

Auftrag an einen KI-Agenten mit Grenzen

Ein Blatt je Anwendungsfall. Wer alle sechs Blöcke ausfüllen kann, hat einen Agenten, der nichts anrichtet, was vorher niemand gesehen hat.

ANWENDUNGSFALL / ABTEILUNG	DATUM	AUFTRAGGEBER
----------------------------	-------	--------------

01 ZIEL UND AUSLÖSER ☐ Ziel in einem Satz: ☐ Auslöser (Postfach, Ticket, Zeitplan): ☐ Erwartetes Ergebnis (Entwurf, Meldung, Bericht): ☐ Häufigkeit je Woche (Schätzung):	02 ERLAUBTE QUELLEN UND SYSTEME ☐ System 1: lesend / schreibend ☐ System 2: lesend / schreibend ☐ System 3: lesend / schreibend ☐ Berechtigung im Namen von (Person / Rolle):
03 VERBOTENE HANDLUNGEN ☐ Nichts nach außen senden: ja / Ausnahme: ☐ Keine Werte schreiben in: ☐ Keine Preise / Zusagen unter oder über: ☐ Weitere Verbote:	04 ABBRUCHBEDINGUNGEN ☐ Abbruch bei fehlender Angabe: ☐ Abbruch ab Betrag / Umfang: ☐ Abbruch bei Unsicherheit zu: ☐ Abbruch wird gemeldet an:
05 FREIGABE UND PROTOKOLL ☐ Freigabe durch (Name / Rolle): ☐ Frist für die Freigabe: ☐ Protokoll enthält: Lesezugriffe / Annahmen / Abbrüche / Quellen ☐ Protokoll abgestimmt mit Betriebsrat am:	06 KENNZAHLEN NACH 4 WOCHEN ☐ Übernahmequote ohne Änderung (%): ☐ Häufigster Änderungsgrund: ☐ Abbruchquote (%) und berechtigt (ja / nein): ☐ Zeit je Vorgang vorher / nachher:

ABLAUF JE VORGANG

1 Auftrag	2 Plan	3 Ausführung	4 Freigabe	5 Protokoll
-----------	--------	--------------	------------	-------------



TIPPS VOM BYTE

- Geben Sie dem ersten Agenten einen Auftrag, den Sie in einem Satz aufschreiben können. Alles andere ist zu früh.
- Jede Handlung mit Außenwirkung, etwa eine Mail an einen Kunden, bekommt vorher einen Menschen zum Freigeben.
- Messen Sie erledigte Vorgänge pro Woche, vorher und nachher. Ein Agent, der Arbeit macht statt abnimmt, fliegt raus.

GEGENZEICHNUNG

Fachbereich	IT	Betriebsrat
-------------	----	-------------

Regel: Wer den Auftrag liest, muss ohne Rückfrage sagen können, was der Agent im schlimmsten Fall tut.