

Dữ liệu lớn và Data engineering

TÌM HIỂU VỀ CÔNG VIỆC CỦA KỸ SƯ DỮ LIỆU

V.A
Phát triển nội dung

Giới thiệu về khóa học

- Giới thiệu về các khái niệm
- Không yêu cầu viết code
- Mục tiêu

Trao đổi được với kỹ sư dữ liệu

Cung cấp các nền tảng kiến thức cơ bản

Chương 1

Ngành kỹ sư dữ liệu làm gì?

1. Dữ liệu lớn và ngành kỹ sư dữ liệu
2. Sự khác biệt giữa kỹ sư dữ liệu và chuyên viên khoa học dữ liệu (KHDL)
3. Data pipelines

Chương 2

Cách lưu trữ dữ liệu

1. Dữ liệu có cấu trúc v/s dữ liệu không cấu trúc
2. SQL
3. Kho dữ liệu và hồ dữ liệu

**Cụm từ tiếng anh: Structured data, unstructured data, data warehouse, data lake*

Chương 3

Cách di chuyển và xử lý dữ liệu

1. Xử lý dữ liệu
2. Lên lịch
3. Tính toán song song
4. Điện toán đám mây

**Cụm từ tiếng anh: parallel computing, cloud computing*

Thực hành

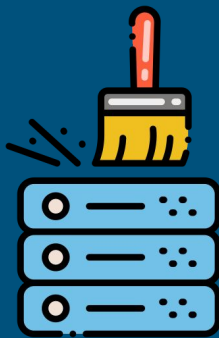


ZinMP3

Data workflow



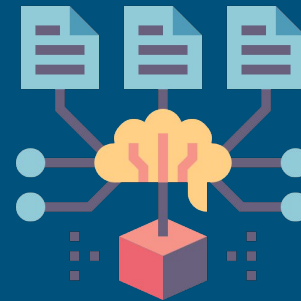
Thu thập và
Lưu trữ dữ liệu



Làm sạch dữ liệu



Khám và
Trực quan hoá dữ liệu



Thực nghiệm và
Dự đoán

Kỹ sư dữ liệu

Kỹ sư dữ liệu đưa dữ liệu:

- ✓ Đúng
- ✓ Chuẩn form
- ✓ Cho đúng người
- ✓ Một cách hiệu quả nhất có thể

Nhiệm vụ của Kỹ sư dữ liệu

- ✓ Tiêu thụ/thu thập dữ liệu từ nhiều nguồn
- ✓ Tối ưu hoá cơ sở dữ liệu (CSDL) cho việc phân tích dữ liệu
- ✓ Loại bỏ những dữ liệu gián đoạn
- ✓ Phát triển, lên cấu trúc, kiểm tra và bảo trì kiến trúc của CSDL

Dữ liệu lớn và kỹ sư dữ liệu

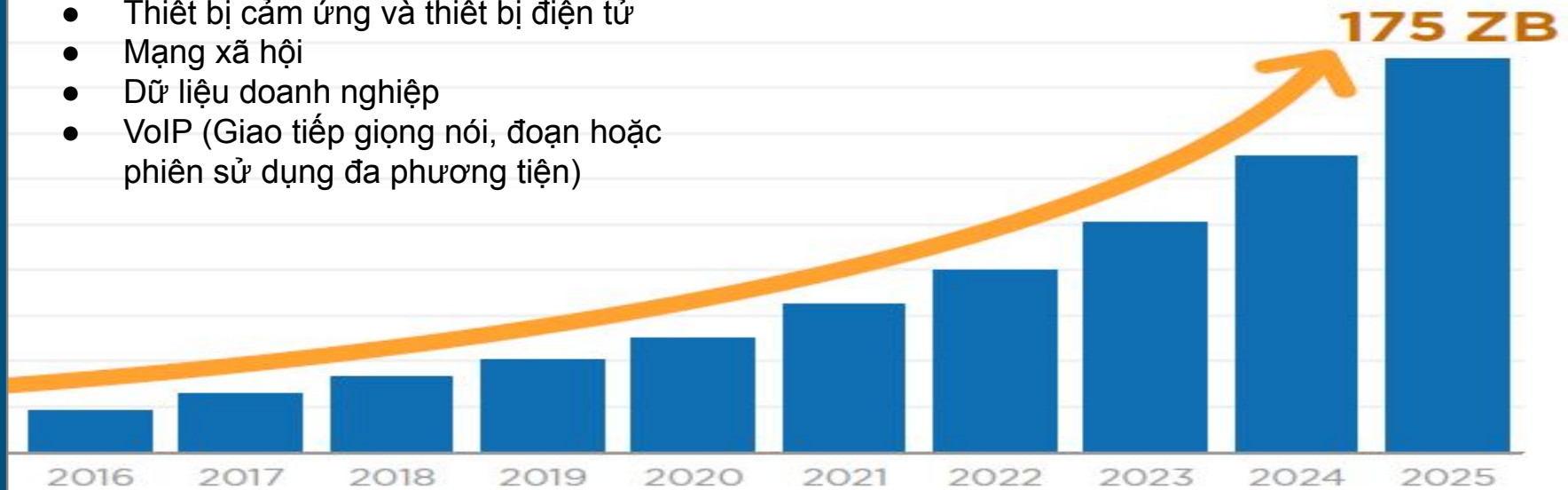
- Dữ liệu lớn đang trở thành tiêu chuẩn, và nhu cầu nhân lực cho ngành kỹ sư dữ liệu càng cấp thiết
- Dữ liệu lớn:

Cách để đối phó với kích thước

Lớn mà các phương pháp truyền thống không làm được

Tăng trưởng của dữ liệu lớn

- Thiết bị cảm ứng và thiết bị điện tử
- Mạng xã hội
- Dữ liệu doanh nghiệp
- VoIP (Giao tiếp giọng nói, đoạn hoặc phiên sử dụng đa phương tiện)



175 ZB sẽ được tạo ra vào năm 2025, theo IDC

5 yếu tố

- ✓ Khối dữ liệu (bao nhiêu?)
- ✓ Đa dạng (kiểu dữ liệu nào?)
- ✓ Tốc độ (mức độ thường xuyên?)
- ✓ Độ tin cậy (mức độ chính xác?)
- ✓ Giá trị (có ý nghĩa?)

Tóm tắt

- ✓ Data engineering
- ✓ Dữ liệu chảy trong tổ chức/công ty
- ✓ Khi nào thì kỹ sư dữ liệu can thiệp vào
- ✓ Nhiệm vụ của kỹ sư giữ liệu là gì
- ✓ Cách kỹ sư dữ liệu liên quan tới dữ liệu lớn.