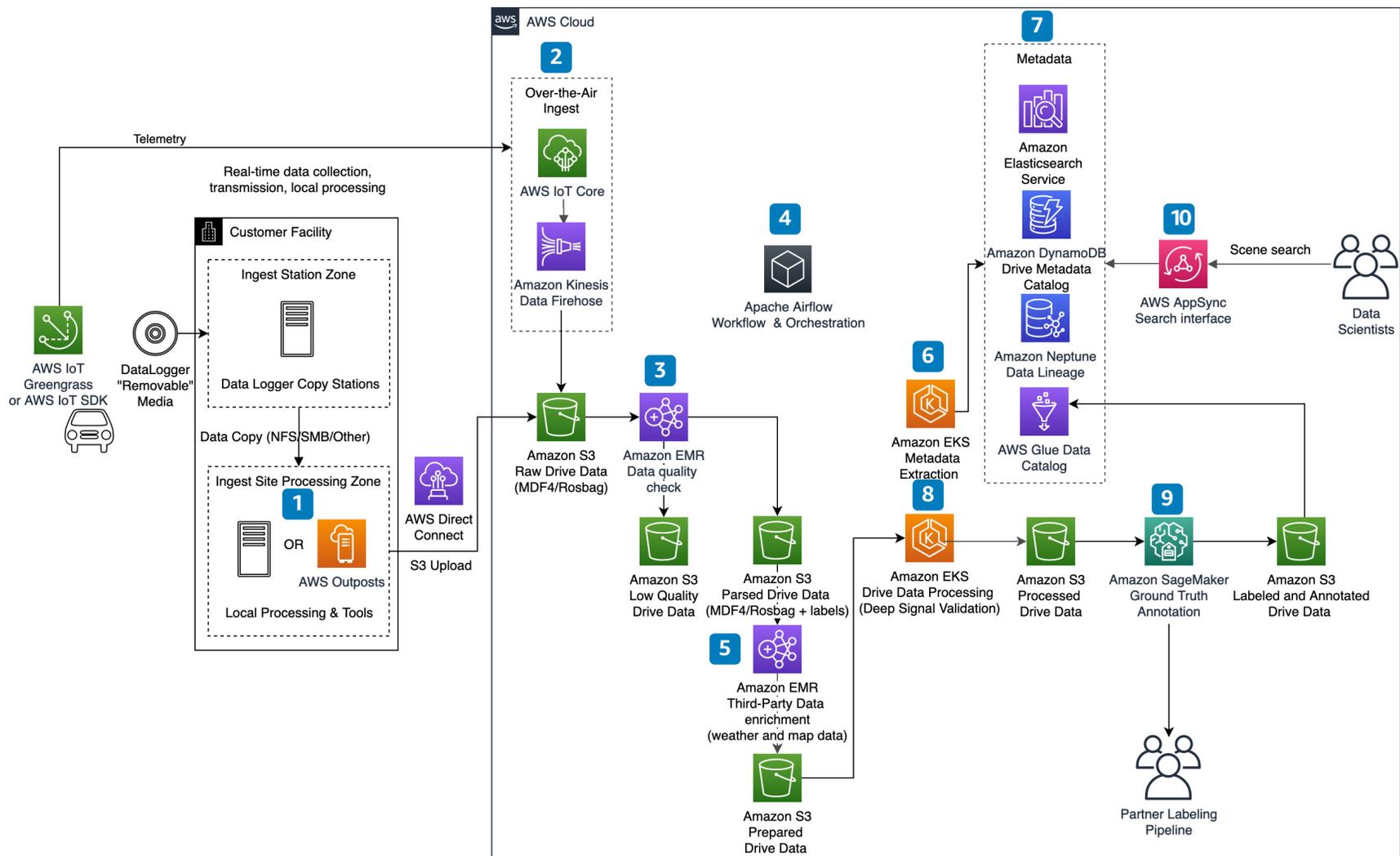


# Autonomous Driving Data Lake

Build an MDF4/Rosbag-based data ingestion and processing pipeline for Autonomous Driving and Advanced Driver Assistance Systems (ADAS).



- 1** Ingest data from autonomous fleet with **AWS Outposts** for local data processing.
- 2** Ingest vehicle telemetry data in real time using **AWS IoT Core** and **Amazon Kinesis Data Firehose**.
- 3** Remove and transform low quality data.
- 4** Schedule the extract, transform, load (ETL) jobs using **Apache Airflow**.
- 5** Enrich data with weather conditions based on GPS location and timestamp.
- 6** Extract metadata using **ASAM OpenSCENARIO** and store data in **Amazon DynamoDB** and **Amazon Elasticsearch Service**.
- 7** Store data lineage in **Amazon Neptune** and catalog data using **AWS Glue Data Catalog**.
- 8** Process drive data and perform deep signal validation.
- 9** Perform automated labeling using **Amazon SageMaker Ground Truth**.
- 10** Provide a search function for particular scenarios using **AWS AppSync**.

