ENTERPRISE ARTIFICIAL INTELLIGENCE

Aufbau einer vertrauenswürdigen KI in der Souveränen Cloud

Shannon Bell
David Fraser
Tom Jenkins

Enterprise Artificial Intelligence

Aufbau einer vertrauenswürdigen KI in der Souveränen Cloud

Von Shannon Bell, David Fraser und Tom Jenkins

Erstveröffentlichung, November 2025

Herausgegeben von

Open Text Corporation

275 Frank Tompa Drive

Waterloo, Ontario, Kanada

N2L 0A1

(519) 888-7111

info@opentext.com | www.opentext.com

Vorwort

James Arroyo OBE
Ehemaliger Direktor für Daten im britischen Außen- und Commonwealth-Ministerium
Büro und Direktor der Ditchley Foundation*

Die Welt befindet sich an einem entscheidenden Wendepunkt, an dem technologische Systeme zunehmend zusammenwachsen und grundlegende Fragen nach Macht, Vertrauen, Souveränität und Governance aufwerfen. Künstliche Intelligenz verändert, wie Entscheidungen getroffen werden, wie Organisationen agieren und wie Gesellschaften funktionieren. Der Fortschritt im Bereich der KI lässt sich jedoch nicht allein an technologischen Fähigkeiten messen. Er muss mit politischer Klarheit, ethischer Disziplin und einem erneuerten Verständnis dessen einhergehen, was Souveränität im digitalen Zeitalter bedeutet.

Einführung: *Künstliche Intelligenz in Unternehmen: Aufbau einer vertrauenswürdigen KI in der Souveränen Cloud* Dieses Buch behandelt sowohl die Herausforderungen als auch die Chancen, die mit dem Einsatz künstlicher Intelligenz verbunden sind. Es betont, dass das kommende Jahrzehnt nicht allein davon bestimmt sein wird, wer die größten Modelle entwickelt, sondern vielmehr davon, wer Daten am effektivsten verwaltet und nutzt. Daten sind zugleich Treibstoff und Grundlage moderner Informationssysteme. Wie Energie oder Währung erfordern sie Regulierung, verantwortungsvolle Verwaltung und vor allem Vertrauen. Die Zukunft der KI hängt davon ab, wie mit der Vertraulichkeit, Herkunft und Souveränität dieser Daten umgegangen wird. Dies wird entscheidend sein, wenn Einzelpersonen, Institutionen und Unternehmen beginnen, ihre persönlichen, firmeneigenen und souveränen Daten in KI-Systeme einzubringen.

Während meiner Zeit als Direktor für Data im britischen Außen- und Commonwealth-Ministerium bestand die Aufgabe nicht nur darin, die Institution zu digitalisieren, sondern auch Richtlinien zu gestalten, die die digitale Transformation über Jahrzehnte hinweg tragfähig machen. Wir mussten schwierige Fragen stellen: *Wem gehören die Daten, auf die wir uns verlassen? Wo befinden sie sich? Wie werden sie weitergegeben, gespeichert und gesichert? Wie lange sollten sie behalten werden?* Mit der zunehmenden Autonomie und Allgegenwärtigkeit von KI-Systemen stellen sich dieselben grundlegenden Fragen heute jedem Unternehmen, jeder Regierung und jedem einzelnen Menschen. Die Fähigkeit von KI, Schlussfolgerungen zu ziehen, bedeutet, dass jedes Datenelement – in Kombination mit anderen – potenziell neue Erkenntnisse liefert. Das verändert alles.

Im Austausch mit führenden Vertreterinnen und Vertretern aus Technologie, Politik und Zivilgesellschaft zeigt sich, dass **souveräne Datenpolitik keine rein technische Debatte ist, sondern eine Frage nationaler und wirtschaftlicher Sicherheit**. In einer Welt, in der die meisten Daten hinter organisatorischen oder staatlichen Firewalls verschlossen sind, reicht Souveränität weit über reine Compliance hinaus. Sie umfasst Kontrolle, Verantwortlichkeit und die Fähigkeit, in einer von algorithmischen Entscheidungen geprägten Welt eigenständig und selbstbewusst zu handeln.

* Die Ansichten von James Arroyo spiegeln nicht unbedingt die der Ditchley Foundation wider.

Dieses Buch liefert überzeugende Argumente für diese Philosophie. Es zeigt, dass vertrauenswürdige Daten und verantwortungsvolle KI untrennbar miteinander verbunden sind: Um eine faire, erklärbare und sichere KI zu entwickeln, müssen die zugrunde liegenden Daten sorgfältig verwaltet, kontextbezogen interpretiert und souverän kontrolliert werden. Ohne Datenschutz und verantwortungsvolle Datenverwaltung riskiert KI, genau die Institutionen zu schwächen, die sie stärken soll. Ohne Innovation hingegen droht Regierungsführung zu einem Hemmnis zu werden, das Fortschritt behindert. Die zentrale Aufgabe besteht darin, Vertrauen und Innovation so in Einklang zu bringen, dass sie sich gegenseitig stärken.

Das Buch zeigt, wie die Landschaft des Enterprise-Computing grundlegend von innen heraus umgestaltet wird. Hyperscale-Infrastrukturen, souveräne Cloud-Lösungen und KI-gestützte Systeme bilden das Rückgrat des industriellen Netzes der digitalen Wirtschaft. Den größten Vorteil werden jedoch jene Akteure haben, die Daten nicht als neutrale Ressource, sondern als konstitutives Prinzip begreifen – als etwas, das geschützt, respektiert und verantwortungsvoll eingesetzt werden muss.

Der darin enthaltene Aufruf zum Handeln ist klar und dringlich: Es muss schnell gehandelt und noch schneller gesteuert werden, mutig innoviert, aber stets mit Ziel und Maß – und jeder digitale Fortschritt muss das Vertrauen der Öffentlichkeit stärken, nicht untergraben. Das ist der Kern von Führung im neuen Zeitalter der künstlichen Intelligenz.

In den kommenden Jahren werden jene Nationen und Unternehmen erfolgreich sein, die KI verantwortungsvoll in ihre Gesellschaften integrieren und eine breite, tief verankerte Akzeptanz fördern – und dabei eine einfache, aber grundlegende Wahrheit erkennen: **Künstliche Intelligenz ohne vertrauenswürdige, souveräne und gut verwaltete Daten bedeutet Macht ohne Verantwortung – und womöglich ohne Legitimität.** Wenn jedoch Innovation und Governance Hand in Hand gehen, wenn Datenschutz, Verantwortlichkeit und Ethik von Anfang an im Design verankert sind, entsteht mehr als nur intelligente Technologie – es entstehen intelligente Gesellschaften.

Das Zeitalter des kognitiven Computing hat begonnen: eine Ära, in der Vertrauen die Grundlage bildet und Innovation der Motor ist. Ihr Erfolg wird nicht allein von Maschinen abhängen, sondern von der gemeinsamen Fähigkeit, die Prinzipien zu definieren, nach denen sie handeln. *Künstliche Intelligenz im Unternehmen – Aufbau einer vertrauenswürdigen KI in der Souveränen Cloud* – bietet einen Fahrplan für diese Reise und eine Einladung, die Zukunft gemeinsam verantwortungsvoll zu gestalten.

Über die Autoren

Shannon Bell

Shannon Bell ist Executive Vice President, Chief Digital Officer und Chief Information Officer bei OpenText. In dieser Funktion verantwortet sie die IT- und digitalen Systeme, Datenplattformen, Netzwerke und Kommunikationssysteme des Unternehmens sowie den kommerziellen und unternehmensweiten Cloud-Betrieb, die Sicherheitsarchitektur und die Compliance. Sie verfügt über mehr als 25 Jahre internationale Erfahrung in Technologietransformation und Großprojektintegration. Vor ihrer Tätigkeit bei OpenText war sie in leitender Funktion bei Rogers Communications tätig, wo sie die technologische Integration der Shaw-Übernahme leitete. Weitere berufliche Stationen sind unter anderem Amdocs, NewStep Networks, MetaSolv Software, Axiom Systems und Newbridge Networks. Shannon Bell hat einen MBA-Abschluss der University of Surrey und einen Bachelor-Abschluss der Carleton University.

David Fraser

Generalmajor a. D. David Fraser ist seit September 2018 Mitglied des Vorstands von OpenText. Er zählt zu den höchstdekorierten Generälen der kanadischen Streitkräfte und ist Träger des Ordens für militärische Verdienste. Im Jahr 2006 war er Kommandeur der multinationalen Brigade des Regionalkommandos Süd in Afghanistan und leitete die Operation Medusa – den größten Kampfeinsatz der kanadischen Streitkräfte seit über fünfzig Jahren. General Fraser diente auch als Kommandant des Canadian Forces Staff College. Nach seinem Ausscheiden aus dem Militär war er als Führungskraft in drei verschiedenen Unternehmen tätig, darunter Blue Goose Pure Foods, und verfügt über umfassende Führungserfahrung – sowohl auf dem Schlachtfeld als auch in der Wirtschaft. David ist zusammen mit Tom Jenkins und Mark J. Barrenechea Autor von *The Anticipant Organization*, einem Überlebensleitfaden für führende Organisationen in einer Welt ständiger Umbrüche.

Tom Jenkins

Tom Jenkins zählt zu den führenden Experten Kanadas im Bereich Innovation und digitale Technologien. Er ist Vorsitzender der OpenText Corporation – des größten Software- und Cloud-Unternehmens in der Geschichte Kanadas und eines der erfolgreichsten Internetunternehmen weltweit. Tom Jenkins war bzw. ist Mitglied der Aufsichtsräte von OpenText Corporation, Manulife Financial, Thomson Reuters, TransAlta Corporation, BMC Corporation und Slater Steel. Darüber hinaus war er Vorsitzender des Nationalen Forschungsrats von Kanada. Er erhielt sein Offizierspatent in den kanadischen Streitkräften und wurde zum Ehrenoberst eines Infanterieregiments sowie eines Jagdgeschwaders der kanadischen Streitkräfte ernannt. Tom Jenkins war der zehnte Kanzler der Universität Waterloo und wurde als Mitglied in die Canadian Business Hall of Fame aufgenommen. Tom ist Träger des Verdienstordens der Bundesrepublik Deutschland (Ritterkreuz) und Offizier des Order of Canada. Er hat zahlreiche wirtschaftswissenschaftliche Publikationen über digitale Innovation verfasst und war gemeinsam mit David Johnston, dem ehemaligen Generalgouverneur von Kanada, Mitautor des Buches *Ingenious: How Canadian Innovators Made the World Smarter, Smaller, Kinder, Safer, Healthier, Wealthier, and Happier (Wie kanadische Innovatoren die Welt intelligenter, kleiner, freundlicher, sicherer, gesünder, wohlhabender und glücklicher gemacht haben)*.

Danksagungen

Die Autoren danken den folgenden Personen für ihre Zeit, ihren Einsatz und ihre wertvollen Beiträge:

Michael Acedo, DeeDee Andrews, James Arroyo, Savinay Berry, Lev Dranikov, Paul Duggan, Lars Drexler, Joe Dwyer, Adam Hennessy, Bita Houshmand Rabiee, Michelle Kelly, Anupam Khazanchi, Edward Kiledjian, Mark L'Heureux, Stephen Ludlow, James McGourlay, Sandy Ono, Sunnie Rothenburger, Hans-Gerd Schaal, Scott Schultz, zusammen mit Elizabeth Chestney-Hanson (Redaktion), Stephen Ksiadz und Kevin Sy (Layout und Design), und Colombo Translation Ltd. für die Übersetzung ins Deutsche.

Inhaltsverzeichnis

Einführung

Willkommen im Zeitalter des kognitiven Computing

Wir erleben eine weitere entscheidende Zeitenwende in der technologischen Entwicklung – den Beginn des Zeitalters des kognitiven Computing, angetrieben durch den Aufstieg der künstlichen Intelligenz in Unternehmen (Enterprise AI, EAI). Was einst als digitale Revolution begann, hat sich zu etwas weit Dynamischerem entwickelt: zu einer Welt, in der Daten nicht nur Entscheidungen beeinflussen, sondern auch die Technologie dazu befähigen, Informationen zu interpretieren, daraus zu lernen und darauf zu reagieren.

In den vergangenen Jahrzehnten hat sich das Fundament der IT-Branche grundlegend gewandelt. Die COVID-19-Pandemie zwang ganze Volkswirtschaften nahezu über Nacht zur Digitalisierung. Cloud-Einführung, Remote-Arbeit und Automatisierung machten in wenigen Monaten Fortschritte, für die zuvor Jahre erforderlich gewesen wären. Das Ergebnis? Eine vollständig neu gestaltete Geschäftslandschaft ist entstanden, in der digitale Infrastruktur nicht mehr nur eine operative Ebene bildet, sondern das zentrale Nervensystem der gesamten Organisation.

Noch vor wenigen Jahren bedeutete Unternehmens-IT provisorische Server in überhitzten Schränken, auf denen eine einzige fehlerhafte Codezeile ein komplettes System zum Absturz bringen konnte. Heute sind diese anfälligen Strukturen durch hyperskalierbare Infrastrukturen ersetzt – global, widerstandsfähig und praktisch unbegrenzt skalierbar.
Auf diesem Rückgrat hat sich eine neue Intelligenz entwickelt. Agentenbasierte künstliche Intelligenz (KI) kann in Echtzeit Schlussfolgerungen ziehen, sich anpassen und reagieren. Aufgaben, für die früher ganze Teams aus Entwicklern, Marketingfachleuten oder Analysten erforderlich waren, lassen sich heute simultan und automatisiert für Millionen von Nutzern orchestrieren.

KLASSISCHE UNTERNEHMENSANWENDUNGEN IM WANDEL

Die Hyperscaler belagern die B2B-Festung mit kostengünstigen Cloud-Diensten.

Der „Burggraben" der Festung sind die GUIs und Workflows (Konfigurationsmanagement) der Anwendung.

Die „Mauer" der Festung sind die historischen Archivdaten, die für das Training der KI benötigt werden.

Das Aufkommen agentenbasierter KI droht, beide Verteidigungslinien zu durchbrechen.

Klassische Unternehmensanwendungen im Wandel

Dies markiert einen neuen Wendepunkt: die Konvergenz von Hyperscale-Computing und agentenbasierter KI. Hyperscaler und souveräne Clouds bilden das industrielle Netz der digitalen Wirtschaft – die Energiequelle, die alle weiteren Prozesse antreibt. Durch die Integration künstlicher Intelligenz entsteht ein Nervensystem für moderne Unternehmen, das Daten nicht nur speichert und verarbeitet, sondern auch vorausschauend interpretiert und dynamisch darauf reagiert. Für Unternehmen bedeutet das keine gewöhnliche Innovationswelle, sondern eine grundlegende Neudefinition von Arbeit, Wertschöpfung und Wissensfluss.

Klassische Unternehmensanwendungen verändern sich rasant, um mit dieser Entwicklung Schritt zu halten. Hyperscaler wie Amazon Web Services, Google Cloud und Microsoft Azure dringen in traditionelle B2B-Domänen vor und bieten leistungsfähige, kostengünstige Cloud-Dienste an, die etablierte Strukturen aufbrechen und neue Maßstäbe setzen. Jahrelang gründete sich die Sicherheit dieser Systeme auf ihrer Komplexität – auf individuellen Arbeitsabläufen, maßgeschneiderten Schnittstellen und gewaltigen Mengen historischer Daten, die tief in Archiven verborgen lagen. Doch genau dort baut sich der neue Druck auf.

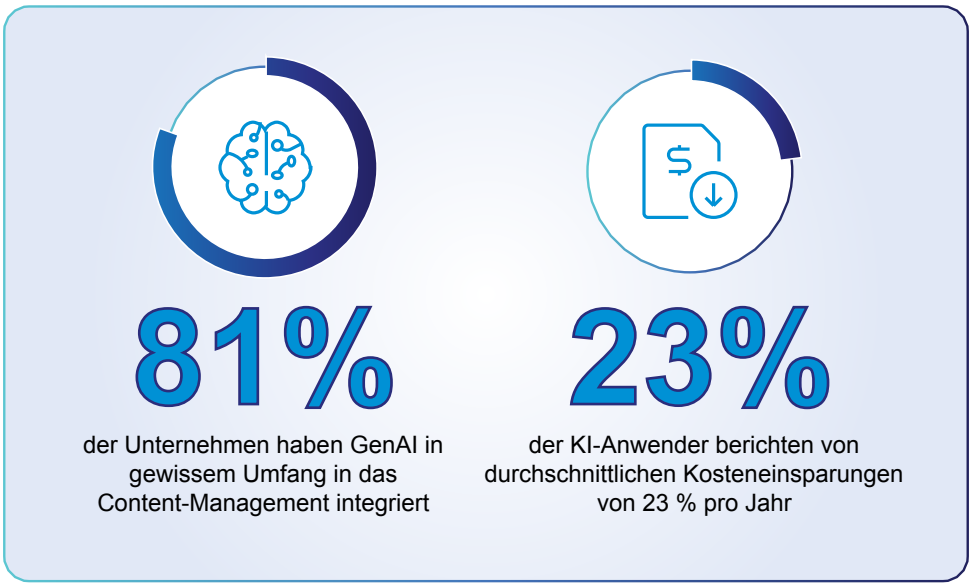
Diese alten Abwehrmechanismen halten nicht mehr so gut wie früher. Benutzeroberflächen und Konfigurationswerkzeuge, die einst einzigartig waren, lassen sich heute in kürzester Zeit replizieren. Und die Daten, die über Jahre ungenutzt in Archiven ruhten – Transaktionen, Interaktionen und Aufzeichnungen –, sind zu wertvollem Treibstoff geworden, aus dem KI-Modelle kontinuierlich lernen. Agentenbasierte KI tritt nicht nur als Ergänzung zu bestehenden Unternehmenssystemen auf, sondern lernt von ihnen, ahmt sie nach – und übertrifft sie zunehmend.

Der wahre Wettbewerbsvorteil liegt heute nicht mehr im Errichten höherer Mauern, sondern im Aufbau intelligenterer Fundamente. Die Widerstandsfähigkeit eines Unternehmens bemisst sich nicht an der Größe seines Software-Stacks, sondern an der Fähigkeit, die darin enthaltenen Daten wirksam zu verwalten, zu schützen und nutzbar zu machen. Hier beginnt echte Intelligenz – und echte Differenzierung.

Führungspersönlichkeiten im Bereich des kognitiven Computing müssen Vertrauen und Innovation miteinander in Einklang bringen. Sie erkennen, dass sichere und gut verwaltete Daten die Grundlage für Vertrauen bilden, während künstliche Intelligenz Fortschritt und Effizienz vorantreibt. Vertrauenswürdige Daten schaffen Zuversicht, Zuverlässigkeit und Compliance; KI eröffnet neue Wege zur Problemlösung und Wertschöpfung. Fehlt eine solide Daten-Governance, kann der schnelle Fortschritt im Bereich der KI Privatsphäre und Sicherheit gefährden – doch ohne Innovation droht Stillstand.

In dieser neuen Ära müssen Vertrauen und Innovation im Gleichgewicht bleiben. Nur die Verbindung robuster Datenpraktiken mit ambitionierten KI-Initiativen ermöglicht es Organisationen, transformative Technologien verantwortungsvoll einzusetzen und sicherzustellen, dass jeder Fortschritt auf öffentlichem Vertrauen und ethischen Prinzipien beruht.

Kurz gesagt: Nur vertrauenswürdige Daten ermöglichen eine wirklich vertrauenswürdige und wirksame KI.



Frühe Anwender profitieren bereits von der Einführung integrierter KI[1]

Das kommende Jahrzehnt wird den Organisationen gehören, die diesen Wandel wirklich verstehen. Denjenigen, die Daten nicht als Nebenprodukt ihres Geschäfts betrachten, sondern als dessen Betriebssystem. Unternehmen, die begreifen, dass Daten nur dann Kraft entfalten, wenn sie vertrauenswürdig, kontrolliert und sicher sind, werden die Zukunft prägen. Künstliche Intelligenz wird jede Branche verändern – doch nur dann, wenn sie auf einem soliden Fundament aufbaut.

Ein prägnantes Beispiel: Im Mai 2025 kündigte Salesforce die Übernahme von Informatica für rund 8 Milliarden US-Dollar an – ein Meilenstein, der deutlich macht, wie sich die Kräfteverhältnisse in der Unternehmenssoftware neu ordnen. *Warum war dieser Deal sinnvoll?* Salesforce war bislang auf externe Systeme angewiesen, um kritische Daten zu verwalten. Die Fähigkeiten zur internen Datenintegration waren begrenzt – und genau dort lag die Schwachstelle.

Durch die Übernahme schließt Salesforce diese Lücke und schafft die Grundlage dafür, dass KI auf kontrollierten, kontextbezogenen und vertrauenswürdigen Informationen aufbauen kann.

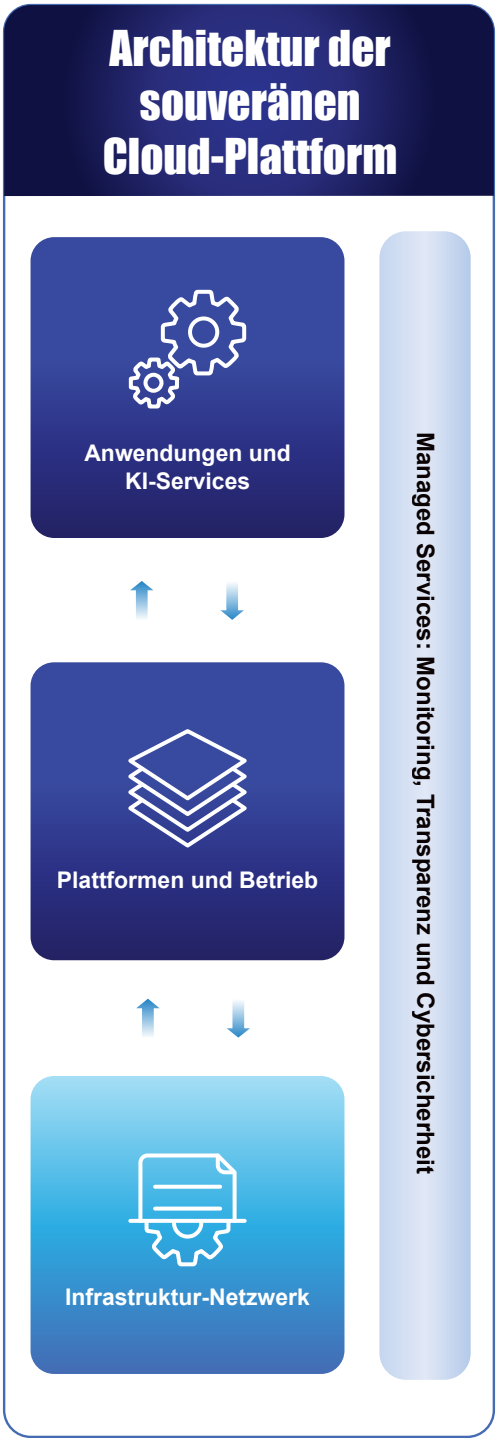
Salesforce erkannte, dass die Zukunft der KI vom Zugriff auf sichere, strukturierte und unstrukturierte Daten abhängt – auf Informationen also, die tief in den Systemen von Organisationen verankert sind und nicht im offenen Web zu finden sind. Das gilt besonders für das Training *agentenbasierter KI*-Systeme, die im kommenden Jahrzehnt zum zentralen Treiber globaler Unternehmensproduktivität werden dürften. Doch genau hier liegt die größte Herausforderung: Die meisten wertvollen Informationen der Welt sind nicht öffentlich zugänglich. Sie liegen in den Datenbeständen von Unternehmen, Regierungen und Institutionen – geschützt, reguliert und durch Governance-Mechanismen abgesichert.

Öffentliche Chatbots wie ChatGPT, Claude oder Perplexity wurden bereits mit nahezu allen frei verfügbaren Inhalten trainiert – darunter Beiträge aus Reddit, Artikel aus Wikipedia und andere offene Repositorien. Wenn diese Modelle nun versuchen, aus spezialisierteren, hochwertigen Informationsquellen zu lernen, stoßen sie unweigerlich an Grenzen: Urheberrechte, Datenschutzvorschriften und Eigentumsrechte an Daten setzen klare Schranken. Diese Informationen sind geschützte Daten – Daten, über die Organisationen Souveränität besitzen. Sie unterliegen zunehmend strengen Datenschutz- und Sicherheitsvorschriften – von kommunalen und föderalen Regelwerken bis hin zu internationalen Standards, die von den Vereinten Nationen und der NATO definiert werden.

Tatsächlich befinden sich über 90 Prozent der weltweiten Daten hinter Firewalls von Unternehmen und Regierungen.[2] Sie sind nicht nur schwer zugänglich, sondern strukturell geschützt. Zugriff erfordert Genehmigungen, die Einhaltung von Vorschriften und eine präzise Steuerung der Datenflüsse. Hier kommt das Enterprise Information Management (EIM) ins Spiel – Systeme und Verfahren zur Dokumenten- und Datensatzverwaltung, mit Workflows und Regelwerken, die genau festlegen, wer was, wann und warum sehen darf. Damit sich agentenbasierte KI in Unternehmen verantwortungsvoll weiterentwickeln kann, muss sie nicht nur aus Daten lernen, sondern innerhalb dieser Vertrauensbarrieren agieren.

Wird ein großes oder kleines Sprachmodell mit Daten trainiert oder feinjustiert, für deren Nutzung keine Rechte bestehen, erhält es im übertragenen Sinn „die Hausaufgaben eines anderen" – und das endet nie gut. Stellen sich diese Daten als vertraulich heraus, ist die Folge kein einfacher Patch, sondern ein vollständiger Neustart. Die betroffene Organisation kann verlangen, dass das Modell „ungeschult" wird – was bedeutet, von vorn zu beginnen. Eine fehlerhafte Datenquelle lässt sich nicht einfach isolieren; das gesamte System muss bis zum Ursprung des Fehlers zurückverfolgt werden. Das kann Millionen kosten und monatelange Verzögerungen verursachen. In einigen Fällen kann dies dazu führen, dass ein vielversprechendes KI-Programm von der Bildfläche verschwindet. Kurz gesagt: Beim Training von KI-Modellen gilt – zweimal prüfen, einmal handeln. Ein tiefes Verständnis von Datensouveränität ist daher entscheidend.

Jede Organisation verfügt über etwas, das zu wertvoll ist, um es zu riskieren: ihr institutionelles Wissen und ihre Daten – die „Schlüssel zur Burg". Sie machen das Unternehmen einzigartig. Wer diese Informationen an das falsche System oder den falschen Partner übergibt, läuft Gefahr, dass sie in veränderter Form wiederverwendet oder weiterverkauft werden. Auch Regierungen erkennen dieses Risiko und arbeiten mit Hochdruck an neuen Schutzmaßnahmen. Gesetze zur Regulierung von KI und Datenschutz stehen bevor – und sie werden die Datenschutz-Grundverordnung (DSGVO) im Rückblick wie einen Vorgeschmack erscheinen lassen. Diese neuen Regelungen werden die Art und Weise, wie Unternehmen Daten speichern, teilen und verarbeiten, grundlegend verändern – insbesondere was personenbezogene oder sensible Informationen betrifft.



Organisationen benötigen souveräne Daten

Um für den Einsatz von KI gerüstet zu sein, müssen Organisationen die drei zentralen Arten von Datensätzen verstehen, die intelligente Systeme antreiben – und sie verantwortungsvoll verwalten:

1. Von Menschen erstellte Inhalte: Dokumente, E-Mails, Präsentationen, Bilder, Videos und Gespräche – das lebendige Zeugnis dessen, wie eine Organisation denkt, kommuniziert und entscheidet.

2. Maschinell generierte Inhalte: Protokolldateien, Warnmeldungen und Telemetriedaten aus IT-Systemen, Netzwerken und Sicherheitstools – das permanente Hintergrundrauschen, das den operativen Zustand einer Organisation widerspiegelt.

3. Datenflüsse zwischen Organisationen: Transaktionen, Lieferantenaustausch und B2B-Integrationen. Das ist das Bindegewebe, das die Wirtschaft am Laufen hält.

Agentenbasierte Unternehmens-KI (EAI) benötigt alle drei Datentypen, um im realen Geschäftskontext effektiv zu funktionieren. Die nächste Entwicklungsstufe verschiebt sich von *Inhalten im Kontext* zu *KI im Kontext*—einer Phase, in der Intelligenz nicht nur Daten verarbeitet, sondern auch Beziehungen, Absichten und Werte versteht, um gezielt handeln und lernen zu können. Doch ebenso wie die Daten, aus denen sie lernt, muss auch die KI selbst sicher, reguliert und regelkonform sein. Das sind nicht bloß technische Anforderungen, sondern die Grundlage des Vertrauens – sie entscheidet darüber, ob KI verantwortungsvoll und sicher in Unternehmen eingesetzt werden kann.

Die entscheidende Frage lautet nicht, *ob* Daten geteilt werden – sondern *wie* die Kontrolle darüber erhalten bleibt.

Wahre Datensouveränität bedeutet, jederzeit zu wissen, wo sich die eigenen Daten befinden, wer darauf zugreift und wie sie genutzt werden – nicht nur einmal, sondern fortlaufend. Organisationen benötigen sichere Wege, um Generative KI (GenAI) dorthin zu bringen, wo ihre Informationen bereits liegen – in ihre kontrollierten und geschützten Systeme. So kann das Personal mit den Inhalten interagieren, sie finden, zusammenfassen und weiterverwenden – ohne gegen bestehende Governance-Regeln zu verstoßen. Auch Analysefunktionen werden dadurch neu definiert: Erkenntnisse lassen sich künftig einfach in natürlicher Sprache abrufen. Zugleich können Lösungen, die dieselbe Intelligenz auf Bereiche wie Cybersicherheit, Anwendungsmanagement und andere Geschäftsprozesse ausweiten, sicher implementiert werden.

Mit sicheren, kontrollierten und souveränen Daten müssen Unternehmen ihre Kronjuwelen nicht preisgeben, um Innovation zu ermöglichen. Sie können KI vertrauensvoll und verantwortungsvoll einsetzen. Denn im Zeitalter der verantwortungsvollen Intelligenz sind die klügsten Organisationen nicht diejenigen, die die KI mit den meisten Daten füttern. Es sind diejenigen, die wissen, welchen Daten sie vertrauen können.

Der Weg nach vorn: Ein Aufruf zur Führung

Wie bereits dargelegt, wird die nächste Innovationsära nicht von den Schnellsten geprägt. Sie wird von denen geprägt, die sich am verantwortungsvollsten verhalten. Sie entsteht an der Schnittstelle von Vertrauen und Innovation.

Jede Managemententscheidung, jede Codezeile und jedes KI-Modell trägt heute eine ethische Dimension in sich – denn Information bedeutet Macht, und die Art, wie diese Macht genutzt wird, wird das kommende Jahrzehnt bestimmen.

Jetzt ist der Moment, in dem Führungskräfte in Wirtschaft, Verwaltung und Industrie gefordert sind, Daten und Informationen nicht nur als Ressource, sondern als Verantwortung zu begreifen. Es gilt, die Datengrundlagen zu stärken, Governance von Beginn an in das Design zu integrieren und Sicherheit zum Standard, statt zur Ausnahme zu machen. Die Zukunft der KI hängt nicht allein davon ab, wie viel automatisiert werden kann, sondern davon, wie sehr wir den Systemen – und den Daten – vertrauen können, auf denen sie beruhen.

Branchenübergreifend zeigt sich dieselbe Herausforderung, sich schnell zu transformieren, ohne die Kontrolle zu verlieren. Erfolgreich werden jene Organisationen sein, die Mut zur Innovation mit Disziplin in der Führung verbinden. Sie werden den Datenschutz ebenso konsequent wahren, wie sie nach Erkenntnissen streben – Transparenz zu einem Wettbewerbsvorteil machen und KI-Systeme entwickeln, die nicht nur logisch denken, sondern auch Vertrauen schaffen können.

Wo digitaler Ehrgeiz zu verantwortungsvoller Intelligenz wird und wo die bahnbrechendsten Technologien auch von Prinzipien geleitet werden trifft Bereitschaft auf Verantwortung.

So wie vertrauenswürdige Daten das Fundament bilden und KI-Innovationen den Weg in die Zukunft weisen, werden die folgenden Kapitel Datenrahmenwerke, KI-Governance-Modelle und zentrale Überlegungen zu souveränen und nicht-souveränen Datenarchitekturen im Detail erläutern. Am Ende jedes Kapitels bietet eine Fünf-Punkte-Zusammenfassung die wichtigsten Erkenntnisse auf einen Blick – als Orientierungshilfe, um mit Zuversicht zu gestalten, zu steuern und Innovation verantwortungsvoll voranzutreiben.

Das Jahrzehnt verantwortungsvoller Aufklärung hat begonnen. Gemeinsam können wir es gestalten – sicher, ethisch und zum Wohl aller.

Kapitel Eins

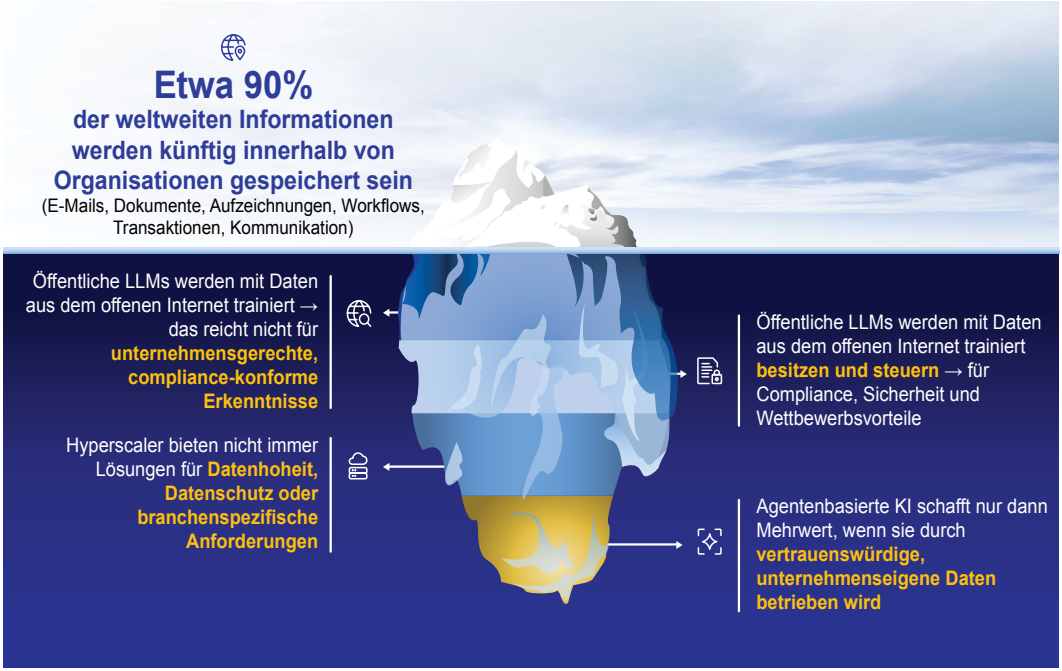
Die Entwicklung von Unternehmensdaten

Noch vor nicht allzu langer Zeit galten Unternehmensdaten als eine Art Dachbodenkiste – gefüllt mit alten Aufzeichnungen, Berichten und Unterlagen zur Erfüllung von Vorschriften. Sie wurden aufbewahrt, aber nicht aktiv genutzt. Nur wenn eine Zahl überprüft, ein Sachverhalt belegt oder ein Prüfer zufriedengestellt werden musste, griff man darauf zurück. Doch heute befinden sich die Daten im Erdgeschoss. Sie sind aktiv, vernetzt und steuern das Geschäft in Echtzeit. Sie beeinflussen jede Entscheidung und steuern jede Transaktion. Werden sie jedoch unbedacht offengelegt, können sie mehr preisgeben, als beabsichtigt ist.

Wer seine Daten hingegen schützt und vertraulich behandelt, verwandelt sie in einen echten Wettbewerbsvorteil.

Künstliche Intelligenz existiert nicht unabhängig von Daten – sie besteht aus Daten, gespeichert, organisiert und in Bewegung gesetzt. Das Unternehmen bleibt im Kern dasselbe Haus, nur dass der Dachboden leergeräumt wurde und das Gehirn nun im Erdgeschoss sitzt. Daraus folgt, dass die Datenschutz- und Sicherheitsgesetze, die bisher für Daten galten, nun auch auf KI angewendet werden müssen. Mit dem Übergang ins kognitive Zeitalter und der rasanten Entwicklung intelligenter Systeme sind Organisationen und Behörden verpflichtet, über ihren gesamten Lebenszyklus Informationen als verwaltetes Vermögen hinweg zu behandeln – nicht als passives Archiv.

Dieses Kapitel erläutert, was Unternehmensdaten sind, wie sie genutzt werden und wie sie mithilfe einer KI-Engine optimiert und durch eine Enterprise-Information-Management-(EIM)-Plattform in der Cloud verwaltet werden können. KI wird dabei als Intelligenzschicht innerhalb der Informationsarchitektur des Unternehmens betrachtet – eingebettet in Content-Dienste und Analysen, verbunden durch die operativen Geschäftsprozesse.



Das verborgene Netz

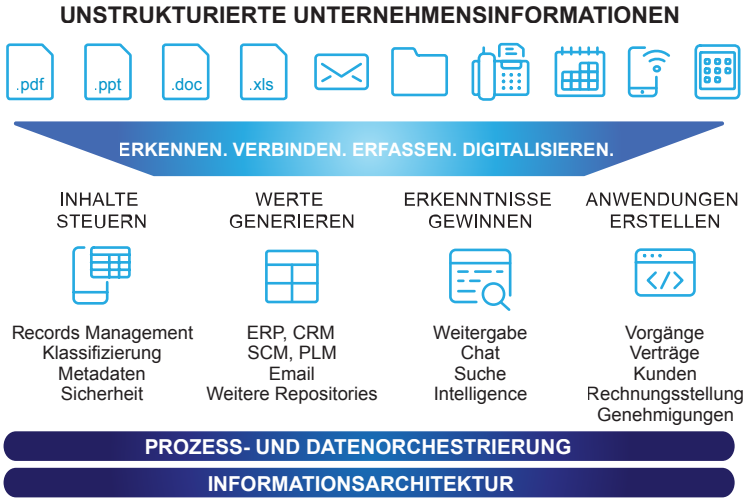
Die reale Landschaft: 10 Teile privat, 1 Teil öffentlich

Der größte Teil der weltweit verfügbaren Unternehmensdaten liegt hinter Firewalls verborgen. Nach Angaben von IDC bestehen nahezu 90 Prozent aller Unternehmensdaten aus unstrukturierten Inhalten wie E-Mails, Berichten, Dokumenten, Bildern oder Aufzeichnungen.[3] Diese privaten Daten übertreffen die öffentlichen Webinhalte, auf denen heutige generative KI-Systeme basieren, bei Weitem. Doch ein Großteil dieser Daten bleiben unstrukturiert, fragmentiert oder in Datensilos eingeschlossen.

Dieses Ungleichgewicht hat weitreichende Folgen. Das Verhältnis von privaten zu öffentlichen Daten liegt bei etwa 10:1, was bedeutet, dass der größte Teil des weltweiten Wissenspotenzials in Systemen verborgen ist, die öffentlichen KI-Modellen unzugänglich bleiben. Der wahre Wettbewerbsvorteil liegt daher im Inneren des Unternehmens – in Verträgen, Konstruktionsdateien, Rechnungen, Wartungsprotokollen, klinischen Notizen und der geschäftlichen Korrespondenz – vorausgesetzt, dass diese Daten geordnet, vernetzt und vertrauenswürdig sind.

Enterprise Information Management (EIM) wurde genau für diese Herausforderung geschaffen. Es vereinheitlicht, sichert und operationalisiert Unternehmensdaten, damit diese verantwortungsvoll und strategisch genutzt werden können. Ein wirksames Informationsmanagement bedeutet, jederzeit zu wissen, wo Informationen gespeichert sind, wem sie gehören, wer Zugriff darauf hat und wann sie verändert wurden.

Aber nicht alle Daten sind gleich. Strukturierte Daten – Zahlen in einer Datenbank – lassen sich problemlos sortieren, abfragen und auswerten. Unstrukturierte Daten – Wörter, Bilder, Videos, Sprache – entziehen sich dieser Ordnung. Sie erfordern Indizierung, Kontext und Klassifizierung. Daher unterscheiden sich die Technologien zur Verwaltung von Zahlen grundlegend von denen, die zur Verwaltung von Sprache und Inhalten erforderlich sind.



Unstrukturierte Unternehmensdaten

Um den Wert unstrukturierter Daten zu begreifen, muss man das Kontext dessen betrachten, was sich hinter jedem einzelnen Geschäftsdatensatz verbirgt. Ein Beschäftigter erscheint in einer Personaldatenbank lediglich als eine Zeile, ist jedoch mit Tausenden von Dokumenten verknüpft – Lebensläufen, Verträgen, Gehaltsabrechnungen, E-Mails und Leistungsbeurteilungen. Ein Vermögenswert, etwa ein Flugzeugtriebwerk oder eine Energieturbine, existiert als einzelner Datensatz im Enterprise-Resource-Planning-(ERP)-System, ist jedoch umgeben von einem dichten Netz aus Handbüchern, Prüfberichten, Inspektionsdaten und Wartungsprotokollen.

Gemeinsam bilden strukturierte und unstrukturierte Daten das tiefe, verborgene Netz des Unternehmens – Daten, die für öffentliche Suchsysteme unsichtbar, aber für den täglichen Betrieb unverzichtbar sind. Jedes digitale Artefakt trägt zu dieser Ebene bei: Jede E-Mail, jeder Bericht, jeder Entwurf, jedes Bild, jeder Chatverlauf. Mit dem Einzug mobiler Technologien und moderner Kollaborationslösungen haben Vielfalt und Geschwindigkeit der Inhalte explosionsartig zugenommen. EIM wurde geschaffen, um diese Komplexität zu erfassen, zu klassifizieren und zu steuern – damit jede Information nachvollziehbar, zugänglich und regelkonform bleibt.

Während EIM in Unternehmensdaten für Struktur und Nachvollziehbarkeit sorgte, eröffnet Künstliche Intelligenz nun ihre nächste Dimension. Generative Modelle, die überwiegend mit öffentlichen Daten trainiert wurden, können schreiben, zusammenfassen und Prognosen erstellen – doch sie können nicht innerhalb eines Unternehmens handeln. Ihnen fehlen die kontrollierten, genehmigten und internen Daten, die für fundierte Entscheidungen notwendig sind. Ohne diese Daten kann KI keine operativen Aufgaben ausführen – keine Rechnungen freigeben, keine Wartungsarbeiten planen und keine technischen Zeichnungen interpretieren.

Um diese Grenze zu überwinden, braucht KI genau das, was Informationsmanagement seit Jahrzehnten bereitstellt: ein System sicherer, regelkonformer und zugriffsberechtigter Unternehmensdaten. Nur unter diesen Voraussetzungen kann KI innerhalb der Firewall verantwortungsvoll arbeiten und echte Intelligenz entfalten – statt nur Muster aus offenen Quellen zu imitieren.

Die Verwaltung solcher Daten bedeutet, jederzeit zu wissen, wo sie gespeichert sind, wem sie gehören, wer Zugriff darauf hat und wann sie geändert wurden. Diese Grundprinzipien des Enterprise Information Management (EIM) – Berechtigungen, Metadaten und Lebenszykluskontrolle – machen Unternehmensinformationen vertrauenswürdig. Diese Funktionen müssen nun konsequent auf KI-Systeme angewendet werden.

Wo soll man anfangen? Werfen wir einen Blick darauf, wo Unternehmensinformationen gespeichert sind.

Wo die Daten gespeichert sind: Arten von Unternehmensdaten

Jede Organisation erzeugt Informationen aus klaren Motiven: um Abläufe zu dokumentieren, Kommunikation zu ermöglichen, Wissen zu bewahren, Vorschriften einzuhalten und Kundennutzen zu schaffen. Was einst als strukturierte Datenerfassung begann – Transaktionen, Rechnungen und Inventarlisten auf Großrechnern –, hat sich zu einem vernetzten Kommunikations- und Kollaborationssystem aus E-Mails, gemeinsam genutzten Dokumenten und digitalen Arbeitsräumen entwickelt.

Heute lassen sich diese Informationsströme in drei Hauptklassen von Unternehmensdaten gliedern, die für KI von unterschiedlicher Bedeutung sind, nämlich in von Menschen erstellte Inhalte, maschinell erstellte Daten und Transaktions- oder Geschäftsnetzwerkdaten. Jede dieser Kategorien weist eine eigene Struktur, eigene Governance-Anforderungen und eine spezifische Rolle beim Training und Einsatz von KI-Modellen auf. Gemeinsam bilden sie die Grundlage agentenbasierter Intelligenz im Unternehmen, das Fundament, auf dem sichere, kontextbezogene und lernfähige KI aufgebaut ist.

Von Menschen erstellte Inhalte: Die Sprache des Unternehmens

Von Menschen erstellte Inhalte umfassen Dokumente, E-Mails, Scans, Multimedia-Dateien, Fallnotizen und andere Kommunikationsformen. Sie sind reich an Bedeutung und Kontext, jedoch von Natur aus unstrukturiert. Dies ist das Gebiet des Content-Managements – dort, wo Informationen persönliche, kontextbezogene und häufig vertrauliche Daten enthalten, die eine präzise Klassifizierung, Metadatenkennzeichnung und Lebenszyklusverwaltung erfordern.

In diesen Inhalten liegen die Richtlinien, Formulierungen und Präzedenzfälle, die bestimmen, wie ein Unternehmen denkt und handelt. Das Trainieren von KI mit solchem Material setzt Anonymisierung und strenge Governance voraus, doch der Nutzen ist erheblich: Hier liegen die Intentionen, Geschäftsregeln und das institutionelle Wissen, die einer Organisation ihre Identität verleihen. Richtig verwaltet, werden unstrukturierte Inhalte zur Grundlage von Retrieval-Augmented Generation (RAG), Prompt-Bibliotheken und Natural Language Processing – Schlüsselkompetenzen, die es agentenbasierter KI ermöglichen, mit Verständnis, statt bloßer Automatisierung zu agieren.

Maschinell erzeugte Daten: Das Nervensystem des Unternehmens

Maschinell generierte Daten entstehen aus den Systemen, die eine Organisation antreiben – Protokolle, Telemetrie, Leistungskennzahlen und Monitor-Daten. Sie zeichnen sich durch hohes Volumen, hohe Geschwindigkeit und klare Struktur aus und zeigen in Echtzeit, was im Unternehmen geschieht. Dies ist der Bereich der Beobachtbarkeit und des operativen Bewusstseins, in dem jedes Ereignis und jede Abweichung eine Spur hinterlässt.

Maschinendaten liefern die kausalen Signale, die KI benötigt, um Infrastrukturen intelligent zu steuern, Muster zu erkennen, Ausfälle vorherzusagen oder Korrekturmaßnahmen einzuleiten, noch bevor das Personal sie bemerkt. Die Herausforderungen liegen im Datenumfang, den Kosten der Verarbeitung und Speicherung sowie in der Notwendigkeit, Rohsignale mit Geschäftskontext zu verknüpfen. In Kombination mit Richtlinien und von Menschen erstellten Inhalten ermöglichen diese Daten einem KI-Agenten, autonom nachvollziehbar und im Einklang mit den geltenden Governance- und Compliance-Vorgaben zu reagieren.

Interne und externe Transaktionsdaten: Die Quelle der Wahrheit

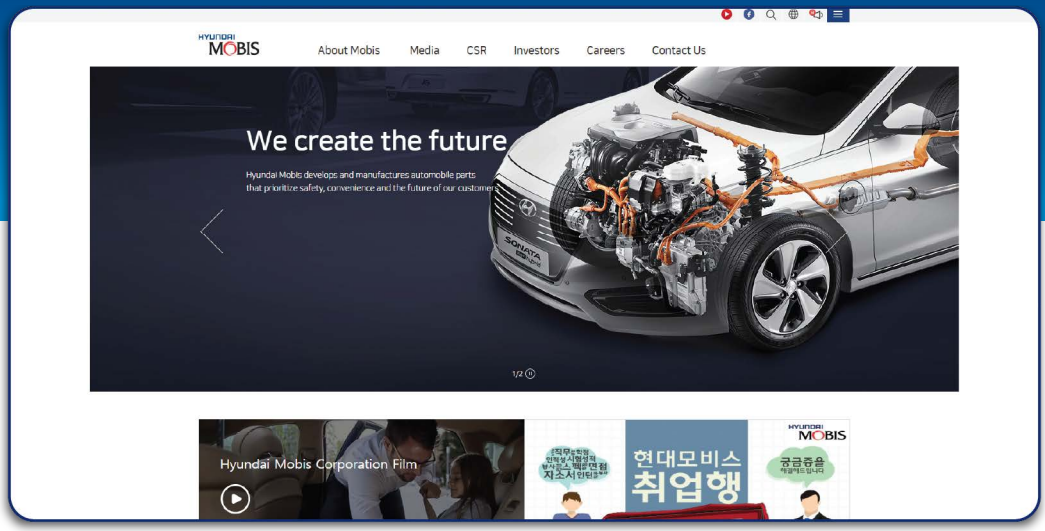
Transaktionsdaten, Bestellungen, Rechnungen, Versandbenachrichtigungen und andere strukturierte Geschäftsnachrichten, bilden die rechtliche und wirtschaftliche Wahrheit der Unternehmensaktivitäten. Sie definieren die gegenseitigen Verpflichtungen zwischen Organisationen und sind unverzichtbar für Compliance, Besteuerung und Wirtschaftsprüfung. Da sie wohldefinierten Schemata folgen und über eine hohe semantische Präzision verfügen, bieten sie eine verlässliche Grundlage für eindeutige Schlussfolgerungen.

Für agentenbasierte KI bilden sie das Fundament faktischer Entscheidungen. Sie ermöglichen es den Mitarbeitern, Finanzdaten abzugleichen, Cashflows zu prognostizieren und Abweichungen in komplexen Lieferketten zu erkennen, ohne dass Ergebnisse manipuliert oder verfälscht werden. Durch die Integration dieser strukturierten Daten mit den unstrukturierten Inhalten und den operativen Signalen entsteht ein ganzheitliches Bild des Unternehmens darüber was geschieht, warum es geschieht und was als Nächstes zu tun ist.

Im folgenden Beispiel zeigt MOBIS, ein Automobilhersteller, wie ein teilebasiertes Produktionssystem mithilfe von Analytik und Business Intelligence die Qualität sichert und Kosteneffizienz entlang der gesamten Lieferkette erzielt.

Case Study

MOBIS



MOBIS

MOBIS hat seinen Hauptsitz in Seoul, Südkorea, verfügt über Tochtergesellschaften in rund 40 Ländern weltweit und steuert die Lieferkette für die Schwergewichte der Automobilindustrie Hyundai Motor Company und Kia Motors. Das Unternehmen hat ein Teileproduktionssystem entwickelt, das Qualität und Kosteneffizienz entlang der gesamten Lieferkette – vom Einkauf über die Lagerhaltung bis hin zu Vertrieb und Logistik – sicherstellen und seinen Kunden helfen soll, sich in der wettbewerbsintensiven Automobilindustrie zu differenzieren. MOBIS Parts Australia Pty Ltd. (MPAU) ist die australische Tochtergesellschaft des Automobilzulieferers.

Die Automobilindustrie steht unter konstantem Wettbewerbsdruck, da andere Marken fortlaufend neue Produkte, Vertriebsstrategien und Preisstrukturen entwickeln. Daher muss MPAU sicherstellen, dass geeignete Systeme und Technologien vorhanden sind, um die Wettbewerbsfähigkeit der Produkte und die Agilität der Abläufe zu sichern.
Um flexibel auf sich ändernde Anforderungen oder Wettbewerbsangebote reagieren zu können, benötigte das Unternehmen ein neues Business-Intelligence-System, das Echtzeitberichte zu Beständen und Händlernetzwerken unterstützt, die Verkaufsleistung und Preisangebote von Lieferanten überwacht und Analysefunktionen zur Vorhersage zukünftiger Verkaufs- und Bestandsanforderungen bereitstellt.

> *Aus logistischer Sicht verschafft uns die Integration der Backend-Informationen in das Frontend des BI-Systems einen klaren Überblick über die Geschäftsprozesse und ermöglicht die Analyse der Daten aus dem Backend-System.*
>
> IT MANAGER, MPAU

Nach der Erprobung mehrerer Systeme entschied sich MPAU schließlich für eine Analytics Suite, die durch robuste Funktionalität und einfache Bedienbarkeit überzeugte. Letzteres war entscheidend, um sicherzustellen, dass das Personal das System intuitiv nutzt und die Zusammenarbeit mit den Vertriebspartnern stärkt, indem es sich nahtlos in den täglichen Betrieb von rund 140 bis 160 Nutzern und Händlern einfügt. Die Lösung ließ sich mit Datenquellen im gesamten MPAU-Betrieb und Händlernetzwerk verknüpfen, sodass Dashboards jeder Abteilung – von der Lagerhaltung bis zu Vertrieb und Logistik – einen Überblick über die täglichen Aktivitäten boten.

Die Analysefunktionen verschaffen dem Unternehmen einen Wettbewerbsvorteil, da nicht nur historische Umsätze und Lagerbestände einsehbar sind, sondern auch zukünftiger Bedarf prognostiziert werden können. Anstelle umständlicher Berichtsprozesse und ungenauer Prognosen können die Nutzer historische Daten in Echtzeit mit aktuellen Verkaufsinformationen vergleichen und daraus künftige Umsätze ableiten. Das Ergebnis ist eine effizientere Arbeitsumgebung mit fundierter Entscheidungsgrundlage, die verlässlichen Datenzugriff und Interaktion ermöglicht und MPAU zu mehr Agilität im Wettbewerbsumfeld verhilft.

Wenn Informationstypen zusammenarbeiten

Wie die Fallstudie zeigt, entfalten Unternehmensdaten ihr volles Potenzial erst, wenn unterschiedliche Informationstypen zusammengeführt werden. Wenn menschliches Wissen, maschinelle Telemetrie und Geschäftstransaktionen gesteuert, vernetzt und in den richtigen Kontext gesetzt werden, entsteht die lebendige Architektur eines intelligenten Unternehmens.

Ein Beispiel ist eine Shared-Services-Abteilung im Finanzwesen, die einen KI-Assistenten nutzt, um nicht übereinstimmende Bestellungen und Rechnungen abzugleichen. Der Assistent analysiert Rechnungsbilder und OCR-Ausgaben (von Menschen erzeugte Inhalte), gleicht sie mit ERP-Transaktionsdatensätzen (Transaktionsdaten) ab und prüft Systemprotokolle, die den Empfang oder die Genehmigung von Rechnungen dokumentieren (Maschinendaten). Mit einer einheitlichen Metadatenebene, die Dokumentenherkunft, Zugriffsrechte und Zeitstempel erfasst, kann der KI-Assistent eine nachvollziehbare Empfehlung erstellen, die die Bearbeitungszeit verkürzt und die Prüfungsintegrität sicherstellt.

In einem weiteren Beispiel greift ein Beschäftigter der Rechtsabteilung, der Schreiben zur Beweissicherung vorbereitet, auf Richtliniendokumente und frühere Kommunikation zurück (von Menschen erstellte Inhalte), nutzt Fristen und Falldaten (Transaktionsdaten) und prüft Serverzugriffsprotokolle, um die Verwahrung zu bestätigen (Maschinendaten). Das Ergebnis ist ein präziser, regelkonformer Entwurf, der in Minuten statt Tagen erstellt wird. Dasselbe Prinzip findet sich in großem Maßstab auch in realen Organisationen wieder, wie die folgende Fallstudie zeigt.

Fallstudie

Ein unabhängiges Energieunternehmen

Ein unabhängiges Energieunternehmen, das zugleich öffentlicher Versorger ist, beliefert mehr als 150.000 Kunden. Das Unternehmen benötigte eine Lösung, um die stetig wachsenden Informationsmengen zu bewältigen und gleichzeitig die Einhaltung zahlreicher regulatorischer Vorgaben sicherzustellen. „Als leistungsstarkes Unternehmen müssen wir gewährleisten, dass die richtigen Informationen zur richtigen Zeit am richtigen Ort sind, damit wir fundierte Entscheidungen treffen können“, erklärt der Leiter der Unternehmensdokumentation. „Dafür müssen wir Compliance-Vorgaben berücksichtigen, Herausforderungen der Informationsorganisation meistern und Geschäftsprozesse fortlaufend optimieren.“

Durch die Erweiterung des Enterprise Information Management um KI-gestützte Suche und Automatisierung kann das Unternehmen Informationen nun schneller finden und effizienter darauf reagieren. Intelligente Suchfunktionen ermöglichen den Zugriff auf strukturierte und unstrukturierte Inhalte, E-Mails, Tabellen, Berichte und PDFs, innerhalb einer sicheren Umgebung. KI-gestützte Zusammenfassungen unterstützen das Personal bei der Interpretation umfangreicher technischer Daten und behördlicher Dokumente, während maschinelle Lernmodelle Aufbewahrungsprobleme aufdecken und Compliance-Prüfungen automatisieren. Das Unternehmen steuert nun Compliance und Performance parallel und nutzt KI, um die Unternehmensführung proaktiver und weniger manuell zu gestalten.

EIM und Enterprise AI ermöglichen es, regulatorische Anforderungen mit operativer Agilität in Einklang zu bringen und Daten-Governance von einer Pflicht in eine Quelle von Erkenntnissen zu verwandeln. Diese Beispiele – von automatisierter Finanzabstimmung bis zu unternehmensweitem Energiemanagement – verdeutlichen eine zentrale Erkenntnis: KI und Automatisierung erzeugen Mehrwert nicht aus einzelnen Datensätzen, sondern aus ihren Beziehungen. Wenn Daten vereinheitlicht, als vertrauenswürdig eingestuft und in den richtigen Kontext gesetzt werden, werden sie zu mehr als bloßen Informationen – sie werden zu Intelligenz.

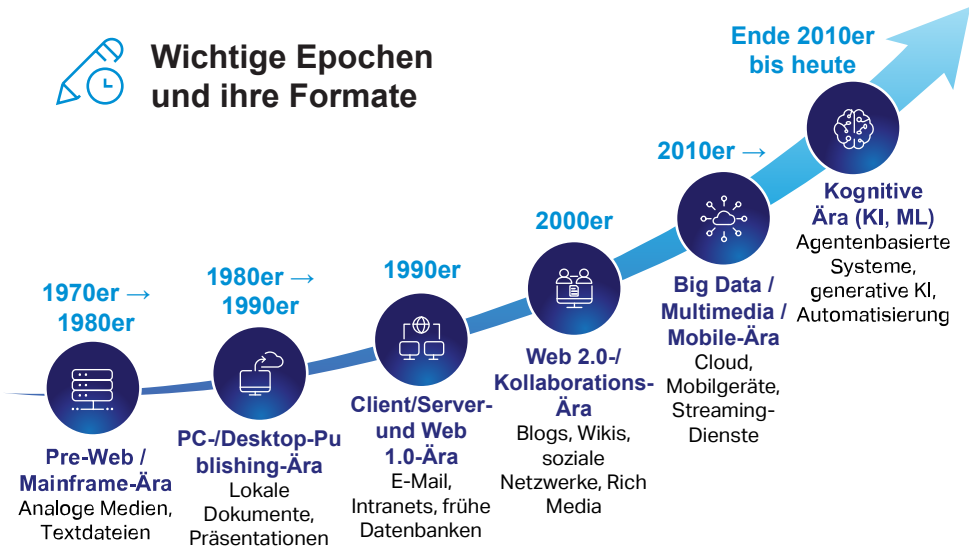
Die digitale Geschichte jeder Organisation beginnt mit ihren Formaten. Dateien, Datensätze und Container, die zur Informationsspeicherung verwendet werden, spiegeln von Lochkarten und Druckspooldateien bis zu API-basierten Strukturen in JavaScript Object Notation (JSON) stets die Technologien und Prioritäten ihrer Zeit wider die heute KI-fähig sind.
Ein Blick auf fünf Jahrzehnte Unternehmensinformation zeigt eine klare Entwicklung: Jeder technologische Fortschritt führte zu einem entsprechenden Fortschritt im Umgang mit Inhalten.

Die Ära vor dem Internet und den Großrechnern (1970er–1980er Jahre)

Die erste Generation von Unternehmensdaten entstand im Mainframe. Sie war strukturiert, starr und auf maximale Maschineneffizienz ausgelegt. Textdatensätze mit fester Breite, COBOL-Datendateien und ISAM-Formate (Indexed Sequential Access Method) ersetzten handgeschriebene Kontobücher und Papierjournale. Da Speicherplatz knapp war, zählte jedes Byte.

Die Verarbeitung erfolgte stapelweise, der Fokus lag auf Durchsatz statt Interaktion. Systeme wurden so konzipiert, dass sie „die Daten über Nacht verarbeiteten“ und am nächsten Morgen gedruckte Berichte und Spool-Dateien lieferten. Kodierungsverfahren wie EBCDIC (Extended Binary Coded Decimal Interchange Code) hielten die Systeme proprietär und eng an die jeweilige Hardware gebunden.

Diese frühen Systeme legten die Grundlage für eine strukturierte Datendisziplin. Sie führten Schema-Kontrolle, versionierte Datensatzformate und erste Ansätze von Metadaten ein. Die Idee dahinter war, dass jedes Feld eine eigene Bedeutung erhält. Die Starrheit der Mainframes zwang Unternehmen dazu, Daten als Vermögenswert zu behandeln, lange bevor der Begriff Data Governance überhaupt existierte.



Wichtige Epochen und die von ihnen geprägten Formate

Das PC- und Desktop-Publishing-Zeitalter (1980er–1990er Jahre)

Mit dem Aufkommen des Personalcomputers kam die Freiheit – und zugleich die Dezentralisierung. Mitarbeiter konnten Inhalte erstmals unabhängig vom Mainframe erstellen. Programme wie WordPerfect und frühe Versionen von Microsoft Word®, Tabellenkalkulationen und Desktop-Publishing-Tools verbreiteten sich rasant in Büros und auf Disketten.

Die Erstellung von Informationen verlagerte sich von zentralen Rechenzentren auf individuelle Arbeitsplätze. Berichte, Memos und Präsentationen entstanden in neuen Formaten wie DOC, XLS und PostScript. Zum ersten Mal waren Dokumente visuell, editierbar und massenhaft druckfähig.

Diese Demokratisierung der Inhalte erhöhte die Produktivität, führte jedoch zu Kontrollverlust. Daten, die früher in zentralisierten Systemen gespeichert waren, verteilten sich auf lokale Festplatten. Damit begann die Content-Explosion, die später die Entwicklung unternehmensweiter Content-Management-Systeme notwendig machte.

Client/Server und Web 1.0 (1990er Jahre)

Mit der zunehmenden Vernetzung verlagerte sich der Informationsfluss in Richtung Internet. Das Web brachte HTML-Seiten, GIF- und JPG-Bilder sowie frühe XML-Formate für strukturierten Datenaustausch hervor. Interne Intranets spiegelten öffentliche Websites wider, wodurch sich die Herausforderung von der Erstellung hin zur Auffindbarkeit verschob.

Client/Server-Architekturen ermöglichten den gemeinsamen Zugriff auf Datenbanken und Anwendungen, während Browser die universelle Veröffentlichung von Informationen etablierten. Der Bedarf an Indexierung und Suchfunktionen führte zur Entwicklung der ersten Metadatenmodelle und Dokumentenmanagementsysteme. Damit begann die Suchrevolution, geprägt vom Lebenszyklusdenken: Erstellung, Speicherung, Abruf und Entsorgung von Informationen.

Web 1.0 machte Unternehmenswissen dynamisch, vernetzt und zunehmend komplex. Damit wurde auch der Grundstein für die Datenverwaltung gelegt: Sobald man Daten gefunden hat, muss man entscheiden, wer sonst noch Zugriff darauf haben sollte.

Web 2.0 und das Zeitalter der Zusammenarbeit (2000er Jahre)

Anfang der 2000er Jahre entstand ein stärker auf Interaktion und Beteiligung ausgerichtetes Internet. Rich Media, PDF/A-Archivierungsstandards, XML- und JSON-Formate sowie Multimedia-Codecs wie MP3 und MP4 wurden zum Standard.

Nutzergenerierte Inhalte und Kollaborationsplattformen hielten Einzug in Unternehmen. E-Mail-Archive, Portale und Wikis ergänzten klassische Datenverwaltungssysteme. Das Content-Management entwickelte sich weiter, um persistente, teilbare Dokumente über Abteilungs- und Organisationsgrenzen hinweg zu verwalten.

Gleichzeitig traf die wachsende Regulierung auf die digitale Transformation. Gesetze wie der Sarbanes-Oxley Act und der Health Insurance Portability and Accountability Act (HIPAA) zwangen Unternehmen, nachzuweisen, was sie wussten – und wann und wie sie es wussten. Diese Verbindung von Regulierung und Zusammenarbeit schuf die Grundlage für Records Management, Versionskontrolle und richtlinienbasierte Archivierung – die Kernpfeiler des modernen Enterprise Information Management.

Das Zeitalter von Big Data, Multimedia und Mobilgeräten (2010er Jahre →)

In den 2010er Jahren übertraf die Informationsmenge erstmals die Zahl klassischer Dokumente. Datenströme, Telemetrie und Zeitreihendaten flossen aus mobilen Geräten, Sensoren und Apps. Neue analytische Formate – Avro, Parquet, ORC (Optimized Row Columnar) – wurden für Skalierbarkeit und Geschwindigkeit optimiert und zum Standard in Data Lakes und Cloud-Speichern.

Mit der Verbreitung von Smartphones und digitalen Erlebnissen wuchsen Video-, Sprach- und Bildarchive exponentiell. Objektspeicher und Content Delivery Networks definierten den Begriff „Datei“ neu als adressierbaren Blob mit Metadatenhülle.

Diese Datenflut führte zum Entstehen des versteckten Netzes, eines riesigen Universums unstrukturierter Inhalte hinter der Firewall. Es stellte sowohl eine Chance als auch eine Belastung für Unternehmen dar; eine Chance, mithilfe von Analysen Daten zu gewinnen und KI-Training anzuwenden, aber gleichzeitig ein Risiko in Bezug auf die Kosten und Compliance.

Das kognitive Zeitalter (Ende der 2010er Jahre → Gegenwart)

Die heutige Informationslandschaft ist dynamisch, vernetzt und multimodal. Daten fließen nicht mehr nur zwischen Menschen und Systemen, sondern auch zwischen den Maschinen. APIs – insbesondere REST (Representational State Transfer) und GraphQL – verbinden Microservices. Container, Notebooks und strukturierte Protokolle schaffen neue hybride Dokumenttypen.

Diese Ära ist geprägt von Interoperabilität, Automatisierung. Informationen werden heute ebenso von KI, maschinellem Lernen und digitalen Agenten erzeugt wie konsumiert. Tokenisierte Textkorpora, eingebettete Metadaten und semantische Verschlagwortung ermöglichen abrufgestützte Generierung und kontextbezogenes Schließen.

Während frühere Epochen auf Formateffizienz ausgerichtet waren, steht nun die Bedeutung im Mittelpunkt. Das moderne Unternehmen muss strukturierte und unstrukturierte Daten, wie Sprache, Bild, Text und Transaktion, über alle Modalitäten hinweg in kontrollierten Ökosystemen vereinen, die sowohl Analysen als auch intelligentes Handeln ermöglichen.

Im kognitiven Zeitalter sind Formate keine statischen Container mehr, sondern Schnittstellen zwischen menschlicher Absicht und maschinellem Verständnis. Der Lebenszyklus, der einst für Dokumente galt – Erfassung, Verwaltung, Verarbeitung, Suche, Archivierung – erstreckt sich heute auf Wissen selbst.

Über alle Epochen hinweg bleibt das Grundprinzip bestehen: Die Technologie verändert sich, doch das Bedürfnis nach Vertrauen und Kontext bleibt unverändert. Von COBOL-Berichten bis zu Cloud-APIs definiert jedes neue Format nicht nur, wie Daten gespeichert werden, sondern auch, wie sie verwaltet, geteilt und verstanden werden. Die Lehre daraus ist eindeutig: Informationsmanagement entwickelt sich mit dem Medium. Was einst der Kontrolle über Dateien diente, umfasst heute die Steuerung von Wissensflüssen und Intelligenzsystemen. Die Formate der Vergangenheit dienten der Lesbarkeit; die der Zukunft dienen dem Lernen.

Warum gute Governance an erster Stelle steht

Enterprise-Information-Management war schon immer ein System des Vertrauens. Es strukturiert Informationen entlang ihres gesamten Lebenszyklus – von der Erfassung über Verwaltung, Verarbeitung und Suche bis zur Archivierung – und verknüpft diese Prozesse direkt mit den Geschäftsabläufen, die sie hervorbringen. Richtig umgesetzt, schafft Governance Transparenz, minimiert Risiken und senkt die Kosten für Compliance. Sie ist kein Zusatz zur KI-Strategie, sondern ihre Grundlage.

Die folgende Fallstudie zeigt, wie die UBS ihre Informationen mithilfe einer zentralen EIM-Plattform verwaltet und gleichzeitig regulatorische Anforderungen erfüllt.

Fallstudie

UBS

Zertifizierungsprozess bei UBS

Als Reaktion auf die Compliance-Anforderungen der Abschnitte 302 und 906 des Sarbanes-Oxley Act führte UBS, eines der weltweit führenden Finanzinstitute, einen internen Zertifizierungsprozess für Finanzberichte ein, bei dem leitende Angestellte ihre Finanzzahlen und -prozesse formell mittels eines „Subconfirmation"-Verfahrens bestätigen.

Während dieses Prozesses werden die verantwortlichen Personen per E-Mail benachrichtigt, sobald ihre Mitwirkung erforderlich ist, und erhalten anschließend personalisierten Zugriff auf die relevanten Dokumente im UBS-Intranet. Alle Vorgänge werden automatisch protokolliert und revisionssicher archiviert. Der CEO und der Group Controller – in der Regel der CFO – erteilen der Securities Exchange Commission erst dann eine endgültige Bescheinigung, wenn alle internen Prozesse abgeschlossen sind.

Das Corporate-Governance-Portal der UBS ermöglicht es Führungskräften weltweit, interne und externe Geschäftsberichte gemeinsam zu erstellen. Die zuständigen Abteilungen behalten dabei jederzeit den Überblick über den Status der Zertifizierungsprozesse. Durch die Automatisierung und Vereinfachung der Abläufe wurde der gesamte Prozess erheblich beschleunigt.

Abbildung von EIM auf KI: Internet, Intranet und Extranet

Enterprise Information Management (EIM) bietet ein anschauliches Modell, um den Reifegrad von KI-Systemen zu verstehen. So wie Informationen vom öffentlichen in den privaten Bereich übergehen, entwickelt sich auch KI von generischer zu kontextbezogener Intelligenz.

- **Internet = Generative KI**
  Die äußerste Schicht steht für öffentliches Wissen – offene Daten und generalisierte Sprachmodelle, die sich für Ideenfindung, erste Entwürfe und Exploration eignen. Generative KI ähnelt dem Internet selbst: weitreichend, vernetzt und kreativ, aber mit begrenztem organisatorischem Kontext und eingeschränkter Präzision.
- **Intranet = Agentenbasierte KI**
  Die mittlere Schicht entspricht dem internen Netzwerk einer Organisation. Hier sind Daten privat, zugriffsbeschränkt und in Workflows eingebettet. Agentenbasierte KI analysiert interne Systeme, führt genehmigte Prozesse aus und trifft kontrollierte Entscheidungen im Rahmen festgelegter Governance-Regeln. Hier hört die KI auf, nur zu beschreiben, und fängt an zu handeln, Aufgaben zu automatisieren, Personal aufzustocken und Richtlinien durch konkrete Maßnahmen durchzusetzen.
- **Extranet = Allgemeine Künstliche Intelligenz (AGI)**
  Die innerste Schicht stellt die Zukunftsgrenze dar, in der KI sicher organisations- und systemübergreifend zusammenarbeitet. Wie ein Extranet, das vertrauenswürdige Partner miteinander verbindet, würde AGI fließend über die Grenzen hinweg denken und Erkenntnisse weitergeben, während gleichzeitig das Vertrauen und die Compliance zwischen den Entitäten gewahrt bleiben.

Mit der Ausbreitung von KI nach innen – von öffentlichen über private hin zu gemeinsam genutzten Bereichen – steigen Kontexttiefe, Präzision und Wert. Gleichzeitig wächst die Notwendigkeit verantwortungsvoller Governance. Je stärker Intelligenz in den Kerndaten eines Unternehmens verankert ist, desto größer ist die Pflicht, sie zu sichern, zu prüfen und mit menschlichen sowie regulatorischen Grenzen in Einklang zu bringen.

In Zahlen: Die Energiekosten des KI-Trainings

Das Training eines Basismodells wie GPT-3 verbrauchte rund **1,287 MWh** Strom und verursachte etwa **502 Tonnen CO_2**–das entspricht in etwa den Jahresemissionen von 112 Benzinfahrzeugen.

Eine Studie aus dem Jahr 2024 ergab, dass bis zu **30 % des Stroms** der beim Training großer Sprachmodelle verbraucht wird, durch ineffiziente Planung und Hardware-Nutzung verloren gehen. Das bedeutet, dass dieselbe Leistung mit deutlich weniger Energie erzielt werden könnte.

Prognosen zufolge wird der Energiebedarf für das Training von KI-Systemen im Jahr 2024 etwa **acht Terawattstunden (TWh)** betragen und bis 2030 auf **652 TWh** steigen – eine **mehr als 80-fache Zunahme** des Stromverbrauchs innerhalb von nur sechs Jahren.

Datenqualität und die Physik des Lernens

„Wenn man Müll hineingibt, kommt auch wieder Müll heraus" – selten war dieser Satz zutreffender als im Zeitalter der KI. Die Genauigkeit eines jeden Modells hängt direkt von der Qualität der verwendeten Daten ab. Statistisch gesehen erhöhen mehr Daten die Wahrscheinlichkeit besserer Ergebnisse – vorausgesetzt, sie sind relevant, konsistent und sauber. Modernes maschinelles Lernen arbeitet nicht bloß mit Datenmengen, sondern mit Signalen. Durch wiederholtes Training und Feedback lernt das System, bestimmten Mustern Gewicht zu geben. Mit der Zeit entwickelt es ein Gespür dafür, was sinnvoll und was bloß Rauschen ist – ähnlich wie ein Mensch durch Erfahrung lernt.

Sind die Eingangsdaten unvollständig, widersprüchlich oder schlecht gepflegt, füllt das Modell die Lücken selbstständig. Das Ergebnis sind sogenannte Halluzinationen: selbstsichere, überzeugende Antworten, die jedoch völlig falsch sind. Mit zunehmender Modellkomplexität steigt das Risiko solcher Fehlschlüsse. Das wirksamste Gegenmittel ist die Kuratierung – die Verankerung von KI in kontrollierten, hochwertigen und vertrauenswürdigen Daten.

Dabei spielt auch das Verhältnis von Kosten zu Nutzen eine zentrale Rolle. Je besser die Daten, desto weniger Energie und Rechenzeit werden beim Training und bei der Inferenz verschwendet. Da KI-Modelle immer größer und rechenintensiver werden, wird die Physik des Lernens zunehmend eine Frage der Effizienz – nicht nur der Genauigkeit. Sorgfältig aufbereitete Daten verringern Redundanzen, reduzieren den Nachbearbeitungsaufwand und senken die Betriebskosten von KI-Systemen erheblich. Das neue Leistungsmaß liegt im Gleichgewicht zwischen Datenqualität, Trainingszeit und Energieverbrauch – ein klarer Hinweis darauf, dass bessere Governance nicht nur sicherer, sondern auch intelligenter und nachhaltiger ist.

Warum KI den Regeln der Daten folgen muss

Herkömmliche Software verarbeitete Daten und machte an dem Punkt halt. KI tut das nicht – sie erinnert sich. Jede Information, die sie aufnimmt, wird Teil ihrer inneren Landschaft und prägt ihr zukünftiges Denken, Reagieren und Handeln. Dieses Gedächtnis unterscheidet KI von allen bisherigen Technologien – und macht ihre Steuerung unverzichtbar.

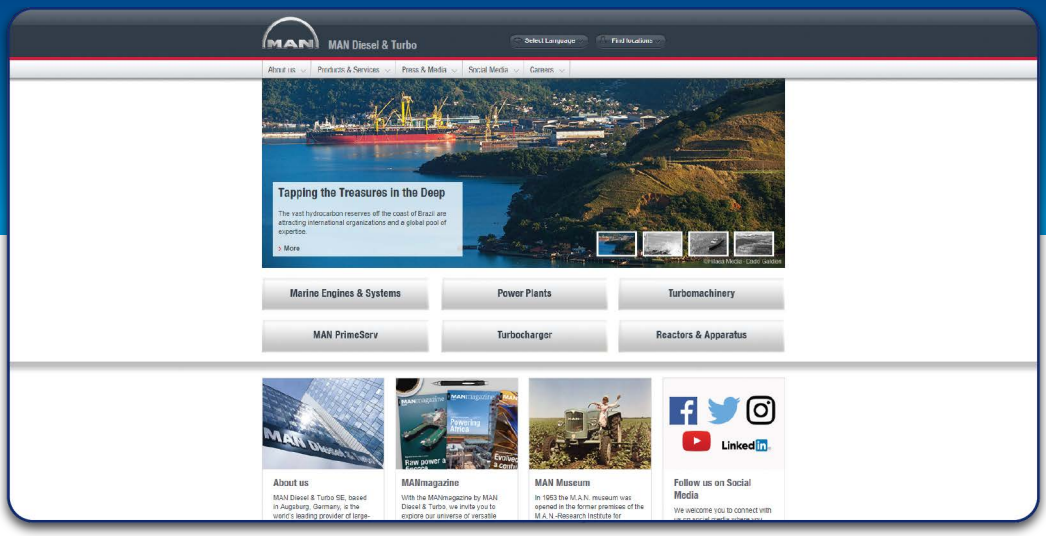
Wenn Daten schon immer Regeln brauchten, dann erweitert KI diese Regeln nun auf ein neues Gebiet. Der gleiche Lebenszyklus, der für Unternehmensinformationen gilt – Erfassung, Verwaltung, Verarbeitung, Suche und Archivierung oder Entsorgung – muss auch auf intelligente Systeme angewendet werden. Organisationen müssen festlegen, was ein Modell lernen darf, was es behalten soll, was es vergessen muss und wie sein Wissen überprüft oder auditiert werden kann.

Ohne klare Grenzen wird das Gedächtnis zur Belastung. Unkontrollierte Informationsansammlung verwandelt Wissen in Risiko – diszipliniertes Lebenszyklusmanagement dagegen in Erkenntnis und Wert. Governance bedeutet nicht mehr nur den Schutz von Daten – es geht darum, der Intelligenz beizubringen, sich verantwortungsvoll zu erinnern.

In der folgenden Fallstudie untersuchen wir, wie MAN Diesel & Turbo EIM als Grundlage für seine Governance und zur Erreichung der Compliance einsetzt.

Fallstudie

MAN Diesel & Turbo



MAN Diesel & Turbo

MAN Diesel & Turbo mit Hauptsitz in Augsburg, Deutschland, ist der weltweit führende Hersteller von großvolumigen Dieselmotoren und Turbomaschinen. Das Unternehmen beschäftigt rund 14.900 Mitarbeiter an mehr als 100 Standorten weltweit, vorwiegend in Deutschland, Dänemark, Frankreich, der Schweiz, Tschechien, Indien und China.

Die Dieselmotoren des Unternehmens – ob in Containerschiffen oder Luxuslinern – gehören zu den größten und langlebigsten technischen Produkten der Welt. Sie müssen über Jahrzehnte hinweg zuverlässig funktionieren und in festgelegten Intervallen gewartet werden. Als einer der weltweit führenden Hersteller in diesem Bereich muss MAN Diesel & Turbo wichtige technische Dokumente mindestens 30 Jahre lang – in manchen Fällen sogar unbefristet – aufbewahren. Das Unternehmen benötigte daher eine Informationsmanagementlösung, die sicherstellt, dass sowohl die Wartungsqualität gewährleistet bleibt als auch eine wirksame Abwehr potenzieller Haftungsansprüche aufgrund vermeintlicher Baumängel möglich ist.

Zur Sicherstellung der Compliance setzte MAN Diesel & Turbo auf die in einer erweiterten EIM-Lösung integrierten Funktionen zur Datensatzverwaltung sowie auf Application Governance & Archiving (AGA). Diese kombinierten Systeme führen verschiedene Anwendungen zusammen und sorgen dafür, dass Informationen im jeweiligen Kontext erhalten bleiben. Rund 1.000 Servicemitarbeiter in Deutschland und Dänemark nutzen die Lösung, um täglich mehr als 4.000 prozessbezogene Transaktionsdateien zu archivieren. In vielen Serviceprozessen gehören papierbasierte Transaktionen inzwischen der Vergangenheit an, da bestehende Akten vollständig digitalisiert wurden.

So spart MAN Diesel & Turbo wertvolle Zeit, die zuvor für das Suchen, Pflegen und Verwalten sowohl physischer als auch zahlreicher digitaler Archive erforderlich war. Die integrierte Lösung reduziert zusätzlich den Wartungsaufwand, ersetzt veraltete Systeme und ermöglicht es MAN Diesel & Turbo, seine Infrastruktur zu modernisieren, zentrale Prozesse zu digitalisieren und gesetzliche Vorgaben zuverlässig zu erfüllen.

Das Datenplateau der KI – und wie man es überwindet

Bis 2026 werden mehr als 80 Prozent der Unternehmen generative KI-Modelle oder APIs in der Produktion einsetzen.[4] Dieses Ausmaß an Implementierung erhöht den Druck auf die Governance-Strukturen erheblich. Die Ausgaben für KI-Governance-Tools werden sich bis 2030 voraussichtlich mehr als vervierfachen, da Unternehmen daran arbeiten, Risiken, Zugriffsrechte, Datenherkunft und Modellüberwachung beim Übergang von der Experimentierphase zur vollständigen Integration zu steuern.[5]

Obwohl der Einsatz generativer KI in Unternehmen inzwischen weit verbreitet ist, basieren viele dieser Anwendungen weiterhin auf öffentlichen Daten und generischen Modellen, die hervorragend zur Inhaltserstellung, aber unzureichend für operative Umsetzung sind. Im Jahr 2024 nutzten rund 65 Prozent der Unternehmen regelmäßig GenAI – ein Anteil, der bis 2025 weiter gestiegen ist. Doch viele verharren auf dem Niveau „guter Demos“, ohne messbare operative Wirkung zu erzielen.[6] Die Kluft ist nicht fehlender Enthusiasmus; es sind die Daten. Damit KI im Unternehmenskontext echten Nutzen entfalten kann, benötigt sie kontrollierten Zugriff auf private und freigegebene Informationen, um kontextbezogen argumentieren und Workflows sicher ausführen zu können.

Die Erfahrung zeigt: Skalierung ohne solide Grundlage führt selten zu nachhaltigen Ergebnissen. Unternehmen mit ausgereiften Daten- und KI-Kompetenzen übertreffen ihre Mitbewerber beständig.[7] Der Unterschied liegt darin, dass sie Datenstrategie und Governance als Grundprinzipien begreifen – nicht als nachträgliche Ergänzung. In der Praxis bedeutet das, private Inhalte unter Kontrolle zu halten, Zugriffsrechte konsequent durchzusetzen und richtliniengestützte Metadaten anzuwenden. So können Modelle relevante Informationen gezielt abrufen, innerhalb festgelegter Vorgaben handeln und Verantwortlichkeit nachweisen.

Wenn KI innerhalb der Firewall arbeitet – also mit dem verwalteten Datenbestand des Unternehmens verbunden ist –, hört sie auf zu raten und beginnt, gezielt zu handeln. Sie bietet keine allgemeinen Antworten, sondern fundierte, nachvollziehbare Maßnahmen: Eine Rechnungsabweichung kann durch Bezugnahme auf Bestellung, Lieferantenbedingungen und Genehmigungshistorie geklärt werden. Eine Richtlinienaktualisierung lässt sich automatisch entwerfen und weiterleiten, wobei Berechtigungen, Aufbewahrungsfristen und regulatorische Anforderungen berücksichtigt werden. Supportanfragen in natürlicher Sprache können auf Basis interner Wissensbestände beantwortet und gleichzeitig im System protokolliert werden.

Jede dieser Funktionen basiert auf demselben EIM-Backbone, der Fragmentierung reduziert, Zugriff regelt und unstrukturierte Informationen mit den zugrunde liegenden Geschäftsprozessen verbindet. Vereinfacht gesagt – eine EIM-Plattform sorgt für Ordnung. Sie steuert den Informationsfluss über Systeme, Silos und Regionen hinweg.
KI liefert den Kontext. Sie lernt aus diesen kontrollierten Daten, um Erkenntnisse, Automatisierung und Entscheidungsunterstützung zu ermöglichen.

Im kognitiven Zeitalter wird KI die nächste Generation des Informationsmanagements prägen. Die Schlussfolgerung ist eindeutig: Ohne private, genehmigte Daten – und ohne Governance, die deren verantwortungsvolle Nutzung sicherstellt – stößt generative KI an ihre Grenzen. Sie kann das Internet zusammenfassen, aber keine Rechnung genehmigen, Reparatur planen oder Kundenanfrage im System bearbeiten. Der Weg liegt in katalogisierten, klassifizierten und zugriffskontrollierten Unternehmensdaten, unterstützt durch prüfbare Datenpipelines, die KI-Agenten befähigen, Fakten abzurufen, definierte Aktionen auszuführen und dabei eine nachvollziehbare Spur zu hinterlassen. So verwandeln Unternehmen breite Akzeptanz in nachhaltigen Geschäftswert.

Die fünf Merksätze

1. **KI-Reife beginnt mit Datenreife.**
   Künstliche Intelligenz ist nur so leistungsfähig wie die Daten, aus denen sie lernt. Generative Modelle, die auf öffentlichen Daten beruhen, stoßen an ihre Grenzen; agentenbasierte KI erfordert kontrollierte, private und genehmigte Informationen. Investition in Datengrundlagen ist daher die Voraussetzung, bevor KI-Fähigkeiten sinnvoll ausgebaut werden können. EIM hilft, hochwertige private Datensätze zu identifizieren und KI dort einzusetzen, wo der Prozessnutzen eindeutig ist.

2. **Governance ist die neue Infrastruktur.**
   Die Prinzipien, die für Datenkonformität sorgen – Metadaten, Berechtigungen, Lebenszykluskontrolle und Prüfbarkeit – sind heute die Voraussetzung für KI. Governance definiert, wie Intelligenz innerhalb der Organisation und über ihre Grenzen hinaus lernt, sich erinnert und sicher handelt.

3. **Für mehr Wert nach innen gehen: Internet → Intranet → Extranet.**
   Öffentliche Daten bilden die Grundlage für generative KI (Inhalte), interne Daten für agentenbasierte KI (Aktionen), und vernetzte Ökosysteme werden eines Tages AGI (Kollaboration) ermöglichen. Jeder Schritt nach innen steigert Genauigkeit, Verantwortlichkeit und Wert – und erfordert zugleich stärkere Kontrollen.

4. **Datenqualität bestimmt die Leistungsfähigkeit von KI – und ihren ökologischen Fußabdruck.**
   Sauber aufbereitete Daten reduzieren Fehlalarme, verbessern Zuverlässigkeit und senken den Rechenaufwand. Die neue Leistungskennzahl berücksichtigt Datenqualität, Trainingszeit und Energieverbrauch gleichermaßen.
   Bessere Datenverwaltung heute bedeutet schnellere, präzisere und ressourcenschonendere KI in der Zukunft.

5. **KI muss den Regeln der Daten folgen.**
   Im Gegensatz zu herkömmlicher Software merkt sich KI, was sie sieht. Sie macht damit ihr Gedächtnis zu einem Teil der Governance-Landschaft. Das Lernen von KI folgt einem Lebenszyklus: Entscheiden, welche Modelle lernen dürfen, was sie behalten, was sie vergessen müssen und wie ihr Wissen überprüft wird.

Kapitel Zwei

Der Aufstieg der Enterprise Artificial Intelligence

Mit dem Fortschreiten der technologischen Entwicklung hat künstliche Intelligenz das Potenzial, zahlreiche Lebensbereiche grundlegend zu verändern. Von der Steigerung der Produktivität am Arbeitsplatz bis hin zur Veränderung der Art, wie Menschen mit Informationen und miteinander interagieren wird KI zunehmend zu einem zentralen Bestandteil des persönlichen und beruflichen Alltags.

Enterprise Artificial Intelligence (EAI) definiert Unternehmensleistung neu, indem sie Intelligenz in Kontext verwandelt. Mit der zunehmenden Verbreitung von Technologien in den Bereichen Kundenerlebnis, Betriebsabläufe und Content-Publishing liegt der eigentliche Vorteil nicht allein in der Automatisierung, sondern im Kontextverständnis. In diesem Kapitel werden die wichtigsten Konzepte der KI-Technologie, ihre grundlegenden Prinzipien und ihre Anwendungen in verschiedenen Sektoren erläutert.

> *Die Zahl der Unternehmen, die ihre Prozesse vollständig modernisiert und KI-gestützte Verfahren eingeführt haben, hat sich von 9 % im Jahr 2023 auf 16 % im Jahr 2024 nahezu verdoppelt. Im Vergleich zu ähnlichen Unternehmen erzielen diese Organisationen ein 2,5-fach höheres Umsatzwachstum, eine 2,4-fach höhere Produktivität und einen 3,3-fach höheren Erfolg bei der Skalierung von Anwendungsfällen für generative KI.*[8]

Kontextuelle Intelligenz ermöglicht es der KI, die geschäftliche Absicht zu erfassen, indem sie nicht nur Daten interpretiert, sondern auch die Strukturen, Arbeitsabläufe und Ziele, die festlegen, wie eine Organisation Mehrwert schafft. Durch die Abbildung von Beziehungen zwischen Kennzahlen, Prozessen und Geschäftslogik kann KI Ergebnisse antizipieren, Abhängigkeiten modellieren und Maßnahmen empfehlen, die mit strategischen Prioritäten im Einklang stehen. So schließt sie die Lücke zwischen Analyse und Umsetzung und verwandelt Erkenntnisse in messbare Entscheidungen, die Wachstum, Effizienz und Differenzierung fördern.

Neben den Vorteilen müssen jedoch auch einige wichtige ethische Überlegungen und Implikationen berücksichtigt werden. Indem wir die duale Natur der KI verstehen – ihr Innovationspotenzial und ihre transformative Kraft –, können wir uns besser auf eine Zukunft vorbereiten, in der diese Technologie eine zentrale Rolle spielt. Nachdem im ersten Kapitel der Fokus auf Daten lag, geht es hier darum, wie vertrauenswürdige und sichere Informationen die Grundlage für eine effektive KI bilden. Gute Daten sind der Schlüssel, um Innovationen zu fördern und gleichzeitig die negativen Auswirkungen einer unkontrollierten KI-Verbreitung zu vermeiden.

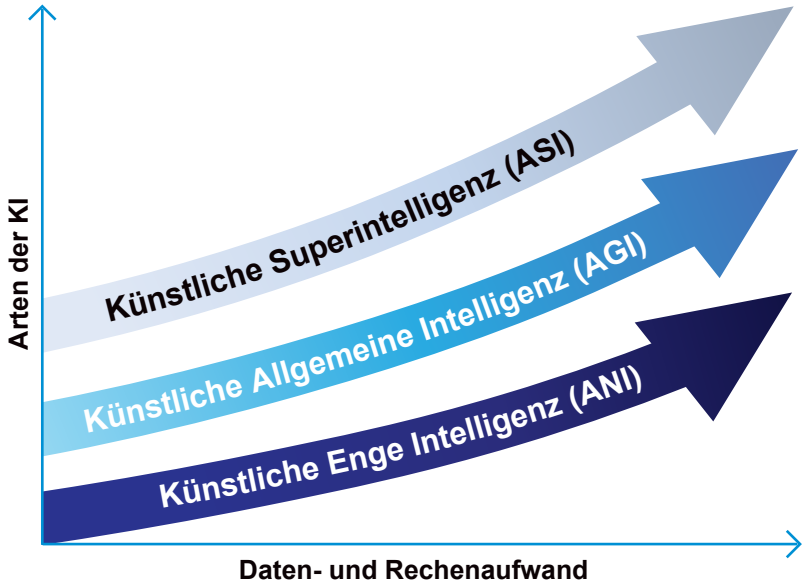
Definition von KI

Es gibt zahlreiche Definitionen von KI, doch laut der Internationalen Organisation für Normung (ISO) ist „Künstliche Intelligenz (KI) ein Zweig der Informatik, der Systeme und Software entwickelt, die in der Lage sind, Aufgaben zu bewältigen, die einst als ausschließlich menschlich galten." KI ermöglicht es Maschinen, aus Erfahrung zu lernen, sich an neue Informationen anzupassen und mithilfe von Daten, Algorithmen und Rechenleistung komplexe Situationen zu interpretieren und Entscheidungen mit minimalem menschlichem Eingriff zu treffen."[9]

KI ist kein einheitliches Konzept und keine einzelne Technologie. Vielmehr handelt es sich um ein breites Feld mit verschiedenen Teilgebieten, von denen jedes eigene Ziele und Spezialisierungen hat. Unter dem Oberbegriff „Künstliche Intelligenz" werden mehrere Technologien zusammengefasst – darunter maschinelles Lernen, Deep Learning und die Verarbeitung natürlicher Sprache (Natural Language Processing, NLP).[10]

Künstliche Intelligenz lässt sich im Allgemeinen in drei Hauptkategorien einteilen, nämlich in spezialisierte Intelligenz, allgemeine Intelligenz und Superintelligenz. Diese unterscheiden sich vor allem darin, wie sie lernen und welche Aufgaben sie bewältigen können. Die künstliche domänen spezifische Intelligenz (Artificial Narrow Intelligence, ANI) lässt sich am besten mit einem Schüler vergleichen, der in einem bestimmten Fachgebiet herausragend ist, dessen Wissen sich jedoch nicht auf andere Bereiche übertragbar lässt. Ein System dieser Art mag zum Beispiel hervorragend Schach spielen, Gesichter erkennen oder Verkehrsmuster vorhersagen können – doch außerhalb seines Spezialgebiets stößt es schnell an Grenzen. DeepMinds AlphaGo und sein Nachfolger AlphaZero markierten bedeutende Meilensteine in der KI-Entwicklung, indem sie komplexe Spiele mithilfe von Selbstlernverfahren und verstärkendem Lernen meisterten. Sie demonstrierten eine beeindruckende Fähigkeit, innerhalb klar definierter Bereiche zu generalisieren, blieben jedoch weiterhin fest im Bereich der engen KI (ANI) und erreichten keine echte künstliche allgemeine Intelligenz (AGI).

Künstliche allgemeine Intelligenz (AGI) lässt sich mit einem vielseitig gebildeten Doktoranden vergleichen. Sie kann, ähnlich wie ein Mensch, neue Themen erfassen, Ideen miteinander verknüpfen und Probleme in unterschiedlichen Disziplinen lösen. Darüber hinaus bezeichnet die künstliche Superintelligenz (ASI) eine hypothetische Form der Intelligenz, die den Menschen in nahezu allen Bereichen übertreffen würde. Sie wäre einem Genie vergleichbar, das in logischem Denken, Kreativität und Selbstverbesserung überlegen ist, eine Vorstellung, die zwar faszinierend, bislang jedoch rein theoretisch bleibt.



KI lässt sich in drei Kategorien einteilen

Der *AI-Watch*-Bericht der Europäischen Kommission beschreibt künstliche spezialisierte Intelligenz (ANI) als Systeme, das „eine spezifische Aufgabe ausführen und innerhalb einer vordefinierten Umgebung operieren können." ANI kann Daten mit hoher Geschwindigkeit verarbeiten und dadurch Produktivität und Effizienz in zahlreichen praktischen Anwendungen steigern. Obwohl ANI in spezialisierten Bereichen leistungsfähig ist, fehlt ihr die Fähigkeit zur Generalisierung, also zur Übertragung erlernter Kenntnisse auf andere Domänen."[11]

Der Bericht führt weiter aus: „AGI bezieht sich auf Maschinen, die menschliche Intelligenz aufweisen." Mit anderen Worten: AGI soll in der Lage sein, jede intellektuelle Aufgabe zu bewältigen, die auch ein Mensch ausführen kann."[12] Derzeit befinden wir uns jedoch nicht im Stadium der AGI, da die Intelligenz heutiger KI-Systeme keine echte menschliche, sondern eine simulierte Form von Intelligenz ist. Um AGI vollständig zu erreichen, müssten KI-Systeme in der Lage sein, neue Aufgaben ohne erneutes Training zu erlernen, autonom zu denken und Ursache-Wirkungs-Zusammenhänge im Kontext zu verstehen.

ASI ist in bestehenden Standards noch nicht definiert und gilt als zukünftiger Zustand jenseits der AGI: „Künstliche Superintelligenz (ASI) ist ein hypothetisches, softwarebasiertes System künstlicher Intelligenz mit einem intellektuellen Umfang, der über die menschliche Intelligenz hinausgeht." Auf ihrer grundlegendsten Ebene würde eine solche superintelligente KI über fortgeschrittene kognitive Fähigkeiten und Denkvermögen verfügen, die weit über das menschliche Maß hinausgehen."[13]

Künstliche Intelligenz lässt sich auf zwei Arten klassifizieren: nach ihren Fähigkeiten – also danach, wie nahe sie der menschlichen Kognition kommt – und nach ihrer Funktionalität, also danach, wie sie sich verhält und mit Daten interagiert. Der KI-Forscher Arend Hintze entwickelte ein weithin anerkanntes funktionales Rahmenwerk, das beschreibt, wie Systeme Informationen verarbeiten und auf ihre Umgebung reagieren.

Nach Hintzes Modell arbeiten reaktive Maschinen auf der grundlegendsten Ebene ausschließlich mit aktuellen Eingaben und sind nicht in der Lage, aus vergangenen Erfahrungen zu lernen. Ein bekanntes Beispiel ist IBMs Schachsystem Deep Blue. KI mit begrenztem Speicher kann kurzfristige Daten speichern, um daraus fundierte Entscheidungen abzuleiten, und bildet die Grundlage nahezu aller modernen KI-Systeme – von autonomen Fahrzeugen bis zu Empfehlungssystemen. Darüber hinaus bewegt sich das Feld in theoretisches Terrain: Die „Theory of Mind"-KI beschreibt Systeme, die menschliche Überzeugungen, Emotionen und Absichten verstehen könnten, während die selbstbewusste KI eine hypothetische Stufe darstellt, auf der Maschinen über echtes Bewusstsein und Selbstwahrnehmung verfügen. Auch wenn diese höheren Ebenen bislang spekulativ bleiben, hilft das Verständnis dieses Spektrums, den heutigen Stand der Unternehmenssysteme einzuordnen. Danach bewegen sie sich überwiegend im Bereich des begrenzten Speichers, wo Wertschöpfung durch verantwortungsvolle Datennutzung, gesteuertes Lernen und disziplinierte, skalierbare Bereitstellung entsteht.[14]

Mit der fortschreitenden Entwicklung der KI und der Einteilung in unterschiedliche Kategorien hat sich ein zentrales Prinzip herausgebildet: Daten, Rechenleistung und Governance bilden die Kernbestandteile aller KI-Formen. Modelle und Fähigkeiten mögen sich verändern – doch die Fähigkeit, Informationen zu organisieren, abzusichern und operativ nutzbar zu machen, bleibt der entscheidende Wettbewerbsvorteil. An dieser Stelle kommt das Enterprise Information Management (EIM) ins Spiel – als Grundlage für die Anwendung von Künstlicher Intelligenz im Unternehmenskontext, also für Enterprise Artificial Intelligence (EAI).

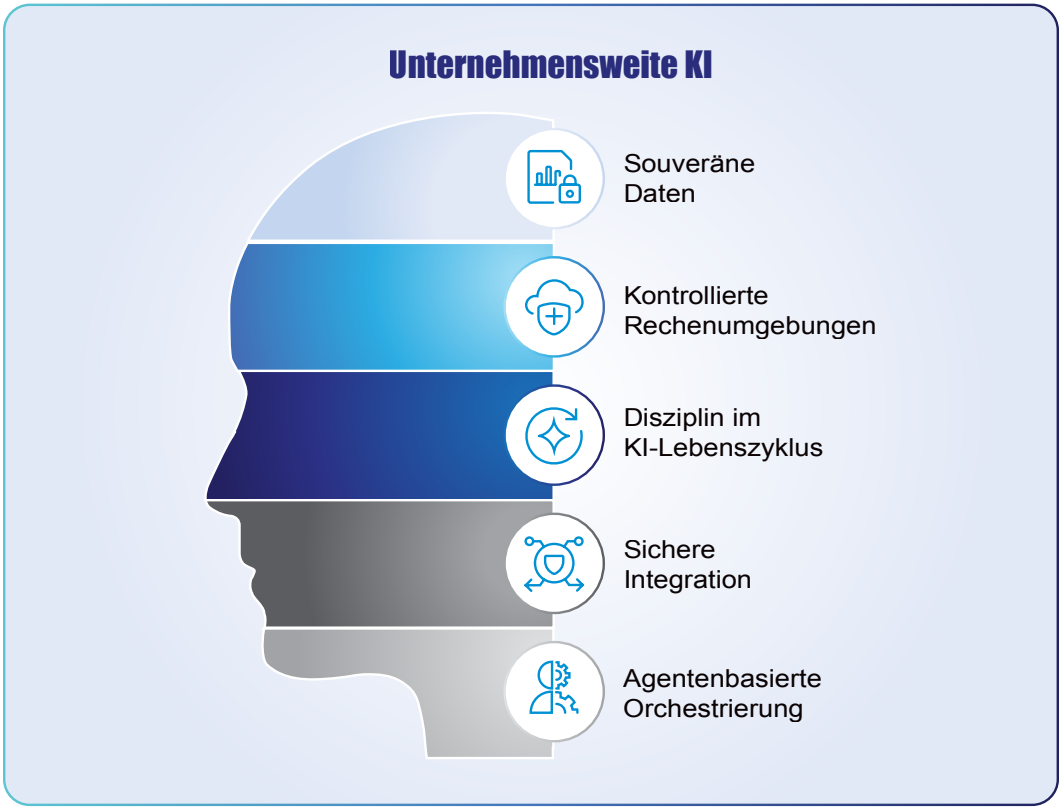
Künstliche Intelligenz für Unternehmen

Enterprise-KI ist keine eigene Intelligenzklasse, sondern beschreibt die strategische Anwendung und Integration verschiedener KI-Technologien und -Fähigkeiten innerhalb einer Organisation. Ziel ist es, spezifische Probleme zu lösen, Prozesse zu automatisieren und Entscheidungen datenbasiert zu unterstützen. EAI fällt überwiegend unter den Bereich der Künstlichen engen Intelligenz (ANI), da diese Systeme für spezialisierte Aufgaben konzipiert sind, um Abläufe zu optimieren, nicht um eine menschenähnliche Allgemeinintelligenz oder ein Bewusstsein zu entwickeln.

Enterprise-KI stützt sich auf:

- Vertrauenswürdige und kontrollierte Daten (die „souveräne Datenebene“)
- KI-Lebenszyklusmanagement-Plattformen (z. B. MLOps, LLMOps)
- Hybride oder souveräne Cloud-Infrastrukturen
- Sichere APIs und Orchestrierungsschichten
- Agentenbasierte KI-Systeme, die mehrere spezialisierte Modelle koordinieren

Enterprise-KI ist somit eine geregelte Architektur – kein einzelnes Modell. Sie vereint souveräne Daten, kontrollierte Rechenleistung, Disziplin im KI-Lebenszyklus, sichere Integration und agentenbasierte Orchestrierung, um vertrauenswürdige Automatisierung im großen Maßstab zu ermöglichen. Enterprise-KI ist damit der Einsatzkontext, in dem bewährte KI-Technologien und -Fähigkeiten operationalisiert werden.



Enterprise-KI

Um echten Mehrwert zu schaffen, stützen sich Enterprise-AI-Lösungen (EAI) auf ein breites Spektrum spezialisierter Technologien:

- **Maschinelles Lernen (ML)** ermöglicht prädiktive Analysen – etwa zur Vorhersage von Geräteausfällen, Nachfrageprognosen oder Optimierung von Lagerbeständen.
- **Verarbeitung natürlicher Sprache (NLP)** unterstützt intelligente Chatbots, die Zusammenfassung von Dokumenten und die Analyse von Kundenstimmungen.
- **Computer vision** automatisiert Prozesse in Bereichen wie Fertigungsinspektion oder Sicherheitsüberwachung.
- **Robotergestützte Prozessautomatisierung (RPA)** optimiert strukturierte, wiederkehrende Aufgaben wie Dateneingabe oder Rechnungsabgleich. Generative KI gewinnt zunehmend an Bedeutung bei der Inhaltserstellung, Codegenerierung und Wissensunterstützung.

Was KI im Unternehmen auszeichnet, ist nicht der Modelltyp, sondern die gesteuerte Integration dieser Technologien in Geschäftsprozesse, Datensysteme und Entscheidungsstrukturen. Erfolg entsteht nicht durch isolierte Modelle, sondern durch deren verantwortungsvolle Orchestrierung im großen Maßstab – auf Basis von vertrauenswürdigen Daten, sicherer Infrastruktur und starker Informationsgovernance. KI für Unternehmen unterscheidet sich sowohl im Zweck als auch im Design grundlegend von KI für Endverbraucher. Consumer AI zielt darauf ab, individuelle Erlebnisse zu verbessern: etwa durch Empfehlungen von Filmen, Unterstützung bei persönlichen Aufgaben oder virtuelle Assistenten. Diese Systeme arbeiten in der Regel im kleinen Maßstab, nutzen öffentliche oder benutzergenerierte Daten und erfordern nur minimale Integration mit anderen Tools. Ihr Wert liegt in Bequemlichkeit und Personalisierung für den einzelnen Nutzer.

Enterprise-KI hingegen ist auf Skalierbarkeit, Sicherheit und strategische Wirkung ausgelegt. Sie arbeitet mit sensiblen, firmeneigenen Daten, die in CRM-Systemen, ERP-Lösungen und anderen operativen Datenbanken gespeichert sind, und muss strenge Governance-, Compliance- und Cybersicherheitsanforderungen erfüllen. Enterprise-KI ist tief in bestehende Systeme und Arbeitsabläufe integriert, automatisiert komplexe, abteilungsübergreifende Prozesse und liefert messbare Ergebnisse wie Effizienzsteigerung, Risikominimierung, Kostensenkung und Innovation. Im Wesentlichen ist Enterprise-KI die industrielle Anwendung moderner KI-Technologien, entwickelt für den Einsatz in groß angelegten Umgebungen, in denen Genauigkeit, Verantwortlichkeit und Vertrauen ebenso wichtig sind wie Intelligenz selbst.

In der folgenden Fallstudie sorgt ein internationaler Flughafen mithilfe von KI im Unternehmen dafür, dass 90 Millionen Passagiere weltweit reibungslos befördert werden.

Fallstudie

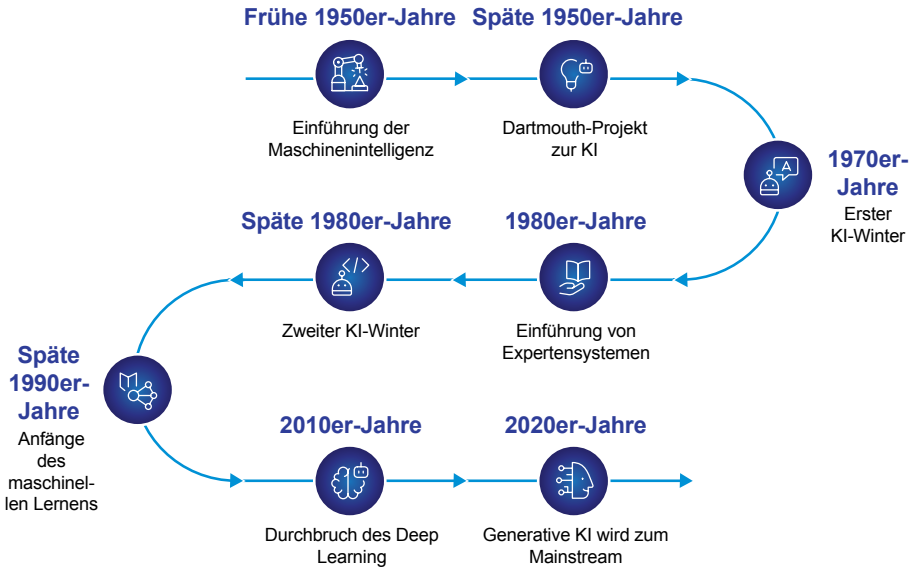
Ein internationaler Flughafen

> *Ein Großteil unserer Daten war über mehrere Systeme verteilt, und die Gewährleistung der Genauigkeit – insbesondere bei der Erfassung von Passagierströmen und Bearbeitungszeiten – stellte eine echte Herausforderung dar. Dem Personal fehlte oft Echtzeit-Einblicke oder Prognoseinstrumente, um Warteschlangen, Personalbesetzung und Engpässe proaktiv zu steuern.*
>
> Leiter IT-Servicemanagement am Flughafen

Mit mehr als 90 Millionen Reisenden pro Jahr ist dieser Flughafen einer der weltweit verkehrsreichsten im internationalen Passagierverkehr – und einer der modernsten im digitalen Bereich. Er hat sich zu einem globalen Drehkreuz entwickelt, das für Innovation, Effizienz und hohen Servicestandard steht. Seit seiner Eröffnung im Jahr 1960 wurde er kontinuierlich erweitert und zusätzlich zu den Terminals wurden modernisierte Abflughallen sowie neue Start- und Landebahnen hinzugefügt.

Eine digitale Grundlage für dieses Wachstum zu schaffen, war eine anspruchsvolle Aufgabe. Da die Beteiligten von Fluggesellschaften über Polizei und Zoll bis hin zu Dienstleistern reichen, ist die Kommunikation komplex. Für eine koordinierte Entscheidungsfindung benötigen alle denselben Zugriff auf relevante Daten. Im Rahmen einer umfassenden Initiative zur Verbesserung des Servicemanagements ging der Flughafen eine Partnerschaft mit einem Technologieanbieter ein, um seine Überwachungsmöglichkeiten auszubauen. Eine KI-gestützte Komponente für das Betriebsmanagement ermöglicht die zentrale und intelligente Steuerung komplexer IT-Umgebungen. Sie verbessert die Beobachtbarkeit, reduziert Fehlalarme, prognostiziert Probleme und unterstützt die Sicherstellung der Betriebszeit.

Echtzeit-Einblicke und intelligentes Monitoring haben die IT von einer reinen Backend-Funktion zu einem Faktor für Servicequalität, Vertrauen und Kundenzufriedenheit gemacht. Durch den Einsatz von KI auf Basis einer EIM-Infrastruktur konnte der Flughafen 30 Prozent der Vorfälle durch proaktive Überwachung vermeiden, den IT-Betrieb enger an die Geschäftsanforderungen anpassen und den Kundenservice durch IT-Exzellenz verbessern.



Die Entwicklung der modernen KI

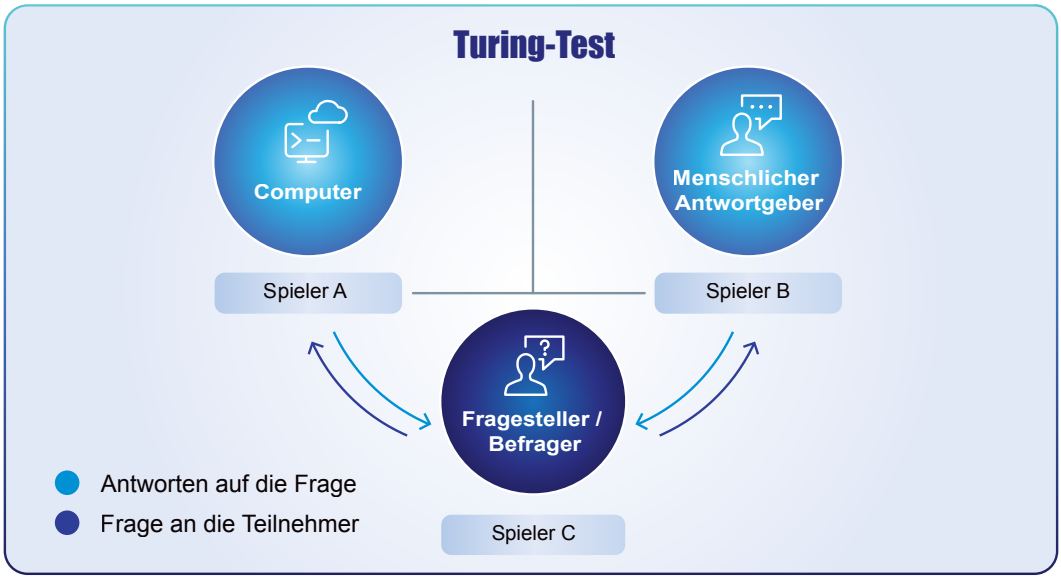
In den vergangenen 75 Jahren hat sich die Künstliche Intelligenz zu dem entwickelt, was wir heute kennen – von Phasen bahnbrechender Innovation bis zu Zeiten überzogener Erwartungen, während sich die Technologie stetig weiterentwickelte. Ein Blick zurück zeigt, wie dieser Weg begann.

Anfang der 1950er Jahre: Alan Turing gilt als einer der frühen Wegbereiter der KI. In seinem 1950 veröffentlichten Artikel „Computing Machinery and Intelligence" stellte er die These auf, dass Maschinen menschliches Denken simulieren könnten, und führte als Maßstab für maschinelle Intelligenz den nach ihm benannten Turing-Test ein.[15]

Ende der 1950er Jahre: Im Jahr 1956 wurde der Begriff künstliche Intelligenz im Rahmen des Dartmouth Summer Research Project on Artificial Intelligence geprägt. Diese von John McCarthy, Marvin Minsky, Nathaniel Rochester und Claude Shannon organisierte Konferenz gilt als Ausgangspunkt der KI-Forschung. Die Veranstaltung brachte Forscher zusammen, um das Ziel zu formalisieren, Maschinen zu schaffen, die in der Lage sind, ähnlich wie Menschen zu denken und zu lernen.[16]

1970er Jahre: Die anfängliche Euphorie über die Fortschritte der KI flaute jedoch bald ab. Die Forschung kam nur langsam voran, die Ergebnisse blieben hinter den Erwartungen zurück, und viele frühe Projekte erwiesen sich als nicht umsetzbar. Der Begriff „KI-Winter" bezeichnet eine Phase, in der die Kritik am mangelnden Fortschritt – unter anderem durch den Lighthill Report im Vereinigten Königreich im Jahr 1973 – dazu führte, dass staatliche Fördermittel gestrichen wurden.[17]

1980er Jahre: Nach dem KI-Winter der 1970er Jahre erlebte die Forschung in den 1980er Jahren mit dem Aufkommen sogenannter Expertensysteme einen neuen Aufschwung. Diese Programme basierten auf Wenn-dann-Regeln und sollten menschliches Entscheidungsverhalten nachbilden. Expertensysteme fanden in vielen Bereichen Anwendung, und erstmals begannen Unternehmen, den kommerziellen Nutzen von KI zu erkennen. Dennoch blieb diese Form der KI eng begrenzt und aufgabenspezifisch – weit entfernt von einer echten allgemeinen Intelligenz.[18]



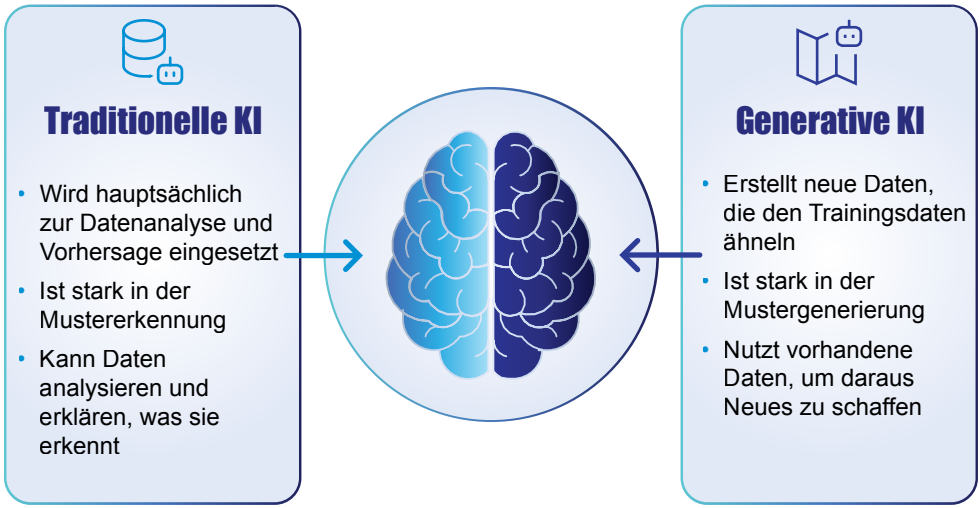
Der Turing-Test

Ende der 1980er/Anfang der 1990er Jahre: Nachdem Entwicklungen zunächst vielversprechend aussahen, folgte ein zweiter KI-Winter. Gegen Ende der 1980er-Jahre ließ die Begeisterung nach, da Unternehmen erkannten, dass Aufbau und Wartung der Systeme hohe Kosten verursachten. Auch die Grenzen der programmierten Logik traten zutage. Mit dem Rückgang der Finanzierung kam die Forschung erneut zum Erliegen.[19]

Ende der 1990er Jahre: Trotz dieses Abschwungs wurde die Arbeit an KI nicht aufgegeben. Forschende begannen, sich von starren Wenn-dann-Regeln zu lösen und entwickelten neue Ansätze, bei denen Maschinen aus Daten lernen konnten. Diese Phase markierte den Beginn des modernen maschinellen Lernens. Zu den wichtigsten Fortschritten dieser Zeit gehörten neuronale Netzwerke und Entscheidungsbäume. In dieser Zeit wurde das Lernen aus Daten auch zu einem entscheidenden Faktor in der Entwicklung der KI.[20]

2010er Jahre: Ein weiterer bedeutender Durchbruch erfolgte 2012 mit dem Aufstieg des Deep Learning. In diesem Jahr gewann ein neuronales Netzwerk namens AlexNet den ImageNet-Wettbewerb in der Bilderkennung und senkte die Fehlerrate deutlich. Dieser Erfolg markierte einen Wendepunkt, da er zeigte, dass KI den Menschen bei visuellen Erkennungsaufgaben übertreffen kann. Technologiekonzerne wie Google, Facebook und später OpenAI knüpften an diesen Fortschritt an und entwickelten Systeme, die nicht nur Bilder erkennen, sondern auch Sprachen übersetzen und Texte generieren konnten.[21]

2020er Jahre: Heute ist generative KI (GenAI) im Mainstream angekommen. Große Sprachmodelle (LLMs) wie GPT, Claude und Gemini entwickeln sich stetig weiter und erweitern die Möglichkeiten künstlicher Intelligenz. Dabei hat sich gezeigt, dass Modellgröße, Trainingsdaten und Rechenleistung entscheidende Faktoren für die Leistungsfähigkeit sind. Dank dieser Fortschritte machen immer mehr Menschen alltägliche Erfahrungen mit KI, was die Akzeptanz deutlich beschleunigt. Auch im Unternehmensumfeld wird KI zunehmend zu einem entscheidenden Differenzierungsfaktor für die Geschäftsleistung.



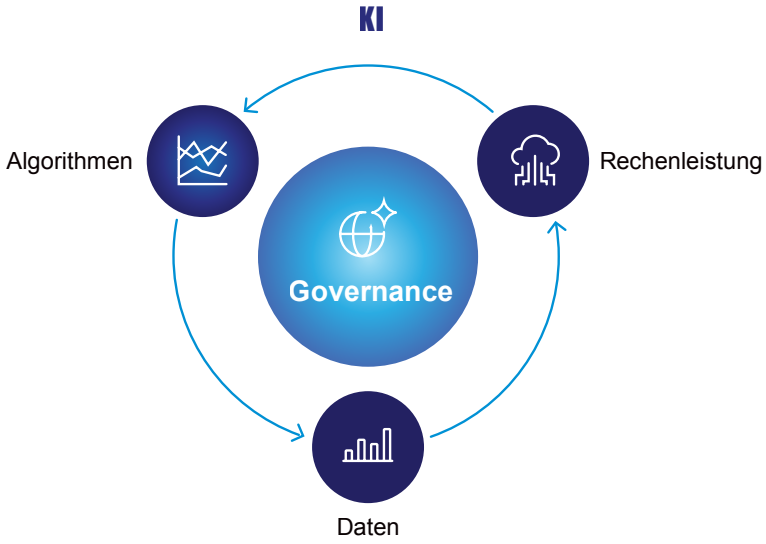
Die Unterschiede zwischen traditioneller und generativer KI[22]

Daten, Rechenleistung, Algorithmen und Governance als Treibstoff für die moderne KI-Engine im Unternehmen

In den 2020er-Jahren hat sich die künstliche Intelligenz durch das Zusammenspiel von Daten, Rechenleistung und Algorithmen – unter der übergeordneten Disziplin der Governance – entscheidend weiterentwickelt. Daten bilden den grundlegenden Treibstoff, der KI-Systeme befähigt, zu lernen, sich anzupassen und Wissen über verschiedene Bereiche hinweg zu verallgemeinern. Ihre Qualität, Kennzeichnung und Integration bestimmen, wie wirksam die Modelle arbeiten: Vielfältige und gut verwaltete Datensätze führen zu präziseren und stabileren Ergebnissen.

Rechenleistung ermöglicht Skalierung. Fortschritte bei der Hardware – insbesondere Grafikprozessoren (GPUs), Tensor-Prozessoren (TPUs) und elastische Cloud-Infrastrukturen – ermöglichen Trainingsvolumen, die noch vor wenigen Jahren undenkbar waren.

Parallel dazu hat sich die algorithmische Innovation rasant entwickelt und zur Entstehung von Grundlagenmodellen geführt, die heutigen generativen und multimodalen KI-Systemen zugrunde liegen. Diese drei Komponenten – Daten, Rechenleistung und Algorithmen – bilden gemeinsam den technischen Kern der modernen KI. Doch erst Governance verleiht dieser Architektur die notwendige Integritätsebene, die verantwortungsvolles Handeln sicherstellt. Wenn alle Kräfte zusammenwirken, entsteht ein Gleichgewicht aus Leistungsfähigkeit, Vertrauen, Compliance und nachhaltiger Effizienz.



KI = Die Kombination aus Daten, Rechenleistung und Algorithmen

Wie die Geschichte zeigt, verlief die Entwicklung der künstlichen Intelligenz selten geradlinig. Der Fortschritt bewegte sich in Zyklen aus Optimismus und Korrektur – auf Phasen intensiver Innovation folgten stets Phasen der Neubewertung. Frühe Durchbrüche führten oft zu überhöhten Erwartungen, die in Ernüchterung umschlugen, sobald die Ergebnisse hinter den Prognosen zurückblieben. Doch gerade diese Schwankungen waren entscheidend für die Reifung des Fachgebiets. Sie zwangen Forschung und Industrie dazu, Vision und Realität miteinander in Einklang zu bringen.

Mit der Zeit setzte sich eine zentrale Erkenntnis durch: Nachhaltige KI-Fähigkeiten von Unternehmen hängen von einem Gleichgewicht ab. Echter Fortschritt entsteht aus dem Zusammenspiel von strukturierten Daten, skalierbarer Rechenleistung und fortlaufend optimierten Algorithmen – eingebettet in einen klar definierten Governance-Rahmen. Organisationen, die diese Komponenten miteinander verbinden, sind in der Lage, Experimente in messbare Geschäftsergebnisse zu überführen.

Im folgenden Fallbeispiel zeigt sich, wie ein medizinisches Forschungsunternehmen mithilfe von Enterprise-KI-Lösungen klinische, finanzielle und patientenbezogene Ergebnisdaten auf nationaler Ebene verknüpft. Das Ergebnis ist eine ganzheitliche Analyse, die den Wandel hin zu einer wertorientierten Gesundheitsversorgung unterstützt.

Fallstudie

Eine landesweite Gesundheitsplattform

Ein niederländisches medizinisches Forschungsunternehmen ermöglicht es seinen Nutzern – darunter Krankenhäuser, Regierungsbehörden, Pharmaunternehmen und Versicherungsgesellschaften –, unter strengen Datenschutzbestimmungen ihre Leistung, Patientenerfahrungen und Behandlungsergebnisse über verschiedene Einrichtungen hinweg zu vergleichen. Durch die Kombination von klinischen, finanziellen und patientenbezogenen Ergebnisdaten unterstützt die Organisation medizinische Fachkräfte bei der Bereitstellung einer wertorientierten Gesundheitsversorgung. Gesucht wurde eine Lösung, die ein intuitives Online-Dashboard bereitstellt und sich skalieren lässt. Die Entscheidung fiel auf eine Kombination aus Künstlicher Intelligenz (KI) und Business-Intelligence-(BI)-Reporting.

Mehr als 5.000 Nutzer greifen regelmäßig darauf zu, um Leistungen zu bewerten und Vergleiche mit anderen Einrichtungen anzustellen. Über detaillierte Analyse-Dashboards lassen sich Kennzahlen aufschlüsseln und Zusammenhänge erkennen. Durch den direkten Zugriff auf diese analytischen Informationen konnten Kliniker Bereiche identifizieren, die gut funktionieren, und Bereiche, die weiterer Betrachtung bedürfen. Sie können dann bei Bedarf entsprechende Verbesserungen vornehmen. So konnte die Zahl der Komplikationen nach Darmkrebsoperationen innerhalb von vier Jahren um mehr als die Hälfte gesenkt werden.

Da inzwischen alle niederländischen Krankenhäuser die Lösung nutzen, ist die Ausweitung auf weitere klinische Bereiche geplant. Das ganze Ausmaß der Möglichkeiten, die KI bietet, wird gerade erst deutlich. Neue Anwendungsfelder wie die Entscheidungsunterstützung, bei der Ärzte und Patienten gemeinsam optimale Behandlungswege wählen, eröffnen zusätzliche Möglichkeiten für Forschung und Praxis.

Die wachsende Rolle von KI-Agenten

KI-Systeme entwickeln sich rasant vom Konzept zur Kernfunktion. Sie verändern branchenübergreifend die Art, wie Arbeit erledigt wird – sie automatisieren wiederkehrende Aufgaben, beschleunigen strategische Abläufe und verstärken menschliches Fachwissen. Im Kundenservice erkennen Teams frühzeitig Verhaltensmuster und Abwanderungsrisiken und können Bindungsmaßnahmen einleiten, bevor ein Support-Ticket entsteht. Im Vertrieb qualifizieren KI-Agenten Leads, automatisieren Nachfassaktionen und liefern Echtzeit-Einblicke, die den Verkaufsprozess beschleunigen. Marketingabteilungen nutzen KI, um Zielgruppen präziser zu segmentieren und Kampagnen in großem Umfang zu personalisieren. In der Produktentwicklung analysieren intelligente Agenten Feedback, vergleichen Wettbewerber und unterstützen schnellere Roadmap-Entscheidungen. Überall dort, wo Daten und Wiederholungen zusammentreffen, kommt KI zum Einsatz – nicht, um den Menschen zu ersetzen, sondern um seine Leistungsfähigkeit zu erweitern.

Drei Säulen der agentenbasierten Enterprise-KI

Große Sprachmodelle (LLMs) sind Algorithmen, die durch ihr außergewöhnliches Verständnis natürlicher Sprache, ihre Fähigkeit zum gezielten Informationsabruf und zur präzisen, dialogischen Antwortgenerierung überzeugen. Doch für die Umsetzung realer Aufgaben – etwa das Konfigurieren von Marketingkampagnen, das Erstellen von User-Journeys oder das Testen von Preismodellen – genügt Sprachverständnis allein nicht. Es braucht Kontext: ein tiefes Verständnis darüber, wie Unternehmenssysteme tatsächlich funktionieren. Hier setzen Retrieval-Augmented Generation (RAG)- und Model Context Protocol (MCP)-Architekturen an und verschieben die Grenzen der Leistungsfähigkeit erneut.

Retrieval-Augmented Generation (RAG) ist ein Verfahren, das die Leistungsfähigkeit großer Sprachmodelle (LLMs) erhöht, indem in jede Antwort relevantes, domänenspezifisches Wissen einfließt. Proprietäre Dokumentationen, Code-Repositorien und Prozessanweisungen, die über RAG sicher integriert sind, ermöglichen einem LLM den Zugriff auf das „Wie“ der Aufgabe. Anstatt sich ausschließlich auf öffentliche Datenquellen zu stützen, greift es auf die verwaltete Wissensdatenbank der Organisation zu, um präzise und regelkonforme Anleitungen zu erzeugen.

Der Model Context Protocol (MCP)-Server schließt den Kreislauf. Ein MCP-Server dient als Kommunikationsschicht und verbindet generative KI mit Unternehmenssystemen, Datenbanken und APIs. So kann die KI über die Konversation hinaus agieren – Daten abrufen, Transaktionen ausführen oder Arbeitsabläufe in Echtzeit auslösen. Moderne Softwareumgebungen können Hunderte von MCP-Endpunkten enthalten, die es der KI jeweils ermöglichen, spezifische Operationen unter Richtlinienkontrolle auszuführen.

Zusammen bilden diese drei Säulen – LLMs, RAG und MCP – die Grundlage agentenbasierter KI im Unternehmen. Sie verwandeln Sprache in Logik, Absicht in Ausführung und Erkenntnis in messbare Ergebnisse. Dies markiert die nächste Evolutionsstufe intelligenter Systeme – gesteuert, kontextbezogen und fähig zur Zusammenarbeit mit Menschen, um in allen Funktionsbereichen Veränderungen voranzutreiben.

Ein globaler Bergbaukonzern hat genau das umgesetzt – sein Forschungsprojekt beschleunigt Abläufe mit KI, wie die folgende Fallstudie zeigt.

Fallstudie

Ein globales Bergbauunternehmen

Ein global tätiges Bergbauunternehmen mit Hauptsitz in Brasilien produziert Eisen, Nickel, Kupfer, Mangan und weitere Rohstoffe. Das Unternehmen führt umfassende Untersuchungen durch, um die wirtschaftlichen, sozialen und ökologischen Auswirkungen seiner Aktivitäten zu bewerten. Seine Forschungsprojekte können sich über bis zu zehn Jahre erstrecken. Typischerweise verbringen funktionsübergreifende Teams Wochen oder sogar Monate mit dem manuellen Sammeln und Zusammenführen von Informationen, um neue Produktmöglichkeiten, Marktschwankungen oder Umweltstandards zu überprüfen.
Diese aufwändige Vorgehensweise sowie ein unzureichend skalierbarer KI-Assistent behinderten die Forschungsprojekte erheblich.

Das Unternehmen, das seit jeher auf technologische Innovationen setzt, suchte nach einer Lösung zur Reduzierung wiederkehrender manueller Aufgaben und zur Beschleunigung der Forschungsphasen, sowohl für neue als auch für bestehende Minen. Ein Sprecher des Unternehmens erklärte: „Informationen zu jedem Bergbauprojekt werden typischerweise in unterschiedlichen Formaten und isolierten Systemen gespeichert. Der Zugriff darauf beansprucht viel wertvolle Zeit, da die Mitarbeiter Berge von Daten durchsuchen müssen.
Künstliche Intelligenz bot mit ihrer Fähigkeit, große Datenmengen schnell zu erfassen und zu analysieren, die ideale Möglichkeit, diese manuelle Arbeit deutlich zu reduzieren."

Das Bergbauunternehmen entwickelte gemeinsam mit einem KI-Anbieter einen Machbarkeitsnachweis und konnte durch Expertenschulungen und den Einsatz bewährter KI-Methoden die Genauigkeit der KI-Antworten um 47 Prozent erhöhen. Durch die Integration von KI in die Rechercheprozesse für Bergbauprojekte gelang es, mehrmonatige Routinearbeiten erheblich zu verkürzen und Wachstumspotenziale zu beschleunigen. Wenn beispielsweise die Machbarkeit der Erzgewinnung aus einer bestehenden Mine bewertet werden muss, eine Aufgabe, für die ein Geologe rund zwei Monate benötigen würde, können die relevanten Informationen heute innerhalb weniger Stunden zusammengestellt werden. Dies ermöglicht es dem Unternehmen, tragfähige Investitionsmöglichkeiten schneller zu erkennen und Marktentwicklungen proaktiv zu begegnen.

Ein Weg nach vorn für die KI

Wenn man auf die vergangenen 75 Jahre der künstlichen Intelligenz zurückblickt, zeigt sich eine zentrale Erkenntnis: Der Erfolg von KI beruhte nie allein auf Technologie. Während Fortschritte in Daten, Rechenleistung und Algorithmen bemerkenswerte Entwicklungen ermöglichten, entstand der nachhaltigste Einfluss stets durch ihre Verknüpfung mit solider Unternehmensführung, ethischen Prinzipien und menschlicher Zusammenarbeit.

Enterprise-KI steht damit für weit mehr als die nächste Phase künstlicher Intelligenz – sie markiert eine grundlegende Neugestaltung der Interaktion zwischen Mensch und Technologie. Im Unterschied zur traditionellen generativen KI, die auf Eingaben reagiert, zeigen agentenbasierte Systeme Autonomie, Eigeninitiative und adaptives Denken. Sie können planen, handeln und lernen, agieren kontextbezogen und überführen Konversationen direkt in Umsetzung – ohne ständige menschliche Steuerung.

Dieser Wandel kündigt den stillen Abschied der grafischen Benutzeroberfläche (GUI) als dominierendes Interaktionsmodell an. Schaltflächen, Registerkarten und Menüs treten in den Hintergrund, während intelligente Agenten über Sprache und Kontext interagieren. An die Stelle der Navigation durch Software tritt die Ausdrucksabsicht des Nutzers – das System interpretiert, entscheidet und handelt.

Die Schnittstelle verschwindet nicht, sie entwickelt sich weiter: Sichtbare Oberflächen weichen intelligenter Interaktion, die auf Konversation, Kontext und Vertrauen beruht. In diesem neuen Paradigma wird Produktivität nicht mehr in Klicks pro Minute, sondern in der Qualität der erzielten Ergebnisse gemessen – durch die Partnerschaft von Mensch und KI.

Agentenbasierte Intelligenz verändert bereits, wie Organisationen Inhalte erstellen, verwalten und personalisieren. Recherchezyklen werden automatisiert, Entwürfe generiert und Erlebnisse in Echtzeit kuratiert, abgestimmt auf Verhalten und Präferenzen des Publikums. KI reagiert nicht nur, sie denkt voraus und bewertet die langfristigen Folgen jeder Entscheidung. Für Unternehmen liegt darin eine klare Chance: Kontextuelle Intelligenz ersetzt nicht das menschliche Urteilsvermögen – sie verstärkt und skaliert Fachwissen, Governance und Kreativität im gesamten Unternehmen.

Die kommende Ära der KI wird geprägt sein von größerer Autonomie, Transparenz und Verantwortlichkeit. Systeme werden intelligent handeln, zugleich erklärbar bleiben und sich an menschlichen Werten orientieren. Das Verständnis, was KI ist, woher sie kommt und wie sie funktioniert, bildet die Grundlage für eine verantwortungsvolle Gestaltung der nächsten Entwicklungsphase.

Das nächste Kapitel widmet sich daher der Schnittstelle von Daten und KI – und untersucht, wie die Verschmelzung von Information und Intelligenz neue Möglichkeiten für Innovation, Vertrauen und Wertschöpfung im modernen Unternehmen eröffnet.

Die fünf Merksätze

1. **KI ist vielfältig, dynamisch und grundlegend.**
   Künstliche Intelligenz im Unternehmen ist kein einzelnes System, sondern ein Oberbegriff für Technologien wie maschinelles Lernen, Deep Learning und Sprachverarbeitung (NLP). Sie reicht von engen, aufgabenspezifischen Systemen (ANI) bis hin zum konzeptionellen Ziel superintelligenter Systeme (ASI). Das Verständnis dieser Unterschiede ist entscheidend für informierte Entscheidungen.

2. **Datenqualität und Governance sind ausschlaggebend.**
   Die Wirksamkeit und Vertrauenswürdigkeit von KI hängt von qualitativ hochwertigen, kontrollierten Daten ab. Daten sind der „Treibstoff" für KI-Innovationen; ohne zuverlässige, sichere und gut verwaltete Daten sind KI-Systeme anfällig für Fehler, Verzerrungen und operationelle Risiken.

3. **Technologischer Fortschritt folgt Hype-Zyklen.**
   Die Geschichte der KI zeigt einen Rhythmus aus Innovationsschüben und Ernüchterungsphasen (KI-Wintern). Diese Zyklen haben die Branche reifen lassen und deutlich gemacht, dass ein nachhaltiger Wert aus dem Gleichgewicht zwischen technologischen Fortschritten, realistischen Erwartungen und umsichtigen Investitionen entsteht.

4. **Rechenleistung und Algorithmen treiben moderne KI.**
   Bahnbrechende Fortschritte bei der Hardware (GPUs, TPUs, Cloud-Infrastruktur) und dem Algorithmen-Design haben die heutigen groß angelegten, generativen KI-Systeme ermöglicht. Die Synergie zwischen Daten, Rechenleistung und Algorithmen ist das, was die führenden Unternehmen im Bereich der KI auszeichnet.

5. **Governance, Ethik und menschliche Ausrichtung sind entscheidend.**
   TDie Zukunft der KI verlangt Transparenz, Nachvollziehbarkeit und ethische Orientierung. Nur mit starker Governance und menschlicher Aufsicht kann KI Wert schaffen und Vertrauen sichern. Erfolg erfordert starke Governance, ethische Rahmenbedingungen und menschliche Aufsicht, um sicherzustellen, dass KI den Geschäftswert steigert und gleichzeitig Vertrauen aufbaut.

Kapitel Drei

Die Schnittstelle von Daten und künstlicher Intelligenz

In diesem Kapitel wird die Schnittstelle zwischen Daten und künstlicher Intelligenz untersucht, wobei der Schwerpunkt darauf liegt, wie Daten zu Intelligenz werden.

Aufbauend auf den Grundlagen aus Kapitel 1 (Daten) und Kapitel 2 (KI) wird gezeigt, wie Daten und Intelligenz eine kontinuierliche Wertschöpfungskette bilden. Daten treiben die KI-Engine an; KI wiederum erschließt den verborgenen Wert der Daten. Kontinuierliches Lernen schließt den Kreislauf und fördert im Laufe der Zeit Genauigkeit, Anpassungsfähigkeit und Erkenntnisgewinn. Governance verbindet diese Welten und stellt sicher, dass Unternehmensintelligenz mit ihrem Wachstum erklärbar, überprüfbar und im Einklang mit dem Vertrauen in die Organisation bleibt. Abschließend werden die strategischen und wirtschaftlichen Auswirkungen der Zusammenführung von Daten und KI betrachtet.

Wo Daten und Intelligenz aufeinandertreffen

Kapitel 1 zeichnete die Entwicklung der Daten nach, wie sie zur Grundlage des Enterprise Information Management wurden und wie ihre Struktur, Verwaltung und Zugänglichkeit geschäftlichen Mehrwert erzeugen. Kapitel 2 vertiefte die Rolle der künstlichen Intelligenz als Motor der Automatisierung und Intelligenz und verfolgte ihre technologische Entwicklung über den Hype-Zyklus hinaus bis zum realen, agentenbasierten Einsatz.

Die zentrale These dieses Buchs lautet: Daten und KI sind symbiotisch. Daten geben der KI den Kontext und das Lernpotenzial, während KI Daten in umsetzbare Erkenntnisse verwandelt. Gemeinsam treiben sie Innovationen im modernen Unternehmen voran.

Hochleistungsfähige KI benötigt qualitativ hochwertige Daten. Diese Daten sind gut verwaltet, strukturiert und kontextreich. Nicht alle Daten sind gleichwertig. Unterschiedliche Datensätze stellen unterschiedliche Anforderungen, insbesondere wenn öffentliche und private Daten kombiniert werden. Öffentliche Datensätze trainieren große Sprachmodelle (LLMs), die Tools wie ChatGPT zugrunde liegen. Für Unternehmen sind jedoch private, kundenspezifische Daten der entscheidende Wettbewerbsvorteil. Strategien, die Vertraulichkeit und Souveränität dieser Daten wahren und zugleich der KI erlauben, daraus zu lernen, sind essenziell.

An diesem Schnittpunkt – wo Unternehmensdaten auf intelligente Systeme treffen – entfaltet sich das wahre Potenzial von KI im Unternehmen. Hier wird aus Kontext Fähigkeit und aus Information Erkenntnis. Organisationen, die diese Beziehung verantwortungsvoll nutzen, werden die nächste Ära digitaler Leistungsfähigkeit prägen.

Kann KI Daten- und Unternehmensinformationsmanagement ersetzen?

Mit zunehmender Verbreitung von KI stellt sich oft die Frage, ob sie bestehende Daten- und Informationsmanagementlösungen ersetzen kann. Die kurze Antwort lautet: nein. Entscheidend ist das Zusammenspiel von Informationsmanagement und KI – das eine kann nicht ohne das andere wirken. KI automatisiert datenspezifische Aktionen wie Extraktion und Klassifizierung, während Informationsmanagement sichere, organisierte Inhalte sowie Struktur, Governance- und Compliance-Regeln bereitstellt, die KI allein nicht liefert.

Während Daten- und Informationsmanagement den Treibstoff für KI liefern, verändert KI zugleich das Informationsmanagement. Dies geschieht durch:

- **Automatisierung** - Dokument-Tagging, Extraktionen, Zusammenfassungen und Verringerung menschlicher Fehler in zentralen Abläufen.
- **Einblicke** - KI liefert wertvolle Einblicke in die zu verwaltenden Inhalte, einschließlich der Ableitung wichtiger Erkenntnisse, Stimmungen und anderer wichtiger Hinweise.
- **Suche und Abruf** - KI in Kombination mit den Metadaten im Content-Management macht Suchoberflächen genauer, effizienter und benutzerfreundlicher.

Informationsmanagement ist der Wächter vertrauenswürdiger Daten; Datenqualität bestimmt die Glaubwürdigkeit jeder KI-Entscheidung. Beide Disziplinen sind eng verflochten: Effektive KI beruht auf kontrollierten, hoch integren Daten, während Informationsmanagement durch KI-gestützte Automatisierung an Tempo und Intelligenz gewinnt. Die Integration beider Ansätze sichert Konsistenz, Konformität und Kontext über den gesamten Informationslebenszyklus hinweg (siehe Kapitel 5). Künstliche Intelligenz kann die Art und Weise verbessern, wie Organisationen Inhalte verwalten – sie kann jedoch die Disziplin und Governance, die Informationen vertrauenswürdig machen, nicht ersetzen.

Diese Wechselbeziehung zwischen KI und Informationsmanagement wird in der folgenden Fallstudie eines globalen Lebensmittelproduzenten verdeutlicht, der KI auf seine Geschäftsdaten anwendet, um Abläufe zu modernisieren und Leistung zu steigern.

Fallstudie

Ein globaler Lebensmittelhersteller

Feldbeobachtung mit Drohnen

Ein weltweit tätiger Lebensmittelhersteller hat verschiedene Strategien umgesetzt, um sich als Branchenführer bei der Anwendung von KI zu positionieren. Im Folgenden finden sich Auszüge aus einem Interview mit dem EIM-Direktor des Unternehmens:

„Im Rahmen unseres Transformationsprojekts untersuchen wir, wie künstliche Intelligenz uns bei der Modernisierung unserer Betriebsabläufe unterstützen kann. Derzeit werden etwa zehn Prozent unserer Daten in der Cloud gespeichert. Das ist noch keine große Zahl. Die bisher eingesetzten Cloud-Lösungen waren alle privat, um die Sicherheit unserer firmeneigenen Daten zu gewährleisten. Da sich die Technologie jedoch rasant weiterentwickelt, werden wir offener für den Einsatz öffentlicher Cloud-Dienste, sofern Governance und Datenhoheit gewährleistet bleiben.

Künstliche Intelligenz wird zunehmend zu einem zentralen Bestandteil der Art und Weise, wie wir Daten verwalten und daraus Wert schöpfen. Wir nutzen KI-gestützte Systeme, um Erkenntnisse aus unseren Betriebsdaten zu gewinnen und so das Management bei einer effizienteren Steuerung unserer Werke zu unterstützen. Die Ergebnisse sind durch qualitativ hochwertigere Produkte, nachhaltigere Praktiken und eine deutliche Verbesserung der Geschäftsergebnisse messbar.

KI hat auch völlig neue Arbeitsweisen inspiriert. Wir erproben die drohnengestützte Überwachung von Nutzpflanzen – die Weiterentwicklung einer Praxis, die wir seit Jahren mithilfe von Satellitenbildern anwenden. Satelliten haben uns bei der Beurteilung des Pflanzenzustands unterstützt, stoßen jedoch bei Bewölkung an ihre Grenzen. Drohnen dagegen können so programmiert werden, dass sie über ganze Felder fliegen, hochauflösende Bilder aufnehmen und diese Daten direkt in unsere KI-Modelle einspeisen. Nach der Verarbeitung prognostizieren die Modelle den Ertrag, erkennen Stressfaktoren oder Krankheiten und geben Empfehlungen für gezielte Anpassungen bei Bewässerung oder Düngung. Diese Erkenntnisse werden in automatisierte Düngerstreuer integriert, die die optimale Menge an Behandlungsmittel an den richtigen Stellen ausbringen, wodurch Verschwendung reduziert und der Ertrag gesteigert wird.

Darüber hinaus machen wir Fortschritte im Bereich der vorausschauenden Landwirtschaft. Durch die Kombination von KI-Modellen mit jahrzehntelangen historischen Wetter- und Erntedaten können Wachstumsbedingungen in bestimmten Regionen zwei bis drei Jahre im Voraus vorhergesagt werden. Diese Modelle sind zwar nicht perfekt, aber sie werden immer genauer und sind für die Planung unglaublich nützlich.

Jede Region der Welt ist anders, hat eigene Böden, Wetterbedingungen, Anbauprodukte und landwirtschaftliche Praktiken. Die Herausforderung besteht darin, sich an all diese Unterschiede anzupassen, und KI hilft dabei, dies in großem Maßstab zu bewerkstelligen. Die Technologie ermöglicht ein tiefes Verständnis der lokalen Gegebenheiten und unterstützt Entscheidungen in Echtzeit, die Produktivität, Nachhaltigkeit und Widerstandsfähigkeit verbessern. Was früher wochenlange manuelle Analysen erforderte, geschieht heute kontinuierlich. Künstliche Intelligenz

ist nicht nur ein Werkzeug zur Gewinnung von Erkenntnissen geworden, sondern ein Partner bei der Art und Weise, wie angebaut, produziert und die Welt ernährt wird."

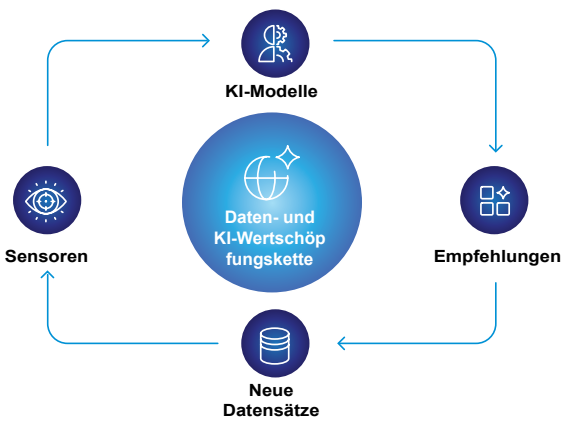
Die Wertschöpfungskette für Daten und KI

Die Wertschöpfung aus Daten beginnt mit dem Verständnis der Daten- und KI-Wertschöpfungskette. Dieser Prozess startet mit der Datengenerierung und -erfassung. Dabei ist der Zugriff auf Unternehmensdaten eine grundlegende Voraussetzung.
Die erfassten Daten müssen integriert, bereinigt und verwaltet werden, um Qualität und Zuverlässigkeit sicherzustellen. Organisationen, die in solide Informationsmanagementpraktiken investiert haben, sind besser aufgestellt, um die Einführung von KI zu beschleunigen, da sie die Grundlagenarbeit für die Nutzung ihrer Daten bereits geleistet haben.

Daten allein genügen jedoch nicht. Ohne strukturierte Prozesse und Arbeitsabläufe bleiben sie ungenutzt. Künstliche Intelligenz entfaltet ihren Wert erst, wenn sie auf reale geschäftliche Herausforderungen angewendet und in diese Arbeitsabläufe integriert wird, um messbare Ergebnisse zu erzielen. An dieser Stelle kommen das Training, die Feinabstimmung und die Validierung der KI-Modelle ins Spiel. LLMs werden zunächst mit öffentlichen Datensätzen trainiert. Doch Unternehmen können ihren Wert steigern, indem diese Daten mit privaten Informationen feinabgestimmt oder Retrieval-Augmented Generation (RAG)-Pipelines verwendet werden, die KI mit internen Wissensquellen verbinden. Die richtige Strategie hängt von den Zielen, Ressourcen und dem Reifegrad der Organisation ab. Unabhängig vom gewählten Ansatz sind jedoch Datenqualität und Modell-Governance entscheidende Voraussetzungen.

Der letzte – und oft übersehene – Schritt in der Wertschöpfungskette ist die Rückkopplungsschleife. Viele Organisationen führen KI-Funktionen ein, ohne Mechanismen für kontinuierliches Lernen und Optimieren einzurichten. Gerade hier zeigt sich der wahre Mehrwert: Durch iteratives Feintuning lässt sich die Modellgenauigkeit im Laufe der Zeit steigern und zu wirksameren Ergebnissen führen.

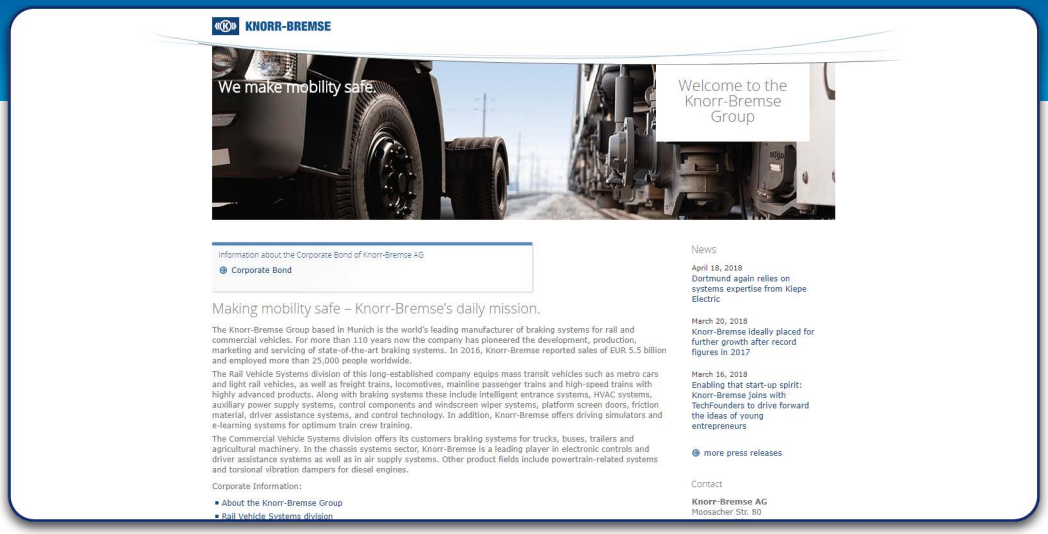
Um die Wertschöpfungskette von Daten und KI in einen praktischen Kontext zu setzen, betrachten wir ein Beispiel aus der Fertigungsindustrie: Sensoren in der Fabrikhalle sammeln Daten und speisen sie in KI-Modelle ein. Diese Modelle geben Empfehlungen zur Leistungsoptimierung. Dies wiederum erzeugt neue Datensätze, die kontinuierlich und iterativ verbessert werden können.



Dieser Ansatz wird in der folgenden Fallstudie veranschaulicht, die beschreibt, wie Knorr-Bremse mit vorausschauender Wartung, die auf umsetzbaren Erkanntnissen basiert, den reibungslosen Ablauf der Prozesse gewährleistet.

Fallstudie

Knorr-Bremse

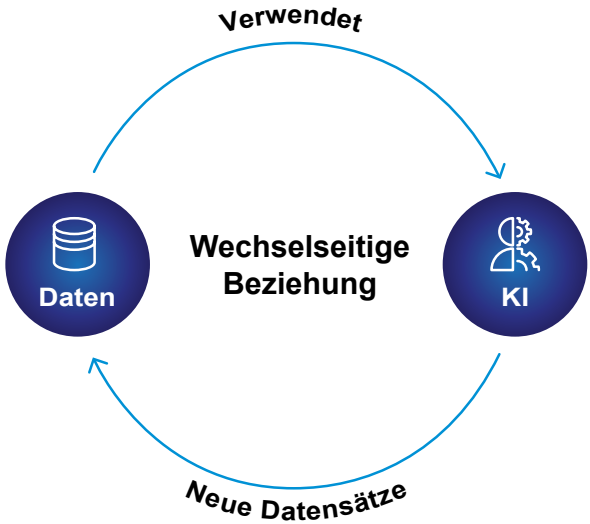


Knorr-Bremse Group

Die Knorr-Bremse Group mit Sitz in München ist der weltweit führende Hersteller von Bremssystemen für Schienen- und Nutzfahrzeuge. Seit über 110 Jahren ist das Unternehmen Vorreiter bei Entwicklung, Produktion, Vermarktung und Wartung modernster Bremssysteme.

Mit der iCOM-Plattform (intelligent Condition Oriented Maintenance) digitalisiert Knorr-Bremse den Bahnsektor, indem drahtlos ausgestattete Sensoren an Bord von Zügen mit einem cloudbasierten Backoffice-Netzwerk auf Basis eines IoT-Modells (Internet of Things) verbunden werden. Diese Plattform übermittelt detaillierte Zustandsdaten, die Prognosen zu Wartungs- und Ersatzbedarf ermöglichen. Die iCOM-Plattform benötigte eine leistungsstarke, zugleich benutzerfreundliche Analysekomponente, um die eingehenden Daten effizient auszuwerten und Nutzern fundierte, datengestützte Entscheidungen zu ermöglichen.

Die Fähigkeit, vorausschauend zu handeln, führt zu effizienteren und kostengünstigeren Instandhaltungsprozessen. Da die Sensoren kontinuierlich Daten erfassen, entstehen innerhalb einer Flotte enorme Datenmengen. Kunden können diese Daten nun unabhängig von der IT-Abteilung mithilfe interaktiver grafischer Dashboards visualisieren. Beispielsweise lassen sich Heatmaps zustandsbezogener Ereignisse anzeigen, etwa überhitzte Bremsen an bestimmten Steigungen. So können Kunden gezielte Maßnahmen ergreifen, um Komponentenausfälle zu vermeiden, die Lebensdauer ihrer Systeme zu verlängern und Wartungskosten deutlich zu senken.



KI und Daten bilden eine wechselseitige Beziehung

Daten als Treibstoff für KI

Daten sind der Treibstoff, der die KI-Engine antreibt. Quantität, Qualität und Vielfalt der Daten sind entscheidender als die Komplexität der Modelle selbst. Hochwertige, vielfältige Datensätze liefern KI-Systemen den Kontext, den sie für wirksames Lernen benötigen. Einfache Modelle können mit hochwertigen, vielfältigen Daten beeindruckende Ergebnisse erzielen, während komplexe Modelle auf minderwertigen, homogenen Datensätzen an Leistung verlieren.

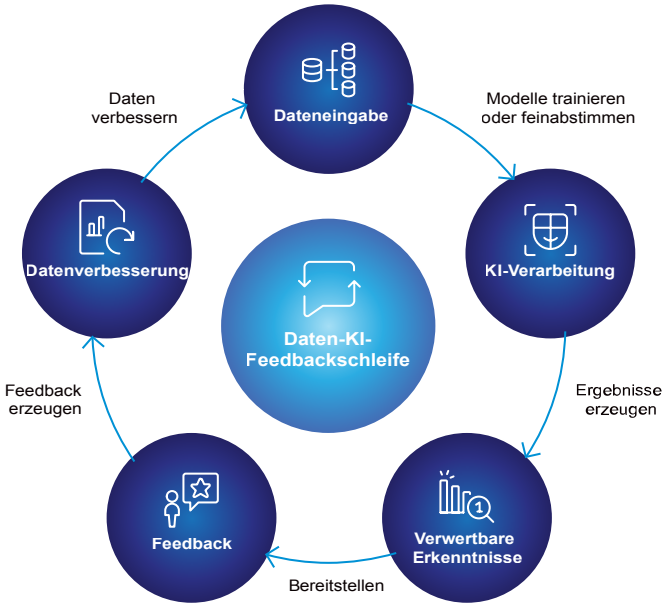
Wie in Kapitel 1 beschrieben, speisen strukturierte und unstrukturierte Daten diesen Mix. Die sichere und verantwortungsvolle Nutzung dieser Daten ist der Schlüssel zu einer sinnvollen KI-Einführung im Unternehmenskontext. Für die meisten Organisationen liegt der Erfolg nicht in der Entwicklung großer öffentlicher Modelle, sondern in der strategischen Nutzung privater Daten innerhalb bestehender Rahmenbedingungen. Sobald diese privaten Daten für KI bereitgestellt werden, wird ihr Schutz zu einem entscheidenden Faktor. Dies ist ein zentraler Aspekt des Souveränitätsprinzips, das später im Buch vertieft wird. Die klare Trennung zwischen öffentlichen und privaten Datensätzen sowie der Schutz dieser privaten Daten hat höchste Priorität.

Künstliche Intelligenz konsumiert Daten nicht einfach, sondern interpretiert, verknüpft und organisiert sie für die unternehmensweite Nutzung. Damit entsteht eine wechselseitige Beziehung: KI lernt aus Daten und steigert gleichzeitig deren Wert. Indem sie Struktur, Integrität und Zugänglichkeit verbessert, organisiert sie diese für die unternehmensweite Nutzung.

Die kontinuierliche Rückkopplungsschleife

Jedes wirksame KI-System im Unternehmen beruht auf einem kontinuierlichen Feedback-Kreislauf: Daten trainieren Modelle, die Modelle erzeugen Erkenntnisse, und diese Erkenntnisse generieren neue Daten, die sowohl das Modell als auch die zugrunde liegenden Datensätze verfeinern. Intelligenz entwickelt sich nicht linear, sondern in wiederkehrenden Lernzyklen.

Empfehlungssysteme beispielsweise lernen fortlaufend aus dem Nutzerverhalten. Jede Interaktion erzeugt neue Datenpunkte, die dem System helfen, genauere Vorhersagen zu treffen. Durch diese iterative Optimierung verbessern sich im Laufe der Zeit Genauigkeit, Personalisierung und Effizienz. Auch auf einer Online-Shopping-Plattform zeigt sich dieser Prozess deutlich: Jeder Klick, jeder Kauf und jede Pause liefert neue Signale, die das Verständnis des Systems für die jeweilige Absicht verändern. Die nachfolgenden Empfehlungen spiegeln wider, was das Modell seit dem letzten Zyklus gelernt hat. Während weitere Interaktionen stattfinden, nutzt das System die neuen Daten, um kontinuierlich dazuzulernen und die Relevanz seiner Vorschläge zu erhöhen.



KI und Daten bilden eine wechselseitige Beziehung

Ebenso wichtig für diesen Prozess sind Beobachtbarkeit und Überwachung. Der Kreislauf muss gesteuert und die Modelle verantwortungsvoll weiterentwickelt werden. Die kontinuierliche Überwachung der Modellleistung und des Datenflusses stellt sicher, dass KI-Systeme zuverlässig, nachvollziehbar und auf die Geschäftsziele abgestimmt bleiben. Wie später im Buch noch erläutert wird, darf das operative Management von KI-Systemen nicht erst im Nachhinein bedacht werden. Es muss als strategische Fähigkeit verstanden werden, die die Grundlage für langfristigen Erfolg bildet.

Governance an der Schnittstelle

Governance steht im Mittelpunkt der Verbindung zwischen Daten und KI im Unternehmen. Im Datenbereich liegt der Schwerpunkt auf Datenschutz, Datenherkunft, Zugriffskontrolle und der Einhaltung von Vorschriften wie der DSGVO. Im Bereich der KI richtet sich die Governance auf Fairness, Transparenz, Verantwortlichkeit und Erklärbarkeit.

Diese beiden Disziplinen vereinen sich zunehmend unter gemeinsamen Prinzipien wie Ethik, Überprüfbarkeit und Vertrauen. Neue Vertrauensrahmen für KI sowie internationale Standards wie ISO/IEC 42001 für KI-Management und ISO/IEC 38505 für Daten-Governance verdeutlichen diese Konvergenz. Mit zunehmender Reife dieser Rahmenwerke werden sie die Art und Weise prägen, wie Organisationen KI verantwortungsvoll entwickeln, einsetzen und überwachen. In den folgenden Kapiteln 5 bzw. 6 gehen wir näher auf die Governance von Daten und KI ein.

Die Integration von Daten und KI schafft Wettbewerbsvorteile und ihre verantwortungsvolle Nutzung verwandelt sie in nachhaltigen wirtschaftlichen Wert.

Strategische und wirtschaftliche Auswirkungen

Die Integration von Daten und KI schafft sowohl strategische Vorteile als auch wirtschaftliche Chancen. Organisationen, die diese Fähigkeiten wirksam aufeinander abstimmen, sind besser positioniert, um Innovation zu fördern, Abläufe zu optimieren und sich in wettbewerbsintensiven Märkten zu differenzieren.

Angesichts der hohen Erwartungen von Führungsebenen an das Potenzial von KI ist nachvollziehbar, dass das Tempo des Wandels und seine tatsächlichen Auswirkungen mitunter als zu gering empfunden werden. Diese Wahrnehmung hat den Blick auf frühe KI-Pilotprojekte gelenkt und deren Erfolg kritisch beleuchtet. Viele dieser Initiativen blieben hinter den Erwartungen zurück, weil sie sich auf öffentlich trainierte Modelle stützten, ohne diese mit unternehmensspezifischen Daten zu kontextualisieren. Der entscheidende Wettbewerbsvorteil liegt darin, eigene Daten sicher zu nutzen und in Erkenntnisse zu überführen. Die nächste Erfolgsphase beruht auf einem datenzentrierten Ansatz, der die Verbesserung der Datenqualität und der Prozessgestaltung über die Entwicklung immer komplexerer Modelle stellt.

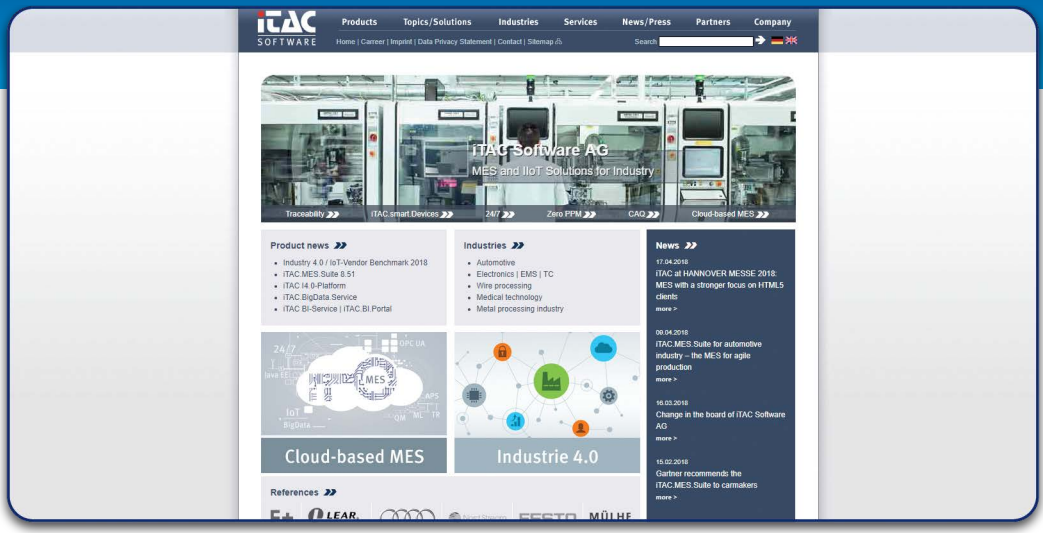
Zuverlässige Daten und klar definierte Prozesse führen zu belastbaren KI-Ergebnissen.

Während umfangreiche Rechenkapazitäten für das Training grundlegender Modelle weiterhin eine Rolle spielen, können die meisten Unternehmen durch gezielte, kleinere Implementierungen bereits erheblichen Mehrwert erzielen. Ein präzises Verständnis der eigenen Datenanforderungen ermöglicht es, den tatsächlichen Investitionsbedarf in Rechenleistung realistisch zu bestimmen, unnötige Ausgaben zu vermeiden und KI-Initiativen konsequent an messbarem Geschäftswert auszurichten. Dies trägt auch dazu bei, Bedenken gegenüber KI bei Führungskräften und Beschäftigten abzubauen, die sich noch im Prozess befinden, die Technologie vollständig zu verstehen.

In der folgenden Fallstudie wird ersichtlich, wie die iTAC Software AG mithilfe von Intelligenz intelligente Fabriken ermöglicht.

Fallstudie

iTAC Software AG



iTAC Software

Seit ihrer Gründung hat sich die iTAC (Internet Technologies and Consulting) Software AG auf die Bereitstellung von Internettechnologien für die Fertigungsindustrie spezialisiert. Das Unternehmen entwickelt Standardsoftware und Produkte für unternehmensübergreifende IT-Anwendungen und gilt als führender System- und Lösungsanbieter von Manufacturing Execution Systems (MES) für die gesamte Lieferkette.

Um seinen Kunden maximale Transparenz und Entscheidungsfähigkeit bei der Produktionssteuerung zu bieten und den wachsenden Anforderungen im Zusammenhang mit dem Internet der Dinge (IoT) gerecht zu werden, wollte iTAC Business-Intelligence- und Analysesoftware in seine MES-Suite integrieren. Dies sollte den Anforderungen an Fertigungsintelligenz, Qualitätskontrolle und Rückverfolgbarkeit entsprechen. Neben einer schnellen und effektiven Implementierung sowie einer nahtlosen Integration erforderte iTAC individuell anpassbare Berichte, Analysen und Dashboards mit voller Interaktivität und Sicherheit. Alle Lösungen mussten webbasiert sein, eine transparente Personalisierung für verschiedene Anwendungen ermöglichen und über unterschiedliche Kanäle zugänglich bleiben.

iTAC verfügt nun über die Business-Intelligence-, Betriebs- und Analysefunktionen, die erforderlich sind, um die steigende Kundennachfrage nach Intelligenz, Qualitätskontrolle und Rückverfolgbarkeit im gesamten Fertigungsprozess zu unterstützen. Die Lösung gewährleistet Transparenz im Kennzahlenmanagement und unterstützt Produktlebenszyklusmanagement, Budgetkontrolle, Qualitätssicherung und Außendienstmanagement. Die Kunden des Unternehmens können zentral auf große Datenmengen zugreifen und diese analysieren. Dank der skalierbaren Architektur und der Erweiterbarkeit für zukünftige Anforderungen verschafft sich das Unternehmen einen nachhaltigen Wettbewerbsvorteil.

Wie in diesem Kapitel beschrieben, sind Daten und KI untrennbar miteinander verbunden. Daten sind der Treibstoff des KI-Motors. KI ohne Daten bleibt richtungslos, und Daten ohne KI bleiben ungenutzt. Erst im Zusammenspiel entstehen intelligente, handlungsrelevante Unternehmensentscheidungen.

Mit der zunehmenden Verlagerung auf KI-gestützte Entscheidungsrahmen werden eine starke Governance und eine klare strategische Ausrichtung unverzichtbar. Die Schnittstelle von Daten und KI steht nicht nur für einen operativen Wandel, sondern markiert eine neue Innovationsfront, die die Art und Weise, wie Organisationen denken, entscheiden und konkurrieren, grundlegend verändert. Diese Konvergenz leitet ein neues Kapitel der digitalen Transformation ein – ein Kapitel, in dem Information tatsächlich zu Intelligenz wird.

Die fünf Merksätze

1. **Datenqualität und Governance haben Priorität.**
   Datenverfügbarkeit sollte als organisatorische Verpflichtung und nicht als Projektergebnis verstanden werden. Teams müssen umfassende Datenprüfungen durchführen und Governance-Richtlinien umsetzen, die Genauigkeit, Sicherheit und Zugänglichkeit aller kritischen Informationsbestände gewährleisten. Um die Wirksamkeit von KI zu maximieren, hat Datenqualität auf Vorstandsebene höchste Priorität.

2. **KI in reale Geschäftsprozesse integrieren.**
   Es sollten zwei bis drei geschäftskritische Bereiche identifiziert werden – etwa Kundensupport, Lieferkettenoptimierung oder Risikomanagement –, in denen unmittelbare Vorteile erzielt werden können. Funktionsübergreifende Teams sind damit zu beauftragen, KI-Lösungen einzusetzen, die firmeneigene Daten nutzen, um konkrete Herausforderungen zu adressieren.

3. **Kontinuierliche Feedbackschleifen für die KI-Verbesserung einrichten.**
   KI-Leistung ist nie statisch; sie erfordert ständige Überwachung und Training. Eine Organisationsrichtlinie ist einzuführen, die die Leistungsüberwachung von KI-Modellen, die Einrichtung von Feedbackschleifen mit Nutzern sowie das automatisierte Nachtraining mit neuen Daten umfasst. Klare Verantwortlichkeiten müssen festgelegt werden, um sicherzustellen, dass Modelle präzise, kontextbezogen und auf die Geschäftsziele abgestimmt bleiben.

4. **Daten und KI-Governance hinsichtlich Vertrauen und Compliance aufeinander abstimmen.**
   Daten und KI sind unter einem einheitlichen Governance-Rahmen zusammenzuführen. Eine funktionsübergreifende Arbeitsgruppe sollte eingerichtet werden, die Datenschutz, Sicherheit, Compliance und ethische Aufsicht vereint und für die Informationsnutzung einheitliche Standards definiert. Um rechtliche, reputationsbezogene und operative Risiken proaktiv zu managen, sollten neue Standards (wie ISO/IEC 42001 und 38505) übernommen oder als Benchmark verwendet werden.

5. **Bei KI-Investitionen ist ein datenzentrierter Ansatz zu verfolgen.**
   Alle Investitionsentscheidungen sind an den Datenwert und die Geschäftsergebnisse zu koppeln. Vor der Freigabe neuer KI-Projekte muss dargelegt werden, wie die Initiative den Wert der Unternehmensdaten erschließt und messbare Geschäftsergebnisse erzielt. Investitionen in groß angelegte KI-Modelle sollten auf Fälle beschränkt werden, die sich auf einzigartige Datenbestände stützen und einen klaren Weg zum Return on Investment aufweisen.

Kapitel Vier

Sicherheit gewährleisten – Die Bedeutung von Cybersicherheit

Innovation muss durch Vertrauen ergänzt werden, um sicherzustellen, dass die von uns geschaffene Intelligenz nicht gegen uns verwendet werden kann. Dieses Kapitel untersucht, wie sich Cybersicherheit parallel zur Entwicklung künstlicher Intelligenz weiterentwickeln muss. Da intelligente Systeme die Arbeitsweise von Unternehmen grundlegend verändern, entstehen neue Risiken, die fortschrittliche Abwehrmechanismen erfordern. Im Fokus stehen dabei aufkommende Bedrohungen und die Strategien, mit denen Daten, Modelle und KI-gesteuerte Operationen geschützt werden können.

> *62 % der Organisationen waren in den letzten zwölf Monaten Opfer eines Deepfake-Angriffs, bei dem Social Engineering oder die Ausnutzung automatisierter Prozesse zum Einsatz kamen, während 32 % angaben, einen Angriff auf KI-Anwendungen erlebt zu haben, bei dem die Eingabeaufforderung manipuliert wurde.*[23]

In den vergangenen Jahren haben sich Cyberbedrohungen von einfachen Sicherheitslücken zu hochentwickelten Angriffen entwickelt, die gezielt auf KI-Systeme von Unternehmen abzielen – und es steht mehr auf dem Spiel als je zuvor. Mit der zunehmenden Einführung von Technologien wie generativer KI (GenAI) steigt die Zahl der Angriffe, die KI für Phishing, Deepfakes oder fortgeschrittenes Social Engineering nutzen. Gleichzeitig entstehen neue Schwachstellen: böswillige Akteure nutzen GenAI-Infrastrukturen aus, manipulieren Eingabeaufforderungen oder kompromittieren verkettete KI-Workflows, um in Unternehmenssysteme einzudringen oder sie zu stören.

Das vorangegangene Kapitel hat gezeigt, wie die Schnittstelle von Daten und KI Innovation und betriebliche Effizienz ermöglicht. Doch jede Chance birgt Risiken. Da Organisationen zunehmend auf private Daten angewiesen sind, um ihre KI-Systeme zu trainieren und zu betreiben, setzen sie sich gleichzeitig neuen, dynamischen Cyberrisiken aus. Der Schutz von Unternehmensdaten und KI muss daher mit der technologischen Entwicklung Schritt halten, denn Daten und Modelle sind attraktive Ziele für Bedrohungsakteure.

In der folgenden Fallstudie wird gezeigt, wie ein Energieunternehmen den Grundstein für KI und fortgeschrittene Analytik innerhalb eines sicheren EIM-Systems legt und eine Unternehmensarchitektur aufbaut, die Daten mit Governance- und Cybersicherheitsprozessen verbindet.

Fallstudie

Ein nordisches Energieunternehmen

Ein Energieerzeuger aus dem nordischen Raum arbeitet in einem stark regulierten Umfeld und verwaltet umfangreiche technische Dokumentationen, die für Sicherheit und Betrieb entscheidend sind. Angesichts der Herausforderung, über 900 Mitarbeitern einen zuverlässigen Zugriff auf die jeweils aktuellen genehmigten Versionen dieser Dokumente in Büro- und Produktionsumgebungen zu ermöglichen, erkannte das Unternehmen, dass veraltete, fragmentierte Systeme nicht die für ein modernes Risiko- und Vertrauensmanagement erforderliche Governance und Transparenz boten. In einer Zeit, in der Daten sowohl ein Wert als auch eine potenzielle Schwachstelle sind, wurde der Aufbau eines robusten Sicherheitskonzepts zur zentralen Voraussetzung.

Um eine belastbare Grundlage zu schaffen, implementierte das Unternehmen eine einheitliche Content-Management-Umgebung auf Basis strenger Identitäts- und Zugriffskontrollen, automatisierter Workflows und klarer Dokumentenlebenszyklus-Governance. Durch die Zentralisierung der Zugriffsrechte und die Durchsetzung richtlinienbasierter Kontrolle wurde gewährleistet, dass sensible Betriebsdaten nur im richtigen Kontext und zur richtigen Zeit von autorisierten Personen eingesehen werden konnten. Automatisierte Workflows führten Dokumente sicher durch die Phasen der Prüfung, Genehmigung und Archivierung und stärkten die Datensicherheit, ohne die Benutzerfreundlichkeit für Mitarbeiter im Büro und den Außendienst einzuschränken. Mit dieser Architektur wurde im Bewusstsein, dass diese nur auf sicheren und gut verwalteten Informationen aufbauen können die Grundlage für fortgeschrittene Analyse- und KI-gestützte Funktionen geschaffen.

Die Ergebnisse waren bahnbrechend. Das Unternehmen erreichte eine hohe Systemstabilität und steigerte gleichzeitig die Produktivität der Beschäftigten erheblich. Der nahezu in Echtzeit verfügbare Zugriff auf geschäftskritische Inhalte stärkte sowohl die Sicherheit als auch die operative Integrität. Noch entscheidender ist die nun vorhandene vertrauenswürdige Informationsinfrastruktur, die den sicheren und verantwortungsvollen Einsatz von KI-gestützten Such-, Analyse- und Entscheidungsunterstützungssystemen ermöglicht. Indem das Unternehmen Cybersicherheit, Daten-Governance und KI-Bereitschaft als ineinandergreifende Elemente behandelte, entwickelte es sich von einer reinen Dokumentenverwaltung zu einer modernen Intelligence-Plattform – getragen von Vertrauen, Transparenz und Automatisierung.

> *Wir verbringen mehr als 70 Prozent unserer Zeit damit, uns aus technologischer Sicht zu verteidigen, sei es gegen regulatorische oder Cybersicherheitsbedrohungen. Wir müssen wachsam bleiben, um die Daten der Bank und unserer Kunden zu schützen und über die neuesten Änderungen und Patches, die Sicherheitslücken schließen, auf dem Laufenden zu bleiben.*
>
> CTO und Geschäftsführer einer globalen Bank

Die Cyberbedrohungslandschaft für Daten und KI

Laut dem *Global Cybersecurity Outlook 2025* des Weltwirtschaftsforums „verändern GenAI-Tools die Landschaft der Cyberkriminalität, indem sie es Kriminellen ermöglichen, ihre Methoden zu verfeinern und ihre Techniken zu automatisieren und zu personalisieren. 47 Prozent der Organisationen nennen die Weiterentwicklung von Angriffstechniken im Zusammenhang mit generativer KI als größte Sorge - und Cyberkriminelle nutzen die Effizienz von KI, bereits um irreführende Kommunikation zu automatisieren und zu personalisieren.
Rund 42 Prozent der Unternehmen wurden im vergangenen Jahr Opfer erfolgreicher Social-Engineering-Angriffe – eine Zahl, die mit der fortschreitenden Entwicklung und dem missbräuchlichen Einsatz von KI weiter steigen wird.[24]

Cybersicherheit für KI in Unternehmen muss aus einer multidimensionalen Perspektive betrachtet werden – einer Perspektive, die das gesamte Spektrum der Bedrohungen einbezieht: Daten, Modelle, Infrastruktur und menschliche Interaktion. Die Cyberbedrohungslandschaft für Daten und KI umfasst sowohl technologische als auch organisatorische und verhaltensbezogene Risiken. Dabei bleiben klassische Bedrohungen wie unbefugter Zugriff, Insider-Risiken und Ransomware bestehen und gefährden weiterhin kritische Unternehmensdaten.
Da immer mehr Daten und Prozesse in Cloud-Umgebungen verlagert werden, wächst die gesamte Angriffsfläche stetig. Bedrohungsakteure nutzen weiterhin Schwächen im Identitätsmanagement, in der Netzwerksegmentierung und in anfälliger Software aus. Diese etablierten Angriffsformen bilden die Grundlage für noch komplexere Methoden, die gezielt die zunehmende Abhängigkeit von KI ausnutzen.

Zu den sich wachsenden Angriffsflächen gehören die Datenpipelines, die für das Training großer Sprachmodelle (LLMs) verwendet werden, sowie die Modelle selbst. Forscherinnen und Forscher von IBM und der Carnegie Mellon University stellten fest: *„Die zunehmende Verwendung großer Sprachmodelle (LLMs), die von Dritten trainiert werden, gibt angesichts der Sicherheitslücken von LLMs Anlass zu ernsthaften Bedenken.“ Nachweislich können Angreifer diese Schwachstellen verdeckt durch Vergiftungsangriffe (Data Poisoning) ausnutzen, um unerwünschte oder manipulative Ergebnisse zu erzeugen.“*[25]

Neben Modellvergiftung treten immer häufiger weitere sicherheitsrelevante Risiken wie Datenexfiltration und Prompt-Injection auf, wobei letztere zu den größten Herausforderungen für die Sicherheit großer Sprachmodelle zählt.

Neu entstehende KI-spezifische Bedrohungen schaffen Schwachstellen, die über herkömmliche Datenpannen hinausgehen. Da sich diese Angriffsmethoden kontinuierlich weiterentwickeln, ist es kaum möglich, eine vollständige und stets aktuelle Liste aller Risiken zu erstellen. Zu den häufigsten Arten gehören jedoch:

- Data Poisoning-Angriffe
- Hintertürangriffe
- Angriffe des Gegners (Ausweichangriffe)
- Modellinversionsangriffe
- Angriffe zur Ausnutzung von Vorurteilen

Das folgende Diagramm zeigt den Lebenszyklus eines KI-Modells, dargestellt mit den jeweiligen Angriffstypen, die in jeder Phase auftreten können. Der Zyklus beginnt mit der **Datenerfassung** und -aufbereitung, die in die **Trainingsphase** einfließt, in der das Modell lernt. Daraus entsteht ein trainiertes **Modell**, das in der **Inferenzphase** eingesetzt wird – dem Prozess, bei dem Vorhersagen, Klassifizierungen, Entscheidungen oder generierte **Antworten** erzeugt werden.

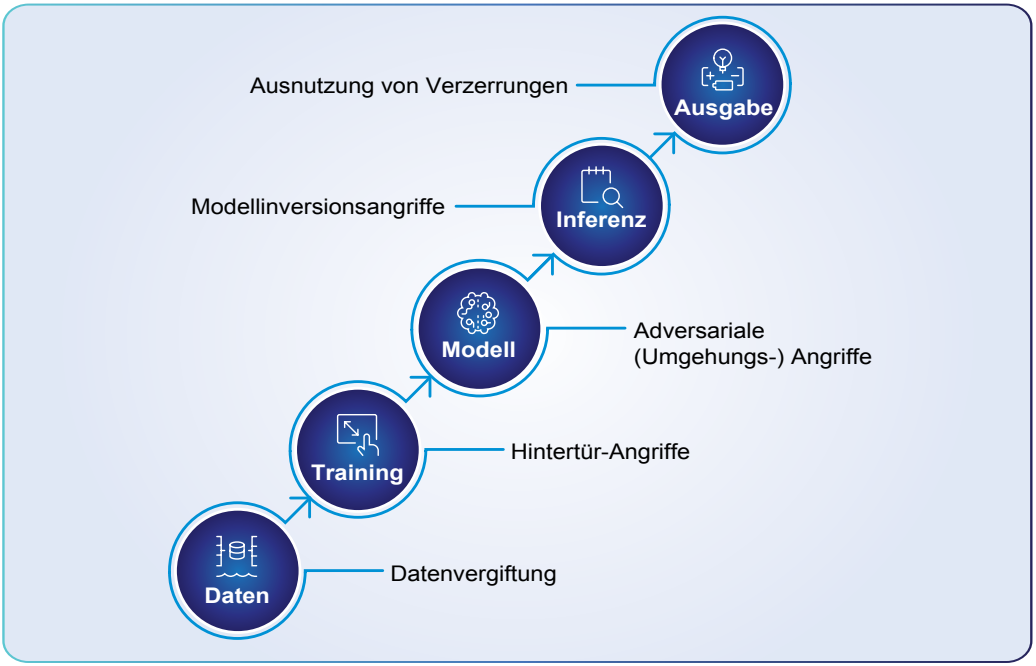


Abbildung von Cyberangriffen auf KI-Modelle

Das Verständnis dafür, wie verschiedene Cyberangriffe mit dem Lebenszyklus von KI-Modellen zusammenhängen, verdeutlicht, an welchen Punkten Bedrohungen entstehen können. Die folgenden Angriffstypen zeigen, wie Schwachstellen in unterschiedlichen Phasen des Modells ausgenutzt werden.

1. Data Poisoning

So genannte Data-Poisoning-Angriffe treten häufig in der Phase vor dem Training der Daten auf – also während der Datenerfassung und -verarbeitung. Dabei schleusen Angreifer bösartige Eingaben in den Trainingsdatensatz ein, wodurch die Lernprozesse des Modells gestört und seine Integrität und Zuverlässigkeit beeinträchtigt werden. Das grundlegende Problem liegt in einer fehlerhaften Annahme: Die meisten Lernalgorithmen setzen voraus, dass die Trainingsdaten sauber und repräsentativ die Realität abbilden. In sicherheitssensiblen Umgebungen trifft diese Annahme schlichtweg nicht zu.[26]

2. Hintertürangriffe

Hintertürangriffe sind eine spezielle Form der Data Poisoning, bei der während des Trainings ein Auslösemuster im Modell verborgen wird. Das Modell verhält sich bei regulären Eingaben unauffällig, gibt jedoch eine manipulierte Ausgabe zurück, sobald der Auslöser erkannt wird. Diese Angriffe sind schwer zu entdecken, da sie inaktiv bleiben, bis die Auslösebedingung eintritt. Ein Angreifer kann so ein bösartig trainiertes neuronales Netz (BadNet) erzeugen, das bei Trainings- und Validierungsdaten hervorragende Ergebnisse liefert, sich aber bei bestimmten, gezielt gewählten Eingaben absichtlich fehlerhaft verhält.[27]

3. Gegnerischer Angriffe

Eine weitere häufige Angriffsart ist der gegnerische Angriff. Der entsteht, wenn Angreifer die Eingaben eines KI-Modells gezielt manipulieren, um falsche oder irreführende Ergebnisse zu provozieren. Manchmal sind diese Veränderungen so gering, dass sie nicht erkennbar sind, aber sie können das Verhalten verändern und die Sicherheit in KI-Anwendungsfällen wie der medizinischen Bildgebung oder der autonomen Navigation beeinträchtigen.[28]

4. Modellinversionsangriffe

Modellinversionsangriffe stellen eine Bedrohung für die Privatsphäre und personenbezogene Daten dar. Dabei wird versucht, *„sensible Eingangsdaten aus Modellparametern, Ausgaben oder Zwischenrepräsentationen zu rekonstruieren.“*[29] Mit anderen Worten: Im Kern wird das Modell rückwärts analysiert, um private Trainingsdaten offenzulegen, die ursprünglich zur Modellerstellung verwendet wurden.

5. Angriffe zur Ausnutzung von Verzerrungen

Diese Angriffsform basiert auf der gezielten Ausnutzung vorhandener Verzerrungen (Bias) in Datensätzen. Angreifer verstärken bestehende Verzerrungen, um die Entscheidungsfindung des Modells zu manipulieren oder zu verfälschen. Diese Angriffe unterscheiden sich von Data Poisoning dadurch, dass sie keine neuen Daten in den Datensatz einfügen. Stattdessen nutzen sie inhärente Ungleichheiten, die bereits in den Daten vorhanden sind, um einen Angriff durchzuführen.[30]

Sowohl im öffentlichen als auch im privaten Sektor gehen die Risiken inzwischen über die technische Kompromittierung (z. B. den Zugriff auf das System) hinaus und umfassen auch die Manipulation von Daten (z. B. das Verändern oder Vergiften von Daten). Diese fünf Beispiele verdeutlichen, wie Bedrohungsakteure unterschiedliche Phasen des Lebenszyklus von KI-Modellen angreifen. Regierungen im öffentlichen Sektor waren mit Ransomware-Angriffen auf kritische Infrastrukturen konfrontiert, die zentrale Dienste beeinträchtigten. Unternehmen im Privatsektor wiederum sahen sich mit Modellstörungen konfrontiert, bei denen Bedrohungsakteure Websites und Empfehlungssysteme manipulierten.

Darüber hinaus wird generative KI zunehmend genutzt, um Fehlinformationen zu verbreiten. Diese Angriffe nutzen Modellverzerrungen aus und untergraben das Vertrauen der Öffentlichkeit in KI. Diese Entwicklungen machen deutlich, dass Cybersicherheit für Unternehmensdaten und KI über den bloßen Schutz von Systemen hinausgeht. Im Mittelpunkt steht der Schutz der Integrität von Daten und Entscheidungen sowie die Wahrung des Vertrauens der Öffentlichkeit im kognitiven Zeitalter.

Grundlagen der Datensicherheit

Diese Analyse der Bedrohungen verdeutlicht ein zentrales Prinzip: Die Daten, die in Unternehmen zum Trainieren von KI-Modellen verwendet werden, müssen sicher und vertrauenswürdig sein. Mit der zunehmenden Nutzung von KI steigt nicht nur das Datenvolumen, sondern auch die Sensibilität der Informationen, die Unternehmen verwalten. Je weiter sich Systeme in Richtung agentenbasierter und allgemeiner Intelligenz entwickeln, desto stärker wächst die Abhängigkeit von privaten, geschützten Datensätzen. Für Organisationen, die eigene Modelle auf Basis proprietärer Daten entwickeln, ist die Gewährleistung von Vertraulichkeit und Integrität über den gesamten Lebenszyklus hinweg entscheidend. Die Grundlage dafür bildet ein Security-by-Design-Ansatz, der robuste technische Kontrollen mit klar definierten Governance-Mechanismen verbindet. Diese Strategie ist unerlässlich, um Informationen zu sichern, regulatorische Anforderungen zu erfüllen und das Vertrauen in KI-gestützte Systeme zu erhalten.

Schutz des Datenlebenszyklus

Der Datenlebenszyklus umfasst mehrere Phasen – Erfassung, Speicherung, Übertragung, Verarbeitung, Archivierung und Entsorgung. Der Schutz in jeder dieser Phasen erfordert eine abgestimmte Kombination aus präventiven, aufspürenden und korrektiven Maßnahmen, um Bedrohungen frühzeitig zu erkennen, abzuwehren und die Resilienz der Systeme zu stärken.

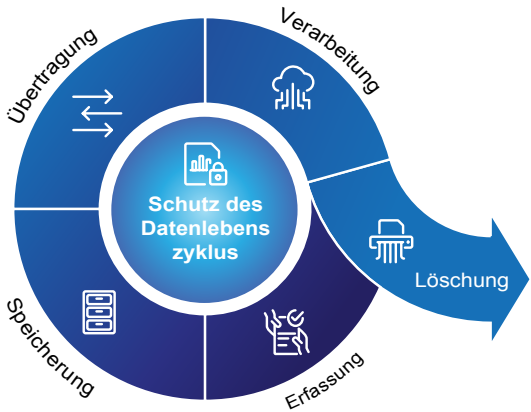


Abbildung von Cyberangriffen auf KI-Modelle

Erfassung

Die Datenerfassung muss kontrolliert und zielgerichtet erfolgen, da bereits in dieser Phase das Risiko einer Kompromittierung entstehen kann. Entscheidend ist, welche Daten erfasst werden und zu welchem Zweck. Der internationale Standard ISO/IEC 27001:2022 bietet einen Rahmen, der Organisationen dabei unterstützt, Informationen über ihren gesamten Lebenszyklus hinweg zu schützen. Er definiert Kontrollkategorien, die sicherstellen, dass Datenerfassung und -verarbeitung rechtmäßig, fair und transparent erfolgen.[31]

Speicherung

Nach der Erfassung müssen Daten sicher gespeichert werden. Dies kann entweder in lokaler Infrastruktur oder in Cloud-Umgebungen sein. Zu den zentralen Datenschutzmaßnahmen gehören die Verschlüsselung ruhender Daten, strikte Zugriffskontrollen und die Trennung sensibler Daten. Ergänzend kann eine unveränderliche Speicherung Teil einer umfassenden Datensicherungsstrategie sein, die Schutz vor Cyberangriffen wie Ransomware bietet. Diese Maßnahmen bilden zugleich einen wesentlichen Bestandteil einer Zero Trust-Datenschutzarchitektur, die im weiteren Verlauf näher erläutert wird.

Übertragung

Daten, die zwischen Systemen übertragen oder verteilt werden, sind anfällig für das Abfangen. Zum Schutz dieser Daten während der Übertragung werden technische Methoden wie Verschlüsselung eingesetzt. Zwar kann Verschlüsselung das Abfangen nicht verhindern, sie stellt jedoch sicher, dass abgefangene Daten ohne den entsprechenden Schlüssel unbrauchbar bleiben.[32]

Verarbeitung

Die Datenverarbeitungsphase ist besonders kritisch, da hier Daten abgefangen, manipuliert oder unbefugt eingesehen werden können. Strenge Zugriffskontrollen und klar definierte Berechtigungen sind unerlässlich, um Datenschutzverletzungen zu verhindern. Ein zentrales Risiko besteht in der Verwendung sensibler Datensätze für das Training oder die Analyse von KI-Modellen, wenn keine geeigneten Schutzmechanismen vorhanden sind. Um diese Risiken abzuschwächen, wurden neue Berechnungsmethoden entwickelt. Homomorphe Verschlüsselung ermöglicht Berechnungen auf verschlüsselten Daten, ohne diese zuvor entschlüsseln zu müssen, und schützt so die Vertraulichkeit während der Verarbeitung. Ebenso fördert föderiertes Lernen einen sicheren, verteilten Ansatz für das KI-Training. Dies ermöglicht das lokale Trainieren von Modellen über mehrere dezentrale Datensätze hinweg. Dieser Ansatz, „den Code zu den Daten zu bringen“, minimiert die Notwendigkeit, sensible Daten zentral zu speichern, verringert dadurch das Risiko einer Offenlegung und gewährleistet gleichzeitig eine stabile Modellleistung.[33]

Datenentsorgung oder -löschung

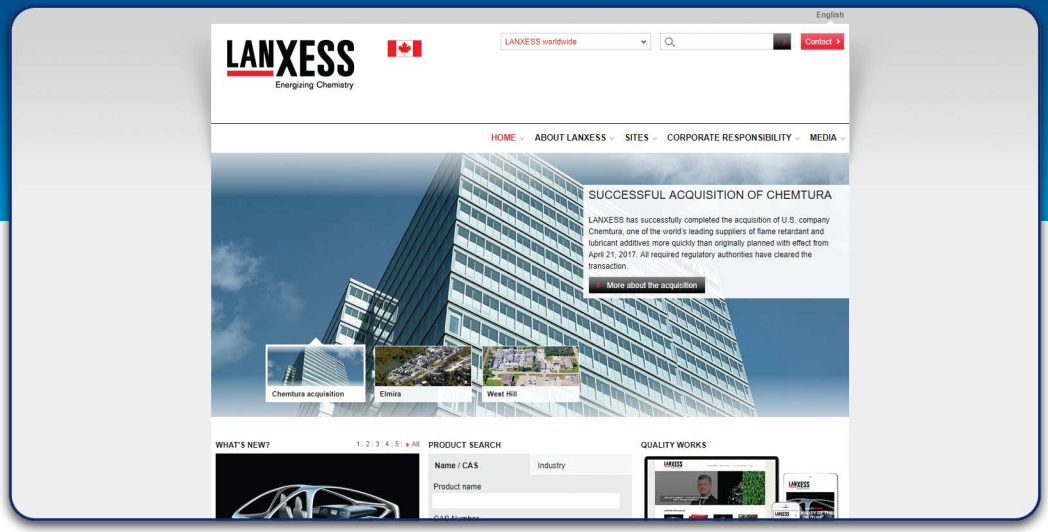
Die sichere Löschung von Daten bildet die letzte Phase des Lebenszyklus und stellt sicher, dass veraltete oder redundante Informationen endgültig und nachvollziehbar entfernt werden. Gemäß den Datenschutzbestimmungen wie Artikel 17 der DSGVO („Recht auf Löschung“ bzw. „Recht auf Vergessenwerden“) müssen Organisationen nachweisen können, dass Löschanforderungen ordnungsgemäß und vollständig umgesetzt wurden.[34]

Die einzelnen Phasen des Datenlebenszyklus sind eng miteinander verknüpft – eine Schwachstelle in einer Phase kann die Sicherheit des gesamten Zyklus gefährden. Ein ganzheitliches Verständnis der Risiken über alle Phasen hinweg ist daher entscheidend, um eine Zero Trust-Datenschutzstrategie zu entwickeln, die präventiv, nachvollziehbar und überprüfbar ist.

In der folgenden Fallstudie wird gezeigt, wie ein führendes Chemieunternehmen ein Enterprise Information Management (EIM)-System einsetzt, um seinen Datenlebenszyklus zentral zu steuern, regulatorische Anforderungen einzuhalten und Datensicherheit über Prozesse, Partner und Standorte hinweg zu gewährleisten.

Fallstudie

LANXESS



LANXESS

Das Kerngeschäft von LANXESS umfasst die Entwicklung, Herstellung und Vermarktung chemischer Zwischenprodukte, Additive, Spezialchemikalien und Kunststoffe. Im Gespräch mit dem Prozess-Experten des Unternehmens (ECM) wird deutlich, wie zentral die Rolle des Informationsmanagements für die globale Compliance und Effizienz ist:

„Angesichts der Komplexität unseres Produktportfolios hinterlassen unsere Herstellungsprozesse – von chemischen Zwischenprodukten über Additive und Spezialchemikalien bis hin zu Kunststoffen – umfangreiche Dokumentationsspuren, die sich von der wissenschaftlichen Forschung bis in den Vertrieb und das Marketing erstrecken.

Ein typischer Ausgangspunkt für neue Forschungsaktivitäten ist die Anfrage eines Kunden nach einer spezifischen Produktfunktion. In vielen Fällen werden externe Partner einbezogen, was hohe Anforderungen an sicheren Zugriff, Nachvollziehbarkeit und Zusammenarbeit stellt. Da wir weltweit produzieren und vertreiben, müssen alle Produkte, Prozesse und Dokumentationen internationalen Vorschriften entsprechen.

Eine Enterprise Content Management (ECM)-Plattform gewährleistet dabei die Informationskonformität – von der Forschung über die von Ingenieuren definierten Verfahren bis hin zur Massenproduktion, dem Anlagenbetrieb, dem Vertrieb und der Vermarktung.
Wir verarbeiten täglich große Mengen Papier. Jeder Schritt im Produktionsprozess muss den Vorschriften entsprechen und nachvollziehbar dokumentiert sein – insbesondere, da wir in 25 Ländern tätig sind, die jeweils unterschiedliche regulatorische Anforderungen haben.

Neben der Compliance führt ein effektives Informationsmanagement zu messbaren Vorteilen in Effizienz und Produktivität – insbesondere durch den schnelleren Zugriff auf relevante Informationen. Damit diese Vorteile realisiert werden können, müssen wir unseren internen Kunden aufzeigen, wie der Einsatz der Technologie ihre Arbeit erleichtern wird. ECM liefert uns die Werkzeuge, die wir benötigen, um Compliance und Sicherheit mit Benutzerfreundlichkeit in Einklang zu bringen."

Zero Trust-Architektur für Enterprise-KI

Nachdem die zentralen Cyberbedrohungen für Daten und KI analysiert wurden – insbesondere entlang des Datenlebenszyklus und der kritischen Angriffspunkte –, stellt sich die Frage, wie Unternehmen diesen Bedrohungen systematisch begegnen können. Das National Institute of Standards and Technology (NIST) definiert in seiner Publikation SP 800-207 die Zero Trust-Architektur (ZTA) als einen strategischen Sicherheitsansatz, der kein implizites Vertrauen innerhalb eines Netzwerks voraussetzt. Das Modell basiert auf der Idee, dass man nie vertrauen und immer überprüfen sollte, und diese Philosophie muss jede Zugangsentscheidung bestimmen. Anstatt sich ausschließlich auf traditionelle Schutzmechanismen wie Firewalls oder VPNs zu verlassen, setzt das Zero Trust-Modell auf kontinuierliche Verifizierung, Zugriffskontrolle und kontextabhängige Bewertung über alle Assets, Benutzer und Datenflüsse hinweg.

Gemäß NIST SP 800-207 verlagert Zero Trust den Schwerpunkt von der reaktiven Verteidigung hin zu einem identitäts- und datenorientierten Schutzmodell mit folgenden Kernprinzipien:

- **Identitätszentrierter Schutz**: Jede Zugriffsanfrage muss in Echtzeit authentifiziert, autorisiert und überprüft werden.
- **Prinzip der minimalen Berechtigungen**: Benutzer, Anwendungen und Systeme erhalten ausschließlich die Zugriffsrechte, die sie unbedingt benötigen, um ihre Aufgaben zu erfüllen.
- **Dynamische Richtliniendurchsetzung**: Zugriffsentscheidungen basieren auf aktuellen Kontextfaktoren wie Nutzerverhalten, Gerätezustand, Standort und Datensensibilität und werden kontinuierlich angepasst.
- **Mikrosegmentierung**: Netzwerke werden in kleine, klar abgegrenzte Zonen unterteilt, um die laterale Bewegung von Bedrohungsakteuren im Falle einer Kompromittierung zu verhindern.
- **Transparenz und Analysen**: Eine fortlaufende Überwachung und Bedrohungserkennung gewährleisten, dass abweichendes Verhalten automatisch erkannt und Gegenmaßnahmen ausgelöst werden.

Zero Trust ist keine einheitliche Lösung, sondern wird durch eine Kombination von Technologielösungen erreicht, darunter Identitäts- und Zugriffsmanagement (IAM), Multi-Faktor-Authentifizierung (MFA), Verschlüsselung, kontinuierliche Überwachung und automatisierte Richtliniendurchsetzung.[35] KI-gestütztes IAM wird zu einer kritischen Sicherheitskomponente eines jeden Unternehmenssystems werden.

Die folgende Fallstudie zeigt, wie ein lateinamerikanisches Unterhaltungsunternehmen diese Technologielösungen in seinem Plan zur Weiterentwicklung hin zu einem Zero Trust-Sicherheitsmodell kombiniert.

Fallstudie

Ein lateinamerikanisches Unterhaltungsunternehmen

Mit Millionen von Kunden und Tausenden Beschäftigten in mehreren Ländern stand ein führendes lateinamerikanisches Unterhaltungsunternehmen bei der Verwaltung von Identitäten und Zugriffsrechten für eine große, verteilte Belegschaft vor zunehmenden Herausforderungen. Im Laufe der Zeit erschwerten fragmentierte Systeme und manuelle Prozesse die Kontrolle über mehr als 15.000 Benutzeridentitäten und über 400 Anwendungen. Das Fehlen einer zentralen Governance verlangsamte die Reaktionszeiten, schuf Sicherheitslücken und erschwerte den Übergang zu einem Zero Trust-Modell.

Um diese Probleme zu lösen, führte das Unternehmen ein umfassendes Identity-Governance-and-Administration-(IGA)-Framework ein, das globale Identitätsdaten in einer einzigen Quelle der Wahrheit und einem zentralen Kontrollpunkt bündelt. Die Plattform ist in HR-Systeme, Active Directory und zahlreiche Unternehmensanwendungen integriert und automatisiert die Bereitstellung, Deaktivierung und Überprüfung von Zugriffsrechten, wodurch der manuelle Aufwand halbiert wurde. Intelligente Warnmeldungen, kontinuierliche Verifizierung und rollenbasierte Zugriffskontrollen stärkten die Einhaltung von Vorschriften, reduzierten Risiken und setzten das Prinzip der minimalen Rechtevergabe um.

Die Ergebnisse ließen nicht lange auf sich warten. Das Unternehmen erhielt einen durchgängigen Einblick in mehr als 15.000 Identitäten, optimierte die Zugriffsverwaltung und stärkte seine Sicherheitslage in einem globalen Umfeld. Mit Identity Governance als zentralem Bestandteil der Cybersicherheitsstrategie ist das Unternehmen nun in der Lage, sein Zero Trust-Modell weiterzuentwickeln und denselben Standard an Strenge und Automatisierung auch auf den Schutz seiner Daten, Anwendungen und KI-gesteuerten Abläufe im gesamten digitalen Umfeld auszudehnen.

KI-Sicherheit und Modellschutz

Zero-trust, as we just reviewed, provides a philosophy and strategy for protecting access, but we Zero Trust bietet, wie zuvor erläutert, eine Philosophie und Strategie zum Zugriffschutz. Darüber hinaus muss jedoch auch die Sicherheit der KI und der Schutz von Modellen in einem breiteren Kontext betrachtet werden. KI-Modelle unterscheiden sich von traditionellen IT-Systemen, da sie Logik mit Lernfähigkeit verbinden und sich durch kontinuierliche Anpassung weiterentwickeln. Zuvor wurden die Angriffsflächen im gesamten Lebenszyklus des KI-Modells betrachtet.

Zum Schutz vor Risiken kombinieren Organisationen bewährte Cybersicherheitsmethoden mit neuen, KI-spezifischen Ansätzen. Dazu zählen Schulungen für Teams und Modelle zu Angriffstechniken, der Einsatz von Modellwasserzeichen und Red-Team-Tests vor der Bereitstellung. Ein Red Team simuliert reale Cyberangriffe, um Schwachstellen in Systemen, Netzwerken oder menschlichen Prozessen aufzudecken.

Durch gezieltes Training lässt sich die Modellleistung verbessern, etwa indem das Modell während des Trainings gegnerischen Eingaben ausgesetzt wird, um seine Widerstandsfähigkeit gegenüber Angriffen zu erhöhen.[36] Modellwasserzeichen unterstützen die Nachverfolgbarkeit und helfen, unautorisierte Wiederverwendung zu erkennen.[37] Red-Team-Tests dienen dazu, Schwachstellen bereits vor dem produktiven Einsatz zu identifizieren.

In Kombination mit einem Zero Trust-Ansatz bilden diese Methoden wirkungsvolle Maßnahmen, um KI-Systeme vor Angriffen zu schützen. Sie sollten jedoch stets als Bestandteil einer umfassenden, unternehmensweiten KI-Sicherheitsstrategie definiert und kontinuierlich weiterentwickelt werden.

Blick in die Zukunft: Die Zukunft der Cybersicherheit für die KI

In diesem Kapitel wurde die zunehmende Bedeutung der Cybersicherheit im Zusammenhang mit KI in Unternehmen untersucht und die wachsende Zahl gezielter Angriffe auf KI-Systeme hervorgehoben. Berichte, wonach 62 Prozent der Organisationen von Deepfake-Angriffen betroffen waren, sowie die steigenden Sorgen über die feindlichen Fähigkeiten von GenAI verdeutlichen die Dringlichkeit, KI- und datenbezogene Risiken entschlossen anzugehen. Während Unternehmen private Daten nutzen, um ihre Effizienz zu steigern, setzen sie sich zugleich komplexen Schwachstellen aus, die sowohl ihre Modelle als auch ihre Daten gefährden.

Die Analyse verschiedener Angriffstypen hat gezeigt, wie diese Angriffe funktionieren, und Bedrohungen wie Data Poisoning, Backdoor-Angriffe und Modellinversionsangriffe beschrieben. Diese Risiken verdeutlichen einige der Grenzen traditioneller Cybersicherheitsansätze beim Schutz fortschrittlicher KI-Systeme. Ein tiefes Verständnis der Angriffspunkte entlang des Lebenszyklus von KI-Modellen ermöglicht es, potenzielle Schwachstellen frühzeitig zu erkennen und gezielte Strategien für Datenschutz und Modellsicherheit zu entwickeln.

Mit Blick auf die Zukunft müssen Organisationen proaktive, adaptive Sicherheitsrahmen aufbauen, die KI-gestützte Verteidigungsmechanismen nutzen, um KI-basierten Angriffen wirksam zu begegnen. Dazu gehören intelligente Systeme zur Bedrohungserkennung sowie dynamische Modelle zur Risikobewertung. Ebenso wichtig wird die enge Zusammenarbeit zwischen öffentlichem und privatem Sektor sein, um durch gemeinsame Investitionen in innovative Sicherheitslösungen die sichere Integration von KI-Technologien in Unternehmensumgebungen zu gewährleisten.

Während Unternehmen ihre Cyberabwehr verstärken, wird eine Wahrheit deutlich: Sicherheit und Vertrauen sind untrennbar miteinander verbunden. Der Schutz von KI-Systemen bedeutet nicht nur, Angriffe abzuwehren, sondern sicherzustellen, dass die zugrunde liegenden Daten korrekt, ethisch und zuverlässig bleiben. Im nächsten Kapitel steht die Grundlage vertrauenswürdiger KI im Mittelpunkt – die Datenverwaltung.

Die fünf Merksätze

1. **Ein eine Zero Trust-Architektur für alle Daten- und KI-Systeme einführen.**
   Ein Zero Trust-Sicherheitsmodell muss ohne implizites Vertrauen im Netzwerk aufgebaut werden. Um das Risiko sowohl interner als auch externer Sicherheitsverletzungen zu minimieren, sollten kontinuierlich Identitätsprüfung, Zugriffsrechte nach dem Prinzip der minimalen Berechtigungen, dynamische Richtliniendurchsetzung und Mikrosegmentierung durchgesetzt werden.

2. **Den gesamte Datenlebenszyklus mit integrierten Kontrollmechanismen absichern.**
   Alle Daten – von der Erfassung über Speicherung, Übertragung und Verarbeitung bis zur Entsorgung – müssen durch mehrstufige Sicherheitsmaßnahmen geschützt werden. Dazu gehören Verschlüsselung ruhender und übertragener Daten, strenge Zugriffskontrollen, unveränderliche Speicherung und die Einhaltung von Vorschriften wie DSGVO und ISO/IEC 27001:2022. Schwachstellen in einer Phase können den gesamten Schutz untergraben.

3. **KI-Modelle gegen neue Bedrohungen absichern.**
   Protokolle zur Abwehr KI-spezifischer Angriffe wie Data Poisoning, Ausnutzung von Hintertüren, Manipulation von Eingaben, Modellinversion und Ausnutzung von Verzerrungen sollten eingerichtet werden. Schulungen zu gegnerischen Angriffen, Modellierung von Wasserzeichen und regelmäßige „Red Team“-Tests sind durchzuführen, um vor der Bereitstellung Schwachstellen zu erkennen und zu beheben.

4. **Sicherheit von Anfang an in KI-Initiativen integrieren.**
   Sicherheit und Datenschutz müssen von Beginn an in jede KI- und Dateninitiative integriert werden. Funktionsübergreifende Teams aus den Bereichen Daten, IT, Compliance und Sicherheit sollten gemeinsam sicherstellen, dass technische Schutzmaßnahmen und Governance-Kontrollen fest in Entwicklung und Betrieb von KI-Modellen verankert sind.

5. **In proaktive, KI-gestützte Cybersicherheitslösungen investieren.**
   Es müssen Ressourcen bereitgestellt werden, um intelligente, adaptive Cybersicherheitslösungen auf der Grundlage von KI zu entwickeln und einzusetzen. Dazu sollten automatisierte Bedrohungserkennung, Risikobewertungsinstrumente und Echtzeitüberwachung gehören, um mit der Weiterentwicklung KI-gestützter Angriffsmethoden Schritt halten zu können. Die Zusammenarbeit mit Branchenkollegen und Partnern aus dem öffentlichen Sektor ist notwendig, um neuen Bedrohungen einen Schritt voraus zu sein.

Kapitel Fünf

Daten-Governance – Die Grundlage von vertrauenswürdiger Enterprise-KI

Bevor eine künstliche Intelligenz denken kann, muss sie den Informationen vertrauen, auf denen sie basiert. Die Governance macht dies möglich. Sie ist die Disziplin, die verstreute Inhalte in ein kohärentes, konformes und nutzbares Asset verwandelt, ein Asset, das intelligente Systeme sicher aufnehmen kann, ohne Sicherheit oder Integrität zu beeinträchtigen.

Enterprise Information Management betrachtet Governance als operatives Prinzip, nicht als Checkliste. Es ruht auf vier ineinander greifenden Säulen: Metadaten, Berechtigungen und Zugriffskontrolle, Aufbewahrung und Lebenszyklusmanagement sowie Revisionsfähigkeit. Jede dieser Säulen legt fest, wie sich Daten über ihren gesamten Lebenszyklus hinweg verhalten, und gemeinsam bilden sie das Rückgrat vertrauenswürdiger Intelligenz. In diesem Kapitel werden diese Säulen und ihre Beziehung zu optimierter, konformer und sicherer KI erläutert.

> *Forrester prognostiziert, dass sich die Ausgaben für KI-Governance-Software bis 2030 vervierfachen werden.*[38]

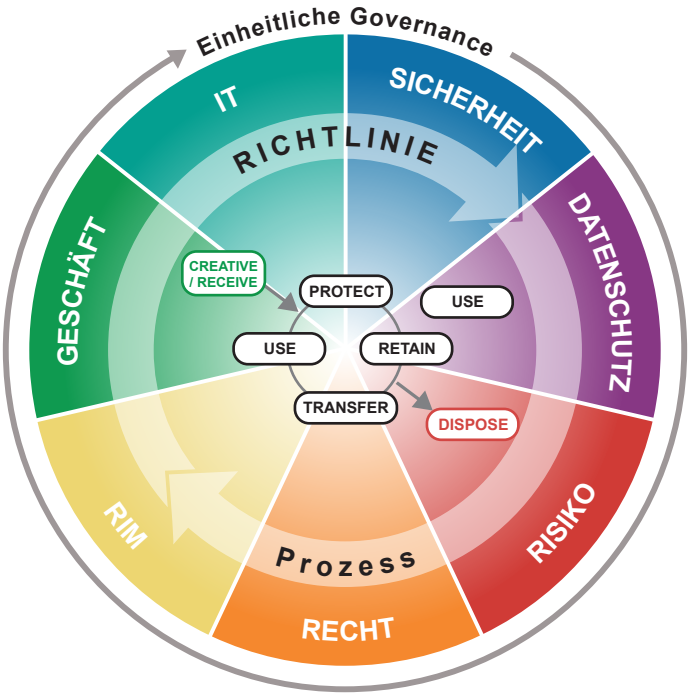
Gute Governance ist ein gutes Geschäft

Informations-Governance bezeichnet die Praxis, Richtlinien, Prozesse und Kontrollen zu implementieren, um Informationen im Einklang mit regulatorischen, rechtlichen, risikobezogenen, umweltbezogenen und betrieblichen Anforderungen zu verwalten. Mit wachsendem Umfang der Unternehmensdaten steigt auch der Bedarf an digitaler Governance, um sicherzustellen, dass die Daten verwaltet, gesichert und durchsuchbar bleiben.
Aus technologischer Sicht beruht Governance auf dem effektiven Management von Daten über ihren gesamten Lebenszyklus hinweg, von der Erstellung oder Erfassung und Klassifizierung bis zur langfristigen Archivierung oder Löschung.

Erfolgreiche Programme zur Informationsgovernance erfordern, dass Unternehmen die Prioritäten zur Minderung rechtlicher und geschäftlicher Risiken mit den Kosten für die Verwaltung sowohl unstrukturierter als auch strukturierter Daten in Einklang bringen. Damit eine Strategie zur Informations-Governance wirksam ist, müssen zentrale Ressourcen und Interessengruppen identifiziert, befähigt und unterstützt werden, Richtlinien in relevante Prozesse integriert sein, Schulungen und Weiterbildungen stattfinden, die technologische Infrastruktur optimiert werden und geeignete Lösungen einen sicheren und zuverlässigen Betrieb gewährleisten.

Der folgende Beitrag zeigt, wie gute Governance ASR Nederland zugutekommt, indem sie die Einhaltung von Vorschriften ermöglicht und durch verbesserten Kundenservice einen strategischen Vorteil schafft.

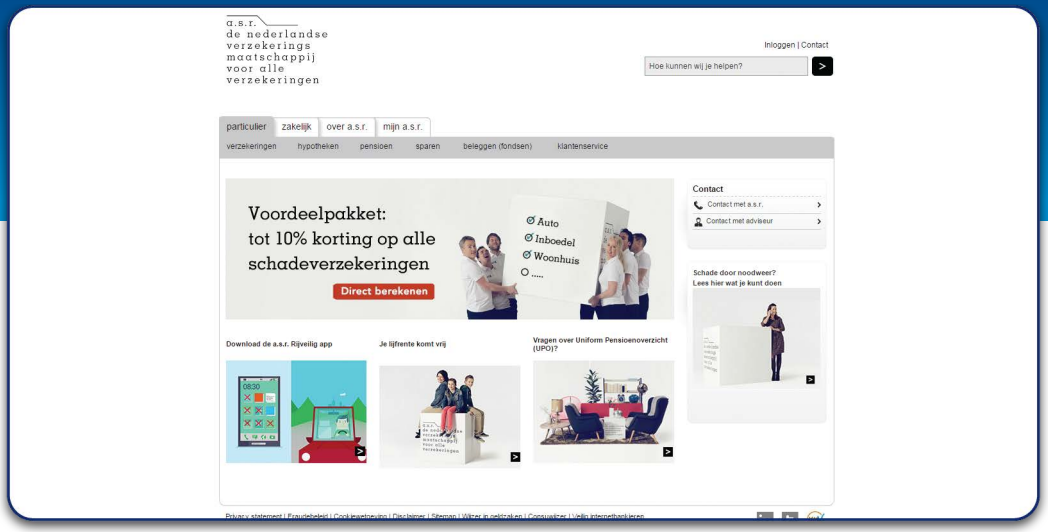
Wert, Risiko und Kosten in Einklang bringen



Referenzmodell für Informations-Governance[39]

Fallstudie

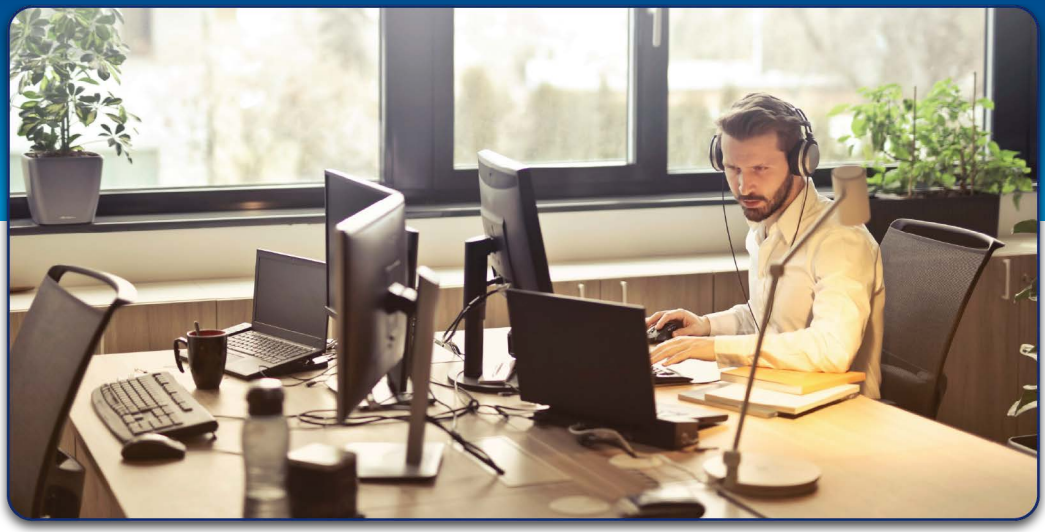
ASR Nederland



ASR Nederland

Eines der zentralen Geschäftsfelder von ASR ist die Erwerbsunfähigkeitsversicherung. Der Schadensregulierungsprozess erfolgte bisher auf Papierbasis. Medizinische und technische Informationen wurden zusammen in physischen in Ordnern abgelegt, die auch für nicht autorisiertes Personal zugänglich waren, was gegen das niederländische Datenschutzgesetz verstieß. Gleichzeitig nahm der Speicherbedarf für die stetig wachsende Zahl an Akten kontinuierlich zu. ASR erkannte die Notwendigkeit einer Lösung, die die Geschäftsprozesse verbessert, die Zusammenarbeit zwischen Abteilungen erleichtert, die Kosten im gesamten Unternehmen senkt und zugleich sicherstellt, dass nur befugte Personen auf vertrauliche Informationen zugreifen können.

Durch die Verbindung von Geschäftsprozessmodellierung mit Lösungen zur Prozessoptimierung gelang es, bestehende Abläufe zu modernisieren und auf gesetzliche Änderungen flexibel zu reagieren. Heute werden medizinische und technische Daten zu Leistungsanträgen getrennt verarbeitet und sind ausschließlich qualifiziertem Personal zugänglich, wodurch ASR die Anforderungen des Datenschutzes zuverlässig erfüllt. Gleichzeitig lässt sich der gesamte Schadenmanagementprozess messen, was der Unternehmensführung präzise Einblicke in Abläufe und Engpässe ermöglicht. Die flexible Systemumgebung schafft die Grundlage für ein modernes Benchmarking, das eine übergreifende Überwachung aller Geschäftsaktivitäten erlaubt.

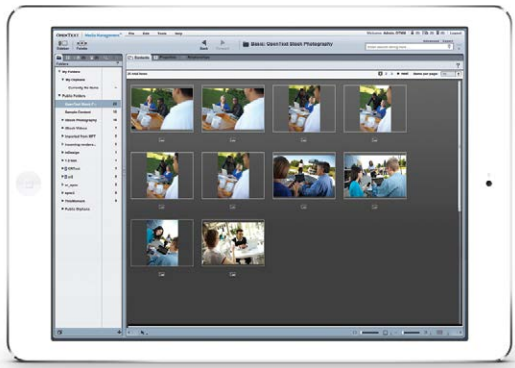


Die implementierte Lösung bietet für die Schadensbegrenzung ein unternehmensweites Standardsystem, das die interne Effizienz erheblich verbessert und die Produktivität spürbar gesteigert hat. Inzwischen werden 80 Prozent aller Schadensfälle fristgerecht bearbeitet, was es möglich macht, das Schadensbearbeitungs-Team um 25 % zu reduzieren. Auch Servicekosten und Entschädigungszahlungen konnten deutlich gesenkt werden, sodass ASR neue Produkte schneller einführen, regulatorische Anforderungen erfüllen und den Kundenservice insgesamt verbessern kann.

Nachdem der konkrete Nutzen von Data Governance deutlich geworden ist, werden im nächsten Abschnitt die vier tragenden Säulen einer soliden Governance-Struktur vorgestellt, die bereits zu Beginn des Kapitels skizziert wurden.

Säule 1: Metadaten – Der Kontext hinter jeder Entscheidung

Metadaten sind die DNA digitaler Informationen, der verborgene Kontext, der Systemen mitteilt, was etwas ist, woher es stammt und wie es verwendet werden soll. Sie verknüpfen Inhalte mit dem Geschäftszweck und verwandeln Rohdaten in steuerbare, durchsuchbare und bedeutungsvolle Informationen.



Metadaten in einem Digital Asset Management System

Daten gelangen auf vielen Wegen in ein Unternehmen. Ein Teil entsteht digital, erzeugt von Menschen mithilfe von Textverarbeitungsprogrammen, Tabellenkalkulationen, CAD-Software oder E-Mail-Clients. Andere Metadaten entstehen automatisch innerhalb von Geschäftssystemen und werden von Enterprise-Resource-Planning-Systemen (ERP), Customer-Relationship-Management-Systemen (CRM) oder Datenbanken mit klar definierten Schemata und relationalen Strukturen generiert. Weitere Informationen stammen aus analogen Quellen, die über Scanner erfasst werden. Hinzu kommen Maschinendaten aus Sensoren, Protokollen und Telemetrie, ergänzt durch Web- und Social-Media-Inhalte aus Intranets, Kollaborationstools und Unternehmensportalen. Darüber hinaus produzieren Organisationen eine wachsende Anzahl von Multimedia-Inhalten – etwa Schulungsvideos, Marketingmaterialien oder aufgezeichnete Besprechungen –, die alle sowohl betrieblichen als auch rechtlichen Wert besitzen.

Die Bewältigung dieser Vielfalt erfordert eine Instanz, die Ordnung schafft und Kontrolle ermöglicht. Metadaten, also Klassifizierungen, Aufbewahrungskennzeichnungen, Herkunftsangaben und Sensibilitätsstufen, übernehmen diese Steuerungsfunktion. Sie ermöglichen automatisierte Berechtigungsvergabe, gezielte Suche, Versionskontrolle und die Einhaltung von Aufbewahrungsrichtlinien. Gleichzeitig bilden sie die Grundlage für verantwortungsvolle KI in Unternehmen, da ein Modell ohne Metadaten nicht zwischen Entwurf und Endfassung oder zwischen öffentlicher Broschüre und vertraulicher Kundendatei unterscheiden kann.

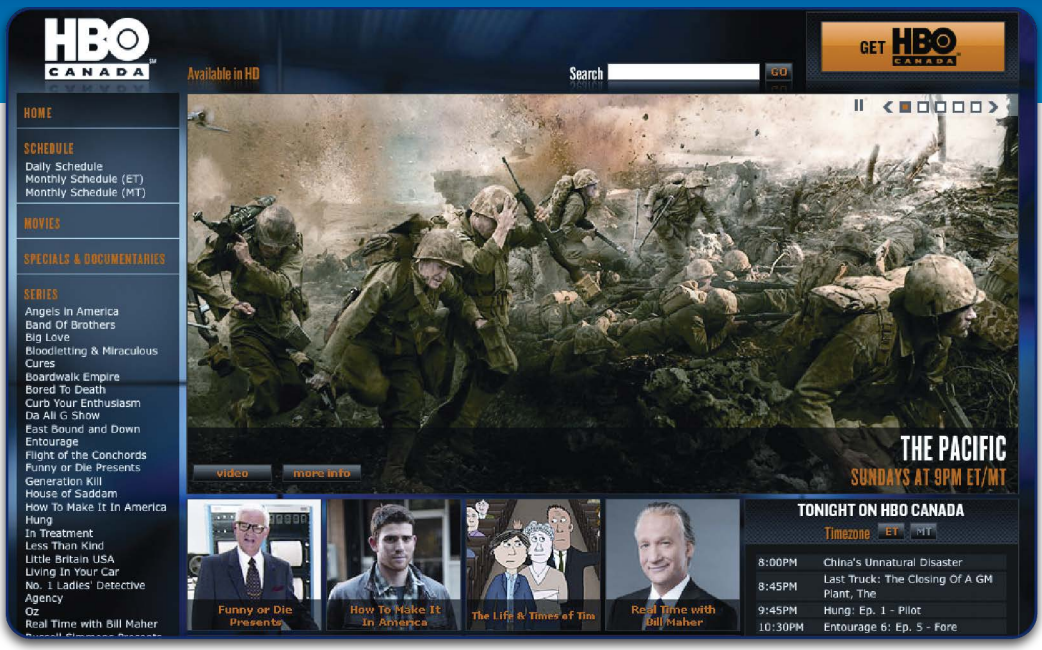
Metadaten liefern den entscheidenden Kontext darüber, wer eine Datei erstellt hat, wann sie geändert wurde, woher sie stammt und wie sensibel ihr Inhalt ist. Diese Details verleihen Rohdaten Bedeutung und helfen sowohl Menschen als auch intelligenten Systemen zu erkennen, welche Informationen vertrauenswürdig sind, welche geteilt werden dürfen und welche geschützt bleiben müssen.

Fehlen einheitliche Metadaten, wird das Training von KI-Systemen unzuverlässig und potenziell unsicher. Enterprise Information Management (EIM) bietet daher den Rahmen, um Governance direkt auf Informationsmodelle abzubilden und sicherzustellen, dass Automatisierung und KI dieselben geschäftlichen, rechtlichen und ethischen Grenzen respektieren, die auch für vertrauenswürdige Daten gelten.

Metadaten sind keine statische Kennzeichnung, sondern ein lebendiges System, das Richtlinien durchsetzt und maschinelles Lernen steuert. Mit wachsender Reife der KI werden sie zum entscheidenden Bindeglied zwischen verwalteten Daten und intelligentem Handeln. Wie dieses Prinzip in der Praxis funktioniert, zeigt HBO. Das Unternehmen nutzt Metadaten, um digitale Assets über ihren gesamten Lebenszyklus hinweg zu konsolidieren und effizient zu verwalten.

Fallstudie

HBO



Medienmanagementsystem von HBO

HBO, Amerikas führender Premium-Fernsehsender, bietet ein breites Spektrum digitaler Medieninhalte – von Blockbuster-Filmen und innovativen Eigenproduktionen bis zu provokanten Dokumentarfilmen, Konzerten und Boxveranstaltungen. Um diese Vielfalt effizient zu verwalten, suchte HBO nach einer Lösung, mit der sich digitale Inhalte sowohl innerhalb des Unternehmens als auch im Verbund der Time-Warner-Gruppe einfach abrufen und teilen lassen. Das neue System sollte große Mengen an Inhalten verarbeiten können und zugleich die unterschiedlichen Datenbanken, Arbeitsabläufe und Anwendungsfälle der beteiligten Organisationen berücksichtigen. Ebenso wichtig waren Benutzerfreundlichkeit und Skalierbarkeit, um die alltägliche Arbeit in allen Bereichen zu unterstützen.

Die Implementierung des Medienmanagementsystems umfasste sämtliche digitalen Fotos von HBO und diente als gemeinsame Plattform für Marketing, Promotion, Werbung und Vertrieb. Die gespeicherten Assets reichen von Drehortaufnahmen aus Filmproduktionen bis zu professionellen Porträts der HBO-Stars.

Ein zentraler Bestandteil der Gesamtstrategie war die präzise Verwaltung der Metadaten. Bereits in einem frühen Stadium werden alle Assets mit relevanten Metadaten versehen – etwa mit Vertragsinformationen –, damit diese Angaben den gesamten Lebenszyklus begleiten. Der Prozess des Meta-Taggings wird durch eine integrierte Workflow-Komponente sichergestellt, die eine konsistente Metadatenvergabe erzwingt. Das digitale Asset-Management-System von HBO wird heute von allen Regionalbüros genutzt und umfasst mehr als 325.000 Assets. Es sorgt dafür, dass Inhalte schnell auffindbar, rechtlich abgesichert und strategisch nutzbar bleiben.

> *Wenn Governance die Regeln definiert, sind Genehmigungen das Instrument, das sie durchsetzt – Entscheidung für Entscheidune.*

Säule 2: Berechtigungen und Zugriffskontrolle – Wer darf was, wann und warum sehen?

Berechtigungen markieren die Grenzen des Vertrauens. Sie legen fest, wer Informationen einsehen, bearbeiten oder weitergeben darf und unter welchen Bedingungen dies geschieht. Seit Jahrzehnten bilden diese Prinzipien den Kern des Schutzes von Unternehmens- und Personendaten. im Zeitalter der künstlichen Intelligenz gewinnen sie jedoch eine neue Bedeutung und Dringlichkeit. Jede Entscheidung, die ein intelligentes System trifft, hängt vom Umfang seines Zugriffs ab: welche Daten es lesen darf, aus welchen Quellen es lernt und welche Aktionen es ausführen kann.

In einer modernen Informationslandschaft sind Berechtigungen weit mehr als technische Schalter. Sie bilden die operative Ebene der Governance, auf der Richtlinien tatsächlich umgesetzt werden. Mechanismen wie Versionskontrolle, Workflow-Genehmigungen, Datensatzsperrungen und selektive Veröffentlichungen beruhen auf klar definierten Berechtigungsmodellen. Solche Modelle regeln nicht nur, was Benutzer tun dürfen, sondern auch, wann und aus welchem Grund. Ein Dokument, das heute bearbeitet werden kann, könnte morgen durch ein reguliertes Verfahren oder eine rechtliche Sperre blockiert sein. Diese dynamische Steuerung sorgt dafür, dass Informationen nachvollziehbar und vertrauenswürdig bleiben, auch wenn sie komplexe Lebenszyklen durchlaufen oder in kollaborativen Umgebungen genutzt werden.

Berechtigungs- und Zugriffskontrolle

Moderne Systeme für Enterprise Information Management erreichen ihre Genauigkeit durch fein abgestufte Berechtigungen. Jedes Objekt – ob Dokument, Ordner, Workflow oder Bild – besitzt ein eigenes Sicherheitsprofil, das festlegt, welche Benutzer und Gruppen im gesamten System Zugriff erhalten. In großen Organisationen, in denen die Zahl der Benutzer leicht in die Hunderttausende gehen kann, entstehen daraus Milliarden einzigartiger Berechtigungskombinationen. Diese Komplexität ist unvermeidlich und notwendig. Denn ohne die Möglichkeit, Sicherheitsmaßnahmen bis ins kleinste Detail zu definieren, kann kein Informationssystem als wirklich sicher gelten. Gerade diese Flexibilität erlaubt es Unternehmen, die physischen Kontrollmechanismen eines geschützten Arbeitsumfelds digital und in großem Maßstab abzubilden.

Da künstliche Intelligenz zunehmend zu einem weiteren „Nutzer“ innerhalb dieser Systeme wird, müssen dieselben Berechtigungsstrukturen auch auf intelligente Agenten angewendet werden. Wenn ein Dokument vertraulich ist, muss die KI das auch wissen. Berechtigungen sind damit nicht länger nur ein Instrument der Kontrolle, sondern ein Ausdruck von Vertrauen. Sie stellen sicher, dass jede Person und jedes System innerhalb klar definierter Grenzen, mit gezieltem Zugriff und unter nachvollziehbarer Aufsicht mit Informationen arbeitet. So entsteht Rechenschaftspflicht auf allen Ebenen. Durch diese präzise Steuerung schützen Organisationen die Privatsphäre, sichern ihre Wettbewerbsvorteile und gewährleisten, dass KI-Systeme ausschließlich in den Bereichen operieren, aus denen sie lernen dürfen.

Wie sich dieses Prinzip in der Praxis umsetzen lässt, zeigt die nachfolgende Fallstudie einer europäischen Investmentbank. Sie veranschaulicht, wie gezielte Berechtigungsstrukturen zur Dokumentklassifizierung an unterschiedlichen Standorten dazu beitragen kann zugleich operative sowie Governance-Ziele zu erfüllen.

Fallstudie

Eine europäische Bank

Eine große europäische Investmentbank finanziert Kapitalprojekte, die mit den politischen Zielen der Europäischen Union im Einklang stehen – die buchstäbliche Infrastruktur eines stärker integrierten Europas. Da ihre Geschäftstätigkeit rund 150 Länder außerhalb der EU umfasst, ist ein sicherer und effizienter Fernzugriff auf Dokumente nicht bloß eine Annehmlichkeit, sondern eine geschäftskritische Voraussetzung. Um dies zu gewährleisten, implementierte die Bank ein System für Enterprise Information Management als Bestandteil einer umfassenden IT-Modernisierung, die sämtliche Kernprozesse – von der Kreditaufnahme über die Kreditvergabe bis zur Verwaltung – transformieren sollte. Das System wurde vollständig in die bestehende IT-Landschaft integriert, sodass Inhalte, Daten und Arbeitsabläufe nahtlos zwischen den Plattformen übertragen werden können. Dadurch bleibt die Konsistenz gewahrt, und alle Prozesse erfüllen die geltenden Compliance-Anforderungen.

Governance stehtm Zentrum dieser Modernisierung . Die Bank entwickelte eine unternehmensweite Taxonomie, die nicht nur die Kategorisierung von Inhalten festlegt, sondern auch deren Verbindung zu Geschäftsprozessen und regulatorischen Rahmenbedingungen. Aufbauend auf internationalen Best Practices, einschließlich der DIRKS-Methodik und der ISO-15489-Normen, wurde die Taxonomie mit einem ausgefeilten Zugriffskontrollmodell kombiniert, das auf den höchsten Klassifizierungsebenen angewendet wird. Gemeinsam bilden diese Modelle ein dynamisches Governance-Rahmenwerk. Die Taxonomie definiert, welche Informationen existieren und wo sie verortet sind, während das Berechtigungsmodell regelt, wer zu welchem Zeitpunkt und zu welchem Zweck darauf zugreifen darf. So entsteht eine digitale Wissenslandkarte, die Struktur, Verantwortlichkeiten und Entscheidungsbefugnisse der Organisation abbildet.

Diese disziplinierte Informationsarchitektur bildet die Grundlage für den gezielten Einsatz künstlicher Intelligenz. Dank der einheitlichen Taxonomie und der präzise abgestuften Berechtigungen kann die Bank KI-Tools trainieren, um Dokumente sicher zu identifizieren, zusammenzufassen und zu klassifizieren – im Bewusstsein, dass jede Aktion eines intelligenten Agenten denselben Regeln und Zugriffsbeschränkungen unterliegt wie bei einem menschlichen Benutzer. Governance sorgt damit nicht nur für Automatisierung, sondern stellt sicher, dass KI-Systeme innerhalb derselben Vertrauensgrenzen operieren, die die Organisation auch für ihre Mitarbeiter definiert.

Die Ergebnisse bestätigen den Ansatz. Zwei Monate nach der Einführung lag die Nutzerakzeptanz um 20 Prozent über den Prognosen, und sämtliche relevanten Datensätze aus neuen Kredit- und Darlehensgeschäften waren bereits vollständig im System erfasst. Innerhalb weniger Wochen wuchs das zentrale Repository auf über 600.000 Dokumente an und wurde wöchentlich um etwa 100.000 weitere ergänzt. Dieser Erfolg zeigt, dass Governance, Taxonomie und Zugriffskontrolle Innovationen nicht bremsen, sondern sie sicher und skalierbar machen.

Säule 3: Aufbewahrung und Lebenszyklusmanagement – Wissen, wann man behalten und wann man löschen sollte

Informations-Governance bedeutet weit mehr als die bloße Speicherung von Daten, es geht um die verantwortungsvolle Verwaltung des gesamten Informationslebenszyklus. Jeder Inhalt hat eine eigene Geschichte – von der Erstellung über die Nutzung und Überarbeitung bis hin zur Aufbewahrung und letztlich zur kontrollierten Löschung. Die Steuerung dieses Lebenszyklus ermöglicht es Organisationen, gesetzeskonform, effizient und nachhaltig zu handeln.

Gute Governance sichert auf diese Weise Sicherheit, Compliance und Geschäftskontinuität

Regulierte Informationen und Aufzeichnungen entstehen in allen Bereichen eines Unternehmens: in ERP- und CRM-Systemen, in E-Mails und Dokumenten, in gescannten Papieren, in Telemetriedaten, in medizinischen Geräten oder in Wartungssystemen der Luftfahrt. Der erste Schritt einer wirksamen Steuerung ist die Erfassung – also die kontrollierte Aufnahme dieser Informationen über digitale Poststellen, Systemschnittstellen oder APIs. Jeder Datensatz muss mit seinen Metadaten, Zeitstempeln und Herkunftsnachweisen erfasst werden, damit Authentizität und rechtliche Nachvollziehbarkeit gewährleistet bleiben. Von Beginn an müssen Metadaten, Zeitstempel und Herkunftsnachweise erfasst werden, um Authentizität und rechtliche Absicherung zu gewährleisten. Governance beginnt nicht mit der Speicherung von Informationen, sondern in dem Moment, in dem diese in das System gelangen, dann wenn die Vertrauensbasis geschaffen wird.

Die erfassten Datensätze werden anschließend über gestaffelte Speicherumgebungen verteilt, die jeweils ihrem Zweck und Risikoprofil entsprechen. Operative Systeme verwalten aktive Datensätze, Content-Management-Systeme sichern Versionierung, Klassifizierung und Aufbewahrung, während unveränderliche Archive rechtlich oder behördlich relevante Kommunikation und Nachweise aufbewahren. In Analyseumgebungen können regulierte Daten tokenisiert oder maskiert werden, um die Privatsphäre zu schützen und gleichzeitig fundierte Auswertungen zu ermöglichen. Langzeitarchive – ob auf Band, in Objektspeichern oder in souveränen Clouds – gewährleisten eine nicht löschbare Aufbewahrung, wenn dies gesetzlich erforderlich ist.

Eine moderne EIM-Plattform integriert Aufbewahrungsrichtlinien direkt in die Systeme, in denen die Inhalte gespeichert werden. Dadurch lässt sich sicherstellen, dass ein Dokument, das heute eine geschäftliche Entscheidung unterstützt, morgen archiviert oder automatisch gelöscht werden kann, sobald seine operative oder rechtliche Bedeutung entfällt.

Die gleichen Grundsätze gelten heute auch für künstliche Intelligenz. Da KI-Systeme zunehmend Unternehmensinhalte erzeugen, verarbeiten und aus ihnen lernen, müssen ihre Eingaben und Ausgaben mit derselben Sorgfalt behandelt werden wie regulierte Daten. Jede Interaktion eines Modells wird zu einem eigenen Datensatz, der erfasst, klassifiziert, aufbewahrt und nachvollziehbar gemacht werden kann. Governance sorgt dafür, dass KI-Systeme ausschließlich aus vertrauenswürdigen Quellen lernen, innerhalb klar definierter Grenzen agieren und Ergebnisse liefern, die erklärbar, revisionsfähig und im Einklang mit den Unternehmensrichtlinien stehen. Auf diese Weise werden die bewährten Prinzipien der Daten-Governance zu den Leitplanken einer verantwortungsvollen und vertrauenswürdigen künstlichen Intelligenz.

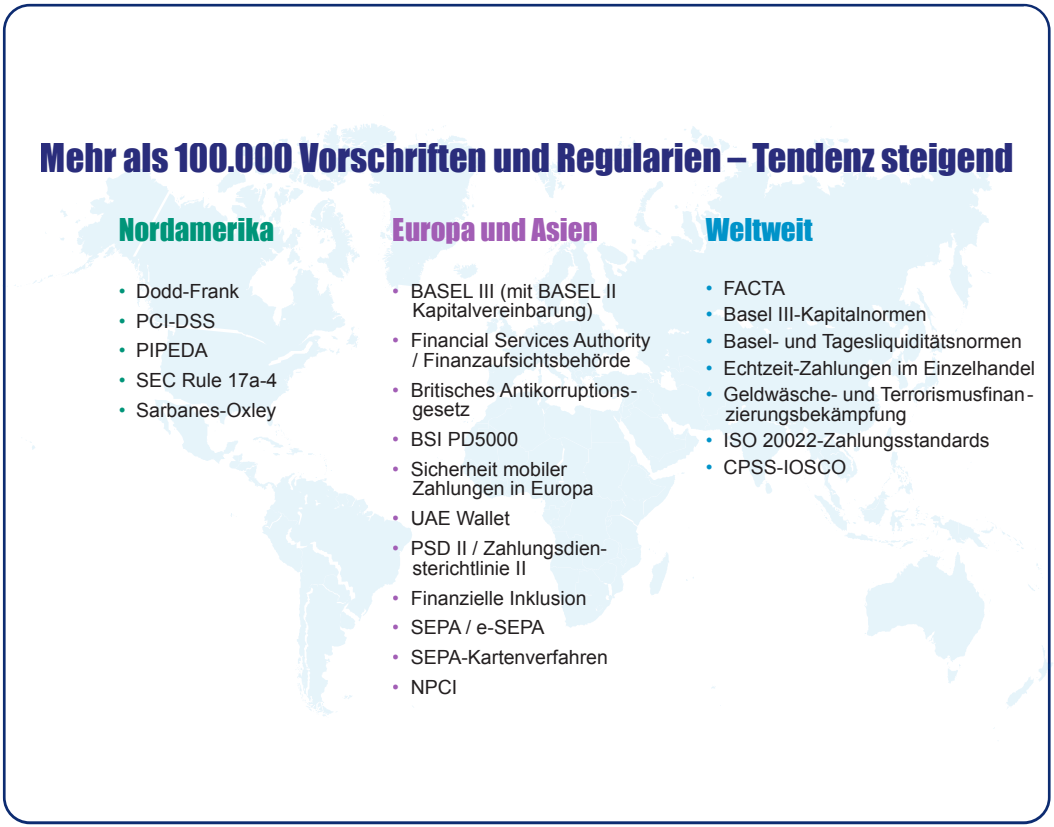
Säule 4: Revisionsfähigkeit – Der Beweis, dass Governance funktioniert

Im traditionellen Dokumentenmanagement bedeutete Revisionsfähigkeit vor allem Protokolle, Versionsverläufe und physische Nachweise. Im Zeitalter der künstlichen Intelligenz umfasst sie weit mehr: Modelltransparenz – also das Verständnis darüber, welche Informationen die Entscheidungen eines Systems geprägt haben.

Nachvollziehbarkeit sollte fest im Lebenszyklus der Inhalte verankert sein. Jedes Dokument, jede Transaktion und jedes Systemereignis besitzt eine lückenlose Historie aller Änderungen und Genehmigungen. Übertragen auf KI-Systeme bedeutet dieses Prinzip Erklärbarkeit: Es zeigt nicht nur, welche Entscheidung ein Modell getroffen hat, sondern auch, auf welchen Daten diese Entscheidung beruht. Diese Transparenz verwandelt Governance in Vertrauen. Sie gibt Aufsichtsbehörden, Führungskräften und der Öffentlichkeit die Sicherheit, dass Automatisierung innerhalb klar definierter Grenzen abläuft. Dadurch wird Compliance nicht länger als reaktiver Prozess verstanden, sondern als überprüfbarer Standard für verantwortungsvolles Handeln.

Alle diese Governance-Funktionen sind in einer Cloud-basierten EIM-Plattform gebündelt. Mit dem Fortschritt der künstlichen Intelligenz muss Governance stets an erster Stelle stehen, denn sie legt die Regeln für das Zusammenspiel zwischen Menschen, Daten und intelligenten Systemen fest. Metadaten schaffen die Orientierung, Berechtigungen definieren den Zugriff, das Lebenszyklusmanagement sorgt für Ausgewogenheit, und Revisionsfähigkeit liefert den Nachweis verantwortlicher Entscheidungsprozesse.

Ohne diese Grundlagen tappt die KI im Dunkeln. Mit ihnen wird sie Teil eines disziplinierten Informationsökosystems – eines Systems, das lernt, argumentiert und innerhalb klarer Grenzen handelt, sodass Informationen sicher, gesetzeskonform und menschlich verantwortbar bleiben.

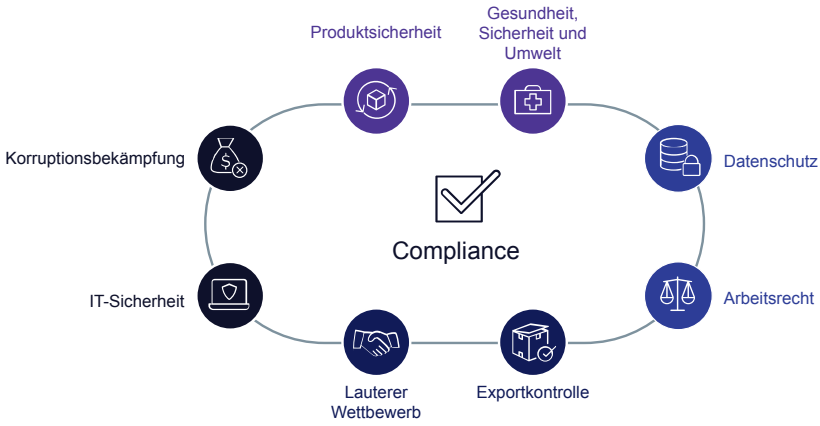


Globaler und regionaler Regulierungsdruck[40]

Eine komplexe Governance-Landschaft

Im heutigen globalen Markt ist das regulatorische Umfeld hochkomplex, insbesondere für international tätige Unternehmen. Organisationen müssen eine Vielzahl branchenspezifischer Vorschriften und Normen ebenso beachten wie regionale und nationale Bestimmungen. Gemäß diesen Bestimmungen tragen sie Verantwortung für ihr Handeln und müssen jederzeit auf umfangreiche historische Daten zugreifen können, um auf Informationsanfragen reagieren zu können.

Die Beziehung zwischen Compliance und Governance ist wechselseitig. Compliance dient als Treiber für Informationsgovernance, und Informationsgovernance wiederum kann Compliance vereinfachen. Angesichts wachsender Datenmengen besteht ein hoher Bedarf an Governance-Programmen, die Unternehmen bei der Transformation unterstützen, damit sie von einer besseren Verwaltung ihrer Informationen profitieren können. Unternehmen, die EIM als Governance-Plattform einsetzen, erkennen die Möglichkeiten, die sich ihnen dadurch bieten, die Geschäftstransformation durch optimierte Intelligenz und KI effizient und erfolgreich voranzutreiben.



Compliance ist vielschichtig

Compliance, Souveränität und die Gestaltung moderner Governance

Datensouveränität hat sich dabei von einer reinen Regelkonformität zu einem Gestaltungsprinzip entwickelt. Sie bestimmt, wo Daten gespeichert werden, wer darauf zugreifen darf und welcher Gerichtsbarkeit diese Vorgänge unterliegen. In einer von künstlicher Intelligenz geprägten Welt ist das von entscheidender Bedeutung, denn Modelle, die in einer Region trainiert oder gehostet werden, können dennoch Einflüssen aus anderen Rechtsräumen unterliegen.
Souveränität ist somit keine abstrakte Rechtsidee mehr, sondern eine architektonische Vorgabe. Bei jeder Speicherlösung, jeder API und jedem Trainingsdatensatz muss nun der jeweilige regulatorische Rahmen berücksichtigt werden.

Datenschutzgesetze und regionale Regelungen

Mit der Datenschutz-Grundverordnung (DSGVO) hat Europa weltweit Maßstäbe gesetzt. Sie hat den grenzüberschreitenden Datenverkehr von einer technischen Selbstverständlichkeit zu einer juristischen Herausforderung gemacht. Die DSGVO verankert die Grundsätze von Rechtmäßigkeit, Zweckbindung, Datenminimierung und Rechenschaftspflicht. Sie schreibt klare Einwilligungen, Folgenabschätzungen und den Schutz der Rechte betroffener Personen verbindlich vor. Die konsequente Durchsetzung dieser Verordnung hat die Systemarchitektur von Unternehmen grundlegend verändert. Dateninventare, metadatengesteuerte Workflows und automatisierte Löschrichtlinien gehören heute zu den zentralen Funktionen moderner Governance-Plattformen und sind längst keine optionalen Kontrollmechanismen mehr.

Das EU-Datenschutz-Grundgesetz (EU Data Act) erweitert diese Ideen auf die Cloud-Mobilität. Anbieter sind nun verpflichtet, Interoperabilität zu gewährleisten und Kunden ohne Abhängigkeiten oder überhöhte Transfergebühren den Wechsel zwischen Cloud-Umgebungen zu ermöglichen. Für Systemarchitekten bedeutet das, Portabilität von Beginn an einzuplanen: offene Standards, reversible Formate und Cloud-unabhängige Strukturen werden zur Voraussetzung. In Europa ist die Souveränität per Gesetz verankert.

Während Europa die Datensouveränität gesetzlich verankert, entsteht in den Vereinigten Staaten ein dezentraler Ansatz, bei dem einzelne Bundesstaaten schrittweise einen faktischen Bundesstandard formen. Inzwischen haben mehr als zwanzig Bundesstaaten eigene Datenschutzgesetze verabschiedet, die Begriffe wie Einwilligung, sensible Daten und Nutzerrechte jeweils unterschiedlich definieren. Dieser regulatorische Flickenteppich erfordert einen programmatischen Ansatz: Richtlinien müssen als Code umgesetzt werden, um sich automatisch an regionale Besonderheiten anzupassen und sicherzustellen, dass für jeden Datensatz, Benutzer und jede Transaktion das jeweils gültige Recht Anwendung findet. Der US CLOUD Act verkompliziert die Angelegenheit noch weiter, indem er den US-Behörden Zugang zu Daten gewährt, die sich im Besitz von in den USA ansässigen Anbietern befinden, selbst dann wenn die Daten im Ausland gespeichert sind.

Kanadas mehrstufiges Modell

Kanada verfolgt das Ziel der Datensouveränität durch ein mehrstufiges System der Verantwortlichkeit. Auf Bundesebene bildet der Personal Information Protection and Electronic Documents Act (PIPEDA) die Grundlage für den verantwortungsvollen Umgang mit personenbezogenen Daten. Er erlaubt grenzüberschreitende Datenübermittlungen, stellt jedoch sicher, dass Organisationen von der Erhebung bis zur Löschung für den Schutz der Daten verantwortlich bleiben. Der Nachfolger – der Digital Charter Implementation Act (Gesetzentwurf C-27) – führt drei zentrale Elemente ein: den Consumer Privacy Protection Act (CPPA), ein Gericht für den Datenschutz, das Data Protection Tribunal, sowie den Artificial Intelligence and Data Act (AIDA), der die verantwortungsvolle Entwicklung und Nutzung von KI regeln soll.

Auf Provinzebene werden die Vorschriften differenzierter umgesetzt. Die Reformen des Freedom of Information and Protection of Privacy Act (FOIPPA) in British Columbia haben die Anforderungen an die Datenresidenz für öffentliche Daten gelockert, während das Gesundheitsdatenschutzgesetz von Ontario, der Personal Health Information Protection Act (PHIPA), weiterhin strenge Standards für Gesundheitsdaten vorschreibt. Finanzaufsichtsbehörden wie das Office of the Superintendent of Financial Institutions (OSFI) haben mit der Richtlinie B-10 die Verantwortung für Datenstandorte und Cloud-Überwachung auf Vorstandsebene angehoben. Die zentrale Botschaft lautet: Souveränität in Kanada ist sowohl praktisch als auch auf Provinzebene verankert und erfordert eine genaue Kenntnis darüber, wo Daten gespeichert sind und wer Zugriff darauf hat.

Globale Compliance- und branchenspezifische Regeln

Über Nordamerika und Europa hinaus ist der Druck zur Datensouveränität global. Das chinesische Gesetz zum Schutz personenbezogener Daten (PIPL) schreibt umfassende Sicherheitsprüfungen für den grenzüberschreitenden Transfer „wichtiger Daten“ vor. Indiens Digital Personal Data Protection Act (DPDP) enthält Lokalisierungsvorgaben und neue Übertragungsbedingungen, die direkten Einfluss darauf haben, wo und wie KI-Workloads betrieben werden dürfen. Jede dieser Regelungen verschärft die Anforderungen an rechtmäßige Datenübermittlungen, ausdrückliche Einwilligungen und die Einhaltung der Aufbewahrungspflichten.

Branchenspezifische Vorschriften verstärken diese Anforderungen zusätzlich. Der Health Insurance Portability and Accountability Act (HIPAA) in den Vereinigten Staaten schafft einen verbindlichen Rahmen für elektronische Gesundheitsakten, der strenge Zugriffskontrollen, Verschlüsselung und Meldepflichten von Datenschutzverletzungen vorschreibt. Die Food and Drug Administration (FDA) legt mit 21 CFR Part 11 verbindliche Standards für vertrauenswürdige elektronische Aufzeichnungen und Signaturen in regulierten Produktions- und klinischen Umgebungen fest. Die Federal Trade Commission (FTC) überwacht die Einhaltung der Sorgfaltspflichten im Bereich der Verbraucherdatensicherheit, während die Federal Aviation Administration (FAA) die Anforderungen an Authentizität und Rückverfolgbarkeit digitaler Wartungsaufzeichnungen in der Luftfahrt festlegt. Gemeinsam bestätigen sie eine zentrale Wahrheit, die über alle Branchen hinweg gilt: Regulierte Daten müssen während ihres gesamten Lebenszyklus erfasst, aufbewahrt und revisionsfähig bleiben.

Souveränität und KII

Für künstliche Intelligenz markieren diese Gesetze operative Grenzen und architektonische Entscheidungen. Unternehmensweite KI kann nur aus den Informationen lernen, auf die sie rechtmäßig zugreifen darf. Souveräne Clouds, die Speicherung, Verarbeitung und Zugriff innerhalb nationaler Grenzen ermöglichen, sind die technologische Antwort auf die zunehmende regulatorische Fragmentierung. Sie erlauben es Organisationen, KI dort einzusetzen, wo die Daten liegen, wobei die Grenzen der Gerichtsbarkeit gewahrt bleiben und zugleich die Kontrolle über Compliance und Vertrauen gesichert wird.

Mit zunehmender Handlungsfähigkeit agentenbasierter Modelle bestimmt die Souveränität ihren Wirkungsbereich, welche Daten sie lesen, welche sie speichern und wie ihre Aktionen protokolliert und nachvollziehbar gemacht werden. Compliance wird damit zu einem dynamischen System, das unter rechtlicher Aufsicht denkt, lernt und handelt. Eine Enterprise-Information-Management-Plattform mit integrierter Governance bildet heute die operative Grundlage dieser neuen Form von Intelligenz.

Governance als Betriebssystem des Vertrauens

In einer Welt, in der Informationen Grenzen, Clouds und Algorithmen überschreiten, legt Governance die Spielregeln fest. Sie stellt sicher, dass Daten korrekt, nachvollziehbar und belastbar bleiben – unabhängig davon, wohin sie gelangen oder wie sie genutzt werden.

Wirksame Governance bedeutet mehr als bloße Dokumentation. Sie verlangt die Unterstützung durch die Unternehmensleitung, klar definierte Prozesse, automatisierte Richtliniendurchsetzung und identitätszentrierte Sicherheitsmechanismen. Jede Maßnahme im Umgang mit Daten – von der Erfassung bis zur Löschung – muss sichtbar, nachvollziehbar und mit rechtlichen wie ethischen Standards vereinbar sein.

Moderne Informationsmanagement-Frameworks verankern diese Kontrollmechanismen direkt im Arbeitsalltag. Metadaten, Berechtigungen und Aufbewahrungsfristen sind keine nachgelagerten Überlegungen, sondern eingebaute Logik, die dafür sorgt, dass Systeme verlässlich agieren und KI verantwortungsvoll bleibt. Wenn Organisationen Informationen als verwaltetes Gut mit klarer Herkunft, definiertem Zweck und festgelegtem Lebenszyklus behandeln, verwandeln sie Governance von einer Pflichtaufgabe in eine Quelle echter Wettbewerbsvorteile.

Letztlich schafft Governance die Grundlage dafür, dass Unternehmen Ihren Informationen und Analysen vertrauen können. Sie verbindet die ethische Verantwortung im Umgang mit Daten mit den Lernmechanismen künstlicher Intelligenz. Richtig umgesetzt, hemmt Governance Innovation nicht – sie macht sie belastbar und zukunftsfähig.

Die Governance von KI wird im folgenden Kapitel ausführlich behandelt.

Die fünf Merksätze

1. **Governance muss zum Unternehmensauftrag werden.**
   Ein funktionsübergreifender Governance-Gremium ist einzurichten, dem Führungskräfte aus den Bereichen Wirtschaft, IT, Recht und Compliance angehören. Dieses Gremium erhält die Befugnis, unternehmensweite Datenrichtlinien festzulegen, KI-Anwendungsfälle zu prüfen und die Einhaltung zu überwachen. Governance ist kein IT-Projekt, sondern eine Managementdisziplin.

2. **Berechtigungen und Zugriffskontrolle realisieren.**
   Zugriffsrechte sollten nicht statisch, sondern dynamisch und richtliniengesteuert sein. Es muss festgelegt werden, wer welche Informationen einsehen oder verwenden darf, und diese Kontrollmechanismen sind auf KI-Systeme auszuweiten. Jede KI-Interaktion ist mit Prüfprotokollen, Ablaufdaten und klarer Verantwortlichkeit als geregeltes Ereignis zu behandeln.

3. **Kritische Datenbestände kartieren und klassifizieren.**
   Unternehmensweite Dateninventuren müssen durchgeführt werden, um wertvolle, regulierte und sensible Daten zu identifizieren. EIM-Tools sind einzusetzen, um Inhalte mit Metadaten wie Verantwortlichkeit, Vertraulichkeit und Aufbewahrungsfristen zu versehen, damit sie für KI-Training, Automatisierung und Analysen sicher verwendet werden können.

4. **Konformität und Souveränität in die Architektur integrieren.**
   Von Beginn an ist die Komplexität der unterschiedlichen Rechtsordnungen zu berücksichtigen. Souveräne oder regionale Cloud-Konfigurationen sind zu wählen, bei denen der Datenstandort eine Rolle spielt. Die Einhaltung von Vorschriften ist durch Metadaten und Richtlinien als Code zu automatisieren, sodass Regeln über Speicherung und Übertragung von Daten im Design verankert sind und nicht erst durch Audits durchgesetzt werden.

5. **KI mit denselben Prinzipien steuern wie Daten.**
   Modelle sind wie verwaltete Vermögenswerte zu behandeln, mit dokumentierter Herkunft, Lebenszykluskontrolle und Governance für Nachschulungen. Es ist sicherzustellen, dass jede KI-Initiative vor ihrer Skalierung den rechtmäßigen Umgang mit Daten, nachvollziehbare Entscheidungen und einen messbaren ROI nachweist.

Kapitel Sechs

Die Governance von Unternehmens-KI (Enterprise AI, EAI)

Wie im vorangegangenen Kapitel erörtert, geht mit dem technologischen Fortschritt der Bedarf an effektiver Steuerung und Kontrolle einher. Während sich die Daten-Governance historisch und organisatorisch bereits weiterentwickelt hat, holt die Notwendigkeit einer umfassenden KI-Governance schnell auf.
KI-Governance bezeichnet die Richtlinien, Prozesse und Kontrollen, die sicherstellen, dass EAI-Technologien mit den Unternehmenszielen und regulatorischen Anforderungen übereinstimmen.

In diesem Kapitel wird untersucht, wie sich die Governance der Unternehmens-KI von einer technischen Fähigkeit in ein vertrauenswürdiges, strategisches Gut verwandelt.
Die Bedeutung der KI-Governance erstreckt sich auf den privaten wie auf den öffentlichen Sektor, mit Schwerpunkten auf Anwendungsbereich, Ethik, Compliance, Risikomanagement und Verantwortlichkeit.

Laut Gartner verfügen nur 12 Prozent der Unternehmen über ein implementiertes KI-Governance-Framework, während 55 Prozent angeben, noch keines eingeführt zu haben.[41]

Gartner-Umfrage unter IT- und Daten- und Analyseverantwortlichen zur KI-Strategie[42]

Da KI zunehmend in sämtliche Geschäftsbereiche integriert wird, kristallisiert sich Governance als ihre entscheidende Grundlage heraus. Eine wirksame KI-Governance weist Verantwortlichkeiten zu, definiert Aufsicht und stellt sicher, dass intelligente Systeme ethisch, sicher und transparent agieren. Ohne klare Steuerungsmechanismen riskieren Organisationen Voreingenommenheit, Datenschutzverletzungen und Reputationsverluste. Viele Unternehmen stehen weiterhin vor Herausforderungen, da es an Fachwissen, Koordination und konsistenten Daten fehlt, die für eine skalierbare Steuerung von KI erforderlich sind.

Künftige Wettbewerbsfähigkeit hängt davon ab, wie Governance gestaltet und Vertrauen aufgebaut wird. Zukunftsorientierte Organisationen bringen ihre Governance frühzeitig in Einklang mit neuen regulatorischen Rahmenbedingungen und integrieren verantwortungsvolle KI-Praktiken von Beginn an in ihre Abläufe.

Die folgende Fallstudie zeigt, wie ein globales Beratungsunternehmen, das zu den 12 Prozent gehört, die Gartner im Zusammenhang mit KI-Governance erwähnt hat, agiert.

Fallstudie

Ein globales Beratungsunternehmen

Das Unternehmen zählt zu den führenden Anbietern in den Bereichen Online- und Mobile-Strategie, Design, Entwicklung und Cybersicherheit und bündelt Fachwissen und Ressourcen eines weltweit renommierten Beratungsnetzwerks. Es erkennt den aktuellen digitalen Wendepunkt an, an dem künstliche Intelligenz, Automatisierung und Cloud-Technologien die Geschäftsmodelle, Belegschaftsstrukturen und Unternehmenskulturen grundlegend verändern. Ein leitender Technologieanalyst des Unternehmens beschreibt diese Entwicklung so:

„Parallel zur digitalen Transformation haben sich auch die Daten selbst weiterentwickelt. Es geht nicht mehr nur um Transaktionen, sondern um den Kontext. Der Wert liegt heute in den Beziehungen zwischen strukturierten und unstrukturierten Daten – in Gesprächen, Bildern und Signalen, die gemessenen Informationen Bedeutung verleihen. KI und Analytik ermöglichen es, diese Bedeutung in großem Umfang zu erschließen und unstrukturierte Informationen in umsetzbare Erkenntnisse zu verwandeln. Die Kombination von KI-gestützten Analysen auf einer konsolidierten Informationsplattform eröffnet ein exponentielles Potenzial und deckt Zusammenhänge und Risiken auf, die zuvor verborgen blieben.

Doch mit dieser Chance geht Verantwortung einher. Da KI zunehmend in unternehmerische Entscheidungen eingebunden ist, steigt die Bedeutung der Cybersicherheit weiter. Jedes intelligente System erfordert vertrauenswürdige Daten und eine geschützte Infrastruktur. Eine robuste Sicherheitsarchitektur, die Governance, Compliance und präventive Schutzmaßnahmen integriert, bildet die Grundlage für den Schutz von Unternehmensdaten. Wir verstehen, dass die Grundlagen unabhängig von der technologischen Entwicklung konstant bleiben: klare Rahmenbedingungen, durchgesetzte Richtlinien und wachsame Aufsicht, konsolidiert auf einer EIM-Plattform.

Da Organisationen verstärkt auf Hybrid- und Multi-Cloud-Modelle setzen, stellt sich nicht mehr nur die Frage, wo Daten gespeichert werden, sondern wie sie zu schützen sind. Künstliche Intelligenz verstärkt sowohl die Möglichkeiten als auch die Risiken der digitalen Transformation und macht Cybersicherheit damit in einem intelligenten Unternehmen zu einer tragenden Säule des Vertrauens.“

Was ist der Anwendungsbereich der EAI-Governance?

KI-Governance definiert, wie ein Unternehmen KI im Einklang mit strategischen Zielen und regulatorischen Verpflichtungen entwickelt, einsetzt und verwaltet. Als Erweiterung der Unternehmens- und IT-Governance adressiert EAI-Governance zentrale Herausforderungen wie Modellüberwachung, BIAS-Vermeidung, Datenmanagement, Cybersicherheitskontrolle und die Einhaltung gesetzlicher Vorgaben. Zunehmend legen Organisationen Richtlinien für den verantwortungsvollen Einsatz von KI fest, die Prinzipien wie Fairness, Transparenz und Rechenschaftspflicht über den gesamten KI-Lebenszyklus hinweg verankern. Diese Leitplanken sind essenziell, um eine vertrauenswürdige und verantwortungsvolle Implementierung sicherzustellen und das Vertrauen der Öffentlichkeit in moderne intelligente Systeme zu wahren.

Zu den maßgeblichen Rahmenwerken zählen die KI-Prinzipien der Organisation für wirtschaftliche Zusammenarbeit und Entwicklung (OECD), der von der Europäischen Union vorgeschlagene KI-Gesetzentwurf, der AI Risk Management Framework (AI RMF) des National Institute of Standards and Technology (NIST) sowie die Normen der Internationalen Organisation für Normung (ISO) im Bereich KI.

Die EAI-Governance erstreckt sich zunehmend über die unmittelbaren Systeme eines Unternehmens hinaus auf die zugrunde liegenden Modelle und Infrastrukturen. Neue regulatorische Rahmenbedingungen – darunter der EU AI Act (2024) und die US Executive Order 14110 on AI Safety (2023) – unterscheiden zwischen Governance auf Systemebene (dem internen Umgang mit KI) und Governance auf Modellebene (den Pflichten jener, die allgemeine oder Frontier-KI-Modelle entwickeln oder anpassen). Unternehmen müssen künftig vor der Integration externer Modelllieferanten eine sorgfältige Revision durchführen (NIST 2023; Europäische Kommission 2024). Dadurch entsteht ein höheres Maß an Verantwortlichkeit in der gesamten KI-Lieferkette.

> *Bei Datenethik geht es um mehr als nur um die Einhaltung von Vorschriften: Es geht darum, auch dann das Richtige zu tun, wenn das Gesetz es nicht vorschreibt.*[43]

Eine ethische und verantwortungsvolle KI gewährleisten

Ethische Governance stellt sicher, dass KI-Systeme fair, transparent und im Einklang mit den Menschenrechten gestaltet sind. Auch wenn der Begriff „ethische KI“ modern wirkt, reichen seine Wurzeln mehrere Jahrzehnte zurück – bis zu den grundlegenden Diskussionen über Computerethik. In seinem bahnbrechenden Aufsatz „Was ist Computerethik?“ aus dem Jahr 1985 heißt es:

> „Ein typisches Problem in der Computerethik besteht darin, dass es ein politisches Vakuum gibt, wie Computertechnologie eingesetzt werden sollte. Computer bieten uns neue Möglichkeiten, und diese wiederum eröffnen uns neue Handlungsoptionen. Oftmals existieren für solche Situationen keine angemessenen Verhaltensrichtlinien, oder bestehende Richtlinien erweisen sich als unzureichend. Eine zentrale Aufgabe der Computerethik besteht darin, festzulegen, was in solchen Fällen zu tun ist – das heißt, Richtlinien zu formulieren, die unser Handeln leiten. Natürlich gibt es ethische Situationen,

> mit denen wir als Einzelne und als Gesellschaft konfrontiert werden. Die Ethik der Computerwissenschaft umfasst dabei sowohl persönliche als auch gesellschaftliche Richtlinien für den verantwortungsvollen Einsatz von Computertechnologie."[44]

Spätere Schriften befassten sich speziell mit der KI-Ethik, darunter die „Principles on Artificial Intelligence (Grundsätze zur künstlichen Intelligenz) (2019) der OECD und in jüngerer Zeit die „Recommendation on the Ethics of Artificial Intelligence" (Empfehlung zur Ethik der künstlichen Intelligenz), die von der Organisation der Vereinten Nationen für Erziehung, Wissenschaft und Kultur (UNESCO) im Jahr 2022 veröffentlicht wurde. Angesichts der 194 Mitgliedstaaten der UN stellt dieser Empfehlungskatalog den bislang umfassendsten ethischen Rahmen dar.

Die Empfehlungen unterstreichen die Notwendigkeit verbindlicher ethischer Leitlinien:

> „KI-Systeme werfen neue ethische Fragen auf, die sich auf Entscheidungsfindung, Beschäftigung, soziale Interaktion, Gesundheitswesen, Bildung, Medien, Datenzugang und digitale Technologien beziehen. Sie betreffen zudem Themen wie Datenschutz, Verbraucherschutz, Umwelt, Demokratie, Rechtsstaatlichkeit, Sicherheit, Dual-Use-Güter sowie Menschenrechte und Grundfreiheiten, einschließlich Meinungsfreiheit, Privatsphäre und Nichtdiskriminierung.
> Darüber hinaus entstehen neue Herausforderungen durch das Potenzial von KI-Algorithmen, bestehende Vorurteile zu reproduzieren und zu verstärken, wodurch Diskriminierung und Stereotypisierung weiter gefestigt werden können."[45]

In den Empfehlungen heißt es weiter, dass die Auswirkungen der KI auf die Menschheit zunehmen werden, wenn sie immer mehr Aufgaben übernimmt, die bisher von Menschen ausgeführt wurden. Sie hat das Potenzial, die Art und Weise, wie wir die Welt um uns herum verstehen, und auch unser Selbstverständnis tiefgreifend zu verändern.[46]

Zu den Empfehlungen für ethische KI gehören (unter anderem) die folgenden Grundsätze:

- **Verhältnismäßigkeit und nicht schaden**: Eine angemessene Nutzung und Minimierung potenzieller Schäden im jeweiligen Anwendungskontext sicherstellen
- **Sicherheit und Schutz**: Schäden, einschließlich Sicherheitsrisiken vermeiden
- **Fairness und Nichtdiskriminierung**: Einschließlich soziale Gerechtigkeit und Wahrung der Gleichbehandlung fördern
- **Nachhaltigkeit**: Einschließlich menschliche, soziale, kulturelle, wirtschaftliche und ökologische Auswirkungen berücksichtigen
- **Recht auf Privatsphäre und Datenschutz**: Regelung der Datennutzung für KI
- **Menschliche Aufsicht und Entscheidungsfindung**: Die menschliche Kontrolle über KI aufrechterhalten
- **Transparenz und Erklärbarkeit**: Einschließlich das Verständnis und die Nachvollziehbarkeit von KI-Anwendungen fördern
- **Verantwortung und Rechenschaftspflicht**: Menschenrechte und Grundfreiheiten gewährleisten

- **Bewusstsein und Lesefähigkeit**: Kompetenzen in Bildung, Ausbildung und Medien stärken, um den bewussten Umgang mit KI zu fördern
- **Multi-Stakeholder- und adaptive Governance und Zusammenarbeit**: Das Völkerrecht und die nationale Souveränität achten[47]

Da eine ethische KI-Governance den Grundsatz „Do no harm" institutionalisiert, ist die Verankerung von Werten wie Nichtdiskriminierung, Verantwortlichkeit und Transparenz für Unternehmen von zentraler Bedeutung und für Organisationen des öffentlichen Sektors unverzichtbar. Die Verlagerung des Diskurses von einer bloßen Checkliste hin zu einem durchdachten System bedeutet, Ethik als Teil der Infrastruktur zu begreifen. Ethische Leitplanken müssen in jede Phase des KI-Lebenszyklus eingebettet werden. Auf diese Weise wird Ethik zu einem Bestandteil der operativen Architektur, der definiert, wofür eine Organisation steht und wie sie handelt – nicht nur, was sie produziert. Da verantwortungsvolle Innovation zunehmend zur Grundlage unternehmerischen Handelns wird, sind ethische und verantwortungsvolle KI für nachhaltiges Wachstum und Vertrauen entscheidend.

Umfassender Rahmen für die KI-Governance

Risikomanagement und Vertrauenssicherung

Künstliche Intelligenz birgt Risiken, die von herkömmlicher IT-Governance nicht vollständig abgedeckt werden. Eine Stärkung von Rechenschaftspflicht und Aufsicht ist daher unerlässlich. Für Zuverlässigkeit sind strenge Tests und belastbare Fallback-Mechanismen erforderlich, da KI-Systeme auf unvorhersehbare Weise versagen können. Kontinuierliche Überwachung trägt dazu bei, Risiken zu minimieren und die Qualität langfristig zu sichern. Vorrangig ist die Gewährleistung von Sicherheit, Datenschutz und Schutz im allgemeinen, da Vorfälle das Vertrauen der Öffentlichkeit und die Reputation eines Unternehmens erheblich beeinträchtigen können. Die Integration des KI-Risikomanagements in umfassendere Enterprise-Risk-Management-(ERM)-Prozesse stellt sicher, dass KI-Risiken mit derselben Strenge behandelt werden wie finanzielle oder operative Risiken.

Folgende bewährte Verfahren bilden die Grundlage eines stabilen Governance-Rahmens:

Integration mit der Governance

EAI-Governance überschneidet sich mit bestehenden Governance-Strukturen. IT-Governance und Risikomanagement erfordern klar definierte Rollen, Verantwortlichkeiten und Aufsichtsmechanismen auf allen Ebenen. Vertrauenswürdige KI-Governance wird durch aufgabenbasiertes Richtlinienmanagement als Erweiterung der bestehenden rollenbasierten Zugriffskontrolle (Role Based Access Controls, RBAC) erreicht – für Menschen ebenso wie für Agenten. Der Aufbau von Vertrauen erfordert funktionsübergreifende Teams, die Fachleute aus Technologie, Ethik, Recht und teils auch Kundenperspektiven einbeziehen. Da KI als vergleichsweise junge Technologie gilt, sollte ihr Risikoprofil als hoch eingeschätzt werden, solange das volle Risikospektrum nicht quantifiziert ist.

Richtlinien und Standards

Klare Richtlinien, Verhaltenskodizes und interne Standards schaffen Transparenz über Erwartungen und ermöglichen deren Durchsetzung. Diese müssen regelmäßig überprüft und in die Geschäftsregeln der KI-Agenten integriert werden, damit autonome Entscheidungen dem festgelegten ethischen Rahmen folgen. Besonders wichtig ist der geschützte Umgang mit Unternehmensdaten und die klare Regelung ihrer Verwendung in Verbindung mit öffentlichen KI-Modellen.

Prüfbarkeit

Um eine gute Governance zu gewährleisten, sind umfassende Dokumentation, Protokollierung und Audit-Trails aller KI-Systeme erforderlich. Sie unterstützen sowohl interne Prüfungen als auch externe Kontrollen Eine proaktive Festlegung dieser Grundsätze ermöglicht es, potenzielle Probleme frühzeitig zu erkennen und zu beheben. Diese Aspekte der Daten-Governance wurde im vorangegangenen Kapitel beschrieben.

Transparenz

Transparenz bildet die Grundlage einer vertrauenswürdigen KI-Governance. Sie stellt sicher, dass Entscheidungen intelligenter Systeme nachvollziehbar und revisionsfähig bleiben.
Das Prinzip der Transparenz, Erklärbarkeit und Anfechtbarkeit bietet hierfür einen strukturierten Ansatz. Organisationen sollten regelmäßige Überprüfungen durchführen, um zu bewerten, wie verständlich KI-Systeme ihre Entscheidungslogik darstellen und wie fair sie operieren. Durch offene Dokumentation der Entscheidungswege, transparente Datennutzung und die Möglichkeit, Ergebnisse zu hinterfragen, wird KI von einer Blackbox zu einem nachvollziehbaren, menschenzentrierten System weiterentwickelt, in dem Fairness und Vertrauen in jede Entscheidung eingebettet sind.

Entwicklung und Betrieb

Ein Privacy-by-Design-Ansatz in Entwicklung und Betrieb gewährleistet, dass der Schutz personenbezogener Daten von der Konzeption über die Bereitstellung bis zur Nachbereitung und darüber hinaus über den gesamten Systemlebenszyklus hinweg verankert ist. Anstatt den Datenschutz als Compliance-Anforderung oder nachträgliche Überlegung zu behandeln, wird er zu einem architektonischen Prinzip, das die Art und Weise der Datenerfassung, -verarbeitung und -speicherung bestimmt. Dies bedeutet, die Datennutzung auf das unbedingt Notwendige zu beschränken, Anonymisierung und Verschlüsselung standardmäßig anzuwenden und Mechanismen zur Nutzereinwilligung und -kontrolle direkt in die Arbeitsabläufe einzubetten. Kontinuierliche Überwachung und Datenschutzfolgenabschätzungen gewährleisten die

Verantwortlichkeit der Systeme im Zuge ihrer Weiterentwicklung. Durch Ausrichtung von Entwicklung und Betrieb an den Prinzipien von Privacy-by-Design reduzieren Unternehmen nicht nur das regulatorische Risiko, sondern schaffen auch Vertrauen, Widerstandsfähigkeit und einen Wettbewerbsvorteil, der auf ethischer Innovation beruht.

Reaktion auf Vorfälle

In Kapitel 11 wird die Notwendigkeit eines anderen Ansatzes für den Betrieb behandelt. Im Kontext von KI-Governance und Risikomanagement ist dabei hervorzuheben, dass Protokolle für die schnelle Reaktion auf KI-Fehler – einschließlich menschlicher Beschwerden und Abhilfemaßnahmen – für die Wahrung der Verantwortlichkeit unerlässlich sind. Kontinuierliche Verbesserung bildet ein zentrales Element jeder Governance-Struktur.

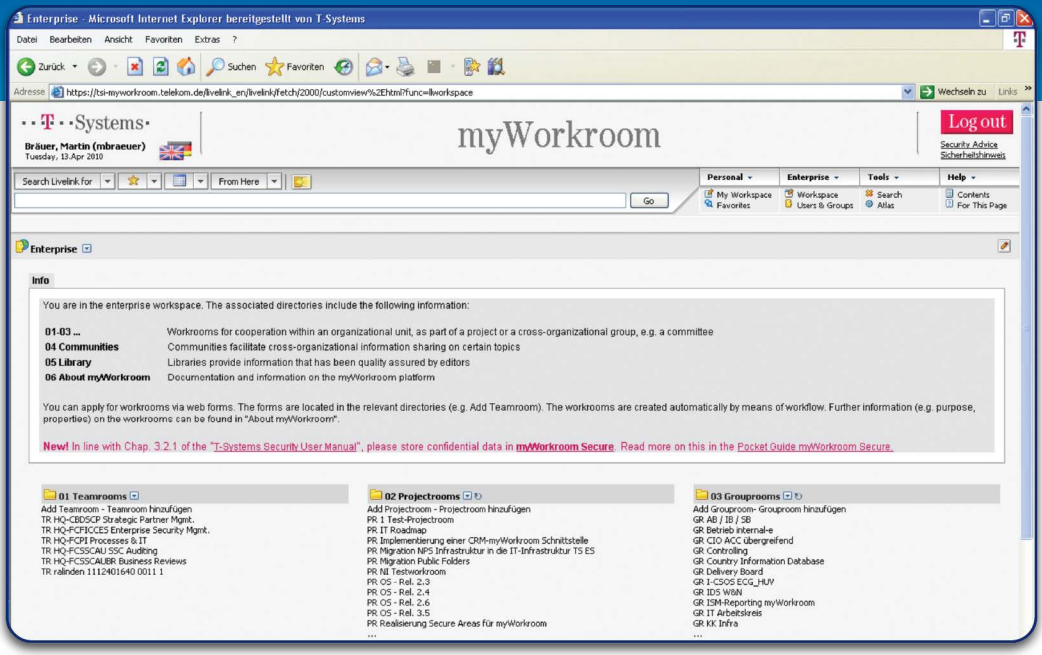
Alle diese Komponenten sind entscheidend für das Risikomanagement und die Aufrechterhaltung von Vertrauen. Da viele Organisationen KI noch nicht vollständig operationalisiert haben, ist es ratsam zu prüfen, wie sich diese Elemente auf die Gesamtstrategie auswirken.

Die heutige KI-Governance erfordert zudem spezifische Sicherheits- und Resilienz-Maßnahmen, die auf generative KI zugeschnitten sind. Herkömmliche IT-Sicherheitsrahmen decken häufig keine Bedrohungen ab, die mit generativen Systemen verbunden sind – etwa Prompt Injection, Halluzination oder Ausgabemanipulation. Das Generative AI Profile des NIST (Generative AI Profile) (NIST) (National Institute of Standards and Technology) und die Secure by Design-Leitlinien der CISA (Secure by Design Guidelines) (CISA) (Cybersecurity and Infrastructure Security Agency) empfehlen daher KI-spezifische Bedrohungsmodellierung, Adversarial Red Teaming, Überwachung auf Datenexfiltration und Herkunftsnachweise von Modellausgaben (CISA 2024; NIST 2024). Organisationen sollten entsprechende Maßnahmen implementieren, um Sicherheit, Zuverlässigkeit und Verantwortlichkeit über den gesamten Lebenszyklus generativer KI sicherzustellen.

Im Anschluss wird beschrieben, wie ein europäisches Telekommunikationsunternehmen den Lebenszyklus seiner Informationen verwaltet, um die Anforderungen an Compliance und Daten-Governance zu erfüllen. Dabei gelten für Aufbewahrungsdauer und Löschfristen strenge Regeln, die gewährleisten, dass Informationen nur so lange gespeichert werden, wie es rechtlich und betrieblich erforderlich ist.

Fallstudie

T-Systems



Unternehmensweites Informationsmanagementsystem

Mit Niederlassungen in mehr als 20 Ländern ist T-Systems – die meistgenutzte Kundenmarke der Deutschen Telekom – für zahlreiche große europäische Unternehmen der bevorzugte Partner bei der Abwicklung globaler Geschäftsprozesse. Rund 160.000 Unternehmen und öffentliche Einrichtungen nutzen die integrierten Dienstleistungen von T-Systems, die von der Verwaltung von Rechenzentren und globalen Internetprotokolldiensten bis zur Entwicklung und Administration von Anwendungen reichen.

Um die Zusammenarbeit über Abteilungen, Standorte und Projekte hinweg zu optimieren, benötigten die Teams von T-Systems eine Plattform, die schnelle Kommunikation, effizienten Informationsaustausch und eine reibungslose Durchführung von Kundenprojekten ermöglicht. Heute greifen rund 40.000 Beschäftigte auf eine unternehmensweite Enterprise Content Management (ECM)-Plattform zu, die Funktionen für Zusammenarbeit, Dokumentenmanagement und Wissensmanagement bündelt.

T-Systems erweitert diese Kollaborationsplattform kontinuierlich – unter anderem um ein Extranet-Gateway, das die Zusammenarbeit mit Kunden und Partnern erleichtert, sowie um ein Lebenszyklusmanagementsystem für Projekträume mit Speicherdauern von bis zu zehn Jahren. Diese Erweiterung unterstützt die Einhaltung von Compliance-Anforderungen im Rahmen der Corporate Governance und stellt sicher, dass wertvolle, aber inaktive Projektinformationen langfristig auffindbar und nutzbar bleiben.

Führende Rahmenwerke, Vorschriften und Standards

Die regulatorischen Rahmenbedingungen für KI entwickeln sich rasant weiter, und die Einhaltung der Vorschriften ist inzwischen ein zentraler Bestandteil moderner Governance. Dabei besteht ein Zusammenspiel aus freiwilligen Rahmenwerken – wie dem Risk Management Framework (RMF) des National Institute of Standards and Technology (NIST) – und verpflichtenden Regelwerken, darunter der EU AI Act, der in Kapitel 5 behandelt wird. Auch freiwillige Standards werden für viele Organisationen zunehmend zur Voraussetzung, um einen strukturierten Ansatz für vertrauenswürdige KI umzusetzen. Governance-Rahmenwerke dienen dazu, abstrakte Anforderungen in konkrete Kontroll-, Richtlinien- und Aufsichtspflichten zu übersetzen, die Unternehmen bei der Implementierung unterstützen.

Obwohl sich die regulatorischen Vorgaben laufend weiterentwickeln, stehen heute verschiedene etablierte Rahmenwerke zur Verfügung, die eine systematische Strukturierung der KI-Governance ermöglichen.

OECD-KI-Prinzipien: Die von 46 Ländern im Jahr 2019 übernommenen OECD-KI-Prinzipien bildeten einen der ersten internationalen Standards für die Steuerung von KI. Sie definieren fünf grundlegende Leitprinzipien für verantwortungsvolle KI-Governance:[48]

1. Künstliche Intelligenz sollte den Menschen und dem Planeten zugutekommen, indem sie inklusives Wachstum und Wohlbefinden fördert.
2. KI-Systeme müssen so konzipiert sein, dass sie Menschenrechte, demokratische Werte und Vielfalt respektieren.
3. KI-Systeme sollten transparent und nachvollziehbar sein.
4. KI-Systeme müssen während ihres gesamten Lebenszyklus robust, sicher und zuverlässig sein.
5. Organisationen und Einzelpersonen, die KI entwickeln, einsetzen oder betreiben, sind für die Auswirkungen und Ergebnisse ihrer Systeme verantwortlich.

EU-KI-Gesetz: Die Verordnung des Europäischen Parlaments und des Rates zur Festlegung harmonisierter Vorschriften für künstliche Intelligenz (KI-Gesetz) wurde im Jahr 2024 verabschiedet. Sie stellt einen der ersten umfassenden Rechtsrahmen für KI dar und legt Anforderungen auf der Grundlage unterschiedlicher Risikostufen fest – von minimal bis inakzeptabel. Die wichtigsten Anforderungen des Gesetzes lauten wie folgt:[49]

- Transparenz bei KI-generierten Inhalten und biometrischen Systemen
- Strenge Prüf-, Dokumentations- und Compliance-Pflichten für KI-Systeme mit hohem Risiko (etwa im Gesundheitswesen, in kritischen Infrastrukturen oder der öffentlichen Verwaltung)
- Verbot des Einsatzes von KI, die Verhalten manipuliert oder Sicherheitslücken ausnutzt
- Verpflichtende menschliche Aufsicht, umfassende Risikomanagementsysteme und Angleichung an die EU-Vorschriften im Bereich digitale und datenbezogene Governance

NIST AI RMF: Das 2023 veröffentlichte NIST Risk Management Framework (AI RMF) bietet einen freiwilligen, international anerkannten Rahmen zur Identifizierung, Bewertung und Steuerung von Risiken in KI-Systemen. Es umfasst die Kontextanalyse und Zieldefinition, die Messung von Risiken, das Management mithilfe definierter Kontrollmechanismen sowie die Steuerung von KI-Systemen über ihren gesamten Lebenszyklus hinweg. Das Rahmenwerk wurde so gestaltet, dass es mit bestehenden Sicherheitsrichtlinien, insbesondere für Zero-Trust-Architekturen, kompatibel ist.

Es erkennt an, dass sich die KI-Technologie kontinuierlich weiterentwickelt: „Das KI-Risikomanagementmodell (AI RMF) soll praxisorientiert sein, sich an die sich weiterentwickelnde KI-Landschaft anpassen und von Organisationen in unterschiedlichem Umfang und mit verschiedenen Kapazitäten umgesetzt werden können, damit die Gesellschaft von KI profitieren und gleichzeitig vor ihren potenziellen Gefahren geschützt werden kann."[50]

ISO/IEC 42001:2023: Die Internationale Organisation für Normung (ISO) und die Internationale Elektrotechnische Kommission (IEC) haben ergänzende Standards für das Management und die Steuerung von KI entwickelt. Dazu zählen Normen für KI-Managementsysteme (AIMS) (ISO/IEC 42001:2023), KI-Konzepte und -Terminologie (ISO/IEC 22989:2022) sowie Lebenszyklusprozesse von KI-Systemen (ISO/IEC 23053:2022).

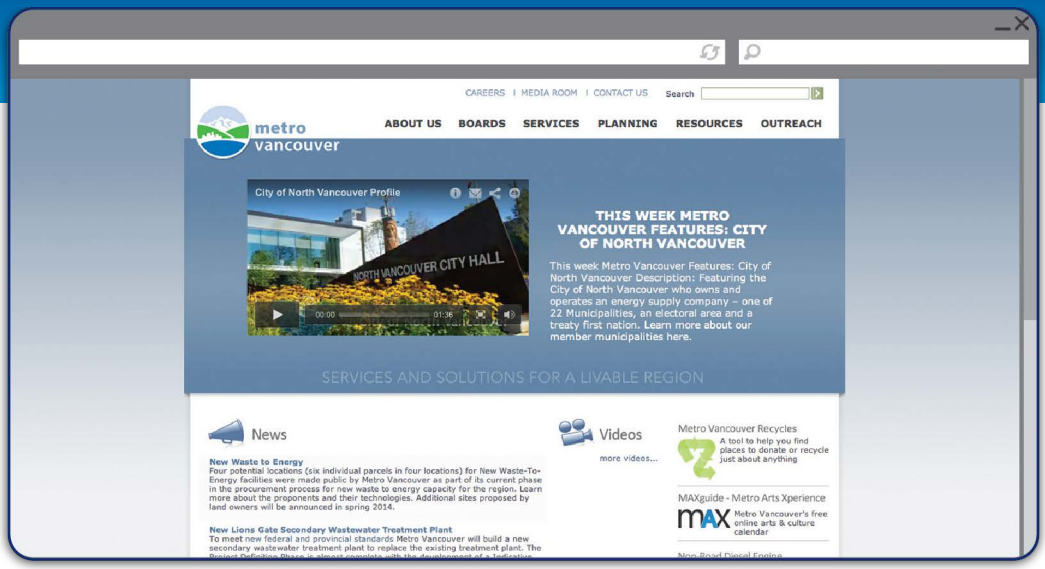
Gemeinsam bieten diese Normen eine strukturierte Grundlage für verantwortungsvolle Innovation. Sie übersetzen zentrale Prinzipien wie Fairness, Transparenz und Rechenschaftspflicht in praxisnahe Governance-Mechanismen, die mit globalen Erwartungen und rechtlichen Standards übereinstimmen. Durch die Verankerung ihrer KI-Programme in diesen Rahmenwerken können Organisationen Systeme entwickeln, die nicht nur konform, sondern auch konsistent, erklärbar und grenzüberschreitend vertrauenswürdig sind.

Es gibt noch weitere Normen und Rahmenwerke, aber diese sind die am weitesten verbreiteten und angenommenen. Mit der Zeit und der Weiterentwicklung der Technologie werden neue Rahmenbedingungen entstehen. Daher ist ein anpassungsfähiger Ansatz bei der Definition von Kontrollen der Schlüssel zur Vermeidung zukünftiger Nacharbeiten.

Im folgenden Beispiel beweist Metro Vancouver, dass gute Regierungsführung auch wirtschaftlich sinnvoll ist. Ein EIM-Backbone trägt dazu bei, dass die Region durch Audits nachweisen kann, dass die Dokumente im System vertrauenswürdige Aufzeichnungen sind und dass sie den Gesetzen und Vorschriften entsprechen und gleichzeitig eine gute Geschäftspraxis fördern.

Fallstudie

Metro Vancouver



Metro Vancouver

Metro Vancouver ist einer von 29 Regionalbezirken, die von der Provinzregierung eingerichtet wurden, um sicherzustellen, dass alle Einwohner von British Columbia gleichen Zugang zu zentralen öffentlichen Dienstleistungen haben. Regionale Parks, bezahlbarer Wohnraum, Arbeitsbeziehungen und regionale Stadtplanung sind wichtige Dienstleistungen, die direkt für die Öffentlichkeit erbracht werden. Die Region beschäftigt mehrere Tausend Vollzeitkräfte und versorgt eine Bevölkerung von über drei Millionen Menschen.

Zur Verwaltung ihrer umfangreichen Administrations- und Archivdaten benötigte die Region ein zentrales, sicheres System zur Speicherung und Verteilung elektronischer Datensätze. Eine E-Government-Lösung sollte es ermöglichen, Aufbewahrungsfristen und Vernichtungsregeln auf der Grundlage vordefinierter Zeiträume automatisiert durchzusetzen. Ziel war es, Risiken zu minimieren, Lagerkosten zu senken und die Einhaltung gesetzlicher Vorgaben sicherzustellen. Darüber hinaus sollte das System die Benutzerfreundlichkeit verbessern – insbesondere bei der Profilerstellung von Dokumenten – und gleichzeitig Automatisierung und höhere Genauigkeit gewährleisten.

Aktuell umfasst das System nahezu zwei Millionen Dokumente. Die automatisierte Lösung vereinfacht die elektronische Dokumentenverwaltung und macht den Prozess für die Nutzer transparent. Sie ordnet Datensatzklassifizierungen automatisch den entsprechenden Aufbewahrungsfristen zu und gewährleistet, dass Datensätze so lange aufbewahrt werden, wie es gesetzlich erforderlich ist, bevor sie nach Ablauf der Frist automatisch gelöscht werden.
Die Einhaltung der Governance-Richtlinien wird in der gesamten Region sichergestellt. Jede der 14 Abteilungen ist für die Befolgung der vom Corporate Records Team festgelegten Richtlinien,

Verfahren und Best Practices verantwortlich. Das System stellt sicher, dass Metro Vancouver durch Audits jederzeit nachweisen kann, dass die gespeicherten Dokumente vertrauenswürdige Aufzeichnungen sind, die den geltenden Gesetzen und Vorschriften entsprechen und zugleich eine effiziente und verantwortungsbewusste Verwaltungspraxis fördern.

Der Weg für die EAI-Governance nach vorne

Die Steuerung von KI ist sowohl im öffentlichen als auch im privaten Sektor unerlässlich. Zwar ähneln sich Umsetzungszeiträume und Umfang, doch unterscheiden sich die Prioritäten deutlich. Im öffentlichen Sektor liegt der Schwerpunkt auf Transparenz und Bürgervertrauen. Gute Governance basiert hier auf öffentlicher Rechenschaftspflicht, Ethik und der Einhaltung der Menschenrechte. Im privaten Sektor hingegen steht die Förderung von Innovation, das Management von Geschäftsrisiken und die Einhaltung gesetzlicher Vorschriften im Vordergrund. Die Governance ist eng mit der Corporate Social Responsibility (Soziale Verantwortung des Unternehmens) und der ESG-Agenda (Umwelt, Soziales und Governance) verknüpft, wobei Agilität und Wirkung in Einklang gebracht werden müssen. Beide Sektoren profitieren von einer Angleichung an internationale Rahmenwerke und Standards und sollten die KI-Governance als dynamisches, sich fortlaufend weiterentwickelndes Programm begreifen.

Insgesamt gewährleistet die EAI-Governance, dass ethische Grundsätze, regulatorische Konformität, Risikomanagement und Verantwortlichkeit über den gesamten KI-Lebenszyklus hinweg und über Organisationsgrenzen hinaus fest verankert sind. Eine erfolgreiche KI-Governance verbindet übergeordnete Prinzipien mit konkreten Prozessen und Werkzeugen und wird durch eine Kultur der Verantwortung auf allen Ebenen getragen. Da sich sowohl die Technologie als auch die Regulierung rasant weiterentwickeln, werden kontinuierliche Investitionen in Governance entscheidend sein – nicht nur zur Risikominderung, sondern auch, um Vertrauen zu schaffen, nachhaltige Innovation zu fördern und langfristige Wettbewerbsvorteile zu sichern.

Die Steuerung von KI geht über technische Erwägungen hinaus und bezieht ethische sowie gesellschaftliche Dimensionen mit ein. Immer mehr Organisationen führen formale Richtlinien für verantwortungsvolle KI ein und verankern zentrale Prinzipien wie Fairness, Transparenz, Rechenschaftspflicht und Achtung der Menschenrechte im gesamten KI-Lebenszyklus. Diese Leitplanken sind nicht nur für die Einhaltung gesetzlicher Vorschriften unerlässlich, sondern auch für die Wahrung des öffentlichen Vertrauens in KI-gestützte Prozesse. Ohne sie droht das Potenzial intelligenter Systeme durch ethische Fehlentwicklungen, Datenschutzverletzungen oder mangelhafte Governance untergraben zu werden.

Die nächste Entwicklungsstufe der KI-Governance betrifft die Ausrichtung und Kontrolle autonomer und agentenbasierter Systeme, die eigenständig Entscheidungen treffen oder Aktionen auslösen können. Die Anforderungen an Governance werden sich dadurch erweitern und künftig auch Autonomiegrenzen, Echtzeitüberwachung und Eskalationsverfahren umfassen, insbesondere wenn Modelle täuschendes oder zielorientiertes Verhalten zeigen (UK AI Safety Institute, 2024). Zudem werden Rechen- und Leistungsschwellenwerte zunehmend als politische Steuerungsinstrumente genutzt, um festzulegen, wann die Entwicklung oder der Betrieb von KI-Systemen eine externe Überprüfung auslösen sollte (NIST, 2024; CISA, 2024). Für Unternehmen bedeutet dies eine Weiterentwicklung der EAI-Governance: weg von statischer Richtlinientreue hin zu einem System kontinuierlicher Überwachung, Qualitätssicherung und adaptivem Risikomanagement. Organisationen, die Governance als lebendiges, mehrstufiges System institutionalisieren, das aus Kontrollen, Aufsicht und externer Validierung besteht, werden am besten aufgestellt sein, um Innovationen im Grenzbereich verantwortungsvoll voranzutreiben.

Die folgende Fallstudie zeigt, wie ein führender Anbieter von Unternehmenssoftware die Anzahl der Dokumenttypen um 96 Prozent reduziert hat, um sich auf KI-Innovationen und Automatisierung vorzubereiten.

Fallstudie

Ein globaler ERP-Anbieter

> *Die Möglichkeiten der Mitarbeiter-Selbstbedienung sind unbegrenzt. Wenn ein Mitarbeiter beispielsweise ein Dokument einreicht, um seine Adresse oder seinen Familienstand ändern zu lassen, kann eine KI-gestützte Automatisierung die Personalakte eigenständig ganz ohne Eingriff durch die Personalabteilung aktualisieren.*
>
> Leiter der globalen Personalabteilung

Die Verwaltung von Datensätzen Millionen von Mitarbeitern einer globalen Belegschaft stellte erhebliche Herausforderungen dar. Manuelle, zeitaufwändige Prozesse erschwerten die konsequente Einhaltung regulatorischer Anforderungen, etwa der Datenschutz-Grundverordnung (DSGVO), während veraltete Systeme nicht mit modernen HR-Technologien kompatibel waren. Zur Unterstützung der Modernisierung setzte sich die Organisation das Ziel, ihren Rahmen für die Informationsverwaltung grundlegend zu erneuern, die Einhaltung gesetzlicher Vorschriften zu automatisieren und Sicherheit sowie von Anfang an in jede Phase des HR-Datenmanagements den Datenschutz zu integrieren.

Das neue Governance-Modell vereinheitlichte Richtlinien für Dokumentenaufbewahrung, Vernichtung und Zugriff über alle Regionen hinweg und ersetzte Tausende inkonsistente Vorlagen durch standardisierte globale Formate. Automatisierte Aufbewahrungs- und Löschrichtlinien gewährleisten nun die kontinuierliche Einhaltung der Vorschriften, reduzieren operationelle Risiken und entlasten die HR-Teams von manueller Überwachung. Verschlüsselung, Zugriffskontrollen und Verifizierungsprotokolle sichern die Datenintegrität, während die Automatisierung der Governance eine schnellere und verlässlichere Entscheidungsfindung ermöglicht.

Mit dieser Grundlage bereitet sich die Organisation auf die nächste Phase vor – den Einsatz von KI zur Verbesserung der Dokumentenklassifizierung, zur Automatisierung der Datensatzverwaltung und zur Stärkung der Governance im großen Maßstab. Durch die Verbindung technischer Kontrollen mit klaren Aufsichts- und Rechenschaftsmechanismen entwickelt sich das Unternehmen von reiner Compliance hin zu einer proaktiven, verantwortungsbewussten Governance – und schafft damit ein sicheres, datengesteuertes Umfeld, das für intelligente Innovation bereit ist.

Die fünf Merksätze

1. **KI-Governance als strategisches Gebot verankern.**
   Die Unterstützung der Geschäftsleitung ist sicherzustellen und klare Verantwortlichkeiten für die KI-Governance festzulegen, um zu gewährleisten, dass alle Initiativen mit ethischen, rechtlichen und organisatorischen Zielen im Einklang stehen.

2. **Ethik und Verantwortlichkeit in den KI-Lebenszyklus integrieren.**
   Ethische Richtlinien, Risikokontrollen, regulatorische Compliance und Rechenschaftsmechanismen sind in jede Phase zu verankern, von der Konzeption über die Implementierung bis zur Überwachung, um Voreingenommenheit und unbeabsichtigten Schaden proaktiv zu vermeiden.

3. **Führende Rahmenwerke und Standards aktivieren.**
   Rahmenwerke wie die OECD-Prinzipien, den EU AI Act, das NIST AI RMF und die ISO 42001 müssen implementiert werden, um bewährte Verfahren und regulatorische Anforderungen in umsetzbare Kontroll- und Aufsichtsmechanismen zu übersetzen.

4. **KI-Governance im gesamten Unternehmen integrieren.**
   Die KI-Governance ist mit der Unternehmens-, IT- und Daten-Governance auszurichten, indem Rollen, Verantwortlichkeiten und Prozesse klar definiert und eine umfassende Aufsicht von der Projektplanung bis zur Stilllegung sichergestellt werden.

5. **Kontinuierliche Verbesserung und Vertrauen fördern.**
   Regelmäßige Audits sind einzuführen, Governance-Protokolle an die Weiterentwicklung von Technologien und Vorschriften anzupassen und eine Kultur des Lernens und der Verantwortlichkeit zu verankern, um Vertrauen und langfristigen Wert zu sichern.

Kapitel Sieben

Die Architektur souveräner EAI-Implementierungen

Wie in früheren Kapiteln beschrieben, befinden sich rund 90 Prozent der weltweiten Daten hinter Firewalls und sind in privaten, proprietären oder sensiblen Umgebungen gespeichert. Nur etwa zehn Prozent sind öffentlich zugänglich, und diese geringe Anzahl hat die erste Welle der generativen KI maßgeblich vorangetrieben. Um das volle Potenzial von GenAI, Agentenbasierte KI (Agentic AI) und nicht zuletzt Künstlicher Allgemeiner Intelligenz (Artificial General Intelligence, AGI) auszuschöpfen, müssen Organisationen sichere und souveräne Mechanismen entwickeln, um ohne die Privatsphäre, Sicherheit oder nationale Kontrolle zu gefährden, auf die 90 Prozent dieser Daten zuzugreifen und sie zu nutzen. Dieses Kapitel zeigt, wie sich dies mit einem hybriden Ansatz erreichen lässt, der souveräne Daten und EAI auf einer EIM-Plattform integriert.

> *Neue Risiken, etwa die Ermächtigung ausländischer Regierungen, kritische Schaltstellen zu kontrollieren, haben weltweit Besorgnis ausgelöst.*[51]

In der digitalen Wirtschaft sind Daten das zentrale Gut. Sie befeuern Innovationen, steigern die Produktivität und bilden die Grundlage für die nationale Sicherheit. Da KI alle Sektoren verändert und sich geopolitische Rahmenbedingungen rasch weiterentwickeln, ist es für Führungsebenen unerlässlich, die Vertraulichkeit und den Schutz von Daten, Infrastruktur und KI-Fähigkeiten zu sichern. Diese Verantwortung reicht von der Unternehmensebene bis zur nationalen Ebene, wo souveräne Strategien für die KI-Vorreiterschaft entscheidend sind.

Die Fähigkeit eines Landes, im Zeitalter der KI führend zu sein, hängt von seiner Kontrolle über seine wertvollste digitale Ressource ab, den Daten. Ohne vollständige Souveränität riskieren Staaten, dass ihre digitale Infrastruktur technisch oder rechtlich unter fremde Gerichtsbarkeit gerät – ein Risiko, das nicht nur Innovationskraft, sondern auch nationale Sicherheit betrifft.

IT-Verantwortliche sehen die Kontrolle über Infrastruktur und Daten zunehmend als strategische Notwendigkeit. In einer Welt geopolitischer Spannungen, Handelsrestriktionen und komplexer Regulierungen ist die Abhängigkeit von weit entfernten oder politisch eingeschränkten Anbietern zu einem zentralen Geschäftsrisiko geworden. Vorausschauende Unternehmen kommen nicht nur ihren Compliance-Verpflichtungen nach, sondern führen robuste, rechtssichere Architekturen ein, die Störungen widerstehen, Rechtssicherheit gewährleisten und die betriebliche Kontinuität unter allen Umständen aufrechterhalten können.[52]

Definitionen der digitalen Souveränität

Weltweit rückt die Bedeutung der digitalen Souveränität immer stärker in den Fokus. Sie beschreibt die Fähigkeit eines Staates oder einer Organisation, die Kontrolle über digitale Vermögenswerte, Daten, Systeme und Abläufe zu behalten, um Unabhängigkeit von externen Einflüssen und die Einhaltung nationaler Vorschriften sicherzustellen.

Je nach Sensibilität der Daten kann digitale Souveränität verschiedene Komponenten umfassen:

- **Datensouveränität**: Sicherstellen, dass Daten innerhalb einer bestimmten Gerichtsbarkeit mit strengen Kontrollen gespeichert, verarbeitet und verwaltet werden, um den Zugriff oder die Weitergabe an ausländische Stellen oder gemäß ausländischem Recht zu verhindern.
- **Operative Souveränität**: Sicherstellen, dass die Geschäftstätigkeit in einem bestimmten Rechtsraum stattfindet und die Mitarbeiter, die die digitalen Vermögenswerte verwalten, Staatsbürger dieses Rechtsraums sind und mit entsprechenden Sicherheitsfreigaben ausgestattet sind.
- **Technologische Souveränität**: Kontrolle über Infrastruktur und Steuerungsebenen beibehalten, einschließlich physischer Sicherheit, Zugriffsrechten sowie der Verwaltung von Hardware, Software und Verschlüsselungsschlüsseln. Dazu gehört auch die Souveränität über die Steuerungsebene, also jene Dienste, die für die Integration von Anwendungen mit der zugrunde liegenden Infrastruktur entscheidend sind.
- **Rechtliche Souveränität**: Sicherstellen, dass Technologieanbieter und Cloud-Service-Anbieter ausschließlich dem Recht einer bestimmten Gerichtsbarkeit unterliegen.

Ein ausgewogener Hybridansatz

Um im KI-Zeitalter wettbewerbsfähig zu bleiben, müssen Nationen die Innovationskraft und Flexibilität globaler öffentlicher Cloud-Dienste nutzen. Diese Notwendigkeit führt jedoch zu einem grundlegenden Widerspruch zum Sicherheitsgebot, die souveräne Kontrolle zu wahren. Ein Hybridmodell ist die entscheidende Lösung für diese Herausforderung. Dieser Ansatz bringt beide Anforderungen in Einklang, indem er anerkennt, dass nicht alle Daten denselben Schutzbedarf haben. Sensible staatliche Daten werden auf national betriebenen Plattformen geschützt, während öffentliche Datensätze und bürgernahe Dienste über globale Hyperscaler skaliert werden können.

Ein führendes Technologie- und Dienstleistungsunternehmen nutzte GenAI in einem hybriden Ansatz, um historische Fälle mit Millionen Dokumenten und Terabytes an Daten zu analysieren.

Fallstudie

Ein globaler Technologieführer

Ein weltweit führender Anbieter von Technologien und Dienstleistungen bestimmt bei Automatisierung, Elektrifizierung, Digitalisierung und Vernetzung weltweit die Trends.

Als globales Unternehmen sind rechtliche Herausforderungen ein unvermeidlicher Bestandteil der im großen Maßstab ausgeführten Geschäftstätigkeit. Langwierige interne Untersuchungen und schwerfällige frühzeitige Fallbewertungen beeinträchtigten jedoch die Fähigkeit des Unternehmens, die nächsten Schritte zu definieren, und lenkten es von der Innovation und Wertschöpfung ab. Ineffiziente Prozesse führten zu erhöhten Kosten und Risiken, da es an Technologien fehlte, die das Fallwissen und die Kontrolle in der frühen Phase des Prozesses unterstützten. Das Unternehmen suchte nach einer technologischen Lösung, die es in die Lage versetzte, durch die schnelle Verarbeitung und Analyse großer Mengen interner Daten bessere und schnellere Entscheidungen zu ermöglichen und die zugleich dazu beitragen sollte, den Verlauf des Rechtsstreits aktiv mitzugestalten.

Das Unternehmen entschied sich für einen hybriden Ansatz. Das Team der Rechtsabteilung nutzte diesen Ansatz in Verbindung mit GenAI, um in der Phase der Fallbewertung umfangreiche Datensätze zu verarbeiten. Anschließend verwendeten sie, um ihre Fallstrategie festzulegen ein Large Language Model, um relevante Fragen zu stellen und innerhalb weniger Minuten Antworten zu erhalten. Auf diese Weise gelang es ihnen, juristischen Arbeitsabläufe zu transformieren, was schnellere datengestützte Entscheidungen und proaktive Untersuchungen ermöglichte. Durch Integration von KI und durch entsprechende Schulungen waren die Teams der Rechtsabteilung in der Lage, einen überlegenen Service zu bieten.

Hochrangige Architektur für souveräne Daten und KI

Architektur für Souveränität

Das Architekturmodell der EIM-Plattform zeigt die zentralen Komponenten für souveräne Daten- und KI-Funktionen auf einer EIM-Plattform für den öffentlichen und privaten Sektor. Es gewährleistet sichere, effiziente Bereitstellung von Diensten und schützt kritische Daten.

Zu den wichtigsten Aspekten dieser Architektur gehören:

- **Duale Datenarchitektur**: Sensible Daten werden innerhalb einer souveränen Schicht geschützt, während öffentlich verfügbare Daten in einer hybriden Cloud-Umgebung (d. h. einer Umgebung, die öffentliche und private Cloud integriert) verarbeitet werden.
- **Multiagenten-KI-Modell**: „Private KI-Agenten“ arbeiten innerhalb einer souveränen Infrastruktur, wohingegen „öffentliche KI-Agenten“ Dienste auf Basis einer Hybrid-Cloud bereitstellen, die sichere Grenzen und Datenintegrität gewährleisten.
- **Erweiterbarkeit**: Entwickelt, um zusätzliche Datensätze einzubinden.
- **Datensicherheit und Governance**: Einhaltung der Richtlinien und Kontrollen in Bezug auf Schutz und Datennutzung.
- **Grundprinzipien**: Vertrauen, Sicherheit, nationale Kontrolle und Resilienz.

Im Folgenden wird die oben beschriebene Architektur detailliert erläutert.

Die Infrastruktur- und Netzwerkschicht

Sensible Daten liegen in Infrastrukturen, die von vertrauenswürdigen Anbietern von Telekommunikation und Rechenzentrumbetrieben werden. Diese Umgebungen sind so aufgebaut, dass sie die höchsten Sicherheits- und Souveränitätsanforderungen erfüllen. Sie nutzen Zero-Trust-Protokolle – Sicherheitsframeworks, die jede Verbindung, jedes Gerät und jeden Benutzer fortlaufend überprüfen, anstatt vorauszusetzen, dass etwas sicher ist – sowie Air-Gap-Konfigurationen, bei denen kritische Systeme physisch oder logisch von öffentlichen Netzwerken getrennt bleiben, um unbefugten Zugriff oder Datenlecks zu verhindern. Alle Einsätze erfolgen innerhalb klar definierter nationaler oder regionaler Grenzen. Das operative Personal ist sicherheitsüberprüft und arbeitet ausschließlich innerhalb dieser Rechtsordnung, um die Einhaltung aller relevanten Gesetze, Vorschriften und Verteidigungsstandards zu gewährleisten.

Für Daten und Workloads, die keine vollständige Souveränität erfordern – etwa öffentlich zugängliche Datensätze oder bürger- und kundenorientierte digitale Anwendungen – bindet das Framework globale Hyperscaler ein. Diese Plattformen stellen Skalierbarkeit, Flexibilität und fortschrittliche Werkzeuge bereit, um Innovation, Reaktionsfähigkeit und Kosteneffizienz zu fördern und gleichzeitig unter strengen Governance-Auflagen zu arbeiten, die eine Offenlegung souveräner Daten verhindern.

In beiden Zonen bleibt die Architektur durch einen gemeinsamen Technologie-Stack verbunden, der Daten- und Informationsmanagement, KI-Modelle und KI-Anwendungen integriert und so eine konsistente, sichere Betriebsbasis schafft.

Die Ebene der Betriebstechnologie

Die Betriebstechnologie bildet die Brücke zwischen Infrastruktur, Netzwerk und Anwendungen. Sie stellt sicher, dass Daten, KI-Modelle, Plattformen, Anwendungen und KI-Dienste reibungslos bereitgestellt werden.

In einer Multi-Cloud- und Hybridumgebung ist die Standardisierung dieser Ebene unverzichtbar. Offene Protokolle und interoperable Frameworks ermöglichen es Unternehmen, Workloads flexibel zu verschieben – also Anwendungen und Daten nahtlos zwischen lokaler, privater und souveräner Cloud zu verlagern, ohne Anpassungen am Code vorzunehmen oder Kompromisse bei der Sicherheit einzugehen. Gerade für KI-Workloads ist dies entscheidend, da hier hohe Rechenlast, große Datenmengen und regulatorische Anforderungen sowohl Flexibilität als auch Kontrolle erfordern.

Die Steuerung der Betriebstechnologien umfasst Überwachung, Beobachtbarkeit und Automatisierung. Einheitliche Steuerungsebenen und Orchestrierungswerkzeuge sorgen für konsistente Konfigurationen, laufendes Patch-Management und kontinuierliche Compliance-Prüfungen. So stellt die operative Ebene sicher, dass jede KI-Anwendung – vom Modelltraining bis zur Inferenz – innerhalb vertrauenswürdiger Grenzen läuft, Zuständigkeiten klar geregelt bleiben und Prozesse sicher skaliert werden können.

Die Datenschicht

Die Datenschicht bildet die Grundlage für EAI und muss die Anforderungen sowohl des öffentlichen als auch des privaten Sektors erfüllen. Sie ermöglicht ein sicheres, intelligentes und skalierbares Datenmanagement, das sich im gesamten Regierungs- und Verwaltungsbereich einsetzen und zugleich flexibel erweitern lässt, um auch die Bedürfnisse privater Organisationen abzudecken.

Die Architektur stützt sich auf eine Kombination expliziter und impliziter Datenhierarchien. Zu den expliziten Strukturen gehören Ordnerhierarchien, Taxonomien, Schemata, Versionskontrolle und Prüfprotokolle. Zu den impliziten Strukturen gehören Metadatenfelder, semantische Beziehungen, Ontologien, Tags und nutzungsbasiertes Clustering. Durch metadatengesteuerte Orchestrierung und semantische Engines werden diese beiden Ebenen miteinander verknüpft, sodass KI-Systeme sowohl strukturierte als auch unstrukturierte Daten analysieren und interpretieren können.

Agentenbasierte Anwendungsfälle

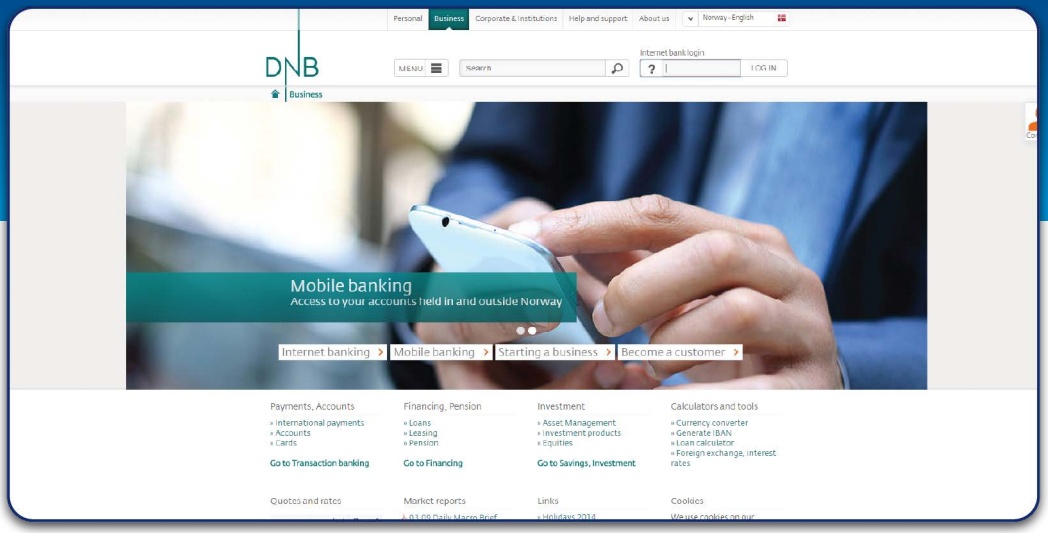
Agenten steigern die Produktivität und Effizienz und liefern gleichzeitig einen besseren Kundenservice und bessere Geschäftsergebnisse. Beispiele für agentenbasierte Anwendungsfälle sind unter anderem:

- **Gesundheitswesen**: Personalisierte Gesundheitsnavigation, Revision der Leistungsansprüche und virtuelle Triage
- **Immobilien**: Antragsbearbeitung, Genehmigungen, Revision der Anspruchsberechtigung und Subventionsverwaltung
- **Bankwesen**: Proaktive Betrugserkennung, personalisierte Finanzberatung und automatisierte Kreditbearbeitung
- **Transport**: Dynamische Routenoptimierung, autonomes Flottenmanagement und vorausschauende Wartung
- **Besteuerung**: Prüfungskennzeichnung, Betrugserkennung und Unterstützung bei der Erstellung von Steuererklärungen

In der folgenden Fallstudie wird gezeigt, wie DNB Finans KI und Daten einsetzt, um die Verwaltung von Fahrzeugflotten zu optimieren, Betrugsversuche zu erkennen und die Zufriedenheit der Leasingkundinnen und -kunden zu erhöhen.

Fallstudie

DNB Finans



DNB Finans

Die DNB Bank Group in Norwegen ist die zweitgrößte Bank Skandinaviens. Sie beschäftigt rund 13.430 Personen und verwaltet ein Gesamtvermögen von 250 Milliarden Euro (273 Milliarden US-Dollar). Ihre Tochtergesellschaft DNB Finans zählt zu den größten Finanzunternehmen in der nordischen Region. Im Privatkundensegment nimmt DNB Finans im Markt für Autofinanzierungen eine führende Position ein und betreut mehr als 300 000 finanzierte Fahrzeuge.

Das Unternehmen sucht kontinuierlich nach neuen Wegen, seinen Kundinnen und Kunden einen spürbaren Mehrwert zu bieten. Besonders wertvoll sind dabei Dienstleistungen, die Unternehmen helfen, ihre Kosten zu kontrollieren, indem sie Transparenz über sämtliche Ausgaben schaffen. Um dies zu erreichen, wollte die Autoleasing-Abteilung von DNB Finans Business-Intelligence-Funktionen bereitstellen, die den Kundinnen und Kunden jederzeit aktuelle Daten liefern.
Das System sollte beispielsweise Statistiken zu allen fahrzeugbezogenen Kosten anzeigen – von Kraftstoffverbrauch und CO_2-Emissionen über Leasingkosten bis hin zu Schadensmeldungen und Betrugshinweisen. Zudem mussten individuelle Kostenstellenstrukturen eingerichtet werden, um Aktivitäten nach Geschäftsbereichen auswerten zu können. Besonderen Wert legte DNB Finans auf eine Software, die sich ohne Schulung intuitiv bedienen lässt. Ziel war ein Nutzererlebnis, das in seiner Einfachheit und Übersichtlichkeit sozialen Netzwerken wie Facebook ähnelt.

Das Unternehmen führte daraufhin eine Business-Intelligence- und Reporting-Lösung ein, die inzwischen von mehr als 30 000 Leasingkundinnen und -kunden genutzt wird. Die Lösung überzeugt durch ein übersichtliches Design mit farbigen Visualisierungen, nutzerfreundlichen Dashboards und logischen Kontrollmechanismen zur Betrugserkennung und zur effizienten Verwaltung von Fahrzeugflotten. Seit der Einführung konnte DNB Finans den Kundenzufriedenheitswert für die Qualität der Reporting-Lösung auf einer Skala von 1 bis 6 von 4,4 auf 5,1 steigern.

Zugleich stiegen die Nutzeranmeldungen um 31 Prozent, was zu einer höheren Aktivität im Bereich Autofinanzierung führte. Das Unternehmen erwartet innerhalb von nur zweieinhalb Jahren eine Kapitalrendite. Die Kundinnen und Kunden erkennen nun frühzeitig Probleme wie übermäßige Kilometerleistung oder betrugsverdächtige Kraftstoffkosten und können die verantwortlichen Geschäftsbereiche eindeutig zuordnen. Dadurch verbessern sie ihre Handlungsfähigkeit und erhöhen ihre Loyalität durch den gezielten Einsatz wertvoller Geschäftsinformationen. Diese Lösung verschafft DNB Finans in einem stark umkämpften Markt einen deutlichen Wettbewerbsvorteil.

Detaillierte Architektur für souveräne Daten und KI

Duale Datenarchitektur

Das Diagramm zeigt eine detaillierte Architektur für eine sichere Daten- und KI-Plattform, die in zwei Zonen gegliedert ist und zwischen einer nicht souveränen beziehungsweise öffentlichen Zone und einer souveränen beziehungsweise privaten Zone unterscheidet.

Ziel dieses Aufbaus ist es, sensible Unternehmens- und Regierungsdaten sowie interne Abläufe klar von öffentlichen oder weniger sensiblen Bereichen zu trennen und zugleich kontrollierte Interaktionen mit der öffentlichen Zone zu ermöglichen, wo dies erforderlich ist. Die Architektur erfüllt die Anforderungen an sichere Daten- und KI-Umgebungen und bietet gleichzeitig die Flexibilität, Implementierungen effizient und kostengünstig zu gestalten – eine Voraussetzung für verbesserte Kundenerlebnisse.

So kann es aufgeschlüsselt werden.

Die nicht-souveräne/öffentliche Zone

In einer Architektur mit dualer Souveränität für Daten und KI fungiert die öffentliche Zone als kontrollierte Schnittstelle zwischen öffentlichem Wissen und interner Unternehmensintelligenz. Sie ermöglicht es Organisationen, nicht sensible und öffentlich verfügbare Daten sowie KI-Dienste zu nutzen, ohne dabei die interne Souveränität oder Compliance-Verpflichtungen zu gefährden. Durch die Isolierung öffentlicher Interaktionen über sichere Zugangswege und Bereinigungs-/Filterprotokolle schafft diese Zone einen geschützten Rahmen, in dem Innovation und externe Vernetzung stattfinden können, ohne die Grenzen der Datensicherheit zu überschreiten.

Diese Zone besteht aus:

Nicht-sensible und öffentliche KI-Agenten: Dieser Bereich beinhaltet Schnittstellen wie Large Language Model-APIs (LLM), die keine sensiblen Daten verarbeiten. Die Public-Agent-API-Schicht stellt Endpunkte für den Datenzugriff bereit und umfasst zugleich Kontrollmechanismen zur Verwaltung der Nutzung, zur Sicherung von Sitzungen und zur Speicherung von Sitzungsdaten.

Nicht-staatliche/öffentliche Datenquellen: Das System greift hier auf nicht sensible und öffentliche Datenbanken, veröffentlichte Vorschriften und Servicehandbücher zu. Darüber hinaus nutzt es öffentliche Wissensdatenbanken, um die Genauigkeit und Relevanz der generierten Informationen zu erhöhen.

Sicherheitskontrollen: Zu den Sicherheitsmaßnahmen gehört die Bereinigung personenbezogener Daten (PII) mithilfe von Modellen zur Erkennung benannter Entitäten (Named Entity Recognition, NER), um sensible Informationen wie das Geburtsdatum vor der Gateway-Validierung zu entfernen.

Die souveräne/private Zone

Die souveräne oder private Zone bildet den intellektuellen Kern einer dual-souveränen Daten- und KI-Architektur, in dem sensible und missionskritische Operationen unter vollständiger organisatorischer Kontrolle ablaufen. Diese Zone ist für regulierte und hochsichere Umgebungen konzipiert und steuert den Einsatz privater Agenten, vertraulicher Datenquellen und sicherer Recheninfrastruktur. Jeder Prozess – vom Modelltraining bis zur Inferenz – wird innerhalb eines Zero-Trust- und Air-Gap-Frameworks ausgeführt, wodurch nationale, unternehmensbezogene oder institutionelle Daten jederzeit souverän, regelkonform und revisionsfähig bleiben.

Diese Zone besteht aus:

Private Agenten: Diese Funktionen stehen ausschließlich Nutzern und Agenten des öffentlichen oder privaten Sektors zur Verfügung, die auf sensible Daten zugreifen. Sie basieren auf einer agentenbasierten Plattform, die eine vom Internet getrennte Bereitstellung ermöglicht. Die Private-Agent-API-Schicht ist nur intern zugänglich und mit einer Zero-Trust-Sicherheitsarchitektur ausgestattet.

Souveräne Datenquellen: Hierzu gehören geschützte Datenbanken, die vertrauliche Informationen wie Personal- oder Finanzdaten sowie sensible Abteilungsunterlagen enthalten. In diesem Zusammenhang nutzt die Retrieval-augmented Generation (RAG)-Pipeline geschützte Wissensquellen, darunter juristische Präzedenzfälle, um präzise KI-Ergebnisse zu erzeugen.

RAG ist in der Unternehmens-KI unverzichtbar, da Modelle damit zur Laufzeit kontrollierten Zugriff auf relevantes Wissen erhalten, statt sich ausschließlich auf ihre Trainingsdaten zu stützen. Frühe Implementierungen – häufig als naive RAG-Systeme bezeichnet – beschränkten sich darauf, Textblöcke zu extrahieren und in Eingabeaufforderungen einzufügen, was bei unvollständigem Kontext zu Ungenauigkeiten oder Halluzinationen führen konnte. Graph-basiertes RAG (Graph-RAG) bildet die nächste Entwicklungsstufe: Es strukturiert das Unternehmenswissen in Form von Beziehungen und Entitäten und ermöglicht es dem Modell dadurch, nicht nur Dokumente, sondern auch die zutreffende kontextuelle Bedeutung abzurufen. Auf diese Weise steigert Graph-RAG Präzision, Nachvollziehbarkeit und Vertrauenswürdigkeit und verringert zugleich den Bedarf an überdimensionierten Eingabeaufforderungen und fehleranfälligem „Context Stuffing“.

Unternehmen verfügen derzeit über drei zentrale Ansätze, um KI-Modellen Kontext bereitzustellen: groß angelegte, sorgfältig formulierte Eingabeaufforderungen innerhalb erweiterter Kontextfenster, RAG- bzw. Graph-RAG-Abrufpipelines sowie Modelloptimierung durch Feinabstimmung oder Einbettungsoptimierung. Die Zukunft der Unternehmens-KI liegt in der intelligenten Orchestrierung dieser Verfahren – im Übergang von der manuellen Prompt-Erstellung zu gesteuerten, strukturierten und skalierbaren Kontextpipelines, die KI-Systeme befähigen, sicher und zuverlässig mit Unternehmenswissen zu arbeiten.

Verarbeitungsinfrastruktur: Diese Komponente stellt beschleunigte Rechenkapazitäten für sichere Arbeitslasten bereit. Sie umfasst eine unveränderliche Audit-Trail-Datenbank zur Überwachung der Regelkonformität sowie die optionale Feinabstimmung von LLM-Modellen, um eine kontrollierte und nachvollziehbare Verarbeitung zu gewährleisten.

Gemeinsame Komponenten und Sicherheitsmaßnahmen

Zwischen beiden Zonen liegt ein API-Gateway, das strenge Authentifizierungs- und Autorisierungsprotokolle durchsetzt. Dazu gehören Identitätsprüfungen über Zugriffskontrollföderationen und die Nutzung von Multi-Faktor-Authentifizierung (MFA). Datenklassifizierungsmechanismen kennzeichnen Inhalte automatisch nach ihrer Sensitivitätsstufe, während Tools zur Verhinderung von Datenverlusten sicherstellen, dass sensible Informationen nicht über die Unternehmensgrenzen hinausgelangen.

Ein Abfrage-Router leitet eingehende Anfragen entsprechend ihrer Klassifizierung an die jeweilige Zone weiter. Nur für die Öffentlichkeit freigegebene Antworten dürfen wieder in den nicht-souveränen Bereich zurückgeführt werden, nachdem die geschützten Daten entfernt wurden.

Die Infrastruktur- und Netzwerkschicht

Beide Zonen basieren auf robusten Infrastruktur- und Netzwerkschichten, die bei Bedarf eine physische oder logische Trennung gewährleisten, um in allen Betriebsumgebungen Sicherheit, Stabilität und Datenintegrität sicherzustellen.

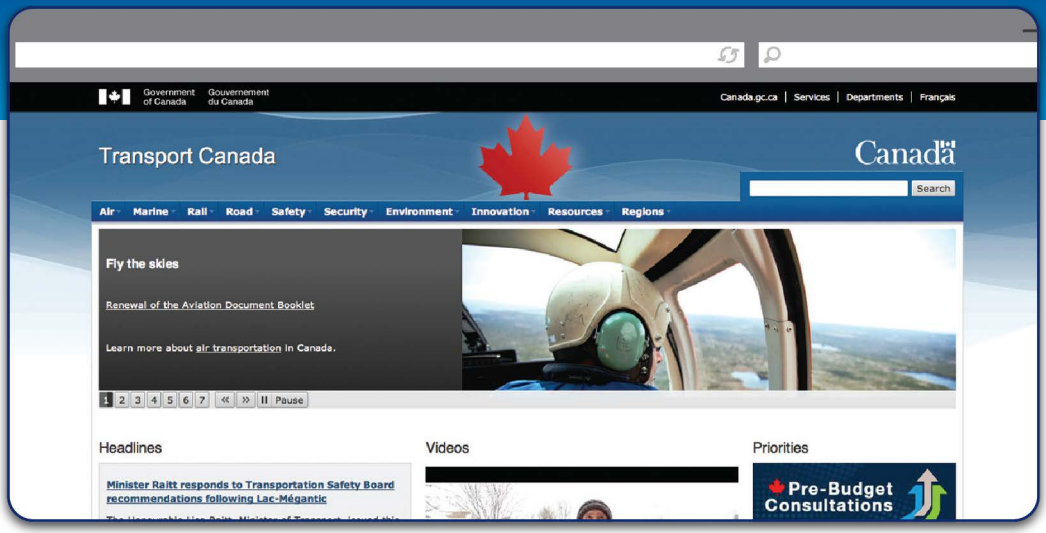
Souveräne Daten sind kein „neues“ Konzept: Eine Geschichte der Anwendungsfälle im Privatsektor

Anwendungsfälle aus dem Privatsektor – sowohl in regulierten als auch in unregulierten Branchen – können dazu beitragen, die souveräne Architektur weiterzuentwickeln und zu präzisieren. Der Schutz privater und sensibler Daten ist seit Jahrzehnten unverzichtbar, und ein solides Informationsmanagement spielt dabei eine entscheidende Rolle.

Der folgende Beitrag über Transport Canada zeigt, wie ein wirksames Informationsmanagement sichere Inhalte für die Anwender bereitstellt und zugleich zentrale Informationen in die Geschäftsprozesse integriert.

Fallstudie

Transport Canada



Transport Canada

Die Mission von Transport Canada besteht darin, dem öffentlichen Interesse zu dienen, indem in Kanada ein sicheres, effizientes und umweltverträgliches Verkehrssystem gefördert wird. Dies erfordert ein leistungsfähiges Informationsmanagement, das eine fundierte und zeitnahe Entscheidungsfindung innerhalb eines breiten Netzwerks von Partnern ermöglicht – darunter 15 Unternehmen im Besitz der Krone, 17 Hafenbehörden, 21 Flughafenbehörden und weitere Organisationen der gemeinsamen Verwaltung.

Angesichts der zunehmenden elektronischen Informationsverbreitung, der Anforderungen an Datenschutz und Datensicherheit, des drohenden Verlusts von Unternehmenswissen durch Personalfluktuation sowie des wachsenden Bedarfs an Echtzeitzugriff auf Daten für Anfragen und Rechtsverfahren setzte die kanadische Regierung auf eine E-Government-Lösung auf Basis von Aufzeichnungs-, Dokumenten- und Informationsmanagement. Transport Canada war die erste Regierungsbehörde Kanadas, die mit über vier Millionen Datensätzen in einer einzigen Bibliothek und rund 5.200 Nutzern an mehr als 117 Standorten eine solche Lösung implementierte.

Damit handelt es sich um die größte Implementierung einer zentralen Bibliothek im öffentlichen Sektor des Landes. Das System dient als integriertes Werkzeugset, das von der Erfassung und Speicherung über Organisation, Abruf, Weitergabe, Wiederverwendung und Schutz bis hin zur sicheren Entsorgung die umfassende Nutzung elektronischer Dokumente ermöglicht. Die Informationslösung entwickelte für die Leitung und Mitarbeiter von Transport Canada sich zu einer geschäftskritischen Anwendung. Sie half der Organisation, die Genauigkeit ihrer Unterlagen sicherzustellen, eine geografisch verteilte Belegschaft zu vernetzen, rechtliche Verpflichtungen – einschließlich der Anforderungen an elektronische Beweissicherung – zu erfüllen, die Produktivität zu steigern und das Informationsmanagement mit der Initiative Government On-Line (GOL) abzugleichen. Durch den Einsatz dieses Systems konnte Transport Canada seine Produktivitätseinsparungen verdreifachen und Einsparungen von bis zu 4,6 Millionen Dollar erzielen. Das System amortisierte sich bereits nach 1,17 Jahren, und die Organisation geht weiterhin von einer jährlichen Kostenvermeidung in ähnlicher Höhe aus.

Mit der zunehmenden Nutzung von Cloud-Technologien arbeiten heute die meisten Institutionen mit hybriden Bereitstellungsmodellen, die On-Premises-Systeme, Rechenzentren und öffentliche Cloud-Umgebungen kombinieren. Mit der wachsenden Verbreitung von KI ist zu erwarten, dass sich diese etablierten Modelle weiterentwickeln, um künftig hybride KI-Strukturen zu unterstützen.

Ein Beispiel liefert das Bundesrechenzentrum (BRZ) in Österreich, das sich für einen hybriden Ansatz entschied. Es nutzt ein cloudbasiertes Informationsmanagementsystem, um sensible Daten von zwölf Regierungskunden, vierzig Regierungsanwendungen sowie über zehn ERP- und Mailsystemen zu konsolidieren und zentral zu verwalten.

Fallstudie

Bundesrechenzentrum (BRZ)

Cloudbasiertes Informationsmanagement beim BRZ

Das Bundesrechenzentrum (BRZ) ist der zentrale IT-Dienstleister der österreichischen öffentlichen Verwaltung. Mit rund 1.200 Beschäftigten und einem jährlichen Umsatz von 265,3 Millionen Euro entwickelt und betreibt das BRZ erfolgreich E-Government-Dienste für Ministerien, Universitäten, Sozialversicherungsträger und öffentliche Einrichtungen. Es steuert 320 IT-Prozesse, versorgt über 1.200 Standorte in ganz Österreich mit Infrastruktur und betreut etwa 30.000 Arbeitsplätze.

Im Jahr 2000 stellten die Grund- und Handelsregister des österreichischen Justizministeriums ein Beispiel für stark fragmentierte Prozesse dar. Während die Daten des Grundbuchs bereits seit den 1980er-Jahren digital verwaltet wurden, blieben die Originaldokumente in physischen Gerichtsarchiven und waren in laufenden Verfahren nicht zugänglich. Dadurch entstanden hohe Archivierungskosten und das Risiko des Dokumentenverlusts.

Um diese Probleme zu lösen, implementierte das BRZ eine Enterprise-Content-Management-(ECM)-Lösung. Nach der erfolgreichen Pilotphase im Grundbuchamt erhielt das BRZ vermehrt Anfragen anderer Behörden zur elektronischen Dokumentenverwaltung und Prozessintegration. Daraufhin entwickelte es die skalierbare ECM-Infrastruktur „eGov Archive Service“ – den ersten privaten ECM-Cloud-Service Österreichs.

Diese Lösung bietet eine stabile Plattform für zwölf Regierungskunden, vierzig Regierungsanwendungen, mehr als zehn ERP-Systeme sowie verschiedene Mailsysteme. Der eGov-Archivdienst verwaltet 45 Terabyte (TB) Daten bzw. 400 Millionen Objekte, verarbeitet täglich rund eine Million Transaktionen und wird von 30.000 Nutzern – darunter Steuerprüfer, Richter, Polizisten, Zollbeamte, Personalverantwortliche und Buchhalter – sowie potenziell allen österreichischen Bürgerinnen und Bürgern genutzt. Die angebotenen Dienste reichen von Verwaltung, Zugriff, Routing und Suche über die rechtskonforme Archivierung bis hin zur Integration mit Fachanwendungen und ERP-Systemen. So entstand eine umfassende, cloudbasierte Lösung für das Dokumenten- und Informationsmanagement der österreichischen Verwaltung.

Grundlage der agentenbasierten KI

Die Erschließung der privaten Datenbestände weltweit stellt die zentrale Herausforderung für die nächste Entwicklungsphase der künstlichen Intelligenz dar. Dieses Kapitel hat den architektonischen Rahmen aufgezeigt, mit dem sich dieser Prozess sicher gestalten lässt. Die Lösung besteht in einem hybriden Dual-Zonen-Modell, das eine geschützte private Zone für sensible Daten und eine öffentliche Zone für andere Workloads bereitstellt. Auf diese Weise entsteht eine sichere Umgebung, in der private KI-Agenten ohne das Risiko eines Datenlecks vertrauliche Unternehmens- oder Regierungsdaten analysieren und verarbeiten können, während öffentliche KI-Agenten nicht-souveräne Aufgaben übernehmen. Dadurch entsteht ein Gleichgewicht zwischen Kontrolle, Skalierbarkeit und Wettbewerbsfähigkeit. Diese Architektur bildet die entscheidende Grundlage für den Einsatz fortschrittlicher KI und schafft das Vertrauen und die Steuerbarkeit, die notwendig sind, um agentenbasierte KI im großen Maßstab zu nutzen, was das Thema des nächsten Kapitels ist.

Die fünf Merksätze

1. **Mandat zur souveränen Datenkontrolle.**
   Die vollständige Kontrolle über nationale und organisatorische Daten sowie über die digitale Infrastruktur muss oberste Priorität haben und aktiv durchgesetzt werden. Richtlinien und technische Maßnahmen sind festzulegen, um externe Einflussnahme auszuschließen, den Datenstandort zu sichern und die Einhaltung nationaler Vorschriften zu gewährleisten.

2. **Dual-Zone-Hybrid-KI-Architektur implementieren.**
   Sensible Daten und Arbeitslasten sollten auf einer sicheren, im Inland betriebenen Infrastruktur ausgelagert werden. Öffentliche Cloud-Plattformen sind nur für unkritische, skalierbare Anwendungen zu nutzen, um Innovation und Sicherheit in Einklang zu bringen.

3. **Multiagenten-KI-Modelle strategisch einsetzen.**
   Private KI-Agenten sollten in geschützten Zonen aktiviert werden, um sensible Daten sicher zu analysieren und darauf zu reagieren. Öffentliche KI-Agenten sind für nicht-souveräne Aufgaben zu nutzen, um Organisationen die Skalierung von KI-Innovationen zu ermöglichen, ohne geschützte Assets zu gefährden.

4. **Strenge Governance und Sicherheit durchsetzen.**
   Es sollten strenge Authentifizierungsverfahren, fortschrittliche Datenklassifizierung und robuste Tools verwendet werden, um Datenverlust zu verhindern. Es muss sichergestellt werden, dass alle kritischen Aktivitäten mit unveränderlichen Prüfprotokollen erfasst und gemäß Zero-Trust-Protokollen betrieben werden, um maximale Transparenz zu gewährleisten.

5. **Die Einführung bewährter Hybridarchitekturen beschleunigen.**
   Erfolgsmodelle aus regulierten Branchen, die hybride Umgebungen bereits effektiv nutzen, sind einzuführen. Es sollte in die sichere Integration von privater und öffentlicher Cloud-Infrastruktur investiert werden, um das Potenzial der 90 Prozent privaten Daten zu erschließen – der entscheidende Schritt für die KI der nächsten Generation.

Kapitel Acht

Einsatz agentenbasierter KI in der Praxis

In diesem Kapitel wird untersucht, wie KI im gesamten Unternehmen eingesetzt werden kann. Ein wirksames Rahmenkonzept setzt das Verständnis der drei Ebenen der KI voraus: generative KI, die erstellt und synthetisiert; agentenbasierte KI, die Entscheidungen trifft und handelt, sowie künstliche allgemeine Intelligenz, die Denken und Verstehen über verschiedene Bereiche hinweg ähnlich dem menschlichen Geist erweitert. Gemeinsam bilden diese Ebenen die Basis für intelligente und anpassungsfähige Unternehmenssysteme.

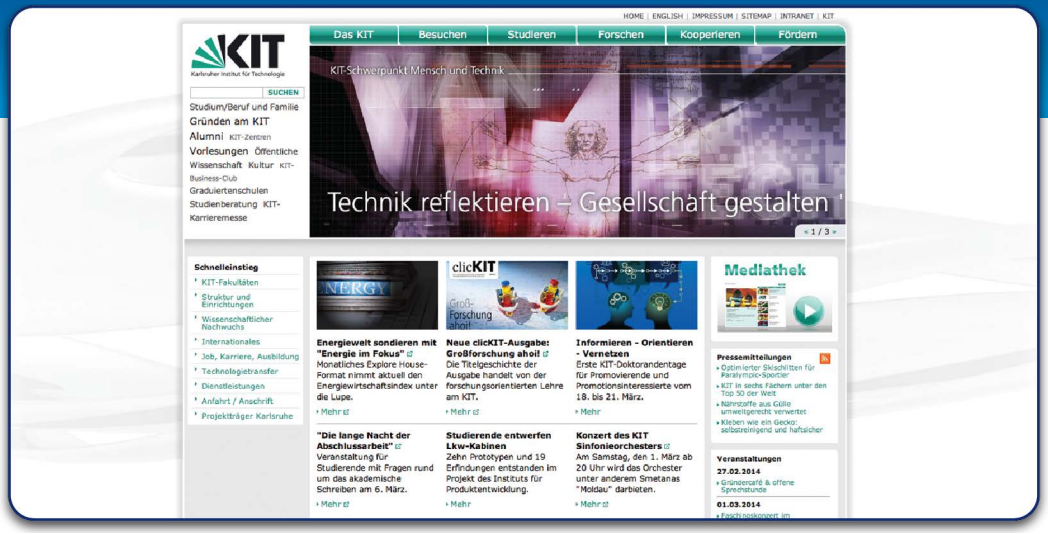
KI im Unternehmensumfeld hat die Phase des Neuen längst überwunden. Unternehmen experimentieren nicht mehr mit einfachen Textgeneratoren oder nutzen KI nur zum Erstellen der Betreffzeilen von E-Mails. Der Schwerpunkt liegt nun auf autonomen Partnern, die eng mit Menschen zusammenarbeiten. EAI-Agenten sind intelligente Softwareeinheiten, die nicht nur Ergebnisse erzeugen, sondern in dynamischen Geschäftsumgebungen eigenständig handeln, reagieren und sich anpassen. Im Gegensatz zu statischen Systemen oder klassischen Eingabeaufforderungen agieren sie wie digitale Kollegen: Sie führen mehrstufige Prozesse aus, lösen kontextbezogene Aufgaben und unterstützen den täglichen Betrieb schnell, beständig und präzise. Der Unterschied ist fein, aber entscheidend – diese neue Generation von KI beantwortet nicht nur Fragen, sondern übernimmt Aufgaben.

EAI-Agenten entwickeln sich von bloßen Effizienzwerkzeugen zu echten Leistungsfaktoren für Unternehmen. Stets verfügbar, konstant präzise und zunehmend fähig, komplexe Zusammenhänge zu erfassen, liefern sie messbare Ergebnisse ohne zusätzlichen Aufwand. Ob bei der Automatisierung von Prozessen, der Verwaltung von Kundeninteraktionen, der Analyse umfangreicher Datensätze oder der Einarbeitung von Nutzern EAI-Agenten bringen Präzision und Skalierbarkeit in wiederkehrende Abläufe. Ihre Zuverlässigkeit sorgt für einheitliche Nutzererlebnisse, während ihre Geschwindigkeit große Datenmengen in verwertbare Erkenntnisse verwandelt. Das Resultat geht über reine Kosteneinsparungen hinaus: Es entstehen neue Freiräume für strategische Planung, Innovation und Wachstum.

Durch die Verbindung von EIM und EAI gewinnen Organisationen des öffentlichen Sektors wie das Karlsruher Institut für Technologie (KIT) Zugang zu modernen Analysefunktionen, die darauf ausgerichtet sind, den tatsächlichen Informationswert zu erfassen, aufzubereiten und nutzbar zu machen – zur Verbesserung von Forschung und Analyse. Mehr dazu in der folgenden Fallstudie.

Fallstudie

Karlsruhe Institute of Technology



Das Karlsruher Institut für Technologie (KIT)

Das Karlsruher Institut für Technologie (KIT), eine der weltweit führenden Forschungseinrichtungen im Ingenieurwesen, entstand 2009 aus dem Zusammenschluss des Forschungszentrums Karlsruhe und der Universität Karlsruhe. Als Mitglied der Helmholtz-Gemeinschaft, der größten Wissenschaftsorganisation Deutschlands, trägt das KIT maßgeblich zur nationalen und internationalen Spitzenforschung bei. Im Einklang mit seiner Mission konzentriert sich das Institut auf drei strategische Schwerpunkte: Forschung, Lehre und Innovation. Mit derzeit rund 9.000 Beschäftigten und 24.000 Studierenden suchte das KIT nach einer modernen Lösung, um Forschern, Studierenden und der Öffentlichkeit einen schnelleren Zugang zu Informationen auf über 600 Websites und rund 200.000 Unterseiten zu ermöglichen.

Im Backend sollte eine leistungsfähige Website-Management-Plattform die Arbeit der weltweit rund 1.300 Redakteure unterstützen – durch automatisierte Vergabe von Metadaten, Schlüsselwörtern und Textauszügen. Zudem sollte eine kollaborative Plattform geschaffen werden, die den Austausch zwischen Forschern, Wissenschaftlern und Studierenden fördert.

Das KIT setzt E-Government-Technologien wie semantische Navigation und Inhaltsanalyse in Verbindung mit einem integrierten Website-Management ein, um Seiteninhalte zu optimieren und relevante Suchergebnisse bereitzustellen. Manuelle, zeitaufwendige Aufgaben wurden durch eine automatisierte Lösung ersetzt, die Metadaten vergibt und mithilfe von Entitätsextraktion automatisch Teaser-Texte für neue Seiten erstellt. Dadurch sparen die Nutzer Zeit und vermeiden Fehler. Besucher erhalten durch Facettensuche und verwandte Treffer personalisierten Zugriff auf besonders relevante Informationen – was das Nutzungserlebnis deutlich verbessert. Dank des vereinfachten Informationszugangs und der Möglichkeit, mit Forschern ähnlicher Fachrichtungen in Kontakt zu treten, hat sich die Website zu einem modernen Forschungsnetzwerk entwickelt, das die Anforderungen aller Beteiligten erfolgreich erfüllt.

Drei Ebenen der KI

Generative KI (GenAI) hat mit Modellen wie ChatGPT von OpenAI und Gemini von Google große Popularität bei Verbrauchern erlangt.
Diese und andere große Sprachmodelle (LLMs) sowie KI-Anwendungen werden darauf trainiert, Vorhersagen und Empfehlungen auf Basis öffentlich zugänglicher Datenquellen zu treffen – etwa aus Websites, Nachrichten, Reddit oder Wikipedia. Zwar liefern GenAI-Modelle wertvolle allgemeine Erkenntnisse, bleiben jedoch auf übergreifende Aufgaben beschränkt. Der Grund liegt im fehlenden Zugang zu privaten, aktuellen und unternehmensspezifischen Daten, die für konkrete Geschäftsanwendungen entscheidend sind.

Agentenbasierte KI beschreibt Systeme, die als autonome Agenten agieren. Im Gegensatz zu Modellen, die lediglich auf Eingaben reagieren, können solche Agenten ihre Umgebung wahrnehmen, mehrstufige Handlungspläne entwickeln, eigenständig Entscheidungen treffen und Werkzeuge einsetzen, um aktiv auf definierte Ziele hinzuarbeiten.

Im Unternehmensumfeld bilden sie damit einen zentralen Produktivitätsmotor, gespeist durch Daten als Treibstoff. Sie können auf vertrauliche Unternehmensinformationen und interne Anwendungen zugreifen und dadurch komplexe Abläufe automatisieren, die bislang menschliches Urteilsvermögen erforderten. Agentenbasierte KI wird von einem „digitalen Gehirn" angetrieben – einem zentralen, leistungsfähigen Modell, das jahrzehntelange Daten und menschliche Rückmeldungen verarbeiten kann.

Diese Technologie verändert bereits ganze Branchen, und Unternehmen, die es versäumen, agentenbasierte KI einzuführen und zu orchestrieren, werden den Anschluss verlieren. Sie ist zugleich ein Weg zur Künstlichen Allgemeinen Intelligenz (AGI).

AGI bezeichnet eine Form der KI, die – ähnlich wie der Mensch – Wissen verstehen, erlernen und in einem breiten Spektrum komplexer Aufgaben anwenden kann.[53] Eine solche Technologie wäre imstande, nicht nur einzelne Branchen, sondern ganze Gesellschaften neu zu gestalten.

Wie bei allen modernen KI-Systemen hängen die Fähigkeiten einer potenziellen AGI wesentlich von der Qualität und dem Umfang der Trainingsdaten ab.
Die Grundlagen von AGI werden wahrscheinlich aus der Orchestrierung Tausender spezialisierter agentenbasierter KI-Instanzen innerhalb eines sicheren und souveränen Rahmens entstehen, wie in Kapitel 7 beschrieben.

Entwicklung agentenbasierter KI für das Unternehmen

Gute Daten und wirksame Geschäftsprozesse bilden die Grundlage für optimale KI-Ergebnisse – und umgekehrt gilt das ebenso. Mit gesicherten Unternehmensdaten und auf die jeweilige Domäne abgestimmten großen Sprachmodellen (LLMs) entsteht die Basis für den Aufbau agentenbasierter Fähigkeiten, die echten Geschäftswert schaffen.

Agentenbasierte KI ist nicht nur Technologie, sondern ein Ansatz zur Entwicklung von Fähigkeiten, die auf bestehenden Geschäftsprozessen aufbauen und das Zusammenspiel von Menschen, Abläufen, Unternehmenskultur und Veränderungsmanagement fördern. Richtig umgesetzt wird sie zu weit mehr als einem Experiment in generativer Programmierung – sie wird zu einem Instrument messbarer Wertschöpfung. Anstatt ein Modell lediglich dazu aufzufordern, ein Dokument zusammenzufassen, kann agentenbasierte KI einen Engpass im Arbeitsablauf erkennen, Aufgaben in Teilaufgaben zerlegen, APIs oder andere Systeme zur Ausführung aufrufen, Ergebnisse überwachen, daraus lernen und anschließend den nächsten Schritt mit minimaler menschlicher Unterstützung optimieren.

Ein Beispiel liefert die NANDA-Initiative des Massachusetts Institute of Technology: Laut der Studie „State of AI in Business 2025“ erzielten nur rund 5 % der Pilotprojekte für generative KI in Unternehmen ein schnelles Umsatzwachstum, während die übrigen 95 % keinen messbaren Einfluss auf Gewinn und Verlust verzeichneten. Den Forschern zufolge lag das Problem nicht im Modell oder in der Hardware, sondern in der Lernlücke – also in der Unfähigkeit von KI-Systemen und organisatorischen Arbeitsabläufen, sich gemeinsam weiterzuentwickeln.[54]

Die wichtigsten Erkenntnisse: Investitionen in KI sollten dort erfolgen, wo sie direkt auf konkrete Prozesse ausgerichtet sind, und nicht allein auf sichtbarkeitsstarke Projekte. Feedbackschleifen müssen integriert werden, damit das System fortlaufend lernt. Eine enge Einbettung in bestehende Abläufe ist entscheidend, anstatt ein generisches Tool lediglich hinzuzufügen und ein aktives Veränderungsmanagement ist erforderlich, damit Beschäftigte und Unternehmenskultur den Wandel mittragen. Kurz gesagt: Wenn Datenqualität, Prozessausrichtung, organisatorische Bereitschaft und domänenspezifische Modellierung ineinandergreifen, kann ein Unternehmen den Schritt vom Pilotprojekt zur echten, EAI-gestützten Wertschöpfung vollziehen.

Die Bedeutung agentenbasierter KI

Organisationen setzen zunehmend KI ein, um dem Druck des rasanten wirtschaftlichen und technologischen Wandels zu begegnen – von der digitalen Transformation bis zur Entwicklung neuer Geschäftsmodelle. Echtzeit-Entscheidungsfindung, globale Reichweite und Anpassungsfähigkeit an Störungen sind dabei entscheidend.[55] KI-Agenten, die Aufgaben halb- oder vollautonom ausführen, helfen Unternehmen, wettbewerbsfähig zu bleiben, Informationsflüsse zu skalieren, die kognitive Belastung zu verringern und die Agilität zu steigern.

Der Einsatz solcher Systeme ist jedoch erst der Anfang. Die größere Herausforderung besteht darin, ihre Leistungsfähigkeit dauerhaft zu sichern – sicherzustellen, dass sie kontinuierlich Wert schaffen, mit den Unternehmenszielen im Einklang bleiben und sich im Zusammenspiel mit dem Mensch-Maschine-Ökosystem weiterentwickeln.[56] Entscheidend ist, mit vertrauten Geschäftsprozessen zu beginnen, wie im folgenden Kapitel erläutert wird. Am besten wird ein standardbasierter Ansatz gewählt, der damit beginnt, klare, einfache Aufgaben für die ersten Agenten zu identifizieren. Diese Vorgehensweise legt die Grundlage für ein künftig komplexeres, orchestriertes Modell.

Der folgende Beitrag über den Europäischen Gerichtshof für Menschenrechte zeigt, wie EIM zur Konsolidierung von Informationen beiträgt, um Genauigkeit sicherzustellen, und wie KI in zentrale Prozesse integriert wird – wodurch der Verwaltungsaufwand sinkt und die Leistungsfähigkeit steigt. So sind die Behörden besser in der Lage, ihren Auftrag zum Schutz der Bürger zu erfüllen.

Fallstudie

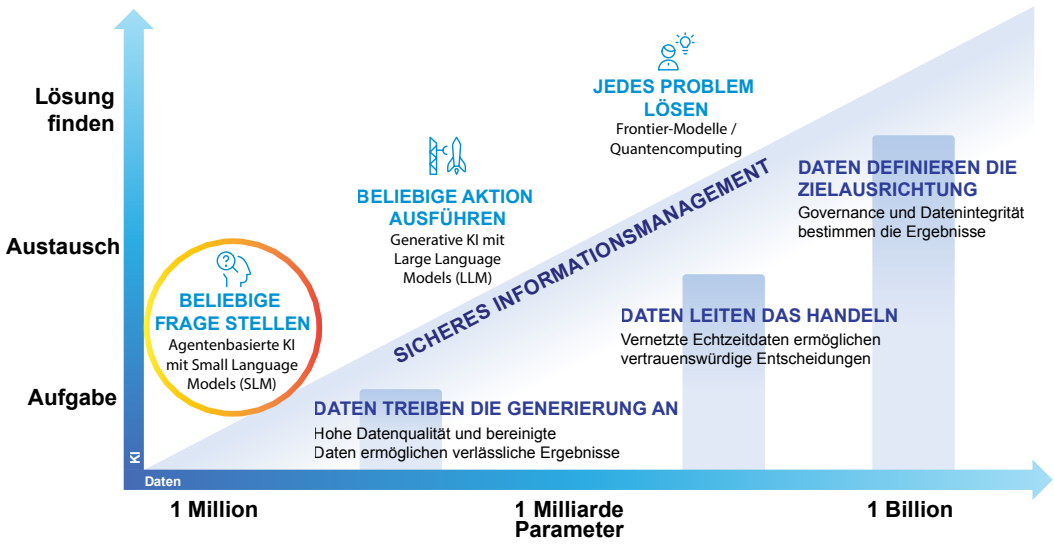
Der Europäische Gerichtshof

Der Gerichtshof ist Teil des Europarats, einer zwischenstaatlichen Organisation, die 1949 gegründet wurde, um Demokratie, Menschenrechte, sozialen Fortschritt und kulturelle Identität in ganz Europa zu fördern.

In den letzten zehn Jahren ist die Zahl der Fälle stark angestiegen – von rund 14.000 auf mehr als 50.000 Anträge. Um diesen Anstieg zu bewältigen, entwickelte die IT-Abteilung des Gerichts eine automatisierte Workflow-Lösung, die den Genehmigungsprozess für Ausschuss- und Kammerverfahren optimierte. Was als Initiative zur digitalen Transformation begann, hat sich inzwischen zu einem intelligenten Informationsökosystem entwickelt, das auf Analytik, EIM und nun auch auf agentenbasierter KI basiert.

Die Arbeitsabläufe des Gerichts gehen heute weit über die reine Weiterleitung von Dokumenten hinaus – sie analysieren, reagieren und handeln selbstständig. Das System stützt sich auf kontrollierte Daten und nutzt Analysen, um Prozessengpässe zu erkennen, während agentenbasierte KI diese in Echtzeit optimiert. Die Plattform kann beispielsweise feststellen, dass eine Fallakte zu lange in der Prüfung verbleibt, sie automatisch zur Eskalation kennzeichnen und die Arbeitslast auf verschiedene Abteilungen verteilen, um den Durchsatz zu sichern. Rechtsanwaltsgehilfen müssen keine Papierunterlagen mehr nachverfolgen. Die KI überwacht den Fortschritt, erstellt dynamische Berichte und empfiehlt nächste Schritte, wobei sie kontinuierlich aus Ergebnissen lernt, um zukünftige Entscheidungen zu verbessern.

Die Ergebnisse sprechen für sich: kürzere Bearbeitungszeiten, weniger administrative Engpässe und mehr Freiraum für die Rechtsexperten des Gerichts, um sich auf die juristische Auslegung statt auf Verwaltung zu konzentrieren. Durch das Zusammenspiel von Analytik, EIM und agentenbasierter KI hat sich der Gerichtshof von einer reaktiven zu einer proaktiven Institution entwickelt – bereit, zu skalieren, sich anzupassen und Gerechtigkeit in dem Tempo zu gewährleisten, das die moderne Arbeitswelt erfordert.



Diese Informationen trainieren alle Formen von KI

Starke Datengrundlagen

Die zentrale Herausforderung besteht darin, sichere und hochwertige Daten zu nutzen, um zuverlässige KI-Ergebnisse zu erzielen und das volle Potenzial der Technologie auszuschöpfen. Wie das oben dargestellte Diagramm zeigt, besteht ein direkter Zusammenhang zwischen einer soliden Datengrundlage und erfolgreichen KI-Ergebnissen.

Neben den Daten muss eine wirksame Strategie auch die technische Infrastruktur berücksichtigen. Da unterschiedliche KI-Modelle unterschiedliche Anforderungen haben, sollte die Daten- und KI-Architektur so aufgebaut sein, dass jeweils das passende Modell für die entsprechende Geschäftsaufgabe eingesetzt wird.

Freigabe vertraulicher Daten zur Unterstützung von agentenbasierte KI (Agentic AI)

Angesichts der Menge und Qualität interner Unternehmensdaten, die sich hinter der Firewall befinden, ist es entscheidend, diese geschützten Informationen gezielt nutzbar zu machen. Die Feinabstimmung oder Anpassung eines LLM mit diesem domänenspezifischen Wissen ist die Voraussetzung dafür, dass agentenbasierte KI sinnvolle, reale Anwendungsfälle im Geschäftsleben und Anwendungen bewältigen kann.

Datenpipelines, Datenherkunft und Datenflüsse werden dabei zu zentralen Faktoren. Jedes Unternehmen muss sich in Richtung eines Data-Warehouse-Unternehmens entwickeln. Bei der Betrachtung des gesamten Datenbestands wird deutlich, dass die Ausschöpfung des KI-Potenzials eine Strategie erfordert, die öffentliche und private Datensätze gleichermaßen einbezieht – beide sind heute wesentliche Bestandteile der Arbeitsweise von Organisationen im privaten wie im öffentlichen Sektor. Das folgende Diagramm veranschaulicht Beispiele für diese Datensätze.

Beispiele für Datensätze

Nutzung privater Daten zur Feinabstimmung von LLMs

Große Sprachmodelle benötigen erhebliche Datenmengen zum Trainieren. Sie lernen aus Mustern in diesen Daten – einschließlich Wörtern, Phrasen, Syntax und semantischen Beziehungen.[57]

Während Qualität und Umfang der anfänglichen Trainingsdaten wichtig sind, ist ihre Relevanz für den konkreten Einsatz im Unternehmen entscheidend. Wie bereits erwähnt, eignen sich öffentlich trainierte LLMs – etwa von OpenAI, Cohere oder Anthropic – hervorragend für allgemeine Aufgaben, doch es fehlt ihnen an tiefem Kontext, wenn es um spezifische Unternehmensanforderungen geht. Um diese Lücke zu schließen, passen Unternehmen ihre Modelle zunehmend an ihre eigenen Daten und Umgebungen an. Die verbreitetste Methode ist die Feinabstimmung, während Verfahren wie Context Engineering zusätzliche Flexibilität und Geschwindigkeit bieten.

Bei der Feinabstimmung wird ein vortrainiertes Basismodell mit einem kleineren, domänenspezifischen oder proprietären Datensatz weitertrainiert. Dadurch kann das Modell das unternehmensspezifische Vokabular, die Daten und Prozesse erlernen – ohne dass diese vertraulichen Informationen der Öffentlichkeit offengelegt werden. Diese Feinabstimmung führt zu einem maßgeschneiderten, abgeleiteten Modell, das bei unternehmens- und domänenspezifischen Aufgaben bessere Ergebnisse erzielt, da es Muster aus privaten Daten verinnerlicht hat.[58]

Parallel dazu ermöglicht Context Engineering Unternehmen, das Verhalten des Modells dynamisch zu verfeinern, indem der Eingabekontext – etwa Eingabeaufforderungen, Beispiele oder Metadaten – strukturiert und laufend angepasst wird, anstatt das Modell neu zu trainieren. Dieser Ansatz bietet eine schnellere und kostengünstigere Anpassung, unterstützt das selektive Verlernen im Hinblick auf Datenschutz oder regulatorische Vorgaben und ist besonders nützlich bei der Arbeit mit proprietären Modellen.[59]

In der Praxis kombinieren Organisationen häufig beide Ansätze: Feinabstimmung für tiefgreifende Domänenanpassung und langfristige Leistungsfähigkeit sowie Context Engineering für agile Echtzeitanpassungen. Zusammen bilden sie eine vielschichtige, nachhaltige Strategie zur Ausrichtung von LLMs an Unternehmenszielen und Compliance-Standards.[60]

Im folgenden Beispiel zeigt ein führender Anbieter von Reisetechnologie für dynamische Pauschalreisen, wie Reisende Zugang zu Millionen von Echtzeit-Kombinationen erhalten. Er nutzt Technologie, um das Reiseerlebnis der Kunden zu vereinfachen, zu personalisieren und zu verbessern.

Fallstudie

Ein Reise-Technologieunternehmen

Da das Unternehmen seine Geschäftstätigkeit auf mehrere europäische Märkte ausgeweitet hatte und die Nachfrage nach Echtzeit-Urlaubspaketen stark zunahm, stieg das Volumen und die Vielfalt der Daten kontinuierlich an. Buchungen, Webinteraktionen, Partnerfeeds, Marketingaktivitäten und Kundensupportanfragen erzeugten Hochgeschwindigkeitsdaten mit unterschiedlichen Strukturen und Latenzanforderungen. Dieses Wachstum beeinträchtigte die Nachverfolgbarkeit von Governance-Initiativen, Datenzugriffsmustern und die Zeit bis zur Erkenntnisgewinnung.

Im Laufe der Zeit teilte sich der Datenbestand in zwei Silos auf: ein Data Warehouse mit ETL- (Extrahieren, Transformieren, Laden) und Reporting-Tools sowie einen Data Lake für die Rohdatenaufnahme. Diese Trennung führte zu doppeltem Aufwand, höherem Wartungsbedarf und langsameren Analyseprozessen. Die Teams arbeiteten mit unterschiedlichen Tools und individuell geschriebenem „Klebstoffcode“, um die Systeme miteinander zu verbinden – was eine deutliche Kluft schuf, die das Wachstum hemmte und die Erkenntnisgewinnung verzögerte.

Das Unternehmen implementierte eine einheitliche Datenzugriffsschicht, wodurch separate Datenaufnahmepipelines und der Einsatz komplexer Verbindungscodes überflüssig wurden. Mit einer gemeinsamen Schnittstelle und einem standardisierten Tool Set konnten Ingenieure und Analysten Daten konsistent über alle Systeme hinweg abfragen und verarbeiten. Durch die Konsolidierung der Umgebungen und die Entkopplung von Datenproduzenten und -Konsumenten – bei gleichzeitiger Wahrung vollständiger Kompatibilität – entstand eine zentrale, verlässliche Datenquelle. Diese Integration ermöglichte umfassendere Analysen, indem BI-Daten mit Erkenntnissen aus Marketing, CRM und Machine Learning verknüpft wurden, um fortschrittliche, prädiktive Kundenmodelle zu entwickeln.

Anschließend migrierte das Unternehmen zu einer containerisierten Cluster-Umgebung auf Basis von Kubernetes, um Bereitstellung, Skalierung und Workload-Management zu automatisieren. Abfrageaufträge können nun bedarfsgesteuert gestartet und nach Abschluss wieder beendet werden, wobei gemeinsam genutzter Speicher und für jede Aufgabe optimal dimensionierte Rechenleistung genutzt werden – von ETL-Prozessen bis zu Analyse-Dashboards. Das Ergebnis: höhere Skalierbarkeit und Effizienz bei geringeren Rechenkosten und niedrigerem Energieverbrauch, was auch den CO_2-Fußabdruck reduziert.

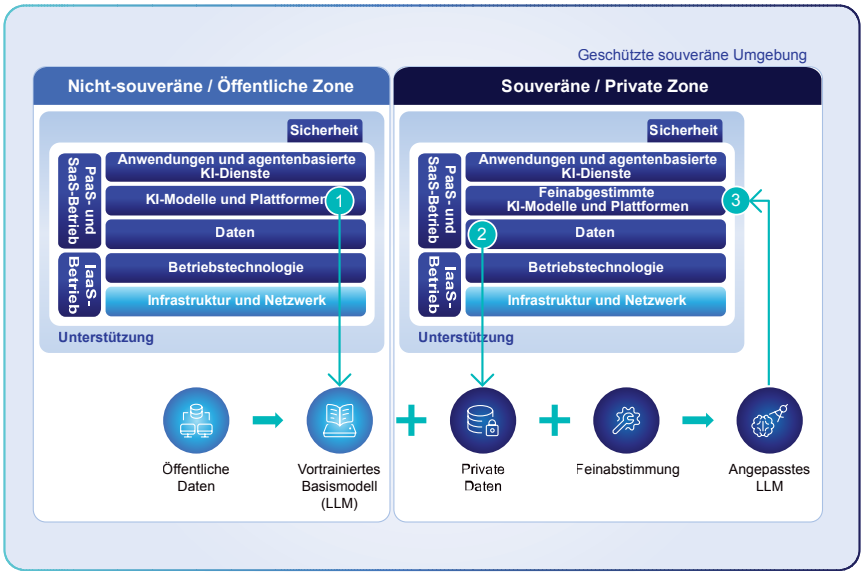
Das Unternehmen kann das Kundenverhalten über alle Kanäle hinweg analysieren und jede Phase der Customer Journey – von der ersten Suche bis zur finalen Zahlung – nachvollziehen. Die Implementierung verbessert sowohl den Return on Investment (ROI) der Marketingkampagnen als auch den Gesamtumsatz. Zudem tragen KI-gestützte Algorithmen zur Attributions- und Gebotsautomatisierung bei, wodurch die Marketingausgaben optimiert und die Gewinnmargen erhöht werden.

Doch kann die Souveränität privater Daten und feinabgestimmter Modelle tatsächlich garantiert werden?

Genau hier liegt die zentrale Herausforderung. Daten – und die daraus entwickelte KI – sind der entscheidende Wettbewerbsvorteil eines Unternehmens. Die Qualität dieser Daten bestimmt unmittelbar die Effektivität und Integrität eines LLM. Sobald ein Modell jedoch mit einem bestimmten Datensatz trainiert wurde, kann es diese Informationen nicht einfach wieder „verlernen".

Das ist einer der Hauptgründe, warum private Daten und private KI unverzichtbar sind. Während des Trainings verinnerlicht das Modell die zugrunde liegenden Muster und hinterlässt damit eine unauslöschliche Spur. Diese internen Repräsentationen werden mit anderen Daten verknüpft – sowohl mit konkreten Beispielen als auch mit der allgemeinen Datenverteilung.
Aus diesem Grund erfordert der Versuch, bestimmte Trainingsdaten zu verlernen, eine aufwendige Bereinigung der Parameter und Verknüpfungen, die das Verhalten des Modells steuern. In „Machine Unlearning Doesn't Do What You Think" (Das Verlernen von Maschinen funktioniert nicht so, wie Sie es sich vorstellen), betonen die Autoren: „Das Löschen von Informationen aus einem ML-Modell [maschinelles Lernen] ist nicht gut definiert. Erstens können Informationen aus einem ML-Modell nicht auf die gleiche Weise gelöscht werden wie aus einer Datenbank."[61]

Auch wenn weiter an Methoden zum „maschinellen Verlernen" geforscht wird, stellt dies für Unternehmen des privaten wie auch des öffentlichen Sektors keinen praktikablen oder einfachen Weg dar. Der Schutz souveräner und privater Daten sowie der darauf aufbauenden Modelle hat daher höchste Priorität.



Souveräne Daten- und KI-Architektur

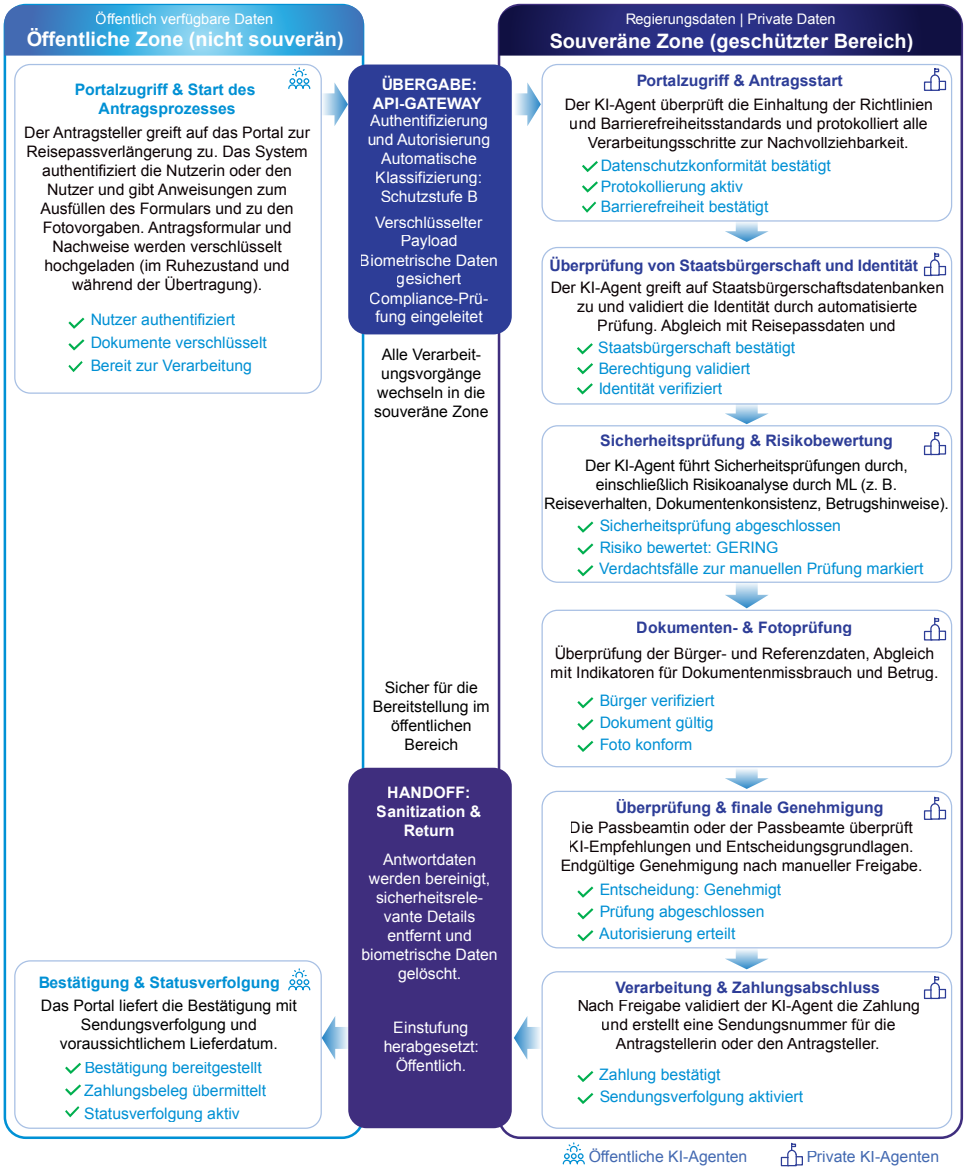
Genau deshalb ist eine souveräne Daten- und KI-Architektur von zentraler Bedeutung. Anstatt auf das Verlernen zu setzen, beruht dieser Ansatz auf Prävention. Er ist darauf ausgelegt, Daten und feinabgestimmte Modelle von Beginn an privat zu halten. Dazu muss das Modell in einer zertifizierten souveränen Zone betrieben werden, in der Ihre feinabgestimmten Modelle geschützt sind – wie in der Abbildung oben dargestellt.

In der obigen Abbildung werden öffentliche Daten genutzt, um ein vortrainiertes Basis-LLM (1) bereitzustellen, das im nicht-souveränen/öffentlichen Bereich operiert. In dieser Umgebung kann ein vollständiger Satz agentenbasierter KI-Funktionen bereitgestellt werden. Auf der Seite der privaten Zone wird das Basismodell mit privaten Daten (2) feinabgestimmt. Das Ergebnis ist ein maßgeschneidertes LLM (3) mit domänenspezifischem Wissen, das ein einzigartiges Unterscheidungsmerkmal für das Unternehmen darstellen kann. Dieses LLM (3) wird ausschließlich von der agentenbasierten KI innerhalb der privaten Zone genutzt.
Eine Rückübertragung von Daten an das LLM in der öffentlichen Zone findet nicht statt.

Wird die Infrastruktur vollständig kontrolliert, kann das Modell vertraulich behandelt werden. Bei Hosting in einer öffentlichen Cloud sind dagegen verbindliche Zusicherungen sowohl vom Cloud-Anbieter als auch vom Modellanbieter erforderlich – insbesondere im Hinblick auf die Datenhoheit und Souveränität.

Agentenbasierte KI im souveränen Kontext: Ein Anwendungsfall

Wenn agentenbasierte KI in einen souveränen Kontext eingeordnet wird, lässt sich dies gut anhand eines Beispiels verdeutlichen, das Bürgerinnen und Bürger vieler Länder kennen: die Beantragung eines neuen Reisepasses. Dieses Beispiel zeigt, dass sowohl private als auch öffentliche Datensätze erforderlich sind – und wie agentenbasierte KI in beiden Umgebungen agieren kann. Feinabgestimmte Modelle im souveränen (privaten) Bereich arbeiten dabei mit Modellen im nicht-souveränen (öffentlichen) Bereich zusammen. Wenn beide Umgebungen reibungslos ineinandergreifen, profitiert letztlich der Endkunde am meisten und erlebt ein deutlich verbessertes Nutzererlebnis.



Anwendungsfall: Passbearbeitung mit Bezug auf souveräne und nicht-souveräne Zonen

Schritt 1: Portalzugang und Anwendungsstart (Öffentlicher Bereich)
Der Antragsteller greift über ein Portal auf den Dienst zur Passerneuerung zu. Das System authentifiziert die Person, stellt Anleitungen zum Ausfüllen bereit und ermöglicht das sichere Absenden des Antrags. Alle Passdaten und Begleitdokumente werden sowohl im Ruhezustand als auch während der Übertragung verschlüsselt. In dieser Phase umfasst der Prozess Statusmeldungen zu Benutzerauthentifizierung, Dokumentenverschlüsselung und Verarbeitungsbereitschaft.

Schritt 2: Übergabe – API-Gateway
Das API-Gateway übernimmt Authentifizierung und Autorisierung und klassifiziert die Daten automatisch als geschützt. Die Nutzdaten werden verschlüsselt, die biometrischen Daten gesichert und die Konformitätsprüfung eingeleitet.

Schritt 3: Konformitäts- und Richtlinienvalidierung (Souveräne Zone)
Der KI-Agent der Regierung überprüft die Einhaltung aller relevanten Datenschutz- und Informationsvorschriften und -richtlinien.

Das System bestätigt die Sprachdienstleistungsfähigkeit sowie die Barrierefreiheitsstandards und protokolliert sämtliche Verarbeitungsschritte zu Prüf- und Compliance-Zwecken. Zu diesem Zeitpunkt umfasst die Statusprüfung die Konformitätsvalidierung, die aktive Audit-Protokollierung und die bestätigte Zugänglichkeit.

Schritt 4: Staatsbürgerschafts- und Identitätsprüfung (Souveräne Zone)
Der KI-Agent der Regierung greift auf die Staatsbürgerschaftsdatenbank zu, um die im Passantrag angegebenen Identitätsdaten abzugleichen. Er vergleicht bestehende Passdatensätze, um die Berechtigung anhand von Staatsangehörigkeit und weiteren Kriterien zu bestätigen. Die Statusmeldungen umfassen die bestätigte Staatsbürgerschaft, die Option einer manuellen Korrektur bei Bedarf sowie die verifizierte Berechtigung.

Schritt 5: Sicherheitsüberprüfung und Risikobewertung (Souveräne Zone)
Der KI-Agent der Regierung führt Sicherheitsüberprüfungen unter Verwendung von Datenbanken der Strafverfolgungsbehörden durch und berücksichtigt dabei Reisemuster sowie Risikofaktoren. Maschinelle Lernmodelle bewerten das Antragsrisiko anhand dieser Muster und gleichen es mit Warnlisten ab. Falls erforderlich, wird eine manuelle Überprüfung eingeleitet. Die Statusmeldungen umfassen eine abgeschlossene Sicherheitsprüfung und eine Risikobewertung, die als niedrig, mittel oder hoch eingestuft wird.

Schritt 6: Dokumenten- und Fotovalidierung (Souveräne Zone)
Der KI-Agent der Regierung validiert die Angaben des Bürgers durch einen Datenbankabgleich, überprüft die Passhistorie auf Unregelmäßigkeiten oder Anzeichen von Reisebetrug und nutzt Modelle des maschinellen Lernens, um das eingereichte Foto anhand staatlicher Ausweisdatensätze und Behördendatenbanken zu authentifizieren. Die Statusmeldungen umfassen die Validierung der Bürgernachweise sowie die Bestätigung der Dokumentenintegrität.

Schritt 7: Überprüfung durch Beamte und endgültige Genehmigung (Souveräne Zone)
Ein menschlicher Experte überprüft alle vom System markierten Informationen und nutzt dabei KI als Entscheidungshilfe. Der menschliche Bearbeiter erteilt die endgültige Genehmigung auf Grundlage einer vollständigen Prüfung aller validierten Informationen. Dieser Prozess stellt sicher, dass sensible personenbezogene Daten sicher behandelt, überprüft und in Übereinstimmung mit Datenschutzvorschriften verarbeitet werden, wobei eine klare Trennung zwischen öffentlichen und souveränen Bereichen gewahrt bleibt.

Auch wenn die konkreten Anwendungsfälle variieren können und es sich hier um ein vereinfachtes Beispiel handelt, verdeutlicht es, wie agentenbasierte KI in souveränen und nicht-souveränen Zonen interagiert und dabei personenbezogene Daten innerhalb eines hybriden Systems verarbeitet. Die Implementierung und Verwaltung des Einsatzes agentenbasierter KI erfordert eine umfassende Betrachtung der Zusammenarbeit zwischen digitalen Agenten und menschlichen Arbeitskräften.
Agentenbasierte KI führt Aufgaben auf Basis ihrer integrierten Logik aus, während die menschliche Aufsicht die Qualität und Verlässlichkeit der Ergebnisse sicherstellt.

Im nächsten Kapitel wird untersucht, wie digitale und menschliche Arbeitskräfte zusammenwirken müssen, um das Potenzial der KI im Unternehmen voll auszuschöpfen.

Der folgende Beitrag zeigt, wie der Generalrat der Justiz ein vertrauenswürdiges EIM-System einsetzt, um öffentliche und private Informationen sicher zu konsolidieren – und so die Dienstleistungen für Bürgerinnen und Bürger in Spanien verbessert.

Fallstudie

Allgemeiner Justizrat

PoderJudicial.es

> *Analysen helfen uns, den Erfolg unserer Dienste und die Gesamtleistung der öffentlichen Website zu messen und statten uns mit den Werkzeugen aus, die wir benötigen, um den Nutzern ein relevantes und responsives Erlebnis zu bieten, das durch Multimedia-Inhalte unterstützt wird.*

Der Generalrat der Justiz (Consejo General del Poder Judicial, CGPJ) wurde 1978 durch die spanische Verfassung als oberstes Verwaltungsorgan der Justiz eingerichtet. Die CGPJ wollte ihre Systeme in einem Online-Portal zusammenfassen, um den Bürgern einen personalisierten Zugang zu den von ihnen benötigten Informationen und Dienstleistungen zu bieten.

Das neue Portal sollte mehrsprachig und über verschiedene Kommunikationskanäle nutzbar sein. Im Backend musste es sämtliche Unternehmensdienste des Justizrats integrieren, um die Zusammenarbeit zu verbessern, integrierte Online-Anwendungen bereitzustellen, eine sichere Informationsverwaltung zu gewährleisten und die gesetzlichen Anforderungen an Transparenz, Zugänglichkeit, Mehrsprachigkeit und das Gesetz 11/2007 zu erfüllen.

Als Grundlage für die Website und das Justiz-Extranet entschied man sich für eine E-Government-Lösung, die dem CGPJ eine technologisch stabile und zukunftssichere Plattform bietet. Das mehrsprachige Portal unterstützt eine große Zahl gleichzeitiger Zugriffe und lässt sich problemlos skalieren. Durch die neuen Self-Service-Funktionen wurde der Web-Veröffentlichungsprozess deutlich effizienter, und die Zeit bis zur Veröffentlichung aktueller Inhalte konnte erheblich verkürzt werden.

Das System wurde intern mit 6.500 aktiven Nutzern und 5.400 ausgetauschten Nachrichten in den Foren eingeführt. Mitglieder der Justiz können über virtuelle Arbeitsumgebungen, 45 Fachgemeinschaften und gemeinsam genutzte Dateien aktiv teilnehmen und zusammenarbeiten. Ein sicherer Zugriff auf alle integrierten Anwendungen und Dienste wird durch Single Sign-On und ein zentrales Identitätsmanagement gewährleistet. Darüber hinaus ist das System flexibel anpassbar, sodass Benutzer ihre Arbeitsumgebung individuell personalisieren und konfigurieren können.

Die fünf Merksätze

1. **Agentenbasierte KI einsetzen, um den Unternehmenswert zu steigern.**
   Steigern der Produktivität und Anpassungsfähigkeit durch den Einsatz agentenbasierter KI-Anwendungen, die autonom wahrnehmen, planen, entscheiden und handeln – so werden komplexe Arbeitsabläufe automatisiert und die Abhängigkeit von manuellen Eingriffen verringert.

2. **Private, domänenspezifische Daten zur Differenzierung nutzen.**
   Wettbewerbsvorteile werden erzielt, indem KI-Modelle mit internen Unternehmensdaten feinabgestimmt werden. Dadurch kann agentenbasierte KI domänenspezifische Herausforderungen lösen, die generische Modelle nicht bewältigen, und auf einer sicheren, hochwertigen Datengrundlage operieren.

3. **Souveräne KI-Architekturen implementieren, um geistiges Eigentum und Privatsphäre zu schützen.**
   Sensible Daten und geistigen Eigentums sollten geschützt werden, indem KI-Modelle in souveränen Umgebungen eingesetzt werden – für Compliance, Datenschutz und volle Kontrolle über die eigenen KI-Ressourcen.

4. **KI-Initiativen auf zielgerichtete, lernfähige Systeme ausrichten.**
   Der ROI wird maximiert, indem agentenbasierte KI auf klar definierte Geschäftsprozesse ausgerichtet wird und in lernende, adaptive Systeme investiert wird – zur Vermeidung generischer Lösungen und disruptiver Veränderungen.

5. **Mensch-KI-Kollaboration für nachhaltige Wirkung fördern.**
   Nachhaltiger Mehrwert kann sichert werden, indem Standards für menschliche Aufsicht und Abstimmung etabliert werden, beginnend mit einfachen Handlungsaufgaben. So entsteht eine effektive Zusammenarbeit zwischen Mensch und Maschine, die sich kontinuierlich mit den Geschäftsanforderungen weiterentwickelt.

Kapitel Neun

Das Management von EAI-Anwendungens

In modernen Unternehmen geht es bei der Verwaltung von KI-Anwendungen nicht mehr um die Bereitstellung eines einzelnen Modells, sondern um die Orchestrierung des Zusammenspiels intelligenter Agenten, die in öffentlichen und souveränen Zonen, privaten Clouds und offenen Netzwerken sowie in einer Vielzahl von Arbeitsabläufen agieren. Das ist Unternehmens-KI (Enterprise AI, EAI), und um mit dieser Komplexität umzugehen, bedarf es mehr als nur technischer Fähigkeiten. Es erfordert einen ausgeklügelten organisatorischen Ansatz, der auf soliden Geschäftsprozessen, Änderungsmanagement und Governance basiert. Wie aktuelle Forschungsergebnisse zeigen, scheitern Unternehmen bei der Implementierung von KI selten an den Modellen selbst, sondern an unzureichend vorbereiteten Strukturen, Rollen und Abläufen.

Dieses Kapitel stellt die zentralen Prinzipien vor, die gewährleisten, dass eine Organisation nicht nur agentenbasierte KI – Systeme, die denken, planen, handeln und zusammenarbeiten – einsetzen kann, sondern sie auch so steuert, dass sie im Einklang mit strategischen Zielen, Risikorahmen und einer menschenzentrierten Governance steht.

Folgende vier Prinzipien schaffen die Voraussetzungen für den erfolgreichen Einsatz agentenbasierter KI-Systeme im Unternehmen:

1. **Organisationsmodell für agentenbasierte EAI-Implementierungen**: Sicherstellen, dass Eigentumsverhältnisse und Verantwortlichkeiten sowie Rollen und Zuständigkeiten klar definiert sind.

2. **Entwicklung agentenbasierter EAI-Anwendungen**: Einführung eines strukturierten Ansatzes zur Priorisierung und zum Aufbau von Fähigkeiten für die Organisation.

3. **Zusammenarbeit zwischen menschlichen und agentenbasierten EAI-Teams**: Sicherstellen, dass die Teams KI mit klaren Funktionen und Zielen annehmen und einsetzen.

4. **Leistungsmanagement und -messung**: Schließen der Feedback-Schleife durch Messung der Ergebnisse und Leistungssteuerung der agentenbasierten KI-Belegschaft.

Im Folgenden werden für jedes dieser Prinzipien Rahmenbedingungen, bewährte Praktiken und praxisnahe Überlegungen vorgestellt, um skalierbare Enterprise KI-Systeme erfolgreich zu gestalten und zu betreiben. Von der Definition des Organisationsmodells über die Gestaltung gemeinsamer Arbeitsabläufe zwischen Menschen und Agenten bis hin zur Wertmessung und kontinuierlichen Anpassung – dieses Kapitel bereitet auf den Übergang von isolierten KI-Experimenten zu unternehmensweiten Implementierungen vor. Zunächst wird untersucht, wie das Organisationsmodell die Grundlage für verantwortungsvolle und skalierbare KI-Anwendungen schafft.

1. Ein Organisationsmodell für den Einsatz agentenbasierter KI

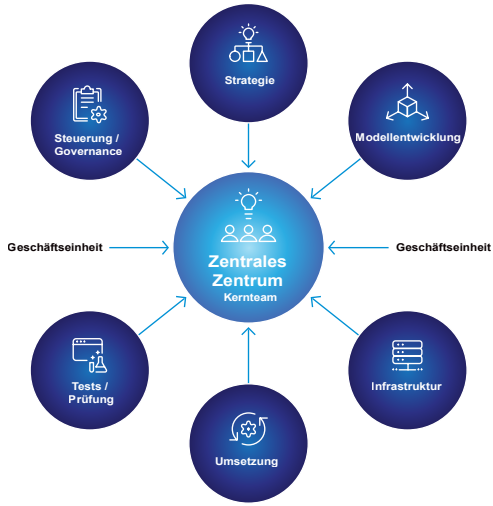
Wie das richtige Modell ausgewählt wird

Unternehmen, die den Einsatz von KI vorantreiben und insbesondere agentenbasierte KI-Funktionen implementieren wollen, müssen entscheiden, wie sie ihre KI-Anwendungen verwalten möchten. Es existieren mehrere gängige Modelle, die jeweils spezifische Vor- und Nachteile bieten. Entscheidend ist die Auswahl des Modells, das zur Organisationsstruktur, den Zielen und dem Regulierungsumfeld passt.

Die folgenden Beispiele zeigen vier Ansätze, die sich je nach Branche und Regulierungsgrad bewährt haben: Zentralisiertes Modell (AI Center of Excellence/CoE), Hub-and-Spoke-Modell, föderiertes Modell und Hybridmodell.

Das zentralisierte Modell

In diesem Modell verfügt ein zentralisiertes Team über umfassende Kompetenzen und ist für Strategie, Modellentwicklung, Infrastruktur, Implementierung, Tests und Governance verantwortlich. Dies gewährleistet Konsistenz und Kontrolle sowie eine einheitliche Governance, kann aber auch dazu führen, dass die Geschäftsbereiche nicht mehr für die Ergebnisse einer KI-Einführung verantwortlich sind. Viele Organisationen, die erst seit Kurzem mit neuen Technologien arbeiten, sowie Unternehmen in regulierten Branchen oder im öffentlichen Sektor bevorzugen dieses Modell.

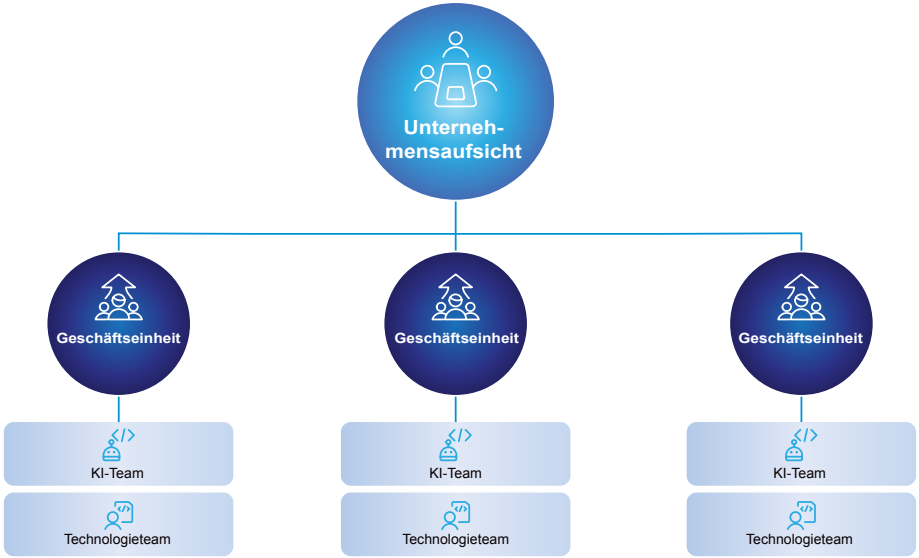


Das zentralisierte Modell

Das Hub-and-Spoke-Modell

In diesem Modell definiert ein zentrales Team die Strategie und stellt Tools, Richtlinien und Rahmenbedingungen bereit, während die einzelnen Geschäftsbereiche als operative Knotenpunkte fungieren, die Projekte in ihrem Verantwortungsbereich umsetzen. Dieses Modell ist von Natur aus agiler und eignet sich besonders für Organisationen, die bereits über technologische Kompetenz in verschiedenen Geschäftsbereichen verfügen. Es ermöglicht zudem, mehrere Projekte gleichzeitig und parallel voranzutreiben. Ein wiederkehrendes Diskussionsthema betrifft die Zuständigkeit für das LLM – insbesondere, wem es gehört, wer es trainiert und wer für die Feinabstimmung verantwortlich ist. In der Regel liegt diese Verantwortung beim zentralen Hub, während die Geschäftsbereiche die agentenbasierten KI-Anwendungen betreuen.

Das Naben-Speichen-Modell



Das föderierte Modell

Das föderierte Modell

In diesem Modell existiert keine zentrale Funktion: Jede Geschäftseinheit verfügt über ein eigenes Team, das für den Betrieb und die Verwaltung ihrer KI-Systeme und -Technologien verantwortlich ist. In einigen Fällen übernimmt ein kleines zentrales Team eine übergeordnete Aufsichtsfunktion. Grundsätzlich bietet dieses Modell den Geschäftsbereichen maximale Kontrolle und ermöglicht eine schnellere Umsetzung von Projekten – jedoch auf Kosten einer konsistenten Governance und eines höheren Sicherheitsrisikos. Es eignet sich daher vor allem für Organisationen mit einem hohen technologischen Reifegrad.

Hybridmodell

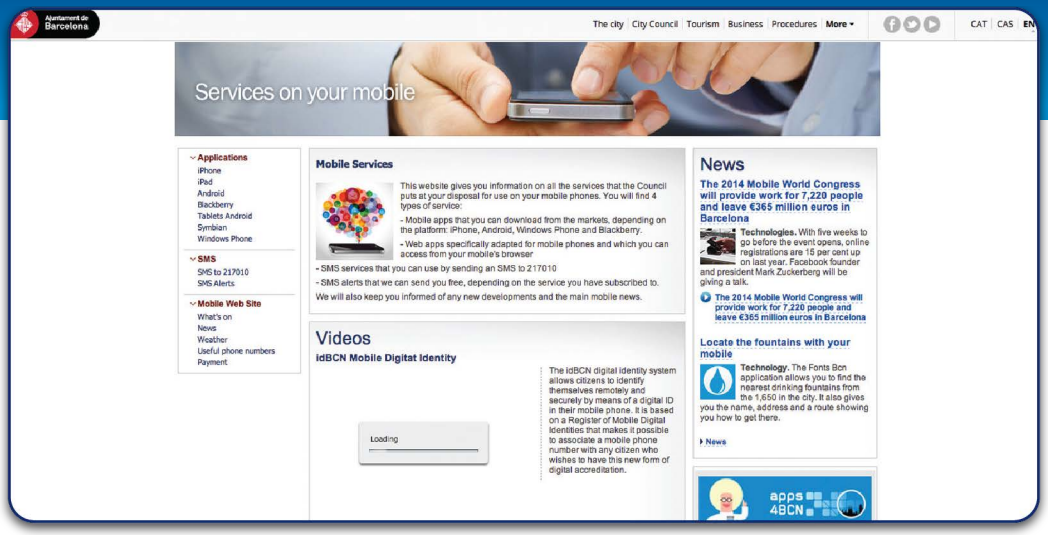
Hybridmodelle kombinieren Elemente der anderen Ansätze. Ihr größter Vorteil liegt in der Möglichkeit, Skalierungseffekte zu nutzen, indem standardisierte Services zentral bereitgestellt werden, während die Geschäftsbereiche gleichzeitig die notwendige Autonomie für ihre operativen Abläufe behalten.

Die Wahl des passenden Modells hängt letztlich von den spezifischen Anforderungen und der Struktur der jeweiligen Organisation ab. Erfolg setzt voraus, dass klar definiert ist, wie die Teams zusammenarbeiten sollen und welche Leitlinien für Benutzer und Verantwortliche gelten. Diese Klarheit ist entscheidend, um skalierbare, ergebnisorientierte Anwendungen agentenbasierter KI aufzubauen.

Im folgenden Fallbeispiel zeigt die Stadt Barcelona, wie ein Hub-and-Spoke-Modell eingesetzt werden kann, um Daten und Dienste über jedes Gerät und von jedem Ort aus zugänglich zu machen – mit dem Ziel, den Bürgerservice und die Lebensqualität nachhaltig zu verbessern.

Fallstudie

Stadt Barcelona



Die Stadt Barcelona – Bürgerservices sind „nur einen Klick entfernt“

Mit über anderthalb Millionen Einwohnern ist Barcelona die zweitgrößte Stadt Spaniens. Um ihre Vision einer intelligenten Stadt zu verwirklichen, setzt die Stadtverwaltung auf mobile und cloudbasierte E-Government-Lösungen, die den Bürgerinnen und Bürgern eine einfachere Beteiligung an Verwaltungsprozessen und öffentlichen Dienstleistungen ermöglichen.

Ziel der Einführung des E-Government-Systems war es, Daten und Services für alle Einwohner jederzeit und von jedem Gerät aus zugänglich zu machen, um dadurch die Lebensqualität spürbar zu verbessern. Ein erster Schritt in diese Richtung war die digitale Bereitstellung von Daten des Stadtrats und anderen Stellen, was gleichzeitig dazu beitrug, die Wiederverwendung dieser Daten zu fördern, um durch Innovationsmöglichkeiten das Wirtschaftswachstum anzukurbeln.

Um ihre Informationssysteme zu standardisieren, konsolidierte die Stadtverwaltung ihre Infrastruktur auf der Grundlage interoperabler, offener Normen und ersetzte schrittweise ihre veralteten Systeme. Die Stadt beschloss, ihre Lösungen in die Cloud zu verlagern. Ein cloudbasiertes Content-Management-System bot eine zuverlässige und flexible Alternative, die langfristig auch wirtschaftliche Vorteile mit sich bringt. Das Ergebnis war die erste Barcelona Open Data-Website mit 510 Datensätzen. Die Lösung folgt den Grundprinzipien Mobilität, Smart City, Verwaltung, Informationssysteme und Innovation und unterstützt inzwischen rund 150 Portale mit über 4 Millionen Nutzern und mehr als 65 Millionen Seitenaufrufen pro Monat.

Ist eine Organisation bereit?

Getrieben von der Angst, den Anschluss zu verlieren, löste der KI-Hype einen Wettlauf um die Einführung der Technologie aus. Die wirklich innovativen Organisationen aber gaben nicht Vollgas, sondern bremsten. Sie erkannten, dass obwohl die Daten der Treibstoff für die KI-Maschine sind, derjenige, der ohne klare Richtung startet, riskiert, dass die KI ausbrennt, bevor das Ziel erreicht ist.

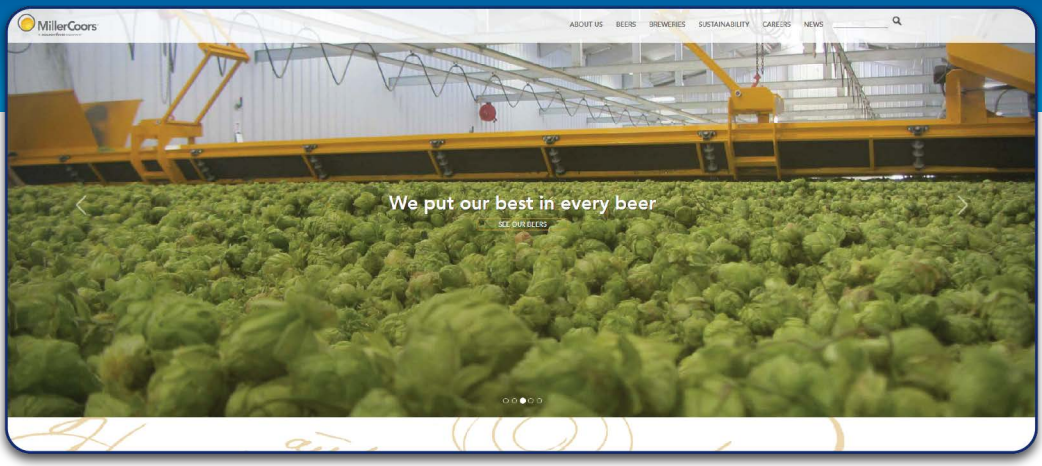
Diejenigen, die sich Zeit nahmen, um zu lernen, zu planen und sich vorzubereiten, erzielten deutlich größere Erfolge. Indem sie sicherstellten, dass ihre Daten aufbereitet und die Governance-Strukturen vorhanden waren, legten sie den Grundstein für den Erfolg. Sie wussten genau, wie sie KI einsetzen wollten, bevor sie das Tempo erhöhten – ein entscheidender Schritt, den viele übergingen, die zu schnell starteten.

Selbst Microsoft drosselte das Tempo bei der unternehmensweiten Einführung von Copilot. Als eines der ersten Unternehmen, das KI in großem Umfang implementierte, gliederte Microsoft die Einführung in mehrere Phasen – beginnend mit einer begrenzten Pilotphase für ausgewählte Gruppen, über eine gezielte Einführung für bestimmte Teams, bis hin zur vollständigen Implementierung nach Benutzergruppen. Microsoft erklärte dazu: „Wir haben unsere Einführung entlang zweier Vektoren strukturiert – nach internen Organisationen, etwa Rechtsabteilung oder Vertrieb und Marketing, sowie nach Regionen, wie Nordamerika oder Europa.“ Verschiedene Benutzergruppen haben unterschiedliche Schwerpunkte, aber die Strategie ist ähnlich.“[62] Dieser Gruppen-basierte Ansatz wurde von vielen Organisationen als entscheidender Erfolgsfaktor bei der Einführung von KI genannt. Er ermöglichte es, bestimmten Gruppen und Nutzern genau die Technologien bereitzustellen, die ihren jeweiligen Anforderungen entsprachen – und so die Akzeptanz im gesamten Unternehmen zu fördern.

Im folgenden Fallbeispiel übernimmt MillerCoors eine zentrale Rolle in seiner Lieferkette, indem das Unternehmen seine Zulieferer direkt überwacht und dadurch Effizienz und Transparenz entlang der gesamten Wertschöpfungskette steigert.

Fallstudie

MillerCoors



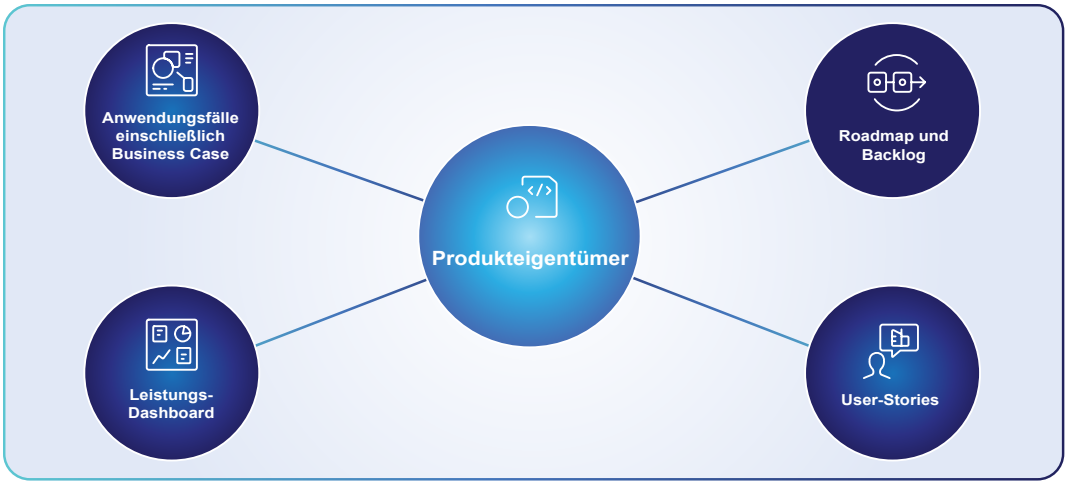
MillerCoors

MillerCoors ist ein Joint Venture der US-amerikanischen Niederlassungen von SABMiller und Molson Coors. Mit mehr als 450 Jahren gemeinsamer Brautradition verfügt das Unternehmen über ein beeindruckendes Portfolio führender Biermarken. Mit einem Marktanteil von fast 30 % am US-amerikanischen Bierabsatz ist MillerCoors das zweitgrößte Brauereiunternehmen der Vereinigten Staaten und betreibt acht Großbrauereien sowie mehrere Craft-Brauereien.

Die frühere Miller Brewing Company, ein Vorgängerunternehmen von MillerCoors, stellte fest, dass ihre Lieferkette – vom Großhändler bis zum Einzelhändler – weder den Branchenstandards noch den Erwartungen der Verbraucher entsprach. Das Unternehmen musste seine ineffizienten, papierbasierten Prozesse modernisieren und standardisieren, um in einem zunehmend komplexen und verbraucherorientierten Markt wettbewerbsfähig zu bleiben.

Durch den Einsatz von B2B Managed Services gelang es Miller Brewing, mehr als 400 Vertriebspartner und 25 unterschiedliche Geschäftssysteme in einer einheitlichen EDI-Plattform (Electronic Data Interchange) zu integrieren. So konnte das gesamte Vertriebsnetz mit allen Einzelhändlern zusammenarbeiten, die eine EDI-Fähigkeit benötigten.

Die B2B Managed Services bilden heute das technische Rückgrat für eine nahtlose, durchgängige EDI-Plattform, die sämtliche Lieferanten- und Bankverbindungen von MillerCoors unterstützt. Wichtige Dokumente gehen ein, werden verarbeitet und nahtlos ausgetauscht – was zu Effizienzsteigerungen, Kosteneinsparungen und, ganz nebenbei, zu mehr Bierumsatz führt. Bereits im ersten Jahr nach der Umstrukturierung der Geschäftsprozesse konnten 1,2 Millionen Arbeitsstunden bei den Distributoren und 1,3 Millionen Stunden bei den Einzelhändlern eingespart werden – insgesamt also 2,5 Millionen Stunden entlang der gesamten Lieferkette vom Distributor bis zum Point of Sale. Diese Zeitersparnis entspricht einer Umverteilung von rund 1.200 Vollzeitäquivalenten (VZÄ) auf andere Tätigkeiten und führt zu einer geschätzten Arbeitskosteneinsparung von etwa 50 Millionen US-Dollar.



Breiter Aufgabenbereich des Produktverantwortlichen

2. Entwicklung agentenbasierter EAI-Anwendungen

Klein anfangen und es einfach halten

Über einen Benutzergruppen-spezifischen Bereitstellungsansatz hinaus ist es ein zentrales Prinzip erfolgreicher KI-Einführungen, klein anzufangen und es einfach zu halten. Dahinter steht ein rücksichtsloser Fokus darauf, Geschäftsergebnissen zu erzielen. Technologie, um ihrer selbst willen zu entwickeln, mag für IT-Teams spannend sein, führt im Umgang mit KI jedoch selten zum Erfolg.

Klein anzufangen bedeutet, einen klar abgegrenzten Anwendungsfall zu wählen und mit einem Unternehmensbereich zusammenzuarbeiten, der die Technologie und ihr Potenzial versteht, über gut dokumentierte Geschäftsprozesse und hohe Datenqualität verfügt und die Akzeptanz fördern kann. Die Bedeutung von Veränderungsagenten und Fürsprechern, die diesen Wandel unterstützen, darf nicht unterschätzt werden. In unserem Fall fiel die Wahl auf das Personalteam. Da das Team eine starke Technologieorientierung aufwies und bereit war, die Veränderung aktiv mitzugestalten und Verantwortung sowohl auf der Business- als auch auf der Produktebene zu übernehmen.*

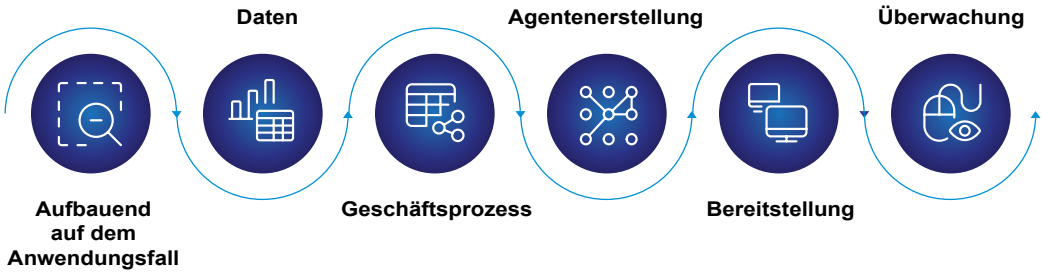
Dabei spielen Produktverantwortliche, auch als Product Owner bezeichnet, eine zentrale Rolle, da sie für die Konzeption und Umsetzung spezifischer KI-Anwendungsfälle verantwortlich sind. Wesentlich ist, dass sie nicht nur über die nötigen Fach- und Prozesskenntnisse verfügen, sondern auch als Veränderungsagenten und Befürworter agentenbasierter KI-Einführungen wirken können. Die Rolle des Produktverantwortlichen umfasst im Wesentlichen die oben gezeigten Aufgaben.

* Zur Information: Das Scaled Agile Rahmenwerk definiert den Produktverantwortlichen als „die Stimme des Kunden und des Unternehmens, die das Team-Backlog verwaltet und priorisiert, die Arbeit des Teams an der Strategie und den Bedürfnissen der Stakeholder ausrichtet und dazu beiträgt, die geschäftliche und technische Integrität der Lösung zu wahren."[63]

Die Bedeutung guter Prozesse und Daten für die Entwicklung agentenbasierter KI-Anwendungen

Sobald ein Team und ein Zielanwendungsfall festgelegt sind, wird ein tiefes Verständnis der Daten und Geschäftsprozesse zum entscheidenden Faktor für die Steuerung der Entwicklung und des Verhaltens eines Agenten. Der folgende Abschnitt beschreibt diesen Ablauf im Detail.

Lebenszyklusmanagement-Ansatz für agentenbasierte KI



Der Lebenszyklus agentenbasierter KI

Ausgehend vom definierten **Anwendungsfall** beginnt der Prozess mit einer Analyse **der Daten** und **Geschäftsprozesse.** Diese Analyse fließt direkt in die Entwicklung der agentenbasierten KI-Anwendung ein.

Diese Entwicklungsphase, die oft vom Produktverantwortlichen gesteuert wird, ist typischerweise eine Zusammenarbeit zwischen der IT-Abteilung und der jeweiligen Geschäftseinheit. Um die Einstiegshürde zu senken, verwenden viele agentenbasierte Rahmenwerke mittlerweile Low-Code-/No-Code-Ansätze, die es auch Nicht-Entwicklern ermöglichen, mitzuwirken.

Nach der **Erstellung des Agenten** wird die Anwendung **bereitgestellt**. Dieser Schritt erfordert ein standardisiertes Vorgehen, um sicherzugehen, dass die Organisation über die richtigen Steuerungsmechanismen und Schutzvorkehrungen verfügt. Abschließend wird der Prozess durch eine kontinuierliche **Überwachung** zur Steuerung der Leistung abgeschlossen.

Der Erfolg eines Agenten hängt oft von einem klaren Verständnis seiner Rolle sowie der von ihm genutzten Daten, Prozesse, Eingaben und Ausgaben ab. Ebenso wichtig ist eine präzise Dokumentation der erwarteten Verhaltensweisen, um eine konsistente und überprüfbare Leistung zu gewährleisten. Der beste Weg, dies zu erreichen, besteht darin, klein anzufangen. Das bedeutet, mit einfachen, klar abgegrenzten Funktionsagenten zu beginnen und komplexe Szenarien mit zahlreichen Randfällen zunächst zu vermeiden. Sobald diese einfachen Agenten einen Mehrwert schaffen, können anspruchsvollere Szenarien entweder durch abgestimmte Orchestrierungsabläufe oder durch komplexere Agenten umgesetzt werden.

3. Zusammenarbeit zwischen menschlicher und KI-gestützter Belegschaft

Digitale Agenten als Erweiterung Ihrer Belegschaft behandeln

Sollten digitale Agenten – ähnlich wie menschliche Mitarbeiter – von der Personalabteilung verwaltet werden? Diese Frage wird zunehmend diskutiert, da agentenbasierte KI-Anwendungen immer stärker in Organisationen integriert werden. In einem Bericht über agentenbasierte Organisationen untersuchte McKinsey & Company, wie Unternehmen menschliche Arbeitskräfte und digitale Agenten kombinieren, um bessere Ergebnisse zu erzielen:

> „Wenn Agenten die Ausführung übernehmen, werden die Menschen zunehmend Ziele definieren, Kompromisse eingehen und die Ergebnisse steuern. Dies wird die Art und Weise verändern, wie Unternehmen eine hybride Belegschaft planen, wen sie einstellen (oder ausleihen) und wie sie die menschliche Belegschaft oder KI-Talente einsetzen, und wie sie den Erfolg messen. HR-Systeme erfassen nicht nur menschliche Beschäftigte, sondern dienen auch als Depot für Agenten und agentenbasierte Arbeitsabläufe."[64]

Ein wirksamer Ansatz, den einige Organisationen verfolgen, besteht darin, Stellenbeschreibungen für ihre Agenten zu erstellen (siehe unten). Diese Praxis wurde auch intern übernommen: „Unsere Teams verfassen während der anfänglichen Entwicklungsphase Stellenbeschreibungen, die denjenigen für menschliche Rollen ähneln. So ist bei der Einführung und Leistungsbewertung klar, welche Erwartungen bestehen. Kulturell trägt dieser Ansatz außerdem dazu bei, dass das Team seine eigene Rolle im Verhältnis zu den neuen digitalen Agenten besser versteht."

Stellenbeschreibung Mensch: Fachkraft für Personalwesen

Zusammenfassung der Rolle
Der Fachkraft für Personalwesen ist dafür verantwortlich, im HR-Ticketsystem die Mitarbeiterfragen mit Schwerpunkt auf Leistungs- und Vorteilswahl, Onboarding-Prozessen und Richtlinienberatung zu bearbeiten und zu lösen. Diese Rolle arbeitet eng mit einem agentenbasierten KI-Assistenten zusammen, um eine zeitnahe, präzise und personalisierte Unterstützung sicherzustellen.

Hauptverantwortlichkeiten

- Die von Beschäftigten eingereichten Anträge auf Auswahl von Sozialleistungen prüfen und bestätigen.
- Individuelle Beratung auf Grundlage von Berechtigung, Standort und Rolle der Beschäftigten bieten.
- Komplexe oder ausnahmebedingte Fälle an die Personalabteilung weiterleiten.
- Mit dem KI-Agenten zusammenarbeiten, um Ticket- Warteschlangen zu überwachen und dringende Fälle zu priorisieren.
- Leistungswahlen auf Einhaltung interner Richtlinien und gesetzlicher Vorgaben prüfen.
- Schulen und Feinabstimmen des KI-Agenten durch Überprüfung seiner Empfehlungen und Rückmeldeschleifen.

Fähigkeiten & Qualifikationen

- 3 Jahre und mehr Erfahrung im Personalwesen oder in der Personalverwaltung.
- Fundiertes Verständnis von HRIS-Systemen und Vergütungsplattformen für Unternehmen.
- Ausgezeichnete Kommunikations- und Entscheidungsfähigkeiten.
- Sicher im Umgang mit KI-Agenten und digitalen Arbeitsabläufen.

Zusammenarbeit mit KI-Agenten

- Die vom Agenten erstellten Leistungsempfehlungen überwachen und genehmigen.
- Kontext und Nuancen für Grenzfälle bereitstellen, die der Agent als mehrdeutig kennzeichnet.
- Mitwirken an der kontinuierlichen Verbesserung agentenbasierter Arbeitsabläufe und Trainingsdaten.

Stellenbeschreibung für agentenbasierte KI: Agent für die Bearbeitung von HR-Tickets

Zusammenfassung der Rolle
Der agentenbasierte KI-Assistent ist darauf ausgelegt, HR-bezogene Tickets eigenständig zu bearbeiten, mit einem Schwerpunkt auf die Auswahl von Leistungen und Vergünstigungen. Er arbeitet eng mit der Fachkraft für Personalwesen zusammen, um Genauigkeit, Compliance und Mitarbeiterzufriedenheit sicherzustellen.

Hauptverantwortlichkeiten

- Automatische Klassifizierung und Weiterleitung eingehender HR-Tickets mithilfe von natürlichem Sprachverständnis.
- Abrufen und Analysieren von Mitarbeiterdaten (z. B. Betriebszugehörigkeit, Standort, Beschäftigungsgrad), um geeignete Leistungspakete zu empfehlen.
- Erstellen personalisierter Leistungsübersichten und FAQs für Beschäftigte.
- Kennzeichnen von Tickets, die menschliches Urteilsvermögen oder Ausnahmen von der Richtlinie erfordern.
- Aus dem Feedback der Nutzer lernen und Entscheidungsmodelle entsprechend aktualisieren.
- Prüfprotokolle führen und die Nachverfolgbarkeit aller durchgeführten Aktionen gewährleisten.

Fähigkeiten

- Integration mit unternehmensweiten HRIS-, Gehaltsabrechnungs- und Sozialleistungssystemen unter Verwendung sicherer, revisionssicherer und zugelassener APIs.
- Nutzung von Richtliniendokumenten und historischen Ticketdaten, um fundierte Entscheidungen zu treffen.
- Betrieb rund um die Uhr mit Echtzeit-Reaktionsfähigkeit.
- Kontinuierliche Verbesserung durch feedbackgesteuerte Kalibrierung und Modelloptimierung.

Zusammenarbeit mit Menschen

- Übermittlung von Empfehlungen zu Sozialleistungen an den HR-Operations-Spezialisten zur Genehmigung.
- Einholen von Feedback zu abgelehnten oder angepassten Empfehlungen, um zukünftige Ergebnisse zu verbessern.
- Meldung von Anomalien, fehlenden Daten oder Richtlinienkonflikten an das zuständige Personal.

Governance und Aufsicht

- Alle Aktionen werden protokolliert und können von automatisierten Prüfagenten und Fachkräfteteam für Personalwesen überprüft werden.
- Das System arbeitet autonom, unterliegt aber einer ständigen Überwachung und Prüfbarkeit und wird regelmäßigen Audits unterzogen, um die Einhaltung der Vorschriften und die Leistungsfähigkeit des Modells zu bestätigen.

Schritt	Agentischer KI-Assistent	Menschlicher Spezialist
1. Ticket erhalten	Klassifiziert das Ticket als „Leistungsauswahl“ und extrahiert relevante Beschäftigtendaten.	Überwacht die Warteschlange und prüft markierte Tickets.
2. Empfehlung	Schlägt ein Leistungspaket basierend auf Richtlinien und Beschäftigtenprofil vor.	Prüft die Empfehlung auf Genauigkeit und Kontext.
3. Kommunikation	Sendet eine Zusammenfassung mit Links zu Anmeldeformularen und FAQs an die Beschäftigten.	Kontaktiert die Beschäftigten, wenn Klärung oder Eskalation erforderlich ist.
4. Ausnahmebehandlung	Markiert Tickets mit fehlenden Daten oder Richtlinienkonflikten.	Löst Ausnahmen und aktualisiert die Trainingsdaten des agentischen Systems.
5. Prüfung und Rückmeldung	Protokolliert Maßnahmen und lernt aus menschlichem Feedback.	Überprüft agentische Entscheidungen und gibt Rückmeldung zur Verbesserung.

Wie Mensch und KI-Assistent zusammenarbeiten, um Leistungen für einen Mitarbeiter auszuwählen

Bei der Entwicklung einer agentenbasierten KI geht es nicht darum, Menschen durch Automatisierung zu ersetzen, sondern um eine neu definierte Zusammenarbeit von Mensch und KI als komplementäre Partner. Branchenübergreifend gestalten führende Organisationen Rollen und Arbeitsabläufe neu, um KI in redaktionelle, kreative und strategische Funktionen zu integrieren. Anstatt KI als Bedrohung zu sehen, setzen sie auf ihre Rolle als Katalysator für Effizienz und Innovation. Der Unterschied liegt in der Denkweise, wobei Unternehmen erfolgreich sind, die die Einführung von KI zielgerichtet und vorausschauend gestalten, Teams weiterentwickeln, Prozesse anpassen und KI dort einsetzen, wo sie menschliche Fähigkeiten erweitert.

Im Kern bewirkt diese Transformation einen Wandel von einer emotional aufgeladenen Technologie zu einem operativen Vorteil. Kreative Fachkräfte wie Autoren, Redakteure und Designer werden nicht ersetzt, sondern in ihrer Arbeit erweitert. Künstliche Intelligenz beschleunigt heute die Forschung, erstellt erste Entwürfe und automatisiert Routineproduktion. Dadurch entstehen Freiräume, um sich auf Kernkompetenzen zu konzentrieren, wie Strategieentwicklung, Wahrung der Markenintegrität und den gezielten Einsatz von menschlichem Urteilsvermögen dort, wo es am wichtigsten ist.

Im folgenden Beispiel orchestrieren KI und menschliche Agenten zusammen mit bestehenden Arbeitsabläufen sichere Datenflüsse zwischen öffentlichen und privaten Datensätzen innerhalb einer privaten Cloud, automatisieren Ansprüche, schützen sensible Informationen und setzen die Einhaltung von Vorschriften durch. Dieses Muster vereint die Agilität der Public Cloud mit lokaler Governance und durchgängiger Datenintegrität.

Fallstudie

Gerichtsschreiber des Kreisgerichts in einem US-Bezirk

Der Gerichtsschreiber des Kreisgerichts in einem US-Bezirk überwacht ein komplexes Justizsystem, führt Gerichtsakten, sichert Beweismittel, kassiert Geldstrafen und verwaltet die Dokumentation in 24 Gemeinden und nicht eingemeindeten Gebieten. Jahrzehntelang führten papierbasierte Systeme zu Engpässen, da Beschäftigte Vorladungen manuell überprüften, Richter auf Gerichtsaktenangewiesen waren und jeder Schritt der Dateneingabe den Justizprozess verlangsamte. Die Herausforderung lag nicht nur in der Ineffizienz, sondern auch im Umfang. Mit dem beschleunigten Bevölkerungswachstum drohte die Vielzahl der Verkehrsfälle die Kapazitäten des Personals zu überfordern und die Gerichtsverfahren zu verzögern.

Durch die Einführung KI-gestützter Automatisierung innerhalb einer gesteuerten Fallmanagementplattform wurden die Gerichtsabläufe des Bezirks grundlegend verändert. Die agentenbasierte KI überwacht den Fallstatus, validiert Datensätze und leitet Dokumente sicher systemübergreifend weiter, während menschliche Sachbearbeiter weiterhin eingebunden bleiben, um Ausnahmen zu genehmigen und Sonderfälle zu prüfen. Richter können über ein einheitliches Dashboard sofort digitale Fallakten abrufen, auf integrierte staatliche Verkehrsdatenbanken zugreifen und frühere Verwarnungen einsehen.

Im Hintergrund verbinden KI-Agenten ein im Fahrzeug integriertes digitales Ticketsystem des Landkreises mit Dokumentenarchiven und Gerichtsdatenbanken und automatisieren dadurch die Erfassung, Klassifizierung und Validierung von Informationen. Sensible Daten verlassen niemals die private Cloud des Landkreises, und jeder Entscheidungspunkt wird zu Prüfungs- und Compliance-Zwecken protokolliert. Was früher manuelle Dateneingabe erforderte, geschieht heute in Sekundenschnelle, wodurch Sachbearbeiter sich auf höherwertige Aufgaben konzentrieren können und das Risiko menschlicher Fehler sinkt.

Mit diesem hybriden KI-Modell, einer Kombination aus agentenbasierter Automatisierung und menschlicher Aufsicht, hat das Büro des Gerichtsschreibers sein gesamtes Informationsökosystem modernisiert. Das Ergebnis sind schnellere Fallbearbeitung, höhere Datenintegrität und vollständige Übereinstimmung mit den staatlichen Vorgaben zum Informationsaustausch. Es ist ein Entwurf für verantwortungsvolle KI in der Regierung, eines intelligenten Systems, das mit menschlichem Urteilsvermögen zusammenwirkt, um einen vertrauenswürdigen und effizienteren öffentlichen Dienst zu gewährleisten.

4. Leistungsmanagement und -messung

Den Erfolg einer agentenbasierten KI-Anwendung messen

Die Messung der Ergebnisse ist der entscheidende Faktor für den Erfolg oder Misserfolg einer agentenbasierten KI-Anwendung. Daher müssen diese Messgrößen im Vorfeld eindeutig in einer Stellenbeschreibung definiert werden, die Ziele, Aufgaben und Kennzahlen des **Beschäftigten** umreißt. Diese im Rahmen der Erstellung des Business Case entwickelte Stellenbeschreibung ist maßgeblich für die Festlegung des geeigneten Aufgabenbereichs des Agenten.

Heute existiert kein allgemein definierter oder anerkannter Standard für KPIs bei agentenbasierten EAI-Anwendungen, auch wenn in vielen Ländern und bei zahlreichen Aufsichtsbehörden Standards im Bereich KI vorhanden sind, etwa zu Ethik, Transparenz und Risiko. Mangels eines universellen Standards können Organisationen auf die vielen öffentlich zugänglichen Rahmenwerke zurückgreifen, indem sie eines direkt übernehmen oder es je nach ihrer spezifischen Umsetzung anpassen.

Ein im *International Journal of Scientific Research and Modern Technology* vorgestelltes Rahmenwerk untersuchte fünf zentrale Dimensionen, nämlich Modellqualität, Systemleistung, geschäftliche Auswirkungen, Interaktion zwischen Mensch und KI sowie ethische und ökologische Überlegungen. Im Hinblick auf die Dimension der Modellqualität untersuchte die Studie Genauigkeit, Präzision, Aufgabenerfüllung, Halluzination und Output. Die operativen Leistungskennzahlen konzentrierten sich auf Latenz, Durchsatz und Ressourcennutzung. Business Impact bewertete den ROI, die Kosteneinsparungen, die Produktivitätssteigerungen und die Marktauswirkungen. Die Interaktion zwischen Mensch und KI wurde anhand von Nutzerzufriedenheit, Vertrauen, Akzeptanz und Engagement analysiert. Abschließend betrachtete die Studie im Hinblick auf ethische und ökologische Aspekte Voreingenommenheit, Fairness, Transparenz, Umweltauswirkungen und ethische Abweichungen.[65]

Auch wenn nicht alle diese KPIs für jeden Einsatz agentenbasierter KI relevant sind, ist es wichtig, dass Teams bei der Auswahl der verschiedenen Dimensionen einen ganzheitlichen Ansatz verfolgen. Nachdem die relevanten KPIs festgelegt sind, können sie in Scorecards formalisiert werden. Dieser Ansatz entspricht dem Verfahren, das bei der robotergestützten Prozessautomatisierung erfolgreich eingesetzt wurde, bei der Scorecards maßgeblich zur Automatisierung und Effizienzsteigerung beitrugen. Die gleiche Strenge sollte hier mit täglichen Beurteilungen anhand der wichtigsten Dimensionen des Agenten angewendet werden.

Ebenso wichtig ist ein klar definierter Rahmen mit einem transparenten Prozess zur Behebung von Mängeln. Nicht jede KI-Anwendung in Unternehmen wird erfolgreich sein, und für die Projekte, die keinen tragfähigen Weg finden, ist dieses Rahmenwerk notwendig, um die Ursache zu identifizieren und eine datengestützte Entscheidung darüber zu treffen, ob eine Korrektur sinnvoll ist oder die Anwendung außer Betrieb genommen werden sollte. Dieser Ablauf unterstützt unmittelbar eine Kultur des schnellen Abbruchs, eine Umstellung, die vielen Organisationen nach wie vor schwerfällt. Teams müssen darauf vorbereitet werden, zu akzeptieren, dass es sinnvoller ist, ein nicht tragfähiges Projekt einzustellen, als eine Lösung zu erzwingen, die das erwartete Ergebnis niemals erreichen wird.

Ebenso ist zu berücksichtigen, dass EAI auf sicheren, kontextbezogenen Daten basiert und nicht auf reiner Menge. Durch die Verbindung strukturierter Systeme, unstrukturierter Inhalte und organisationsübergreifender Informationsflüsse innerhalb der EIM Cloud können Organisationen Agenten entwickeln und bereitstellen, die verantwortlich agieren, ihre Entscheidungen nachvollziehbar machen und Compliance-Anforderungen von Beginn an erfüllen. Im nächsten Kapitel folgt ein Blick auf die Entwicklung von agentenbasierter KI hin zu AGI.

Die fünf Merksätze

1. **Das richtige Organisationsmodell für den KI-Einsatz auswählen.**
   Die Wahl eines geeigneten Bereitstellungsmodells, ob zentralisiert, Hub-and-Spoke, föderiert oder hybrid, ist grundlegend für die erfolgreiche Einführung agentenbasierter KI. Dieses Modell sollte Eigentumsverhältnisse, Verantwortlichkeiten und Governance klar definieren. Gleichzeitig muss es Kontrolle und Agilität mit den Anforderungen der jeweiligen Branche und dem regulatorischen Umfeld in Einklang bringen.

2. **Eine stufenweise und Benutzergruppen basierte Einführungsstrategie anwenden.**
   Eine überstürzte Einführung von KI führt häufig zu Fehlentscheidungen. Erfolgreiche Organisationen setzen agentenbasierte KI-Anwendungen in klar strukturierten Phasen ein, die auf spezifische Geschäftsbereiche oder Regionen ausgerichtet sind. Dieser Benutzergruppen orientierte Ansatz gewährleistet Einsatzbereitschaft, stärkt die Nutzerakzeptanz und ermöglicht eine passgenaue Unterstützung für jede Gruppe.

3. **Mit kleinen, klar definierten Anwendungsfällen beginnen und auf Erfolgen aufbauen.**
   Der Einstieg in einfache und eindeutig abgegrenzte Anwendungsfälle schafft ein belastbares Fundament für weitere Entwicklungen. Geschäftsbereiche mit ausgereiften Prozessen und hoher Datenqualität bieten hierfür die besten Voraussetzungen. Veränderungsagenten wie Produktverantwortliche tragen dazu bei, Akzeptanz zu fördern und auf frühen Erfolgen aufzubauen, bevor komplexere Szenarien skaliert werden.

4. **Digitale Agenten in die Personalplanung und -verwaltung integrieren.**
   KI-Agenten sind als Erweiterung der Belegschaft zu behandeln. Digitale Agenten sollten als Erweiterung der Belegschaft betrachtet werden. Klare Rollen, Erwartungen und HR-Elemente wie Stellenbeschreibungen und Leistungsziele schaffen Transparenz und stärken die Abstimmung zwischen menschlichen und digitalen Teammitgliedern. Diese Einbettung fördert das Verständnis der Technologie und die Akzeptanz innerhalb der Organisation.

5. **Strenge Leistungsmessung und Korrekturmaßnahmen implementieren.**
   Der Erfolg hängt von klar definierten KPIs in mehreren Dimensionen ab (Modellqualität, Geschäftsauswirkungen, Mensch-KI-Interaktion, Ethik und Systemleistung). Scorecards dienen der regelmäßigen Bewertung der Ergebnisse und unterstützen fundierte Entscheidungen über die Überarbeitung oder Außerdienststellung leistungsschwacher Anwendungen. Eine Kultur, die aus Fehlschlägen lernt und schnell iteriert, stärkt die kontinuierliche Weiterentwicklung.

Kapitel Zehn

Die Entstehung von AGI aus agentenbasierter KI

Wie in diesem Buch gezeigt, ist die Geschichte der KI eine Geschichte des Strebens und der Innovationen, geprägt von dem fortwährenden Interesse, Systeme zu entwickeln, die rechnen und verstehen, reagieren und schlussfolgern. Agentenbasierte KI eröffnet einen Einblick in künftige Möglichkeiten und führt zu Spitzenmodellen, die den Weg zur Allgemeinen KI ebnen. Allgemeine KI wird die Entwicklung domänenspezifischer Systeme zu einer Form von Intelligenz fortsetzen, die menschenähnliche Generalisierung, domänenübergreifendes Lernen, die Bildung abstrakter Konzepte und autonomes Handeln ermöglicht.

Unternehmen, die intelligente Systeme in großem Maßstab einsetzen wollen, müssen die Steuerung, Orchestrierung und das Lebenszyklusmanagement vernetzter Agentensysteme planen, nicht nur einzelner Modelle. Während sich diese Rahmenwerke weiterentwickeln, dienen sie als Bausteine auf einem Weg von handlungsfähiger KI zu umfassenderen und adaptiveren kognitiven Systemen, die klassische Architektur-, Betriebs- und Governance-Modelle infrage stellen.

Dieses Kapitel untersucht den Übergang von agentenbasierter KI, die Ziele verfolgt und unter Anleitung lernt, hin zum Potenzial der Allgemeinen KI, einer Form von Intelligenz, die menschenähnliches Verhalten, Verständnis, logisches Denken und Anpassungsfähigkeit in unterschiedlichen Kontexten ermöglicht.
Es beleuchtet die technischen und ethischen Grundlagen für diesen Übergang und greift die laufende Debatte darüber auf, ob Allgemeine KI durch die bloße Skalierung bestehender Modelle oder durch neue Architekturen, erweiterte Denkfähigkeiten und emotionale Intelligenz entstehen wird.

Der Weg von domänenspezifischer KI zu Allgemeiner KI (AGI)

Agentenbasierte KI: Die Brücke zur AGI

Agentenbasierte KI kann aufgabenspezifische Aktionen ausführen, bleibt jedoch an bestimmte Domänen gebunden und ist nicht in der Lage, domänenübergreifend zu lernen. Ihre Argumentation ist Kontext aber nicht Konzept bezogen, sodass weder Prinzipien abstrahiert noch Lernprozesse verallgemeinert werden können. Im Unterschied dazu kann Allgemeine KI lernen, schlussfolgern und sich an ein breites Spektrum von Aufgaben anpassen. AGI nähert sich jener flexiblen Intelligenz an, die für die menschliche Kognition charakteristisch ist.

Der Übergang von agentenbasierter KI zu Allgemeiner KI beruht nicht allein auf einer größeren Datenmenge, sondern auf einem veränderten architektonischen Ansatz. Forschende diskutieren, ob Allgemeine KI durch die bloße Skalierung aktueller grundlegender Modelle entstehen oder ob ein neuer Zugang zum Denken, zur Erfahrung und emotionaler Intelligenz erforderlich sein wird. Zwei Ansätze prägen die frühen Debatten zu diesem Thema:

1. Skalierungshypothese: Die Skalierungshypothese geht davon aus, dass Allgemeine KI durch die konsequente Erweiterung heutiger Large Language Models und multimodaler Modelle entsteht, ohne dass neue Algorithmen nötig sind. In Scaling *Laws for Neural Language Models* (Skalierungsgesetze für neuronale Sprachmodelle) wird der Zusammenhang zwischen Modellleistung und folgenden drei Faktoren beschrieben: Modellgröße, Datensatzumfang und Rechenkapazität. Die Autorinnen und Autoren kommen zu dem Schluss, dass größere Modelle weiterhin besser abschneiden und effizienter mit Stichproben umgehen als bislang angenommen, sodass große Modelle wichtiger sein könnten als große Datenmengen. Vor diesem Hintergrund sind weiterführende Untersuchungen zur Modellparallelität gerechtfertigt. Tiefe Modelle können mithilfe von Pipelining trainiert werden, wodurch die Parameter schichtweise auf Geräte verteilt werden, was jedoch bei einer steigenden Anzahl von Geräten größere Batchgrößen erfordert."[66]

2. Diskontinuitätshypothese: Die Diskontinuitätshypothese geht davon aus, dass Allgemeine KI nicht allein durch die Vergrößerung bestehender großer Sprachmodelle oder neuronaler Architekturen entstehen wird, sondern grundlegende neue Paradigmen, Architekturen oder Formen der Kognition notwendig sind. Einige Fachleute argumentieren, dass die Skalierung neuronaler Netze nicht zu Allgemeiner KI führt, da strukturiertes Denken, Kausalmodelle und kompositionelle Generalisierung fehlen, wesentliche Merkmale menschlicher Kognition. Gary Marcus beispielsweise hat die Bedeutung von Symbolen gegenüber neuronalen Netzen für die Weiterentwicklung der KI betont:

> „Symbole [computerinterne Codierungen, wie Zeichenketten aus Binärbits, die für komplexe Ideen stehen] sind den heutigen neuronalen Netzen in vielen grundlegenden Aspekten der Datenverarbeitung immer noch weit überlegen. Sie sind deutlich besser geeignet, komplexe Szenarien logisch zu durchdenken, können grundlegende Operationen wie Arithmetik systematischer und zuverlässiger ausführen und präziser darstellen, wie Teile und Ganzes zueinanderstehen, was sowohl für die Interpretation der dreidimensionalen Welt als auch für das Verständnis menschlicher Sprache entscheidend ist. Sie sind robuster und flexibler in ihrer Fähigkeit, groß angelegte Datenbanken darzustellen und abzufragen. Symbole eignen sich außerdem besser für formale Verifikationsverfahren, die für einige Sicherheitsaspekte von zentraler Bedeutung und beim Entwurf moderner Mikroprozessoren allgegenwärtig sind. Es wäre wenig sinnvoll, auf diese Stärken zu verzichten, anstatt sie in einer Form von Hybrid-Architektur zu nutzen."[67]

Aller Wahrscheinlichkeit nach wird der Weg von agentenbasierter KI zu AGI nicht einfach sein, sondern vielmehr eine Mischung aus Modell- und Fähigkeitserweiterung, angetrieben durch Verbesserungen bei der Datenqualität, der Modellinterpretierbarkeit, der Verfügbarkeit von Rechenleistung und der Nutzung symbolischer Argumentation jenseits traditioneller neuronaler Netze.

Die Rolle von agentenbasierter KI und Unternehmensorchestrierung

Da der Einsatz agentenbasierter KI immer ausgereifter wird, stellt er einen ersten Schritt hin zu hybriden Ansätzen für die Entwicklung einer AGI dar. Mit agentenbasierter KI lassen sich komplexe Aufgaben auf einem Niveau zerlegen und verteilen, das an menschliche Arbeitsweisen erinnert. Spezialisierte Agenten können für Funktionen wie Planung, Recherche, Codierung, Überprüfung, Simulation und Verwaltung eingesetzt werden. Sie führen diskrete Funktionen aus, können parallel arbeiten und kooperieren, wodurch die kognitive und rechnerische Belastung für einzelne Modelle sinkt.

Parallel dazu lassen sich spezialisierte Agenten unabhängig weiterentwickeln und für unterschiedliche Aufgaben wiederverwenden. Sie können verschiedene Lernansätze verfolgen und dabei bestärkendes Lernen zur Optimierung, symbolisches Denken zur Logik oder unüberwachtes Lernen zur Entdeckung neuer Muster nutzen. Diese Modularität ermöglicht den Aufbau hybrider Lernsysteme.

In Kombination mit den Agenten übernimmt die Orchestrierungsebene eine zentrale Koordinierungsfunktion. Der Orchestrator verwaltet Aufgaben durch Zerlegung und Zuweisung, übernimmt Planung und Ressourcenallokation, steuert die Kommunikation durch Weitergabe relevanter Kontextinformationen und bietet Aufsicht, indem er Ausgaben validiert, die Leistung überwacht und den Systemlebenszyklus steuert.

Entscheidend ist zudem, dass die Orchestrierungsebene auch das Lernen von Orchestrierung ermöglicht, sodass sich das gesamte agentenbasierte System im Laufe der Zeit auf Grundlage verteilter Erfahrungen verbessern kann. Gleichzeitig bleibt zu berücksichtigen, dass Orchestrierung nicht mit Kognition gleichzusetzen ist, da vielen gegenwärtigen Systemen langfristige Planung, persistentes Gedächtnis, logisches Denken und modellbasiertes Kausalverständnis fehlen. Aus gestalterischer Sicht sollte sich der Orchestrator von einem einfachen Aufgabenplaner zu einem Meta-Controller weiterentwickeln, dessen Kernfunktionen Zielmanagement, dynamische Ressourcen- und Rollenzuweisung an Agenten, reflektierende Fehler- und Feedbackschleifen sowie kontinuierliche Evaluierungs- und Sicherungspipelines umfassen.
Dieser modulare und orchestrierte Ansatz bildet eine tragfähige Grundlage für das Training hybrider Systeme auf dem Weg zur Allgemeinen KI.

Im folgenden Artikel wird dargestellt, wie eine portugiesische Gemeinde mithilfe KI-gestützter Automatisierung die Produktivität gesteigert und die Betriebskosten gesenkt hat.

Fallstudie

Eine portugiesische Gemeinde

Als Teil des städtischen Ballungsraums Groß-Lissabon beschäftigt die Gemeinde mehr als tausend Mitarbeiter. Für die Organisation des öffentlichen Sektors stellte die zunehmende Komplexität digitaler Abläufe eine Herausforderung dar, da fragmentierte Systeme und manuelle Arbeitsprozesse die Klassifizierung und Priorisierung von Anfragen erschwerten, die Nachverfolgung von Interventionen behinderten und eine präzise Dokumentation verhinderten. Daten waren vorhanden, aber Erkenntnisse fehlten. Ohne integrierte Aufsicht waren die Teams zu reaktiven Entscheidungen gezwungen. Sie hatten Schwierigkeiten, die Einhaltung der sich fortlaufend ändernden regulatorischen Standards aufrechtzuerhalten. Das Ergebnis war ein Mangel an Transparenz und Verantwortlichkeit. Dies stellte ein Hindernis dar, das moderne KI-Governance-Modelle besonders gut angehen können.

Um diese Probleme zu lösen, entwickelte die Organisation ein einheitliches KI-gestütztes Management-Framework, das Arbeitsabläufe orchestriert, Anfragen klassifiziert und den Servicebedarf anhand von Mustererkennung vorhersagt. Durch strukturierte Governance und intelligente Automatisierung entstanden Echtzeit-Einblicke in Leistung, Ressourcenzuweisung und regulatorische Compliance. Anstelle manueller Steuerung arbeitet die Organisation nun mit einem sich selbst weiterentwickelnden System, das aus jeder Interaktion lernt. Komponenten agentenbasierter KI analysieren kontinuierlich Leistungsdaten, optimieren Arbeitsabläufe und identifizieren entstehende Ineffizienzen, wodurch ein adaptives Servicemanagement entsteht, das sich an der Unternehmensstrategie orientiert.

Diese Entwicklungen gehen über operative Verbesserungen hinaus und markieren einen Schritt hin zu einer unternehmerischen Intelligenz, die Elemente Allgemeiner KI vorwegnimmt. Durch die vollständige Integration von KI in das Servicemanagement reagiert das System nicht nur auf Anfragen, sondern antizipiert sie, indem es Kontext analysiert, Ergebnisse prognostiziert und die Zusammenarbeit zwischen Teams koordiniert. Mit der fortschreitenden Entwicklung dieser agentenbasierten Systeme entsteht eine Grundlage für Organisationen, die lernen, argumentieren und sich in großem Umfang weiterentwickeln können. Was als Optimierung der IT-Abläufe begann, hat sich zu einem Modell dafür entwickelt, wie intelligente Governance, menschliche Aufsicht und adaptive KI in einer Struktur zusammenwirken, in der die Automatisierung den Mensch nicht ersetzt, sondern seine Fähigkeit stärkt, fundierter zu denken und zu handeln.

Die Zahlen im Überblick: Die Auswirkungen intelligenter Governance

- **60 % schnellere Bearbeitung** von Anfragen durch die Umstellung von manuellen Arbeitsabläufen auf KI-gestützte Klassifizierung und Priorisierung.
- **Bis zu 40 % Reduzierung der Betriebskosten** durch Automatisierung des Vertrags-, Anlagen- und Servicemanagements.
- **Vollständige Transparenz** bei Leistungskennzahlen und Einhaltung von Vorschriften in allen Servicebereichen.
- **50 % weniger Prozessfehler** dank automatisierter Validierung, Revisionsprotokollen und Echtzeit-Überwachung bei der Erkennung von.
- **Etablierter kontinuierlicher Lernkreislauf:** Agentenbasierte KI-Systeme optimieren nun Arbeitsabläufe autonom auf Grundlage von Datentrends und Feedback.
- **Die abteilungsübergreifende Zusammenarbeit** wurde verbessert und Silos wurden durch transparente, adaptive Arbeitsabläufe reduziert.

Daten als Wegbereiter für Allgemeine KI: Kognitive Skalierung vorantreiben

Die Grundlage der Entwicklung hin zu Allgemeiner KI bilden weiterhin Daten, wobei Datenqualität, Diversität und Governance die wichtigsten Faktoren darstellen. Agentenbasierte KI profitiert von strukturierten und semistrukturierten Daten innerhalb definierter operativer Grenzen. Allgemeine KI benötigt jedoch einen umfangreicheren und repräsentativeren Datensatz, der Schlussfolgerungen höherer Ordnung unterstützt. Aus diesem Grund gewinnt der Bedarf an Daten mit der Weiterentwicklung der Technologie kontinuierlich an Bedeutung.

Daten für Allgemeine KI müssen Kontext, Kausalität und Ethik erfassen. Dies erfordert ein Daten-Framework, föderierte Datenaustauschmodelle und souveräne Dateninfrastrukturen, die verantwortungsvollen Zugriff und verantwortungsvolle Nutzung sicherstellen. Die KI-Prinzipien (2019) der Organisation für wirtschaftliche Zusammenarbeit und Entwicklung (Organisation for Economic Co-operation and Development, OECD) und die Norm ISO/IEC 42001:2023 betonen beide, dass KI-Systeme unter klar definierten Daten-Governance-Mechanismen betrieben werden sollten, die Fairness, Verantwortlichkeit und Rückverfolgbarkeit gewährleisten.

Technologien zur Verbesserung der Privatsphäre spielen ebenfalls eine entscheidende Rolle. Da sich KI immer weiter der allgemeinen Kognition annähert, wird die Wahrung von Datenethik und Datenintegrität ebenso wichtig wie Leistungskennzahlen.

Governance und Ethik: Autonomie und Verantwortlichkeit in Einklang bringen

Da agentenbasierte KI-Systeme immer unabhängiger werden, rücken Fragen der Kontrolle, Rechenschaftspflicht und Aufsicht in den Vordergrund, wie in Kapitel 4 erläutert wurde. Die Herausforderung wächst jedoch, je näher sich die Entwicklung der Allgemeinen KI annähert, die über einen Grad an Autonomie verfügen könnte, der menschliche Fähigkeiten übersteigt. Ethische Rahmenbedingungen für KI in Unternehmen konzentrierten sich in der Vergangenheit auf Fairness, Transparenz und Erklärbarkeit. Wenn Systeme jedoch beginnen, komplexe Entscheidungen zu treffen, wird es wichtig sein, sicherzustellen, dass die Ziele von KI-Systemen mit menschlichen Werten im Einklang stehen.

Wie in Kapitel 6 erörtert, konzentrieren sich die internationalen Bemühungen um bessere Governance in Bezug auf diese Herausforderung. Der EU-Gesetzentwurf zur künstlichen Intelligenz (2024) legt abgestufte Risikokategorien für KI fest und schreibt eine strenge Überwachung von Anwendungen mit hohem Risiko vor. Die UNESCO-Empfehlung zur Ethik der künstlichen Intelligenz (2021) fordert eine auf Menschenrechten basierende Governance, während das Nationale Institut für Standards und Technologie (National Institute of Standards and Technology , NIST) AI RMF (2023) Vertrauenswürdigkeit als messbares Element einführt. Diese Initiativen gewährleisten, dass KI-Systeme auch bei der Erweiterung ihrer Fähigkeiten weiterhin unter menschlicher Intention und Kontrolle bleiben.

Aus politischer Sicht könnte Allgemeine KI ein neues Maß an Autonomie ermöglichen, in dem sich Systeme auf unvorhersehbare Weise anpassen oder selbst verbessern könnten. In Erwartung dessen arbeiten Forschende im Bereich der KI-Sicherheit an neuen Standards. Auch dies muss parallel zur technologischen Entwicklung erfolgen, sodass die Standards vorbereitet sind, sobald Allgemeine KI einsatzbereit ist.

Die Unternehmensrelevanz von AGI-ähnlichen Systemen

Da Unternehmen zunehmend autonome und vernetzte agentenbasierte KI-Systeme einsetzen, zeigen sich in der Praxis erste Anzeichen von Intelligenzformen, die über spezialisierte Automatisierung hinausgehen – auch wenn sie noch nicht als Allgemeine KI gelten. Systeme, die wir als „AGI-ähnlich" bezeichnen, entstehen organisch in Organisationen: intelligente Agenten, die aus verschiedenen Datenquellen lernen, Entscheidungen abteilungsübergreifend koordinieren und ihr Verhalten an sich ändernde Ziele oder externe Gegebenheiten anpassen. Es handelt sich hierbei nicht um isolierte Werkzeuge, sondern um sich selbst optimierende Netzwerke von Intelligenz, die die Arbeitsweise kontinuierlich verfeinern.

Der Wandel ist schleichend, aber tiefgreifend. Während frühere Automatisierung klar umrissene Routineaufgaben ersetzte, integriert agentenbasierte KI heute Schlussfolgerungen, Planung und Selbstkorrektur. Im Finanzdienstleistungssektor können Agenten etwa Kundenstimmungen analysieren, Abwanderungswahrscheinlichkeiten berechnen und eigenständig Strategien zur Kundenbindung entwerfen – Tätigkeiten, die Marketing, Risiko und Compliance in einem gemeinsamen Feedbackkreislauf vereinen. In der Fertigung analysieren KI-Systeme bereits Sensordaten, verteilen Ressourcen innerhalb der Lieferkette neu und erkennen ethische Risiken in der Beschaffung, bevor menschliche Teams eingreifen. Die Grenze zwischen spezialisierter Automatisierung und kognitiver Intelligenz auf Unternehmensebene wird zunehmend unscharf.

Diese Entwicklung eröffnet erhebliche strategische Chancen, birgt jedoch auch Risiken. Ohne wirksame Steuerungsmechanismen könnten Unternehmen Systeme betreiben, die über ihre ursprüngliche Programmierung hinaus lernen und handeln. Laut Gartner nennen über 80 % der Unternehmen, die KI in großem Maßstab einsetzen möchten, Governance und Transparenz als größte Hindernisse bei der Einführung.[68]

Die Steuerung solcher Systeme erfordert neue Modelle der Verantwortlichkeit und Aufsicht. Das traditionelle IT-Management war auf statische Strukturen ausgerichtet, während moderne KI-Ökosysteme adaptive Steuerungsformen verlangen, die sich parallel zur KI selbst weiterentwickeln. Erklärbarkeit, Nachvollziehbarkeit und Feedbackschleifen müssen feste Bestandteile der Architektur sein, nicht nachträgliche Ergänzungen. Da diese Systeme beginnen, domänenübergreifend zu agieren, bleibt menschliches Urteilsvermögen unverzichtbar – nicht, um Prozesse zu verlangsamen, sondern um Richtung und Maß vorzugeben.

Allgemeine KI ist damit kein fernes Konzept, sondern eine sich abzeichnende Realität in der Unternehmenswelt. Die Organisationen, die in dieser neuen Phase führend sein werden, sind jene, die Intelligenz als Infrastruktur begreifen – als Ressource, die wie Daten oder Cybersicherheit gesteuert, integriert und fortlaufend verbessert werden muss. Das Ergebnis sind keine Maschinen, die Menschen ersetzen, sondern Systeme, die deren Fähigkeiten erweitern, den Einblick in Unternehmensprozesse vertiefen, den Wandel beschleunigen und Vertrauen als Grundlage künftiger Entwicklung schaffen.

Im folgenden Beispiel wird dargestellt, wie eine führende südafrikanische Universität einen KI-gestützten automatisierten Servicedesk einsetzt, um das Studentenerlebnis zu verbessern und die Geschäftskontinuität in Krisensituationen sicherzustellen.

Fallstudie

Eine südafrikanische Universität

> *Der Nutzen von maschinellem Lernen ist in unserer Studentengemeinschaft enorm, wie die breite Verwendung unserer virtuellen Agenten und Wissensartikel beweist. Ohne maschinelles Lernen und KI gäbe es für uns absolut keine Möglichkeit, unsere Endnutzer mit den wenigen uns zur Verfügung stehenden Mitarbeitern zu unterstützen.*
>
> Änderungs- und Konfigurationsmanager, Universität in Südafrika

Eine der führenden Universitäten Afrikas betreibt Forschung, um Lösungen für drängende gesellschaftliche Probleme zu entwickeln. Die Lehre erfolgt sowohl im Hörsaal als auch online und ist eng mit den lokalen Gemeinschaften verbunden. Angesichts des wachsenden Umfangs und der bestehenden digitalen Ungleichheit beschloss die Universität, die Art und Weise, wie sie Bildung und Unterstützung in einer zunehmend hybriden Welt vermittelt, grundlegend zu überdenken. Bei über 130.000 Studierenden und nur einer kleinen Zahl engagierter Supportmitarbeiter waren manuelle Prozesse nicht länger tragfähig. Der eingeschränkte Zugang zu Geräten und Netzverbindungen vertiefte die digitale Kluft, während nicht-technische Abteilungen Schwierigkeiten hatten, sich an die neue Form der Fernarbeit anzupassen. Die Herausforderung war nicht nur technologischer, sondern struktureller Natur. Die Universität benötigte ein intelligentes Betriebsmodell, das akademische und administrative Kontinuität sicherstellt und zugleich allen Studierenden die aktive Teilhabe an einem vernetzten Ökosystem ermöglicht.

Um dieser Herausforderung zu begegnen, wurde das digitale Service-Framework vollständig im Hinblick auf Automatisierung, Governance und künstliche Intelligenz neugestaltet. KI-basierte Systeme klassifizieren Anfragen, prognostizieren den Bedarf und leiten Aufgaben automatisch weiter. Dadurch kann ein kleines Team ein großes Volumen an Supportinteraktionen in Echtzeit bearbeiten. Während der COVID-Pandemie beschleunigten maschinelle Lernverfahren die Umstellung auf Fernunterricht. Automatisierte Prozesse ermöglichten den VPN-Zugang, die Bereitstellung von Laptops und die Optimierung der Konnektivität für Tausende von Studierenden. Agentenbasierte KI wurde zum unsichtbaren Rückgrat der digitalen Infrastruktur der Universität. Sie orchestrierte Arbeitsabläufe über Abteilungen hinweg und erweiterte die Intelligenz auf nicht-technische Bereiche wie Einschreibung, Studierendenservice und Bibliotheksbetrieb. Jedes System lernte kontinuierlich aus Interaktionen und steigerte dadurch Genauigkeit, Reaktionsgeschwindigkeit und Fairness innerhalb der gesamten Institution.

Was als Krisenmaßnahme begann, entwickelte sich zu einem Modell für eine Bildung, die auf Allgemeine KI vorbereitet ist – eine Form des Lernens, in der adaptive Intelligenz sowohl Reichweite als auch Chancengleichheit verbessert. Heute ist das Volumen der Serviceanfragen deutlich gestiegen, gleichzeitig haben sich Effizienz und Transparenz verbessert. KI unterstützt nicht nur Mitarbeitende und Studierende, sondern arbeitet aktiv mit ihnen zusammen. Sie lernt aus Mustern, Kontexten und Feedback, um Bedürfnisse vorherzusehen und Entscheidungsprozesse zu optimieren. Die Transformation der Universität zeigt, wie intelligente Steuerung und agentenbasierte Automatisierung menschliche und digitale Fähigkeiten miteinander verbinden. Sie schafft die Grundlage für ein akademisches Ökosystem, in dem Intelligenz verteilt, kollaborativ und fortlaufend selbstverbessernd ist.

Die Zukunft der AGI neu definieren: Jenseits des technischen Horizonts

Das Streben nach Allgemeiner KI ist ebenso sehr eine philosophische und gesellschaftliche Entwicklung wie eine technische. Während einige sie als logische Folge der Skalierung bestehender Architekturen betrachten (Skalierungshypothese), sehen andere darin eine grundlegende Neudefinition von Intelligenz – einen Schritt hin zu Systemen mit Intentionalität, moralischem Urteilsvermögen und Selbstbewusstsein (Diskontinuitätshypothese).

Wenn agentenbasierte KI die Automatisierung von Handlungen darstellt, verkörpert Allgemeine KI die Automatisierung des Verstehens. Zukünftige KI-Systeme müssen nicht nur denken und lernen, sondern sich zugleich an kollektiven menschlichen Werten orientieren. Dies ist eine Herausforderung, die das kommende Jahrzehnt der KI-Politik und -Innovation in Unternehmen prägen wird. Der Übergang von agentenbasierter KI zu Allgemeiner KI bedeutet nicht nur, dass Maschinen menschliche Fähigkeiten übertreffen. Er beschreibt vielmehr, wie sich menschliche und maschinelle Intelligenz gemeinsam weiterentwickeln.

Mit wachsender Leistungsfähigkeit von KI-Systemen werden sich auch die Rollen menschlicher Arbeitskräfte verändern. Aus diesem Grund muss die menschliche Aufsicht weiterhin Priorität haben – sowohl aus Governance-Perspektive als auch aus operativer und betrieblicher Sicht. Rollen, die heute selbstverständlich erscheinen, könnten künftig entfallen, während neue Aufgabenfelder entstehen. Diese reichen von der Schulung und Steuerung agentenbasierter KI-Funktionen über die Sicherstellung von Datenqualität und Daten-Governance bis zur Entwicklung neuer Produkte und Fähigkeiten, die heute noch nicht vorhersehbar sind. Jeder technologische Fortschritt schafft neue Möglichkeiten, und Allgemeine KI wird keine Ausnahme darstellen.

Die Zukunft der Arbeit wird nicht durch den Ersatz des Menschen, sondern durch die Partnerschaft zwischen Mensch und Maschine geprägt sein. Die Entwicklung wirksamer Kollaborationsmodelle zwischen menschlichen Fachkräften und generalistischen KI-Systemen erfordert eine präzise Definition von Rollen, Befugnissen und Verantwortlichkeiten innerhalb gemeinsamer Arbeitsstrukturen. Dazu gehört auch, kontinuierlich in Weiterbildung und digitale Kompetenz zu investieren, damit Teams verstehen, wie sie KI-Ergebnisse verantwortungsvoll hinterfragen, interpretieren und steuern können. In einer Welt, in der KI schneller lernt, bleibt Vertrauen die Grundlage jeder Form von Intelligenz. Ethische Kontrolle stellt sicher, dass technologische Fortschritte den Zielen der Menschheit dienen – und nicht umgekehrt.

In diesem Kapitel wurde die Entwicklung von agentenbasierter KI hin zu einer möglichen Allgemeinen KI nachgezeichnet, die menschenähnliches Denken, Lernen und Anpassungsfähigkeit über verschiedene Kontexte hinweg ermöglicht. Dieser Übergang erfordert nicht nur die Skalierung von Daten, sondern ebenso die Integration höherer Datenqualität, besserer Modellinterpretierbarkeit und symbolischen Denkens. Dadurch entsteht der Weg für hybride EAI-Systeme, die komplexe Aufgaben durch Spezialisierung und Zusammenarbeit bewältigen können.

Mit Blick auf die Zukunft wird Allgemeine KI nicht als Ersatz, sondern als Verstärker menschlicher Intelligenz verstanden. Daten, die im Enterprise Information Management (EIM) verwaltet und gesteuert werden, spielen dabei als Wegbereiter eine zentrale Rolle. Sie liefern das Wissen, das erforderlich ist, damit fortschrittliche Systeme lernen, fundierte Entscheidungen treffen und sich an komplexe, dynamische Umgebungen anpassen können. Investitionen in agentenbasierte KI werden sich im Übergang zur Allgemeinen KI als entscheidend erweisen.

Im folgenden Beispiel wird dargestellt, wie ein mexikanischer Anbieter von Einzelhandelsanalysen Verkaufsdaten in verwertbare Erkenntnisse umwandelt, um das Umsatzwachstum zu fördern.

Fallstudie

Ein mexikanisches Unternehmen für Einzelhandelsanalysen

> *Durch den Einsatz von KI haben wir bei der Leistungsfähigkeit einen Quantensprung erzielt und die Antwortzeiten auf Anfragen drastisch reduziert. Unsere Kunden verfügen nun über die entscheidenden Daten, die sie benötigen, um ihre Umsätze zu optimieren.*

Geschäftsführer, Unternehmen für Einzelhandelsanalysen

Das Unternehmen betreibt eine KI-gestützte Analyseplattform, die Einzelhändlern ermöglicht, ihre Vertriebsleistung und Entscheidungsprozesse zu optimieren. Das cloudbasierte System unterstützt mehr als 130 führende Konsumgütermarken in ganz Lateinamerika, integriert Verkaufs- und Bestandsdaten, analysiert das Kaufverhalten in Echtzeit und generiert prädiktive Erkenntnisse, um fundierte operative Entscheidungen zu ermöglichen.

Ziel des Unternehmens ist es, eine der zentralen Herausforderungen des Handels zu lösen: Angebot und Nachfrage präzise in Einklang zu bringen. In einem Umfeld, in dem sich Verbraucherpräferenzen fortlaufend verändern, ist die Fähigkeit, Vertriebs-, Lager- und Preisentscheidungen zu synchronisieren, von entscheidender Bedeutung. Ältere Systeme, die auf traditionellen Datenbanken basierten, konnten mit der wachsenden Menge und Geschwindigkeit der generierten Daten nicht Schritt halten. Mit dem starken Anstieg der Transaktionslast stieß das Unternehmen an seine Leistungsgrenzen. Abfragen verzögerten sich, Erkenntnisse kamen zu spät, und die Möglichkeit, Anpassungen in Echtzeit vorzunehmen, ging verloren. Für ein datenabhängiges Unternehmen stellte dies nicht nur ein technisches Problem dar, sondern ein Risiko für seine Wettbewerbsfähigkeit.

Um über reaktive Berichterstattung hinauszugehen, wurde die Analyseplattform mithilfe von KI vollständig neugestaltet. Maschinelles Lernen und adaptive Algorithmen verarbeiten nun täglich Milliarden von Datensätzen, erkennen neue Nachfragemuster und optimieren die Verteilung im großen Maßstab. Die Einführung KI-gestützter Simulationstools ermöglichte es, Preisstrategien, deren Berechnung zuvor Stunden dauerte, in Sekundenschnelle zu modellieren. Diese Systeme lernen kontinuierlich aus historischen Daten, testen neue Variablen und verfeinern Elastizitätsmodelle für Tausende von Produkten und Regionen. So entstand eine Plattform, die nicht nur beschreibt, was geschehen ist, sondern vorhersieht, was als Nächstes geschehen wird. Statische Analysen wurden in ein dynamisches, lernendes System der Intelligenz überführt.

Diese Entwicklung markiert den Übergang von reiner Analytik zu Kognition – einen Schritt in Richtung Allgemeiner KI auf Unternehmensebene. Durch die Integration von logischem Denken, Vorhersagen und Selbstoptimierung in ihre Abläufe hat die Organisation ein digitales Nervensystem geschaffen, das sich in Echtzeit an das Marktverhalten anpassen kann. Jede neue Transaktion wird zu einem Lernsignal und stärkt die Rückkopplungsschleifen, die die künftige Strategie bestimmen. Mit jeder Iteration nähert sich das System menschlichen Entscheidungsprozessen weiter an. Es ergänzt und verstärkt sie, anstatt sie zu ersetzen. Was als Suche nach schnelleren Erkenntnissen begann, hat sich zu einer Vision des kognitiven Unternehmens entwickelt – eines Unternehmens, in dem Intelligenz verteilt, kollaborativ und kontinuierlich selbstverbessernd ist.

Die fünf Merksätze

1. **Agentenbasierte KI als Unternehmensgrundlage.**
   Agentenbasierte KI bildet den entscheidenden Zwischenschritt auf dem Weg zur Allgemeinen KI, indem sie die Aufteilung von Aufgaben, die Spezialisierung und die Zusammenarbeit autonomer Agenten ermöglicht. Dieser modulare Ansatz schafft die Basis für flexiblere, skalierbare und menschenähnlichere Intelligenzsysteme.

2. **Konkurrierende Wege zur Allgemeinen KI.**
   Der Übergang zur Allgemeinen KI wird von zwei zentralen Hypothesen bestimmt: der Skalierungshypothese, nach der Allgemeine KI durch die Ausweitung bestehender Modelle entsteht, und der Diskontinuitätshypothese, die grundlegende neue Architekturen und Denkansätze erfordert. Führungskräfte sollten beide Entwicklungspfade in ihre strategische Planung und Investitionsentscheidungen einbeziehen.

3. **Eine Datenstrategie ist entscheidend.**
   Der Fortschritt hin zur Allgemeinen KI hängt von der Datenqualität, der Datenvielfalt und der Governance ab. Organisationen müssen robuste Datenrahmen, Technologien zum Schutz der Privatsphäre und die Einhaltung internationaler Standards priorisieren, um fortschrittliche KI-Fähigkeiten zu ermöglichen.

4. **Governance und Ethik im Zentrum.**
   Mit zunehmender Autonomie von KI-Systemen wird es entscheidend, ihr Handeln an menschlichen Werten auszurichten. Führungskräfte müssen auf starke Governance-Strukturen, ethische Aufsicht und wirksames Risikomanagement setzen, um regulatorischen Anforderungen und gesellschaftlichen Erwartungen gerecht zu werden.

5. **Sich auf Veränderungen in Arbeitswelt und Politik vorbereiten.**
   Der Übergang von agentenbasierter KI zu Allgemeiner KI wird die Rollenverteilung in der Arbeitswelt und die politischen Rahmenbedingungen neu definieren. Frühzeitige Investitionen in Qualifizierung, Veränderungsmanagement und kontinuierliche menschliche Kontrolle sind notwendig, um das Potenzial der Allgemeinen KI zu nutzen und Risiken zu begrenzen.

Kapitel Elf

Die Zukunft von EAI und Operations-Management

Das oft übersehene operative Management spielt im Unternehmen eine entscheidende Rolle, da es sämtliche Geschäftsbereiche unterstützt und das Wachstum maßgeblich vorantreibt. Mit der fortschreitenden Entwicklung betrieblicher Abläufe hat sich künstliche Intelligenz in diesem Bereich zu einem zentralen Bestandteil entwickelt. Dieses Kapitel beleuchtet die Entwicklung von KI im Unternehmenskontext und im operativen Management und zeigt, warum deren Verständnis unerlässlich ist, um die Vorteile agentenbasierter KI-Implementierungen vollständig zu erschließen.

In diesem Buch wurde zuvor die Konvergenz vertrauenswürdiger Daten und KI bei der Bereitstellung innovativer Erlebnisse sowie bei der Operationalisierung agentenbasierter Systeme untersucht. Wir haben zudem die Auswirkungen auf die Belegschaft und die internen Strukturen sowie Strategien zur Steuerung und Weiterentwicklung einer KI-gestützten Organisation betrachtet.

Klare Rollenbeschreibungen und messbare Leistungskennzahlen für menschliche wie digitale Arbeitskräfte sind dabei für den Erfolg entscheidend. Die Zuweisung klar umrissener, logisch strukturierter Aufgaben an die digitale Belegschaft ermöglicht effiziente Abläufe, während die Zusammenarbeit mehrerer spezialisierter Agenten die Bewältigung komplexerer Aufgaben erlaubt. Gleichzeitig bleibt die menschliche Aufsicht unverzichtbar, um Probleme oder Abweichungen frühzeitig zu erkennen und zu beheben. Diese Dynamik ähnelt dem Dirigieren eines Orchesters: Wenn jedes Instrument seinen Part spielt, erkennt der Dirigent sofort Disharmonien und sorgt für ein stimmiges Zusammenspiel.

Im weiteren Verlauf dieses Kapitels geht es um alle Aspekte des EAI-gestützten Betriebsmanagements – im Kern darum, wie sich Infrastruktur, Plattformen, Daten, Personal und digitale Agenten zu einem kontinuierlichen, harmonisierten Betriebsrhythmus verbinden lassen.

Ein anschauliches Beispiel liefert ein weltweit führendes Unternehmen im Bereich Gesundheitstechnologie, das basierend auf fortschrittlichen KI- und maschinellen Lernmodellen seine Abläufe durch die Einführung einer hochentwickelten, automatisierten Plattform für vorausschauende Wartung optimiert hat.

Fallstudie

Ein weltweit führendes Unternehmen im Gesundheitswesen

> *„Unser auf umfangreichen Datenmengen und fortschrittlichen KI-Modellen basierendes System für vorausschauende Wartung ermöglicht es uns, potenzielle Probleme zu identifizieren und zu beheben, bevor sie den klinischen Betrieb beeinträchtigen. Dadurch steigt die Zuverlässigkeit unserer Geräte, und sowohl die Behandlungsergebnisse als auch die Patientenzufriedenheit verbessern sich deutlich."*

Leitender Architekt, Service

Ein führendes Unternehmen im Bereich Gesundheitstechnologie stand vor wachsenden Herausforderungen bei der Instandhaltung seiner hochentwickelten medizinischen Bildgebungssysteme – MRT- und CT-Scanner, die für Diagnose und Behandlung von Patientinnen und Patienten unverzichtbar sind. Ein einzelnes MRT-Gerät kann täglich über eine Million Ereignisse protokollieren und rund 200.000 Sensormesswerte generieren – das entspricht Zehntausenden von Datenpunkten. In einem so komplexen und streng regulierten Umfeld dauert jedoch die Entwicklung und Zertifizierung von Medizinprodukten oft mehrere Jahre. Trotz der enormen Menge an Betriebsdaten waren diese bislang nicht so strukturiert, dass eine vorausschauende Wartung möglich gewesen wäre.

Die Umstellung war daher nicht nur technologisch, sondern auch operativ: Die Organisation musste ihre bestehenden Prozesse neu denken und modernisieren. Dies erforderte die Integration umfangreicher Datensätze aus Medizinprodukten, fortschrittlicher Analytik und Modellen des maschinellen Lernens, um potenzielle Ausfälle frühzeitig zu erkennen und zu verhindern, bevor sie die Patientenversorgung beeinträchtigen konnten.

Die Verbesserung der Behandlungsqualität steht dabei im Mittelpunkt. Mithilfe von KI können nun komplexe Datensätze effizient analysiert und Muster erkannt werden, die auf mögliche Probleme hinweisen, wodurch präventive Maßnahmen rechtzeitig eingeleitet werden. Das Unternehmen hat über 200 Datenströme – darunter Echtzeitprotokolle, Fehlerberichte und Leistungskennzahlen – in einem zentralen Data Warehouse zusammengeführt, das mehr als ein Jahrzehnt an Datenhistorie und 1,5 Petabyte kontinuierlich aktualisierter Informationen enthält. Prognosemodelle werten nun diesen riesigen Datensatz aus, um Anomalien frühzeitig zu erkennen, was eine proaktive Wartung ermöglicht und kostspielige Ausfallzeiten der Anlagen um 30 % reduziert. Das System hat dazu geführt, dass 50 % der CT-Servicefälle per Ferndiagnose behandelt und gelöst werden konnten und dass bei Problemen mit Geräten vor Ort eine Erstlösungsquote von 84 % erreicht wurde – was die Serviceeffizienz des Unternehmens steigert und die Patientenversorgung insgesamt verbessert.

178 / ENTERPRISE INTELLIGENCE: BUILDING THE SOVEREIGN CLOUD

Die Zukunft der AGI neu definieren: Jenseits des technischen Horizonts

Die Zukunft des Betriebsmanagements wird maßgeblich durch den Einsatz von KI geprägt sein und durch deren tiefgreifende Wirkung auf die Art, wie operative Prozesse gesteuert, überwacht und optimiert werden. Mehrere Schlüsselfaktoren sind von entscheidender Bedeutung:

1. **Übergang von reaktiven zu autonomen Operationen**
   Im operativen Bereich ist reaktives Handeln keine Option mehr. Zunehmende Cyberbedrohungen sowie die wachsende Komplexität vernetzter Systeme verlangen einen autonomen 24/7-Betrieb, der Probleme erkennt und behebt, bevor sie Auswirkungen auf Kunden oder Geschäftsprozesse haben.

2. **Entwicklung des Betriebsmanagements**
   In den letzten zwei Jahrzehnten haben sich die Abläufe grundlegend verändert und sind durch Fortschritte bei der Datenerfassung, -verwaltung und -analyse wesentlich effizienter geworden.

3. **Kernelemente der KI im operativen Bereich**
   Der Betrieb von EAI basiert auf fünf Schlüsselkomponenten: Datennutzung, Intelligenzformulierung, Entscheidungsprozesse, menschliches Eingreifen im Betriebskreislauf und ein umfassender Feedback-Lebenszyklus. Dieser ganzheitliche Ansatz ermöglicht es den Betriebsteams, den Übergang von manuellen zu automatisierten Methoden effektiv zu bewältigen.

4. **Anwendung von KI im Netzwerk- und Sicherheitsbetrieb**
   Agentenbasierte KI (Agentic AI) besitzt das Potenzial, Netzwerk- und Sicherheitsoperationen grundlegend zu transformieren. Im weiteren Verlauf dieses Kapitels werden praxisnahe Anwendungsfälle vorgestellt, die verdeutlichen, wie agentenbasierte Systeme die operative Effektivität und Reaktionsfähigkeit steigern.

5. **Transformationseffekte auf operative Kennzahlen**
   Diese Übergänge wirken sich direkt auf die wichtigsten Leistungskennzahlen aus – etwa die mittlere Wiederherstellungszeit (MTTR), die Serviceverfügbarkeit, die Anzahl von Ausfällen und Vorfällen sowie die Fähigkeit, eine Verfügbarkeit von „Five Nines“ (99,999 %) zu erreichen – den Goldstandard des operativen Geschäfts, denn 99,999 % Verfügbarkeit bedeutet nur etwa 5–6 Minuten Ausfallzeit pro Jahr. Die Einführung von KI bringt in all diesen Bereichen spürbare Verbesserungen und nachhaltige Effizienzgewinne.

Nachdem die Grundlagen nun umrissen sind, folgt im nächsten Abschnitt eine detaillierte Betrachtung der fünf zentralen Übergänge, die diesen Wandel prägen.

1. Von reaktiven zu autonomen Operationen

Die zunehmende Größe und Komplexität von Netzwerken und Unternehmensabläufen hat die Zeiten manueller, rein reaktiver Überwachung hinter sich gelassen. Früher griffen Einsatzteams erst ein, wenn ein Problem bereits aufgetreten war – ein Vorgehen, das eine überwiegend reaktive Haltung im operativen Betrieb zur Folge hatte. In den vergangenen Jahren hat sich dieses Paradigma deutlich verschoben: Verantwortliche im operativen Management bemühen sich zunehmend, die Fähigkeiten ihrer Teams zu erweitern und einen proaktiven Ansatz zu etablieren. Dieser Ansatz umfasst nicht nur das frühzeitige Erkennen potenzieller Zwischenfälle, sondern auch die aktive Identifizierung und Minderung von Risiken, bevor sie sich zu größeren Störungen entwickeln. Moderne Überwachungs- und Analysewerkzeuge haben dabei erhebliche Fortschritte gemacht. Sie ermöglichen die Erkennung minimaler Veränderungen in Systemleistung, Latenz oder betrieblichen Kennzahlen und dienen als Frühwarnsystem für potenzielle Störungen. Wie Feuermelder signalisieren sie kleinste Unregelmäßigkeiten, sodass Teams eingreifen können, bevor sich ein lokales Problem zu einem großflächigen Ausfall ausweitet.

Trotz dieser Entwicklungen bleiben viele operative Abläufe weiterhin reaktiv geprägt. Der Übergang zu autonomen Systemen stellt deshalb einen bedeutenden Paradigmenwechsel dar. Durch die Integration von Enterprise Artificial Intelligence (EAI) gewinnen Betriebsteams ein tieferes Verständnis der Systemdynamik, einschließlich der komplexen Wechselwirkungen zwischen Aktivitäten und Ereignissen, die zu Zwischenfällen führen können. Diese tiefergehenden Einblicke sind entscheidend, um die proaktiven Fähigkeiten der Teams zu stärken und zugleich Selbstheilungsmechanismen in Betriebsprozesse einzubinden. Autonome Systeme können bestimmte Probleme eigenständig erkennen und beheben, wodurch die operative Arbeitslast sinkt.

So entsteht Freiraum für Aufgaben mit höherer Priorität und für präventive Maßnahmen, die das Risiko künftiger Zwischenfälle deutlich reduzieren. Die Einführung von EAI bietet damit die Möglichkeit, das Betriebsmanagement grundlegend neu zu definieren – weg von einer reaktiven Haltung, hin zu einem strategisch ausgerichteten, proaktiven Rahmenwerk, das den Anforderungen eines dynamischen und komplexen Betriebsumfelds gerecht wird.

2. Die Evolution des Betriebsmanagements

Traditionell war das Betriebsmanagement eng mit Teams verbunden, die in isolierten Kontrollräumen arbeiteten und permanent Systemanzeigen und Warnmeldungen überwachten. Die Realität des operativen Managements hat sich jedoch deutlich weiterentwickelt. In der Anfangsphase basierten operative Prozesse weitgehend auf manueller Überwachung und Fehlerbehebung. Ein erster großer Fortschritt war die Einführung von Automatisierung durch Skripte, die es Teams ermöglichten, wiederkehrende Aufgaben zu standardisieren, Prozesse zu beschleunigen und Fehlerquoten zu senken.

Mit dem Aufkommen von Künstlicher Intelligenz (KI) und der Entwicklung agentenbasierter Systeme hat das operative Management ein neues Niveau erreicht. EAI-Tools (Enterprise Artificial Intelligence) können heute große Datenmengen zusammenführen, analysieren und miteinander in Beziehung setzen, um Anomalien und Musterveränderungen in Echtzeit zu erkennen. Dieses Verfahren ermöglicht es den Anwendern, verschiedene Datensätze miteinander zu korrelieren, was für die Ermittlung der Ursachen von Problemen zunehmend unerlässlich geworden ist. Die Fähigkeit, diese Ursachen schnell zu erkennen, ist entscheidend für die Minimierung von Betriebsausfällen.

Darüber hinaus eröffnet EAI neue Möglichkeiten der prädiktiven Analytik. Durch die Analyse von Trends, zeitlichen Abläufen und kausalen Zusammenhängen zwischen Symptomen und Ursachen lassen sich betriebliche Herausforderungen frühzeitig vorhersagen – und häufig sogar vermeiden. Damit vollzieht sich ein grundlegender Wandel: Das operative Management wird von einer reaktiven Disziplin zu einer strategisch-intelligenten Steuerungsebene.

Die fortschreitende Digitalisierung und der zunehmende Einsatz von KI verändern grundlegend, welche Fähigkeiten im Betriebsmanagement gefragt sind. Es gibt einen zunehmenden Trend dahin, dass Entwickler in operative Teams wechseln und Rollen wie Site Reliability Engineers (SREs) übernehmen. Diese Fachkräfte bündeln technisches Wissen, analytische Kompetenz und Programmierkenntnisse, um Systemvorfälle zu beheben, Ursachen zu identifizieren und in Echtzeit Lösungen zu entwickeln, die ein erneutes Auftreten derselben Probleme verhindern. Die Kombination aus Engineering-Know-how, Automatisierung und KI-gestützten Tools wird damit zur zentralen Qualifikation in modernen Betriebsabläufen.

Das Operationszentrum entwickelt sich zunehmend zu einem strategischen Lern- und Innovationsfeld – einem Ort, an dem KI-Anwendungen getestet, Prozesse optimiert und komplexe Herausforderungen unter realen Bedingungen bewältigt werden. Unternehmen, die ihre Betriebszentren gezielt als Innovations- und Talentzentren positionieren, schaffen sich einen klaren Wettbewerbsvorteil. Sie sind besser in der Lage, hochqualifizierte Fachkräfte zu gewinnen, weiterzubilden und langfristig zu halten. Diese Zentren dienen als Testumgebung, in der zukünftige Führungskräfte ihre Entwicklungs- und Problemlösungsfähigkeiten verbessern können, um so einen robusten Talentpool zu gewährleisten, der bereit ist, Führungsrollen innerhalb der Organisation zu übernehmen.

Da sich das Betriebsmanagement kontinuierlich weiterentwickelt, wird es zunehmend zu einem wertvollen Trainingsfeld für leitende Entwickler in Produkt- und Technologieunternehmen. Diese Fachkräfte sammeln dort unmittelbare Erfahrungen mit betrieblichen Vorfällen, die sowohl Kundenbeziehungen als auch Geschäftsprozesse betreffen, und können dieses Wissen gezielt in Produktentwicklung und strategische Entscheidungsprozesse einfließen lassen. Ein tiefes Verständnis für modernes, KI-gestütztes Betriebsmanagement prägt damit nachhaltig ihre Fähigkeit, technologische Innovationen mit operativer Exzellenz zu verbinden.

Was früher als reine Überwachungsfunktion galt, hat sich zu einem zentralen Nervensystem der Organisation entwickelt – einem Bereich, der Innovation, Stabilität und Lernkultur miteinander verknüpft. Das Betriebsmanagement ist heute ein Kompetenzzentrum für datenbasierte Weiterentwicklung, in dem operative Erfahrung und technologische Intelligenz zusammenfließen.

Ein Beispiel dafür liefert ein südafrikanischer Einzelhändler, der systematisch analysiert, wie seine Teams KI einsetzen, um Prozesse zu verfeinern, Abläufe zu beschleunigen und die Gesamtleistung zu optimieren.

Fallstudie

Führender Einzelhändler für Konsumgüter in Afrika

> *Durch die Integration von KI in die Produkttestprozesse lässt sich für jede getestete Anwendung im gesamten Omnichannel-Portfolio innerhalb von Sekunden auf Sprint-, Release- und Feature-Ebene nachvollziehen, an welchem Punkt sich jedes Projekt im Entwicklungszyklus befindet.*
>
> SQA Manager

Mit Tausenden von Filialen in Südafrika und Präsenzen in sieben weiteren Ländern betreibt dieser führende afrikanische Einzelhändler für Konsumgüter ein umfangreiches Netzwerk digitaler Omnichannel-Anwendungen. Neben dem stationären Handel verfügt das Unternehmen über eine starke Online-Präsenz für Lebensmittel, Haushaltswaren und Bekleidung sowie über eine mobile App für die ultraschnelle Lieferung von Lebensmitteln.

In einem hart umkämpften Markt ist es entscheidend, sowohl die Verfügbarkeit in den Filialen als auch die Stabilität der digitalen Systeme sicherzustellen. Um die Markteinführungszeiten zu verkürzen und die Freigabezyklen zu beschleunigen, hat der Einzelhändler KI in seine Test- und Freigabeprozesse integriert. Die Einführung von KI-gestützter Testfallerstellung hat es ermöglicht, die Automatisierung bereits während der zweiwöchigen Sprints einzusetzen. Das Ergebnis war eine Steigerung der Automatisierungsabdeckung von rund 65 % auf etwa 95 %.

Innerhalb von nur acht Wochen wurden für fast zwanzig Anwendungen im Omnichannel-Bereich vier bis fünf Releases pro Woche abgeschlossen. Dadurch konnten die Zykluszeiten um 43 % verkürzt, die Release-Frequenz um das 60-fache erhöht und die Testabdeckung signifikant erweitert werden – mit direkter Wirkung auf schnellere Erkenntnisse und verkürzte Markteinführungszeiten. Angesichts dieser Ergebnisse prüft das Unternehmen jetzt, wie KI auch in anderen Geschäftsbereichen eingesetzt werden kann, um Effizienz, Agilität und Innovationskraft weiter zu steigern.

3. Kernkomponenten KI-gesteuerter Operationen

Da Unternehmen ihre Betriebsabläufe modernisieren und fortschrittliche KI-Funktionen integrieren, müssen die zentralen technischen und strukturellen Grundlagen sorgfältig berücksichtigt werden.

A. Die Datenebene

Der Aufbau einer einheitlichen und zugänglichen Datenebene ist die Voraussetzung für effiziente, KI-gestützte Betriebsmodelle. Traditionelle Betriebsumgebungen litten häufig unter fragmentierten Daten, die über verschiedene Systeme, Service-Management-Tools und organisatorische Silos verteilt waren. Durch die Konsolidierung dieser Datensätze und ihre Nutzung durch KI-Agenten werden Transparenz, Kontextverständnis und Analysetiefe erheblich verbessert. Ein Beispiel ist die Integration von Sicherheits- und Netzwerkbetriebsdaten, wodurch sich Sicherheitsprotokolle gemeinsam mit Netzwerkkennzahlen auswerten lassen. Diese kombinierte Sicht ermöglicht es den Teams, Abläufe ganzheitlich zu verstehen, Anomalien schneller zu erkennen und Probleme in Echtzeit zu beheben.

B. Die Intelligenzebene

In dieser Ebene kommen Sprachmodelle, maschinelles Lernen sowie generative und agentenbasierte KI-Systeme zum Einsatz. Hier werden Daten korreliert, Wissen aufgebaut und Handlungsempfehlungen abgeleitet, sodass KI-Anwendungen aktiv mit menschlichen Analysten zusammenarbeiten. Generative KI trägt wesentlich dazu bei, das Lagebewusstsein zu verbessern, komplexe Muster zu interpretieren und schnellere, fundierte Entscheidungen im Betriebszentrum zu unterstützen. Agentenbasierte KI-Anwendungen ermöglichen autonome Betriebsabläufe.

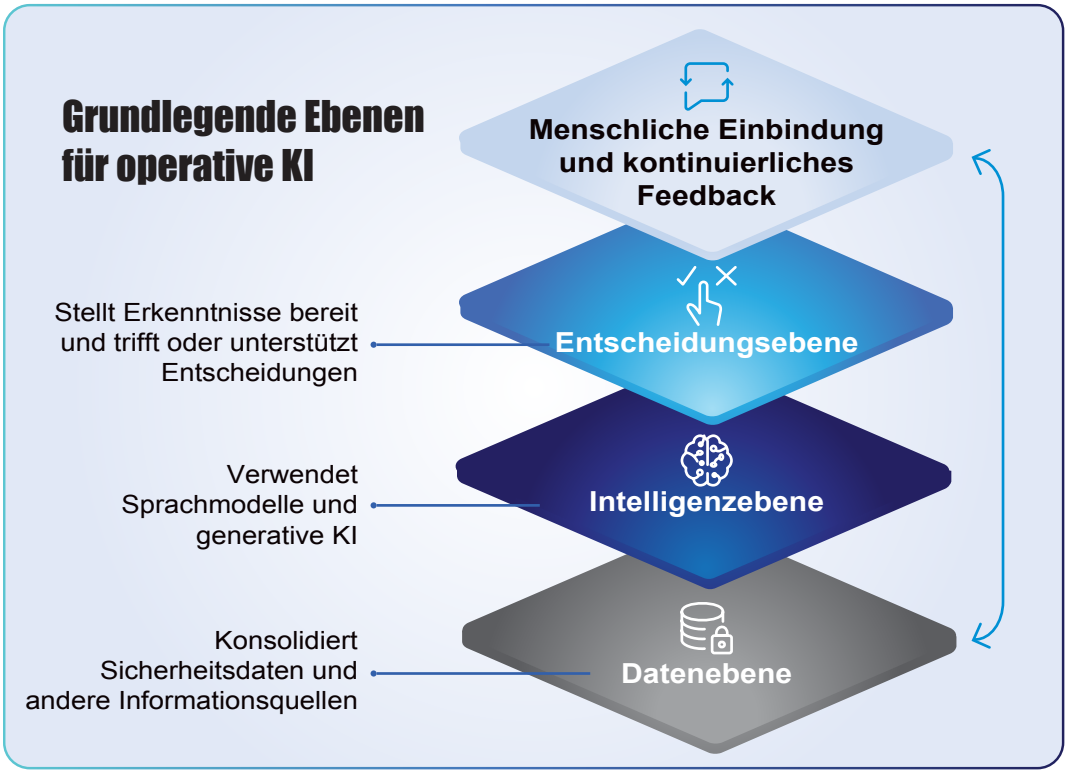
C. Die Entscheidungsebene

Daten zu analysieren ist das eine – auf dieser Grundlage eigenständig Entscheidungen zu treffen, das andere. In der frühen Phase der KI-Integration in Betriebsumgebungen lieferten KI-Systeme in erster Linie analytische Erkenntnisse, während die endgültige Entscheidungsverantwortung bei den menschlichen Bedienern lag. Mit zunehmender Reife und dem Aufbau klarer Governance-Strukturen beginnen Organisationen, ihren KI-Systemen zu gestatten, vordefinierte oder risikoarme Entscheidungen autonom zu treffen. Je weiter sich diese Systeme entwickeln, desto besser werden sie in der Lage sein, innerhalb der von Menschen gesetzten Rahmenbedingungen komplexe, wiederkehrende Entscheidungen konsistent und nachvollziehbar zu übernehmen.

D. Human-in-the-Loop und kontinuierliches Feedback

Auch in stark automatisierten Betriebsumgebungen bleibt der Mensch ein zentraler Faktor. Seine Rolle verschiebt sich hin zu Aufsicht, Kontextinterpretation, strategischer Steuerung und Eingriff in Ausnahme- oder Grenzfälle, die Urteilsvermögen erfordern.

Ebenso wichtig ist die Feedback-Schleife. Bei der Behebung von Vorfällen und der Analyse ihrer Ursachen werden die gewonnenen Erkenntnisse in das System zurückgespeist, um einen kontinuierlichen Lernprozess zu gewährleisten. Dieser Mechanismus führt dazu, dass sich Wiederherstellungszeiten stetig verkürzen und die mittlere Reparaturzeit (MTTR) sinkt. Gleichzeitig werden Wiederholungsvorfälle reduziert und die Einheitlichkeit der Reaktionen im gesamten Betrieb gestärkt. Wenn Betriebsteams Ursachen nachhaltig beseitigen, sinkt die Gesamtzahl der Vorfälle, wodurch Kapazitäten für proaktive und innovative KI-Arbeit frei werden – Arbeit, die die Widerstandsfähigkeit, Effizienz und Anpassungsfähigkeit des gesamten Betriebszentrums weiter erhöht. Im Zusammenspiel bilden diese zentralen Komponenten die Grundlage für KI-gesteuerte Betriebsabläufe, die notwendig sind, um agentenbasierte KI-Systeme effektiv zu unterstützen und weiterzuentwickeln.



Grundlegende Ebenen, die behandelt werden müssen

4. Agentenbasierte KI im Netzwerk- und Sicherheitsbetrieb

The role of the Network Operations Center (NOC) operator is one of high stress and tight deadlines Die Arbeit im Network Operations Center (NOC) ist geprägt von hohem Druck, engen Fristen und der ständigen Herausforderung, in einem Meer von Daten die sprichwörtliche Nadel im Heuhaufen zu finden. Der Erfolg eines Operators wird häufig daran gemessen, wie schnell ein Problem erkannt, analysiert und behoben wird. Vor der Einführung agentenbasierter KI bestand die Aufgabe der NOC-Teams darin, manuell Korrelationen zwischen Vorfällen zu suchen, ähnliche Probleme in umfangreichen Datenbanken zu identifizieren und aus diesen Mustern Hinweise für eine mögliche Lösung abzuleiten. Das folgende Beispiel verdeutlicht die Vorteile einer Partnerschaft mit einer agentenbasierten KI-Anwendung. Der KI-Agent durchsucht Datenbanken automatisch, findet Zusammenhänge zwischen früheren und aktuellen Ereignissen, leitet Ursachen ab und unterstützt den Operator bei der schnelleren und gezielteren Problemlösung – nicht mehr in Stunden, sondern in Minuten.

5. Transformationsauswirkungen auf den Unternehmensbetrieb

Ein prägnantes Beispiel: Es ist 2:47 Uhr morgens, als ein globales E-Commerce-NOC (Network Operations Center) Datenbank-Latenzspitzen feststellt, die die Zahlungsabwicklung beeinträchtigen. Das Überwachungssystem löst automatisch einen Alarm aus, aber die agentenbasierte KI verändert alles.

Von Stunden zu Minuten mit KI-Agenten

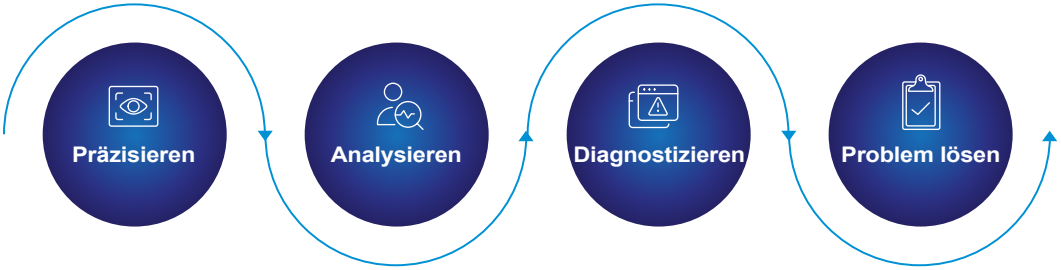
Klarstellung: Der KI-Agent führt sofort eine automatisierte Ereigniskorrelation in der Flut eingehender Warnmeldungen durch und trennt präzise zwischen der eigentlichen Ursache – einem Batch-Job-Konflikt – und den zahlreichen Symptomereignissen wie Latenzspitzen, Timeout-Fehlern und Warteschlangen-Rückstau. Stattdessen entsteht aus zwanzig unübersichtlichen Alarmmeldungen entsteht für den Operator ein klares, verständliches Lagebild:
„Datenbankkonflikt festgestellt; wahrscheinliche Ursache identifiziert."

Analysieren: Der KI-Agent fragt Sicherheitsprotokolle, Netzwerktelemetrie, Anwendungsspuren und Änderungsmanagement-Datenbanken ab und durchsucht gleichzeitig jahrelange historische Vorfallsdaten in einer Vektordatenbank. Innerhalb weniger Sekunden identifiziert das System 847 ähnliche Muster und markiert eine 89-prozentige Übereinstimmung mit zwei früheren Vorfällen, deren Symptome, Zeitpunkte und Systemverhalten nahezu identisch waren. Es liefert den entscheidenden Beweis: Mit dem Änderungsticket CHG-45209 wurde ein Batch-Job-Zeitplan geändert, der nun während der Spitzenzeiten der Transaktionen ausgeführt wird.

Diagnose: Die KI begnügt sich nicht damit, bloße Korrelationen herzustellen – sie identifiziert aktiv die Ursache, indem sie die Zeitpunkte der Batch-Jobs mit den Transaktionsvolumenmustern abgleicht und präzise angibt, wann und wo die Störung aufgetreten ist. Sie präsentiert die Diagnose: „Batch-Job user_analytics_aggregation_v2 steht im Konflikt mit der Echtzeit-Zahlungsverarbeitung; gleiches Muster wie bei den Vorfällen #3421 und #11203"

Auflösung: An diesem Punkt tritt wieder der Mensch in den Prozess ein.
Der KI-Agent schlägt drei Lösungsoptionen vor, basierend auf früheren erfolgreichen Eingriffen, sortiert nach Erfolgswahrscheinlichkeit und Risiko. Der Operator bewertet die Vorschläge, berücksichtigt den aktuellen Kontext – es ist 2:52 Uhr morgens, außerhalb der geplanten Wartungsfenster – und trifft die Entscheidung: Die KI führt die genehmigten Aktionen unter kontinuierlicher Überwachung aus.

Gesamtzeit vom Alarm bis zur Lösung: 14 Minuten. Ohne agentenbasierte KI hätte derselbe Prozess über vier Stunden gedauert.



Agentenbasierte KI verkürzt die Zeit bis zur Lösung

Die agentenbasierte KI übernahm in jeder Phase die Hauptarbeit – von der automatisierten Korrelation über die intelligente Analyse bis hin zur präzisen Diagnose und geführten Problemlösung. Das führte zu einer deutlichen Reduzierung der mittleren Reparaturzeit (MTTR). Zudem fließt jeder so behandelte Vorfall in die historische Datenbank ein und ermöglicht künftig noch präzisere Reaktionen.

Jedes Betriebszentrum wird anhand definierter Leistungskennzahlen (KPIs) bewertet. Traditionelle Kennzahlen konzentrieren sich in der Regel auf die Dienstverfügbarkeit – viele Organisationen streben 99,999 % (Five Nines) an – sowie auf das Volumen und die Schwere von Vorfällen, seien sie groß oder klein. 99,999 % Verfügbarkeit bedeutet nur etwa 5–6 Minuten Ausfallzeit pro Jahr. Weitere wichtige Kennzahlen sind die mittlere Reparaturzeit (Mean Time to Repair, MTTR), die maximal tolerierbare Wiederherstellungszeit (Recovery Time Objective, RTO) und das Wiederherstellungspunktziel (Recovery Point Objective, RPO).

Diese Kennzahlen sind zwar zentral, messen aber letztlich nur die Reaktion auf äußere Einflüsse. Sie zeigen, wie gut ein Team auf Störungen reagiert, nicht, wie effektiv es diese verhindert. Da KI zunehmend in die Betriebsabläufe integriert wird, müssen Unternehmen beginnen, proaktive Kennzahlen einzuführen. Zum Beispiel:

- Wie viele Vorfälle wurden durch KI erkannt oder abgemildert, bevor sie eskalierten?
- Wie schnell erkannten agentenbasierte KI-Systeme frühe Anzeichen von Fehlern oder Kompromittierungen?
- Welcher Prozentsatz aller Vorfälle wird autonom von KI-Agenten bearbeitet?

Die separate Erfassung der Auswirkungen agentenbasierter KI hilft Organisationen, den tatsächlichen Transformationseffekt ihrer Einführung zu verstehen. Der Erfolg sollte sich in messbaren Verbesserungen widerspiegeln, wie zum Beispiel:

- Eine Verringerung der Gesamtzahl der Vorfälle, insbesondere der schwerwiegenden.
- Schnellere Erkennung und Behebung von Vorfällen
- Höhere betriebliche Resilienz durch vorausschauende Überwachung und automatisierte Reaktion

Generative und agentenbasierte KI bieten die Möglichkeit, die sprichwörtliche Nadel im Heuhaufen zu finden – Ursachen schnell zu erkennen und ein früheres Eingreifen zu ermöglichen.

Die Weiterentwicklung der Betriebsabläufe muss dabei stets mit der Einführung von KI einhergehen. Allzu oft konzentrieren sich Organisationen auf den Einsatz von KI-Modellen, während sie die Modernisierung ihrer Betriebszentralen vernachlässigen. Ein Team, das weiterhin auf manuelle Überwachung und traditionelle Arbeitsabläufe setzt, kann mit einem Unternehmen, das auf intelligente, KI-gestützte Prozesse umstellt, nicht Schritt halten.

Da Cyberkriminelle zunehmend KI nutzen, um Cyberangriffe zu automatisieren und zu verstärken, müssen sich auch Sicherheitszentralen parallel weiterentwickeln und KI einsetzen, um Parität zu wahren und in Maschinengeschwindigkeit zu reagieren. Dieser Wandel ist nicht optional – er ist unerlässlich, um im Zeitalter KI-gesteuerter Unternehmensabläufe Wettbewerbsfähigkeit, Widerstandsfähigkeit und Vertrauen zu sichern.

In dieser neuen Ära ist die Fähigkeit, Organisationsdaten effektiv zu verwalten, der entscheidende Faktor, um das volle Potenzial der KI auszuschöpfen. Der Einsatz von KI zur Verwaltung, Steuerung und kontinuierlichen Optimierung von Datenoperationen ist dafür unverzichtbar. Da sich operative Zentren zunehmend zu Nervenzentren für Innovation und Resilienz entwickeln, müssen Datenmanagement und Datenstrategie im Mittelpunkt jeder Managementagenda stehen. Unternehmen, die Daten nicht nur als Ressource, sondern als strategisches Gut begreifen und KI nutzen, um Integrität und Verfügbarkeit zu gewährleisten, sind am besten positioniert, um das volle transformative Potenzial von KI in großem Maßstab zu realisieren.

Im folgenden Fallbeispiel nutzt ein führender Hersteller das Potenzial seiner Daten, um tiefere Einblicke in seine Prozesse zu gewinnen, Kosten zu senken und die Abhängigkeit von manueller Arbeit zu reduzieren.

Fallstudie

North Star BlueScope Steel

North Star BlueScope Steel

North Star BlueScope Steel, eine Tochtergesellschaft des australischen Unternehmens BlueScope, produziert und liefert warmgewalzte Stahlbänder für Coil-Verarbeiter, Kaltbandhersteller, Rohr- und Schlauchproduzenten, Erstausrüster und Stahlservicezentren. Das 1997 gegründete Unternehmen ist der größte Stahlschrott-Recycler in Ohio und recycelt jährlich fast 1,5 Millionen Tonnen Stahlschrott.

Um seine Kostendaten und Arbeitsabläufe präziser zu verstehen, benötigte North Star BlueScope Steel ein effizienteres Werkzeug, das marktorientierte Analysen und detaillierte Einkaufsübersichten ermöglicht. Die neue Technologie sollte den bisher manuellen Aufwand eliminieren und Daten automatisch aus einer Vielzahl von Quellen erfassen – darunter interne Datenbanken und die Elektrolichtbogenöfen (EAF) des Unternehmens – um Personal gezielter einzusetzen, Ressourcen zu optimieren und die Kundenbedürfnisse besser zu erfüllen.

Das Unternehmen entschied sich für den Einsatz intelligenter Datenverarbeitung und Analytik, um Informationen automatisch abzurufen, zu kombinieren, zu prüfen und zu analysieren. Die eingeführte Lösung ermöglicht es North Star BlueScope Steel, Algorithmen auf die extrahierten Daten anzuwenden und daraus automatisiert monatliche Abschlussberichte zu erstellen – wodurch die Abhängigkeit von manuellen Eingriffen deutlich reduziert wird. Durch die Nutzung von Daten und Analysen kann das Unternehmen nun monatliche Kennzahlen vergleichen und untersuchen, wie sich Ereignisse wie Produktionsverzögerungen oder Engpässe auf die Rentabilität auswirken. Darüber hinaus setzt North Star BlueScope Steel auf das Internet der Dinge (IoT), um Analysen direkt in die von seinen Sensoren erfassten Datenpunkte zu integrieren. Dies umfasst den Stromverbrauch, Wettermuster, Materialeinsatz und Stahlpreise – mit dem Ziel, ein besseres Verständnis für künftige Anforderungen und Absatzpotenziale zu gewinnen.

Die fünf Merksätze

1. **Den Wandel hin zu autonomen Betriebsabläufen vorantreiben.**
   Das Unternehmen muss über die reine Überwachung hinausgeführt werden, indem in KI-gestützte, selbstheilende Systeme investiert wird, die Störfälle proaktiv verhindern und die Leistung optimieren.

2. **Mensch-KI-Zusammenarbeit stärken.**
   Teams sollten gezielt auf die Zusammenarbeit mit der KI vorbereitet werden, Governance-Frameworks sind zu schaffen, menschliche Kontrolle zu sichern und kontinuierliche Feedbackschleifen zu etablieren, um Sicherheit und Innovation zu fördern.

3. **Betriebszentren zu Innovationszentren entwickeln.**
   Das Betriebszentrum muss sich von einer Supportfunktion zu einem strategischen Motor für Innovation und Talententwicklung entwickeln, um Dreh- und Angelpunkt für den Einsatz und die Erprobung von KI zu werden.

4. **KI muss auf Vorstandsebene Priorität erhalten.**
   Die Integration von KI in Betriebsabläufe ist als entscheidender Wettbewerbsvorteil zu betrachten. Die Aufmerksamkeit der Führungskräfte und die verfügbaren Ressourcen sollten darauf ausgerichtet werden, die Einführung zu beschleunigen, die Sicherheit zu erhöhen und das Vertrauen in die Organisation stärken.

5. **Kennzahlen für das KI-Zeitalter neu definieren.**
   Von traditionellen, reaktiven Kennzahlen sollte sich das Unternehmen verabschieden und neue Kennzahlen einführen, die den tatsächlichen Einfluss von KI erfassen, wie etwa verhinderte Störfälle, Reaktionsgeschwindigkeit und autonome Maßnahmen, um den Wert präzise zu messen und kontinuierliche Verbesserungen zu fördern.

Fußnoten

[1] Foundry Research sponsore by OpenText, „MarketPulse Survey: The Role of GenAI in Modernizing Content Management“ (MarketPulse-Umfrage: Die Rolle von GenAI bei der Modernisierung des Content-Managements), Mai 2025.

[2] Philip Miller, „Unlocking Unstructured Data: Fueling AI with Insights“ (Potenziale unstrukturierter Daten erschließen: KI mit Erkenntnissen speisen), *Dataversity*, 3. Juni 2025, https://www.dataversity.net/articles/unlocking-unstructured-data-fueling-ai-with-insights/.

[3] Ibid.

[4] „More Than 80% of Enterprises Will Have Used Generative AI APIs or Deployed Generative AI Applications by 2026“ (Mehr als 80 % der Unternehmen werden bis 2026 Generative-AI-APIs genutzt oder generative KI-Anwendungen implementiert haben), *Gartner Press Release*, 11. Oktober 2023, www.gartner.com/en/newsroom/press-releases/2023-10-11-gartner-says-more-than-80-percent-of-enterprises-will-have-used-generative-ai-apis-or-deployed-generative-ai-enabled-applications-by-2026.

[5] „AI Governance Software Spend Will See 30% CAGR From 2024 to 2030“ (Ausgaben für KI-Governance-Software wachsen von 2024 bis 2030 mit 30 % jährlich), *Forrester Blog*, 13. November 2024, www.forrester.com/blogs/ai-governance-software-spend-will-see-30-cagr-from-2024-to-2030/.

[6] McKinsey & Company, „The State of AI in Early 2024: Gen AI Adoption Spikes and Starts to Generate Value“ (Der Stand der KI Anfang 2024: Gen-AI-Nutzung steigt sprunghaft und beginnt Wert zu schaffen), *QuantumBlack by McKinsey*, 30. Mai 2024, www.mckinsey.com/capabilities/quantumblack/our-insights/the-state-of-ai-2024.

[7] Ibid.

[8] Accenture, „New Accenture Research Finds that Companies with AI-Led Processes Outperform Peers“ (Neue Accenture-Studie zeigt: Unternehmen mit KI-gesteuerten Prozessen übertreffen ihre Wettbewerber), *Accenture*, 10. Oktober 2024, https://newsroom.accenture.com/news/2024/new-accenture-research-finds-that-companies-with-ai-led-processes-outperform-peers.

[9] „What is Artificial Intelligence (AI)?“ (Was ist Künstliche Intelligenz (KI)?), *International Organization for Standardization*, 31. Januar 2024, https://www.iso.org/artificial-intelligence/what-is-ai?.

[10] Melissa Russell, „How can I learn artificial intelligence?“ (Wie kann ich Künstliche Intelligenz lernen?), *Harvard*, 8. April 2025, https://extension.harvard.edu/blog/how-can-i-learn-artificial-intelligence/#What-is-Artificial-Intelligence.

[11] Sofia Samoili, Montserrat Lopez Cobo, Blagoj Delipetrev, Fernando Martinez-Plumed, Emilia Gomez Gutierrez, and Giuditta De Prato, „AI Watch, Defining Artificial Intelligence 2.0: Towards an operational definition and taxonomy for the AI Landscape“ (AI Watch, Definition von Künstlicher Intelligenz 2.0: Auf dem Weg zu einer operativen Definition und Taxonomie der KI-Landschaft), *Publications Office of the European Union*, 2021.

[12] Ibid.

[13] Tim Mucci and Cole Stryker, „What is artificial superintelligence?“ (Was ist künstliche Superintelligenz?), IBM, 22. Juli 2025, https://www.ibm.com/think/topics/artificial-superintelligence.

[14] Arend Hintze, „Understanding the four types of AI, from reactive robots to self-aware beings“ (Die vier Typen von KI verstehen – von reaktiven Robotern bis zu selbstbewussten Wesen), *The Conversation*, 13. November 2016, https://theconversation.com/understanding-the-four-types-of-ai-from-reactive-robots-to-self-aware-beings-67616

[15] A. M. Turing, „Computing Machinery and Intelligence“ (Rechenmaschinen und Intelligenz), *Mind*, Volume LIX, Issue 236, Oktober 1950, https://doi.org/10.1093/mind/lix.236.433.

[16] J. McCarthy, M. L. Minsky, N. Rochester, & C. E. Shannon, „A proposal for the Dartmouth summer research project on artificial intelligence“ (Ein Vorschlag für das Dartmouth-Sommerforschungsprojekt zur Künstlichen Intelligenz), *Dartmouth College*, 1955, https://ojs.aaai.org/aimagazine/index.php/aimagazine/article/view/1904.

[17] Ben Lutkevich, „What is AI Winter? Definition, History and Timeline“ (Was ist der KI-Winter? Definition, Geschichte und Zeitachse), *Tech Target*, 26. August 2024, https://www.techtarget.com/searchenterpriseai/definition/AI-winter.

[18] D. Crevier, „*AI: The tumultuous history of the search for artificial intelligence*“ (KI: Die bewegte Geschichte der Suche nach künstlicher Intelligenz), Basic Books, 1993.

[19] Ibid.

[20] S. J. Russell & P. Norvig, Artificial intelligence: *A modern approach (Künstliche Intelligenz: Ein moderner Ansatz)*, 4th ed., Pearson, 2021.

[21] Yann LeCun, Yoshua Bengio, and Geoffrey Hinton, „Deep learning“ (Deep Learning), *Nature*, 521(7553), 436–444, 2015, https://doi.org/10.1038/nature14539.

[22] Melissa Russell, „How can I learn artificial intelligence?“ (Wie kann ich Künstliche Intelligenz lernen?), *Harvard*, 8. April 2025, https://extension.harvard.edu/blog/how-can-i-learn-artificial-intelligence/#What-is-Artificial-Intelligence.

[23] „Gartner Survey Reveals GenAI Attacks Are on the Rise“ (Gartner-Umfrage zeigt: Angriffe mit GenAI nehmen zu), *Gartner Inc.*, 22. September 2025, https://www.gartner.com/en/newsroom/press-releases/2025-09-22-gartner-survey-reveals-generative-artificial-intelligence-attacks-are-on-the-rise.

[24] Akshay Joshi, Giulia Moschetta, and Ellie Winslow, „Global Cybersecurity Outlook 2025 Insight Report“ (Globaler Ausblick auf die Cybersicherheit 2025 – Insight Report), *World Economic Forum in Collaboration with Accenture*, Januar 2025, https://reports.weforum.org/docs/WEF_Global_Cybersecurity_Outlook_2025.pdf.

[25] Shuli Jiang, Swanand Ravindra Kadhe, Yi Zhou, Ling Cai, and Nathalie Baracaldo, „Forcing Generative Models to Degenerate Ones: The Power of Data Poisoning Attacks“ (Generative Modelle zu degenerierten Modellen zwingen: Die Wirkung von Data-Poisoning-Angriffen), Cornell University, arXiv:2312.04748, 7. Dezember 2023, https://arxiv.org/abs/2312.04748.

[26] B. Biggio, B. Nelson, and P. Laskov, „Poisoning Attacks Against Support Vector Machines“ (Vergiftungsangriffe auf Support Vector Machines), *Proceedings of the 29th International Conference on Machine Learning (ICML)*, 2012.

[27] Tianyu Gu, Brendan Dolan-Gavitt, and Siddharth Garg, „BadNets: Identifying Vulnerabilities in the Machine Learning Model Supply Chain“ (BadNets: Schwachstellen in der Lieferkette von Machine-Learning-Modellen erkennen), Cornell University, arXiv:1708.06733, 11. März 2019, https://arxiv.org/abs/1708.06733.

[28] Ian J. Goodfellow, Jonathon Shlens, and Christian Szegedy, „Explaining and Harnessing Adversarial Examples“ (Adversarielle Beispiele erklären und nutzbar machen), Cornell University, arXiv:1412.6572, 20 März 2015, https://arxiv.org/abs/1412.6572.

[29] Wencheng Yang, Song Wang, Di Wu et al., „Deep Learning Model Inversion Attacks and Defenses: A Comprehensive Survey“ (Model-Inversion-Angriffe und -Abwehr bei Deep-Learning-Modellen: Ein umfassender Überblick), Cornell University, arXiv:2501.18934, 30. April 2025, https://arxiv.org/abs/2501.18934.

[30] Ninareh Mehrabi, Fred Morstatter, Nripsuta Saxena, Kristina Lerman, and Aram Galstyan, „A Survey on Bias and Fairness in Machine Learning“ (Ein Überblick über Bias und Fairness im maschinellen Lernen), Cornell University, arXiv:1908.09635, 25. Januar 2022, https://arxiv.org/abs/1908.09635.

[31] ISO/IEC 27001:2022, „Information Security, Cybersecurity and Privacy Protection—Information Security Management Systems—Requirements“ (Informationssicherheit, Cybersicherheit und Datenschutz – Informationssicherheits-Managementsysteme – Anforderungen), International Organization for Standardization.

[32] Sandeep Kumar Jangam, „Importance of Encrypting Data in Transit and at Rest Using TLS and Other Security Protocols and API Security Best Practices“ (Bedeutung der Verschlüsselung von Daten bei der Übertragung und im Ruhezustand mit TLS und anderen Sicherheitsprotokollen sowie Best Practices für API-Sicherheit), *International Journal of AI, BigData, Computational and Management Studies*, 4(3), 82–91, 2023, https://ijaibdcms.org/index.php/ijaibdcms/article/view/242/.

[33] Qiang Yang, Yang Liu, Tianjian Chen, and Yongxin Tong, „Federated Machine Learning: Concept and Applications“ (Föderiertes maschinelles Lernen: Konzept und Anwendung), *ACM Digital Library*, 28. Januar 2019, https://dl.acm.org/doi/10.1145/3298981.

[34] Europäisches Parlament und Rat, „Verordnung (EU) 2016/679 des Europäischen Parlaments und des Rates vom 27. April 2016 zum Schutz natürlicher Personen bei der Verarbeitung personenbezogener Daten – Artikel 17 (Recht auf Löschung)“, 2016, https://eur-lex.europa.eu/eli/reg/2016/679/oj.

[35] Scott Rose, Oliver Borchert, Stu Mitchell, and Sean Connelly, „Zero Trust Architecture“ (Zero-Trust-Architektur), *NIST Special Publication 800-207*, August 2020, https://doi.org/10.6028/NIST.SP.800-207.

[36] Alexey Kurakin, Ian Goodfellow, and Samy Bengio, „Adversarial Machine Learning at Scale“ (Adversariales maschinelles Lernen im großen Maßstab), Cornell University, arXiv:1611.01236, 11. Februar 2017, https://arxiv.org/abs/1611.01236.

[37] Yusuke Uchida, Yuki Nagai, Shigeyuki Sakazawa, and Shin'ichi Satoh, „Embedding Watermarks into Deep Neural Networks“ (Wasserzeichen in tiefe neuronale Netze einbetten), Cornell University, arXiv:1701.04082, 20. April 2017, https://arxiv.org/abs/1701.04082.

[38] Forrester, „AI Governance Software Spend Will See 30 % CAGR From 2024 to 2030“ (Ausgaben für KI-Governance-Software werden von 2024 bis 2030 um 30 % jährlich wachsen), *Forrester Research*, 13. November 2024, https://www.forrester.com/blogs/ai-governance-software-spend-will-see-30-cagr-from-2024-to-2030/.

[39] „Information Governance Reference Model“ (Information-Governance-Referenzmodell), EDRM, http://www.edrm.net/projects/igrm.

[40] „Key Regulatory and Industry Initiatives“ (Zentrale regulatorische und branchenspezifische Initiativen), *Capgemini*, https://web.archive.org/web/20141105171058/https://www.worldpaymentsreport.com/kriis#Heat-Map-of-KRIIs-Global-and-Regional.

[41] Gartner, Inc., „Gartner Poll Finds 55% of Organizations Have an AI Board“ (Gartner-Umfrage: 55 % der Unternehmen verfügen über ein KI-Gremium), Press Release, 26. Juni 2024, https://www.gartner.com/en/newsroom/press-releases/2024-06-26-gartner-poll-finds-55-percent-of-organizations-have-an-ai-board

[42] Ibid.

[43] Alex Edquist, Liz Grennan, Sian Griffiths, and Kayvaun Rowshankish, „Data ethics: What it means and what it takes“ (Datenethik: Was sie bedeutet und was sie erfordert), *McKinsey & Company*, 23. September 2022, https://www.mckinsey.com/capabilities/mckinsey-digital/our-insights/data-ethics-what-it-means-and-what-it-takes.

[44] James Moor, „What is Computer Ethics?“ (Was ist Computerethik?), *Metaphilosophy*, 16(4), 266–275, 1985, https://doi.org/10.1111/j.1467-9973.1985.tb00173.x.

[45] Unesco, „Recommendation on the Ethics of Artificial Intelligence“ (Empfehlung zur Ethik der Künstlichen Intelligenz), *UNESCO.org*, 2022, https://www.unesco.org/en/articles/recommendation-ethics-artificial-intelligence.

[46] Ibid.

[47] Ibid.

[48] OECD, „Recommendation of the Council on Artificial Intelligence“ (Empfehlung des Rates zur Künstlichen Intelligenz), OECD/LEGAL/0449, 2019.

[49] European Commission, „Regulation (EU) 2024/1689 on Artificial Intelligence“ (Verordnung (EU) 2024/1689 über Künstliche Intelligenz), 2024.

[50] NIST, „Artificial Intelligence Risk Management Framework (AI RMF 1.0)“ (Rahmenwerk zum Risikomanagement für Künstliche Intelligenz (AI RMF 1.0)), *NIST Special Publication AI 100-1*, 2023.

[51] Dario Maisto, „From Digital Sovereignty Platforms To Sovereign Cloud Platforms: Three Reasons For A Title Change“ (Von digitalen Souveränitätsplattformen zu souveränen Cloud-Plattformen: Drei Gründe für die Umbenennung), *Forrester Blogs*, 11. August 2025, www.forrester.com/blogs/from-digital-sovereignty-platforms-to-sovereign-cloud-platforms-three-reasons-for-a-title-change.

[52] McKinsey & Company, „Future-Proofing the IT Function Amid Global Trends and Disruptions“ (Die IT-Funktion zukunftssicher machen angesichts globaler Trends und Disruptionen), *McKinsey Digital*, 11. Juni 2024, www.mckinsey.com/capabilities/mckinsey-digital/our-insights/tech-forward/future-proofing-the-it-function-amid-global-trends-and-disruptions.

[53] Sébastien Bubeck, Varun Chandrasekaran, Ronen Eldan et al., „Sparks of artificial general intelligence: Early experiments with GPT-4“ (Funken Allgemeiner KI: Frühe Experimente mit GPT-4), Cornell University, 13. April 2023, https://arxiv.org/abs/2303.12712.

[54] Aditya Challapally, Chris Pease, Ramesh Raskar, et al., „The GenAI Divide: State of AI in Business 2025“ (Die GenAI-Kluft: Stand der KI im Unternehmensumfeld 2025), *MIT NANDA*, Juli 2025, https://mlq.ai/media/quarterly_decks/v0.1_State_of_AI_in_Business_2025_Report.pdf.

[55] Mark J. Barrenechea, Tom Jenkins, and David Fraser, The Anticipant Organization (Die antizipierende Organisation), OpenText Corporation, 2022.

[56] Ibid.

[57] Ibomoiye Domor Mienye, Nobert Jere, George Obaido, Oyindamola Omolara Ogunruku, Ebenezer Esenogho and Cameron Modisane, „Large language models: an overview of foundational architectures, recent trends, and a new taxonomy“ (Große Sprachmodelle: Überblick über grundlegende Architekturen, aktuelle Trends und eine neue Taxonomie), *Discover Applied Sciences*, 7, 1027, 2. September 2025, https://link.springer.com/article/10.1007/s42452-025-07668-w.

[58] Ruei-Shan Lu, Ching-Chang Lin, and Hsiu-Yuan Tsao, „Empowering Large Language Models to Leverage Domain-Specific Knowledge in E-Learning“ (Große Sprachmodelle befähigen, domänenspezifisches Wissen im E-Learning zu nutzen), *Applied Sciences*, 14(12), 5264, 18, Juni 2024, https://doi.org/10.3390/app14125264.

[59] Qizheng Zhang, Changran Hu, Shubhangi Upasani et al., „Agentic Context Engineering: Evolving Contexts for Self-Improving Language Models“ (Agentenbasierte Kontextgestaltung: Evolvierende Kontexte für sich selbst verbessernde Sprachmodelle), arXiv preprint arXiv:2510.04618, 2025, https://www.arxiv.org/pdf/2510.04618.

[60] Lingrui Mei, Jiayu Yao, Yuyao Ge et al., „A survey of context engineering for large language models“ (Überblick über Kontext-Engineering für große Sprachmodelle), Cornell University, arXiv:2507.13334, 21. Juli 2025, https://arxiv.org/abs/2507.13334.

[61] A. Feder Cooper, Christopher A. Choquette-Choo, Miranda Bogen et al., „Machine Unlearning Doesn’t Do What You Think: Lessons for Generative AI Policy, Research, and Practice“ (Maschinelles Vergessen wirkt anders als gedacht: Lehren für Politik, Forschung und Praxis im Bereich generativer KI), SSRN, 6. Februar 2025, https://papers.ssrn.com/sol3/papers.cfm?abstract_id=5060253.

[62] K. Boyd, „Microsoft 365 copilot for executives: Sharing Our Customer Zero Deployment and adoption journey at Microsoft“ (Microsoft 365 Copilot für Führungskräfte: Unsere interne Einführung und Nutzung als „Customer Zero“), *Microsoft Inside Track Blog*, 5. Dezember 2024, https://www.microsoft.com/insidetrack/blog/copilot-for-microsoft-365-for-executives-sharing-our-internal-deployment-and-adoption-journey-at-microsoft/.

[63] „Product owner“ (Produktverantwortlicher), Scaled Agile Framework, 25. Februar 2025, https://framework.scaledagile.com/product-owner.

[64] Alexander Sukharevsky, Alexis Krivkovich, Arne Gast et al., „The agentic organization: Contours of the next paradigm for the AI era“ (Die agentenbasierte Organisation: Konturen des nächsten Paradigmas im KI-Zeitalter), *McKinsey & Company*, September 26, 2025, https://www.mckinsey.com/capabilities/people-and-organizational-performance/our-insights/the-agentic-organization-contours-of-the-next-paradigm-for-the-ai-era.

[65] V. L. Sunkara, „KPIs for AI agents and Generative AI: A rigorous framework for evaluation and Accountability“ (KPIs für KI-Agenten und generative KI: Ein strenges Rahmenwerk für Bewertung und Verantwortlichkeit), *International Journal of Scientific Research and Modern Technology*, 22–29, 2024, https://doi.org/10.38124/ijsrmt.v3i4.572.

[66] Jared Kaplan, Sam McCandlish, Tom Henighan et al., „Scaling Laws for Neural Language Models“ (Skalierungsgesetze für neuronale Sprachmodelle), Cornell University, arXiv:2001.08361, 23. Januar 2020, https://arxiv.org/abs/2001.08361.

[67] Gary Marcus, „Deep learning is hitting a wall“ (Deep Learning stößt an Grenzen), *Communications of the ACM*, 65(8), 36–43, 2022, https://nautil.us/deep-learning-is-hitting-a-wall-238440/.

[68] Annette Zimmermann and Danielle Casey, „Emerging Tech Impact Radar: Generative AI“ (Impact Radar für neue Technologien: Generative KI), Gartner, 14. Februar 2025.

Glossar

Agentenbasierte künstliche Intelligenz: Agentenbasierte KI ist ein Paradigma der künstlichen Intelligenz, bei der Systeme als autonome Agenten fungieren. Im Gegensatz zu Modellen, die nur auf Eingaben reagieren, können solche Systeme ihre Umgebung wahrnehmen, mehrstufige Pläne erstellen, eigenständig Entscheidungen treffen und Werkzeuge einsetzen, um aktiv auf ein definiertes Ziel hinzuarbeiten.

Im Unternehmenskontext bilden diese Agenten einen leistungsfähigen Motor für Produktivität, wobei Daten als zentraler Treibstoff dienen. Durch den Zugriff auf private Unternehmensdaten und interne Anwendungen lassen sich komplexe Arbeitsabläufe automatisieren, die zuvor menschliches Urteilsvermögen erforderten. Agentenbasierte KI wird von einem „digitalen Gehirn" gesteuert – einem leistungsfähigen Modell, das jahrzehntelanges menschliches Wissen und Erfahrungen verarbeiten kann.

Air gap: Eine Netzwerksicherheitsmaßnahme, bei der ein oder mehrere Computer physisch von allen ungesicherten Netzwerken getrennt sind. Ziel ist es, sicherzustellen, dass ein sicheres Computernetzwerk physisch von allen ungesicherten Netzwerken in Reichweite getrennt ist. Ungesicherte Netzwerke können das Internet oder andere lokale Netzwerke umfassen. Diese Methode wird häufig in Hochsicherheitsumgebungen wie im militärischen Bereich eingesetzt. Dieser Begriff wird manchmal auch als Luftwand bezeichnet.

Analytik: Der systematische Prozess, Daten zu erfassen, zu verarbeiten und zu interpretieren, um Muster, Trends und Erkenntnisse zu identifizieren. In der Informatik wird Analytik eingesetzt, um die Entscheidungsfindung zu unterstützen, die Leistung zu optimieren und durch Techniken wie statistische Analytik, Datenvisualisierung und maschinelles Lernen zukünftige Ergebnisse vorherzusagen.

Automatisierung: Der Einsatz von Technologie zur Ausführung von Aufgaben mit minimalem oder keinem menschlichen Eingreifen. In der Computertechnik rationalisiert die Automatisierung sich wiederholende oder regelbasierte Prozesse wie z. B. die Softwarebereitstellung, Systemüberwachung oder Datenverarbeitung, um die Effizienz, Konsistenz und Zuverlässigkeit zu verbessern.

Berechtigungen: Verwaltung und Steuerung der Zugriffsrechte in Computer- und Netzwerksystemen. Eine Zugriffskontrollliste (Access Control List, ACL) enthält die Datensätze, die einer Datei, einem Verzeichnis oder einer anderen Ressource zugeordnet sind, und definiert, welche Benutzer, Gruppen, Prozesse oder Geräte darauf zugreifen dürfen und in welchem Umfang.

Big Data: Extrem große und komplexe Datensätze, die zur Verarbeitung und Gewinnung von Erkenntnissen fortschrittliche Tools und Analytik benötigen.

Bots: Softwareprogramme zur Automatisierung von Aufgaben, häufig durch Simulation menschlicher Aktivitäten. In der Computertechnik können Bots eine Vielzahl von Funktionen erfüllen, von hilfreichen Aktivitäten wie Chatbots für den Kundensupport, Suchmaschinenindizierung und Workflow-Automatisierung bis hin zu böswilligen Anwendungen wie Spamming, Credential Stuffing (automatisches Ausprobieren gestohlener Zugangsdaten) oder der Verbreitung von Malware. Bots arbeiten in der Regel mit hoher Geschwindigkeit und Skalierung, was sie sowohl für legitime als auch für schädliche Zwecke zu leistungsstarken Tools macht.

Business Intelligence (BI): Die Technologien, Tools und Verfahren, mit denen Geschäftsdaten erfasst, integriert und analysiert werden, um die Entscheidungsfindung zu unterstützen. BI-Plattformen wandeln Rohdaten in Dashboards, Berichte und Visualisierungen um und helfen Unternehmen dabei, Trends zu erkennen, die Leistung zu messen und fundiertere strategische Entscheidungen zu treffen.

Change Management (Änderungsmanagement): Eine umfassendere Disziplin, die sich auf die menschliche und organisatorische Seite von Veränderungsprozessen konzentriert und sicherstellt, dass neue Systeme, Prozesse oder Technologien erfolgreich eingeführt werden. Während sich das Konfigurationsmanagement mit der technischen Konsistenz befasst, befasst sich das Change Management mit Kommunikation, Schulung und Abstimmung der Interessengruppen, um den Widerstand zu minimieren und den Wert von Änderungsinitiativen zu maximieren.

Cloud/Die Cloud: Ein Computermodell, das über das Internet einen bedarfsgerechten Zugriff auf gemeinsam genutzte Ressourcen wie Server, Speicher, Datenbanken, Netzwerke, Software und Analytik ermöglicht. Die Cloud ermöglicht es Benutzern und Organisationen, Ressourcen schnell zu skalieren, nur für das zu zahlen, was sie nutzen, und auf Dienste zuzugreifen, ohne physische Hardware oder Infrastruktur vor Ort zu unterhalten.

CloudOps: CloudOps steht für Cloud Operations und konzentriert sich auf die Verwaltung und Optimierung von Anwendungen und Infrastrukturen, die in Cloud-Umgebungen ausgeführt werden. Es erweitert die DevOps-Prinzipien auf die Cloud und legt den Schwerpunkt auf Skalierbarkeit, Leistungsüberwachung, Sicherheit und Kosteneffizienz in dynamischen, verteilten Systemen.

Compliance oder Einhaltung gesetzlicher Vorschriften (in der Technik): Die Praxis, sicherzustellen, dass Systeme, Prozesse und Datenmanagement den geltenden Gesetzen, Vorschriften, Standards und internen Richtlinien entsprechen. Im technischen Bereich deckt die Compliance häufig Bereiche wie Datenschutz (z. B. DSGVO, CCPA), Cybersicherheit, Barrierefreiheit und branchenspezifische Regeln ab und hilft Unternehmen dabei, Risiken zu reduzieren, Vertrauen zu wahren und rechtliche oder finanzielle Strafen zu vermeiden. (Siehe auch: Datensicherheit; Cybersicherheit; Datenschutz).

Content-Lifecycle-Management (CLM): Die Kombination aus Dokumentenmanagement, Datensatzverwaltung, Workflow, Archivierung und Bildverarbeitung zu einer integrierten Lösung, die den gesamten Lebenszyklus von Inhalten – von der Erstellung über die Speicherung und Archivierung bis zur endgültigen Löschung – effizient verwaltet.

Content-Management-System (CMS): Eine Softwareplattform, die es Benutzern ermöglicht, digitale Inhalte – normalerweise für Websites – zu erstellen, zu bearbeiten, zu organisieren und zu veröffentlichen, ohne dass umfassende technische Kenntnisse erforderlich sind. CMS-Plattformen wie WordPress, Drupal oder Adobe Experience Manager bieten Vorlagen, Workflows und Integrationen, die die Verwaltung von Online-Inhalten effizienter und kollaborativer machen. Ein Content-Management-System (CMS) verwaltet Website-Inhalte zur Veröffentlichung, während ein Enterprise Content Management (ECM)-System alle Organisationsinformationen über ihren gesamten Lebenszyklus hinweg für Governance, Compliance und Geschäftsprozesse verwaltet.

Customer Relationship Management, CRM: Eine Strategie und eine Reihe von Softwaretools, mit denen Unternehmen die Interaktionen mit aktuellen und potenziellen Kunden verwalten können. CRM-Systeme zentralisieren Kundendaten, verfolgen Verkäufe und Kommunikation und unterstützen Marketing, Service und den Aufbau von Beziehungen, um die Kundenbindung zu verbessern und das Wachstum voranzutreiben.

Cybersicherheit: Die Praxis, Systeme, Netzwerke, Software und Daten vor digitalen Angriffen, unbefugtem Zugriff oder Beschädigung zu schützen. Sie umfasst Technologien, Prozesse und Richtlinien zum Schutz von Vertraulichkeit, Integrität und Verfügbarkeit und hilft Unternehmen, sich gegen Bedrohungen wie Malware, Phishing, Ransomware und Insiderrisiken zu verteidigen. (Siehe auch: Datensicherheit; Compliance; Datenschutz).

Data Lake: Ein zentraler Speicher, der große Mengen an Rohdaten in ihrem ursprünglichen Format enthält, unabhängig davon, ob sie strukturiert, halbstrukturiert oder unstrukturiert sind. Im Gegensatz zu herkömmlichen Datenbanken oder Data Warehouses ermöglichen Data Lakes Unternehmen, Daten zuerst zu speichern und sie später zu organisieren oder zu analysieren. Dies unterstützt Big-Data-Analytik, maschinelles Lernen und Echtzeit-Einblicke.

Data Mapping oder Datenzuordnung: Der Prozess des Abgleichs von Datenfeldern aus einem System, einem Format oder einer Datenbank mit einem anderen, um die Integration, Migration oder Analytik zu ermöglichen. Eine effektive Datenzuordnung sorgt für Konsistenz und Genauigkeit und ermöglicht es, Informationen plattformübergreifend zu konsolidieren, die Einhaltung von Vorschriften zu unterstützen und Daten für Analytik oder maschinelles Lernen vorzubereiten.

Data Warehouse: Ein strukturiertes Speichersystem, das für die Abfrage und Berichterstattung über kuratierte Daten optimiert ist. Im Gegensatz zu Data Lakes, die Rohdaten enthalten, speichern Data Warehouses bereinigte, organisierte und integrierte Daten – in der Regel aus mehreren Quellen – wodurch Unternehmen leichter Analytik durchführen, Dashboards erstellen und Entscheidungsprozesse unterstützen können.

Datengesteuert: Der Einsatz von Datenanalyse und Erkenntnissen als Grundlage für Geschäftsentscheidungen und -strategien.

Datenresidenz (Data Residency): Der physische oder geografische Ort, an dem die Daten eines Unternehmens gespeichert und verarbeitet werden. Die Datenresidenz wird häufig durch rechtliche, behördliche oder geschäftliche Anforderungen bestimmt und stellt sicher, dass Informationen innerhalb bestimmter Rechtsräume verbleiben, was sich auf Compliance, Sicherheit und Leistung auswirken kann.

Datenschutz (Governance): Die Disziplin der Verwaltung und des Schutzes personenbezogener Daten, um sicherzustellen, dass diese auf eine Weise erfasst, gespeichert und verwendet werden, bei der die Rechte und Erwartungen des Einzelnen respektiert wird. Beim Datenschutz stehen Transparenz, Einwilligung und die Einhaltung gesetzlicher Vorschriften im Mittelpunkt, um sicherzustellen, dass Unternehmen verantwortungsvoll mit sensiblen Daten umgehen und gleichzeitig das Vertrauen der Benutzer wahren. (Siehe auch: Datensicherheit; Cybersicherheit; Compliance).

Datenschutz (technisch): Der Schutz und der ordnungsgemäße Umgang mit Benutzerdaten, um sicherzustellen, dass Einzelpersonen die Kontrolle darüber behalten, wie ihre personenbezogenen Daten in Computerumgebungen erfasst, verwendet, weitergegeben und gespeichert werden. Datenschutzmaßnahmen verhindern den unbefugten Zugriff auf oder Missbrauch von Daten und orientieren sich an rechtlichen, ethischen und regulatorischen Standards. (Siehe auch: Persönlich identifizierbare Informationen (PII).

Datenschutz-Folgenabschätzung (Data Protection Impact Assessment, DPIA): Ein Prozess, der gemäß Vorschriften wie der DSGVO erforderlich ist, um Risiken für personenbezogene Daten zu identifizieren und zu minimieren, bevor ein Projekt gestartet wird, bei dem Daten erfasst oder verarbeitet werden. Eine Datenschutz-Folgenabschätzung bewertet, wie Daten verwendet werden, bewertet mögliche Auswirkungen auf den Datenschutz und dokumentiert Schutzmaßnahmen, um die Einhaltung der Vorschriften sicherzustellen und die Rechte von Einzelpersonen zu schützen.

Datenschutz-Grundverordnung (DSGVO): Ein umfassendes Datenschutzgesetz, das 2018 von der Europäischen Union verabschiedet wurde. Die DSGVO regelt, wie Unternehmen personenbezogene Daten erheben, verarbeiten, speichern und weitergeben, wobei Transparenz, Nutzereinwilligungen und individuelle Rechte im Vordergrund stehen und die Nichteinhaltung erhebliche Strafen nach sich zieht.

Datensicherheit: Eine Reihe von Praktiken, Technologien und Richtlinien, die zum Schutz digitaler Informationen vor unbefugtem Zugriff, Beschädigung oder Verlust eingesetzt werden. Es umfasst Maßnahmen wie Verschlüsselung, Zugriffskontrollen, Backups und Überwachung, um die Vertraulichkeit, Integrität und Verfügbarkeit von Daten während ihres gesamten Lebenszyklus zu gewährleisten. (Siehe auch: Cybersicherheit; Compliance; Datenschutz).

Datensilos: Isolierte Datensammlungen, die von einer Abteilung, einem System oder einer Plattform kontrolliert werden und für andere innerhalb einer Organisation nicht ohne weiteres zugänglich sind. Datensilos schränken die Zusammenarbeit ein, verringern die Sichtbarkeit und können zu Ineffizienzen oder Inkonsistenzen führen, was es schwieriger macht, eine einheitliche Sicht auf die Informationen im gesamten Unternehmen zu erhalten.

Datensouveränität: Der Grundsatz, dass digitale Daten den Gesetzen und Verwaltungsstrukturen des Landes oder der Region unterliegen, in dem sie erhoben, gespeichert oder verarbeitet werden. Sie stellt sicher, dass der Umgang mit Daten den lokalen Vorschriften entspricht, wie z. B. der Datenschutz-Grundverordnung (DSGVO) in der EU oder den Gesetzen zur Datenlokalisierung in anderen Ländern, was sich darauf auswirkt, wie Unternehmen die Speicherung, Sicherheit und grenzüberschreitende Übertragungen verwalten.

Deep Learning: Ein spezialisierter Zweig des maschinellen Lernens, der auf tiefen neuronalen Netzwerken basiert – mehrschichtigen Strukturen miteinander verbundener „Neuronen", deren Gewichte und Parameter durch Training angepasst werden können. Dieser Ansatz zeichnet sich durch das Extrahieren von Mustern und Erkenntnissen aus unstrukturierten Daten wie Bildern, Text, Audio und Video aus und ist damit das Rückgrat vieler moderner KI-Anwendungen wie Bilderkennung, Sprachübersetzung und Sprachverarbeitung.

Deklarative oder Zero-Party-Daten: Von Benutzern freiwillig weitergegebene Informationen, wie etwa Präferenzen oder Absichten, die im Rahmen von Umfragen oder Fragebögen erhoben wurden.

DevOps: Eine Reihe von Verfahren, die Softwareentwicklung (Dev) und IT-Betrieb (Ops) kombinieren, um Entwicklungszyklen zu verkürzen, die Zusammenarbeit zu verbessern und Updates zuverlässiger bereitzustellen. Durch die Automatisierung von Tests, Integration und Bereitstellung hilft DevOps den Teams, Software schneller zu veröffentlichen und gleichzeitig die Qualität zu erhalten.

Digitale Arbeitskräfte: Eine Sammlung automatisierter Softwaresysteme, die als „digitale Arbeitskräfte" bezeichnet werden (z. B. KI-Agenten, Bots und virtuelle Assistenten), die Aufgaben ausführen, die traditionell von Menschen erledigt werden.

Digitale Governance: Der Rahmen aus Richtlinien, Rollen, Prozessen und Standards, der die Art und Weise bestimmt, wie ein Unternehmen seine digitalen Ressourcen, Technologien und Daten verwaltet. Sie sorgt für Rechenschaftspflicht, Compliance, Sicherheit und Ausrichtung auf die Unternehmensziele und schafft ein Gleichgewicht zwischen Innovation und Risikomanagement im digitalen Betrieb.

Domänen spezifische Intelligenz (ANI): Sie bezeichnet KI-Systeme, die dafür entwickelt und trainiert werden, spezifische Aufgaben auszuführen oder klar definierte Probleme zu lösen – oft auf oder über dem Leistungsniveau des Menschen –, jedoch ohne allgemeines Denkvermögen oder Verständnis, das über ihren programmierten Aufgabenbereich hinausgeht.

Echtzeit-Analytik: Der Prozess des Erfassens, Verarbeitens und Analysierens von Daten unmittelbar nach deren Entstehung, der es Unternehmen ermöglicht, Erkenntnisse zu gewinnen und ohne Verzögerung Entscheidungen zu treffen. Im Computerbereich unterstützt die Echtzeitanalyse Anwendungsfälle wie Betrugserkennung, personalisierte Empfehlungen, Systemüberwachung und Live-Performance-Tracking. (Siehe auch: Analytik).

E-Commerce: Der Kauf und Verkauf von Waren oder Dienstleistungen über das Internet, einschließlich Aktivitäten wie Online-Shopping, elektronische Zahlungen, digitale Marktplätze und mobiler Handel. E-Commerce ermöglicht es Unternehmen, Kunden direkt über Websites, Apps und Plattformen zu erreichen, was die Art und Weise, wie Produkte vermarktet, gekauft und geliefert werden, verändert.

Edge Computing: Eine verteilte Computerarchitektur, bei der sich Verarbeitungsleistung, Speicher und Datenanalyse näher am Ort der Datengenerierung befinden – auf Geräten, Gateways oder lokalen „Edge"-Servern – und nicht zentral in entfernten Cloud-Rechenzentren. Dieser Ansatz verbessert die Reaktionszeit, reduziert die Latenzzeit und die Bandbreitennutzung und ermöglicht Echtzeit- oder echtzeitnahe Anwendungen, insbesondere im Internet der Dinge, in autonomen Systemen und in Umgebungen, in denen Geschwindigkeit oder lokale Entscheidungsfindung wichtig sind.

Elektronischer Datenaustausch (EDI): Eine standardisierte Methode für den elektronischen Austausch von Geschäftsdokumenten – wie Bestellungen, Rechnungen und Versandbenachrichtigungen – zwischen Organisationen, wodurch die manuelle Dateneingabe entfällt und Geschwindigkeit, Genauigkeit und Konsistenz der Transaktionen verbessert werden.

Enterprise Content Management (ECM): Eine Plattform zur Speicherung, Verwaltung und Bereitstellung von Inhalten auf Unternehmensebene. Dazu gehören Dokumente, Bilder, Videos und andere geschäftsrelevante Inhaltsformen. Eine ECM-Plattform sollte sich nahtlos in wichtige Unternehmensanwendungen und -systeme integrieren lassen (z. B. Lösungen für Enterprise Resource Planning, Customer Relationship Management, Human Capital Management und Supply Chain Management), um Geschäftsprozesse zu beschleunigen und die von ihnen generierten Daten zu nutzen. ECM umfasst auch Cloud Content Management, das eine schnelle Bereitstellung und Zusammenarbeit in Cloud-Umgebungen ermöglicht.

Enterprise Information Management (EIM): Lösungen für die Verwaltung der Erstellung, Erfassung, Nutzung und des gesamten Lebenszyklus von strukturierten und unstrukturierten Informationen. Sie unterstützen Organisationen dabei, den Wert ihrer Informationen zu erschließen, diese zu schützen und gesetzliche sowie regulatorische Anforderungen zu erfüllen.

Enterprise Resource Planning (ERP): Ein integriertes Softwaresystem, das zentrale Geschäftsprozesse – wie Finanzen, Lieferkette, Fertigung, Personalwesen und Kundenbeziehungen – auf einer einheitlichen Plattform verwaltet. ERP zentralisiert Daten und Arbeitsabläufe abteilungsübergreifend, verbessert die Effizienz, Zusammenarbeit und Entscheidungsfindung und bietet gleichzeitig eine einzige zuverlässige Informationsquelle für das Unternehmen.

Extraterritorialer Datenzugriff: Die Fähigkeit von Regierungen, Organisationen oder Einrichtungen, auf Daten zuzugreifen, die außerhalb ihrer eigenen nationalen Gerichtsbarkeit gespeichert sind, und diesen Zugriff zu verlangen. Im Bereich der Technologie- und Datenverwaltung wirft der extraterritoriale Zugang Bedenken hinsichtlich der Souveränität, des Datenschutzes und der Einhaltung von Vorschriften auf, da Gesetze wie der U.S. CLOUD Act oder ähnliche Rahmenwerke die Offenlegung von im Ausland gespeicherten Daten erzwingen können.

FinOps: Kurzform für Cloud Financial Operations – ein Rahmenwerk zur Verwaltung und Optimierung von Cloud-Kosten durch die Zusammenarbeit von Entwicklungs-, Finanz- und Geschäftsteams. FinOps schafft finanzielle Transparenz, ermöglicht die Optimierung von Ausgaben und unterstützt fundierte Entscheidungen im Hinblick auf Kosten und Leistung in Cloud-Umgebungen.

Generative KI (GenAI): Bezeichnet KI-Systeme, die mithilfe von Modellen des maschinellen Lernens neue und originelle Inhalte erzeugen. Modelle wie ChatGPT, Claude, Gemini oder DeepSeek werden mit öffentlich zugänglichen Datenquellen wie Websites, Nachrichtenartikeln, Reddit oder Wikipedia trainiert. Generative KI eignet sich zur Gewinnung allgemeiner Erkenntnisse, ist jedoch auf allgemeine Aufgaben beschränkt. Dies liegt daran, dass ihnen der Zugang zu den privaten, in Echtzeit verfügbaren und unternehmenszentrierten Daten fehlt, die für bestimmte Geschäftsanwendungsfälle erforderlich sind.

Geopolitik: Die Untersuchung, wie geografische Faktoren – wie Standort, Ressourcen, physisches Terrain, Bevölkerung und wirtschaftliche oder demografische Trends – die politische Macht, die Außenpolitik und die Entscheidungsfindung zwischen Staaten oder anderen politischen Akteuren beeinflussen. Es wird untersucht, wie die Kontrolle über Territorien, strategische Regionen und geografische Merkmale Beziehungen, Konflikte und Zusammenarbeit auf globaler Ebene prägt.

Geschäftsprozessmanagement (BPM): Bezeichnet die Ausrichtung von Prozessen an den strategischen Zielen einer Organisation, die Entwicklung und Implementierung prozessorientierter Werkzeuge und Architekturen sowie die Definition von Messsystemen zur Unterstützung eines effektiven Prozessmanagements.

Gesetz zum Schutz personenbezogener Daten und elektronischer Dokumente (PIPEDA): Das kanadische Bundesdatenschutzgesetz regelt, wie Organisationen des privaten Sektors personenbezogene Daten im Rahmen kommerzieller Aktivitäten erheben, verwenden und weitergeben dürfen. PIPEDA gewährt Einzelpersonen das Recht, auf ihre Daten zuzugreifen und sie zu korrigieren, verlangt von Organisationen, dass sie eine aussagekräftige Einwilligung einholen, und schreibt Schutzmaßnahmen zum Schutz personenbezogener Daten vor.

Grafische Benutzeroberfläche (Graphical User Interface, GUI): Eine visuelle Oberfläche, die es Benutzern ermöglicht, anstatt über textbasierte Befehle über grafische Elemente wie Fenster, Symbole, Schaltflächen und Menüs mit Software oder Geräten zu interagieren. GUIs machen Technologie intuitiver und zugänglicher, indem sie die Navigation per Mausklick, Drag-and-Drop-Aktionen und visuelles Feedback ermöglichen.

Große Sprachmodelle (Large Language Models, LLMs): Eine Art Basismodell, das mit Deep Learning erstellt und mithilfe selbstüberwachter Methoden anhand riesiger Textkorpora vortrainiert wurde. Diese Modelle verarbeiten Eingaben in Form von „Token“ (Wort- oder Zeichenteile), lernen die statistischen Beziehungen zwischen ihnen und verwenden diese Beziehungen, um Texte in menschlicher Sprache zu verstehen, zu generieren, zusammenzufassen oder zu transformieren. Aufgrund ihrer Größe und Architektur (oft transformerbasiert) können LLMs viele Sprachaufgaben sofort ausführen, können aber auch durch Prompt-Engineering für bestimmte Anwendungen feinabgestimmt oder gesteuert werden. Beispiele für LLMs sind GPT-4 und die Llama-Modelle von Meta.

Grundlagenmodelle (Foundational Models): Große Deep-Learning-Modelle, die anhand riesiger Mengen unstrukturierter, nicht gekennzeichneter Daten trainiert wurden und entweder direkt oder durch Feinabstimmung für bestimmte Anwendungen für die Ausführung einer großen Bandbreite von Aufgaben konzipiert sind. Grundlagenmodelle (Foundational Models) können auch für generative oder nicht-generative Zwecke verwendet werden (zum Beispiel zur Klassifizierung der Nutzerstimmung als negativ oder positiv auf der Grundlage von Anrufprotokollen).

Human-in-the-Loop (HITL): Ein Ansatz der künstlichen Intelligenz und Automatisierung, bei dem Menschen aktiv in den Entscheidungs- oder Trainingsprozess eines Systems eingebunden bleiben. HITL wird eingesetzt, um den Überblick zu behalten, die Genauigkeit zu verbessern, Fehler zu korrigieren und Grenzfälle zu behandeln, um sicherzustellen, dass automatisierte Systeme mit menschlichem Urteilsvermögen, ethischen Grundsätzen und dem realen Kontext übereinstimmen.

Hyperscaler: Ein großer Cloud-Service-Anbieter – wie Amazon Web Services (AWS), Microsoft Azure oder Google Cloud – der in globalen Rechenzentren Rechen-, Speicher- und Netzwerkdienste in großem Maßstab bereitstellt. Hyperscaler sind bekannt für ihre Fähigkeit, Ressourcen sofort nach oben oder unten zu skalieren, Multi-Tenant-Workloads zu unterstützen und das Rückgrat für Cloud native, KI- und datenintensive Anwendungen zu bilden.

Identitäts- und Zugriffsmanagement (IAM): Ein Sicherheitsrahmen, der sicherstellt, dass die richtigen Personen zum richtigen Zeitpunkt den passenden Zugriff auf die entsprechenden Ressourcen erhalten. IAM verwaltet digitale Identitäten, Authentifizierung und Berechtigungen system- und anwendungsübergreifend, um sensible Daten zu schützen und die Einhaltung regulatorischer Vorgaben zu gewährleisten.

Identitätsmanagement (IdM): Ein zentraler Bestandteil des Identitäts- und Zugriffsmanagements (IAM), der sich auf die Erstellung, Pflege und Löschung von Benutzeridentitäten und deren Attributen konzentriert. Während IAM sowohl Identitäts- als auch Zugriffskontrolle umfasst, befasst sich IdM speziell mit dem Lebenszyklus der Benutzeridentität – von der Kontoerstellung über Rollenaktualisierungen bis zur Sicherstellung korrekter und geschützter Identitätsdaten.

Infrastructure-as-a-Service (IaaS): Ein Cloud-Computing-Modell, das virtualisierte Computerressourcen – wie Server, Speicher und Netzwerke – über das Internet auf Pay-as-you-go-Basis bereitstellt. IaaS ermöglicht es Unternehmen, ihre Infrastruktur schnell zu skalieren, ohne in physische Hardware investieren zu müssen, was Flexibilität, Kosteneffizienz und schnelle Bereitstellung unterstützt.

Intelligenz-Ebene: Die Analytik- und Entscheidungsebene innerhalb eines Technologie-Stacks, die Rohdaten in umsetzbare Erkenntnisse umwandelt. Die Intelligenzebene, die häufig auf KI, maschinellem Lernen oder fortschrittlicher Analytik basiert, liegt über den Datenspeicher- und -verarbeitungssystemen und ermöglicht Personalisierung, Vorhersagen und Automatisierung, die zu intelligenteren Geschäftsergebnissen führen.

Internet der Dinge (IoT): Bezeichnet ein Netzwerk physischer Geräte – etwa Sensoren, Haushaltsgeräte, Fahrzeuge oder Maschinen –, die mit dem Internet verbunden sind und Daten erfassen, austauschen und darauf reagieren können. Dadurch werden Automatisierung, Überwachung und datenbasierte Entscheidungsfindung in Echtzeit ermöglicht.

Interoperabilität: Die Fähigkeit verschiedener Systeme, Anwendungen oder Plattformen, Daten nahtlos auszutauschen und zu verwenden. Da die Interoperabilität auf gemeinsamen Normen und Protokollen beruht, werden technische Hindernisse abgebaut und eine effiziente Zusammenarbeit verschiedener Systeme ermöglicht.

Kalifornisches Verbraucherdatenschutzgesetz (California Consumer Privacy Act, CCPA): Ein Datenschutzgesetz, das 2020 in Kalifornien in Kraft getreten ist und den Bürgern mehr Kontrolle über ihre persönlichen Daten gibt. Der CCPA verpflichtet Unternehmen, offenzulegen, welche Daten sie erfassen, wie diese verwendet und mit wem sie geteilt werden, und gleichzeitig den Verbrauchern das Recht einzuräumen, auf ihre Daten zuzugreifen, sie zu löschen und dem Verkauf ihrer Daten zu widersprechen.

Kapitalrendite (ROI): Eine Leistungskennzahl, die die Rentabilität oder Effizienz einer Investition misst und durch Vergleich des Nettogewinns oder -nutzens mit den Kosten berechnet wird. Im Computer- und Geschäftskontext wird der ROI verwendet, um den Wert von Technologieinitiativen, -projekten oder -käufen zu bewerten, indem die finanziellen Erträge im Verhältnis zu den investierten Ressourcen quantifiziert werden.

KI-Agent: Ein spezifisches Softwaresystem, das für eine definierte Aufgabe entwickelt wurde und seine Umgebung autonom wahrnehmen, Aufgaben planen und analysieren sowie Maßnahmen ergreifen kann, um festgelegte Ziele zu erreichen. Es arbeitet mit einem gewissen Maß an Unabhängigkeit (aber in der Regel innerhalb der vom Menschen gesetzten Einschränkungen), lernt oder passt sich im Laufe der Zeit an, verwendet bei Bedarf Tools oder externe Datenquellen und unterstützt die Entscheidungsfindung mit minimaler direkter Aufsicht. KI-Agenten sind die Bausteine des Frameworks der agentenbasierten KI.

KI-fähige Plattform / KI-fähiger Stack: Eine Technologieumgebung, die entwickelt wurde, um die Entwicklung, Bereitstellung und Skalierung von Anwendungen für künstliche Intelligenz zu unterstützen. Sie vereint skalierbare Infrastruktur, starke Datenverwaltung und modulare Tools – wie APIs, Modellmanagement und Integrationspipelines –, um KI-Projekte schneller, zuverlässiger und einfacher von der Experimentierphase in die Produktion zu überführen.

KI-gestützte Analytik: Die Anwendung von künstlicher Intelligenz und maschinellem Lernen zur Automatisierung der Datenerfassung, -aufbereitung und -analyse. Durch das Erkennen von Mustern und das Erfassen von Erkenntnissen in Echtzeit oder nahezu in Echtzeit unterstützt es Unternehmen dabei, Ergebnisse vorherzusagen, Trends zu erkennen und schnellere, fundiertere Entscheidungen zu treffen.

Konfigurationsmanagement: Der Prozess der systematischen Steuerung und Nachverfolgung von Änderungen an einem System, um dessen Integrität, Konsistenz und Rückverfolgbarkeit über den gesamten Lebenszyklus sicherzustellen.

Kontextuelle Intelligenz: Die Fähigkeit von Systemen, Organisationen oder Einzelpersonen, Daten, Ereignisse oder Verhaltensweisen im jeweiligen Kontext zu interpretieren und entsprechend zu handeln. In der Technologie bezieht es sich auf die Verwendung von Echtzeitsignalen – wie Standort, Verhalten, Vorlieben oder Zeitpunkt – durch KI, um relevantere Erkenntnisse, Empfehlungen und Maßnahmen zu liefern. In den Bereichen Medien und Marketing unterstützt kontextbezogene Intelligenz die Personalisierung und die Orchestrierung der Customer Journey, indem sie sicherstellt, dass jede Interaktion zeitnah, sinnvoll und auf die Kundenbedürfnisse abgestimmt ist.

Künstliche allgemeine Intelligenz (AGI): AGI bezeichnet eine hochentwickelte Form künstlicher Intelligenz, die in der Lage ist, nicht nur einzelne Bereiche, sondern ganze Gesellschaften grundlegend zu verändern. Sie verfügt – ähnlich wie der Mensch – über die Fähigkeit, Wissen zu verstehen, zu erlernen und auf ein breites Spektrum von Aufgaben anzuwenden. Trotz ihres weitreichenden Potenzials wird AGI weiterhin von der Qualität und dem Umfang der zu ihrer Entwicklung verwendeten Trainingsdaten bestimmt.

Künstliche Intelligenz (KI): Technologie, die Maschinen oder Software in die Lage versetzt, Aufgaben auszuführen, die normalerweise menschliche Intelligenz erfordern – Dinge wie Lernen, logisches Denken, Wahrnehmen, Problemlösen, Entscheidungsfindung, Verstehen natürlicher Sprache und Erkennen von Mustern. KI-Systeme imitieren oder simulieren die kognitiven Funktionen des Menschen. Dabei nutzen sie Algorithmen, große Datensätze und Rechenleistung. Da KI eine Vielzahl von Methoden umfasst (von maschinellem Lernen über neuronale Netze bis hin zu Deep Learning), können ihre Anwendungen von der Analytik von Sprache und Text über Vorhersagen und Übersetzungen bis hin zur Generierung kreativer Inhalte oder der Optimierung komplexer Prozesse reichen.

Künstliche Superintelligenz (ASI): Eine hypothetische Form künstlicher Intelligenz, die die menschliche Intelligenz in allen Bereichen übertrifft – einschließlich logischem Denken, Kreativität, sozialem Verständnis und strategischer Entscheidungsfindung – und in der Lage wäre, die besten menschlichen Köpfe in jedem Fachgebiet zu übertreffen.

Kubernetes: Eine Open-Source-Plattform zur Automatisierung der Bereitstellung, Skalierung und Verwaltung von containerisierten Anwendungen. Kubernetes wurde ursprünglich von Google entwickelt und orchestriert Cluster von Containern und übernimmt die Planung, den Lastausgleich und die Ausfallsicherung, sodass Anwendungen in der Cloud, vor Ort oder in hybriden Umgebungen zuverlässig und effizient ausgeführt werden.

Legacy-Plattformen: Veraltete oder ältere Technologiesysteme, Software oder Infrastruktur, die weiterhin verwendet werden, obwohl sie durch neuere Alternativen ersetzt wurden. Ältere Plattformen unterstützen zwar immer noch wichtige Geschäftsabläufe, stellen jedoch häufig Herausforderungen wie eingeschränkte Kompatibilität, höhere Wartungskosten, Sicherheitslücken und Schwierigkeiten bei der Integration in moderne Lösungen dar.

Lizenz- und Nutzungsrechte: Bezeichnen die Bedingungen, unter denen ein bestimmter Vermögenswert (Asset) verwendet werden darf. Dazu gehören Angaben zum rechtlichen Eigentümer, zu den zulässigen Verwendungsmedien (z. B. Web, Print, TV) sowie zu den finanziellen Verpflichtungen, die durch die Nutzung des Assets entstehen.

Managed Services oder Verwaltete Dienste: Ausgelagerte IT-Operationen und Verantwortlichkeiten, die von einem Drittanbieter bereitgestellt werden. Im Computerbereich umfassen Managed Services in der Regel proaktive Überwachung, Wartung, Sicherheit, Updates und Support für Infrastruktur, Anwendungen oder Cloud-Umgebungen. Dieser Ansatz ermöglicht es Unternehmen, die Belastung interner Teams zu verringern, die Zuverlässigkeit des Systems zu gewährleisten und auf spezialisiertes Fachwissen zuzugreifen, während sie sich auf ihre Kerntätigkeiten konzentrieren.

Maschinelles Lernen (ML): Ein Bereich der künstlichen Intelligenz (KI), der sich auf die Entwicklung von Algorithmen und Modellen konzentriert, die es Systemen ermöglichen, aus Daten zu lernen und ihre Leistung im Laufe der Zeit zu verbessern, ohne explizit programmiert zu werden. ML wird verwendet, um Muster zu erkennen, Vorhersagen zu treffen und die Entscheidungsfindung in Anwendungen wie Empfehlungsmaschinen, Betrugserkennung, Bilderkennung und natürlicher Sprachverarbeitung zu unterstützen.

Metadaten: Daten, die Informationen über andere Daten enthalten und in ein Dokument oder System eingebettet sind. Sie ermöglichen Funktionen der Datensatzverwaltung wie Klassifizierung, Suche, Versions- und Änderungsverfolgung (z. B. Erstellungs-, Änderungs- oder Abrufdatum), Benutzeridentifizierung (Autor und Bearbeiter) sowie zahlreiche weitere Angaben zu den Eigenschaften und dem Kontext eines Dokuments oder Datensatzes.

Middleware-Schicht: Software, die als Brücke zwischen Betriebssystemen, Datenbanken und Anwendungen fungiert und es ihnen ermöglicht, effizient zu kommunizieren und Daten auszutauschen. Die Middleware-Schicht bietet allgemeine Dienste wie Messaging, Authentifizierung, API-Management und Transaktionsverarbeitung und vereinfacht so die Integration und Interoperabilität zwischen komplexen Systemen.

Mikrodienste: Ein architektonischer Ansatz für die Softwareentwicklung, bei dem Anwendungen als Sammlung kleiner, unabhängiger Dienste strukturiert sind, die über APIs kommunizieren. Jeder Microservice konzentriert sich auf eine bestimmte Geschäftsfunktion, kann unabhängig entwickelt und bereitgestellt werden und ist separat skalierbar. Dieses Design verbessert die Flexibilität, Belastbarkeit und Wartbarkeit im Vergleich zu monolithischen Architekturen.

Modernisierung (in der Software): Der umfassendere Prozess der Aktualisierung älterer Systeme, Anwendungen oder Infrastrukturen, um die Vorteile moderner Technologien, Architekturen und Praktiken zu nutzen. Während sich die Migration häufig auf die Verlagerung vorhandener Systeme konzentriert, kann die Modernisierung Refactoring, Replatforming oder Neugestaltung umfassen, um die Skalierbarkeit, Agilität und den langfristigen Geschäftswert zu verbessern.

Monolithische Architektur: Ein traditionelles Softwaredesign, bei dem alle Komponenten einer Anwendung – wie Benutzeroberfläche, Geschäftslogik und Datenverwaltung – eng integriert und als eine Einheit bereitgestellt werden. Obwohl der Aufbau monolithischer Systeme zunächst einfacher ist, kann es im Vergleich zu zusammensetzbaren Alternativen schwieriger sein, sie zu skalieren, zu aktualisieren oder anzupassen.

Multi-Cloud: Eine Cloud-Strategie, bei der ein Unternehmen Dienste von zwei oder mehr Cloud-Anbietern gleichzeitig nutzt. Dieser Ansatz trägt dazu bei, die Abhängigkeit von Anbietern zu verringern, die Widerstandsfähigkeit zu erhöhen, die Leistung zu optimieren und regulatorische oder geografische Anforderungen zu erfüllen, indem Workloads auf mehrere Plattformen wie AWS, Microsoft Azure und Google Cloud verteilt werden.

Multi-Region-Modell: Eine Cloud-Computing-Architektur, in der Anwendungen und Daten in mehreren geografischen Regionen eines Cloud-Anbieters bereitgestellt werden. Dieses Modell verbessert die Verfügbarkeit, Leistung und Notfallwiederherstellung, indem es die Arbeitslasten näher an den Endbenutzern verteilt und Redundanz gewährleistet, falls es in einer Region zu Ausfällen oder Latenzproblemen kommt.

Öffentliche Cloud oder Public Cloud: Ein Cloud-Computing-Modell, bei dem Infrastruktur, Ressourcen und Dienste einem Drittanbieter gehören und von diesem betrieben werden und über das Internet bereitgestellt werden. Öffentliche Cloud-Umgebungen werden von mehreren Organisationen (Mandanten) gemeinsam genutzt, wobei Daten und Workloads logisch voneinander getrennt bleiben. Sie bieten Skalierbarkeit, Flexibilität und Kosteneffizienz. Zu den gängigen Beispielen gehören Amazon Web Services (AWS), Microsoft Azure und Google Cloud. (Siehe auch: Private Cloud).

Operationen mit großen Sprachmodellen (LLMOps): Eine spezialisierte Erweiterung von MLOps, die auf die besonderen Anforderungen im Lebenszyklus großer Sprachmodelle ausgerichtet ist. LLMOps steuert die Auswahl, Feinabstimmung, Bereitstellung, Überwachung und Steuerung von Basismodellen, einschließlich Schnellmanagement, Sicherheitskontrollen, Zugriffssteuerung, Vermeidung von Fehlalarmen und Kosteneffizienz. Es gewährleistet, dass LLMs sicher mit Unternehmensdaten arbeiten, Richtlinien und regulatorische Anforderungen erfüllen und eine zuverlässige Leistung erbringen. Im Unternehmenskontext ist LLMOps entscheidend, um generative KI verantwortungsvoll zu skalieren und gleichzeitig Datenintegrität und Vertrauen zu gewährleisten.

Operatives maschinelles Lernen (MLOps): Ein strukturierter Rahmen für das Management des gesamten Lebenszyklus von Machine-Learning-Modellen in der Produktion. MLOps überträgt die Prinzipien von DevOps auf KI und stellt sicher, dass Modelle zuverlässig, skalierbar und reproduzierbar entwickelt, bereitgestellt, überwacht und gesteuert werden. Es umfasst Datenpipelines, Modellversionierung, Leistungsüberwachung, Erkennung von Verzerrungen und Abweichungen, Sicherheitskontrollen sowie Prüfbarkeit. Ziel von MLOps ist es, den KI-Einsatz zu operationalisieren – um konsistente Ergebnisse zu erzielen, Risiken zu verringern und kontinuierliche Verbesserungen in Unternehmenssystemen zu ermöglichen.

Optische Zeichenerkennung (OCR): Technologie, die gedruckten oder handschriftlichen Text in gescannten Dokumenten oder Bildern in maschinenlesbaren digitalen Text umwandelt und so die Suche, Bearbeitung und automatisierte Verarbeitung ermöglicht.

Orchestrator: Ein System, Werkzeug oder eine Instanz, die mehrere Komponenten oder Prozesse koordiniert und steuert, damit diese gemeinsam komplexe Aufgaben ausführen können. (Siehe auch: Query Router im Kontext von Sovereign Data und KI.)

Orchestrierungsebene: Eine Verwaltungsebene in der Datenverarbeitung, die die Koordination, Planung und Ausführung komplexer Aufgaben über mehrere Systeme, Anwendungen oder Dienste hinweg automatisiert. Die Orchestrierungsebene stellt sicher, dass die Komponenten nahtlos zusammenarbeiten, indem sie Arbeitsabläufe, Ressourcenzuweisung, Skalierung und Abhängigkeiten handhabt. Sie wird häufig in Cloud-Umgebungen, containerisierten Anwendungen und Microservices verwendet, um den Betrieb zu rationalisieren und manuelle Eingriffe zu reduzieren.

Persönlich identifizierbare Informationen (PII): Alle Daten, die entweder allein oder in Kombination mit anderen Informationen zur Identifizierung einer Person verwendet werden können. Beispiele hierfür sind Namen, Adressen, Telefonnummern, E-Mail-Adressen, Sozialversicherungs- oder Reisepassnummern sowie Finanz- oder Krankenakten. Im Bereich Computer- und Datenschutz ist der Schutz personenbezogener Daten von entscheidender Bedeutung, um Vorschriften einzuhalten und Einzelpersonen vor Identitätsdiebstahl oder -missbrauch zu schützen.

Plattform: Eine Computerumgebung, die die zugrunde liegende Infrastruktur, Software und Tools bereitstellt, die für die Ausführung von Anwendungen oder Diensten erforderlich sind. Zu den Plattformen können Betriebssysteme, Cloud-Umgebungen oder Anwendungs-Frameworks gehören, die Entwicklung, Bereitstellung und Integration unterstützen. Sie dienen als Grundlage, auf der Benutzer, Entwickler oder Organisationen digitale Lösungen entwickeln und verwalten.

Platform-as-a-Service (PaaS): Ein Cloud-Computing-Modell, das Entwicklern zum Erstellen, Testen und Bereitstellen von Anwendungen eine einsatzbereite Plattform – einschließlich Infrastruktur, Betriebssystemen und Entwicklungstools – bietet. PaaS abstrahiert die Hardware- und Systemverwaltung, sodass sich Teams auf die Programmierung und Innovation konzentrieren können und gleichzeitig Skalierbarkeit, Sicherheit und Integration mit anderen Cloud-Diensten gewährleistet werden.

Plattform-Ökosystem: Ein komplexes System, in dem mehrere miteinander verbundene Geschäftsmodelle, Inhaltsformate und Einnahmequellen nebeneinander existieren.

Privacy-by-Design oder Datenschutz durch Design: Ein Rahmen, der Datenschutz- und Datenschutzprinzipien direkt in die Gestaltung und den Betrieb von Technologien, Prozessen und Systemen einbettet. Der Schwerpunkt liegt auf proaktiven Maßnahmen – wie der Minimierung der Datennutzung, dem standardmäßigen Schutz und der Gewährleistung von Transparenz –, sodass Datenschutz nicht nachträglich berücksichtigt wird, sondern ein zentrales Designprinzip darstellt.

Private Cloud: Eine Cloud-Computing-Umgebung, die einer einzelnen Organisation vorbehalten ist und exklusiven Zugriff auf Infrastruktur, Ressourcen und Dienste bietet. Private Clouds können vor Ort oder von einem Drittanbieter gehostet werden und sind so konzipiert, dass sie im Vergleich zu öffentlichen Cloud-Umgebungen mehr Kontrolle, Sicherheit und Anpassung bieten. (Siehe auch: Öffentliche Cloud).

Programmierschnittstelle (API): Eine Reihe von Softwareregeln und Protokollen, die es zwei Anwendungen ermöglichen, miteinander zu kommunizieren. APIs bieten KI-Systemen eine strukturierte Möglichkeit, sich programmgesteuert mit externen Modellen, Datensätzen oder anderen Softwarekomponenten zu verbinden.

Prompt Engineering: Die Praxis des Erstellens, Verfeinerns und Optimierens von Eingabeaufforderungen, um ein generatives KI-Modell dazu zu bringen, genaue, relevante und nützliche Ergebnisse zu liefern. Effektives Prompt Engineering verbessert die Effizienz, Konsistenz und Zuverlässigkeit von KI-Systemen und ermöglicht es ihnen, Aufgaben wie Zusammenfassung, Übersetzung, Codierung oder kreative Generierung mit höherer Qualität auszuführen. Da KI-Modelle immer leistungsfähiger werden, ist Prompt Engineering zu einer Schlüsseldisziplin geworden, um die Ergebnisse an den Absichten der Benutzer und den Geschäftszielen auszurichten.

Prozessmanagement: Die Automatisierung von Geschäftsprozessen mithilfe eines regelbasierten Expertensystems, das die jeweils geeigneten Werkzeuge aufruft und dem Benutzer die erforderlichen Informationen, Checklisten, Beispiele und Statusberichte bereitstellt.

Query-Router: Eine intelligente Funktion, die mithilfe einer Regel-Engine steuert, wie souveräne und nicht-souveräne Daten auf Grundlage der Anfrage eines agentenbasierten KI-Workflows weitergeleitet werden.

Reasoning-KI (schlussfolgernde KI): KI-Systeme, die logisches Denken, schrittweise Planung, Problemlösung und Entscheidungsfindung mithilfe strukturierter oder unstrukturierter Daten durchführen und dabei über die Mustererkennung hinausgehen, um Schlussfolgerungen zu ziehen und komplexe Probleme zu lösen.

Recommendation-Engines (Empfehlungssysteme): Sie analysieren Daten und Nutzerverhalten, um relevante Produkte, Dienstleistungen oder Inhalte vorzuschlagen. In der Informatik nutzen Empfehlungssysteme Techniken wie maschinelles Lernen, kollaboratives Filtern oder inhaltsbasiertes Filtern, um die Nutzererfahrungen zu personalisieren, wie sie im E-Commerce, auf Streaming-Plattformen und im digitalen Publishing üblich sind.

Revision oder Audit: Eine systematische Überprüfung und Bewertung von Systemen, Prozessen oder Daten in der Informatik, um Genauigkeit, Sicherheit, Compliance und ordnungsgemäßen Betrieb sicherzustellen. Prüfungen oder Audits können Protokolle, Zugriffskontrollen, Konfigurationen und Richtlinien überprüfen, um Unregelmäßigkeiten aufzudecken, die Einhaltung von Standards zu bestätigen und verbesserungswürdige Bereiche zu identifizieren.

Revisionsfähiger Zugriff: Die Fähigkeit eines Systems, zu protokollieren und zu überprüfen, wer wann und wie auf Ressourcen zugegriffen hat. Es sorgt für Rechenschaftspflicht und Compliance, indem es sicherstellt, dass Zugriffsaktivitäten überprüft und bei Bedarf nachverfolgt werden können.

Revisionspfade: Chronologische Aufzeichnungen, die Systemaktivitäten verfolgen, einschließlich Benutzeraktionen, Datenänderungen und Zugriffsereignisse. Sie sorgen für Transparenz, unterstützen die Einhaltung gesetzlicher Vorschriften und helfen bei Sicherheitsuntersuchungen, indem sie zeigen, wer innerhalb eines Systems was, wann und wie getan hat.

Repository: Ein Speicherort, der häufig mit Tools wie GitHub oder GitLab verwaltet wird und an dem eine Codebasis und ihr Änderungsverlauf aufbewahrt werden. Während sich die Codebasis auf den eigentlichen Code selbst bezieht, verfolgt das Repository auch Revisionen, Zweige und Beiträge, sodass Teams den Code effektiv verwalten und daran zusammenarbeiten können.

Robotergestützte Prozessautomatisierung (RPA): Eine Softwaretechnologie zur Automatisierung strukturierter, regelbasierter Geschäftsprozesse, die menschliche Handlungen in digitalen Systemen nachbildet. RPA interagiert mit Anwendungen, Formularen und Daten auf dieselbe Weise wie ein Benutzer – durch Klicken, Tippen, Kopieren und Verschieben von Informationen – jedoch schneller, konsistenter und fehlerfreier. Sie wird häufig eingesetzt, um administrative Prozesse mit hohem Volumen, etwa Dateneingabe, Rechnungsverarbeitung oder Datensatzverwaltung, zu optimieren. Im Unternehmenskontext wird RPA besonders leistungsfähig, wenn sie mit KI und Workflow-Orchestrierung kombiniert wird, wodurch skalierbare Aufgabenautomatisierung und intelligente Entscheidungsunterstützung ermöglicht werden.

Rules Engine: Ein Softwaresystem, das vordefinierte logische Regeln verwendet, um Entscheidungen automatisch zu treffen oder Prozesse zu steuern.

Skalierbarkeit: Die Fähigkeit eines Systems, einer Anwendung oder einer Infrastruktur, steigende Arbeitslasten oder Anforderungen durch Hinzufügen von Ressourcen wie Verarbeitungsleistung, Arbeitsspeicher oder Speicherplatz zu bewältigen. In der Computertechnik sorgt Skalierbarkeit für gleichbleibende Leistung und Zuverlässigkeit bei zunehmender Nutzung und kann sowohl nach oben (vertikale Skalierung) als auch nach unten (horizontale Skalierung) skaliert werden.

Stimmungs- und Emotionsanalyse: Softwareanwendungen, die natürliche Sprachverarbeitung (NLP), maschinelles Lernen oder statistische Methoden verwenden, um Meinungen oder Emotionen zu identifizieren und zu kategorisieren, die in Text, Sprache oder anderen Daten ausgedrückt werden. Mit diesen Tools können Unternehmen feststellen, ob die Stimmung positiv, negativ oder neutral ist, und sie werden häufig in Bereichen wie Kundenfeedback, Überwachung sozialer Medien und Marktforschung eingesetzt.

Strukturierte Daten: Daten, die in einem festgelegten Format gespeichert sind, meist in klar definierten Feldern oder Spalten eines Datensatzes oder einer Datei. Typische Beispiele sind relationale Datenbanken oder Tabellenkalkulationen, in denen Informationen nach festen Schemata organisiert werden.

Unstrukturierte Daten: Daten ohne einheitliche oder vordefinierte Struktur, die nicht an festen Positionen gespeichert sind. Beispiele sind Freitext in Textverarbeitungsdokumenten, E-Mails, Bilder, Audio- oder Videodateien.

Unternehmens-KI (Enterprise AI, EAI): Bezeichnet die strukturierte Anwendung künstlicher Intelligenz innerhalb einer Organisation zur Lösung konkreter Geschäftsprobleme, zur Verbesserung der Entscheidungsfindung und zur sicheren, skalierbaren Automatisierung von Arbeitsabläufen. Unternehmens-KI ist keine eigenständige Form von Intelligenz, sondern der kontrollierte Einsatz bestehender KI-Fähigkeiten – einschließlich maschinellen Lernens, Verarbeitung natürlicher Sprache, Computer Vision, Automatisierung und generativer KI – innerhalb einer regulierten Daten- und Compliance-Umgebung. Sie basiert auf vertrauenswürdigen Informationen, souveräner Datenkontrolle, Lebenszyklusmanagement (MLOps und LLMOps), hybrider oder souveräner Cloud-Infrastruktur und sicheren Orchestrierungsschichten, um KI im gesamten Unternehmen verantwortungsvoll, transparent und zuverlässig zu betreiben.

Verarbeitung natürlicher Sprache (Natural Language Processing, NLP): Ein Bereich der künstlichen Intelligenz (KI), der es Computern ermöglicht, menschliche Sprache zu verstehen, zu interpretieren und zu erzeugen. NLP kombiniert Linguistik, maschinelles Lernen und Computertechniken, um Anwendungen wie Chatbots, Übersetzung, Stimmungsanalyse und Spracherkennung zu unterstützen.

Verschlüsselung: Der Prozess der Konvertierung von Daten in ein verschlüsseltes Format mithilfe von Algorithmen und kryptografischen Schlüsseln, um unbefugten Zugriff zu verhindern. Bei der Datenverarbeitung stellt die Verschlüsselung sicher, dass nur autorisierte Parteien mit dem richtigen Schlüssel die Informationen entschlüsseln und lesen können, wodurch sensible Daten während der Speicherung oder Übertragung geschützt werden.

Vordenker: Ein Vordenker ist eine vorausschauende Person, die in der Lage ist, mögliche zukünftige Ereignisse und Entwicklungen frühzeitig zu erkennen. Diese Personen minimieren die mit potenziellen Ereignissen und Ergebnissen verbundenen Risiken. Ein Vordenker wird notwendige Änderungen im Voraus vornehmen oder Protokolle erstellen, die bei Bedarf umgesetzt werden können. Sie sind auf fast alles vorbereitet.

Wichtige Leistungsindikatoren (Key Performance Indicators, KPIs): Quantifizierbare Kennzahlen, anhand derer bewertet wird, wie effektiv eine Organisation, ein Team oder ein Prozess ihre Ziele erreicht. KPIs verfolgen den Fortschritt bei der Erreichung strategischer Ziele, leiten die Entscheidungsfindung und können von finanziellen Kennzahlen wie Umsatzwachstum bis hin zu operativen Kennzahlen wie Kundenbindung oder Systemverfügbarkeit reichen.

Workflow-Automatisierung: Eine Untergruppe der Automatisierung, die sich auf die Koordinierung und Ausführung einer Reihe von Aufgaben oder Prozessen über Systeme, Anwendungen oder Teams hinweg konzentriert. Die Workflow-Automatisierung bildet die Schritte in einem geschäftlichen oder technischen Prozess ab und nutzt die Automatisierung, um sicherzustellen, dass sie in der richtigen Reihenfolge und mit einem Minimum an manuellen Eingaben ausgeführt werden.

Zero-Trust: Ein Sicherheitsframework, das davon ausgeht, dass standardmäßig keinem Benutzer, Gerät oder System vertraut werden sollte, sei es innerhalb oder außerhalb des Unternehmensnetzwerks. Im Computerbereich erfordert Zero-Trust eine kontinuierliche Überprüfung der Identität, strenge Zugriffskontrollen und die Überwachung aller Aktivitäten, um Risiken zu minimieren und sensible Daten zu schützen.

Quellenangabe

Agrawal, A., Gans, J., und Goldfarb, A. *„Prediction Machines: The Simple Economics of Artificial Intelligence."* (Vorhersagemaschinen: Die einfache Ökonomie der künstlichen Intelligenz) Harvard Business Review Press, 2022.

„AI Governance Framework: Transparency, Explainability, and Contestability (TEC)." (Rahmenwerk für KI-Governance: Transparenz, Erklärbarkeit und Anfechtbarkeit) AI-Governance.eu, 2024. https://ai-governance.eu/ai-governance-framework/tec/. (Zugriff Okt. 2025).

„AI Governance Software Spend Will See 30% CAGR From 2024 to 2030." (Ausgaben für KI-Governance-Software werden von 2024 bis 2030 jährlich um 30 % steigen) Forrester Blog, 13. Nov. 2024. www.forrester.com/blogs/ai-governance-software-spend-will-see-30-cagr-from-2024-to-2030/. (Zugriff Okt. 2025).

Barrenechea, Mark J. und Tom Jenkins. *„Digital Financial Services."* (Digitale Finanzdienstleistungen) OpenText Corporation, 2016.

Barrenechea, Mark J. und Tom Jenkins. *„Digital Manufacturing."* (Digitale Fertigung) OpenText Corporation, 2018.

Barrenechea, Mark J. und Tom Jenkins. *„e-Government or Out of Government."* (E-Government oder kein Government) OpenText Corporation, 2014.

Barrenechea, Mark J. und Tom Jenkins. *„Enterprise Information Management: The Next Generation of Enterprise Software."* (Enterprise Information Management: Die nächste Generation von Unternehmenssoftware) OpenText Corporation, 2013.

Barrenechea, Mark J., Tom Jenkins und David Fraser. *„The Anticipant Organization."* (Die antizipierende Organisation) OpenText Corporation, 2022.

Biggio, B., Nelson, B. und P. Laskov. *„Poisoning Attacks Against Support Vector Machines."* (Vergiftungsangriffe auf Support-Vector-Machines) Proceedings of the 29th International Conference on Machine Learning (ICML), 2012.

Boyd, K. *„Microsoft 365 Copilot for Executives: Sharing Our Customer Zero Deployment and Adoption Journey at Microsoft."* (Microsoft 365 Copilot für Führungskräfte: Einblicke in die interne Einführung und Implementierungsreise bei Microsoft) Microsoft Inside Track Blog, 5. Dez. 2024. https://www.microsoft.com/insidetrack/blog/copilot-for-microsoft-365-for-executives-sharing-our-internal-deployment-and-adoption-journey-at-microsoft/. (Zugriff Okt. 2025).

Bubeck, Sébastien, Chandrasekaran, Varun, Eldan, Ronen u. a. *„Sparks of Artificial General Intelligence: Early Experiments with GPT-4."* (Funken künstlicher allgemeiner Intelligenz: Frühe Experimente mit GPT-4) Cornell University, arXiv:2303.12712, 13. Apr. 2023. https://arxiv.org/abs/2303.12712. (Zugriff Okt. 2025).

Challapally, Aditya, Pease, Chris, Raskar, Ramesh et al. *„The GenAI Divide: The State of AI in Business 2025."* (Die GenAI-Kluft: Der Stand der KI in der Wirtschaft 2025) MIT NANDA Report. MIT Sloan School of Management, Juli 2025. https://mlq.ai/media/quarterly_decks/v0.1_State_of_AI_in_Business_2025_Report.pdf. Zugriff Okt. 2025).

Crevier, Daniel. *„AI: The tumultuous history of the search for artificial intelligence."* (KI: Die bewegte Geschichte der Suche nach künstlicher Intelligenz) Basic Books, 1993.

„Cyber Risks Associated with Generative Artificial Intelligence." (Cyberrisiken im Zusammenhang mit generativer künstlicher Intelligenz) Monetary Authority of Singapore (MAS), Circular No. TRPD-G01-2024, August 2024. https://www.mas.gov.sg/-/media/mas-media-library/regulation/circulars/trpd/cyber-risks-associated-with-generative-artificial-intelligence.pdf. (Zugriff Oktober 2025).

Daugherty, Paul, Ghosh, Bhaskar, Narain, Karthik, et al. *„A new generative era of AI for everyone."* (Eine neue generative Ära der KI für alle) Accenture, 2023.

„Data Sovereignty as Your Foundation Layer." (Datensouveränität als Fundamentebene) Katonic Blog, 13. Oktober 2025. https://www.katonic.ai/blog/building-your-ai-stack-data-sovereignty-as-your-foundation-layer. (Zugriff November 2025).

„Deepfake and AI Phishing Statistics (2024)." (Deepfake- und KI-Phishing-Statistiken (2024)) ZeroThreat.ai. ZeroThreat, 2024. https://zerothreat.ai/blog/deepfake-and-ai-phishing-statistics. (Zugriff Oktober 2025).

Edquist, Alex, Grennan, Liz, Griffiths, Sian et al. *„Data ethics: What it means and what it takes."* (Datenethik: Bedeutung und Voraussetzungen) McKinsey & Company, 23. September 2022. https://www.mckinsey.com/capabilities/mckinsey-digital/our-insights/data-ethics-what-it-means-and-what-it-takes. (Zugriff Oktober 2025).

European Commission. *„Regulation (EU) 2024/1689 on Artificial Intelligence."* (Verordnung (EU) 2024/1689 über künstliche Intelligenz) 2024.

European Parliament & Council. *„Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data."* (Verordnung (EU) 2016/679 des Europäischen Parlaments und des Rates vom 27. April 2016 zum Schutz natürlicher Personen bei der Verarbeitung personenbezogener Daten) Artikel 17 (Recht auf Löschung), 2016. https://eur-lex.europa.eu/eli/reg/2016/679/oj. (Zugriff Oktober 2025).

Cooper, A. Feder, Choquette-Choo, Christopher A., Bogen, Miranda et al. *„Machine Unlearning Doesn't Do What You Think: Lessons for Generative AI Policy, Research, and Practice."* (Maschinelles Vergessen funktioniert nicht wie gedacht: Erkenntnisse für Politik, Forschung und Praxis der generativen KI) SSRN, 6. Februar 2025. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=5060253. (Zugriff Oktober 2025).

Foundry Research, gesponsert von OpenText. *„MarketPulse Survey: The Role of GenAI in Modernizing Content Management."* (MarketPulse-Umfrage: Die Rolle von GenAI bei der Modernisierung des Content-Managements) Mai 2025.

„Gartner Poll Finds 55% of Organizations Have an AI Board." (Gartner-Umfrage: 55 % der Organisationen verfügen über ein KI-Gremium) Gartner, Inc. Pressemitteilung, 26. Juni 2024.

„Gartner Says More Than 80% of Enterprises Will Have Used Generative AI APIs or Deployed Generative AI-Enabled Applications by 2026." (Gartner prognostiziert: Über 80 % der Unternehmen werden bis 2026 generative KI-APIs oder KI-fähige Anwendungen einsetzen) Gartner, Inc. Pressemitteilung, 11. Oktober 2023. www.gartner.com/en/newsroom/press-releases/2023-10-11-gartner-says-more-than-80-percent-of-enterprises-will-have-used-generative-ai-apis-or-deployed-generative-ai-enabled-applications-by-2026. (Zugriff Oktober 2025).

„Gartner Survey Reveals GenAI Attacks Are on the Rise." (Gartner-Umfrage zeigt: GenAI-Angriffe nehmen zu) Gartner, Inc., 22. September 2025. https://www.gartner.com/en/newsroom/press-releases/2025-09-22-gartner-survey-reveals-generative-artificial-intelligence-attacks-are-on-the-rise. (Zugriff Oktober 2025).

Goodfellow, Ian J., Shlens, Jonathon und Christian Szegedy. *„Explaining and Harnessing Adversarial Examples."* (Erklärung und Nutzung adversarialer Beispiele) Cornell University, arXiv:1412.6572, 20. März 2015. https://arxiv.org/abs/1412.6572. (Zugriff Oktober 2025).

Gownder, J. P. *„The Artificial Intelligence Pathway to the Future of Work."* (Der Weg der künstlichen Intelligenz in die Zukunft der Arbeit) Forrester Research, Juni 2023.

Gu, Tianyu, Dolan-Gavitt, Brendan und Siddharth Garg. *„BadNets: Identifying Vulnerabilities in the Machine Learning Model Supply Chain."* (BadNets: Aufdeckung von Schwachstellen in der Lieferkette von Machine-Learning-Modellen) Cornell University, arXiv:1708.06733, 11. März 2019. https://arxiv.org/abs/1708.06733. (Zugriff Oktober 2025).

Hintze, Arend. *„Understanding the four types of AI, from reactive robots to self-aware beings."* (Die vier Arten von KI: Von reaktiven Robotern bis zu selbstbewussten Wesen) The Conversation, 13. November 2016. https://theconversation.com/understanding-the-four-types-of-ai-from-reactive-robots-to-self-aware-beings-67616. (Zugriff Oktober 2025).

„How we built our multi-agent research system." (Wie wir unser Multiagenten-Forschungssystem aufgebaut haben) Anthropic, 2024. https://www.anthropic.com/engineering/multi-agent-research-system. (Zugriff Oktober 2025).

„Information Governance Reference Model." (Referenzmodell für Informationsgovernance) EDRM http://www.edrm.net/projects/igrm. (Zugriff Oktober 2025).

International Organization for Standardization. *„Artificial Intelligence Standards Portfolio."* (Normenportfolio für künstliche Intelligenz) ISO/IEC JTC 1/SC 42, 2023.

International Organization for Standardization. *„Information Security, Cybersecurity and Privacy Protection—Information Security Management Systems—Requirements."* (Informationssicherheit, Cybersicherheit und Schutz der Privatsphäre – Informationssicherheits-Managementsysteme – Anforderungen) ISO/IEC 27001:2022.

International Organization for Standardization. *„What is Artificial Intelligence (AI)?"* (Was ist künstliche Intelligenz (KI)?) 31. Januar 2024. https://www.iso.org/artificial-intelligence/what-is-ai?. (Accessed Oct. 2025).

Jangam, Sandeep Kumar. *„Importance of Encrypting Data in Transit and at Rest Using TLS and Other Security Protocols and API Security Best Practices."* (Die Bedeutung der Datenverschlüsselung bei Übertragung und Speicherung unter Verwendung von TLS, anderer Sicherheitsprotokolle und bewährter API-Sicherheitspraktiken) *International Journal of AI, BigData, Computational and Management Studies*, 4(3), 82–91, 2023. https://ijaibdcms.org/index.php/ijaibdcms/article/view/242/. (Zugriff Oktober 2025).

Jenkins, Tom. *„Behind the Firewall: Big Data and the Hidden Web: The Path to Enterprise Information Management."* (Hinter der Firewall: Big Data und das versteckte Web – Der Weg zum Enterprise Information Management) OpenText Corporation, 2012.

Jenkins, Tom. *„Enterprise Content Management: What You Need to Know."* (Enterprise Content Management: Was Sie wissen müssen) OpenText Corporation, 2004.

Jenkins, Tom. *„Managing Content in the Cloud: Enterprise Content Management 2.0."* (Content-Verwaltung in der Cloud: Enterprise Content Management 2.0) OpenText Corporation, 2011.

Jiang, Shuli, Kadhe, Swanand Ravindra, Zhou, Yi et al. *„Forcing Generative Models to Degenerate Ones: The Power of Data Poisoning Attacks."* (Generative Modelle zur Degeneration zwingen: Die Macht von Datenvergiftungsangriffen) Cornell University, arXiv:2312.04748, 7. Dezember 2023. https://arxiv.org/abs/2312.04748. (Zugriff Oktober 2025).

Joshi, Akshay, Moschetta, Giulia und Ellie Winslow. *„Global Cybersecurity Outlook 2025 Insight Report."* (Globaler Cybersicherheitsausblick 2025 – Insight Report) World Economic Forum in Zusammenarbeit mit Accenture, Januar 2025. https://reports.weforum.org/docs/WEF_Global_Cybersecurity_Outlook_2025.pdf. (Zugriff Oktober 2025).

Kandogan, Eser, Bhutani, Nikita, Zhang, Dan et al. *„Orchestrating Agents and Data for Enterprise: A Blueprint Architecture for Compound AI."* (Koordination von Agenten und Daten im Unternehmen: Eine Referenzarchitektur für zusammengesetzte KI) arXiv Preprint, 10. April 2025. https://arxiv.org/abs/2504.08148. (Accessed Nov. 2025).

Kaplan, Jared, McCandlish, Sam, Henighan, Tom et al. *„Scaling Laws for Neural Language Models."* (Skalierungsgesetze für neuronale Sprachmodelle) Cornell University, arXiv:2001.08361, 23. Januar 2020. https://arxiv.org/abs/2001.08361. (Zugriff Oktober 2025).

„Key Regulatory and Industry Initiatives." (Wichtige regulatorische und branchenweite Initiativen) *Capgemini*. https://web.archive.org/web/20141105171058/https://www.worldpaymentsreport.com/kriis#Heat-Map-of-KRIIs-Global-and-Regional. (Zugriff Oktober 2025).

„Key Terms for AI Governance." (Schlüsselbegriffe für KI-Governance) International Association of Privacy Professionals (IAPP), 2024. https://iapp.org/resources/article/key-terms-for-ai-governance/. (Zugriff Oktober 2025).

Kourinian, Arsen und Mayer Brown. *„Addressing Transparency & Explainability When Using AI Under Global Standards."* (Transparenz und Nachvollziehbarkeit bei der Nutzung von KI nach globalen Standards sicherstellen) Bloomberg Law, 2024. https://www.mayerbrown.com/-/media/files/perspectives-events/publications/2024/01/addressing-transparency-and-explainability-when-using-ai-under-global-standards.pdf. (Zugriff Oktober 2025).

Kurakin, Alexey, Goodfellow, Ian und Samy Bengio. *„Adversarial Machine Learning at Scale."* (Adversariales maschinelles Lernen im großen Maßstab) Cornell University, arXiv:1611.01236, 11. Februar 2017. https://arxiv.org/abs/1611.01236. (Zugriff Oktober 2025).

LeCun, Yann, Bengio, Yoshua und Geoffrey Hinton. „Deep learning." (Tiefes Lernen) Nature, 521(7553), 436–444, 2015. https://doi.org/10.1038/nature14539. (Zugriff Oktober 2025).

Lu, Ruei-Shan, Lin, Ching-Chang und Hsiu-Yuan Tsao. *„Empowering Large Language Models to Leverage Domain-Specific Knowledge in E-Learning."* (Große Sprachmodelle für die Nutzung domänenspezifischen Wissens im E-Learning befähigen) Applied Sciences, 14(12), 5264, 18. Juni 2024. https://doi.org/10.3390/app14125264. (Zugriff Oktober 2025).

Lutkevich, Ben. *„What is AI Winter? Definition, History and Timeline."* (Was ist der KI-Winter? Definition, Geschichte und Zeitstrahl) Tech Target, 26. August 2024. https://www.techtarget.com/searchenterpriseai/definition/AI-winter. (Zugriff Oktober 2025).

Marcus, Gary. *„Deep learning is hitting a wall."* (Deep Learning stößt an seine Grenzen) Communications of the ACM, 65(8), 36–43, 2022. https://nautil.us/deep-learning-is-hitting-a-wall-238440/. (Zugriff Oktober 2025).

Maisto, Dario. *„From Digital Sovereignty Platforms to Sovereign Cloud Platforms: Three Reasons for a Title Change."* (Von digitalen Souveränitätsplattformen zu souveränen Cloud-Plattformen: Drei Gründe für eine Titeländerung) Forrester Blogs, 11. August 2025. www.forrester.com/blogs/from-digital-sovereignty-platforms-to-sovereign-cloud-platforms-three-reasons-for-a-title-change/. (Zugriff Oktober 2025).

McCarthy, J., Minsky, M. L., Rochester, N. et al. *„A proposal for the Dartmouth summer research project on artificial intelligence."* (Vorschlag für das Dartmouth-Sommerforschungsprojekt zur künstlichen Intelligenz) Dartmouth College, 1955. https://ojs.aaai.org/aimagazine/index.php/aimagazine/article/view/1904. (Zugriff Oktober 2025).

Mienye, I. D., Jere, N., Obaido, G. et al. *„Large language models: an overview of foundational architectures, recent trends, and a new taxonomy."* (Große Sprachmodelle: Überblick über grundlegende Architekturen, aktuelle Trends und eine neue Klassifizierung) Discover Applied Sciences, 7, 1027, 2025. https://doi.org/10.1007/s42452-025-07668-w. (Zugriff Oktober 2025).

McKinsey & Company. *„Future-Proofing the IT Function Amid Global Trends and Disruptions."* (Zukunftssicherung der IT-Funktion angesichts globaler Trends und Umbrüche) McKinsey Digital, 11. Juni 2025. www.mckinsey.com/capabilities/mckinsey-digital/our-insights/tech-forward/future-proofing-the-it-function-amid-global-trends-and-disruptions. (Zugriff Oktober 2025).

McKinsey & Company. *„The State of AI in Early 2024: Gen AI Adoption Spikes and Starts to Generate Value."* (Der Stand der KI Anfang 2024: GenAI-Adoption steigt sprunghaft und beginnt, Wert zu schaffen) QuantumBlack by McKinsey, 30. Mai 2024. https://www.mckinsey.com/capabilities/quantumblack/our-insights/the-state-of-ai-2024. (Zugriff Oktober 2025).

Mehrabi, Ninareh, Morstatter, Fred, Saxena, Nripsuta et al. *„A Survey on Bias and Fairness in Machine Learning."* (Eine Übersicht über Verzerrungen und Fairness im maschinellen Lernen) Cornell University, arXiv:1908.09635, 25. Januar 2022. https://arxiv.org/abs/1908.09635. (Zugriff Oktober 2025).

Mei, Lingrui, Yao, Jiayu, Ge, Yuyao et al. *„A Survey of Context Engineering for Large Language Models."* (Eine Umfrage zum Context Engineering für große Sprachmodelle) Cornell University, arXiv:2507.13334, 21. Juli 2025. https://arxiv.org/abs/2507.13334. (Zugriff Oktober 2025).

Miller, Philip. *„Unlocking Unstructured Data: Fueling AI with Insights."* (Erschließung unstrukturierter Daten: Wie KI durch Erkenntnisse angetrieben wird) Dataversity, 3. Juni 2025. https://www.dataversity.net/articles/unlocking-unstructured-data-fueling-ai-with-insights/. (Zugriff Oktober 2025).

Moor, James. *„What is Computer Ethics?"* (Was ist Computerethik?) Metaphilosophy, 16(4), 266–275, 1985, https://doi.org/10.1111/j.1467-9973.1985.tb00173.x. (Zugriff Oktober 2025).

„More Than 80% of Enterprises Will Have Used Generative AI APIs or Deployed Generative AI Applications by 2026." (Mehr als 80 % der Unternehmen werden bis 2026 generative KI-APIs oder -Anwendungen nutzen) Gartner Inc. Pressemitteilung, 11. Oktober 2023. www.gartner.com/en/newsroom/press-releases/2023-10-11-gartner-says-more-than-80-percent-of-enterprises-will-have-used-generative-ai-apis-or-deployed-generative-ai-enabled-applications-by-2026. (Zugriff Oktober 2025).

Mucci, Tim und Cole Stryker. *„What is Artificial Superintelligence?"* (Was ist künstliche Superintelligenz?) IBM, 22. Juli 2025. https://www.ibm.com/think/topics/artificial-superintelligence. (Zugriff Oktober 2025).

„New Accenture Research Finds that Companies with AI-Led Processes Outperform Peers." (Neue Accenture-Studie zeigt: Unternehmen mit KI-gesteuerten Prozessen übertreffen ihre Wettbewerber) Accenture, 10. Oktober 2024. https://newsroom.accenture.com/news/2024/new-accenture-research-finds-that-companies-with-ai-led-processes-outperform-peers. (Zugriff Oktober 2025).

NIST. *„Artificial Intelligence Risk Management Framework (AI RMF 1.0)."* (Rahmenwerk für das Risikomanagement künstlicher Intelligenz – AI RMF 1.0) NIST Special Publication AI 100-1, 2023.

Organisation for Economic Co-operation and Development. *„Recommendation of the Council on Artificial Intelligence."* (Empfehlung des Rates zur künstlichen Intelligenz) OECD/LEGAL/0449, 2019.

O'Grady, Michael und Michele Goetz et al. *„Global Commercial AI Software Governance Market Forecast, 2024 to 2030."* (Weltweiter Prognosebericht zum Markt für Governance von kommerzieller KI-Software, 2024–2030) *Forrester Research*, 1. November 2024.

„Predicts 2025: Data and Analytics Strategy—Unlocking Value with AI and Governance." (Prognosen 2025: Daten- und Analyse-Strategien – Wertschöpfung durch KI und Governance) *Gartner Inc.*, 2024.

„Product Owner." (Produktverantwortlicher) Scaled Agile Framework, 25. Februar 2025. https://framework.scaledagile.com/product-owner. (Zugriff Oktober 2025).

Rabot, M. *„Winning in the Autonomous AI Agents Race."* (Erfolgreich im Rennen um autonome KI-Agenten) Medium, 4. April 2025. https://rabot.medium.com/winning-in-the-autonomous-ai-agents-race-a0c03d52acad. (Zugriff Oktober 2025).

Rose, Scott, Borchert, Oliver, Mitchell, Stu et al. *„Zero Trust Architecture."* (Zero-Trust-Architektur) NIST Special Publication 800-207, August 2020. https://doi.org/10.6028/NIST.SP.800-207. (Zugriff Oktober 2025).

Rowe, Adam. *„MIT Finds 95% of Enterprise AI Pilots Fail to Deliver Revenues."* (MIT stellt fest: 95 % der KI-Pilotprojekte in Unternehmen generieren keine Umsätze) Tech.co, 20. August 2025. https://tech.co/news/mit-enterprise-ai-pilots-fail-revenues. (Zugriff Oktober 2025).

Russell, Melissa. *„How Can I Learn Artificial Intelligence?"* (Wie kann man künstliche Intelligenz lernen?) Harvard, 8. April 2025. https://extension.harvard.edu/blog/how-can-i-learn-artificial-intelligence/#What-is-Artificial-Intelligence. (Zugriff Oktober 2025).

Russell, S. J. und P. Norvig. *Artificial Intelligence: A Modern Approach*. (Künstliche Intelligenz: Ein moderner Ansatz) 4. Auflage, Pearson, 2021.

Sanchez, Jarvy. *„Enterprise AI Architecture | Components & Best Practices."* (Enterprise-KI-Architektur – Komponenten und bewährte Verfahren) *Leanware*, 28. August 2025. https://www.leanware.co/insights/enterprise-ai-architecture. (Zugriff November 2025).

Samoili, S., López Cobo, M., Delipetrev, B. et al. *„AI Watch, Defining Artificial Intelligence 2.0: Towards an Operational Definition and Taxonomy for the AI Landscape."* (AI Watch: Definition von Artificial Intelligence 2.0 – Hin zu einer operativen Definition und Taxonomie der KI-Landschaft) *Publications Office of the European Union*, 2021.

Semba, Kurt. *„Artificial Intelligence, Real Consequences: Confronting AI's Growing Energy Appetite."* (Künstliche Intelligenz, reale Folgen: Den wachsenden Energiehunger der KI bewältigen) Extreme Networks, 15. August 2024. https://www.extremenetworks.com/resources/blogs/confronting-ai-growing-energy-appetite-part-1. (Zugriff Oktober 2025).

Singla, Alex, Sukharevsky, Alexander, Yee, Lareina et al. *The State of AI: How Organizations Are Rewiring to Capture Value*. (Der Stand der KI: Wie Organisationen sich neu aufstellen, um Wert zu schaffen) McKinsey & Company, 2025.

Stanford University. *AI Index Report 2025*. (AI Index Bericht 2025) Stanford Institute for Human-Centered Artificial Intelligence, 2025.

Sukharevsky, Alexander, Krivkovich, Alexis, Gast, Arne et al. *„The Agentic Organization: Contours of the Next Paradigm for the AI Era."* (Die agentenbasierte Organisation: Konturen des nächsten Paradigmas im KI-Zeitalter) McKinsey & Company, 26. September 2025. https://www.mckinsey.com/capabilities/people-and-organizational-performance/our-insights/the-agentic-organization-contours-of-the-next-paradigm-for-the-ai-era#/. (Zugriff Oktober 2025).

Sunkara, V. L. *„KPIs for AI Agents and Generative AI: A Rigorous Framework for Evaluation and Accountability."* (KPIs für KI-Agenten und generative KI: Ein strenges Rahmenwerk für Bewertung und Verantwortlichkeit) International Journal of Scientific Research and Modern Technology, 3(4), 22–29, 2024. https://doi.org/10.38124/ijsrmt.v3i4.572. (Zugriff Oktober 2025).

Tegmark, M. *Life 3.0: Being Human in the Age of Artificial Intelligence*. (Leben 3.0: Mensch sein im Zeitalter künstlicher Intelligenz) Vintage Books, Penguin Random House LLC, 2018.

„The State of AI in Early 2024: Gen AI Adoption Spikes and Starts to Generate Value." (Der Stand der KI Anfang 2024: GenAI-Adoption steigt sprunghaft und beginnt, Wert zu schaffen) McKinsey & Company, 30. Mai 2024. https://www.mckinsey.com/capabilities/quantumblack/our-insights/the-state-of-ai-2024. (Zugriff Oktober 2025).

„Towards a Unified Agent with Foundation Models." (Auf dem Weg zu einem vereinheitlichten Agenten mit Foundation Models) Reddit – r/MachineLearning, 2024. https://www.reddit.com/r/MachineLearning/comments/155wa2p/r_towards_a_unified_agent_with_foundation_models/. (Zugriff Oktober 2025).

Turing, A. M. *„Computing Machinery and Intelligence."* (Rechenmaschinen und Intelligenz) Mind, Band LIX, Ausgabe 236, Oktober 1950. https://doi.org/10.1093/mind/lix.236.433. (Zugriff Oktober 2025).

Uchida, Yusuke, Nagai, Yuki, Sakazawa, Shigeyuki et al. *„Embedding Watermarks into Deep Neural Networks."* (Einbettung von Wasserzeichen in tiefe neuronale Netze) Cornell University, arXiv:1701.04082, 20. April 2017. https://arxiv.org/abs/1701.04082. (Zugriff Oktober 2025).

UNESCO. *„Recommendation on the Ethics of Artificial Intelligence."* (Empfehlung zur Ethik der künstlichen Intelligenz) UNESCO.org, 2022. https://www.unesco.org/en/articles/recommendation-ethics-artificial-intelligence. (Zugriff Oktober 2025).

University of Michigan College of Engineering. *„Up to 30% of the Power Used to Train AI Is Wasted — Here's How to Fix It."* (Bis zu 30 % der für das Training von KI verwendeten Energie wird verschwendet – so lässt sich das beheben) Michigan Engineering News, 12. November 2024. https://news.engin.umich.edu/2024/11/up-to-30-of-the-power-used-to-train-ai-is-wasted-heres-how-to-fix-it/. (Zugriff Oktober 2025).

„What is Artificial Intelligence (AI)?“ (WAS IST KÜNSTLICHE INTELLIGENZ (KI)?) *International Organization for Standardization*, 31. Januar 2024. https://www.iso.org/artificial-intelligence/what-is-ai?

Wharton School of the University of Pennsylvania. *„The Hidden Cost of AI: Energy Consumption.“* (Die versteckten Kosten der KI: Energieverbrauch) Knowledge@Wharton, 25. April 2024. https://knowledge.wharton.upenn.edu/article/the-hidden-cost-of-ai-energy-consumption/. (Zugriff Oktober 2025).

Yang, Wencheng, Wang, Song, Wu, Di et al. *„Deep Learning Model Inversion Attacks and Defenses: A Comprehensive Survey.“* (Inversionsangriffe und Abwehrmechanismen bei Deep-Learning-Modellen: Eine umfassende Übersicht) Cornell University, arXiv:2501.18934, 30. April 2025. https://arxiv.org/abs/2501.18934. (Zugriff Oktober 2025).

Yang, Qiang, Liu, Yang, Chen, Tianjian et al. *„Federated Machine Learning: Concept and Applications.“* (Föderiertes maschinelles Lernen: Konzept und Anwendungen) ACM Digital Library, 28. Januar 2019. https://dl.acm.org/doi/10.1145/3298981. (Zugriff Oktober 2025).

Zhang, Qizheng, Hu, Changran, Upasani, Shubhangi et al. *„Agentic Context Engineering: Evolving Contexts for Self-Improving Language Models.“* (Agentisches Kontext-Engineering: Dynamische Kontexte für selbstverbessernde Sprachmodelle) arXiv Preprint arXiv:2510.04618, 2025. (https://www.arxiv.org/pdf/2510.04618. Zugriff Oktober 2025).

Zimmermann, Annette und Danielle Casey. *Emerging Tech Impact Radar: Generative AI*. (Aufkommender Technologie-Impact-Radar: Generative KI) Gartner, 2025.

Index

A

B

C

D

E

F

G

H

I

K

L

M

R

S

T

U

V

W

Z

Enterprise Artificial Intelligence: Aufbau einer vertrauenswürdigen KI in der Souveränen Cloud

In der Geschichte der Technologie gab es immer wieder Momente, in denen sich alles auf einmal veränderte – die Einführung des Webs, der Aufstieg des Cloud Computing, die mobile Revolution. Jede dieser Wellen hat die Art und Weise, wie Unternehmen arbeiten, wie Branchen konkurrieren und wie Menschen leben und arbeiten, grundlegend verändert.

Doch keine dieser Umwälzungen lässt sich mit dem Aufstieg der KI im kognitiven Zeitalter vergleichen – weder in ihrer Geschwindigkeit noch in ihrem Ausmaß oder in ihren Auswirkungen.

Künstliche Intelligenz hat sich vom Rand der Unternehmensstrategie in ihr Zentrum bewegt. Sie ist keine Forschungsinitiative oder experimentelle Ergänzung mehr, sondern der neue Motor für Produktivität, Innovation und Wettbewerbsfähigkeit.

Enterprise Artificial Intelligence: Aufbau einer vertrauenswürdigen KI in der Souveränen Cloud zeigt, warum die nächste Innovationsära jenen Organisationen gehört, die Informationen als ihren strategisch wichtigsten Vermögenswert begreifen. Aufbauend auf jahrzehntelanger Erfahrung im sicheren und regelkonformen Informationsmanagement wird dargelegt, wie Unternehmen und Regierungen KI entwickeln können, die nicht nur leistungsfähig, sondern auch kontrolliert und sicher ist.

Von Hyperscale-Clouds bis hin zu nationalen Datenrichtlinien untersucht das Buch, wie sich Geschwindigkeit mit Verantwortung, Automatisierung mit Rechenschaftspflicht und Intelligenz mit Integrität in Einklang bringen lassen. Denn Erfolg wird nicht davon abhängen, wer die intelligentesten Systeme entwickelt, sondern wer die Systeme schafft, denen man vertrauen kann.

Die Zukunft der Unternehmen ist intelligent. Die Zukunft der Intelligenz ist kontrolliert. Und die Arbeit daran beginnt jetzt.

12.25 | 243-DE0119-002