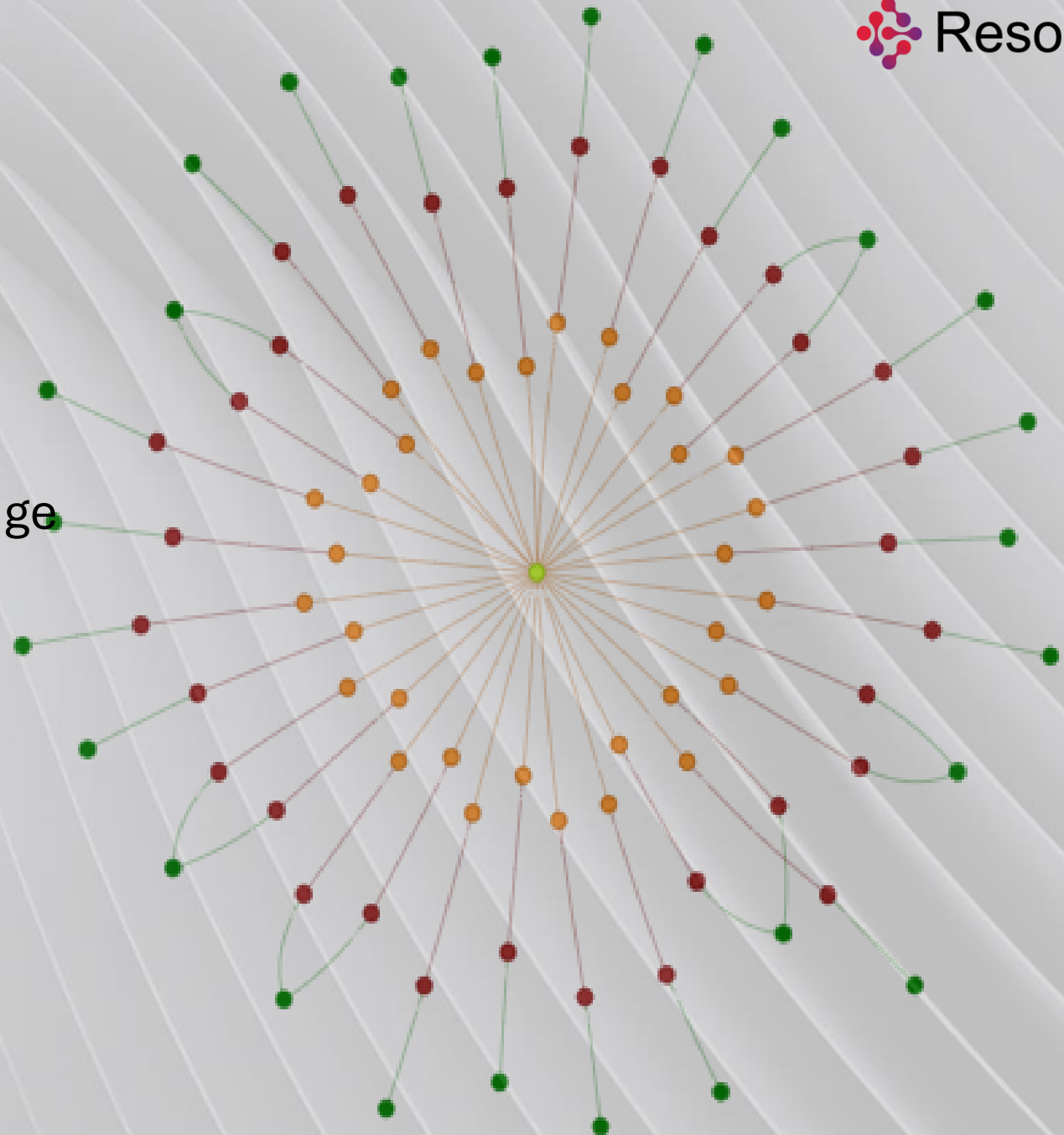# ResolveR -
Extract, Combine, Match and Manage

Key Business Challenges and Asks

Fragmented and inconsistent data across multiple systems and sources

Inefficient data matching and deduplication processes leading to duplicate and inconsistent data

Difficulty in establishing a single source of truth for critical data entities

Inability to proactively monitor and address data quality issues
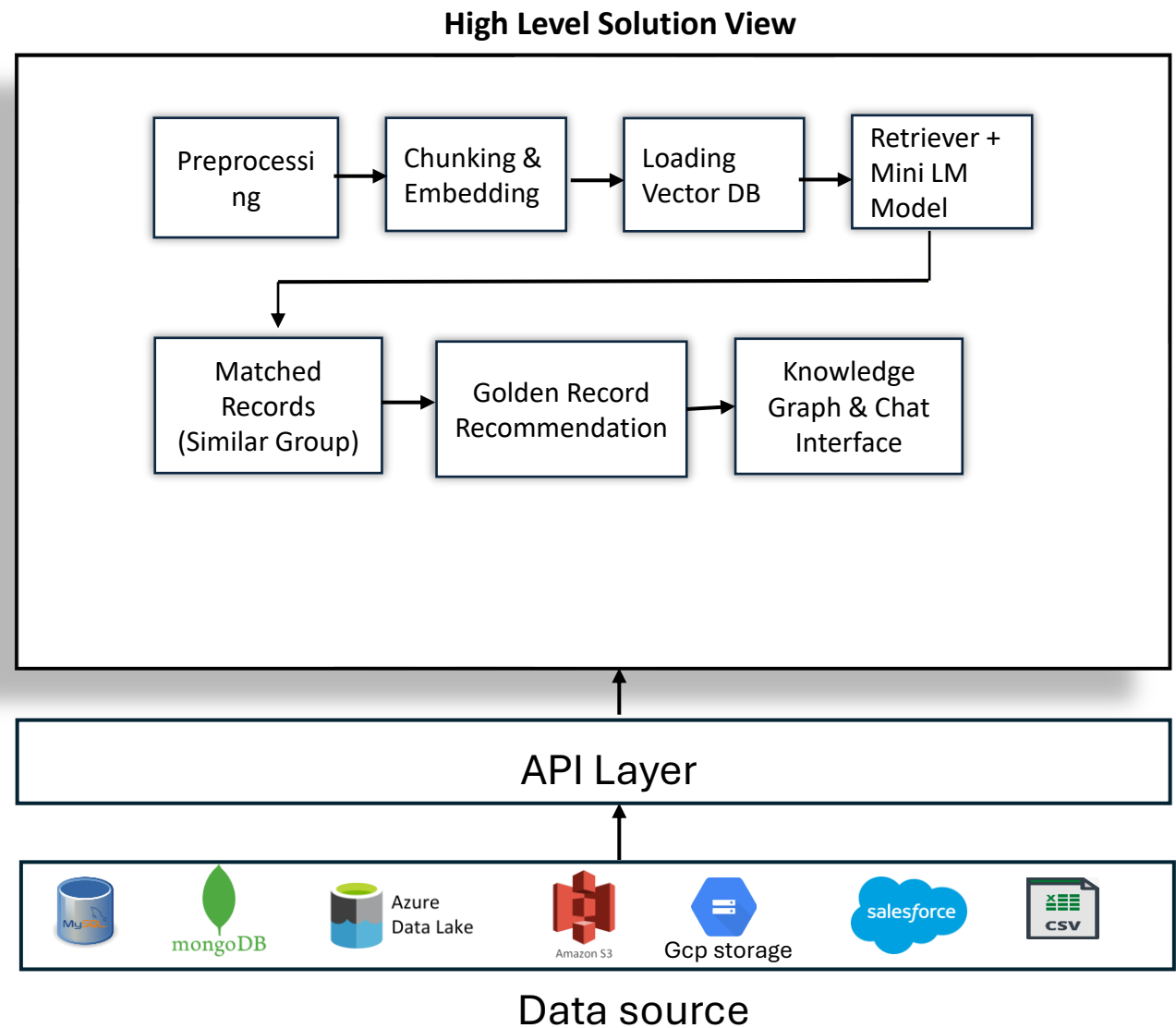
Traceability of the golden Records

Complex process for searching records and enabling data discovery

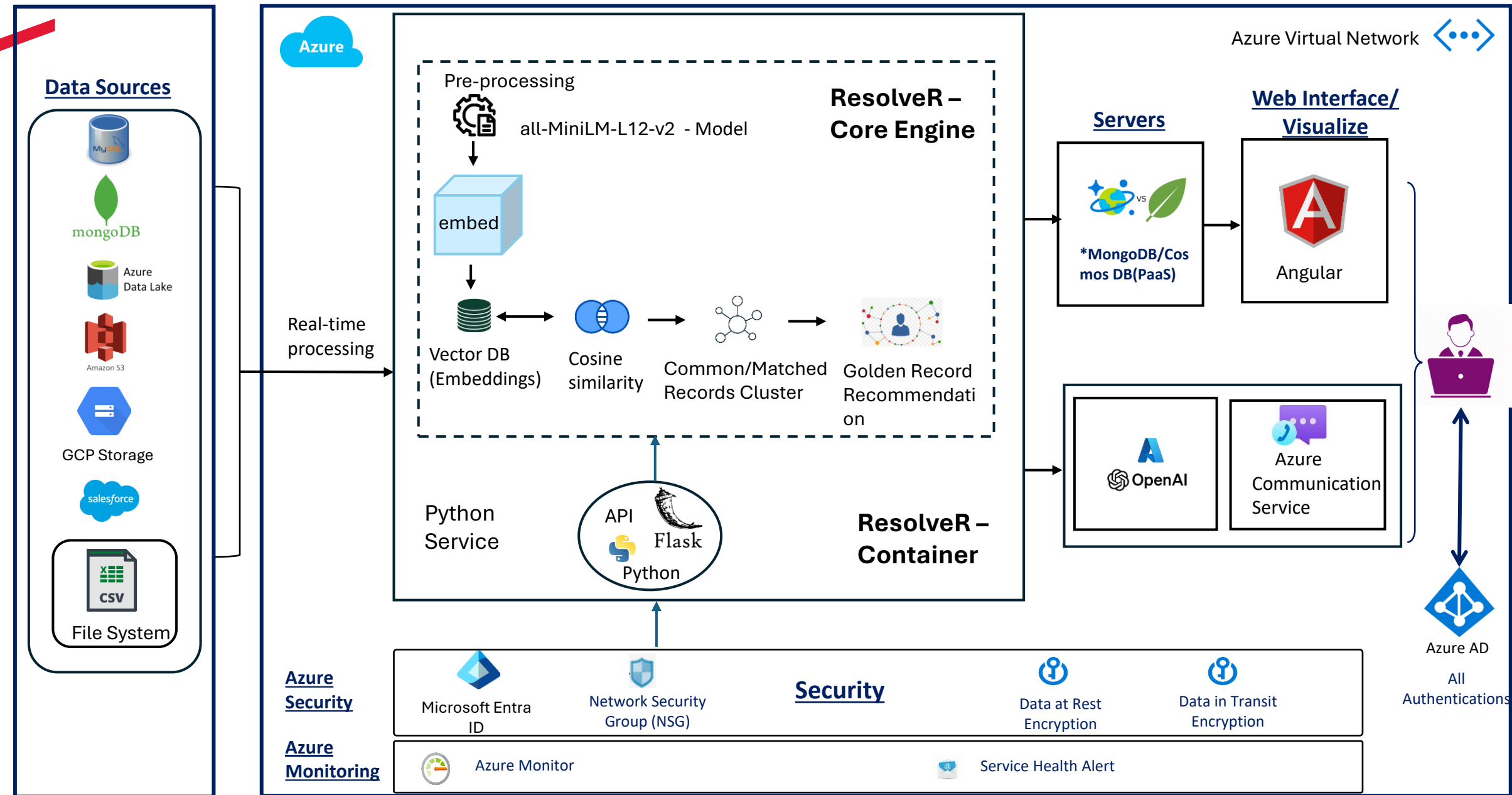Support of multi-tenant projects for entity resolutions

Maintaining history and tracking performance report

Solution Approach

A comprehensive strategy leveraging data integration, matching AI-ML algorithms, and Generative AI to accurately resolve entities, ensure data consistency, and improve golden records traceability across multiple sources

High Level Solution View

ResolveR on Azure - Solution Architecture

Data Sources

- MySQL
- mongoDB
- Azure Data Lake
- Amazon S3
- GCP Storage
- salesforce
- File System (CSV)

Real-time processing

Azure Virtual Network

ResolveR – Core Engine

Pre-processing

all-MiniLM-L12-v2 - Model

embed

Vector DB (Embeddings) → Cosine similarity → Common/Matched Records Cluster → Golden Record Recommendation

Python Service — API Flask Python

ResolveR – Container

Servers

*MongoDB/Cosmos DB(PaaS)

Web Interface/ Visualize

Angular

OpenAI | Azure Communication Service

Azure Security

Security

Microsoft Entra ID | Network Security Group (NSG) | Data at Rest Encryption | Data in Transit Encryption

Azure Monitoring

Azure Monitor | Service Health Alert

Azure AD

All Authentications

Solution Features

Gen AI based potential column match recommendation

Data profile to reflect null % for match column selection

Configurable weightages and threshold % for AI match algorithm

Golden records recommendation for individual records across sources

Standardizing the attributes (Combine/Unify)

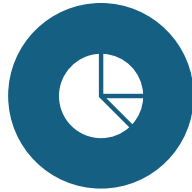Knowledge graph for golden records traceability

LLM based Chat interface

Store golden records recommendation & performance trend

Support for multiple database with pre-built adaptors

Solution Benefits

Minimize errors from duplicate or inaccurate data

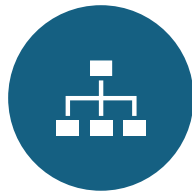Golden record recommendations to drive data driven decision-making

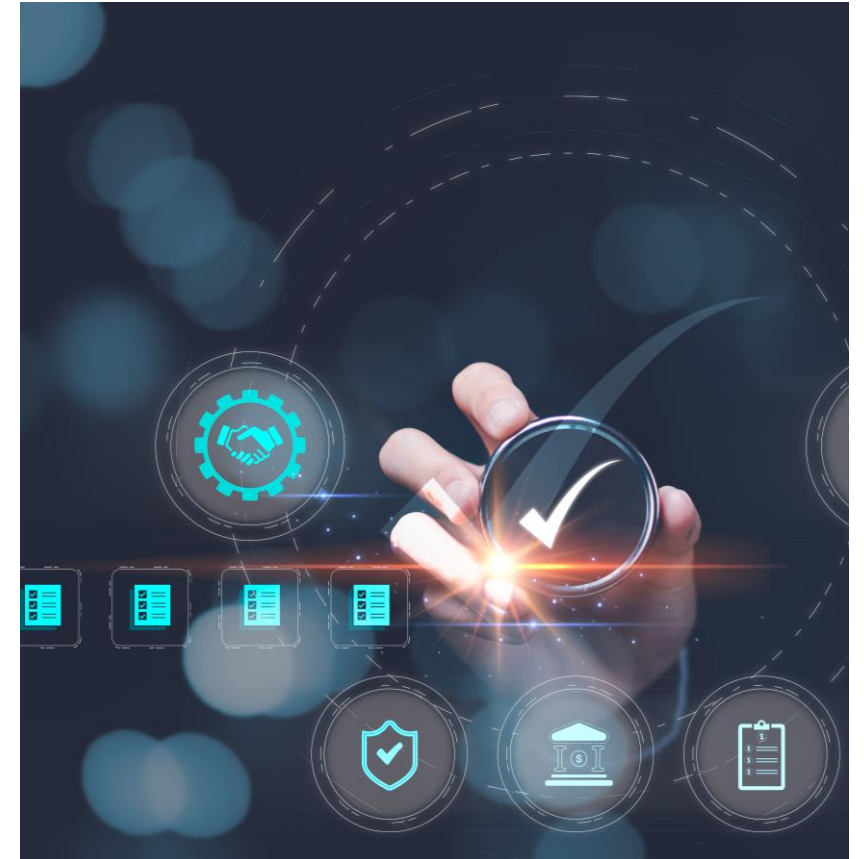Faster time to market for entity resolution

Harmonize records across systems for reliable operations

Better 360-degree customer/product/supplier views

Access to accurate, consolidated records

TECH
mahindra