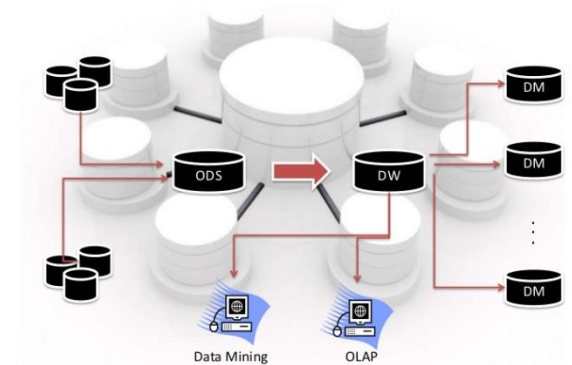


A decorative graphic on the left side of the slide consisting of a grid of squares in various shades of blue and purple, arranged in a stepped pattern.

Fundamentos de *Business Intelligence* (3º Encontro)

Prof. MSc. Fernando Siqueira

Projeto de *Data Warehouse*





Projeto de Data Warehouse

Dicas de Revisão

- Granularidade
- Dimensão Tempo
- *Slow Changing Dimension*
- Conformidade dos Dados
- Chaves Substitutas (SKs)
- Modelo Dimensional
- Qualidade de Dados



Granularidade

- O grão é uma das mais importantes definições na modelagem de dados do DW.
- Trata-se do menor nível da informação e é definido de acordo com as necessidades levantadas no início do projeto.
- Quanto maior for a granularidade, menor será o detalhe (ou maior será a sumarização).
- Quanto menor for a granularidade, maior será o detalhamento (ou menor será a sumarização)



Dimensão Tempo

- A Dimensão temporal (ou dimensão data para alguns) é a mais importante perspectiva para a análise dos dados.
- Sem ela é difícil fazer a averiguação descritiva dos fatos.
- É imprescindível para a análise das ocorrências durante o tempo.
- É importante que seja avaliado a conformidade, analisando os níveis de detalhes exigidos para essa dimensão, mantendo a coerência de definição entre os diversos Data Marts.

Slow Changing Dimension

- O SCD tem papel fundamental para a visualização de informações históricas e armazenamento de versões dos dados de acordo com variações no tempo.
- Retrata as Dimensões que sofrem atualizações em seus campos e os classifica pelo tipo de mudança existente em cada uma delas.
- Exemplo:
 - Em modelo dimensional você tem a possibilidade de avaliar as vendas de cada filial e de cada vendedor. Se um vendedor é transferido para uma outra filial, deve-se ter o tratamento dessa mudança para não ter uma avaliação incorreta do montante de uma venda para cada filial.



Conformidade dos Dados

- A conformidade propicia o compartilhamento semântico de uma mesma Dimensão entre vários *Data Marts*, mantendo a consistência das informações geradas pelas diferentes análises.
- Centraliza as alterações, e permite uma otimização no uso dos recursos computacionais.
- A conformidade de Dimensões é imprescindível para um projeto evolutivo de DW exitoso.



Chaves Substitutas (*Surrogate Keys*)

- Chaves substitutas ou artificiais (surrogate keys), são chaves desprovidas de inteligência de negócio.
- Utilizadas como artifício para o versionamento dos dados e junção entre tabelas Fatos e de Dimensões.



Modelo Dimensional

- Revise também o modelo dimensional resultante na busca de erros que só são evidentes após a implementação do DW.
- Verifique a existência de relacionamentos desnecessários e duplicidade de informações sem propósito e que oneram a performance das análises e consultas.



Qualidade de Dados

- A qualidade dos dados é um dos aspectos mais importante no processo de construção do DW.
- Devemos analisar os dados armazenados a procura de inconsistências e deformidades.
- Caso seja detectado alguma anormalidade nos dados, o processo de ETL deverá ser verificado e revisado para garantir que as informações armazenadas possuam valor para a utilização na tomada de decisão empresarial.



Qualidade de Dados

- A qualidade dos dados é um dos aspectos mais importante no processo de construção do DW.
- Devemos analisar os dados armazenados a procura de inconsistências e deformidades.
- Caso seja detectado alguma anormalidade nos dados, o processo de ETL deverá ser verificado e revisado para garantir que as informações armazenadas possuam valor para a utilização na tomada de decisão empresarial.

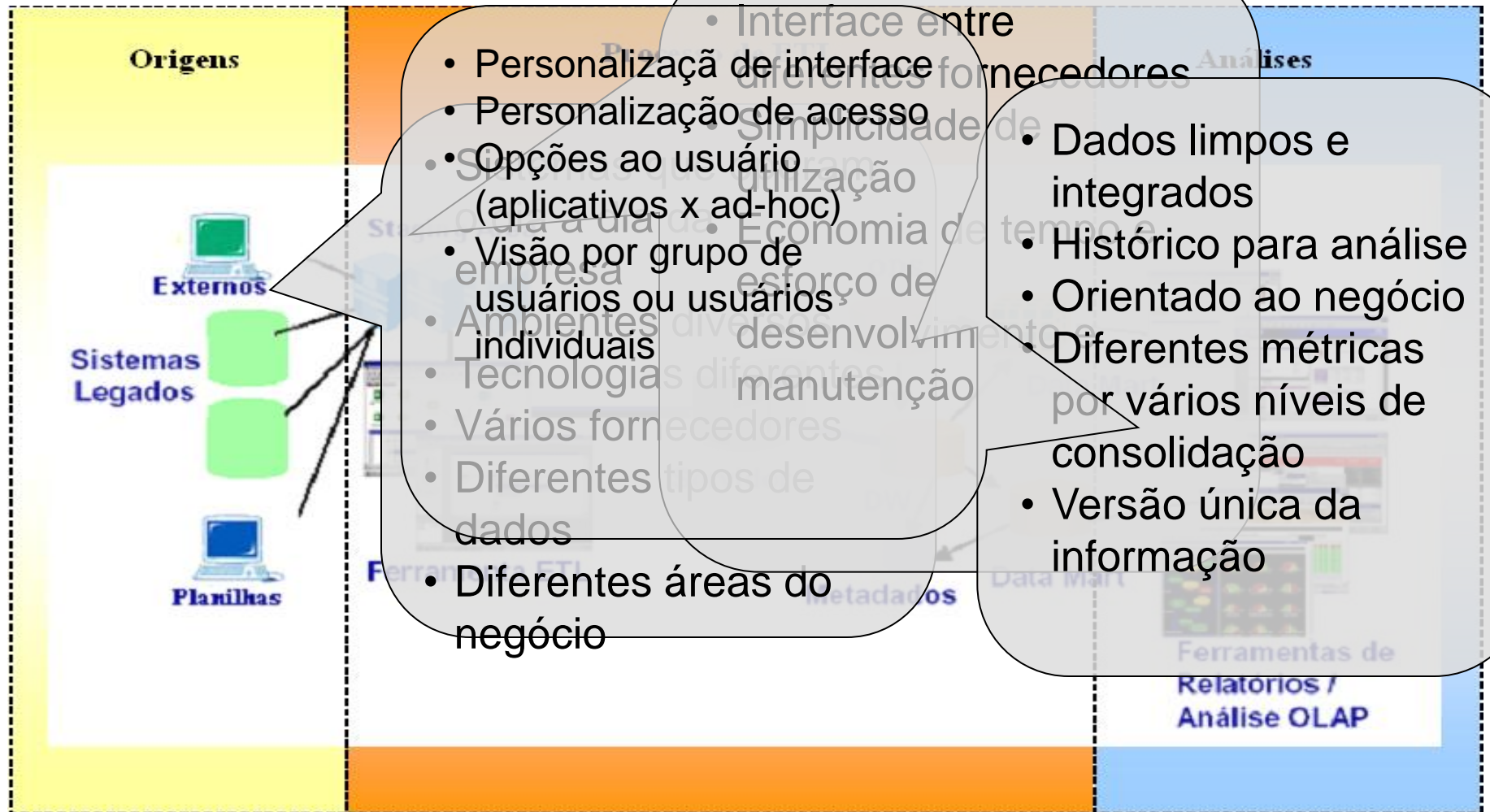


Business Intelligence

O Processo BI ...



O Processo BI ...



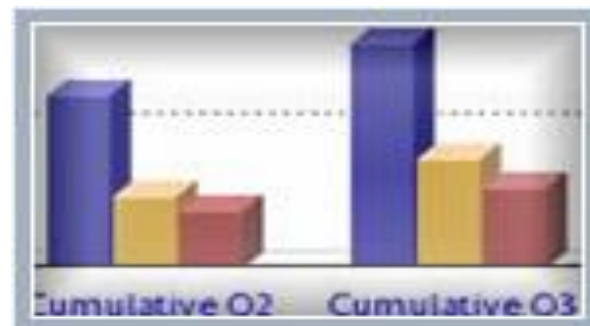


OLAP

- Conjunto de ferramentas voltadas para acesso e análise *ad-hoc* de dados
- Objetivo de uma ferramenta OLAP
 - “Transformar dados em informações capazes de dar suporte a decisões gerenciais de forma amigável e flexível ao usuário e em tempo hábil”

OLAP

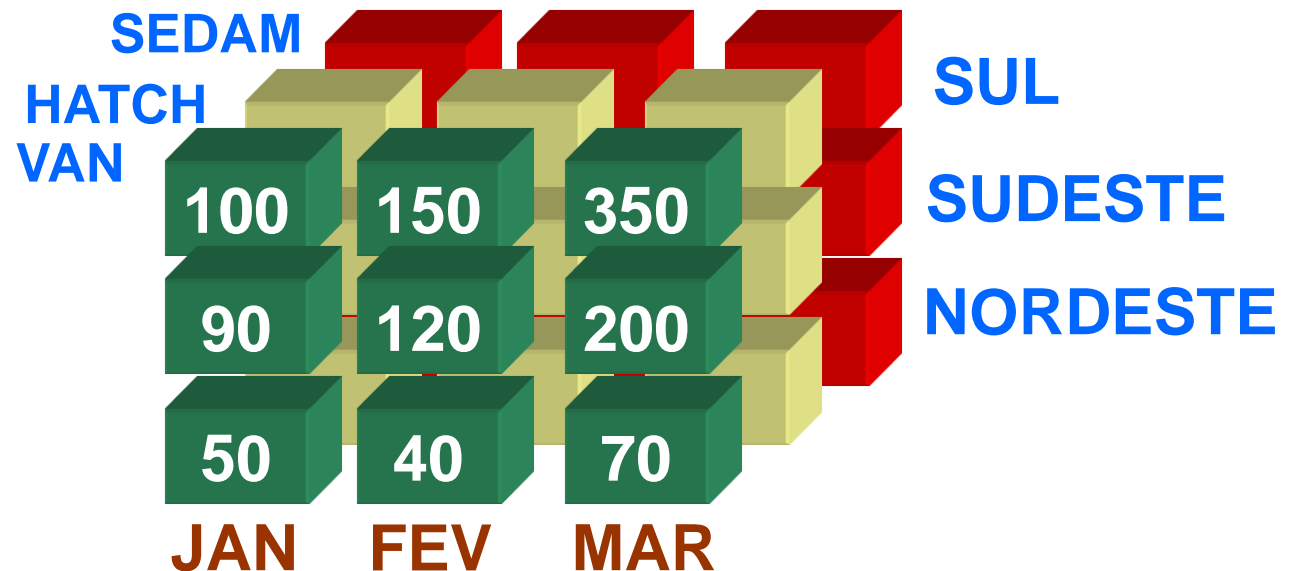
Características



- OLAP – *On-Line Analytic Processing*: fornece para organização um método de acessar, visualizar, e analisar dados corporativos com alta flexibilidade e performance.
- Disponibiliza os dados em forma de cubos para a análise dos dados em diversos ângulos de visão, sobre os aspectos de negócio. Trata as informações como Dimensões e Indicadores.

OLAP

- Cubo OLAP – Cubo de dados
 - Estrutura que facilita aos usuários visualizar os dados segundo suas dimensões



- Operações de navegação OLAP
 - Slice-dice / drill-down / roll-up / drill-up



OLAP

Operações

- Principais operações
 - Pivot
 - Drill-down
 - Driw-up (Roll-up)
 - Slice-and-Dice
- Operações podem ser combinadas para se obter a visualização desejada



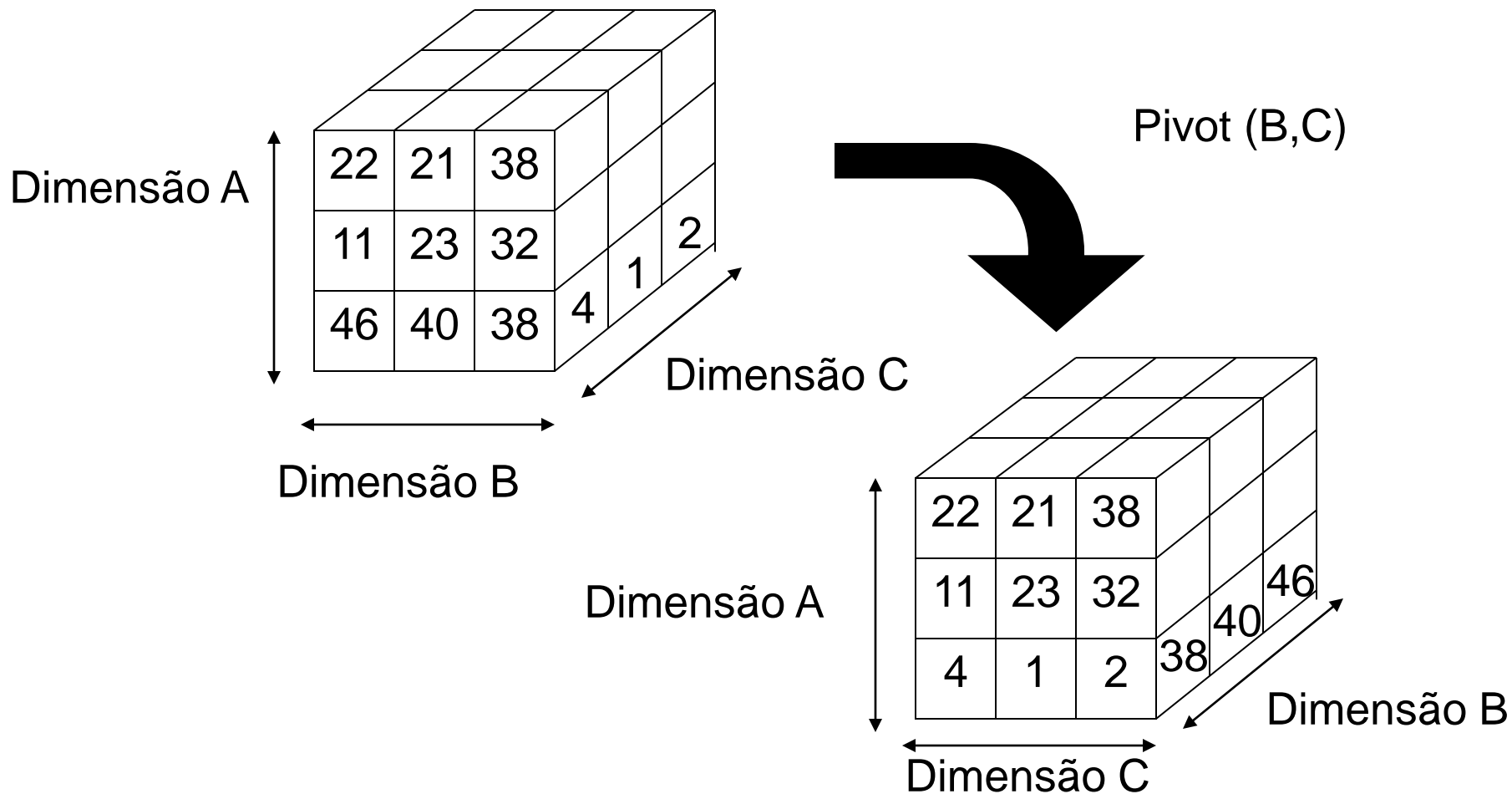
OLAP

Operações - PIVOT

- Rotação do hipercubo
- A forma mais comum de visualização
- Corresponde a inversão dos eixos das dimensões para, por exemplo, uma posterior rolagem.

OLAP

Operações - PIVOT

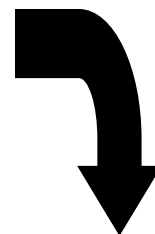


OLAP

Operações – PIVOT - Exemplo

Quantidade Vendida (milhares)	2017	
	Liquidificador	Micro-Ondas
SUL	77	152

Pivot (local, produto)



Quantidade Vendida (milhares)	2017
	SUL
Liquidificador	77
Micro-ondas	152



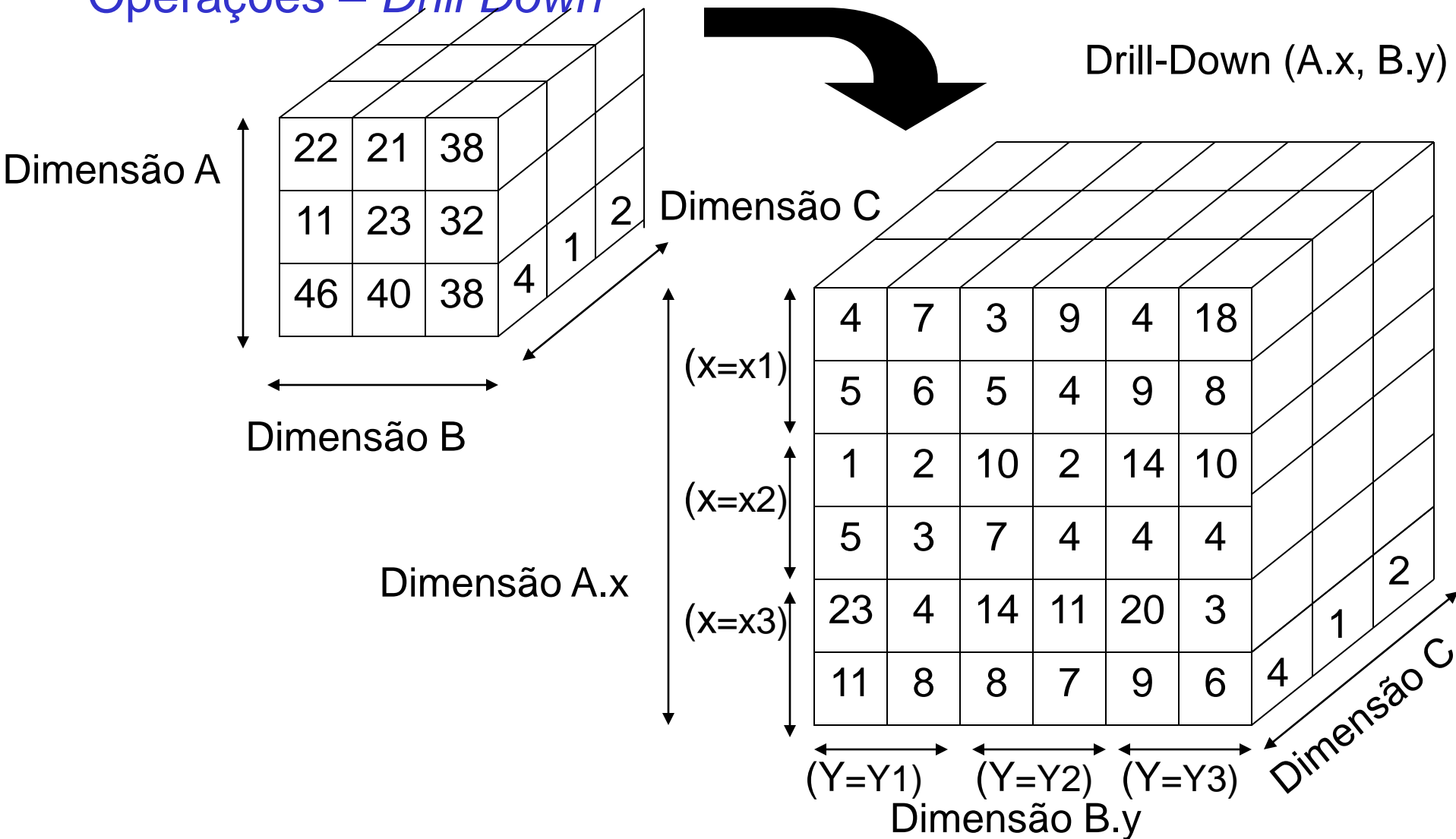
OLAP

Operações – *Drill Down*

- Determinam o detalhamento de um consulta
- As consultas são mais restritas se existirem mais detalhes nos critérios de seleção
 - Ex. País x Cidade
- Os resultados são mais detalhados

OLAP

Operações – *Drill Down*

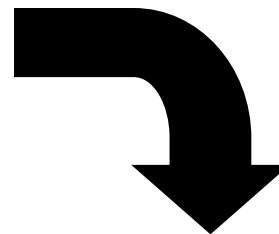


OLAP

Operações – *Drill Down* - Exemplo

Quantidade Vendida (milhares)	2017	
	Liquidificador	Micro-Ondas
SUL	77	152

Drill-Down
(*local.região.estado*,
produto.tipo.marca)



Quantidade Vendida (milhares)		2017			
		Liquidificador		Micro-Ondas	
		Walita	Faet	Sharp	Consul
SUL	RS	14	16	40	25
	SC	9	10	12	20
	PR	15	13	30	25



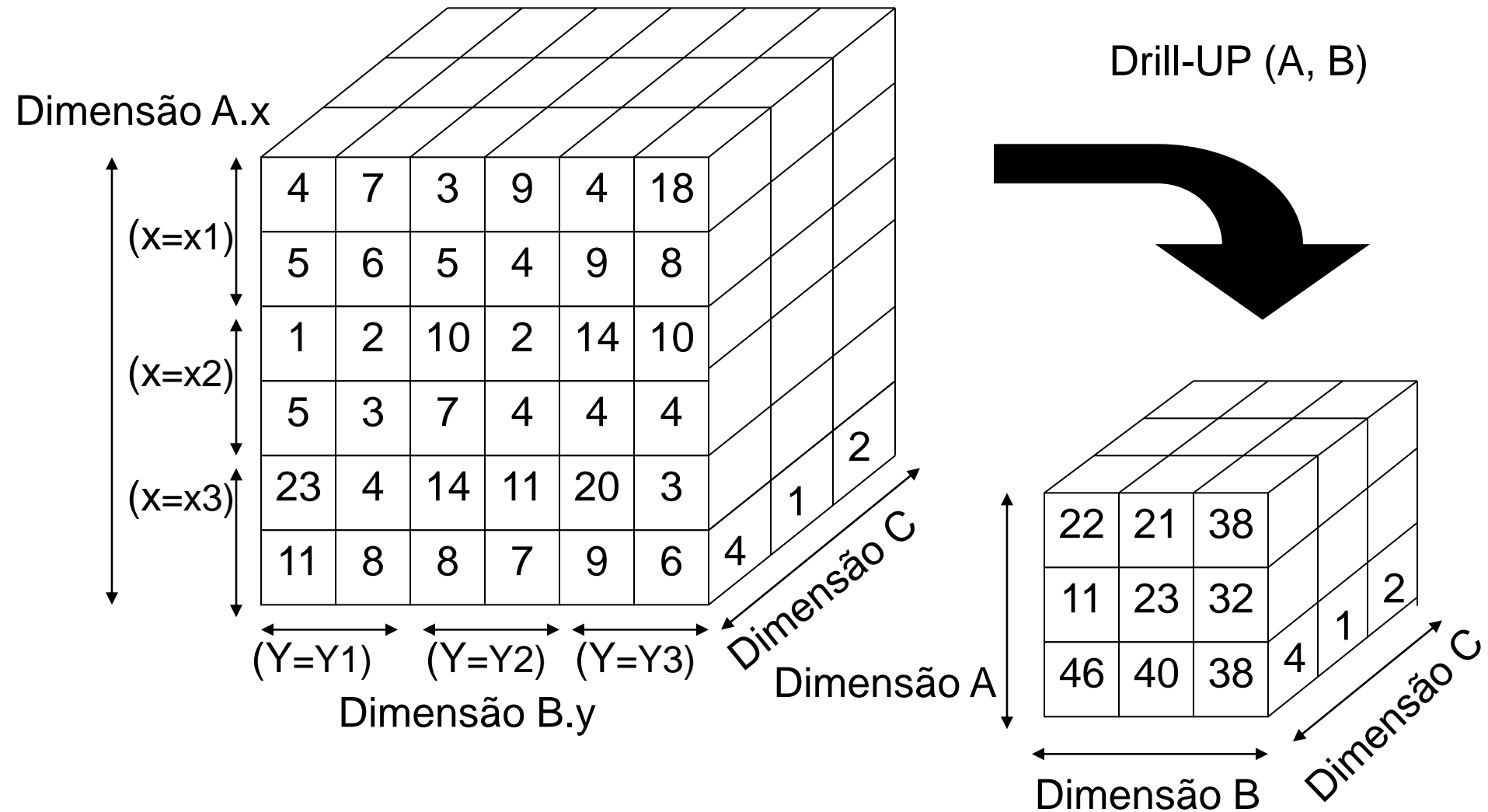
OLAP

Operações – *Drill Up*

- Determinam o detalhamento de um consulta
- As consultas são mais abrangentes se existirem menos detalhes
 - Ex. Cidade x pais
- O resultado é mais sumarizado

OLAP

Operações – *Drill Up*

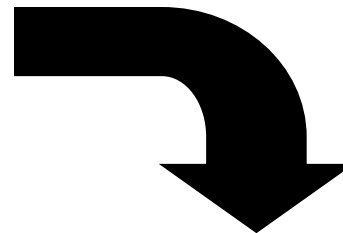


OLAP

Operações – *Drill Up*- Exemplo

Quantidade Vendida (milhares)		2017			
		Liquidificador		Micro-Ondas	
		Walita	Faet	Sharp	Consul
SUL	RS	14	16	40	25
	SC	9	10	12	20
	PR	15	13	30	25

Drill-up (produto.tipo)



Quantidade Vendida (milhares)		2017	
		Liquidificador	Micro-ondas
SUL	RS	30	65
	SC	19	32
	PR	28	55



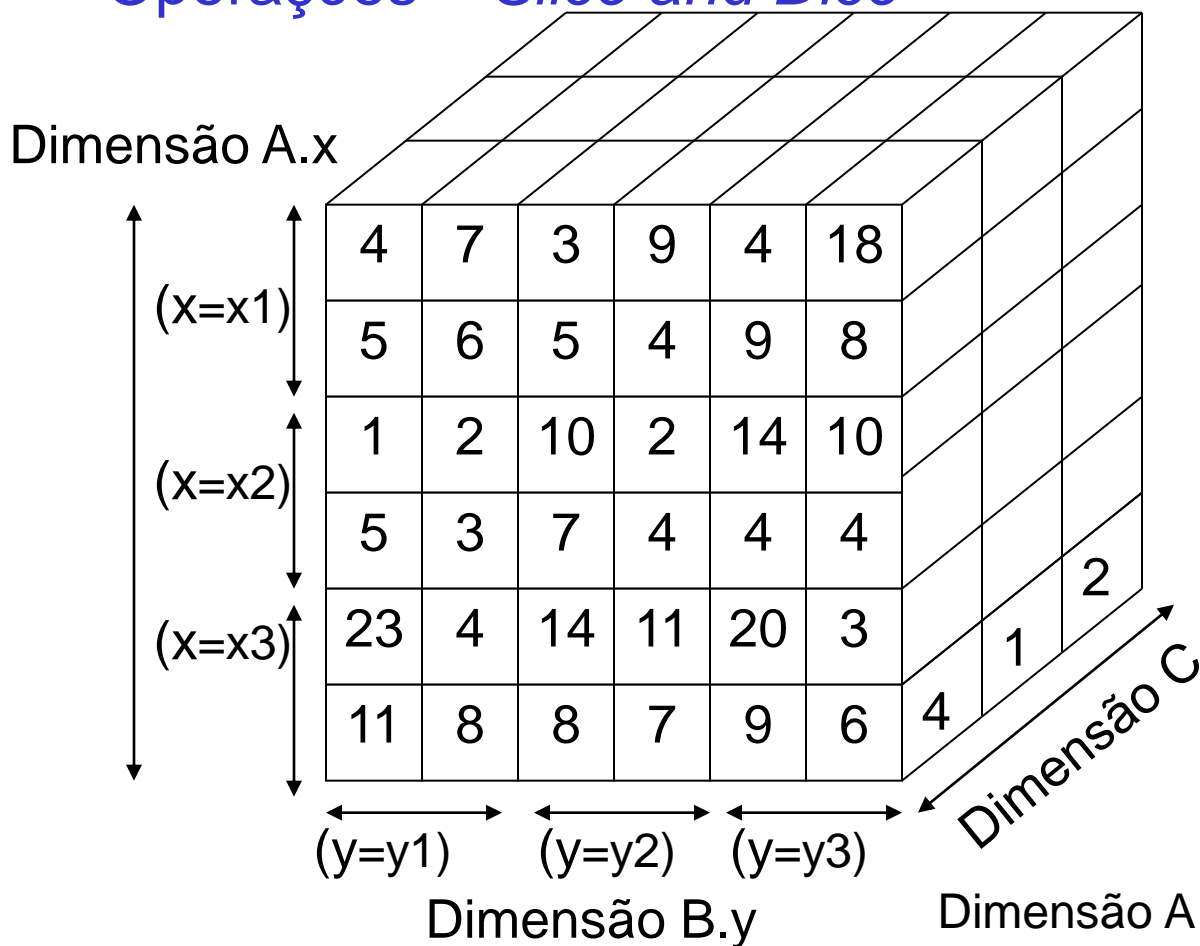
OLAP

Operações – *Slice and Dice*

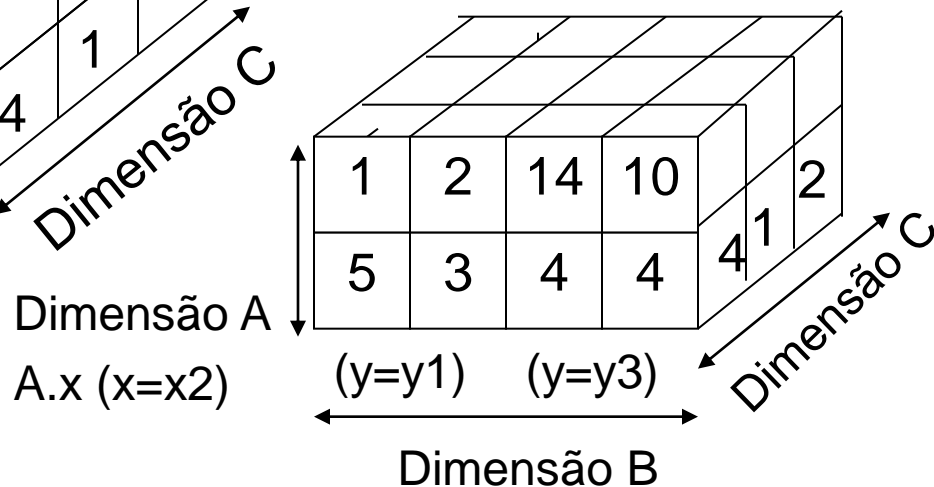
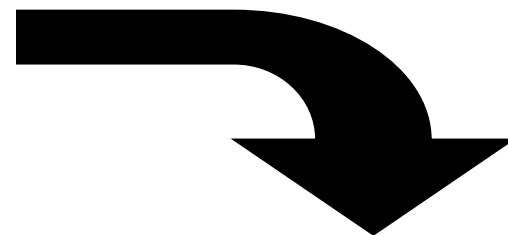
- As consultas são realizadas de acordo com as restrições impostas
- As restrições correspondem a atributos de dimensões com determinado valor de consulta Quanto mais atributos de dimensão mais restrito é o conjunto de resposta.
- Uma escolha de atributos de dimensões representa uma fatia(Slice) na visualização das dimensões
- A troca dos atributos de dimensões é uma operação de rolagem (dice).

OLAP

Operações – *Slice and Dice*



Slice-and-Dice
(A.x = x2), B.y = y1 v B.y = y3)



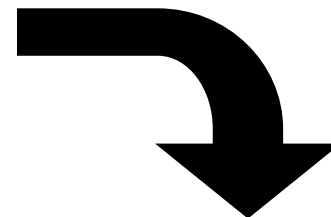
OLAP

Operações – *Slice and Dice*

Quantidade Vendida (milhares)		2017			
		Liquidificador		Micro-Ondas	
		Walita	Faet	Sharp	Consul
SUL	RS	14	16	40	25
	SC	9	10	12	20
	PR	15	13	30	25

Slice-and-Dice

*(local.região.estado = 'SC',
produto.tipo.marca = 'Walita',
produto.tipo.marca = 'Sharp')*



Quantidade Vendida (milhares)		2017	
		Liquidificador	Micro-ondas
		Walita	Sharp
SUL	SC	9	12

Ferramentas de BI

MicroStrategy

 Microsoft

 birst

SAP

 .sas

alteryx

 + a b | e a u
S O F T W A R E

PROGNOZ

**Information
Builders**

Logi 
ANALYTICS

TIBCO 

 DATAWATCH

 pentaho[®]
A Hitachi Group Company

Qlik 

ORACLE[®]



Self-service BI



Self-Service BI

Conceito

- Criar e implantar suas próprias análises apoiado na estruturação de uma arquitetura corporativa e de ferramentas que entreguem autonomia na concepção e modelagem para a implementação de relatórios e *dashboards*.



Self-Service BI

Contexto

- A **complexidade do negócio** cresceu além da capacidade de organizações e indivíduos para compreendê-lo e gerenciá-lo.
- A necessidade de velocidade **analítica**, flexibilidade e inovação exigidas pelo negócio digital mudou o investimento analítico para unidades de negócios dos departamentos de TI tradicionais, mas a mudança em si acrescenta ainda mais complexidade.
- O negócio digital exige análises efetivas e inteligência de negócios para permitir decisões de negócios rápidas e confiáveis. Muitas empresas lutam para conseguir esse objetivo aparentemente simples.



Self-Service BI

Desafios

- A **complexidade do negócio** cresceu além da capacidade de organizações e indivíduos para compreendê-lo e gerenciá-lo.
- A necessidade de velocidade **analítica**, flexibilidade e inovação exigidas pelo negócio digital mudou o investimento analítico para unidades de negócios dos departamentos de TI tradicionais, mas a mudança em si acrescenta ainda mais complexidade.
- O negócio digital exige análises efetivas e inteligência de negócios para permitir decisões de negócios rápidas e confiáveis. Muitas empresas lutam para conseguir esse objetivo aparentemente simples.

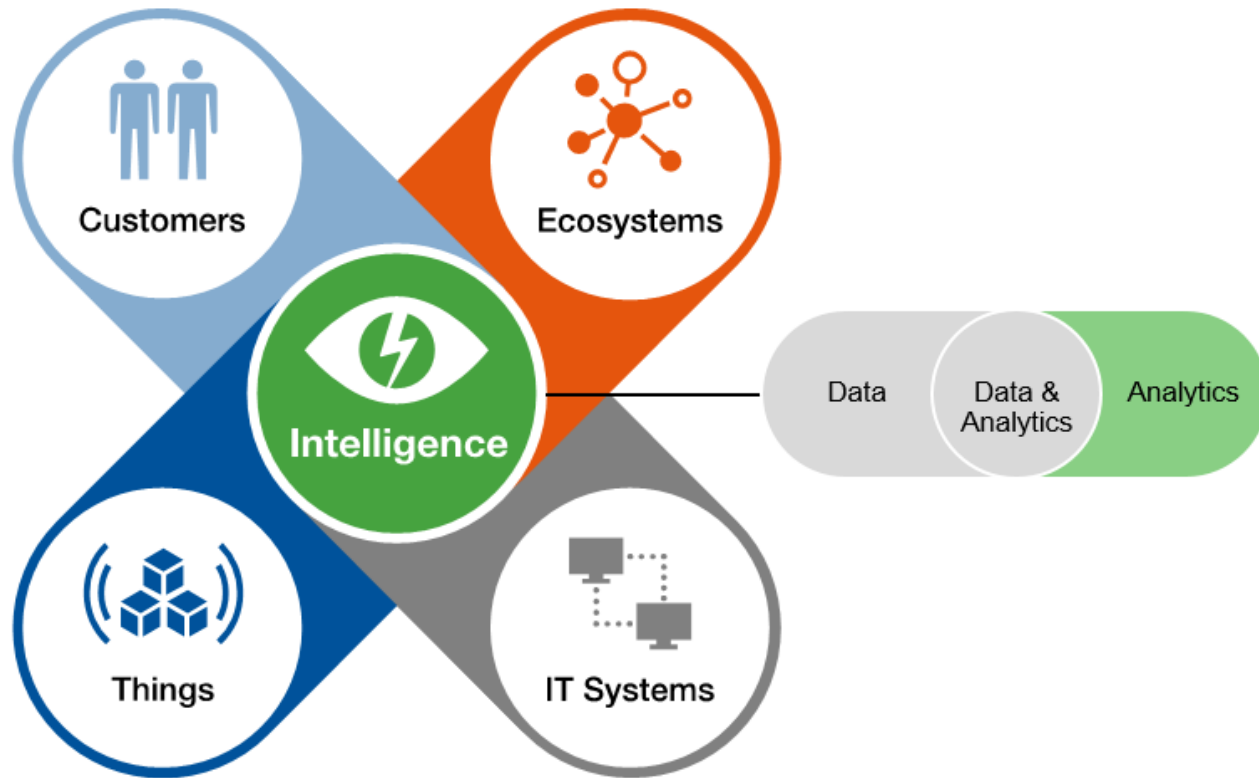


Self-Service BI

Desafios

- Líderes de dados e análise identificaram três desafios principais:
 - Primeiro, eles precisam verificar quais recursos de análise são essenciais para oferecer valor comercial.
 - Em segundo lugar, eles devem gerenciar os papéis, habilidades e conhecimentos emergentes que permitirão que as equipes de análise façam seu próprio trabalho e treinem usuários empresariais para realizar suas próprias análises quando apropriado.
 - Em terceiro lugar, eles são dificultados por processos e tecnologias existentes na criação de soluções de análise ágeis que suportam diretamente iniciativas empresariais digitais.

Self-Service BI

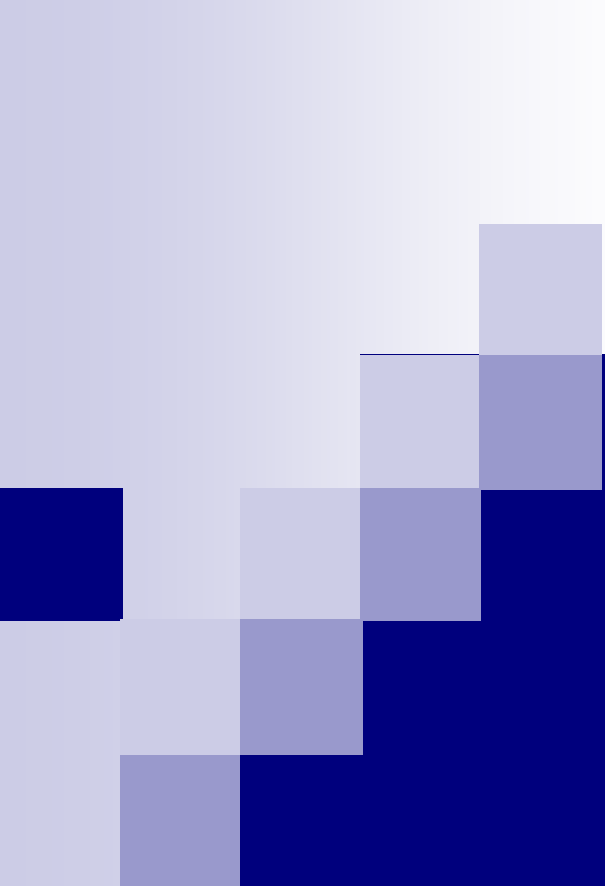




Self-Service BI

Pontos de Atenção

- As necessidades de cada empresa é que direcionarão a solução e as questões como segurança, integridade e disponibilização da informação, que devem ser cuidadosamente analisadas.
- É preciso planejar a adoção de soluções self-service, assim como implantar processos de utilização e governança consistentes. Porque a competitividade de sua organização no futuro depende da definição dessas políticas hoje.



Ferramentas de Visualização de Dados



Visualização de Dados

"O uso de representações visuais para explorar, dar sentido e comunicar dados".

- Visualização de dados vs. Visualização de informações
 - Informação = agregação, resumo e contextualização de dados
- Relacionado a gráficos de informação, visualização científica e gráficos estatísticos
- Muitas vezes inclui gráficos, ilustrações, ...

Visualização Analítica

- Um termo recentemente cunhado
 - Visualização de informações + análise preditiva
- Visualização de informações
 - Descritivo, orientado para trás
 - "O que aconteceu" o que está acontecendo "
- Análise preditiva
 - Preditivo, focado no futuro
 - "O que vai acontecer" "por que isso vai acontecer"
- Há um forte movimento em direção à análise visual

visual analytics

O Crescimento da *Visualização de Dados e Analítica*

- Quadrante Mágico para Plataformas de *Business Intelligence e Analytics*
- Muitas empresas de visualização de dados estão no 4º quadrante
- Existe um movimento para a visualização



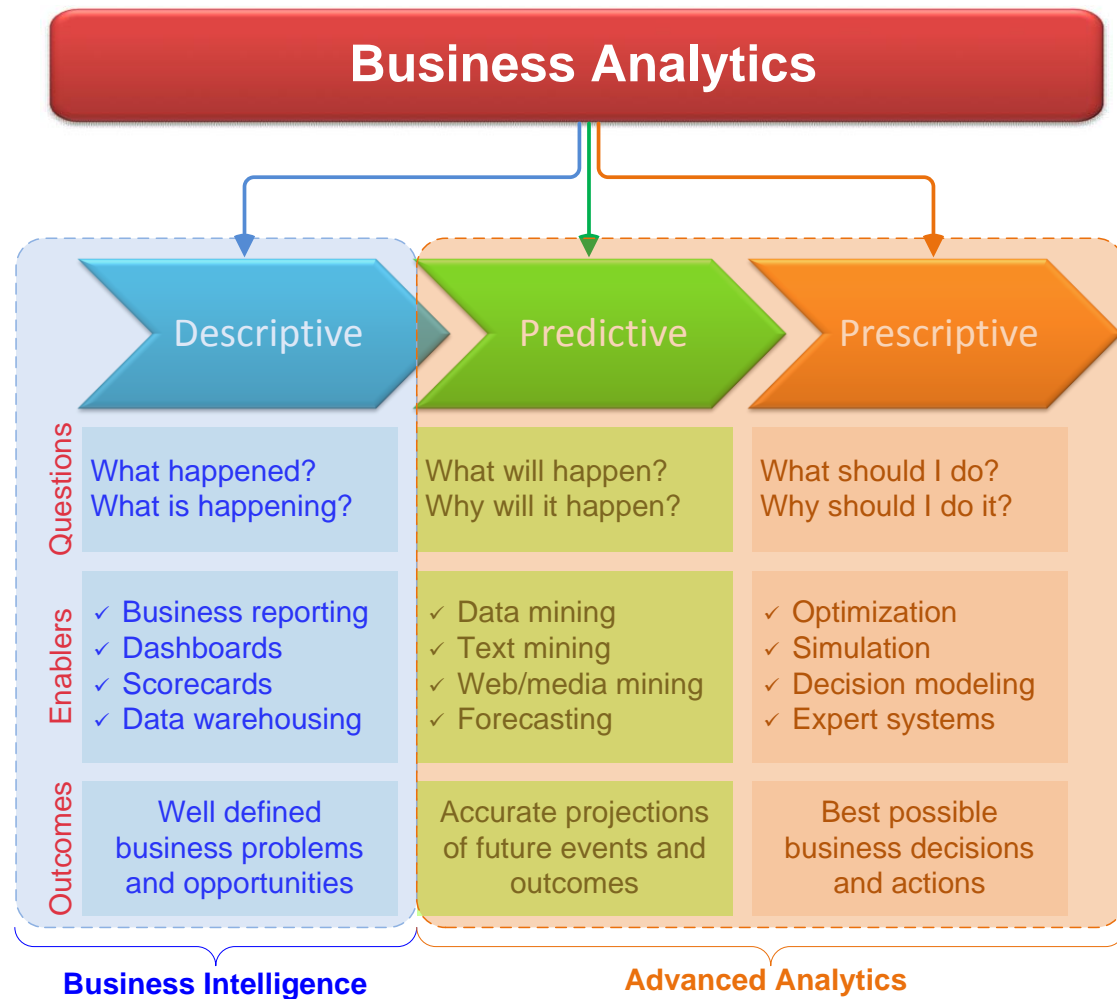


O Crescimento da *Visualização de Dados e Analítica*

- Novos *players*
 - Tableau, QlikView, Spotfire, ...
- Aumento do foco dos *big players*
 - MicroStrategy melhorando Visual Insight
 - Microsoft reforçando PowerPivot with Power View
 - IBM lançando Cognos Insight
 - Oracle adquirindo Endeca
 - SAP lançando Visual Intelligence
 - SAS lançando Visual Analytics

Business Intelligence e Data Warehouse

- O BI costumava ser tudo relacionado ao uso de dados para suporte de decisão gerencial
- Agora, é uma parte do *Business Analytics*
 - BI = Análise descritiva





LDW – Logical Data Warehouse



LDW

Introdução

- Nas últimas décadas, o DW se tornou uma ferramenta essencial para que as organizações compreendam os clientes, melhorem os processos e até executem as operações diárias.
- No entanto, forças convergentes aceleram as mudanças de negócios e tecnologia que desafiam a eficácia do DW tradicional.
- Como resultado, muitos pressupostos do DW e práticas comuns não podem mais ser considerados como adquiridos.

LDW

Motivação

- Os seguintes pontos propuseram oportunidades e desafios para a evolução do DW:
 - As mídias sociais: embora a mídia social seja uma mina de ouro para análise de sentimentos, o DW não consegue processar e integrar redes sociais devido a barreiras causadas por dados não estruturados e volumes de dados maciços. —
 - Computação em nuvem: infraestrutura como serviço (IaaS) e plataforma como serviço (PaaS) menor custo de entrada, diminui os requisitos de habilidades e entrega soluções mais rápidas. Em contraste, as análises "out-of-the-box" oferecidas pelo software como aplicações de serviço (SaaS) reduzem a necessidade do DW.

LDW

Motivação

- Os seguintes pontos propuseram oportunidades e desafios para a evolução do DW:
 - Mobilidade: novos dispositivos móveis podem capturar instantaneamente dados de alta fidelidade - por exemplo, tempo, localização e identidade. Eles também permitem que *business intelligence* (BI) atinja uma base de usuários mais ampla. O desafio é processar dados capturados rapidamente e fornecer insights relevantes para os contextos dos usuários.

LDW

Motivação

- Os seguintes pontos propuseram oportunidades e desafios para a evolução do DW:
 - Informações: grande análise de dados apresenta vantagens competitivas e novas oportunidades de negócios. No entanto, 3Vs - de características importantes dos dados impõem grandes desafios ao DW.
 - Primeiro, o volume de dados está crescendo exponencialmente a cada ano, e o TDW muitas vezes não consegue atender a SLAs.
 - Em segundo lugar, as organizações devem trabalhar com uma variedade cada vez maior de formatos de dados - por exemplo, feeds, mídia social e informações contextuais; O TDW não pode processar uma variedade de formatos de dados.
 - Em terceiro lugar, o negócio de hoje requer informações em tempo real (velocidade de dados), mas a maioria dos TDWs ainda operam em um mundo em lote.



LDW

Conceito

- “O *Logical Data Warehouse* (LDW) é uma arquitetura de gerenciamento de dados para análise que combina os pontos fortes dos armazéns de repositório tradicionais com gerenciamento de dados alternativo e estratégia de acesso.”



LDW

Benefícios

- Os principais benefícios do LDW são melhorar a tomada de decisões e proporcionar uma vantagem competitiva.
- O LDW aproveita o potencial da inteligência de negócios em dados em repouso e em movimento, a fim de fornecer informações poderosas para ambientes operacionais e estratégicos.
- Novos paradigmas de computação, como Hadoop e CEP, utilizados pelo LDW, têm potencial para ajudar as organizações a atender às crescentes demandas de 3Vs.



LDW

Benefícios

- O LDW também melhora a agilidade e reutilização do *data warehouse* seguindo uma arquitetura em camadas.
- Os sistemas podem ser alterados de forma mais fácil e rápida usando novos padrões de integração de dados, como a federação de dados e a integração orientada para mensagens.
- Esses padrões são mais flexivelmente acoplados e flexíveis do que o tradicional ETL.



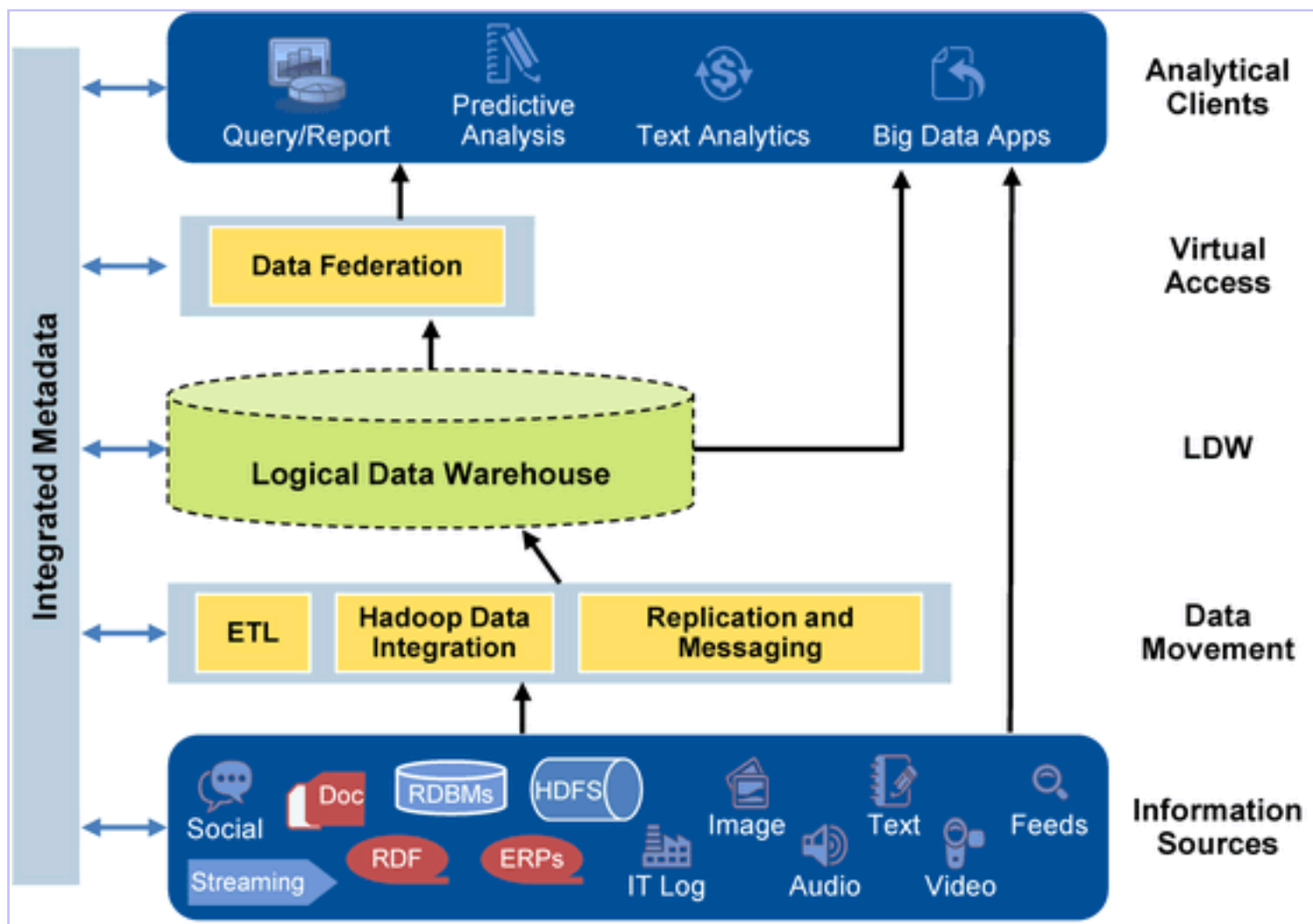
LDW

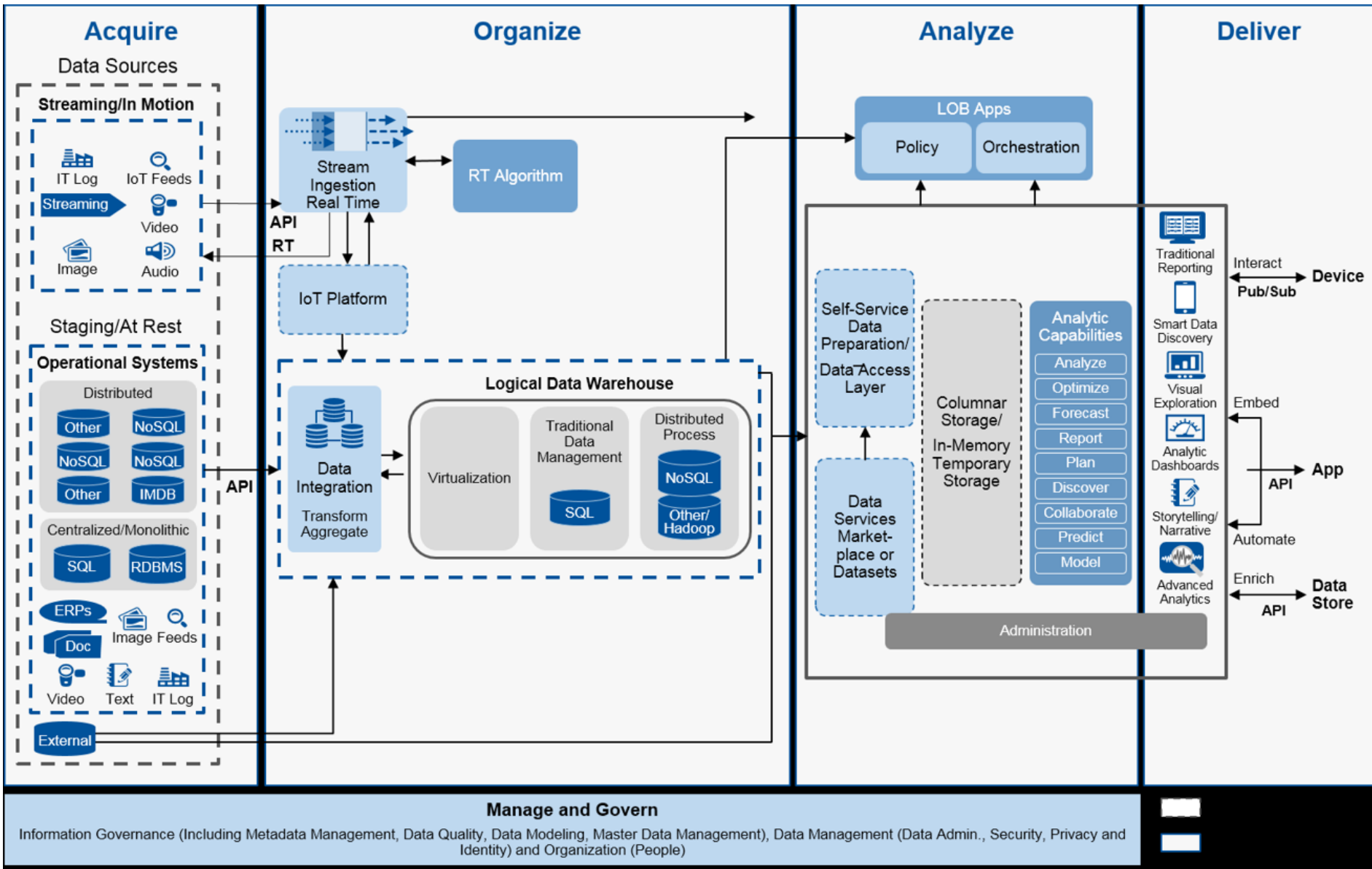
Componentes

- Repository Management
- Data Virtualization
- Distributed Processes
- Auditing Statistics and Performance Evaluation Services
- Metadata Management
- SLA Management

LDW

Arquitetura







Ferramenta de Visualização Prática



Prática

- Tableau

- <https://www.tableau.com/pt>

- Power BI

- <https://powerbi.microsoft.com/pt-br/>

- Portal da Transparência

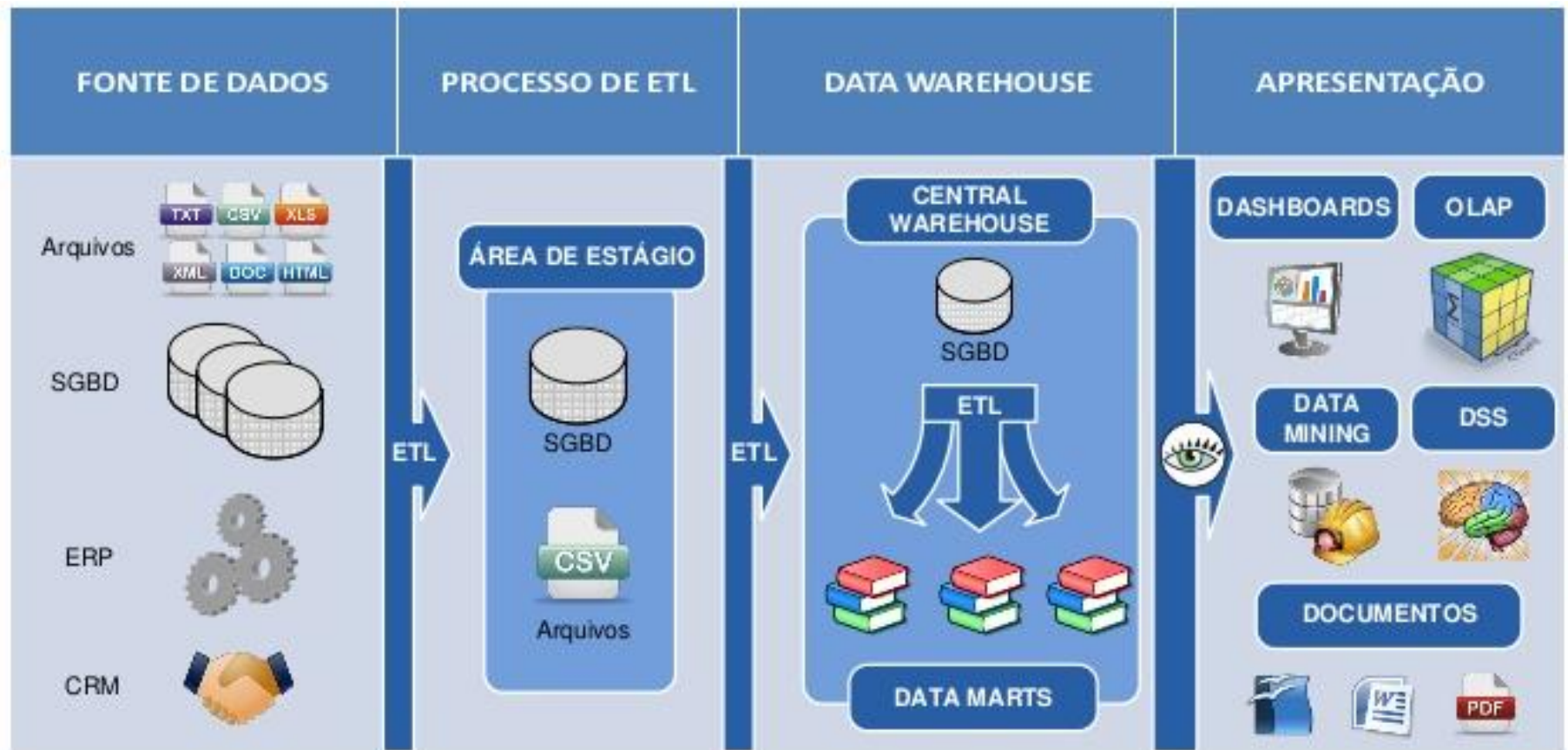
- <http://www.portaltransparencia.gov.br/downloads/>



Avaliação

Avaliação

1. Propor uma arquitetura de BI para uma empresa



Avaliação

1. Selecionar um estudo de caso (sugestões nos links abaixo)
 - ☐ www.portaltransparencia.gov.br/downloads/
 - ☐ www.kaggle.com/
2. Elaborar a matriz de necessidades contendo os indicadores e as perspectivas de análise.
3. Construir o diagrama multidimensional
 - ☐ Dica: utilizar a ferramenta MySQL Workbench
4. Construir um *dashboard* no Power BI ou Tableau para ajudar os diretores do estudo de caso a tomarem alguma decisão.

Avaliação


Orientações:

- Poderá ser em equipe (2 integrantes) ou individual
- Deverá ser entregue por e-mail os seguintes itens:
 - ☐ Projeto da Arquitetura BI da empresa
 - ☐ Matriz de necessidades
 - ☐ Diagrama multidimensional em uma ferramenta de modelagem de dados
 - ☐ Arquivo do Power BI ou Tableau com o dashboard
- Data de entrega: 11/11/2017 23:59 (1,0 a menos por cada dia de atraso)
 - ☐ Trabalho entregue antecipadamente(até 04/11/2017) valerá até 11,00 pontos
- Email: fssiqueira@gmail.com
 - ☐ Assunto: [UNI7-CD 2017 T2] FBI - Entrega Trabalho

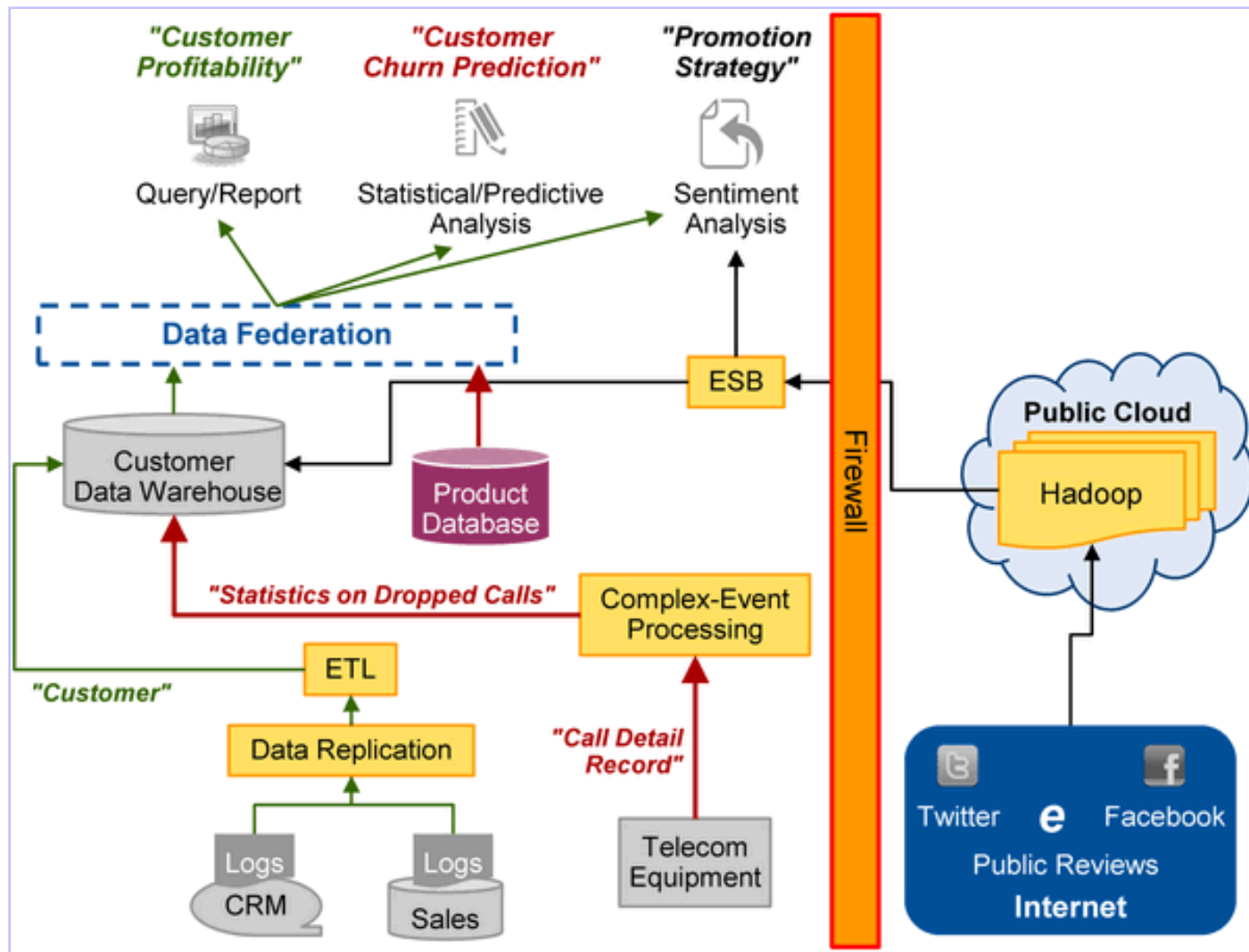


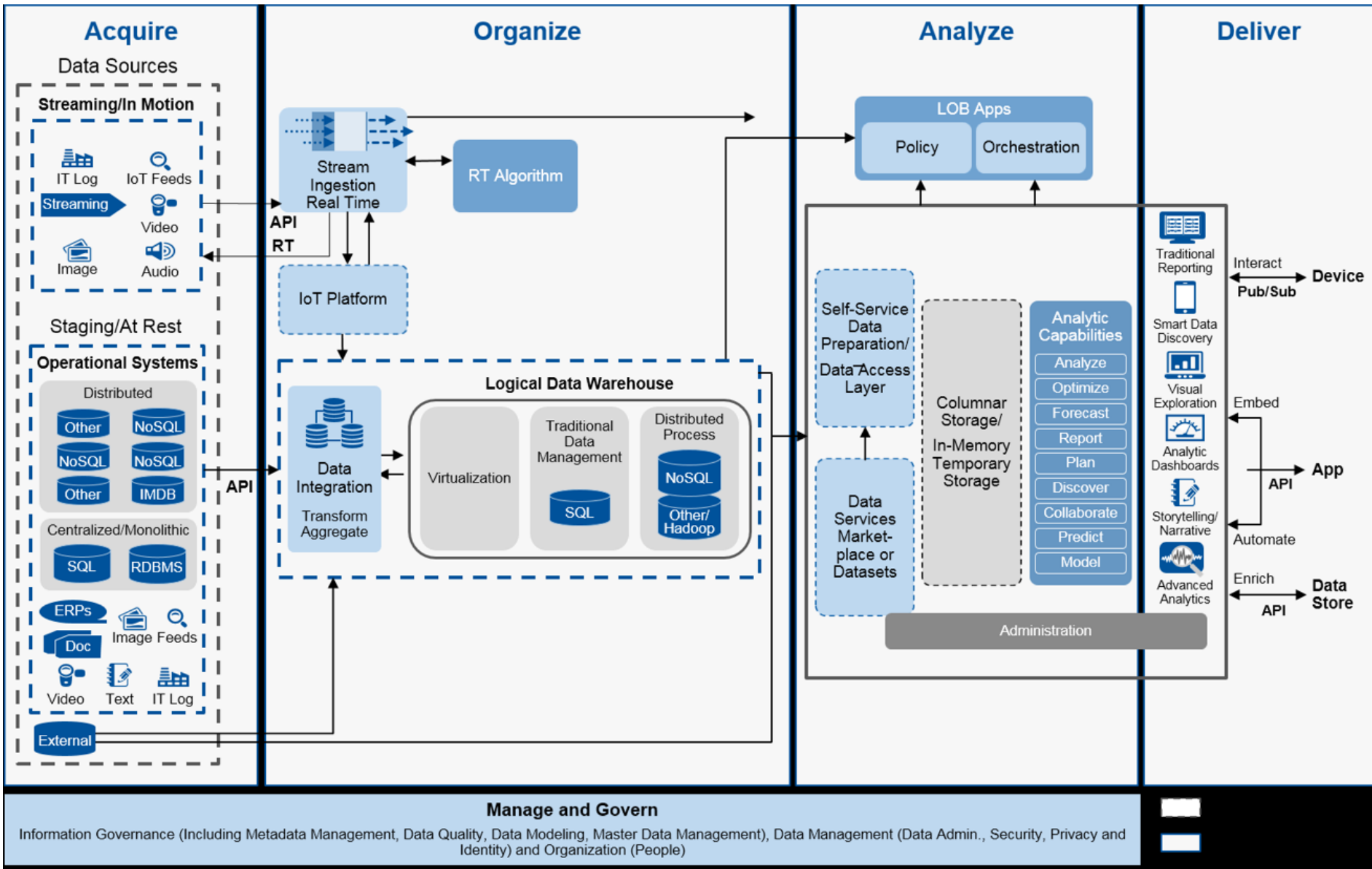
Considerações Finais

A Nova Plataforma do BI

 Analytic Workflow Component	Traditional BI Platform	Modern BI Platform
Data source	Upfront dimensional modeling required (IT-built star schemas)	Upfront modeling not required (flat files/flat tables)
Data ingestion and preparation	IT-produced	IT-enabled
Content authoring	Primarily IT staff, but also some power users	Business users
Analysis	Predefined, ad hoc reporting, based on predefined model	Free-form exploration
Insight delivery	Distribution and notifications via scheduled reports or portal	Sharing and collaboration, storytelling, open APIs

A Nova Plataforma do BI







Gestão de Dados

- “Controlar e alavancar eficazmente o uso dos ativos de dados e sua missão e objetivos são atender e exceder às necessidades de informação de todos os envolvidos (stakeholders) da empresa em termos de disponibilidade, segurança e qualidade.”

(DMBOK, 2009)

Gestão de Dados



Especialização em Ciência de Dados com Big Data, BI e *Data Analytics*



Prof. Fernando Siqueira
fssiqueira@gmail.com

Especialização em Ciência de Dados com Big Data, BI e *Data Analytics*



NUNCA DESISTA DOS SEUS SONHOS



Prof. Fernando Siqueira

fssiqueira@gmail.com