The Review of Financial Studies

How to Talk When a Machine Is Listening: Corporate Disclosure in the Age of AI

Sean Cao
Robert H. Smith School of Business, University of Maryland, USA

Wei Jiang
Goizueta Business School, Emory University, USA

Baozhong Yang
J. Mack Robinson College of Business, Georgia State University, USA

Alan L. Zhang
College of Business, Florida International University, USA

Growing AI readership (proxied for by machine downloads and ownership by AI-equipped investors) motivates firms to prepare filings friendlier to machine processing and to mitigate linguistic tones that are unfavorably perceived by algorithms. Loughran and McDonald (2011) and BERT available since 2018 serve as event studies supporting attribution of the decrease in the measured negative sentiment to increased machine readership. This relationship is stronger among firms with higher benefits to (e.g., external financing needs) or lower cost (e.g., litigation risk) of sentiment management. This is the first study exploring the *feedback effect* on corporate disclosure in response to technology. (*JEL* D83, G14, G30)

Received January 14, 2021; editorial decision January 16, 2023 by Editor Tarun Ramadorai. Authors have furnished an Internet Appendix, which is available on the Oxford University Press Web site next to the link to the final published paper online.

The authors thank Tarun Ramadorai (editor) and three anonymous referees for their detailed and constructive comments. The authors have benefited from discussions with Rui Albuquerque, Elizabeth Blankespoor (discussant), Emilio Calvano (discussant), Lauren Cohen (discussant), Will Cong (discussant), Ilia Dichev, Arup Ganguly (discussant), Jillian Grennan, Bing Han, Kathleen Hanley (discussant), Rebecca Hann, Gerard Hoberg (discussant), Byoung-Hyoun Hwang (discussant), Chris Hennessy, Alan Huang (discussant), Bin Ke (discussant), Michael Kimbrough, Leonid Kogan, Augustin Landier (discussant), Tim Loughran (discussant), Song Ma, Tarun Ramadorai (editor), Ville Rantala (discussant), Max Rohrer (discussant), Gustavo Schwenkler (discussant), Kelly Shue, Suhas Sridharan, Isabel Wang (discussant), Teri Yohn, Gwen Yu, Dexin Zhou, and three anonymous referees and comments and suggestions from participants in seminars and conferences at Columbia, ECB, EDHEC, Emory, Georgia State, Harvard, University of Hong Kong, London Business School, Maryland, Michigan, Michigan State, Peking University, Q-group, the Pacific Center for Asset Management, Renmin University, Stockholm Business School, Toronto, Utah, Washington, the NBER Economics of Artificial Intelligence Conference, the NBER Big Data and Securities Markets Conference, AFA 2022, the SOAR Symposium at Singapore Management University, the Third Bergen FinTech Conference at the NHH Norwegian School of Economics, Machine Learning and Business Conference at University of Miami, RCFS Winter Conference 2021, 11th Financial Markets and Corporate Governance Conference, the China FinTech Research Conference 2021, the Adam Smith Workshop 2021, the Conference on Financial Innovation at Stevens Institute of Technology, FIRS 2021, the Cambridge Alternative Finance Sixth Annual Conference, CAPANA Research Conference 2021, CICF 2021, and NFA 2021. Supplementary data can be found on *The Review of Financial Studies* web site. Send correspondence to Wei Jiang, wei.jiang@emory.edu.

The Review of Financial Studies 36 (2023) 3603–3642

https://doi.org/10.1093/rfs/hhad021 Advance Access publication March 27, 2023

The annual report (and other regulatory filings) is more than a legal requirement for public companies; it provides an opportunity to communicate financial health, promote the culture and brand, and engage with a full spectrum of stakeholders. How those readers process this wealth of information significantly affects their perception of and, hence, participation in the business. Increasingly, companies realize that the target audience of their mandatory and voluntary disclosures no longer solely consists of human analysts and investors. A substantial amount of buying and selling of shares is triggered by recommendations made by robots and algorithms that process information with machine learning tools and natural language processing kits.[1]

Both the technological progress and the sheer volume of disclosures make the trend inevitable: Cohen, Malloy, and Nguyen (2020) document that the length of 10-Ks increases by five times from 2005 to 2017. Companies that wish to accomplish the desired outcome of communication and engagement with stakeholders need to adjust how they talk about their finances, brands, and forecasts in the age of AI. In other words, they should heed the unique logic and techniques underlying the rapidly evolving analysis of language and sentiment facilitated by large-scale machine-learning techniques, such as automated computational processes that identify positive, negative, and neutral opinions in a whole corpus of firm disclosures that is beyond the processing ability of human brains. While the literature is catching up with and guiding investors' rising aptitude to apply machine learning and computational tools to extract qualitative information from disclosures and news, there has not been an analysis exploring the *feedback effect*: how companies adjust the way they talk knowing that machines are listening. This paper fills this void.

Our analysis starts with a diagnostic test that connects how machine-friendly a company composes its disclosures (measured by *Machine readability* following Allee, DeAngelis, and Moon 2018) and the expected extent of machine readership for a company's SEC filings on EDGAR, for which we develop multiple proxies. The first variable, *Machine downloads*, is constructed by tracking IP addresses that conduct downloads in large batches. Machine request is a precursor and a necessary condition for machine reading, and the sheer volume of machine-downloaded documents makes it unlikely for them to be processed by human readers alone. Because the SEC Log files used to construct *Machine downloads* became available to the public in 2015, our analyses implicitly assume that firms were aware of the extent of machine readership before the exact numbers of machine downloads became public. To relax the assumption, we also construct a measure based on share ownership by institutional investors with AI capabilities, *AI ownership*, tracked

1 See, for example, Gara (2018). The Man Group, a leading hedge fund, has begun to manage substantial portions of its assets using AI and algorithmic trading (Satariano and Kumar 2017).

from their AI-related job postings. Finally, we proxy investor technology capacity by calculating the ownership-weighted *AI talent supply* available to institutional investors, based on the state-year-level proportion of the working-age population with IT degrees where the investors are headquartered. Because asset manager headquarters were mostly chosen before the AI era and bear no direct relation to portfolio firms, the last variable is likely to be orthogonal to omitted variables explaining *Machine readability*.

We show that, in the cross-section of filings with firm and year fixed effects, a one-standard-deviation change in expected machine downloads is associated with a 0.24-standard-deviation increase in the *Machine readability* of the filing. On the other hand, other (nonmachine) downloads do not bear a meaningful correlation with machine readability, validating *Machine downloads* as a proxy for machine readership. The alternative proxies *AI ownership* and *AI talent supply* bear similar economic and statistical significance. We further validate the economic mechanism underlying our main variables by showing that trades follow more quickly after a filing becomes public when *Machine downloads* is higher, with even stronger interactive effect with better *Machine readability*. Such a result demonstrates the real impact of machine processing on information dissemination.

After establishing a positive association between a higher AI reader base and machine-friendlier disclosure documents, we next explore how firms manage the "sentiment" and "tone" perceived by machines. It is well documented that corporate disclosures attempt to strike the right sentiment and tone with (human) readers without being explicitly dishonest or overtly noncompliant (Loughran and McDonald 2011; Kothari, Shu, and Wysocki 2009). Hence, we expect a similar strategy catering to machine readers. While researchers and practitioners have long relied on the Harvard Psychosociological Dictionary (especially the Harvard-IV-4 TabNeg file) to count and contrast "positive" and "negative" words to construct "sentiment" as perceived by (mostly human) readers, the publication of Loughran and McDonald (2011, "LM" hereafter) presents an instrumental event to test our hypothesis pertaining to machine readers. This is not only because the paper presented a specialized finance dictionary of positive/negative words and words that are informative about prospects and uncertainty but also because the word lists that came with the paper have served as a leading lexicon for algorithms to sort out sentiments in both the industry and academia.[2] The differences in both the timeline and the context of the new dictionary allow us to trace out the impact of AI readership on sentiment management by corporations.

As a first step, we establish that firms which expect high machine downloads avoid LM-negative words but only post-2011 (the publication year of the LM dictionary). Such a structural change is absent with respect to words deemed

[2] For examples of industry uses, see Marinov (2019) and Adusumilli (2020).

negative by the Harvard dictionary. As a result, the difference, *LM – Harvard sentiment*, follows the same path as *LM sentiment*. For a tighter identification, we further confirm a parallel pre-trend in *LM – Harvard sentiment* between firms with high and low (top and bottom terciles of) machine downloads up to 2010. Post-2011 saw a clear divergence where the "high" group significantly reduced, relative to the "low" group, the use of negative words from the LM dictionary as opposed to those from the Harvard dictionary. Given the quasi-randomness of the exact timing of publication, the difference-in-differences in the sentiment expression is more likely to be attributable to firms' catering to their AI readers than to an alternative hypothesis that the publication was a side show of a preexisting and continuing trend.

The documented relation raises intriguing equilibrium implications. If firms can "positify" language without cost and constraint in order to impress machine and human readers, the signals would quickly lose relevance. To remain in an equilibrium in which investors extract information from disclosures, we hypothesize that firms derive and incur heterogeneous benefits and costs from managing sentiment and tone. On the benefit side, we find that firms facing imminent external financing needs are more likely to suppress LM (2011) negative words and to disclose in more machine-readable format so as to ensure that the positive signals are well received. On the cost side, firms facing higher litigation risk are more moderated in their word-mincing.

The rapid evolution of AI technology, even during the writing and revision of this paper, provides "out-of-sample" tests to affirm that the relation we identified off the publication of LM (2011) is not a lone incidence. First, we resort to the emergence of Bidirectional Encoder Representations from Transformers (BERT) developed by Google in 2018 (Devlin et al., 2018), the state-of-the-art for machine processing of textual data. We show that BERT-measured negative sentiment drops more post-2018 for firms with higher AI readership, measured by *AI ownership* and *AI talent supply*. Second, we take the study about "how to talk when a machine is listening" literally into the speech setting. Earlier work (Mayew and Ventakachalam 2012) finds that managers' vocal expressions, as assessed by vocal analytic software, can convey incremental information valuable to analysts covering the firm. Thus, managers should recognize that their speeches need to impress bots as well as humans. Applying the software to extract two emotional features well-established in the psychology literature, valence and arousal (corresponding to positivity and excitedness of voices), from managerial speeches in conference calls, we find that managers of firms with higher expected machine readership exhibit more positivity and excitement in their vocal tones, echoing the anecdotal evidence that managers increasingly train or even seek professional help to improve their vocal performances along the quantifiable metrics (Wong 2012; Dizik 2017).

Our study builds on an expanding literature on information acquisition and dissemination via SEC-filing downloads (Bernard, Blackburne, and Thornock

2020; Chen et al. 2020; Cao et al. 2021; Crane, Crotty, and Umar 2022), opting for a new angle on the consequences of and human reactions to machine processing. A central theme from the rapidly growing literature on textual analysis is that qualitative information from and the writing quality of disclosures predict asset returns and corporate performance.[3] The computational textual analyses have been steadily advanced by more-modern machine-learning techniques (see a survey article by Cong et al. 2021), and have been extended to nontext data, such as the audio of conference calls (Mayew and Ventakachalam 2012) and video of startup pitch presentations (Hu and Ma 2021). Our study departs from the extant literature as we explore managerial disclosure strategies in response to the growing presence of AI analytical tools in both the industry and academia.

Our study thus connects to a distinct literature on the "feedback effect": while the financial markets reflect firm fundamentals, market perception also influences managers' information sets and decision-making (see a survey by Bond, Edmans, and Goldstein 2012). As long as the encoded rules are not completely opaque—and thus are transparent, observable, or reverse engineerable to at least some degree—agents affected by machine learning decisions have the incentive to manipulate inputs in order to game a more desirable outcome. Though a relation between evaluation metrics and agent behavior including disclosure is not new (Bushee 1998; Bushee and Noe 2000; Graham, Harvey, and Rajgopal 2005; Dhaliwal et al. 2011), it is fairly recent that the machine learning community formalizes the matter as one of "strategic classification" (Hardt et al. 2016; Dong et al. 2018; Milli et al. 2019) and that anecdotal evidence surfaces that companies' investor relations departments resort to algorithmic systems to test draft versions of disclosures for optimal effects.[4] While some adaptive behavior, such as making disclosures more machine-reading friendly, is innocuous or even welcome, other algorithm-induced changes, such as the expression of sentiment and tone, highlight the increasing challenge on machine learning to be "manipulation proof" in that the algorithms learn to anticipate the strategic behavior of informed agents without observing it in training samples (see theoretical analyses in Bjorkegren, Blumenstock, and Knight 2020; Hennessy and Goodhart 2021).

---

[3] Tetlock (2007), Tetlock, Saar-Tsechansky, and Macskassy (2008), and Hanley and Hoberg (2010) pioneered applying psychological dictionaries to financial texts to give content to sentiments. LM (2011) developed capital-market-specific dictionaries which have since been applied to the large-scale computation of tone and sentiment in financial texts, for example, Dow Jones newswires (Da, Engelberg, and Gao 2011), *New York Times* financial articles (Garcia 2013), 10-K and IPO prospectuses (Jegadeesh and Wu 2013), corporate press releases (Ahern and Sosyura 2014), earnings conference calls (Jiang et al. 2019), and wire news from Factiva (Huang, Tan, and Wermers 2020). Hwang and Kim (2017) directly connect the writing quality of filings to valuation in the context of closed-end funds. See also the survey article by Loughran and McDonald (2016).

[4] LM (2011) acknowledged, without providing evidence, the theoretical possibility that "[k]nowing that readers are using a document to evaluate the value of a firm, writers are likely to be circumspect and avoid negative language." A few news articles, for example, Hunter (2020), Naughton (2020), and Wigglesworth (2020) featured our research in the context of the new phenomenon.

1. Hypothesis Development

The experience of Man Group chief executive, Luke Ellis, provides a fitting motivation to our hypothesis development. Realizing that his speech could be systematically and instantaneously scraped by quant investors with natural language processing tools, Mr. Ellis decided to be coached to avoid certain words and phrases that algorithms could pick up on and thus affect Man's stock price. He was quoted as saying, "There's always been a game of cat and mouse, in CEOs trying to be clever in their choice of words. But the machines can pick up a verbal tick that a human might not even realise is a thing."[5] The episode suggests that some firms are adjusting their external communications in order for the right message to be sent to, or the right impression to be made on, a machine audience.

To formalize the hypothesis, we develop a stylized model (see Internet Appendix 1) that connects firm disclosures targeting machine readers to securities trading and pricing. In disclosures, a firm manages two additive terms to the true quality of firm fundamentals. The first is "tone." A more positive tone, other things equal, elicits a higher perception of firm fundamentals. The second is "noise" picked up by machine readers and capturing information lost because of imperfect machine readability. The higher the machine readability, the lower the signal's noise. Costly technology means an increasing marginal cost to reach higher levels of machine readability.

The trading game consists of a "machine trader" (i.e., an AI-equipped speculator who trades on machine-parsed information from the disclosure), a noise trader, and a market maker who sets the price according to the Kyle (1985) model (see also Kim and Verrecchia 1994; Foster and Viswanathan 1996). The firm's utility is a sum of three terms. The first is increasing in the current stock price, capturing the reality that managerial payoffs or firms' gains from external financing tend to be an increasing function of stock price.

The second term captures the cost of manipulating tones in disclosure, which can result in reputation and litigation risk. The last term reflects the costs to maintain a given level of machine readability. Such costs could be technology driven. Note that higher machine readability or more precise machine signals lead to more machine-driven trades, which in turn increase the impact of tones on prices. Therefore, under such an objective function, the firm desires, from an initial level, to adopt more positive tones and higher machine readability but is eventually constrained by the costs in mispricing (including reputation concerns and litigation risk.) and technology upgrades.

Empirical tests in Sections 3 and 4 demonstrate these first-order effects. In Section 4.3, we further test the empirical relation between machine-targeted disclosure management and the proxies for costs (e.g., litigation risk).

[5] For the full story, see Wigglesworth (2020).

After extending the model to multiple human and machine traders, we show that firms are motivated to maintain higher levels of tone management and machine readability when the machine traders are more numerous. Our model further shows that stock liquidity (market depth) decreases with the increasing presence of machine readers. The intuition here is that providing machine traders a more accurate signal increases the information asymmetry between the machine traders and the market maker, forcing the latter to increase price sensitivity while trading in order to avoid being taken advantage of by the machine traders. We present the empirical test on this relation in Section 3.3.

2. Data, Variable Construction, and Sample Overview

2.1 Data sources

The primary data source of this study is the Securities and Exchange Commission's (SEC) Electronic Data Gathering, Analysis, and Retrieval (EDGAR) system and the associated Log File Data Set. Since 1994, the SEC has provided the public with access to securities filings containing value-relevant and market-moving information through its EDGAR system, available through the SEC's website and WRDS SEC Analytics Suite.

While EDGAR is a content archive, its Log File tracks requests and downloads. More specifically, it comprises all records of the requests of SEC filings from EDGAR from January 2003 to June 2017 (after which the SEC stopped updating the Log Files). Each observation in the original data set contains information on the visitor's Internet Protocol (IP) address, timestamp, and the unique accession number of the filing that the visitor downloads. In preprocessing the raw Log File, we exclude requests that land on index pages because such requests do not download actual company filings. We then match the accession number with the SEC master filing index to select all the 10-K and 10-Q filings. This procedure yields a total of 438,752 filings (119,135 10-K and 319,617 10-Q). After matching to CRSP/Compustat, our final sample of raw filings consists of 359,819 filings (90,437 10-K and 269,382 10-Q), filed by 13,763 unique CIKs, between 2003 and 2016.

Needless to say, regulatory filings are one of the venues through which firms communicate to the marketplace. Alternatively, firms can host corporate events, such as conference calls, corporate presentations, and nondeal roadshows. Regulatory filings have the advantage that their audience composition is mostly exogenous to firms' own decisions, which is less true in other settings. For example, managers can invite a selected audience to corporate events, while regulatory filings are open to everyone (Cohen, Lou, and Malloy 2020). For these considerations, we focus on these two most important SEC filings for public companies.

2.2 Construction of main variables

2.2.1 Proxies for machine readership. Several constructed variables are fundamental to our analyses; we describe those in detail here. The first key variable measures the frequency of machine downloads of corporate filings, which serves as an upper bound as well as a proxy for the presence of "machine readers." Despite the advent of multiple data sources, the SEC EDGAR website remains the earliest and most authoritative source for company filings to be publicly released.[6] With the advances in computing power and data availability, some large hedge funds and asset managers have started big-data-driven programs to process and analyze unstructured data, including corporate filings and news. Recent academic studies also provide evidence that investment companies rely on machine downloads of EDGAR filings for some of their trading strategies. Crane, Crotty, and Umar (2022) find that hedge funds that employ robotic downloads perform better than those that do not. Cao et al. (2021) show that machine downloaders exhibit skills in identifying profitable copycat trades from their peers' disclosures.

To measure machine downloads, we identify an IP address downloading more than 50 unique firms' filings on any given date as a machine (i.e., robot) visitor and classify its requests on that day as machine downloads, the same criterion as used by Lee, Ma, and Wang (2015).[7] In addition, we include requests that are attributed to web crawlers in the SEC Log File Data as machine initiated. All remaining requests are labeled as "other" requests. Finally, we aggregate machine requests and other requests, respectively, for each filing within 7 days (i.e., days [0, 7]) after it becomes available on EDGAR; the majority of requests occur during this period.

Figure 1 shows an exponential growth of machine downloads since 2003. The number of machine downloads of corporate 10-K and 10-Q filings increased from 360,861 in 2003 to 165,318,719 in 2016. Other filings, notably 8-K, are also of strong interest to the market, but we do not include 8-K filings mainly because they, unlike 10-K/Qs, do not follow a standard structure, making it difficult to compare readability and writing styles in the cross-section. During the same period, machine downloads also have become the predominant force among all EDGAR requests: the number of machine downloads as a fraction of all downloads increased from 39% in 2003 to 78% in 2016. The dip in 2016 appears to be temporary. The fraction recovers to 92% during the first half of 2017, namely, the last time period (but incomplete year) for which the SEC log information is available.

[6] A multiyear episode of early leakage was largely resolved in mid-2015. See Bolandnazar et al. (2020).

[7] Loughran and McDonald (2017) proposed an alternative and more aggressive approach to classify those daily IP addresses having more than 50 requests as robot visitors. Because this approach tends to classify almost all downloads as machine driven in the most recent years, we resort to the more stringent measure by Lee, Ma, and Wang (2015). We nevertheless present the results using the Loughran and McDonald (2017) classification, which is qualitatively similar, in sensitivity checks.

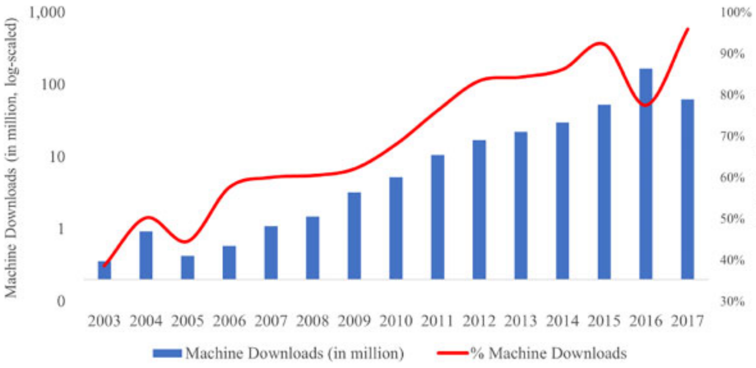

Figure 1
Trend of machine downloads
This figure plots the annual number of machine downloads (blue bars and left axis) and the annual ratio of machine downloads to total downloads (red line and right axis) across all 10-K and 10-Q filings from 2003 to the first half of 2017 (after which the SEC Log File Data Set stopped coverage). Machine downloads are defined as downloads from an IP address downloading more than 50 unique firms' filings daily. The number of machine downloads or total downloads for each filing is recorded as the respective downloads within 7 days after the filing becomes available on EDGAR.

The variable *Machine downloads* measures the propensity of machine downloads of a particular filing using ex ante information only. For a firm's (indexed by i) filing (indexed by j) at time t, *Machine downloads* is the natural logarithm of the average number of machine downloads of firm i's filings during the four quarters prior to time t (we only include the machine downloads of a historical filing within 7 days of posting on EDGAR, as explained earlier). *Other downloads* (the remainder) and *Total downloads* (the sum) are constructed analogously. Further, results using *% machine downloads*, defined as the ratio of *Machine downloads* to *Total downloads* (without taking the natural logarithm for either variable), are reported in Table IA.1 in the Internet Appendix.

It is worth noting that the SEC log files became available to the public in 2015 with retrospective information from 2003 and quarterly updates forward. Our analyses implicitly assume that firms were aware of the extent of machine readership before the exact numbers of machine downloads became public. We address the limitation of this assumption in three ways. First, market participants were able to obtain near real-time information about downloading activities via FOIA requests.[8] Second, we argue that companies have other ways learn about the interests of AI-equipped investors in real time, and the

[8] SEC log files are posted on a quarterly basis with a 6-month delay (Chen et al. 2020); we were also informed by the SEC that they later changed the delay to a quarter, and quicker discovery could be made with a Freedom of Information Act (FOIA) request. The list of people/entities who requested the SEC Log files from the SEC website regarding FOIA logs is available at https://www.sec.gov/foia/docs/foia-logs. Interestingly, a significant drop in such requests started in 2015. The coincidence with the publication of the EDGAR Log Dataset suggests that a substantial number of earlier requests could have been directed at the downloading logs.

downloads available to researchers ex post could serve as a proxy for such information. For example, companies use web analytic tools extensively to track and analyze usage data.[9] Third, we expect firms to be informed to various degrees about the AI capacity of their institutional shareholders, for which we construct alternative measures.

The first such measure is *AI ownership*, which is the percentage of shares outstanding held by investment companies with AI capabilities. We classify an investment company as such if it has AI-related job postings in the past 5 years according to data from Burning Glass, following Abis and Veldkamp (2022). *AI ownership* is the aggregate ownership measured at the firm level and in the quarter before the firm's current filing. The *AI ownership* variable is available from 2011 to 2019 since the Burning Glass data are available after 2010.

Both *Machine downloads* and *AI ownership* involve choices made by investors; those choices could be jointly determined with firms' disclosure choices. To form a sharper causal inference from investor base to disclosure choices, we construct a third proxy for machine readership, *AI talent supply*, based on local AI talent supplies where investors are headquartered, which is mostly exogenous to firms and investors. In the first step, we retrieve the number of people between 18 and 64 with college or graduate school degrees in information technology, scaled by the population at the state-year level, using data from Integrated Public Use Microdata Series (IPUMS) surveys, over the period from 2010 to 2019.[10] Second, for each firm and during the quarter prior to the current filing, we aggregate *AI talent supply* over all states based on the headquarters of the investors, weighted by their ownership. Because the headquarters locations for most investors were determined before the AI era and bear no inputs from the portfolio firms, the resultant *AI talent supply* should be exogenous to the omitted variables in firm disclosures.

2.2.2 Machine readability. The second key variable pertains to the "machine readability" of a 10-K or 10-Q filing, which measures the ease with which a filing can be "understood," that is, processed and parsed, by an automated program. Recent literature in accounting and finance has studied various concepts of (e.g., Hodge, Kennedy, and Maines 2004; Blankespoor 2019; Blankespoor, deHaan, and Marinovic 2020; Gao and Huang 2020) and proposed metrics for (Allee, DeAngelis, and Moon 2018) information processing costs related to either machine or human processing (or both). After reviewing the existing research, we adopt multiple metrics developed in Allee, DeAngelis, and Moon (2018) that we believe to best summarize the important

9 The first log file analysis program, Analog, that analyzes usage patterns on companies' web servers became available in 1995. Google Analytics first appeared in 2005. Firms also can identify the types of visitors (e.g., human or web-crawlers) on their own websites and make reasonable inferences about the composition of human versus machine among visitors of their SEC filings (Angwin 2011; Burnham 2014).

10 A recent paper by Jiang et al. (2021) describes the data in detail.

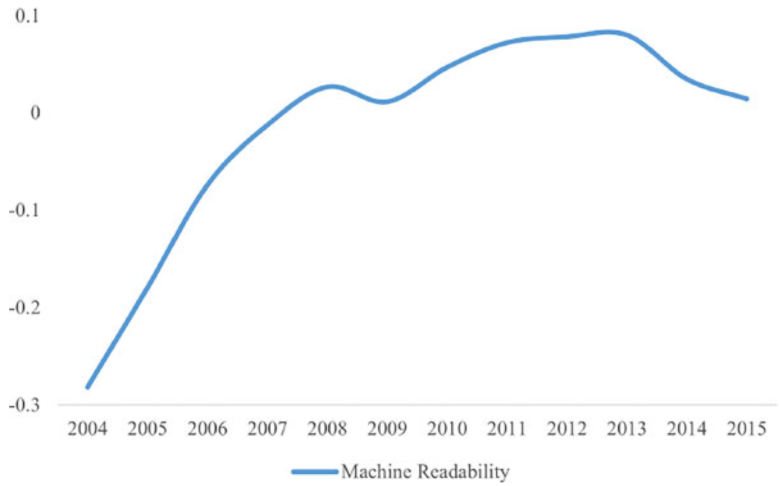

Figure 2
Trend of machine readability
This figure plots the annual *Machine readability* across all 10-K and 10-Q filings from 2004 to 2015. *Machine readability* is the average of five standardized filing attributes: *Table extraction*, *Number extraction*, *Table format*, *Self-containedness*, and *Standard characters*. All attributes are defined in the appendix.

attributes distinctly related to machine readability:[11] (a) *Table extraction*, the ease of separating tables from the text; (b) *Number extraction*, the ease of extracting numbers from the text; (c) *Table format*, the ease of identifying the information contained in the table (e.g., whether a table has headings, column headings, row separators, and cell separators); (d) *Self-containedness*, whether a filing includes all needed information (i.e., without relying on external exhibits); and (e) *Standard characters*, the proportion of characters that are standard ASCII (American Standard Code for Information Interchange) characters. In our main specification, each attribute is standardized to a Z-score before being averaged to form a single-index *Machine readability* measure. We present sensitivity checks (and demonstrate robustness) using the first principal component (see Table IA.1 in the Internet Appendix) of the five attributes as well as the individual underlying attributes. Figure IA.1 in the Internet Appendix provides a visualization of the *Machine readability* variable by showing two sample filings with a low and high score, with explanations of how features of the filings are related to the machine readability scoring.

Figure 2 shows the trend of *Machine readability* from 2004 to 2015. *Machine readability* saw steep ascendance till 2008, followed by modest growth before leveling off around 2011. The increasing trend *per se* is prima

[11] We thank Allee, DeAngelis, and Moon (2018) for sharing these component variables from their paper. We adopt a subset of the measures developed therein as we solely focus on components that matter mostly for machine readability (e.g., whether numbers and tables are parsable) and do not include components that may affect both machine parsing and human understanding (e.g., whether a document is separated into different sections).

facie evidence that companies are not following a fixed template for financial filings, but instead have been adapting the format of their filings to a changing environment.[12]

2.2.3 (Negative) sentiment and tones. The third class of key variables aims at measuring "sentiments," which broadly refer to the use of natural language processing, text analysis, and computational linguistics to systematically identify, extract, and quantify subjective information. Because a primary interest of this study is to contrast the sentiment as perceived by human and machine readers, we resort to two established lexica that guide sentiment classification by the two types of readers. The first lexicon is the Harvard General Inquirer IV-4 psychological dictionary. This comprehensive dictionary assigns 77 psychological intonations or categories to English words. For each corporate filing, we count the number of words that fall into the "Negative" category and normalize it by the total number of words in the textual part of a 10-K/Q filing with all tags, tables, and exhibits removed. This procedure follows the common practice in the literature, for example, LM (2011) and Cohen, Malloy, and Nguyen (2020). The resultant measure, expressed in percentage points, is termed *Harvard sentiment*. The average filing in our sample contains four Harvard General Inquirer negative words per 100 words. The second lexicon is developed by LM (2011), who create dictionaries of positive and negative words that are specific to the context of financial documents. We count the number of LM-negative words and scale it by the length of the document. The resultant measure, expressed in percentage points, is the *LM sentiment*. We consider only the negative sentiment related to both dictionaries because the previous literature, including Tetlock (2007), LM (2011), and Cohen, Malloy, and Nguyen (2020), finds that positive sentiment is not as informative. An average (median) filing uses 1.63 (1.54) LM-negative words in every 100 words. The interquartile range is from 1.19 to 1.98 words per 100 words. Finally, we form the difference, *LM – Harvard sentiment*, to capture the contrast. The LM (2011) list of measures for sentiment goes broader to include litigiousness, uncertainty, weak modal and strong modal words, all in financial contexts. More specifically, *Litigious* is the number of litigation-related words (such as "claimant" and "tort") divided by the length of the document, expressed in percentage points. The other measures are constructed analogously. Uncertainty words capture a general notion of imprecision (such as "approximate" and "contingency"), and weak modal and strong modal words convey levels of confidence (such as "always" and "must" as strong, and "possibly" and "could" as weak). In an average filing, every 100

12 On April 13, 2009, the SEC released a mandate on "Interactive Data to Improve Financial Reporting" (see https://www.sec.gov/info/smallbus/secg/interactivedata-secg.htm) as a regulatory effort in adapting disclosures to machine readers. This mandate applies to the financial reports of all companies and was implemented over the period from 2009 to 2011, and could explain some of the variations around that period.

words contain 0.97 (1.43, 0.52, and 0.30) litigious (uncertainty, weak modal, and strong modal) words. We confirm LM (2011) findings that the frequency of words in these categories in firm filings is associated with stock market reactions and real outcomes, hence constitute a motive for firms to manage the wording that could lead to tone inference.

The emergence of Bidirectional Encoder Representations from Transformers (BERT), a transformer-based machine-learning technique for natural language processing developed by Google in 2018, offers us an additional—and recent—setting to test the same economic mechanism. The BERT model provides an integral treatment of sentences that takes into account the meaning, order, and interactions of words. More specifically, we use FinBERT (Huang, Wang, and Yang 2022), a version of BERT trained with financial disclosure data (including 10-K, conference call transcripts, and analyst reports) and thus more tailored to our setting, to classify the sentiment of individual sentences in 10-Ks to be positive or negative. We construct the *BERT sentiment* measure as the ratio of the number of BERT-negative sentences to the total number of sentences (or the total number of words) in 10-K sections. To economize on computation time, we focus on the key 10-K section most relevant to our context: Item 7 ("Management Discussion & Analysis (MD&A)"). We also conduct a sensitivity check which also includes Item 1 ("Business (a description of the company's operation)").

2.2.4 Vocal emotions. Though the focus of this study rests on 10-K and 10-Q filings, we extend to conference calls between firms and the public. The last set of key variables thus concerns audio quality. We build a web crawler using *Selenium-Python* to obtain the audios of conference calls from 2010 to 2016 from EarningsCast.[13] After matching with CRSP/Compustat, our sample consists of 43,462 audio files from 3,290 unique firms (*gvkey*).

Anecdotal evidence suggests that executives have become aware that their speech patterns and emotions, evaluated by humans or software, affect their assessment by investors and analysts. A pioneer academic study by Mayew and Ventakachalam (2012) finds that when analysts make stock recommendations, they incorporate managers' emotions during conference calls. One of the most prominent models of emotion, the circumplex model, originally developed by Russell (1980), suggests that emotions are distributed in a two-dimensional space defined by valence and arousal. Following Hu and Ma (2021), we rely on a pretrained Python machine learning package *pyAudioAnalysis*[14] (Giannakopoulos 2015) to code the vocal emotion of each conference call.

[13] EarningsCast is a commercial aggregator for company earnings calls, calendar feeds, and podcast feeds. Its website is https://earningscast.com. *Selenium-Python* is an open-source software package that allows us to program a specific mouse-clicking sequential pattern for a particular website so that we can automate web browsing and internet data retrieval from the website; see https://selenium-python.readthedocs.io.

[14] The open-source *pyAudioAnalysis* is available at https://github.com/tyiannak/pyAudioAnalysis.

Emotion valence describes the extent to which an emotion is positive or negative, with a larger value indicating greater positivity. *Emotion arousal* refers to the intensity or strength of the associated emotion state; a greater (lower) value suggests that the speaker is more excited (calmer). Both measures are bounded between –1 and 1.

2.2.5 Firm characteristics. As usual, the firm characteristics variables (serving as control variables) are retrieved or based on information from standard databases accessed via WRDS, such as CRSP/Compustat and Thomson Reuters Ownership Database. In this category of variables, *Size* is the market capitalization in the natural logarithm. *Tobin's q* is the natural logarithm of the ratio of the sum of market value of equity and book value of debt to the sum of book value of equity and book value of debt. *ROA* is the ratio of EBITDA to assets. *Leverage* is the ratio of total debt to assets at book value. *Growth* is the average sales growth of the past 3 years. *Industry adjusted return* is the monthly average SIC three-digit industry-adjusted stock returns over the past year. *Institutional ownership* is the ratio of the total shares of institutional ownership to shares outstanding. *Analyst coverage* is the natural logarithm of one plus the number of IBES analysts covering the stock. *Idiosyncratic volatility* is the annualized idiosyncratic volatility (using daily data) from the Fama-French three-factor model. *Turnove*r is the monthly average of the ratio of trading volume to shares outstanding. *Segment* is the number of business segments and measures the complexity of business operations, following Cohen and Lou (2012). All control variables are constructed annually using information available at the previous year-end. All potentially unbounded variables are winsorized at the 1% extremes.

The appendix defines all variables, and Table 1 reports summary statistics. Because some variables require historical information, the sample for our regression analyses starts in 2004 and consists of a total of 324,607 filings (81,075 10-K and 243,532 10-Q).

3. AI Readership and Machine Readability of Disclosures

3.1 Validating *Machine downloads* as proxy for AI readership

Our analyses critically depend on *Machine downloads* being an effective proxy for the presence of AI readership. We thus conduct two tests that support the validity of this key empirical proxy. First, tracing the downloads to the identities of the downloaders would help ascertain that the large-batch downloads are indeed a likely precursor for machine processing. To this end, we use the ARIN Whois database to manually match the IP addresses that have the highest volumes of machine downloads to the universe of investors who ever appear as a 13F filer in the Thomson Reuters 13F database during the sample period. Table 2 reports the identities of the top-20 machine downloaders and the types of institutions they are. Half of the top-10 on the list are prominent quantitative

Table 1
Summary statistics

Variables	Mean	Median	SD	P25	P75	N
			Filing level			
Machine downloads	4.729	4.508	1.763	3.296	6.377	324,607
Other downloads	3.448	3.474	1.378	2.615	4.363	324,607
Total downloads	5.090	4.915	1.609	3.829	6.535	324,607
% machine downloads	0.742	0.775	0.179	0.623	0.892	324,231
Machine readability	−0.020	0.125	0.584	−0.224	0.359	199,421
AI ownership	0.041	0.014	0.048	0.000	0.077	79,567
AI talent supply	0.475	0.470	0.403	0.017	0.834	95,643
LM – Harvard sentiment	−2.413	−2.385	0.544	−2.747	−2.047	324,589
LM sentiment	1.625	1.543	0.599	1.185	1.982	324,589
Harvard sentiment	4.038	4.021	0.697	3.561	4.492	324,589
Litigious	0.965	0.82	0.537	0.593	1.177	324,589
Uncertainty	1.425	1.377	0.398	1.146	1.652	324,589
Weak modal	0.521	0.427	0.304	0.314	0.634	324,589
Strong modal	0.295	0.271	0.133	0.202	0.359	324,589
			Conference call level			
Emotion valence	0.331	0.375	0.261	0.227	0.498	43,462
Emotion arousal	0.647	0.650	0.138	0.557	0.740	43,462
			Firm-year-level control variables			
Size	6.238	6.220	2.022	4.804	7.617	43,764
Tobin's q	0.672	0.557	0.718	0.178	1.064	43,764
ROA	0.049	0.101	0.271	0.028	0.163	43,764
Leverage	0.221	0.160	0.244	0.008	0.337	43,764
Growth	0.152	0.0736	0.42	−0.005	0.191	43,764
Industry adjusted return	0.000	−0.001	0.039	−0.021	0.019	43,764
Institutional ownership	0.482	0.528	0.359	0.080	0.816	43,764
Analyst coverage	1.498	1.609	1.193	0.000	2.485	43,764
Idiosyncratic volatility	0.463	0.386	0.289	0.263	0.576	43,764
Turnover	2.150	1.619	1.960	0.826	2.791	43,764
Segment	5.323	5.000	3.564	2.000	7.000	43,764

This table provides summary statistics. Filing-level variables are based on the sample of SEC EDGAR 10-K and 10-Q filings from 2004 to 2016. Conference-call-level variables are based on the sample of the audios of corporate conference calls from 2010 to 2016. Firm-year-level control variables are calculated annually using information available at the previous year-end. The appendix defines the variables.

hedge funds: Renaissance Technologies, Two Sigma, Point 72, Citadel, and D.E. Shaw. This revelation confirms the anecdotal evidence that quant funds are major players in integrating big data and unstructured data analyses in making investment decisions. The remaining institutions are mostly brokers and investment banks with significant asset management businesses.

Second, we connect *Machine downloads* to its primary suspect, hedge funds that adopt AI strategies. Following Guo and Shi (2020), we classify a hedge fund to be AI-prone if at least one employee has been involved in AI projects based on their LinkedIn profiles.[15] We then define *AI hedge fund* to be the percentage of shares outstanding held by such hedge funds at the firm-quarter level, based on the 13F filings via the Thomson Reuters Ownership Database. We find that *AI hedge fund* significantly (at the 5% level) predicts *Machine downloads* inclusive of all the control variables introduced in Section 2.2.5 (see Table IA.3 in the Internet Appendix).

15 We thank Norman Xuxi Guo and Zhen Shi for sharing the data of hedge funds with AI-experienced employees. AI projects are identified based on both job titles and descriptions of experience and responsibilities.

Table 2
Top machine downloaders

Rank	Name of institution	#MD	Type of institution
1	Renaissance Technologies	536,753	Quantitative hedge fund
2	Two Sigma Investments	515,255	Quantitative hedge fund
3	Barclays Capital	377,280	Financial conglomerate with asset management
4	JPMorgan Chase	154,475	Financial conglomerate with asset management
5	Point72 Asset Management	104,337	Quantitative hedge fund
6	Wells Fargo	94,261	Financial conglomerate with asset management
7	Morgan Stanley	91,522	Investment bank with asset management
8	Citadel LLC	82,375	Quantitative hedge fund
9	RBC Capital Markets	79,469	Financial conglomerate with asset management
10	D. E. Shaw Co.	67,838	Quantitative hedge fund
11	UBS AG	64,029	Financial conglomerate with asset management
12	Deutsche Bank AG	55,825	Investment bank with asset management
13	Union Bank of California	50,938	Full-service bank with private wealth management
14	Squarepoint Ops	48,678	Quantitative hedge fund
15	Jefferies Group	47,926	Investment bank with asset management
16	Stifel, Nicolaus Company	24,759	Investment bank with asset management
17	Piper Jaffray	18,604	Investment bank with asset management
18	Lazard	18,290	Investment bank with asset management
19	Oppenheimer Co.	15,203	Investment bank with asset management
20	Northern Trust Corporation	11,916	Financial conglomerate with asset management

This table lists the 20 13F-filing institutional investors with the highest number of machine downloads (#MD) during our sample period of 2004 to 2016.

3.2 Relation between *Machine downloads* and *Machine readability*

As more and more investors use AI tools, such as natural language processing and sentiment analyses, we hypothesize that companies adjust the way they talk in order to communicate effectively to readers what they put in the reports. Specifically, our model in Section 1 shows that, other things equal, a larger presence of machine readers in the market will lead firms to increase the readability of their disclosure with respect to machines. A first test is thus to relate *Machine readability* to *Machine downloads* in the cross-section and over time. The first four columns of Table 3 report the results from the following regression at the filing level, indexed by firm(i)-filing(j)-date(t), with both year and firm (or industry) fixed effects, in addition to the slew of control variables ($Control$, as introduced in Section 2.2.5):

$$MachineReadability_{i,j,t} = \beta MachineReadership_{i,j,t} + \delta OtherDownloads_{i,j,t} + \gamma Control_{i,year} + \alpha_i(\alpha_{SIC3}) + \alpha_{year} + \epsilon_{i,j,t}. \quad (1)$$

In Table 3, panel A, the variable *Machine downloads* serves as the proxy for machine readership.[16] It shows that the expected machine downloads for the company's filing, whether measured as the volume or percentage of machine

[16] Table IA.3 in the Internet Appendix reports regressions for the determinants of *Machine downloads*. Results show that machine downloads tend to be higher for large firms with more firm-specific developments (e.g., high trading turnover, or high idiosyncratic volatility). Because our research question concerns the consequence of machine readership, the magnitude of machine downloads (instead of the percentage) is the more pertinent metric and hence our default measure.

downloads, significantly (at the 1% level) and positively predicts machine-reading friendliness across all specifications. With the standard deviations of *Machine downloads* and *Machine readability* being 1.763 and 0.584, respectively (see Table 1), the first four columns show that a one-standard-deviation increase in *Machine downloads* is associated with a 0.18- to 0.24-standard-deviation increase in *Machine readability*. If we calibrate the effect to incorporate firm fixed effects, a one-standard-deviation increase in within-firm variation of *Machine downloads* is associated with a 0.20-standard-deviation increase in within-firm variation in *Machine readability*. The effects are almost invariant with or without the control variables, indicating that other firm characteristics have little confounding effect. To ensure that the intertemporal persistence of *Machine downloads* is not affecting statistical inference, we adopt the Driscoll and Kraay (1998) standard errors to account for serial dependence. In addition, we present a sensitivity check for standard errors double clustered by industry and time. Results, reported in Table IA.4 in the Internet Appendix, are robust. Presumably, nonmachine downloads could serve as a natural placebo test. Indeed, all four coefficients for *Other downloads* (columns 1 to 4) turn out to be indistinguishable from zero, economically and statistically.

In reality, firms are unlikely to manage the level of machine readability of their disclosures back and forth from year to year. Instead, increasing machine readability is usually an outcome of a technology upgrade which firms conduct every once in a while when they observe the rise of machine readership of their published filings. To capture such a mechanism, we present a new machine readability *upgrade* analysis based on intertemporal differencing (instead of firm fixed effects). More specifically, we define an "upgrade" event at the filing (i,j,t) level if $Machine\,Readability_{i,j,t}$ incurs a significant (i.e., one standard deviation of the full sample) increase over the previous year's $Machine\,Readability_{i,t-1}$. We then regress the indicator variable *MR upgrade*$_{i,j,t}$ on lagged changes in *Machine downloads* from $t-2$ to $t-1$, $\Delta Machine\,Downloads_{i,t-1}$.

The last two columns of panel A in Table 3 report the results. We show that past growth in machine downloads is a significant predictor of machine readability upgrades. Such a dynamic upgrading analysis affords a byproduct of tighter causal identification: While a regression with firm fixed effects (columns 2 and 4) helps with identification when endogeneity due to firm-level heterogeneity is time invariant, the intertemporal differencing (i.e., *MR upgrade*$_{i,j,t}$ and $\Delta Machine\,Downloads_{i,t-1}$) relaxes the assumption such that the unobserved firm-level heterogeneity is only required to be stable during the differencing window, or 2 years, which is plausible. Moreover, this specification also mitigates the concerns for the intertemporal persistence of the key independent variable *Machine downloads* in levels, because the upgrades do not exhibit persistence in our sample.

Table 3
Machine downloads and machine readability

A. Machine readability

	(1)	(2)	(3)	(4)	(5)	(6)
Dependent variable	*Machine readability*				*MR upgrade*	
Machine downloads	0.076***	0.075***	0.060***	0.078***		
	(13.89)	(17.45)	(10.33)	(15.93)		
ΔMachine downloads					0.005***	0.006***
					(2.90)	(3.40)
Other downloads	0.005	0.002	−0.007	−0.006	0.000	−0.001
	(1.15)	(0.47)	(−1.44)	(−1.33)	(0.20)	(−0.44)
Size			0.004	0.021***	−0.002	−0.001
			(1.05)	(2.66)	(−1.27)	(−0.27)
Tobin's q			−0.006	−0.008	−0.002	−0.000
			(−0.92)	(−1.00)	(−0.94)	(−0.03)
ROA			0.056***	0.009	0.006	0.026**
			(3.15)	(0.49)	(1.15)	(2.52)
Leverage			−0.087***	−0.037*	0.017***	0.016*
			(−4.62)	(−1.67)	(3.02)	(1.66)
Growth			−0.017**	0.010	0.006**	−0.001
			(−2.34)	(1.27)	(2.29)	(−0.26)
Industry adjusted return			0.033	0.013	0.024	0.004
			(0.52)	(0.20)	(0.82)	(0.13)
Institutional ownership			0.050***	−0.038	−0.001	0.008
			(2.69)	(−1.50)	(−0.21)	(0.73)
Analyst coverage			0.005	0.000	−0.003*	−0.003
			(0.79)	(0.02)	(−1.74)	(−0.76)
Idiosyncratic volatility			−0.072***	0.015	0.009	0.004
			(−3.81)	(0.86)	(1.36)	(0.40)
Turnover			−0.002	−0.007***	−0.000	−0.001
			(−1.17)	(−3.16)	(−0.68)	(−0.69)
Segment			0.004***	−0.003	0.001*	0.001
			(3.05)	(−1.42)	(1.95)	(1.09)
Observations	198,358	199,241	150,425	150,346	135,146	135,068
R-squared	.082	.363	.084	.357	.025	.144
Firm FE	No	Yes	No	Yes	No	Yes
Industry FE	Yes	No	Yes	No	Yes	No
Year FE	Yes	Yes	Yes	Yes	Yes	Yes

B. Components of Machine readability

	(1)	(2)	(3)	(4)	(5)
	Machine readability				
Dependent variable	*Table extraction*	*Number extraction*	*Table format*	*Self-containedness*	*Standard characters*
Machine downloads	0.051***	0.028***	0.026***	0.161***	0.125***
	(6.02)	(3.47)	(2.88)	(21.80)	(14.68)
Other downloads	0.018**	−0.011	0.022**	−0.036***	−0.040***
	(2.37)	(−1.49)	(2.51)	(−6.69)	(−6.08)
Observations	149,484	150,346	149,484	150,245	140,061
R-squared	.471	.389	.439	.306	.344
Control variables	Yes	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes

Table 3
(Continued)

C. Alternative machine-readership measures

	(1)	(2)	(3)	(4)
Dependent variable		*Machine readability*		
AI ownership	0.553***	0.400***		
	(8.53)	(9.56)		
AI talent supply			0.240***	0.349***
			(14.85)	(21.01)
Observations	58,720	58,674	70,969	70,912
R-squared	.091	.375	.091	.366
Control variables	Yes	Yes	Yes	Yes
Firm FE	No	Yes	No	Yes
Industry FE	Yes	No	Yes	No
Year FE	Yes	Yes	Yes	Yes

This table examines the relation between the machine readability of a firm's filing and the machine downloads of the firm's past filings. *Machine downloads* measures the expected machine readership of a filing. Panel A reports a single-index *Machine readability* score that measures the ease at which a filing can be processed by an automated program. *MR upgrade* indicates an upgrade event, that is, when a filing incurs a one-standard-deviation increase over the previous-year average *Machine readability*. Δ*Machine downloads* measures the change of machine readership. Panel B reports the underlying components of *Machine readability*: *Table extraction* (the ease of separating tables from the text), *Number extraction* (the ease of extracting numbers from the text), *Table format* (the ease of identifying the information contained in the table), *Self-containedness* (whether a filing includes all needed information), and *Standard characters* (the proportion of characters that are standard ASCII characters). Panel C reports alternative machine-readership measures. *AI ownership* is the aggregate ownership of a firm by AI-equipped investment company shareholders. *AI talent supply* measures the local talent supplies to a firm's institutional shareholders, weighted by their ownership; the local talent supply is the available workforce with IT degrees in the state where an investor is headquartered. Both *AI ownership* and *AI talent supply* are available for the sample period from 2011 to 2019. Control variables include *Size*, *Tobin's q*, *ROA*, *Leverage*, *Growth*, *Industry adjusted return*, *Institutional ownership*, *Analyst coverage*, *Idiosyncratic volatility*, *Turnover*, and *Segment*. The appendix defines all variables. In all panels, the t-statistics, in parentheses, are based on standard errors clustered by firm. *p <.1; **p <.05; ***p <.01.

Panel B of Table 3 breaks down *Machine readability* into its five components: *Table extraction*, *Number extraction*, *Table format*, *Self-Containedness*, and *Standard characters*. Results show that high expected machine downloads increase all five submetrics of machine readability significantly (at the 1% level). Again, the coefficients for *Other downloads* do not have consistent signs across the five attributes.

Panel C of Table 3 examines the relation between machine readability with the two alternative measures for machine readership. *AI ownership* is the percentage of shares outstanding of a given firm during the quarter before the filing that are owned by "AI-equipped" institutional investors based on their job postings. *AI talent supply* is the state-level information technology talent (as percentage of population) aggregated at the firm level based on the headquarter locations of its investors. Both variables are described in Section 2.2.1). Results show that a one-standard-deviation increase in *AI ownership* (*AI talent supply*) is associated with a 0.04- (0.12-)standard-deviation increase in *Machine readability* (all significant at the 5% level). The consistent relation using all machine-readership proxies provides confidence in the inferences. Moreover, the results associated with *AI talent supply* are particularly helpful for causal inferences, as state-level AI talent supply where firms' investors are headquartered, mostly decided before the AI era, is likely to be exogenous to any omitted variables in the regression.

Finally, a strand of accounting literature documents that sometimes firms may want to downplay bad news with obfuscated language (Asay, Libby, and Rennekamp 2018). To demonstrate a consistent incentive, we verify a correlation between linguistic obfuscation and complexity (Loughran and McDonald 2014; Kim, Wang, and Zhang 2019) and low *Machine readability*, which could be interpreted as technical/formatting obfuscation; moreover, firms exhibiting greater linguistic complexity are less likely to have an upgrade in machine readability (see Table IA.5 in the Internet Appendix.)

3.3 The effect of *Machine downloads* and *Machine readability* on trading and information dissemination

The primary advantage machines enjoy is their capacity and information processing speed. When disclosures are read more by machines, and when filings are made more machine readable, we hypothesize that trades motivated by the information in the disclosures should materialize faster, and the speed of information dissemination should be faster. The testing of such a hypothesis is operationalized into a duration analysis connecting "time to trade" and "time to quote change" to the key independent variables. Using high-frequency data in NYSE Trade and Quote (TAQ) databases, we first conduct the following regression at the filing level, indexed by firm(i)-filing(j)-date(t), with year and firm (or industry) fixed effects:

$$\begin{aligned} TimetoTrade_{i,j,t} = {} & \beta_1 MachineDownloads_{i,j,t} \times MachineReadability_{i,j,t} \\ & + \beta_2 MachineDownloads_{i,j,t} + \beta_3 MachineReadability_{i,j,t} \\ & + \delta OtherDownloads_{i,j,t} + \gamma Control_{i,year} + \alpha_i(\alpha_{SIC3}) + \alpha_{year} + \epsilon_{i,j,t}. \end{aligned} \tag{2}$$

The dependent variable has two versions: *Time to the first trade* and *Time to the first directional trade*, the construction of which follows Bolandnazar et al. (2020). *Time to the first trade* is the length of time, in seconds, between the time stamps of the EDGAR posting and the first subsequent trade of the issuer's stock. *Time to the first directional trade* adds a requirement that the trade needs to be profitable (before any transaction cost) based on the price at the end of the 15th minute post-filing. That is, the first directional trade is the first buy (sell) trade at a price below (above) the "terminal value," where buy- and sell-initiated trades are classified by the Lee and Ready (1991) algorithm. As in Bolandnazar et al. (2020), we focus on the 15-minute window in order to isolate the effect of the filing; hence, the duration variables are censored at the end of the time window.

The results, reported in Table 4, panel A, support the prediction that high *Machine downloads* are associated with faster trades after a filing becomes publicly available. A one-standard-deviation increase in *Machine downloads*

Table 4
Effects of machine downloads

A. Time to the first trade

	(1)	(2)	(3)	(4)
Dependent variable	*Time to first trade*		*Time to first directional trade*	
Machine downloads	−4.857*	−3.398	−7.540***	−7.258**
	(−1.68)	(−1.14)	(−2.71)	(−2.55)
Machine downloads × *Machine readability*		−3.887***		−2.127*
		(−2.84)		(−1.67)
Machine readability		−5.980		−8.709
		(−0.92)		(−1.46)
Other downloads	3.499	1.304	3.885*	2.336
	(1.42)	(0.51)	(1.72)	(1.00)
Observations	161,664	144,193	161,664	144,193
R-squared	.269	.272	.285	.286
Control variables	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes

B. Effects of machine readership: Bid-ask spread

	(1)	(2)	(3)	(4)
Dependent variable	*Bid-ask spread*		*Bid-ask spread*	
Groups	Entire sample		Low turnover	High turnover
Machine downloads × *After*	0.055***	0.081***	0.080***	0.089***
	(8.46)	(10.91)	(7.18)	(8.97)
Machine readability × *After*		0.023	0.010	0.030
		(1.15)	(0.33)	(1.10)
Observations	2,673,992	2,416,151	1,203,653	1,212,498
R-squared	.720	.732	.738	.715
Firm FE	Subsumed	Subsumed	Subsumed	Subsumed
Filing FE	Yes	Yes	Yes	Yes
Minute FE	Yes	Yes	Yes	Yes

This table examines the effects of *Machine downloads* on trading and information dissemination. *Machine downloads* measures the expected machine readership of a filing. *Machine readability* measures the ease at which a filing can be processed by an automated program. Panel A reports the relation between the time to the first trade after a firm's filing is publicly released and the expected machine readership of the filing, and how the machine readability of the filings affects such a relation. *Time to the first trade* is the length of time, in seconds, between the EDGAR publication time stamp and the first trade of the issuer's stock since the publication. *Time to the first directional trade* is defined analogously, where the first directional trade is the first buy (sell) trade at a price below (above) the terminal value at the end of a 15-minute window. Panel B reports the relation between *Machine downloads* and *Bid-ask spread*, where the sample consists of filing-minute-level observations from 15 minutes before to 15 minutes after the posting of the filings. *Bid-ask spread* is the difference between the ask price and the bid price scaled by the midpoint, calculated at the minute level following the NBBO rule. *After* is an indicator variable equal to one if the time is after a filing is publicly released and zero otherwise. The sorting variable *Turnover*, the ratio of trading volume to shares outstanding, separates firms into two subsamples by the median. Control variables include *Size*, *Tobin's q*, *ROA*, *Leverage*, *Growth*, *Industry adjusted return*, *Institutional ownership*, *Analyst coverage*, *Idiosyncratic volatility*, *Turnover*, and *Segment*. The appendix defines all variables. The t-statistics, in parentheses, are based on standard errors clustered by firm in panel A and by filing in panel B. $^{*}p<.1$; $^{**}p<.05$; $^{***}p<.01$.

saves 8.6 to 14.7 seconds for the first trade and 13.3 to 21.8 seconds for the first directional trade. All coefficients associated with directional trades (in the last two columns) are significant at the 1% level, while the coefficients lose significance with *Time to the first trade* when firm fixed effects are included. Moreover, the relation between *Machine downloads* and the Time to Trade variables is indeed significantly stronger when *Machine readability* is higher.

In addition to trades, we examine how *Machine downloads* affects the quote changes around filings, a more direct test for information dissemination. We define a directional quote change as an increase (decrease) in the ask (bid) price if the price at the end of the 15th minute post-filing is higher (lower) than the latest price prior to filing. We then replace the dependent variable in Equation (2) to be *Time to the first directional quote change*, classified as the first increase in ask price upon favorable news or the first decrease in the bid price upon unfavorable news, where the direction is determined by stock price 15 minutes post-filing. We find similar but statistically weaker results.[17]

Our model discussed in Section 1 demonstrates that stock liquidity decreases with the increasing presence of machine readers. The intuition is that providing machine traders with a more accurate signal increases the information asymmetry between the machine traders and the market maker (analogous to Kim and Verrecchia 1994, 1997), forcing the market maker to increase their price sensitivity to trades to avoid trading losses against the machine traders. Following the common practice in the market microstructure literature, we test the impact of machine readers on information asymmetry and hence trading liquidity by exploring the bid-ask spread before and after a filing. Specifically, we conduct the following regression at the firm(i)-filing(j)-minute(m) level with both filing and minute fixed effects:

$$Bid\text{-}Ask\,Spread_{i,j,m} = \beta Machine\,Downloads_{i,j} \times After_{i,j,m} + \gamma Machine\,Readability_{i,j} \times After_{i,j,m} + \alpha_{i,j} + \alpha_m + \epsilon_{i,j,m}. \qquad (3)$$

The samples cover from 15 minutes before each filing to 15 minutes afterward. The dependent variable, *Bid-ask spread*, is constructed using the latest pair of lowest ask price and highest bid price within each minute following the National Best Bid and Offer (NBBO) rule, and is scaled by the midpoint of the bid price and ask price. *After* is a dummy variable equal to one if minute m occurs after the filing is posted. When both filing ($\alpha_{i,j}$) and minute-level time (α_m) fixed effects are included, the single-variable terms (including *Machine downloads* and *Machine readability*) and the control variables are all subsumed because firm characteristics do not change during the 30-minute window.

The most important coefficient from the results, reported in Table 4, panel B, is the coefficient associated with *Machine downloads* $\times$ *After*. Panel B shows that *Bid-ask spread* widens more for filings with higher expected *Machine downloads* after filings become publicly available. The coefficient is significant at the 1% level across all specifications. From the result in column 2, the incremental increase in the spread associated with a one-standard-deviation increase of *Machine downloads* amounts to 14 basis points, or about 19%

[17] Table IA.6 in the Internet Appendix reports detailed results. It is worth noting that the relation we study herein is different from the setting in Allee, DeAngelis, and Moon (2018), who combine the information processing costs of both humans and machines. We make more strict empirical choices to focus on machine readability. Such a difference could explain why Allee, DeAngelis, and Moon (2018) show limited evidence on the speed of news dissemination.

(3.3%) of the median (average) spread in our sample. However, files that score higher on *Machine readability* do not experience significant spread expansion post-filing, despite positive coefficients for *Machine readability* × *After*.

Because firm characteristics variables are subsumed by high-dimensional fixed effects, we explore the cross-sectional effects by sorting firms into two subsamples by the median value of *Turnover* (defined in the appendix), an important variable characterizing a firm's trading environment. The last two columns of Table 4, panel B, show results that the trading environment has little impact on the relation between *Machine downloads* and *Bid-ask spread*. The two coefficients are not materially different from each other, economically or statistically.

The overall evidence is consistent with the prediction that machine-equipped (hence quicker-informed) investors are able to update their judgments about a firm's fundamentals more efficiently than others, which worsens information asymmetry.

4. Managing Sentiment and Tone with Machine Readers

4.1 Textual sentiment

While truthfulness in disclosure reports is expected and required, managers usually want to portray their business activities and prospects in a positive light to attract or gain from stakeholders (creditors, employees, suppliers, and customers). Earlier literature has quantified the information content from sentiment by counting positive and negative words in corporate reports, based on respectable lexicons, such as the Harvard Psychosociological Dictionary, specifically, the Harvard-IV-4 TagNeg (H4N) file. Such word lists were originally developed for human readers and for general purposes, and over time they have come to serve as an objective standard for researchers to analyze the sources and consequences of tone and sentiment, as perceived by the general readership, in corporate disclosures and new media (Tetlock 2007; Tetlock, Saar-Tsechansky, and Macskassy 2008; Hanley and Hoberg 2010). However, the meaning and tone of English words are highly context- and discipline-specific, and a general word categorization scheme might not translate effectively into a specialized field, such as finance. This motivated the influential work by LM (2011), which presented a specialized dictionary of positive and negative words that fits the unique text of financial situations. According to LM (2011), almost three-fourths of the words identified by the Harvard dictionary as negative (such as "liability") are words typically not considered negative in financial contexts. The LM (2011) dictionary has since become the leading lexicon used in algorithms for sentiment calibration.[18]

[18] Though Loughran and McDonald (2011) was in public circulation earlier (posted on the SSRN since 2009), its publication generated discrete growth in the impact of the dictionary: Google citation counts rose from 10 times prior to 2011 to 243 times by 2013, and had grown exponentially to 3,700 as of April 2022. Their word list has

The timeline of the Harvard General Inquirer dictionary (existing since 1996) and the Loughran-McDonald dictionary (since 2011) and their differential adoption by human versus machine readers provide a unique setting for us to test how the writing of corporate filings adjusts to AI readers. Our model discussed in Section 1 predicts that firms tend to increase tone management when more machine readers are present. We consider the following regression at the filing level, indexed by firm(i)-filing(j)-date(t), with year and firm (or industry) fixed effects:

$$\begin{aligned} NegativeSentiment_{i,j,t} = &\ \beta_1 MachineDownloads_{i,j,t} \times Post_t \\ &+\beta_2 MachineDownloads_{i,j,t} \\ &+\delta OtherDownloads_{i,j,t} + \gamma Control_{i,year} \\ &+\alpha_i(\alpha_{SIC3}) + \alpha_{year} + \epsilon_{i,j,t}. \end{aligned} \tag{4}$$

The equation above uses three versions of the dependent variable *Negative sentiment*: the *LM sentiment*, the *Harvard sentiment*, and their difference *LM – Harvard sentiment*, as defined in Section 2.2.3. We only consider the prevalence of negative words because earlier research (Tetlock 2007; LM, 2011; Cohen, Lou, and Malloy 2020) indicates that positive words are not informative of firm future outcomes or stock returns. *Post* is an indicator variable for years that came after the publication of LM (2011) and is equal to one for filings in 2012 onward, and zero otherwise. Filings in 2011 are excluded from the analysis. The year fixed effect subsumes the variable *Post* on its own.

Given the model's predictions that AI readers shape the style and quality of corporate writing, we expect the difference-in-differences coefficient β_1 to be significantly negative for *LM sentiment*, but not for *Harvard sentiment*. That is, there should be a differential relation between *LM sentiment* and *Machine downloads* during the *Post* period (after the publication of LM (2011)) relative to before, but a similar change around 2011 should be absent for *Harvard sentiment*. Such an exclusive set of effects is confirmed by results in Table 5.

Table 5 shows an unambiguous contrast before and after 2011, the year when the paper was published, on the effect of measures related to LM (2011). Post-2011, a one-standard-deviation increase in *Machine downloads* is associated with a 9- to 11-basis-point incremental decrease in *LM sentiment*, on top of an insignificant (column 3 with industry fixed effects) or much smaller (column 4 with firm fixed effects) effect during the pre-2011 period. The incremental effect post-2011, significant at the 1% level, represents about 5% of the sample mean of *LM sentiment*, or 0.15 standard deviations. In contrast, the coefficient

been adopted for the WRDS SEC Sentiment Data. The dictionary has been frequently featured in industry white papers and technical reports, such as in Marinov (2019) by the Man Group.

Table 5
Machine downloads and sentiment: Loughran and McDonald (2011) publication

	(1)	(2)	(3)	(4)	(5)	(6)
Dependent variable	*LM – Harvard sentiment*		*LM sentiment*		*Harvard sentiment*	
Machine downloads	−0.072***	−0.079***	−0.062***	−0.050***	0.010	0.029***
× *Post*	(−6.95)	(−8.94)	(−4.98)	(−4.99)	(0.76)	(2.65)
Machine downloads	−0.007	−0.011**	−0.009	−0.019***	−0.002	−0.008
	(−1.17)	(−2.46)	(−1.18)	(−3.72)	(−0.23)	(−1.43)
Observations	158,578	158,515	158,578	158,515	158,578	158,515
R-squared	.217	.568	.241	.632	.208	.590
Control variables	Yes	Yes	Yes	Yes	Yes	Yes
Firm FE	No	Yes	No	Yes	No	Yes
Industry FE	Yes	No	Yes	No	Yes	No
Year FE	Yes	Yes	Yes	Yes	Yes	Yes

This table reports the impact of the publication of Loughran and McDonald (2011) on the relation between the negative sentiment of a firm's filing and the machine downloads of the firm's past filings. *Machine downloads* measures the expected machine readership of a filing. *LM sentiment* (*Harvard sentiment*) is the number of Loughran-McDonald finance-related (Harvard General Inquirer) negative words in a filing, scaled by the total number of words in the filing. *LM – Harvard sentiment* is the difference between *LM sentiment* and *Harvard sentiment*. *Post* is an indicator variable equal to one for filings in 2012 onward, and zero for filings in 2010 and before. Control variables include *Other downloads*, *Size*, *Tobin's q*, *ROA*, *Leverage*, *Growth*, *Industry adjusted return*, *Institutional ownership*, *Analyst coverage*, *Idiosyncratic volatility*, *Turnover*, and *Segment*. The appendix defines all variables. The t-statistics, in parentheses, are based on standard errors clustered by firm. *p <.1; **p <.05; ***p <.01.

for *Harvard sentiment* is positive in both columns 5 and 6, and even statistically significant in column 6 with refined firm fixed effects. This evidence is suggestive of a substitution effect; that is, managers use negative words from the Harvard dictionary in place of synonyms from the LM list. Finally, columns 1 and 2 show that the relation between *LM – Harvard sentiment* and *Machine downloads* conforms to that of *LM sentiment*, confirming that the differential effect is mainly driven by reduced *LM sentiment*.

Results in Table 5 keep the possibility open that the publication of LM (2011) merely reflects a general trend of a strengthening relation between the machine downloads and avoiding using words that are perceived to have negative connotations in the finance context. Such a possibility still supports the general thesis that machine readership affects disclosure quality; nevertheless, a "parallel pre-trend" would allow a sharper identification on the impact of a new lexicon available to machine reading. Figure 3 illustrates the structural break, instead of a preexisting and continuing trend, around 2011. More specifically, we aggregate the *LM – Harvard sentiment* at the annual level separately for filings that are in the top and bottom terciles of *Machine downloads* in each year. Figure 3 plots the time series of the incremental tendency to use LM-negative words over Harvard-negative words by the two groups of filings.

Figure 3 shows a parallel pre-trend of the two groups until 2011 and then a clear divergence afterward. Before 2011, filings in the top and bottom terciles of *Machine downloads* exhibit clustered movements in the *LM – Harvard sentiment*. Afterward, the top tercile's sentiment trends down relative to that of the bottom tercile. We note a general trend, among all firms, to use

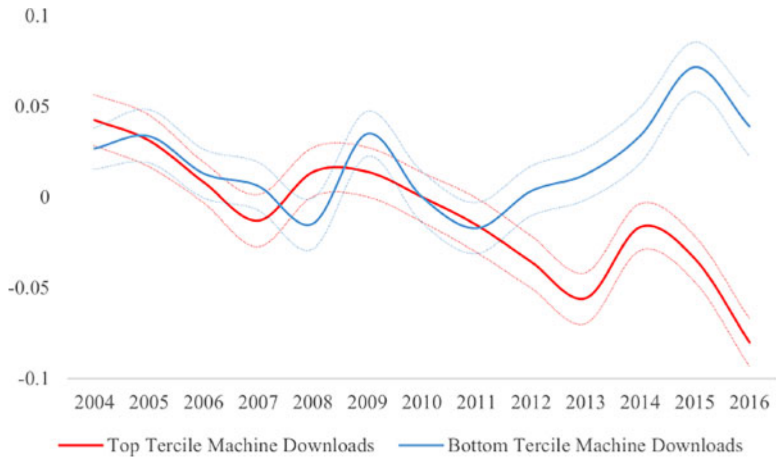

Figure 3
Sentiment trend and machine downloads
This figure plots *LM – Harvard sentiment* of 10-K and 10-Q filings and compares the sentiment of firms with high machine downloads with that of the low group. *LM – Harvard sentiment* is the difference of *LM sentiment* and *Harvard sentiment*. *LM sentiment* is defined as the number of Loughran-McDonald (LM) finance-related negative words in a filing divided by the total number of words in the filing. *Harvard sentiment* is defined as the number of Harvard General Inquirer negative words in a filing divided by the total number of words in the filing. Filings are sorted into top tercile or bottom tercile based on *Machine downloads*, defined in the appendix. *LM sentiment* and *Harvard sentiment* are normalized to one, respectively, in 2010 within each group, one year before the publication of Loughran and McDonald (2011). The dotted lines represent the 95% confidence limits.

fewer negative words in disclosures, which may reflect a growing awareness among firms of the perception induced by linguistic sentiments after the first generation of textual research. After the LM (2011) list was published, clearer and more practical guidance became available. Figure 3 suggests that firms with high machine readership were more motivated to avoid negative words that could feed into machine reading, leading to divergence.

Given the quasi-randomness of the event year 2011 due to the long and unpredictable time period for finance research to appear in print,[19] it is unlikely that the publication of LM (2011) perfectly timed a structural break in the tone management by corporations that would have materialized in the paper's absence. In other words, it is implausible that the LM dictionary summarizes the practice that was already in place, and that it serves as a coincidentally concurrent sideshow. Table 5 and Figure 3 thus provide more support to the hypothesis that corporate writing has been adjusted to serve machine readers, and this shift was affected by the availability of the LM dictionary.

Given the aggregate evidence that firms avoid words that are likely to be classified as negative by algorithms, we are curious to further uncover which words have become the least welcome. Out of all words classified

19 A recent paper by Dai et al. (2023) shows that the typical eventually published finance paper takes about 3 years to come to publication fruition, with a standard deviation of 1.8 years.

as negative by the LM dictionary, but not the Harvard dictionary, we are able to compare the frequencies they appear in filings pre- (2004–2010) and post-2011 (2012–2016). Sorted by the reduction in the average frequency per filing, the ten most avoided words are: "restructuring," "termination," "restatement," "declined," "correction," "misstatement," "terminated," "late," "alleged," and "omitted." The reduction amounts to 0.15 times to 0.35 times per filing. Sorted by the percentage reduction, that is, the reduction in frequency scaled by the frequency in the pre-2011 period,[20] the ten most avoided words are "restatement," "declined," "misstatement," "closure," "late," "dismissed," "inquiry," "alleged," "omitted," and "restructuring." The reduction in these words amounts to 10% to 35%.

To refute the alternative hypothesis that reduction in the LM-negative words could be part of the time trend that firms experienced fewer negative events coming out of the Financial Crisis since 2010, we present a comparison for pre- and post-2011 of major "negative" events, forced CEO turnover (based on data provided by authors of Peters and Wagner 2014), restatement, bankruptcy, operating losses, and default.

We show that firms in the high and low machine downloads groups do not exhibit significant preexisting differences in all outcome variables except $Restatement$. We find it plausible that a firm with frequent restatements tends to attract a lower level of machine readership as information contained in restatements is usually difficult to standardize and often involves external reference links. More importantly, none of the variables exhibits a divergence between the two groups of firms post-2011 (see Table IA.7 in the Internet Appendix). Further, in Table IA.8, we replicate the main results in Tables 3 and 5 with additional control variables that proxy for all main aspects of firm fundamentals (that might be correlated with the occurrence of negative events), including size, Tobin's q, ROA, and the Altman Z-score, as well as these negative economic outcomes. The key variables $Machine\ Downloads$ and $Machine\ Downloads \times Post$ retain coefficients that are qualitatively similar.

4.2 Managing other textual tones with machine readers

In addition to providing lists of sentiment-related words, LM (2011) also constructs lists of "tone" words, tailored to the financial context, aiming to capture litigiousness, uncertainty, and weak and strong modality. The expanded dictionary allows machines to assess more dimensions of a document's connotations. LM (2011) discovers that the stock market responds less positively to disclosures using more negative, uncertain, strong modal, and weak modal words, and that firms with a high proportion of negative or strong modal words are more likely to report material weakness. Given the market reaction, it is reasonable to expect managers to adjust tone along these

[20] Some words which show up infrequently before 2011 but never appear after 2011 would have a percentage reduction of -100%. We only consider words with an average frequency per filing of no less than 0.5 times.

Table 6
Machine downloads and other tones: Loughran and McDonald (2011) publication

Dependent variable	(1) *Litigious*	(2) *Uncertainty*	(3) *Weak modal*	(4) *Strong modal*
Machine downloads × *Post*	−0.057***	−0.021***	−0.034***	−0.007***
	(−6.02)	(−3.49)	(−8.86)	(−4.39)
Machine downloads	0.007	−0.009***	−0.021***	−0.004***
	(1.44)	(−3.05)	(−10.05)	(−4.98)
Observations	158,515	158,515	158,515	158,515
R-squared	.509	.600	.624	.571
Control variables	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes

This table reports the impact of the publication of Loughran and McDonald (2011) on the relation between the various tones of a firm's filing and the machine downloads of the firm's past filings. *Machine downloads* measures the expected machine readership of a filing. *Litigious*/*Uncertainty*/*Weak modal*/*Strong modal* is the number of Loughran-McDonald litigation-related/uncertainty-related/weak modal/strong modal words in a filing, scaled by the total number of words in the filing. *Post* is an indicator variable equal to one for filings in 2012 onward, and zero for filings in 2010 and before. Control variables include *Other downloads*, *Size*, *Tobin's q*, *ROA*, *Leverage*, *Growth*, *Industry adjusted return*, *Institutional ownership*, *Analyst coverage*, *Idiosyncratic volatility*, *Turnover*, and *Segment*. The appendix defines all variables. The t-statistics, in parentheses, are based on standard errors clustered by firm. *p <.1; **p <.05; ***p <.01.

dimensions after the methodology became publicly known. We reestimate Equation (4) by replacing the dependent variable with *Litigious*, *Uncertainty*, *Weak modal*, and *Strong modal*, which are all defined in Section 2.2.3 as well as in the appendix:

$$\begin{aligned} Tone_{i,j,t} = {} & \beta_1 MachineDownloads_{i,j,t} \times Post_t + \beta_2 MachineDownloads_{i,j,t} \\ & + \delta OtherDownloads_{i,j,t} + \gamma Control_{i,year} + \alpha_i(\alpha_{SIC3}) + \alpha_{year} + \epsilon_{i,j,t}. \end{aligned} \tag{5}$$

If managers have adjusted the frequency of LM-negative words based on their knowledge about investor reaction to sentiment, they should then be expected to also understand the impact of other tones documented in LM (2011). Given LM (2011) discovery that the frequency of all four tones was met with negative stock market reactions, we conjecture that managers of firms with high expected machine readership should moderate these words after 2011. Results in Table 6 support such a prediction. The coefficients associated with *Machine downloads* × *Post* are significant (at the 5% level or better) for all four dependent variables. That is, post-2011 corporate reports expecting more machine readers are more likely to avoid conveying a sentiment, as evaluated by an algorithm, that is predictive of legal liabilities, that is indicative of uncertain prospects, or that exhibits too little or too much confidence and surety. Taking the coefficient from column 1, a one-standard-deviation increase in *Machine downloads* predicts a 0.19-standard-deviation decrease in the *Litigious* tone.

4.3 Equilibrium and cross-sectional effects

The empirical findings in the previous sections generate intriguing equilibrium implications. For corporate disclosures to remain informative to investors in equilibrium, the language used must be, to some extent, constrained to honesty and transparency. If firms can "positify" language unlimitedly in order to impress machine and human readers, the signals would quickly lose relevance, resulting in a babbling equilibrium (Crawford and Sobel 1982).[21] To remain in an equilibrium in which investors extract information from disclosures, we hypothesize that firms face heterogeneous costs and derive heterogeneous benefits when deviating from truthful and transparent language. Our model discussed in Section 1 predicts that the higher the benefits (and the lower the costs) of modifying tones and machine readability, the greater the firm will change its disclosure along these dimensions.

We test these predictions in two cross-sectional settings. The first test explores motives underlying positive disclosures by sorting firms by upcoming external financing needs, defined as the net total issuance in a given year in excess of that in the last year. The net total issuance is calculated as the sum of the net debt issuance (change in current and long-term debt) and the net equity issuance, scaled by book assets. We single out firms that fall into the top quartile of external financing needs and compare them with the rest of the sample firms. Results in columns 1 and 2 of Table 7 show that firms facing high external financing needs, which presumably present greater incentives to convey clear and positive communications to investors, are indeed more likely to increase machine readability. They are also more likely to economize on words that would be perceived negatively by textual analyzers (columns 3 and 4).

The second test builds on the premise that firms under tighter regulatory scrutiny or higher litigation risk are more constrained in mincing words. To sort on litigation risk, we follow Bertomeu et al. (2021), who developed a measure of machine-learning-predicted probability of litigation at the industry level using a broad set of variables capturing accounting, capital markets, governance, and auditing conditions.[22] Based on the predicted probability, we classify firms in the top-quartile industries as embodying high litigation risk, while the rest of the firms serve as controls. Columns 5 and 6 in Table 7 show that the reduction in the use of negative words after 2011 is significantly less pronounced among high litigation risk firms, presumably because such firms are more constrained in manipulating language in disclosures.

5. Out-of-Sample Tests: Recent Technology and Audio Tone

Despite the extensive tests conducted based on LM (2011), we have results based on a single event. Fortunately, the rapid evolution of AI technology

---

[21] Indeed, we find that the return predictability based on LM sentiment diminishes after 2011, consistent with an evolving "cheap talk" effect. However, such diminishing returns are also commonly associated with the publication of return predictability based on publicly observable signals (McLean and Pontiff 2016).

[22] We gather the data from Jeremy Bertomeu's website.

Table 7
Machine readability and sentiment: Cross-sectional effects in terms of costs and benefits

	(1)	(2)	(3)	(4)	(5)	(6)
Dependent variable	*Machine readability*		*LM – Harvard sentiment*		*LM – Harvard sentiment*	
	External financing needs				Litigation risk	
Groups	Top quartile	Other	Top quartile	Other	Top quartile	Other
Machine downloads × *Post*			−0.103***	−0.075***	−0.054***	−0.090***
			(−6.62)	(−7.55)	(−3.46)	(−8.81)
Machine downloads	0.107***	0.076***	−0.024***	−0.011**	−0.018**	−0.012**
	(10.28)	(13.37)	(−2.89)	(−1.96)	(−2.54)	(−2.16)
Difference of coefficients	0.031***		−0.028*		0.036**	
p-value	.004		.065		.027	
Observations	35,014	101,242	36,984	106,468	48,457	102,467
R-squared	.439	.365	.635	.572	.598	.591
Control variables	Yes	Yes	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes

This table explores the cross-sectional variation in the relation between machine readability (first two columns)/sentiment (last four columns) and the machine downloads of the firm's past filings. *Litigation Risk*, the machine learning-predicted probability of litigation at a firm's industry, and *External Financing Needs*, the excess net total issuance of a firm, are the sorting variables that separate the sample into the top quartile and the rest. *Machine downloads* measures the expected machine readership of a filing. *Machine readability* measures the ease at which a filing can be processed by an automated program. *LM – Harvard sentiment* measures the difference in sentiments based on Loughran-McDonald finance-related negative words and Harvard General Inquirer negative words. *Post* is an indicator variable equal to one for filings in 2012 onward, and zero for filings in 2010 and before. Control variables include *Other downloads*, *Size*, *Tobin's q*, *ROA*, *Leverage*, *Growth*, *Industry adjusted return*, *Institutional ownership*, *Analyst coverage*, *Idiosyncratic volatility*, *Turnover*, and *Segment*. The appendix defines all variables. Difference of coefficients compares the coefficients for variables of interest *Machine downloads* (first two columns) and *Machine downloads* × *Post* (last four columns) between the top-quartile group and the rest. The t-statistics, in parentheses, are based on standard errors clustered by firm. *p <.1; **p <.05; ***p <.01 (for the regression coefficients [two-tailed] and for the difference of coefficients [one-tailed]).

provides us "out-of-sample tests" to support the inferences developed in earlier sections. This section explores disclosure adaptation to newer natural language processing technology and AI audio analyzers.

5.1 Managing sentiment in response to recent technology (BERT)

In the first test, we study managerial disclosure adaptation to the Bidirectional Encoder Representations from Transformers (BERT), the current state-of-art for machine processing of text data. BERT was introduced in 2018 by a group of researchers at Google (Devlin et al. 2018), who also open-sourced the associated codes and model. BERT considers the sequential relations of words inside sentences and produces superior results in understanding the meanings of sentences.

Because the EDGAR Log File Data Set stopped in 2017 and BERT was published in 2018, our *Machine downloads* variable is not available for this test. Instead, we resort to *AI ownership* and *AI talent supply* developed in Section 2.2.1 as the key independent variables, both of which are proxies for the percentage of firm stocks held by investment companies with high potential for AI capabilities. The coverage of our key independent variables ends in 2019; hence we focus on a relatively close window, between 2016 and 2019, around the publication of BERT. We consider the following regression at the firm-year level, indexed by firm(i)-year(t), with year and firm fixed effects:

$$\begin{aligned} BERT\,Sentiment_{i,t} \;=\; & \beta AI\,Readership_{i,t} \times Post\text{-}BERT_t \\ & + \delta AI\,Readership_{i,t} + \gamma Control_{i,t} + \alpha_i + \alpha_t + \epsilon_{i,t}. \end{aligned} \tag{6}$$

The dependent variable, *BERT sentiment*, is the ratio of the number of negative sentences (based on BERT) to the total number of sentences in the key 10-K section most relevant to our context: Item 7 ("Management Discussion & Analysis (MD&A)"). That section is considered to be the focal place where management provides investors with its view of the financial performance and condition of the company. It is a common practice for researchers to focus on this item of 10-K for textual analysis so as to optimize on the ratio of informative disclosure to boilerplate language, as well as economizing on computation time (Loughran and McDonald 2011; Cohen, Malloy, and Nguyen 2020). A sensitivity check which also includes Item 1 ("Business (a description of the company's operations)") (see Table IA.9 in the Internet Appendix) shows that results that are indistinguishable from those in the main specification. The key independent variable *AI Readership* in (6) is either *AI ownership* or *AI talent supply*. In a difference-in-differences setting, reported in Table 8, we find that firms with higher *AI ownership* or *AI talent supply* reduce the representation of negative sentences significantly, relative to firms with lower AI-equipped investors, after the introduction of BERT in 2018.

5.2 Managing audio quality in conference calls with machine readers

Though the textual quality of disclosures is this study's focus, voice analytics, enabled by the development of modern machine-learning methods, provides an out-of-sample test for our hypothesis that corporate disclosure caters to machines. Starting around 2008, voice analytic software, such as the commercial Layered Voice Analysis (LVA) software and open-source software on GitHub, have gained attention among investors looking for an edge in information processing. Such software has enabled researchers to study managers' vocal expressions and their implications on capital markets (Mayew and Ventakachalam 2012; Hu and Ma 2021). If managers are aware that their disclosure documents could be parsed by machines, they should have realized that their machine readers may be also using voice analyzers to extract signals from vocal patterns and emotions contained in managers' speeches.

This section explores whether management adjusts the way they talk (on conference calls) when they expect that machines are listening, based on a sample of audio data of earnings-related conference calls from 2010 to 2016, as described in Section 2.2.4. The choice of the sample is motivated by two factors. First, conference calls are staged events that allow firms to interact with stock analysts and institutional investors. Importantly, Huang and Wermers (2022) find that institutional investors significantly react to the tone of calls

Table 8
Managing sentiment in response to recent technology (BERT)

	(1)	(2)	(3)	(4)
Dependent variable	*BERT sentiment*			
	NegSent/TotalSent		NegSent/TotalWords	
AI ownership × *Post-BERT*	−4.953**		−0.212***	
	(−2.49)		(−2.68)	
AI ownership	2.313		0.103	
	(1.26)		(1.39)	
AI talent supply × *Post-BERT*		−0.983***		−0.041***
		(−3.61)		(−3.98)
AI talent supply		−0.522		−0.010
		(−1.18)		(−0.65)
Observations	6,627	6,627	6,627	6,627
R-squared	.796	.796	.804	.804
Control variables	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes

This table examines the impact of the publication of BERT on the relation between the negative sentiment of a firm's 10-K filing and the machine readership on the firm's filing. *BERT sentiment* is defined as the number of negative sentences, scaled by the total number of sentences in columns 1 and 2, and scaled by the total number of words in columns 3 and 4, respectively. *AI ownership* is a firm's aggregate ownership of AI-equipped investment company shareholders. *AI talent supply* measures the local talent supplies to a firm's institutional shareholders, weighted by their ownership; the local talent supply is the available workforce with IT degrees in the state in which an investor is headquartered. *Post-BERT* is an indicator variable equal to one for filings after 2018, and zero before 2018. Control variables include *Size*, *Tobin's q*, *ROA*, *Leverage*, *Growth*, *Industry adjusted return*, *Institutional ownership*, *Analyst coverage*, *Idiosyncratic volatility*, *Turnover*, and *Segment*. The appendix defines all variables. The t-statistics, in parentheses, are based on standard errors clustered by firm. *p <.1; **p <.05; ***p <.01.

in their trades and holdings of stocks, and hence these calls should be the right venue to test any feedback effect. Second, vocal tones are inevitably affected by fundamentals: managers are more likely to exhibit positivity and excitement when firm fundamentals are strong and outlooks bright. By analyzing earnings calls, we can control for the underlying fundamentals by including earnings surprise in the regressions.

Since there are no data on downloads of conference calls, we keep *Machine downloads* of a firm's filings as the proxy for the prevalence of "machine listeners," based on the premise that *Machine downloads* represents investors' propensity to deploy AI tools in analyzing corporate disclosures. Table 9 reports the results from the following regression at the conference call level, indexed by firm(i)-call(k)-date(t), with year and firm (or industry) fixed effects:

$$Emotion_{i,k,t} = \beta Machine Downloads_{i,k,t} + \delta Other Downloads_{i,k,t} + \gamma Control_{i,year} + \alpha_i(\alpha_{SIC3}) + \alpha_{year} + \epsilon_{i,k,t}. \tag{7}$$

We measure emotion along two dimensions developed in psychology, *Valence* and *Arousal*, that capture positivity and intensity of vocal tones, respectively (Russell 1980).

The first three columns of Table 9 show that higher *Machine downloads* is associated with higher *Valence*, or positivity in vocal emotion. A one-standard-deviation increase in *Machine downloads* is associated with a 0.28-standard-deviation higher *Valence*. The last three columns of Table 9 indicate a positive,

Table 9
Machine Downloads and Managers' Emotion during Conference Calls

	(1)	(2)	(3)	(4)	(5)	(6)
Dependent variable		*Emotion valence*			*Emotion arousal*	
Machine downloads	0.043***	0.042***	0.042***	0.004*	0.005**	0.007**
	(11.40)	(11.14)	(8.84)	(1.79)	(2.28)	(2.49)
Other downloads	−0.017***	−0.017***	−0.012***	−0.006***	−0.006***	−0.006***
	(−5.74)	(−5.67)	(−3.12)	(−3.65)	(−3.71)	(−2.92)
Observations	43,336	41,224	27,437	43,336	41,224	27,437
R-squared	0.389	0.383	0.388	0.395	0.395	0.469
Control Variables	No	Yes	Yes	No	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes

This table examines the relation between the manager's speech emotion during conference calls and the machine downloads of the firm's past filings. *Machine downloads* measures the expected machine readership of the most recent filing before a firm's conference call. *Emotion valence* and *Emotion arousal* measure the positivity and excitedness, respectively, of the conference call speech emotion. Control variables include *Size*, *Tobin's q*, *ROA*, *Leverage*, *Growth*, *Industry adjusted return*, *Institutional ownership*, *Analyst coverage*, *Idiosyncratic volatility*, *Turnover*, and *Segment* as in the previous tables. Columns 3 and 6 further include *EarningsSurprise* as an additional control. The appendix defines all variables. The sample consists of audio of conference calls between January 2010 and December 2016. The t-statistics, in parentheses, are based on standard errors clustered by firm. *p <.1; **p <.05; ***p <.01.

but much weaker, relation between *Machine downloads* and *Arousal*, a more exciting emotion in conference calls. In columns 3 and 6, $Control$ further includes *Earnings surprise*, defined as the difference between actual earnings and the median analyst forecast. Calculating the *Earnings surprise* variable requires analyst coverage (tracked by the IBES analyst data), which results in a much smaller sample. The coefficients associated with *Machine downloads* barely change.

Based on videos of entrepreneurs pitching investors for funding, Hu and Ma (2021) show that venture capitalists are more likely to invest in start-ups whose founders give pitches that are rated high in valence and arousal. Reactions by VC investors to vocal emotion may well apply to the general capital markets. Our findings support the hypothesis that managers are motivated to manipulate their vocal expressions to achieve a more favorable effect on investors that rely on machine processing, and also justify the anecdotal evidence that managers increasingly seek professional coaching in order to improve vocal performances (Wong 2012; Dizik 2017).

6. Concluding Remarks

This paper presents the first study showing how corporate disclosure in writing and speaking has been reshaped by machine readership employed by algorithmic traders and quantitative analysts. Our findings indicate that increasing AI readership motivates firms to prepare filings that are friendlier to machine parsing and processing, highlighting the growing roles of AI in the financial markets and their potential impact on corporate decisions. Firms manage sentiment and tone perception that caters to AI readers by, for example, differentially avoiding words perceived as negative by algorithms, as compared to those perceived as such by human readers. CEOs also aim to present with vocal qualities that are favorably rated by software. While the literature has shown how investors and researchers apply machine learning and computational tools to extract information from disclosures and news,[23] our study is the first to identify and analyze the *feedback effect*: how companies adjust the way they talk knowing that machines are listening. Such a feedback effect can lead to unexpected outcomes, such as manipulation and collusion (Calvano et al. 2020). The technological advancement calls for more studies to understand the impact of and induced behavior by AI in financial economics and in the broad society.[24]

Appendix

Table A1
Definitions of variables

Variable	Definition
After	An indicator variable equal to one if the time *m* occurs after a filing is publicly released on EDGAR. It is defined within the [-15, 15]-minute window, where minute 0 is the filing time
AI hedge fund	The percentage of shares outstanding owned by AI hedge funds, classified based on employees' work experience in AI-related projects disclosed on their LinkedIn profiles (Guo and Shi 2020). It is computed at the stock-quarter level from 13F holdings of hedge funds
AI ownership	The firm-year-level aggregate ownership of AI-equipped investment company shareholders in the quarter before the firm's current filing. We classify an investment company as having AI capacity if it has AI-related job postings in the past 5 years using the job posting data between 2011 and 2018 from Burning Glass
AI talent supply	We first retrieve the number of people between 18 and 64 with college or graduate school degrees in information technology, scaled by the population at the state-year level, using data between 2011 and 2018 from Integrated Public Use Microdata Series (IPUMS) surveys. Second, for each firm and during the quarter prior to the current filing, we aggregate state-year-level AI talents over all states based on the headquarters of the investors, weighted by their ownership
Analyst coverage	The natural logarithm of one plus the number of IBES analysts covering the stock
BERT sentiment	The number of negative sentences in Item 7 of a 10-K filing, scaled by the total number of sentences (or the total number of words), expressed in percentage points
Bid-ask spread	The difference between the ask price and the bid price scaled by the midpoint of them, expressed in percentage points and calculated at the minute level following the NBBO rule
Earnings surprise	The difference between the actual quarterly earnings and the median earnings forecast of IBES analysts, scaled by the stock price
Emotion arousal	The excitedness of speech emotion, calculated from a pretrained Python machine learning package *pyAudioAnalysis*

[23] Applications of more recent machine learning techniques in finance research include support vector regressions (Manela and Moreira 2017), word embedding and latent Dirichlet allocation (Li et al. 2020; Hanley and Hoberg 2019; Cong, Liang and Zhang 2019), and neural networks as well as ensemble models (Chen, Wu, and Yang 2019; Cao et al. 2022; Cao, Yang, and Zhang 2022).

[24] Sports provide an analogous example in a nonfinance setting. The English Premier League decided not to let Video Assistant Referee (VAR) overpower referee judgment. One main reason is that players will reverse-engineer and play to the rules underlying the VAR decisions, which will likely lead to undesirable outcomes, such as more "low-grade" (to the machine) but atrocious (to humans) fouls. See Reade (2020).

Table A1
(continued)

Variable	Definition
Emotion valence	The positivity of speech emotion, calculated from a pretrained Python machine learning package *pyAudioAnalysis*
External financing needs	The net total issuance in a given year in excess of that in the previous year. The net total issuance is calculated as the sum of the net debt issuance (change in current and long-term debt) and the net equity issuance, scaled by book assets
Growth	The average sales growth over the past 3 years
Harvard sentiment	The number of Harvard General Inquirer negative words in a filing divided by the total number of words in the filing, expressed in percentage points
Idiosyncratic volatility	The annualized idiosyncratic volatility (using daily data) from the Fama-French three-factor model
Industry adjusted return	The monthly average SIC3-adjusted stock returns over the past year
Institutional ownership	The ratio of the total shares of institutional ownership to shares outstanding
Leverage	The ratio of total debt to assets
Litigation risk	The machine-learning-predicted probability of litigation at the Fama-French 48-industry level using a broad set of variables capturing accounting, capital markets, governance, and auditing conditions, developed by Bertomeu et al. (2021)
Litigious	The number of Loughran-McDonald (LM) litigation-related words in a filing divided by the total number of words in the filing, expressed in percentage points
LM sentiment	The number of Loughran-McDonald (LM) finance-related negative words in a filing divided by the total number of words in the filing, expressed in percentage points
LM – Harvard sentiment	*LM sentiment* minus *Harvard sentiment*
Machine downloads	For a firm's filing at time t, *Machine downloads* is the natural logarithm of the average number of machine downloads of the firm's historical filings during the $[t-4, t-1]$ quarters. To measure machine downloads, we identify an IP address downloading more than 50 unique firms' filings daily as a machine visitor, the same criterion used by Lee, Ma, and Wang (2015). In addition, we include requests attributed to web crawlers in the Log File Data as machine initiated. Machine requests are aggregated for each filing within 7 days (i.e., days [0, 7]) after it becomes available on EDGAR
Δ*Machine downloads*	For a firm's filing at time t, the change in *Machine downloads* (before taking the natural logarithm) from the previous-year average. Δ*Machine downloads* is the natural logarithm of the change (A constant is added to ensure the number is positive before taking the natural logarithm).
Machine readability	The average of five filing attributes, including (a) *Table extraction*, the ease of separating tables from the text; (b) *Number extraction*, the ease of extracting numbers from the text; (c) *Table format*, the ease of identifying the information contained in the table (e.g., whether a table has headings, column headings, row separators, and cell separators); (d) *Self-containedness*, whether a filing includes all needed information (i.e., without relying on external exhibits); and (e) *Standard characters*, the proportion of characters that are standard ASCII (American Standard Code for Information Interchange) characters. Each attribute is standardized to a Z-score before being averaged to form a single-index *Machine readability* measure.
MR upgrade	An "upgrade" event at the filing (i, j, t) level equal to one if *Machine readability*, $MR_{i,j,t}$, incurs a significant (i.e., one-standard-deviation) increase over the previous-year average, $MR_{i,t-1}$, and zero otherwise.
Other downloads	For a firm's filing on day t, *Other downloads* is the natural logarithm of the average number of nonmachine downloads of the firm's historical filings during the $[t-4, t-1]$ quarters.
Post	An indicator variable equal to one for filings in 2012 onward, and zero for filings in 2010 and before (filings in 2011 are excluded from the analysis).

Table A1
(continued)

Variable	Definition
Post-BERT	An indicator variable equal to one for filings after 2018, and zero otherwise (filings in 2018, when BERT was published, are excluded from the analysis).
ROA	The ratio of EBITDA to assets
Segment	The number of business segments, following Cohen and Lou (2012). It measures the complexity of business operations
Size	The natural logarithm of the market capitalization
Strong modal	The number of Loughran-McDonald (LM) strong modal words in a filing divided by the total number of words in the filing, expressed in percentage points
Time to first directional trade	The length of time, in seconds, between the EDGAR publication time stamp and the first directional trade after a filing is publicly released, censored at the end of a 15-minute window. The first directional trade is the first buy (sell) trade at a price below (above) the terminal value at the end of the window, where buy- and sell-initiated trades are classified by the Lee and Ready (1991) algorithm
Time to first trade	The length of time, in seconds, between the EDGAR publication time stamp and the first trade of the issuer's stock, censored at the end of a 15-minute window
Tobin's q	The natural logarithm of the ratio of the sum of market value of equity and book value of debt to the sum of book value of equity and book value of debt
Turnover	The monthly average of the ratio of trading volume to shares outstanding, multiplied by 12
Uncertainty	The number of Loughran-McDonald (LM) uncertainty-related words in a filing divided by the total number of words in the filing, expressed in percentage points
Weak modal	The number of Loughran-McDonald (LM) weak modal words in a filing divided by the total number of words in the filing, expressed in percentage points

References

Abis, S., and L. Veldkamp. 2022. The changing economics of knowledge production. Working Paper, Columbia University.

Adusumilli, R. 2020. NLP in the stock market. *Medium*, February 1.

Ahern, K. R., and D. Sosyura. 2014. Who writes the news? Corporate press releases during merger negotiations. *Journal of Finance* 69:241–91.

Allee, K. D., M. D. DeAngelis, and J. R. Moon, Jr. 2018. Disclosure "scriptability." *Journal of Accounting Research* 56:363–430.

Angwin, J. 2011. Privacy study: Top U.S. websites share visitor personal data. *Wall Street Journal*, October 11.

Asay, H. S., R. Libby, and K. Rennekamp. 2018. Firm performance, reporting goals, and language choices in narrative disclosures. *Journal of Accounting and Economics* 65:380–98.

Balakrishnan, K., M. B. Billings, B. Kelly, and A. Ljungqvist. 2014. Shaping liquidity: On the causal effects of voluntary disclosure. *Journal of Finance* 69:2237–78.

Bernard, D., T. Blackburne, and J. Thornock. 2020. Information flows among rivals and corporate investment. *Journal of Financial Economics* 136:760–79.

Bertomeu, J., E. Cheynel, E. Floyd, and W. Pan. 2021. Using machine learning to detect misstatements. *Review of Accounting Studies* 26:468–519.

Björkegren, D., J. E. Blumenstock, and S. Knight. 2020. Manipulation-proof machine learning. Working Paper, Brown University.

Blankespoor, E. 2019. The impact of information processing costs on firm disclosure choice: Evidence from the XBRL mandate. *Journal of Accounting Research* 57:919–67.

Blankespoor, E., E. deHaan, and I. Marinovic. 2020. Disclosure processing costs, investors' information choice, and equity market outcomes: A review. *Journal of Accounting and Economics* 70:101344.

Bolandnazar, M., R. J. Jackson Jr, W. Jiang, and J. Mitts. 2020. Trading against the random expiration of private information: A natural experiment. *Journal of Finance* 75:5–44.

Bond, P., A. Edmans, and I. Goldstein. 2012. The real effects of financial markets. *Annual Review of Financial Economics* 4:339–60.

Burnham, K. 2014. LinkedIn sues after scraping of user data. *InformationWeek*, January 8.

Bushee, B. J. 1998, The influence of institutional investors on myopic R&D investment behavior. *Accounting Review* 73:305–33.

Bushee, B. J., and C. F. Noe. 2000. Corporate disclosure practices, institutional investors, and stock return volatility. *Journal of Accounting Research* 38:171–202.

Calvano, E., G. Calzolari, V. Denicolò, and S. Pastorello. 2020. Artificial intelligence, algorithm pricing, and collusion. *American Economic Review* 100:3267–97.

Cao, S. S., K. Du, B. Yang, and A. L. Zhang. 2021. Copycat skills and disclosure costs: Evidence from peer companies' digital footprints. *Journal of Accounting Research* 59:1261–302.

Cao, S. S., W. Jiang, J. Wang, and B. Yang. 2022. From man vs. machine to man + machine: The art and analyses of stock analyses. Working Paper, University of Maryland.

Cao, S. S., B. Yang, and A. L. Zhang, 2022. Managerial risk assessment and fund performance: Evidence from textual disclosure. Working Paper, University of Maryland.

Chen, H., L. Cohen, U. Gurun, D. Lou, and C. Malloy. 2020. IQ from IP: Simplifying search in portfolio choice. *Journal of Financial Economics* 138:118–37.

Chen, M. A., Q. Wu, and B. Yang. 2019. How valuable is FinTech innovation? *Review of Financial Studies* 32:2062–106.

Cohen, L., and D. Lou. 2012. Complicated firms. *Journal of Financial Economics* 104:383–400.

Cohen, L., D. Lou, and C. J. Malloy. 2020. Casting conference calls. *Management Science* 66:5015–39.

Cohen, L., C. Malloy, and Q. Nguyen. 2020. Lazy prices. *Journal of Finance* 75:1371–415.

Cong, L. W., T. Liang, B. Yang, and X. Zhang. 2021. Analyzing textual information at scale. In *Information to facilitate efficient decision making: Big data, blockchain and relevance*, ed. Kashi Balachandran, 239–72. New Jersey: World Scientific Publishers.

Cong, L. W., T. Liang, and X. Zhang. 2019. Textual factors: A scalable, interpretable, and data-driven approach to analyzing unstructured information. Working Paper, Cornell University.

Crane, A. D., K. Crotty, and T. Umar. 2022. Hedge funds and public information acquisition. *Management Science*. Advance Access published June 22, 2022, 10.1287/mnsc.2022.4466.

Crawford, V. P., and J. Sobel. 1982. Strategic information transmission. *Econometrica* 50:1431–51.

Da, Z., J. Engelberg, and P. Gao. 2011. In search of attention. *Journal of Finance* 66:1461–99.

Dai, R., L. Donohue, Q. Drechsler, and W. Jiang. 2023. Dissemination, publication, and impact of finance research: When novelty meets conventionality. *Review of Finance* 27:79–141.

Devlin, J., M. Chang, K. Lee, and K. Toutanova. 2018. BERT: Pre-training of deep bidirectional transformers for language understanding. Working Paper, Google AI Language.

Dhaliwal, D. S., O. Z. Li, A. Tsang, and Y. G. Yang. 2011. Voluntary nonfinancial disclosure and the cost of equity capital: The initiation of corporate social responsibility reporting. *Accounting Review* 86:59–100.

Diamond, D. W., and R. E. Verrecchia. 1991. Disclosure, liquidity, and the cost of capital. *Journal of Finance* 46:1325–59.

Dizik, A. 2018. How to listen for the hidden data in earnings calls. *Chicago Booth Review*, May 25.

Dong, J., A. Roth, Z. Schutzman, B. Waggoner, and Z. S. Wu. 2018. Strategic classification from revealed preferences. *Proceedings of the 2018 ACM Conference on Economics and Computation*, 55–70.

Driscoll, J. C., and A. C. Kraay. 1998. Consistent covariance matrix estimation with spatially dependent panel data. *Review of Economics and Statistics* 80:549–60.

Easley, D., and M. O'Hara. 2004. Information and the cost of capital. *Journal of Finance* 59:1553–83.

Foster, F. D., and S. Viswanathan. 1996. Strategic trading when agents forecast the forecasts of others. *Journal of Finance* 51:1437–78.

Gao, M., and J. Huang, 2020. Informing the market: The effect of modern information technologies on information production. *Review of Financial Studies* 33:1367–411.

Gara, A. 2018. Wall Street tech spree: With Kensho acquisition S&P Global makes largest A.I. deal in history. *Forbes*, March 6.

García, D. 2013. Sentiment during recessions. *Journal of Finance* 68:1267–300.

Giannakopoulos, T. 2015. pyAudioAnalysis: An open-source Python library for audio signal analysis. PloS ONE 10, e0144610.

Goldstein, I., and L. Yang. 2017. Information disclosure in financial markets. *Annual Review of Financial Economics* 9:101–25.

Graham, J. R., C. R. Harvey, and S. Rajgopal. 2005. The economic implications of corporate financial reporting. *Journal of Accounting and Economics* 40:3–73.

Guo, N. X., and Z. Shi. 2020. The impact of AI talents on hedge fund performance. Working Paper, Saint Louis University.

Hanley, K. W., and G. Hoberg. 2010. The information content of IPO prospectuses. *Review of Financial Studies* 23:2821–64.

Hanley, K. W., and G. Hoberg. 2019. Dynamic interpretation of emerging risks in the financial sector/ *Review of Financial Studies* 32:4543–603.

Hardt, M., N. Megiddo, C. Papadimitriou, and M. Wootters. 2016. Strategic classification. *Proceedings of the 2016 ACM Conference on Innovations in Theoretical Computer Science*, 111–122.

Hennessy, C., and C. Goodhart. 2021. Goodhart's Law and machine learning: A structural approach. Working Paper, London Business School.

Hodge, . D., J. J. Kennedy, and L. A. Maines. 2004. Does search-facilitating technology improve the transparency of financial reporting? *Accounting Review* 79:687–703.

Hu, A., and S. Ma, 2021, Persuading Investors: A Video-Based Study. Working paper, Yale University.

Huang, A. G., H. Tan, and R. Wermers. 2020. Institutional trading around corporate news: Evidence from textual analysis. *Review of Financial Studies* 33:4627–75.

Huang, A. G., and R. Wermers. 2022. Who listens to corporate conference calls? The effect of "soft information" on institutional trading. Working Paper, University of Waterloo.

Huang, A. H., H. Wang, and Y. Yang. 2022. FinBERT: A large language model for extracting information from financial text. *Contemporary Accounting Research*. Advance Access published 29 September, 2022, 10.1111/1911-3846.12832.

Hunter, G. S. 2020. Sweet-talking CEOs are starting to outsmart the robot analysts. *Bloomberg*, October 20.

Hwang, B., and H. H. Kim. 2017. It pays to write well. *Journal of Financial Economics* 124:373–94.

Jegadeesh, N., and D. Wu. 2013. Word power: A new approach for content analysis. *Journal of Financial Economics* 110:712–29.

Jiang, F., J. Lee, X. Martin, and G. Zhou. 2019. Manager sentiment and stock returns. *Journal of Financial Economics* 132:126–49.

Jiang, W., Y. Tang, R. J. Xiao, and V. Yao. 2021. Surviving the FinTech disruption. Working Paper, Emory University.

Kim, C., K. Wang, and L. Zhang. 2019. Readability of 10-K reports and stock price crash risk. *Contemporary Accounting Research* 36:1184–216.

Kim, O., and R. E. Verrecchia. 1994. Market liquidity and volume around earnings announcements. *Journal of Accounting and Economics* 17:41–67.

Kim, O., and R. E. Verrecchia. 1997. Pre-announcement and event-period private information. *Journal of Accounting and Economics* 24:395–419.

Kothari, S.P., S. Shu, and P. D. Wysocki. 2009. Do managers withhold bad news? *Journal of Accounting Research* 47:241–76.

Kyle, A. S. 1985. Continuous auctions and insider trading. *Econometrica* 53:1315–35.

Lee, C. M., P. Ma, and C. C. Wang. 2015. Search-based peer firms: Aggregating investor perceptions through internet co-searches. *Journal of Financial Economics* 116:410–31.

Lee, C. M., and M. J. Ready. 1991. Inferring trade direction from intraday data. *Journal of Finance* 46:733–46.

Li, K., F. Mai, R. Shen, and X. Yan. 2020. Measuring corporate culture using machine learning. *Review of Financial Studies* 34:3265–315.

Loughran, T., and B. McDonald. 2011. When is a liability not a liability? Textual analysis, dictionaries, and 10-Ks. *Journal of Finance* 66:35–65.

Loughran, T., and B. McDonald. 2014. Measuring readability in financial disclosures. *Journal of Finance* 69:1643–71.

Loughran, T., and B. McDonald. 2016. Textual analysis in accounting and finance: A survey. *Journal of Accounting Research* 54:1187–230.

Loughran, T., and B. McDonald. 2017. The use of EDGAR filings by investors. *Journal of Behavioral Finance* 18:231–48.

Manela, A., and A. Moreira. 2017. News implied volatility and disaster concerns. *Journal of Financial Economics* 123:137–62.

Marinov, S. 2019. Natural language processing in finance: Shakespeare without the monkeys. *Man Institute*, July.

Mayew, W. J., and M. Venkatachalam. 2012. The power of voice: Managerial affective states and future firm performance. *Journal of Finance* 67:1–43.

McLean, R. D., and J. Pontiff. 2016. Does academic research destroy stock return predictability? *Journal of Finance* 71:5–32.

Milli, S., J. Miller, A. D. Dragan, and M. Hardt. 2019. The social cost of strategic classification. *Proceedings of the Conference on Fairness, Accountability, and Transparency*, 230–9.

Naughton, J. 2020. Companies are now writing reports tailored for AI readers – and it should worry us. *The Guardian*, December 5.

Peters, F. S., and A. F. Wagner. 2014. The executive turnover risk premium. *Journal of Finance* 69:1529–63.

Reade, J. 2020. Why has the introduction of video technology gone so badly in soccer? *Forbes*, December 10.

Russell, J. A. 1980. A circumplex model of affect. *Journal of Personality and Social Psychology* 39:1161–78.

Satariano, A., and N. Kumar. 2017. The massive hedge fund betting on AI. *Bloomberg*, September 27.

Tetlock, P. C. 2007. Giving content to investor sentiment: The role of media in the stock market. *Journal of Finance* 62:1139–68.

Tetlock, P. C., M. Saar-Tsechansky, and S. Macskassy. 2008. More than words: Quantifying language to measure firms' fundamentals. *Journal of Finance* 63:1437–67.

Verrecchia, R. E. 2001. Essays on disclosure. *Journal of Accounting and Economics* 32:97–180.

Wigglesworth, R. 2020. Robo-surveillance shifts tone of CEO earnings calls. *Financial Times*, December 5.

Wong, S. 2012. Listening without prejudice: How the experts analyze earnings calls for lies, bluffs, and other flags. *Minyanville*, April 18.