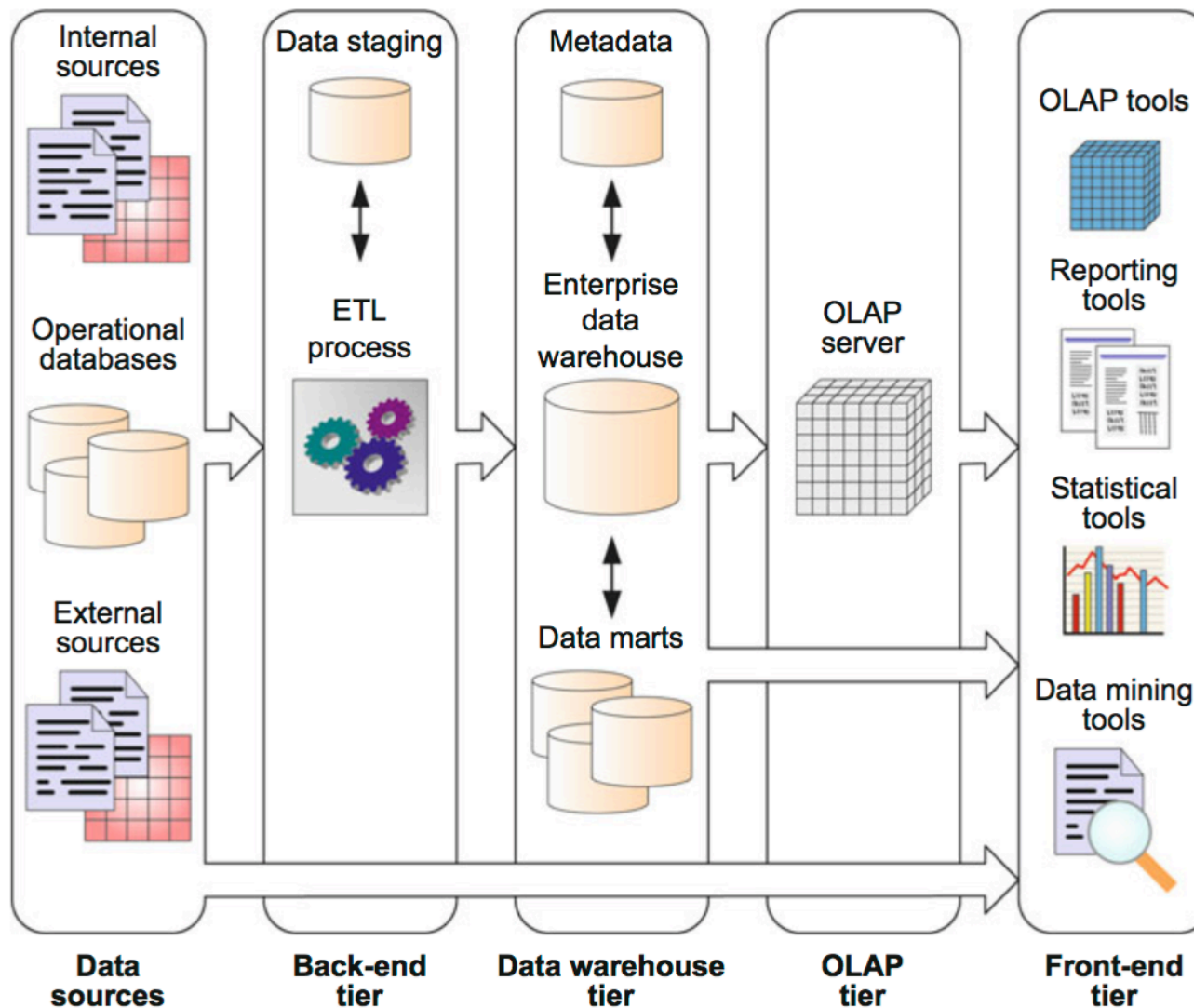


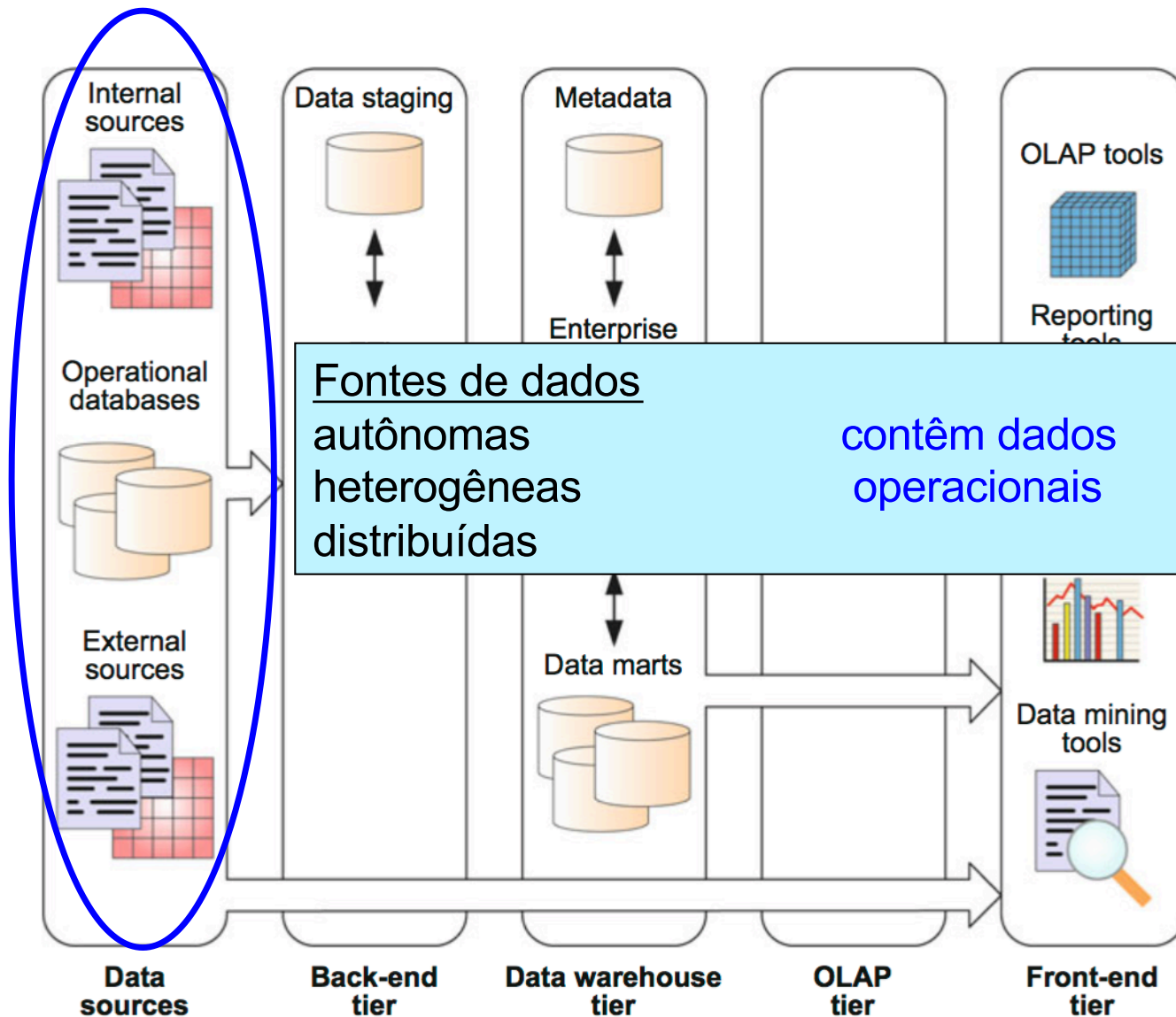
Arquitetura de um Ambiente de Data Warehousing

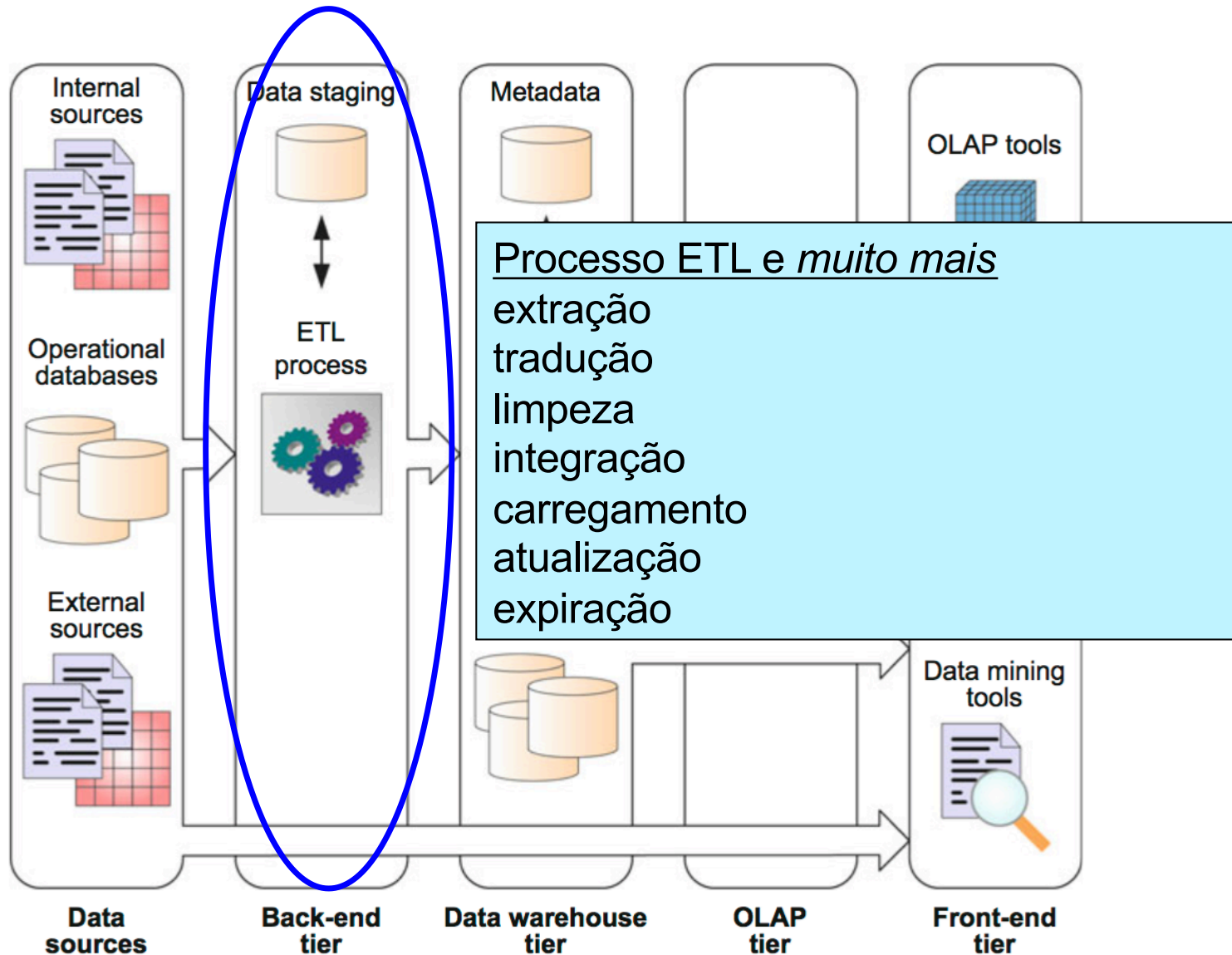
Processamento Analítico de Dados
(em Larga Escala)

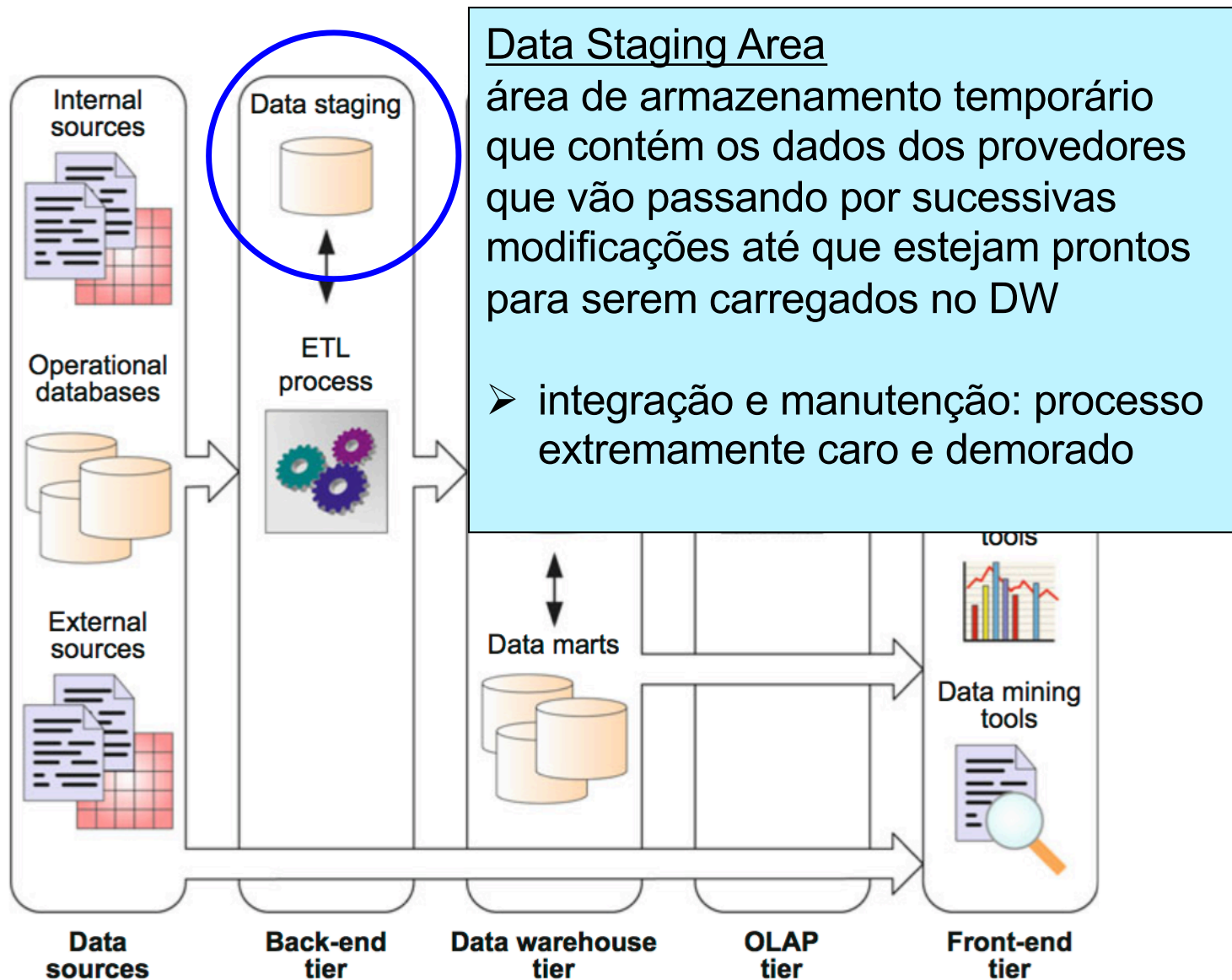
Profa. Dra. Cristina Dutra de Aguiar Ciferri



Fonte: Vaisman, A., Zimányi, E. Data Warehouse Systems: Design and Implementation. Springer, 2014.

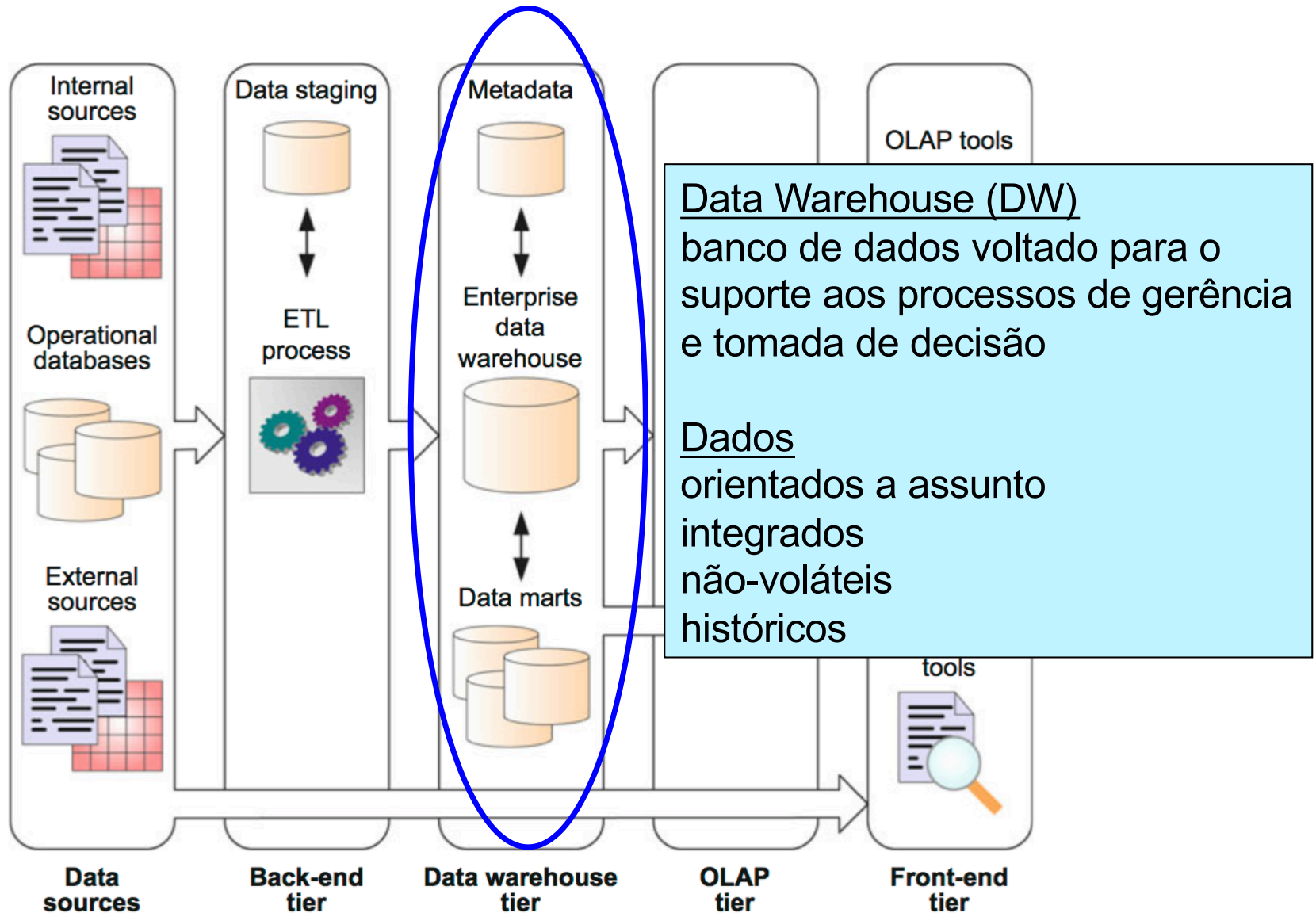






Data Staging Area
 área de armazenamento temporário que contém os dados dos provedores que vão passando por sucessivas modificações até que estejam prontos para serem carregados no DW

- integração e manutenção: processo extremamente caro e demorado

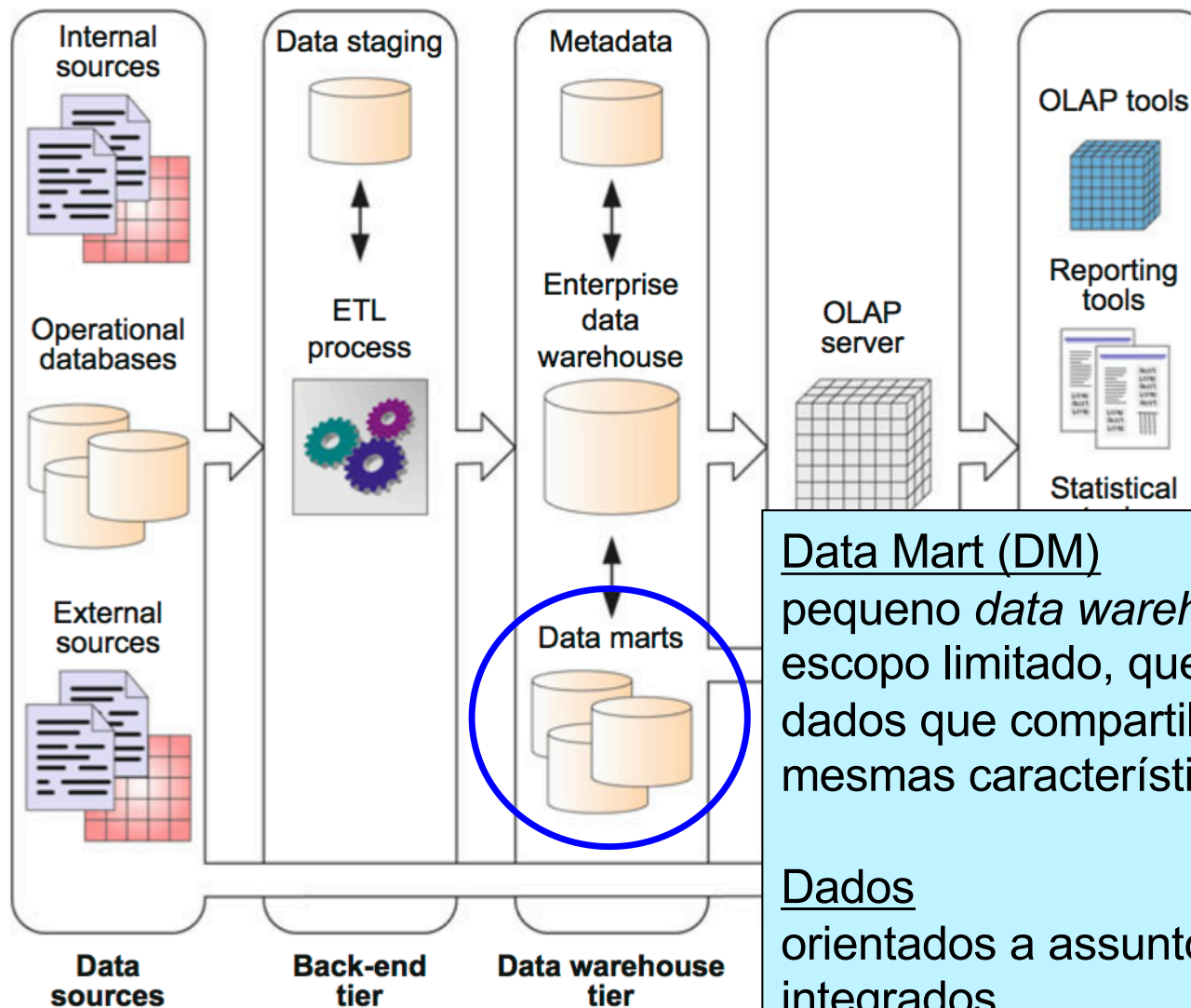


Características dos Dados

- Orientados a assunto
 - relativos aos temas de negócio de maior interesse da corporação
 - *exemplos*: clientes, produtos, promoções, contas e vendas
- Integrados
 - dados obtidos dos provedores de informação corrigidos para eliminar possíveis inconsistências

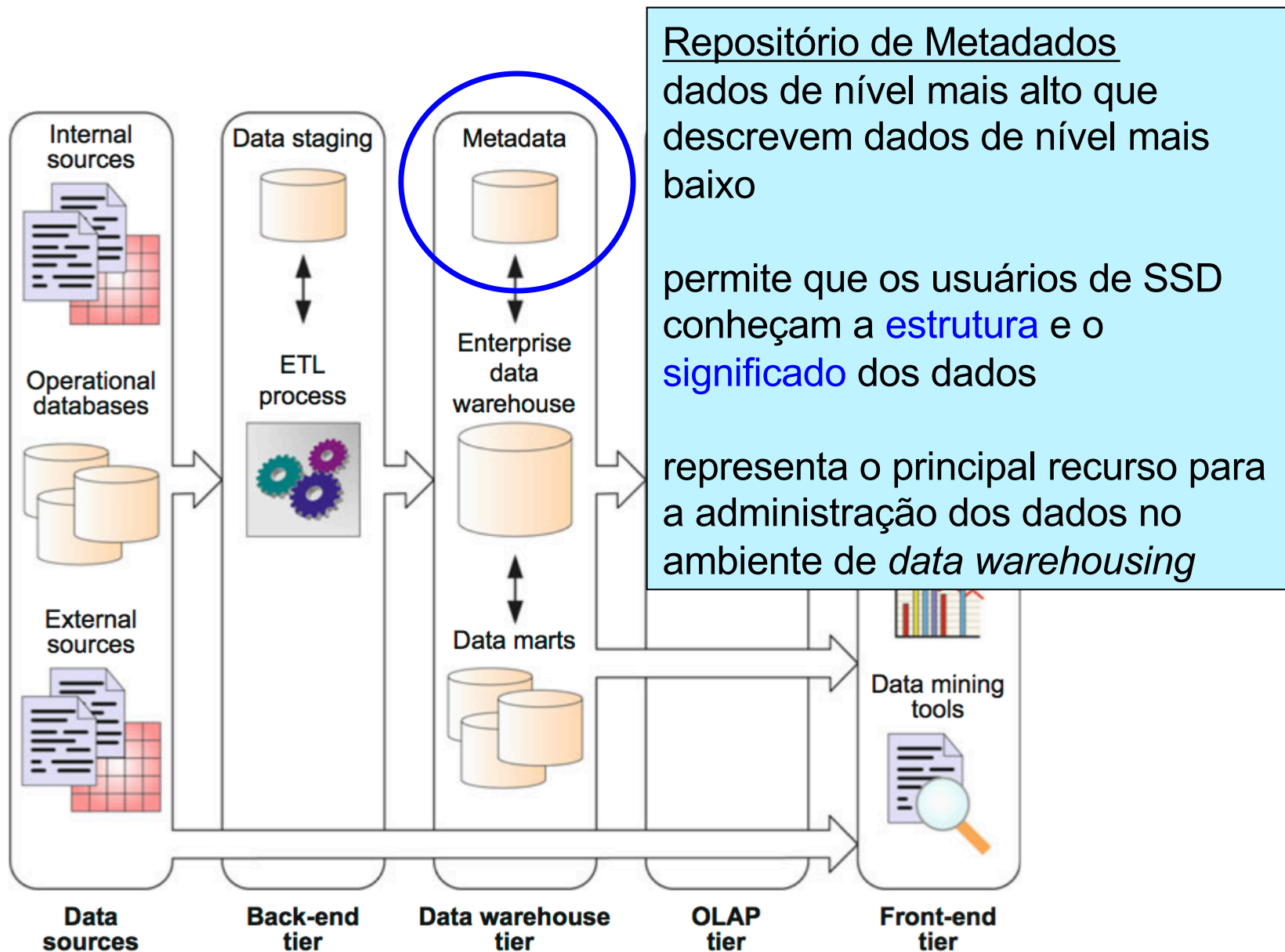
Características dos Dados

- Não-voláteis
 - o conteúdo do DW permanece estável por longos períodos de tempo
- Históricos
 - relevantes a algum período de tempo
 - *exemplo*: usualmente dados relativos a um grande espectro de tempo (5 a 10 anos) encontram-se disponíveis

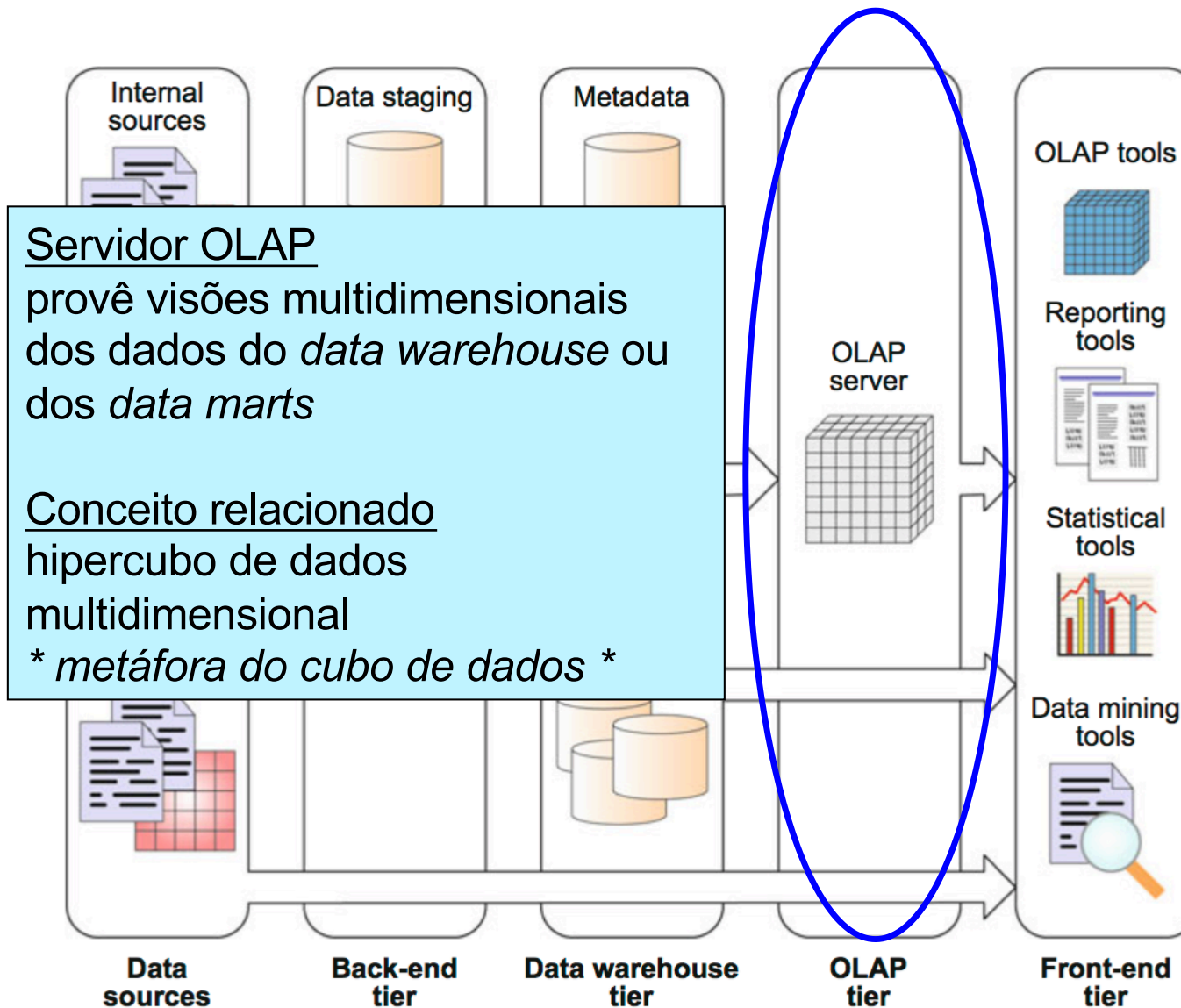


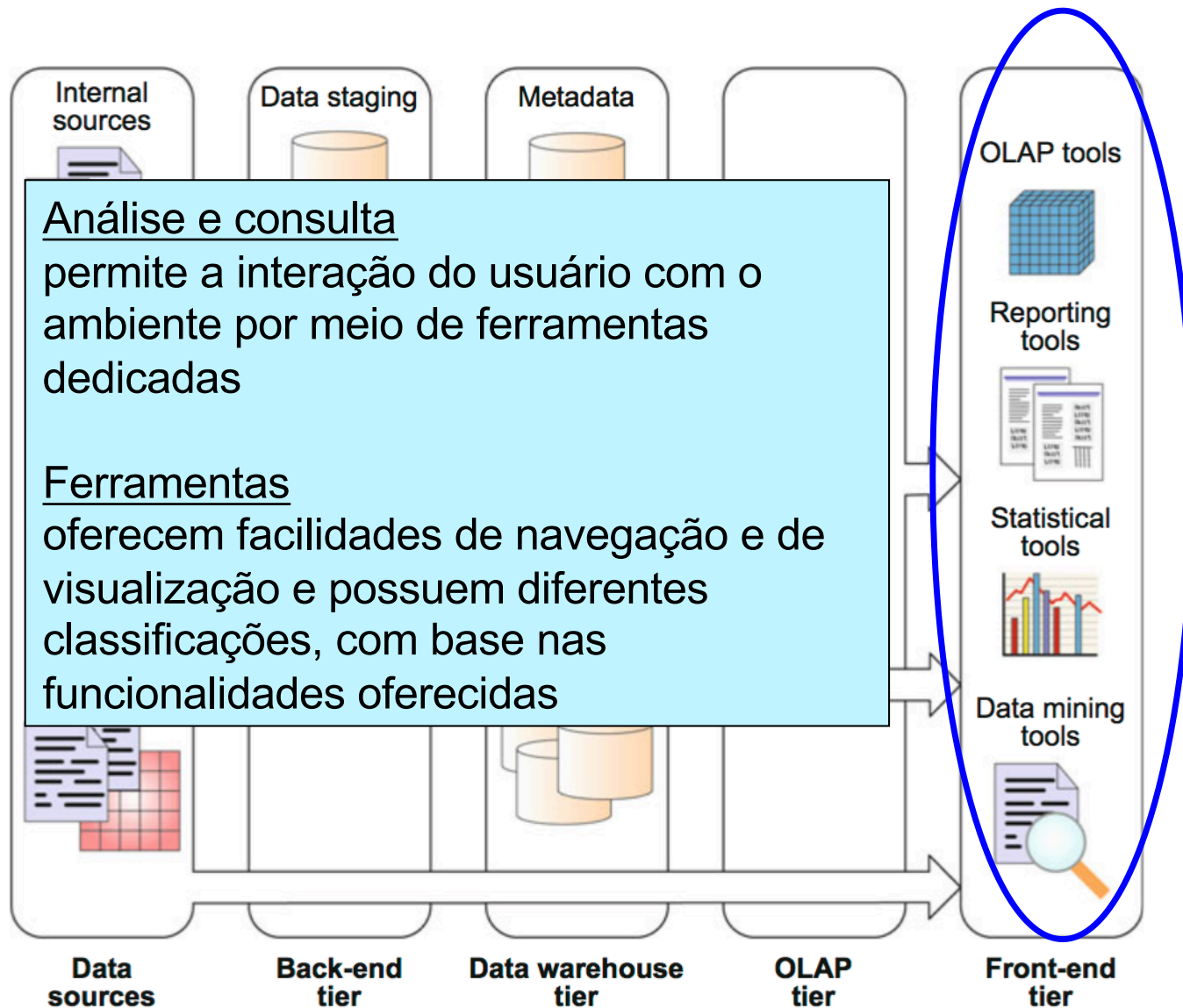
Data Mart (DM)
 pequeno *data warehouse* com escopo limitado, que armazena dados que compartilham as mesmas características

Dados
 orientados a assunto
 integrados
 não-voláteis
 históricos



Repositório de Metadados
 dados de nível mais alto que descrevem dados de nível mais baixo
 permite que os usuários de SSD conheçam a **estrutura** e o **significado** dos dados
 representa o principal recurso para a administração dos dados no ambiente de *data warehousing*





Ferramentas

front-end tier: reporting tools

- De consulta gerenciáveis e geradores de relatório
 - tipos mais simples de ferramentas
 - têm como objetivo produzir relatórios periódicos
 - permitem que os usuários realizem consultas independentemente da estrutura do banco de dados e/ou da linguagem de consulta

Ferramentas

front-end tier: statistical tools

- Estatísticas
 - analisam e visualizam os dados usando métodos estatísticos
 - oferecem visualização gráfica simplificada, por exemplo representando exceções a atividades normais de negócio ou a regras por meio de diferentes cores

Ferramentas

front-end tier: OLAP tools

- OLAP
 - oferecem capacidades analíticas sofisticadas, permitindo que os dados sejam analisados usando visões multidimensionais complexas e elaboradas
 - oferecem navegação facilitada nessas visões
 - exemplo: usuários de SSD podem analisar os dados sob diferentes perspectivas e determinar tendências por meio da navegação entre diferentes níveis de hierarquias de agregação

Ferramentas

front-end tier: data mining tools

- De mineração de dados
 - permitem que informações, padrões e tendências de negócio “escondidas” nos dados sejam descobertas

IMPORTANTE: Independentemente da ferramenta utilizada, um fator primordial refere-se à **visualização dos resultados obtidos**. Técnicas de visualização dos dados devem determinar a melhor forma de se exibir relacionamentos e padrões complexos em um monitor bidimensional, de modo que o problema inteiro e/ou a solução sejam claramente visíveis usuários de SSD