

AI Compliance und Ethics

Orientierung für Unternehmen im Zeitalter von Sprachmodellen und autonomen Agenten



Management Summary

Die Nutzung generativer KI und autonomer Agenten entwickelt sich zum festen Bestandteil betrieblicher Prozesse. Gleichzeitig wird mit dem schrittweisen Wirksamwerden des EU AI Acts der regulatorische Rahmen verbindlich. Unternehmen aller Größen stehen vor der Herausforderung, Innovationspotenziale zu erschließen und gleichzeitig Transparenz, Verantwortlichkeit und regulatorische Konformität sicherzustellen.

Viele Organisationen betrachten Compliance, ethische Leitlinien und technische Schutzmaßnahmen noch getrennt. In der Praxis zeigt sich: Erst ihr Zusammenspiel bildet eine tragfähige Grundlage. AI Governance wird zur Verbindung zwischen regulatorischen Anforderungen, unternehmerischen Werten und dem operativen Einsatz von KI.

Dieses Whitepaper richtet sich an Verantwortliche in den Bereichen KI, Compliance, Risikomanagement, Enterprise Architecture und Governance sowie an Betreiber von KI-Systemen. Es zeigt, welche Entwicklungen die aktuelle Diskussion prägen und wie Unternehmen die Grundlagen für eine verantwortungsvolle und zukunftsfähige KI-Nutzung schaffen.

Drei Perspektiven für verantwortungsvolle KI



Compliance

Regulatorische Anforderungen frühzeitig berücksichtigen und strukturiert umsetzen.



Ethik

Transparenz, Fairness und menschliche Verantwortung als Leitlinien etablieren.



AI Governance

Verantwortlichkeiten, Prozesse und Kontrollen für den sicheren KI-Einsatz schaffen.

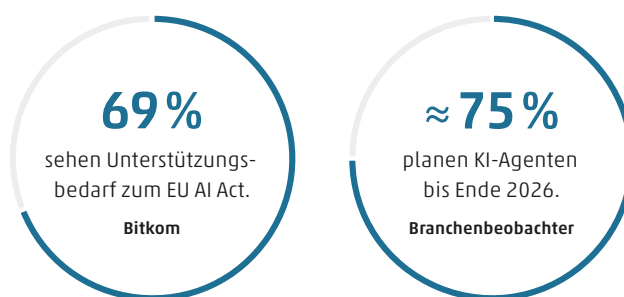
AI Compliance und Ethics – Warum diese Themen relevanter sind denn je

Die Integration generativer KI und autonomer Agenten in Unternehmensprozesse fällt zeitlich mit dem vollständigen Wirksamwerden des EU AI Acts zusammen. Wer verantwortungsvolle KI ernsthaft umsetzen will, bewegt sich dabei nicht mehr nur im Spannungsfeld von Ethik und Compliance, sondern zunehmend im Feld der AI Governance, also der laufenden Steuerung von KI über ihren gesamten Lebenszyklus. Diese drei Perspektiven gemeinsam zu denken, ist die naheliegendste Antwort auf eine Lage, die sich in den kommenden Monaten rasch verändern wird.

Große Sprachmodelle und zunehmend auch autonome KI-Agenten prägen die betriebliche Realität. Sie erstellen Texte, beantworten Anfragen, werten Daten aus und treffen, wenn es sich um Agenten handelt, im Rahmen vorgegebener Ziele eigenständige Entscheidungen. Die Geschwindigkeit dieser Entwicklung ist beachtlich, ihre Einbettung in Prozesse oft aber noch zögerlich.

Parallel dazu wird der europäische Rechtsrahmen nun verbindlich. Seit Februar 2025 gelten die Verbote bestimmter KI-Praktiken, seit August 2025 die Pflichten für Anbieter großer Sprachmodelle. Bereits zum 2. August 2026 tritt der Großteil des EU AI Acts in Kraft, darunter die Anforderungen an Hochrisiko-Systeme. Der Rahmen ist also gesetzt; was fehlt, ist in vielen Unternehmen die strukturelle Übersetzung in eigene Prozesse und ethische Grundlagen.

Die Größenordnung lässt sich beziffern: Laut Bitkom sehen 69 % der Unternehmen Unterstützungsbedarf bei der Auseinandersetzung mit dem AI Act. Zugleich gehen Branchenbeobachter davon aus, dass rund drei Viertel der Unternehmen bis Ende 2026 den Einsatz von KI-Agenten planen. Beide Entwicklungen verlaufen nicht nacheinander, sondern zeitgleich, wodurch sich Adaption und Normalisierung überlagern.



Drei Entwicklungen, die das Thema prägen

■ Regulatorik

Der EU AI Act wird 2026 in seiner Breite wirksam, verbundene Sanktionen reichen bis hin zu 35 Mio. € oder 7 % des weltweiten Jahresumsatzes. Normen wie ISO 42001 etablieren sich als Managementstandard und ergänzen bestehende Rahmen für Qualität und Informationssicherheit.

■ Technologie

Sprachmodelle und Agentensysteme bringen neue Risikoarten mit sich, die klassische IT-Sicherheit nicht vollständig erfasst. Dazu gehören etwa fehlerhafte Ausgaben mit verdeckten Datenabflüssen, manipulierte Eingaben und autonome Handlungen jenseits der intendierten Grenzen. Die Risikotaxonomie verschiebt sich mit jedem neuen Modell bzw. Einsatzkontext erneut.

■ Markt und Gesellschaft

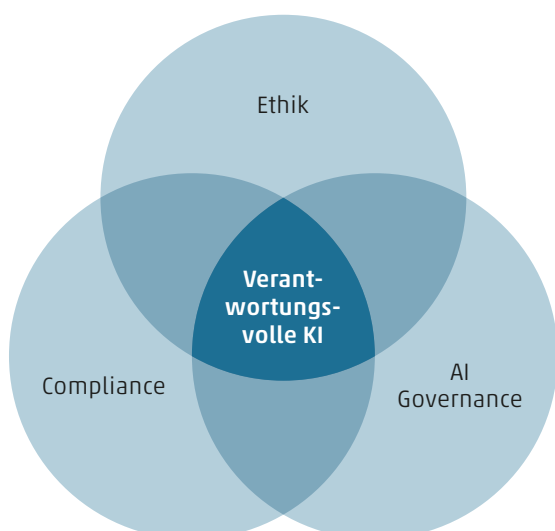
Kunden, Partner, Mitarbeitende und Aufsichtsgremien erwarten Nachvollziehbarkeit und Fairness. Vertrauen gilt dabei als Qualitätsmerkmal und entsteht dabei durch die Einhaltung von Standards und die Etablierung starker Strukturen, nicht durch bloße Kommunikation.

Von Trustworthy AI zu AI Governance

Die Debatte um vertrauenswürdige KI verändert ihren Schwerpunkt. Während Trustworthy AI lange als Eigenschaft eines einzelnen Systems verstanden wurde – einmal geprüft, einmal zertifiziert –, rückt mit der zunehmenden Einbettung von KI in den betrieblichen Alltag eine andere Perspektive in den Vordergrund: AI Governance als laufende Steuerung entlang des gesamten Lebenszyklus. Der Unterschied ist nicht semantisch, sondern organisatorisch. Eine einmalige Bewertung genügt nicht, wenn Modelle sich weiterentwickeln, Anwendungsfälle wachsen und neue Agenten hinzukommen.

Die Einhaltung des AI Act ist nicht nur eine rein rechtliche Frage. Compliancerichtlinien und ein Code of Conduct beschreiben, was ein Unternehmen machen soll – die Ethikgrundlagen aber geben wieder, wofür es in Bezug auf seine Werte und sein Handeln stehen will. Governance verbindet beides mit dem betrieblichen Alltag und hält es beweglich. Erst alle drei gemeinsam ergeben eine KI-Nutzung, welche rechtssicher und auch in Zukunft tragfähig ist.

Bemerkenswert ist, wie weit sich die Perspektiven überschneiden. Die sieben Leitprinzipien der EU für vertrauenswürdige KI, wie menschliche Aufsicht, Transparenz, Fairness, technische Robustheit, Datenschutz, gesellschaftliches Wohlergehen und Rechenschaftspflicht, decken einen Großteil der Pflichten des AI Acts ab und lassen sich zugleich auf drei operative Dimensionen bündeln: Security, Explainability und Fairness. Wer diese Dimensionen sauber operationalisiert, erfüllt regulatorische Anforderungen beinahe en passant mit.



Drei Dimensionen, die gängige Ansätze oft unterschätzen

Viele verfügbare Leitfäden beschränken sich auf Checklisten zur Umsetzung des AI Acts. Das ist hilfreich, greift jedoch an drei Stellen zu kurz:

■ Werte vor Architektur

Bevor die Frage beantwortet werden kann, wie ein KI-System gestaltet werden soll, ist zu klären, welche Werte es tragen muss. Der Standard IEEE 7000-2021 liefert dafür ein strukturiertes Verfahren mit definierten Rollen – Value Lead, Risk Lead, Transparency Manager – und überführt abstrakte Prinzipien in konkrete Anforderungen an das Systemdesign.

■ Governance als Querschnittsaufgabe

AI Governance funktioniert nur, wenn sie in etablierte Disziplinen wie Data Governance, Enterprise Architecture und Risikomanagement eingebunden wird. Sie erfordert zugleich interdisziplinäre Kompetenz: Entwickler und Produktverantwortliche benötigen ein Grundverständnis von Recht, Ethik und Risikodokumentation; Risikofachleute brauchen mehr Kenntnis über Modelle, Daten und Infrastruktur als in klassischen IT-Kontexten. Ohne diese Verzahnung entstehen Parallelstrukturen und Lücken, etwa durch Shadow AI, also den unkontrollierten Einsatz von KI-Werkzeugen außerhalb etablierter Prozesse.

■ Agentic AI als eigene Kategorie

Autonome Systeme stellen Anforderungen, die klassische Governance-Modelle nicht vorsehen: eindeutige Identitäten je Agent, durchgängige Protokollierung mehrstufiger Handlungsketten, definierte Eingriffspunkte für Menschen sowie die Möglichkeit, einzelne Agenten unverzüglich zu deaktivieren.



Ethik

Werte vor Technologie



Compliance

Standards in die Praxis übersetzen



AI Governance

Verantwortung dauerhaft verankern

Fünf Schritte als Ausgangspunkt

■ KI-Bestand erheben

Alle eingesetzten KI-Systeme erfassen und den vier Risikoklassen des AI Acts zuordnen. Diese Inventur schließt explizit auch informell genutzte Werkzeuge ein, um Shadow AI sichtbar zu machen.

■ Wertebasis klären

In einem moderierten Verfahren – etwa nach IEEE 7000 – die Werte bestimmen, an denen sich KI-Entscheidungen im Unternehmen messen lassen sollen, und sie als Teil des Kontexts der Organisation gemäß ISO 42001 dokumentieren.

■ Verantwortlichkeiten verankern

Zuständigkeiten für Aufsicht, Transparenz und Risikobewertung in bestehende Governance-Strukturen integrieren und interdisziplinär besetzen, statt neue Silos zu schaffen.

■ Agentensysteme gesondert betrachten

Für jeden produktiv eingesetzten Agenten Identität, Berechtigungen, Protokollierung und Abschaltwege definieren, bevor er an Prozesse angeschlossen wird.

■ KI-Kompetenz aufbauen

AI-Literacy ist seit Februar 2025 Pflicht und zugleich die wirksamste Maßnahme gegen Fehlanwendungen im operativen Alltag.

Unser Ansatz

Materna verbindet drei Perspektiven, die sonst selten zusammen gedacht werden: ethische Werteerhebung nach IEEE 7000 und 7010, regulatorische Compliance entlang des AI Acts und die technische Einbettung in Data Governance, Enterprise Architecture und Risikomanagement – zusammengeführt in einer AI Governance, die entlang des gesamten Lebenszyklus wirkt. Die Ergebnisse sind zertifizierbar, etwa nach ISO 42001, IEEE CertifAIEd oder ISACA AAIA, und damit auch nach außen belegbar.

Ganz gleich, wo Sie im Bereich der Künstlichen Intelligenz heute stehen: Unsere Experten setzen dort an, wo Ihr Unternehmen direkten Mehrwert und die gewünschte Transparenz über vorhandene KI-Tools, Integrationen und regulatorische Herausforderungen erzielen kann. Ob mittels Tagesworkshop, initialer Bestimmung des Reifegrades oder einer Integration in laufende Projekte: Den Umfang bestimmen Sie!

Sprechen Sie uns unverbindlich an und finden Sie heraus, welche nächsten Schritte in Ihrem Unternehmen nun den entscheidenden Vorteil bringen.

„Materna verbindet drei Perspektiven, die sonst selten zusammen gedacht werden: ethische Werte, regulatorische Compliance und technische Steuerung.“

Die fünf Handlungsfelder im Überblick



Die Materna-Gruppe realisiert komplexe und geschäftskritische Digitalisierungsprojekte für Konzerne, Mittelstand und öffentliche Verwaltung. Das Unternehmen erzielte im Jahr 2025 einen vorläufigen Gruppenumsatz von 795 Millionen Euro und beschäftigt weltweit rund 4.500 Mitarbeitende.

Materna deckt das gesamte Leistungsspektrum eines Full-Service-IT-Dienstleisters im Premium-Segment ab: von der Beratung über Implementierung bis zum Betrieb. Kunden sind IT-Organisationen sowie Fachabteilungen in Privatwirtschaft und Verwaltung.

Kontaktieren Sie uns:

Materna Information & Communications SE
Robert-Schuman-Straße 20, 44263 Dortmund
Tel.: +49 231 - 5599 - 00
E-Mail: sales@materna.group
www.materna.de