



# Our Data Practice



डेटा प्रैक्टिस  
ओ सी आई



---

# Data practice – what we do

We offer comprehensive set of services for data-driven business to increase their productivity and enhance customer experience using intelligent solutions.

## Data Engineering

Build complex data ingestion and transformation pipelines

## Data Integration

Data integration across Enterprise on-premise, cloud apps and third party solutions

## Data Warehousing

Build and maintain Data Warehouse – cloud and on-premise

## Data Analytics

Build solutions based on machine learning, text analytics & vision analytics to gain business insights and improve outcomes.

## Data Visualization

Develop BI dashboards to present insights and actionable information

# Our clients

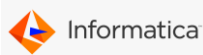


# Data skills

## Data Engineering



aws



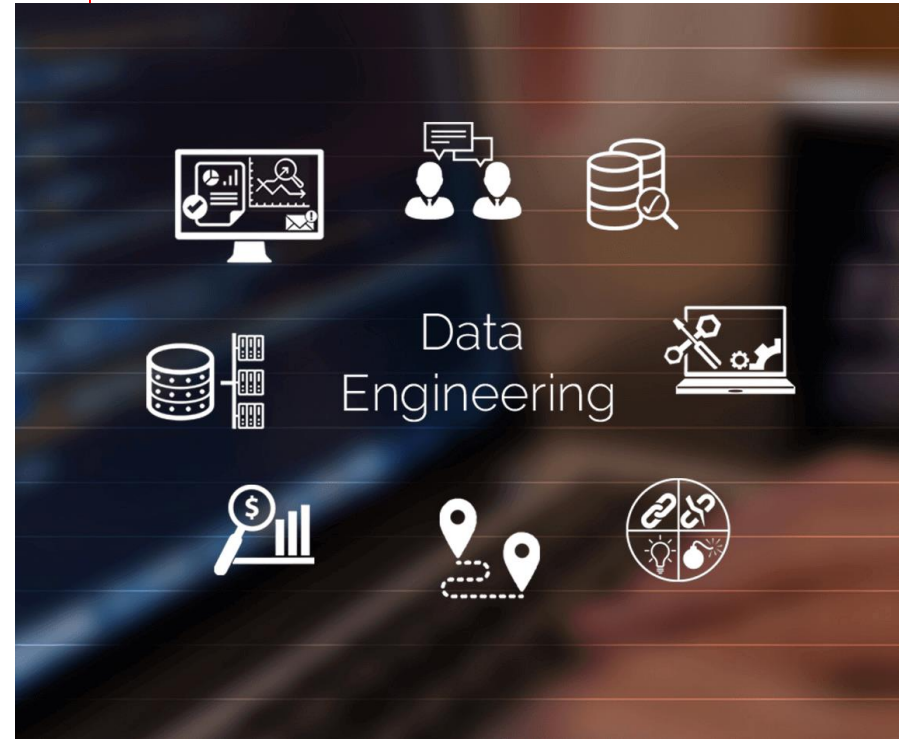
## Data Analytics




























## Data Visualization



# Data Engineering



# Our cloud data engineering skills

	aws	Azure	Google Cloud Platform
<b>Data Ingestion</b>	 Glue  Lambda  AWS DMS	 Data Factory  IR  Synapse Pipeline	 Data Fusion
<b>Data Orchestration, Scheduling, Logging and Notification</b>	 Step Functions  SNS  Cloud Watch  Event Bridge	 Data Factory  Monitor	 Cloud Composer
<b>Data Processing and Querying</b>	 Lambda  Athena	 Synapse Analytics  Functions	 Big Query  Cloud Dataflow
<b>Data Storage and Warehouse</b>	 S3  Redshift	 ADLS Gen2  SQL Data Warehouse	 Cloud Storage

# Data engineering, warehousing & Business Intelligence

Cloud, On-Prem & Hybrid

Data mart

Descriptive analysis

Data lake

Diagnostic analysis

AZURE

Predictive analysis

AWS

BI Dashboards

DW

ETL



Building complex data ingestion and curation pipelines for one of the world's largest pulp and paper major, to populate a data lake. Data is ingested from legacy, on-prem and cloud based DBs, 3<sup>rd</sup> party APIs, flat files, json, and real time streaming. The data gets further transformed to create a Delta Lake that facilitated upserts & columnar data format in file storage.

The Data lake is further used for building a massive parallel processing cloud Data warehouse that serves the reporting needs, performance tracking, marketing initiatives, demand predictions and customer analysis.



Helping a digital media marketing aggregator platform to ingest data from multiple Ad platforms – Twitter, Google, Facebook etc. into a unified Data Warehouse. The data is then used to provide a common aggregation of their ad related data across multiple platforms.

We are technical partners for one of the leading natural herbal products brand in the US. The client's existing system architecture was fractured and not scalable.



We are designing and building the new azure synapse data pipeline and the Data Warehouse that will serve as a single source of truth for the organization. The data is ingested from various source systems like CMS, PIM, CRM, CS(Zen-desk), Promo, E-commerce and ERP system. The scope of work includes extraction, cleansing, aggregating, transforming and validating the data to ensure accuracy and consistency.

# Data engineering, warehousing & Business Intelligence

Cloud, On-Prem & Hybrid

Data mart

Descriptive analysis

Data lake

Diagnostic analysis

AZURE

Predictive analysis

AWS

BI Dashboards

DW

ETL



We helped a leading Global MRO integrated supply chain services provider by creating a Data Warehouse which unifies transactional data from different client procurements. Currently we support the business and IT by managing day-to-day activities, building BI dashboards and resolving user requests thus enabling accurate financial reporting to achieve operational efficiency. This DW became a launch pad for several dashboards and apps for MIS, sales predictions, KPI monitoring and efficiency improvement.



We utilized SSIS to build an On Premise DWH in SQL Server for the Sports and Entertainment division of leading food services provider. SSIS Tasks were used to enable dynamic connections to the POS data sources, to extract incremental data, to archive old data and for data cleansing and transformation. SSIS tasks were further used to run dimensional modelling procedures in parallel to add data to the Type 2 SCD Data Model in the DWH.



Impressico has helped with solutions for data ingestion, transformation, and storage of data generated by IoT sensors and BI dashboards to identify actionable insights based on real-time data. The actionable insights are presented through alerts, notifications, and visualizations. The platform allows building managers to remotely monitor an endless array of building performance data.

# Data engineering, warehousing & Business Intelligence

Cloud, On-Prem & Hybrid

Data mart

Descriptive analysis

Data lake

Diagnostic analysis

AZURE

Predictive analysis

AWS

BI Dashboards

DW

ETL

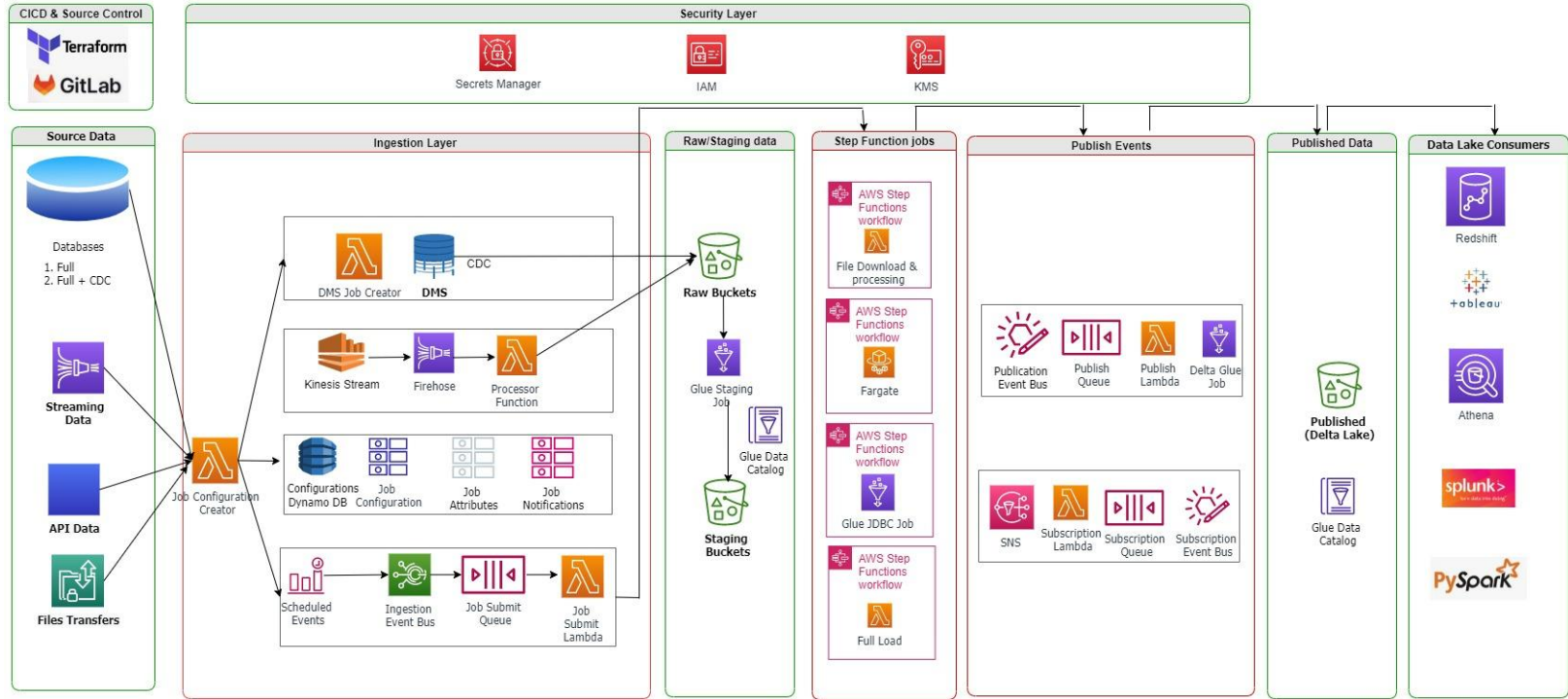


For Aramark, we created and managed data pipelines to harmonize disparate data sources into cohesive datasets using Palantir Foundry. This process encompasses data ingestion, transformation, harmonization using Hugging face Semantic Search – cosine similarity, and validation, ensuring efficient inventory management in manufacturing and supply chain companies. We also design and create Ontology objects and their actions for business requirements and objectives. Additionally, we monitor health checks to ensure the stability and performance of our data pipelines, create and analyze data using complex workshop apps tailored to business needs, and export data from Palantir to Snowflake, facilitating seamless integration with external data warehousing solutions.

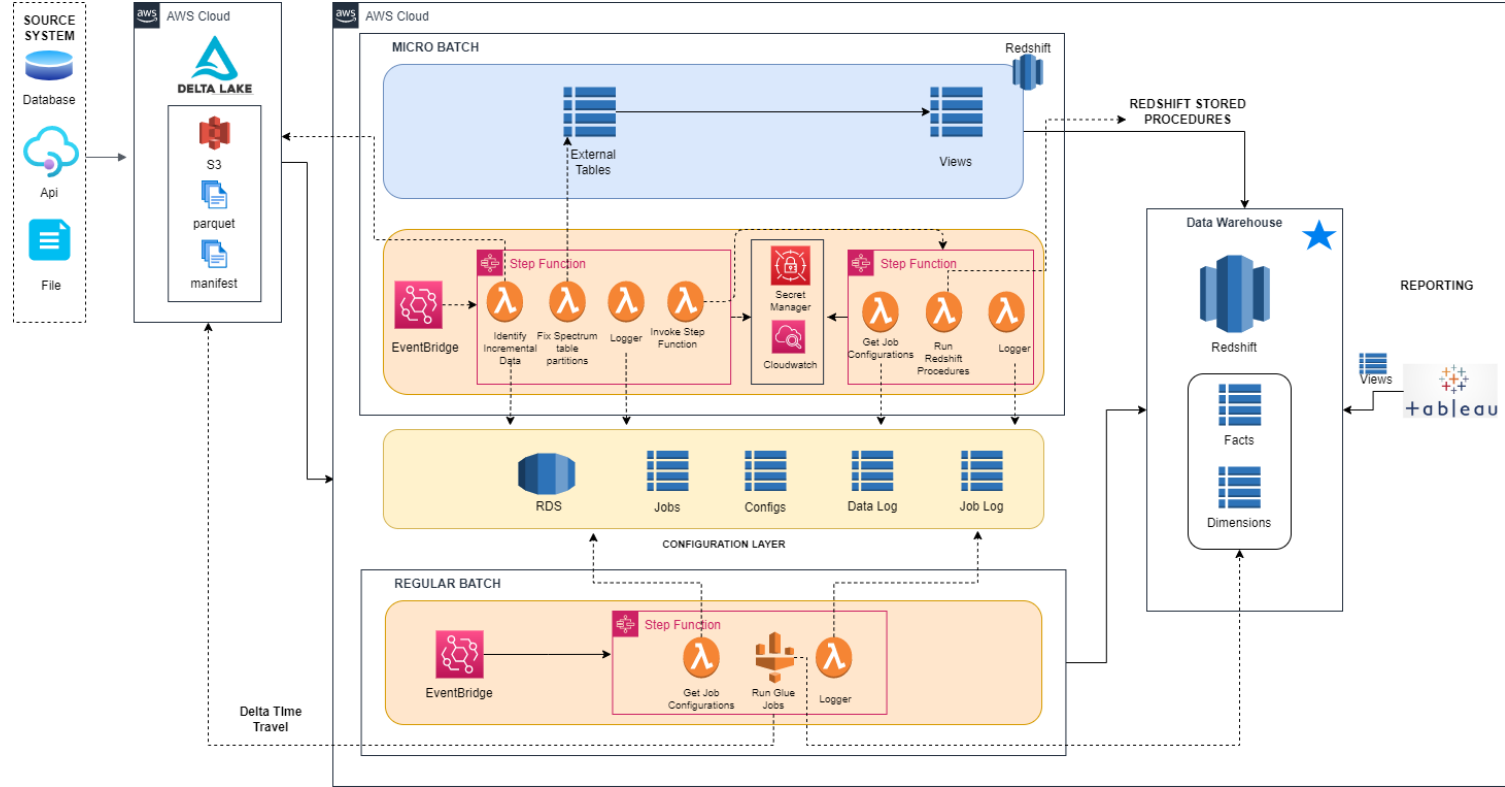


At Mhymatch, we streamlined user conversation storage by integrating with GCP CloudSQL, ensuring secure and efficient data management. Additionally, we connected WhatsApp via Twilio with GCP Pub/Sub to enhance recruitment messaging. This integration enables real-time, reliable communication, significantly improving our recruitment process including job seekers and recruiters.

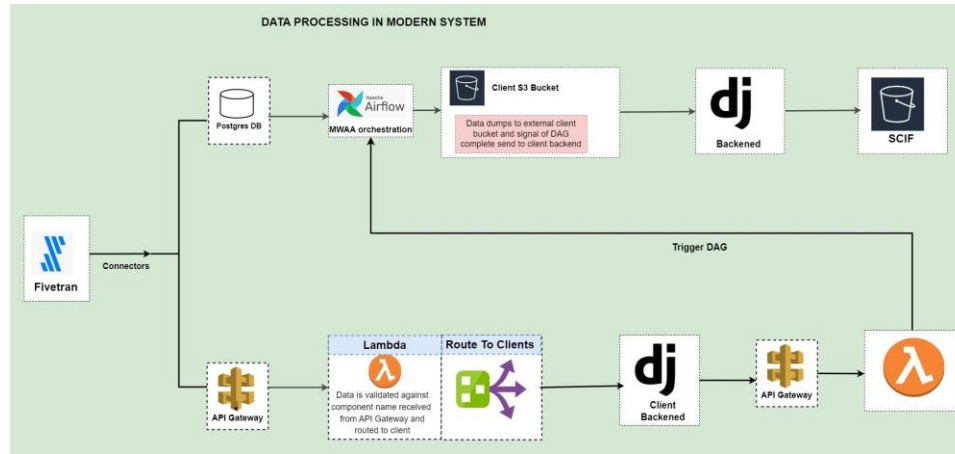
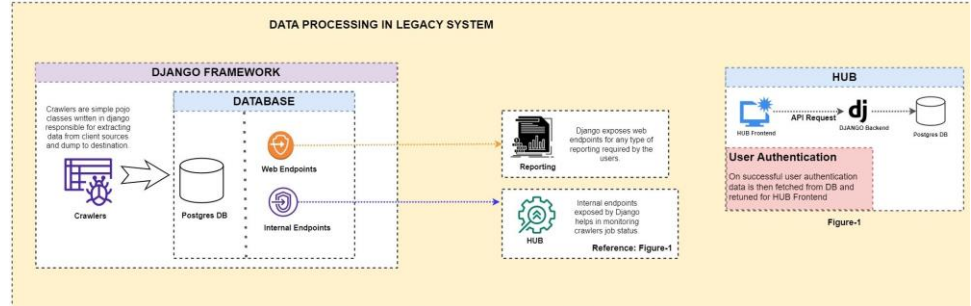
# Generic AWS ingestion architecture



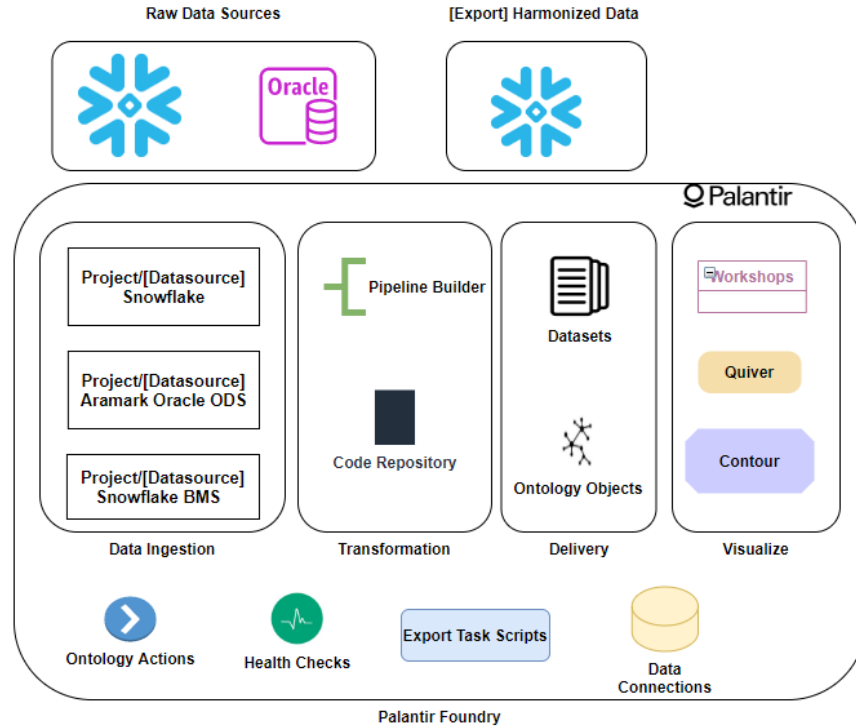
# AWS DWH architecture: ingestion, transformation and visualization



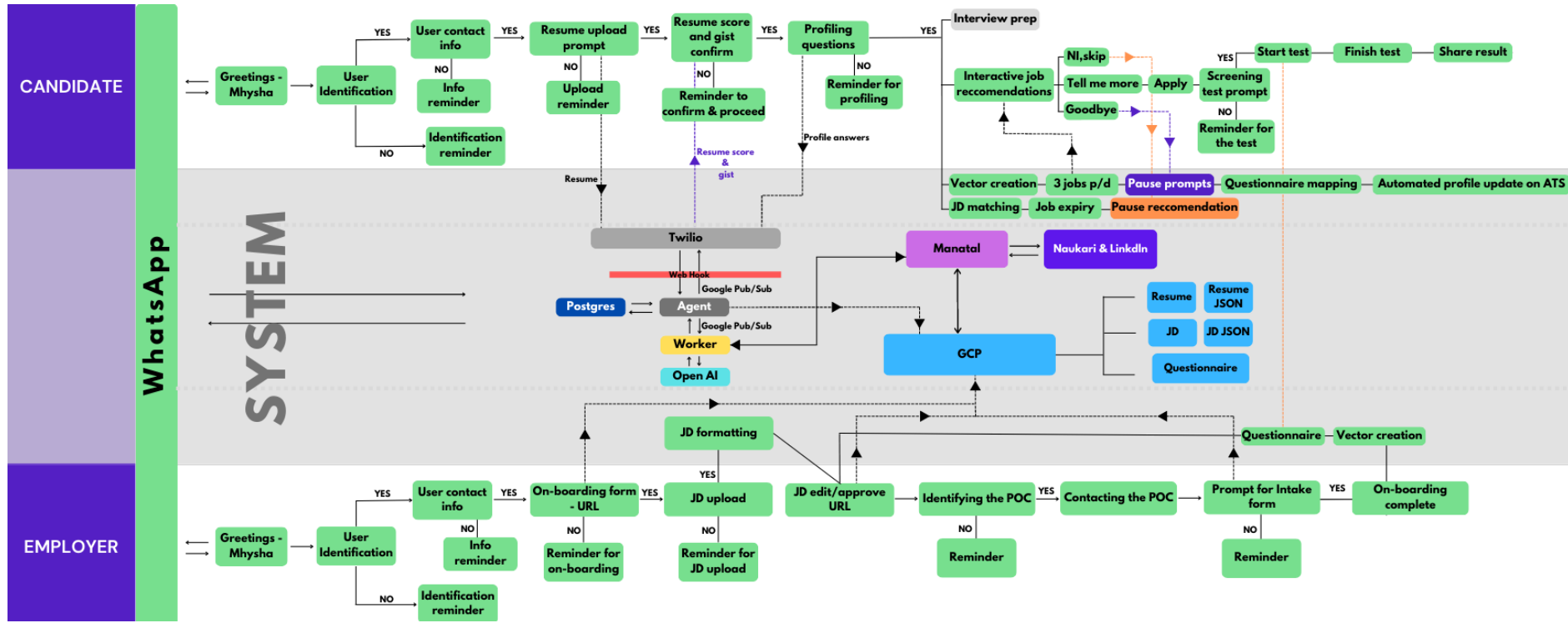
# Marketing Data: ingestion, transformation and visualization



# Palantir Data Harmonization



# Recruitment Analysis Architecture



# Data Science



---

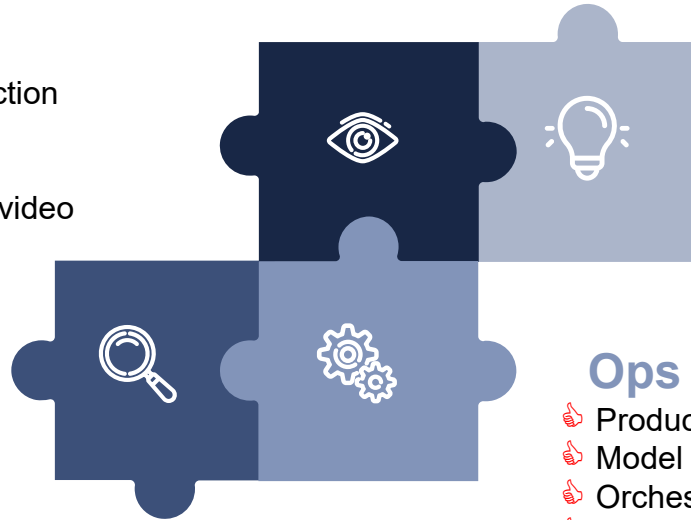
# Data science skills

## Vision analytics

- 👍 Object Localization and Detection
- 👍 Image Segmentation
- 👍 Face recognition
- 👍 Public Safety solutions using video analytics

## Text analytics

- 👍 Topic Modelling
- 👍 Text Mining
- 👍 Text /Document Clustering
- 👍 Named Entity Detection and linkage with knowledge base
- 👍 Sentiment Analytics



## AI/ ML

- 👍 Predictive Analytics
- 👍 Sales Forecasting
- 👍 Recommender system
- 👍 Anomalies detection
- 👍 Claim Fraud Detection
- 👍 Credit Risk Score
- 👍 Next Best Action
- 👍 Price Optimization

## Ops

- 👍 Productionize ML model
- 👍 Model Versioning
- 👍 Orchestration
- 👍 CI/CD
- 👍 Monitoring

# Data science domains

## HR Analytics

### ❑ Face Recognition

To recognize the employees via multiple CCTCs installed in the premises.

Computer Vision

Deep Learning

Low latency solution based on premise



## Public Safety

- ❑ Fence Trespassing
- ❑ Vehicle Over Speeding

An array of problem facilitating client to maintain order at sea port.

Computer Vision

Deep Learning

Analyzing 16 camera on cluster of machine



## Supply Chain

### ❑ Text Classification

To identify the item based on its textual description to provide the client an opportunity to automate the manual process intelligently.

NLP

Text Processing

Deep Learning

Used Azure ML and ML Ops as deployment platform.



# Data science domains

## Banking and Insurance

- ❑ Fraud Detection
- ❑ Recommendation

To detect the fraudulent transaction in insurance claims and recommend best possible policy which insurer can buy.

Machine Learning

Successfully detected ~90% frauds



## Consumer Electronics

- ❑ Smart Homes (Voice driven appliances)

To develop a system facilitating end user to control and monitor their appliances via their voice commands.

Speech Analysis

IoT

Dozens of products launched complying WWA standards



## Facility Management

- ❑ Smart Facilities

To generate reports powered by tons of IoT sensors installed in multiple premise, enabling client to measure occupancy level, service patterns etc.

IoT

Data platform

Reporting

Monitoring ~1200 sensors and still growing





# Thank you

## **Delhi-NCR, India**

C-1, Sector 7, Noida, UP 201301

Phone : +91-120-419-0000

## **Hyderabad, India**

Ground Floor, Plot No. 33, Silicon Valley Image Gardens Road, Madhapur,  
Hyderabad, Telangana 500081

Phone: +91-40-4853-2316

## **Dallas, USA**

4512 Legacy Drive, Suite 100 Plano, TX 75024

Phone : +1-512-772-3193

## **London, United Kingdom**

83 Victoria Street, London SW1H0HW

Phone : +44-207-096-9578

## **Toronto, Canada**

1 Dundas St W #2500, Toronto, ON M5G 1Z3

Phone : +1-416-619-9258

