



Big Data Fundamentos 3.0

Big Data Fundamentos – Versão 3.0

Soluções Comerciais



Vemos atualmente ocorrer com as arquiteturas modernas de Big Data, o mesmo fenômeno que ocorreu cerca de 10 anos atrás com o próprio Big Data: uma explosão de soluções comerciais que facilitam a implementação das arquiteturas. Aqui vão algumas dicas.

Para o Data Lakehouse, a solução mais comum é o Databricks. A DSA foi pioneira no Brasil ao trazer projetos de Databricks em nossos treinamentos e o último curso da Formação Engenheiro de Dados terá diversos novos projetos sobre a plataforma. Conheça o Databricks:

<https://databricks.com/>

Data Mesh é um conceito de arquitetura e sua implementação pode envolver diversas tecnologias. Considerando o ambiente de Cloud Computing do Microsoft Azure, o *Azure Purview* seria seu ponto de partida para descobrir dados. Se você precisar fazer consultas entre domínios, também chamadas de consultas federadas, use o Synapse sem servidor com o *Azure Virtual Network Peering* se estiver consultando dados de contas de armazenamento (vinculando as contas de armazenamento em cada espaço de trabalho do Synapse). Se consultar dados de pools dedicados relacionais do Synapse, isso exigiria atualmente trabalho extra, como usar notebooks Synapse Spark, Databricks, Power BI ou fluxos de dados do *Azure Data Factory* para chamar vários bancos de dados hospedados em pools dedicados separados (mas há soluções mais fáceis).

As Formações Engenheiro de Dados e Arquiteto de Dados fornecem o conhecimento necessário para essas implementações.