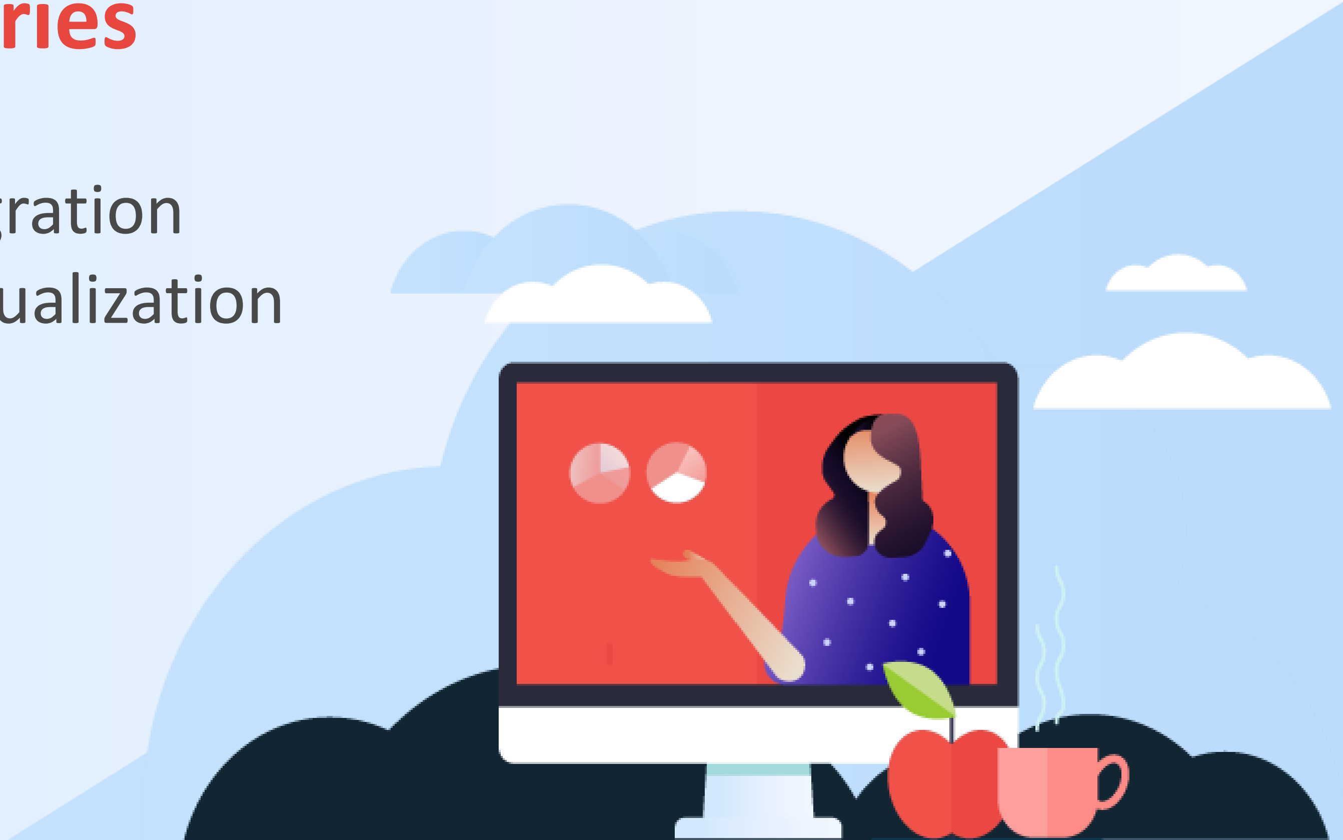
denodo

DATA VIRTUALIZATION
Packed Lunch Webinar Series

DATA VIRTUALIZATION
Packed Lunch Webinar Series

Sessions Covering Key Data Integration Challenges Solved with Data Virtualization



denodo

DATA VIRTUALIZATION
Packed Lunch Webinar Series

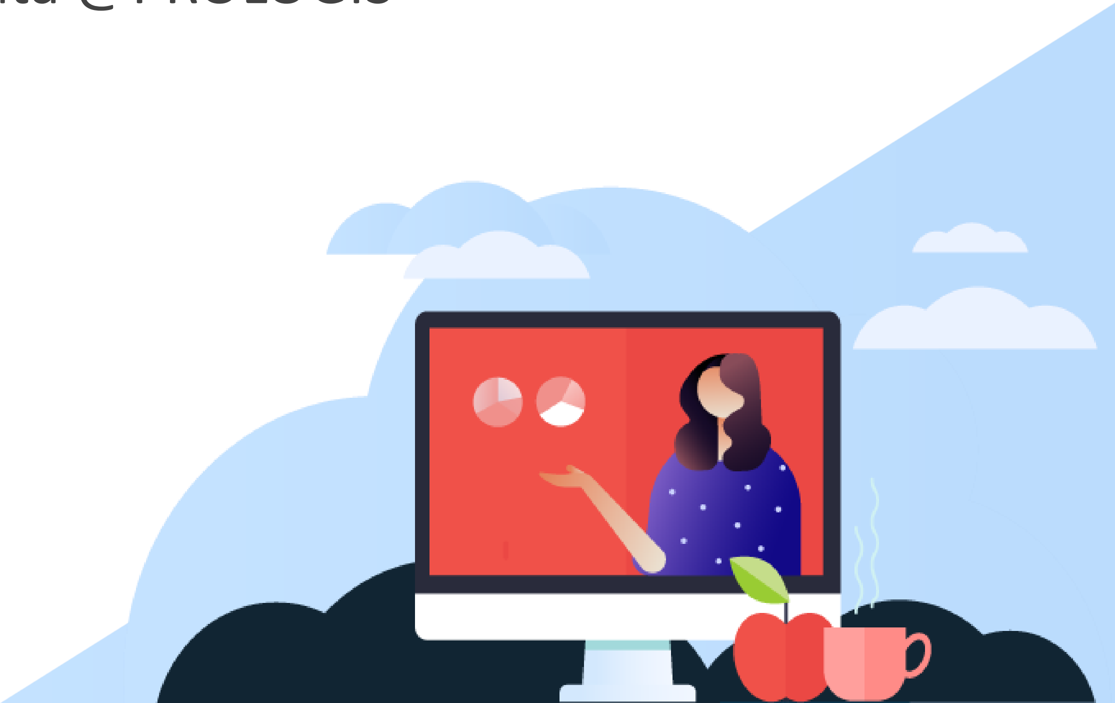
Augmentation, Collaboration and Governance are Defining the Future of Self-Service BI

Saptarshi Sengupta

Director of Product Marketing

Agenda

1. Broadening Horizon of Self-Service BI/Analytics
2. Problems with Self-Service BI/Analytics
3. The Role of Logical Data Fabric in Self-Service BI/Analytics
4. Customer Success Story – Democratizing Data @PROLOGIS
5. Conclusion



Broadening Horizon of Self-Service BI/Analytics

Case in Point

Data Democratization at LeasePlan

denodo

DATA VIRTUALIZATION
Packed Lunch Webinar Series

About LeasePlan | World Leader in Fleet Management

1963
Founded

Present in
32
countries

#1 EU reseller

LP Group B.V.
Investor consortium
Shareholder

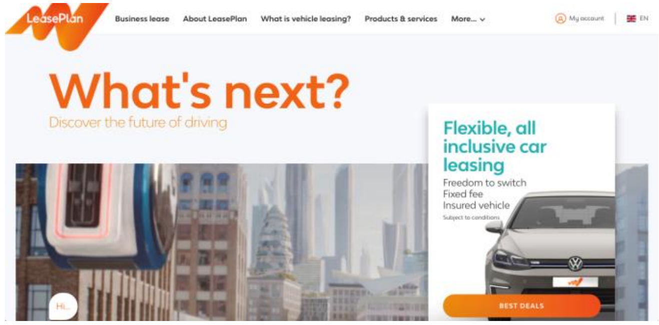
6,600+
Employees

HQ
Amsterdam

175,000+
Worldwide customers

Size of fleet
1.9 million

Vehicle Lifecycle | End-to-end Services Focus


LeasePlan
Business lease
About LeasePlan
What is vehicle leasing?
Products & services
More...
My account
EN

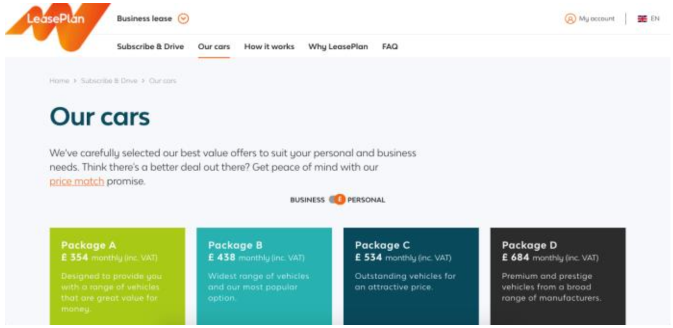
What's next?

Discover the future of driving

Flexible, all inclusive car leasing

Freedom to switch
Fixed fee
Insured vehicle
Subject to conditions

BEST DEALS
Hi...


LeasePlan
Business lease
My account
EN
Subscribe & Drive
Our cars
How it works
Why LeasePlan
FAQ
Home > Subscribe & Drive > Our cars

Our cars

We've carefully selected our best value offers to suit your personal and business needs. Think there's a better deal out there? Get peace of mind with our price match promise.

BUSINESS
PERSONAL

Package A
£ 354 monthly (inc. VAT)
Designed to provide you with a range of vehicles that are great value for money.

Package B
£ 438 monthly (inc. VAT)
Widest range of vehicles and our most popular option.

Package C
£ 534 monthly (inc. VAT)
Outstanding vehicles for an attractive price.

Package D
£ 684 monthly (inc. VAT)
Premium and prestige vehicles from a broad range of manufacturers.

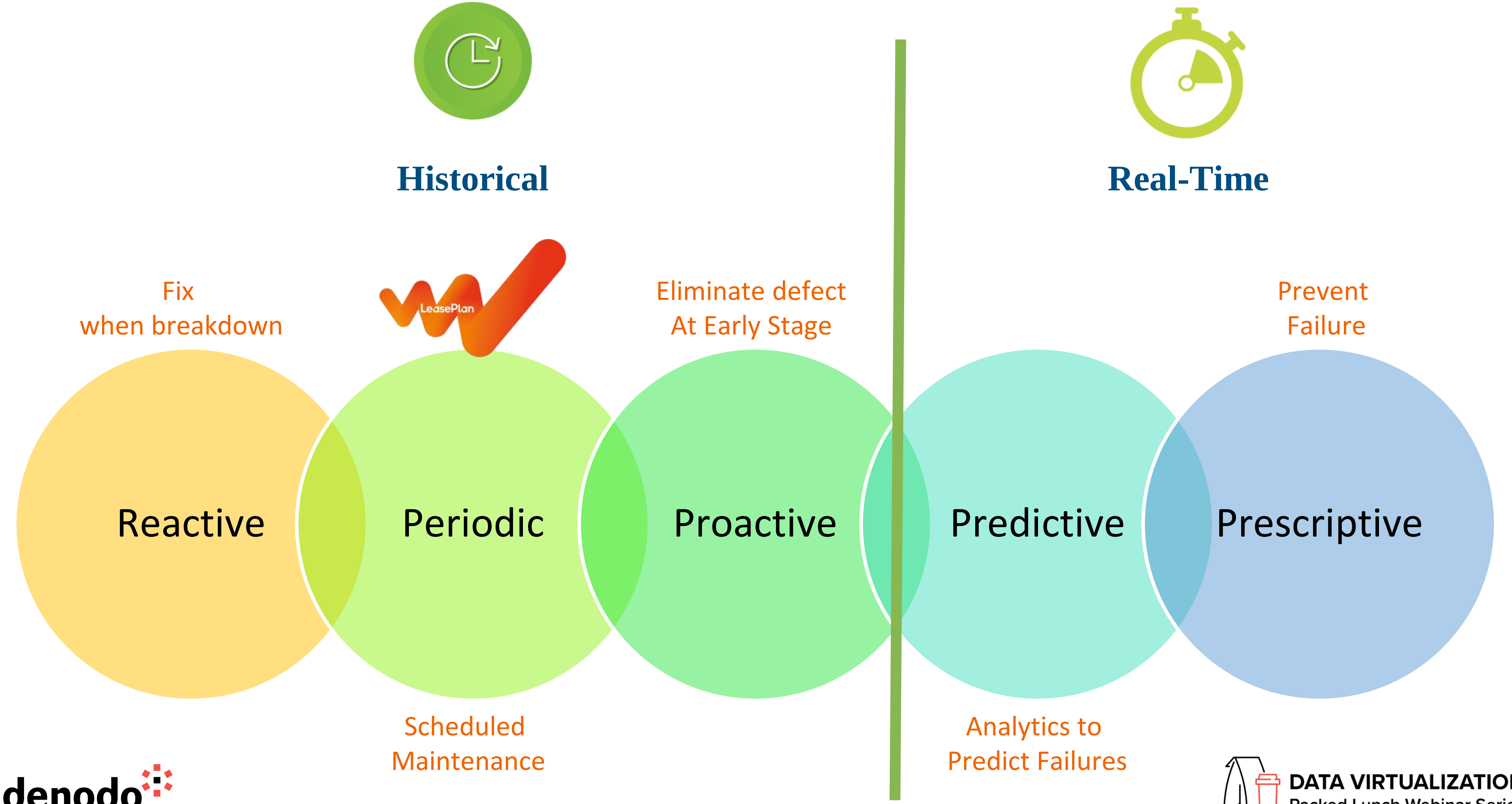
CarNext
by LeasePlan

LeasePlan typically owns the car and therefore owns the value chain

- Purchase
- Financing
- Insurance
- Maintenance Management
- Fuel Management
- Accident Management
- Rental Management
- High-quality used car resale & leasing

Car-as-a-Service → **CarNext.com**

Data in Action | Predictive Maintenance



Historical

Real-Time

Fix
when breakdown

Eliminate defect
At Early Stage

Prevent
Failure

Reactive

Periodic

Proactive

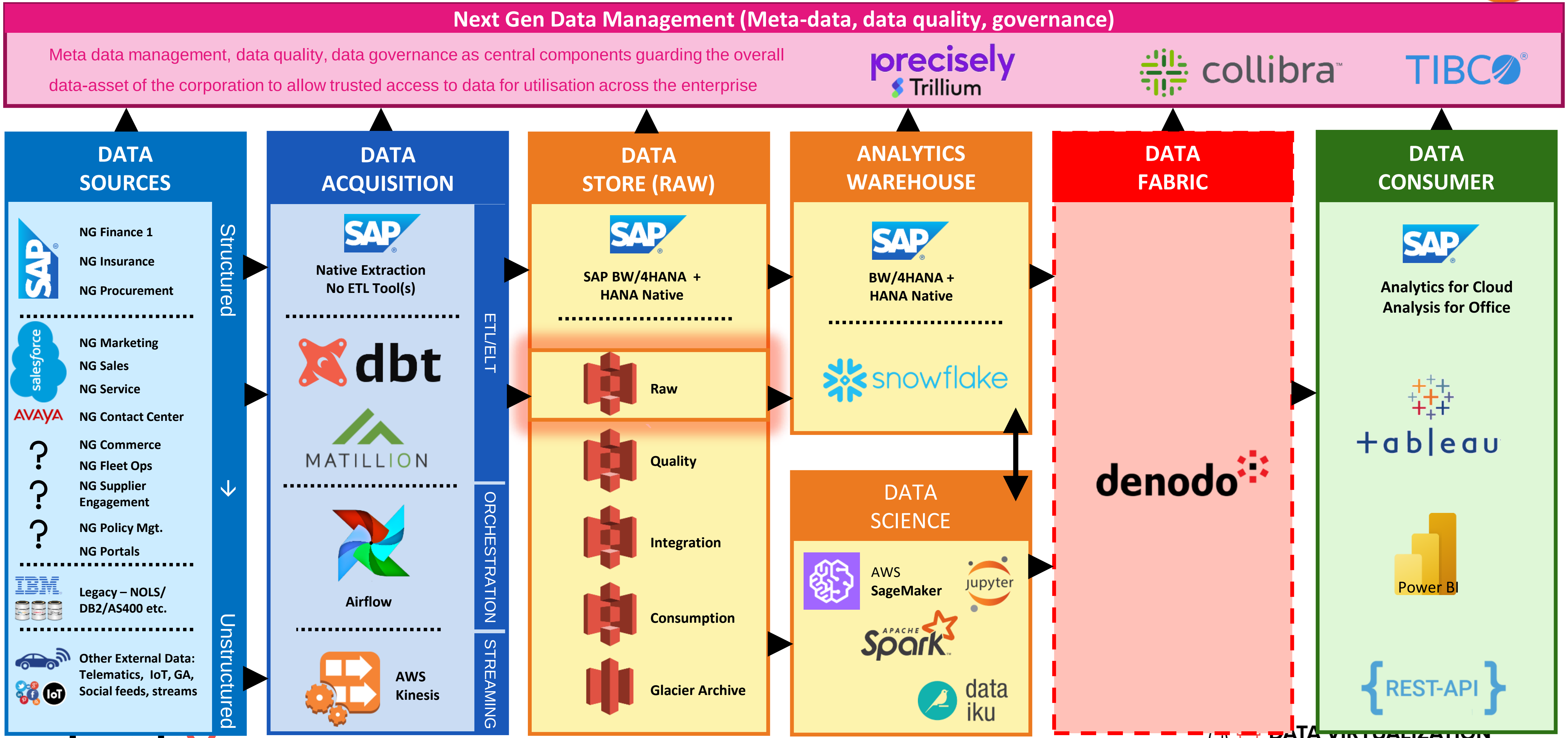
Predictive

Prescriptive

Scheduled
Maintenance

Analytics to
Predict Failures

Global Data Hub | Reference Architecture



Next Gen Data Management (Meta-data, data quality, governance)

Meta data management, data quality, data governance as central components guarding the overall data-asset of the corporation to allow trusted access to data for utilisation across the enterprise

precisely
Trillium
collibra
TIBCO
DATA SOURCES
NG Finance 1
NG Insurance
NG Procurement
salesforce
NG Marketing
NG Sales
NG Service
AVAYA
NG Contact Center
NG Commerce
NG Fleet Ops
NG Supplier Engagement
NG Policy Mgt.
NG Portals
IBM
Legacy – NOLS/ DB2/AS400 etc.
Other External Data: Telematics, IoT, GA, Social feeds, streams
IoT
Structured
Unstructured
DATA ACQUISITION
SAP
Native Extraction No ETL Tool(s)
dbt
MATILLION
Airflow
AWS Kinesis
ETL/ELT
ORCHESTRATION
STREAMING
DATA STORE (RAW)
SAP
SAP BW/4HANA + HANA Native
Raw
Quality
Integration
Consumption
Glacier Archive
ANALYTICS WAREHOUSE
SAP
BW/4HANA + HANA Native
snowflake
DATA SCIENCE
AWS SageMaker
jupyter
APACHE Spark
data iku
DATA FABRIC
denodo
DATA CONSUMER
SAP
Analytics for Cloud Analysis for Office
+ a b | e a u
Power BI
{REST-API}

> Tools will change, data will stay – Need a scalable, easily maintainable, high performing data platform that keeps all together, and serve as a data fabric.”

Tekin Mentes, Head of Enterprise Data Management, LeasePlan

Problems with Self-Service BI/Analytics

The Self-Service Promise

1. Faster Data Delivery to the Business

- More people needs access to data to make decisions
- Users can solve their own information needs

2. Unlock Data for Advanced, Specialized Needs

- New users with advanced analytical skills and needs e.g., data scientists and citizen analysts can create customized analysis and reports

3. Avoid the BI/IT Bottleneck

- Huge data growth, operation costs → IT is looking for cheaper and more flexible solutions without disrupting the business

Self-Service: Problems

Problems with Self-Service Initiatives (Source: Eckerson Group)

- 64% of users rate their experience as “average” or “lower”
- +70% says “it requires more training than expected”
- +67% says “self-service tools are difficult to use”
- +60% says they “create informational chaos”

What is the Problem ?

Complexity:

- No unified infrastructure (multiple data sources and analysis / visualization tools)
- Integrate, transform and combine data is hard

Self-Service vs Governance

- Inconsistent reports / Single Source of Truth
- Compliance with company glossaries and policies
- How to enforce consistent security, data quality and governance policies across multiple systems

Do Data Governance Tools Solve the Problem ?

DG Tools allow:

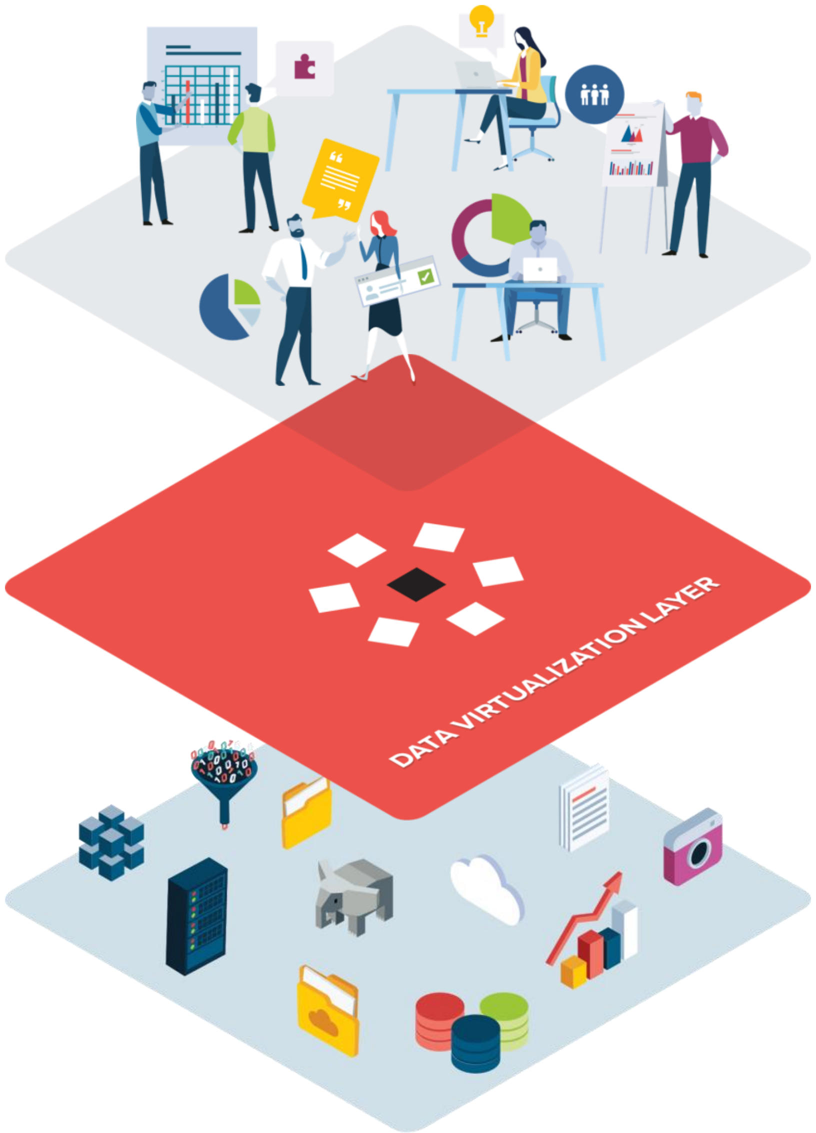
- Informing about data assets and their level of quality
- Defining unified glossaries and terminology
- Defining data quality and data governance policies, and managing/tracking changes

Disconnected from the data delivery process

- Do not ensure delivered data conforms to glossaries
- Do not enforce security, data quality and governance policies in the data delivery process
  - The problem of how to enforce these policies across multiple data sources and consumption tools remain

The Role of Logical Data Fabric in Self-Service BI/Analytics

Denodo's Logical Data Fabric Enables Information Self-Service



1. Single Access Point to all Data at any location

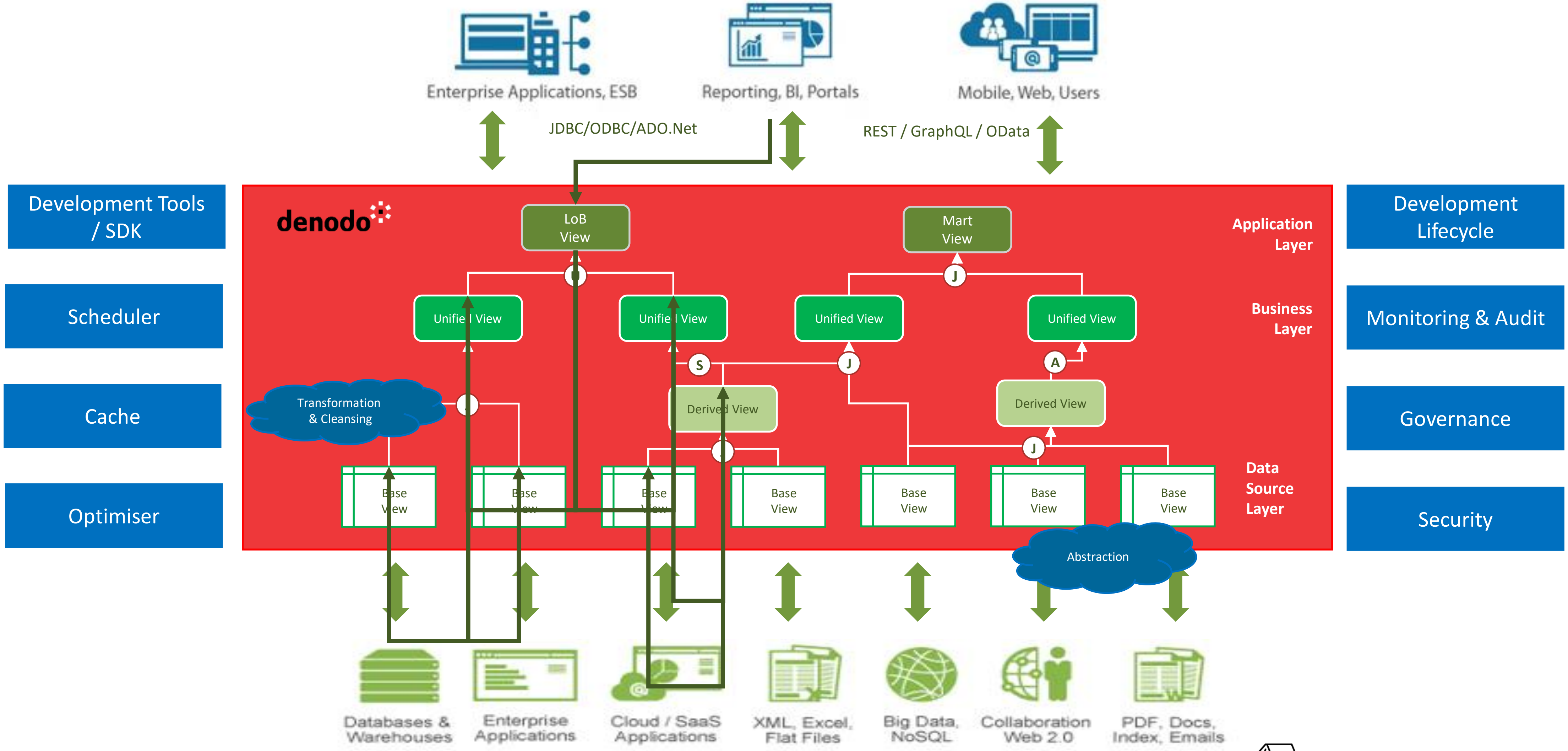
2. Semantic Layer – Expose Data in Business-Friendly form, adapted to the needs of each consumer

3. Up to 80% reduction in integration costs, in terms of resources and technology data

4. Consume data with any tool and access technology (SQL, REST, GraphQL, OData,…)

5. Single entry point to apply security and governance policies

Data Virtualization: Logical Data Delivery for the Business


Enterprise Applications, ESB
Reporting, BI, Portals
Mobile, Web, Users
JDBC/ODBC/ADO.Net
REST / GraphQL / OData
Development Tools / SDK
Scheduler
Cache
Optimiser
Development Lifecycle
Monitoring & Audit
Governance
Security
denodo
LoB View
Mart View
Application Layer
Unified View
Unified View
Unified View
Unified View
Business Layer
J
J
S
J
A
J
J
Transformation & Cleansing
Derived View
Derived View
Base View
Base View
Base View
Base View
Base View
Base View
Base View
Data Source Layer
Abstraction
Databases & Warehouses
Enterprise Applications
Cloud / SaaS Applications
XML, Excel, Flat Files
Big Data, NoSQL
Collaboration Web 2.0
PDF, Docs, Index, Emails

Data Virtualization for Data Governance

Single Entry Point for Enforcing Security and Governance Policies

Data on-premises and off, combined through the same governed virtual layer

Single Source of Truth / Canonical Views

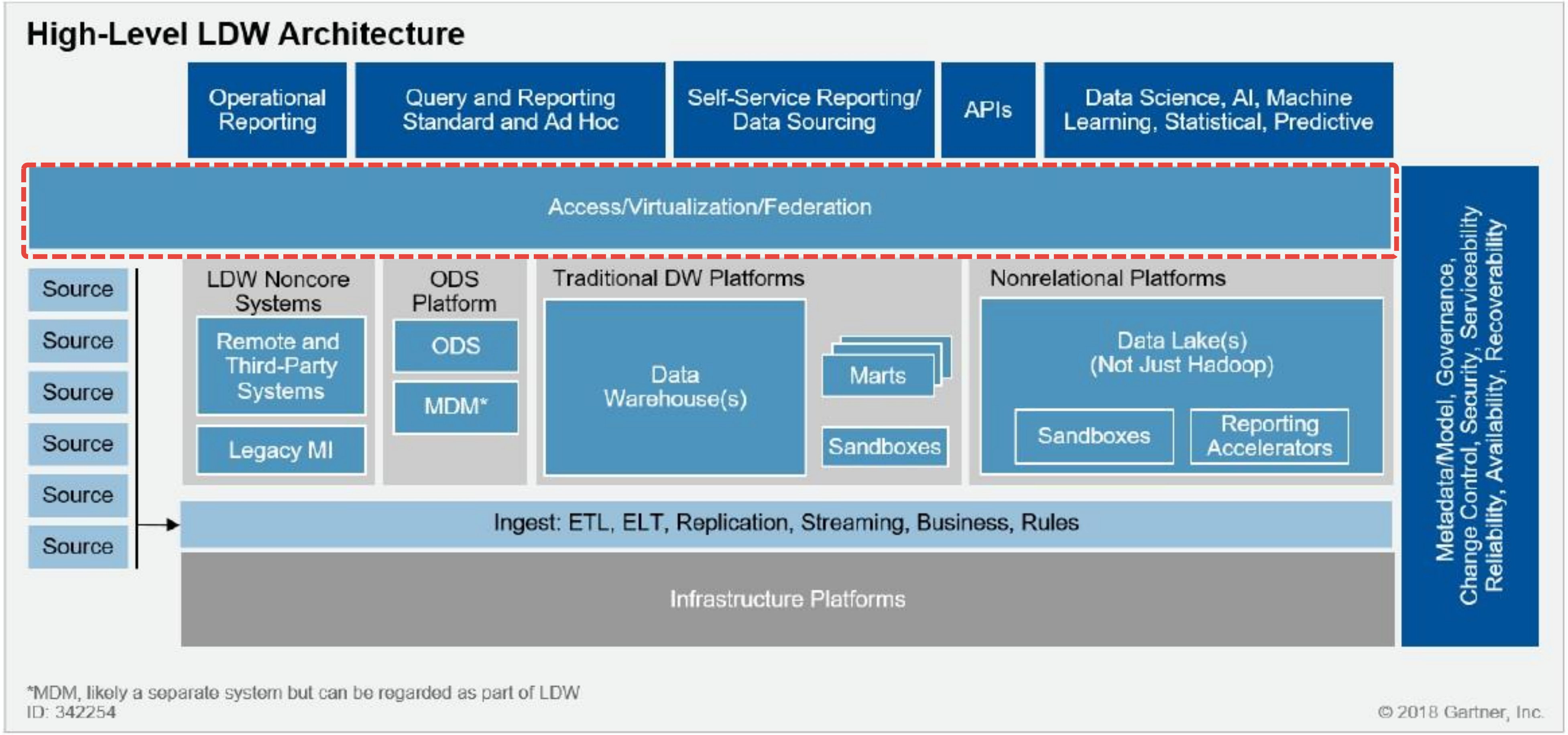
Who is Doing / Accessing What, When and How

Fewer copies of personal data. Lineage of copies is available.

Logical Data Warehouse – The path to the future

Gartner

Figure 1. An Overview of a Typical LDW Architecture


High-Level LDW Architecture
Operational Reporting
Query and Reporting Standard and Ad Hoc
Self-Service Reporting/ Data Sourcing
APIs
Data Science, AI, Machine Learning, Statistical, Predictive
Access/Virtualization/Federation
Source
Source
Source
Source
Source
Source
LDW Noncore Systems
Remote and Third-Party Systems
Legacy MI
ODS Platform
ODS
MDM*
Traditional DW Platforms
Data Warehouse(s)
Marts
Sandboxes
Nonrelational Platforms
Data Lake(s) (Not Just Hadoop)
Sandboxes
Reporting Accelerators
Ingest: ETL, ELT, Replication, Streaming, Business, Rules
Infrastructure Platforms
Metadata/Model, Governance, Change Control, Security, Serviceability, Reliability, Availability, Recoverability

*MDM, likely a separate system but can be regarded as part of LDW
ID: 342254

© 2018 Gartner, Inc.

Source: Gartner (April 2018)

Personas

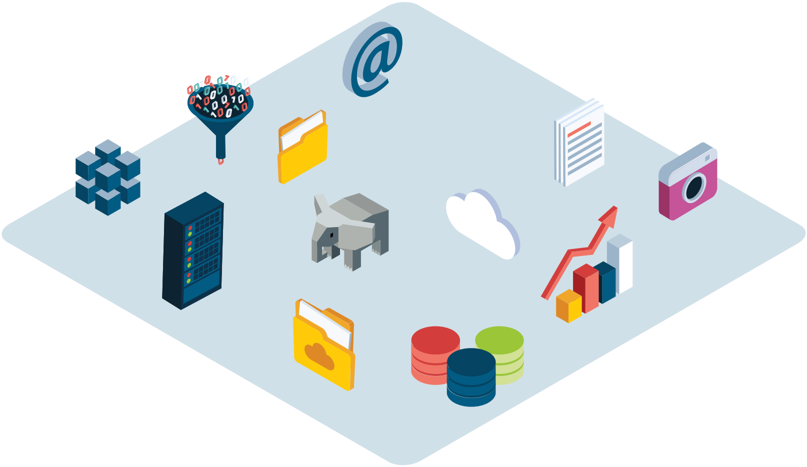
BI Analyst

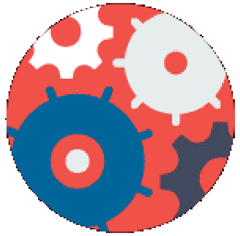
Data Scientist

Administration & Operations

Business Users

Data Engineers

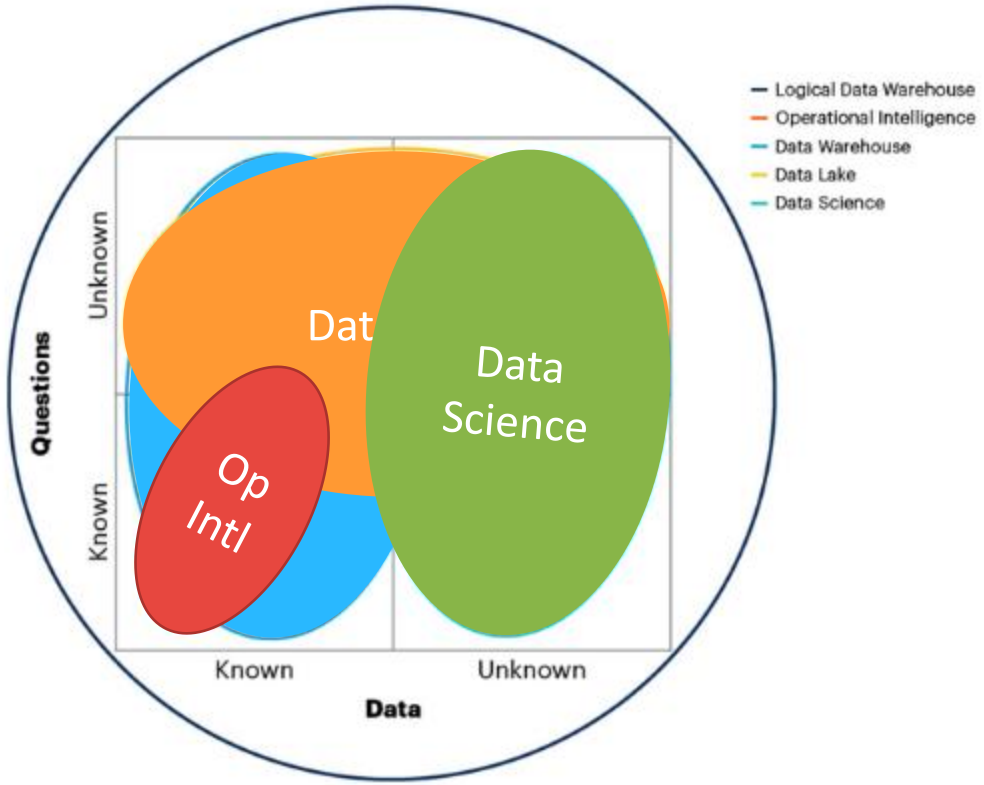




Application-to-Application

One Size Doesn't Fit All: Industry Analyst

Logical Data Warehouse Overlay for the Data and Analytics Infrastructure Model



Logical Data Warehouse
Operational Intelligence
Data Warehouse
Data Lake
Data Science
Questions
Unknown
Known
Data
Known
Unknown
Dat
Data Science
Op Intl

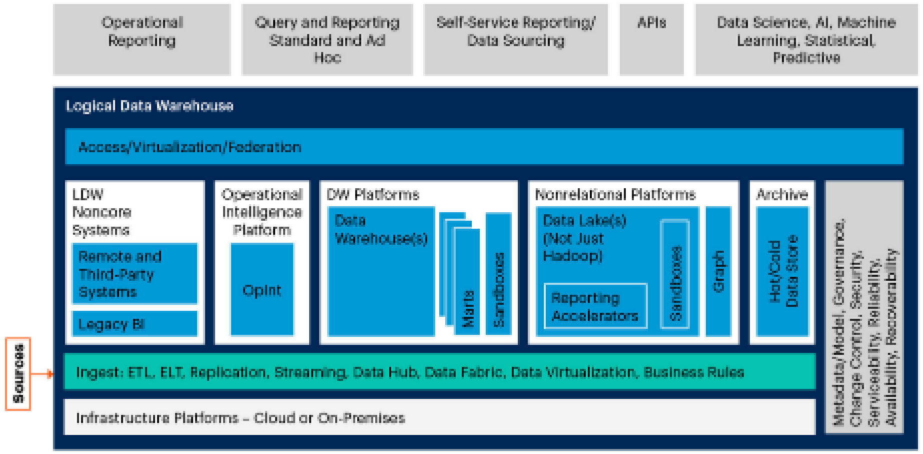
Source: Gartner
719608_C

The Practical Logical Data Warehouse (Dec 2020) by Henry Cook, Rick Greenwald and Adam Ronthal

Distributed and Logical Architectures

- **Distributed**: Data resides at multiple systems / locations
  - Data today is too big and distributed
  - Modern analytics needs are too diverse: one size never fits all
  - Hybrid and multicloud
- **Logical**: Consumers access data through semantic models, decoupled from data location and physical schemas
  - Semantic models adapt to consumer needsand enforce common policies
  - Not necessarily virtual/federated access: any delivery style including replication when needed
  - Allows for technology evolution and infrastructure changes (e.g. cloud transition)

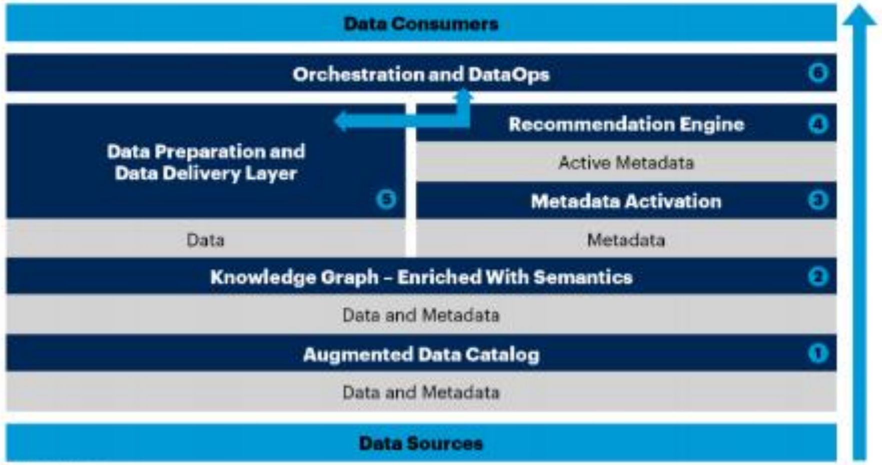
Practical LDW Architecture in Context



Operational Reporting
Query and Reporting Standard and Ad Hoc
Self-Service Reporting/ Data Sourcing
APIs
Data Science, AI, Machine Learning, Statistical, Predictive
Logical Data Warehouse
Access/Virtualization/Federation
LDW Noncore Systems
Remote and Third-Party Systems
Legacy BI
Operational Intelligence Platform
OpInt
DW Platforms
Data Warehouse(s)
Marts
Sandboxes
Nonrelational Platforms
Data Lake(s) (Not Just Hadoop)
Reporting Accelerators
Sandboxes
Graph
Archive
Hot/Cold Data Store
Metadata/Model, Governance, Change Control, Security, Serviceability, Reliability, Availability, Recoverability
Sources
Ingest: ETL, ELT, Replication, Streaming, Data Hub, Data Fabric, Data Virtualization, Business Rules
Infrastructure Platforms – Cloud or On-Premises

Source: Gartner

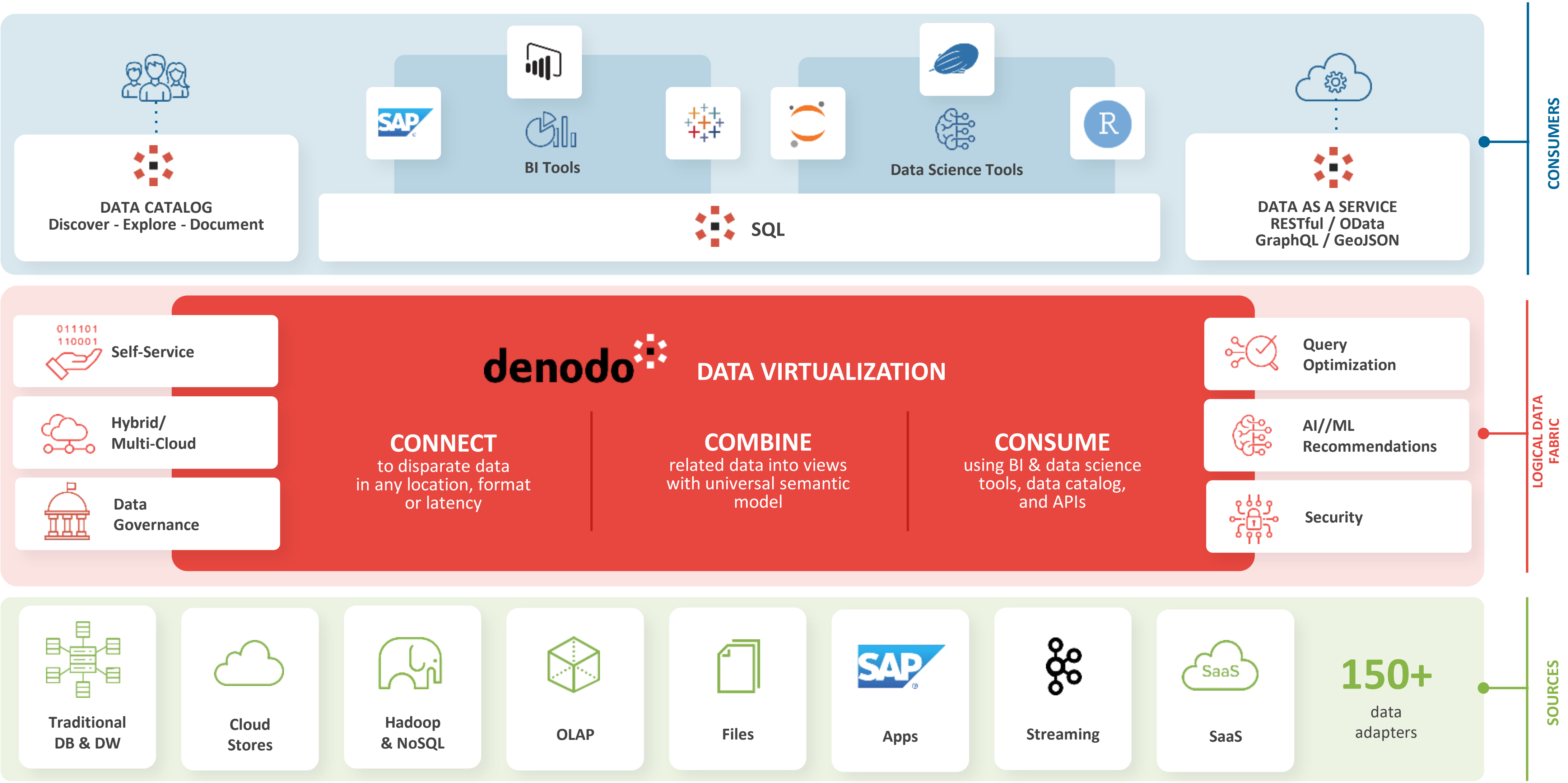
Data Fabric Pillars



Data Consumers
Orchestration and DataOps ⑥
Recommendation Engine ④
Active Metadata
Data Preparation and Data Delivery Layer ⑤
Metadata Activation ③
Data
Metadata
Knowledge Graph – Enriched With Semantics ②
Data and Metadata
Augmented Data Catalog ①
Data and Metadata
Data Sources

Source: Gartner
720486_C

Denodo Platform



CONSUMERS

DATA CATALOG
Discover - Explore - Document

BI Tools

Data Science Tools

SQL

DATA AS A SERVICE
RESTful / OData
GraphQL / GeoJSON

LOGICAL DATA FABRIC

Self-Service

Hybrid/
Multi-Cloud

Data
Governance

denodo DATA VIRTUALIZATION

CONNECT
to disparate data
in any location, format
or latency

COMBINE
related data into views
with universal semantic
model

CONSUME
using BI & data science
tools, data catalog,
and APIs

Query
Optimization

AI//ML
Recommendations

Security

SOURCES

Traditional
DB & DW

Cloud
Stores

Hadoop
& NoSQL

OLAP

Files

Apps

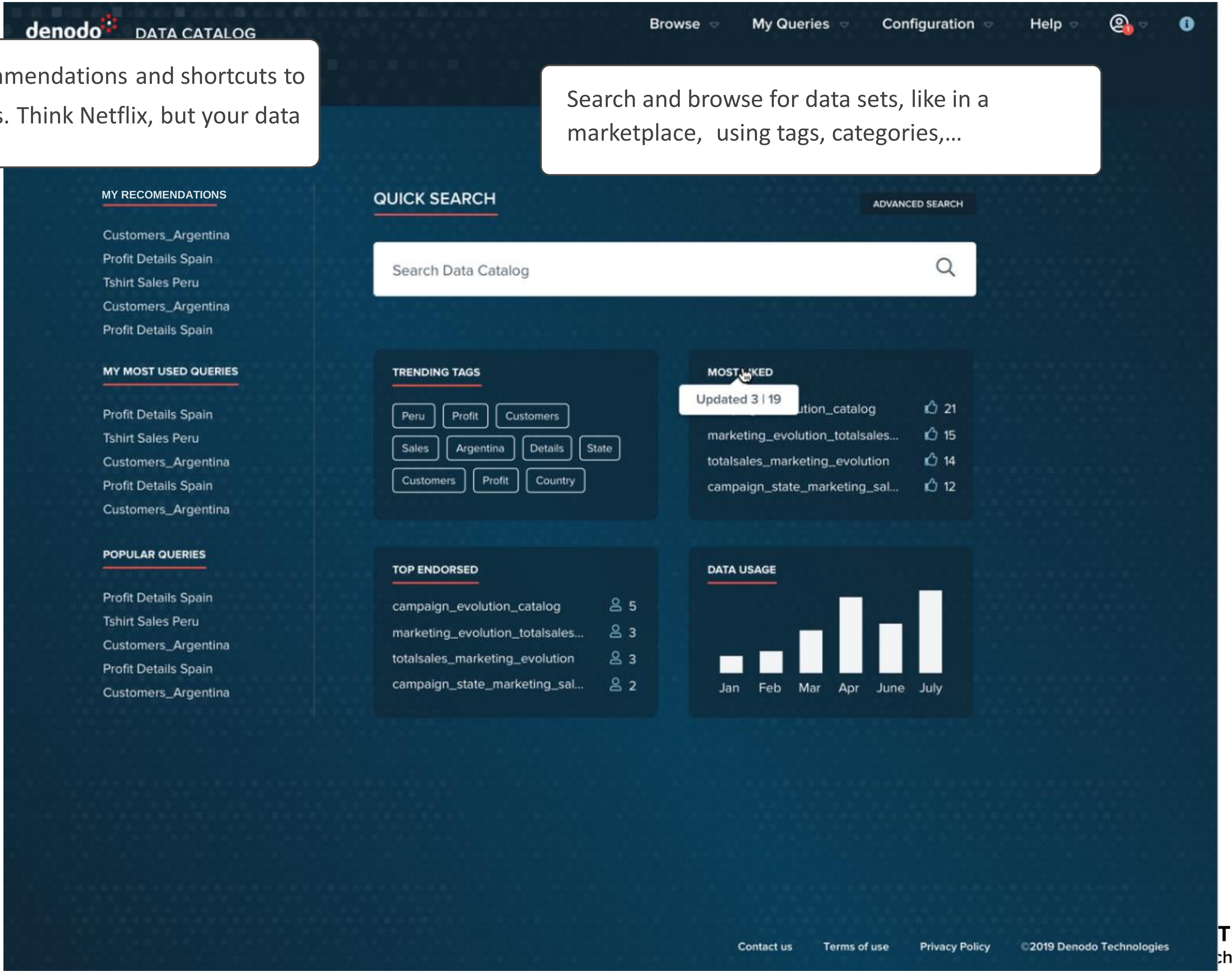
Streaming

SaaS

150+
data
adapters

Personalized recommendations and shortcuts to most used datasets. Think Netflix, but your data

Search and browse for data sets, like in a marketplace, using tags, categories,…



MY RECOMENDATIONS

- Customers_Argentina
- Profit Details Spain
- Tshirt Sales Peru
- Customers_Argentina
- Profit Details Spain

MY MOST USED QUERIES

- Profit Details Spain
- Tshirt Sales Peru
- Customers_Argentina
- Profit Details Spain
- Customers_Argentina

POPULAR QUERIES

- Profit Details Spain
- Tshirt Sales Peru
- Customers_Argentina
- Profit Details Spain
- Customers_Argentina

QUICK SEARCH

ADVANCED SEARCH

Search Data Catalog

TRENDING TAGS

Peru | Profit | Customers | Sales | Argentina | Details | State | Customers | Profit | Country

MOST LIKED

Updated 3 | 19

ution_catalog	21
marketing_evolution_totalsales...	15
totalsales_marketing_evolution	14
campaign_state_marketing_sal...	12

TOP ENDORSED

campaign_evolution_catalog	5
marketing_evolution_totalsales...	3
totalsales_marketing_evolution	3
campaign_state_marketing_sal...	2

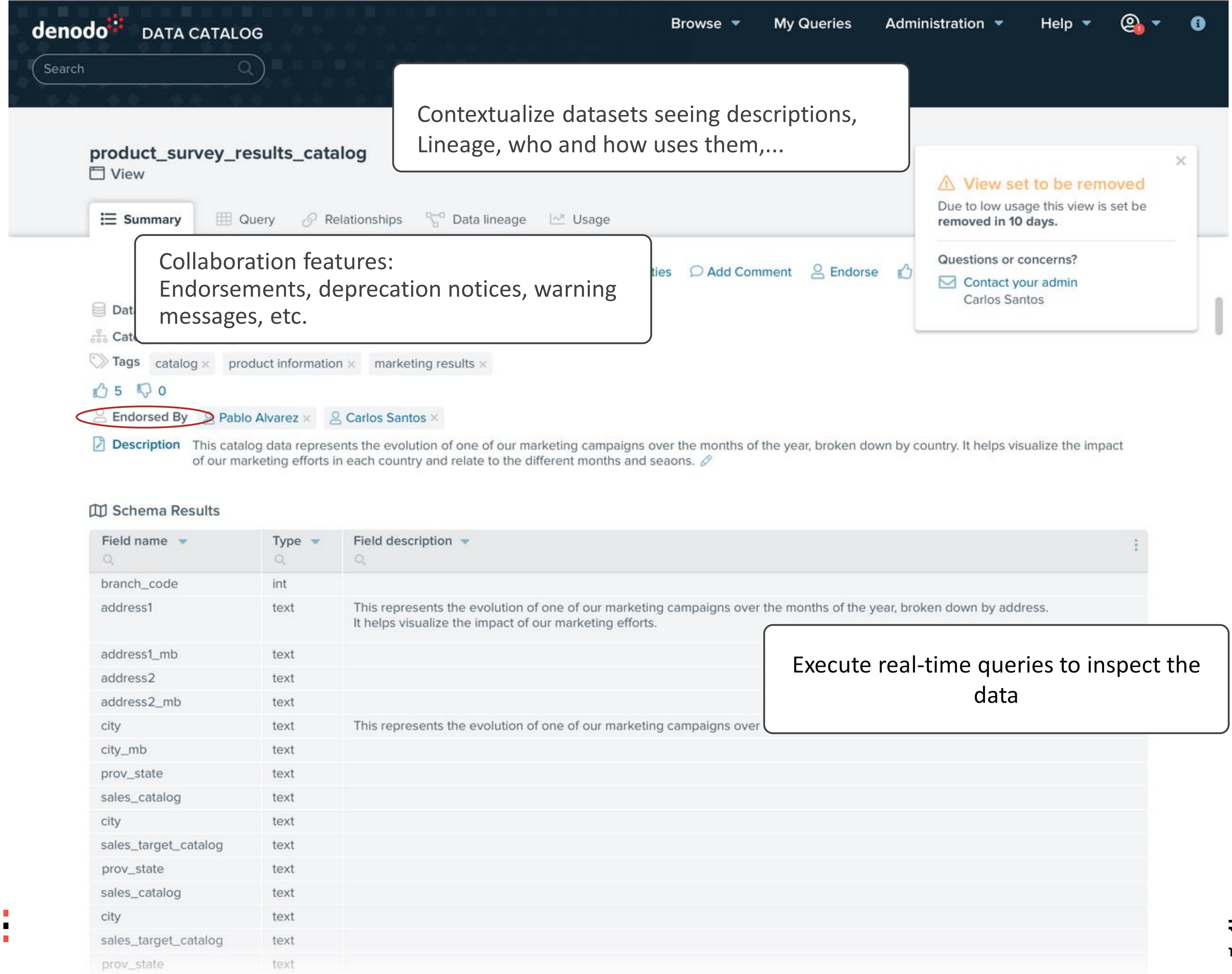
DATA USAGE

Contact us   Terms of use   Privacy Policy   ©2019 Denodo Technologies

Contextualize datasets seeing descriptions, Lineage, who and how uses them,...

Collaboration features:
Endorsements, deprecation notices, warning messages, etc.

Execute real-time queries to inspect the data


denodo DATA CATALOG
Search
Browse
My Queries
Administration
Help

product_survey_results_catalog
View

Summary
Query
Relationships
Data lineage
Usage
ties
Add Comment
Endorse
Dat
Cat

View set to be removed
Due to low usage this view is set be **removed in 10 days.**

Questions or concerns?
Contact your admin
Carlos Santos

Tags: catalog, product information, marketing results

5 0

Endorsed By: Pablo Alvarez, Carlos Santos

Description This catalog data represents the evolution of one of our marketing campaigns over the months of the year, broken down by country. It helps visualize the impact of our marketing efforts in each country and relate to the different months and seaons.

Schema Results

Field name	Type	Field description
branch_code	int	
address1	text	This represents the evolution of one of our marketing campaigns over the months of the year, broken down by address. It helps visualize the impact of our marketing efforts.
address1_mb	text	
address2	text	
address2_mb	text	
city	text	This represents the evolution of one of our marketing campaigns over
city_mb	text	
prov_state	text	
sales_catalog	text	
city	text	
sales_target_catalog	text	
prov_state	text	
sales_catalog	text	
city	text	
sales_target_catalog	text	
prov_state	text	

Consume data from the Catalog or
Your favorite visualization tool

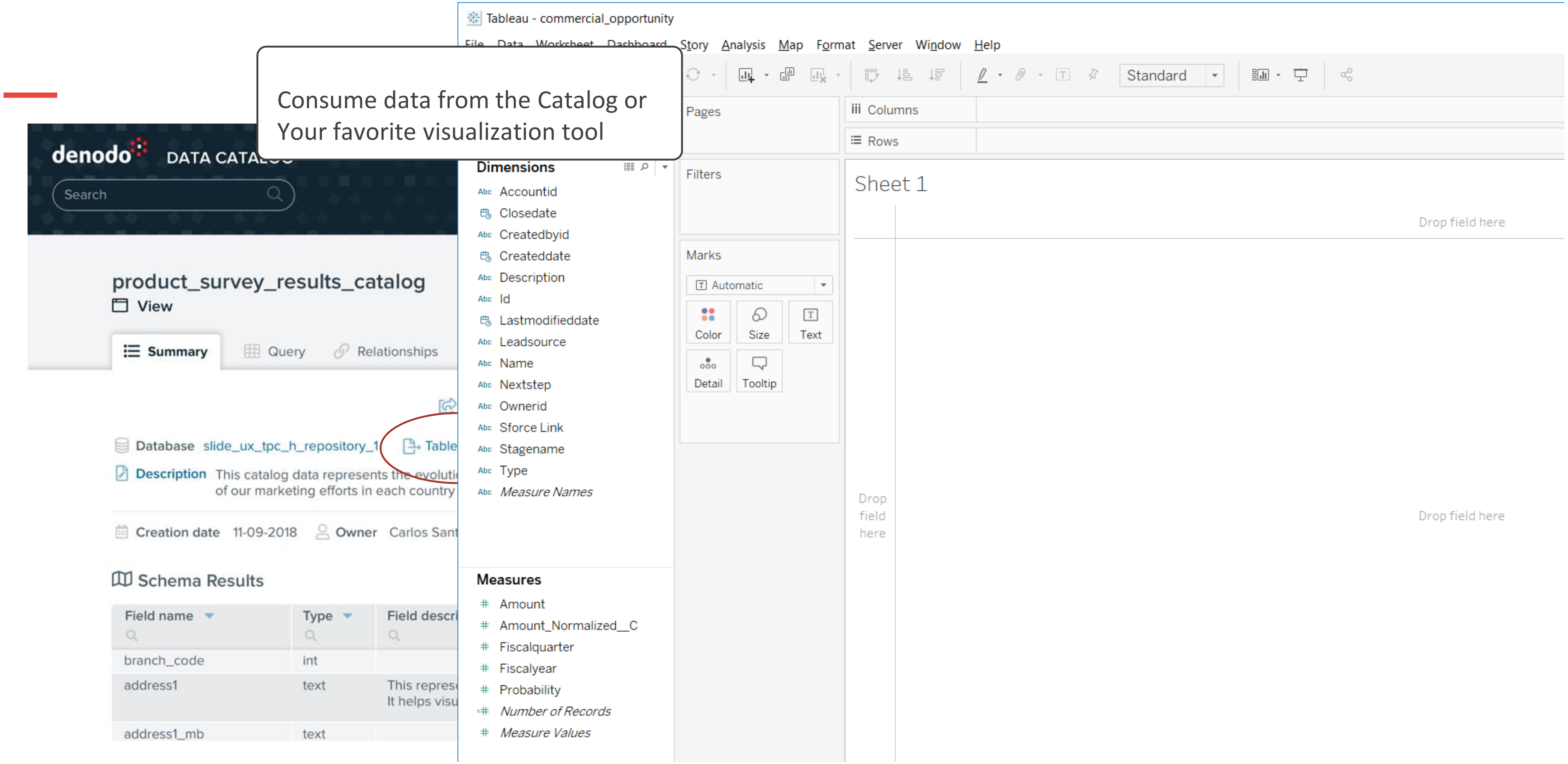


Tableau - commercial_opportunity

product_survey_results_catalog
View

Database slide_ux_tpc_h_repository_1
Description This catalog data represents the evoluti
of our marketing efforts in each country
Creation date 11-09-2018 Owner Carlos Sant

Schema Results

Field name	Type	Field descri
branch_code	int	
address1	text	This repres It helps visu
address1_mb	text	

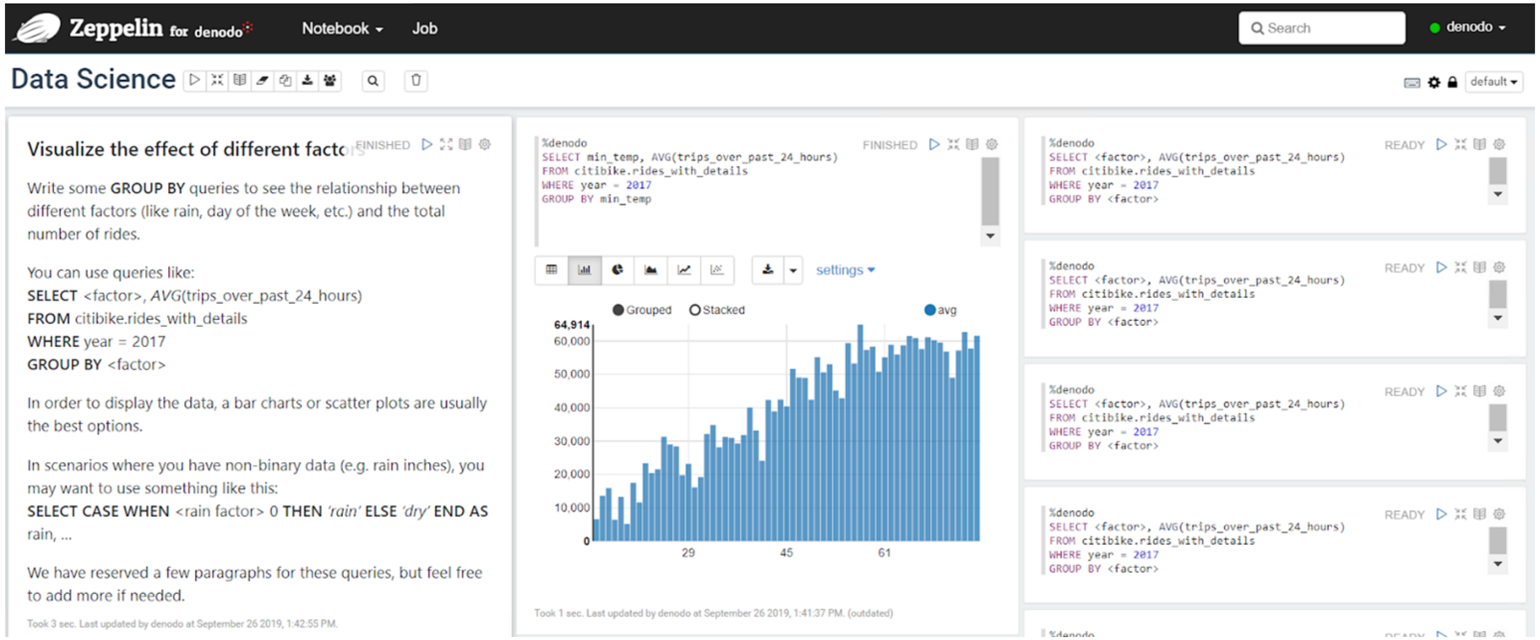
Dimensions
- Accountid
- Closedate
- Createdbyid
- Createddate
- Description
- Id
- Lastmodifieddate
- Leadsource
- Name
- Nextstep
- Ownerid
- Sforce Link
- Stagename
- Type
- *Measure Names*

Measures
- Amount
- Amount_Normalized__C
- Fiscalquarter
- Fiscalyear
- Probability
- *Number of Records*
- *Measure Values*

denodo
DATA CATALOG
Search
Summary
Query
Relationships
Table
File
Data
Worksheet
Dashboard
Story
Analysis
Map
Format
Server
Window
Help
Standard
Pages
Columns
Rows
Filters
Marks
Automatic
Color
Size
Text
Detail
Tooltip
Sheet 1
Drop field here
Drop field here
Drop field here

Denodo Notebook for Data Science

- Combine queries, scripts, text and graphics to build “narratives” and share results
- Based on Apache Zeppelin
- Denodo users can create, save, and share their own notebooks
- Fully integrated with Denodo’s security system and SSO capabilities
- Accessible from the catalog



Data Science

Visualize the effect of different factors

Write some **GROUP BY** queries to see the relationship between different factors (like rain, day of the week, etc.) and the total number of rides.

You can use queries like:
SELECT <factor>, *AVG*(trips_over_past_24_hours)
FROM citibike.rides_with_details
WHERE year = 2017
GROUP BY <factor>

In order to display the data, a bar charts or scatter plots are usually the best options.

In scenarios where you have non-binary data (e.g. rain inches), you may want to use something like this:
SELECT CASE WHEN <rain factor> 0 **THEN** *'rain'* **ELSE** *'dry'* **END AS** rain, ...

We have reserved a few paragraphs for these queries, but feel free to add more if needed.

Took 3 sec. Last updated by denodo at September 26 2019, 1:42:55 PM.

```
%denodo
SELECT min_temp, AVG(trips_over_past_24_hours)
FROM citibike.rides_with_details
WHERE year = 2017
GROUP BY min_temp
```

Took 1 sec. Last updated by denodo at September 26 2019, 1:41:37 PM. (outdated)

```
%denodo
SELECT <factor>, AVG(trips_over_past_24_hours)
FROM citibike.rides_with_details
WHERE year = 2017
GROUP BY <factor>
```

```
%denodo
SELECT <factor>, AVG(trips_over_past_24_hours)
FROM citibike.rides_with_details
WHERE year = 2017
GROUP BY <factor>
```

```
%denodo
SELECT <factor>, AVG(trips_over_past_24_hours)
FROM citibike.rides_with_details
WHERE year = 2017
GROUP BY <factor>
```

```
%denodo
SELECT <factor>, AVG(trips_over_past_24_hours)
FROM citibike.rides_with_details
WHERE year = 2017
GROUP BY <factor>
```

Customer Success Story

Democratizing Data @PROLOGIS

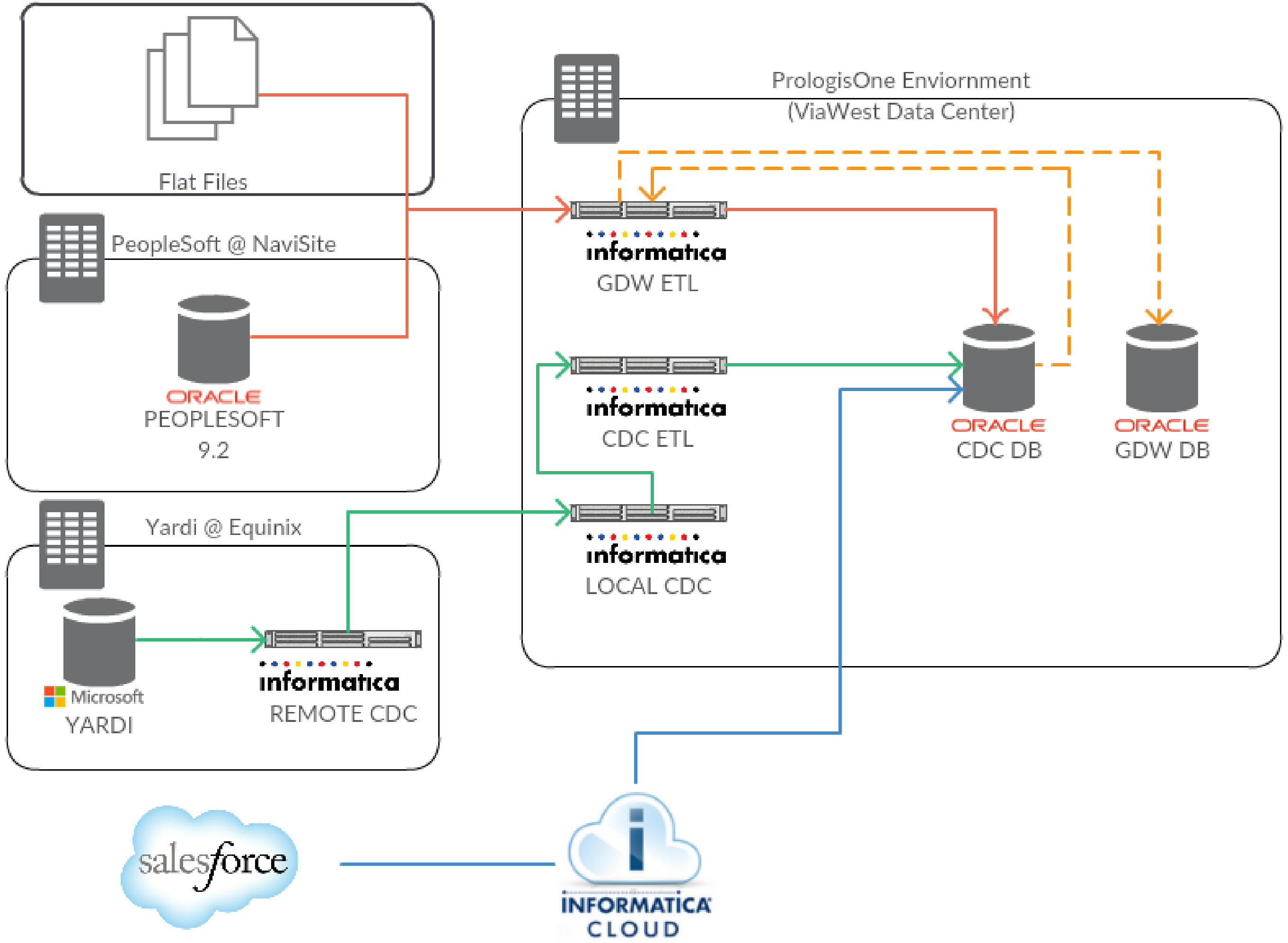
Logical Data Fabric Architecture @Prologis

Original State Architecture

27 Servers

- 8 Database
- 12 Integration
- 5 Reporting
- 2 Virtual Desktop

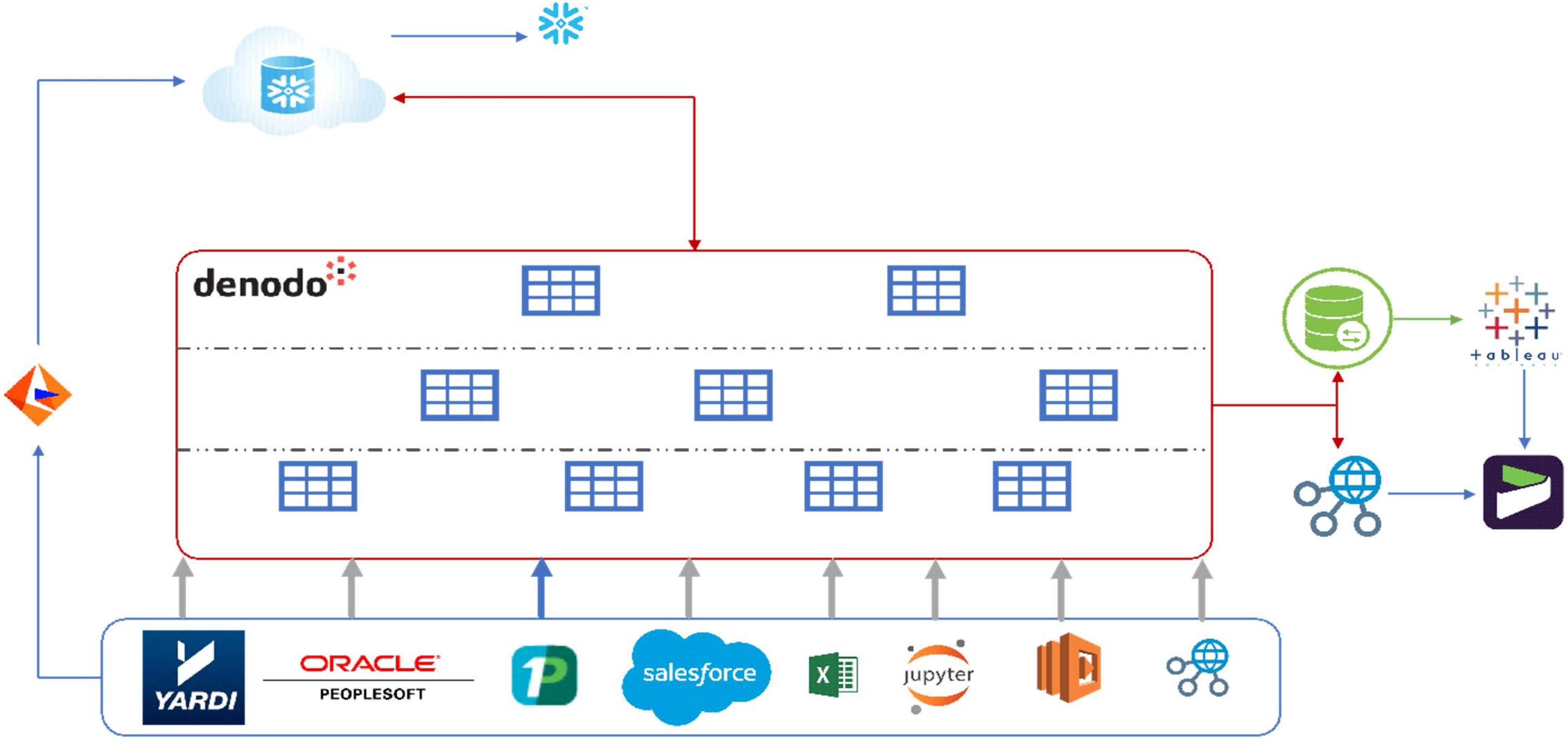
4 Environments


Flat Files
PeopleSoft @ NaviSite
ORACLE PEOPLESOFT 9.2
Yardi @ Equinix
Microsoft YARDI
informatica REMOTE CDC
PrologisOne Enviornment (ViaWest Data Center)
informatica GDW ETL
informatica CDC ETL
informatica LOCAL CDC
ORACLE CDC DB
ORACLE GDW DB
salesforce
INFORMATICA CLOUD

denodo

DATA VIRTUALIZATION
Packed Lunch Webinar Series

Current State Architecture


denodo
YARDI
ORACLE
PEOPLESOFT
salesforce
jupyter
tableau

Building Blocks of Analytics

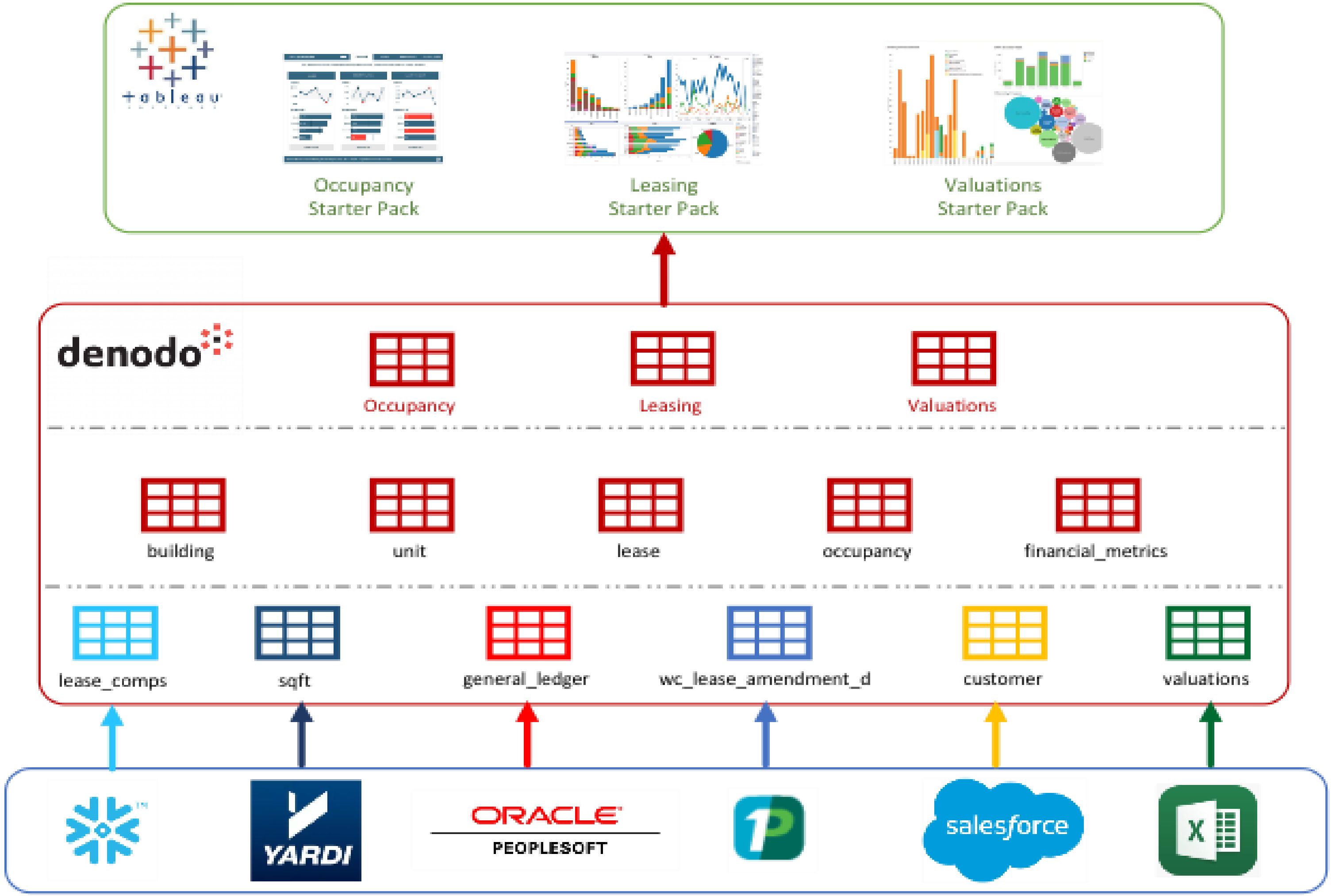
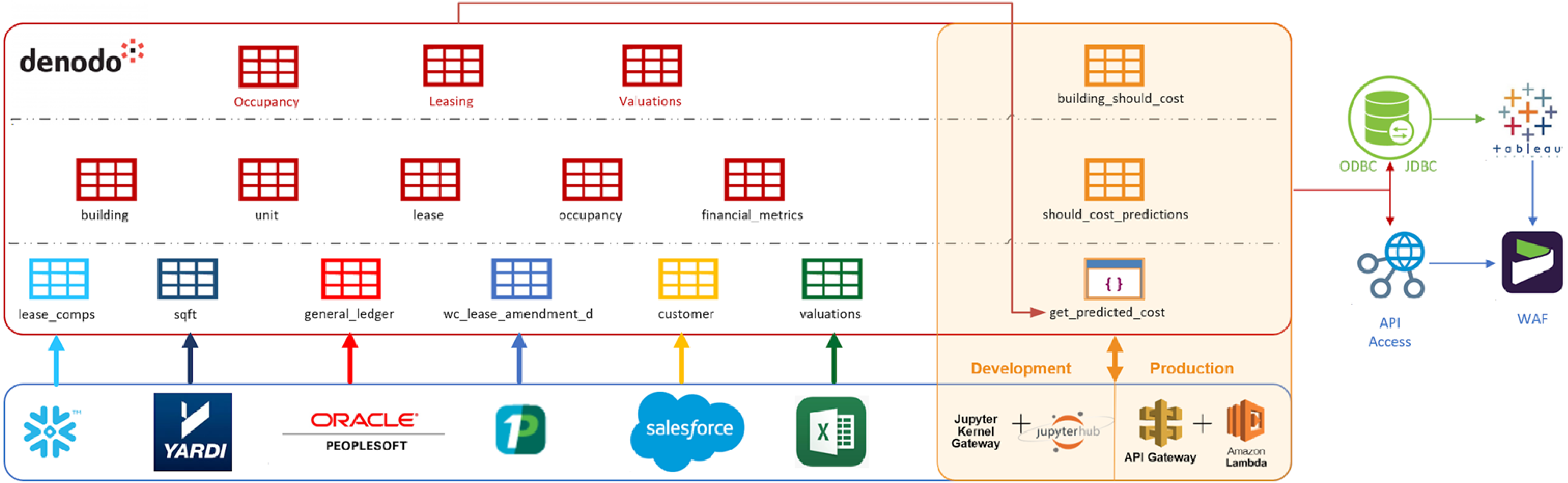


tableau
Occupancy Starter Pack
Leasing Starter Pack
Valuations Starter Pack
denodo
Occupancy
Leasing
Valuations
building
unit
lease
occupancy
financial_metrics
lease_comps
sqft
general_ledger
wc_lease_amendment_d
customer
valuations
YARDI
ORACLE PEOPLESOFT
salesforce

Data Science ToolKit


denodo
Occupancy
Leasing
Valuations
building_should_cost
building
unit
lease
occupancy
financial_metrics
should_cost_predictions
lease_comps
sqft
general_ledger
wc_lease_amendment_d
customer
valuations
get_predicted_cost
Development
Production
YARDI
ORACLE
PEOPLESOFT
salesforce
Jupyter Kernel Gateway
jupyterhub
API Gateway
Amazon Lambda
ODBC
JDBC
tableau
API Access
WAF

Key Takeaways

Logical Data Fabric architectures based on data virtualization are crucial for agile and governed self-service data delivery

An active Data Catalog on top of the Logical Data Fabric enables a ‘Data Marketplace’ where citizen analysts can ‘shop’ for data

denodo

DATA VIRTUALIZATION
Packed Lunch Webinar Series

Q&A

Get Started Today

Try the Denodo Standard 30-day free trial in the cloud marketplaces

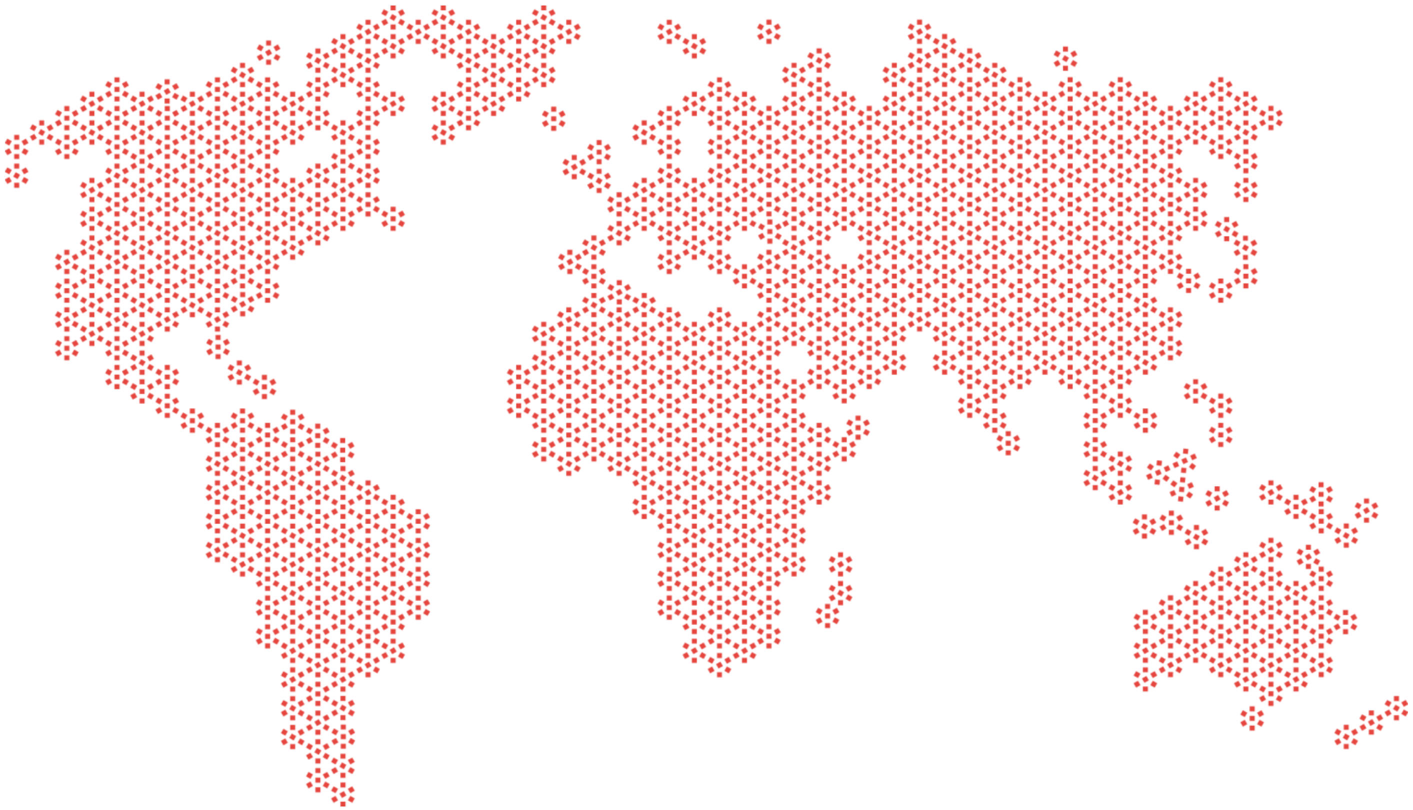
CHOICE
Under your cloud account

SUPPORT
Community forum AND remote sales engineer

OPPORTUNITY
30 minutes free consultation with Denodo Cloud specialist

denodo.com/free-trials

Thanks!



denodo | www.denodo.com | info@denodo.com

© Copyright Denodo Technologies. All rights reserved
Unless otherwise specified, no part of this PDF file may be reproduced or utilized in any for or by any means, electronic or mechanical, including photocopying and microfilm, without prior the written authorization from Denodo Technologies.