

GEFÖRDERT VOM



Bundesministerium
für Bildung
und Forschung

Vernetzungsworkshop KI-Exam

Katharina Hähn, Katja Buntins & Lutz Goertz

Vernetzungsworkshop

Tamed.AI

Christiani
Wir bilden voraus

mmb Institut

UNIVERSITÄT
DUISBURG
ESSEN

Offen im Denken

Herzlich Willkommen zum Vernetzungsworkshop von KI-Exam!

- Inhalt:
 - Gegenseitiges Kennenlernen
 - Vorstellung & Diskussion der Ergebnisse
 - Geplante Umsetzung



Studien Design & Ergebnisse

Vorgehen bei der Prüfungskorrektur und Bedarfe

- 20 Interviews mit Prüfenden und Aufgabenersteller:innen im Februar 2023
- Interviewdauer 20 – 40 Minuten
- Interviews wurden von 3 Personen durchgeführt

Ablauf Workshop

Handlungsroutinen Prüfungskorrektur

Schwierigkeiten Prüfungskorrektur & Besonderheiten

Befürchtungen und Potenziale

Wünsche an konkrete Umsetzung

1. Ergebnisdarstellung
2. Ergänzungen und Diskussion mit Prüfenden
3. Rückfragen und Diskussion für die konkrete Umsetzung

Themenblöcke:

- Vorbereitung/Nachbereitung Prüfungsprozesse
- Aufgabe bewerten/Korrekturverhalten
- Erfassen Aufgabeneinhalt/Textstrukturierung
- Auswertungsstrategien

Handlungsroutinen Prüfungskorrektur - Vorbereitung/Nachbereitung Prüfungsprozesse

- Auseinandersetzen mit den Aufgaben, indem man den Erwartungsraum definiert:
 - Orientierung an Fragestellung und Punkten
 - Eigene Lösung
 - Lösungshinweise oft erst als weiterer Schritt
 - Lösungshinweise oft nicht ausreichend
 - sind aber z. T. trotzdem Ratgeber

Orientierung an Fragestellung und Punkten:

Das Wording der Aufgabenstellung erfordert unterschiedliche Antworten, z. B. Benennung versus Beschreiben oder Erläuterung. Hierdurch ergibt sich eine sehr unterschiedliche Länge der Texte.

Schaue mir die Lösungshinweise an, bei hochgewichteten Aufgaben 8 Punkte und mehr lege ich mir ein Schema zurecht, wie ich die Punkte vergabe. Beispiel der letzten Korrektur Aufgabe zur SPS-Schaltung. Niemand wird 12 Punkte haben. 6 Kategorien, max. 2 Punkte pro Kategorie.

Handlungsroutinen Prüfungskorrektur - Vorbereitung/Nachbereitung Prüfungsprozesse

Eigene Lösung

IP löst alle Aufgaben erst einmal selbst, da die IHK-Lösungshinweise nicht alle Antwortmöglichkeiten enthalten, überlegt sie sich, welche Punkte sie dort aufgeführt haben möchte.

Selbst Lösung für Aufgabe finden und diese als Grundlage nehmen. Diese wird dann je nach Antworten der Prüflinge erweitert.

Als erstes versuche ich die Aufgabe selbst zu lösen, ohne Lösungshinweis, Musterlösungshinweis ansehen – ist das plausibel?

Lösungshinweise oft erst als weiterer Schritt

Die Lösungshefte sind nur Hinweise, der Prüfer ist autark, beruft sich auf sein Fachwissen, versuche anhand der Lösungen abzugleichen, ob ich da etwas Sinnvolles erkennen kann.

Da die IHK-Lösungshinweise nicht alle Antwortmöglichkeiten enthalten, überlegt sie sich, welche Punkte sie dort aufgeführt haben, möchte.

Handlungsroutinen Prüfungskorrektur - Vorbereitung/Nachbereitung Prüfungsprozesse

Lösungshinweise oft nicht ausreichend, sind z. T. aber trotzdem ein wichtiger Ratgeber

Lösungshinweise können nur einen ersten Hinweis geben – Unterscheiden sich je nach Disziplin erheblich

Lösungen ganz anders als Lösungsvorschlag. Hier braucht es viel Expertise und viel Zeit dies nachzuvollziehen..

Man hat Hilfsmittel zur Orientierung: sein eigenes Know-how, den Lösungshinweise der Aufgabenersteller – hierüber wird gewisse Erwartungshaltung vorgegeben. In der Regel deckt sich das mit meinen Erwartungen. Falls abweichend, schaue ich nach.

Prüfungsausschuss schaut sich Aufgabenstellungen und Lösungshinweise an

Themenblöcke:

- Vorbereitung/Nachbereitung Prüfungsprozesse
- **Aufgabe bewerten/Korrekturverhalten**
- Erfassen Aufgabeneinhalt/Textstrukturierung
- Auswertungsstrategien

Handlungsroutinen Prüfungskorrektur – **Aufgabe bewerten/Korrekturverhalten**

- Das Bewerten von Aufgaben ist oft ein **kollaborativer Prozess**
- Hierbei findet ein Austausch entweder im Diskurs ODER
- durch Kommentare für Zweitprüfende statt

Kollaboratives Auswerten:

Gemeinsames Betrachten
und besprechen im
Konsortium

Korrekturen sollten immer diskursiv
und kollaborativ stattfinden, so
werden Fehler vermieden

- Pro Durchlauf sind es 13 bis 15 Prüfungen
- Der Ausschuss blockt sich Tage für einen Termin und trifft sich in Präsenz
- Zwei bis drei Prüfer gehen einen Unterrichtsblock durch

Zusammenarbeit mit Zweitprüfenden:

Abstimmung mit
Zweitkorrektor, ich bin
meistens Erstkorrektor,
selten Abstimmung
notwendig. Es gibt auch die
Prüfer-App. Da sieht der
Zweitprüfer die Ergebnisse.

Kommunikation zwischen Erst-
und Zweitprüfer auch im
Korrekturprozess, z. B. mit
Post-its, die in die Bögen
geklebt werden

Handlungsroutinen Prüfungskorrektur – Aufgabe bewerten/Korrekturverhalten

Beim Vorgehen während der Korrektur:

- im Sinne des Prüflings
- Aufgabenweise
- iteratives Vorgehen beim Korrigieren

Neben berufsfachlichen Kenntnissen bei den Prüfern spielt auch der „Faktor Mensch“ eine große Rolle

Frage 1 bei Prüfung 1 dann bei Prüfung 2 etc.
→ bessere Variante, weil Antworten vergleichbarer sind
→ wichtig, um die Antworten gleich zu behandeln/zu bewerten

Handlungsroutinen Prüfungskorrektur

Erfassen Aufgabeninhalt/Textstrukturierung

Themenblöcke:

- Vorbereitung/Nachbereitung Prüfungsprozesse
- Aufgabe bewerten/Korrekturverhalten
- **Erfassen Aufgabeninhalt/Textstrukturierung**
- Auswertungsstrategien

Handlungsroutinen Prüfungskorrektur

Erfassen Aufgabeninhalt/Textstrukturierung

- Zusammenfügen der Textteile oft zeitaufwendiger und erster Schritt
- Suchen nach Fachbegriffen und Kernaussagen
- Erfassen der Struktur
- Nachvollziehen von Lösungswegen/Varianten

Handlungsroutinen Prüfungskorrektur – Erfassen Aufgabeninhalt/Textstrukturierung

- Suchen nach Fachbegriffen und Kernaussagen

Text wird nach Schlüsselwörtern durchsucht und je Schlüsselwort gibt es schonmal einen Punkt. Das wird mit Häkchen markiert. Wenn sich Aussagen doppeln, kommt eine Klammer darum, damit verständlich ist, dass es kein Häkchen gibt, weil das schonmal gesagt wurde.

Themenblöcke:

- Vorbereitung/Nachbereitung Prüfungsprozesse
- Aufgabe bewerten/Korrekturverhalten
- Erfassen Aufgabeneinhalt/Textstrukturierung
- **Auswertungsstrategien**

Handlungsroutinen Prüfungskorrektur

Auswertungsstrategien

- Vergleichsweise wenig Erkenntnisse
- Ergibt sich möglicherweise aus dem Zusammenspiel von *Vorbereitung/Nachbereitung Prüfungsprozesse* und *Erfassen Aufgabeninhalt/Textstrukturierung*

Handlungsroutinen Prüfungskorrektur

Schwierigkeiten Prüfungskorrektur & Besonderheiten

Befürchtungen und Potenziale

Wünsche an konkrete Umsetzung

1. Ergebnisdarstellung
2. Ergänzungen und Diskussion mit Prüfenden
3. Rückfragen und Diskussion für die konkrete Umsetzung

Themenblöcke:

- Bei der Korrektur/dem Verständnis des Textes
- Bei der Korrektur der inhaltlichen Antwort

Schwierigkeiten Prüfungskorrektur & Besonderheiten – Bei der Korrektur/dem Verständnis des Textes

Schwierigkeiten:

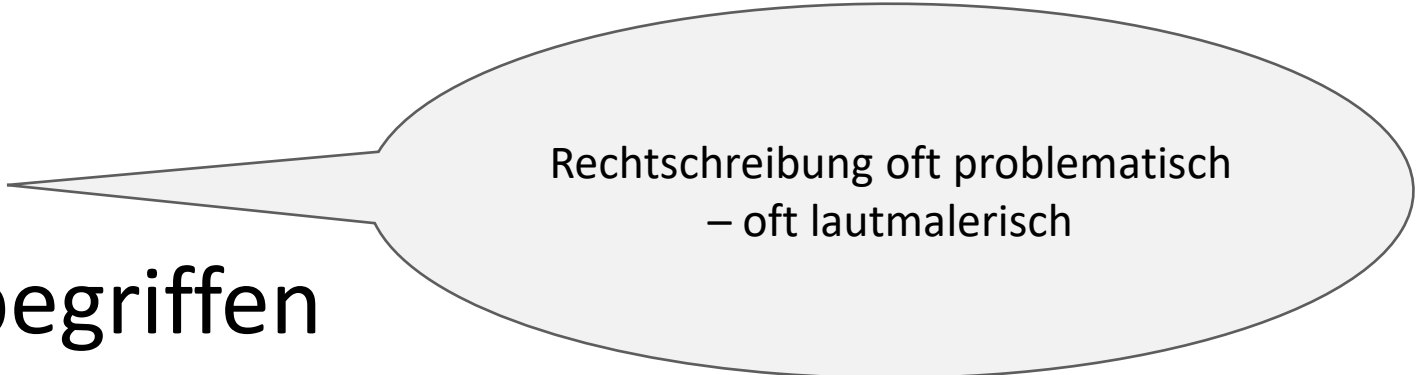
- Lesbarkeit
- Rechtschreibung
- Synonyme von Fachbegriffen
- Identifizieren des Inhaltes
- Grammatik

Lesbarkeit erfassen – Interpretation
Texten in Bezug auf das, was man
lesen kann, Versuchen es aus dem
Kontext zu erschließen

Schwierigkeiten Prüfungskorrektur & Besonderheiten – Bei der Korrektur/dem Verständnis des Textes

Schwierigkeiten:

- Lesbarkeit
- Rechtschreibung
- Synonyme von Fachbegriffen
- Identifizieren des Inhaltes
- Grammatik



Rechtschreibung oft problematisch
– oft lautmalerisch

Schwierigkeiten Prüfungskorrektur & Besonderheiten – Bei der Korrektur/dem Verständnis des Textes

Schwierigkeiten:

- Lesbarkeit
- Rechtschreibung
- Synonyme von Fachbegriffen
- Identifizieren des Inhaltes
- Grammatik

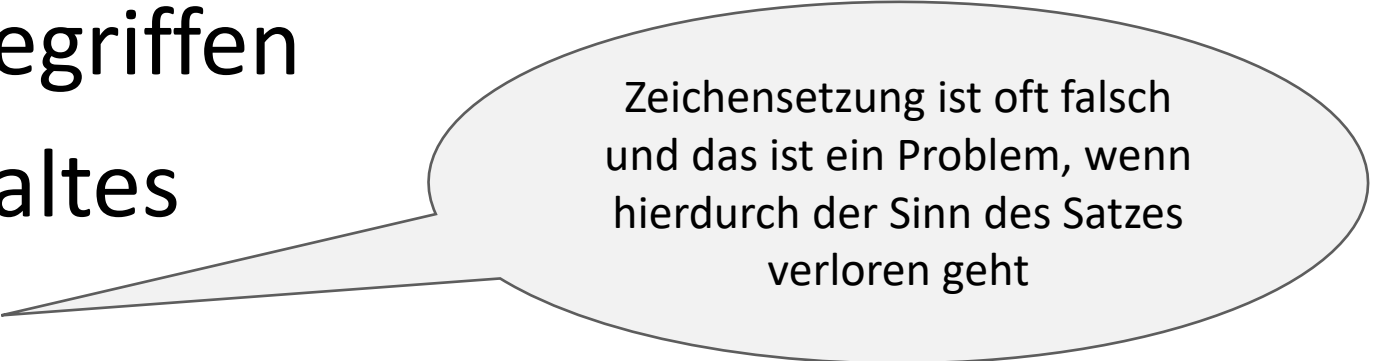
Anstatt berufsfachlicher Begriffe und Normen konkret anzuführen werden häufig Umschreibungen hierfür aufgeschrieben

Fachbegriffe spielen eine Rolle, wenn der Eigenname eines Herstellers anstatt der Bezeichnung eines z. B. Werkzeuges genutzt wird, gilt das als richtige Antwort. (Synonyme sind richtige Antworten)

Schwierigkeiten Prüfungskorrektur & Besonderheiten – Bei der Korrektur/dem Verständnis des Textes

Schwierigkeiten:

- Lesbarkeit
- Rechtschreibung
- Synonyme von Fachbegriffen
- Identifizieren des Inhaltes
- Grammatik



Zeichensetzung ist oft falsch
und das ist ein Problem, wenn
hierdurch der Sinn des Satzes
verloren geht

Themenblöcke:

- Bei der Korrektur/dem Verständnis des Textes
- **Bei der Korrektur der inhaltlichen Antwort**

Schwierigkeiten Prüfungskorrektur & Besonderheiten – Bei der Korrektur der inhaltlichen Antwort

Schwierigkeiten:

- Unpräzise Fragestellungen
- Ausreichendes Fachwissen notwendig
- Falsche Elemente innerhalb der Antwort

Manchmal sind die Fragen nicht klar definiert. Dies ermöglicht viel Spielraum bei der Antwort. Hier ist die Bewertung oft sehr divers zwischen verschiedenen Prüfern

Schwierigkeiten Prüfungskorrektur & Besonderheiten – Bei der Korrektur der inhaltlichen Antwort

Schwierigkeiten:

- Unpräzise Fragestellungen
- Ausreichendes Fachwissen notwendig
- Falsche Elemente innerhalb der Antwort

Antworten sind nicht im Lösungshinweis der IHK vorhanden:
Um Richtigkeit dieser Lösungen zu erkennen, müssen Prüfer vom Fach sein.

Schwierigkeiten Prüfungskorrektur & Besonderheiten – Bei der Korrektur der inhaltlichen Antwort

Schwierigkeiten:

- Unpräzise Fragestellungen
- Ausreichendes Fachwissen notwendig
- Falsche Elemente innerhalb der Antwort

Eine Schwierigkeit ergibt sich, wenn die Texte die richtigen Antworten und zusätzliche falsche Bestandteile enthalten. unterschiedliche Meinungen zur Punktevergabe im Prüfungsausschuss; IP zieht Punkte ab

Schwierigkeiten Prüfungskorrektur & Besonderheiten – Bei der Korrektur der inhaltlichen Antwort

Rechenaufgaben/ Formeln:

- Viele verschiedene Lösungswege
- Einheiten und Formeln werden oft nicht angegeben
- So schwer Rechenfehler zu erkennen

Stichpunkte:

- Oft mehr Nennungen als gefordert
- Nennungen doppeln sich
- Mehr als eine Nennung pro Stichpunkt

Abbildungen und Zeichnungen:

- Beschriftung oft ungenau
- Zeichnungen sehr unterschiedlich
- Eher selten, wenn vor allem Zeichnungen

Handlungsroutinen Prüfungskorrektur

Schwierigkeiten Prüfungskorrektur & Besonderheiten

Befürchtungen und Potenziale

Wünsche an konkrete Umsetzung

1. Ergebnisdarstellung
2. Ergänzungen und Diskussion mit Prüfenden
3. Rückfragen und Diskussion für die konkrete Umsetzung

Themenblöcke:

Befürchtungen

- An KI an sich
- An KI-Exam:
 - Kontrolle verlieren
 - Umgang mit Handschriften
 - Datengrundlage
 - Umgang mit Diversität/seltenen Lösungen

Potenziale

- Zeitgewinn
- Fairness
- Vereinfachung
- Kontrolle

Befürchtungen

- An KI an sich
- An KI-Exam:
 - Kontrolle verlieren
 - Umgang mit Handschriften
 - Datengrundlage
 - Umgang mit Diversität/seltenen Lösungen

Wenn man den Text liest und es fehlt ein Begriff, dann kann man sagen „komm, er hat es aber umschrieben und weiß, worum es geht“. Klar muss man als Fachkraft Fachbegriffe kennen. In Punkten kann ich das nicht ausdrücken, weil es aus dem Zusammenhang kommt. Deshalb weiß ich nicht, wie KI sowas machen kann.

Die soziale Komponente ist wichtig: Wir verteilen auch Sozialchancen. Insbesondere bei Wackelkandidaten. Aber ob das KI kann?

Befürchtungen

- An KI an sich
- An KI-Exam:
 - Kontrolle verlieren
 - Umgang mit Handschriften
 - Datengrundlage
 - Umgang mit Diversität/seltenen Lösungen

Schwierigkeit: Welche Daten nutzt KI? Viele Halbwahrheiten im Netz

Wer füttert die KI? Woher kommen die Antworten?

Befürchtungen

- An KI an sich
- An KI-Exam:
 - Kontrolle verlieren
 - Umgang mit Handschriften
 - Datengrundlage
 - Umgang mit Diversität/seltenen Lösungen

Soziale Komponente fehlt, d. h. Kontext der Aufgabe wird nicht verstanden, wegen der es zu unterschiedlichen Antworten kommt. Das gibt es heute auch schon bei Prüfern, wenn sie den Kontext der Aufgabe nicht verstehen. Dann machen sie eventuell auch eine Fehlbewertung.

Potenziale

- Zeitgewinn
- Fairness
- Vereinfachung
- Kontrolle

Ein bundeweiter Vergleich zu Bestehens-/Durchfallquoten wäre eventuell erleichtert und würde einen solchen Vergleich befördern.

Potenziale

- Zeitgewinn
- Fairness
- Vereinfachung
- Kontrolle

Ablauf Workshop

Handlungsroutinen Prüfungskorrektur

Schwierigkeiten Prüfungskorrektur & Besonderheiten

Befürchtungen und Potenziale

Wünsche an konkrete Umsetzung

1. Ergebnisdarstellung
2. Ergänzungen und Diskussion mit Prüfenden
3. Rückfragen und Diskussion für die konkrete Umsetzung

Themenbereiche:

- Vorbereitung/Nachbereitung Prüfungsprozesse
- Vorbereitung Prüfung für Kontrolle
- Unterstützung Verständnis/Erfassung der Aufgabe
- Aufgabe bewerten

Wünsche an konkrete Umsetzung – Vorbereitung/Nachbereitung Prüfungsprozesse

- Plagiatserkennung
- Erstellen von Lösungshinweise
- Einfließen von Lösungshinweisen
- Dokumentation des Prüfungsprozesses

Wünsche an konkrete Umsetzung – Vorbereitung Prüfung für Kontrolle

- Lesen von Handschriften
- Struktur im Dokument erfassen
- Digitalisierung der Prüfungen

Wünsche an konkrete Umsetzung – Unterstützung Verständnis/Erfassung der Aufgabe

- Kerninhalte identifizieren und markieren
- Text strukturieren
- Verschiedene Lösungswege erkennen/verstehen

Wünsche an konkrete Umsetzung – Aufgabe bewerten

- Korrekturvorschlag
- Vergleich mit Lösungshinweisen
- Bepunktung
- Vorbewertung vornehmen



Bedeutung für die Entwicklung und Projektrealisation

- Mehrwert
- Aufwand
- Risiken
- Besondere Relevanz für KI-Exam
- Relevanz der Prüfenden

Bewertungsschema

1
...
5

Besondere Relevanz
für KI-Exam

ID	Relevanz entsprechender Prüfende (1 = selten genannt – 5 = häufig genannt)	Probleme	Aufgaben für KI-Exam:	Einschätzung Tech-Runde (Christiani, TAL, UDE)
Vorverarbeitung, Vorkorrektur, Sprachkorrektur				
2	4	Rechtschreibfehler,	- Integration automatischer Rechtschreibkorrektur. - System sollte kennzeichnen, dass Rechtschreibung angepasst wurde und beide Versionen zur Verfügung stellen - Fachwörter etc. sollten erkannt werden und ggf. nicht korrigiert werden.	Scheint sinnvoll + einfach umsetzbar
3	3	Grammatikfehler, falsche Interpunktion, fehlende Wörter	- Ausgleich von Grammatikfehlern - Unsicherheiten angeben - Erkennen wann relevante Teile zum Lösen der Aufgabe fehlen - Beide Versionen sollten zur Verfügung stellen	Scheint sinnvoll + umsetzbar; Erkennen... eher schwierig
14	3	Es werden mehr Lösungen angegeben als gefordert	- System sollte Zahl der erlaubten Lösungen kennzeichnen (also wenn bei vier geforderten Lösungen sechs genannt werden, dann die letzten beiden streichen)	- Wäre machbar; - Ggf. eher Aufgabendesign (z.B. einzelne Textfelder)
10	2	Prüfkandidat:innen scheitern am Wording, z.B. „benennen“, „erläutern“, „nennen Sie Beispiele für...“	- System sollte diese Keywords berücksichtigen und damit den Erwartungshorizont bestimmen und danach bereits den Umfang und den Charakter der Antworten beurteilen (enthält es Beispiele? Ist „erläutern“ gefragt und die Antwort besteht nur aus einem Satz?).	Schwierig?! Eher hinten anstellen
17	2	Extreme und originelle (und gleichfalls richtige) Lösungen	- Sicherstellen, dass genügend passende Fachtexte in das System eingegeben werden, die auch ein Verständnis des Kontextes ermöglichen - Längeres Training des Systems, dass solche Lösungen toleriert werden - Im Notfall explizit Unsicherheit angeben, anstatt zu halluzinieren	Manueller Input > Die Lösung hier ist kreativ; Mensch, bitte bewerte das mal; als „unsicher“ kennzeichnen -> nice-to-have: Kreativität hat keine Grenzen...

? Ungewiss
1 Niedriger Mehrwert
2 Mittlerer Mehrwert
3 Hoher Mehrwert

Mehrwert

? Ungewiss
1 Geringer Aufwand
2 Mittlerer Aufwand
3 Hoher Aufwand

Aufwand

! Riskant!

Risiko

3	2	2	2	2	1	1	2	1			
3	3	2	2	2	2	3	2	2	1		
2	1	2	1	1	2	2	1	1	1		
3	2	1	2	1	3	3	3	3	3	3	
3	3	3	2	3	2	3	3	3			

Berechnung Priorisierung

Bewertung von

- ... Mehrwert
- ... Aufwand
- ... Risiko
- ... Relevanz für KI-Exam
- ... Relevanz der Prüfenden

Formel

$$\frac{(\text{Mehrwert} \times (\text{Relevanz der Prüfenden} + \text{Relevanz für KI-Exam})) - (\text{Aufwand} \times \text{Risiko})}{\text{= Prioritäts-Score}}$$

Priorisierung

ID	Score	Feature / Qualitätsanforderung	Art
5	95	Ähnlichkeit zwischen Prüfungsantworten ermitteln	Feature
22	93	Suche nach inhaltlich ähnlichen Klausurantworten	Feature
2	83	Rechtschreibfehler vorkorrigieren	Feature
16	81	Robustheit gegenüber anders formulierten + kreativen Antworten	Qualitätsanforderung
15	48		
17	45		
16	81	Fachbegriffe und Synonyme als Schlagwörter bei Musterlösung ergänzen	Feature
7	69	Visualisierung der KI-Ergebnisse zur Weiterverarbeitung durch Prüfende	Feature
27	45	Abweichung in Bepunktung zwischen Erst- und Zweitgutachten anzeigen	Feature
6	44	Ähnlichkeit zwischen Prüfungsantworten und Musterlösung ermitteln	Feature
21	42	Prüfende warnen, wenn es systematische Abweichungen in Prüfungsantworten gibt	Feature
3	38	Grammatikfehler vorkorrigieren	Feature

Vielen Dank

<https://www.ki-exam.de>