WHITEPAPER

Agentic Work ohne Shadow AI

Warum Organisationen jetzt einen Rahmen für KI-Agenten brauchen

Begleitendes Whitepaper zum Agentic Work Operating Canvas · digital-matters.io · Jens Arne Lück

Kurzfassung

KI-Agenten – Systeme, die nicht nur antworten, sondern eigenständig Aufgaben ausführen – ziehen gerade in jede Organisation ein. Meist ungeplant, an der IT vorbei, über private Accounts und kostenlose Tools. Das ist **Shadow AI**: KI-Nutzung, die niemand überblickt, niemand verantwortet und niemand absichert.

Das Problem ist nicht die Technologie. Das Problem ist das fehlende Betriebssystem darum herum: Wer darf was, mit welchen Daten, mit wie viel Autonomie, unter welcher Kontrolle? Der **Agentic Work Operating Canvas** gibt genau diesen Rahmen – in einem halben Tag, mit dem Team, das die Arbeit tatsächlich macht.

Dieses Whitepaper erklärt das Warum. Der Canvas und das Workshop-Paket liefern das Wie.

1. Was „Agentic Work" bedeutet

Klassische KI-Tools warten auf einen Prompt und liefern eine Antwort. Ein KI-**Agent** geht einen Schritt weiter: Er verfolgt ein Ziel, plant Zwischenschritte, ruft Tools und Systeme auf, trifft Entscheidungen und handelt – teils ohne dass ein Mensch jeden Schritt bestätigt.

Konkret heißt das: ein Agent, der eigenständig E-Mails beantwortet, Angebote erstellt, Daten aus mehreren Systemen zusammenführt, Tickets bearbeitet oder Code schreibt und ausrollt. Der Unterschied zwischen „Chatbot" und „Agent" ist der Unterschied zwischen einem Werkzeug und einem Mitarbeiter mit Handlungsspielraum.

Dieser Handlungsspielraum – die **Autonomie** – ist der Kern der Chance und des Risikos zugleich.

2. Warum Shadow AI entsteht

Shadow AI ist keine Böswilligkeit. Sie ist die logische Folge von drei Kräften:

Der Nutzen ist sofort spürbar. Mitarbeitende, die mit einem Agenten in Minuten erledigen, wofür sie sonst Stunden brauchten, warten nicht auf ein Freigabeverfahren. Sie handeln.

Die Hürde ist null. Leistungsfähige Agenten sind kostenlos oder für wenige Euro im Monat verfügbar, ohne IT, ohne Einkauf, ohne Vertrag. Ein privater Account genügt.

Der Rahmen fehlt. Solange die Organisation keine klare Antwort auf „Was ist erlaubt?" hat, füllt jeder Einzelne die Lücke mit seiner eigenen Interpretation.

Das Ergebnis: Vertrauliche Daten landen in Systemen, die niemand geprüft hat. Entscheidungen werden von Agenten getroffen, die niemand dokumentiert hat. Und wenn etwas schiefgeht, kann niemand rekonstruieren, was passiert ist.

3. Was auf dem Spiel steht

Datenschutz und Vertraulichkeit. Kundendaten, Verträge, Personaldaten, Betriebsgeheimnisse – einmal in einem ungeprüften Modell verarbeitet, sind sie außer Kontrolle.

Compliance und Haftung. DSGVO, der EU AI Act und branchenspezifische Vorgaben verlangen Nachvollziehbarkeit und Verantwortlichkeit. „Wir wussten nicht, dass jemand das nutzt" ist keine Verteidigung.

Qualität und Vertrauen. Ein Agent, der falsch handelt, tut das schnell und skaliert. Fehlerhafte Angebote, falsche Auskünfte, unbemerkte Automatisierungsfehler beschädigen Kundenbeziehungen.

Verpasste Chancen. Wer Agentic Work nur verbietet, verliert doppelt: Der Schatten bleibt, und der legitime Nutzen wird nicht gehoben. Ein Verbot ohne Alternative erzeugt genau die Heimlichkeit, die man vermeiden wollte.

4. Warum Verbieten nicht funktioniert – und Ignorieren auch nicht

Organisationen reagieren typischerweise mit einem von zwei Extremen. **Das Verbot** treibt die Nutzung in den Untergrund; sie hört nicht auf, sie wird nur unsichtbar. **Das Laissez-faire** überlässt jede Risikoentscheidung dem Einzelnen, der weder die Daten-Klassifizierung noch die rechtliche Lage überblicken kann.

Der dritte Weg ist ein **bewusst gesetzter Rahmen**: klar definierte Zonen, in denen Agentic Work erwünscht, erlaubt oder tabu ist. Nicht als Bremse, sondern als Landebahn – damit die Energie, die ohnehin da ist, in geordnete Bahnen fließt.

5. Warum kein bestehendes Framework diese Lücke schließt

Es gibt bereits reichlich Werkzeuge rund um KI – nur keines, das genau diese Frage beantwortet. **Data- & AI-Strategy-Canvas** helfen bei der übergeordneten Strategie, sind aber nicht auf agentische Arbeit, Autonomie und Toolzugriff zugeschnitten. **AI-Use-Case-Canvas** bewerten einzelne Ideen, beantworten aber nicht, welche Art von agentischer Arbeit die Organisation überhaupt zulassen will. **AI-Governance-Canvas** setzen meist einen konkreten Use Case voraus. **NIST AI RMF** und **ISO/IEC 42001** sind belastbare Referenzen für Risikomanagement und Managementsysteme, aber zu abstrakt für einen Workshop mit Fachbereichen. Und **Agentic-AI-Governance-Ansätze** treffen das Problem, sind aber häufig vendor-, security- oder plattformlastig.

Was fehlt, ist ein Format *zwischen* Strategie, Governance und Use-Case-Bewertung. Viele Organisationen springen direkt von „Wir wollen KI-Agenten nutzen" zu „Welche Use Cases bauen wir?" – und überspringen die entscheidende Zwischenfrage:

> **Welche Form von agentischer Arbeit wollen wir als Organisation überhaupt ermöglichen, begrenzen und steuern?**

Genau das ist der **Agentic Work Operating Canvas**: kein weiteres Use-Case-Template, sondern ein *Organisations-Canvas für agentische Arbeit*. Er schafft nicht den einzelnen Agenten – er schafft den Rahmen, in dem gute Agenten-Ideen sicher entstehen können.

---

6. Fünf Leitprinzipien

1 · Agenten sind keine Tools. Agenten übernehmen Arbeit. Ein Tool wird benutzt; ein Agent übernimmt Teile eines Arbeitsprozesses. Deshalb muss nicht nur Technologie bewertet werden, sondern Arbeit, Verantwortung und Kontrolle.

2 · Erst der Rahmen. Dann die Agenten. Einzelne Ideen lassen sich nur sinnvoll bewerten, wenn vorher geklärt ist, welche Daten, Tools und Autonomiegrade zulässig sind, wer verantwortet und welche Bereiche tabu sind.

3 · Governance soll ermöglichen, nicht nur verhindern. Wer KI-Agenten nur verbietet oder überreguliert, erzeugt Shadow AI. Der Rahmen muss sichere Spielräume schaffen, nicht nur Grenzen ziehen.

4 · Nicht jede Agenten-Idee ist ein Agenten-Use-Case. Manches ist als klassischer Workflow, Automatisierung, Softwarefeature, RPA oder Schulung besser aufgehoben. Agenten gehören dort hin, wo agentische Arbeit wirklich Mehrwert schafft – nicht dorthin, wo der Hype hinzeigt.

5 · Der Mensch bleibt verantwortlich, aber nicht automatisch wirksam. „Human-in-the-Loop" ist keine Floskel: Kontrolle funktioniert nur, wenn der Mensch fachlich qualifiziert ist, genug Kontext und Zeit hat, den Output nachvollziehen kann, klare Freigabekriterien kennt und Fehler erkennen darf.

---

7. Der Agentic Work Operating Canvas

Der Canvas übersetzt „bewusster Rahmen" in eine konkrete, an einem Vormittag bearbeitbare Struktur. Zehn Bausteine in vier Schritten:

Schritt 1 – Strategic Intent. *Wohin wollen wir?* Baustein 01 **Ambition & Zielbild** und 02 **Agentic Work Scope** klären die strategische Absicht: Welche Rolle sollen Agenten spielen, welche Arbeit sollen sie unterstützen dürfen?

Schritt 2 – Boundaries & Risk. *Wo sind die Grenzen?* Baustein 03 **Risk Appetite**, 04 **Data Boundaries** und 05 **Tool & Platform Guardrails** definieren die Leitplanken: wie viel Risiko, welche Daten in welchen Kontexten, welche Tools und Aktionen erlaubt sind.

Schritt 3 – Control & Accountability. *Wer hat die Kontrolle?* Baustein 06 **Human Control Principles**, 07 **Ownership & Roles** und 08 **Lifecycle & Registry** regeln Kontrolle und Verantwortung: wo der Mensch eingreift, wer verantwortlich ist, wie Agenten dokumentiert und abgeschaltet werden.

Schritt 4 – Enablement & Decision. *Wie kommen wir ins Handeln?* Baustein 09 **Enablement & Adoption** und 10 **Decision Model** sorgen für Befähigung und Entscheidung: wie Mitarbeitende sicher befähigt werden und wie über neue Agenten-Ideen entschieden wird.

Das Ergebnis ist kein Regelwerk aus der Rechtsabteilung, sondern eine gemeinsam getragene Betriebsvereinbarung für KI-Agenten – erarbeitet von den Menschen, die damit arbeiten.

8. Was ein guter Rahmen leistet

Er macht **Erlaubtes sichtbar**, damit Nutzung nicht mehr heimlich passieren muss. Er macht **Grenzen konkret**, damit niemand raten muss. Er macht **Verantwortung eindeutig**, damit im Ernstfall klar ist, wer entscheidet. Und er **hält sich selbst lebendig** über einen Review-Rhythmus, weil sich die Technologie schneller ändert als jede Richtlinie.

9. Kernbotschaften auf den Punkt

— Erst der Rahmen. Dann die Agenten.

— Shadow IT war Excel. Shadow AI werden Agenten.

— Agenten sind keine Tools. Sie übernehmen Arbeit.

— Wer Agenten nur verbietet, bekommt Shadow AI.

— Wer Agenten ohne Operating Model einführt, verliert Kontrolle.

— Governance darf Innovation nicht verhindern. Sie muss sichere Spielräume schaffen.

— Der Unterschied zwischen Produktivität und Risiko liegt nicht im Prompt. Er liegt im Rahmen.

— Unternehmen brauchen nicht nur KI-Kompetenz. Sie brauchen agentische Arbeitsfähigkeit.

10. Der nächste Schritt

Der Canvas ist frei nutzbar. Das begleitende Workshop-Paket – Facilitator Guide, Leitfragen, ausgefülltes Beispiel, Vorlagen – macht ihn ohne externe Begleitung durchführbar. Für Organisationen, die es gemeinsam mit uns erarbeiten wollen, bietet digital-matters.io einen moderierten Workshop an.

Der wichtigste Schritt ist der erste: den Rahmen bewusst zu setzen, bevor der Schatten ihn für einen setzt.