

Analyse der Potenziale von Künstlicher Intelligenz in Anerkennungs- und Anrechnungsprozessen

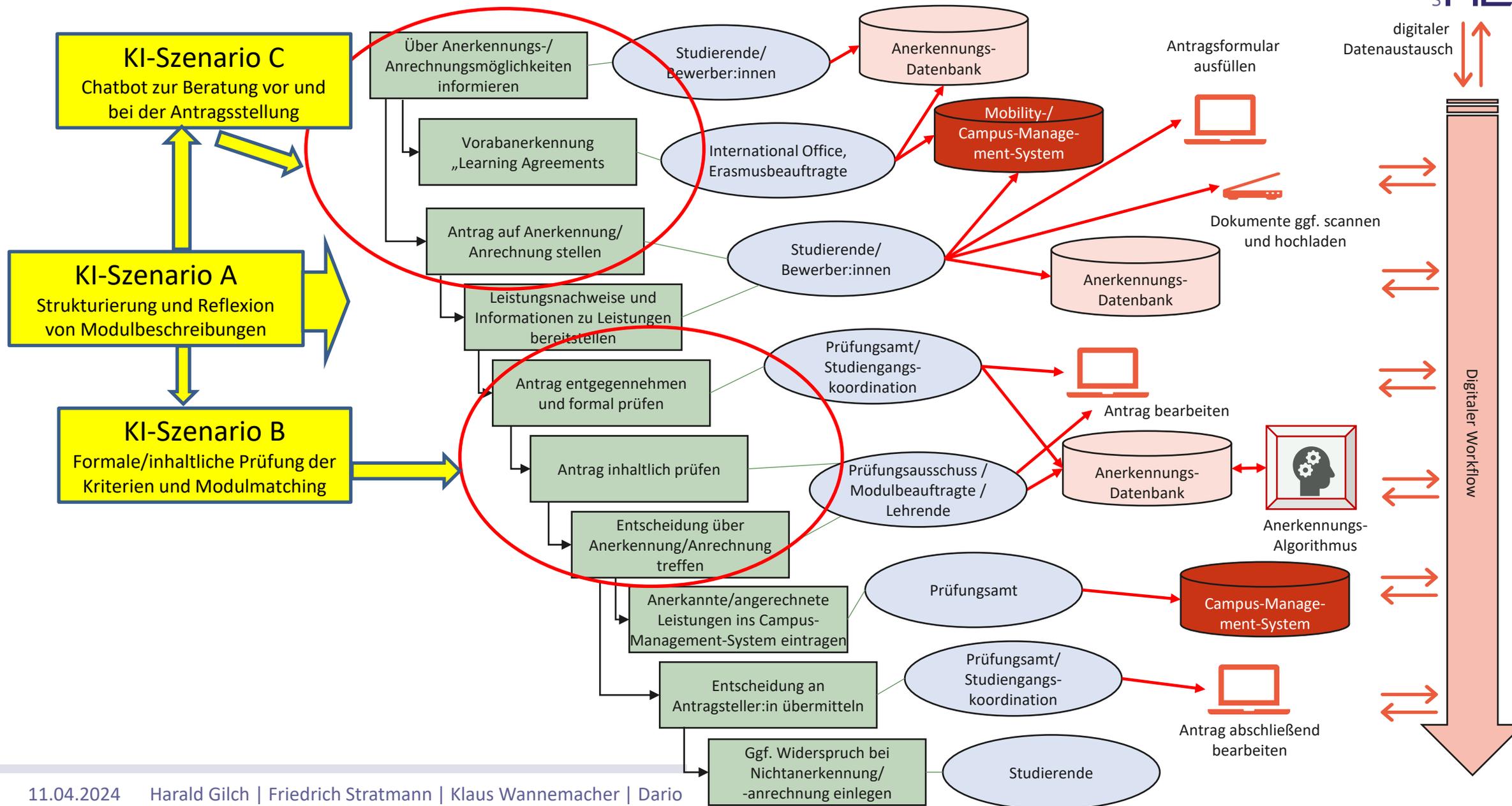
Vorstellung der Studie im Auftrag des HRK-Projekts MODUS

Digitalisierung weiterdenken – Künstliche Intelligenz in Lehrorganisation und Hochschulverwaltung, Karlsruhe, 10. und 11. April 2024

Harald Gilch | Friedrich Stratmann | Klaus Wannemacher | Dario Macey

Ziel der Studie ist es, für Verfahren der Anerkennung und Anrechnung studentischer Leistungen Unterstützungspotential durch den Einsatz von KI-Technologien zu ermitteln. Dafür hat die HRK aus dem Verfahrensablauf folgende drei Szenarien ausgewählt, die HIS-HE näher untersucht hat:

- **Szenario A:** Strukturierung und Reflexion von Modulbeschreibungen in Bezug auf Stringenz, Vollständigkeit und Kompetenzorientierung mit Hilfe von KI-basierten Tools.
- **Szenario B:** Unterstützung des akademischen Bereichs durch KI-basierte Tools bei der Prüfung inhaltlicher Kriterien als Matching von erbrachten Leistungen und Leistungsanforderungen für die Anrechnung bzw. Anerkennung studentischer Leistungen.
- **Szenario C:** Entwicklung eines Chatbots zur Unterstützung der Studierenden bei der Informationsbereitstellung bzw. der administrativen Bearbeitung der Hochschule im Rahmen von Anrechnungs- und Anerkennungsverfahren.



Ergebnisse der Expert:innen-Interviews und der Sichtung lfd. KI-Projekte zur Hochschulbildung

- (Gute) Allgemeinkenntnis über KI und KI-Tools, Erfahrung im Einsatz von generativer KI jedoch nur durch persönliche Nutzung; hier allerdings auch überrascht über die „Güte“ der Ergebnisse.
- Rollenkonflikt zwischen Hochschullehrer:in und IT-Expert:in in Bezug auf KI-Tauglichkeit in Anerkennungs- und Anrechnungsverfahren.
- Anerkennung und Anrechnung wird als komplexe Beurteilungssituation beschrieben, deshalb auch andere Lösungen wie Learning Agreements, Kooperationsvereinbarungen zwischen Hochschulen, Studienberatung in den Blick nehmen.
- Validität der Ergebnisse von generativer KI wird skeptisch eingeschätzt, da Halluzinationen entstehen, und das „Lösungsverfahren“ selbst eine „Blackbox“, Ergebnisse nur als Inspiration in der Entscheidungsvorbereitung nutzbar.
- Chance für KI-bezogene Chatbots in der Beratung bei geeignetem Datenkorpus
- *Lfd. KI-Projekte zur Hochschulbildung bieten Ansätze übergreifend als Technologietransfer und gegenstandsbezogen bei der Entwicklung von Chatbots und Textanalysen mittels KI*

Generative KI



z.B. ChatGPT+ (4.0)

- Verfahren basiert auf statistischen Korrelationen und Mustern.
- Training mittels großem Datenbestand von Lernmodellen (Generator, Diskriminator).
- Blackbox über „Entscheidungsregeln“, deshalb auch kein Wissen über konkrete Lösungswege.
- Optimierung der Ergebnisse durch Prompt-Engineering (Feintuning).
- Chat-Kommunikation in natürlicher (menschlicher) Sprache.

Expertensysteme (symbolische KI)

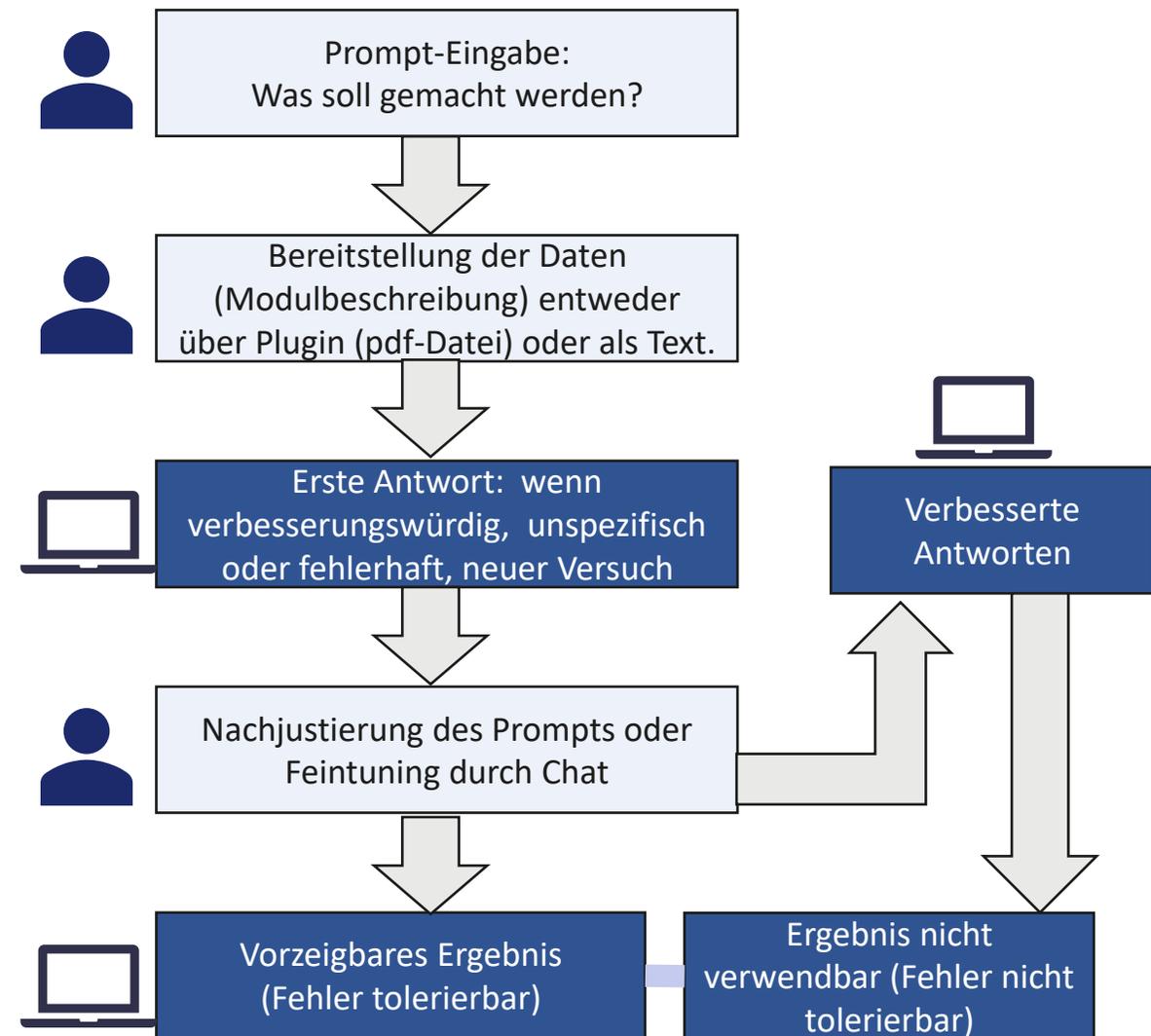


z. B. SemaLogic®

- Symbolische KI wird „händisch“ durch Feedback schlau gemacht.
- Festlegung von Entscheidungsregeln durch menschliche Expert:innen, deshalb Nachvollzug und Wissen über konkrete Lösungswege.
- Strukturierung (Regeln) benötigen Ontologien und Taxonomien (Cluster, Relationen)
- Eignet sich als Designmodell.
- Validierung der Repräsentation des Modells in Echtzeit.

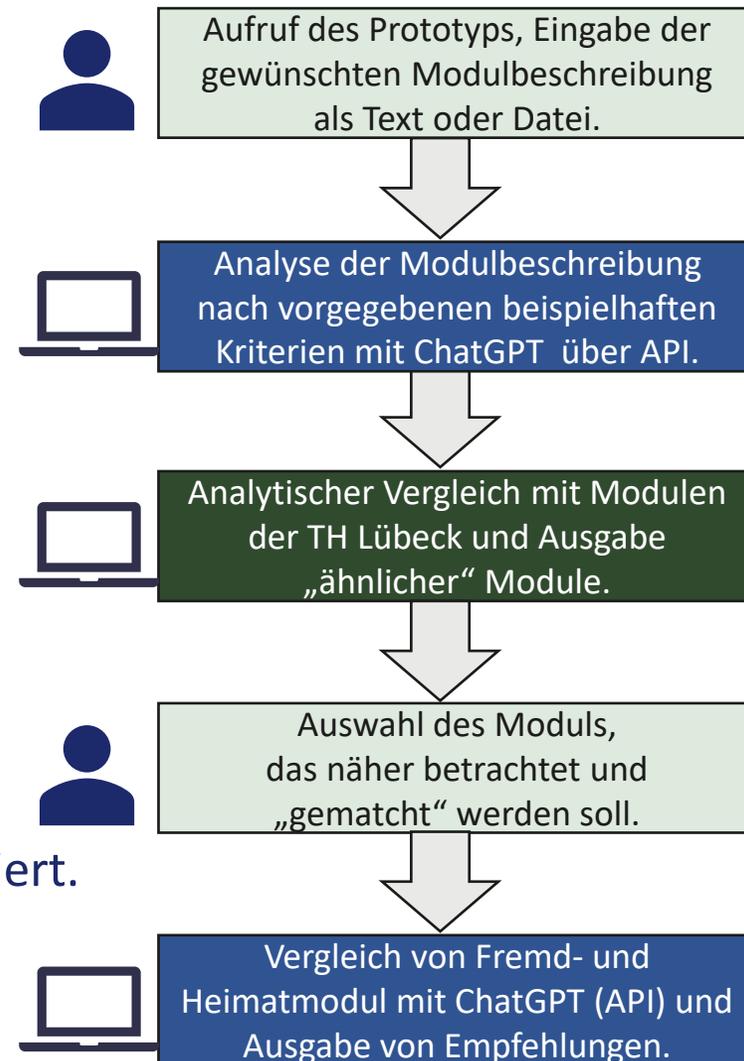
I. Proof of Concept Demonstration (ChatGPT+ 4.0) durch HIS-HE

- Problemstellung in Aufgabe (Prompt) übersetzt; Datenkorpus einschl. Zugriffsart ausgewählt.
- Art der Fragestellung an ChatGPT+ ist nicht eingeschränkt, deshalb Promptoptimierung und Feintuning durch Chat notwendig.
- Praxis: Um „gute“ Antworten zu erhalten, müssen die Fragen präzise gestellt werden und die Daten, auf die sich die Fragen beziehen, möglichst strukturiert und im Umfang begrenzt vorliegen.
- Im Dialog zwischen Mensch und ChatGPT auf der Basis natürlicher Sprache wird in Schleifen versucht, akzeptable Ergebnisse in Bezug auf die Problemstellung zu erreichen.
- Wenn dauerhaft Fehler und Halluzinationen die Validität der Ergebnisse beeinträchtigen, erweist sich der (eingesetzte) Prompt als ungeeignet.



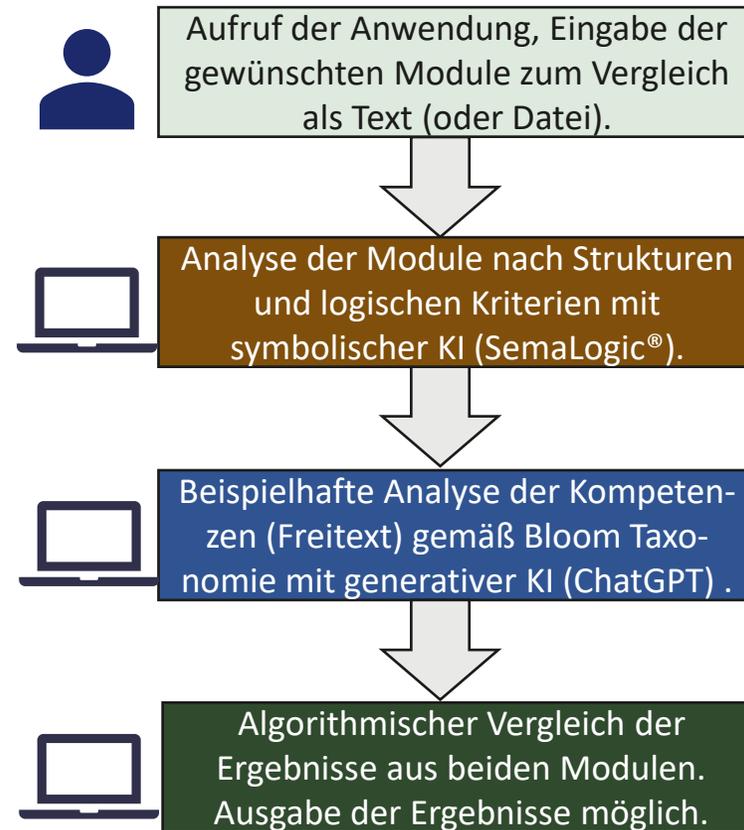
II. Proof of Concept Demonstration (Vektordatenbank + ChatGPT+ 4.0) durch TH Lübeck

- Standardisiertes Vorgehen, feste Eingabefelder, hinterlegte beispielhafte Kriterien und optimierte Prompts im Prototyp legen den Ablauf fest.
- Beliebige Modulbeschreibungen, unabhängig von der Sprache (z. B. auch Arabisch) wird mit dem kompletten Modulkatalog der TH Lübeck verglichen. Die Modulbeschreibung wird entweder als Text eingegeben oder als Datei eingelesen.
- Nach festen Kriterien analysiert der Prototyp mit Hilfe von ChatGPT die Modulbeschreibung und schreibt das Ergebnis als „Vektor“ in die Moduldatenbank.
- Das Fremdmodul wird mit den Modulen der TH Lübeck algorithmisch verglichen und ähnliche Module identifiziert.
- Für ein ausgewähltes Heimatmodul werden anschließend mit ChatGPT Übereinstimmungen und Abweichungen ermittelt und Empfehlungen ausgegeben.



III. Proof of Concept Demonstration (SemaLogic® + ChatGPT+ 4.0)

- Standardisiertes Vorgehen, symbolische KI zur Analyse von Modulbeschreibungen nach Strukturen und logischen Kriterien (Bestandteile, Credits, Workloads etc.).
- Zu vergleichende Module werden mit symbolischer KI strukturiert und gegenüber gestellt. Zur beispielhaften Analyse der Kompetenzen nach Bloom'scher Taxonomie werden die Freitextfelder an ChatGPT über API und optimierte Prompts übergeben.
- Die Ergebnisse für die zu vergleichenden Module werden anschließend algorithmisch verglichen. Gemeinsamkeiten und Abweichungen können als Ergebnisse des Modulvergleichs ausgegeben werden, sofern die Kriterien dafür definiert sind.
- Der Einsatz generativer KI beschränkt sich auf die Analyse von Freitextfeldern (z. B. für die zu erzielenden Kompetenzen). Sonst wird symbolische KI verwendet.



Ergebnisse: Gestaltungspotentiale mit LLM (z. B. ChatGPT+) in Anwendungsszenarien für Anerkennung und Anrechnung

■ Prüfung von Modulhandbüchern/-beschreibungen (Szenario A)

- Es liegen hier nur Tests von HIS-HE vor: Modulhandbuch einer Hochschule wurde mit den Anforderungen einer Schreibhilfe für Modulhandbücher (TU Darmstadt) verglichen. Insofern gelten die für den Modulvergleich eingeforderten Optimierungen beim Einsatz generativer KI (z. B. Vektordatenbank).
- Strukturierung von Kriterien und Inhalten gelingt von der Form her, allerdings nicht von den Inhalten. Auswahl der analysierten Module bleibt unbegründet und wechselt je nach Feintuning (Optimierung über Promptengineering und Abschichtung des Datenkorpus). Beurteilungsergebnisse bleiben vergleichsweise vage und entsprechen allenfalls manchem „Evaluierungs- und Akkreditierungs-Small Talk“.
- Die vom Umfang und der Beurteilungsintensität hohe Komplexität dieses Anwendungsszenarios wäre bei Weiterverfolgung ein gutes Beispiel, die Kombination von generativer KI und semantischer bzw. symbolischer KI (z. B. SemaLogic, AIStudyBuddy), zu erproben und Erfahrungen beim Einsatz von KI in der Studiengangsplanung zu nutzen.

Ergebnisse: Gestaltungspotentiale mit LLM (z. B. ChatGPT+) in Anwendungsszenarien für Anerkennung und Anrechnung

■ Modulvergleich/Matching (Anwendungsszenario B)

- Strukturierung von Modulvergleichen aus Hochschulen im Kontext von Anerkennung gelingt, wenn geeignetes Prompting (Ein- und Ausgabe, Steuerung) und LLM geeigneter Datenkorpus (Vektorisierung) vorliegen.
- Strukturierung von Vergleichen zwischen Ausbildungsordnungen und Hochschulmodulen gelingt für eine Anrechnung nicht, weil die Struktur der Dokumente zu unterschiedlich ist.
- Bewertung der Anerkennung als Entscheidungsunterstützung ist möglich, wenn Kriterien der Beurteilung transparent, nachvollziehbar und replizierbar sind (Gefahr des „Ankereffekts“ der KI für die formal noch durchzuführende menschliche Entscheidung).
- Generelle Anforderung: Programmtechnische Einhegung von Halluzinationen
- Als generelle Weiterentwicklung: **Einbindung eines optimierten Prototyps (Beispiel TH Lübeck) in den Anerkennungs- und Anrechnungsworkflow von PIM;** Überführung von Modulhandbüchern in eine vektorisierte Form; Datenkorpus für RAG oder eigenes LLM nutzen und mit Feedback- bzw. Memoryfunktionen ausstatten

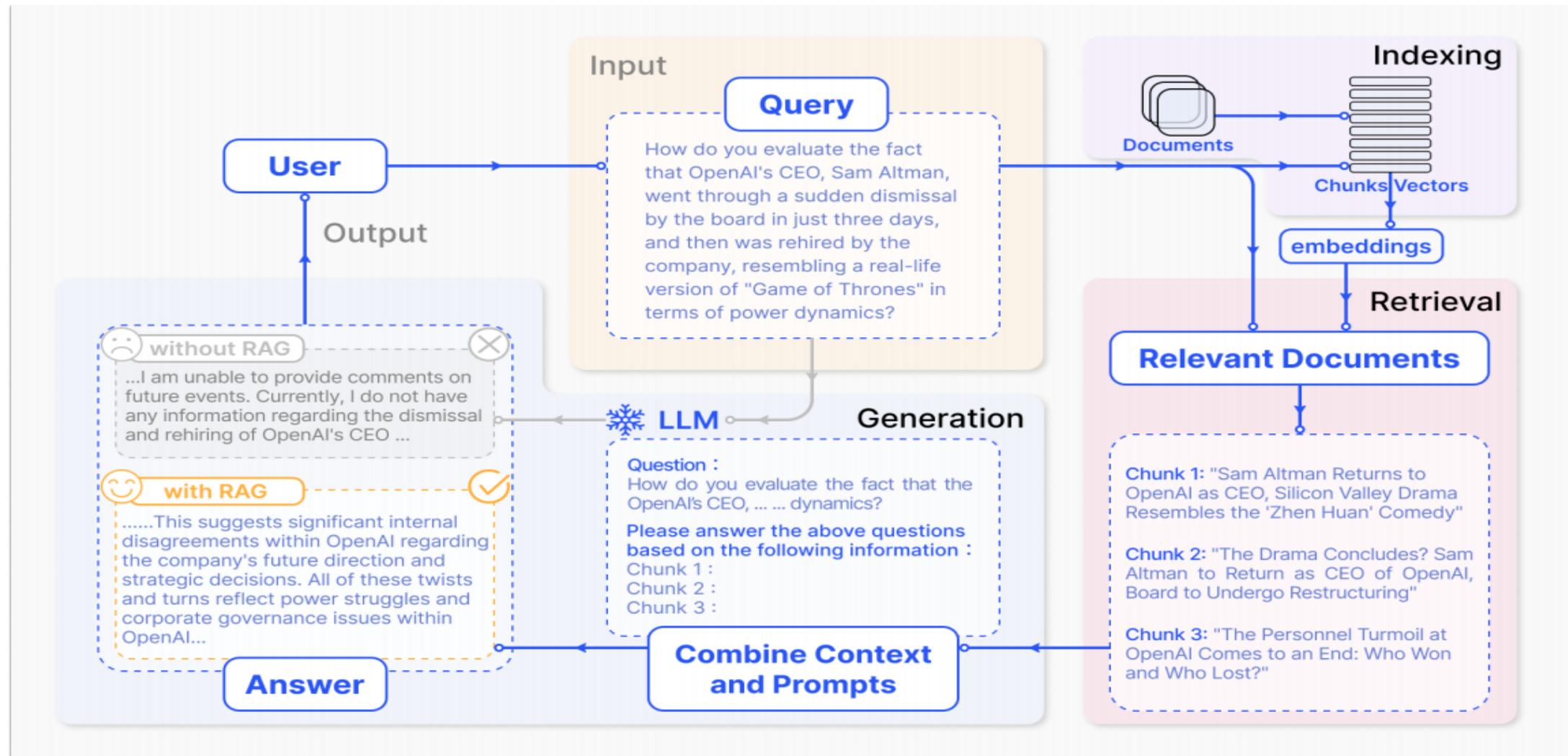
Ergebnisse: Gestaltungspotentiale mit LLM (z. B. ChatGPT+) in Anwendungsszenarien für Anerkennung und Anrechnung

■ Chatbot (Anwendungsszenario C)

- Strukturierung von allgemeinen Verfahrensbeschreibungen und Ausfüllhilfen für eine Hochschule ist möglich, wenn der Datenkorpus für generative KI vom Umfang bearbeitbar bleibt; Informationen behalten allerdings immer einen hohen Grad von Allgemeinheit.
- Den hohen Erwartungen, die zahlreiche Expert:innen und Interviewpartner:innen in die generative KI für künftig verlässlichere chatbot-basierte Beratung gegenüber HIS-HE geäußert haben, kann zumindest ChatGPT+ (noch) nicht einlösen. Allerdings zeigen KI-Projekte (RWTH Aachen, TU Berlin) erste Erfolge mit eigenen und speziell trainierten Systemen.
- Die Nutzung hochschuleigener KI für Beratungszwecke hat das Potential, die bisherige Form der Chatbots mit definierten Frage-/Antwort-Katalogen und Verweisen auf persönliche Hochschulberatung nicht nur dialogfreundlicher, sondern grundsätzlich individueller zu gestalten.
- Die Idee eines chatbot-basierten Auskunftssystems auf Grundlage einer vektorisierten, für Anfragen nutzbare Modul- oder Anerkennungs-/Anrechnungsdatenbank (z. B. ZeDoLa von HSBI) wäre zu prüfen.

- **KI-Chatbot:** Im Hochschulbereich sind KI-Chatbots im Rahmen der Studienberatung im Einsatz bzw. in der Entwicklung. Perspektive: in Kürze qualitativ hochwertige und verlässliche Anwendungen, die unterschiedliche KI-Tools (symbolische /generative KI, RAG) einsetzen.
- **KI-Modul-„Analyzer“:** KI-basierte Analyse von komplexen Texten nicht nur im Hochschulbereich, sondern „überall“ (s. LegalTech). „Automatisierte“ Analyse und Strukturierung von Modulen wäre Modellprojekt für Textanalysen mit KI im Hochschulbereich (Transferpotential für andere Anwendungen); Potential hätte auch ein **KI-Modul-„Designer“** bei der Erstellung von Modulbeschreibungen (siehe U Potsdam).
- **KI-Modul-„Matcher“:** Praktische Anwendung ist umstritten, betrifft sie doch das Selbstverständnis einer wissensbasierten Entscheidungsfindung durch „Mensch und Akademia“. Ein KI-Modul-„Matcher“ wird zumindest zur Entscheidungsvorbereitung genutzt werden können (s. Projekt PIM, Einbindung von eigener Wissensdatenbank (RAG))

RAG-Konzept: Verknüpfung von LLM und eigener Wissensdatenbank



Quelle: Gao, Yunfan et. al. (2023): Retrieval-Augmented Generation for Large Language Models: A Survey. arXiv:2312.10997 [cs.CL] Link: <http://arxiv.org/pdf/2312.10997v4>

Dr. Harald Gilch | Dr. Klaus Wannemacher | Dr. Friedrich Stratmann

HIS-Institut für Hochschulentwicklung e.V.
Goseriede 13a | D-30159 Hannover | www.his-he.de

+49 511 169929-32 | +49 174 8816073 | gilch@his-he.de
+49 511 169929-23 | +49 151 18731047 | wannemacher@his-he.de