

# Agentic Recommender System Konzept für nachhaltiges Wissensmanagement

## KI-befähigtes Wissensmanagement in Unternehmen

Simon Wilbers<sup>1,\*</sup> , Ron van de Sand<sup>2</sup> , Bastian Prell<sup>1</sup>  und Jörg Reiff-Stephan<sup>1</sup> 

<sup>1</sup>Technische Hochschule Wildau, Germany

<sup>2</sup>Causality GmbH, Germany

\*Korrespondenz: [simon.wilbers@th-wildau.de](mailto:simon.wilbers@th-wildau.de)

**Abstract.** Unternehmen stehen vor erheblichen Herausforderungen beim Erhalt, bei der Nutzung und Weitergabe zentralen Wissens. Demografischer Wandel, zunehmender Mitarbeiterfluktuation und steigende organisatorische Komplexität erschweren klassische Dokumentations- und Austauschprozesse, welche auf erhebliche Mitarbeitendenmitwirkung angewiesen sind. Mithilfe generativer KI und agentischer Recommender-Systeme lassen sich diese Hindernisse überwinden, indem wichtiges Wissen „on-the-fly“ und weitestgehend automatisiert erfasst wird. Ein zentrales Element bildet hierbei die Kombination aus semantischen Embeddings, einer Graphdatenbank und spezialisierten KI-Agenten, die Dokumente und Chatverläufe analysieren, Mitarbeiter gezielt mit Vorschlägen unterstützen und Wissen so kontinuierlich aktualisieren. So entsteht eine effektive Wissenskultur mit niedrigen Barrieren für die Nutzung dank minimalem Mehraufwand für die menschlichen Nutzer.

**Keywords:** Wissensmanagement, Graphdatenbank, Semantische Suche, Recommender Systeme

## 1. Einführung

Unternehmen sehen sich zunehmend mit der Aufgabe konfrontiert, wertvolles Wissen unter sich stetig ändernden Bedingungen erfolgreich zu bewahren, aufzubauen und weiterzugeben. Dieser Druck entsteht unter anderem durch demografische Veränderungen, bei denen erfahrene Wissensträger in den Ruhestand gehen und ihr Know-how nur unzureichend dokumentiert hinterlassen [1]. Zugleich wechseln Mitarbeitende häufiger den Arbeitsplatz, sodass fachliche Kompetenzen das Unternehmen verlassen, während neue Teammitglieder einen schnellen Zugang zu bereits vorhandenem Wissen benötigen [2], [3]. Hinzu kommt, dass die zunehmende Komplexität moderner Organisationsprozesse ein effizientes Wissensmanagement weiter erschwert, wenn Informationen nur sporadisch erfasst und geteilt werden [4]. Bis zu 73 Prozent, der Mitarbeitende, erkennen den Bedarf an Wissensmanagement-Tools, scheuen jedoch oft deren Nutzung aufgrund von Komplexität und Pflegeaufwand [3]. Gleichzeitig eröffnen Fortschritte bei der generativen KI neue Möglichkeiten, um Wissen in

Unternehmen effizienter zu erfassen und aufzubereiten. Automatische Textgenerierung und semantische Analysen können den Aufwand für die manuelle Dokumentation erheblich verringern und Barrieren für die Mitarbeitenden abbauen. Damit stellt sich die Frage, wie solch ein KI-gestütztes Wissensmanagement mit minimalem Zusatzaufwand in den betrieblichen Alltag integriert werden kann, sodass kein unverhältnismäßiger Dokumentationsaufwand, für die Mitarbeitenden entsteht [5], [6] und Chancen von *Knowledge Mining* genutzt werden. Knowledge Mining kombiniert Methoden aus dem Data-Mining und maschinellen Lernen, um wertvolle Informationen aus großen Datenmengen zu extrahieren [7], [8]. Ziel ist die Erstellung strukturierter Repräsentationen von Daten, die das Verständnis komplexer Zusammenhänge sowie die Entwicklung anwendungsorientierter Lösungen erleichtern [7]. Diese Arbeit adressiert diese Frage, indem ein Konzept für ein Agentic Recommender System vorgestellt wird. Ziel ist es, durch eine „On-the-fly“-Wissensdokumentation das Potenzial generativer KI so einzusetzen, dass Organisationswissen nicht nur erhalten, sondern ohne großen Mehraufwand laufend erweitert und nutzbar gemacht werden kann. Abbildung 2 zeigt schematisch einen möglichen KI-basierten Wissensmanagementprozess: Über Knowledge Mining werden unstrukturierte Daten analysiert, in Form von Embeddings in eine Graphdatenbank überführt und stehen anschließend über eine semantische Suche in einem Chatbot-basierten Interface zur Verfügung. Somit entsteht eine nutzerzentrierte, selbstlernende Umgebung, in der Mitarbeitende Informationen kontextbezogen abrufen und Schritt für Schritt erweitern können [9].

## 2. Stand der Technik

Bei klassischen, *dokumentenorientierten* Wissensmanagement-Systemen, wie sie in Form von Wikis, Intranets oder gängigen internen Datenbanken verbreitet sind, besteht ein erheblicher Pflegeaufwand. Mitarbeitende, abgeschreckt vom Aufwand, investieren selten Zeit, um Dokumentationen zu erstellen oder zu aktualisieren, was zu veralteten Beständen, mangelhafter Nutzerakzeptanz und sogenannten „Wissenssilos“ führt. Als Konsequenz bleibt wertvolles Unternehmenswissen häufig ungenutzt, während mit der Zeit neue Informationsanforderungen entstehen.

Die Nutzung generativer KI durch Office- und Knowledge-Worker nimmt stetig zu. Insbesondere Large Language Models (LLMs) wie ChatGPT werden für verschiedenste Aufgaben wie Inhaltserstellung, Informationssuche und Prozessautomatisierung eingesetzt [10], [11], [12]. Es zeigt sich, dass durch den Einsatz solcher Systeme signifikante Produktivitätssteigerungen erzielt werden können. So berichtet Noy [13], dass sich die Bearbeitungszeit für bestimmte Aufgaben um 40% verringert, während die Qualität der Ergebnisse um 18% steigt.

Ein besonders relevanter Vorteil dieser KI-gestützten Systeme liegt in ihrer Fähigkeit, repetitive Aufgaben zu übernehmen. Dies ermöglicht es den Nutzern, sich stärker auf kreative und strategische Tätigkeiten zu fokussieren [14], [15]. Ferner steigern LLMs die Motivation, ihrer menschlichen Nutzer sich mit Schreibaufgaben zu beschäftigen, indem sie als inspirierende und unterstützende Schreibpartner wahrgenommen werden [16]. Gleichzeitig ergeben sich jedoch Herausforderungen bei der Integration in bestehende Arbeitsprozesse. So kann sich Teamdynamik verändern und neue Kompetenzen erforderlich werden, um das Potenzial der Technologie vollständig auszuschöpfen [17].

Interessanterweise zeigt sich in der Wahrnehmung vieler Arbeitnehmer ein eher konservatives Verständnis der Technologie. Während einige Forscher transformative Effekte auf wissensbasierte Industrien prognostizieren, betrachten viele Nutzer KI nach

wie vor hauptsächlich als unterstützendes Werkzeug für einfache oder repetitive Aufgaben, das stets unter menschlicher Aufsicht bleibt [18]. Diese Dichotomie unterstreicht die Notwendigkeit einer gezielten Systemintegration, um sowohl die Produktivität als auch die Akzeptanz generativer KI im Unternehmenskontext zu steigern.

Recommender-Systeme im Wissenskontext könnten hier Abhilfe schaffen. Durch Social-Media-Plattformen sind automatisierte Vorschläge lange bekannt, etwa wenn ein System Content-basiert oder kollaborativ ähnliche Inhalte empfiehlt. Diese Ideen lassen sich auf organisationale Daten übertragen, um Mitarbeitenden zum richtigen Zeitpunkt passende Informationen anzubieten [19], [20]. In Unternehmen bedeutet dies beispielsweise, relevante Dokumente, Ansprechpersonen oder Prozesse zu empfehlen, sobald ein bestimmtes Thema in einer Arbeitsgruppe aktuell wird. Aktuelle Trends in generativer KI unterstreichen die Vorteile der Kombination von LLMs mit Graphdatenbanken. Hierdurch lassen sich semantische Beziehungen zwischen Inhalten abbilden, sodass eine kontextbezogene effektive Suche und Empfehlung möglich wird [21], [22]. Während einzelne Ansätze ChatGPT-artige Lösungen ins Unternehmen „importieren“, geht die *systemische Integration* einen Schritt weiter: Es entsteht ein *nahtloser Prozess*, bei dem LLMs zusammen mit semantischen Analysemodulen den Arbeitsalltag begleiten und *Wissensobjekte* laufend aktualisieren [23]. So werden nicht nur starre Dokumentenablagen besser organisiert, sondern ein aktiver Wissensaustausch gefördert, der sich an wechselnden Projektbedarfen und Nutzenden orientiert.

Bereits der alltägliche Gebrauch von LLM-Chatbots durch Mitarbeitende – seien es nun Recherchefragen, E-Mail-Entwürfe oder Prozessbeschreibungen – hat das Potential, in einem Agentic Recommender System (ARS) zu wertvollen Wissensbausteinen zu werden. Während ohne ARS jegliche Interaktion isoliert bleibt, erlaubt die gemeinsame Wissensbasis eine automatisierte und kontextbezogene Wiederverwendung von Prompts und bereits generierten Inhalten. Dadurch profitiert das gesamte Team von „digitalen Gedächtnisfunktionen“, die den Informationsfluss nachhaltig verbessern und den Dokumentationsaufwand für Einzelne merklich reduzieren. Mit dieser zentralisierten Sicht auf Wissen steigert das ARS die *Nachhaltigkeit* von betrieblichem Know-how und legt zugleich das Fundament für das *Integrationsmodell* im folgenden Kapitel.

### 3. Konzept

Das hier vorgestellte ARS Wissensmanagementkonzept erweitert klassische Retrieval-Augmented Generation (RAG)-Ansätze, indem es nicht nur einzelne Nutzeranfragen isoliert verarbeitet, sondern den gesamten Interaktionsprozess in eine adaptive Wissensstruktur einbettet. Das Ziel dieses Ansatzes besteht in der proaktiven Unterstützung von Nutzenden zur Aufwands-minimierten Beteiligung *On-the-fly* am Wissensmanagementprozess. Hieraus ergeben sich zwei Vorteile:

1. Durch das Vorschlagsystem verringert sich der Aufwand zur Pflege einer unternehmensweiten Wissensbasis für jedes Teammitglied und führt somit zu einer erhöhten Akzeptanz, was wiederum stärkere Nutzung und dadurch höhere Qualität des enthaltenen Wissens bedingt.
2. Durch den Einsatz eines Agentensystems kann die Knowledge-Management-Strategie eines Unternehmens zu wesentlichen Teilen durch gezieltes Prompt-Engineering implementiert werden.

Abbildung 1 illustriert das hier vorgestellte Prinzip. Im Gegensatz zu einem klassischen RAG System, mit einer unidirektionalen Anreicherung von Daten erfolgt,

ermöglichen die durch spezifisch angepasste Vorschlagsagenten eine bidirektionale Anreicherung von Daten, die die Nutzenden ins Zentrum der Datenkontrolle stellt. Die Grundidee basiert somit auf einem iterativen Prozess, der es den Nutzenden nicht nur ermöglicht, eine Datenbasis natürlichsprachlich zu konsumieren, sondern über konkrete Vorschläge diese auch mit zu gestalten und zu aktualisieren. Sowohl durch das System ausgeführte als auch vorgeschlagene, jedoch nicht umgesetzte Aktionen werden in das System zurückgeführt. Dies dient der gezielten und kontinuierlichen Verbesserung von Qualität und Relevanz. Auf diese Weise entsteht ein System, das sich iterativ verbessert, indem es nicht nur bestehende Inhalte abrufen, sondern auch neue Verknüpfungen zwischen Wissensseinheiten herstellt.

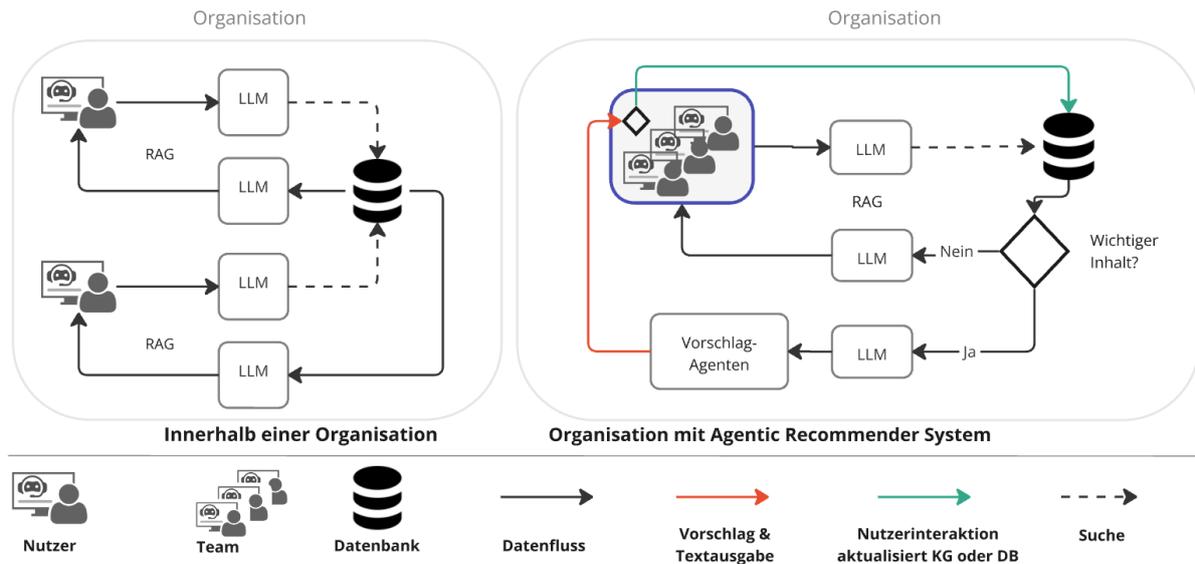


Abbildung 1. Klassisches RAG in Organisationen (links) versus Wissensmanagement mit ARS (rechts).

### 3.1 Grundprinzipien des Systems

Die Grundlage dieses Ansatzes bildet eine Kombination aus Knowledge Mining [7], [8], semantischen Beziehungen und einem Vorschlagsmechanismus, der darauf ausgelegt ist, den Nutzenden relevante Informationen kontextabhängig bereitzustellen. Der Prozess beginnt mit der Interaktion zwischen Mensch und System, in der Nutzerinnen und Nutzer beispielsweise Texte generieren, Suchanfragen stellen oder Dokumente bearbeiten. Basierend auf diesen Eingaben analysiert das System die enthaltenen Informationen, extrahiert semantisch relevante Inhalte und prüft deren Bezug zu bereits vorhandenem Wissen. Dabei kommt eine Entscheidungsfunktion **D** auf Grundlage eines *Knowledge Graphs*(KG) zum Einsatz, die bestimmt, ob neue Informationen für eine Speicherung oder zur Generierung von Vorschlägen geeignet sind. Falls eine hohe semantische Übereinstimmung mit bestehenden Wissensseinträgen vorliegt oder neue relevante Zusammenhänge erkannt werden, generiert das System einen Vorschlag für die Erweiterung der Wissensbasis. Die finale Entscheidung darüber, ob eine Information tatsächlich übernommen wird, verbleibt jedoch stets beim Menschen auch: *Human-in-the-Loop*. Dieses Vorgehen reduziert die Wahrscheinlichkeit, dass sensible oder fehlerhafte Inhalte automatisiert gespeichert werden. Auch dienen sie der Akzeptanz des Systems, da Nutzende stets die Kontrolle über ihre Daten behalten. Außerdem werden somit Regularien, wie z.B. DSGVO-Konformität und der zuletzt eingeführte AI Act innerhalb der Europäischen Union eingehalten.

### 3.2 Agentic Workflow

Wie in Abbildung 2 gezeigt, erfolgt der avisierte Agentic Workflow nach Abschluss der Inhaltserstellung (Text, Programmcode, etc.). Hier wird in einem Entscheidungsknoten **D** geprüft, ob die jeweiligen Inhalte für eine Wissensspeicherung oder weitere Empfehlungen in Frage kommen. Dieser Ansatz basiert auf vergangenen Interaktionen, bzw. gewählten Aktionen, sowohl auf Nutzerebene als auch auf Team-Ebene. Dazu werden Ähnlichkeitsmaße und Wahrscheinlichkeiten berechnet. Teilweise kommen auch zufallsbasierte Verfahren zum Einsatz, um das Modell zu diversifizieren und einseitige Lerneffekte zu vermeiden. Dieser Kompromiss wird, in der wissenschaftlichen Literatur, oft diskutiert als *Trade-off between Exploration und Exploitation*, also dem Spannungsfeld zwischen Erkundung und Nutzung [24].

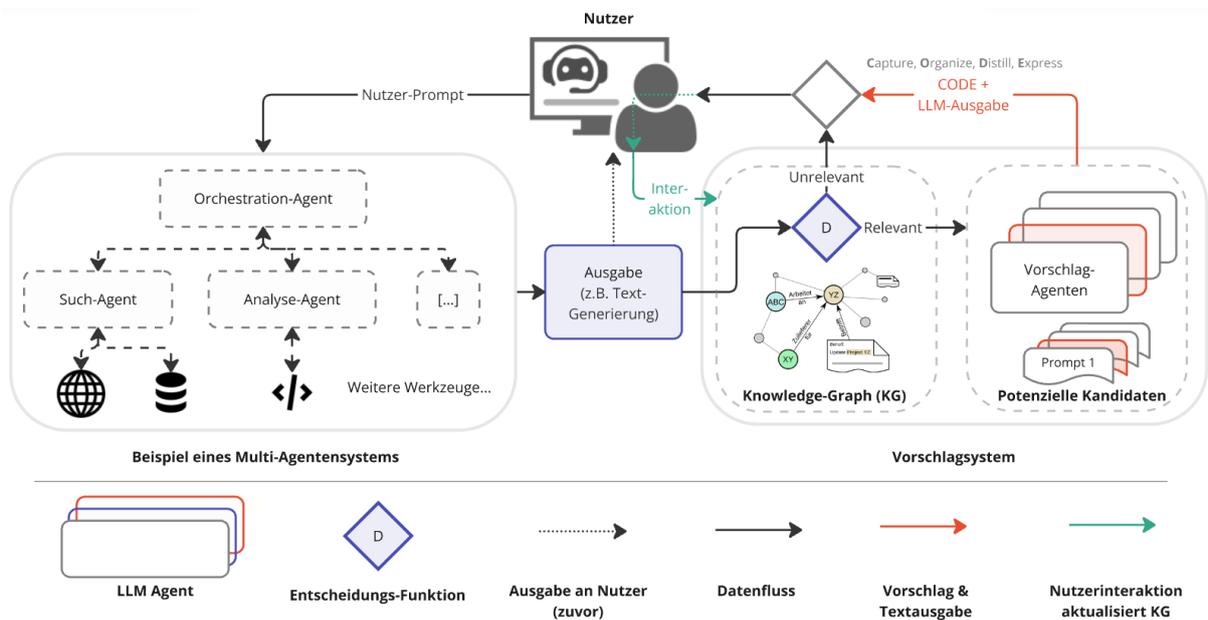


Abbildung 2. Agentic Recommender System

Durch die Implementierung entsprechender Metriken können potenzielle Kandidaten für Vorschlagsagenten identifiziert werden, die anschließend konkrete Vorschläge generieren. In dieser Generierungsphase werden relevante Informationen, einschließlich der berechneten Wahrscheinlichkeiten, als Prompt in das Modell eingespeist (Grounding), um ein fundiertes Ergebnis bereitzustellen. Die vom Nutzer gewählte Aktion wird wiederum in den KG zurückgeführt und optimiert kontextbasiert zukünftige Vorschläge. Das vorgestellte System bildet einen iterativen Wissensverarbeitungszyklus ab und orientiert sich dabei am CODE-Prinzip [25]: **C**apture, **O**rganize, **D**istill, **E**xpress.

### 4. Implementierung

Die Umsetzung des beschriebenen Systems lässt sich in die folgenden Schritte unterteilen:

1. Bestimmung der semantischen Ähnlichkeit zu vorherigen Kontexten (z.B. Chatverläufe, Dokumente) und Extraktion relevanter Textkandidaten.
2. Identifikation geeigneter Vorschlagsagenten und zugehöriger Aktionen.
3. Erzeugung von Vorschlägen durch spezifisch angepasste Vorschlagsagenten.

4. Überprüfung durch den Nutzer mit der Möglichkeit zur Annahme oder Ablehnung der Vorschläge.
5. Aktualisierung des KGs zur kontinuierlichen Optimierung zukünftiger Vorschläge.

Hierbei sei anzumerken, dass sowohl die Vorschlagsagenten als auch die zugehörigen Aktionen auf die spezifischen Nutzendenprozesse abgestimmt sind. Darüber hinaus können Aktionen sowohl intern (z.B. Aktualisierung einer Datenbank oder eines Dokuments) als auch extern (z.B. Eintrag in ein CRM-System) ausgeführt werden. Aktionen werden außerdem vorbestimmt, um die Schritte bei der Datenverarbeitung durch das ARS zu lenken.

Das hier beschriebene Konzept lässt sich also als Klassifizierungsproblem ableiten und ist in der Literatur als *Link-Prediction-Task*, also Kanten-Vorhersage-Aufgabe, bekannt.

### 4.1 Embeddings

Für die Implementierung der Entscheidungsfunktion werden zwei Arten von Embeddings (siehe Abbildung 3), also eine semantisch-vektorielle Abbildung von Daten genutzt. In dieser Arbeit werden zwei Arten genutzt und unterschieden:

1. Text-Embeddings: Semantische Beziehungen werden aus dem reinen Text-Kontext, z.B. durch Transformer-Modelle [26], extrahiert.
2. Graph-Embeddings: Beziehungen werden aus der Topologie des KGs extrahiert. Die Erzeugung dieser Embeddings erfolgt oftmals induktiv, z.B. durch sogenannte Random-Walk-Algorithmen [27] oder aber Random Projection-Techniken [28].

Durch die Kombination beider Ansätze können so neue, bisher nicht zugeordnete Daten klassifiziert werden.

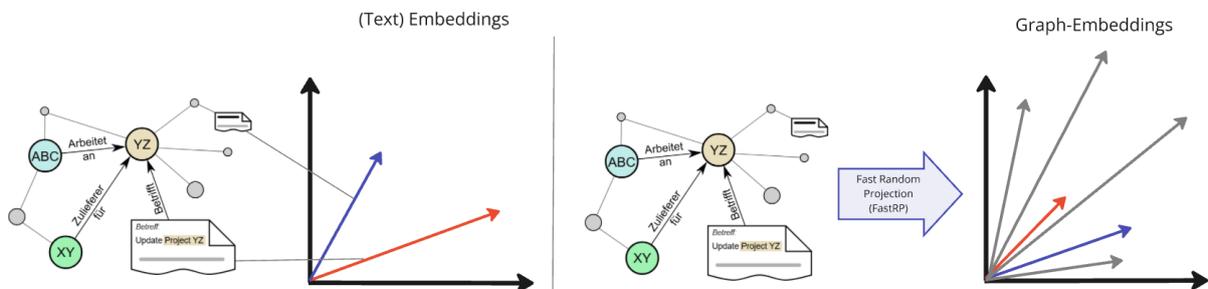


Abbildung 3. Text- und Graph-Embeddings

### 4.2 Vorschlagsgenerierung

Wie in Abbildung 4 veranschaulicht, ermöglicht die Kombination von Text- und Graph-Embeddings eine Ähnlichkeitsanalyse sowohl auf rein textbasierter als auch auf relevanzbasierter Ebene. Letztere nutzt die Topologie des KG sowie vergangene Interaktionen, um kontextbezogene Aktionskandidaten und geeignete Vorschlagsagenten zu identifizieren. Dennoch spielen Text-Embeddings eine zentrale Rolle, da neue Konversationen – abgesehen von der Nutzendenzuordnung – zunächst keine bestehenden Beziehungen im KG aufweisen und daher semantisch zugeordnet werden müssen. Die Erzeugung konkreter Vorschläge nach CODE [25] und deren Interaktion durch den Nutzenden führt zunehmend zur Optimierung des KG und somit zur

Entscheidungsfunktion **D** und dient einerseits zur Verbesserung der Nutzendenerfahrung und andererseits zur besseren und zielgerichteten Wissensverarbeitung im Unternehmen.

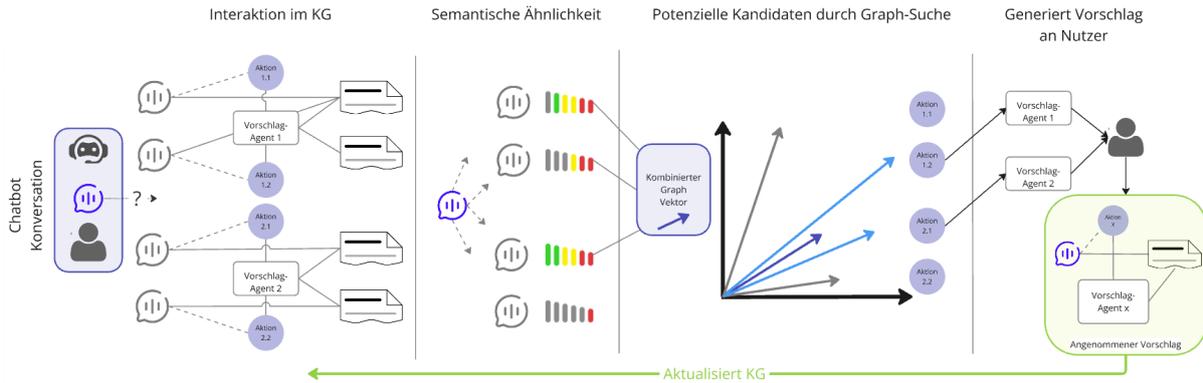


Abbildung 4. Prozessdarstellung der Aktualisierung des Knowledge-Graphs(KG)

### 4.3 Aktualisierung des Knowledge-Graphs

Bei der Umsetzung des Systems, insbesondere bei der Aktualisierung des KG, spielt die zeitliche Abfolge der Nutzerinteraktionen eine entscheidende Rolle. Es wird davon ausgegangen, dass ältere Interaktionen eine abnehmende Bedeutung im Vorschlagssystem haben, während neuere Interaktionen höher gewichtet werden.

Um diese Dynamik abzubilden, können bestehende Verbindungen im KG mit einer gewichteten Abklingrate versehen werden. Die Relevanz einer Verbindung *link* zum Zeitpunkt *t* kann durch die folgende Exponentialfunktion beschrieben werden:

$$Relevance(link, t) = Relevance(link, 0) \times e^{-\lambda \cdot t} \tag{1}$$

wobei:

*Relevance(link, 0)* den Ausgangswert der Relevanz unmittelbar nach der Erstellung der Verbindung bezeichnet,

*t* die vergangene Zeit seit der Erstellung (z.B. in Tagen) repräsentiert und

$\lambda$  eine Konstante ist, die den Abfall der Relevanz pro Zeiteinheit steuert.

Durch diese Gewichtung wird sichergestellt, dass aktuelle Nutzerinteraktionen einen stärkeren Einfluss auf zukünftige Vorschläge haben, während ältere Interaktionen allmählich an Bedeutung verlieren. Für bestehende Verbindungen werden zusätzlich Time-to-Live-Mechanismen eingeführt, um veraltete Vorgänge zu entfernen. Die Erzeugung der Graph-Embeddings erfolgt in voreingestellten Zeitabständen (z.B. täglich) um neue topologische Veränderungen innerhalb des KG adäquat vektorieLL abbilden zu können.

### 4.4 Entscheidungsfunktion

Durch fortlaufende Aktualisierungen des KGs sowie den Einsatz von Embeddings wird das System kontinuierlich erweitert. Jeder Interaktionspunkt

$$\langle u_t, c_t, a_t \rangle \tag{2}$$

wobei

$u_t$  einen Benutzer zum Zeitpunkt  $t$  beschreibt,

$c_t$  einen Inhalt repräsentiert und

$a_t$  eine gewählte Aktion angibt.

geht in die Entscheidungsfunktion ein und wird wie folgt zusammengefasst:

$$X = \{(u_t, c_t, a_t)\}_{t=1}^T, \quad (3)$$

Für den nachfolgenden Schritt  $(u_{T+1}, c_{T+1})$  bestimmt das System potentielle Kandidaten, für die Aktionen  $a_{T+1}$  mit assoziierten Wahrscheinlichkeiten zur Auslösung der Aktion. Hierbei werden sowohl Text-Embeddings  $E_c(c)$  als auch Graph-Embeddings  $E_g(n)$  (wobei  $n$  beliebige Knoten im KG beschreibt) genutzt, die zum Beispiel durch FastRP erzeugt werden können [28].

#### 4.5 Bestimmung geeigneter Aktionskandidaten

Um für einen neuen Inhalt  $(u_{T+1}, c_{T+1})$  passende Aktionen zu ermitteln, wird ein zweistufiger Prozess durchlaufen:

1. **Semantische Kandidatenauswahl:** Zunächst wird anhand der *cosine Ähnlichkeit* zwischen  $E_c(c_{T+1})$  und  $E_c(c_i)$  für bereits vorhandene Inhalte  $c_i$  eine Rangliste gebildet. Anschließend werden die  $K$  Inhalte

$$\{c_i \mid i = 1, \dots, K\} \quad (4)$$

ausgewählt, die dem neuen Inhalt  $c_{T+1}$  semantisch am ähnlichsten sind.

2. **Erzeugung eines Graph-Embedding-Prototyps:** Für jedes der  $K$  Inhalte  $c_i$  ist im KG ein zugehöriger Knoten  $n_i$  hinterlegt, dessen Embedding  $E_g(n_i)$  bekannt ist. Um den Kontext dieser  $K$  Inhalte im Graphen zu repräsentieren, wird ein *Durchschnittsvektor* im Graph-Embedding-Raum gebildet:

$$E_g(n_{\bar{c}}) = \frac{1}{K} \sum_{i=1}^K E_g(n_i). \quad (5)$$

Dieser *fiktive Knoten*  $n_{\bar{c}}$  spiegelt die kombinierte graphbasierte Repräsentation der am besten passenden Inhalte wider.

#### Ranking und Auswahl der Aktionen

Um aus dem KG relevante Aktionen (und Vorschlagsagenten) abzuleiten, werden nun sämtliche Aktionskandidaten  $(a_i)$  im Graphen gesichtet. Jeder Kandidat ist einem Knoten  $n_{a_i}$  zugeordnet, für den ein Embedding  $E_g(n_{a_i})$  existiert. Die Relevanz dieses Knotens im Kontext von  $E_g(n_{\bar{c}})$  wird über die *cosine Ähnlichkeit* bestimmt:

$$\cos(E_g(n_{\bar{c}}), E_g(n_{a_i})) \quad (6)$$

Basierend auf diesen Werten lassen sich die Aktionskandidaten in absteigender Reihenfolge ihrer Relevanz sortieren. Üblicherweise werden nur die  $K$  Kandidaten (z. B.  $K = 3$ ) als Vorschläge präsentiert. Überschreitet ein Kandidat zudem einen bestimmten Schwellenwert, kann er als besonders relevanter Vorschlag markiert werden.

## Einbettung in die Entscheidungsfunktion

Da jeder Inhalt  $c_i$  genau einem Agenten zugeordnet wird, entsteht so eine Liste konkreter Handlungsempfehlungen (jeweils bestehend aus *Inhalt* und *Aktion*), aus denen der Nutzer eine Option auswählen kann. Die gewählte Aktion wird anschließend in die Wissensbasis (z. B. Wiki, CRM, etc.) übertragen und dort verknüpft, sodass sich der KG sowie die Embeddings durch diesen neuen Interaktionspunkt

$$(u_{T+1}, c_{T+1}, a_{T+1})$$

kontinuierlich weiterentwickeln.

## 5. Diskussion

Eine zentrale Herausforderung bei der Einführung des ARS wird darin liegen, Mitarbeitende von dessen Mehrwert zu überzeugen und den erforderlichen Kulturwandel zu gestalten. Würde das System Inhalte gänzlich automatisch erfassen und empfehlen, müssten Datenschutz und sensible Informationen noch stärker beachtet werden. Durch die Anwendung des Human-in-the-loop-prinzips, wird dem Rechner getragen. Auch um eine Überautomatisierung zu vermeiden, bleibt der Mensch wichtig: Wissen zur manuellen Durchführung des Knowledge Minings muss erhalten bleiben, damit Beschäftigte bei Bedarf aktiv eingreifen und effektiv Korrekturen vornehmen können. Allerdings wird die automatisierte Erfassung von Wissen auch zu einer gewissen Abhängigkeit führen: Greift die KI zu stark ein, könnten Beschäftigte Kompetenzen verlieren, die bislang zum manuellen Dokumentieren oder zum Überblick über komplexe Prozesse nötig waren. Diese „Ironie der Automatisierung“ [29] verlangt, dass das System keinesfalls als Ersatz für das Fachwissen der Mitarbeitenden verstanden wird. Stattdessen sollten Human-in-the-Loop-Prinzipien etabliert sein, sodass stets der menschliche Nutzer, seine Fähigkeit erhält, kritisch prüfen und notfalls eingreifen zu können. Beispielsweise wenn das automatische System ausfällt oder niedagewesene Situationen auftreten. Langfristig ist somit ein ausgewogenes Zusammenspiel von Automation und menschlichem Know-how entscheidend, um den Nutzen des Systems zu maximieren, ohne essenzielle Fähigkeiten im Unternehmen zu verlieren. Gelingt diese Balance zwischen Automatisierung und menschlicher Kontrolle, kann sich das System langfristig zu einem selbstlernenden Bestandteil des betrieblichen Alltags entwickeln. Perspektivisch könnten spezialisierte Integrationen in Kollaborationstools oder CRM-Systeme den Nutzen weiter steigern und das Wissensmanagement als elementare Ressource im digitalen Transformationsprozess etablieren.

## 6. Fazit

Vor dem Hintergrund des demografischen Wandels, der ein effizienteres Wissensmanagement nahezu erzwingt, zeigt sich das Potenzial eines Agentic Recommender Systems. Insbesondere dann, wenn es gelingt, den manuellen Dokumentationsaufwand auf ein Minimum zu reduzieren. Sowie Nutzende einen zusätzlichen Mehrwert bemerkten, etwa die automatische Erstellung relevanter Textbausteine. Dies würde die Akzeptanz von Wissensmanagement im Arbeitsalltag fördern und eine lebendige Wissenskultur in Organisationen ermöglichen. Forschungsbedarf besteht in der Weiterentwicklung semantischer Embeddings für spezifische Branchen und in der Analyse von Human Factors im Einsatz KI-basierter Wissensassistenten. Langfristig ließe sich, durch Mensch-Maschine-Kollaboration, ein sehr effektives und nachhaltigeres Wissensmanagement für Organisationen realisieren.

## Author contributions

Die vorgestellten Ergebnisse sind das Resultat der Entwicklungsarbeit in der Forschungsgruppe iC3@Smart Production, sowie der Causality GmbH. Der Artikel wurde verfasst, um die gewonnenen Erkenntnisse so zu vermitteln, dass sie in praktischen Anwendungen genutzt werden können.

Simon Wilbers: Verfassen – Originalentwurf, Konzeption, Methodik, Untersuchung.  
Ron van de Sand: Konzeption, Methodik, Visualisierung, Formale Analyse, Validierung.  
Bastian Prell: Verfassen – Überprüfung & Bearbeitung, Konzeption, Betreuung.  
Jörg Reiff-Stephan: Projektadministration, Verfassen – Überprüfung & Bearbeitung, Finanzierungseinwerbung, Betreuung.

## Competing interests

Alle Autoren erklären, dass kein Interessenkonflikt besteht.

## Danksagung

Unser Dank gilt, der Technischen Hochschule Wildau und dem Bundesministeriums für Wirtschaft und Klimaschutz für die Unterstützung bei der Verwirklichung dieser Vision mit dem EXIST-Gründungsstipendium.

## Referenzen

- [1] E. Latniak and J. Hentrich, "Herausforderungen des demografischen wandels für fertige betriebe und deren beschäftigte", in Springer, 2013, pp. 9–23.
- [2] E. Commission, *The impact of demographic change—in a changing environment*, 2023.
- [3] K. D. Solutions, *Wissensmanagement im mittelstand: Mit dokumentenmanagement wissen besser verfügbar machen*, Whitepaper, 2018.
- [4] G. Schuh and V. Stich, *Enterprise-Integration: auf dem Weg zum kollaborativen Unternehmen*. Springer, 2014.
- [5] S. I. ein Sektor mit Zukunft, "Informationswirtschaft: Ein sektor mit zukunft: Symposium; 4.-5. september 2003, wien, österreich", Ges. für Informatik, 2003.
- [6] M. John, S. Schmidt, and B. Decker, "Community-management in unternehmen mit wiki- und weblogtechnologien", in *Gemeinschaften in Neuen Medien*, 2005.
- [7] Y. Rui et al., "Knowledge mining: A cross-disciplinary survey", *Machine Intelligence Research*, vol. 19, no. 2, pp. 89–114, 2022.
- [8] P. Markellou, M. Rigou, and S. Sirmakessis, "Knowledge Mining: A Quantitative Synthesis of Research Results and Findings", in *Studies in Fuzziness and Soft Computing*. Springer Berlin Heidelberg, 2005, pp. 1–11, ISBN: 9783540250708.
- [9] W. "Bünnagel, "Angewandte künstliche intelligenz", in *Künstliche Intelligenz und Unternehmenswissen: Betriebliches Wissensmanagement auf morgen ausrichten*. Springer Berlin Heidelberg, 2024, pp. 179–212, ISBN: 978-3-662-68779-6.
- [10] M. Brachman, A. El-Ashry, C. Dugan, and W. Geyer, "How knowledge workers use and want to use llms in an enterprise context", in *Extended Abstracts of the CHI Conference on Human Factors in Computing Systems*, 2024, pp. 1–8.
- [11] P. Ritala, M. Ruokonen, and L. Ramaul, "Transforming boundaries: How does chatgpt change knowledge work?", *Journal of Business Strategy*, vol. 45, no. 3, pp. 214–220, 2023.
- [12] J. Retkowsky, E. Hafermalz, and M. Huysman, "Managing a chatgpt-empowered workforce: Understanding its affordances and side effects", *Business Horizons*, vol. 67, no. 5, pp. 511–523, 2024.

- [13] N. Shakked and Z. Whitney, "Experimental evidence on the productivity effects of generative artificial intelligence", *Science*, 2023.
- [14] R. Ulfsnes, N. B. Moe, V. Stray, and M. Skarpen, "Transforming Software Development with Generative AI: Empirical Insights on Collaboration and Workflow", *arXiv.org*, 2024.
- [15] R. Laavanya, R. P., and R. Mika, "Creational and conversational AI affordances: How the new breed of chatbots are revolutionizing the knowledge industries", *Business Horizons*, 2024.
- [16] S. Wilbers, J. Gröpler, B. Prell, and J. Reiff-Stephan, "Overall writing effectiveness: Exploring students' use of llms, pushing the limits of automated text generation", in *International Conference on Smart Technologies & Education*, Springer, 2024, pp. 11–22.
- [17] D. Daniel-Florin and T. Emanuela, "Knowledge Workers and the Rise of Artificial Intelligence: Navigating New Challenges", *SEA - Practical Application of Science*, 2024.
- [18] A. Woodruff, R. Shelby, P. G. Kelley, S. Rousso-Schindler, J. Smith-Loud, and L. Wilcox, "How knowledge workers think generative ai will (not) transform their industries", in *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems*, 2024, pp. 1–26.
- [19] I. Guy, N. Zwerdling, I. Ronen, D. Carmel, and E. Uziel, "Social media recommendation based on people and tags", in *Proceedings of the 33rd International ACM SIGIR Conference on Research and Development in Information Retrieval*, 2010.
- [20] I. Ronen, I. Guy, E. Kravi, and M. Barnea, "Recommending social media content to community owners", in *Proceedings of the 37th International ACM SIGIR Conference on Research and Development in Information Retrieval*, 2014.
- [21] L. Chen et al., "Collaboration-aware hybrid learning for knowledge development prediction", in *The Web Conference*, 2024.
- [22] K. Christidis, G. Mentzas, and D. Apostolou, "Using latent topics to enhance search and recommendation in enterprise social software", *Expert Systems with Applications*, vol. 39, no. 10, pp. 9297–9307, 2012.
- [23] C.-H. Lai and D.-R. Liu, "Integrating knowledge flow mining and collaborative filtering to support document recommendation", *Journal of Systems and Software*, vol. 82, no. 12, pp. 2023–2037, 2009.
- [24] M. Karimzadehgan and C. Zhai, "A learning approach to optimizing exploration–exploitation tradeoff in relevance feedback", *Information retrieval*, vol. 16, pp. 307–330, 2013.
- [25] T. Forte, *Building a second brain: A proven method to organize your digital life and unlock your creative potential*. Simon and Schuster, 2022.
- [26] N. Reimers and I. Gurevych, "Sentence-bert: Sentence embeddings using siamese bert-networks", *arXiv preprint arXiv:1908.10084*, 2019.
- [27] F. Xia, J. Liu, H. Nie, Y. Fu, L. Wan, and X. Kong, "Random walks: A review of algorithms and applications", *IEEE Transactions on Emerging Topics in Computational Intelligence*, vol. 4, no. 2, pp. 95–107, 2019.
- [28] H. Chen, S. F. Sultan, Y. Tian, M. Chen, and S. Skiena, *Fast and accurate network embeddings via very sparse random projection*, 2019. arXiv: [1908.11512 \[cs.SI\]](https://arxiv.org/abs/1908.11512).
- [29] L. Bainbridge, "Ironies of automation", in *Analysis, design and evaluation of man–machine systems*, Elsevier, 1983, pp. 129–135.