

LIÊM CHÍNH HỌC THUẬT TRONG ĐÀO TẠO VÀ HỌC TẬP NGÀNH CÔNG NGHỆ THÔNG TIN: TỪ NHẬN THỨC, KỸ NĂNG ĐẾN CÔNG CỤ HỖ TRỢ

Trịnh Thị Xuân^{1}, Mai Thị Thuý Hà¹, Nguyễn Thuỳ Linh¹, Lê Thị Thanh Thuỳ¹*
**Tác giả liên hệ, Email: trinxuan@hou.edu.vn*

Ngày tòa soạn nhận được bài báo: 02/06/2025

Ngày phản biện đánh giá: 15/08/2025

Ngày bài báo được duyệt đăng: 29/08/2025

DOI: 10.59266/houjs.2025.669

Tóm tắt: Trong bối cảnh hiện nay khi công nghệ phát triển nhanh chóng thì vấn đề đạo văn mã nguồn, giả mạo dữ liệu và sử dụng nội dung do trí tuệ nhân tạo sinh ra ngày càng phổ biến, trở lên tinh vi và khó phát hiện đặc biệt trong môi trường học tập của sinh viên Khoa Công nghệ Thông tin. Bài viết này phân tích về các hình thức đạo văn mã nguồn khác nhau từ mức độ đơn giản đến mức độ sao chép nguyên văn mã nguồn. Đồng thời, nhóm tác giả trình bày phân biệt giữa sử dụng hợp pháp, tham khảo và hành vi đạo văn trong môi trường mã nguồn mở.

Với vấn đề về liên chính dữ liệu trong bài viết cũng có đề xuất đến sai lệch dữ liệu có thể từ mức vô ý đến mức cố ý, đồng thời có trình bày về nguyên tắc FAIR và tiêu chí ALCOA+ để đảm bảo liên chính dữ liệu.

Nhóm tác giả đã đề xuất mô hình tích hợp “Nhận thức - Kỹ năng - Công cụ” nhằm xây dựng văn hoá liên chính học thuật trong môi trường đào tạo Công nghệ Thông tin. Cuối cùng, nhóm tác giả cũng đã đề xuất khung liên chính học thuật gồm 4 bước khác nhau để hỗ trợ giảng viên trong việc có thể đánh giá khách quan, minh bạch và công bằng với kết quả học tập cũng như sản phẩm nghiên cứu của sinh viên.

Từ khoá: đạo văn mã nguồn, giả mạo dữ liệu, kỹ năng, liên chính học thuật, nhận thức

I. Đặt vấn đề

Hiện nay, trong kỷ nguyên bùng nổ thông tin, khoa học và công nghệ phát triển không ngừng đã tạo ra một cơ hội cũng như đã mở ra một không gian rộng lớn cho nghiên cứu. Tuy nhiên, song song

với đó vấn đề liên chính học thuật càng ngày càng có nhiều thách thức lớn, đặc biệt trong lĩnh vực Công nghệ thông tin vì trong lĩnh vực này dữ liệu được tạo ra, sau đó được thực hiện xử lý và phân tích với tốc độ nhanh chóng. Đồng thời, trong lĩnh

¹ Khoa Công nghệ Thông tin, Trường Đại học Mở Hà Nội

vực Công nghệ thông tin thì mã nguồn cũng như các thuật toán được coi là những tài sản trí tuệ cốt lõi vì vậy việc duy trì liên chính học thuật là yêu cầu về mặt đạo đức đồng thời cũng là yếu tố sống còn với uy tín cá nhân các nhà khoa học, là chất lượng của nghiên cứu cũng như là sự tiến bộ chung của toàn xã hội.

Hiện nay, do sự toàn cầu hoá mạnh nên biên giới trong nghiên cứu khoa học đã dần dần được gỡ bỏ. Các nhà nghiên cứu khoa học có xu hướng công bố kết quả trên các tạp chí, trên các hội thảo trong nước cũng như quốc tế uy tín. Từ đó giữ cho việc tăng cường hợp tác, chia sẻ tri thức và nâng cao sự ảnh hưởng của nghiên cứu giữa các nhà khoa học. Đồng thời, dẫn đến áp lực về việc được công bố hay bị đào thải ngày càng lớn của các nhà khoa học. Với áp lực này nhiều khi dẫn đến một số hành vi thiếu liên chính trong khoa học, từ việc đạo văn, rồi làm giả dữ liệu, dẫn đến việc chia nhỏ kết quả nghiên cứu để có nhiều công bố bài báo hơn.

Công nghệ thông tin giữ vai trò trung tâm trong hầu hết các lĩnh vực nghiên cứu. Từ đó đặt ra một số thách thức liên quan đến liên chính học thuật như: Tính toàn vẹn của dữ liệu, Đạo văn mã nguồn, Tính tái tạo của kết quả, Trí tuệ nhân tạo tạo sinh.

Tại Việt Nam, vấn đề liên chính học thuật đã được một số tác giả nghiên cứu ở các góc độ khác nhau. Như nghiên cứu của Đặng Hùng Vũ và Nguyễn Thành Long (năm 2021) tập trung đánh giá nhận thức của sinh viên về môi trường học thuật và hành vi không trung thực. Nguyễn Kiều Dung (năm 2022) đã phân tích các nguyên tắc cơ bản của liên chính học thuật và đề

xuất áp dụng bối cảnh ở Việt Nam. Tuy nhiên các nghiên cứu này mới đang tập trung vào khía cạnh nhận thức và chính sách chung cho các sinh viên. [3][5][7]

II. Phương pháp nghiên cứu

Nhóm tác giả đã thực hiện khảo sát với nhóm đối tượng chính là sinh viên Khoa công nghệ thông tin - Trường Đại học Mở Hà Nội, bao gồm các sinh viên ngành Công nghệ thông tin năm thứ hai và năm thứ ba.

Phiếu khảo sát được thực hiện thông qua Google Form, với khoảng 100 phản hồi từ phía sinh viên, phiếu khảo sát tập trung vào ba nhóm vấn đề chính là Nhận thức, Kỹ năng và Công cụ.

Từ kết quả khảo sát, mặc dù sinh viên bước đầu đã nhận thức được sao chép y nguyên mã nguồn là hành vi vi phạm nghiêm trọng nhưng đến hành vi tinh vi hơn như “thay đổi tên biến, thay đổi tên hàm khi sao chép” thì số lượng sinh viên nhận ra đó là hành vi sai được giảm đi đáng kể. Vậy, điều này cho thấy vẫn còn “vùng xám”, “vùng chưa rõ” trong nhận thức của sinh viên với các cấp độ đạo văn.

III. Kết quả nghiên cứu và thảo luận

3.1. Đề xuất mô hình giải pháp “Nhận thức, Kỹ năng và Công cụ”

3.1.1. Nâng cao nhận thức [3]

Để đối phó với các nhận thức về đạo văn mã nguồn, liên chính dữ liệu và quyền tác giả thì việc nâng cao nhận thức về liên chính học thuật trong môi trường đào tạo Công nghệ thông tin bao gồm sinh viên và giảng viên là một nhiệm vụ cấp bách cần thực hiện.

Để thực hiện giáo dục về liên chính học thuật thì chương trình đào tạo cần tích hợp nội dung về liên chính học thuật, được thiết kế đặc thù riêng của ngành Công nghệ thông tin và lồng ghép xuyên suốt toàn bộ quá trình học tập của sinh viên. Việc tích hợp thực hiện từ năm thứ nhất đến năm cuối của sinh viên với mục đích thường xuyên củng cố kiến thức và nâng cao nhận thức - kỹ năng về liên chính học thuật. Chương trình đào tạo có tích hợp liên chính học thuật cần có đầy đủ các nội dung lý thuyết cốt lõi như các khái niệm về liên chính học thuật, các hành vi vi phạm, các quy định quyền sở hữu trí tuệ, các vấn đề đạo đức dữ liệu và các quy trình trong quá trình công bố các nghiên cứu.

Các nội dung về khái niệm liên chính học thuật cũng như thực hành kỹ năng liên chính học thuật của sinh viên Công nghệ thông tin có thể tích hợp vào một số môn học như Kỹ thuật lập trình cơ sở, Cấu trúc dữ liệu và giải thuật, Cơ sở dữ liệu, Nhập môn khai phá dữ liệu, Đồ án/Khoá luận tốt nghiệp, ... Thường xuyên hướng dẫn sinh viên sử dụng các công cụ quản lý tài liệu tham khảo và các công cụ trích dẫn khi sinh viên viết báo cáo Bài tập lớn các môn trong Chương trình đào tạo cũng như khi sinh viên hoàn thiện báo cáo của Khoá luận/Đồ án tốt nghiệp. Đồng thời cũng chỉ dẫn và hướng dẫn cho sinh viên sử dụng công cụ kiểm tra đạo văn mã nguồn cũng như đạo văn văn bản để sinh viên tránh đạo văn trong quá trình học tập tại khoa.

Tại Khoa Công nghệ Thông Tin - Trường Đại học Mở Hà Nội đã có văn bản hướng dẫn cách viết và cấu trúc báo cáo

Bài tập lớn, Đồ án tốt nghiệp/Khoá luận tốt nghiệp trong đó có hướng dẫn chi tiết cách viết tài liệu tham khảo và cách viết tham chiếu tài liệu tham khảo trong báo cáo. Trong quá trình hướng dẫn sinh viên triển khai thực hiện Bài tập lớn cũng như Đồ án tốt nghiệp/Khoá luận tốt nghiệp các giảng viên hướng dẫn đều yêu cầu chặt chẽ về nguồn gốc tài liệu tham khảo và cách trích dẫn tài liệu tham khảo. Nguyên tắc, cách trình bày và sử dụng tài liệu tham khảo cũng là một trong các tiêu chí đánh giá điểm của Bài tập lớn, Đồ án tốt nghiệp/Khoá luận tốt nghiệp.

3.1.2. Trang bị kỹ năng thực hành liên chính

Kỹ năng nhận thức là đạo đức chỉ hướng cho nhà nghiên cứu, cho người học trong xác định và thực hiện liên chính khoa học. Nhận thức cần đi kèm là năng lực thực hành để có thể thực hiện một cách trung thực và chuyên nghiệp. Khi đó, trang bị kỹ năng giúp cho việc ngăn ngừa các hành vi vi phạm không chủ ý đồng thời xây dựng tác phong làm việc chuyên nghiệp, được coi là yêu cầu nên bắt buộc đối với các nhà nghiên cứu và sinh viên Công nghệ thông tin.

Có bốn kỹ năng cốt lõi hỗ trợ thực hành liên chính gồm: Kỹ năng trích dẫn và ghi nhận nguồn trong môi trường số; Kỹ năng quản lý quy trình làm việc rõ ràng minh bạch; Kỹ năng tương tác có trách nhiệm với Trí tuệ nhân tạo; Kỹ năng tư duy phản biện và giải quyết vấn đề.

Kỹ năng trích dẫn và ghi nhận nguồn trong môi trường số. Trong ngành Công nghệ thông tin, các tài liệu tham khảo không chỉ là sách hay bài báo khoa học mà bao gồm mã nguồn mở, đoạn mã có

sẵn, bộ dữ liệu sẵn, các API, các tài liệu kỹ thuật, ... Vì vậy sinh viên cần được hướng dẫn để ghi nhận nguồn của các mã nguồn được tham khảo hoặc sử dụng. Tương tự khi sử dụng các bộ dữ liệu có sẵn để huấn luyện hay phân tích, sinh viên được hướng dẫn cách tìm và trích dẫn theo định dạng theo yêu cầu của người tạo ra dữ liệu. Đối với việc sử dụng các thư viện hay các khung làm việc thì cần có liệt kê và quản lý các tệp tin của dự án một cách rõ ràng.

Kỹ năng quản lý quy trình làm việc rõ ràng minh mạch. Kỹ năng này giúp cho giảng viên thấy được quá trình sinh viên thực hiện triển khai dự án một cách rõ ràng, thấy được sinh viên đã nỗ lực thực hiện và thể hiện được tiến trình tư duy của sinh viên trong quá trình thực hiện. Để hoàn thành kỹ năng này sinh viên sử dụng Git để ghi nhận nhật ký quá trình làm việc, mỗi khi hoàn thành một chức năng nhỏ hay sửa một lỗi thành công sinh viên thực hiện một commit để lưu lại lịch sử làm việc chi tiết, đồng thời sinh viên cũng cần viết dòng văn bản mô tả commit đầy đủ rõ ràng.

Kỹ năng tương tác có trách nhiệm với trí tuệ nhân tạo. Hiện nay, trí tuệ nhân tạo ngày càng phát triển mạnh mẽ và rộng rãi, việc cấm sử dụng AI là phản tác dụng và khó thực hiện. Vậy cần trang bị kỹ năng cho sinh viên để trở thành người khai thác AI thông thái. Sinh viên cần có một khuôn khổ rõ ràng để phân biệt giữa việc sử dụng AI trong hỗ trợ học tập nghiên cứu và việc lạm dụng AI để gian lận. Sinh viên cũng cần luôn có kiểm chứng và kiểm thử kỹ lưỡng với các đoạn mã nhận được do AI tạo ra, đảm bảo tính chính xác.

Kỹ năng Tư duy phản biện và giải

quyết vấn đề. Kỹ năng này được đánh giá là kỹ năng nền tảng để hạn chế tối đa việc đạo văn, đạo mã nguồn. Bởi vì khi gặp một vấn đề khó khăn phức tạp sinh viên chưa biết cách giải quyết dẫn đến có hành vi gian lận. Vì vậy, để hạn chế tiếp cận hành vi gian lận sinh viên cần thực hiện chia bài toán, vấn đề lớn và phức tạp thành các bài toán con, vấn đề con nhỏ hơn và dễ quản lý hơn. Khi bài toán, vấn đề mà sinh viên có thể tự giải quyết được thì sự tiếp cận sử dụng hành vi gian lận sẽ bị giảm đi nhiều.

3.1.3. Các công cụ hỗ trợ giảng viên phát hiện, phòng ngừa và đánh giá

Trong lĩnh vực Công nghệ thông tin do sự phát triển mạnh mẽ và nhanh chóng của công nghệ cũng như các nguồn tài nguyên trực tuyến thì việc đảm bảo tính liêm chính học thuật trở thành một thách thức từ nhiều phía. Để xác định và đối phó được với các vấn đề đạo văn mã nguồn, liêm chính dữ liệu, gian lận trong thi cử thì giảng viên cần được trang bị các công cụ hỗ trợ thiết yếu. Các công cụ hỗ trợ cho giảng viên bao gồm việc phát hiện các hành vi vi phạm liêm chính học thuật đến việc phòng ngừa vi phạm để có thể đánh giá công bằng và khách quan, chính xác kết quả học tập của sinh viên. [6]

*** Nhóm công cụ phát hiện gian lận**

Các công cụ phát hiện gian lận giúp giảng viên sàng lọc và có thể nhận diện được các dấu hiệu bất thường trong khi chấm một lượng lớn các bài tập của các sinh viên ở các học kỳ các năm học khác nhau. Nhóm công cụ phát hiện gian lận được chia thành hai nhóm chính bao gồm: nhóm công cụ phát hiện đạo văn mã nguồn và nhóm công cụ phát hiện đạo văn

văn bản/nội dung do AI sinh tạo ra. Một số công cụ tiêu biểu để phát hiện đạo văn mã nguồn như MOSS, Jplag, Codequiry, PlagScan,...

Để thực hiện phát hiện gian lận đạo văn mã nguồn, Giảng viên tham khảo các bước[1]:

Bước 1: Chuẩn bị, tập hợp toàn bộ các bài nộp của sinh viên vào một cấu trúc thư mục tương ứng

Bước 2: Chạy công cụ để có kết quả báo cáo tương ứng

Bước 3: Đọc kết quả báo cáo, trong báo cáo có con số tỷ lệ phần trăm tương đồng, giảng viên mở báo cáo song song để đọc đánh giá các vị trí đáng nghi ngờ.

Trong trường hợp để phát hiện đạo văn văn bản/nội dung do AI sinh ra thì sử dụng một số công cụ tiêu biểu như Turnitin/iThenticate, GPTZero, Copyleaks. Để phát hiện đạo văn văn bản/nội dung do AI sinh ra thì giảng viên có thể thực hiện tích hợp sẵn các công cụ hỗ trợ vào các hệ thống hỗ trợ học tập tương ứng đang được triển khai tại khoa và nhà trường như Moodle, Canvas, ... Khi đã tích hợp các công cụ này có thể thiết lập thêm mỗi khi sinh viên nộp bài thì có thể xem trước báo cáo tương đồng bài báo cáo của mình để sinh viên có thể điều chỉnh trước khi giảng viên tiến hành chấm đánh giá. Sau khi đã có báo cáo đánh giá, giảng viên ngoài xem xét con số tỷ lệ phần trăm tương đồng thì cần xem xét chi tiết các đoạn nội dung tương đồng đến từ đâu, khuyến cáo với sinh viên và thực hiện thêm quá trình phân tích sâu kết quả hoặc tổ chức vấn đáp để đánh giá thêm.

*** Nhóm công cụ phòng ngừa và hỗ trợ thực hành liên chính**

Ngoài việc phát hiện vi phạm thì việc sử dụng các công cụ cũng như các nền tảng công nghệ để chủ động phòng ngừa hành vi gian lận và hỗ trợ người học, giảng viên là một khía cạnh không kém phần quan trọng trong việc thực hành liên chính một cách hiệu quả. Các công cụ phòng ngừa này ngoài những giá trị sư phạm mang đến thì mục tiêu là tạo thành thói quen về việc minh bạch cũng như có trách nhiệm từ phía sinh viên và giảng viên.

*** Hệ thống quản lý phiên bản**

Trong quy trình phát triển phần mềm chuyên nghiệp thì hệ thống quản lý phiên bản là công cụ không thể thiếu, bao gồm các Git. Thông qua Git cho phép theo dõi và quản lý mọi thay đổi trong mã nguồn theo thời gian, mỗi một thay đổi đều được lưu trữ đầy đủ thông tin người thực hiện, thời điểm thực hiện, nội dung mô tả thay đổi. Các nền Git phổ biến như GitHub, GitLab, ...

Trong quá trình thực hiện và triển khai công tác giảng dạy, Giảng viên yêu cầu sinh viên nộp bài tập, nộp đồ án môn học, bài tập lớn môn học thông qua nền tảng Git sẽ đảm bảo được liên chính học thuật, bởi vì: Minh bạch hóa quá trình làm việc, cho phép giảng viên truy cập và kiểm tra toàn bộ quá trình lịch sử nộp của một dự án, lịch sử sẽ giúp giảng viên thấy được quá trình phát triển phần mềm của sinh viên, từ khi bắt đầu dự án, các lần chỉnh sửa lỗi, quá trình bổ sung thêm các chức năng đến sản phẩm nộp hoàn thiện cuối cùng. Từ đó cho phép giảng viên đánh giá được tiến độ thực hiện cũng như quá trình

nỗ lực thực hiện của sinh viên thay vì chỉ đánh giá sản phẩm cuối cùng. Đánh giá được đóng góp cá nhân sinh viên trong dự án nhóm, thông qua quá trình nộp bài thì giảng viên có thể đánh giá được mức độ tham gia cũng như đóng góp của từng thành viên nhóm triển khai. Đồng thời các nền tảng Git cung cấp các công cụ thống kê trực quan về sự tham gia đóng góp thành viên từ đó giảng viên có thể nhìn nhận rõ ràng về sự phân chia công việc cũng như trách nhiệm của từng sinh viên trong quá trình thực hiện triển khai dự án, bài tập lớn. Phát hiện các hành vi bất thường và xác định dấu hiệu tiềm ẩn vi phạm, thông qua lịch sử nộp bài thừa thớt hoặc lịch sử nộp đầy đủ hoàn thiện ngay trước khi thời hạn hoàn thành dự án, bài tập lớn kết thúc thì có thể thấy được dấu hiệu của việc sao chép mã nguồn từ nguồn khác cũng như thể hiện quá trình làm việc không có tính kế hoạch cũng như không có sự thực hiện triển khai liên tục, ngược lại khi nhìn thấy quá trình triển khai nộp bài đều đặn với trình tự lần lượt các thay đổi phù hợp đồng thời có sự mô tả rõ ràng thể hiện quá trình triển khai thực tế của sinh viên cho dự án, bài tập lớn. Ngoài ra, việc so sánh về mặt cấu trúc cũng như các thời điểm nộp bài của các dự án, bài tập lớn khác nhau cũng là một căn cứ để tìm ra trường hợp thông đồng hay sao chép cả quá trình làm việc của người khác. Khuyến khích thực hành tốt trong phát triển phần mềm, việc sử dụng Git để thực hiện toàn bộ quá trình thực hiện dự án làm bài tập lớn cho phép sinh viên tiếp cận với việc chia nhỏ dự án bài tập lớn, triển khai mã theo module, thực hiện nộp bài thường xuyên có ý nghĩa có giá trị.

*** Môi trường lập trình tích hợp và các plugin hỗ trợ kiểm tra-thực hiện liên chính**

Hiện nay, các môi trường tích hợp lập trình hiện đại như Visual Studio Code, Eclipse,... là một hệ sinh thái đầy đủ với các plugin được cài thêm vào. Có nhiều plugin có thể hỗ trợ đánh giá liên chính học thuật một cách gián tiếp hay trực tiếp.

Plugin hỗ trợ trích dẫn và ghi nhận nguồn, có thể hỗ trợ các bình luận theo một định dạng chuẩn để cho phép ghi nhận được nguồn gốc của các đoạn mã được tham khảo cũng như tái sử dụng từ các thư viện, diễn đàn, ...

Plugin kiểm tra giấy phép mã nguồn, mỗi khi làm việc với các thư viện cũng như các framework mã nguồn mở thì các giấy phép cần được tuân thủ. Khi đó một số plugin cho phép quét xác định các thành phần phụ thuộc của dự án bài tập lớn để nhận diện được giấy phép được sử dụng và cảnh báo về xung đột giấy phép cũng như các nghĩa vụ pháp lý. Từ đó, giúp sinh viên, giảng viên tránh vô tình vi phạm quyền sở hữu trí tuệ cũng như các điều khoản sử dụng.

Plugin phân tích mã và phát hiện trùng lặp cục bộ, giúp phát hiện các đoạn mã trùng lặp trong một dự án hoặc giữa các file nguồn khác nhau mỗi khi sinh viên nộp bài, cho phép nâng cao nhận thức về tái sử dụng mã nguồn có chủ ý và từ đó khuyến khích viết mã tường minh.

Plugin tích hợp công cụ kiểm tra đạo văn mã nguồn, một số plugin có thể tích hợp các dịch vụ kiểm tra mã nguồn bên ngoài để sinh viên biết được kết quả sản phẩm dự án bài tập lớn có vi phạm liên chính học thuật hay không.

*** Nền tảng học tập và đánh giá tương tác hỗ trợ liên chính**

Hiện nay các nền tảng hỗ trợ học tập và đánh giá trực tuyến đặc biệt trong lĩnh vực lập trình và công nghệ thông tin ngày càng phát triển mạnh mẽ, các nền tảng này cũng cấp tính năng hỗ trợ giảng viên trong việc giao bài, chấm bài tự động cũng như phân tích mã nguồn. Một số nền tảng phổ biến như Gradescope, LeetCode, Hệ thống quản lý học tập trực tuyến (LMS).

Tất cả các nền tảng này giúp cho giảng viên chuẩn hóa được quy trình đánh giá, đồng thời các nền tảng này cũng hỗ trợ giảng viên nhanh chóng có phản hồi đối với sinh viên, cũng giúp giảng viên ban đầu xác định được các trường hợp vi phạm liên chính học thuật.

3.2. Đề xuất khung đánh giá liên chính cho giảng viên

Các công cụ trình bày ở trên hỗ trợ giảng viên trong việc phát hiện các trường hợp vi phạm liên chính học thuật tuy nhiên các công cụ đó đang làm nhiệm vụ chỉ ra được tín hiệu, chỉ ra được các trường hợp có nguy cơ vi phạm liên chính học thuật. Để hỗ trợ giảng viên chúng tôi đề xuất Khung đánh giá liên chính học thuật, khung đánh giá gồm 4 bước [2].

Bước 1: Thiết lập môi trường và quy định minh bạch

Giảng viên đưa ra các quy định rõ ràng về liên chính học thuật trong khi sinh viên thực hiện dự án bài tập lớn, bao gồm: quy định hợp tác làm việc nhóm, quy định trích dẫn nguồn tài nguyên tham khảo, quy định sử dụng AI (nếu có).

Giảng viên cũng đưa ra các yêu cầu bắt buộc mỗi khi nộp bài để có căn

cứ đánh giá xác định liên chính học thuật của sinh viên khi triển khai dự án bài tập lớn, thông thường yêu cầu sinh viên ngoài nộp sản phẩm mã nguồn của bài tập lớn thì nộp kèm đường dẫn toàn bộ kho mã nguồn của quá trình thực hiện dự án bài tập lớn, nộp file mô tả đầy đủ cách thức thực hiện và cấu trúc tổ chức thư mục dự án, nộp file văn bản ngắn gọn chứa tuyên bố liên chính học thuật sinh viên cam kết thực hiện

Bước 2: Sàng lọc tự động và thu thập dữ liệu

Sau khi sinh viên nộp bài, giảng viên sử dụng các công cụ công nghệ để sàng lọc bước đầu các tài nguyên nộp của toàn bộ sinh viên. Tại bước này, ban đầu xác định được các trường hợp có khả năng vi phạm liên chính học thuật trong thực hiện dự án bài tập lớn và giảng viên cần xem xét kỹ lưỡng cẩn thận hơn các trường hợp dự án bài tập lớn này.

Kết quả giảng viên có được một báo cáo kết quả sơ bộ, gồm các dữ liệu thô là căn cứ cho đánh giá phân tích chuyên sâu bước tiếp theo.

Bước 3: Phân tích sâu kết quả sàng lọc

Giảng viên dựa vào báo cáo sơ bộ với các dữ liệu thô ở trên thực hiện quá trình đánh giá kết quả bao gồm việc đối chiếu báo cáo từ công cụ, rồi thực hiện phân tích lịch sử phiên bản thực hiện, đánh giá phong cách thực hiện của sinh viên cũng như mức độ phức tạp của mã nguồn

Bước 4: Đối chất, phản hồi và đưa ra quyết định

Giảng viên chuẩn bị buổi gặp trực tiếp với các sinh viên được đánh giá nguy

cơ vi phạm liêm chính học thuật. Thực hiện trao đổi đối chất với các sinh viên có nguy cơ vi phạm. Căn cứ trên các kết quả trả lời của sinh viên để giảng viên đưa ra kết luận cuối cùng về liêm chính học thuật trong triển khai dự án bài tập lớn đối với sinh viên được xác định có nguy cơ. Và cuối cùng giảng viên phải đưa ra kết luận cuối cùng dành cho sinh viên về vấn đề liêm chính học thuật đối với dự án bài tập lớn đó.

Để có thể đưa ra quyết định và hạn chế các sai sót không đáng có, giảng viên nên lập một Rubric các câu hỏi đánh giá tình huống để xác định tính liêm chính học thuật.

Hiện nay, khi mà Công nghệ thông tin ngày càng bùng nổ thì vấn đề liêm chính học thuật trở thành một thách thức đa chiều và rất cấp bách đặc biệt trong lĩnh vực ngành Công nghệ thông tin nói chung và đào tạo giảng dạy Công nghệ thông tin nói riêng. Việc xây dựng một hệ sinh thái về liêm chính học thuật trong môi trường đào tạo cũng như nghiên cứu của ngành Công nghệ thông tin là yêu cầu cần được đề cao để đảm bảo tính tin cậy cũng như giá trị của một dự án nghiên cứu hay công trình nghiên cứu.

Bài viết này bước đầu đã có những phân tích ban đầu về liêm chính học thuật trong ngành Công nghệ thông tin như đạo văn mã nguồn, liêm chính dữ liệu. Để hạn chế liêm chính học thuật trong đào tạo giảng dạy ngành Công nghệ thông tin, nhóm tác giả đã trao đổi về mô hình giải pháp “Nhận thức - Kỹ năng và Công cụ” với mục đích ban đầu là nâng cao nhận thức cũng như trang bị kỹ năng cơ bản cần thiết cho cả sinh viên và giảng viên thông qua làm quen tiếp cận với một số công cụ hỗ trợ. Bên cạnh đó, để hỗ trợ cho giảng viên có thể đưa ra đánh giá tin cậy khi thực hiện đánh giá sinh viên trong quá trình

giảng dạy, nhóm tác giả đã tổng hợp đưa ra Khung đánh giá liêm chính học thuật gồm 4 bước được đề xuất, khung này giúp hệ thống hoá các bước thực hiện cũng như đảm bảo được tính khách quan và công bằng trong đánh giá.

Tài liệu tham khảo:

- [1]. Aiken, A. (2022). *A system for detecting software similarity*. Stanford University.
- [2]. Bretag, T., et al. (2019). Contract cheating and assessment design: Exploring the relationship. *Assessment & Evaluation in Higher Education*, 44(5), 676 -691. <https://doi.org/10.1080/002602938.2018.1527892>
- [3]. Dung, N. K. (2022). Liêm chính học thuật và những nguyên tắc cơ bản đối với Việt Nam, *Tạp chí Khoa học và Công nghệ Việt Nam*.
- [4]. Mehse, R. S. (2022). Review of source code plagiarism detection techniques. In *Proceedings of the 4th International Conference on Communication Engineering and Computer Science*.
- [5]. Nghĩa, L. T. (2023). Các nguyên tắc dữ liệu FAIR - chuẩn mực cho dữ liệu số để đảm bảo việc chia sẻ, dùng chung mức quốc tế. In *Báo cáo toàn văn hội thảo “Xây dựng trung tâm Tri thức số và mô hình thư viện đại học thông tin”, Giáo dục mở - Tài nguyên giáo dục mở ứng dụng và phát triển*. <https://doi.org/10.5281/zenodo.7949540>
- [6]. Nga, N. T., & Trang, H. T. (2019). Thực hiện liêm chính trong nghiên cứu khoa học tại Việt Nam hiện nay - Thực trạng và giải pháp. *Tạp chí Cộng sản*.
- [7]. Vũ, Đ. H., & Long, N. T. (2021). Đánh giá liêm chính học thuật của sinh viên qua nhận thức của sinh viên về môi trường học thuật và hành vi không trung thực học thuật. *Tạp chí Khoa học Đại học Mở Thành phố Hồ Chí Minh*.

ACADEMIC INTEGRITY IN INFORMATION TECHNOLOGY EDUCATION AND LEARNING: FROM AWARENESS, SKILLS TO SUPPORT TOOLS

Trinh Thi Xuan², Mai Thi Thuy Ha², Nguyen Thuy Linh², Le Thi Thanh Thuy²

Abstract: *In today's rapidly evolving technological landscape, source code plagiarism, data fabrication, and the use of AI-generated content are increasingly common, sophisticated, and hard to detect, particularly among Information Technology students. This article analyzes source code plagiarism, from verbatim copying to subtle modifications, and clarifies the distinction between legitimate use, referencing, and plagiarism within open-source contexts.*

This article addresses both data integrity and potential data bias, whether unintentional or intentional, and introduces the FAIR principles and ALCOA+ criteria to ensure data integrity.

The authors propose an "Awareness - Skills - Tools" model to cultivate academic integrity in Information Technology education. The authors propose a four-step academic integrity framework to help lecturers evaluate student work objectively, transparently, and fairly.

Keywords: *academic integrity, awareness, data fabrication, skills, source code plagiarism*

² Faculty of Information Technology, Hanoi Open University