



Desvendando o Universo de Dados

19/05 21h Online

Centro de Tecnologias Digitais
Universidade FEEVALE

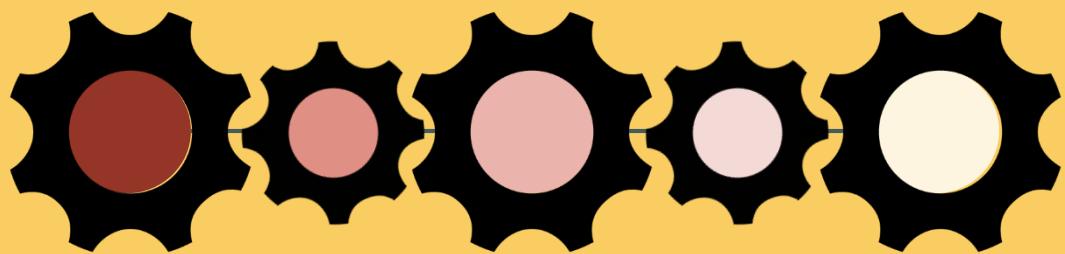
Lourenço Taborda
Oracle





Foto de Michel Stockman no Unsplash

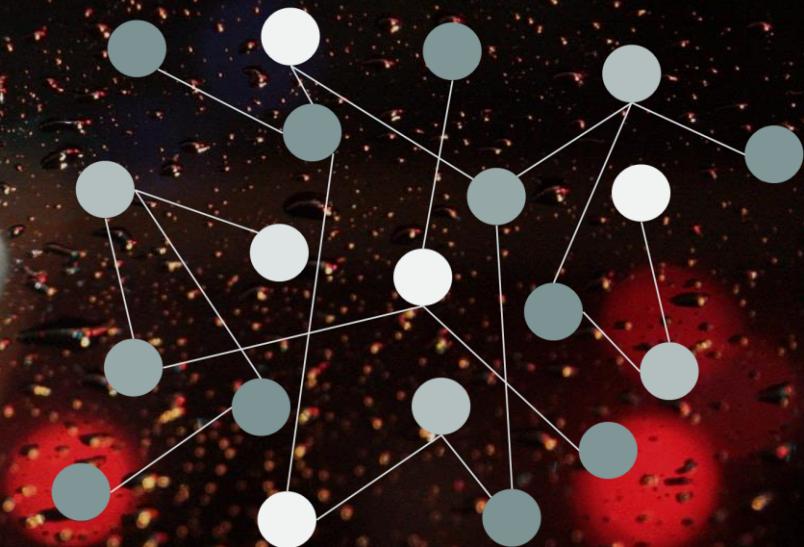
Terceira Revolução Industrial



Modelo de Negócio Linear

Projetar, Produzir, Distribuir, Consumir.

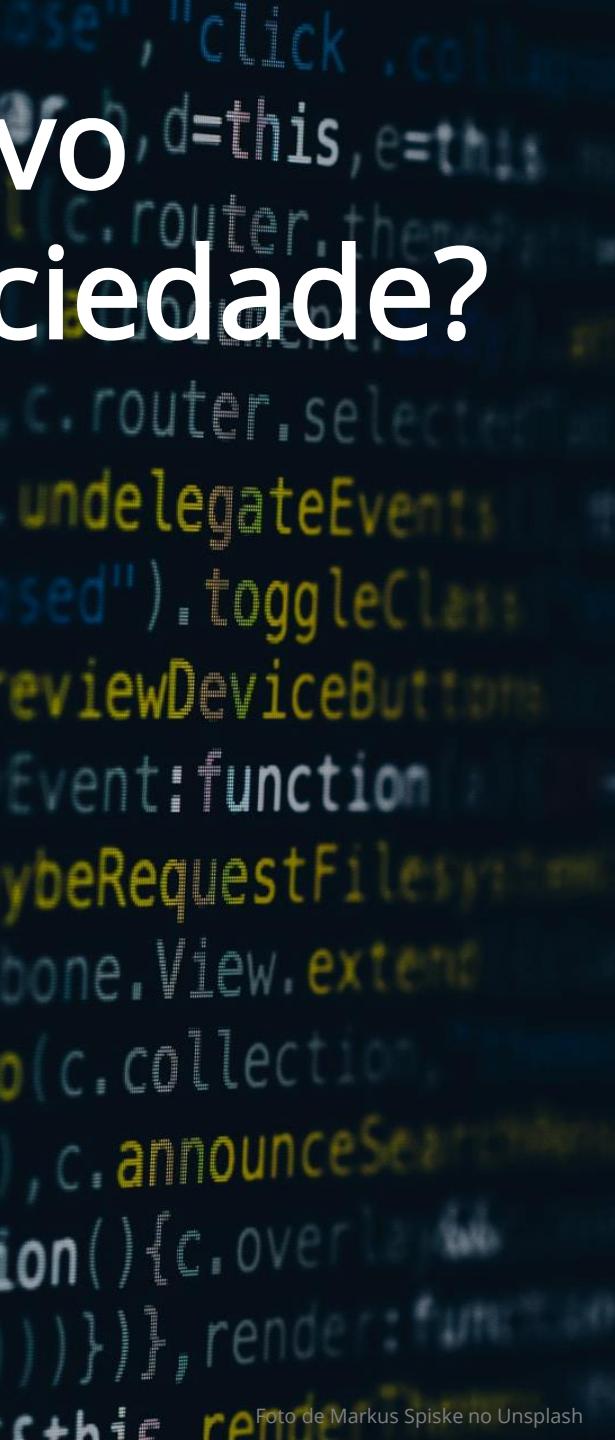
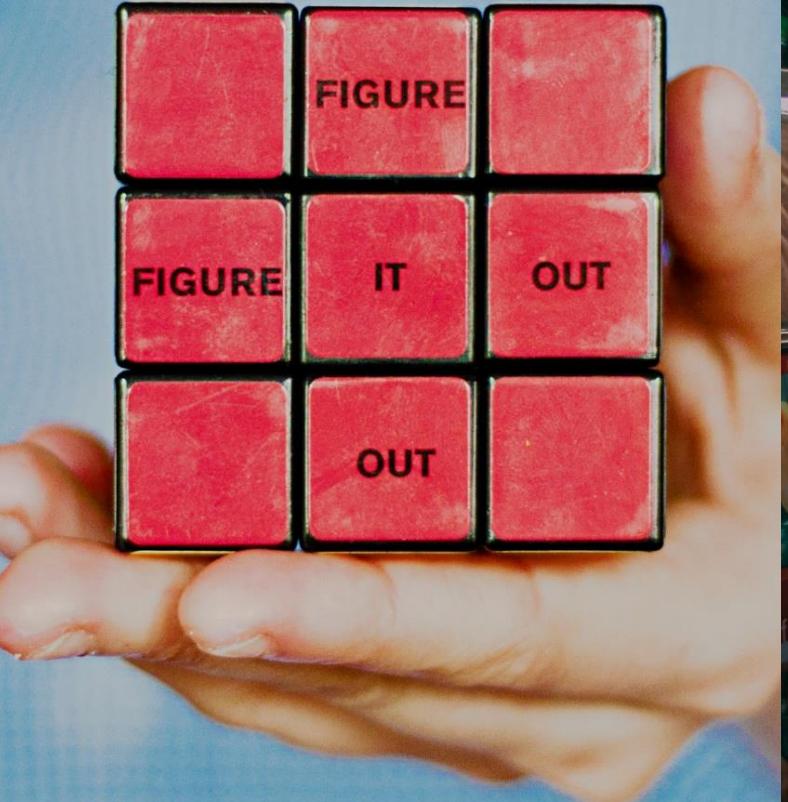
Quarta Revolução Industrial ou Transformação Digital



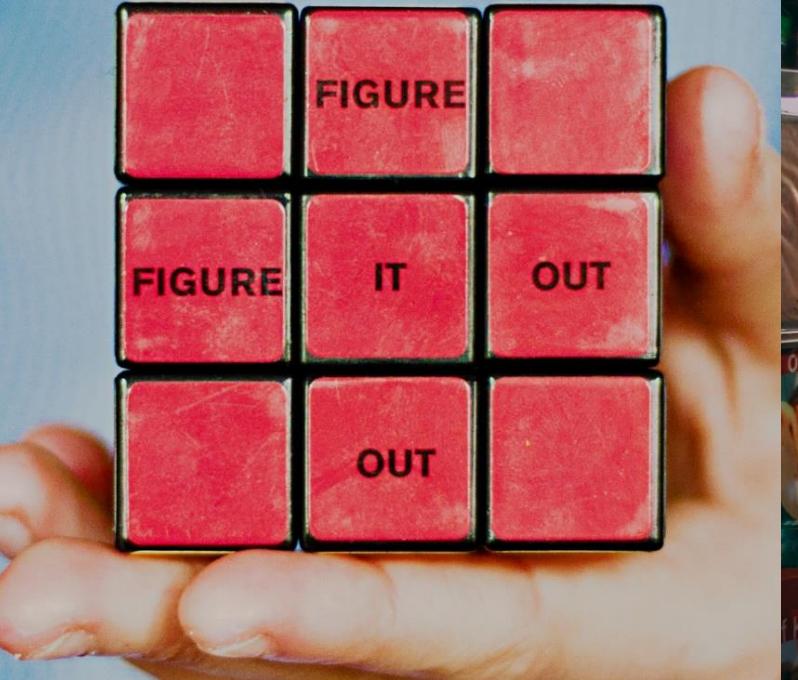
Modelo de Negócio Exponencial

Ecossistemas, Plataformas, Parceiros Coopetidores.

Como alcançar o impacto positivo da Transformação Digital na Sociedade?

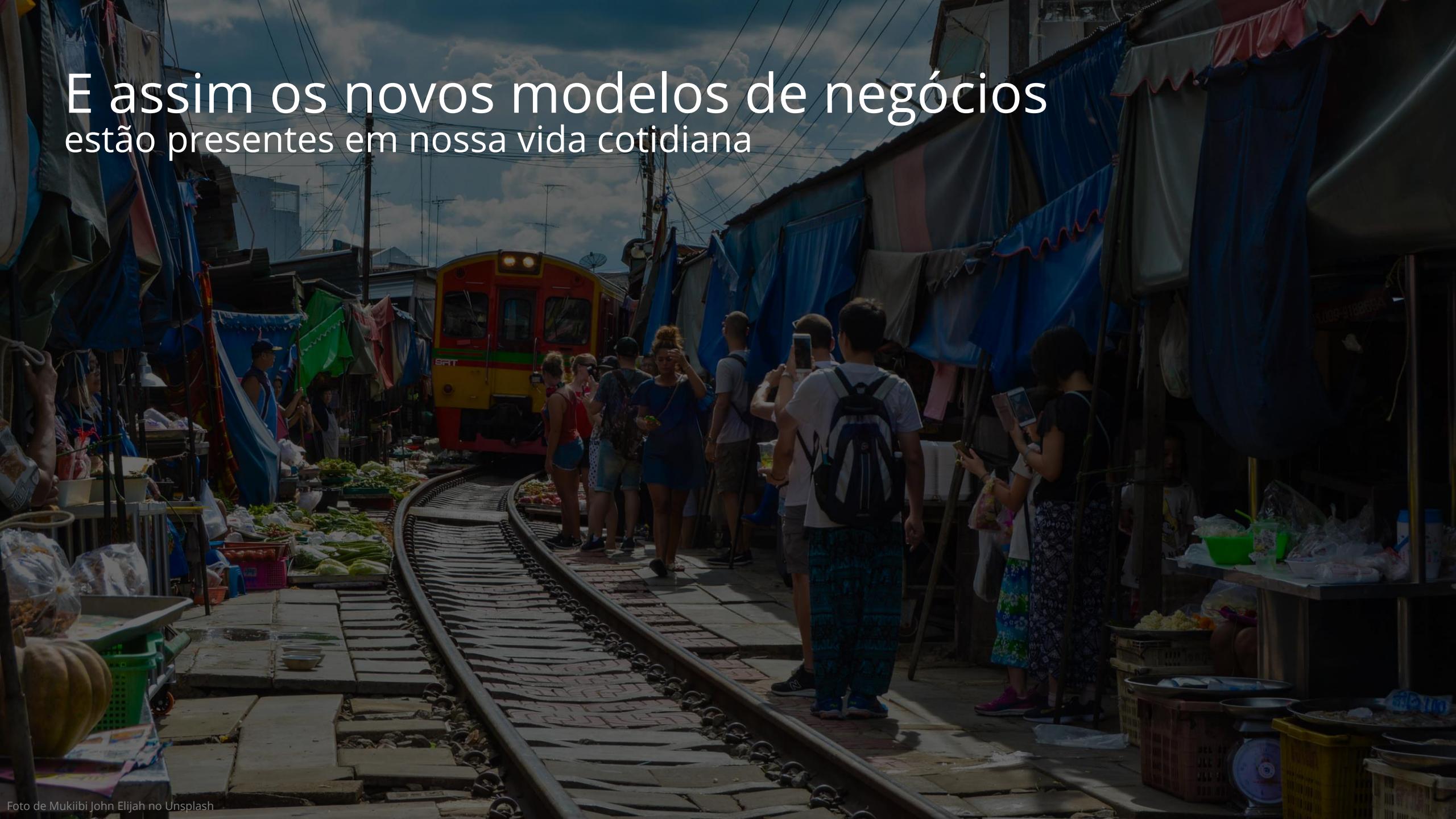


RESOLVER PROBLEMAS



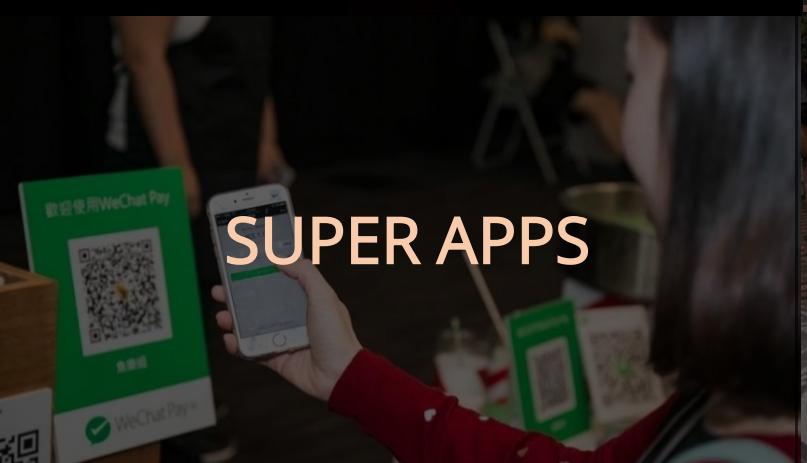
HUMANO CENTRADO

APRENDER COM DADOS



E assim os novos modelos de negócios
estão presentes em nossa vida cotidiana

E assim os novos modelos de negócios
estão presentes em nossa vida cotidiana



Então, vamos aprender com dados
para encontrar as possibilidades de inovação nos negócios





Foto de Edgar Chaparro no Unsplash



Foto de Oscar Nord no Unsplash



Foto de Sebastian Bjune no Unsplash



EXPERIÊNCIA

Foto de Edgar Chaparro no Unsplash



ÚNICO

Foto de Oscar Nord no Unsplash



NÃO-RIVAL

Foto de Sebastian Bjune no Unsplash

Por que construímos Data Warehouses, Data Lakes e Lakehouses para inovar nos negócios?



**PORQUE DADOS
GERAM VANTAGEM
COMPETITIVA PARA
OS NEGÓCIOS.**



**DADOS SÃO FATORES
DE PRODUÇÃO DE
BENS E SERVIÇOS
DIGITAIS.**



**DADOS SÃO UM
BEM DE EXPERIÊNCIA,
INFUNGÍVEL E NÃO-
RIVAL**

Estratégia de Dados





O recipiente mágico entrega:

Escalabilidade

Desempenho

Transação ACID | Atomicidade | Consistência | Isolamento | Durabilidade

Formatos diversos | Estruturado | Semiestruturado | Não estruturado

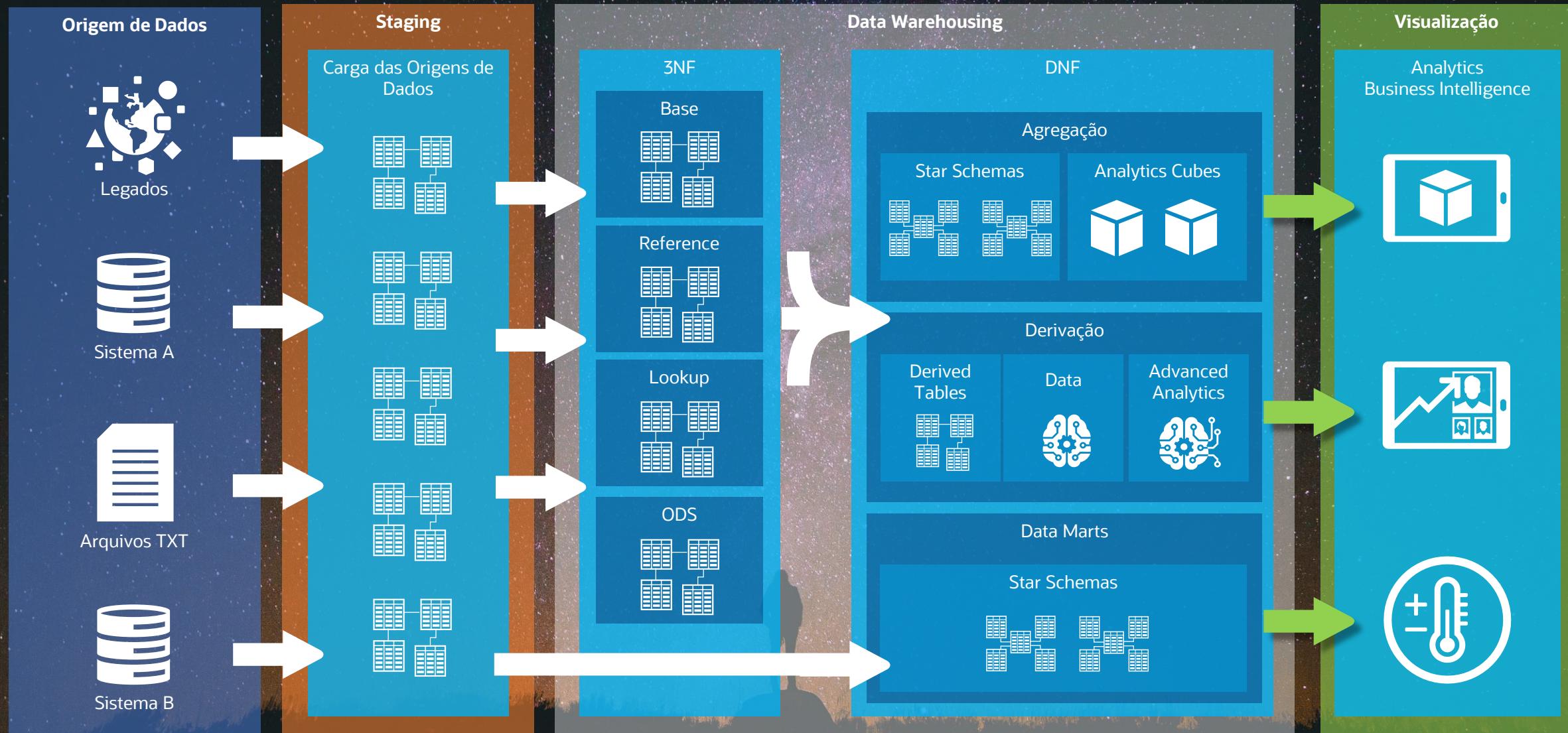
Cargas mistas | SQL para BI | Batch de ETL | Streaming | AI e ML

Acessibilidade

Referência: Learning Spark, 2nd Edition. O'Reilly Media. 2020. ISBN 9781492050049.



Arquitetura Clássica do Data Warehouse



A fama do Data Warehouse



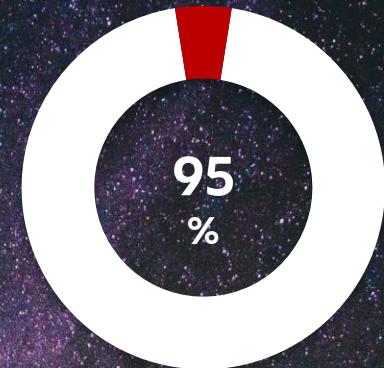
Manual



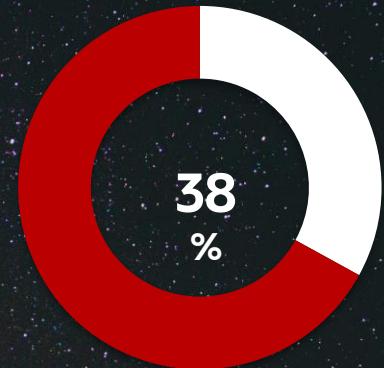
Complexo e caro



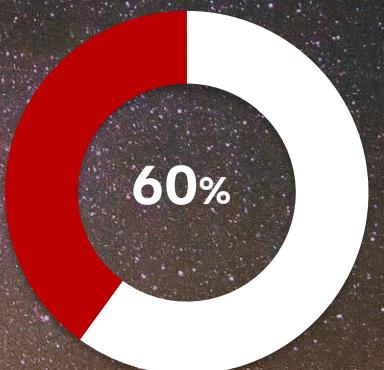
Lento



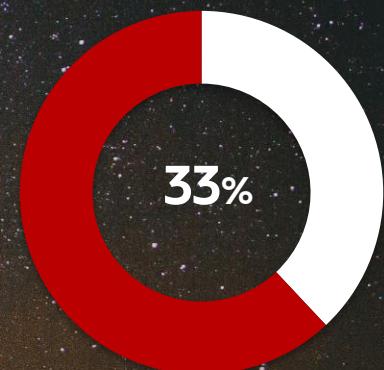
Requer extenso envolvimento manual



Aquisição inicial e custos contínuos com manutenção



Muito complexo de gerenciar



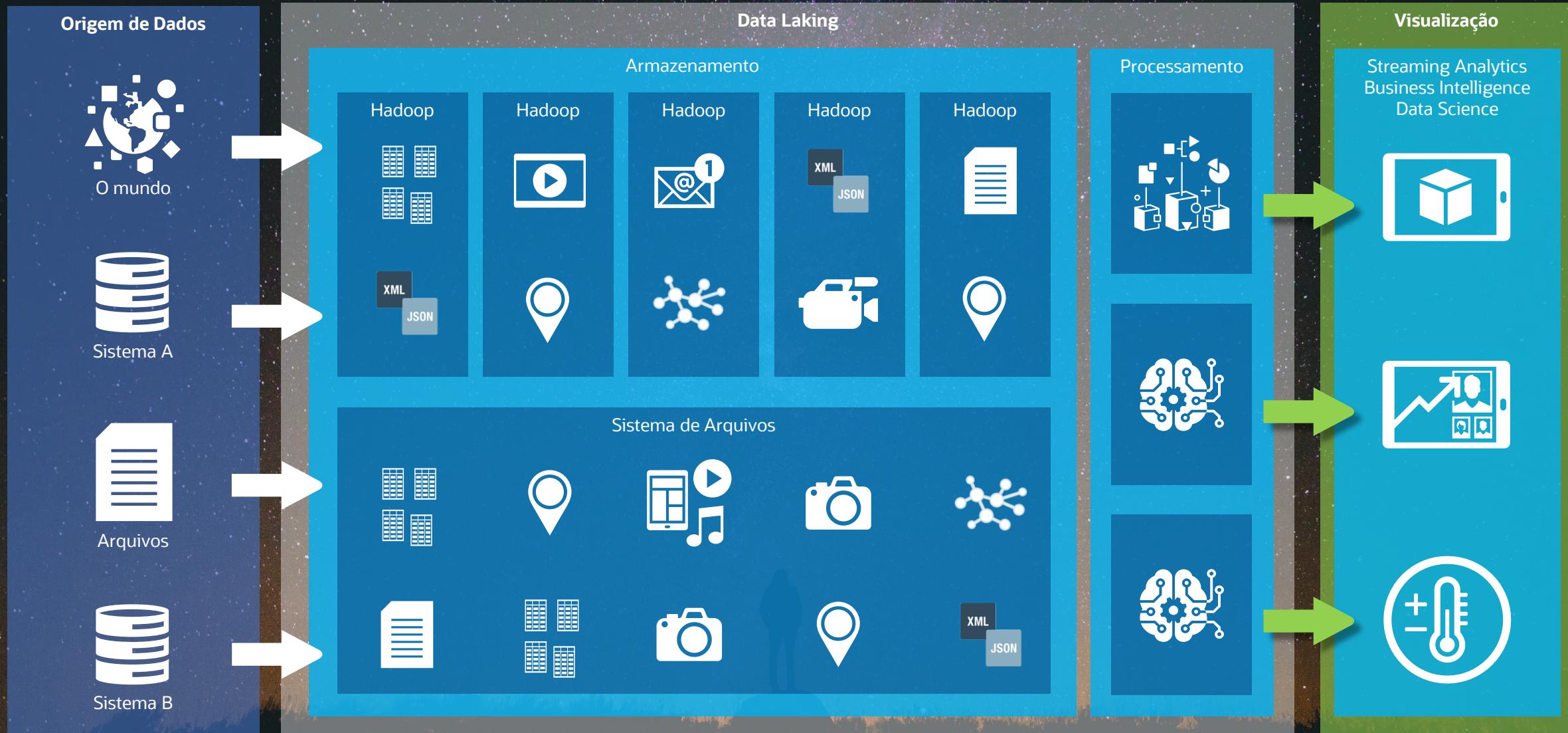
Lento e demorado de implementar

Fonte: Dimensional Research – The State of the Data Warehouse

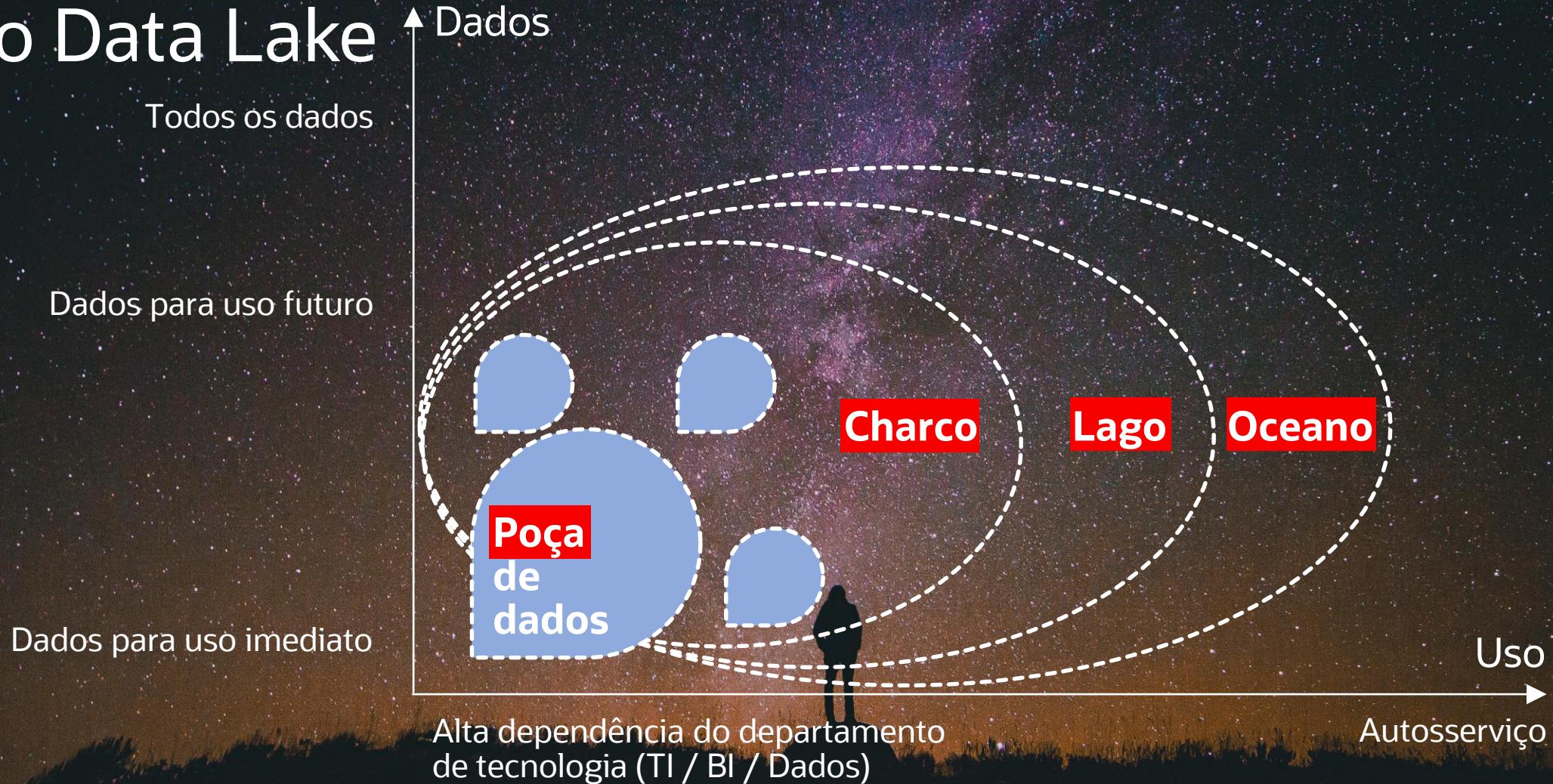


Foto de Felix M. Dorn no Unsplash

Arquitetura Clássica do Data Lake



Maturidade do Data Lake



A fama do Data Lake



Falta de atomicidade e isolamento transacional



Inconsistência de dados e qualidade de dados reduzida



Caótico e complexo

Referência: Learning Spark, 2nd Edition. O'Reilly Media, 2020. ISBN 9781492050049.



Arquitetura de Solução Cloud Data Warehouse & Lake

ORIGEM DE DADOS



INTEGRAÇÃO DE DADOS



DATA LAKING



DATA WAREHOUSING



DATA WAREHOUSE MODERNO

DATA LAKE MODERNO

APRESENTAÇÃO



Abordagem para Data Warehouse Moderno

ORIGEM DE DADOS



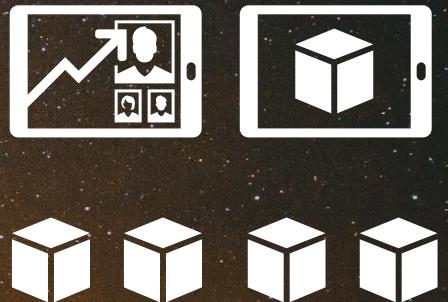
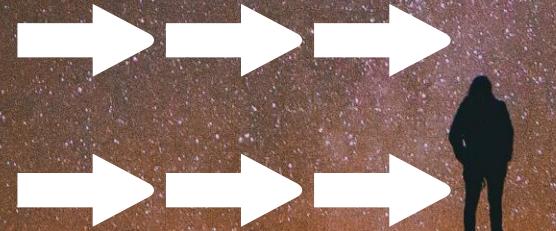
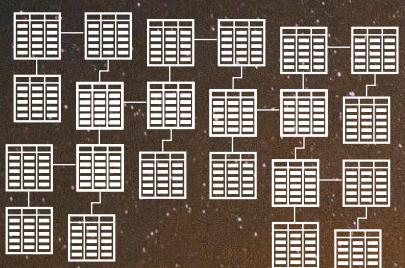
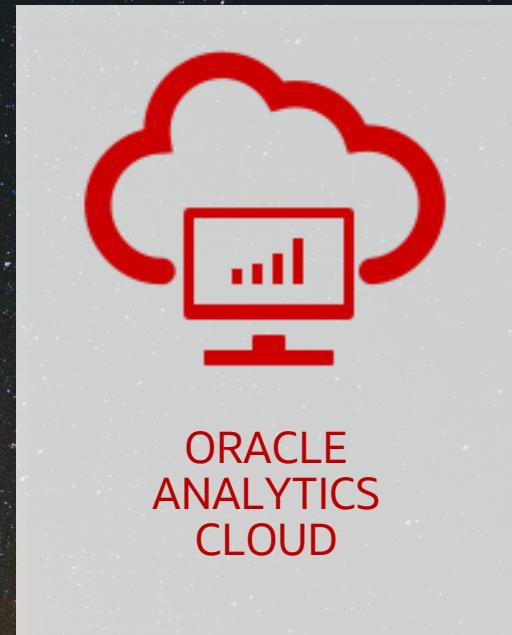
INTEGRAÇÃO DE DADOS



DATAWAREHOUSING



APRESENTAÇÃO

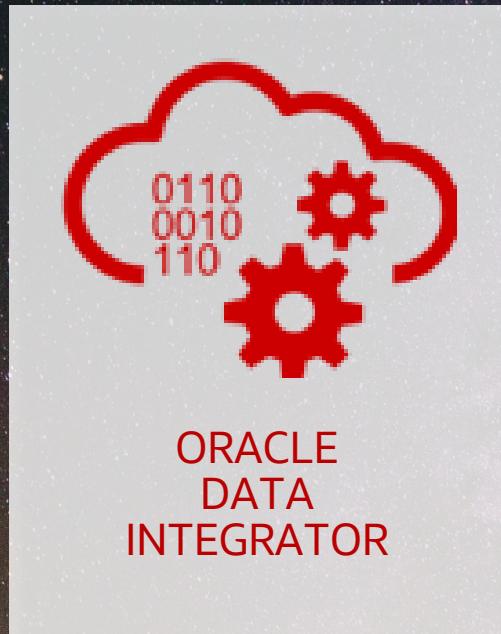


Abordagem para Data Lake Moderno

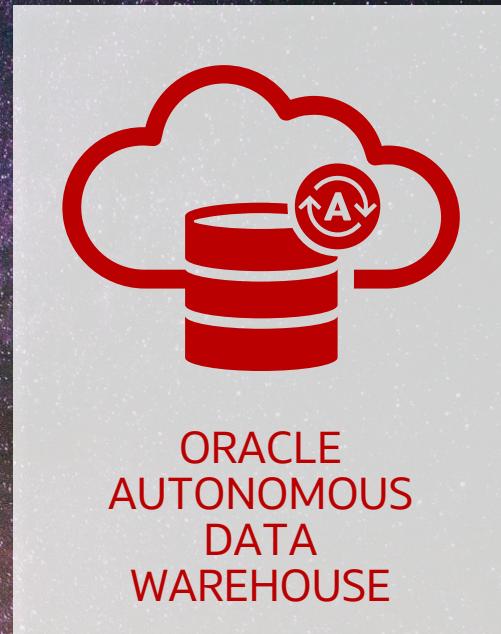
ORIGEM DE DADOS



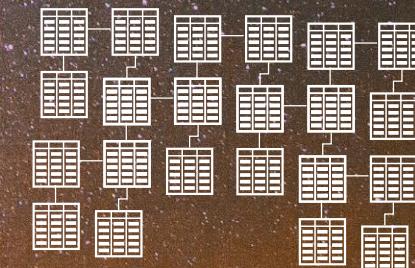
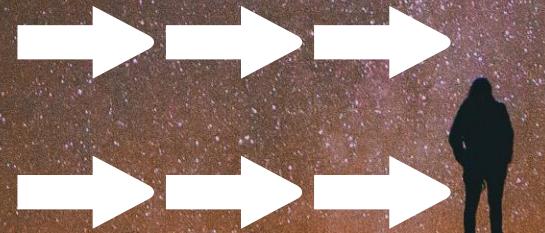
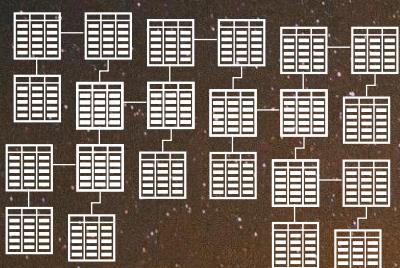
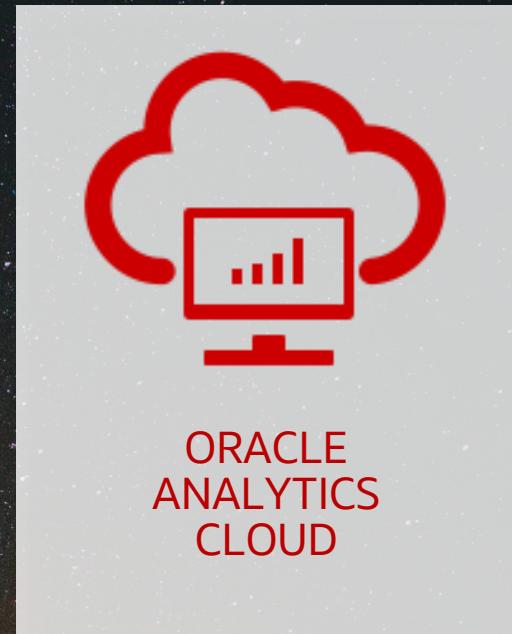
INTEGRAÇÃO DE DADOS



DATA LAKING



APRESENTAÇÃO



E se... eu quiser um Data Lake “mesmo”?

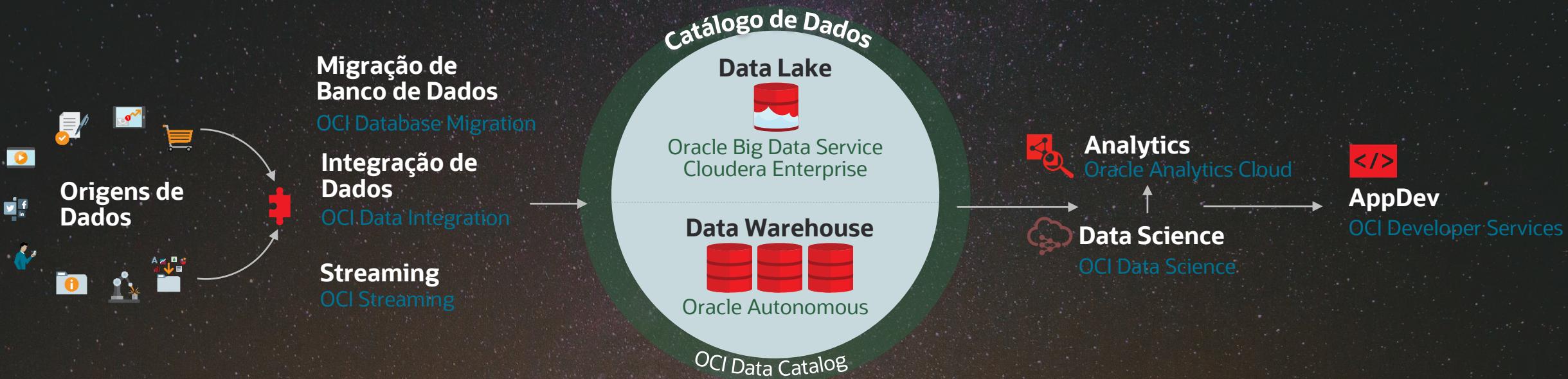




Foto de Evelyn Paris no Unsplash

Lakehouse: um novo paradigma que combina elementos de Data Lake e Data Warehouse.



TRANSAÇÃO
ACID



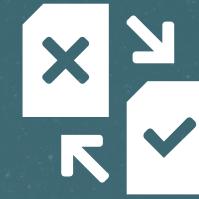
CONFORMIDADE
DE ESQUEMA



FORMATOS
DIVERSOS E
ABERTOS



CARGAS
MISTAS



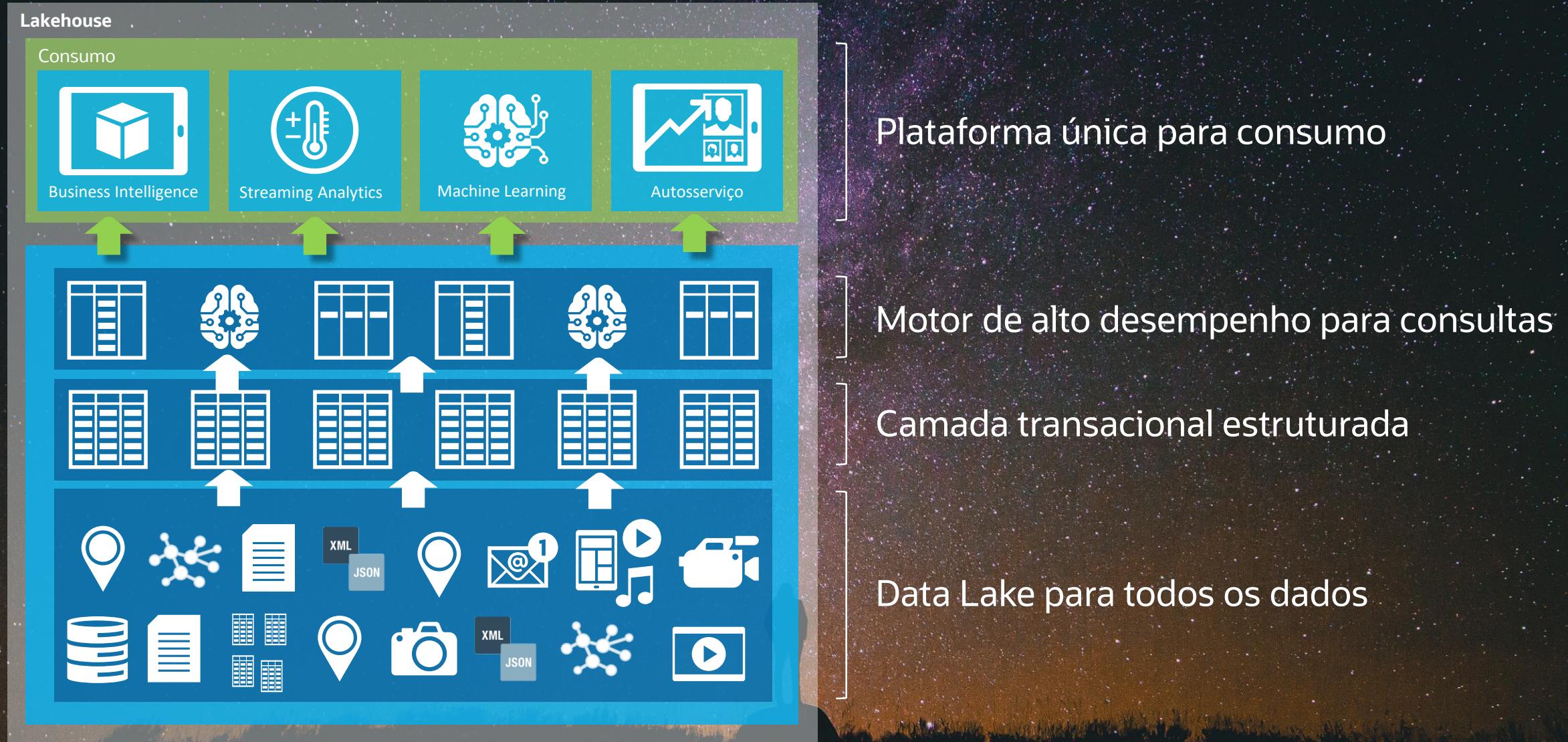
UPSERT &
DELETE
PARALELOS



GOVERNANÇA
DE DADOS

Referência: Learning Spark, 2nd Edition. O'Reilly Media: 2020. ISBN 9781492050049.

Arquitetura de Dados Lakehouse



Projetos Lakehouse

APACHE HUDI

Hadoop Update Delete and
Incremental

Focado em upserts e deletes
em Chave-Valor

Combina formatos colunares e
lineares

APACHE ICEBERG

Focado em propósito geral de
armazenamento em tabelas
únicas de grande tamanho

Evolução de esquema e
particionamento,
versionamento e isolamento
serializado

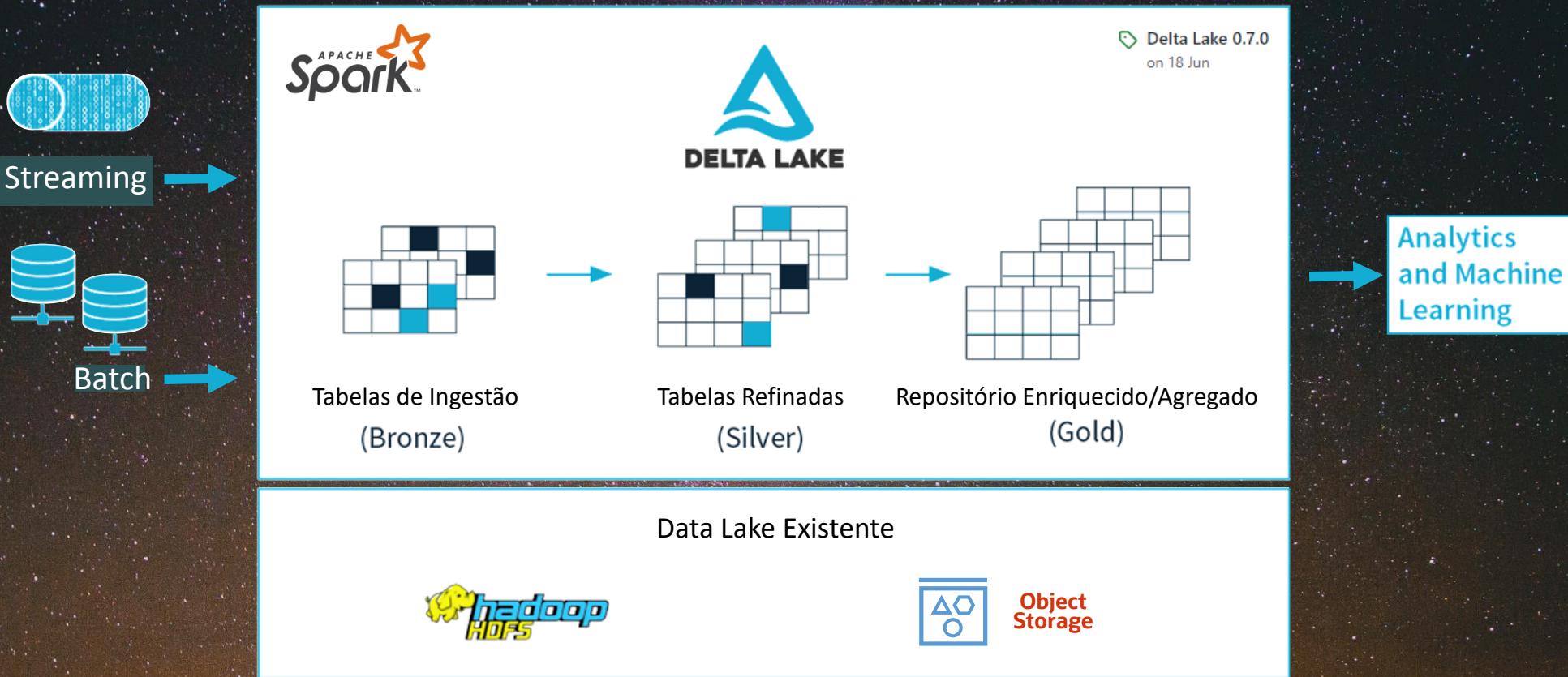
DELTA LAKE

Mantido pela Linux
Foundation e construído pelos
criadores do Apache Spark

Formato aberto de
armazenamento com suporte
transacional e evolução de
esquema

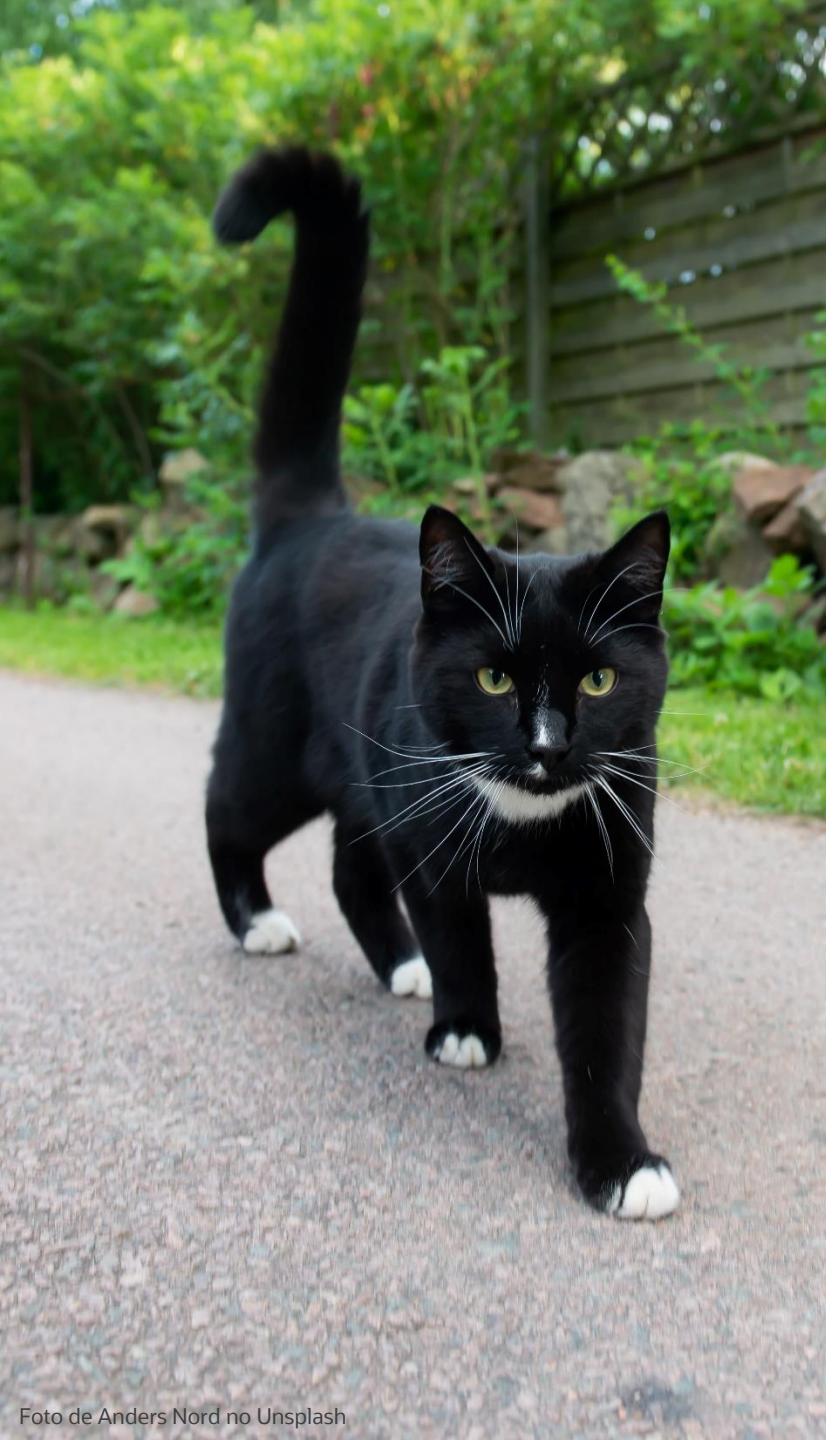
Referência: Learning Spark, 2nd Edition. O'Reilly Media: 2020. ISBN 9781492050049.

Arquitetura de Referência Delta Lake



Referência: <https://delta.io/>





Componentes do Data Warehouse moderno



INTEGRAÇÃO

Streaming,
batch data, on-premises e cloud



DATA WAREHOUSE

Autonomous,
self-driving,
self-securing,
self-repairing



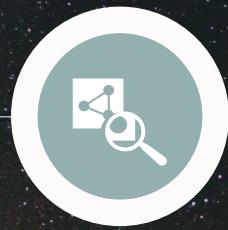
DATA LAKE

Baseado em
Object storage e
integrado com o
Data Warehouse



ANALYTICS

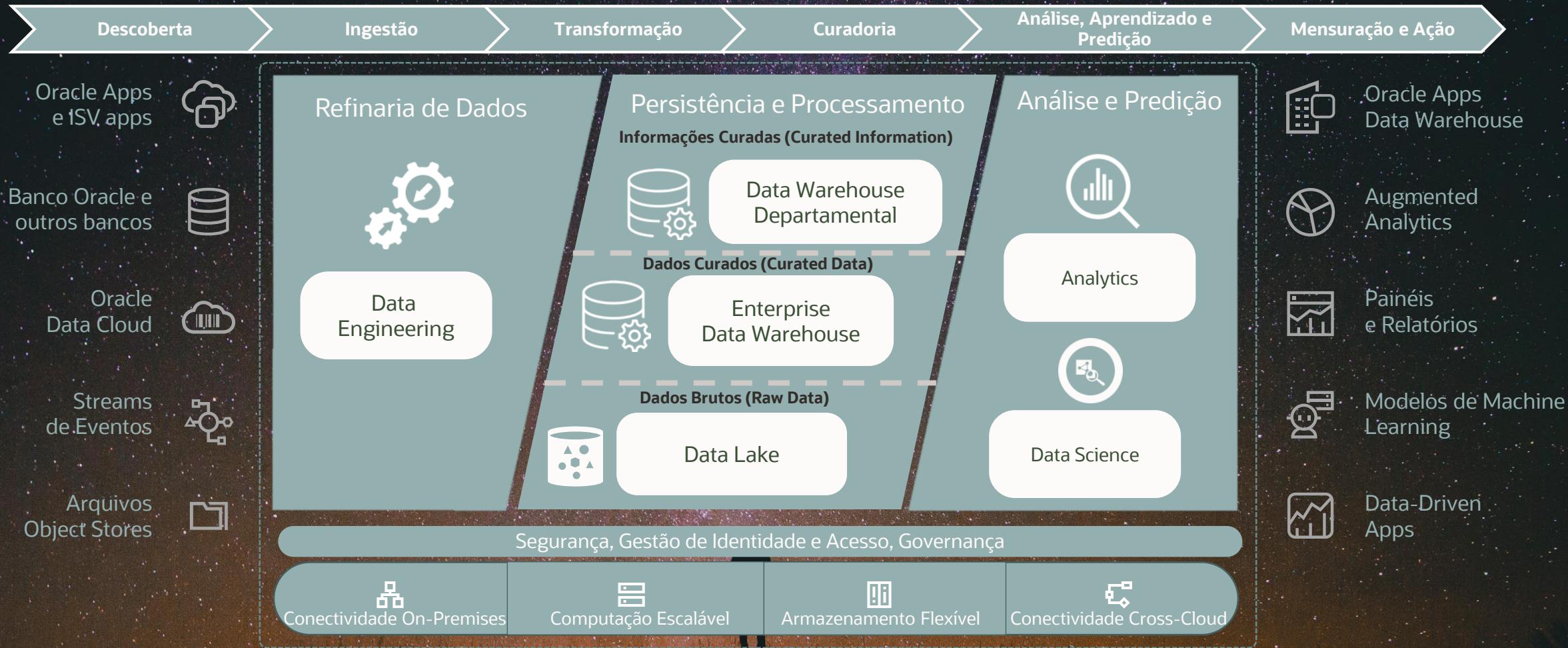
Visualização e
inteligência
analítica baseadas
em Machine
Learning



DATA SCIENCE

Machine learning
de propósito geral
e in-database

Arquitetura de Solução do Data Warehouse moderno

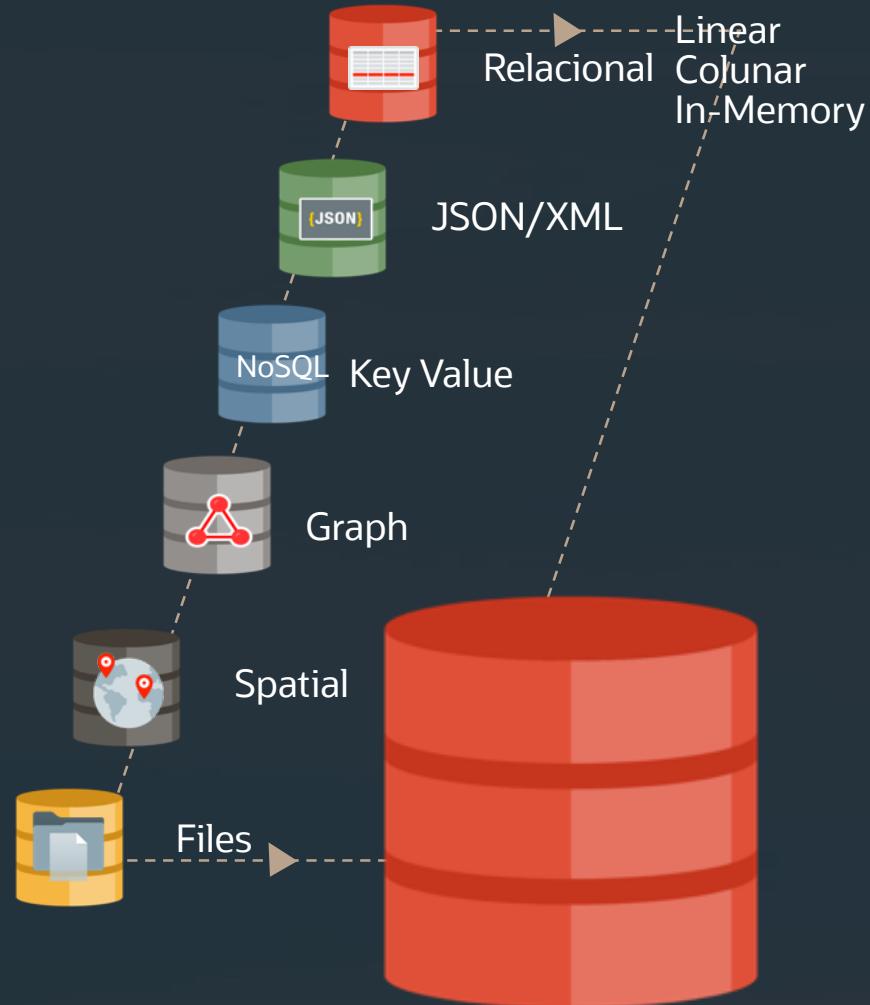


Gravidade
de dados...

Qual é a sua?



Oracle: Base de Dados Convergente



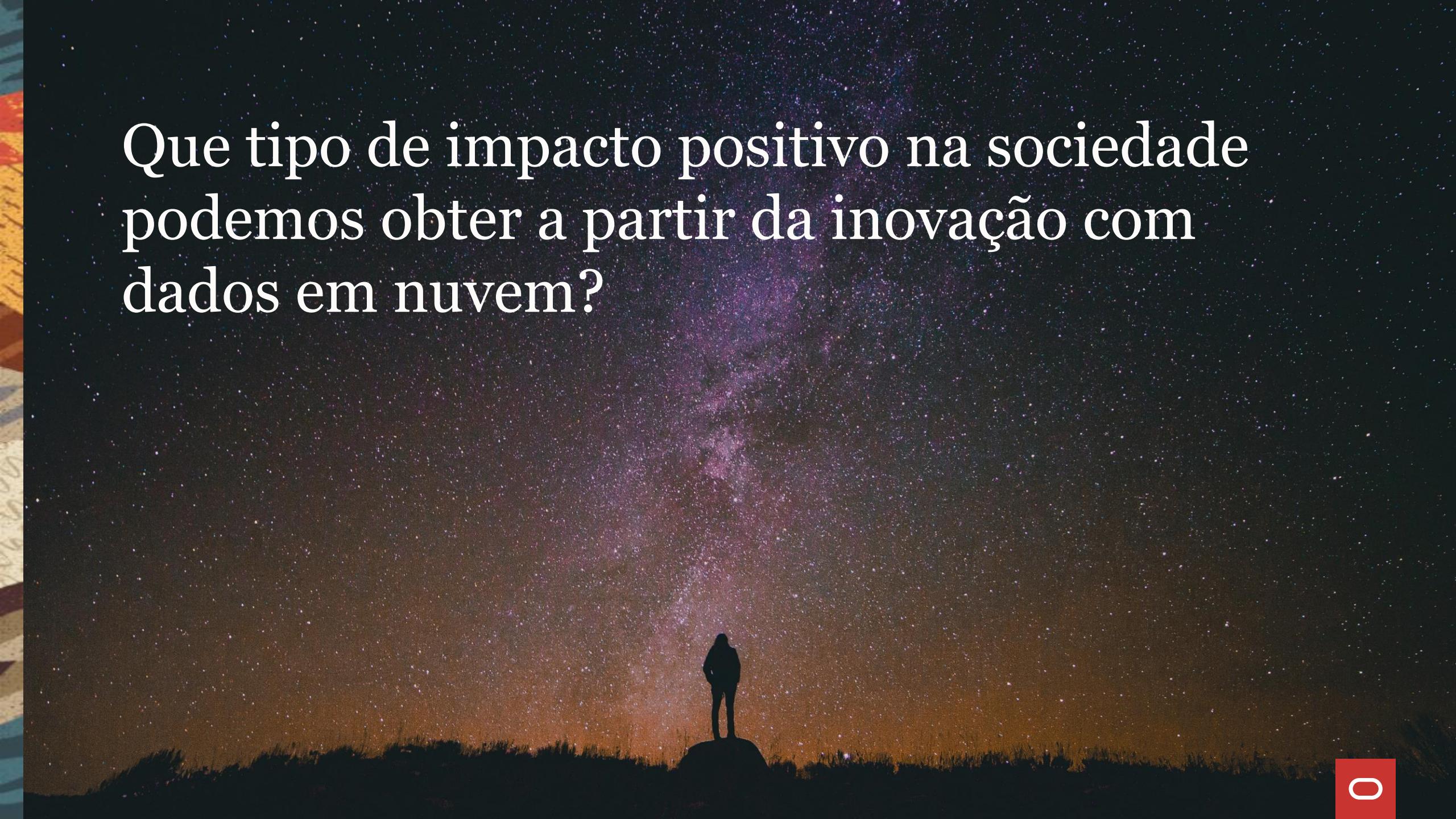
Resumindo as Arquiteturas de Dados...



Conheça a gravidade de dados do seu negócio e use todos os seus ativos de dados com as ferramentas maduras que você domina.

Data Warehouse e Data Lake são paradigmas complementares, relevantes ao seu negócio e permanecem com alta demanda pelo mercado.

Lakehouse é o conjunto de práticas de sucesso conhecidas no manejo de dados e sua implementação pode ser aderente a diferentes tecnologias.



Que tipo de impacto positivo na sociedade
podemos obter a partir da inovação com
dados em nuvem?



AgroScout
Detection Done Right

<https://www.oracle.com/customers/infrastructure/agroscout/>

O



linkedin.com/groups/8984009



youtube.com/c/Inovaçãocomdadosemnuvem



anchor.fm/inova-dados-nuvem



github.com/taborda-cbip/novacao-com-dados-em-nuvem



linkedin.com/in/taborda-cbip

