

RỦI RO TỪ ỨNG DỤNG TRÍ TUỆ NHÂN TẠO TẠO SINH TRONG DỊCH VỤ TÀI CHÍNH

● NGUYỄN THỊ KIM THOA

Khoa Tài chính Ngân hàng và Bảo hiểm, Trường Đại học Kinh tế - Kỹ thuật công nghiệp

TÓM TẮT:

Việc ứng dụng trí tuệ nhân tạo tạo sinh (GenAI) trong dịch vụ tài chính đang mở ra nhiều cơ hội nâng cao hiệu quả hoạt động, cá nhân hóa dịch vụ và hỗ trợ ra quyết định. Tuy nhiên, đi kèm với những lợi ích đó là các rủi ro ngày càng rõ nét liên quan đến độ tin cậy của thông tin, bảo mật dữ liệu, trách nhiệm pháp lý và đạo đức nghề nghiệp. Bài viết tập trung làm rõ những rủi ro tiềm ẩn khi các tổ chức tài chính triển khai GenAI trong thực tiễn, đồng thời phân tích tác động của các rủi ro này đến hoạt động kinh doanh và niềm tin của khách hàng. Trên cơ sở đó, bài viết nhấn mạnh sự cần thiết của việc xây dựng cơ chế quản trị phù hợp nhằm đảm bảo việc ứng dụng công nghệ mới được kiểm soát hiệu quả và phát triển theo hướng bền vững.

Từ khóa: Generative AI (GenAI), rủi ro dịch vụ tài chính, bảo mật dữ liệu, minh bạch và giải trình, quản trị công nghệ.

1. Đặt vấn đề

Sự phát triển nhanh chóng của trí tuệ nhân tạo (AI), đặc biệt là GenAI, đang tạo ra những thay đổi sâu sắc trong cách thức vận hành của ngành dịch vụ tài chính. Các ngân hàng, công ty bảo hiểm và tổ chức tài chính ngày càng sử dụng công nghệ này trong nhiều hoạt động như chăm sóc khách hàng, phân tích dữ liệu lớn, đánh giá tín dụng, phát hiện gian lận và hỗ trợ ra quyết định. Nhờ khả năng xử lý khối lượng thông tin lớn và tạo ra nội dung mới trong thời gian ngắn, GenAI được kỳ vọng sẽ góp phần nâng cao năng suất lao động và giảm chi phí vận hành.

Tuy nhiên, cùng với tốc độ ứng dụng ngày càng

nhANH, các rủi ro liên quan đến công nghệ này cũng dần bộc lộ rõ nét. Không giống như các hệ thống tự động hóa truyền thống, GenAI hoạt động dựa trên các mô hình học sâu với cấu trúc phức tạp, khó giải thích và phụ thuộc lớn vào dữ liệu huấn luyện. Điều này đặt ra những thách thức đáng kể đối với tính minh bạch, độ tin cậy của thông tin và trách nhiệm giải trình, đặc biệt trong lĩnh vực tài chính vốn chịu sự giám sát chặt chẽ của cơ quan quản lý và có mức độ rủi ro cao. Trong bối cảnh đó, việc phân tích một cách hệ thống các rủi ro từ ứng dụng GenAI trong dịch vụ tài chính là cần thiết nhằm hỗ trợ các tổ chức tài chính nhận diện và quản lý hiệu quả các nguy cơ tiềm ẩn.

2. Thực trạng rủi ro từ ứng dụng GenAI trong dịch vụ tài chính

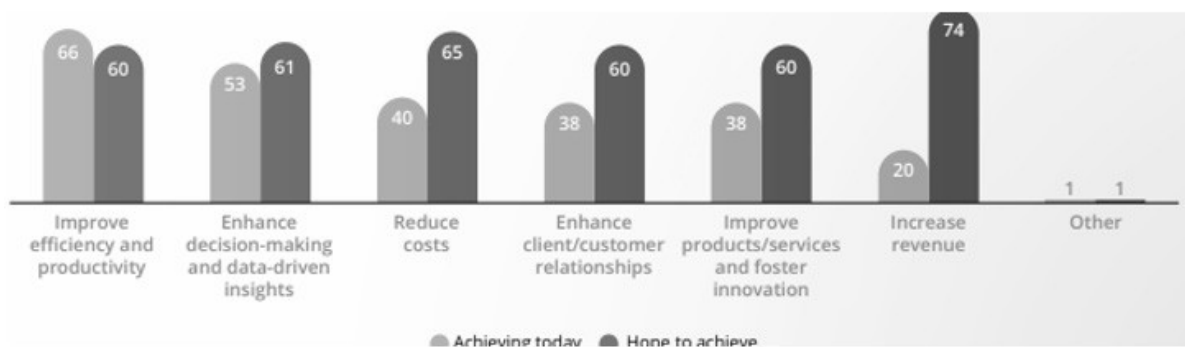
Trong bối cảnh chuyển đổi số diễn ra mạnh mẽ, GenAI đang được các tổ chức tài chính triển khai ngày càng rộng rãi nhằm nâng cao hiệu quả vận hành, cải thiện trải nghiệm khách hàng và hỗ trợ ra quyết định. GenAI được ứng dụng trong nhiều khâu của chuỗi giá trị dịch vụ tài chính như chăm sóc khách hàng thông qua chatbot thông minh, phân tích hồ sơ tín dụng, phát hiện gian lận, hỗ trợ tư vấn đầu tư và xây dựng báo cáo tự động. Theo báo cáo Deloitte's State of Generative AI in the Enterprise Quarter four report, 40% doanh nghiệp được khảo sát cho rằng họ đã đạt được lợi ích cải thiện hiệu quả và năng suất khi ứng dụng GenAI. Báo cáo State of AI in the Enterprise The untapped edge vào tháng 1/2026, tác động thực tiễn của AI đối với kinh doanh đang tăng nhanh, với 25% các nhà lãnh đạo hiện báo cáo rằng AI đang có tác động chuyển đổi đối với công ty của họ, gấp hơn 2 lần so với 12% một năm trước.

Tuy nhiên, song hành với tốc độ ứng dụng nhanh là sự gia tăng đáng kể các rủi ro tiềm ẩn, đặt ra thách thức lớn đối với quản trị rủi ro, tuân thủ pháp luật và bảo vệ lợi ích của khách hàng. Các nghiên cứu gần đây của Deloitte cho thấy phần lớn tổ chức tài chính thừa nhận GenAI mang lại tiềm năng lớn nhưng đồng thời cũng xác định đây là một trong những công nghệ có mức độ rủi ro cao nhất nếu thiếu cơ chế kiểm soát phù hợp.

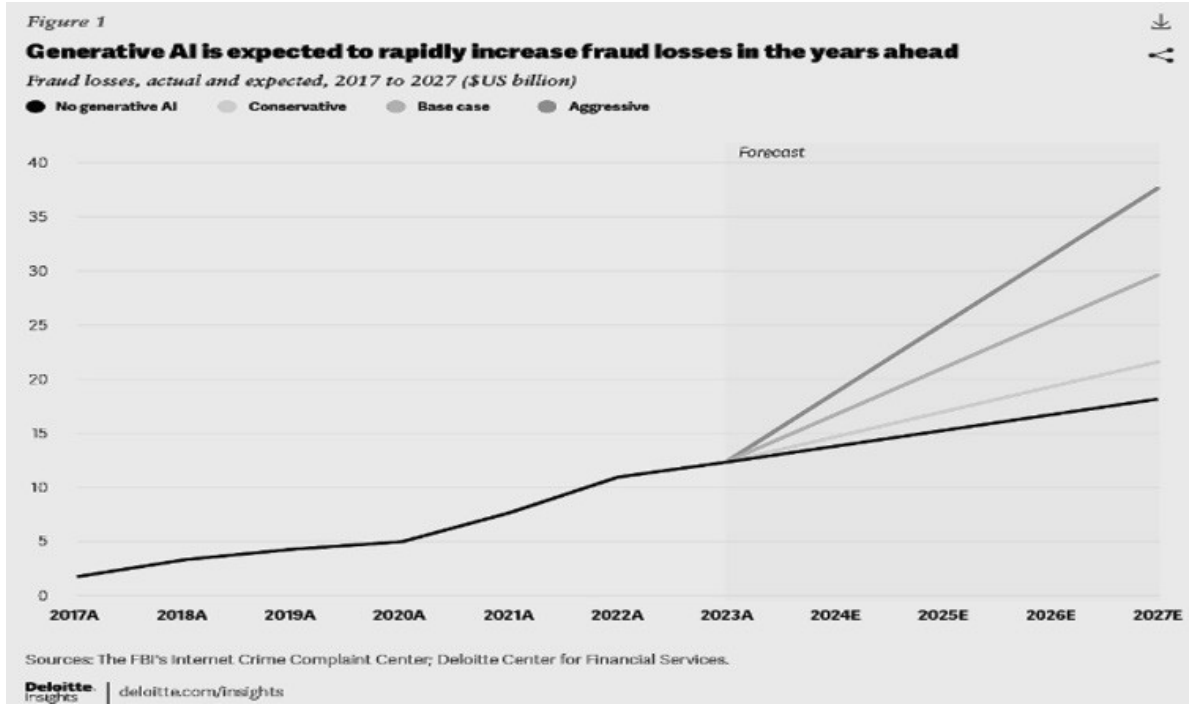
Báo cáo “New Deloitte survey finds expectations for Gen AI remain high, but many are feeling pressure to quickly realize value while managing risks” từ Deloitte cho biết khoảng 75% lãnh đạo doanh nghiệp cho rằng GenAI sẽ chuyển đổi hoạt động của tổ chức trong vòng 3 năm tới, nhưng chỉ khoảng 25% tin rằng tổ chức đã chuẩn bị đầy đủ để quản lý rủi ro liên quan tới GenAI. Ngoài ra, hơn một nửa người được khảo sát lo ngại rằng thiếu khả năng quản trị GenAI có thể làm tăng bất bình đẳng hoặc tạo ra rủi ro không mong muốn.

Một trong những rủi ro nổi bật nhất trong thực tiễn ứng dụng GenAI là rủi ro về độ chính xác và tính tin cậy của thông tin. Các mô hình GenAI hoạt động dựa trên xác suất thống kê và khả năng suy diễn ngôn ngữ, do đó không được thiết kế để kiểm chứng tính đúng sai của nội dung đầu ra. Điều này dẫn đến hiện tượng mô hình tạo sinh thông tin sai lệch, thiếu căn cứ hoặc không phản ánh đúng dữ liệu thực tế. Trong lĩnh vực dịch vụ tài chính, nơi thông tin là nền tảng cho mọi quyết định, việc sử dụng đầu ra thiếu chính xác của GenAI có thể gây ra tổn thất tài chính, làm sai lệch chiến lược kinh doanh và ảnh hưởng nghiêm trọng đến uy tín tổ chức. Cụ thể vào tháng 1 năm 2024, một nhân viên tại một công ty có trụ sở tại Hồng Kông đã chuyển 25 triệu USD cho những kẻ lừa đảo sau khi được giám đốc tài chính chỉ đạo làm như vậy trong một cuộc gọi video có cả các đồng nghiệp khác tham

Hình 1: Lợi ích của AI đạt được ngày hôm nay so với hy vọng đạt được trong tương lai



Nguồn: Deloitte

Hình 2: GenAI gia tăng thiệt hại do gian lận trong những năm tới

Nguồn: Deloitte

gia. Những kẻ lừa đảo đã tạo ra một video giả mạo (deepfake) sao chép hình ảnh để lừa chuyển tiền. Trung tâm Dịch vụ Tài chính của Deloitte dự đoán rằng GenAI có thể khiến thiệt hại do gian lận đạt 40 tỷ USD tại Hoa Kỳ vào năm 2027, từ 12,3 tỷ USD vào năm 2023, với tốc độ tăng trưởng kép hàng năm là 32% như hình 2.

Bên cạnh rủi ro về độ tin cậy của thông tin, rủi ro liên quan đến bảo mật dữ liệu và quyền riêng tư của khách hàng cũng đang gia tăng. GenAI cần được huấn luyện và vận hành dựa trên khối lượng dữ liệu lớn, trong đó bao gồm nhiều dữ liệu nhạy cảm như thông tin cá nhân, lịch sử giao dịch, hồ sơ tài chính và dữ liệu định danh khách hàng. Báo cáo của Deloitte về ứng dụng GenAI trong lĩnh vực tài chính tại khu vực châu Á - Thái Bình Dương chỉ ra rằng nhiều tổ chức chưa xây dựng đầy đủ các quy trình kiểm soát dữ liệu đầu vào, dẫn đến nguy cơ rò rỉ thông tin hoặc sử dụng dữ liệu vượt quá mục đích ban đầu được khách hàng cho phép. Trong bối cảnh các quy định về bảo vệ dữ liệu cá nhân

ngày càng chặt chẽ, rủi ro này không chỉ gây tổn thất về mặt uy tín mà còn tiềm ẩn các chế tài pháp lý nghiêm khắc.

Rủi ro về thiên lệch dữ liệu và phân biệt đối xử cũng là một thực trạng đáng chú ý trong việc ứng dụng GenAI vào dịch vụ tài chính. Các mô hình tạo sinh được huấn luyện trên dữ liệu lịch sử, trong khi dữ liệu này có thể phản ánh các định kiến xã hội hoặc sai lệch mang tính hệ thống tồn tại từ trước. Deloitte cảnh báo rằng nếu không được giám sát và hiệu chỉnh phù hợp, GenAI có thể khuếch đại những thiên lệch này trong quá trình đánh giá khách hàng, đặc biệt trong các hoạt động như chấm điểm tín dụng hoặc xét duyệt vay vốn. Hệ quả là một số nhóm khách hàng có thể bị đánh giá bất lợi một cách không công bằng, làm gia tăng rủi ro đạo đức và gây ra xung đột với các nguyên tắc về công bằng và bao trùm trong dịch vụ tài chính.

Một rủi ro khác ngày càng được ghi nhận là thiếu minh bạch và khó giải trình trong các quyết

định do GenAI hỗ trợ. Phần lớn các mô hình GenAI hiện nay hoạt động như các “hộp đen”, khiến tổ chức tài chính gặp khó khăn trong việc giải thích vì sao mô hình đưa ra một kết quả hoặc khuyến nghị cụ thể. Deloitte nhận định rằng đây là thách thức lớn đối với các tổ chức chịu sự giám sát chặt chẽ của cơ quan quản lý, bởi yêu cầu giải trình và minh bạch là yếu tố cốt lõi trong hoạt động tài chính. Khi các quyết định quan trọng không thể được giải thích rõ ràng, niềm tin của khách hàng và nhà đầu tư có nguy cơ bị suy giảm.

Ngoài ra, rủi ro pháp lý và tuân thủ cũng đang trở nên rõ nét hơn khi hành lang pháp lý đối với GenAI vẫn trong quá trình hoàn thiện tại nhiều quốc gia. Deloitte chỉ ra rằng các tổ chức tài chính đang phải đối mặt với tình trạng bất định pháp lý, khi các quy định liên quan đến trách nhiệm pháp lý, sở hữu trí tuệ và sử dụng dữ liệu cho GenAI chưa được thống nhất. Việc triển khai GenAI trong bối cảnh này có thể khiến doanh nghiệp vô tình vi phạm các quy định hiện hành hoặc gặp khó khăn trong việc thích ứng với các yêu cầu pháp lý mới được ban hành trong tương lai.

Nhìn chung, thực trạng rủi ro từ ứng dụng GenAI trong dịch vụ tài chính cho thấy công nghệ này đang phát triển nhanh hơn khả năng quản trị và kiểm soát của nhiều tổ chức. Các rủi ro không chỉ xuất phát từ yếu tố kỹ thuật mà còn gắn chặt với dữ liệu, con người, quy trình và môi trường pháp lý. Việc nhận diện đầy đủ và đánh giá đúng mức các rủi ro này là cơ sở quan trọng để các tổ chức tài chính xây dựng chiến lược ứng dụng GenAI an toàn, hiệu quả và bền vững trong giai đoạn tới.

3. Giải pháp hạn chế rủi ro từ ứng dụng GenAI trong dịch vụ tài chính

Để giảm thiểu các rủi ro phát sinh từ việc ứng dụng GenAI trong dịch vụ tài chính, các tổ chức cần tiếp cận theo hướng quản trị tổng thể, kết hợp đồng bộ giữa công nghệ, con người và thể chế.

Trước hết, việc xây dựng khung quản trị trí tuệ nhân tạo là giải pháp nền tảng nhằm kiểm soát toàn bộ vòng đời của hệ thống GenAI, từ khâu lựa chọn mô hình, huấn luyện dữ liệu, triển khai ứng dụng đến giám sát và đánh giá kết quả. Khung quản trị này cần xác định rõ trách nhiệm của từng bộ phận liên quan, đặc biệt là mối liên kết giữa bộ phận công nghệ thông tin, quản lý rủi ro và pháp chế, nhằm đảm bảo các quyết định dựa trên GenAI có thể được giám sát và giải trình khi cần thiết.

Bên cạnh đó, tăng cường kiểm soát chất lượng dữ liệu và đầu ra của mô hình là giải pháp quan trọng để hạn chế rủi ro về độ tin cậy thông tin. Các tổ chức tài chính cần thiết lập cơ chế rà soát, xác thực và hiệu chỉnh kết quả do GenAI tạo ra trước khi đưa vào các quy trình nghiệp vụ trọng yếu. Việc áp dụng mô hình “con người trong vòng kiểm soát” đối với các quyết định có tác động lớn đến khách hàng và tài chính được xem là cần thiết nhằm giảm thiểu nguy cơ sai lệch và thiên lệch thuật toán.

Giải pháp tiếp theo là tăng cường bảo mật dữ liệu và bảo vệ quyền riêng tư trong quá trình triển khai GenAI. Các tổ chức tài chính cần xây dựng chính sách quản lý dữ liệu phù hợp với đặc thù của GenAI, bao gồm phân loại dữ liệu, kiểm soát quyền truy cập, mã hóa và giám sát việc sử dụng dữ liệu huấn luyện. Đồng thời, cần đảm bảo rằng việc sử dụng dữ liệu cá nhân cho GenAI tuân thủ đầy đủ các quy định pháp luật hiện hành về bảo vệ dữ liệu và quyền riêng tư của khách hàng.

Ngoài ra, nâng cao năng lực con người và văn hóa quản trị rủi ro cũng đóng vai trò then chốt. Nhân sự trong lĩnh vực tài chính cần được trang bị kiến thức cơ bản về GenAI, bao gồm cả lợi ích và rủi ro, để có thể nhận diện sớm các vấn đề phát sinh trong quá trình sử dụng công nghệ. Việc đào tạo liên ngành giữa công nghệ, tài chính và pháp lý sẽ góp phần nâng cao khả năng phối hợp và quản lý rủi ro một cách hiệu quả hơn.

TÀI LIỆU THAM KHẢO:

Deloitte (2022). Generative AI: Application and Regulation in Asia Pacific. Truy cập từ: <https://www.deloitte.com/content/dam/assets-zone1/au/en/docs/industries/financial-services/2023/deloitte-au-fs-acrs-publication-on-ai-regulation-in-apac-141223.pdf>

Deloitte (2025). Deloitte's State of Generative AI in the Enterprise Quarter four report . Truy cập từ: <https://www.deloitte.com/content/dam/assets-shared/docs/about/2025/quarter-4.pdf>

Deloitte (2024). New Deloitte survey finds expectations for Gen AI remain high, but many are feeling pressure to quickly realize value while managing risks. Truy cập từ:

<https://www.deloitte.com/global/en/about/press-room/gen-ai-survey.html>.

Deloitte (2026). State of AI in the Enterprise The untapped edge. Truy cập từ: <https://www.deloitte.com/content/dam/assets-zone3/us/en/docs/services/consulting/2026/state-of-ai-2026.pdf>

Deloitte (2024). Generative AI is expected to magnify the risk of deepfakes and other fraud in banking. Truy cập từ: <https://www.deloitte.com/us/en/insights/industry/financial-services/deepfake-banking-fraud-risk-on-the-rise.html>

Ngày nhận bài: 9/1/2026

Ngày phản biện đánh giá và sửa chữa: 21/1/2026

Ngày chấp nhận đăng bài: 13/2/2026

RISKS ASSOCIATED WITH THE APPLICATION OF GENERATIVE AI IN FINANCIAL SERVICES

● NGUYEN THI KIM THOA

Faculty of Finance, Banking and Insurance
University of Economics - Technology for Industries

ABSTRACT:

The application of generative artificial intelligence (GenAI) in financial services is creating significant opportunities to enhance operational efficiency, personalize customer services, and support data-driven decision-making. However, these benefits are accompanied by increasingly prominent risks related to information reliability, data security, legal liability, and professional ethics. This study examines the potential risks associated with the practical implementation of GenAI in financial institutions and analyzes their implications for business operations and customer trust. The study highlights the growing importance of establishing appropriate governance frameworks to ensure that the adoption of emerging technologies is effectively managed, responsibly implemented, and aligned with the goal of sustainable development in the financial sector.

Keywords: Generative AI (GenAI), financial services risks, data security, transparency and accountability, technology governance.