

Kiến thức dữ liệu cho bài học khoa học dữ liệu

Chào mừng! Bảng thuật ngữ được sắp xếp theo thứ tự bảng chữ cái này chứa nhiều thuật ngữ trong khóa học này. Những điều khoản này rất quan trọng để bạn nhận ra khi làm việc trong ngành, tham gia các nhóm người dùng và tham gia các chương trình chứng chỉ khác.

Thuật ngữ	Sự định nghĩa	Video giới thiệu thuật ngữ này
Tuân thủ ACID	Đảm bảo tính chính xác và nhất quán của dữ liệu thông qua Nguyên tử, Tính nhất quán, Cách ly và Độ bền (ACID) trong các giao dịch cơ sở dữ liệu.	Hệ thống quản lý cơ sở dữ liệu quan hệ
Nền tảng tích hợp dựa trên đám mây dưới dạng dịch vụ (iPaaS)	Nền tảng tích hợp được lưu trữ trên đám mây cung cấp dịch vụ tích hợp thông qua đám mây riêng ảo hoặc mô hình đám mây lai, mang lại khả năng mở rộng và tính linh hoạt.	Nền tảng tích hợp dữ liệu
Cơ sở dữ liệu dựa trên cột	Một loại cơ sở dữ liệu NoSQL tổ chức dữ liệu trong các ô được nhóm thành cột, thường được sử dụng cho các hệ thống yêu cầu khối lượng yêu cầu ghi cao và lưu trữ dữ liệu chuỗi thời gian hoặc dữ liệu IoT.	NoSQL
Dữ liệu ở trạng thái nghỉ	Dữ liệu được lưu trữ và không chuyển động, thường nằm trong cơ sở dữ liệu hoặc hệ thống lưu trữ cho nhiều mục đích khác nhau, bao gồm cả sao lưu.	Những cân nhắc khi lựa chọn kho lưu trữ dữ liệu
Tích hợp dữ liệu	Một nguyên tắc liên quan đến thực tiễn, kỹ thuật kiến trúc và công cụ cho phép các tổ chức tiếp thu, chuyển đổi, kết hợp và cung cấp dữ liệu trên nhiều loại dữ liệu khác nhau, được sử dụng cho các mục đích như tính nhất quán của dữ liệu, quản lý dữ liệu chủ, chia sẻ dữ liệu và di chuyển dữ liệu.	Nền tảng tích hợp dữ liệu
Hồ dữ liệu	Kho lưu trữ dữ liệu để lưu trữ khối lượng lớn dữ liệu có cấu trúc, bán cấu trúc và không cấu trúc ở định dạng gốc, tạo điều kiện thuận lợi cho việc khám phá và phân tích dữ liệu linh hoạt.	Siêu thị dữ liệu, Hồ dữ liệu, ETL và Đường ống dữ liệu
Dữ liệu Mart	Một tập hợp con của kho dữ liệu được thiết kế cho các chức năng kinh doanh hoặc cộng đồng người dùng cụ thể, cung cấp hiệu suất và bảo mật riêng biệt cho các phân tích tập trung.	Siêu thị dữ liệu, Hồ dữ liệu, ETL và Đường ống dữ liệu
Đường dẫn dữ liệu	Một quy trình di chuyển dữ liệu toàn diện bao gồm toàn bộ hành trình của dữ liệu từ hệ thống nguồn đến hệ thống đích, bao gồm tích hợp dữ liệu như một thành phần chính.	Nền tảng tích hợp dữ liệu
Kho dữ liệu	Thuật ngữ chung đề cập đến dữ liệu đã được thu thập, sắp xếp và tách biệt cho hoạt động kinh doanh hoặc phân tích dữ liệu. Nó có thể bao gồm cơ sở dữ liệu, kho dữ liệu và kho dữ liệu lớn.	Thu thập và tổ chức dữ liệu
Kho dữ liệu	Kho lưu trữ trung tâm hợp nhất dữ liệu từ nhiều nguồn khác nhau thông qua quy trình Trích xuất, Chuyển đổi và Tải (ETL), giúp dữ liệu có thể truy cập được để phân tích và thông tin kinh doanh.	Thu thập và tổ chức dữ liệu
Cơ sở dữ liệu dựa trên tài liệu	Một loại cơ sở dữ liệu NoSQL lưu trữ từng bản ghi và dữ liệu liên quan của nó trong một tài liệu duy nhất, cho phép lập chỉ mục, truy vấn đặc biệt và phân tích linh hoạt trên các bộ sưu tập tài liệu.	NoSQL
quá trình ETL	Quá trình Trích xuất, Chuyển đổi và Tải để tích hợp dữ liệu bao gồm việc trích xuất dữ liệu từ nhiều nguồn khác nhau, chuyển đổi dữ liệu sang định dạng có thể sử dụng được và tải dữ liệu vào kho lưu trữ.	Siêu thị dữ liệu, Hồ dữ liệu, ETL và Đường ống dữ liệu

Thuật ngữ	Sự định nghĩa	Video giới thiệu thuật ngữ này
Cơ sở dữ liệu dựa trên đồ thị	Một loại cơ sở dữ liệu NoSQL sử dụng mô hình đồ họa để biểu diễn và lưu trữ dữ liệu, lý tưởng để trực quan hóa, phân tích và khám phá các kết nối giữa các điểm dữ liệu được kết nối với nhau.	NoSQL
Kho lưu trữ khóa-giá trị	Một loại cơ sở dữ liệu NoSQL trong đó dữ liệu được lưu trữ dưới dạng cặp khóa-giá trị, với khóa đóng vai trò là mã định danh duy nhất và giá trị chứa dữ liệu, có thể đơn giản hoặc phức tạp.	NoSQL
Tính di động	Khả năng sử dụng các công cụ tích hợp dữ liệu trong nhiều môi trường khác nhau, bao gồm các kịch bản đám mây đơn, nhiều đám mây hoặc đám mây lai, mang lại sự linh hoạt trong các tùy chọn triển khai.	Nền tảng tích hợp dữ liệu
Đầu nối dựng sẵn	Bộ kết nối và bộ điều hợp được phân loại giúp đơn giản hóa việc kết nối và xây dựng các luồng tích hợp với các nguồn dữ liệu đa dạng như cơ sở dữ liệu, tệp phẳng, phương tiện truyền thông xã hội, API, ứng dụng CRM và ERP.	Nền tảng tích hợp dữ liệu
Cơ sở dữ liệu quan hệ (RDBMS)	Cơ sở dữ liệu sắp xếp dữ liệu thành định dạng bảng với các hàng và cột, tuân theo cấu trúc và lược đồ được xác định rõ ràng.	Thu thập và tổ chức dữ liệu
Khả năng mở rộng	Khả năng kho lưu trữ dữ liệu phát triển và mở rộng khả năng của nó để xử lý khối lượng dữ liệu và nhu cầu khối lượng công việc ngày càng tăng theo thời gian.	Những cân nhắc khi lựa chọn kho lưu trữ dữ liệu
Lược đồ	Cấu trúc được xác định trước mô tả tổ chức và định dạng dữ liệu trong cơ sở dữ liệu, chỉ ra các loại dữ liệu được phép và mối quan hệ của chúng.	Những cân nhắc khi lựa chọn kho lưu trữ dữ liệu
Truyền dữ liệu	Dữ liệu được tạo và truyền liên tục trong thời gian thực đòi hỏi phải có cách xử lý và xử lý chuyên biệt để thu thập và phân tích.	Những cân nhắc khi lựa chọn kho lưu trữ dữ liệu
Các trường hợp sử dụng cho cơ sở dữ liệu quan hệ	Các ứng dụng như Xử lý giao dịch trực tuyến (OLTP), Kho dữ liệu (OLAP) và các giải pháp IoT nơi cơ sở dữ liệu quan hệ vượt trội.	Hệ thống quản lý cơ sở dữ liệu quan hệ
Khóa nhà cung cấp	Tình huống trong đó người dùng trở nên phụ thuộc vào công nghệ và giải pháp của một nhà cung cấp cụ thể, khiến việc chuyển sang các nền tảng khác trở nên khó khăn.	Nền tảng tích hợp dữ liệu