

Bewerbung für studentischen DINI-Wettbewerb:
"Mit KI lernen: Intelligente Bildung für eine intelligente Zukunft"

**Hallo,
ich bin**

**(\ /)
A · I · K A**

**KI-gestützte Integrations-App für Migrant:innen
Deutsch, rechtliche & politische Bildung**

28.06.2024

KONTAKT:

StAI Hub - student AI Hub Marburg
ai.lab@uni-marburg.de





ÜBERSICHT

Wir entwickeln eine KI-gestützte Integrationsapp, die darauf abzielt, die frühen Stufen des Integrationsprozesses für Migrant:innen zu unterstützen. Unser Curriculum gliedert sich in drei Bereiche: Deutsch, Recht und Politik. Durch den Einsatz künstlicher Intelligenz werden alle Lerninhalte und die dazugehörenden Aufgaben erstellt, sowie die Fragen zum Lernstoff in einem Dialog mit dem System beantwortet. Dies ermöglicht es, sprachliche Defizite effektiv zu adressieren und ein tiefergehendes Verständnis für rechtliche und politische Themen zu fördern.

Das intelligente System unterstützt Mehrsprachigkeit und kann Deutsch, Englisch, Russisch, Ukrainisch, Arabisch und Türkisch verstehen. Darüber hinaus kann die KI sprachliche Unterstützung in alltäglichen Situationen und rechtliche Hinweise für spezifische Fragen anbieten. Dies hilft den Teilnehmenden, sich nicht nur sprachlich, sondern auch kulturell und politisch in Deutschland zu orientieren.

TEAM

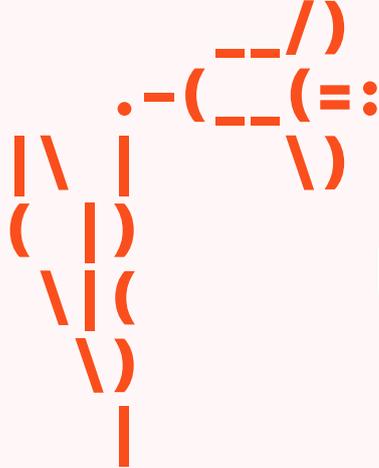
Wir sind ein interdisziplinäres studentisches Team von der Philipps Universität Marburg, das Teil der **stAI Hub** Hochschulgruppe an der Universität Marburg ist. Unser Team besteht aus:

- **Ali Kholmovaia** (B.Sc. Informatik)
- **Tim Ushakov** (B.Sc. Informatik)
- **Daria Gubina** (M.Sc. Politikwissenschaft)

SAFE & OPEN SOURCE AI

Wir, **stAI Hub**, setzen uns für **die Demokratisierung** und **die sichere Anwendung** von KI ein.

Für uns ist es wichtig, nur LLM-Modelle zu verwenden, die Open Source sind und auf lokalen europäischen Servern laufen. Darüber hinaus schätzen wir die potentiellen Risiken der Anwendung von AIED (AI in Education) als hoch ein und ergreifen Maßnahmen, um diese zu minimieren. So verhindern wir z.B. die Generierung potenziell toxischer oder nicht themenbezogener Agentenantworten. Außerdem wollen wir, dass das Endprodukt allen Migrant:innen kostenlos zur Verfügung stehen wird.



MOTIVATION

Der Anstoß für unser Projekt kam aus unseren eigenen Erfahrungen als internationale Studierende. Wir wünschten uns eine Anwendung, die sowohl Einblicke in das deutsche Rechts- und Politiksystem bietet als auch beim Erlernen der deutschen Sprache unterstützt. Uns wurde klar, dass der Mangel an spezialisierten Bildungsangeboten für Migrant:innen in den Bereichen Recht und Politik sowie zusätzlicher Sprachförderung eine Barriere darstellt, die Menschen daran hindert, nicht nur über ihre Rechte und Pflichten informiert zu sein, sondern sich auch als Teil der Gesellschaft zu fühlen und aktiv am gesellschaftlichen Leben teilzunehmen.

Unser Engagement in der Flüchtlingshilfe hat uns dazu bewogen, uns zunächst auf die Geflüchteten zu konzentrieren. Aus eigener Erfahrung wissen wir, dass gerade für diese Gruppe der Mangel an sprachlicher Förderung und rechtlicher Bildung katastrophale, wenn nicht gar lebensbedrohliche Folgen haben kann.

Offizielle Zahlen bestätigen unsere Erfahrungen: Die durchschnittliche Wartezeit auf einen Integrationskurs betrug im vergangenen Jahr mehr als fünf Monate*. Dies ist die Wartezeit, in der die Migrant:innen auf ein Kursangebot warten. Auch im Bereich der rechtlichen und politischen Bildung besteht ein Mangel an Fachpersonal und zugänglichen Angeboten, was die Integration erschwert und auch die Arbeit von Sozialarbeiter:innen belastet.

Bevor wir mit der Entwicklung der App begannen, haben wir verschiedene Sprachlern-Apps getestet. Wir stellten fest, dass es zwar viele Angebote gibt, aber keines ein strukturiertes Curriculum bietet und gleichzeitig auf Sprachtests vorbereitet, vergleichbar mit traditionellen Sprachkursen. Im Gegensatz dazu wollen wir mit unserer App sowohl ein strukturiertes Curriculum und eine Sprachtestvorbereitung anbieten, als auch die Menschen auf das Leben in Deutschland, auf rechtliche und sprachliche Alltagssituationen vorbereiten.

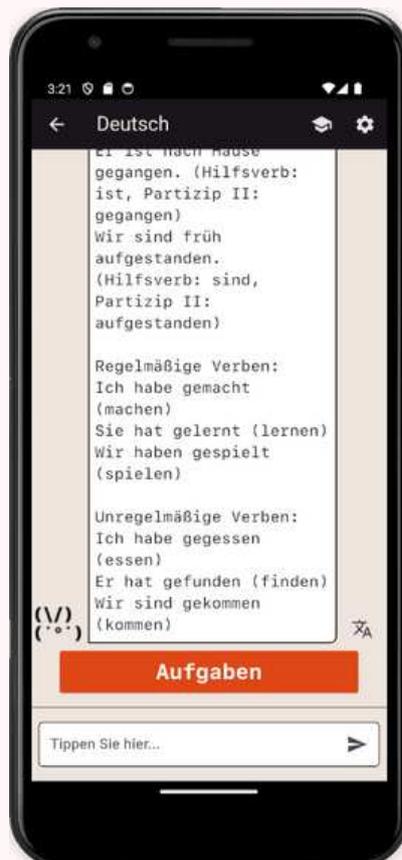
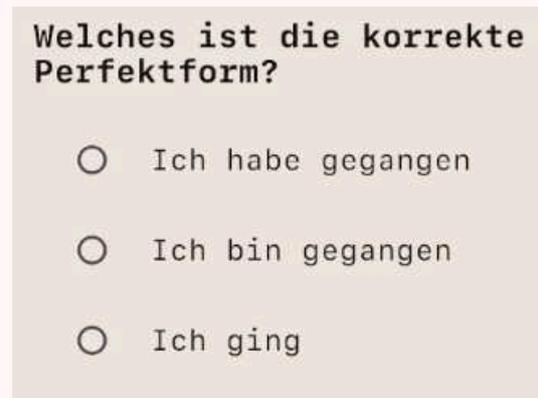
*Bundesrechnungshof. (2023). Bericht nach § 88 Absatz 2 BHO an den Haushaltsausschuss des Deutschen Bundestages: Förderung von Kriegsflüchtlingen aus der Ukraine in Integrationskursen. Abgerufen von <http://www.bundesrechnungshof.de>



DEUTSCH

Unser Modul zum Deutschlernen konzentriert sich darauf, die Schreib-, Lese- und Hörverständnisfähigkeiten der Nutzer durch eine Serie von interaktiven und adaptiven Lektionen zu fördern. Das von uns entwickelte intelligente Tutorsystem bietet nicht nur traditionelle Lerninhalte wie Grammatikerklärungen, Lese- und Hörtexte, sondern auch erstellt die Aufgaben zu jeder Lektion. Diese Aufgaben umfassen Single-Choice-Fragen, Lückentexte und offene Fragen, allesamt generiert durch das Sprachmodell.

Das intelligente System kann Deutsch, Englisch, Russisch, Ukrainisch, Arabisch und Türkisch verstehen. Nutzer:innen haben auch die Möglichkeit, alle Systemnachrichten in ihre bevorzugte Sprache zu übersetzen.



DEUTSCH

Zusätzlich ermöglicht das Tutorsystem einen dialogorientierten Lernansatz. Nutzer können in Echtzeit Fragen zu grammatikalischen Details, Bedeutungen von Ausdrücken oder nach weiteren Beispielen stellen. Das System bietet auch situative Formulierungshilfen an, die den Nutzern helfen, ihre Sprachkenntnisse in verschiedenen Alltagssituationen anzuwenden und zu vertiefen.

Dieser interaktive und benutzerorientierte Ansatz stellt sicher, dass jeder Lernende individuell unterstützt wird und effektiv gemäß dem Lehrplan des Sprachniveaus A2 voranschreiten kann.

Schreiben Sie drei Sätze im Perfekt über Ihre gestrigen Erlebnisse. Verwenden Sie sowohl "haben" als auch "sein" als Hilfsverben.

Type your answer here...

Höre die Aufnahme an und löse im Anschluss die Aufgaben.



Aufgaben



RECHT & POLITIK

Unser Curriculum Recht und Politik ist als interaktiver Fragenkatalog konzipiert. Die Benutzer:innen können aus einer Liste von Fragen auswählen, die thematisch sowohl rechtliche als auch politische Inhalte abdecken. Nach der Auswahl einer Frage werden detaillierte Informationen zum jeweiligen Thema angezeigt. Darüber hinaus besteht die Möglichkeit, mit unserer KI in einen Dialog zu treten und das Thema zu vertiefen.



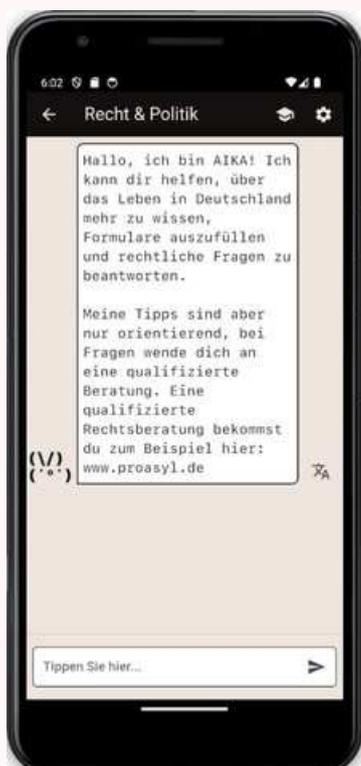
Die KI antwortet dabei ausschließlich auf Basis von Informationen aus einer Wissensdatenbank, die von uns manuell auf Basis von offiziellen Leitfäden und Beratungshilfen zusammengestellt wurde. Dadurch sollen präzise und verlässliche Antworten gewährleistet werden, die den Lernenden helfen, ein fundiertes Verständnis der rechtlichen Grundlagen und politischen Strukturen in Deutschland zu entwickeln.

Eine Ausbildungsduldung ist eine Form der Duldung für Menschen, die in Deutschland eine Ausbildung machen. Sie müssen dafür seit mindestens 3 Monaten eine Duldung haben, eine staatlich anerkannte Ausbildung machen, die mindestens 2 Jahre dauert, und Ihre Identität muss geklärt sein. Es gibt Ausnahmen und weitere Details. Wenn Sie im Dublin-Verfahren sind, ein Arbeitsverbot haben oder wegen einer vorsätzlichen Straftat zu einem bestimmten Strafmaß verurteilt wurden, können Sie keine Ausbildungsduldung bekommen.

Quellen:

- Berufliche Perspektiven für Geflüchtete in Thüringen
- Duldung + Arbeit = Aufenthaltserlaubnis? 04/2023 Mit den Änderungen des ChAR-G

Beispiel von Antworten

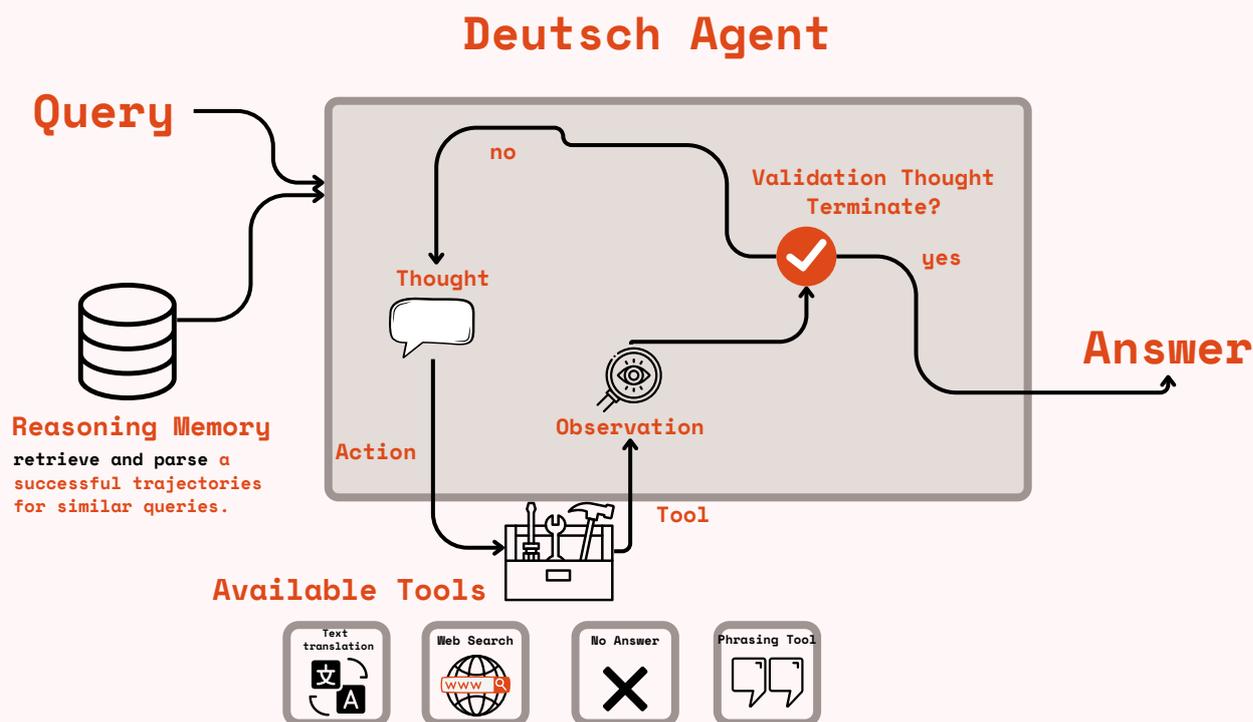


Für das Deutschlernen setzen wir das REACT-Agent-Pattern* in Verbindung mit dem Large Language Model (LLM) „mixtral-8x7b“ ein. Der Agent startet mit der Produktion von „Thoughts“, basierend auf Nutzeranfragen und vorgegebenen Prompts. Dies führt zur Auswahl eines spezifischen Tools aus unserer Liste, für das der Agent notwendige Eingaben definiert und ausführt. Die daraus resultierende „Observation“ informiert den nächsten Schritt.

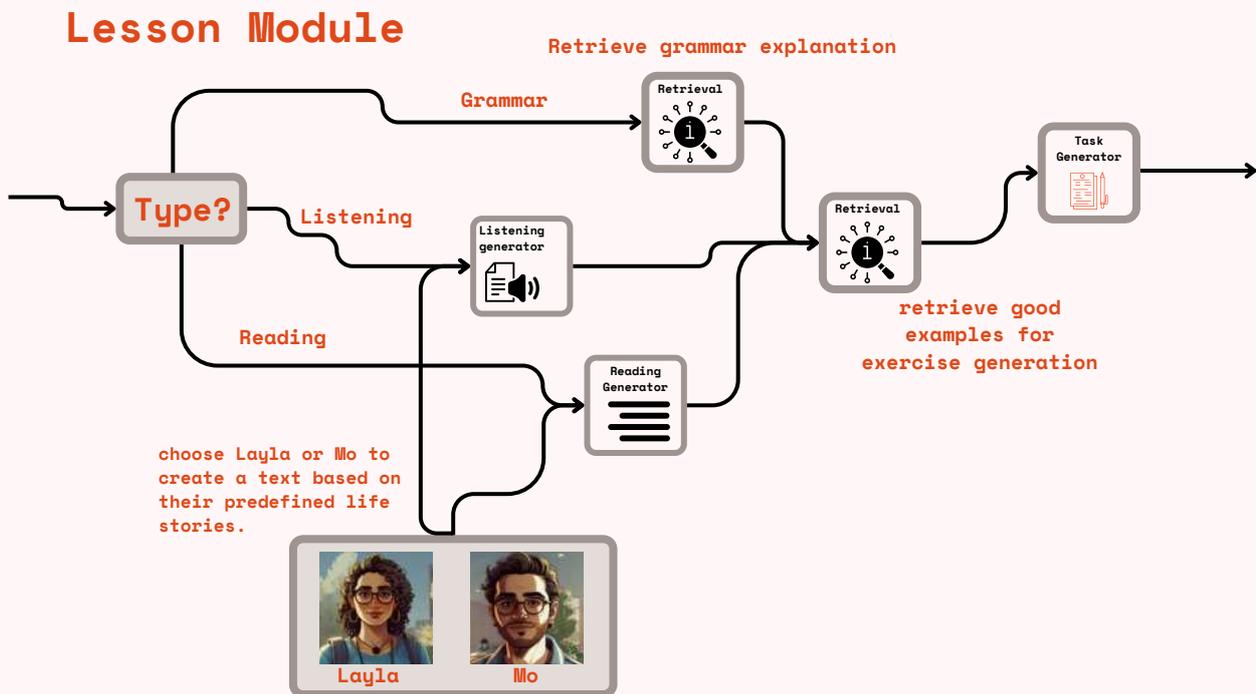
Der Agent erstellt daraufhin einen Validierungsgedanken, der die Observation analysiert und entscheidet, ob eine finale Antwort gegeben werden kann oder ob aufgrund von Irrelevanz oder toxischem Inhalt das „no answer tool“ zur Anwendung kommt, welches eine Standardmeldung ausgibt.

Als Tools können verschiedene Aktionen konzipiert werden. Z.B. externe Anwendungen wie DeepTranslate oder Websuche, spezifische Anfragen an das LLM, z.B. Phrasing Tool, das Formulierungsvorschläge liefern soll.

Beispiel: Bei der Nutzeranfrage „Übersetze für mich 'Im Rahmen von dieser Arbeit lernen wir viel über intelligente Agenten'“ generiert das LLM einen Gedanken, entscheidet sich für das Übersetzungstool und führt es mit der korrekten Eingabe aus. Der Übersetzer liefert den übersetzten Text und das wird zur endgültigen Antwort des Systems.



*S. Yao et al., "ReAct: Synergizing Reasoning and Acting in Language Models," arXiv.org, Mar. 09, 2023. <https://arxiv.org/abs/2210.03629>

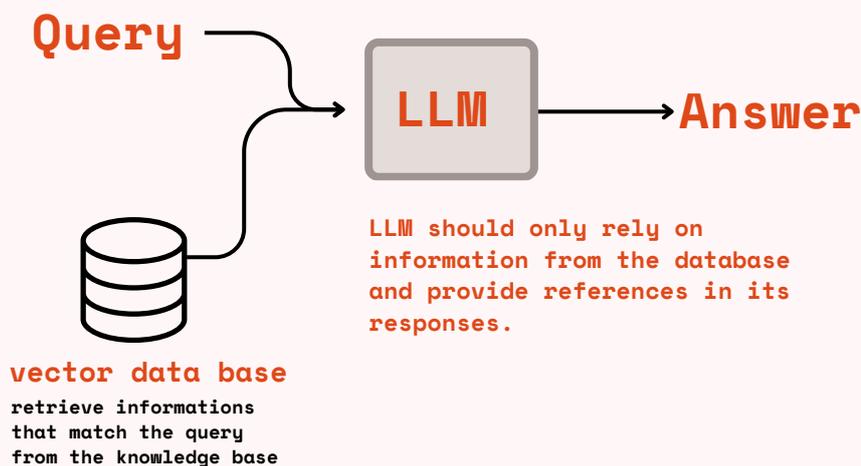


Unser Lektionsmodul ermöglicht es, je nach Lektionstyp unterschiedliche Inhalte zu generieren. Das Modul unterstützt drei Haupttypen von Lektionen: Grammatik, Hörverstehen und Leseverstehen. Je nach gewähltem Typ wird eine entsprechende Anfrage an das LLM gesendet. Für grammatikalische Erklärungen haben wir uns jedoch gegen eine Generierung entschieden, um eine gute Qualität der Erklärungen zu gewährleisten.

Für Lesetexte beispielsweise erhält das LLM eine vordefinierte Biografie: von Layla oder Mo, Texte, die wir erstellt haben. Das LLM verwendet diese Biografien, um kontextbezogene, relevante Geschichten zu generieren, die thematisch zur Lektion passen. Bei Hörtexten wird zusätzlich das Audio aus einem Text2Speech Open Source Modell generiert.

Nach der Generierung des Textes erstellt das LLM Aufgaben, die sich auf den Inhalt des Textes oder auf spezifische grammatikalische Erklärungen beziehen. Die Aufgabentypen sind: Single-Choice-Fragen, Lückentexte und offene Fragen.

Recht und Politik (RAG)



Für das Modul Recht und Politik verwenden wir das Retrieval Augmented Generation* (RAG) System. Diese Architektur extrahiert die für die Benutzeranfrage relevanten Informationen aus der Wissensdatenbank und generiert auf dieser Basis eine Antwort:

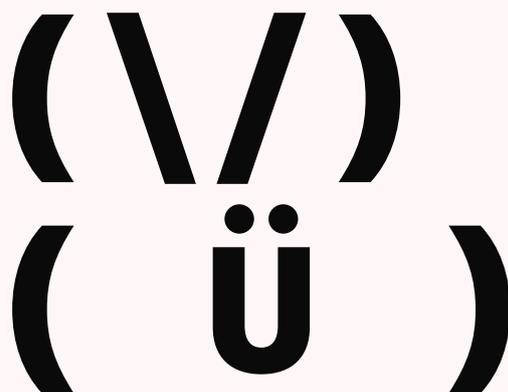
Retriever: Hier wird eine Vektor-Datenbank nach Informationen durchsucht, die am besten zur Anfrage des Benutzers passen (basierend auf der Ähnlichkeit von Vektoren). Wir haben dafür die milvus Datenbank mit multilingual embedding model BGE-M3 verwendet. Die Datenbank enthält manuell zusammengestellte Informationen aus Handreichungen, Beratungshilfen sowie verifizierte Informationen aus vertrauenswürdigen Webportalen wie basiswissen.asyl.net und einfach POLITIK.

Generator: Nach der Informationsbeschaffung verwendet das LLM die abgerufenen Informationen, um eine Antwort zu formulieren. Dabei soll sich das Modell ausschließlich auf die aus der Datenbank extrahierten Informationen stützen und stets die Quellen dieser Informationen angeben.

*P. Lewis et al., "Retrieval-Augmented Generation for Knowledge-Intensive NLP Tasks," arXiv.org, Apr. 12, 2021. <https://arxiv.org/abs/2005.11401>

Um unsere Lernapp zu verbessern, planen wir die Implementierung und Verbesserung mehrerer Schlüsselkomponenten:

- **Progress Tracker Modul:** Die Einführung eines erweiterten Progress Tracking Moduls wird es uns ermöglichen, den Lernfortschritt der Nutzer:innen genauer zu erfassen, um die Lerninhalte dynamisch anzupassen.
- **Verbesserung der Aufgabengenerierung:** Wir streben an, die Qualität der Aufgaben zu verbessern.
- **Generierung von Aufgaben auf Anfrage:** Die Entwicklung einer Funktion, die es den Nutzer:innen ermöglicht, spezifische Aufgaben generieren zu lassen - beispielsweise „Erstelle 5 Aufgaben zum Thema Perfekt mit ‚sein‘ für mich“.
- **Erweiterung des Curriculums:** Wir planen, unser Deutsch-Curriculum um zusätzliche Themen aus den Niveaustufen A2 und B1 zu erweitern.
- **Probe-Sprachtests:** Die Einführung von Probe-Sprachtests wird den Nutzer:innen helfen, sich auf offizielle Sprachprüfungen vorzubereiten und ihren Lernfortschritt zu evaluieren.
- **Erweiterung des Themenbereichs Recht und Politik:** Zusätzlich wollen wir unser Angebot an rechtlichen und politischen Inhalten erweitern, um mehr spezifische Inhalte für verschiedene Migrantengruppen anbieten zu können.



Wir würden uns sehr freuen, wenn wir mit Preisgeldern unterstützt werden könnten, um unsere App intensiver weiterzuentwickeln. Darüber hinaus sind wir sehr an einem Austausch mit den Mitgliedern des DINI-Netzwerks interessiert, um weitere Perspektiven und Expertise in unser Projekt einfließen zu lassen.

Wir freuen uns auf Ihr Feedback und hoffen auf eine positive Rückmeldung!