

KI-Agenten brauchen wie neue Mitarbeitende klare Aufgaben, Zugriffsrechte und verlässliche Prozesse

Deep Dive · KI-Tag der Wirtschaft 2026

Ufkun-Bayram Menderes · Ferdinand-Steinbeis-Institut

Was dieses Format NICHT ist

- Kein Prompt-Workshop
- Kein Bauen am eigenen Agenten
- Kein Compliance-Seminar, kein Tool-Pitch
- **Stattdessen: eine Einstellungsentscheidung.**
 - Wann reicht ChatGPT?
 - Wann braucht KI einen Arbeitsplatz?

1. Wer nutzt ChatGPT (oder ähnlich) mindestens wöchentlich in der Arbeit?
2. Wer hat Dokumente, die in einem Pflichtprozess landen?
 - Richtlinie
 - Vertrag
 - Bericht
 - Rundschreiben
 - Etc.
- **Wir stellen heute niemanden ein. Wir entscheiden, ob wir würden.**

These – Agenten sind Digitale Angestellte

KI-Agenten sind meistens normale Angestellte → Eine Neueinstellung bekommt nicht nur einen Laptop:



Stellenbeschreibung

Was ist der Job?
Wann ist der Job
„fertig“?



Zugriffsrechte

Was darf gelesen,
geändert, gesendet
werden?
Welche Daten dürfen
eingesehen werden?



Prozess

Übergaben
Ausnahmen
Vier-Augen
Protokolle



Vorgesetzte / Verantwortung

Wer steht für Fehler
ein?
Wer überwacht
Ergebnisse und
Abläufe?



Probezeit

Nachvollziehbar?
Gut genug?
Wird von der Arbeit
abgedrückt?

Aufgabe · Rechte · Ablauf · Rechenschaft · Qualität → wie bei menschlichen Angestellten

Einstellung nicht nur für einen Prozess

Text von außen (Beispiel NIS-2 / BSI)	Interne Richtlinie (Stand 2023, Mock)
Frühwarnung bei erheblichem Vorfall innerhalb von 24 Stunden	Vorfälle der IT-Leitung innerhalb von 5 Werktagen
Danach: 72-Stunden-Bericht	

Zwei Texte, ein Widerspruch. Das ist die Arbeit.

Nicht: NIS-2 erklären, dass kann auch das LLM im Browser. Sondern: die Lücke finden, vorschlagen, einen Menschen entscheiden lassen. NIS-2 ist nur das Aktenzeichen. Wir bauen keinen NIS-2-Agenten: wir üben die Einstellungsfrage.

1
24 h

2
5 Tage

Fristdifferenz = konkreter Jobbedarf → Stelle heißt Richtlinienprüfung

Digitaler Mitarbeitender für die Richtlinienprüfung bei neuen Vorgaben

Beispiel-Aktenzeichen heute: NIS-2. Die Stellenbezeichnung bleibt Richtlinienprüfung.

A

Aufgabe



Rundschreiben lesen · gilt
das für uns?
Abgleich
Lücken benennen
Änderung vorschlagen

R

Rechte



Öffentliche Texte frei
Richtlinie nur lesen, nicht
ändern
Nicht an die Behörde senden

P

Prozess



Ergebnis immer an die
nächste Stelle im Prozesse
Finale Freigabe &
Anwendung bleibt
menschlich

M

Probezeit / Mensch



Nachvollziehbar (Dokument,
Version, Vorschlag)
Ein Mensch steht für
Ergebnisse (und Fehler) ein

Dieselbe Arbeit in fünf Schritten



Schritte 1–4 = möglicher Arbeitsplatz
Schritt 5 = nicht verhandelbar
Das gilt für jedes Aktenzeichen. Heute üben wir es an NIS-2.

Drei Bewerber auf dieselbe Stelle

	Bewerber	Darf in dieser Stelle
A	Prompt im Browser	Schritt 1, öffentlicher Text, einmalig
B	Automatisierter Arbeitsablauf	Schritte 1+3, wiederkehrend, lesen, Ticket an Menschen
C	Mehrere Agenten	Schritte 2–4 wie ein Fachteam, Freigabe bleibt Mensch

Dieselbe Richtlinienprüfung. Drei mögliche Einstellungen.

A
Prompt

B
Ablauf
Einstellungen, keine Produktstufen

C
Fachteam

Wann der Prompt nicht mehr reicht

Aus Prompt wird Arbeitsplatz, wenn mindestens eines zutrifft:



Häufigkeit

Das Rundschreiben
kommt zyklisch immer
wieder



Vertraulichkeit

Interne Richtlinie, nicht
die öffentliche Website
→ darf nur von „innen
gesehen“ werden



Nachvollziehbarkeit

Welche Version wurde
geprüft?
Welche Handlung bzw.
Handlungsempfehlung
folgt daraus?



Systeme

Text liegt in der
internen Ablage



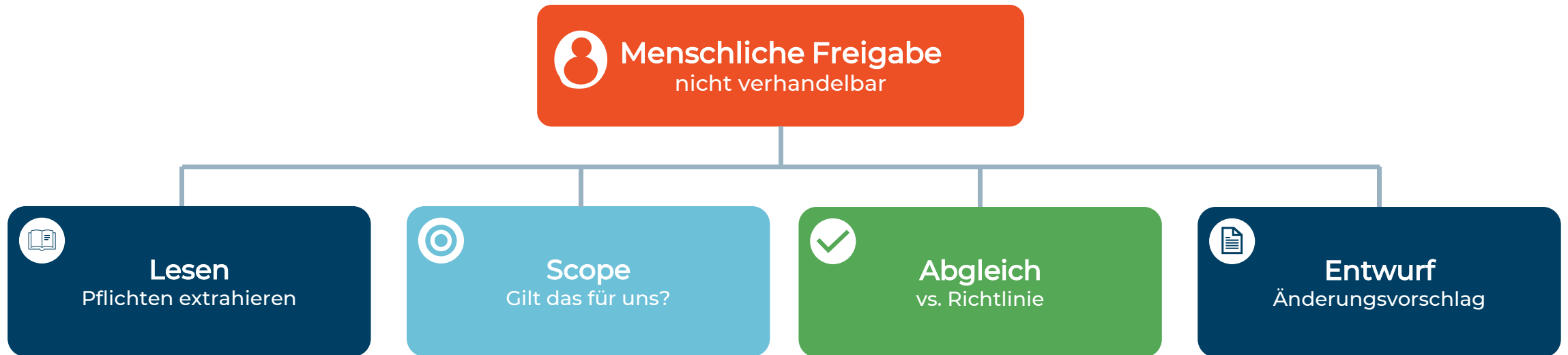
Freigabe

Ergebnis stößt eine
Änderung an

Das ist keine bessere Frage, sondern Infrastruktur...oder eben nicht.
Wiederkehrend + intern + nachvollziehbar = Arbeitsplatz!
Einmalig + extern + öffentliche Daten = „Praktikant“!

Multi-Agenten sind ein KI-Organigramm

Eine Person, Donnerstagnachmittag → der Arbeitsablauf reicht
Ein zweiter Agent ist eine zweite Rolle, kein stärkeres Modell.



Vier Rollen → Scope · Lesen · Abgleich · Entwurf + menschliche Freigabe

Nur einstellen, wenn diese Rollen bei den Menschen schon existieren UND auf KI-Agenten abbildbar sind.

-
- Wir lassen zwei Bewerber denselben Beispiel-Job machen.
 1. Vertrag zeigen (Aufgabe, Rechte, Prozesse)
 2. Bewerber 1: öffentlicher Text (NIS-2 / BSI), eine Frage
 3. Bewerber 2: derselbe Text gegen die interne Richtlinie (Mock)
 4. Was der Vertrag nicht erlaubt: keine interne Freigabe, keine Behördenmeldung
 - Mock-Richtlinie, echter öffentlicher Text

Sie sind der Einstellungsausschuss

- Einstellen
- Nur mit Aufsicht
- Nicht einstellen
- Dieselbe Stelle (Richtlinienprüfung). Drei Bewerber. Einmal Hände.



Einstellen



Aufsicht



Nicht

Bewerber A: Prompt im Browser

- Ein Mitarbeiter öffnet ChatGPT, fügt die öffentliche Seite zur Meldepflicht ein und fragt nach den Fristen.
- Die Antwort wird einmal gelesen. Nichts speichern, nichts intern, nichts nach außen.
- **Einstellen / nur mit Aufsicht / nicht einstellen?**



Bewerber B: Automatisierter Arbeitsablauf

- Jedes neue Rundschreiben wird gegen die interne Richtlinie geprüft.
 - Vertraulich
 - Versioniert
 - Automatisiert
- Lesen, nicht ändern → Lückenliste als Ticket an die Compliance-Leitung
- Freigabe bleibt beim Menschen
 - KI-Agent kann maximal Änderungsvorschläge geben
- **Einstellen / nur mit Aufsicht / nicht einstellen?**



Bewerber C: Vier Agenten, ein Prozess

- Scope : Lesen → Abgleich → Entwurf → Weitergabe an Menschen
 - Pro Aufgabe ein Agent
 - Jeder Agent ist Spezialist für seine Aufgabe → sieht nur relevante Daten
 - Agenten kommunizieren Daten, Outputs & Ergebnisse untereinander
 - Finale Freigabe bleibt beim Menschen.
- In Ihrem Haus macht diese Arbeit heute eine Person, nebenbei.
- **Einstellen / nur mit Aufsicht / nicht einstellen?**



In Keinem der drei: Richtlinie publizieren, Meldung an die Behörde.

- A: Praktikant reicht (ChatGPT)
- B: Arbeitsplatz, sobald der Job wiederkehrt und intern wird
- C: Fachteam nur, wenn das Organigramm schon da ist
 - Divide & Conquer → Prozess **SINNVOLL** in **Unterschritte** einteilen
 - Orchestratoragenten und Sub-Agenten einstellen
- **Als Nächstes: dieselbe Einstellungsfrage, anderer Dokumentenprozess.**

A

A · Praktikant

B

B · Arbeitsplatz

C

C · Fachteam

Die Stelle ist bereits besetzt: Demo ESG-Radar

Kein Themenwechsel von NIS-2 zu ESG. Dieselbe Einstellungsfrage an einem echten System.

Forschungskooperation Curacon & Steinbeis
ESG-Radar als Plattform zur **Nachhaltigkeitsberatung**
KIO-Agent: Liest Regularotik, Branchenspezifika und CSRD-Berichte → gezielte Fragestellungen beantworten
Benchmarking: erlaubt Einrichtungen, sich vergleichen zu lassen



1 Aufgabe

Dokument rein, strukturierte
Bewertung raus
KI-Agent nutzt Dokumente
zur Nachhaltigkeitsberatung



2 Rechte

Reines Leserecht bei
Dokumenten
Keine Schreibrechte →
Agent liest Daten nur aus,
kann sie nicht ändern



3 Prozess

KI-Agent ersetzt
Nachhaltigkeitsberater nicht,
er **unterstützt** sie



4 Vorgesetzte

Nachhaltigkeitsberater
tragen Haftung → KI-
Ergebnisse mit Bedacht
benutzen

Was Sie gesehen haben

	Beispiel-Job Richtlinienprüfung	ESG-Radar
Aufgabe	Lücke zwischen Vorgabe und Richtlinie	strukturierte Dokumentenbewertung
Rechte	lesen ja, ändern/sendern nein	begrenzter Dokumentenzugriff
Prozess	Ticket / menschliche Freigabe	menschlicher Checkpoint
Verantwortung	Compliance-Leitung / meldende Person	benannte fachliche Verantwortung

Zwei Prozesse. Dieselbe Einstellungsfrage. NIS-2 war nur das Übungsaktenzeichen.



Gleiche Einstellungsfrage, unterschiedliche Anwendung

- Der Agent sitzt am bestehenden Schreibtisch (Compliance, IT-Security, QM)
- Keine parallele „KI-Abteilung“
- Keine neuen Rollen erfinden, nur weil das Modell es kann
- Prozess zuerst, Agent danach
- **Wenn Sie die Stellenbeschreibung nicht schreiben können, bauen Sie keine Infrastruktur.**



- Bestes Beispiel für die technikaffinen unter Ihnen: **Software-Agenten!**

- ☐ Aufgabe ist **klar** definiert
- ☐ Rechte sind klar **eng & klar**: was darf nicht gelesen, geändert, gesendet werden?
- ☐ Prozess **existiert bereits**: Trigger, Übergabe, menschliche Freigabe
- ☐ **Human-in-the-responsibility**: Ein Mensch ist verantwortlich
- ☐ **Probezeit**: wie merken wir, dass es gut genug ist?
- ☐ ChatGPT reicht, solange: einmalig, öffentlich, Mensch stärker im Loop
- ☐ **Arbeitsablauf** lohnt sich, sobald: wiederkehrend, intern, nachvollziehbar
- ☐ Mehrere Agenten nur, wenn die **menschlichen Rollen schon geteilt sind**
- ☐ Rückkopplung: Menschliches Anpassen & Ändern von generierten Entwürfen

