

WHITEPAPER

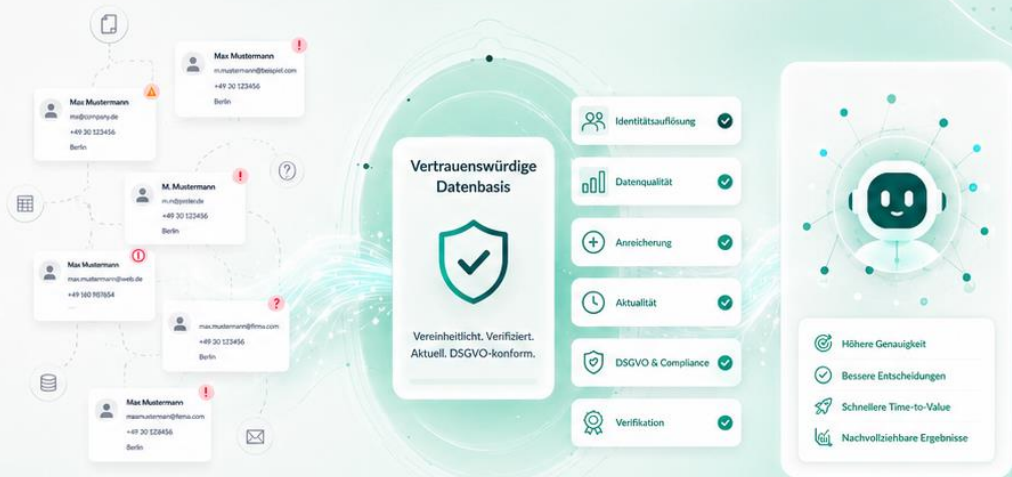
KI-Projekte scheitern nicht an KI

Warum Datenqualität, Identitätsauflösung und Compliance über den Erfolg von KI entscheiden - und weshalb Trusted Data zur Grundlage wirtschaftlich nutzbarer KI wird.

von Sylvia Türcke, Geschäftsführerin Capaneo GmbH
veröffentlicht am 19.06.2026

Die meisten KI-Projekte scheitern nicht an KI

... sondern an schlechter **Datenbasis**, fehlender **Identitätsauflösung**, mangelnder **Datenqualität**, fehlender **Aktualität** und **DSGVO-Risiken**.



Trusted Data for Agentic AI



Datenchaos

Fragmentierte, doppelte und inkonsistente Daten bremsen Ihre KI-Projekte aus.



Trusted Data Foundation

Saubere, angereicherte und verifizierte Daten bilden das Fundament für verlässliche KI.



Nutzbare KI

Verlässliche Daten. Intelligente Agenten. Messbarer Impact für Ihr Unternehmen.

Executive Summary

Künstliche Intelligenz ist heute leichter zugänglich als je zuvor. Unternehmen können leistungsfähige Modelle, Agenten und Automatisierungslösungen vergleichsweise schnell testen. Trotzdem bleiben viele KI-Projekte hinter den Erwartungen zurück. Der Grund liegt häufig nicht in der KI selbst, sondern in der Datenbasis, auf der sie arbeiten soll.

Kernaussage

Der Wettbewerbsvorteil entsteht zunehmend nicht durch den bloßen Zugang zu KI-Modellen, sondern durch die Fähigkeit, verlässliche, aktuelle, angereicherte und rechtssicher nutzbare Daten bereitzustellen.

- KI benötigt eindeutige Identitäten, aktuelle Informationen, Zielgruppenlogik, Anreicherung und Qualitätskontrolle.
- Fehlende Datenqualität führt zu falschen Scores, unzuverlässigen Agenten, ineffizienten Kampagnen und Compliance-Risiken.
- Trusted Data wird zur Grundlage für wirtschaftlich nutzbare KI in Marketing, Vertrieb, Service und Analyse.
- Unternehmen sollten KI-Projekte deshalb nicht nur technologisch, sondern datenstrategisch vorbereiten.

Für wen dieses Whitepaper relevant ist

Dieses Whitepaper richtet sich an Entscheiderinnen und Entscheider, die KI nicht nur testen, sondern produktiv nutzen möchten - insbesondere in Marketing, Vertrieb, CRM, Data Management, Business Development, Analyse und Customer Experience.

- Marketing- und Vertriebsverantwortliche, die Zielgruppen, Kampagnen und Lead-Prozesse datenbasiert verbessern möchten.
- CRM- und Data-Teams, die Datenqualität, Dubletten, Golden Records und Datenfusion verantworten.
- Geschäftsführungen und Digitalverantwortliche, die KI-Initiativen skalierbar und compliance-sicher aufsetzen wollen.
- Fachbereiche, die KI-Agenten oder Automatisierungslösungen mit belastbaren Daten versorgen müssen.

Inhaltsverzeichnis

1. Warum KI-Projekte häufig an der Datenbasis scheitern
2. Data beats AI: Der Wettbewerbsvorteil liegt nicht im Modell
3. Trusted Data for Agentic AI
4. Identitätsauflösung als unterschätzter Erfolgsfaktor
5. Datenanreicherung macht KI kontextfähig
6. Compliance ist kein Zusatzthema, sondern Voraussetzung
7. Framework: Die sechs Grundlagen KI-fähiger Daten
8. Checkliste: Ist Ihre Datenbasis bereit für KI?
9. Handlungsempfehlungen für den Einstieg
10. Fazit und Autorin

1. Warum KI-Projekte häufig an der Datenbasis scheitern

Künstliche Intelligenz ist längst kein Zukunftsthema mehr. Unternehmen experimentieren mit Chatbots, automatisierten Workflows, KI-Agenten, Prognosemodellen, Lead Scoring, Personalisierung und generativer Content-Erstellung. Die verfügbaren Modelle werden leistungsfähiger, günstiger und leichter zugänglich. GPT, Claude, Gemini, Mistral und andere Systeme lassen sich heute vergleichsweise schnell testen, integrieren und produktiv einsetzen.

Trotzdem bleiben viele KI-Projekte hinter den Erwartungen zurück. Nicht, weil die Modelle zu schwach wären. Nicht, weil die Technologie grundsätzlich nicht funktioniert. Sondern weil die Grundlage fehlt, auf der KI sinnvoll arbeiten kann: verlässliche, strukturierte, aktuelle und rechtssichere Daten.

Viele Unternehmen starten KI-Projekte mit einer technologischen Fragestellung: Welches Modell nutzen wir? Welche Plattform ist die richtige? Welche Agentenarchitektur brauchen wir? Wie integrieren wir KI in CRM, Marketing Automation, Vertrieb oder Service? Diese Fragen sind wichtig. Sie greifen aber zu kurz, wenn die darunterliegende Datenbasis unsauber ist.

Typische Datenprobleme vor KI-Projekten

Dubletten, veraltete Informationen, unvollständige Datensätze, fehlende Identitätsauflösung, Daten-Silos, uneinheitliche Formate und ungeklärte DSGVO- oder Compliance-Anforderungen.

Das Problem zeigt sich besonders deutlich in Marketing und Vertrieb. Ein KI-Agent kann keine sinnvolle Kundenpriorisierung vornehmen, wenn nicht klar ist, welche Unternehmen tatsächlich aktiv sind. Ein Scoringmodell kann keine Kaufwahrscheinlichkeit bewerten, wenn Branchen, Unternehmensgrößen, Standorte oder Entscheidungsstrukturen fehlen. Eine automatisierte Kampagne kann nicht effizient angesteuert werden, wenn Dubletten, veraltete Kontaktdaten und uneinheitliche Profile im System liegen.

2. Data beats AI: Der Wettbewerbsvorteil liegt nicht im Modell

Eine provokante, aber zunehmend realistische These lautet: Der Wettbewerbsvorteil liegt nicht im KI-Modell, sondern in den Daten. Denn die großen Modelle sind für viele Unternehmen gleichermaßen verfügbar. Fast jeder kann heute leistungsfähige Sprachmodelle, Agentenframeworks oder Automatisierungstools einsetzen. Der technologische Zugang demokratisiert sich.

Was sich nicht demokratisiert, ist die Qualität der eigenen Datenbasis. Kaum ein Unternehmen verfügt automatisch über kuratierte Unternehmensdaten, adressierbare Zielgruppen, belastbare Identitäten, verifizierte B2B-Informationen und aktuelle Marktmerkmale. Genau diese Daten entscheiden jedoch darüber, ob KI nur generische Antworten produziert oder konkrete wirtschaftliche Wirkung entfaltet.

Ein Unternehmen, das seine Daten versteht, bereinigt, strukturiert, anreichert und kontrolliert, kann KI wesentlich gezielter einsetzen. Es weiß, wen es anspricht, welche Profile relevant sind, welche Signale belastbar sind und welche Daten für welche Zwecke verwendet werden dürfen.

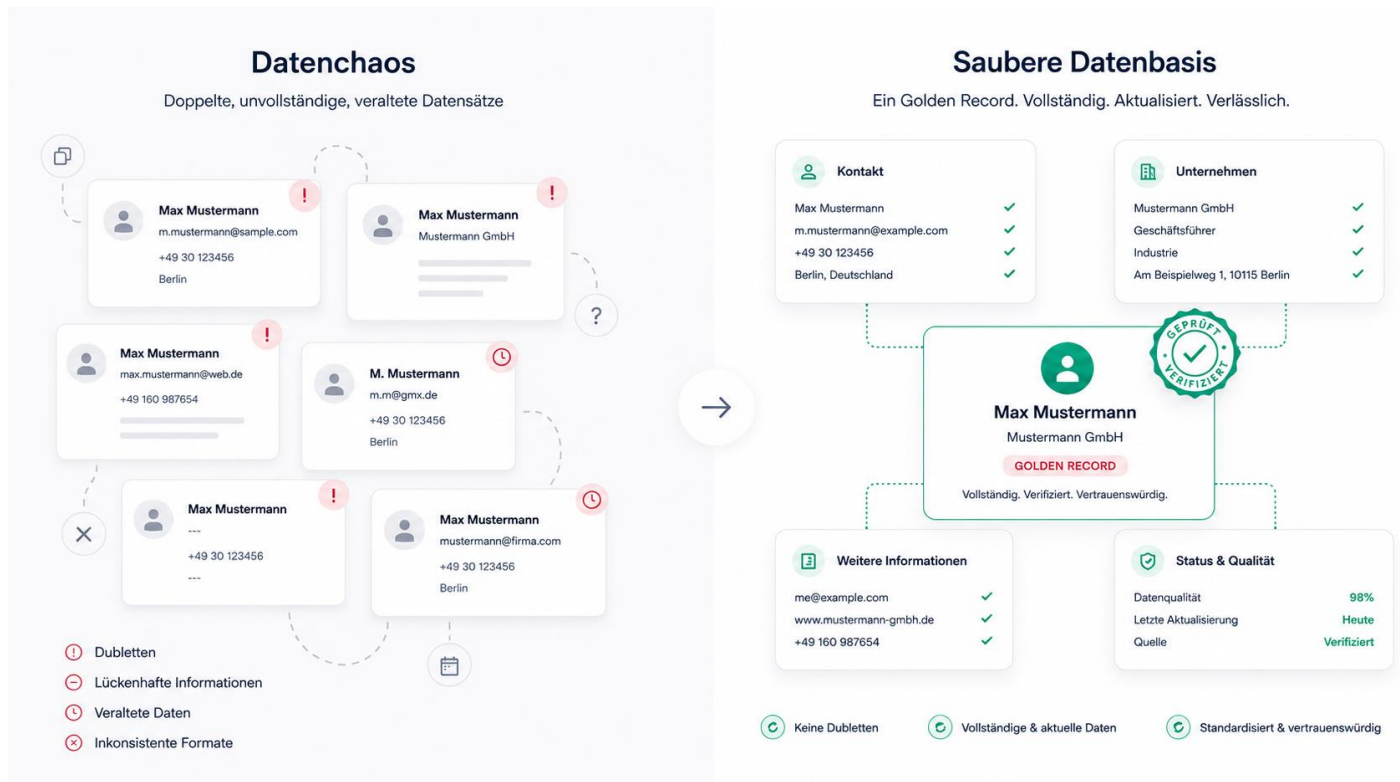


Abbildung: Von Datenchaos zu einer sauberen, verlässlichen Datenbasis.

3. Trusted Data for Agentic AI

Mit dem Aufkommen von KI-Agenten wird die Datenfrage noch wichtiger. Agentic AI verspricht Systeme, die nicht nur antworten, sondern Aufgaben ausführen: recherchieren, bewerten, priorisieren, segmentieren, Kampagnen vorbereiten, Kundenpotenziale erkennen oder nächste Schritte empfehlen.

Doch Agenten können nur so gut handeln wie die Daten, auf denen sie arbeiten. Viele entscheidende Fragen sind keine reinen LLM-Probleme, sondern Datenprobleme:

- Ist dieses Unternehmen noch aktiv?
- Welche Branche ist tatsächlich relevant?
- Wer ist der richtige Ansprechpartner?
- Welche Niederlassung gehört zu welcher Unternehmensstruktur?
- Welche Zielgruppe ist adressierbar?
- Welche Daten dürfen für welchen Zweck verwendet werden?
- Welche Merkmale sind aktuell, geprüft und belastbar?
- Welche Kaufwahrscheinlichkeit lässt sich ableiten?

Ein Sprachmodell kann solche Fragen formulieren, interpretieren oder in Workflows einbetten. Die Antwortqualität hängt jedoch von Daten ab, nicht von Sprachfähigkeit allein. Für wirtschaftlich nutzbare KI braucht es deshalb Trusted Data: geprüfte, aktualisierte, konsolidierte und compliance-sichere Daten, die in operativen Prozessen verwendet werden können.

4. Identitätsauflösung als unterschätzter Erfolgsfaktor

Ein besonders kritischer Baustein ist die Identitätsauflösung. In vielen Systemen existieren Unternehmen, Personen oder Kunden mehrfach: mit unterschiedlichen Schreibweisen, alten Adressen, verschiedenen E-Mail-Domains, mehreren Telefonnummern oder abweichenden Firmierungen.

Für klassische Prozesse ist das bereits problematisch. Für KI wird es zum massiven Qualitätsrisiko. Wenn ein KI-System nicht erkennt, dass mehrere Datensätze zur gleichen Einheit gehören, entstehen falsche Zusammenhänge. Potenziale werden doppelt gezählt, Kampagnen falsch priorisiert, Reports verzerrt und Kundenprofile unvollständig interpretiert.

Identitätsauflösung schafft hier die Grundlage für einen Golden Record: ein konsistentes, eindeutiges und nutzbares Profil, das verschiedene Informationen zusammenführt und Widersprüche reduziert. Erst dadurch können KI-Systeme auf einer stabilen Datenbasis arbeiten.

5. Datenanreicherung macht KI kontextfähig

Neben der Bereinigung und Zusammenführung spielt auch Datenanreicherung eine zentrale Rolle. Viele interne Datenbestände enthalten nur einen Ausschnitt der Realität: Name, Adresse, E-Mail, Kaufhistorie oder CRM-Status. Für intelligente Segmentierung, Scoring oder automatisierte Empfehlungen reicht das häufig nicht aus.

KI benötigt Kontext. Dazu gehören zum Beispiel Brancheninformationen, Unternehmensgrößen, Standorte, Entscheidungsstrukturen, regionale Merkmale, Zielgruppenattribute, Marktpotenziale oder Signale zur Relevanz eines Unternehmens. Erst solche Zusatzinformationen machen aus einem Datensatz ein verwertbares Profil.

Ohne Anreicherung bleibt KI oft oberflächlich. Mit angereicherten, geprüften Daten kann sie deutlich präzisere Empfehlungen geben, Zielgruppen besser unterscheiden und operative Maßnahmen sinnvoller vorbereiten.

6. Compliance ist kein Zusatzthema, sondern Voraussetzung

Je mehr KI in Marketing, Vertrieb und Kundenprozesse integriert wird, desto wichtiger wird Compliance. Daten dürfen nicht beliebig verwendet werden, nur weil sie technisch verfügbar sind. Unternehmen müssen nachvollziehen können, woher Daten stammen, wofür sie genutzt werden dürfen, wie aktuell sie sind und welche rechtlichen Anforderungen gelten.

Gerade bei KI-Anwendungen ist Transparenz entscheidend. Wenn Daten in Modelle, Agenten oder Automatisierungsprozesse einfließen, müssen Unternehmen sicherstellen, dass Datenschutzvorgaben, Zweckbindungen und interne Richtlinien eingehalten werden.

Compliance ist daher kein Bremsklotz für KI. Sie ist die Voraussetzung dafür, dass KI dauerhaft produktiv, vertrauenswürdig und skalierbar eingesetzt werden kann.

7. Framework: Die sechs Grundlagen KI-fähiger Daten

Viele Unternehmen betrachten Datenqualität noch immer als operatives Aufräumthema: Dubletten entfernen, Adressen korrigieren, Telefonnummern prüfen, Felder vereinheitlichen. Das bleibt wichtig. Im KI-Zeitalter bekommt Datenqualität jedoch eine strategische Dimension.

Diese sechs Grundlagen bilden die Brücke zwischen Technologie und wirtschaftlicher Wirkung. Sie sorgen dafür, dass KI nicht auf isolierten Dateninseln arbeitet, sondern auf einer konsistenten, kontextfähigen und aktivierbaren Informationsbasis.



8. Checkliste: Ist Ihre Datenbasis bereit für KI?

Die folgende Checkliste hilft bei einer ersten Einordnung. Sie ersetzt kein Daten-Audit, macht aber sichtbar, ob ein KI-Projekt auf einer belastbaren Grundlage starten kann oder ob zunächst Datenqualität, Identitätsauflösung, Anreicherung und Compliance geklärt werden sollten.

Prüffeld	Leitfrage	Status		
1 Datenqualität	Sind Dubletten, Fehlformate und unvollständige Felder bekannt und quantifiziert?	☐ offen	☐ teilweise	☐ erfüllt
2 Identität	Gibt es eindeutige Kunden-, Unternehmens- und Kontaktprofile über Systeme hinweg?	☐ offen	☐ teilweise	☐ erfüllt
3 Aktualität	Werden Adressen, Firmenstatus, Ansprechpartner und Kontaktdaten regelmäßig aktualisiert?	☐ offen	☐ teilweise	☐ erfüllt
4 Anreicherung	Sind relevante Zusatzmerkmale für Segmentierung, Scoring und Priorisierung vorhanden?	☐ offen	☐ teilweise	☐ erfüllt
5 Nutzungskontext	Ist klar definiert, welche Daten für welchen KI-Use-Case benötigt werden?	☐ offen	☐ teilweise	☐ erfüllt
6 Compliance	Sind Herkunft, Zweckbindung, Datenschutz und Zugriffskonzepte dokumentiert?	☐ offen	☐ teilweise	☐ erfüllt
7 Aktivierung	Können die Daten in CRM, Marketing Automation, Vertrieb oder Agentenprozessen genutzt werden?	☐ offen	☐ teilweise	☐ erfüllt
8 Monitoring	Gibt es einen laufenden Prozess zur Qualitätskontrolle und Aktualisierung?	☐ offen	☐ teilweise	☐ erfüllt



9. Handlungsempfehlungen für den Einstieg

Unternehmen, die KI erfolgreich einsetzen wollen, sollten nicht nur in Tools und Modelle investieren. Sie sollten zuerst klären, ob ihre Daten aktuell, vollständig, eindeutig, anreicherbar, aktivierbar und rechtssicher nutzbar sind. Für den Einstieg empfiehlt sich ein pragmatisches Vorgehen:

1. Use Cases priorisieren: Welche Prozesse, Entscheidungen oder Kampagnen sollen durch KI konkret verbessert werden?
2. Datenquellen erfassen: Welche internen und externen Daten werden dafür benötigt?
3. Datenqualität prüfen: Wo bestehen Dubletten, Lücken, veraltete Informationen oder Formatprobleme?
4. Identitäten auflösen: Welche Kunden-, Unternehmens- und Kontaktprofile müssen zusammengeführt werden?
5. Daten anreichern: Welche Zusatzinformationen erhöhen Relevanz, Kontext und Entscheidungsqualität?
6. Compliance absichern: Welche Daten dürfen für welchen Zweck verwendet werden?
7. Dauerhafte Qualität etablieren: Wie werden Aktualität, Konsistenz und Verwendbarkeit laufend überwacht?

Praxisprinzip

Erfolgreiche KI-Projekte beginnen nicht mit maximaler Komplexität, sondern mit einer klar abgegrenzten Datenbasis, einem konkreten Use Case und messbaren Qualitätskriterien.

10. Fazit: Die eigentliche Frage lautet nicht, welche KI wir nutzen

Die wichtigere Frage lautet: Auf welcher Datenbasis soll unsere KI arbeiten? Wer diese Frage nicht beantwortet, riskiert teure Pilotprojekte ohne nachhaltigen Effekt. Wer sie beantwortet, schafft die Voraussetzung für KI, die nicht nur beeindruckt, sondern wirtschaftlich wirkt.

KI braucht Daten. Sie braucht Identitäten. Sie braucht Zielgruppen. Sie braucht Anreicherung. Sie braucht Qualitätskontrolle. Und sie braucht Compliance. Ohne diese Grundlage bleibt KI beliebig.

Genau hier wird spezialisierte Datenkompetenz entscheidend. Unternehmen benötigen Partner, die Daten nicht nur technisch verarbeiten, sondern fachlich verstehen: Datenqualität, Identitätsauflösung, Datenfusion, Anreicherung, Zielgruppenlogik, Zusatzinformationen, Datenschutz und Aktivierung.

Denn wirtschaftlich nutzbare KI entsteht nicht allein durch den Einsatz eines Modells. Sie braucht eine Datenbasis, die geprüft, aktuell, eindeutig, angereichert und rechtssicher verwendbar ist. Erst dann können KI-Systeme in Marketing, Vertrieb und Analyseprozessen belastbare Ergebnisse liefern.

Der entscheidende Punkt ist: KI braucht Treibstoff. Dieser Treibstoff sind keine beliebigen Daten, sondern geprüfte, aktuelle, angereicherte und vertrauenswürdige Informationen.



Über die Autorin

Sylvia Türcke

Geschäftsführerin Capaneo GmbH

Sylvia Türcke beschäftigt sich mit Datenqualität, Zielgruppenlogik und der Frage, wie Daten für Marketing, Vertrieb, Analyse und KI nutzbar werden. Bei Capaneo arbeitet sie unter anderem an Lösungen, die verstreute Daten bereinigen, anreichern und in eine compliance-sichere Grundlage überführen. Dazu gehört auch das eigene Datenuniversum mit B2B- und B2C-Daten - als Basis für bessere Zielgruppen, belastbare Analysen und produktiv nutzbare KI.

Kontakt: sylvia.tuercke@capaneo.de

www.capaneo.de



Hinweis: Dieses Whitepaper dient der fachlichen Einordnung und ersetzt keine rechtliche Beratung. Datenschutz- und Compliance-Anforderungen sollten im konkreten Einzelfall geprüft werden.