



Künstliche Intelligenz mit SAP®

- › Der strategische Überblick über KI bei SAP
- › Von Joule bis SAP Business AI: die SAP-KI-Toollandschaft praxisnah erklärt
- › KI-Lösungen und -Funktionen verstehen und einsetzen

Kapitel 3

Embedded AI und Joule

Fertige KI-Lösungen von SAP sparen vielen Unternehmen Zeit und Aufwand. Erfahren Sie hier, wie Sie Embedded AI einsetzen können und welche Vorteile Joule bietet.

Die Integration von KI in Unternehmenssoftware hat in den letzten Jahren große Fortschritte gemacht, doch oft bleibt der Zugang zu dieser Technologie für viele Unternehmen komplex. Bisher waren spezielle Ressourcen, IT-Infrastruktur sowie Expert*innen erforderlich, um KI-Anwendungen direkt im Unternehmen zu etablieren. Durch die fortschreitende Standardisierung von Prozessen und Systemlandschaften sowie die steigende Qualität von Machine-Learning-Modellen ergibt sich für SAP die Möglichkeit, KI-Funktionen zentral zu entwickeln und für alle Kunden anzubieten.

Embedded AI und *Joule* verfolgen hier genau diesen Ansatz: KI wird nahtlos in bestehende Geschäftsprozesse eingebettet, sodass sie ohne separate Implementierung oder aufwendige Konfiguration genutzt werden kann. In diesem Kapitel werden Sie tiefere Einblicke sowohl in Embedded AI als auch in Joule erhalten. Wir werden mit den Möglichkeiten von Embedded AI starten und klären, was Sie sich darunter vorstellen können, worin die Vor- und Nachteile liegen und welche Use Cases möglich sind bzw. wo Sie sich darüber informieren können.

Danach werden Sie mehr über Joule erfahren, einerseits über die Funktionen, die geplant sind, andererseits über das Anwendungsbeispiel des Document Grounding. Am Ende dieses Kapitels werden Sie erfahren, wie Sie selbst Joule nutzen und einrichten können und wie die Bezahlstruktur aussehen wird.

Bitte beachten Sie, dass diese Themen sehr neu sind und kontinuierlicher Entwicklung unterliegen. Insbesondere was den Funktionsumfang und die Anwendungsfälle angeht, kann dieses Kapitel keinen vollständigen Überblick geben, sondern lediglich zum Verständnis darüber beitragen, wie die einzelnen Themen zusammenhängen und was wir vielleicht in Zukunft von den Tools erwarten können.

3.1 KI in SAP-Cloud-Produkten

Mit *Embedded AI* profitieren Sie von vorgefertigten KI-Funktionen, die direkt in SAP-Anwendungen integriert sind. Der Begriff *Embedded* steht hier für »eingebettet« oder »integriert« und hat nichts mit dem Thema *Embedding-Modelle* zu tun, das Sie noch näher in Kapitel 6, »Generative KI: Sprachmodelle einsetzen«, kennenlernen werden.

Die Funktionen von Embedded AI sind also direkt in SAP-Cloud-Produkte integriert, sodass Sie mit diesen Funktionen direkt in Ihrer gewohnten SAP-Umgebung interagieren können. Es muss erwähnt werden, dass Embedded AI für absolute Standardprozesse entwickelt wird. Eine Übersicht dieser Prozessschritte, die aktuell durch Embedded AI unterstützt werden, finden Sie im SAP Discovery Center unter der Kachel **SAP Business AI Features** (den Link finden Sie im Kasten »Hilfreiche Links zum Thema SAP Embedded AI«).

Diese Standardintegration ermöglicht es, dass die KI-Tools ohne großen Implementierungsaufwand genutzt werden können. Je nach Anwendungsfall muss jedoch unterschieden werden, ob eine Embedded-AI-Lösung ausreichend ist oder doch eher eine eigene KI-Extension entwickelt werden sollte (dabei bezieht sich KI-Extension auf alle im Folgenden genannten Möglichkeiten von *Custom AI* auf der SAP BTP, siehe Kapitel 4, »AI Services: KI-Lösungen für Geschäftsprozesse«, bis Kapitel 6, »Generative KI: Sprachmodelle einsetzen«). Tabelle 3.1 zeigt die Unterschiede zwischen Embedded-AI- und Custom-AI-Lösungen sowie deren Vor- und Nachteile.

	Embedded AI	Custom AI Extensions
Definition	KI ist nahtlos in SAP-Anwendungen integriert.	Unternehmen entwickeln eigene KI-Lösungen auf Basis von SAP BTP.
Eigenschaften	Nutzer*innen profitieren von fertigen KI-Funktionen innerhalb der gewohnten SAP-Umgebung.	höchste Flexibilität, aber komplexe Entwicklung nötig
Kosten und Ressourcen	kosteneffizient und sofort nutzbar, aber begrenzte Anpassungsmöglichkeiten	maximale Anpassbarkeit, aber hohe Entwicklungsressourcen erforderlich
Implementierung	schnelle Implementierung, da bereits in SAP-Umgebungen integriert	abhängig von IT-Infrastruktur und SAP-BTP-Anbindung, aber generell hoher Aufwand nötig
Vorteile	schnell in Standardprozessen einsetzbar, kosteneffizient, flexibel	spezifisch auf Use Case implementiert, anpassbar, individuell zugeschnitten
Nachteile	kaum bis nicht individualisierbar, intransparente Kosten- und Modellgestaltung	höherer Kostenaufwand für Entwicklung und Integration, Wissen und Kompetenzen erforderlich

Tabelle 3.1: Gegenüberstellung von Embedded AI vs. Custom AI

Falls Sie sich also im SAP-Standard bewegen und es einen Embedded AI Case gibt, der zu Ihrem Prozess passt, kann es sich lohnen, über die Nutzung der Embedded-AI-Funktion nachzudenken. Mit Embedded-AI-Lösungen profitieren Sie von der simplen und fertigen Integration von KI-Lösungen direkt in der Anwendung und müssen diese nicht erst selbst aufwendig entwickeln. Allerdings müssen Sie Abstriche bei der Flexibilität und Anpassbarkeit machen.

Technologisch bildet SAP Embedded AI die Brücke zwischen der AI Foundation, wo die AI Features von SAP entwickelt und bereitgestellt werden, und Joule sowie anderen SAP-Systemen, in denen die Features dann für die Nutzer*innen verfügbar gemacht werden (siehe auch [Abbildung 3.1](#)). Joule dient dabei vor allem als Schnittstelle zwischen den Nutzer*innen und den jeweiligen SAP-Anwendungen. Mittels natürlicher Sprache kann man dabei von diversen Funktionen Gebrauch machen. Für sehr spezifische Funktionen gibt es aber auch direkt in den jeweiligen SAP-Anwendungen Embedded AI Tools wie etwa im SAP Sustainability Control Tower.

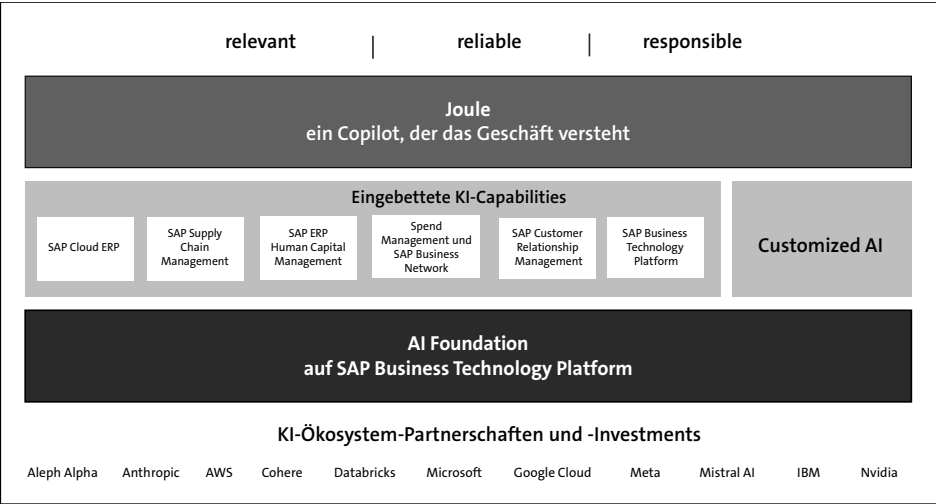


Abbildung 3.1: Embedded AI im Kontext der gesamten SAP Business AI

3.1.1 ESG Report Generation im SAP Sustainability Control Tower

Ein anschauliches Beispiel für Embedded AI in SAP-Lösungen ist die Erstellung eines ESG-Reports (*ESG Report Generation*) als Teil des *SAP Sustainability Control Tower*.

Hintergrund: SAP Sustainability Control Tower

Der SAP Sustainability Control Tower unterstützt Unternehmen dabei, ihre Nachhaltigkeitsziele zu verfolgen und ihre ESG-Daten (Environmental, Social, Governance) zentral zu verwalten und zu analysieren.

Mit dem SAP Sustainability Control Tower können Sie ESG-Berichte erstellen, die Fortschritte im Bereich Nachhaltigkeit messbar und transparent machen. Da das Erstellen solcher umfangreichen Berichte sehr zeit- und kostenintensiv ist, hilft das ESG Report Generation Tool dabei, Nachhaltigkeitsberichte automatisiert zu erstellen. Auf Basis der im SAP Sustainability Control Tower bereits erfassten ESG-Daten generiert die Lösung mithilfe von SAP Business AI vollständige Berichte, die den gängigen Berichtsstandards entsprechen. Dabei kommen vordefinierte Templates sowie generative KI zum Einsatz, um verständliche und strukturierte Texte zu erstellen. Die generierten Berichte können direkt als PDF heruntergeladen und für interne oder externe Stakeholder genutzt werden. Das ESG Report Generation Tool nutzt dabei Embedded AI, um relevante ESG-Metriken zu sammeln, Tabellen und Diagramme automatisch zu generieren und mithilfe generativer KI einen Bericht aus den gesammelten Informationen zu erstellen. Durch diese KI-gestützte Automatisierung reduziert sich der Zeitaufwand für die Berichterstellung erheblich.

SAP gibt an, dass die Zeit, die für die Erstellung von Tabellen und Grafiken benötigt wird, um 96% reduziert wird und die Zeit, die für das Verfassen und Schreiben des Reports benötigt wird, um 85% sinkt (<https://discovery-center.cloud.sap/ai-feature/c2b939fa-8655-446c-8d24-81e945c7ecd6/>). Sie profitieren von der KI-Funktion in der gewohnten SAP-Umgebung, ohne ein

weiteres Tool nutzen zu müssen. Alles passiert »embedded«, also direkt im SAP-System. Die einzige Voraussetzung ist, dass Sie über eine aktive SAP-AI-Units-Lizenz verfügen. Mehr zu den Lizenzen erfahren Sie auch in Kapitel 7, »Praktische Umsetzung«. Die Embedded-AI-Funktionen werden, genau wie auch die Capability-Typen von Joule, über *SAP AI Units* abgerechnet. Diese Units werden je nach Use Case, Nutzung und Komplexität verbraucht und in Rechnung gestellt. Eine genaue Aufstellung aller Embedded AI Features und SAP-AI-Unit-Kosten finden Sie im SAP Discovery Center.

Hilfreiche Links zum Thema SAP Embedded AI

Weitere hilfreiche Informationen zum Thema SAP Embedded AI finden Sie unter folgenden Links:

- AI Feature Catalog: <https://discovery-center.cloud.sap/ai-catalog/>
- AI Estimator: <https://discovery-center.cloud.sap/ai-estimator>
- Überblicks-Blogbeitrag über SAP Business AI: <https://community.sap.com/t5/technology-blogs-by-sap/sap-business-ai-overview-for-all/ba-p/13871365>

3.2 Joule – der KI-Copilot von SAP

Anders als bei Embedded AI Tools wie dem ESG Report Generation Tool, das direkt in die Anwendung integriert ist, bietet der generative KI-Assistent *Joule* eine allgemeinere, dialogbasierte Schnittstelle. Joule soll nach und nach in viele SAP-Anwendungen ausgerollt werden (siehe [Abbildung 3.2](#) für eine Beispielansicht). Die Idee ist dabei, dass Joule Ihr ständiger Begleiter ist, dem Sie alle möglichen Fragen stellen oder Aufgaben abgeben können, um Ihre Arbeit zu erleichtern. Joule versteht dabei Geschäftsprozesse, kann Daten interpretieren und Vorschläge zu Aktionen machen.

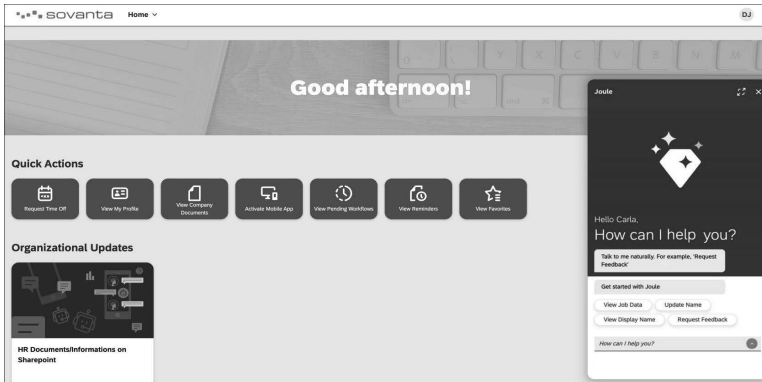


Abbildung 3.2: Joule wird in Zukunft nahtlos in SAP-Oberflächen integriert.

Da Joule in sämtliche SAP-Lösungen eingebettet ist und zusätzlich auch firmeneigene Daten mittels Document Grounding mit einbezogen werden können, werden Antworten in einem unternehmensbezogenen Kontext generiert. Dabei nutzt Joule als Kontext für das LLM sämtliche Daten der angebundenen SAP-Anwendung, die firmeninternen Dokumente (falls über Document Grounding aktiviert) sowie die Rolle und aktiven Rechte der Nutzer*innen im System, um die Anfrage der Nutzer*innen zu bearbeiten.

Welches LLM genau verwendet wird, ist allerdings nicht ganz klar, da dies auch von der Anfrage abhängig ist. Joule ist in der Lage, zwischen verschiedenen Modellen zu wählen, und entscheidet für den jeweiligen Anwendungsfall separat, welches Modell ausreichend ist und die besten Ergebnisse erzielt. Je nach Anwendungsfall nutzt Joule außerdem gezielte Fähigkeiten, auch *Capabilities* genannt, um bei verschiedenen Prozessen zu unterstützen. Diese Fähigkeiten lassen sich in drei zentrale Capability-Typen unterteilen: *Transactional*, *Navigational* und *Informational*. Diese Typen bestimmen, wie Joule mit den Nutzer*innen interagiert und welche Aufgaben innerhalb der SAP-Systeme übernommen werden.

3.2.1 Transactional Capabilities

Joule erlaubt es, über *Transactional Capabilities* direkt mit den Geschäftsdaten zu arbeiten, ohne manuell zwischen verschiedenen Anwendungen wechseln zu müssen. Anstatt sich durch mehrere komplexe Benutzerflä-

chen zu navigieren, können Nutzer*innen ihr Anliegen in natürlicher Sprache kommunizieren, und dieses wird dann von Joule übernommen. So kann Ihnen Joule z. B. in SAP SuccessFactors nicht nur generelle HR-Informationen über Sie bereitstellen, sondern auch direkt Aufgaben übernehmen. Bei einer Frage wie »Ich benötige Feedback meines Vorgesetzten« kann Joule diese Anfrage z. B. direkt zusammen mit Ihrem Anliegen an Ihren Vorgesetzten senden.

3.2.2 Navigational Capabilities

Gerade in großen Unternehmen mit vielen genutzten SAP-Anwendungen kann es herausfordernd und zeitaufwendig sein, sich durch diverse Menüs zu navigieren, um die entsprechende Funktion zu finden. Hier helfen *Navigational Capabilities*, indem sie Nutzer*innen direkt zur relevanten Funktion oder Anwendung führen. Sie können so eine Navigationsanfrage an Joule senden, wie z. B. »Wo kann ich eine Bestellung für das Material AB0123 anlegen?«. Joule wird daraufhin die relevantesten SAP-Anwendungen identifizieren und Ihnen eine Vorschlagsliste anzeigen, aus der Sie wählen können. Über den vorgeschlagenen Link können Sie anschließend direkt auf die entsprechende Seite navigieren.

3.2.3 Informational Capabilities

Neben der Unterstützung bei Transaktionen und Navigation bietet Joule auch eine *Informational Capability*. Dabei durchsucht Joule verfügbare Dokumentationen, das SAP Help Portal und eigenes Unternehmenswissen, um präzise Antworten auf Nutzeranfragen zu liefern, und durchläuft dazu die folgenden Schritte:

1. Die Nutzer*innen stellen eine Frage in natürlicher Sprache.
2. Joule durchsucht die eigene Wissensbasis, das SAP Help Portal und, falls eingerichtet, auch eigene Unternehmensdokumente.
3. Die Antwort wird als zusammengefasste Information bereitgestellt, ergänzt durch die drei relevantesten Dokumente oder Links.

Damit Joule auch unternehmenseigene Informationen nutzen kann, muss, wie bereits erwähnt, Document Grounding aktiviert sein.

3.3 Document Grounding in Joule

Da die meisten großen Sprachmodelle (auch die, die Joule im Hintergrund verwendet) nur auf allgemeine Daten trainiert sind, fehlt ihnen unternehmensspezifisches Wissen, etwa interne Prozessdokumentationen oder spezifische, industrierelevante Richtlinien. Mit *Document Grounding* können auch eigene Dokumente in Joule eingebunden werden. Auf diese Weise kann sichergestellt werden, dass unternehmensrelevante Antworten generiert werden können. Document Grounding basiert auf der Technik *Retrieval-Augmented Generation* (RAG).

Die grundlegende Idee ist dabei, die relevantesten Dokumente für die jeweilige Anfrage aus einer Datenbank zu erheben und mithilfe dieser Informationen eine Antwort zu generieren. Dies ermöglicht es Joule, z. B. unternehmensspezifisches Wissen abzufragen, ohne dass das genutzte LLM vorher auf diesen Daten trainiert wurde. Weitere Informationen zu RAG finden Sie in Kapitel 6, »Generative KI: Sprachmodelle einsetzen«.

Um Document Grounding nutzen zu können, sind einige technische und organisatorische Voraussetzungen notwendig. Zum einen muss die Joule-Integration in SAP BTP bereits aktiv sein. Außerdem benötigen Sie Zugriff auf *Microsoft SharePoint*, das als Dokumentenablage dient. In SharePoint können für Document Grounding maximal 2.000 Dokumente hochgeladen werden. Dabei werden die folgenden Dateitypen unterstützt: HTML, JPEG, JPG, Microsoft Word, PDF, PNG und TIFF. Allerdings wird aus diesen Dokumenten nur der Klartext genutzt – Tabellen und Bilder werden aktuell nicht unterstützt.

3.4 Einrichtung von Joule in der SAP BTP

Sie haben bereits gesehen, wie Joule eingesetzt werden kann, um verschiedene Arten von Fähigkeiten zu nutzen. Dazu müssen jedoch einige grundlegende Voraussetzungen erfüllt und bestimmte Konfigurationsschritte durchgeführt werden, damit Joule erfolgreich integriert werden kann. Eine Übersicht über mögliche Einsatzgebiete von Joule innerhalb des SAP-Portfolios zeigt Abbildung 3.3, etwa in Bereichen wie Personalwesen, Vertrieb oder IT.

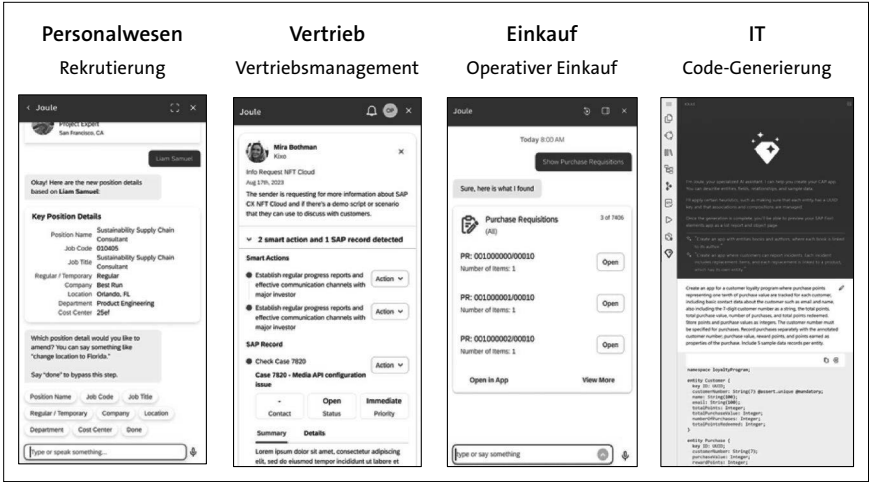


Abbildung 3.3: Einsatzgebiete von Joule im SAP-Portfolio

3.4.1 Voraussetzungen für die Nutzung von Joule

Da Joule in die jeweiligen SAP-Produkte wie SAP SuccessFactors oder SAP S/4HANA eingebettet ist, benötigen Sie zunächst eine gültige Lizenz für das jeweilige Produkt sowie einen Global Account in der SAP BTP. Da Joule auf die Cloud Identity Services setzt, muss die Integration mit dem *Identity Authentication Service* erfolgen. Für Joule benötigen Sie außerdem noch einen Subaccount in der SAP BTP. Es ist nur eine einzige Subskription nötig, um Joule in mehrere Produkte zu integrieren. Zusätzlich erfordert die Verbindung zwischen Joule und anderen SAP-Produkten eine Trust-Konfiguration, die über das SAP BTP Cockpit eingerichtet werden kann.

3.4.2 Joule-Booster starten

SAP bietet einen *Booster* an, der die Einrichtung von Joule weitgehend automatisiert. Eine detaillierte Beschreibung über die Nutzung von Boostern finden Sie in Kapitel 7, »Praktische Umsetzung«. Für den Joule Booster umfasst das Booster-Paket folgende Schritte:

1. **Prüfung der Voraussetzungen:** Der Booster validiert, ob alle notwendigen Berechtigungen und Integrationen vorhanden sind.
2. **Erstellung des Subaccounts:** Ein neuer Subaccount für Joule wird eingerichtet, falls noch keiner existiert.
3. **Auswahl der Integrationen:** Es wird festgelegt, mit welchen SAP-Produkten Joule verbunden werden soll.
4. **Zuweisung von Berechtigungen:** Die relevanten Systeme und deren Zugriffsrechte werden konfiguriert.
5. **Validierung und Abschluss:** Nach einer finalen Prüfung wird die Bereitstellung abgeschlossen, und Joule steht zur Nutzung bereit.

Hilfreiche Links zum Thema Joule

Weitere hilfreiche Informationen zu Joule finden Sie unter folgenden Links:

- Joule Onboarding Guide: <https://help.sap.com/docs/JOU-LE/6189c8655c484916bb8eb767126a653a/48e8f655470048c29555272578fbc2ce.html>
- Joule for Developers: <https://community.sap.com/t5/technology-blogs-by-sap/joule-in-sap-build-code/ba-p/13576765>
- Joule for Consultants: <https://www.sap.com/products/artificial-intelligence/ai-assistant/sap-consulting-capability.html>
- Joule Capabilities in SAP Build Process Automation: <https://community.sap.com/t5/application-development-and-automation-blog-posts/how-to-enable-joule-capabilities-in-sap-build-process-automation-and-sap/ba-p/14050593>

3.5 Zusammenfassung

Joule wird in Zukunft eine immer größere Rolle spielen, da es als zentrale Interaktionsfläche für Nutzer*innen innerhalb der SAP-Systeme ausgerollt wird. Die sprachbasierte Interaktion gewinnt zunehmend an Bedeutung, da durch die fortschreitende Kontextualisierung von Daten, Aktionen und Nutzer*innen immer mehr Prozesse automatisiert und personalisiert ablaufen können.

Gleichzeitig wird sich das Spektrum der Joule-Fähigkeiten kontinuierlich erweitern. Auch die Integration in externe Nicht-SAP-Systeme wird weiter voranschreiten. Durch die Fortschritte in der KI-Welt – von leistungsfähigeren Sprachmodellen mit besseren Reasoning-Fähigkeiten über multimodale Anwendungsfelder (Text, Ton, Bild) bis hin zu Agenten-Funktionen – wird auch Joule immer mächtiger und vielseitiger. Außerdem können Unternehmen in Zukunft selbst Skills für Joule entwickeln und integrieren, mehr dazu erfahren Sie in Kapitel 6, »Generative KI: Sprachmodelle einsetzen«.

Damit entwickelt sich Joule nicht nur zu einer zentralen Schnittstelle innerhalb der SAP-Welt, sondern auch zu einem intelligenten Begleiter, der Prozesse vereinfacht, Entscheidungen unterstützt und die Art und Weise, wie wir mit Systemen interagieren, grundlegend verändert.

Inhalt

Kapitel 1

Strategische Bedeutung von KI im Unternehmen 15

1.1	KI zwischen Hype und Realität: Wo stehen wir?	15
1.2	Die SAP Business Technology Platform als Fundament für KI	16
1.3	Die vier strategischen Zugänge zu KI im SAP-Ökosystem	18
1.3.1	Embedded AI: Intelligenz direkt im Standard	18
1.3.2	SAP AI Services: vortrainierte Modelle per Plug and Play	18
1.3.3	Generative KI: LLMs orchestrieren, verknüpfen und betreiben	19
1.3.4	Kundenspezifische KI: der Weg zur differenzierenden Intelligenz	20
1.4	Warum KI auf der SAP BTP ein strategischer Erfolgsfaktor ist	21
1.5	Herausforderungen im Unternehmen: Was KI-Initiativen heute bremst	22
1.6	Zusammenfassung	23

Kapitel 2

Kurze Einführung in künstliche Intelligenz 25

2.1	Durchblick im Begriffsdschungel: Was ist KI überhaupt?	26
2.1.1	Künstliche Intelligenz	28
2.1.2	Maschinelles Lernen	29
2.1.3	Deep Learning	34
2.2	Die neue Ära der Sprachmodelle: Was sind eigentlich LLMs?	39
2.2.1	Large Language Models (LLMs)	39
2.2.2	KI-Agenten	43
2.3	Angewandte KI: Was bedeutet Data Science?	44
2.3.1	Classification (Klassifikation)	46
2.3.2	Prediction (Vorhersage)	47
2.3.3	Clustering (Gruppierung ohne Labels)	48

2.3.4	Extraction (Merkmalsextraktion/Informationsextraktion)	49
2.3.5	Traditionelle KI vs. generative KI	50
2.4	Zusammenfassung	52

Kapitel 3

Embedded AI und Joule **55**

3.1	KI in SAP-Cloud-Produkten	56
3.1.1	ESG Report Generation im SAP Sustainability Control Tower	59
3.2	Joule – der KI-Copilot von SAP	60
3.2.1	Transactional Capabilities	61
3.2.2	Navigational Capabilities	62
3.2.3	Informational Capabilities	62
3.3	Document Grounding in Joule	63
3.4	Einrichtung von Joule in der SAP BTP	64
3.4.1	Voraussetzungen für die Nutzung von Joule	64
3.4.2	Joule-Booster starten	65
3.5	Zusammenfassung	66

Kapitel 4

AI Services: KI-Lösungen für Geschäftsprozesse **69**

4.1	SAP Document AI	70
4.1.1	Base Edition	71
4.1.2	Premium Edition	72
4.1.3	Nutzung und erste Schritte	74
4.1.4	Einsatzempfehlungen	75
4.2	Business Entity Recognition Service	76
4.2.1	Vortrainiertes Modell	77
4.2.2	Spezifisches Modell	78
4.2.3	Nutzung und erste Schritte	78
4.2.4	Pakete und Abrechnung	79
4.2.5	Einsatzempfehlungen	80
4.3	Data Attribute Recommendation Service	80
4.3.1	Funktionen	82

4.3.2	Nutzung und erste Schritte	82
4.3.3	Pakete und Abrechnung	83
4.3.4	Einsatzempfehlungen	84
4.4	Personalized Recommendation Service	84
4.4.1	Funktionen	85
4.4.2	Nutzung und erste Schritte	87
4.4.3	Pakete und Abrechnung	88
4.4.4	Einsatzempfehlungen	89
4.5	SAP Translation Hub	89
4.5.1	Funktionen	90
4.5.2	Nutzung und erste Schritte	91
4.5.3	Pakete und Abrechnung	92
4.5.4	Einsatzempfehlungen	93
4.6	Referenzarchitektur	93
4.7	Zusammenfassung	95

Kapitel 5

Custom AI: kundenspezifische KI entwickeln **97**

5.1	Anwendungsszenarien im SAP-Kontext	98
5.1.1	Image Recognition: In Bildern steckt oft die Wahrheit	99
5.1.2	Time Series Analysis: ein Blick in die Zukunft	99
5.1.3	Natural Language Processing: das Spiel mit der Sprache	100
5.1.4	Weitere spezifische ML-Anwendungsfälle	100
5.2	Was benötigt man für Custom AI?	101
5.2.1	Wozu werden Metriken benötigt?	101
5.2.2	Ist ein Goldstandard erforderlich?	102
5.2.3	Ist ein Machine-Learning-Modell die beste Wahl?	103
5.2.4	Gibt es vortrainierte Modelle für Ihren Anwendungsfall?	104
5.3	Anforderungen an eine Arbeitsumgebung für Custom AI	105
5.3.1	Ort für Experimente	105
5.3.2	Versionskontrolle	106

- 5.3.3 Populäre Programmiersprachen 107
 - 5.3.4 Arbeitsplattformen 108
- 5.4 Custom AI auf der SAP BTP mit dem SAP AI Launchpad und SAP AI Core 110
 - 5.4.1 Workspaces in SAP AI Core und Übersicht der SAP-AI-Launchpad-Infrastruktur 111
 - 5.4.2 Ansteuerung des SAP AI Core 112
 - 5.4.3 Mögliche Benutzergruppen 113
 - 5.4.4 Best Practices für die Arbeit mit SAP AI Core 114
 - 5.4.5 Training auf SAP AI Core 116
 - 5.4.6 Deployment auf SAP AI Core 120
 - 5.4.7 Technischer Zugriff auf SAP AI Core über die API 122
 - 5.4.8 Rechenkapazitäten und Kosten für SAP AI Core 124
- 5.5 Custom AI auf einem Kyma-Cluster auf der SAP BTP 127
- 5.6 Cloud Foundry auf der SAP BTP 128
- 5.7 Beispiel-Set-up für die Nutzung einer Custom AI 131
 - 5.7.1 Kontext des Beispiels 131
 - 5.7.2 Beispielarchitektur 132
- 5.8 Zusammenfassung 135
- 5.9 Anhang 136
 - 5.9.1 Workflow Template 136
 - 5.9.2 Serving Template 138

Kapitel 6

Generative KI: Sprachmodelle einsetzen 141

- 6.1 Large Language Models und ihre Unzulänglichkeiten 142
 - 6.1.1 Few- und Zero-Shot Prompting 143
 - 6.1.2 Feintuning 144
 - 6.1.3 Retrieval-Augmented Generation 145
 - 6.1.4 Knowledge Graph Retrieval 148
 - 6.1.5 Function Calling und Joule als digitaler Assistent 149
 - 6.1.6 KI-Agenten 151
 - 6.1.7 Zusammenfassung und Vergleich 152

6.2	Generative AI Hub: Einsatz von LLMs und Embedding-Modellen	155
6.2.1	Generative AI Hub und AI Core	155
6.2.2	SAP AI Launchpad	157
6.2.3	Verfügbare LLMs	160
6.2.4	Embedding-Modelle	162
6.2.5	Python SDK	164
6.3	SAP HANA Cloud Vector Engine und Knowledge Graph DB	166
6.3.1	SAP HANA Built-in Embeddings	167
6.3.2	Effizientere Suche mit HNSW-Index	168
6.3.3	SAP HANA und Knowledge Graphs	169
6.4	Joule Studio	170
6.4.1	Tools	172
6.4.2	Best Practices	173
6.4.3	Beispielanwendungen	174
6.5	Referenzarchitektur	175
6.6	Zusammenfassung	177

Kapitel 7

Praktische Umsetzung 179

7.1	Kriterien eines guten Use Cases	180
7.1.1	Daten	180
7.1.2	Ziel	181
7.1.3	Prozess	181
7.1.4	Aufgabe	182
7.2	Datenaufbereitung auf der SAP BTP	184
7.3	Arbeitsmethoden	187
7.3.1	Machine Learning Lifecycle	187
7.3.2	Rollen	190
7.4	Konkrete Implementierungsbeispiele	192
7.4.1	Global Account und Subaccounts	192
7.4.2	Booster	193
7.4.3	SAP AI Services: Rechnungen verarbeiten	196
7.4.4	Custom AI: Zahlungsdatum vorhersagen mit SAP Business Data Cloud und Databricks oder Cloud Foundry	198

- 7.4.5 Custom AI: Unterschriften erkennen mit SAP AI Core ... 200
 - 7.4.6 Generative KI: Mitarbeiter-Onboarding 201
- 7.5 Bezahlmodelle 203
 - 7.5.1 Credits 203
 - 7.5.2 AI Units 208
 - 7.5.3 Berechnungshilfen 210
- 7.6 Ethische Aspekte von KI 212
 - 7.6.1 Bias 215
 - 7.6.2 Energieverbrauch 217
 - 7.6.3 Leitlinien 218
- 7.7 Governance und Best Practices 218
- 7.8 Zusammenfassung 222

Kapitel 8

Ausblick: Wohin geht die Reise mit KI auf der SAP BTP? 225

- 8.1 Von Funktionen zu Erlebnissen: Joule als zentrales Interface 225
- 8.2 KI-Agenten: vom Use Case zur intelligenten Prozessinstanz 227
- 8.3 SAP Business Data Cloud: die Datenbasis für intelligente Systeme 228
- 8.4 Zusammenführung: eine neue Generation von Unternehmenssoftware 229
- 8.5 Was Sie in Ihrem Unternehmen jetzt tun sollten 230
 - 8.5.1 Erstens: Schaffen Sie technologische Bereitschaft 230
 - 8.5.2 Zweitens: Kompetenzen und Governance etablieren 231
 - 8.5.3 Drittens: Pilotprojekte mit Agenten und generativer KI initiieren 231
- 8.6 Zusammenfassung 231

- Das Autorenteam 233

- Index 235

Index

A

agentenbasierte Architektur	21
AI Catalog	208
AI Feature Estimator	208
AI Unit	208
ANN → neuronales Netz	
Artificial Intelligence → KI	
Artificial Neural Network	
→ neuronales Netz	

B

Bias	215
Big Data	45
Blackbox	36
Booster	193, 65
Buildpack	128
Business AI	18
Business Entity Recognition	76
Abrechnung	79
Einsatzempfehlung	80
erste Schritte	78
Business Process Consultant	192

C

Capability	61
ChatGPT	40
Chunk	145
Chunks	146
Classification → Klassifikation	
Cloud Foundry	197, 199, 128
Clustering	48
Cold Start	88
Consumption-based Commercial	
Model	204
Context Window	144

Credit	203
CRISP-ML(Q)	187
Custom AI	198, 98, 50, 97
Arbeitsplattformen	108
Arbeitsumgebung	105
Programmiersprache	107
Versionskontrolle	106

D

Data Attribute Recommendation	80
Abrechnung	83
Einsatzempfehlung	84
erste Schritte	82
Data Attribute Recommendation	
Service	81
Databricks	109
Data Drift	103, 131
Data Engineer	191
Data Leakage	102
Data Mining	46
Data Preparation	188
Data Science	44
Data Scientist	190, 44
Datenschutz	214
Deep Learning	27, 34
Deployment	189
Docker-Datei	200
Document Grounding	63

E

Early Stopping	173
Effort-Effect-Matrix	183
Embedded AI	18, 55, 56
Embedding	145, 146
Embedding-Modell	39

Energieverbrauch 217
ESG Report Generation 208
Estimator 80
EU AI Act 214
Evaluation 189
Explainability 87
Extraction 49

F

Feature Engineering 188
Feintuning 104, 144
Few-Shot Prompting 143
Foundation Model 42
Function Calling 149

G

Generative AI Developer 191
Generative AI Hub 19, 155
Generative AI hub in SAP AI Core
 Calculator 211
generative KI 201, 50, 39
Git 107
Goldstandard 103

H

Halluzination 142
Hierarchical Navigable Small World 168
Hyperscaler 19

I

Image Recognition 99
Inference 189
Informational Capability 62

J

Joule 17, 60, 55, 150
 Booster 65
 einrichten 64
 for Consultants 226

Voraussetzungen 64
Joule Skill 172
Joule Studio 170
Jupyter Notebook 108

K

KI 28
KI → künstliche Intelligenz
KI-Agent 43, 170
 Beispielanwendungen 174
 Tools 172
KI-Ethikrat 215
KI-Governance 219
KI-Leitlinie 218
KI-Literacy 221
Klassifikation 46
Knowledge Graph 148, 169
Konfiguration 117
Kubernetes 127
künstliche Intelligenz 15
 kundenspezifische 20
Kyma 199, 127

L

Labeling 188
LangChain 154
Large Language Model 39, 72, 142
LlamaIndex 154
LLM 19
LLM → Large Language Model

M

Machine Learning 29
maschinelles Lernen → Machine
 Learning
Master Data Management 81
Metrik 181, 101
ML → Machine Learning
ML Engineer 191

MLOps 189
 Modelling 188
 Monitoring 189

N

Natural Language Processing ... 39, 100
 Navigational Capability 62
 Neuron 35
 neuronales Netz 34
 NLP → Natural Language
 Processing 39, 100

O

Object Store 200, 175
 OCR → Optical Character
 Recognition
 ONNX 129
 ONNX → Open Neural Network
 Exchange
 Open Neural Network Exchange ... 129
 Open Source 104
 Optical Character Recognition 80,
 202, 71
 Orchestration Service 156

P

Pay as you go 204
 Personalized Recommendation 85
 Abrechnung 88
 Einsatzempfehlung 89
 erste Schritte 87
 Prediction 47
 Predictive Maintenance 100
 Prompt Engineer 192
 Proxy 101
 Python 107

R

RAG → Retrieval-Augmented
 Generation
 Reasoning 151
 Reinforcement Learning 32
 Retrieval-Augmented Generation ... 19,
 201, 63, 175, 145
 RRR-Prinzip 213

S

SAP AI Core 200, 110, 155
 Kosten 124
 SAP AI Core Standard Cost
 Calculator 211, 124
 SAP AI Launchpad 157
 SAP AI Services 18, 196
 SAP BTP 16
 SAP BTPEA 204
 SAP Build 226
 SAP Build Process Automation 197
 SAP Business AI 69
 SAP Business Data Cloud 228, 109
 SAP Business Technology Platform
 → SAP BTP 16
 SAP CPEA 204
 SAP Data Intelligence 109
 SAP Document AI 197, 70, 176
 Einsatzempfehlung 75
 erste Schritte 74
 SAP HANA Cloud Knowledge Graph
 Engine 148
 SAP HANA Cloud Vector Engine 19, 167
 SAP Sustainability Control Tower 59
 SAP Translation Hub 89
 Abrechnung 92
 Einsatzempfehlung 93
 erste Schritte 91