

# ÁP DỤNG LLM VÀO ĐÁNH GIÁ VÀ KIỂM ĐỊNH CÁC TIÊU CHUẨN VỀ DỮ LIỆU THEO GDPR, ND13, ISO 27701

**Giảng viên hướng dẫn**

Giảng viên: Nguyễn An Khương

**Trình bày**

Nguyễn Lê Quốc Anh - CHAT3P01

Tô Duy Hình - CHAT3P06

# NỘI DUNG

# Bối cảnh & Nhu cầu

# Đặc điểm của LLM và RAGs

# Giải pháp đề xuất

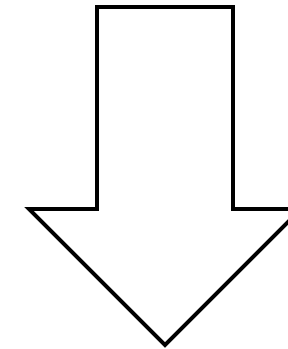
# Thực nghiệm

# Khó khăn & Hạn chế

# Bài học

# # BỐI CẢNH & NHU CẦU

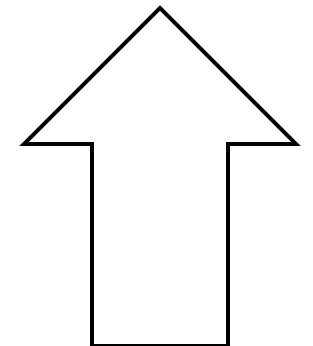
	<b>LUẬT DỮ LIỆU</b> <ul style="list-style-type: none"><li>• Quy định về sử dụng dữ liệu</li><li>• Đang dự thảo</li></ul>
	<b>NGHỊ ĐỊNH</b> <ul style="list-style-type: none"><li>• Nghị định bảo vệ thông tin cá nhân</li><li>• Ban hành năm 2023</li></ul>
	<b>LUẬT AN NINH MẠNG</b> <ul style="list-style-type: none"><li>• Quy định chung về An ninh mạng</li><li>• Ban hành năm 2018</li></ul>



Chính phủ  
siết chặt quản  
lý



Doanh nghiệp  
còn thiếu hiểu  
biết về pháp  
lý, hạn chế về  
công nghệ.  
Nguồn lực  
bảo mật còn  
thiếu

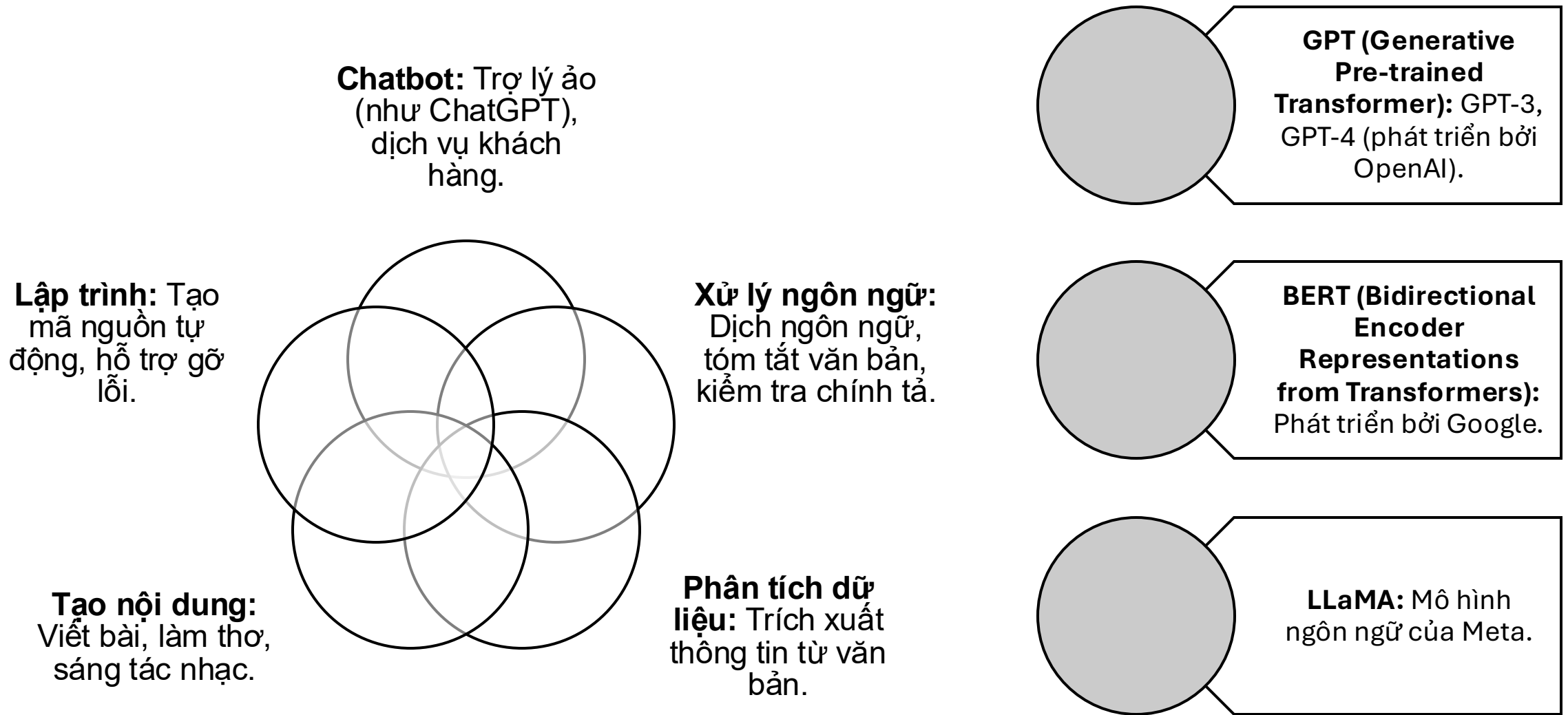


# # Đặc điểm của LLM và RAGs

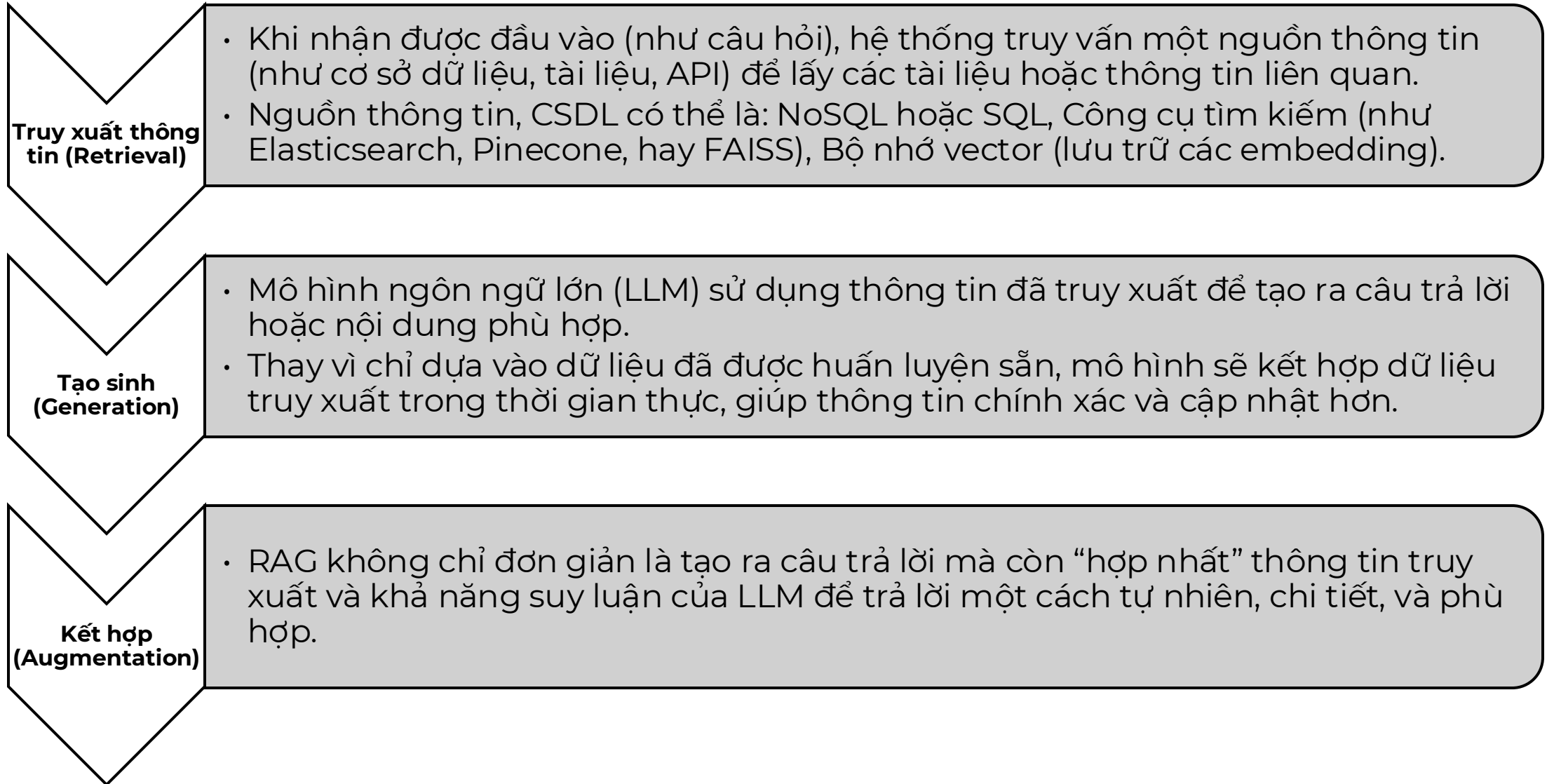
**LLM** là viết tắt của **Large Language Model** (Mô hình ngôn ngữ lớn). Đây là một loại mô hình trí tuệ nhân tạo (AI) được huấn luyện trên lượng dữ liệu văn bản khổng lồ để hiểu và tạo ra ngôn ngữ tự nhiên (như tiếng Anh, tiếng Việt, v.v.)



# # Đặc điểm của LLM và RAGs



# # Đặc điểm của LLM và RAGs



# # Đặc điểm của LLM và RAGs

## **Tokenization**

- Tokenization là quá trình chia nhỏ một đoạn văn bản thành các đơn vị nhỏ hơn (tokens). Tokens có thể là từ, cụm từ, hoặc thậm chí là ký tự.

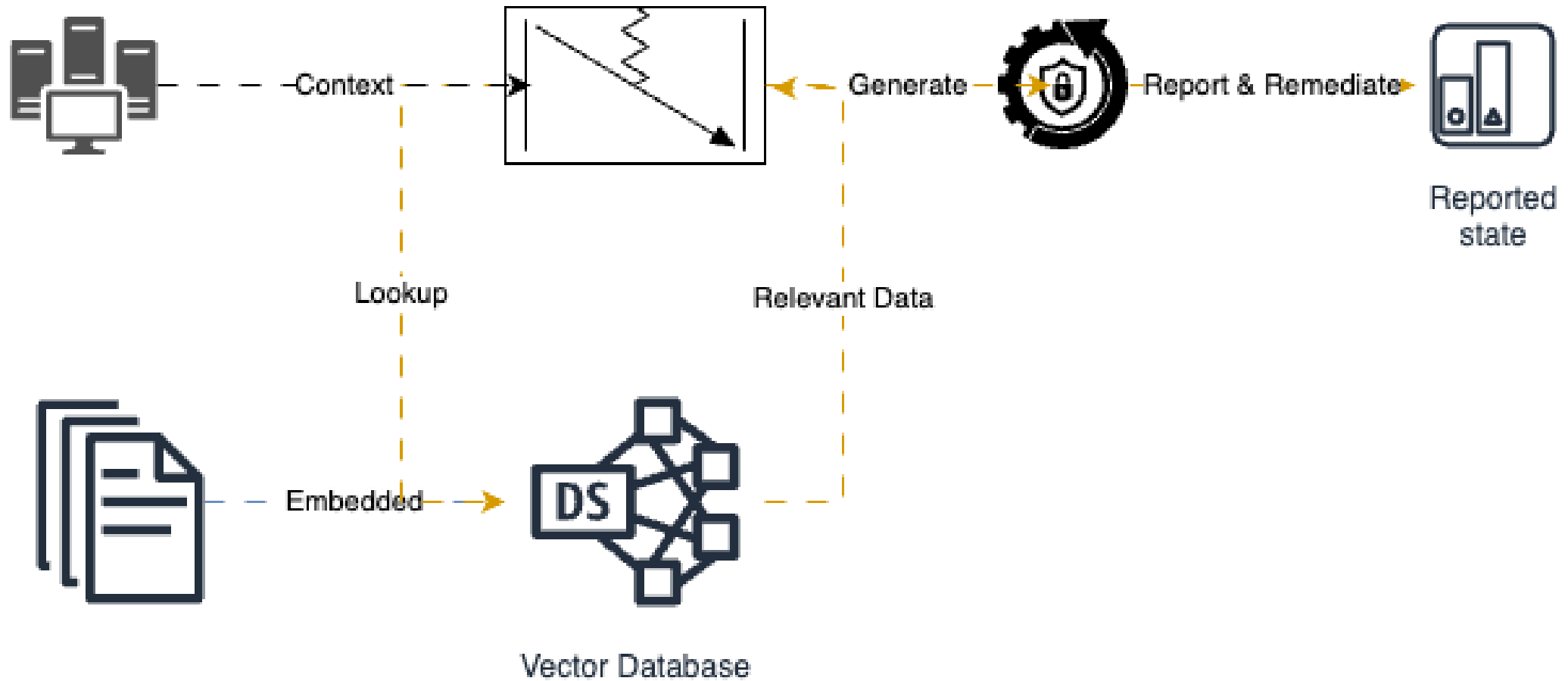
## **Embedding**

- Embedding là một cách để biểu diễn dữ liệu (đặc biệt là văn bản, từ, hoặc câu) dưới dạng các vector số, trong không gian nhiều chiều. Mục tiêu của embedding là mã hóa ý nghĩa hoặc ngữ nghĩa của dữ liệu sao cho những dữ liệu tương tự sẽ có vector gần nhau trong không gian đó

## **Vector Search**

- Vector Search (hay Approximate Nearest Neighbor Search - ANN Search) là phương pháp tìm kiếm các đối tượng tương tự dựa trên vector biểu diễn của chúng trong không gian nhiều chiều.

# # Giải pháp đề xuất





# # THỰC NGHIỆM

## Thử nghiệm trên ChatGPT

Dựa vào hiện trạng, tôi sẽ phân tích các vi phạm, tên thiết bị, tiêu chuẩn liên quan và nguy cơ về mất an toàn dữ liệu theo Luật Việt Nam (Decree 13/2023/ND-CP, Luật An toàn thông tin mạng).

### 1. Tổng hợp vi phạm

#### Vi phạm 1: Sử dụng TLS phiên bản cũ (TLS 1.0)

##### •Chi tiết vi phạm:

- TLS 1.0 không đáp ứng tiêu chuẩn bảo mật hiện tại, dễ bị tấn công đánh chặn (MITM) hoặc giải mã.
- Các cổng bị ảnh hưởng: 2083, 443, 2053, 2096, 2087, 8443.

•**Tiêu chuẩn liên quan:** PCI DSS yêu cầu sử dụng TLS v1.1 trở lên.

##### •Nguy cơ:

- Thông tin nhạy cảm như dữ liệu thẻ thanh toán có thể bị đánh cắp.
- Vi phạm quy định về bảo vệ dữ liệu cá nhân trong Luật An toàn thông tin mạng Việt Nam.

#### Vi phạm 2: SQL Injection

##### •Chi tiết vi phạm:

- Máy chủ web bị phát hiện dễ bị tấn công SQL Injection.

•**Tiêu chuẩn liên quan:** PCI DSS yêu cầu kiểm tra và bảo vệ chống lại tấn công SQL Injection.

##### •Nguy cơ:

- Hacker có thể truy xuất, sửa đổi hoặc xóa dữ liệu nhạy cảm.
- Vi phạm Điều 18 của Nghị định 13/2023/ND-CP về bảo vệ dữ liệu cá nhân nhạy cảm.

#### Vi phạm 3: CVSS >= 4.0 (lỗ hổng có mức nguy hiểm cao)

##### •Chi tiết vi phạm:

- Báo cáo phát hiện lỗ hổng có CVSS >= 4.0 trên hệ thống.
- Lỗ hổng mức độ cao chưa được chỉ rõ, cần phân tích thêm.

•**Tiêu chuẩn liên quan:** PCI DSS yêu cầu không được có lỗ hổng cao trong hệ thống trước khi được phê duyệt.

##### •Nguy cơ:

- Lỗ hổng nghiêm trọng có thể bị khai thác để xâm nhập trái phép.
- Vi phạm Điều 27 Luật An toàn thông tin mạng về đảm bảo an toàn thông tin.

#### Vi phạm 4: 6 lỗ hổng mức trung bình khác

##### •Chi tiết vi phạm:

- Các lỗ hổng mức trung bình được liệt kê tại các URL liên quan.

##### •Nguy cơ:

- Tích lũy lỗ hổng nhỏ có thể tạo điều kiện cho các cuộc tấn công chuỗi.
- Vi phạm yêu cầu kiểm soát rủi ro liên tục trong PCI DSS và Luật Việt Nam.

### 2. Tên thiết bị

•Không có thông tin rõ ràng về tên thiết bị trong báo cáo.

•Dựa trên cổng và dịch vụ, có thể liên quan đến máy chủ web hoặc cơ sở dữ liệu.

### 3. Tiêu chuẩn liên quan

•**PCI DSS (Payment Card Industry Data Security Standard):** Tiêu chuẩn chính trong báo cáo.

•**Luật An toàn thông tin mạng Việt Nam:** Điều 18, Điều 27.

•**Nghị định 13/2023/ND-CP:** Bảo vệ dữ liệu cá nhân, đặc biệt liên quan đến dữ liệu thẻ thanh toán.

### 4. Nguy cơ về mất an toàn dữ liệu

•**Nguy cơ bị đánh cắp dữ liệu cá nhân:** Vi phạm TLS 1.0, SQL Injection.

•**Nguy cơ bị tấn công mạng:** Lỗ hổng mức cao hoặc trung bình chưa được vá.

•**Rủi ro không tuân thủ luật pháp:** Có thể bị phạt hành chính, đình chỉ hoạt động xử lý dữ liệu hoặc bồi thường thiệt hại.

### 5. Khuyến nghị

•**Với TLS 1.0:** Nâng cấp lên TLS 1.2 hoặc TLS 1.3.

•**Với SQL Injection:** Kiểm tra và sửa lỗ hổng trong mã nguồn, áp dụng các cơ chế bảo vệ (parameterized queries, WAF).

•**Với lỗ hổng CVSS cao:** Ưu tiên khắc phục ngay và kiểm tra các biện pháp giảm thiểu rủi ro.

•**Tuân thủ PCI DSS:** Thực hiện kiểm tra định kỳ và báo cáo lại sau khi vá lỗi.

Nếu cần hỗ trợ chi tiết hơn về cách vá các lỗ hổng hoặc tư vấn pháp lý, hãy cho tôi biết!

# # THỰC NGHIỆM

```markdown

## Báo cáo phân tích vi phạm an toàn dữ liệu theo luật Việt Nam

**Mô hình nghiên cứu áp dụng RAGs**

### I. Trích dẫn điều khoản cụ thể từ văn bản

#### Luật An ninh mạng 2018

- Điều 5, khoản 1

- Điều 26

#### Nghị định số 53/2022/NĐ-CP

- Điều 24, khoản 7

### II. Phân tích và giải thích rõ ràng

#### A. Vi phạm phát hiện được

1. Sử dụng các phiên bản SSL/TLS cũ (TLS v1.0)

2. Lỗ hổng SQL Injection

3. Tài khoản mặc định hoặc backdoor applications (malware)

#### B. Thiết bị liên quan

Các thiết bị qua các cổng: port `443`, `2083`, `2053`, `2096`, `2087`, `8443`

#### C. Tiêu chuẩn áp dụng

PCI DSS Compliance

### III. Nguy cơ mất an toàn dữ liệu

1. Rò rỉ thông tin khách hàng do lỗ hổng SQL Injection.

2. Nguy cơ mất mát danh tiếng và thiệt hại tài chính do sử dụng giao thức mã hóa lỗi thời.

3. Khả năng mất quyền kiểm soát hệ thống do malware hoặc backdoor applications.

```

# # THỰC NGHIỆM

Legal Assistant

System Information

System Info

- Python: 3.8.10
- PyTorch: 2.4.1+cu124
- OS: Windows 10

CUDA Details

- CUDA Version: 12.4
- cuDNN Version: 90100
- Device: NVIDIA GeForce RTX 2080
- Capability: 7.5

GPU Memory

Total	Used
8.00 GB	1.02 GB

cuDNN Status

- Enabled: ☒
- Benchmark: ☐
- Deterministic: ☐

Resource Usage

CPU Usage %

50

Legal Assistant 🤖 ⚖️

Nhập câu hỏi của bạn...

# # THỰC NGHIỆM

Legal Assistant

localhost:8501

Deploy

System Information

System Info

- Python: 3.8.10
- PyTorch: 2.4.1+cu124
- OS: Windows 10

CUDA Details

- CUDA Version: 12.4
- cuDNN Version: 90100
- Device: NVIDIA GeForce RTX 2080
- Capability: 7.5

GPU Memory

Total	Used
8.00 GB	0.50 GB

cuDNN Status

- Enabled: ✓
- Benchmark: ✗
- Deterministic: ✗


Resource Usage

CPU Usage %

0

50

100

Legal Assistant 

Dựa vào dữ liệu từ các dự thảo luật, nghị định dữ liệu và luật dữ liệu, luật an ninh mạng của Việt Nam, hãy chỉ ra những sai phạm về dữ liệu (xử lý, lưu trữ) từ thông tin (địa chỉ nguồn, pattern phát hiện định danh cá nhân, địa chỉ đích):

Source IP	Destination IP	Ports	Pattern PII	PII Detected	Action
172.31.61.106	43.156.222.216	57124 → 80	<code>\b\d{8,12}\b</code>	00001326	<button>View Graph</button>
172.31.61.216	43.152.112.190	61259 → 80	<code>\b\d{8,12}\b</code>	00001371	<button>View Graph</button>

Thêm câu hỏi của bạn...

Type here to search

27°C Nắng nhiều nơi

10:12 AM 12/28/2024

# # KHÓ KHĂN VÀ HẠN CHẾ

## **Tài nguyên**

- Đòi hỏi nhiều tài nguyên tính toán để huấn luyện và vận hành.

## **Độ chính xác**

- Có thể đưa ra thông tin không chính xác hoặc thiên lệch.

## **Đạo đức**

- Rủi ro lạm dụng (như tạo tin giả, lừa đảo).

## **Đặc thù các tiêu chuẩn**

- Đa số thực hiện theo phương pháp phỏng vấn

**?** Nhận xét – Tư vấn cải thiện  
từ Giảng viên

Q&A

**CẢM ƠN!**