CME-zertifizierte Fortbildung

KI in der Arztpraxis

Modul 4: Technologie, Automatisierung, Wirtschaftlichkeit und Zukunft

Teilnehmer-Handout

Referent: Dr. med. Christian Mauch
Facharzt für Orthopädie und Unfallchirurgie
Gründer und Geschäftsführer, SummitGO

Maybach Medical Akademie | November 2026
Landesärztekammer Baden-Württemberg | 10 CME-Punkte (5 Module)

Interessenkonflikt: Der Referent ist Gründer und Geschäftsführer von SummitGO, einer KI-Plattform für Arztpraxen (noch nicht am Markt, Start nach C5-Testat, geplant 2027). Die Fortbildung ist anbieterübergreifend und wissenschaftlich ausgewogen konzipiert. SummitGO wird im Vortrag als offengelegtes Fallbeispiel genannt und mit einem Sternchen (*) gekennzeichnet.

Kapitel I – Die heutige Arztpraxis

Dieses Kapitel beschreibt den Ausgangspunkt: Wie viel Arbeitszeit in einer niedergelassenen Praxis heute für Verwaltung aufgewendet wird, wo die Engpässe liegen und an welchen Stellen technische Lösungen bereits ansetzen.

Bürokratie und Dokumentation in Zahlen

Der Bürokratieindex der Kassenärztlichen Bundesvereinigung (KBV) beziffert den Aufwand, den eine Vertragsarztpraxis allein für Informationspflichten gegenüber der Selbstverwaltung aufbringen muss, auf rechnerisch 61 Arbeitstage pro Praxis und Jahr. Erfasst sind dabei ausschließlich Pflichten wie Bescheinigungen und Formulare, nicht die klinische Dokumentation. Die tatsächliche Gesamtlast liegt folglich deutlich höher. Die Daten stammen aus dem Datenstand 2020 der KBV.

Aus den Vereinigten Staaten liefert die Beobachtungsstudie von Sinsky und Kollegen (Annals of Internal Medicine, 2016) einen weiteren Bezugspunkt: In 16 Praxen aus vier Fachgebieten verbrachten 57 beobachtete Ärzte 49,2 Prozent ihrer Praxiszeit mit Aktenführung und Schreibtischarbeit, aber nur 27 Prozent im direkten Patientenkontakt. Diese Zahlen sind auf deutsche Verhältnisse nur bedingt übertragbar, weil US-amerikanische Aktensysteme und Abrechnungsregeln sich erheblich unterscheiden. Beide Erhebungen machen jedoch deutlich, dass administrative Tätigkeiten in der ambulanten Medizin einen großen Teil der verfügbaren Arbeitszeit binden.

Die Patientenreise – vom Anruf bis zur Abrechnung

Ein einzelner Patientenkontakt durchläuft in einer typischen Praxis eine Kette von Schritten: Anruf, Terminvergabe, Anamnese, Untersuchung, Dokumentation, Arztbrief und Abrechnung. An jedem dieser Schritte ist Personal gebunden. Für einige dieser Schritte – Telefonassistenz, Spracherfassung, Briefentwurf – gibt es heute technische Werkzeuge, die einen Teil der Arbeit übernehmen können. Die Prüfung und Freigabe bleiben beim Menschen.

In der Präsentation wird der Weg des Patienten in neun Stationen unterteilt: (1) Anruf mit KI-Telefonassistent, (2) intelligente Terminvergabe, (3) digitale Anamnese und Check-in, (4) KI-unterstütztes Arztgespräch, (5) automatisierte Dokumentation und Arztbrief, (6) OP-Planung, (7) klinische Entscheidungsunterstützung (CDSS), (8) Check-out mit Verordnungen und (9) Abrechnung und Kodierung. Jede Station wird in Kapitel IV im Detail behandelt.

Der Weg des Patienten durch die Praxis – fünf Stationen vor dem Sprechzimmer

Bereits bevor der Patient das Sprechzimmer betritt, durchläuft er fünf Stationen, an denen jeweils Personal gebunden ist und an denen technische Antworten existieren: (1) Erreichen – niemand geht ans Telefon; Antwort: Online-Rezeption und KI-Telefonassistenz. (2) Ankommen – Schlange am Empfang; Antwort: Self-Check-in per eGK am Kartenterminal mit Versichertenstammdatenabgleich über die Telematikinfrastruktur. (3) Warten – unklar wohin; Antwort: Zuweisung per Push-Nachricht oder Bon. (4) Aufruf – bisher durch Rufen; Antwort: Aufrufsystem mit Push-Nachricht und Raumnummer. (5) Sprechzimmer – alles wird neu erzählt; Antwort: Die Angaben aus Telefonat, Check-in und Anamnese liegen bereits vor.

Rechtlich ist seit dem 2. August 2026 zu beachten, dass sich eine KI-Telefonassistenz nach Art. 50 der KI-Verordnung (EU) 2024/1689 als KI zu erkennen geben muss. Vor der Inbetriebnahme sollte die Praxis Notfallsätze wie „Brustschmerz“ oder „Atemnot“ testen und eine Durchwahl zu einem Menschen offenhalten.

Was kostet ein verpasster Anruf?

Die Präsentation stellt eine Modellrechnung vor, keine Messung: 30 nicht angenommene Anrufe pro Tag, multipliziert mit einem angenommenen Anteil von 20 Prozent, der zur Konkurrenz abwandert, multipliziert mit einem durchschnittlichen Fallwert von 60 Euro und 220 Arbeitstagen pro Jahr, ergibt einen rechnerischen Umsatzausfall von 79.200 Euro pro Jahr. Alle Werte sind Annahmen und lassen sich durch die eigenen Praxiszahlen ersetzen.

Die Präsentation weist darauf hin, dass eine unabhängige, begutachtete Messung verpasster Anrufe in deutschen Praxen nicht vorliegt. Die verfügbaren Quellen stammen fast ausschließlich von Anbietern: AGFEO (20–40 Prozent), aiagens (30–40 Prozent), digital-rezeption.de (30–40 Prozent, Spitze 50 Prozent), medi-one (30–50 Prozent für Physiotherapie). Keine dieser Quellen legt eine Methodik offen. Unabhängig ist nur die GKV-Versichertenbefragung 2024 des GKV-Spitzenverbands, nach der 44 Prozent der Versicherten mit der telefonischen Erreichbarkeit unzufrieden sind; das ist eine Zufriedenheitsangabe, keine Anrufquote.

Die technische Prozesskette hinter einer KI-Anwendung

Hinter jeder scheinbar einfachen KI-Funktion stehen mehrere technische Komponenten, die zuverlässig zusammenarbeiten müssen. Die typische Architektur einer medizinischen KI-Anwendung umfasst sechs Schritte: (1) Spracheingabe über Mikrofon (Diktiermikrofon, Smartphone oder Raummikrofon), (2) Automatic Speech Recognition (ASR), also die Umwandlung von Sprache in Text durch Modelle wie Whisper, Azure Speech oder Google Speech, (3) De-Identifikation, also das Entfernen oder Ersetzen von Patientennamen, Geburtsdaten und Versichertennummern durch Platzhalter, bevor der Text an ein externes Modell gesendet wird (Pseudonymisierung, nicht Anonymisierung), (4) Verarbeitung durch ein Sprachmodell (LLM) wie GPT, Claude, Gemini, Llama oder MedGemma, (5) strukturierte Ausgabe in Formaten wie JSON-Schema, HL7 FHIR oder KBV-MIOs und (6) Integration in das Praxisverwaltungssystem über Schnittstellen wie GDT oder HL7 FHIR.

Jede Komponente verursacht eigene Kosten, erfordert Wartung und erzeugt potenzielle Fehlerquellen. Die Prüffrage an jeden Anbieter lautet: Was verlässt die Praxis, und in welcher Form?

Kapitel II – Infrastruktur, Cloud und C5

Bevor über KI-Funktionen gesprochen werden kann, müssen die Grundlagen der Infrastruktur geklärt sein: Wo laufen medizinische KI-Systeme, welcher Rechtsrahmen gilt und was bedeutet das C5-Testat?

Der Rechtsrahmen – drei Fragen, fünf Gesetze

Vor jedem KI-Einsatz helfen drei Fragen, den Rechtsrahmen zu ordnen. Frage 1 betrifft das Werkzeug selbst: Darf es das, was es tun soll? Maßgeblich sind die Medizinprodukteverordnung (MDR, EU 2017/745) und die KI-Verordnung (KI-VO, EU 2024/1689). Zuständig ist der Hersteller, der den Zweck festlegt, die Risikoklasse bestimmt und das CE-Zeichen erhält. Frage 2 betrifft die Daten: Darf die Praxis Patientendaten hineingeben? Die Datenschutz-Grundverordnung (DSGVO, insbesondere Art. 9 und Art. 28) schützt die Daten; § 203 des Strafgesetzbuchs schützt das ärztliche Berufsgeheimnis. Beide gelten nebeneinander, weshalb ein Verstoß sowohl ein Bußgeld nach der DSGVO als auch eine strafrechtliche Sanktion nach sich ziehen kann. Zuständig ist die Praxis als Verantwortliche. Frage 3 betrifft den Ort: Wo liegen die Daten, und wer kommt heran? Hier greifen § 393 SGB V mit der Pflicht zum C5-Testat sowie die Datentransferregeln der DSGVO (Kapitel V). Der Anbieter weist nach, die Praxis prüft.

Risikoklassen der KI-Verordnung

Die KI-Verordnung verfolgt einen risikobasierten Ansatz und teilt KI-Systeme in vier Klassen ein:

Klasse 1 – Verboten (Art. 5, seit 2. Februar 2025): Darunter fallen Social Scoring, Emotionserkennung am Arbeitsplatz und in Schulen sowie Manipulation, die Schwächen von Kindern, Älteren oder Kranken ausnutzt. Für die Praxis bedeutet das: Kein Werkzeug, das Mitarbeitende per Kamera oder Stimme auf Gefühle auswertet. Sogenannte Nudify-Apps sind ab dem 2. Dezember 2026 verboten.

Klasse 2 – Hochrisiko (Art. 6, Anhänge I und III): KI in Medizinprodukten, Kraftfahrzeugen und Maschinen; KI für Bewerberauswahl, Kreditwürdigkeit und Prüfungsnoten. Für die Praxis relevant ist vor allem der Bereich Diagnose mit Patientendaten, etwa eine Hautkrebs-App, ein EKG-Analyseprogramm oder eine Röntgen-KI. Pflicht: Anleitung befolgen, Ergebnis ärztlich prüfen. Geltung: 2. August 2028 für Anhang I, 2. Dezember 2027 für Anhang III.

Klasse 3 – Begrenzt (Art. 50, Transparenz, seit 2. August 2026): Chatbots, Sprachassistenten und KI-erzeugte Bilder, Videos oder Stimmen (Deepfakes). Für die Praxis: Ein Telefon- oder Terminassistent muss sich als KI zu erkennen geben (Kennzeichnungspflicht).

Klasse 4 – Minimal (keine Zusatzpflichten, seit 2. Februar 2025): Spamfilter, Übersetzung, Rechtschreibprüfung, Navigations-App. Einzige Pflicht: Das Team muss geschult werden (Art. 4 KI-VO, KI-Kompetenzpflicht seit 2. Februar 2025).

Die Klasse richtet sich nach dem Zweck des Werkzeugs, nicht nach den Daten. Werden Patientendaten verarbeitet, gelten in jeder Klasse zusätzlich DSGVO Art. 9, § 203 StGB und ein Auftragsverarbeitungsvertrag. Wer ein Werkzeug zweckfremd nutzt – etwa ChatGPT für Diagnosen – kann nach Art. 25 der KI-VO selbst als Anbieter gelten. Die Fristen gelten nach dem Digital Omnibus on AI (Rat, 29. Juni 2026).

Hosting – wo KI-Programme laufen

Hosting bedeutet, dass ein Programm nicht auf dem Praxisrechner läuft, sondern auf gemieteten Rechnern in einem Rechenzentrum. Die Praxis oder ihr Softwareanbieter zahlt für Rechenleistung, anstatt Server selbst zu kaufen. Es gibt vier grundlegende Optionen:

Lokal (On-Premise): Volle Datenkontrolle, keine Cloud-Abhängigkeit, aber hohe Hardwarekosten, Wartungsaufwand und begrenzte Skalierung. Beispiele: Dell Technologies, HPE, Lenovo, NVIDIA DGX.

Deutsches Rechenzentrum: DSGVO-konform, Datenhaltung in Deutschland, professioneller Betrieb, aber höhere Kosten als Hyperscaler und begrenzte GPU-Verfügbarkeit. Beispiele: Hetzner, IONOS, STACKIT (Schwarz-Gruppe), T Cloud Public (Telekom).

Europäische Cloud: Skalierbar, GPU-Verfügbarkeit, Managed Services, aber vertragliche Absicherung nötig und Abhängigkeit vom Anbieter. Beispiele: Exoscale (Schweiz), OVHcloud (Frankreich), Scaleway (Frankreich), UpCloud (Finnland).

Internationale Hyperscaler: Maximale Skalierung, modernste GPU-Hardware, breites Service-Ökosystem, aber US Cloud Act Risiko, EU-Region erforderlich und komplexe Verträge. Beispiele: Amazon Web Services (AWS), Google Cloud, Microsoft Azure.

Meist bucht nicht die Praxis selbst, sondern der Softwareanbieter. Die entscheidende Frage lautet: In welchem Rechenzentrum hostet mein Anbieter, und wer ist der Vertragspartner?

Drei deutsche Cloud-Anbieter im Vergleich

	IONOS	STACKIT	Hetzner
Betreiber	IONOS SE, Montabaur	Schwarz-Gruppe, Neckarsulm	Hetzner Online, Gunzenhausen
Rechenzentren	Deutschland; KI in Berlin	Deutschland, Österreich	Nürnberg, Falkenstein, Helsinki
Sicherheitsnachweis	BSI C5, ISO 27001	BSI C5 Typ 2, ISO 27001	ISO 27001, kein C5
KI-Angebot	Offene Modelle (Llama, Mistral)	Offene Modelle (Llama, Gemma)	GPU-Server, Modell selbst betreiben
Preis (netto)	0,15/0,65 € je Mio. Token	0,45/0,65 € je Mio. Token	ca. 232 €/Monat + 114 € Einrichtung

Rechenbeispiel: Ein Arztbrief (etwa 2.000 Token Eingabe, 800 Token Ausgabe) kostet rund 0,14 Cent; 50 Briefe am Tag ergeben etwa 16 Euro im Jahr. Teuer ist die Software, nicht die Rechenleistung. Wichtig: Das C5-Testat der Infrastruktur gilt nicht automatisch für die Anwendung, die darauf läuft. Preisstand: Oktober 2026.

Hyperscaler und Modellanbieter

Der Begriff „Hyperscaler" bezeichnet Betreiber von Rechenzentren, die ihre Kapazität sehr schnell und nahezu beliebig erweitern können. Dazu gehören Amazon (AWS), Microsoft (Azure), Google Cloud sowie im weiteren Sinne Meta und Oracle. Sie bauen Rechenzentren, kaufen Chips und sichern sich Strom. Die vier größten planen für 2026 Investitionen von rund 730 Milliarden Dollar (2025: rund 410 Milliarden Dollar).

Davon zu trennen sind die Modellanbieter wie OpenAI (ChatGPT), Anthropic (Claude), Google (Gemini) und xAI (Grok). Sie entwickeln die Sprachmodelle, besitzen aber meist keine eigenen großen Rechenzentren und mieten die Rechenleistung bei den Hyperscalern. Die Hyperscaler sind häufig zugleich als Investoren beteiligt: Microsoft bei OpenAI, Amazon und Google bei Anthropic.

Als Beispiel für die wirtschaftliche Entwicklung nennt die Präsentation Anthropic: Der Umsatz betrug im ersten Halbjahr 2026 laut Börsenprospekt (S-1 bei der SEC) 16,3 Milliarden Dollar – das 14-Fache des Vorjahreswerts von 1,2 Milliarden Dollar. Der operative Verlust sank auf 0,7 Milliarden Dollar (Gesamtjahr 2025: 8,1 Milliarden). Der in den Medien genannte Nettoverlust von 80,7 Milliarden Dollar ist zu rund 72 Milliarden ein Buchungseffekt aus der Neubewertung von Wandelanleihen; Geld ist dabei nicht abgeflossen.

Für die Praxis bedeutet das: Die heutigen Preise müssen nicht dauerhaft gelten. Wer KI einführt, sollte darauf achten, dass Modell und Anbieter austauschbar sind und die eigenen Daten exportiert werden können.

Das C5-Testat – jedes Glied der Kette

Nach § 393 SGB V (seit 1. Juli 2025) darf eine Praxis Gesundheitsdaten nur dann über einen Cloud-Dienst verarbeiten, wenn die datenverarbeitende Stelle ein aktuelles C5-Typ-2-Testat des Bundesamts für Sicherheit in der Informationstechnik (BSI) vorweisen kann. C5 steht für „Cloud Computing Compliance Criteria Catalogue". „Typ 2" bedeutet, dass ein Wirtschaftsprüfer über mehrere Monate geprüft hat, ob die Sicherheitsmaßnahmen tatsächlich wirken – nicht nur, ob sie beschrieben sind (Typ 1).

Das Testat wird für jedes Glied der Kette benötigt: (1) Das Rechenzentrum, in dem Programm und Daten liegen. (2) Der Anbieter des Sprachmodells, falls ein externer Dienst das Modell betreibt. (3) Der

Softwareanbieter selbst, also die Anwendung, die die Praxis nutzt. Das Testat des Hosters reicht für die Anwendung nicht aus; auch ein kleines Unternehmen braucht ein eigenes Testat für seine Software. Für neue Systeme genügt in den ersten 18 Monaten ein Typ-1-Testat. Der Standort muss im Inland, in der EU oder im EWR liegen; in Drittstaaten nur mit Angemessenheitsbeschluss, und der Anbieter braucht eine Niederlassung in Deutschland.

Welche Unterlagen die Praxis braucht

Vor dem ersten Einsatz mit Patientendaten sollte die Praxis die folgenden Unterlagen anfordern, prüfen und ablegen:

Unterlage	Rechtsgrundlage	Woher
C5-Typ-2-Testat jedes Glieds	§ 393 SGB V	vom Anbieter anfordern
Auftragsverarbeitungsvertrag (AVV) mit Unterauftragnehmer-Liste	Art. 28 DSGVO	vom Anbieter
Technische und organisatorische Maßnahmen (TOM)	Art. 32 DSGVO	Anlage zum AVV
Verpflichtung zur Verschwiegenheit	§ 203 StGB	im Vertrag mit dem Anbieter
Eintrag im Verzeichnis der Verarbeitungstätigkeiten (VVT)	Art. 30 DSGVO	erstellt die Praxis
Datenschutz-Folgenabschätzung (DSFA)	Art. 35 DSGVO	Praxis, mit Datenschutzbeauftragtem
Patienteninformation; Einwilligung bei Gesprächsaufzeichnung	Art. 13 DSGVO, § 201 StGB	erstellt die Praxis
Schulungsnachweis des Teams	Art. 4 KI-VO	erstellt die Praxis

Fehlt das Testat eines Glieds, darf die Anwendung für Patientendaten nicht genutzt werden. Keine Rechtsberatung; im Einzelfall den Datenschutzbeauftragten einbeziehen.

Kapitel III – Sprachmodelle und KI-Komponenten

Offene und geschlossene Sprachmodelle

Bei offenen Modellen ist die Modelldatei frei verfügbar; bezahlt wird nur die Rechenleistung, entweder bei einem Hoster oder auf dem eigenen Server. Beispiele sind Llama (Meta), Mistral, Gemma (Google) und gpt-oss (OpenAI). Bei geschlossenen Modellen gibt der Hersteller die Modelldatei nicht heraus; man nutzt sie per App oder Schnittstelle, und der Preis enthält Modell und Rechenleistung. Beispiele sind ChatGPT (OpenAI), Claude (Anthropic) und Gemini (Google).

Der Kostenunterschied pro Arztbrief ist gering: Selbst die Spitzenklasse kostet pro Brief nur etwa sechs Cent. In der eigenen Messung des Referenten mit Wissensbasis lag ein Brief in der Mittelklasse bei 3,9 Cent. Entscheidend sind deshalb Qualität, Datenweg und Vertrag, nicht der Tokenpreis.

Benchmarks – Medizin-Examenstest (MedQA)

MedQA ist ein Testsatz aus 1.273 Multiple-Choice-Fragen im Stil des US-Staatsexamens (USMLE) mit jeweils vier Antwortmöglichkeiten. Zufälliges Raten ergibt 25 Prozent. Die in der Präsentation genannten Ergebnisse: MedGemma 27B (Google, offene Gewichte) 87,7 Prozent; GPT-4o (OpenAI, geschlossen) 86,5 Prozent; Meditron 70B (EPFL, 2023, offene Gewichte) 70,2 Prozent; MedGemma 1.5 4B (Google, 2026, offene Gewichte) 69,1 Prozent.

MedGemma 1.5 4B ist besonders bemerkenswert, weil es mit nur 4 Milliarden Parametern klein genug für den Praxisrechner ist. Es läuft ohne Cloud auf eigener Hardware, ist multimodal (Text plus Röntgen, CT/MRT, Dermatologie und Pathologie) und bietet mit MedASR eine durchgehende Pipeline von der Sprache zur Analyse.

Benchmark-Warnung: Prüfungsfragen messen Faktenwissen, nicht klinische Sicherheit. Die Werte stammen überwiegend aus Herstellerberichten.

Eigene Messung – neun Modelle im Arztbrief-Test

Der Referent ließ neun Sprachmodelle dieselben zehn erfundenen orthopädischen Fälle als Arztbrief schreiben und benotete je zehn Briefe verdeckt (das Modell war dem Bewerter nicht bekannt). Insgesamt wurden 90 Briefe bewertet, im September 2026.

Modell	Art	Note	Einsen	Kosten/Brief
Gemini 2.5 Pro	geschlossen	1,30	7/10	4,6 Ct
Claude Sonnet 5	geschlossen	1,30	7/10	3,9 Ct
Claude Sonnet 5, Frankfurt	geschlossen	1,30	7/10	5,2 Ct
Claude Haiku 4.5	geschlossen	1,60	4/10	1,0 Ct
GPT-5.4-mini	geschlossen	1,80	4/10	1,0 Ct
Gemini 2.5 Flash	geschlossen	1,80	3/10	1,1 Ct
Claude Opus 5	geschlossen	1,90	2/10	9,9 Ct
gpt-oss-120b (OpenAI)	offen	2,40	2/10	nur Server
Qwen3-14B (Alibaba)	offen	2,60	0/10	nur Server

Die Rangfolge insgesamt ist statistisch gesichert (Friedman-Test p = 0,00025), der Abstand zwischen zwei einzelnen Modellen jedoch nicht. Claude Sonnet lieferte über die USA und über Frankfurt dieselbe Note – der Standort änderte die Qualität hier nicht. Der teuerste Brief (Opus, 9,9 Cent) war nicht der beste. Grenzen der Messung: ein Bewerter, erfundene Fälle, ein Sprecher, und der Bewerter ist selbst Anbieter (Interessenkonflikt offengelegt).

Ein offenes Modell kam an Claude Sonnet heran

Im Oktober 2026 testete der Referent das offene Modell Gemma (Google) auf einem eigenen Server in Deutschland gegen Claude Sonnet. Bei den Dosisfehlern in 60 Briefen lag Gemma gleichauf (9 vs. 8–9 Fehler), bei der falschen Wirkstärke sogar besser (0 vs. je 4). Beim Nachlesen von 36 echten Diktaten war Gemma 12-mal besser, Claude 7-mal, der Unterschied statistisch nicht belegt. Gemma war schwächer beim OP-Bericht (nur halb so lang) und langsamer (25 vs. 18 Sekunden). Der eigene Server kostet rund 1.500 Euro im Monat; für denselben Betrag gäbe es rund 38.000 Briefe mit Claude Sonnet.

Warnung: Die sparsamere Fassung von Gemma schrieb in allen drei Durchläufen Pantoprazol 200 statt 20 mg – eine zehnfache Dosis. Wer auf ein offenes Modell setzt, muss es selbst prüfen. Insgesamt wurden 20 offene Modelle versucht, 14 liefen auf dem eigenen Server. In einer Probe mit verhörter Wirkstärke (vier Diktate) fing Gemma 4 von 4 ab, zwei Modelle je 1, die übrigen 11 keines.

De-Identifikation – drei Generationen

Bevor Patientendaten an ein externes Sprachmodell gesendet werden, müssen identifizierende Merkmale entfernt oder durch Platzhalter ersetzt werden. Die Präsentation unterscheidet drei Generationen dieses Verfahrens:

Generation 1 (feste Regeln): Erkennt nur Formate wie Daten oder Telefonnummern, übersieht aber Namen im Freitext.

Generation 2 (Erkennungsmodell, z. B. Microsoft Presidio): Erkennt Namen mit einem trainierten Modell (F1-Score über 0,95 im i2b2-Test), übersieht aber indirekte Hinweise. Beispiel: „Die Bürgermeisterin von Bort" ist auch ohne Namensnennung identifizierbar.

Generation 3 (Sprachmodell, lokal betrieben): Versteht den Zusammenhang und entfernte in der NEJM-AI-Studie 2025 99,24 Prozent der Personendaten, benötigt aber starke Hardware.

Wichtig: Auch Generation 3 ist nur eine Pseudonymisierung – die Daten bleiben personenbezogen. Nur echte Anonymisierung befreit von DSGVO-Pflichten.

Die Kostenstruktur einer KI-Lösung – fünf Bausteine

Ein Diktat durchläuft fünf Schritte, jeder mit eigenen Kosten:

Baustein	Funktion	Kosten je Nutzung	Anteil an Gesamtkosten
1. Transkription (ASR)	Sprache zu Text	0,02–0,10 €/Min.	15 %
2. De-Identifikation	Namen entfernen	0,01–0,05 €/Dok.	5 %
3. Sprachmodell (LLM)	Text zum Arztbrief	0,01–0,30 €/Aufruf	25 %
4. Hosting & Infrastruktur	Server/Cloud rund um die Uhr	450–1.830 €/Monat	35 %
5. Integration & Support	PVS-Anbindung, Updates, Hilfe	100–500 €/Monat	20 %

Das Sprachmodell kostet nur Cent-Beträge je Aufruf; die Monatsrechnung bestimmen Hosting und Integration (zusammen 55 Prozent). Die Empfehlung: Fragen Sie Anbieter nach den monatlichen Gesamtkosten, nicht nach dem Preis pro Aufruf. Eigene Modellrechnung, Evidenzkategorie D.

Kapitel IV – Die automatisierte Patientenreise

Dieses Kapitel begleitet einen fiktiven Patienten – Herrn Müller, 58 Jahre, Knieschmerzen rechts seit drei Monaten, Verdacht auf Gonarthrose – durch alle neun Stationen einer orthopädischen Gemeinschaftspraxis (drei Ärzte, sechs MFA, rund 120 Patienten am Tag).

Station 1: KI am Telefon

Die KI-Telefonassistenz nimmt Anrufe entgegen und leitet sie strukturiert weiter. Die Herausforderung liegt im Dialog an sieben Stellen: (1) Hören – Dialekt (etwa Schwäbisch), Mischsprache, Akzent, Handy und Störgeräusche. (2) Identifizieren – Name buchstabieren, Geburtsdatum, Abgleich mit der Patientenkartei. (3) Wiedererkennen – bei Mehrfachanrufen Dubletten vermeiden, Erreichbarkeit per Mobilnummer oder E-Mail sichern. (4) Verstehen – Laienwörter („es zieht", „Ischias"), Umgangssprache, mehrere Anliegen in einem Satz. (5) Einordnen – Rezept (einfach), Termin (welcher Arzt, welches Fach), Dringlichkeit (Triage); wer triagiert, betreibt ein Medizinprodukt (MDR). (6) Weitergeben – Schnittstelle zum PVS, Online-Terminkalender. (7) Rückmelden – SMS, Rückruf oder Buchungslink, Terminbestätigung.

KI-Telefonassistenten am Markt (alphabetisch, ohne Wertung): 321 MED, Aaron (Doctolib), CGM, DocMedico, Dr.wait, fonio.ai, medflex, Mediform, PraxisConcierge, VITAS.

Vor dem Start sollte die Praxis Testanrufe im Dialekt der eigenen Patienten durchführen, Notfallsätze testen und eine Durchwahl zu einem Menschen offenhalten.

Station 2: Intelligente Terminvergabe

Ein Termin bringt sechs Merkmale mit: Wer (Neu-/Bestandspatient), Versicherung (GKV, PKV, BG-Unfall, Selbstzahler), bei wem (Arzt, Fachrichtung), was (Erstvorstellung, Verlaufskontrolle, MRT-Besprechung etc.), wann (geplant, ungeplant, akut, Notfall) und womit (Dauer, Raum, Gerät, Vorbereitung). Herr Müller: Neupatient, GKV, Knie, Dr. Weber, 30-Minuten-Slot Erstvorstellung, Röntgen vorher.

Zwei Wege führen in den Kalender: Weg A führt von außen in freie Slots über Online-Rezeptionen (321 MED, arzt-direkt, CGM, DocMedico, Doctolib, MediDesk, Mediform, PraxisConcierge) und Online-Terminbuchung (Doctena, Doctolib, Dr. Flex, jameda, Samedi, Termed). Weg B führt über den Terminkalender im PVS (CGM/Medistar/ClickDoc, dubidoc, medatixx/x.time, tomedo). Das Hauptproblem: Beide Systeme müssen dieselben Terminregeln kennen, sonst bucht ein Neupatient eine 10-Minuten-Wiedervorstellung statt eines 30-Minuten-Slots.

Station 3: Digitale Anamnese und Check-in

Das Ziel ist eine lückenlose Informationskette, wenn der Arzt den Raum betritt. In der Realität entsteht häufig eine Papierlücke: Die Terminbestätigung bittet den Patienten, einen Fragebogen vorab auszufüllen, was die wenigsten tun. In der Praxis wird dann auf Papier (Abtippen nötig), auf einem Praxis-iPad (fremdes Gerät, neue Umgebung) oder auf dem eigenen Handy (kleiner Bildschirm, Fehleingaben) nachgeholt. Im Anamnese-Dialog führt die MFA oder eine KI das Gespräch, und die Antworten landen strukturiert im PVS.

Die Informationskette umfasst: den letzten Besuch (Diagnosen, Medikation, Verlauf), das Telefonat (Anliegen aus Station 1, KI-Notiz), vorab gesendete Unterlagen (Überweisung, Medikationsplan) und Befunde (Labor, Röntgen, MRT). Wenn der Arzt den Raum betritt, sollen Anliegen, Vorgeschichte, Befunde und offene Fragen auf einer Seite liegen. Keine Frage wird zweimal gestellt.

Anbieter in drei Gruppen: Anamnese-Spezialisten (Anaboard, Idana, infoskop, Medudoc, Mediform, myMedax, Nelly, Simpleprax), Plattformen mit Anamnese-Modul (DocMedico, Doctolib, Samedi, SummitGO*) und PVS mit eigenem Modul (CGM/Medistar, medatixx, tomedo).

Station 4: Das KI-unterstützte Arztgespräch

Die KI nimmt vier Quellen auf: (1) Das mitgehörte Gespräch (Ambient AI: Patient, MFA und Arzt, Sprecher erkannt). (2) Den diktierten Befund (klinische Untersuchung direkt in den Befundtext). (3) Die Aufklärung (OP-Aufklärung mit Inhalt, Risiken und Fragen). (4) Die Vorinformationen aus der Anamnese (PC, HPC, PMH).

Der Eintrag folgt einer Karteikartenlogik: A (Anamnese), B (Befund), D (Diagnose), P (Prozedere), T (Therapie). Die Abrechnung richtet sich nach dem Versicherungsstatus: GKV nach EBM, PKV nach GOÄ, IGeL nach GOÄ plus Vertrag, BG nach UV-GOÄ, Facharztverträge (FAV) nach § 140a SGB V (seit 2015, früher § 73c). Der Arzt prüft und gibt frei. Eine Patienteneinwilligung vor der Aufzeichnung ist erforderlich.

KI-Dokumentation am Markt (alphabetisch): Dragon Copilot (Microsoft), Heidi Health, Idana Flow Pro, ORPHEUS, Tandem Health, tomedo, SummitGO*.

Station 5: Arztbrief und strukturierte Dokumentation

Aus dem Karteieintrag (A, B, D, P, T aus Station 4) entstehen Briefe an den Hausarzt, an den Patienten (in einfacher Sprache) und gegebenenfalls an die Berufsgenossenschaft. Der Arzt prüft und signiert jeden Brief.

Studien zeigen drei typische Fehlerarten:

Auslassung: Williams und Kollegen (JAMA Internal Medicine, 2025, n = 100) fanden, dass KI-Entwürfe knapper und klarer, aber lückenhafter waren: 2,91 statt 1,82 Fehler je Brief, wobei das Schadenspotenzial in beiden Gruppen niedrig war. Gegenmaßnahme: Pflichtfelder prüfen (Diagnose, Medikation, Prozedere).

Erfindung (Halluzination): Song und Kollegen (JAMA Network Open, 2025, Notaufnahme) berichteten von einem Entwurf, der eine Schiene und einen Verband nannte, die nie angelegt wurden. Fünf von sechs Ärzten strichen die Passage, der unerfahrenste nicht. Gegenmaßnahme: Jede Maßnahme im Brief gegen das Gespräch abgleichen.

Konjunktiv-Verwechslung: In einem Charité-Pilot 2025 erkannte die KI Möglichkeitsformen unsicher; aus einem Verdacht konnte im Brief eine gesicherte Diagnose werden. Gegenmaßnahme: „V. a." (Verdacht auf) und gesicherte Diagnosen gezielt kontrollieren.

Station 7: Klinische Entscheidungsunterstützung (CDSS)

Ein klinisches Entscheidungsunterstützungssystem wird nach der Medizinprodukteverordnung (MDR, Regel 11) dann zum Medizinprodukt, wenn es einen Diagnose- oder Therapievorschlag zum konkreten Patienten macht (Klasse IIa oder höher). Reines Nachschlagen in Leitlinien oder Lehrbüchern ist in der Regel kein Medizinprodukt. Ein allgemeiner Chatbot wie ChatGPT oder Claude mit einem Patientenfall gilt ebenfalls nicht als Medizinprodukt – das Risiko trägt dann der Arzt. Für die Behandlung haftet der Arzt nach § 630a BGB; die Haftung für KI-Fehler ist juristisch ungeklärt.

Zur Zuverlässigkeit liefern drei Studien Orientierung:

Wu und Kollegen (Nature Communications, 2025): 50 bis 90 Prozent der KI-Antworten werden von den genannten Quellen nicht vollständig gestützt. Ärzte fanden nur 40 Prozent voll belegt. Gegenmaßnahme: Quelle öffnen und den Satz suchen (1–2 Minuten je Quelle).

Vishwanath und Kollegen (Nature Medicine, 2026): Fachwerkzeuge wie OpenEvidence (89,6 Prozent) und UpToDate Expert AI (88,4 Prozent) erreichten im MedQA-Examenstest hohe, aber nicht perfekte Werte. Allgemeine Modelle lagen sogar höher. Examensfragen sind jedoch kein Praxisfall.

NOHARM-Benchmark (Stanford/Harvard, 2026): 31 Systeme wurden an echten Konsilfragen getestet. Der beste Wert: 86 Prozent korrekt und vollständig, 2,9 Prozent schwere Fehler. Gegenmaßnahme: Bei Therapieentscheidungen die Leitlinie selbst nachlesen.

CDSS-Anbieter am Markt (alphabetisch): AMBOSS (LiSA), DynaMed (EBSCO), Elsevier ClinicalKey AI, Isabel DDx, OpenEvidence (nicht in der EU), UpToDate Expert AI, SummitGO*.

Station 8: Verordnungen und Check-out

Der Arzt legt das Prozedere im PVS fest und signiert. Die Software soll den Rest erledigen: Drucken (Physio-Verordnung, Patienteninfo, Terminzettel), TI-Dienste (eRezept, eAU, eÜberweisung), ePA (Arztbrief und Befunde), Termin und Recall (Kontrolle in 6 Wochen). Durchgängig gibt es das noch nicht.

Drei Hürden auf dem Weg zum automatischen Check-out: (1) Die Prüflogik muss stimmen: Bei Heilmitteln müssen ICD-10 und Diagnosegruppe zur Richtlinie passen; ohne zweiten ICD-Code für den besonderen

Verordnungsbedarf zählt die Verordnung ins Budget (§ 106b SGB V). (2) Die ePA ist schwer zu pflegen: Nach dem Stecken der Karte sieht die Praxis 90 Tage alles; Verbergen geht nur für alle Praxen, nicht für eine einzelne; viele Dokumente sind PDFs, gesucht wird nur in Metadaten. (3) Ein Prozedere-Vorschlag für einen konkreten Patienten ist ein Medizinprodukt nach MDR Regel 11.

Station 9: Abrechnung und Kodierung

Abgerechnet wird nur, was dokumentiert ist. Drei Angaben müssen zusammenpassen: die Leistung (Therapieeintrag in der Kartei), die Gebührenordnung (abhängig vom Kostenträger: EBM, GOÄ oder UV-GOÄ) und die Diagnose (ICD-10-GM mit Sicherheit G/V/A/Z und Seite R/L/B). Passt eine Angabe nicht, ist die Abrechnung angreifbar.

Ein Patient kann am selben Tag drei Kostenträger haben: GKV nach EBM (Konsultation und Sonografie), Selbstzahler nach GOÄ (z. B. Hyaluronsäure-Injektion als IGeL, schriftliche Vereinbarung nach § 18 Abs. 8 BMV-Ä erforderlich) und BG nach UV-GOÄ (Verdrehtrauma als eigener Unfallfall). Die KI kann die Gebührenordnung je Leistung vorschlagen, Diagnosen kodieren und Lücken melden (dokumentiert, aber nicht abgerechnet und umgekehrt).

Kapitel V – Anbieter, Kosten und Wirtschaftlichkeit

Der deutsche KI-Anbietermarkt

Die Präsentation zeigt eine Nachweistabelle für 14 Anbieter (einschließlich SummitGO) mit den Kriterien: eigenes C5-Testat, Hosting, Auftragsverarbeitungsvertrag (AVV), Verpflichtung nach § 203 StGB und Monatskosten. Die Tabelle unterscheidet drei Kategorien: belegt (✓), eingeschränkt (✓ in Klammern, z. B. nur Typ 1 oder nur Infrastruktur) und nicht öffentlich belegt (? anfragen).

Hervorzuheben ist, dass seit dem 1. Juli 2025 nach § 393 SGB V ein C5-Typ-2-Testat des Dienstes selbst erforderlich ist; ein Testat nur des Rechenzentrums reicht nicht. Die Verpflichtung auf § 203 StGB ist Pflicht der Praxis und wird durch den AVV nicht ersetzt. Die Präsentation betont ausdrücklich, dass diese Maßstäbe auch für SummitGO gelten.

Was KI die Beispielpraxis kostet

Die Beispielpraxis: drei Orthopäden, sechs MFA (zwei in Teilzeit), rund 120 Patienten und 80 Anrufe am Tag. MFA-Stunde rund 23 Euro (Median 3.129 Euro brutto plus rund 21 Prozent Arbeitgeberanteil, 167 Stunden im Monat). Laufende KI-Kosten: Dokumentation 240–840 Euro (3 Ärzte × 79–279 Euro), Telefon 40–230 Euro (je Praxis), zusammen 300–1.100 Euro im Monat.

Nicht auf der Rechnung stehen: Prüfzeit (jeder Entwurf wird gelesen und freigegeben), Einführung (Schulung, Vorlagen, Probelauf: einige Wochen), PVS-Schnittstelle (kann je nach System extra kosten) und Datenschutz (AVV, § 203 StGB, Einwilligung). Für die Bilanz rechnet die Präsentation mit rund 700 Euro im Monat.

Was KI der Beispielpraxis bringt

Die Präsentation identifiziert fünf Nutzenposten: (1) Weniger MFA-Stunden am Telefon: rund 900 Euro (40 Stunden × 22,70 Euro), aber nur wenn Überstunden oder eine Stelle wegfallen. (2) Arztzeit für Patienten: 300–2.250 Euro (Studienwert 6–10 Fälle mehr × 50 Euro; Obergrenze bei 15 Stunden = 45 Fälle). (3) Weniger verpasste Anrufe: rund 1.000 Euro (Annahme: 1 Fall pro Tag zurückgewonnen × 20 Tage × 50 Euro). (4) Feierabend statt Nacharbeit: 15 Arztstunden Freizeit, kein Geldwert. (5) Vollständigere Dokumentation: In den USA +5,8 Prozent abgerechnete Leistung (Holmgren, JAMA Network Open, 2026), in der GKV durch das Budget gedeckelt.

Geld bringt KI nur, wenn die Praxis aktiv etwas ändert: weniger Stunden einplanen, freie Termine füllen, Anrufe annehmen lassen. Sonst bleibt es Entlastung ohne finanzielle Wirkung.

Studienwerte – nur eine Untergrenze

Die drei zitierten Studien messen eine Untergrenze, nicht die Erwartung:

Rotenstein und Kollegen (JAMA, 2026): 8.581 Behandler, 5 US-Kliniken. Ergebnis: minus 16 Minuten Dokumentation je 8-Stunden-Sprechstunde. Einschränkungen: Nur 21 Prozent nutzten das Werkzeug überhaupt, höchstens 5 Monate beobachtet.

Lukac und Kollegen (NEJM AI, 2025): 238 Ärzte, randomisiert, 2 Monate. Ergebnis: minus 44 Sekunden je Notiz (minus 9,5 Prozent). Einschränkungen: Werkzeug nur in 30–34 Prozent der Besuche genutzt, Vertrag begrenzte Dauer und Nutzerzahl, Zeit im KI-Programm selbst nicht erfasst.

Holmgren und Kollegen (JAMA Network Open, 2026): 1.565 Ärzte, 1 Klinik. Ergebnis: plus 0,8 Patienten je Woche und Arzt. Einschränkungen: Freiwillige Nutzung, Frühwender, Selbstauswahl, nur eine Klinik.

Nicht gemessen wurden: Telefon, Anmeldung, Arztbrief, Kodierung, Verordnung – gerade dort entsteht in der Praxis die meiste Arbeit. Erfahrungsberichte aus Praxen nennen 2 bis 4 Stunden Teamzeit am Tag allein durch den Telefonassistenten (geschätzt, nicht gemessen).

ROI – drei Szenarien

Die Präsentation rechnet drei Szenarien bei rund 700 Euro monatlichen KI-Kosten:

Szenario A – Mehr Freizeit: Ertrag 0 Euro. Die 15 Arztstunden werden Freizeit. Bilanz: minus 700 Euro. Die Praxis bezahlt die Freizeit mit rund 47 Euro je gewonnener Stunde (700 Euro geteilt durch 15).

Szenario B – Mehr Patienten: Ertrag 300–500 Euro (Studienwert: 6–10 Fälle mehr × 50 Euro). Bilanz: minus 200 bis minus 400 Euro. Werden alle 15 Stunden mit Patienten gefüllt (rund 45 Fälle à 20 Minuten), ergibt sich rund plus 1.550 Euro, in der GKV aber durch das Budget begrenzt.

Szenario C – Weniger MFA-Stunden: Ertrag rund 900 Euro (40 Stunden × 22,70 Euro = 0,25 Stelle). Bilanz: plus 200 Euro. Das ist die eigentliche Ersparnis, aber nur wenn Überstunden entfallen oder eine Stelle nicht nachbesetzt wird.

A und C lassen sich verbinden: Die Ärzte gewinnen Feierabend, das Team gewinnt Zeit am Telefon. Ergebnis: rund plus 200 Euro und 5 Stunden je Arzt und Monat. Drei Zahlen der eigenen Praxis entscheiden: Anrufe am Tag, Anteil der Briefe mit KI, Wert eines zusätzlichen Falls.

Einzelwerkzeuge oder integrierte Plattform?

Einzelwerkzeuge erlauben das jeweils beste Werkzeug pro Aufgabe, geringere Abhängigkeit und flexibleren Wechsel, haben aber den Nachteil, dass kein Datenfluss zwischen ihnen besteht, mehrere Verträge und Zugänge nötig sind und der Schnittstellenaufwand bei der Praxis liegt. Integrierte Plattformen bieten durchgängigen Datenfluss, einen Vertrag und eine konsistente Datenschutzlösung, bedeuten aber Abhängigkeit von einem Anbieter, nicht überall das beste Werkzeug und einen höheren Einstiegspreis. SummitGO verfolgt den Plattform-Ansatz (offengelegtes Eigeninteresse). Ein dritter Weg: Das PVS bleibt, Einzelwerkzeuge docken über offene Schnittstellen an.

PVS-Schnittstellen

Das PVS entscheidet, welche KI-Werkzeuge genutzt werden können, denn jedes Werkzeug muss Patientendaten aus dem PVS lesen und Ergebnisse zurückschreiben. Die Präsentation beschreibt acht Schnittstellenarten:

GDT (Gerätedatentransfer): Patient hinaus, Befund als Dokument zurück – der Standardweg für Zusatzprogramme. BDT/xDT: Stamm- und Behandlungsdaten, vor allem beim PVS-Wechsel. LDT: Laborbefunde in die Akte. HL7 v2: Nachrichten zwischen Klinik, Radiologie und PVS. HL7 FHIR/REST-API: Liest und schreibt einzelne Felder, Grundlage moderner Anbindung. AW-SST (§ 371 SGB V): Pflicht seit 2021, Archiv und Systemwechsel auf FHIR-Basis. TI-Dienste: KIM, eRezept, eAU, ePA über den Konnektor, für Drittanbieter nicht offen. Herstellerschnittstelle: Eigene API oder Partnerprogramm, Zugang nur per Vertrag.

Drei Stufen der Anbindung: Stufe 1 – Text kopieren. Stufe 2 – Dokument ablegen (GDT). Stufe 3 – Felder in die Karteikarte schreiben (FHIR, API). Die Frage an jeden Anbieter lautet: Welche Stufe läuft mit meinem PVS, und wer bezahlt die Freischaltung?

Kapitel VI – Recht, Sicherheit und ärztliche Verantwortung

Was der KI-Anbieter liefern muss

Vier Unterlagen muss der Anbieter vor Vertragsschluss liefern: (1) Nutzungsvereinbarung – regelt Funktionsumfang, Pflichten, Haftung und Laufzeit. (2) Auftragsverarbeitungsvertrag (AVV) nach Art. 28 DSGVO mit TOM, Subunternehmer-Liste und Weisungsbefugnissen. (3) Verschwiegenheitsverpflichtung nach § 203 StGB für alle Personen beim Anbieter mit Zugang zu Patientendaten. (4) C5-Testat nach § 393 SGB V.

Dabei ist nicht nur der KI-Anbieter selbst nachweispflichtig, sondern die gesamte Lieferkette: Auch der Cloud- oder Serveranbieter (z. B. Hetzner, AWS, Azure) als Sub-Auftragsverarbeiter braucht ein eigenes C5-Testat, einen AV-Vertrag mit dem KI-Anbieter, einen aktuellen Eintrag in der Subunternehmer-Liste und eine TOM-Dokumentation.

Fünf Pflichten der Praxis

Auch wenn der Anbieter alle Unterlagen mitbringt, bleiben fünf Pflichten bei der Praxis:

(1) Verarbeitungsverzeichnis ergänzen (VVT, Art. 30 DSGVO): Neuen Eintrag für KI-Dokumentation anlegen. (2) Datenschutzinformation aktualisieren (DSI, Art. 13 DSGVO): Patienten über KI-Einsatz informieren (Aushang, Website). (3) Datenschutz-Folgenabschätzung prüfen (DSFA, Art. 35 DSGVO): Bei Gesundheitsdaten und KI in der Regel erforderlich. (4) Einwilligungen einholen (EW, § 201 StGB/DSGVO): Aufnahmeeinwilligung und vor jedem Gespräch bei Ambient AI. (5) KI-Schulung durchführen (Art. 4 KI-VO): Alle Mitarbeiter mit KI-Kontakt schulen, Pflicht seit 2. Februar 2025.

Checkliste: Bevor Sie unterschreiben

Vom Anbieter erhalten?	Selbst erledigt?
☐ Nutzungsvereinbarung	☐ VVT-Eintrag angelegt
☐ AV-Vertrag mit TOM-Anlage	☐ Patienteninformation aktualisiert
☐ Subunternehmer-Liste	☐ DSFA geprüft / durchgeführt
☐ § 203 StGB Verpflichtung	☐ Einwilligungsformulare erstellt
☐ C5-Testat (§ 393 SGB V)	☐ Mitarbeiter geschult (KI-VO)
☐ Datenschutzerklärung des Anbieters	

Erst wenn alle Punkte erfüllt sind, sollte der Vertrag unterschrieben werden.

KI-Risiken

Die Präsentation benennt sechs Risiken beim Einsatz von KI in der Arztpraxis:

Halluzinationen (Risikostufe HOCH): Plausible, aber inhaltlich falsche Aussagen. Das Sprachmodell erzeugt Text, der grammatisch korrekt und fachlich plausibel klingt, aber sachlich falsch sein kann.

Unvollständige Befunde (HOCH): Fehlende relevante Informationen. Das Modell kann Teile des Diktats oder der Vorinformation übergehen, insbesondere wenn sie am Rand einer langen Eingabe stehen.

Falsche Negationen (KRITISCH): Pathologie wird nicht erkannt oder nicht erwähnt. Dies ist das gefährlichste Risiko, weil es beim flüchtigen Lesen am wenigsten auffällt.

Bias (MITTEL): Verzerrungen durch Trainingsdaten. Modelle können systematisch bestimmte Bevölkerungsgruppen, Geschlechter oder Altersgruppen benachteiligen, wenn die Trainingsdaten nicht repräsentativ sind.

Datenschutzverletzung (HOCH): Personenbezogene Gesundheitsinformationen werden bei der Cloud-Übertragung nicht vollständig entfernt.

Automation Bias (HOCH): Unkritische Übernahme von KI-Ergebnissen. Dies ist nach Einschätzung der Präsentation das tückischste Risiko: Der Arzt übernimmt den KI-Output, ohne ihn gründlich zu prüfen, weil er davon ausgeht, dass das System zuverlässig ist.

Die Lösung: Ein ärztlicher Freigabe- und Qualitätssicherungsprozess als unverzichtbarer Bestandteil jeder KI-Integration. Jeder Entwurf wird gelesen, geprüft und freigegeben, bevor er die Praxis verlässt.

Glossar und Abkürzungsverzeichnis

Ambient AI
Technologie, bei der ein KI-System das Arzt-Patienten-Gespräch im Hintergrund mithört, die Sprecher erkennt und daraus automatisch einen strukturierten Dokumentationsentwurf erzeugt.

ASR (Automatic Speech Recognition)
Automatische Spracherkennung; wandelt gesprochene Sprache in maschinenlesbaren Text um. Beispiele: Whisper (OpenAI), Azure Speech (Microsoft), Google Speech.

AV-Vertrag / AVV (Auftragsverarbeitungsvertrag)
Vertrag nach Art. 28 DSGVO, der die Verarbeitung personenbezogener Daten durch einen Auftragnehmer regelt, einschließlich technisch-organisatorischer Maßnahmen (TOM) und Unterauftragnehmer.

AW-SST (Archiv- und Wechselschnittstelle)
Pflichtschnittstelle nach § 371 SGB V (seit 2021), die den Datenaustausch und Systemwechsel zwischen PVS auf FHIR-Basis ermöglicht.

BDT / xDT
Datenaustauschformate für Stamm- und Behandlungsdaten in der ambulanten Medizin, vorrangig beim PVS-Wechsel eingesetzt.

BMV-Ä
Bundesmantelvertrag für Ärzte; Rahmenvertrag zwischen Kassenärztlicher Bundesvereinigung und GKV-Spitzenverband.

BSI
Bundesamt für Sicherheit in der Informationstechnik, zuständig u. a. für den C5-Prüfkatalog.

C5 (Cloud Computing Compliance Criteria Catalogue)
Prüfkatalog des BSI für Cloud-Dienste. Typ 1 beschreibt die Sicherheitsmaßnahmen, Typ 2 bestätigt deren Wirksamkeit über einen Prüfzeitraum durch einen Wirtschaftsprüfer.

CDSS (Clinical Decision Support System)
Klinisches Entscheidungsunterstützungssystem; Software, die Ärzte bei Diagnose- oder Therapieentscheidungen unterstützt, z. B. durch Leitlinien-Abgleich.

CE-Kennzeichnung
Kennzeichnung, dass ein Produkt die Anforderungen der EU-Verordnungen (z. B. MDR für Medizinprodukte) erfüllt.

Cloud Act (Clarifying Lawful Overseas Use of Data Act)
US-Gesetz, das US-Behörden Zugriff auf Daten ermöglicht, die von US-Unternehmen gespeichert werden, auch wenn die Server außerhalb der USA stehen.

CME (Continuing Medical Education)
Gesetzliches Pflichtfortbildungssystem für Ärzte in Deutschland. 250 CME-Punkte in 5 Jahren sind für Vertragsärzte verpflichtend.

COI (Conflict of Interest)
Interessenkonflikt; in CME-Fortbildungen muss der Referent wirtschaftliche Verbindungen offenlegen.

De-Identifikation / Pseudonymisierung
Verfahren, bei dem identifizierende Merkmale (Name, Geburtsdatum, Versichertennummer) durch Platzhalter ersetzt werden, bevor Daten an externe Systeme gesendet werden. Die Daten bleiben personenbezogen.

DSFA (Datenschutz-Folgenabschätzung)
Risikoanalyse nach Art. 35 DSGVO, bei der systematisch geprüft wird, ob eine Datenverarbeitung hohe Risiken für Betroffene birgt.

DSGVO (Datenschutz-Grundverordnung)
Verordnung (EU) 2016/679 zum Schutz personenbezogener Daten. Regelt u. a. Auftragsverarbeitung (Art. 28), besondere Kategorien (Art. 9, Gesundheitsdaten), Informationspflichten (Art. 13) und Verarbeitungsverzeichnis (Art. 30).

EBM (Einheitlicher Bewertungsmaßstab)
Gebührenordnung für die Abrechnung ärztlicher Leistungen in der GKV.

eGK (Elektronische Gesundheitskarte)
Versicherungskarte der gesetzlichen Krankenversicherung mit Chip, die am Kartenterminal eingelesen wird.

eHBA (Elektronischer Heilberufsausweis)
Ausweis für Ärzte zur elektronischen Signatur, u. a. für eRezept und eAU.

ePA (Elektronische Patientenakte)
Die gesetzliche, krankenkassenübergreifende digitale Patientenakte nach § 341 SGB V.

Evidenzkategorien (in der Präsentation)
A = unabhängige wissenschaftliche Studie, B = Beobachtungsstudie/Herstellerbericht, C = Herstellerangabe ohne Studie, D = eigene Modellrechnung des Referenten.

F1-Score
Maß für die Genauigkeit eines Erkennungsmodells; harmonisches Mittel aus Präzision und Trefferquote. Werte nahe 1,0 sind sehr gut.

FHIR (Fast Healthcare Interoperability Resources)
Moderner HL7-Standard für den Austausch von Gesundheitsdaten über REST-APIs; Grundlage der AW-SST und zunehmend für PVS-Integrationen.

GDT (Gerätedatentransfer)
Standardisiertes Schnittstellenformat in der ambulanten Medizin für die Kommunikation zwischen PVS und Zusatzgeräten/-programmen.

GKV
Gesetzliche Krankenversicherung.

GOÄ (Gebührenordnung für Ärzte)
Gebührenordnung für die Abrechnung ärztlicher Leistungen bei Privatpatienten und IGeL.

GPU (Graphics Processing Unit)
Grafikprozessor, der wegen seiner parallelen Rechenarchitektur für das Training und den Betrieb von Sprachmodellen eingesetzt wird.

Halluzination (im KI-Kontext)
Eine Ausgabe des Sprachmodells, die grammatisch und fachlich plausibel klingt, aber sachlich falsch ist – z. B. eine Maßnahme, die nie stattgefunden hat.

HL7 (Health Level 7)
Internationale Normenreihe für den Datenaustausch im Gesundheitswesen. Version 2 nutzt Nachrichtenformate, FHIR nutzt REST-APIs.

Hyperscaler
Betreiber von Rechenzentren, die ihre Kapazität sehr schnell und nahezu beliebig erweitern können (Amazon AWS, Microsoft Azure, Google Cloud).

i2b2
Referenzdatensatz für die Evaluation von De-Identifikationsverfahren im klinischen Bereich.

ICD-10-GM
Internationale statistische Klassifikation der Krankheiten, deutsche Modifikation; Pflichtkodiersystem in der ambulanten und stationären Versorgung.

IGeL (Individuelle Gesundheitsleistung)
Medizinische Leistung, die nicht von der GKV übernommen wird und vom Patienten selbst bezahlt werden muss. Abrechnung nach GOÄ.

KBV
Kassenärztliche Bundesvereinigung; Dachorganisation der 17 Kassenärztlichen Vereinigungen.

KI-VO (KI-Verordnung)
Verordnung (EU) 2024/1689 zur Regulierung künstlicher Intelligenz. Risikobasierter Ansatz mit vier Klassen (verboten, Hochrisiko, begrenzt, minimal). Art. 4: KI-Kompetenzpflicht seit 02/2025. Art. 50: Kennzeichnungspflicht seit 08/2026.

KIM (Kommunikation im Medizinwesen)
Sicherer E-Mail-Dienst der Telematikinfrastruktur für den Austausch medizinischer Dokumente.

LDT (Labordatentransfer)
Standardisiertes Format für die Übermittlung von Laborbefunden an das PVS.

LLM (Large Language Model)
Großes Sprachmodell; KI-System, das auf riesigen Textmengen trainiert wurde und natürliche Sprache verstehen und erzeugen kann (z. B. GPT, Claude, Gemini, Llama).

MDR (Medical Device Regulation)
Verordnung (EU) 2017/745 über Medizinprodukte. Regel 11 bestimmt, wann Software (einschließlich KI) als Medizinprodukt eingestuft wird.

MedQA
Benchmark-Testsatz mit 1.273 Multiple-Choice-Fragen im Stil des US-Staatsexamens (USMLE) zur Bewertung medizinischer KI-Modelle.

MFA (Medizinische Fachangestellte)
Ausgebildete Fachkraft in der Arztpraxis, früher Arzthelferin.

MIO (Medizinische Informationsobjekte)
Von der KBV definierte Standards für strukturierte medizinische Dokumente (Impfpass, Mutterpass etc.) in der ePA.

NOHARM
Benchmark von Stanford und Harvard (2026) zur Bewertung klinischer Entscheidungsunterstützungssysteme an echten Konsilfragen.

PKV
Private Krankenversicherung.

Presidio
Open-Source-Werkzeug von Microsoft zur Erkennung und Entfernung personenbezogener Daten in Texten (De-Identifikation der Generation 2).

PVS (Praxisverwaltungssystem)
Die zentrale Software einer Arztpraxis für Patientenakte, Terminverwaltung, Abrechnung und Kommunikation. Beispiele: CGM Medistar, medatixx, tomedo, T2med.

§ 203 StGB
Verletzung von Privatgeheimnissen; stellt die ärztliche Schweigepflicht unter Strafschutz. Alle Personen beim KI-Anbieter mit Zugang zu Patientendaten müssen auf § 203 StGB verpflichtet werden.

§ 393 SGB V
Seit 1. Juli 2025: Verarbeitung von GKV-Gesundheitsdaten über Cloud-Dienste nur mit aktuellem C5-Typ-2-Testat der datenverarbeitenden Stelle.

TI (Telematikinfrastruktur)
Das geschlossene digitale Netz des deutschen Gesundheitswesens, betrieben von der gematik. Über die TI laufen u. a. eRezept, eAU, ePA und KIM.

Token
Kleinste Verarbeitungseinheit eines Sprachmodells; ungefähr ein Wortstück oder Satzzeichen. Ein typischer Arztbrief umfasst rund 2.000 Token Eingabe und 800 Token Ausgabe.

TOM (Technische und organisatorische Maßnahmen)
Maßnahmen nach Art. 32 DSGVO zum Schutz personenbezogener Daten (z. B. Verschlüsselung, Zugangskontrollen, Protokollierung).

Triage (im KI-Kontext)
Einschätzung der Dringlichkeit eines Patientenanliegens. Wenn eine KI dies für einen konkreten Patienten tut, kann sie als Medizinprodukt eingestuft werden (MDR).

UV-GOÄ
Gebührenordnung für die Abrechnung ärztlicher Leistungen bei Arbeitsunfällen und Berufskrankheiten (Berufsgenossenschaft).

VSDM (Versichertenstammdatenmanagement)
Abgleich der Versichertendaten auf der eGK mit der Krankenkasse beim Einlesen am Kartenterminal.

VVT (Verzeichnis der Verarbeitungstätigkeiten)
Dokumentationspflicht nach Art. 30 DSGVO; jede Praxis muss darin alle Verarbeitungen personenbezogener Daten auflisten.

Quellenverzeichnis

Die folgenden Quellen werden in der Präsentation zitiert oder referenziert:

Chen Z et al. MEDITRON-70B: Scaling Medical Pretraining for Large Language Models. arXiv:2311.16079, 2023.

Charité Berlin, Pilot 2025: Erprobung KI-gestützter Arztbriefe (unveröffentlicht).

Google Research. MedGemma 1.5 Technical Report. arXiv:2604.05081, 2026.

Holmgren AJ et al. Association Between Use of an Ambient AI Scribe and Patient Throughput. JAMA Netw Open 2026.

Jin D et al. What Disease does this Patient Have? A Large-Scale Open Domain Question Answering Dataset from Medical Exams. Appl Sci 2021 (MedQA, n = 1.273).

Kassenärztliche Bundesvereinigung. Bürokratieindex für die vertragsärztliche Versorgung. Datenstand 2020. gesundheitsdaten.kbv.de.

Lukac PJ et al. AI-Generated Clinical Documentation: A Randomized Trial. NEJM AI 2025.

NOHARM Benchmark. Stanford/Harvard 2026 (Preprint). 31 Systeme an echten Konsilfragen.

Rotenstein LS et al. Association of AI Ambient Documentation with Physician Time Allocation. JAMA 2026.

Sellergren A et al. MedGemma Technical Report. arXiv:2507.05201, 2025.

Sinsky C et al. Allocation of Physician Time in Ambulatory Practice: A Time and Motion Study in 4 Specialties. Ann Intern Med 2016;165:753–760.

Song H et al. Hallucinated Procedures in AI-Generated Discharge Summaries. JAMA Netw Open 2025.

Vishwanath A et al. Clinical Decision Support AI: Benchmarking Medical Knowledge Systems. Nat Med 2026.

Verordnung (EU) 2024/1689 (KI-Verordnung). Europäisches Parlament und Rat.

Verordnung (EU) 2017/745 (Medizinprodukteverordnung, MDR). Europäisches Parlament und Rat.

Williams CYK et al. Quality of AI-Generated Clinical Letters. JAMA Intern Med 2025.

Wu C et al. Accuracy of AI-Generated Citations in Medical Literature. Nat Commun 2025.

Hawaii Pacific Health. Prospective Evaluation of AI-Assisted Documentation. 2025 (n = 79).

Stanford Health Care. Patient Perspectives on AI-Assisted Documentation. 2025 (n = 2.202).

GKV-Spitzenverband. Versichertenbefragung 2024. 90 Prozent, Ausgabe 40.

BSI: Cloud Computing Compliance Criteria Catalogue (C5:2020).

BVMed: Infoblatt C5, Stand Januar 2026.

Anthropic. Börsenprospekt (S-1) bei der SEC, Oktober 2026 (nach Berichten von Wall Street Journal und Reuters).

Modulübersicht der Fortbildungsreihe

Modul 1 (22. Juli 2026): Grundlagen, Regulierung und Verantwortung

Modul 2 (16. September 2026): Prompting, Tools und Arbeitsabläufe

Modul 3 (7. Oktober 2026): Datenschutz und Sicherheit

Modul 4 (4. November 2026): Technologie, Automatisierung, Wirtschaftlichkeit und Zukunft

Modul 5 (25. November 2026): KI-gestützte Dokumentation und MD-Prüfung

Alle 5 Module: KI-Kompetenz nach Art. 4 KI-VO und 10 CME-Punkte (LÄK BW).

Dieses Handout fasst die Inhalte der Präsentation zusammen, erklärt die verwendeten Fachbegriffe und ordnet die zitierten Studien ein. Es ersetzt keine Rechtsberatung. Stand: November 2026.