

Car Domain — End-to-End Walkthrough (Single Page)

CAR DOMAIN — END-TO-END WALKTHROUGH (1 TB SCALE, TINY SAMPLE SHOWN)

A) SAMPLE DATA (ONE-TWO ROWS EACH)

- (1) Sales(DB)
S1,T101,CA,2023-03-15,120
S2,T102,CA,2023-03-15,55
- (2) R&D PDF (text): RND-42 "Falcon X ACC Tuning" → "Improved Adaptive Cruise Control (ACC)..."
- (3) OCR PDF (text): MAN-9 "Falcon X Owner Manual" → "...supports ACC as optional..."
- (4) Image caption: IMG-7 "Exploded view of Falcon X front radar module"
- (5) JSON features (2023 Falcon X): T101:[ACC,LaneKeep], T102:[ACC,AutoPark]
- (6) Review: TW1 "ACC on my Falcon X Pro handled traffic flawlessly."
- (7) Parts: P-001 "Front Radar Unit" fits T102; Feature ACC="Adaptive Cruise Control"
- (8) Patent: PAT-9 "Thermal control for radar modules" mentions radar for Falcon X

B) LOGICAL FLOW (TEXT DIAGRAM)

- [CQs] → [Ontology (OWL/RDFS + SKOS)]
- [Ingest 8 sources → Bronze → Silver (clean) → Gold (conformed)]
 - [Map Gold/XML/JSON to RDF (R2RML/RML) → Load Graph DB]
 - [Validate (SHACL) + Reasoner]
 - [Index unstructured (BM25 + Vector)]
 - [Q&A Router: SPARQL (facts) + Vector (evidence) + Federated SPARQL]
 - [Re-rank + Compose] → [Answer + citations]

C) PHYSICAL (AWS-FIRST; OSS IN PARENS)

S3 + Delta/Iceberg; Glue/Spark + MWAA/Airflow (Spark/NiFi/Airbyte); Texttract/Tesseract + Tika/GROBID;
Protégé(OWL/RDFS/SKOS) + SHACL; Graph DB: Neptune (GraphDB/Jena/Neo4j);
OpenSearch (BM25+kNN)/Qdrant/Milvus; API Gateway + FastAPI/Lambda;
Great Expectations, OpenLineage, Hermit reasoner.

D) GOLD TABLES (CONFORMED)

dim_vehicle_model: (M1, FX, "Falcon X")
dim_trim: (T101,T101,M1,2023), (T102,T102,M1,2023)
dim_feature: (F1,ACC,"Adaptive Cruise Control")
bridge_trim_feature: (T101,F1), (T102,F1)
dim_part: (P1,"P-001","Front Radar Unit")
bridge_part_fitment: (P1,T102)
dim_market_region: (CA,"California")
fact_sales: (S1,T101,CA,2023-03-15,120), (S2,T102,CA,2023-03-15,55)
dim_document: (RND-42,"Falcon X ACC Tuning",RND,2023-03-10), (MAN-9,"Falcon X Owner Manual",Manual,2023-02-01)
bridge_doc_entity_mentions: (RND-42,Feature,F1), (RND-42,Model,M1), (MAN-9,Trim,T101)