

PHÂN TÍCH CÁC BÁO CÁO ESG CỦA CITIBANK THÔNG QUA MÔ HÌNH LATENT DIRICHLET ALLOCATION (LDA)

ANALYZING CITIBANK'S ESG REPORTS USING THE LATENT DIRICHLET ALLOCATION (LDA) MODEL

Trương Công Đoàn^{1,*}, Phạm Thị Phương Thảo¹,
Hà Thị Hồng Ánh¹, Đỗ Thanh Vân¹, Nguyễn Bích Diệp¹

DOI: <https://doi.org/10.57001/huih5804.2025.103>

TÓM TẮT

Bài báo này trình bày việc áp dụng phương pháp Latent Dirichlet Allocation (LDA), một kỹ thuật học máy không giám sát, để phân tích các báo cáo về Môi trường, Xã hội và Quản trị (ESG) của Citibank trong giai đoạn từ năm 2019 đến 2023. Mục tiêu là khám phá các cấu trúc chủ đề cơ bản và hiểu cách truyền thông về tính bền vững đã thay đổi theo thời gian. Sáu câu hỏi nghiên cứu được phát triển để trích xuất và so sánh các chủ đề chính trong mỗi báo cáo hàng năm, sau đó là phân tích theo năm. Các phát hiện cho thấy sự thay đổi dần dần trong trọng tâm chủ đề, từ các vấn đề rộng như quản trị doanh nghiệp, sự tham gia của nhân viên và hỗ trợ cộng đồng trong những năm trước sang các mối quan tâm cụ thể hơn như rủi ro tài chính liên quan đến khí hậu, tài chính bền vững và xác minh phát thải trong những năm gần đây. Những phát hiện này chứng minh sự nhấn mạnh ngày càng tăng của Citibank vào hiệu suất phát triển bền vững có thể đo lường được và tuân thủ quy định. Nghiên cứu chứng minh hiệu quả của mô hình chủ đề trong các đánh giá theo chiều dọc về các tiết lộ ESG của doanh nghiệp, cung cấp thông tin chi tiết về xu hướng truyền thông chiến lược và thay đổi các ưu tiên của tổ chức.

Từ khóa: Latent Dirichlet Allocation (LDA), mô hình chủ đề, ESG.

ABSTRACT

This article applies Latent Dirichlet Allocation (LDA), an unsupervised machine learning technique, to analyze Citibank's Environmental, Social, and Governance (ESG) reports from 2019 to 2023. The goal is to discover underlying topic structures and understand how sustainability communication has changed over time. Six research questions are developed to extract and compare the major themes in each yearly report, followed by a cross-year analysis. The findings show a progressive shift in theme emphasis, from broad issues like corporate governance, employee involvement, and community support in previous years to more particular concerns like climate-related financial risk, sustainable financing, and emission verification in recent years. These findings demonstrate Citibank's growing emphasis on measurable sustainability performance and regulatory compliance. The study demonstrates the efficacy of topic modeling in longitudinal assessments of corporate ESG disclosures, providing insights into strategic communication trends and altering organizational priorities.

Keywords: Latent Dirichlet Allocation (LDA), topic modeling, ESG.

¹Trường Quốc tế, Đại học Quốc gia Hà Nội

*Email: tcdoan@vnu.edu.vn

Ngày nhận bài: 02/3/2025

Ngày nhận bài sửa sau phản biện: 12/4/2025

Ngày chấp nhận đăng: 25/4/2025

1. GIỚI THIỆU

Trong bối cảnh kinh doanh hiện đại, nơi sự thay đổi diễn ra với tốc độ chóng mặt, tính bền vững và quản trị đạo đức không còn là những yếu tố thứ yếu mà đã trở thành những kỳ vọng cốt lõi từ các bên liên quan, giới đầu tư và các cơ quan quản lý. Báo cáo Môi trường, Xã hội và

Quản trị (ESG) đã nổi lên như một kênh truyền thông quan trọng, giúp các doanh nghiệp thể hiện một cách hiệu quả hiệu suất hoạt động phi tài chính và định hướng kiến tạo giá trị lâu dài. Những tài liệu này mang đến một cái nhìn sâu sắc về cách các công ty tiếp cận các vấn đề toàn cầu cấp bách như biến đổi khí hậu, công bằng xã hội, sự đa dạng và trách nhiệm doanh nghiệp [1].

Tuy nhiên, dù mang ý nghĩa to lớn, các báo cáo ESG thường có xu hướng dài dòng, nặng về mô tả và thiếu sự nhất quán trong cấu trúc, gây khó khăn cho việc phân tích một cách hệ thống, đặc biệt là khi theo dõi sự thay đổi qua nhiều năm. Điều này mở ra một tiềm năng lớn cho việc ứng dụng các kỹ thuật Xử lý Ngôn ngữ Tự nhiên (NLP) để khám phá các mô hình ẩn và trích xuất thông tin có cấu trúc từ các báo cáo dưới dạng văn bản tự do.

Dự án nghiên cứu này tập trung vào Citibank, một tổ chức tài chính toàn cầu với lịch sử báo cáo ESG lâu dài và minh bạch. Nghiên cứu tiến hành phân tích năm báo cáo ESG liên tiếp từ năm 2019 đến năm 2023 nhằm làm sáng tỏ sự thay đổi trong các ưu tiên, ngôn ngữ sử dụng và trọng tâm chủ đề của ngân hàng qua từng năm. Thông qua việc ứng dụng kỹ thuật NLP, đặc biệt là phương pháp phân bổ tiềm ẩn Dirichlet (Latent Dirichlet Allocation - LDA), dự án này trích xuất các chủ đề chính trong mỗi báo cáo. Từ đó, nghiên cứu đóng góp một phương pháp phân tích có khả năng mở rộng và tái sử dụng cho các tài liệu ESG. Cách tiếp cận này không chỉ cung cấp cái nhìn tổng quan về nội dung từng năm mà còn mang lại những so sánh giá trị về sự phát triển và điều chỉnh trong chiến lược phát triển bền vững của doanh nghiệp theo thời gian [2].

Dựa trên các mục tiêu đó, nghiên cứu đặt ra các câu hỏi nghiên cứu sau:

Câu hỏi 1: Những chủ đề chính trong Báo cáo ESG năm 2019 của Citibank là gì?

Câu hỏi 2: Những chủ đề chính trong Báo cáo ESG năm 2020 của Citibank là gì?

Câu hỏi 3: Những chủ đề chính trong Báo cáo ESG năm 2021 của Citibank là gì?

Câu hỏi 4: Những chủ đề chính trong Báo cáo ESG năm 2022 của Citibank là gì?

Câu hỏi 5: Những chủ đề chính trong Báo cáo ESG năm 2023 của Citibank là gì?

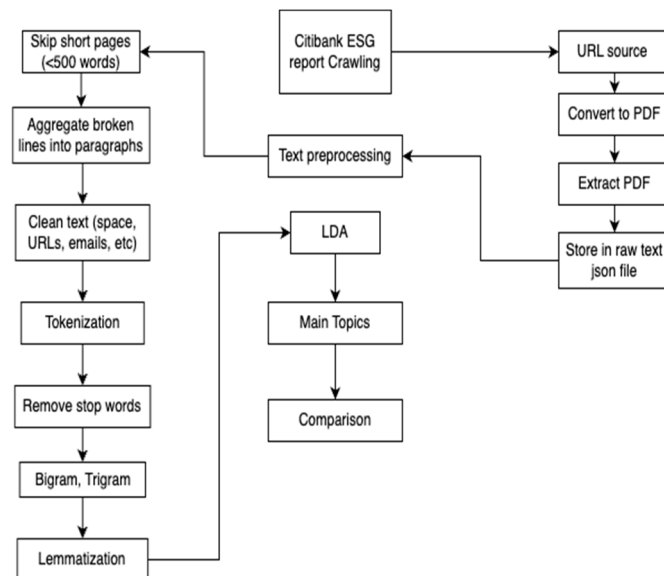
Câu hỏi 6: Những điểm khác biệt nổi bật giữa các báo cáo ESG của Citibank qua các năm là gì?

2. PHƯƠNG PHÁP NGHIÊN CỨU

2.1. Thiết kế nghiên cứu

Trọng tâm của thiết kế nghiên cứu này xoay quanh Latent Dirichlet Allocation (LDA), một mô hình xác suất sinh nổi tiếng được sử dụng để khai thác các chủ đề tiềm ẩn từ văn bản phi cấu trúc. Trong nghiên cứu này, LDA được áp dụng nhằm xác định và so sánh nội dung chủ đề trong các báo cáo ESG của Citibank qua nhiều năm. Việc triển khai ban đầu dựa trên một kho mã nguồn mở trên GitHub [2], vốn được áp dụng cho báo cáo ESG năm 2019

của Citibank. Kho mã này cung cấp một quy trình nền tảng cho việc tiền xử lý dữ liệu và xây dựng mô hình chủ đề, và đã được rà soát, điều chỉnh cẩn thận để phù hợp với cấu trúc và ngôn ngữ đặc thù của các báo cáo ESG. Dựa trên khuôn khổ đó, nghiên cứu đã mở rộng phương pháp và ứng dụng cho các báo cáo ESG của những năm tiếp theo.



Hình 1. Khung nghiên cứu

2.2. Mô hình chủ đề và Latent Dirichlet Allocation

Mô hình chủ đề là một phương pháp hiệu quả trong việc phân tích tập hợp lớn các tài liệu bằng cách khám phá cấu trúc chủ đề tiềm ẩn trong văn bản. Phương pháp này nhận diện các chủ đề dưới dạng những nhóm từ ngữ thường xuyên xuất hiện cùng nhau, phản ánh các ngữ cảnh chung [3].

Latent Dirichlet Allocation (LDA) là kỹ thuật chính được sử dụng trong nghiên cứu này để thực hiện topic modeling trên dữ liệu văn bản phi cấu trúc. LDA là một mô hình xác suất sinh (generative probabilistic model), được thiết kế nhằm phát hiện cấu trúc chủ đề tiềm ẩn trong một tập hợp tài liệu lớn [3]. Ban đầu được phát triển cho các tác vụ xử lý ngôn ngữ tự nhiên, LDA đã được ứng dụng rộng rãi trong nhiều lĩnh vực như phân tích cảm xúc, khai phá tri thức, trích xuất thông tin học thuật và gần đây là phân tích báo cáo phát triển bền vững.

LDA đặc biệt phù hợp với việc phân tích các báo cáo ESG vốn thường dài, phi cấu trúc và sử dụng ngôn ngữ chuyên môn. Việc áp dụng LDA giúp nghiên cứu có thể hệ thống hóa và tự động hóa việc xác định các chủ đề lặp lại qua nhiều năm, từ đó rút ra được các thông tin có giá trị mà phương pháp đọc thủ công khó có thể đạt được. Đồng thời, mô hình còn giúp phát hiện các thay đổi tinh tế trong cách sử dụng từ ngữ, phản ánh rõ nét cách

Citibank truyền đạt các ưu tiên về môi trường, xã hội và quản trị theo thời gian.

2.3. Quy trình xử lý dữ liệu

Trước khi tiến hành mô hình hóa chủ đề, một quy trình tiền xử lý dữ liệu toàn diện đã được áp dụng để chuẩn bị các báo cáo ESG cho quá trình phân tích. Các bước này đóng vai trò thiết yếu trong việc giảm nhiễu, chuẩn hóa dữ liệu văn bản, và nâng cao chất lượng đầu vào cho mô hình Latent Dirichlet Allocation (LDA).

Quá trình tiền xử lý bắt đầu bằng việc chuẩn hóa văn bản, bao gồm chuyển tất cả ký tự thành chữ thường, loại bỏ dấu câu, chữ số và các ký tự không thuộc bảng chữ cái. Việc chuẩn hóa này đảm bảo tính nhất quán trên toàn bộ tập dữ liệu và giảm thiểu ảnh hưởng từ sự đa dạng về mặt cú pháp [4].

Tiếp theo, các từ dừng, là những từ phổ biến nhưng không mang nhiều giá trị thông tin, được loại bỏ bằng cách sử dụng danh sách từ dừng tiêu chuẩn từ thư viện Natural Language Toolkit (NLTK) trong Python. Bước này đóng vai trò quan trọng trong việc loại bỏ các từ có tần suất cao nhưng không đóng góp đáng kể cho ý nghĩa ngữ cảnh, từ đó cải thiện độ rõ ràng của chủ đề [5].

quá trình phân tích. Ví dụ, từ "reporting" sẽ được chuyển về "report" [6].

Cuối cùng, một ma trận tài liệu và từ vựng được xây dựng, biểu diễn định lượng tần suất xuất hiện của các từ trong từng tài liệu. Ma trận này đóng vai trò là đầu vào cơ bản cho mô hình LDA, cho phép phát hiện các chủ đề tiềm ẩn thông qua việc xác định các mẫu đồng xuất hiện của từ ngữ trong toàn bộ tập dữ liệu [7].

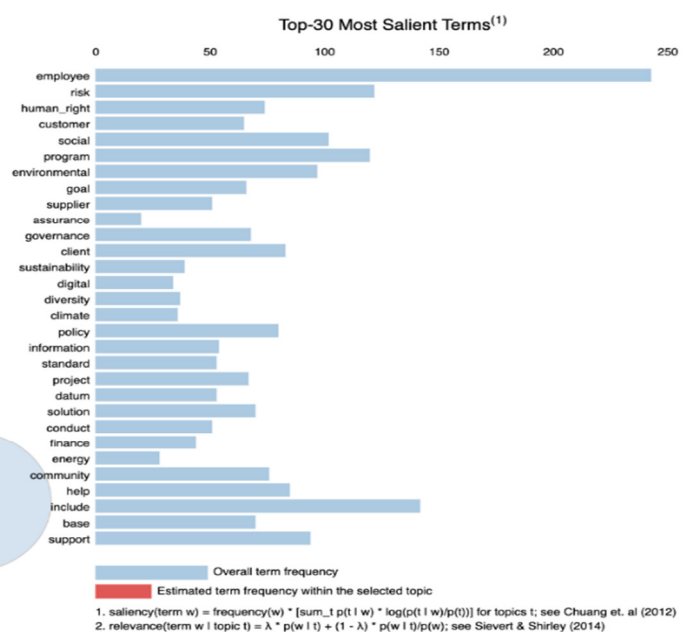
3. KẾT QUẢ

3.1. Kết quả cho câu hỏi nghiên cứu 1

Trong hình 2 thì bản đồ khoảng cách liên chủ đề từ Báo cáo ESG toàn cầu Citibank 2019 trực quan hóa quan hệ giữa các chủ đề ESG. Vòng tròn biểu thị từng chủ đề, kích thước tương ứng độ quan trọng. Danh sách 30 thuật ngữ nổi bật chỉ ra vấn đề cốt lõi, liên kết nhiều từ khóa, giúp người dùng hiểu nội dung chính. Ví dụ, Chủ đề 1 tập trung vào sự gắn bó của nhân viên và hợp tác cộng đồng, quyết định hiệu quả ESG. Bản đồ còn hiển thị tỷ lệ và làm nổi bật chủ đề chính, làm rõ các tương tác giữa các chủ đề và nhấn mạnh phương pháp tiếp cận tích hợp cho tính bền vững.



Những phần nội dung bị lặp lại và các đoạn văn bản dư thừa cũng được loại bỏ để tránh làm lệch phân bố chủ đề do sự lặp lại không cần thiết giữa các tài liệu. Sau bước này, văn bản được phân tách thành các đơn vị từ riêng lẻ thông qua quá trình tách từ. Các từ sau đó được đưa về dạng gốc thông qua kỹ thuật lemmatization, nhằm đảm bảo các biến thể của một từ được xử lý thống nhất trong



Hình 2. Biểu đồ khoảng cách giữa các chủ đề của Báo cáo ESG Citibank 2019

Xác minh và Chứng nhận trong Quản lý nhà kính (23,5%): Với tỷ lệ cao nhất là 23,5%, chủ đề này cho thấy sự nhấn mạnh đáng kể vào việc đảm bảo về các chứng chỉ và tiêu chuẩn về quản lý khí nhà kính. Chỉ ra rằng công ty cam kết mức độ chính xác về dữ liệu nhà kính đã công khai, đảm bảo hoạt động tuân thủ các nguyên về môi trường bền vững.

Tính bền vững & Quản trị (16,5%): Chủ đề này sử dụng 16,5%, làm nổi bật nhu cầu ngày càng tăng về quản lý rủi ro và các hoạt động bền vững. Tổ chức được coi là tập trung xây dựng các chính sách giảm thiểu tác động đến môi trường và tăng cường trách nhiệm xã hội trong toàn bộ hoạt động của mình.

Quyền con người & Sự đa dạng (13,7%): Với 13,7%, chủ đề này nhấn mạnh tầm quan trọng của việc bảo vệ quyền con người và cung cấp sự đa dạng trong tổ chức. Điều này cho thấy cam kết xây dựng một văn hóa tôn trọng và coi trọng sự bình đẳng đẳng cấp cho tất cả các bên liên quan.

Trải nghiệm khách hàng (13,7%): Tương tự, với 13,7%, việc cải tiến trải nghiệm khách hàng được xác định là yếu tố then chốt để duy trì lòng trung thành của khách hàng và xây dựng uy tín thương hiệu. Hãy tìm sự đầu tư vào công nghệ và đào tạo để nâng cao chất lượng dịch vụ khách hàng.

Mục tiêu Tài chính & Môi trường (12,8%): Việc thiết lập các mục tiêu tài chính và môi trường với tỷ lệ 12,8% có thể thực hiện cam kết của tổ chức để phát triển bền vững. Điều này không chỉ giúp cải thiện hình ảnh thương hiệu mà còn thu hút các nhà tư vấn có trách nhiệm.

Hỗ trợ Nhân viên và Phát triển cộng đồng (12,34%): Mặc dù có tỷ lệ thấp nhất (12,34%), việc tăng cường gắn kết của nhân viên vẫn được coi là rất quan trọng để cải thiện hiệu suất và sự hài lòng trong công việc và tạo ra phát triển bền vững thông qua các chương trình về cộng đồng.

Bảng 1. Phân loại chủ đề và từ khóa trích xuất từ Báo cáo ESG năm 2019 của Citibank

Chủ đề	Từ khóa	Tên chủ đề đề xuất	Tỷ lệ
0	risk, social, environmental, include, governance, client, project, policy, information, base	Tính bền vững và Quản trị	16,5%
1	employee, provide, support, work, program, business, help, community, financial, new	Hỗ trợ nhân viên và Phát triển cộng đồng	12,34%
2	human_right, employee, supplier, diversity, pay, compliance, policy, code, culture, director	Quyền con người và sự đa dạng	15,8%
3	customer, digital, program, skill, solution, customer, workforce, technology, learn, future	Trải nghiệm khách hàng	13,7%

4	assurance, verification, greenhouse, register, opinion, auditor, certificate, depend, text, assure	Xác minh và Chứng nhận trong Quản lý nhà kính	23,5%
5	goal, climate, energy, firm, finance, banking, maintain, activity, global, client	Mục tiêu tài chính và môi trường	12,8%

Sự phân bố tỷ lệ giữa các chủ đề ESG cho thấy một sự tập trung nhất quán vào các trụ cột cốt lõi. Sự nổi bật của "Xác minh và Chứng nhận trong Quản lý nhà kính" cho thấy tổ chức coi trọng việc hoàn thành sứ mệnh vì môi trường và đảm bảo tuân thủ các quy tắc như nền tảng cho hoạt động lâu dài. Sự tương đồng về mức độ chú trọng giữa "Tính bền vững & Quản trị", "Quyền con người & Đa dạng", "Trải nghiệm Khách hàng" và "Mục tiêu Tài chính & Môi trường" cho thấy một sự cân bằng trong việc theo đuổi trách nhiệm xã hội, thực hành hướng đến khách hàng và kế hoạch tài chính bền vững. Mặc dù, "Hỗ trợ nhân viên và Phát triển cộng đồng" có tỷ lệ thấp hơn, nó vẫn là một ưu tiên nội bộ quan trọng. Nhìn chung, báo cáo này thể hiện một chiến lược ESG toàn diện, trong đó các yếu tố khác nhau được tích hợp để đạt được các mục tiêu bền vững rộng lớn hơn.

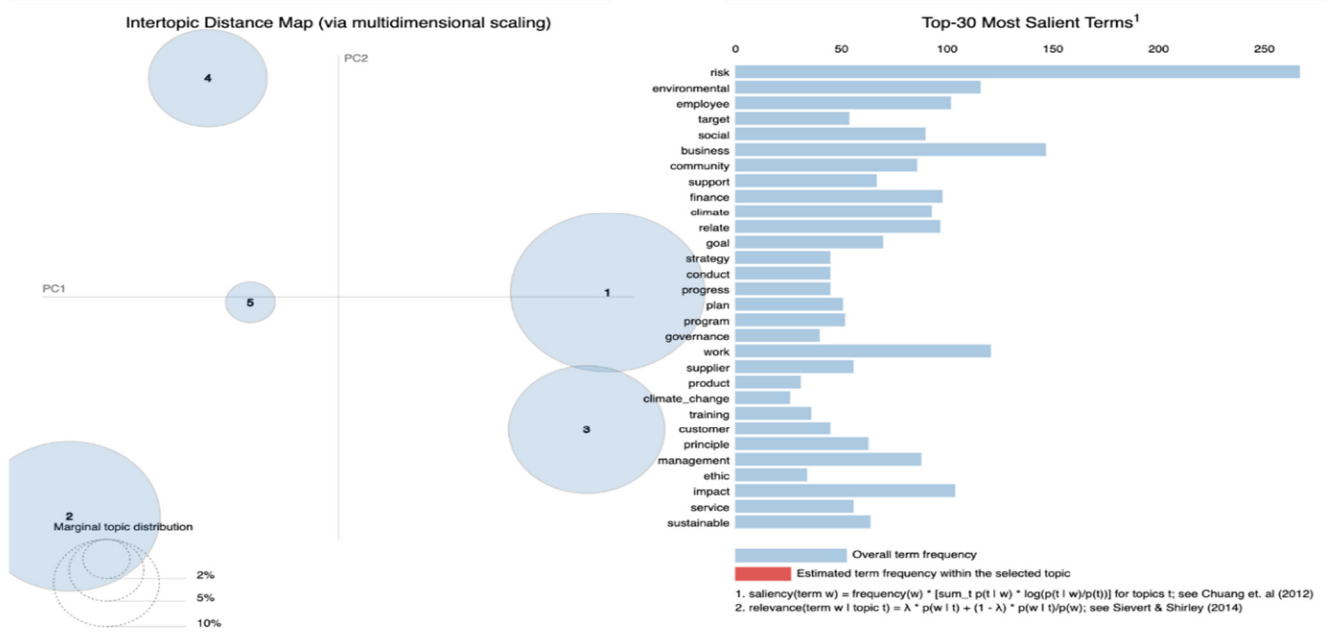
3.2. Kết quả cho câu hỏi nghiên cứu 2

Trong hình 3, bản đồ khoảng cách liên chủ đề từ Báo cáo ESG toàn cầu Citi 2020 và Danh sách 30 thuật ngữ nổi bật, trực quan hóa quan hệ giữa các chủ đề ESG.

Năm chủ đề được phân tích cho thấy một chiến lược ESG cân bằng, với sự nhấn mạnh đặc biệt vào tính bền vững và việc giám sát rủi ro.

Tính bền vững và hành động khí hậu (15%): Chủ đề này phản ánh các nỗ lực chiến lược của tổ chức trong việc giải quyết biến đổi khí hậu và thúc đẩy tính bền vững. Cho thấy cam kết giảm thiểu tác động đến môi trường thông qua kế hoạch hành động cụ thể, đồng thời khẳng định đây là một lĩnh vực trọng tâm trong chương trình nghị sự ESG của tổ chức.

Quản lý rủi ro môi trường và xã hội (17,6%): Với tỷ lệ cao nhất là 17,6%, chủ đề này làm nổi bật cách tiếp cận của tổ chức trong việc xác định và giảm thiểu các rủi ro liên quan đến ESG. Nó nhấn mạnh sự tích hợp của yếu tố bền vững vào các quyết định tài chính và hoạt động, thể hiện một cam kết mạnh mẽ đối với việc giám sát rủi ro một cách có trách nhiệm.



Hình 3. Biểu đồ khoảng cách giữa các chủ đề của Báo cáo ESG Citibank 2020

Quyền riêng tư và bảo mật dữ liệu (7,9%): Chủ đề này tập trung vào việc bảo vệ dữ liệu của các bên liên quan và đảm bảo an ninh hệ thống. Tỷ lệ 7,9% phản ánh nỗ lực của tổ chức trong việc xây dựng lòng tin và ngăn chặn các hành vi xâm phạm dữ liệu, đồng thời tuân thủ các quy định và nhấn mạnh tầm quan trọng của việc bảo vệ quyền riêng tư.

Bảng 2. Phân loại chủ đề và từ khóa trích xuất từ Báo cáo ESG năm 2020 của Citibank

Chủ đề	Từ khóa	Tên chủ đề đề xuất	Tỷ lệ
0	target, goal, plan, strategy, progress, product, climate_change, commitment, action, service	Tính bền vững và hành động khí hậu	15%
1	risk, environmental, relate, social, climate, impact, finance, client, sustainable, supplier	Quản lý rủi ro môi trường và xã hội	17,6%
2	privacy, breach, exercise, identity, incident, benefit, president, qualified, center, card	Quyền riêng tư và bảo mật dữ liệu	7,9%
3	include, conduct, policy, risk, principle, governance, training, client, ethic, management	Quản trị và tiêu chuẩn đạo đức	11,4%

4	business, employee, work, community, support, program, customer, provide, access, company	Gắn kết nhân viên và cộng đồng	1,4%
---	---	--------------------------------	------

Quản trị và tiêu chuẩn đạo đức (11,4%): Chiếm tỷ lệ 11,4%, chủ đề này nhấn mạnh các thông lệ quản trị vững chắc và đạo đức kinh doanh trên toàn tổ chức. Thể hiện các nỗ lực nhằm đảm bảo tính minh bạch, trách nhiệm giải trình và sự phù hợp của chính sách, khẳng định quản trị là một trụ cột chính trong khuôn khổ ESG.

Gắn kết nhân viên và cộng đồng (1,4%): Mặc dù có tỷ lệ thấp nhất là 1,4%, chủ đề này thể hiện các sáng kiến của tổ chức trong việc hỗ trợ nhân viên và cộng đồng. Nó phản ánh các giá trị về tính hòa nhập, trách nhiệm xã hội và sự tương tác tích cực. Dù tỷ lệ nhỏ, nó vẫn làm nổi bật những nỗ lực có ý nghĩa hướng đến con người.

Năm chủ đề phản ánh một chiến lược ESG cân bằng, nhấn mạnh tính bền vững và quản lý rủi ro. Hai lĩnh vực trọng tâm, “Quản lý rủi ro môi trường và xã hội” cùng “Tính bền vững và hành động khí hậu”, thể hiện cam kết với trách nhiệm môi trường. “Quản trị và tiêu chuẩn đạo đức” hỗ trợ tính trách nhiệm nội bộ, trong khi “Quyền riêng tư và bảo mật dữ liệu” thể hiện sự chú trọng đến niềm tin số. Mặc dù “Gắn kết của nhân viên và cộng đồng” ít nổi bật hơn, nhưng vẫn đại diện cho những giá trị xã hội thiết yếu. Báo cáo thể hiện cam kết dựa trên dữ liệu đối với tính bền vững, khả năng phục hồi và đạo đức.

3.3. Kết quả cho câu hỏi nghiên cứu 3

Kết quả mô hình hóa chủ đề năm 2021 cho thấy, sự tập trung mạnh mẽ vào một chủ đề thống lĩnh duy nhất (Chủ đề 1). Những kết quả này chỉ ra rằng vào năm 2021, Citibank chủ yếu tập trung vào việc đặt nền tảng chiến lược cho các sáng kiến ESG, đặt ra các mục tiêu, chính sách và khuôn khổ hành động nhằm thúc đẩy tính bền vững. Hơn nữa, các chủ đề còn lại tương đối nhỏ và hơi chồng chéo, cho thấy việc tích hợp ESG trên nhiều hoạt động kinh doanh khác nhau vẫn còn hạn chế, chủ yếu ở giai đoạn đầu. Điều này phản ánh giai đoạn đầu của quá trình chuyển đổi ESG, khi công ty vẫn đang định hình định hướng nội bộ và các chiến lược ưu tiên của mình.



Hình 4. Biểu đồ khoảng cách giữa các chủ đề của Báo cáo ESG Citibank 2021

Đạo đức và Chống tham nhũng (10,7%): Citi khẳng định các nỗ lực chống hối lộ, chống rửa tiền đồng thời tập trung vào quản trị điều hành có đạo đức, việc làm toàn diện và thù lao cho nhà điều hành.

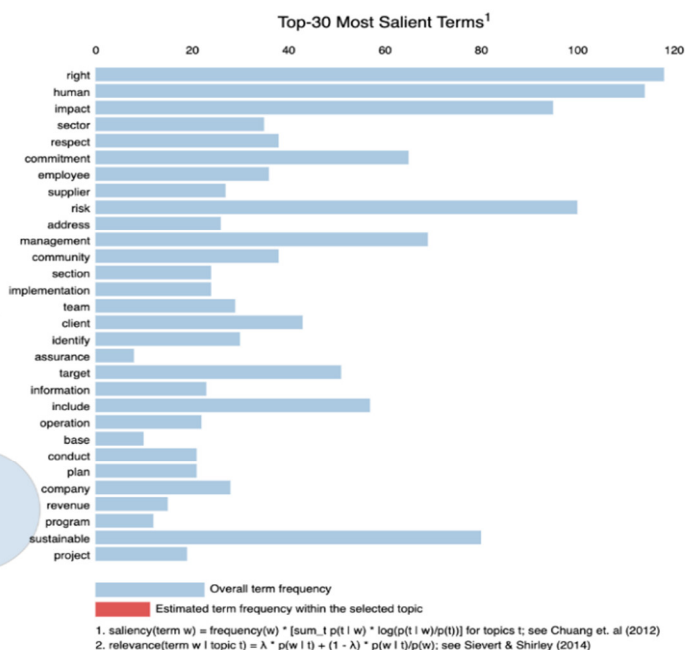
Nhân quyền và Giám sát chuỗi cung ứng (39,9%): Chủ đề thống lĩnh này vào năm 2021 phản ánh sự nhấn mạnh của Citi vào việc tôn trọng nhân quyền và đảm bảo quản lý nhà cung cấp có trách nhiệm, phù hợp với Tuyên bố nhân quyền và đánh giá rủi ro ESG của Citibank.

Đảm bảo độc lập và Nguyên tắc Equator (9,8%): Citi nhấn mạnh tính minh bạch thông qua đảm bảo của bên thứ ba, phạm vi kiểm toán và cam kết tuân thủ Nguyên tắc Xích đạo trong việc tài trợ cho các dự án nhạy cảm với môi trường và xã hội.

Tác động ngành và Rủi ro tài chính (20,9%): Chủ đề này đề cập đến cách Citi xem xét các rủi ro ESG theo ngành và các tác động tài chính, chẳng hạn như tổn thất dự kiến, tác động đến tài sản và rủi ro hoạt động.

Cam kết và quản trị rủi ro (21,3%): Citi nhấn mạnh cách tiếp cận rủi ro của mình thông qua các tương tác với khách hàng, cơ chế quản trị và tích hợp ESG trong các nhóm. Các nỗ lực triển khai và công bố thông tin được nhấn mạnh.

Cam kết cộng đồng và tiếp cận khách hàng (13,9%): Bao gồm các nỗ lực cải thiện khả năng tiếp cận tài chính, thu hút các cộng đồng chưa được phục vụ đầy đủ và tài trợ cho các sáng kiến hỗ trợ hòa nhập kinh tế.



Tài chính bền vững (15,7%): Citi nhấn mạnh các hành động được thực hiện hướng tới các mục tiêu tài chính bền vững của mình, được hướng dẫn bởi các chính sách nội bộ và theo dõi tiến độ hướng tới các mục tiêu về khí hậu và tác động dài hạn.

Báo cáo ESG năm 2021 của Citi cho thấy cam kết mạnh mẽ đối với quản trị đạo đức, nhân quyền và các hoạt động bền vững. Sự tập trung đáng kể vào quyền con người và giám sát chuỗi cung ứng làm nổi bật sự tận tâm của tổ chức đối với việc quản lý nhà cung cấp có trách nhiệm. Tính minh bạch được ưu tiên thông qua đảm bảo độc lập và Nguyên tắc Equator, trong khi việc xem xét các rủi ro ESG theo ngành cho thấy sự hiểu biết chiến lược về các tác động tài chính. Citi cũng nhấn mạnh vào sự tham gia và quản trị của khách hàng, củng cố phương pháp quản

lý rủi ro của mình. Hơn nữa, các sáng kiến hướng đến sự tham gia của cộng đồng và tài chính bền vững cho thấy một chiến lược toàn diện hỗ trợ cho sự hòa nhập kinh tế và các mục tiêu khí hậu dài hạn.

Bảng 3. Phân loại chủ đề và từ khóa trích xuất từ Báo cáo ESG năm 2021 của Citibank

Chủ đề	Từ khóa	Tên chủ đề đề xuất	Tỷ lệ
0	employee, program, remuneration, bribery, anti, executive, disability, director, responsibly, launder.	Đạo đức và Chống tham nhũng	10,7%
1	right, human, respect, management, supplier, employee, risk, statement, evaluation, operation.	Nhân quyền và Giám sát chuỗi cung ứng	39,9%
2	assurance, equator, verification, base, opinion, scope, auditor, comprise, contain, register.	Đảm bảo độc lập và Nguyên tắc Equator	9,8%
3	sector, impact, expect, asset, loss, revenue, minor, operation, company, increase	Tác động ngành và Rủi ro tài chính	20,9%
4	risk, commitment, client, address, team, section, implementation, information, include, conduct.	Cam kết và quản trị rủi ro	21,3%
5	community, identify, customer, impact, business, fund, work, number, access, network.	Cam kết cộng đồng và tiếp cận khách hàng	13,9%
6	sustainable, finance, impact, target, action, principle, progress, activity, include, policy.	Tài chính bền vững	15,7%

3.4. Kết quả cho câu hỏi nghiên cứu 4

Kết quả năm 2022 cho thấy, một tập hợp các chủ đề đa dạng hơn và phân bổ đều hơn. Những phát hiện này chỉ ra rằng Citibank đã chuyển từ giai đoạn lập kế hoạch chiến lược ESG sang triển khai thực tế và tích hợp kinh doanh. Ngân hàng không chỉ theo đuổi các mục tiêu bền vững cho chính mình mà còn tích cực hỗ trợ khách hàng trong quá trình chuyển đổi ESG của họ. Biểu đồ khoảng cách liên chủ đề năm 2022 cũng cho thấy sự tách biệt lớn hơn giữa các chủ đề, phản ánh phạm vi mở rộng của các

hoạt động liên quan đến ESG và kết nối chặt chẽ hơn với các dịch vụ kinh doanh, sản phẩm tài chính và các nỗ lực hướng đến thị trường.

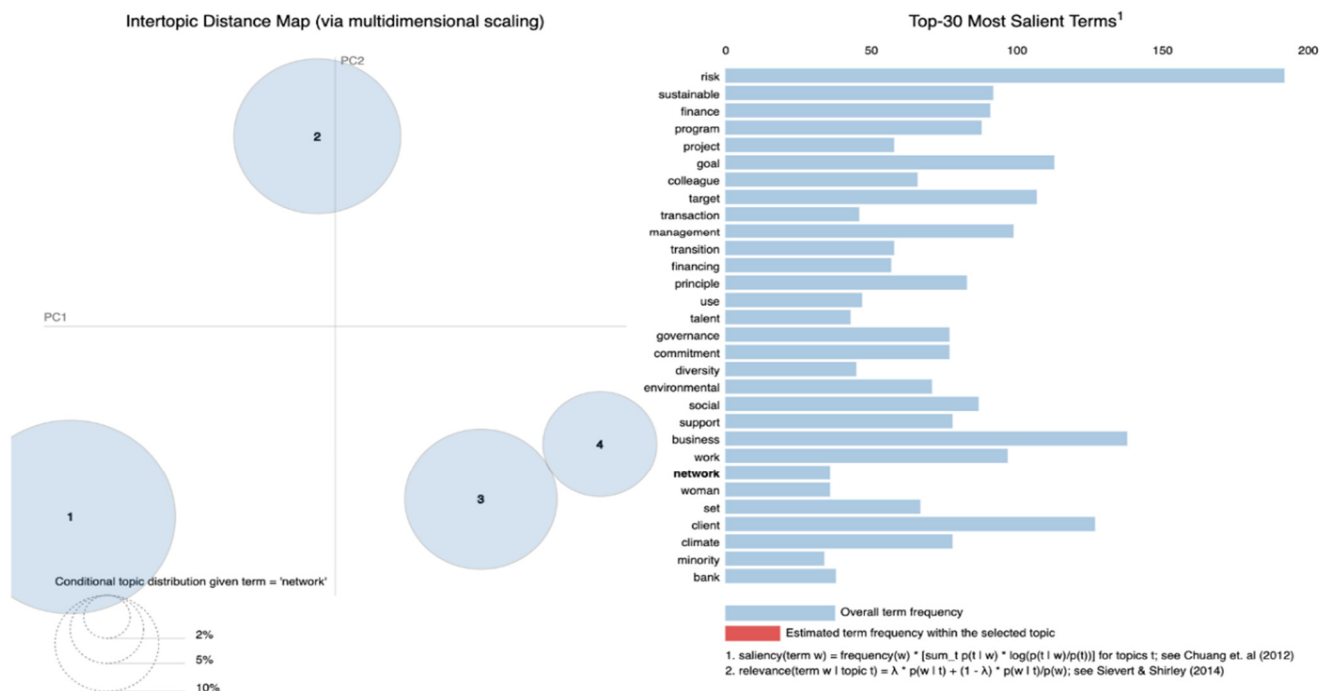
Bảng 4. Phân loại chủ đề và từ khóa trích xuất từ Báo cáo ESG năm 2022 của Citibank

Chủ đề	Từ khóa	Tên chủ đề đề xuất	Tỷ lệ
0	risk, target, management, include, principle, governance, commitment, client, climate, sector	Quản lý rủi ro và quản trị	18,5%
1	project, transaction, use, financing, environmental, social, diligence, mining, labor, facility	Tài chính bền vững và trách nhiệm xã hội	14,2%
2	goal, sustainable, finance, business, transition, impact, client, service, climate, development	Phát triển kinh tế và phát triển bền vững	12,7%
3	program, colleague, work, diversity, talent, support, network, woman, minority, diverse	Đa dạng, công bằng và hòa nhập	10,6%

Quản lý rủi ro và quản trị (18,5%): Chủ đề này phản ánh tầm quan trọng của quản trị doanh nghiệp và quản lý rủi ro trong chiến lược ESG của Citibank. Các yếu tố như cam kết, nguyên tắc và tác động của biến đổi khí hậu đóng vai trò quan trọng.

Tài chính bền vững và trách nhiệm xã hội (14,2%): Citibank tập trung mạnh vào tài trợ tài chính bền vững, trách nhiệm xã hội và quản lý lao động trong chuỗi cung ứng. Điều này cho thấy vai trò của ngân hàng trong việc hỗ trợ các dự án có tác động tích cực đến môi trường và xã hội.

Phát triển kinh tế và phát triển bền vững (12,7%): Các mục tiêu tài chính bền vững và hỗ trợ kinh doanh là trọng tâm, nhấn mạnh tầm quan trọng của quá trình chuyển đổi kinh tế xanh, tác động của khí hậu và hỗ trợ khách hàng trong các chiến lược ESG của họ.



Hình 5. Biểu đồ khoảng cách giữa các chủ đề của Báo cáo ESG của Citibank 2022

Đa dạng, công bằng và hòa nhập (10,6%): Chủ đề này phản ánh cam kết của Citibank trong việc thúc đẩy môi trường làm việc đa dạng và hỗ trợ các nhóm dễ bị tổn thương, đặc biệt là phụ nữ và nhóm thiểu số.

Báo cáo năm 2022 phản ánh cách tiếp cận toàn diện và chiến lược đối với tính bền vững, tập trung mạnh vào quản lý rủi ro, quản trị doanh nghiệp, tài chính bền vững và trách nhiệm xã hội. Các chủ đề chính như chuyển đổi kinh tế xanh, tác động của khí hậu, đa dạng và hòa nhập làm nổi bật cam kết của Citibank trong việc lồng ghép các giá trị ESG vào hoạt động của mình. Những nỗ lực này không chỉ đáp ứng kỳ vọng của các bên liên quan mà còn đóng vai trò quan trọng trong việc định hình một tương lai công bằng, bền vững và có trách nhiệm hơn.

3.5. Kết quả cho câu hỏi nghiên cứu 5

Hình 6 là biểu đồ khoảng cách liên chủ đề của báo cáo ESG toàn cầu của Citibank 2023, minh họa mối quan hệ giữa các chủ đề ESG khác nhau.

Quản lý rủi ro & quản trị ESG (13,8%): Chủ đề này tập trung vào việc đưa các yếu tố môi trường, xã hội và quản trị (ESG) vào cách tổ chức quản lý rủi ro và vận hành bộ máy. Mục tiêu là xây dựng các nguyên tắc và chính sách để đảm bảo rằng các vấn đề ESG được xem xét cẩn trọng trong quá trình ra quyết định và mọi hoạt động, đặc biệt là việc ứng phó với những thách thức và tiềm năng từ biến đổi khí hậu.

Xác minh khí thải nhà kính & báo cáo môi trường (19,2%): Chủ đề này tập trung vào quá trình đảm bảo tính chính xác và minh bạch của dữ liệu về lượng khí thải nhà kính cũng như các thông tin liên quan đến môi trường. Mục tiêu hàng đầu là tuân thủ nghiêm ngặt các tiêu chuẩn báo cáo, đồng thời cung cấp thông tin đáng tin cậy và dễ hiểu về tác động môi trường của tổ chức, bao gồm cả phạm vi và cách thức trình bày trong các báo cáo.

Dự án tác động & tài chính bền vững (19,4%): Chủ đề này đề cập đến các hoạt động và kế hoạch cụ thể mà tổ chức triển khai nhằm tạo ra những ảnh hưởng tích cực đến môi trường và cộng đồng. Vai trò của các công cụ tài chính trong việc hỗ trợ những hoạt động này được đặc biệt chú trọng, song song với việc đánh giá một cách khách quan và hiệu quả các kết quả tác động mang lại.

Kinh doanh bền vững & hỗ trợ xã hội (11,8%): Chủ đề này hướng tới việc xây dựng một mô hình kinh doanh có trách nhiệm và lâu dài, trong đó các yếu tố môi trường, xã hội và quản trị (ESG) được tích hợp một cách sâu sắc vào các hoạt động kinh doanh cốt lõi. Bên cạnh đó, chủ đề này cũng bao gồm các hoạt động hỗ trợ cộng đồng, phát triển nguồn nhân lực và các chương trình trách nhiệm xã hội khác, cùng với việc cung cấp các sản phẩm và dịch vụ thân thiện với môi trường và mang lại lợi ích cho xã hội.

Nhìn vào các lĩnh vực chính, phân tích báo cáo ESG năm 2023 của Citibank cho thấy một chiến lược ESG toàn diện và có chiều sâu. Hai chủ đề cốt lõi là xác minh khí thải & báo cáo môi trường và dự án tác động & tài chính bền

vững được đặc biệt nhấn mạnh, cho thấy cam kết mạnh mẽ của tổ chức đối với trách nhiệm bảo vệ môi trường, cũng như vai trò chủ động trong việc tài trợ và thực hiện các sáng kiến mang lại những thay đổi tích cực cho cả môi trường và cộng đồng.



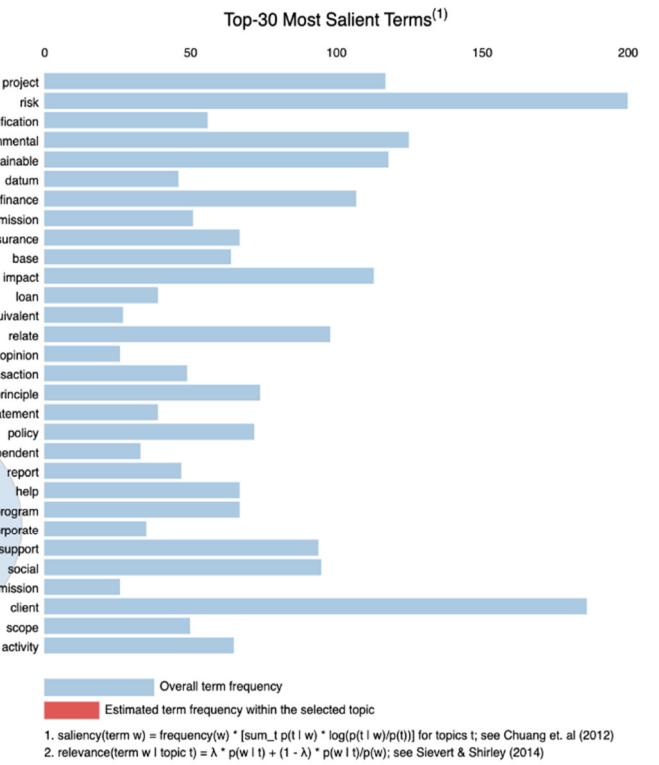
Hình 6. Biểu đồ khoảng cách giữa các chủ đề của Báo cáo ESG Citibank 2023

Bảng 5. Phân loại chủ đề và từ khóa trích xuất từ báo cáo ESG năm 2023 Citibank

Chủ đề	Từ khóa	Tên chủ đề đề xuất	Tỷ lệ
0	risk, client, include, management, relate, principle, policy, climate, information, provide	Quản lý rủi ro & Quản trị ESG	13,8%
1	verification, datum, equivalent, opinion, emission, base, ghg_emission, statement, scope, report	Xác minh phát thải nhà kính & Báo cáo môi trường	19,2%
2	project, environmental, assurance, impact, loan, finance, transaction, social, independent, corporate	Dự án tác động & Tài chính bền vững	19,4%
3	sustainable, business, support, work, finance, service, help, program, product, provide	Kinh doanh bền vững & Hỗ trợ xã hội	11,8%

3.6. Kết quả cho câu hỏi nghiên cứu 6

Phân tích so sánh các báo cáo ESG của Citibank từ năm 2019 đến 2023 cho thấy, sự tiến triển rõ rệt trong các ưu tiên chủ đề. Theo thời gian, các báo cáo thể hiện mức độ



chuyên sâu và chi tiết kỹ thuật ngày càng cao, đồng thời vẫn duy trì sự nhất quán đối với các lĩnh vực cốt lõi như phát triển bền vững, quản trị doanh nghiệp và quản lý rủi ro. Quá trình này phản ánh mức độ trưởng thành ngày càng cao của Citibank trong lĩnh vực ESG, cũng như sự phù hợp ngày càng chặt chẽ với các tiêu chuẩn công bố thông tin quốc tế.

Năm 2019, các nội dung công bố ESG của Citibank tập trung vào những chủ đề rộng như phát triển bền vững, sự tham gia của nhân viên, quản trị doanh nghiệp và các mục tiêu tài chính - môi trường. Đây là những chủ đề nền tảng, đặt cơ sở cho các sáng kiến cụ thể hơn trong những năm tiếp theo.

Đến năm 2020, ngân hàng bắt đầu đưa vào những chủ đề mang tính chuyên biệt hơn như hành động khí hậu và quản lý rủi ro môi trường - xã hội. Một điểm mới đáng chú ý là sự nhấn mạnh vào quyền riêng tư và an ninh dữ liệu, phản ánh quá trình chuyển đổi số cũng như mối quan tâm ngày càng tăng đối với bảo vệ dữ liệu trong bối cảnh đại dịch COVID-19.

Năm 2021 chứng kiến sự chuyển dịch đáng kể sang các chủ đề liên quan đến quyền con người và giám sát chuỗi cung ứng, vốn trở thành chủ đề nổi bật nhất trong năm. Các chủ đề như chống tham nhũng, tiếp cận cộng đồng và tài chính bền vững cũng phản ánh mức độ gắn kết sâu sắc hơn với các tiêu chuẩn ESG toàn cầu và chuẩn mực đạo đức trong kinh doanh.

Báo cáo ESG năm 2022 thể hiện sự phân bố chủ đề cân bằng hơn, với các nội dung chính gồm quản lý rủi ro, tài trợ dự án bền vững và đa dạng, công bằng và hòa nhập. Điều này cho thấy nỗ lực tích hợp ESG không chỉ ở cấp độ vận hành mà còn trong văn hóa tổ chức của doanh nghiệp.

Đến năm 2023, báo cáo ESG của Citibank thể hiện rõ độ chính xác kỹ thuật và nhận thức pháp lý cao hơn, với trọng tâm là xác minh phát thải nhà kính, tài chính bền vững và quản trị rủi ro. Đồng thời, ngân hàng tiếp tục chú trọng việc hỗ trợ các hoạt động kinh doanh phản ánh các giá trị ESG, cho thấy một cách tiếp cận toàn diện nhằm tích hợp phát triển bền vững vào chiến lược và vận hành.

Sự thay đổi theo hướng chuyên sâu này cho thấy khả năng của Citibank trong việc điều chỉnh cách truyền thông ESG nhằm đáp ứng các yêu cầu toàn cầu đang thay đổi và ưu tiên nội tại. Việc tập trung vào các kết quả có thể đo lường phản ánh một sự chuyển dịch có chủ đích hướng đến tính minh bạch và tích hợp phát triển bền vững dài hạn. Khi báo cáo ESG trở thành một công cụ chiến lược thay vì chỉ là một yêu cầu tuân thủ, quá trình phát triển của Citibank có thể được xem như một ví dụ điển hình về cách thức gắn kết hiệu quả với các bên liên quan trong bối cảnh doanh nghiệp hiện đại.

4. KẾT LUẬN

Nghiên cứu sử dụng mô hình Latent Dirichlet Allocation (LDA) để xác định các chủ đề có trong các báo ESG của Citibank từ năm 2019 đến năm 2023. Kết quả chỉ ra mức độ hiệu quả của LDA trong việc tìm từ khóa, xác định và phân nhóm các từ khóa liên quan đến nhau. Các chủ đề được gán cho nhóm các từ khóa và dựa trên tần suất xuất hiện để phản ánh chiến lược của Citibank đối với chỉ số ESG. Mô hình LDA cung cấp một phương pháp phân tích chủ đề dựa trên dữ liệu văn bản, đem lại những thông tin giá trị cho các bên liên quan, nhà nghiên cứu và những người ra quyết định trong tổ chức của Citibank.

TÀI LIỆU THAM KHẢO

- [1]. Eccles R.G., Krzus M.P., Solano C., "A Comparative Analysis of Integrated Reporting in Ten Countries," *SSRN*, 2019. <https://doi.org/10.2139/ssrn.3345590>.
- [2]. Edgetrader, *Analysing ESG report using Natural Language Processing*, 2025, <https://github.com/edgetrader/esg-nlp>.
- [3]. Barde B.V., Bainwad A.M., "An overview of topic modeling methods and tools," in *2017 International Conference on Intelligent Computing and Control Systems (ICICCS)*, 745-750, 2017. <https://doi.org/10.1109/ICCONS.2017.8250563>.
- [4]. Manning C.D., Raghavan P., Schütze H., *Introduction to Information Retrieval*. Cambridge University Press, 2008.
- [5]. Steven B., Ewan K., Edward L., *Natural Language Processing with Python*. O'Reilly Media, Inc., 2009.
- [6]. Daniel J., James H. M., *Speech and Language Processing: An Introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition*. Online manuscript released January 12, 2025. https://web.stanford.edu/~jurafsky/slp3/ed3book_Jan25.pdf
- [7]. Blei D.M., Ng A.Y., Jordan M.I., "Latent dirichlet allocation," *Journal of Machine Learning Research*, 3, 993-1022, 2003.

AUTHORS INFORMATION

Truong Cong Doan, Pham Thi Phuong Thao, Ha Thi Hong Anh, Do Thanh Van, Nguyen Bich Diep

International School, Vietnam National University, Hanoi, Vietnam