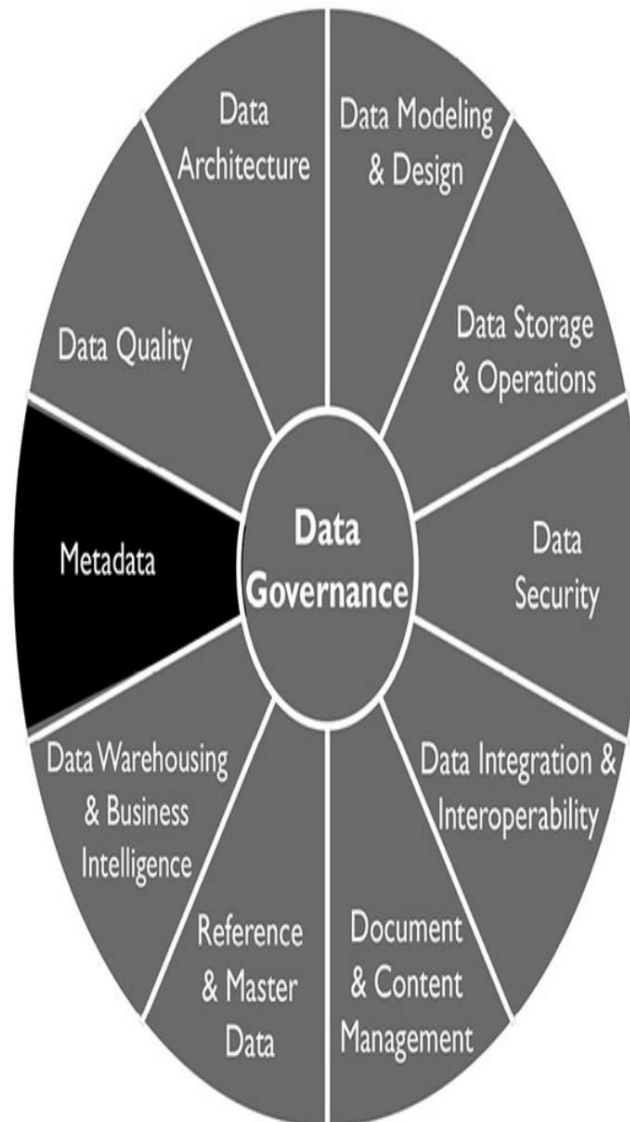


## CAPÍTULO 12

### Gerenciamento de Metadados

---



### DAMA-DMBOK2 Data Management Framework

Copyright © 2017 by DAMA International

#### 1. Introdução

---

A definição mais comum de *Metadados*, “dados sobre dados”, é enganosamente simples. O tipo de informação que pode ser classificada como Metadados é abrangente. Metadados incluem informações sobre processos técnicos e comerciais, regras e restrições de dados e estruturas de dados lógicas e físicas. Eles descrevem os dados em si (por exemplo, bancos de dados, elementos de dados, modelos de dados), os conceitos que os dados representam (por exemplo, processos comerciais, sistemas de aplicativos, código de software, infraestrutura de tecnologia) e as conexões (relacionamentos) entre os dados e os conceitos. Metadados ajudam uma organização a entender seus dados, seus sistemas e seus fluxos de trabalho. Eles permitem a avaliação da qualidade dos dados e são essenciais para o gerenciamento de bancos de dados e outros aplicativos. Eles contribuem para a capacidade de processar, manter, integrar, proteger, auditar e governar outros dados.

Para entender o papel vital dos Metadados no gerenciamento de dados, imagine uma grande biblioteca, com centenas de milhares de livros e revistas, mas sem um catálogo de fichas. Sem um catálogo de fichas, os leitores podem nem saber como começar a procurar um livro específico ou mesmo um tópico específico. O catálogo de fichas não apenas fornece as informações necessárias (quais livros e materiais a biblioteca possui e onde eles estão nas prateleiras), mas também permite que os usuários encontrem materiais usando diferentes pontos de partida (área de assunto, autor ou título). Sem o catálogo, encontrar um livro específico seria difícil, se não impossível. Uma organização sem Metadados é como uma biblioteca sem um catálogo de fichas.

Os metadados são essenciais para o gerenciamento e uso de dados (veja diversas referências a metadados no DAMA-DMBOK).

Todas as grandes organizações produzem e usam muitos dados. Em uma organização, diferentes indivíduos terão diferentes níveis de conhecimento de dados, mas nenhum indivíduo saberá tudo sobre os dados.

Essas informações devem ser documentadas ou a organização corre o risco de perder conhecimento valioso sobre si mesma. Metadados fornecem os principais meios de capturar e gerenciar conhecimento organizacional sobre dados.

No entanto, o gerenciamento de metadados não é apenas um desafio de gerenciamento de conhecimento; é também uma necessidade de gerenciamento de riscos.

Os metadados são necessários para garantir que uma organização possa identificar dados privados

ou dados confidenciais e que pode gerenciar o ciclo de vida dos dados para seu próprio benefício e para atender aos requisitos de conformidade e minimizar a exposição ao risco.

Sem Metadados confiáveis, uma organização não sabe quais dados possui, o que os dados representam, de onde se originam, como se movem pelos sistemas, quem tem acesso a eles ou o que significa para os dados serem de alta qualidade. Sem Metadados, uma organização não pode gerenciar seus dados como um ativo. De fato, sem Metadados, uma organização pode não ser capaz de gerenciar seus dados de forma alguma.

Conforme a tecnologia evoluiu, a velocidade com que os dados são gerados também aumentou. Metadados técnicos se tornaram parte integrante da maneira como os dados são movidos e integrados. O Padrão de Registro de Metadados da ISO, ISO/IEC 11179, tem como objetivo permitir a troca de dados orientada por Metadados em um ambiente heterogêneo, com base em definições exatas de dados. Metadados presentes em XML e outros formatos permitem o uso dos dados. Outros tipos de marcação de Metadados permitem que os dados sejam trocados, mantendo significantes de propriedade, requisitos de segurança, etc. (Consulte [o Capítulo 8.](#))

---

Assim como outros dados, os Metadados exigem gerenciamento. À medida que a capacidade das organizações de coletar e armazenar dados aumenta, o papel dos Metadados no gerenciamento de dados cresce em importância. Para ser orientada por dados, uma organização deve ser orientada por Metadados.

## Figura 84 Diagrama de Contexto: Metadados

## Metadata Management

**Definition:** Planning, Implementation, and control activities to enable access to high quality, integrated metadata

**Goals:**

1. Provide organizational understanding of business terms and usage.
2. Collect and integrate metadata from diverse sources.
3. Provide a standard way to access metadata.
4. Ensure metadata quality and security.

Business  
Drivers

**Inputs:**

- Business Requirements
- Metadata Issues
- Data Architecture
- Business Metadata
- Technical Metadata
- Process Metadata
- Operational Metadata
- Data Governance Metadata

**Activities:**

1. **Define Metadata Strategy (P)**
2. **Understand Metadata Requirements (P)**
  1. Business User Requirements
  2. Technical User Requirements
3. **Define Metadata Architecture (P)**
  1. Create MetaModel (D)
  2. Apply Metadata Standards (C)
  3. Manage Metadata Stores (C)
4. **Create and Maintain Metadata (O)**
  1. Integrate Metadata (O)
  2. Distribute and Deliver Metadata (O)
5. **Query, Report and Analyze Metadata**

**Deliverables:**

- Metadata Strategy
- Metadata Standards
- Metadata Architecture
- MetaModel
- Unified Metadata
- Metadata Stores
- Data Lineage
- Impact Analysis
- Dependency Analysis
- Metadata Control Process

**Suppliers:**

- Business Data Stewards
- Data Managers
- Data Governance Bodies
- Data Modelers
- Database

**Participants:**

- Data Stewards
- Project Managers
- Data Architects
- Business Analysts
- System Analysts

**Consumers:**

- Application Developers
- Analyst
- Data Integrators
- Business Users
- Knowledge Workers
- Customers & Collaborators
- Data Scientists

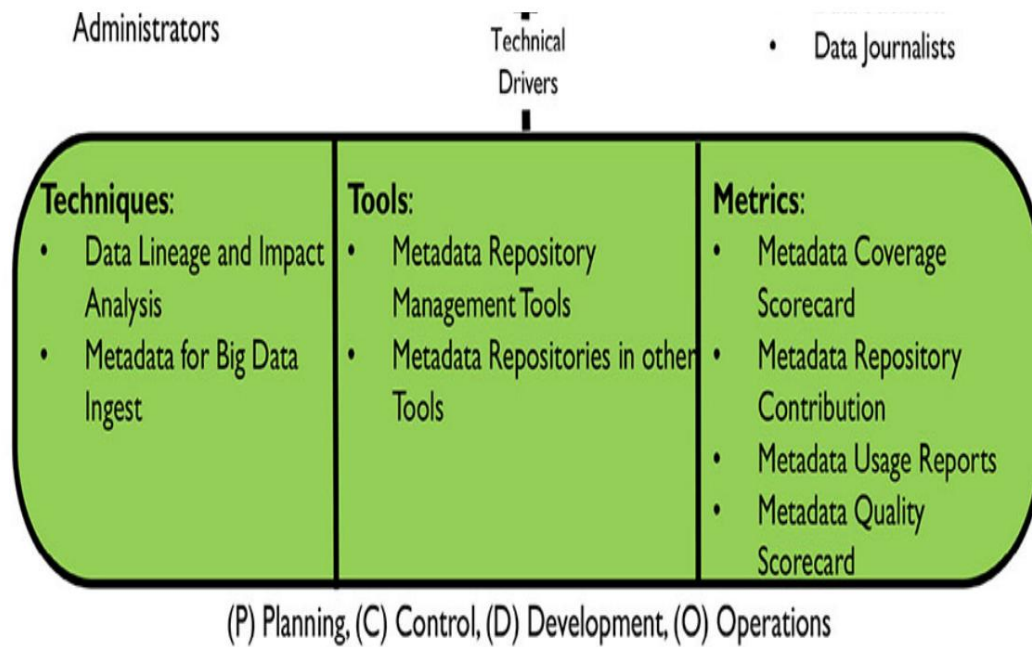


Figure 84 Context Diagram: Metadata

## 1.1 Drivers de negócios

Dados não podem ser gerenciados sem Metadados. Além disso, os próprios Metadados devem ser gerenciados. Metadados confiáveis e bem gerenciados ajudam a:

- Aumentar a confiança nos dados fornecendo contexto e permitindo a medição da qualidade dos dados Aumentar o valor das
- informações estratégicas (por exemplo, Master Data) permitindo múltiplos usos Melhorar a eficiência operacional
- identificando dados e processos redundantes
- Evite o uso de dados desatualizados ou incorretos
- Reduza o tempo de pesquisa orientada a dados
- Melhore a comunicação entre consumidores de dados e profissionais de TI
- Crie uma análise de impacto precisa, reduzindo assim o risco de falha do projeto

- Melhore o tempo de colocação no mercado reduzindo o tempo do ciclo de vida do desenvolvimento do sistema
- Reduza os custos de treinamento e diminua o impacto da rotatividade de pessoal por meio da documentação completa do contexto, histórico e origem dos dados
- Apoiar a conformidade regulamentar

Os metadados auxiliam na representação consistente de informações, otimizando os recursos de fluxo de trabalho e protegendo informações confidenciais, principalmente quando a conformidade regulatória é necessária.

As organizações obtêm mais valor de seus ativos de dados se seus dados forem de alta qualidade. Dados de qualidade dependem de governança. Como explicam os dados e processos que permitem que as organizações funcionem, os metadados são essenciais para a governança de dados. Se os metadados são um guia para os dados em uma organização, eles devem ser bem gerenciados. Metadados mal gerenciados levam a:

- Dados redundantes e processos de gerenciamento de dados
- Dicionários, repositórios e outros armazenamentos de metadados replicados e redundantes
- Definições inconsistentes de elementos de dados e riscos associados ao uso indevido de dados
- Fontes e versões concorrentes e conflitantes de Metadados que reduzem a confiança dos consumidores de dados
- Dúvida sobre a confiabilidade dos metadados e dados

O gerenciamento de metadados bem executado permite uma compreensão consistente dos recursos de dados e um desenvolvimento interorganizacional mais eficiente.

## 1.2 Objetivos e Princípios

Os objetivos do gerenciamento de metadados incluem:

- Documentar e gerenciar o conhecimento organizacional da terminologia empresarial relacionada a dados para garantir que as pessoas entendam o conteúdo dos dados e possam usá-los de forma consistente
- Coletar e integrar metadados de diversas fontes para garantir que as pessoas entendam as semelhanças e diferenças entre dados de diferentes partes da organização
- Garantir a qualidade, consistência, atualidade e segurança dos metadados
- Fornecer formas padronizadas para tornar os metadados acessíveis a Consumidores de metadados (pessoas, sistemas e processos)
- Estabelecer ou impor o uso de padrões técnicos de metadados para permitir a troca de dados

A implementação de uma solução de metadados bem-sucedida segue estes princípios orientadores:

- **Comprometimento organizacional:** Garanta o comprometimento organizacional (suporte e financiamento da alta gerência) com o gerenciamento de metadados como parte de uma estratégia geral para gerenciar dados como um ativo empresarial.
- **Estratégia:** Desenvolva uma estratégia de Metadados que considere como os Metadados serão criados, mantidos, integrados e acessados. A estratégia deve direcionar requisitos, que devem ser definidos antes de avaliar, comprar e instalar produtos de gerenciamento de Metadados. A estratégia de Metadados deve estar alinhada com as prioridades do negócio.
- **Perspectiva empresarial:** adote uma perspectiva empresarial para garantir extensibilidade futura, mas implemente por meio de entrega iterativa e incremental para agregar valor.
- **Socialização:** Comunicar a necessidade de Metadados e

o propósito de cada tipo de Metadados; a socialização do valor dos Metadados incentivará o uso comercial e, igualmente importante, a contribuição da expertise empresarial.

- **Acesso:** certifique-se de que os membros da equipe saibam como acessar e usar metadados.
- **Qualidade:** reconheça que os metadados são frequentemente produzidos por meio de processos existentes (modelagem de dados, SDLC, definição de processos de negócios) e responsabilize os proprietários dos processos pela qualidade dos metadados.
- **Auditoria:** defina, aplique e audite padrões para metadados para simplificar a integração e permitir o uso.
- **Melhoria:** Criar um mecanismo de feedback para que os consumidores possam informar a equipe de Gerenciamento de Metadados sobre Metadados incorretos ou desatualizados.

## 1.3 Conceitos Essenciais

### 1.3.1 Metadados vs. Dados

Conforme declarado na introdução do capítulo, Metadados são um tipo de dado e devem ser gerenciados como tal. Uma questão que algumas organizações enfrentam é onde traçar a linha entre dados que não são Metadados e dados que são Metadados. Conceitualmente, essa linha está relacionada ao nível de abstração representado pelos dados. Por exemplo, ao relatar a divulgação da vigilância da Administração de Segurança Nacional dos EUA sobre o uso de telefones por pessoas nos EUA, números de telefone e horários de chamadas eram rotineiramente chamados de "Metadados", o que implicava que os dados "reais" compreendiam apenas o conteúdo das conversas telefônicas. O senso comum reconhece que números de telefone e duração de chamadas telefônicas também são apenas dados simples.<sup>70</sup> Uma regra prática pode ser que os Metadados de uma pessoa são os dados de outra. Mesmo algo que parece Metadados (por exemplo, uma lista de nomes de colunas) pode ser apenas dados simples - se, por exemplo, esses dados fossem a entrada para uma análise destinada a entender o conteúdo dos dados em



diferentes organizações.

Para gerenciar seus Metadados, as organizações não devem se preocupar com as distinções filosóficas. Em vez disso, elas devem definir requisitos de Metadados focados no que elas precisam de Metadados (para criar novos dados, entender dados existentes, habilitar movimentação entre sistemas, acessar dados, compartilhar dados) e dados de origem para atender a esses requisitos.

### *1.3.2 Tipos de Metadados Os*

metadados são frequentemente categorizados em três tipos: comercial, técnico e operacional. Essas categorias permitem que as pessoas entendam a gama de informações que se enquadram no guarda-chuva geral dos Metadados, bem como as funções por meio das quais os Metadados são produzidos. Dito isso, as categorias também podem levar à confusão, especialmente se as pessoas ficarem presas em perguntas sobre a qual categoria um conjunto de Metadados pertence ou quem deve usá-lo. É melhor pensar nessas categorias em relação à origem dos Metadados, em vez de como eles são usados. Em relação ao uso, as distinções entre os tipos de Metadados não são rígidas. A equipe técnica e operacional usa Metadados 'comerciais' e vice-versa.

Fora da tecnologia da informação, por exemplo, em biblioteconomia ou ciência da informação, os metadados são descritos usando um conjunto diferente de categorias:

- Metadados descritivos (por exemplo, título, autor e assunto) descrevem um recurso e permitem identificação e recuperação.
- Metadados estruturais descrevem relacionamentos dentro e entre recursos e suas partes componentes (por exemplo, número de páginas, número de capítulos).
- Metadados administrativos (por exemplo, números de versão, datas de arquivamento) são usados para gerenciar recursos ao longo de seu ciclo de vida.

Essas categorias podem ajudar a informar o processo de definição de requisitos de metadados.

#### 1.3.2.1 Metadados de negócios

*Metadados de Negócios* focam amplamente no conteúdo e condição dos dados e incluem detalhes relacionados à governança de dados. Metadados de Negócios incluem nomes e definições não técnicos de conceitos, áreas de assunto, entidades e atributos; tipos de dados de atributos e outras propriedades de atributos; descrições de intervalo; cálculos; algoritmos e regras de negócios; valores de domínio válidos e suas definições. Exemplos de Metadados de Negócios incluem:

- Definições e descrições de conjuntos de dados, tabelas e colunas
- Regras de negócios, regras de transformação, cálculos e derivações
- Modelos de dados
- Regras de qualidade de dados e resultados de medição
- Cronogramas pelos quais os dados são atualizados
- Proveniência e linhagem de dados
- Padrões de dados
- Designações do sistema de registro para elementos de dados
- Restrições de valor válidas
- Informações de contato das partes interessadas (por exemplo, proprietários de dados, administradores de dados)
- Nível de segurança/privacidade dos dados
- Problemas conhecidos com dados
- Notas sobre o uso de dados

#### 1.3.2.2 Metadados Técnicos

*Os metadados técnicos* fornecem informações sobre os detalhes técnicos

de dados, os sistemas que armazenam dados e os processos que os movem dentro e entre sistemas. Exemplos de Metadados Técnicos incluem:

- Nomes de tabelas e colunas do banco de dados físico
- Propriedades da coluna
- Propriedades do objeto de banco de dados
- Permissões de acesso
- Regras de dados CRUD (criar, substituir, atualizar e excluir)
- Modelos de dados físicos, incluindo nomes de tabelas de dados, chaves e índices
- Relacionamentos documentados entre os modelos de dados e os ativos físicos
- Detalhes do trabalho ETL
- Definições de esquema de formato de arquivo
- Documentação de mapeamento de origem para destino
- Documentação da linhagem de dados, incluindo informações de impacto de alterações upstream e downstream
- Nomes e descrições de programas e aplicativos
- Cronogramas e dependências do ciclo de atualização de conteúdo
- Regras de recuperação e backup
- Direitos de acesso a dados, grupos, funções

#### 1.3.2.3 Metadados Operacionais *Metadados*

*Operacionais* descrevem detalhes do processamento e acesso de dados. Por exemplo:

- Logs de execução de tarefas para programas em lote
- Histórico de extratos e resultados
- Anomalias de programação

- Resultados de auditoria, balanço, medições de controle
- Registros de erros
- Padrões de acesso a relatórios e consultas, frequência e tempo de execução
- Plano e execução de manutenção de patches e versões, nível de patch atual
- Provisões de backup, retenção, data de criação e recuperação de desastres
- Requisitos e disposições do SLA
- Padrões volumétricos e de uso
- Regras de arquivamento e retenção de dados, arquivos relacionados
- Critérios de purga
- Regras e acordos de compartilhamento de dados
- Funções e responsabilidades técnicas, contatos

### *1.3.3 Padrão de Registro de Metadados ISO/IEC 11179 O Padrão*

de Registro de Metadados da ISO, ISO/IEC 11179, fornece uma estrutura para definir um registro de Metadados. Ele é projetado para permitir a troca de dados orientada por Metadados, com base em definições exatas de dados, começando com elementos de dados. O padrão é estruturado em várias partes:

- Parte 1: Estrutura para a geração e padronização de elementos de dados
- Parte 3: Atributos básicos dos elementos de dados
- Parte 4: Regras e Diretrizes para a Formulação de Dados  
Definições
- Parte 5: Princípios de nomenclatura e identificação de dados  
Elementos
- Parte 6: Registro de Elementos de Dados

### *1.3.4 Metadados para Dados Não Estruturados*

Por sua natureza, todos os dados têm alguma estrutura, embora nem todos sejam formalmente estruturados nas linhas, colunas e registros familiares de bancos de dados relacionais. Quaisquer dados que não estejam em um banco de dados ou arquivo de dados, incluindo documentos ou outras mídias, são considerados dados não estruturados. (Ver Capítulos [9](#) e [14](#).)

Os metadados são tão essenciais para o gerenciamento de dados não estruturados quanto para o gerenciamento de dados estruturados – talvez até mais.

Pense novamente sobre a analogia do catálogo de fichas da introdução do capítulo. Livros e revistas em uma biblioteca são bons exemplos de dados não estruturados. O uso principal dos Metadados em um catálogo de fichas é encontrar os materiais que se está procurando, qualquer que seja o formato.

Metadados para dados não estruturados incluem Metadados descritivos, como informações de catálogo e palavras-chave de tesouros; Metadados estruturais, como tags, estruturas de campo, formato; Metadados administrativos, como fontes, cronogramas de atualização, direitos de acesso e informações de navegação; Metadados bibliográficos, como entradas de catálogo de biblioteca; Metadados de manutenção de registros, como políticas de retenção; e Metadados de preservação, como armazenamento, condição de arquivamento e regras de conservação. (Consulte [o Capítulo 9](#).)

Embora a maioria das afirmações sobre metadados para dados não estruturados esteja conectada a preocupações tradicionais de gerenciamento de conteúdo, novas práticas estão surgindo em torno do gerenciamento de dados não estruturados em data lakes. Organizações que desejam aproveitar os data lakes, usando plataformas de Big Data como Hadoop, estão descobrindo que precisam catalogar os dados ingeridos para permitir acesso posterior. A maioria coloca em prática processos para coletar Metadados como parte da ingestão de dados. Um conjunto mínimo de atributos de Metadados precisa ser coletado sobre cada objeto ingerido no data lake (por exemplo, nome, formato, fonte, versão, data de recebimento, etc.). Isso produz um catálogo de conteúdos do data lake.

### *1.3.5 Fontes de Metadados*

Como deve ficar claro pelos tipos de Metadados, os Metadados podem ser coletados de muitas fontes diferentes. Além disso, se os Metadados de aplicativos e bancos de dados foram bem gerenciados, eles podem ser simplesmente

colhidos e integrados. No entanto, a maioria das organizações não gerencia bem os Metadados no nível do aplicativo, porque os Metadados são frequentemente criados como um subproduto do processamento do aplicativo, em vez de um produto final (ou seja, não são criados com o consumo em mente). Assim como acontece com outras formas de dados, há muito trabalho na preparação dos Metadados antes que eles possam ser integrados.

A maioria dos Metadados operacionais é gerada conforme os dados são processados. A chave para usar esses Metadados é coletá-los em um formato utilizável e garantir que os responsáveis por interpretá-los tenham as ferramentas necessárias para isso. Tenha em mente que interpretar dados em lugares como os próprios logs de erros requer Metadados que descrevam os logs.

Da mesma forma, uma grande parte dos metadados técnicos pode ser coletada de objetos de banco de dados.

É possível fazer engenharia reversa do conhecimento sobre dados de sistemas existentes e coletar Metadados de negócios de dicionários de dados, modelos e documentação de processos existentes (Loshin, 2001; Aiken, 1995), mas há riscos em fazer isso. O maior risco é não saber quanto cuidado foi tomado para desenvolver e refinar as definições em primeiro lugar. Se as definições forem subdesenvolvidas ou ambíguas, elas não fornecerão aos consumidores de dados as informações de que precisam para entender os dados que estão usando.

É melhor ser intencional sobre o desenvolvimento de definições do que simplesmente aceitar as existentes. O desenvolvimento de definições leva tempo e o conjunto de habilidades certo (por exemplo, habilidades de escrita e facilitação). É por isso que o desenvolvimento de Metadados de negócios requer administração. (Veja o [Capítulo 3](#).)

---

Grande parte dos Metadados técnicos necessários para gerenciar bancos de dados e os Metadados comerciais necessários para usar dados podem ser coletados e desenvolvidos como parte do trabalho do projeto. Por exemplo, o processo de modelagem de dados requer discussões sobre o significado dos elementos de dados e a relação entre eles. O conhecimento compartilhado durante essas discussões deve ser capturado e preparado para uso em Dicionários de Dados, Glossários Comerciais e outros repositórios. Os próprios modelos de dados incluem detalhes importantes sobre as características físicas dos dados. O tempo deve ser alocado para garantir que o projeto

Os artefatos contêm metadados de alta qualidade que estão alinhados aos padrões empresariais.

Metadados de negócios bem definidos são reutilizáveis de projeto para projeto e podem impulsionar um entendimento consistente de como os conceitos de negócios são representados em diferentes conjuntos de dados. Como parte do desenvolvimento intencional de Metadados para que possam ser reutilizados, uma organização também pode planejar a integração de Metadados. Por exemplo, ela pode desenvolver um inventário de sistemas, e todos os Metadados relacionados a um sistema específico podem ser marcados com o mesmo identificador de sistema.

Criar Metadados por si só raramente funciona bem. A maioria das organizações não financiará esse tipo de esforço e, mesmo quando o faz, é improvável que implemente processos para manutenção. Nesse aspecto, como em outros, Metadados são como outros dados: devem ser criados como o produto de um processo bem definido, usando ferramentas que darão suporte à sua qualidade geral. Administradores e outros profissionais de gerenciamento de dados devem garantir que haja processos em vigor para manter Metadados relacionados a esses processos. Por exemplo, se uma organização coleta Metadados críticos de seus modelos de dados, ela deve garantir que haja um processo de gerenciamento de mudanças em vigor para manter os modelos atualizados.

Para dar uma ideia da amplitude dos metadados em qualquer organização, uma variedade de fontes é descrita aqui, em ordem alfabética e não de prioridade.

#### 1.3.5.1 Repositórios de Metadados de Aplicativos Um

repositório de Metadados se refere às tabelas físicas nas quais os Metadados são armazenados. Frequentemente, eles são incorporados em ferramentas de modelagem, ferramentas de BI e outros aplicativos. Conforme uma organização amadurece, ela desejará integrar Metadados de repositórios nesses aplicativos para permitir que os consumidores de dados olhem através da amplitude das informações.

#### 1.3.5.2 Glossário de negócios O

objetivo de um glossário de negócios é documentar e armazenar os conceitos e a terminologia de negócios de uma organização, as definições e os relacionamentos entre esses termos.

Em muitas organizações, o glossário de negócios é apenas uma planilha.

No entanto, conforme as organizações amadurecem, elas frequentemente compram ou criam glossários que contêm informações robustas e a capacidade de gerenciá-las ao longo do tempo. Como acontece com todos os sistemas orientados a dados, os glossários de negócios devem ser arquitetados para levar em conta hardware, software, banco de dados, processos e recursos humanos com diferentes funções e responsabilidades. O aplicativo de glossário de negócios é estruturado para atender aos requisitos funcionais dos três públicos principais:

- **Usuários empresariais:** analistas de dados, analistas de pesquisa, gerentes e executivos usam o glossário empresarial para entender terminologia e dados.
- **Data Stewards:** Os Data Stewards usam o glossário de negócios para gerenciar o ciclo de vida de termos e definições e para aprimorar o conhecimento empresarial associando ativos de dados a termos do glossário; por exemplo, vinculando termos a métricas de negócios, relatórios, análise de qualidade de dados ou componentes de tecnologia. Os Data Stewards levantam questões de terminologia e uso e ajudam a resolver diferenças em toda a organização.
- **Usuários técnicos:** Usuários técnicos usam o glossário de negócios para tomar decisões de arquitetura, design de sistemas e desenvolvimento, além de conduzir análises de impacto.

O glossário de negócios deve capturar atributos de termos de negócios, como  
como:

- Nome do termo, definição, sigla ou abreviação e quaisquer sinônimos
- Unidade de negócios e ou aplicação responsável por gerenciar os dados associados à terminologia
- Nome da pessoa que identifica o termo e data de atualização
- Associação de categorização ou taxonomia para o termo (associação funcional empresarial)



- Definições conflitantes que precisam de resolução, natureza do problema, cronograma de ação
- Mal-entendidos comuns em termos
- Algoritmos que suportam definições
- Linhagem
- Fonte oficial ou autorizada para os dados que dão suporte à prazo

Cada implementação de glossário de negócios deve ter um conjunto básico de relatórios para dar suporte aos processos de governança. É recomendado que as organizações não "imprimam o glossário" porque o conteúdo do glossário não é estático. Os administradores de dados são geralmente responsáveis pelo desenvolvimento, uso, operações e relatórios do glossário. Os relatórios incluem o rastreamento de novos termos e definições que ainda não foram revisados, aqueles em status pendente e aqueles que estão sem definições ou outros atributos. (Consulte [a Seção 6.4.](#))

A facilidade de uso e a funcionalidade podem variar amplamente. Quanto mais simples e fácil for a busca no glossário de negócios, maior a probabilidade de o conteúdo do glossário ser usado. No entanto, a característica mais importante de um glossário é que ele contém conteúdo robusto.

#### 1.3.5.3 Ferramentas de Business Intelligence (BI) As

ferramentas de Business Intelligence produzem vários tipos de metadados relevantes para o design de Business Intelligence, incluindo informações gerais, classes, objetos, itens derivados e calculados, filtros, relatórios, campos de relatório, layout de relatório, usuários de relatórios, frequência de distribuição de relatórios e canais de distribuição de relatórios.

#### 1.3.5.4 Ferramentas de gerenciamento de configuração

Ferramentas de gerenciamento de configuração ou bancos de dados (CMDB) fornecem a capacidade de gerenciar e manter Metadados especificamente relacionados aos ativos de TI, os relacionamentos entre eles e os detalhes contratuais do ativo. Cada ativo no banco de dados CMDB é chamado de item de configuração (CI).

Metadados padrão são coletados e gerenciados

para cada tipo de CI. Muitas organizações integram o CMDB com os processos de gerenciamento de mudanças para identificar os ativos ou aplicativos relacionados impactados por uma mudança em um ativo específico. Os repositórios fornecem mecanismos para vincular os ativos no repositório de Metadados aos detalhes de implementação física reais no CMDB para dar uma imagem completa dos dados e das plataformas.

#### 1.3.5.5 Dicionários de Dados

Um dicionário de dados define a estrutura e o conteúdo de conjuntos de dados, geralmente para um único banco de dados, aplicativo ou warehouse. O dicionário pode ser usado para gerenciar nomes, descrições, estrutura, características, requisitos de armazenamento, valores padrão, relacionamentos, exclusividade e outros atributos de cada elemento de dados em um modelo. Ele também deve conter definições de tabela ou arquivo. Dicionários de dados são incorporados em ferramentas de banco de dados para a criação, operação e manipulação de dados contidos neles. Para tornar esses Metadados disponíveis para consumidores de dados, eles devem ser extraídos do banco de dados ou ferramentas de modelagem. Dicionários de dados também podem descrever em terminologia empresarial quais elementos de dados estão disponíveis para a comunidade, provisionados sob quais restrições de segurança e aplicados em qual processo empresarial. Tempo pode ser economizado ao definir, publicar e manter uma camada semântica para relatórios e análises, aproveitando o conteúdo diretamente do modelo lógico. No entanto, conforme observado anteriormente, as definições existentes devem ser usadas com cautela, especialmente em uma organização com baixo nível de maturidade em torno do gerenciamento de metadados.

Muitos processos, relacionamentos e terminologias de negócios importantes são explicados durante o desenvolvimento do modelo de dados. Essas informações, capturadas no modelo de dados lógico, geralmente são perdidas quando estruturas físicas são implantadas na produção. Um dicionário de dados pode ajudar a garantir que essas informações não sejam perdidas inteiramente para a organização e que os modelos lógico e físico sejam mantidos em acordo após a implantação da produção.

#### 1.3.5.6 Ferramentas de integração de

dados Muitas ferramentas de integração de dados são usadas para executáveis para mover dados de um sistema para outro ou entre vários módulos dentro do mesmo.

mesmo sistema. Muitas dessas ferramentas geram arquivos transitórios, que podem conter cópias ou cópias derivadas dos dados. Essas ferramentas são capazes de carregar dados de várias fontes e, em seguida, operar nos dados carregados, por meio de agrupamento, correção, reformatação, junção, filtragem ou outras operações e, em seguida, gerar dados de saída, que são distribuídos para os locais de destino. Elas documentam a linhagem como dados à medida que se movem entre os sistemas. Qualquer solução de Metadados bem-sucedida deve ser capaz de usar os Metadados de linhagem à medida que se movem pelas ferramentas de integração e expô-los como uma linhagem holística das fontes reais aos destinos finais.

Ferramentas de integração de dados fornecem interfaces de aplicativos (API) para permitir que repositórios externos de Metadados extraiam as informações de linhagem e os Metadados de arquivos transitórios. Uma vez que o repositório de Metadados coleta as informações, algumas ferramentas podem gerar um diagrama de linhagem holístico para qualquer elemento de dados. Ferramentas de integração de dados também fornecem Metadados sobre a execução dos vários trabalhos de integração de dados, incluindo a última execução bem-sucedida, duração e status do trabalho. Alguns repositórios de Metadados podem extrair as estatísticas de tempo de execução de integração de dados e Metadados e expô-los junto com os elementos de dados. (Consulte os Capítulos 6 e 8.)

#### 1.3.5.7 Gerenciamento de Banco de Dados e Catálogos de Sistema Catálogos

de banco de dados são uma fonte importante de Metadados. Eles descrevem o conteúdo de bancos de dados, juntamente com informações de dimensionamento, versões de software, status de implantação, tempo de atividade da rede, tempo de atividade da infraestrutura, disponibilidade e muitos outros atributos operacionais de Metadados. A forma mais comum de banco de dados é relacional. Bancos de dados relacionais gerenciam os dados como um conjunto de tabelas e colunas, onde uma tabela contém uma ou mais colunas, índices, restrições, visualizações e procedimentos. Uma solução de Metadados deve ser capaz de se conectar aos vários bancos de dados e conjuntos de dados e ler todos os Metadados expostos pelo banco de dados. Algumas das ferramentas de repositório de Metadados podem integrar os Metadados expostos das ferramentas de gerenciamento do sistema para fornecer uma imagem mais holística sobre os ativos físicos capturados.

#### 1.3.5.8 Ferramentas de gerenciamento de mapeamento de

dados As ferramentas de gerenciamento de mapeamento são usadas durante a análise e o design

fase de um projeto para transformar requisitos em especificações de mapeamento, que podem então ser consumidas diretamente por uma ferramenta de integração de dados ou usadas pelos desenvolvedores para gerar código de integração de dados. A documentação de mapeamento também é frequentemente mantida em documentos Excel em toda a empresa. Os fornecedores agora estão considerando repositórios centralizados para as especificações de mapeamento com recursos para executar controle de versão e análise de alterações entre versões. Muitas ferramentas de mapeamento se integram com ferramentas de integração de dados para automatizar a geração dos programas de integração de dados e a maioria pode trocar dados com outros repositórios de Metadados e Dados de Referência. ([Consulte o Capítulo 8.](#))

#### 1.3.5.9 Ferramentas de qualidade de

dados As ferramentas de qualidade de dados avaliam a qualidade dos dados por meio de regras de validação. A maioria dessas ferramentas oferece a capacidade de trocar as pontuações de qualidade e padrões de perfis com outros repositórios de metadados, permitindo que o repositório de metadados anexe as pontuações de qualidade aos ativos físicos relevantes.

#### 1.3.5.10 Diretórios e catálogos Enquanto

dicionários de dados e glossários contêm informações detalhadas sobre terminologia, tabelas e campos, um diretório ou catálogo contém informações sobre sistemas, fontes e locais de dados dentro de uma organização. Um diretório de Metadados é particularmente útil para desenvolvedores e superusuários de dados, como equipes de administração de dados e analistas de dados, para entender o escopo dos dados na empresa, seja para pesquisar problemas ou para encontrar informações sobre o fornecimento de novos aplicativos.

#### 1.3.5.11 Ferramentas de Mensagens de

Eventos Ferramentas de mensagens de eventos movem dados entre sistemas diversos. Para fazer isso, elas exigem muitos Metadados. Elas também geram Metadados que descrevem esse movimento. Essas ferramentas incluem interfaces gráficas por meio das quais gerenciam a lógica do movimento de dados. Elas podem exportar os detalhes de implementação das interfaces, lógica de movimento e estatísticas de processamento para outros repositórios de Metadados.

#### 1.3.5.12 Ferramentas de modelagem e repositórios

Ferramentas de modelagem de dados são usadas para construir vários tipos de modelos de dados: conceituais, lógicos e físicos. Essas ferramentas produzem Metadados relevantes para o design do aplicativo ou modelo de sistema, como áreas de assunto, entidades lógicas, atributos lógicos, relacionamentos de entidade e atributo, supertipos e subtipos, tabelas, colunas, índices, chaves primárias e estrangeiras, restrições de integridade e outros tipos de atribuição dos modelos. Os repositórios de metadados podem ingerir os modelos criados por essas ferramentas e integrar os Metadados importados no repositório. As ferramentas de modelagem geralmente são a fonte do conteúdo do dicionário de dados.

#### 1.3.5.13 Repositórios de Dados de Referência

Os Dados de Referência documentam os valores comerciais e as descrições dos vários tipos de dados enumerados (domínios) e seu uso contextual em um sistema. As ferramentas usadas para gerenciar Dados de Referência também são capazes de gerenciar relacionamentos entre os vários valores codificados dentro do mesmo domínio ou entre domínios. Esses conjuntos de ferramentas normalmente fornecem recursos para enviar os Dados de Referência coletados para um repositório de Metadados, que por sua vez fornecerá mecanismos para associar os Dados de Referência ao glossário comercial e aos locais onde são fisicamente implementados, como colunas ou campos.

#### 1.3.5.14 Registros de serviço Um

registro de serviço gerencia e armazena as informações técnicas sobre serviços e pontos de extremidade de serviço de uma perspectiva de arquitetura orientada a serviços (SOA). Por exemplo, definições, interfaces, operações, parâmetros de entrada e saída, políticas, versões e cenários de uso de amostra. Alguns dos metadados mais importantes relacionados a serviços incluem versão do serviço, local do serviço, data center, disponibilidade, data de implantação, porta do serviço, endereço IP, porta de estatísticas, tempo limite de conexão e tempo limite de nova tentativa de conexão. Os registros de serviço podem ser interrogados para satisfazer várias necessidades, como exibir uma lista de todos os serviços disponíveis, serviços com uma versão específica, serviços obsoletos ou detalhes sobre um serviço específico. Os serviços também podem ser revisados para possível reutilização. As informações contidas nesse

fornece fatos importantes sobre quais dados existem e como eles se movem entre vários sistemas ou aplicativos. Metadados em repositórios de serviço podem ser extraídos e incorporados com Metadados coletados de outras ferramentas para fornecer uma imagem completa de como os dados estão se movendo entre os vários sistemas.

#### 1.3.5.15 Outros armazenamentos de metadados

Outros armazenamentos de metadados incluem listas especializadas, como registros de eventos, listas de fontes ou interfaces, conjuntos de códigos, léxicos, esquemas espaciais e temporais, referências espaciais e distribuição de conjuntos de dados geográficos digitais, repositórios de repositórios e regras de negócios.

#### 1.3.6 *Tipos de Arquitetura de Metadados Como*

outras formas de dados, os Metadados têm um ciclo de vida. Conceitualmente, todas as soluções de gerenciamento de Metadados incluem camadas arquitetônicas que correspondem a pontos no ciclo de vida dos Metadados:

- Criação e fornecimento de metadados
- Armazenamento de metadados em um ou mais repositórios
- Integração de metadados
- Entrega de metadados
- Uso de metadados
- Controle e gerenciamento de metadados

Diferentes abordagens arquitetônicas podem ser usadas para obter, armazenar, integrar, manter e tornar metadados acessíveis aos consumidores.

##### 1.3.6.1 Arquitetura de Metadados Centralizada

Uma arquitetura centralizada consiste em um único repositório de metadados que contém cópias de metadados de várias fontes.

Organizações com recursos de TI limitados, ou aquelas que buscam automatizar o máximo possível, podem optar por evitar essa opção de arquitetura.

Organizações que buscam um alto grau de consistência dentro do

Um repositório comum de metadados pode se beneficiar de uma arquitetura centralizada.

As vantagens de um repositório centralizado incluem:

- Alta disponibilidade, pois é independente dos sistemas de origem
- Recuperação rápida de metadados, já que o repositório e a consulta residem juntos
- Estruturas de banco de dados resolvidas não afetadas pela natureza proprietária de sistemas comerciais ou de terceiros
- Os metadados extraídos podem ser transformados, personalizados ou aprimorados com metadados adicionais que podem não residir no sistema de origem, melhorando a qualidade

Algumas limitações da abordagem centralizada incluem:

- Processos complexos são necessários para garantir que as alterações nos metadados de origem sejam rapidamente replicadas no repositório
- A manutenção de um repositório centralizado pode ser dispendiosa
- A extração pode exigir módulos personalizados ou middleware
- A validação e manutenção de código personalizado podem aumentar as demandas tanto da equipe interna de TI quanto dos fornecedores de software

A Figura 85 mostra como os Metadados são coletados em um repositório de Metadados autônomo com seu próprio repositório interno de Metadados. O repositório interno é preenchido por meio de uma importação programada (setas) dos Metadados das várias ferramentas. Por sua vez, o repositório centralizado expõe um portal para os usuários finais enviarem suas consultas. O portal de Metadados passa a solicitação para o repositório centralizado de Metadados. O repositório centralizado atenderá à solicitação dos Metadados coletados. Neste

tipo de implementação, a capacidade de passar a solicitação do usuário para várias ferramentas diretamente não é suportada. A busca global pelos Metadados coletados das várias ferramentas é possível devido à coleta de vários Metadados no repositório centralizado.

Figura 85 Metadados centralizados  
Arquitetura

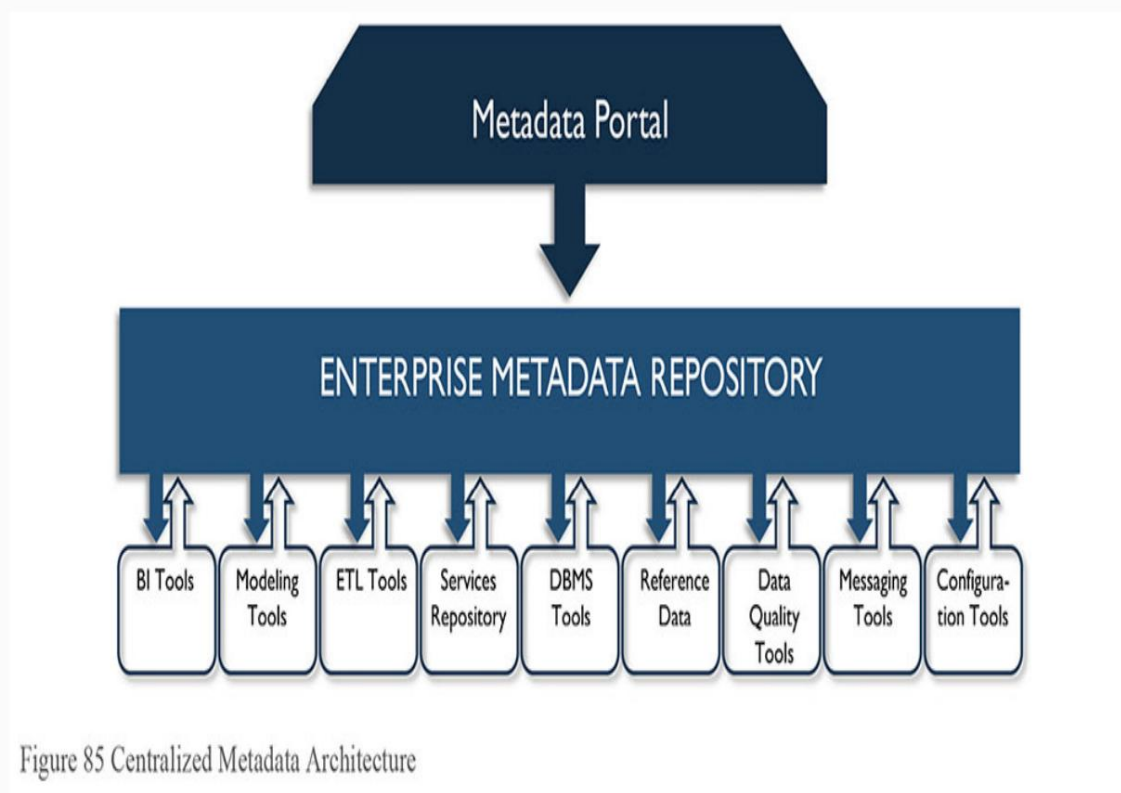


Figure 85 Centralized Metadata Architecture

#### 1.3.6.2 Arquitetura de Metadados Distribuídos

Uma arquitetura completamente distribuída mantém um único ponto de acesso. O mecanismo de recuperação de metadados responde às solicitações do usuário recuperando dados de sistemas de origem em tempo real; não há repositório persistente. Nessa arquitetura, o ambiente de gerenciamento de metadados mantém os catálogos do sistema de origem e as informações de pesquisa necessárias para processar consultas e pesquisas do usuário de forma eficaz. Um corretor de solicitação de objeto comum ou protocolo de middleware semelhante acessa esses sistemas de origem.



As vantagens da arquitetura de metadados distribuída incluem:

- Os metadados são sempre tão atuais e válidos quanto possível porque são recuperados de sua fonte
- As consultas são distribuídas, possivelmente melhorando o tempo de resposta e processamento
- As solicitações de metadados de sistemas proprietários são limitadas ao processamento de consultas, em vez de exigir uma compreensão detalhada das estruturas de dados proprietárias, minimizando assim o esforço de implementação e manutenção necessário
- O desenvolvimento do processamento automatizado de consultas de metadados é provavelmente mais simples, exigindo intervenção manual mínima
- O processamento em lote é reduzido, sem processos de replicação ou sincronização de metadados

Arquiteturas distribuídas também têm limitações:

- Nenhuma capacidade de oferecer suporte a dados definidos pelo usuário ou inseridos manualmente  
Entradas de metadados, pois não há repositório onde colocar essas adições
- Padronização da apresentação de Metadados de vários sistemas
- Os recursos de consulta são diretamente afetados pela disponibilidade dos sistemas de origem participantes
- A qualidade dos metadados depende exclusivamente dos sistemas de origem participantes

Figura 86 Metadados Distribuídos

## Arquitetura

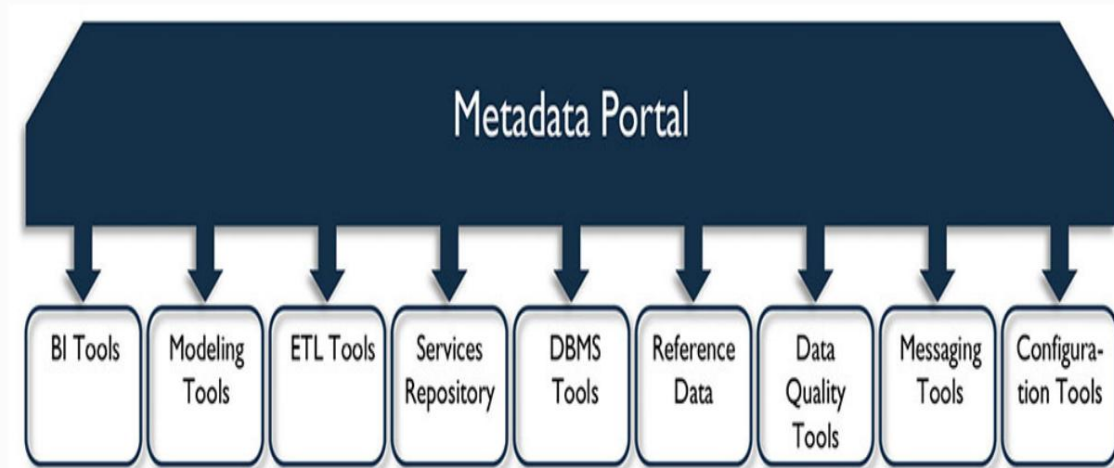


Figure 86 Distributed Metadata Architecture

A Figura 86 ilustra uma arquitetura de Metadados distribuída. Não há um repositório de Metadados centralizado e o portal passa as solicitações dos usuários para a ferramenta apropriada para execução. Como não há um repositório centralizado para os Metadados serem coletados das várias ferramentas, cada solicitação tem que ser delegada às fontes; portanto, não existe capacidade para uma pesquisa global nas várias fontes de Metadados.

### 1.3.6.3 Arquitetura de Metadados Híbrida Uma

arquitetura híbrida combina características de arquiteturas centralizadas e distribuídas. Os metadados ainda se movem diretamente dos sistemas de origem para um repositório centralizado. No entanto, o design do repositório considera apenas os Metadados adicionados pelo usuário, os itens padronizados críticos e as adições de fontes manuais.

A arquitetura se beneficia da recuperação quase em tempo real de Metadados de sua fonte e Metadados aprimorados para atender às necessidades do usuário de forma mais eficaz, quando necessário. A abordagem híbrida reduz o esforço para intervenção manual de TI e funcionalidade de acesso codificado personalizado a sistemas proprietários. Os Metadados são tão atuais e válidos quanto possível no momento do uso, com base nas prioridades e requisitos do usuário. A arquitetura híbrida não melhora a disponibilidade do sistema.

A disponibilidade dos sistemas de origem é uma limitação, porque a natureza distribuída dos sistemas de back-end manipula o processamento de consultas. Sobrecarga adicional é necessária para vincular esses resultados iniciais com o aumento de Metadados no repositório central antes de apresentar o conjunto de resultados ao usuário final.

Muitas organizações podem se beneficiar de uma arquitetura híbrida, incluindo aquelas que têm Metadados operacionais em rápida mudança, aquelas que precisam de Metadados consistentes e uniformes e aquelas que experimentam crescimento substancial em Metadados e fontes de Metadados. Organizações com Metadados mais estáticos e perfis de crescimento de Metadados menores podem não ver o potencial máximo dessa alternativa de arquitetura.

#### 1.3.6.4 Arquitetura de metadados bidirecionais

Outra abordagem arquitetônica avançada é a arquitetura de metadados bidirecional, que permite que os metadados mudem em qualquer parte da arquitetura (fonte, integração de dados, interface do usuário) e, em seguida, o feedback é coordenado do repositório (corretor) para seu original fonte.

Vários desafios são aparentes nessa abordagem. O design força o repositório de Metadados a conter a versão mais recente da fonte de Metadados e o força a gerenciar alterações na fonte também.

As mudanças devem ser capturadas sistematicamente e depois resolvidas. Conjuntos adicionais de interfaces de processo para vincular o repositório às fontes de metadados devem ser criados e mantidos.

## Figura 87 Metadados Híbridos Arquitetura

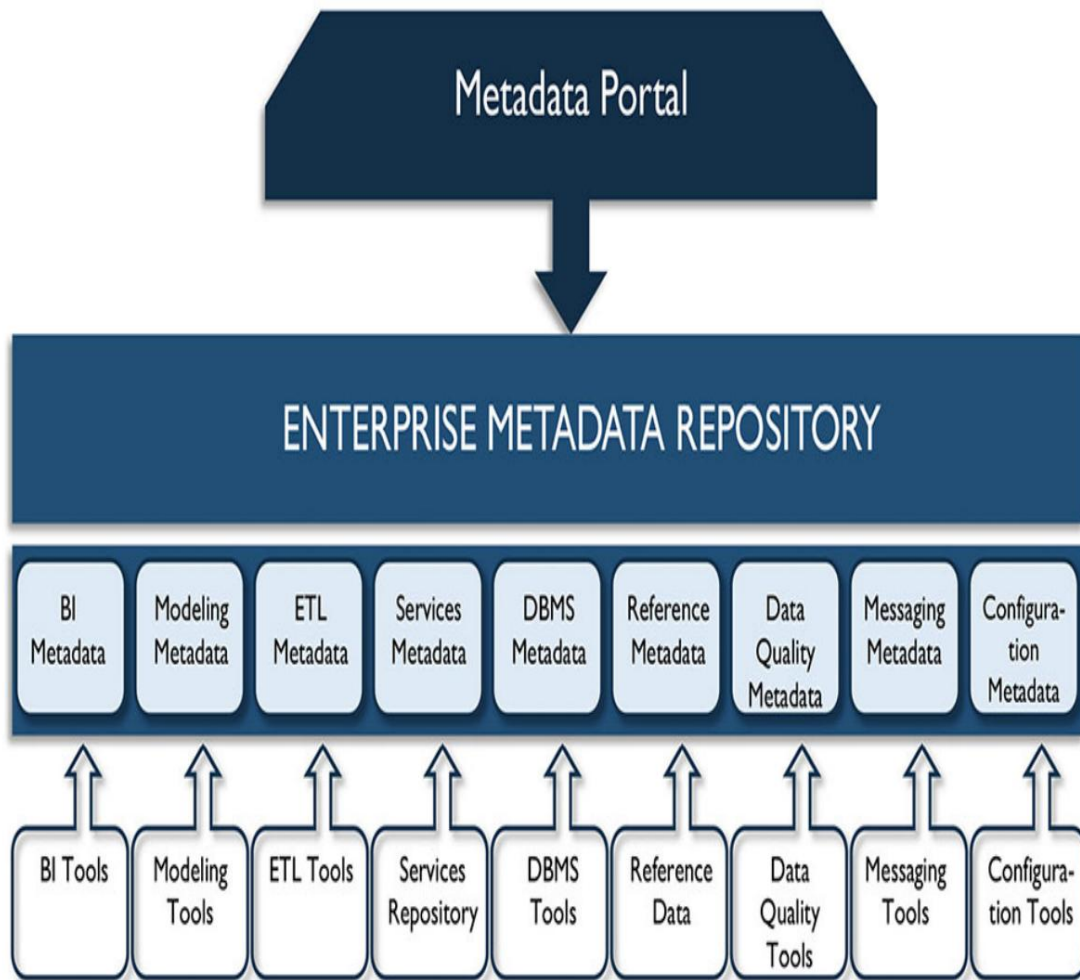


Figure 87 Hybrid Metadata Architecture

A Figura 87 ilustra como Metadados comuns de diferentes fontes são coletados em um repositório de Metadados centralizado. Os usuários enviam suas consultas ao portal de Metadados, que passa a solicitação para um repositório centralizado. O repositório centralizado tentará atender à solicitação do usuário a partir dos Metadados comuns coletados inicialmente de várias fontes. À medida que a solicitação se torna mais específica ou o usuário precisa de Metadados mais detalhados, o repositório centralizado delegará à fonte específica para pesquisar os detalhes específicos.

A pesquisa global entre as diversas ferramentas está disponível devido aos metadados comuns coletados no repositório centralizado.

## 2. Atividades

## 2.1 Definir Estratégia de Metadados Uma

estratégia de Metadados descreve como uma organização pretende gerenciar seus Metadados e como ela passará do estado atual para as práticas de estado futuro.

Uma estratégia de Metadados deve fornecer uma estrutura para equipes de desenvolvimento melhorarem o gerenciamento de Metadados. O desenvolvimento de requisitos de Metadados ajudará a esclarecer os drivers da estratégia e a identificar potenciais obstáculos para sua implementação.

A estratégia inclui definir a arquitetura de metadados corporativos do estado futuro da organização e as fases de implementação necessárias para atender aos objetivos estratégicos. As etapas incluem:

- **Iniciar planejamento estratégico de Metadados:** O objetivo da iniciação e do planejamento é permitir que a equipe de estratégia de Metadados defina suas metas de curto e longo prazo. O planejamento inclui a elaboração de um estatuto, escopo e objetivos alinhados com os esforços gerais de governança e o estabelecimento de um plano de comunicações para dar suporte ao esforço. As principais partes interessadas devem estar envolvidas no planejamento.
- **Conduzir entrevistas com as principais partes interessadas:** entrevistas com partes interessadas comerciais e técnicas fornecem uma base de conhecimento para a estratégia de metadados.
- **Avalie as fontes de metadados existentes e a arquitetura de informações:** a avaliação determina o grau relativo de dificuldade na resolução de problemas de metadados e sistemas identificados nas entrevistas e na revisão da documentação.  
  
Durante esta etapa, conduza entrevistas detalhadas com a equipe principal de TI e revise a documentação das arquiteturas do sistema, modelos de dados, etc.
- **Desenvolver arquitetura futura de Metadados:** Refine e confirme a visão futura e desenvolva a arquitetura de destino de longo prazo para o ambiente de Metadados gerenciado neste estágio. Esta fase deve levar em conta componentes estratégicos, como estrutura organizacional, alinhamento com governança e administração de dados, Metadados gerenciados

arquitetura, arquitetura de entrega de metadados, arquitetura técnica e arquitetura de segurança.

- **Desenvolva um plano de implementação em fases:** Valide, integre e priorize descobertas das entrevistas e análises de dados. Documente a estratégia de Metadados e defina uma abordagem de implementação em fases para mover do ambiente de Metadados gerenciado existente para o futuro.

A estratégia evoluirá ao longo do tempo, à medida que os requisitos de metadados, a arquitetura e o ciclo de vida dos metadados forem melhor compreendidos.

## **2.2 Entenda os requisitos de metadados** Os requisitos de metadados

começam com o conteúdo: Quais metadados são necessários e em que nível. Por exemplo, nomes físicos e lógicos precisam ser capturados para colunas e tabelas. O conteúdo de metadados é amplo e os requisitos virão de consumidores de dados comerciais e técnicos. (Consulte [a Seção 1.3.2.](#))

---

Há também muitos requisitos focados em funcionalidade associados a uma solução abrangente de metadados:

- **Volatilidade:** Com que frequência os atributos e conjuntos de metadados serão atualizados
- **Sincronização:** Tempo de atualizações em relação às alterações de origem
- **Histórico:** se as versões históricas dos metadados precisam ser mantidas
- **Direitos de acesso:** quem pode acessar os metadados e como eles acessam, juntamente com a funcionalidade específica da interface do usuário para acesso
- **Estrutura:** Como os metadados serão modelados para armazenamento
- **Integração:** O grau de integração de metadados de diferentes fontes; regras para integração

- **Manutenção:** Processos e regras para atualização de Metadados (registro e encaminhamento para aprovação)
- **Gestão:** Funções e responsabilidades para gerenciar Metadados
- **Qualidade:** Requisitos de qualidade de metadados
- **Segurança:** Alguns metadados não podem ser expostos porque isso revelará a existência de dados altamente protegidos

## 2.3 Definir Arquitetura de Metadados

Um sistema de gerenciamento de metadados deve ser capaz de extrair metadados de muitas fontes. Projete a arquitetura para ser capaz de escanear as várias fontes de metadados e atualizar periodicamente o repositório. O sistema deve suportar atualizações manuais de metadados, solicitações, pesquisas e consultas de metadados por vários grupos de usuários.

Um ambiente de Metadados gerenciado deve isolar o usuário final das várias e distintas fontes de Metadados. A arquitetura deve fornecer um único ponto de acesso para o repositório de Metadados. O ponto de acesso deve fornecer todos os recursos de Metadados relacionados de forma transparente ao usuário. Os usuários devem ser capazes de acessar Metadados sem estar cientes dos diferentes ambientes das fontes de dados. Em soluções de análise e Big Data, a interface pode ter funções amplamente definidas pelo usuário (UDF) para extrair vários conjuntos de dados, e a exposição de Metadados ao usuário final é inerente a essas personalizações. Com menos dependência de UDF em soluções, os usuários finais estarão reunindo, inspecionando e usando conjuntos de dados mais diretamente e vários Metadados de suporte geralmente são mais expostos.

O design da arquitetura depende dos requisitos específicos da organização. Três abordagens técnicas de arquitetura para construir um repositório comum de Metadados imitam as abordagens para projetar data warehouses: centralizado, distribuído e híbrido (consulte [a Seção 1.3.6](#)).

---

Todas essas abordagens levam em consideração a implementação do repositório e como os mecanismos de atualização operam.

### *2.3.1 Criar MetaModelo*

Crie um modelo de dados para o repositório de Metadados, ou metamodelo, como uma das primeiras etapas de design após a estratégia de Metadados ser concluída e os requisitos de negócios serem compreendidos. Diferentes níveis de metamodelo podem ser desenvolvidos conforme necessário; um modelo conceitual de alto nível, que explica os relacionamentos entre sistemas, e um metamodelo de nível inferior que detalha as atribuições, para descrever os elementos e processos de um modelo. Além de ser uma ferramenta de planejamento e um meio de articular requisitos, o metamodelo é em si uma fonte valiosa de Metadados.

A Figura 88 descreve um metamodelo de repositório de metadados de amostra. As caixas representam as principais entidades de alto nível, que contêm os dados.

## Figura 88 Exemplo de metadados Metamodelo de Repositório



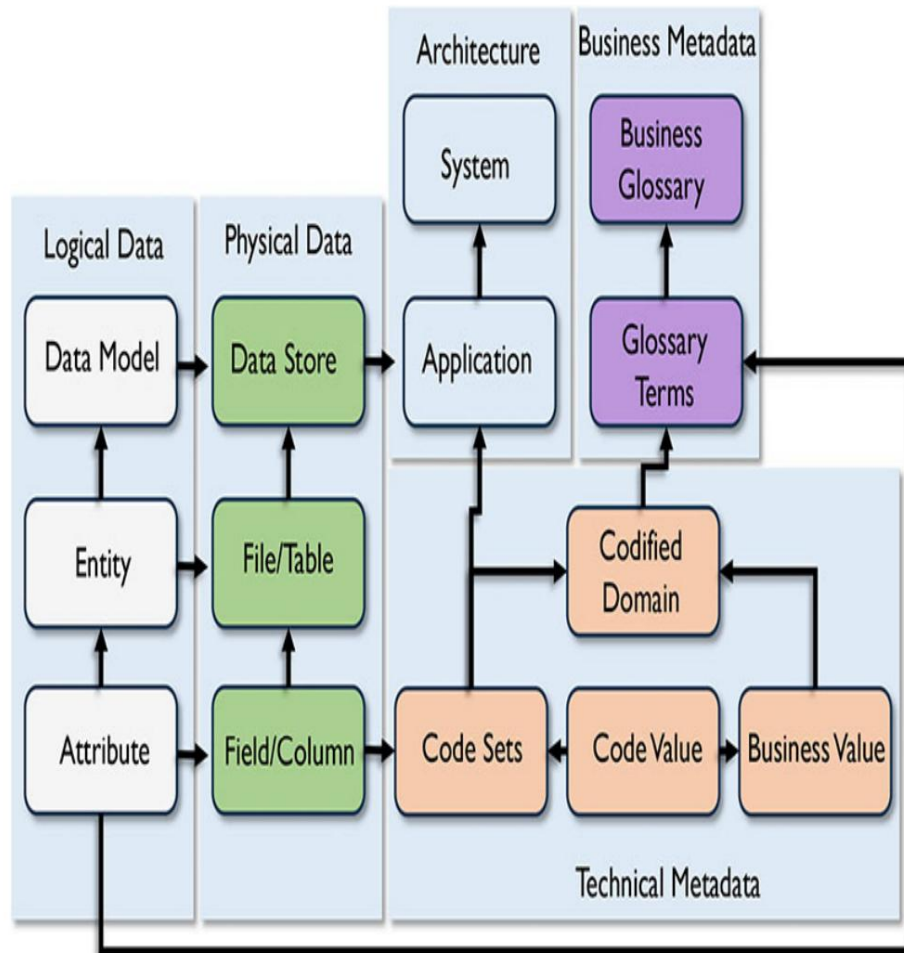


Figure 88 Example Metadata Repository Metamodel

### 2.3.2 Aplicar Padrões de Metadados

A solução de Metadados deve aderir aos padrões internos e externos acordados conforme identificados na estratégia de Metadados. Os metadados devem ser monitorados para conformidade por atividades de governança.

Os padrões de Metadados internos da organização incluem convenções de nomenclatura, atribuições personalizadas, segurança, visibilidade e documentação de processamento. Os padrões de Metadados externos da organização incluem os formatos de troca de dados e o design de interfaces de programação de aplicativos.

### 2.3.3 Gerenciar repositórios de

*metadados* Implemente atividades de controle para gerenciar o ambiente de metadados.

O controle de repositórios é o controle do movimento de Metadados e das atualizações de repositórios realizadas pelo especialista em Metadados. Essas atividades são de natureza administrativa e envolvem o monitoramento e a resposta a relatórios, avisos, logs de tarefas e a resolução de vários problemas no ambiente de repositório implementado. Muitas atividades de controle são padrão para operações de dados e manutenção de interface.

As atividades de controle devem ter supervisão de governança de dados.

As atividades de controle incluem:

- Agendamento e monitoramento de tarefas
- Carregar análise estatística
- Backup, recuperação, arquivamento, limpeza
- Modificações de configuração
- Ajuste de desempenho
- Análise de estatísticas de consulta
- Geração de consulta e relatório
- Gestão de segurança
- As atividades de controle de qualidade incluem:
- Garantia de qualidade, controle de qualidade
- Frequência de atualização de dados – correspondência de conjuntos com períodos de tempo
- Relatórios de metadados ausentes
- Relatório de metadados de envelhecimento
- As atividades de gerenciamento de metadados incluem:
- Carregando, digitalizando, importando e marcando ativos
- Mapeamento e movimentação de fontes
- Controle de versão
- Gerenciamento da interface do usuário
- Vinculando conjuntos de dados Manutenção de metadados – para provisionamento NOSQL
- Vinculando dados à aquisição interna de dados – links personalizados e

Licenciamento de

- metadados de trabalho para fontes de dados externas e feeds
- Aprimoramento de dados Metadados, por exemplo, Link para GIS
- E treinamento, incluindo: Educação
- e treinamento de usuários e administradores de dados Geração e
- análise de métricas de gerenciamento Treinamento nas
- atividades de controle e consulta e relatórios

## 2.4 Criar e manter metadados

Conforme descrito na [Seção 1.3.5](#), os Metadados são criados por meio de uma série de processos e armazenados em muitos lugares dentro de uma organização. Para serem de alta qualidade, os Metadados devem ser gerenciados como um produto. Bons Metadados não são criados por acidente. Eles exigem planejamento. (Consulte [o Capítulo 13](#).)

---

Vários princípios gerais de gerenciamento de metadados descrevem os meios para gerenciar metadados para qualidade:

- **Responsabilidade:** reconheça que os metadados são frequentemente produzidos por meio de processos existentes (modelagem de dados, SDLC, definição de processos de negócios) e responsabilize os proprietários dos processos pela qualidade dos metadados.
- **Padrões:** defina, aplique e audite padrões para metadados para simplificar a integração e permitir o uso.
- **Melhoria:** Crie um mecanismo de feedback para que os consumidores possam informar a equipe de Gerenciamento de Metadados sobre metadados incorretos ou desatualizados.

Assim como outros dados, os metadados podem ser perfilados e inspecionados quanto à qualidade. Sua manutenção deve ser programada ou concluída como parte auditável do trabalho do projeto.

#### 2.4.1 Integrar Metadados Os

processos de integração reúnem e consolidam Metadados de toda a empresa, incluindo Metadados de dados adquiridos fora da empresa. O repositório de Metadados deve integrar Metadados técnicos extraídos com Metadados relevantes de negócios, processos e administração. Os Metadados podem ser extraídos usando adaptadores, scanners, aplicativos de ponte ou acessando diretamente os Metadados em um armazenamento de dados de origem. Os adaptadores estão disponíveis com muitas ferramentas de software de fornecedores terceirizados, bem como de ferramentas de integração de Metadados. Em alguns casos, os adaptadores serão desenvolvidos usando as APIs da ferramenta.

Surtem desafios na integração que exigirão governança.

A integração de conjuntos de dados internos, dados externos, como estatísticas governamentais, e dados provenientes de formulários não eletrônicos, como white papers, artigos em revistas ou relatórios, pode levantar inúmeras questões sobre qualidade e semântica.

Realize a varredura do repositório em duas abordagens distintas.

- **Interface proprietária:** Em um processo de escaneamento e carregamento de etapa única, um scanner coleta os Metadados de um sistema de origem e, em seguida, chama diretamente o componente carregador específico do formato para carregar os Metadados no repositório. Nesse processo, não há saída de arquivo específica do formato e a coleta e o carregamento dos Metadados ocorrem em uma única etapa.
- **Interface semiproprietária:** Em um processo de duas etapas, um scanner coleta os Metadados de um sistema de origem e os gera em um arquivo de dados de formato específico. O scanner produz apenas um arquivo de dados que o repositório receptor precisa ser capaz de ler e carregar adequadamente. A interface é uma arquitetura mais aberta, pois o arquivo é legível por muitos métodos.

Um processo de digitalização usa e produz vários tipos de arquivos durante o processo.

- **Arquivo de controle:** contendo a estrutura de origem do modelo de dados
- **Arquivo de reutilização:** contendo as regras para gerenciar a reutilização de cargas de processo
- **Arquivos de log:** produzidos durante cada fase do processo, um para cada varredura ou extração e um para cada ciclo de carga
- **Arquivos temporários e de backup:** Use durante o processo ou para rastreabilidade

Use uma área de preparação de Metadados não persistente para armazenar arquivos temporários e de backup. A área de preparação suporta processos de reversão e recuperação e fornece uma trilha de auditoria provisória para auxiliar os gerentes de repositório ao investigar problemas de origem ou qualidade de Metadados. A área de preparação pode assumir a forma de um diretório de arquivos ou um banco de dados.

Ferramentas de integração de dados usadas para armazenamento de dados e aplicativos de Business Intelligence são frequentemente usadas de forma eficaz em processos de integração de metadados. (Consulte [o Capítulo 8.](#))

#### *2.4.2 Distribuir e entregar metadados*

Os metadados são entregues aos consumidores de dados e aos aplicativos ou ferramentas que exigem feeds de metadados. Os mecanismos de entrega incluem:

- Sites de intranet de metadados para navegação, pesquisa, consulta, geração de relatórios e análise
- Relatórios, glossários e outros documentos Data
- warehouses, data marts e ferramentas de BI (Business Intelligence)
- Ferramentas de modelagem e desenvolvimento de software Mensagens e
- transações Serviços da Web e
- Interfaces de Programação de Aplicativos (APIs)
- Soluções de interface de organização externa (por exemplo, cadeia de suprimentos

soluções)

A solução Metadata frequentemente vincula-se a uma solução de Business Intelligence, de modo que tanto o escopo quanto a moeda dos Metadata sincronizam com o conteúdo do BI. Um link fornece um meio de integração na entrega do BI ao usuário final. Da mesma forma, algumas soluções de CRM (Customer Relationship Management) ou outras ERP (Enterprise Resource Planning) podem exigir integração de Metadata na camada de entrega do aplicativo.

Os metadados são trocados com organizações externas usando arquivos (simples, XML ou estruturados em JSON) ou por meio de serviços web.

## **2.5 Consultar, relatar e analisar metadados** Os metadados

orientam o uso de ativos de dados. Use metadados em Business Intelligence (relatórios e análises), decisões de negócios (operacionais, táticas, estratégicas) e em semântica de negócios (o que eles dizem, o que eles significam – jargão empresarial). Um repositório de metadados deve ter um aplicativo front-end que suporte a funcionalidade de busca e recuperação necessária para toda essa orientação e gerenciamento de ativos de dados. A interface fornecida aos usuários empresariais pode ter um conjunto diferente de requisitos funcionais do que para usuários técnicos e desenvolvedores. Alguns relatórios facilitam o desenvolvimento futuro, como análise de impacto de mudanças, ou solucionam problemas de definições variadas para projetos de data warehouse e Business Intelligence, como relatórios de linhagem de dados.

## **3. Ferramentas**

A ferramenta primária usada para gerenciar Metadados é o repositório de Metadados. Isso incluirá uma camada de integração e, frequentemente, uma interface para atualizações manuais. Ferramentas que produzem e usam Metadados se tornam fontes de Metadados que podem ser integradas a um repositório de Metadados.

### **3.1 Ferramentas de gerenciamento de repositório de metadados**

Ferramentas de gerenciamento de metadados fornecem recursos para gerenciar metadados em um local centralizado (repositório). Os metadados podem ser inseridos manualmente ou extraídos de várias outras fontes por meio de conectores especializados. Repositórios de metadados também fornecem recursos para trocar metadados com outros sistemas.

Ferramentas de gerenciamento de metadados e repositórios em si também são uma fonte de Metadados, especialmente em um modelo de arquitetura de Metadados híbrido ou em grandes implementações empresariais. Ferramentas de gerenciamento de metadados permitem a troca dos Metadados coletados com outros repositórios de Metadados, permitindo a coleta de vários e diversos Metadados de diferentes fontes em um repositório centralizado, ou permitindo o enriquecimento e a padronização dos diversos Metadados conforme eles se movem entre os repositórios.

## 4. Técnicas

### 4.1 Linhagem de Dados e Análise de Impacto Um benefício

fundamental de descobrir e documentar Metadados sobre os ativos físicos é fornecer informações sobre como os dados são transformados à medida que se movem entre os sistemas. Muitas ferramentas de Metadados carregam informações sobre o que está acontecendo com os dados dentro de seus ambientes e fornecem recursos para visualizar a linhagem em todo o intervalo dos sistemas ou aplicativos com os quais fazem interface. A versão atual da linhagem com base no código de programação é chamada de 'Linhagem Conforme Implementada'. Em contraste, a linhagem descrita em documentos de especificação de mapeamento é chamada de 'Linhagem Conforme Projetado'.

As limitações de uma construção de linhagem são baseadas na cobertura do sistema de gerenciamento de metadados. Repositórios de metadados específicos de função ou ferramentas de visualização de dados têm informações sobre a linhagem de dados dentro do escopo dos ambientes com os quais interagem, mas não fornecerão visibilidade sobre o que está acontecendo com os dados fora de seus ambientes.

Os sistemas de gerenciamento de metadados importam a linhagem 'Conforme implementado' das várias ferramentas que podem fornecer esses detalhes de linhagem e, em seguida,

aumentar a linhagem de dados com o 'As Designed' dos lugares onde os detalhes reais da implementação não são extraíveis. O processo de conectar as partes da linhagem de dados é conhecido como *stitching*. Ele resulta em uma visualização holística dos dados conforme eles se movem de seus locais originais (fonte oficial ou sistema de registro) até chegarem ao seu destino final.

A Figura 89 mostra uma linhagem de elemento de dados de exemplo. Ao ler isso, o elemento de dados de negócios 'Total Backorder', que é fisicamente implementado como coluna `zz_total`, depende de 3 outros elementos de dados: 'Custo de Unidades em Centavos' fisicamente implementado como `'yy_unit_cost'`, 'Imposto no Estado de Envio' implementado em `'yy_tax'` e 'Quantidade de Pedidos em Espera' implementado em `'yy_qty'`.

Embora um gráfico de linhagem, como na Figura 89, descreva o que está acontecendo com um elemento de dados específico, nem todos os usuários empresariais o entenderão. Níveis mais altos de linhagem (por exemplo, 'Linhagem do Sistema') resumem o movimento no nível do sistema ou do aplicativo. Muitas ferramentas de visualização fornecem capacidade de zoom-in/zoom-out, para mostrar a linhagem do elemento de dados no contexto da linhagem do sistema. Por exemplo, a Figura 90 mostra uma linhagem de sistema de amostra, onde, à primeira vista, o movimento geral de dados é compreendido e visualizado em um nível de sistema ou aplicativo.

Figura 89 Dados de amostra

Diagrama de fluxo de linhagem de elementos



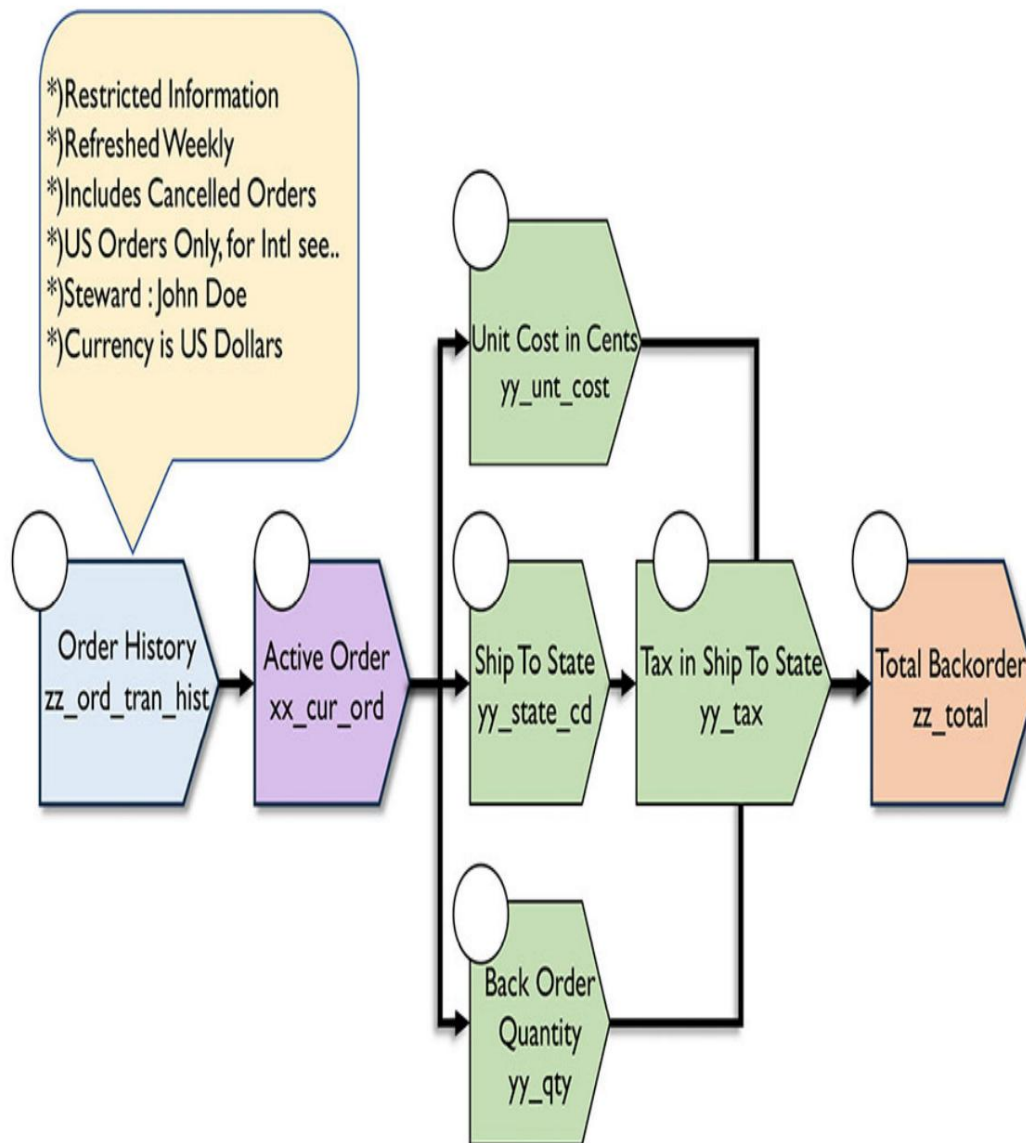


Figure 89 Sample Data Element Lineage Flow Diagram

Figura 90 Sistema de Amostra  
Diagrama de fluxo de linhagem

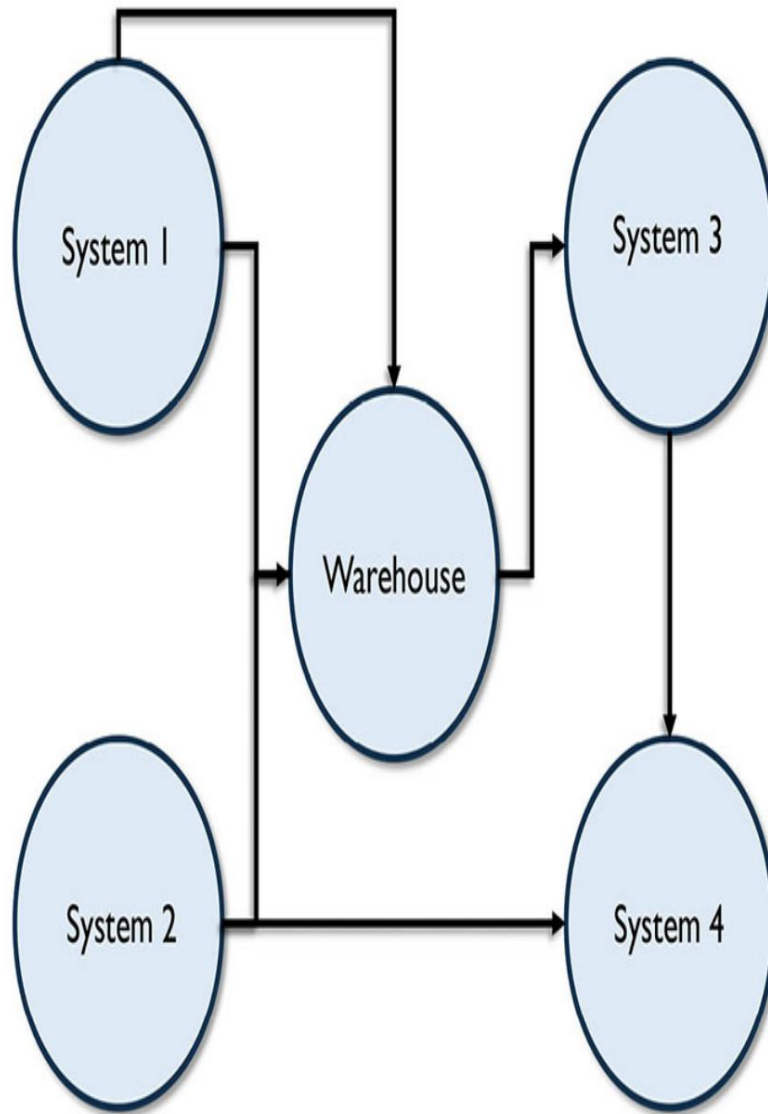


Figure 90 Sample System Lineage Flow Diagram

À medida que o número de elementos de dados em um sistema cresce, a descoberta de linhagem se torna complexa e difícil de gerenciar. Para atingir com sucesso os objetivos de negócios, uma estratégia para descobrir e importar ativos para o repositório de metadados requer planejamento e design. A descoberta de linhagem bem-sucedida precisa levar em conta o foco comercial e técnico:

- **Foco do negócio:** Limite a descoberta de linhagem a elementos de dados priorizados pelo negócio. Comece pelo alvo

localizações e rastrear de volta aos sistemas de origem onde os dados específicos se originam. Ao limitar os ativos escaneados àqueles que movem, transferem ou atualizam os elementos de dados selecionados, essa abordagem permitirá que os consumidores de dados empresariais entendam o que está acontecendo com o elemento de dados específico conforme ele se move pelos sistemas. Se acoplado com medições de qualidade de dados, a linhagem pode ser usada para identificar onde o design do sistema impacta negativamente a qualidade dos dados.

- **Foco técnico:** comece nos sistemas de origem e identifique todos os consumidores imediatos, depois identifique todos os consumidores subsequentes do primeiro conjunto identificado e continue repetindo essas etapas até que todos os sistemas sejam identificados. Usuários de tecnologia se beneficiam mais da estratégia de descoberta de sistema para ajudar a responder às várias perguntas sobre os dados. Essa abordagem permitirá que usuários de tecnologia e negócios respondam a perguntas sobre a descoberta de elementos de dados em toda a empresa, como "Onde está o número da previdência social?" ou gerem relatórios de impacto como "Quais sistemas são impactados se a largura de uma coluna específica for alterada?" Essa estratégia pode, no entanto, ser complexa de gerenciar.

Muitas ferramentas de integração de dados oferecem análise de linhagem que considera não apenas o código populacional desenvolvido, mas também o modelo de dados e o banco de dados físico. Algumas oferecem interfaces da web voltadas para o usuário empresarial para monitorar e atualizar definições. Elas começam a parecer glossários empresariais.

A linhagem documentada ajuda tanto as pessoas de negócios quanto as técnicas a usar dados. Sem ela, muito tempo é desperdiçado na investigação de anomalias, potenciais impactos de mudanças ou resultados desconhecidos. Procure implementar uma ferramenta integrada de impacto e linhagem que possa entender todas as partes móveis envolvidas no processo de carga, bem como relatórios e análises do usuário final. Os relatórios de impacto descrevem quais componentes são afetados por um

potencial mudança agilizando e simplificando tarefas de estimativa e manutenção.

#### **4.2 Metadados para Ingestão de Big Data Muitos**

profissionais de gerenciamento de dados estão familiarizados e confortáveis com armazenamentos de dados estruturados, onde cada item pode ser claramente identificado e marcado. Hoje em dia, porém, muitos dados vêm em formatos menos estruturados. Algumas fontes não estruturadas serão internas à organização e algumas serão externas.

Em ambos os casos, não há mais necessidade de trazer fisicamente os dados para um lugar. Por meio das novas tecnologias, o programa irá até os dados em vez de mover os dados para o programa, reduzindo a quantidade de movimentação de dados e acelerando a execução do processo.

No entanto, o gerenciamento bem-sucedido de dados em um data lake depende do gerenciamento de metadados.

As tags de metadados devem ser aplicadas aos dados na ingestão. Os metadados podem ser usados para identificar o conteúdo dos dados disponível para acesso no data lake. Muitos mecanismos de ingestão criam perfis de dados à medida que são ingeridos. O perfil de dados pode identificar domínios de dados, relacionamentos e problemas de qualidade de dados. Ele também pode habilitar a marcação. Na ingestão, as tags de metadados podem ser adicionadas para identificar dados confidenciais ou privados (como informações de identificação pessoal - PPI), por exemplo. Os cientistas de dados podem adicionar confiança, identificadores textuais e códigos que representam clusters de comportamento. ([Consulte o Capítulo 14.](#))

### **5. Diretrizes de implementação**

Implemente um ambiente de Metadados gerenciado em etapas incrementais para minimizar riscos à organização e facilitar a aceitação. Implemente repositórios de Metadados usando uma plataforma de banco de dados relacional aberta. Isso permite o desenvolvimento e a implementação de vários controles e interfaces que podem não ser previstos no início de um projeto de desenvolvimento de repositório.

O conteúdo do repositório deve ser genérico em design, não refletindo meramente os designs do banco de dados do sistema de origem. Conteúdos de design em

alinhamento com os especialistas da área de assunto da empresa e com base em um modelo de Metadados abrangente. O planejamento deve levar em conta a integração de Metadados para que os consumidores de dados possam ver em diferentes fontes de dados. A capacidade de fazer isso será um dos recursos mais valiosos do repositório. Ele deve abrigar versões atuais, planejadas e históricas dos Metadados.

Frequentemente, a primeira implementação é um piloto para provar conceitos e aprender sobre o gerenciamento do ambiente de Metadados. A integração de projetos de Metadados na metodologia de desenvolvimento de TI é necessária. Haverá variações dependendo da arquitetura e dos tipos de armazenamento.

## **5.1 Avaliação de prontidão / Avaliação de risco**

Ter uma estratégia sólida de Metadados ajuda todos a tomar decisões mais eficazes. Primeiro e mais importante, as pessoas devem estar cientes dos riscos de não gerenciar Metadados. Avalie o grau em que a falta de Metadados de alta qualidade pode resultar em:

- Erros de julgamento devido a suposições incorretas, incompletas ou inválidas ou falta de conhecimento sobre o contexto dos dados
- Exposição de dados sensíveis, que podem colocar clientes ou funcionários em risco, ou impactar a credibilidade do negócio e levar a despesas legais
- Risco de que o pequeno grupo de PME que conhece os dados saia e leve consigo o seu conhecimento

O risco é reduzido quando uma organização adota uma estratégia sólida de Metadados. A prontidão organizacional é abordada por uma avaliação formal da maturidade atual nas atividades de Metadados. A avaliação deve incluir os elementos críticos de dados de negócios, glossários de Metadados disponíveis, linhagem, criação de perfil de dados e processos de qualidade de dados, maturidade de MDM (Master Data Management) e outros aspectos. As descobertas da avaliação, alinhadas com os negócios

prioridades, fornecerão a base para uma abordagem estratégica para melhoria das práticas de Gerenciamento de Metadados. Uma avaliação formal também fornece a base para um caso de negócios, patrocínio e financiamento.

A estratégia de Metadados pode ser parte de uma estratégia geral de governança de dados ou pode ser o primeiro passo na implementação de uma governança de dados eficaz. Uma avaliação de Metadados deve ser conduzida por meio de inspeção objetiva de Metadados existentes, juntamente com entrevistas com as principais partes interessadas. Os resultados de uma avaliação de risco incluem uma estratégia e um roteiro.

## **5.2 Mudança Organizacional e Cultural** Assim como outros

esforços de gerenciamento de dados, as iniciativas de Metadados frequentemente encontram resistência cultural. Passar de um ambiente de Metadados não gerenciado para um gerenciado exige trabalho e disciplina. Não é fácil de fazer, mesmo que a maioria das pessoas reconheça o valor de Metadados confiáveis. A prontidão organizacional é uma grande preocupação, assim como os métodos de governança e controle.

O gerenciamento de metadados é uma baixa prioridade em muitas organizações. Um conjunto essencial de metadados precisa de coordenação e comprometimento em uma organização. Podem ser estruturas de dados de identificação de funcionários, números de apólices de seguro, números de identificação de veículos ou especificações de produtos, que, se alterados, exigiriam grandes revisões de muitos sistemas empresariais. Procure aquele bom exemplo em que o controle colherá benefícios imediatos de qualidade para os dados na empresa. Construa o argumento a partir de exemplos concretos relevantes para os negócios.

A implementação de uma estratégia de governança de dados corporativos precisa de suporte e engajamento da alta gerência. Ela requer que as equipes de negócios e tecnologia sejam capazes de trabalhar em conjunto de forma multifuncional.

## **6. Governança de Metadados**

---

As organizações devem determinar seus requisitos específicos para o gerenciamento do ciclo de vida dos metadados e estabelecer governança

processos para habilitar esses requisitos. É recomendado que funções e responsabilidades formais sejam atribuídas a recursos dedicados, especialmente em áreas grandes ou críticas para os negócios. Os próprios processos de governança de metadados dependem de metadados confiáveis, então a equipe encarregada de gerenciar metadados pode testar princípios nos metadados que eles criam e usam.

## 6.1 Controles de Processo

A equipe de governança de dados deve ser responsável por definir os padrões e gerenciar as mudanças de status dos metadados – geralmente com software de fluxo de trabalho ou colaboração – e pode ser responsável por atividades promocionais e desenvolvimento de treinamento ou treinamento real em toda a organização.

Uma governança de metadados mais madura exigirá que termos e definições de negócios progridam por meio de várias mudanças de status ou portões de governança; por exemplo, de um termo candidato para aprovado, para publicado e para um ponto final no ciclo de vida de substituído ou aposentado. A equipe de governança também pode gerenciar associações de termos de negócios, como termos relacionados, bem como a categorização e o agrupamento de

termos.

A integração da estratégia de Metadados no SDLC é necessária para garantir que os Metadados alterados sejam coletados quando forem alterados. Isso ajuda a garantir que os Metadados permaneçam atualizados.

## 6.2 Documentação de Soluções de Metadados

Um catálogo mestre de Metadados incluirá as fontes e alvos atualmente em escopo. Este é um recurso para usuários de TI e negócios e pode ser publicado para a comunidade de usuários como um guia para "o que está onde" e para definir expectativas sobre o que eles encontrarão:

- Status de implementação de metadados
- Armazenamento de metadados de origem e destino
- Informações de programação para atualizações
- Retenção e versões mantidas

- Conteúdo
- Declarações ou avisos de qualidade (por exemplo, valores ausentes)
- Sistema de registro e outros status de fonte de dados (por exemplo, cobertura do histórico de conteúdo de dados, desativação ou substituição de sinalizadores)
- Ferramentas, arquiteturas e pessoas envolvidas Informações
- confidenciais e estratégia de remoção ou mascaramento da fonte

Em documentos e gerenciamento de conteúdo, os mapas de dados mostram informações semelhantes. Visualizações do panorama geral dos sistemas de integração de Metadados também são mantidas como parte da documentação de Metadados. (Consulte [o Capítulo 9](#).)

---

## 6.3 Padrões e Diretrizes de Metadados

Os padrões de metadados são essenciais na troca de dados com parceiros comerciais operacionais. As empresas percebem o valor do compartilhamento de informações com clientes, fornecedores, parceiros e órgãos reguladores. A necessidade de compartilhar metadados comuns para dar suporte ao uso ideal de informações compartilhadas gerou muitos padrões baseados em setores.

Adote padrões de Metadados baseados na indústria e sensíveis ao setor no início do ciclo de planejamento. Use os padrões para avaliar tecnologias de Gerenciamento de Metadados. Muitos fornecedores líderes oferecem suporte a vários padrões, e alguns podem ajudar a personalizar padrões baseados na indústria e sensíveis ao setor.

Os fornecedores de ferramentas fornecem suporte a XML e JSON ou REST para trocar dados para seus produtos de gerenciamento de dados. Eles usam a mesma estratégia para vincular suas ferramentas em conjuntos de soluções. Tecnologias, incluindo integração de dados, bancos de dados relacionais e multidimensionais, gerenciamento de requisitos, relatórios de Business Intelligence, modelagem de dados e regras de negócios, oferecem recursos de importação e exportação para dados e metadados usando XML. Os fornecedores mantêm seus esquemas XML proprietários e definições de tipo de documento (DTD) ou mais



comumente as definições de esquema XML (XSD). Elas são acessadas por meio de interfaces proprietárias. O desenvolvimento personalizado é necessário para integrar essas ferramentas em um ambiente de gerenciamento de metadados.

As diretrizes incluem modelos e exemplos associados e treinamento sobre entradas e atualizações esperadas, incluindo regras como "não defina um termo usando o termo" e declarações de completude. Diferentes modelos são desenvolvidos para diferentes tipos de Metadados e são conduzidos em parte pela solução de Metadados selecionada. O monitoramento contínuo das diretrizes para eficácia e atualizações necessárias é uma responsabilidade de governança.

Os padrões ISO para metadados fornecem orientação para desenvolvedores de ferramentas, mas provavelmente não serão uma preocupação para organizações que implementam usando ferramentas comerciais, já que as ferramentas devem atender aos padrões. De qualquer forma, pode ser útil ter uma boa compreensão desses padrões e suas repercussões.

## 6.4 Métricas

É difícil medir o impacto dos Metadados sem primeiro medir o impacto da falta de Metadados. Como parte da avaliação de risco, obtenha métricas sobre a quantidade de tempo que os consumidores de dados gastam buscando informações, a fim de mostrar melhorias após a solução de Metadados ser colocada em prática. A eficácia da implementação de Metadados também pode ser medida em termos da completude dos próprios Metadados, das rotinas de gerenciamento associadas a eles e do uso de Metadados. Métricas sugeridas em ambientes de Metadados incluem:

- **Completude do repositório de metadados:** Compare a cobertura ideal dos metadados corporativos (todos os artefatos e todas as instâncias dentro do escopo) com a cobertura real. Referencie a estratégia para definições de escopo.
- **Maturidade do gerenciamento de metadados:** métricas desenvolvidas para julgar a maturidade dos metadados da empresa, com base na abordagem do Capability Maturity Model (CMM-DMM) para

avaliação de maturidade. (Ver [Capítulo 15.](#))

- **Representação de administradores:** comprometimento organizacional com metadados, conforme avaliado pela nomeação de administradores, cobertura da administração em toda a empresa e documentação das funções nas descrições de cargos.
- **Uso de metadados:** a aceitação do usuário no uso do repositório de metadados pode ser medida pelas contagens de login no repositório. A referência a Metadados por usuários na prática empresarial é uma medida mais difícil de rastrear. Medidas anedóticas em pesquisas qualitativas podem ser necessárias para capturar isso medir.
- **Atividade do Glossário de Negócios:** Uso, atualização, resolução de definições, cobertura.
- **Conformidade de dados do serviço Master Data:** Mostra a reutilização de dados em soluções SOA. Metadados nos serviços de dados auxiliam os desenvolvedores a decidir quando um novo desenvolvimento pode usar um serviço existente.
- **Qualidade da documentação de metadados:** Avalie a qualidade da documentação de metadados por meio de métodos automáticos e manuais. Os métodos automáticos incluem executar lógica de colisão em duas fontes, medindo o quanto elas correspondem e a tendência ao longo do tempo. Outra métrica mediria a porcentagem de atributos que têm definições, tendências ao longo do tempo. Os métodos manuais incluem pesquisa aleatória ou completa, com base em definições empresariais de qualidade. Medidas de qualidade indicam a integridade, confiabilidade, atualidade, etc., dos metadados no repositório.
- **Disponibilidade do repositório de metadados:** tempo de atividade, tempo de processamento (lote e consulta).

## 7. Trabalhos Citados / Recomendados

Aiken, Peter. *Engenharia reversa de dados: matando o dragão legado.*