DRESDNER DATENBANKFORUM · HTW DRESDEN · 2026

BIKE24
BUILD YOUR RIDE

Das Modell ist nicht das Problem

Was KI vom DWH verlangt

Matthias Wendt · Head of BI, Data & Analytics · BIKE24 GmbH

BIKE24
BUILD YOUR RIDE

Wir reden erstaunlich viel über Modelle

GPT | Claude | Gemini | Llama | Cortex

Welches Modell sollen wir nehmen?

Vielleicht ist das die falsche Frage.

UNSERE JOURNEY

Fünf Jahre KI bei BIKE24

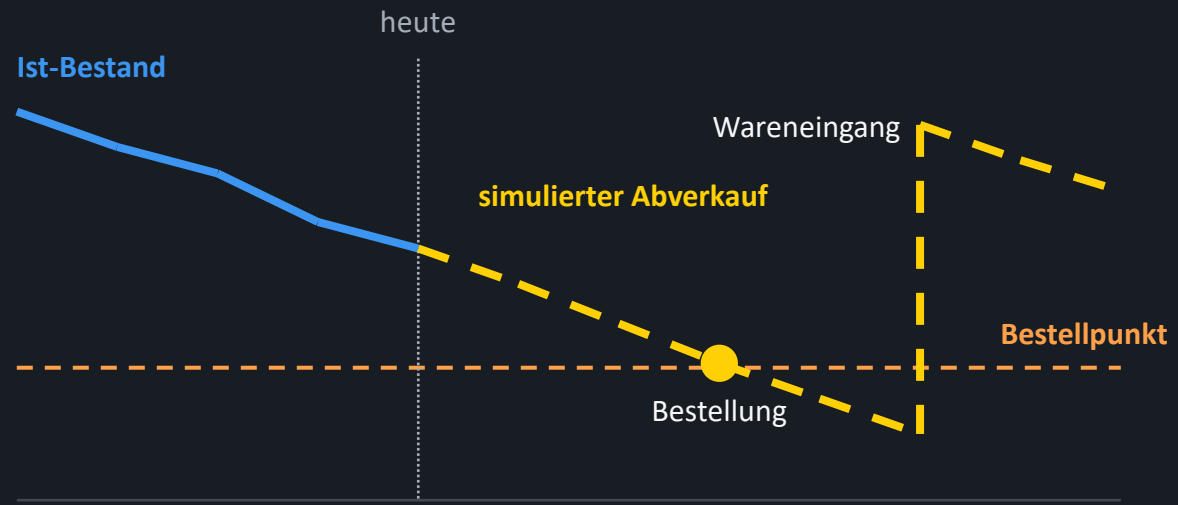
2021 Forecasting → **2023** Produktberater → **2025** Große Modelle → **2026** Agentic Analytics

Vier Stationen. Und jedes Mal war der Engpass woanders.

„Können wir das nicht vorhersagen?“

UNSER ERSTER USE CASE

SKU-basiertes Replenishment mit Abverkaufssimulation



illustrativ · eine SKU

× 230.000–250.000 SKUs

Die Fragen waren klassisch:

Welche Features?

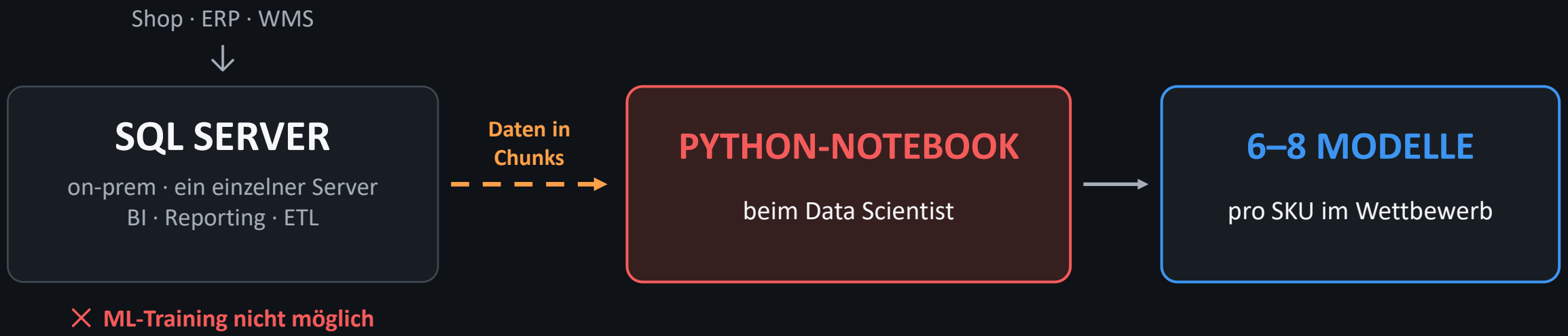
Welcher Algorithmus?

Welche Accuracy?

Die Lektion kam von einer anderen Seite.

BIKE24 BUILD YOUR RIDE

Unsere Architektur hatte eine Meinung



bis zu 1 h je SKU × **230–250 T** SKUs

Ein Jahr hat 8.760 Stunden.

Training ist auch nur eine Query. Unser DWH konnte sie nicht ausführen.

BIKE24
BUILD YOUR RIDE

LEARNING #1

Das Modell skalierte.
Unsere Plattform nicht.

WORKLOAD ISOLATION

ELASTIC COMPUTE

STORAGE / COMPUTE

BIKE24 BUILD YOUR RIDE

Was wir geändert haben

2021
Start der Snowflake-Einführung

Ende 2022 / Anfang 2023
Erste große Welle: klassische Reporting-Use-Cases migriert

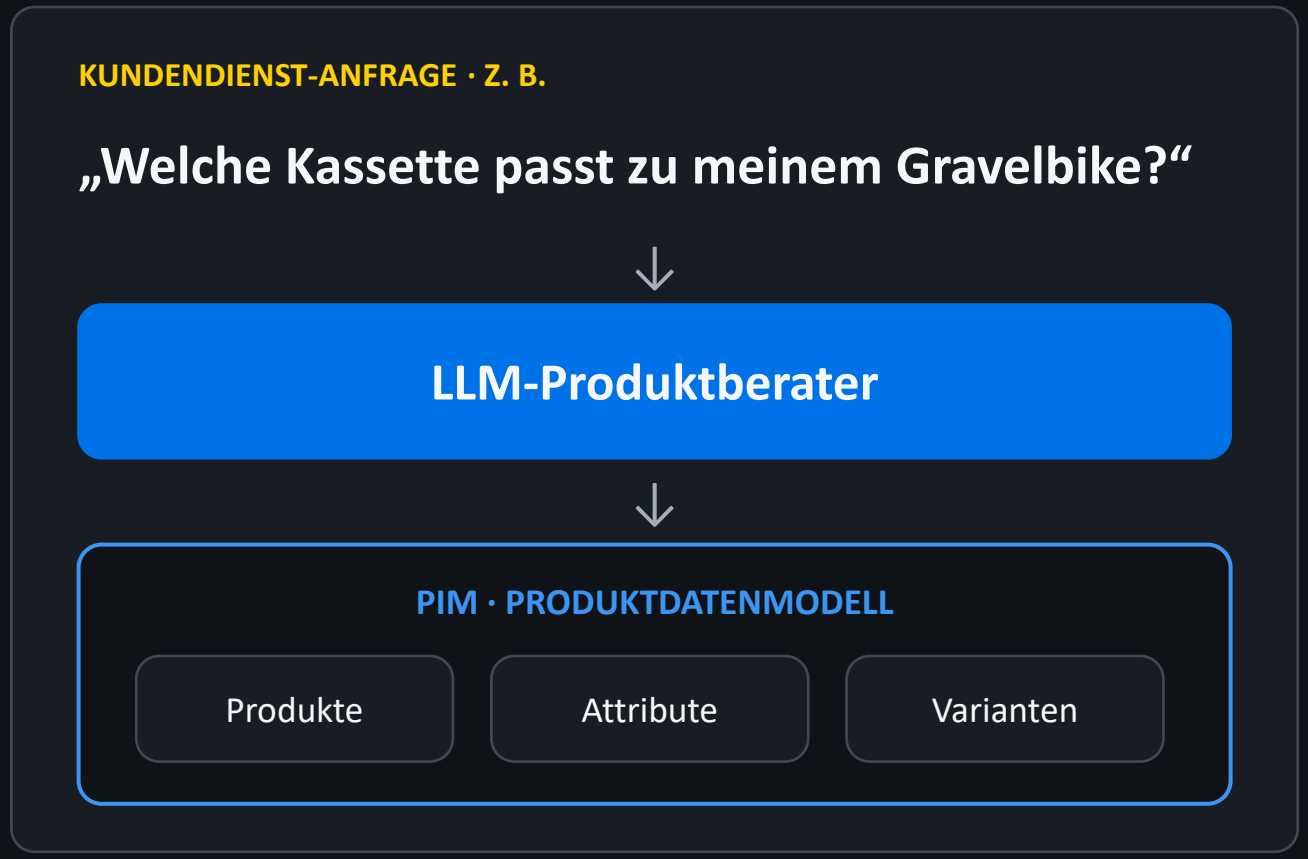
Danach
Forecasts problemlos rechenbar

	VORHER		NACHHER
Workload-Isolation	Ein SQL Server für BI, Reporting und ETL	→	**Eigene Rechenleistung pro Workload**
Elastic Compute	Feste Kapazität	→	**Nach Bedarf hoch- und runterskalieren**
Storage / Compute	Speicher und Rechenleistung gekoppelt	→	**Unabhängig voneinander skalierbar**

Anlass war ML. Migriert haben wir zuerst das Reporting.

BIKE24 BUILD YOUR RIDE

„Kann der nicht Kundenfragen beantworten?“



Wir dachten:

Der Kundendienst bekommt einen Produktexperten.

SPOILER **So einfach war es nicht.**

BIKE24 BUILD YOUR RIDE

Drei Gründe, warum es nicht funktionierte

MODELL 01

Zu kleines Kontextfenster

Das komplette Produktdatenmodell passte nicht hinein.

MODELL 02

Fehlendes Weltwissen

Dem Modell fehlte das Wissen, das ein Experte mitbringt.

DATEN 03

Kompatibilitäten fehlten

Das PIM wusste zu wenig darüber, welche Teile zusammenpassen.

Zwei davon lösen heutige Modelle. Den dritten löst kein Modell.

DAS EISBERGPROBLEM

BIKE24
BUILD YOUR RIDE

Das PIM kennt Produkte.
Der Experte weiß, was zusammenpasst.

WAS DER CHATBOT SIEHT

PRODUCT	ATTRIBUTE	VARIANT	CATEGORY	BRAND

WAS DER PRODUKTEXPERTE WEISS

Welche Teile kompatibel sind	Einbau- und Achsstandards	Herstellerangabe vs. Praxis	Was der Kunde eigentlich meint
Typische Fehlkäufe	Nachfolger- und Ersatzteile	Was zusammen sinnvoll ist	**„Frag mal jemanden aus der Werkstatt.“**

BIKE24
BUILD YOUR RIDE

LEARNING #2

Ein Schema beschreibt Daten.
Es erklärt nicht das Unternehmen.

DATA MODEL ≠ BUSINESS MODEL

BIKE24 BUILD YOUR RIDE

Dann wurden die Modelle richtig gut



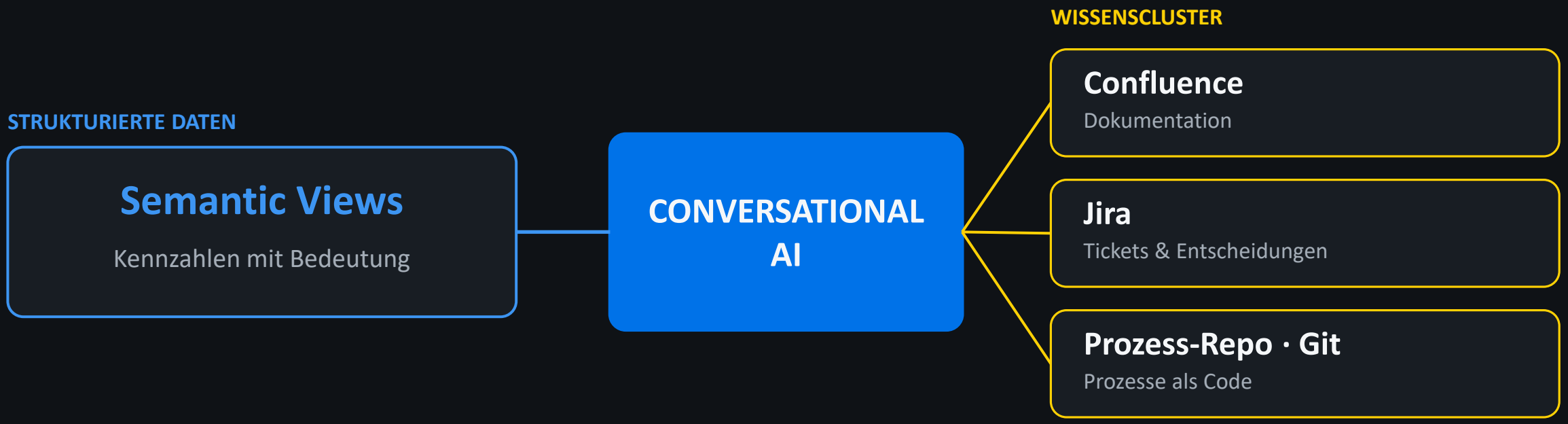
ZURÜCK ZUM PRODUKTBERATER VON 2023

✓	**Zu kleines Kontextfenster**	heute bis zu 1 Mio. Tokens
✓	**Fehlendes Weltwissen**	Modelle wissen heute deutlich mehr
✗	**Kompatibilitäten fehlten**	müssen erst in die Daten

Problem gelöst? **Nein. Problem verschoben: vom Modell zu den Daten.**

BIKE24
BUILD YOUR RIDE

„Gib dem Modell mehr Kontext.“



Welcher Kontext? Wann?

BIKE24
BUILD YOUR RIDE

Context Window ≠ Context

CONTEXT WINDOW

1.000.000

Tokens

USEFUL CONTEXT

Eine Handvoll

für genau diese Frage

Context Engineering wird zur Datenarchitekturaufgabe.

BIKE24 BUILD YOUR RIDE

Wenn der Konsument kein Mensch mehr ist



BIKE24 BUILD YOUR RIDE

Menschen sind erstaunlich gute Middleware

WAS MAN HEUTE HÖRT		WAS DER AGENT STATTDESSEN BRAUCHT
„Die Tabelle nehmen wir nicht mehr.“	→	**Status & Gültigkeit**
„Finance rechnet das anders.“	→	**Verbindliche Definitionen**
„Vor Weihnachten ist das verzerrt.“	→	**Annotierte Zeiteffekte**
„Das Feld ist seit Go-live kaputt.“	→	**Datenqualität**
„Bei Barcelona musst du aufpassen.“	→	**Dokumentierte Ausnahmen**
„Das weiß eigentlich nur eine Person im Marketing.“	→	**Ownership**

Ein Agent weiß davon nichts.

Implizites Wissen muss maschinenzugänglich werden.

BIKE24 BUILD YOUR RIDE

Dieselbe Lektion, neue Frage

„Wie hoch war unsere Marge in Deutschland letzten Monat?“

WAS EIN ANALYST ZURÜCKFRAGT		WAS DIE SEMANTIK BEANTWORTEN MUSS
„Welche Marge?“	→	**Definition: Umsatz netto – Wareneinsatz**
„Bestell- oder Rechnungsdatum?“	→	**Zeitbezug explizit festgelegt**
„Retouren, Stornos, Rabatte, Versand?“	→	**Abzüge benannt**
„Welcher Wechselkurs?“	→	**Währung & Kurs-Stichtag**
„Lieferland oder Shop?“	→	**Zuordnung eindeutig**
„Wen frage ich, wenn es nicht passt?“	→	**Owner & gültig seit**

Ohne Definition rät der Agent. Mit Definition nennt er sie mit.

BIKE24 BUILD YOUR RIDE

Read → Reason → Act

READ

Der Agent empfiehlt einen Preis.

Fehler = falsche Antwort

REASON →

ACT

Der Agent setzt den Preis. Live im Shop.

Fehler = falsche Aktion

BEVOR EIN AGENT HANDELN DARF, MUSS DIE PLATTFORM BEANTWORTEN:

Woher kommt die Zahl?	Darf er das?	Was hat er getan?	Wo ist die Grenze?
Herkunft	Berechtigung	Nachvollziehbarkeit	Guardrails

Data Quality reicht nicht mehr.

BIKE24
BUILD YOUR RIDE

Was KI vom DWH verlangt

		ANFORDERUNG	GELERNT AN
01	SCALE	Compute folgt dem Workload	2021 · Forecast für 230–250 T SKUs
02	EXPLAIN	Bedeutung wird maschinenlesbar	2023 · Produktberater heute · Margenfrage
03	CONTEXTUALIZE	Relevanter Kontext ist auffindbar	heute · Semantic Views und Wissenscluster
04	TRUST	Herkunft, Rechte und Grenzen sind explizit	Ausblick · Agenten, die handeln

AI-ready wird man nicht durch ein besseres Modell, sondern durch bessere Daten.

BIKE24
BUILD YOUR RIDE

Brauchen wir noch Data People?

WENIGER

Dashboards bauen

SQL schreiben

Ad-hoc-Fragen beantworten

→

MEHR

Bedeutung modellieren
Definitionen, Semantic Views

Kontext kuratieren
Wissen, geprüfte Beispielfragen

Vertrauen absichern
Herkunft, Rechte, Guardrails

Die Arbeit, die KI Heute braucht, haben wir nie richtig gemacht.

DAS MODELL IST NICHT DAS PROBLEM

Das DWH wird zur Wissensinfrastruktur.

Wir haben Daten für Menschen aufbereitet.
Jetzt bereiten wir Wissen für Maschinen auf.

Das wird unsere zukünftige Aufgabe.