



KI als Schreibcoach statt
Ghostwriter:

TextAid zur Förderung
wissenschaftlicher
Schreibkompetenz

TextAid: KI als Schreibcoach für Studierende



Input: Text +
Features



Verarbeitung
durch LLMs



Feedback-
Kommentare

- StiL-Forschungsprojekt TextAid (2025-2027) an der THM (Fachbereich MNI)
- Weiterentwicklung des bestehenden Feedbacksystems
- Begleitung des Lernprozesses statt Ersetzen eigener Leistung
- TextAid: Intelligentes Tutorensystem, das Studierende beim Verfassen von **wissenschaftlichen Texten** wie Seminar- und Abschlussarbeiten mit **zielgerichtetem, lernförderlichem Feedback** unterstützt.

→ Bei anderen Chat-Bots (wie ChatGPT) liegt der Fokus auf der Bereitstellung von fertigen Lösungen, während bei TextAid der Lernprozess im Vordergrund steht

Ausgangssituation: KI im Studienalltag

- Studie des Centrums für Hochschulentwicklung (CHE) [1]:
- Ca. 2/3 der Studierenden nutzen KI-Tools mindestens wöchentlich
- Ca. 1/4 nutzt KI täglich
- Problem: Abnehmende Fähigkeit zum eigenständigen Schreiben
- Herausforderung: KI sinnvoll in Lernprozesse integrieren

[1] <https://www.che.de/2025/ein-viertel-der-studierenden-nutzt-taeglich-kuenstliche-intelligenz>

Motivation / Ziele des Projektes

- In frühen Schreibphasen treten oft ähnliche, einfach zu erkennende Fehler auf, die zeitaufwendig mit “Standardkommentaren” von Lehrenden korrigiert werden
- Oft benötigen Korrekturen mehrere Iterationsrunden, in der Praxis → nicht umsetzbar
- TextAid unterstützt bei diesen Anfängerfehlern, ggf. auch wiederholt, und entlastet so Lehrende

- Lehrende haben durch TextAid mehr Zeit sich komplexeren Fragestellungen zu widmen
- Gründe warum wissenschaftliches Schreiben herausfordernd ist sind vielfältig und individuell → Einführungskurs in wiss. Arbeiten oft nicht ausreichend
- TextAid setzt am aktuellen Lernstand an und bietet wiederholt ausführliches, passgenaues Feedback
- Ziel ist es, Grundlagen zu festigen und individuelle Herausforderungen adressieren

Mehrwert für Studierende

Sofortiges Feedback, 24/7 verfügbar, kann frühzeitig und beliebig häufig verwendet werden

Rückfragen sind möglich

Erste Unterstützung beim präzisen Formulieren von Forschungsfragen & Überschriften

Vorabüberprüfung und Verbesserungshinweise zur Argumentationsstruktur

Lesbarkeit wird verbessert

- Einfache/Erste Hinweise zu Übergängen, rotem Faden, Nachvollziehbarkeit und Verständlichkeit

Formale Qualität sichern

- Prüfung von Abkürzungen, Rechtschreibung und Zeichensetzung
- Verbesserung von Inhaltsverzeichnis, Literaturverzeichnis (auch spezifische Vorgaben mögl.)



Chat

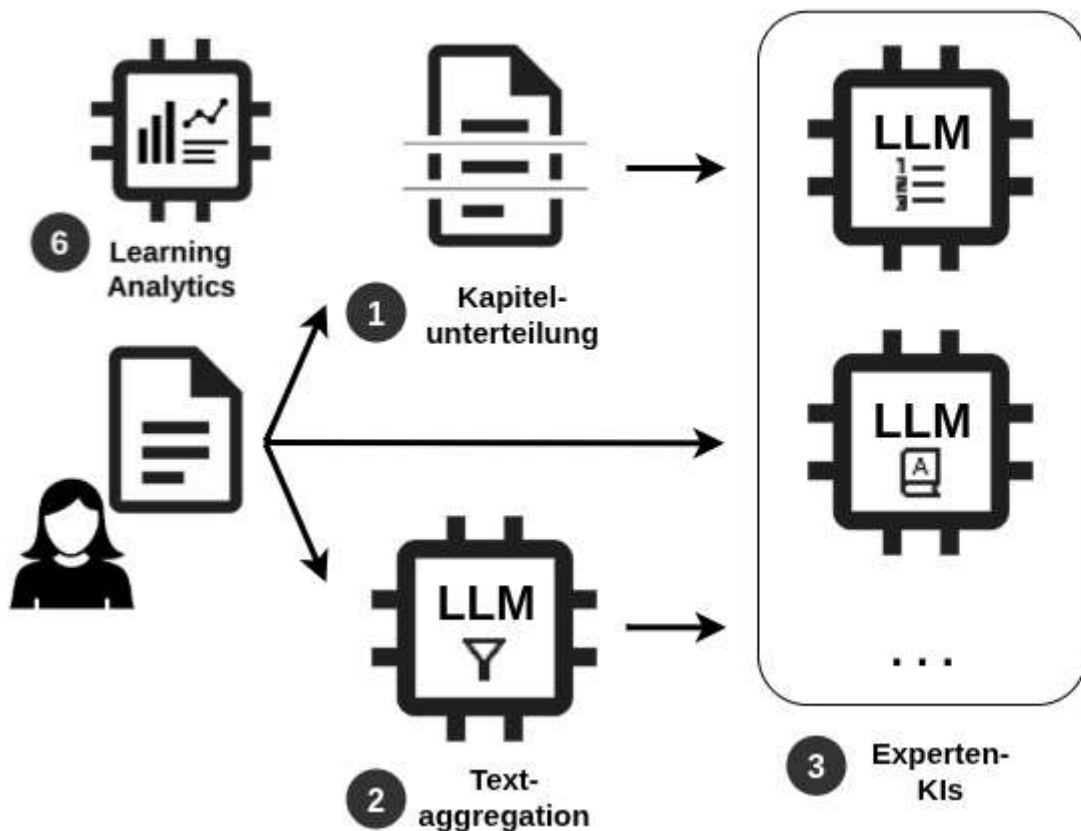
Selected text: Wortwiederholungen: Begriffe wie „Prozessketten“ und „Produktentwicklung“ werden häufig wiederholt. Das kann den Text trocken wirken lassen.

Ask something...

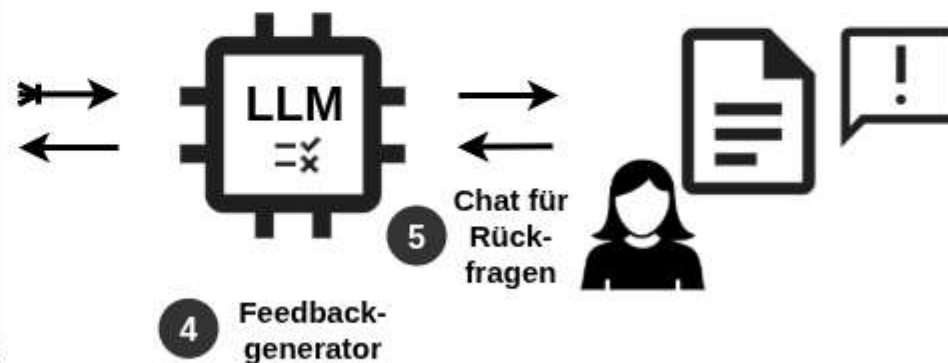
Send

user:
welche begriffe würden sich alternativ eignen?

Technische Funktionsweise

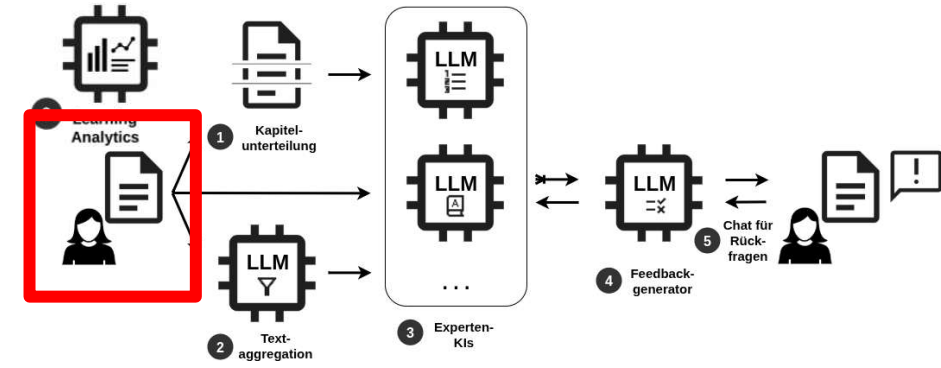


1. Kapitelunterteilung: Das Dokument wird in Kapitel, Unterkapitel, Absätze und Sätze zerlegt abgespeichert.
2. Textaggregation: Der Text wird zusammengefasst und getaggt.
3. Experten-KIs: Spezialisierte Modelle analysieren die Blöcke auf Themenbezug, Nachvollziehbarkeit, Sprache usw. und liefern Scores und gefundene Fehler.



4. Feedbackgenerator: Fasst die Fehler zusammen, priorisiert Probleme und erzeugt konkrete, lernförderliche Rückmeldungen.
5. Chatfunktion: Möglichkeit Rückfragen zu stellen
6. Learning Analytics: Analyse der Leistungen

Konvertieren von docx und pdf Dokumenten



Ausgangssituation:

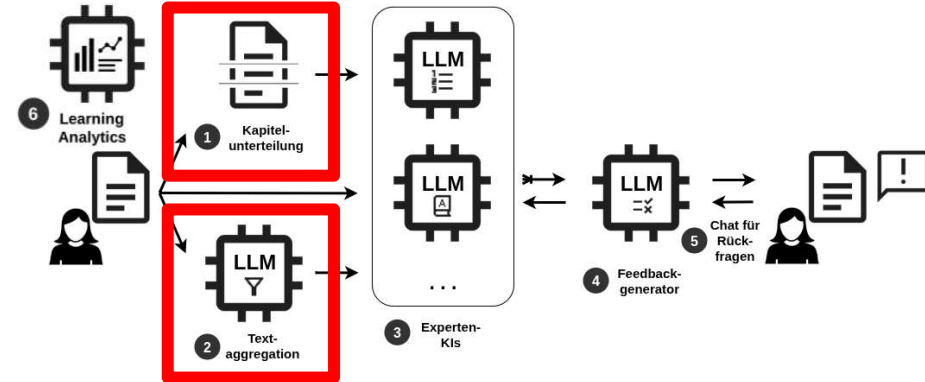
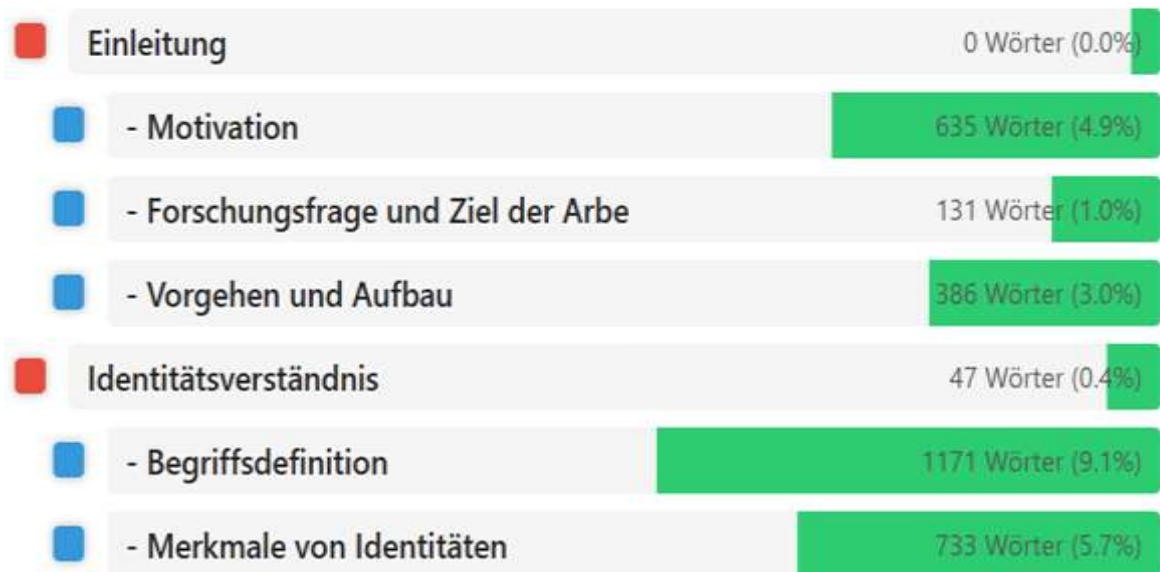
- Viele Seminararbeiten liegen in Formaten wie DOCX und PDF vor
- Diese Formate enthalten mehr als nur Fließtext (z. B. Struktur, Formatierung, Metadaten)

Schwachstelle:

- Bestehende Systeme (darunter Chatbots) erkennen oft keine Formatfehler
- Viele Fehler entstehen jedoch innerhalb der Formatierung



Spezial-LLM: Textaggregation



```

{
  "status": "ok",
  "file": "Vorlage_studentische_Arbeiten.docx",
  "document": {
    "type": "document",
    "id": 81,
    "children": [
      {
        "type": "section",
        "id": 1,
        "feedback": [
          {
            "coherence": 1.0,
            "flesch": 40.5,
            "outline": {
              "Schluss": 0.0,
              "Hauptteil": 0.0,
              "Einleitung": 1.0
            }
          }
        ]
      }
    ]
  }
}
  
```

Spezial-LLM: Analyse der Textabschnitte

Body

Textabschnitt

Readability: 29.8

Lesbarkeitsindex

Coherent

Bewertung der Nachvollziehbarkeit

Missing outline

Komponentenprüfung

Claim

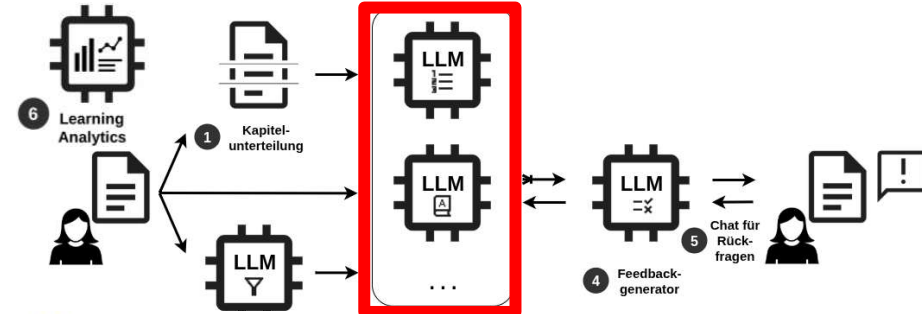
Satzarten

Definition

Definitionserkennung

Use abbreviation: ITS

Abkürzungsprüfung



Schreiben

Body

Readability: 29.8

Coherent

Missing outline

Missing introduction

Section 0, Paragraph 0

TextAid ist ein intelligentes Tutorensystem (ITS), das Studierende beim Verfassen wissenschaftlicher Texte mit gezieltem, lernförderlichem Feedback unterstützt.

Claim

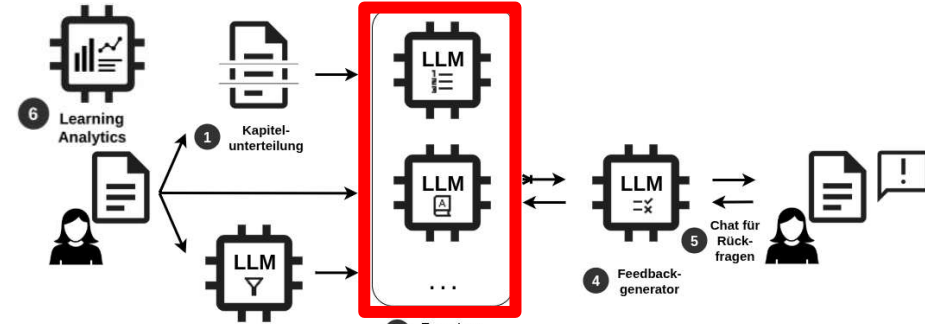
Definition

Das intelligente Tutorensystem soll als vorgeschaltete Instanz vor Dozierenden fungieren, indem es typische Probleme abfängt und Studierenden jederzeit beim Lernen hilft.

Future

Use abbreviation: ITS

Spezial-LLM: Satzarten erkennen



Zusammenfassend lässt sich festhalten, dass die in [1] dargestellte Literatur einen detaillierten Überblick über den aktuellen Stand existierender Ontologien für die Industrie 4.0 sowie über den Fortschritt in der Standardisierung liefert.

CONC
(97.97 %)

Der Fokus liegt dabei auf einer mikroskopischen Betrachtung der Domänen Robotik und Fertigung.

OTH (84.68 %)

In Kapitel III wurden exemplarisch die Ontologien SUMO, CORA, ROA und ORArch analysiert, aus denen die übergeordnete Ontologie O4I4 hervorgeht.

LINK (99.27 %)

Diese verfolgt das Ziel, verschiedene Domänen der industriellen Fertigung semantisch zu verbinden und damit Interoperabilität zu ermöglichen.

CLAIM (36.91 %)

Jedoch zeigen sich in den untersuchten Anwendungsszenarien (Kapitel IV) Grenzen hinsichtlich der architektonischen Umsetzung und industriellen Breitenanwendung.

COUNTERCLAIM
(63.38 %)

Insbesondere das Smart-RapidPrototyping-Szenario macht deutlich, dass eine umfassende, cloudbasierte Vernetzung der beteiligten Akteure bislang eher hypothetischer Natur ist.

COUNTERCLAIM
(93.84 %)

Auf makroskopischer Ebene stellen die Arbeiten zur German Standardization Roadmap Industrie 4.0 [2] sowie die Veröffentlichungen der Plattform Industrie 4.0 [4], [6], [8] umfassende Konzepte zur Digitalisierung industrieller Prozesse vor.

OTH
(79.74 %)

Im Zentrum dieser Beiträge steht die Vision einer semantisch integrierten Smart Factory entlang des RAMI-4.0 Referenzmodells.

OTH (97.97 %)

Dabei werden allgemeingültige Anforderungen formuliert, um langfristig einen standardisierten, interoperablen Rahmen für die Industrie 4.0 zu schaffen.

FUT (66.60 %)

Spezial-LLM: Feedback zu Definitionen

1. Erkennen von Definitionen im Text

2. Zerlegen der Definition in Bestandteile

Was wird definiert? (Begriff)

Wie wird es beschrieben? (Erklärung)

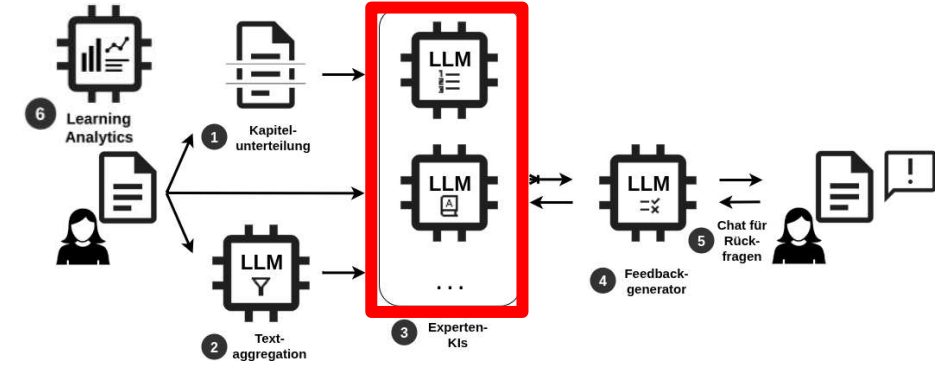
Werden Beispiele oder Einordnungen genannt?

3. Qualitätsprüfung der Definition

Sind alle Fachbegriffe definiert?

Wird der Begriff sinnvoll in den Text eingebettet?

Ist die Formulierung verständlich?



Künstliche Intelligenz (KI) bezeichnet die Fähigkeit von Maschinen, menschenähnliche Denkprozesse nachzuahmen. Definition

Maschinelles Lernen (ML) ist ein Teilbereich der KI, der es Systemen ermöglicht, aus Daten zu lernen und Vorhersagen zu treffen. Definition KI wird

mittlerweile in vielen Branchen zur Optimierung von Prozessen eingesetzt. Maschinelles Lernen kommt etwa bei der Analyse großer Datensätze im Gesundheitswesen zum Einsatz. In dieser Arbeit wird untersucht, wie KI-Systeme zur automatischen Analyse wissenschaftlicher Texte verwendet

werden können. Ein neuronales Netz ist ein Modell, das aus künstlichen Neuronen besteht und komplexe Muster erkennt. Definition Viele moderne

KI-Anwendungen basieren auf neuronalen Netzen.

Spezial-KI: Formalia-Prüfung

- Überprüfung von Zitation und Quellenarbeit
- Analyse des Literaturverzeichnisses (Vollständigkeit, Konsistenz, Format)
- Kontrolle der formalen Anforderungen (Schriftart, -größe, Formatierung)
- Prüfung der Einhaltung vorgegebener Formalia
- Analyse der Struktur der Arbeit (z.B. Aufbau, Gliederung)
- Rechtschreib- und Sprachprüfung
- Flexibel konfigurierbar

Auszug des Feedbacks:

 = Konsistenz Zitationen

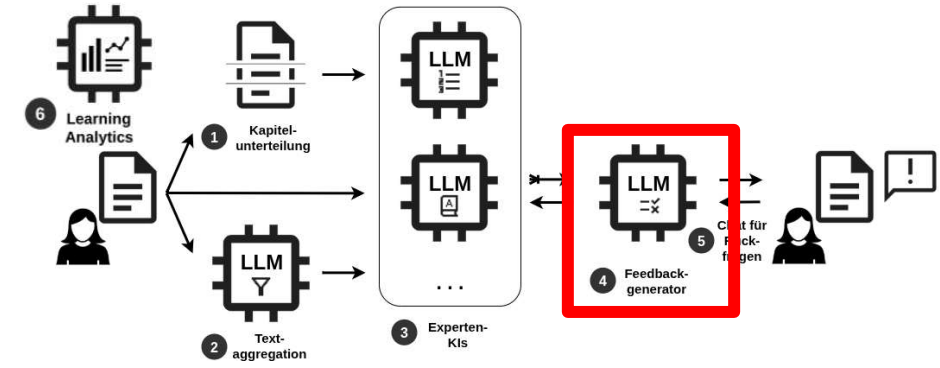
Wurden Zitationen konstant einheitlich gehalten?

| | |
|-------------------------------|---------------------|
| Formalia | |
| Struktur | |
| Zitation | |
| Zitationsstil | Harvard |
| Zitierweise | In-Text-Zitation |
| Zitation in Klammern | Runde Klammern "()" |
| Trennzeichen zwischen Autoren | and |
| Jahr in Klammern | Jahr in Klammern |
| Sprache | ▼ |

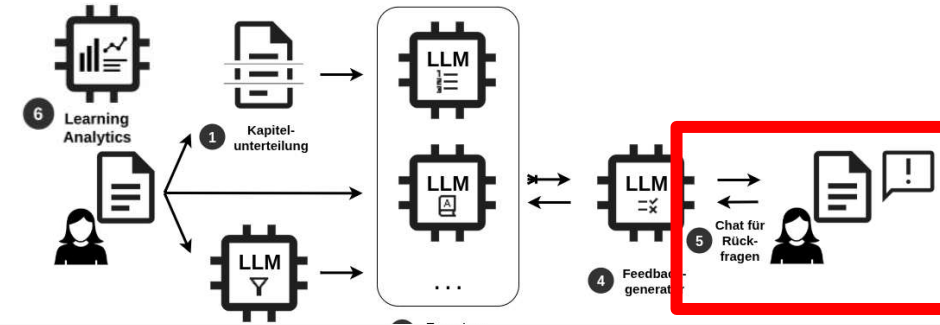
Korrekt ✓

Feedback formulieren

- Hauptproblem beim Generieren des Feedbacks:
 - Zusammenbringen der Ergebnisse aus allen Experten KIs
 - Nur so viel Feedback wie nötig
 - Konkrete Probleme der Studierenden angehen
- Qualitätsmerkmale guten Feedbacks:
 - Spezifisch: Bezieht sich auf klar identifizierbare Aspekte des Produkts oder Prozesses
 - Hilfreich: Gibt handlungsleitende, konstruktive Impulse statt bloßer Bewertung
 - Valide: In fachlicher und didaktischer Hinsicht korrekt, nachvollziehbar und begründet



Chat-Assistent



Schwächen:

- **Sprachliche Unschärfe:** Phrasen wie „Prozessketten werden häufig meist unbeabsichtigt in die Länge gezogen“ sind umständlich und unklar formuliert.
- **Wiederholung und redundante Formulierungen:** „...schnellstmöglich und auf dem besten Wege zu bewältigen gilt“ ist eine überflüssige Wiederholung.

Chat

user:

Worauf muss bei klarer Formulierung geachtet werden?

assistant:

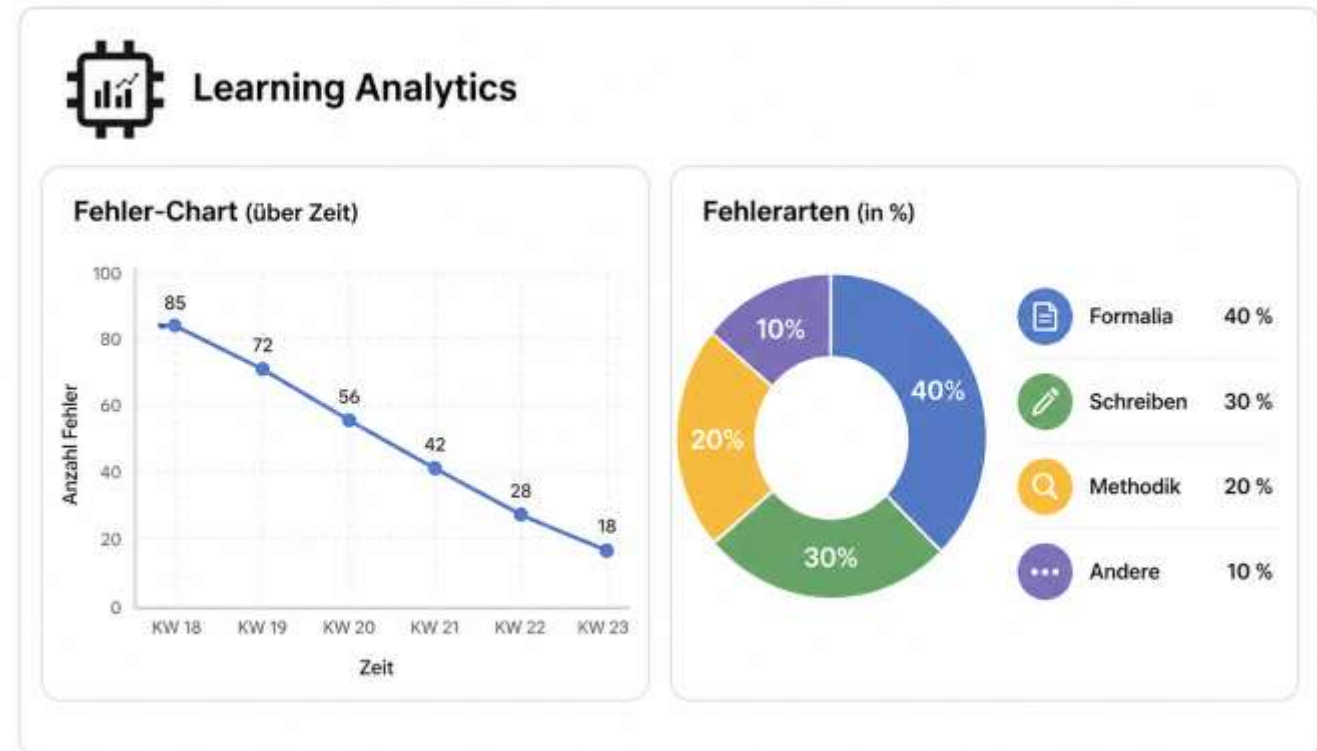
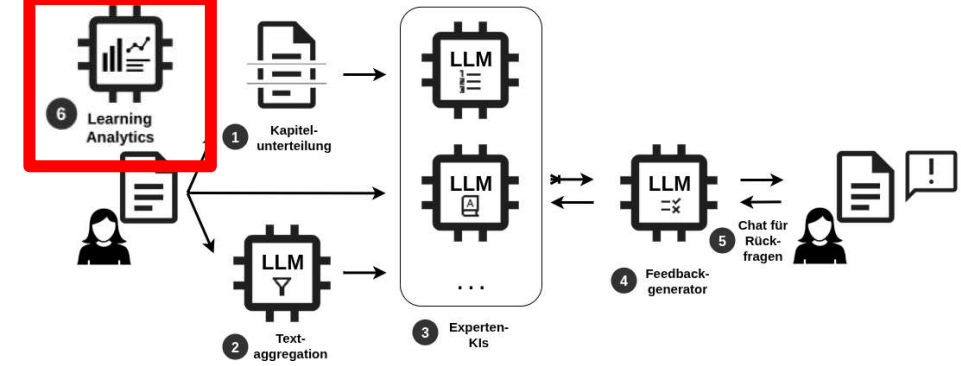
Hallo! Ich bin wieder da. Es ist großartig, dass du dich mit **klarer Formulierung** beschäftigst – das ist ein wichtiger Teil des Textens.

Ziel:

Du sollst verstehen, **welche Aspekte bei klarer Formulierung wichtig sind**, ohne dass ich dir eine Lösung gebe, sondern du selbst lernst, wie du deine Sätze präziser und verständlicher formulierst.

Learning Analytics / Rückmeldung an Lehrende

- KI-Nutzung vermindert die Sichtbarkeit von Lernprozessen und Leistungsständen für Lehrende
- Herausforderung: tatsächlichen Kompetenzstand trotz KI-Unterstützung erkennen
- TextAid analysiert Schreibprozesse statt nur Endprodukte
- Ermöglicht Sichtbarmachung individueller und aggregierter Muster (z. B. Definitionsqualität, Struktur, Feedbackverläufe)
- Transparente, erklärbare Analysen zur Unterstützung von Reflexion



Nur Entwurf, nicht final

Wie funktioniert es technisch?

1. Grundlage: Sprachmodelle (Large Language Models)

- Wir nutzen vortrainierte KI-Sprachmodelle.
- Diese Modelle wurden mit sehr großen Textmengen trainiert und „verstehen“ dadurch Sprachmuster.

2. Spezialisierung durch Fine-Tuning

- Statt Modelle neu zu trainieren, passen wir es gezielt an.
- Dafür nutzen wir Trainingsdaten: z. B. Wikipedia-Artikel mit markierten Definitionen.

Ergebnis:

- Ein Modell, das speziell Definitionen erkennt.

3. KI-gestützte Feedback-Generierung

- Die erkannten Strukturen werden analysiert.
- Daraus wird gezieltes Feedback mittels Prompting (spezifische Instruktionen an das Sprachmodell) erzeugt, z.B.:
„Der Begriff wird genannt, aber nicht definiert.“

Ressourcenverbrauch pro Auswertung

| Auswertung | Methode | Ressourcenverbrauch |
|-----------------------------------|---------------------------|---------------------|
| Schlüsselwörter extrahieren | BERT-Modell | 0.1B |
| Ähnlichkeit von Texten messen | BERT-Modell | 0.1B |
| Kohärenz klassifizieren | BERT-Modell | 0.3B |
| Textteile klassifizieren | BERT-Modell | 0.3B |
| Definitionen erkennen | BERT-Modell | 0.3B |
| Bewertung der Nachvollziehbarkeit | GEMMA3, Prompting | 27B |
| Rechtschreibprüfung | Externes Open-Source-Tool | Minimal |
| Prüfung auf Füllwörter | Regelbasiert | Minimal |
| Zeichensetzung prüfen | Externes Open-Source-Tool | Minimal |
| Lesbarkeitsscore | Regelbasiert | Minimal |
| Zum Vergleich | GPT-5 | ~1700B |

Einfaches, skalierbares System

Datenschutz & Infrastruktur

- On-Premise, keine Abhängigkeit von externen KI-Diensten
- Transparenz & Nachvollziehbarkeit im Fokus
- Ab Release: offen zugänglich und von der Community weiterentwickelbar



**DSGVO-konform,
On-Premise**



**Kein externer
Cloud-Zugriff in die USA**



Open-Source



**Ziel:
Vertrauenswürdiger
KI-Schreibcoach**

Integrationsmöglichkeiten in andere Systeme

TextAid ist ein offener, kostenfreier und DSGVO-konformer Schreibkompetenzcoach. Dieser Coach soll auf Rückfragen der Studierenden reagieren können und selbstständig weiterlernen.

Kooperationen mit anderen StIL Projekten ist geplant.

Bereits im Austausch mit dem NeDa Netzwerk (neda.digital).

TextAid Backend 0.1 OAS 3.1

| | | | |
|------|--|-----------------------------------|---|
| GET | <code>/node/{doc_id}/{path_id}</code> | Retrieve single node content | ▼ |
| GET | <code>/feedback_node/{doc_id}/{path_id}</code> | Retrieve node-specific feedback | ▼ |
| GET | <code>/document/{doc_id}/nodes</code> | Retrieve paginated document nodes | ▼ |
| POST | <code>/feedback/generate</code> | Generate Feedback | ▲ |

Übersicht des Entwicklungsstands



- Entwicklungsstand einsehbar unter:
 - <https://frits.mni.thm.de/>
 - <https://transferkiosk.net/p/101519>