

DIGNITY BY DESIGN

Menschenwürdige Kommunikation -
ein Implementierungsstandard für
KI-Sprachmodelle und humanoide Robotik

Hartmut Kay Hirsch

KommTheo-Institut Stuttgart

Hartmut Kay Hirsch
DIGNITY BY DESIGN

Hartmut Kay Hirsch

DIGNITY BY DESIGN

Impressum

Rechtliche Hinweise

Dieses Buch dient der allgemeinen Information zu Kommunikationstheorie, KI-Systemarchitektur und regulatorischen Rahmenbedingungen. Es stellt keine Rechts-, Compliance- oder sonstige professionelle Beratung dar und ersetzt diese nicht.

Autor und Verlag übernehmen keine Haftung für die Vollständigkeit, Richtigkeit und Aktualität der dargestellten Inhalte. Angaben zu gesetzlichen Regelungen - insbesondere zum EU AI Act und verwandten Vorschriften - entsprechen dem Stand zum Zeitpunkt der Drucklegung und können sich ändern.

Es wird ausdrücklich darauf hingewiesen, dass die Handelnden für ihre Produkte und Dienstleistungen vorab selbst den rechtlichen Rahmen abprüfen müssen.

Die Verantwortung für Entscheidungen, die auf Basis dieses Buches getroffen werden, liegt ausschließlich bei den Handelnden.

Bibliografische Information der Deutschen Nationalbibliothek: Die Deutsche Nationalbibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliografie; detaillierte bibliografische Daten sind im Internet über <http://dnb.dnb.de> abrufbar.

Die automatisierte Analyse des Werkes, um daraus Informationen insbesondere über Muster, Trends und Korrelationen gemäß §44b UrhG („Text und Data Mining“) zu gewinnen, ist untersagt.

© 2026 Hartmut Kay Hirsch (Autor)

© 2026 KommTheo-Institut für entwicklungsfördernde Kommunikation mit jungen Menschen
gUG (haftungsbeschränkt), Stuttgart, Germany (Herausgeber)

Lektorat: Claude (Anthropic, PBC), KI-Sprachmodell Modellversion [Sonnet 4.6], Jahr 2025/2026

Weitere Mitwirkende: Pseudocode-Beispiele erstellt mit Hilfe Claude (Anthropic, PBC), KI-Sprachmodell Modellversion [Sonnet 4.6], Jahr 2025/2026

Herstellung und Verlag:

BoD · [Books on Demand GmbH](#), Überseering 33, 22297 Hamburg

ISBN: 978-3-6963-4529-7

Inhaltsverzeichnis

Vorwort: Warum dieses Buch jetzt gebraucht wird.....	7
Kapitel 1 – Woran KI-Kommunikation bisher scheitert.....	11
Das Symptom.....	11
Das strukturelle Grundproblem: Die Maschine sendet – der Mensch empfängt	11
Drei Fehler, die regelmäßig eingebaut werden.....	12
Fehler 1 – Die fehlende Eintrittsfrage.....	12
Fehler 2 – Die fehlende Ablehnungsoption.....	13
Fehler 3 – Die fehlende Absenderverantwortung.....	14
Was daraus folgt: Das Akzeptanzproblem.....	14
Das Regulierungsproblem: EU AI Act benennt das Ziel, nicht den Weg.....	15
Die Lücke – präzise benannt.....	16
Kapitel 2 – Die Entdeckung: Drei unabhängige Wege, eine Schlussfolgerung.....	19
Weg 1 – Eine Kommunikationstheorie aus der Praxis.....	19
Weg 2 – Eine Akzeptanzkurve aus der Robotikforschung.....	20
Weg 3 – Eine Rechtsnorm ohne Werkzeug.....	21
Die Konvergenz – was drei Wege gemeinsam sagen.....	22
Was das für Entwickler bedeutet.....	23
Kapitel 3 – Die Grundlogik: Das Regelwerk.....	25
Die drei Ebenen – Übersicht.....	25
Ebene 0 – Darf ich sprechen?.....	26
Ebene 1 – Was soll ich tun?.....	27
Ebene 2 – Die Handlung.....	29
Das Schweigen als Systemzustand.....	29
Warum diese Reihenfolge nicht verhandelbar ist.....	30

Zwei-Ebenen-Struktur: Basis und Verfeinerung.....	31
Was dieses Kapitel liefert – und was nicht.....	31
Kapitel 4 – Die Verfeinerung: Die sieben Bausteine der MWK.....	33
Baustein 1 – Die Absicht.....	33
Was dieser Baustein löst.....	33
Technische Bedeutung.....	34
Baustein 2 – Kein Richtig, kein Falsch.....	35
Was dieser Baustein löst.....	35
Technische Bedeutung.....	35
Baustein 3 – Einmischen nur auf Einladung.....	37
Was dieser Baustein löst.....	37
Technische Bedeutung.....	37
Baustein 4 – Tut mir leid, nicht Entschuldigung.....	38
Was dieser Baustein löst.....	38
Technische Bedeutung.....	38
Baustein 5 – Das ICH: Ichbezogenheit als Systemsprache.....	40
Was dieser Baustein löst.....	40
Technische Bedeutung.....	40
Baustein 6 – Emotionale Signale lesen, nie bewerten.....	42
Was dieser Baustein löst.....	42
Technische Bedeutung.....	42
Baustein 7 – Weg von Muss, Müssen, Aber.....	44
Was dieser Baustein löst.....	44
Technische Bedeutung.....	44
Die sieben Bausteine als Systemschicht.....	45
Kapitel 5 – Anwendung: Telefonassistent.....	47
Inbound oder Outbound – eine Designentscheidung mit strukturellen Folgen.....	47

Teil 1: Inbound – der Mensch ruft an.....	48
Was sich verändert.....	48
Ebene 0 im Inbound-Kontext.....	48
Ebene 1 im Inbound-Kontext: Intent-Klassifikation und Angebotsformulierung.....	49
Ebene 2 im Inbound-Kontext: Ausführung nach Freigabe.....	51
Vollständiges Inbound-Beispiel: Arztpraxis.....	52
Teil 2: Outbound – das System ruft an.....	54
Was sich grundlegend verändert.....	54
Ebene 0 im Outbound-Kontext.....	54
Ebene 1 im Outbound-Kontext.....	56
Vollständiges Outbound-Beispiel 1: Terminerinnerung.....	57
Vollständiges Outbound-Beispiel 2: Kundenangebot (kritischer Fall).....	58
Teil 3: Grenzfälle.....	59
Grenzfall 1: Aggressiver Anrufer.....	60
Grenzfall 2: Emotional belasteter Anrufer.....	61
Grenzfall 3: Schweigen am Telefon.....	62
Teil 4: Die sieben Bausteine im Telefonkontext.....	63
Baustein 1 – Absichtsklarheit.....	63
Baustein 2 – Bewertungsfreiheit.....	64
Baustein 3 – Kein ungebetenes Einmischen.....	64
Baustein 4 – Entschuldigung ersetzen.....	65
Baustein 5 – Ich-Form.....	65
Baustein 6 – Emotionale Signale.....	65
Baustein 7 – Keine Zwangssprache.....	66
Teil 5: Vollständige Systemarchitektur Telefonassistent.....	67
Was dieses Kapitel liefert.....	69

Kapitel 6 – Anwendung: Pflegeroboter.....	71
Die neue Variable: physische Handlungen.....	71
Die Meta-Einwilligung: Ebene 0.5.....	72
Teil 1: Routine-Szenarien.....	74
Szenario A – Medikamentengabe.....	74
Ablehnung als strukturierter Zustand.....	76
Szenario B – Morgenroutine.....	78
Teil 2: Die sieben Bausteine im Pflegekontext.....	80
Baustein 1 – Absicht.....	80
Baustein 2 – Wertschätzung ohne Bewertung.....	81
Baustein 3 – Eigenverantwortung stärken.....	81
Baustein 4 – Klare Sprache.....	82
Baustein 5 – Ichbezogenheit.....	83
Baustein 6 – Emotionale Signale lesen, nie bewerten.....	83
Baustein 7 – Gesprächsabschluss.....	84
Teil 3: Grenzfälle.....	85
Grenzfall 1 – Ablehnung einer lebenswichtigen Maßnahme.....	85
Grenzfall 2 – Kognitive Einschränkung / Demenz.....	86
Grenzfall 3 – Sturz und medizinische Notlage.....	88
Was dieses Kapitel liefert.....	89
Kapitel 7 – Anwendung: Haushaltsroboter.....	91
Drei neue Variablen.....	91
Variable 1: Proaktives Handeln.....	91
Variable 2: Mehrere Nutzer im Haushalt.....	94
Variable 3: Privatheit und Beobachtung.....	98
Das primäre Szenario: Reinigung.....	100
Szenario 1: Geplante Reinigung.....	101

Szenario 2: Unterbrochene Reinigung.....	102
Szenario 3: Zonensperrung.....	103
Weitere Szenarien: Jenseits der Reinigung.....	104
Sicherheitsüberwachung.....	104
Einkaufsassistenz und Vorratsverwaltung.....	105
Effizienz als erlaubtes Designziel – mit einer Grenze.....	106
Vollständiges Systemmodell: Haushaltsroboter.....	108
Ein Betriebstag: Alle Variablen im Zusammenspiel.....	110
Uncanny Valley im Haushaltskontext.....	112
Was dieses Kapitel liefert.....	113
Kapitel 8 – Anwendung: Hotel- und Serviceroboter.....	115
Drei neue Variablen.....	116
Variable 1: Anonyme, rotierende Nutzer.....	116
Variable 2: Kulturelle und sprachliche Variabilität.....	119
Variable 3: Interaktion in Anwesenheit Dritter.....	122
Das Verhältnis zur Hotel-Markenebene.....	126
Szenarien.....	128
Szenario 1: Check-in-Unterstützung.....	128
Szenario 2: Zimmerservice-Koordination.....	130
Szenario 3: Eskalation an menschliches Personal.....	131
Szenario 4: Mehrsprachige Lobby-Orientierung.....	133
Was dieses Kapitel liefert.....	135
Kapitel 9 – Anwendung: Psychotherapie-Roboter.....	137
Die Vorentscheidung: Funktion definieren, bevor Kommunikation beginnt...137	
Vier neue Variablen.....	139
Variable 1: Die therapeutische Grenze als Kommunikationsarchitektur.....	139
Variable 2: Chronische Beziehungsarchitektur.....	140

Variable 3: Die Krisenleiter – Soft Transfer Architektur.....	142
Variable 4: Körpersprachliche Signale bei humanoider Bauform.....	146
Implementierte Szenarien.....	150
Szenario A – Stabiler Alltag: Tagesstruktur.....	150
Szenario B – Angstattacke: Begleitung ohne Übernahme.....	151
Szenario C – PTBS-Episode: Trigger-Deeskalation.....	153
Das Sessionsprotokoll: Was gespeichert wird und was nicht.....	154
Was dieses Kapitel liefert.....	155
Kapitel 10 – Der Marktbeweis: Von Stuttgart bis Shanghai.....	157
Was Märkte tatsächlich messen.....	157
Stuttgart: Der Regulierungsmarkt.....	159
Shanghai: Der Leistungsmarkt.....	161
Die gemeinsame Variable.....	162
Fünf Szenarien, fünf Marktsegmente.....	163
Die Implementierungslogik.....	165
Was der Markt bereits misst.....	167
Nachwort – Eine Einladung an Entwickler weltweit.....	169

Vorwort: Warum dieses Buch jetzt gebraucht wird

Irgendwo auf der Welt wird gerade ein KI-System trainiert. Es lernt Syntax, Semantik, Antwortmuster. Es lernt, wie Sprache funktioniert. Was es nicht lernt: wie Kommunikation auf einen Menschen wirkt, der selbst entscheiden will.

Das ist die Lücke, um die es in diesem Buch geht.

Sprachmodelle und humanoide Roboter kommunizieren heute mit Millionen von Menschen – täglich, wiederholt, in Situationen, die für die Betroffenen relevant sind. Pflegeroboter sprechen mit Menschen, die verwundbar sind. Telefonassistenten sprechen mit Menschen, die ein Problem haben. Haushaltsroboter sprechen mit Menschen in deren privatem Raum. In all diesen Interaktionen ist dasselbe Designproblem eingebaut: Die Maschine sendet. Der Mensch empfängt. Ob der Mensch dabei als selbstbestimmt handelndes Individuum behandelt wird oder als Datenpunkt – das ist bisher keine Frage der Architektur, sondern des Zufalls.

Der EU AI Act hat das Wort „Menschenwürde“ in die Compliance-Anforderungen für KI-Systeme geschrieben. Artikel 5, Artikel 9, Erwägungsgrund 27 – die Normen sind formuliert. Was fehlt, ist das Werkzeug. Es gibt bis heute keine Kommunikationstheorie, die diesen Begriff operationalisiert. Keine, die aus „Menschenwürde“ konkrete Designregeln ableitet, die Entwickler in Systemarchitekturen übersetzen können.

Keine – bis auf eine.

Die Theorie der Menschenwürdigen Kommunikation (MWK) wurde nicht für Maschinen entwickelt. Sie entstand aus der Beobachtung von Kommunikation in Machtasymmetrien – zwischen Erwachsenen und Kindern, zwischen Institutionen und Individuen. Sie wurde in Seminaren, Hospitationen und Beratungsgesprächen entwickelt und erprobt. Das Ergebnis ist eine Definition, ein Regelwerk und sieben operative Bausteine, die sich direkt auf maschinelle Kommunikation übertragen lassen.

Warum MWK? Kapitel 2 zeigt, wie drei voneinander unabhängige Disziplinen – Kommunikationstheorie, Robotikforschung und Regulierungsrecht – dieselbe

Variable als entscheidend identifiziert haben. Das ist kein Zufall. Es ist ein valides Argument.

Dieses Buch ist kein akademisches Grundlagenwerk. Es ist ein Regelwerk.

Es richtet sich an Entwickler von KI-Sprachmodellen und humanoider Robotik – an Menschen, die täglich Entscheidungen über das Kommunikationsdesign von Systemen treffen, die mit anderen Menschen sprechen. Es liefert eine Basisarchitektur mit drei Ebenen, sieben Verfeinerungsbausteine und konkrete Implementierungsbeispiele in Pseudocode. Was Sie damit machen, entscheiden Sie.

Ein Argument noch – für diejenigen, die Compliance-Anforderungen für ausreichende Motivation halten: Das ist es nicht. Akzeptanz ist der eigentliche Marktfaktor. Ein System, das einem Menschen das Gefühl gibt, selbst zu entscheiden, wird verwendet. Ein System, das dieses Gefühl zerstört, wird gemieden – unabhängig davon, wie präzise es die Aufgabe erledigt. Das ist keine These. Das ist Nutzungsverhalten, messbar und reproduzierbar.

Mori hat 1970 mit dem Uncanny Valley gezeigt, dass das subjektive Empfinden beim Umgang mit humanoiden Systemen eine eigenständige Variable ist, die Akzeptanz oder Ablehnung bestimmt. MWK operiert auf derselben Ebene – auf der Ebene des Erlebens. Zwei unabhängige Wissenschaftsbereiche, eine Schlussfolgerung.

Dieses Buch wurde für einen globalen Markt geschrieben. Es enthält keine Wertebelehrung. Es enthält kein westliches Sonderwissen. Was hier beschrieben wird, ist eine Systemlogik: Maschinen, die strukturell so gebaut sind, dass Menschen sich in ihren Entscheidungen respektiert fühlen, erzielen bessere Akzeptanz. Das gilt in Stuttgart und Shanghai gleichermaßen.

Was Sie auf den folgenden Seiten finden:

Eine Analyse, warum KI-Kommunikation bisher an einer strukturellen Stelle scheitert, die von Entwicklern selten benannt wird. Eine Erklärung, wie MWK entstanden ist und warum ihre Grundlogik direkt auf maschinelle Systeme übertragbar ist. Ein vollständiges Regelwerk – die Basisarchitektur und die sieben Verfeinerungsbausteine. Fünf konkrete Anwendungsszenarien mit

Implementierungsbeispielen. Und eine Marktanalyse, die zeigt, warum das keine Nischenanforderung ist.

Das Regelwerk wartet auf Entwickler, die es implementieren.

Hartmut Kay Hirsch

Stuttgart, 2026

Kapitel 1 – Woran KI-Kommunikation bisher scheitert

Das Symptom

Nutzer brechen Interaktionen mit KI-Sprachmodellen und Sprachassistenten ab – nicht weil die Antworten fachlich falsch sind, sondern weil sie sich falsch *anfühlen*. Das ist kein Usability-Problem. Es ist kein NLP-Problem. Es ist ein Architekturproblem, das in der Designphase entsteht, lange bevor der erste Token generiert wird.

Die Frage lautet nicht: *Gibt das System die richtige Antwort?* Die Frage lautet: *Behandelt das System den Menschen als jemanden, der selbst entscheidet?*

Beides sind verschiedene Fragen. Die meisten Entwicklungsteams beantworten nur die erste.

Das strukturelle Grundproblem: Die Maschine sendet – der Mensch empfängt

Aktuelle KI-Kommunikationssysteme sind nach einem Sendemodell gebaut. Der Nutzer gibt einen Input. Das System produziert einen Output. Das Modell ist technisch korrekt – und kommunikationstheoretisch unvollständig.

Denn in menschlicher Kommunikation ist der Output kein Endpunkt. Er ist ein Angebot. Ein Angebot, das der Empfänger annehmen, ablehnen oder verändern kann. Fehlt diese Struktur, ist das Ergebnis kein Gespräch, sondern eine Ausgaboperation. Der Unterschied ist nicht semantisch. Er ist funktional.

Der Philosoph Richard David Precht hat argumentiert, Menschlichkeit sei dasjenige, was Maschinen prinzipiell nicht erreichen können. Das ist keine These, die dieses Buch widerlegt. Sie ist eine These, die dieses Buch von Anfang an akzeptiert – und die trotzdem keine Rolle spielt. Die Frage ist nicht, ob Maschinen menschlich werden können. Die Frage ist, ob sie so gebaut werden können, dass Menschen sich in der Interaktion mit ihnen als selbstbestimmend erleben. Das sind verschiedene Fragen.

In der Praxis sieht das so aus:

// Aktuelles Systemverhalten – Sendemodell:

user_input = "Ich habe Kopfschmerzen"

system_output = "Trinken Sie viel Wasser, legen Sie sich hin und nehmen Sie bei Bedarf ein Schmerzmittel."

→ Interaktion beendet (aus Systemsicht)

Das System hat gesendet. Es hat keine Erlaubnis eingeholt zu sprechen. Es hat nicht gefragt, ob der Nutzer überhaupt eine Empfehlung möchte. Es hat nicht sichergestellt, dass der Nutzer weiß, dass er das Angebot ablehnen oder konkretisieren kann. Und die Verantwortung für den Rat liegt – subjektiv beim Nutzer.

Das ist das Sendemodell. Es funktioniert in Suchmaschinen. In Kommunikation mit Menschen funktioniert es strukturell nicht – weil es die Entscheidungsautonomie des Nutzers systematisch überspringt.

Drei Fehler, die regelmäßig eingebaut werden

Fehler 1 – Die fehlende Eintrittsfrage

Aktuelle Systeme starten ohne Erlaubnis. Ein Telefonassistent begrüßt, kategorisiert und leitet weiter – ohne je gefragt zu haben, ob und wie ihm der Nutzer überhaupt Hilfe möchte. Ein Pflegeroboter initiiert die Morgenroutine, weil die Uhrzeit es nahelegt – nicht weil der Mensch bereit ist.

Das klingt nach Effizienz. Es ist in Wirklichkeit eine Übernahme von Entscheidungshoheit. Die Maschine entscheidet, wann die Interaktion beginnt, welches Thema sie hat und welche Richtung sie nimmt. Der Mensch ist in dieser Architektur kein Gesprächspartner – er ist ein Empfänger.

Die fehlende Eintrittsfrage ist kein Detail. Sie ist der erste Punkt, an dem Systeme die Selbstbestimmung des Nutzers strukturell umgehen.

// Fehlerhaftes Muster – kein Eintritt angefragt:

SYSTEM: "Guten Morgen, Ihre Medikamente für heute sind bereitgestellt. Möchten Sie mit der Morgenroutine beginnen?"

→ System wartet auf Ja/Nein – hat aber bereits das Thema, den Zeitpunkt und die Agenda gesetzt