KI & GENERATIVE KI LÖSUNGEN FÜR UNTERNEHMEN

ÜBERSICHT

Das exponentielle Wachstum der generativen KI in den letzten Monaten verändert die Arbeitswelt erheblich und hat Auswirkungen auf Unternehmen weltweit. Dieser Trend spiegelt sich in Berichten von McKinsey, Forbes und dem Weltwirtschaftsforum wider und zeigt sich auch in den strategischen Entscheidungen führender Technologieunternehmen.

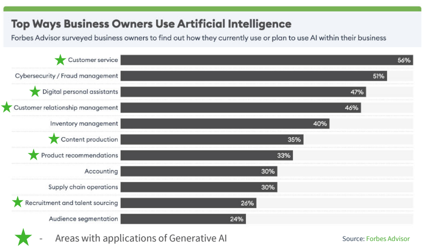
Hier stellen wir ein Portfolio vor, das Unternehmen helfen kann, mit den sich schnell entwickelnden Technologien Schritt zu halten und sie zu nutzen. Es skizziert verschiedene Möglichkeiten, wie das Unternehmen KI und generative KI in seine Geschäftsprozesse integrieren kann.

Das Dokument besteht aus drei Hauptteilen. Der erste Teil veranschaulicht die Bedeutung von KI und generativer KI für Unternehmen: Was ist generative KI, wie kann sie einem Unternehmen helfen, was sind ihre Risiken und Grenzen, und wie können diese gemildert werden? Der zweite Abschnitt skizziert mögliche Anwendungen von KI und generativer KI für kundenorientierte, interne und hybride Kontexte und bietet ausführliche illustrative Beispiele. Der dritte und letzte Abschnitt schließt mit einem **Ausblick auf mögliche zukünftige Schritte.**

KI IM BUSINESS

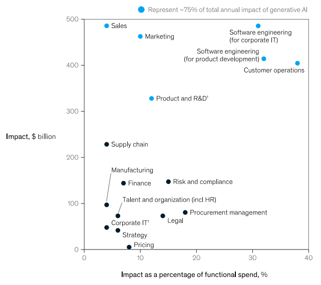
Die KI kommt den Unternehmen bereits in mehreren Anwendungsbereichen zugute (siehe Abbildung rechts).

Das Aufkommen der generativen KI hat zu Anwendungen auch in kreativen Bereichen wie der Content-Generierung geführt, was wiederum einen großen Einfluss auf Vertrieb und Marketing haben kann. Die Abbildung von McKinsey (siehe unten) veranschaulicht die finanziellen Auswirkungen auf eine breite Palette von Branchen.



Source: Forbes Advisor

Possible impact of generative AI on business functions



Note: Impact is averaged.
¹Excluding software engineering.
Source: Comparative Industry Service (CIS), IHS Markit; Oxford Economics; McKinsey Corporate and Business Functions database; McKinsey Manufacturing and Supply Chain 360; McKinsey Sales Navigator; Ignite, a McKinsey database; McKinsey analysis

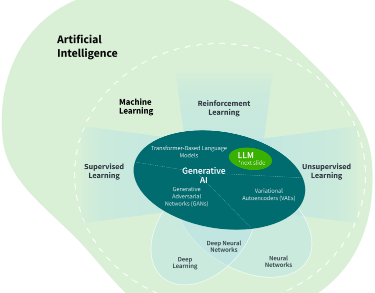
Source: McKinsey

Zu den üblichen Anwendungen der generativen KI gehören:

- Organisation von Support-Tickets
- Meinungsforschung und Berichterstattung
- Bewertung von Mitbewerbern
- Automatisierte Beantwortung von mehrsprachigen Anfragen
- Unterstützung bei der Recherche
- Berichterstattung über Ausgaben

WAS IST GENERATIVE KI?

- **KI** ist ein Oberbegriff, der mehrere Teilbereiche und Techniken umfasst.
- **Generative KI** ist ein Teilbereich der KI, der sich auf die Erstellung **von neuem, einzigartigem Content** auf der Grundlage fortschrittlicher Architekturen wie GANs, VAEs und Transformers spezialisiert hat.
- Generative KI überschneidet **mehrere Teilbereiche des maschinellen Lernens** wie Supervised-, Unsupervised-, und Reinforcement-Lernens.
- Es kann mit **verschiedenen Datentypen** umgehen, z. B. mit Bildern, Text und strukturierten Tabellendaten.
- Zu den Anwendungsbereichen der generativen KI gehören:
  - Synthetische Datengenerierung,
  - Simulation von Zukunftsszenarien,
  - Hyper-Personalisierung von Erlebnissen und mehr.



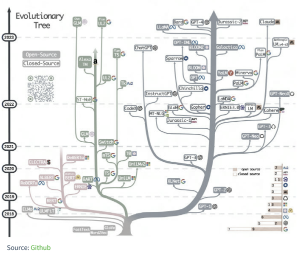
Source: Blog post:Jeff Winter

✓ ***Grundlage intelligenter Systeme***

✓ ***Breiter branchenübergreifender Einsatz***

✓ ***Unterstützt Automatisierung und Entscheidungsfindung***

Large Language Models



Source: Github

Entwicklung & Beschränkungen

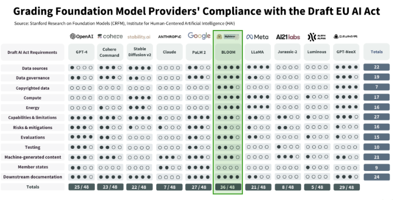
Große Sprachmodelle **(Large Language Models, LLMs)** haben sich in den letzten Jahren dank der zunehmenden Rechenressourcen und der Verfügbarkeit großer Datenmengen rasch weiterentwickelt (siehe Abbildung links). Obwohl ChatGPT, das Closed-Source-Modell von OpenAI, zu den bekannteren Akteuren in der KI-Landschaft zählt, gilt es zu bedenken, dass es nur Teil eines **viel größeren, vielfältigen Ökosystems** ist, das viele Open-Source-Modell in mehreren Sprachen umfasst.

LLMs bieten viele Vorteile und haben das Potenzial, die Produktivität sowohl im geschäftlichen als auch im alltäglichen Leben erheblich zu verbessern. Es ist jedoch wichtig, ihre Grenzen zu verstehen. Eine dieser Einschränkungen ist das Phänomen der **Halluzinationen**, d. h. die Tendenz der Modelle, Texte zu produzieren, die zwar richtig zu sein scheinen, aber in Wirklichkeit falsch sind. Glücklicherweise gibt es Möglichkeiten, Halluzinationen abzuschwächen und zu minimieren - zum Beispiel **durch weiteres Finetuning der LLMs auf domänenspezifische Daten.**

Datenschutz und EU-KI-Gesetz

Ein weiterer wichtiger Punkt ist, dass es in der Regel einen Kompromiss zwischen Modellleistung und Datenschutz gibt. Closed-Source-Modelle wie ChatGPT mögen bei bestimmten Aufgaben besser abschneiden, aber es ist nicht bekannt, **mit welcher Art von Daten** das Modell trainiert wurde und **was mit den Nutzerdaten geschieht**. Daher sollten sensible, vertrauliche Informationen nicht in solche Systeme eingespeist werden. Diese Punkte wurden auch im **EU-KI-Gesetz** angesprochen.

Die Abbildung auf der rechten Seite zeigt eine Bewertung verschiedener LLMs anhand der Rubriken des EU-KI-Gesetzes. Man sieht zum Beispiel, dass das Open-Source-Modell BLOOM von Hugging Face im Vergleich zu anderen bekannteren Closed-Source-Modellen die höchste Punktzahl bei der Einhaltung des EU-KI-Gesetzes erreicht. Diesen Punkt sollte man bei der Entwicklung und dem Einsatz von KI-Systemen im Auge behalten. Es ist möglich, diese Überlegungen zu berücksichtigen und die Einhaltung der Datenschutzbestimmungen durch die Implementierung **eines lokal gehosteten Open-Source-Modells** zu gewährleisten.



Grading Foundation Model Providers' Compliance with the Draft EU AI Act

Source: Stanford Research on Foundation Models (CRFM), Institute for Human-Centered Artificial Intelligence (HAI)

	OpenAI	cohere	stability.ai	ANTHROPIC	Google		Meta	AI21labs		EleutherAI	
Draft AI Act Requirements	GPT-4	Cohere Command	Stable Diffusion v2	Claude	PaLM 2	BLOOM	LLaMA	Jurassic-2	Luminous	GPT-NeoX	Totals
Data sources	●○○○	●●●○	●●●●	○○○○	●●○○	●●●●	●●●●	○○○○	○○○○	●●●●	22
Data governance	●●○○	●●●○	●●○○	○○○○	●●●○	●●●●	●●○○	○○○○	○○○○	●●●○	19
Copyrighted data	○○○○	○○○○	○○○○	○○○○	○○○○	●●●○	○○○○	○○○○	○○○○	●●●●	7
Compute	○○○○	○○○○	●●●●	○○○○	○○○○	●●●●	●●●●	○○○○	●○○○	●●●●	17
Energy	○○○○	●○○○	●●●○	○○○○	○○○○	●●●●	●●●●	○○○○	○○○○	●●●●	16
Capabilities & limitations	●●●●	●●●○	●●●●	●○○○	●●●●	●●●○	●●○○	●●○○	●○○○	●●●○	27
Risks & mitigations	●●●○	●●○○	●○○○	●○○○	●●●○	●●○○	●○○○	●●○○	○○○○	●○○○	16
Evaluations	●●●●	●●○○	○○○○	○○○○	●●○○	●●●○	●●○○	○○○○	●○○○	●○○○	15
Testing	●●●○	●●○○	○○○○	○○○○	●●○○	●●○○	○○○○	●○○○	○○○○	○○○○	10
Machine-generated content	●●●○	●●●○	○○○○	●●●○	●●●○	●●●○	○○○○	●●●○	●○○○	●●○○	21
Member states	●●○○	○○○○	○○○○	●●○○	●●●●	○○○○	○○○○	○○○○	●○○○	○○○○	9
Downstream documentation	●●●○	●●●●	●●●●	○○○○	●●●●	●●●●	●●○○	○○○○	○○○○	●●●○	24
Totals	25 / 48	23 / 48	22 / 48	7 / 48	27 / 48	36 / 48	21 / 48	8 / 48	5 / 48	29 / 48	

Source: Stanford

Was bedeutet das für ein Unternehmen?

Sicherstellung des Datenschutzes

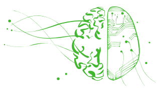
- Entwicklung von KI-Systemen in Übereinstimmung mit dem EU-KI-Gesetz.
- Verwendung eines lokal gehosteten Open-Source-Modells zur Gewährleistung der Vertraulichkeit von Kundeninformationen oder sensiblen Unternehmensdaten.

Die Fallstricke von LLMs Vermeiden

- Minimierung von Halluzinationen durch Finetuning des LLM auf bereichsspezifische Daten.

Vorteile von LLMs

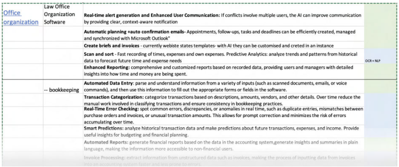
- Steigerung der Effizienz in der Steuerverwaltung, bei der Kundenbetreuung, bei internen Abläufen und mehr.
- Ausführlichere Beispiele von Anwendungsfällen folgen im nächsten Abschnitt.



ANWENDUNGSBEREICHE

Wir haben eine **umfassende Liste von generativen KI-Anwendungen** zusammengestellt, die die Prozesse des Unternehmens ergänzen könnten. Die Ideen basieren auf einer sorgfältigen Prüfung der Dienstleistungen des Unternehmens sowie der allgemeinen Arbeitsabläufe in bestimmten Branchen. Wir haben das **Potenzial von KI zur Verbesserung oder Automatisierung breiten Palette von Aufgaben** bewertet und anhand dieses Beispiels 100 vielversprechende Szenarien ermittelt. Gemeinsam mit unseren Experten für maschinelles Lernen haben wir die Ideen weiter verfeinert und kategorisiert, wobei wir zwischen Anwendungsfällen für generative KI allein und solchen, die zusätzliche KI-Technologien erfordern, unterschieden haben.

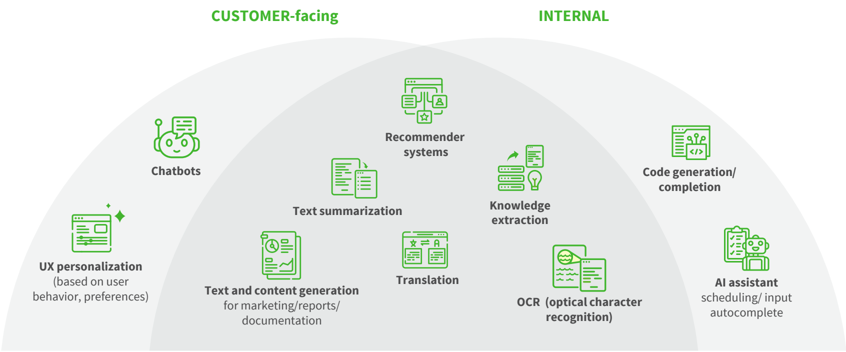
Der Screenshot auf der rechten Seite zeigt eine Momentaufnahme der identifizierten Anwendungsfälle.



Office organization	Law Office Organization Software	**Real-time alert generation and Enhanced User Communication:** If conflicts involve multiple users, the AI can improve communication by providing clear, context-aware notification	
		Automatic planning +auto confirmation emails- Appointments, follow-ups, tasks and deadlines can be efficiently created, managed and synchronized with Microsoft Outlook"	
		Create briefs and invoices - currently webiste states templates- with AI they can be customised and creted in an instance	
		Scan and sort - Fast recording of times, expenses and own expenses. Predictive Analytics: analyze trends and patterns from historical data to forecast future time and expense needs	OCR + NLP
		Enhanced Reporting: comprehensive and customized reports based on recorded data, providing users and managers with detailed insights into how time and money are being spent.	
	-- bookkeeping	**Automated Data Entry**: parse and understand information from a variety of inputs (such as scanned documents, emails, or voice commands), and then use this information to fill out the appropriate forms or fields in the software.	
		Transaction Categorization: categorize transactions based on descriptions, amounts, vendors, and other details. Over time reduce the manual work involved in classifying transactions and ensure consistency in bookkeeping practices.	
		Real-Time Error Checking: spot common errors, discrepancies, or anomalies in real time, such as duplicate entries, mismatches between purchase orders and invoices, or unusual transaction amounts. This allows for prompt correction and minimizes the risk of errors accumulating over time.	
		Smart Predictions: analyze historical transaction data and make predictions about future transactions, expenses, and income. Provide useful insights for budgeting and financial planning.	
		Automated Reports: generate financial reports based on the data in the accounting system,generate insights and summaries in plain language, making the information more accessible to non-financial users.	
		Invoice Processing: extract information from unstructured data such as invoices, making the process of inputting data from invoices into an accounting system faster and less prone to errors.	

Die vollständige Liste der potenziellen Anwendungsfälle könnten wir in einem kleinen Workshop zwischen dem Kunden und QualityMinds erläutern. Bitte beachten Sie, dass sich hierbei um Ausgangspunkte für weitere Untersuchungen handelt und nicht um endgültige Ergebnisse. Dies spiegelt die dynamische und sich ständig weiterentwickelnde Natur der KI-Technologie wider.

Darüber hinaus veranschaulicht das nachstehende **Venn-Diagramm potenzielle Anwendungen von allgemeiner und generativer KI** sowohl für kundenorientierte als auch für interne Wissensmanagement-Szenarien.


CUSTOMER-facing
INTERNAL
Recommender systems
Chatbots
Code generation/ completion
Text summarization
Knowledge extraction
UX personalization (based on user behavior, preferences)
Text and content generation for marketing/reports/ documentation
Translation
OCR (optical character recognition)
AI assistant scheduling/ input autocomplete

Generative KI Beispiele

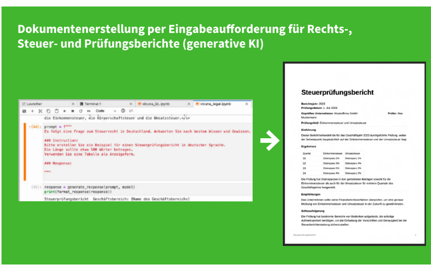
Beispiel 1: AI Doc

Automatische Generierung neuer Inhalte (z.B. Finanzberichte, Rechtsdokumente, Rechnungen) auf der Grundlage von Eingabeinformationen

Zusammenfassen von Dokumenten

Verringerung menschlicher Fehler, Verbesserung der Effizienz und Konsistenz sowie Freisetzung von Mitarbeitenden für komplexere Aufgaben und Analysen

Dokumentenerstellung per Eingabeaufforderung für Rechts-, Steuer- und Prüfungsberichte (generative KI)



```
die Einkommensteuer, die Körperschaftsteuer und die Umsatzsteuer.</s>

[44]: prompt = f"""
Es folgt eine Frage zum Steuerrecht in Deutschland. Antworten Sie nach besten Wissen und Gewissen.

### Instruction:
Bitte erstellen Sie ein Beispiel für einen Steuerprüfungsbericht in deutscher Sprache.
Die Länge sollte etwa 500 Wörter betragen.
Verwenden Sie eine Tabelle als Anzeigeform.

### Response:

"""

[45]: response = generate_response(prompt, model)
print(format_response(response))

Steuerprüfungsbericht  Geschäftsbereich: [Name des Geschäftsbereichs]
```

Steuerprüfungsbericht

Berichtsjahr: 2023
Prüfungsdatum: 1. Juli 2023
Geprüftes Unternehmen: Musterfirma GmbH **Prüfer:** Max Mustermann
Prüfungsfeld: Einkommensteuer und Umsatzsteuer

Einführung

Dieser Bericht behandelt die für das Geschäftsjahr 2023 durchgeführte Prüfung, wobei der Schwerpunkt hauptsächlich auf der Einkommensteuer und der Umsatzsteuer liegt.

Ergebnisse

Quartal	Einkommensteuer	Umsatzsteuer
Q1	Diskrepanz 2%	Diskrepanz 1%
Q2	Diskrepanz 3%	Diskrepanz 4%
Q3	Diskrepanz 5%	Diskrepanz 3%
Q4	Diskrepanz 4%	Diskrepanz 2%

Die Prüfung hat Diskrepanzen in den gemeldeten Beträgen sowohl für die Einkommensteuer als auch für die Umsatzsteuer für mehrere Quartale des Geschäftsjahres festgestellt.

Empfehlungen

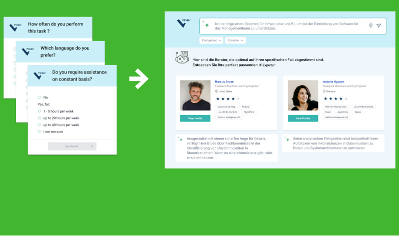
Das Unternehmen sollte seine Finanzberichtsverfahren überprüfen, um eine genaue Meldung von Einkommensteuer und Umsatzsteuer in der Zukunft zu gewährleisten.

Schlussfolgerung

Die Prüfung hat bestimmte Bereiche von Bedenken aufgedeckt, die sofortige Aufmerksamkeit benötigen, um die Einhaltung der Vorschriften und Genauigkeit bei der Steuerberichterstattung sicherzustellen.

Steuerprüfungsbericht 1

Chatbasierte Interaktionen für Kunden und interne Anfragen (generative KI + KI-Empfehlungssystem)



How often do you perform this task ?

Which language do you prefer?

Do you require assistance on constant basis?

○ No
Yes, for:
○ 1 - 5 hours per week
○ up to 20 hours per week
○ up to 40 hours per week
○ I am not sure

Ich benötige einen Experten für Infrastruktur und KI, um bei der Einrichtung von Software für das Managementteam zu unterstützen.

Hier sind die Berater, die optimal auf Ihren spezifischen Fall abgestimmt sind. Entdecken Sie Ihre perfekt passenden IT-Experten

Marcus Brose
Freelance Machine Learning Engineer
United States
Machine Learning, Android, Linux RHEL/CentOS, Algorithms, Artificial Intelligence (AI)

Isabella Nguyen
Freelance Machine Learning Engineer
Germany
RESTful Web Services, Linux RHEL/CentOS, React, Algorithms, Redux, Artificial Intelligence (AI)

Ausgestattet mit einem scharfen Auge für Details, verfügt Herr Brose über Fachkenntnisse in der Identifizierung von Unstimmigkeiten in Steuerberichten. Wenn es eine Inkonsistenz gibt, wird er sie entdecken.

Seine analytischen Fähigkeiten sind beispielhaft beim Aufdecken von Inkonsistenzen in Datenmustern zu finden und Systemarchitekturen zu optimieren

Dies sind nur zwei Beispiele aus mehreren Möglichkeiten

Beispiel 2: virtuellen Assistent

Optimierung von mehrstufigen Arbeitsabläufen mit einem virtuellen Assistenten/Chatbot, der auf unternehmensspezifische Daten trainiert ist. Zusätzlich wird der Assistent über ein Empfehlungssystem verfügen. Funktionalitäten umfassen:

Vorschlagen relevanter Produkte und Dienstleistungen entsprechend einer bestimmten Kundenanfrage

Effiziente Suche nach fallspezifischer Übereinstimmung mit Inhaltszusammenfassung

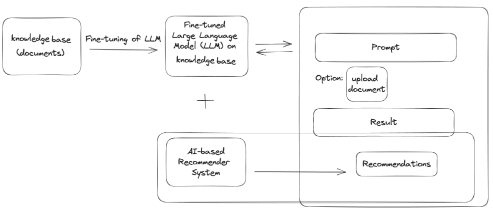
Beantwortung von Anfragen durch Extraktion von Wissen aus dem Unternehmenskontext (statt über Schlagworte oder Filter)

Finden von fallrelevanten Informationen mit Verweis auf Quelldokumente

Alle oben genannten Funktionen können sowohl für kundenorientierte als auch für interne Wissensmanagement-Szenarien entwickelt werden.

End-to-End-Prozesslandkarte des KI-Systems

Nachfolgend haben wir den Gedanken einer high-level Prozesslandkarte visualisiert, die für die Ableitung einer KI-Architektur nützlich sein kann


knowledge base (documents)
Fine-tuning of LLM
Fine-tuned Large Language Model (LLM) on knowledge base
Prompt
Option: upload document
Result
+
AI-based Recommender System
Recommendations

MÖGLICHE ZUKÜNFTIGE SCHRITTE

Generative KI bietet verschiedene Möglichkeiten zur Verbesserung von Arbeitsabläufen und Produkten. Um diese Möglichkeiten voll auszuschöpfen, wäre es von Vorteil, zunächst die folgenden Schlüsselfragen zu untersuchen, bevor eine bestimmte Anwendung implementiert wird.

Erstens: Welches ist die gesamte Bandbreite der Anwendungsfälle für die Anwendung?

Zweitens: Welche Ressourcen werden für die Umsetzung benötigt?

Drittens: Wie können wir den Nutzen der Anwendung maximieren?

Um diese Fragen zu beantworten, schlagen wir die folgenden Schritte vor:

Discovery Workshops organisieren, um ...

1 ... die verschiedenen internen Arbeitsabläufe und kundenorientierten Dienstleistungen des Unternehmens abzubilden.

2 ... Prozesse innerhalb eines Unternehmens zu identifizieren, die mit KI & generativer KI verbessert werden können.

3 ... KI-Szenarien zu identifizieren, die die größte Auswirkung auf ein Unternehmen mit der höchsten Anzahl von Anwendungsfällen haben würden.

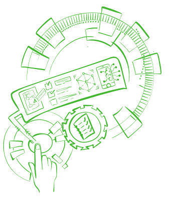
4 ... die Ressourcenzuweisung gegenüber den Auswirkungen abzuwägen und die zu implementierenden Anwendungen.

5 ... potenzielle Risiken und Maßnahmen zur Risikominderung zu identifizieren.

Wenn die Landkarte priorisiert ist, fahren wir fort mit ...

6 ... der Entwicklung und Implementierung eines Prototyps - unter Verwendung des QM-eigenen KI-Clusters oder der firmeneigenen KI-Infrastruktur - mit kontinuierlichem Feedback der relevanten Stakeholder.

7 ... der weiteren Entscheidungen über Tests und dem Einsatz.



qualityminds

Fazit

Die Nutzung von KI und generativen KI-Funktionen kann für Unternehmen eine Vielzahl von Vorteilen mit sich bringen. Wir haben einen Einblick in die Möglichkeiten gegeben, die sich für beschriebene Fälle ergeben könnten.

Unsere vorläufigen Analysen und Anwendungsbeispiele zeigen, dass die Integration intelligenter generativer KI-Anwendungen sowohl machbar als auch vielversprechend für die Zukunft sein kann.

- ✓ ***Werkzeuge zur Unterstützung der menschlichen Entscheidungsfindung***
- ✓ ***KI und generative KI ermöglichen es Ihren Mitarbeitenden, sich auf hochwertige Tätigkeiten zu konzentrieren***
- ✓ ***Automatisierung von Prozessen in verschiedenen Geschäftsbereichen***
- ✓ ***Gesteigerte Produktivität für mehrere verschiedene Kategorien***

Copyright by QualityMinds