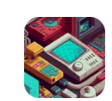




# Erfahrungen mit ChatGPT in Lehre und Prüfungen

Prof. Dr. Dominik Herrmann  
@herdom · herdom.net



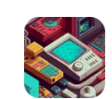
# The Homework Apocalypse

Fall is going to be very different this year. Educators need to be ready.



ETHAN MOLLICK

1 JUL 2023



# The Homework Apocalypse

Fall is going to be very different this year. Educators need to be ready.



ETHAN MOLLICK

1 JUL 2023

Provide up-to-date academic paper references for studies about the question whether phishing trainings are effective



Searching for: **phishing trainings effectiveness academic papers**



Generating answers for you...

Sure, I can help you with that. Here are some up-to-date academic paper references for studies about the question whether phishing trainings are effective:

- **Don't click: towards an effective anti-phishing training. A comparative literature review** by Daniel Jampen, Gürkan Gür, Thomas Sutter & Bernhard Tellenbach (2020). This paper surveys and categorizes works that consider different elements of anti-phishing training programs and identifies key findings in the technical literature. It also describes how a well-founded anti-phishing training program should be designed and parameterized with a set of proposed research directions 1



# The Homework Apocalypse

Fall is going to be very different this year. Educators need to be ready.



ETHAN MOLLICK

1 JUL 2023



Login



Show changes

rizes



This paper reviews and categorizes works that consider different elements of anti-phishing training programs and identifies key findings in the technical literature. It also describes how a sound anti-phishing training program should be designed and parameterized, with a number of suggested research directions.



Provide up-to-date academic paper references for studies about the question whether phishing trainings are effective



Searching for: **phishing trainings effectiveness academic papers**



Generating answers for you...

Sure, I can help you with that. Here are some up-to-date academic paper references for studies about the question whether phishing trainings are effective:

- **Don't click: towards an effective anti-phishing training. A comparative literature review** by Daniel Jampen, Gürkan Gür, Thomas Sutter & Bernhard Tellenbach (2020). This paper surveys and categorizes works that consider different elements of anti-phishing training programs and identifies key findings in the technical literature. It also describes how a well-founded anti-phishing training program should be designed and parameterized with a set of proposed research directions 1

Additionally, and most importantly: THERE IS NO WAY TO DETECT THE OUTPUT OF GPT-4. A couple rounds of prompting remove the ability of any detection system to identify AI writing. And, even worse, detectors have high false positive rates, accusing people (and especially non-native English speakers) of using AI when they are not. You

Additionally, and most importantly: THERE IS NO WAY TO DETECT THE OUTPUT OF GPT-4. A couple rounds of prompting remove the ability of any detection system to identify AI writing. And, even worse, detectors have high false positive rates, accusing people (and especially non-native English speakers) of using AI when they are not. You

Betroffen: **Arbeiten ohne Aufsicht** = „kritische und vertiefte Auseinandersetzung“

Additionally, and most importantly: THERE IS NO WAY TO DETECT THE OUTPUT OF GPT-4. A couple rounds of prompting remove the ability of any detection system to identify AI writing. And, even worse, detectors have high false positive rates, accusing people (and especially non-native English speakers) of using AI when they are not. You

Betroffen: **Arbeiten ohne Aufsicht** = „kritische und vertiefte Auseinandersetzung“

**Texte zu Hause lesen und vorbereiten ...**

... geht schneller, wenn sie ChatGPT zusammenfasst und kritisch würdigt.

— Also **Lesen unter Aufsicht?**

Additionally, and most importantly: THERE IS NO WAY TO DETECT THE OUTPUT OF GPT-4. A couple rounds of prompting remove the ability of any detection system to identify AI writing. And, even worse, detectors have high false positive rates, accusing people (and especially non-native English speakers) of using AI when they are not. You

Betroffen: **Arbeiten ohne Aufsicht** = „kritische und vertiefte Auseinandersetzung“

**Texte zu Hause lesen und vorbereiten ...**

... geht schneller, wenn sie ChatGPT zusammenfasst und kritisch würdigt.

— Also **Lesen unter Aufsicht?**

**Essays/Seminare/Abschlussarbeiten**

Ideen, Struktur, Beispiele, Quellen und Texte lassen sich von ChatGPT erzeugen

— Also **Schreiben unter Aufsicht?**



Additionally, and most importantly: THERE IS NO WAY TO DETECT THE OUTPUT OF GPT-4. A couple rounds of prompting remove the ability of any detection system to identify AI writing. And, even worse, detectors have high false positive rates, accusing people (and especially non-native English speakers) of using AI when they are not. You

Betroffen: **Arbeiten ohne Aufsicht** = „kritische und vertiefte Auseinandersetzung“

**Texte zu Hause lesen und vorbereiten ...**

... geht schneller, wenn sie ChatGPT zusammenfasst und kritisch würdigt.

— Also **Lesen unter Aufsicht?**

**Essays/Seminare/Abschlussarbeiten**

Ideen, Struktur, Beispiele, Quellen und Texte lassen sich von ChatGPT erzeugen

— Also **Schreiben unter Aufsicht?**

**Übungsaufgaben** lassen sich lösen — also unter Aufsicht? **Flipped Classroom.**

# Rules for Tools

Version 2.1 (March 15<sup>th</sup> 2023)

(German version on page 2)

## 1 English Version

1. **All media and tools are allowed.** You may use any media (texts, videos, ...) and tools (apps, calculators, ...) in my course that you find useful. This also applies to AI tools such as ChatGPT, which can be very helpful in generating ideas and writing texts, for example. So these tools are available to you in my course just as they are now in your everyday life and later in your job. By the way, the use of tools will also be subject of our course. Let's find out together how to use tools in a meaningful way to solve tasks!
2. **You are responsible for your results.** All tools have their limitations. Information in media can be wrong. Calculators cannot work with real numbers. And AI language models like ChatGPT can produce well formulated texts, but they make errors and reproduce biases. So before you proceed with results and impulses, you need to check them and revise them if necessary. The tool does not think *for you*, but *you think with the help of the tool*. The theoretical background is the approach of *distributed cognitions*: „Cognitions become ‘distributed’ in the sense that the tool and its human partner think jointly. Whatever is produced is product of the joint system, resulting from the pooling together of the intelligences of both partners [...]“ (Salomon, 1993, p. 182). In the end, however, you stand up for your solution. You have to be able to explain your solution to others. And you (not the tool) are responsible for errors in the solution.
3. **Indicate aids used.** You must list all media you have used as sources. Likewise, you must list all tools used, except trivial tools such as pens or word processors. If applicable, also add how the tools were used (for example, writing down the prompt in the case of AI tools).
4. **Special rules apply to social media.** If you use social media (Instagram, TikTok, Mastodon, Twitter, ...) as part of your course and, for example, share your work

Neue Erklärung zur  
Eigenständigkeit nötig?

Lizenzen

Datenschutz

# Bisherige Nutzung in der Lehre und für Prüfungen

# Bisherige Nutzung in der Lehre und für Prüfungen

- ▶ **Freundliche Antworten** formulieren

## Bisherige Nutzung in der Lehre und für Prüfungen

- ▶ **Freundliche Antworten** formulieren
- ▶ Ausführliche Erläuterungen (**Forum**)

## Bisherige Nutzung in der Lehre und für Prüfungen

- ▶ **Freundliche Antworten** formulieren
- ▶ Ausführliche Erläuterungen (**Forum**)
- ▶ **Grob- und Feinplanung** von Inhalten

## Bisherige Nutzung in der Lehre und für Prüfungen

- ▶ **Freundliche Antworten** formulieren
- ▶ Ausführliche Erläuterungen (**Forum**)
- ▶ **Grob- und Feinplanung** von Inhalten
- ▶ Neue Multiple-Choice-Case-Studies mit **verschiedenen richtigen und falschen Antworten**

# Bisherige Nutzung in der Lehre und für Prüfungen

- ▶ **Freundliche Antworten** formulieren
- ▶ Ausführliche Erläuterungen (**Forum**)
- ▶ **Grob- und Feinplanung** von Inhalten
- ▶ Neue Multiple-Choice-Case-Studies mit **verschiedenen richtigen und falschen Antworten**
- ▶ Welche **Missverständnisse** könnten Studierende nach dem Lesen des folgenden Textes haben, welche inhaltlichen **Lücken** gibt es?



# Bisherige Nutzung in der Lehre und für Prüfungen

- ▶ **Freundliche Antworten** formulieren
- ▶ Ausführliche Erläuterungen (**Forum**)
- ▶ **Grob- und Feinplanung** von Inhalten
- ▶ Neue Multiple-Choice-Case-Studies mit **verschiedenen richtigen und falschen Antworten**
- ▶ Welche **Missverständnisse** könnten Studierende nach dem Lesen des folgenden Textes haben, welche inhaltlichen **Lücken** gibt es?
- ▶ **Interaktion im Hörsaal:** spontane Beantwortung von Fragen in der Vorlesung; Aufgabe, genau lesen und Korrektheit überprüfen

## Bisherige Nutzung in der Lehre und für Prüfungen

- ▶ **Freundliche Antworten** formulieren
- ▶ Ausführliche Erläuterungen (**Forum**)
- ▶ **Grob- und Feinplanung** von Inhalten
- ▶ Neue Multiple-Choice-Case-Studies mit **verschiedenen richtigen und falschen Antworten**
- ▶ Welche **Missverständnisse** könnten Studierende nach dem Lesen des folgenden Textes haben, welche Inhaltlichen **Lücken** gibt es?
- ▶ **Interaktion im Hörsaal:** spontane Beantwortung von Fragen in der Vorlesung; Aufgabe, genau lesen und Korrektheit überprüfen
- ▶ **Seminar:** Erzeugung bewusst fehlerbehafteter Essays: alle bekommen ein eigenes GPT-generiertes Essay zu ihrem Thema und müssen im Rahmen der Seminararbeit alle inhaltlichen und argumentativen Fehler darlegen; zusätzlich: Mini-Essay unter Aufsicht

## Bisherige Nutzung in der Lehre und für Prüfungen

- ▶ **Freundliche Antworten** formulieren
- ▶ Ausführliche Erläuterungen (**Forum**)
- ▶ **Grob- und Feinplanung** von Inhalten
- ▶ Neue Multiple-Choice-Case-Studies mit **verschiedenen richtigen und falschen Antworten**
- ▶ Welche **Missverständnisse** könnten Studierende nach dem Lesen des folgenden Textes haben, welche Inhaltlichen **Lücken** gibt es?
- ▶ **Interaktion im Hörsaal:** spontane Beantwortung von Fragen in der Vorlesung; Aufgabe, genau lesen und Korrektheit überprüfen
- ▶ **Seminar:** Erzeugung bewusst fehlerbehafteter Essays: alle bekommen ein eigenes GPT-generiertes Essay zu ihrem Thema und müssen im Rahmen der Seminararbeit alle inhaltlichen und argumentativen Fehler darlegen; zusätzlich: Mini-Essay unter Aufsicht
- ▶ **Denkbare Falschantworten** für Übungs- und Klausuraufgaben

## Bisherige Nutzung in der Lehre und für Prüfungen

- ▶ **Freundliche Antworten** formulieren
- ▶ Ausführliche Erläuterungen (**Forum**)
- ▶ **Grob- und Feinplanung** von Inhalten
- ▶ Neue Multiple-Choice-Case-Studies mit **verschiedenen richtigen und falschen Antworten**
- ▶ Welche **Missverständnisse** könnten Studierende nach dem Lesen des folgenden Textes haben, welche Inhaltlichen **Lücken** gibt es?
- ▶ **Interaktion im Hörsaal:** spontane Beantwortung von Fragen in der Vorlesung; Aufgabe, genau lesen und Korrektheit überprüfen
- ▶ **Seminar:** Erzeugung bewusst fehlerbehafteter Essays: alle bekommen ein eigenes GPT-generiertes Essay zu ihrem Thema und müssen im Rahmen der Seminararbeit alle inhaltlichen und argumentativen Fehler darlegen; zusätzlich: Mini-Essay unter Aufsicht
- ▶ **Denkbare Falschantworten** für Übungs- und Klausuraufgaben
- ▶ **Bewertungsschemata** bzw. Rubriks für Klausuraufgaben und Texte

## Bisherige Nutzung in der Lehre und für Prüfungen

- ▶ **Freundliche Antworten** formulieren
- ▶ Ausführliche Erläuterungen (**Forum**)
- ▶ **Grob- und Feinplanung** von Inhalten
- ▶ Neue Multiple-Choice-Case-Studies mit **verschiedenen richtigen und falschen Antworten**
- ▶ Welche **Missverständnisse** könnten Studierende nach dem Lesen des folgenden Textes haben, welche Inhaltlichen **Lücken** gibt es?
- ▶ **Interaktion im Hörsaal:** spontane Beantwortung von Fragen in der Vorlesung; Aufgabe, genau lesen und Korrektheit überprüfen
- ▶ **Seminar:** Erzeugung bewusst fehlerbehafteter Essays: alle bekommen ein eigenes GPT-generiertes Essay zu ihrem Thema und müssen im Rahmen der Seminararbeit alle inhaltlichen und argumentativen Fehler darlegen; zusätzlich: Mini-Essay unter Aufsicht
- ▶ **Denkbare Falschantworten** für Übungs- und Klausuraufgaben
- ▶ **Bewertungsschemata** bzw. Rubriks für Klausuraufgaben und Texte
- ▶ Brainstorming und Ausarbeitung neuer **Klausuraufgaben(typen)**

# Erzeugung neuer Klausuraufgaben

## Task 5 Automated Decisions (10 P.)

b) Consider a hate speech detection system that deletes hate speech on social media platforms.

The system assigns a “hate speech score” to each post based on the likelihood of it being hate speech. Moreover, every hour the system is supposed to delete the 10 posts with the highest hate speech scores. However, the system is also supposed to act fairly, i. e., it should not discriminate against minorities by handling their posts differently. Here, the sensitive attribute

is the author’s country of origin. The table below shows the hate speech score assigned by the detection system for the 20 posts that were posted to the social media site in a particular hour. The distribution of authors from Country X and Country O in the table is representative of the demographics of the whole population.

\_\_\_ / 2  
□  
*I don't know.*

Determine which posts will be deleted by the system when it is programmed to aim for “Fairness via Unawareness”. Mark the deleted posts with an “X” in the column “FU”.

Rank	Post ID	Author	Country	Score	FU	DP
1	P1	Alice F.	X	95		
2	P2	Bob G.	X	94		
3	P3	Carol H.	O	93		



# Erfahrungen mit ChatGPT in Lehre und Prüfungen

Prof. Dr. Dominik Herrmann  
@herdom · herdom.net