

KI-Optimierung für alle: So profitiert das ganze Unternehmen

5.10.2026 - Florian Kunzke | SAP SE

Es gibt eine klassische Harvard-Fallstudie, die folgende Situation beschreibt. Der Trainer einer Rudergruppe wählt seine besten acht Ruderer für das Spitzenteam und seine schwächsten acht für das Nachwuchsteam. Entgegen aller Erwartungen verliert das Spitzenteam, das aus den schnellsten und stärksten Ruderern besteht, regelmäßig gegen das Nachwuchsteam.

Die Ruderer des Spitzenteams konzentrierten sich nur darauf, ihre individuelle Leistung zu maximieren. Wenn das Boot langsamer wurde, ruderten sie isoliert kräftiger, störten damit den synchronen Rhythmus und erzeugten Wasserwiderstand. Die Ruderer des Nachwuchsteams wussten, dass sie als einzelne Personen schwächer waren. Und so ruderten sie gleichmäßig.

Unternehmen sind mit unzähligen Varianten dieses Problems konfrontiert: Sie führen Systeme ein, die die Produktivität auf Einzel- oder Teamebene maximieren, doch die Arbeit im gesamten Unternehmen wird dadurch behindert. Viele Unternehmen erleben heute etwas Ähnliches mit KI.

KI und Suboptimierung

Suboptimierung ist ein systemisches Versagen, das auftritt, wenn die Leistung eines bestimmten Teils des Systems maximiert und dabei die Leistung des Gesamtsystems beeinträchtigt wird. Eine zentrale Rolle dabei spielen drei miteinander verflochtene Themen: **Intensität, Kontext und Prognose**. Anhand dieser Faktoren lässt sich erklären, wie KI im Unternehmen zu einer Suboptimierung führen kann, nämlich indem sie Einzelpersonen und lokale Systeme stärkt, dabei aber die Gesamtorganisation belastet.

Aktuelle Forschungsergebnisse belegen die Diskrepanz zwischen einer KI-Einführung auf individueller Ebene und auf Unternehmensebene und dem Mehrwert, den sie tatsächlich bietet. Der McKinsey-Bericht [State of AI](#) zeigt, dass 89 Prozent der Unternehmen KI bereits in mindestens einem Geschäftsbereich einsetzen. Doch nur 37 Prozent berichten von einem messbaren Einfluss auf das Ergebnis vor Zinsen und Steuern (EBIT) auf Unternehmensebene. Noch ernüchternder: Lediglich sechs Prozent der Unternehmen können als leistungsstarke Unternehmen (High Performer) eingestuft werden, die bereits im gesamten Unternehmen bedeutenden Nutzen aus der Nutzung von KI ziehen.

Was macht es so schwierig, von der KI-Einführung zu einer messbaren Wertschöpfung zu gelangen? Ich glaube, es gibt drei Faktoren, die beeinflussen, ob ein Unternehmen KI zum Nutzen des gesamten Betriebs einsetzen kann.

Intensität

KI-Tools reduzieren die Arbeit meist nicht – sie intensivieren sie. Eine [Studie der University of California](#) in Berkeley ergab, dass Beschäftigte, die KI intensiv nutzen, schneller arbeiteten, ein breiteres Aufgabenspektrum übernahmen und mehr Arbeitsstunden leisteten – oft ohne explizite Aufforderung. Was kurzfristig wie höhere Produktivität wirkt, ist in Wirklichkeit schleichende Arbeitsverdichtung und wachsender Druck: Mitarbeitende müssen neue KI-Workflows verwalten und

mit weniger mehr leisten. [Eine weitere Studie zeigte](#), dass Überwachung beim Einsatz von KI mental besonders herausfordernd ist. KI-Tools, die direkte Beaufsichtigung benötigen, verschärfen das Problem der Informationsflut im Arbeitsalltag. Viele Beschäftigte fühlen sich dadurch zunehmend überfordert.

Es ist leicht nachzuvollziehen, warum KI als intensiv empfunden werden kann: Aufgaben, die früher Tage erforderten, lassen sich heute nahezu sofort per Prompt erledigen. Dies führt dazu, dass Mitarbeitende mehr Projekte in Angriff nehmen, mehr Analysen erstellen und mehr Berichte verfassen. Doch wie eine achtspurige Autobahn, die sich plötzlich an einer Mautstelle zu einer einspurigen Straße verengt, müssen Personen all die Informationen, die ausgegeben wurden, noch verarbeiten. Auf Unternehmensebene verstärkt sich dieser Engpass. Alle Mitarbeitenden leisten mehr denn je, wodurch sowohl die einzelnen Personen als auch das Unternehmen als Ganzes Schwierigkeiten haben, Schritt zu halten. Die Erstellung von Inhalten wurde skaliert, die Verarbeitung der Inhalte jedoch nicht.

Die Lösung liegt nicht zwingend darin, weniger KI einzusetzen, sondern darin, wo und wie KI zum Einsatz kommt. Die KI sollte die Absicht der Anwender verstehen und die Informationen zu Tage fördern, die zur Beantwortung der Frage notwendig sind, anstatt nur statische Antworten zu liefern. Außerdem sollte sie sicherstellen, dass Anwender nicht zwischen verschiedenen Apps und Systemen wechseln müssen.

Wenn in der Antwort auf eine Frage zu viele Informationen erzeugt werden, wird die Aufnahmefähigkeit des Nutzers herausgefordert. Die Antwort von SAP ist [Joule Work](#), ein zentraler Arbeitsbereich für SAP- und Drittanbietersysteme, der KI-Agenten einsetzt, die Aufgaben eigenständig erledigen.

Stellt man die Frage „In welchen Filialen gehen in den nächsten 72 Stunden die 65-Zoll-Fernseher aus, und wo gibt es Bestände, die ich umlagern kann?“, ruft das System systemübergreifend Daten ab und beauftragt Agenten, im Auftrag des Nutzenden zu handeln. Hier setzt SAP auf autonomes Handeln und eine individuelle Benutzererfahrung, die speziell auf diese Situation zugeschnitten ist, ohne dass weitere Inhalte verarbeitet werden müssen. Wenn Mitarbeitende nicht mehr zwischen Systemen hin- und herspringen und nicht mehr so viele Inhalte konsumieren müssen, bleibt ihnen mehr Zeit, ihr Urteilsvermögen einzusetzen und Maßnahmen zu ergreifen, die wirklich wichtig sind. So kann die KI dazu beitragen, dass die Arbeitsintensität verringert wird.

Kontext

Die meisten KI-Systeme verstehen die Welt – aber nicht das Unternehmen, in dem sie eingesetzt werden. Es gibt einen Unterschied zwischen einem System of Record (Transaktionen, Stammdaten, Prozesslogik) und implizitem Wissen (E-Mails, Chats, ungeschriebene Regeln). Und in Unternehmen findet man beides. Wenn die KI nur das System of Record sieht, können ihre Antworten technisch korrekt, aber je nach Kontext inhaltlich falsch sein, weil sie die Praxis nicht widerspiegeln, die im Unternehmen gelebt wird.

Auch bei der Beantwortung selbst scheinbar einfacher Fragen spielt der Unternehmenskontext eine entscheidende Rolle. Die Frage „Von welchen Lieferanten kann ich Kokosnüsse beziehen?“ erfordert Kenntnis der Prozesslandschaft eines Unternehmens, vor allem in den Bereichen Beschaffung, Lieferkette, Compliance und Finanzwesen. Dieses Unternehmenswissen ist meist über Prozessmodelle, Richtlinien, Chats, Tabellenkalkulationen und Anwendungen verteilt und daher schwer aktuell zu halten. Selbst wenn die KI-Agenten dieses Wissen finden, können sie es ohne das Prozesswissen der beteiligten Personen nicht in Handlungen umsetzen. Denn es sind die ungeschriebenen Regeln, Entscheidungen und Schritte, die einen Prozess ausführbar machen.

[SAP Company Memory](#) – derzeit als Preview verfügbar – erfasst kontinuierlich internes Know-how und macht es sowohl für Menschen als auch für Agenten nutzbar. Schriftliche Eingaben aus E-Mails, Chats, Prozesswissen, Richtlinien und Anwendungslogik werden als Wissensgrundlage in wiederverwendbaren Bausteinen hinterlegt, die dann von KI-Agenten genutzt werden können. Die Bausteine werden einmalig zusammengesetzt, zentral verwaltet und im gesamten Unternehmen wiederverwendet. Wenn jemand Joule Work nach Kokosnüssen fragt, basiert die Antwort auf dem Company Memory und spiegelt die tatsächlichen Regeln wider, auf die sich die Prozessverantwortlichen geeinigt haben – zum Beispiel: „Nur aus Brasilien beziehen; andere Quellen erfordern eine formelle Ausnahmegenehmigung.“

SAP Company Memory ist keine einmalige Implementierung, sondern ein kontinuierlicher Prozess. Es handelt sich um eine Wissensschicht, die als Infrastruktur für KI-Agenten dient. Durch dieses Wissen wird sichergestellt, dass die Agenten nicht nur richtig, sondern auch kontextbezogen handeln, selbst dann, wenn Richtlinien und Teams sich ändern.

Prognose

Geschäftsentscheidungen sind im Grunde Vorhersageprobleme, die auf strukturierten Daten beruhen. Die meisten Unternehmen verwenden Large Language Models (LLMs), die sich hervorragend für die Verarbeitung von unstrukturierten Daten wie Text eignen. Aus architektonischen Gründen sind diese Modelle jedoch nicht so gut dafür geeignet, strukturierte numerische Daten, die fundierte Vorhersagen ermöglichen, zu generieren und zu verarbeiten. Lieferverzögerungen vorhersagen, Prognosen erstellen, Anomalien erkennen, Lagerbestände optimieren und Kreditrisiken bewerten, all dies sind alltägliche betriebliche Fragen, für die strukturierte, tabellarische Daten und eine vorausschauende Beurteilung maßgebend sind.

LLMs dafür zu nutzen, zuverlässige Vorhersagen aus Tabellendaten zu berechnen, ist schlicht der falsche Ansatz. Die herkömmlichen Verfahren des maschinellen Lernens sind für viele zeitkritische Fragestellungen zu langsam. Die Daten werden zunächst extrahiert und an Spezialisten gesendet. Bis die Antwort vorliegt, vergehen mehrere Wochen, und die ursprüngliche Frage hat sich bis dahin oft längst geändert. Dies bedeutet, Vorhersageanalysen werden entweder von spezialisierten Fachleuten erstellt oder von generischen LLMs, die aber unzuverlässige Ergebnisse liefern. In beiden Fällen kommt es zu Verzögerungen bei wichtigen betrieblichen Entscheidungen.

Verlässliche Prognosen und Risikobewertungen sollten integraler Bestandteil des Systems sein und nicht individuelle Ad-hoc-Lösungen. SAP-RPT-1.5 und TabPFN 3 sind speziell für tabellarische Daten optimierte Modelle, die in Joule Work und SAP integriert werden, um zukunftsgerichtete Fragen sofort zu beantworten – direkt auf den aktuellen Daten, ohne Verzögerung.

SAP-RPT-1.5 für SAP-Daten und TabPFN 3 für tabellarische Daten sind spezielle Prognose-Engines für strukturierte Daten. Entscheidungsträger, die in den Kernsystemen arbeiten, haben die Möglichkeit zu fragen: „Soll ich das Volumen umleiten? Wie hoch ist die Wahrscheinlichkeit, dass die Lieferung pünktlich erfolgen wird? Wie groß ist der Kostenunterschied zwischen den unterschiedlichen Szenarien?“ Und die Antworten, die sie erhalten, basieren auf echten Unternehmensdaten.

Durch die unternehmensweite Verfügbarkeit lassen sich Engpässe bei Spezialisten vermeiden. Und das Unternehmen ist dank fundierter Prognosen besser in der Lage, mit Unsicherheiten umzugehen. So ist die Geschäftsführung in der Lage, informierte Entscheidungen zu treffen und die Probleme wirklich anzugehen.

KI zum Nutzen des gesamten Unternehmens

Es wurde bereits bewiesen, dass KI das Potenzial des Einzelnen steigern kann, aber der Gesamtorganisation nicht automatisch zugutekommt. Bei der Nutzung von KI sollte es nicht darum gehen, einzelne Aufgaben zu verbessern. Es sollte vielmehr darum gehen, wie Arbeitsprozesse, Wissen und Entscheidungen in einem Unternehmen besser miteinander verbunden werden können.

SAP zeigt mit drei Beispielen – Joule Work, SAP Company Memory und SAP-RPT-1.5/TabPFN 3 – wie das funktioniert: Diese Tools bieten eine einheitliche Bedienoberfläche, hausinterne Dokumentation und eine Prognose-Engine für strukturierte Geschäftsdaten. Dadurch wird KI von isolierten Einzellösungen zu einem Werkzeug, das dem ganzen Unternehmen nutzt.

Diese KI-Systeme schließen die Lücke zwischen einzelnen Mitarbeitenden und dem Gesamtunternehmen. Sie ermöglichen es, KI nicht einfach nur einzuführen, sondern sie so einzusetzen, dass das gesamte Unternehmen davon profitiert und Mehrwert entsteht.

<https://news.sap.com/germany/2026/10/ki-optimierung>