

Embarcadero Conference

Um único esforço, uma única base de código, múltiplas
plataformas, múltiplos dispositivos



Caetano Silva

Big Models, Big Data e Migração com ER/Studio XE6

Quem sou eu pra falar sobre este assunto?



Caetano de Andrade Silva

MBA em Gestão Estratégica da Tecnologia da Informação

Bacharel Física – IF USP

22 anos de experiência de TI

15 anos em Projetos de TI

Projetos relevantes em Gestão de Dados Corporativos:

- Metodologia de AD - Bradesco
- ER/Studio e Ferramentas de BI - Ipiranga
- Desenvolvimento de ferramentas de AD e DBA
- Gerenciamento de projetos de BI e DW
- Implantação de Métodos Ágeis para desenvolvimento de sistemas

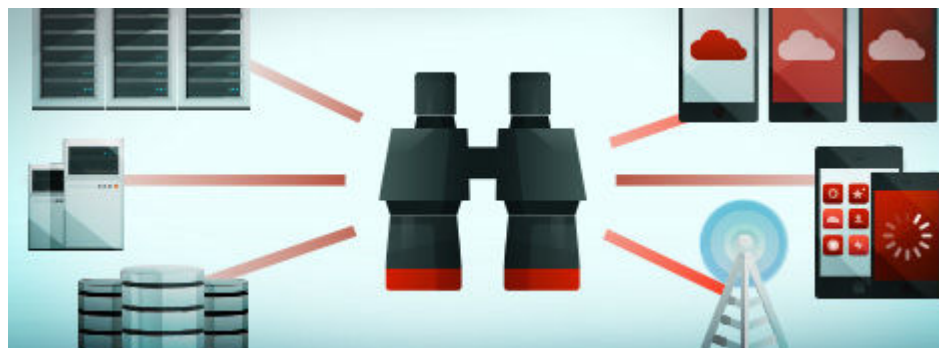
Treinamentos:

- Modelagem Relacional
- Modelagem Multi Dimensional BI e DW
- Modelagem de dados com Método Ágil
- Gestão de Dados com Método Ágil
- Modelagem com ER/Studio
- Excel PowerBI
- Data Mining com SQL Server 2014

Big Models – Gestão Corporativa de Dados

Visão da Gestão Corporativa de Dados

- Dados como ativos da Empresa
- Usando as melhores práticas como guia de trabalho
- Os processos de Gestão de Dados
- Administrador de Dados, DBA e Analistas de Sistemas na Gestão de Dados Corporativo



Foco da Gestão de Dados

“As organizações que não entenderem a enorme importância da gestão de dados e informações como ativos tangíveis na nova economia não sobreviverão.”

Tom Peters, 2001.



Resultado: Construção e compartilhamento de grandes modelos de dados corporativos.

Requisitos para Big Models

- Capacidade de fazer Engenharia Reversa de vários bancos de dados
- Capacidade de armazenar e compartilhar
- Controlar versões e o trabalho em paralelo
- Gerenciar o trabalho de equipes que atualizam e consultam os modelos de dados
- Importar grande volume de metadados de diversas fontes





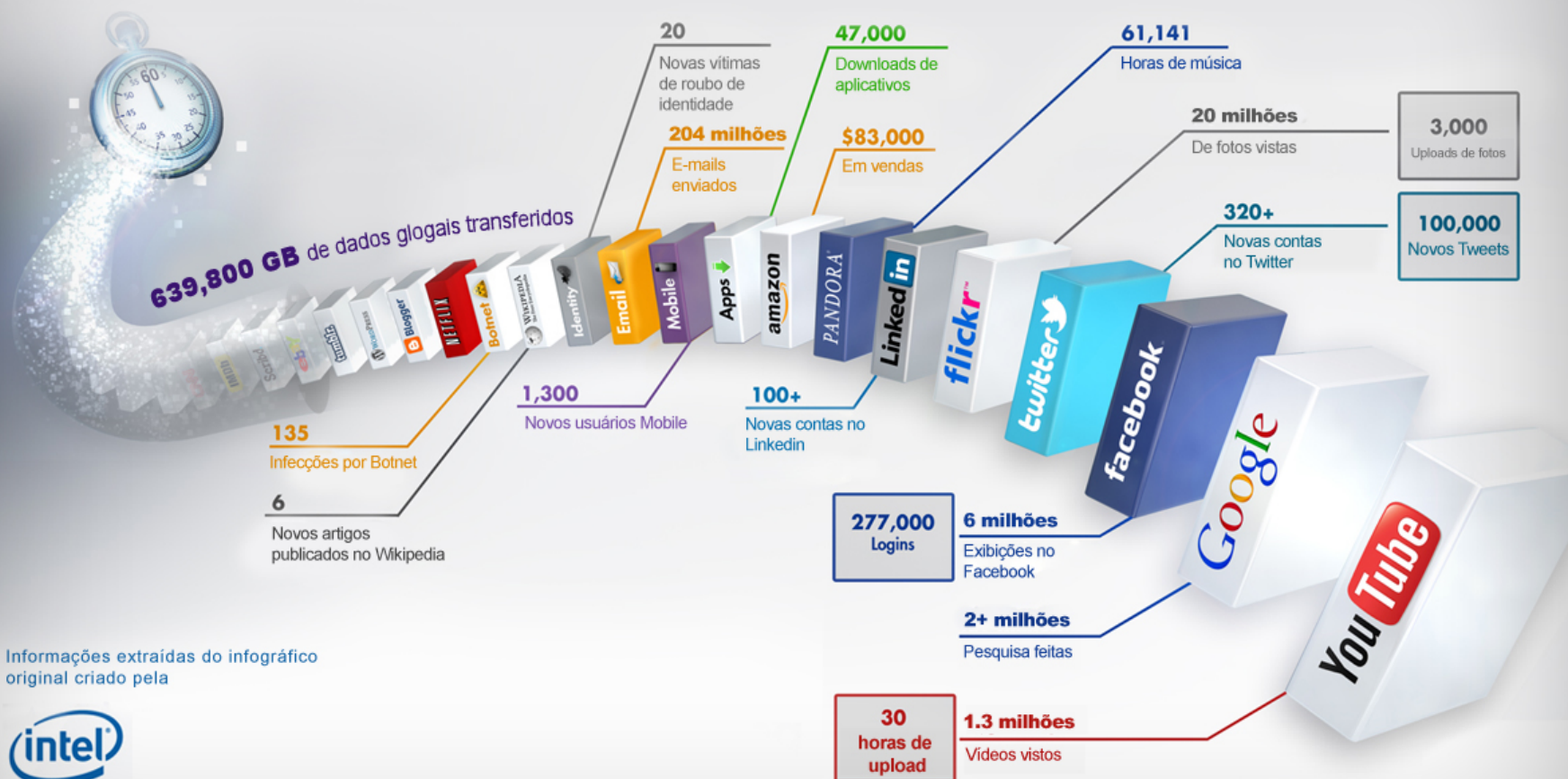
ER/Studio XE6 = Canivete

- Extended Memory Model – agora 64 bits
- Melhoria de performance nas operações do Repositório
- Instalação completa em S.O. 64 bits
- Na versão Enterprise Team
 - Registro de Data Source centralizado
 - Acesso ilimitado para o portal de metadados e modelos
- Na versão Data Architect Professional
 - Inclui o Repositório
 - Inclui recursos de importação de metadados (26 ferramentas)



Big Data – “Dados e dados por todos os lados e nenhum conhecimento para beber.”

O que acontece em um **minuto** na internet?



Informações extraídas do infográfico original criado pela



Tradução feita pela **Agência Social Tag - Mídias Sociais**

Big Data – Da onde vem? Pra onde me leva?

Big Data se baseia em 5 Vs :
Velocidade, Volume, Variedade, Veracidade e Valor.

Hadoop - Apache Foundation

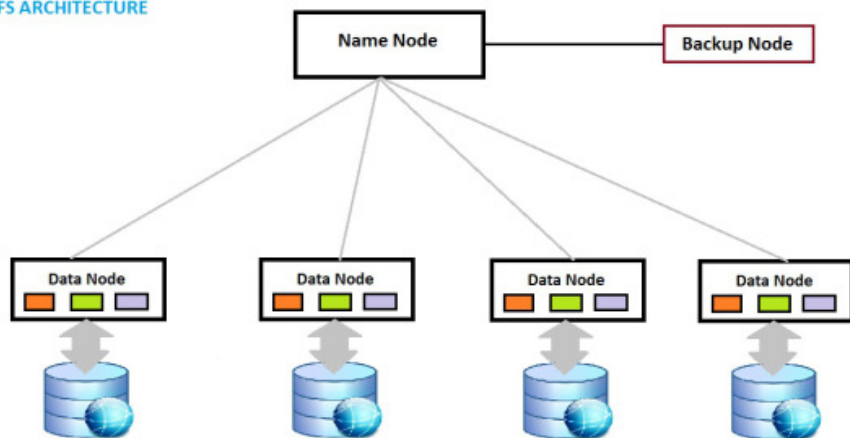
Usa conceitos de SW e HW distribuídos para aumentar a capacidade de armazenamento, velocidade de leitura de dados e processamento.

Hive – Camada de definição das estrutura de dados e de consulta a dados com linguagem SQL.

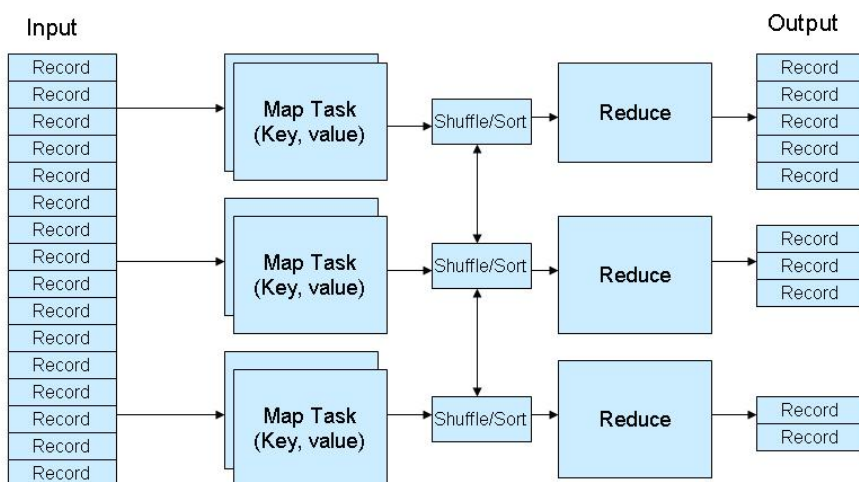


Hadoop = HDFS + MapReduce

HDFS ARCHITECTURE



HDFS - Aumenta a performance de acesso a grandes volumes de dados dividindo os arquivos em blocos de 64Mb, replicando em 3 nós cada bloco.



Função *Mapper* é feita para filtrar e transformar os dados que serão agregados pela função *Reducer*.

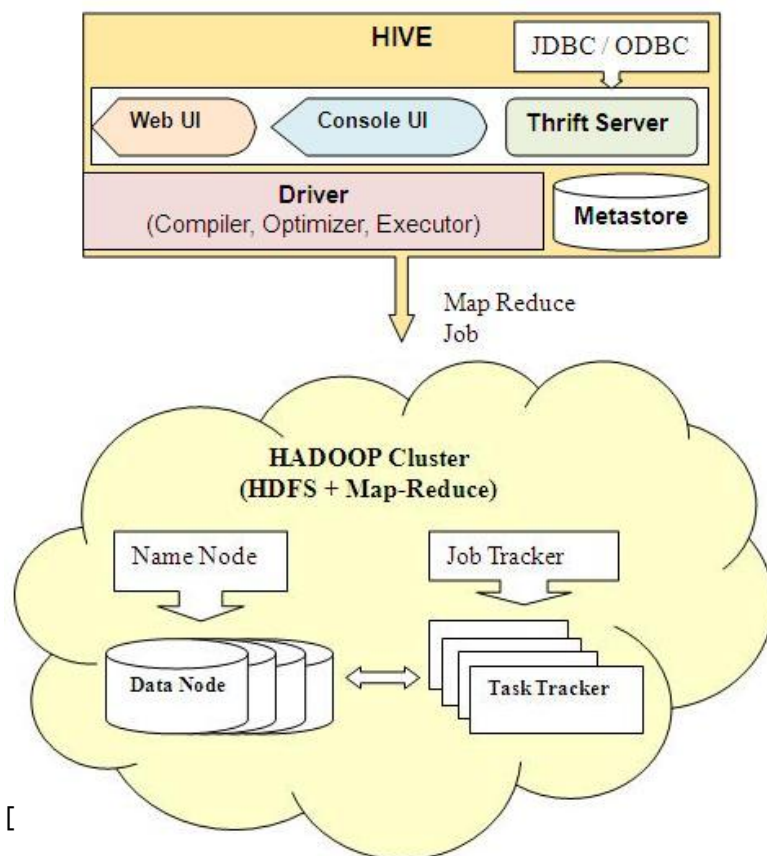
Hive = Big Data acessível

DML – Data Manipulation Language Select, Insert, Update e Delete

```
SELECT [ALL | DISTINCT] select_expr, select_expr, ...
FROM table_reference
[WHERE where_condition]
[GROUP BY col_list]
[CLUSTER BY col_list
 | [DISTRIBUTE BY col_list] [SORT BY col_list]
]
[LIMIT number]
```

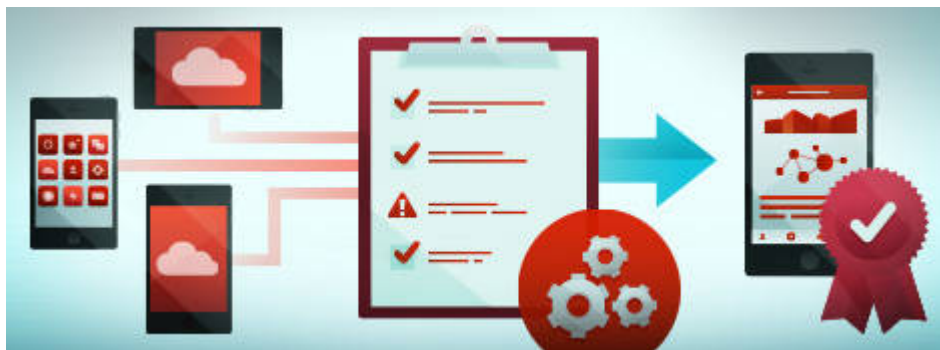
DDL – Data Definition Language Create, Drop, Truncate e Alter

```
CREATE [TEMPORARY] [EXTERNAL] TABLE [IF NOT EXISTS]
[db_name.]table_name
[(col_name data_type [COMMENT col_comment], ...)]
[COMMENT table_comment]
[PARTITIONED BY (col_name data_type [COMMENT col_comment], ...)]
[CLUSTERED BY (col_name, col_name, ...) [SORTED BY (col_name [ASC|
DESC], ...)] INTO num_buckets BUCKETS]
[SKEWED BY (col_name, col_name, ...) ON ((col_value,
col_value, ...), ...|col_value, col_value, ...)]
[STORED AS DIRECTORIES] -- [
[ROW FORMAT row_format] [STORED AS file_format]
| STORED BY 'storage.handler.class.name' [WITH SERDEPROPERTIES
(...)]
]
[LOCATION hdfs_path]
[TBLPROPERTIES (property_name=property_value, ...)]
[AS select_statement];
```



ER/Studio XE6 = Descrever e Entender Big Data

- Suporte ao Hadoop Hive
 - Engenharia reversa do Hive
 - Criação de novas estruturas de dados no Hive (Hadoop)
- Suporte ao MongoDB (noSQL)
 - Engenharia reversa da estrutura hierárquica
 - Descrição das estruturas de dados de Aplicativos Mobile e Serviços Web



Migração e Integração de dados e metadados

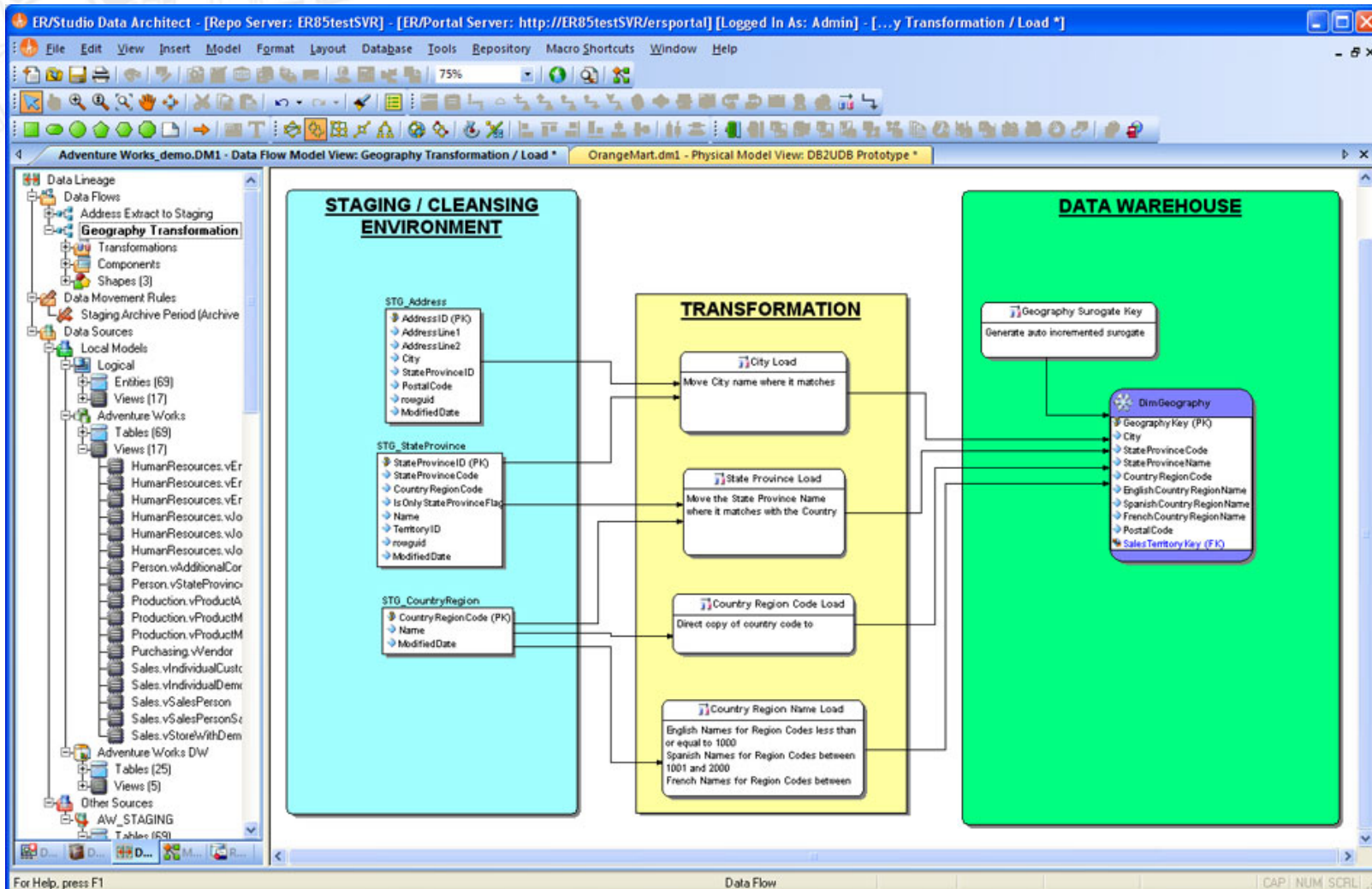
Ambientes diversificados

- Sistemas diversos (ERP, CRM, Legados, Mainframe, Web, Mobile, DW, BI)
- Bancos de dados (muitos fornecedores)
- Ferramentas de ETL (vixe!)
- Ferramentas de BI (nó!)

Projetos recorrentes de integrações de dados entre sistemas e migrações de dados entre versões.



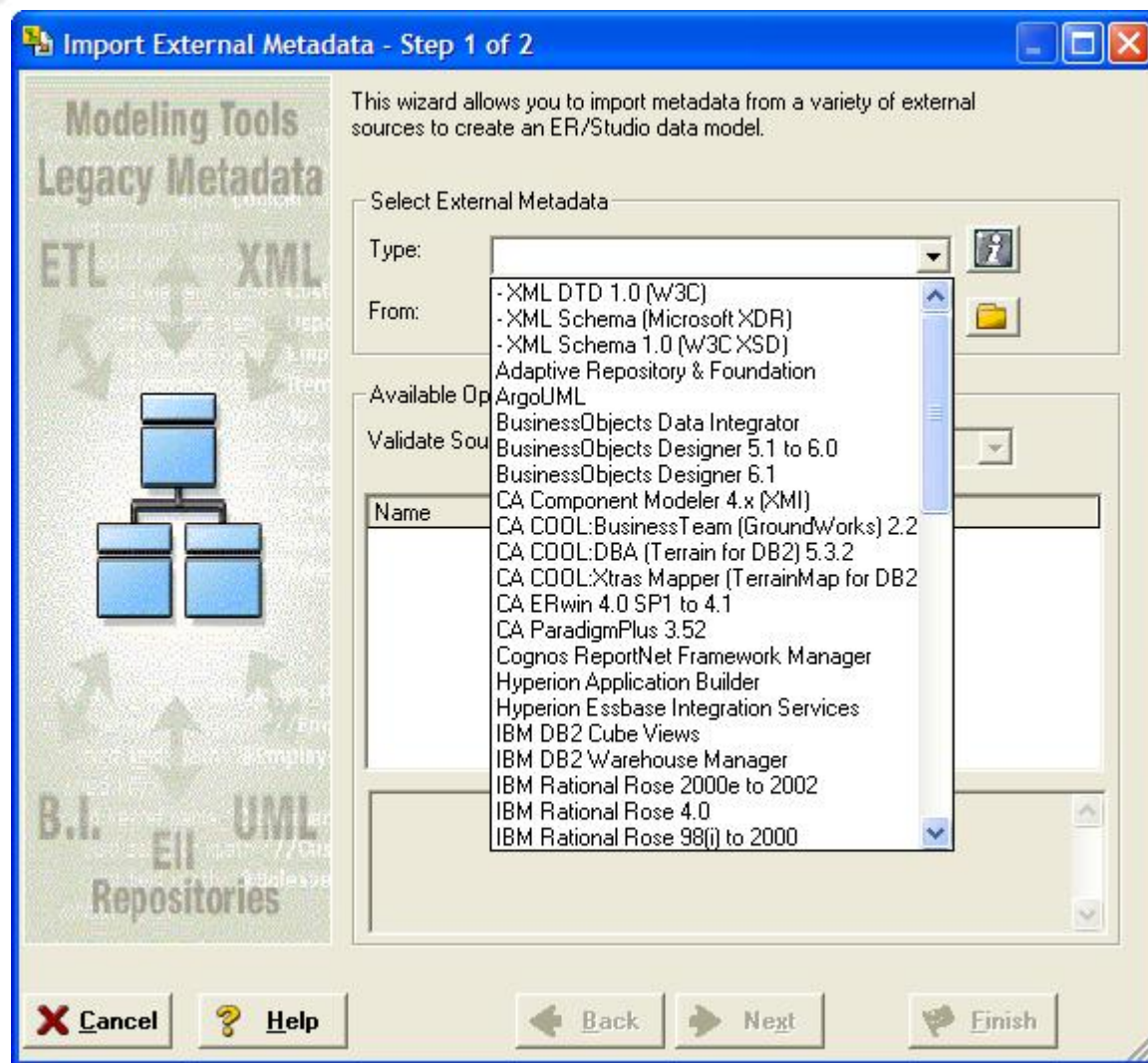
Data Lineage = Descrever o fluxo de dados



Migração e Integração de Metadados

MetaWizard

- Importação
- Exportação
- CASE Tools
- ETL
- Repositórios
- UML
- DW/BI



Obrigado! 

caetano@cstreina.com.br 
www.cstreina.com.br