

Welche LLM-Anwendungsfälle sind für produzierende Unternehmen besonders interessant?

Studie zum Einsatz von LLMs im Produktionsumfeld

N. Fjodorovs, A. Gütermann

ZUSAMMENFASSUNG Large Language Models (LLMs) haben sich in kurzer Zeit zu einem zentralen Bestandteil der digitalen Transformation entwickelt. Während ihr Einsatz in Kundenservice, Marketing und Vertrieb bereits weit verbreitet ist, stellt sich zunehmend die Frage, welches Potenzial LLMs im Produktionsumfeld bieten. In der vorliegenden Studie wurde anhand von Interviews mit produzierenden Unternehmen untersucht, welche Anwendungsfälle im Produktionsumfeld besonders vielversprechend sind.

STICHWÖRTER

Digitalisierung, Künstliche Intelligenz (KI), Produktionsmanagement

Study on the use of LLMs in the production environment

ABSTRACT Large Language Models (LLMs) have quickly become a central component of digital transformation. While their use is widespread in customer service, marketing and sales, the question arises as to what potential LLMs offer in the production environment. This study makes use of interviews with manufacturing companies to investigate which use cases are particularly promising in the production environment.

1 Einführung

Large Language Models (LLMs) entwickeln sich rasant in der Leistung als auch in der Variation angebotener Modelle [1, 2] und gewinnen in hohem Tempo Millionen von Nutzenden, sowohl im privaten als auch im gewerblichen Bereich [3–5].

Für Unternehmen werden LLMs und generative künstliche Intelligenz (KI) im Allgemeinen zu wichtigen Werkzeugen in der digitalen Transformation. Aktuelle Studien zeigen, dass LLMs von immer mehr Mitarbeitenden und Unternehmen verwendet werden und über ein hohes wirtschaftliches Potenzial verfügen [5, 6]. Besonders verbreitet sind LLMs in den Bereichen Kundenservice, Marketing und Vertrieb. Sie zeigen dort eine sehr große Leistungsfähigkeit [2, 7], da in diesen Bereichen viele und teilweise unstrukturierte Textdateien verarbeitet werden (etwa E-Mails, Berichte, Kundenanfragen). Die zahlreichen erfolgreichen Anwendungsfälle in Dienstleistungsbereichen werfen die Frage auf, welches Potenzial LLMs für die Produktion bergen.

Vor diesem Hintergrund wurde die vorliegende Studie initiiert. Durch Interviews mit Unternehmen der produzierenden Industrie konnte systematisch erfasst werden, welche konkreten Anwendungsfälle für LLMs in der Produktion als vielversprechend erachtet und wie diese von Unternehmen bewertet werden.

2 Aktueller Stand der Technologie

2.1 Fähigkeiten von LLMs

Generative KI ist eine Unterart der künstlichen Intelligenz, die sich auf die autonome Generierung neuer Inhalte oder Informa-

tionen konzentriert. Ihr Ziel ist es, auf Basis der ihr zur Verfügung gestellten Daten neue und einzigartige Inhalte zu erzeugen. Sowohl Input- als auch Output-Format reichen von Text, Bildern, Audio bis hin zu Videos [8].

LLMs sind ein Teilbereich dieser generativen KI, der auf die Verarbeitung und Generierung natürlicher Sprache fokussiert. LLMs basieren auf den Prinzipien des Deep Learnings und setzen künstliche neuronale Netzwerke (KNN) zur Verarbeitung natürlicher Sprache ein. LLMs werden durch selbstüberwachtes Lernen mit enormen Mengen unstrukturierter Textdaten trainiert, umfassen riesige Parametergrößen und haben die transformative Fähigkeit, die Wahrscheinlichkeit von Wortfolgen vorherzusagen oder neuen Text anhand eines gegebenen Inputs zu generieren. Sogenannte N-Gramm-Modelle, der häufigste Typ von LLMs, schätzen Wortwahrscheinlichkeiten auf der Grundlage des vorangehenden Kontextes. Die Fähigkeiten moderner LLMs umfassen ein breites Spektrum: Sie sind in der Lage, Texte mit einem hohen Maß an Kontextverständnis zu generieren, was sich in Anwendungen wie der Erstellung von Texten, dem Verfassen von Code, der Textübersetzung sowie der Zusammenfassung komplexer Inhalte zeigt [9].

Während sich große Anbieter wie OpenAI, Google, Meta, Anthropic oder Mistral auf die Entwicklung leistungsfähiger Sprachmodelle konzentrieren, die sowohl für den privaten als auch den allgemeinen Unternehmenseinsatz verfügbar sind, entstehen vermehrt spezialisierte Lösungen für die Industrie. Unternehmen wie Cohere oder Deepset richten sich gezielt an Geschäftskunden und bieten Werkzeuge zur Integration von LLMs in Geschäftsprozesse an, etwa zur Dokumentensuche, zur

mehrsprachigen Datenverarbeitung oder für domänenspezifische Anwendungen mit Fokus auf Datenhoheit und Anpassbarkeit.

2.2 Limitierungen von LLMs

Bei der Nutzung von LLMs sind bestimmte technologische Einschränkungen zu berücksichtigen. So sind LLMs ausschließlich auf die Verarbeitung von Textdaten ausgelegt und sind nicht in der Lage, Bilder eigenständig zu interpretieren oder zu generieren [10]. Kommerzielle LLMs, die diese Fähigkeiten auf den ersten Blick besitzen, wie etwa ChatGPT oder Gemini, greifen auf externe generative Modelle der Bildverarbeitung (beispielsweise DALL-E oder Imagen) zurück. Die Bildverarbeitung erfolgt dabei nicht durch die LLMs selbst, sondern durch spezialisierte, angebundene Subsysteme.

Da LLMs Inhalte auf Basis statistischer Wahrscheinlichkeiten der nächsten Symbole oder Symbolfolgen (sogenannte Tokens) generieren, sind sie nur bedingt für arithmetische Berechnungen oder die Ausführung logischer Operationen geeignet [11]. Ihre Generalisierungskapazität macht sie weniger prädestiniert für spezialisierte Rechenaufgaben, bei denen derzeit konventionelle Algorithmen oder dedizierte Softwarelösungen eine höhere Präzision und Effizienz bieten [12].

Eine weitere wesentliche Einschränkung betrifft die Zuverlässigkeit der generierten Inhalte. LLMs neigen unter bestimmten Bedingungen zu sogenannten Halluzinationen – also zur Ausgabe inhaltlich falscher oder verzerrter Informationen. Dieser Effekt tritt insbesondere bei domänenspezifischen oder stark regulierten Fragestellungen auf und stellt im industriellen Umfeld ein erhebliches Risiko dar [9, 13]. Vor allem bei streng formalen Ausgabeanforderungen oder konstanten Wenn-dann-Beziehungen kann nicht durchgängig von einer korrekten und konsistenten Modellantwort ausgegangen werden.

2.3 Weitere Entwicklungen

Moderne kommerzielle LLMs besitzen die Fähigkeit der Intermodalität. Dies bedeutet, dass sie auch Bilder, Audio, Videos und verschiedenste weitere Datenformate verarbeiten und kreieren können. Dabei leitet zum Beispiel ChatGPT beim Erkennen von Aufgaben außerhalb des reinen Textspektrums die Anfrage an hauseigene spezialisierte KI wie DALL-E für Bilder oder Whisper für Transkription um. Diese Kombination aus LLMs und anderen KI-Modellen wird als Multi-Agenten-Ansatz bezeichnet [14].

Dadurch eröffnen sich in der Industrie zahlreiche Anwendungsfälle. So unterstützen sie beispielsweise Predictive Maintenance, indem sie durch Datenanalysen Fehler in Produktionsprozessen frühzeitig etwa durch Bildverarbeitung erkennen. Auch im Bereich der Unternehmenssoftware leisten sie zum Beispiel bei der Optimierung von ERP- und SCM-Systemen durch intelligente Datenanalyse oder bei der automatischen Code-Generierung für die Softwareentwicklung gute Dienste. Weiterhin ermöglicht die Fähigkeit der autonomen Interaktion den Einsatz als Chatbots, die als digitaler Assistent oder als Agenten in Softwaresystemen eingesetzt werden können [15].

Die Entwicklungsfortschritte von LLMs über die letzten drei Jahre sind enorm: Eine Studie zeigt, dass sich der Rechenaufwand, um ein bestimmtes Leistungsniveau zu erreichen, etwa alle acht Monate halbiert hat [12], was einerseits auf schnelle algorithmische Verbesserungen und andererseits auf sich immer

neu ergebende Use-Cases hindeutet. Monatlich erscheinen Updates zu den großen LLMs, welche die Grenzen der Nutzung weiter verschieben. So hat OpenAI die Funktion „Tiefenrecherche“ für ChatGPT eingeführt, die es ermöglicht, tausende aktuelle Online-Quellen zu analysieren und zusammenzufassen. Diese Funktion kann in Minuten leisten, wofür Menschen Stunden benötigen würden [16]. Weiterhin wird für die Einführung der neuesten Generation GPT-5 eine multimodale Oberfläche mit Echtzeitzugriff aufs Internet geplant. Google arbeitet an Gemini 1.5, einem Modell, welches das Langzeitgedächtnis des Modells verbessert, um die Interaktion zu personalisieren und autonome Agenten kontextbezogener agieren zu lassen [17]. Auch das chinesische Unternehmen DeepSeek hat mit der Einführung des Modells R1 und anschließend V2 die KI-Welt aufgerüttelt. Trotz deutlich geringerer Investitionen und der Nutzung älterer Hardware erreicht V2 vergleichbare Leistungen wie etablierte Modelle [18]. Dies stellt die bisherigen Annahmen über die notwendigen Ressourcen für leistungsfähige KI-Modelle stark infrage. Die kontinuierlichen Fortschritte machen deutlich, dass sich LLMs weiter verbessern und damit für den Einsatz in hochkomplexen industriellen Anwendungen zunehmend attraktiver werden.

Das Zukunftsbild ist eindeutig: ein LLM als Kernbestandteil eines autonomen Agenten, der Aufgaben selbstständig durchführt. Ein LLM allein ist dabei noch kein Agent, sondern benötigt eine Umgebung und Werkzeuge, um die Autonomie zu gewährleisten. Merkmale eines solchen LLM-Agenten sind zum einen die Wahrnehmung und Verarbeitung von Daten, zum anderen das Treffen einer Entscheidung auf der Basis des Dateninputs und abschließend die Handlung mittels Interaktion mit der Umgebung [1]. Es gibt einige Bemühungen hin zu solchen LLM-Agenten. So nutzt etwa das Open-Source-Projekt „Auto-GPT“ GPT-4 zur autonomen Bewältigung komplexer Aufgaben, wie der Durchführung von Marketing-Analysen oder zur Geschäftsoptimierung [19]. Weiterhin ist „CognitionAI“ von Devin der erste autonome KI-Softwareentwickler, der Code nicht nur generiert, sondern auch testet, Bugs behebt und Projekte selbstständig verwaltet.

3 Existierende Ansätze für Anwendungsfälle in der Literatur

LLMs bieten ein hohes Potenzial für verschiedene wissenschaftliche und berufliche Anwendungsbereiche [20]. In der vorliegenden Studie liegt der Fokus jedoch auf der Produktion und angrenzenden betrieblichen Funktionen, sodass nicht alle in der Literatur verfügbaren Anwendungsfälle genutzt werden können. Die aktuell identifizierbaren Einsatzmöglichkeiten von LLMs im Produktionsumfeld lassen sich in folgende Anwendungskategorien einordnen:

- Kommunikationsbasierte Assistenzsysteme und menschenähnliche Kommunikation: LLMs ermöglichen die intuitive und sprachbasierte Interaktion mit unterschiedlichen Objekten, wie Prozessbeschreibungen, Anweisungen oder sogar Maschinen. Im Bereich der internen Kommunikation helfen LLMs bei der Bearbeitung natürlichsprachlicher Anfragen, dienen als Assistenzlösungen für Mitarbeitende und als Werkzeuge zur Erschließung von Unternehmenswissen [21–25].
- Unterstützung bei der Dokumentation unter Einhaltung geregelter Rahmenwerke: LLMs können bei der Erstellung technischer Dokumentationen unterstützen, insbesondere wenn standardisierte Verfahren oder klare Strukturvorgaben einge-

A	Menschliche Kommunikation	B	Mehrere Textquellen verbinden	C	Textgenerierung	D	Textinterpretation
1.	Beantwortung von Fragen zu einer Anweisung oder einem Prozess (mit der optionalen Aufnahme von Daten)	1.	Erstellung von Testfällen für IT-Systeme	1.	Generierung von ausführlichen Anweisungen für z. B. Montage, Maschinenbedienung, Prüfungsdurchführung, Verpackung, IT-Systembedienung	1.	Vereinfachung und Erklärung langer oder komplexer Anweisungen
2.	Unterstützung bei der Durchführung einer Analyse basierend auf definierten Frameworks (z. B. 8D-Bericht, 5S-Report, FMEA, SWOT usw.)	2.	Erstellung von Lastenheften und Pflichtenheften	2.	Generierung von Produktbeschreibungen und -broschüren	2.	Zusammenfassen von Kundenanforderungen
		3.	Fehler- bzw. Abweichungsanalyse, die textbasierte Informationen berücksichtigt	3.	Generierung von Berichten	3.	Zusammenfassen von Berichten und Protokollen
		4.	(Stamm-)Datenanalyse und -qualitätserhöhung für Freitextfelder	4.	Generierung von Protokollen (z. B. Produkt- oder Prozessänderungen, IT-Funktionen, Maßnahmenarbeit)	4.	Zusammenfassung von Besprechungen (z. B. Stand-up-Meetings, Schichtübergabe)
		5.	Machbarkeitsprüfungen / Abfrage historischer Erfahrungen	5.	Generierung von gesetzlich angeforderten Unterlagen (z. B. Konformitätserklärung, Auditberichte)	5.	Unterstützung bei der Deutung von gesetzlichen Unterlagen (z. B. Gesetze, Zertifizierungen, Normen)
						6.	Analyse von KVP-Vorschlägen seitens Mitarbeitenden

Bild 1 Übersicht der kategorisierten Anwendungsfälle für das Produktionsumfeld. Grafik: eigene Darstellung

halten werden müssen. So können sie beispielsweise bei der Durchführung von FMEA-Analysen helfen oder strukturierte Beschreibungen für Verwaltungsschalen generieren, indem sie Informationen aus mehreren unstrukturierten Quellen zu einem Asset oder einer Maschine zusammenführen [26].

- Gestaltung von Texten und Anweisungen: Sprachmodelle unterstützen bei der Erstellung von Anweisungen oder Prozessbeschreibungen, zum Beispiel bei der Interpretation von Fertigungsvorgaben oder der Ableitung von Maßnahmen aus internen Daten und externen Quellen [25, 27].
- Produktentwicklung und -design: Sprachmodelle unterstützen kreative Prozesse wie die Ideenfindung oder die Beschreibung von Produktkonzepten [15, 28]. Für bildbezogene Aufgaben werden multimodale Modelle eingesetzt, die Text- und Bildverarbeitung kombinieren. Da der Fokus dieser Studie auf textbasierten Anwendungen liegt, wird diese Kategorie nicht weiter vertieft.

4 Methodisches Vorgehen der Studie

4.1 Ableitung und Kategorisierung der Anwendungsfälle

Die Entwicklung konkreter Anwendungsfälle erfolgte durch ein mehrstufiges methodisches Vorgehen. Ausgangspunkt war die systematische Identifikation relevanter textbasierter Datenquellen in produzierenden Unternehmen (wie etwa Arbeitsanweisungen, Materialzeugnisse, Schichtprotokolle, Lastenhefte, Kundenreklamationen und Begleitscheine), basierend auf einer Kombination aus Literaturrecherche und Praxiserfahrung der Autoren. Diese Analyse erlaubte eine erste Einschätzung dazu, welche textuellen Datenbestände prinzipiell für den Einsatz von LLMs geeignet sind.

Zudem wurden bereits publizierte Anwendungsbeispiele (siehe Kapitel 3) hinsichtlich ihrer Übertragbarkeit auf das Produktionsumfeld analysiert und anhand funktionaler Prinzipien kategorisiert. In einem weiteren Schritt wurden die charakteristischen Fähigkeiten aktueller LLMs berücksichtigt (Kapitel 2), um die technische Machbarkeit und Anwendungsrelevanz einzugrenzen.

Aus der Integration dieser drei Perspektiven konnten vier übergeordnete Funktionskategorien mit insgesamt 18 exemplarischen Anwendungsfällen abgeleitet werden (Bild 1):

- Imitierung menschlicher Kommunikation: Automatisierte Korrespondenz und Interaktion
- Semantische und logische Verbindung mehrerer Textquellen: Verknüpfung und Analyse verschiedener Dokumente
- Textgenerierung: Erstellung neuer Inhalte auf Basis vorhandener Daten
- Textinterpretation: Analyse und Extraktion von Informationen aus bestehenden Texten

4.2 Bewertungssystem

Zur systematischen Bewertung der identifizierten Anwendungsfälle wurde ein Bewertungsschema mit neun Kriterien entwickelt:

- Nutzenaspekte
 - Relevanz: bewertet die strategische Bedeutung des Anwendungsfalls für das jeweilige Unternehmen
 - Effizienzsteigerung: misst das Potenzial für Zeiteinsparungen durch Automatisierung von Aufgaben
 - Qualitätssteigerung: erfasst die Verbesserung der Ergebnisqualität durch konsistentere und präzisere Verarbeitung
 - Kostensenkung: quantifiziert die Einsparung von konkreten Kosten, vor allem externen wie beispielsweise Beratungs- oder Übersetzungsleistungen.
- Umsetzbarkeitsaspekte
 - Datenverfügbarkeit: prüft das Vorhandensein und die Qualität der für den Anwendungsfall erforderlichen Daten
 - Mitarbeiterakzeptanz: schätzt die voraussichtliche Akzeptanz der Lösung durch die Mitarbeitenden im befragten Unternehmen ein
 - Technische Eignung: analysiert, ob ein LLM die optimale technische Lösung darstellt oder kostengünstigere Alternativen existieren
- Risikoaspekte
 - Datenkritikalität: beurteilt die Sensibilität der verarbeiteten Informationen bezüglich Datenschutz und Betriebsgeheimnissen
 - Fehlertoleranz: untersucht die Auswirkungen möglicher LLM-Halluzinationen und Fehlaussagen auf Geschäftsprozesse

Tabelle. Übersicht der beteiligten Unternehmen und Experten.

Nr.	Branche	Größe (Mitarbeitende weltweit)	Fertigungstiefe	Abteilung der interviewten Person
1	Maschinenbau	120	hoch	Digitalisierung
2	Lebensmittel	3.000	hoch	IT
3	Landtechnik	5.400	hoch	Digitalisierung
4	Haushaltsgeräte	2.200	hoch	Prozessmanagement
5	Maschinenbau	8.800	mittel	Innovationsmanagement
6	Pharmaindustrie	1.000	hoch	Produktionsplanung
7	Baustoffindustrie	500	hoch	IT
8	Landtechnik	12.000	hoch	Digitalisierung
9	Haushaltsgeräte	23.300	hoch	Digitalisierung
10	Maschinenbau	92.000	mittel	Digitalisierung
11	Maschinenbau	14.000	niedrig	IT

Kategorie	Nr.	Anwendungsfall	Nennungen	Ø Relevanz
A	1	Beantwortung von Fragen zu einer Anweisung oder einem Prozess	6	3,17
	2	Unterstützung bei der Durchführung einer Analyse basierend auf definierten Frameworks	4	2,75
B	1	Erstellung von Testfällen für IT-Systeme	1	4
	2	Erstellung von Lastenheften und Pflichtenheften	1	2
	3	Fehler- bzw. Abweichungsanalyse, die textbasierte Informationen berücksichtigt	5	3,4
	4	(Stamm-)Datenanalyse und -qualitätserhöhung für Freitextfelder	4	3,75
	5	Machbarkeitsprüfungen / Abfrage historischer Erfahrungen	1	2
C	1	Generierung von ausführlichen Anweisungen für Prozesse und Aufgaben	3	3
	2	Generierung von Produktbeschreibungen und -broschüren	1	4
	3	Generierung von Berichten	1	3
	4	Generierung von Protokollen (z. B. Produkt- oder Prozessänderungen)	3	2,67
	5	Generierung von gesetzlich angeforderten Unterlagen	1	2
D	1	Vereinfachung und Erklärung langer oder komplexer Anweisungen	2	2,5
	2	Zusammenfassen von Kundenanforderungen	2	2,5
	3	Zusammenfassen von Berichten und Protokollen	2	3
	4	Zusammenfassung von Besprechungen	1	4
	5	Unterstützung bei der Deutung von gesetzlichen Unterlagen	2	3
	6	Analyse von KVP-Vorschlägen seitens der Mitarbeitenden	-	-

Bild 2 Übersicht der Anwendungsfälle mit der Anzahl der Nennungen des Anwendungsfalls und der durchschnittlichen Relevanz. Grafik: eigene Darstellung

4.3 Bewertung der Anwendungsfälle mit Unternehmen durch Interviews

Die entwickelten Anwendungsfälle wurden in Experteninterviews mit elf produzierenden Unternehmen (**Tabelle**) validiert, priorisiert und bewertet. Jedes Unternehmen hat bis zu vier für sie besonders relevante Anwendungsfälle ausgewählt und diese in absteigender Relevanz priorisiert. Anschließend erfolgte eine detaillierte Bewertung anhand zuvor definierter Bewertungskriterien. Diese praxisorientierten Einschätzungen bildeten die Grundlage für das finale Relevanzranking sowie die systematische Bewertung sämtlicher Anwendungsfälle.

5 Bewertung der Anwendungsfälle

Bild 2 zeigt das Relevanzranking aller 18 identifizierten Anwendungsfälle basierend auf den Bewertungen der elf befragten Unternehmen. Bei der Bewertung konnte jedes Unternehmen bis zu vier für sie relevante Anwendungsfälle auswählen und nach Priorität ordnen. Das Bewertungssystem basiert auf einer inversen Punkteskala: Der relevanteste Anwendungsfall erhält vier Punkte, der viertwichtigste einen Punkt. Da nicht alle Anwendungsfälle von jedem Unternehmen ausgewählt wurden (einige Fälle erhielten nur eine oder gar keine Bewertung), basiert das finale Ranking auf einer gewichteten Kombination aus Bewer-

tungshäufigkeit (mindestens drei Nennungen) und durchschnittlicher Relevanz (4 ist die höchste Bewertung). Diese Methodik verhindert, dass hochspezifische Anwendungsfälle einzelner Branchen das Gesamtranking verzerren, während gleichzeitig branchenübergreifend relevante Szenarien entsprechend ihrer tatsächlichen Bedeutung für die produzierende Industrie gewichtet werden.

5.1 (Stamm-)Datenanalyse und -qualitätserhöhung für Freitextfelder

Unternehmen erstellen und sammeln Daten, die teilweise auch als Freitext eingegeben werden. Solche Eingaben können Tipp- und Formulierungsfehler enthalten, was für die aktuell verfügbaren Analysetools oder Filterverfahren Probleme bereithält. LLMs können hier sowohl während als auch nach der Eingabe helfen, die Eingabe zu standardisieren und die entstandenen Fehler zu beheben.

Die Effizienzsteigerung dieses Anwendungsfalls wird als gering eingeschätzt, da die Freitexteingabe zwar präzisere Formulierungen erlaubt, aber mehr Zeit erfordert als vordefinierte Drop-down-Eingaben. Dennoch erwarten Unternehmen eine Qualitätssteigerung, da sie sich von der Nutzung von Freitextfeldern und -eingaben eine höhere inhaltliche Präzision und Detailtiefe versprechen. Hinsichtlich der Kostensenkung gehen die Unternehmen nicht von einer direkten Reduktion expliziter Kosten aus.

Bei den Daten existieren in den befragten Unternehmen unterschiedliche Verfügbarkeiten: Während die Hälfte der befragten Unternehmen über eine ausreichende Datenbasis verfügt, fehlen der anderen Hälfte oft essenzielle Metadatenbeschreibungen, passende Vorschläge oder vollständige Formulierungen. Weiterhin ist die Akzeptanz bei den Mitarbeitenden grundsätzlich gegeben, solange die Taktzeiten nicht beeinträchtigt werden und die Auswertungen schnelle sowie sinnvolle Erkenntnisse liefern.

In Bezug auf die Lösungseignung zeigen sich in diesem Anwendungsfall vordefinierte Gründe, Kategorien, Textbausteine oder Drop-down-Eingaben als Alternativen zu LLMs. Allerdings sind LLMs für nachgelagerte Auswertungsmöglichkeiten aktuell alternativlos. Die Kritikalität der Daten ist jedoch hoch, da es sich um präzise Produkt- und Materialbeschreibungen oder um Rückmeldungen zu Prozessen und Qualität handelt, die eine besonders sorgfältige Behandlung erfordern. Ebenso ist die Kritikalität von Fehlaussagen nicht zu unterschätzen: Während Echtzeiteingaben und -ergänzungen stets zuverlässig sein müssen, sind Fehler in nachgelagerten Analysen oder Anpassungen bis zu einem gewissen Grad tolerierbar.

Ein wesentlicher Vorteil ist demnach die erhöhte Qualität der erfassten Daten. Durch mehr Standardisierung, eine gezielte Ausfüllunterstützung und die Reduzierung von Copy-&-Paste-Fehlern wird sichergestellt, dass die Eingaben präziser und konsistenter erfolgen. Dies trägt dazu bei, die erfassten Informationen verlässlicher und besser nutzbar zu machen. Zusätzlich eröffnen sich neue Analysemöglichkeiten. Große Mengen an Textdaten können durch automatisierte Auswertungen schnell analysiert werden, wodurch eine verbesserte Transparenz über die Inhalte der Freitextfelder entsteht. Dies erlaubt es Unternehmen, aus unstrukturierten Daten wertvolle Erkenntnisse zu gewinnen und gezieltere Entscheidungen zu treffen.

Den Vorteilen stehen Herausforderungen gegenüber, vor allem der erhöhte Zeitaufwand. Da die Eingabe in Freitextfeldern sonst

oft ausbleibt, führt die für den Anwendungsfall notwendige Nutzung zu einer Verlangsamung des Prozesses. Dieser Mehraufwand muss durch eine spürbare Qualitätssteigerung der erfassten Daten ausgeglichen werden, um eine sinnvolle Umsetzung zu gewährleisten. Zudem besteht das Risiko des Strukturverlusts. Wenn benötigte Daten nicht mehr vorstrukturiert gespeichert werden, droht Verlust von Strukturen und Übersichtlichkeit von Daten. Langfristig kann dies erhebliche Probleme in der Datenhaltung und Arbeitsweise verursachen, da unstrukturierte Daten schwerer auszuwerten und in bestehende Systeme zu integrieren sind.

Zur Verbesserung der Datenqualität und Fehlererkennung von Freitextfeldern eignen sich LLMs gut. Jedoch sollte eine Balance zwischen erhöhter Datenqualität und dem Zeitaufwand für die Eingabe gewährleistet sein, um die Akzeptanz der Mitarbeitenden zu gewährleisten. Für manche Rückmeldungen eignen sich nach wie vor am besten vorformulierte Kategorien und die Möglichkeit zum Anhängen von Bildern.

5.2 Fehler- oder Abweichungsanalyse, die textbasierte Informationen berücksichtigt

In Produktionsprozessen treten regelmäßig Fehler und Abweichungen auf, deren Ursachen oft in textbasierten Informationen, beispielsweise in Wartungsprotokollen, Qualitätsberichten oder Fehlermeldungen, verborgen sind. Mithilfe von LLMs können in diesem Anwendungsfall Analysen durchgeführt werden, welche schnell Fehlerursachen aus vielen textbasierten Daten und Unterlagen identifizieren können.

Die somit von den Unternehmen erwartete Effizienzsteigerung ist hoch, da solche Aufgaben derzeit manuell durchgeführt werden. Eine Automatisierung durch LLMs könnte den Prozess erheblich beschleunigen und die Arbeitsbelastung der Mitarbeitenden reduzieren. Auch die Qualitätssteigerung ist gegeben, da weitere, bislang verborgene Fehlerursachen schneller identifiziert und behoben werden können. Dies ermöglicht eine präzisere Analyse und trägt dazu bei, potenzielle Probleme frühzeitig zu erkennen. Eine direkte Kostensenkung wird dagegen nicht erwartet. Zwar können Effizienzgewinne erzielt werden, doch eine Reduktion expliziter Kosten durch den Einsatz von LLMs ist nicht unmittelbar absehbar.

Die Datenverfügbarkeit stellt in diesem Fall keinen grundlegenden Engpass dar. Für die Nutzung und Analyse sind relevante Daten wie Berichte bereits vorhanden, müssen jedoch manuell zusammengesucht werden. Vor allem für das Modelltraining ist eine zusätzliche Prüfung der Aussage- und Formqualität dieser Textdateien erforderlich, um eine sinnvolle Grundlage für LLMs zu schaffen. Die Mitarbeitendenakzeptanz wird als hoch eingeschätzt, da der Einsatz von LLMs monotone und repetitive Tätigkeiten reduzieren kann. Durch die Automatisierung von Routineaufgaben wird den Beschäftigten mehr Zeit für wertschöpfende Tätigkeiten eingeräumt.

Die Lösungseignung ist ebenfalls hoch, da die betroffenen Textdaten und Berichte bislang nur manuell ausgewertet werden können. Eine automatisierte Analyse mittels LLMs könnte diesen Prozess erheblich erleichtern und beschleunigen. Die Kritikalität der Daten ist hoch, da es sich um Rückmeldungen zu Prozessen oder zur Produktqualität handelt. Daher müssen diese besonders geschützt und sorgfältig verarbeitet werden, um Fehlschlüsse oder falsche Bewertungen zu vermeiden. Die Kritikalität von Fehlaussagen variiert je nach Branche. In den meisten Fällen sind

Fehlaussagen in nachgelagerten Analysen tolerierbar. Eine Ausnahme bildet die Pharmaindustrie, in der Fehlinformationen nicht akzeptabel sind. Falls Fehlaussagen auftreten, muss dort eine manuelle rückwirkende Prüfung aller Daten erfolgen, was die potenziellen Effizienzgewinne erheblich reduzieren könnte.

Ein großer Vorteil dieses Anwendungsfalls ist die Möglichkeit, verschiedene Datenquellen zusammenzuführen. Informationen aus Berichten, Schichtübergabeprotokollen und Freitexteingaben aus IT-Systemen können gemeinsam verarbeitet werden, wodurch Zusammenhänge besser erkennbar und Analysen umfassender werden. Zusätzlich bieten LLMs die Chance, umfangreiche Textdaten effizient zu analysieren. Automatisierte Auswertungen ersetzen zeitaufwendige manuelle Tätigkeiten und ermöglichen eine schnellere Gewinnung relevanter Erkenntnisse. Somit lassen sich Abläufe optimieren und Fachkräfte können sich auf stärker wertschöpfende Aufgaben konzentrieren.

Dennoch gibt es Herausforderungen, vor allem im Umgang mit kritischen Daten. Prozessinformationen müssen besonders geschützt werden, sodass die Verarbeitung möglichst lokal erfolgen sollte, um Sicherheitsrisiken zu minimieren. Eine weitere Schwierigkeit liegt in der Auswahl geeigneter Datenquellen. Um verlässliche und präzise Ergebnisse zu erhalten, ist es essenziell, vorab festzulegen, welche Quellen für die Analyse relevant sind. Ohne eine klare Struktur besteht die Gefahr, dass wichtige Informationen ungenutzt bleiben oder unpassende Daten in die Auswertung einfließen.

Dieser Anwendungsfall erlaubt die Analyse von bisher selten analysierten Unterlagen oder die Verknüpfung mehrerer Datenquellen (sowohl Zahlenwerte als auch Texte) zur Generierung neuer Erkenntnisse zur Prozess- oder Qualitätsverbesserung.

5.3 Beantwortung von Fragen zu einer Anweisung oder einem Prozess

Dieser Anwendungsfall ermöglicht es den Mitarbeitenden, schnell Antworten auf ihre spezifischen Fragen zu erhalten, die in Arbeitsanweisungen und Prozessen beschrieben sind. Dies reduziert Suchzeiten und stellt sicher, dass die Anweisungen korrekt und effizient befolgt werden.

Der Einsatz von LLMs kann die Effizienz steigern, da die manuelle Suche nach den richtigen Antworten erheblich reduziert wird. Mitarbeitende erhalten schneller präzise Informationen, wodurch Arbeitsprozesse beschleunigt und unnötige Verzögerungen vermieden werden. Besonders positiv wirkt sich die Technologie auf die Qualität der Prozesse aus. Da relevante Antworten schneller zur Verfügung stehen, können Fehler reduziert und Abläufe optimiert werden. Dies führt zu einer spürbaren Verbesserung der Prozessqualität und einer höheren Zufriedenheit der Mitarbeitenden. Ein weiteres Potenzial liegt in der möglichen Kostensenkung, vor allem durch die Reduktion von Supportzeiten. Können Anfragen durch ein LLM automatisiert beantwortet werden, reduziert sich der Bedarf an personellem Support, was gerade in Zeiten mit geringer Personalverfügbarkeit – etwa an Feiertagen oder während Nachtschichten – von Vorteil ist.

Die Datenverfügbarkeit ist jedoch eine Herausforderung: Viele Unternehmen verfügen bereits über umfassende Prozessbeschreibungen und Arbeitsanweisungen, die als Grundlage für ein LLM dienen könnten. Allerdings sind diese Dokumente oft veraltet oder widersprüchlich, sodass eine vorherige Aktualisierung und Harmonisierung erforderlich sind. Zudem sind wertvolle Erfah-

rungswerte und implizites Wissen der Mitarbeitenden oft nicht schriftlich festgehalten. Weiterhin hängt die Akzeptanz unter den Mitarbeitenden stark vom wahrgenommenen Nutzen der Lösung ab. Während einige die Erleichterung bei der Informationsbeschaffung begrüßen, gibt es auch Bedenken hinsichtlich einer möglichen Überwachung, was in einem Fall explizit als Sorge genannt wurde.

Die Lösung eignet sich besonders gut für die gezielte Suche nach Informationen. Allerdings gibt es alternative Methoden wie klassische FAQ-Bereiche, Keyword-Suchen oder visuelle Anleitungen in Form von Bildern und Videos, die je nach Anwendungsfall sogar sinnvoller sein können. Bei der Kritikalität der Daten ist das Risiko in den meisten Fällen gering. Nur wenn Arbeitsanweisungen oder Prozessdokumentationen spezifisches, geschütztes Know-how enthalten, muss auf eine angemessene Absicherung geachtet werden. Besonders kritisch ist die Verlässlichkeit der generierten Antworten. Während kleinere Ungenauigkeiten in allgemeinen Anfragen tolerierbar sein können, dürfen im laufenden Produktionsprozess keine Fehlaussagen auftreten. Falsche oder missverständliche Anweisungen könnten dazu führen, dass Prozessschritte fehlerhaft ausgeführt werden, was direkte Auswirkungen auf die Produktqualität sowie nachfolgende Abläufe hätte.

Ein wesentlicher Vorteil des Anwendungsfalls ist der schnelle Zugang zu Informationen. Statt lange Anweisungen zu lesen oder zeitaufwendig nach relevanten Dokumenten im Intranet zu suchen, können Mitarbeitende gezielt Fragen stellen und erhalten sofort passende Antworten. Dies macht die Informationsbeschaffung deutlich effizienter und attraktiver. Außerdem bieten LLMs neue Trainingsmöglichkeiten für das Personal. Durch individualisierte Inhalte, die sich sowohl thematisch als auch zeitlich an den individuellen Lernfortschritt anpassen, können neue Mitarbeitende effektiver eingearbeitet werden. Dies ermöglicht flexibles und bedarfsgerechtes Lernen, ohne dass feste Schulungstermine erforderlich sind.

Zu den Herausforderungen zählen vor allem Fehlaussagen, die ein erhebliches Risiko darstellen, insbesondere wenn falsche oder missverständliche Informationen im laufenden Produktionsprozess zu Qualitätsproblemen oder Verzögerungen führen. Eine sorgfältige Validierung der generierten Antworten ist daher essenziell. Auch kann ein LLM den persönlichen Austausch mit erfahrenen Mitarbeitenden nicht ersetzen. In vielen Unternehmen steckt das entscheidende Prozesswissen ausschließlich in den Köpfen erfahrener Fachkräfte. Daher bleibt der direkte Austausch mit Vorgesetzten sowie Kolleginnen und Kollegen weiterhin unerlässlich, um praktisches Wissen zu vermitteln und kritische Fragen zu klären.

Eine weitere Schwierigkeit ergibt sich aus aufwendigen oder nicht offensichtlichen Beschreibungen. Manche Fragestellungen lassen sich nur schwer in Worte fassen, besonders wenn sie sich auf sensorische Wahrnehmungen wie Geräusche oder bestimmte visuelle Merkmale beziehen. Außerdem müssen Mitarbeitende wissen, wie sie ihre Fragen formulieren können, um zielführende Antworten zu erhalten. Eine intuitive und gut strukturierte Nutzung des Systems ist daher entscheidend, um Missverständnisse und Frustration zu vermeiden.

LLMs bieten in diesem Anwendungsfall eine Möglichkeit, schnell auf individuell benötigte Informationen zuzugreifen. Es bedarf jedoch qualitativ hochwertiger Anweisungen und Prozessbeschreibungen. Da Fehlaussagen im laufenden Produktions-

prozess sehr kritisch sind, eignet sich der Anwendungsfall vor allem für Trainingszwecke oder für Situationen, in denen sonst keine Hilfe verfügbar wäre (zum Beispiel an Feiertagen oder während Nachtschichten).

6 Zusammenfassung

Die Studie zeigt, dass Unternehmen ein hohes Potenzial in der Automatisierung manueller Texterfassung und -auswertung durch LLMs erkennen, was vor allem zu Effizienz- und Qualitätssteigerungen führen kann. Die Akzeptanz bei Mitarbeitenden ist höher, wenn LLMs vorbereitende statt finaler Arbeitsschritte übernehmen, um Risiken durch mögliche Fehlaussagen zu minimieren und Doppelarbeit zu vermeiden. Obwohl Datensicherheit als essenziell bewertet wird, fehlt es vielen Unternehmen noch an einer klaren Governance im Umgang mit produktionsbezogenen Daten.

Bezogen auf die Frage der Studie lässt sich festhalten, dass LLMs im Produktionsumfeld vorrangig bei unterstützenden Tätigkeiten sinnvoll eingesetzt werden können – etwa in den Bereichen Dokumentation, Schulung, Wissensmanagement oder bei der Auswertung textbasierter Informationen. In den direkten Produktionsprozessen hingegen, die häufig stark strukturiert, zeitkritisch und auf präzise Eingaben angewiesen sind, besteht aktuell noch ein zu hohes Risiko durch fehlerhafte oder ungenaue Ausgaben. Der Prüf- und Anpassungsaufwand von LLM-Vorschlägen nivelliert in schnell getakteten und zeitkritischen Produktionsprozessen die Vorteile von LLMs.

Angesichts der dynamischen Weiterentwicklung von LLMs erscheint es zwar denkbar, dass operative Produktionsaufgaben künftig direkt unterstützt werden können. Gegenwärtig sollten Unternehmen jedoch vorrangig auf etablierte IT-Systeme zurückgreifen und zugleich eine klare, strukturierte Datenbasis schaffen, um LLMs gezielt für textbasierte Assistenzfunktionen zu nutzen.

LITERATUR

- [1] Minaee, S.; Mikolov, T.; Nikzad, N. et al.: Large Language Models: A Survey. ArXiv 2024, doi.org/10.48550/arXiv.2402.06196
- [2] Roland Berger (ed.): GenAI-driven transformation. Preparing your company for success with GenAI on every level. Stand: 2024. Internet: content.rolandberger.com/hubfs/07_presse/Roland%20Berger%20GenAI-driven%20transformation%20final.pdf. Zugriff am 22.08.2025
- [3] Buchholz, K.: Threads Shoots Past One Million User Mark at Lightning Speed. One Million Users. Statista. Stand: 07.07.2023. Internet: www.statista.com/chart/29174/time-to-one-million-users/. Zugriff am 22.08.2025
- [4] Fletcher, R.; Nielsen, R. K.: What Does the Public in Six Countries Think of Generative AI in News? Stand: 28.05.2024. Internet: reutersinstitute.politics.ox.ac.uk/what-does-public-six-countries-think-generative-ai-news. Zugriff am 22.08.2025
- [5] Singla, A.; Sukharevsky, A.; Yee, I. et al.: The state of AI in early 2024. Gen AI adoption spikes and stats to generate value. McKinsey. Stand: 2024. Internet: www.mckinsey.com/capabilities/quantumblack/our-insights/the-state-of-ai. Zugriff am 22.08.2025 11.06.2025
- [6] Bick, A.; Blandin, A.; Deming, D. J.: The Rapid Adoption of Generative AI. National Bureau of Economic Research. Stand: 04.11.2024 Internet: papers.ssrn.com/sol3/papers.cfm?abstract_id=4965142. Zugriff am 22.08.2025
- [7] Chui, M.; Roberts, R.; Yee, L. et al.: The economic potential of generative AI. The next productivity frontier. McKinsey & Company. Stand: 2023. Internet: www.mckinsey.com/capabilities/mckinsey-digital/our-insights/the-economic-potential-of-generative-ai-the-next-productivity-frontier#. Zugriff am 22.08.2025
- [8] Kurpicz-Briki, M.: Generative AI: was ist das und was kann sie bereits? SocietyByte – Wissenschaftsmagazin der Berner Fachhochschule. (2023), doi.org/10.24451/arbor.19044
- [9] Bharathi Mohan, G.; Prasanna Kumar, R.; Vishal Krishh, P. et al.: An analysis of large language models: their impact and potential applications. Knowledge and Information Systems 66 (2024) 9, pp. 5047–5070
- [10] Rahmzadehgervi, P.; Bolton, L.; Taesiri, M. R. et al.: Vision language models are blind. ArXiv.2025, doi.org/10.48550/arXiv.2407.06581
- [11] Ma, J.; Dai, D.; Sha, L. et al.: Large Language Models Are Unconscious of Unreasonability in Math Problems. ArXiv 2024, arxiv.org/html/2403.19346v2
- [12] Ahn, J.; Verma, R.; Lou, R. et al.: Large Language Models for Mathematical Reasoning: Progresses and Challenges. ArXiv 2024, doi.org/10.48550/arXiv.2402.00157
- [13] Xu, Z.; Jain, S.; Kankanhalli, M.: Hallucination is Inevitable: An Innate Limitation of Large Language Models. ArXiv 2024, doi.org/10.48550/arXiv.2401.11817
- [14] Huang, D.; Yan, C.; Li, Q. et al.: From Large Language Models to Large Multimodal Models: A Literature Review. Applied Sciences 14 (2024) 12, #5068
- [15] Li, Y.; Zhao, H.; Jiang, H. et al.: Large Language Models for Manufacturing. ArXiv 2024, doi.org/10.48550/arXiv.2401.11817
- [16] Milmo, D.: OpenAI launches 'deep research' tool that it says can match research analyst. The Guardian. Stand: 03.02.2025. Internet: www.theguardian.com/technology/2025/feb/03/openai-deep-research-agent-chatgpt-deepseek. Zugriff am 22.08.2025
- [17] Pichai, S.; Hassabis, D.: Our next-generation model: Gemini 1.5. Stand: 15.02.2024. Internet: blog.google/technology/ai/google-gemini-next-generation-model-february-2024/. Zugriff am 22.08.2025
- [18] Liu, A.; Feng, B.; Wang, B. et al.: DeepSeek-V2: A Strong, Economical, and Efficient Mixture-of-Experts Language Model. ArXiv 2024, doi.org/10.48550/arXiv.2405.04434
- [19] Firat, M.; Kuleli, S.: What if GPT4 Became Autonomous: The Auto-GPT Project and Use Cases. Journal of Emerging Computer Technologies 3 (2024) 1, pp. 1–6
- [20] Bruscia, M.; Manduzio, G. A.; Galatolo, F. A. et al.: An Overview On Large Language Models Across Key Domains: A Systematic Review. 2024 IEEE International Conference on Metrology for eXtended Reality, Artificial Intelligence and Neural Engineering (MetroXRAINE), St Albans, United Kingdom, 2024, pp. 125–130
- [21] Wang, T.; Fan, J.; Zheng, P.: An LLM-based vision and language cobot navigation approach for Human-centric Smart Manufacturing. Journal of Manufacturing Systems 75 (2024), pp. 299–305
- [22] Lim, J.; Patel, S.; Evans, A. et al.: Enhancing Human-Robot Collaborative Assembly in Manufacturing Systems Using Large Language Models. Stand: ArXiv 2024, doi.org/10.48550/arXiv.2406.01915
- [23] Finkel, P.; Wurster, P.; Radler, R.: Large Language Models (LLM) im Produktionsumfeld. Eine Potenzialanalyse zur Transformation von Produktionsabläufen in modernen Fabriken. Industry 4.0 Science (2024) 40, S. 48–55
- [24] Kernan Freire, S.; Wang, C.; Foosherian, M. et al.: Knowledge sharing in manufacturing using LLM-powered tools: user study and model benchmarking. Frontiers in artificial intelligence 7 (2024), #1293084, doi.org/10.3389/frai.2024.1293084
- [25] Mustapha, K. B.: A survey of emerging applications of large language models for problems in mechanics, product design, and manufacturing. Advanced Engineering Informatics 64 (2025), #103066
- [26] Xia, Y.; Jazdi, N.; Weyrich, M.: Applying Large Language Models for Intelligent Industrial Automation. ATP Magazin 66 (2024) 6–7, pp. 62–71
- [27] Khanghah, K. N.; Patel, A.; Malhotra, R. et al.: Large Language Models for Extrapolative Modeling of Manufacturing Processes. ArXiv 2025, doi.org/10.48550/arXiv.2502.12185
- [28] Makatura, L.; Foshey, M.; Wang, B. et al.: Large Language Models for Design and Manufacturing. An MIT Exploration of Generative AI (27.03.2024), doi.org/10.21428/e4baedd9.745b62fa

Nikita Fjodorovs M.Sc. 
nikita.fjodorovs@fir.rwth-aachen.de

Alan Gütermann B.Sc.
FIR e. V. an der RWTH Aachen
Campus-Boulevard 55, 52074 Aachen
www.fir.rwth-aachen.de

LIZENZ



Dieser Fachaufsatz steht unter der Lizenz Creative Commons
Namensnennung 4.0 International (CC BY 4.0)