



Team Name : Elint Data

Problem Statement : Unified Data Platform for Verification

Brief about the Idea:

To build a Unified Data Platform using latest DE Frameworks capable of scaling up responsively to any data volume and data velocity at source side in a secured way and make the data available in UDP for the State Police Department with significant reduction in data loading cost, data loading time, data flow latency, data storage cost, data querying cost, data querying time, thus enabling easy Data Verification Process across different different sets of data.

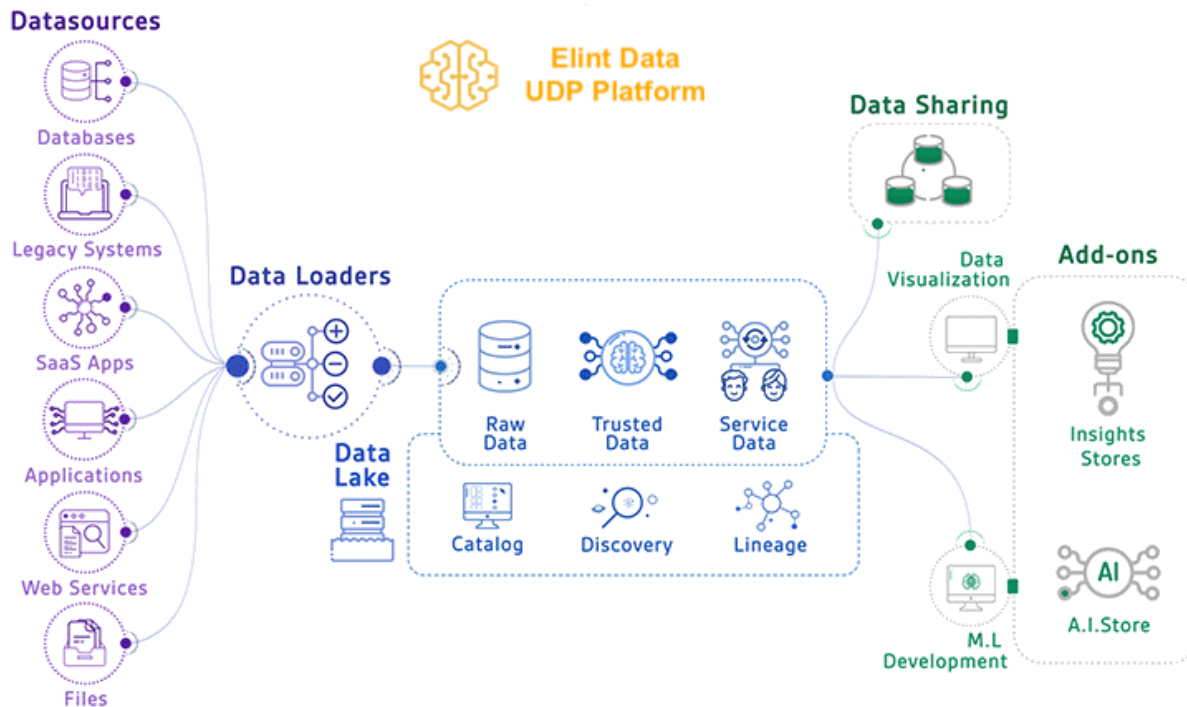
Opportunity:

Data across different sources, formats need to be brought at one place to create that “Single Source of Truth” to enable the holistic Data Verification. As per our understanding, currently the various data related to the citizens reside on different sources, hence any verification would be very time and labor intensive.

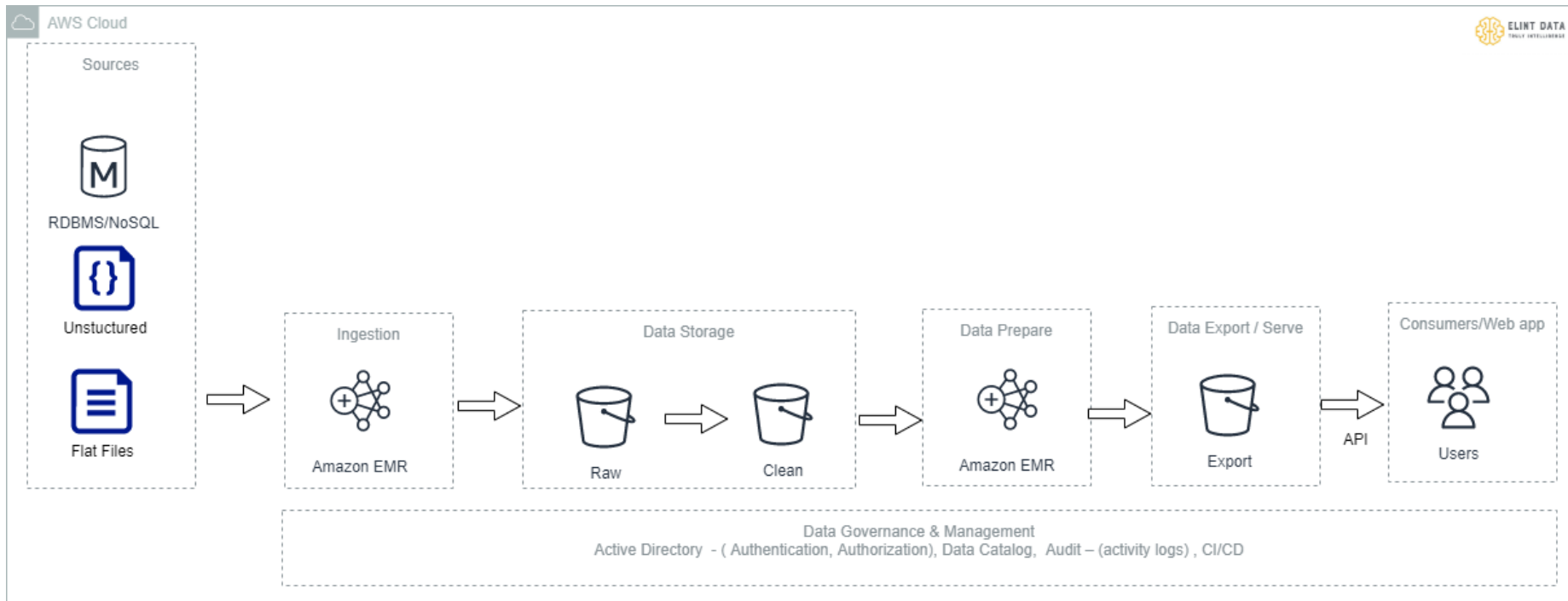
Our low-code/ no-code platform would bring all the data from different sources at one Data Lake and will empower all the Verification activities by the State Police, hence reducing time and money, and increasing effectiveness across.

This is an indigenous solution to solve the disparate data sources problem of the State Police.

List of features offered by the solution :



Business Logic of the solution :





Technology used :

Apache HUDI/ Delta Files

Databricks

PySpark

Azure/ AWS/ GCP

Debezium

Kinesis/ Kafka

Usage of this platform:

1. Create a “Single Source of Truth”
2. Various visualization dashboard of top of this data, having real-time accessibility of events
3. The platform also works on high frequency image data as well, this can be used for real time verification across different locations
4. Various utility mobile apps on top of this platform to serve different purpose and objectives like – historical location traverse of a person of interest

Estimated cost of/after implementing the solution :

The solution development cost would be roughly estimated around 20 lakhs, but is hugely governed by following two factors:

1. Number of different data sources
2. Volume of the data



POLICE HACKATHON

H2S
HACK2SKILL

THANK YOU

