Neue Akteure in der Hochschultransformation:

KI-Agenten im Einsatz

Stefan Göllner (KI-Campus), Andreas Sexauer (KIT)

Zentrum für Mediales Lernen

AI Campus
powered by Stifterverband

Ablauf Zeitplan (65 Minuten)

- Ankommen, Vorstellung Workshopleiter (5 Min)
- Mini Intro und Live Demo (5 Min)
- Input: Agenten (15 Min)
- Einstieg Gruppenarbeit (3 Min)
- Gruppenarbeit (15 Min)
- Ergebnisse Gruppenarbeit besprechen (wir machen Rundgang) in Bewertung durch TN durch 3 Klebepunkt gleich zu Beginn des Rundgangs begleitend (10 Min)
- Abrundung und Fragen (10 Min)

AI Campus
powered by Stifterverband

KIT

1. »KI-Agenten« was ist das?
2. Beispiele
3. Workflow oder agentischer Prozess?
4. Agentische Plattformen für den Einstieg
5. Dokumentation und Reflektion

KIT

AI Campus
powered by Stifterverband

Theorie und mehr zum Nachlesen vorweg …
Organisationale Perspektive

Publikation: Agentische KI im Hochschulsystem

07.05.2026

https://hochschulforumdigitalisierung.de/news/publikation-agentische-ki/

Wir stehen an einer Schwelle. Es gilt das Chatfenster zu verlassen und KI bei der Ausführung komplexer Abläufe zu orchestrieren. Ähnlich wie das Aussteuern einer wissenschaftlichen Hilfskraft.

PROMPT:

»Im Ordner "Aufbereitete-Inhalte" liegt die Aufbereitung eines Diskussionspapiers zu agentischer KI im Hochschulkontext. Ich will ein Projekt vorschlagen, um an meiner Hochschule das Thema voranzubringen, wir haben bisher noch keine strategischen Aktivitäten dazu.
Erstelle ein Verzeichnis, in dem eine erste Projektskizze und jeweils Dokumente für relevante Stakeholder sein sollen. Diese Dokumente sollen jeweils auch eine Executive Summary und Hintergrundinformationen beinhalten, mit dem ich die Stakeholder gezielt ansprechen kann.«

LIVE DEMO

Mehr Tools lösen nicht das Problem

Quellen: LinkedIn Harish kumar, 04.05.2026, https://www.linkedin.com/posts/harishkumar-sh_artificialintelligence-aitools-futureofwork-share-7457014872862613504-15em und AI for Enterprise, https://www.linkedin.com/posts/ai-tools-2025-vs-2026-%F0%9D%97%95%F0%9D%97%B2%F0%9D%97%B0%F0%9D%97%BC%F0%9D%97%BA%F0%9D%97%B2-%F0%9D%97%AF%F0%9D%97%B2-share-7454665691665879040-cbGG

Hinter vielen Tools stehen „nur“ Agenten …

Perplexity Deep Research

- **Research Agent**
- **Ziel**: Bericht zu einem Thema zu erstellen
- **Analysiert** das Thema führt eine **Suche** im Internet und Fachdatenbanken durch.
- **Bewertet** die Suchergebnisse
- **Passt** bei Bedarf die Suche mehrfach **an** und verwendet dazu bereits ermittelte Informationen
- **Erstellt** einen Bericht dazu
- **Wandelt** diesen in eine Infografik, Word oder PDF-Datei **um**.

Notebook LM

- Ist eine **App mit verschiedenen Agenten**
- Verschiedene Agenten für z.B.:
  - Analyse und Zusammenfassung von Dokumenten
  - Erstellen von Audio- und Videozusammenfassungen
  - Interaktive Mindmaps
  - Erstellung von Präsentationen

Claude Desktop

- **KI-Agent** System von Anthropic das direkt auf dem Desktop arbeitet
- Führt unterschiedlichste Aufgaben eigenständig und mehrstufig aus
- Z.B. Lesen, Bearbeiten und Organisieren von Dateien
- Ist ein Beispiel für einen Agent-Harness

Lösung ist: Selbst auf der individuellen Ebene generisch einsteigen, …

- **Ein universelles Agenten-Tool** (Harness) statt vieler Spezialwerkzeuge
- **Selbst steuern** statt Blackbox mit eingeschränkten Anpassungsmöglichkeiten
- **Ein System** statt zahlreicher Lizenzen oder Abos
- Verwendung **hochschuleigener oder lokaler Modelle** statt fehlenden Datenschutzes und fehlender Informationssicherheit

Den Einstieg wie dies geht, bekommt ihr heute.

Abgrenzung KI-Chatbot, KI-Assistent, KI-Agent

KI-Chatbot

- **Eine Frage, eine Antwort**
- Generiert Text, Bilder, Code, …
- Einsatz für Einzelanliegen
- Mensch initiiert und steuert den Dialog
- *Beispiel: Erläuterung eines Sachverhaltes*

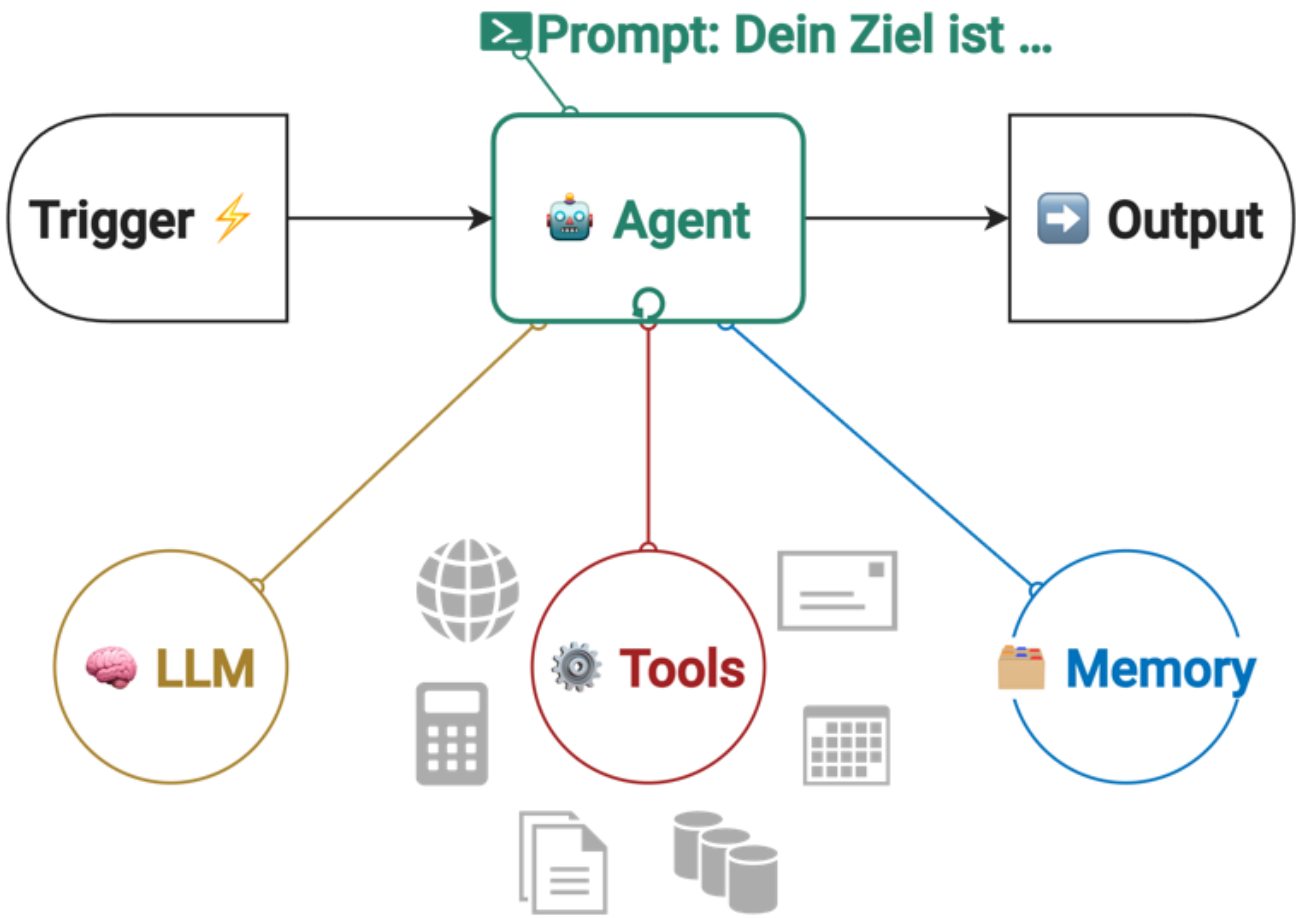
KI-Assistent

- **Komplexere vordefinierte Tätigkeiten**
- Generiert Berichte, Präsentationen, Diagramme …
- Einsatz für wiederkehrende komplexere Anliegen nach bestimmtem Schema
- Mensch konfiguriert, gibt Daten, nutzt Ausgabe
- *Beispiele: Deep Research von Perplexity, NotebookLM*

KI-Agent

- **Generische Bearbeitung von Aufgaben**
- Trifft Entscheidungen, ändert Dateien und Daten, agiert in anderen Systemen
- Einsatz für Aufgaben mit offenen Lösungswegen
- Mensch definiert den Auftrag, die Ziele, prüft und nutzt Ergebnisse
- *Beispiel: Analyse und Organisation von Datei- und Datenbeständen*

Ein Agent verfolgt autonom Ziele, statt Workflows abzuarbeiten.



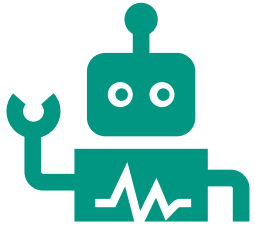
- Nimmt seine Umgebung wahr.
- Verfolgt komplexe Ziele.
- Trifft auf Basis von Beobachtungen und Zielen eigenständig Entscheidungen über: nächste Schritte; Nutzung verfügbarer Werkzeuge
- Lernt potenziell aus Interaktionen.
- Arbeitet oft in einer Schleife (z.B. Beobachten – Orientieren – Entscheiden – Handeln).
- Nutzt interne Speicher ("Gedächtnis").
- Orchestriert eine Reihe von Tools (Software, APIs, Datenbanken), um Ziele zu erreichen.

KI-Agenten: unterschiedlich eingesetzt

Klassische Workflows mit Agenten-Unterstützung

Agenten **helfen bei der Erstellung** von Prozessen/Anwendungen als klassischer Business-Prozess in vorhandenen Systemen.

*SAP Joule Studio: Ein **KI-Agent generiert Workflows** für die automatisierte Übertragung von Testergebnissen aus dem E-Prüfungssystem ins Prüfungsmanagement."*



Agentenanwendungen (selbstständig agierend)

Agenten **sind das Ergebnis** und werden eingesetzt. Die Agenten selbst können wiederum mit einem Agenten erstellt werden (siehe links).

*Ein **KI-Agent analysiert Rechnungen**, extrahiert Daten und schlägt Buchungssätze vor.*

Agentensysteme in der Organisation

Lokale, generische Agentensysteme (individuelle Verwendung)

Agentic Harness **läuft lokal als generische Anwendung** und wird situativ und individuell in Tätigkeiten verwendet.

*VSCodium + Cline/Hermes: **Lokale KI-Agenten** analysieren Meeting-Protokolle, führen Code-Reviews durch oder extrahieren Wissen aus Projekt-Dokumentationen und erstellen strukturierte Zusammenfassungen– direkt auf dem Dateisystem.*

Persönlicher Einsatz

Die Zutaten für den Einstieg auf der individuellen Ebene

Inferenz per API

- **Zugriff** auf ein leistungsstarkes **Modell** mit Tool-Calling **per API**
  - Es muss nicht zwingend Claude Sonnet/Opus sein
  - Qwen 3.6, Minimax 2.7, Gemma-4
  - Kosten im Blick behalten
  - Datenschutz und Informationssicherheit beachten
- API besteht aus
  - **Endpoint URL**
  - **API-Key**: sk-8d9f0a1b2c3d…
  - Model-ID
- Alternativ Token bei Openrouter.ai für den Einstieg kaufen.

Agentische Umgebung (Harness)

- Tools kommen häufig aus dem Vibe-Coding, sind aber nicht darauf eingeschränkt
  - Claude Desktop
  - **Cline in VsCodium**
  - Hermes Agent
  - Pi Agent
  - …
- Der **Harness verwaltet alles**, was nicht das Modell selbst ist: Werkzeuge, Speicher, Statusverwaltung und Ausführungslogik

Vorgehensweisen und Tools

- Gezielte Verwendung, z.B. Plan und Ausführungsmodus
- Gestaltung über z.B.
  - **Skills** als Fähigkeiten die häufiger benötigt werden, z.B. bestimmte Datenanalysen und für wiederkehrende feste Abläufe, z.B. Dokument konvertieren, analysieren, Analyseergebnisse als Bericht ablegen
  - **Rules** als immer einzuhaltende Regeln, z.B. Projektkonventionen für Dateinamen und Verzeichnisstrukturen

Use Cases in Breakouträumen weiterdenken:

Der perfekte HIWI

Was ist Dein Traum-Agent? ... und was braucht dieser, um seine Aufgaben perfekt zu erledigen?

Arbeitsphase: 15 Minuten

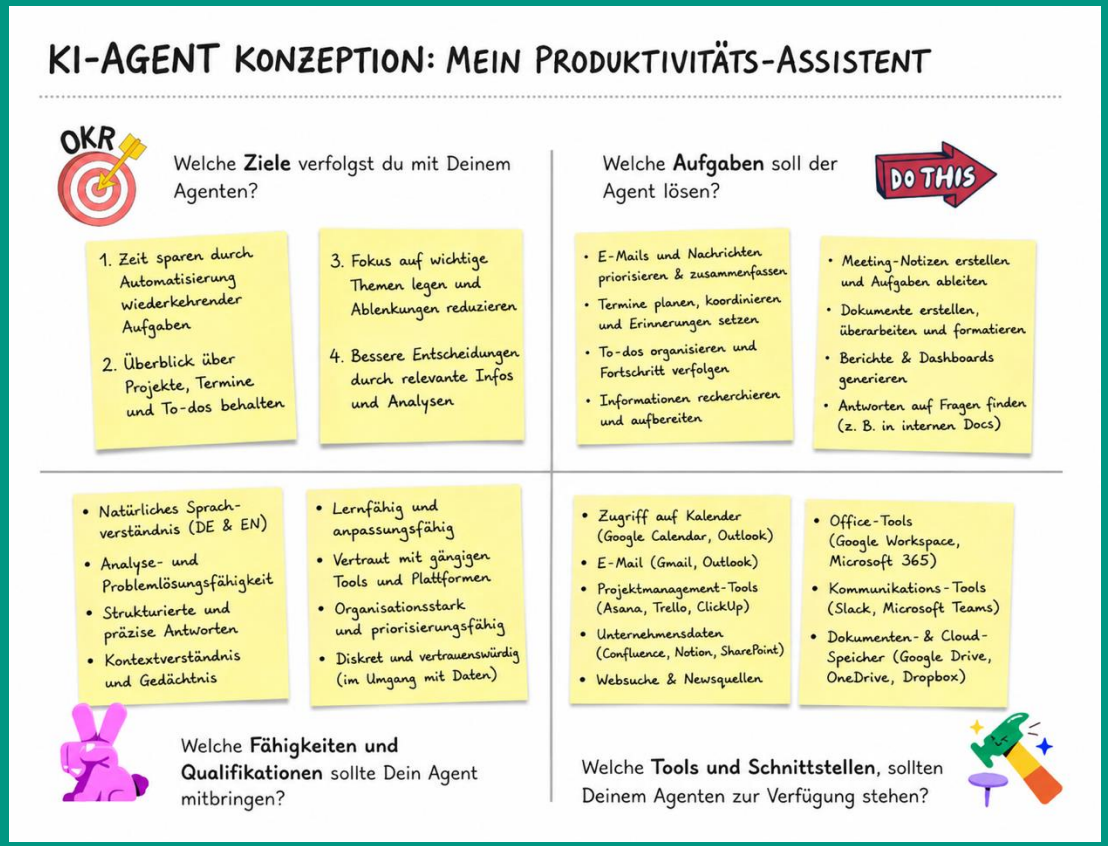
Titel Deines Agenten

Welche **Ziele** verfolgst du mit Deinem Agenten?

Welche **Aufgaben** soll der Agent lösen?

Welche **Fähigkeiten und Qualifikationen** sollte Dein Agent mitbringen?

Welche **Tools und Schnittstellen**, sollten Deinem Agenten zur Verfügung stehen?



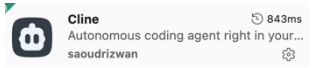
Rundgang: Ergebnisse der Arbeitsphase

AI Campus
powered by Stifterverband

Schritt für Schritt Anleitung Cline in VSCodium

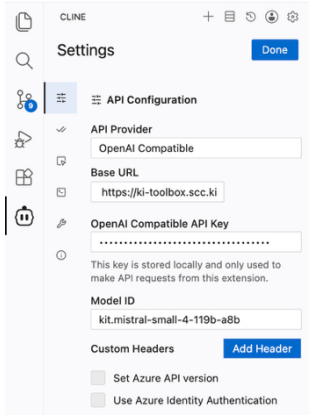
Installieren

- VSCodium (https://vscodium.com/) oder Visual Studio Code (Microsoft App Store) installieren
- In VSCodium bzw. Visual Studio Code den Extensions Marketplace öffnen und nach Cline suchen.
- Die Extension Cline installieren



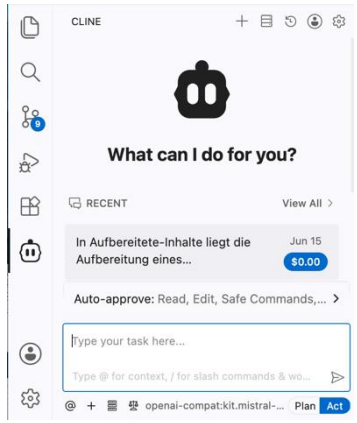
Einrichten

- Cline erscheint in der Seitenleiste als Roboter Symbol . Dieses öffnen und dort in den Einstellungen (Zahnrad oben rechts) die API-Konfiguration vornehmen.
- Entweder API der Hochschule mit entsprechendem Datenschutz und Informationssicherheit …
- oder zum Testen (mit unverfänglichen Inhalten) einen kostenlosen Cline-Account anlegen und freie Modelle von dort verwenden.



Verwenden

- Aufträge an den Agenten werden im Chatfenster eingegeben.
- Die Modi unten rechts gezielt verwenden:
  - Plan: Plant erst, ohne zu ändern
  - Act: Führt aus und Bearbeitet Dateien
- Weitere Hinweise zur Verwendung: https://docs.cline.bot



AI Campus powered by Stifterverband

KIT

Agentisches Arbeiten weiter ausbauen, typische Elemente

Rules

*Regeln definieren die **Spielregeln**, **Rahmenbedingungen** und Leitplanken für KI-Agenten.*

Sie legen fest, was ein Agent tun darf oder nicht tun soll, um zielgerichtet und sicher zu arbeiten.

Sie werden bei jedem Vorgang vollständig gelesen.

Skills

Sind spezielle **Fähigkeiten** oder **Abläufe**, die **situativ** zum Einsatz kommen. Der KI-Agent entscheidet, ob und wann er diese braucht.

Sie legen fest, wie etwas bestimmtes erfolgen soll, damit dieses wiederholbar nach einem bestimmten Muster erfolgt.

Der KI-Agent kennt nur eine Übersicht und lädt die vollständigen Anweisungen bei Bedarf nach.

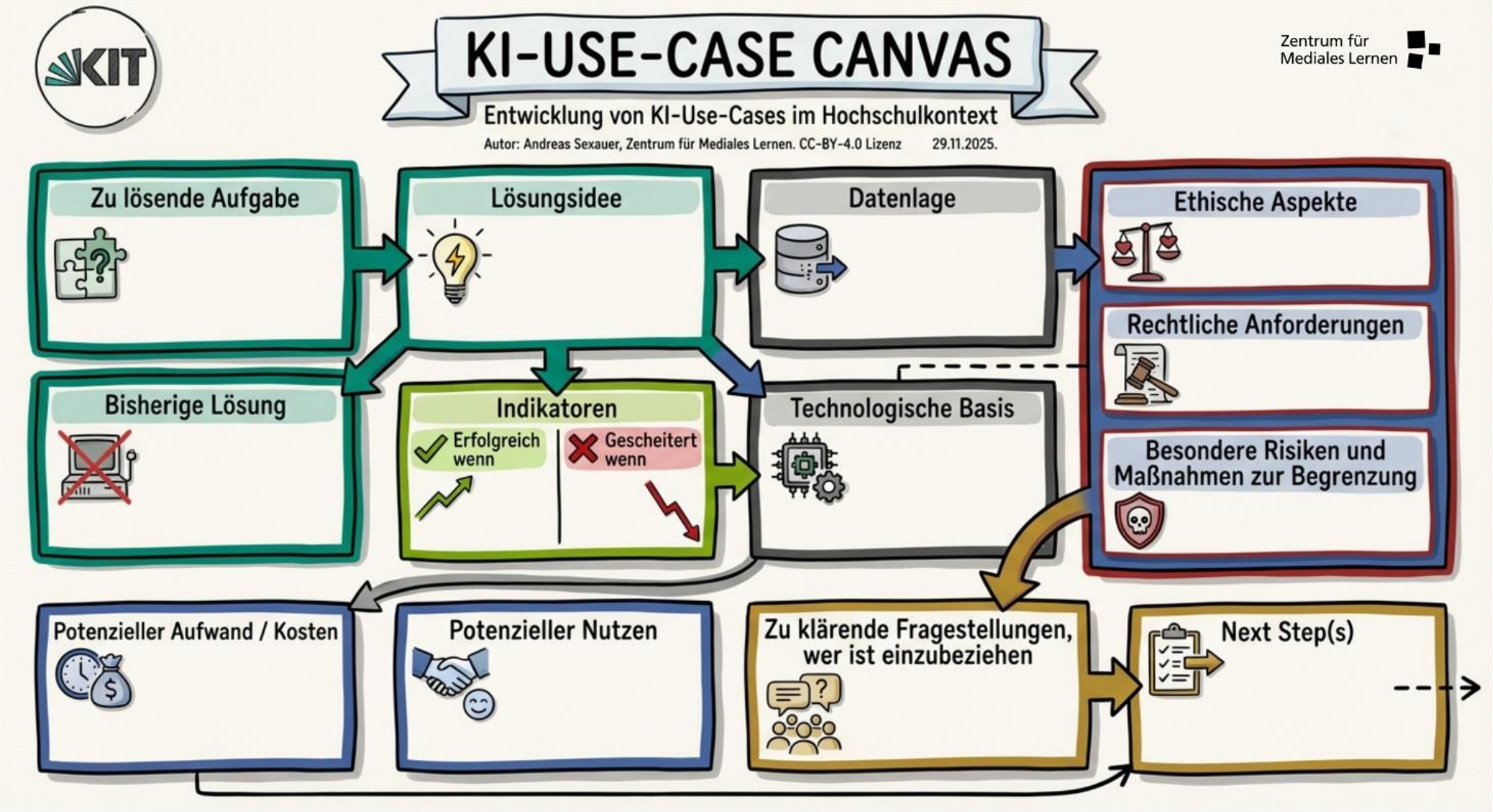
Tools (z.B. per MCP)

MCP steht für Modularen Konnektor – also die **Schnittstelle**, über die ein KI-Agent **auf externe Werkzeuge** oder Datenquellen zugreift.

Jedes Tool erweitert die Fähigkeiten des Agenten und ermöglicht spezifische Funktionen, z.B. der Zugriff auf eine Datenbank oder ein anderes System.

Der KI-Agent weiß welche Tools zur Verfügung stehen und nutzt diese bei Bedarf.

Den Use-Case vollständig betrachten



Entwicklung von KI-Agenten: Verantwortungsvoll gestalten

Die neuen Risiken (Vom Chatbot zum Agenten)

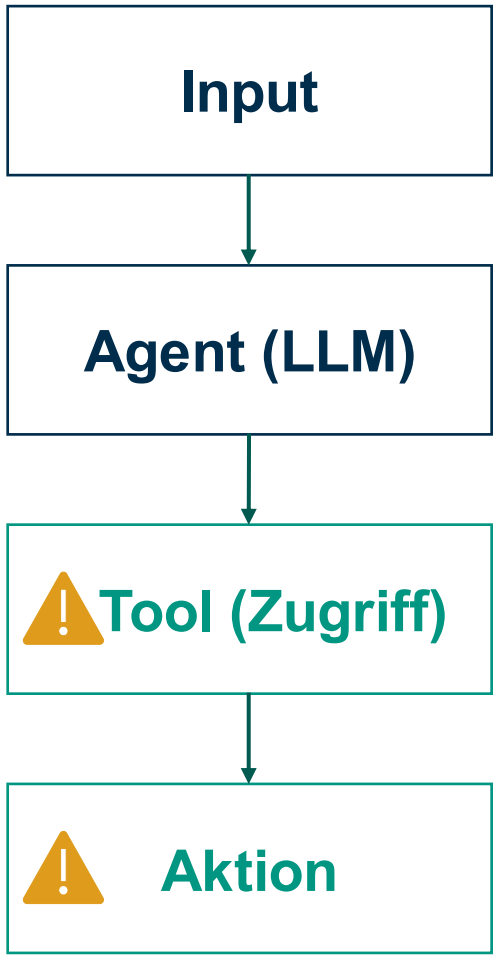
Verlust der Kontrolle (Autonomie): Schleifen, ungewollte API-Aufrufe, halluzinierte Befehle

Datenabfluss an Dritte: externe LLMs, insbesondere bei sensiblen Forschungs- oder Personaldaten.

Indirect Prompt Injection: Verarbeitete Inhalte werden als Prompt interpretiert

Rechte-Eskalation: API-Key hat zu viel Zugriff.

Unkontrollierbare Kosten durch kontinuierliches "verbrennen" von Tokens



Best Practices (Security by Design)

Human-in-the-Loop: Kritische Aktionen benötigen menschliche Bestätigung.

Principle of Least Privilege: Agenten haben nur die *absolut minimalen* Rechte, dedizierte Service-Accounts.

Sandboxing & Isolation: Agenten in Containern (Docker), ohne Zugriff auf interne Ressourcen, sofern nicht zwingend nötig.

Monitoring & Logging: Protokollieren Sie jeden „Thought“ und jede „Action“ des Agenten, um Fehlverhalten sofort zu erkennen.

Fragen?