

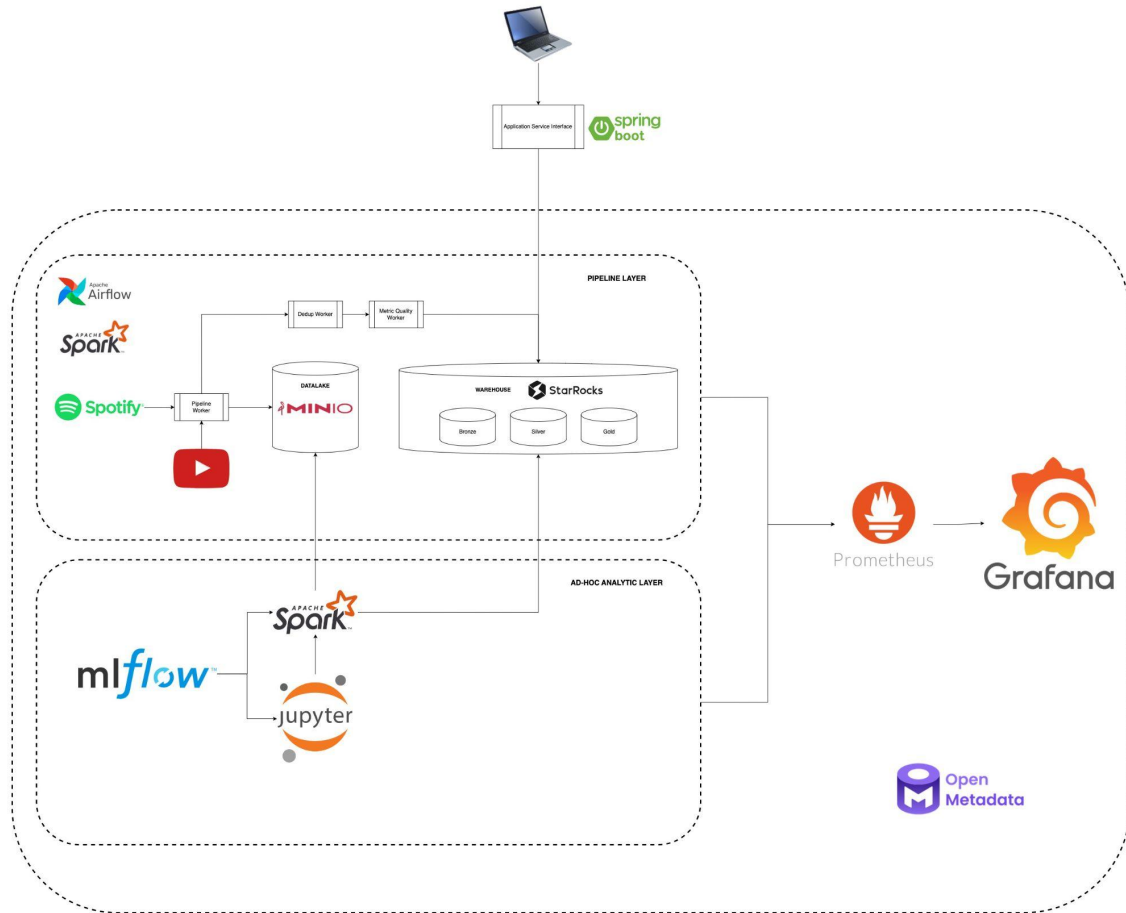


Data Architecture and Strategy

By : Rahul M Soemitro



Architecture Design



Data Quality Strategy

Completeness Checks

- Required fields presence
- Acceptable null ratio per column
- Record count validation

Accuracy Checks

- Data type validation
- Range validation
- Pattern matching and Similarity algorithm for (Track Name and Singer Name)

Consistency Checks

- Cross-reference validation
- Duplicate detection
- Time series continuity

Monitoring & Maintenance

Pipeline Health

- Task success rate
- Processing latency
- Error rates
- Data freshness

Data Quality

- Validation failure rate
- Missing data percentage
- Duplicate record count
- Schema violation count

System Performance

- Processing time
- Resource utilization
- Queue lengths
- Storage usage