

Feedback-Copilot: VW-Präsentation

Präsentation für VW-Meeting | Januar 2026

Slide 1: Problem & Lösung

Das Problem

Wie analysiert man tausende In-Car-Kundenfeedbacks schnell, verlässlich und datenschutzkonform?

Herausforderungen bei VW:

-  Große Datenmengen aus Sprachassistent, Touch-Events, Fehlermeldungen
-  Personenbezogene Daten (E-Mails, Telefon, VINs, Kennzeichen)
-  Manuelle Analyse ist zeitaufwändig und inkonsistent

Die Lösung: Feedback-Copilot

Ein **RAG-basiertes Analyse-Tool**, das:

1. Feedbacks automatisch anonymisiert (PII-Filter)
 2. Fragen per Chat beantwortet mit Quellenangabe
 3. Probleme identifiziert und Trends visualisiert
 4. Tickets für Entwicklerteams generiert
-

Slide 2: Architektur & Wissenschaftliche Basis

RAG-Pipeline (Literatur-basiert)

Frage → Hybrid Retrieval → LLM → Antwort mit Quellen
(BM25 + Vector) (GPT-4)

Komponente	Literatur	Implementierung
Hybrid Search	Praneeth et al. 2025	BM25 + Vector + RRF
Guardrails	Wu & Wu 2025	Zitationspflicht
PII-Filter	Forschungslücke	Regex + Hashing

Zahlen

- **83 wissenschaftliche Papers** analysiert
- **Forschungslücke identifiziert:** PII-Anonymisierung in RAG-Pipelines
- Kein existierendes Paper behandelt automotive-spezifische PII (VIN, Kennzeichen)

Slide 3: Live-Demo (Features)

1. Dashboard

- KPIs auf einen Blick: Gesamt, Sprache, Touch, Fehler
- Verteilung nach Fahrzeugmodell und Markt

2. Analytics

- Trend-Charts nach Datum
- Filter nach Modell (ID.4, ID.5, Golf 8, Passat, Tiguan)
- Verteilung nach Quell-Typ und Sprache

3. RAG-Chat

- **Frage:** "Welche Probleme gibt es mit dem Sprachassistenten?"
- **Antwort:** Mit Quellenangabe zu konkreten Feedbacks
- **Unanswerable-Guardrail:** Sagt "keine Daten vorhanden" statt zu halluzinieren

4. Ingest & Export

- CSV/JSON-Upload mit PII-Preview
- Ticket-Export als JSON/CSV/Markdown

Slide 4: Status & Ausblick

 Fertig implementiert

Feature	Status
Dashboard mit KPIs	
Analytics Charts (5 Typen)	
RAG-Chat mit Quellenangabe	
Hybrid Retrieval (BM25+Vector)	
PII-Anonymisierung	
Ingest (CSV/JSON)	
Export (JSON/CSV/Markdown)	
Persistenter VectorStore	

 Nächste 3 Wochen (bis finale Präsentation)

1. **Evaluation:** Recall@k, nDCG, Latenz-Messung
2. **NER für Personennamen:** spaCy-Integration
3. **PDF-Export:** Professionelle Berichte
4. **Docker-Compose:** One-Click Deployment

PII-Anonymisierung in RAG-Pipelines für automotive Feedback-Daten

Keine existierende Literatur behandelt die Kombination von:

- Automotive-spezifische PII (VIN, Kennzeichen)
- Anonymisierung VOR Indizierung
- Integration in RAG-Pipeline

Demo-Flow (empfohlen)

1. **Dashboard öffnen** → KPIs zeigen (40 Feedbacks)
 2. **Analytics** → Trend-Chart, Filter nach ID.4
 3. **Chat** → "Welche Probleme gibt es mit dem Sprachassistenten?" → Quellen-Badges zeigen
 4. **Export** → Markdown-Bericht generieren → Download
-

Technologie-Stack

Layer	Technologie
Frontend	Next.js 14, Mantine UI
Backend	FastAPI (Python)
RAG	LangChain, GPT-4 Turbo
Vector DB	ChromaDB (persistent)
Keyword Search	rank_bm25

URLs für Demo:

- Frontend: <http://localhost:3000>
- API Docs: <http://localhost:8000/docs>

