SALLENBACH

AI & Innovation Consulting

Muster-KI-Weisung

für Aufsichtsräte, Verwaltungsräte und Board Committees

Board Edition: klare Guardrails für verantwortungsvollen KI-Einsatz, strategische Kontrolle und wirksame Aufsicht.

Ziel: KI nutzen, ohne Vertraulichkeit, Datenschutz, Reputation und Verantwortlichkeit zu verlieren.

Ansatz: Wertschöpfung statt Hype: pragmatische Regeln, Board-Cockpit und handhabbare Mindestkontrollen.

Status: Showcase-Draft; unternehmens-, branchen- und tool-spezifisch anzupassen.

Prepared for Fachzeitschrift Aufsichtsrat Ausgabe Juli/August 2026 | www.sallenbach.org; email: vero@sallenbach.org

TEXTVORLAGE FÜR MANDATSFIRMEN

Begleittext an die Belegschaft

Liebe Mitarbeitende

Künstliche Intelligenz ist inzwischen Teil unseres Arbeitsalltags: in E-Mails, Office-Tools, Recherche, Präsentationen, Protokollen, Datenanalyse und zunehmend auch in automatisierten Workflows. Wir möchten diese Möglichkeiten nutzen, aber mit klaren Leitplanken.

Der Aufsichtsrat hat deshalb eine kompakte Muster-KI-Weisung als Orientierung verabschiedet. Sie beschreibt, welche Daten in welche Tools eingegeben werden dürfen, wann eine Freigabe nötig ist und weshalb KI-Ergebnisse immer menschlich geprüft werden müssen.

Bitte nutzt KI pragmatisch, aber bewusst: keine vertraulichen, personenbezogenen, marktpreisrelevanten oder geschützten Informationen in nicht freigegebene öffentliche KI-Tools eingeben; Outputs kritisch prüfen; Unsicherheiten früh melden.

Diese Weisung ist ein Startpunkt. Fragen, Tool-Vorschläge oder Unsicherheiten können an [KI-Verantwortliche/r / Compliance / Datenschutz] gerichtet werden.

Beste Grüsse
[Verwaltungsrat / Aufsichtsrat / Geschäftsleitung]

> **Hinweis für Aufsichtsräte:** Dieser Text kann zusammen mit der Weisung an Mandatsfirmen weitergeleitet und auf die jeweilige Organisation, Branche und Tool-Landschaft angepasst werden.

1. Zweck und Geltungsbereich

Diese Muster-KI-Weisung regelt den verantwortungsvollen Einsatz von künstlicher Intelligenz in Unternehmen, die durch einen Aufsichtsrat, Verwaltungsrat oder ein vergleichbares Board überwacht werden. Sie richtet sich an Mitglieder des Aufsichtsgremiums, Geschäftsleitung, Mitarbeitende, externe Beratende sowie weitere Personen, die im Auftrag des Unternehmens mit KI-Systemen arbeiten.

Ziel ist es, KI-Nutzung zu ermöglichen, ohne Vertraulichkeit, Datenschutz, Geschäftsgeheimnisse, Urheberrechte, regulatorische Pflichten, Reputation oder Entscheidungsqualität zu gefährden.

2. Board-Prinzipien

Verantwortung bleibt menschlich: KI unterstützt Analyse, Entwurf und Strukturierung. Entscheidungen des Aufsichtsgremiums, der Ausschüsse und der Geschäftsleitung werden nicht an KI delegiert.

Vertraulichkeit zuerst: Vertrauliche Board-Unterlagen, Strategie-, Finanz-, HR-, M&A- und Kundendaten dürfen nur in freigegebenen, vertraglich geprüften Umgebungen verarbeitet werden.

Wertschöpfung vor Hype: KI wird dort eingesetzt, wo sie bessere Vorbereitung, schnellere Orientierung, höhere Qualität oder robustere Kontrollen ermöglicht.

Transparenz und Nachvollziehbarkeit: Wesentliche KI-Nutzungen werden dokumentiert: Tool, Zweck, Datenklasse, verantwortliche Person, Freigabe und Kontrollschritte.

Security und Datenschutz by Design: Neue KI-Anwendungen werden vor Einsatz auf Datenflüsse, Hosting, Training, Zugriff, Protokollierung, Löschung und Exit geprüft.

3. Rollen und Verantwortlichkeiten

Aufsichtsrat / Verwaltungsrat: genehmigt Grundsätze, Risikobereitschaft, wesentliche KI-Vorhaben und jährliche Berichterstattung. Er verlangt ein KI-Register und überprüft kritische Use Cases.

Ausschüsse: prüfen die für ihren Auftrag relevanten KI-Risiken. Audit/Risk/Compliance achten besonders auf Kontrollen, Datenschutz, Informationssicherheit und Berichtsqualität.

Geschäftsleitung: setzt die Weisung um, benennt Verantwortliche, führt Tool- und Use-Case-Register und stellt Schulung sowie Kontrollen sicher.

KI-Verantwortliche/r: prüft Tools, koordiniert Freigaben, dokumentiert Risiken und betreut Incident Response.

Nutzende und externe Mitwirkende: halten Datenzonen und Freigaben ein, prüfen Resultate kritisch und melden Fehlverhalten, Datenabfluss oder kritische Outputs sofort.

Daten- und Nutzungszonen

Je sensibler Daten oder Wirkung, desto höher sind Freigabe, Kontrolle und Dokumentation.

Zone Grün | öffentlich / unkritisch

Öffentliche Informationen, allgemeine Ideen, erste Strukturvorschläge, nicht vertrauliche Markttrends. Nutzung freigegebener Standardtools möglich; Faktenprüfung bleibt Pflicht.

Zone Gelb | intern / moderat

Interne Vorlagen, anonymisierte Prozessdaten, nicht sensible Auswertungen, Trainingsmaterial. Nur freigegebene Tools; keine personenbezogenen Details ohne Zweck und Schutzmassnahmen.

Zone Rot | vertraulich / board-relevant

VR-Unterlagen, Protokolle, Strategie, Finanzdaten, M&A, Rechtsfragen, Kunden- und Mitarbeiterdaten. Nur geprüfte Enterprise- oder Private-Umgebung; kein Upload in öffentliche Consumer-Accounts.

Zone Schwarz | verboten / hochsensibel

Berufsgeheimnisse ohne Freigabe, Insiderinformationen, laufende Verfahren, Gesundheitsdaten, Bankdaten, vertrauliche Krisenlagen. Keine Nutzung ohne spezifische Freigabe durch Legal/Compliance, Datenschutz und zuständiges Board-Gremium.

Board-Cockpit: vier Mindestfragen vor jedem KI-Einsatz

Diese vier Fragen reichen oft, um riskante KI-Nutzung früh zu erkennen.

Dokumentieren:	Wird Tool, Zweck, Datenzone und verantwortliche Person im KI-Register festgehalten?
Absichern:	Ist das Tool für diese Datenklasse freigegeben und sind Hosting, Training, Human Review und Zugriff geklärt?
Prüfen:	Wer kontrolliert Fakten, Quellen, Rechts- oder Finanzbehauptungen vor Weitergabe an das Board?
Eskalieren:	Wann muss Geschäftsleitung, Legal/Compliance, Datenschutz oder ein Board-Ausschuss beigezogen werden?

5. Erlaubte Board-Use-Cases

KI darf als Denk-, Strukturierungs- und Qualitätssicherungswerkzeug eingesetzt werden. Die folgenden Use Cases eignen sich für Pilotierung und Board-Schulung, sofern Datenzone und Tool-Freigabe eingehalten werden.

Board-Briefings: Zusammenfassung regulatorischer, technologischer oder marktbezogener Entwicklungen. Quellen prüfen; keine vertraulichen Inputs in nicht freigegebene Tools.

Sitzungsvorbereitung: Agendaentwurf, Fragenkatalog, Risikomatrix und Entscheidungsoptionen. Finale Unterlagen durch verantwortliche Person freigeben.

Policy- und Reglementsentwürfe: Erste Fassungen für KI-Weisung, Datenschutz, Informationssicherheit oder Delegationsordnung. Juristische Prüfung bleibt zwingend.

Risikoregister und Kontrollfragen: Identifikation von Schwachstellen, Massnahmen und Monitoring-Fragen. Risiken priorisieren und im Board-Protokoll nachvollziehbar festhalten.

Schulung und Simulation: Board-Szenarien zu Cyber, Krise, Medien, Haftung, EU AI Act oder Datenschutz. Keine echten Personendaten in Trainingssimulationen verwenden.

6. Unzulässige oder besonders freigabepflichtige Nutzung

Keine Eingabe vertraulicher, personenbezogener, marktpreisrelevanter oder berufsgeheimer Informationen in nicht freigegebene öffentliche KI-Tools.

Keine automatisierte Entscheidung über Personal, Vergütung, Bonität, Kredit, Versicherung, Leistung, Kündigung oder Zugang zu Dienstleistungen ohne gesonderte rechtliche Prüfung.

Keine verdeckte Transkription, Aufzeichnung oder Analyse von Board-Sitzungen ohne vorgängige Information, Rechtsgrundlage und klare Löschfristen.

Keine Veröffentlichung von KI-generierten Aussagen, Zahlen, Rechtsbehauptungen oder Finanzinformationen ohne unabhängige Prüfung.

Keine autonome Kommunikation von KI-Agenten an Kunden, Behörden, Investoren oder Medien ohne definierte Freigabegrenzen und menschliche Endkontrolle.

7. Tool-Freigabe und KI-Register

Vor dem produktiven Einsatz eines KI-Tools ist eine Freigabe erforderlich. Der Aufsichtsrat verlangt mindestens ein schlankes KI-Register, das jährlich oder bei wesentlichen Änderungen überprüft wird.

Zweck: Welcher konkrete Geschäfts- oder Kontrollnutzen wird verfolgt?

Daten: Welche Datenklassen werden eingegeben, gespeichert, abgerufen oder erzeugt?

Training / Human Review: Werden Eingaben oder Outputs zum Training, zur Produktverbesserung oder durch Menschen beim Provider geprüft?

Hosting / Transfer: Wo werden Daten verarbeitet und gespeichert? Bestehen Drittland-, CLOUD-Act- oder Subprozessor-Risiken?

Zugriff: Wer erhält Zugriff? Sind Rollen, Protokollierung, MFA und Löschung geregelt?

Kontrolle: Welche menschliche Prüfung, Quellenprüfung, Qualitätssicherung und Eskalation gelten?

8. Qualitätssicherung und Human-in-the-loop

KI-Outputs sind Arbeitsentwürfe. Wer KI nutzt, bleibt für Richtigkeit, Vollständigkeit, Aktualität, Angemessenheit, Tonalität und rechtliche Tragfähigkeit verantwortlich.

Fakten, Zahlen, Zitate, Rechtsaussagen und Quellen werden unabhängig geprüft.

Bei rechtlichen, finanziellen, regulatorischen oder reputationsrelevanten Themen erfolgt eine qualifizierte Zweitprüfung.

Entscheidungsunterlagen für das Aufsichtsgremium enthalten keine ungeprüften KI-Behauptungen.

Wesentliche Annahmen, Unsicherheiten und Datenlücken werden transparent gemacht.

Die finale Verantwortung liegt bei der zuständigen Person und beim entscheidenden Organ, nicht beim Tool.

9. Board-Sitzungen, Protokolle und vertrauliche Räume

Für Sitzungen des Aufsichtsgremiums gelten erhöhte Anforderungen. KI-gestützte Transkription, Zusammenfassung, Übersetzung oder Analyse von Sitzungen ist nur zulässig, wenn Zweck, Teilnehmerinformation, Tool-Freigabe, Datenhaltung, Löschung und Zugriff klar geregelt sind.

> **Board-Regel:** Keine KI-Aufzeichnung von Sitzungen ohne vorherige Zustimmung und definierte Lösch- oder Aufbewahrungslogik. Für Krisen-, M&A-, Personal-, Rechts- und Insider-Themen gilt grundsätzlich Zone Rot oder Schwarz.

10. Agentische Workflows und Automatisierung

KI-Agenten, die selbständig recherchieren, E-Mails vorbereiten, Daten abfragen, Termine buchen, Dokumente verändern oder externe Systeme auslösen, benötigen eine separate Freigabe.

Least-privilege-Zugriff: Agenten erhalten nur die minimal erforderlichen Rechte.

Freigabeschwellen: externe Kommunikation, Zahlungen, Vertragsänderungen und Datenexporte benötigen menschliche Bestätigung.

Protokollierung: Aktionen, Eingaben, Ausgaben und Freigaben müssen nachvollziehbar sein.

Stop-/Kill-Switch: Verantwortliche können den Agenten jederzeit deaktivieren.

Pilotierung: neue Agenten werden zuerst mit Testdaten, begrenztem Scope und klaren Erfolgskriterien geprüft.

11. Incident Response

Jede Person meldet unverzüglich, wenn vertrauliche Daten in ein falsches Tool eingegeben wurden, ein KI-System unbefugt Zugriff erhalten hat, ein kritischer Fehloutput verwendet wurde oder Anzeichen für Datenabfluss, Manipulation, Bias oder Sicherheitsverletzungen bestehen.

> **Stoppen:** Tool-Nutzung, Veröffentlichung oder Agenten-Aktion sofort anhalten.
>
> **Sichern:** Zeitpunkt, Tool, Prompt, Output, betroffene Daten und beteiligte Personen dokumentieren.
>
> **Eskalieren:** KI-Verantwortliche/r, Legal, Datenschutz, IT-Security und bei Bedarf Board-Präsidium informieren.

12. Umsetzung: 30-60-90-Tage-Plan

30 Tage | KI-Lagebild: Wo wird KI bereits genutzt? Welche Tools, Daten, Risiken und Quick Wins bestehen? Ergebnis: initiales KI-Register, Datenzonen, Sofortregeln und Kommunikationsnotiz.

60 Tage | Freigabe- und Kontrollmodell: Tool-Check, Verantwortlichkeiten, Schulung und Pilot-Use-Cases. Ergebnis: genehmigte Tool-Liste, Pilotplan, Board-Kontrollfragen und Schulungsnachweis.

90 Tage | Governance verankern: Reporting, Audit Trail, Review-Rhythmus, Krisen- und Incident-Prozess. Ergebnis: Board-Beschluss, jährlicher Review, KPIs/KRIs und Integration in Risk/Compliance.

13. Beschlussvorlage für das Aufsichtsgremium

Das Aufsichtsgremium nimmt die Chancen und Risiken des KI-Einsatzes zur Kenntnis und beschliesst:

Die vorliegende KI-Weisung wird als Mindeststandard für den verantwortungsvollen KI-Einsatz verabschiedet und unternehmensspezifisch angepasst.

Die Geschäftsleitung erstellt und pflegt ein KI-Register mit Tool-, Use-Case-, Daten- und Risikoinformationen.

Vertrauliche oder hochsensible Daten dürfen nur in freigegebenen, vertraglich geprüften KI-Umgebungen verarbeitet werden.

Wesentliche KI-Vorhaben mit rechtlichem, finanziellem, reputativem oder personenbezogenem Risiko werden dem zuständigen Ausschuss vorgelegt.

Die Weisung wird mindestens jährlich oder bei wesentlichen regulatorischen, technischen oder organisatorischen Änderungen überprüft.

Board Quick-Check: 12 Fragen

Screenshot-Modul für Board-Diskussionen, KI-Workshops oder AI Readiness Checks.

Nr.	Kontrollfrage
1	Welche KI-Tools werden heute bereits genutzt, offiziell oder inoffiziell?
2	Welche Datenklassen dürfen in welche Tools eingegeben werden?
3	Gibt es eine freigegebene Tool-Liste und einen Prozess für neue KI-Anwendungen?
4	Wo entstehen durch KI neue Haftungs-, Datenschutz-, Cyber- oder Reputationsrisiken?
5	Welche KI-Use-Cases betreffen Personal, Kunden, Finanzen, Recht, Sicherheit oder kritische Entscheidungen?
6	Wie wird geprüft, ob KI-Outputs richtig, aktuell und quellenbasiert sind?
7	Welche KI-Agenten dürfen externe Systeme auslösen oder nach aussen kommunizieren?
8	Wie werden Board-Sitzungen, Protokolle und vertrauliche Unterlagen geschützt?
9	Wann muss der Aufsichtsrat informiert oder ein Ausschuss beigezogen werden?
10	Welche Schulung brauchen Board, Geschäftsleitung und Mitarbeitende?
11	Wie werden Vorfälle gemeldet, untersucht und in Verbesserungen übersetzt?
12	Welche drei Quick Wins liefern Nutzen ohne unnötiges Risiko?

SALLENBACH | www.sallenbach.org | vero@sallenbach.org

Anhang A: KI-Register Vorlage

Minimaler Pflichtsatz für Tool- und Use-Case-Transparenz.

Registerfeld	Eintrag / Mindestinhalt
Tool / Anbieter	Name, Version, Vertragspartei und verantwortliche Fachstelle.
Zweck und Use Case	Konkreter Einsatz, erwarteter Nutzen, betroffene Prozesse.
Datenzone	Grün / Gelb / Rot / Schwarz; inkl. Beispiele der Datenarten.
Hosting / Datenresidenz	Ort der Verarbeitung, Subprozessoren, Drittlandtransfer.
Training / Human Review	Ja / Nein / unklar; vertragliche Zusicherungen dokumentieren.
Freigabestatus	Pilot / freigegeben / eingeschränkt / gesperrt.
Kontrollpunkt	Human Review, Quellenprüfung, Vier-Augen-Prinzip, Audit Trail.
Review-Datum	Nächste Überprüfung und verantwortliche Person.

Anhang B: Prompt-Check vor dem Absenden

Fünf Fragen, bevor ein Prompt in ein KI-System eingegeben wird.

- ☐ Enthält mein Prompt Namen, Kundendaten, Finanzdaten, Personalinformationen, Verträge, Protokolle oder Geheimnisse?
- ☐ Kann ich die Daten anonymisieren, aggregieren oder durch Platzhalter ersetzen?
- ☐ Ist das Tool für diese Datenzone freigegeben?
- ☐ Brauche ich für diesen Use Case eine menschliche oder juristische Vorprüfung?
- ☐ Muss dokumentiert werden, dass KI eingesetzt wurde?

Praxisregel: Wenn eine Frage nicht sicher beantwortet werden kann, keine Eingabe vertraulicher Daten. Rückfrage bei KI-Verantwortlichen, Legal/Compliance oder Datenschutz.

SALLENBACH | www.sallenbach.org | vero@sallenbach.org

Rechtliche Anker und Quellen

Diese Musterweisung basiert auf öffentlich zugänglichen Leitlinien und Rechtsankern. Sie ersetzt keine unternehmens-, branchen- oder fallbezogene juristische Prüfung.

SAV-Wegleitung für den Umgang mit künstlicher Intelligenz, verabschiedet am 14. Juni 2024.

EDÖB: KI und Datenschutz, veröffentlicht am 24. September 2025; DSG ist auf KI-gestützte Datenbearbeitung direkt anwendbar.

Schweizer Datenschutzgesetz (DSG), insbesondere Transparenz, Zweckbindung, Datensicherheit und Datenschutz-Folgenabschätzung bei hohen Risiken.

Art. 321 StGB als besonderer Anker für Berufsgeheimnisse; für Board- und Mandatskontexte zusätzlich vertragliche und gesellschaftsrechtliche Geheimhaltungspflichten.

EU AI Act / Verordnung (EU) 2024/1689, soweit Unternehmen, Anbieter, Nutzer oder Märkte mit EU-Bezug betroffen sind.

> **Hinweis:** Dieser Draft ist ein Showcase und Startpunkt für Board-Gespräche. Vor produktivem Einsatz sollte die Weisung auf Branche, Datenlage, Tool-Stack, Konzernvorgaben, regulatorische Pflichten und konkrete Verantwortlichkeiten angepasst werden.

www.sallenbach.org | vero@sallenbach.org