

Ist Ihr Unternehmen organisatorisch reif für autonome KI?

Vom Prompt zum Kontext: Weshalb künstliche Intelligenz nicht nur die Technologie, sondern auch die Organisation, ihre Verantwortlichkeiten und ihre Sicherheitsarchitektur herausfordert.

Künstliche Intelligenz übernimmt zunehmend Aufgaben, die bisher Menschen steuern und kontrollieren. Doch sind unsere Organisationen darauf vorbereitet?

Der Mensch als Engpass – und als Kontrollinstanz

KI-Systeme entwickeln sich vom unterstützenden Werkzeug zu Lösungen, die Aufgaben planen und innerhalb vorgegebener Grenzen ausführen können. Unternehmen erwarten schnellere Entscheidungen, weniger manuelle Arbeit und effizientere Geschäftsprozesse. Damit verändern sich jedoch auch Kontroll- und Entscheidungsfunktionen: Menschen erkennen Ausnahmen, interpretieren Regeln und handeln innerhalb definierter Kompetenzen.

Der Fortschritt der KI muss deshalb von einer entsprechenden Entwicklung der organisatorischen und technischen Kontrollen begleitet werden.

Vom Prompting zum Context Engineering

Beim klassischen Prompting formuliert ein Mensch eine Anweisung, liefert Informationen und überprüft das Ergebnis. Context Engineering ermöglicht es dagegen, einer KI gezielt Daten aus Dokumenten, Anwendungen und Geschäftsprozessen zur Verfügung zu stellen. Context Graphs können Beziehungen zwischen Personen, Informationen und geschäftlichen Vorgängen abbilden. KI-Agenten nutzen solche Kontexte, um Werkzeuge aufzurufen und Aufgaben auszuführen; mehrere Agenten können sich Teilaufgaben gegenseitig übertragen.

Die KI verarbeitet damit nicht mehr ausschliesslich Informationen, sondern wird Teil betrieblicher Abläufe. Das NIST AI Risk Management Framework betont, dass KI-Risikomanagement Rollen, Verantwortlichkeiten und Governance über den gesamten Lebenszyklus benötigt. [1]

Die Organisation kennt Regeln – kennt sie die KI auch?

Unternehmen besitzen Organisationsreglemente, Stellenbeschreibungen, Freigabeprozesse und IT-Berechtigungskonzepte. Dennoch liegen Regeln oft verteilt in unterschiedlichen Systemen, Dokumenten und in gelebter Praxis vor. Ein KI-Agent kennt solche Grenzen nicht automatisch.

Selbst wenn Regeln in seinem Kontext enthalten sind, garantiert das noch keine zuverlässige Durchsetzung.

Technische Integrationsfähigkeit ist nicht gleich organisatorische Handlungserlaubnis.

Wer darf was – und in welchem Kontext?

Ein Verkaufsleiter lässt durch einen KI-Agenten die Profitabilität seiner Kunden analysieren. Der Agent verbindet dafür CRM-Daten, ERP-Informationen und Finanzaufzeichnungen. Vielleicht wären auch individuelle Personalkosten oder vertrauliche Einkaufskonditionen für die Berechnung nützlich. Daraus folgt jedoch nicht, dass diese Informationen im Auftrag des Verkaufsleiters verwendet oder offengelegt werden dürfen. Selbst abgeleitete Ergebnisse können schützenswert sein.

Bewährte Verfahren wie rollenbasierte Zugriffskontrolle bleiben relevant. Mit Attribute-Based Access Control (ABAC) lassen sich zusätzlich Merkmale der Identität, der Information, der vorgesehenen Aktion und der Umgebung für Autorisierungsentscheidungen berücksichtigen. NIST beschreibt diesen Ansatz in SP 800-162. [2] Entscheidend bleibt: Die Autorisierung muss ausserhalb des KI-Modells technisch durchgesetzt werden.

Wenn KI-Agenten Aufgaben delegieren

Ein Agent soll einen Beschaffungsprozess optimieren und beauftragt weitere Agenten mit Lieferantenbewertung und Angebotsanalyse. Damit entsteht eine Kette von Aufgaben und technischen Zugriffen. Nachgelagerte Agenten dürfen dabei keine zusätzlichen Rechte erhalten, nur weil sie für den ursprünglichen Auftrag nützlich erscheinen. Besonders sensibel sind Bestellungen, Zahlungen, Vertragsabschlüsse oder Änderungen an produktiven Systemen.

Die OWASP Top 10 for Agentic Applications 2026 dokumentieren Risiken durch manipulierte Agentenziele, missbräuchliche Werkzeugnutzung und fehlerhafte Vertrauens- oder Berechtigungsgrenzen. [3] Sinnvolle Gegenmassnahmen sind getrennte Agentenidentitäten, minimale Rechte, eng begrenzte Schnittstellen, Protokollierung und gegebenenfalls menschliche Freigaben.

Effizienz ersetzt keine Governance

Autonome KI ist nicht nur ein IT-Projekt. Die Geschäftsleitung legt Risikobereitschaft und zulässige Automatisierungsgrade fest. Fachbereiche definieren Geschäftszwecke und Entscheidungskompetenzen. Der CIO verantwortet Integration und Betrieb, der CISO die Beurteilung von Sicherheitsrisiken und Kontrollen. Datenschutz und Compliance berücksichtigen einschlägige rechtliche Anforderungen. Die Zuständigkeiten müssen gemeinsam funktionieren, nicht nebeneinander.

Die NIST-Funktionen Govern, Map, Measure und Manage geben eine Struktur für KI-Risikomanagement. [1] Die CIS Controls v8.1 unterstützen grundlegende Cybersicherheitsmassnahmen wie den Schutz von Daten, Zugriffskontrollen, sichere Konfiguration und Audit-Logs. [4] Sie sind jedoch kein vollständiges Framework für KI-Governance: Autonomie und Modellverhalten erfordern zusätzliche Prüfungen.

Schweizer Unternehmen bleiben verantwortlich

Der Eidgenössische Datenschutz- und Öffentlichkeitsbeauftragte bestätigt, dass das Schweizer Datenschutzgesetz auf KI-gestützte Bearbeitungen von Personendaten anwendbar ist. [5] Für Unternehmen mit entsprechendem EU-Bezug kann auch der europäische AI Act relevant sein; seine Bestimmungen werden stufenweise anwendbar. [6]

Rechtliche Konformität allein ersetzt aber keine organisatorische Reife. Ein Unternehmen muss Entscheidungen nachvollziehen, delegierte Befugnisse begrenzen und kritische Abläufe unterbrechen können.

Sechs Dimensionen organisatorischer KI-Reife

01 · Organisation	Sind Aufgaben, Kompetenzen, Verantwortung und Stellvertretungen klar definiert?
02 · Prozesse	Sind Genehmigungen und die Grenzen automatisierter Entscheidungen dokumentiert?
03 · Informationen	Sind schützenswerte Daten klassifiziert und ihre zulässigen Verwendungszwecke bekannt?
04 · Berechtigungen	Werden Zugriffs- und Handlungsregeln auch zwischen KI-Agenten durchgesetzt?
05 · Governance	Sind Betrieb, Überwachung und KI-spezifisches Risikomanagement zugeordnet?
06 · Kontrolle	Werden Aktionen protokolliert, überprüft und bei Bedarf sicher gestoppt?

Diese sechs Dimensionen sind ein praxisorientierter Orientierungsrahmen, kein eigenständiger Zertifizierungsstandard.

Fazit: Die nächste Grenze liegt in der Organisation

Context Engineering und autonome Agenten eröffnen erhebliche Möglichkeiten. Mit wachsender Autonomie steigen jedoch die Anforderungen an Verantwortlichkeiten, Berechtigungen und Kontrollmechanismen. Der Mensch kann als operativer Engpass entlastet werden. Seine Kontrollfunktion und die Verantwortung des Unternehmens dürfen nicht unbemerkt verloren gehen.

Sichere autonome KI beginnt mit einer Organisation, die ihre eigenen Prozesse, Entscheidungsbefugnisse und Sicherheitsgrenzen kennt und beherrscht.

Fachliche Grundlagen und Referenzen

[1] NIST – Artificial Intelligence Risk Management Framework (AI RMF 1.0), NIST AI 100-1 (2023). doi.org/10.6028/NIST.AI.100-1

[2] NIST – Guide to Attribute Based Access Control, SP 800-162. doi.org/10.6028/NIST.SP.800-162

[3] OWASP – Top 10 for Agentic Applications 2026. genai.owasp.org

[4] CIS – CIS Critical Security Controls v8.1. [cisecurity.org/controls/v8-1](https://www.cisecurity.org/controls/v8-1)

[5] EDÖB – Geltendes Datenschutzgesetz ist auf KI direkt anwendbar. edoeb.admin.ch

[6] Europäische Kommission – AI Act Implementation Timeline. [AI Act Service Desk](#)

ÜBER SEC4IT

sec4it – Unabhängige Cybersecurity-Beratung für Schweizer Unternehmen. Ein strukturiertes Assessment auf Basis der CIS Controls macht Sicherheitsrisiken sichtbar und unterstützt die Priorisierung wirksamer Massnahmen. Fragen der organisatorischen KI-Reife ergänzen diesen Blick um Entscheidungen, Zuständigkeiten und KI-spezifische Risiken.

www.sec4it.ch