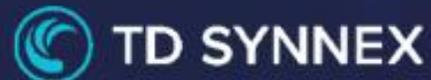


Diamante



Platina



DISCOVER

[No Title]
Ouro



VERTICA
by opentext™

Prata

TRACES



Apoio

FIAP



Data Mesh

Uma nova abordagem para expandir o valor dos dados

Data Mesh

Uma nova abordagem para expandir o valor dos dados

Luciano Verissimo

Master Principal Solution Engineer – Oracle Brasil



+20 anos de experiência em gestão de dados, modernização de aplicações, e desenho de soluções que visam ajudar os clientes da Oracle a alcançarem seus objetivos estratégicos através de inovações em armazenamento, processamento, exploração e segurança dos dados.



Por que é difícil
ser orientado a dados?

Alguns desafios de uma empresa que deseja ser guiada por dados

Como ser ágil na exploração dos dados?

Como reduzir o tempo dos projetos?

Como responder e agir em tempo real?

Como reduzir as barreiras entre os times?

Como aumentar o reuso de conhecimento?

Como ter uma visão concisa e unificada das coisas?

Como lidar com repositórios descentralizados?

Como garantir a consistência das transações de negócio?

Como alcançar conhecimento relevante e correto?

Empresa guiada por dados



75%

Algumas informações importantes

“75% das empresas não conseguem se tornar guiadas por dados”

[2] – New Vantage Partners Big Data and AI Executive Survey 2021



50%

“O trabalho de integração representará 50% do tempo e custo de construção de uma plataforma digital”

[1] – Gartner - Use a Hybrid Integration Approach to Empower Digital Transformation

Arquitetura Data Mesh

Abordagem alternativa para vencer estes desafios:



Trend No. 6: Data Fabric

Data fabric enables frictionless access and sharing of data in a distributed data environment. It enables a single and consistent data management framework, which allows seamless data access and processing by design across otherwise siloed storage.

Through 2022, bespoke data fabric designs will be deployed primarily as a static infrastructure, forcing organizations into a new wave of cost to completely re-design for more dynamic data mesh approaches.



REPORT

Data Mesh Makes Connected Intelligence Experiences A Reality

Explainable, Trusted, Secure, And Scaled AI Is Possible With The Right Data Architecture

July 7th, 2021



DATA INNOVATION SUMMIT 2021

Panel on Data Fabric, Data Mesh and other trends in Data Management

3 months ago 1 min read

towards
data science

What is a Data Mesh — and How Not to Mesh it Up

A beginner's guide to implementing the latest industry trend: a data mesh.

Barr Moses Jul 28, 2020 · 9 min read ★

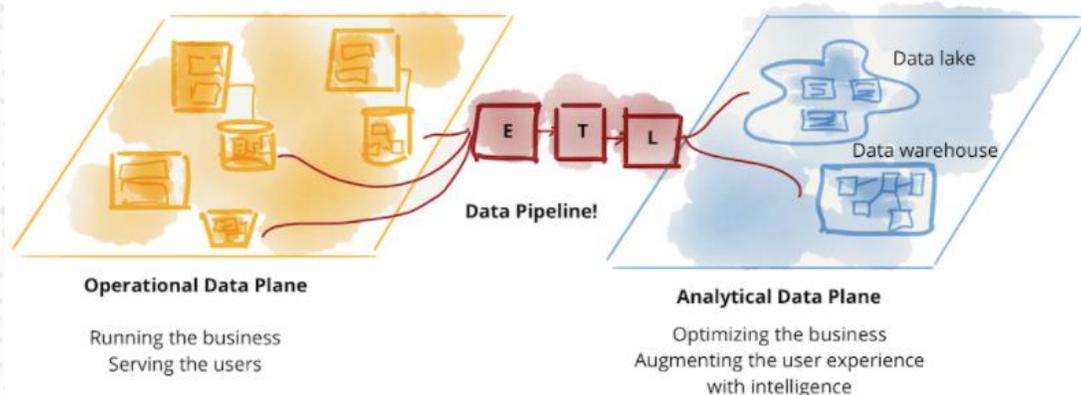


Mas antes de entrarmos
no Data Mesh...

Uma Jornada Analítica típica de muitas empresas

Jornada analítica dos clientes

Estágio 1 – DW e Data Lake Padrão



Cenário Comum

- A TI possui times generalistas responsáveis pela coleta, preparação e disponibilização de dados para análise.
- Basicamente temos:
 - DW para **dados estruturados e relatórios de BI**
 - Data Lake para **dados não estruturados e suporte à ciência de dados (preparação para ML, IA, etc.)**

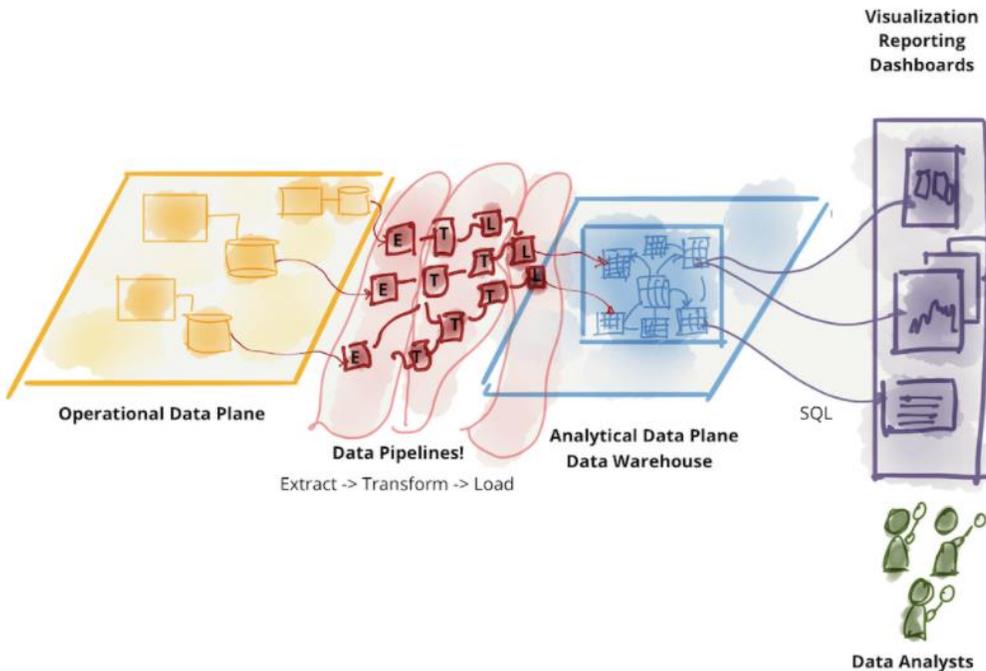
Algumas Dificuldades

- A falta de padronização e documentação sobre a origem dos dados compromete a qualidade
Exemplos: o que é ser cliente? Quais dados devo coletar sobre ele? Que formato de dados usar?
- Normalmente há pouco conhecimento sobre o negócio, reduzindo o valor prático dos dados preparados



Jornada analítica dos clientes

Estágio 2 – Adicionando significado



Cenário Comum

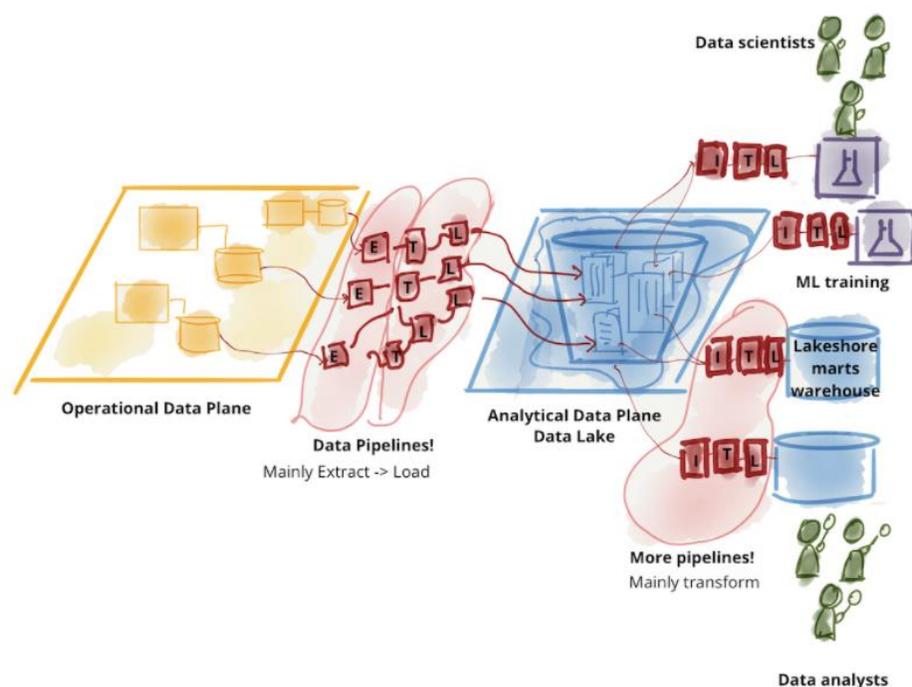
- Busca ser mais efetiva para o negócio
- Analistas de dados pressionam por mais dados
- Maior quantidade de coleta e transformação
- Eventualmente existem Data Marts e Sand-Boxes (ideia de descentralização)
- Foco na análise descritiva

Algumas Dificuldades

- Gestão de centenas ou até milhares de processos de ETL/ELT
- Ferramentas de transformação muitas vezes operadas por técnicos que não conhece a fundo a operação do negócio
- Problemas de qualidade e linhagem
- Projetos mais lentos. TI normalmente é o gargalo.

Jornada analítica dos clientes

Estágio 3 – Expandindo casos de uso



Cenário Comum

- Foco na análise preditiva
- Usa-se mais técnicas de análise (ex: ML, IA)
- Demanda novas tarefas, novos cargos, novas equipes

Algumas Dificuldades

- Silos de dados e conhecimento (normalmente por área de negócio)
- Áreas de negócio buscando as mesmas respostas
- Resultados inconstantes entre áreas para o mesmo assunto
- Difícil escalar para nível corporativo (baixo reuso de conhecimento)
- Os sistemas da empresa obtêm pouco benefício do trabalho analítico

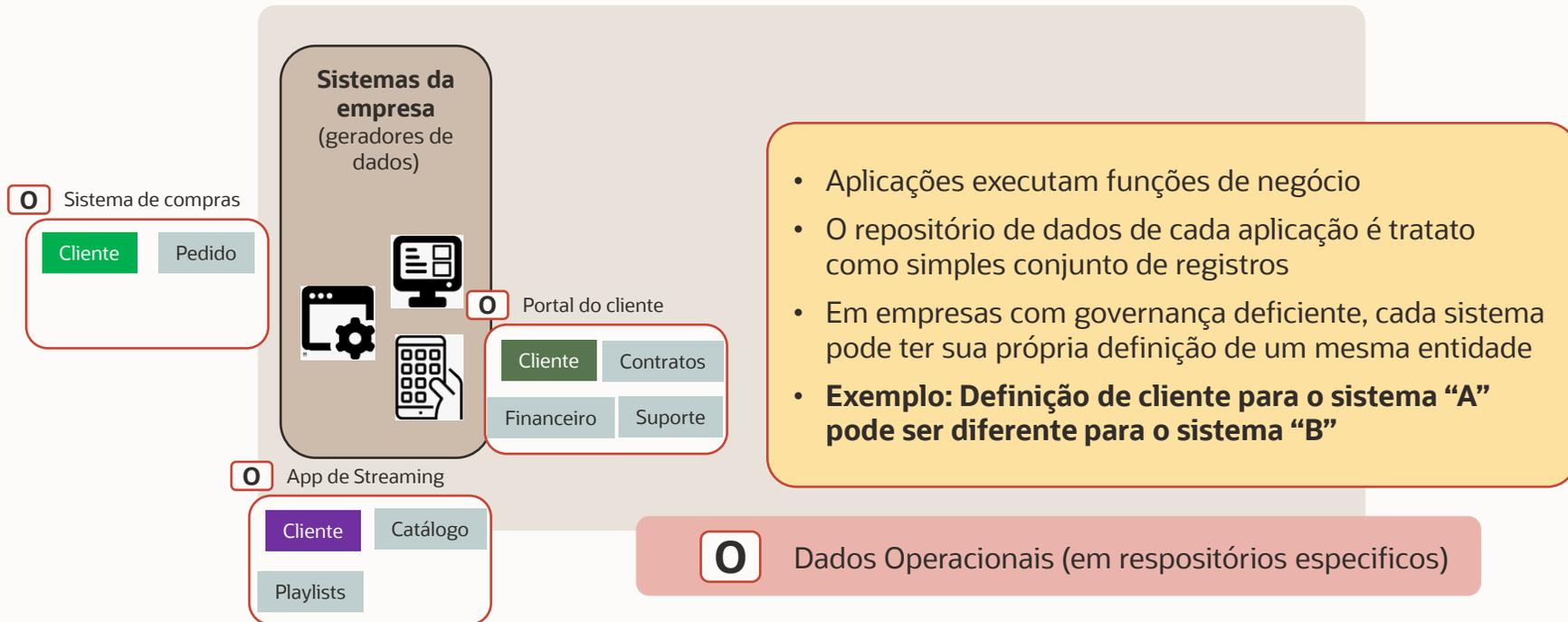
Com expandir daqui?

Empresa guiada por Dados



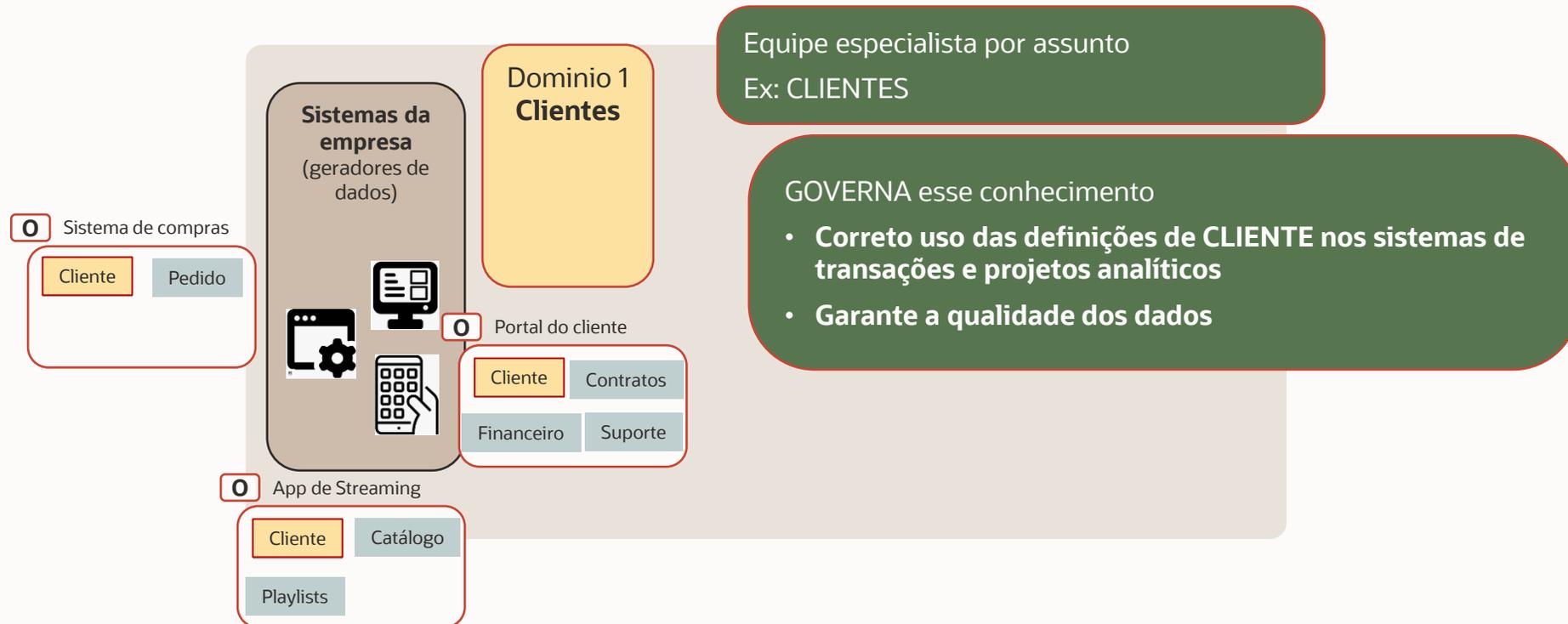
Introdução ao Data Mesh

Como os dados são normalmente tratados?



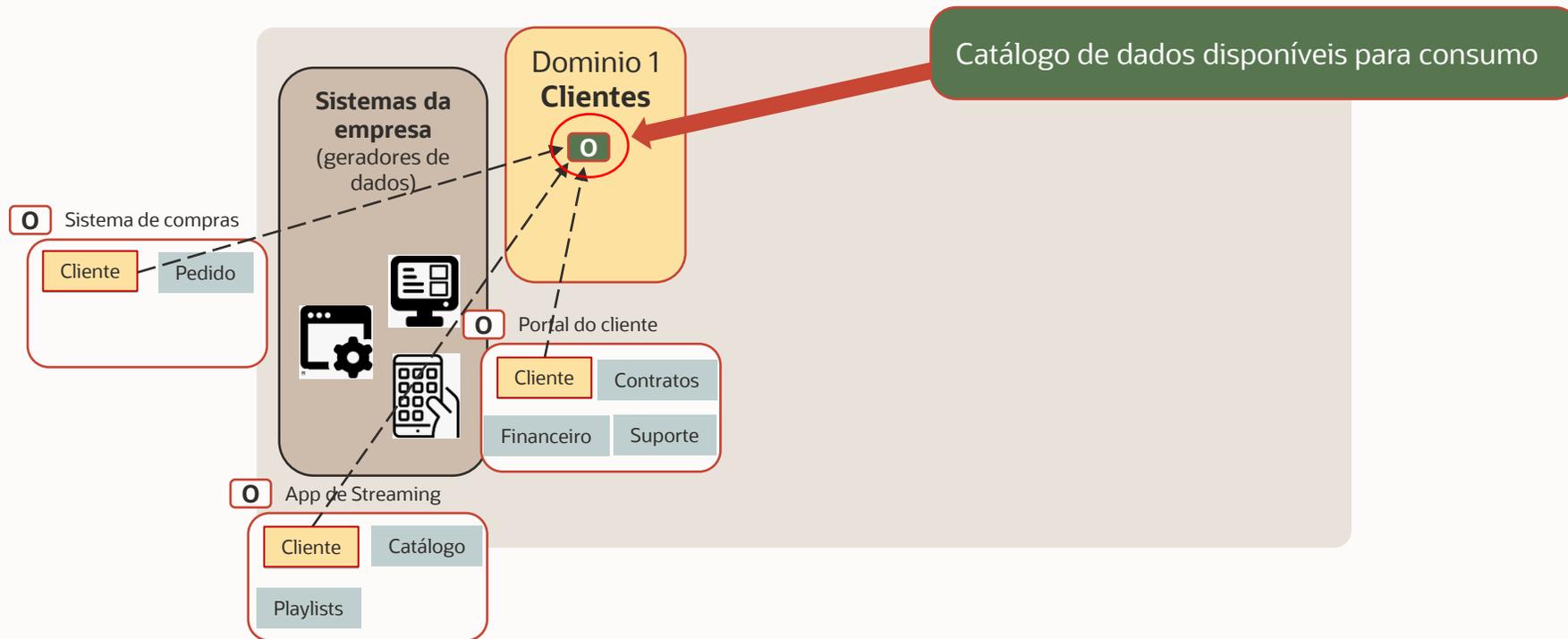
E se criássemos....

— Um conceito de Dominio de Dados

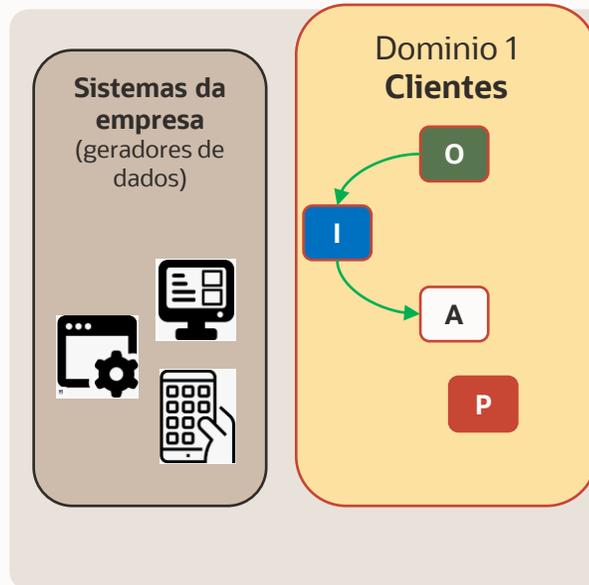


E se criássemos....

— Um conceito de Dominio de Dados



Explorando os dados dentro do Dominio



- Acessar dados diretamente na origem
- Interceptar transações e eventos (CDC - CI/CD)
 - Uso imediato (tempo real)
 - Uso posterior
- Usa ferramentas de ETL / ETL
- Realiza transformações e enriquecimento
- Usar diversos tipos e modelos de dados

Criar produtos de dados

Exemplos

- Uma API Rest que responde a visão única do cliente
- Uma tabela para exploração do perfil dos clientes
- Um algoritmo de ML para classificar os clientes
- Um grafo para descobrir comunidades

O Dados Operacionais

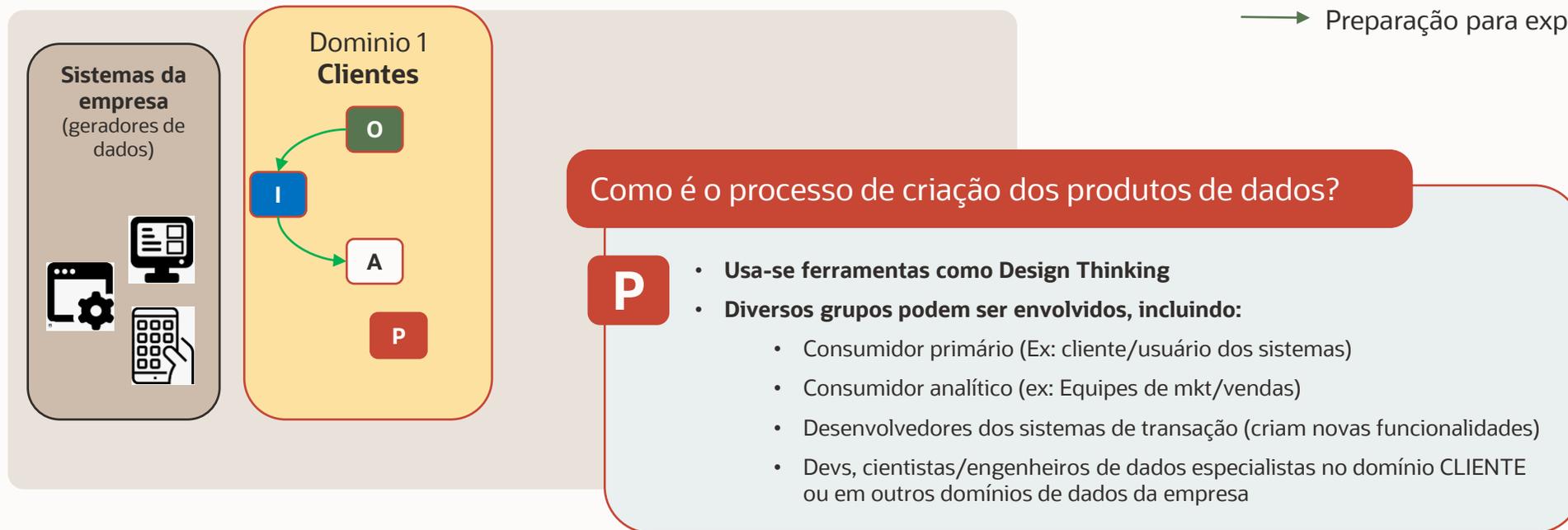
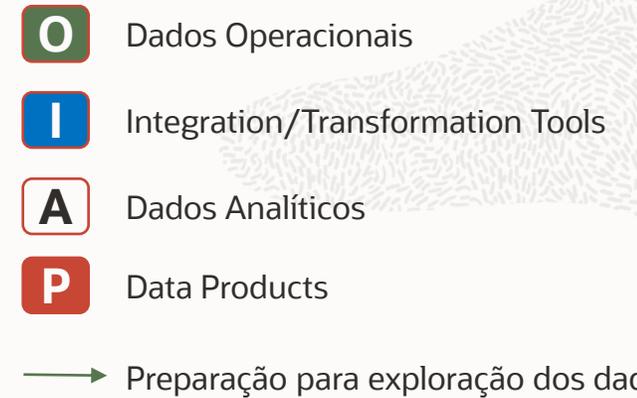
I Integration/Transformation Tools

A Dados Analíticos

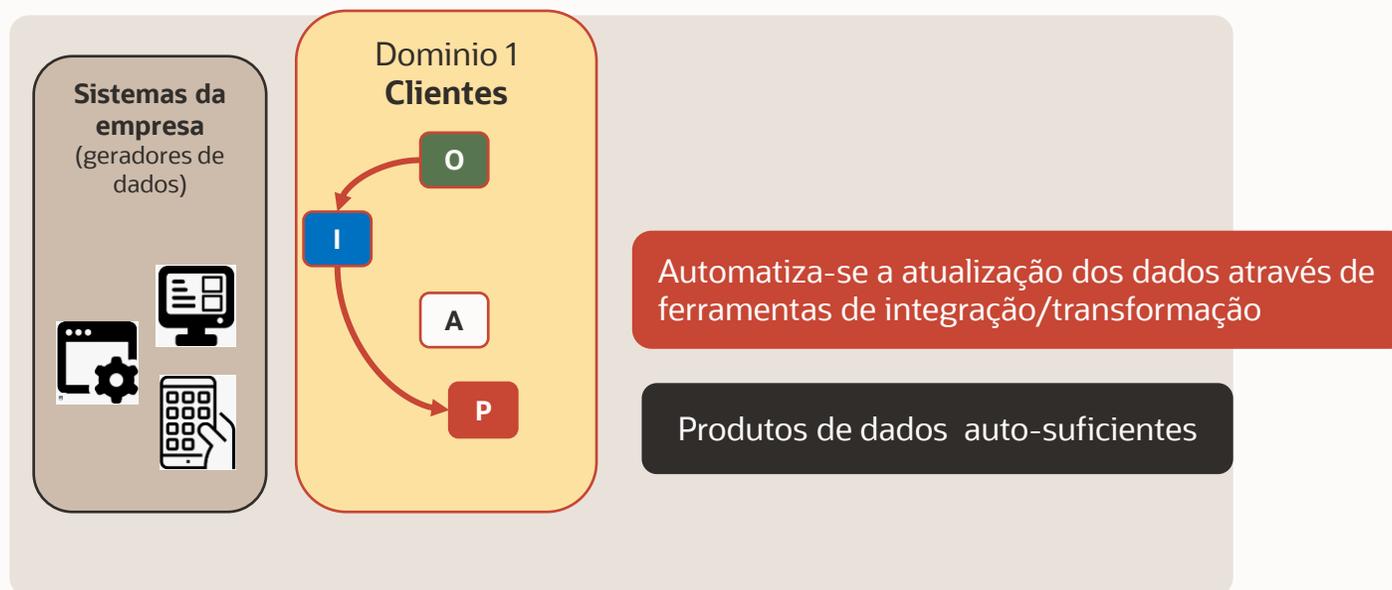
P Data Products

→ Preparação para exploração dos dados

Construindo produtos de dados



Atualização dos produtos de dados

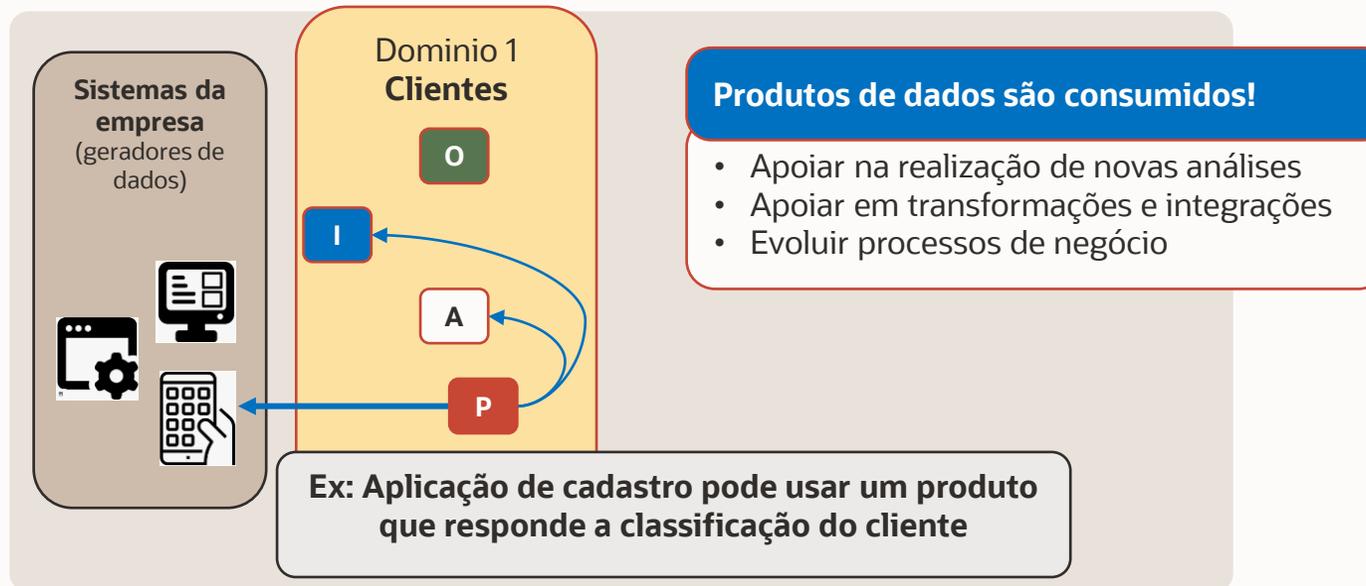


- O** Dados Operacionais
- I** Integration/Transformation Tools
- A** Dados Analíticos
- P** Data Products

- Preparação e Exploração dos dados
- Atualização dos produtos de dados



Produtos desejados pelos consumidores



O Dados Operacionais

I Integration/Transformation Tools

A Dados Analíticos

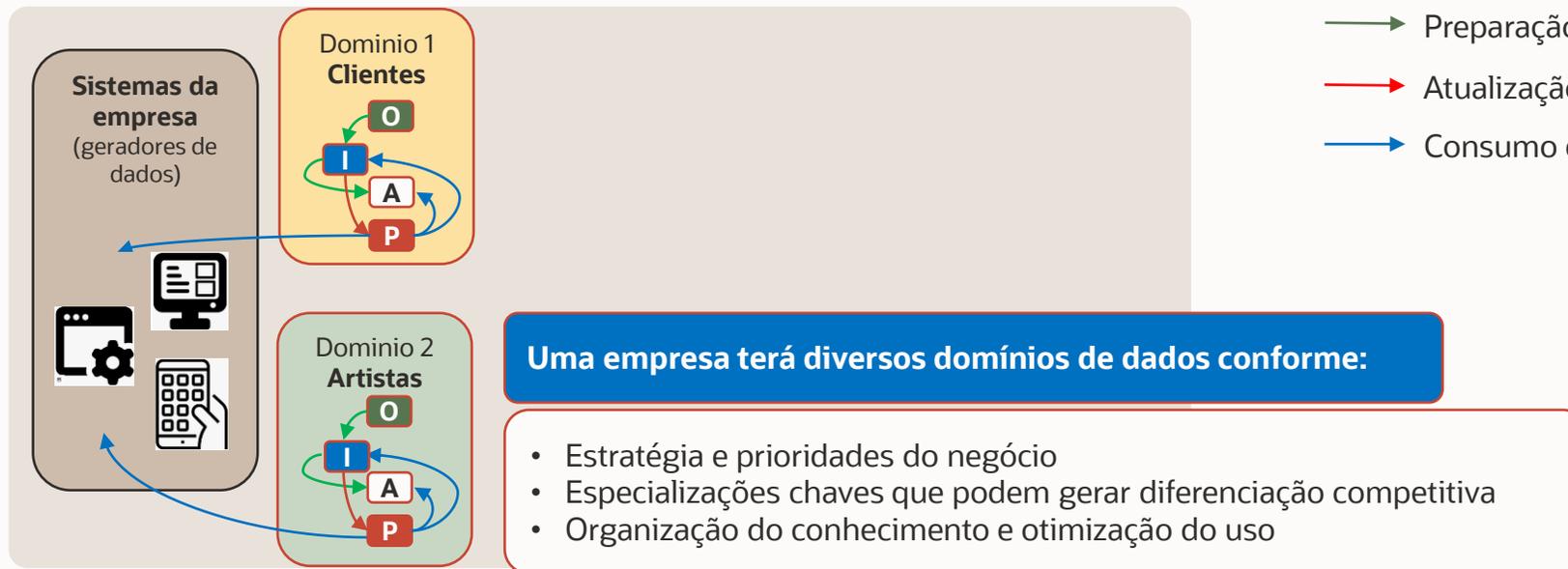
P Data Products

→ Preparação e Exploração dos dados

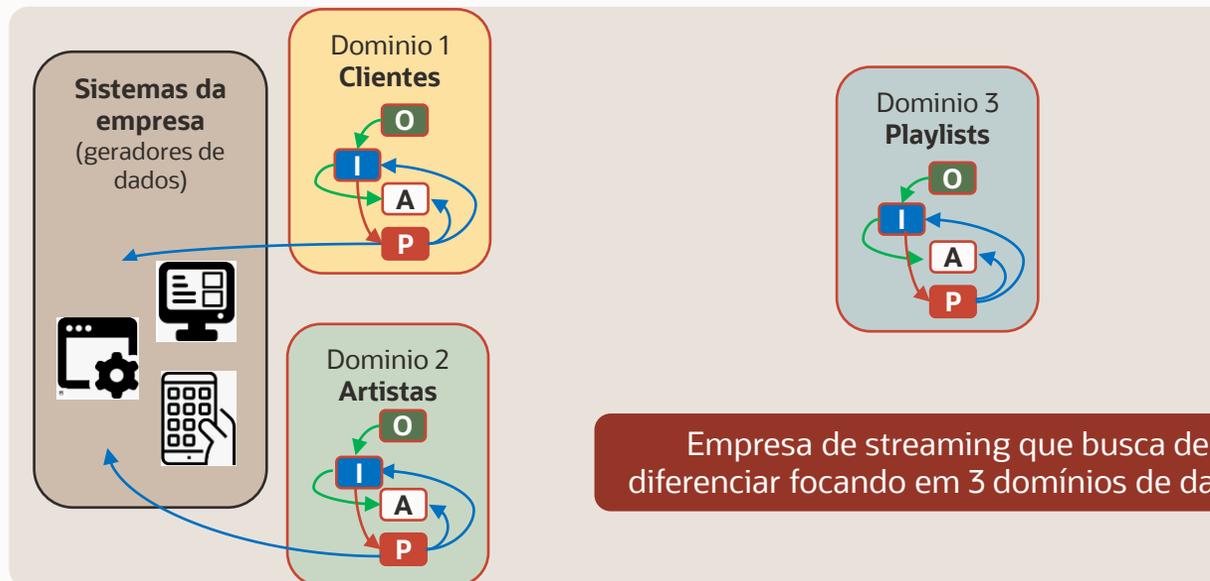
→ Atualização dos produtos de dados

→ Consumo do produto

Multiplos domínios de dados



Multiplos domínios de dados



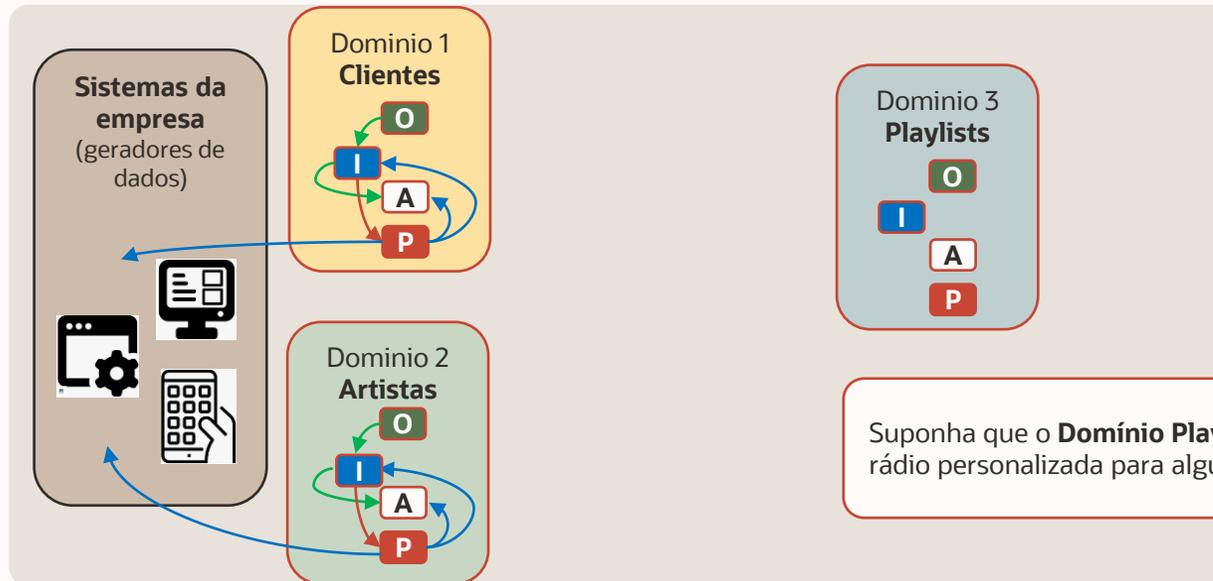
- O** Dados Operacionais
- I** Integration/Transformation Tools
- A** Dados Analíticos
- P** Data Products
- Preparação e Exploração dos dados
- Atualização dos produtos de dados
- Consumo do produto



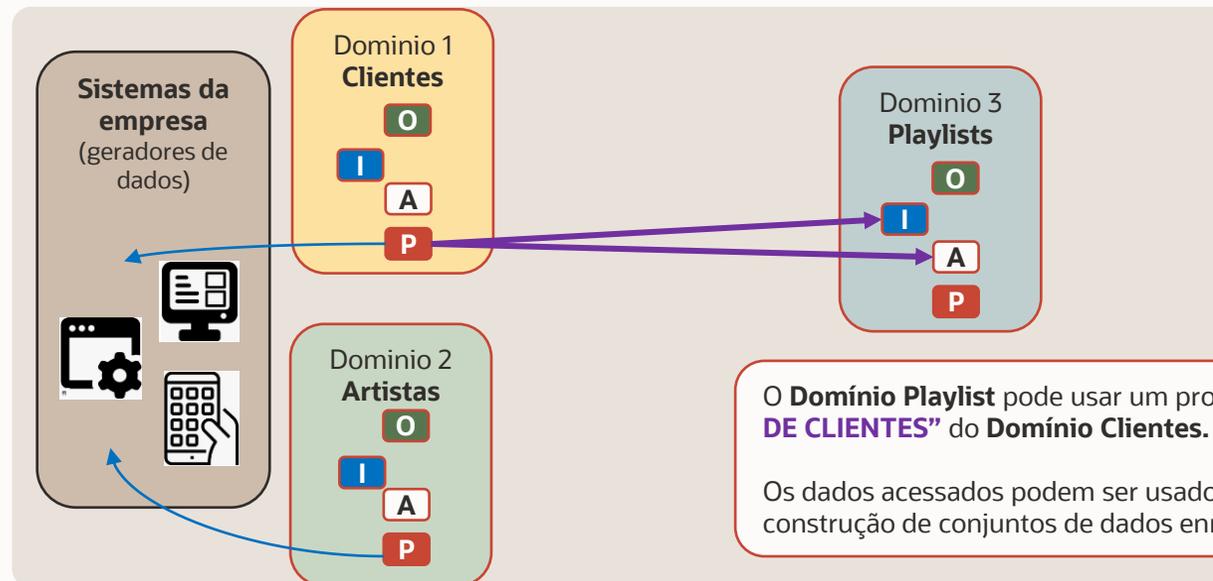
Multiplos domínios de dados

- O** Dados Operacionais
- I** Integration/Transformation Tools
- A** Dados Analíticos
- P** Data Products

- Preparação e Exploração dos dados
- Atualização dos produtos de dados
- Consumo do produto



Produtos de dados consumidos entre domínios



- O** Dados Operacionais
- I** Integration/Transformation Tools
- A** Dados Analíticos
- P** Data Products

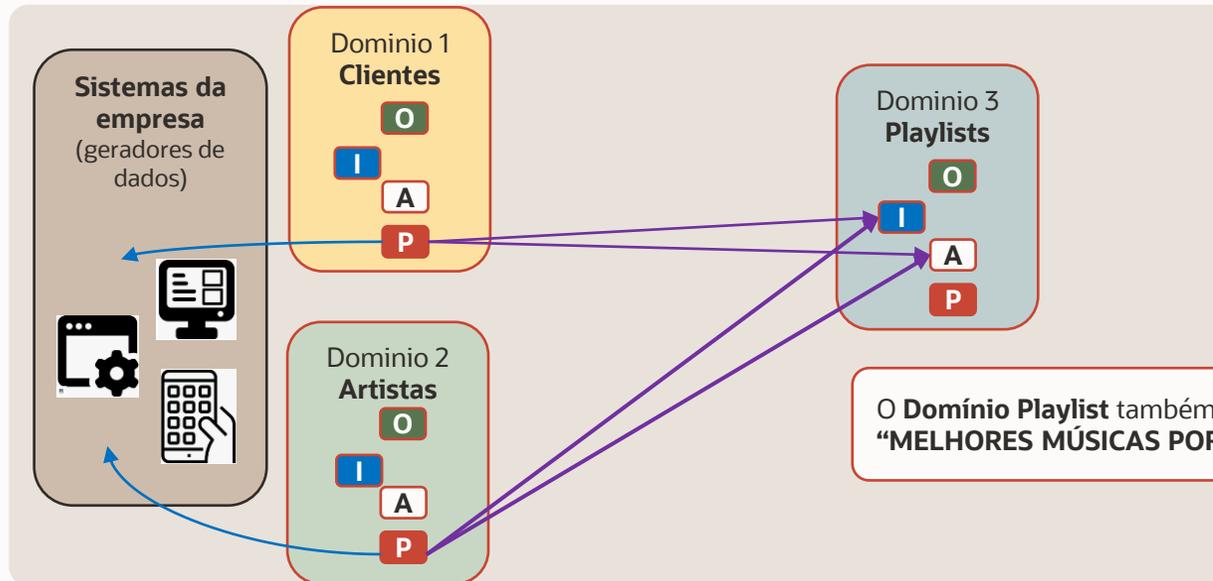
- Preparação e Exploração dos dados
- Atualização dos produtos de dados
- Consumo do produto
- Consumo entre domínios



Produtos de dados consumidos entre domínios

- O** Dados Operacionais
- I** Integration/Transformation Tools
- A** Dados Analíticos
- P** Data Products

- Preparação e Exploração dos dados
- Atualização dos produtos de dados
- Consumo do produto
- Consumo entre domínios



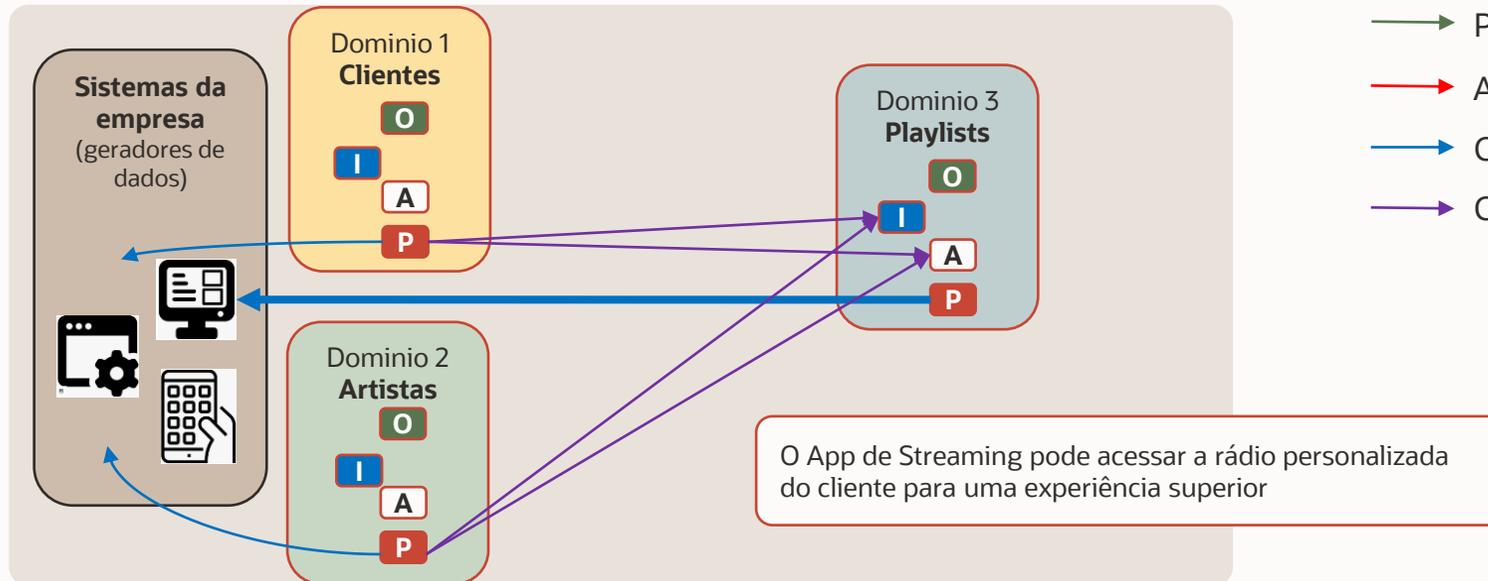
O Domínio Playlist também pode usar um produto de dados chamado "MELHORES MÚSICAS POR GENERO" do Domínio Artistas.



Aplicações consumindo produtos de dados

- O** Dados Operacionais
- I** Integration/Transformation Tools
- A** Dados Analíticos
- P** Data Products

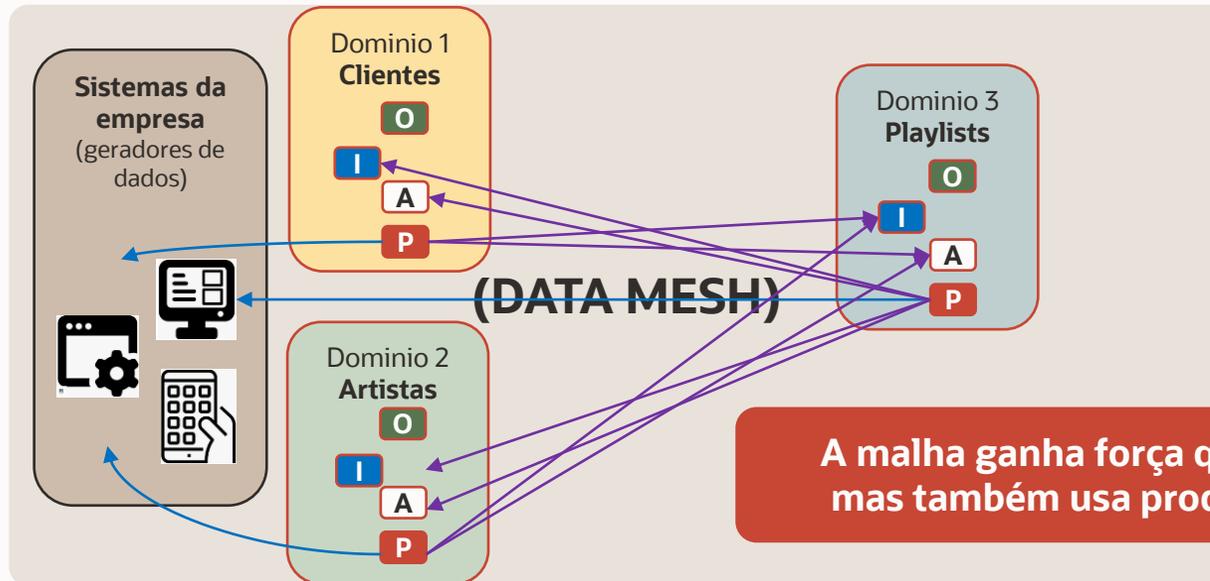
- Preparação e Exploração dos dados
- Atualização dos produtos de dados
- Consumo do produto
- Consumo entre domínios



Expansão do consumo dos produtos de dados

- O** Dados Operacionais
- I** Integration/Transformation Tools
- A** Dados Analíticos
- P** Data Products

- Preparação e Exploração dos dados
- Atualização dos produtos de dados
- Consumo do produto
- Consumo entre domínios

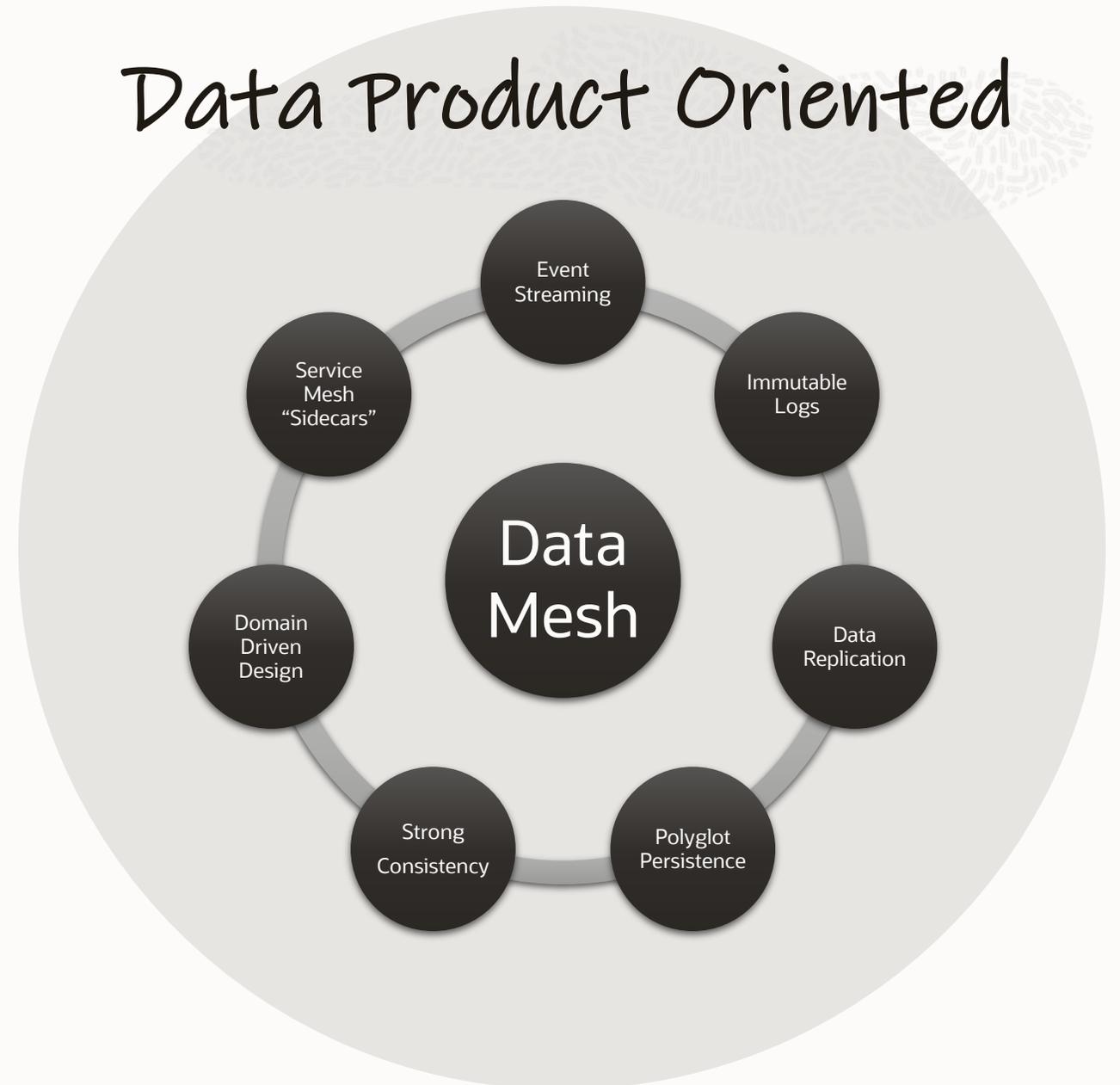


Fechando as ideias por trás do Data Mesh

Data Mesh

- Entrega **maior valor dos dados** e **reuso de conhecimento** por toda a companhia
- Melhor **governança e especialização** por assuntos de conhecimento (**Domínios**)
- Criação **produtos de dados**, valiosos e estratégicos
- **Entrega agilidade** aos times através de **automação e simplificação** no uso dos dados
- Permite usar **persistência poliglota** (múltiplos tipos e modelos de dados)
- **Transações de negócio e eventos monitorados** são as principais fontes de dados (pipeline)
- Apoia-se no **processamento de dados e na análise em tempo real**

Data Product Oriented



Soluções de Apoio

Mesh de serviço descentralizado (para integração de dados):

- Abordagem CI/CD/Agile DevOps moderna aplicadas em ferramentas de integração
- Podem ser executadas em qualquer lugar como Edge ou Multi-cloud.

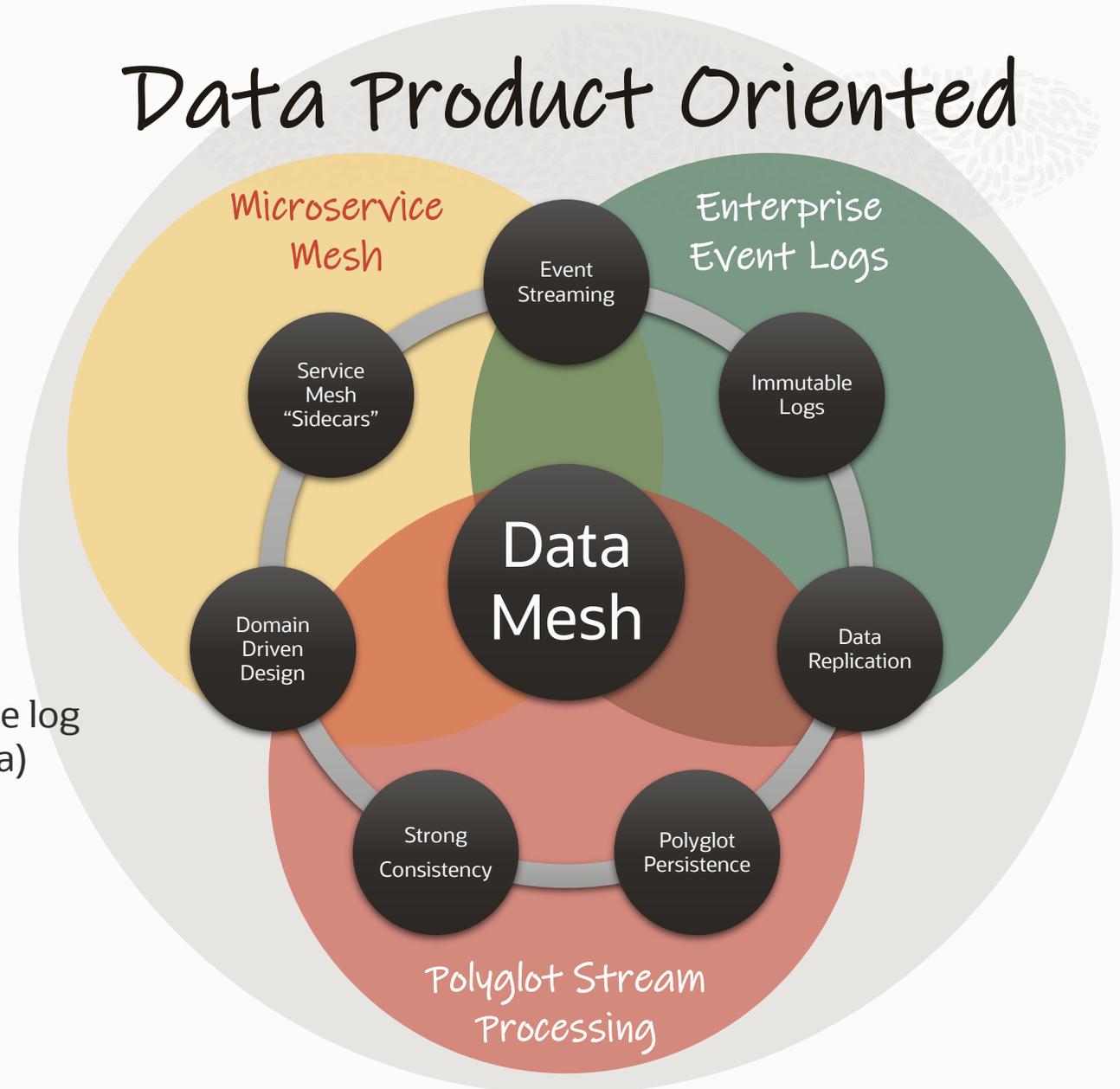
Logs de eventos corporativos (registro de todos os eventos de dados):

- Consumo contínuo a partir de aplicativos e eventos de log de banco de dados (infraestrutura moderna ou legada)

Processamento de fluxo poliglota (pipelines de produtos de dados):

- Filtre, transforme e mantenha dados confiáveis continuamente em produtos de dados consumíveis

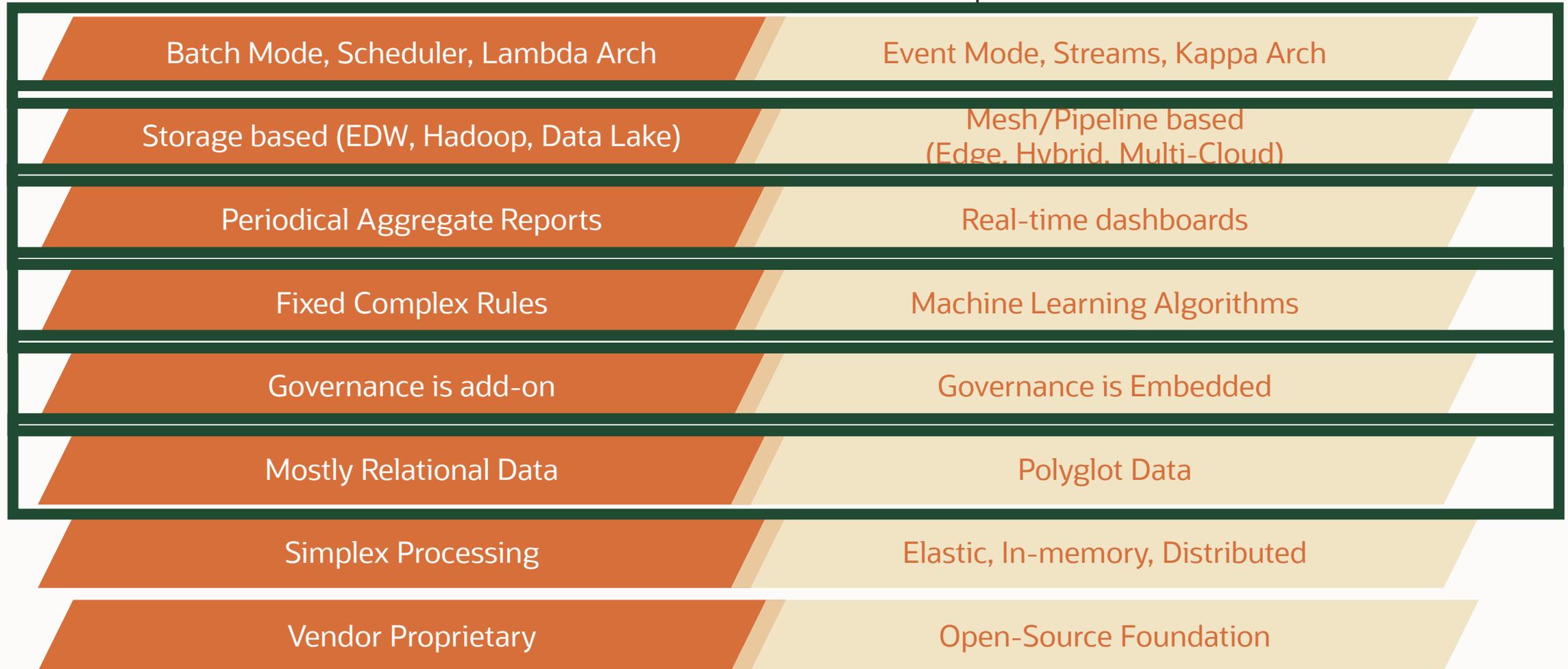
Data Product Oriented



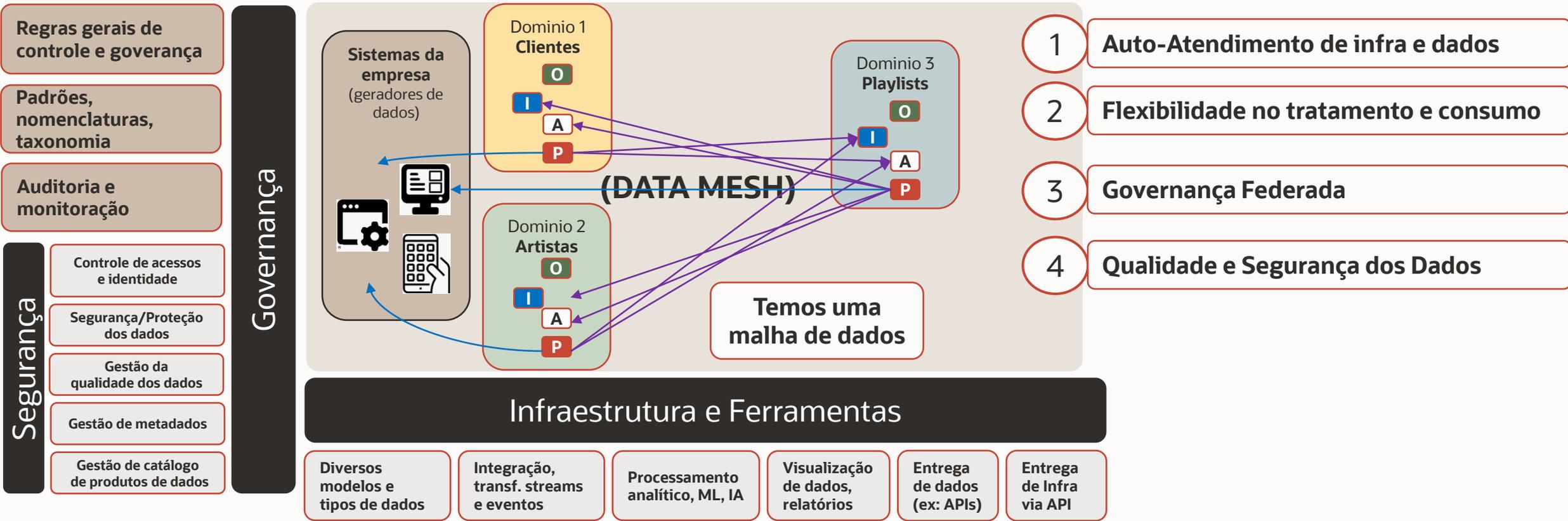
Malha de dados em tempo

... da Indústria 3.0

... para a Indústria 4.0



Princípios do Data Mesh (malha de dados) e Demandas Geradas



Gostaram do Data Mesh?

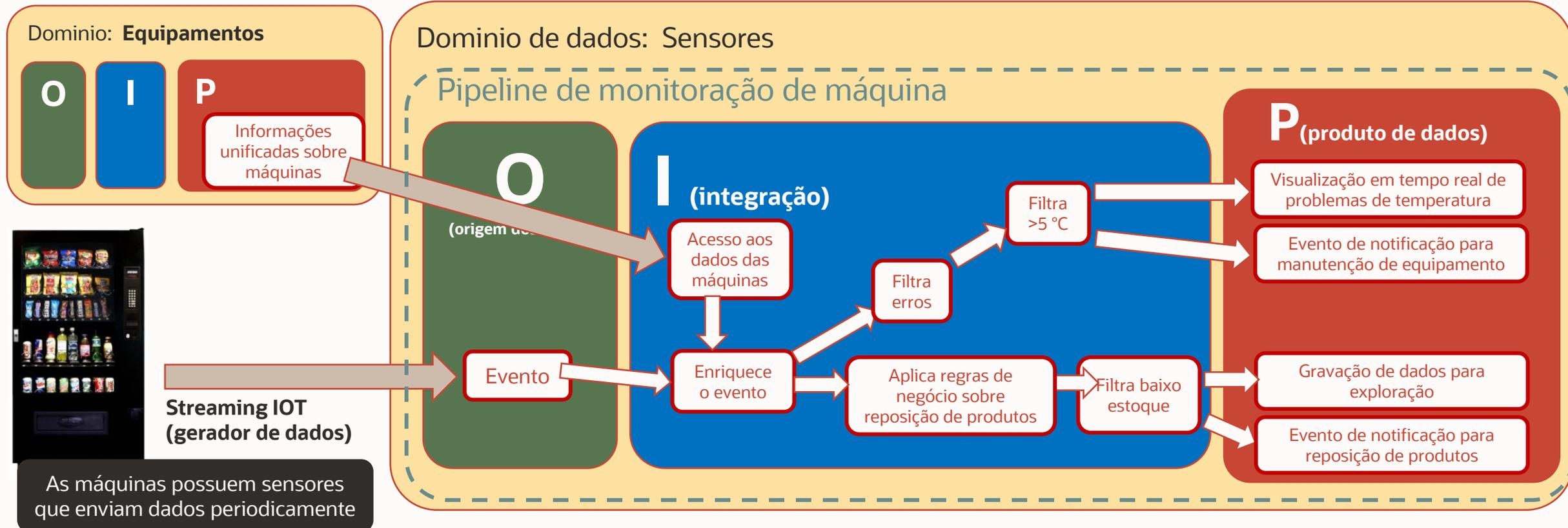
Exemplo Data Mesh (DEMO)

Análise e operação em tempo real

Empresa de venda de produtos através de máquinas de auto-atendimento

Objetivos de Negócio

- Garantir a qualidade dos produtos vendidos
- Não perder vendas por erros de operação dos equipamentos
- Não perder vendas por quebra de estoque nas máquinas



Empresa de venda de produtos através de máquinas de auto-atendimento

Objetivos de Negócio

- Garantir a qualidade dos produtos vendidos
- Não perder vendas por erros de operação dos equipamentos
- Não perder vendas por quebra de estoque nas máquinas

Dominio: Equipamentos

O I P

Informações unificadas sobre máquinas

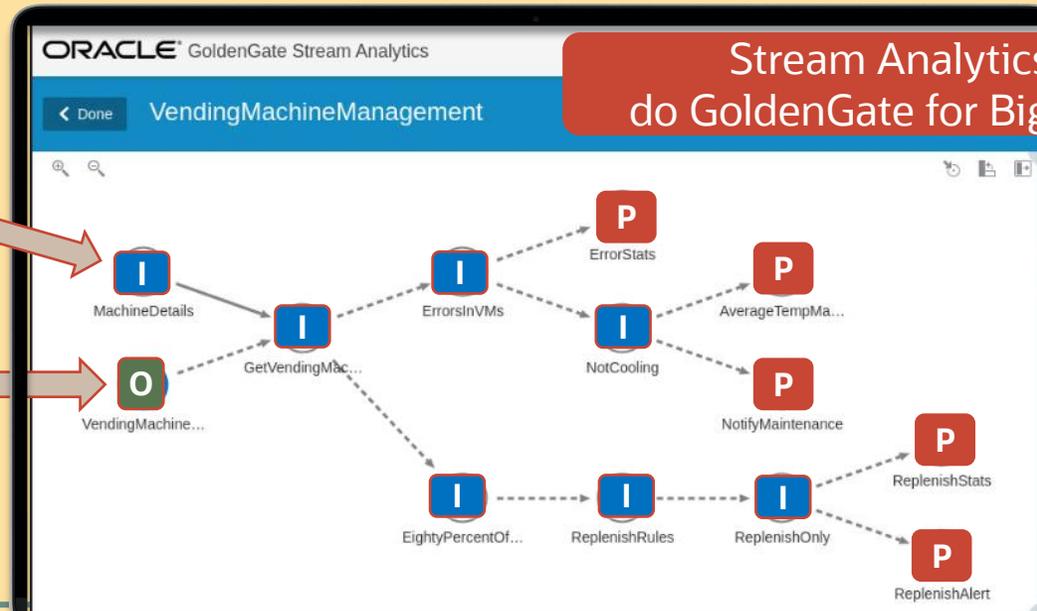


Streaming IOT (gerador de dados)

As máquinas possuem sensores que enviam dados periodicamente

Dominio de dados: Sensores

Pipeline de monitoração de máquina



Stream Analytics do GoldenGate for Big Data

Conheça: Oracle Live Lab

Recomendo: **Oracle GoldenGate Analytics With Lab Exercises**

Siga o Lab e criei um pipeline para implementar produtos de dados seguindo **princípios de Data Mesh**

The screenshot displays the Oracle LiveLabs website interface. At the top, the 'LiveLabs' logo is on the left, and a search bar contains the text 'Oracle GoldenGate Stream Analytics'. Below the search bar, there's a 'Clear Search & Filters' button and a 'Sort By' dropdown menu set to 'Most Popular'. A filter section shows 'Number of Workshops: 2' and 'Level' options: 'Beginner (2)', 'Intermediate', and 'Advanced'. Another filter section for 'Workshop Type' is partially visible. The main content area, titled 'Workshops and Sprints', features a card for 'Oracle GoldenGate Stream Analytics' with a database icon, a description 'Explore Oracle Stream Analytics with Lab Exercises.', a duration of '3 hrs', and '11278 Views'. To the right, another card titled 'Create a Pipeline in minutes with Stream Analytics' shows a duration of '1 hr'. A modal window is open over the second card, showing a 'Share' button, a 'Start' button, and two options: 'Run on Your Tenancy' and 'Run on LiveLabs Sandbox'. The modal text explains that the 'Run on Your Tenancy' button provides step-by-step instructions for running the workshop on a personal tenancy, while the 'Run on LiveLabs' button dynamically creates resources in an Oracle-owned tenancy for free. It also includes links for 'Oracle account help' and 'Oracle account signup'. A list of prerequisites is partially visible, including 'An Oracle Free T...'.

LiveLabs

Oracle GoldenGate Stream Analytics

Clear Search & Filters

Sort By
Most Popular

Number of Workshops: 2

Level

Beginner (2)

Intermediate

Advanced

Workshop Type

Workshops and Sprints

Oracle GoldenGate Stream Analytics

Explore Oracle Stream Analytics with Lab Exercises.

3 hrs 11278 Views

Create a Pipeline in minutes with Stream Analytics

This workshop shows you how to create a pipeline. Filter to it, and you'll see the details.

1 hr

3 hours

Outline

- Login and Navigation
- Streaming IoT Data
- Streaming Retail Data
- Publishing Dashboard

Prerequisites

- An Oracle Free Tenancy

The **Run on Your Tenancy** button provides step-by-step instructions so you can run this workshop on your personal tenancy!

Run on Your Tenancy

The **Run on LiveLabs** button will dynamically create resources in an Oracle-owned tenancy for you to use for free!

[Oracle account help](#) | [Oracle account signup](#)

Run on LiveLabs Sandbox

Use o ambiente SandBox gratuito que a Oracle já preparou para você.



Obrigado!

[Return Main](#)

