

Fundamentos de *Business Intelligence*

2º Encontro

Prof. MSc. Fernando Siqueira

fernando.siqueira@uni7.edu.br



Business Intelligence

Visão Geral

Conceitos



- Processos inteligentes de coleta, organização, análise, compartilhamento e monitoração de dados contidos em **Data Warehouse** (DW) ou **Data Mart** (DM), gerando informações para o suporte à tomada de decisões no ambiente empresarial.
- **BI não é um produto** novo ou uma tecnologia, mas sim **um conjunto de conceitos, metodologias, ferramentas, banco de dados** que, fazendo uso de dados extraídos de uma organização, apoia a tomada de decisões.

Conceitos



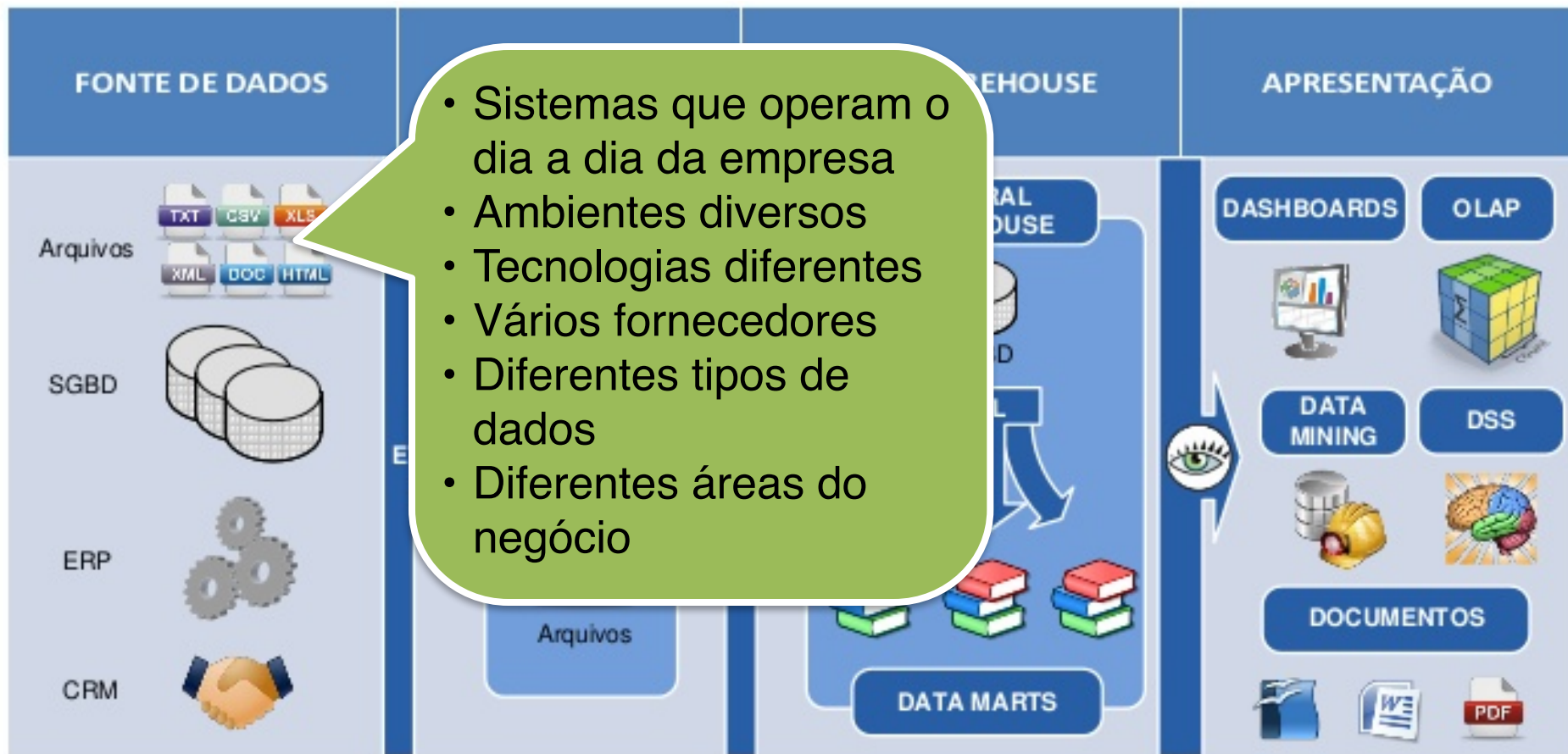
- **Business Intelligence** é um conceito muito atual que vai além da gestão empresarial. Entre outras coisas, envolve a utilização de **produtos** e **soluções** com tecnologia analíticas de ponta que permitem transformar dados armazenados em bases de dados em informações que auxiliam os diversos níveis de uma empresa na tomada de decisões.
- **Inteligência** é o produto da transformação de dados em informação, após ser analisada ou inserida em um determinado ambiente. Esta informação transformada, aplicada a um determinado processo de decisão, gera **vantagem competitiva** para a organização.

Objetivos

- Disponibilizar informações da maneira e formato correto e no tempo certo para que a empresa possa tomar decisões melhores e mais rápidas.
- Permitir a realização de análises e projeções, de forma a melhorar os processos relacionados às tomadas de decisão e ao ganho de vantagens competitivas.



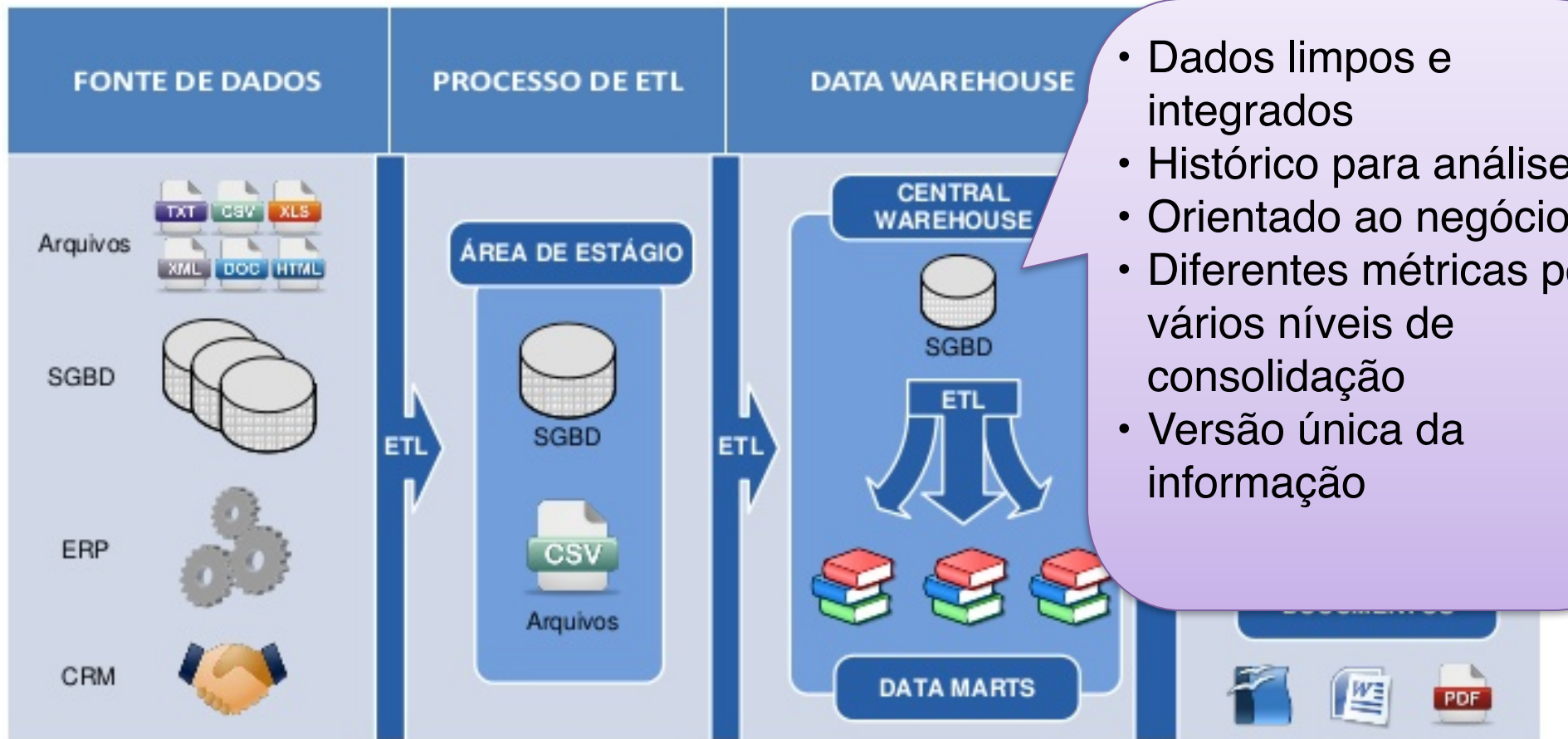
O Processo BI



O Processo BI

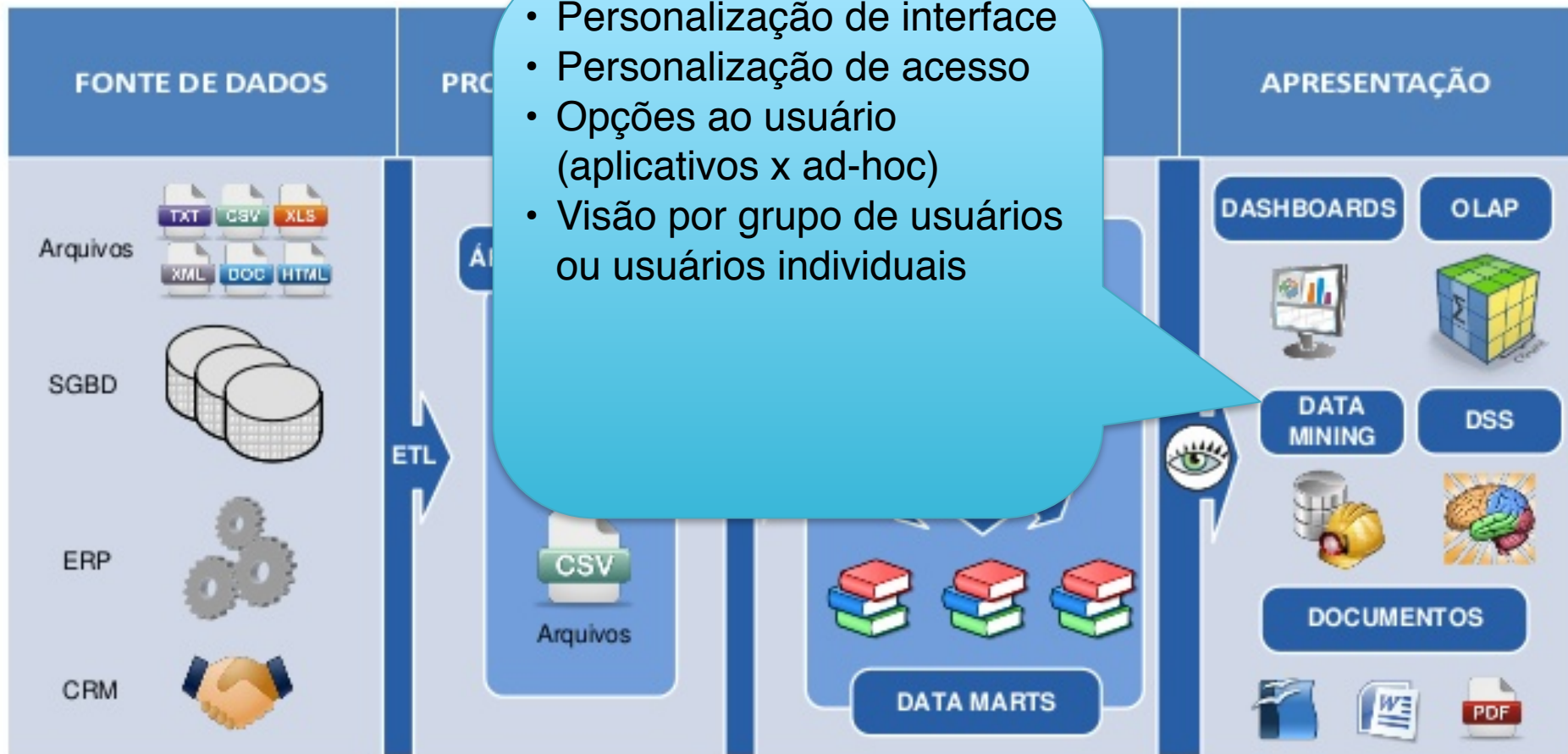


O Processo BI



- Dados limpos e integrados
- Histórico para análise
- Orientado ao negócio
- Diferentes métricas por vários níveis de consolidação
- Versão única da informação

O Processo BI

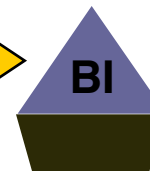


Hierarquia

Dos Dados Operacionais para
Informações gerenciais



Business Intelligence



Dados organizados para os modelos de
negocio

DATA WAREHOUSE



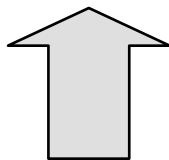
Integração e Validação dos dados



**Banco de Dados
Operacional - ERP**



Dados Operacionais

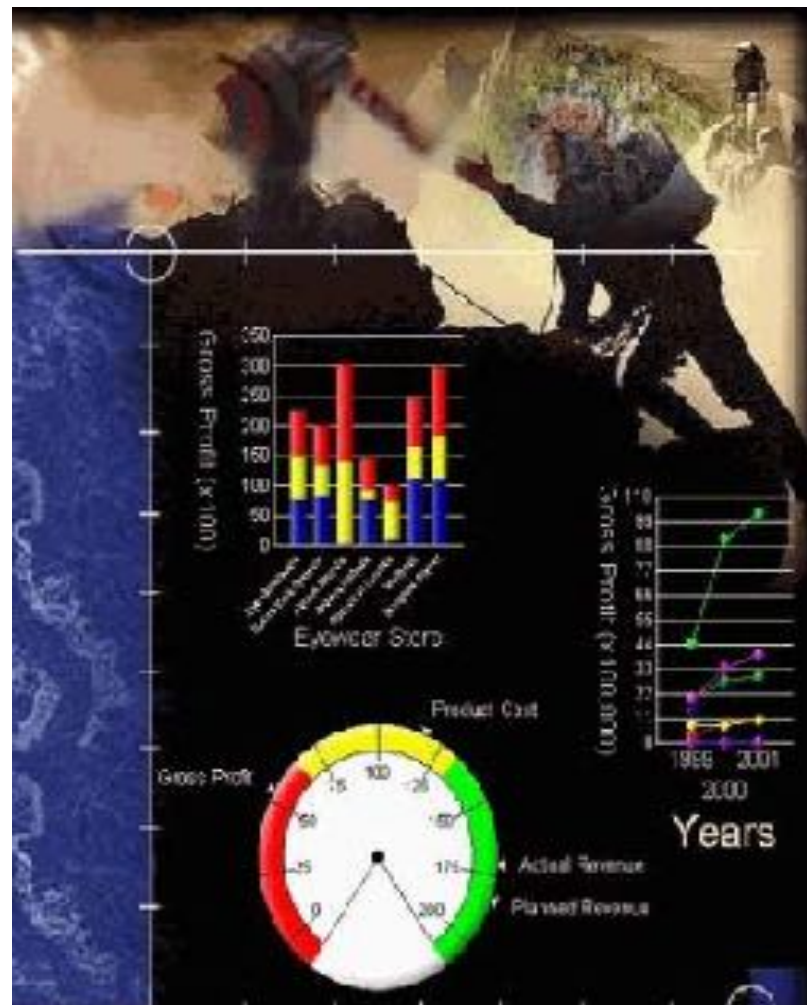


**Dados Externos
Fontes diversas**



Benefícios e Vantagens

- As ferramentas de *Business Intelligence* possibilitam a busca e interpretação de informações armazenadas na corporação, garantindo maior exatidão nas tomadas de decisão;
- Permite à Gerência estabelecer uma aproximação integrada e colaborativa para capturar, criar, organizar e usar todos os ativos de informação de uma empresa;
- Visão clara sobre novos negócios;
- Conhecimento sobre o negócio;
- Antecipação às mudanças bruscas no mercado;
- Antecipação às ações sobre os competidores;
- Aprendizado através do sucesso e falhas dos concorrentes;
- Auxílio na implementação de novas ferramentas gerenciais;
- O **Business Intelligence**, como interface, interfere, transforma e torna verdadeira todas estas informações e as transforma em conhecimento estratégico.

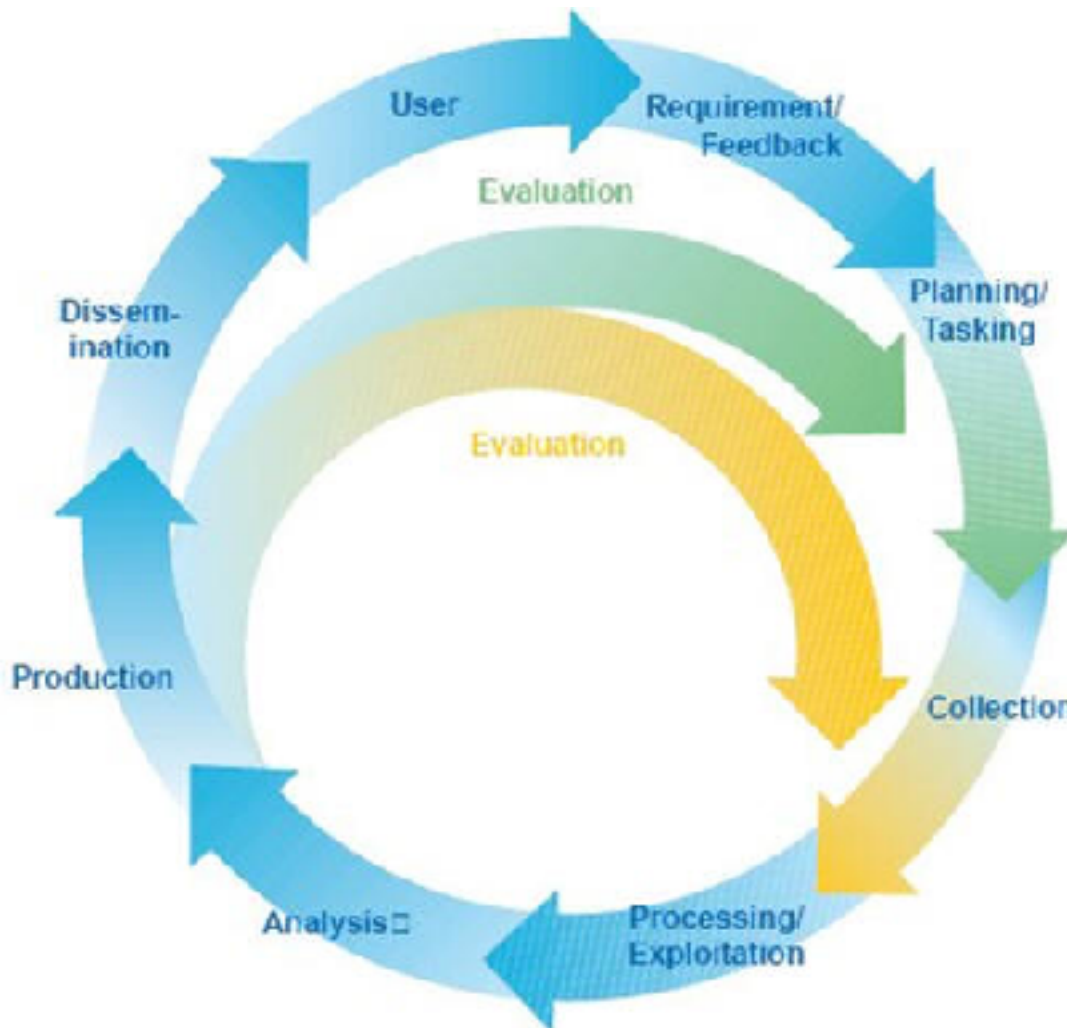


Benefícios e Vantagens

- Antecipa mudanças de mercado
- Antecipa ações dos competidores
- Descobre novos ou potenciais clientes
- Entra em novos negócios
- Possibilita a revisão de suas práticas de negócios
- Aumenta seu grau de competitividade

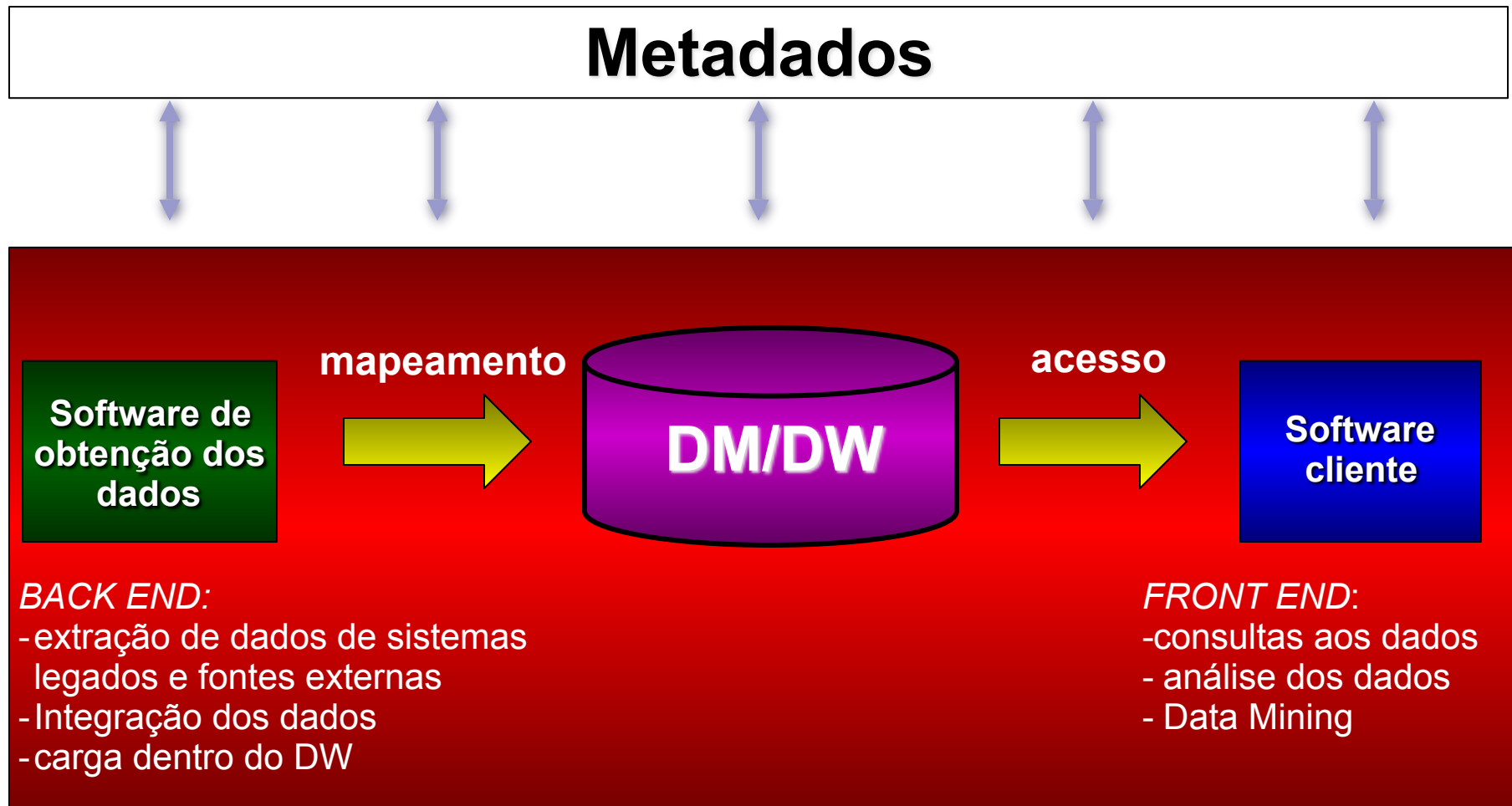


Como Criar e Utilizar a Inteligência

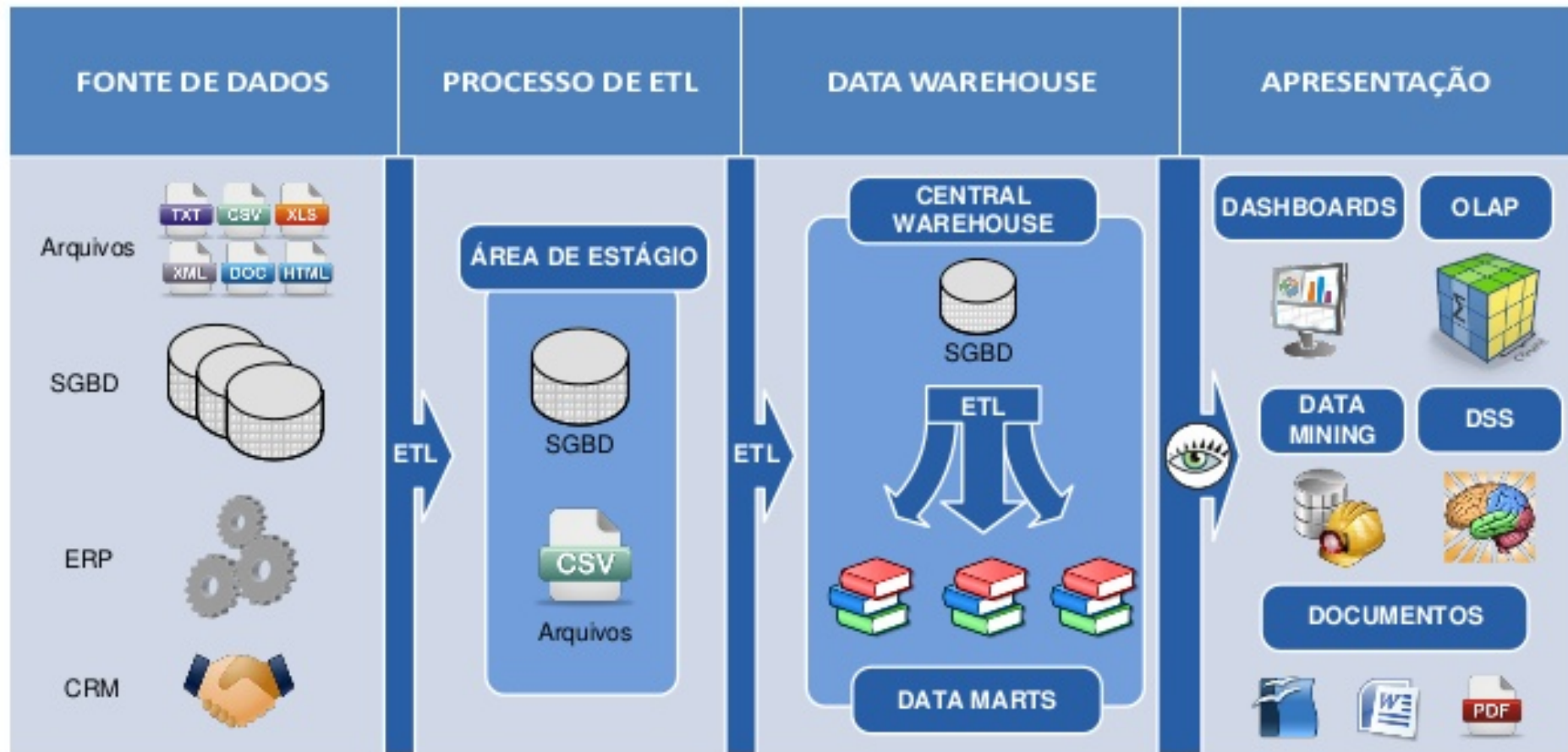


1. Análise
2. Produção
3. Disseminação
4. Uso
5. Exigência/Feedback
6. Planejamento e atribuição das tarefas
7. Coleta
8. Processamento e Exploração

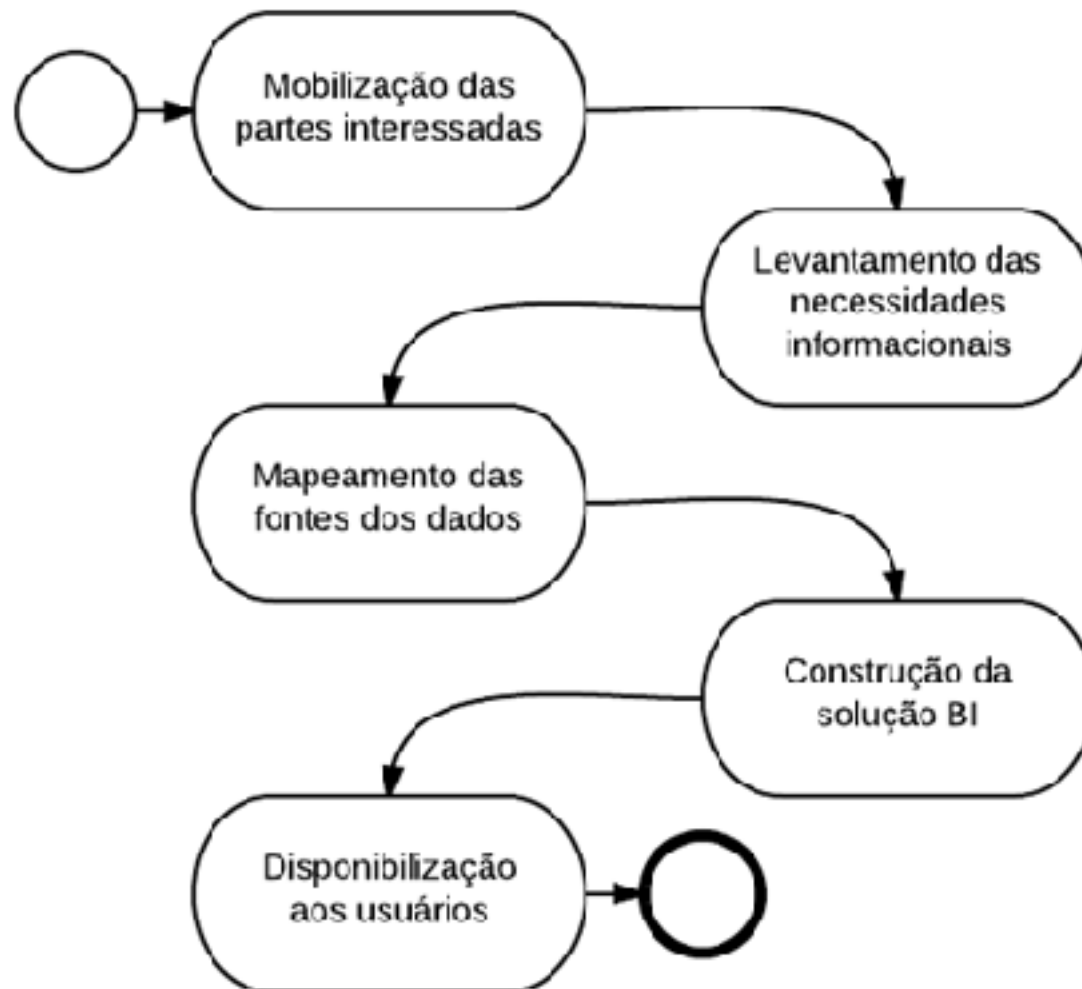
Principais Componentes do BI



Arquitetura Genérica de Dados



Processo de Desenvolvimento

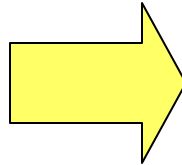
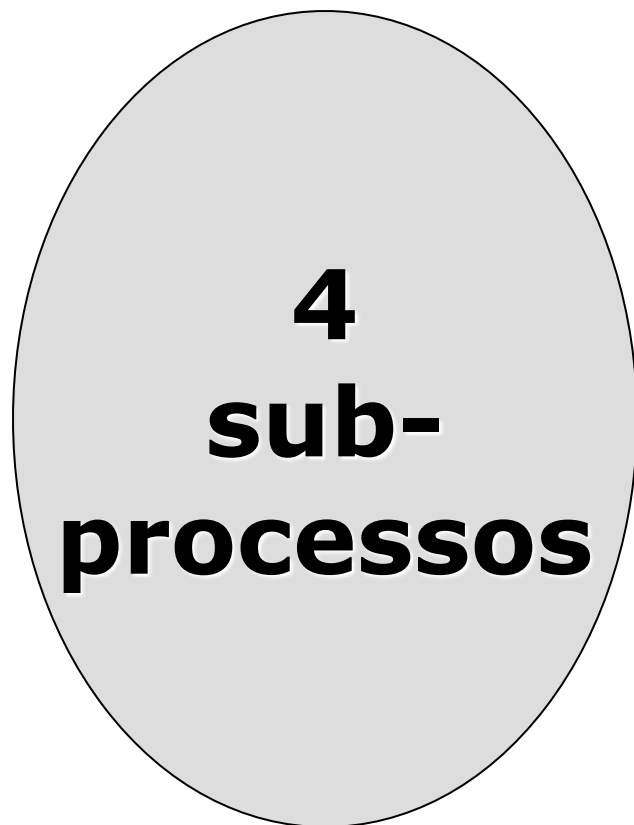


Processo de Desenvolvimento



KickOff	Planejamento	Indicadores	Modelo Lógico	Modelo Físico	Desenvolvimento
Avaliação Inicial	Análise situação atual	Análise indicadores existentes	Identificação das métricas	Mapeamento de fontes de dados	Desenvolvimento
Consenso sobre a necessidade Entendimento de conceitos	Coleta de materiais Análise processo gestão	Revisão dos indicadores Avaliação de metas	Identificação das dimensões Mapeamento de processos	Avaliação de Plataforma de BI Critérios de Data Quality	Gestão de projeto
Definir Sponsor	Definição key users e equipe	Detalhamento dos indicadores	Mapeamento de sistemas e funcionalidades	Estratégia de Atualização	
Identificar benefícios Aprovação Metodologia	Definição escopo	Workshop de revisão	Nível de detalhamento por métrica Contextualização de métricas e atributos		
Entregáveis:	Entregáveis:	Entregáveis:	Entregáveis:	Entregáveis:	Entregáveis:
Workshop com líderes	Plano de desenvolvimento Macro cronograma	Indicadores e conceitos	Modelo lógico Cronograma detalhado	Modelo físico dos dados Dicionário de Dados	Escopo desenvolvido Treinamento
			Escopo definido		

Processo de Desenvolvimento



Modelagem Dimensional

Construção dos DM's / DW

Extração e Transformação
(fontes internas e externas)

Desenvolvimento das Aplicações

Pontos de Atenção

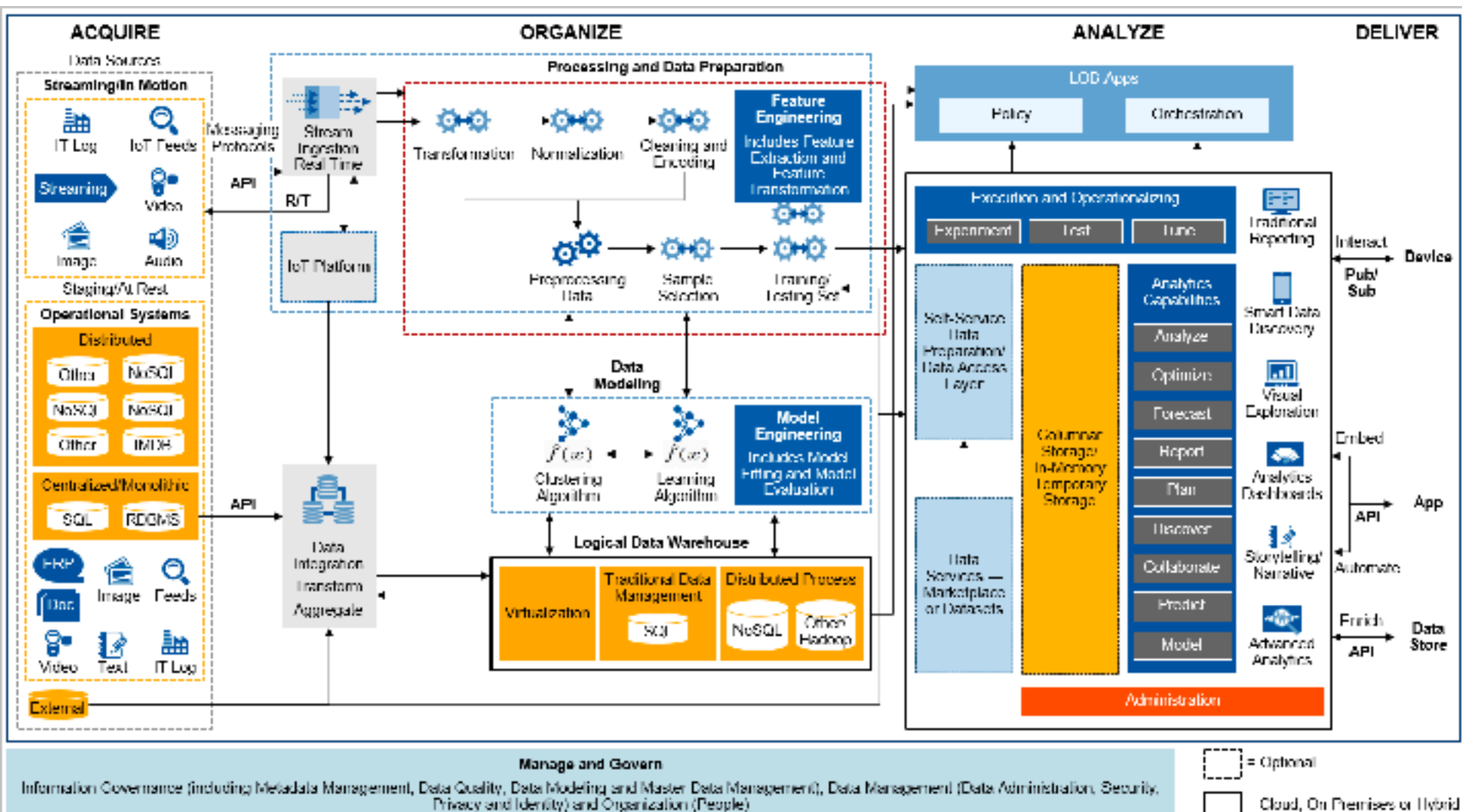


- Integração de dados e metadados de várias fontes
- Qualidade dos dados: limpeza e refinamentos
- Sumarização e agregação de dados
- Sincronização das fontes com o *data warehouse* para assegurar atualidade
- Problemas de desempenho relacionados ao compartilhamento do mesmo ambiente computacional para abrigar os BDs corporativos operacionais e o *data warehouse*.

Fatores de Sucesso!

- Compromisso da alta administração
- Um time dedicado de analistas de BI
- Um sólido modelo de dados (DW/DM/LDW)
- Um plano de implementação bem elaborado
- Administração dinâmica do sistema
- Ferramenta de acesso aos dados de fácil uso e intuitiva pelos usuários finais







Evolução das Plataformas de BI e Analytics





As of February 2017



COMPLETENESS OF VISION



As of February 2018

© Gartner, Inc



As of January 2019

© Gartner, Inc



Evolução das Plataformas de Data Science e ML







As of November 2018

© Gartner, Inc



Evolução das Plataformas de Data Management for Analytics

Figure 1. Magic Quadrant for Data Management Solutions for Analytics





Data Warehouse

DW & BI





Data Warehouse

Conceitos

“É a separação física dos sistemas de dados operacionais de uma organização, de seus sistemas de suporte à decisão”. Harry Singh

“É um banco de dados, voltado para suporte à decisão, não volátil, variante no tempo e orientado a assuntos”. Bill Inmon

“Não consiste apenas de dados, mas também em um conjunto de ferramentas para consultar, analisar e apresentar informações”.

Ralph Kimball



Data Warehouse

Histórico

- Criado pela IBM na década de 60 com o nome de *Information Warehouse*
- Tornou-se viável com o surgimento de novas tecnologias para armazenar e processar uma grande quantidade de dados
- O nome Data Warehouse foi dado por Willian H. Inmon, considerado o pai dessa tecnologia, em 1990
- Aproximadamente em 1994 ficou mais conhecido

Data Warehouse

Histórico

- ✓ Mainframe computers
- ✓ Simple data entry
- ✓ Routine reporting
- ✓ Primitive database structures
- ✓ Teradata incorporated

- ✓ Centralized data storage
- ✓ Data warehousing was born
- ✓ Inmon , *Building the Data Warehouse*
- ✓ Kimball , *The Data Warehouse Toolkit*
- ✓ EDW architecture design

- ✓ Big Data analytics
- ✓ Social media analytics
- ✓ Text and Web Analytics
- ✓ Hadoop , MapReduce , NoSQL
- ✓ In-memory , in-database



- ✓ Mini / personal computers (PCs)
- ✓ Business applications for PCs
- ✓ Distributer DBMS
- ✓ Relational DBMS
- ✓ Teradata ships commercial DBs
- ✓ *Business Data Warehouse* coined

- ✓ Exponentially growing data Web data
- ✓ Consolidation of DW / BI industry
- ✓ Data warehouse appliances emerged
- ✓ Business intelligence popularized
- ✓ Data mining and predictive modeling
- ✓ Open source software
- ✓ SaaS , PaaS , Cloud Computing



Data Warehouse

Para que serve?

- Armazenar dados históricos usados no processo de tomada de decisão
- Criar uma visão única e centralizada dos dados que estavam dispersos em diversos bancos de dados
- Permitir que os usuários finais executem consultas, gerem relatórios e façam análises



Data Warehouse

Principais Características

- De acordo com a definição de Inmon, um DW deve ser:
 - Orientado a assunto
 - Integrado
 - Não-volátil
 - Variável com o tempo

Data Warehouse

Principais Características

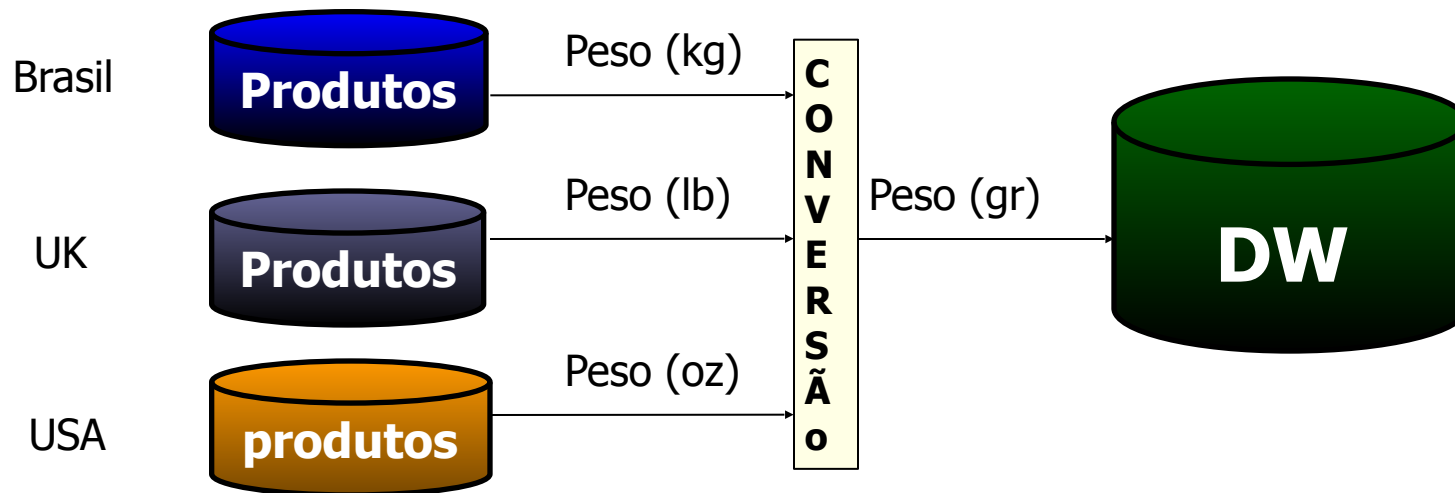
- Orientação a assunto
 - As informações são organizadas de modo a facilitar a análise dos dados, por exemplo: vendas, marketing, etc
 - Os dados são organizados por assunto e não por aplicação, como em BDs operacionais



Data Warehouse

Principais Características

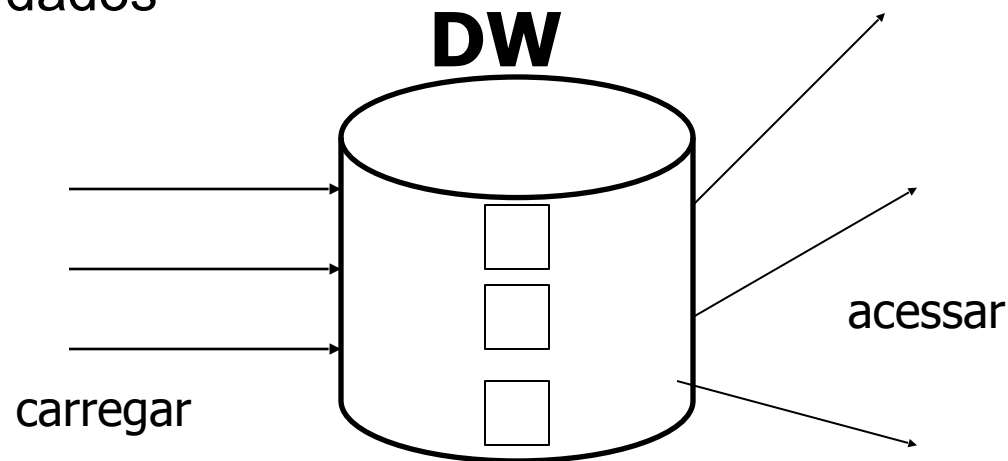
- Integração
 - Dados do DW provêm de diversas fontes
 - Dados podem ser sumarizados ou eliminados
 - Formato dos dados deve ser padronizado para uniformizar nomes, unidades de medidas, etc.



Data Warehouse

Principais Características

- Não-volátil
 - Uma vez que corretamente gravados no DW, os dados não podem ser alterados ou eliminados
 - Com isso, garante-se que consultas subsequentes a um dado produzirão o mesmo resultado
 - As operações possíveis são as de carga dos dados e de acesso aos dados



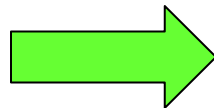
Data Warehouse

Principais Características

- Variável com o tempo
 - Os dados no DW são relativos a um determinado instante de tempo



Produto	Preço
Caneta	0,50
Lápis preto	0,30
...	...

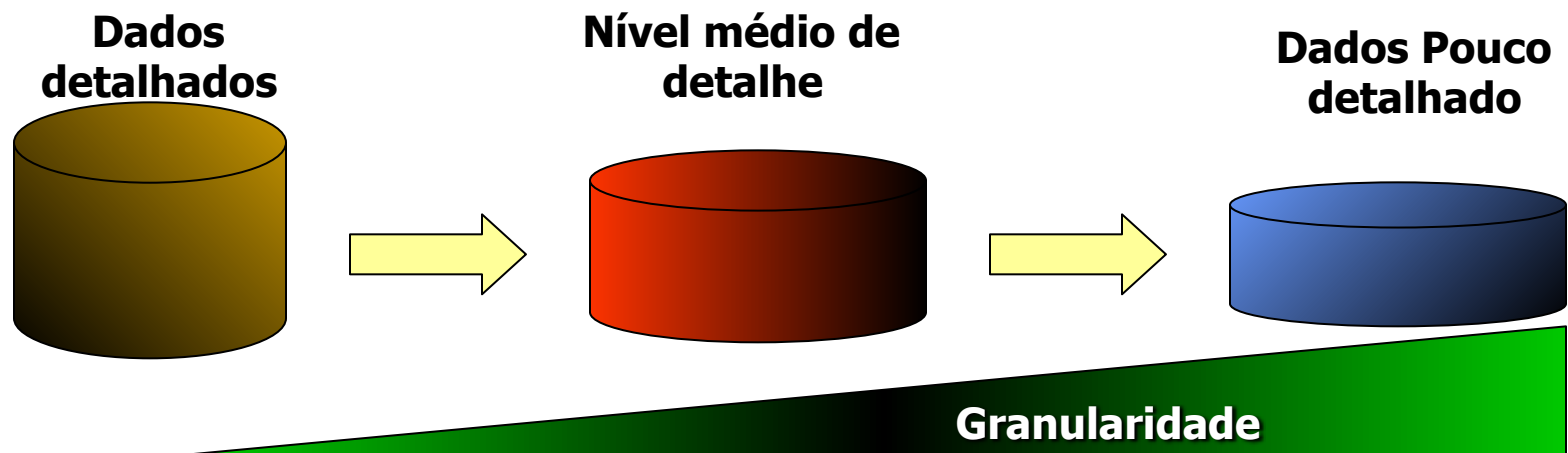


Produto	Jan/2019	Fev/2019	Mar/2019
Caneta	0,40	0,45	0,50
Lápis preto	0,25	0,28	0,30
...

Data Warehouse

Principais Características

- Granularidade
 - É o nível de detalhe, ou sumarização, ou resumo dos dados no DW
 - O nível de granularidade
 - É de extrema importância no projeto DW
 - Tem efeito direto no tamanho do BD e no tipo de análise que o BD pode suportar





Data Warehouse

Granularidade

- Definir a granularidade adequada é vital para que o DW atenda aos seus objetivos
 - **Menor Granularidade:**
 - Mais detalhes → Mais dados → Análises mais longas → Informação mais detalhada
 - **Maior Granularidade:**
 - Menos detalhes → Menos dados → Análises mais curtas → Informação menos detalhada
- Normalmente, para evitar que se perca informação, são criados vários níveis de granularidade



Data Warehouse

Granularidade

- Procurar um equilíbrio entre o gerenciamento do volume dos dados e a armazenagem dos dados
 - Cuidado com um nível muito alto que impossibilite a utilização detalhada
- Passos:
 - Usar bom senso → criar uma pequena parte de DW, deixar o usuário acessar os dados, colher feedback
 - Estimar o número de linhas de dados e do dispositivo de armazenamento.
 - Planejar → quanto espaço de disco é necessário e se são necessários mais de um nível de granularidade



Data Warehouse

Granularidade

- Exemplo: rastreamento de compras com cartão de crédito
 - 3 anos -> 36 meses
 - número de contas -> 50 milhões
 - número médio de compras por mês/conta -> 30
 - número de registros -> $36 \times 50.000.000 \times 30 = 54$ bilhões
 - tamanho do registro -> 8 campos de 4 bytes
 - tamanho básico da tabela -> $54 \text{ bilhões} \times 8 \times 4 = 1730 \text{ Gb} = 1.73 \text{ TB}$

[Inmon]



Data Warehouse

Granularidade

- Processo de sumarização
 - Aplica um novo esquema de modo a condensar os dados
 - ✓ Ex.: armazenar totais, médias, etc.
- Processo de envelhecimento
 - Transfere os dados antigos do HD para fita, CD, etc.
 - Mantém o nível de detalhe para que nenhuma informação seja perdida

Data Warehouse

Granularidade – Exemplo Cenário Bancário

Operacional – 60 dias

- número conta
- data ocorrência
 - valor
 - caixa
 - local
 - identificação (débito/crédito)
 - para quem
 - saldo da conta

DW – até 10 anos

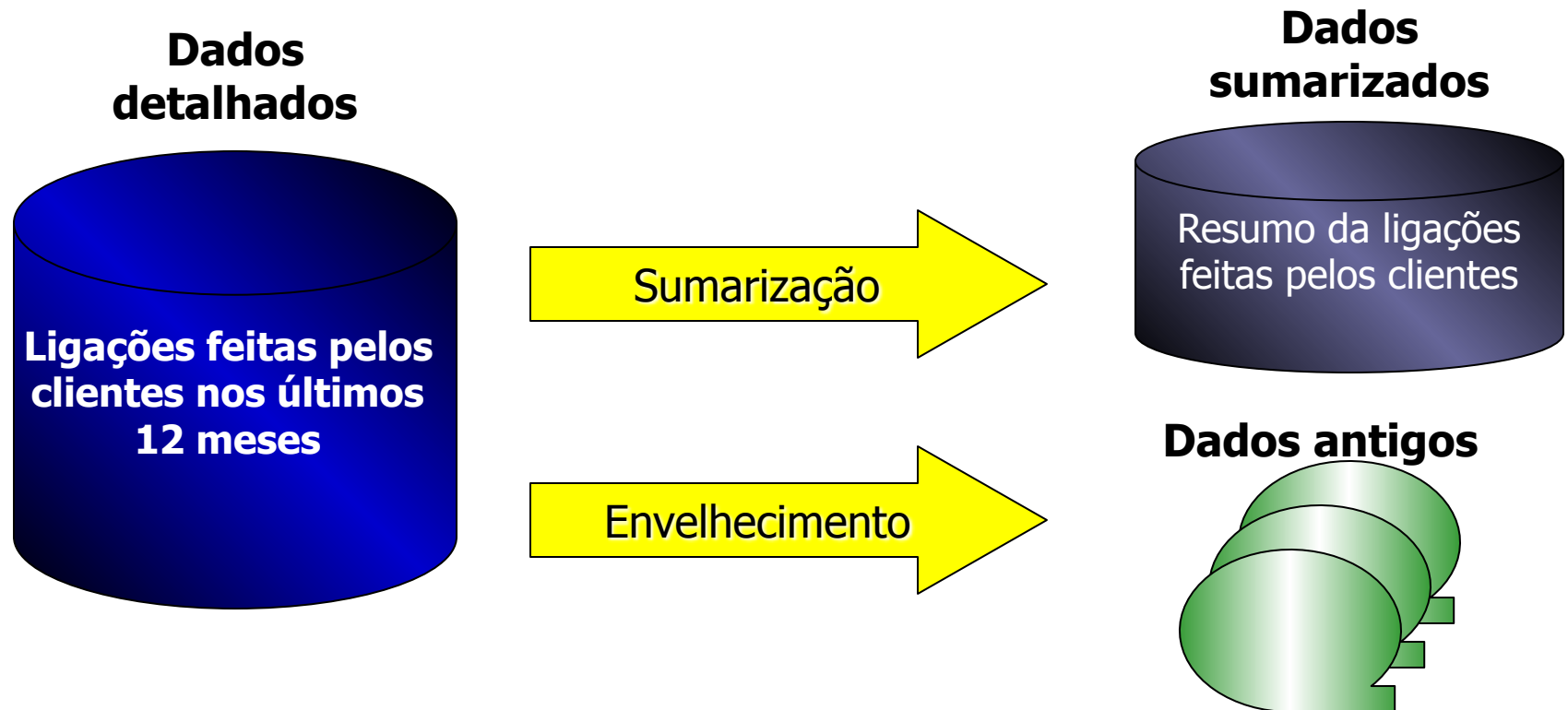
- número conta
- mês
 - número de transações
 - saques
 - depósitos
 - saldo inicial
 - saldo final
 - maior saldo da conta
 - menor saldo da conta
 - saldo médio da conta

DW – até 10 anos

- número conta
- data ocorrência
 - valor
 - identificação (débito/crédito)
 - saldo da conta

Data Warehouse

Granularidade – Exemplo Cenário Empresa Telefônica



Data Warehouse

Granularidade – Exemplo Cenário Empresa Telefônica

Dados Detalhados

Ligações
Origem
Destino
Início
Fim
Tarifa
Status

Num. Registros:

Ligações nos últimos
12 meses

Dados Sumarizados

Ligações
Cliente
Mês
Pulsos
Valor conta
Longa dist.
Vencimento

Num. Registros:

Contas emitidas pela
empresa

Dados Antigos

Ligações
Origem
Destino
Início
Fim
Tarifa
Status

Num. Registros:

Ligações efetuadas
pela empresa

Data Warehouse

Granularidade – Exemplo Cenário Empresa Telefônica

- Quanto menor a granularidade, mais detalhada é a informação disponível
 - No exemplo anterior, poderíamos determinar se o cliente A ligou para B na semana passada
 - Também poderíamos verificar se A faz muitas chamadas de longa distância

Dados Detalhados

Ligações
Origem
Destino
Início
Fim
Tarifa
Status

Data Warehouse

Granularidade – Exemplo Cenário Empresa Telefônica

- Durante o processo de sumarização, algumas informações podem ser perdidas
 - Não seria possível saber se A ligou para B
 - É possível verificar o padrão de consumo de A

Dados Sumarizados

Ligações

Cliente

Mês

Pulsos

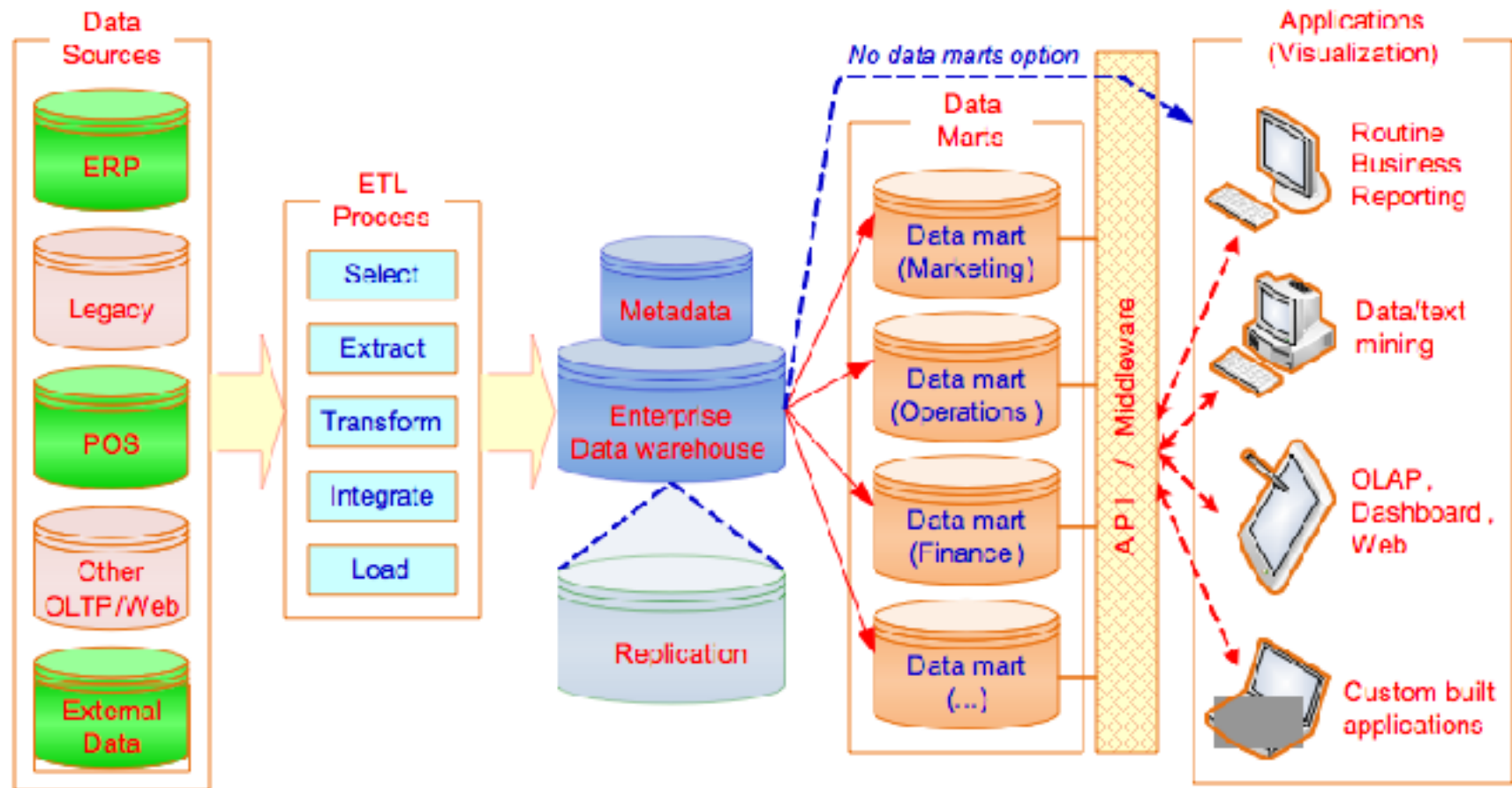
Valor conta

Longa dist.

Vencimento

Data Warehouse

Arquitetura Genérica





Data Warehouse

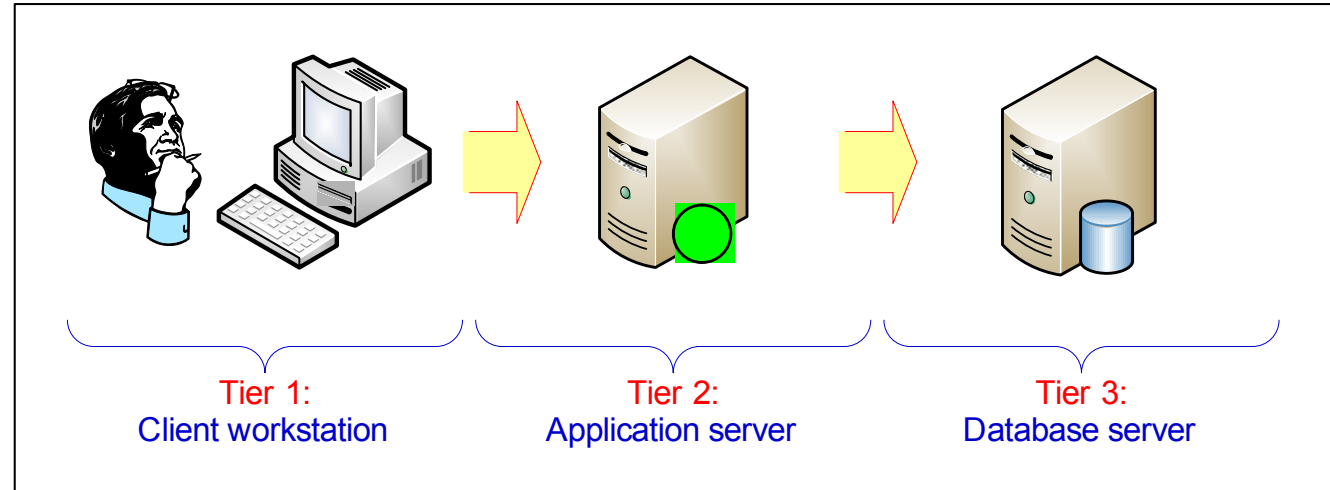
Arquiteturas

- 3 Camadas
 - Software de aquisição de dados (back-end)
 - O data warehouse contém dados e o software
 - Software cliente (front-end) permite os usuários acessar e analisar dados do *data warehouse*
- 2 Camadas
 - As duas primeiras camadas na arquitetura de três camadas são combinadas em uma
 - ... às vezes há apenas uma camada?

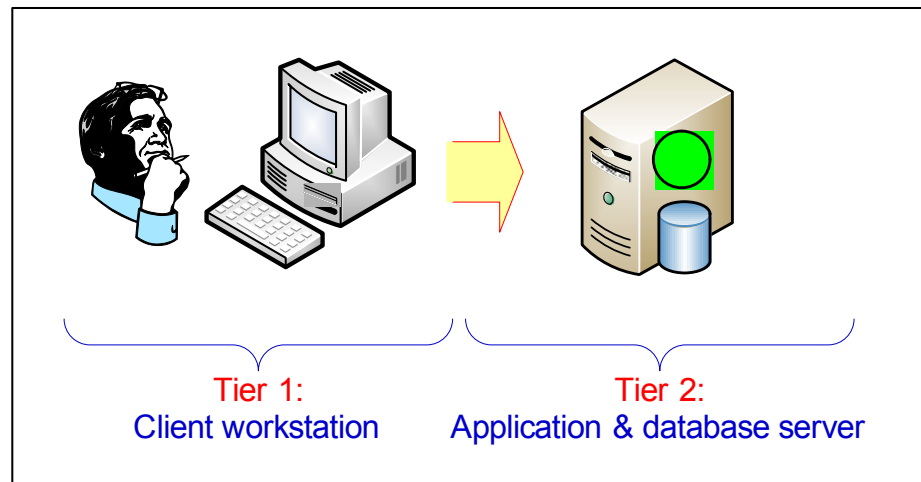
Data Warehouse

Arquiteturas

3-tier architecture



2-tier architecture



1-tier Architecture ?



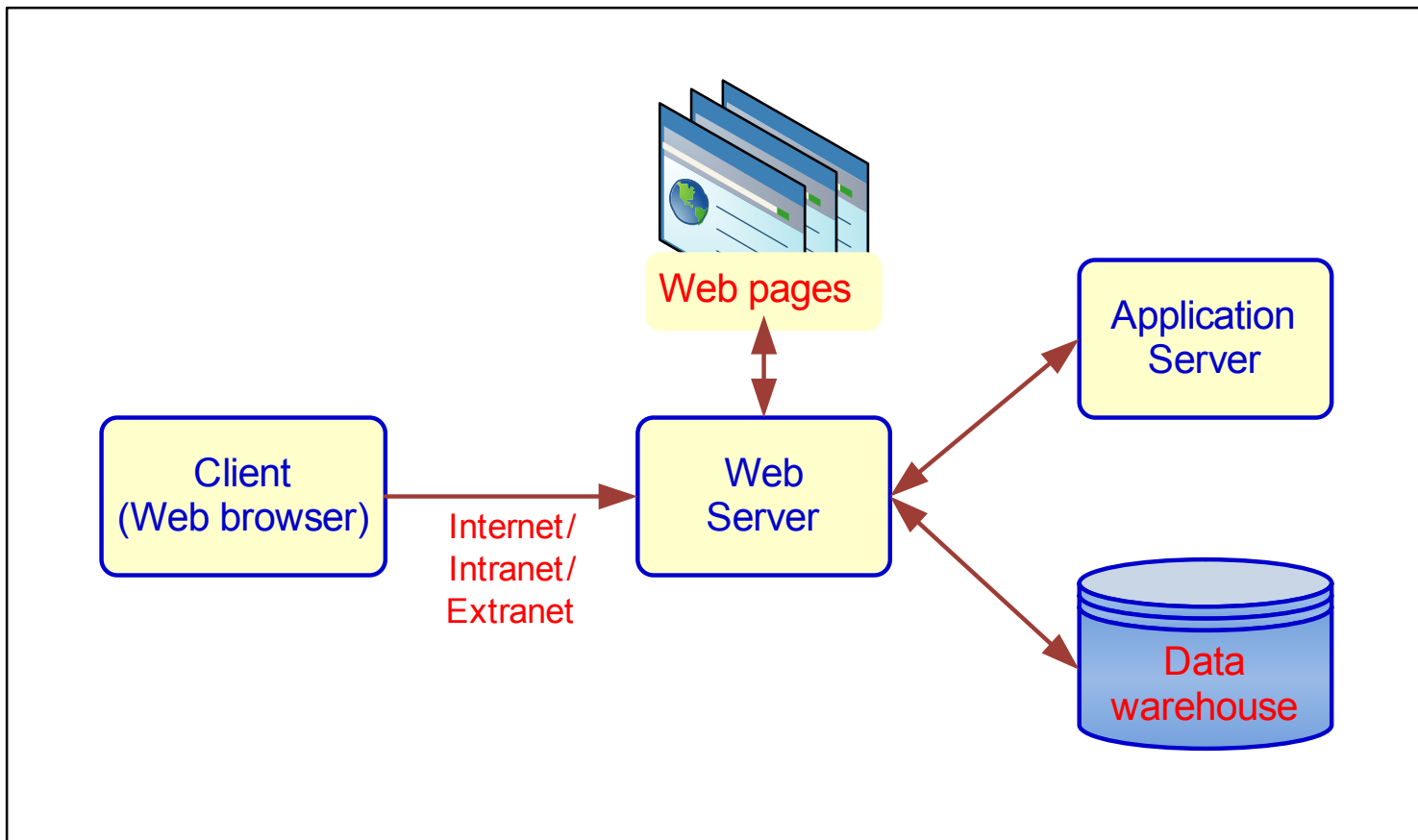
Data Warehouse

Arquiteturas

- Questões a serem consideradas ao decidir qual arquitetura usar:
 - Qual sistema de gerenciamento de banco de dados (DBMS) deve ser usado?
 - O processamento será paralelo e/ou particionamento será utilizado?
 - As ferramentas de migração de dados serão usadas para carregar o data warehouse?
 - Que ferramentas serão usadas para apoiar a recuperação e análise de dados?

Data Warehouse

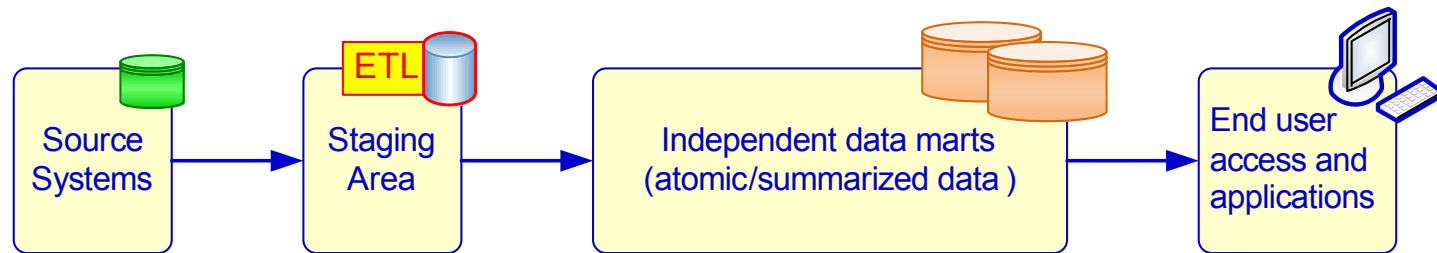
Arquitetura baseada na Web



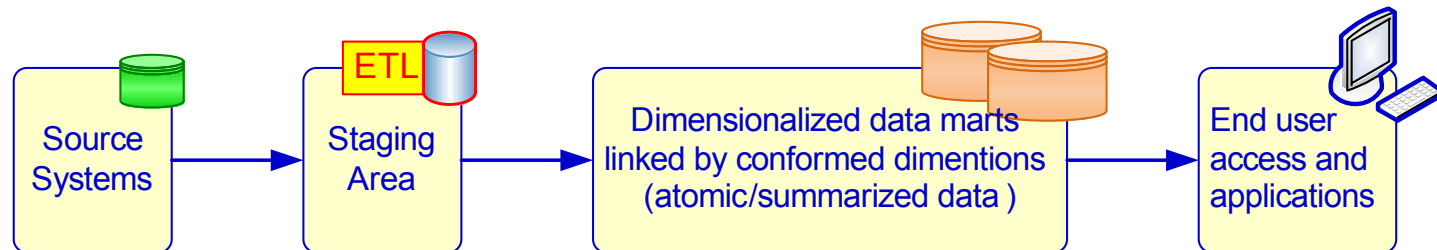
Data Warehouse

Arquiteturas Alternativas

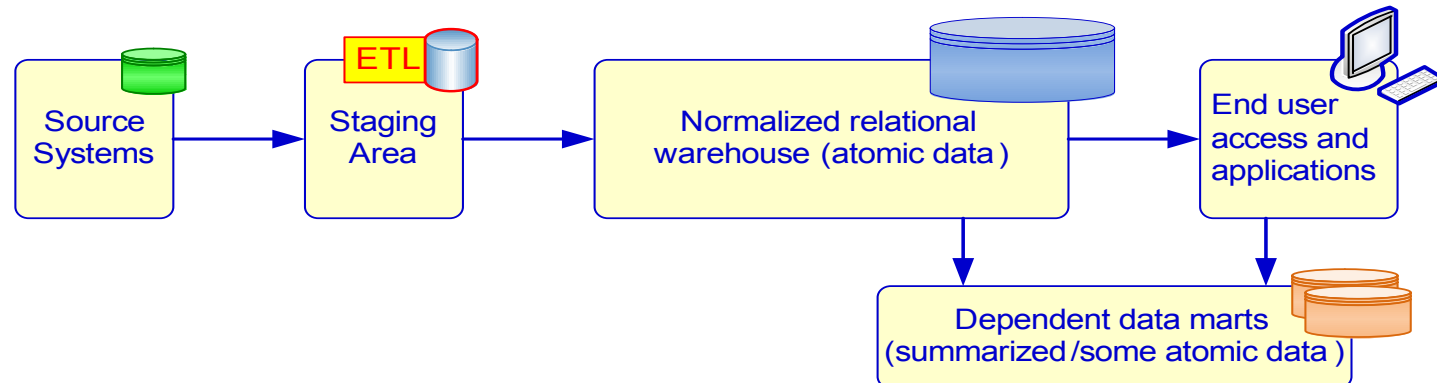
(a) Independent Data Marts Architecture



(b) Data Mart Bus Architecture with Linked Dimensional Datamarts



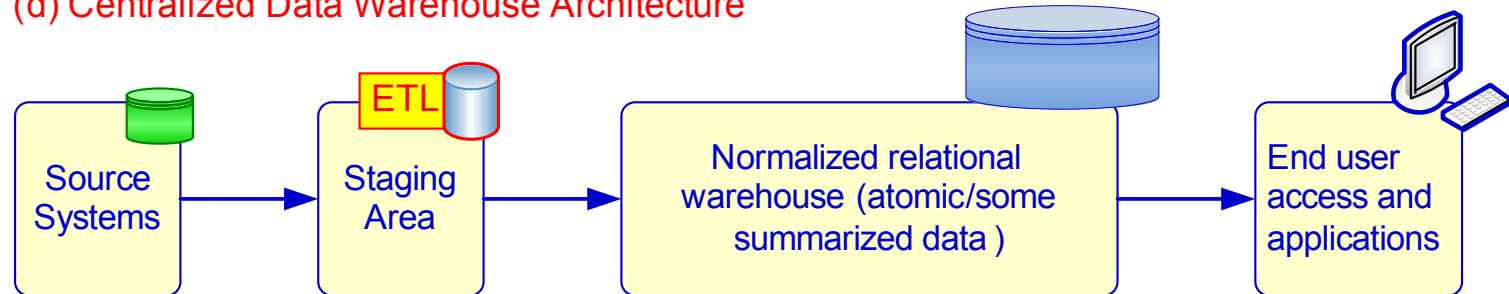
(c) Hub and Spoke Architecture (Corporate Information Factory)



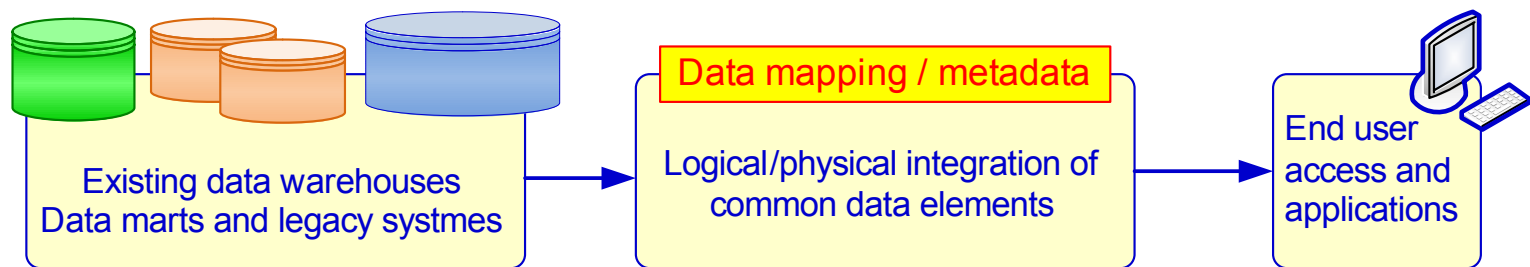
Data Warehouse

Arquiteturas Alternativas

(d) Centralized Data Warehouse Architecture

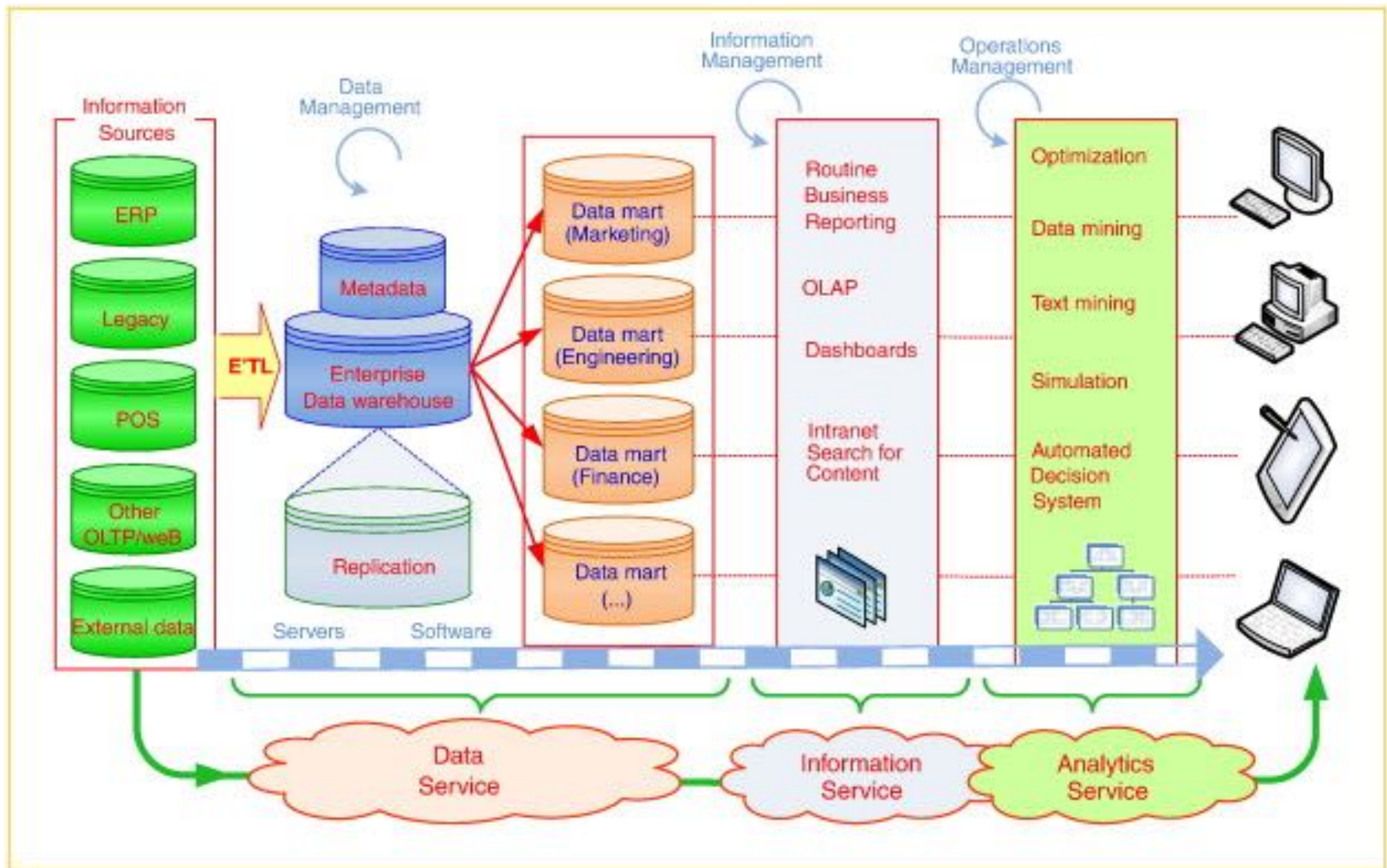



(e) Federated Architecture



- Cada arquitetura tem vantagens e desvantagens!
- Qual arquitetura é a melhor?

BI Orientado a Serviço





Dez fatores que potencialmente afetam a decisão de seleção de arquitetura DW

1. Interdependência de informação entre unidades organizacionais
2. Necessidades de informação da alta administração
3. Urgência da necessidade de um data warehouse
4. Natureza das tarefas do utilizador final
5. Restrições sobre os recursos
6. Visão estratégica do data warehouse antes da implementação
7. Compatibilidade com sistemas existentes
8. Percepção da capacidade da equipe de TI
9. Problemas técnicos
10. Fatores sociais / políticos



Data Warehouse

Outros Componentes

- *Operational data stores (ODS)*
 - Um tipo de banco de dados usado frequentemente como uma área intermediária para um data warehouse.
- *Oper marts*
 - Um data mart operacional.
- *Enterprise data warehouse (EDW)*
 - Um data warehouse corporativo.
- Metadata – “*data about data*”
 - Em DW, os metadados descrevem o conteúdo de um data warehouse e sua aquisição e uso.

BD Operacional x *Data Warehouse*

	OLTP System Online Transaction Processing (Operational System)	OLAP System Online Analytical Processing (Data Warehouse)
Source of data	Operational data; OLTPs are the original source of the data.	Consolidation data; OLAP data comes from the various OLTP Databases
Purpose of data	To control and run fundamental business tasks	To help with planning, problem solving, and decision support
What the data	Reveals a snapshot of ongoing business processes	Multi-dimensional views of various kinds of business activities
Inserts and Updates	Short and fast inserts and updates initiated by end users	Periodic long-running batch jobs refresh the data
Queries	Relatively standardized and simple queries Returning relatively few records	Often complex queries involving aggregations
Processing Speed	Typically very fast	Depends on the amount of data involved; batch data refreshes and complex queries may take many hours; query speed can be improved by creating indexes
Space Requirements	Can be relatively small if historical data is archived	Larger due to the existence of aggregation structures and history data; requires more indexes than OLTP
Database Design	Highly normalized with many tables	Typically de-normalized with fewer tables; use of star and/or snowflake schemas
Backup and Recovery	Backup religiously; operational data is critical to run the business, data loss is likely to entail significant monetary loss and legal liability	Instead of regular backups, some environments may consider simply reloading the OLTP data as a recovery method

A Evolução do Data Warehouse

- *Fontes de dados...*

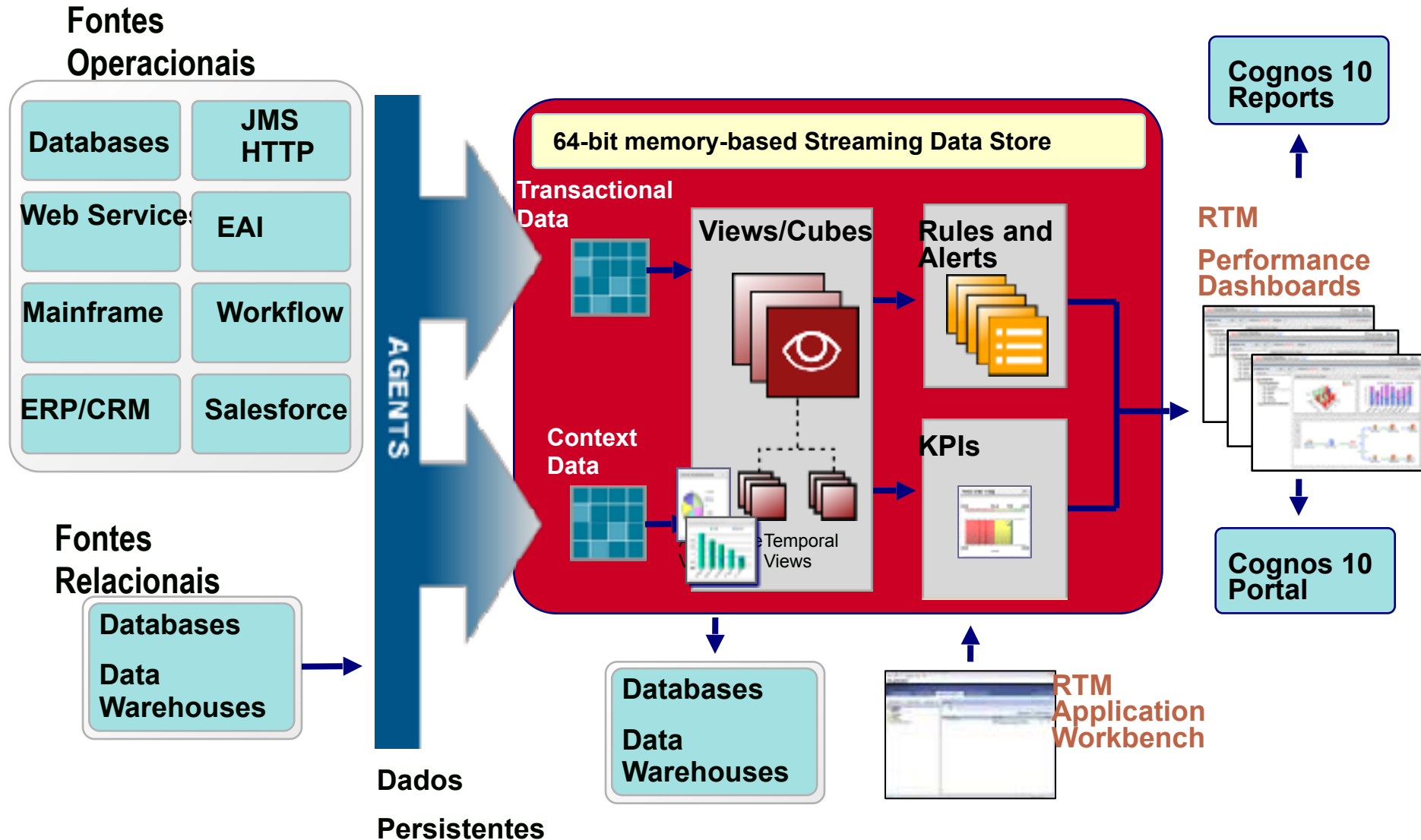
- Web, social media, and Big Data
- Open source software
- SaaS (software as a service)
- Cloud computing
- Data lakes

- *Infraestruturas...*

- Columnar
- Real-time DW
- Data warehouse appliances
- Data management practices/technologies
- In-database & In-memory processing New DBMS
- New DBMS, Advanced analytics, ...



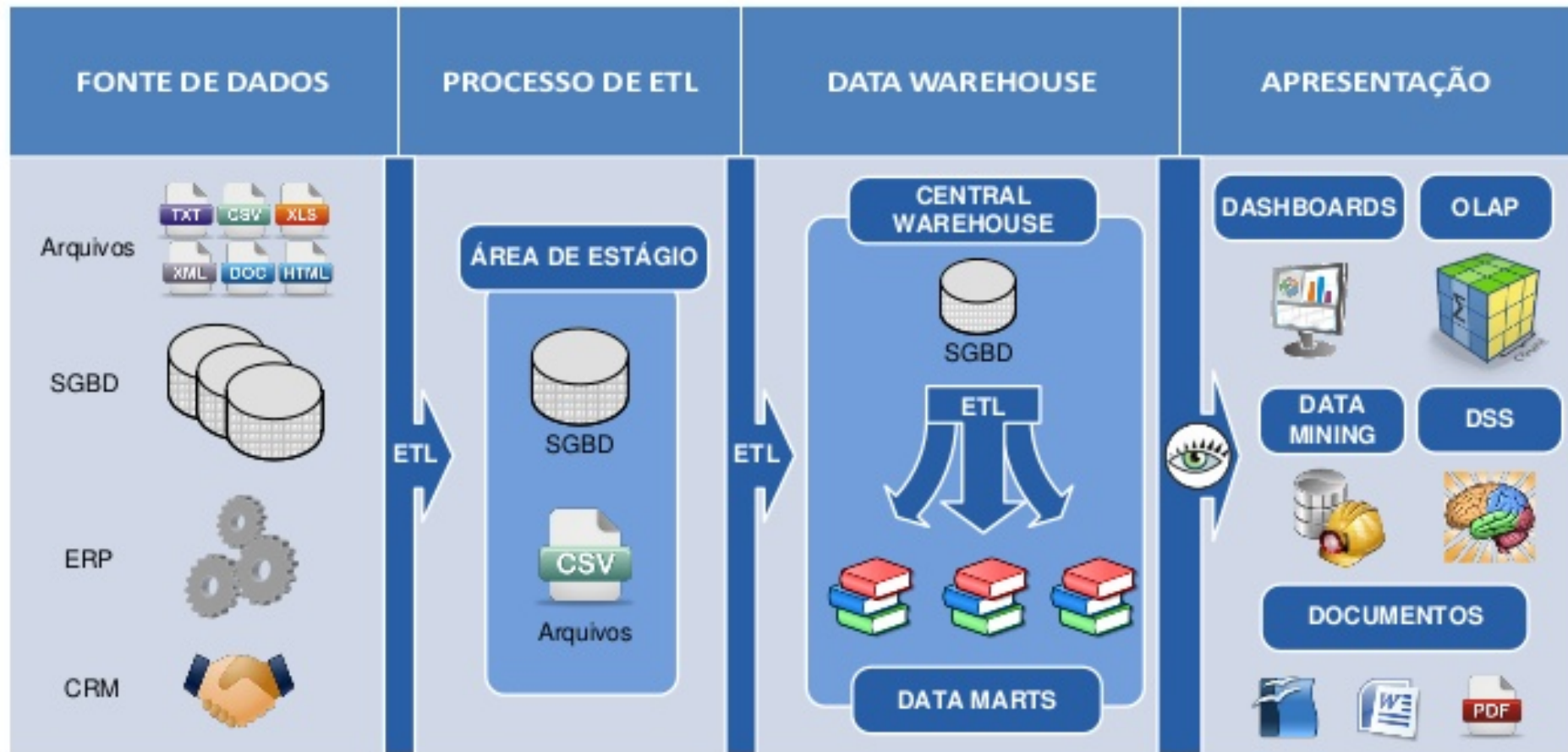
Arquitetura *Real Time* – IBM Cognos RTM (Exemplo)



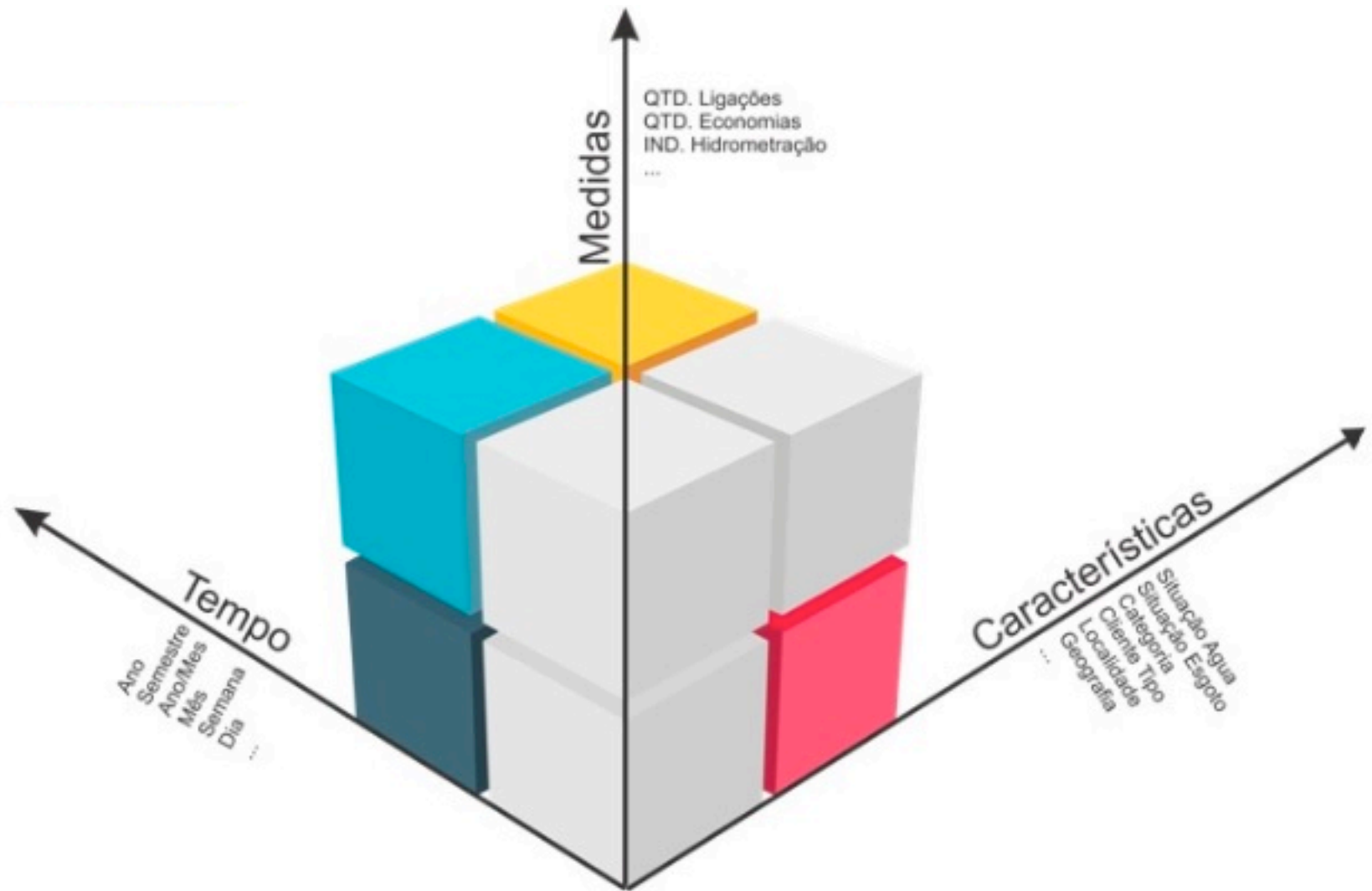
Representação dos Dados no *Data Warehouse*



Arquitetura Genérica de Dados



Modelo de Dados de um DW



Modelo de Dados de um DW

- Modelo Dimensional
 - Um modelo de dados baseado que suporta acesso de consulta de alto volume
- *Star schema*
 - Tipo de modelo dimensional mais comumente usado e o mais simples de modelagem.
 - Contém uma tabela de fato conectada a várias tabelas de dimensão
- *Snowflakes schema*
 - Uma extensão do esquema em estrela (*star schema*) onde o diagrama se assemelha a um floco de neve.



Modelo Multidimensional

- Multidimensional
 - Utiliza dimensões para armazenar e/ou visualizar os dados armazenados em um BD.
- Modelagem Multidimensional
 - Gera um modelo de dados simples de ser utilizado, principalmente por usuários que não são profissionais de informática
 - Fundamental para poder se trabalhar com ferramentas OLAP



Modelo Multidimensional

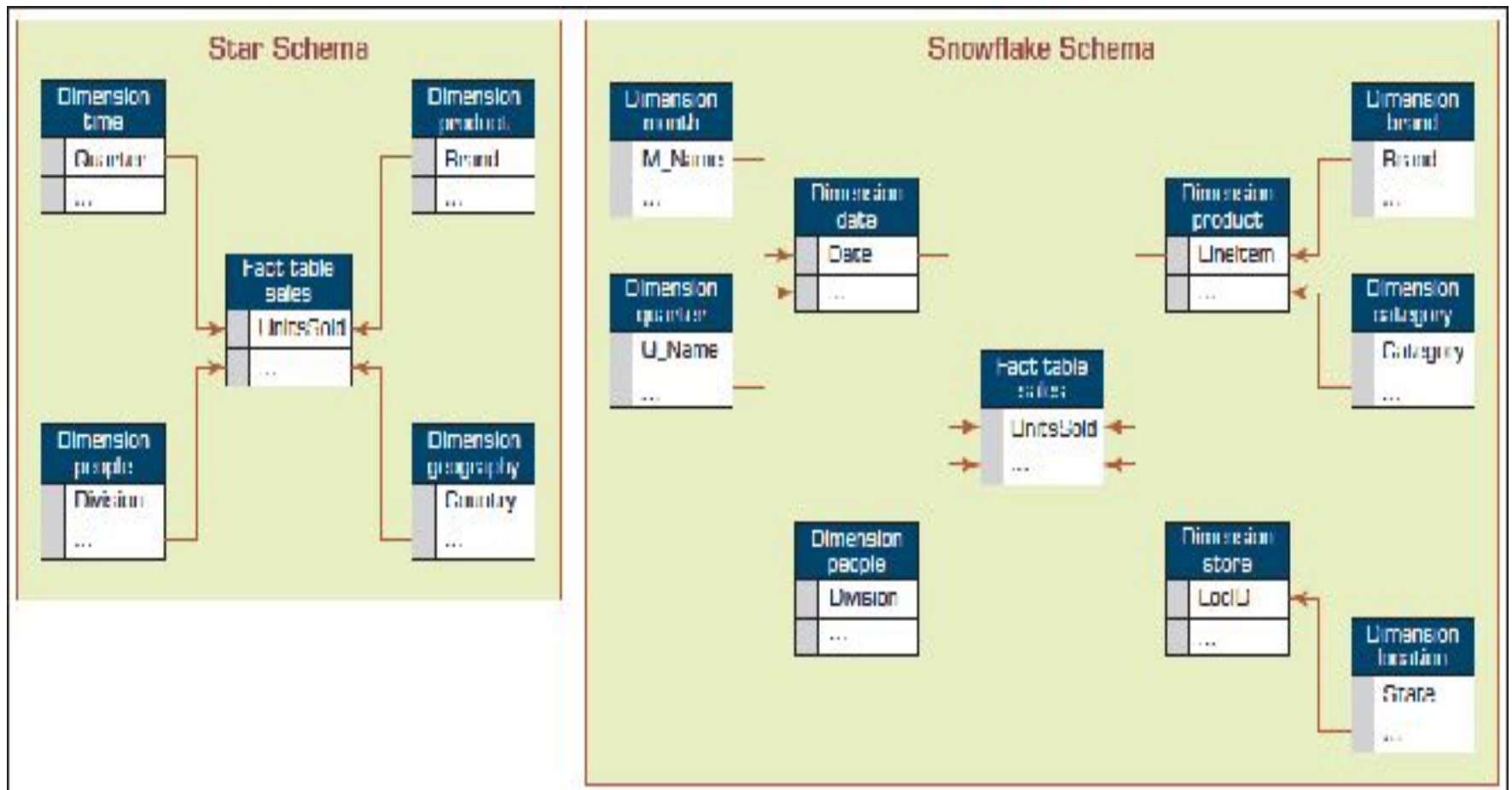
- Perspectivas
 - mês
 - modelo
 - loja
 - fabricante
- Visão histórica do volume de vendas sob múltiplas perspectivas:
 - volume de vendas por modelo
 - volume de vendas por loja
 - volume de vendas por período de tempo
 - volume de vendas por fabricante



Modelo Multidimensional

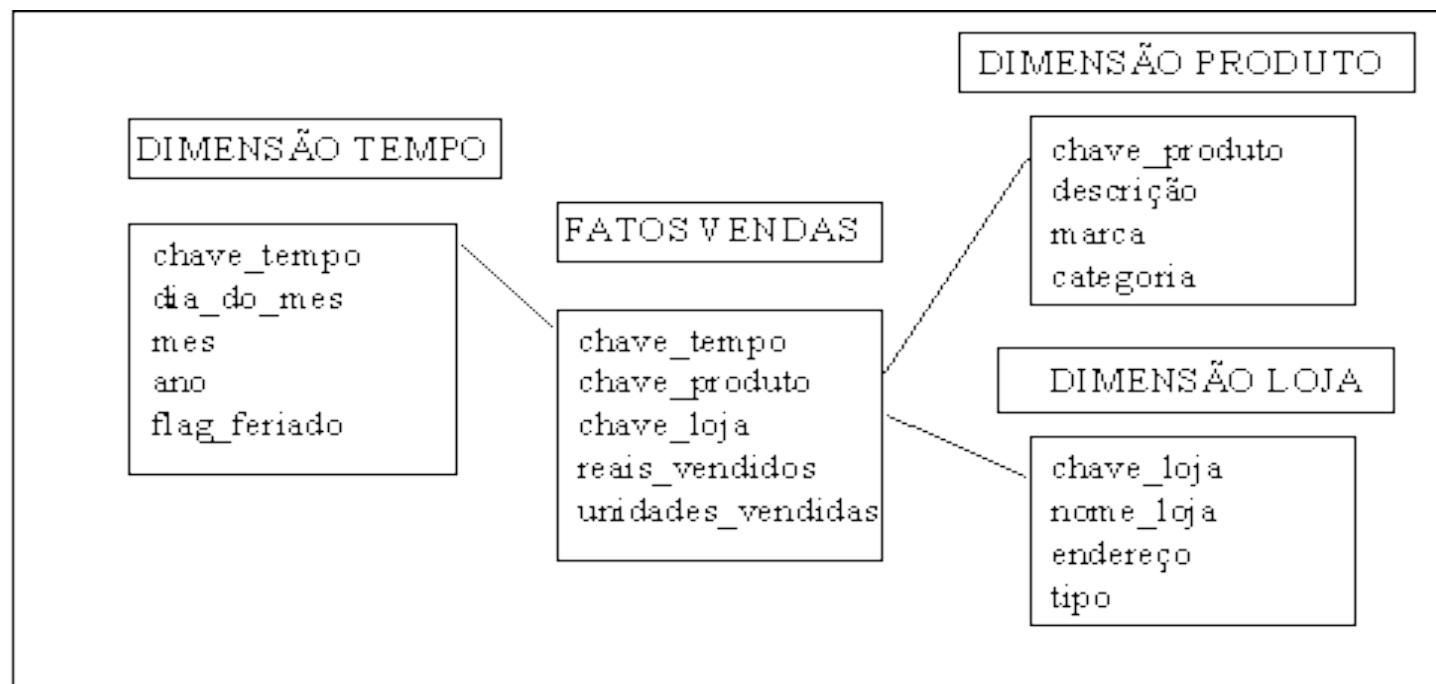
- Tabelas fato
 - Armazenam medições numéricas do negócio
 - Ex. valores, qtd. de ocorrências
- Tabelas dimensão
 - Armazenam as descrições textuais das dimensões do negócio
 - Ex. produto: identificador e nome

Modelo Multidimensional



Modelo Multidimensional

- Modelo Estrela – *Star Schema*
 - Representação de um modelo dimensional em um banco de dados relacional



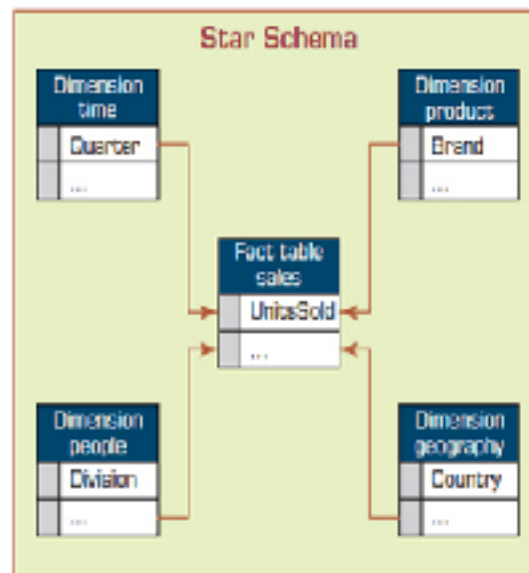


Modelo Multidimensional

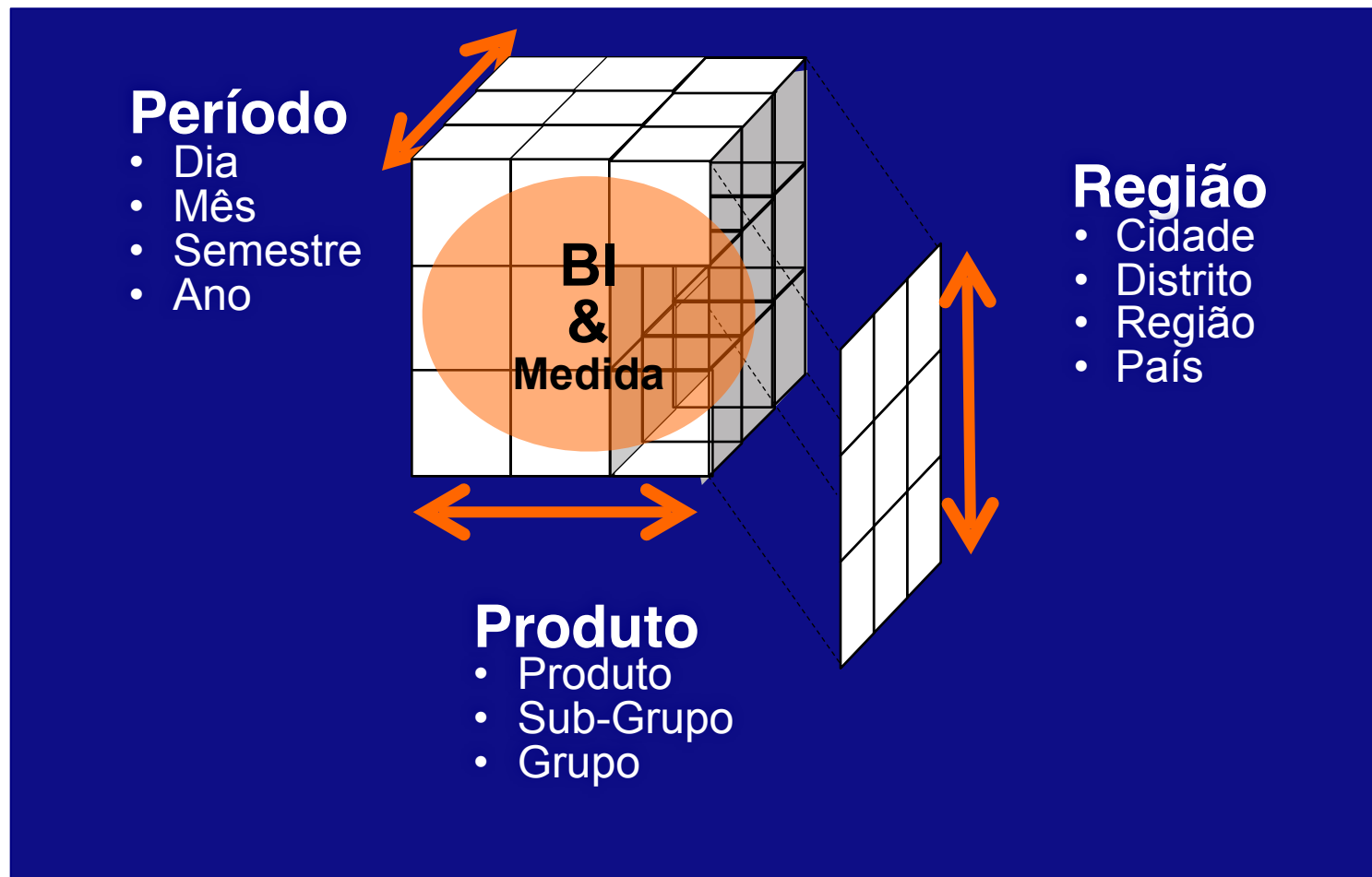
- Modelo Estrela – *Star Schema*
 - A Modelagem *Snowflake* não é recomendada, pois dificulta o entendimento do modelo dimensional por parte do usuário e resulta em decréscimo de performance porque mais tabelas precisam ser unidas para satisfazer as consultas.

Modelo Multidimensional

- Modelo Estrela – *Star Schema*
 - Oferece a capacidade de organizar, apresentar e analisar dados por várias dimensões, como vendas por região, por produto, por vendedor e por tempo (quatro dimensões)

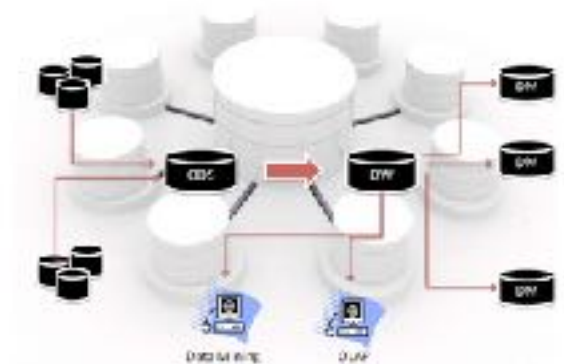


Análise Dimensional

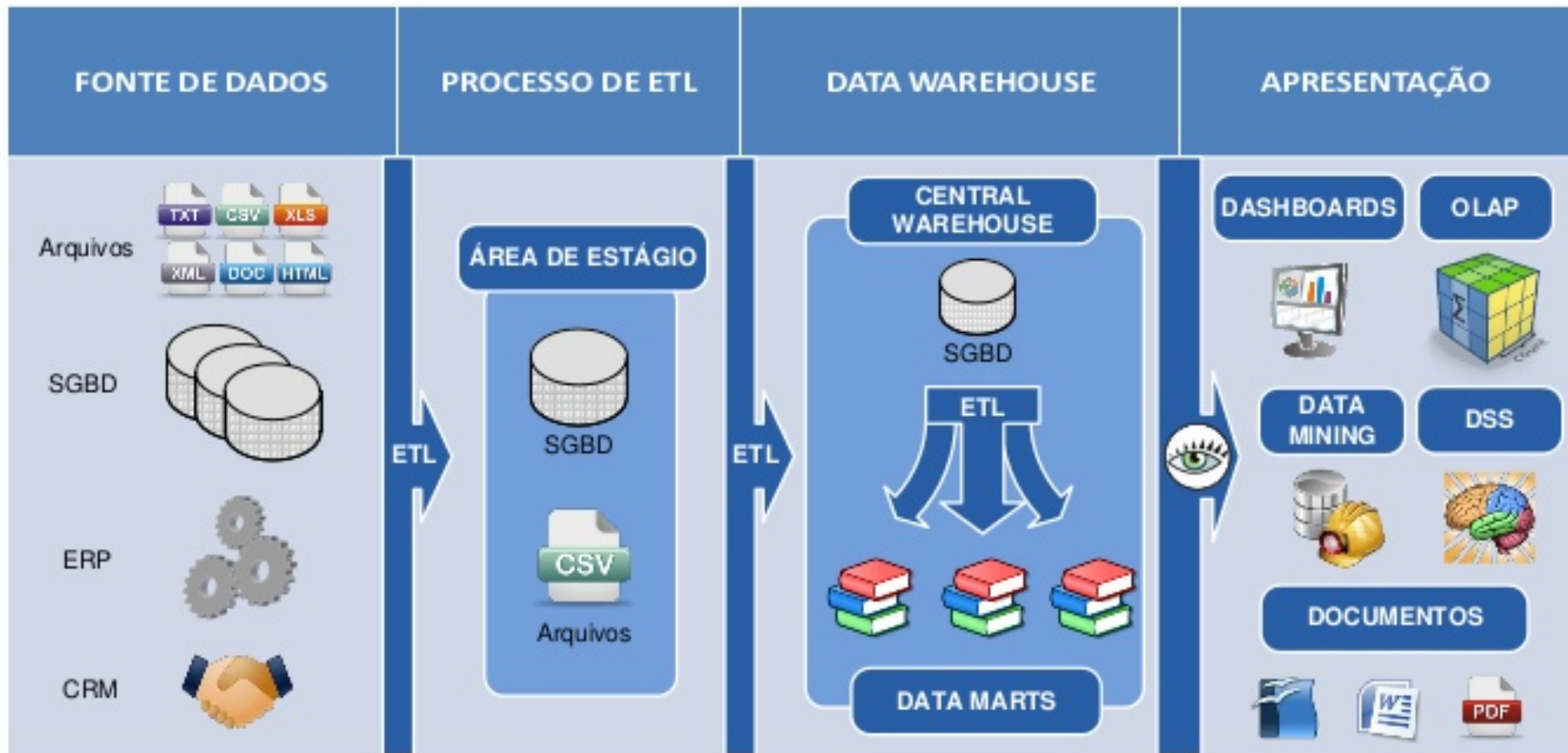


Indicadores de Desempenho (Medida)
Dimensões (Níveis de Análise)

Projeto de *Data Warehouse*



Arquitetura Genérica da Plataforma BI

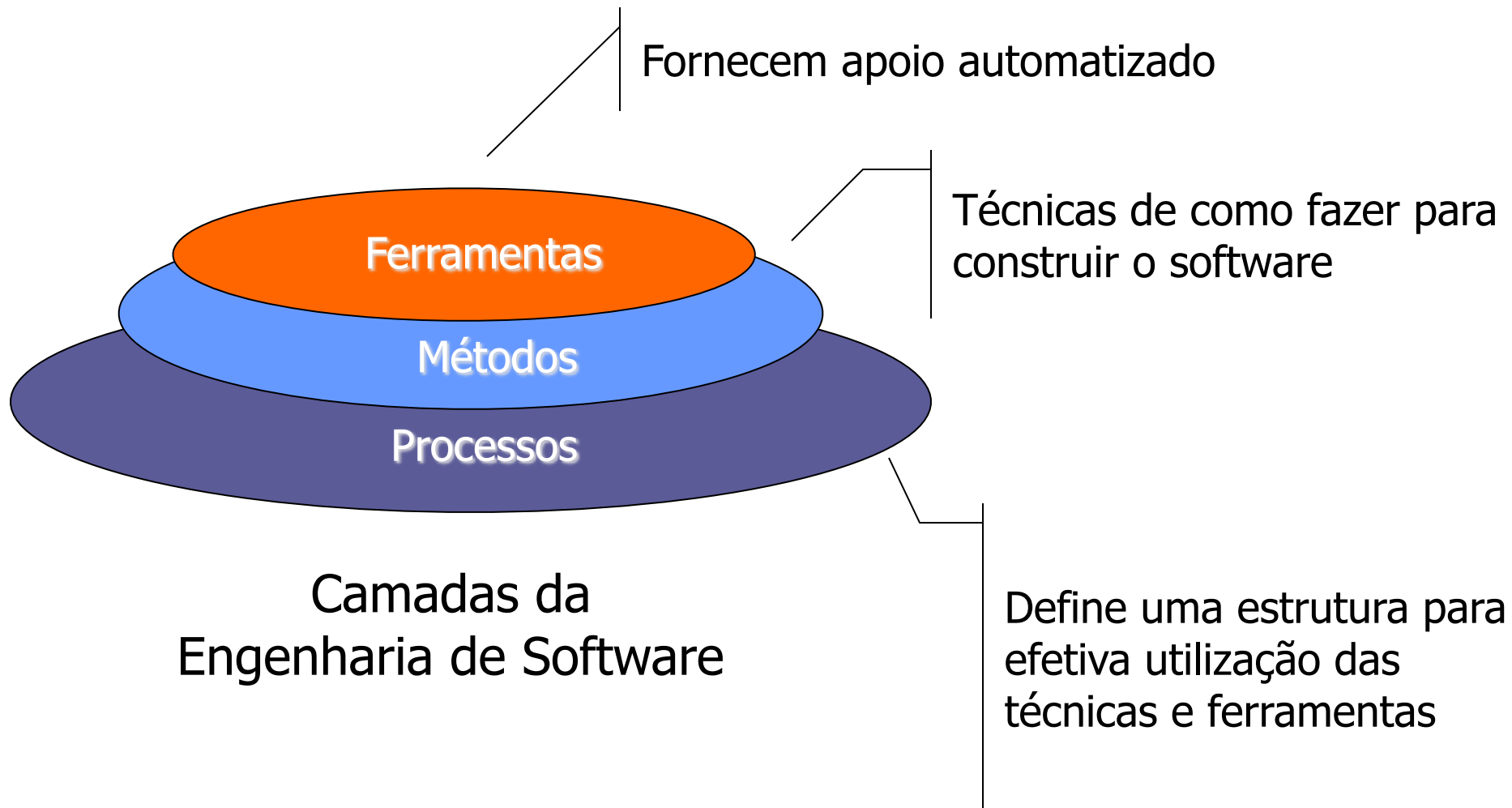




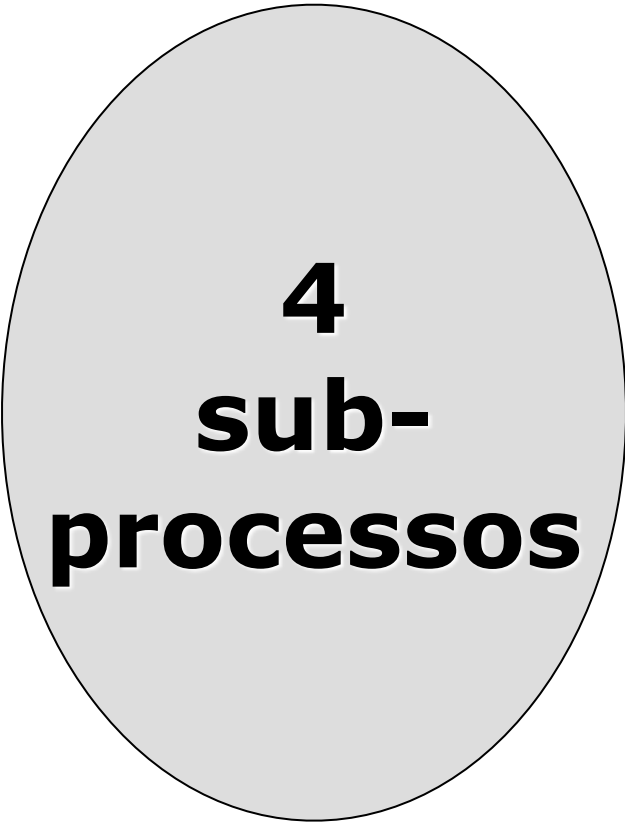
Processo de Desenvolvimento de Software

- O que é?
 - Uma série de passos, um roteiro, que se segue para criar, “a tempo” um resultado de alta qualidade
 - Resultado → Produto ou Sistema
- Por que é importante?
 - Fornece estabilidade, controle e organização

Processo de Desenvolvimento de Software



Processo de Desenvolvimento do BI



**4
sub-
processos**

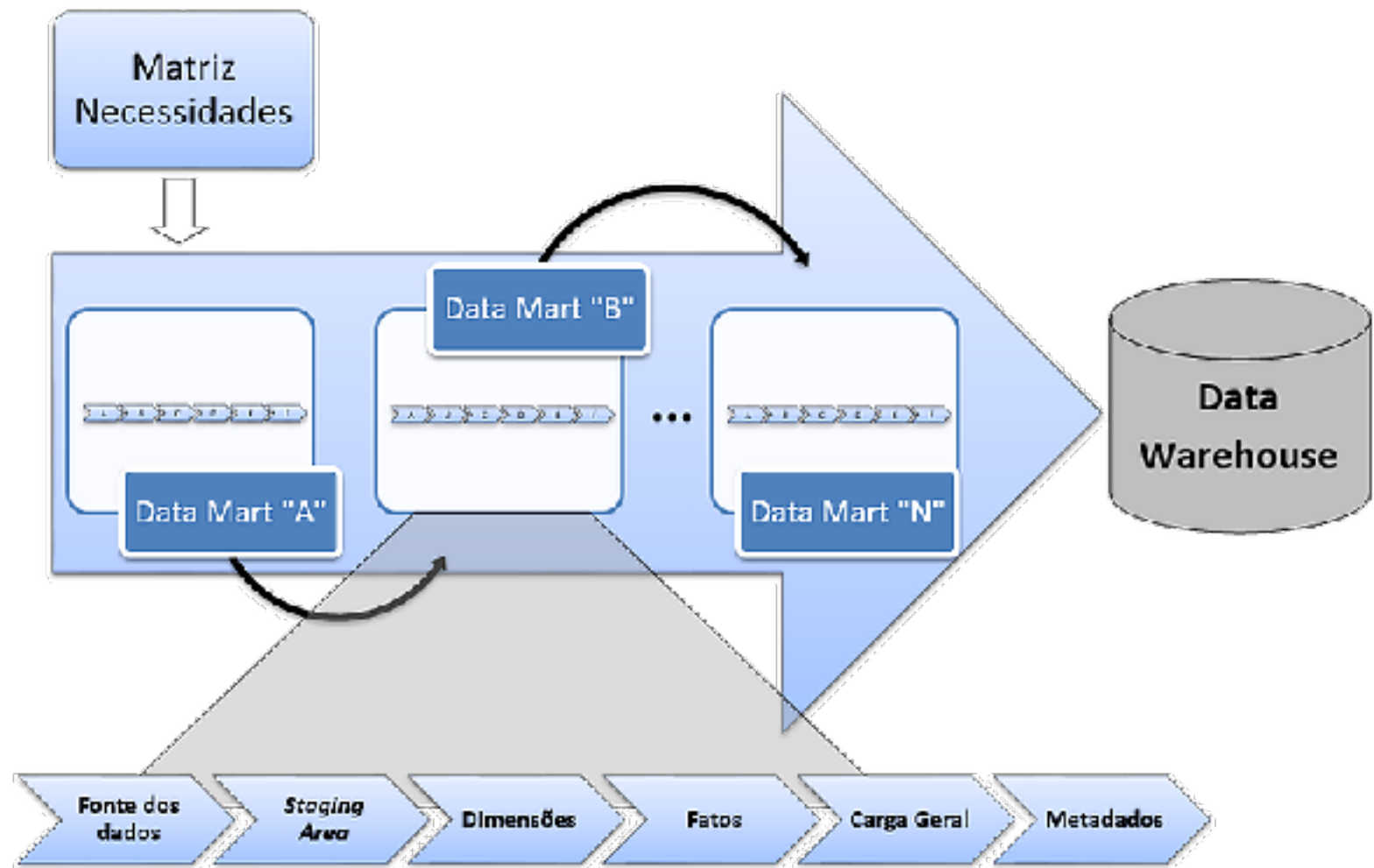
Modelagem Dimensional

Construção dos DM's / DW

Extração e Transformação
(fontes internas e externas)

Desenvolvimento das Aplicações

Processo de Construção do DW





Processo de Construção do DW

Etapas:

1. Especificação da Matriz de Necessidades
2. Elaboração do Modelo Dimensional
3. Projeto dos Bancos de Dados BI (SA/ODS/DW/DM)
4. ETL
5. Dashboards/Relatórios (aplicações/front-end)



Matriz de Necessidades

- Identificar Métricas e Dimensões
 - O objetivo desta atividade é gerar uma definição de alto nível dos indicadores que devem ser acompanhados pelo cliente, juntamente com as formas de visualização (dimensões) dos mesmos . Esse documento deve ser avaliado e aceito pelo solicitante do projeto
- Etapas:
 - Analisar Relatórios e Gráficos
 - Identificar Indicadores
 - Analisar Demandas provenientes de outras áreas
 - Definir Matriz de Indicadores e Dimensões

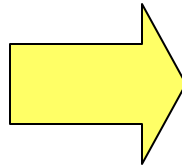
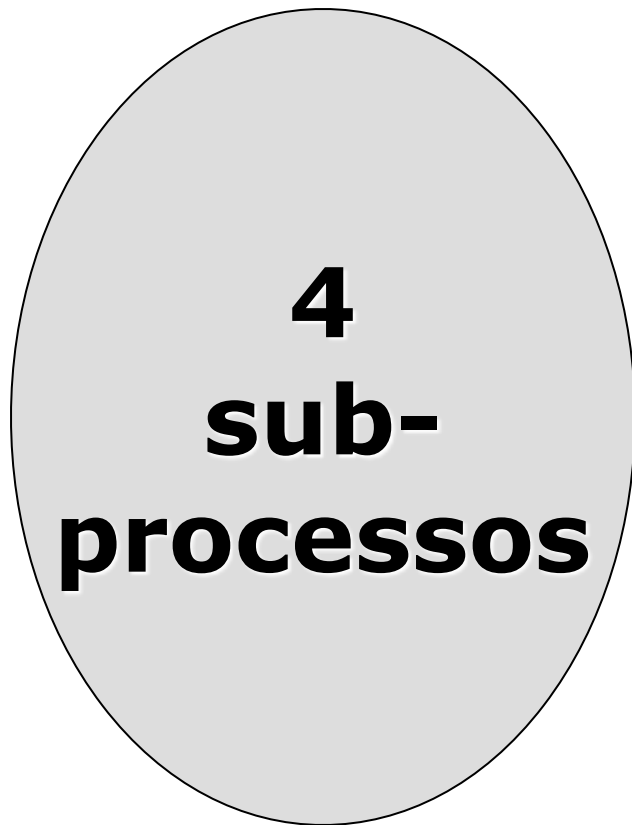
Matriz de Necessidades

Matriz Necessidade	Dimensões	Sexo	Raça	Faixa Etária	Profissão	Estado	Cidade	Bairro	Tempo	
Indicadores										
Quantidade de Portadores de Deficiência		X	X	X	X	X	X	X	X	
Quantidade de Empregados		X	X	X	X	X	X	X	X	
Total da Renda de Empregados		X	X		X	X	X	X	X	
Total da Renda de Portadores de Deficiência		X	X		X	X	X	X	X	
% Desempregados		X	X	X		X	X		X	

Matriz de Necessidade



Processo de Desenvolvimento do BI



Modelagem Dimensional

Construção dos DM's / DW

Extração e Transformação
(fontes internas e externas)

Desenvolvimento das Aplicações



Modelagem Dimensional

- É uma técnica de projeto lógico que procura apresentar dados em uma forma comum que é intuitiva e permita acesso de alto desempenho
- Técnica antiga para criar BD simples e compreensíveis
- Modelo manipulado pelas ferramentas OLAP
- Modelo para a representação de assuntos passíveis de análise
- Contém as mesmas informações do MER (entidades, atributos e relacionamentos)

Modelagem Dimensional

- Os dados são visualizados através de uma estrutura simples de “cubo de dados”
 - Visão mais simplificada do domínio
 - Facilmente extensível
 - Aumenta desempenho
 - Também chamado de “star schema”





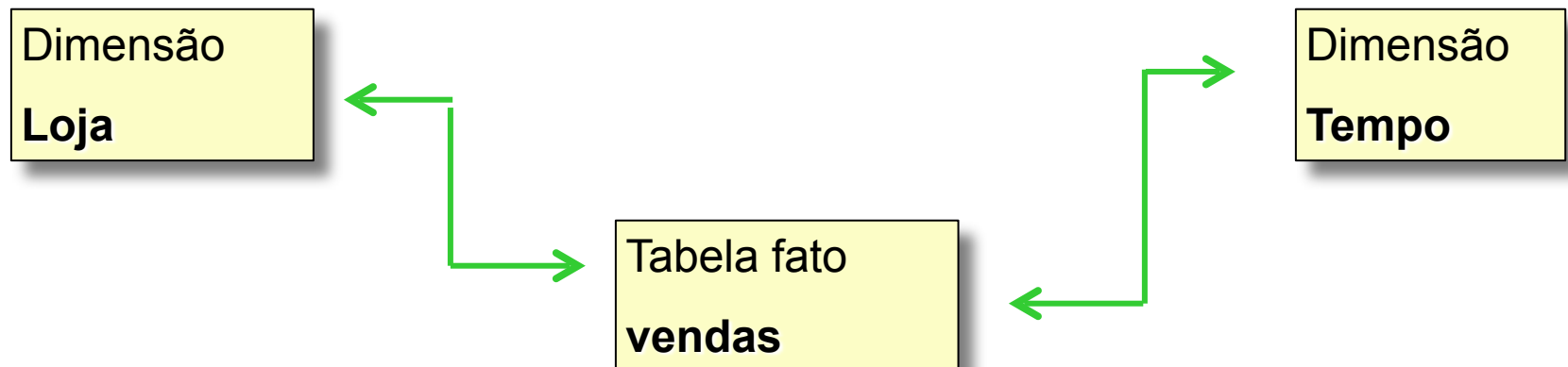
Modelagem Dimensional

- Modelo manipulado pelas ferramentas OLAP
- Modelo para a representação de assuntos passíveis de análise
- Busca apresentar os dados em uma estrutura padronizada que é intuitiva e permite alto desempenho

Modelagem Dimensional

Fatos

- São observações do negócio
- Um foco de interesse da empresa
 - Exemplos (domínio Loja):
 - Vendas, Promoções, Compras de fornecedores, ...
- Representação de um assunto
- Um assunto pode ser
 - Dado operacional, transação do negócio ou evento





Modelagem Dimensional

Dimensões

- Representação de contextos relevantes para a análise de um fato
- Granularidade adotada para representar fatos
- Dimensões exemplos para o Fato Vendas:
 - Clientes, Produtos, Tempo, Locais,
- Uma dimensão pode:
 - Conter membros
 - Ser organizada em hierarquias



Modelagem Dimensional

Medidas

- Representação de atributos (variáveis) relevantes para a análise de um fato
- São normalmente valoradas e aditivas
- Indicadores de desempenho para análise
- Praticamente todas as consultas são construídas através de adições das medições

Modelagem Dimensional

Medidas

- São três tipos:
 - Aditivas
 - Medidas que podem ser sumarizadas independente das dimensões utilizadas. Pode-se somar ao longo de qualquer dimensão
 - Semi-aditivas
 - Medidas que podem ser sumarizadas em alguns casos. Isso porque a depender da situação empregada à métrica, pode-se perder sentido para a análise caso seja agregada. Isto é, permite ser somada ao longo de algumas dimensões.
 - Não aditivas
 - Medidas que não podem ser sumarizadas ao longo das dimensões. Essas métricas não podem ter agregações, pois perdem a veracidade do valor.



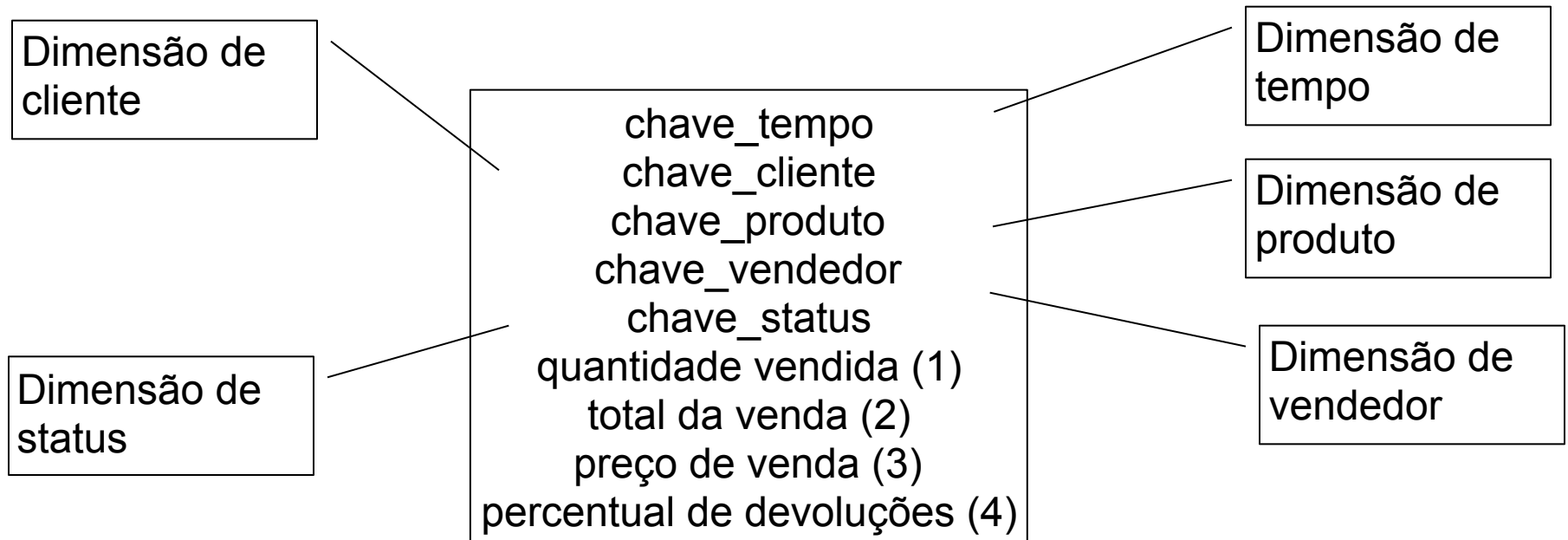
Modelagem Dimensional

Medidas

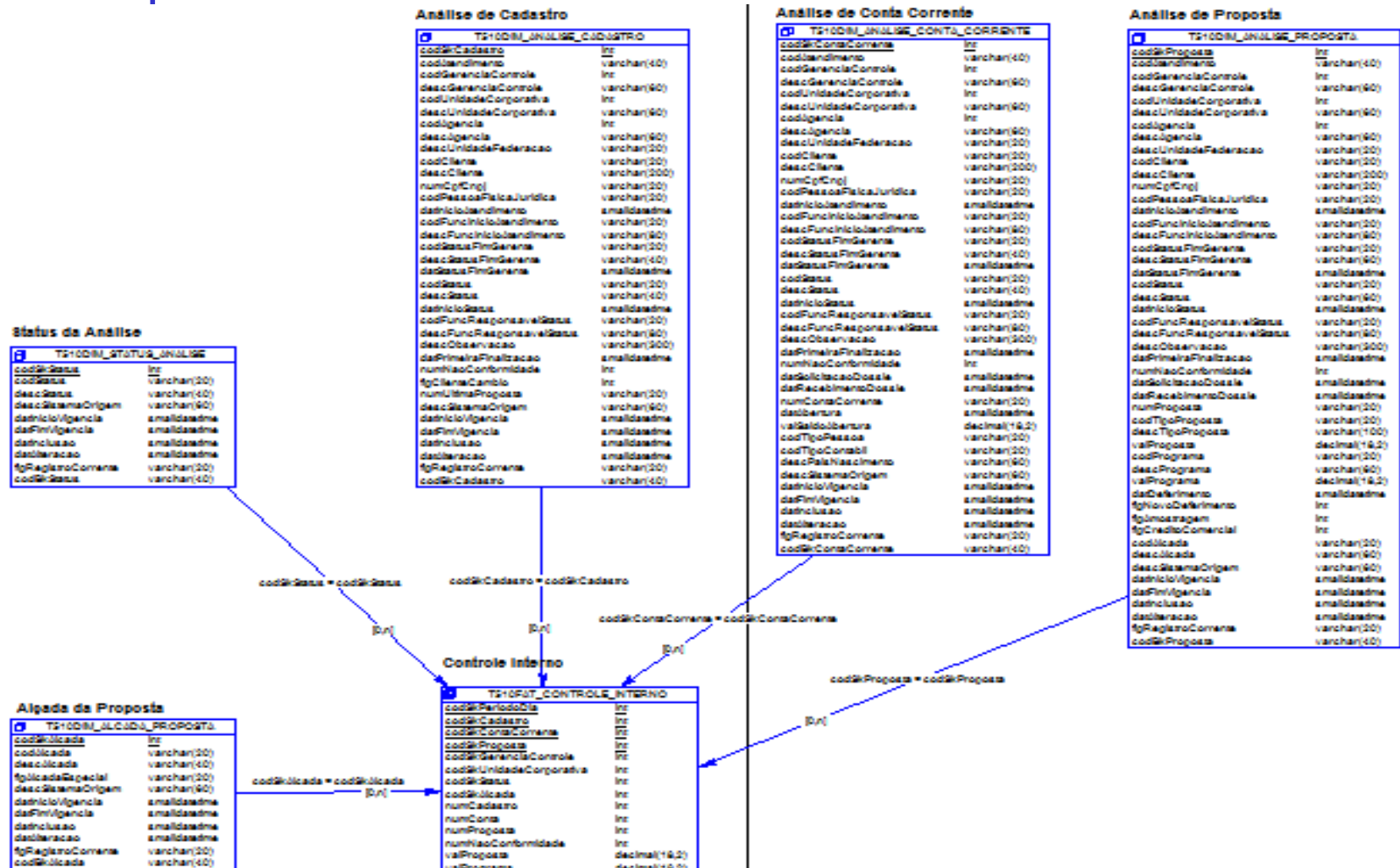
- Uma medida é determinada pela associação de dimensões
 - Produto X Tempo.ano
 - Local.estado X Cliente.classe X Tempo.mês
- Exemplo para o Fato Vendas
 - Quantidade vendida por ano e produto;
 - Valor venda por estado e classe de cliente;
 - Percentual de devoluções por mês e vendedor.

Modelagem Dimensional

Exemplo



Exemplo



Modelagem Dimensional

<div> <div>Linhas:</div> <div> <div> <div> <div></div> <div>Tipo de Pessoa</div> <div></div> </div> </div> <div>Colunas:</div> <div> <div> <div></div> <div>2010</div> <div></div> </div> </div> <div>Filtro de contexto:</div> <div> <div> <div></div> <div>Ativa</div> <div></div> </div> </div> </div> </div>			
Quantidade de Contas	2010/Jan	2010/Fev	2010
Conta Simplificada Pessoa Física	560.582	563.599	563.599
Conta Normal Crediamigo Pessoa Física	494.416	506.928	506.928
Demais Casos de Pessoas Físicas	128.244	129.464	129.464
Pessoa Jurídica - Cheque Empresa - Comércio	82.068	82.804	82.804
Beneficiário do INSS	80.309	80.629	80.629
Pessoa Jurídica para os Demais Tipos Contábeis Maiores que 04	28.200	28.665	28.665
Pessoa Física para os Tipos Contábeis 11, 20, 21, 22, 24, 29, 31, 32,	21.279	21.351	21.351
Funcionário do Banco do Nordeste	11.239	11.090	11.090
Conta Padrão Pessoa Física - Tomadores de Crédito	10.811	11.065	11.065
Funcionário Aposentado do Banco do Nordeste	6.198	6.112	6.112
Cliente que Recebe seus Vencimentos Através de Crédito em Conta	4.793	4.729	4.729
Pessoa Jurídica Optante do Simples - Cheque Empresa - Comércio	2.942	2.966	2.966
<i>Mais</i>			
Tipo de Pessoa	1.440.834	1.459.207	1.459.207

Modelagem Dimensional

Medida, Indicador e Indicador Chave de Desempenho

Qual a diferença entre essas três palavras?

Medida: Dado que isolado não tem significado relevante, qualquer formalização de eventos observados.

Ex.: 50 litros de combustível

Indicador: Medida aplicada dentro de um determinado contexto de negócio. Ajuda a entender o que está acontecendo. Responde a uma pergunta específica. A sua forma de visualização tem um significado próprio.

Ex.: Medidor de combustível de um automóvel. Responde a pergunta: Quanto de combustível ainda posso utilizar? ou, Quanto de combustível consumi até agora?

Indicador Chave de Desempenho (KPI): É um indicador que nos ajuda a entender como estamos indo em relação a uma determinada meta. Ele nos leva à uma ação. Assim como o indicador, a sua forma de apresentação é parte crítica da sua definição.

Ex.: Sinalizador de reserva de combustível de um carro.

Modelagem Dimensional





Modelagem Dimensional

Melhores Práticas

- Todas as chaves do modelo devem ser chaves sem significado (Surrogate keys):
- Não se deve usar as chaves originais de produção (estas devem aparecer com atributos de dimensões).
- Um campo inteiro de quatro bytes pode conter mais de dois bilhões de valores, o suficiente para qualquer dimensão.
- Tabelas fatos não devem conter atributos. Apenas Chaves e Métricas



Modelagem Dimensional

Melhores Práticas

- Não misture assuntos diferentes em uma mesma tabela fato
- Não misture granularidades diferentes em uma mesma tabela fato
- Na dimensão deve existir um único atributo que se relaciona com uma tabela Fato. O relacionamento entre esse atributo e a fato é sempre um para n



Modelagem Dimensional

As 10 Regras Essenciais para a Modelagem de Dados Dimensional by Ralph Kimball

1. Carregue dados detalhados para as estruturas dimensionais
2. Estructure os modelos dimensionais em torno dos processos de negócios
3. Tenha certeza de que cada tabela fato tenha uma dimensão de data associada.
4. Certifique-se que todos os fatos em uma única tabela fato estão na mesma granularidade ou nível de detalhe
5. Resolva relacionamentos muitos-para-muitos em tabelas fato



Modelagem Dimensional

As 10 Regras Essenciais para a Modelagem de Dados Dimensional by Ralph Kimball

6. Resolva os relacionamentos muitos-para-um nas tabelas de dimensões.
7. Gravar nomes de relatórios e valores de domínios de filtros em tabelas dimensão.
8. Tenha certeza de que as tabelas dimensão usam uma chave artificial (surrogate key).
9. Crie dimensões padronizadas para integrar os dados na empresa.
10. Avalie requisitos e realidade continuamente para desenvolver uma solução de DW/BI que seja aceita pelos usuários de negócios e suporte seu processo de tomada de decisões.

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



Prof. Fernando Siqueira
fernando.siqueira@uni7.edu.br