

Bernd Oestereich (unter Mitarbeit von Claudia Schröder)

KI in der Organisation

Wie Organisationen Künstliche Intelligenz (ein-)führen, gestalten und verantworten



- Teil 1: KI-Grundbegriffe und -Konzepte
- Teil 2: Herausforderungen für Organisationen
- Teil 3: Wozu KI in der Organisation einsetzen?
- Teil 4: Das KI-System-Lagebild (zu klärende Kernfragen)
- Teil 5: Wo ist KI organisatorisch zu integrieren?
- Teil 6: KI-integrierende Organisationsprinzipien
- Teil 7: Der KI-Verantwortungskreis

Download: <https://kollegiale-fuehrung.de/kio-buch/>



Werkstatt
für kollegiale Führung



KI in der Organisation

Künstliche Intelligenz ist nicht einfach eine weitere Technologie im Werkzeugkasten der Organisation. Sie kommuniziert menschenähnlich, kann eigenständig handeln und wird in Organisationen zunehmend wie ein Mitglied des Teams wahrgenommen.

Gleichzeitig kann sie keine Verantwortung übernehmen, haftet nicht und verfügt – ohne externe Vorgaben – weder über eigene Intentionen noch über ein eigenes moralisches Urteilsvermögen. Wie Menschen produziert sie überzeugende Ergebnisse, die jedoch nicht immer belastbar sind. Allerdings tut sie dies mit einer Geschwindigkeit und Skalierbarkeit, die jede bisherige Technologie übertrifft.

KI überschreitet mühelos Bereichsgrenzen, Hierarchieebenen und etablierte Compliance-Strukturen. Sie entwickelt sich schneller, als viele Organisationen Schritt halten können, und wirft rechtliche, gesellschaftliche und ordnungspolitische Fragen auf, deren Dynamik sich unserer Kontrolle zunehmend entzieht.

Mit diesem Buch klären wir die zentralen Begriffe und Konzepte von KI-Systemen, zeigen, was sie leisten können, wo ihre Grenzen liegen und wodurch sie sich grundlegend von Menschen und klassischen Maschinen unterscheiden.

Das Buch bietet einen fundierten Überblick über die technischen, organisatorischen, rechtlichen und kulturellen Herausforderungen, die mit der Einführung von KI verbunden sind.

Es behandelt vor allem auch die Frage, wie Organisationen geführt und gestaltet werden müssen, wenn KI zu einem selbstverständlichen Teil der Wertschöpfung wird. Welche bewährten Prinzipien und Praktiken tragen weiter? Welche verlieren ihre Wirksamkeit? Und was kommt an ihre Stelle?

Dort finden sich dann praktische Anregungen, neue Organisations- und Führungsideen, Beispiele und Vorlagen, mit denen das Buch Orientierung für eine Zukunft gibt, die bereits begonnen hat.

Inhaltsverzeichnis

1	Grundlegende Begriffe und Konzepte.....	15
1.1	Was ist KI und wie funktioniert KI?	15
1.1.1	Künstliche neuronale Netze.....	15
1.2	Unterschiede zwischen KI und anderer Technik	18
1.2.1	Warum KI kein Algorithmus ist.....	18
1.2.2	KI ist keine bessere Software	18
1.2.3	Unterschied Digitalisierung – KI-sierung.....	19
1.2.4	KI – vom Werkzeug zur kommunikativen Rolle.....	21
1.3	Unterschiede zwischen Menschen und KI	22
1.3.1	Art des Verstehens – das biologische System	23
1.3.2	Art des Denkens – existenzielle Selbstbezogenheit	24
1.3.3	Art des Lernens – die Informationsbasis.....	26
1.3.4	Art des Verhaltens – die Zielgrößen.....	26
1.3.5	Art der Existenz – Zeitlichkeit und Kontinuität.....	28
1.3.6	Art des Weltbilds – strukturelle Voreingenommenheit.....	29
1.4	Systemtypen	31
1.4.1	Analytische KI	31
1.4.2	Generative KI	32
1.4.3	KI-Agent	33
1.4.4	Multiagentensysteme	35
1.4.5	KI-Assistent.....	38
1.4.6	Abgrenzung: Assistent – Agent – Multiagentensystem	39
1.5	Modellarchitekturen	40
1.5.1	Großes Sprachmodell (LLM – Large Language Model).....	40
1.5.2	Diffusionsmodell.....	42
1.5.3	Diskriminationsmodell	44
1.5.4	Hybride und spezielle KI-Architekturen.....	45

1.6	Erweiterungstechniken	45
1.6.1	Verhaltensprägung durch Anweisungstexte.....	45
1.6.2	Wissensanreicherung (RAG)	47
1.6.3	Gedächtnissysteme.....	49
1.6.4	Domänentraining (Finetuning).....	52
1.6.5	Werkzeugaufrufe und Verbindungsstandards.....	53
1.7	KI-Verarbeitungsstrategien.....	56
1.7.1	Schlussfolgerungsprozesse (Reasoning)	56
1.7.2	Explizite Denkschrittfolgen (Chain-of-Thought).....	56
1.7.3	Explizite Aufgabenplanung.....	57
1.7.4	Denk-Handlungs-Schleifen (ReAct-Ansatz)	57
1.7.5	Denkbaumverfahren (Tree of Thoughts).....	57
1.7.6	Selbstreflexion und Selbstkorrektur	57
1.8	Prinzipielle Funktionsweise von KI-Systemen.....	59
1.8.1	Modell und Modellanbieterin	59
1.8.2	Das KI-System um das Modell herum	60
1.8.3	Was alles im Kontextfenster steht	61
1.8.4	Wer gibt was vor: Anbieterin, Betreiberin, Nutzerin.....	63
1.8.5	Die Grenzen des Kontextfensters	63
1.8.6	Unteragenten und Orchestrierungs-Funktionen.....	64
1.8.7	Konnektoren.....	65
1.8.8	Sicherheits- und Filterfunktionen.....	65
1.8.9	Was das Prinzipbild organisatorisch bedeutet.....	66
1.9	Strukturformate lesen, verstehen und nutzen.....	67
1.9.1	Markdown – Struktur für Texte.....	67
1.9.2	YAML – Konfiguration für Menschen.....	67
1.9.3	JSON – Daten zwischen Systemen.....	68
1.9.4	Das Gemeinsame.....	68

2 Herausforderungen beim Einsatz von KI 69

2.1	Immanente KI-Eigenschaften	69
2.1.1	Nichtdeterminismus als Führungsproblem	70

2.1.2	Nicht-Erklärbarkeit von Entscheidungen.....	72
2.1.3	Konfabulationen (Halluzinationen).....	74
2.1.4	Bullshitting – Reden ohne Wahrheitsanspruch.....	77
2.1.5	Unwahre Statusmeldungen von KI-Agenten.....	78
2.1.6	Absichtsvolle Täuschung durch KI-Systeme.....	80
2.1.7	Datenvergiftung – wenn das Fundament bröckelt.....	81
2.1.8	Vortrainiertes vs. nachgeladenes LLM-Wissen.....	83
2.1.9	LLM-Wesensveränderungen (Model drift).....	86
2.1.10	Unbrauchbare Systemvergleiche.....	89
2.2	Verantwortung, Haftung und Entscheidung	90
2.2.1	Kaskadierende Verantwortungsverdünnung.....	91
2.2.2	Verantwortungsdiffusion durch KI-Vorschläge	92
2.2.3	(Wieder-)Trennung von Denken und Handeln.....	94
2.2.4	Haftbarkeit für KI-Handlungen	96
2.2.5	Datenschutz, Datensicherheit, Urheberrecht.....	99
2.2.6	Verantwortungsübergreifende KI-Entscheidungen.....	103
2.2.7	Verschwimmende Systemgrenzen.....	104
2.2.8	Verantwortung für die KI-Datenbasis	106
2.2.9	Systemverfügbarkeit externer Anbieter	107
2.3	Systemtheoretische Herausforderungen	110
2.3.1	Luhmanns Sicht auf Organisationen.....	110
2.3.2	Wer ist Organisationsmitglied?	112
2.3.3	Die Grenzfälle: Wer gehört wirklich dazu?.....	114
2.3.4	Die sozial-kommunikative Zurechnung zerfällt	117
2.3.5	Juristische Mitgliedschaft? Ist auch kein Ausweg.....	119
2.3.6	KI als funktionales Organisationselement	120
2.4	Wissen, Kompetenz und Lernen	124
2.4.1	Was KI mit Können macht	125
2.4.2	Wenn KI die Lernräume schließt.....	127
2.4.3	Der Verstärkereffekt und der Sog zur Mitte.....	130
2.4.4	Karriere und Können im KI-Zeitalter	131
2.4.5	Implizites menschliches Wissen	132
2.4.6	Unternehmenswissen KI-nutzbar machen	134
2.4.7	Unklarheit über Nutzung kontextspezifischer Daten.....	135

2.5	Psychologische Herausforderungen	137
2.5.1	Vermenschlichte KI-Kommunikation.....	138
2.5.2	Die verwerkzeuglichte KI als Kategorienfehler	141
2.5.3	Verantwortungs- und Wirksamkeitsillusion	143
2.5.4	Die Arbeitsverdichtungsfall.....	144
2.5.5	KI-Scheinarbeit: Wenn die Nacharbeit bei anderen landet	146
2.5.6	Schädlicher Produktivitätsdruck.....	146
2.5.7	Manipulation und Abhängigmachung	149
2.5.8	Das stille Gift der Harmonie	151
2.6	Menschliche (Meta-)Kommunikation im KI-Kontext	154
2.6.1	Kommunikative Grundfertigkeiten – auch ein KI-Thema....	154
2.6.2	Über das Gespräch sprechen: Metakommunikation	157
2.6.3	Vier Seiten einer Nachricht von Schulz von Thun	157
2.6.4	Die Axiome von Watzlawick als Kontrastfolie	158
2.6.5	Einseitige Metakommunikation.....	160
2.7	Organisation, Führung und Struktur	161
2.7.1	Problem: funktional und vorgesetzt geführt	162
2.7.2	Trennung von Arbeits- und Führungsorganisation	163
2.7.3	KI als Werkzeug oder als Ersatz für Menschen?	164
2.7.4	Globale Dysfunktion durch lokale Optimierungen.....	166
2.7.5	Der Verantwortungsengpass an der Spitze	169
2.7.6	Die Eisberg-Illusion (Strategie trifft Strukturen).....	170
2.7.7	Das Tempoproblem.....	172
2.7.8	Ziele als Scheuklappen.....	174
2.8	Sicherheit und Missbrauch	177
2.8.1	Bösartige Anweisungs-Injektionen	178
2.8.2	Lieferkettenangriffe mit Slopsquatting.....	182
2.8.3	Destruktive Handlungen von KI-Agenten.....	184
2.8.4	Unklare Identität von KI-Agenten	186
2.8.5	Cyberangriffe und Malware	188
2.8.6	Deepfakes, Menschen- und Ereignisfälschungen.....	190
2.8.7	Verschiebung krimineller Ressourcenökonomie	192
2.8.8	CBRN-Wissen.....	194
2.8.9	CSAM und übergreifige Inhalte	195

2.9 Außenwirkung, Markt und Gesellschaft	197
2.9.1 Die Kundschaft wird zum Engpass	198
2.9.2 Die Kundschaft schlägt zurück: KI als Gegenmacht.....	199
2.9.3 Reputationsrisiken.....	202
2.9.4 Gesellschaftspolitische Indoktrinationen	203
2.9.5 Der stille Raub an der Wissensarbeit	205
2.10 Regulatorischer Rahmen: EU KI-Verordnung.....	208
2.10.1 Das Risikoklassenmodell	208
2.10.2 Fast alle Organisationen sind sogenannte Betreiber	209
2.10.3 Pflichten für Organisationen	209
2.10.4 Die freiwillige Norm ISO/IEC 42001.....	212

3 Wofür KI in der Organisation einsetzen? 215

3.1 Braucht es hier überhaupt KI?	215
3.1.1 Digitalisiere wo möglich, KI-siere wo nötig	215
3.1.2 Eigene Handschrift oder brauchbarer Durchschnitt?	217
3.2 Ohne Vision für Wirkung steigt nur die Effizienz	219
3.2.1 KI-Einführung ist eine strategische Entscheidung.....	219
3.2.2 Wirkrichtung nach innen oder nach außen?.....	221
3.3 Nach innen: Was mit KI besser wird	222
3.3.1 Was mit KI konkret besser wird	222
3.3.2 Erfahrungswissen wird teilbar	223
3.3.3 Die entlastete Fachabteilung.....	223
3.3.4 Naheliegend: Effizienz und Automatisierung	224
3.3.5 Büro- und Wissensarbeit.....	224
3.4 Nach außen: Was bisher nicht ging	225
3.4.1 Suchrichtung Sofortigkeit	225
3.4.2 Suchrichtung Zugang	226
3.4.3 Suchrichtung Individualisierung.....	226
3.4.4 Suchrichtung Verlässlichkeit	227
3.4.5 Suchrichtung Vorausschau.....	227

3.4.6	Suchrichtung Erkenntnis	228
3.4.7	Zwei Fragen an jede dieser Richtungen.....	228
3.5	KI verändert die Zusammenarbeit mit dem Umfeld.....	230
3.6	Sich eine bessere Zukunft vorstellen	230

4 Das KI-Lagebild (zu klärende Kernfragen)..... 231

4.1	Was wissen wir über das konkrete KI-System?	232
4.1.1	Welches Basismodell liegt zugrunde?	232
4.1.2	Wo läuft das Modell physisch?	232
4.1.3	Welches Architekturmuster wurde gewählt?	233
4.1.4	Welche Schnittstellen hat das System?	233
4.1.5	Wo ist das Verhalten vorhersehbar, wo nicht?	233
4.1.6	Welche Teile bauen wir selbst, welche mieten wir?	233
4.2	Was weiß das KI-System?	234
4.2.1	Was steckt in den Trainingsdaten?	234
4.2.2	Welche Kontextdaten stellen wir zur Verfügung?	234
4.2.3	Welche Echtzeitdaten werden abgerufen?	235
4.2.4	Wie groß ist das Kontextfenster?	235
4.2.5	Was merkt sich das System in einer Konversation?	235
4.2.6	Was merkt es sich über Konversationen hinweg?	235
4.2.7	Was wird das System nicht wissen?	236
4.3	Was darf das KI-System tun?	236
4.3.1	Welche Daten darf es lesen?	236
4.3.2	Welche Daten darf es verändern?	237
4.3.3	Welche Funktionen und Werkzeuge kann es aufrufen?	237
4.3.4	In welcher Umgebung läuft das, was die KI tut?	237
4.3.5	Welche Freigaben sind eingebaut?	237
4.3.6	Was haben wir der KI ungewollt erlaubt?	238
4.4	Wer kontrolliert das KI-System?	238
4.4.1	Wer beobachtet, was hineingeht?	238
4.4.2	Wer beobachtet, was das System zwischendrin tut?	239

4.4.3	Wer beobachtet, was herauskommt?.....	239
4.4.4	Wo sitzt ein Mensch im Entscheidungspfad?.....	239
4.4.5	Welche Leitplanken sind eingebaut?	239
4.4.6	Welche regelmäßigen Tests gibt es?	240
4.4.7	Wer ist (in welcher Rolle) verantwortlich?.....	240
4.5	Was passiert, wenn etwas schief läuft?	240
4.5.1	Was wird protokolliert und wie lange?.....	240
4.5.2	Können wir Aktionen im Nachhinein rekonstruieren?	241
4.5.3	Was lässt sich (nicht) rückgängig machen?	241
4.5.4	Wer prüft regelmäßig auf Muster?.....	241
4.5.5	Gibt es automatische Sicherheitsmechanismen?.....	241
4.5.6	Wer erfährt wann was bei Überraschungen?	242
4.6	Wer ist mit welchen Interessen beteiligt?	242
4.6.1	Wer hat das Basismodell gebaut?	242
4.6.2	Wer bietet die konkrete Anwendung an?.....	242
4.6.3	Wer hat die KI in unsere Systeme eingebunden?.....	243
4.6.4	Woher kommen die Daten?.....	243
4.6.5	Welcher Rechtsrahmen gilt und wer gestaltet ihn mit?	243
4.6.6	Wer verändert das Modell und wie oft?	243
4.6.7	Welche fremden Interessen fließen ein?	243
4.6.8	Was, wenn der Anbieter wegbricht oder sich verändert?	244
4.6.9	Wie ausweichfähig sind wir?	244

5 Wo ist KI organisatorisch zu integrieren? 245

5.1	Unterscheidung: KI-Nutzungsstufen.....	245
5.1.1	Übersicht.....	246
5.1.2	Stufen 0 und 1: Keine oder einfache Werkzeugnutzung	246
5.1.3	Stufen 2 bis 4: Sozial-kommunikative Wirksamkeit	246
5.1.4	Stufe 5: Autonome KI und verteilte Verantwortung	247
5.1.5	Stufe 6: Unbeobachtete KI – ein Organisationsversagen.....	248
5.2	Unterscheidung: Datenkritikalität.....	249
5.2.1	Schutzwürdige Daten.....	249

5.2.2	Folgenrelevante Verarbeitungssituationen	250
5.3	Unterscheidung: KI-Wirkungsreichweite.....	251
5.3.1	Mensch (lokal)	251
5.3.2	Team.....	251
5.3.3	Bereich oder Organisation.....	251
5.3.4	Extern.....	251
5.4	Die Organisationsmuster-Matrix.....	252
5.4.1	Menschliche Verantwortung in gegebenen Rahmen.....	252
5.4.2	Die drei Achsen der Matrix.....	253
5.4.3	Die Matrix	253
5.4.4	Sechs kumulative Organisationsmuster.....	255
5.5	Die Organisationsmuster und ihre Sprünge.....	256
5.5.1	Muster 0: im Handlungsrahmen einer Person.....	256
5.5.2	Muster 1: Im Handlungsrahmen eines Teams	257
5.5.3	Muster 2: Mit Governance-Rolle im Team und Handlungsrahmen fürs KI-System.....	259
5.5.4	Muster 3: Mit KI-Verantwortungskreis	261
5.5.5	Muster 4: Teamübergreifende Verantwortung	263
5.5.6	Muster 5: Mit übergeordneter Freigabe.....	265
5.5.7	Abgrenzung	267
5.6	Menschliche Verantwortungsstufen.....	268
5.6.1	0) Unselbstständig, angeleitet	269
5.6.2	1) Selbstständig, unterstützt	270
5.6.3	2) Selbstständig, beobachtet	270
5.6.4	3) Eigenverantwortlich, kontextbewusst	270
5.6.5	4) Vollumfänglich verantwortlich	271
5.6.6	5) Strategisch verantwortlich	271
5.6.7	6) Visionär verantwortlich	272
5.6.8	7) Zwecksetzend mit unternehmerischen Risiken.....	272

6 KI-integrierende Organisationsprinzipien 273

6.1	Vorhersehbare und überraschende Führungsbedarfe	273
6.1.1	Wofür überhaupt Führung?.....	273
6.1.2	Warum tayloristische Führung nicht mehr passt.....	274
6.1.3	Wenn KI ins Spiel kommt.....	275
6.1.4	Agile Führung als ergänzendes Lösungsprinzip.....	276
6.1.5	Agile Führung ist Metaführung.....	277
6.2	Integrale, interdisziplinäre Teams.....	279
6.3	Rollen und Kreise.....	281
6.3.1	Unterscheidung Team und Kreis.....	281
6.3.2	Rollen als Verantwortungsbereiche	281
6.3.3	Arten von Rollen	282
6.3.4	Arten von Kreisen.....	283
6.4	Handlungsrahmen für Rollen und Kreise	284
6.4.1	Organisationen bestehen aus verschachtelten Handlungsrahmen.....	284
6.4.2	Übergreifende Verantwortungsstrukturen bilden	286
6.5	Konstitution von Teams und Rollen.....	288
6.5.1	Kreis-Konstitution	288
6.5.2	Rollen-Konstitution	293
6.6	Handlungsrahmen für KI-Systeme.....	294
6.6.1	Basis (unbedingt für jedes KI-System)	294
6.6.2	Standard (für die meisten KI-Systeme).....	295
6.6.3	Erweitert (für kritische oder stark autonome KI-Systeme).....	295
6.6.4	Handlungsrahmen für KI im Markdown-Format.....	298
6.6.5	Ein Beispiel: Einkaufsanfragen, teilautonom.....	298
6.7	Organisatorische KI-Konstitution.....	300
6.7.1	Basis-Elemente (für jedes formal eingeführte KI-System)	300
6.7.2	Standard-Elemente	301
6.7.3	Erweiterte Elemente	303
6.7.4	KI-Konstitution im Markdown-Format	304

6.8	Teamboard	304
6.9	Fallweise Ermächtigungen.....	306
6.10	Empirisch-rückgekoppeltes Vorgehen	308
6.10.1	1. Einsichten gewinnen	309
6.10.2	2. Hypothesen bilden.....	309
6.10.3	3. Handlungsoptionen erfinden	309
6.10.4	4. Veränderungen erproben	310
6.10.5	5. Entwicklungen etablieren.....	310

7 Der KI-Verantwortungskreis 311

7.1	KI-fachliche Rollen.....	312
7.1.1	Governance und Rahmensetzung.....	312
7.1.2	Kritische Beobachtung	313
7.1.3	Governance-Themen.....	314
7.1.4	Eingreifen	319
7.1.5	Vorfallsbearbeitung und Folgenbehebung	319
7.1.6	Datenbasis	320
7.2	Selbstorganisationsrollen.....	321
7.2.1	Kreishüterin	322
7.2.2	Lernbegleiterin (Teamcoach).....	324
7.2.3	Teamboard-Verantwortliche.....	326
7.2.4	Dokumentarin	327
7.2.5	Kreisökonomin.....	329
7.3	Was nicht in den KI-Verantwortungskreis gehört.....	331
7.4	Sonderfall: Die All-in-one-Governance-Rolle	331

8 Anhang	333
8.1 Über dieses Buch.....	333
8.1.1 Warum dieses Buch?	333
8.1.2 Für wen und wozu ist dieses Buch?	333
8.1.3 Wer hat dieses Buch geschrieben?	334
8.1.4 Wie ist dieses Buch geschrieben?	335
8.1.5 Wie ist das Buch aufgebaut?	337
8.1.6 Welchen Nutzen soll es haben?	337
8.2 Unterschiede zur 1. Auflage	338
8.2.1 Tabellarische Übersicht	338
8.2.2 Was sich in den Kapiteln 1 bis 4 geändert hat	339
8.2.3 Was neu dazugekommen ist: Kapitel 5 bis 7	340
8.2.4 Lohnt sich die 2. Auflage?	341
8.3 Quellen.....	342
8.4 Glossar	344