Khuyến nghị về Việc Sử dụng Trí tuệ Nhân tạo trong Truyền thông Học thuật

Bối cảnh: Sự phát triển gần đây của các Mô hình Ngôn ngữ Lớn (LLMs) và Trí tuệ Nhân tạo Tạo sinh (GenAI) đã mang đến những thách thức và cơ hội mới trong truyền thông học thuật. Điều này dẫn đến sự đa dạng trong các chính sách của các tạp chí, nhà xuất bản và các tổ chức tài trợ liên quan đến việc sử dụng các công cụ AI. Các nghiên cứu, bao gồm cả khảo sát, cho thấy các nhà nghiên cứu đã sử dụng AI ở quy mô đáng kể để tạo hoặc chỉnh sửa bản thảo và báo cáo phản biện. Tuy nhiên, độ chính xác, hiệu quả và khả năng tái lập của AI vẫn còn chưa rõ ràng. Bộ công cụ này nhằm thúc đẩy việc sử dụng AI có trách nhiệm và minh bạch bởi các biên tập viên, tác giả và phản biện, đồng thời cung cấp các liên kết đến các ví dụ về chính sách và thực tiễn hiện tại. Vì AI đang phát triển rất nhanh, chúng tôi sẽ theo dõi và cập nhật các khuyến nghị này khi có thông tin mới. Vui lòng liên hệ với chúng tôi và chia sẻ bất kỳ ý kiến, chính sách hoặc ví dụ nào có thể giúp chúng tôi cải thiện hướng dẫn này.

Khuyến nghị dành cho Biên tập viên, Tạp chí và Nhà xuất bản

- **Xây dựng, công bố và giám sát việc tuân thủ các chính sách liên quan đến AI**
- **Thu thập và công bố các tuyên bố về việc sử dụng AI của tạp chí, tác giả và phản biện**

Chúng tôi khuyến nghị mạnh mẽ rằng các biên tập viên, tạp chí và nhà xuất bản xây dựng các chính sách liên quan đến việc sử dụng AI trong quy trình xuất bản của họ và công bố các chính sách này trên trang web của tạp chí. Ví dụ, các chính sách dành cho tác giả nên được liệt kê trong phần "Hướng dẫn cho Tác giả" hoặc "Hướng dẫn Nộp bài", trong khi các chính sách dành cho phản biện nên nằm trong phần "Hướng dẫn dành cho phản biện". Các chính sách cần được truyền đạt rõ ràng đến tác giả và phản biện thông qua email và trong hệ thống nộp bài trực tuyến. Các chính sách này cũng phải bao gồm thông tin về cách các bên có thể nêu mối lo ngại về khả năng vi phạm chính sách, hậu quả mà bất kỳ bên nào có thể phải đối mặt khi vi phạm, và các phương thức có thể sử dụng để khiếu nại các quyết định của tạp chí liên quan đến những chính sách đó. Ngoài ra, các chính sách nên được bổ sung bằng tài liệu giáo dục hoặc liên kết đến các nguồn thông tin về việc sử dụng AI một cách có trách nhiệm (ví dụ: Hướng dẫn về việc sử dụng AI tạo sinh một cách có trách nhiệm trong nghiên cứu do Diễn đàn Khu vực Nghiên cứu Châu Âu phát triển). Biên tập viên cũng nên cân nhắc công bố thông báo về việc ban hành hoặc cập nhật chính sách AI thông qua các bài xã luận đã xuất bản.

Chúng tôi ghi nhận rằng có thể tồn tại những khó khăn theo lĩnh vực hoặc theo quy trình (ví dụ: hạn chế của hệ thống nộp bài) có thể ảnh hưởng đến việc xây dựng và triển khai các chính sách liên quan đến AI. Tuy nhiên, có thể xây dựng các biểu mẫu kiểm tra và tuyên bố rõ ràng để thu thập thông tin về việc sử dụng AI cũng như bất kỳ xung đột lợi ích nào liên quan đến việc sử dụng đó (ví dụ: tiết lộ nếu công cụ AI được sử dụng do chính nhà xuất bản phát triển).

Chúng tôi khuyến nghị mạnh mẽ rằng mọi thông tin liên quan đến việc sử dụng AI cho một bản thảo cụ thể đều phải được khai báo trong các mục phù hợp của bản thảo, hoặc trong một mục riêng dành cho tuyên bố về AI (xem ví dụ tại đây), hoặc thông qua việc sử dụng nhãn thông tin xuất bản (xem các khuyến nghị chi tiết bên dưới). Cuối cùng, chúng tôi khuyến nghị mạnh mẽ việc giám sát mức độ tuân thủ chính sách AI của tạp chí và cung cấp các báo cáo định kỳ về mức độ tuân thủ đó cũng như bất kỳ trường hợp vi phạm nào.

Khuyến nghị của EASE về các nội dung cần có trong chính sách AI dành cho truyền thông học thuật và cách thức khai báo việc sử dụng AI:

- **Quyền tác giả/Quyền đóng góp:** Chúng tôi khuyến nghị mạnh mẽ rằng AI không nên được liệt kê là đồng tác giả trong các ấn phẩm. Biên tập viên nên tham khảo tài liệu hướng dẫn của Hiệp hội Biên tập Y khoa Thế giới (WAME), Ủy ban Quốc tế về Biên tập Tạp chí Y khoa (ICMJE), hoặc Hiệp hội STM (STM) về chủ đề này. Ví dụ, ICMJE nêu rõ rằng AI không thể là tác giả: *"vì chúng không thể chịu trách nhiệm về độ chính xác, tính toàn vẹn và tính nguyên bản của công trình, và những trách nhiệm này là yêu cầu bắt buộc đối với tư cách tác giả"*.

- **Trích dẫn và tổng quan tài liệu:** Chúng tôi khuyến nghị mạnh mẽ rằng các đầu ra của AI không nên được trích dẫn như nguồn chính để hỗ trợ cho các luận điểm cụ thể. Nghiên cứu cho thấy trích dẫn và thông tin do AI cung cấp có thể không chính xác hoặc bị bịa đặt. Tác giả và phản biện cần được nhắc nhở luôn đọc và kiểm tra nguồn mà họ trích dẫn, vì họ mới là người chịu trách nhiệm về thông tin được trình bày. Ngoài ra, đầu ra AI có thể không tái tạo được tại thời điểm khác, nên biên tập viên có thể cân nhắc việc yêu cầu tác giả lưu lại và đánh dấu thời gian các đầu ra mà họ đề cập hoặc trích dẫn (xem ví dụ về tác giả trình bày phản hồi của ChatGPT cho một chủ đề cụ thể).

- **Thu thập dữ liệu, làm sạch, phân tích và giải thích:** Chúng tôi khuyến nghị mạnh mẽ rằng mọi việc sử dụng AI cho việc thu thập, làm sạch, phân tích hoặc diễn giải dữ liệu phải được khai báo trong phần Phương pháp của bản thảo hoặc trong một mục khai báo AI (xem ví dụ tại đây). Những tuyên bố này nên đi kèm với các chỉ số về độ tin cậy và độ bền vững, cũng như các bước nhằm đảm bảo khả năng tái lập.

- **Tạo dữ liệu và mã:** Chúng tôi khuyến nghị mạnh mẽ rằng mọi việc sử dụng AI để tạo dữ liệu hoặc mã nguồn phải được khai báo trong phần Phương pháp, cũng như trong các mục khai báo dữ liệu và mã nguồn mà một số tạp chí yêu cầu, hoặc trong mục khai báo AI (xem ví dụ tại đây). Biên tập viên cần lưu ý rằng dữ liệu hoặc mã được tạo ra bằng AI có thể là tài nguyên tuyệt vời cho mục đích giáo dục, nhưng cũng có thể bị lạm dụng để tạo dữ liệu giả phục vụ kiểm định giả thuyết hoặc các phân tích khác. Hơn nữa, có thể khó phân biệt giữa tác giả sử dụng AI để tạo (một phần) dữ liệu hoặc mã rồi tự chỉnh sửa và tác giả tự tạo dữ liệu hoặc mã, rồi dùng AI để chỉnh sửa.

- **Trực quan hóa - tạo bảng, hình ảnh, video hoặc các định dạng khác:** Chúng tôi khuyến nghị mạnh mẽ rằng mọi việc sử dụng AI để tạo trực quan hóa phải được khai báo trong phần Phương pháp của bản thảo và trong phần chú thích hoặc mô tả của các đầu ra đó. Các sản phẩm trực quan do AI tạo ra có thể cần được kiểm tra bổ sung để đảm bảo tính hợp lệ, cũng như các bước để bảo đảm khả năng tái lập. Biên tập viên nên tham khảo ví dụ về chính sách cấm sử dụng AI cho các mục đích này và ví dụ về thu hồi bài báo do hình ảnh "vô nghĩa".

- **Viết, chỉnh sửa ngôn ngữ và văn phong:** Chúng tôi khuyến nghị mạnh mẽ rằng cần quy định rõ tác giả có phải khai báo hay không và khai báo như thế nào khi sử dụng AI cho việc viết, chỉnh sửa ngôn ngữ hoặc chỉnh sửa phong cách. Biên tập viên có thể khuyến nghị tác giả khai báo việc sử dụng AI trong phần Lời cảm ơn hoặc một mục khai báo AI (xem ví dụ tại đây). Ngoài ra, biên tập viên có thể quy định rằng việc sử dụng AI tương tự phần mềm kiểm tra chính tả và không cần khai báo. Biên tập viên cần lưu ý rằng rất khó phân biệt giữa: tác giả tạo một phần văn bản bằng AI rồi tự chỉnh sửa, và tác giả tự viết bản nháp ban đầu rồi sử dụng AI để chỉnh sửa. Để tham khảo thêm, có thể xem bài tổng hợp chính sách của các nhà xuất bản trong Scholarly Kitchen (biên soạn mùa xuân 2024).

- **Các mục đích nghiên cứu khác:** Chúng tôi khuyến nghị mạnh mẽ rằng mọi việc sử dụng AI khác, ngay cả khi không nằm trong các mục ở trên, đều phải được khai báo trong phần phù hợp (ví dụ: Lời cảm ơn, Phương pháp, hoặc mục khai báo AI). Ví dụ: tác giả có thể sử dụng các công cụ tự kiểm tra bằng AI để kiểm tra khuyến nghị báo cáo nghiên cứu, tính toàn vẹn của hình ảnh, mã, hoặc dữ liệu.

- **Phản biện:** Chúng tôi khuyến nghị mạnh mẽ rằng cần quy định rõ người phản biện có được phép sử dụng AI hay không, và phân biệt giữa sử dụng AI để chỉnh sửa ngôn ngữ/phong cách cho nội dung phản biện và sử dụng AI để tạo toàn bộ nội dung phản biện. Tạp chí cũng cần quy định liệu họ có sử dụng công cụ kiểm tra xem phần nào của bản phản biện được tạo bởi AI hay không, hậu quả có thể xảy ra, và tác giả nên làm gì khi nghi ngờ việc sử dụng AI trong phản biện. Ví dụ: xem trường hợp tác giả lo ngại rằng phản biện được viết bằng AI, và tài liệu thảo luận của COPE về AI trong việc ra quyết định. Nhiều tạp chí, nhà xuất bản và cơ quan tài trợ đã cấm người phản biện sử dụng AI (ví dụ: Royal Society, National Institutes of Health, Elsevier) do rủi ro thiên lệch, lo ngại bảo mật, và độ chính xác, hiệu quả, và độ tin cậy chưa được kiểm chứng. Tuy nhiên, các lệnh cấm này khó thực thi, và chưa rõ hậu quả cụ thể khi vi phạm. Xem thêm các cân nhắc về sử dụng AI trong phản biện tại đây.

- **Công tác biên tập:** Chúng tôi khuyến nghị mạnh mẽ rằng mọi việc sử dụng AI bởi biên tập viên hoặc nhân viên biên tập phải được công bố trên trang web của tạp chí và trong các thông báo gửi đến tác giả và phản biện, bao gồm việc sử dụng các công cụ kiểm tra xem bản thảo hoặc báo cáo phản biện có được tạo hoặc chỉnh sửa bởi AI hay không. Chúng tôi cũng khuyến nghị các biên tập viên, tạp chí và nhà xuất bản cân nhắc khai báo các kiểm tra được thực hiện bằng AI thông qua nhãn thông tin xuất bản. Cuối cùng, chính sách AI của tạp chí phải bao gồm thông tin về cách các bên có thể nêu lo ngại về khả năng vi phạm chính sách AI, hậu quả khi vi phạm, và các cách thức khiếu nại đối với các quyết định của tạp chí liên quan đến những chính sách này.

Ví dụ về những chính sách về AI (theo thứ tự bảng chữ cái)

- **Nhà xuất bản:** Elsevier, Springer-Nature, Taylor-Francis, Wiley
- **Hiệp hội:** ICMJE, WAME
- **Tạp chí:** PLOS ONE

Đề xuất cho các tác giả:

- **Kiểm tra các chính sách của tạp chí hoặc nhà xuất bản về việc sử dụng AI.**
- **Tuyên bố mọi hình thức sử dụng AI trong quá trình soạn thảo bản thảo, biên tập và chỉnh sửa (bao gồm cả khi soạn phản hồi đối với ý kiến phản biện).**

Trước khi nộp bản thảo, chúng tôi đặc biệt khuyến nghị các tác giả kiểm tra các chính sách của tạp chí hoặc nhà xuất bản liên quan đến việc sử dụng AI trong giao tiếp học thuật (ví dụ: xem trên trang web của tạp chí hoặc liên hệ trực tiếp với tòa soạn). Khi tạp chí chưa có chính sách về AI, hoặc chính sách hiện có không bao quát các khía cạnh sử dụng AI mà bên liên quan quan tâm, chúng tôi đề xuất các tác giả tham khảo và tuân thủ đề xuất của EASE về cách thức tuyên bố việc sử dụng AI. nếu chính sách của tạp chí không phù hợp với cách sử dụng AI của tác giả, chúng tôi khuyên không nên che giấu với tạp chí, mà hãy liên hệ để yêu cầu giải thích hoặc ngoại lệ, hoặc cân nhắc nộp sang các tạp chí khác. Tác giả cần lưu ý rằng tạp chí hoặc đồng tác giả có thể sử dụng các công cụ phát hiện một phần bản thảo được tạo ra hoặc chỉnh sửa bởi công cụ AI. Như một bảng kiểm tự đánh giá, tác giả cũng có thể chủ động sử dụng các công cụ này.

Khi tác giả nghi ngờ có vi phạm chính sách AI của tạp chí, hoặc nghi ngờ báo cáo phản biện hay ý kiến của biên tập viên được tạo bởi AI, chúng tôi đề xuất các tác giả liên hệ với biên tập viên với mô tả rõ ràng về nghi vấn. Tác giả cũng có thể xem xét đưa ra các nhận xét qua công cụ phát hiện AI và đính kèm báo cáo như một dạng bằng chứng về khả năng có sử dụng AI (xem ví dụ một nhà nghiên cứu nêu quan ngại về việc nhận báo cáo phản biện do AI soạn, cũng như tài liệu thảo luận của COPE về AI trong ra quyết định). Nếu việc trao đổi tiếp theo với biên tập viên và văn phòng tòa soạn không được hồi đáp, chúng tôi khuyến nghị tác giả liên hệ nhà xuất bản hoặc hội học thuật của tạp chí, kèm bản sao trao đổi trước đó; nếu không có, có thể liên hệ Ủy ban Đạo đức Xuất bản (COPE) hoặc STM Integrity Hub.

Đề xuất cho các phản biện

- **Kiểm tra các chính sách về AI của tạp chí hoặc nhà xuất bản**
- **Tuyên bố mọi hình thức sử dụng AI trong quá trình soạn thảo và chỉnh sửa báo cáo phản biện.**

Khi nhận được lời mời phản biện từ một tạp chí, phản biện viên nên kiểm tra các chính sách của tạp chí hoặc nhà xuất bản về việc sử dụng AI trong quá trình phản biện (ví dụ: bằng cách xem trang web của tạp chí hoặc nội dung email mời phản biện). Khi tạp chí không có chính sách về AI, hoặc chính sách hiện hành không bao quát các khía cạnh cụ thể liên quan đến việc sử dụng AI, chúng tôi đặc biệt đề xuất phản biện viên nên tham khảo và tuân thủ các đề xuất của EASE về cách thức tuyên bố và sử dụng AI trong quá trình phản biện. Nếu chính sách của tạp chí không phù hợp với cách sử dụng AI của phản biện viên, chúng tôi khuyến nghị không nên che giấu hoặc cố tình vi phạm, mà nên liên hệ với tạp chí để yêu cầu giải thích hoặc ngoại lệ, hoặc xem xét từ chối lời mời phản biện. phản biện viên cần nhận thức rằng tạp chí hoặc tác giả có thể sử dụng các công cụ phát hiện AI để kiểm tra xem một phần báo cáo phản biện có được tạo hoặc chỉnh sửa bằng công cụ AI hay không. Vì vậy, như một bảng kiểm tự đánh giá, phản biện viên cũng có thể tự sử dụng các công cụ này để kiểm tra báo cáo của mình.

Khi phản biện viên nghi ngờ có hành vi vi phạm chính sách AI của tạp chí, hoặc cho rằng bản thảo hoặc nhận xét của một phản biện viên khác được tạo ra bằng công cụ AI, chúng tôi khuyến nghị phản biện viên liên hệ với biên tập viên và mô tả rõ ràng nghi vấn của mình. phản biện viên cũng có thể chạy các nội dung đó qua các công cụ phát hiện việc sử dụng AI và đính kèm báo cáo kết quả trong quá trình trao đổi như bằng chứng tiềm năng về việc sử dụng AI (xem ví dụ về một nhà nghiên cứu đã nêu quan ngại về việc nhận được báo cáo phản biện do AI viết, cũng như tài liệu thảo luận của COPE về AI trong ra quyết định). Nếu việc trao đổi tiếp theo với biên tập viên và văn phòng tòa soạn không nhận được phản hồi, chúng tôi khuyến nghị phản biện viên liên hệ với nhà xuất bản hoặc hiệp hội của tạp chí, kèm theo bản sao nội dung trao đổi trước đó; nếu không có, có thể liên hệ với Ủy ban Đạo đức Xuất bản (COPE) hoặc STM Integrity Hub để được xem xét và hỗ trợ.

Ủy ban phản biện của EASE

Xuất bản ngày Thứ Tư, 25 tháng 9 năm 2024