

WHITEPAPER

Generative KI verantwortungsvoll einsetzen

Impulse für Unternehmen und Industrie

Detelf Houdeau, Matthias Peissner et al.

Inhalt

| | |
|---|-----------|
| Zusammenfassung | 3 |
| 1. Einleitung | 4 |
| 2. Generative KI im Einsatz | 6 |
| 2.1 Die Frage nach der Rolle: Anwender oder Anbieter?..... | 6 |
| 2.2 Prozessoptimierung – Zwischen Effizienzgewinn und Entlastung..... | 9 |
| 2.3 Fokus: Neue Potenziale durch Small Language Models (SLMs)..... | 9 |
| 3. Neue Chancen und Anforderungen für IT-Sicherheit, Leitlinien und Compliance | 12 |
| 3.1 Neue Chancen für die IT-Sicherheit in Unternehmen..... | 12 |
| 3.2 Leitlinien für den sicheren und verantwortungs- vollen Einsatz..... | 13 |
| 3.3 Compliance und rechtliche Rahmenbedingungen..... | 14 |
| 4. Handlungsempfehlungen für den Einsatz von generativer KI in Unternehmen | 16 |
| 5. Mit SWOT-Analysen den Einsatz von generativer KI ausloten | 18 |
| 5.1 Generative KI im Wissensmanagement..... | 19 |
| 5.2 Generative KI in industriellen Anwendungen und in der Produktion..... | 22 |
| 5.3 Generative KI in der Softwareentwicklung..... | 25 |
| Literatur | 28 |
| Über dieses Whitepaper | 30 |

Empfohlene Zitierweise

Houdeau, Peissner et al. (2026): Generative KI verantwortungsvoll einsetzen – Impulse für Unternehmen und Industrie. Whitepaper aus der Plattform Lernende Systeme.
DOI: https://doi.org/10.48669/pls_2026-2

Zusammenfassung

Generative KI (engl. GenAI) bezeichnet eine Klasse Lernender Systeme, mit der neue Inhalte in Form von Texten, Bildern, Videos, Musik oder Code erstellt werden können. Ihre Grundlage bilden Basismodelle (Foundations Models), unter denen große Sprachmodelle (Large Language Models, kurz: LLMs) mit Anwendungen wie ChatGPT, Le Chat oder Copilot gegenwärtig eine rasante Verbreitung erfahren. Generative KI-Modelle lernen aus umfangreichen Datensätzen, um originelle Ergebnisse zu erzeugen, die die menschliche Kreativität nachahmen. So lässt sich generative KI für unterschiedlichste Anwendungen einsetzen, von der Wissensarbeit über Produktionsprozesse bis hin zu Gesundheitsforschung und Kreativwirtschaft.

Das rasante Tempo, mit dem generative KI in unterschiedliche Anwendungsbereiche Einzug hält, unterstreicht ihren bahnbrechenden Charakter. Unternehmen integrieren in ihre Prozesse LLMs und generative KI vor allem für Effizienzgewinne. Die Sicherheit der Systeme und die Anpassung an Unternehmenswerte stellen dabei eine große Herausforderung dar. Doch wie können deutsche Unternehmen von generativer KI profitieren, und welche Risiken sollten sie im Blick behalten?

Das Whitepaper, das in Zusammenarbeit der drei Arbeitsgruppen „IT-Sicherheit und Privacy“, „Arbeit/Qualifikation, Mensch-Maschine-Interaktion“ sowie „Innovation, Geschäftsmodelle und -prozesse“ der Plattform Lernende Systeme entstand, zeigt Potenziale und Herausforderungen des Einsatzes generativer KI in Unternehmen und in der industriellen Produktion auf. Zudem thematisiert es dabei auch neue Anforderungen für die IT-Sicherheit sowie regulatorische Rahmenbedingungen. Mithilfe von SWOT-Analysen zu generativer KI – in industriellen Anwendungen, in der Softwareentwicklung sowie im Wissensmanagement – werden beispielhaft Strategien abgeleitet, wie Potenziale gehoben und Herausforderungen angegangen werden können.

1. Einleitung

Anwendungen für generative KI (engl. GenAI) wie ChatGPT, Le Chat oder Midjourney sind sowohl im privaten Umfeld als auch in Unternehmen zunehmend in Verwendung. Sie sind leicht zugänglich, einfach zu nutzen und liefern in sehr kurzer Zeit Antworten. Mittlerweile sind generative KI-Chat-Assistenten bereits in Büroanwendungen enthalten (wie Microsofts Copilot) oder als KI-Assistenten in Apps und Browsern (wie Monica), die gleich auf mehrere KI-Modelle, beispielsweise DeepSeek, ChatGPT, Claude und Gemini, zurückgreifen. Es sind vor allem jüngere Generationen sowie Menschen im akademischen Umfeld, die generative KI-Anwendungen nutzen (acatech, 2025). Mittels Prompts assistiert generative KI den Menschen dabei, auch ohne tiefgreifendes technisches oder künstlerisches Fachwissen, in kurzer Zeit neue Inhalte zu schaffen. Basierend auf dem Modell verfasst sie Texte, erstellt Videos, erschafft Bilder, komponiert Musik oder generiert Code. Dies eröffnet für eine Vielzahl von Anwendungsbereichen unterschiedliche Einsatzmöglichkeiten.

Die Entwicklung von generativer KI lässt sich auf die Grundlagenforschung im Bereich des maschinellen Lernens und der neuronalen Netze zurückverfolgen. Ursprünglich konzentrierten sich Anwendungen maschinellen Lernens auf eng begrenzte Anwendungen. Fortschritte bei der Rechenleistung und der Entwicklung von Algorithmen und Datenstrukturen haben jedoch die Nutzung von großen und tiefen neuronalen Netzen und ähnlichen Repräsentationsformen in praktischen Anwendungen ermöglicht. Ein weiterer wesentlicher Faktor für das Aufkommen und den Betrieb großer Basismodelle ist zudem der Zugang zu den riesigen Datenmengen von Text, Bild und Video, die im Internet verfügbar sind (Löser et al., 2023).

Generative KI in Unternehmen

Die Implementierung von generativer KI in kleinen und mittleren Unternehmen (KMU) und großen Unternehmen erfordert ein hohes Maß an Schnelligkeit und Anpassungsfähigkeit. Denn die Technologie entwickelt sich zügig weiter – neue Funktionen und Anwendungsbereiche entstehen in hohem Tempo. Gerade in innovationsgetriebenen Branchen wie der Automobilindustrie, dem Maschinenbau, dem Gesundheitswesen, dem Finanz- und Versicherungswesen sowie der Medien- und Kreativwirtschaft können generative KI-Anwendungen Effizienz- und Wettbewerbsvorteile ermöglichen. Unternehmen, die sich frühzeitig mit KI-basierten Lösungen auseinandersetzen, sind daher besser darauf vorbereitet, Potenziale zu erkennen und gezielt zu nutzen. Dies setzt sowohl eine technologische Infrastruktur als auch eine agile Unternehmenskultur voraus, die bereit ist, sich schnell auf neue Entwicklungen einzulassen.

Dieses Papier fasst in Kürze zusammen, welche Potenziale und neuen Anforderungen generative KI mit sich bringt, und spricht insbesondere Unternehmen und Industrien an, die im Feld generativer KI Orientierung suchen. Es gliedert sich in zwei Teile: eine allgemeine Einordnung des Einsatzes generativer KI in Unternehmen (Kap. 2–4) sowie eine SWOT-Analyse zu spezifischen Branchen (Kap. 5).

Die Kapitel 2 und 3 geben einen allgemeinen Überblick zu neuen wirtschaftlichen Potenzialen durch den Einsatz generativer KI-Technologien sowie zu den damit einhergehenden Anforderungen an IT-Sicherheit, Compliance und rechtliche Rahmenbedingungen. Kapitel 4 fasst dies in Handlungsempfehlungen für einen verantwortungsvollen und sicheren Einsatz von generativer KI zusammen.

Der zweite Teil des Papiers zeigt beispielhaft, wie die Potenziale und Herausforderungen beim Einsatz generativer KI in bestimmten Anwendungsbereichen mittels SWOT-Analysen untersucht werden können. Im Fokus stehen die Branchen Softwareentwicklung, industrielle Anwendungen und Produktion sowie Wissensmanagement im Allgemeinen. Kapitel 5 kann als Handreichung in Unternehmen genutzt werden, um für das eigene Geschäftsfeld SWOT-Analysen durchzuführen.¹

1 Es sei angemerkt, dass dieses Papier lediglich die Einsatzmöglichkeiten von generativer KI auslotet. In allen Anwendungsbereichen ist der Einsatz herkömmlicher, deskriptiver KI-Tools ebenfalls meist ausreichend oder zielführend. Untersuchungen zu generativer KI in der Arbeitswelt, in der Medizin sowie der Vergleich zu agentischer KI (KI-Agenten) und ethischen Implikationen finden sich in weiteren Veröffentlichungen der PLS.

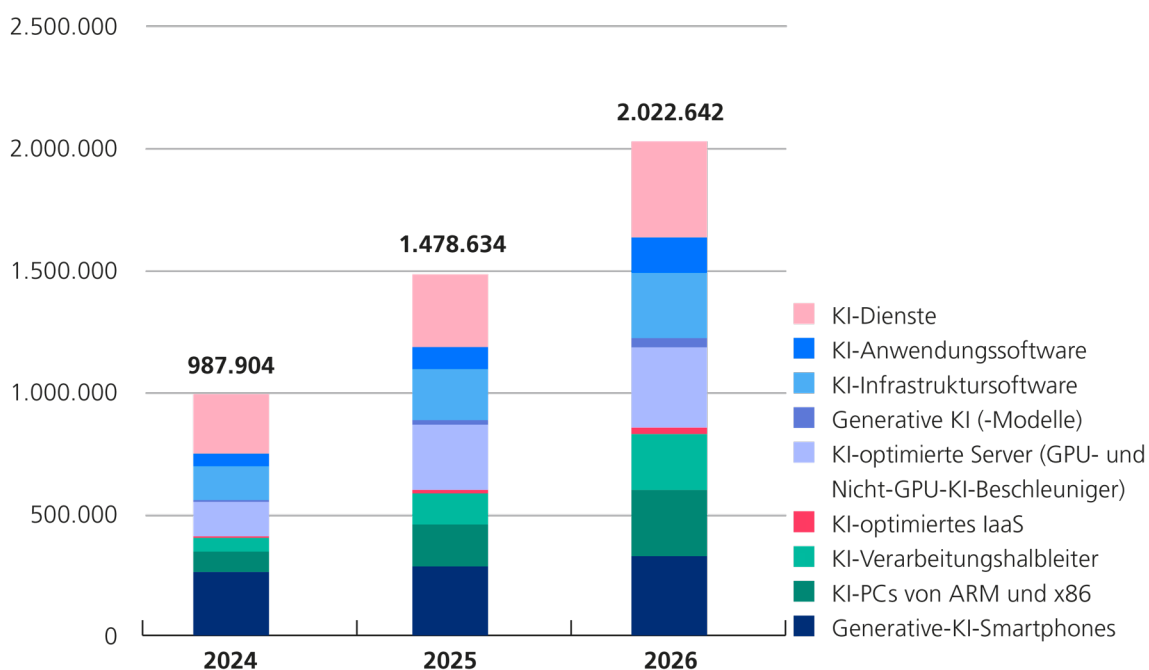
2. Generative KI im Einsatz

2.1 Die Frage nach der Rolle: Anwender oder Anbieter?

Seit 2023 sind die Ausgaben europäischer Unternehmen für generative KI massiv gestiegen, insbesondere deutsche Unternehmen sind neben Frankreich europaweit führend bei Investitionen in generative KI-Technologien (Masood et al., 2023). Die Frage, ob Unternehmen dabei generative KI primär als Anwender nutzen oder selbst als Anbieter entsprechender Lösungen auftreten, ist von zentraler strategischer Bedeutung. Während deutsche Unternehmen in vielen Branchen – insbesondere in der Industrie – über exzellentes Fachwissen, große Datenbestände und langjährige Erfahrung verfügen, bestehen allerdings Defizite bei der Entwicklung eigener Basismodelle, der effizienten Nutzung vorhandener Daten und der Weiterbildung der Belegschaft.² Sollte dieser Rückstand bestehen bleiben, droht Deutschland vor allem im Hinblick auf die industrielle Nutzung generativer KI dauerhaft in die Rolle des bloßen Anwenders zu geraten.



Abbildung 1: Ausgaben für KI auf den IT-Märkten weltweit, 2024–2026
(in Millionen US-Dollar)

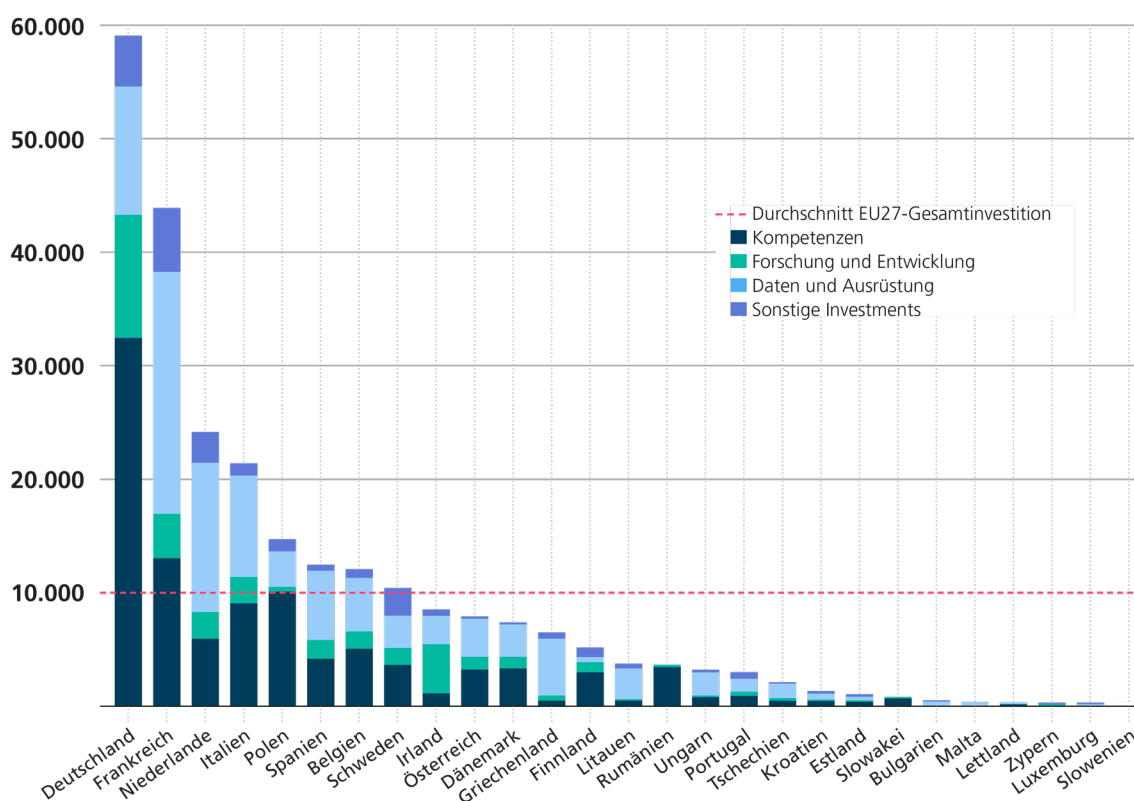


Quelle: Gartner (09/2025)

² Grundsätzlich ist Deutschland im EU-Vergleich führend beim Investment in KI. Der Anteil an Ausgaben für Schulungen macht allerdings im Vergleich zu Ausgaben für F&E sowie zur Vergütung von IKT-SpezialistInnen nur einen kleinen Teil aus (Fonteneau et al., 2025, S. 29).



Abbildung 2: KI-Investitionen in den EU-Mitgliedstaaten nach Kategorie
in Millionen Euro



Anmerkung: Basisszenario, 2023. Sonstige Investments bezieht sich hier auf Design-, Marken- und Organisationskapital.

Quelle: Berechnungen der OECD auf der Grundlage der Volkswirtschaftlichen Gesamtrechnungen von Eurostat 2025, S.29
(https://www.oecd.org/content/dam/oecd/en/publications/reports/2025/09/advancing-the-measurement-of-investments-in-artificial-intelligence_7f58ff65/13e0da2f-en.pdf)

Eine Anbieterrolle eröffnet dagegen nicht nur die Möglichkeit, Wertschöpfungsketten stärker zu kontrollieren, sondern auch neue Geschäftsmodelle zu etablieren. Gleichzeitig können Unternehmen neue Geschäftsmodelle entwickeln, indem sie generative KI anderer Anbieter nutzen – entweder indem sie selbst erstmals als Anbieter solcher Leistungen auftreten oder indem sie ihren Kunden auf dieser Basis zusätzliche, bislang noch nicht mögliche Geschäftsmodelle eröffnen. Diese können etwa in Form spezialisierter, branchenspezifischer KI-Dienste, in der Entwicklung kleiner Sprachmodelle (Small Language Models, kurz: SLMs) für datensensible Anwendungen oder in der Bereitstellung von Infrastruktur- und Beratungsleistungen entstehen.

Insbesondere SLMs bieten deutschen und europäischen Unternehmen die Chance, auf vorhandenen Infrastrukturen ressourcenschonende, maßgeschneiderte KI-Lösungen zu entwickeln und gleichzeitig Datenschutzanforderungen zu erfüllen.

Ob ein Unternehmen selbst Ressourcen hat, eigene KI-Modelle zu entwickeln – und damit zugleich als Anbieter auftritt-, oder lediglich Anwender ist, wird stark durch Unternehmensgröße, Ressourcen und interne Strukturen bestimmt. **Start-ups**, die sich ausschließlich auf generative KI fokussieren, sind häufig beides zugleich: Sie entwickeln eigene Modelle oder spezialisierte Anwendungen und setzen diese unmittelbar bei sich und ihren Kunden ein. Dadurch können sie Innovationen schnell in den Markt bringen, haben aber zugleich ein hohes Risiko bei der Finanzierung, Skalierung und Regulierung. **Mittelständische Unternehmen** treten hingegen meist ausschließlich als Anwender auf. Sie stehen vor besonderen Herausforderungen: Der Einkauf von Technologie ist oft mit hohen Lizenzkosten verbunden, während gleichzeitig die Anpassung an spezifische Unternehmensprozesse, die Sicherstellung von Datensouveränität und die Erfüllung regulatorischer Anforderungen erhebliche Investitionen in Know-how und Infrastruktur erfordern. **Großunternehmen** wiederum verfügen in der Regel über die Ressourcen, generative KI sowohl von globalen Anbietern einzukaufen als auch intern auf die eigenen Bedürfnisse hin weiterzuentwickeln. Sie können eigene Teams für Datenaufbereitung, Modellanpassung und Governance aufbauen und so eine hybride Rolle zwischen Anwender und Anbieter einnehmen (BSI, 2025, S. 13, und Turan, Straßer & Zürn, 2024, S. 52).

KI-Agenten vs. generative KI

KI-Agenten sind weitgehend autonom agierende KI-Systeme, die nach Zielvorgabe eigenständig Aufgaben ausführen. Sie können Aufgaben planen, bewerten, sich anpassen und eigenständig Schlussfolgerungen ableiten. KI-Agenten sind demnach in der Lage, Entscheidungen zu treffen und Aktionen an neue Informationen anzupassen: Sie interagieren mit ihrer digital vernetzten Umgebung (Menschen, anderen KI-Agenten/-Systemen, Cloud). Diese Interaktion unterscheidet KI-Agenten von reinen Chatbots, die auf Konversations-KI basieren, und von herkömmlichen generativen KI-Anwendungen. Grundlage für KI-Agenten sind meist maschinelles Lernen und Basismodelle. Die darauf basierenden Anwendungen können Text, Sprache, Bilder oder Codes verarbeiten (Jäger & Jensen, 2025; Neuburger, 2025; Gartner, 2024).

Eine weitere Lösung bieten neue Geschäftsmodelle, die entlang der gesamten Wertschöpfungskette entstehen können: vom „AI-as-a-Service“-Ansatz mit cloudbasierten Modellen über On-Premise-Lösungen für hochsensible Anwendungen bis hin zu Modellen, die die Vorteile beider Welten kombinieren. Für viele Unternehmen wird es zudem strategisch entscheidend sein, nicht nur KI-gestützte Produkte und Prozesse anzubieten, sondern auch datengetriebene Services, Plattformmodelle und Ökosysteme zu entwickeln.

Die Implikationen sind deutlich: Unternehmen, die sich auf die reine Anwenderrolle beschränken, laufen Gefahr, in starke Abhängigkeiten von wenigen globalen Anbietern zu geraten, sofern keine Auswahlmöglichkeiten bestehen – mit möglichen Nachteilen bei Kosten, Datensouveränität und Innovationsgeschwindigkeit. Eine aktive Anbieter- oder Hybridrolle kann dagegen zur digitalen Souveränität beitragen, regionale Innovationscluster stärken und Wettbewerbsvorteile sichern. Entscheidend ist, frühzeitig strategische Entscheidungen zu treffen, Partnerschaften aufzubauen und Kompetenzen in den Bereichen KI-Entwicklung, Datenmanagement und Geschäftsmodellinnovation gezielt auszubauen.

2.2 Prozessoptimierung – Zwischen Effizienzgewinn und Entlastung

Viele Unternehmen führen generative KI-Anwendungen in ihren Unternehmen für Effizienzgewinne und zur Entlastung der Mitarbeitenden ein. Nach aktuellen Studien wird der Einsatz generativer KI in Unternehmen zunehmend als zentraler Faktor für Effizienz, Wettbewerbsfähigkeit und Innovation genannt (KPMG, 2025). Dies gelingt durch die Prozessoptimierungen, die mit generativer KI möglich sind. So können Abläufe in Unternehmen effizienter, kostengünstiger und flexibler gestaltet werden. Generative KI kann helfen, bisher manuell durchgeführte Schritte zu reduzieren, indem sie Ergebnisvorschläge macht. In der Wissensarbeit werden Routineaufgaben übernommen, wie etwa das Beantworten von E-Mails in natürlicher Sprache, das Protokollieren und Übersetzen bis hin zur Erstellung von Projektplänen basierend auf historischen Daten und Projektzielen. Durch die rasante Verarbeitung, auch heterogener Daten, können durch den Einsatz von generativer KI neue Erkenntnisse gewonnen oder neue Impulse gesetzt werden (Damyanov, Tsankov und Nedyalkov, 2024). Dabei besteht jedoch stets das Risiko von Halluzinationen, sodass vermeintlich neue Erkenntnisse grundsätzlich kritisch geprüft und mit verlässlichen Quellen verifiziert werden müssen.

Im Global GenAI Report (NTT Data, 2024, S. 46) wird davon ausgegangen, dass sich der Fokus in den kommenden Jahren stärker auf die Anwendung von generativer KI zur Optimierung von Back-Office- und Middle-Office-Workflows verlagern wird. Für die Anwendungen in Back-Office- und Middle-Office-Prozessen sind generative-KI-gesteuerte Chatbots bereits die beliebteste Anwendung, gefolgt von Tools zur Workflow-Optimierung und Aufgabenmanagement-Plattformen. Die Bedeutung der Anwendungen variiert je nach Branche: Das Bankwesen und der Einzelhandel setzen verstärkt auf generative KI-Chatbots, während die industrielle Fertigung generativer KI zur Optimierung von Arbeitsabläufen einsetzt.

Die Effizienzgewinne, die der Einsatz generativer KI verspricht, sind jedoch nicht mit einer automatischen Entlastung der Mitarbeitenden gleichzusetzen. Vielmehr verändert generative KI die Extensität und Intensität von Arbeit – Prozesse werden beschleunigt, Aufgaben verdichtet, neue Anforderungen entstehen. Für einige Beschäftigte können Routinen, die sie gut erledigen, auch entlastend sein, und diese Beschäftigten könnten sich durch die neuen Prozesse schnell überfordert fühlen. Damit generative KI nicht zu einer Verschärfung von Belastungen führt, sondern tatsächlich zur qualitativen Verbesserung von Arbeit beiträgt, braucht es klare arbeitsgestalterische Leitlinien hinsichtlich Transparenz, Qualifizierung und Partizipation. Die Arbeitsgruppe „Arbeit/Qualifikation, Mensch-Maschine-Interaktion“ der Plattform Lernende Systeme erörtert dieses Thema im kürzlich erschienenen Whitepaper „KI in Unternehmen: Perspektiven auf den Kulturwandel“ (Neuburger et al., 2025).

2.3 Fokus: Neue Potenziale durch Small Language Models (SLMs)

Große Sprachmodelle (Large Language Models, kurz: LLMs) sind KI-Systeme mit Milliarden von Parametern, die mit ihrem statistischen Funktionswissen typischerweise allgemeines Wissen abdecken und vielseitig einsetzbar sind, etwa für Texterzeugung, Übersetzung oder Analyse. Ihr Einsatz – und insbesondere ihr Training – erfordert hohe Rechenressourcen und sie sind weniger effizient bei spezifischen, anpassungsintensiven Aufgaben. Kleine Sprachmodelle (Small Language Models, kurz: SLMs) sind dagegen deutlich kompakter, benötigen weniger Ressourcen, lassen

sich schneller anpassen und eignen sich besonders für mobile Geräte oder spezialisierte Lösungen, in denen Effizienz und Datenschutz wichtig sind. Der Hauptunterschied ist die Modellgröße: LLMs verfügen oft über Hunderte Milliarden bis zu einer Billion Parameter, SLMs meist zwischen einigen Millionen und wenigen Milliarden. Während LLMs breite Aufgaben meistern, sind SLMs für domänenspezifische, effiziente Anwendungen optimiert. SLMs können lokal ausgeführt werden, ohne dass eine umfangreiche, für LLMs geeignete Recheninfrastruktur eingerichtet werden muss. Sie sind unter anderem für mobile oder Edge-Geräte ideal. Führende Anbieter sind Microsoft (Phi-3), OpenAI (GPT-4o Mini), Google (MobileBERT, Gemma-2), Meta (Llama 3), Alibaba (Qwen 2.5), IBM (Granite) und Mistral AI (Nemo).

Tabelle 1: Charakteristika von LLMs und SLMs im Vergleich

| | Large Language Models (LLMs) | Small Language Models (SLMs) |
|-------------------------------------|---|--|
| Anwendungsfälle | <ul style="list-style-type: none"> ▪ Komplexe, allgemeine Aufgaben ▪ Tiefes Verständnis ▪ Groß angelegte Anwendungen, bspw. in Forschung und Entwicklung³ | <ul style="list-style-type: none"> ▪ Spezialisierte, domänenspezifische Aufgaben, wie Sprachübersetzung oder textbasierte Datenanalyse, und in Fahrzeugsystemen (z. B. intelligente Navigation) |
| Anwendungsbeispiel Gesundheitswesen | <ul style="list-style-type: none"> ▪ Analyse riesiger Datensätze für medizinische Forschung | <ul style="list-style-type: none"> ▪ Einsatz in datenschutzkonformen Diagnosetools für Patienten |
| Ausführung | <ul style="list-style-type: none"> ▪ Benötigt Internetverbindung ▪ Läuft meist in der Cloud ▪ Erzeugt entsprechende Latenz | <ul style="list-style-type: none"> ▪ Kann lokal ohne Internetverbindung mit sehr niedriger Latenz ausgeführt werden |
| Einsatz | <ul style="list-style-type: none"> ▪ Erfordert leistungsstarke Infrastruktur | <ul style="list-style-type: none"> ▪ Geräte mit begrenzten Ressourcen ▪ Mobilgeräte ▪ Edge Computing |
| Größe und Modellkomplexität | <ul style="list-style-type: none"> ▪ Bis zu Hunderte Milliarden Parameter ▪ Höhere Latenz | <ul style="list-style-type: none"> ▪ I.d.R. bis zu 10 Milliarden Parameter ▪ Niedrigere Latenz |
| Kontextverständnis & Fähigkeiten | <ul style="list-style-type: none"> ▪ Vielseitig: kann für viele Aufgaben angepasst und verbessert werden | <ul style="list-style-type: none"> ▪ Eingeschränktes Allgemeinwissen: Sehr gut in ihrem spezifischen Anwendungsbereich |
| Ressourcenbedarf | <ul style="list-style-type: none"> ▪ Sehr hoch (GPU, Cluster nötig) ▪ Kostspielig und ressourcenintensiv | <ul style="list-style-type: none"> ▪ Gering (läuft lokal) ▪ Nachhaltiger: Kann mit deutlich geringeren Ressourcen trainiert werden |
| Antwortgeschwindigkeit | <ul style="list-style-type: none"> ▪ Langsamer | <ul style="list-style-type: none"> ▪ Sehr schnell |
| Kosten | <ul style="list-style-type: none"> ▪ Hohe Trainings-/Betriebskosten | <ul style="list-style-type: none"> ▪ Niedrige Kosten |
| Kontrolle | <ul style="list-style-type: none"> ▪ Typischerweise größere Abhängigkeit vom Anbieter des Modells | <ul style="list-style-type: none"> ▪ Unabhängig entwickelbar ▪ Mehr Kontrolle über Daten und Anpassungen |

Eigene Darstellung (basierend auf Duricic, 2024; Lu et al., 2025; Caballar, 2025)

³ Vgl. bspw. Krähnke, U., Pehl, T. & Dresing, T. (2025): Hybride Interpretation textbasierter Daten mit dialogisch integrierten LLMs: Zur Nutzung generativer KI in der qualitativen Forschung. <https://nbn-resolving.org/urn:nbn:de:0168-ssoar-99389-7>

Für deutsche und europäische Unternehmen können SLMs ein Weg sein, um ihre Geschäftsmodelle mit generativer KI datenschutzkonform und energie- und kostenschonend zu erweitern. Denn SLMs sind schlanker, benötigen weniger Rechenleistung und können auf vorhandener Infrastruktur betrieben werden. SLMs sind ideal, um mit lokal verfügbaren, branchenspezifischen oder sensiblen Daten trainiert zu werden, die nicht in außereuropäische LLMs eingespeist werden sollen oder dürfen. In Deutschland und der EU sind viele dieser Daten dezentral in Unternehmen, Behörden und Forschungseinrichtungen angesiedelt. SLMs lassen sich nicht nur als Cloud-Lösungen, sondern auch vor Ort trainieren und betreiben.⁴ So können Unternehmen die Datenhoheit behalten und zudem gezielt steuern, wie die Modelle geschult und aktualisiert werden. Dies ist nicht nur für die Wahrung von Geschäftsgeheimnissen vorteilhaft, sondern erleichtert auch die Einhaltung regulatorischer Anforderungen aus der Datenschutzgrundverordnung (DSGVO) und dem AI Act.

Anwendungsnahe SLMs lassen sich zudem rasch für spezifische Aufgaben trainieren (z. B. juristische Textanalyse, technische Dokumentation, Branchenkommunikation). Dadurch entstehen passgenaue KI-Lösungen, die direkt in bestehende Prozesse eingebettet werden können – ohne monatelanges „Finetuning“ großer Modelle. Der Fokus auf SLMs eröffnet Chancen für regionale KI-Hubs, die Entwicklung, Hosting und Anwendung bündeln. Das belebt nicht nur ländliche Räume, sondern schafft auch resilientere, dezentrale KI-Infrastrukturen im Einklang mit europäischen Werten. Darüber hinaus könnte der Aufbau eines europäischen Ökosystems für SLMs die digitale Unabhängigkeit vom US-dominierten LLM-Markt stärken. Mittelständische Unternehmen, Forschungseinrichtungen und Start-ups könnten eigenständig Innovation betreiben – ohne auf die Infrastruktur von Tech-Giganten angewiesen zu sein.

Teuken 7B

Teuken 7B ist ein großes Open-Source-KI-Sprachmodell, entwickelt im [Forschungsprojekt OpenGPT-X](#) (Januar 2022 bis März 2025) des Bundeswirtschaftsministeriums unter [Leitung des Fraunhofer IAIS](#) (mit Fraunhofer IIS). Mit rund sieben Milliarden Parametern besitzt es aber SLM-Größe und ist aufgrund eines eigens entwickelten Tokenizers auch bei mehrsprachigen Anwendungen besonders effizient. Die Teuken 7B-v0.6-Modelle wurden mit 6 Billionen Tokens vortrainiert. Teuken 7B wurde von Grund auf in allen 24 EU-Sprachen trainiert, wobei etwa die Hälfte der Trainingsdaten nicht aus dem Englischen stammt. Das Modell kann von Forschenden und Unternehmen aller Branchen frei genutzt, angepasst und in proprietäre Anwendungen integriert werden und eignet sich besonders für mehrsprachige, transparente und datenschutzkonforme KI-Anwendungen ([Teuken 7B-Grundprinzipien](#)). Das KI-Sprachmodell steht auf Hugging Face für die lokale Weiternutzung zum [Download/Teuken 7B](#) bereit.

Im [Video](#) erklärt Dr. Mehdi Ali vom Lamarr-Institut, wie sein Team Teuken 7B aufgebaut hat.

⁴ Das Open-Source-Tool Ollama ermöglicht es bspw., große Sprachmodelle lokal auszuführen und so SLMs zu trainieren.

3. Neue Chancen und Anforderungen für IT-Sicherheit, Leitlinien und Compliance

Der Einsatz von generativer KI verändert bestehende Sicherheits-, Governance- und Compliance-Strukturen in Unternehmen grundlegend. Während sich neue Potenziale für die IT-Sicherheit eröffnen, entstehen zugleich Risiken, denen mit klaren Leitlinien und technischen Schutzmaßnahmen begegnet werden muss.

3.1 Neue Chancen für die IT-Sicherheit in Unternehmen

In Unternehmen gewinnt generative KI zunehmend auch für die IT-Sicherheit an Bedeutung. Im Unterschied zu klassischen IT-Sicherheitsansätzen, die primär auf regelbasierten Verfahren, Signaturerkennung und statischen Kontrollmechanismen beruhen, ermöglichen generative Modelle eine kontextbezogene Analyse großer und heterogener Datenmengen und stellen ein weiteres Werkzeug zur Unterstützung sicherheitsrelevanter Prozesse dar (BSI, 2023; Boutemour et al., 2025). Das Bundesamt für Sicherheit in der Informationstechnik (BSI) ordnet generative KI als leistungsfähige, zugleich aber sicherheitskritische Technologie ein. Ihr Einsatz eröffnet neue Möglichkeiten zur Stärkung der IT-Sicherheit, bringt jedoch auch zusätzliche Risiken mit sich, da generative KI sowohl als Angriffsvektor als auch als Instrument der Sicherheitsanalyse fungieren kann (BSI, 2025).

Mittelfristig bietet generative KI potenziell Mehrwerte bei der Identifikation und Minderung von Sicherheitsrisiken. Durch die Analyse von Logdaten, Netzwerkereignissen oder Schwachstelleninformationen könnten perspektivisch Muster und Anomalien erkannt werden, die mit traditionellen Verfahren nur mühsam oder eingeschränkt identifizierbar sind – hierzu sind jedoch noch erhebliche Forschungs- und Entwicklungsanstrengungen notwendig (Tabassi, 2023; Eckert et al., 2025). Dabei könnte generative KI nicht nur klassische Bedrohungsszenarien adressieren, sondern in einigen Szenarien auch Risiken erfassen, die aus dem Einsatz von KI selbst resultieren, etwa durch Prompt-Manipulationen, unkontrollierte Datenabflüsse oder fehlerhafte Modellausgaben mit potenziellen gravierenden Folgen.

Bereits heute bietet generative KI Möglichkeiten für die realitätsnahe Simulation von Angriffsszenarien, beispielsweise im Rahmen von „Red-Teaming“-Übungen oder strukturiertem Threat Modeling. Generative Modelle können Angriffsstrategien variieren und gezielt an spezifische Unternehmenskontexte anpassen, ohne dabei produktive Systeme zu beeinträchtigen (Lella et al., 2023). Solche Simulationen unterstützen die systematische Identifikation von Schwachstellen und tragen zu einer fundierten Priorisierung von Schutzmaßnahmen bei.

Ein weiteres zentrales Einsatzfeld liegt in der Automatisierung operativer Sicherheitsaufgaben. Generative KI kann die Analyse und Korrelation von Sicherheitsereignissen sowie die Priorisierung von Warnmeldungen unterstützen und so zur Reduktion von Alarmmüdigkeit beitragen (IBM, 2025). Auch im Bereich von Compliance- und Kontrollprüfungen eröffnen sich Potenziale, etwa bei der Auswertung interner Richtlinien, technischer Dokumentationen oder Auditnachweisen (BSI, 2023). Dadurch werden IT-Sicherheitsverantwortliche entlastet und erhalten mehr Handlungsspielraum für strategische und Governance-bezogene Aufgaben.

Neben operativen Prozessen bietet generative KI zudem Chancen im Wissensmanagement und in der Sicherheitskommunikation. Dazu zählen etwa die Erstellung zielgruppenspezifischer Awareness-Materialien, die Unterstützung von Schulungs- und Compliance-Kampagnen oder der Einsatz KI-basierter Sicherheitsassistenten, beispielsweise in Form von IT-Sicherheits-Chatbots (siehe bspw. [↗ Sofie von Sosafe](#), ein KI-gestützter Sicherheits-Copilot).

Gleichzeitig verdeutlicht insbesondere die BSI-Perspektive, dass diese Potenziale nur dann realisiert werden können, wenn generative KI klar geregelt, kontrolliert und regulatorisch sauber in bestehende Sicherheitsstrukturen eingebettet wird (BSI, 2025). Damit bilden die dargestellten Chancen die Grundlage für die nachfolgende Betrachtung von Leitlinien und Compliance-Anforderungen.

3.2 Leitlinien für den sicheren und verantwortungsvollen Einsatz

Mit dem zunehmenden Einsatz von generativer KI in Unternehmen entstehen neue, teils schwerwiegende Risiken für IT-Sicherheit, Datenschutz und Governance. Diese Risiken sind nicht ausschließlich technologischer Natur, sondern resultieren häufig aus fehlenden Leitplanken, unzureichender Transparenz oder mangelnder Sensibilisierung im Umgang mit KI-Systemen.

Vor diesem Hintergrund gewinnt die Entwicklung unternehmensinterner Leitlinien für den Einsatz von generativer KI zunehmend an Bedeutung. Ziel solcher Leitlinien ist es, den sicheren, verantwortungsvollen und unternehmenskonformen Einsatz generativer KI sicherzustellen. Sie adressieren sowohl IT-sicherheitsbezogene als auch ethische und Compliance-relevante Fragestellungen. Dazu gehören beispielsweise Regelungen zur sicheren Verarbeitung von Daten, zum Schutz vor Manipulationen oder zur Vermeidung der Preisgabe sensibler Informationen. Ebenso umfassen Leitlinien Anforderungen an den Output der Anwendungen, etwa im Hinblick auf die Vermeidung von Halluzinationen, Verzerrungen oder Abweichungen von Corporate Identity.

In vielen Unternehmen fehlen bislang geeignete Sicherheitsarchitekturen zur systematischen Einbindung generativer KI. Daher ist es erforderlich, Bedrohungsszenarien frühzeitig und strukturiert zu analysieren. Ohne Maßnahmen wie die Validierung von Eingaben, Red Teaming⁵ oder den Einsatz geprüfter und sicherer Trainingsdaten entstehen potenzielle Einfallstore für Manipulationen und Missbrauch. Ein konsequenter „Security-by-Design“-Ansatz, wie er in anderen IT-Bereichen bereits etabliert ist, sollte daher auch auf generative KI-Systeme angewendet werden (BSI, 2025).

Darüber hinaus benötigen Unternehmen klare Nutzungsrichtlinien, technische Zugriffsbeschränkungen sowie praxisnahe Schulungen für Mitarbeitende. Diese Maßnahmen sind insbesondere deshalb relevant, weil generative KI derzeit teilweise ohne Wissen der IT-Abteilungen eingesetzt wird – häufig auch auf privaten Endgeräten. Dieser sogenannte Schatteneinsatz führt zu einem erheblichen Kontrollverlust über Unternehmens- oder Forschungsdaten. Werden vertrauliche

⁵ In der IT-Sicherheit meint „Red Teaming“ eine Methode, bei der ein Team von SicherheitsexpertInnen einen simulierten, realistischen Angriff auf die KI- oder IT-Systeme bzw. auf die physische Infrastruktur einer Organisation durchführt.

Informationen unbedacht in öffentlich zugängliche KI-Tools eingegeben, werden diese dauerhaft an externe Anbieter übertragen und können nicht mehr zurückgeholt werden. Betroffen sind sowohl personenbezogene Daten als auch geschäftskritisches Wissen.

3.3 Compliance und rechtliche Rahmenbedingungen

Neben internen Leitlinien stellen rechtliche und regulatorische Anforderungen einen zentralen Rahmen für den Einsatz von generativer KI dar. Öffentlich zugängliche KI-Anwendungen übertragen Daten häufig automatisiert über Landesgrenzen hinweg oder erfüllen nicht in allen Fällen datenschutzrechtliche Mindeststandards. Daraus ergeben sich erhebliche Risiken für unbeabsichtigte Verstöße gegen europäische Verordnungen oder branchenspezifische Regulierungen.

Seit August 2024 ist der AI Act als zentrale europäische Verordnung zur Regulierung von KI-Systemen in Kraft und wurde ab August 2025 schrittweise verbindlich umgesetzt. Unternehmen sind verpflichtet, ihre KI-Anwendungen anhand definierter Risikoklassen einzuordnen – von minimalem bis hin zu unannehmbarem Risiko – und je nach Einstufung bestimmte Anforderungen zu erfüllen. Dazu zählen unter anderem Vorgaben zu Transparenz, Dokumentation und menschlicher Aufsicht. Für Basismodelle und generative KI gelten zusätzliche Pflichten, insbesondere wenn sie als sogenannte General-Purpose AI (GPAI) klassifiziert werden.⁶ Anbieter und Betreiber müssen in diesem Fall unter anderem technische Dokumentationen bereitstellen, Informationen zu Trainingsdaten offenlegen und Urheberrechte beachten (Plattform Lernende Systeme, 2024).

Ergänzend zum AI Act greifen weiterhin die DSGVO sowie nationale Regelungen wie das Bundesdatenschutzgesetz (BDSG).⁷ Unternehmen müssen sicherstellen, dass Mitarbeitende im Umgang mit KI-Systemen ausreichend geschult sind und keine sensiblen Daten unbedacht eingegeben werden. Der Einsatz nicht genehmigter KI-Tools („Schatten-KI“) wird in diesem Zusammenhang zunehmend kritisch diskutiert. Die rechtliche Verantwortung liegt dabei nicht allein bei der IT-Abteilung, sondern ausdrücklich auch bei der Geschäftsleitung. Alle eingesetzten KI-Anwendungen müssen geprüft werden. Zudem muss die Auswahl der Plattformen konsequent an rechtlichen Vorgaben ausgerichtet werden. Juristische Schulungen des Personals im Themenfeld „KI & Recht“ sind daher unverzichtbar.

Darüber hinaus besteht die Gefahr, dass fehlerhafte, voreingenommene oder halluzinierte KI-Antworten unkritisch als Entscheidungsgrundlage genutzt werden. Dem muss durch eine Kultur der kritischen Reflexion begegnet werden: KI soll unterstützen, darf aber menschliche Verantwortung nicht ersetzen. Verantwortlichkeiten für den Einsatz und die Qualitätsprüfung von KI-Ergebnissen müssen klar definiert und entsprechende Kompetenzen aufgebaut werden.

6 Der AI Act (KI-VO) regelt die Anforderungen an Hochrisiko-KI-Systeme insbesondere in den Artikeln 8–15. Darüber hinaus hilft der Verhaltenskodex für Allzweck-KI (GPAI) Unternehmen und Industrie dabei, die gesetzlichen Verpflichtungen des KI-Gesetzes in Bezug auf Sicherheit, Transparenz und Urheberrecht von Allzweck-KI-Modellen einzuhalten. <https://digital-strategy.ec.europa.eu/en/policies/contents-code-gpai>

7 Die Studie „KI-Verordnung, NIS-2-Richtlinie und Cyber Resilience Act: Auswirkungen auf KMU“ untersucht, wie die EU-Regelwerke KI-Verordnung, NIS-2-Richtlinie und Cyber Resilience Act den digitalen Wandel kleiner und mittlerer Unternehmen prägen, welche Chancen und Herausforderungen daraus entstehen und welcher Unterstützungsbedarf sich ableiten lässt. [KI-Verordnung, NIS-2-Richtlinie und Cyber Resilience Act: Auswirkungen auf KMU](#).

Grundsätzlich bedarf die Einführung von KI-Tools einer KI-förderlichen Unternehmenskultur, damit die Einführung gelingen kann, aber auch damit entsprechende (kritische) Kompetenzen aufgebaut werden können. Gleichzeitig agiert KI auch als Treiber einer sich ständig wandelnden Unternehmenskultur, da sich beispielsweise Interaktionen und Führungskultur verändern (Neuburger et al., 2025). Diese Veränderung äußert sich etwa darin, dass Entscheidungsprozesse demokratisiert werden und schneller ablaufen, eine kollaborative, lernende Haltung zwischen Mensch und Maschine gefördert wird und dass sich eine wertebasierte KI-Nutzung auch in der Führungskultur wiederfindet (Stowasser & Neuburger et al., 2022).

4. Handlungsempfehlungen für den Einsatz von generativer KI in Unternehmen

Produktivitätssteigerung und Kostensenkung sind wichtige Treiber zur Implementierung von generativer KI in Unternehmen und für industrielle Anwendungen. Dabei ist nicht nur Schnelligkeit gefragt, sondern auch die Ausgestaltung einer KI-assistierten Arbeitskultur, die sich an Unternehmenswerte anpasst, und die Adaption an neue Anforderungen an IT-Sicherheit und Rechtskonformität. Die Autorinnen und Autoren resümieren folgende Handlungsempfehlungen für einen sicheren und verantwortungsvollen Einsatz von generativer KI:

1. Verantwortungsbewusste KI-Kultur etablieren: KI ist kein Ersatz für menschliches Urteilsvermögen. Es braucht eine Kultur in den Unternehmen, in der kritische Prüfung, Transparenz und Verantwortlichkeit die Regel sind – nicht das blinde Vertrauen in maschinelle Ergebnisse. Die partizipative Einführung von generativen KI-Systemen, bei der Beschäftigte frühzeitig eingebunden werden, kann dabei hilfreich sein (Plattform Lernende Systeme, 2025).

2. Transparente KI-Nutzung sicherstellen: Inoffizielle oder nicht genehmigte Tools bergen erhebliche Risiken. Durch die Bereitstellung zugelassener Anwendungen und gezieltes Monitoring schaffen Unternehmen eine sichere und kontrollierte KI-Nutzung. Zudem ist eine transparente Aufgabenverteilung zwischen Mensch und Maschine gefragt, um Überforderung oder Entfremdung zu vermeiden.

3. Schnelligkeit und Dynamik adressieren: Organisationen sollten die Dynamik rund um generative KI nutzen, indem sie Pilotprojekte starten, Erfahrungen systematisch auswerten und erfolgreiche Ansätze skalieren. Dafür braucht es klare Verantwortlichkeiten, sichere Testumgebungen und Leitlinien, wann und wie KI-Anwendungen in den Regelbetrieb überführt werden. Ziel ist ein balanciertes Vorgehen, um Innovationschancen zu nutzen und zugleich Qualität, Sicherheit und Compliance sicherzustellen.

4. Kostensenkung und Produktivitätssteigerung ermöglichen: Generative KI kann sowohl Produktivitätssteigerungen als auch Kostensenkungen ermöglichen – entscheidend ist, beide Ziele in ein ausgewogenes Verhältnis zu bringen. Wo generative KI Prozesse beschleunigt oder Freigaben verkürzt, braucht es jedoch angepasste Prüf- und Compliance-Mechanismen, damit Sicherheit, Sorgfalt und Transparenz nicht unter Effizienzdruck geraten. Verantwortungsvolle Produktivitätsstrategien machen daher sichtbar, wie Kosten- und Leistungsgewinne gemeinsam erreicht und nachhaltig genutzt werden.

5. Auswirkungen auf die Arbeit prüfen und gestalten: Es muss verhindert werden, dass der Einsatz von KI zu ungewolltem Wissens- und Kompetenzverlust und zu neuen Risiken und Belastungen (z.B. Überforderung, durch Intensivierung und Extensivierung) führt. Auch muss das Verhältnis zwischen Produktivitätsgewinnen und Qualität beobachtet werden. So sollten beispielsweise Produktivitätsgewinne nicht auf Kosten notwendiger Überprüfungen der Ergebnisse generativer KI erzielt werden.

6. Sicherheit als Gestaltungsprinzip verankern: Der gesamte Lebenszyklus von KI-Systemen – von der Datenaufbereitung über Modellauswahl und Training bis hin zum Betrieb – muss durchdacht abgesichert werden. Security-by-Design, kontinuierliche Bedrohungsanalysen und qualitativ hochwertige Trainingsdaten sind hierbei unerlässlich. Neuartige Ansätze zur nachweisbaren Erhöhung der Sicherheit von KI-Systemen beispielsweise in den Bereichen Nachvollziehbarkeit/Erklärbarkeit, Datenverdichtung und -generierung, Leitplanken und Verifikation von Modellverhalten sollten bei Verfügbarkeit schnellstmöglich adaptiert werden. Wichtig ist dabei, auch die klassischen technischen und organisatorischen Maßnahmen der IT-Sicherheit, wie beispielsweise regelmäßige Schulungen der Mitarbeitenden, automatisierte Patch-Verwaltung⁸ oder Mehr-Faktor-Authentifizierung (MFA), umzusetzen.

7. Datenschutz und Datensicherheit von Anfang an mitdenken: Verantwortlichkeiten bei der Nutzung von Fremdsoftware und fremder Daten, die mittels generativer KI verarbeitet werden, sollten klar definiert sein. Sensible oder personenbezogene Informationen dürfen nicht unkontrolliert in KI-Systeme gelangen. Organisationen müssen klare Richtlinien für die Datennutzung erstellen und ihre Mitarbeitenden befähigen, sicher mit KI-Anwendungen umzugehen. Auch die Datenspeicherung ist mit Sicherheitsrisiken verbunden, insbesondere bei der Nutzung öffentlicher Cloud-Dienste. Diesen Risiken kann mit einer guten Data-Governance-Strategie und zusätzlichen Sicherheitsmaßnahmen begegnet werden.

8. KI-Governance etablieren: KI-Governance bezieht sich auf den gesamten Lebenszyklus einer KI-Lösung – das heißt Entwicklung, Einführung und Betrieb – und umfasst Leitlinien, Prozesse, Organisationsstrukturen und technische Lösungen. Damit soll sichergestellt werden, dass die eingesetzte KI-Lösung die erwarteten Ziele erfüllt und entsprechende Ergebnisse liefert, sicher und überprüfbar ist und die mit der Lösung verbundenen Risiken möglichst minimiert bzw. sogar vermeidet.

9. Rechtliche Rahmenbedingungen prüfen und konsequent einhalten: Der Einsatz von KI muss jederzeit mit den geltenden Compliance-Vorgaben vereinbar sein. Nur geprüfte und Compliance-konforme generative KI-Lösungen sollten eingesetzt werden. Des Weiteren kann Unsicherheiten im Zusammenhang mit der Anwendung des AI Act durch entsprechende Schulungen entgegengewirkt werden, auch im Sinne einer konsequenten Einbindung der Beschäftigten.

⁸ Identifizierung, Beschaffung, Prüfung und Installation von Software-Updates (Patches)

5. Mit SWOT-Analysen den Einsatz von generativer KI ausloten

Die im Folgenden beschriebene SWOT-Analyse wurde von den Autorinnen und Autoren in zwei Workshops entwickelt – mit dem Ziel, beispielhaft damit auszuloten, wie sich der Einsatz von generativer KI in einem bestimmten Anwendungs- oder Geschäftsbereich gestalten lässt.⁹

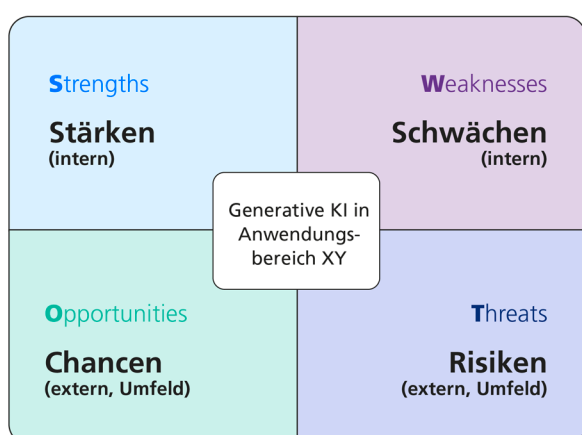
Zunächst wurden mittels Brainstorming Branchen festgehalten, in denen der Einsatz von generativer KI große Effizienzgewinne und Entlastung für Mitarbeitende verspricht. In industriellen Anwendungen und der Produktion sowie in der Softwareentwicklung sehen die Autorinnen und Autoren das größte Potenzial. Darüber hinaus lassen sich über alle Branchen hinweg mittels generativer KI beim Wissensmanagement in den Unternehmen große Potenziale ausschöpfen. Zu diesen drei Anwendungsbereichen wurden mittels SWOT-Analysen Impulse abgeleitet, die Unternehmen berücksichtigen sollten, um resiliente generative KI-Anwendungen zu implementieren. In einem ersten Schritt wurden interne Stärken und Schwächen des Einsatzes von generativer KI in den jeweiligen Anwendungsbereichen sowie externe Chancen und Bedrohungen aus dem Umfeld des Anwendungsbereichs identifiziert. In einem zweiten Schritt wurden die Ergebnisse der SWOT-Analyse zur Ableitung von Impulsen genutzt.

Die Analyse zeigt: Die SWOT-Analysen offenbaren auch konträre Charakteristika einer Technologie. Werden diese sinnvoll kombiniert, lassen sich nützliche Impulse oder Strategien für den Umgang mit generativer KI ableiten. Ein Vergleich der abgeleiteten Impulse mit den drei folgenden Fallbeispielen zeigt, dass sich in unterschiedlichen Geschäftsbereichen oft ähnliche Strategien zur erfolgreichen Implementierung von generativer KI eignen.



Abbildung 3: SWOT-Analyse: Vorgehen in 2 Schritten

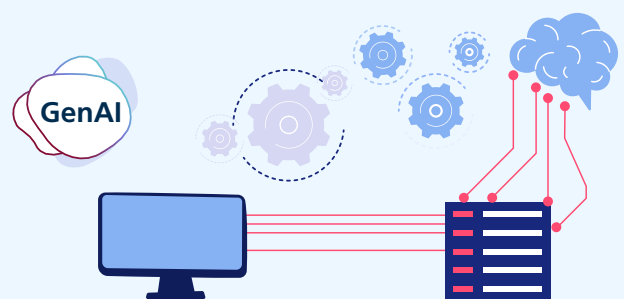
Schritt 1: SWOT identifizieren



Schritt 2: Strategien zum Einsatz generativer KI



⁹ Methodische Notiz: Mittels SWOT-Analysen werden in komprimierter Form Vor- und Nachteile einer Technologie festgehalten. Dieses Vorgehen offenbart auch Widersprüchlichkeiten. Ziel ist es, aus der Kombination von Stärken und Chancen sowie Schwächen und Risiken Strategien abzuleiten, mit denen bspw. bei der Implementierung einer Technologie den Risiken entgegengewirkt werden kann.



5.1 Generative KI im Wissensmanagement

In der Wissensarbeit wie Marketing, Personalwesen, Kundenbetreuung, Buchhaltung, IT-Support, Rechtsabteilungen, Forschung & Entwicklung – generative KI kann branchenübergreifend in Back-offices etwa für die Dokumentenverwaltung, zur Erstellung von Unterlagen, in Workflow-Optimierungstools bis hin zu Prognosen oder Strategiebildung eingesetzt werden. Generative KI bietet neue Möglichkeiten zur Sicherung und Bündelung von Fach- und Erfahrungswissen, insbesondere von alternden und scheidenden Fachkräften. Es speichert Wissen und erweitert Wissensbasen, insbesondere wenn es mit Wissenssilos und großen unternehmensinternen Datenmengen geschickt verknüpft wird. Unter dieser Voraussetzung ermöglicht es schnellere, effizientere Recherchen.

Der Einsatz von generativer KI im Wissensmanagement birgt auch Bedrohungen und Herausforderungen. Besonders bei sensiblem Wissen wie Industriegeheimnissen oder medizinischen Daten können Sicherheits- und Datenschutzprobleme auftreten, etwa durch den unbefugten Zugriff auf internes Wissen oder durch Datenabfluss über öffentliche LLMs. Ein übermäßiger Fokus auf digital erfasstes Wissen kann neue Abhängigkeiten erzeugen und zudem mit Wissens- und Kompetenzverlust einhergehen. Denn generative KI hat Schwierigkeiten, Erfahrungswissen, also praktisch erworbenes und implizites Wissen, vollständig zu erfassen. Weitere potenzielle Schwachstellen sind Halluzinationen und Fehlinformationen, die unerkannt in die Anwendung gelangen und nachfolgend kaum automatisch korrigiert werden können.

| Stärken Strengths | Schwächen Weaknesses |
|--|--|
| <ul style="list-style-type: none"> • Kann Wissensbasen über RAG¹⁰ oder andere Mechanismen in generative KI integrieren, auch mit Kontrolle von Informationsflüssen • Breite Anwendung möglich – auch für andere Anwendungsgebiete (z. B. Pharmaindustrie, Materialforschung) • Verknüpfung von bisherigen Wissenssilos ist möglich • Rasche Verfügbarkeit von Wissen • Einfache Recherche ermöglicht schnelles Lernen/ Wissensaufbau | <ul style="list-style-type: none"> • Halluzinationen können nicht leicht automatisch korrigiert oder entdeckt werden • Wissensrepräsentation abhängig von "Weltbild" und Wertvorstellung des LLM (s. DeepSeek) • Abbau von Kompetenzen durch Nutzung von KI; Halluzinationen und Fehlinformation gelangen unerkannt in die Anwendung • (Zu starker) Fokus auf das digital erfasste Wissen • Silent Truncation (lückenhafte Erfassung von Gelerntem) möglich • Mangelhafte Erfassung von Erfahrungswissen |
| Wissensmanagement | |
| Chancen Opportunities | Bedrohungen Threats |
| <ul style="list-style-type: none"> • Stärkung des Industriestandorts • Erkenntnisgewinn • Implizites Wissen einfangen, insbesondere von erfahrenen Fachkräften • Bewahrung des Wissens (Speicher für scheidende ExpertInnen/Erfahrungsträger) • Erweiterung der Wissensbasis • Mehr Wissen durch einfaches Recherchieren • Entlastung von unproduktiver Tätigkeit, z. B. Suche | <ul style="list-style-type: none"> • Falsche Informationen • Sicherheits- und Datenschutzprobleme bei sensiblem Wissen (Industriegeheimnisse, Medizin) • Gefahr des externen unbefugten Zugriffs auf internes Wissen • Homogenisierung von Wissen • Reduktion von Innovationspotenzial • Zwänge/Fremdsteuerung von Arbeit • Neue Abhängigkeiten und Wissens-/Kompetenzverlust |

¹⁰ Retrieval-Augmented Generation (RAG) ist ein KI-Verfahren, das externe Daten mit großen Sprachmodellen kombiniert, um genauere und kontextbezogene Antworten zu liefern. Dadurch kann das Modell aktuelleres Wissen und relevante Informationen aus einer Datenbank oder Wissensbasis abrufen und somit das Risiko durch veraltete Informationen oder „Halluzinationen“ verringern. Dies, ohne neu trainiert werden zu müssen. Siehe Glossar der Plattform Lernende Systeme.

Aus der SWOT-Analyse lassen sich folgende Impulse für Unternehmen ableiten, um den Einsatz von generativer KI im **Wissensmanagement** effektiv zu gestalten.

Vertrauenswürdige generative KI-Tools nutzen

- Vertrauenswürdige Anbieter wählen (z. B. auf Datenschutz achten)
- Zugriffsbeschränkte, lokal oder regional gehostete Wissenssysteme aufbauen (z. B. für sensible Informationen)
- Belastbare, nachhaltige Partnerschaften suchen und generative KI-Werkzeuge modular einbinden, um kritische Abhängigkeiten zu reduzieren

Schulungen anbieten & KI-Kompetenzen aufbauen

- Schulungen zur Nutzung von KI-Tools anbieten, um schneller Vorteile nutzen zu können sowie kritisches Denken im Umgang mit KI zu fördern
- Austausch von Wissen zwischen Menschen fördern, um Falschinformationen aufzudecken
- KI-Kompetenzen zum informationssicherheits- und datenschutzbewussten Umgang mit generativer KI aufbauen
- Aufklären über Schwächen und Grenzen von KI/LLMs, Menschen hinsichtlich Falschinformationen (weiter) in die Verantwortung nehmen

Halluzinationen begegnen & Datenschutz adressieren

- Kompetenz zur Identifikation von Halluzinationen und Bias als USP/Standortvorteil/Exportartikel aufbauen
- Kuratierte Wissensdatenbanken aufbauen, um Halluzinationen systematisch zu begrenzen → SLM
- Sensible Daten nicht zum Training breit nutzbarer Machine-Learning-Modelle nutzen, sondern modell-extern einbinden, um Datenschutz zu ermöglichen
- Trainingsdaten anonymisieren, um Datenschutzproblemen vorzubeugen

Use Cases identifizieren

- Breites Abfragen in der Organisation, um werthaltige Use Cases für KI-basiertes Wissensmanagement zu identifizieren
- Anwendungsfälle kategorisieren nach Fokus auf explizitem und dokumentiertem vs. implizitem und undokumentiertem Wissen
- Eigene Strategien zur Extrahierung von implizitem Wissen verfolgen, z. B. mittels ExpertInnen oder datenbasierter Wissensextraktion
- Wissensrepräsentationen nutzen, die unabhängig vom eingesetzten LLM sind (z. B. Dokumentenfeatures in Vektordatenbanken)

Zugang zu generativen KI-Tools für Mitarbeitende regeln

- Geregelter Zugang für die eigenen Mitarbeitenden ermöglichen, um Zugang zu allgemeinem Wissen und allgemeiner KI-Assistenz zu vereinfachen
- Schnelle Recherchemöglichkeiten zur Entlastung produktiver Mitarbeit fördern

Wissensmanagement mit RAG aktiv gestalten

- Verknüpfung von Silos, um Erfahrungswissen alternder Belegschaft aktiv zu sichern
- Fortgeschrittene RAG-Systeme innerhalb der eigenen Organisation in die Anwendung bringen, um Wissenssuche zu beschleunigen
- RAG-gestützte Wissensportale als zukunftssichere Wissensspeicher für projektübergreifende Arbeit einsetzen
- Externe Wissensbasen (RAG) und neu verknüpfte Wissenssilos zur Identifikation von Fehlinformationen durch Abgleich von Datenquellen gezielt nutzen
- Informationsflüsse in RAG-Systemen kontrollieren, damit Informationssicherheit gewährleistet bleibt
- Bedeutung von organisationspezifischen Datenquellen hochhalten und diese weiter pflegen (trotz und wegen Einsatz von LLMs/RAG-Systemen)
- Innovationen stets fördern und schnellen Wissenszugang nutzen, um Lücken im State of the Art schneller zu finden
- Kombination mit symbolischer KI zu neurosymbolischer KI (→ Graph RAG)
- Sicherheitskonzepte direkt in Wissensschnittstellen integrieren (z. B. über Rollen/Policies bei Retrieval)

Qualitätskontrolle etablieren

- Domänenspezifische Modelle oder RAG-Architekturen zur besseren Kontextkontrolle nutzen
- Validierungsmechanismen kontinuierlich und regelmäßig einsetzen
- Kombination aus Mensch-und-Maschine-Verifikation zur Qualitätskontrolle von Wissensausgaben
- RAG-Systeme mit Validierungsschleifen ausstatten, um Halluzinationen von Wissen zu unterscheiden und falsche Informationen zu erkennen

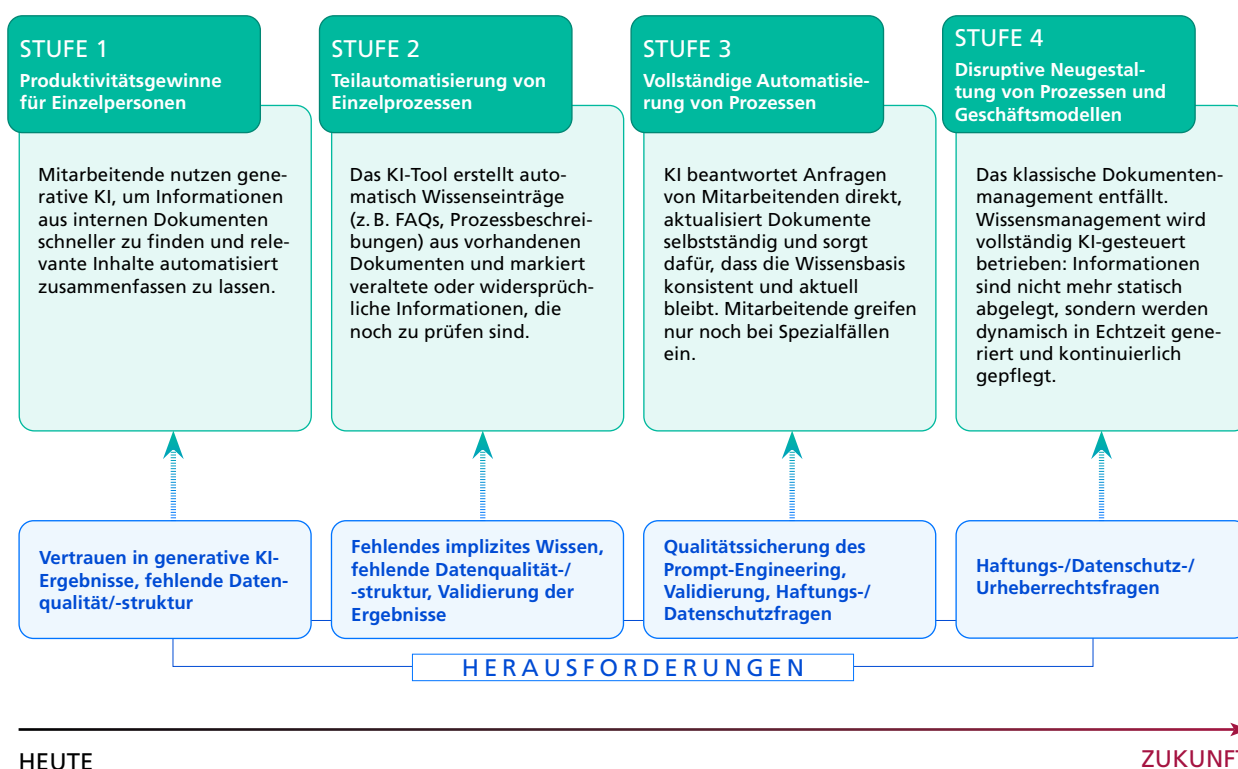


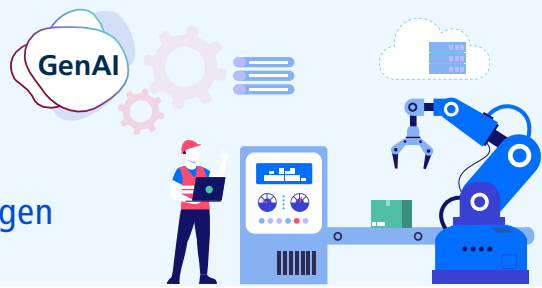
Abbildung 4: Stufen der Prozessautomatisierung mittels generativer KI Beispiel 1: Maintenance / Wissensmanagement

Mittels generativer KI könnten Maintenance-Prozesse in Unternehmen, beispielsweise im Wissensmanagement, zunehmend automatisiert ablaufen. Wie könnte der Einsatz von generativer KI zur Automatisierung der Prozesse beitragen? Das Anwendungsszenario zeigt auf, wie sich die Automatisierung der Prozesse stufenweise vollziehen könnte, und welche Herausforderungen dabei überwunden werden müssten.

Unter anderem noch zu klären:

- Welche wirtschaftlichen Potenziale sind denkbar, wenn der Einsatz von generativer KI in Maintenance-Prozessen beim Wissensmanagement stetig zunimmt?
- Sind Akzeptanz und Vertrauen in Technologie und Mensch-Maschine-Interaktion gegeben?
- Werden andere technische Lösungen diese Entwicklung aufheben oder ersetzen?





5.2 Generative KI in industriellen Anwendungen und in der Produktion

Generative KI bietet vielfältige Potenziale für industrielle Anwendungen und die Produktion, beispielsweise die Unterstützung bei der Entwicklung von CNC-Programmen oder bei der Prozess- und Fertigungsplanung sowie im Requirements Engineering. Die hohe Standardisierung industrieller Prozesse, eine beherrschbare Komplexität sowie große Mengen strukturierter Daten begünstigen den KI-Einsatz. Generative KI kann zudem dazu genutzt werden, um implizites Erfahrungswissen erfahrener Mitarbeitender zu erfassen, aufzubereiten und für andere zugänglich zu machen – etwa in Form von Assistenzsystemen oder erklärenden Entscheidungshilfen (vgl. 5.1).

Demgegenüber stehen neue Risiken und Herausforderungen wie Abhängigkeiten von einzelnen Anbietern, mögliche Produktionsstörungen durch fehlerhafte oder unzureichend validierte Modelle sowie zusätzliche Angriffsflächen für Cyberangriffe. Die bislang begrenzte Verfügbarkeit ausgereifter Basismodelle für sensorische Daten, Aktuatorik oder Bewegungssteuerung schränkt den Einsatz von generativer KI in sicherheitskritischen Bereichen derzeit noch ein. Zudem erschwert das teilweise intransparente Entscheidungsverhalten („Black Box“) vieler Modelle die Nachvollziehbarkeit, etwa bei automatisierten Freigaben in der Qualitätssicherung. Dies kann zu Fehlentscheidungen führen und langfristig Kompetenzen der Mitarbeitenden schwächen.

| Stärken Strengths | Schwächen Weaknesses |
|---|--|
| Chancen Opportunities <ul style="list-style-type: none"> • Standardisierbare Prozesse und beherrschbare Komplexität • Zugriff auf viele (teils strukturierte) Daten • Verkürzte Time-to-Market & Innovationszyklen • Verknüpfung von Industrie-/KI-Kompetenz • KI-basierte Auswertungen großer Datenmengen: schnelle Ergebnisse durch/mit KI • Fehleranalyse, effizientere Fehlerbehebung • Unterstützung bei der Aus-/Weiterbildung • Ergänzung physischer Produkte mit KI-basierten Services • Einfache Verfügbarkeit über generative KI-Oberflächen • Bestehende Anlage "revitalisieren" (Retro-Fitting) | Bedrohungen Threats <ul style="list-style-type: none"> • Fokus nur auf Prozessoptimierungen und nicht auf neue Produktentwicklung oder Features • Verlust der Innovationsfähigkeit, z. B. durch ständige Reproduktion • Verlust geistigen Eigentums bei falscher Anbieterwahl • Neue Sicherheitsrisiken und Angriffspunkte möglich • Abhängigkeit von externen Anbietern • Eingeschränkter Blick auf Qualifizierung sowie Fach- und Erfahrungswissen • Verlust der MA-Expertise bei KI-gesteuerten Prozessen • Black-Box-Verhalten führt zu Wissensverlust |

Industrie/Produktion

Aus den Chancen und Risiken lassen sich folgende Impulse für Unternehmen ableiten, um zügig resiliente generative KI-Anwendungen in der **industriellen Anwendung und Produktion** zu implementieren.

Industrielle Transformation mit KI angehen

- Industrielle Stärken mit KI verknüpfen
- Bestehende Anlagen "revitalisieren" (Retro-Fitting) ohne große Investitionen
- Steigender Bedarf an Sensorik und Datenerfassung als Geschäftsmöglichkeit verstehen

Digitale Souveränität stärken

- „Digital souveräne“ Entscheidungen treffen nach sorgfältiger Analyse der verfügbaren Anbieter von generativer KI
- Einsatz von Modellen großer US-Tech-Konzerne in die Produktion kann zu neuen technologischen Abhängigkeiten führen -> Entwicklung eigener KI-Lösungen oder Nutzung europäischer Alternativen

Innovation und Effizienz zusammendenken

- Fokus auf zielgenaue Anwendungen mit schnellem praktischen Nutzen in der Arbeit setzen
- Neue Geschäftsfelder/-möglichkeiten in den Bereichen Datenbereitstellung und -harmonisierung ausloten
- Unterschied zwischen Effizienzsteigerungen und Innovationspotenzialen strategisch abbilden

Kompetenzen entwickeln & Wissen integrieren

- Generative KI als unterstützendes Werkzeug gestalten, das menschliche Expertise ergänzt und nicht ersetzt
- Co-Evolution: Lösungen zur Rationalisierung und Kompetenzentwicklung bzw. Neu-Einbindung von Fach-/Erfahrungswissen ausloten
- Gezielte Weiterbildung/Kompetenzausbildung in den notwendigen analogen Wissens-/Erfahrungsbereichen anbieten und durchführen
- KI- und Produktionskompetenz gezielt aufbauen, um neue Abhängigkeiten zu vermeiden

IT-Sicherheit verankern

- Cybersicherheitskompetenz zur Schaffung resilienterer KI-Systeme gezielt nutzen, dazu das Personal qualifizieren
- Sicherheitsprinzipien von Anfang an berücksichtigen („Security by Design“), um KI-Systeme von der Konzeption bis zum Betrieb robust gegen Angriffe und Missbrauch zu machen
- Klare Governance-Strukturen schaffen und frühzeitig Haftungs-, Sicherheits- und Verantwortungsfragen klären

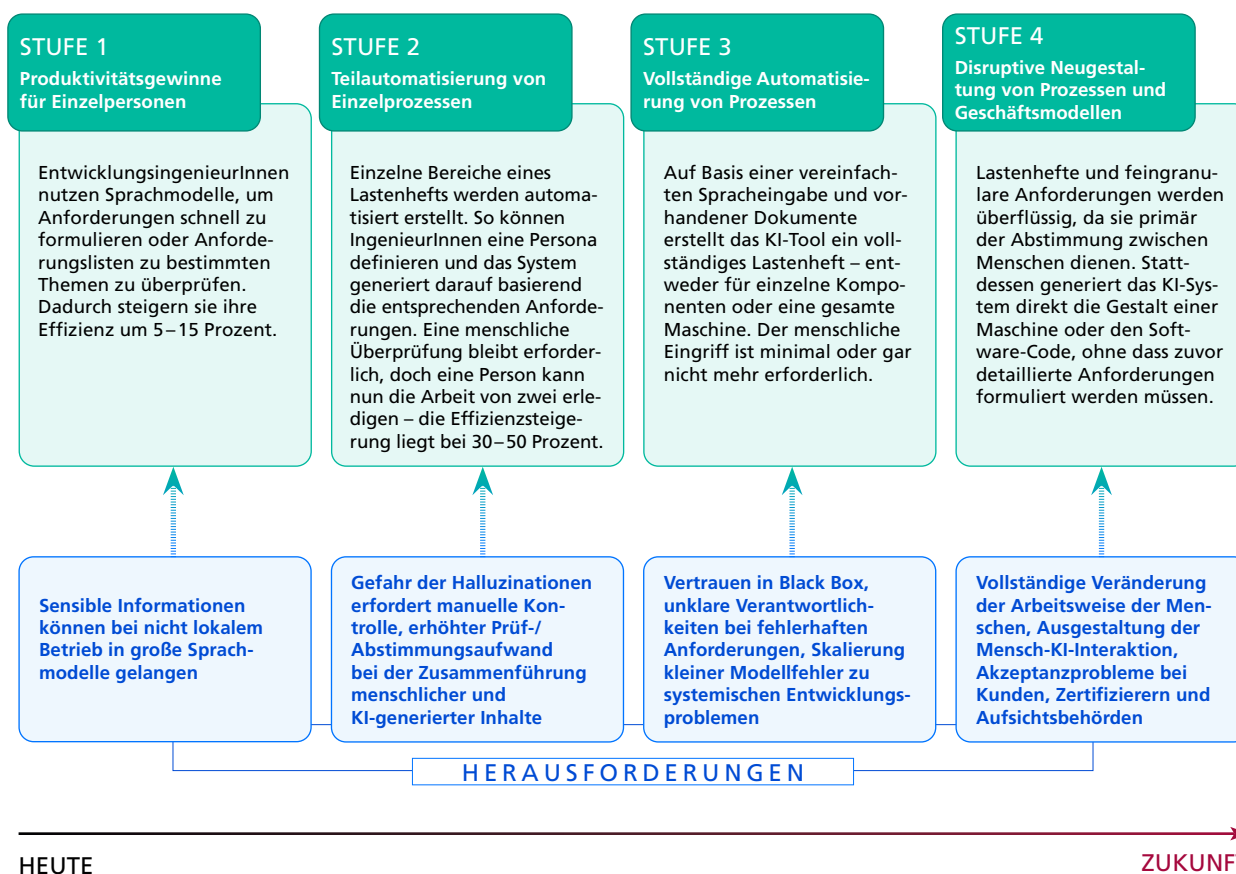


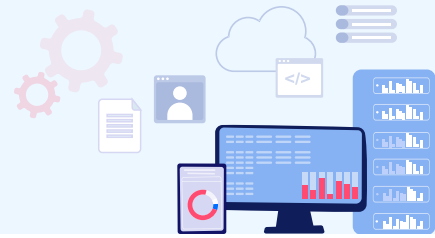
Abbildung 5: Stufen der Prozessautomatisierung mittels generativer KI Beispiel 2: Arbeit eines Ingenieurs im Bereich Requirements Engineering

Mittels generativer KI könnte die Arbeit von IngenieurInnen, beispielsweise im Bereich Requirements Engineering, zunehmend automatisiert ablaufen. Wie könnte der Einsatz von generativer KI zur Automatisierung der Prozesse beitragen? Das Anwendungsszenario zeigt auf, wie sich die Automatisierung der Prozesse stufenweise vollziehen könnte, und welche Herausforderungen dabei überwunden werden müssten.

Unter anderem noch zu klären:

- Welche wirtschaftlichen Potenziale sind denkbar, wenn EntwicklungsingenieurInnen zunehmend generative KI in ihre Arbeit adaptieren?
- Sind Akzeptanz und Vertrauen in Technologie und Mensch-Maschine-Interaktion gegeben?
- Werden andere technische Lösungen diese Entwicklung aufheben oder ersetzen?
- Wie können Verlässlichkeitsgarantien gegeben werden und welcher Grad von Verlässlichkeit kann als akzeptabel hinsichtlich Funktionalität, Sicherheit und Akzeptanz anerkannt werden?





5.3 Generative KI in der Softwareentwicklung

Die SWOT-Analyse zeigt: Der Einsatz generativer KI in der Softwareentwicklung wirkt vor allem als Produktivitäts- und Effizienztreiber. Durch schnellere Code-Erstellung, effizientere Formulierung von Anforderungen sowie automatisierte Dokumentation sind deutliche Produktivitätsgewinne möglich. Routineaufgaben können teilweise übernommen werden, wodurch hochqualifizierte Entwicklerinnen und Entwickler entlastet werden und sich somit stärker auf konzeptionelle und kreative Tätigkeiten konzentrieren können. Dies steigert nicht nur die Effizienz, sondern auch die Attraktivität des Tätigkeitsfelds Softwareentwicklung. Zugleich entstehen Potenziale für neue Lösungsansätze und einen schnelleren Kompetenzaufbau, auch außerhalb klassischer IT-Bereiche. Low-Code- und No-Code-Ansätze können darüber hinaus Kapazitätsengpässe mindern und Innovationsprozesse beschleunigen.

Zur Realisierung der genannten Potenziale ist jedoch eine höhere technologische Reife gegenüber dem heutigen Stand der Technik notwendig; einige Studien weisen derzeit sogar Produktivitätsverluste nach (Becker et al., 2025). Denn generative KI ersetzt kein erfahrenes Personal und kann ineffizienten oder fehlerhaften Code erzeugen, einschließlich sicherheitskritischer Schwachstellen. Zudem besteht die Gefahr eines schleichenden Know-how-Verlusts in Bereichen wie Softwarearchitektur, Best Practices und Qualitätsprinzipien sowie einer Reduktion von Innovationspotenzial durch die Reproduktion bestehender Lösungen. Auf strategischer Ebene drohen neue Cybersecurity-Risiken, Lizenz- und Urheberrechtskonflikte sowie ein Abfluss von Wertschöpfung, insbesondere durch die Abhängigkeit von außereuropäischen generativen KI-Plattformen.

| | |
|---|---|
| <p>Stärken Strengths</p> <ul style="list-style-type: none"> • Effiziente Code-Erstellung • Schnellere Resultate • Übernahme von Routinearbeiten teils von hochbezahlten Personen • Effizienzsteigerung (erlebbar) • Schnelligkeit in Softwareentwicklung erhöht • Mehr Nutzung von Open-Source-Software • Effiziente Formulierung von Anforderungen • Effiziente Erstellung von Dokumentation • Schnellere Einarbeitung in Codebasis und Sprachen | <p>Schwächen Weaknesses</p> <ul style="list-style-type: none"> • Kein Ersatz für erfahrene EntwicklerInnen • Möglichkeit des Erzeugens von ineffizientem Code sowie Code mit Sicherheitslücken • Keine 100-prozentige Verlässlichkeit (stochastische Funktionsweise, Halluzinationen, Missverständnisse) • Aufwendige Fehlerbehebung • Know-how-Verlust (z.B. bzgl. SW-Architekturen, Best Practices, Qualitätsprinzipien) • Reduzierte Qualität und Innovationspotenzial durch reproduzierte Lösungen • Viel Code/Code-Schwemme • Intensivierung von Zeitdruck |
| Softwareentwicklung | |
| <p>Chancen Opportunities</p> <ul style="list-style-type: none"> • Attraktivitätssteigerung des Tätigkeitsfelds • Produktivitätsgewinne • Neuartige Lösungen • Kosteneinsparung und weniger Outsourcing • Kompensation fehlender Kompetenzen und Kapazitäten durch automatische oder KI-unterstützte Entwicklung • Demokratisierung der Software-Entwicklung mit No-Code/ Low-Code • Schnellerer Kompetenzaufbau in nicht IT-affinen Bereichen, z. B. industrielle Produktion | <p>Bedrohungen Threats</p> <ul style="list-style-type: none"> • Schwächung des Softwarestandorts durch Know-how-Verlust • Cybersecurity: neue Angriffsflächen • Abfluss von Wertschöpfung durch fehlende deutsche und europäische Foundation-Modelle: KI-unterstützte Tools werden sehr viel kosten • Hoher Energieverbrauch bei ineffizientem Code • Weniger Regulatorik außerhalb der EU • Hohe Folge-/Lebenszykluskosten durch SW-Qualitätsprobleme • Lizenzkonflikte durch Reproduktion geschützten Codes • Streitigkeiten zu Urheberrecht |

Aus der SWOT-Analyse lassen sich für den Einsatz von generativer KI in der **Softwareentwicklung** für Unternehmen folgende Impulse ableiten.

Generative KI-Tools einführen und nutzen

- Generative KI-Tools einführen, Nutzung erlernen, Chancen realisieren
- Generative KI-Tools erfahrenen EntwicklerInnen an die Hand geben
- EntwicklerInnen verschiedener Erfahrungsstufen (und KI-Agenten) für Know-how-Transfer zusammenarbeiten lassen
- Vorschläge von generativer KI nutzen, um Open-Source-Software (OSS) schneller zu finden und bewusst einzusetzen
- Schnelle und effiziente Code-Erstellung zur agilen Entwicklung neuartiger Cybersecurity-Tools mit kurzen Entwicklungszyklen nutzen
- Experimente und Messungen durchführen
 - Welche Nutzenden werden wie viel schneller?
 - Steigt der Anteil von OSS?
 - Empfinden Nutzende ihr Tätigkeitsfeld attraktiver?

Kompetenzen aufbauen & Weiterbildung fördern

- Effiziente Code-Erstellung und Toolnutzung gezielt für Kompetenzaufbau in KMU und nicht-IT-affinen Abteilungen einsetzen (Citizen Developer)
- Fortbildungsprogramme zur Stärkung architektureller Kompetenzen einführen (Know-how-Verlust vermeiden)
- Entwickler-Assistenzsysteme für Berufseinsteiger zur schrittweisen Kompetenzentwicklung einsetzen
- Hochqualifiziertes Personal in Richtung Cybersecurity und Code-Effizienz weiterbilden, um Bedrohungen zu erkennen und zu beseitigen
- Hochentwickeltes Personal in Richtung KI-Kompetenz weiterbilden, um freie Basismodelle¹² nutzen und weiterentwickeln zu können

Qualitätssicherung in Softwareentwicklung verankern

- Qualitätssicherung durch menschliche EntwicklerInnen allen Stufen der Softwareentwicklung nutzen
- Software-Testing und Codeanalyse mit Qualitätssiegel für LLM-generierten Code durchführen
- Qualitätssicherungstools integrieren, um ineffizienten oder sicherheitskritischen Code frühzeitig zu erkennen
- Neue Modellversionen von Basismodellen mit verbesserten Eigenschaften nutzen
- „Human-in-the-Loop“-Systeme zur Vermeidung von Sicherheitslücken einführen
- Bewusst entscheiden: soll kleines Softwaretool (SLM, mit begrenzter Komplexität und Anspruch) oder umfangreiches Softwaresystem (LLM, mit hoher Komplexität und hohem Kompetenzanspruch) entwickelt werden?
- CI/CD-Pipelines erweitern mit Tools zur Erkennung von Cybersecurity-Problemen, Codequalitäts-Problemen, Lizenzkonflikten
- Kontrollmechanismen gegen übermäßigen Code-Output („Code-Schwemme“) etablieren

Standardisierung, Templates & Governance etablieren

- Standardisierung nutzen, um Produktivitätsgewinne bei gleichzeitiger Qualitätssicherung zu erreichen
- Code-Standards und Templates entwickeln, die Sicherheit und Energieeffizienz systematisch einfordern
- KI-Governance einführen und Geschäftsführung, JuristInnen, IT, AnwenderInnen involvieren
- Toolanbieter so auswählen, dass regulatorische Konformität gegeben und Vertrauen ausreichend ist, z. B. Nutzung europäischer, getesteter und ggf. angepasster Basismodelle (Mistral, Teuken, PhariaAI etc.), Bedeutung von Cybersecurity ernst nehmen, mehr Ressourcen dafür bereitstellen

Demokratisierung & Innovation zusammendenken

- Demokratisierung durch No-/Low-Code mit Open-Source-Tools fördern -> Innovationspotenzial erhöhen
- Effizienzgewinne nutzen, um Ressourcen in Sicherheitsprüfungen zu reinvestieren
- Menschliche Stärken wie Kreativität einbeziehen, um Innovation zu fördern

12 Foundation Models, dt. Basismodelle. Siehe [↗ Glossar der Plattform Lernende Systeme](#)



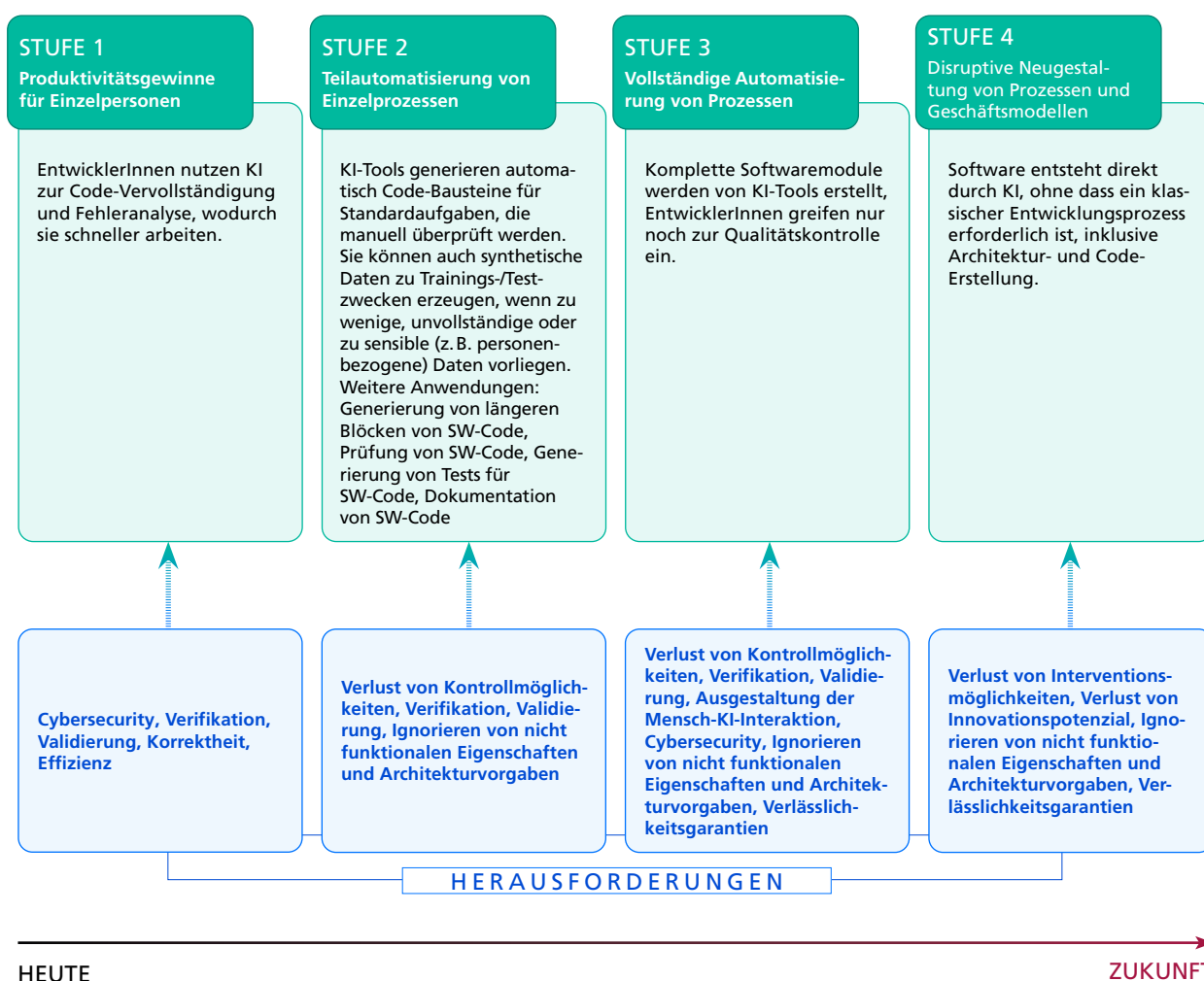
Abbildung 6: Stufen der Prozessautomatisierung mittels generativer KI

Beispiel 3: Generative KI in der Softwareentwicklung

Mittels generativer KI könnten die Prozesse in der Softwareentwicklung zunehmend assistiert ablaufen und bei fortschreitender Reifung der Technologie möglicherweise hohe Autonomisierungsgrade erreichen. Das Anwendungsszenario zeigt auf, wie sich die Automatisierung der Prozesse stufenweise vollziehen könnte, und welche Herausforderungen dabei überwunden werden müssten.¹³

Unter anderem noch zu klären:

- Welche wirtschaftlichen Potenziale sind denkbar, wenn EntwicklungsingenieurInnen zunehmend generative KI in ihre Arbeit adaptieren?
- Sind Akzeptanz und Vertrauen in Technologie und Mensch-Maschine-Interaktion gegeben?
- Werden andere technische Lösungen diese Entwicklung aufheben oder ersetzen?
- Welche Methodenkombinationen und Architekturen eignen sich für verlässliche Ergebnisse?
- Kann ein semantisches Verständnis der Codebasis durch KI-Systeme erreicht werden, um verlässliche und nachvollziehbare Ergebnisse zu erzielen?
- Wie können Verlässlichkeitsgarantien gegeben werden und welcher Grad von Verlässlichkeit kann als akzeptabel hinsichtlich Funktionalität, Sicherheit und Akzeptanz anerkannt werden?



13 Es sei an dieser Stelle darauf hingewiesen, dass es sich um hypothetische Szenarien handelt. „Ob“ und „wie“ KI eingesetzt werden kann, um Menschen bei der Softwareentwicklung zu unterstützen, ist ein aktives Forschungsgebiet mit allen daraus folgenden Implikationen.

Literatur

- acatech (2025):** TechnikRadar 2025. Schwerpunkt Digitale Transformation und KI, München.
 ↗ https://doi.org/10.48669/aca_2025-4
- Avocat Gros, A., Bois, S. & Malosse, A. (2025):** SLM vs LLM: Dans quelle mesure les SLM sont-ils plus efficaces énergétiquement et financièrement que les LLM? Université Côte d'Azur.
 Online unter: ↗ https://rimel-uca.github.io/chapters/2025/SLM_vs_LLM-Team_E/content (letzter Zugriff: 13.11.2025)
- Becker, J., Rush, N., Barnes, B. & Rein, D. (2025):** Measuring the Impact of Early-2025 AI on Experienced Open-Source Developer Productivity. ↗ <https://arxiv.org/pdf/2507.09089>
- Boutemeur, J. et al. (2025):** ENISA Threat Landscape 2025: July 2024 to June 2025, European Union Agency for Cybersecurity (ENISA), Heraklion. Online unter: ↗ https://www.enisa.europa.eu/sites/default/files/2026-01/ENISA%20Threat%20Landscape%202025_v1.2.pdf (abgerufen am 19.02.2026)
- BSI – Bundesamt für Sicherheit in der Informationstechnik (2023):** Die Lage der Sicherheit in Deutschland 2023. Online unter: ↗ https://www.bsi.bund.de/SharedDocs/Downloads/DE/BSI/Publikationen/Lageberichte/Lagebericht2023.pdf?__blob=publicationFile&v=7 (letzter Zugriff: 19.02.2026)
- BSI – Bundesamt für Sicherheit in der Informationstechnik (2025):** Generative KI-Modelle. Chancen und Risiken für Industrie und Behörden. Online unter: ↗ https://www.bsi.bund.de/SharedDocs/Downloads/DE/BSI/KI/Generative_KI-Modelle.pdf?__blob=publicationFile&v=7 (letzter Zugriff: 19.02.2026)
- Caballar, R. D. (2025):** Was sind kleine Sprachmodelle? IBM Think.
 Online unter: ↗ <https://www.ibm.com/de-de/think/topics/small-language-models> (letzter Zugriff: 13.11.2025)
- Damyanov, I., Tsankov, N. & Nedyalkov, I. (2024):** Applications of Generative Artificial Intelligence in the Software Industry, TEM Journal, Vol. 13, Nr. 4, S. 2724-2733. ↗ <https://doi.org/10.18421/TEM134-10>
- Duricic, A. (2024):** Small language models: A beginner's guide. Ataccama.
 Online unter: ↗ <https://www.ataccama.com/blog/small-language-models> (letzter Zugriff: 13.11.2025)
- Eckert et al. (2025):** Generative künstliche Intelligenz und ihre Auswirkungen auf die Cybersicherheit. Impulspapier. Wissenschaftliche Arbeitsgruppe Nationaler Cyber-Sicherheitsrat Juni 2025.
 Online unter: ↗ <https://www.forschung-it-sicherheit-kommunikationssysteme.de/dateien/forschung/2025-06-impulspapier-generative-kuenstliche-intelligenz.pdf> (letzter Zugriff: 19.02.2026)
- Fonteneau, F. et al. (2025):** Advancing the measurement of investments in artificial intelligence, 47. Aufl., OECD Artificial Intelligence Papers. ↗ <https://doi.org/10.1787/13e0da2f-en>
- Gartner (2024):** 4 Ways Generative AI Will Impact CISOs and Their Teams.
 Online unter: ↗ <https://www.tevora.com/wp-content/uploads/2024/02/AI-Gartner-Reprint.pdf> (letzter Zugriff: 08.04.2026)
- IBM (2025):** Cost of a Data Breach Report 2025. Die Lücke bei der KI-Aufsicht. Online unter: ↗ <https://www.ibm.com/reports/data-breach?p1=Display&p2=428388885&p3=227599223> (letzter Zugriff: 19.02.2026)
- Jäger, K. & Jensen, J.-C. (2025):** Next Level Künstliche Intelligenz: Warum KI-Agenten die Spielregeln verändern, Bundesverband Digitale Wirtschaft (BVDW) e.V. Online unter: ↗ <https://www.bvdw.org/wp-content/uploads/2025/06/Definition-KI-Agenten.pdf> (letzter Zugriff: 19.02.2026)
- KPMG (2025):** Generative KI in der deutschen Wirtschaft 2025. Wo stehen deutsche Unternehmen bei der Implementierung von generativer KI? Online unter: ↗ <https://kpmg.com/de/de/themen/digital-transformation/kuenstliche-intelligenz/studie-generative-ki-in-der-deutschen-wirtschaft-2025.html> (letzter Zugriff: 19.02.2026)
- Krähnke, U., Pehl, T. & Dresing, T. (2025):** Hybride Interpretation textbasierter Daten mit dialogisch integrierten LLMs: Zur Nutzung generativer KI in der qualitativen Forschung.
 Online unter: ↗ <https://nbn-resolving.org/urn:nbn:de:0168-ssoar-99389-7> (letzter Zugriff: 19.02.2026)
- Lella, I. et al. (2023):** ENISA Threat Landscape 2023: July 2022 to June 2023, European Union Agency for Cybersecurity (ENISA), Heraklion. ↗ <https://doi.org/10.2824/782573>
- Lu, Z. et al. (2025):** Small language models: Survey, measurements, and insights (arXiv preprint arXiv:2409.15790).
 Online unter: ↗ <https://arxiv.org/pdf/2409.15790> (letzter Zugriff: 13.11.2025)

Löser, A. et al. (2023): Große Sprachmodelle – Grundlagen, Potenziale und Herausforderungen für die Forschung. Whitepaper aus der Plattform Lernende Systeme, München. [↗ https://doi.org/10.48669/pls_2023-3](https://doi.org/10.48669/pls_2023-3)

Masood, S. et al. (2023): Generative AI Radar Europe, Infosys Knowledge Institute, London, Dallas. Online unter: [↗ https://www.infosys.com/services/data-ai-topaz/gen-ai-radar-eu.pdf](https://www.infosys.com/services/data-ai-topaz/gen-ai-radar-eu.pdf) (letzter Zugriff: 19.02.2026)

Neuburger, R. (2025): Künstliche Intelligenz als strategische Führungsaufgabe, in: Badura, B. et al. (Hrsg.), Fehlzeiten-Report 2025, Springer Verlag, S. 283–299. [↗ https://doi.org/10.1007/978-3-662-71885-8_21](https://doi.org/10.1007/978-3-662-71885-8_21)

Neuburger, R. et al. (2025): KI in Unternehmen: Perspektiven auf den Kulturwandel. Impulspapier aus der Plattform Lernende Systeme, München. [↗ https://doi.org/10.48669/pls_2025-3](https://doi.org/10.48669/pls_2025-3)

NTT Data (2024): Global GenAI Report. How organizations are mastering their GenAI destiny in 2025. Online unter: [↗ https://services.global.ntt/en-us/campaigns/global-genai-report](https://services.global.ntt/en-us/campaigns/global-genai-report) (letzter Zugriff: 19.02.2026)

Plattform Lernende Systeme (2024): AI Act der Europäischen Union. Regeln für vertrauenswürdige KI (Publikationsreihe KI Kompakt). Online unter: [↗ https://www.plattform-lernende-systeme.de/files/Downloads/Publikationen/KI_Kompakt/KI_Kompakt_AI_Act_Plattform_Lernende_Systeme_2024.pdf](https://www.plattform-lernende-systeme.de/files/Downloads/Publikationen/KI_Kompakt/KI_Kompakt_AI_Act_Plattform_Lernende_Systeme_2024.pdf) (letzter Zugriff: 07.04.2026)

Plattform Lernende Systeme (2025): Zukunft gestalten! Mit generativer KI. Gesellschaftliche Auswirkungen und Handlungsansätze. [↗ https://doi.org/10.48669/pls_2025-7](https://doi.org/10.48669/pls_2025-7)

Stowasser, S. & Neuburger, R. et al. (2022): Führung im Wandel: Herausforderungen und Chancen durch KI. Whitepaper aus der Plattform Lernende Systeme. [↗ https://doi.org/10.48669/pls_2022-4](https://doi.org/10.48669/pls_2022-4)

Tabassi, E. (2023): Artificial Intelligence Risk Management Framework (AI RMF 1.0), NIST Trustworthy and Responsible AI, National Institute of Standards and Technology, Gaithersburg, MD. [↗ https://doi.org/10.6028/NIST.AI.100-1](https://doi.org/10.6028/NIST.AI.100-1)

Turan, B., Straßer, S. & Zürn, S. (2024): GenAI – Einsatz von künstlicher Intelligenz in der Projektleitung Qualität von Entwicklungsprojekten, Projektmanagement aktuell, Vol. 35, Nr. 5, S. 50–56. [↗ https://doi.org/10.1007/978-3-658-46749-4](https://doi.org/10.1007/978-3-658-46749-4)

Weiterführende Literatur

Abendroth, D. et al. (2025): Generative AI Outlook Report. Exploring the Intersection of Technology, Society and Policy, Publications Office of the European Union, Luxemburg. [↗ https://doi.org/10.2760/1109679](https://doi.org/10.2760/1109679)

Bolz, T. & Schuster, G. (2024): Generative Künstliche Intelligenz im Marketing und Sales. Innovative Unternehmenspraxis: Insights, Strategien und Impulse, Springer Verlag. [↗ https://doi.org/10.1007/978-3-658-45132-5](https://doi.org/10.1007/978-3-658-45132-5)

Eloundou, T. et al. (2023): GPTs are GPTs: An Early Look at the Labor Market Impact Potential of Large Language Models. [↗ https://doi.org/10.48550/arXiv.2303.10130](https://doi.org/10.48550/arXiv.2303.10130)

Engels, B., Lang, T. & Scheufen, M. (2025): KI-Verordnung, NIS-2-Richtlinie und Cyber Resilience Act: Auswirkungen auf KMU, Kurzstudie des Instituts der deutschen Wirtschaft Köln e.V. und der IW Consult GmbH im Rahmen der Begleitforschung des Förderschwerpunkts Mittelstand-Digital, gefördert vom Bundesministerium für Wirtschaft und Energie (BMWE), Berlin, Köln. Online unter: [↗ https://www.mittelstand-digital.de/MD/Redaktion/DE/Publikationen/kurzstudie-ki-verordnung.pdf?__blob=publicationFile&v=11](https://www.mittelstand-digital.de/MD/Redaktion/DE/Publikationen/kurzstudie-ki-verordnung.pdf?__blob=publicationFile&v=11) (letzter Zugriff: 04.03.2026)

Hawighorst, J. (2024): Die Europäische Verordnung über Künstliche Intelligenz, in: Ulbricht, C. et al. (Hrsg.), Praxishandbuch KI und Recht, Haufe Gruppe, S. 155-213, Freiburg. [↗ https://doi.org/10.34157/9783648177037](https://doi.org/10.34157/9783648177037)

Kintz, M. et al. (2024): Potenziale Generativer KI für den Mittelstand. Wie große KI-Modelle die Arbeitswelt verändern, Fraunhofer Institut. [↗ https://doi.org/10.24406/publica-2246](https://doi.org/10.24406/publica-2246)

Ndiaye et al. (2025): Generative AI in Software Engineering: Transforming the Software Development Process. Online unter: [↗ https://www.dfki.de/fileadmin/user_upload/DFKI/Medien/News/2025/Wissenschaftliche_Exzellenz/Generative_AI_in_Software_Engineering_Transforming_the_Software_Development_Process_2025.pdf](https://www.dfki.de/fileadmin/user_upload/DFKI/Medien/News/2025/Wissenschaftliche_Exzellenz/Generative_AI_in_Software_Engineering_Transforming_the_Software_Development_Process_2025.pdf) (letzter Zugriff: 08.04.2026)

Schaller, D. et al. (2023): Künstliche Intelligenz: Chance oder Gefahr? Wie verändert der Einsatz von KI unsere Gesellschaft? ifo Schnelldienst, Vol. 76, Nr. 8. Online unter: [↗ https://hdl.handle.net/10419/279729](https://hdl.handle.net/10419/279729) (letzter Zugriff: 19.02.2026)

TÜV-Verband (2025): Cybersicherheit in deutschen Unternehmen. Neue Bedrohungslage – besserer Schutz. Online unter: [↗ https://www.tuev-verband.de/studien/tuev-cybersecurity-studie-2025](https://www.tuev-verband.de/studien/tuev-cybersecurity-studie-2025) (letzter Zugriff: 04.03.2026)

Über dieses Whitepaper

Das Whitepaper wurde in Zusammenarbeit der drei Arbeitsgruppen „IT-Sicherheit und Privacy, Ethik und Recht“, „Arbeit/Qualifikation, Mensch-Maschine-Interaktion“ sowie „Innovation, Geschäftsmodelle und -prozesse“ der Plattform Lernende Systeme erstellt. Federführend waren Detlef Houdeau für die Unterarbeitsgruppe „IT-Sicherheit und Privacy“ sowie Matthias Peissner für die Arbeitsgruppe „Arbeit/Qualifikation, Mensch-Maschine-Interaktion“.

Hauptautoren

Detlef Houdeau, Infineon Technology (vormals)

Matthias Peissner, Fraunhofer-Institut für Arbeitswirtschaft und Organisation IAO

Autorinnen und Autoren der Arbeitsgruppe „IT-Sicherheit und Privacy“

Prof. Dr. Michael Huth, Technische Universität Nürnberg, Xayn AG

Andrea Martin, IBM Watson Center Munich

Dr. Dirk Wacker, Giesecke+Devrient GmbH

Prof. Dr. Konrad Rieck, Technische Universität Berlin

Autorinnen und Autoren der Arbeitsgruppe „Arbeit/Qualifikation, Mensch-Maschine-Interaktion“

Dr. Andreas Angerer, XITASO GmbH

Klaus Bauer, TRUMPF Werkzeugmaschinen GmbH & Co. KG

Prof. Dr. Andreas Dengel, DFKI GmbH

Prof. Dr. Michael Heister, Bundesinstitut für Berufsbildung (BIBB)

Dr. Norbert Huchler, Institut für Sozialwissenschaftliche Forschung e. V. (ISF München)

Dr. Rahild Neuburger, Ludwig-Maximilians-Universität München

Prof. Dr. Dr. h. c. Christoph M. Schmidt, RWI, Leibniz-Institut für Wirtschaftsforschung

Andrea Stich, Infineon Technologies AG

Autorinnen und Autoren der Arbeitsgruppe „Innovation, Geschäftsmodelle und -prozesse“

Olga Mordvinova, incontext.technology GmbH

Dr. Martin Rabe, Fraunhofer-Institut für Entwurfstechnik Mechatronik IEM

Autoren mit Gaststatus

Dr. Daniel Gille, Cyberagentur

Christoph Maerz, Deutsches Forschungsinstitut für Künstliche Intelligenz

Redaktion

Dr. Erduana Wald, Geschäftsstelle der Plattform Lernende Systeme

Christine Wirth, Geschäftsstelle der Plattform Lernende Systeme

Redaktionelle Mitarbeit

Sami Badr, Geschäftsstelle der Plattform Lernende Systeme

Gefördert durch:



Bundesministerium
für Forschung, Technologie
und Raumfahrt



Impressum

Herausgeber

Lernende Systeme –
Die Plattform für Künstliche Intelligenz
Geschäftsstelle | c/o acatech
Karolinenplatz 4 | 80333 München
➔ www.plattform-lernende-systeme.de

Gestaltung

PRpetuum GmbH, München

Stand

Mai 2026

Bildnachweis

iStock/gorodenkoff/Titel

Empfohlene Zitierweise

Houdeau, Peissner et al. (2026): Generative KI verantwortungsvoll einsetzen – Impulse für Unternehmen und Industrie. Whitepaper aus der Plattform Lernende Systeme.
DOI: ➔ https://doi.org/10.48669/pls_2026-2

Dieses Werk ist urheberrechtlich geschützt. Die dadurch begründeten Rechte, insbesondere die der Übersetzung, des Nachdrucks, der Entnahme von Abbildungen, der Wiedergabe auf fotomechanischem oder ähnlichem Wege und der Speicherung in Datenverarbeitungsanlagen, bleiben – auch bei nur auszugsweiser Verwendung – vorbehalten.

Bei Fragen oder Anmerkungen zu dieser Publikation kontaktieren Sie bitte
Dr. Thomas Schmidt
(Leiter der Geschäftsstelle):
info@plattform-lernende-systeme.de



Über die Plattform Lernende Systeme

Die Plattform Lernende Systeme ist ein Netzwerk von Expertinnen und Experten zum Thema Künstliche Intelligenz (KI). Sie bündelt vorhandenes Fachwissen und fördert als unabhängiger Makler den interdisziplinären Austausch und gesellschaftlichen Dialog. Die knapp 200 Mitglieder aus Wissenschaft, Wirtschaft und Gesellschaft entwickeln in Arbeitsgruppen Positionen zu Chancen und Herausforderungen von KI und benennen Handlungsoptionen für ihre verantwortliche Gestaltung. Damit unterstützen sie den Weg Deutschlands zu einem führenden Anbieter von vertrauenswürdiger KI sowie den Einsatz der Schlüsseltechnologie in Wirtschaft und Gesellschaft. Die Plattform Lernende Systeme wurde 2017 vom Bundesforschungsministerium auf Anregung von acatech – Deutsche Akademie der Technikwissenschaften gegründet und wird von einem Lenkungskreis gesteuert. Die Leitung der Plattform liegt bei Dorothee Bär (Bundesministerin für Forschung, Technologie und Raumfahrt) und Claudia Eckert (Präsidentin acatech).

Weiteres unter [↗ www.plattform-lernende-systeme.de](http://www.plattform-lernende-systeme.de)