

Show Me the Use Case: Internal AI Adoption in Corporate Earnings Calls

Luis Sacouto
Luis.Sacouto@qu.edu

Guido Lang
Guido.Lang@qu.edu

Johannes January
Johannes.January@qu.edu

Tan Gürpınar
Tan.Gurpinar@qu.edu

Quinnipiac University
Hamden, CT, USA

Abstract

Organizational artificial intelligence (AI) adoption is one of the most impactful technological transitions of the decade, yet existing evidence on organizations comes from surveys that capture stated intent rather than disclosed operational deployments. This study examines how S&P 500 companies disclose AI deployed within their own operations during quarterly earnings calls from Q1 2020 through Q1 2026. Using a two-stage extraction process combining rule-based keyword identification and large language model classification with a three-part qualification test, we identify 3,907 internal AI use cases from 2,037 earnings calls across 420 S&P 500 companies and code each along dimensions including use-case category, workforce impact, and workforce group. Among calls that mention AI, the share including at least one internal AI deployment rose from 43% in 2020 to 75% by Q1 2026, with a post-generative-AI inflection beginning in 2023. Enterprise productivity and workflow automation accounts for 31% of internal AI use cases, followed by supply chain and operations (20%), sales and marketing personalization (13%), customer service and support automation (11%), and code generation and software development (5%). While augmentation framing predominates at 91%, displacement framing appears in 3.1% of use cases, concentrated in customer service and frontline workers, knowledge workers, and finance and back-office roles. These patterns characterize the investor-facing narrative of organizational AI adoption and carry implications for information systems research, management practice, and workforce policy discussions that rely on corporate voluntary disclosure as evidence.

Keywords: artificial intelligence, earnings calls, internal AI adoption, workforce framing, corporate disclosure

AI in the Workplace: How S&P 500 Companies Disclose Internal AI Adoption in Quarterly Earnings Calls

Luis Sacouto, Guido Lang, Johannes January and Tan Gürpınar

1. INTRODUCTION

Quarterly earnings calls are among the most information-dense moments in the corporate calendar, combining prepared executive remarks with unscripted analyst questions in a setting where disclosure choices directly shape investor expectations. As artificial intelligence has risen to the top of the corporate agenda, these calls have become a primary source on how companies characterize their AI programs, not as reported in after-the-fact surveys but as communicated by executives under market scrutiny. Within earnings call AI disclosures, use cases in which companies describe AI deployed in their own operations carry particular significance for questions about labor and organizational transformation.

The distinction matters because the workforce consequences of AI are most direct in internal deployments, where the same organization that adopts the technology also employs the workers whose roles it may augment or displace. Survey-based evidence on organizational AI adoption (McElheran et al., 2024; Jöhnk et al., 2021) captures intentions and capabilities rather than specific operational use cases and cannot show how companies characterize the labor implications of the systems they deploy, whereas earnings call transcripts offer a complementary perspective, produced every quarter across hundreds of companies, addressed to investors with a direct stake in productivity outcomes, and increasingly detailed about what AI is doing and to whom.

Existing applications of natural language processing to earnings calls aggregate all AI-related mentions without isolating deployments within the reporting company's own operations (Hassan et al., 2019; Cao et al., 2024), while prior firm-level measures rely on occupational exposure scores (Eisfeldt et al., 2023) or résumé-based proxies for AI investment (Babina et al., 2024) that estimate where AI is likely to hit or how much firms invest in AI talent rather than what firms describe deploying inside their operations. By focusing exclusively on internal AI deployments and extracting structured use-case records with workforce framing, we analyze a more precisely defined corpus and address a question about organizational AI adoption that prior firm-level proxies cannot directly answer. Survey-based measures of adoption intent (McElheran et al., 2024) and aggregate AI-mention counts that do not isolate a firm's own operations (Hassan et al., 2019; Cao et al., 2024) do not extract individual, workforce-coded internal AI use cases at the level of a single disclosed claim, validated against human coding, as this study does

Our central research question asks: *How do large publicly traded companies disclose internal AI adoption in investor-facing communications, with particular attention to the types of AI use cases deployed, the workforce groups most frequently implicated, and the framing of workforce impact across augmentation, displacement, and related dimensions.* We address this question using 3,907 internal AI use cases drawn from 2,037 earnings calls across 420 S&P 500 companies spanning Q1 2020 through Q1 2026.

The study contributes to three streams of research. First, it advances the literature on organizational AI adoption (McElheran et al., 2024; Davenport & Ronanki, 2018) by providing a longitudinal inventory of deployment types drawn from primary executive communications rather than surveys or retrospective case studies (Hassan et al., 2019; Cao et al., 2023). Second, it contributes to the literature on large language model-based extraction from corporate text (Bochkay et al., 2023; Siano, 2025) by demonstrating that structured multi-dimensional use-case records can be reliably extracted at scale across a broad cross-section of industries. Third, it informs the AI-and-workforce literature (Acemoglu & Restrepo, 2019; Felten et al., 2021) by documenting that displacement framing, while a minority pattern, concentrates in occupational groups with known automation exposure at a rate substantially above what aggregate AI disclosure analyses report.

2. RELATED LITERATURE

2.1 Organizational AI Adoption and Diffusion

Research on technology adoption at the organizational level provides the theoretical backdrop for interpreting why internal AI deployment patterns follow the trajectory we observe. Rogers (2003) identifies relative advantage, compatibility, and observability as primary drivers of adoption velocity, with observability offering a natural explanation for the acceleration of internal AI adoption following the high-visibility launch of ChatGPT, which made AI capabilities legible to executives across all sectors simultaneously. Venkatesh et al. (2003) establish that performance expectancy is the strongest individual-level adoption driver, a dynamic whose organizational analog is evident in how companies frame the returns from internal AI deployment in operational efficiency terms when communicating with investors.

McElheran et al. (2024), drawing on U.S. Census data covering 850,000 firms, find that fewer than 6% used any AI technology as of 2018, with adoption concentrated at very large firms whose earnings calls form our corpus, making this dataset a well-targeted window into AI adoption at the frontier. Brynjolfsson et al. (2021) argue that AI investments precede measurable productivity gains because complementary organizational investments must co-develop in a Productivity J-Curve dynamic, a framework that explains why companies describe deployed systems before reporting quantified outcomes and why ROI evidence rates in earnings calls rise gradually rather than immediately upon adoption. Davenport and Ronanki (2018) propose a three-category AI use-case taxonomy covering process automation, cognitive insight, and customer engagement that maps closely onto the categories we observe in internal AI disclosures and reflects vocabulary executives employ.

2.2 Textual Analysis of Corporate Communications

The use of corporate text as a data source has a substantial tradition in accounting and information systems research. Loughran and McDonald (2011) show that finance-domain word lists outperform general-purpose sentiment dictionaries for corporate text analysis, and their subsequent survey (2016) documents the scalability limitations of word-count approaches that large language model-based extraction now addresses. Hassan et al. (2019) demonstrate that text-based measures of earnings call content provide valid cross-firm comparisons that predict subsequent firm behavior, laying the methodological groundwork for treating transcript content as structured, analyzable data. Bochkay et al. (2023) survey transformer-based approaches applied to accounting texts and identify the contextual sensitivity gains these architectures offer over earlier methods, while Siano (2025) shows that large language model-based processing of earnings calls explains three times more variation in short-window stock returns than dictionary-based methods.

Cao et al. (2023) demonstrate that firms strategically manage textual disclosure content in response to machine readership, reducing negative sentiment in filings accessed more frequently by algorithmic readers, a finding that implies the framing patterns we observe in earnings call transcripts reflect deliberate strategic choices rather than neutral operational reporting. Cao et al. (2024) find that voluntary AI disclosures with clear implementation plans generate significant valuation benefits while speculative mentions do not, placing direct market stakes on the disclosure framing choices we analyze. Jia et al. (2026) examine managerial discussions of generative AI in earnings conference calls following ChatGPT's release and find that disclosures increased substantially, became more action-oriented, and varied systematically across firm size, growth profile, and sector exposure, establishing earnings-call AI discussion as a meaningful and increasingly structured channel of corporate communication.

2.3 AI and Workforce Transformation

Acemoglu and Restrepo (2019) provide the conceptual framework distinguishing automation effects, in which capital replaces labor in existing tasks, from the creation of new tasks where labor retains a comparative advantage, with net employment effects depending on the relative rates of both processes. This displacement-versus-reinstatement duality is the theoretical anchor for our workforce impact coding and for interpreting why augmentation framing predominates in internal AI disclosures despite the cost-reduction outcomes companies simultaneously report. Felten et al. (2021) construct an

occupational AI exposure index linking AI application categories to occupational task demands, supporting cross-occupational comparison of disclosure patterns in our workforce category analysis. Building on this exposure approach, Eisfeldt et al. (2023) construct firm-level scores of workforce exposure to generative AI from firms' occupational composition, while Babina et al. (2024) measure firm-level AI investment from the share of AI-skilled workers each firm employs. These approaches predict where AI is likely to hit or capture how much firms invest in AI talent, but neither measures what firms describe as their actual internal deployments. Shrestha et al. (2019) provide a contingency framework for when organizations should delegate decisions to AI versus retain human judgment, informing the distinction between augmentation and displacement framings across different use-case categories.

Announcing productivity gains attracts investors, while announcing workforce reduction creates legal, labor, and reputational risk. These asymmetric incentives create systematic pressure toward augmentation framing regardless of actual outcomes (Cao et al., 2023). The 3.1% displacement rate in internal AI disclosures suggests that the efficiency reality of operational AI breaks through this framing pressure in some categories and occupational groups more than others, reaching levels that aggregate AI disclosure analyses are unlikely to detect.

3. METHODOLOGY

The analysis covers S&P 500 quarterly earnings calls from Q1 2020 through Q1 2026. Transcript collection targeted S&P 500 member firms over this period, sourcing calls from financial data providers. Processing followed a two-stage pipeline. The first stage applied a rule-based keyword scan using terms including "artificial intelligence," "machine learning," "large language model," "generative AI," "foundation model," "neural network," "deep learning," and specific AI product name variants to identify calls with meaningful AI content. The scan also incorporated compound patterns for terms such as "natural language processing," "computer vision," and "predictive analytics."

The second stage passed each identified call to Claude Sonnet (claude-sonnet-4-6) via the Anthropic Messages Batches API with a structured extraction prompt targeting AI use cases in which the reporting company deploys AI within its own operations. This extraction approach extends an emerging line of work using large language models to extract structured records from heterogeneous unstructured text corpora at scale (Sa-Couto et al., 2025; SaCouto et al., 2026), where GenAI-based extraction has been shown to replicate manual coding with comparable accuracy at substantially greater scale. Each use case was required to satisfy a three-part qualification test: an identifiable AI or machine learning capability, a specific task or process it performs, and an identifiable beneficiary within the firm's own workforce or operations. Use cases without an identifiable internal workforce or operational beneficiary were excluded, as were vague aspirational mentions lacking a named capability or task. The model returned structured JSON records coding up to 15 qualifying use cases per call, each coded on workforce impact (augmentation, displacement, reskilling, new_hiring, or none), workforce group, and use-case category. The use-case categories were derived inductively from the data rather than specified in advance: new categories were added as novel use-case types emerged across the growing corpus and merged with existing ones when sufficiently similar patterns recurred.

To assess extraction quality, two of the authors independently hand-coded a stratified random sample of 100 use cases drawn proportionally across sectors and across the three-part test dimensions. Coders evaluated each LLM-extracted record on three criteria: whether the use case qualified as a concrete internal AI deployment, whether the assigned use-case category was correct, and whether the workforce impact label matched the quoted evidence. Cohen's kappa between human coders and the LLM output was 0.78 for the inclusion decision, 0.71 for use-case category, and 0.69 for workforce impact, all in the substantial agreement range (Landis & Koch, 1977). Inter-coder agreement between the two human raters was 0.82 across the three dimensions.

The entire analysis process yielded 3,907 internal AI use cases from 2,037 earnings calls across 420 companies that had AI mentions. All figures and statistics below are calculated on this dataset.

4. FINDINGS

Among earnings calls that mention AI, the share that included at least one internal deployment rose from 42.9% in 2020 to 74.9% by Q1 2026 (Figure 1), following a relatively stable period from 2020 through 2022 (42.9%, 44.9%, and 45.0%, respectively) before accelerating in 2023 (57.7%) and continuing upward through 2025 (69.1%) and into 2026. The pre-2023 plateau reflects the period when internal AI disclosure was concentrated among technology sector leaders, while the acceleration beginning in 2023 corresponds to the diffusion of generative AI tools that could be deployed inside organizations within a single fiscal quarter and announced in the same earnings season. The continued rise through 2025 and 2026 indicates that adoption was still spreading across the S&P 500, with industrial, energy, and consumer sector companies entering the internal AI disclosure record in growing numbers.

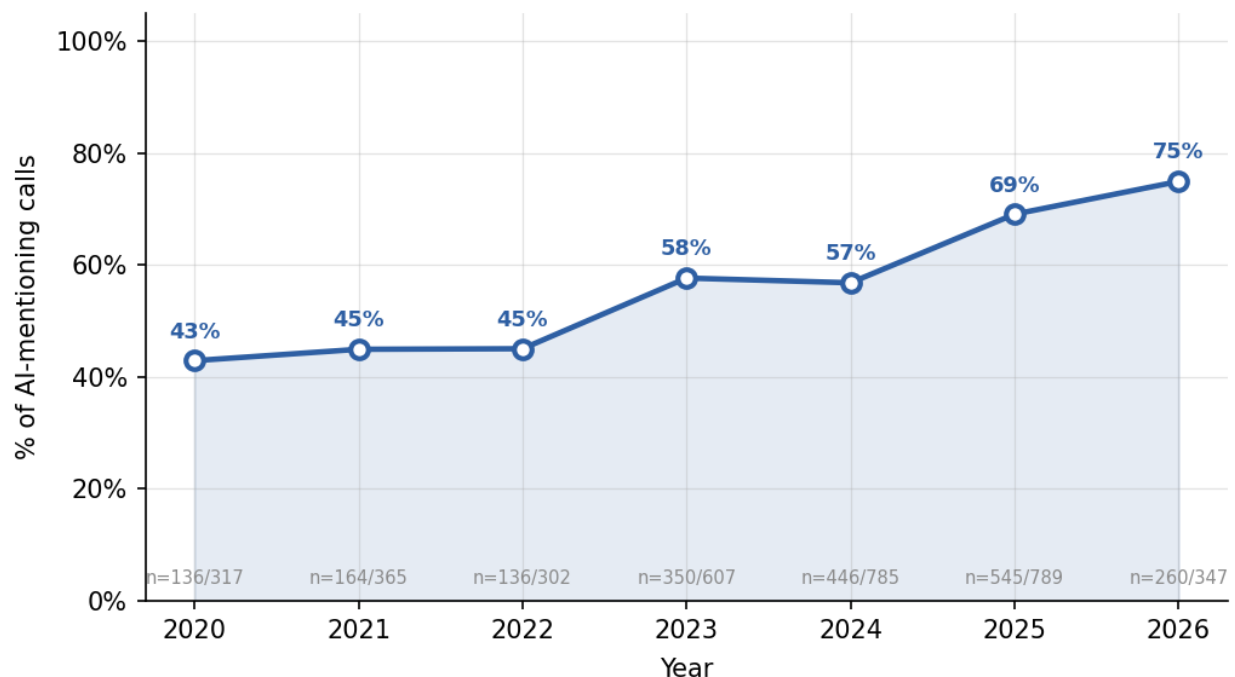


Figure 1. Share of Calls That Mention AI, by Year.

4.1 Use-Case Categories

Enterprise productivity and workflow automation leads the internal use-case distribution, accounting for 1,213 of the 3,907 use cases (31.0%) across 16 categories (Figure 2). Supply chain, operations, and manufacturing follows at 772 (19.8%), with the remaining top-five positions going to sales and marketing personalization (516, 13.2%), customer service and support automation (410, 10.5%), and code generation and software development (192, 4.9%), and the other eleven categories representing 20.6% combined. The supply chain and operations share substantially exceeds what a technology-sector-focused sample would produce, reflecting the breadth of the S&P 500 across industrial, healthcare, energy, and consumer sectors where physical and operational AI applications are concentrated.

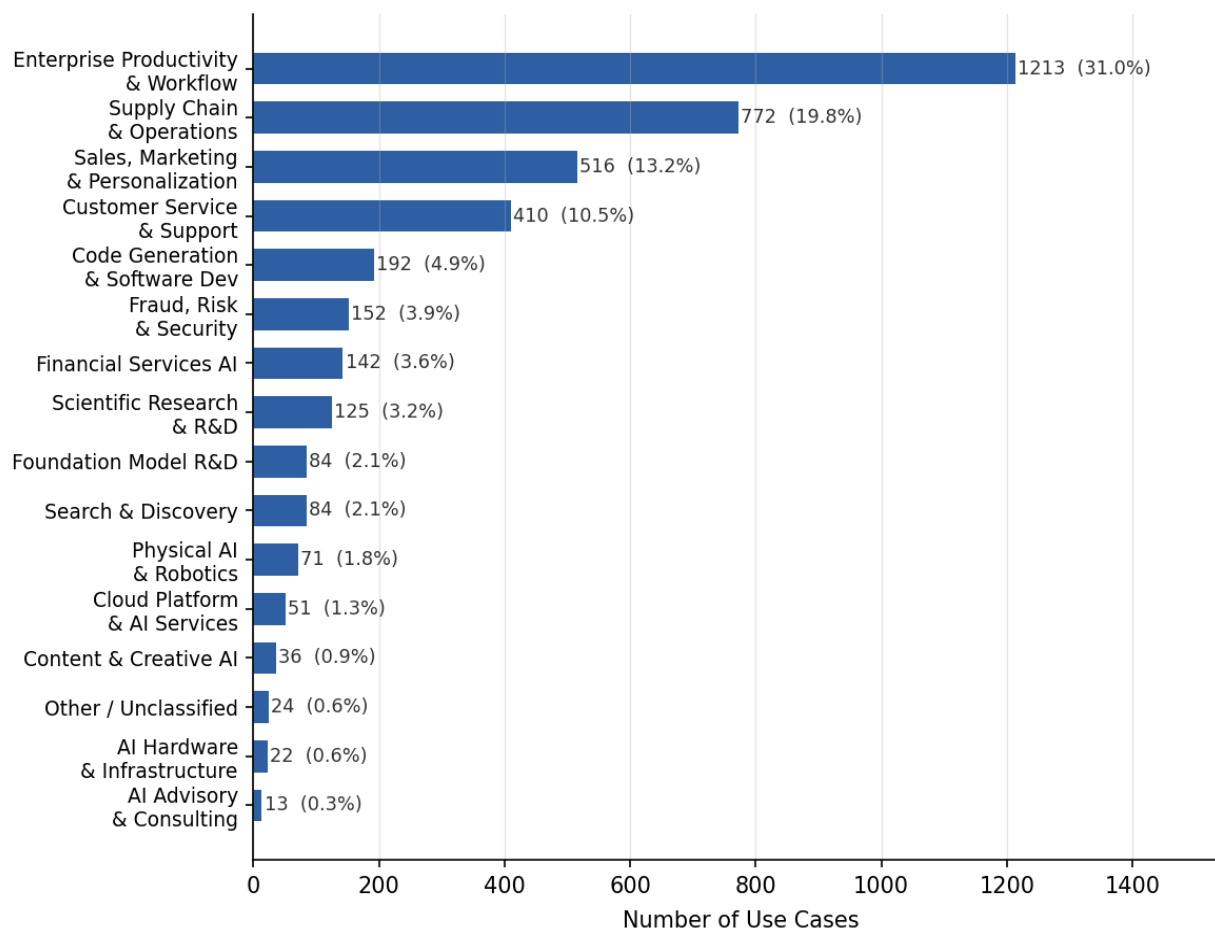


Figure 2. Distribution of AI Use-Case Categories.

Category composition varies substantially across sectors (Figure 3). Enterprise productivity and workflow automation ranks first in eight of eleven sectors, reaching its highest shares in Financials (43.9%) and Information Technology (42.6%), where knowledge-work automation across large white-collar workforces predominates, whereas supply chain, operations, and manufacturing leads in Energy (66.7%), Utilities (45.8%), Materials (41.6%), and Consumer Staples (34.8%), where AI optimizes physical processes rather than cognitive work. Health Care carries the highest scientific research and R&D share (19.3%), reflecting clinical-trial support, drug-discovery, and genomic analysis applications, while Information Technology shows the highest code generation share (12.6%). The following examples draw on verbatim executive disclosures to examine the operational specifics of the five largest categories.

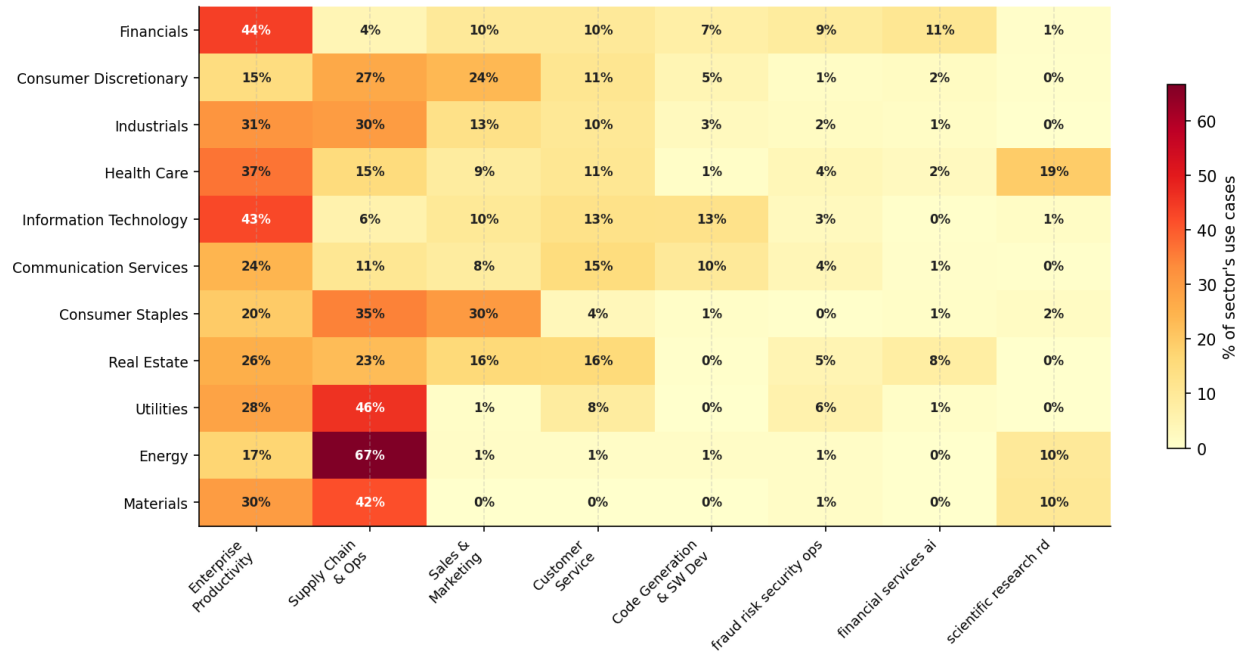


Figure 3. Use-Case Category Composition by GICS Sector.

The largest category spans AI tools that reshape how knowledge workers perform tasks across functions, ranging from document drafting and data synthesis to scheduling, compliance, and internal query handling. Bank of New York Mellon, a custodial bank and investment services firm with operations spanning securities settlement, asset management, and treasury services, described how AI consolidates delivery infrastructure across its business units:

"You might want to be able to respond to an RFP. You might want to generate summaries of documents. You might want to be able to get a head start on a research report. But actually, when you look under the heading, those three things sound different. But the AI that's actually powering them, some of the sort of mini platforms that are required are very, very similar. So we're using our AI hub to collect these different use cases and then be able to sort of deliver solutions and many AI platforms that can then be used in multiple places around the company."

The second category covers AI in physical operations including predictive maintenance, logistics optimization, energy management, quality control, and process automation across industrial, energy, healthcare, and consumer sectors. Norfolk Southern, a Class I freight railroad operating across the eastern United States, described its machine vision deployment for detecting rail car component failures before they occur:

"We are using machine vision technology to detect component failures before they occur. We're in the process of deploying fully automated inspection corridors, which will cover more than 90% of the cars moving across our network using a variety of systems to detect signs and symptoms of pending failures before they occur. Equally as important as deploying the hardware is developing the next-generation AI algorithms that detect these failures with edge computing and procedures for intervening quickly."

The third category covers AI applied to sales force productivity, lead scoring, creative testing, and campaign optimization within the reporting company's own commercial operations. Omnicom, a global advertising and marketing services holding company, described how agentic AI has changed creative concept development within its agency network:

"Historically, creative teams would typically put two to three different concepts in front of our clients for a specific campaign. The reason is because it takes time... Today, with the use of the tools that John is talking about, with the agentic capabilities that we put in place, our teams can now test 20 concepts, can test 50 concepts. More importantly is that they can test them synthetically so that we can understand what the impact and value of that work could be. We can predict that before we even spend a single dollar on media."

The fourth category captures AI in internal customer-facing operations, encompassing contact center tools, agent-assist systems, and self-service platforms managed by the reporting company for its own customers. Universal Health Services, a for-profit hospital management company operating acute care and behavioral health facilities across the United States, described an AI deployment in its post-discharge patient follow-up process:

"One of the early experiments we've done is using AI to follow up with patients on their post-discharge instructions. So, you know, an AI-generated entity will call a patient and make sure that they've followed up with the appropriate doctor appointments and they filled their prescriptions, and they're following their diet, you know, whatever the post-discharge instructions are. And we have found in the early stages that's been very well received and efficient. And then again, it frees up a clinical person, you know, generally a nurse who would otherwise be making that call."

The fifth category captures AI tools used by software engineers and development teams to accelerate code production, review, and testing. Alphabet, the parent company of Google and DeepMind, described AI-assisted coding practices across its engineering organization:

"In terms of how we are using AI for coding, we are continuing to make a lot of progress there in terms of people using coding suggestions. I think the last time I had said the number was, like, 25% of code that's checked in. It involves people accepting AI suggested solutions. That number is well over 30% now. But more importantly, we have deployed more deeper flows and particularly with the newer models, I think we are working on early agentic workflows."

4.2 Workforce Impact

Augmentation accounts for 3,570 of the 3,907 internal AI use cases (91.4%), with displacement framing in 120 (3.1%), no workforce impact in 282 (7.2%), new hiring in 22 (0.6%), and reskilling in 51 (1.3%) (Figure 4). These categories are not mutually exclusive: a single use case may combine multiple workforce-impact framings, such as displacement in one function alongside new hiring for a specialized AI team, which is why the percentages sum to slightly more than 100%. The 3.1% displacement rate, achieved against strong institutional pressure toward augmentation framing, reflects a disclosure context in which companies deploying AI in their own operations face direct analyst scrutiny of headcount implications, giving investors standing to ask operational questions and executives more pressure to characterize efficiency outcomes than earnings call AI disclosures on average would suggest.

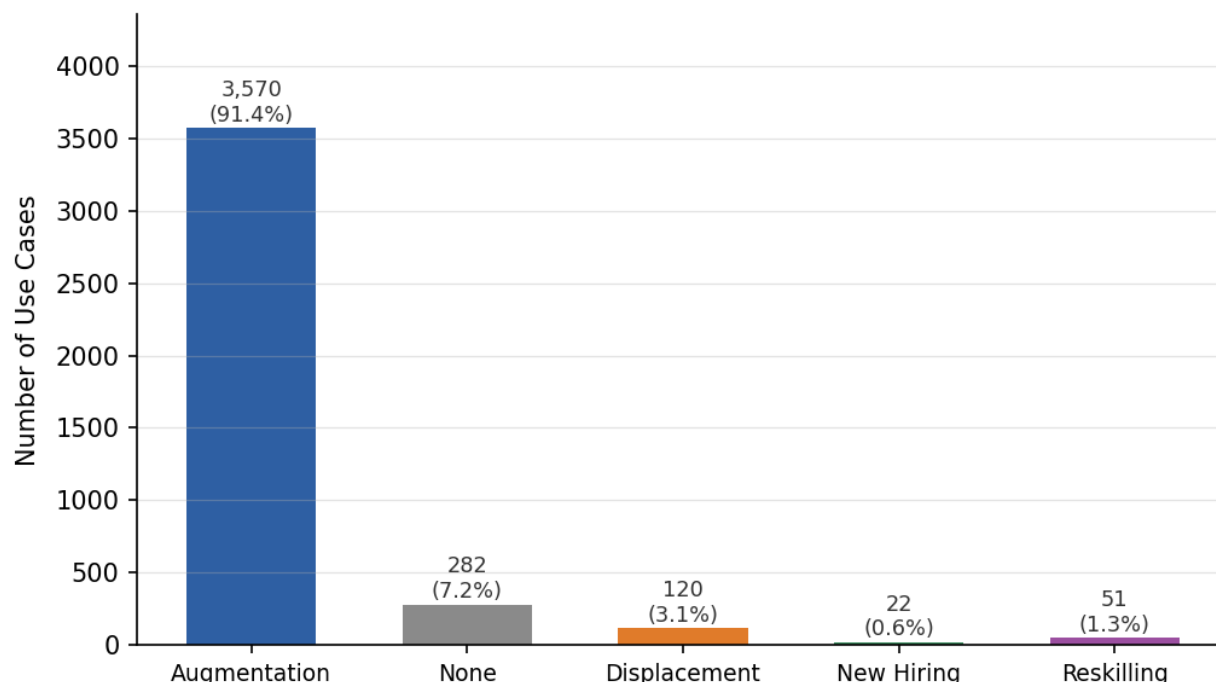


Figure 4. Workforce Impact Distribution.

Among the 3,777 use cases naming a specific workforce group (96.7% of the total), knowledge workers and general enterprise employees account for the largest share at 1,030 (27.3%) (Figure 5), because enterprise productivity and copilot tools span roles rather than target a single occupational group. Manufacturing and field operations workers follow at 656 use cases (17.4%), a share that reflects the significant presence of supply chain and operations use cases that implicate workers on plant floors, in distribution centers, and in field maintenance roles. Sales and marketing professionals account for 475 use cases (12.6%), customer service and frontline workers for 412 (10.9%), and finance and operations back-office staff for 360 (9.5%). Software engineers and developers appear in 234 use cases (6.2%), scientific and clinical research staff in 178 (4.7%), and AI and data specialists in 161 (4.3%). The remaining five categories account for the balance: investment and financial advisory (139, 3.7%), security and IT operations (65, 1.7%), creative content and media (38, 1.0%), other/unclassified workforce impact (27, 0.7%), and supply chain operations (2, <0.1%).

The AI and data specialist group carries the highest rate of new hiring of any workforce group in the dataset, as companies build dedicated AI teams while simultaneously deploying AI tools that reshape other roles, creating demand for AI specialists even as they reduce labor requirements in adjacent functions.

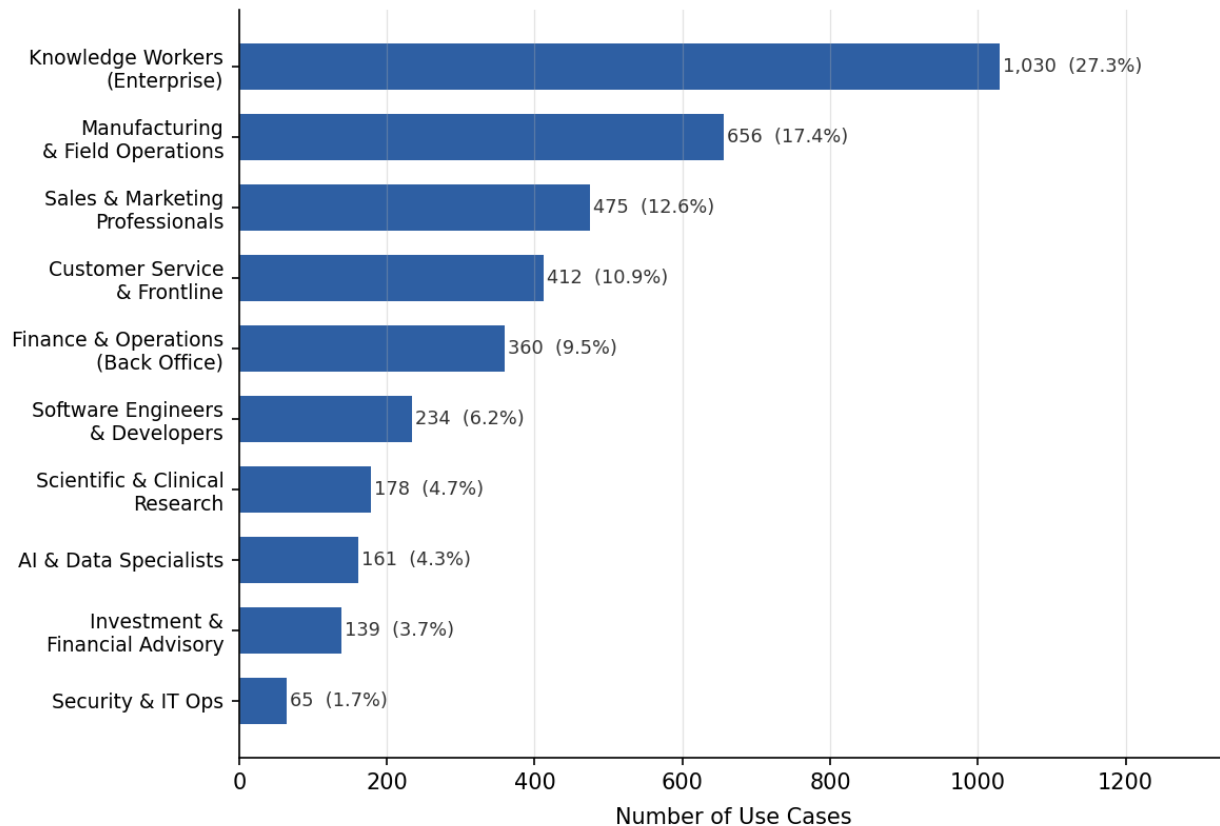


Figure 5. Distribution of Workforce Category.

The augmentation-displacement balance differs substantially across occupational groups (Figure 6). Knowledge workers show the highest absolute augmentation count among the major groups, because enterprise productivity tools are consistently framed as capability amplifiers that allow the same workforce to handle greater volume or complexity. Customer service and frontline workers carry a substantially higher displacement share, at 8.0% of use cases in that group, a result that follows from the contact-reduction objective that characterizes most customer service automation, given that when the stated goal is to reduce call volume or agent handle time, augmentation framing becomes harder to sustain.

Finance and operations back-office staff show a displacement share (4.2%) expected of a group whose routinized, rule-bound tasks including compliance processing, invoice reconciliation, and data entry present clear substitution opportunities. Knowledge workers carry a 4.3% displacement share despite their dominant augmentation count, the scale of the group meaning even a modest rate produces a substantial absolute count, while software engineers appear at 2.1%, below customer service and back-office staff but above categories with more variable task structures. Internal AI programs thus create specialized talent demand in one part of the workforce while reducing labor requirements in adjacent roles, enacting within a single organizational context the displacement-reinstatement dynamic described by Acemoglu and Restrepo (2019).

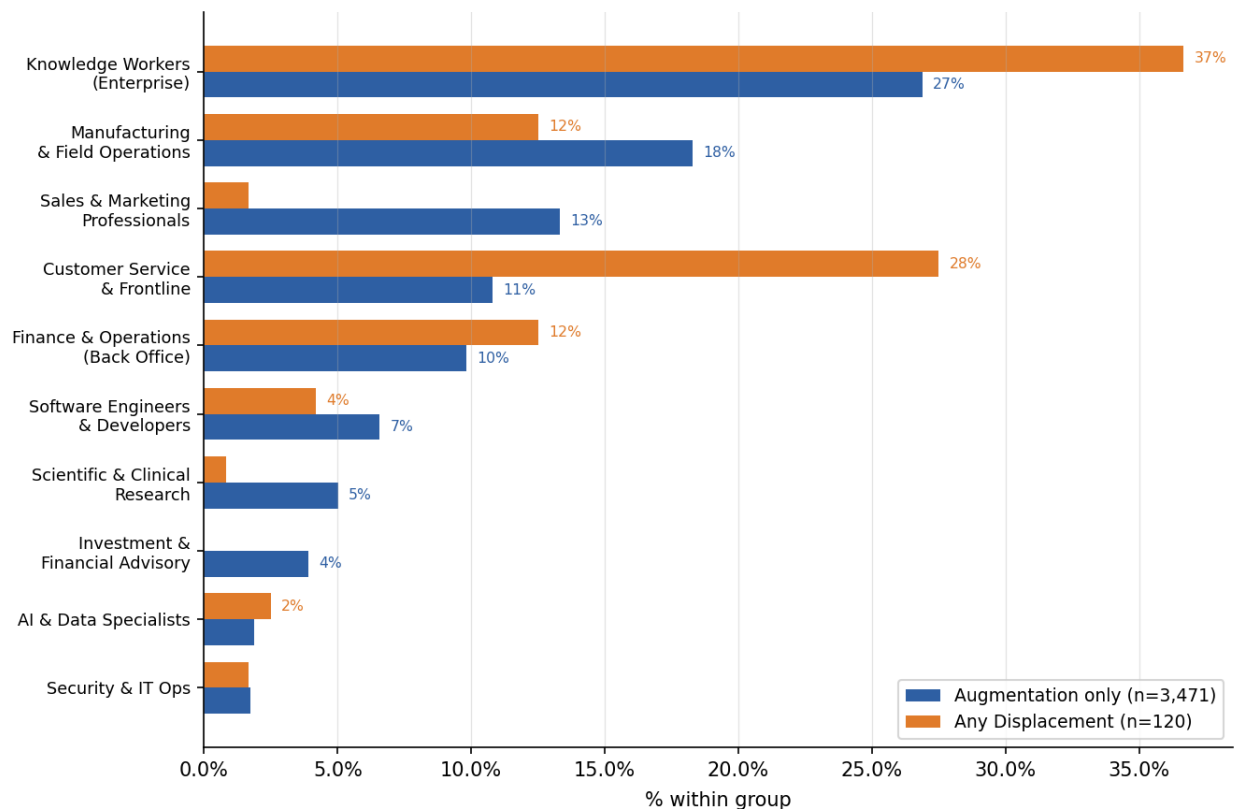


Figure 6. Workforce Category Distribution by Workforce Impact.

5. DISCUSSION

The analysis of 3,907 internal AI use cases from 2,037 earnings calls across 420 S&P 500 companies yields a picture of organizational AI adoption that goes considerably further than aggregate AI mention rates can capture, and three findings deserve particular attention.

We interpret the patterns below through two complementary theoretical lenses. Task-based automation theory (Acemoglu & Restrepo, 2019; Felten et al., 2021) predicts that displacement framing should concentrate in occupations performing routinized, rule-bound tasks, while skill-complementarity accounts of technology adoption (McElheran et al., 2024) predict that AI deployment simultaneously increases demand for specialized technical talent even as it reduces labor requirements elsewhere. These two mechanisms jointly explain why a low aggregate displacement rate coexists with concentrated displacement in specific occupational groups and simultaneous hiring growth in AI-specialist roles.

The diffusion trajectory in Figure 1, rising from 42.9% of AI-mentioning calls in 2020 to 74.9% by Q1 2026 and still climbing through the most recent periods, shows internal AI disclosure spreading beyond technology-sector leaders into industrial, energy, and consumer sector companies whose programs entered the record later. The acceleration from 2023 onward confirms that generative AI tools sharply lowered the barrier to internal deployment, with organizations converting emerging capabilities into disclosed operational programs within a single fiscal year of broad availability. This aligns with Jia et al. (2026), who document that generative AI discussions in earnings calls increased sharply and became more action-oriented after ChatGPT's release, with the strongest acceleration among larger, growth-oriented, high-tech firms. This convergence between conference-call-level and use-case-level evidence supports reading the post-2023 inflection as a real shift in operational AI activity, not just a change in disclosure style. Our broader sectoral coverage extends their finding to industrial, energy, and consumer sectors, where AI discussion was previously sparse.

The 3.1% displacement rate carries direct methodological implications for IS research using earnings call data. Studies that aggregate all AI-related mentions without isolating a company's own deployments will systematically underestimate workforce displacement in the operational record. Treating all AI mentions as equivalent mismatches the data to the research question it is meant to answer. That a 3.1% displacement rate persists despite the institutional pressure toward augmentation framing documented by Cao et al. (2023) suggests the actual displacement dynamic in organizational AI programs is larger than voluntary disclosure captures.

The occupational patterns in Figure 6 follow the task-level predictions of Acemoglu and Restrepo (2019) and Felten et al. (2021). Groups performing routinized, transactional tasks, such as customer service agents and finance back-office staff, show higher displacement rates, while groups doing knowledge-intensive, variable work are more consistently framed as augmented. The concentration of new-hiring indications among AI and data specialists (Figure 5) aligns with McElheran et al.'s (2024) finding that AI adoption is highly complementary with specialized skill acquisition. Internal AI programs appear to create demand for AI expertise even as they reshape labor requirements in adjacent roles. The manufacturing and field operations group, the second-largest workforce category at 17.4% of internal use cases, deserves attention in future research. Its prominence is driven by the supply chain and operations category, which contains specific disclosures about AI embedded in physical processes. Yet workforce displacement framing is relatively modest in this group, at just 2.3% of use cases, a pattern consistent with AI functioning as real-time decision support in settings where physical execution still requires a human presence.

For IS practice, enterprise productivity's leading share (31%) indicates that the most readily achievable near-term return on internal AI investment lies in knowledge-work automation, matching the process automation category identified by Davenport and Ronanki (2018). The supply chain and operations category (20%) tells a complementary story: physical-operations AI has matured into a routinely deployable technology across a broad range of industries, not just among industrial-sector leaders. For computing and IS curricula, two signals matter together: concentrated new-hiring around AI and data specialists, and the breadth of enterprise productivity deployments across knowledge-work functions. Near-term skill demand is therefore twofold: specialized AI and data engineering expertise, and AI literacy distributed widely across general business roles.

The sector distribution in Figure 3 confirms that internal AI adoption extends well beyond the technology sector and varies substantially by sector context. Financials leads by use-case volume, combining the highest enterprise productivity share (43.9%) with a distinctive financial services AI category, while in Energy (66.7%), Utilities (45.8%), and Materials (41.6%) supply chain and operations AI predominates in the mix. Analyses relying on aggregate category distributions will therefore understate the operational AI exposure of workers in asset-intensive sectors and overstate the prevalence of knowledge-work automation relative to what those companies actually deploy.

Taken together, these findings support three contributions. Methodologically, isolating internal AI deployments from product and customer mentions through a three-part qualification test produces a more precisely defined evidence base than aggregate AI mention measures permit. Empirically, the longitudinal use-case-level account of internal AI adoption across the S&P 500, including the post-ChatGPT inflection in 2023, extends prior conference-call-level evidence to the level of individual operational deployments. Substantively, characterizing workforce framing across occupational groups at a level of detail not previously available from earnings-call corpora reveals that voluntary AI disclosure, despite its augmentation bias, still reflects the task-level structure of automation pressure and provides a credible indication of where labor substitution is occurring.

Several limitations apply beyond those noted above. The S&P 500 corpus over-represents large publicly traded firms relative to the broader economy and cannot be taken as representative of organizational AI adoption among smaller firms or private companies. The hand-coded validation sample of 100 use cases reported in Section 3 indicates substantial agreement between LLM output and human coders, but a larger validation effort with multiple independent coders would more precisely quantify classification error rates across sectors and use-case categories. The positive framing bias inherent in investor-facing

communications creates a systematic floor on disclosed displacement rates that the extraction methodology cannot correct, meaning the 3.1% displacement figure should be treated as a lower bound on actual displacement framing rather than a complete accounting of it. Future research linking these disclosure patterns to establishment-level employment data would help test the gap between the augmentation framing that predominates in investor communications and the actual labor effects of organizational AI deployment.

6. CONCLUSION

Quarterly earnings calls from Q1 2020 through Q1 2026 document a broad and accelerating transformation in how S&P 500 companies discuss AI deployed in their own operations, rising from 43% of calls that mention AI in 2020 to 75% by Q1 2026. The main findings are a diffusion trajectory that accelerated following the emergence of generative AI tools in 2023 and that continued rising through 2026 as adoption spread beyond technology-sector leaders, a use-case distribution led by enterprise productivity automation at 31% and supply chain and operations at 20%, and a displacement framing rate of 3.1% concentrated in occupational groups whose routinized tasks present the clearest substitution opportunities.

Together these findings establish internal AI disclosure as a partially transparent record of organizational transformation, specific enough to distinguish use-case types, occupational groups, and operational outcomes while remaining selective in what it surfaces. The evidence positions quarterly earnings calls as a credible longitudinal source for tracking how AI reshapes work within S&P 500 companies, even as the augmentation framing characteristic of investor communications sets a floor on what the record reveals about labor substitution.

7. REFERENCES

- Acemoglu, D., & Restrepo, P. (2019). Automation and new tasks: How technology displaces and reinstates labor. *Journal of Economic Perspectives*, 33(2), 3–30. <https://doi.org/10.1257/jep.33.2.3>
- Babina, T., Fedyk, A., He, A., & Hodson, J. (2024). Artificial intelligence, firm growth, and product innovation. *Journal of Financial Economics*, 151, 103745. <https://doi.org/10.1016/j.jfineco.2023.103745>
- Bochkay, K., Brown, S. V., Leone, A. J., & Tucker, J. W. (2023). Textual analysis in accounting: What's next? *Contemporary Accounting Research*, 40(2), 765–805. <https://doi.org/10.1111/1911-3846.12825>
- Brynjolfsson, E., Rock, D., & Syverson, C. (2021). The productivity J-curve: How intangibles complement general purpose technologies. *American Economic Journal: Macroeconomics*, 13(1), 333–372. <https://doi.org/10.1257/mac.20180386>
- Cao, S., Jiang, W., Yang, B., & Zhang, A. L. (2023). How to talk when a machine is listening: Corporate disclosure in the age of AI. *Review of Financial Studies*, 36(9), 3603–3642. <https://doi.org/10.1093/rfs/hhad021>
- Cao, Y., Liu, M., Qiu, J., & Zhao, R. (2024). Information in disclosing emerging technologies: Evidence from AI disclosure. SSRN Working Paper 4987085. <https://ssrn.com/abstract=4987085>
- Davenport, T. H., & Ronanki, R. (2018). Artificial intelligence for the real world. *Harvard Business Review*, 96(1), 108–116.
- Eisfeldt, A. L., Schubert, G., & Zhang, M. B. (2023). Generative AI and firm values (NBER Working Paper No. 31222). National Bureau of Economic Research. <https://doi.org/10.3386/w31222>

- Felten, E. W., Raj, M., & Seamans, R. (2021). Occupational, industry, and geographic exposure to artificial intelligence: A novel dataset and its potential uses. *Strategic Management Journal*, 42(12), 2195–2217. <https://doi.org/10.1002/smj.3286>
- Hassan, T. A., Hollander, S., van Lent, L., & Tahoun, A. (2019). Firm-level political risk: Measurement and effects. *Quarterly Journal of Economics*, 134(4), 2135–2202. <https://doi.org/10.1093/qje/qjz021>
- Jia, N., Li, N., Ma, G., & Xu, D. (2026). Corporate responses to generative AI: Early evidence from conference calls. *Review of Accounting Studies*, 31(2), 649–703. <https://doi.org/10.1007/s11142-025-09916-1>
- Jöhnk, J., Weissert, M., & Wyrski, K. (2021). Ready or not, AI comes: An interview study of organizational AI readiness factors. *Business & Information Systems Engineering*, 63(1), 5–20. <https://doi.org/10.1007/s12599-020-00676-7>
- Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, 33(1), 159–174. <https://doi.org/10.2307/2529310>
- Loughran, T., & McDonald, B. (2011). When is a liability not a liability? Textual analysis, dictionaries, and 10-Ks. *Journal of Finance*, 66(1), 35–65. <https://doi.org/10.1111/j.1540-6261.2010.01625.x>
- Loughran, T., & McDonald, B. (2016). Textual analysis in accounting and finance: A survey. *Journal of Accounting Research*, 54(4), 1187–1230. <https://doi.org/10.1111/1475-679X.12123>
- McElheran, K., Li, J. F., Brynjolfsson, E., Kroff, Z., Dinlersoz, E., Foster, L., & Zolas, N. (2024). AI adoption in America: Who, what, and where. *Journal of Economics & Management Strategy*, 33(2), 375–415. <https://doi.org/10.1111/jems.12576>
- Rogers, E. M. (2003). *Diffusion of innovations* (5th ed.). Free Press.
- Sa-Couto, L., Gürpınar, T., Jia, S. J., Quadri, S., & McKenna, A. (2025). Automating curriculum intelligence: Scalable extraction and categorization of university courses using GenAI. In *Proceedings of the ISCAP Conference*.
- SaCouto, L., Gürpınar, T., Jia, S., & Quadri, S. (2026). Preparing the XR workforce: Curricular patterns of extended reality in higher education. *Journal of Information Systems Education*. Forthcoming.
- Shrestha, Y. R., Ben-Menahem, S. M., & von Krogh, G. (2019). Organizational decision-making structures in the age of artificial intelligence. *California Management Review*, 61(4), 66–83. <https://doi.org/10.1177/0008125619862257>
- Siano, F. (2025). The news in earnings announcement disclosures: Capturing word context using LLM methods. *Management Science*, 71(11), 9831–9855. <https://doi.org/10.1287/mnsc.2024.05417>
- Venkatesh, V., Morris, M. G., Davis, G. B., & Davis, F. D. (2003). User acceptance of information technology: Toward a unified view. *MIS Quarterly*, 27(3), 425–478. <https://doi.org/10.2307/30036540>