Smart Claims

Insurance Lakehouse Demo

Claims Solution Accelerator

Agenda

- Challenges & Emerging Trends in Insurance Industries
- A typical claims workflow & need for ‘smart claims’
- How does Databricks enable ‘Smart Claims’
- Reference Architecture
- Workflow & DLT Pipeline
- ML & Rule Inferencing for Insight generation
- DB SQL parameterized queries & dashboard

State of the Industry

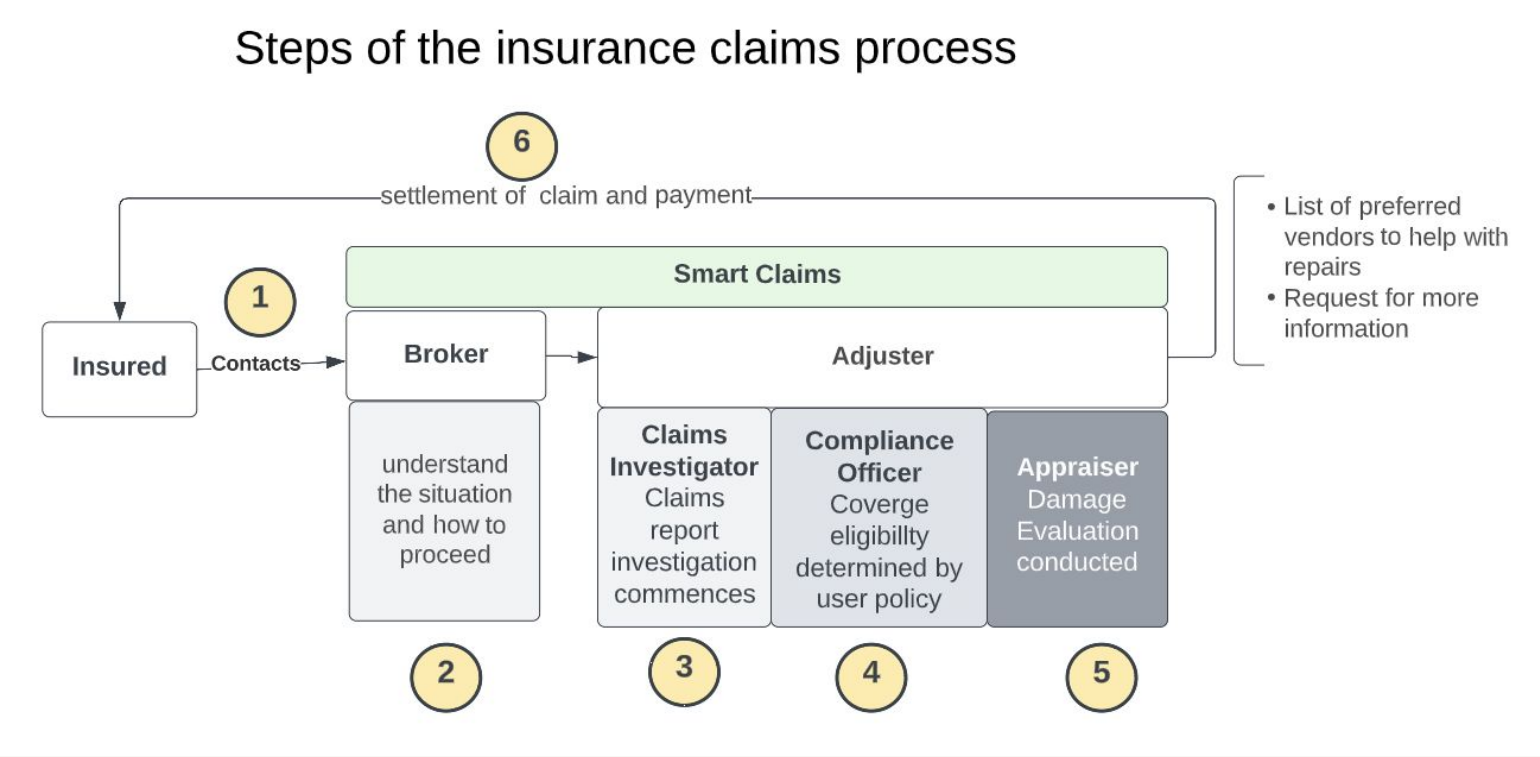
Challenges

- Insurance companies have to constantly innovate to beat competition
- Customer Retention & Loyalty can be a challenge as people are always shopping for more competitive rates leading to churn
- Fraudulent tractions can erode profit margins
- Processing Claims which can be very time consuming at times
- Tech debt, Skill gap, Rising costs, inflation, challenge from InsurTechs couple with demands from compliance & regulatory bodies
- How to improve the Claims Management process for faster claims settlement, lower claims processing costs and quicker identification of possible fraud.

Emerging Trends in Insurance

- According to EY: 'It is given that the future of insurance will be **data-driven** and **analytics-enabled**. But tomorrow's top-performing insurers will also excel at making **human connections** and applying the personal touch at the right time.'
- Deloitte in its '2023 Insurance outlook' states 'Technology infrastructure has improved, but focus needs to shift to **value realization**, and broaden historical focus from risk and cost reduction to prioritize greater levels of experimentation and risk-taking that drives ongoing **innovation**, competitive differentiation, and profitable growth.' with increased focus on ESG as value differentiator & DEI to broaden offerings.
- **Personalization** is not about bothering the customer with multiple touchpoints but wowing them with relevant insights that suit their need in a timely manner.
- Apart from courage and conviction, Innovation requires patience because no worthy change is delivered overnight. Hence the need to be on a platform that enables fast paced innovation and an architecture that is open, extensible and pluggable so that technology is never a constraint nor a hindrance to execution of novel ideas.

A typical claims process flow



1. The **Insured** contacts the broker who is the primary contact w.r.t. policy
2. The **Broker** examines the data to ensure that relevant details of the claim situation have been captured

The **Adjuster** takes over the investigation and may collaborate with internal/external experts to determine the amount of loss or damages covered by the insurance policy.

3. The **Claims Investigator** does due diligence on the paperwork
4. The **Compliance Officer** checks eligibility of coverage and ensure no foul play is involved
5. The **Appraiser** conducts a damage evaluation to determine the severity of the claim
6. The **Adjuster** will ensure payment is approved and released and communicates back to the Insured

Increasing Efficiency in Claim Processing

Asset Damage Assessment

- **What**
  - There is a lot of customer churn in insurance companies.
    - How can customer loyalty & retention be improved?
  - Processing Claims is time consuming
    - How can funds and resources be released in a timely manner to deserving parties?
  - Fraudulent transactions erodes profit margins
    - How can suspicious activities be flagged for further investigation?
- **Why**
  - Faster approvals, Lower Operating expenses
  - Detect & Prevent fraudulent scenarios, Lower Leakage ratio
  - Improve customer satisfaction, Lower Loss ratio
- **How**
  - Claim Automation
    - What aspects of the claims processing pipeline can be automated
  - Augmenting Info to claims data to aid Investigation – Recommend Next Best Action
    - Explainability for the human workflow
    - Claims Role (Adjustor, skill set, tenure time - who should take it based on claim characteristics)

How can Databricks help?

Databricks for Insurance Claims Lakehouse

Platform capabilities

- How
  - Databricks features used
    - Delta, DLT, Multitask-workflows, ML & ML Flow, DBSQL Queries & Dashboards
  - Unified Lakehouse architecture for
    - All data personas to work collaboratively on a single platform contributing to a single pipeline
    - All big data architecture paradigms including streaming, ML, BI, DE & Ops
  - Workflow Pipelines are easier to create, monitor and maintain
    - Multi-task Workflows accommodate multiple node types (notebooks, DLT, ML tasks, QL dashboard and support repair & run & compute sharing)
    - DLT pipelines offer quality constraints and faster path to flip dev workloads to production
    - Robust, Scalable and fully automated via REST APIs thereby improving team agility and productivity
  - BI & AI workloads
    - Created, managed with MLFlow for easy reproducibility and auditability
    - Supports any model either created or ported
    - Parameterized Dashboards that can access all data in the Laake and can be setup in minutes
- Built Upon earlier work done in
  - Damaged Car Classification Model (Github) & GTM Insurance Claims Ingestion (GitLab, Demo)

Reference Architecture

Datasets

Data characterization

Volume

High

- Telematic data

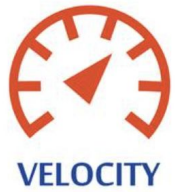
Medium

- Claim data

Low

- Policy, Policy-Holder data
- Accident data

Velocity



Near Real time

- Telematic data

Streaming

- Claim data, Accident data

Batch

- Policy, Policy-Holder data

Variety

Structured

- Policy, Policy-Holder, Telematic

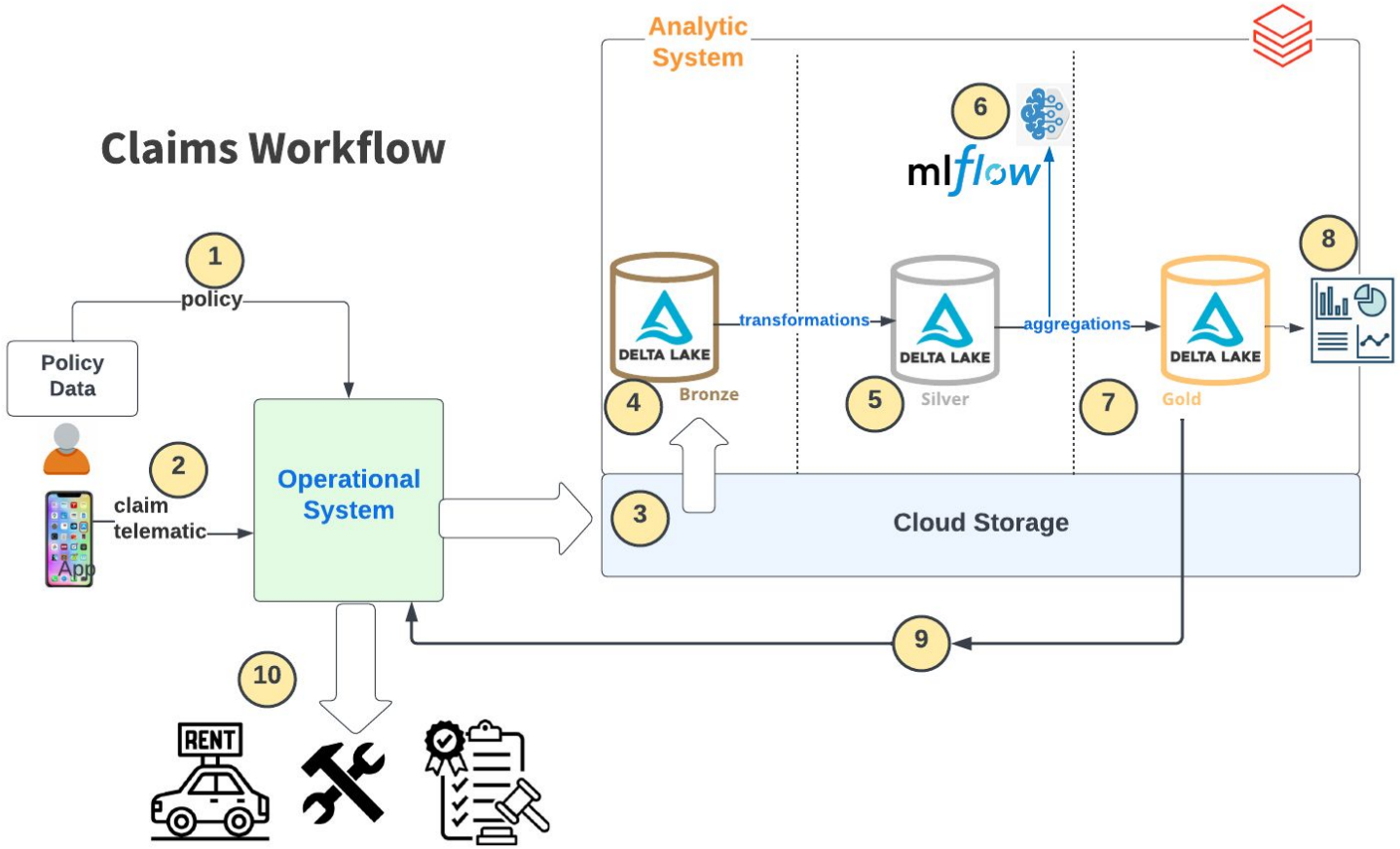
Semi-Structured

- Claim data

Unstructured

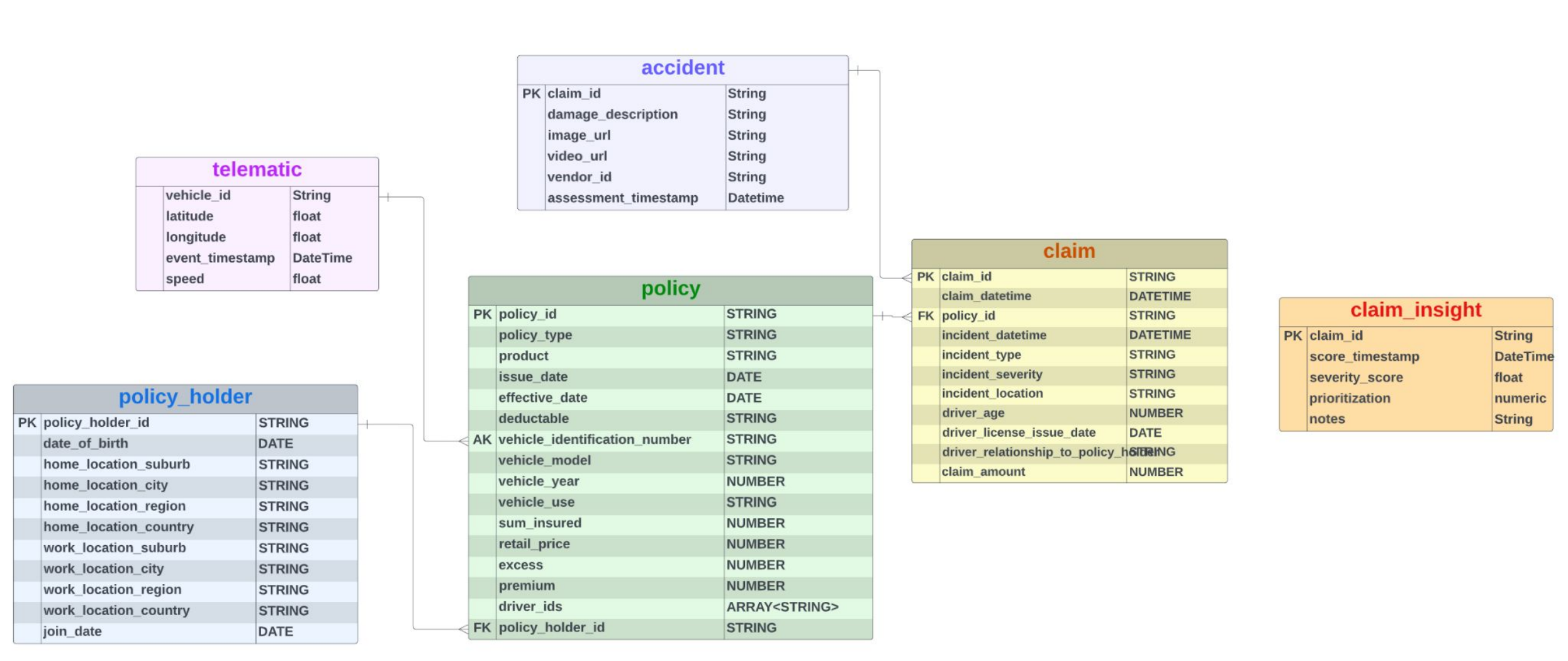
- Accident data
  - Image and Video & text

Claims Workflow



1) Policy data ingestion
2) Claims and telematics data ingestion
3) Ingest all data sources to the cloud storage
4) Incrementally Load Raw data to Delta Bronze table
5) Transform and Manipulate data
6) Model scoring (and model training in the training pipeline)
7) Load predictions to a gold table and perform aggregations
8) Dashboard visualization
9) Feed the results back to the operational system
10) Claims routing based on decision

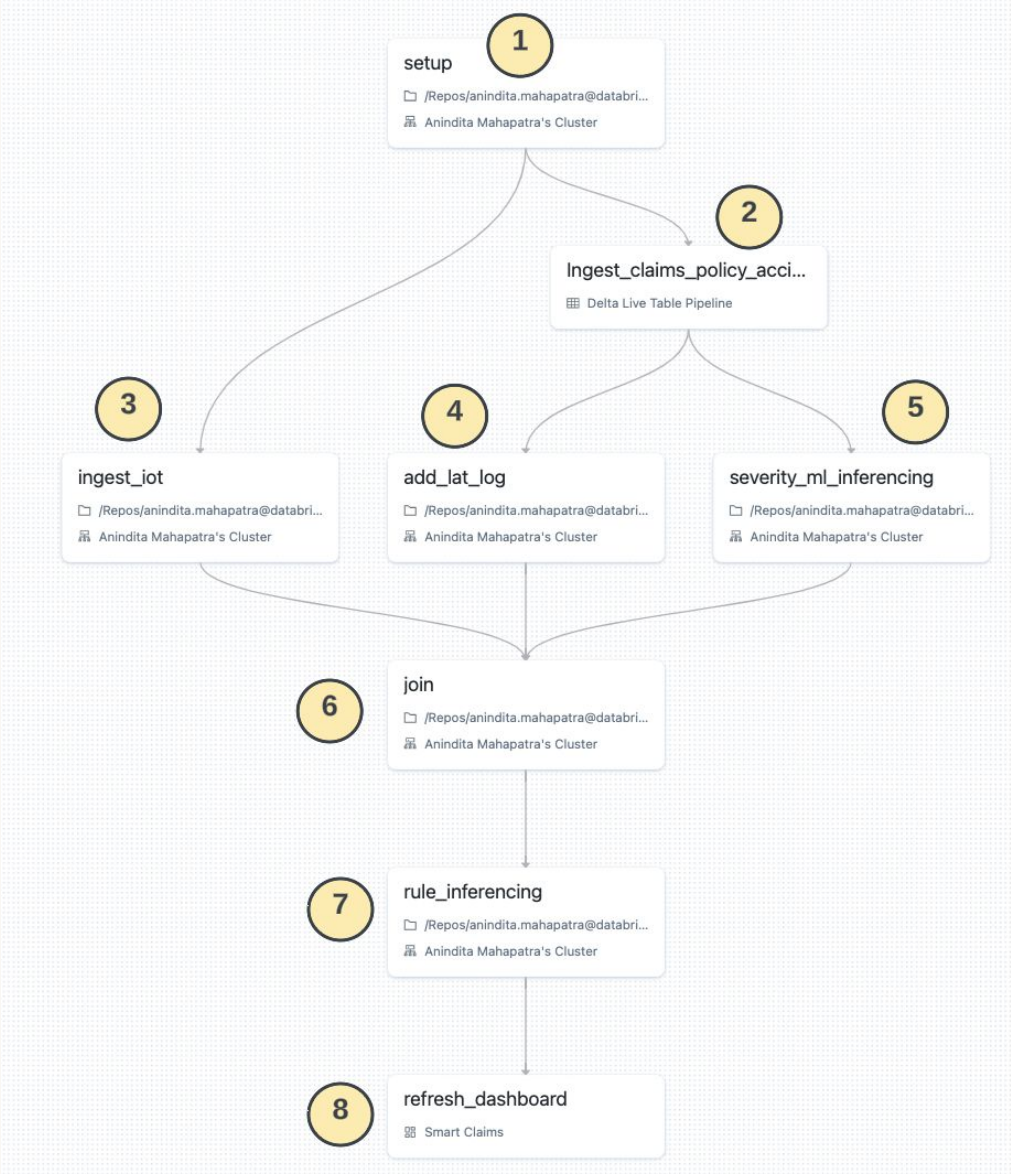
Domain Data Model



Workflows

Multi-task Workflows

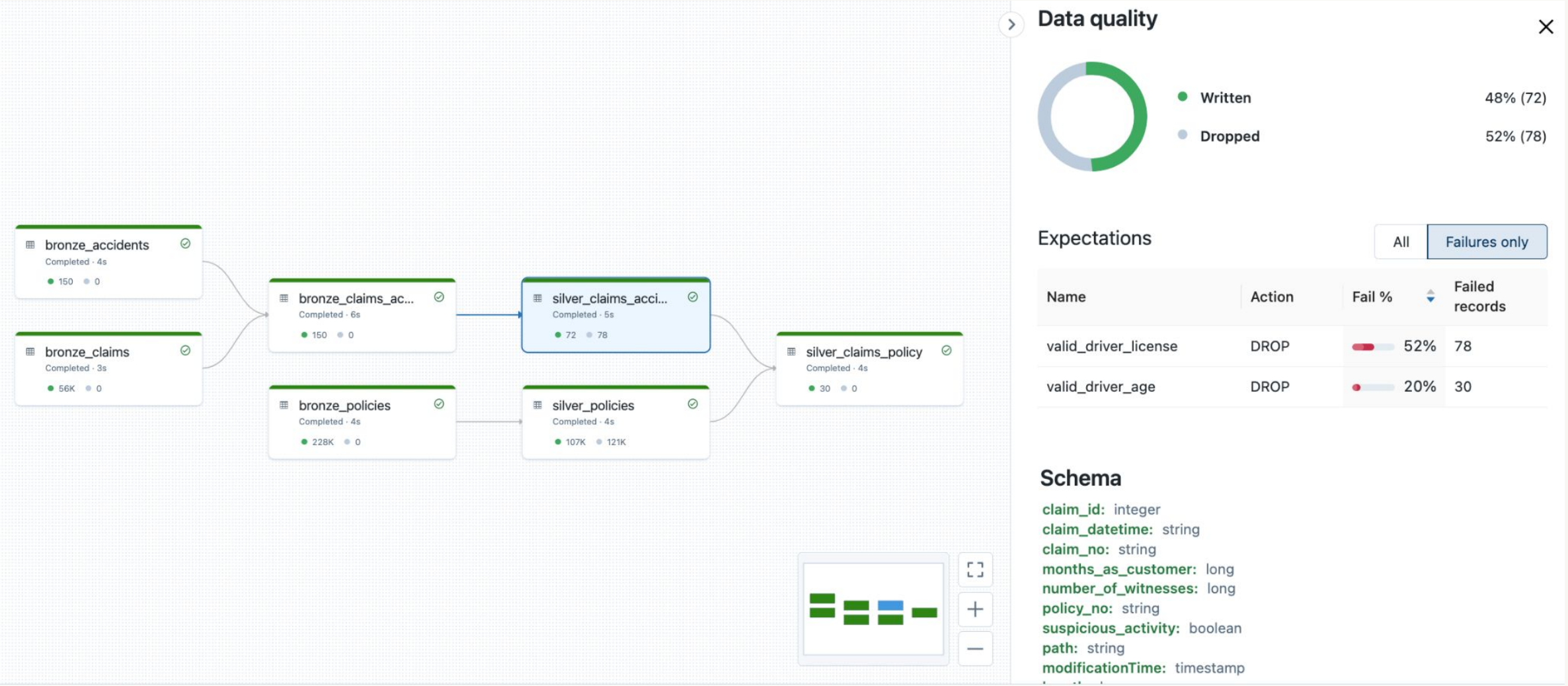
Orchestration of pipelines that is inclusive of all tasks (DE, AI & BI)



1) Policy data ingestion
2) Claims and telematics data ingestion
3) Ingest all data sources to the cloud storage
4) Incrementally Load Raw data to Delta Bronze table
5) Transform and Manipulate data
6) Model scoring (and model training in the training pipeline)
7) Load predictions to a gold table and perform aggregations
8) Dashboard visualization
9) Feed the results back to the operational system
10) Claims routing based on decision

Delta Live Tables (DLT)

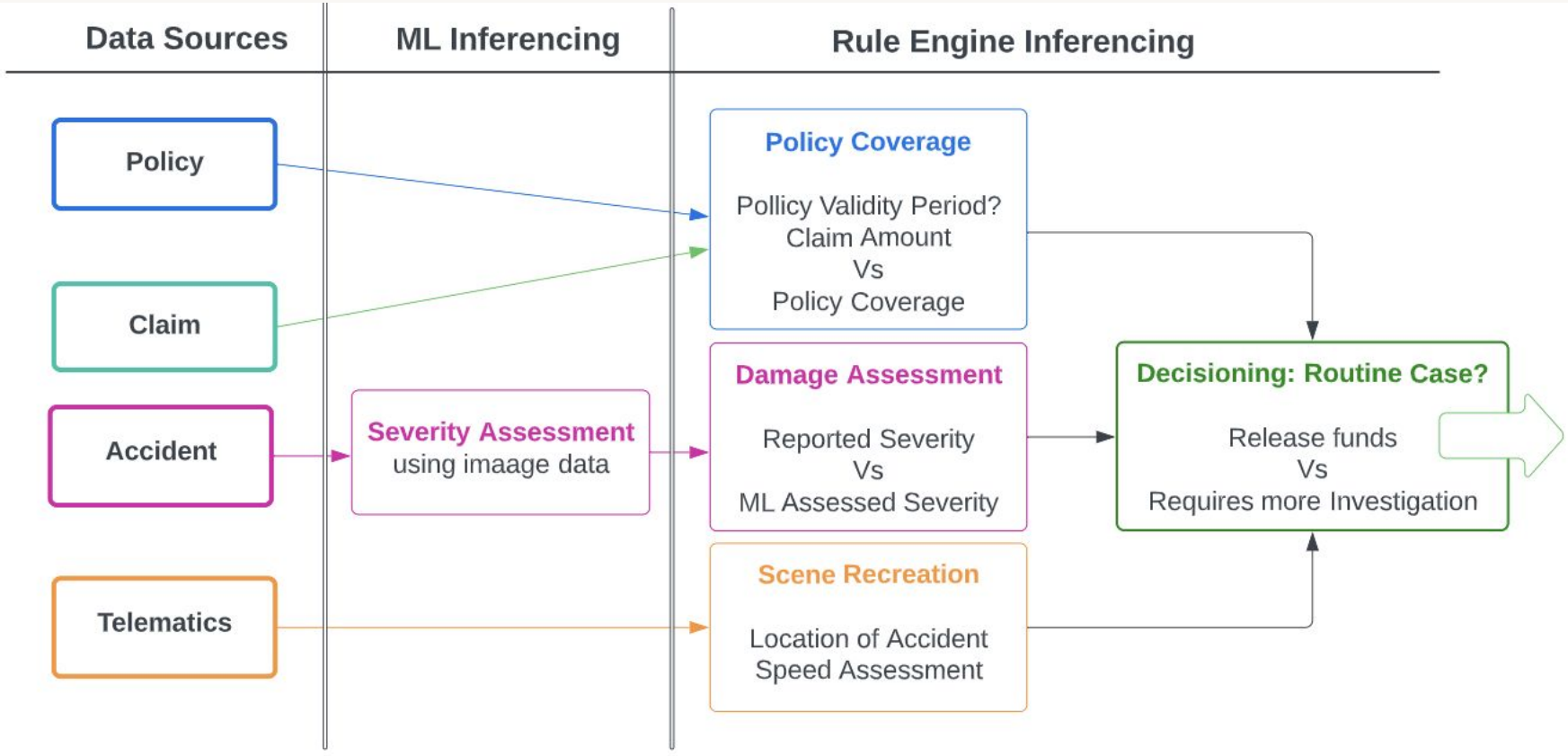
Quality Constraints via Expectations



Insight Generation

Data to Insights

ML & Configurable Rules Engine



- A pre-trained **ML Model** is used to score the image attached in the claims record to assess the severity of damage.
- A **Rule Engine** is a flexible way to define known operational static checks that can be applied without requiring a human in the loop, thereby speeding up '*routine cases*'. When the reported data does not comply with auto detected info, flags are raised to involve additional human investigation
- This additional info helps a claims investigator by **narrowing down** the number of cases that need intervention a well as by narrowing down the specific areas that need additional follow up and scrutiny

Results Dashboard

Loss Summary

- Loss Ratio is computed by insurance claims paid plus adjustment expenses divided by total earned premiums
- Damage is captured in 2 categories - property & liability - their loss ratios are tracked separately
- Incident type refers to damage on account of theft, collision
- Damage Severity is categorized as trivial, minor, major, total loss
- Analyzing recent trends helps to prepare for handling similar claims in the near future

Results Dashboard

Claims Investigation helps investigation officer narrow down scrutiny areas

Claims Investigation

14,635	3,515	17	4	15	23
#claims	#Suspicious	#Expired Policies	#Severity Mismatch	#ExceedsClaimAmount	Needs Investigation
19 hours ago	19 hours ago	19 hours ago	19 hours ago	19 hours ago	19 hours ago

Insights

#	claim_no	claim_datetime	policy_no	reported_severity	valid_policy	correct_severity_assessment	approve_claim_amount	claim_amount_total	claim_amount_injury	claim_amount_property	claim_amount_vehicle
20	00ecfeed-4083-4a2a-8bde-45a74eef64e2	2017-11-06 15:26:00	102097638	Major Damage	✗	✓	✗	47790	5310	5310	37170
21	00c5bcfb-4bfa-4d54-a5e4-49be9f699ce7	2019-10-21 15:33:00	102120018	Total Loss	✓	✗	✗	79750	14500	14500	50750
22	00c5e7c3-3f98-41c4-a0e1-52122dedcf2a	2016-11-22 09:48:00	102106539	Major Damage	✗	✗	✗	78500	15700	7850	54950
23	00df4270-3464-4eda-a641-75d5d2fa3e23	2015-09-09 08:34:00	102059460	Major Damage	✗	✓	✗	64200	10700	10700	42800
24	00d46108-12eb-423e-8513-6c3b0798e8e6	2017-11-20 12:04:00	102146876	Major Damage	✗	✓	✓	61290	6810	6810	47670
25	00b36194-1d4e-4646-985d-54b768ffef4f	2018-08-15 13:36:00	102085418	Minor Damage	✓	✓	✓	32480	4060	4060	24360

1 2 3 >

20 hours ago

Claim Details

claim_no	to_date(claim_datetime)	incident_type	incident_date	collision_type	collision_number_of_vehicles_involved	driver_age	driver_insured_relationship	driver_license_issu
09b42db4-096b-42f2-84b6-b089817449ce	2016-06-21	Multi-vehicle Collision	2016-06-18	NULL	4	51.00	other-relative	20

5 days ago

Policy Details

CUST_ID	POLICY_NO	POLICYTYPE	pol_issue_date	pol_eff_date	pol_expiry_date	BODY	MAKE	MODEL	MODEL_YEAR	CHASSIS_NO	USE_OF_VEHICLE	D
3304.0	102122697	COMP	2018-09-17	2018-09-17	2019-10-16	COUPE	MINI	COOPER COUPE	2018.0	WMWXM5104J2F37698	PRIVATE	1

20 hours ago

Driver Details

CUST_ID	POLICY_NO	POLICYTYPE	pol_issue_date	pol_eff_date	pol_expiry_date	BODY	MAKE	MODEL	MODEL_YEAR	CHASSIS_NO	USE_OF_VEHICLE	D
3304.0	102122697	COMP	2018-09-17	2018-09-17	2019-10-16	COUPE	MINI	COOPER COUPE	2018.0	WMWXM5104J2F37698	PRIVATE	1

20 hours ago