Determinants and Information Content of Voluntary Artificial Intelligence Disclosure

Yang Cao[1], Miao Liu[2], Jiaping Qiu[3], Ran Zhao[4]

Abstract

We study the determinants and information content in the voluntary disclosure of artificial intelligence (AI) in corporate annual reports. Using a combination of traditional keyword searching based textual analysis and large language models (LLMs) like ChatGPT, we extract and classify AI-related disclosures into revenue generation and cost reduction activities. We find that AI disclosure for both purposes has grown significantly across all industry sectors since 2010 and is highly related to a firm's real activities in AI, measured by AI-related job postings. However, AI disclosure includes additional information than a firm's current real AI activities in predicting a firm's growth, investment, and operational efficiency. Corporate AI risk disclosure predicts a higher firm tail risk. Our results highlight that voluntary AI disclosures not only reflect a firm's present engagement in AI activities but also provide material forward-looking information about the risks and opportunities associated with AI activities, which is useful for external stakeholders to gauge a firm's AI engagement.

Keywords: Artificial intelligence, 10-K disclosure, firm growth, investment

JEL Classification: D22, E22, G14, G32, M41

[1] Carroll School of Management, Boston College. Email: caoyf@bc.edu
[2] Carroll School of Management, Boston College. Email: liueu@bc.edu
[3] DeGroote School of Business, McMaster University. Email: qiu@mcmaster.ca.
[4] Fowler College of Business, San Diego State University. Email: rzhao2@sdsu.edu.

1. Introduction

Artificial Intelligence (AI) has become the most transformative technology in today's business landscape. A recent survey of 240 CEOs and senior executives shows that almost 60% of business leaders consider AI to "fundamentally transform the landscape."[5] Such an emerging, widespread, and rapidly evolving technology introduces significant uncertainty for investors when assessing its impact on firm value. On one hand, AI has the potential to enhance future cash flows by driving revenue growth through the creation of innovative products and more precise customer targeting. Additionally, AI can reduce operational costs by automating routine tasks, improving supply chain management, and optimizing production processes, thereby boosting overall efficiency. However, alongside these benefits, AI adoption also introduces heightened risks. One major source of risk is the evolving regulatory landscape surrounding AI, as governments grapple with ethical concerns, privacy issues, and the potential societal impact of widespread automation. Furthermore, the implementation of AI technology is often accompanied by adoption risks, including the possibility of integration failures and the emergence of new risks, such as cybersecurity vulnerabilities.

These uncertainties make it essential for investors to have access to high-quality information regarding firms' AI initiatives and the associated risks. The current disclosure regulatory framework, however, lacks explicit guidelines or standards for AI reporting, making such disclosures voluntary. The SEC has raised significant concerns about the credibility of AI-related disclosures by companies, warning against the practice of "AI washing," where firms overstate their AI capabilities to appear more advanced than they are.[6] Investors share these concerns, becoming increasingly wary of exaggerated claims, which has contributed to a rise in securities

[5] https://www.kornferry.com/institute/humans-still-wanted-the-future-of-work-in-an-ai-driven-world.

[6] SEC Chair Gary Gensler, in a September 2024 statement, emphasized the need for companies to ensure that their AI-related claims align with their actual capabilities, signaling the regulator's increasing focus on this issue. See https://www.thecorporatecounsel.net/blog/2024/09/sec-chair-addresses-ai-washing-by-public-companies.html.

class action lawsuits.[7] Against this backdrop, there is a dearth of evidence on the determinants and information content of voluntary AI disclosure, presenting substantial challenges for external stakeholders seeking to assess corporate AI initiatives and engagements. The purpose of this paper is to fill in this important gap by providing a systematic analysis of corporate voluntary AI disclosure in their annual reports.

Specifically, we ask three questions regarding corporate voluntary AI disclosure. First, to what extent do publicly traded companies voluntarily disclose AI-related information in annual reports? How does AI disclosure evolve over time and vary across industries? Second, what are the determinants of AI disclosure? Are companies using AI disclosure merely to appease investor interests without substantial involvement in AI activities, or do they reflect a firm's active engagement in AI? Third, what is the information content of AI disclosures? Do they contain forward-looking material information that is associated with a firm's future performance? We believe answering these questions is the first important step in understanding the implications of corporate voluntary AI disclosure.

The voluntary nature of AI disclosure in annual reports and the lack of standardized reporting frameworks create significant measurement challenges. Disclosures are often dispersed across various sections of firms' reports and may lack a consistent format, making it difficult to accurately and consistently measure AI disclosure across firms. To address this challenge, we develop a two-step procedure that combines traditional keyword searching based textual analysis with the latest large language models (LLMs).

[7] AI disclosure-related lawsuit filings are increasing in recent years as reported by the Stanford Law School Securities Class Action Clearinghouse: https://securities.stanford.edu/current-trends.html#collapse1. For instance, a recent case was filed against GitLab, alleging that the company misled investors about its AI capabilities and their expected impact on market demand: https://www.dandodiary.com/wp-content/uploads/sites/893/2024/09/GitLab-complaint-2.pdf.

In the first step, we utilize a traditional keyword-based textual analysis approach to gauge the voluntary AI disclosure in U.S. public firms' 10-K filings (from SEC EDGAR). We started with only "artificial intelligence" as the keyword, ensuring clarity in identifying AI disclosure. Subsequently, we expand the keyword list by manually reviewing 10-K statements that disclose AI, summarizing the related and frequently mentioned words and bigrams. In the validation process, we compare algorithm-identified AI disclosure instances with human reading outcomes in randomly selected 10-K samples. Expansion of the keyword list halts once the accuracy ratio reaches 99% in validation. Finally, we use the finalized keyword list to identify firms with AI disclosures and measure their intensity.

While this method aids in identifying sections of the 10-K that discuss AI, it has two notable limitations. First, traditional textual analysis techniques may overlook crucial contextual information by focusing solely on sentences containing specific keywords. As a result, discussions about AI might be present in surrounding sentences or implied rather than explicitly stated, leading to inaccurate assessments of a firm's AI usage based purely on keyword counts. Second, traditional textual analysis methods struggle to capture the nuanced context within the identified sentences. This limitation makes it challenging to determine the specific nature of AI usage being discussed. For instance, it is difficult to discern whether the mention of AI pertains to revenue-generation activities, such as product development and pricing optimization, or cost-reduction activities, such as inventory management or operational efficiency improvement. As a result, traditional textual analysis techniques are limited in their ability to fully understand and categorize the nature of AI usage discussed.

To overcome these limitations, we incorporate advanced large language models (LLMs), specifically ChatGPT, in the second step of our analysis. For each AI-related keyword identified

in the first step, we extract 400 words before and after the keyword and prompt ChatGPT to assess the relevance of the information to AI usage. ChatGPT then classifies the types of AI usage mentioned into five categories: Product Development (leveraging AI to design or enhance new products), Pricing Optimization (using AI to set or adjust prices dynamically based on market conditions), AI Product Provider (offering AI-based solutions to customers as part of the firm's product portfolio), Inventory Management (utilizing AI to forecast demand and optimize stock levels), and Operational Efficiency (applying AI to streamline and automate internal processes). By integrating ChatGPT's ability to understand and interpret context, we aim to provide a more accurate and detailed analysis of firms' AI-related activities.

Using these new measures of AI disclosure, we first document patterns associated with voluntary AI disclosure in the 10-K statements of U.S. public firms between 2010 and 2023. We observe that only 2.36% of U.S. public firms disclosed AI-related matters in 2010, a figure that surged to 20.02% in 2023. This rapid growth in AI disclosure over the past decade is pervasive across all industry sectors. When breaking down 10-K sections, we find that AI disclosure is concentrated in the business description (section Item 1), risk factors (section Item 1A), and management discussion and analysis (section Item 7). This concentration is not surprising as the business descriptions (section Item 1) typically offer a general overview of how AI relates to corporate operations. Furthermore, risk factors (section Item 1A) and management discussion and analysis (section Item 7) tend to provide information on risks and opportunities associated with AI.

One crucial concern regarding the surge in AI voluntary disclosure is whether it merely signifies a firm's superficial interest in aligning with fashion or genuinely reflects its active involvement in AI endeavors. To gain insights into the motives of AI disclosure, we assess the impact of a firm's AI activities along with other key firm characteristics on the likelihood of AI

disclosure. To gauge a firm's real activities in AI, we use a metric proposed by Babina et al. (2024): the fraction of AI employees, which represents the proportion of employees possessing AI-related skills (identified in their resumes) relative to the total workforce. AI development requires recruiting employees with AI expertise because they possess specialized knowledge and skills essential for tasks such as algorithm design, data analysis, and model development. Therefore, the fraction of AI-related employees within a firm reflects the sources firms commit to AI activities.

Our results reveal that a firm's AI disclosure is significantly related to its AI activities. Specifically, one percentage in the fraction of AI employees results in a 0.6% probability increase in AI disclosure. The results indicate that voluntary AI disclosures are not superficial presentations but reflect firms' actual engagements in AI initiatives. However, apart from the fraction of AI employees, we also find that larger firms, firms with higher market valuations, and younger firms are more likely to have AI disclosure. This inclination suggests that these firms could have greater financial resources, competitive pressures, and a culture of innovation and risk-taking. Consequently, they are better positioned to invest in AI technologies and more inclined to affirm their commitment to innovation publicly.

Corporate disclosure should not only reflect the present state of activities but also encompass forward-looking information pertaining to both opportunities and risks (Merkley, 2014; Noh, So, and Weber, 2019; Aghamolla and An, 2021; Kremer, Schreiber, and Skrzypacz, 2024). Therefore, although AI employee share is one of the key determinants of AI disclosure, the AI disclosure in the 10-K statement could contain risks and forward-looking management insights associated with AI technology adoption. To shed light on the information content of voluntary AI disclosure, we extend Babina et al. (2024) by comparing AI disclosure with the fraction of AI employees in predicting firm growth and innovations. We find that AI disclosure in the present year is positively

associated with the following year's sales growth, employment growth, capital investment, and R&D input. In addition, AI disclosure in the present year is negatively associated with the following year's operation efficiency as measured by COGS over sales and operation expenses over the number of employees. The predictability of AI disclosure remains significant even after controlling for the fraction of AI employees, suggesting that AI disclosure is not merely a sideshow of a firm's current AI engagement but also provides additional material forward-looking information associated with the risk and opportunities of AI.

We utilize ChatGPT to classify AI disclosures into five categories: product development, pricing optimization, AI product provision, inventory management, and operational efficiency. To better understand the impact of these activities, we group the first three categories—product development, pricing optimization, and AI product provision—as revenue generation activities, while inventory management and operational efficiency are categorized as cost reduction activities. Our analysis reveals that the number of firms employing AI for both revenue generation and cost reduction has grown at a similar pace over time. Additionally, we find that both revenue generation and cost reduction through AI play a significant role in explaining future growth and operational efficiency, highlighting the dual importance of AI in driving firm performance.

Our work adds to the literature that aims to gauge firm-level AI engagement and its impacts. Eisfeldt, Schubert, and Zhang (2023) find a positive equity premium for firms whose employees' occupations having a greater exposure to generative AI. Babina et al. (2024) use job posting information to measure a firm's fraction of AI-related employees and find it is positively associated with a firm's growth and innovations. Acemoglu, Autor, Hazell, and Restrepo (2022) use online job postings to measure occupational AI exposures and find that, when AI-exposed establishments use AI, the hiring in non-AI positions is dampened, with a change of skill requirements in the

remaining positions. Our study proposes a new measure of a firm's AI disclosure, which is significantly related to a firm's present engagement in AI activities yet includes additional predictive power on firm growth, innovation, and operation efficiency. This new measure complements other firm-level AI engagement metrics and offers a new approach to analyzing the firm's AI activities.

Our study contributes to the growing literature on extracting information from corporate voluntary disclosures. Li, Lundholm, and Minnis (2013) use annual reports to measure competition, while Kim, Wang, and Wu (2023) and Kölbel, Leippold, Rillaerts, and Wang (2024) examine climate risk exposure in the risk factor sections of 10-K statements. Sautner, van Lent, Vilkov, and Zhang (2023), along with Li, Shan, Tang, and Yao (2024), extract firm-level climate risk measures from earnings call transcripts, and Bourveau et al. (2022) and Zhang (2022) analyze human capital disclosures in annual reports. Additionally, Hassan, Hollander, van Lent, and Tahoun (2019) provide firm-level political risk measures using textual analysis of quarterly earnings calls. By applying textual analysis to annual reports to examine AI disclosures, our research introduces a novel dimension to corporate disclosure analysis. Specifically, our study of AI activities offers deeper insights into firms' engagement with AI, thus contributing to a broader understanding of the evolving corporate disclosure landscape and the role of advanced technologies in shaping corporate strategies.

Furthermore, our study offers policy implications, particularly in light of increasing regulatory efforts in artificial intelligence. President Biden's Executive Order on AI in October 2023 underscores the importance of developing standards and practices to enhance consumer safety and privacy.[8] By analyzing AI disclosure through textual analysis in annual reports, regulators can gain

[8] https://www.whitehouse.gov/briefing-room/presidential-actions/2023/10/30/executive-order-on-the-safe-secure-and-trustworthy-development-and-use-of-artificial-intelligence/.

valuable insights into firms' AI strategies, potential risks, and approaches to safeguarding consumer interests. This information can inform the development of regulatory frameworks tailored to the rapidly evolving AI landscape.

2. Measuring AI Disclosure in SEC Filings

2.1. AI-related keywords identification

Our first task is to compile a list of AI-related keywords that allow us to pinpoint potential AI disclosure. We obtain 10-K filings from 2010 to 2023 using Stage One 10-X Parse Files from the Notre Dame Software Repository for Accounting and Finance, established by Loughran and McDonald (2016). In the first step, following the textual analysis approach by Kim, Wang, and Wu (2023), we identify all 10-Ks that mention "artificial intelligence."[9] After this pre-selection, we randomly choose 25 reports with "artificial intelligence" mentioned in the 10-Ks for each year from 2010 to 2023 (350 reports in total). We manually read through this set of 10-Ks and create the initial AI-related keyword list by including the frequently mentioned AI-related terminology words and bigrams in these reports. To ensure the high relevance of keywords, we limit the initial list to 25 keywords.

In the next step, we improve this initial AI-related keyword list using an out-of-sample learning approach. The out-of-sample learning approach compares classification outcomes using the keyword list approach versus the manual review approach for 50 randomly selected firms yearly for the 14-year sample period (700 reports). A report is automatically classified as AI-related if it includes words in the AI-related keyword list. Then, we use human review to classify if a report is AI-related. The accuracy of using the AI-related keyword list in classification is measured as the

[9] Sample 10-K report paragraphs with "artificial intelligence" explicitly referred to are shown in Appendix A1.

percentage of automatically identified AI-related reports by keyword list that are consistent with manually identified AI-related reports. If accuracy is less than 99%, we read manually identified AI-related reports that differ from keyword list classification. We then add three new keywords in each improvement round. We repeat this improvement process and stop once the accuracy ratio reaches 99%. The final AI disclosure keyword list is presented in Table A2, with the frequency of the keyword showing in all 10-K filings and AI-disclosing ones.

Table A3 tracks the top 3 most frequently mentioned AI keywords in the final list from 2010 to 2023. In the early years (2010 to 2013) within our sample, the AI keywords with the highest appearance frequency in the 10-K reports are "business intelligence," "AI/ML," and "data mining." In the following years (2014 to 2017), "big data" picks up the trend and becomes the top 3 highest referenced AI keywords, while "business intelligence" remains the highest cited AI keyword. Starting in 2018, "artificial intelligence" and "machine learning" are the top two highest-mentioned AI keywords in the corporate annual reports. The keyword "business intelligence" loses its third position and is replaced by "data science" in 2022. The evolution of AI technologies over time is also reflected in the first appearance of keywords in the annual reports. For example, "big data" first showed up in 2011, suggesting the focus of AI applications shifted to data science. A more recent AI advancement, "deep learning," appeared in the 2014 annual reports. In 2023, the reference to the "large language model" quickly occupies the reports, coinciding with the introduction of ChatGPT in late 2022.

2.2. AI disclosure identification and classification

The keyword list approach has a few potential limitations. First, it is susceptible to human bias and inconsistency in classifying 10-K reports. Second, the same keywords can have different meanings in varying contexts. For example, "data mining" might refer to applications of AI or

machine learning algorithms, but it could also be irrelevant to AI when discussing data processes related to analyzing customer behavior. Third, using keywords alone makes it challenging to accurately identify the purpose of a firm's engagement with AI. To address these issues, we utilize the latest large language model, ChatGPT-4o, to improve the precision of classifying AI-related disclosures and identify the motivations behind a firm's AI initiatives.

We first apply the aforementioned keyword list to identify potential AI disclosures in the 10-K filings. Since we focus on AI applications, we search for AI-related keywords in Item 1 (Business) or Item 7 (Management's Discussion and Analysis). When a keyword is identified, we extract an 800-word window around it (400 words before and after) and employ the ChatGPT-4o model to review this text.[10] The model then determines whether the disclosure pertains to AI (*AI Disclosure*) and, if so, classifies it into one of five categories: Product Development, Pricing Optimization, AI Product Provider, Inventory Management, and Operational Efficiency.

We further classify Product Development, Pricing Optimization, and AI Product Provider as revenue-related AI disclosures (*AI Disclosure-Revenue*), while Inventory Management and Operational Efficiency fall under cost-related AI disclosures (*AI Disclosure-Cost*). For this classification task, we apply a "zero-shot prompting" approach due to the challenge of providing a solid and comprehensive definition of AI. Additionally, we require ChatGPT to provide a probability score for the AI-related classification and for each application category separately. The model also provides an explanation for these scores, allowing an audit of its reasoning. Table A4 details the GPT prompts used for classification and categorization. This procedure results in 6,028 firm-year observations with AI disclosures from 1,985 unique firms.

2.3. Measuring AI intensity

[10] In most cases, the length of the extracted corpus exceeds 800 words to ensure the sentences remain complete.

To measure AI intensity in disclosures, we utilize ChatGPT to identify AI-related content within the 800-word corpus. Our identification process employs the "few-shot prompting" strategy, which generates clear and accurate responses when the task is well-defined. This strategy involves three steps: 1) System Prompt Creation: We create a system prompt that outlines the setup of the task. 2) Example User-Assist Prompts: We develop example prompts to illustrate the expected output. In these examples, the user prompt asks ChatGPT to remove content unrelated to AI disclosure, while the assist prompt provides the sample answers. 3) Task Execution: For each task, we provide the GPT model with the system prompt, the example user-assist prompts, and a new user prompt that includes the AI disclosure text extracted from the 10-K filings. The model then produces responses similar to the assist prompt as instructed.

Table A5 details our prompts and classification procedure. *AI intensity* is measured as the number of words in AI disclosures processed by GPT, scaled by the total number of words in the corpus extracted from 10-K filings. Similar to the AI disclosure indicator, we also create two measures, *AI Intensity-Revenue,* and *AI Intensity-Cost*, to capture AI disclosure intensity in these two dimensions.

3. Trends and Patterns in AI Disclosure

Table 1 reports the number of AI-disclosing firms each year in the sample, demonstrating a rapid growth in AI disclosure among U.S. public firms over the past decade. In 2010, only 83 firms (2.36% of firms) disclosed AI-related issues in their 10-K statements. This number surged to 834 firms (20.02%) in 2023. Figure 1 illustrates this trend, clearly revealing an accelerated growth regime post-2017. This pattern aligns with the significant advancements in AI technologies in 2017

and 2018, driven by breakthroughs in deep learning, particularly in natural language processing and image recognition tasks.[11]

[Insert Figure 1 and Table 1 about here]

We calculated the proportions of AI disclosure firms within each industry sector using the two-digit NAICS codes to investigate the pattern of AI disclosure across different industries. Figure 2 illustrates the evolution of AI disclosure in various industries from 2010 to 2023. Two main observations emerge. First, AI disclosure has become ubiquitous, with all industry sectors reporting AI disclosures by 2023. This was not the case in 2010 when many industries, such as healthcare, construction, utilities, mining, and agriculture, had no AI disclosures. In addition, in 2023, over 30% of firms in sectors like information, education, professional and business services, agriculture, administrative support, and healthcare reported AI disclosures in their 10-K statements. Clearly, the impact of AI technology has permeated the entire economy.

[Insert Figure 2 about here]

Second, industry sectors with the highest AI disclosure have shifted over time. Panel A of Figure 2 shows that, in 2010, the three industries with the most intensive AI disclosure were professional and business services (11.18%), information (11.11%), and administrative and entertainment (9.52%). By 2023, the top three industries had shifted to information (56.47%),

[11] Notable contributions include the introduction of the transformer architecture in "Attention is All You Need" by Vaswani et al. (2017), which revolutionized natural language processing, and the development of BERT, a groundbreaking language representation model, as outlined in "BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding" by Devlin et al. (2018). These years also saw increased adoption of AI across various industries and heightened discussions surrounding AI ethics, regulation, and the development of autonomous systems, as documented in the "2018 AI Index Report" by Stanford University.

education (52.94%)[12], and professional and business services (49.23%).[13] The high percentage of firms in the education industry discussing AI in their annual reports is somewhat surprising. This underscores how AI has evolved into a highly sophisticated educational technology, presenting both significant challenges and opportunities for the education sector.

Panels B and C of Figure 2 break down AI disclosure into revenue and cost categories. We find that revenue and cost disclosure follow a similar distribution to overall disclosure, with the top three industries in 2023 being information (42.46%), education (35.29%), and professional and business services (29.23%) for revenue, while the top three industries for cost disclosure are education (29.41%), information (27.59%), and professional and business services (26.92%).

In Panel D of Figure 2, the growth of the AI risk disclosure is significant among different industries. In 2010, rare industrial sectors have the AI risk disclosure. In 2023, the top three industrial sectors with AI risk disclosure are information (22.20%), administrative support (18.42%) and education (17.65). Almost of the industrial sectors have firms with AI risk disclosure (except for agriculture, entertainment, and other services).

4. Data and Summary Statistics

[12] We take a closer look at AI reporting in the education sector and read all the 2023 10-K reports from the education sector. The annual reports show that educational institutions actively use artificial intelligence and machine learning to provide learning resources to their customers. For example, Nerdy Inc. stated in the 10-K: "Our purpose-built proprietary platform leverages technology, including *artificial intelligence* (AI), to connect students, users, parents, guardians, and purchasers…". Coursera Inc. included the following in the 10-K statement: "Coursera's data and *machine learning* systems drive personalized learning, and skills benchmarking."

[13] This is partially due to the number of firms in this industry sector being relatively small (15 firms in 2023 in our sample). We also investigate the industry sectors (by 2-digit NAICS code) with the most companies disclosing AI-related matters in 2010 and 2023. The top 3 industry sectors with the most firms disclosing AI were manufacturing, information, and professional and business services in 2010. The top 3 industry sectors with the most firms disclosing AI switch to information, manufacturing, and finance/insurance in 2023. The number of AI disclosing firms may be biased by the total number of firms in that industry but represents the industry-wise AI disclosure intensity.

We merge AI disclosure data with multiple datasets to analyze AI disclosure's determinants and information content. Firm characteristics and financial information are from the Compustat database. Information on stock performance is from CSRP. Finally, we obtain institutional ownership from Thomson Reuters 13F. Combining these databases creates a final sample containing 4,290 unique firms from 2010 to 2023, with about 32,000 firm-year observations.

The summary statistics of the main variables used in this study are reported in Table 2, with variable definitions provided in Table A6. The mean of AI Disclosure is 9.4%, indicating that 9.4% of 10-K statements during the sample period disclose AI information. AI Disclosure-Revenue (Cost) has an average value of 5.9% (4.5%), suggesting that 5.9% (4.5%) of 10-K filings disclose revenue (cost)-related AI information. The AI risk disclosure is relatively sparse compared with the AI usage information. 1.2% of the sample contains the AI risk disclosure in item 1A of the 10-K statement. Another key variable in this study is the AI Employee Share, developed by Babina et al. (2024), representing the proportion of employees with AI-related skills (as identified in their resumes) relative to the total workforce. On average, our sample firms have 0.1% of employees with AI-specific knowledge, 3,300 employees on average, $404.7 million in gross sales, $849.8 million in total assets, cash holdings at 22.4% of total assets, a market-to-book ratio of about 2.2, and an average of 13.8 years since IPO. Table 3 shows the pairwise correlation between the main variables. It shows that the likelihood of having AI disclosure is positively correlated with the share of employees with AI knowledge.

[Insert Tables 2 and 3 about here]

5. Determinants of Corporate AI Disclosure

The public generally perceives AI technologies as a novel component that could provide firms with a significant competitive edge by enhancing operation efficiency, decision-making, and

innovation. Firms may choose to disclose their AI initiatives as a strategic move to showcase their commitment to innovation and technological advancement, thereby enhancing investor confidence and attracting potential stakeholders. However, it is also possible that some firms may disclose AI initiatives to attract investors, even if they lack substantial AI activities, as a means of leveraging investor interest in AI technologies. Therefore, the voluntary nature of AI disclosure raises the question of what factors influence the firm to disclose, particularly whether such disclosure is related to a firm's actual engagement in AI.

To gain insights into the determinants of AI disclosure, we investigate how a firm's AI disclosure is related to its AI activities and other key firm characteristics by estimating the following regression:

$$AI\ Disclosure\ Vars_{i,t+1}/AI\ Intensity\ Vars_{i,t+1} = \beta \cdot AI\ Employee\ Share_{i,t} + \lambda \cdot Other\ Firm\ Chacteristics_{i,t} + D_j + D_t + \epsilon_{i,t+1}$$

where the dependent variable is either *AI Disclosure* or *AI Intensity*. We use the logistic regression models when the dependent variable is AI Disclosure or AI Disclosure-Revenue (Cost). When the dependent variable is *AI Intensity* or *AI Intensity-Revenue (Cost)*, we use OLS regression. The independent variable $AI\ Employee\ Share_{i,t}$ is the fraction of AI employees developed by Babina et al. (2024), gauging a firm's real activities in AI. The advancement of AI necessitates the recruitment of employees proficient in AI, as they hold specialized expertise crucial for tasks like algorithm design, data analysis, and model development. Consequently, the proportion of AI-skilled employees within a company signifies the resources dedicated by firms to AI endeavors. Other firm characteristics include firm size (logarithm of total assets), return on assets, cash holdings, leverage ratio, net PP&E, market-to-book ratio, and (logarithm of) firm age.

The estimation results are shown in Table 4. Columns (1), (2), and (3) report the results when the dummy *AI Disclosure* is the dependent variable. In Column (1), the coefficient on $AI\ Employee\ Share_{i,t}$ is positive and significant at a 1% level. It shows that the share of employees with AI skills positively affects the likelihood of AI disclosure. Economically, one percentage increase in AI Employee Share leads to a 0.6% increase[14] in the likelihood of AI disclosure, which represents a 6.4% increase relative to the sample mean (0.094). The results show that a firm's real engagement in AI is a key factor affecting AI disclosure.

The results also show that firms with larger sizes, larger cash holdings, lower leverage, lower tangibility, higher growth opportunities, and younger ages are more likely to have AI disclosures. Larger firms are more likely to engage in AI disclosure due to their extensive resources and capabilities to invest in cutting-edge technologies like AI, signaling their commitment to innovation and maintaining market leadership. Cash holdings and leverage are indicative of a firm's financial resources and associated risks. The coefficients show that firms with greater financial capability and lower risk profiles are more likely to disclose AI, which suggests that AI projects are resource-intensive and risky. Lower tangibility firms, relying more on intangible assets like technology, might use AI disclosure to signal their future growth prospects. Similarly, firms with higher growth opportunities and younger age might prioritize innovation to capitalize on growth potential. Thus, disclosing their AI initiatives serves to signal their commitment to technological advancement, aligning with their growth-oriented strategies.

Columns (2) and (3) report the determining factors of revenue-related and cost-related AI disclosures, respectively. The coefficients in column (2) remain largely unchanged, except that ROA is negatively correlated with revenue-related AI disclosure. This suggests that companies

[14] When calculating the marginal effect, we assume the rest of the covariates taking the sample mean value.

tend to disclose more AI information related to boosting revenue during periods of lower performance. In column (3), the coefficient for AI Employee Share has a lower magnitude and loses significance, indicating that most firms hire AI employees primarily to enhance revenue rather than reduce costs. Columns (4), (5), and (6) present the results for AI Intensity measures. The coefficients for AI Employee Share and other characteristics are largely similar to those for AI Disclosure. Overall, the results indicate that AI disclosure extends beyond merely catering to investors; it is intricately linked to a firm's engagement with AI technologies.

6. Information Content of AI Disclosure

Corporate disclosure should not only reflect the present state of activities but also encompass forward-looking information pertaining to both opportunities and risks (Merkley, 2014; Noh, So, and Weber, 2019; Aghamolla and An, 2021; Kremer, Schreiber, and Skrzypacz, 2024). However, a firm's AI voluntary disclosure may not necessarily include material information useful for investors to predict its future performance. Some firms may disclose comprehensive details about their AI initiatives, including strategies, investments, risks, and outcomes, providing investors with valuable insights into the firm's prospects. Conversely, other firms may opt for minimal disclosure, withholding sensitive information to maintain a competitive advantage or mitigate perceived risks. Additionally, the complexity and novelty of AI technologies present challenges in accurately quantifying their potential impact on future performance. Therefore, even if a firm provides AI disclosure related to its present AI activity, it might not include adequate information that is forward-looking of a firm's future performance. In this section, we assess the information content of AI disclosure and investigate if AI disclosure is related to a firm's future growth and corporate costs.

6.1. AI disclosure and firm growth

We estimate the following regression to investigate the information content of AI disclosure about a firm's future growth and innovation.

$$Firm\ Growth_{i,t+1} = \beta_1 \cdot AI\ Disclosure\ Vars_{i,t} + \beta_2 \cdot AI\ Employee\ Share_{i,t} + \lambda \cdot Controls_{i,t} + D_j + D_t + \epsilon_{i,t+1}.$$

where the dependent variable $Firm\ Growth_{i,t+1}$ is a firm's performance in year t+1. Similar to Babina et al. (2024), we use three different dependent variables to measure a firm's growth, including the logarithm of the number of employees, the logarithm of sales and investment. $AI\ Disclosure_{i,t}$ and $AI\ Employee\ Share_{i,t}$ are two main independent variables. Both variables are measured in year t, allowing us to examine its predictability of performance in year t+1. Other control variables include the logarithm of total assets, return on assets, cash holdings scaled by total assets, leverage ratio, net PP&E scaled by total assets, market-to-book ratio, and the logarithm of firm age. D_j and D_t represent the industry sector (2-digit NAICS code) fixed effect and year fixed effect.

Columns (1) and (2) in Panel A of Table 5 report the results with the logarithm of employment as the dependent variable. The coefficient on $AI\ Disclosure_{i,t}$ is positive and statistically significant, suggesting that AI disclosure is associated with an increase in total employees. Economically, in column (2), the coefficient on $AI\ Disclosure_{i,t}$ is 0.139 (t-stats. = 6.653), suggesting that firms with AI disclosure generally have a 13.9% increase in employee population compared to the firms without AI disclosure. These results demonstrate that AI disclosure positively predicts the growth in the number of employees within a firm, whereas the AI employee

share does not have a similar effect. This evidence implies that AI disclosure provides additional explanatory information regarding firms' employment and staffing.

[Insert Table 5 about here]

Columns (3) and (4) present the results with the logarithm of sales in year t+1 as the dependent variable. The coefficients for AI disclosure are consistently positive and significant across both regression specifications, with a coefficient of 0.16. Economically, firms with AI disclosure experience a 16% increase in gross sales in the following year. Columns (5) and (6) of Table 5 report the results with firm investment as the dependent variable. When all control variables are included in column (6), the coefficient for AI disclosure is 0.002 (t-stats. = 3.588). This coefficient is also economically significant, indicating that AI-disclosing firms tend to have 6.1% higher capital expenditures in the next year relative to the sample mean (0.033) than firms without AI disclosure. Columns (7) and (8) show the impact of AI disclosure on R&D expenditures in year t+1. With all control variables included in column (8), the coefficient for AI disclosure is 0.464 (t-statistic: 11.371), suggesting that firms disclosing AI information tend to have 15.3% higher capital expenditures compared to non-disclosing firms relative to the sample mean (2.806).

Panel B of Table 5 presents the results by separating AI Disclosure into revenue and cost components. In columns (1), (4), (7), and (10), we include only *AI Disclosure-Revenue*, while in columns (2), (5), (8), and (11), we only retain *AI Disclosure-Cost.* In columns (3), (6), (9), and (12), we put both into the regressions. The coefficients of *AI Disclosure-Revenue* and *AI Disclosure-Cost* are mainly positive and statistically significant across all columns, but the coefficient of *AI Disclosure-Revenue* is insignificant in *log(Employees), log(Sales),* and *CAPX/Assets.*

Panel C reports the results from estimating the predictability of AI intensity variables on firm growth. Columns (1), (4), (7), and (10) present the results for overall AI intensity. Columns (2), (5), (8), and (11) show the results for revenue-related AI intensity, while columns (3), (6), (9), and (12) present the results for cost-related AI intensity. We find that firms with higher AI intensity in SEC filings tend to experience increases in the number of employees, sales, investment in capital expenditures, and R&D. However, firms disclosing more intensive revenue-related AI information are only associated with higher R&D in the following year. In contrast, firms with more intensive cost-related AI information are more likely to have a higher number of employees, increased sales, and larger investments in capital expenditures and R&D.

6.2. AI disclosure and costs

A key function of AI disclosure is to report a firm's commitment to reducing costs and enhancing efficiency. In this section, we investigate whether AI disclosure is indeed followed up by cost reduction. We focus on the cost of goods sold (COGS) and operating expenses, as these constitute a substantial portion of corporate costs and are closely related to AI applications. We scale COGS by total sales (*COGS/Sales*) and number of employees (*COGS/Employees*). For operating expenses, we scale by the number of employees (*Operating Expense/Employees*).

Panel A of Table 6 reports the regression results estimating the predictability AI disclosure on corporate costs. Columns (1) - (4) use COGS as the dependent variable, with COGS scaled by sales in columns (1) and (2), and by the number of employees in columns (3) and (4). The coefficient of AI Disclosure indicates that AI disclosure is negatively associated with next year's COGS. Economically, firms with AI disclosure in SEC filings experience a 23.04% (14.28%) decrease in COGS scaled by sales (number of employees) in the following year. Additionally, AI disclosure is associated with a 13.36% drop in *Operating Expenses/Employees*. Panel B presents the results by separating AI disclosure into revenue-related and cost-related components. Both *AI*

Disclosure-Revenue and *AI Disclosure-Cost* are associated with a significant decline in corporate costs in the subsequent year across all columns. Panel C shows the results for AI intensity variables. Columns (1), (4), and (7) use overall AI Intensity. Columns (2), (5), and (8) use revenue-related AI intensity, while columns (3), (6), and (9) use cost-related AI intensity. We find that all AI intensity variables are negatively correlated with corporate costs for all proxies, including COGS/Sales, COGS/Employees, and Operating Expenses/Employees. Overall, our findings suggest that AI disclosure and AI intensity signal firms' commitment to reducing corporate costs.

[Insert Table 6 about here]

6.3. AI risk disclosure and firm risk

The risk factor disclosure in item 1A of the 10-K statement provides the forward-looking risk assessment of the underlying business from the firm managers. We consider the AI risk-related disclosure to have distinguished effects on the financial consequences of the firm. Given the forward-looking risk related natural in the item 1A disclosure, we estimate the following regression to investigate the information content of AI risk disclosure about a firm's risk profile.

$$Firm\ Risk_{i,t+1} = \beta_1 \cdot AI\ Risk\ Disclosure\ Vars_{i,t} + \lambda \cdot Risk\ Controls_{i,t} + D_j + D_t + \epsilon_{i,t+1}.$$

where the dependent variable $Firm\ Risk_{i,t+1}$ is a firm's risk related proxies in year t+1. We assess the firm risk profile using the stock inferred and option implied information. We select the idiosyncratic volatility estimated from the annualized residual volatility from the Fama and French (1993) three factor model to represent the stock volatility risk. We follow the option implied volatility literature (Bakshi, Kapadia, and Madan, 2003; Ilhan, Sautner, and Vilkov, 2021; Kelly, Pástor, and Veronesi, 2016) and construct the slopeness of option-implied volatility on option delta (SlopeD), the variance risk premium (VRP), the model-free option implied volatility (MFIV), and

the model-free option implied skewness (MFIS). The option-implied risk metrics reflect the forward-looking firm tail risk and volatility risk.

As for the control variable, we include the logarithm of total assets, the logarithm of firm age, Tobin's Q, book leverage ratio, the cash holding scaled by the total assets, intangible assets scaled by the total assets, the stock return sensitivity with respect to the market return, the operating expense scaled by the total assets. The control variables cover the firm characteristics that may impact the firm's risk profile. The main independent variable is the AI risk disclosure dummy variable. The coefficient of the variable estimates the incremental contribution of the AI risk disclosure on the firm risk.

Using the firm level stock and option implied risk measures as the main dependent variable, we perform the regression analysis with the AI risk disclosure quantified from section item 1A. Panel A of Table 7 presents the regression results using aggregate level AI risk disclosure dummy variable as the main independent variable. We find that the AI risk disclosure in the 10-K statement predicts a higher level of firm tail risk. The coefficient of the AI risk disclosure variable is statistically and economically significant for stock inferred and option implied volatility measures. This indicates that the AI related risk factors are priced in the stocks and options of the firms. The investors gain incremental risk assessment and react to the corporate AI risk disclosure.

[Insert Table 7 about here]

Additionally, we decompose the topics of the AI risk disclosure into six categories: regulation risk, operational risk, competition risk, cybersecurity risk, ethical risk, and third-party risk. We rerun the firm risk regressions using the categorical AI risk disclosure dummy variable as the main independent variable. The empirical results are summarized in Panel B of Table 7. Each line in the table represents a standalone regression, and we report only the coefficients of the main independent variables (categorical AI risk disclosure) for brevity. Overall, the categorical AI risk disclosure forecasts higher firm tail risk. Among the magnitude of the categorical AI risk disclosure

variables, regulation risk, competition risk, and ethical risk seem to be the major AI related risk factors that raise the investors' concern on the stock and option markets.

7. Conclusion

AI has emerged as the most transformative technology in today's business world. AI offers significant potential to enhance future cash flows by enabling innovative products, more precise customer targeting, and operational efficiency through automation and optimization of supply chains and production processes. However, despite these benefits, AI adoption brings substantial uncertainty to investors, arising from evolving regulations, concerns about privacy and ethics, and the potential for cybersecurity vulnerabilities. Additionally, the success of AI implementation is not guaranteed, as it involves risks of integration failure and unforeseen consequences. Given these uncertainties, investors require high-quality information about firms' AI initiatives and the associated risks. Yet, the current regulatory framework lacks specific guidelines for AI reporting, making disclosures largely voluntary. This paper takes the first step in analyzing voluntary AI disclosure.

We use a combination of traditional textual analysis techniques and the latest LLMs like ChatGPT to measure and analyze AI disclosure in corporate annual reports. In the first step, keyword-based analysis is used to identify AI-related information in firms' 10-K filings. However, traditional textual analysis methods have limitations in capturing the full context and nature of AI usage. To address this, we use ChatGPT to classify AI-related disclosures into five categories: product development, pricing optimization, AI product provision, inventory management, and operational efficiency. This approach allows for a deeper understanding of how firms engage with

AI, grouping the first three categories as revenue generation activities and the latter two as cost reduction activities.

Our findings reveal that the number of firms adopting AI for both revenue generation and cost reduction has grown substantially since 2010. Furthermore, both revenue generation and cost reduction through AI are associated with future growth and operational efficiency, indicating that AI plays a crucial role in shaping firm performance. These results underscore the importance of accurate AI disclosures for investors to assess not only current engagements but also the future opportunities and risks associated with AI adoption.

Reference

Aghamolla, C., & An, B. J. (2021). Voluntary disclosure with evolving news. *Journal of Financial Economics, 140*(1), 21-53.

Babina, T., Fedyk, A., He, A., & Hodson, J. (2024). Artificial intelligence, firm growth, and product innovation. *Journal of Financial Economics, 151*, 103745.

Bakshi, G., Kapadia, N., & Madan, D. (2003). Stock return characteristics, skew laws, and the differential pricing of individual equity options. *Review of Financial Studies, 16*(1), 101-143.

Bourveau, T., Chowdhury, M., Le, A., & Rouen, E. (2022). Human capital disclosures. *Working paper*.

Eisfeldt, A. L., Schubert, G., & Zhang, M. B. (2023). Generative AI and firm values. *Working paper*.

Fama, E. F., & French, K. R. (1993). Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics, 33*(1), 3-56.

Giczy, A. V., Pairolero, N. A., & Toole, A. A. (2022). Identifying artificial intelligence (AI) invention: A novel AI patent dataset. *Journal of Technology Transfer, 47*(2), 476-505.

Hassan, T. A., Hollander, S., van Lent, L., & Tahoun, A. (2019). Firm-level political risk: Measurement and effects. *Quarterly Journal of Economics, 134*(4), 2135-2202.

Ilhan, E., Sautner, Z., & Vilkov, G. (2021). Carbon tail risk. *Review of Financial Studies, 34*(3), 1540–1571.

Kelly, B., Pástor, Ľ., & Veronesi, P. (2016). The price of political uncertainty: Theory and evidence from the option market. *Journal of Finance, 71*(5), 2417-2480.

Kim, J. B., Wang, C., & Wu, F. (2023). The real effects of risk disclosures: evidence from climate change reporting in 10-Ks. *Review of Accounting Studies, 28*(4), 2271-2318.

Kölbel, J. F., Leippold, M., Rillaerts, J., & Wang, Q. (2024). Ask BERT: How regulatory disclosure of transition and physical climate risks affects the CDS term structure. *Journal of Financial Econometrics, 22*(1), 30-69.

Kremer, I., Schreiber, A., & Skrzypacz, A. (2024). Disclosing a random walk. *Journal of Finance*, forthcoming.

Li, F., Lundholm, R., & Minnis, M. (2013). A measure of competition based on 10-K filings. *Journal of Accounting Research, 51*(2), 399-436.

Li, Q., Shan, H., Tang, Y., & Yao, V. (2024). Corporate climate risk: Measurements and responses. *Review of Financial Studies*, forthcoming.

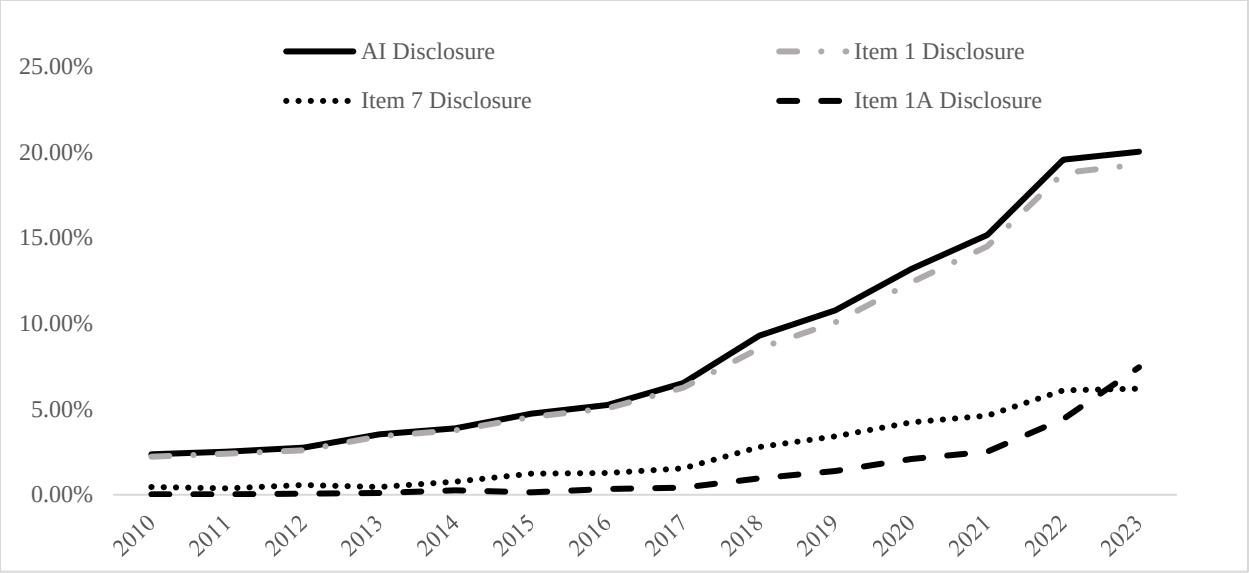
Merkley, K. J. (2014). Narrative disclosure and earnings performance: Evidence from R&D disclosures. *The Accounting Review, 89*(2), 725-757.

Noh, S., So, E. C., & Weber, J. P. (2019). Voluntary and mandatory disclosures: Do managers view them as substitutes? *Journal of Accounting and Economics, 68*(1), 101243.

Sautner, Z., van Lent, L., Vilkov, G., & Zhang, R. (2023). Firm-level climate change exposure. *Journal of Finance, 78*(3), 1449-1498.

Zhang, M. (2022). Determinants and consequences of human capital management disclosure. *Working paper*.

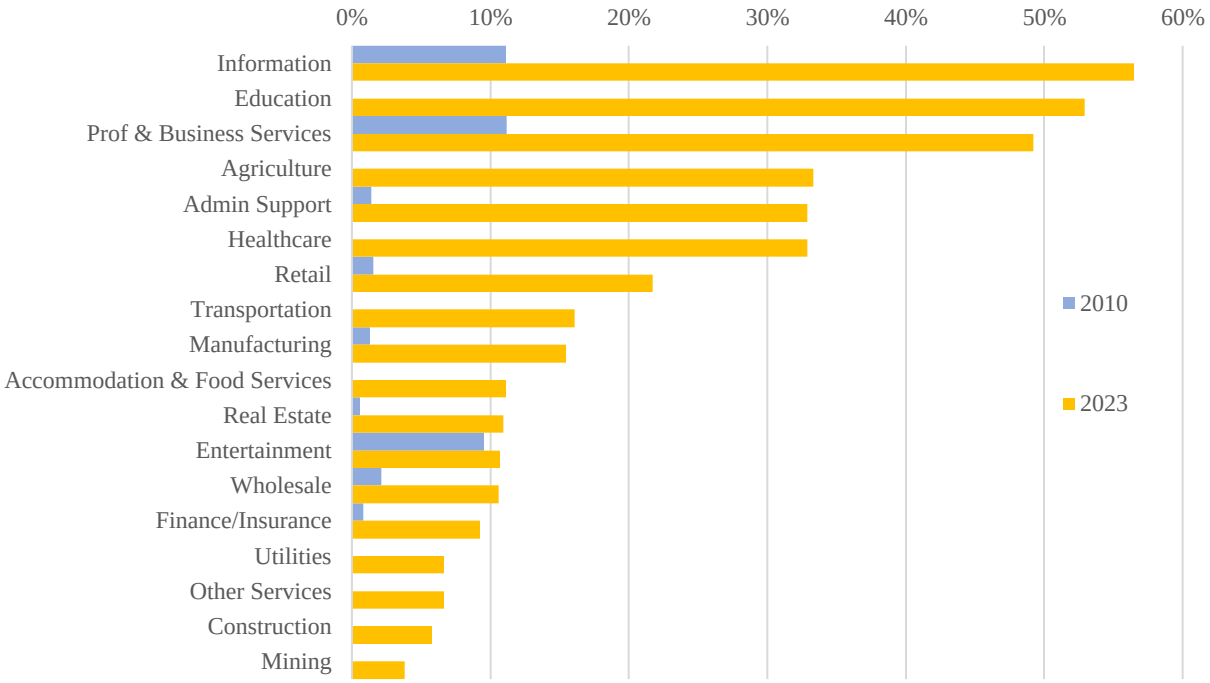
Figure 1: Trends of AI Disclosure



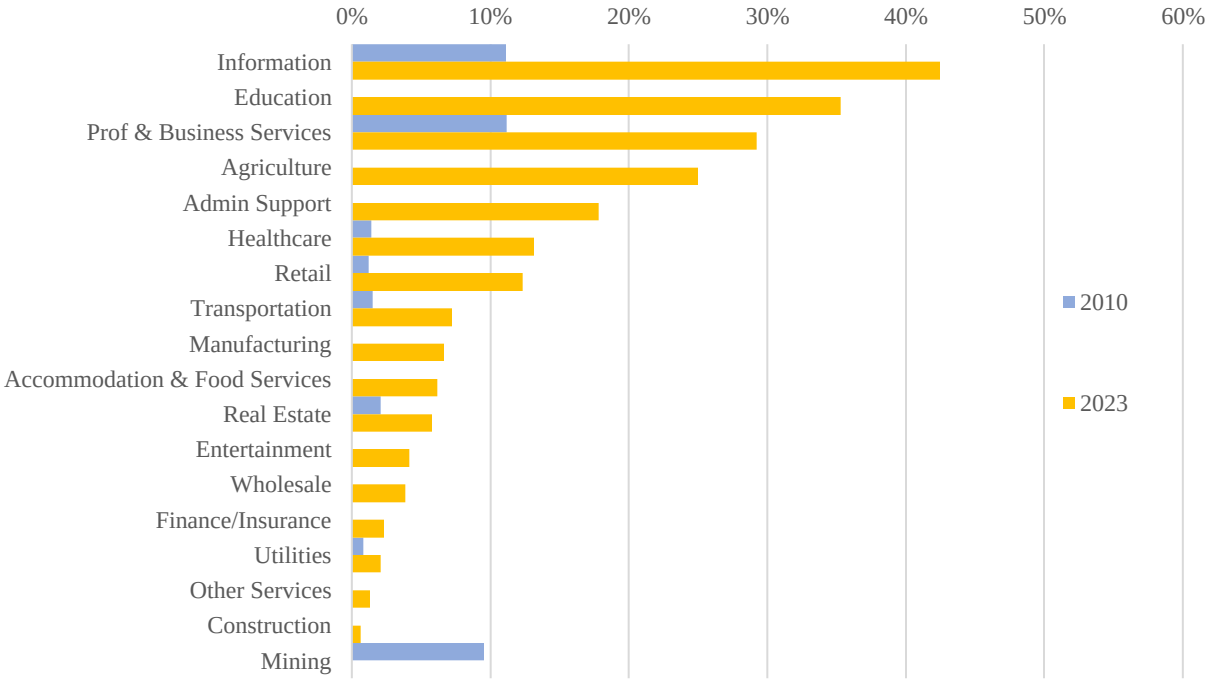
Note: Trends of categorical AI disclosure in the 10-K statement. The graph illustrates the fraction of firms with AI disclosure over the total number of firms. The solid line plots the trend of AI disclosure shown in the full 10-K filings. The dot-dashed line plots the trend of AI disclosure in section Item 1 business. The dotted line plots the trend of AI disclosure in section Item 7, management discussion and analysis. The dashed line plots the trend of the AI risk disclosure in Item 1A risk factors.

Figure 2: Industrial Distribution of AI Disclosure

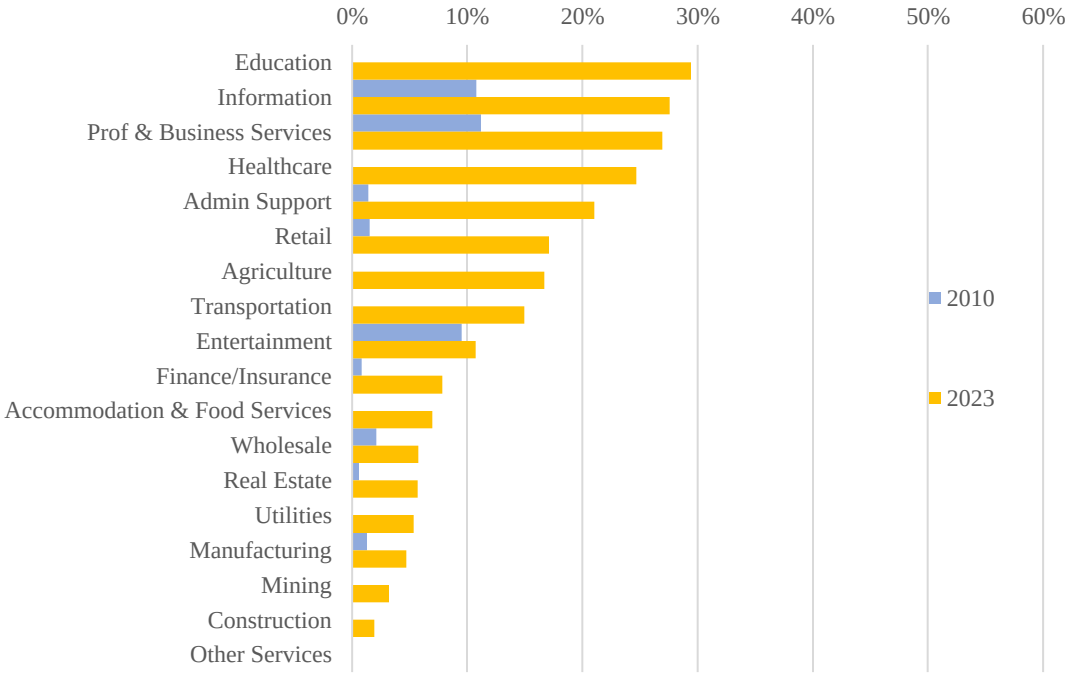
Panel A: AI Disclosure



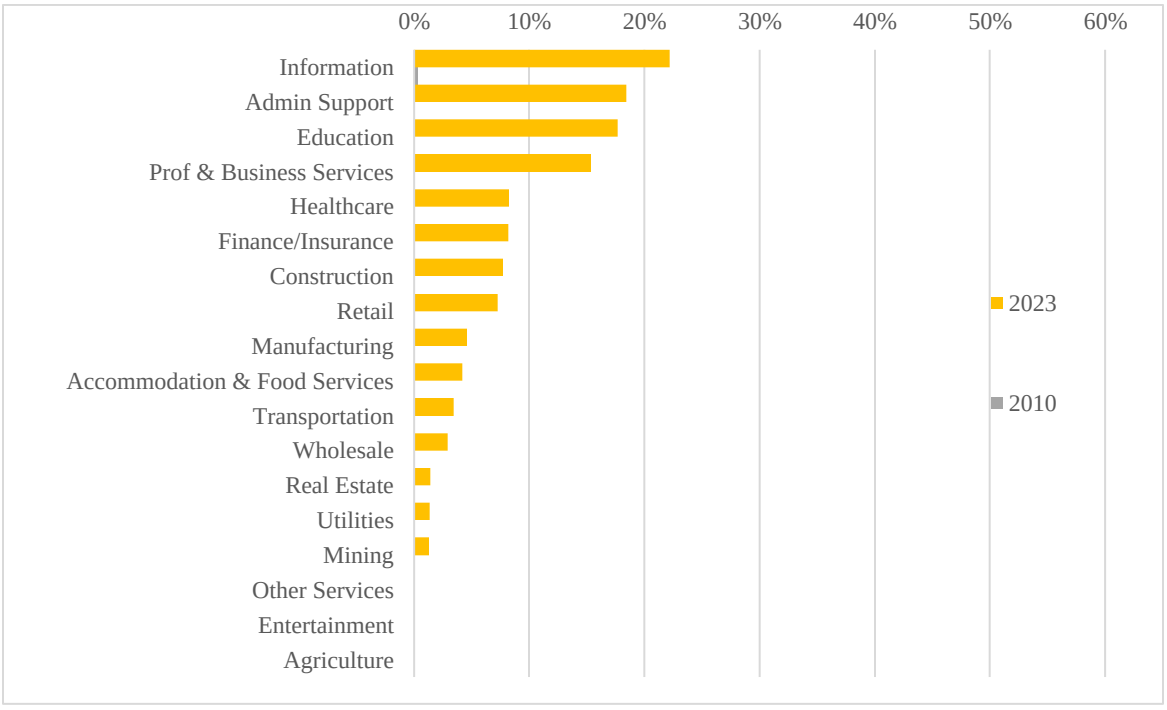
Panel B: AI Disclosure-Revenue



Panel C: AI Disclosure-Cost



Panel D: AI Risk Disclosure



Note: This figure shows industrial distribution of AI disclosures. Panel A, B, C and D report the distribution based on overall AI disclosure, AI revenue disclosure, AI cost disclosure, and AI risk disclosure respectively.

Table 1: Trends of AI Disclosure

Panel A: AI disclosure in item 1 and item 7

Year	Total Firms	AI Disclosure	AI Revenue Disclosure	AI Cost Disclosure	Product Development	Pricing Optimization	AI Product Provider	Inventory Management	Operation Efficiency
2010	3,511	83	42	43	17	1	30	3	42
2011	3,460	87	44	43	19	1	29	2	41
2012	3,435	94	51	50	21	1	33	3	47
2013	3,444	121	71	55	26	3	48	2	53
2014	3,550	138	77	68	35	3	44	2	66
2015	3,642	172	102	81	38	4	69	1	80
2016	3,650	191	122	86	50	5	82	1	85
2017	3,606	235	139	111	76	7	74	2	109
2018	3,600	334	212	151	124	4	118	6	148
2019	3,652	393	255	169	155	8	147	6	165
2020	3,670	483	303	229	186	12	182	8	225
2021	3,950	599	380	282	231	21	218	9	275
2022	4,156	812	508	407	335	25	274	12	399
2023	4,166	834	547	411	362	22	284	11	403

Panel B: AI risk disclosure in item 1A

Year	Total Firms	AI Disclosure	AI Risk Disclosure	Regulatory Risk	Operational Risk	Competition Risk	Cybersecurity Risk	Ethical Risk	Third-party Risk
2010	3,511	83	1	0	1	0	0	0	0
2011	3,460	87	1	0	1	0	0	0	0
2012	3,435	94	2	0	2	0	1	0	0
2013	3,444	121	3	1	1	1	1	1	0
2014	3,550	138	9	0	4	3	3	0	0
2015	3,642	172	5	2	2	1	2	0	0
2016	3,650	191	12	4	5	0	1	1	0
2017	3,606	235	15	0	7	4	3	0	0

2018	3,600	334	34	3	14	9	9	1	0
2019	3,652	393	51	4	26	12	12	2	3
2020	3,670	483	77	8	40	14	18	3	3
2021	3,950	599	99	19	49	20	21	13	4
2022	4,156	812	184	51	106	30	29	29	10
2023	4,166	834	310	91	153	46	56	50	22

Note: Panel A of this table summarizes the number of firms in the sample, the number of firms with AI disclosures (AI Disclosure) in SEC filings, and the categories of AI disclosure for each calendar year. AI revenue (cost) indicates that AI is primarily applied to enhance revenue(cost). AI revenue disclosure includes categories of product development, pricing optimization, and AI product provider. AI cost disclosure includes categories of inventory management and operational efficiency. Panel B of this table summarizes the number of firms in the sample, the number of firms with AI risk disclosures in item 1A (AI Risk Disclosure) in SEC filings, and the categories of AI risk disclosure for each calendar year. The AI risk disclosure categories include regulatory risk, operational risk, competition risk, cybersecurity risk, ethical risk, and third-party risk. The sample covers U.S. public firms from 2010 to 2023.

Table 2: Summary Statistics

Panel A: AI Related Variables	N	Mean	S.D.	P25	P50	P75
AI Disclosure	47,870	0.094	0.291	0	0	0
AI Disclosure-Revenue	47,870	0.059	0.235	0	0	0
AI Disclosure-Cost	47,870	0.045	0.207	0	0	0
AI Disclosure-Product Development	47,870	0.035	0.183	0	0	0
AI Disclosure-Inventory Management	47,870	0.001	0.037	0	0	0
AI Disclosure-Operational Efficiency	47,870	0.044	0.205	0	0	0
AI Disclosure-Pricing Optimization	47,870	0.002	0.049	0	0	0
AI Disclosure-AI Product Provider	47,870	0.033	0.180	0	0	0
AI Intensity	47,870	0.015	0.049	0	0	0
AI Risk Disclosure	47,870	0.012	0.108	0	0	0
AI Risk Disclosure-Regulation	47,870	0.004	0.063	0	0	0
AI Risk Disclosure-Operation	47,870	0.009	0.094	0	0	0
AI Risk Disclosure-Competition	47,870	0.005	0.071	0	0	0
AI Risk Disclosure-Cybersecurity	47,870	0.004	0.061	0	0	0
AI Risk Disclosure-Ethical	47,870	0.004	0.062	0	0	0
AI Risk Disclosure-Third Party	47,870	0.002	0.044	0	0	0
AI Employee Share	29,130	0.101	0.489	0	0	0

Panel B: Firm Growth and Risk	N	Mean	S.D.	P25	P50	P75
log(Employees)	47,187	1.184	1.279	0.143	0.700	1.902
log(Sales)	45,623	6.003	2.445	4.478	6.205	7.700
CAPX/Assets	47,870	0.033	0.046	0.003	0.017	0.042
log(R&D)	23,716	2.806	2.147	0.742	2.881	4.331
log(AI Patents)	47,870	0.006	0.083	0	0	0
log(Product Patents)	47,870	0.015	0.144	0	0	0
COGS/Sales	45,655	1.528	6.990	0.360	0.608	0.778
COGS/Employees	45,630	4.243	8.387	0.724	1.644	4.071
Operating Expense/Employees	45,630	6.101	9.573	1.828	3.120	6.341
IVOL	47,870	0.023	0.016	0.013	0.019	0.029
SlopeD	47,870	0.586	0.510	0.209	0.422	0.859
VRP	47,870	0.049	0.248	-0.089	-0.034	0.091
MFIV	47,870	0.447	0.418	0.163	0.301	0.584
MFIS	47,870	-0.437	0.483	-0.755	-0.460	-0.140

Panel C: Control Variables	N	Mean	S.D.	P25	P50	P75
log(Assets)	47,870	6.745	2.247	5.288	6.892	8.290
ROA	47,870	-0.070	0.304	-0.049	0.015	0.060
Cash	47,870	0.224	0.262	0.036	0.110	0.312
Leverage	47,685	0.245	0.244	0.030	0.179	0.387

Net PPE	46,178	0.202	0.239	0.025	0.102	0.285
MtB	47,779	2.217	2.057	1.051	1.460	2.419
log(Firm Age)	47,464	2.624	1.003	1.946	2.773	3.367
Tobin's Q	47,870	2.862	1.762	1.732	2.392	3.320
Intan/Assets	47,870	0.162	0.204	0.003	0.062	0.272
β_{MKT}	47,870	1.017	0.528	0.664	1.008	1.346
XOPR/Assets	47,870	0.706	0.674	0.206	0.538	0.986

Note: This table presents summary statistics of the main variables. The sample period is from 2010 to 2023. All variables are defined in Appendix A. All continuous and variables are winsorized at the 1st and 99th percentile.

Table 3: Correlation Matrix

Panel A: AI Dislosure

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
(1) AI Disclosure	1.00									
(2) AI Disclosure Intensity	0.92	1.00								
(3) AI Revenue Disclosure	0.78	0.72	1.00							
(4) AI Cost Disclosure	0.67	0.60	0.14	1.00						
(5) AI Disclosure - Product Development	0.59	0.54	0.76	0.12	1.00					
(6) AI Disclosure - Inventory Management	0.11	0.12	0.02	0.17	0.01	1.00				
(7) AI Disclosure - Operational Efficiency	0.67	0.59	0.14	0.99	0.12	0.04	1.00			
(8) AI Disclosure - Pricing Optimization	0.15	0.13	0.20	0.09	0.04	0.07	0.08	1.00		
(9) AI Disclosure - AI Product Provider	0.58	0.55	0.75	0.15	0.31	0.01	0.16	0.03	1.00	
(10) AI Employee Share	0.24	0.24	0.26	0.09	0.26	0.00	0.09	0.06	0.20	1.00

Panel B: AI Risk Disclosure

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
(1) AI Risk Disclosure	1.00						
(2) AI Risk Disclosure – Regulation	0.58	1.00					
(3) AI Risk Disclosure – Operation	0.87	0.39	1.00				
(4) AI Risk Disclosure – Competition	0.65	0.25	0.68	1.00			
(5) AI Risk Disclosure – Cybersecurity	0.56	0.28	0.54	0.26	1.00		
(6) AI Risk Disclosure – Ethical	0.57	0.43	0.50	0.30	0.23	1.00	
(7) AI Risk Disclosure – Third Party	0.41	0.22	0.41	0.27	0.18	0.28	1.00

Note: This table presents Pearson correlation coefficients across our main independent variables. The sample period is from 2010 to 2023. All variables are defined in Appendix A. All continuous and variables are winsorized at the 1st and 99th percentile.

Table 4: Determinants of the AI Disclosure.

Spec.	Disclosure Dummy Logit			Disclosure Intensity Panel		
Dep. Var.	*AI Disclosure*	*AI Disclosure -Revenue*	*AI Disclosure -Cost*	*AI Intensity*	*AI Intensity -Revenue*	*AI Intensity -Cost*
	(1)	(2)	(3)	(4)	(5)	(6)
AI Employee Share	0.588***	0.630***	0.122	0.045***	0.042***	0.009*
	(6.344)	(7.394)	(1.317)	(7.967)	(7.846)	(1.794)
log(Assets)	0.321***	0.249***	0.226***	0.003**	0.002**	0.001*
	(6.207)	(4.554)	(3.965)	(2.535)	(2.121)	(1.888)
ROA	-0.351	-0.555*	0.402	0.004	-0.004	0.008**
	(-1.258)	(-1.893)	(1.037)	(0.849)	(-0.743)	(2.369)
Cash	0.926**	0.625	0.855*	0.006	0.004	0.001
	(2.472)	(1.574)	(1.918)	(0.588)	(0.541)	(0.226)
Leverage	-0.947***	-1.359***	-0.552	-0.026***	-0.019***	-0.010**
	(-2.597)	(-3.156)	(-1.305)	(-4.000)	(-4.448)	(-2.358)
Net PPE	-2.812***	-2.818***	-2.522***	-0.035***	-0.021***	-0.015***
	(-4.951)	(-4.155)	(-3.963)	(-3.590)	(-3.435)	(-2.622)
MtB	0.069**	0.043	0.065	0.003**	0.001	0.001**
	(1.990)	(1.152)	(1.546)	(2.170)	(1.431)	(2.213)
log(Firm Age)	-0.357***	-0.197*	-0.354***	-0.003	-0.001	-0.003**
	(-3.430)	(-1.755)	(-3.091)	(-1.591)	(-0.394)	(-2.234)
N	27,878	25,986	27,358	27,957	27,957	27,957
Adj. R2	0.2066	0.1878	0.1490	0.116	0.1009	0.0411
Industry FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Cluster SE	Yes	Yes	Yes	Yes	Yes	Yes

Note: This table presents results on the determinants of AI disclosure in 10-K filings by estimating following regression at the firm-year level:

$$AI\ Disclosure\ Vars_{i,t+1}/AI\ Intensity\ Vars_{i,t+1} = \beta \cdot AI\ Employee\ Share_{i,t} + \lambda \cdot Other\ Firm\ Chacteristics_{i,t} + D_j + D_t + \epsilon_{i,t+1}$$

In columns (1), (2), and (3), the regression specification is a logistic model, and the dependent variables (*AI Disclosure Vars*) are *AI Disclosure, AI Disclosure-Revenue*, and *AI Disclosure-Cost*. *AI Disclosure* is equal to one if one of the AI keywords is referenced in the 10-K filings and identified as AI-related by ChatGPT. *AI Disclosure-Revenue (Cost)* is equal to one if the AI Disclosure is primarily related to enhancing revenue (cost). In columns (4), (5), and (6), the model specification is a panel regression, and the dependent variables (*AI Disclosure Vars*) are *AI Intensity, AI Intensity-Revenue*, and

AI Intensity-Cost. *AI Intensity* is measured by the number of AI disclosure words processed by GPT, scaled by the total number of words in the corpus extracted from 10-K filings. *AI Intensity-Revenue (Cost)* is the AI Intensity if the AI Disclosure is primarily related to enhancing revenue (cost). All independent variables are lagged by one period. The sample is from 2010 to 2023. Variable definitions are provided in Table A6. The regression coefficients are reported, followed by the robust t-statistics (in parentheses) based on standard errors clustered on the 5-digit NAICS code. *, **, and *** indicate statistical significance at the 10%, 5%, and 1% levels, respectively. The industry-fixed effect is based on the 2-digit NAICS code.

Table 5: AI Disclosure and Firm Growth.

Panel A: The Predictability of AI Disclosure on Firm Growth

Dep. Var.	*log(Employees)*		*log(Sales)*		*CAPX/Assets*		*log(R&D)*	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
AI Disclosure	0.234***	0.139***	0.307***	0.160***	-0.002	0.002***	0.871***	0.464***
	(6.648)	(6.653)	(5.205)	(6.315)	(-1.564)	(3.588)	(12.950)	(11.371)
AI Employee Share	-0.015	-0.042***	-0.003	0.036***	-0.000	0.001**	0.337***	0.156***
	(-0.896)	(-3.850)	(-0.084)	(3.050)	(-1.116)	(2.156)	(8.353)	(8.161)
log(Assets)		0.489***		0.891***		0.000		0.727***
		(76.626)		(142.946)		(0.123)		(56.192)
ROA		-0.565***		1.247***		0.006***		-0.629***
		(-20.951)		(28.563)		(3.674)		(-13.087)
Cash		-0.306***		-1.244***		-0.004***		1.456***
		(-7.163)		(-8.112)		(-2.901)		(22.099)
Leverage		-0.331***		-0.249***		-0.019***		-1.032***
		(-8.891)		(-4.546)		(-11.697)		(-12.419)
Net PPE		-0.007		0.063		0.134***		-1.025***
		(-0.126)		(1.031)		(47.299)		(-9.544)
MtB		0.037***		0.070***		0.002***		0.082***
		(7.347)		(8.166)		(10.344)		(17.168)
log(Firm Age)		0.202***		0.135***		-0.004***		0.002
		(27.520)		(11.972)		(-11.724)		(0.176)
N	28,863	27,747	28,405	27,280	29,130	27,957	15,903	15,206
Adj. R2	0.1665	0.7228	0.1151	0.8687	0.3054	0.5274	0.4077	0.7558
Industry FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Cluster SE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Panel B: The Predictability of AI Disclosure-Revenue and AI Disclosure-Cost on Firm Growth

Dep. Var.	*log(Employees)*			*log(Sales)*			*CAPX/Assets*			*log(R&D)*		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
AI Disclosure-Revenue	0.021		0.014	0.058**		0.051**	0.002**		0.001*	0.502***		0.497***
	(0.896)		(0.598)	(2.279)		(2.028)	(2.013)		(1.901)	(12.096)		(12.029)
AI Disclosure-Cost		0.257***	0.257***		0.252***	0.250***		0.003***	0.003***		0.267***	0.252***
		(8.808)	(8.804)		(7.730)	(7.772)		(3.381)	(3.325)		(5.207)	(4.986)
AI Employee Share	-0.032***	-0.034***	-0.036***	0.046***	0.047***	0.042***	0.001**	0.001***	0.001**	0.154***	0.192***	0.151***
	(-2.926)	(-3.230)	(-3.270)	(3.773)	(3.747)	(3.497)	(2.389)	(2.683)	(2.282)	(7.909)	(9.382)	(7.893)
log(Assets)	0.490***	0.489***	0.489***	0.892***	0.891***	0.891***	0.000	0.000	0.000	0.729***	0.731***	0.728***
	(76.454)	(76.872)	(76.609)	(142.986)	(143.543)	(143.013)	(0.199)	(0.174)	(0.134)	(56.751)	(57.521)	(56.355)
ROA	-0.563***	-0.567***	-0.567***	1.249***	1.244***	1.245***	0.006***	0.006***	0.006***	-0.624***	-0.633***	-0.626***
	(-20.825)	(-21.126)	(-21.120)	(28.416)	(28.442)	(28.476)	(3.696)	(3.666)	(3.672)	(-12.946)	(-13.555)	(-13.008)
Cash	-0.305***	-0.305***	-0.305***	-1.243***	-1.243***	-1.243***	-0.004***	-0.004***	-0.004***	1.457***	1.459***	1.457***
	(-7.037)	(-7.125)	(-7.130)	(-8.065)	(-8.089)	(-8.098)	(-2.895)	(-2.884)	(-2.890)	(21.939)	(21.796)	(22.032)
Leverage	-0.338***	-0.333***	-0.333***	-0.256***	-0.252***	-0.250***	-0.019***	-0.019***	-0.019***	-1.028***	-1.049***	-1.028***
	(-9.009)	(-8.956)	(-8.926)	(-4.660)	(-4.627)	(-4.573)	(-11.749)	(-11.725)	(-11.707)	(-12.422)	(-12.586)	(-12.404)
Net PPE	-0.018	-0.009	-0.008	0.053	0.059	0.062	0.134***	0.134***	0.134***	-1.035***	-1.054***	-1.027***
	(-0.297)	(-0.154)	(-0.143)	(0.861)	(0.973)	(1.012)	(47.316)	(47.346)	(47.301)	(-9.590)	(-9.675)	(-9.556)
MtB	0.038***	0.037***	0.037***	0.070***	0.070***	0.070***	0.002***	0.002***	0.002***	0.083***	0.083***	0.082***
	(7.488)	(7.413)	(7.386)	(8.198)	(8.140)	(8.148)	(10.391)	(10.403)	(10.354)	(17.298)	(17.340)	(17.168)
log(Firm Age)	0.201***	0.203***	0.203***	0.133***	0.135***	0.135***	-0.004***	-0.004***	-0.004***	-0.000	0.001	0.001
	(27.468)	(27.527)	(27.547)	(11.874)	(12.024)	(12.020)	(-11.786)	(-11.730)	(-11.716)	(-0.013)	(0.078)	(0.087)
N	27,747	27,747	27,747	27,280	27,280	27,280	27,957	27,957	27,957	15,206	15,206	15,206
Adj. R2	0.7222	0.7233	0.7233	0.8684	0.8688	0.8688	0.5273	0.5274	0.5274	0.7554	0.7531	0.7558
Industry FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Cluster SE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Panel C: The Predictability of AI Intensity on Firm Growth

Dep. Var.	*log(Employees)*			*log(Sales)*			*CAPX/Assets*			*log(R&D)*		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
AI Intensity	0.131***			0.153***			0.002*			0.607***		
	(4.253)			(3.824)			(1.845)			(9.554)		
AI Intensity-Revenue		0.001			0.034			0.001			0.663***	
		(0.023)			(0.816)			(0.662)			(9.421)	
AI Intensity-Cost			0.280***			0.296***			0.003**			0.395***
			(6.098)			(5.603)			(2.301)			(4.956)
AI Employee Share	-0.036***	-0.030***	-0.032***	0.044***	0.049***	0.048***	0.001**	0.001***	0.001***	0.170***	0.168***	0.193***
	(-3.314)	(-2.781)	(-3.060)	(3.613)	(4.000)	(3.885)	(2.504)	(2.629)	(2.738)	(8.679)	(8.499)	(9.374)
log(Assets)	0.489***	0.490***	0.489***	0.892***	0.892***	0.892***	0.000	0.000	0.000	0.730***	0.731***	0.732***
	(76.641)	(76.594)	(76.808)	(143.205)	(143.247)	(143.539)	(0.202)	(0.233)	(0.214)	(57.061)	(57.383)	(57.723)
ROA	-0.564***	-0.563***	-0.566***	1.248***	1.248***	1.246***	0.006***	0.006***	0.006***	-0.630***	-0.626***	-0.633***
	(-20.858)	(-20.834)	(-20.973)	(28.474)	(28.366)	(28.416)	(3.685)	(3.695)	(3.676)	(-13.158)	(-13.016)	(-13.566)
Cash	-0.306***	-0.305***	-0.305***	-1.243***	-1.242***	-1.243***	-0.004***	-0.004***	-0.004***	1.454***	1.457***	1.457***
	(-7.089)	(-7.026)	(-7.083)	(-8.079)	(-8.057)	(-8.075)	(-2.897)	(-2.891)	(-2.890)	(21.987)	(21.866)	(21.783)
Leverage	-0.335***	-0.339***	-0.336***	-0.254***	-0.257***	-0.255***	-0.019***	-0.019***	-0.019***	-1.037***	-1.034***	-1.049***
	(-8.969)	(-9.030)	(-9.002)	(-4.630)	(-4.692)	(-4.670)	(-11.737)	(-11.769)	(-11.744)	(-12.471)	(-12.441)	(-12.604)
Net PPE	-0.014	-0.019	-0.014	0.055	0.050	0.054	0.134***	0.134***	0.134***	-1.039***	-1.044***	-1.058***
	(-0.237)	(-0.313)	(-0.242)	(0.905)	(0.828)	(0.890)	(47.320)	(47.333)	(47.350)	(-9.590)	(-9.617)	(-9.685)
MtB	0.038***	0.038***	0.037***	0.070***	0.071***	0.070***	0.002***	0.002***	0.002***	0.083***	0.083***	0.083***
	(7.445)	(7.517)	(7.458)	(8.198)	(8.206)	(8.159)	(10.392)	(10.422)	(10.417)	(17.198)	(17.316)	(17.329)
log(Firm Age)	0.202***	0.201***	0.202***	0.134***	0.133***	0.134***	-0.004***	-0.004***	-0.004***	0.000	-0.001	0.001
	(27.495)	(27.448)	(27.491)	(11.897)	(11.872)	(11.949)	(-11.776)	(-11.798)	(-11.769)	(0.006)	(-0.101)	(0.041)
N	27,747	27,747	27,747	27,280	27,280	27,280	27,957	27,957	27,957	15,206	15,206	15,206
Adj. R2	0.7224	0.7222	0.7226	0.8685	0.8684	0.8685	0.5273	0.5273	0.5273	0.7546	0.7542	0.7530
Industry FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Cluster SE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Note: This table presents regression results on predicting the firm growth using AI disclosure measures at the firm-year level. The dependent variables (*Firm Growt*h) are *log(Employees)*, *log(Sales)*, *CAPX/Assets* and *log(R&D)*. *log(Employees)* is the logarithm of the number of employees. *log(Sales)* is the logarithm of the gross sales. *CAPX/Assets* is Capital expenditures scaled by the total assets. *log(R&D)* is the logarithm of the research and development expenditure. Panel A and B report the results estimating the predictability of AI disclosure variables on the firm growth by estimating following regression:

$$Firm\ Growth_{i,t+1} = \beta_1 \cdot AI\ Disclosure\ Vars_{i,t} + \beta_2 \cdot AI\ Employee\ Share_{i,t} + \lambda \cdot Controls_{i,t} + D_j + D_t + \epsilon_{i,t+1}$$

In Panel A*, AI disclosure* is equal to one if the AI keyword is referenced in the 10-K statement and identified as AI-related by ChatGPT. *AI Employee Share* is the fraction of employees classified as AI-related over the total number of employees. In Panel B, *AI Disclosure-Revenue (Cost)* is equal to one if AI Disclosure is primarily related to enhancing revenue (cost). Panel C reports the results of the predictability of AI intensity on the firm growth by estimating following regression:

$$Firm\ Growth_{i,t+1} = \beta_1 \cdot AI\ Intensity\ Vars_{i,t} + \beta_2 \cdot AI\ Employee\ Share_{i,t} + \lambda \cdot Controls_{i,t} + D_j + D_t + \epsilon_{i,t+1}$$

AI intensity is the number of words in AI disclosure words processed by GPT scaled by the total number of words in the corpus extracted from 10-K filings. *AI Intensity-Revenue (Cost)* is the AI Intensity if the AI Disclosure is primarily related to enhancing revenue (cost). All independent variables are lagged by one period. The sample spans from 2010 to 2023. Variable definitions are provided in Table A6. The regression coefficients are reported, followed by the robust t-statistics (in parentheses) based on standard errors clustered by the 5-digit NAICS code. *, **, and *** indicate statistical significance at the 10%, 5%, and 1% levels, respectively. The industry-fixed effect is based on the 2-digit NAICS code.

Table 6: AI Disclosure and Costs

Panel A: The Predictability of AI Disclosure on Corporate Costs

Dep. Var.	*COGS/Sales*		*COGS/Employees*		*Operating Expense/Employees*	
	(1)	(2)	(3)	(4)	(5)	(6)
AI Disclosure	-0.211**	-0.352***	-0.667***	-0.606***	-0.717***	-0.815***
	(-2.481)	(-4.593)	(-5.964)	(-5.725)	(-5.711)	(-7.019)
AI Employee Share	-0.025	-0.389***	-0.112***	-0.081**	0.061	-0.067
	(-0.549)	(-5.367)	(-2.830)	(-2.207)	(1.073)	(-1.195)
log(Assets)		0.138***		0.419***		0.447***
		(4.306)		(9.165)		(9.401)
ROA		-5.809***		-0.807**		-2.543***
		(-8.507)		(-2.362)		(-7.231)
Cash		4.980***		1.708***		3.412***
		(5.464)		(5.133)		(9.349)
Leverage		-0.589***		1.802***		1.611***
		(-3.619)		(4.740)		(4.089)
Net PPE		0.158		0.488		-0.623
		(0.745)		(0.716)		(-0.839)
MtB		-0.182***		-0.196***		-0.169***
		(-2.906)		(-7.913)		(-6.604)
log(Firm Age)		-0.191***		-0.591***		-0.825***
		(-4.316)		(-7.531)		(-10.870)
N	28,408	27,283	28,342	27,255	28,342	27,255
Adj. R2	0.0809	0.2119	0.1512	0.1341	0.1640	0.1618
Industry FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Cluster SE	Yes	Yes	Yes	Yes	Yes	Yes

Panel B: The Predictability of AI Disclosure-Revenue and AI Disclosure-Cost on Corporate Costs

Dep. Var.	*COGS/Sales*			*COGS/Employees*			*Operating Expense/Employees*		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
AI Disclosure-Revenue	-0.388***		-0.381***	-0.466***		-0.448***	-0.564***		-0.538***
	(-3.929)		(-3.871)	(-4.284)		(-4.149)	(-4.368)		(-4.177)
AI Disclosure-Cost		-0.247***	-0.234***		-0.668***	-0.653***		-0.960***	-0.942***
		(-3.807)	(-3.682)		(-5.539)	(-5.450)		(-7.453)	(-7.345)
AI Employee Share	-0.388***	-0.416***	-0.384***	-0.097**	-0.124***	-0.087**	-0.093*	-0.123**	-0.079
	(-5.344)	(-5.576)	(-5.318)	(-2.566)	(-3.315)	(-2.357)	(-1.667)	(-2.186)	(-1.425)
log(Assets)	0.137***	0.136***	0.138***	0.417***	0.417***	0.419***	0.444***	0.444***	0.447***
	(4.286)	(4.267)	(4.305)	(9.120)	(9.123)	(9.161)	(9.329)	(9.348)	(9.394)
ROA	-5.815***	-5.808***	-5.811***	-0.817**	-0.804**	-0.806**	-2.556***	-2.538***	-2.540***
	(-8.511)	(-8.500)	(-8.508)	(-2.392)	(-2.354)	(-2.359)	(-7.269)	(-7.211)	(-7.219)
Cash	4.978***	4.976***	4.979***	1.707***	1.703***	1.705***	3.410***	3.405***	3.408***
	(5.460)	(5.454)	(5.463)	(5.111)	(5.096)	(5.118)	(9.306)	(9.297)	(9.326)
Leverage	-0.584***	-0.575***	-0.589***	1.816***	1.819***	1.803***	1.632***	1.633***	1.614***
	(-3.605)	(-3.561)	(-3.624)	(4.779)	(4.786)	(4.742)	(4.143)	(4.144)	(4.094)
Net PPE	0.168	0.177	0.159	0.516	0.514	0.492	-0.582	-0.591	-0.617
	(0.789)	(0.832)	(0.752)	(0.758)	(0.755)	(0.722)	(-0.787)	(-0.798)	(-0.832)
MtB	-0.182***	-0.183***	-0.182***	-0.198***	-0.197***	-0.196***	-0.172***	-0.171***	-0.169***
	(-2.914)	(-2.918)	(-2.905)	(-7.977)	(-7.985)	(-7.918)	(-6.695)	(-6.668)	(-6.600)
log(Firm Age)	-0.189***	-0.189***	-0.190***	-0.586***	-0.589***	-0.591***	-0.819***	-0.824***	-0.825***
	(-4.280)	(-4.287)	(-4.304)	(-7.491)	(-7.509)	(-7.528)	(-10.806)	(-10.848)	(-10.871)
N	27,283	27,283	27,283	27,255	27,255	27,255	27,255	27,255	27,255
Adj. R2	0.2119	0.2117	0.2119	0.1338	0.1339	0.1341	0.1613	0.1616	0.1618
Industry FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Cluster SE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Panel C: The Predictability of AI Intensity on Corporate Costs

Dep. Var.	*COGS/Sales*			*COGS/Employees*			*Operating Expense/Employees*		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
AI Intensity	-0.471***			-0.810***			-1.018***		
	(-3.296)			(-5.227)			(-5.679)		
AI Intensity-Revenue		-0.522**			-0.736***			-0.850***	
		(-2.569)			(-4.404)			(-3.928)	
AI Intensity-Cost			-0.416***			-0.845***			-1.154***
			(-3.915)			(-4.549)			(-5.950)
AI Employee Share	-0.399***	-0.398***	-0.417***	-0.099***	-0.104***	-0.128***	-0.094*	-0.104*	-0.130**
	(-5.414)	(-5.377)	(-5.579)	(-2.658)	(-2.754)	(-3.390)	(-1.694)	(-1.868)	(-2.292)
log(Assets)	0.137***	0.136***	0.136***	0.417***	0.416***	0.416***	0.444***	0.442***	0.443***
	(4.279)	(4.266)	(4.261)	(9.126)	(9.103)	(9.102)	(9.339)	(9.306)	(9.311)
ROA	-5.811***	-5.815***	-5.809***	-0.811**	-0.818**	-0.808**	-2.549***	-2.557***	-2.544***
	(-8.507)	(-8.508)	(-8.500)	(-2.375)	(-2.394)	(-2.366)	(-7.250)	(-7.270)	(-7.230)
Cash	4.978***	4.977***	4.976***	1.708***	1.706***	1.704***	3.412***	3.409***	3.407***
	(5.459)	(5.456)	(5.453)	(5.117)	(5.105)	(5.096)	(9.311)	(9.295)	(9.286)
Leverage	-0.582***	-0.579***	-0.574***	1.812***	1.819***	1.825***	1.627***	1.637***	1.642***
	(-3.588)	(-3.580)	(-3.556)	(4.772)	(4.789)	(4.805)	(4.132)	(4.156)	(4.170)
Net PPE	0.170	0.175	0.180	0.510	0.523	0.526	-0.591	-0.573	-0.572
	(0.799)	(0.823)	(0.844)	(0.749)	(0.769)	(0.774)	(-0.798)	(-0.775)	(-0.773)
MtB	-0.182***	-0.183***	-0.183***	-0.197***	-0.198***	-0.198***	-0.171***	-0.172***	-0.172***
	(-2.917)	(-2.925)	(-2.919)	(-7.960)	(-7.998)	(-8.006)	(-6.668)	(-6.721)	(-6.704)
log(Firm Age)	-0.189***	-0.188***	-0.188***	-0.587***	-0.585***	-0.587***	-0.820***	-0.817***	-0.820***
	(-4.281)	(-4.263)	(-4.279)	(-7.501)	(-7.480)	(-7.490)	(-10.822)	(-10.792)	(-10.811)
N	27,283	27,283	27,283	27,255	27,255	27,255	27,255	27,255	27,255
Adj. R2	0.2118	0.2118	0.2117	0.1339	0.1338	0.1338	0.1615	0.1613	0.1613
Industry FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Cluster SE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Note: This table presents regression results on predicting the cost of goods sold (COGS) using AI disclosure measures at the firm-year level. The dependent variables are *COGS/Sales, COGS/Employees*, and *Operating Expense/Employees*. *COGS/Sales* is the COGS scaled by total sales. *COGS/Employees* is the COGS scaled by the number of employees. *Operating Expense/Employees* is the operating expenses scaled by the number of employees. Panel A and B report the results estimating the predictability of AI disclosure variables on the COGS by estimating following regression:

$$COGS_{i,t+1} = \beta_1 \cdot AI\ Disclosure\ Vars_{i,t} + \beta_2 \cdot AI\ Employee\ Share_{i,t} + \lambda \cdot Controls_{i,t} + D_j + D_t + \epsilon_{i,t+1}$$

In Panel A*, AI disclosure* is equal to one if the AI keyword is referenced in the 10-K statement and identified as AI-related by ChatGPT. *AI Employee Share* is the fraction of employees classified as AI-related over the total number of employees. In Panel B, *AI Disclosure-Revenue (Cost)* is equal to one if AI Disclosure is primarily related to enhancing revenue (cost). Panel C reports the results of the predictability of AI intensity on the COGS by estimating following regression:

$$COGS_{i,t+1} = \beta_1 \cdot AI\ Intensity\ Vars_{i,t} + \beta_2 \cdot AI\ Employee\ Share_{i,t} + \lambda \cdot Controls_{i,t} + D_j + D_t + \epsilon_{i,t+1}$$

AI intensity is the number of words in AI disclosure words processed by GPT scaled by the total number of words in the corpus extracted from 10-K filings. *AI Intensity-Revenue (Cost)* is the AI Intensity if the AI Disclosure is primarily related to enhancing revenue (cost). All independent variables are lagged by one period. The sample spans from 2010 to 2023. Variable definitions are provided in Table A6. The regression coefficients are reported, followed by the robust t-statistics (in parentheses) based on standard errors clustered by the 5-digit NAICS code. *, **, and *** indicate statistical significance at the 10%, 5%, and 1% levels, respectively. The industry-fixed effect is based on the 2-digit NAICS code.

Table 7: AI Risk Disclosure and Firm Risk

Panel A: AI Risk Disclosure at the Aggregated Level

Dep. Var.	(1) *IVOL*	(2) *SlopeD*	(3) *VRP*	(4) *MFIV*	(4) *MFIS*
AI Risk Disclosure	0.002***	0.156***	0.098***	0.190***	0.039**
	(3.600)	(7.039)	(10.137)	(12.668)	(2.000)
log(Assets)	-0.003***	-0.011***	-0.049***	-0.101***	-0.106***
	(-67.600)	(-3.388)	(-36.913)	(-46.016)	(-57.001)
log(Firm Age)	-0.002***	-0.017***	-0.011***	-0.051***	-0.027***
	(-15.913)	(-3.330)	(-3.075)	(-8.329)	(-7.694)
Tobin's Q	-0.001***	0.014***	-0.026***	-0.048***	-0.068***
	(-20.220)	(5.192)	(-12.130)	(-13.003)	(-27.569)
Leverage	0.006***	0.062*	0.064***	0.191***	0.138***
	(11.005)	(1.766)	(4.396)	(8.238)	(5.853)
Cash	0.008***	-0.301***	0.173***	0.362***	0.279***
	(12.558)	(-8.819)	(4.750)	(5.882)	(14.457)
Intangible/Assets	-0.001***	0.048*	0.003	-0.056***	-0.178***
	(-3.072)	(1.896)	(0.236)	(-3.005)	(-11.002)
Market Beta	0.004***	-0.104***	-0.058***	0.043***	0.029***
	(21.075)	(-8.121)	(-8.364)	(3.708)	(4.178)
Operating Expense	0.001***	-0.065***	-0.051***	-0.060***	-0.029***
	(7.596)	(-7.056)	(-10.402)	(-7.660)	(-5.849)
Intercept	0.041***	0.869***	0.564***	1.365***	0.643***
	(61.444)	(20.795)	(24.233)	(37.955)	(24.639)
N	47,870	47,870	47,870	47,870	47,870
Adj. R2	0.3522	0.0685	0.2176	0.3393	0.2419
Industry FE	Yes	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes	Yes
Clustered SE	Yes	Yes	Yes	Yes	Yes

Panel B: Categorical AI Risk Disclosures

Dep. Var.	(1) *IVOL*	(2) *SlopeD*	(3) *VRP*	(4) *MFIV*	(4) *MFIS*
AI Risk Disclosure - Regulation	0.003***	0.171***	0.140***	0.243***	0.070**
	(3.112)	(4.708)	(8.898)	(8.932)	(2.343)
AI Risk Disclosure - Operation	0.002***	0.147***	0.098***	0.200***	0.042*
	(3.932)	(5.433)	(8.876)	(11.700)	(1.843)
AI Risk Disclosure - Competition	0.001	0.191***	0.095***	0.191***	-0.033
	(1.518)	(5.035)	(7.295)	(9.189)	(-1.155)
AI Risk Disclosure - Cybersecurity	0.002**	0.115***	0.069***	0.162***	0.046
	(2.544)	(3.386)	(4.711)	(7.408)	(1.544)
AI Risk Disclosure - Ethical	0.003***	0.174***	0.118***	0.240***	0.068*
	(2.895)	(4.374)	(7.221)	(7.419)	(1.946)
AI Risk Disclosure - Third Party	0.003**	0.052	0.099***	0.216***	0.062*
	(2.054)	(1.106)	(3.495)	(5.524)	(1.750)
N	47,870	47,870	47,870	47,870	47,870
Controls	Yes	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes	Yes
Clustered SE	Yes	Yes	Yes	Yes	Yes

Note: Panel A of this table presents regression results on predicting the firm stock and option-implied risk using AI risk disclosure measures at the firm-year level. The dependent variables are the stock idiosyncratic volatility (*IVOL*), the slopeness of the out-the-money put options' implied volatility on moneyness (*SlopeD*), the variance risk premium defined as the difference the implied variance from the 30-day options and realized variance of daily stock returns over the same period (*VRP*), the model-free implied volatility for options with a 30-day maturity, computed as the second central moment of the risk-neutral distribution (*MFIV*), and the model-free implied skewness for options with a 30-day maturity, computed as the third central moment of the risk-neutral distribution (*MFIS*). Panel A report the results estimating the predictability of AI risk disclosure variables on the firm risk variables by estimating following regression:

$$Firm\ Risk_{i,t+1} = \beta_1 \cdot AI\ Risk\ Disclosure_{i,t} + \lambda \cdot Risk\ Controls_{i,t} + D_j + D_t + \epsilon_{i,t+1}$$

Panel B report the results estimating the predictability of AI risk disclosure variables on the firm risk variables by estimating following regression, where $AI\ Risk\ Disclosure^n_{i,t}$ is the AI risk disclosure for the *n*th category. Each line represents a standalone regression using the categorical AI risk disclosure as the main independent variable.

$$Firm\ Risk_{i,t+1} = \beta_1 \cdot AI\ Risk\ Disclosure^n_{i,t} + \lambda \cdot Risk\ Controls_{i,t} + D_j + D_t + \epsilon_{i,t+1}$$

Appendix

Table A1: Examples of AI-related disclosure in 10-K statements.

International Business Machines (IBM), 2022, 10-K Item 1 Business

This was signaled almost 100 years ago, in 1924, when C-T-R changed its name to International Business Machines Corporation. And it continues today we create sustained value for clients by helping them leverage the power of hybrid cloud and **artificial intelligence** (AI). Our hybrid cloud platform and AI technology support clients' digital transformations and helps them reimagine critical workflows, at scale, and modernize applications to increase agility, drive innovation and create operational efficiencies. Our offerings draw from leading IBM capabilities in software, consulting services capability to deliver business outcomes, and deep incumbency in mission-critical infrastructure, all bolstered by one of the world's leading research organizations.

Blackberry, Ltd., 2022, 10-K Item 1A Risk Factors

The process of developing new technology is complex and uncertain, and involves time, substantial costs and risks, which are further magnified when the development process involves multiple operating platforms. The development of next-generation technologies that utilize new and advanced features, including **artificial intelligence** and machine learning, involves making predictions regarding the willingness of the market to adopt such technologies over legacy solutions. The Company may be required to commit significant resources to developing new products, software and services before knowing whether such investment will result in products or services that the market will accept.

Codexis, Inc., 2018, 10-K Item 7 Management Discussion and Analysis

Over the last fifteen years, we have made substantial investments in the development of our CodeEvolver protein engineering technology platform, the primary source of our competitive advantage. Our technology platform is powered by proprietary, **artificial intelligence**-based, computational algorithms that rapidly mine our large and continuously growing library of protein variants performance attributes. These computational outputs enable increasingly reliable predictions for next generation protein variants to be engineered, enabling delivery of targeted performance enhancements in a time-efficient manner.

Table A2: Keywords of AI Disclosure in the 10-K Filings

Keyword	(%) All 10-K Filings	(%) AI Disclosing Transcripts
artificial intelligence	5.91%	40.42%
machine learning	5.13%	35.09%
business intelligence	4.29%	29.35%
big data	2.40%	16.45%
data science	1.84%	12.57%
AI/ML	1.48%	10.10%
data mining	0.98%	6.70%
data scientist	0.93%	6.34%
deep learning	0.49%	3.35%
computer vision	0.45%	3.08%
natural language processing	0.40%	2.75%
chatbot	0.34%	2.33%
image recognition	0.18%	1.21%
object recognition	0.06%	0.39%
large language model	0.04%	0.31%
machine translation	0.04%	0.25%
support vector machine	0.03%	0.23%
classification algorithm	0.03%	0.23%
generative ai	0.03%	0.19%
A.I.	0.02%	0.16%
neutral network	0.02%	0.12%
supervised learning	0.01%	0.07%
computational linguistic	0.01%	0.07%
clustering algorithms	0.01%	0.05%
recommender system	0.01%	0.04%
dimensionality reduction	0.00%	0.03%
information extraction	0.00%	0.03%
kernel method	0.00%	0.01%
unsupervised learning	0.00%	0.01%

Note: The table lists all keywords in the textual analysis for identifying the AI disclosure in the corporate annual report in the first column. The second column summarizes the fraction of occurrence of the corresponding keyword in all 10-K filings in our sample. The second column summarizes the fraction of occurrence of the corresponding keyword in transcripts with AI disclosure. The sample contains U.S. public firms spanning from 2010 to 2023.

Table A3: Most frequently mentioned AI keywords.

Year	Top #1 Frequent	Top #2 Frequent	Top #3 Frequent
2010	business intelligence	AI/ML	data mining
2011	business intelligence	AI/ML	data mining
2012	business intelligence	AI/ML	data mining
2013	business intelligence	AI/ML	data mining
2014	business intelligence	big data	AI/ML
2015	business intelligence	big data	AI/ML
2016	business intelligence	big data	AI/ML
2017	business intelligence	big data	machine learning
2018	artificial intelligence	machine learning	business intelligence
2019	artificial intelligence	machine learning	business intelligence
2020	artificial intelligence	machine learning	business intelligence
2021	artificial intelligence	machine learning	business intelligence
2022	artificial intelligence	machine learning	data science
2023	artificial intelligence	machine learning	data science

Note: The table lists the top 3 keywords with the highest frequency being mentioned in the corporate 10-K report each year. The sample contains U.S. public firms spanning from 2010 to 2023.

Table A4: GPT prompt for the AI Disclosure classification and categorization

Objective:

As an expert financial analyst, review the company's annual 10-K financial statement provided below. Your task is to assess the relevance of the disclosure to artificial intelligence (AI). Determine whether the disclosure is "Related to AI" or "Not Related to AI" and provide a probability score between 0 and 1 indicating the likelihood of the disclosure being related to AI.

Instructions:

Classification:

Label the disclosure as either "Related to AI" or "Not Related to AI".

Probability Score:

Assign a score from 0 to 1 estimating the likelihood of the disclosure being related to AI. A score of 0 means no relation, while 1 indicates a high certainty of relation to AI. The probability score should reflect how clearly the disclosure outlines the company's AI usage.

General Assessment:

If classified as "Not Related to AI", provide reasons supporting this classification.

If classified as "Related to AI", proceed with the following steps:

Application Categorization: Categorize the AI usage into one or more of the following categories based on the description. :

Product Development

Inventory Management

Operational Efficiency

Pricing Optimization

AI Product Provider

Detailed Reasons: For each category selected, explain how the disclosure aligns with these specific AI applications. Don't infer the application, just show the categories that are explicitly mentioned.

Probability Scores for Subcategories: Provide a score from 0 to 1 for each selected subcategory indicating the likelihood that the disclosure falls into that category.

Definitions:

Related to AI:

This encompasses any mention or implication of AI technologies, such as machine learning, neural networks, deep learning, natural language processing, or AI-driven analytics, and their application in the company's operations, products, or services.

Example: "The company integrates machine learning algorithms to enhance customer personalization on its platforms."

Product Development:

The use of AI in the research, creation, design, or enhancement of the company's products or services. This includes employing AI for R&D, product testing, or accelerating the innovation process.

Example: "AI is utilized in the R&D phase to simulate product performance under various conditions, improving design and reducing time-to-market."

Inventory Management:

AI applications focused on optimizing inventory processes, including managing stock levels, predicting future inventory needs, or streamlining supply chain operations.

Example: "Implementing AI-driven demand forecasting to automate inventory replenishment and reduce overstock."

Operational Efficiency:

Leveraging AI to improve internal processes and overall productivity. This includes automating routine tasks, predictive maintenance, optimizing workflows, and enhancing day-to-day operational activities.

Example: "Using AI to automate customer service interactions, reducing response times and operational costs."

Pricing Optimization:

Utilizing AI to develop and implement strategic pricing models. This involves dynamic pricing, competitive pricing analysis, and AI-driven algorithms to adjust prices in real-time to maximize revenue or market share.

Example: "AI algorithms adjust product pricing dynamically based on real-time market data to maintain competitive advantage."

AI Product Provider:

The company's role as a provider of AI-based products or services. This includes developing, selling, or licensing AI technologies, platforms, or solutions to other businesses or consumers.

Example: "The company offers an AI-powered analytics platform that enables clients to gain actionable insights from their data."

Output in JSON Response Format:

If Not Related to AI:

```
{
  "classification": "Not Related to AI",
  "probability_score": [score],
  "reasons": "[explanation]"
}
```

If Related to AI:

```
{
  "classification": "Related to AI",
  "probability_score": [score],
  "summary": "[general application summary]",
  "applications": [
    {
      "category": "[Category Name]",
      "probability_score": [score],
      "reasons_of_application": "[explanation for this categorization]"
    }
    // Additional categories as applicable
  ]
}
```

Table A5: ChatGPT prompt for AI intensity

System prompt

As an expert financial analyst, review the texts from company's annual 10-K financial statement provided below. Your task is to remove contents that are not related to artificial intelligence (AI). When editing, assume that sentences between two segments related to AI also part of this category and should not be removed. Edits should only affect the beginning or end of the provided text. Output back text related to AI. Do not change the text.

User prompt (Example of feed):

strategies have helped CareAdvantage clients to more effectively: Identify and quantify disease burden and associated risk with their entire population and sub-populations; Improve member care quality through the defensible evaluation of health care providers and facilities; Facilitate provider cooperation and collaboration based on case mix and severity-adjusted data; Forecast resource consumption based on disease burden; and Optimize allocation of resources. Recently, the Company has also assisted health plan clients with respect to further validating the value that they bring to purchasers and have had a demonstrable impact on sales retention and attraction of new sales. The RPNavigator suite of services that the Company offers our state government customers provides opportunities in transparency, improving quality, controlling costs and providing information to consumers and healthcare policymakers to make informed decisions. In addition, this expertise has been applied to help employers assess the efficiency and effectiveness of their present health care insurers, carriers and supporting vendors. CareAdvantage achieves this by empowering employers with the information, skills and guidance necessary to facilitate future purchasing decisions and optimize managerial and administrative practices. In order to deliver these solutions, CareAdvantage utilizes experienced health plan executives and medical directors as well as a wide range of care management operations, clinical data analysis and information technology (IT) subject-matter experts. It is this vast array of experience that enables CareAdvantage to benefit clients with objective and quantifiable insight to develop the strategies and tactical initiatives that combine care management processes with a deep understanding of medical and health care insurance-related best practices. Operations The Company utilizes a multi-disciplinary team approach in providing its management, data analysis and consulting services. The Company, through its employees and independent contractors, assesses care management operations, systems resources, integration and outcomes. Typically, assessment occurs on the client s site, through interviews and data analysis. At the center of CareAdvantage s data-driven analyses is RPNavigator, the next generation software solution. RPNavigator categorizes and quantifies a population s disease burden and provides a clear picture of the health status and severity associated with its clients member populations. RPNavigator s underlying infrastructure incorporates classification methodologies from 3M Health Information Systems along with various analytical techniques to stratify the population and describe the individual member s associated risks in intuitive ways. It also enables the valid assessment of existing health care quality and cost as well

as projection of future risk from a resource consumption, disease progression and mortality perspective. RPNavigator includes a data mining tool (RPN 3 that employs multi-dimensional cubes (data structures) for online analytic processing (OLAP). RPN 3 references the same data set within RPNavigator and allows power users additional flexibility in querying that data. RPNavigator utilizes this information to stratify its clients members, groups and providers through the use of a wide range of clinical and demographic descriptors to quantify their risk as well as evaluate the impact of key interventions and programs. These descriptors and the underlying logic increase the associated transparency of the resulting analyses and support the new direction of the industry. Among the benefits of this solution is the ability of CareAdvantage clients to: Access meaningful information via an Internet-based portal; Track population and member-related changes in disease status and severity over time; Compare client sub-populations; Profile provider using case mix and severity-adjusted techniques; Select and prioritize members who would best benefit from care management interventions; 4 Understand adverse selection associated with existing and/or newly-obtained business as well as understand the impact of a plan s overall turnover in terms of stayers and leavers; and Reduce the dependence on internal resources to develop and produce required reports to accomplish these tasks. There are several related efforts that result in additional sources of income for CareAdvantage, including: Developing new standardized analyses on a client-specific basis to meet a particular need for that client; Undertaking broader analytic consulting projects, using the methodology and logic within RPNavigator, on behalf of clients that need CareAdvantage s expertise in analyzing and interpreting the data; Offering experienced health care executives for care management program leadership, internal physician review services, and mentoring of less experienced health plan staff Executive and Clinical Management Services). CareAdvantage also supports its clients through the provision of a wide range of consulting services to develop and implement the right solutions. These solutions provide the health care industry with strategies and tactical initiatives for effectively managing health care consumption, reducing costs and improving the quality and cost benefit of care. For its services, the Company seeks to be compensated either (i) on a fee-for-service basis; (ii) per member per month (PMPM); or (iii) on the basis of a combination of both fee-for-service and PMPM.

Assistant prompt (Expected response):

The Company, through its employees and independent contractors, assesses care management operations, systems resources, integration and outcomes. Typically, assessment occurs on the client s site, through interviews and data analysis. At the center of CareAdvantage s data-driven analyses is RPNavigator, the next generation software solution. RPNavigator categorizes and quantifies a population s disease burden and provides a clear picture of the health status and severity associated with its clients member populations. RPNavigator s underlying infrastructure incorporates classification methodologies from 3M Health Information Systems along with various analytical techniques to stratify the population and describe the individual member s associated risks in intuitive ways. It also enables the valid assessment of existing health care quality and cost as well as projection of future risk from a resource consumption, disease progression and mortality

perspective. RPNavigator includes a data mining tool (RPN 3 that employs multi-dimensional cubes (data structures) for online analytic processing (OLAP). RPN 3 references the same data set within RPNavigator and allows power users additional flexibility in querying that data. RPNavigator utilizes this information to stratify its clients members, groups and providers through the use of a wide range of clinical and demographic descriptors to quantify their risk as well as evaluate the impact of key interventions and programs. These descriptors and the underlying logic increase the associated transparency of the resulting analyses and support the new direction of the industry. Among the benefits of this solution is the ability of CareAdvantage clients to: Access meaningful information via an Internet-based portal; Track population and member-related changes in disease status and severity over time; Compare client sub-populations; Profile provider using case mix and severity-adjusted techniques; Select and prioritize members who would best benefit from care management interventions; Understand adverse selection associated with existing and/or newly-obtained business as well as understand the impact of a plan s overall turnover in terms of stayers and leavers; and Reduce the dependence on internal resources to develop and produce required reports to accomplish these tasks.

Table A6: Variable Definitions.

Variable	Definition
AI Disclosure	An indicator with value of one when there is AI-related disclosure in the item 1 and item 7 sections from the 10-K statement and zero otherwise. The AI disclosure is measured by at least one keyword and processed by the ChatGPT.
AI Disclosure-Revenue	An indicator with value of one when AI disclosure in the 10-K statement is related to revenue enhancement.
AI Disclosure-Cost	An indicator with value of one when AI-related disclosure in the 10-K statement is related to cost enhancement.
AI Intensity	The number of words in AI disclosure words processed by GPT scaled by the total number of words in the corpus extracted from 10-K filings.
AI Intensity-Revenue	AI intensity in the AI Revenue disclosure sample.
AI Intensity-Cost	AI intensity in the AI Cost disclosure sample.
AI Risk Disclosure	An indicator with value of one when there is AI-related risk factor disclosure in the item 1A section from the 10-K statement and zero otherwise. The AI risk disclosure is measured by at least one keyword and processed by the ChatGPT.
AI Employee Share	Following Babina et al. (2024), the fraction of employees classified as AI-related over the total number of employees, scaled by 100.
log(Employees)	The logarithm of the number of employees (EMP on Compustat) in thousands in the fiscal year.
log(Sales)	The logarithm of the gross sales level (SALE on Compustat) in millions in the fiscal year.
CAPX/Assets	Capital expenditures (CAPX on Compustat) scaled by the beginning period of total assets (AT on Compustat) in the fiscal year.
log(R&D)	The logarithm of the research and development expenditure (XRD on Compustat) in millions in the fiscal year.
COGS/Sales	The COGS (COGS on Compustat) scaled by total sales (SALE on Compustat).
COGS/Employees	The COGS (COGS on Compustat) scaled by the number of employees (EMP on Compustat).
Operating Expense/Employees	The operating expenses (XOPR on Compustat) scaled by the number of employees (EMP on Compustat).
IVOL	
SlopeD	The slopeness of the function that relates implied volatility to the Black-Scholes delta for OTM put options (with deltas between -0.5 and -0.1) with a 30-day maturity. The annual measure is constructed by averaging daily values.
VRP	Variance risk premium is defined as the daily difference between the implied variance of returns from t to t+30 calendar days and realized variance of daily returns over the same period (t, t+30). The annual measure is constructed by averaging daily values.
MFIV	Model-free implied volatility for options with a 30-day maturity, calculated as the second central moment of the risk-neutral distribution.

	We follow Bakshi, Kapadia, and Madan (2003) to construct the variable. The variable is constructed at the annual level by taking the average of daily values.
MFIS	Model-free implied skewness for options with a 30-day maturity, calculated as the third central moment of the risk-neutral distribution, normalized by the risk-neutral variance (raised to the power of 3/2). We follow Bakshi, Kapadia, and Madan (2003) to construct the variable. The variable is constructed at the annual level by taking the average of daily values.
log(Assets)	The logarithm of total assets (AT on Compustat) in millions.
ROA	The ratio of net income (NI on Compustat) over the total assets (AT on Compustat) in the fiscal year.
Cash	Cash holdings (CHE on Compustat), scaled by total assets (AT on Compustat).
Leverage	The book value of debt (DLC + DLTT on Compustat) divided by total assets (AT on Compustat).
Net PPE	Net Property, plant, and equipment (PPENT on Compustat), scaled by total assets (AT on Compustat).
MtB	The ratio of the equity market value over the difference between common equity and preferred stock capital at fiscal year-end.
log(Firm Age)	The logarithm of the years between the firm's first year of data in Compustat and the current fiscal year.
Tobin's Q	(Total assets (AT on Compustat) - total common equity (CEQ on Compustat) + share price (PRCC on Compustat) × common shares outstanding (CSHO on Compustat)) / total assets (AT on Compustat).
Intan/Assets	The intangible assets (INTAN on Compustat) scaled by the beginning of the period total assets (AT on Compustat).
β_{MKT}	Sensitivity of daily stock returns to daily S&P returns. We run daily regressions of excess returns on a constant and the market factor for each firm and year. For each firm-by-year combination, β_{MKT} (market beta) corresponds to the estimated regression coefficient.
XOPR/Assets	The operating expense (XOPR on Compustat) scaled by the beginning of the period total assets (AT on Compustat).

Note: The table reports the definitions of all variables.