

# RAG und Metadaten-Architekturen im B2B-E-Commerce

Aufbereitung technischer Produktdaten für strukturierte KI-Beratungsdialoge

## 1. Das Problem mit komplexen B2B-Produktdaten

Der digitale Vertrieb von technisch anspruchsvollen B2B-Produkten steht vor einer bekannten Hürde: Produktinformationen liegen meist gut strukturiert in Datenbanken, PIM- oder ERP-Systemen – Kunden formulieren ihre Anliegen aber zielorientiert und in Alltagssprache.

Klassische Online-Shops geraten dabei schnell an ihre Grenzen:

- **Unflexible Datenbankfilter:** Wer nach einem Produkt sucht, muss oft die exakten Parameter oder Teilenummern bereits kennen. Eine freie Anfrage wie *„Ich brauche eine Lösung für viskose Medien im Reinraumeinsatz“* lässt sich über standardmäßige Suchmasken kaum abbilden.
- **Unterschiedliche Ansprechpartner und Interessen:** Je nach Rolle im Unternehmen stehen ganz unterschiedliche Fragen im Vordergrund. Während Instandhalter und Ingenieure prüfen, ob Normen und technische Schnittstellen passen, schaut der Einkauf vor allem auf Preise, Garantien, Lieferzeiten und Betriebskosten (TCO).

In der Praxis müssen Kaufinteressenten die Lücke zwischen ihrer konkreten Aufgabenstellung und dem passenden Produkt selbst schließen. Wenn Informationen schwer zu finden sind oder die Suche zu kompliziert wird, springen Interessenten häufig ab.

## 2. RAG als Verbindung zwischen Kundenanfrage und Produktdaten

Um diese Lücke zu schließen, bieten sich Retrieval-Augmented Generation (RAG) Systeme an. Ein solches System verbindet unstrukturierte Dokumente (wie Handbücher oder Datenblätter) mit den strukturierten Stamm- und Bestandsdaten des Unternehmens.

Für den Einsatz im B2B-Umfeld sind dabei vor allem drei Punkte relevant:

- **Verlässliche Antworterzeugung (Grounding):** Die generierten Antworten werden strikt an die tatsächlich hinterlegten Produktinformationen gebunden. Dadurch kann das Risiko nicht belegbarer Aussagen („Halluzinationen“) reduziert werden.
- **Direkte Anbindung an Quellsysteme:** Bestände, technische Spezifikationen und tagesaktuelle Preise können über Schnittstellen direkt aus den ERP- oder PIM-Systemen abgefragt werden, ohne dass das KI-Modell neu trainiert werden muss.
- **Kombination zweier Suchverfahren (Hybride Suche):** Die Suche kombiniert zwei Ansätze: Die Vektorsuche versteht den Sinngehalt einer freien Textanfrage, während Metadaten genutzt werden, um konkrete technische und kaufmännische Vorgaben exakt einzuschränken.

## 3. Welche Rolle Metadaten bei der Produktsuche spielen

Damit dieses Verfahren zuverlässig funktioniert, reicht es allerdings nicht aus, die vorhandenen Dokumente einfach in eine Vektordatenbank zu übertragen. Gerade bei technischen Produkten müssen bestimmte Eigenschaften exakt eingehalten werden. Deshalb werden die Produktinformationen zusätzlich über Metadaten strukturiert.

In der Praxis werden die Daten dafür zweigeteilt: Längere Beschreibungen fließen als Freitext-Textabschnitte (Chunks) in die Vektorsuche ein, während klare Werte als harte Filterkriterien dienen:

| Metadaten-Kategorie           | Beispiel-Attribute                                                       | Funktion im Beratungsdialog                                                       |
|-------------------------------|--------------------------------------------------------------------------|-----------------------------------------------------------------------------------|
| <b>Zustand &amp; Logistik</b> | zustand (Neu/Generalüberholt), baujahr, betriebsstunden, lieferzeit_tage | Filtert Produkte nach Verfügbarkeit, Budgetrahmen und Dringlichkeit.              |
| <b>Technische Standards</b>   | durchsatz_stk_h, normen_cert (CE, FDA), anschlusswert_kw                 | Beantwortet technische Detailfragen und prüft Schnittstellenkompatibilität.       |
| <b>Kaufmännische Daten</b>    | preis_netto, leasing_rate, staffelpreise                                 | Ermöglicht zielgerichtete Preis- und Wirtschaftlichkeitsbetrachtungen.            |
| <b>Zielsegment</b>            | zielrolle (Technik, Einkauf), industrie (Pharma, Food)                   | Steuert, wie detailliert und mit welchem Schwerpunkt die Antwort formuliert wird. |

## Die Datenstruktur an einem konkreten Beispiel (JSON)

Wie eine solche Datenstruktur im Hintergrund aussieht, zeigt das folgende Beispiel eines generalüberholten (refurbished) Industrieguts. Das Schema verdeutlicht, wie sowohl Fließtext für das allgemeine Verständnis als auch präzise Metadaten für harte Filterregeln zusammengeführt werden:

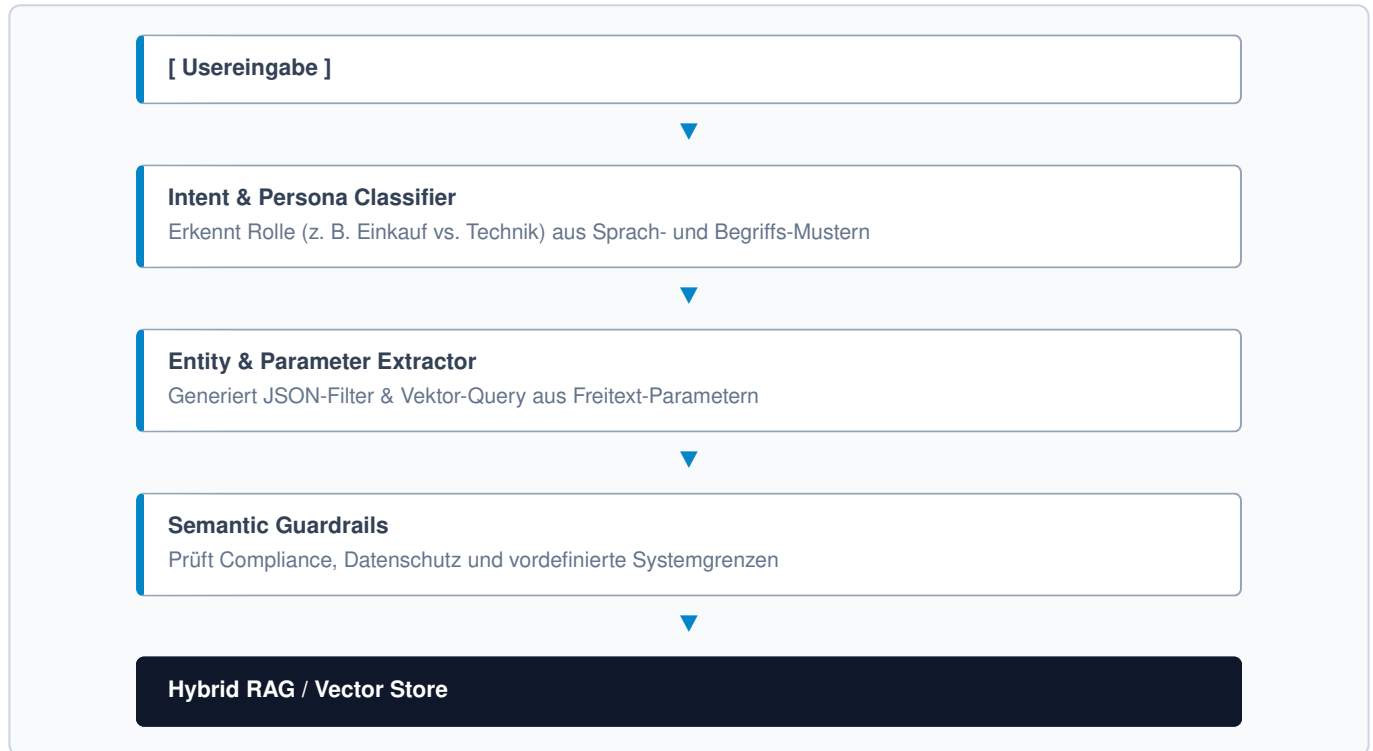
```
{
  "product_id": "FILL-MASTER-V3",
  "name": "Rotations-Flüssigkeitsbefüller (Generalüberholt)",
  "chunk_text": "Generalüberholte 8-Kopf-Befüllanlage für viskose Medien. Neue Dichtungen verbaut, Steuerung auf moderne Standard-SPS aufgerüstet. Einsatzbereit für CIP-Reinigung. Geeignet für Pharma- und Kosmetik-Medium-Batch-Produktion.",
  "metadata": {
    "zustand": "Gebraucht / Generalüberholt",
    "baujahr": 2019,
    "betriebsstunden": 4200,
    "durchsatz_stk_h": 3600,
    "viskositaet_max_pas": 15,
    "cert_fda": true,
    "cert_cleanroom": "ISO Class 7",
    "steuerung": "Standard-SPS (Industriestandard)",
    "preis_netto_eur": 45000.00,
    "neupreis_aequivalent_eur": 110000.00,
    "garantie_monate": 12,
    "lieferzeit_tage": 5,
    "kompatibel_mit_ids": ["CONVEYOR-BELT-2M", "CAPPER-SEAL-X"],
    "dokumente_verfuegbar": ["IQ_OQ_Protokoll.pdf", "CE_Erklaerung.pdf"]
  }
}
```

Diese Kombination stellt sicher, dass das System später beispielsweise genau weiß:

- Welche harten Grenzwerte eingehalten werden müssen ( viskositaet\_max\_pas: 15 ).
- Welche Dokumente der Kunde direkt als Anhang erhalten kann ( dokumente\_verfuegbar ).
- Welche kaufmännischen Argumente gegenüber dem Neupreis sprechen.

## 4. Verarbeitung der Kundenanfrage

Bevor das System in der Vektordatenbank nach passenden Inhalten sucht, wird die Anfrage des Nutzers in einem vorgeschalteten Prozess zerlegt und aufbereitet. Das stellt sicher, dass die Suchanfrage präzise formuliert wird und das System innerhalb der vorgegebenen Grenzen bleibt.



- **Intent & Persona Classifier:** Erkennt aus der Art der Fragestellung, wer gerade sucht. Fachbegriffe wie „ISO Class 7“ deuten auf einen technischen Ansprechpartner hin. Wörter wie „Lieferzeit“, „Garantie“ oder „TCO“ sprechen eher für jemanden aus dem Einkauf.
- **Entity & Parameter Extractor:** Übersetzt unstrukturierte Alltagssprache in harte Suchfilter. Aus der Aussage „Maximal 50.000 €, und ich benötige das Gerät innerhalb von zwei Wochen“ erzeugt die Komponente automatisch den Datenbankfilter: `{"preis_netto_eur": {"$lte": 50000}, "lieferzeit_tage": {"$lte": 14}}`.
- **Semantic Guardrails:** Diese Schutzmechanismen prüfen, ob die Anfrage zulässig ist, und reduzieren das Risiko, dass das System unbelegte Aussagen erzeugt oder vertrauliche Informationen preisgibt.

## 5. Beispiel: Von der Kundenfrage zur Produktempfehlung

Wie das Zusammenspiel aus Metadaten, Freitextsuche und Rollenerkennung in der Praxis aussieht, zeigen die folgenden zwei Szenarien für die oben gezeigte Anlage ( FILL-MASTER-V3 ).

### Szenario A: Technische Validierung (Ansprechpartner: Technik / Instandhaltung)

**Kunde:** „Wir befüllen hochviskose Creme (ca. 12 Pa·s) in ISO-7-Reinräumen. Passt die Fill-Master V3 bezüglich Viskosität und Steuerung?“

**System (RAG):** Die **Fill-Master V3** erfüllt diese Anforderungen:

- **Viskosität & Reinraum:** Die Anlage ist für Viskositäten bis 15 Pa·s ausgelegt und für ISO Class 7 zertifiziert (FDA-Konformität liegt vor).
- **Steuerung:** Ausgestattet mit einer aktuellen Standard-SPS (Industriestandard).
- **Dokumentation:** Das IQ/OQ-Qualifizierungsprotokoll sowie der Schaltplan stehen direkt als PDF zum Download bereit.

### Szenario B: Kaufmännische Bewertung (Ansprechpartner: Einkauf)

**Kunde:** „Gebrauchtgeräte bergen ein Ausfallrisiko. Welche kaufmännischen Argumente sprechen gegen einen Neukauf?“

**System (RAG):** Für den Einkauf sind vor allem drei Punkte relevant:

- **Gewährleistung:** Das generalüberholte Gerät wird mit 12 Monaten Garantie geliefert.
- **Kostenvergleich:** Bei 4.200 Betriebsstunden liegt der Preis bei 45.000 € netto – gegenüber 110.000 € Neupreis (eine Ersparnis von 59 %).
- **Verfügbarkeit:** Die Anlage ist innerhalb von 5 Werktagen lieferbar, während marktübliche Neuanlagen oft mehrere Monate Lieferzeit haben.

## 6. Fazit

Die Beispiele zeigen, dass RAG-Systeme im B2B-E-Commerce besonders dann stark werden, wenn Produktdaten aus unterschiedlichen Quellen zusammengeführt werden. Erst die Verknüpfung von unstrukturierten Texten mit strukturierten Metadaten macht es möglich, komplexe Anfragen präzise zu beantworten – egal ob die Frage aus der Technik oder aus dem Einkauf kommt. So wird aus einer starren Produktsuche ein direkter digitaler Beratungsprozess.