

TECHNOLOGY LAB

04.07.2025

TASK 3

Alex Casapu, Fabian Röhrle, Diar Sulejmani





Gliederung

- 01 Aufgabenstellung
- 02 Infrastruktur
- 03 Fragen Generierung & Antwort Validierung
- 04 Analytics
- 05 Gamification
- Of Studentmodule und Gamification
- 07 Live-Demo
- 08 Fazit und Lessons Learned
- 09 Ausblick

Aufgabenstellung

Implement and evaluate the "student module" of the application and refine the "teacher module".

Student Module:

- Design UI of student module including screen to answer
- questions and screens to emphasize the gamification aspect of
- the application (e.g. ranks, coins, levels)
- Integrate UI with the questions / answers generated in Task2
- In case of open text answers, think of a way to validate them
- Evaluate student module

Teacher Module:

- Design a screen to perform summary analysis about the learning progress
- of the students using the application
- Data must be anonymized
- Additionally:
- Question/Answer Validation

Cloud Setup & Infrastruktur

Technology Stack

Frontend:

- React
- Next.js
- Tailwind CSS
- shadcn/ui

Backend:

- Node.js - Next.js
- SupabaseAWS

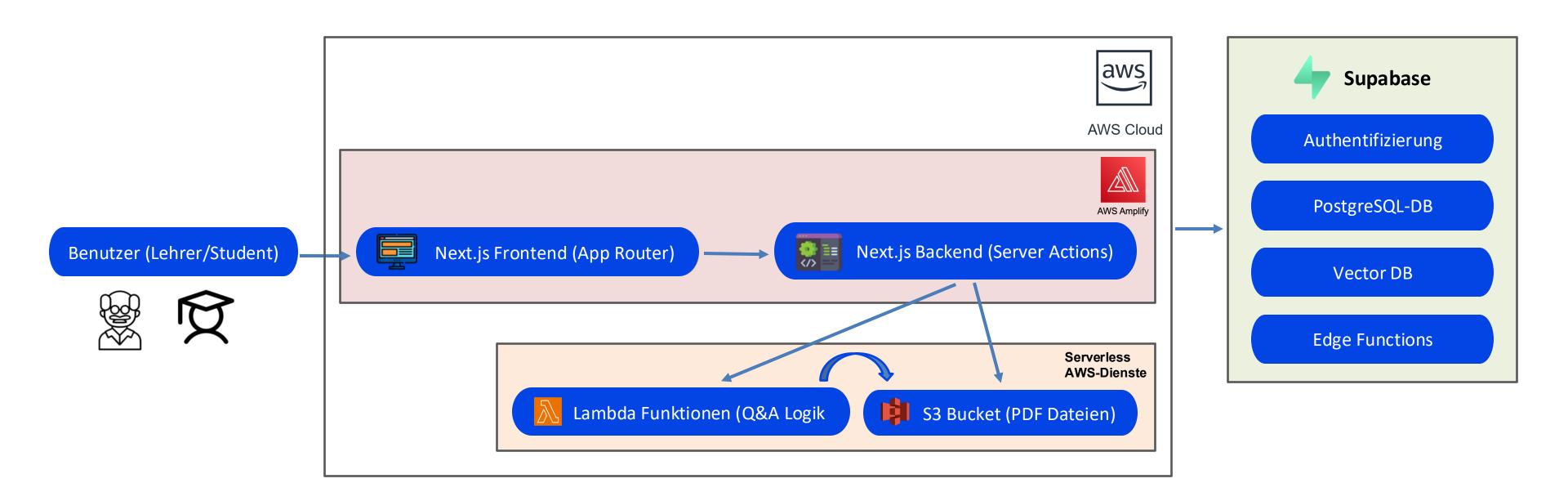
DB / Auth: Al:

- Agno
- OpenAl - Gemini
- AWS



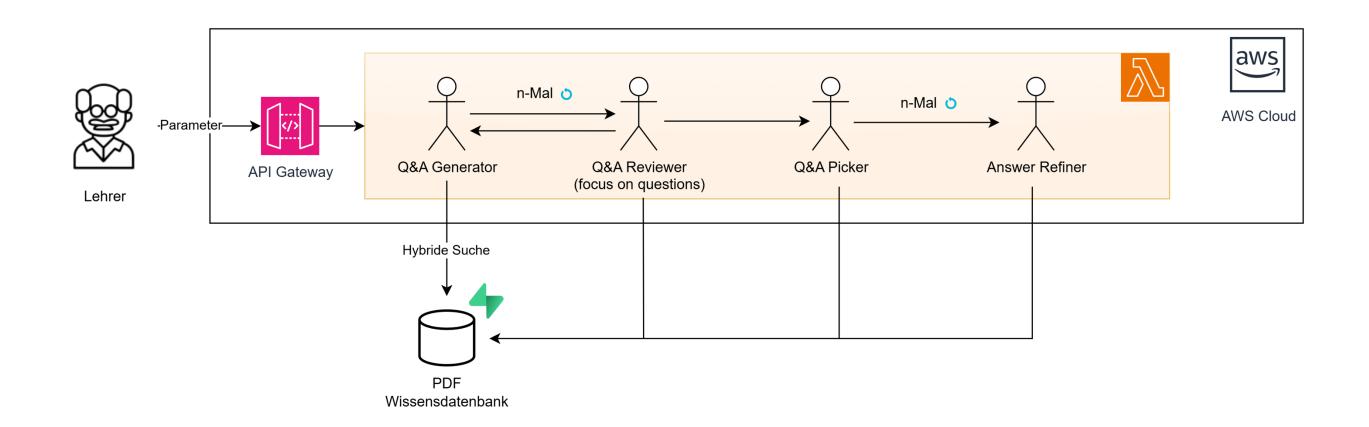
Cloud:





Generierung von Q&A-Paaren (Konzept)

- **Grundprinzip:** "Generiere → Kritisiere → Wähle aus → Überarbeite"
- Qualität von Q&A-Paaren wird fortlaufend in mehrstufigem Prozess verbessert
 - Feedback-Schleife: Iterative Verbesserung von Q&A-Paaren (Feed-Forward-Prinzip)
 - Best-of-Selektion: Auswahl der besten Q&A-Paare aus Gesamtpool (Kuratierungs- bzw. Filterprinzip)
 - o Ggf. Überarbeitung und Anreicherung der Antworten am Ende
- Alle Prozessschritte berücksichtigen definierte Qualitätskriterien bzgl. Fragen & Antworten



Generierung von Q&A-Paaren (Demo)

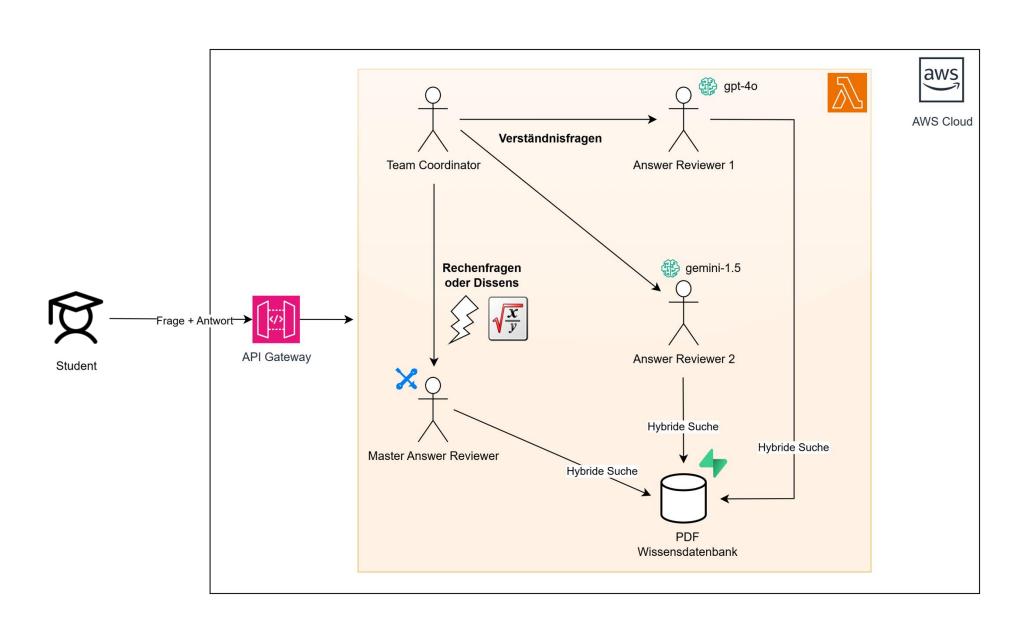


Validierung von Studentenantworten (Konzept)

Interrater-Reliabilität

Warum zwei unterschiedliche LLM-Reviewer?

- Unabhängige Urteile GPT-40 und Gemini 1.5 bewerten dieselbe Antwort getrennt
- Statistische Zuverlässigkeit Übereinstimmung der Modelle ⇒ größeres Vertrauen in das Ergebnis
- Bias-Reduktion Unterschiedliche Trainingsdaten/Architekturen gleichen individuelle Modellfehler aus
- Escalation-Mechanismus Bei Abweichungen übernimmt "Master Reviewer"
- → Überträgt menschliche Peer-Review-Prinzipien (Mehrgutachter, Konsensfindung)



Validierung von Studentenantworten (Demo)



Statistiken zur Fragengenerierung

>50 Fragen

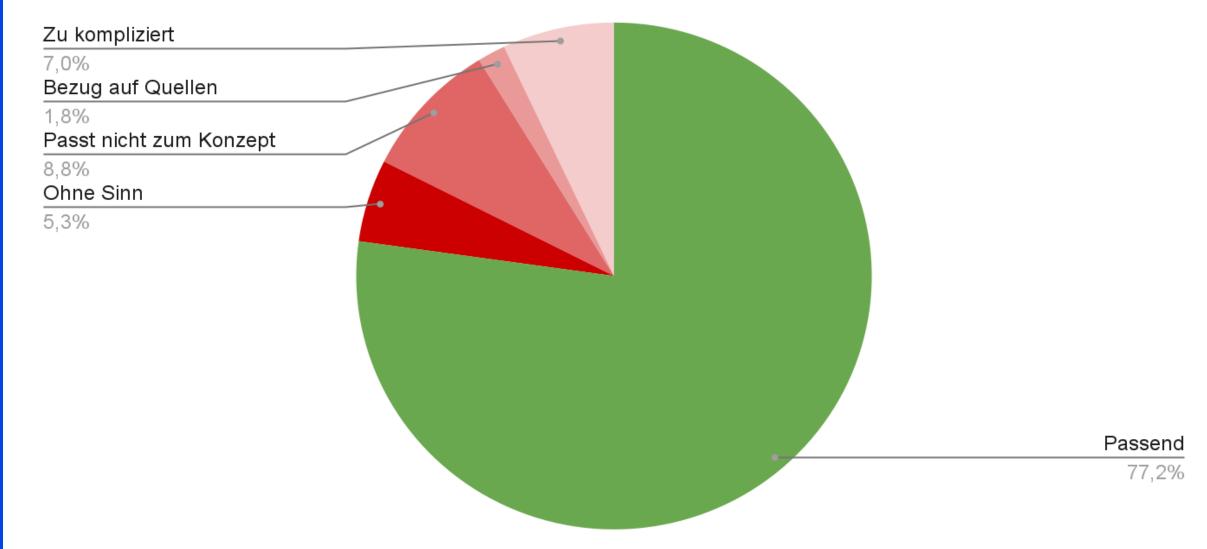
Wichtige Ausschlusskriterien

7 Zeit

2 Aufgabenart

Sinn der Frage

Angemessenheit der Fragen





Statistiken zur Antwortgenerierung

> 50 Fragen

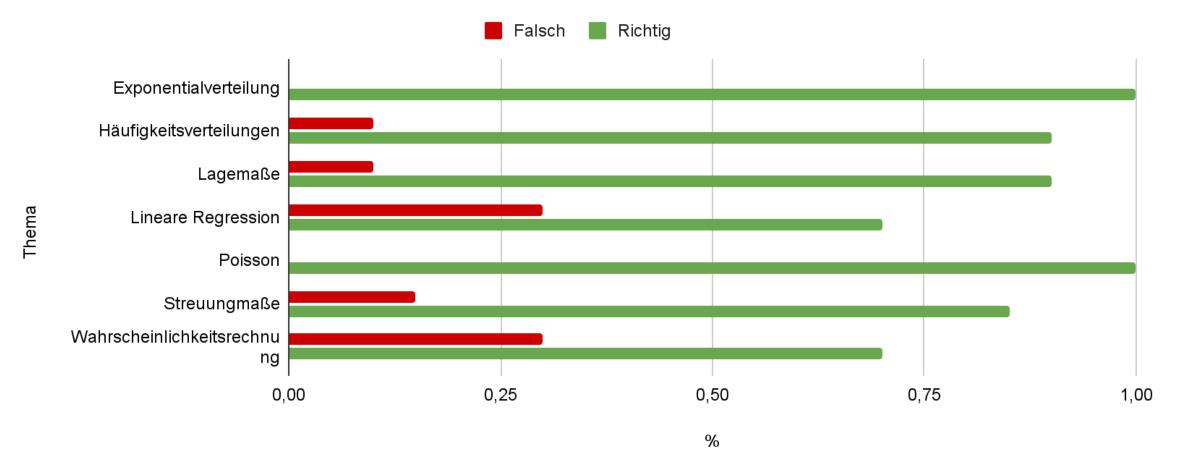
Beobachtungen

Textantworten waren zu 98% richtig

2 38% der falschen Antworten enthielten Rechnungen und Zahlen

7% der Antworten waren als "Falsch" markiert als Folge von irreführende Fragen

Antwortkorrektheit nach Themen







Kurs Analytics

Umfassende Statistiken zu Ihren Kursen



Eingeschriebene Studenten

4



Quiz-Versuche

 ∇ Alle Kurse

Versuche:

Neueste

Perfekte Abschlüsse \odot



Alle Fragen korrekt beantwortet

Abschlussrate

50%

der Schüler haben ihre Kurse abgeschlossen

Ø Session-Dauer

3min

Durchschnittliche Zeit pro Quiz-Session

Kurs-Schwierigkeitsranking (Multi-Faktor)

Normalverteilung

2 Einschreibungen • 0 Abschlüsse

Sehr schwer 75 Schwierigkeits-Score

Statistik

1 Einschreibungen • 1 Abschlüsse

20 Mittel Schwierigkeits-Score

Wahrscheinlichkeitsrechnung

1 Einschreibungen • 1 Abschlüsse

Einfach Schwierigkeits-Score





Normalverteilung

2 Einschreibungen • 0 Abschlüsse

Abschlussrate

0%

Falsche Antworten

66.7%

Schwierigkeitsfaktoren (Beitrag zum Score):

Abschlussrate (25%)

Perfekte Abschlüsse (20%)

Versuche pro Frage (25%)

Falsche Antworten (20%)

Abbrüche (10%)

Perfekte Abschlüsse

0%

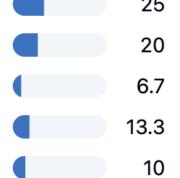
Abbruchrate

100%

Ø Versuche/Frage

2.3







Abbruchpunkte

Normalverteilung

Frage 2

Abbrüche

Normalverteilung

Frage 1

Abbrüche



⚠ Top 3 Schwierigste Fragen

#1

Normalverteilung

Erklären Sie die 68-95-99.7 Regel in Bezug auf die Normalverteilung.

Falsche Antworten: 100% Ø Versuche: 3 Gesamt Antworten: 1

#2

Normalverteilung

Welche Eigenschaft hat die Standardnormalverteilung im Vergleich zur allgemeinen Normalverteilung?

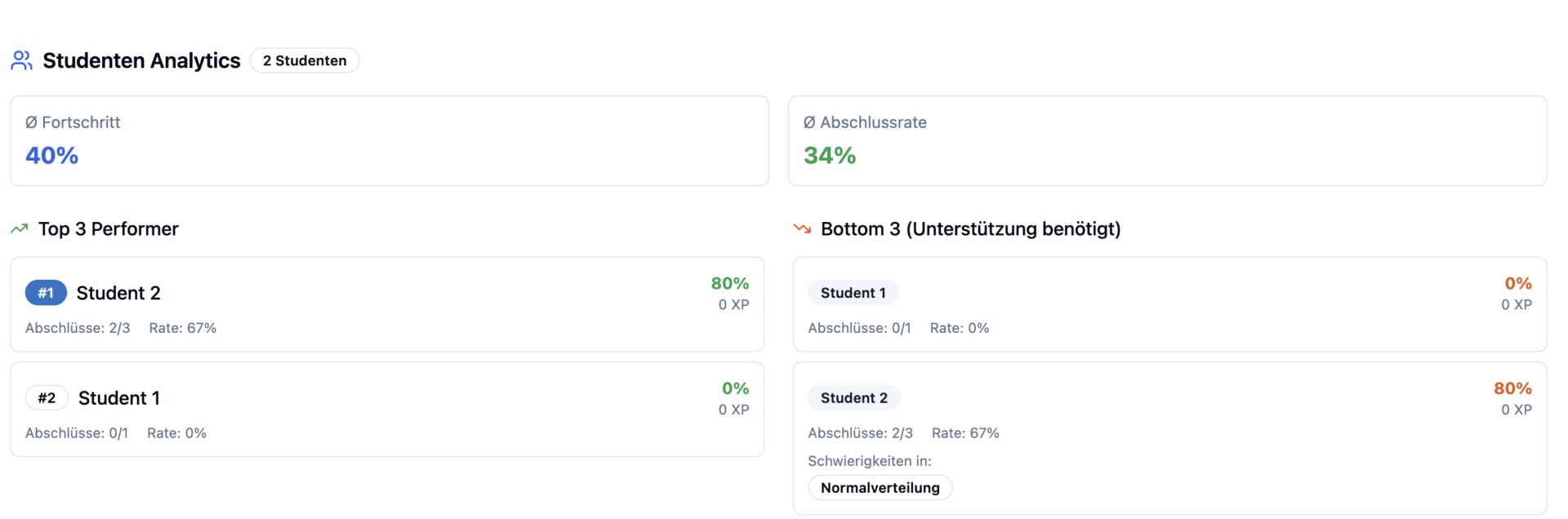
Falsche Antworten: 100% Ø Versuche: 2 Gesamt Antworten: 1

Normalverteilung

Was sind die zwei Parameter, die eine Normalverteilung definieren?

Falsche Antworten: 0% Ø Versuche: 2 Gesamt Antworten: 1





Gamification









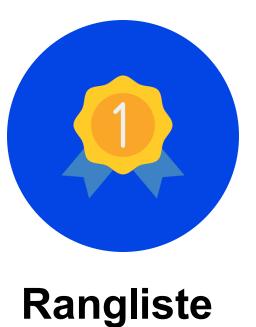
XP System

Economy

Shop System



Streaks





Farmville

Live Demo

Duo Quanto Logo





VIELEN DANK

FÜR IHRE AUFMERKSAMKEIT

04.07.2025

Alex Casapu, Fabian Röhrle, Diar Sulejmani



Quellen

- ❖ Bloom, B. S., Englehart, M. D., Furst, E. J., Hill, W. H., & Krathwohl, D. R. (1956). Taxonomy of educational objectives, handbook: The cognitive domain.
- Wilson, L. O. & Anderson and Krathwohl. (2016). Bloom's taxonomy revised. https://quincycollege.edu/wp-content/uploads/Anderson-and-Krathwohl Revised-Blooms-Taxonomy.pdf
- * LLM evaluation metrics and methods. (n.d.). https://www.evidentlyai.com/llm-guide/llm-evaluation-metrics?utm_source=chatgpt.com
- Schedler, M. (2020). Mit Gamification spielend die Schulen verändern. In *Pädagogische Hochschule Vorarlberg* | *F&E Edition 26*. https://www.ph-vorarlberg.ac.at/fileadmin/user-upload/RED_SOZ/PDFs/F_E_26/FE26_03_Schedler.pdf
- Li, C., Liang, L., Fryer, L. K., & Shum, A. (2024). The use of leaderboards in education: A systematic review of empirical evidence in higher education. *Journal of Computer Assisted Learning*. https://doi.org/10.1111/jcal.13077
- https://www.clearinghouse.edu.tum.de/lehrstrategien/quizze-im-unterricht-und-in-der-lehre-wie-haeufiges-testenzum-lernen-beitraegt/