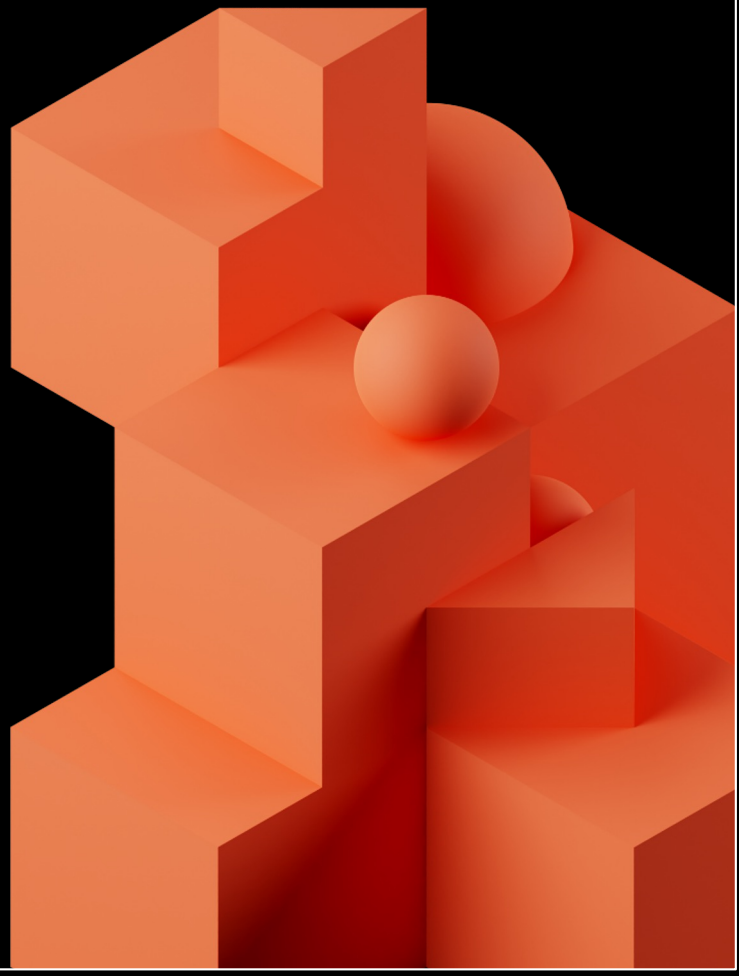
DATA+AI SUMMIT
BY databricks

Data Democratization with Lakehouse

An Open Banking application case



Databricks
2023

Who we are

Msc Pedro Antonio Boareto
Data Engineer

LinkedIn:

Fábio Luis Correia da Silva
Architecture Specialist

LinkedIn:

History

About us

- Founded in 1943
- Focus on small traders, civil servants and people with modest means
- Became the fifth private bank in Brazil in just eight years

Our first branch (Marilia, São Paulo, Brazil)

History

Innovation is our business

In 1970s Bradesco was the first company to have a computer in Latin America.

Buying the Bank's first computer (1970s)

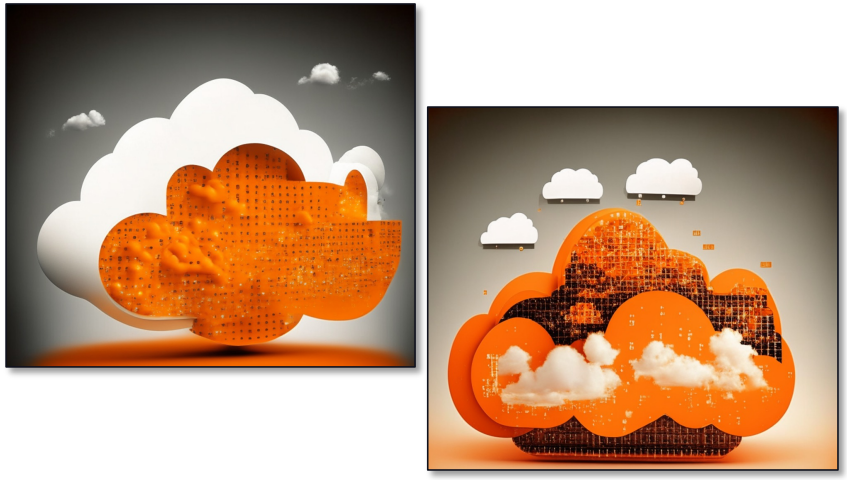
Digitalization culture slogan (1990s)

In 1990s Bradesco expanded the online services, mobile banking and ATMs all over Brazil

History

Now

We are living in a big data and cloud revolution

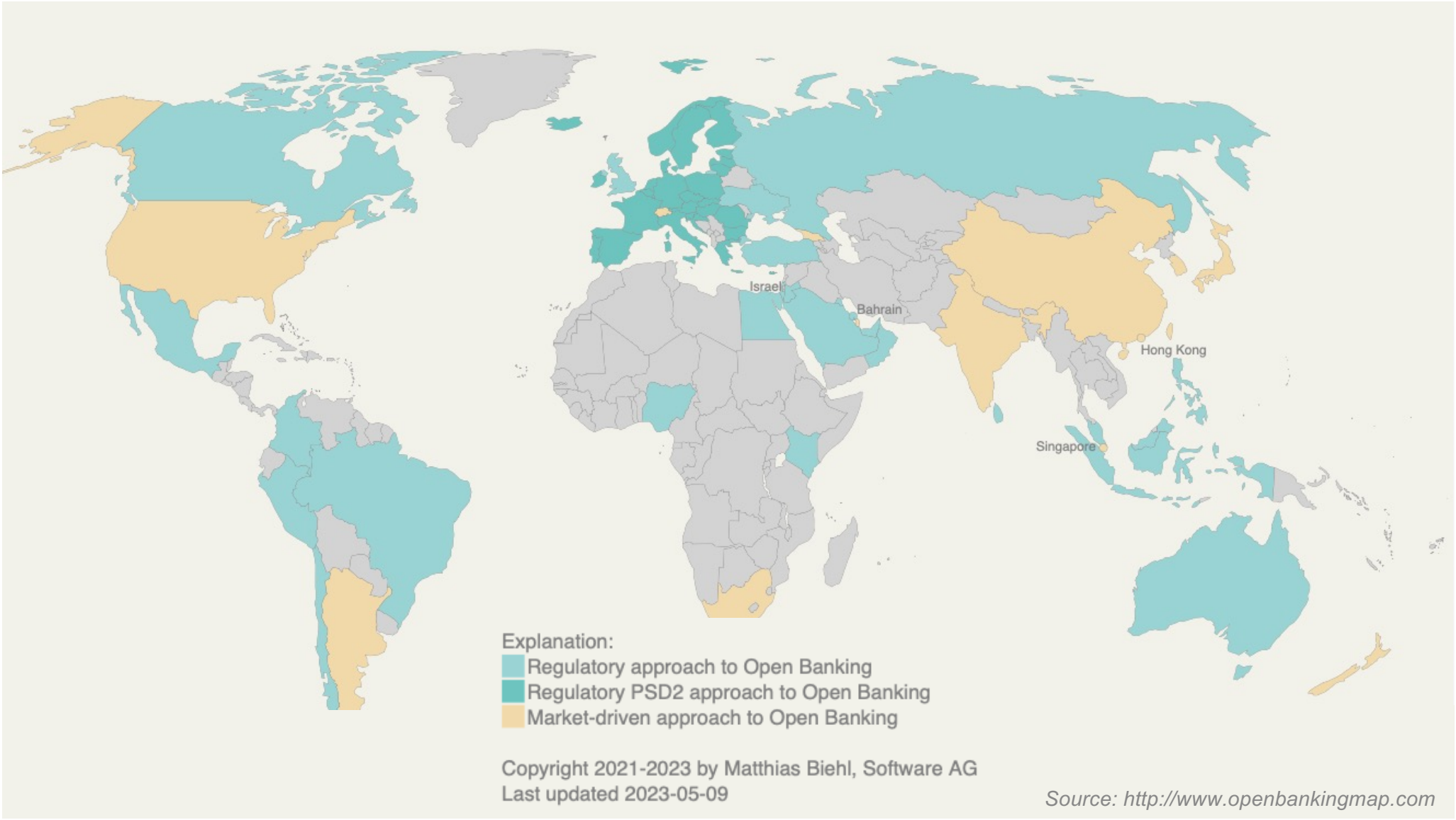


AI generated images

OPEN BANKING BRASIL

Open Banking – Our challenge

Open Banking lies on balancing security, privacy, data sharing, competition, and innovation in the financial industry.


Israel
Bahrain
Hong Kong
Singapore
Explanation:
Regulatory approach to Open Banking
Regulatory PSD2 approach to Open Banking
Market-driven approach to Open Banking
Copyright 2021-2023 by Matthias Biehl, Software AG
Last updated 2023-05-09
Source: http://www.openbankingmap.com

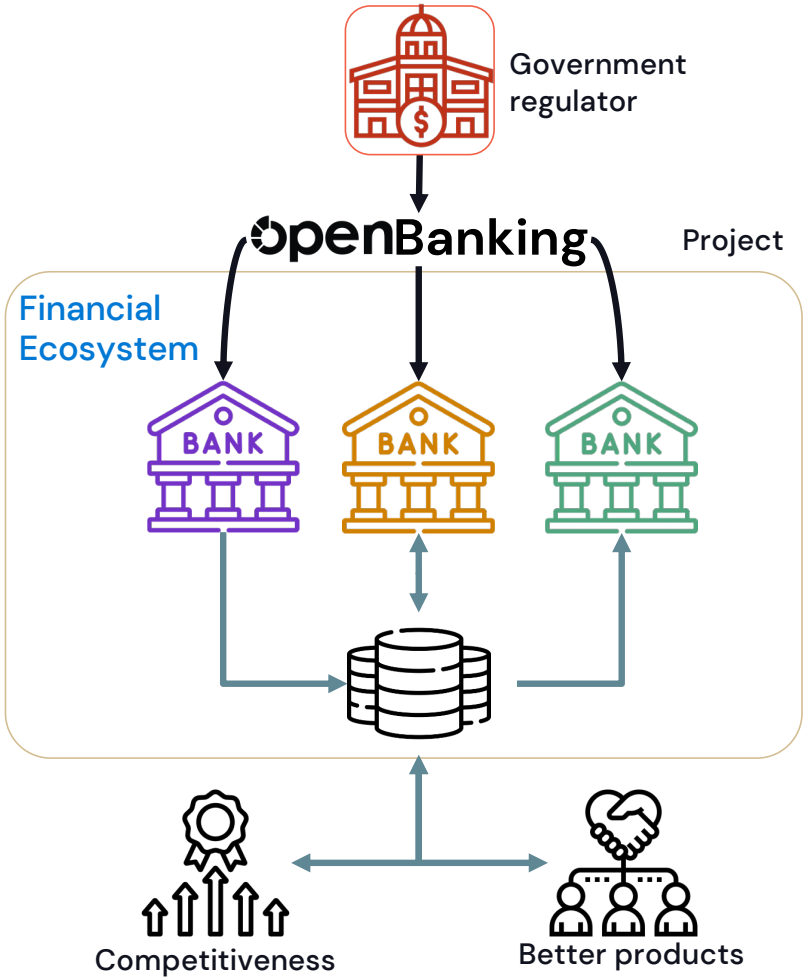
OpenBanking

A set of rules that enable the sharing of customer data and services between financial institutions through the integration of their systems

OpenBanking

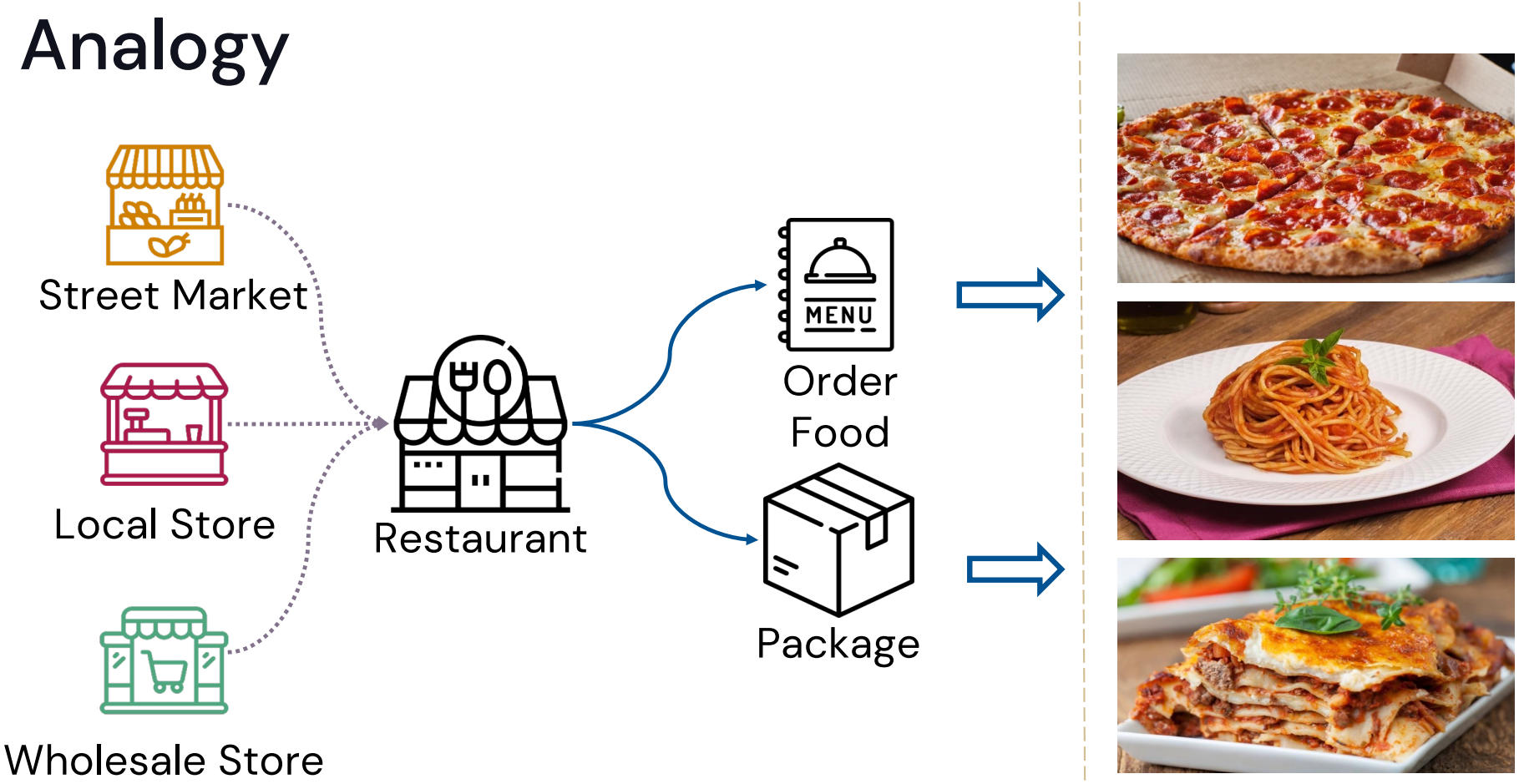
What is the Open Banking case!

- Known in Brazil as “Open Finance”
- Mandatory requirement of the government financial regulator
- Improves financial competitiveness
- Incentivizes companies to develop better products




WHAT

Analogy


Street Market
Local Store
Wholesale Store
Restaurant
MENU
Order Food
Package

Phases of implementation

Phase 1 – Public Data (Feb/21)

- Service channels
- Fees and taxes

Phase 2 - Customer Data (Aug/21)

- Personal
- Banking account
- Credit Cards
- Others...

Phase 3 - Instant Payments Data (Oct/22)

- PIX (same as Fed Flow)

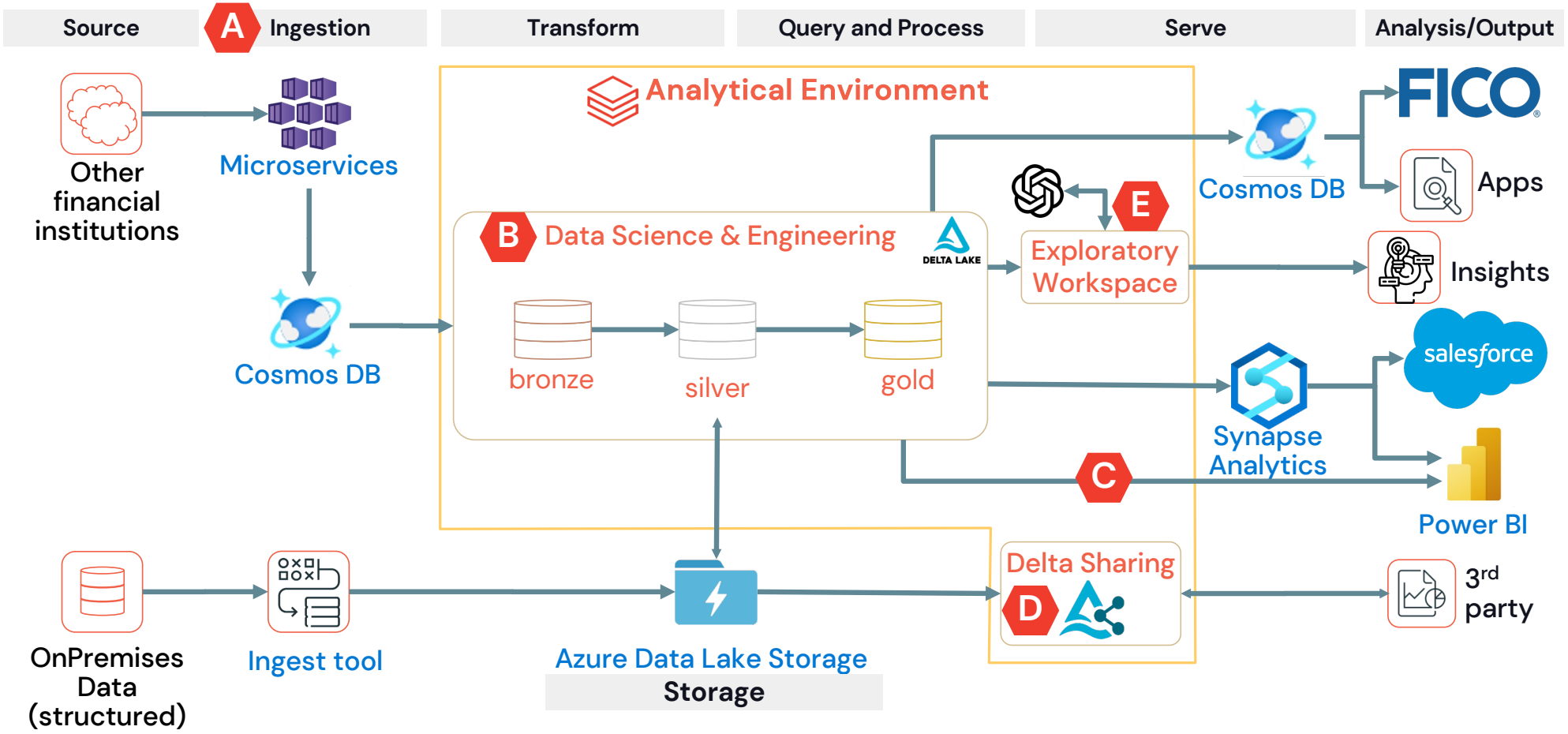
Phase 4 - Investment Data (Mar/23)

- Exchange
- Insurance
- Pension
- Investment

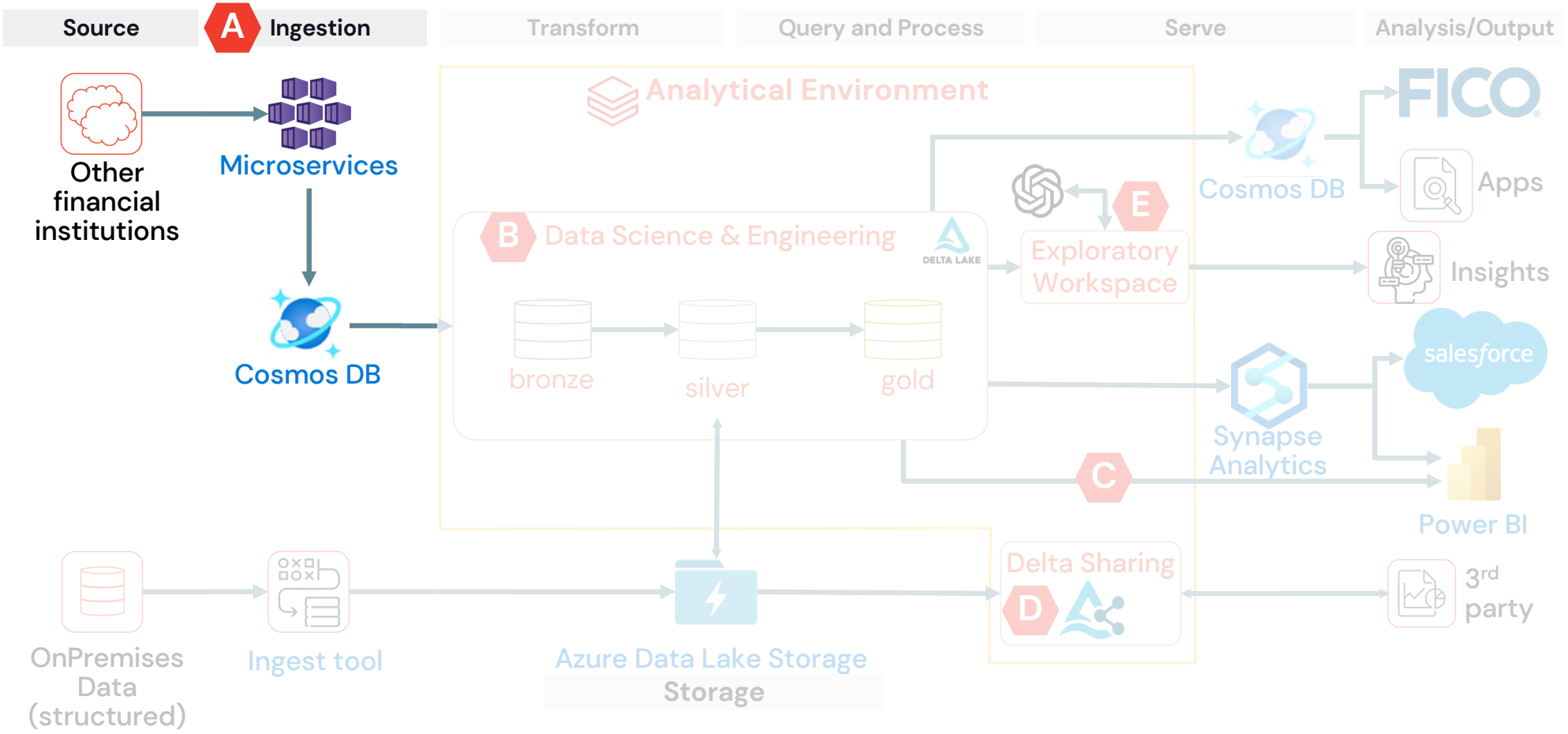

HOW

DATA DEMOCRATIZATION

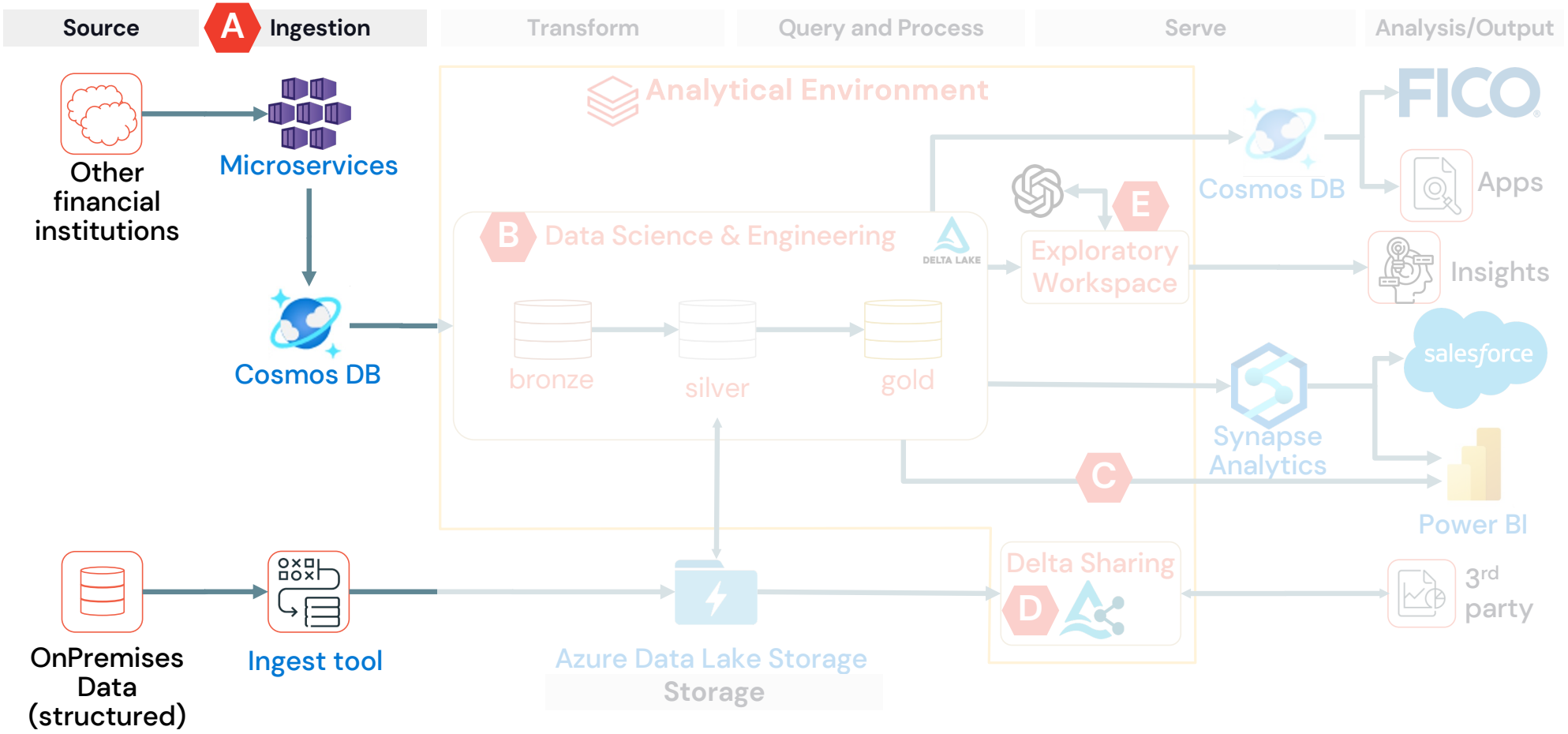
Lakehouse Open Banking Architecture



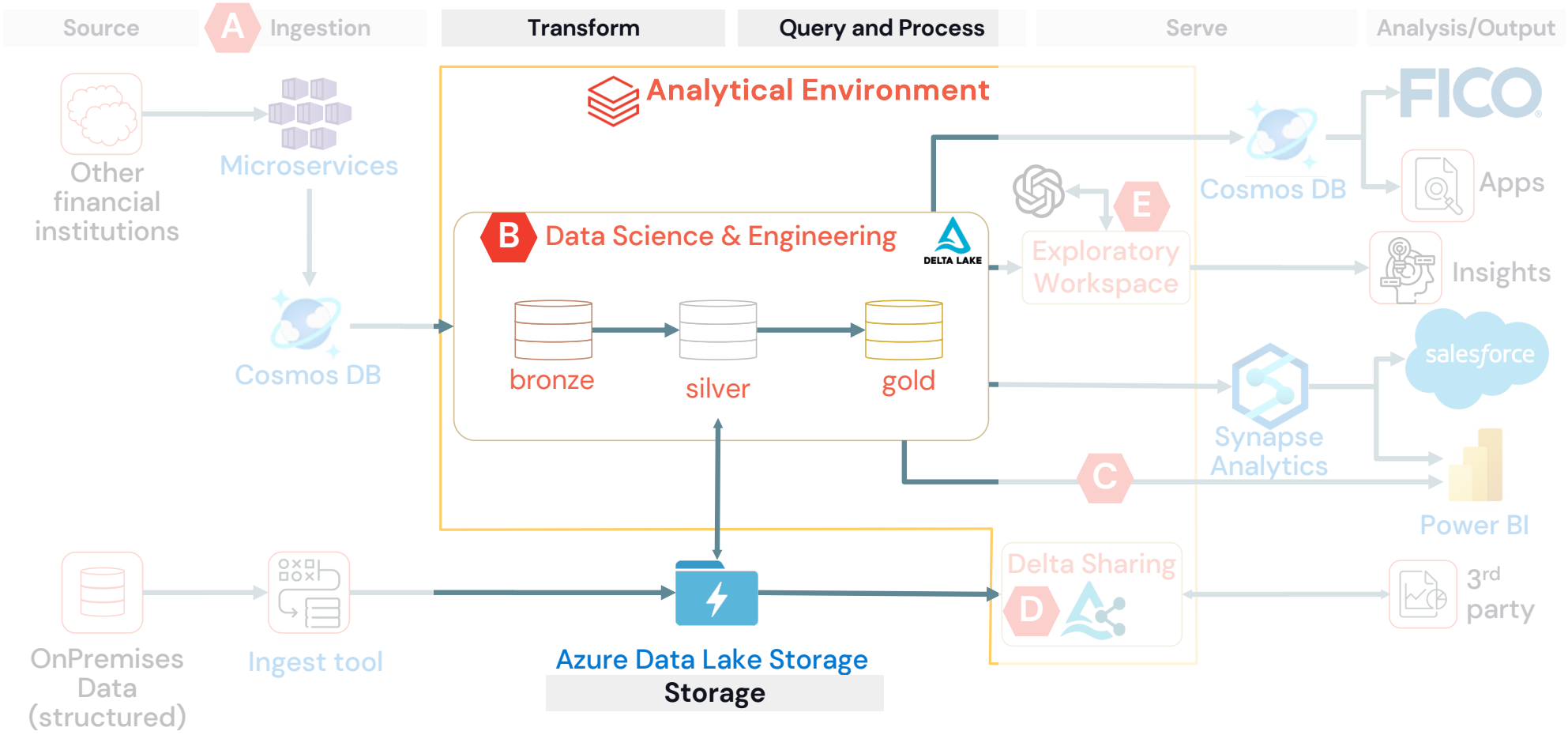
Lakehouse Open Banking Architecture



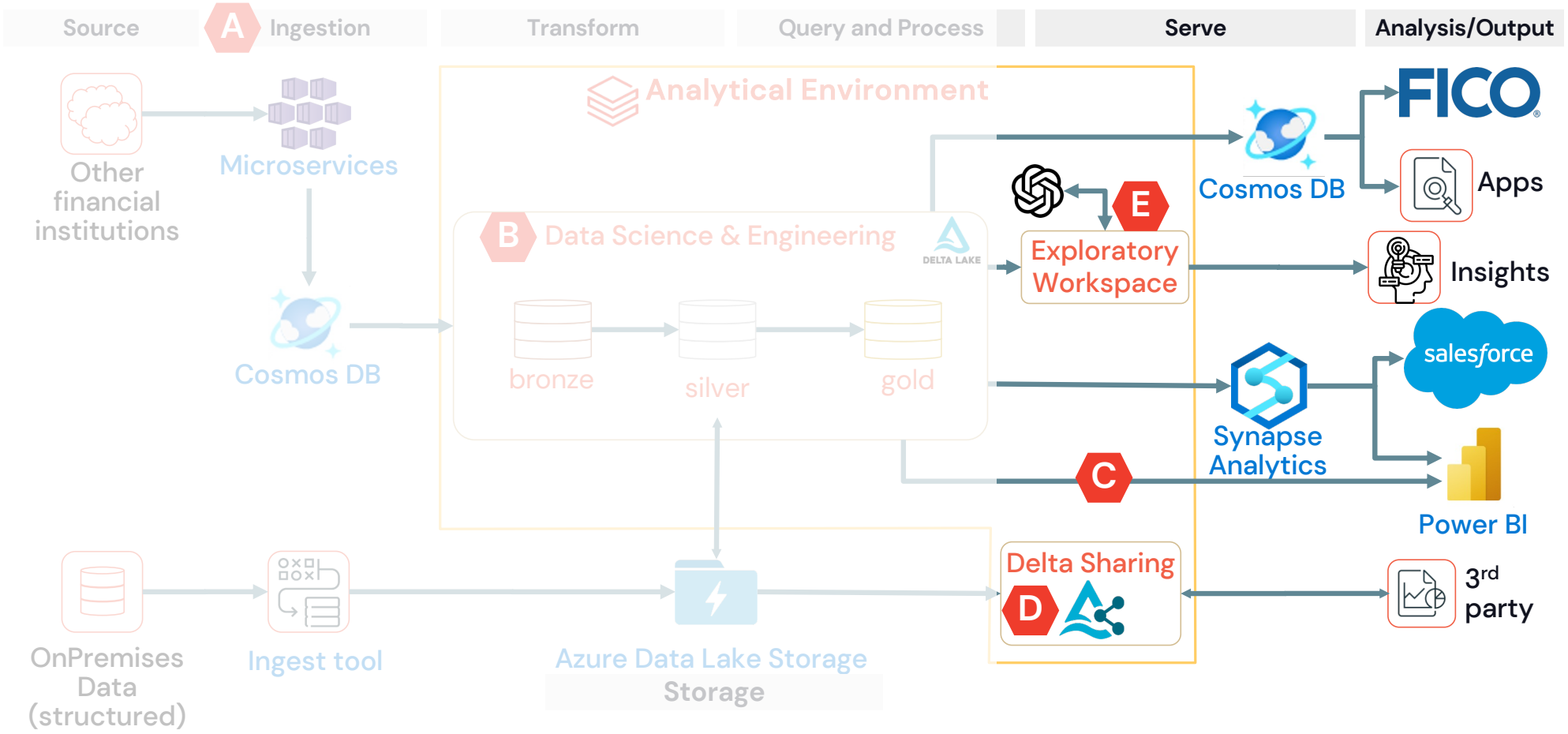
Lakehouse Open Banking Architecture



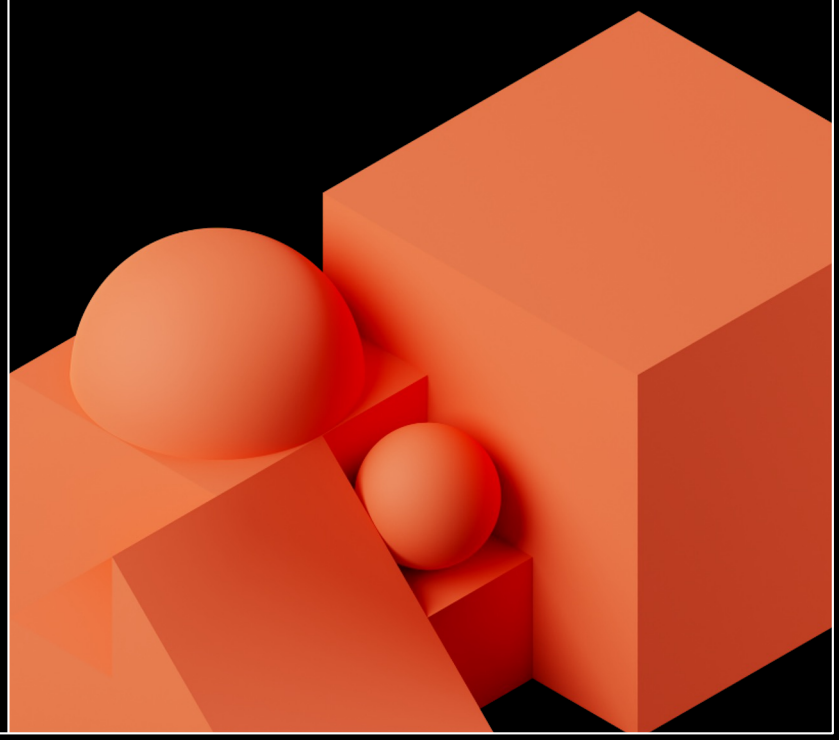
Lakehouse Open Banking Architecture



Lakehouse Open Banking Architecture



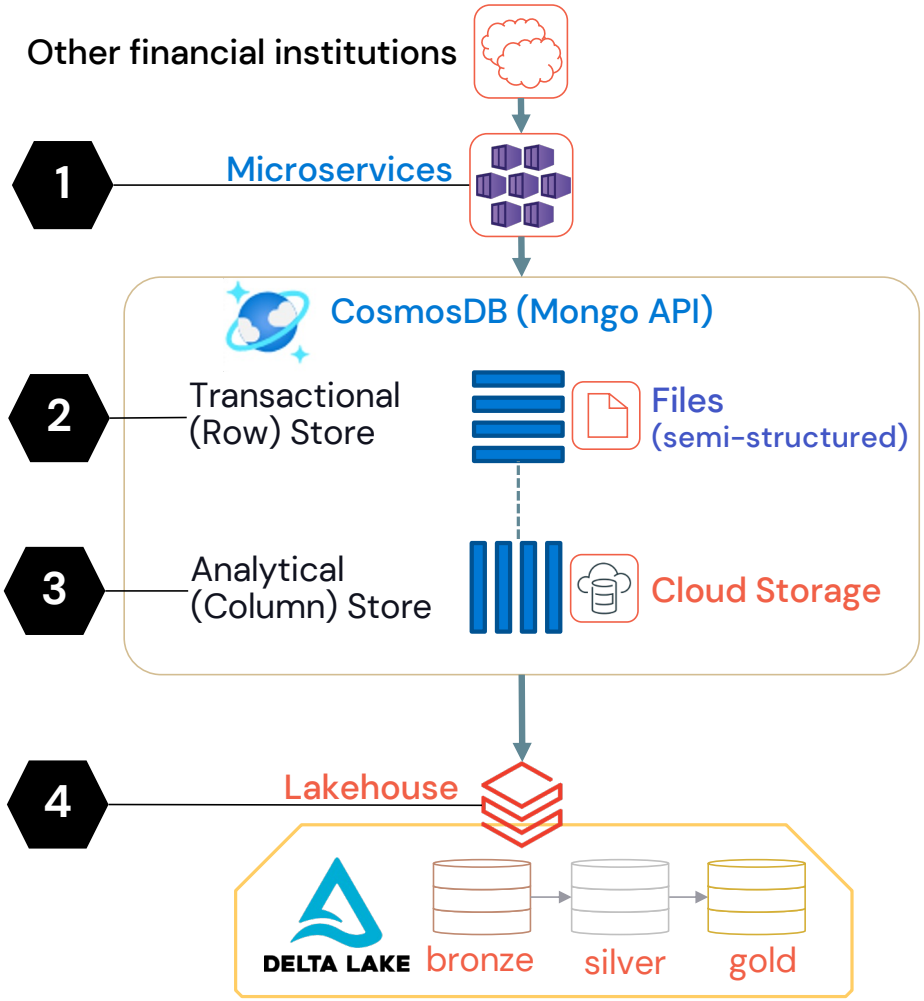
Data Ingestion



A Ingestion

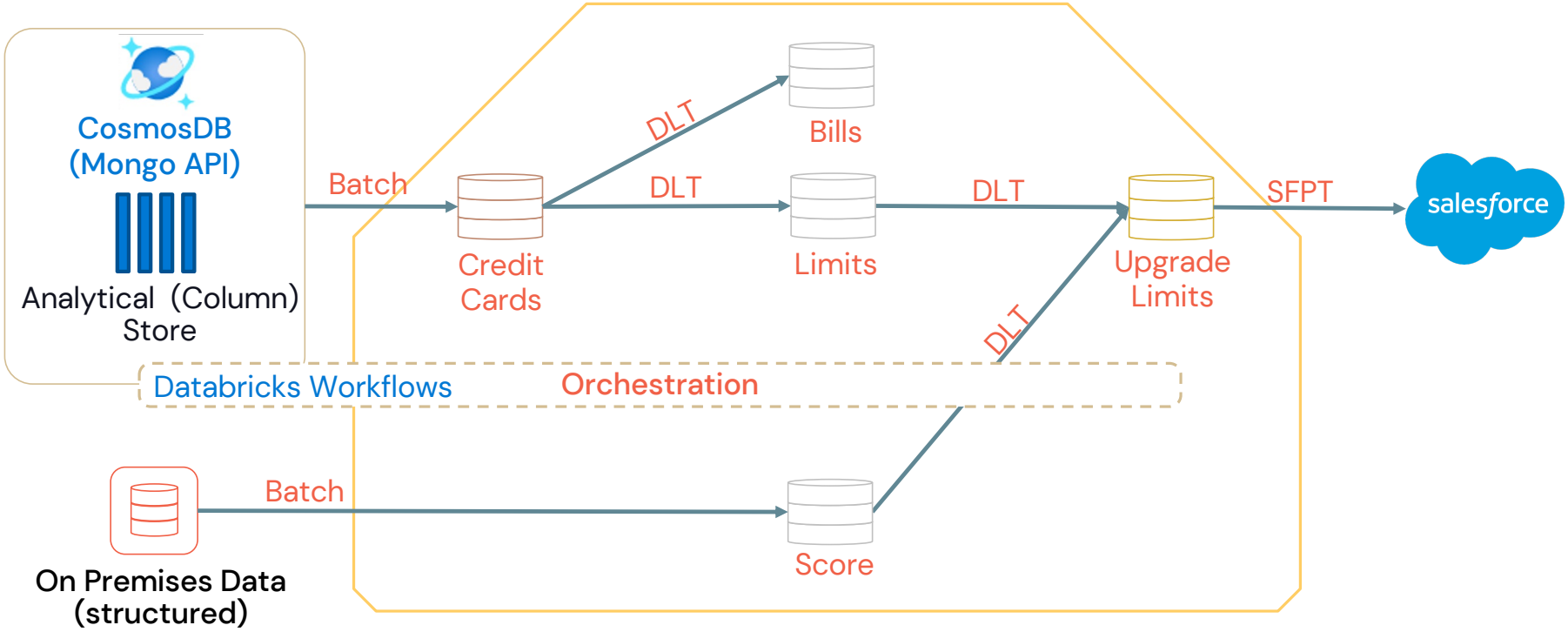
From data source to analytical

1. Data is received through APIs
2. Data is kept in a CosmosDB (.json file)
3. Enables reading data without competing with transactional
4. Json structure is transformed into relational data



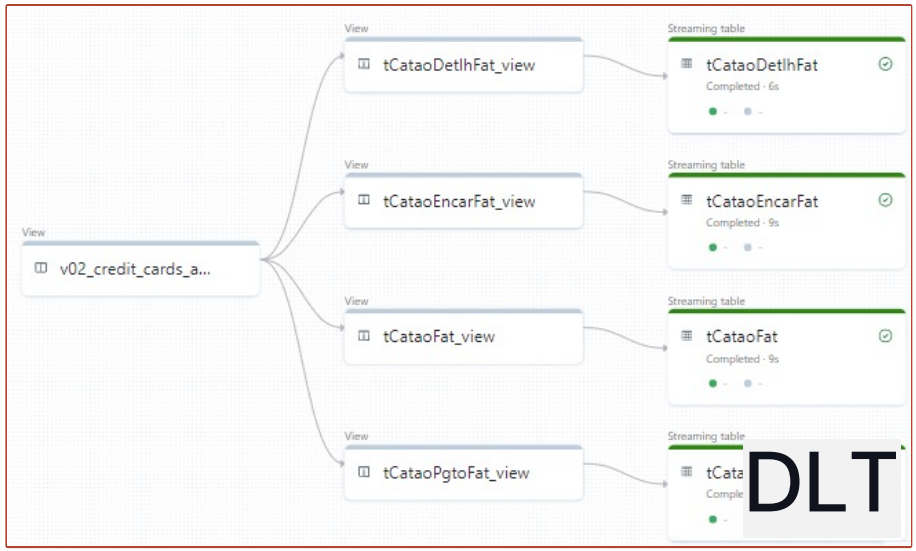
Example

Delivering credit card limits to Bradesco customers



id	loggedUserRequest
Nj	{"object
	{"identif
	{"string'
	{"string'
	{"branch
	{"string'
	{"string'
Nj	{"object
	{"identif
	{"string'
	{"string'
	{"branch
	{"string'
	{"string'
Nj	{"object
	{"identif
	{"string'
	{"string'
	{"branch
	{"string'
	{"string'
Nj	{"object
Nj	{"object
Nj	{"object
Nj	
Nj	

Bronze (Json)



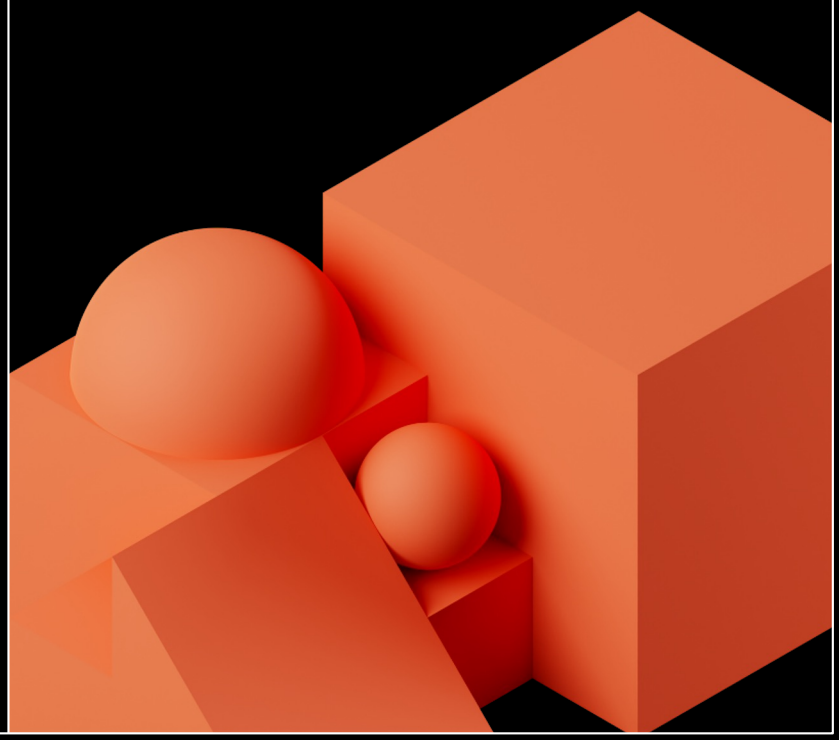

```
dlt.create_streaming_live_table(name="tcontoreceb", comment = "Tabel
partition_cols = ["dCriacConto"])

dlt.expect("pessoa_juridica", "cCnpjEmprContoReceb IS NOT NULL")
dlt.expect_all_or_drop({"valid_id": "cIdtfdContoReceb IS NOT NULL",

dlt.apply_changes(
  target = "tcontoreceb",
  source = "tcontoreceb_view",
  keys = ["cIdtfdContoReceb"],
  column_list = sorted_consent_columns,
  sequence_by = col("hIgtaoAnlStg"),
  track_history_column_list = track_history_colum
  stored_as_scd_type = "2"
)
```

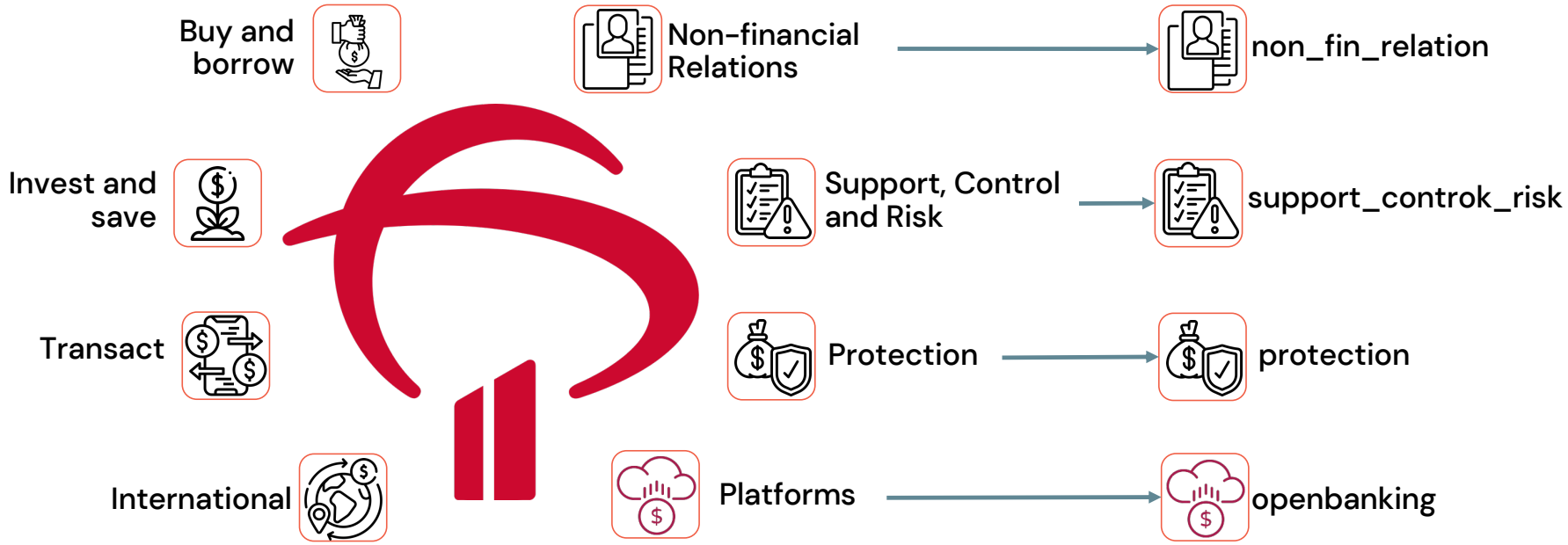
SCD2

Democratization in DATA MODELLING



Data Domains

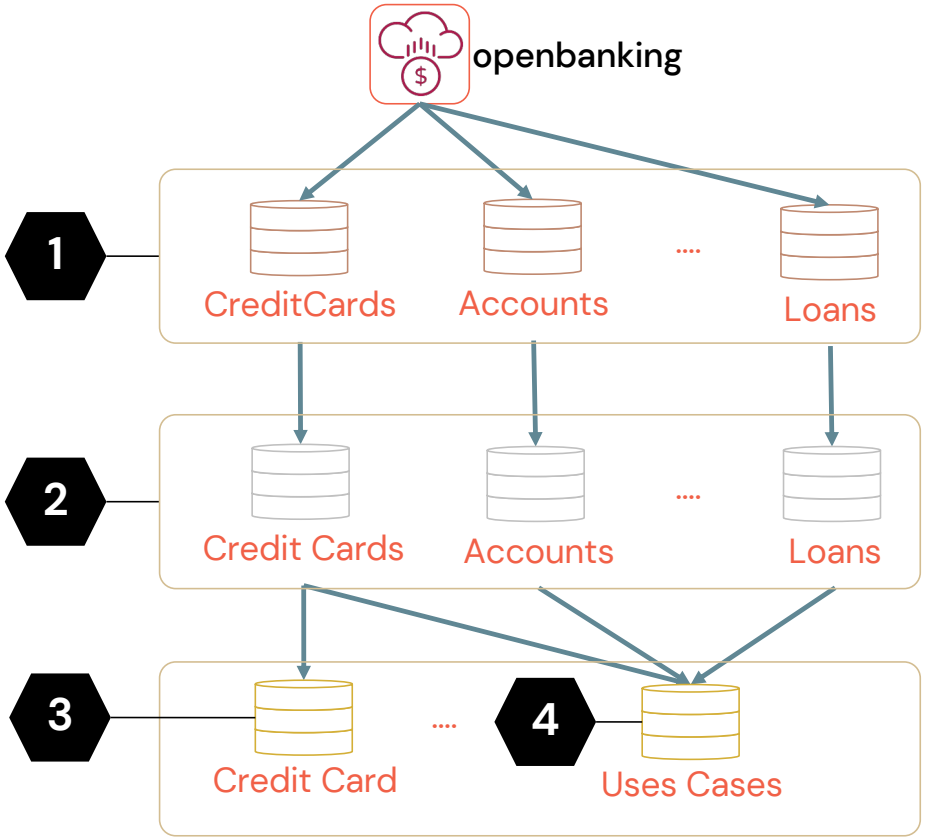
Unity Catalogs



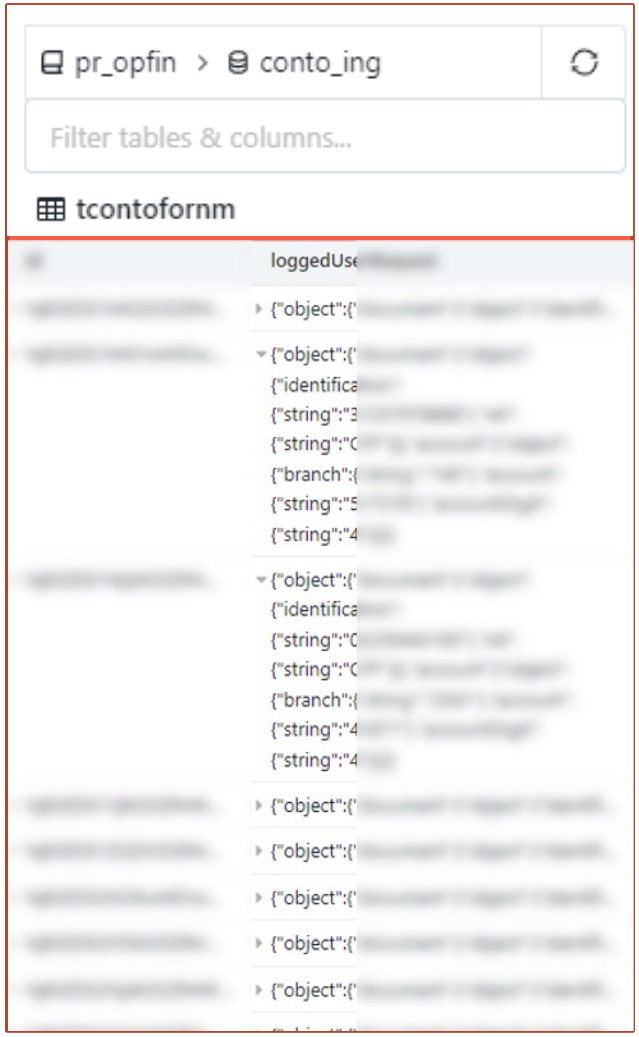
Open Banking Catalog

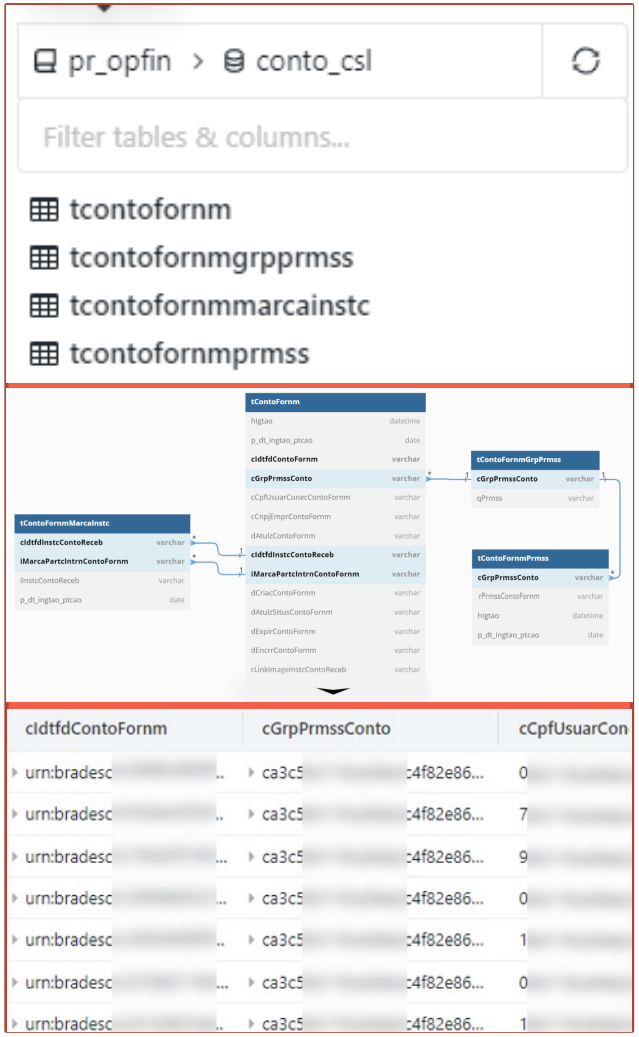
Highlights

1. Raw struct with .json files inside a column
2. Data is modelled into a relational model
3. Aggregated data
4. Use cases (cross) for specific cases.



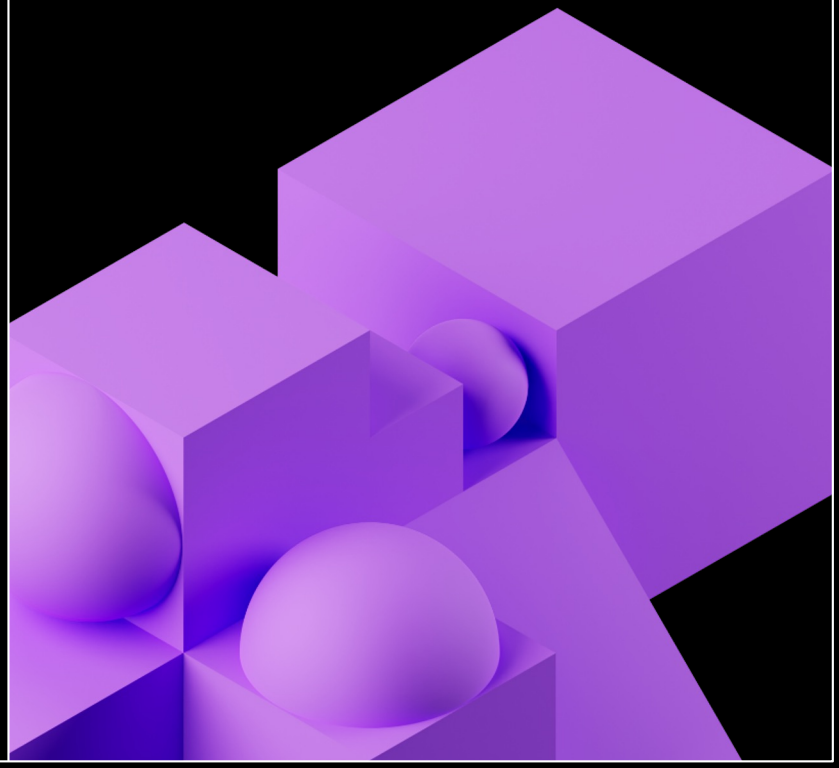
Bronze





Silver

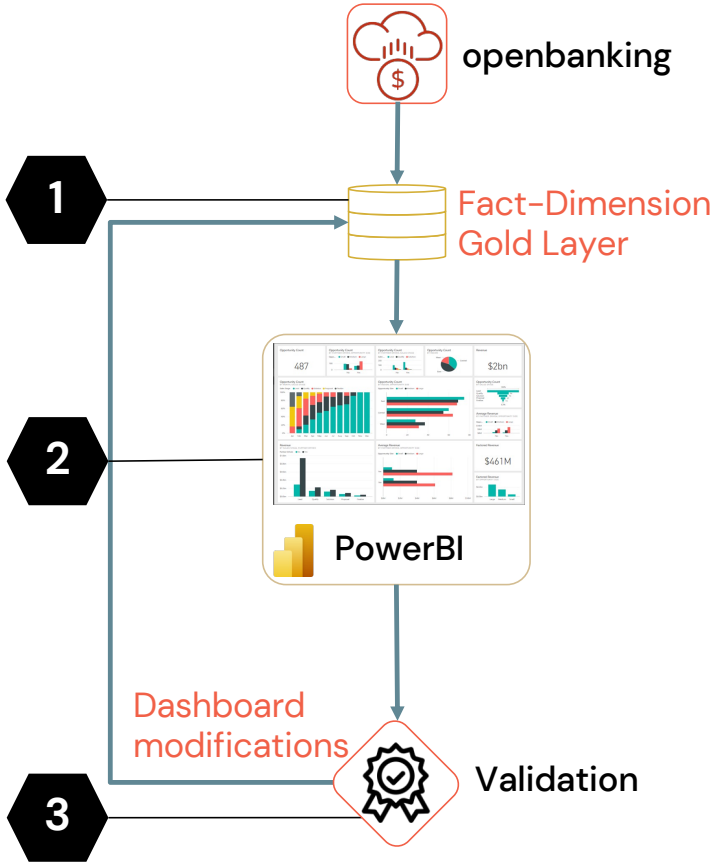
Democratization in DATA VISION



C Building dashboards

Our process before

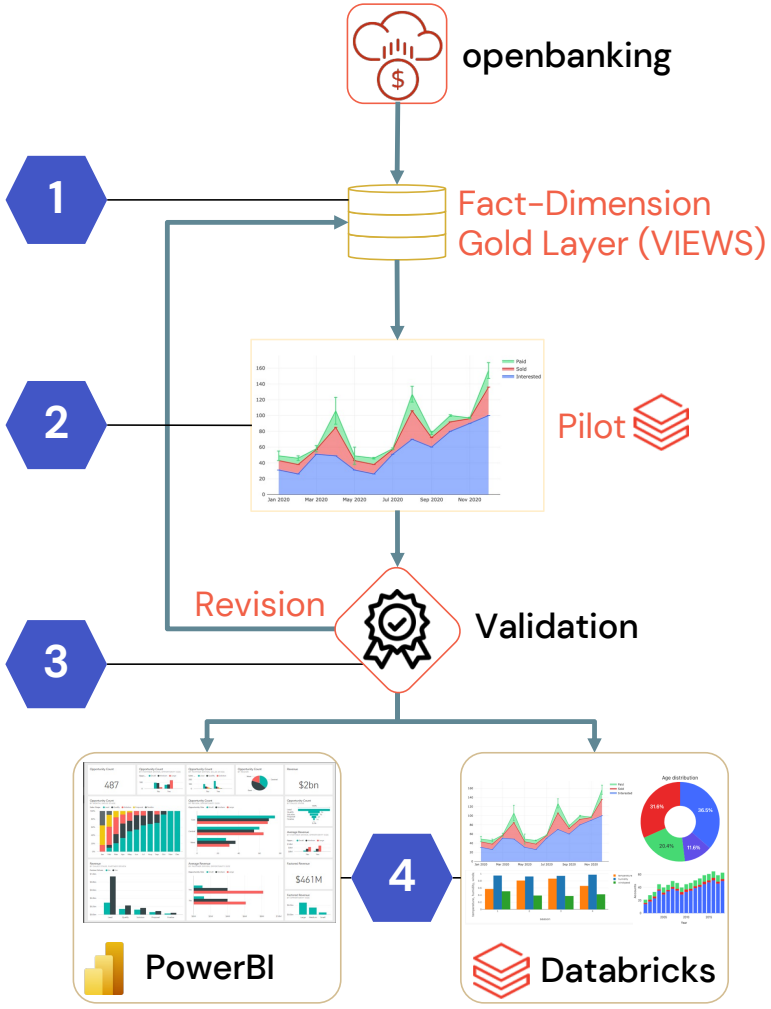
1. Creation of gold tables with aggregated data
2. Dashboard experimentation with PowerBI
3. Dashboard validation



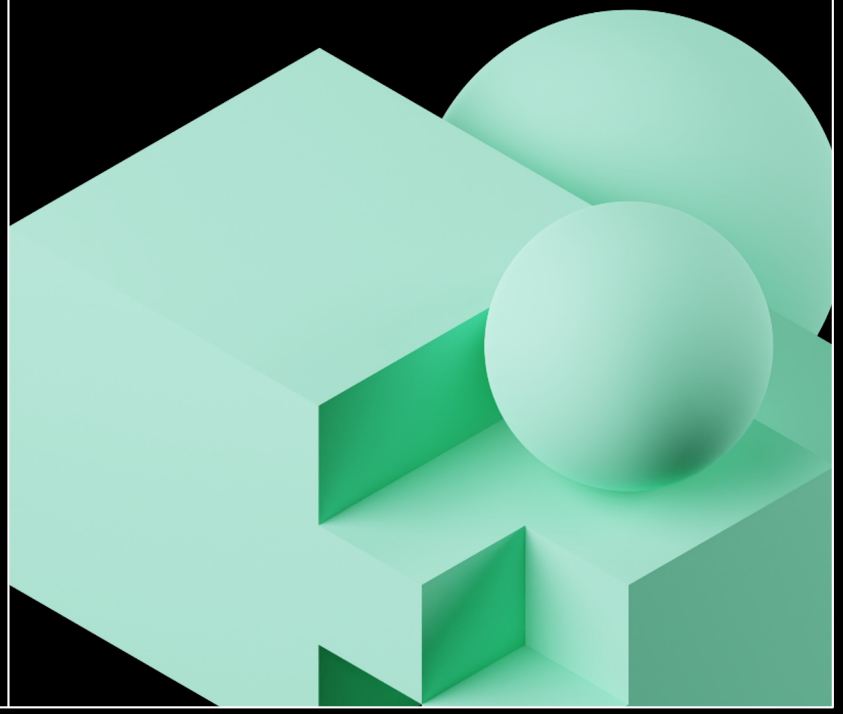
C Building dashboards

A more thoughtful way to get there

1. Gold views with aggregated data
2. Dashboard with notebook/SQL queries
3. Validation of pilot with businesses
4. Choice of environment for productization



Democratization WITH PARTNERS



Bradesco as a group of companies

And many others

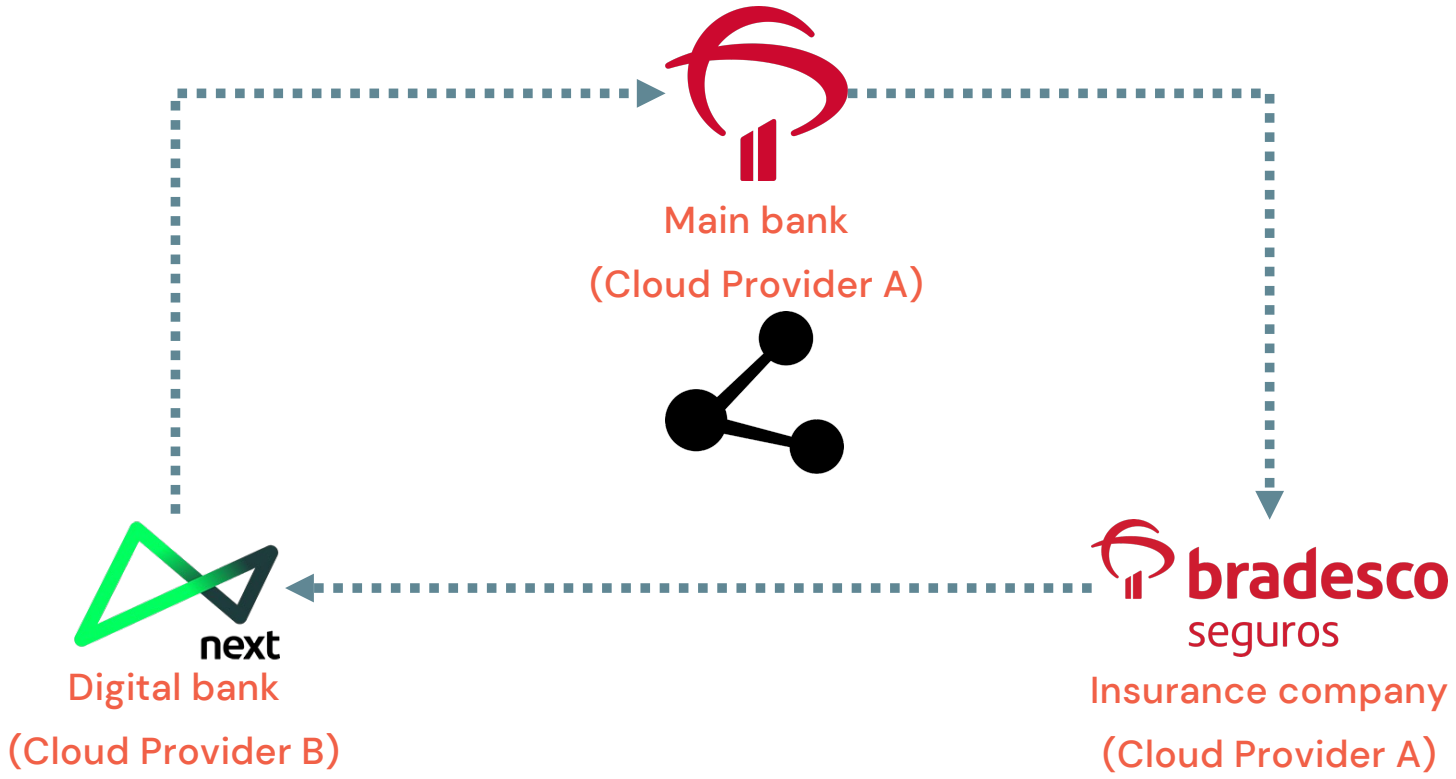
+10 companies

+55 million customers

SO MUCH

DATA!

The need to share data

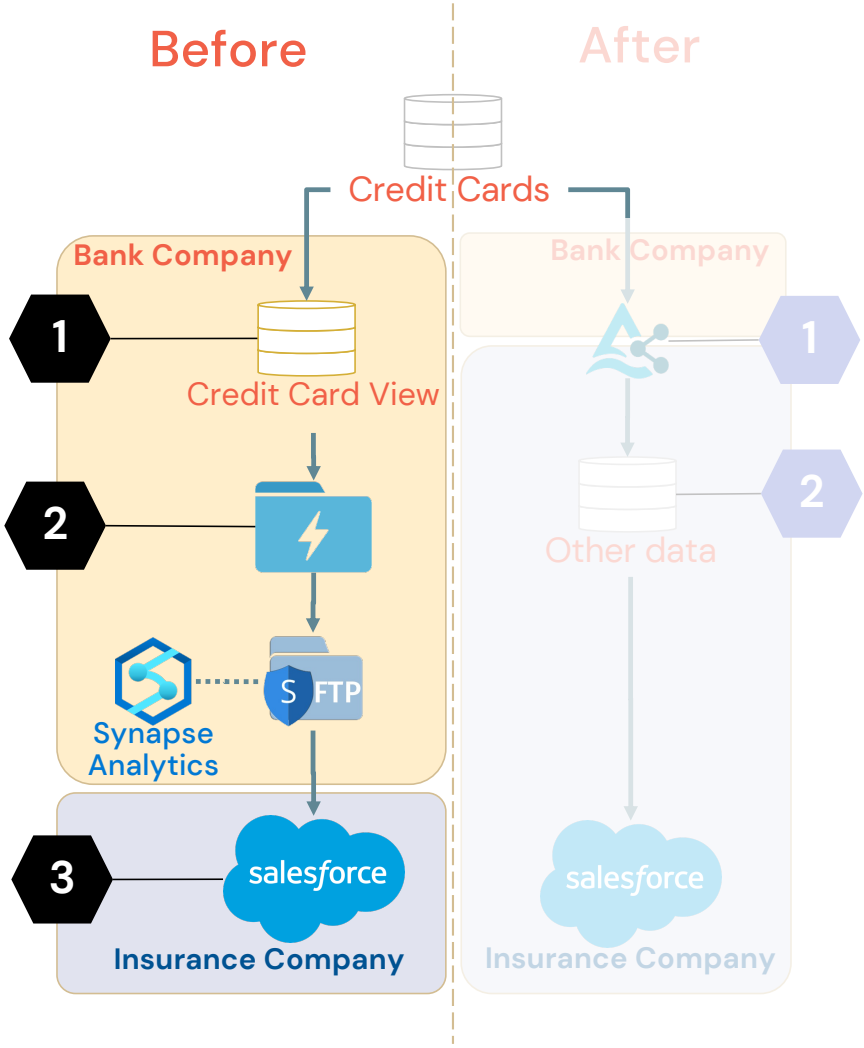


Solution

Simplifying data journey

1. Applied data views
2. Send to Storage in .zip files
3. SFTP procedure directly to Marketing Cloud

1. Delta Sharing creation
2. Freedom to create views and enhancement of data

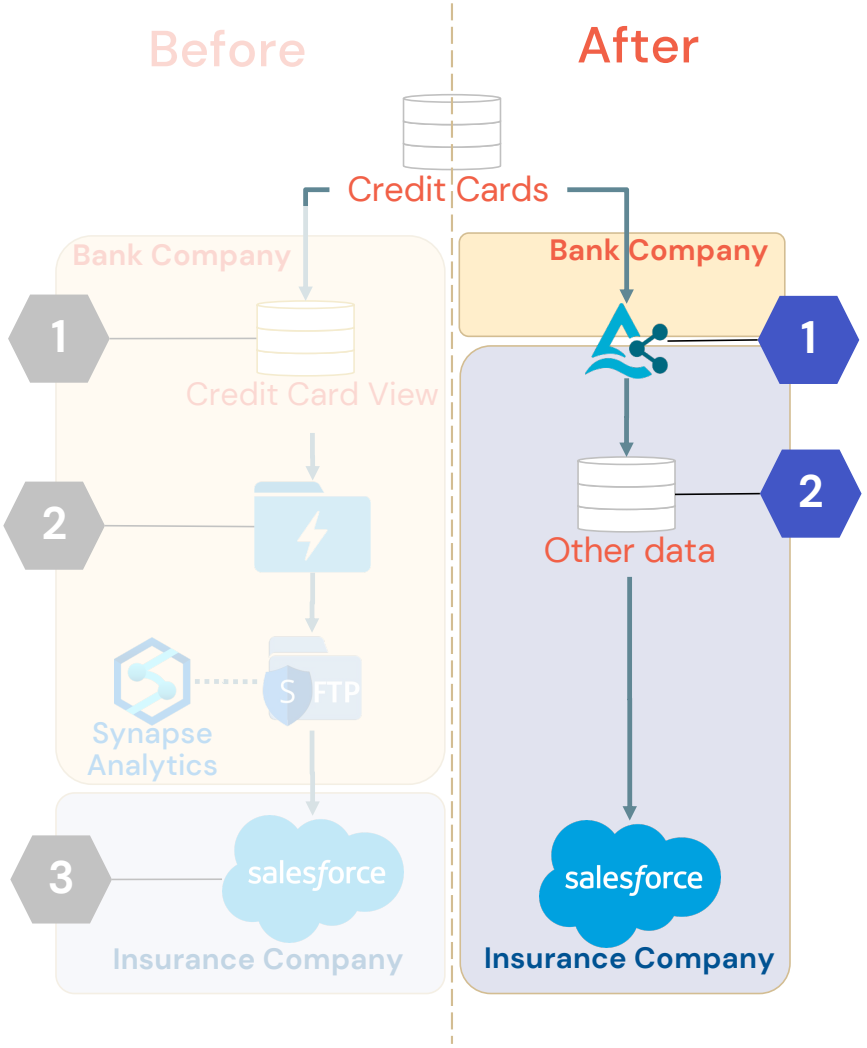


Solution

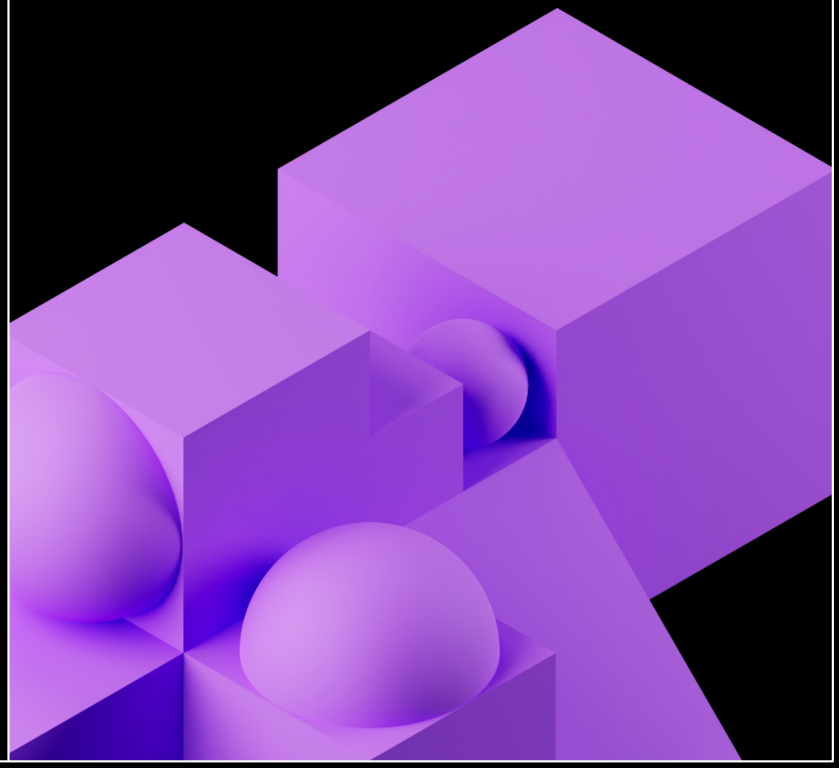
Simplifying data journey

1. Applied data views
2. Send to Storage in .zip files
3. SFTP procedure directly to Marketing Cloud

1. Delta Sharing creation
2. Freedom to create views and enhancement of data



Democratization WITH BUSINESS

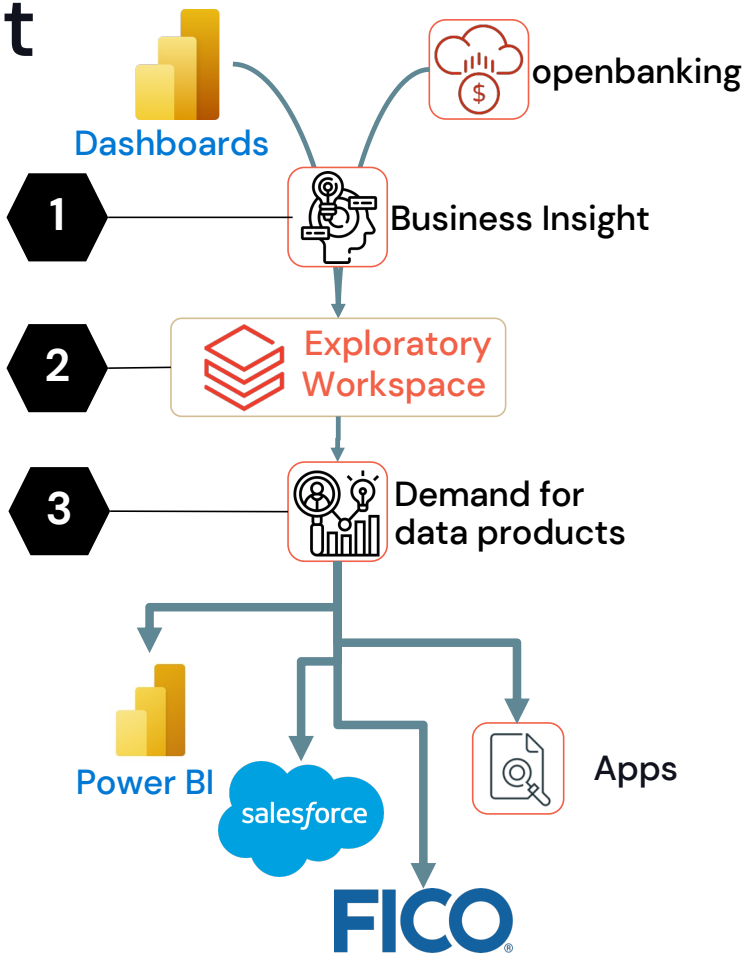


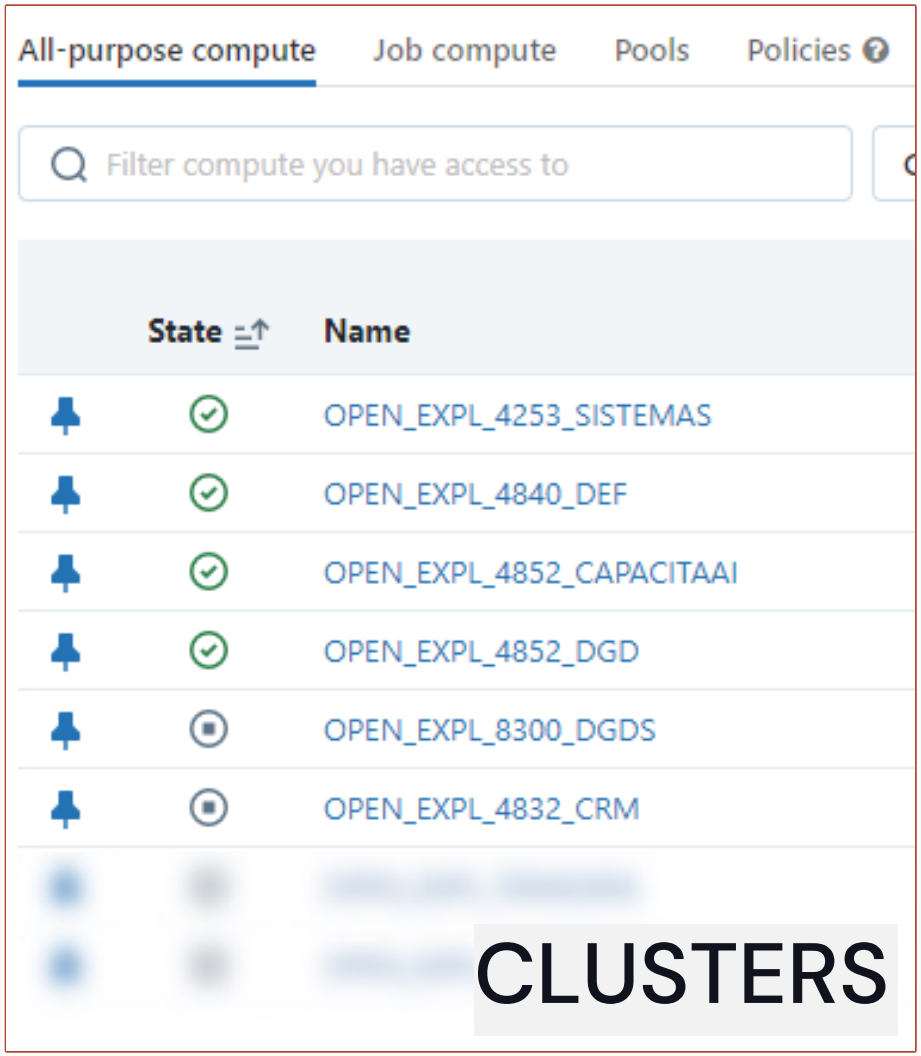
Where the magic happens!

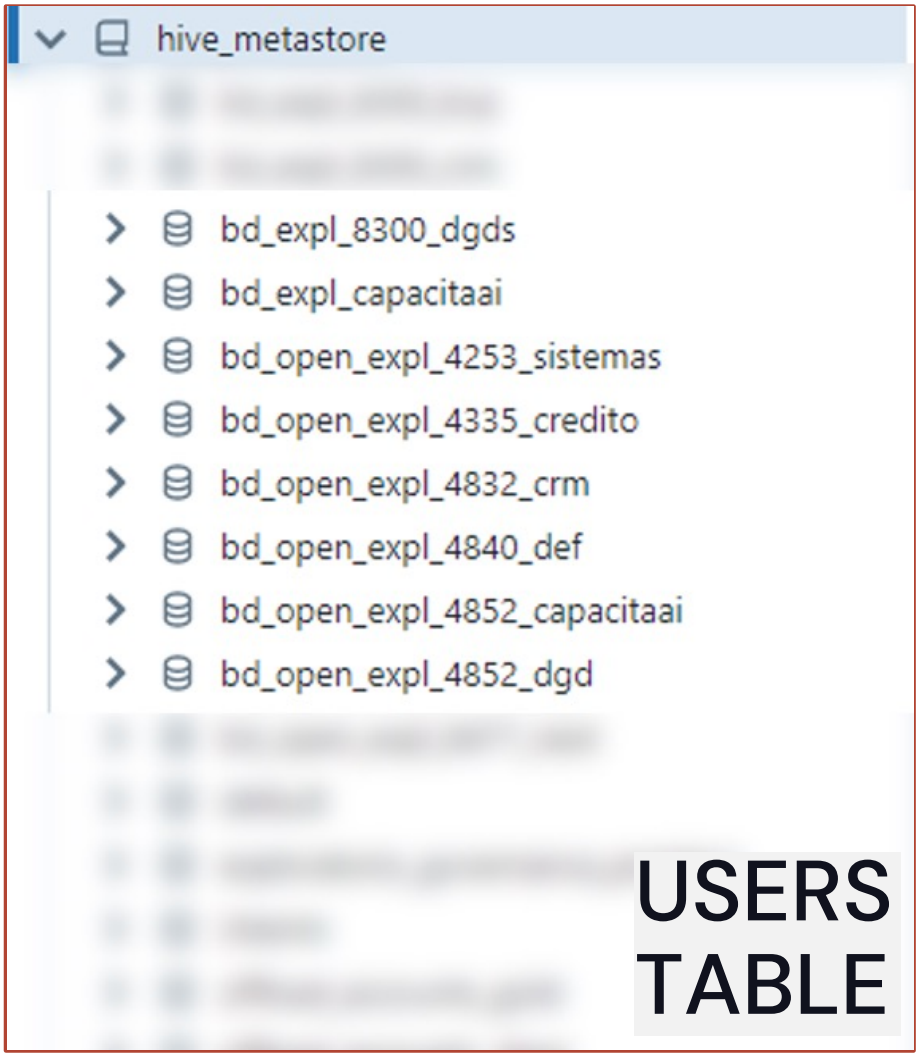
E Exploratory Environment

A place to leverage business

1. Business creation of demands
2. Dedicated space for business
   - 1 Cluster per business unity
   - UC Governance
   - Space for data testing
   - Not allowed sending data to external locations
3. Faster time-to-market




All-purpose compute
Job compute
Pools
Policies
Filter compute you have access to
State
Name
OPEN_EXPL_4253_SISTEMAS
OPEN_EXPL_4840_DEF
OPEN_EXPL_4852_CAPACITAAI
OPEN_EXPL_4852_DGD
OPEN_EXPL_8300_DGDS
OPEN_EXPL_4832_CRM
CLUSTERS


hive_metastore
bd_expl_8300_dgds
bd_expl_capacitaai
bd_open_expl_4253_sistemas
bd_open_expl_4335_credito
bd_open_expl_4832_crm
bd_open_expl_4840_def
bd_open_expl_4852_capacitaai
bd_open_expl_4852_dgd
USERS
TABLE

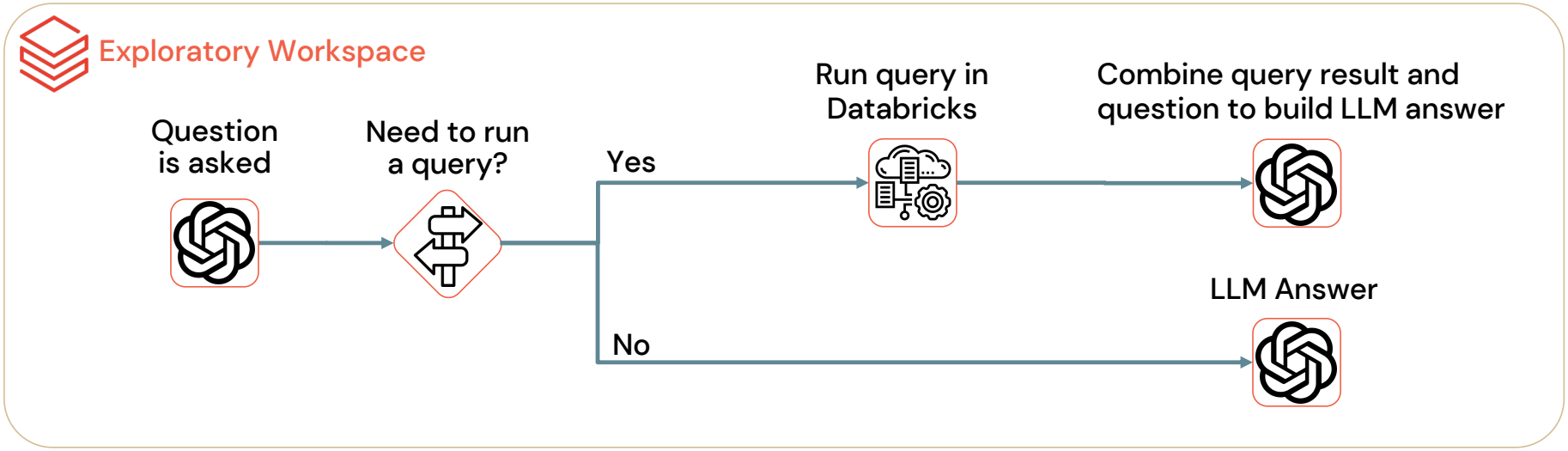
A NEW HORIZON

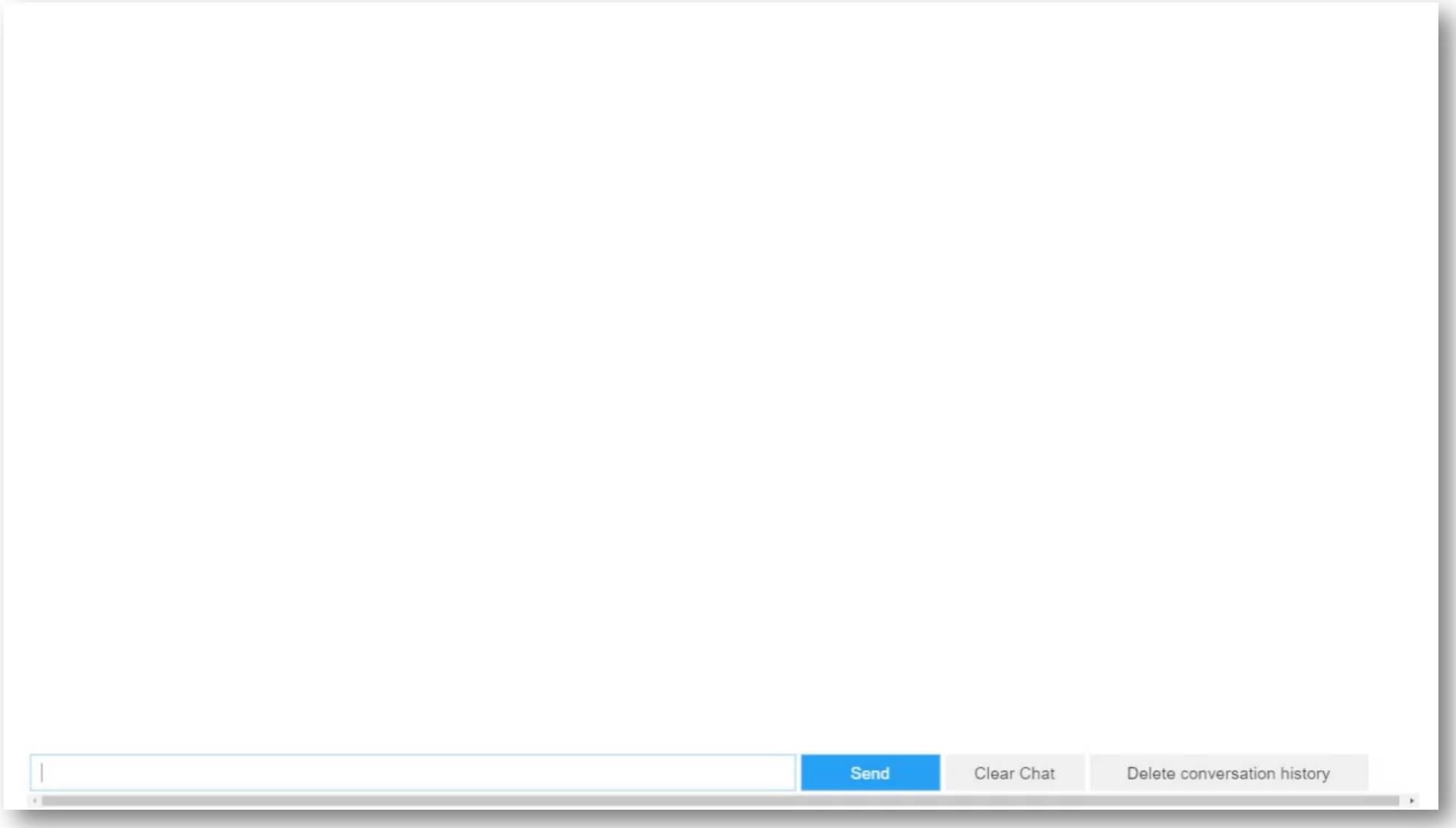
Exploratory Environment *powered by OpenAI*

Our problem: Data is hard to understand and needs technical skills to query inside exploratory.

Our solution: Use generative AI to help business users generate insights

How we did that:




Send
Clear Chat
Delete conversation history

BIG NUMBERS

300%

More speed sending new data to production

Allow

Cost segregation

Simplify

Data sharing

6

Business units using exploratory environment

200

Users in exploratory environment

10

Dashboards

+180

Use cases mapped

CUSTOMER STORY

Leading Latin America Into an Open Finance Future

24x
Faster time-to-market for new features

180+
Analytics and AI use cases

200+
Users accessing data insights

INDUSTRY: Financial services
SOLUTION: Customer Segmentation, Customer Entity Resolution
PLATFORM USE CASE: Lakehouse, Databricks SQL, Unity Catalog
CLOUD: Azure

Acknowledgment

OpenBanking Data Intelligence Team


Rafael


Vanessa


Ana


Augusto

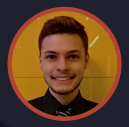

Phelipi


Helder


Fernando

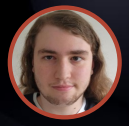

Pedro


Thiago T.


Thiago A.


Rubens


Rodrigo

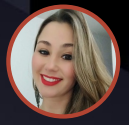

Gabriel


Luciano


Maria


Andreia


Karina

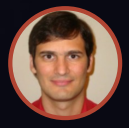

Dayane

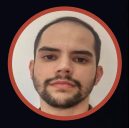

Vitor

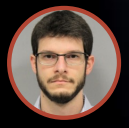

Luce


Karla


Alexandre


Reginaldo


Danilo


Caio

Databricks


Clecio


Guilherme


Jorge

Architecture team


Fabio


Arthur

And more...

Systems team

Governance team

Networking team

Thank you!

Msc Pedro Antonio Boareto
Data Engineer
Linkedin:

Fábio Luis Correia da Silva
Architecture Specialist
Linkedin:

Our customer story