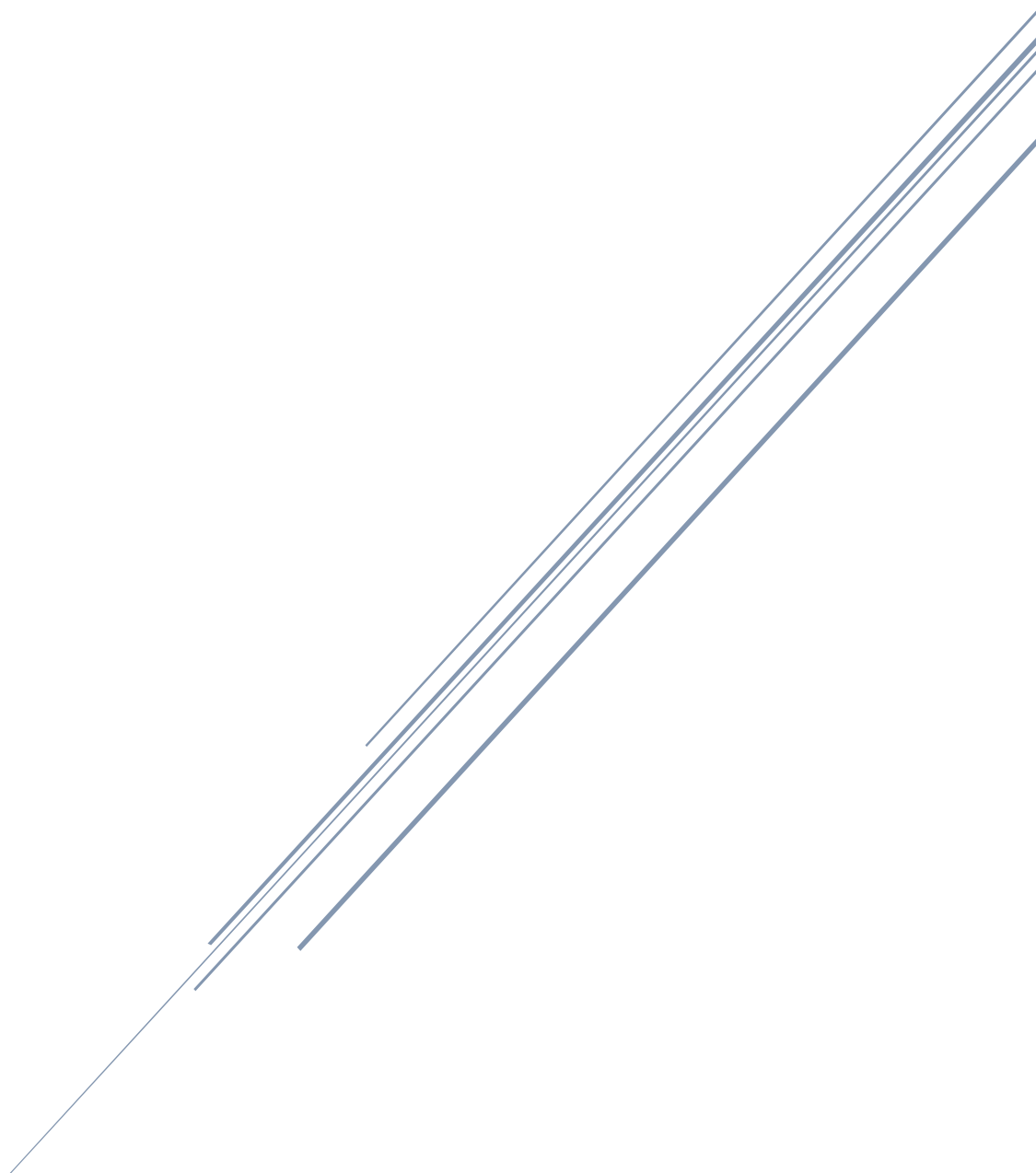


# MẠNG XÃ HỘI

HƯỚNG DẪN THỰC HIỆN ĐỒ ÁN



Biên soạn: ThS. Nguyễn Thị Anh Thư

## MỤC LỤC

1. Mục tiêu .....	2
2. Yêu cầu .....	2
3. Danh sách đề tài .....	2
4. Hướng dẫn thực hiện.....	2
4.1. Hình thức trình bày.....	2
4.2. Nội dung thực hiện .....	3

## 1. Mục tiêu

- Ôn tập lại các khái niệm cơ bản.
- Nắm vững kiến thức cơ bản về bài toán khai thác dữ liệu mạng.
- Các kỹ năng giải quyết vấn đề.
- Kỹ năng tìm kiếm tài liệu, tổng hợp và viết báo cáo.
- Khả năng thuyết trình.

## 2. Yêu cầu

- Thực hiện đồ án theo nhóm, sinh viên tự lập nhóm, mỗi nhóm có **3~5 thành viên** (*sĩ số*  $\leq 50$ ) **6~7 thành viên** (*sĩ số*  $> 50$ ). Sau khi lập nhóm, mỗi nhóm bầu ra **1 nhóm trưởng** chịu trách nhiệm điều hành hoạt động: *phân công công việc và đại diện nhóm nộp bài*.
- Nhiệm vụ của nhóm là “**Tìm hiểu kiến trúc dữ liệu lớn để xử lý dữ liệu mạng xã hội**”.
- Sản phẩm đồ án: **01 file báo cáo word, 01 file thuyết trình powerpoint, 01 video báo cáo đồ án và 01 phần mềm thực nghiệm (nếu có)**.
- Cách thức nộp:
  - Tạo drive theo Tên nhóm → lưu toàn bộ sản phẩm và bảng phân công công việc lên drive → nộp link trên <https://courses.uit.edu.vn/> .
  - Lịch báo cáo đồ án trên <https://courses.uit.edu.vn/> .

## 3. Danh sách đề tài

Tìm hiểu về một công nghệ triển khai kiến trúc dữ liệu lớn phổ biến để xử lý dữ liệu mạng xã hội. Mỗi nhóm tiến hành thảo luận và chọn một đề tài để thực hiện.

**Lưu ý:** Mỗi nhóm chọn một đề tài riêng và không trùng lặp. Ưu tiên cho nhóm đăng ký trước.

## 4. Hướng dẫn thực hiện

### 4.1. Hình thức trình bày

#### File báo cáo word

- Trang bìa: Các thông tin cơ bản về trường, môn học, tên đề tài, tên giảng viên hướng dẫn, tên nhóm và thông tin của các thành viên (*MSSV + Họ và tên*).

- Mục lục: Các nội dung chính của cuốn báo cáo và mục lục nên được tạo tự động.
- Nội dung:
  - Giới thiệu
  - Đánh giá ưu nhược điểm, ứng dụng, cách thức triển khai các thành phần theo kiến trúc dữ liệu lớn (*data warehouse, data lake, lakehouse, data sources, ingest, process, enrich*), ví dụ demo.
  - Kết luận và hướng phát triển
  - Tài liệu tham khảo (*theo chuẩn IEEE*).
- Định dạng: Khổ giấy A4, font chữ Times New Roman, size 13, cách dòng 1.5, canh lề trên dưới trái phải là 1 inch.

#### **File thuyết trình powerpoint**

- Bố cục:
  - Slide 1: Thông tin về đề tài, giảng viên hướng dẫn, nhóm thực hiện và các thành viên của nhóm.
  - Slide 2: Giới thiệu những nội dung sẽ trình bày.
  - Slide 3: Giới thiệu đề tài sẽ thực hiện.
  - Slide 4~(n-2): Nội dung trình bày.
  - Slide (n-1): Kết luận
  - Slide n: Cảm ơn thầy/cô và các bạn đã lắng nghe.
- Định dạng slides:
  - Đánh số slide, size chữ phải đủ lớn để người ngồi xa cũng có thể đọc được và hình ảnh sử dụng trên slide phải rõ nét.
  - Nội dung slide nên hạn chế quá nhiều chữ, thay vào đó tận dụng mô hình và hình ảnh minh họa để thể hiện nội dung.

#### **4.2. Nội dung thực hiện**

Chọn một công nghệ triển khai kiến trúc dữ liệu lớn phổ biến để xử lý dữ liệu mạng xã hội và tìm hiểu các nội dung sau:

1. Giới thiệu
2. Đánh giá ưu nhược điểm
3. Ứng dụng

4. Cách thức triển khai các thành phần theo kiến trúc dữ liệu lớn:
  - a. Data house: data warehouse, data lake, lakehouse. Thực hiện ví dụ demo cho mỗi chức năng tìm hiểu.
  - b. Các components: data sources, ingest, process, enrich. Thực hiện ví dụ demo cho mỗi chức năng tìm hiểu.
  - c. Triển khai thành một hệ thống hoàn thiện để xử lý dữ liệu mạng xã hội (nếu có).
5. Kết luận và hướng phát triển
6. Tài liệu tham khảo