

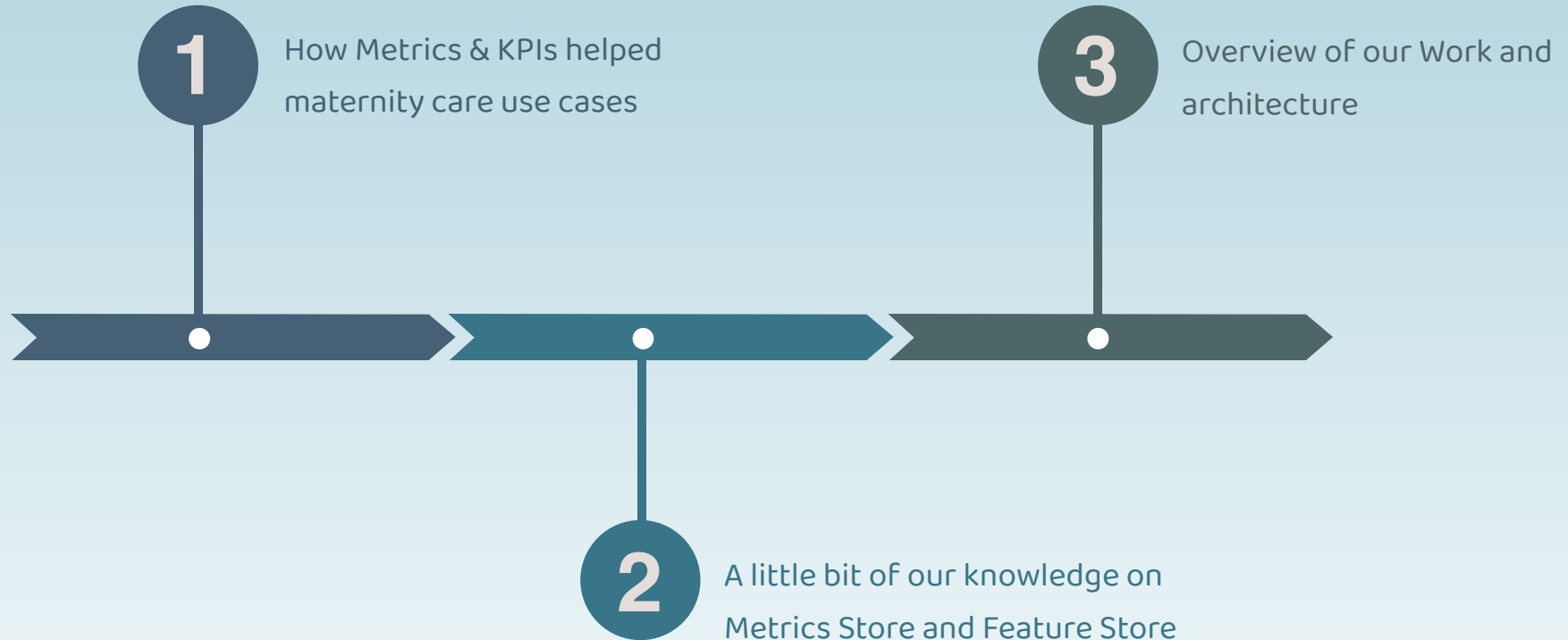


COMATRIX

**Using KPI & Feature Store
with spark for self served
data sharing and analysis**



Summary of this talk



Our Team



Sameer M

Co-Founder, experienced data practitioner

Education: Stanford, Rutgers

Apple: Analytics for business integrations

Kohl's online: Recommendation engine

Twilio: Customer engagement marketing platform

Published: Multiple papers on methodologies of

Privacy-preserving techniques

5 Data Engineers



Neeru P

Co-Founder, Associate Professor

Education: Doctorate in marketing from
Harvard Business School.

Georgetown University McDonough School of
Business: Marketing, Privacy, Consumer
behavior, Sustainability

2 ML Engineers



Sasidhar K

Founding Member, entrepreneur

10+ years of experience running
enterprises, advising startups, and
organizing first customer sales.

2 Software Engineers

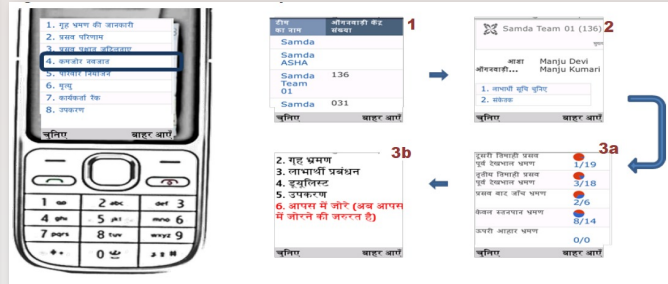
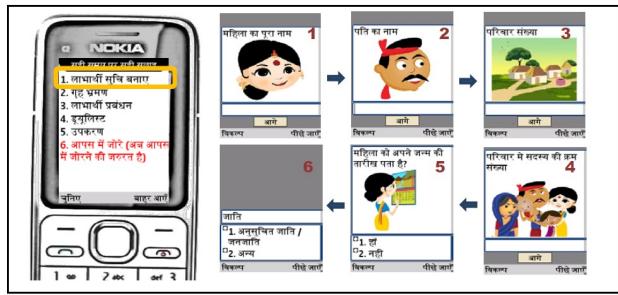
01

Why did we start building KPIs and Features?

Our customer health care service provider wanted to use data from 600 health care centers for better maternity care in low-income towns.



Data collected at various care centers



MTP/Abortion, Migration, Death and Other Events

Growth Monitoring

Family Planning

- 9 Months -6 -3 0 Birth +6 +12 +24 +3 Years +6 Years

Birth Preparedness

PNC

Exclusive Breast Feeding

Complementary Feeding

Preschool and Spot Feeding

Mother Nutrition Take-Home Ration

Child Nutrition Take-Home Ration

Immunization

Angawaadi Centre Management



Saurbazar block, Saharsa district

Benefits of the survey for improved maternity

- KPIs of healthcare services to healthcare improvements for low-income maternity care
- Childbirth disease anomalies based on demographics for a given timeline
- Improve maternity care based on vaccination rates, and health data in low-income neighborhoods

02

Metrics Store + Feature Stores

Metrics Stores are like the maestro, Feature Stores are like the conductor of an orchestra and you want good synchronization between maestro and conductor for better music



What are Metrics Stores and Feature Stores

Metrics and KPIs are the "shared language" for organizations in today's world to make data-driven decisions

- ❖ Data measures, frequencies aggregated over time
- ❖ Standardize goals across organizations
- ❖ Used for measuring success or failure of products, services, revenue as well as companies

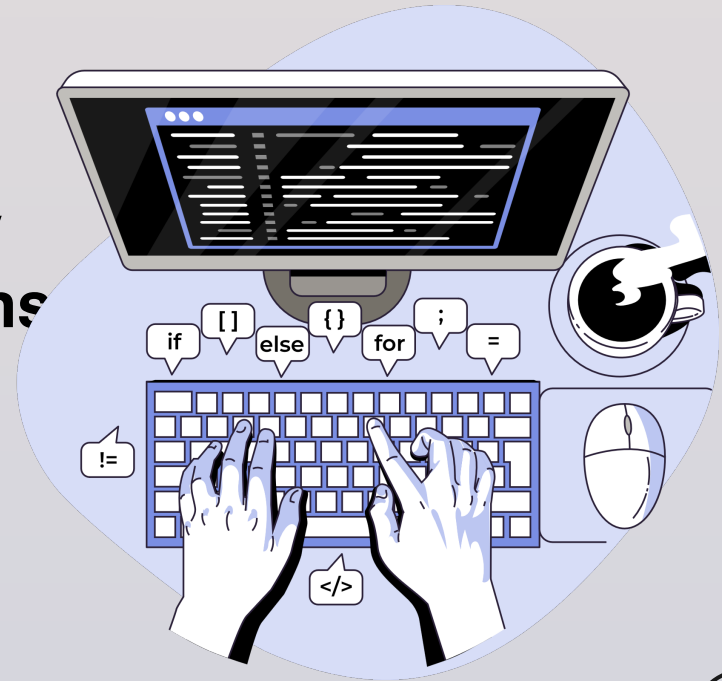
Features are the fuel of our machine learning models that are often used to provide value added services to customers

- ❖ Represent raw data from the real world
- ❖ A numerical representation of metrics and KPIs
- ❖ To understand customer behavior, improve products and services

Metrics, Features & KPIs are everywhere



Rise in governance, data security challenges and privacy regulations is making the data platform infrastructure increasingly complex



Challenges to Metrics, KPIs and Features

Organization Challenges

- ❖ Data Strategies are hard
- ❖ Multiple stakeholders with slightly different results
- ❖ Data sharing is a challenge

Technical Challenges

- ❖ Governance, Compliance, Security and Privacy
 - Sensitive data is everywhere
 - Complex Data tooling ecosystem
 - 9 Copies on average of the same data
 - 70% of data is unstructured
- ❖ Data silos
- ❖ Complex tooling

03

Collaborative Information Sharing approach

For small-to-medium enterprises, Information Sharing is a valid instrument to improve their key performance indexes (KPI) for improving revenues, helping with better patient care and many more use cases.



Objectives for Metrics and KPIs Infrastructure

Business analysts, data scientists, Insight apps and Reporting need access to consistent metrics

- ❖ Secure, Compliant and Privacy Preserving
- ❖ Data Producers are responsible for data
- ❖ A common infrastructure that security is given
- ❖ Based on the principles of data mesh
- ❖ Unified data access patterns



Features of CoMatrixLabs KPI & Feature Platform

Discovery

Data sharing Agreements

KPIs and Metrics

Automated Analysis



- ❖ Sensitive data discovery using sampling
- ❖ Semi Structured data support
- ❖ Rest API support
- ❖ SQL Server

- ❖ PII Classification
- ❖ Policy & Rule Engine
- ❖ Data Sharing Agreements

- ❖ Counts, Measures and Top counts
- ❖ Cubes and rollups
- ❖ Preview, metadata & recommendations for data cookbooks

- ❖ Comparison Analysis
- ❖ Basket Analysis
- ❖ Engagement Analysis
- ❖ Clustering, Segmentation & classification analysis
- ❖ Automated Record of activity processing



Our approach for collaborative information sharing

A multi-party collaborative analysis and information sharing infrastructure for producers of data to share privacy-preserving KPIs and Features for Machine learning use cases

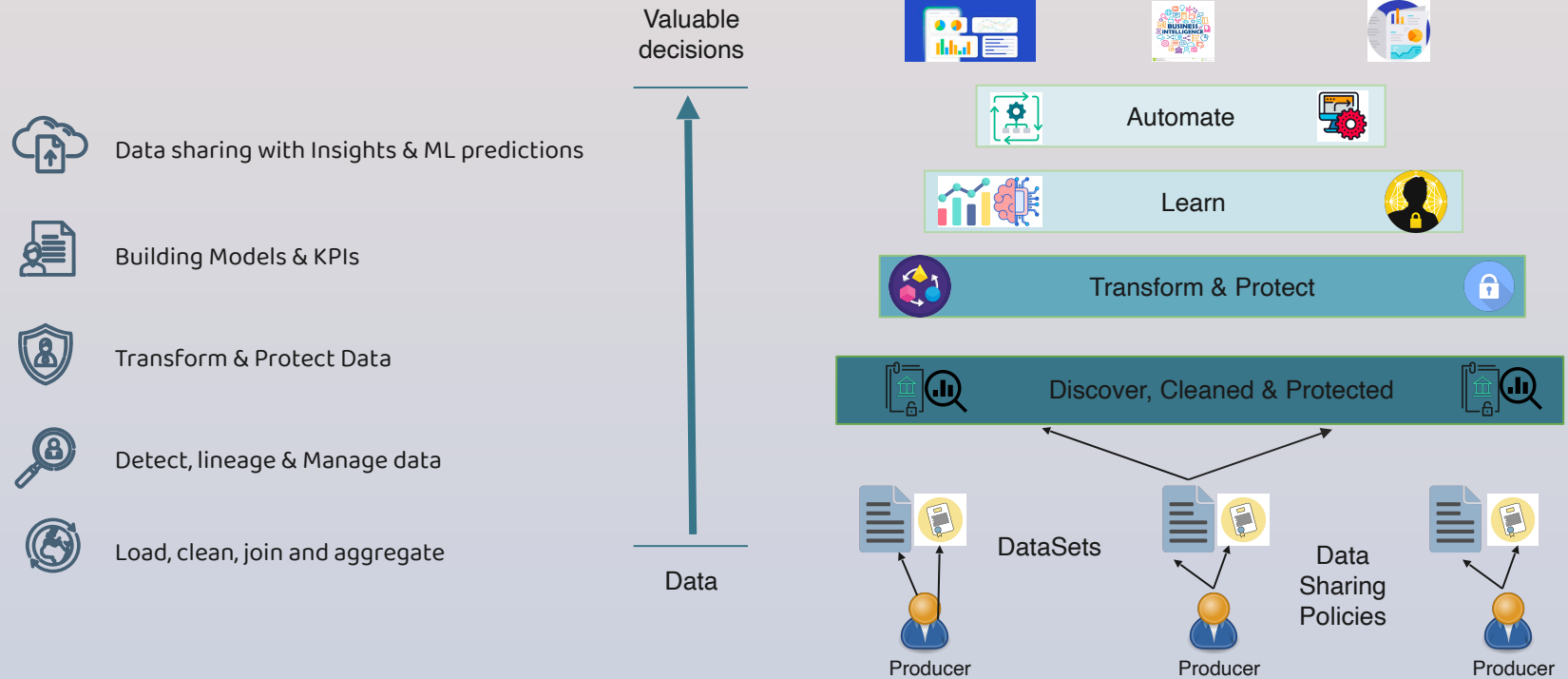
A Data Sharing Cookbook to automate the privacy-preserving, compliant KPIs and Metrics



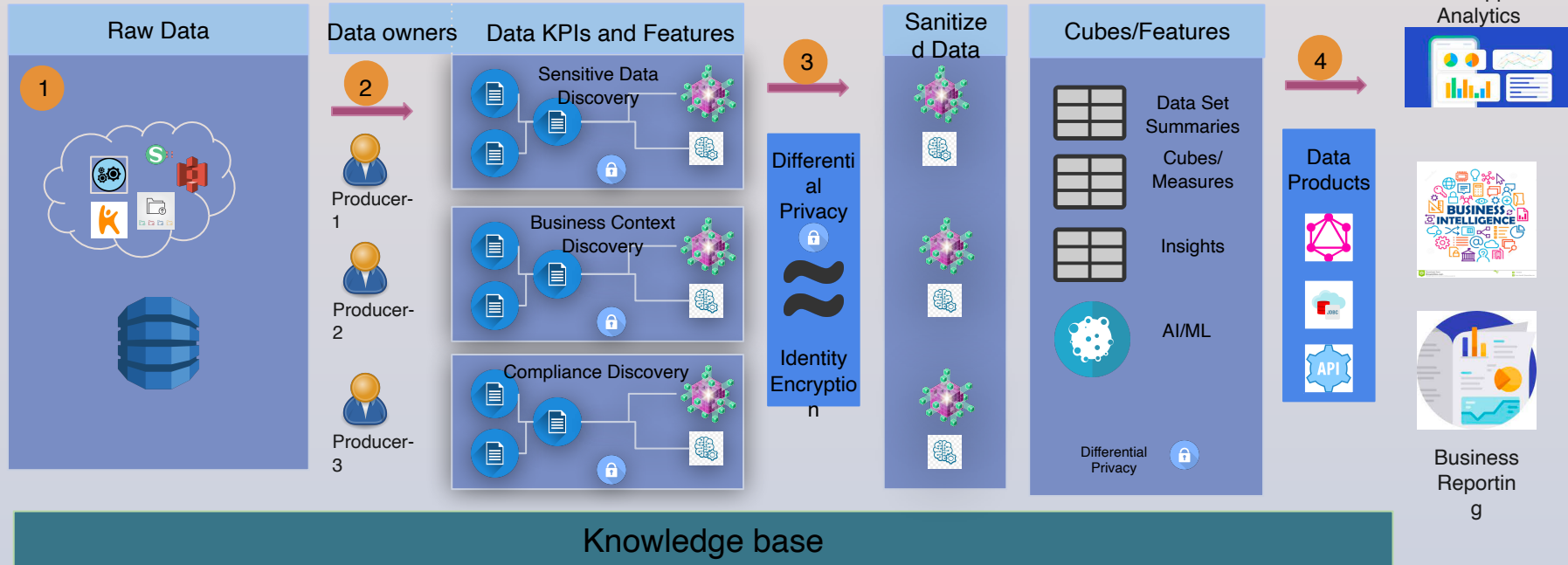
Continuously identify and classify data targets to provide business context

A framework of data sharing agreements using policy-based mechanisms

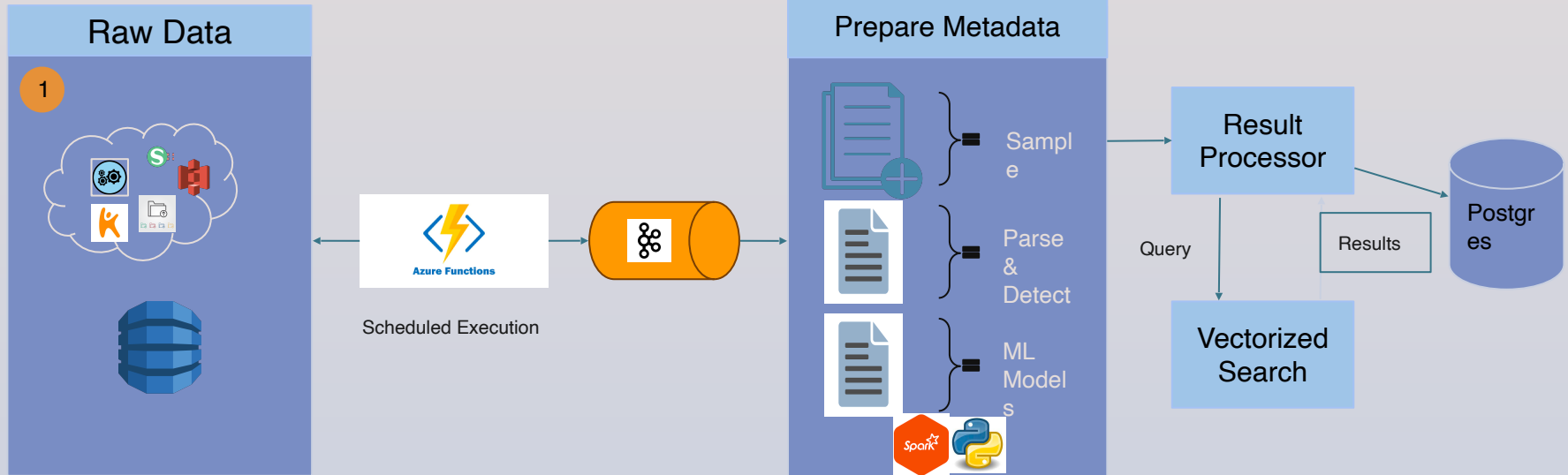
From Data to Data driven insights



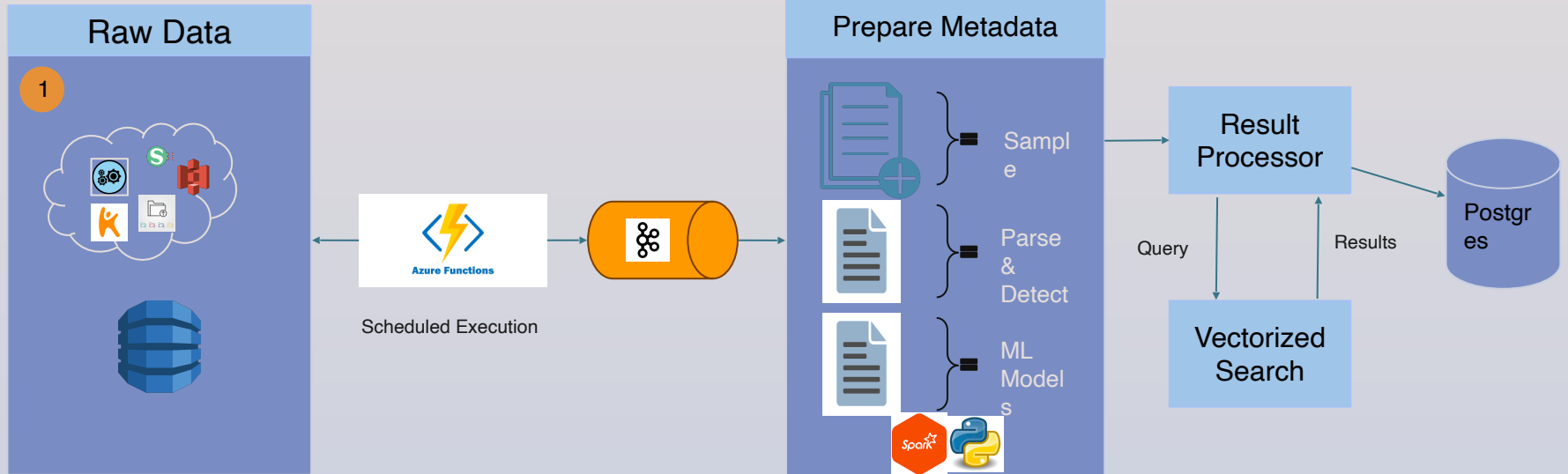
Raw data to Privacy & Security Preserving Insights



Data Discovery Architecture

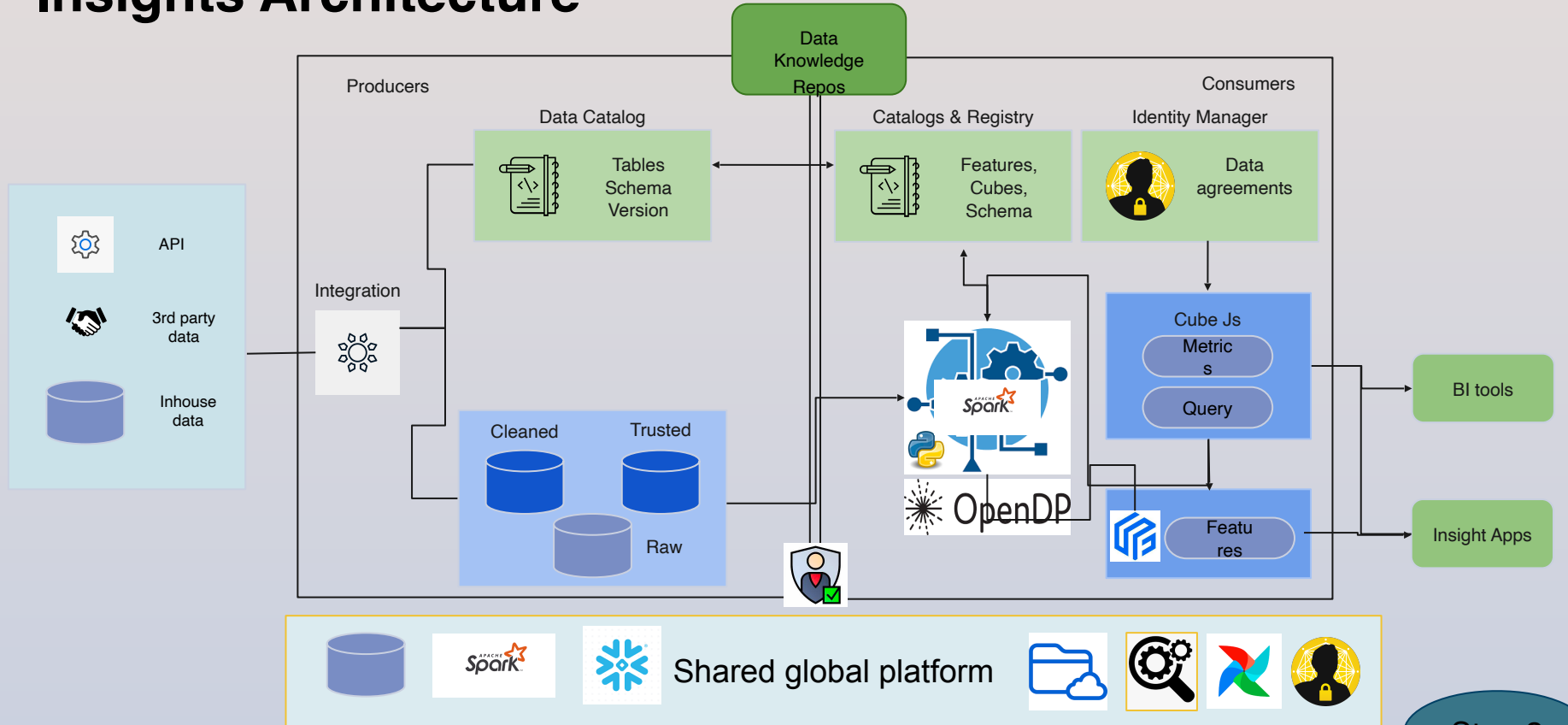


Data Discovery Architecture



Step 1

Insights Architecture



A 2 Min Demo

A secure, compliant, privacy aware Metrics,
KPIs and Feature stores can help enterprises
conduct data driven decisions like a well
organized orchestra

Thank you

References

- Opendp smart noise
- Airbnb Guo & Capital One data profiler
- Privacy preserving data sharing and analysis for edge-based architectures by “Mina Sheikhalishahi¹ · Andrea Saracino² · Fabio Martinelli² · Antonio La Marra³”
- Many more....



Thanks!

Do you have any questions?

info@comatrixlabs.io

www.comatrixlabs.io

