BOTFORCE

Praxisleitfaden

Vom Drehbuch zum Ziel.

Agentic AI im Unternehmen: Wie KI-Agenten Arbeit übernehmen – und was sie im Zaum hält

BOTFORCE GmbH, Wien

Mit einer durchgehenden Fallgeschichte, eingeordnet aus Sicht der Wirtschaftsinformatik

Stand

09/2026

6 KAPITEL

botforce-technology.com

ZU DIESEM TEXT

Warum ein Leitfaden – und wie er aufgebaut ist

Über KI-Agenten wird meist in zwei Tonlagen geschrieben: als Versprechen, dass Software bald ganze Abteilungen ersetzt, oder als Warnung vor Systemen, die niemand mehr kontrolliert. Dieser Text wählt eine dritte Tonlage, die der Wirtschaftsinformatik. Sie fragt nicht zuerst, was eine Technologie kann, sondern welche Aufgabe sie in einem Unternehmen übernimmt, wie sie mit Menschen, Prozessen und bestehenden Systemen zusammenspielt und woran man erkennt, ob sich ihr Einsatz lohnt.

Der Text folgt den sechs Blöcken des **Agentic AI Cheat Sheets**. Was dort in wenigen Zeilen steht, wird hier ausführlich erklärt. Jedes Kapitel beginnt mit einer **Szene** aus einer durchgehenden Fallgeschichte. Danach folgt die Sache selbst, eingeordnet in einem Kasten **Aus Sicht der Wirtschaftsinformatik**. Am Ende stehen ein **Merksatz** und **Fragen für euer Unternehmen**; Denkanstöße dazu finden sich im Anhang.

Die Fallgeschichte

Die **Aurelia Haustechnik GmbH** ist ein Großhändler für Sanitär- und Heizungstechnik mit Sitz in St. Pölten, 420 Mitarbeitenden, gut 900 Lieferanten und rund 4.000 Eingangsrechnungen im Monat. Durch die Geschichte führen:

Miriam Leitner	leitet das Rechnungswesen und wird Owner des ersten Agenten.
Jonas Hofer	ist Wirtschaftsinformatiker in der IT und verantwortet Prozesse und Automatisierung.
Sabine Kraus	bearbeitet seit elf Jahren Kreditorenrechnungen und kennt jeden schwierigen Lieferanten.
Karl Weninger	ist kaufmännischer Geschäftsführer und muss am Ende entscheiden.

Unternehmen, Personen, Lieferanten und Zahlen der Geschichte sind frei erfunden. Sie veranschaulichen Zusammenhänge und sind weder Messwerte noch Referenzen. Die Angaben zum EU AI Act geben den Stand nach dem Digital Omnibus wieder (September 2026) und sind keine Rechtsberatung.

Inhalt

PROLOG

Montag, 7:40 Uhr

Als Miriam Leitner an einem Montag im Februar ihren Rechner startet, sind übers Wochenende 312 Rechnungen eingegangen. Es ist kein schlechter Montag. Es ist ein gewöhnlicher.

Aurelia bezieht Ware von gut 900 Lieferanten, und jeder schickt seine Rechnungen auf seine Weise: als PDF im Anhang einer Mail, über ein Lieferantenportal, gelegentlich noch auf Papier. Seit 2019 holt ein Software-Roboter die Rechnungen jede Nacht aus den Portalen ab. Seit 2021 liest eine Dokumentenerkennung Lieferant, Betrag und Bestellnummer aus. Gut sechs von zehn Rechnungen buchen sich seither, ohne dass ein Mensch sie anfasst. Die übrigen landen bei Miriams Team – rund 1.500 im Monat.

„Die einfachen Fälle haben wir längst automatisiert", sagt Miriam, wenn man sie danach fragt. „Übrig bleiben die, bei denen man nachdenken muss." Eine fehlende Bestellnummer. Ein Lieferant, der seine Bankverbindung geändert hat. Eine Teillieferung, die zur Bestellung passt, aber nicht zum Lieferschein. Diese Fälle kosten Zeit, sie kosten Skonto, und sie kosten die Geduld von Menschen, die eigentlich anderes zu tun hätten.

An diesem Montag steht Jonas Hofer um kurz nach acht in ihrer Tür. Er hat einen Kaffee in der Hand und eine Frage im Kopf, über die er seit Wochen nachdenkt: „Was wäre, wenn wir die schwierigen Fälle nicht mehr nur weiterreichen, sondern bearbeiten lassen?"

Miriam lacht. „Von wem? Von deinem Roboter?"

„Nein", sagt Jonas. „Von etwas, das ein Ziel versteht."

Dieser Text erzählt, was in den folgenden zwölf Monaten geschieht – und erklärt an jeder Station, was dahintersteckt.

KAPITEL 01

Was ein Agent ist

WORUM ES GEHT

- Einen KI-Agenten präzise von Chatbot und Copilot abgrenzen.
- Die Schleife aus Wahrnehmen, Folgern, Planen, Handeln und Prüfen an einem Geschäftsvorfall erklären.
- Verstehen, warum die Delegation von Zielen eine organisatorische und nicht nur eine technische Frage ist.

SZENE · DIE RECHNUNG OHNE BESTELLNUMMER

Jonas legt eine Rechnung auf Miriams Schreibtisch, die das Team am Freitag liegen lassen musste: Kranzl Armaturen, 2.380 Euro, keine Bestellnummer. Sabine Kraus erklärt, wie sie den Fall lösen würde. Lieferantennummer im ERP nachschlagen. Die offenen Bestellungen bei Kranzl ansehen – es sind zwei. Betrag und Lieferdatum vergleichen. Die passende Bestellung zuordnen, buchen, eine Notiz hinterlassen, warum. Zehn Minuten, wenn niemand stört.

„Das sind sechs Schritte“, sagt Jonas, „und keiner davon steht so in einer Arbeitsanweisung. In der Arbeitsanweisung steht: Rechnung sachlich prüfen und buchen. Sabine kennt das Ziel und findet den Weg selbst.“

„Unser Roboter kennt auch einen Weg“, sagt Miriam.

„Einen einzigen“, sagt Jonas. „Den, den wir ihm aufgeschrieben haben.“

Ziel statt Schrittfolge

Ein **KI-Agent** ist ein Softwaresystem, das ein Ziel statt einer Schrittfolge erhält und dieses Ziel in einer Schleife verfolgt. Klassische Automatisierung arbeitet prozedural: Jemand beschreibt vorab jeden Schritt, und die Software führt ihn aus. Ein Agent arbeitet zielorientiert: Er bekommt einen gewünschten Endzustand – „die Rechnung ist richtig gebucht und belegt“ – und bestimmt die Schritte dorthin für jeden Fall selbst, innerhalb der Grenzen, die ihm gesetzt wurden.

Möglich ist das erst, seit große Sprachmodelle Dokumente lesen, Absichten erkennen und Zwischenschritte formulieren können (Kapitel 3). Das Modell allein ist aber noch kein Agent. Zum Agenten wird es durch Werkzeuge, mit denen es in echten Systemen handelt, und durch die Schleife, in der es sein eigenes Ergebnis prüft.

Die Schleife

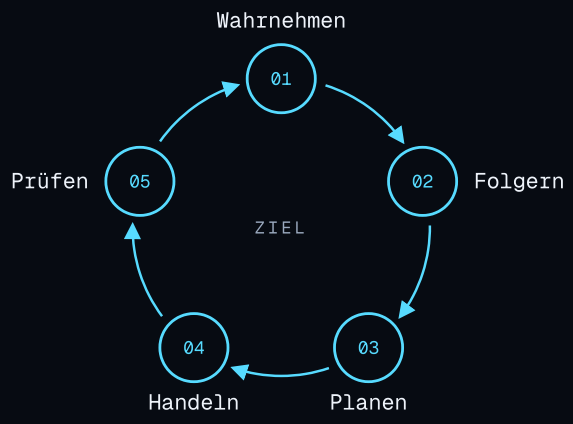

ABB. 1 - DIE AGENTEN-SCHLEIFE: FÜNF SCHRITTE RUND UM EIN ZIEL.

Die fünf Schritte beschreiben, wie jede zielgerichtete Sachbearbeitung funktioniert. Am Fall aus der Szene:

01	**Wahrnehmen**	Der Agent liest die Rechnung und stellt fest, dass die Bestellnummer fehlt.
02	**Folgern**	Über die Lieferantennummer lassen sich die offenen Bestellungen des Lieferanten finden.
03	**Planen**	Er fragt die offenen Bestellungen im ERP ab und legt fest, woran er die richtige erkennt: Betrag, Menge, Lieferdatum.
04	**Handeln**	Er ordnet die passende Bestellung zu und bereitet die Buchung vor – oder führt sie aus, wenn er das darf.
05	**Prüfen**	Er gleicht das Ergebnis mit dem Lieferschein ab und hält fest, warum er sich für diese Bestellung entschieden hat.

Passt das Ergebnis nicht, beginnt die Schleife von vorn. Sie endet, wenn das Ziel erreicht ist oder wenn eine Regel den Agenten anhalten lässt – etwa weil er zwei gleich gute Kandidaten findet und nicht raten soll. Dann geht der Fall mit seiner Begründung an einen Menschen.

Chatbot, Copilot, Agent

Im Alltag werden drei Begriffe oft vermischt. Sie beschreiben unterschiedliche Grade von Selbstständigkeit:

	WAS ER TUT	BEISPIEL BEI AURELIA	WAS ER ZUSÄTZLICH BRAUCHT
Chat-bot	beantwortet Fragen, handelt nicht	„Welche Skontofrist gilt bei Kranzl?“	Wissen und Quellen
Copilot	bereitet vor, der Mensch führt aus	markiert die passende Bestellung, Sabine bucht	Lesezugriff auf Daten
Agent	führt selbst aus und meldet sich mit Ergeb-nis oder Ausnahme	ordnet zu, bucht und dokumen-tiert die Wahl	eigene Identität, Grenzen, Protokoll

TAB. 1 - DREI GRADE VON SELBSTSTÄNDIGKEIT.

Zwei Fragen helfen bei der Einordnung. **Wer legt die Schritte fest?** Wählt das System Weg und Werkzeuge selbst, ist es ein Agent. **Wer drückt auf „Ausführen“?** Solange ein Mensch jede Aktion bestätigt, trägt er die Verantwortung für jeden Schritt; sobald das System selbst handelt, muss die Organisation diese Verantwortung auf andere Weise absichern. Davon handeln die Kapitel 4 und 5.

AUS SICHT DER WIRTSCHAFTSINFORMATIK

Der Agent als Beauftragter

Das Wort Agent stammt nicht aus der Informatik. In der Betriebswirtschaftslehre beschreibt die **Prinzipal-Agent-Theorie** Beziehungen, in denen ein Auftraggeber (Prinzipal) eine Aufgabe an einen Beauftragten (Agent) delegiert. Der Beauftragte weiß mehr über seine eigene Arbeit als der Auftraggeber, und seine Ziele decken sich nicht automatisch mit dessen Zielen. Die Theorie kennt dafür Gegenmittel: klare Verträge, Kontrollen, Berichtspflichten.

Bei Software-Agenten kehren beide Probleme wieder. Niemand kann jeden Zwischenschritt beobachten, und ein Sprachmodell optimiert auf plausible Antworten, nicht auf die Ziele des Unternehmens. Die Gegenmittel heißen hier Policy, Audit-Trail und menschliche Aufsicht. Wer einen Agenten einführt, gestaltet deshalb nicht nur ein technisches System, sondern eine **Delegationsbeziehung** – ein Gedanke, den die Wirtschaftsinformatik mit ihrem Blick auf Mensch, Technik und Organisation als Einheit gut fassen kann.

MERKSATZ

Ein Agent bekommt ein Ziel, keinen Ablauf. Damit verschiebt sich Verantwortung – und die muss die Organisation neu regeln.

FRAGEN FÜR EUER UNTERNEHMEN

1.1 Grenzt Chatbot, Copilot und Agent anhand der Frage „Wer führt aus?“ voneinander ab und nennt je ein Beispiel aus dem Rechnungswesen.

1.2 Beschreibt die fünf Schritte der Agenten-Schleife an einem Vorgang aus eurem eigenen Arbeitsumfeld.

1.3 Welche Probleme der Prinzipal-Agent-Theorie treten bei Software-Agenten auf, und welche Mechanismen begegnen ihnen?

KAPITEL 02

Skript, RPA oder Agent?

WORUM ES GEHT

- Skripte und Schnittstellen, RPA und Agenten als Formen der Automatisierung unterscheiden.
- Die Prozessmerkmale benennen, die über die passende Technik entscheiden.
- Erklären, warum Process Mining und Prozesskennzahlen vor jeder Automatisierungsentscheidung stehen.

SZENE · DER BOT, DER STEHEN BLIEB

Jonas erinnert sich gut an den Oktober davor. Das größte Lieferantenportal hatte über Nacht sein Anmeldeformular umgebaut. Der Roboter, der dort jede Nacht Rechnungen abholte, fand das Passwortfeld nicht mehr und brach ab. Er meldete den Fehler korrekt – in ein Postfach, das niemand las. Drei Nächte lang. Rund 400 Rechnungen kamen verspätet in die Buchhaltung, bei 23 davon war die Skontofrist verstrichen.

„Ich dachte, das ist automatisiert“, hatte Karl Weninger damals gesagt.

„Ist es“, hatte Jonas geantwortet. „Nur die Überraschung nicht.“

Jetzt, im Februar, will Karl wissen, ob ein Agent dieselbe Überraschung wäre, nur teurer.

Ein Spektrum, keine Wahl zwischen Alt und Neu

Automatisierung ist keine einzelne Technik, sondern ein Spektrum mit drei Bereichen.

Skripte und Schnittstellen verbinden Systeme über definierte Wege (APIs). Sie sind schnell, günstig und robust – vorausgesetzt, die beteiligten Systeme bieten solche Wege an und die Regeln lassen sich vollständig beschreiben.

Robotic Process Automation (RPA) setzt dort an, wo Schnittstellen fehlen. Software-Roboter bedienen Anwendungen über deren Oberfläche, wie ein Mensch es täte: Sie klicken, tippen, lesen Bildschirminhalte. Das macht RPA schnell einführbar. Es macht sie aber auch empfindlich gegen jede Änderung einer Maske und blind für jeden Fall, den das Drehbuch nicht vorsieht.

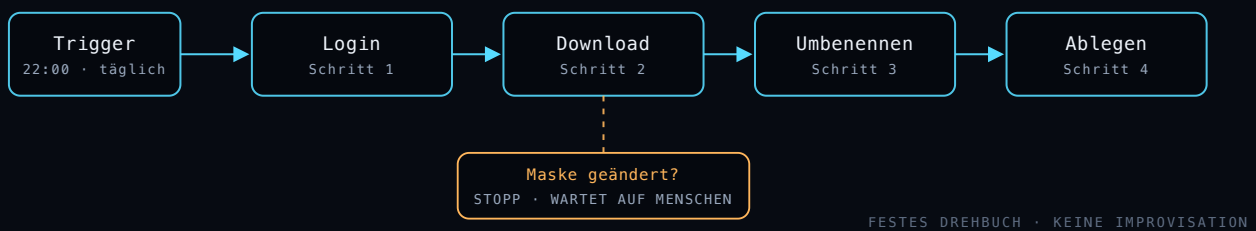

ABB. 2 – EIN RPA-BOT SPIELT SEIN DREHBUCH AB; DIE UNERWARTETE MASKE STOPPT IHN.

Agenten verfolgen Ziele statt Drehbücher. Sie kommen mit Varianten, unstrukturierten Dokumenten und unvollständigen Angaben zurecht. Dafür sind sie im Betrieb teurer, arbeiten probabilistisch und brauchen eine eigene Governance.

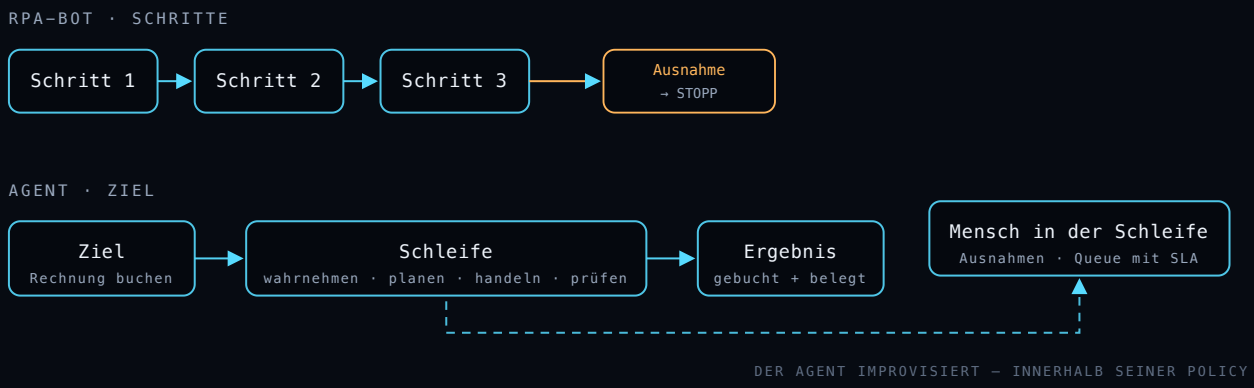

ABB. 3 - DREHBUCH GEGEN ZIEL: WO DER BOT STOPPT, ESKALIERT DER AGENT KONTROLLIERT.

Welche Technik passt?

Die Wahl folgt aus den Merkmalen des Prozesses, nicht aus der Begeisterung für eine Technik:

MERKMAL	SKRIPT / API	RPA	AGENT
Schnittstellen	vorhanden	fehlen	vorhanden oder über Werkzeuge erreichbar
Regeln	fest und vollständig beschreibbar	fest	teilweise; Urteil im Einzelfall nötig
Varianten	wenige	wenige, Oberfläche stabil	viele
Daten	strukturiert	strukturiert, am Bildschirm	auch unstrukturiert: Mails, PDFs, Freitext
Ausnahmen	selten	selten	häufig, aber mit erkennbaren Mustern
Laufende Kosten je Fall	sehr niedrig	niedrig	höher: Modellaufrufe, Aufsicht
Typisches Risiko	Schnittstelle ändert sich	Oberfläche ändert sich	Fehlurteil ohne Kontrolle

TAB. 2 - PROZESSMERKMALE UND PASSENDE AUTOMATISIERUNG.

In der Praxis schließen sich die Formen nicht aus, sie ergänzen sich. Bei Aurelia bleibt die Übergabe an den Zahlungsverkehr ein Skript. Der Download aus zwei Portalen ohne Schnittstelle bleibt RPA – diesmal mit einer Überwachung, die jemand liest. Der Agent übernimmt die Fälle, an denen beide scheitern, und ruft den RPA-Roboter bei Bedarf sogar als Werkzeug auf.

Wann kein Agent?

Drei Warnzeichen sprechen gegen einen Agenten, zumindest vorerst:

- **Keine echten Fälle zum Testen.** Ohne historische Vorgänge mit bekannten richtigen Ergebnissen lässt sich die Qualität nicht messen (Kapitel 3).
- **Niemand ist für Ausnahmen zuständig.** Ein Agent gibt Fälle an Menschen zurück. Liegen sie dort ungelesen, ist nichts gewonnen – das Postfach aus dem Oktober lässt grüßen.
- **Der Prozess ist selten.** Wo zehn Fälle im Monat anfallen, lohnen sich weder Aufbau noch Aufsicht.

AUS SICHT DER WIRTSCHAFTSINFORMATIK

Erst messen, dann automatisieren

Die Wirtschaftsinformatik beurteilt Automatisierung über **Prozesskennzahlen**. Für die Rechnungsbearbeitung sind das vor allem die Durchlaufzeit, die Kosten je Vorgang und die **Dunkelverarbeitungsquote** (Straight-Through Processing, STP): der Anteil der Fälle, die ohne menschliche Berührung durchlaufen. Bei Aurelia liegt sie bei 62 Prozent. Jede Automatisierungsstufe ist ein Versuch, diese Quote zu erhöhen, ohne das Fehler- und Kontrollrisiko zu steigern.

Wie ein Prozess tatsächlich läuft, zeigt **Process Mining**: Es rekonstruiert den Ablauf aus den Ereignisprotokollen der Systeme – Zeitstempel, Statuswechsel, Bearbeiter. Als Jonas die Rechnungsdaten eines Jahres auswertet, findet er, dass 71 Prozent der manuellen Fälle auf fünf Ausnahmemuster zurückgehen: fehlende Bestellnummer, Preisabweichung, Teillieferung, geänderte Stammdaten, unleserliche Belege. Die Ausnahmen sind vielfältig, aber nicht beliebig. Das ist die eigentliche Grundlage des Business Case.

RECHENBEISPIEL · ANNAHMEN DER FALLGESCHICHTE

1.500 manuelle Rechnungen im Monat × 9 Minuten = 225 Arbeitsstunden. Schließt ein Agent 60 Prozent dieser Fälle ab, entfallen rund 135 Stunden im Monat. Dem stehen Modellaufrufe, Betrieb, Evals, die verbleibende Aufsicht und die Einführung selbst gegenüber. Die wichtigere Frage stellt Karl Weninger: „Wofür nutzen wir die Zeit?“

MERKSATZ

Die Technik folgt dem Prozess: Schnittstelle, wo es geht; RPA, wo nichts anderes geht; ein Agent, wo Urteil gefragt ist – und nur dort, wo jemand die Ausnahmen verantwortet.

FRAGEN FÜR EUER UNTERNEHMEN

2.1 Warum ist RPA schnell einzuführen, aber anfällig im Betrieb?

2.2 Ordnet drei Prozesse aus eurem Unternehmen mithilfe von Tabelle 2 einer Automatisierungsform zu und begründet die Wahl.

2.3 Welche Rolle spielen Process Mining und die Dunkelverarbeitungsquote für den Business Case eines Agenten?

KAPITEL 03

Die Bausteine

WORUM ES GEHT

- Die vier Bausteine Modell, Kontext, Werkzeuge und Evaluation beschreiben.
- Erklären, warum Sprachmodelle keine Wahrheit garantieren – auch nicht mit Reasoning – und wie Retrieval-Augmented Generation damit umgeht.
- Die Werkzeugschicht als wichtigste Kontrollfläche eines Agenten begreifen.

SZENE · DAS WHITEBOARD

Karl Weninger will verstehen, was er genehmigen soll, bevor er es genehmigt. Jonas zeichnet vier Kästen an das Whiteboard im Besprechungsraum.

„Das hier ist das Modell. Es versteht Sprache, aber es weiß nichts über Aurelia. Hier kommt der Kontext hinein: unsere Verträge, Bestellungen, Arbeitsanweisungen. Das hier sind die Hände – Werkzeuge, mit denen der Agent im ERP nachsehen und buchen darf, und sonst nichts." Er tippt auf den vierten Kasten. „Und das ist der Grund, warum ich nachts schlafen kann. Wir messen, bevor wir ihm etwas anvertrauen."

Karl sieht eine Weile auf die Kästen. „Und welcher davon macht Fehler?"

„Der erste", sagt Jonas. „Die anderen drei sind dafür da, dass wir es merken."

Das Modell: Verständnis ohne Garantie

Den Kern bildet ein **großes Sprachmodell** (Large Language Model, LLM). Es wurde auf riesigen Textmengen darauf trainiert, das jeweils nächste Wort vorherzusagen, und hat dabei gelernt, Sprache zu verstehen, Wissen abzurufen und Anweisungen zu folgen. Heutige Spitzenmodelle werden danach zusätzlich mit Reinforcement Learning trainiert, Aufgaben in Zwischenschritten zu durchdenken – das macht sie bei mehrstufigen Aufgaben deutlich stärker. Genauso wichtig ist, was ein Sprachmodell nicht ist: keine Datenbank, kein Rechenwerk und keine Quelle von Garantien. Wo Wissen fehlt, ist eine plausible Vermutung für das Modell oft naheliegender als ein „weiß ich nicht".

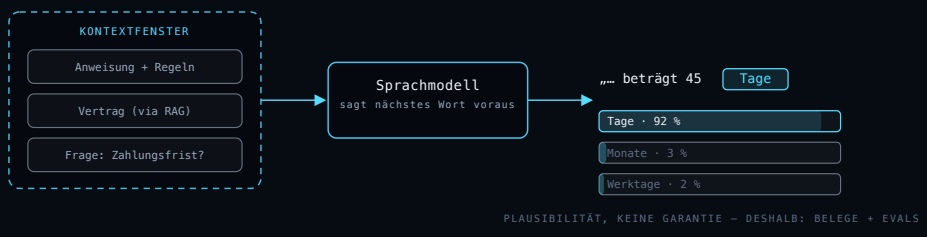

ABB. 4 – EIN LLM ERZEUGT DIE ANTWORT WORT FÜR WORT AUS WAHRSCHEINLICHKEITEN ÜBER SEINEM KONTEXT.

Kontext und RAG: die Fakten des Unternehmens

Alles, was das Modell für eine Aufgabe zu sehen bekommt – Anweisung, Regeln, Dokumente –, steht in seinem **Kontextfenster**. Was dort nicht steht, existiert für das Modell in diesem Moment nicht. Die Fenster fassen heute oft eine Million Token und mehr, verlässlich genutzt wird davon aber weniger – deshalb heißt die

Disziplin **Context Engineering**: das Richtige zeigen, nicht möglichst viel. Fragt man ein Modell ohne Zugriff auf den Liefervertrag nach der Zahlungsfrist, antwortet es womöglich „30 Tage“, weil diese Frist in seinen Trainingsdaten häufig vorkommt. Eine solche überzeugt vorgetragene Falschaussage heißt **Halluzination**.

Retrieval-Augmented Generation (RAG) begegnet dem, indem das System vor jeder Antwort die passenden Passagen aus den eigenen Dokumenten sucht und dem Modell in den Kontext legt. Dann zitiert es die echte Klausel: 45 Tage, Abschnitt 7.2. Für Unternehmen gilt als Faustregel: erst Anweisungen und RAG verbessern, ein Modell auf eigenen Daten nachtrainieren (Fine-Tuning) erst dann, wenn es nachweislich nötig ist.

Werkzeuge: die Hände des Agenten

Ein Modell allein kann nur Text erzeugen. **Werkzeuge** (Tools) geben ihm definierte Fähigkeiten: eine Bestellung im ERP abfragen, eine Buchung vorschlagen, einen RPA-Roboter starten, eine Mail an einen Lieferanten senden. Beim **Function Calling** wird jedes Werkzeug mit Namen, Zweck und Parametern beschrieben; das Modell entscheidet, wann es welches mit welchen Werten aufruft. Das **Model Context Protocol** (MCP), seit Dezember 2025 herstellerneutral unter dem Dach der Linux Foundation, vereinheitlicht, wie Systeme ihre Werkzeuge anbieten; mit A2A übergeben Agenten einander Aufgaben.

Für die Governance ist diese Schicht zentral. Was nicht als Werkzeug freigegeben ist, kann der Agent nicht tun, und eine Grenze, die ein Werkzeug technisch durchsetzt, kann er nicht wegdiskutieren. Vorsicht gilt bei Allzweck-Werkzeugen wie Bildschirm- oder Browsersteuerung („Computer Use“) und Code-Ausführung: Sie öffnen alles, was das Konto des Agenten darf.

Evals: Qualität als Messwert

Weil Sprachmodelle probabilistisch arbeiten, tritt die **Evaluation** (Eval) an die Stelle des klassischen Softwaretests: eine Sammlung realer Fälle mit bekannten richtigen Ergebnissen, gegen die das System nach jeder Änderung automatisch geprüft wird – nach jeder neuen Anweisung, jedem Modellwechsel, vor jeder Freigabe. Bei Agenten zählt dabei auch der Weg – welche Werkzeuge mit welchen Werten aufgerufen wurden –, und jeder Fall läuft mehrfach, weil derselbe Fall unterschiedlich ausgehen kann.

Jonas stellt aus den Rechnungen des Vorjahres ein solches **Golden Set** aus 1.200 Fällen zusammen, bewusst mit allen fünf Ausnahmemustern. Die erste Version des Agenten bearbeitet 96,4 Prozent davon korrekt. Das klingt gut, bis man die 43 Fehler einzeln ansieht: Bei 19 davon hätte er eine Teillieferung falsch zugeordnet. Die Liste wird zur Arbeitsgrundlage der nächsten Wochen.

AUS SICHT DER WIRTSCHAFTSINFORMATIK

Ein Agent als Schichtenarchitektur

Betriebliche Informationssysteme werden in der Wirtschaftsinformatik gern in Schichten beschrieben. Das hilft auch hier:

- **Wissensschicht** – Kontext, RAG, Stammdaten: worüber der Agent urteilt.
- **Verarbeitungsschicht** – Modell und Planung: wie er urteilt.
- **Integrationsschicht** – Werkzeuge, Schnittstellen, RPA: was er bewirken kann.
- **Kontrollschicht** – Evals, Policy, Protokoll, Aufsicht: ob man ihm trauen darf.

Die Trennung hat einen handfesten Nutzen. Sprachmodelle werden in kurzen Abständen ersetzt. Wer Wissen, Werkzeuge und Kontrollen sauber vom Modell trennt, kann das Modell tauschen und mit denselben Evals nachweisen, dass der Nachfolger mindestens so gut arbeitet. Das Modell wird zur austauschbaren Komponente, das Unternehmenswissen bleibt im Haus.

MERKSATZ

Das Modell liefert Verständnis, der Kontext die Fakten, die Werkzeuge die Wirkung und die Evals den Nachweis. Fehlt einer der vier Bausteine, fehlt dem Agenten etwas Wesentliches.

FRAGEN FÜR EUER UNTERNEHMEN

3.1 Warum halluzinieren Sprachmodelle, und wie verringert RAG dieses Risiko?

3.2 Weshalb gilt die Werkzeugschicht als wichtigste Kontrollfläche eines Agenten?

3.3 Wie würdet ihr ein Golden Set für einen Agenten in eurem Fachbereich zusammenstellen?

KAPITEL 04

Der Governance-Ring

WORUM ES GEHT

- Die vier Kontrollen Identität, Policy, Audit und Aufsicht erklären.
- Begründen, warum Regeln technisch durchgesetzt werden müssen und nicht nur im Prompt stehen dürfen.
- Human-in-the-Loop als Gestaltung eines Prozesses verstehen, nicht als Pauschalkontrolle.

SZENE · 12.400 EURO

Dienstag, 14:02 Uhr, dritte Woche im Pilotbetrieb. Eine Rechnung von Vettori Pumpen geht ein: 12.400 Euro. Der Agent erkennt Lieferant, Bestellung und Lieferschein, alles passt zusammen. Er könnte buchen.

Er bucht nicht. Eine Sekunde später steht der Fall in Sabine Kraus' Arbeitsliste – mit der zugeordneten Bestellung, dem Buchungsvorschlag und einem Satz Begründung: „Betrag über dem Limit von 10.000 Euro für autonome Buchungen."

Sabine öffnet den Fall nach ihrer Besprechung, sieht sich den Lieferschein an und gibt frei. Der Agent bucht. Im Protokoll stehen vier Zeilen. Zwischen der ersten und der letzten liegen knapp 30 Minuten, und niemand musste jemanden anrufen.

„Früher", sagt Sabine später zu Miriam, „hätte ich die Rechnung erst gesucht, dann geprüft, dann gebucht. Jetzt prüfe ich nur noch. Das ist der Teil, für den ich eigentlich da bin."

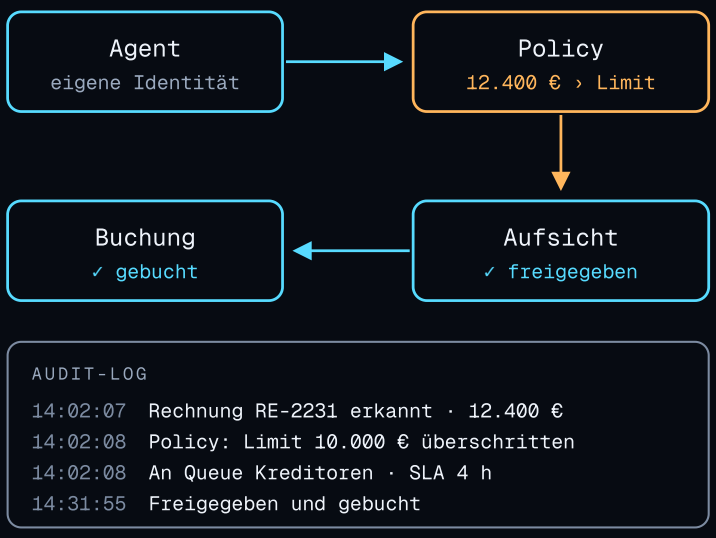

ABB. 5 – DIE POLICY GREIFT, EIN MENSCH GIBT FREI, JEDER SCHRITT STEHT IM AUDIT-LOG.

Handlungsfähigkeit braucht Grenzen

Ein Agent, der in echten Systemen handelt, ist so wertvoll wie riskant. Unternehmenstaugliche Agenten stehen deshalb in einem Ring aus vier Kontrollen, die einander ergänzen.

Identität

Jeder Agent erhält eine eigene Identität im Identitätssystem des Unternehmens – die Anbieter führen Agenten inzwischen als eigene Identitätsart – mit einer verantwortlichen Person, nie einen geteilten Service-Account. Seine Berechtigungen sind eng geschnitten: Der Kreditoren-Agent darf Kreditorenbelege lesen und Buchungen anlegen, aber keine Lieferantenstammdaten ändern. Seine Zugangsdaten sind kurzlebig und verfallen nach Minuten statt nach Monaten. So bleibt jederzeit nachvollziehbar, wer gehandelt hat, und Berechtigungsprüfungen funktionieren wie bei menschlichen Mitarbeitenden.

Policy

Die Grenzen des Agenten sind als Regeln festgelegt: Buchungen bis 10.000 Euro autonom, darüber Vier-Augen-Freigabe; Änderungen von Bankverbindungen nie autonom; Zugriff nur auf Kreditoren-Vorgänge. Entscheidend ist, **wo** diese Regeln durchgesetzt werden. Eine Regel im Prompt ist eine Bitte an ein probabilistisches Modell, die es missverstehen kann. Eine Regel in der Werkzeugschicht ist eine Tatsache: Das Buchungswerkzeug lehnt ab, egal wie überzeugend der Agent argumentiert. Geändert werden Policies wie Software – mit Review und Versionsgeschichte.

Besonders wichtig ist das wegen **Prompt Injection**: Anweisungen, die in Mails, PDFs oder Webseiten versteckt sind, die der Agent liest. Sprachmodelle können Anweisungen und Daten nicht zuverlässig trennen, und Fachstellen halten das Problem für nicht vollständig lösbar. Entscheidend ist deshalb, was ein getäuschter Agent tun kann – und das legen Werkzeuge und Policy fest, nicht der Prompt.

Audit

Jede Eingabe einschließlich fremder Inhalte, jede Entscheidung samt Begründung, jeder Werkzeugaufruf und jedes Ergebnis wird manipulationssicher protokolliert, idealerweise hash-verkettet, sodass nachträgliche Änderungen auffallen. Die Begründung allein genügt nicht – was ein Modell als Grund ausgibt, ist nicht immer der wahre Grund. Der Audit-Trail beantwortet die Frage, die Wirtschaftsprüfung, Revision und Geschäftsführung stellen werden: *Warum hat das System das getan?* Zugleich ist er das beste Werkzeug zur Fehleranalyse, weil sich jeder Fall Schritt für Schritt nachvollziehen lässt.

Aufsicht

Human-in-the-Loop heißt nicht, dass ein Mensch alles ansieht. Gut gestaltet landen Ausnahmen in einer Warteschlange mit Service-Level (SLA), zusammen mit dem Fall, der Begründung und dem Vorschlag des Agenten. Menschen entscheiden die Fälle, die Urteilskraft brauchen, statt Routine abzutippen. Dazu gehört ein Stopp-Schalter je Agent, der außerhalb des Modells greift – Identität sperren, Tokens entziehen – und den der Fachbereich selbst bedienen darf, nicht erst die IT nach einem Ticket.

AUS SICHT DER WIRTSCHAFTSINFORMATIK

Ein internes Kontrollsystem für einen neuen Akteur

Für Rechnungswesen und Wirtschaftsinformatik ist der Governance-Ring kein Fremdkörper. Er überträgt bewährte Prinzipien des **internen Kontrollsystems** (IKS) auf eine neue Art von Akteur:

- **Funktionstrennung:** Wer bucht, ändert keine Stammdaten.
- **Vier-Augen-Prinzip:** ab definierten Schwellen verpflichtend.
- **Belegfunktion:** Jeder Vorgang ist lückenlos dokumentiert.
- **Berechtigungsprüfung:** regelmäßig, auch für Agenten.

Neu ist das Tempo. Ein Agent kann in einer Stunde so viele Vorgänge anstoßen wie ein Team in einer Woche. Kontrollen, die bei Menschen als Stichprobe genügen, müssen bei Agenten deshalb vollständig und automatisch wirken.

MERKSATZ

Vertrauen in einen Agenten entsteht nicht aus seiner Intelligenz, sondern aus seinen Grenzen: eigene Identität, Regeln in der Werkzeugschicht, ein lückenloses Protokoll und Menschen, die Ausnahmen entscheiden.

FRAGEN FÜR EUER UNTERNEHMEN

4.1 Warum reicht es nicht, eine Buchungsgrenze im Prompt festzuhalten?

4.2 Erklärt anhand der Szene, wie die vier Kontrollen zusammenwirken.

4.3 Welche Prinzipien des internen Kontrollsystems lassen sich direkt auf Agenten übertragen, und wo braucht es Anpassungen?

KAPITEL 05

Autonomie in Stufen

WORUM ES GEHT

- Die Autonomiestufen L0 bis L4 unterscheiden.
- Kennzahlen nennen, die den Wechsel auf eine höhere Stufe rechtfertigen.
- Die organisatorischen Voraussetzungen für den Betrieb von Agenten beschreiben.

SZENE · DIE ERSTEN ACHT WOCHEN

Karl Weninger hat für den Pilotbetrieb eine Bedingung gestellt: „Am Anfang bucht der Agent nichts, was nicht ein Mensch gesehen hat." Jonas ist das recht, Miriam auch.

Acht Wochen lang schlägt der Agent deshalb nur vor. Sabine und zwei Kolleginnen bestätigen jede Buchung mit einem Klick oder korrigieren sie. Jede Korrektur wandert als neuer Fall ins Golden Set. In Woche drei stimmen die Vorschläge in 94 Prozent der Fälle mit der Entscheidung der Menschen überein, in Woche acht in 98,1 Prozent. Die Ausnahmen in der Warteschlange werden im Schnitt nach zweieinhalb Stunden bearbeitet; das vereinbarte Service-Level liegt bei vier.

Dann sitzen Miriam, Jonas und Karl zusammen und entscheiden: Rechnungen bis 10.000 Euro mit eindeutig zugeordneter Bestellung bucht der Agent ab sofort selbst. Rechnungen mit Preisabweichung schlägt er weiter nur vor. „Wir schalten nicht die Autonomie ein", sagt Miriam. „Wir schalten eine Fallgruppe frei."

Eine Treppe, keine Schwelle

Autonomie ist kein Schalter, der einmal umgelegt wird. Sie wächst in Stufen:

L0	**Manuelle Arbeit**	Menschen erledigen alle Schritte.
L1	**Skriptbasierte Automatisierung**	RPA und Skripte führen feste Abläufe aus, ohne zu urteilen.
L2	**Intelligente Automatisierung**	Machine Learning im Prozess, etwa zum Auslesen von Dokumenten.
L3	**Assistierte Agenten**	Der Agent schlägt vor, ein Mensch bestätigt jede Aktion.
L4	**Kontrollierte Autonomie**	Der Agent handelt in Grenzen, Menschen entscheiden die Ausnahmen.

TAB. 3 - AUTONOMIESTUFEN L0 BIS L4.

Der Weg nach oben ist kein Technik-, sondern ein Vertrauensaufbau. Er erfolgt je Prozess und oft je Fallgruppe: Bei Aurelia arbeitet der Agent für eindeutige Rechnungen bis 10.000 Euro auf L4, für Rechnungen mit Preisabweichung weiter auf L3. Eine Stufe höher geht es erst, wenn Evals, Audit-Trail und Ausnahmequote es rechtfertigen – nicht, weil ein Projektplan es vorsieht. Die Stufen sind ein Reifemodell für Prozesse, keine Norm.

Woran man den richtigen Zeitpunkt erkennt

- **Übereinstimmungsquote:** Wie oft entscheidet der Agent im L3-Betrieb so wie die Menschen?
- **Eval-Ergebnis:** stabil auf dem Golden Set, ohne Rückschritt gegenüber der Vorversion.
- **Ausnahmequote und Durchlaufzeit:** Kann das Team die Warteschlange innerhalb des Service-Levels bearbeiten?
- **Fehler nach der Buchung:** Stornos oder Korrekturen je 1.000 Fälle, verglichen mit dem manuellen Ablauf.

Betrieb braucht Rollen

Ein Agent ohne Verantwortliche ist ein Vorfall in Wartestellung. Jeder Agent hat deshalb einen benannten **Owner** im Fachbereich, der für die Ergebnisse geradesteht – bei Aurelia ist das Miriam. Die Kapazität für Human-in-the-Loop wird eingeplant wie jede andere Arbeit. Fehlerfälle werden in einem festen Takt besprochen und ins Golden Set übernommen. Und das Team übt den Wechsel auf ein neues Modell, bevor ein Anbieter ihn erzwingt.

AUS SICHT DER WIRTSCHAFTSINFORMATIK

Reife ist so hoch wie die schwächste Dimension

Reifegradmodelle sind in der Wirtschaftsinformatik ein vertrautes Werkzeug, um Entwicklungspfade von Organisationen zu beschreiben. Für den Betrieb von Agenten zählen fünf Dimensionen: **Evals**, **Governance**, **Betrieb**, **Rollen** und **kontinuierliche Verbesserung**. Eine Organisation ist nur so reif wie ihre schwächste Dimension: Der beste Agent nützt wenig, wenn niemand die Ausnahmen bearbeitet.

Ebenso wichtig ist die Gestaltung der Arbeit. Sachbearbeitung verschiebt sich vom Erfassen zum Entscheiden. Das ist für viele attraktiver, verlangt aber andere Kompetenzen und muss begleitet werden – durch Schulung, klare Zuständigkeiten und eine ehrliche Kommunikation darüber, was sich für wen ändert.

MERKSATZ

Autonomie wird verdient, nicht eingeschaltet: Stufe für Stufe, je Prozess und Fallgruppe, belegt durch Evals, Protokoll und eine tragbare Ausnahmequote.

FRAGEN FÜR EUER UNTERNEHMEN

5.1 Worin unterscheiden sich L3 und L4, und welche Nachweise braucht der Wechsel?

5.2 Warum kann ein und derselbe Prozess gleichzeitig auf zwei Stufen laufen?

5.3 Welche Rollen braucht der Betrieb eines Agenten, und was geschieht, wenn eine davon fehlt?

KAPITEL 06

EU AI Act: die Einordnung

WORUM ES GEHT

- Den risikobasierten Ansatz des EU AI Act und seine vier Stufen erklären.
- Die Rollen Anbieter und Betreiber unterscheiden.
- Die wichtigsten Termine nach dem Digital Omnibus kennen.

SZENE · „BETRIFFT UNS DAS?“

Im April liest Karl Weninger einen Zeitungsartikel über KI-Regulierung und leitet ihn mit einer einzigen Zeile an Miriam und Jonas weiter: „Betrifft uns das?“

Jonas antwortet nicht mit einer Einschätzung, sondern mit einem Dokument, das er ohnehin führt: dem **Agenten-Steckbrief**. Darin steht, welchen Zweck der Kreditoren-Agent hat, welche Daten er verarbeitet, welche Entscheidungen er trifft, wen diese Entscheidungen betreffen, welche Grenzen gelten und wer ihn beaufsichtigt. Dazu schreibt er: „Die rechtliche Bewertung muss die Kanzlei machen. Aber ohne diese Fakten kann sie es nicht.“

Eine Woche später liegt der Steckbrief bei der Kanzlei, mit der Aurelia zusammenarbeitet.

Risiko statt Technik

Der **EU AI Act** (Verordnung (EU) 2024/1689) ist das erste umfassende KI-Gesetz. Er gilt unmittelbar in allen Mitgliedstaaten und erfasst alle, die KI-Systeme in der EU anbieten oder einsetzen, auch Anbieter von außerhalb der EU. Sein Kern ist ein risikobasierter Ansatz: Je größer das Risiko eines Systems für Menschen, desto strenger die Pflichten.

ABB. 6 - DIE RISIKOPYRAMIDE DES EU AI ACT: JE HÖHER, DESTO STRENGER DIE PFLICHTEN.

- **Verbotene Praktiken** (Art. 5): etwa Social Scoring oder manipulative Systeme. Der Digital Omnibus hat KI-Systeme ergänzt, die nicht einvernehmliche intime Darstellungen oder Missbrauchsmaterial erzeugen; dieses Verbot gilt ab 2. Dezember 2026.
- **Hochrisiko** (Anhang I und III): KI in Bereichen wie Personalauswahl, Kreditvergabe, kritischer Infrastruktur oder regulierten Produkten. Erlaubt, aber mit strengen Pflichten zu Risikomanagement, Datenqualität,

Dokumentation und menschlicher Aufsicht (Art. 14, für Betreiber Art. 26), anwendbar ab 2. Dezember 2027 bzw. 2. August 2028.

- **Transparenzpflichten** (Art. 50): Anbieter sorgen dafür, dass Menschen erkennen, dass sie mit einem KI-System interagieren, und markieren KI-erzeugte Inhalte maschinenlesbar; Betreiber legen u. a. Deepfakes offen. Die Kommission hat dazu am 20. Juli 2026 Leitlinien veröffentlicht.
- **Minimales Risiko:** der große Rest, ohne besondere Auflagen.

Die Einstufung hängt am **Einsatzzweck**, nicht an der Technik. Ein Agent, der Rechnungen zuordnet und Buchungen vorschlägt, liegt typischerweise in den unteren Stufen; beantwortet er Anfragen von Lieferanten, kommen Transparenzpflichten hinzu. Derselbe Agent, eingesetzt zum Vorsortieren von Bewerbungen, wäre ein Hochrisiko-System nach Anhang III.

Rollen: Anbieter und Betreiber

Der AI Act verteilt Pflichten nach Rollen. **Anbieter** ist, wer ein KI-System entwickelt und in Verkehr bringt. **Betreiber** ist, wer es in eigener Verantwortung einsetzt – bei Aurelia also das Unternehmen selbst. Wichtig für die Praxis: Wer ein System wesentlich verändert oder unter eigenem Namen einsetzt, kann in die Anbieterrolle geraten, samt deren Pflichten. In Projekten mit Dienstleistern gehört die Rollenfrage deshalb in den Vertrag: Wer dokumentiert, wer überwacht, wer meldet Vorfälle?

Zwei weitere Punkte betreffen fast alle Betreiber. Für die großen Basismodelle (General-Purpose AI) gelten seit 2. August 2025 eigene Anbieterpflichten; die Dokumentation des Modellanbieters gehört in die eigene Akte des KI-Systems. Und seit 2. Februar 2025 ergreifen Anbieter und Betreiber Maßnahmen, um die **KI-Kompetenz** ihres Personals zu unterstützen (Art. 4), passend zu Rolle und eingesetzten Systemen; seit dem Digital Omnibus müssen sie kein bestimmtes Niveau garantieren. Am Kern ändert das wenig: Wer die Warteschlange bearbeitet, muss verstehen, was der Agent kann, wo er irrt und wann man ihn stoppt.

Der Zeitplan

DATUM	WAS GILT
2. Februar 2025	Verbote (Art. 5) und KI-Kompetenz (Art. 4)
2. August 2025	Pflichten für Anbieter von General-Purpose-AI-Modellen
2. August 2026	Transparenzpflichten nach Art. 50
2. Dezember 2026	Frist für die maschinenlesbare Markierung bei Systemen, die vor August 2026 in Verkehr gebracht wurden (Anbieter); neues Verbot KI-erzeugter intimer Darstellungen
2. Dezember 2027	Hochrisiko-Pflichten nach Anhang III
2. August 2028	Hochrisiko-Pflichten für KI in regulierten Produkten (Anhang I)

TAB. 4 - ZEITPLAN NACH DEM DIGITAL OMNIBUS, IN KRAFT SEIT 27. JULI 2026.

Verstöße gegen die Verbote können mit bis zu 35 Millionen Euro oder 7 Prozent des weltweiten Jahresumsatzes geahndet werden, bei den meisten übrigen Pflichten mit bis zu 15 Millionen Euro oder 3 Prozent; für KMU gilt der niedrigere Wert. Überwacht wird national: In Deutschland ist nach dem KI-Marktüberwachungsgesetz vom Juli 2026 die Bundesnetzagentur zuständig, in Österreich ist bislang keine Marktüberwachungsbehörde benannt – die KI-Servicestelle der RTR informiert und berät. Für die Basismodelle ist EU-weit das AI Office zuständig. Realistisch beginnt Aufsicht mit Fragen – und die beste Antwort ist eine gepflegte Akte.

AUS SICHT DER WIRTSCHAFTSINFORMATIK

Compliance beginnt in der Architektur

Bemerkenswert ist, wie viele Anforderungen des AI Act eine technische Entsprechung im Governance-Ring haben. Menschliche Aufsicht entspricht der Warteschlange mit Stopp-Schalter, Dokumentationspflichten dem Audit-Trail, Transparenz einer Kennzeichnungsfunktion der Plattform, das Inventar dem Agenten-Steckbrief. Wer diese Funktionen von Anfang an einbaut, schafft die Faktenbasis, die eine rechtliche Einordnung braucht.

Eine sinnvolle Reihenfolge im Unternehmen: zuerst ein **Inventar** aller KI-Systeme mit Steckbriefen, dann die **Einordnung** durch Juristinnen und Juristen, dann die **Kennzeichnung**, schließlich – falls nötig – die Unterlagen für Hochrisiko-Systeme.

KEINE RECHTSBERATUNG

Dieses Kapitel gibt einen Überblick nach dem Stand September 2026. Ob und welche Pflichten für ein konkretes System gelten, gehört in die Hände von Juristinnen und Juristen.

MERKSATZ

Der AI Act regelt Einsatzzwecke, nicht Technologien. Wer Inventar, Protokoll und Aufsicht sauber baut, gibt der rechtlichen Einordnung eine belastbare Grundlage.

FRAGEN FÜR EUER UNTERNEHMEN

6.1 Warum kann derselbe Agent je nach Einsatz in unterschiedliche Risikostufen fallen?

6.2 Wann kann ein Unternehmen, das ein KI-System nur einsetzt, in die Anbieterrolle geraten?

6.3 Welche Termine sind für ein Unternehmen mit Backoffice-Agenten bis 2028 noch relevant?

EPILOG

Ein Jahr später

Zwölf Monate nach jenem Montag im Februar sieht Miriams Arbeitsliste anders aus.

Die Dunkelverarbeitungsquote liegt bei 88 Prozent. In der Warteschlange landen im Schnitt 480 Rechnungen im Monat, und die meisten davon sind tatsächlich schwierig. Sabine Kraus bearbeitet heute vor allem Preisabweichungen und Klärungen mit Lieferanten – Arbeit, für die man die Lieferanten kennen muss. Zwei Kolleginnen haben ins Cash-Management gewechselt, und Aurelia nutzt Skonto so konsequent wie nie zuvor.

Nicht alles lief glatt. Im Juni ordnete der Agent zwei Wochen lang auffällig oft Teillieferungen falsch zu, nachdem ein großer Lieferant sein Rechnungslayout geändert hatte. Die Evals hätten es früher gezeigt, wären neue Layouts schneller ins Golden Set gewandert. Seitdem gibt es dafür einen festen Termin im Monat. Als der Modellanbieter im Herbst eine neue Version ankündigte, lief sie zuerst drei Wochen parallel gegen inzwischen 1.900 Testfälle, bevor sie den Kreditoren-Agenten übernahm.

Und der Roboter aus dem Oktober? Er holt noch immer jede Nacht Rechnungen aus zwei Portalen. Wenn er stehen bleibt, bemerkt es jetzt der Agent als Erster – und meldet es in eine Warteschlange, die jemand liest.

Als ein befreundeter Geschäftsführer Karl Weninger fragt, was die KI denn nun gebracht habe, überlegt er kurz. „Weniger Tippen, mehr Entscheiden“, sagt er. „Und wir wissen jederzeit, wer was getan hat.“

Was bleibt

01 **Mit dem Prozess beginnen, nicht mit der Technik.** Process Mining und Kennzahlen zeigen, wo Urteil gefragt ist.

02 **Ziele zu delegieren heißt, Verantwortung neu zu regeln.** Ein Agent ist eine Delegationsbeziehung.

03 **Grenzen gehören in die Werkzeugschicht.** Was dort durchgesetzt wird, gilt.

04 **Autonomie wird Stufe für Stufe verdient.** Belegt durch Evals, Protokoll und Ausnahmequote.

05 **Technik liefert Fakten, Juristen die Einordnung.** Ein gepflegter Steckbrief verbindet beides.

ANHANG

Kurzfassung, Denkanstöße, Weiterlesen

A · Das Cheat Sheet in sechs Sätzen

01 Ein Agent bekommt ein Ziel statt einer Schrittfolge und arbeitet in einer Schleife, bis das Ziel erreicht ist oder eine Regel ihn stoppt.

02 Schnittstelle, wo es geht; RPA, wo nichts anderes geht; ein Agent, wo Urteil gefragt ist – aber nie ohne echte Testfälle und Verantwortliche für Ausnahmen.

03 Modell, Kontext, Werkzeuge und Evals bilden die Bausteine; das Modell allein garantiert nichts.

04 Identität, Policy, Audit und Aufsicht halten den Agenten im Zaum.

05 Autonomie steigt Stufe für Stufe von L0 bis L4, belegt durch Evals, Audit-Trail und Ausnahmequote.

06 Der EU AI Act ordnet nach Einsatzzweck und Risiko; die meisten Backoffice-Agenten liegen in den unteren Stufen.

Die grafische Kurzfassung steht unter botforce-technology.com/glossar/cheat-sheet.

B · Denkanstöße zu den Fragen

KAPITEL 1

1.1 Kriterium ist, wer die Aktion im System auslöst. Chatbot: beantwortet die Frage nach einer Skontofrist. Copilot: schlägt die passende Bestellung vor, ein Mensch bucht. Agent: ordnet selbst zu, bucht innerhalb seiner Grenzen und dokumentiert die Entscheidung.

1.2 Wahrnehmen: Eingang und Lücke erkennen. Folgern: ableiten, wo die fehlende Information liegt. Planen: Schritte und Prüfkriterien festlegen. Handeln: über Werkzeuge ausführen. Prüfen: Ergebnis gegen die Kriterien abgleichen, Begründung festhalten, bei Unsicherheit eskalieren.

1.3 Informationsasymmetrie – der Auftraggeber sieht nicht jeden Schritt: Audit-Trail. Zielkonflikt – das Modell verfolgt nicht automatisch die Ziele des Unternehmens: Policy und Evals. Kontrolle: menschliche Aufsicht mit Eskalation und Stopp-Schalter.

KAPITEL 2

2.1 RPA braucht keine Schnittstellen und ist deshalb schnell eingeführt. Weil sie Oberflächen bedient und ein festes Drehbuch abspielt, bricht sie bei geänderten Masken und unvorhergesehenen Fällen ab.

2.2 Individuelle Antwort. Prüfkriterien: Schnittstellen, Beschreibbarkeit der Regeln, Varianten, Datenstruktur, Ausnahmen, laufende Kosten, typisches Risiko. Häufig ergibt sich eine Kombination der Formen.

2.3 Process Mining zeigt reale Abläufe, Ausnahmemuster und deren Häufigkeit. Daraus folgt, welchen Anteil der manuellen Fälle ein Agent adressieren kann und wie stark die Dunkelverarbeitungsquote steigen könnte. Die Muster liefern zugleich Testfälle für die Evals.

KAPITEL 3

3.1 Sprachmodelle erzeugen plausible Antworten und füllen fehlende Fakten mit häufigen Mustern aus ihren Trainingsdaten. RAG legt die relevanten Belege in den Kontext; Aussagen werden dadurch begründbar und prüfbar.

3.2 Ein Agent wirkt nur über Werkzeuge. Dort durchgesetzte Grenzen können Eingaben nicht aushebeln, und alle Aufrufe lassen sich zentral protokollieren.

3.3 Reale historische Fälle mit bekannten richtigen Ergebnissen, repräsentativ einschließlich der wichtigen Ausnahmemuster, versioniert und laufend um neue Fehlerfälle ergänzt.

KAPITEL 4

4.1 Ein Prompt ist eine Anweisung an ein probabilistisches Modell und kann missverstanden oder durch Eingaben ausgehebelt werden. Nur eine in der Werkzeugschicht durchgesetzte Grenze wirkt verlässlich und lässt sich gegenüber Prüfern nachweisen.

4.2 Identität: Der Agent handelt unter eigener, eng berechtigter Identität. Policy: Das Limit von 10.000 Euro verhindert die autonome Buchung. Aufsicht: Der Fall landet mit Vorschlag und Begründung in Sabines Warteschlange. Audit: Vier Protokollzeilen dokumentieren den Ablauf mit Zeitstempeln.

4.3 Direkt übertragbar: Funktionstrennung, Vier-Augen-Prinzip, Belegfunktion, Berechtigungsprüfung. Anpassung: Wegen Tempo und Volumen müssen Kontrollen vollständig und automatisiert statt stichprobenartig wirken.

KAPITEL 5

5.1 Auf L3 bestätigt ein Mensch jede Aktion, auf L4 handelt der Agent innerhalb von Grenzen, und Menschen entscheiden nur Ausnahmen. Für den Wechsel braucht es eine hohe Übereinstimmungsquote, stabile Evals ohne Rückschritt, einen vollständigen Audit-Trail und eine Ausnahmequote, die das Team tragen kann.

5.2 Die Stufe gilt je Fallgruppe. Eindeutige Fälle mit geringem Risiko können früher auf L4, unsichere oder riskantere Fälle bleiben auf L3, bis die Nachweise reichen.

5.3 Ein Owner im Fachbereich, eingeplante Kapazität für die Warteschlange, technischer Betrieb, Verantwortliche für Evals und Fehlerreview. Fehlt eine Rolle, bleiben Ausnahmen liegen, Fehler unentdeckt oder die Verantwortung ungeklärt.

KAPITEL 6

6.1 Die Einstufung richtet sich nach dem Einsatzzweck und danach, wie stark Entscheidungen Menschen betreffen. Rechnungen zuordnen ist etwas anderes als Bewerbungen vorsortieren, das unter Anhang III fällt.

6.2 Wenn es das System wesentlich verändert oder unter eigenem Namen einsetzt. Die Rollenverteilung sollte vertraglich geklärt sein.

6.3 Am 2. Dezember 2026 endet die Frist für die maschinenlesbare Markierung bei Bestandssystemen; Hochrisiko-Pflichten folgen am 2. Dezember 2027 (Anhang III) und am 2. August 2028 (Anhang I), sofern ein Einsatz darunter fällt. Transparenzpflichten und KI-Kompetenz gelten bereits. Die Einordnung konkreter Systeme ist juristisch zu prüfen.

C · Weiterlesen

Automatisierung & RPA	botforce-technology.com/glossar/automation
KI & Machine Learning	botforce-technology.com/glossar/ai
Agentic AI	botforce-technology.com/glossar/agentic-ai
EU AI Act	botforce-technology.com/glossar/eu-ai-act
Agentic AI Cheat Sheet	botforce-technology.com/glossar/cheat-sheet

ZITIERVORSCHLAG BOTFORCE GmbH (2026): Vom Drehbuch zum Ziel. Agentic AI im Unternehmen – ein Praxisleitfaden mit Fallgeschichte. Wien.

HERAUSGEBER BOTFORCE GmbH, Wien · Stand: September 2026. Die Fallgeschichte ist frei erfunden. Die Angaben zum EU AI Act sind keine Rechtsberatung.