universität wien

MASTERARBEIT | MASTER'S THESIS

Titel | Title

Measuring the Actual Adoption of AI and Its Determinants: A Refined Textual Analysis Approach to Overcome Limitations in Large-Scale Empirical Analyses

verfasst von | submitted by

Frederik Ivo Stolba B.Sc.

angestrebter akademischer Grad | in partial fulfilment of the requirements for the degree of

Master of Science (MSc)

Wien | Vienna, 2025

Studienkennzahl lt. Studienblatt | Degree programme code as it appears on the student record sheet:	UA 066 915
Studienrichtung lt. Studienblatt | Degree programme as it appears on the student record sheet:	Masterstudium Betriebswirtschaft
Betreut von | Supervisor:	Univ.-Prof. Dipl.-Chem. Dr. Markus Georg Reitzig MBR

Abstract

Recent advancements in artificial intelligence (AI) have accelerated the discourse around the topic in both academia and the economy. The increased applicability and business value of the technology have raised awareness of its capabilities, leading more companies to adopt AI into their operations. The impact of automation, augmentation, displacement, or new opportunities through AI is not only reflected in potential cost savings or increased revenue, but also holds severe implications for workforce composition and potential unemployment. To continue releasing relevant insights on how AI will determine our future, academia must keep up with its current fast-paced development. However, limitations in data availability and existing empirical methods call for additional approaches to capturing a company's adoption of artificial intelligence. This thesis aims to provide a new textual analysis-based measure that captures a firm's actual adoption of artificial intelligence, as well as its determinants. Building on prior advancements and findings of this research method, the measurement considers the word counts of a more narrowly defined AI dictionary within the context of adoption. Through the added sophistication, I established significant differences from existing measures and was able to demonstrate an improved relationship between investment effects and this adoption-centered AI variable. Ultimately, this thesis contributes to the existing body of empirical research on AI by providing a replicable measure of actual AI Adoption, which can be employed to produce large-scale panel data.

Table of contents

1 Introduction

Artificial intelligence (AI) has been an ever-evolving topic area for decades. Continuous advancements in supporting technologies, such as processing power, big data, and cloud computing, contributed to the increasing maturity of AI applications (Goldfarb et al., 2023; Zhang & Lu, 2021). During the last decade, AI has become increasingly applicable for a wider range of business scenarios, even for non-tech companies (Bass, 2018). Recent advancements in natural language processing and generative AI and their respective announcements not only accelerated the discourse around the topic but also shifted the focus of firms towards the adoption of AI. With potential competitive advantages through cost reductions, enhanced efficiency, or access to new business fields, AI is estimated to generate multiple trillion dollars in value across different industries (McKinsey & Company, 2023a; Stanford University, 2024). Possible use cases of AI within a business context are widespread, with applications across the value chain. Examples include AI benefitting businesses through augmentation or substitution in decision making processes, being capable of evaluating business models and strategic decisions (Balasubramanian et al., 2022; Choudhary et al., 2025; Di Vaio et al., 2020; Doshi et al., 2024), maximizing value in an automated purchasing process enriched by a smart recommendation system (Cui et al., 2022), refining the understanding of consumer behaviors and assisting in retail (Guha et al., 2021; Syam & Sharma, 2018), or by increasingly augmenting and ultimately replacing human labor in services (M.-H. Huang & Rust, 2018). For more detailed insights, Loureiro et al. (2021) provide a structured overview of existing studies covering the multifaceted business applications of artificial intelligence. In conjunction with this broad range of business applications, Goldfarb et al. (2023) found support for the assumption that AI qualifies as the next general-purpose technology (GPT), falling in line with past innovations such as the steam engine, electric power, internet, and computers (Lipsey et al., 2005). While AI can create new tasks for humans, in the past, its focus was more centered around the advancements in automation, resulting in decreased labor demand, productivity growth, and higher inequality (Acemoglu & Restrepo, 2020). This inequality is also reflected in the increasing wage gaps that result from the impact of AI, yielding higher wages for high-wage occupations and lower wages for low-wage occupations (Felten et al., 2019). These opportunities and potential threats arising from AI underscore the urgent need for further quantitative studies that provide empirical evidence of its various

implications. However, the current body of empirical literature on AI suffers from several limitations that can predominantly be traced back to its underlying data sources. One of these limitations is the scarce availability of large data volumes over an extended time horizon. The textual analysis of Form 10-K reports, which publicly listed US companies must submit annually to the Securities and Exchange Commission (SEC), offers a promising solution to this problem. The relatively large and multi-subject text corpus has offered researchers the opportunity to analyze various business-relevant topics in prior studies, including the focus and use of technologies in the company (Chen & Srinivasan, 2024; Mishra et al., 2022). However, a comprehensive data source alone is not enough. To be able to draw conclusions about AI activities in companies, a more nuanced metric is required that identifies only those terms from the large pool of words that represent the actual adoption of AI.

This thesis aims to fill this gap. Over the next chapters, I will first discuss AI in more detail, explain the research method of textual analysis in the field of economic literature, and provide an overview of existing empirical studies on AI. Building on the insights of prior literature, I will proceed to describe the approach of developing an AI Adoption measurement that uses advanced semantics to identify and process relevant AI-related terms from the 10-K reports of the years 2017 to 2023. Through a series of tests, I aim to demonstrate that this newly developed measure constitutes a novelty and a contribution to the existing approaches used in previous studies and enables researchers to make inferences about a firm's actual use of AI.

2 About artificial intelligence

Prior literature produced various definitions for artificial intelligence. Some approach the subject in a broader, less technical way, while others provide a more detailed differentiation. As a continually evolving technology, it is therefore only natural that the definition of AI evolves too. In a proposal for the Dartmouth summer research project on artificial intelligence back in 1955, the subject was approached under the conjecture that "every aspect of learning or any other feature of intelligence can in principle be so precisely described that a machine can be made to simulate it." (McCarthy et al., 2006, p. 12). This raises the question for the meaning of intelligence. I follow Kaplan and Haenlein (2019) in referring to Gardner (1999), who defined intelligence as the "biopsychological potential to process information . . . to solve problems or create products that are of value in a culture" (Gardner, 1999, pp. 33–34). Zhang and Lu (2021) provide detailed information on how the research field of artificial intelligence has developed over the last decades since its founding period. In the contemporary age of artificial intelligence, Kaplan and Haenlein (2019) defined AI "as a system's ability to interpret external data correctly, to learn from such data, and to use those learnings to achieve specific goals and tasks through flexible adaptation". They further outline the three evolutionary stages of artificial intelligence: first, "Narrow AI", which can match or beat human competence within the limits of certain areas (Kaplan & Haenlein, 2019). As this form of AI relies on learning from data inputs to enhance its capabilities, a widely used alternative labeling for narrow AI is "Machine Learning" (Felten et al., 2021). Second, according to Kaplan and Haenlein (2019), "General AI" has the ability to solve problems autonomously while matching or beating human competence in a wider range of areas. And third, "Super AI" surpasses human competence in all areas and can quickly solve problems in other areas. Additionally, the authors provide a "classification of AI into analytical AI, human-inspired AI, and humanized AI" (Kaplan & Haenlein, 2019, p. 19), which is bound to the AI's capabilities in different types of intelligence. According to this classification, analytical AI is limited to cognitive intelligence, human-inspired AI is capable of analytical and emotional intelligence, while humanized AI covers cognitive, emotional, and social intelligence. These AI systems are being trained with preexisting data along three different learning processes: "supervised learning, unsupervised learning, and reinforcement learning" (Kaplan & Haenlein, 2019, p. 19). These training inputs often stem from a vast range of web pages, articles, news

sites, books and other text corpora and when combined with the billions of parameters in a neural network form a so called large language model (LLM), like ChatGPT from OpenAI or Gemini from Google (Brown et al., 2020; Chang et al., 2023). A human is able to interact with the LLM and provide necessary input by entering a so-called “prompt”, which can yield a certain output (Liu et al., 2023).

In the same way, it is essential to establish a common understanding of what AI actually resembles, I argue for the necessity to delineate its meaning from other complementary or supporting technologies to refine the definition. Zhang and Lu (2021) describe ‘Big Data’ and the processing capabilities provided by graphic processing unit (GPU) hardware as enabling technologies, since the development of AI highly relies on large amounts of multi-dimensional data, which can be parallelly processed on efficient GPU cores. As mentioned above, Kaplan and Haenlein (2019) differentiate AI from ordinary expert systems. These expert systems would neither have capabilities related to cognitive, emotional, or social intelligence nor to artistic creativity. Instead, the authors describe expert systems as “collections of rules programmed by humans in form of if-then statements” (Kaplan & Haenlein, 2019, p. 18), missing the autonomous learning component that would classify it as AI. This delineation appears less obvious, since an expert system is designed “to simulate the thinking process of human experts and uses knowledge and reasoning to solve complex problems that only domain experts can solve” (Zhang & Lu, 2021, p. 4). This is similar to the field of robotics, where the first two generations of robots perform tasks defined in a program and might be additionally equipped with sensors that collect information about the working environment, enabling the robot to adapt. Only the third generation of robots is able to intelligently sense and process information obtained by its highly sensitive sensors, resulting in an analytical, self-controlled, and reactive machine (Zhang & Lu, 2021). Therefore, one might argue whether robotics in general should fall under the definition of artificial intelligence, or if even an intelligent robot would rather be the product of delicate engineering combined with artificial intelligence capabilities. This consideration aligns with the argumentation of Davenport et al. (2020), who describe AI within robots as the physical embodiment of AI through which artificial intelligence transitions from pure virtuality to a real form.

3 Textual analysis in economic literature

Textual analysis has been utilized across multiple economic studies for various purposes. This section aims to provide an overview of the variety of different research topics and how different applications of textual analysis have been able to produce empirical insights. I will touch upon the key elements to consider and conclude with an argument for the method's credibility.

3.1 Text corpora

Prior studies have accessed a variety of different sources to analyze economic topics and relationships, whereby the choice of text body naturally depends on the subject of the research. To link the attributes of merger and acquisition (M&A) deals to market and firm performance measures, Hu et al. (2021) used transcripts from M&A conference calls to obtain information on the presence of the acquiring firms' executives during the call, the tone of the language used, the percentage of numbers in the text as well as the frequency of financial vs. strategic terms, enabling the authors to predict market reactions. Hanley and Hoberg (2010) decompose the text of initial public offering (IPO) prospectuses into standard and informative content to analyze for relationships with price adjustments and initial returns. By comparing the different sections of the IPO prospectus (Summary, Risk Factors, Use of Proceeds, Management Discussion and Analysis (MD&A)) to each other and the entire document, the authors highlight differences in economic significance of the relationships. Agarwal et al. (2016) demonstrate the content value of credit rating action reports, which appear to provide a useful source of information to predict abnormal returns and rating changes, based on the linguistic tone of the disclosures. Another often analyzed text corpus stems from earnings announcement press releases, which companies can disclose voluntarily for firm-investor communication without underlying any formal requirements (Henry, 2008). Hence, this type of text corpus provides several research opportunities, but also limitations. Analyzing the sentiment of earnings conference calls linked to data about a firm's ownership through institutions enabled Amoozegar et al. (2020) to demonstrate how institutional ownership affects the overall tone of the conference call, and that short-term institutional investors favor a more optimistic tone as compared to long-term institutional investors. Engelberg (2008) employed earnings announcement article texts to distinguish between easy-to-process information and qualitative, processing-intensive information contents about the predictability of asset prices and long-term returns.

Findings from Henry (2008) suggest that the market reacts positively to a more positive tone of an earnings press release, even if the actual financial results are controlled for. Even though additional results from Huang et al. (2014) have shown a delayed negative reaction, the overall literature suggests that managers use disclosure tone to intentionally mislead investors, making earnings press releases a questionable text corpus for certain textual analysis applications.

Another widely used text corpus is the Form 10-K report, which has to be filed annually by most public US companies at the Securities and Exchange Commission (SEC). Besides financial figures, the SEC requires qualitative disclosure of the filing companies' business activities and matters divided into several items, such as a business description, risk factors, legal proceedings, or MD&A (Securities and Exchange Commission, 2025). The area of finance and accounting research has produced numerous studies examining the tone and readability of the entire report text linked to financial performance and stock price trends (Li, 2008; Li & Zhang, 2015; Loughran & McDonald, 2014). In addition, numerous other studies within and outside the financial research field use Form 10-K reports to analyze a wide variety of topics. In their study about the focus of firms on AI, Mishra et al. (2022) consider the full body of the Form 10-K report to look for relevant keywords. On the contrary, other authors extract only a specific text section that matches the underlying topic, e.g. looking for disclosures of cybersecurity risk in 'Item 1A. Risk Factors' of the 10-K report (Cheong et al., 2021), predicting bankruptcy by analyzing managerial statements for sentiment and the specific mention of "going concern" in 'Item 7. Management's Discussion and Analysis of Financial Condition and Results of Operations' (Mayew et al., 2015), or determining value and performance implications from digital activities of firms by analyzing 'Item 1. Business Description' for occurrences of predefined digital terms (Chen & Srinivasan, 2024).

3.2 Linguistic tone measurements

As mentioned above, many studies from the domain of financial and accounting research employed sentiment analysis to determine the impacts of disclosure tone. To derive the sentiment, there are essentially two options: word-frequency tone measures and machine learning tone measures (Henry & Leone, 2016). Word-frequency tone measures aim to capture the sentiment of a text corpus by counting positive or negative words that have been previously defined in word lists, or 'dictionaries'. A popular

word list that has been used in the earlier days of textual analysis in accounting literature is the Harvard IV-4 psychological dictionary within the General Inquirer (GI) tool, originally developed to analyze psychological themes in group discussions (Stone et al., 1962). Tetlock (2007) employed GI's Harvard IV-4 list of negative words (H4N) to measure negativity in daily stock market media articles of the Wall Street Journal (WSJ) column, revealing high pessimism to force downward pressure on stock prices and trading volume to increase at high and low pessimism rates. Additionally, Tetlock et al. (2008) found that smaller earnings of firms can be predicted by the fraction of negative words in WSJ news stories. The prediction quality was found to be higher the more news stories focused their reporting on actual earnings. Engelberg (2008) followed this approach to construct a measurement of the fraction of negative words as part of his analysis of news stories about earnings announcements, indicating that negative press reporting diminishes the returns after the earnings announcement.

Investigating whether a psychological dictionary is actually capable of capturing the subtleties of sentiment in a financial context, Loughran and McDonald (2011) found that 73.8% of the words in H4N are typically not associated with negativity in financial statements. Moreover, they found crucial sentiment-determining financial terms to be missing in the Harvard dictionary. After developing "five other word classifications (positive, uncertainty, litigious, strong modal, and weak modal words)" (Loughran & McDonald, 2011, p. 62) tailored to financial texts, a large sample analysis of Form 10-K reports demonstrated vastly improved explanatory power of negative sentiment when predicting excess return during the filing period. The authors further propose a weighting logic to lessen the impact of frequently used terms, aiming to reduce misclassification bias. Weighting appeared to enhance the explanatory quality of both the H4N and the negative Loughran and McDonald (LM) dictionary. Recognizing this advancement in financial disclosure tone measurement, other studies employed the LM dictionary to demonstrate relation between short window market reactions and disclosure tone changes in MD&A sections of 10-K filings (Feldman et al., 2010), abnormally high optimism in earnings announcements for firms under indictment (Rogers et al., 2011), or the association between negativity in form 10-K reports and firm's strategies to avoid taxes (Law & Mills, 2015).

An alternative approach to capturing the sentiment of a text relies on machine learning-based measurements. Lewis (1998) pointed towards the applicability of the Naive Bayes Classifier algorithm

in textual analysis, which assumes independence between the occurrences of words, hence naive. Antweiler and Frank (2004) approached their sentiment analysis using the Naive Bayes machine learning technique to construct a measurement for bullish tone on internet stock message boards. Findings suggest that stock messages are associated with market volatility and disagreement among messages to increase trading volume. Li (2010) measured positivity in forward-looking statements of 10-Q and 10-K filings and found significant relations to financial and firm characteristics. A comparison to the conventional tone measurements mentioned earlier indicated that only the machine learning-based measure was able to predict future performance.

In an exhaustive analysis of different disclosure tone measurements for 10-K filings, Henry and Leone (2016) have compared general vs. domain-specific, equal vs. inverse-document weighted, as well as word-frequency vs. machine learning based measures. By testing the more general wordlists of the 'General Inquirer Harvard IV' and 'Diction' programs against the more financial language tailored wordlists from Henry (2006) and Loughran and McDonald (2011), the results support their presumption, that word lists, which consider the text type, perform better in a financial context, by including and excluding sentiment determining words. Further, the authors argue for equally weighted measures over inverse-document weighted measures as they perform similarly when using a domain-specific wordlist, but in addition are far easier to replicate. Lastly, equivalent results were found when comparing a Naive Bayesian machine learning algorithm and a word-frequency-based measure as predictor variables, leading to the authors' final suggestion of using domain-specific wordlists for equally weighted word-frequency-based measures to analyze the tone of financial disclosures (Henry & Leone, 2016). However, a comparison from Frankel et al. (2022) of dictionaries-based measures, i.e. General Inquirer Harvard IV, and Loughran and McDonald's (2011) dictionary, to different machine learning based measures, produced results strongly favoring the machine learning-based random-forest-regression-tree method over simple word counts.

3.3 Topic measurements

Besides calculating tone measurements, a later wave of research found textual analysis of financial disclosures to be a helpful method to investigate the dynamics of various topics. Chen and Srinivasan (2024) created a measurement for digital activities by counting information technology-related

keywords in the Form 10-K business description sections of non-tech companies across the years 2010–2020. These words were primarily identified from IT-specific articles and glossaries, and their occurrence counts in the 10-Ks aggregate the final measurement for a firm's digital activity. The authors found that digitally active firms are more highly valued when considering their market-to-book ratio, while other accounting performance variables were only weakly associated with digital activities.

Li et al. (2013) constructed a measure for competition based on the counts of the five narrowly defined keywords "competition, competitor, competitive, compete, competing" (Li et al., 2013, p. 433) and their variations containing an appended 's'. An occurrence is excluded from the count if at least one of the preceding three words equals one of four predefined terms that would negate or mitigate the presence of competition. Applying this measurement to form 10-K filings of the years 1995–2009, findings suggest a negative effect of competition levels on a firm's return on net operating assets. This measurement of competition was also employed later by Shi et al. (2018), who link information about competitiveness from 10-K filings to accounting and auditing enforcement releases combined with securities class action lawsuits. Their analysis revealed that competitive pressure is positively related to the existence of shareholder lawsuits, accounting irregularities, and accrual earnings management, while being negatively related to real earnings management. Andreou et al. (2020) also employed a dictionary-based approach to measure firms' market orientation derived from Form 10-K filings. The authors developed lists of core words combined with contextual words for both customer orientation and competitor orientation. A total of 21 keywords were taken over from the MKTOR survey instrument of measuring market orientation and have been enriched with an additional 86 synonyms identified from Princeton University's WorldNet Lexical Database and the Harvard IV-4 Psychosocial Dictionary. The count considers the occurrence of the core word, if it is surrounded by at least one of the corresponding contextual words within the radius of four words before and after the core word. Similarly to Li et al. (2013), counts are excluded if one of the core word's four preceding words is a negating or mitigating term. In a subsequent empirical analysis, this measurement indicated a positive effect of market orientation on firm performance, reinforced when a firm operates in a competitive environment.

In the domain of marketing research, Balvers et al. (2016) applied a measurement of form 10-K filing dictionary-counts from a concise list of 24 word variations covering the topic of customer services

and satisfaction, derived from Michalisin and White (2000). Besides this conventional word-count measurement, the authors also computed two indicator variables for when a customer satisfaction-related term is mentioned in a negative or monitoring context. Accounting for the length of one sentence in a 10-K filing, the 300 characters before and after the customer satisfaction target keyword delineate the context boundaries. If at least one term from a predefined monitoring (negativity) wordlist occurs in this context, the respective indicator is equal to one. Mishra et al. (2020) derived a concise list of 10 words from marketing literature that are related to customer value proposition (CVP). After an evaluation by three marketing academics, a final set of five words remained, which all jurors agreed reflected CVP. Analyzing 10-K reports of B2B firms from the years 2004–2017 for the occurrence counts of these keywords, the authors demonstrated a small negative impact of CVP on future advertising expenses, but substantial positive effects on future brand-related investments and sales.

In the field of risk research, Campbell et al. (2014) created an extensive list of risk-related keywords, classified into five categories of risk types. Their dictionary is composed of existing keywords of prior literature, enriched by their own identification of words by reviewing sample text corpora, and further inclusion of missing words via applying machine learning-based document clustering. This latter approach is called Latent Dirichlet Allocation (LDA), a probabilistic Bayesian model that identifies hidden topics in a collection of texts based on patterns of word co-occurrence (Blei et al., 2003). Applied to extracted risk factors, MD&A and market risk disclosure sections of firms' annual 10-K reports for the years 2005–2008, their findings reveal that firms with more risk exposure also disclose more risk statements, supporting the quality of the, at the time, newly introduced risk factors section in 10-K reports. In their approach to create a measurement for cybersecurity risks, Cheong et al. (2021) employ LDA to identify cybersecurity topics in text corpora, together with word embedding, an identification technique of surrounding words based on vectorization. This combination is called LDA2Vec and was originally developed by Moody (2016), who demonstrated its enhanced precision for content analysis over conventional approaches.

Around the topic of corporate culture, Pandey and Pandey (2019) created a measurement based on a natural language processing approach of generating a coding dictionary, which covers six common dimensions of organizational culture. In a first step, a working definition and dimensionality of the

construct serve as a basis for deductively obtaining a comprehensive list of words (uni-grams) and phrases (n-grams) of synonyms from a thesaurus. Next, the application 'OpenNLP' extracted around 10,000 final verb and noun phrases from annual shareholder letters of 50 Fortune 500 companies to inductively generate data dictionary items. In a last step, the deductive and inductive items were coded following the dimensions and measurement instruments of organizational culture to obtain a final measure. Li et al. (2021) measured corporate culture by using the machine learning-based approach 'word embedding'. This approach represents words as numeric vectors and works under the assumption that focal words surrounded by the same neighbor words are related to each other (Harris, 1954). The natural language processing application 'word2vec', developed by Mikolov et al. (2013), is able to efficiently assign and process such a vectorization via a neural network, which is based on the source text corpus. The tool can be configured for vector size, radius of surrounding words, number of text processing iterations, minimum word frequency consideration, and underlying training method. Oriented on the five most often stated cultural values of S&P 500 firms, Li et al. (2021) define and further qualify a set of seed words for each cultural value, which serve as input for word2vec to identify related words, resulting in an extensive dictionary for corporate cultural dimensions. The final measurement is based on the occurrence counts of dictionary items in earnings call transcripts over the years 2001–2018 and reveals positive relations between a strong corporate culture and several firm characteristics and performance outcomes.

3.4 Credibility

The previously mentioned studies around textual analysis all performed content or construct validity tests for their respective measurement before proceeding with their actual analysis. While these tests alone certainly contribute to the overall credibility of textual analysis, additional comparisons further underline the applicability of this method. Li et al. (2013) compare their textual analysis-based measure for competition against the Herfindahl-Hirschman-Index (HHI), which reflects the level of industry concentration based on total industry sales. While being weakly related to the HHI, their competition measure appears to be successful in predicting a firm's return on net operating assets. Balvers et al. (2016) have compared their Customer Satisfaction keyword measurement against the American Customer Satisfaction Index (ACSI), developed by Fornell et al. (1996). This score reflects aggregated

US economy, industry and sector information about customer satisfaction trends, collected from telephone survey responses. Comparing the two measures revealed that textual analysis-based counts of Customer Satisfaction keywords are negatively related to ACSI scores. However, when the term was mentioned in the context of monitoring and measuring, the relation to ASCI became positive for retail firms. Enache et al. (2022) address the valid concern that innovation-related intentions disclosed in annual reports are rather vague and might not necessarily reflect the firms' actual realization of such investments, especially with a long-term focus. To investigate whether these statements are trustworthy, the authors analyze a sample of 67,173 Form 10-K reports over the years 1996–2017 for textual emphasis on innovation (TEI) and how capital expenditure, changes in property, plant and equipment, as well as patent applications, are affected. The results indicate that TEI can be associated with actual long-term investments, qualifying the disclosed innovation intentions as a trustworthy signal to investors and researchers. Further evidence suggests that this trustworthiness increases if a firm is subject to scrutiny, and that the signal through TEI typically comes at the cost of future over-investing in innovation.

4 Empirical research on artificial intelligence

In recent years, several studies investigated the nature of AI Adoption by using different data sources for their empirical analysis. The following sections will provide an overview of studies employing these different data sources and outline their advantages and limitations.

4.1 Survey data

A variety of studies conducted surveys or accessed existing survey data, in which respondents provided information about whether and how their company implemented AI-related processes or applications. After developing a framework to define AI capabilities within companies covering eight types of tangible, intangible, and human resources, Mikalef and Gupta (2021) conducted a survey among 143 technology managers to identify the extent to which these capabilities are present in their companies. Their findings suggest a positive relation between the presence of AI capabilities and both organizational creativity and performance. Chatterjee et al. (2021) aim to explain both the internal and external factors of AI Adoption by combining a technology-organization-environment (TOE) framework with a technology acceptance model (TAM). Their analysis of 340 responses from professionals in small, medium, and large digital manufacturing and production companies yielded insight into how socio-environmental factors affect an individual's perception of the technology's usefulness and ease-of-use, which both positively relate to the overall intention to adopt AI. Additional evidence suggests that these relations are further reinforced by the support of leadership. Analyzing a dataset of 1,882 hospital-year observations, evenly divided in AI-using an non-AI-using by applying a binary measurement, Pham et al. (2024) were able to detect a positive impact of hospital market share, measured by the number of patient days, on AI Adoption as well as a positive effect from AI Adoption on hospital performance, considered as outpatient revenue, inpatient revenue and occupancy. However, results did not indicate a positive relation between AI Adoption and hospital performance when considering return on assets. The authors explain this with their sample composition of mostly non-profit and government hospitals, which likely prioritize care and treatment over conventional performance outcomes. In a study about the link between AI Adoption intensity and firm performance, Lee et al. (2022) collected responses from high-tech venture companies about the extent of their AI investments, yielding a cross-sectional dataset of 160 observations. Their results indicate that only a certain level of AI Adoption intensity, i.e. being at

least 25% invested in an AI-related technology, leads to a significantly positive impact on the firm's revenue growth rate. Additional findings suggest, that this influence is being amplified if a firm priorly invested in complementary technologies related to cloud computing, data, and research and development (R&D). Rammer et al. (2022) used a large-scale dataset of German firms within the Community Innovation Survey conducted by the European Commission for the year 2018, in which technology managers gave insights into the range of actively used AI applications in their firms. Out of the final 6,738 observations, 573 firms stated active usage of AI. Analyzing relationships demonstrated a positive impact from AI usage on cost-reducing process innovation and on sales through world-first product innovation. Results further indicate that innovation is higher for companies with a larger variety of AI utilization and a longer track record of AI experience. Czarnitzki et al. (2023) used the very same data source both for a cross-sectional analysis of the year 2018 and as panel data for the years 2017 and 2018. The cross-sectional analysis provided evidence for significantly positive relationships between productivity measured in sales and both the existence of AI usage and the intensity of AI usage. The effects remain relatively consistent for the panel regression analysis, where the results suggest a sales increase between 4.4% and 6.2% with regard to AI usage and on average 5.2% higher sales when considering the firm's AI intensity. Considering value-added instead of sales produced similar findings and yielded a 14.6% increase through the usage of AI. Lastly, a large pool of 447,000 firms that responded to the Annual Business Survey of 2018 in the United States, covering start-ups, small and medium-sized enterprises (SMEs) as well as large firms, served McElheran et al. (2024) as a data source to analyze several dynamics relating to AI Adoption. Their findings revealed that younger companies in the manufacturing, information, and health care sectors with a high number of employees are the leading adopters of AI. Large businesses and high-growth start-ups seem to adopt AI at a higher intensity, suggesting the formation of a gap to SMEs. The adoption of AI in fast-growing start-ups was found to be fueled by the characteristics of their entrepreneurship, reflected by younger age and advanced education, and prior experience of founders. Geographically, start-ups were found to be located closer to known technology areas and locations with already increased AI activities. Further, the presence of digitization and cloud technologies indicated a significantly positive effect on AI usage as well as the firms' strategic orientation towards growth, process innovation, and product innovation.

The key advantage of survey data is the enhanced precision over other research methods. With respondents typically being employees with a technical background or managers who oversee technologies employed within their area of responsibility, researchers may not only obtain information about whether or not AI was adopted, but also to what extent and for which purpose. This ultimately allows for a more in-depth analysis of the subject while maintaining higher confidence in accuracy. Further, it is possible to assemble a more comprehensive sample that is generalizable for a larger range of company sizes and industry affiliations. However, conducting surveys usually comes at the cost of time and resources, which limits the number of observations and thus the possibilities for more complex analyses. Moreover, the samples are, in most cases, cross-sectional, which may lead to possible time-specific distortions. Especially in the case of AI, which gradually emerges over time into more and more technology portfolios, panel data could provide valuable insights regarding its adoption. Lastly, various survey-related biases can only be accounted for to some extent. Among other factors, the accuracy of responses is affected by the objectiveness of respondents, or their ability to provide correct information about a complex topic, such as details regarding AI Adoption.

4.2 Human resources data

Being a more accessible alternative to surveys and allowing for large sample sizes, some studies have utilized information from job postings to identify AI-related job descriptions. Goldfarb et al. (2020) have been among the first to utilize this type of data by composing a dataset of 4,556 US hospitals and their 1,840,784 job postings during the years 2015 to 2018, which have been archived in the database of 'Burning Glass Technologies'. AI Adoption was found to be rather low, with only 1,479 AI-related job postings from 126 unique hospitals, and jobs are predominantly covering administrative tasks. Using a binary measurement of AI Adoption, depending on whether AI skills have been required in the job posting, findings suggest that AI Adoption is positively influenced by larger hospital sizes, larger counties, and the utilization of an integrated salary model. Bäck et al. (2022) assembled a dataset from the Finnish job posting database 'Oikotie Oy' linked to financial data from 'Bureau van Dijk Electronic Publishing (BvD) ORBIS', comprising 2,476 observations of 721 distinct firms from 2016 to 2019. Statistically significant findings have only been observed for larger firms, while still indicating a positive impact from AI Adoption on productivity. Further findings suggest a delayed effect of at least three years

until the investment in AI yields an increase in productivity. Acemoglu et al. (2022) also extracted over one million job postings in the US from the Burning Glass database covering the years 2010–2018. By dropping the information sector as well as the professional and business services sector, the authors aim to exclude job postings of companies that either develop AI applications or implement them in other businesses. Applying measurements of AI occupational exposure on establishment-level developed in previous literature (Felten et al., 2019; Webb, 2020), results indicate a rise in job postings demanding AI-related skills after 2015, especially for establishments with high exposure to AI. At the same time, these AI-exposed businesses appear to reduce their overall hiring and non-AI hiring activities. Findings further suggest a shift in required tasks and competencies, as firms adapting to AI stop seeking certain traditional skills and start requiring new ones. Babina et al. (2024) constructed a large dataset around employee profiles from 'Cognism', containing information about the AI-relatedness of current jobs. Linking this source to firm-level data from 'Compustat' left the authors with 1,993 firms over the period 2010–2018. Using employee profiles allowed for the construction of a new AI intensity measurement by calculating the ratio of AI workers to total employees within a firm. Overall, their extensive dataset allowed for a wide range of analyses contributing new evidence to previous findings. First, a self-reinforcing loop was identified as large firms tend to invest more in AI and consequently appear to increase in size, sales, market share, and number of employees. Second, the observed growth related to AI seems to stem from product innovations and larger product portfolios rather than reduced costs.

In contrast to survey-based studies, human resources-related data like job postings allow for larger sample sizes and panel data analyses. Earlier evidence from Tambe and Hitt (2012) already supported the assumption that a firm's hiring activities allow for inferring that their adoption of technologies is associated with demanded skills. Considering the number of job postings per company and year, as well as the frequency of AI-related terms in the description, additionally enables researchers to measure AI Adoption not only binary, but at an intensity rate. However, when compared to survey data, human resource-related data is limited to the firm level, without delivering any insight regarding the diffusion or maturity of AI within the company. Further, the case of Goldfarb et al. (2020) has shown that applying this method to a narrower research subject could result in data limitations. Only three percent of hospitals in their sample had posted any AI-related jobs in the four-year time frame, impeding

more in-depth statistical analyses. Lastly, it is questionable whether the description of roles and tasks reflects the company's current needs based on past technology investments, or rather, its future technology considerations. Especially when competing for a group of high-skilled talents, who are essential to employ AI and realize the desired return on investment, firms might start recruiting before the investment into AI is actually made. Additionally, the description might be phrased to appear more attractive to applicants by potentially overemphasizing the subject technology in a way that would distort the final measure.

4.3 Patent data

Large databases of patent applications serve researchers as a convenient data source for determining innovation activities across longer time spans, at a higher scale, and cross-regional. Webb et al. (2018) have shed light on the research opportunities around measuring AI Adoption via related patent activity. Parsing for relevant keywords in the titles and abstracts of all patents filed between 1970 and 2015 in the 'United States Patent and Trademark Office (USPTO)' database allowed the authors to identify patent activities for several technologies, "such as cloud, machine learning, and neural networks" (Webb et al., 2018, p. 11). Findings based on the count of patents including relevant keywords suggest innovations related to neural networks to kick off in 1986 and machine learning in 1998, with large patenting activity by US-based technology giants like Microsoft, IBM, or Google. Alderucci et al. (2020) linked a large set of patents from the USPTO database to US census microdata over the time period 1997 to 2016. Employing machine learning techniques allowed the authors to identify AI-related patents at a large scale after establishing a training dataset of 1,000 non-AI and 1,000 AI-related patents. An event study revealed a positive effect of 8.1% from AI patent activity on employment and a 4.15% increase in revenue per employee. A distinct analysis of manufacturing firms produced positive effects of AI patent activity on the workers' total factor productivity between 7% and 8.9%. Damioli et al. (2021) accessed patents from the European Patent Office's database 'PATSTAT' and linked them to business data from the ORBIS database. Their final panel data set of 40,717 observations across the years 2000 to 2016 consists of 5,257 firms that have filed at least one AI-related patent in the respective time frame. Their results indicate a positive impact of AI patent activity on labor productivity, mainly driven by patent applications in the latter half of the observed time period. Further analysis of this relationship for

the subsample of the years 2009–2016 only revealed significant positive effects for the services sector and SMEs. Yang (2022) created a dataset of 10,651 observations of firms filing patents within the time 2002–2018 at the Taiwanese Patent Office, out of which 1,353 filed an AI-related patent. The results display a positive effect of the binary measurement of AI patent activity on a firm's total factor productivity. When considering the continuous measurement of AI-related patent counts, the significant positive effects remain, yet at a smaller degree. The same dynamic holds for the influence of AI patent activity on labor productivity as well as on employment, overall confirming the findings of previous studies (Alderucci et al., 2020; Damioli et al., 2021). Additional findings suggest significant negative effects of AI patenting on the ratio of employees with less than a college education (Yang, 2022).

Similarly to human resources data, observing the contents of filed patents can produce larger sample sizes over an extensive time period, while still limiting the analysis to the firm-level. Moreover, patents resemble the results of already performed activities and investments into AI-related developments, providing more confidence for inferences about actual AI Adoption. On the other side, this excludes firms from the observation that do not perform patenting activities, which might limit the interpretability of findings. Patents mentioning AI further only indicate AI usage, which stems from proprietary developments. Therefore, AI Adoption through third-party vendor procurement is completely left out of the analysis.

4.4 Annual report data

While already heavily used in previous economic literature, annual reports of publicly listed companies have only been utilized in recent years as a source to find evidence for AI Adoption activities. Wang and Qiu (2023) performed a textual analysis of annual reports from the 'WinGo' database filed by listed firms in China during the years 2006–2020, linked to financial data and employee education data, producing a final dataset of 28,463 observations. Aiming to explain the effect of AI Adoption on labor cost stickiness, they constructed a measurement for AI Adoption based on counts of AI-related terms, which were defined based on the machine learning technique 'word2vec'. This technique enriches a set of pre-defined seed words with similar terms based on their vectorized relatedness. Results indicate a positive effect of AI Adoption on labor cost stickiness, suggesting that AI Adoption in the context of decreasing sales rather has an augmentation effect on the workforce than a displacement effect. Yet,

these firms have to account for increased labor adjustment costs as their investment in AI requires more educated staff. Zhang et al. (2024) used a combination of annual reports from 2012 until 2022, retrieved from the CNINFO database, filed by listed Chinese companies, together with patent and financial data. As in Wang and Qiu (2023) the authors employed the machine learning approach word2vec to define AI-related keywords. Analyzing their final dataset of 22,057 observations among 3,784 unique firms provided evidence for a significant positive effect of AI on firm-level value creation, measured as Tobin's Q. Additional findings suggest a higher benefit from AI Adoption for manufacturing firms and firms located in the eastern region of China. Aiming to answer how AI delivers these positive effects, patent data was used to confirm the positive impacts that AI, in conjunction with sensing, integrating, and restructuring mechanisms of digital agility, had on a firm's value creation. As a third study utilizing annual reports of Chinese listed companies, Zhong et al. (2024) created a dataset of 25,453 firm-year observations from 2007–2020, enriched with financial data, patent data, and employment data. Their measurement of AI is also based on keyword counts, but the terms revolve around the technical categories 'Intelligence', 'Cloud', 'Machine Learning', 'Data', and 'IoT (Internet of Things)'. This results in a broader definition of AI as compared to previous studies, certainly explaining the high share of 88.17% of companies that were reported to have implemented AI in 2020. Investigating the influence of AI on total factor productivity (TFP) revealed significant positive effects, which seem to apply for the high-tech manufacturing sector and for the services sector, slightly favoring consumer services firms. Additional analyses on innovation and labor characteristics moderating the relation between AI and TFP imply that high- and low-quality patenting activity as well as high-skilled labor enhance the positive effect of AI on TFP, while the deployment of low-skilled labor appears to hinder AI-fueled productivity gains. Lastly, I identified Mishra et al. (2022) as the first study to analyze AI based on annual reports filed by listed US companies. Drawing on a dataset of 19,000 firm-year observations across the years 2005–2019, the authors aim to explain how the intensity of a firm's focus on AI relates to its performance. Their word-count-based measure of AI-related terms was developed inductively by a group of subject experts. Results provide evidence for AI Focus to decrease the costs of advertisement and to increase return on sales, return on marketing investment, and net income per employee. In addition, AI appears to cause increased employment but decreased gross operating efficiency. The

authors argue that AI Focus not only drives automation, but also requires new positions and higher salaries, suggesting that efficiency improvements can be rather observed long term.

Obtaining empirical insights on AI Adoption by textually analyzing annual reports appears to be a valuable addition to prior approaches. This method also allows for the generation of a larger sample size over a long time frame, deriving observations at the firm level. A key advantage is the standardization through regulators. This ensures comparable data points across the observed time period, since the report is filed each year and has to follow a certain structure. Further, regulating authorities demand an objectively written report, mitigating possible alterations of the truth. This overall raises confidence when drawing conclusions on the mentions of AI in these report texts. However, in contrast to the aforementioned data sources, annual reports are predominantly filed by publicly listed companies, which usually resemble large enterprises. Therefore, small- and midsized firms are excluded from the sample, which limits certain insights. Further, as elaborated in Chapter 3, textual analysis requires cautious handling. Researchers have to consider the interpretation of the subject of analysis, dependent on the specific text section of the annual report. For example, if AI is mentioned within the risk disclosure section of the report, the meaning is likely different, as compared to AI mentioned in the business description section. Additionally, even when mentioned within the appropriate text section, it is not ensured that the disclosure of AI also means AI Adoption. In order to reduce Type I error in the sample, these subtleties need to be accounted for.

4.5 Research gap

In light of the inarguable relevance of the subject and considering the possibilities and limitations of the empirical analysis of AI in previous literature, further advancements and refinements of existing empirical methods are critical to better understand the different facets of artificial intelligence in a business context. While surveys already present a reliable source of information, one can see the need for improved large-scale empirical research. Being a rather novel approach to analyzing topics like technology adoption, the textual analysis of annual reports presents a valuable addition to existing empirical methods by delivering panel data insights at a larger scale. However, the few studies that previously employed this method to analyze AI activities fall short regarding some key aspects. First, all studies did not consider AI within a particular text section, but observed the entire document for AI-

related terms. This neglects the differentiation between risk, legal, business, or other report sections, which, depending on the research subject, limits the analysis of AI to very generic insights at best, or at worst, produces Type I errors in large quantities. Second, these studies observed AI within a broader scope by including generic terms in their dictionaries that can be related to AI, but also to other technologies or functions. This likely results in an inflated overall count of dictionary items and presents another source of possible false positives, as some counts might capture non-AI-related disclosures. Third, none of the existing measurements of AI considered the context in which a dictionary item was mentioned. Again, this limits the analysis to either a less-specific understanding of artificial intelligence or potentially false assumptions about whether a company actually adopted AI, as opposed to simply mentioning it. As described earlier, such an additional contextual consideration improved the measurement of customer services and satisfaction, proposed by Balvers et al. (2016), significantly. Combined, all these limitations yield a questionable measurement of AI and hence might lead to serious misunderstandings about how certain factors are related to firms' AI activities. Besides the lack of measurement sophistication, artificial intelligence is currently evolving at a rapid speed, with multiple changes and additions over the last few years (McKinsey & Company, 2023a; Stanford University, 2024). Even when assuming a perfect measurement of AI at a certain point in time by having defined a highly accurate context-dependent dictionary, this dictionary is unlikely to capture AI terms with a similar accuracy in future reports. Hence, it is crucial to constantly advance the AI dictionary with updated terms that fit the time-based evolution status of AI. It also has to be acknowledged that over the first two decades of the 21st century, the empirical method of textually analyzing annual reports would likely not have captured a meaningful enough quantity of terms, when authors would have applied a narrowly defined and context-dependent measurement of AI. With more recent breakthroughs in the technology, resulting in increased business value, disclosures of AI in annual reports should have spiked significantly. This development should allow for a more selective approach to measuring the topic. This thesis aims to do exactly that. By leveraging the possibilities of computerized textual analysis of annual reports combined with the recently gained traction of the research subject, a more sophisticated measurement of AI is being developed that should address the limitations of prior empirical research.

5 Creating a textual analysis-based measurement for AI Adoption

In this chapter, I will describe the process of developing a model which is capable of measuring the actual adoption of AI. This measurement follows a dictionary-based approach consisting of two separate word lists: one including keywords related to AI and one with keywords related to adoption. I will outline how the dictionary for AI-related target terms has been derived, how the surrounding terms for the adoption context were determined, and how this dictionary approach is applied in an overarching model for parsing SEC filings. Before proceeding with the empirical analysis of my research subject, the chapter concludes with a validation of the textual analysis model by replicating the results of prior related studies.

5.1 AI dictionary development

Following the priorly discussed studies that employed word lists in their textual analysis, my dictionary for artificial intelligence contains both unigrams and n-grams, the latter accounting for approximately 90% of the entire dictionary. This is necessary, as in most cases, only the pairing of words constitutes the AI subject. For example, as uni-grams, the words ‘deep’ and ‘learning’ have a different meaning than their bi-gram combination ‘deep learning’. For simplicity, I will use the terminology ‘word’, ‘keyword’, ‘item’, or ‘term’ from here on to refer to uni-grams, bi-grams, and n-grams alike. The overall purpose of the AI dictionary is to cover an exhaustive list of items indicating AI activity, while avoiding the possibility that these words might describe a different technology, even if it is related to AI. For example, words like ‘data analytics’ or ‘modeling’ can be associated with AI in a broader sense, but could also appear outside of a narrow AI definition, such as in ‘Big Data’ or ‘Construction planning’. In this spirit, I aim only to consider items that fall within the boundaries of AI as outlined in Chapter 2. The deductive dictionary assembly follows three steps. First, I extracted items from dictionaries used in previous studies (Chen & Srinivasan, 2024; Van Roy et al., 2020; Yang, 2022) that met the specified conditions, resulting in an initial set of 37 items. For the second step, I enriched this preliminary list by parsing a limited sample of Form 10-K reports from the most recent filing year 2024 for the initial AI terms, which allowed me to identify filings with high counts of AI-related keywords. I manually scanned the 20 filings with the highest counts for additional terms that were not included in the preliminary list and appended missing items that also meet the conditions from above. As the Form 10-K filing primarily addresses

investors, firms tend to disclose basic AI terminology over more technical terms. Hence, I argue for selecting the filings with the highest counts, because firms with more AI mentions may be more likely to disclose additional non-basic terms compared to firms with only a few mentions of AI. In a third step, I applied an approach used by Chen and Srinivasan (2024), identifying relevant keywords from technological practitioner-oriented reports and glossaries covering the subject, while ensuring the conditional fit (the respective sources can be found in Appendix B). This enrichment procedure added 41 items to the initial list, resulting in a total set of 78 AI dictionary items, which are displayed in Appendix C. Notably, I consciously refrain from inductively supplementing AI-related dictionary items by applying a machine learning-based approach, such as word embedding. Previously described studies measuring AI Adoption with textual analysis (Wang & Qiu, 2023; Q. Zhang et al., 2024) employed Word2Vec based on a short list of seed words, producing extensive dictionaries with many items that fall outside the scope of my second chapter's outline of artificial intelligence. Inductive determination of keywords through experts (Mishra et al., 2022) was not feasible due to the limited resources available for this thesis, and this approach also appears to have yielded keywords that did not meet my criteria.

5.2 Adoption dictionary development

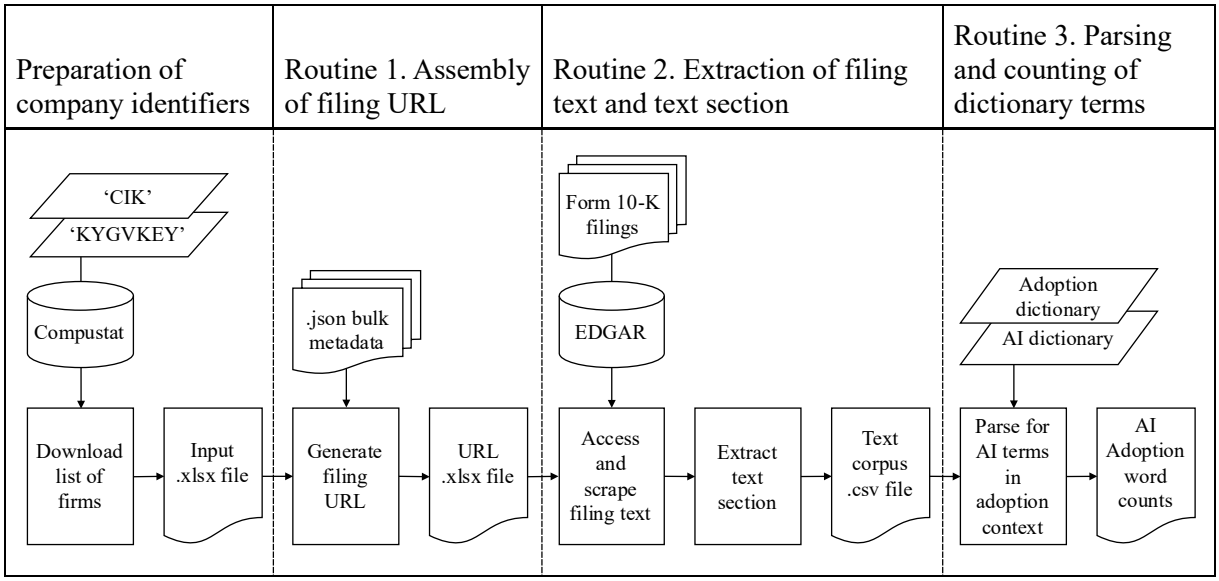
For itself, the AI dictionary is only able to capture how often an AI-related keyword was mentioned in the text corpus, making it uncertain whether the actual adoption of AI was disclosed or just the firm's awareness of the technology. To address this uncertainty, previous studies have either specified their measures as AI Focus (Mishra et al., 2022) or performed a robustness check by controlling for the sentiment of the sentence in which the AI-related term appeared (Wang & Qiu, 2023). Similar to Balvers et al. (2016), who enhanced their measure of customer service and satisfaction by identifying contextual words related to monitoring, I address this issue by defining an additional contextual dictionary with words that are positively associated with adoption. I approach the development of this dictionary by first defining three seed words that can confidently indicate whether a firm has adopted a technology. These seed words are 'adopt', to cover the phenomenon of adoption in general, 'invest', to include the commercial aspects of adopting a technology, and 'implement', to target adoption in a technical sense. Similar to Pandey and Pandey (2019), I expanded this initial set of words by incorporating their synonyms derived from Roget's 21st Century Thesaurus. This step resulted in 57 total dictionary items.

Next, I integrated the preliminary adoption dictionary into the parsing routine, as further explained in Section 5.3, to identify the number of AI-related items found in an adoption context. I then manually inspected the business description sections of the 20 Form 10-K filings with the highest differences between the counts of normal AI mentions and adoption-contextual occurrences of AI terms. I appended adoption-related terms that were missing initially, resulting in a final adoption dictionary of 110 items, including term variations, displayed in Appendix D. Alternatively performing a sentiment analysis for the words surrounding the target AI-term (Wang & Qiu, 2023) was initially a consideration, but this would have still left some uncertainty regarding actual adoption. Since the adoption dictionary only contains terms that relate positively to adoption, this approach combines the qualities of sentiment analysis with the narrower reference of adoption.

5.3 Development of a computerized textual analysis model

To develop a quantitative measure for AI Adoption, that is applicable in large-scale observations of listed US companies, I developed a computerized textual analysis model in Python, that can access filings from the SEC's database EDGAR, extract their textual contents, narrow down the selection to a specific section, tokenize the target text into individual sentences, parse for prior defined dictionary items and perform a simple as well as conditional counting routine. The programming efforts for developing the Python-based routine have been assisted by the use of OpenAI's ChatGPT, a pre-trained transformer capable of generating code based on initially prompted requirements.

Figure 1. Model flow of producing a dataset of AI Adoption word counts



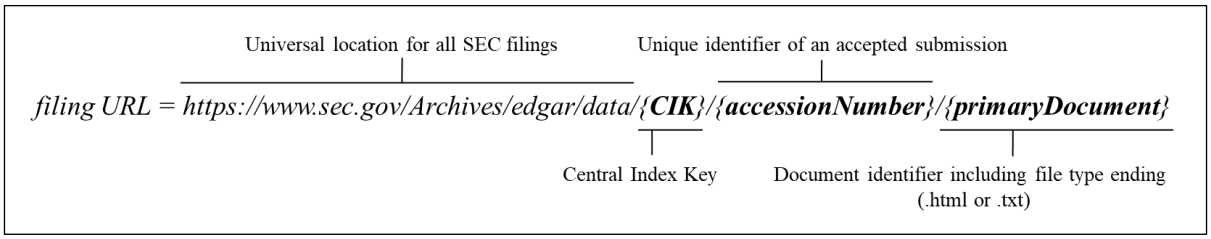
To ensure proper functionality and accuracy, the outputs have been carefully assessed and refined before being implemented in the final model. This section describes each step in the model's development and the creation of the dataset. Figure 1 depicts the entire process of assembling a dataset with AI Adoption word counts. Appendix F provides the corresponding Python code.

For most research applications, the empirical analysis of textual financial disclosures is combined with key financial figures or company characteristics, which are either integral to the research subject or used as control variables. For US companies, for which most of the Form 10-K filings correspond to, the CRSP Compustat Merged (CCM) database, from here on simply 'Compustat', contains an extensive collection of periodic reportable data such as balance sheet figures, stock information, company details, industry classification, geographic location, and many more. While it is possible to compose insights on a business-unit level, linking data to financial disclosures requires almost exclusively firm-level information on an annual basis, which will therefore guide the logic of my computerized textual analysis model. Importantly, Compustat stores multiple company identifiers that enable connections to other data sources, such as the SEC's EDGAR database. Due to the several possible options to narrow down the selection of companies according to specific research needs, e.g. by filtering for industries or stock exchanges, I recommend Compustat as a starting point to construct a simple list of firms. This also ensures time and resource efficiency, since not all filings from SEC-registered companies have to be accessed and processed before being filtered out later. At this point, the company list from Compustat should not contain the periodic unit or any other time-dependent variables. The variables for the final empirical analysis can be appended to the dataset at a later stage, once the textual analysis routine is finished. The initial company list should be saved as an .xlsx file and must contain the columns 'GVKEY', which is the company's unique identifier in Compustat, and 'CIK', which is the company's unique identifier in EDGAR.

In the second step, I execute the first of three Python routines, which generates the filing URLs for the respective report, in this case, the Form 10-K. This first script (see Appendix F) starts by initializing the form type and the filing years. Using the Python module 'pandas', the company list is loaded as a data frame, from which the CIK identifier can be accessed. This data frame is now being expanded to include an entry for every predefined filing year for every individual company. The SEC

maintains a .json metadata file for each company, which can be located via the CIK number. Iterating through all CIK numbers in the data frame, relevant metadata can be located via the initialized form type and the respective year of the iteration. This way, each necessary information about the filing date, reporting date, as well as the filing's accession number and the primary document number can be extracted and saved to the data frame. Figure 2 depicts how these components assemble the link to the respective filing. Lastly, the data frame containing the filing URLs is being exported to an Excel file for further processing.

Figure 2. Assembly of the Form 10-K filing URL



In the third step, a second Python script (see Appendix F) is executed, which begins by loading the Excel file exported in step two into a Pandas data frame. The routine now iterates through every index in the data frame and performs a request for the respective filing URL. Performing a request is defined in one dedicated function, which also accounts for several obstacles. First, at the time of writing this thesis, the SEC allows only ten requests per second to regulate website traffic, which is why I integrated logic that checks how much time has passed since the previous request. If the previous request was less than a tenth of a second ago, the routine pauses until the remaining time has passed, ensuring maximum efficiency. Secondly, despite adhering to this regulation, it is still possible that the SEC may temporarily block the user's IP address for making too many requests. This block should be lifted after approximately ten minutes; hence, the routine pauses for fifteen minutes if access has been denied. Lastly, the function accounts for possible internet connection failure, in which case the routine pauses for two minutes before performing another request. This logic is repeated until the internet connection is restored. Once a filing is successfully accessed, it is checked whether its URL ends with '.txt' or the more commonly used '.html' format. If it is an .html file or a .txt file with HTML structures, the filing text is being extracted using the Python module 'BeautifulSoup'. If the filing is an ordinary .txt file, the extraction works without the use of BeautifulSoup. Both the extracted filing text and its total word count are being saved

to the data frame. Additionally, a function is triggered that is capable of extracting a single text section from the entire filing text, based on predefined start and end words. For the scope of this thesis, I want to extract the 'Item 1. Business Description' section from the full 10-K report, therefore I define the start term as 'Item 1.' and the end term as 'Item 1A.', which marks the beginning of the next section 'Risk Factors'. This approach is an adaptation of a previously used extraction logic from Chen and Srinivasan (2024). I found only using 'Item 1.' or 'Item 1A.' to be most reliable, since instead of 'Business Description', some companies are writing variations like 'Business' or 'About the Business'. I also manually inspected a range of randomly chosen Form 10-K filings to observe how these start and end terms are used in the document. For most filings, each term occurs exactly twice, first in the table of contents and secondly at the beginning of each section. Few exceptions contained more than two occurrences, but I found these additional terms to be only mentioned after the beginning of the text section, e.g. when later sections refer to 'Item 1.' or 'Item 1A.'. In some rare cases, each term occurred only once, when it was not mentioned in the table of contents. As another obstacle, the 'Item 1A. Risk Factors' section was first demanded by the SEC in 2005, which is why older filings would miss the end term. To account for all of these peculiarities, the extraction function first checks if 'Item 1.' occurs more than once, which would define the second occurrence as the start term, or only once, which would define this single occurrence as the start term. If no occurrence was found, the function checks for 'Item 1' (without a dot). The same logic applies to the end term 'Item 1A.', but if no occurrence was found even without the dot, the function uses the respective occurrence of 'Item 2.' or 'Item 2' (without a dot) as end term. Having start and end points defined, the business description section is extracted, along with the word count. If the length is below 2,000 words, the section is excluded from further analysis due to potential extraction error or insufficient information content (for my analysis, the average length of business description sections is above 9,000 words). If the extracted section is qualified, the Python module 'nltk' tokenizes it into individual sentences using a full stop as a delimiter. The untrained module is smart enough to distinguish between sentence-ending stops and other dots used, for example, in numbers or salutations. After tokenization, the sentences are merged again into a string variable, allowing for space efficiency when exporting. By using a special delimiter to merge the sentences, they can be quickly split again into individual sentences for further processing. Optionally, the same

tokenization can be applied to the entire filing text. I wanted to perform the more processing-intensive tokenization with 'nltk' in this script, since it only needs to be executed once. In contrast, the dictionary item counting routine likely requires more cycles. Before the loop starts the next iteration, the extracted section, both as plain text and as merged sentences, is stored in the data frame along with the section's word count. Once the loop is finished, the data frame is exported to a .csv file for further processing. Having also tested a download-based routine, I found that using a .csv file as a storage container for filing texts was more viable than downloading and storing each file in a separate folder, as Campbell et al. (2014) did. Downloading the filings takes at least as long as extracting only the text contents, and the routine has to open each document to access its contents. In contrast, the .csv file allows loading multiple filing texts at once into a data frame. Overall, I observed superior efficiency when using the CSV-storage option, particularly for the dictionary parsing routine. Before proceeding further with the model's development, I performed a manual review of 50 extracted 'Item 1. Business Description' sections from different randomly chosen companies and years, and resolved arising issues. The initial inspection revealed that 88% of the sections were extracted accurately. The 12% falsely identified sections appear to have emerged from four different peculiarities. First, in three cases, 'Item 1.' or 'Item 1A.' were referenced in the text, before the section started. This led to 'Item 1.' either starting slightly earlier, because the early reference to it after the table of contents already marks the second occurrence of the term, or 'Item 1.' to end sooner, because the reference to 'Item 1A.' within the section counts as second occurrence, hence assuming the beginning of the risk factors section. I aim to solve this issue be defining a condition which excludes an occurrence, if one of the words 'to', 'as', 'in', 'and', 'or', 'see', 'from', 'under' or a comma occurs within 30 characters preceding 'Item 1(A).', suggesting a reference to the section was made. The occurrence is not excluded if a stop (.) occurs closer to the target term than one of the reference-related words. The second problem was caused by a filing using 'Item 1' (without a dot) in the table of contents, but 'Item 1.' (with a dot) when starting the text section. A later reference marked the second occurrence and led the routine to assume the starting point of the text section. I adjust the logic to simultaneously consider occurrences of 'Item 1(A).' and 'Item 1(A) ' (with a space at the end). The third issue occurred because 'Item 1A.' was missing in the text, despite being listed in the table of contents, which caused the routine to define 'Item 2.' as the end point. While this is probably a rare

occasion, I integrated a condition so that 'Item 2.' is only defined as the end point if the year is before 2005, which is when the SEC first mandated the risk factors section (Securities and Exchange Commission, 2005). Lastly, one filing used the formulation 'Item 1.A.' with two dots, which I added to my routine's search pattern. I repeated the section extraction with these changes and performed another manual randomized review of 50 extracted sections. The adjustment improved the accuracy from 88% to 96%, with only two sections, which began in the table of contents, instead of the actual text section. This adds the table's items and an entry note regarding forward-looking statements to the extracted section, which would likely not significantly distort the textual analysis. Therefore, I assess that my extraction routine performs well enough to continue with my research. Overall, from a total of 38,693 Form 10-K reports filed during 2017–2024, my model managed to extract 32,813 'Item 1. Business Description' sections containing at least 2,000 words, yielding a success rate of 84.8%.

In the third routine and final step, a new Python script initializes the predefined dictionaries and sets the global count values for each term to zero. The .csv file, exported in the third step, is loaded into a data frame in chunks of rows. This ensures compatibility with computer systems that have less memory, and allows the setting of the chunk size according to the system's specifications. The first chunk is processed with the dictionary parsing routing and exported to an .xlsx file, including the headers of the data frame. Subsequent chunks are being appended after the last row of the .xlsx file, excluding their headers. Now, a count variable is created for every dictionary and set by default to zero for each filing. Next, a loop iterates through every row of the chunk and parses the filing text or extracted text section for the dictionary items. For my purposes, I first perform a simple parsing and counting routine on the extracted text corpus 'Item 1. Business Description' for both the AI dictionary items and the adoption dictionary items. To measure AI Adoption, the occurrences of AI dictionary items are only counted if they appear in an adoption-related context. First, the extracted text section, which was tokenized and merged in the second routine, is split again into individual sentences and stored as a list variable, allowing an index position to be assigned to each sentence. Now, a loop iterates through each sentence and determines whether an adoption context can be assumed by parsing the current sentence for the keywords from the adoption dictionary. If at least one of these adoption keywords was found in the sentence, the routine parses the current sentence for occurrences of the term from the AI dictionary and

Table 1. Comparison between the 15 most counted terms of AI Adoption and AI Focus

AI Adoption Terms	Count	%	cum. %	AI Focus Terms	Count	%	cum. %
AI	13,264	39.5%	39.5%	Automation	32,124	21.5%	21.5%
Machine learning	5,199	15.5%	55.0%	Data processing	24,645	16.5%	38.0%
Artificial intelligence	4,662	13.9%	68.8%	Modeling	18,173	12.2%	50.1%
Biometric	1,939	5.8%	74.6%	Artificial intelligence	14,139	9.5%	59.6%
Quantum computing	1,190	3.5%	78.2%	Machine learning	11,102	7.4%	67.0%
Predictive analytics	1,143	3.4%	81.6%	Pooling	7,196	4.8%	71.8%
Predictive modeling	624	1.9%	83.4%	Robotic	6,929	4.6%	76.4%
Virtual reality	604	1.8%	85.2%	SAS	4,883	3.3%	79.7%
Computer vision	604	1.8%	87.0%	Prediction	4,242	2.8%	82.5%
Augmented reality	518	1.5%	88.6%	Algorithm	3,568	2.4%	84.9%
Deep learning	502	1.5%	90.1%	Data science	3,210	2.1%	87.1%
Neural network	395	1.2%	91.2%	Data acquisition	2,516	1.7%	88.8%
Machine vision	308	0.9%	92.1%	Predictive analytics	1,890	1.3%	90.0%
Virtual assistant	276	0.8%	93.0%	OCR	1,394	0.9%	90.9%
Natural language processing	242	0.7%	93.7%	Computer vision	1,384	0.9%	91.9%
…				…			
Total count	33,590			Total count	149,546		

adds the sum of counts to the AI Adoption count variable and the global AI Adoption dictionary value. After the loop has iterated through every index of the chunk, the entire filing text and the extracted sections' text are dropped from the data frame, so that it can be exported to an .xlsx file at a manageable

size. To selectively control counting accuracy, the filing text can be opened via the URL in a browser. Once every chunk has been processed and appended to the final .xlsx file, the dictionary items and their corresponding global counts are stored in a separate data frame to be exported to a distinct .xlsx file. This allows tracing how the total word counts are distributed across dictionary items.

This concludes the development of a computerized textual analysis model to parse the business description sections of Form 10-K files. The exported file can now be loaded into an application for statistical analysis to further clean the dataset, calculate the word-count-based measurements for AI Adoption, and investigate relationships. Missing variables from Compustat can now be appended using a combination of the firm's unique 'GVKEY' and the respective 'YEAR'.

5.4 Validation of the model mechanics

To ensure proper functionality of the text extraction and parsing routine, I attempted to replicate the results of prior dictionary-based textual analysis studies by applying each of their dictionaries to the respective text corpora. Mishra et al. (2022) are parsing the entire text of Form 10-K reports for the years 2005–2019 for AI-related terms, using a simple word count. The authors use an application called 'BUTTER' to perform this parsing and counting routine. Their final sample only includes firms that disclosed at least some AI activity. Table 2 of their study provides information on how their AI-related terms are distributed across the twelve Fama-French industries. Figure 3 depicts their results in comparison to my replication attempt. The replication is showing continuously lower numbers, suggesting a systematic difference in how the filings are processed or dictionary items are counted. However, the sum of counts distributed over the observation years reported in Table 3 of their study is 1,216 counts lower than the sum of counts reported in their word-count-industry distribution. This implies that the deviation in my replication might also be explained by other undisclosed data treatments. Still, with the slight exceptions for the industries 'Mines, Constructions', and 'Utilities', my replication appears to produce fairly similar results. Therefore, I evaluate my model to perform well at accessing filing texts and parsing for dictionary items.

With my model's text accessing and parsing abilities confirmed, its performance at extracting a single text section remains to be assessed. For this task, I attempted to replicate the results of Chen and Srinivasan (2024), who analyzed 'Item 1. Business Description' sections of Form 10-K reports over the

years 2010–2020 for disclosures of digital activities. This analysis considered only companies listed at the NYSE, NASDAQ, or NYSE MKT. The authors further excluded tech companies from their sample, which were classified in a comprehensive list of industry codes.

Figure 3. Replication of word-count-industry distribution results of Mishra et al. (2022)

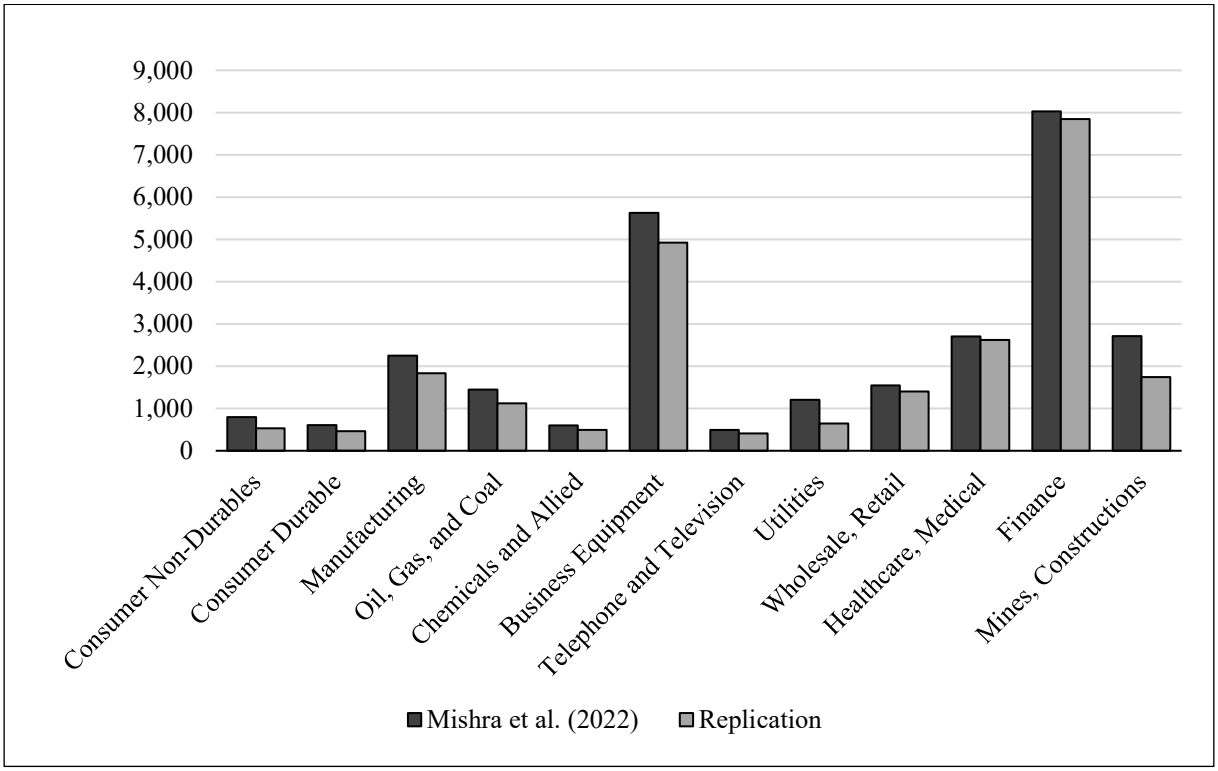
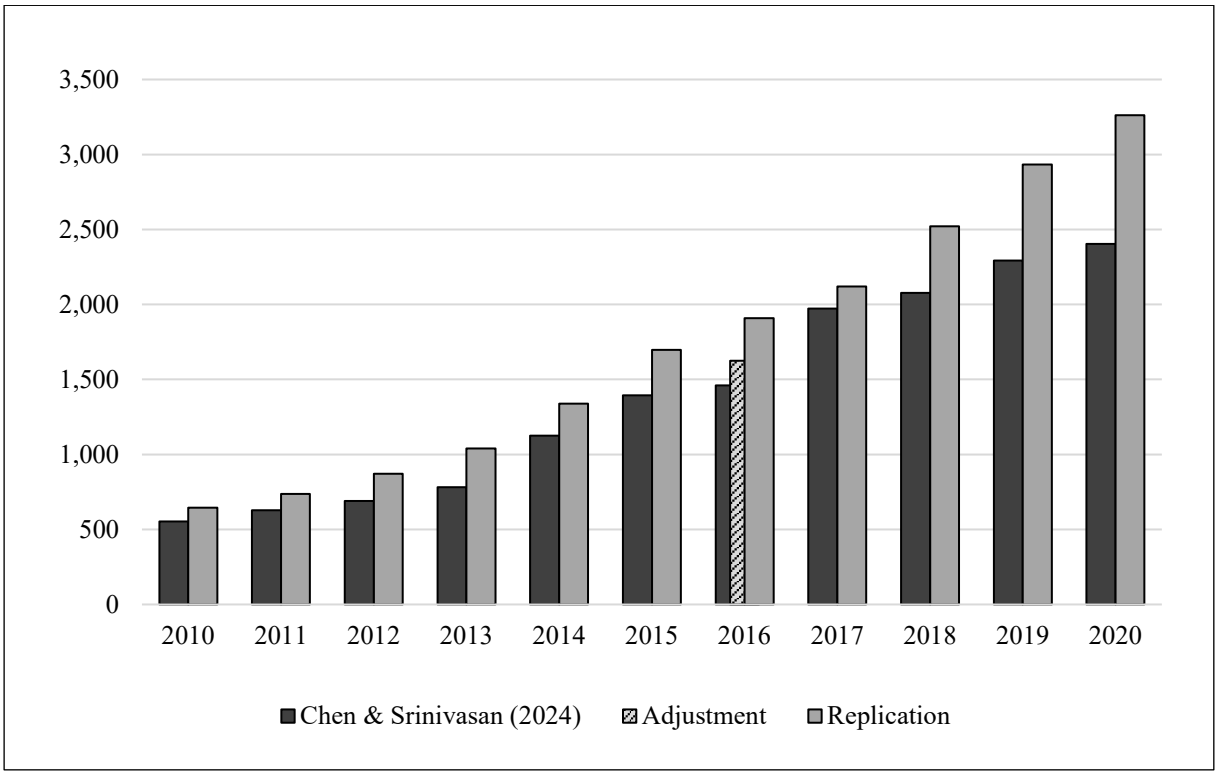


Figure 4 illustrates a comparison between my replication and the word count-year distribution of their study's Table 2, panel A. While displaying a similar trend, my replication yielded continuously higher counts, diverging towards the end of the panel. A possible explanation would be that my model was able to extract more text sections, thereby offering more sources for additional word counts. As described in Section 5.3, my logic for extracting a text section is similar to the one used by Chen and Srinivasan (2024), but perhaps a bit more refined. They describe their routine of defining the text sections' boundaries "by searching for the line with either "Item1" or "Business" and the lines with either "Item1A" or "Risk Factors." (Chen & Srinivasan, 2024, p. 1627). I initially tried to apply this approach, but achieved rather modest results. The divergence during the years 2018–2020 appears reasonable, as the firms' focus on digital technologies likely increased over time, which could have eventually been reflected in more digital-related disclosures per business description section. This exponential dynamic appears to be mitigated due to missing text sections. Another outlier in the comparison is the year 2016,

where a clear flattening is observed in the word count of Chen and Srinivasan. This is certainly driven by the missing word count for their digital activity category ‘Cloud’ for that specific year. Since the previous and following years’ word counts were reported as 109 and 220, the missing value must have been caused by an unknown error. I performed a simple linear forecast based on the counts for the years 2010–2015, which produced 164 counts for the missing value. Added to the existing counts in 2016, the additional pillar ‘Adjustment’ appears to integrate more consistently in the chart’s progression. Overall, the results in Figure 4 suggest that my text section extraction routine performs decently well, while likely being able to extract more text sections. The manual review, described in the second routine of Section 5.3, supports this claim.

Figure 4. Replication of word-count-year distribution results of Chen & Srinivasan (2024)



6 Hypothesis development

In Section 4.5, I addressed the limitations of prior research related to AI by enabling large-scale empirical analyses of a narrower and adoption context-related measurement of artificial intelligence. As this thesis aims to contribute to existing empirical studies on AI by proposing a novel measurement, the main scope is to compare the measure to the status quo and highlight the key differences. Section 4.4 already summarized the body of textual analysis-based studies as well as their advantages and limitations. Out of these studies, only Mishra et al. (2022) obtained their AI-related word counts on Form 10-K reports of publicly listed US companies, and hence qualify for actual comparability. However, I deem the measurement to be representative of the other studies, as all AI dictionaries appeared to be defined rather broadly, the respective word counts were based on the entire filing text, and none of the studies considered the context around the AI term. Along these three elements, I expect to observe significant differences between my proposed measurement of AI Adoption and the AI Focus measurement from Mishra et al. (2022), which represents the current body of dictionary-based textual analysis research on AI. First, both measures should likely differ due to the breadth of the AI dictionary definition. While I aimed to limit the body of terms to only those that would meet the definitions of AI described in Chapter 2, Mishra and colleagues' online appendices III and VI provide insights into their selection of terms and which of these have been counted most frequently. Across different industries, the leading counts stem from terms like 'SAS', 'Automation', 'Data Acquisition', or 'Data Processing', which I also observed as highly counted terms during my replication attempt in Table 1. Regarding the definition of AI outlined in Chapter 2, these terms do not guarantee the intelligent and learning capabilities that characterize artificial intelligence. While their underlying functions within a company can certainly be associated with AI, they do not provide certain evidence about the actual presence of AI. Hence, the overall word count of AI Focus likely includes companies that are active in these areas but have not yet engaged in AI activities, thereby producing false positives on a potentially large scale. Being more strict in the selection of terms, my AI Adoption measure likely mitigates the threat of this Type I error, offering a more passive but potentially more accurate measurement. Granted, this conservatism might exclude some observations of companies that adopted AI without disclosing it in a way that my measure would capture. Still, I assume an overall smaller impact of potentially false

negatives, as companies should have an incentive to signal their focus on innovation to the shareholders, especially when it promises competitive advantages and is currently being regarded as a future-leading topic. Secondly, the selection of the text corpus is likely a further differentiating element of both measures. As all studies discussed in Section 4.4, Mishra et al. (2022) included all sections of the annual report for their observation of AI-related terms. In the case of a Form 10-K report, such an observation also encompasses risk factors, cybersecurity concerns, legal proceedings, disclosures regarding market risks, and other matters. As outlined in Chapter 3, previous studies employing textual analysis of Form 10-K reports have singled out a respective text section, which the authors deemed relevant to their research subject (Campbell et al., 2014; Chen & Srinivasan, 2024; Cheong et al., 2021; Mayew et al., 2015). Following Chen and Srinivasan (2024), who analyzed firm's digital activities, I selected the business description section of the Form 10-K reports as the text corpus for my measurement, since actual adoption activities related to AI are likely found here and other non-adoption related disclosures, such as risk awareness, should be largely excluded from the observation. On average, the entire filing text of a 10-K report comprises about 66,000 words, which is over six times larger than the average length of about 10,000 words in the business description section. While this clear difference in text lengths already suggests deviations between the two measures by counting terms within an increased pool, the various topics of other sections would again entail the risk of false positives when trying to capture AI Adoption. Third and last, another differentiating element between the two measures can be assumed due to the introduction of context dependency. The consideration of surrounding words has already been explored in previous research (Balvers et al., 2016; Campbell et al., 2014), yielding a notable improvement in results for Balvers et al. (2016), as described in Section 3.4. In my case, the already limited scope of observation within the business description section is now additionally restricted to AI terms that occur within the same sentence as at least one adoption-related term. This could result in reduced counts, targeting a more conservative AI Adoption rate, or even the exclusion of certain observations where companies did not mention AI within an adoption context. Overall, this restriction aims to produce an AI word count that, due to its passive nature, can confidently reflect a firm's actual AI Adoption rate, with a limited risk of Type II error. Due to my rather exploratory research

agenda, which is less concerned with judging existing research achievements than with evaluating the contribution of my new measurement method, I cautiously formulate my first hypothesis as:

H1. There is a significant difference between a narrow measurement of AI Adoption and a broad measurement of AI Focus.

Under the presumption that a significant difference between both measures will be established, I further aim to demonstrate that my pursuit of capturing the actual adoption of AI is reflected in investment-related effects, which are less pronounced or even unobserved for the broad AI Focus measurement. In their survey-based study of US hospitals, Pham et al. (2024) found that total expenses are positively related to AI Adoption, providing evidence that the adoption of AI can be observed through expenditure-related variables. Investigating the determinants of corporate investment, An et al. (2016) found a significant effect of latent liquidity on a firm's contemporaneous investment rate. Chen and Srinivasan (2024) found that a firm's lagged capital expenditure is related to digital activities. Enache et al. (2022) produced evidence for a positive relationship between the textual emphasis on innovation, measured by the frequency of innovation-related words in 10-K reports, and long-term investment activities, measured by a firm's capital expenditures. The authors further observed that signaling long-term focus on innovation is costly for firms as it is associated with over-investment, an efficiency measure based on R&D, acquisition, and capital expenditure. I argue that artificial intelligence can be classified as such an innovation that would be textually emphasized in Form 10-K reports. Since this signal to invest in AI could come at the cost of over-investment as mentioned above, this might be reflected by investment-related effects. I further postulate that my measure's explicit focus on AI disclosure within a concrete adoption-related context should therefore capture these most costly signals, leading to a higher association with investment variables than a broader measurement. Of course, a firm's capital expenditure reflects far more than just potential technological investments, and it is uncertain whether AI-related investments are necessarily contained within this figure. However, under the presumption that reinvesting in existing assets signals orientation towards long-term growth (Enache et al., 2022), it can be argued that a company adopting AI is likely also investing in its future. In light of my exploratory research pursuit, I formulate my second hypothesis as:

H2. The actual adoption of AI is more prominently reflected in investment-related effects.

7 Empirical model

Chapter 7 describes the data sources used for my empirical research model and the variables I derive from them. I then discuss the descriptive statistics and pairwise correlation among these variables and conduct a preliminary validity test before proceeding with the actual analysis.

7.1 Data

The final sample for the empirical analysis comprises two distinct sources. To construct the measurement of AI Adoption, I accessed the SEC's database EDGAR to extract the 'Item 1. Business Description' sections of form 10-K filings, which report annual business activities of US companies for the years 2017 to 2023. Within the extracted section, AI-related terms were counted if they occurred within the same sentence as an adoption-related term. The entire extraction and counting process is elaborated in Section 5.3. It is essential to distinguish between the filing date and the report date of a Form 10-K, especially when linking to another panel dataset. While the filing date indicates when the report was submitted to the SEC, the report date specifies up to which point in time the report's contents remain relevant. In some cases, the report date and the filing date were more than two years apart, which entails the risk of substantial temporal distortions. Depending on the company type, the ultimate deadline to file the Form 10-K is May 31st. I assigned filings with report dates before that deadline to the report year t–1, which should reflect the time period that contains the majority of reported business activities. Filings with report dates after the deadline are assigned to the report year t. To ensure that the report year 2023 is fully represented, my sample includes 10-K reports that have been filed up to the end of May 2024. This procedure yielded 32,813 extracted text sections with at least 2,000 words from 6,152 unique companies, out of which 1,847 disclosed AI within an adoption context.

The SEC's EDGAR data is matched via the CIK number and the report year to business data from the CRSP Compustat Merged (CCM) database. 7,708 firm-year observations had to be dropped due to missing business data in Compustat. Further dropping observations with missing NAICS values, NAICS codes that changed during the panel period, or unclassified establishments, resulted in a final sample of 22,193 firm-year observations and 4,683 individual companies. 1,472 of these companies disclosed a total of 25,441 AI-related terms within an adoption context, ranging across 4,409 firm-year observations.

7.2 Variables and model

For the subject of my analysis, the dependent variable in my model is *AI Adoption*. In Chapter 5, I described in detail how I developed a routine to count AI terms in an adoption-related context. The final measurement is calculated by scaling this count by the number of thousand words in the respective text section, as depicted in the equation below. Compared to a binary measurement, this variable includes information about the degree of *AI Adoption*, allowing it to not only distinguish between adopters and non-adopters but also between different levels of usage. Further scaling this degree by the number of words in the text section ensures a more considerate interpretation of the count. This way, companies that use more words to describe their AI-related activities, but within an overall longer disclosure statement, should be more comparable to companies that discuss their entire business activities more briefly, including the application of AI. Multiplying by one thousand simply ensures that the *AI Adoption* count is adequately weighted within the text section, with total lengths varying between 2,000 and 20,000 words for ~90% of the text sections.

$$AI\ Adoption = \frac{No.of\ AI\text{-}related\ terms\ in\ adoption\ context}{Total\ No.\ of\ words\ in\ extracted\ text\ section} \times 1{,}000$$

The investment-related effects addressed in hypothesis two are represented by *CAPEX* – the firm's capital expenditure scaled by its total assets, and *Liquidity* – denoted as the natural logarithm of a firm's total cash. Lastly, I controlled for a range of variables that have been empirically examined in other studies analyzing technology adoption (Chen & Srinivasan, 2024; Mishra et al., 2022). Firm characteristics are captured through *Employees* – the natural logarithm of a firm's number of employees, *Age* – denoted by the number of years since the firm was initially listed in Compustat, *Size* – the firm's market capitalization, *Risk* – the four year moving standard deviation of earnings per share as well as its *Tangible Assets* – the companies' gross property plant and equipment scaled by it's total assets, and *Intangible Assets* – the firm's intangible assets scaled by it's total assets. Firm performance is measured with *ROA* – the company's operating income after depreciation scaled by the current and previous year's averaged total assets. Financial characteristics are included through *Leverage* – the firm's total debt divided by the stockholders' equity, as well as the company's *R&D Expenses* and selling, general and administrative *(SG&A) Expenses*, each scaled by the firm's total assets. Competition is measured using

Table 2. Calculation of variables

Variable name	Calculation	Compustat labels
AI Adoption	$\frac{\textit{AI Adoption terms}}{\textit{Total text section words}} \times 1{,}000$	
AI Focus (Mishra et al., 2022)	$\frac{\textit{AI Focus terms}}{\textit{Total filing words}} \times 1{,}000$	
Employees	*ln(No. of employees + 1)*	*ln(emp + 1)*
Age	*ln(current year – year of first entry +1)*	*ln(year – begyr +1)*
Size	*ln(market capitalization + 1)*	*ln(tcap + 1)*
Risk	$ln\left(\sqrt{\frac{\sum_t (x-\bar{x})^2}{4}} + 1\right)$ *x = earnings per share*	$ln\left(\sqrt{\frac{\sum_t \left(epsfx-\overline{epsfx}\right)^2}{4}} + 1\right)$
Tangible Assets	*ln(gross property, plant & equipment + 1)*	*ln(ppegt + 1)*
Intangible Assets	$\frac{\textit{Intangible Assets}}{\textit{Total assets}}$	$\frac{intan}{at}$
ROA	$\frac{\textit{Operating income after depreciation}_t}{(\textit{Total assets}_t + \textit{total assets}_{t-1})/2}$	$\frac{oiadp}{(at - at1)/2}$
CAPEX	$\frac{\textit{Capital expenditure}}{\textit{Total assets}}$	$\frac{capx}{at}$
Leverage	$\frac{\textit{Total debt}}{\textit{Total assets}}$	$\frac{dlc + dltt}{at}$
Liquidity	*ln(Cash + 1)*	*ln(ch + 1)*
R&D Expenses	$\frac{\textit{Research and development expenses}}{\textit{Total assets}}$	$\frac{xrd}{at}$
SG&A Expenses	$\frac{\textit{Selling, general and administrative expenses}}{\textit{Total assets}}$	$\frac{xsga}{at}$
HHI	*i = firm; j = industry* $\sum_j \left(\frac{Sales_i}{Sales_{i,j}}\right)^2$	$\sum_j \left(\frac{sale_i}{sale_{i,j}}\right)^2$

the Herfindahl-Hirschman-Index (*HHI*), which calculates a firm's market share as the ratio of firm sales to the overall industry sales, assigning the total sum of squared market shares to each three-digit NAICS industry. A low *HHI* value indicates a competitive industry environment with market shares distributed among multiple players, whereas a high value suggests a concentrated environment characterized by oligopolistic structures. The equation below represents the full model, encompassing all the variables described above. Following Chen and Srinivasan (2024), the determinants of the current year's *AI Adoption* are lagged by one year, assuming that the disclosure of *AI Adoption* results from an investment performed during the previous period.

$$\begin{aligned} AI\ Adoption_{i,\,t} = \alpha + \beta_1 CAPEX_{i,\,t-1} + \beta_2 Liquidity_{i,\,t-1} + \beta_3 ROA_{i,\,t-1} + \beta_4 R\&D\ Expenses_{i,\,t-1} + \\ \beta_5 Employees_{i,\,t-1} + \beta_6 Age_{i,\,t-1} + \beta_7 Size_{i,\,t-1} + \beta_8 Leverage_{i,\,t-1} + \beta_9 SG\&A\ Expenses_{i,\,t-1} + \\ \beta_{10} Intangible\ Assets_{i,\,t-1} + \beta_{11} Tangible\ Assets_{i,\,t-1} + \beta_{12} HHI_{i,\,t-1} + \beta_{13} Risk_{i,\,t-1} + \epsilon_{i,t} \end{aligned}$$

7.3 Descriptive statistics

Table 3 reports the descriptive statistics of all variables included in the final model. The average adoption of AI in the sample is relatively low, at 0.147, with a few firms exhibiting significantly higher levels of adoption. Capital expenditure activities are predominantly centered around a small mean of 0.028, with few firms investing higher amounts into their physical assets. Negative *CAPEX* values can occur if companies sell off their fixed assets. Considering *R&D Expenses*, average investment activities are relatively low at 0.099. However, a few firms appear to invest drastically more in R&D compared to others. 52.7% of the sample firms do not invest in R&D at all. While the *Leverage* ratio of a firm's total debt in relation to its total assets is relatively centered around the average value of 0.919, the maximum value indicates that a few firms are considerably more in debt compared to the vast majority. Contained within the total assets, *Intangible Assets* are distributed between zero and one, although low on average. The market concentration, as represented by the *HHI*, is low at 0.123, indicating a competitive environment for most firms, although a few are highly concentrated. The average *Risk* in the sample is moderate at 0.659, with some firms facing substantially higher risk than others. As Table 3 already suggests, some variables are fairly skewed, particularly those scaled by total assets.

Table 3. Descriptive statistics

Variable	Obs.	Mean	Std. Dev.	Min	Max
AI Adoption	22,193	0.147	0.683	0	23.525
CAPEX	22,193	0.028	0.048	-0.186	1.457
Liquidity	22,193	10.881	2.555	0	19.303
ROA	22,193	-0.023	0.213	-10.145	19.104
R&D Expenses	22,193	0.099	0.515	-0.001	55.668
Employees	22,193	6.556	2.663	0	13.266
Age	22,193	2.537	0.979	0	4.304
Size	22,193	13.478	2.413	0	21.751
Leverage	22,193	0.919	22.832	-1,672.55	659.274
SG&A Expenses	22,193	0.264	0.586	-0.042	29.816
Intangible Assets	22,193	0.172	0.222	0	0.991
Tangible Assets	22,193	0.336	0.514	0	15.962
HHI	22,193	0.123	0.125	0	1
Risk	22,193	0.659	0.58	0	9.922

When considering the pairwise correlation of all variables presented in Table 4, one can observe an overall low correlation between *AI Adoption* and other variables. The highest effect is observed for the positive relationship between *AI Adoption* and industry concentration (*HHI*) of 0.124 and *Intangible Assets* of 0.096. Considering the more pronounced correlations among control variables, economically expected patterns seem to hold. A strong positive correlation can be observed between *R&D Expenses* and *SG&A Expenses*, and high *SG&A Expenses* are negatively related to the return on assets (*ROA*). The number of *Employees* is highly correlated to a firm's *Size*, measured in market capitalization, as well as to its *Tangible Assets*. Further, there is a moderate correlation between *CAPEX* and *Tangible Assets*. Overall, the correlation among variables is low, implying reduced risk of multicollinearity, which qualifies the model for further regression analysis. Appendix G contains an overview of how selected variables are distributed with regard to *AI Adoption* and the *AI Focus* measure of Mishra et al. (2022).

Table 4. Pairwise correlation

Variables	(1)	(2)	(3)	(4)	(5)	(6)
(1) AI Adoption	1.000					
(2) CAPEX	-0.010	1.000				
(3) Liquidity	0.037	-0.021	1.000			
(4) ROA	-0.017	0.007	0.070	1.000		
(5) R&D Expenses	0.010	-0.033	-0.073	-0.320	1.000	
(6) Employees	0.044	0.123	0.512	0.180	-0.162	1.000
(7) Age	-0.057	0.024	0.136	0.191	-0.115	0.385
(8) Size	0.038	0.030	0.560	0.198	-0.139	0.631
(9) Leverage	-0.003	0.004	0.002	0.004	-0.006	0.015
(10) SG&A Expenses	0.059	0.012	-0.164	-0.373	0.446	-0.148
(11) Intangible Assets	0.096	-0.082	0.097	0.077	-0.077	0.351
(12) Tangible Assets	-0.049	0.493	-0.066	0.034	-0.048	0.173
(13) HHI	0.124	0.148	0.028	0.050	-0.073	0.183
(14) Risk	-0.027	0.074	0.177	0.021	0.009	0.186

(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)
1.000							
0.274	1.000						
0.010	0.004	1.000					
-0.156	-0.260	-0.012	1.000				
0.063	0.212	0.000	-0.039	1.000			
0.185	-0.022	0.008	-0.018	-0.102	1.000		
-0.024	0.040	0.006	0.042	0.186	0.173	1.000	
0.104	0.122	-0.003	-0.024	0.028	0.125	0.042	1.000

7.1 Validity test

Before continuing with the empirical analysis, I performed a preliminary validity test to see how *AI Adoption* is associated with the variables *R&D Expenses* and *Intangible Assets*, which were also associated with technology adoption in previous research (Chen & Srinivasan, 2024; Mishra et al., 2022). Note that this entire thesis aims to examine the fit of the proposed *AI Adoption* variable. Hence,

this test should be considered within the overall context of results. Due to the large share of non-adopters in my sample, this test only includes companies that invested to some extent in AI. The procedure of the validity test follows Mishra et al. (2022). The upper half of Table 5 shows the mean values of these variables for each quintile of *AI Adoption* intensity, where quintile one represents the least amount of *AI Adoption*, while quintile five consists of observations with the highest levels of *AI Adoption*. The bottom half of Table 5 presents t-tests for the differences between the mean values of each quintile, including their significance levels. Considering *R&D Expenses*, one notices that the mean value for the lowest *AI Adoption* quintile is the highest at 0.163. This contradicts the pattern from the second to the fifth quintile, which displays increasing mean values of *R&D Expenses* for higher levels of *AI Adoption*. This anomaly can be partially attributed to the scaling of *R&D Expenses* by total assets, which relates a firm's R&D activities to its size. Hence, larger firms with high levels of *AI Adoption* and high *R&D Expenses* can be located in the upper quintiles. Still, their ratio of *R&D Expenses* to *Total Assets* might be smaller than that of a smaller firm with lower *AI Adoption* and fewer *R&D Expenses*.

Table 5. Preliminary validity test of AI Adoption quintile comparison

	$\frac{\textit{R\&D Expenses}}{\textit{Total Assets}}$	$\frac{\textit{Intangible Assets}}{\textit{Total Assets}}$
Mean values		
Quintile 1	0.163	0.183
Quintile 2	0.075	0.243
Quintile 3	0.084	0.271
Quintile 4	0.09	0.285
Quintile 5	0.118	0.292
Mean differences and significance		
1 and 2	0.089***	-0.06***
1 and 3	0.08***	-0.088***
1 and 4	0.073***	-0.102***
1 and 5	0.045***	-0.11***
2 and 3	-0.009	-0.028**
2 and 4	-0.015**	-0.042***
2 and 5	-0.044***	-0.049***
3 and 4	-0.007	-0.014
3 and 5	-0.009***	-0.021*
4 and 5	-0.028***	-0.007

Solely considering the natural logarithm of *R&D Expenses* supports this explanation, showing a strongly mitigated effect, with the mean value of quintile one being only slightly larger than that of quintile two, but smaller than quintiles three to five. The t-tests below reveal significance throughout the quintile's mean differences, except for the two closely situated pairs, quintiles two and three, and quintiles three and four. When considering *Intangible Assets*, the expected pattern of increasing mean values along increasing levels of *AI Adoption* can be observed. Further, the differences between the mean values are significant, with the exception of quintile pairs three and four, and four and five. To conclude the preliminary validity test, both variables exhibit a growing pattern of mean values with increased *AI Adoption*. The only outlying mean value of *AI Adoption* quintile one for *R&D Expenses* can partially be explained by the scaling effect of *Total Assets*. Both variables demonstrate significant differences in quintile mean values, particularly in the more contrasting quintile comparisons. While this test alone is surely not sufficient to assess the validity of *AI Adoption*, it qualifies the variable for more in-depth empirical analysis.

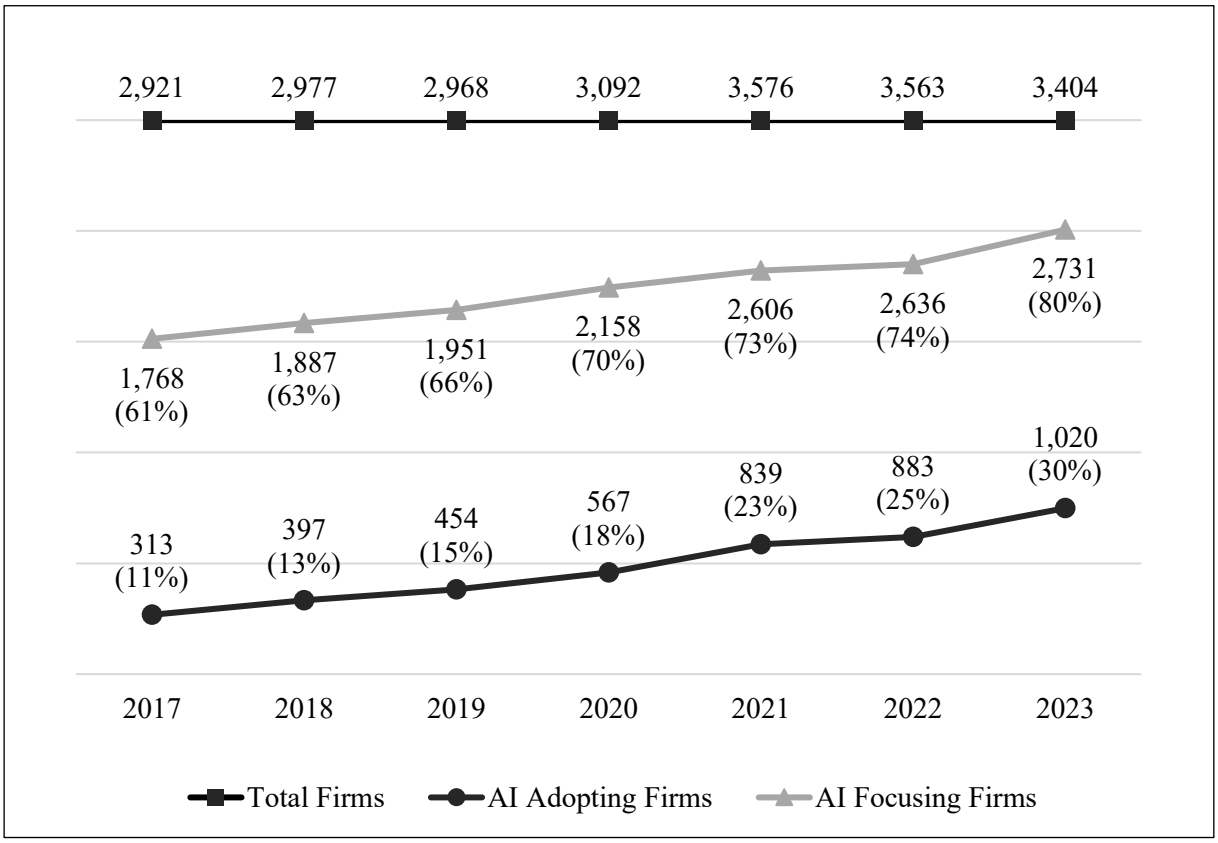
8 Empirical results

As previously mentioned, my empirical analysis is particularly concerned with demonstrating the fundamental difference of my proposed *AI Adoption* variable from the existing *AI Focus* measure of Mishra et al. (2022). First, I present the temporal progression of both variables and their underlying word counts, before examining the differences between the metrics and their potential implications for interpreting AI. Finally, I examine whether and how the different variables might determine *AI Adoption* and whether significant associations with investment-related variables can be identified.

8.1 Temporal development of AI

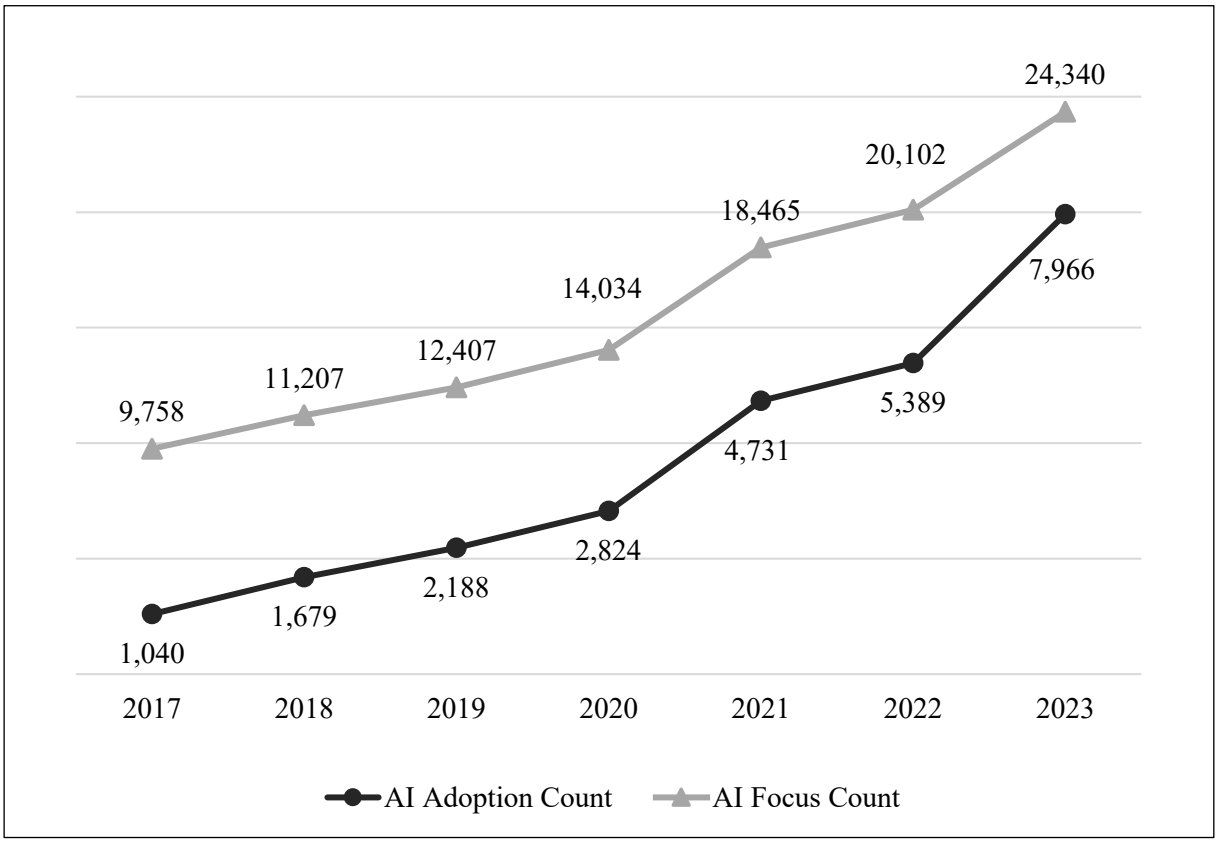
Figure 5 illustrates the evolution of companies' adoption and focus on AI over the sample time frame from 2017 to 2023. The numbers are calculated by assigning an indicator of one to firms that displayed any counts of *AI Focus* or *AI Adoption* dictionary items in the respective year. While both measurements produce a fairly comparable inclining course, it is worth noting that the share of AI adopting firms increased by nearly three times when comparing the years 2017 and 2023. This effect is less pronounced when considering the growth rate of *AI Focus*.

Figure 5. Firms adopting and focusing on AI during 2017-2023



A similar increasing pattern can be observed when comparing the counts of *AI Focus* and *AI Adoption* dictionary items over time, as shown in Figure 6. Note that the graphs for *AI Adoption* counts and *AI Focus* counts were assigned to different scales to enhance the comparability of the courses. Here, the rapid incline of *AI Adoption* becomes even more apparent, with counts nearly eightfold over the time period. Again, the growth of *AI Focus* counts is far more modest, at approximately 2.5 times. Despite their different growth rates, both counts reveal the same two sudden disruptions of the otherwise steadily inclining charts. While *AI Focus* has averaged a growth rate of 13% from 2017 to 2020 and 9% from 2021 to 2022, the counts for 2021 and 2023 increased by 32% and 21%, respectively. The chart of *AI Adoption* overall shows a more intense growth rate, with moderate growth of roughly 30% between 2018 and 2020, and 14% in 2022, as well as clear spikes of 61% in 2018, 68% in 2021, and 48% in 2023. This pattern fits into the historical chain of advancements in artificial intelligence, seeing medical and other industry applications in 2018 (Taulli, 2018), the release of large language models across the globe like OpenAI's ChatGPT-3 in 2021 (Heaven, 2021), as well as the breakout and increasing adoption of generative AI in 2023 (McKinsey & Company, 2023b). Compared to the *AI Adoption* dictionary,

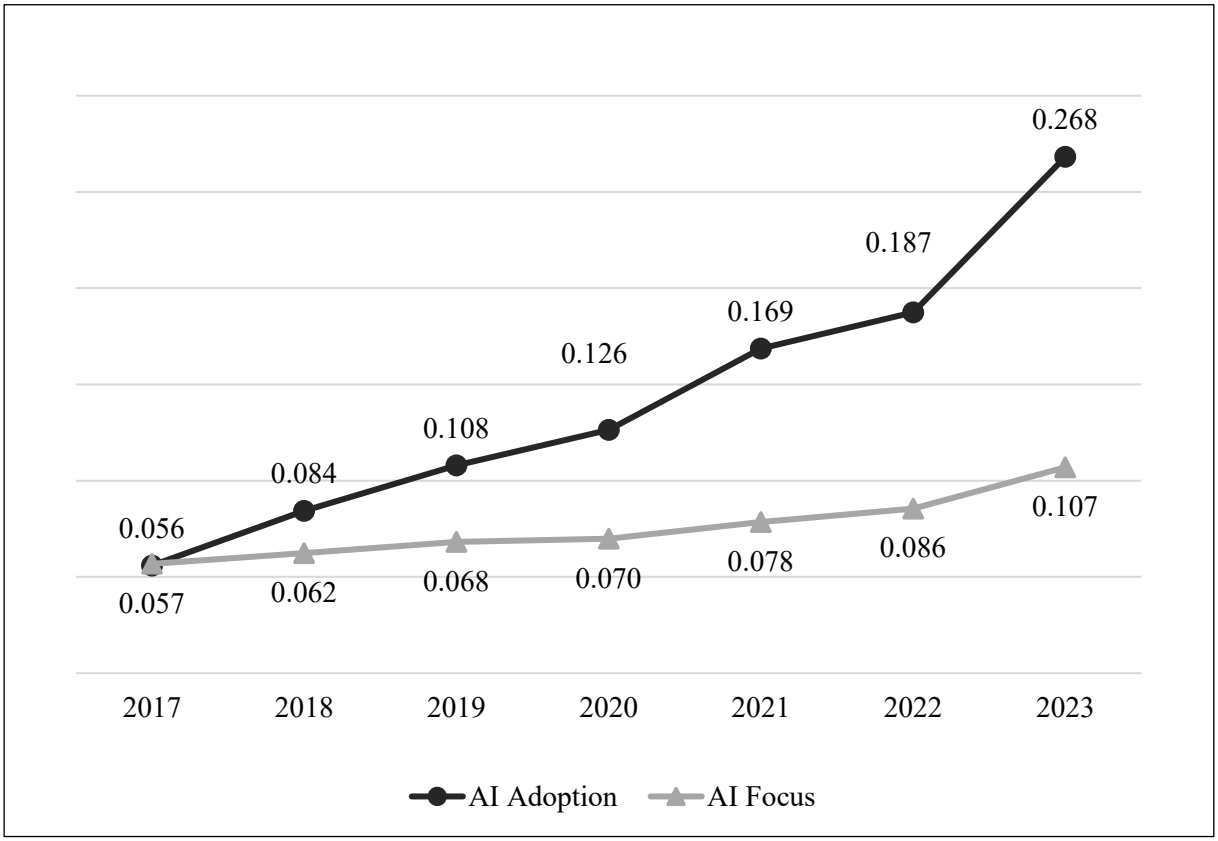
Figure 6. Temporal development of AI Adoption and AI Focus word counts



the broader scope of the *AI Focus* dictionary likely masked the accelerating effect of AI in 2018.

Lastly, Figure 7 illustrates the evolution of the final calculated measurements of *AI Adoption* and *AI Focus* over time. The first noticeable difference is the size of each year's average *AI Adoption* and *AI Focus* intensity. While *AI Adoption* counts significantly fewer terms, the higher count of *AI Focus* is divided by the drastically higher number of words in the entire filings, instead of the reduced number of words in the business description section. Similar to the previous two figures, the growth rate of *AI Focus* is less pronounced, and the points in time of AI acceleration are captured similarly to each measure's count of AI terms. When compared to the AI counts of Figure 6, the trend of growth rates for each AI measure is mitigated in 2020, at 17% (2%) for *AI Adoption* (*AI Focus*). This effect may be explained by the COVID-19 pandemic, which likely compelled companies to address additional business impacts in their reports, potentially resulting in an increased text corpus word count and, consequently, a smaller ratio of AI terms to overall filing words.

Figure 7. Temporal development of AI Adoption and AI Focus full measurement



8.2 Differences in measurements of AI

To test my first hypothesis, that a narrower understanding of AI in an adoption-related context would differ significantly from a broader definition without evidence for adoption, I tested the *AI Adoption* variable developed in this thesis against the *AI Focus* variable proposed by Mishra et al. (2022). Figure 8 presents the histograms of both variables. The distribution is fairly right-skewed, predominantly caused by the large number of firm observations for each measure, which did not disclose any focus or adoption of AI. However, even when these observations are neglected, both measures are distributed more densely around the lower rates of *AI Adoption* and *AI Focus*.

Figure 8. Distribution of AI Adoption and AI Focus variables

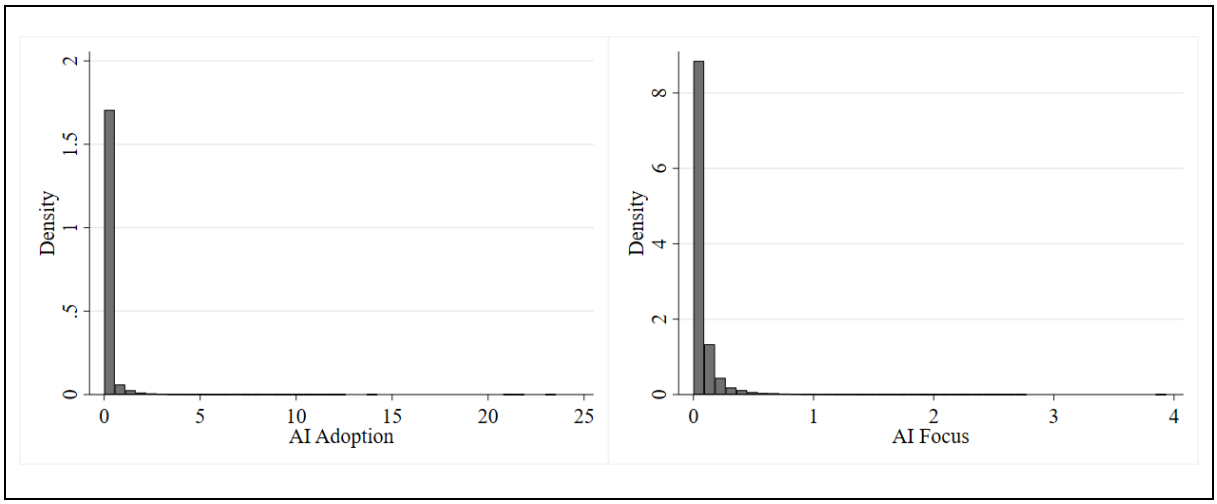


Table 6 presents the results of a t-test analysis comparing the two variables at different levels. The first t-test compares the counts of both measurements when applying their respective dictionaries to the chosen text corpus. While Mishra and colleagues' *AI Focus* considered disclosures of AI within the entire form 10-K filing text, the counts for my *AI Adoption* variable have been solely obtained from the 'Item 1. Business Description' section of the 10-K, considering only AI disclosure in sentences including at least one adoption-related term. This restriction results in, on average, 3.755 fewer AI-related terms, yielding a strongly significant t-value of -62.938. The second and third comparisons suggest that this difference is partly due to the differing compositions of AI dictionaries. When applying both dictionaries to the entire filing text without an additional adoption context, the broader AI dictionary of Mishra et al. (2022) yields, on average, 1.847 more counts than my more narrowly defined AI dictionary. This finding remains significant when applying both dictionaries to the extracted text

section ‘Item 1. Business Description’, although the difference in mean word counts is less pronounced at 0.598. T-tests four and five confirm the assumption for statistically significant deviations when counting the AI dictionary items in the whole filing text as compared to only the ‘Item 1. Business Description’ text section. Applying my AI dictionary to the entire text corpus of the Form 10-K results in, on average, 1.481 more counts than when applying it only to the extracted text section. The text corpora comparison shows the same trend when considering the *AI Focus* counts, but at a significantly higher magnitude. This demonstrates that the broader definition of AI encompasses terms that appear more frequently in other sections of the 10-K filing than within the business description section, which may have significant implications for the interpretation of the count. Lastly, the sixth t-test compares the plain count of my AI dictionary items within the business description text section and further limitation to a one-sentence adoption context. This comparison also indicates a significant difference, with the simple count considering on average 0.427 words more, which do not appear within the refined context of adoption. Overall, the t-test comparison between the two measurements at different levels of adoption refinement provides support for hypothesis one.

Table 6. Comparison between AI Adoption and AI Focus word counts at different restrictions

Comparison	Mean (1)	Mean (2)	Mean diff.	t-test
1. Overall counts of AI Adoption (1) & AI Focus (2)	1.147	4.903	-3.755	-62.938***
2. Full text counts AI Adoption (1) & AI Focus (2)	3.055	4.903	-1.847	-16.786***
3. Text section counts AI Adoption (1) & AI Focus (2)	1.574	2.172	-0.598	-12.79***
4. AI Adoption full text count (1) & AI Adoption text section count (2)	3.055	1.574	1.481	20.338***
5. AI Focus full text count (1) & AI Focus text section count (2)	4.903	2.172	2.73	71.198***
6. AI Adoption text section count (1) & AI Adoption one-sentence adoption context count (2)	1.574	1.147	0.427	24.917***

Having established an overall difference between the two measures, I further identified how these differences would relate to industry affiliation and firm characteristics. Table 7 shows an overview

of how the measures for *AI Adoption* and *AI Focus* are distributed over the twelve Fama and French industries. 'N' denotes how the entire sample is distributed among the industries, 'Count' indicates the number of terms related to *AI Adoption* (1) and *AI Focus* (2) and the columns for 'Mean' show each industries rate of *AI Adoption* (1) and *AI Focus* (2), consider the full measurement. The t-tests highlight between-industry differences by comparing each industry's mean value of the full AI measurement against the combined mean value of the other industries. While on average, the AI dictionary of Mishra et al. (2022) applied to the entire filing text produces 4.27 times more word counts than my AI dictionary applied to the business description text section within a one-sentence adoption context, clear differences between the measurements become especially apparent regarding specific industries. First, despite producing fairly large counts, the average rate of *AI Focus* in the 'Consumer Durables', 'Manufacturing', and 'Oil, Gas and Coal' does not indicate statistically significant differences from the other industries. This can certainly be attributed to these industries' minimal mean differences of 0.0004, -0.0016, and 0.0023 at p-values ranging from 35% to 48%. However, it is worth noting that 'Finance' also shows a small mean difference of 0.0089, but is significant at p = 0.0002.

Table 7. Comparison between each of twelve Fama & French industries and remaining sample

Twelve Fama & French Industries	N	Count (1)	Count (2)	Mean (1)	Mean (2)	t-test (1)	t-test (2)
Consumer Non-Durables	844	189	1,305	0.031	0.025	5.050***	9.592***
Consumer Durables	487	267	2,036	0.080	0.076	2.199**	0.057
Manufacturing	1,644	583	7,317	0.047	0.078	6.174***	-0.390
Energy	625	77	2,430	0.015	0.074	4.938***	0.349
Chemicals and Allied	506	72	945	0.023	0.031	4.164***	6.488***
Business Equipment	3,642	16,139	36,485	0.616	0.171	-47.570***	-40.614***
Telephone and Television	369	149	895	0.050	0.042	2.767***	4.241***
Utilities	373	13	665	0.004	0.020	4.107***	6.930***
Wholesale, Retail	1,726	604	3,434	0.054	0.037	5.930***	10.780***
Healthcare, Medical	4,794	3,693	16,902	0.056	0.049	10.511***	13.533***
Finance	5,108	1,906	27,690	0.042	0.070	12.533***	3.494***
Mines, Constructions	2,383	2,125	10,209	0.129	0.072	1.336*	1.395*

Besides the lack of statistical significance, these three industries, along with 'Utilities' and 'Finance', demonstrate heavily differing counts between the measures' AI dictionaries, exceeding the average factorial difference of 4.27. The especially high counts in the 'Manufacturing' and 'Finance' industries are likely due to the broad dictionary definition of the *AI Focus* variable, which covers terms that do not meet my conditions for AI. The terms 'data processing', 'modeling', and 'algorithm' are likely embedded into the core business activities of most finance companies, but amount to 31.1% of total *AI Focus* word counts. Similarly, 'automation' or 'robotic' contributed another 26.1% of word counts to *AI Focus*, and are likely a frequent subject in Form 10-K reports of manufacturing firms. The inclusion of such broad terms might also incorrectly attribute more AI to the industries 'Energy' and 'Utilities', which are usually less digitally active (Chen & Srinivasan, 2024). Lastly, both measures appear to observe AI intensely in the 'Business Equipment' industry. 91.8% of observations in this industry stem from tech companies, as defined by Chen and Srinivasan (2024). On the one hand, these certainly count as early adopters and heavy users of AI, but on the other hand, they might also integrate AI-related features into their product portfolio. Despite the logical assumption that a vendor of AI probably also employs AI in its own business, the overall adoption rate is most certainly overestimated. When tech companies are deducted from the sample, the number of firms in the 'Business Equipment' industry is reduced to 299, indicating a mean *AI Adoption* rate of 0.067 and a mean *AI Focus* rate of 0.127. Further, the mean difference for the *AI Adoption* in this industry has lost its significance. The AI-related word counts also decreased significantly to 133 for *AI Adoption* and 2,061 for *AI Focus*. Depending on how broadly AI is defined, as well as how tech companies are handled, word counts may systematically overrepresent AI in specific industries, potentially leading to false assumptions.

In an attempt to identify some of these possible misinterpretations, Table 8 presents an analysis of how selected firm characteristics are represented within the full sample as well as for each of the conflicted industries. Firms in the 'Manufacturing' and 'Utilities' industries are both significantly older than the entire sample average age, which could cause a misinterpretation of the relation between company age and AI when potentially overestimated by the *AI Focus* count. An overrepresentation of *AI Focus* counts in the 'Utilities' and 'Finance' industries could also yield false assumptions about how firm size, measured in total assets, is related to AI, as both industries' average company sizes are

significantly larger compared to the remaining sample. These two industries stand out further in terms of their reported capital expenditures. With the average value for 'Utilities' being over ten times higher than the full sample mean, conclusions regarding AI and investment activities in long-term growth could be inflated. Further, as a final variable, capital expenditures are scaled by total assets, which mitigates the extent of the effect in the case of the large-sized 'Utilities' companies.

Table 8. Variable mean values among selected industries and the full sample

	Consumer Dur.	Manu-facturing	Energy	Business Equip.	Utilities	Finance	Full Sample
Age	22	27	19	16	37	18	16
Total Assets	7,519.7	5,018.7	4,267.6	5,235.4	24,147.3	25,608	4,763.5
Total Debt	3,429.6	1,677.3	1,156.5	1,344.1	10,155.2	4,429.5	2,516.9
CapEx	275	146.6	377.3	201.8	1,712.1	43.4	167.6
Operating Income	380.9	415.4	287	548.5	1,093.1	638.9	480.5
Loss	0.302	0.224	0.354	0.457	0.072	0.095	0.483
Leverage	0.658	0.924	2.497	0.445	1.001	1.246	0.919
Growing	0.616	0.65	0.528	0.652	0.874	0.728	0.613
R&D Active	0.832	0.71	0.107	0.835	0.005	0.028	0.535
Int. Assets	0.202	0.229	0.054	0.29	0.067	0.065	0.184
N	487	1,644	625	3,642	373	5,108	22,501

On the contrary, overestimating AI in the 'Finance' industry certainly lessens the effect of capital expenditure activities on AI, especially as the already small CapEx mean value is further reduced when scaling it by the large amount of total assets. Both industries further stand out by the very small proportion of firms registering a loss, measured as the operating income after depreciation. When considering a simple indicator of growth, equaling one, if a firm's total assets have increased compared to the previous year, 87.4% of 'Utilities' firms and 72.8% of Finance firms report growth during the years 2017 to 2023, aligning with the results of profitability. But when overestimating AI for Utilities

and Finance firms, results could potentially reveal a misleading effect of profitability and growth. Lastly, the 'Energy', 'Utilities', and 'Finance' industries share a similar pattern regarding R&D activity, indicated as one, if *R&D Expenses* were observed, and *Intangible Assets*, measured as the ratio of intangible to total assets. All three industries display mean values clearly below the full sample average, which could lead to misinterpretations regarding the effects of innovativeness and digitization on AI. While the overall impact from overestimating AI in the 'Utilities' ('Energy') industry, with only 1.66% (2.78%) of sample observations, might be limited, the 'Finance' industry amounts to almost a quarter of overall observations, hence potentially jeopardizing the interpretation of explanatory effects positive effects on AI.

Applying my AI dictionary in a more conservative measurement of *AI Adoption* should counter a large proportion of potential overestimation due to the significantly reduced word counts. However, similarly to *AI Focus*, my measurement likely overemphasizes AI in the 'Business Equipment' industry, suggesting counts to be driven by vendors of artificial intelligence applications. As these observations may still contain valuable insights into the dynamics of the firm's own *AI Adoption*, I attempted a simple way of handling the *AI Adoption* counts of tech companies, rather than eliminating them from the sample. Firstly, I removed both tech company observations and observations with no *AI Adoption* count from the sample and split the remaining observations into quintiles based on their *AI Adoption* count. This produced an average *AI Adoption* count of 11.742 for the highest quintile. Secondly, I defined a new *AI Adoption* count for the entire sample, which caps the count for tech companies at this rounded-down mean value of eleven. This logic supposedly ensures that a comparatively high level of *AI Adoption* is still represented in these tech companies, but the high number of counts, possibly caused by repetitions of product or service names, is mitigated. This procedure left 9,698 *AI Adoption* counts for the 'Business Equipment' industry, putting this count number closer to the average difference ratio of the *AI Focus* count of Mishra et al. (2022). The t-statistic between the 'Business Equipment' industry mean and the remaining sample mean of the adjusted *AI Adoption* variable keeps statistical significance at -53.604.

To further investigate the difference between the two measures and whether the correction of the *AI Adoption* variable would significantly alter the results, I performed an ANOVA and a Bonferroni

post-hoc analysis to examine their relationship with the two control variables previously used in the preliminary validity test – *R&D Expenses* and *Intangible Assets.* The upper half of Table 9 presents the ANOVA results for R-squared and the F-test for each AI variable, divided into quintiles, and performed on a limited sample that excludes firms with no *AI Adoption* or *AI Focus*. The lower half of Table 9 presents the results of a subsequent Bonferroni post-hoc test, which demonstrates the differences in mean values of *R&D Expenses* and *Intangible Assets* among each of the AI quintiles. Significant differences between quintiles are displayed as bold values. A first finding is that the variable of *AI Adoption*, which was corrected for the high adoption values of tech companies, only appears to differ slightly from the original *AI Adoption* variable. This holds for both control variables, pertaining throughout the similar R-squared and F-test values, as well as the mean differences and their significance. In contrast, the already established difference between *AI Adoption* and *AI Focus* is further demonstrated via this test.

Table 9. ANOVA and Bonferroni post-hoc test of (adjusted) AI Adoption and AI Focus variables

	R&D expenditure / *Total Assets*			*Intangible Assets* / *Total Assets*		
	AI Adoption	AI Adoption (adjusted)	AI Focus	AI Adoption	AI Adoption (adjusted)	AI Focus
Obs.	4,412	4,412	15,550	4,412	4,412	15,550
R^2	0.023	0.022	0.005	0.026	0.026	0.016
F	25.93***	25.12***	19.51***	28.96***	29.14***	64.63***
Bonferroni post-hoc analysis of quintile contrast and significance (bold)						
2 vs 1	**-0.088**	**-0.088**	-0.022	**0.062**	**0.062**	0.002
3 vs 1	**-0.079**	**-0.079**	**-0.058**	**0.088**	**0.087**	0.014
4 vs 1	**-0.072**	**-0.068**	**-0.066**	**0.103**	**0.104**	0.003
5 vs 1	**-0.044**	**-0.048**	**-0.037**	**0.11**	**0.11**	**0.076**
3 vs 2	0.009	0.008	**-0.036**	0.026	0.025	0.012
4 vs 2	0.016	0.020	**-0.043**	**0.041**	**0.041**	0.002
5 vs 2	**0.044**	**0.040**	-0.015	**0.048**	**0.048**	**0.074**
4 vs 3	0.007	0.012	-0.008	0.015	0.017	-0.011
5 vs 3	**0.035**	**0.032**	0.021	0.022	0.023	**0.062**
5 vs 4	**0.028**	0.020	**0.029**	0.007	0.007	**0.073**

Compared to *AI Adoption*, *AI Focus* appears to explain a lesser amount of the variance in *R&D Expenses* and *Intangible Assets*. The F-test results between *AI Adoption* and *AI Focus* deviate notably,

but this could be partially explained by the different sample sizes. The differences in mean values further demonstrate that the rather unusual pattern of average *R&D Expenses* distributed over *AI Adoption* quintiles observed in the preliminary validity test is even more prominent for *AI Focus*. Except for quintile five, mean values appear to descend from quintile one to quintile four. While the mean values of *Intangible Assets* continuously rise throughout the quintiles of *AI Adoption*, the continuity for *AI Focus* is disrupted by the negative difference between quintiles four and three, which lacks significance, however. The significances also behave differently between the two AI measures. Overall, considering both *R&D Expenses* and *Intangible Assets*, more significant differences can be observed between quintiles of *AI Adoption* over quintiles of *AI Focus*. Especially concerning *Intangible Assets*, these significances seem to be distributed differently. For *AI Adoption*, significant differences exist between quintile one and the other quintiles, as well as in the high-contrast comparisons of upper and lower quintiles. On the contrary, *AI Focus* demonstrates significance only over the highest quintile five compared to the others. Overall, the ANOVA and post-hoc Bonferroni tests exposed further differences between the two measures. The comparison of the *AI Adoption* variable to its adjusted version suggested only marginal alterations. Hence, I will continue to use the initially proposed *AI Adoption* measurement for the remainder of this thesis.

8.3 AI Adoption and investment-related effects

Demonstrating how my proposed measurement of *AI Adoption* performs in a parametric empirical analysis, I tested hypotheses two and three using a panel regression analysis, which reflects the empirical model depicted in Section 7.2. Furthermore, this analysis presents another opportunity to compare my proposed AI measure to the *AI Focus* measure of Mishra et al. (2022). For this purpose, Table 10 presents a fixed effects panel regression that is further clustered by three-digit NAICS industries and controls for time fixed effects. As the pairwise correlation matrix in Table 4 already suggests, multicollinearity is not a concern for the regression model, with a mean (maximum) variance inflation factor of 1.34 (2.24). However, strongly significant evidence for heteroskedasticity was found when performing a Breusch–Pagan/Cook–Weisberg test. The use of robust standard errors in the panel regression should ensure consistent coefficient estimates despite heteroskedasticity. The panel regression was performed using both measurements of artificial intelligence, with columns one to three applying my measure of *AI*

Adoption as the dependent variable, and columns four to six applying the *AI Focus* measure. Similarly to Chen and Srinivasan (2024), I used one-year lagged determinant variables to explain how existing firm attributes reported for the past year influence a company's adoption or focus on AI, reported for the current year. The relatively comparable R^2-values of 0.033 for full model three and 0.037 for full model six demonstrate that the variance of both AI measurements is similarly well explained by the model. To comment on the low values, AI is certainly influenced by many other factors that are not included in or represented by Compustat business data. Additionally, the fixed effects regression analysis is likely estimating coefficients at a more conservative level.

For my second hypothesis, I evaluated the assumption that the actual adoption of AI is reflected in investment-related effects. I further expected these effects to be more pronounced than those associated with the broader *AI Focus* measure of Mishra et al. (2022). Considering my *AI Adoption* measurement, the regression analysis yielded a positive coefficient for *CAPEX* in the reduced model in column one, which is significant at a p-value of less than 0.05. Considering the effect within the full model, column three reveals a slightly increased magnitude and significance at a p-value below one percent. The effect of *CAPEX* on *AI Focus* is strongly significant in the reduced model of column four and the full model of column six. However, at about a third of the effect size, when compared to *AI Adoption*. As a second investment-related predictor, the regression revealed a negative effect of *Liquidity* on *AI Adoption*, significant at $p < 0.1$, which was not present for the *AI Focus* measure. This effect supports the intuition that the actual adoption of AI, stemming from either external procurement or internal development expenses, is associated with a firm's cash holdings. Additionally, the negative direction of the effect appears reasonable, as companies investing in long-term growth, including AI, may report reduced amounts of cash on the same year's balance sheet in which the investment was made. In contrast, *Liquidity* appears to be unrelated to a firm's sole focus on AI, which supports the presumption that the *AI Focus* measurement captures less actual investment activities related to AI. Overall, the significant coefficients of *CAPEX* and *Liquidity* suggest a relation between *AI Adoption* and investment-related effects. Regarding *AI Focus*, the reduced magnitude of *CAPEX* and the lack of significance in *Liquidity* support my claim that these investment-related effects arise from measuring the actual adoption of AI.

Table 10. Fixed effects panel regression analysis of AI determinants

	(1) AI Adoption	(2) AI Adoption	(3) AI Adoption	(4) AI Focus	(5) AI Focus	(6) AI Focus
CAPEX $_{i,t-1}$	0.00018**		0.00020***	0.00006***		0.00006***
	(0.000)		(0.000)	(0.000)		(0.000)
Liquidity $_{i,t-1}$		-0.00440*	-0.00441*		0.00018	0.00018
		(0.002)	(0.002)		(0.000)	(0.000)
ROA $_{i,t-1}$	0.00017***	0.00020***	0.00015***	-0.00027***	-0.00026***	-0.00027***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
R&D Expenses $_{i,t-1}$	-0.00093***	-0.00114***	-0.00125***	-0.00028***	-0.00024***	-0.00027***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Employees $_{i,t-1}$	-0.00004	0.00107	0.00107	0.00013	0.00008	0.00009
	(0.002)	(0.002)	(0.002)	(0.001)	(0.001)	(0.001)
Age $_{i,t-1}$	0.00340	0.00472	0.00471	-0.00390	-0.00395	-0.00395
	(0.020)	(0.021)	(0.021)	(0.004)	(0.004)	(0.004)
Size $_{i,t-1}$	0.00155	0.00168	0.00168	0.00060**	0.00060*	0.00060*
	(0.001)	(0.001)	(0.001)	(0.000)	(0.000)	(0.000)
Leverage $_{i,t-1}$	0.00002	0.00002	0.00002	-0.00001	-0.00001	-0.00001
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
SG&A Expenses $_{i,t-1}$	0.00004	0.00001	-0.00004	-0.00005*	-0.00003*	-0.00004
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Intangible Assets $_{i,t-1}$	0.05189***	0.04834**	0.04842**	0.00938	0.00950	0.00952
	(0.019)	(0.020)	(0.020)	(0.011)	(0.011)	(0.011)
Tangible Assets $_{i,t-1}$	-0.01079	-0.01351	-0.01345	0.00116	0.00125	0.00127
	(0.010)	(0.012)	(0.012)	(0.002)	(0.002)	(0.002)
HHI $_{i,t-1}$	0.49945	0.50066	0.50067	0.02790	0.02785	0.02785
	(0.315)	(0.315)	(0.315)	(0.034)	(0.034)	(0.034)
Risk $_{i,t-1}$	-0.00182	-0.00081	-0.00082	-0.00009	-0.00013	-0.00013
	(0.010)	(0.010)	(0.010)	(0.002)	(0.002)	(0.002)
Constant	-0.00901	0.02670	0.02673	0.05678***	0.05534***	0.05535***
	(0.075)	(0.069)	(0.069)	(0.011)	(0.012)	(0.012)
Time FE	Yes	Yes	Yes	Yes	Yes	Yes
Industry Clustering	Yes	Yes	Yes	Yes	Yes	Yes
Observations	21,939	21,939	21,939	21,939	21,939	21,939
R^2	0.032	0.033	0.033	0.037	0.037	0.037
Number of firms	4,630	4,630	4,630	4,630	4,630	4,630

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Regarding additional determinants, both AI measures appear to be explained by a firm's profitability, measured by return on assets (*ROA*). Columns one to three demonstrate a continuously positive effect of *ROA* on AI Adoption at a p-value of below 0.01, suggesting that higher levels of profitability can be associated with increased *AI Adoption* rates. Due to the strongly significant, yet negative, effects of ROA on *AI Focus* observed in columns five and six, this statement cannot be made when considering the broader measurement of artificial intelligence. These different directions of both measures' *ROA* coefficients further underscore the disparity between the narrow *AI Adoption* measurement and the broad *AI Focus* variable, providing further evidence that supports hypothesis one. Additionally, both measurements appear to be related to the negative effects of *R&D expenses*. The prior validity test, as well as the Bonferroni post-hoc test, already suggested this dynamic, as higher levels of *R&D expenses* were observed for the lowest quintiles of *AI Adoption* and *AI Focus*. Chen and Srinivasan (2024) also found that *R&D Expenses* are negatively related to their dependent variable of Digital Activity. A significant relation between firm size and AI was only found at a slight degree for the *AI Focus* measurement. Lastly, the regression analysis revealed a positive relation between *Intangible Assets* and *AI Adoption*, significant at $p < 0.05$. Considering the broader *AI Focus* measurement yielded no significant relation to *Intangible Assets*. Similarly to *R&D Expenses*, this association was already suggested by the preliminary validity test and the Bonferroni post-hoc test. The significant positive coefficient of *Intangible Assets* in the regression analysis further strengthens the validity of the measurement and supports the idea of narrowly defining AI within an adoption context to capture actual firm-level activities of artificial intelligence.

To address the concern that the results are influenced by tech companies, I performed an additional regression analysis on a sample that excludes observations from companies of tech industries as defined by Chen and Srinivasan (2024). The results are presented in Appendix H. The positive effect of *CAPEX* on *AI Adoption* remains stable, indicating the same level of magnitude and significance in full model three. The exclusion of tech companies further introduced significance to the effect of *CAPEX* on *AI Focus*, which was previously not present in the full model six. However, the coefficient's size remains far lower when compared to model three. The negative effects of *R&D Expenses* on AI retain similar levels of magnitude and significance, with additional significance introduced in model four. The

relation between *Size* and *AI Focus* holds with similar magnitude and increased significance at $p < 0.05$. Additional significance was also introduced to the association between increased *AI Focus* and an averagely younger *Age* of companies. However, the exclusion of tech companies appears to have masked the effects of *ROA* on both AI measures, the effect of *Leverage* on *AI Focus*, and the association between *AI Adoption* and *Liquidity,* as well as *Intangible Assets*. Lastly, at an R^2-value of 0.016, the model's explanatory power of the variance in *AI Adoption* appears to have suffered more compared to the *AI Focus* measure at an R^2 of 0.031. While the support for hypothesis two does not seem to depend entirely on the word counts of tech companies, due to the vanished significance of *Liquidity*, caution is advised when interpreting investment-relatedness. Despite similarly pronounced results for *CAPEX* and *R&D Expenses*, the differences in R^2 as well as the significant effects of *Age* and *Size* on *AI Focus* keep both AI measures distinguished, which continues to support the initial confirmation of hypothesis one.

9 Discussion and conclusion

Overall, the results demonstrate a significant difference between the prior existing broadly defined measure of AI Focus of Mishra et al. (2022) and my newly developed, narrowly defined measurement of AI Adoption. The t-tests revealed the sources of the differences across the differently defined AI dictionaries, the varying considerations of text corpora, and the restriction to an adoption context. It was also established that the difference between the two measures leads to different industry-specific weightings, which may have implications for inferences on AI, especially regarding crucial company and financial characteristics. While the temporal development of both measures over the time period revealed a similar pattern of the underlying word counts, it was shown that these word counts highly deviate from each other, underlining the importance of the dictionary definition. The results also provide further support for the idea presented in previous studies, which suggests selectively concentrating textual analysis on a specific text section based on the research subject (Campbell et al., 2014; Chen & Srinivasan, 2024; Cheong et al., 2021; Mayew et al., 2015). In addition, the further restriction to only consider AI terms within an adoption context was also found to differ significantly from the text section-based count. While this has already been done in previous studies that also established an improvement from considering target terms within a character-based surrounding context (Balvers et al., 2016; Campbell et al., 2014), applying the enhanced capability of identifying and splitting distinct sentences with the 'nltk' module in Python certainly improves accuracy, while still accounting for replicability. Combining both of these restrictions has likely led to the enhanced association between the AI Adoption variable and investment-related effects such as capital expenditure or cash amount. While the business data from Compustat is limited in indicating which exact expenditure type includes specific IT expenses, such as investments in artificial intelligence, the relationship between AI Adoption and capital expenditure might still provide a representative estimation. Following Enache et al. (2022) in their interpretation of capital expenditures activities, the results suggest that firms signaling an orientation towards future growth are likely to do so by adopting AI into their operations. At the same time, an actual investment in AI, especially if a company is already spending on existing assets, should be reflected in respective cash movements, which was already found to be associated with investment activities by An et al. (2016). Further, the restriction to adoption-related AI terms could also explain the

significant effect of intangible assets on AI Adoption. This relationship appears reasonable, as it had been suggested before, that AI relies on prior enabling technologies (Zhang & Lu, 2021) and, resembled in its machine learning capabilities, was therefore found to be highly associated with other digital activities such as data science, big data, blockchain, data mining and cloud computing (Goldfarb et al., 2023; Lee et al., 2022). Along the multitude of different tests, the general difference between AI Adoption and AI Focus was demonstrated, as well as the underlying causes. It was also possible to obtain a better understanding of the possible implications resulting from these different properties and how they might affect the interpretation of certain effects. Finally, this difference was further confirmed in a parametric test and remained even when excluding non-tech companies. Thus, hypothesis one is overall supported by empirical evidence. This parametric test also enabled the investigation of investment-related effects on AI Adoption and AI Focus. On the one hand, it provided basic empirical evidence suggesting the existence of investment-related relationships with AI Adoption. On the other hand, it has been shown that a mere focus on AI is associated with a lower level, or even a lack of, investment-related effects. Although these effects only partially hold in a separate sample of non-tech firms, the overall empirical results support a confirmation of hypothesis two.

This thesis, like previous studies, demonstrates the tremendous potential of a textual analysis of Form 10-K reports to provide researchers with a rich and largely objective data source for analyzing different topics and producing large panel datasets. Due to the selection of different topic sections in the report and the relatively large remaining text corpus, it is also possible to incorporate additional refinements, such as considering an adoption context. This, beyond the simple method of counting terms in the entire document, can enhance the overall precision of the final measurement. Thus, this thesis was able to show that, in addition to the narrower definition of the topic-specific AI dictionary, the additional restriction to the business description section and the condition of an existing adoption context in the same sentence led to significantly different results and at the same time enabled a more precise interpretation of the determinants of actual AI Adoption. Furthermore, this thesis has addressed the existing limitations of various research methods for measuring AI and contributes to a reduction of these constraints. The possibility of automated data extraction from a large number of companies over a long time horizon represents an enormous advantage over the mostly cross-sectional and often smaller-sized

survey-based data sets. As the previously discussed results from Lee et al. (2022) suggested, the impact of AI is more notable at higher levels of adoption. As my measurement constitutes an intensity rate of AI adoption, its added value over a simple indicator variable further qualifies the measure for more research applications. However, limitations remain for the empirical method of textual analysis, which cannot be eliminated completely, even with additional sophistication. First, surveys certainly still allow a more precise measurement of the topic and permit a more detailed consideration of aspects beyond the firm level. This also enables the examination of other potential drivers of AI Adoption that cannot be derived from reportable business figures. Second, although the textual analysis of Form 10-K reports permits a large number of observations, these are restricted to large publicly listed companies from the United States. Therefore, only limited inferences can be drawn from the results about the total population of companies in a country or an entire economy. In contrast, McElheran et al. (2024) were able to analyze numerous small and medium-sized companies in addition to their large-scale study on AI in America. Third, as proposed in Section 8.1, the measurement may be exposed to shocks, like the COVID-19 pandemic, which might shift firms' focus from reporting on tech investments to how they handle the current situation and react to market changes. While it can be argued that such events can be classified as systematic risks, affecting the entire economy, they could also unsystematically impact different industries. Integrating the overall word count into the final measurement could lead to decreasing rates, due to possibly less focus on disclosing AI Adoption within an increased text section length. This could lead to a misrepresentation of a firm's actual AI activities. Fourth and last, referring to the importance of domain-specific dictionaries established by Henry (2006) and Loughran and McDonald (2011), the analysis of technical topics such as AI might be limited when considering the initial purpose of the text corpus. In the case of Form 10-K reports, they are primarily aimed at investors. Therefore, the amount of technological jargon is likely kept to a lower level so that the content remains comprehensible to the target group. Consequently, the various technical facets of AI can only be observed to a limited extent, and interpretations need to be kept rather general. The regression analysis in a reduced sample of non-tech companies has also shown that the results are affected by the industry-specific sample composition.

Given the sum of these persisting limitations inherent in this empirical methodology, it is advisable to cautiously consider and interpret the results within the bigger picture of empirical findings

about AI. However, as time progresses and AI matures, AI Adoption counts in 10-K reports can be expected to increase accordingly, which could provide additional reliability and validity to the metric, especially if more non-tech companies report an integration of AI into their operations. As one of many possible further research projects, a comprehensive comparison with the results of survey-based research could therefore be carried out in order to validate the reliability of the measure more thoroughly. Additionally, the metric could and should be continually revisited to enhance its precision further. In particular, the AI dictionary needs to be updated with the terminology of the continuously evolving technology. At the same time, the question arises as to whether every single existing term used to measure AI in earlier reports remains meaningful for capturing AI activities in more recent reports, or if it might increase the degree of noise in the measurement. In both cases, the document type's level of technicality should again be kept in mind. Additional considerations could also be made regarding industries, for example, by assigning them dedicated AI dictionaries that better capture the industry-specific applications of AI. However, when aiming for improvements or additional refinements of the measure, researchers should not neglect the importance of replicability, which has already been stressed by Henry and Leone (2016). In order to keep up with the rapidly developing field of AI, a high degree of collaboration among researchers is required, which would certainly be facilitated by appropriately comprehensible measurement methods. Lastly, given the various economic and social aspects that will continue to be affected by AI in the future, the proposed metric might apply to a wide range of research questions that rely on quantitative data. To conclude, it can be stated that while certain limitations remain, the newly developed AI Adoption measure constitutes a novelty and contribution to existing methods within empirical AI literature. It is now on future research to continue evolving the measure and applying it in various models, in order to continually sharpen our understanding of AI and its impact.

10 References

Acemoglu, D., Autor, D., Hazell, J., & Restrepo, P. (2022). Artificial Intelligence and Jobs: Evidence from Online Vacancies. *Journal of Labor Economics*, *40*(S1), S293–S340. https://doi.org/10.1086/718327

Acemoglu, D., & Restrepo, P. (2020). The wrong kind of AI? Artificial intelligence and the future of labour demand. *Cambridge Journal of Regions, Economy and Society*, *13*(1), 25–35. https://doi.org/10.1093/cjres/rsz022

Agarwal, S., Chen, V. Y. S., & Zhang, W. (2016). The Information Value of Credit Rating Action Reports: A Textual Analysis. *Management Science*, *62*(8), 2218–2240. https://doi.org/10.1287/mnsc.2015.2243

Alderucci, D., Branstetter, L., College, H., Hovy, E., & Runge, A. (2020). *Quantifying the Impact of AI on Productivity and Labor Demand: Evidence from U.S. Census Microdata*.

Amoozegar, A., Berger, D., Cao, X., & Pukthuanthong, K. (2020). Earnings conference calls and institutional monitoring: Evidence from textual analysis. *Journal of Financial Research*, *43*(1), 5–36. https://doi.org/10.1111/jfir.12199

An, H., Chen, Y., Luo, D., & Zhang, T. (2016). Political uncertainty and corporate investment: Evidence from China. *Journal of Corporate Finance*, *36*, 174–189. https://doi.org/10.1016/j.jcorpfin.2015.11.003

Andreou, P. C., Harris, T., & Philip, D. (2020). Measuring Firms' Market Orientation Using Textual Analysis of 10-K Filings. *British Journal of Management*, *31*(4), 872–895. https://doi.org/10.1111/1467-8551.12391

Antweiler, W., & Frank, M. Z. (2004). Is All That Talk Just Noise? The Information Content of Internet Stock Message Boards. *The Journal of Finance*, *59*(3), 1259–1294. https://doi.org/10.1111/j.1540-6261.2004.00662.x

Babina, T., Fedyk, A., He, A., & Hodson, J. (2024). Artificial intelligence, firm growth, and product innovation. *Journal of Financial Economics*, *151*, 103745. https://doi.org/10.1016/j.jfineco.2023.103745

Bäck, A., Hajikhani, A., Jäger, A., Schubert, T., & Suominen, A. (2022). Return of the Solow-paradox in AI? AI-adoption and firm productivity. *Papers in Innovation Studies*, Article 2022/1. https://ideas.repec.org//p/hhs/lucirc/2022_001.html

Balasubramanian, N., Ye, Y., & Xu, M. (2022). Substituting Human Decision-Making with Machine Learning: Implications for Organizational Learning. *Academy of Management Review*, *47*(3), 448–465. https://doi.org/10.5465/amr.2019.0470

Balvers, R. J., Gaski, J. F., & McDonald, B. (2016). Financial Disclosure and Customer Satisfaction: Do Companies Talking the Talk Actually Walk the Walk? *Journal of Business Ethics*, *139*(1), 29–45. https://doi.org/10.1007/s10551-015-2612-6

Bass, A. S. (2018, March 28). *Non-tech businesses are beginning to use artificial intelligence at scale*. The Economist. https://www.economist.com/special-report/2018/03/28/non-tech-businesses-are-beginning-to-use-artificial-intelligence-at-scale

Blei, D. M., Ng, A. Y., & Jordan, M. I. (2003). Latent Dirichlet Allocation. *Journal of Machine Learning Research*, *3*, 993–1022. https://doi.org/10.1162/jmlr.2003.3.4-5.993

Brown, T. B., Mann, B., Ryder, N., Subbiah, M., Kaplan, J., Dhariwal, P., Neelakantan, A., Shyam, P., Sastry, G., Askell, A., Agarwal, S., Herbert-Voss, A., Krueger, G., Henighan, T., Child, R., Ramesh, A., Ziegler, D. M., Wu, J., Winter, C., … Amodei, D. (2020). *Language Models are Few-Shot Learners* (arXiv:2005.14165). arXiv. https://doi.org/10.48550/arXiv.2005.14165

Campbell, J. L., Chen, H., Dhaliwal, D. S., Lu, H., & Steele, L. B. (2014). The information content of mandatory risk factor disclosures in corporate filings. *Review of Accounting Studies*, *19*(1), 396–455. https://doi.org/10.1007/s11142-013-9258-3

Chang, Y., Wang, X., Wang, J., Wu, Y., Yang, L., Zhu, K., Chen, H., Yi, X., Wang, C., Wang, Y., Ye, W., Zhang, Y., Chang, Y., Yu, P. S., Yang, Q., & Xie, X. (2023). *A Survey on Evaluation of Large Language Models* (arXiv:2307.03109). arXiv. https://doi.org/10.48550/arXiv.2307.03109

Chatterjee, S., Rana, N. P., Dwivedi, Y. K., & Baabdullah, A. M. (2021). Understanding AI adoption in manufacturing and production firms using an integrated TAM-TOE model. *Technological Forecasting and Social Change*, *170*, 120880. https://doi.org/10.1016/j.techfore.2021.120880

Chen, W., & Srinivasan, S. (2024). Going digital: Implications for firm value and performance. *Review of Accounting Studies*, *29*(2), 1619–1665. https://doi.org/10.1007/s11142-023-09753-0

Cheong, A., Yoon, K., Cho, S., & No, W. G. (2021). Classifying the Contents of Cybersecurity Risk Disclosure through Textual Analysis and Factor Analysis. *Journal of Information Systems*, *35*(2), 179–194. https://doi.org/10.2308/ISYS-2020-031

Choudhary, V., Marchetti, A., Shrestha, Y. R., & Puranam, P. (2025). Human-AI Ensembles: When Can They Work? *Journal of Management*, *51*(2), 536–569. https://doi.org/10.1177/01492063231194968

Cui, R., Li, M., & Zhang, S. (2022). AI and Procurement. *Manufacturing & Service Operations Management*, *24*(2), 691–706. https://doi.org/10.1287/msom.2021.0989

Czarnitzki, D., Fernández, G. P., & Rammer, C. (2023). Artificial intelligence and firm-level productivity. *Journal of Economic Behavior & Organization*, *211*, 188–205. https://doi.org/10.1016/j.jebo.2023.05.008

Damioli, G., Van Roy, V., & Vertesy, D. (2021). The impact of artificial intelligence on labor productivity. *Eurasian Business Review*, *11*(1), 1–25. https://doi.org/10.1007/s40821-020-00172-8

Davenport, T., Guha, A., Grewal, D., & Bressgott, T. (2020). How artificial intelligence will change the future of marketing. *Journal of the Academy of Marketing Science*, *48*(1), 24–42. https://doi.org/10.1007/s11747-019-00696-0

Di Vaio, A., Palladino, R., Hassan, R., & Escobar, O. (2020). Artificial intelligence and business models in the sustainable development goals perspective: A systematic literature review. *Journal of Business Research*, *121*, 283–314. https://doi.org/10.1016/j.jbusres.2020.08.019

Doshi, A. R., Bell, J. J., Mirzayev, E., & Vanneste, B. S. (2024). Generative artificial intelligence and evaluating strategic decisions. *Strategic Management Journal*, *46*(3), 583–610. https://doi.org/10.1002/smj.3677

Enache, L., Fogel-Yaari, H., & Li, H. (2022). Signalling long-term focus through textual emphasis on innovation: Are firms putting their money where their mouth is? *Accounting & Finance*, *62*(3), 3791–3836. https://doi.org/10.1111/acfi.12905

Engelberg, J. (2008). Costly Information Processing: Evidence from Earnings Announcements. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.1107998

Feldman, R., Govindaraj, S., Livnat, J., & Segal, B. (2010). Management's tone change, post earnings announcement drift and accruals. *Review of Accounting Studies*, *15*(4), 915–953. https://doi.org/10.1007/s11142-009-9111-x

Felten, E., Raj, M., & Seamans, R. (2021). Occupational, industry, and geographic exposure to artificial intelligence: A novel dataset and its potential uses. *Strategic Management Journal*, *42*(12), 2195–2217. https://doi.org/10.1002/smj.3286

Felten, E., Raj, M., & Seamans, R. C. (2019). The Effect of Artificial Intelligence on Human Labor: An Ability-Based Approach. *Academy of Management Proceedings*. https://doi.org/10.5465/AMBPP.2019.140

Fornell, C., Johnson, M. D., Anderson, E. W., Cha, J., & Bryant, B. E. (1996). The American Customer Satisfaction Index: Nature, Purpose, and Findings. *Journal of Marketing*, *60*(4), 7–18. https://doi.org/10.1177/002224299606000403

Frankel, R., Jennings, J., & Lee, J. (2022). Disclosure Sentiment: Machine Learning vs. Dictionary Methods. *Management Science*, *68*(7), 5514–5532. https://doi.org/10.1287/mnsc.2021.4156

Gardner, H. (1999). *Intelligence reframed: Multiple intelligences for the 21st century*. Hachette Uk.

Goldfarb, A., Taska, B., & Teodoridis, F. (2020). Artificial Intelligence in Health Care? Evidence from Online Job Postings. *AEA Papers and Proceedings*, *110*, 400–404. https://doi.org/10.1257/pandp.20201006

Goldfarb, A., Taska, B., & Teodoridis, F. (2023). Could machine learning be a general purpose technology? A comparison of emerging technologies using data from online job postings. *Research Policy*, *52*(1), 104653. https://doi.org/10.1016/j.respol.2022.104653

Guha, A., Grewal, D., Kopalle, P. K., Haenlein, M., Schneider, M. J., Jung, H., Moustafa, R., Hegde, D. R., & Hawkins, G. (2021). How artificial intelligence will affect the future of retailing. *Journal of Retailing*, *97*(1), 28–41. https://doi.org/10.1016/j.jretai.2021.01.005

Hanley, K. W., & Hoberg, G. (2010). The Information Content of IPO Prospectuses. *Review of Financial Studies*, *23*(7), 2821–2864. https://doi.org/10.1093/rfs/hhq024

Harris, Z. S. (1954). Distributional Structure. *Word*, *10*(2–3), 146–162. https://doi.org/10.1080/00437956.1954.11659520

Heaven, W. D. (2021, December 21). *2021 was the year of monster AI models*. MIT Technology Review. https://www.technologyreview.com/2021/12/21/1042835/2021-was-the-year-of-monster-ai-models/

Henry, E. (2006). Market Reaction to Verbal Components of Earnings Press Releases: Event Study Using a Predictive Algorithm. *Journal of Emerging Technologies in Accounting*, *3*(1), 1–19. https://doi.org/10.2308/jeta.2006.3.1.1

Henry, E. (2008). Are Investors Influenced By How Earnings Press Releases Are Written? *Journal of Business Communication*, *45*(4), 363–407. https://doi.org/10.1177/0021943608319388

Henry, E., & Leone, A. J. (2016). Measuring Qualitative Information in Capital Markets Research: Comparison of Alternative Methodologies to Measure Disclosure Tone. *The Accounting Review*, *91*(1), 153–178. https://doi.org/10.2308/accr-51161

Hu, W., Shohfi, T., & Wang, R. (2021). What's really in a deal? Evidence from textual analysis of M&A conference calls. *Review of Financial Economics*, *39*(4), 500–521. https://doi.org/10.1002/rfe.1126

Huang, M.-H., & Rust, R. T. (2018). Artificial Intelligence in Service. *Journal of Service Research*, *21*(2), 155–172. https://doi.org/10.1177/1094670517752459

Huang, X., Teoh, S. H., & Zhang, Y. (2014). Tone Management. *The Accounting Review*, *89*(3), 1083–1113. https://doi.org/10.2308/accr-50684

Kaplan, A., & Haenlein, M. (2019). Siri, Siri, in my hand: Who's the fairest in the land? On the interpretations, illustrations, and implications of artificial intelligence. *Business Horizons*, *62*(1), 15–25. https://doi.org/10.1016/j.bushor.2018.08.004

Law, K. K. F., & Mills, L. F. (2015). Taxes and Financial Constraints: Evidence from Linguistic Cues. *Journal of Accounting Research*, *53*(4), 777–819. https://doi.org/10.1111/1475-679X.12081

Lee, Y. S., Kim, T., Choi, S., & Kim, W. (2022). When does AI pay off? AI-adoption intensity, complementary investments, and R&D strategy. *Technovation*, *118*, 1–15. https://doi.org/10.1016/j.technovation.2022.102590

Lewis, D. D. (1998). Naive (Bayes) at forty: The independence assumption in information retrieval. In C. Nédellec & C. Rouveirol (Eds.), *Machine Learning: ECML-98* (Vol. 1398, pp. 4–15). Springer Berlin Heidelberg. https://doi.org/10.1007/BFb0026666

Li, F. (2008). Annual report readability, current earnings, and earnings persistence. *Journal of Accounting and Economics*, *45*(2–3), 221–247. https://doi.org/10.1016/j.jacceco.2008.02.003

Li, F. (2010). The Information Content of Forward-Looking Statements in Corporate Filings—A Naïve Bayesian Machine Learning Approach. *Journal of Accounting Research*, *48*(5), 1049–1102. https://doi.org/10.1111/j.1475-679X.2010.00382.x

Li, F., Lundholm, R., & Minnis, M. (2013). A Measure of Competition Based on 10-K Filings. *Journal of Accounting Research*, *51*(2), 399–436. https://doi.org/10.1111/j.1475-679X.2012.00472.x

Li, K., Mai, F., Shen, R., & Yan, X. (2021). Measuring Corporate Culture Using Machine Learning. *The Review of Financial Studies*, *34*(7), 3265–3315. https://doi.org/10.1093/rfs/hhaa079

Li, Y., & Zhang, L. (2015). Short Selling Pressure, Stock Price Behavior, and Management Forecast Precision: Evidence from a Natural Experiment. *Journal of Accounting Research*, *53*(1), 79–117. https://doi.org/10.1111/1475-679X.12068

Lipsey, R. G., Carlaw, K. I., & Bekar, C. T. (2005). *Economic Transformations: General Purpose Technologies and Long-Term Economic Growth*. Oxford University PressOxford. https://doi.org/10.1093/oso/9780199285648.001.0001

Liu, P., Yuan, W., Fu, J., Jiang, Z., Hayashi, H., & Neubig, G. (2023). Pre-train, Prompt, and Predict: A Systematic Survey of Prompting Methods in Natural Language Processing. *ACM Computing Surveys*, *55*(9), 1–35. https://doi.org/10.1145/3560815

Loughran, T., & McDonald, B. (2011). When Is a Liability Not a Liability? Textual Analysis, Dictionaries, and 10-Ks. *The Journal of Finance*, *66*(1), 35–65. https://doi.org/10.1111/j.1540-6261.2010.01625.x

Loughran, T., & McDonald, B. (2014). Measuring Readability in Financial Disclosures. *The Journal of Finance*, *69*(4), 1643–1671. https://doi.org/10.1111/jofi.12162

Loureiro, S. M. C., Guerreiro, J., & Tussyadiah, I. (2021). Artificial intelligence in business: State of the art and future research agenda. *Journal of Business Research*, *129*, 911–926. https://doi.org/10.1016/j.jbusres.2020.11.001

Mayew, W. J., Sethuraman, M., & Venkatachalam, M. (2015). MD&A Disclosure and the Firm's Ability to Continue as a Going Concern. *The Accounting Review*, *90*(4), 1621–1651. https://doi.org/10.2308/accr-50983

McCarthy, J., Minsky, M. L., Rochester, N., & Shannon, C. E. (2006). A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence: August 31, 1955. *AI Magazine*, *27*(4), 12–14. ProQuest One Business; Social Science Premium Collection.

McElheran, K., Li, J. F., Brynjolfsson, E., Kroff, Z., Dinlersoz, E., Foster, L., & Zolas, N. (2024). AI adoption in America: Who, what, and where. *Journal of Economics & Management Strategy*, *33*(2), 375–415. https://doi.org/10.1111/jems.12576

McKinsey & Company. (2023a). *The economic potential of generative AI: the next productivity frontier*. QuantumBlack AI by McKinsey. https://www.mckinsey.com/capabilities/mckinsey-digital/our-insights/the-economic-potential-of-generative-ai-the-next-productivity-frontier#/

McKinsey & Company. (2023b). *The state of AI in 2023: Generative AI's breakout year*. QuantumBlack AI by McKinsey. https://www.mckinsey.com/capabilities/quantumblack/our-insights/the-state-of-ai-in-2023-generative-ais-breakout-year

Michalisin, M. D., & White, G. P. (2000). Validity of annual report assertions about quality: An empirical study. *The Mid - Atlantic Journal of Business*, *36*(2/3), 103–122.

Mikalef, P., & Gupta, M. (2021). Artificial intelligence capability: Conceptualization, measurement calibration, and empirical study on its impact on organizational creativity and firm performance. *Information & Management*, *58*(3), 103434. https://doi.org/10.1016/j.im.2021.103434

Mikolov, T., Sutskever, I., Chen, K., Corrado, G., & Dean, J. (2013). *Distributed Representations of Words and Phrases and their Compositionality* (arXiv:1310.4546). arXiv. https://doi.org/10.48550/arXiv.1310.4546

Mishra, S., Ewing, M. T., & Cooper, H. B. (2022). Artificial intelligence focus and firm performance. *Journal of the Academy of Marketing Science*, *50*(6), 1176–1197. https://doi.org/10.1007/s11747-022-00876-5

Mishra, S., Ewing, M. T., & Pitt, L. F. (2020). The effects of an articulated customer value proposition (CVP) on promotional expense, brand investment and firm performance in B2B markets: A text based analysis. *Industrial Marketing Management*, *87*, 264–275. https://doi.org/10.1016/j.indmarman.2019.10.005

Moody, C. E. (2016). *Mixing Dirichlet Topic Models and Word Embeddings to Make lda2vec* (arXiv:1605.02019). arXiv. https://doi.org/10.48550/arXiv.1605.02019

Pandey, S., & Pandey, S. K. (2019). Applying Natural Language Processing Capabilities in Computerized Textual Analysis to Measure Organizational Culture. *Organizational Research Methods*, *22*(3), 765–797. https://doi.org/10.1177/1094428117745648

Pham, P., Zhang, H., Gao, W., & Zhu, X. (2024). Determinants and performance outcomes of artificial intelligence adoption: Evidence from U.S. Hospitals. *Journal of Business Research*, *172*, 114402. https://doi.org/10.1016/j.jbusres.2023.114402

Rammer, C., Fernández, G. P., & Czarnitzki, D. (2022). Artificial intelligence and industrial innovation: Evidence from German firm-level data. *Research Policy*, *51*(7), 104555. https://doi.org/10.1016/j.respol.2022.104555

Rogers, J. L., Van Buskirk, A., & Zechman, S. L. C. (2011). Disclosure Tone and Shareholder Litigation. *The Accounting Review*, *86*(6), 2155–2183. https://doi.org/10.2308/accr-10137

Securities and Exchange Commission. (2005). Securities and exchange commission final rule. *Release No. 33–8591 (FR-75)*. https://www.sec.gov/files/rules/final/33-8591.pdf

Securities and Exchange Commission. (2025, January). *Investor Bulletin: How to Read a 10-K*. Investor.Gov. https://www.investor.gov/introduction-investing/general-resources/news-alerts/alerts-bulletins/investor-bulletins/how-read

Shi, G., Sun, J., & Zhang, L. (2018). Product market competition and earnings management: A firm-level analysis. *Journal of Business Finance & Accounting*, *45*(5–6), 604–624. https://doi.org/10.1111/jbfa.12300

Stanford University. (2024). *AI Index Report 2024*. Stanford Institute for Human-Centered Artificial Intelligence. https://hai.stanford.edu/ai-index/2024-ai-index-report

Stone, P. J., Bales, R. F., Namenwirth, J. Z., & Ogilvie, D. M. (1962). The general inquirer: A computer system for content analysis and retrieval based on the sentence as a unit of information. *Behavioral Science*, *7*(4), 484–498. https://doi.org/10.1002/bs.3830070412

Syam, N., & Sharma, A. (2018). Waiting for a sales renaissance in the fourth industrial revolution: Machine learning and artificial intelligence in sales research and practice. *Industrial Marketing Management*, *69*, 135–146. https://doi.org/10.1016/j.indmarman.2017.12.019

Tambe, P., & Hitt, L. M. (2012). The Productivity of Information Technology Investments: New Evidence from IT Labor Data. *Information Systems Research*, *23*(3-part-1), 599–617. https://doi.org/10.1287/isre.1110.0398

Taulli, T. (2018, December 22). *2018's Biggest Moments In AI (Artificial Intelligence)*. Forbes. https://www.forbes.com/sites/tomtaulli/2018/12/22/2018s-biggest-moments-in-ai-artificial-intelligence/

Tetlock, P. C. (2007). Giving Content to Investor Sentiment: The Role of Media in the Stock Market. *The Journal of Finance*, *62*(3), 1139–1168. https://doi.org/10.1111/j.1540-6261.2007.01232.x

Tetlock, P. C., Saar-Tsechansky, M., & Macskassy, S. (2008). More Than Words: Quantifying Language to Measure Firms' Fundamentals. *The Journal of Finance*, *63*(3), 1437–1467. https://doi.org/10.1111/j.1540-6261.2008.01362.x

Van Roy, V., Vertesy, D., & Damioli, G. (2020). AI and Robotics Innovation. In K. F. Zimmermann (Ed.), *Handbook of Labor, Human Resources and Population Economics* (pp. 1–35). Springer International Publishing. https://doi.org/10.1007/978-3-319-57365-6_12-2

Wang, H., & Qiu, F. (2023). AI adoption and labor cost stickiness: Based on natural language and machine learning. *Information Technology and Management*, *26*(2), 163–184. https://doi.org/10.1007/s10799-023-00408-9

Webb, M. (2020). The Impact of Artificial Intelligence on the Labor Market. *Available at SSRN 3482150*. https://web.stanford.edu/~mww/webb jmp.pdf

Webb, M., Short, N., Bloom, N., & Lerner, J. (2018). *Some Facts of High-Tech Patenting* (24793; p. w24793). National Bureau of Economic Research. https://doi.org/10.3386/w24793

Yang, C.-H. (2022). How Artificial Intelligence Technology Affects Productivity and Employment: Firm-level Evidence from Taiwan. *Research Policy*, *51*(6), 1–12. https://doi.org/10.1016/j.respol.2022.104536

Zhang, C., & Lu, Y. (2021). Study on artificial intelligence: The state of the art and future prospects. *Journal of Industrial Information Integration*, *23*, 1–9. https://doi.org/10.1016/j.jii.2021.100224

Zhang, Q., Wang, A., & Li, R. (2024). Enterprise value creation effects of artificial intelligence technology from the perspective of digital agility: Evidence from China. *Technology Analysis & Strategic Management*, 1–15. https://doi.org/10.1080/09537325.2024.2383605

Zhong, Y., Xu, F., & Zhang, L. (2024). Influence of artificial intelligence applications on total factor productivity of enterprises—Evidence from textual analysis of annual reports of Chinese-listed companies. *Applied Economics*, *56*(43), 5205–5223. https://doi.org/10.1080/00036846.2023.2244246

Appendix A. Lists of tables and figures

List of tables

List of figures

Appendix B. AI dictionary sources

Reports:

Deloitte. (2024). State of Generative AI in the Enterprise 2024 | Deloitte US. https://www.deloitte.com/us/en/what-we-do/capabilities/applied-artificial-intelligence/content/state-of-generative-ai-in-enterprise.html

McKinsey & Company. (2023b). The state of AI in 2023: Generative AI's breakout year. McKinsey & Company. https://www.mckinsey.com/capabilities/quantumblack/our-insights/the-state-of-ai-in-2023-generative-ais-breakout-year

McKinsey & Company. (2024). The state of AI in early 2024 | McKinsey. https://www.mckinsey.com/capabilities/quantumblack/our-insights/the-state-of-ai-2024

PricewaterhouseCoopers. (2017). PwC's Global Artificial Intelligence Study: Sizing the prize. PwC. https://www.pwc.com/gx/en/issues/artificial-intelligence/publications/artificial-intelligence-study.html

Standford University. (2024). AI Index Report 2024. Stanford Institute for Human-Centered Artificial Intelligence. https://hai.stanford.edu/ai-index/2024-ai-index-report

Glossaries:

Council of Europe. (n.d.). Glossary—Artificial Intelligence—Www.coe.int. Artificial Intelligence. https://www.coe.int/en/web/artificial-intelligence/glossary

Coursera. (n.d.). Generative AI Definitions: A to Z Glossary Terms. https://www.coursera.org/collections/generative-ai-terms?utm_medium=sem&utm_source=gg&utm_campaign=B2C_EMEA__coursera_FTCOF_career-academy_pmax-multiple-audiences-country-multi&campaignid=20858198824&adgroupid=&device=c&keyword=&matchtype=&network=x&devicemodel=&adposition=&creativeid=&hide_mobile_promo&gad_source=1&gclid=CjwKCAjwmaO4BhAhEiwA5p4YLz9hnlYFmsyWEYgdF1h64hZwY5v4D3DgzQl8Ah8CNBiejJzChvRrIxoC-6EQAvD_BwE

Nielsen Norman Group. (2024). *Artificial Intelligence: Glossary*. https://www.nngroup.com/articles/artificial-intelligence-glossary/

Sino-German Company. (2020). AI Glossary. Deutsche Gesellschaft für Internationale Zusammenarbeit (GIZ) GmbH, China Center for Information Industry Development (CCID). https://www.plattform-i40.de/IP/Redaktion/DE/Downloads/Publikation/China/ai-glossary.pdf?__blob=publicationFile&v=2

UK Parliament. (2024). Artificial intelligence (AI) glossary. https://post.parliament.uk/artificial-intelligence-ai-glossary/

Appendix C. AI dictionary

AI Term	REGEX	Counts		
		Item 1. + Adoption	**Item 1.**	**Entire Filing Text**
Adversarial learning	\badversarial learn\w*	0	0	0
Adversarial network	\badversarial network\w*	6	7	7
AI	(?<!\w)AI(?![\w(-])|\b\(AI\)\b|\bAI-\w+	13,264	18,696	41,790
AIops	\baiops\b	73	110	128
Artificial general intelligence	\bartificial general intelligence\b	0	1	1
Artificial intelligence	\bartificial intelligen\w*	4,662	6,208	14,158
Artificial reality	\bartificial realit\w*	12	12	12
Augmented intelligence	\baugmented intelligen\w*	95	108	142
Augmented reality	\baugmented realit\w*	518	794	1,373
Automated decision-making	\bautomated decision-making\b	98	106	189
Automated intelligence	\bautomated intelligence\b	11	11	15
Automated reporting	\bautomated reporting\b	9	19	27
Automatic speech recognition	\bautomatic speech recognition\b	35	49	52
Autonomous agent	\bautonomous agent\b	0	1	1
Autonomous intelligence	\bautonomous intelligen\w*	1	1	3
Bayesian model	\bbayesian model\w*	4	4	4
Bayesian network	\bbayesian network\w*	0	1	1
Biometric	\bbiometric\b	1,939	3,130	5,432
Chain-of-thought prompt	\bchain-of-thought prompt\w*	0	0	0
Chatbot	\bchatbot\w*	188	245	363
Cognitive automation	\bcognitive automation\b	14	18	18
Cognitive computing	\bcognitive computing\b	29	46	81
Computational neuroscience	\bcomputational neuroscience\b	0	0	3
Computer vision	\bcomputer vision\b	604	822	1,384
Conversational agent	\bconversational agent\w*	0	2	2
Deep learning	\bdeep learn\w*	502	709	954
Evolutionary computing	\bevolutionary comput\w*	0	0	1

Face recognition	\bface recogni\w*	37	57	63
Facial recognition	\bfacial recogni\w*	198	305	441
Few-shot prompt	\bfew-shot prompt\w*	0	0	0
Generative adversarial network	\bgenerative adversarial network\w*	6	6	6
Generative model	\bgenerative model\w*	0	1	5
Generative pretrained transformer	\bgenerative pretrained transformer\w*	0	0	0
Genetic algorithm	\bgenetic algorithm\w*	3	3	5
Gesture recognition	\bgesture recogni\w*	29	29	43
Image recognition	\bimage recogni\w*	36	59	81
Intelligent data analysis	\bintelligent data analy\w*	11	11	12
Intelligent document processing	\bintelligent document processing\b	3	7	9
Intelligent system	\bintelligent system\w*	97	112	233
Knowledge-based system	\bknowledge-based system\w*	0	0	0
Large language model	\blarge language model\b	13	18	41
Long short-term memory	\blong short-term memory\b	0	0	0
Machine intelligence	\bmachine intelligen\w*	73	92	222
Machine learning	\bmachine learn\w*	5,199	6,885	11,122
Machine vision	\bmachine vision\b	308	394	575
Mlops	\bmlops\b	20	26	37
Multi-shot prompt	\bmulti-shot prompt\w*	0	0	0
Natural language generation	\bnatural language generation\b	1	6	6
Natural language processing	\bnatural language processing\b	242	331	472
Neural network	\bneural network\w*	395	571	667
Neuro-linguistic programming	\bneuro-linguistic programm\w*	3	3	4
Pattern recognition	\bpattern recogni\w*	141	187	253
Predictive analysis	\bpredictive analy\w*	1,143	1,433	2,066
Predictive modeling	\bpredictive model\w*	624	791	1,407
Prescriptive analysis	\bprescriptive analy\w*	58	68	104
Probabilistic modeling	\bprobabilistic model\w*	10	11	38

Prompt engineering	\bprompt engineer\w*	0	1	2
Quantum computing	\bquantum comput\w*	1,190	1,889	3,717
Quantum generative modeling	\bquantum generative model\w*	0	0	0
Recommendation engine	\brecommendation engine\w*	117	189	244
Reinforcement learning	\breinforcement learn\w*	17	22	26
Retrieval-augmented generation	\bretrieval-augmented generation\b	0	0	0
Semi-supervised learning	\bsemi-supervised learn\w*	0	0	0
Sentiment analysis	\bsentiment analy\w*	44	54	61
Single-shot prompt	\bsingle-shot prompt\w*	0	0	0
Speech recognition	\bspeech recogni\w*	206	275	387
Speech synthesis	\bspeech synthesis\b	7	7	7
Statistical learning	\bstatistical learn\w*	3	8	8
Superintelligence	\bsuperintelligen\w*	1	1	1
Supervised learning	\bsupervised learn\w*	4	6	8
Tokenization	\btokenization\b	236	290	583
Transfer learning	\btransfer learning\b	2	3	3
Unsupervised learning	\bunsupervised learn\w*	3	3	6
Variational autoencoder	\bvariational autoencoder\b	0	0	0
Virtual agent	\bvirtual agent\w*	82	129	154
Virtual assistant	\bvirtual assistant\w*	276	406	468
Virtual reality	\bvirtual realit\w*	604	1,003	1,640
Voice recognition	\bvoice recogni\w*	84	106	260
Zero-shot prompt	\bzero-shot prompt\w*	0	0	0
	Total Counts	**33,590**	**46,898**	**91,628**

Appendix D. Adoption dictionary

Adoption Term	REGEX
Accept	\baccept\w*
Achieve	\bachiev\w*
Acquire	\bacquir\w*
Actualize	\bactualiz\w*
Adapt	\badapt\w*
Add	\badd\b
Added	\badded\b
Adding	\badding\b
Adds	\badds\b
Adopt	\badopt\w*
Affiliate	\baffiliat\w*
Affirm	\baffirm\w*
Applie	\bapplie\w*
Apply	\bapply\w*
Appropriate	\bappropriat\w*
Approve	\bapprov\w*
Assent	\bassent\w*
Assume	\bassum\w*
Bargain	\bbargain\w*
Borrow	\bborrow\w*
Carried out	\bcarried out\b
Carries out	\bcarries out\b
Carry out	\bcarry out\b
Carrying out	\bcarrying out\b
Complete	\bcomplet\w*
Contract	\bcontract\w*
Deploy	\bdeploy\w*

Adoption Term	REGEX
Develop	\bdevelop\w*
Effect	\beffect\w*
Embrace	\bembrac\w*
Employ	\bemploy\b
Employed	\bemployed\b
Employing	\bemploying\b
Employs	\bemploys\b
Enable	\benabl\w*
Endorse	\bendors\w*
Enforce	\benforc\w*
Espouse	\bespous\w*
Execute	\bexecut\w*
Focus	\bfocus\w*
Follow	\bfollow\w*
Fulfill	\bfulfill\w*
Get	\bget\b
Gets	\bgets\b
Getting	\bgetting\b
Got	\bgot\b
Imitate	\bimitat\w*
Implement	\bimplement\w*
Include	\binclude\w*
Incorporate	\bincorporat\w*
Integrate	\bintegrat\w*
Invest	\binvest\w*
Invoke	\binvok\w*
Involve	\binvolve\w*

Adoption Term	REGEX
Leverage	\bleverag\w*
Made possible	\bmade possible\b
Maintain	\bmaintain\w*
Make possible	\bmake possible\b
Makes possible	\bmakes possible\b
Making possible	\bmaking possible\b
Materialize	\bmaterializ\w*
Mimic	\bmimic\w*
Obtain	\bobtain\w*
Opt	\bopt\b
Opted	\bopted\b
Opting	\bopting\b
Opts	\bopts\b
Paid for	\bpaid for\b
Pay for	\bpay for\b
Paying for	\bpaying for\b
Pays for	\bpays for\b
Perform	\bperform\w*
Procure	\bprocur\w*
Provide	\bprovid\w*
Purchase	\bpurchas\w*
Put into effect	\bput into effect\b
Puts into effect	\bputs into effect\b
Putting into effect	\bputting into effect\b
Ratify	\bratif\w*
Realize	\brealiz\w*
Redeem	\bredeem\w*
Resolve	\bresolv\w*

Adoption Term	REGEX
Seize	\bseiz\w*
Select	\bselect\w*
Shop for	\bshop for\b
Shopped for	\bshopped for\b
Shopping for	\bshopping for\b
Shops for	\bshops for\b
Sign for	\bsign for\b
Signed for	\bsigned for\b
Signing for	\bsigning for\b
Signs for	\bsigns for\b
Support	\bsupport\w*
Take on	\btake on\b
Take over	\btake over\b
Takes on	\btakes on\b
Takes over	\btakes over\b
Taking on	\btaking on\b
Taking over	\btaking over\b
Tap into	\btap into\b
Tapped into	\btapped into\b
Tapping into	\btapping into\b
Taps into	\btaps into\b
Took on	\btook on\b
Took over	\btook over\b
Use	\buse\b
Used	\bused\b
Uses	\buses\b
Using	\busing\b
Utilize	\butiliz\w*

Appendix E. Examples of AI terms with and without adoption context

Example of AI term occurrences in 'Item 1.' and adoption context

"Advanced analytics and AI to accelerate the pace of learning: We utilize AI to enable various parts of our platform and drug discovery. Examples include:

- Neural networks for protein engineering: One way to optimize the efficacy of the proteins encoded by our mRNA is to engineer the sequence of the protein itself. We use neural networks to analyze and model protein sequences. We train these models by inputting orthologous sequences from thousands of organisms, from which we can generate potential protein sequences optimized for specific attributes.
- Neural networks for mRNA engineering: The redundancy in the genetic code allows for a large number of mRNA sequences that encode the same protein. mRNA sequence may impact translation, thereby impacting the amount of protein produced in circulation. We are developing AI tools to predict mRNA sequences that can enhance protein expression."

Source: Moderna, Inc. (2023) Form 10-K, p.36.
https://www.sec.gov/Archives/edgar/data/1682852/000168285223000011/mrna-20221231.htm

Example of AI term occurrences in 'Item 1.' without adoption context

"The power of big data and machine learning is evident across the modern economy, but there is reason to believe their impact on insurance will run deeper than elsewhere. On e-commerce sites, for example, artificial intelligence can help recommend a new camera, but the device that arrives 2 hours later is indistinguishable from the one being sold on main street. Algorithms propose songs and movies on entertainment apps, but Sgt. Pepper sounds the same today as it did in 1967. In short, most industries ultimately sell either atoms or electrons, and these are largely impervious to machine learning. The core product of insurance is different: it is a probability algorithm, and data and artificial intelligence can absolutely transform algorithms. In insurance, data science is not helpful to the business, it is the business."

Source: Lemonade, Inc. (2022) Form 10-K, p.9.
https://www.sec.gov/Archives/edgar/data/1691421/000169142122000015/lmnd-20211231.htm

Examples of AI term occurrences outside of 'Item 1.' (here: 'Item 1A. Risk Factors')

"The applications of artificial intelligence and machine learning, including those in the investment and financial sectors, continue to develop rapidly, and it is impossible to predict all of the future risks that may arise from such developments. We cannot control the use of artificial intelligence or machine learning in our portfolio companies or third-party products or services and therefore could be exposed to associated risks if our portfolio companies, third-party service providers or any counterparties use artificial intelligence or machine learning in their business activities."

Source: Hercules Capital, Inc. (2024) Form 10-K, p.50.
https://www.sec.gov/Archives/edgar/data/1280784/000128078424000007/htgc-20231231.htm

"In the EU, a number of new laws related to digital data and AI have recently entered into force or have been proposed. For example, on December 8, 2023, EU legislators reached agreement on the AI Act, which imposes regulatory requirements onto AI system providers, importers, distributors, and users of AI systems, in accordance with the level of risk involved with the AI system. The UK has not adopted formal legislation to regulate AI but has adopted guidelines in the form of a White Paper."

Source: Omnicell, Inc. (2024) Form 10-K, p.11.
https://www.sec.gov/Archives/edgar/data/926326/000092632624000008/omcl-20231231.htm

Appendix F. Underlying Python code

Routine 1. Assembly of filing URL

```
import pandas as pd
import numpy as np
import json
import os

# Define form type and filing years
form_type = '10-K'
filing_years = [2024, 2023, 2022, 2021, 2020, 2019, 2018, 2017]

# Load initial Compustat export to dataframe and transform CIK identifier
companyData = pd.read_excel("CCM_Export_2017-2023.xlsx")
companyData['cik_str_zeros'] = companyData['CIK'].astype(str).str.zfill(10)

# Create a cartesian product of companyData and expanded_years
expanded_years = pd.DataFrame(filing_years, columns=['year'])
expanded_years['key'] = 1  # Dummy key for merging
companyData['key'] = 1
companyData = companyData.merge(expanded_years, on='key').drop('key', axis=1)

# Source of .json metadata files (store in same folder)
submissions_path = 'submissions_to12062025'

# Cache for storing fetched data
filing_cache = {}

# Add new columns for the fetched and derived data
new_cols = ['form','filingDate','reportDate','accessionNumber','primaryDocument','filing_url']
for col in new_cols:
    companyData[col] = np.nan

# Iterate through the dataframe rows
for index, row in companyData.iterrows():
    cik = row['CIK']
    cik_zeros = row['cik_str_zeros']
    year = row['year']

    try:
        if cik_zeros not in filing_cache:
            # Get company-specific filing metadata only if not cached
            with open(os.path.join(submissions_path, f'CIK{cik_zeros}.json'), 'r') as file:
                    filingMetadata = json.load(file)
            recentForms = pd.DataFrame.from_dict(filingMetadata['filings']['recent'])
            allFiles = pd.DataFrame.from_dict(filingMetadata['filings']['files'])

            allForms = recentForms.copy()

            for file_idx, file_row in allFiles.iterrows():
                file_name = file_row['name']
                with open(os.path.join(submissions_path, f'{file_name}'), 'r') as file:
                        filingMetadata2 = json.load(file)
                olderForms = pd.DataFrame.from_dict(filingMetadata2)
                allForms = pd.concat([allForms, olderForms], ignore_index=True)

            filing_cache[cik_zeros] = allForms
        else:
            # Use cached metadata
            allForms = filing_cache[cik_zeros]

        # Filter filings within the specific year and form type
        filings_from = f"{year}-01-01"
        filings_to = f"{year}-12-31"
        allForms_filter = allForms[
            (allForms['form'] == form_type) &
            (allForms['filingDate'] <= filings_to) &
            (allForms['filingDate'] >= filings_from)
        ]

        if not allForms_filter.empty:
            filtered_data = allForms_filter.iloc[0]
            for col in new_cols[:-1]:  # Exclude 'filing_url' for now
```

```
                companyData.loc[index, col] = filtered_data[col]

            # Construct and store the filing URL
            clean_accession = filtered_data['accessionNumber'].replace('-', '')
            companyData.loc[index,'filing_url']=
f"https://www.sec.gov/Archives/edgar/data/{cik}/{clean_accession}/{filtered_data['primaryDocumen
t']}"

    except FileNotFoundError:
        print(f"File not found for CIK: {cik_zeros}. Skipping to next record.")
        continue
    except Exception:
        print(f"An error occurred for CIK: {cik_zeros}.")
        continue

companyData.to_excel(f"CCM_SEC_{form_type}_filingurls_{filing_years[-1]}-
{filing_years[0]}.xlsx", index=False)

print("Script executed sucessfully!")
```

Routine 2. Extraction of filing text and text section

```
import requests
import pandas as pd
from bs4 import BeautifulSoup
import re
import time
import nltk
from nltk.tokenize import sent_tokenize

# Define identity to access EDGAR database
headers = {'User-Agent': "email-address"}

# Define form type and filing years
form_type = '10-K'
filing_years = [2024, 2023, 2022, 2021, 2020, 2019, 2018, 2017]

# Protect 10 requests per second rule (current limitation from SEC)
request_interval = 0.1
last_request_time = 0

# Load the existing dataframe that was exported in routine 1
existing_df = pd.read_excel("CCM_SEC_10-K_filing urls_2017-2024.xlsx")

# Define start and end points for text extraction with regular expressions
item_1dot_pattern = re.compile(r'\bITEM\s+1\.\s', re.I)  # strict: Item 1.
item_1space_pattern = re.compile(r'\bITEM\s+1\s', re.I)   # strict: Item 1 (space)
item_1a_dot_pattern = re.compile(r'\bITEM\s+1A\.\s', re.I)
item_1a_space_pattern = re.compile(r'\bITEM\s+1A\s', re.I)
item_1a_double_dot_pattern = re.compile(r'\bITEM\s+1\.A\.', re.I)
item_2_pattern = re.compile(r'\bITEM\s+2[\.]?\s', re.I)

# Define forbidden context words around Item mentions
forbidden_patterns = [r"to", r"as", r"in", r"see", r"from", r"under", r"and", r"or", r","]

# Helper function to check forbidden context
def is_valid_match(text, match_start):
    window_size = 30
    start = max(0, match_start - window_size)
    context = text[start:match_start].lower()
    nearest_dot = context.rfind('.')
    nearest_forbidden = -1

    for pattern in forbidden_patterns:
        match = re.search(pattern, context)
        if match:
            nearest_forbidden = max(nearest_forbidden, match.start())

    # If there's a dot and it appears after any forbidden word -> allow match
    if nearest_dot != -1 and (nearest_forbidden == -1 or nearest_dot > nearest_forbidden):
        return True

    # If forbidden word is closer than any dot or no dot exists -> disallow match
```

```
    return not any(re.search(pattern, context) for pattern in forbidden_patterns)

# Section extraction function
def extract_section(text, year):
    def find_section(start_patterns, end_patterns, text):
        start_matches = []
        for pattern in start_patterns:
            start_matches.extend([m for m in pattern.finditer(text) if is_valid_match(text,
m.start())])

        end_matches = []
        for pattern in end_patterns:
            end_matches.extend([m for m in pattern.finditer(text) if is_valid_match(text,
m.start())])

        if len(start_matches) > 1:
            start_pos = start_matches[1].start()
        elif len(start_matches) == 1:
            start_pos = start_matches[0].start()
        else:
            return "", 0

        if len(end_matches) > 1:
            end_pos = end_matches[1].start()
        elif len(end_matches) == 1:
            end_pos = end_matches[0].start()
        else:
            return "", 0

        if end_pos > start_pos:
            section = text[start_pos:end_pos].strip()
            return section, len(section.split())
        else:
            return "", 0

    start_patterns = [item_1dot_pattern, item_1space_pattern]
    end_patterns = [item_1a_dot_pattern, item_1a_space_pattern, item_1a_double_dot_pattern]

    text_section, text_section_word_count = find_section(start_patterns, end_patterns, text)

    if text_section_word_count >= 100:
        return text_section, text_section_word_count

    fallback_patterns = [item_2_pattern] if year < 2005 else []
    text_section, text_section_word_count = find_section(start_patterns, fallback_patterns,
text)

    if text_section_word_count >= 100:
        return text_section, text_section_word_count

    print(f"Short text section, {text_section_word_count} words.")
    return "", 0

# Iterate over the 'filing_url' in existing_df starting from the specified index
start_index = 1

def make_request_with_retry(url, headers):
    while True:
        try:
            # Check when last request was made
            global last_request_time
            now = time.time()
            elapsed = now - last_request_time

            # Sleep only if the last request was too recent
            if elapsed < request_interval:
                time.sleep(request_interval - elapsed)

            # Access filing via url and identity
            response = requests.get(url, headers=headers)
            if response.status_code == 200:
                last_request_time = time.time()  # update after the request is made
                return response
            elif response.status_code == 429:
                # Wait 15 minutes if being temporarly blocked
```

```
                print("Waiting for access to SEC EDGAR.")
                time.sleep(15 * 60)
            else:
                print(f"Unexpected status code {response.status_code} for URL: {url}")
                return None

        except requests.ConnectionError:
            # Wait 2 minutes if internet connection was interrupted
            print("Waiting for internet connection.")
            time.sleep(2 * 60)

# Iterate through the dataframe rows
for index, row in existing_df.iloc[start_index:].iterrows():
    filing_url = row['filing_url']
    year = row['year']
    # Check if filing_url is nan or empty
    if pd.isna(filing_url) or filing_url.strip() == "":
        print(f"Skipping empty or invalid URL at index {index}")
        continue  # Skip to the next iteration if the URL is not valid

    # call url request function
    response = make_request_with_retry(filing_url, headers)

    if response:
        # Choose the way of extracting the entire text depending on the file type
        if filing_url.endswith('.txt'):
            if '<html>' in response.text.lower():
                soup = BeautifulSoup(response.text, 'html.parser')
                filing_text = soup.get_text(separator=' ', strip=True)
            else:
                filing_text = response.text
        else:
            soup = BeautifulSoup(response.content, 'html.parser')
            filing_text = soup.get_text(separator=' ', strip=True)

        # Store the entire text corpus and its word count
        entire_text_word_count = len(filing_text.split())
        existing_df.loc[index, 'entire_text'] = filing_text
        existing_df.loc[index, 'entire_text_word_count'] = entire_text_word_count

        # Call text section extraction function and get the word count
        text_section, text_section_word_count = extract_section(filing_text, year)

        # Tokenize the text_section into individual sentences
        if text_section:
            section_sentences = sent_tokenize(text_section)
            # Convert the list of sentences into a single string for storage
            section_sentences_str = "\n".join(section_sentences)

        # Output the result for each document
        print(f"Index: {index} | Text section length: {text_section_word_count} of
{entire_text_word_count} words for URL: {filing_url}")

        # ensure proper lenght of text section before storing in the dataframe
        if text_section_word_count >= 2000:
            existing_df.loc[index, 'text_section'] = text_section
            existing_df.loc[index, 'section_sentences'] = section_sentences_str
            existing_df.loc[index, 'text_section_word_count'] = text_section_word_count

        else:
            print(f"Extracted text section is too short ({text_section_word_count} words),
likely incorrect.")
            existing_df.loc[index, 'text_section'] = None

    else:
        print(f"Failed to retrieve data from {filing_url}")
        existing_df.loc[index, 'text_section'] = None

# Export the data frame to .csv for later processing and to .xlsx for direct access
existing_df.to_csv(f"CCM_SEC_{form_type}_Filings_{filing_years[-1]}-{filing_years[0]}_Filing
Text and Item 1 Section.csv", index=False)
existing_df.to_excel(f"CCM_SEC_{form_type}_Filings_{filing_years[-1]}-{filing_years[0]}_Filing
Text and Item 1 Section.xlsx", index=False)

print("Script executed successfully!")
```

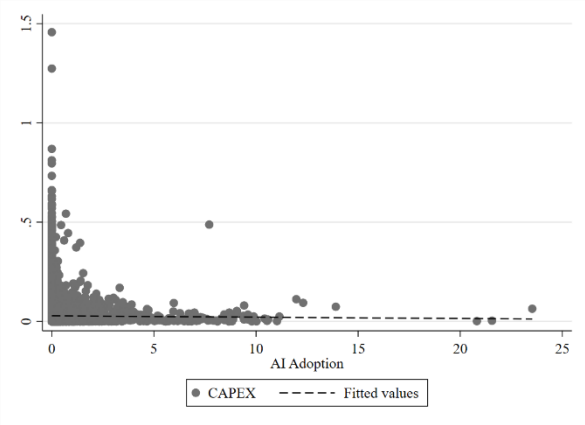
Routine 3. Parsing and counting of dictionary terms

```
import pandas as pd
import re
from openpyxl import load_workbook

# Define form type and filing years
form_type = '10-K'
filing_years = [2024, 2023, 2022, 2021, 2020, 2019, 2018, 2017]

# Define import file, export file, and chunksize
INPUT_FILE = 'CCM_SEC_10-K_Filings_2024-2025_Filing Text and Item 1 Section_15062025.csv'
OUTPUT_FILE = 'CCM_SEC_10-K_Filings_2024-2025_Dict Count_15062025.xlsx'
CHUNKSIZE = 39000

# Define AI and Adoption dictionaries with REGEX and initialize counter with "0"
ai_adoption_dict_1sent = {
    r"(?<!\w)AI(?![\w(-])|\b\(AI\)\b|\bAI-\w+": 0,
    r"\baiops\b": 0,
    r"\bartificial general intelligence\b": 0,
    ...
    r"\brecommendation engine\w*": 0
}

adoption_dict = {
    r"\badopt\w*": 0,
    r"\baccept\w*": 0,
    r"\bapprov\w*": 0,
    ...
    r"\bdeploy\w*": 0
}

# Check at which index previous chunk ended
def get_last_row(filename, sheetname):
    wb = load_workbook(filename)
    ws = wb[sheetname]
    return ws.max_row

# Helper function to count occurrences of terms in a text and update a dictionary
def count_terms(text, term_dict):
    count = 0
    for term in term_dict.keys():
        matches = len(re.findall(term, text, re.IGNORECASE))
        count += matches
        term_dict[term] += matches  # Update the dictionary count dynamically
    return count

# Proccess each chunk
chunk_num = 0
for chunk in pd.read_csv(INPUT_FILE, chunksize=CHUNKSIZE):

    print(f"// Processing chunk {chunk_num}...")

    chunk['adoption_count'] = 0
    chunk['ai_adoption_count_1sent'] = 0
    chunk['section_sentences'] = chunk['section_sentences'].fillna("")

    # Iterate through each row in the dataframe
    for idx, row in chunk.iterrows():
        section_sentences = str(row['section_sentences']) if
pd.notnull(row['section_sentences']) else ""
        entire_text = row['entire_text']
        text_section_word_count = row['text_section_word_count']
        entire_text_word_count = row['entire_text_word_count']

        # Check if section_sentences is nan or empty
        if pd.isna(section_sentences) or section_sentences.strip() == "":
            print(f"Skipping empty or invalid section at index {idx}")
            continue  # Skip to the next iteration if the section_sentences is not valid

        # Check if the word count is at least 2000
        if text_section_word_count < 2000 or entire_text_word_count < 2000:
            print(f"Index: {idx} | Extracted text section is too short
({text_section_word_count}/{entire_text_word_count} words), no dict count.")
            continue  # Skip to the next iteration if the text is too short
```

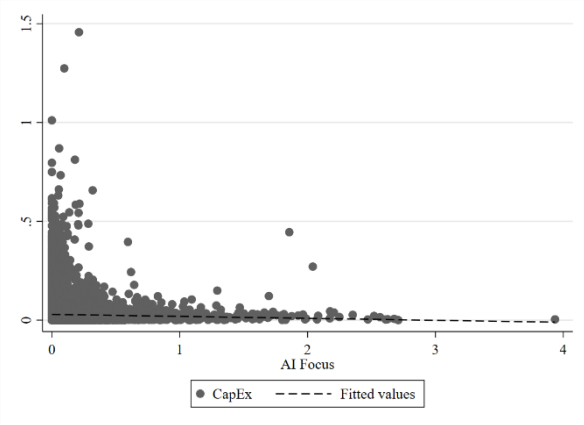
```
        # Count AI Adoption terms, considering a one sentence adoption context
        ai_adoption_count_1sent = 0
        sentences = section_sentences.split('\n')
        for i, sentence in enumerate(sentences):
            # Check if any term from adoption_dict is in the same sentence
            adoption_context_1sent = (
                any(re.search(term, sentence, re.IGNORECASE) for term in adoption_dict.keys())
            )

            # Update the dictionary count dynamically
            if adoption_context_1sent:
                for term in ai_adoption_dict_1sent.keys():
                    matches = len(re.findall(term, sentence, re.IGNORECASE))
                    ai_adoption_count_1sent += matches
                    ai_adoption_dict_1sent[term] += matches

        # Update AI Adoption count in dataframe
        chunk.at[idx, 'ai_adoption_count_1sent'] = ai_adoption_count_1sent
        print(f"Index: {idx} | AI Adoption: {ai_adoption_count_1sent}")

    # drop large text columns for a lighter export
    chunk.drop(columns=['text_section', 'section_sentences', 'entire_text_section'],
inplace=True, errors='ignore')

    # Write chunk to Excel (append after first chunk)
    if chunk_num == 0:
        chunk.to_excel(OUTPUT_FILE, index=False)
    else:
        start_row = get_last_row(OUTPUT_FILE, 'Sheet1')  # Or your correct sheet name
        with pd.ExcelWriter(OUTPUT_FILE, engine='openpyxl', mode='a', if_sheet_exists='overlay')
as writer:
            chunk.to_excel(writer, index=False, header=False, startrow=start_row)
    chunk_num += 1

# Create a separate DataFrame to summarize dictionary counts
dictionary_summary = pd.DataFrame()

# Add each dictionary's terms and counts as columns
for dict_name, dictionary in zip([
    'ai_adoption_dict_1sent'
], [
    ai_adoption_dict_1sent
]):
    dict_df = pd.DataFrame({
        f"{dict_name}_terms": list(dictionary.keys()),
        f"{dict_name}_counts": list(dictionary.values())
    })
    dictionary_summary = pd.concat([dictionary_summary, dict_df], axis=1)

# Export the summary DataFrame to an Excel file
dictionary_summary.to_excel("dictionary_summary2024.xlsx", index=False)

print("Script executed successfully!")
```

Appendix G. Scatter plots of variables

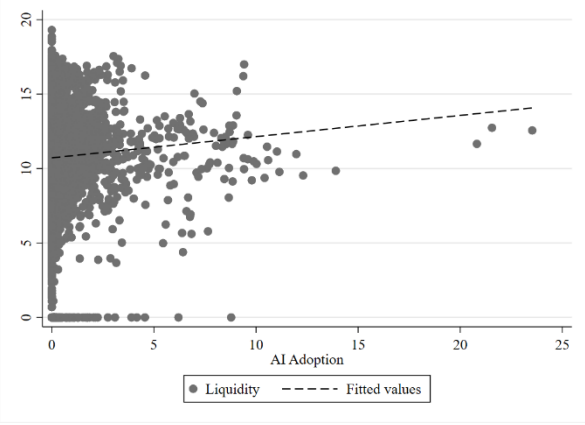
AI Adoption & CAPEX plot



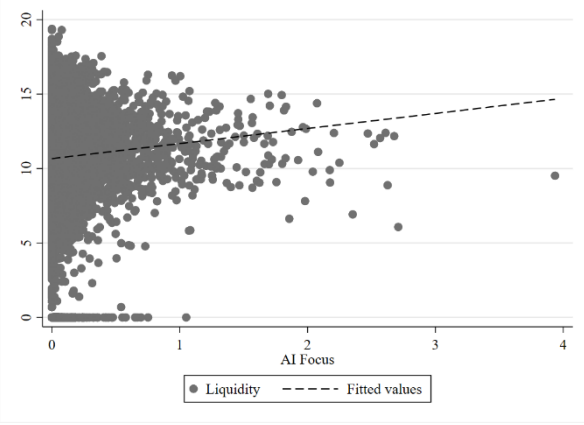
AI Focus & CAPEX plot



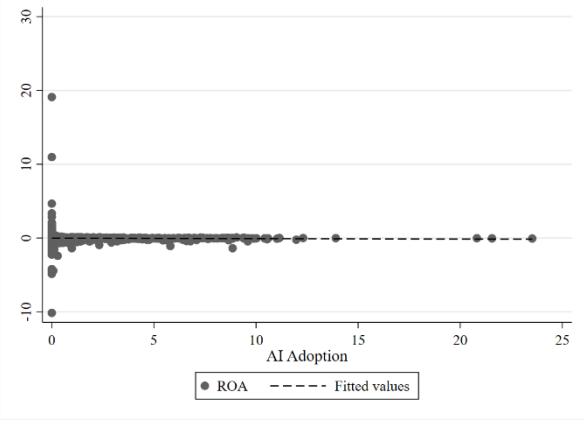
AI Adoption & Liquidity plot



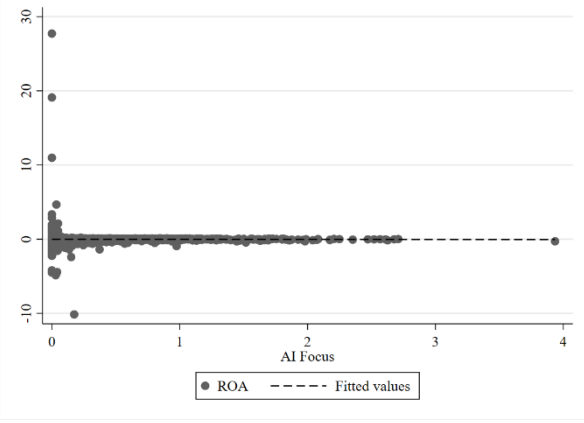
AI Focus & Liquidity plot



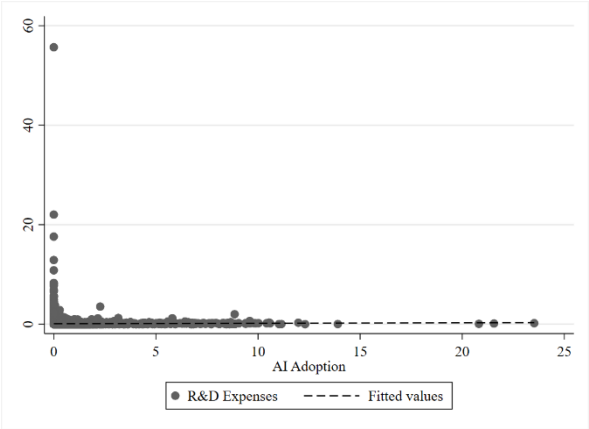
AI Adoption & ROA plot



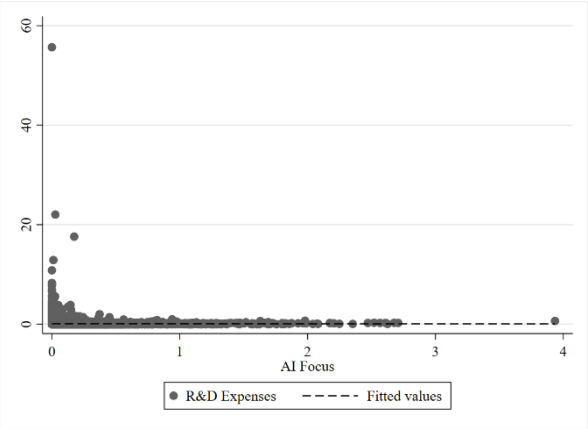
AI Focus & ROA plot



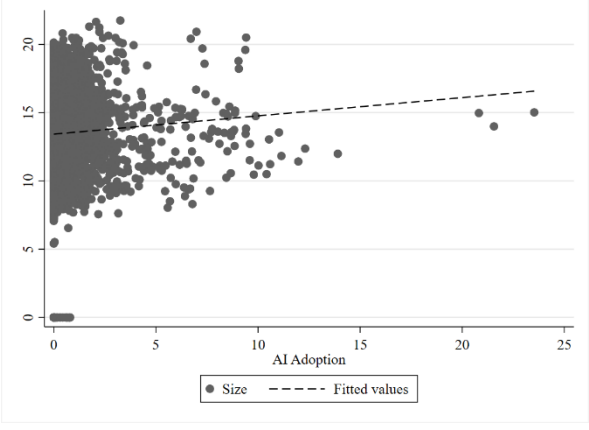
AI Adoption & R&D Expenses plot



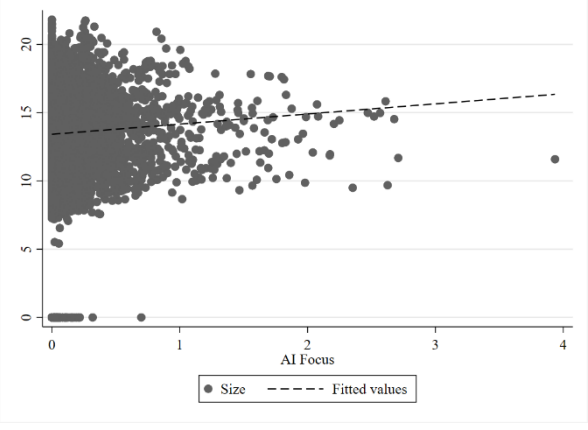
AI Focus & R&D Expenses plot



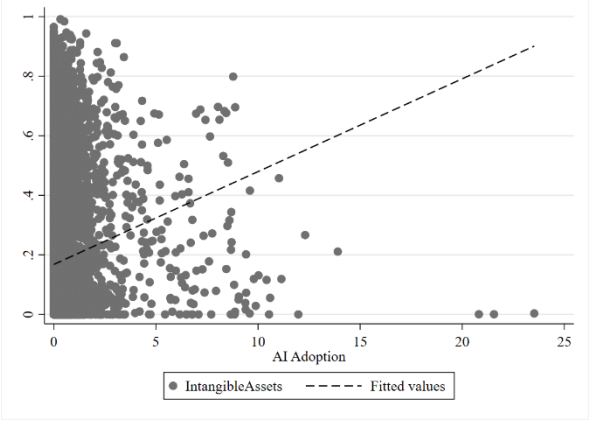
AI Adoption & Size plot



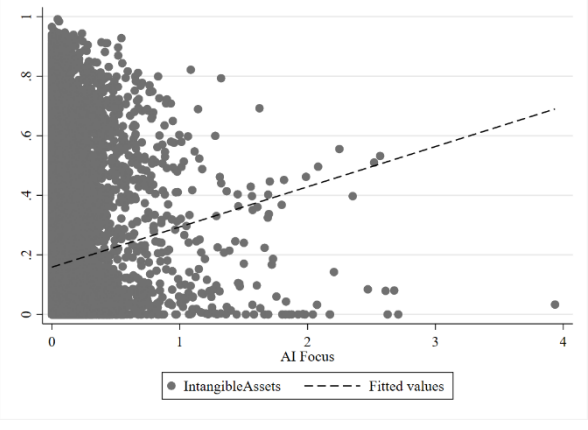
AI Focus & Size plot



AI Adoption & Intangible Assets plot



AI Focus & Intangible Assets plot



Appendix H. Fixed effects panel regression analysis of a non-tech company sample

	(1) AI Adoption	(2) AI Adoption	(3) AI Adoption	(4) AI Focus	(5) AI Focus	(6) AI Focus
CAPEX $_{i,t-1}$	0.00018**		0.00018***	0.00006***		0.00005***
	(0.000)		(0.000)	(0.000)		(0.000)
Liquidity $_{i,t-1}$		-0.00048	-0.00048		0.00024	0.00024
		(0.001)	(0.001)		(0.001)	(0.001)
ROA $_{i,t-1}$	0.00005	0.00015	0.00006	-0.00025	-0.00023	-0.00026
	(0.002)	(0.001)	(0.002)	(0.000)	(0.000)	(0.000)
R&D Expenses $_{i,t-1}$	-0.00131***	-0.00123***	-0.00134***	-0.00033***	-0.00028***	-0.00031***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Employees $_{i,t-1}$	-0.00109	-0.00098	-0.00097	-0.00028	-0.00034	-0.00034
	(0.002)	(0.002)	(0.002)	(0.001)	(0.001)	(0.001)
Age $_{i,t-1}$	-0.01094	-0.01084	-0.01085	-0.00767**	-0.00771**	-0.00771**
	(0.012)	(0.012)	(0.012)	(0.004)	(0.004)	(0.004)
Size $_{i,t-1}$	0.00122	0.00123	0.00123	0.00066**	0.00065**	0.00065**
	(0.001)	(0.001)	(0.001)	(0.000)	(0.000)	(0.000)
Leverage $_{i,t-1}$	0.00000	0.00000	0.00000	-0.00000	-0.00000	-0.00000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
SG&A Expenses $_{i,t-1}$	-0.00005	0.00000	-0.00005	-0.00005	-0.00003	-0.00005
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Intangible Assets $_{i,t-1}$	0.01660	0.01610	0.01620	0.00625	0.00642	0.00645
	(0.027)	(0.027)	(0.027)	(0.013)	(0.013)	(0.013)
Tangible Assets $_{i,t-1}$	-0.00746	-0.00777	-0.00771	0.00145	0.00156	0.00158
	(0.008)	(0.008)	(0.008)	(0.002)	(0.002)	(0.002)
HHI $_{i,t-1}$	0.03139	0.03146	0.03147	-0.02491	-0.02496	-0.02495
	(0.043)	(0.043)	(0.043)	(0.015)	(0.015)	(0.015)
Risk $_{i,t-1}$	-0.00220	-0.00207	-0.00208	-0.00075	-0.00081	-0.00081
	(0.009)	(0.009)	(0.009)	(0.001)	(0.001)	(0.001)
Constant	0.03671	0.04066	0.04067	0.05972***	0.05773***	0.05773***
	(0.022)	(0.026)	(0.026)	(0.006)	(0.009)	(0.009)
Time FE	Yes	Yes	Yes	Yes	Yes	Yes
Industry Clustering	Yes	Yes	Yes	Yes	Yes	Yes
Observations	18,148	18,148	18,148	18,148	18,148	18,148
R^2	0.015	0.015	0.015	0.031	0.031	0.031
Number of firms	3,771	3,771	3,771	3,771	3,771	3,771

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Deutschsprachiges Abstract

Die jüngsten Fortschritte im Bereich der künstlichen Intelligenz (KI) haben den Diskurs über dieses Thema sowohl in der Wissenschaft als auch in der Wirtschaft beschleunigt. Die zunehmende Anwendbarkeit und der geschäftliche Nutzen haben das Bewusstsein für die Fähigkeiten der Technologie geschärft und dazu geführt, dass immer mehr Unternehmen KI in ihren Betrieb integrieren. Die Auswirkungen von Automatisierung, Ergänzung, Verdrängung oder neuen Möglichkeiten durch KI spiegeln sich nicht nur in potenziellen Kosteneinsparungen oder Umsatzsteigerungen wider, sondern bergen auch schwerwiegende Folgen für die Zusammensetzung des Personals und mögliche Arbeitslosigkeit. Um weiterhin relevante Erkenntnisse darüber zu gewinnen, wie KI unsere Zukunft bestimmen wird, muss die Wissenschaft mit der rasanten Entwicklung von KI Schritt halten. Die begrenzte Datenverfügbarkeit und die bestehenden empirischen Methoden erfordern jedoch zusätzliche Möglichkeiten, den Einsatz von künstlicher Intelligenz in Unternehmen zu erfassen. Diese Arbeit zielt darauf ab, eine neue, auf Textanalyse basierende Messgröße zur Verfügung zu stellen, die den tatsächlichen Einsatz von künstlicher Intelligenz in einem Unternehmen sowie dessen Einflussfaktoren erfasst. Aufbauend auf früheren Fortschritten und Ergebnissen dieser Forschungsmethode berücksichtigt die Messung die Wortzahlen eines enger definierten KI-Wörterbuchs im Zusammenhang mit dem Einsatz. Durch die zusätzliche Verfeinerung wurden wesentliche Unterschiede zu bestehenden Messgrößen festgestellt, und es konnte ein verbesserter Zusammenhang zwischen Investitionseffekten und dieser einsatzzentrierten KI-Variable nachgewiesen werden. Insgesamt leistet diese Arbeit einen Beitrag zur bestehenden empirischen Forschung über KI, indem sie eine replizierbare Methode zur Messung des tatsächlichen Einsatzes von KI bereitstellt, die für die Erstellung groß angelegter Paneldaten verwendet werden kann.