



# KI und Prüfungen

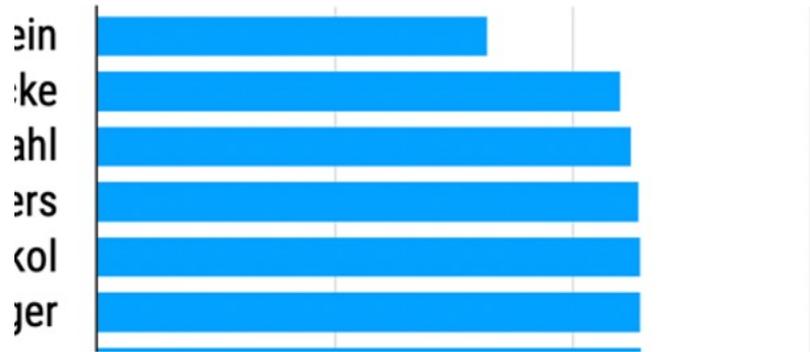
Wie reagieren wir auf „KI als Werkzeug“?  
*nicht: wie sind KI-Kompetenzen zu prüfen*

Prof. Dr. **Dominik Herrmann**

@herdom · herdom.net

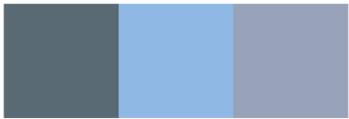
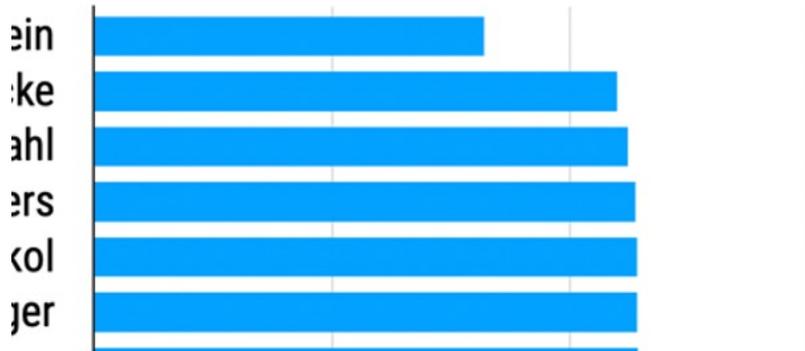
LSt Privatsphäre und Sicherheit in Informationssystemen (PSI) an der Universität Bamberg  
und Co-Sprecher des Projekts DiKuLe

■ Variabilität der Tonlage



*Rückblick auf die Eröffnung  
des Symposiums gestern...*

## Variabilität der Tonlage



You

AttributeError: module 'PIL.Image' has no attribute 'Draw'

the script doesn't really capture the dominant colors per video. something seems wrong. also, in many cases it returns white and black as dominant shads. also, parallelize processing (8 cores)



ChatGPT

To address the issues you're facing, including the incorrect capturing of dominant colors, the presence of white and black as dominant shades, and to parallelize the processing to use multiple cores, we'll make several adjustments to the script.

### Correcting Color Extraction and Filtering White and Black

1. **Filtering White and Black:** We'll enforce stricter filtering for black and white colors in the `find_dominant_colors` function.
2. **Fixing `PIL.Image` Drawing Issue:** The correct attribute for drawing in PIL is `ImageDraw.Draw`.

### Parallelizing Processing

We'll use the `concurrent.futures` module to parallelize the processing of videos, which can significantly speed up the operation on systems with multiple cores.

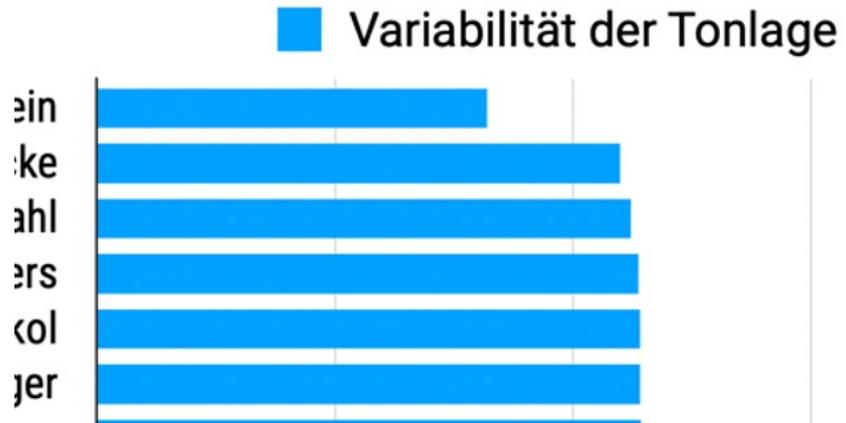
Here's the revised script with these improvements:

```
python
```

Copy code

```
import cv2
import numpy as np
from sklearn.cluster import KMeans
import os
from collections import Counter
```

Prompts hier: <https://chat.openai.com/share/4165bbec-b3ee-4c80-9132-55a8f52a11a1>  
und <https://chat.openai.com/share/85b533e6-0b92-495e-878d-f2b24dc9a9e8>



Oft geäußerte Idee für Umgang mit der  
Behandlung von Generativer KI (GenAI)

**„Process over Product!“**

Forderung nach kleinteiligem  
Monitoring, z.B.

- wöchentliche Zwischenstände
- github-commit-History

***Wollen wir das wirklich?***



15.03.2024, 19:08 Uhr

 Bildbeitrag

 > Mit KI-Hilfe an die Uni: Zwei weitere Bewerber abgelehnt

## Mit KI-Hilfe an die Uni: Zwei weitere Bewerber abgelehnt

Erneut sind an der Technischen Universität in München Bewerber abgelehnt worden. Für ihre Bewerbung sollen sie Künstliche Intelligenz benutzt haben. Ein Bewerber hat deshalb bereits vor Gericht geklagt - ohne Erfolg. Nun folgen weitere Verfahren.

Von  Julia Haderecker  Julia Binder

An der Technischen Universität München (TUM) herrscht Ärger wegen mehrerer Bewerbungssays. Wie ein Sprecher auf BR-Nachfrage bestätigte, wurden zwei weitere Bewerber abgelehnt. Der Verdacht: Sie sollen ihre Aufsätze mithilfe Künstlicher Intelligenz (KI) verfasst haben.

Bereits vor einigen Tagen kursierte der Fall eines Studenten, der sich für einen Masterstudiengang an der TUM beworben hatte und abgelehnt wurde. Die Prüfer waren der Meinung, dass sein Essay mit einer künstlichen Intelligenz wie ChatGPT erstellt worden war. Der Bewerber klagte gegen diesen Entschluss am Bayerischen Verwaltungsgericht. Sein Eilantrag wurde allerdings abgelehnt.

# Dirk Heckmann und Sarah Rachut

## *Rechtssichere Hochschulprüfungen mit und trotz generativer KI*

### I. Einleitung

#### 1. Generative KI „überrollt“ universitäres Lernen, Lehren und Prüfen

- a) Perspektive der Studierenden
- b) Perspektive der Hochschule
- c) Perspektive der Lehrenden
- d) Zwischenfazit

#### 2. Einordnung generativer KI im Prüfungswesen

- a) Anknüpfungspunkt „zugelassene Hilfsmittel“
- b) Anknüpfungspunkt „eigenständige Prüfungsleistung“
- c) Fazit

#### 3. Rechtliche Herausforderungen

- a) *de lege lata*
- b) *de lege ferenda*

### II. Rechtliche Eckpunkte

#### 1. Relevante rechtliche Anforderungen

- a) Grundrechte (Art. 12 Abs. 1, Art. 3 Abs. 1, Art. 5 Abs. 3 GG)
- b) Rechtsstaatliche Anforderungen, insb. Gesetzesvorbehalt
- c) Einfachgesetzliches Prüfungsrecht

a) Einsatz von KI-Detektoren zur Täuschungskontrolle und Aufdeckung des KI-Einsatzes durch Studierende

b) Exkurs: Einsatz von KI zur Klausuraufsicht

c) Exkurs: Einsatz von KI zum Lehr-/Lern-Monitoring

d) Exkurs: Einsatz generativer KI in der Hochschulverwaltung

#### 4. Übersicht der Anwendungsszenarien III – Perspektive der Lehrenden

a) Einsatz von generativer KI zur Unterstützung der Lehre

b) Einsatz von generativer KI zur Erstellung von Prüfungsaufgaben

c) Einsatz von generativer KI als Bestandteil innovativer Lehr- und Prüfungsformate

### IV. Leitgedanken zum rechtssicheren Einsatz generativer KI im Prüfungskontext

1. Disclaimer („work in progress“)

2. Regulierungsbedarf

3. Verhältnismäßiger Ausgleich von Lehr- und Lernfreiheit mit Chancengerechtigkeit

### V. Fazit

The significance of EER lies in its provision of a singular measure that encapsulates the accuracy of a biometric system by pinpointing the threshold at which the system's sensitivity and specificity are optimally balanced. A lower EER value is indicative of a superior system performance, as it suggests a lower rate of both false acceptances and false rejections (Mansfield & Wayman, 2002). This metric is pivotal for comparing the efficacy of different biometric systems or configurations, especially in the absence of a predefined operational point or when an application demands a balance between the acceptance of genuine users and the rejection of impostors (Jain et al., 2016).

The EER is therefore a critical parameter in the domain of biometric system evaluation, offering a straightforward and comprehensive benchmark for system comparison and threshold optimization decisions. Its widespread adoption in academic and practical assessments underscores its utility in conveying the effectiveness of biometric authentication mechanisms in a manner that is both quantifiable and universally understandable.

⋮ WAITING FOR  
⋮ YOUR INPUT



I AGREE TO THE TERMS OF SERVICE  
(NO ACADEMIC MISCONDUCT)

CHECK FOR AI

HUMANIZE

Your content appears human

#### AI DETECTION LIKELIHOOD

GPTZERO OPENAI WRITER CROSSPLAG COPYLEAKS SAPLING CONTENTATSCALE ZEROGPT

KEY: 100% HUMAN 50% HUMAN 0% HUMAN

Sign up to humanize and make your text undetectable.

SIGN UP FOR FREE

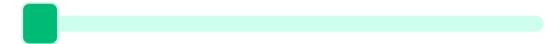
Yearly

Limited Time - 50% Off

\$5.00/month

10,000

300,000



10,000 words / month



Get Started

Cancel Anytime

6 Months Free

Charged annually \$60.00

#### Quick look at features:

- ✓ Passes AI detectors
- ✓ High quality, legible content
- ✓ Watermark and future proof
- ✓ Writing level matching
- ✓ API compatible
- ✓ Unlimited Human Auto Typer

## **Das sind Proxy-Probleme.**

Bequemer als das Angehen der Ursachen und ziehen Aufmerksamkeit auf sich.

*vgl. Bike Shedding  
(Parkinson's Gesetz der Trivialität)*

## Nutzung von GenAI durch Lehrpersonen für Prüfungen

angenehme Arbeit  
*Prokrastinationspotential!*

*Wie könnten non-native Speaker diese Aufgabe missverstehen?*

*Erzeuge drei richtige und drei falsche Antworten für diese Frage, gegeben diese behandelten Inhalte [...].*

*Beschreibe ein Szenario, in dem das Konzept [...] eine Rolle spielt.  
Baue irrelevante Details ein.*

*Ist die Antwort [...] besser oder schlechter als diese [...]?*

## Reaktion auf GenAI durch angepasste Prüfungsformate

anspruchsvolle Arbeit

... für Prüfende und Prüflinge.

### **Methodenkompetenz**

(Lesen/Denken/Schreiben)  
in Präsenz üben und in langen  
Aufsichtsarbeiten (2-3h) prüfen

Neue/alte Ideen für Prüfungen unter  
Aufsicht (angloamerik. Modell):  
*Midterm Exams + unangekündigtes  
Vorrechnen von Übungsaufgaben.*

# (E-)Portfolio einfach komplett unter Aufsicht!

## Regelungsbedarf.

### § 55 Portfolioprüfung

Die Portfolioprüfung setzt sich aus mehreren semester- begleitenden Prüfungselementen in der Regel unterschiedlicher Form zusammen ...

Es kommen hierbei ... die schriftliche Ausarbeitung, der **Multiple-Choice-Test**, der Vortrag, die protokollierte praktische Leistung, der Entwurf, die künstlerische Arbeit, die Rücksprache oder das Poster in Betracht.

Ein **schriftlicher Test** darf die Dauer von 60 Minuten, eine mündliche Rücksprache die Dauer von 20 Minuten nicht überschreiten.

Im Rahmen der Portfolioprüfung dürfen maximal drei schriftliche Tests verlangt werden, deren Bearbeitungszeit in Summe 120 Minuten nicht überschreiten darf; mehrere mündliche Rücksprachen dürfen die Gesamtdauer von 40 Minuten nicht überschreiten.

## Ideen (Vorlesung)

8 unüberwachte Essays mit Peer Feedback; mind. 6 müssen bearbeitet werden, um einen überwachten Essay schreiben zu dürfen (3h), der mit angekündigtem Rubrik bewertet wird.

E-Prüfungen für unbekannte / praktische Problemlösungsaufgaben auf präparierten Rechnern mit vorinstallierten Hilfsmitteln

Idee von Thomas Basbøll: Studierende schreiben vormittags einen überwachten Essay unter Aufsicht zu einem vorher nicht bekannten Thema, nachmittags wird der Essay in einer mündlichen Prüfung besprochen

<https://inframethodology.cbs.dk>

## Idee (Seminar)

Studierende lernen Schreiben in **3 Writing Labs** (mit Präsenz-Übungen): *outlining, paragraphs, coherence, cohesion, conciseness, clarity*

Studierende erhalten **einen von mir mit GenAI präparierten und angepassten** Essay mit Fehlern, falschen Zitaten, ...

**Seminararbeit (unüberwacht):** eine max. fünfseitige Abhandlung, die alles richtig stellt, kritisiert bzw. vervollständigt.



**Ziel:  
Akrasia / Prokrastination  
vermeiden**

**Wir müssen die richtigen  
Anreize setzen!**



# KI und Prüfungen

Prof. Dr. **Dominik Herrmann**

@herdom · herdom.net

LSt Privatsphäre und Sicherheit in Informationssystemen (PSI) an der Universität Bamberg

und Co-Sprecher des Projekts DiKuLe

Folien (später): [dhgo.to/ki-dikule24](https://dhgo.to/ki-dikule24)