 1 (Klassenstufe 5 bis 10). Deine Hauptaufgabe ist es, Schulaufgaben basierend auf spezifischen Metriken zu analysieren und zu bewerten. Bitte gib nur die Bewertungen gemäß den vorgegebenen Metriken zurück, ohne zusätzliche Begründungen oder Kommentare. Hier sind die Metriken:

- **Bloom'sche Taxonomie**:

1. **Erinnern** (Wiedergabe von Fakten, Grundbegriffen und einfachen Konzepten ohne Verständnis)
2. **Verstehen** (Erklärung von Ideen oder Konzepten, Interpretation von Material)
3. **Anwenden** (Verwendung von Informationen in neuen Situationen, Anwendung von Konzepten und Theorien)
4. **Analysieren** (Unterscheidung zwischen verschiedenen Teilen, Erkennung von Mustern)
5. **Bewerten** (Begründung von Entscheidungen, Auswahl unter verschiedenen Alternativen)
6. **Schaffen** (Erstellung von neuem Material oder Ideen, Kombination von Konzepten)

- **Sprachliche Komplexität**:

1. **Sehr einfaches Sprachniveau** (Kurze Sätze, einfacher Wortschatz, keine Fachbegriffe)
2. **Einfaches Sprachniveau** (Einfache Satzstrukturen, gelegentliche Verwendung von Fachbegriffen mit Erklärungen)
3. **Mittleres Sprachniveau** (Komplexere Satzstrukturen, häufigere Verwendung von Fachbegriffen, aber immer noch verständlich)
4. **Fortgeschrittenes Sprachniveau** (Lange Sätze, fortgeschrittener Wortschatz, Fachbegriffe ohne Erklärungen)
5. **Sehr fortgeschrittenes Sprachniveau** (Komplexe Satzstrukturen, reicher Wortschatz, viele Fachbegriffe und möglicherweise abstrakte Konzepte)

- **Beispiel**: Gibt es ein zusätzliches, klar gekennzeichnetes Beispiel? (true/false)

- **Hinweis**: Gibt es zusätzliche Hinweise oder Tipps? (true/false)

Bitte analysiere die folgende Schulaufgabe und gib deine Bewertung im vorgegebenen Format zurück:

Beginn der Analyse

[BLOOM]: X

[SPRACHE]: X

[BEISPIEL]: true/false

[HINWEIS]: true/false

Ende der Analyse.

Wenn du das verstanden hast antworte mit: "Bitte geben Sie mir eine Aufgabe die ich analysieren soll".