

**IFC-ECCBSO-Bank of Spain Workshop on "New insights from financial statements"**

**17 October 2024**

Quantifying material risks from textual disclosures in  
financial statements using LLMs and LLM agents – A  
case study on physical risks<sup>1</sup>

Taejin Park,  
Bank for International Settlements

---

<sup>1</sup> This contribution was prepared for the workshop. The views expressed are those of the authors and do not necessarily reflect the views of the European Committee of Central Balance Sheet Data Offices (ECCBSO), the Bank of Spain, the BIS, the IFC or the other central banks and institutions represented at the event.

# Quantifying Material Risks from Textual Disclosures in Financial Statements using LLM Agents – A Case Study on Physical Risks

Taejin Park<sup>1</sup>

## Abstract

This paper highlights the novel application of large language models (LLMs) in context-aware materiality assessment of risk disclosures, using physical risks as an example. As physical risks like natural disasters increasingly affect financial stability, accurately capturing and evaluating these risks in corporate disclosures is critical. Traditional NLP methods based on keyword analysis struggle to detect risks embedded in complex or indirect language. In contrast, LLMs, through their ability to understand context and semantics, offer significant improvements in extracting and evaluating disclosures. The key innovation explored in this study is the LLM agent's capacity to assess materiality by integrating company-specific financial, operational, and geographic data, providing a more accurate and tailored evaluation of risks. This approach enhances the ability to differentiate between material and non-material disclosures, aiding financial supervisors and investors in more effective decision-making.

Keywords: Physical risks, risk assessment, large language models (LLM), Natural language processing (NLP), language agent, corporate disclosures

JEL classification: G32, C45, Q54

<sup>1</sup> Bank for International Settlements (taejin.park@bis.org). The views expressed here are those of the author only and not necessarily those of the Bank for International Settlements.

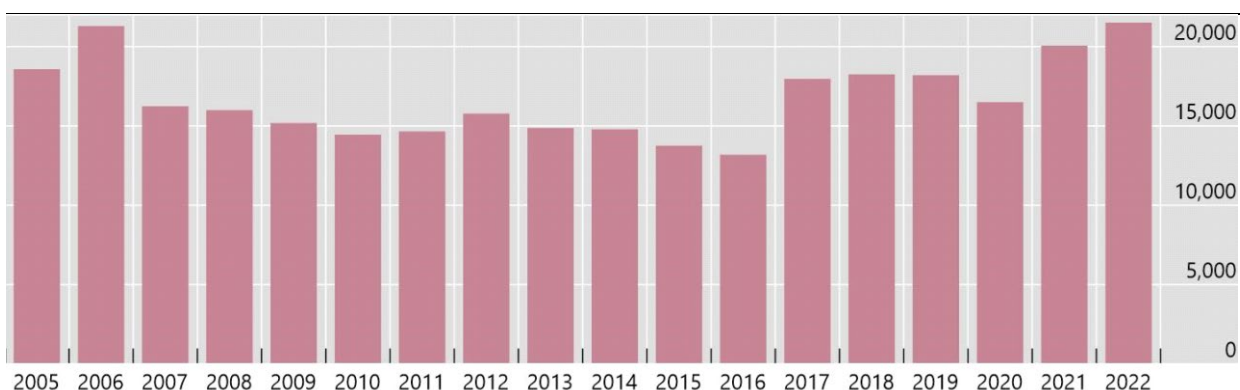
## 1. Introduction

The timely and reliable capture of physical risks is crucial for both investors and financial supervisors. Physical risks, such as natural disasters, pose significant financial stability risks due to their sudden and potentially severe impacts on companies. Unlike transition risks, which develop gradually—such as policy changes or shifts in consumer preferences—physical risks can occur unexpectedly and cause substantial damage.

An increasing number of companies are disclosing their exposure to physical risks (Graph 1). However, extracting relevant information from these disclosures presents challenges. Unlike transition risks, which can often be quantified through metrics such as carbon emissions or energy mix, physical risk disclosures are typically qualitative. As a result, if these disclosures are not thoroughly analyzed, important risks may be overlooked. Furthermore, assessing materiality requires subjective judgment, as readers must weigh short-term versus long-term impacts and relate the risks with various aspects of a company's operation. Given the inherent uncertainty of natural disasters, even well-disclosed risks may be underestimated depending on the risk tolerance of the reader.

Frequency of physical risk-related keywords in SEC 10-K filings

Graph 1



Included keywords: flood, wildfire, hurricane, Flood, typhoon, Hurricane, drought, Wildfire, heat wave, extreme heat, Typhoon, extreme cold, Drought, Extreme heat, heatwave, FLOOD, WildFire, WILDFIRE, HeatWave, HURRICANE, cold wave, Extreme cold, coldwave, Heat Wave, TYPHOON, Heat wave, Extreme Heat, Heatwave, Extreme Cold, DROUGHT and HEAT WAVE.

Sources: S&P; author's calculations.

Recent developments in natural language processing (NLP) and machine learning, especially large language models (LLMs), offer significant potential for efficiently scaling up the capture of qualitative risk exposures from disclosures (BIS, 2024). Despite this progress, challenges remain. Accurately capturing information relevant to a specific user's needs is particularly difficult. The process involves multiple steps: first, capturing the risk information using NLP; second, analyzing the sentiment of that risk; and finally, contextualizing the risk sentiment to ensure it is relevant to the reader's specific concerns. Proper contextualization, like human expert's judgment, is critical for effective risk assessment (Graph 2).

This paper explores how various NLP techniques can be applied to analyze the disclosure of physical risks and presents real-world examples to illustrate their practical applications.



Source: Author's elaboration.

## 2. Extracting Relevant Information

### 1. Traditional keyword-based approaches

Traditional techniques for identifying keywords or topics from large volumes of text primarily rely on word frequency analysis. Methods like Term Frequency-Inverse Document Frequency (TF-IDF; Jones, 1972) and Bag of Words (Harris, 1954) rank words based on how often they appear within a document or across a collection of texts. More advanced techniques, such as Latent Dirichlet Allocation (LDA; Blei et al, 2003) and Latent Semantic Analysis (LSA; Deerwester et al, 1990), group words into topics by identifying patterns in their co-occurrence. Approaches like TextRank (Mihalcea and Tarau, 2004) and basic word frequency counting also emphasize the repetition of words or phrases. Overall, these methods focus on word frequency to extract key terms or themes from text.

Such traditional keyword extraction techniques offer the advantage of simplicity, transparency and computational efficiency, making them accessible for basic text analysis tasks. They are effective for identifying frequently occurring terms and clustering related words into topics, which is useful for quick insights or when processing large text corpora.

However, these methods have notable limitations. They often fail to capture the semantic meaning of words, as they treat words in isolation or based solely on their co-occurrence patterns. Additionally, they struggle with handling synonyms, polysemy (words with multiple meanings), and contextual nuances. Another major limitation is the difficulty in pre-defining an exhaustive list of keywords related to complex topics where descriptions can vary widely. Moreover, risks may be described indirectly, without mentioning any of the expected keywords, making it easy for these methods to overlook critical information (Box 1). As a result, they may miss deeper relationships within the text, especially in more complex language structures or when a deeper understanding of context is required. This restricts their accuracy and limits their ability to provide comprehensive insights in advanced text analysis tasks.

### Examples of disclosures that do not include typical keywords related physical risks

- “We anticipate a modest decrease in harvest volumes versus the prior year as logging conditions normalize following a period of relatively **dry weather**.” (Rayonier Inc., 8-K, 31-01-2024)
- “Because substantially all of our products are manufactured and assembled in and distributed from a few locations, our operations could be interrupted by events beyond our control, including: ... **water damage** or **water shortage** ...” (Acushnet Holdings Corp., 10-K, 01-03-2023)
- “Both parks experienced unusually **inclement weather** during that quarter.” (Cedar Fair, L.P., 10-K, 16-02-2024)
- “FY23 Saleable production levels are lower than prior year due to the impacts from these **poor geotechnical conditions** and the impacts from the rock intrusion in the September quarter, that slowed production rates and impacted yield.” (Coronado Global Resources Inc., 8-K, 22-01-2024)
- “...**warmer weather in the heating season** and **cooler weather in the cooling season** will reduce our customers’ energy demand and our retail operating revenues.” (Avista Corporation, 10-K, 2024-02-21)

## 2. Recent context-based approaches

Recent advances in NLP, particularly in semantic search using embeddings, have significantly improved the ability to overcome the limitations of traditional keyword-based methods. Embeddings, such as those generated by models like Word2Vec (Mikolov et al, 2013), GloVe (Pennington et al, 2014), and more recently, transformer-based models like BERT (Devlin et al, 2019) and GPT (Brown et al, 2020), represent words and phrases as continuous vectors in a high-dimensional space (Box 2). This allows the models to capture the semantic relationships between words, even if they do not explicitly share the same surface form. As a result, semantic search can identify relevant information based on meaning, rather than relying solely on exact word matches. For example, in the context of physical risks, embeddings can recognize descriptions of risks even when they are phrased differently or do not include predefined keywords. This shift from word frequency to meaning-based search allows NLP systems to better handle synonyms, polysemy, and context, leading to more accurate and comprehensive extraction of relevant information from large and complex texts.

### Example of OpenAI’s ‘text-embedding-3-small’ model: translating text into a 1,536-dimension vector

“Prolonged drought conditions increase the intensity of competition for fresh water.” (Energy Transfer LP, 10-K, 2023)

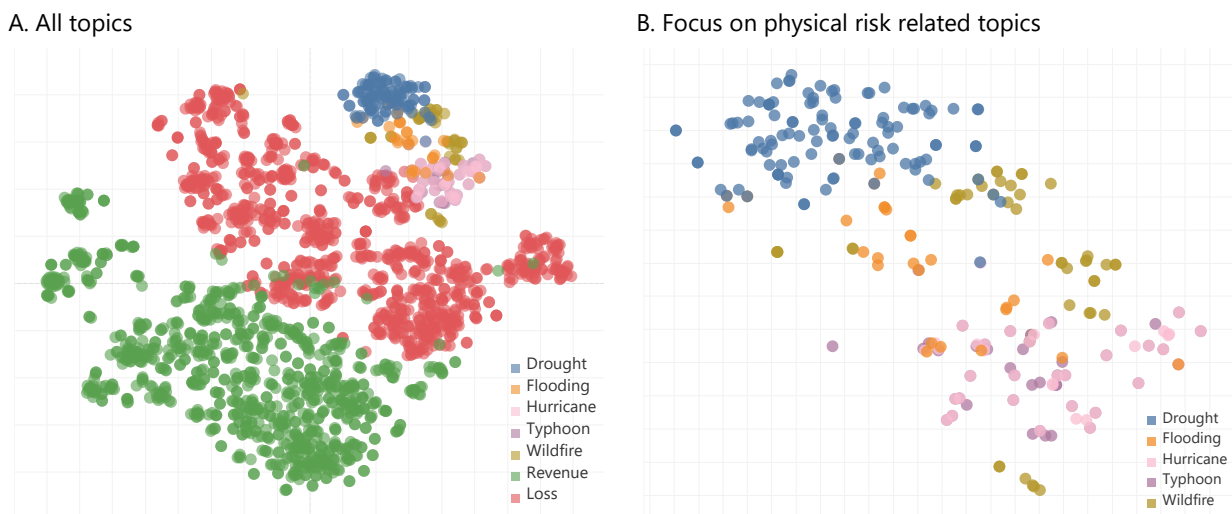


[0.03909442201256752, 0.025339428335428238, ..., 0.03553372994065285, 0.012559967115521431]

To illustrate the use of embeddings, I extracted 2,272 sentences from SEC 10-K and 8-K filings over the past year that are semantically close to physical risk-related keywords such as “drought,” “flooding,” “hurricane,” “typhoon,” and “wildfire.”

Additionally, I included sentences related to general financial topics, such as "revenue" and "loss," to compare them with the physical risk-related topics. These sentences were then translated into 1,536-dimensional vectors, representing their semantic meaning based on OpenAI's 'text-embedding-3-small' model. However, because this high-dimensional space is difficult to interpret visually, I reduced the dimensionality to two using the t-SNE (t-Distributed Stochastic Neighbor Embedding; van der Maaten and Hinton, 2008)<sup>2</sup> method for visualization purposes. After applying t-SNE, the embeddings were plotted in 2D space (Graph 2), enabling a clearer visual comparison between physical risk-related topics and general financial topics, as well as distinctions among the physical risk-related topics.

2D visualization of embeddings for physical risks and financial topics using t-SNE Graph 2



The embeddings are generated using OpenAI's 'text-embedding-3-small' model.

Sources: OpenAI; ProntoNLP; S&P; author's calculations.

Graph 2.A reveals several interesting patterns in the sentence embeddings. Clear clusters emerge, with distinct separation by topic, highlighting the effectiveness of the embeddings in capturing semantic relationships. Notably, the sentences related to physical risks form a well-defined cluster, distinct from the general financial topics like "revenue" and "loss." This suggests that the embeddings effectively distinguish physical risk-related content from broader financial discussions, even after significant dimensionality reduction. Additionally, the cluster of physical risk-related sentences is positioned close to the cluster of "loss"-related sentences, which aligns with expectations, as both topics likely convey negative impacts on a company's performance. This proximity suggests a semantic connection, where physical risks and losses are understood to be closely related in their potential consequences for businesses.

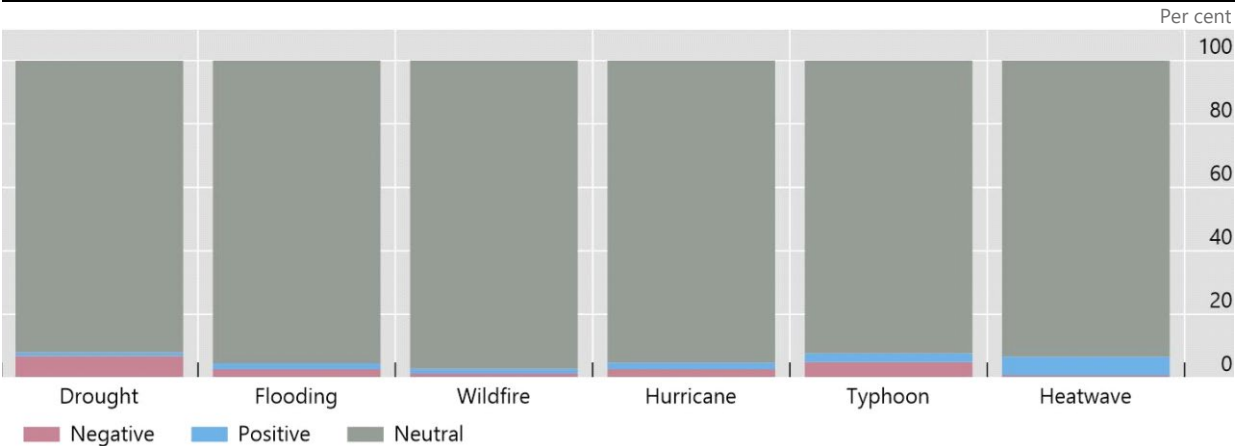
<sup>2</sup> t-SNE is a non-linear dimensionality reduction technique commonly used to visualize high-dimensional data. It works by preserving the local structure of the data, meaning that points close to each other in the high-dimensional space remain close in the lower-dimensional space, making it well-suited for visualizing clusters or relationships in embeddings.

When zooming in on only the physical risk-related sentences, distinct clusters emerge corresponding to different types of events (Graph 2.B). Notably, the embeddings for "hurricane" and "typhoon" show a significant overlap, despite the differences in their surface forms. This highlights the strength of the embeddings in capturing semantic similarity, as these two terms, though different in wording, represent nearly identical meteorological phenomena. Furthermore, their cluster is positioned opposite the "drought" cluster, which also makes sense. While hurricanes and typhoons involve excessive water, droughts represent water scarcity, placing them at opposite extremes. This clear separation demonstrates the model's ability to distinguish between different types of physical risks while effectively grouping semantically similar events, recognizing and organizing concepts based on meaning rather than just word forms.

### 3. Assessing Risk Sentiment

Once information related to physical risks is captured, it is important to assess whether the disclosure is positive, negative, or neutral for the company, as not all disclosures have a negative impact. Some may be neutral descriptions, or even positive if accompanied by negations. In fact, my analysis of 6,000 randomly selected physical risk-related sentences across six different event types from SEC 8-K and 10-K filings over the past year revealed that the vast majority of the captured information was classified as "neutral" according to sentiment analysis by ProntoNLP, an NLP service provider (Graph 3). This underscores the need for sentiment analysis following the capture of physical risk-related text disclosures.

Sentiment analysis of physical risk disclosures Graph 3



The analysis is based on a random selection of 1,000 sentences per category related to physical risk disclosures.

Sources: ProntoNLP; S&P; author's calculations.

Traditional NLP approaches for sentiment classification often rely on counting positive or negative words in a text. While simple, this method tends to miss a lot of nuance, as it fails to account for context, negations, or the overall structure of the sentence. More advanced approaches have shifted toward supervised machine learning, where sentiment classification occurs in the embedding space, allowing

models to capture contextual sentiments. These methods, however, typically require large amounts of labelled data for training, making human annotation important for tasks like fine-tuning. Recent generative LLMs, such as those equipped with zero-shot or few-shot learning capacity, have reduced the need for extensive fine-tuning. These models can classify sentiment with little to no task-specific training, making sentiment analysis more efficient and adaptable, particularly in cases where labelled data is limited.

## 4. Assessing Materiality

Even if risk sentiment is accurately captured, not all negative sentiments are necessarily material to users. The materiality of a given risk is highly subjective and depends on the specific priorities and needs of the information users. Box 3 illustrates examples of disclosures where understanding the broader context is essential. These disclosures cannot be fully understood in isolation, as their significance often hinges on the circumstances surrounding the company, industry, or location. Therefore, assessing the materiality of information relevant to a particular user is critical. However, determining materiality is inherently subjective, varying based on the reader's perspective, objectives, and risk tolerance.

Box 3

### Examples of disclosures that require understanding of the context

- Approximately \$0.01 per diluted share of net expense due to fire related costs.
- As more fully discussed in Note 3, in December 2021, the Company wrote off approximately \$8.7 million of inventory which is held at the Solar Park contract manufacturer's site in Korea and is not expected to be recoverable.
- Current accident year losses in 2022 included \$69.2 million of net losses and loss adjustment expenses attributed to Hurricane Ian and the Russia-Ukraine conflict.
- In recent years, California and Utah have experienced persistent and severe drought conditions.
- We have experienced short-term impacts on our sales due to closures from weather events in recent years, including Hurricane Ian in Florida in 2022.
- The long-term trend is that wildfires are increasing resulting from drought conditions.

Understanding materiality requires a high level of language proficiency and the ability to grasp the overall context, much like a human expert. Recent technologies in LLMs have made this more feasible. Developing LLM agents, especially those specialized in identifying the materiality of risks, could offer a promising solution. An LLM agent is defined by its ability to access and integrate external knowledge provided by users, allowing it to broadly capture the context and interpret information in a human-like manner.

However, just as with human experts, fully understanding the materiality of risks requires extensive knowledge about the specific company. Factors like industry,



location, financials, and business description all play a crucial role in determining materiality. A comprehensive assessment that integrates this information is essential to making informed judgments about the significance of disclosed risks.

To test the capabilities of LLM agents in such tasks, I conducted an analysis using the sample sentences from 8-K and 10-K filings listed in Box 3. In addition to the disclosed text, I provided financial and contextual information to support the LLM agent’s assessment of materiality. The additional data included disclosure date, total assets, total revenue, EBITDA, EBITDA margin, earnings-per-share, price-to-book value, return on assets, net income, business description, industry classification, and location. With this context, the LLM agent was then tasked with evaluating the materiality of each disclosure, as illustrated in Box 4.

Box 4

Example of prompt given to LLM Agent for assessing materiality of physical risk disclosure

**Instruction:** You are a risk analyst assessing firm disclosures for material risks arising from physical climate change risks. When I give you an extract from a firm's disclosure, please assess whether the stated risk is material for the company and investors. Materiality is the major events that shareholders should immediately know about for timely decisions. When you make the assessment, consider all available factors, not only the disclosure in isolation, but also all the relevant information available, such as the nature of the company, business areas, financials, location, etc. The assessment is based on current or future implications.

**Risk statement:** In recent years, California and Utah have experienced persistent and severe drought conditions.

**Company background:** [Company Name] operates as an independent upstream energy company in the western United States. It operates through Exploration and Production (E&P), and Well Servicing and Abandonment segments. The E&P segment engages in the development and production of onshore, low geologic risk, and long-lived conventional oil and gas reserves primarily located in California and Utah. The Well Servicing and Abandonment segment provides wellsite services in California to oil and natural gas production companies with a focus on well servicing, well abandonment services, and water logistics. The company was founded in 1909 and is headquartered in Dallas, Texas. The company is located in Dallas, TX and operates in the Energy; Energy; Oil, Gas and Consumable Fuels; Oil and Gas Exploration and Production; Oil and Gas Production industry. As of 2023-12-31, it had total assets of 1613.925 million USD and total revenue of 222.084 million USD for the quarter. The net income for the quarter was 37.4 million USD. The EBITDA for the quarter was 401.104 million USD with an EBITDA margin of 38.003%. The Basic EPS was 3.186, the price-to-book value was 0.751229, and ROA was 9.889%.

Sources: S&P; author’s elaboration.

Table 1 summarizes the materiality assessments of the LLM agent, which demonstrates a capability to assess the materiality of physical risk disclosures within a given context. It effectively evaluates the implications of risks or damages by considering the company's financial position, the nature of its business, operational processes, and geographic locations.

The agent is able to differentiate between material and non-material impacts depending on the company's size and financial health. For instance, it assessed a \$69.2 million damage as non-material for one company, while determining that an \$8.7 million loss was material for another, reflecting their differing financial circumstances.

Additionally, the LLM agent showed an ability to connect events to a company's operational locations, identifying if risks are more severe based on whether the company operates in a primary or diversified location. It also linked specific risks to business processes. For example, it considered drought as material for companies where water is essential for drilling or production.

The agent also accounts for both short-term and long-term impacts on the business. A generic statement like "The long-term trend is that wildfires are increasing due to drought conditions" was assessed as material for a winemaker operating in fire-prone areas, demonstrating the agent's ability to recognize the significant implications of ongoing risks given the specific nature of business (see Annex for the full assessment).

Summary of LLM agent's materiality assessments for physical risk disclosures

Table 1

Disclosure	Materiality	Main reason
"Approximately \$0.01 per diluted share of net expense due to fire related costs."	No	<ul style="list-style-type: none"> <li>• <b>Strong financial position</b> (eg EPS \$1.43)</li> <li>• <b>Isolated incident</b></li> </ul>
"As more fully discussed in Note 3, in December 2021, the Company wrote off approximately \$8.7 million of inventory which is held at the Solar Park contract manufacturer's site in Korea and is not expected to be recoverable."	Yes	<ul style="list-style-type: none"> <li>• The <b>relatively large size of the write-off</b> 12% of total assets and 35% of quarterly revenue</li> <li>• Weak financial position</li> <li>• <b>Operational inefficiencies and supply chain management issues</b></li> </ul>
"Current accident year losses in 2022 included \$69.2 million of net losses and loss adjustment expenses attributed to Hurricane Ian and the Russia-Ukraine conflict."	No	<ul style="list-style-type: none"> <li>• This loss accounts for approximately <b>3.3% of the company's quarterly net income</b></li> <li>• Isolated incident</li> </ul>
"In recent years, California and Utah have experienced persistent and severe drought conditions."	Yes	<ul style="list-style-type: none"> <li>• The company's <b>primary operation locations.</b></li> <li>• <b>The need for water for drilling and production processes</b></li> </ul>
"We have experienced short-term impacts on our sales due to closures from weather events in recent years, including Hurricane Ian in Florida in 2022."	No	<ul style="list-style-type: none"> <li>• Short-term sales impacts</li> <li>• <b>Diversified operations and geographical spread</b></li> </ul>
"The long-term trend is that wildfires are increasing resulting from drought conditions."	Yes	<ul style="list-style-type: none"> <li>• The company's <b>reliance on vineyards in fire-prone regions</b></li> <li>• <b>Financial and operational vulnerability</b></li> </ul>

Sources: S&P; OpenAI; author's calculations.

## 5. Conclusion

This paper outlines how recent LLM technology can present a valuable opportunity to streamline the analysis of financial disclosures, from identifying physical risk-related events to sentiment classification and materiality assessment. The ability of LLM agents to capture context and assess the materiality of risks can greatly enhance the detection of realized or potential risks, including those linked to climate change.

While assessing materiality is inherently subjective—given that the accounting definition of materiality is loosely defined and one must consider the unique needs of information users—customizing LLM agents to address specific requirements or

interests is essential. This ensures that the assessment is aligned with the particular concerns of stakeholders.

Ultimately, the final judgment on materiality remains with human experts. However, LLM technology can significantly aid this process by providing an initial evaluation, including a contextual summary tailored to the needs of the user. This enables professionals to focus on areas of particular interest, improving both the efficiency and accuracy of the risk assessment process.

## Reference

BIS (2024): *Project Gaia: Enabling climate risk analysis using generative AI*, March.

Blei, DM, AY Ng and MI Jordan (2003): "Latent Dirichlet Allocation", *Journal of Machine Learning Research*, vol 3, pp 993–1022.

Brown, T, B Mann, N Ryder, M Subbiah, J Kaplan, P Dhariwal, A Neelakantan, P Shyam, G Sastry, A Askell, S Agarwal, A Herbert-Voss, G Krueger, T Henighan, R Child, A Ramesh, D Ziegler, J Wu, C Winter, C Hesse, M Chen, E Sigler, M Litwin, S Gray, B Chess, J Clark, C Berner, S McCandlish, A Radford, I Sutskever and D Amodei (2020): "Language models are few-shot learners", *NeurIPS*.

Deerwester, S, ST Dumais, GW Furnas, TK Landauer and R Harshman (1990): "Indexing by latent semantic analysis", *Journal of the American Society for Information Science*, vol 41, no 6, pp 391–407.

Devlin, J, MW Chang, K Lee and K Toutanova (2019): "BERT: Pre-training of deep bidirectional transformers for language understanding", *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics (NAACL)*, June, pp 4171–4186.

Harris, Z (1954): "Distributional structure", *Word*, vol 10, no 2-3, pp 146–162.

Jones, K Sparck (1972): "A statistical interpretation of term specificity and its application in retrieval", *Journal of Documentation*, vol 28, no 1, pp 11–21.

Mihalcea, R and P Tarau (2004): "TextRank: Bringing order into texts", *Proceedings of the 2004 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, July, pp 404–411.

Mikolov, T, C Corrado, K Chen and J Dean (2013): "Efficient estimation of word representations in vector space", *Proceedings of the International Conference on Learning Representations (ICLR)*, May.

Pennington, J, R Socher and C Manning (2014): "GloVe: Global vectors for word representation", *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, October, pp 1532–1543.

van der Maaten, L and G Hinton (2008): "Visualizing data using t-SNE", *Journal of Machine Learning Research*, vol 9, pp 2579–2605.

## Annex – Materiality Assessment Example by LLM Agent

### Material Risk Assessment: Increasing Wildfires Due to Drought Conditions

The risk statement highlights a **long-term trend of increasing wildfires**, exacerbated by drought conditions, which could affect [Company Name]. To assess whether this poses a **material risk**, we will examine the potential operational, financial, and strategic implications of this risk in the context of the company's performance and business model.

#### 1. Company Background and Operations:

- **Industry:** [Company Name] operates in the **wine and spirits industry**, producing and selling wines under various brands, including [Brand A], [Brand B], and [Brand C], among others. In addition to producing wines, the company owns and operates **hospitality facilities**, providing services like bottling, fulfillment, and storage to other companies.
- **Geographical Presence:** The company operates primarily in the **United States** and **Canada**, regions that have seen increasing wildfire activity, particularly in **California** and other key wine-producing areas. Wine production is highly dependent on climatic conditions, and **wildfires** are a known threat to vineyards, wine quality, and overall operations.
- **Financial Position:** [Company Name] has filed for **Chapter 11 bankruptcy** (as of July 24, 2024), indicating financial distress. The company's **net income of -\$190.229 million** for the quarter and **low EBITDA margin of 3.756%** suggest it is operating with very limited profitability and is highly vulnerable to additional financial shocks. Any significant disruption, like wildfires, could further strain its financial position.

#### 2. Nature of the Risk (Wildfires Due to Drought):

- **Operational Impact:**
  - **Damage to Vineyards:** Wildfires pose a direct threat to **vineyards**, which are highly susceptible to damage from fire, heat, and smoke. Even if vineyards are not directly burned, **smoke taint** can spoil grape quality, leading to reduced production volumes or lower-quality wine.
  - **Disruption to Harvesting and Production:** Wildfires can disrupt vineyard operations by **forcing evacuations**, limiting access to properties, and damaging equipment and infrastructure (e.g., irrigation systems). This can lead to delays or cancellations of **harvests** and interruptions in wine production, directly affecting revenues.
  - **Hospitality Facilities:** The company also operates **hospitality facilities**, which could face closures or reduced visitor numbers during wildfire seasons, particularly if air quality deteriorates or if there are safety concerns in affected areas.
  - **Supply Chain Disruptions:** Wildfires could disrupt the **logistics and supply chains** for bottling and distribution, particularly in California and other regions with key infrastructure.
- **Financial and Operational Flexibility:**
  - **Bankruptcy Proceedings:** The company is already in **Chapter 11 bankruptcy**, meaning it is undergoing reorganization to address its financial challenges. Wildfires and the resulting disruptions could significantly hinder

recovery efforts by causing additional **revenue losses** or increasing operational costs (e.g., for repairing damaged vineyards or investing in smoke-taint mitigation technologies).

- **Limited Financial Resilience:** With a **low EBITDA margin of 3.756%**, the company is operating on thin margins. Any additional operating costs due to wildfires (e.g., damaged crops, reduced yields, or lower wine quality) could further reduce profitability, complicating its ability to stabilize financially.
- **Long-Term Climate Change Implications:** Wildfires, exacerbated by drought conditions, are expected to become more frequent and severe due to **climate change**. For a wine producer like [Company Name], this represents a significant long-term risk to vineyard operations, product quality, and regional tourism, all of which are critical to the company's business model.

### 3. Materiality of the Risk:

- **Direct Financial Impact:**
  - **Revenue and Production:** The company's quarterly revenue of **\$64.651 million** is already under pressure, and any reduction in production due to wildfires could further decrease revenues. **Vineyards are capital-intensive**, and rebuilding or rehabilitating vineyards after a wildfire can take years, making the risk of wildfires particularly costly.
  - **Net Income:** With a **net income of -\$190.229 million**, [Company Name] is already facing significant financial challenges. Even minor disruptions from wildfires could exacerbate these losses, given the thin profit margins and ongoing bankruptcy proceedings.
  - **Asset Impairment:** The company has **total assets of \$626.446 million**, and wildfires could lead to asset impairments if vineyards, equipment, or hospitality facilities are damaged. This could reduce the value of the company's assets further, negatively impacting its already low **price-to-book ratio of 0.23**, indicating the market already perceives significant risk.
- **Operational Disruptions:**
  - **Wildfire Frequency:** The risk statement emphasizes that **wildfires are increasing**, suggesting that operational disruptions are not limited to a single event but represent an ongoing and growing threat. This increased frequency could lead to recurring disruptions in production cycles and long-term damage to grapevines, impacting both short-term revenue and long-term viability.
  - **Wine Quality:** Smoke taint can severely affect the **quality of wine**, leading to potential reputational damage if the company produces wine that fails to meet customer expectations. Lower-quality wine might need to be sold at a discount or discarded altogether, further impacting profitability.

### 4. Long-Term Outlook and Strategic Considerations:

- **Climate Change and Sustainability:** Given that wildfires are likely to increase due to **climate change**, [Company Name] may need to invest in **mitigation strategies**, such as enhanced fire protection for its vineyards, more robust irrigation systems to combat drought, or **smoke-taint mitigation technologies** (e.g., air filtration, protective coverings). However, given the company's financial situation, it may lack the capital required to make such investments.
- **Reputation and Market Position:** [Company Name] may face reputational challenges if it is unable to consistently deliver high-quality wines due to wildfire-

related disruptions. This could lead to **loss of market share** in an already competitive industry where brand loyalty and product quality are critical to success.

- **Long-Term Financial Viability:** The combination of bankruptcy proceedings and increased wildfire risk places significant pressure on the company's long-term viability. The ability to successfully emerge from Chapter 11 will depend on how well the company can mitigate climate risks while stabilizing its finances.

#### 5. Investor Considerations:

- **Material Risk for Investors:** The **increasing frequency of wildfires** due to drought conditions represents a **material risk** for [Company Name]. The company's fragile financial position, combined with its reliance on climate-sensitive assets (vineyards), makes it highly vulnerable to wildfire disruptions. Investors should be concerned about the potential for continued operational and financial deterioration if the company cannot manage wildfire risks effectively.
- **Bankruptcy and Recovery:** The company's ongoing **Chapter 11 bankruptcy** adds another layer of complexity. Any significant operational disruptions caused by wildfires could delay or derail the company's recovery plan, potentially leading to further financial distress or even liquidation if the reorganization is unsuccessful.
- **Climate Risk as a Long-Term Threat:** Investors focused on **ESG** factors may see the increasing wildfire risk as part of a broader **climate risk** concern. [Company Name] may need to prioritize sustainability and resilience strategies in order to remain competitive in an industry increasingly impacted by climate change.

#### Conclusion: Material Risk Assessment

- **Material Risk:** The **long-term trend of increasing wildfires** due to drought conditions poses a **material risk** to [Company Name]. Given the company's reliance on vineyards in fire-prone regions, any damage to vineyards, disruptions in production, or quality issues caused by smoke taint could significantly impact its operations and financial performance.
- **Financial and Operational Vulnerability:** [Company Name] is already under financial strain, with a **negative net income** and **low EBITDA margin**. Wildfires could exacerbate this situation, leading to further losses, operational disruptions, and a delayed recovery from Chapter 11 bankruptcy.
- **Long-Term Considerations:** Wildfires and climate risks will likely continue to challenge the company's long-term viability, requiring strategic investments in mitigation efforts that may be difficult to implement given its financial constraints.

In conclusion, **increasing wildfires** due to drought conditions represent a **material risk** for [Company Name], particularly given its current financial instability and reliance on climate-sensitive assets.





## Quantifying Material Risks from Textual Disclosures in Financial Statements using LLM Agents – A Case Study on Physical Risks

Taejin Park

Workshop on New insights from Financial Statements

Madrid, 17 October 2024

## Introduction

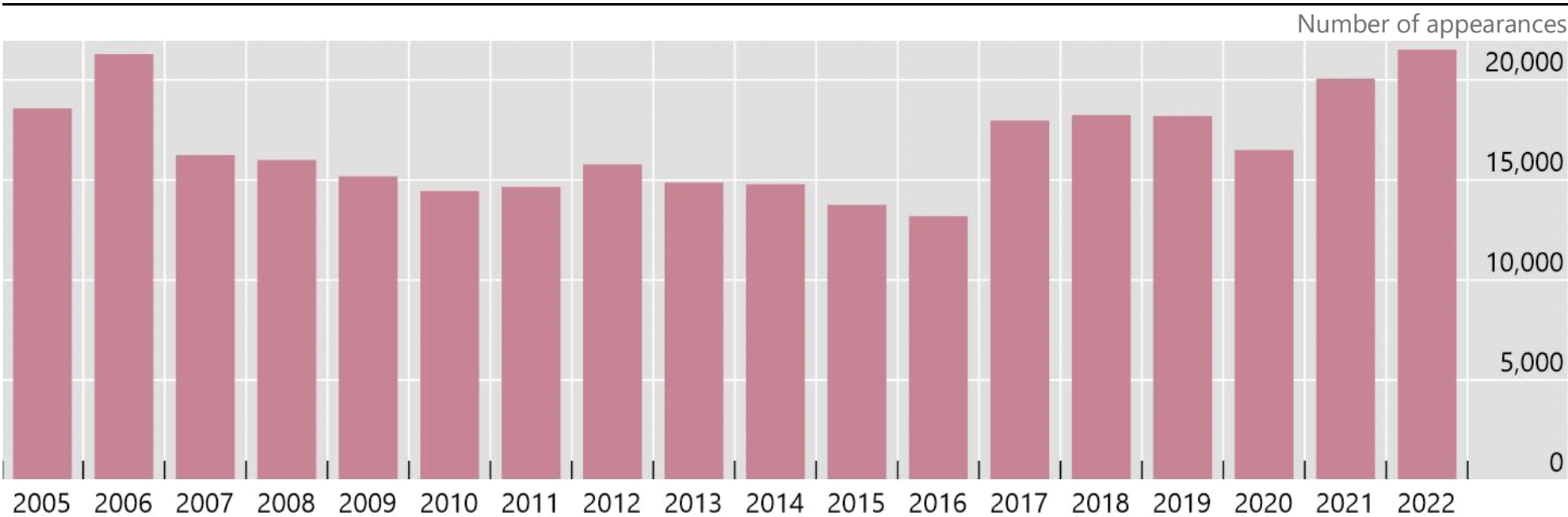
- Data on physical risks largely rely on **disclosures**.
- Gathering and analysing them can be **overwhelming**.
- **Recent LLM technologies** offer scalable solutions.





# Rising Trends in Disclosure

Frequency of physical risk-related keywords in 10-K filings



Included keywords: flood, wildfire, hurricane, Flood, typhoon, Hurricane, drought, Wildfire, heat wave, extreme heat, Typhoon, extreme cold, Drought, Extreme heat, heatwave, FLOOD, WildFire, WILDFIRE, HeatWave, HURRICANE, cold wave, Extreme cold, coldwave, Heat Wave, TYPHOON, Heat wave, Extreme Heat, Heatwave, Extreme Cold, DROUGHT and HEAT WAVE.

Sources: S&P; author's calculations.

# Key Challenges

How to detect info  
embedded in complex or  
indirect language?

Reflecting the **specific**  
**context** and relevance to  
each user.

Extracting **Relevant  
Information**

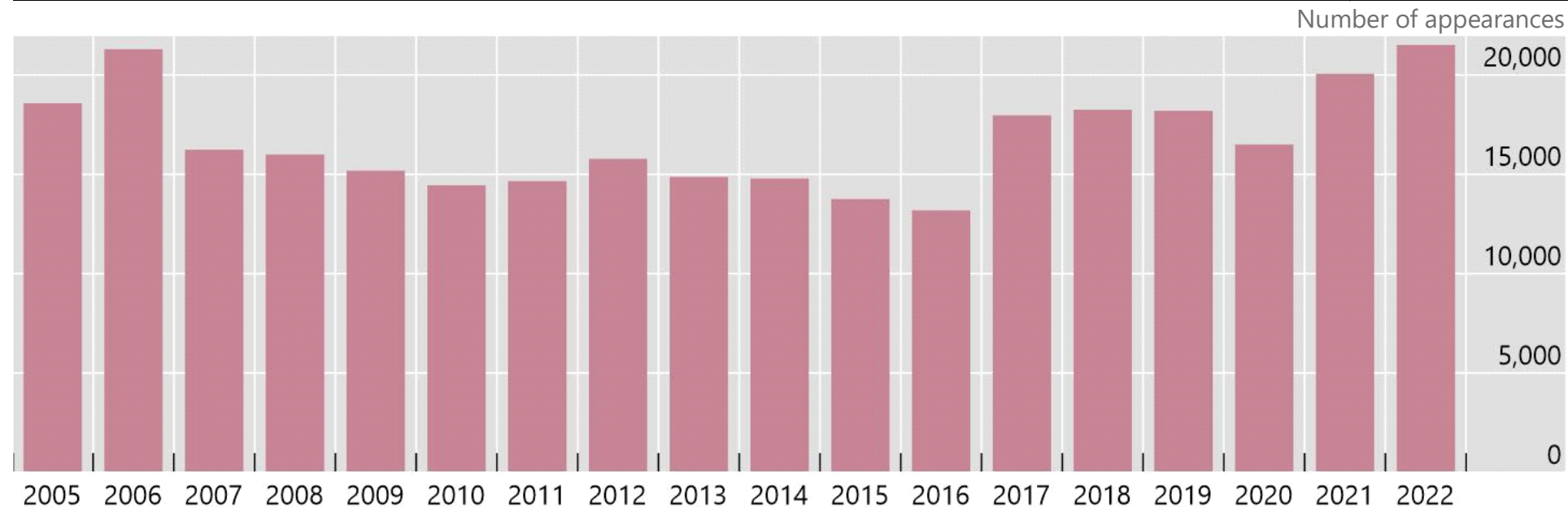
Assessing **Risk  
Sentiment**

Assessing  
**Materiality**

Not all risk disclosures are  
negative.

# Extracting Relevant Information: Traditional Keyword-Based Approach

## Frequency of physical risk-related keywords in 10-K filings



Included keywords: flood, wildfire, hurricane, Flood, typhoon, Hurricane, drought, Wildfire, heat wave, extreme heat, Typhoon, extreme cold, Drought, Extreme heat, heatwave, FLOOD, WildFire, WILDFIRE, HeatWave, HURRICANE, cold wave, Extreme cold, coldwave, Heat Wave, TYPHOON, Heat wave, Extreme Heat, Heatwave, Extreme Cold, DROUGHT and HEAT WAVE.

Sources: S&P; author's calculations.

# Extracting Relevant Information: Traditional Keyword-Based Approach

- Main Challenges:
  - How do you create a truly **exhaustive list** of keywords for risk identification?
    - What about synonyms, polysemy... (eg bank, bond)
  - One can **describe a risk without explicitly mentioning those keywords**.

## Extracting Relevant Information: Traditional Keyword-Based Approach

- Examples of complex language
  - “We anticipate a modest decrease in harvest volumes versus the prior year as logging conditions normalize following a period of relatively **dry weather**.” (Rayonier Inc., 8-K, 31-01-2024)
  - “Because substantially all of our products are manufactured and assembled in and distributed from a few locations, our operations could be interrupted by events beyond our control, including: ... **water damage** or **water shortage** ...” (Acushnet Holdings Corp., 10-K, 01-03-2023)
  - “Both parks experienced unusually **inclement weather** during that quarter.” (Cedar Fair, L.P., 10-K, 16-02-2024)
  - “FY23 Saleable production levels are lower than prior year due to the impacts from these **poor geotechnical conditions** and the impacts from the rock intrusion in the September quarter, that slowed production rates and impacted yield.” (Coronado Global Resources Inc., 8-K, 22-01-2024)
  - “...**warmer weather in the heating season** and **cooler weather in the cooling season** will reduce our customers’ energy demand and our retail operating revenues.” (Avista Corporation, 10-K, 2024-02-21)

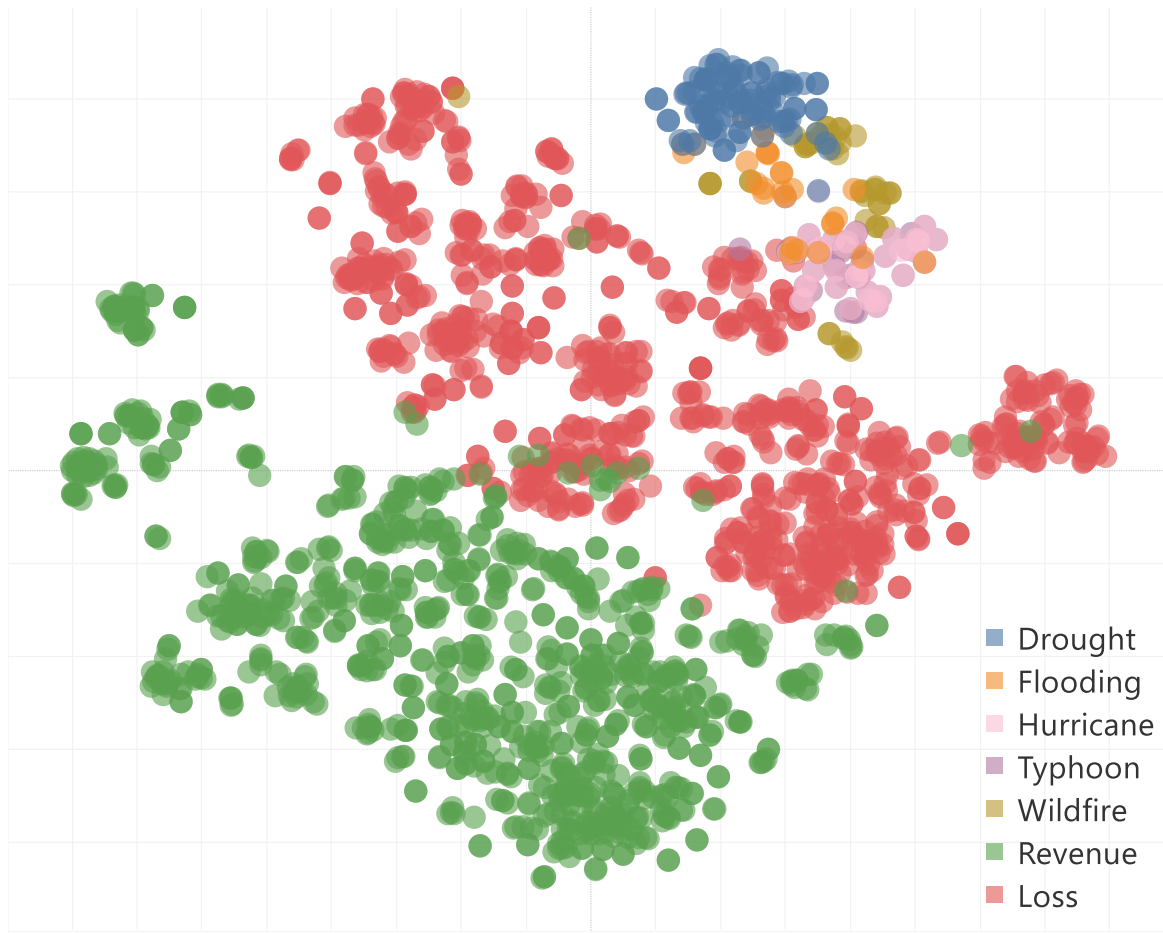


"Prolonged drought conditions increase the intensity of competition for fresh water."

→ 1,536 dimensions

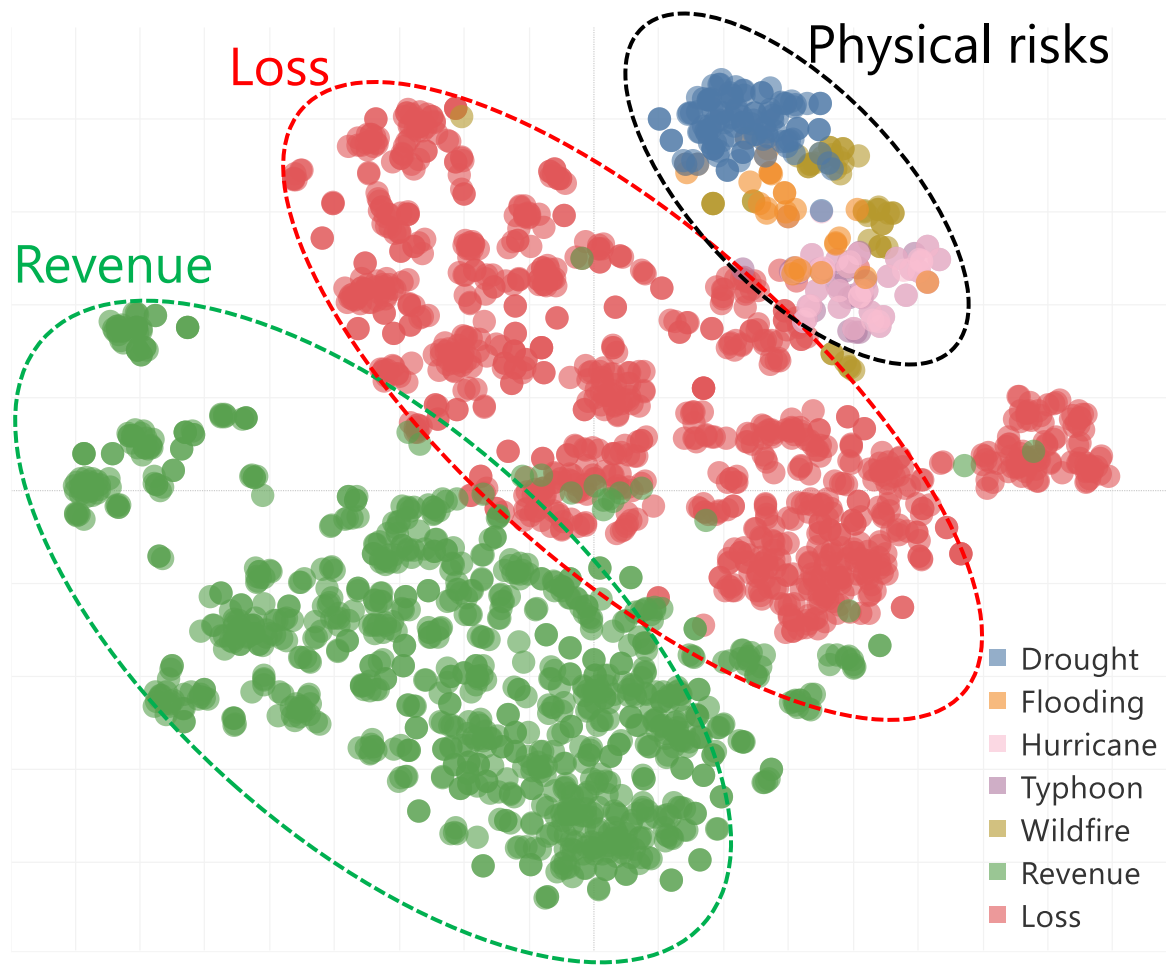
- Enabling semantic search

## Recent NLP approach – Embedding-based approaches



- Physical risks & general financial topics
- Dimension reduction: 1,536  $\rightarrow$  2 using t-SNE

## Recent NLP approach – Embedding-based approaches

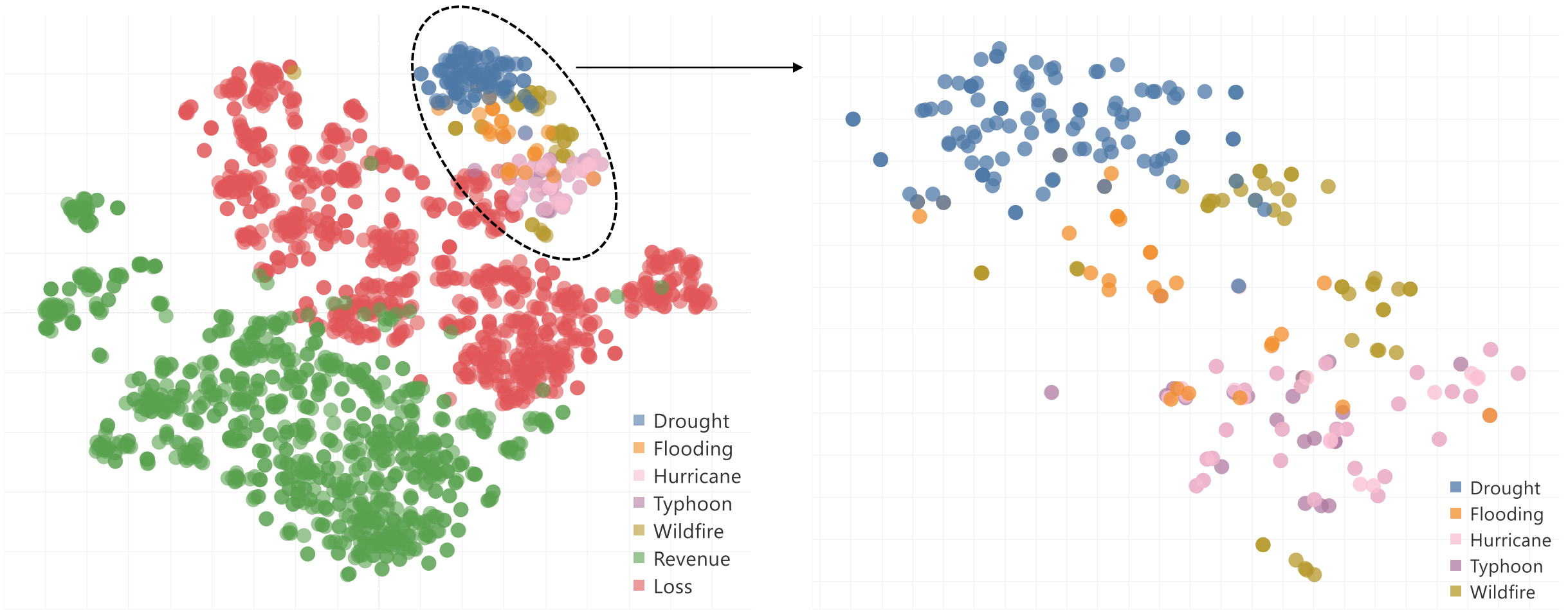


Even with the highly compressed dimensions,

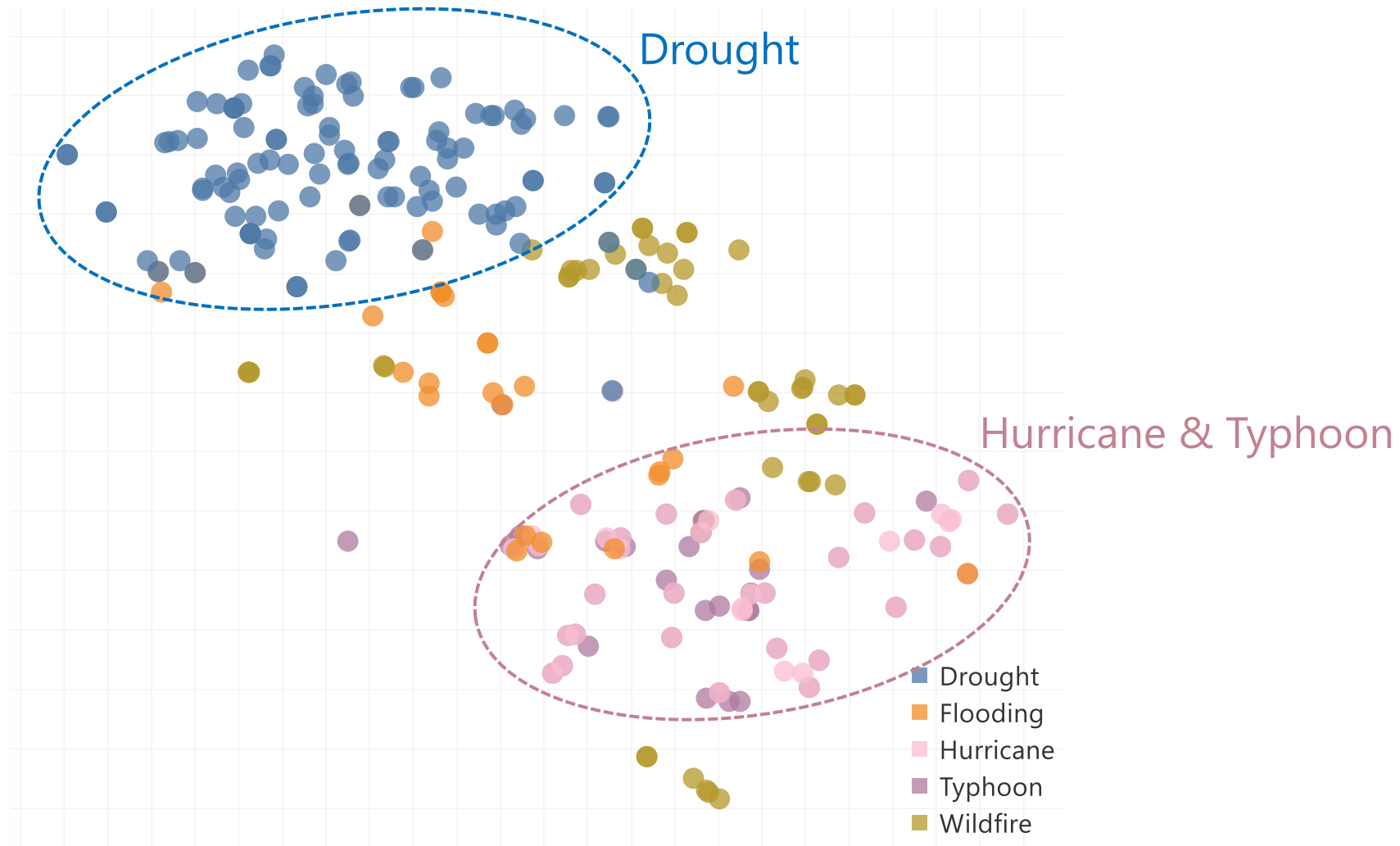
- Clear emergence of clusters by topic
- Physical risk-related cluster
- Physical risks (semantically) close to “loss” cluster



## Recent NLP approach – Embedding-based approaches

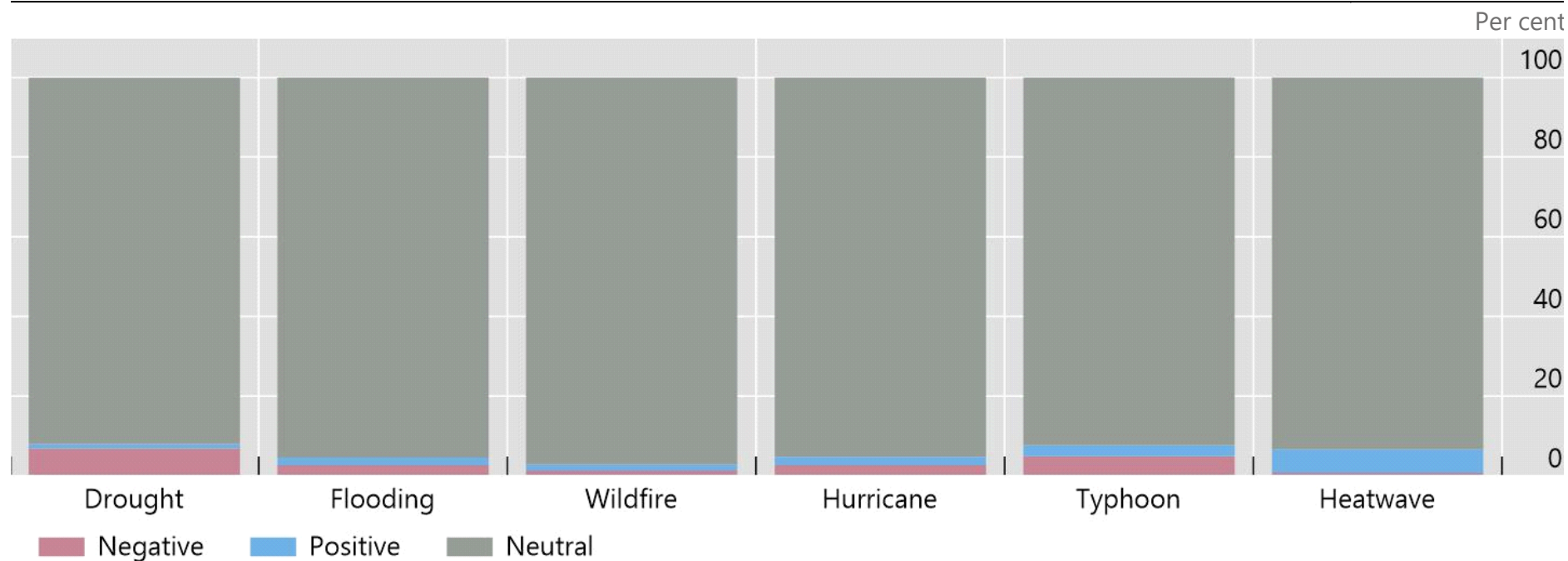


## Recent NLP approach – Embedding-based approaches



# Assessing Risk Sentiment is Essential

## Sentiment analysis of physical risk disclosures



The analysis is based on a random selection of 1,000 sentences per category related to physical risk disclosures.

Sources: S&P; ProntoNLP; author's calculations.

- Fine-tuned LLMs or In-context learning can be used.

## Contextual and Subjective Nature of Materiality Assessment: Examples

- "Approximately **\$0.01 per diluted share** of net expense due to fire related costs."
- "As more fully discussed in Note 3, in December 2021, the Company wrote off approximately **\$8.7 million of inventory** which is held at the Solar Park contract manufacturer's site in Korea and is not expected to be recoverable."
- "Current accident year losses in 2022 included **\$69.2 million of net losses** and loss adjustment expenses attributed to Hurricane Ian and the Russia-Ukraine conflict."
- "In recent years, **California and Utah** have experienced persistent and severe drought conditions."
- "We have experienced **short-term impacts on our sales** due to closures from weather events in recent years, including Hurricane Ian in Florida in 2022."
- "The **long-term trend** is that wildfires are increasing resulting from drought conditions."

# The need for context-aware analysis to assess the materiality

- Disclosure, Disclosure date, Filing type



- Business description
- Industry classification
- Location
- Total assets
- Total revenue
- EBITDA
- EBITDA margin
- Earnings-per-share
- Price-to-book value
- Return on Assets
- Net income, ...



**LLM Agent**  
to assess materiality

```
def generate_description(row):  
    description = (  
        f"{row['Business Description']} "  
        f"The company is located in {row['Location']} and operates in the {row['Industry Classification']} industry. "  
        f"As of {row['Date']}, it had total assets of {row['Total Assets (mn USD)']} million USD and total revenue of {row['Total Revenue (Qtr; mn USD)']} million USD for the quarter. "  
        f"The net income for the quarter was {row['Net Income (Qtr; mn USD)']} million USD."  
        f"The EBITDA for the quarter was {row['EBITDA (Qtr)']} million USD with an EBITDA margin of {row['EBITDA Margin']}%. "  
        f"The Basic EPS was {row['EPS']}, the price-to-book value was {row['Price-to-book value']}, and ROA was {row['ROA']}%. "  
    )  
    return description
```

# Context-aware analysis

Disclosure	Summary of responses by the LLM Agent	
	Materiality	Main reason
"Approximately \$0.01 per diluted share of net expense due to fire related costs."	No	<ul style="list-style-type: none"> <li>Strong financial position (eg EPS \$1.43)</li> <li>Isolated incident</li> </ul>
"As more fully discussed in Note 3, in December 2021, the Company wrote off approximately \$8.7 million of inventory which is held at the Solar Park contract manufacturer's site in Korea and is not expected to be recoverable."	Yes	<ul style="list-style-type: none"> <li>The relatively large size of the write-off 12% of total assets and 35% of quarterly revenue</li> <li>Weak financial position</li> <li>Operational inefficiencies and supply chain management issues</li> </ul>
"Current accident year losses in 2022 included \$69.2 million of net losses and loss adjustment expenses attributed to Hurricane Ian and the Russia-Ukraine conflict."	No	<ul style="list-style-type: none"> <li>This loss accounts for approximately 3.3% of the company's quarterly net income.</li> <li>Isolated incident</li> </ul>
"In recent years, California and Utah have experienced persistent and severe drought conditions."	Yes	<ul style="list-style-type: none"> <li>The company's primary operation locations.</li> <li>The need for water for drilling and production processes.</li> </ul>
"We have experienced short-term impacts on our sales due to closures from weather events in recent years, including Hurricane Ian in Florida in 2022."	No	<ul style="list-style-type: none"> <li>Short-term sales impacts</li> <li>Diversified operations and geographical spread</li> </ul>
"The long-term trend is that wildfires are increasing resulting from drought conditions."	Yes	<ul style="list-style-type: none"> <li>The company's reliance on vineyards in fire-prone regions.</li> <li>Financial and operational vulnerability</li> </ul>

## Conclusion

- **Streamlining Analysis:** LLMs help automate financial disclosure analysis—from identifying physical risks to sentiment and materiality assessment.
- **Context Matters:** LLMs' ability to capture context boosts accuracy in detecting risks, including climate-related ones.
- **Subjectivity in Materiality:** Tailoring LLMs to users' needs is crucial, given the subjective nature of materiality assessments.
- **Augmenting Human Expertise:** LLMs provide initial evaluations, allowing experts to focus on key risks and enhance decision-making.