21 April, 2025

Leveraging Large Language Models for KPI and Causal Relationship Extraction from Social Media Companies' SEC 10-Q Financial Reports

Alex Matrajt Frid
University of Michigan, Ann Arbor
alexmat@umich.edu

Eytan Adar
University of Michigan, Ann Arbor
eadar@umich.edu

1. Introduction

As social media firms continue to grow, reshaping the way people communicate and consume information, they constantly face a complex web of operational, reputational, and financial risks [1]. These risks are often detailed in the "Risk Factors" section of their SEC filings. While these disclosures are intended to inform investors, they also provide unique insights into what these companies deem essential for their success. In this research, we ask a fundamental question: What do social media companies truly care about from a financial performance perspective—and how have these priorities shifted over time?

Manually analyzing such disclosures at scale, however, presents a significant challenge due to the volume and complexity of financial documents. To address this we turn to Large Language Models (LLMs), which have revolutionized Natural Language Processing by offering powerful tools for automating the analysis of extensive documents, such as SEC filings [2]. This research also seeks to investigate the following question: Can these models help uncover meaningful cause-and-effect relationships that reveal corporate priorities within complex financial texts?

To address these questions, we developed a framework utilizing LLMs to extract chains of events affecting Key Performance Indicators (KPIs) from the "Risk Factors" section of social media companies' 10-Q reports. By analyzing these KPIs and chains of events—termed "causal chains" in this study—we gain a deeper understanding of these companies' corporate priorities and address our main research question. Furthermore, by comparing financial reports across years, we uncover how these priorities have shifted over time in response to market trends, user preferences, technological changes, and regulatory pressures.

By automating the analysis of financial documents, we propose an innovative framework that could transform the financial industry's approach to data-driven decision-making.

2. Related Work

Prior work on financial text analysis has largely focused on extracting KPIs and linking them to their associated numerical values or contextual attributes. For example, the KPI-EDGAR system introduces an Electronic Data Gathering, Analysis, and Retrieval (EDGAR) framework designed to extract KPIs from financial documents and link them to their numerical values and other attributes [3]. While this line of research provides a framework for extracting KPIs and their associated data, it fails to explore causal dependencies within the financial report—a limitation our work addresses.

Another notable study applies LLMs to the KPI-EDGAR dataset [3], automating the extraction of KPIs from financial reports and achieving higher accuracy and efficiency in linking KPIs to their corresponding values than the previous EDGAR technique [4]. Although this approach leverages LLMs for KPI extraction, it remains focused on linking KPIs to their values and attributes.

While these methods focus on identifying KPI mentions and linking them to data, our work goes a step further by using LLMs to extract narratives describing how specific conditions impact KPIs. We shift the focus from KPI-value linking to the extraction of embedded relationships within the reports. This allows us to understand the priorities and risk factors of these companies, as well as identifying the KPIs that are most affected, thereby gaining insights into their strategic priorities and vulnerabilities.

3. Foundational Concepts

This research centers around three core concepts: KPIs, 'causal factors', and 'causal chains'. While KPIs are well-established in business literature, the latter two concepts are introduced in this study.

3.1 KPIs

Key Performance Indicators (KPIs) are essential metrics used to evaluate a business's success [5]. They are crucial to this study because they define the financial metrics that a company prioritizes.

We identified 53 KPIs that are particularly relevant to social media companies, drawing from 10-Q reports of Meta, Snap, and Reddit [6, 7, 8, 9, 10]. These KPIs were selected based on their measurability and numerical computability, ensuring that our analysis remains grounded in quantifiable data. A list with the selected KPIs can be found in Appendix 2.

While some KPIs are often explicitly mentioned in these reports (e.g., Revenue, Profit, etc.), others require inference. This is where the use of an LLM becomes invaluable. The LLM's ability to infer context beyond the explicit text enables us to capture implicit KPIs [11]. However, this capability also introduces the challenge of balancing inference with accuracy, as too much inference could lead to fabricated information. We addressed this challenge in our model selection process, which is elaborated upon in the subsequent sections.

3.2 Causal Factors

We introduce the concept of a 'causal factor'—an event or change that potentially impacts one or more KPIs. Since our analysis focuses on the risks section of financial reports, these causal factors typically represent negative impacts on KPIs. Given the high variability of causal factors across different reports and companies, we opted against creating a predefined list of them. Instead, the LLM generates these factors based on the text it analyzes, preserving the report's original context.

3.3 Causal Chains

A causal factor can lead to another causal factor. To chain these events we introduce the 'causal chain' structure. These chains contain sequences of causal factors and the KPIs they affect, and are formatted in a JSON structure. To link causal factors, we add a 'comes from' object, which describes the relationship between them. Figure 1 illustrates this structure through a simple example, showing both the JSON representation and its visualization. As seen in the figure, causal factors leading to other causal factors and a KPI being affected are represented with arrows. A KPI can be affected by multiple causal factors. In the example, 'Customer Trust' was impacted by both 'Unfavorable media coverage' and 'Brand Reputation Damage'.


```
"causal_chain": [
  {
    "causal_factor": "Unfavorable media coverage",
    "KPIs_impacted": ["Customer Trust",
                      "Investor Confidence"],
    "comes_from": []
  },
  {
    "causal_factor": "Brand Reputation Damage",
    "KPIs_impacted": ["Net Promoter Score",
                      "Customer Satisfaction Score",
                      "Customer Trust"],
    "comes_from": ["Unfavorable media coverage"]
  },
  {
    "causal_factor": "Decline in stock price",
    "KPIs_impacted": ["Stock Price",
                      "Market Capitalization",
                      "Trading Volume"],
    "comes_from": ["Brand Reputation Damage"]
  }
]
```


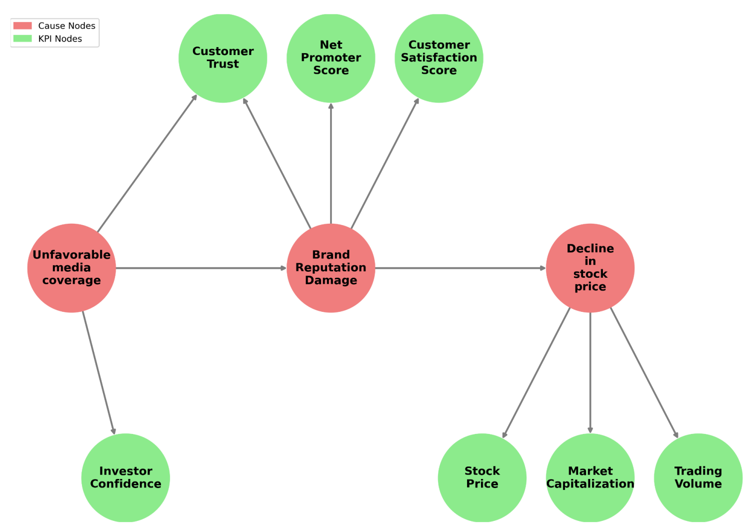

Figure 1. Causal Chain JSON structure and Graph Visualization Example

A single paragraph can generate multiple causal chains, which are collected in a 'causal_chains' object. This structure enables comprehensive analysis of KPI and causal factor relationships within the report

4. Methodology

Our methodology involves extracting the “Risk Factors” section from SEC reports and processing it paragraph by paragraph using a designed prompt to extract the causal chains contained in each paragraph.

4.1 Text extraction

To extract the text, we begin by downloading the SEC report from the U.S. Securities and Exchange Commission website as an HTML file [12]. First, we remove any unnecessary information such as page numbers, headers (e.g., Table of Contents), and links. Once the text is cleaned, we define the start and end patterns for the Risks section. This section consistently begins with the subtitle "Item 1A. Risk Factors" and concludes with the subsequent section titled "Item 2. Unregistered Sales of Equity Securities and Use of Proceeds." After isolating the relevant section, we remove subtitles, as they do not contain cause-effect information and often result in empty causal chains.

It should be noted that most SEC reports include bulleted lists, which may contain complete independent sentences or numerous incomplete and dependent ideas. To process the latter, we identify bulleted lists with semicolons separating each point, indicating they are part of a cohesive idea. In such cases, we merge the entire list into a single paragraph to better reflect their interrelatedness, preserving the dependent nature of these sentences. Once fully processed, the text is saved into a text file, with each line representing a paragraph from the report.

4.2 Generation of causal chains

For generating causal chains, we employ the GPT-4o-2024-08-06 model [13]. This model is better for tasks requiring reasoning and inference compared to previous models, making it ideal for extracting causal chains [14]. While tests with GPT-3.5 Turbo [15] were conducted, the results were inconsistent, and KPI inference proved challenging. GPT-4's ability to produce JSON outputs further aligns with our need to represent causal chains in Python dictionaries [14]. We opted not to fine-tune the model, as this process requires a metric to evaluate output success. Given that our output is a dictionary, evaluating correctness is complex; dictionary outputs are not easily comparable to a training set, especially because of the variability in LLM responses, which are based on probability distributions [16].

The model is accessed via API calls to OpenAI. Each paragraph is individually embedded into the designed prompt and sent to the model. The model returns the output in a JSON structure, which is then saved in a JSON file. This file collects all the causal chains generated from the risks section of that report.

4.3 Prompt Design

The prompt is built to guide the LLM in extracting the KPIs and causal factors from each paragraph and formatting them into causal chains. The complete prompt is detailed in Appendix 1 and follows this formula:

Context and High-Level Task Description + Output Format Specification + Detailed Task Description + Example of Expected Output + List of KPIs + Paragraph from SEC Report

This structure progresses from high-level to low-level details, taking advantage of the tendency of LLMs to better retain information presented at the end of the prompt [11]. During testing, we observed that the LLM struggled most with causal chain generation—especially with recognizing that a single paragraph can yield multiple causal chains—and with remembering the KPI list. Therefore, the detailed task description, comprehensive example, and KPI list are positioned towards the end of the prompt to ensure these elements receive the highest attention. Despite this strategy, the LLM occasionally generates KPIs not included in the provided list, requiring post-generation data cleaning. We chose to remove these extraneous KPIs, as they were a small percentage and often did not meet our measurability criteria. Finally, the paragraph being analyzed is placed at the end to ensure it is accurately processed, as it is the primary focus of the analysis.

Another line of research on LLM prompting has shown that including examples at the end of prompts significantly enhances LLM performance [17]. For this reason, our prompt includes a detailed example that specifies the output format of two causal chains generated from one example paragraph. This comprehensive example clarifies the expected output, reinforcing the desired behavior [11].

Our framework processes paragraphs individually, meaning the LLM does not have access to previously analyzed portions of the report. We experimented with adding contextual information to the prompt, such as sets of already extracted KPIs and causal factors. However, this approach reduced output accuracy by diverting attention from the prompt's critical components and overwhelming it with excessive information. Additionally, it increased the prompt's length, diminishing the importance of information placed at the beginning. It also raised token usage which increased generation costs, making this strategy impractical.

5. Evaluating the Method

To assess the reliability of our extraction process, we present an applied case using Meta's September 30, 2024 10-Q Report [8]. We conducted the causal chain extraction process three times to ensure the framework's repeatability and to evaluate the consistency of the results across runs.

5.1 KPI Stability

Due to the inherent variability in LLM responses, the total number of KPIs generated varied across runs: 1526, 1536, and 1452, respectively, with an average of 1505 KPIs per report. Their distribution can be observed in Figure 2.

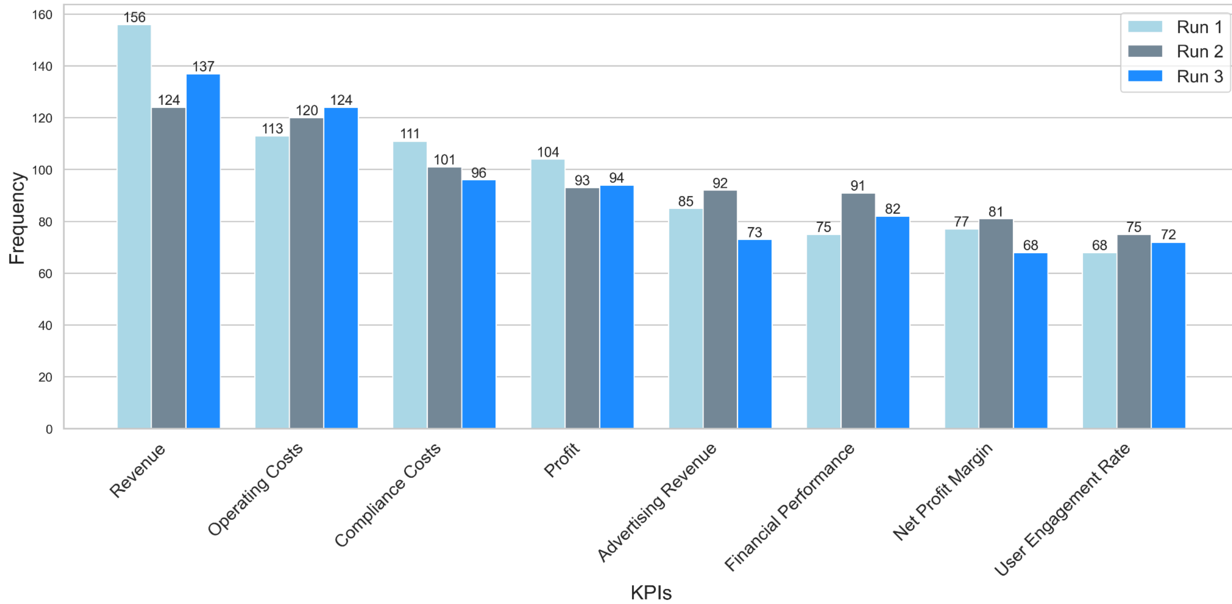

Figure 2: Top 10 KPIs Comparison Across Three Runs of Meta's September 30, 2024 10-Q Report

As shown in Figure 2, the top 10 KPIs were consistent across all runs, although some in slightly different orders. This variation can be attributed to the differing total number of KPIs generated per report. Although only the top 10 KPIs are plotted, the others followed similar trends. This consistency demonstrates the method's reliability in extracting KPIs.

Notably, KPIs related to revenue showed more variability within runs. This is because we included multiple revenue-related KPIs (e.g., 'Revenue', 'Advertising Revenue', "Average Revenue Per User", "Revenue Growth Rate", "International Revenue") to capture different aspects of revenue, especially relevant for social media firms. A limitation of using an LLM is that responses may vary slightly each time, leading to different labeling of similar KPIs across runs.

Out of the 53 KPIs in our set list, only the second run captured all of them. The first run missed 'Debt-To-Equity Ratio,' and the third run missed 'Bounce Rate' and 'Debt Issuance.' This variability is due to the potential for similar KPIs to be labeled differently in each run. However, these three KPIs appeared only once or twice in the second run, making their impact negligible given the average of 1505 KPIs across runs.

With this plot we also gain initial insights into Meta's ten most important KPIs for 2024. With the highest frequency, 'Revenue' emerged as the most critical KPI for the company.

5.2 Causal Factor Stability

The LLM-generated causal factors exhibited high variability in naming. The total number of distinct causal factors found per run was 869, 857, and 841, with the first run capturing the most.

To evaluate causal factor extraction, we categorized them into 10 groups to standardize them and enable cross-report comparisons. The categories are:

1. **Technology & Infrastructure Issues**: Platform reliability, cybersecurity, system outages, or scalability challenges.
2. **Financial & Economic Risks**: Revenue fluctuations, cost pressures, inflation, or macroeconomic instability.
3. **Regulatory & Legal Pressures**: Changing laws, government investigations, compliance burdens, or lawsuits.
4. **User Trust & Engagement Risks**: Loss of user trust, privacy concerns, negative public perception, or declining user activity.
5. **Competitive Pressure**: Threats from new entrants, pricing wars, innovation gaps, or shrinking market share.
6. **Operational & Business Execution Risks**: Internal mismanagement, poor coordination, supply chain issues, or missed deliverables.
7. **Strategic planning and Leadership Risks**: High-level missteps such as failed expansions, leadership turnover, or poor long-term vision that misalign the company with market needs.
8. **Global & Political Risks**: Trade disputes, geopolitical instability, or region-specific regulatory uncertainty.
9. **Product Quality & Reliability Issues**: Defects, underperformance, lack of innovation, or technical shortcomings affecting user satisfaction.
10. **Other**

We used the LLM once again but now as a classifier, employing the GPT-4o-2024-08-06 model [13] to categorize all of the generated causal factors. Each factor was fed individually, with the model returning the corresponding category number. The full prompt can be found in Appendix 3. We built a JSON dictionary with causal factors as keys and group numbers as values, which was used to map the causal factors to their categorized names for comparison.

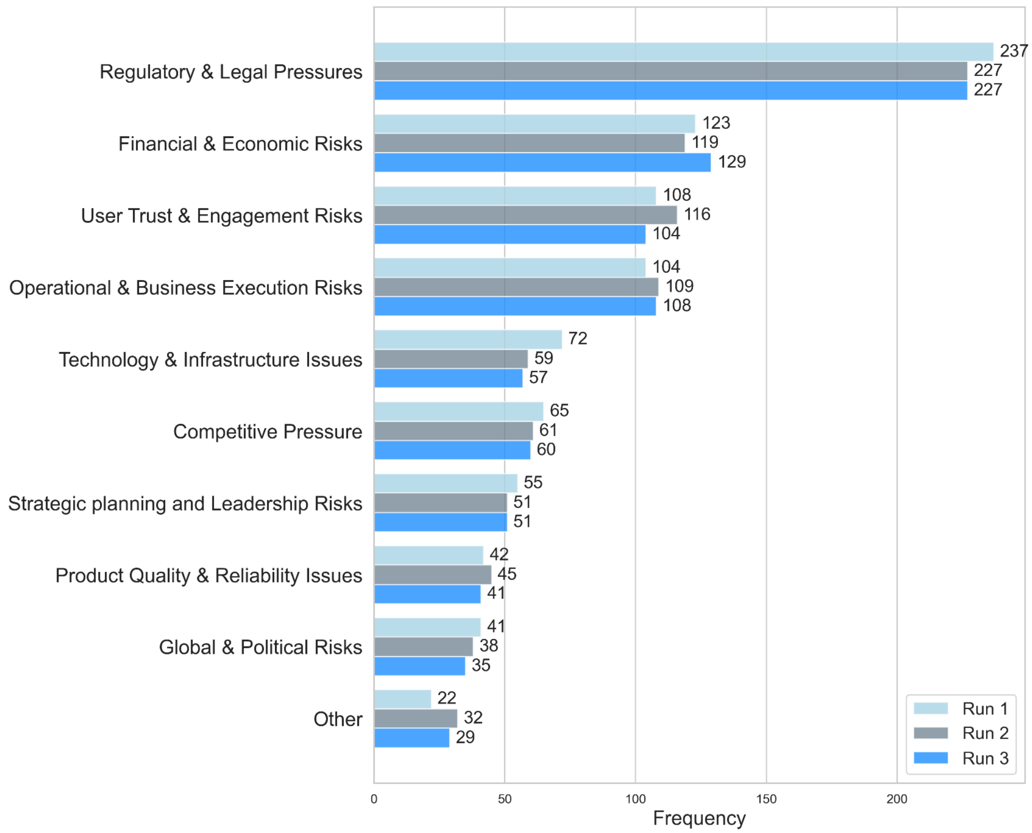

Figure 3. Causal Factor Category Distribution Across Three Runs of Meta's September 30, 2024 10-Q Report

Figure 3 shows that causal factor categories are stable across runs. Apart from ‘Technology and Infrastructure Issues’, distributions are similar. This small variation is likely due to the first run generating more causal factors. The hierarchy of causal factor frequency was also stable across runs, with only a minor flip between run 2 and 3 in the ‘Technology and Infrastructure Issues’ and ‘Competitive Pressure’ categories, demonstrating the methodology's stability and reproducibility. This plot also helps identify the causal factor categories posing the biggest threat to the company, with ‘Regulatory and Legal Pressures’ posing the largest threat to Meta in 2024.

5.3 Causal Chain Stability

We also examined the stability of causal chains by comparing the distribution of causal factor categories impacting each KPI across runs. This ensures the relationships between KPIs and causal factors are reproducible across runs. We present an analysis with the top 5 KPIs, but similar trends were observed for the other KPIs.

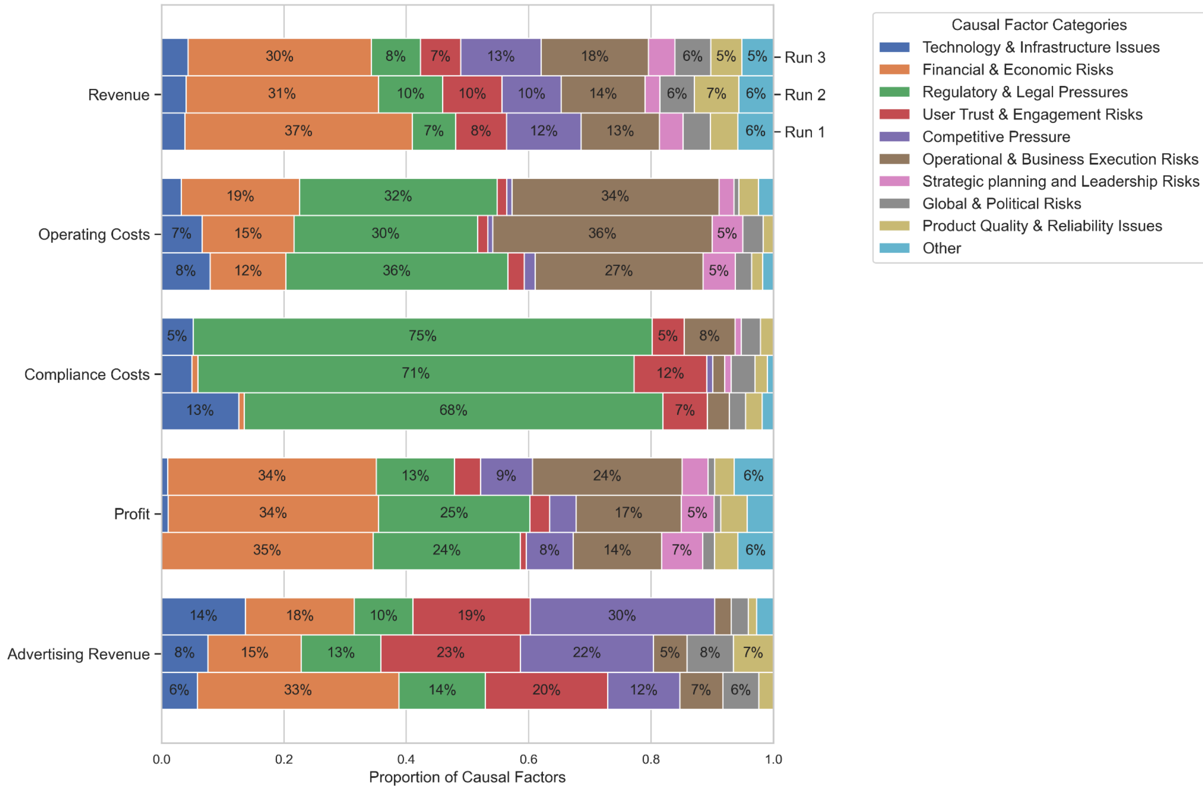

Figure 4: Distribution of Causal Factor Categories Affecting Top 5 KPIs Across Three Runs of Meta's September 30, 2024 10-Q Report

In Figure 4, each top 5 KPI is plotted with its corresponding three runs, showing the percentage of causal factor categories affecting that KPI. For example, in the first run, 30% of the causal factors affecting ‘Revenue’ belonged to the ‘Financial and Economic Risks’ category.

The top 3 KPIs show similar causal factor distributions across runs, with only minor percentage differences. However, discrepancies exist in the other 2 KPIs, particularly 'Advertising Revenue.' This issue stems from categorizing causal factors with the LLM, as some can belong to multiple categories (e.g., 'Reduced user engagement due to competitive products like TikTok' could fit both 'Competitive Pressure' and 'User Trust and Engagement Risks'). For simplicity, we assigned only one category per factor, which explains some variability in causal factor distribution within KPIs, as the LLM may have made different decisions in each run. This suggests that while our framework is a solid foundation, it could benefit from refinement in future research. The interaction between causal factors and KPIs is generally consistent but not perfect, indicating room for improvement.

6. Insights Gained Using the Method

With the KPI and causal factor extraction methodology proven reliable and stable, we present two insightful analyses utilizing causal chains: a network analysis of the causal factor-KPI relationships within the same Meta 10-Q report [8], and a temporal analysis showing how KPIs and causal factors evolved over time. This analysis uses Meta's September 30, 2020, 2022, and 2024 10-Q Reports [6, 7, 8].

6.1 Network Analysis

To visualize all the causal chains, we graph them the same way as in the example from Figure 1, but now including all chains generated from Meta's September 30, 2024 10-Q Report [8]. We focus on the second run, which captured all KPIs in the list. The resulting plot shows a complex web of causal relationships, as shown in Figure 5.

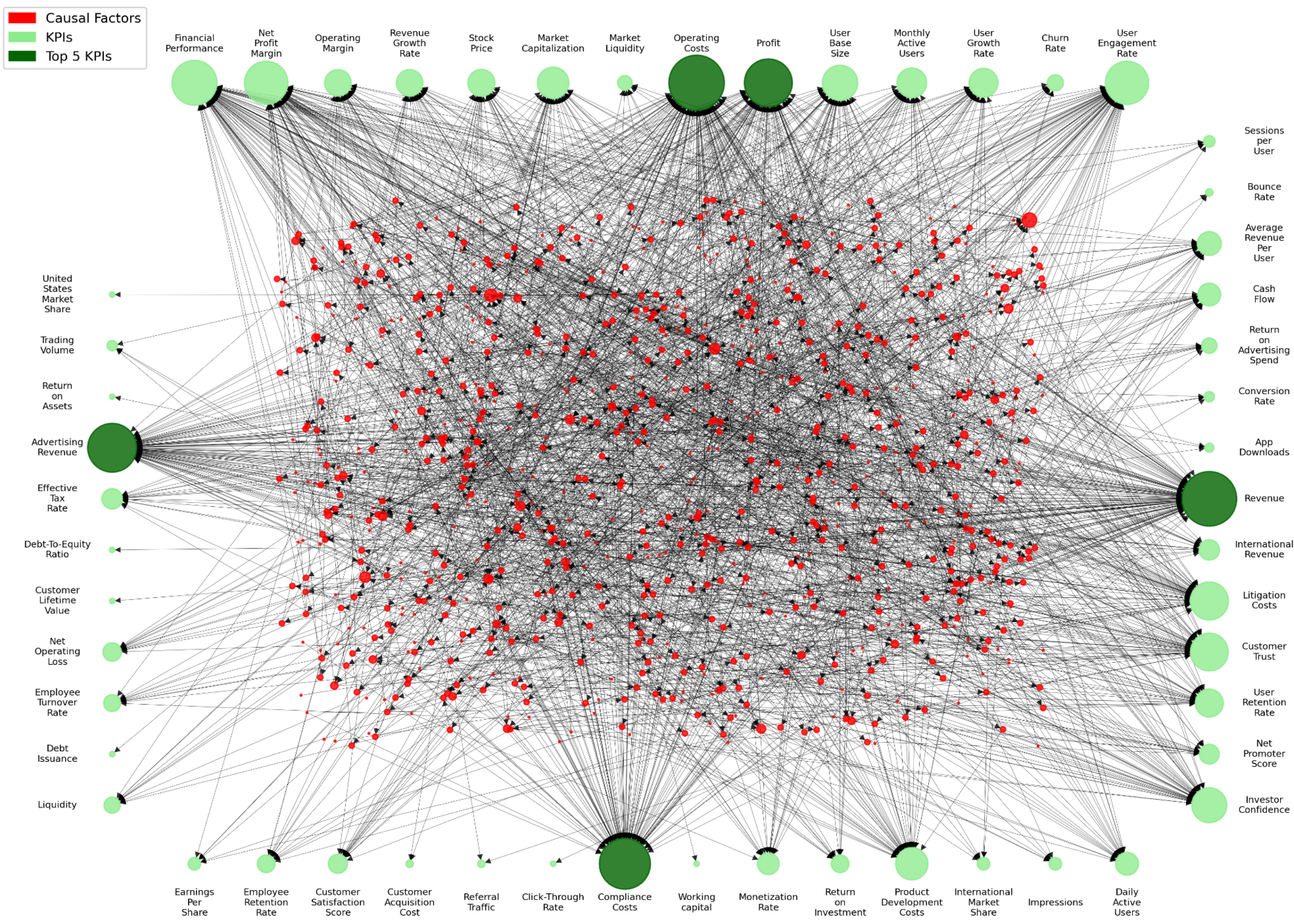

Figure 5: Network of All Causal Chains in Meta's September 30, 2024 10-Q Report

In this network, KPIs are positioned along the border. The causal factors interact in the center, forming causal chains where each step impacts KPIs—interactions that are represented with arrows. Node size indicates the in-degree, while arrow thickness reflects the frequency of interactions. Due to minimal overlap in causal factor names, arrow thickness is less pronounced. The network's complexity arises from the large number of distinct causal factors in the report (857).

The most significant insight comes from the KPI node sizes, which indicate the number of causal factors impacting each KPI. Larger node sizes suggest a KPI is influenced by more causal factors, highlighting its importance and centrality in the network. This insight helps identify which KPIs are most affected by various risks and conditions, providing a clearer understanding of the company's financial priorities and vulnerabilities. The top 5 KPIs are highlighted in dark green.

6.2 Time Analysis

A significant insight from our research comes from the temporal analysis of KPIs and causal factors. By comparing the same company across different time periods, we can observe shifts in their priorities and threats—information that is crucial for informed financial decision-making.

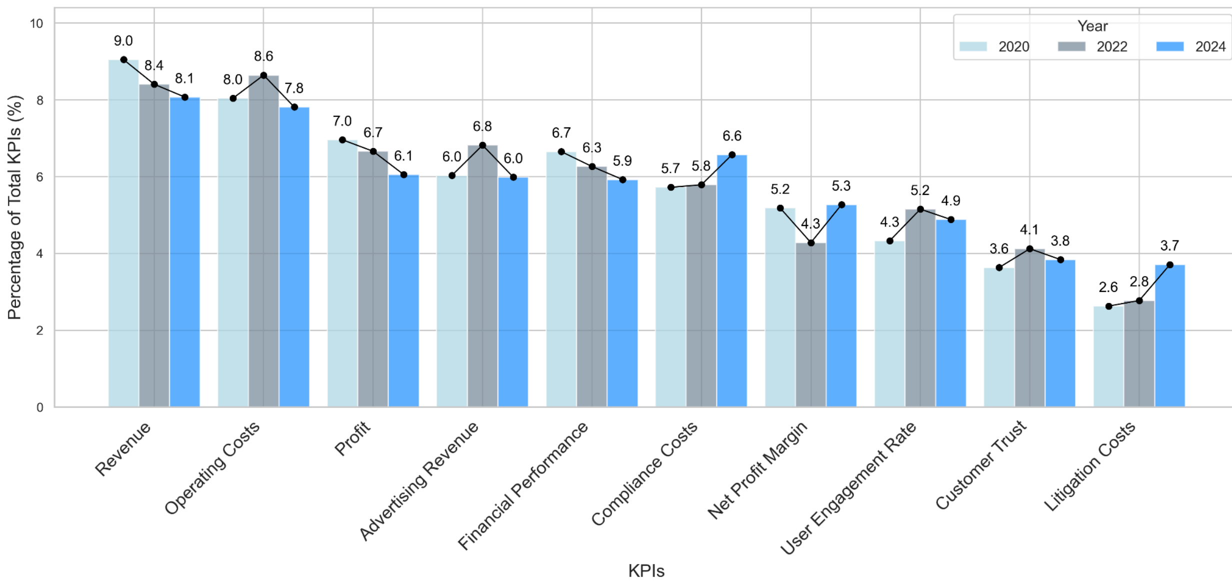

Figure 6: Top KPIs Comparison Across Meta's 2020, 2022, and 2024 10-Q Reports (Normalized by Total KPIs per Report)

Figure 6 compares the top 10 KPIs across Meta's 2020, 2022, and 2024 10-Q reports, providing insights into the relative importance of different KPIs over time. Only the top 10 are included for simplicity, but the same analysis can be done for all 53 KPIs. The data is normalized by total KPIs per report to account for differences in report length, since the 2024 report is significantly longer.

The plot includes line graphs over the bars to illustrate trends over the years, highlighting shifts in the importance of KPIs. We observe a steady decrease in mentions of Revenue, Profit, and Financial Performance, suggesting a reduction in risk factors affecting these KPIs. Conversely, there is a steady increase in the frequency of mentions related to Compliance Costs and Litigation Costs, indicating a rise in associated risk factors.

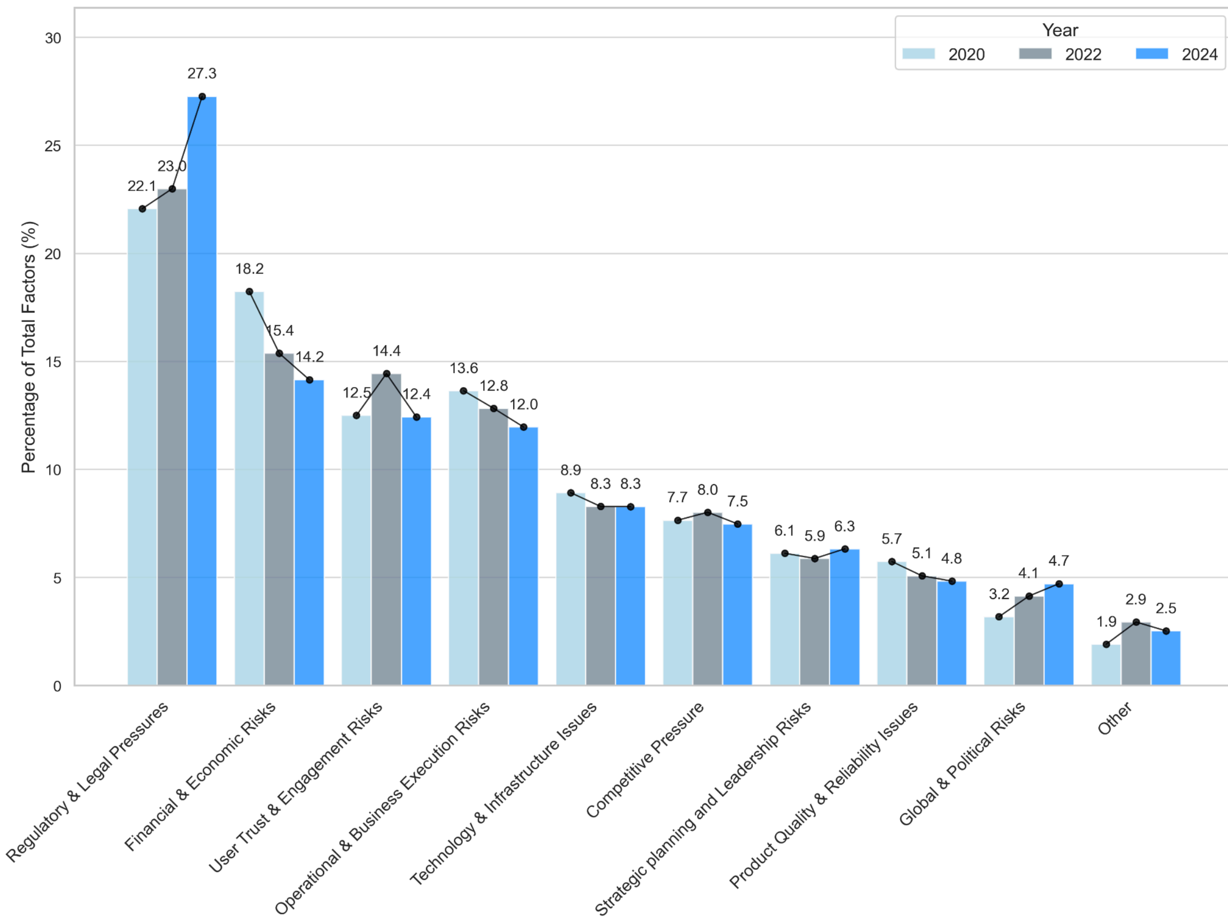

Figure 7: Causal Factor Category Distribution Comparison Across Meta's 2020, 2022, and 2024 10-Q Reports (Normalized by Total Causal Factors per Report)

Figure 7 analyzes causal factor category trends across years, normalized similarly to the KPIs. This plot provides insights into how each causal factor category's threat level has changed. For instance, 'Regulatory and Legal Pressures' increased by 5.2% over the years, indicating a heightened threat level to Meta's KPIs and financial performance, likely due to increased data privacy regulations [18]. Financial analysts can use this method to gain insights into Meta's evolving risks and make more informed decisions.

This framework presents a new approach to tracking how Meta's business priorities and risk factors have evolved, offering valuable information for strategic decision-making.

7. Limitations and Future Work

One significant limitation is the inherent variability in LLM responses, which can lead to inconsistencies in KPI and causal factor extraction across different runs. This variability may affect the reliability of the results. Additionally, the framework is specifically tailored to social media companies, limiting its applicability to other industries without adaptation. Future work could involve modifying the prompt and KPI list to include industry-specific KPIs, thereby broadening the framework's applicability.

Moreover, this study relies on LLMs to infer causal relationships, which may not always reflect the true underlying mechanisms. This could lead to incorrect conclusions about the relationships between KPIs and causal factors. Additionally, the analysis assumes that the frequency with which a KPI is mentioned corresponds to its importance to the company—a simplification that may not always hold true, as some critical metrics might be underreported or discussed indirectly. Future work could incorporate additional data sources or methodologies to validate the inferred causal relationships.

The accuracy of the analysis relies on the quality and completeness of the SEC filings. Missing or incomplete data could impact the findings, and companies may not disclose risks they are unaware of, leading to an incomplete analysis. Future research could explore methods to supplement missing data.

Currently, KPIs generated outside the predefined list are deleted due to their rarity. Future work could explore more sophisticated methods for handling these KPIs, such as integrating them into the analysis if they provide valuable insights.

The current method focuses on direct connections between causal factors and KPIs, without analyzing full causal chains. Future research could explore methods to identify key causal factors that catalyze other factors, providing deeper insights into the network's dynamics.

An additional area for improvement is in the classification of causal factors into groups. Currently, each causal factor is assigned to a single category, which may not fully capture the complexity of their interactions. Allowing causal factors to belong to multiple categories could provide a more complete analysis. Future research could employ vector embeddings to identify clusters of causal factors, enhancing the accuracy and depth of the classification. This approach could reveal more intricate patterns and relationships within the data, offering richer insights into the dynamics of risk factors affecting social media companies.

8. Conclusions

In this study, we developed a framework leveraging LLMs to extract and analyze causal relationships from the "Risk Factors" section of social media companies' SEC filings. By focusing on KPIs and the risk factors affecting them, we aimed to uncover the financial priorities and perceived threats of these companies and how they have evolved over time.

Our results suggest that LLMs are capable of extracting meaningful cause-and-effect relationships from complex financial texts and formatting them into structured causal chains. This approach facilitates the efficient analysis of large volumes of text, such as SEC filings, and allows for a comprehensive network and temporal analysis of causal factors and KPIs. These insights provide investors and stakeholders with a clearer understanding of corporate priorities and risk factors, enabling more informed decision-making.

This study shows the transformative potential of LLMs in automating and enhancing the analysis of financial documents. Future research could focus on refining the methods for classifying and analyzing causal relationships, potentially expanding the framework's applicability to other industries and enhancing its impact across various domains.

References

[1] Cheryl DuBose. 2011. The Social Media Revolution. Radiologic Technology, 83, 2, 112. ISSN: 0033-8397.

[2] Mengting Wan, Tara Safavi, Sujay Kumar Jauhar, Yujin Kim, Scott Counts, Jennifer Neville, Siddharth Suri, Chirag Shah, Ryen W. White, Longqi Yang, Reid Andersen, Georg Buscher, Dhruv Joshi, and Nagu Rangan. 2024. TnT-LLM: Text Mining at Scale with Large Language Models. In Proceedings of the 30th ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD '24). Association for Computing Machinery, New York, NY, USA, 5836–5847. https://doi.org/10.1145/3637528.3671647

[3] T. Deußer, et al. 2022. KPI-EDGAR: A Novel Dataset and Accompanying Metric for Relation Extraction from Financial Documents. In Proceedings of the 21st IEEE International Conference on Machine Learning and Applications (ICMLA), Nassau, Bahamas, 1654–1659. DOI:https://doi.org/10.1109/ICMLA55696.2022.00254.

[4] T. Deußer, et al. 2024. Leveraging Large Language Models for Few-Shot KPI Extraction from Financial Reports. In Proceedings of the 2024 IEEE International Conference on Big Data (BigData), Washington, DC, USA, 4864–4868. DOI:https://doi.org/10.1109/BigData62323.2024.10825458

[5] Bernard Marr. 2012. Key Performance Indicators (KPI): The 75 Measures Every Manager Needs To Know. Ashford Colour Press Ltd, Gosport, Hampshire, Great Britain. ISBN 978-0-273-75011-6.

[6] Facebook, Inc. 2020. Form 10-Q, September 30, 2020. Retrieved from https://www.sec.gov/ix?doc=/Archives/edgar/data/0001326801/000132680120000084/fb-09302020x10q.htm.

[7] Meta Platforms, Inc. 2022. Form 10-Q, September 30, 2022. Retrieved from https://www.sec.gov/ix?doc=/Archives/edgar/data/0001326801/000132680122000108/meta-20220930.htm.

[8] Meta Platforms, Inc. 2024. Form 10-Q, September 30, 2024. Retrieved from https://www.sec.gov/ix?doc=/Archives/edgar/data/0001326801/000132680124000081/meta-20240930.htm.

[9] Reddit, Inc. 2024. Form 10-Q, March 31, 2024. Retrieved from https://www.sec.gov/ix?doc=/Archives/edgar/data/0001713445/000171344524000006/rddt-20240331.htm.

[10] SNAP INC. 2024. Form 10-Q, June 30, 2024. Retrieved from https://www.sec.gov/Archives/edgar/data/1564408/000156440824000050/snap-20240630.htm#if71ca3086f57414a89511ca0e41512c5_136.

[11] Noah MacCallum and Julian Lee. 2025. GPT-4.1 Prompting Guide. OpenAI CookBook. Retrieved from https://cookbook.openai.com/examples/gpt4-1_prompting_guide. Accessed April 19, 2025.

[12] U.S. Securities and Exchange Commission. 2025. Retrieved from https://www.sec.gov/. Accessed January 31, 2025.

[13] OpenAI. 2024. GPT-4o-2024-08-06. Accessed via API.

[14] OpenAI. 2025. GPT-4 is OpenAI's most advanced system, producing safer and more useful responses. Accessed April 20, 2025. https://openai.com/index/gpt-4/

[15] OpenAI. 2023. GPT-3.5 Turbo. Accessed via API.

[16] Avi Bewtra. 2025. Introduction to Large Language Models: Everything You Need to Know for 2025 [+Resources]. LAKERA AI. Last updated March 27, 2025. Accessed April 20, 2025. https://www.lakera.ai/blog/large-language-models-guide\

[17] Huixue Zhou, Mingchen Li, Yongkang Xiao, Han Yang, and Rui Zhang. 2024. LEAP: LLM instruction-example adaptive prompting framework for biomedical relation extraction. Journal of the American Medical Informatics Association 31, 9 (September 2024), 2010–2018. https://doi.org/10.1093/jamia/ocae147

[18] Osano. 2025. Data Privacy Laws: What You Need to Know in 2025. Updated January 7, 2025. Accessed April 20, 2025. https://www.osano.com/articles/data-privacy-laws

Appendix 1 - Full Prompt for Causal Chain Extraction

messages=[
{"role": "system", "content": 'A business has indicated risks in a paragraph of their SEC filing. Describe the causal chains included in the paragraph that make this a risk. For each step in the causal chain, indicate the KPIs that would be affected. Produce the output as a JSON file. The file should be a list of objects that is contained in a list called causal_chains that contains all the causal_chain(s) found in the paragraph. Each causal_chain object should be the causal factor, KPIs impacted, and comes_from which contains the names of previous causal_factor(s) that led to this specific causal factor. Only continue the causal_chain if the previous causal_factor led to this causal_factor, if that is not the case, create a new causal_chain object in the causal_chains list. Each element in the causal chain should point to the previous causal_factor that it is based on by using the same causal_factor names. Causal_factor(s) should be related to context, for example a causal_factor cannot be "Decline in user retention, growth, or engagement", it should explain what caused that decline, for example "Negative media reports or publicity", we then specify the KPIs impacted within that causal_factor. Here\'s an example of 2 causal_chain(s) generated by the following paragraph, keep the same field names. "Our success depends on our ability to provide users of our products and services with valuable content, which in turn depends on the content contributed by Redditors. We seek to foster a broad and engaged Redditor community. If Redditors do not continue to contribute content and otherwise engage with our platform, and we are unable to provide Redditors with valuable and timely content, our user base and their engagement may decline. A large portion of the content on our platform comes from a small number of Redditors contributing to communities (which are also known as subreddits). If prolific Redditors do not continue to contribute content and otherwise engage with our platform, or decide to leave our platform and encourage other Redditors to follow them to a new platform, our user base and their engagement may decline. Our platform may also be used by third parties to disseminate abusive or other harmful content in violation of our terms and applicable law. We may not proactively discover or quickly respond to such content once alerted to it due to our scale and the limitations of existing technology and operational infrastructure. If we are unable to successfully prevent or detect and timely address abusive or other harmful behavior on our platform, our user base and their engagement may decline. Additionally, in keeping with our mission to bring belonging to everyone in the world, our site-wide content policy is designed to be protective, but not intrusive. If Redditors perceive the content available on Reddit to be offensive, inappropriate, hostile, or otherwise objectionable, we may experience a decline in user activity generally, or among certain demographics. We generate a majority of our revenue from the sale of advertising services. If we experience a decline in the number of Redditors, or a decrease in Redditor growth rate or engagement, including as a result of lack of valuable or appealing content, or the loss of influential Redditors or subreddits, advertisers may not view our products and services as attractive for their marketing expenditures, and may reduce their spending with us, which would harm our reputation, business, results of operations, financial condition, and prospects." Example: [{"causal_chain": [{"causal_factor": "Decline in Redditor engagement", "KPIs_impacted": ["Monthly Active Users", "User Growth Rate", " User Engagement Rate", "Advertising Revenue"], "comes_from": []}, {"causal_factor": "Loss of influential Redditors", "KPIs_impacted": ["Monthly Active Users", "User Retention Rate", "User Engagement Rate", "Advertising Revenue"], "comes_from": ["Decline in Redditor engagement"]}, {"causal_factor": "Decrease in valuable content", "KPIs_impacted": ["User

Growth Rate", "User Engagement Rate", "Advertising Revenue"], "comes_from": ["Loss of influential Redditors"]}, {"causal_factor": "Reduced advertiser spending", "KPIs_impacted": ["Advertising Revenue", "Profit"], "comes_from": ["Decrease in valuable content"]}]}, {"causal_chain": [{"causal_factor": "Inability to detect and address harmful content", "KPIs_impacted": ["User Base Size", "User Engagement Rate", "Churn Rate"], "comes_from": []}, {"causal_factor": "Perception of offensive or inappropriate content", "KPIs_impacted": ["User Base Size", "User Retention Rate", "User Engagement Rate"], "comes_from": ["Inability to detect and address harmful content"]}, {"causal_factor": "Decline in user activity and engagement", "KPIs_impacted": ["User Growth Rate", "Daily Active Users", "Advertising Revenue"], "comes_from": ["Perception of offensive or inappropriate content"]}, {"causal_factor": "Reduced attractiveness of advertising services", "KPIs_impacted": ["Advertising Revenue", "Revenue", "Net Profit Margin"], "comes_from": ["Decline in user activity and engagement"]}]}]. Use the following list of KPIs and select the best fitting KPIs from this list as much as possible (use the same name), however if a causal_factor does not affect a KPI and only leads to another causal_factor, leave the KPIs_impacted empty: [Monthly Active Users, Daily Active Users, User Growth Rate, User Retention Rate, App Downloads, User Base Size, Churn Rate, Sessions per User, Referral Traffic, Bounce Rate, User Engagement Rate, Revenue, Advertising Revenue, Average Revenue Per User, Profit, Monetization Rate, Operating Costs, Net Profit Margin, Operating Margin, Return on Advertising Spend, Return on Investment, Customer Acquisition Cost, Customer Lifetime Value, Net Promoter Score, Customer Satisfaction Score, Impressions, Click-Through Rate, Conversion Rate, Stock Price, Earnings Per Share, Financial Performance, Debt Issuance, Debt-To-Equity Ratio, Working capital, Liquidity, Return on Assets, Effective Tax Rate, Net Operating Loss, Employee Turnover Rate, Employee Retention Rate, Compliance Costs, Investor Confidence, Market Liquidity, United States Market Share, International Market Share, Product Development Costs, Market Capitalization, Revenue Growth Rate, Customer Trust, Trading Volume, Cash Flow, Litigation Costs, International Revenue]. Remember, you MUST ONLY use KPIs from the given list'},

{"role": "user", "content": paragraph}

]

Appendix 2 - KPI List

1. Monthly Active Users
2. Daily Active Users
3. User Growth Rate
4. User Retention Rate
5. App Downloads
6. User Base Size
7. Churn Rate
8. Sessions per User
9. Referral Traffic
10. Bounce Rate
11. User Engagement Rate
12. Revenue
13. Advertising Revenue
14. Average Revenue Per User
15. Profit
16. Monetization Rate
17. Operating Costs
18. Net Profit Margin
19. Operating Margin
20. Return on Advertising Spend
21. Return on Investment
22. Customer Acquisition Cost
23. Customer Lifetime Value
24. Net Promoter Score
25. Customer Satisfaction Score
26. Impressions
27. Click-Through Rate
28. Conversion Rate
29. Stock Price
30. Earnings Per Share
31. Financial Performance
32. Debt Issuance
33. Debt-To-Equity Ratio
34. Working capital
35. Liquidity
36. Return on Assets
37. Effective Tax Rate
38. Net Operating Loss
39. Employee Turnover Rate
40. Employee Retention Rate
41. Compliance Costs
42. Investor Confidence
43. Market Liquidity
44. United States Market Share
45. International Market Share
46. Product Development Costs
47. Market Capitalization
48. Revenue Growth Rate
49. Customer Trust
50. Trading Volume
51. Cash Flow
52. Litigation Costs
53. International Revenue

Appendix 3 - Prompt for Classification of Causal Factors

```
messages=[
	{"role": "system", "content": 'You will receive a causal factor extracted from an SEC financial report. Your task is to classify it into one of these categories. Respond ONLY with the number (1-10) that best fits, nothing else:\n\n1. Technology & Infrastructure Issues: Platform reliability, cybersecurity, system outages, or scalability challenges.\n2. Financial & Economic Risks: Revenue fluctuations, cost pressures, inflation, or macroeconomic instability.\n3. Regulatory & Legal Pressures: Changing laws, government investigations, compliance burdens, or lawsuits.\n4. User Trust & Engagement Risks: Loss of user trust, privacy concerns, negative public perception, or declining user activity.\n5. Competitive Pressure: Threats from new entrants, pricing wars, innovation gaps, or shrinking market share.\n6. Operational & Business Execution Risks: Internal mismanagement, poor coordination, supply chain issues, or missed deliverables.\n7. Strategic planning and Leadership Risks: High-level missteps such as failed expansions, leadership turnover, or poor long-term vision that misalign the company with market needs.\n8. Global & Political Risks: Trade disputes, geopolitical instability, or region-specific regulatory uncertainty.\n9. Product Quality & Reliability Issues: Defects, underperformance, lack of innovation, or technical shortcomings affecting user satisfaction.\n10. Other. ONLY return the category number, nothing else.'},
	{"role": "user", "content": factor}
]
```