Im Tandem statt im Blindflug: KI kann Entscheidungen unterstützen – die Verantwortung bleibt beim Menschen.

Vertrauen in KI-gestützte Compliance

Künstliche Intelligenz ist im Compliance Management angekommen. Unternehmen prüfen, wie sie regulatorische Änderungen auswerten, Pflichten aus Gesetzestexten extrahieren, Verträge analysieren, Transaktionen überwachen, Risiken bewerten und Kontrollen testen können. Auch Hinweisgebersysteme und Managementberichte lassen sich unterstützen. Die Erwartungen: mehr Geschwindigkeit, Skalierbarkeit und Datenqualität bei weniger Fehlern.

VON
PROF. DR. BERND IRLENBUSCH

Je stärker KI in Compliance-Prozesse eingreift, desto wichtiger wird die Frage: Wie entsteht Vertrauen?

Dabei geht es nicht nur darum, ob Mitarbeitende einem System vertrauen. Entscheidend ist auch, ob KI das Vertrauen in Compliance-Prozesse stärkt oder schwächt. Vertrauen Verantwortliche den Ergebnissen, ohne sie zu überschätzen? Werden Risiken tatsächlich sichtbar? Und bleibt menschliche Verantwortung erkennbar?

Vertrauen entsteht nicht automatisch durch bessere Technologie. Es braucht nachvollziehbare Prozesse, klare Zuständigkeiten und überprüfbare Ergebnisse. Vor allem braucht es eine Kultur, in der KI menschliche Urteilskraft unterstützt, statt sie zu ersetzen.

Compliance ist auch ein Verhaltensproblem

Klassisches Compliance Management arbeitet mit Richtlinien, Schulungen, Risikoanalysen, Kontrollen, Audits und Hinweisgebersystemen. KI kann dabei Texte zusammenfassen, Nachweise sammeln, Auffälligkeiten erkennen und Hinweise priorisieren.

Doch Regelverstöße entstehen nicht allein aus Unkenntnis. Menschen verdrängen Risiken, delegieren Verantwortung, folgen Gruppennormen oder legen unklare Situationen zu ihren Gunsten aus.

Genau hier verändert KI das Spielfeld. Sie beeinflusst, was Mitarbeitende sehen, wie sie Risiken bewerten, wann sie eskalieren und wie sie Entscheidungen rechtfertigen. KI wird damit Teil der Compliance-Infrastruktur.

Die entscheidende Frage lautet: Fördert sie verantwortliches Handeln – oder erleichtert sie die Abgabe von Verantwortung?

Vertrauen braucht Kalibrierung

Bei KI darf das Ziel nicht maximales, sondern kalibriertes Vertrauen sein. Nutzerinnen und Nutzer müssen wissen, was ein System leisten kann und wo seine Grenzen liegen. Zu wenig Vertrauen lässt nützliche Systeme ungenutzt; zu viel erzeugt Automatisierungsbias, Scheinsicherheit und Verantwortungslücken.

KI sollte daher keine endgültigen Wahrheiten suggerieren, sondern Quellen, Unsicherheiten und Prüfbedarf sichtbar machen. Vertrauen entsteht durch überprüfbare Antworten.

„Vertrauen entsteht dort, wo Verantwortung nicht im System verschwindet."

Moral Licensing: Vertrauen darf nicht zur Selbstberuhigung werden

Moral Licensing bezeichnet die Tendenz, sich nach guten Handlungen spätere Nachlässigkeit zu erlauben. Im KI-Kontext kann schon die Investition in moderne Compliance-Technologie zum moralischen Entlastungssignal werden: Wir nehmen Compliance ernst, wir sind technologisch fortschrittlich.

Die Organisation verlässt sich dann auf die Existenz des Systems statt auf seine Wirksamkeit. KI gilt als Zeichen von Compliance-Reife, obwohl unklar ist, ob sie Risiken tatsächlich besser erkennt.

Auch Einzelne können sich entlastet fühlen. Wenn ein System eine Transaktion nicht markiert oder einen Geschäftspartner nicht auffällig findet, liegt der Schluss nahe: Dann wird es schon in Ordnung sein.

Die Governance-Frage lautet deshalb: Vertrauen wir der KI wegen überprüfbar guter Ergebnisse – oder weil sie existiert?

Moral Wiggle Room: Vertrauen braucht weniger Auslegungsspielraum

Compliance-Situationen bieten viele Mehrdeutigkeiten: Ist ein Geschenk noch angemessen? Reicht die Due Diligence? Muss ein Verdacht eskaliert werden?

KI kann solche Spielräume verkleinern, indem sie Pflichten systematisch erfasst, Fälle vergleicht, Abweichungen erkennt und Entscheidungen dokumentiert. Sie kann aber auch neue Spielräume schaffen. Risikoscores, Wahrscheinlichkeiten und generierte Texte lassen sich selektiv nutzen, um eine gewünschte Entscheidung zu rechtfertigen.

KI-gestützte Entscheidungen müssen deshalb quellenbasiert, reproduzierbar und dokumentiert sein. Wer eine Empfehlung überstimmt oder nur teilweise übernimmt, sollte dies begründen.

Ethical Fading: Vertrauen braucht sichtbare Werte

Ethical Fading beschreibt, wie der ethische Gehalt einer Entscheidung aus dem Blick gerät. Eine moralische Frage erscheint dann nur noch technisch oder administrativ.

KI kann diesen Effekt verstärken: Ein Korruptionsrisiko wird zum Drittanbieter-Score, ein Datenschutzproblem zur Datenflussanomalie, ein Hinweis auf Fehlverhalten zur Fallpriorität. Das ist operativ nützlich, kann aber den eigentlichen Schutzgegenstand verdecken.

Compliance schützt Menschen und Werte: Kunden, Mitarbeitende, Geschäftspartner, Märkte, Daten und Hinweisgeber. Gute Systeme sollten deshalb nicht nur „mittleres Risiko" melden, sondern auch den möglichen Schaden benennen – etwa Bestechungsgefahr, Datenschutzverletzungen oder Nachteile für Hinweisgeber.

Delegation und Blame Shifting: Vertrauen braucht klare Verantwortung

KI erleichtert Delegation. Sie analysiert vor, sortiert Fälle, priorisiert Risiken und schlägt Antworten vor. Das spart Arbeit, kann aber Verantwortung verschieben.

In der Praxis klingt das dann so: „Das System hat den Geschäftspartner freigegeben." „Der Algorithmus hat kein Risiko erkannt." „Wir haben uns auf die Modellbewertung verlassen." Aus einer Entscheidungshilfe wird eine Entscheidungsträgerin.

KI trägt selbst keine Compliance-Verantwortung. Diese bleibt bei Menschen und Organen – in der Geschäftseinheit, bei Compliance, Legal, Risk, Informationstechnik, der Modellverantwortung oder im Vorstand.

„Human in the loop" reicht als Formel nicht. Entscheidend ist, welcher Mensch mit welcher Kompetenz eingreift. Jedes Verfahren braucht klare Antworten: Wer verantwortet Modell und Datenqualität? Wer entscheidet final? Wer darf überstimmen? Wer muss eskalieren?

„Vertrauen setzt nicht Fehlerfreiheit voraus. Es setzt voraus, dass Verantwortung auch im Fehlerfall erkennbar bleibt."

Vertrauen setzt nicht Fehlerfreiheit voraus. Es setzt voraus, dass Verantwortung auch im Fehlerfall erkennbar bleibt.

Overconfidence Bias: Vertrauen nicht mit Gewissheit verwechseln

Menschen überschätzen häufig die Genauigkeit ihrer Einschätzungen. Generative KI verstärkt diesen Effekt: Ihre Antworten wirken flüssig, präzise und überzeugend, können aber Ausnahmen übersehen, Zwischentöne verlieren oder auf lückenhaften Daten beruhen.

Ein KI-generierter Überblick über eine Regulierung kann brillant klingen und dennoch falsch sein. Eine Untersuchungschronologie kann plausibel wirken, obwohl relevante Informationen fehlen. Ein Risikoscore kann exakt erscheinen, obwohl seine Grundlage unsicher ist.

Vertrauenswürdige Compliance-KI muss daher Unsicherheit zeigen, Quellen nennen, Annahmen offenlegen und menschlichen Prüfbedarf markieren.

Die richtige Frage lautet nicht: Können wir der KI vertrauen? Sondern: In welchem Umfang, Kontext und Prüfrahmen ist ihr Ergebnis verwendbar?

Konformität und algorithmische Autorität: Vertrauen braucht Widerspruch

Wenn alle dasselbe Dashboard nutzen, demselben Score folgen und dieselbe KI-Triage akzeptieren, entsteht schnell Konformität. Das kann Standards vereinheitlichen, aber kritisches Denken schwächen.

Hinzu kommt ein Autoritätsbias: KI wirkt objektiv, neutral und fachlich überlegen. Ist sie fest in Arbeitsabläufe eingebettet, gewinnt sie zusätzliche Autorität.

Gute Compliance braucht jedoch nicht nur Standards, sondern auch Zweifel und professionellen Widerspruch. Wer einen Verdacht hat, muss sagen dürfen: Das System liegt hier falsch.

Vertrauen entsteht nicht durch widerspruchsloses Folgen, sondern durch die Möglichkeit, Ergebnisse anzufechten.

Conditional Cooperation: Vertrauen entsteht durch sichtbare Fairness

Menschen kooperieren eher, wenn sie glauben, dass andere ebenfalls kooperieren. Mitarbeitende beobachten deshalb, ob Regeln gelten und Führungskräfte sie ernst nehmen. KI kann sichtbar machen, dass Kontrol-

len stattfinden, Hinweise bearbeitet und Maßnahmen umgesetzt werden. Das stärkt kooperative Normen. Werden Warnungen dagegen regelmäßig geschlossen, Kontrollen verspätet durchgeführt oder Ausnahmen folgenlos akzeptiert, kann Regelbruch normal erscheinen.

Vertrauen braucht daher sichtbare Fairness: gleiche Maßstäbe, nachvollziehbare Konsequenzen und eine erkennbare Reaktion des Managements.

Altruistic Cheating und Verantwortungsdiffusion

Nicht jeder Regelverstoß dient dem eigenen Vorteil. Menschen umgehen Regeln auch, um dem Team, einem Kunden oder dem Unternehmen zu helfen. Dokumentationspflichten werden verkürzt, damit ein wichtiger Kunde rechtzeitig aufgenommen werden kann; Risiken werden heruntergespielt, um Umsatz oder Projekte zu schützen.

KI kann solche Muster erkennen, etwa Ausnahmen vor Fristen, auffällige Freigaben oder Unterschiede zwischen Kundengruppen. Sie kann Regelverstöße aber auch fördern, wenn Effizienz zum dominanten Ziel wird und Warnsignale als Hindernisse gelten.

Hinzu kommt die Verantwortungsdiffusion. An KI-gestützten Prozessen sind häufig Fachbereiche, Compliance, Legal, Risk, Informationstechnik, Anbieter, Revision und Management beteiligt. Je mehr Akteure beteiligt sind, desto leichter nimmt jeder an, jemand anderes habe bereits geprüft.

Saubere Dokumentation von Zuständigkeiten, Prüfhandlungen, Freigaben und Eskalationen kann dem entgegenwirken. Vertrauen entsteht dort, wo Verantwortung nicht im System verschwindet.

Sieben Bedingungen für Vertrauen in KI-gestützte Compliance

Vertrauenswürdige KI braucht sieben Bedingungen:

Erstens: Nachvollziehbarkeit. Es muss erkennbar sein, worauf eine Einschätzung beruht.

Zweitens: Verantwortlichkeit. Zuständigkeiten für Entscheidung, Modell und Fehlerreaktion müssen feststehen.

Drittens: Kalibrierung. Unsicherheit, Grenzen und Prüfbedarf müssen sichtbar sein.

Viertens: Anfechtbarkeit. Menschen müssen Ergebnisse hinterfragen, überstimmen und eskalieren können.

Fünftens: Ethische Salienz. KI darf Compliance nicht auf Scores und Tickets reduzieren. Der Zweck einer Regel muss erkennbar bleiben.

„Bei KI darf das Ziel nicht maximales, sondern kalibriertes Vertrauen sein."

Sechstens: Fairness. KI-gestützte Compliance muss konsistent und diskriminierungsfrei wirken.

Siebtens: Wirksamkeitsprüfung. Systeme brauchen regelmäßige Prüfungen auf Fehler, Verzerrungen, Nutzungsmuster und ihren tatsächlichen Beitrag zur Compliance.

Was Compliance-Verantwortliche jetzt tun sollten

Vor der Einführung sollten Unternehmen KI nicht nur technisch, rechtlich und datenschutzrechtlich prüfen. Sie sollten auch fragen, welches Verhalten sie fördert: Wo erzeugt sie ungerechtfertigtes Vertrauen? Wo erleichtert sie Verantwortungsabgabe? Wo reduziert sie moralische Fragen auf technische Probleme?

Ein Score, eine Ampel, ein Alert oder eine Zusammenfassung verändert Entscheidungen. Deshalb sind Sprache, Darstellung und Prozesseinbettung entscheidend.

Hochrisikoentscheidungen brauchen dokumentierte menschliche Verantwortung. KI kann vorbereiten, strukturieren und warnen, darf aber nicht zur verdeckten Entscheidungsträgerin werden.

Ebenso wichtig ist die Beobachtung der tatsächlichen Nutzung: Werden Empfehlungen fast immer akzeptiert? Werden Abweichungen begründet? Werden rote Flaggen zu schnell geschlossen? Werden Risiken regelmäßig herabgestuft? Solche Muster zeigen, wie eine Organisation mit KI umgeht – und wie ihre Compliance-Kultur funktioniert.

Fazit: Vertrauen entsteht, wenn KI Verantwortung stärkt

KI kann Compliance verbessern: Sie macht Komplexität beherrschbarer, erkennt Muster, prüft Kontrollen und dokumentiert Entscheidungen.

Sie ist jedoch kein neutraler Vertrauensgenerator. Sie stärkt Vertrauen, wenn sie Risiken sichtbar macht, Entscheidungen nachvollziehbar hält und menschliche Urteilskraft unterstützt. Sie schwächt es, wenn sie Scheinsicherheit erzeugt, Verantwortung verwischt oder Widerspruch erschwert.

Die Aufgabe der Compliance-Verantwortlichen besteht darin, KI als verhaltensbezogene Infrastruktur zu gestalten. Sie muss nicht nur leistungsfähig, sicher und rechtskonform sein, sondern so in Prozesse eingebettet werden, dass sie Vertrauen verdient.

Die Zukunft der Compliance liegt nicht in vollständiger Automatisierung, sondern in einer klugen Zusammenarbeit von Mensch und Maschine. KI sollte Compliance nicht entmenschlichen, sondern Menschen helfen, in komplexen Organisationen verantwortlicher zu handeln.

Prof. Dr. Bernd Irlenbusch ist Professor und Direktor des Seminars für Unternehmensentwicklung und Wirtschaftsethik an der Universität zu Köln. Seit 2015 gehört er dem Wissenschaftlichen Beirat des BCM an.

COMPLIANCE MANAGER

MAGAZIN FÜR COMPLIANCE MANAGEMENT
AUSGABE 48 (3/2026)
WWW.COMPLIANCE-MANAGER.NET

MENSCHENRECHTE

8 MENSCHENRECHTE: ZWANGSARBEIT WIRD ZUM MARKTZUGANGSRISIKO **16 ABHILFE:** WARUM RÜCKZUG BETROFFENEN OFT NICHT HILFT **20 KI:** MENSCHENRECHTSRISIKEN ERKENNEN – OHNE BLIND ZU VERTRAUEN **24 INTERVIEW:** EUROPAS ABHÄNGIGKEITEN WERDEN ZUM RISIKO **50 VORSTANDSHAFTUNG:** SICHER ENTSCHEIDEN, WENN DAS RECHT UNSCHARF WIRD