

Modul 3: Datenbasis aufbauen - Qualität vor Quantität

Kostenlose Tools für KI-ready Daten

Agenda & Community-Check

- 17:00-17:05** **Begrüßung & Community-Check**
- 17:05-17:15** **Daten-Realitätsschock**
Warum ChatGPT bei schlechten Daten versagt
- 17:15-17:27** **Datenqualitäts-Check**
Die 5 kritischen Datentypen für KMU
- 17:27-17:42** **Kostenlose Aufbereitungs-Tools**
Tool-Showcase für jeden Datentyp
- 17:42-17:50** **KI-Ready machen**
Von Chaos zu KI-Input
- 17:50-18:00** **Q&A & Ausblick**
Fragen und Vorschau auf Modul 4

Erkenntnisse aus Modul 1 & 2

- ✓ Technische Optimierung: Größere Schrift, kontrastreichere Folien für bessere Lesbarkeit
- ✓ Engagement-Strategie: Direktere Ansprache der 30-40 Teilnehmer
- ✓ Praxis vor Theorie: Sofort mit Beispielen starten
- ✓ Community-Wertschätzung: Bezug auf vorherige Diskussionen
- ✓ Zeitmanagement: 45 Min Kerninhalt, keine Überlastung

Daten-Realitätsschock

Warum ChatGPT bei schlechten Daten versagt

⚠ Praxis-Beispiel: Kundendaten-Chaos

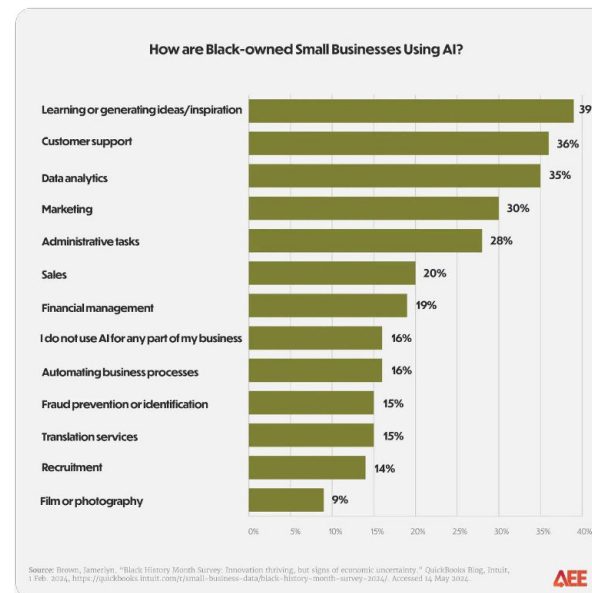
Unstrukturierte Kundendaten führen zu falschen KI-Ausgaben und können erheblichen Imageschaden verursachen.

⚖ Die 80/20-Regel der KI

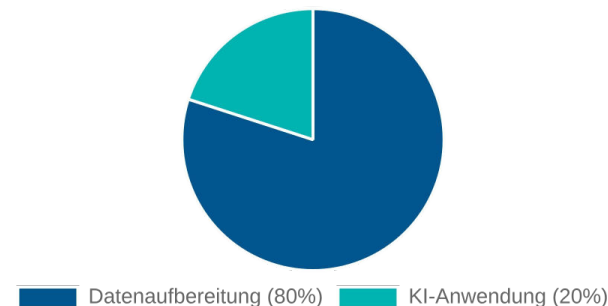
80% der Zeit wird für Datenaufbereitung benötigt, nur 20% für die eigentliche KI-Anwendung.

📊 Quick-Poll

"Wie strukturiert sind Ihre Unternehmensdaten auf einer Skala von 1-10?"



Zeitverteilung bei KI-Projekten



Datenqualitäts-Check

Die 5 kritischen Datentypen für KMU

Kundendaten

CRM-Systeme, Excel-Listen, Kontaktdatenbanken

Kommunikationsdaten

E-Mails, Chat-Verläufe, Kundenservice-Tickets

Dokumenten-Archive

Verträge, Handbücher, Compliance-Dokumente

Produktdaten

Kataloge, Inventarlisten, Preislisten

Finanzdaten

Buchhaltung, Rechnungen, Zahlungsverläufe



? "Welcher Datentyp ist bei Ihnen am chaotischsten?"

Kostenlose Aufbereitungs-Tools

3 KOSTENLOSE DATENBEREINIGUNGS-TOOLS FÜR KLEINE UNTERNEHMEN



OpenRefine

- ✓ Speziell für "messy data" entwickelt
- ✓ Erkennt Muster und Inkonsistenzen
- ✓ Umfangreiche Filterfunktionen

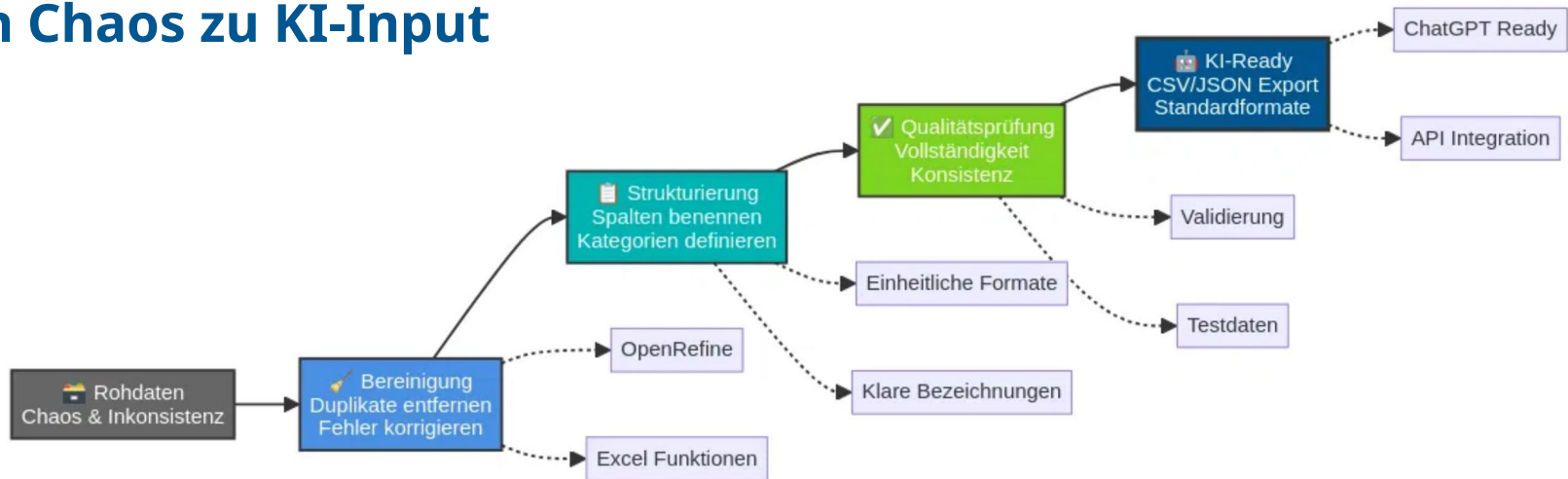
Excel/Google Sheets

- ✓ Duplikate finden & entfernen
- ✓ Datenvalidierung einrichten
- ✓ Bedingte Formatierung nutzen

LibreOffice Base

- ✓ Kostenlose Datenbank-Alternative
- ✓ Relationale Datenstrukturen
- ✓ SQL-Abfragen möglich

Von Chaos zu KI-Input



Rohdaten

Unstrukturierte Daten aus verschiedenen Quellen sammeln und identifizieren. Typische Probleme: Inkonsistenzen, Duplikate, fehlende Werte.

Bereinigung

Duplikate entfernen, Fehler korrigieren, fehlende Werte ergänzen. Tools wie OpenRefine oder Excel-Funktionen nutzen.

Strukturierung

Spalten einheitlich benennen, Kategorien definieren, Datentypen festlegen. Konsistente Formate für Datum, Zahlen und Text.

KI-Ready

Export in CSV oder JSON Format. Validierung der Datenqualität. Vorbereitung für ChatGPT oder andere KI-Tools.

Datenqualitätsstufen

Der Weg zu KI-ready Daten



KI-Ready

Strukturierte, validierte Daten in standardisierten Formaten (CSV, JSON), die direkt von KI-Systemen verarbeitet werden können. Vollständig, konsistent und aktuell.



Strukturiert

Daten mit klaren Spaltenbezeichnungen, definierten Kategorien und einheitlichen Formaten. Beziehungen zwischen Datensätzen sind klar definiert.



Bereinigt

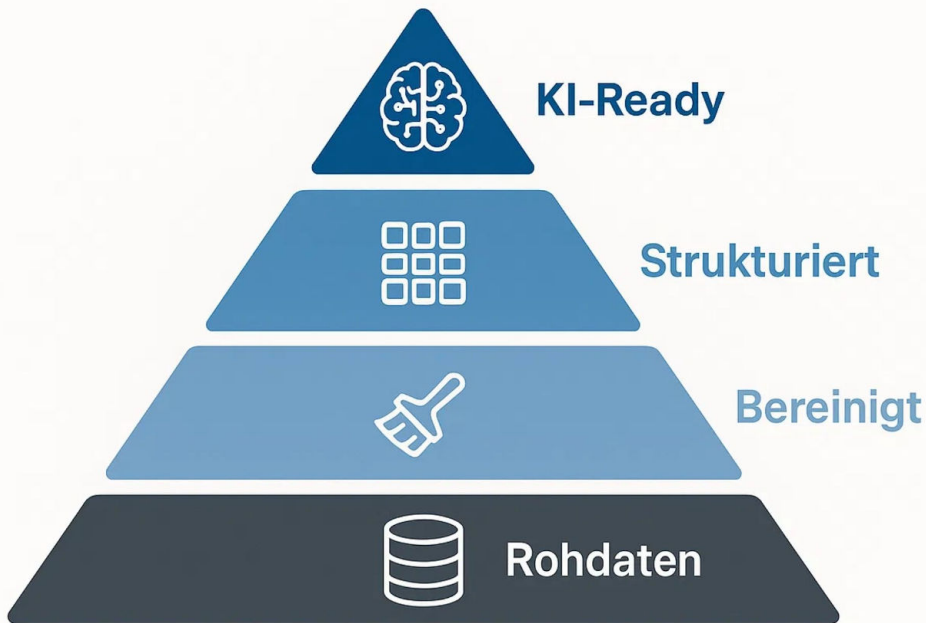
Daten ohne Duplikate und offensichtliche Fehler. Fehlende Werte wurden identifiziert und behandelt. Grundlegende Konsistenz ist hergestellt.



Rohdaten

Unbearbeitete Daten mit Inkonsistenzen, Duplikaten, Fehlern und fehlenden Werten. Oft in verschiedenen Formaten und ohne klare Struktur.

Datenqualitätsstufen



Rechtssichere Umsetzung

Datenschutz und Compliance

AI Act Compliance

- ✓ Transparenz: Dokumentieren Sie, welche Daten wie verarbeitet werden
- ✓ Datenverarbeitungsverzeichnis aktualisieren
- ✓ Risikobewertung für Hochrisiko-Anwendungen durchführen

DSGVO-konforme Tools

- ✓ Bevorzugen Sie Tools mit europäischen Servern
- ✓ Personenbezogene Daten vor Bearbeitung anonymisieren
- ✓ Auftragsverarbeitungsverträge (AVV) auch für kostenlose Tools prüfen
- ✓ Führen Sie eine lückenlose Dokumentation



Branchenspezifische Beispiele

Handwerk (Elektro, SHK)

- ✓ Kundendatenbank mit Wartungsintervallen strukturieren
- ✓ Materiallisten für bessere Lagerhaltung optimieren
- ✓ Servicedokumentation systematisieren
- ✓ Angebotskalkulation durch strukturierte Preisdaten verbessern
- ✓ Einsatzplanung durch historische Auftragsdaten optimieren

Finanzen/Versicherung

- ✓ Kundensegmentierung für gezielte Angebote verfeinern
- ✓ Risikodaten für KI-Modelle aufbereiten
- ✓ Compliance-Dokumentation strukturieren
- ✓ Beratungsprotokolle für bessere Nachvollziehbarkeit standardisieren
- ✓ Schadensmeldungen für schnellere Bearbeitung kategorisieren

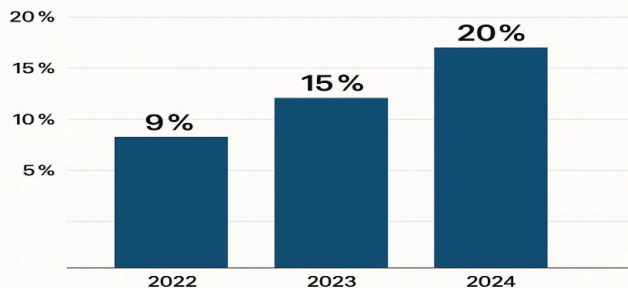
Dienstleistung/Beratung

- ✓ Projektverlaufsdaten sammeln und analysieren
- ✓ Kundenfeedback systematisch erfassen und auswerten
- ✓ Abrechnungsdaten für Effizienzanalysen optimieren
- ✓ Wissensmanagement durch strukturierte Dokumentation verbessern
- ✓ Mitarbeiterkapazitäten durch historische Projektdaten planen

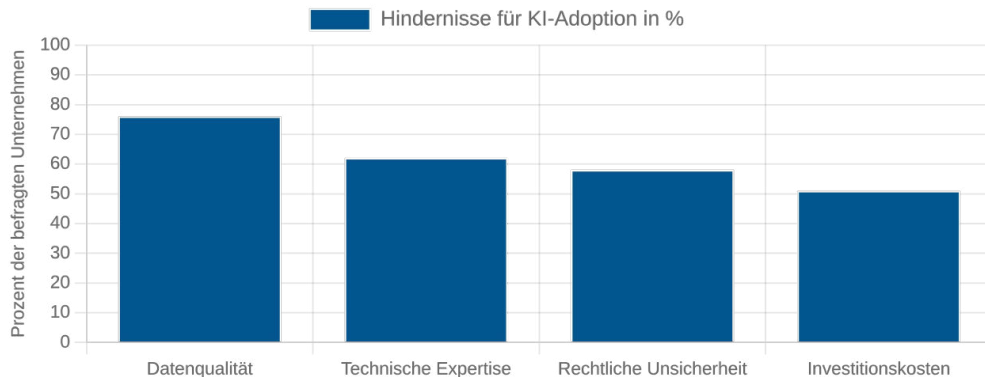
KI-Nutzung in deutschen KMU

KI-Nutzung in deutschen KMU

Steigerung um 8 Prozentpunkte jährlich



Haupthindernisse für KI-Adoption in KMU



Aktuelle Entwicklungen

Kontinuierliches Wachstum

Die KI-Nutzung in deutschen KMU steigt jährlich um etwa 8 Prozentpunkte. Von 9% (2022) auf 20% (2024).

Haupthindernis: Datenqualität

76% der Unternehmen geben an, dass schlechte Datenqualität und -verwaltung den KI-Fortschritt behindert.

Potenzial für KMU

Weitere 37% der Unternehmen planen den KI-Einsatz in naher Zukunft. Der Fokus verschiebt sich von Experimentieren zur produktiven Nutzung.

Quellen: Bitkom-Studie 2024, Destatis 2024, Institut für Mittelstandsforschung Bonn 2025

Interaktivitäts-Elemente

Community-Integration

Namen-Callbacks

Bezug auf Axel Kochs KI-Coach und Thomas Wittlebens Rechtsfragen aus den vorherigen Modulen. Persönliche Ansprache schafft Verbindung und Kontinuität.

Peer-Learning

"Wer hat Erfahrung mit Datenbereinigung?" - Teilnehmer teilen ihr Wissen und lernen voneinander. Aktivierung der kollektiven Intelligenz der Community.

Tool-Tipps aus Community

Teilnehmer teilen bewährte Lösungen und Erfahrungen mit Tools. Praktisches Wissen aus der täglichen Anwendung ist besonders wertvoll.

Engagement-Booster

17:05

Quick-Poll

17:10

Frage

17:15

Chat

17:20

Umfrage

17:25

Feedback

**5-Minuten-Regel: Alle 5 Min eine Frage/Poll**

Praxis-Challenge

"Welche 3 Datenfelder würden Ihr Business sofort verbessern?" - Konkrete Aufgabenstellung mit direktem Bezug zur Geschäftspraxis der Teilnehmer.

Screenshot-Sharing

Teilnehmer zeigen anonymisierte Daten-Beispiele aus ihrem Geschäftsalltag und erhalten direktes Feedback zur Optimierung.

Erfolgsmessung & Follow-up

Immediate Takeaways



Daten-Audit-Checkliste

PDF zum Download mit strukturierter Anleitung zur Bewertung Ihrer aktuellen Datenqualität und Identifikation von Verbesserungspotentialen.



Tool-Liste mit Installations-Anleitungen

Detaillierte Schritt-für-Schritt Anleitungen zur Installation und Nutzung der vorgestellten kostenlosen Tools für Datenbereinigung.



DSGVO-Schnellcheck

Praktische Checkliste zur Überprüfung der DSGVO-Konformität bei der Datenbearbeitung und -verarbeitung für KI-Anwendungen.

Community-Aufgabe bis Modul 4



Peer-Support

WhatsApp/Telegram-Gruppe für technische Hilfe und Austausch. Teilen Sie Ihre Erfahrungen und unterstützen Sie andere Community-Mitglieder bei der Datenbereinigung.



Success Stories

Wer hat bereits Verbesserungen erzielt? Teilen Sie Ihre Erfolgsgeschichten und lernen Sie von den Erfahrungen anderer Teilnehmer.

Challenge bis zum nächsten Modul

"Bereinigen Sie einen Datensatz und teilen Sie Ihre Erkenntnisse mit der Community"

Q&A & Ausblick

Fragen & Diskussion



Haben Sie Fragen zur Datenaufbereitung oder zu den vorgestellten Tools?

Jetzt ist die Zeit für offene Fragen und Diskussion. Teilen Sie auch gerne Ihre eigenen Erfahrungen mit der Community.



Ausblick Modul 4

Termin

Modul 4 findet am 07. Oktober 2025 statt

Thema

"KI-Anwendungen im Alltag - Praktische Implementierung"

Vorbereitung

Bringen Sie Ihre bereinigten Datensätze mit und wir werden gemeinsam erste KI-Anwendungen damit umsetzen

Kontakt & Ressourcen

 bds-akademie-bayern.de

 Community-Forum

 info@arno-schimmelpfennig.de