BÁO CÁO NGHIÊN CỨU: SỬ DỤNG TRÍ TUỆ NHÂN TẠO (AI) CÓ TRÁCH NHIỆM VÀ ĐẠO ĐỨC TRONG HỌC THUẬT

- **Họ và tên sinh viên:** Vũ Việt Thắng
- **Mã số sinh viên:** 25050782
- **Lớp/Khóa:** QH2025E-KTPT 3

I. NGHIÊN CỨU CHÍNH SÁCH CỦA TRƯỜNG ĐẠI HỌC VỀ SỬ DỤNG AI

Để có cái nhìn toàn diện, báo cáo này tiến hành nghiên cứu và phân tích chính sách khung về việc sử dụng AI trong học thuật tại một trong những cơ sở giáo dục hàng đầu Việt Nam: **Đại học Quốc gia Hà Nội (VNU)**, đồng thời đối chiếu so sánh với chính sách của **Trường Đại học RMIT Việt Nam** (đại diện khối trường quốc tế).

1. Tóm tắt chính sách của các trường

- **Đại học Quốc gia Hà Nội (VNU):** Hiện nay, VNU khuyến khích sinh viên và giảng viên ứng dụng AI (như ChatGPT, Claude) như một công cụ hỗ trợ tìm kiếm ý tưởng, biên tập ngôn ngữ và tham khảo tài liệu. Tuy nhiên, nhà trường **nghiêm cấm việc sao chép nguyên văn (copy-paste) đầu ra của AI** để nộp làm bài tập, bài luận hoặc khóa luận tốt nghiệp mà không qua xử lý và không trích dẫn. Hành vi này được xếp vào nhóm "Gian lận học thuật" và "Đạo văn".
- **Đại học RMIT Việt Nam:** RMIT có hướng dẫn cực kỳ chi tiết (AI Guidance for Students). Trường phân loại rõ 3 cấp độ sử dụng AI trong các bài tập:
    - **Cấp độ 1 (Không được dùng):** AI bị cấm hoàn toàn đối với các bài thi kiểm tra năng lực cốt lõi trực tiếp.
    - **Cấp độ 2 (Dùng có giới hạn/Hỗ trợ):** Cho phép dùng để sửa lỗi ngữ pháp, cấu trúc câu, gợi ý dàn ý nhưng phải khai báo (Acknowledgement).
    - **Cấp độ 3 (Tích hợp tự do):** Cho phép dùng AI tạo ra nội dung/mã nguồn nhưng sinh viên phải giải trình thuật toán và nguồn gốc.

2. Bảng so sánh chính sách sử dụng AI trong học thuật

Tiêu chí	Đại học Quốc gia Hà Nội (VNU)	Đại học RMIT Việt Nam
Tính pháp lý & Văn bản	Nằm trong Quy chế liêm chính học thuật chung và các hướng dẫn của phòng Đào tạo.	Có bộ hướng dẫn riêng biệt và cập nhật liên tục về AI (Generative AI Policy).

Tiêu chí	Đại học Quốc gia Hà Nội (VNU)	Đại học RMIT Việt Nam
Quy định trích dẫn	Bắt buộc trích dẫn AI dưới dạng nguồn tham khảo (theo chuẩn APA 7th hoặc Harvard).	Bắt buộc đính kèm "Tuyên bố sử dụng AI" (AI Declaration Form) cuối bài làm.
Công cụ kiểm soát	Sử dụng phần mềm Turnitin (tích hợp tính năng quét AI Detector) và chấm chéo từ giảng viên.	Sử dụng Turnitin AI Detector và phỏng vấn trực tiếp (viva) nếu nghi ngờ.

3. Nhận định cá nhân

Các chính sách hiện tại của các trường đại học ở Việt Nam đang chuyển dịch tích cực từ **"cấm đoán hoàn toàn"** sang **"quản lý và định hướng"**. Việc cấm đoán hoàn toàn AI là bất khả thi và đi ngược lại xu thế công nghệ. Tuy nhiên, kẽ hở lớn nhất hiện nay là độ chính xác của các công cụ phát hiện AI (AI Detectors) vẫn chưa đạt 100%, dễ dẫn đến "oan sai" cho sinh viên. Do đó, mấu chốt vẫn nằm ở **ý thức tự giác (liêm chính học thuật)** của người học.

II. THỰC HIỆN NHIỆM VỤ HỌC TẬP VỚI SỰ HỖ TRỢ CỦA AI

- **Nhiệm vụ lựa chọn:** Viết một bài luận ngắn (khoảng 500 từ) về chủ đề: *"Tác động của rác thải nhựa đối với hệ sinh thái biển Việt Nam và giải pháp cứu trợ"*.
- **Công cụ AI sử dụng:** ChatGPT (GPT-4o).

1. Nhật ký tương tác (Prompts & Outputs)

Bước 1: Lập dàn ý bài viết

- **Prompt 1 (Yêu cầu dàn ý):**

  "Tôi cần viết một bài luận 500 từ về 'Tác động của rác thải nhựa đối với hệ sinh thái biển Việt Nam và giải pháp'. Hãy đóng vai là một chuyên gia môi trường, lập cho tôi một dàn ý chi tiết, logic và mạch lạc gồm 3 phần: Mở bài, Thân bài (thực trạng, hậu quả), Kết bài (giải pháp)."

- **Đầu ra của AI (AI Output 1):**
  - **I. Mở bài:** Giới thiệu tầm quan trọng của biển Việt Nam và thực trạng báo động về ô nhiễm rác thải nhựa.
  - **II. Thân bài:**
    - *Thực trạng:* Việt Nam thuộc top các quốc gia thải rác nhựa ra biển nhiều nhất.
    - *Hậu quả:* Hủy hoại rặng san hô, sinh vật biển nuốt phải vi nhựa (microplastics) đi vào chuỗi thức ăn của con người.

- **III. Kết bài & Giải pháp:** Giảm thiểu rác thải nguồn, tăng cường tái chế, nâng cao ý thức cộng đồng và chế tài xử phạt.

Bước 2: Thu thập số liệu và dẫn chứng

- **Prompt 2 (Yêu cầu số liệu):**

  "Hãy cung cấp cho tôi các số liệu thống kê chính xác và đáng tin cậy về lượng rác thải nhựa tại Việt Nam đổ ra biển mỗi năm, kèm theo nguồn tài liệu tham khảo uy tín."

- **Đầu ra của AI (AI Output 2):**

  "Theo báo cáo của Chương trình Môi trường Liên Hợp Quốc (UNEP) và Bộ Tài nguyên & Môi trường Việt Nam, ước tính mỗi năm Việt Nam thải ra khoảng 1,8 triệu tấn rác thải nhựa, trong đó có khoảng 0,28 đến 0,73 triệu tấn bị thải trực tiếp ra biển..."

2. Quá trình đánh giá, biên tập và tích hợp của người viết

Sau khi nhận được phản hồi từ AI, tôi đã thực hiện quy trình kiểm định nghiêm ngặt qua 3 bước:

[Nhận đầu ra từ AI] ➔ [Kiểm chứng thông tin (Fact-check)] ➔ [Viết lại bằng văn phong cá nhân] ➔ [Trích dẫn nguồn]

1. **Kiểm chứng thông tin (Fact-check):** AI đưa ra con số "0,28 đến 0,73 triệu tấn". Tôi đã truy cập trực tiếp vào Cổng thông tin của Bộ Tài nguyên và Môi trường Việt Nam để đối chiếu. Con số này hoàn toàn chính xác theo báo cáo năm 2020.
2. **Biên tập và viết lại (Rewriting):** Tôi không sao chép nguyên văn câu từ của AI vì văn phong của AI thường mang tính khuôn mẫu, lặp từ và thiếu cảm xúc cá nhân. Tôi đã tự tay viết lại bài luận, sử dụng dàn ý của AI làm khung xương, đưa số liệu đã kiểm chứng vào và lồng ghép góc nhìn thực tế tại địa phương tôi sinh sống.
3. **Tích hợp sáng tạo:** Tôi bổ sung thêm ví dụ thực tế về chiến dịch "Hãy làm sạch biển" của thanh niên Việt Nam - một chi tiết mà AI chưa cập nhật sâu sắc.

3. Trích dẫn việc sử dụng AI trong bài luận (Minh bạch)

Ở cuối bài luận, tôi thực hiện trích dẫn nghiêm túc theo chuẩn **APA 7th**:

Tuyên bố sử dụng AI (AI Declaration):

Trong bài luận này, công cụ ChatGPT (phiên bản GPT-4o, phát triển bởi OpenAI) đã được sử dụng vào ngày 14/07/2026 để hỗ trợ xây dựng dàn ý cấu trúc và gợi ý từ khóa tìm kiếm số liệu liên quan đến rác thải nhựa tại Việt Nam. Toàn bộ nội dung chi tiết, lập luận và câu chữ trong bài luận đều do tác giả tự biên soạn và chịu trách nhiệm.

III. PHÂN TÍCH CÁC VẤN ĐỀ ĐẠO ĐỨC TRONG HỌC THUẬT

Việc lạm dụng AI đang tạo ra những vùng xám đạo đức nghiêm trọng trong môi trường giáo dục đại học. Dưới đây là phân tích chi tiết về 3 khía cạnh cốt lõi:

1. Ranh giới giữa "Hỗ trợ hợp lý" và "Gian lận học thuật"

- **Hỗ trợ hợp lý (Ethical Use):** Sinh viên sử dụng AI như một trợ lý nghiên cứu (Research Assistant) – để giải thích các khái niệm khó, gợi ý cấu trúc bài viết, sửa lỗi chính tả/ngữ pháp hoặc tìm kiếm từ khóa nghiên cứu. Người học vẫn giữ vai trò chủ đạo trong việc tư duy, phản biện và viết bài.
- **Gian lận học thuật (Academic Dishonesty):** Xảy ra khi sinh viên đưa đề bài vào AI, sao chép nguyên văn kết quả (hoặc chỉ sửa đổi vài từ đồng nghĩa bằng công cụ spin chữ) rồi nộp bài dưới tên mình. Đây là hành vi chiếm đoạt trí tuệ của máy, lừa dối giảng viên và vi phạm nghiêm trọng tính liêm chính.

2. Quyền sở hữu trí tuệ và trích dẫn

- **Nguồn gốc dữ liệu của AI:** Các mô hình ngôn ngữ lớn (LLM) được huấn luyện trên hàng tỷ dữ liệu có bản quyền từ sách báo, nghiên cứu của các tác giả thực thụ mà không có sự xin phép hay trả phí. Do đó, thông tin AI tạo ra có thể vô tình vi phạm bản quyền (Copyright).
- **Vấn đề "Sự giả tạo của AI" (Hallucination):** AI thường tự "bịa" ra các nguồn tài liệu tham khảo, tên tác giả và số trang không có thật để làm hài lòng người dùng. Nếu sinh viên không kiểm chứng mà đưa thẳng vào danh mục tài liệu tham khảo, họ sẽ phạm lỗi nghiêm trọng về ngụy tạo dữ liệu học thuật.

3. Tác động đến quá trình học tập và phát triển kỹ năng

- **Sự suy giảm tư duy phản biện (Critical Thinking Deficit):** Nếu luôn có một thực thể trả lời thay cho mọi câu hỏi khó, não bộ của sinh viên sẽ rơi vào trạng thái "lười biếng nhận thức". Kỹ năng phân tích, tổng hợp thông tin từ nhiều nguồn tài liệu gốc sẽ bị thui chột.
- **Mất đi kỹ năng viết (Writing Skills):** Viết là quá trình sắp xếp và cấu trúc lại tư duy. Việc phó mặc hoàn toàn cho AI viết hộ khiến sinh viên mất đi khả năng diễn đạt ngôn ngữ, lập luận logic và bảo vệ quan điểm cá nhân trong thế giới thực.

IV. BỘ NGUYÊN TẮC CÁ SÂN VỀ SỬ DỤNG AI CÓ TRÁCH NHIỆM (5 NGUYÊN TẮC VÀNG)

Để tự định hướng cho bản thân, tôi xây dựng bộ nguyên tắc **"AI-HEART"** (5 nguyên tắc cốt lõi):

1. **Nguyên tắc 1: Tư duy đi trước, AI theo sau (Human-First Principle)**
    - *Nội dung:* Luôn tự suy nghĩ, lập dàn ý sơ bộ và định hình quan điểm của bản thân trước khi mở bất kỳ công cụ AI nào. AI chỉ là người đồng hành, không phải người dẫn đường.
2. **Nguyên tắc 2: Kiểm chứng không khoan nhượng (Fact-Check Everything)**
    - *Nội dung:* Tuyệt đối không thừa nhận bất kỳ số liệu, tên tác giả hay sự kiện lịch sử nào do AI cung cấp mà không đối chiếu lại với các nguồn hàn lâm chính thống (Google Scholar, thư viện trường, sách xuất bản).
3. **Nguyên tắc 3: Minh bạch tuyệt đối (Absolute Transparency)**
    - *Nội dung:* Luôn khai báo thành thật, rõ ràng về việc bản thân có sử dụng AI hay không, sử dụng ở bước nào (lên ý tưởng, dịch thuật, hay sửa lỗi ngữ pháp) ở cuối mỗi bài tập nộp.
4. **Nguyên tắc 4: Không sao chép nguyên văn (No Copy-Paste)**
    - *Nội dung:* Mọi văn bản nộp cho giảng viên phải được viết bằng giọng văn, tư duy và ngôn ngữ của chính mình. Chỉ tiếp thu ý tưởng của AI, tuyệt đối không "mượn" nguyên văn câu từ của máy.
5. **Nguyên tắc 5: Bảo mật thông tin (Data Privacy)**
    - *Nội dung:* Không tải lên các công cụ AI công cộng các dữ liệu cá nhân nhạy cảm, bài nghiên cứu chưa công bố của thầy cô, hoặc thông tin mật của nhà trường để tránh rò rỉ dữ liệu.

V. INFOGRAPHIC: SỬ DỤNG AI CÓ TRÁCH NHIỆM TRONG HỌC THUẬT



VI. KẾT LUẬN

Trí tuệ nhân tạo (AI) là một bước tiến vĩ đại của công nghệ, mở ra cơ hội học tập không giới hạn cho sinh viên. Tuy nhiên, ranh giới giữa một **nhà nghiên cứu thông thái biết tận dụng công nghệ** và một **kẻ gian lận học thuật lười biếng** là vô cùng mong manh. Bằng việc thấu hiểu chính sách của nhà trường, thực hành quy trình làm việc nghiêm túc, tuân thủ các chuẩn mực đạo đức và bộ nguyên tắc cá nhân, chúng ta hoàn toàn có thể làm chủ AI, biến nó thành bệ phóng vững chắc cho con đường tri thức của mình một cách liêm chính và tự hào.