



PHÁP LUẬT VỀ TRÁCH NHIỆM DÂN SỰ TRONG NGHIÊN CỨU, PHÁT TRIỂN VÀ ỨNG DỤNG TRÍ TUỆ NHÂN TẠO

Nguyễn Bích Thảo*

Tóm tắt: Bài viết phân tích một số thách thức đặt ra đối với pháp luật về trách nhiệm dân sự trong nghiên cứu, phát triển và ứng dụng trí tuệ nhân tạo, phân tích kinh nghiệm lập pháp của Liên minh châu Âu về trách nhiệm dân sự trong kỷ nguyên trí tuệ nhân tạo, từ đó đưa ra một số khuyến nghị cho Việt Nam.

Từ khóa: trí tuệ nhân tạo, trách nhiệm dân sự, bồi thường thiệt hại

1. ĐẶT VẤN ĐỀ

Trách nhiệm dân sự trong nghiên cứu, phát triển và ứng dụng trí tuệ nhân tạo (AI) là một chủ đề mang tính thời sự, thu hút sự quan tâm của các nhà lập pháp, các nhà nghiên cứu, phát triển, ứng dụng AI ở cả khu vực công và khu vực tư cũng như sự quan tâm của toàn xã hội bởi quy định pháp luật rõ ràng về trách nhiệm pháp lý sẽ góp phần tạo sự an tâm cho các chủ thể liên quan đến vòng đời của trí tuệ nhân tạo, tạo niềm tin của xã hội vào công nghệ mới nổi. Trong Quyết định số 127/QĐ-TTg ngày 26/01/2021 của Thủ tướng Chính phủ ban hành Chiến lược quốc gia về nghiên cứu, phát triển, ứng dụng trí tuệ nhân tạo, Bộ Tư pháp được giao nhiệm vụ xây dựng và hoàn thiện bổ sung các văn bản pháp luật về trách nhiệm pháp lý của các đối tượng liên quan đến trí tuệ nhân tạo. Do đó, Bộ trưởng Bộ Tư pháp đã ban hành Quyết định số 552/QĐ-BTP ngày 12/04/2022 về kế hoạch thực hiện chiến lược quốc gia về nghiên cứu, phát triển và ứng dụng Trí tuệ nhân tạo (AI) đến năm 2030 và chiến lược quốc gia về cách mạng công nghiệp lần thứ tư đến năm 2030 của Bộ Tư pháp. Mới đây, Bộ Tư pháp đã xây dựng dự thảo báo cáo “*Xây dựng và hoàn thiện bổ sung các văn bản pháp luật về trách nhiệm pháp lý của các đối tượng liên quan đến trí tuệ nhân tạo*” và đưa ra lấy ý kiến (Báo Pháp luật, 2024).

Trong bối cảnh đó, việc tiếp tục nghiên cứu, nhận diện các thách thức và vấn đề đặt

ra đối với pháp luật về trách nhiệm dân sự trong nghiên cứu, phát triển, ứng dụng AI, tham khảo kinh nghiệm quốc tế trong việc điều chỉnh chính sách, pháp luật đáp ứng yêu cầu của thực tiễn trong bối cảnh phát triển, ứng dụng AI ngày càng rộng rãi có ý nghĩa quan trọng đối với Việt Nam.

2. NHẬN DIỆN MỘT SỐ THÁCH THỨC ĐỐI VỚI PHÁP LUẬT VỀ TRÁCH NHIỆM DÂN SỰ TRONG KỶ NGUYÊN TRÍ TUỆ NHÂN TẠO

Trách nhiệm dân sự là một dạng trách nhiệm pháp lý. Ở Việt Nam, trách nhiệm pháp lý được hiểu là “những hậu quả pháp lý bất lợi về vật chất hoặc tinh thần, được áp dụng bởi các cơ quan nhà nước có thẩm quyền đối với các chủ thể vi phạm pháp luật. Những hậu quả pháp lý bất lợi này là những hình thức cưỡng chế pháp lý được quy định trong bộ phận chế tài của các quy phạm pháp luật tương ứng” (Hoàng Thị Kim Quế, 2015). Trách nhiệm pháp lý thường được phân loại thành: trách nhiệm hình sự, trách nhiệm hành chính, trách nhiệm dân sự, trách nhiệm kỷ luật.

Trách nhiệm dân sự (civil liability) là hậu quả pháp lý bất lợi về vật chất hoặc tinh thần mà bên vi phạm nghĩa vụ dân sự phải gánh chịu, thường là dưới hình thức bồi thường thiệt hại, buộc thực hiện nghĩa vụ hoặc biện pháp khắc phục khác. Trách nhiệm dân sự gồm trách nhiệm phát sinh từ hợp đồng hoặc

* TS., Trường Đại học Luật, Đại học Quốc gia Hà Nội

trách nhiệm ngoài hợp đồng. Trong khuôn khổ bài viết này, tác giả sẽ tập trung vào trách nhiệm dân sự ngoài hợp đồng.

AI đang đặt ra những thách thức mới về trách nhiệm dân sự, chủ yếu là trách nhiệm bồi thường thiệt hại ngoài hợp đồng, ví dụ: Dựa vào học thuyết pháp lý nào để quy trách nhiệm bồi thường thiệt hại? Xác định chủ thể nào có trách nhiệm bồi thường thiệt hại khi xảy ra thiệt hại do AI gây ra? Pháp luật về trách nhiệm dân sự cần được điều chỉnh như thế nào cho phù hợp với kỷ nguyên AI?

2.1. Thách thức về áp dụng học thuyết pháp lý nào đối với trách nhiệm bồi thường thiệt hại do trí tuệ nhân tạo gây ra

Trong khoa học luật dân sự, có các học thuyết (nguyên lý) sau đây được áp dụng đối với trách nhiệm bồi thường thiệt hại ngoài hợp đồng: trách nhiệm dựa trên lỗi cố ý (intentional torts), trách nhiệm dựa trên sự bất cẩn (negligence), trách nhiệm nghiêm ngặt (strict liability). Các yếu tố cấu thành trách nhiệm dựa trên sự bất cẩn bao gồm: (1) nghĩa vụ cẩn trọng hợp lý (duty of reasonable care) của bị đơn; (2) bị đơn đã vi phạm nghĩa vụ đó bằng việc không loại trừ các nguy cơ gây thiệt hại có thể dự liệu trước (breach of duty), (3) có thiệt hại xảy ra (damages) và (4) có mối quan hệ nhân quả giữa sự vi phạm nghĩa vụ của bị đơn và thiệt hại xảy ra. Trách nhiệm nghiêm ngặt là trách nhiệm của chủ thể dù không có lỗi. Trách nhiệm bồi thường thiệt hại do AI gây ra nên được xác định dựa trên nguyên lý nào?

Nếu áp dụng nguyên lý trách nhiệm do lỗi cố ý sẽ rất khó để chứng minh ý định gây hại hoặc ít nhất là hiểu biết cơ bản về thiệt hại sẽ xảy ra ở hiện tại đối với cả nhóm nhà phát triển và nhà triển khai. Nếu áp dụng nguyên lý trách nhiệm dựa trên sự bất cẩn cũng khó chứng minh do thiếu hiểu biết đầy đủ về AI và khả năng tự học, tự điều chỉnh của nó. Hơn thế nữa để áp dụng nguyên lý trách nhiệm do lỗi sơ suất thì tiêu chuẩn về mức độ cẩn trọng hợp lý (standard of care) phải được xác định rõ để từ đó xác định có sự vi phạm tiêu chuẩn

hay không, trong khi đối với AI thì những tiêu chuẩn này vẫn chưa được thiết lập đầy đủ, rõ ràng. Như vậy, trách nhiệm nghiêm ngặt có vẻ như là hợp lý, tuy nhiên sẽ gây ra gánh nặng cho nhà phát triển, nhà cung cấp và triển khai hệ thống AI và không khuyến khích được sự phát triển của các công nghệ mới.

2.2. Thách thức về xác định chủ thể chịu trách nhiệm

Liên quan đến vòng đời của AI có nhiều nhóm chủ thể tham gia, ví dụ: (1) Nhà phát triển (Developer) gồm cá nhân, tổ chức nghiên cứu, thiết kế, phát triển hoặc sản xuất hệ thống AI, (2) Nhà triển khai (deployer) là cá nhân, tổ chức cung cấp dịch vụ/hệ thống AI, sử dụng hoặc triển khai một hệ thống AI do chính họ phát triển hoặc do người khác phát triển, (3) Người dùng (user) là cá nhân, tổ chức tương tác với hệ thống AI hoặc dịch vụ được hỗ trợ bởi AI và có thể bị ảnh hưởng bởi các quyết định của hệ thống đó. (4) người không thuộc 3 nhóm trên nhưng bị thiệt hại do hệ thống AI gây ra. Có ý kiến cho rằng còn có một chủ thể nữa là bản thân thực thể AI.

Ở góc nhìn truyền thống, nhóm thứ hai có nhiều khả năng phải chịu trách nhiệm bồi thường thiệt hại do AI gây ra bởi lẽ họ là chủ thể kiểm soát, điều khiển AI và được hưởng lợi từ việc sử dụng hệ thống AI. Tuy nhiên do đặc tính của AI là có tính tự chủ cao, có khả năng tự học và điều chỉnh hành vi một cách chủ động, các quy tắc trách nhiệm dân sự truyền thống dường như không phù hợp. AI có khả năng hành động độc lập và đưa ra quyết định nằm ngoài sự kiểm soát của con người, có thể gây ra những thiệt hại không dự liệu trước. Trong khi đó, với nhóm nhà phát triển, mặc dù là nhóm chủ thể tạo ra AI nhưng sự ảnh hưởng của nhóm này giảm dần theo thời gian vận hành bởi AI có khả năng tự học hỏi, tự điều chỉnh.

Trong trường hợp nhà triển khai (vận hành hệ thống AI và cung cấp dịch vụ được hỗ trợ bởi hệ thống AI) có lỗi, sử dụng sai mục đích, sai hướng dẫn, gây thiệt hại cho người dùng hoặc người khác, thì có thể dễ dàng quy trách

nhệm bồi thường cho họ. Tuy nhiên, nếu không có hành vi trên thì nhóm nhà phát triển có lẽ nên chịu trách nhiệm trên cơ sở nguyên lý về trách nhiệm sản phẩm (trách nhiệm của nhà sản xuất đối với sản phẩm có khuyết tật). Trên thực tế, đã xuất hiện những vụ kiện dân sự mà trong đó người dùng bị thiệt hại bởi hệ thống AI khởi kiện nhà triển khai (nhà cung cấp dịch vụ được hỗ trợ bởi hệ thống AI), nhưng bị đơn bác bỏ trách nhiệm vì cho rằng đó hoàn toàn là lỗi của hệ thống AI, không phải do con người. Điển hình là vụ Moffat kiện Hãng hàng không Air Canada với phán quyết được đưa ra vào tháng 2/2024, trong đó cơ quan giải quyết tranh chấp tuyên Hãng hàng không Air Canada phải bồi thường cho một hành khách hơn 600 USD do chatbot trả lời tự động của họ đưa thông tin sai về chính sách giảm giá đặc biệt. Năm 2022, Jake Moffatt đặt chuyến bay tới Toronto của hãng Air Canada để dự đám tang bà ngoại. Sau khi tham khảo chatbot trả lời trực tuyến của hãng trên trang web, Jake Moffatt thấy thông tin về chính sách giảm giá đặc biệt cho các hành khách di chuyển để dự tang lễ người thân. Nội dung trên ảnh chụp màn hình từ cuộc trò chuyện cho thấy, trợ lý ảo của hãng Air Canada khuyên Jake Moffatt mua vé với giá thông thường và sẽ có tối đa 90 ngày để yêu cầu giảm giá vì lý do dự tang lễ. Jake Moffatt làm theo hướng dẫn, đặt vé đến và đi từ Toronto để dự đám tang. Tuy nhiên, khi ông nộp đơn xin hoàn tiền, hãng Air Canada thông báo rằng họ sẽ không trả lại tiền tang lễ sau khi đã mua vé. Sau đó Jake Moffatt dùng ảnh chụp màn hình với chatbot để khiếu nại với hãng. Đáp lại, hãng hàng không lớn nhất Canada chỉ thừa nhận chatbot đã sử dụng “những từ ngữ gây hiểu lầm” và hứa sẽ cập nhật trợ lý ảo. Không hài lòng, Jake Moffatt đã khởi kiện hãng hàng không về khoản chênh lệch giá vé. Hành động sau đó của Air Canada gây tranh cãi, khi lập luận mặc dù có lỗi nhưng chatbot là một “thực thể pháp lý riêng biệt”, do đó hãng từ chối hoàn tiền. Cơ quan giải quyết tranh chấp dân sự (CRT) của bang British Columbia không đồng ý với Air

Canada, và đồng ý với cáo buộc của nguyên đơn về việc bị đơn, thông qua chatbot của mình, đã có hành vi cung cấp thông tin sai lệch một cách cẩu thả (negligent misrepresentation) - phát sinh theo luật pháp Canada khi người bán không thực hiện sự cẩn trọng hợp lý để đảm bảo các tuyên bố của mình là chính xác và không gây hiểu lầm (European Parliament, 2017a) và tuyên bố hãng hàng không này phải có nghĩa vụ thông tin chính xác cho khách hàng. Phán quyết nhận định rằng chuẩn mực ứng xử cẩn trọng (standard of care) yêu cầu hãng hàng không phải có sự cẩn thận hợp lý để đảm bảo những phát ngôn của họ không gây hiểu lầm. Chatbot là một phần của trang web chứ không phải một chủ thể pháp luật độc lập. Air Canada phải thấy rõ rằng họ chịu trách nhiệm về tất cả thông tin trên trang web của mình, bất kể thông tin đó xuất hiện trên trang web “tĩnh” (static page) hay được cung cấp bởi chatbot. Phán quyết cũng bác bỏ lập luận của bị đơn rằng Moffatt phải xác minh, đối chiếu thông tin do chatbot cung cấp với thông tin đúng trên trang web của Air Canada, bởi như vậy sẽ tạo nên gánh nặng quá mức cho người tiêu dùng. Phán quyết buộc hãng hàng không Air Canada phải bồi thường cho ông Moffatt tổng cộng hơn 600 USD (Tạp chí Tài chính doanh nghiệp, 2024).

Như vậy, trong vụ án trên, cơ quan giải quyết tranh chấp vẫn áp dụng nguyên lý truyền thống về trách nhiệm bồi thường thiệt hại do cẩu thả, sơ suất (negligence) và bác bỏ lập luận mới đổ lỗi cho AI, không công nhận AI là một thực thể pháp lý độc lập. Đây là quan niệm phổ biến hiện nay, dù năm 2017, Nghị viện châu Âu đã có một bước đi khá táo bạo và đột phá khi thông qua Nghị quyết kêu gọi Ủy ban châu Âu thiết lập một “tư cách pháp lý riêng cho robot trong dài hạn, để ít nhất các robot tự trị phát triển ở mức độ tinh vi nhất có thể được thừa nhận tư cách là điện tử nhân, chịu trách nhiệm bồi thường đối với thiệt hại do chúng gây ra, và có thể áp dụng tư cách điện tử nhân đó đối với những trường hợp robot đưa ra các quyết định độc lập hoặc

giao thiệp với bên thứ ba một cách độc lập” (European Parliament, 2017b). Đây là một trong các văn bản có tính quốc tế đầu tiên đặt ra vấn đề trao tư cách chủ thể cho AI, tuy nhiên văn bản này cũng làm dấy lên rất nhiều tranh cãi gay gắt. Tính tự trị, độc lập và khó dự đoán trước của AI đặt ra vấn đề liệu có nên thiết lập một loại chủ thể pháp luật mới, bởi nếu áp dụng các quy tắc pháp lý hiện hành thì không giải quyết được thấu đáo một số trường hợp cụ thể như xác định chủ thể chịu trách nhiệm bồi thường thiệt hại ngoài hợp đồng, tạo ra “khoảng trống về trách nhiệm pháp lý” (responsibility gap). Một số quan điểm cho rằng trao tư cách chủ thể cho AI là một giải pháp khả quan để lấp đầy khoảng trống nói trên (Schirmer, 2020). Khi AI ngày càng mang tính tự trị, chúng sẽ không đơn thuần chỉ là công cụ trong tay các chủ thể khác như người sản xuất, vận hành, sở hữu, hay sử dụng chúng. Do vậy, câu hỏi đặt ra là liệu các quy tắc pháp lý hiện hành về trách nhiệm có đủ sức điều chỉnh trong bối cảnh mới, hay cần thiết lập các nguyên tắc và quy tắc mới để làm rõ trách nhiệm pháp lý của các chủ thể liên quan đối với hành vi của robot trong trường hợp khó có thể truy cứu trách nhiệm đối với một chủ thể cụ thể nào đó.

Tuy nhiên, đa số quan điểm hiện nay khá đồng thuận rằng pháp luật không cần bổ sung một loại chủ thể mới là thực thể AI. Từ xưa đến nay, pháp luật luôn dựa trên nền tảng lấy con người làm trung tâm, đề cao các giá trị của con người. Việc ghi nhận AI có tư cách chủ thể sẽ hạ thấp vai trò, giá trị của con người cá nhân. Chiến lược quốc gia về nghiên cứu, phát triển và ứng dụng AI của Việt Nam đến năm 2030 cũng đã khẳng định quan điểm: *“Phát triển và ứng dụng trí tuệ lấy con người và doanh nghiệp làm trung tâm, tránh lạm dụng công nghệ và xâm phạm quyền, lợi ích hợp pháp của tổ chức, cá nhân”* (Thủ tướng Chính phủ, 2021). Hơn nữa, tiếp cận từ góc độ chủ thể gánh chịu trách nhiệm, đối với thiệt hại do AI gây ra, không cần thiết quy định thêm tư cách chủ thể cho AI bởi có thể quy trách nhiệm cho các chủ thể truyền thống

(cá nhân, pháp nhân) theo khuôn khổ pháp luật hiện hành. Trách nhiệm dân sự chủ yếu là trách nhiệm về tài sản, đòi hỏi chủ thể chịu trách nhiệm phải có tài sản. Tuy nhiên, thiệt hại gây ra bởi các hệ thống công nghệ, kể cả hoàn toàn tự trị, cũng có thể được quy trách nhiệm về thể nhân hoặc pháp nhân thông qua các nguyên tắc trách nhiệm thay thế (vicarious liability) - là một học thuyết quy trách nhiệm bồi thường thiệt hại cho người chủ, người sử dụng lao động đối với hành vi gây thiệt hại do người làm công, người đại diện (agent) của mình gây ra - theo Bộ luật Dân sự số 91/2015/QH13 (Quốc hội, 2015), trách nhiệm nghiêm ngặt (strict liability). Việc thừa nhận tư cách chủ thể của AI đòi hỏi AI phải có tài sản để có thể trở thành chủ thể chịu trách nhiệm bồi thường và từ đó người bị thiệt hại có thể khởi kiện AI ra trước tòa án. Vậy thực chất AI sẽ là chủ thể đứng đằng sau tạo lập khối tài sản cho AI hay đóng góp vào một quỹ, hay mua bảo hiểm để AI có thể thực hiện được trách nhiệm bồi thường của mình? (Expert Group on Liability and New Technologies, 2019). Suy cho cùng, vẫn sẽ quy về trách nhiệm của thể nhân hoặc pháp nhân có mối liên hệ gần nhất với AI.

3. KINH NGHIỆM XÂY DỰNG PHÁP LUẬT VỀ TRÁCH NHIỆM DÂN SỰ CỦA LIÊN MINH CHÂU ÂU TRONG NGHIÊN CỨU, ỨNG DỤNG TRÍ TUỆ NHÂN TẠO

Trong hai năm gần đây, Liên minh châu Âu đã có hai sáng kiến lập pháp đáng chú ý liên quan đến trách nhiệm dân sự trong nghiên cứu, phát triển, ứng dụng AI: Chỉ thị trách nhiệm sản phẩm của EU số 2024/2853 và Dự thảo Chỉ thị về trách nhiệm pháp lý liên quan đến AI.

3.1. Chỉ thị về Trách nhiệm sản phẩm số 2024/2853

Chỉ thị về Trách nhiệm sản phẩm số 2024/2853 áp đặt trách nhiệm pháp lý nghiêm ngặt (trách nhiệm không dựa trên lỗi) đối với

các nhà sản xuất, nhà cung cấp và các thực thể khác đối với các sản phẩm khuyết tật. Chỉ thị này thay thế cho Chỉ thị trước đây về trách nhiệm sản phẩm của EU đã tồn tại từ năm 1985. Sau khi được công bố trên Công báo EU vào ngày 18 tháng 11 năm 2024, Chỉ thị này có hiệu lực vào ngày 8 tháng 12 năm 2024. Các quốc gia thành viên sẽ có thời hạn đến ngày 9 tháng 12 năm 2026 để triển khai mới vào luật quốc gia.

Chỉ thị mới về Trách nhiệm sản phẩm được ban hành nhằm đáp ứng yêu cầu mới của nền kinh tế số và kinh tế tuần hoàn. Chỉ thị được xây dựng theo xu hướng tăng cường bảo vệ quyền lợi người tiêu dùng. Người tiêu dùng bị thiệt hại do sản phẩm khuyết tật không an toàn sẽ dễ dàng hơn trong việc yêu cầu bồi thường thiệt hại tại tòa án. Mặt khác, các nhà sản xuất sẽ được hưởng lợi từ các quy tắc rõ ràng liên quan đến sản phẩm kỹ thuật số và mô hình kinh doanh kinh tế tuần hoàn (The Council of European Union, 2024).

Chỉ thị mới mở rộng định nghĩa về “sản phẩm” sang các tệp và phần mềm sản xuất kỹ thuật số. Các nền tảng trực tuyến cũng có thể phải chịu trách nhiệm đối với sản phẩm lỗi được bán trên nền tảng của họ. Đây là điểm mới căn bản của Chỉ thị nhằm thích ứng với nền kinh tế số, khi hầu hết các sản phẩm hữu hình được tích hợp các phần mềm, trong đó có các hệ thống AI. Trong khi đó, Chỉ thị năm 1985 về trách nhiệm sản phẩm chỉ áp dụng đối với sản phẩm hữu hình.

Khi một sản phẩm được sửa chữa và nâng cấp ngoài tầm kiểm soát của nhà sản xuất ban đầu, công ty hoặc cá nhân đã sửa đổi sản phẩm đó phải chịu trách nhiệm. Đây là quy định mới nhằm khuyến khích phát triển nền kinh tế tuần hoàn, tạo lập quy chế pháp lý rõ ràng cho các nhà sản xuất.

Người tiêu dùng được tạo điều kiện thuận lợi hơn cho việc tiếp cận chứng cứ để chứng minh yêu cầu bồi thường của họ với quy định mới của Chỉ thị, theo đó, người khởi kiện có quyền yêu cầu tiếp cận chứng cứ có liên quan do nhà sản xuất cung cấp.

Để đảm bảo rằng người tiêu dùng được bồi thường cho những thiệt hại do sản phẩm được sản xuất bên ngoài lãnh thổ EU gây ra, công ty nhập khẩu sản phẩm hoặc đại diện có trụ sở tại EU của nhà sản xuất nước ngoài có thể phải chịu trách nhiệm về thiệt hại.

Khi người tiêu dùng bị thiệt hại phải đối mặt với những khó khăn quá mức trong việc chứng minh sản phẩm có khuyết tật hoặc chứng minh mối liên hệ nhân quả giữa khuyết tật của sản phẩm và thiệt hại, tòa án có thể quyết định rằng người yêu cầu bồi thường chỉ phải chứng minh khả năng sản phẩm có thể bị khuyết tật hoặc lỗi của sản phẩm là nguyên nhân có thể gây ra thiệt hại.

3.2. Dự thảo Chỉ thị về trách nhiệm dân sự liên quan đến trí tuệ nhân tạo (Proposed AI Liability Directive - AILD)

Dự thảo Chỉ thị lần đầu tiên được công bố vào tháng 9 năm 2022, đề xuất một chế độ trách nhiệm dân sự dựa trên lỗi (fault-based liability) đối với AI để bổ sung cho Đạo luật về AI của EU và Chỉ thị mới về trách nhiệm sản phẩm. Dự thảo này vẫn đang được EU xem xét. Chỉ thị áp dụng đối với thiệt hại gây ra bởi đầu ra của hệ thống AI hoặc do hệ thống AI không tạo ra được kết quả đầu ra.

Tuy một số ý kiến cho rằng dự thảo Chỉ thị về trách nhiệm dân sự liên quan đến AI chông chéo với Chỉ thị về trách nhiệm sản phẩm, nhưng các ý kiến khác cho rằng hai Chỉ thị này không hoàn toàn đồng nhất về phạm vi áp dụng và có thể bổ sung cho nhau. Trong khi Chỉ thị về trách nhiệm sản phẩm quy định chế độ trách nhiệm pháp lý nghiêm ngặt đối với các sản phẩm có khuyết tật (tức là người yêu cầu bồi thường không cần phải chứng minh lỗi/sự bất cẩn của nhà sản xuất). Phiên bản dự thảo hiện hành cho phép nguyên đơn yêu cầu bồi thường thiệt hại dựa trên lỗi phát sinh ngoài hợp đồng (lỗi bất cẩn) liên quan đến thiệt hại do AI gây ra. Dự thảo được xây dựng dựa trên quan điểm của Ủy ban Châu Âu cho rằng các quy tắc trách nhiệm dân sự hiện hành trong pháp luật quốc gia không đủ sức giải quyết những thách thức mới phát sinh liên quan đến trách nhiệm pháp lý đối với thiệt hại

do các sản phẩm và dịch vụ dựa trên AI gây ra do không thể tính đến các tính chất phức tạp và tự chủ của AI, bao gồm các tính chất không rõ ràng của các hệ thống "hộp đen".

Dự thảo Chỉ thị AILD đưa ra các quy định tạo điều kiện thuận lợi để người yêu cầu bồi thường thực hiện được nghĩa vụ chứng minh vốn khó khăn, phức tạp trong các vụ kiện yêu cầu bồi thường thiệt hại dựa trên nguyên lý trách nhiệm do lỗi bất cẩn (negligence). Dự thảo đưa ra một giả định có thể bác bỏ (rebuttable presumption) về mối quan hệ nhân quả giữa lỗi hệ thống AI và thiệt hại gây ra, giúp nguyên đơn có thể dễ dàng chứng minh yếu tố này. Nguyên đơn phải chứng minh rằng: Có khả năng cao là sự bất cẩn của bị đơn đã ảnh hưởng đến kết quả đầu ra của hệ thống AI hoặc không tạo ra được kết quả đầu ra; Kết quả đầu ra do hệ thống AI tạo ra hoặc việc không tạo ra được kết quả đầu ra đã gây thiệt hại cho nguyên đơn. Có thể phát sinh giả định về mối quan hệ nhân quả khi các nhà cung cấp và triển khai hệ thống AI có rủi ro cao không tuân thủ các yêu cầu theo Đạo luật về AI. Đối với các hệ thống không được coi là “rủi ro cao” theo Đạo luật AI, giả định này sẽ chỉ áp dụng khi tòa án quốc gia cho rằng nguyên đơn khó chứng minh được mối quan hệ nhân quả. Giả định này sẽ không áp dụng nếu bị đơn có thể chứng minh rằng nguyên đơn có đủ chứng cứ và kiến thức chuyên môn để chứng minh mối quan hệ nhân quả giữa lỗi của bị đơn và kết quả đầu ra/không có kết quả đầu ra dẫn đến thiệt hại.

Một điểm mới thứ hai tạo điều kiện thuận lợi cho người yêu cầu bồi thường là nguyên đơn có quyền tiếp cận chứng cứ từ các nhà cung cấp hệ thống AI hoặc người dùng hệ thống AI có rủi ro cao. Quyền này có thể bao gồm nhiều thông tin, bao gồm các tài liệu cụ thể như hướng dẫn sử dụng và yêu cầu ghi nhận ký. Yêu cầu cung cấp chứng cứ phải cần thiết và tương xứng để hỗ trợ cho yêu cầu bồi thường thiệt hại. Tòa án quốc gia sẽ phải xem xét lợi ích hợp pháp của tất cả các bên, bao gồm bảo vệ bí mật kinh doanh hoặc bất kỳ thông tin nào có thể ảnh hưởng đến an ninh quốc gia.

Tuy nhiên, tháng 02/2025, Ủy ban Châu Âu đã quyết định rút dự thảo Quy định này

khỏi chương trình làm việc năm 2025 vì không đạt được sự đồng thuận của các quốc gia thành viên do tác động của làn sóng kêu gọi đơn giản hóa quy định về trí tuệ nhân tạo từ phía các doanh nghiệp công nghệ nhằm thúc đẩy đổi mới sáng tạo.

4. KẾT LUẬN VÀ KHUYẾN NGHỊ ĐỐI VỚI VIỆT NAM

Pháp luật về trách nhiệm dân sự trong nghiên cứu, phát triển, ứng dụng AI đang đối mặt với nhiều thách thức, trong đó nổi bật là thách thức về áp dụng học thuyết (nguyên lý) nào để quy trách nhiệm bồi thường thiệt hại do AI gây ra và xác định chủ thể nào chịu trách nhiệm. Từ thực tiễn lập pháp của Liên minh châu Âu và thực tiễn xét xử ở một số quốc gia, có thể rút ra một số gợi ý mở cho Việt Nam sau đây:

Về cơ bản, cơ quan giải quyết tranh chấp vẫn có thể áp dụng các nguyên lý truyền thống của pháp luật về bồi thường thiệt hại ngoài hợp đồng như trách nhiệm dựa trên lỗi bất cẩn và trách nhiệm nghiêm ngặt (thường được áp dụng đối với lĩnh vực “trách nhiệm sản phẩm” - trách nhiệm của nhà sản xuất, kinh doanh đối với thiệt hại do sản phẩm khuyết tật gây ra cho người tiêu dùng) để giải quyết vấn đề bồi thường thiệt hại do AI gây ra.

Tuy nhiên, pháp luật bồi thường thiệt hại ngoài hợp đồng vẫn cần một số điều chỉnh nhất định để thích ứng với kỷ nguyên AI. Thứ nhất, cân nhắc kết hợp giữa chế độ trách nhiệm do lỗi bất cẩn và chế độ trách nhiệm nghiêm ngặt (không dựa trên lỗi), trong đó trách nhiệm nghiêm ngặt chỉ áp dụng đối với trường hợp sản phẩm có khuyết tật. Thứ hai, đối với pháp luật về trách nhiệm sản phẩm, nên xem xét mở rộng khái niệm “sản phẩm” bao gồm cả các sản phẩm phần mềm độc lập, trong đó có hệ thống AI. Thứ ba, điều chỉnh lại các quy định về nghĩa vụ chứng minh trong các vụ kiện liên quan đến trách nhiệm sản phẩm và AI, nhất là chứng minh mối quan hệ nhân quả, để tạo thuận lợi hơn cho người tiêu dùng.

7 -, /, Ê 8 7 + \$ 0 . + ! 2

%iR 3KiS OX±W *yS ê KRjQ WKLEQ TX\ ýİQK SKiS İÖXİWDQĀ WU
ýĀQ WUt W XEWQİSKQ VEDR S KXS Ç X S WQ YjQ

Council of the European Union. (2024, October 10). EU brings product liability rules in line with digital age and circular economy. <https://www.consilium.europa.eu/en/press/releases/2024/10/10/eu-brings-product-liability-rules-in-line-with-digital-age-and-circular-economy/>

European Parliament. (2017a, February 16). European Parliament resolution of 16 February 2017 with recommendations to the Commission on Civil Law Rules on Robotics (2015/2103(INL)). https://www.europarl.europa.eu/doceo/document/ST-2017-0051_EN.html

(XURSHDQ 3DUOLDPHQW E)HEUXDU\ (XURSHDQ 3DUOLDP
UHFRPPHQGDWLRQV WR WKH &RPPLVVLRQ RQ &LYLO /DZ 5XO
KWWSV ZZZ HXURSDUO HXURSD HX SRUWDO HQ

+RjQJ 7Kİ .LP 4XĀ FKē ELrQ *LiR WUuQ% OĖLOKİRQ 4QKjFQİİDF+Y
4XÖF KÝL %ŶYOX±W 'kQ+ Vō

Schirmer, JE. (2020). \$UWLILFLDO ,QWHOOLJHQFH DQG /HJDO 3HUVRQDOLW\ /HJDO 6WDWXV ODGH LQ *HUPDQ\ 7URQJ 7 :L Regulating Artificial 7 5DG intelligence(tr. 123-142). Springer https://doi.org/10.1007/978-030-323615_6

7Kē WmßQJ &KtQK SKē &KLĀQ OmçF TXÖF JLD YĀ QJKLrQ FíX
9LEW 1DP ýĀQ Q P EDQ KJQK NqP WK7H QXjĀW ýİQ% XÖtQK
WUS TXYF JLD 6y WK5W

7¥S FKt 7jL FKtQK GRDQK QJKLEŠ &KDWERW KíD KmR KmQ
KWWSV WDLFKLQKGRDQKQ\$KLHS QHW YQ

LAW ON CIVIL LIABILITY IN RESEARCH, DEVELOPMENT AND APPLICATION OF ARTIFICIAL INTELLIGENCE

Nguyen Bich Thảo

Abstract This paper analyzes some challenges for the law on civil liability in research, development, and application of artificial intelligence, analyzes the legislative experience of the European Union on civil liability in the era of artificial intelligence and makes some recommendations for Vietnam.

Keywords artificial intelligence, civil liability, compensation for damages

* Dr., University of Law, Vietnam National University, Hanoi