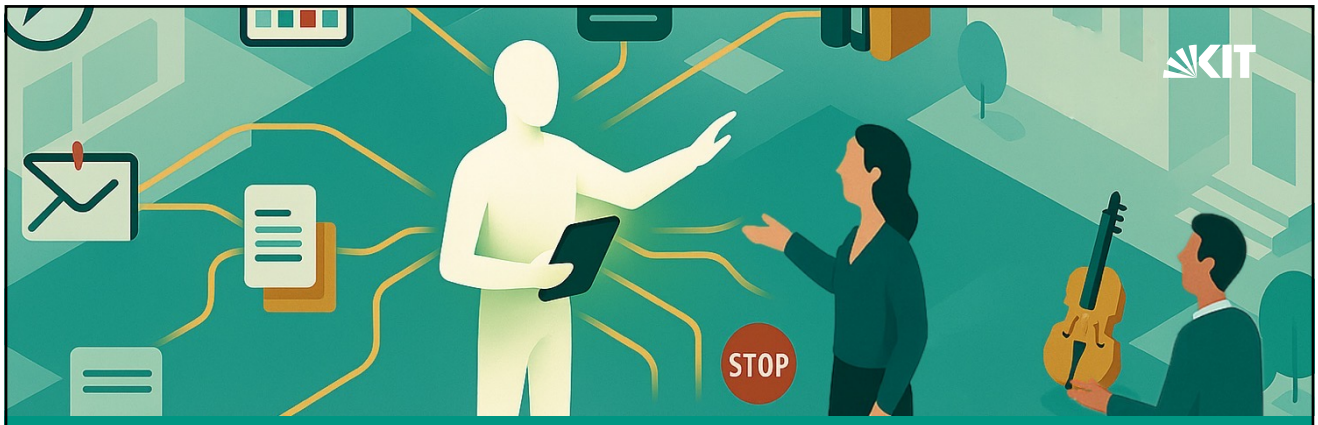
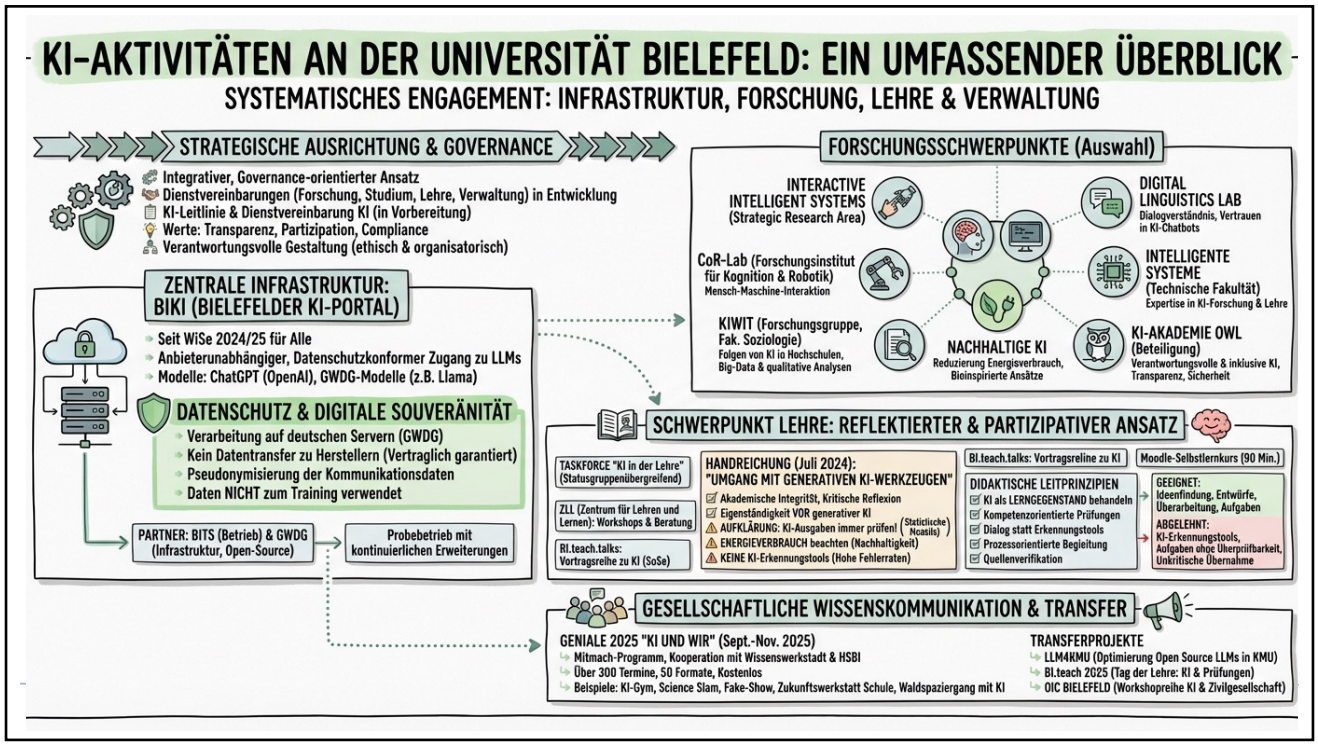


KI-Agenten in der Hochschulpraxis: Von der Potenzialanalyse zur verantwortungsvollen Implementierung

Workshop Universität Bielefeld 11.12.2025

Andreas Sexauer (KIT)

1



2

1. Was macht einen Agenten zum Agent?
2. Prototypische Use-Cases
3. Eigene Use-Cases diskutieren
4. Transformation, Governance, Risiko

3

4

Das Zentrum für Mediales Lernen

5

Das Zentrum für Mediales Lernen am KIT

- Umfangreiche Expertise im Bereich E-Learning und digitale Lehre
- Koordinierungsstelle für digitale Lehre am KIT und Verortung der Geschäftsstelle des HND-BW
- Produzent von multimedialen Inhalten zur Wissenschaftskommunikation
- Projektpartner: KIT-intern, national und international

6 04.12.25

6

Innovationsraum für generative KI am KIT – GenAI@KIT

Entwicklungsrahmen für die Nutzung generativer KI am KIT	Definition eines Serviceangebots für den KI-Einsatz am KIT	Qualifizierungs-angebote und Netzwerke für relevante Akteure	Kontinuierliche Kommunikation ins KIT

7

9

AWS unveils AI agents that code for days without human oversight

Amazon Web Services introduced a new class of AI technology on Tuesday that can work independently for hours or days to develop, secure and maintain software applications—marking a decisive shift from AI assistants to fully autonomous digital workers.

Autonomous Operation Sets New Standard

Unlike existing AI coding tools such as GitHub Copilot that require constant human prompting, frontier agents maintain persistent memory across sessions and can spawn multiple versions of themselves to tackle different aspects of a problem simultaneously. Deepak Singh, AWS vice president of developer agents and experiences, told VentureBeat that these agents are "fundamentally crafted to operate for extended hours and days" without human intervention.

venturebeat +1

The Kiro agent continuously learns from pull requests and feedback, handling tasks from bug triage to multi-repository code changes. It connects to tools like Jira, GitHub and Slack, operating as a virtual team member that proposes changes rather than directly committing code to production. constellationr +3

Human engineers retain final approval authority, with agents unable to commit code directly to production systems.

venturebeat +1

2. Dezember 2025 https://www.perplexity.ai/page/aws-unveils-ai-agents-that-cod-LI9PQ_N2TNuW.fj.zYwQdA

10

Bis 2028 werden voraussichtlich 33 Prozent der Enterprise-Software autonome Agenten integrieren.

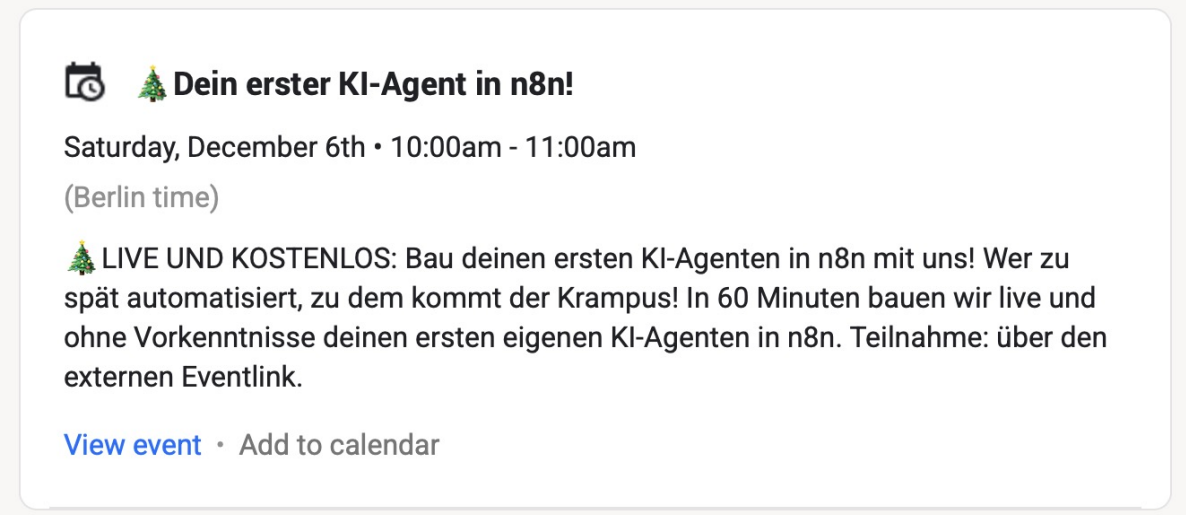
Quelle:
https://www.gartner.com/en/articles/ai-agents

11

Bis 2027 werden fast 50 Prozent aller geschäftlichen Entscheidungen durch intelligente KI-Agenten geprägt oder vollzogen.

Quelle:
https://prwire.com.au/pr/122276/gartner-announces-the-top-data-analytics-predictions

12



13

POLITICO

French political crisis EU-US relations War in Ukraine | Newsletters Podcasts Poll of Polls Policy news Ev

NEWS > TECHNOLOGY

Albania appoints world's first AI-made minister

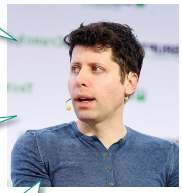
Diella, who is powered by artificial intelligence, will handle public procurement.

Quelle: https://www.politico.eu/article/albania-apppoints-worlds-first-virtual-minister-edi-rama-diella (12.09.2025)

14

Sam Altman Juli 2025

Although the utility is significant, so are the potential risks.



... not something I'd yet use for high-stakes uses or with a lot of personal information until we have a chance to study and improve it in the wild

...we recommend giving agents the minimum access required to complete a task to reduce privacy and security risks.

Sam Altman
7. Juli 2025 · **2,8 Mio.** Mal angezeigt

16

17

Was sind KI-Agenten?

- Ein Agent mit künstlicher Intelligenz (KI) ist ein Softwareprogramm, das mit seiner Umgebung interagieren, Daten sammeln und die Daten verwenden kann, um selbstbestimmte Aufgaben auszuführen, um vorgegebene Ziele zu erreichen. (Amazon)
- … eine künstliche Entität, die dazu in der Lage ist, ihre Umgebung wahrzunehmen, Entscheidungen zu treffen und Aktionen auszuführen.

Charakteristische Eigenschaften von KI-Agenten

Autonomie: Agenten nehmen ihre Umgebung unabhängig wahr, treffen Entscheidungen und führen Aktionen aus, ohne externe Anweisungen zu benötigen.

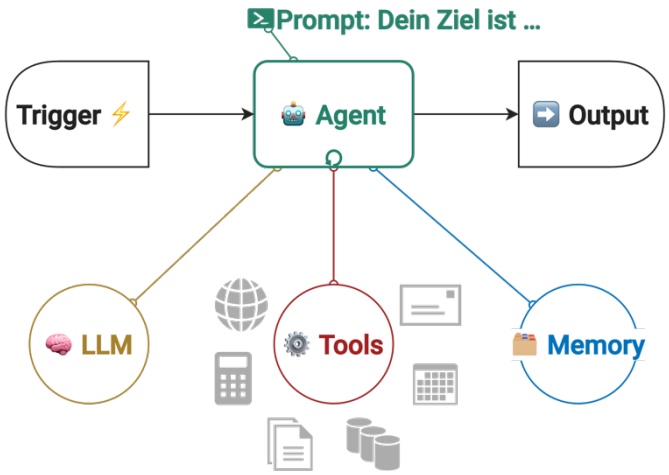
Wahrnehmung: Sie verfügen über sensorische Fähigkeiten, um Informationen über ihre Umgebung zu sammeln.

Entscheidungsfindung: Agenten treffen Entscheidungen basierend auf wahrgenommenen Informationen, um ihre Ziele zu erreichen.

Handlung: Sie führen Aktionen aus, die den Zustand ihrer Umgebung verändern.

20

Ein Agent verfolgt autonom Ziele, statt Workflows abzuarbeiten.



Ein KI-Agent kann seine Umgebung wahrnehmen, komplexe Ziele verfolgen, auf Basis von Beobachtungen und Zielen **eigenständig Entscheidungen über die nächsten Schritte und die Nutzung verfügbarer Werkzeuge treffen**, und potenziell aus Interaktionen lernen. Agenten arbeiten oft in einer Schleife (z.B. Beobachten-Orientieren-Entscheiden-Handeln), nutzen interne Speicher (Gedächtnis) und können eine Reihe von Tools (Software, APIs, Datenbanken) orchestrieren, um ihre Ziele zu erreichen.

21

Demo (n8n)

22

Abgrenzung Workflow oder Agent

23

Die Unterscheidung: Agent vs. Workflow

Charakteristik	**Workflow** (Steuerung)	**KI-Agent** (Regelung)
Kontrollfluss	Statisch, vordefiniert	Dynamisch, vom Agenten bestimmt
Zielorientierung	Folgt festem Prozess	Verfolgt übergeordnetes Ziel bis erreicht
Werkzeugauswahl	Hart codiert	Situativ gewählt
Fehlerbehandlung	try-catch	Adaptiv, emergent

- Das Unterscheidungsmerkmal liegt in der Feedback-Schleife:
- Der Agent folgt dem ReAct-Pattern (Reason + Act): Beobachten → Vergleichen (Soll-Ist) → Planen → Handeln → Loop.
- Ein Workflow hingegen arbeitet eine festgelegte Sequenz ab, unabhängig davon, ob das Ziel erreicht wurde.

25

Ein Agent ist nicht zwingend das richtige Instrument

Chatbots
Individuelle Fragen

Custom Chatbots
Spezifische Fragen zu einem Thema

Workflows
Wiederkehrende Abläufe

Agenten
Komplexe Aufgaben mit verschiedenen Lösungs-möglichkeiten

Case 1: Intelligenter IT-Support Agent (Verwaltung)

Anlass / Auslöser: Ein Drucker funktioniert nicht und ein Benutzer stellt daraufhin eine Service Anfrage an den IT-Support.

Primärziel des agentischen Verhaltens:

Das IT-Problem des Nutzers ist gelöst – nicht nur "Ticket bearbeitet": der Drucker funktioniert wieder und alle beteiligten Akteure sind darüber in Kenntnis gesetzt.

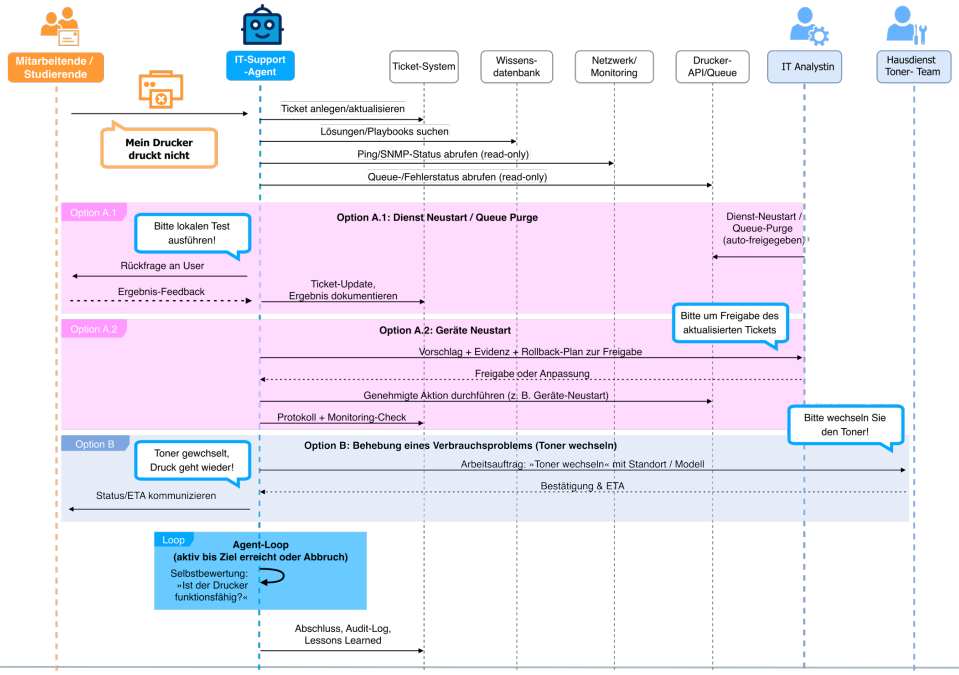
Sekundär Ziele im agentischen Prozess:

„Reduziere Zeit bis zur Lösung und erhöhe First-Contact-Resolution, ohne Sicherheits- oder Compliance-Verstöße."

Warum Agent statt Workflow?

Ein Workflow würde das Ticket kategorisieren und weiterleiten – fertig. Der Agent hingegen iteriert: Er durchsucht die Wissensbasis, schlägt Lösungen vor, prüft, ob sie funktioniert haben, und eskaliert bei Bedarf.

Besonderheit des Use-Case: hybride Mensch-Agenten-Teams



29

Use-Case 2: Intelligenter Tutor als Doppel-Agent (Student:in UND Lehrende:r)

Ziel: Mastery-Learning für die Vorlesung Mathematik im ersten Semester erreichen. Studierende sollen die Materie beherrschen, statt nur das Modul bestehen.

Kernidee

„Verbessere Lernfortschritte über das Semester (z. B. Konzeptbeherrschung, Durchfallquote senken), personalisiert und fair.“

Der Tutor-Agent arbeitet in zwei koordinierten Regelkreisen:

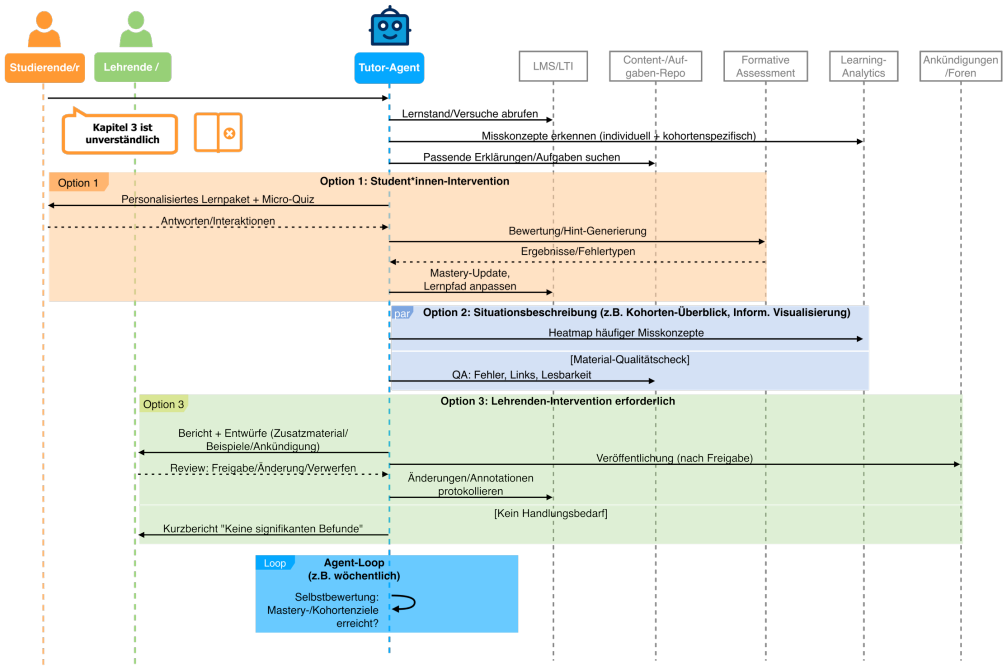
Student-Loop: Diagnostiziert Lernstand, schlägt Inhalte/Übungen vor, gibt Feedback, plant Nudges

Lehrenden-Loop: Prüft Kursmaterialien (Qualität, Fehler), erkennt häufige Misskonzepte, schlägt Kurs-Interventionen vor (Material anpassen, Ankündigungen), erstellt Entwürfe – Veröffentlichung durch Lehrende

Ein Workflow würde starr Aufgaben ausrollen oder wöchentliche Mails versenden. Der Agent betrachtet iterativ Lernstände, wählt dynamisch passende Interventionen, prüft deren Wirkung und greift bei Bedarf auch in der Lehrenden-Ebene ein (Material-QA, Misskonzept-Reports), bis die Ziele erreicht sind.

Besonderheit des Use-Case: Adressierung mehrerer Zielgruppen

30



Case 3: Lernender Research-Monitoring- und Einordnungs-Agent (Forschung)

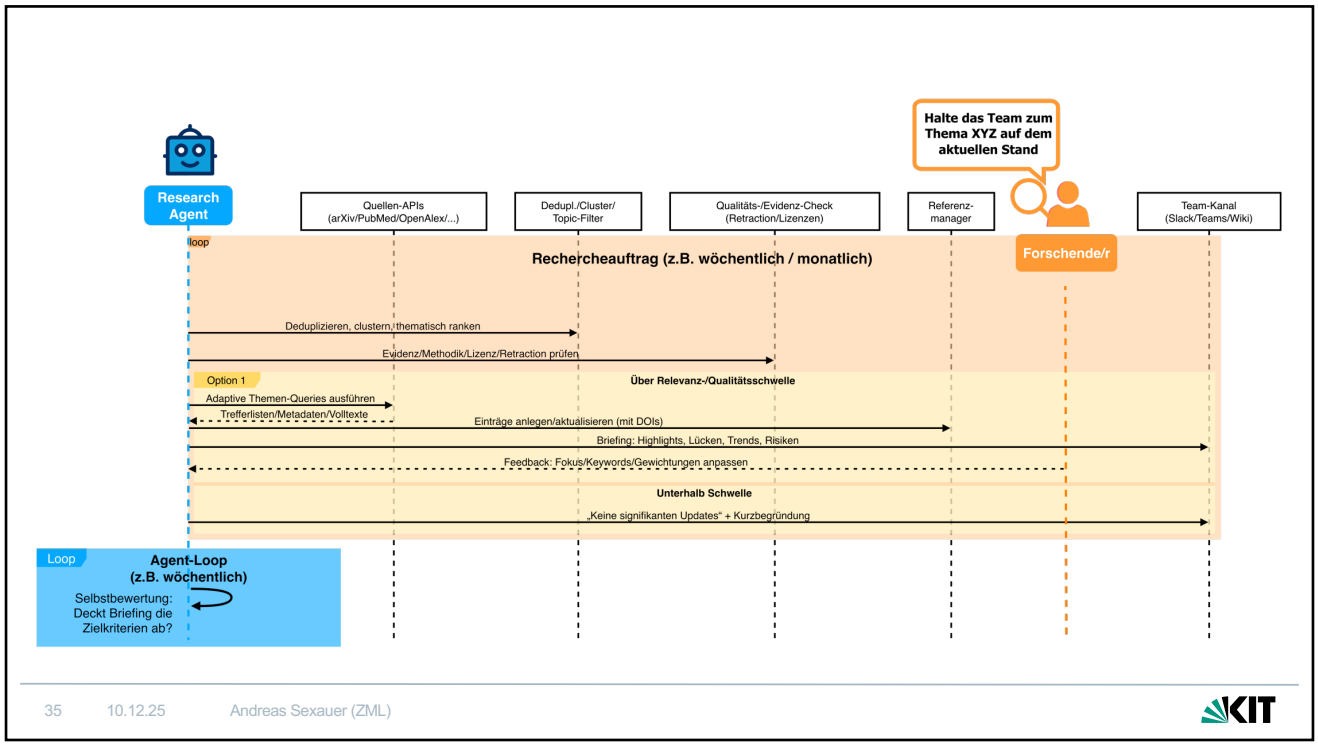
Ziel: Der Agent beobachtet für Forscher ein bestimmtes Forschungsfeld, erstellt dazu periodisch ein aufbereitetes Briefing. Er reagiert adaptiv auf Feedback der Forscher zur thematischen Schwerpunktsetzung oder dem Format der Aufbereitung. Ebenfalls erkennt er Veränderungen im Forschungsfeld und passt seinen Rechercheauftrag daran an. Er ist damit vergleichbar zu der Tätigkeit einer menschlichen Assistenz mit dem Auftrag relevante Forschung rechtzeitig zu erkennen und belastbar einzuordnen – statt nur „Links zu sammeln“

Kernidee:

„Erstelle regelmäßig evidenzbasierte Briefings zu definierten Themen, mit hoher Signal-zu-Rausch-Rate, korrekter Zitierung und klarer Priorisierung.“

Ein Workflow würde vordefinierte Queries fahren und Listen liefern. Der Agent passt Suchstrategien an, bewertet Relevanz/Qualität, dedupliziert, prüft, erstellt kuratierte Einordnungen und justiert den Fokus anhand von Feedback.

Besonderheit des Use-Case: Lernender Agent



35

Idee 1	Idee 2	Idee 3

Brainwriting Use-Cases

36

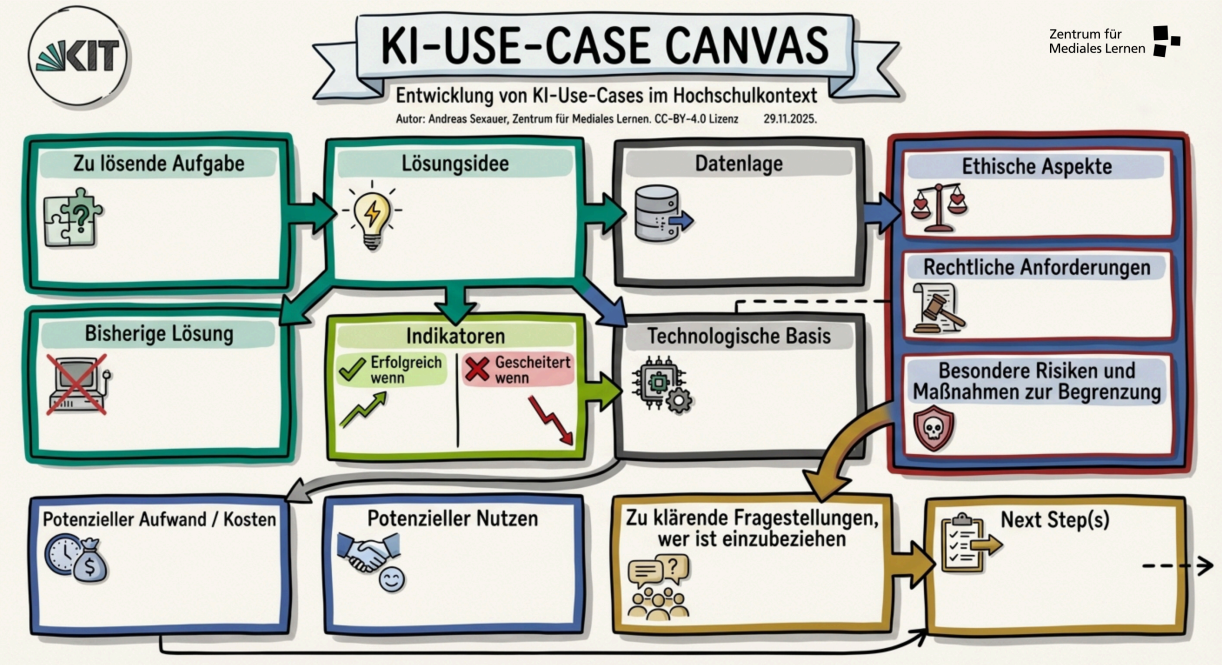
Arbeitsgruppen: Drei Schlüsselfragen beantworten (20 Min)

Fähigkeiten:
Über welche Fähigkeit sollte der KI-Agent verfügen?

Mehrwert:
Welcher Mehrwert ist ausschlaggebend?

Ethische Grenze / Risiko:
Welche Grenzen müssen gesetzt werden, welche Risiken sind zu berücksichtigen?

37

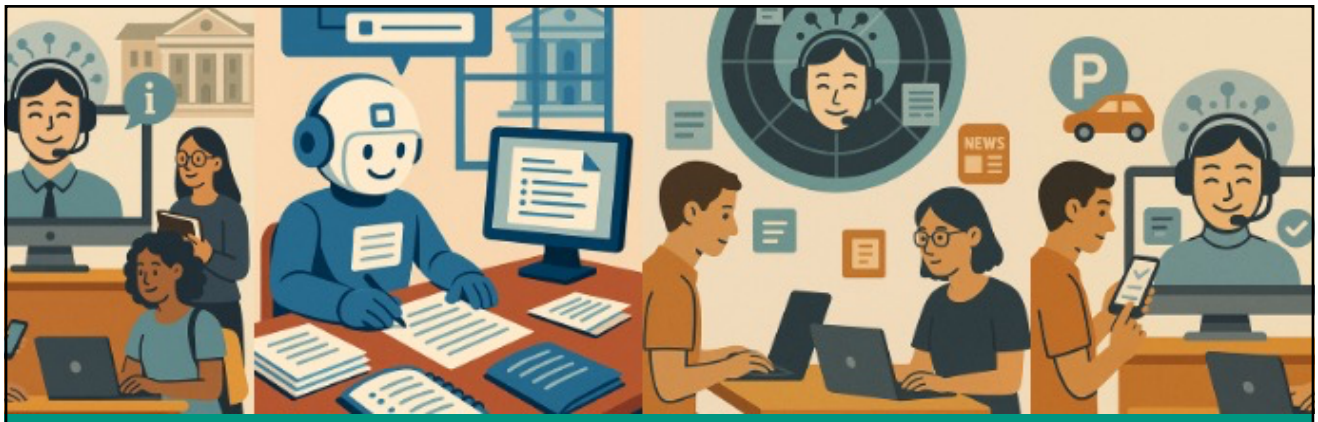


38

Diskussion und Abschluss (20 Min)

- **Kurzvorstellung der Diskussionsergebnisse (3 Min / Gruppe)**
- **Fazit und Überlegungen zum Anschluss**

39



Transformation, Governance, Risiko

02

40

Transformation: Neue Führungsaufgaben

CIO-Transformation: Von IT-Manager zum strategischem Transformationsarchitekt

Die Rolle des Chief Information Officer ändert sich radikal. Salesforce und weitere Firmen empfehlen, einen "AI-Orchestrator" zu ernennen, um eine unternehmensweite Strategie zu treiben, mit einer prominenten strategischen Rolle, die Zusammenarbeit und Veränderungsmanagement über die gesamte Organisation ermöglicht.

https://www.salesforce.com/eu/blog/ai-agents-for-c-suites/

41

Zentrale Fragestellungen

Wahrnehmung:
- Welche Daten nehmen Agenten auf?
- Mit welcher Genauigkeit und Bias?

Entscheidungen von Agenten:
- Wie treffen Agenten Entscheidungen?
- Sind diese nachvollziehbar und kalibriert?

Ausführungsebene:
- Welche Systemzugriffe hat der Agent?
- Mit welchen Sicherheitsmechanismen?

Feedbackdimension:
- Wie lernen Agenten?
- Sind die Feedback-Schleifen robust?

42

Neue Organisationsrollen

AI-Orchestratoren: sind die Brücke zwischen technischen und inhaltlich verantwortlichen Teams und stellen sicher, dass KI mit Zielen und -werten abgestimmt sind.

AI-Governance-Architektinnen: definieren die Prozesse für die Einbindung von Agenten, überwachen das Verhalten und passen Ziele an.

Datenqualitäts-Kuratorinnen: sichern die Qualität und Relevanz der Trainingsdaten und Datenquellen.

Business-Process-Agent-Designer: konzeptualisieren, wie Prozesse für die Zusammenarbeit mit Agenten umgestaltet werden müssen.

https://www.hireborderless.com/post/how-ai-is-reshaping-organizational-design

43

Governance als strategische Imperative

KI-Governance wird **von theoretischer zu operativer Notwendigkeit**.

Mit konsequentem Einsatz von KI-Agenten werden **Prozesse nicht mehr primär für menschliche Benutzer designt**, sondern für eine "digitale Hybrid-Belegschaft" aus Menschen und KI-Agenten.

Diese Agents werden:
- Arbeitsschritte selbständig planen
- Tools und Daten nutzen
- Entscheidungen vorbereiten oder sogar treffen
- Mit Elementen ausgestattet sein, die bisher menschlichen Teams vorbehalten waren

Aspekt	Traditionelle Governance	Agentische Realität
Entscheidungen	Vordefinierte Logikpfade, prüfbar und auditierbar	Kontextabhängiges Reasoning, emergentes Verhalten
Delegation	Einzelne Service-Grenze, klare Verantwortung	Rekursive Agent-Ketten, verteilte Verantwortung
Richtlinien-durchsetzung	Deployment-Zeit-Überprüfungen, periodische Audits	Echtzeit-Durchsetzung im Moment der Aktion
Auditierbarkeit	Statische Codes und Logs	Dynamische Entscheidungsspuren über mehrere Agenten und Tools

44

Rechtliche Rahmenbedingungen (EU-AI-Act)

Risikostufe	Charakterisierung	KI-Agenten-Beispiele	Regulatorische Anforderungen
Minimales Risiko	Keine oder sehr geringe Risiken; keine spezielle Regulierung erforderlich	Einfache Verwaltungsassistenten, Datenklassifizierungssysteme	Keine spezifischen Anforderungen
Begrenztes Risiko	Direkte Interaktion mit Menschen; Transparenzpflichten	Virtuelle Service-Agenten, Recherchesysteme in der Forschung, allgemeine Assistenten	Transparenzpflicht: Nutzer müssen wissen, dass sie mit KI interagieren; keine manipulativen Funktionen; KI-Transparenzkennzeichnung
Hohes Risiko	Erhebliches Schadensrisiko; umfassende Governance erforderlich	Agenten für Personalentscheidungen, Beschaffung, Kreditvergabe, Prüfungen, Bewerbungs- und Zulassungsprozesse, kritische Infrastruktur-Steuerung, komplexe Agentensysteme in der Forschung	Umfassende Risikomanagement, Datengovernance, Dokumentation, menschliche Überwachung, Registrierung, Konformitätsbewertung
Inakzeptables Risiko	Verletzung fundamentaler Grundrechte; bzw. verboten	Systeme für manipulative Verhaltenssteuerung, biometrische Massenüberwachung, Social Scoring,	Absolute Verbote – keine Ausnahmeregelungen

Uses-Cases und Bewertung mit dem Konzept des Akzeptierten Risikos

Geführter Denkprozess für die Diskussion:

Status Quo analysieren: Welche und wie viele Fehler passieren im heutigen manuellen Prozess? Welche davon sind für uns und unsere Kunden (Studierende etc.) unbemerkt und okay? Bei welchen Fehlern knallt es – also gibt es ernsthafte Konsequenzen?

Risikogrenze definieren: Aus dieser Analyse leiten wir ab: Welches Fehlerniveau (Art und Häufigkeit) wäre für den neuen, agentenbasierten Prozess aus institutioneller Sicht akzeptabel?

Agenten gestalten und messen: Wie müssen wir den Agenten und seine Leitplanken (z.B. menschliche Freigabeschleifen) gestalten, damit er diese Risikogrenze einhält? Wie können wir seine Leistung und Fehlerquote messen, um sicherzustellen, dass dies erreicht wird?

Use-Cases zum Starten

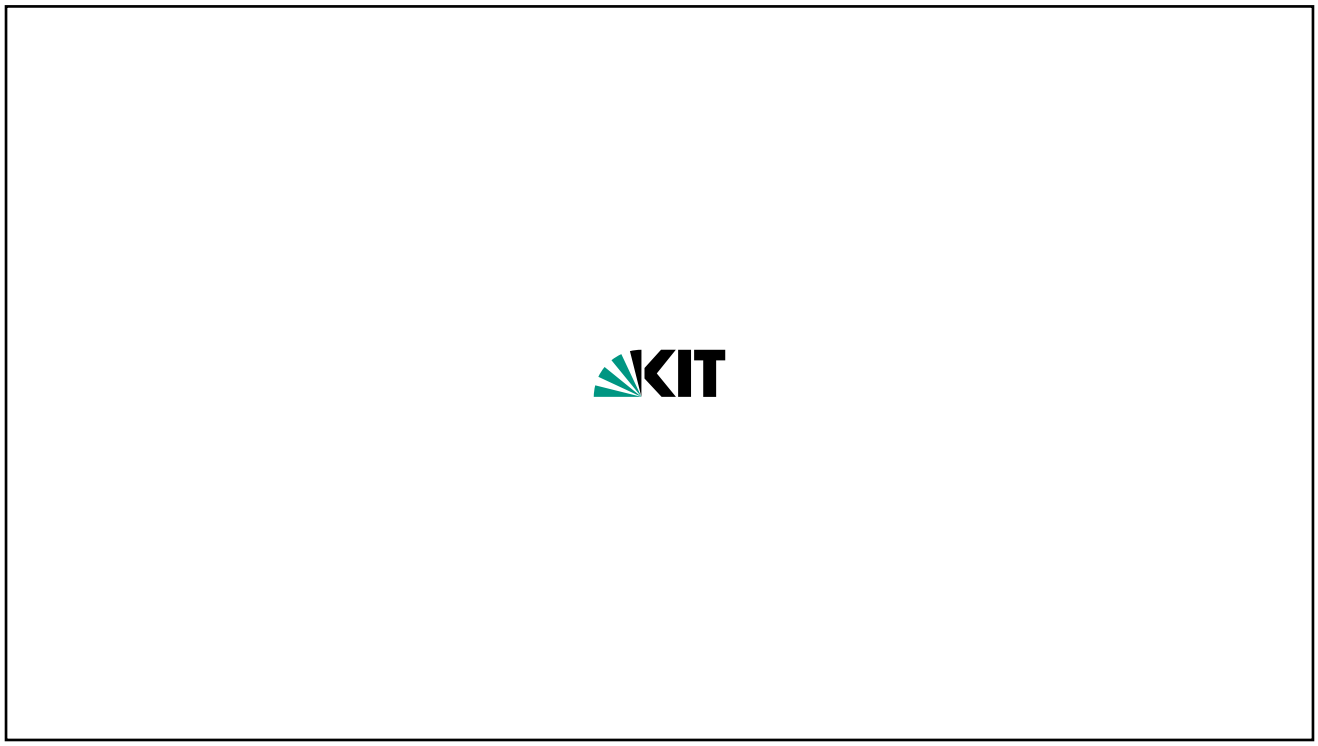
- Low Risk
- High Impact
- Low Complexity

49

Ein Agent ist nicht zwingend das richtige Instrument

Chatbots	Custom Chatbots	Workflows	Agenten
Individuelle Fragen	Spezifische Fragen zu einem Thema	Wiederkehrende Abläufe	Komplexe Aufgaben mit verschiedenen Lösungs-möglichkeiten

50



51