RESEARCH REPORT
STUDY PAPER 300
AUGUST 2026

Artificial Intelligence, Firm Narratives, and Productivity: Evidence from Annual Reports

Nicolai Kaarsen, Anja Rösner, Filip Milak and Anders Sørensen

THE ROCKWOOL FOUNDATION

Artificial Intelligence, Firm Narratives, and Productivity: Evidence from Annual Reports

Published by:

Address:
The ROCKWOOL Foundation
Ny Kongensgade 6
1472 Copenhagen, Denmark

Telephone +45 33 34 48 00
E-mail: kontakt@rff.dk
en.rockwoolfonden.dk

August 2026

Artificial Intelligence, Firm Narratives, and Productivity: Evidence from Annual Reports*

Nicolai Kaarsen†
ROCKWOOL Foundation

Filip Milak‡
ROCKWOOL Foundation

Anja Rösner§
ROCKWOOL Foundation

Anders Sørensen¶
ROCKWOOL Foundation and Copenhagen Business School

August 21, 2026

Abstract

How do firms frame AI as a productivity-relevant technology, and what can these narratives tell us about perceived productivity effects? We study this question using LLM-based text analysis of annual reports from the largest firms in Denmark. Annual reports are a useful source because firms are expected to report matters that are material to current and future performance. Consistent with this, we find that annual reports mainly describe AI applications already in use within the firm, rather than AI as a future possibility or general market trend. We also document a marked increase in AI narratives over the past five years, in line with survey evidence showing a sharp rise in adoption. We then examine how firms describe the theoretical channels through which AI may affect productivity. While existing empirical literature has provided especially detailed evidence on labor augmentation and automation, our results suggest that firms frame AI through a broader set of productivity channels. These two labor channels are prominent, but firms also frequently describe AI as improving product quality and expanding the scope of products and services. Other channels are less prominent but still present, including greater scalability, more effective R&D, more efficient use of non-labor inputs, and the creation of new labor tasks. The findings suggest that AI should be studied as a general-purpose technology that operates through multiple productivity channels, not only through its effects on labor.

Keywords: Artificial intelligence, Annual Reports, Narratives, Productivity

JEL classification: D83

*We are grateful to Anders Humlum for valuable and very helpful comments.

†The ROCKWOOL Foundation Research Unit, Denmark, nk@rff.dk.

‡The ROCKWOOL Foundation Research Unit, Denmark, fim@rff.dk.

§The ROCKWOOL Foundation Research Unit, Denmark, anr@rff.dk.

¶The ROCKWOOL Foundation Research Unit, and Copenhagen Business School (CBS), Department of Economics, as@rff.dk.

1 Introduction

Artificial intelligence (AI) has moved from niche technology to a central component of firms' digital transformation strategies. Advances in machine learning, natural language processing, and data-intensive computing have raised expectations about substantial gains in productivity. If AI increases R&D effectiveness, it could even lead to a permanent increase in productivity growth. This is a particularly salient expectation in advanced economies, where productivity growth has remained weak despite continued investment in digital technologies (e.g., Brynjolfsson et al. 2017; Filippucci et al. 2024b; Filippucci et al. 2024a; OECD 2019). A more troubling possibility is that AI adversely affects labor-market outcomes by replacing tasks previously performed by workers. The economic and social effects of a new technology depend crucially on the productivity channels it operates through, i.e., does it automate tasks, augment labor, increase the number of new tasks, or expand the quality or type of products produced (e.g., Acemoglu and Restrepo 2020).

Empirical evidence on the relative importance of these productivity channels remains limited and inconclusive. Recent experimental work suggests a labor-augmenting channel: Generative AI raises worker productivity on knowledge tasks, but the effects are task-specific and uneven (e.g., Noy and Zhang 2023; Dell'Acqua et al. 2024; Brynjolfsson et al. 2023). Such positive effects are not confirmed by non-experimental labor-market evidence: workers and firms report productivity gains from AI use, but early effects on hours and earnings appear limited (e.g., Humlum and Vestergaard 2025). Other studies investigate the effects of automation, with evidence that AI exposure is associated with reduced hiring in non-AI jobs and declining employment among early-career workers in occupations where AI is more likely to substitute for labor (e.g., Acemoglu and Restrepo 2020; Brynjolfsson et al. 2023). Firm-level evidence also points to broader effects beyond labor. Recent evidence from Germany finds positive associations between AI adoption and productivity (e.g., Czarnitzki et al. 2023). European firm-level evidence similarly suggests positive productivity effects of AI adoption, while emphasizing that these effects depend on firm-level exposure, capabilities, and implementation context (e.g., Aldasoro et al. 2026). Additionally, firms that invest more heavily in AI subsequently experience higher growth and product innovation (e.g., Babina et al. 2024). However, systematic evidence on how firms use AI remains limited (e.g., Allen 2026; OECD 2024). Existing evidence often comes from case studies, surveys, and technology-investment proxies (e.g., Allen 2026; Calvino and Fontanelli 2024; OECD 2024), each capturing only selected dimensions of adoption.

This study contributes to the debate by analyzing AI-related statements in the annual reports of Denmark's largest firms. Prompting an LLM to perform text analysis, we investigate how AI is discussed, whether it is described as currently used or merely anticipated, which applications are emphasized, and what the implied productivity channels of AI in use are. This approach treats annual reports as a signal of firms' expectations about how AI affects productivity. It therefore complements adoption-based measures by capturing not only whether AI appears in firm disclosures, but also how firms explain its role in value creation. Existing firm-level studies provide important evidence on AI adoption, productivity, innovation, and growth, but provide less systematic evidence on the relative prominence of the different mechanisms through which firms themselves expect AI to affect productivity. Additionally, survey-based evidence tends to focus on adoption rates, but to our knowledge, it typically pays less attention to the implied productivity channels (e.g., Allen 2026; OECD 2024; Calvino and Fontanelli 2024).

Additionally, the empirical strategy builds on three complementary literatures. First, the text-as-data methodology provides principles and validation practices for converting unstructured narrative into reproducible measures (e.g., Baker et al. 2016; Gentzkow et al. 2019; Grimmer and Stewart 2013; Grimmer et al. 2022). Second, the accounting and finance literature shows how corporate disclosures, annual reports, earnings releases, and prospectuses, can be used as sources of firm-level signals, emphasizing issues of dictionary construction, readability, tone, and economic interpretation (e.g., Davis et al. 2007; Huang et al. 2014; Li 2008; Li 2010; Loughran and McDonald 2011; Loughran and McDonald 2016). Third, the disclosure and capital-markets literature provides the economic motivation: narrative communication matters because it shapes information asymmetries, liquidity, and the cost of capital (e.g., Botosan 1997; Diamond and Verrecchia 1991; Healy and Palepu 2001; Verrecchia 2001). Thus, we use chunk-level validation, taxonomy-based classification, and conservative semantic validation in this paper.

Our first three results confirm the usefulness of annual reports to study AI usage in firms. First, AI-related references in annual reports were rare until the mid-2010s, but have increased sharply in recent years. This is consistent with the broader diffusion of machine learning and the rise of public attention around generative AI. Second, explicit statements of non-use are persistent and informative. Some firms do not simply omit AI, but they state that AI is not used or is not relevant for specific use cases. These statements often appear over time, suggesting a deliberate and stable positioning rather than a temporary absence of disclosure. Third, most AI narratives describe the firms' present-day usage of AI. Only to a lesser degree do narratives speak of plans of future use or discussions of how AI affects the market and other firms. This suggests that annual reports contain useful information about AI use.

We then investigate the productivity channels through which firms report that AI affects their activities. The results reveal a broad but uneven set of mechanisms. Labor-task augmentation is the most widespread channel, suggesting that firms most commonly frame AI as supporting workers in performing existing tasks more effectively. Automation is also prominent, but it does not dominate the narratives. Firms frequently associate AI with improvements in the quality of existing products and services and with the introduction of new products or services. Scalability and greater R&D effectiveness also appear regularly, whereas improvements in non-labor input efficiency and the creation of new labor tasks are discussed less frequently. Thus, the findings indicate that firms do not primarily present AI as a labor-replacing technology. Instead, they describe it as affecting both the input side of production, through augmentation, automation, and input efficiency, as well as the output side, through product improvement, innovation, and expansion of the product set.

These patterns matter because the economic effects of AI are likely to depend on the channels through which it operates. The productivity effects of general-purpose technologies may also materialize only gradually and depend on complementary investments in skills, data infrastructure, and organizational redesign (e.g., Bresnahan et al. 2002; Aghion et al. 2019; Syverson 2017). We therefore use firm narratives as early evidence on where firms perceive AI-related productivity gains to arise, while remaining explicit that annual reports capture expected or publicly articulated mechanisms rather than realized productivity effects. Recent adoption tracking also suggests that uptake differs substantially by firm size and sector, with larger firms and firms in analytical service industries tending to adopt earlier (e.g., Allen 2026; OECD 2024).

The remainder of this paper is structured as follows. Section 2 provides an overview of the institutional background and context. Section 3 presents the data and explains the empirical approach. Section 4 reports the baseline results, and Section 5 discusses productivity-related findings. Section 6 describes robustness tests before Section 7 concludes.

2 Institutional and theoretical background

2.1 Why are annual reports useful to study firm adoption of AI?

The empirical analysis is based on annual reports from large firms operating in Denmark. Annual reports are useful for studying AI because they combine a regulated reporting format with substantial managerial discretion. On the one hand, Danish firms are required to publish annual reports under the Danish Financial Statements Act (*Årsregnskabsloven*), and listed firms typically prepare consolidated financial statements under International Financial Reporting Standards (IFRS), while large non-listed firms follow Danish GAAP with substantial overlap in disclosure requirements (e.g., Christensen et al. 2013). These rules make annual reports relatively comparable across firms and over time. On the other hand, the management review and related narrative section allow firms to discuss business models, strategy, risks, expected future developments, and long-term value creation in their own terms. Firms are encouraged, but not required, to discuss technological developments insofar as these are deemed relevant for understanding performance, risk, or future prospects (European Commission 2017). This combination of comparability and discretion makes annual reports particularly suitable for analyzing how firms publicly frame emerging technologies. This interpretation is consistent with the disclosure literature, which emphasizes that corporate reporting reduces information asymmetries while leaving room for managerial discretion in what firms choose to emphasize (e.g., Healy and Palepu 2001; Verrecchia 2001).

However, there is no explicit requirement to disclose the use of AI as a distinct category. References to AI therefore usually appear within broader discussions of digitization, automation, data analytics, innovation, IT systems, risk management, or strategic transformation. AI-related disclosure consequently reflects managerial judgments about materiality and relevance rather than compliance with narrowly defined reporting standards. This is central to the interpretation of the evidence. The analysis does not treat annual reports as a complete inventory of firms' AI use. Instead, it treats them as strategic communication documents that reveal which aspects of AI firms choose to make visible, how they describe them, and whether they connect them to productivity-relevant activities. Recent evidence on AI narratives in corporate filings suggests that markets distinguish substantive, actionable AI mentions from vague or speculative ones, which makes this disclosure margin economically meaningful (e.g., Basnet et al. 2025).

Corporate disclosures are shaped by both informal and strategic incentives. Firms face pressure to communicate credibly with investors, lenders, regulators, and other stakeholders. Transparent disclosure can reduce information asymmetries, improve liquidity, lower the cost of capital, and strengthen credibility in financial markets (e.g., Botosan 1997; Diamond and Verrecchia 1991; Healy and Palepu 2001; Verrecchia 2001). At the same time, narrative disclosures are also used to manage impressions. Firms may emphasize favorable developments, frame uncertainty, or downplay risks and underperformance (e.g., Merkl-Davies and Brennan 2007; Li 2010). Emerging technologies such as AI are especially exposed to this dual logic. They are widely associated with innovation and future competitiveness, but their concrete implementation, scale, and payoffs are often difficult for outsiders to verify. Recent evidence on AI narratives in corporate filings shows that markets react differently to actionable and speculative disclosures, reinforcing the interpretation of AI language as a strategic signal rather than boilerplate (e.g., Basnet et al. 2025).

This selectivity is not only a limitation. It is also the identifying logic of the narrative approach. Because firms decide whether to mention AI, which applications to emphasize, and how to describe expected benefits or risks, AI disclosure provides information about expectations, framing, and perceived relevance. A firm that presents AI as part of everyday operations communicates something different from a firm that presents AI as an experimental future opportunity, and both differ from a firm that remains silent. Prior research in accounting and finance shows that the language used in corporate disclosure can contain information beyond quantitative indicators, including forward-looking information about managerial beliefs, strategic intentions, and uncertainty (e.g., Li 2010; Loughran and McDonald 2016). In the context of AI, annual reports may therefore capture early-stage expectations, organizational learning, or strategic positioning before such activities are visible in standard performance measures.

Explicit non-use is particularly important in this setting. Some firms do not merely omit language, but they state that AI is not used, not relevant, or not applied in specific parts of their operations. This distinction matters because explicit non-use is not equivalent to missing disclosure. It may reflect a conscious boundary around implementation, governance, compliance, customer interaction, automated decision-making, or profiling. It may also indicate that management has considered AI and concluded that it is immaterial, inappropriate, or inconsistent with the firm's current operating model. In this sense, explicit non-use can be interpreted as a substantive narrative position. It helps distinguish firms for which AI is absent from the report because it is unnoticed from firms for which non-use is itself part of the public account of technology, risk, or governance.

Annual reports, therefore, provide informative but selective evidence. They do not reveal the full technical architecture of AI systems, the intensity of adoption, or realized productivity effects. Disclosure may lag behind actual implementation, overstate the significance of pilot projects, or use generic language shaped by legal caution and reputational concerns. The evidence should therefore be interpreted through three recurring tensions. First, AI disclosure captures expectations rather than realized productivity outcomes. Second, it may describe short-run operational applications even when the broader economic interest lies in long-run transformation. Third, AI may be framed either as a substitute for labor and routine tasks or as a complement to human judgment, skills, and organizational capabilities. These tensions motivate the conceptual framework, which links AI narratives to economically meaningful productivity channels while remaining agnostic about realized effects.

2.2 Theory: Productivity channels

We build on the view that AI is a general-purpose technology that can affect productivity through multiple channels (e.g., Aghion et al. 2019; Bresnahan et al. 2002; Brynjolfsson et al. 2017).

We use a simple product-level production framework to organize the channels through which AI may affect firm productivity. Let the firm produce a set of products or services $p \in \mathscr{P}$. Each product has quantity y_p and quality q_p. Product p is produced according to

$$y_p = A_p \left[F_p \left(E_p K_p, \{x_{pj}\}_{j \in \mathscr{T}_p} \right) \right]^{\sigma_p},$$

where A_p is product-specific productivity, K_p is non-labor input, and E_p captures the efficiency of non-labor inputs. The parameter σ_p captures scalability: a higher value implies that output can expand with a less than proportional increase in inputs. The set $\mathscr{T}_p$ contains the tasks used to produce product p, and x_{pj} denotes task output for task j. We write task output as

$$x_{pj} = Z_{pj} + B_{pj} L_{pj},$$

where L_{pj} is labor used in task j for product p. The term Z_{pj} captures task automation, while B_{pj} captures labour augmentation.

This framework organizes eight non-mutually-exclusive productivity channels into three groups discussed in the following: labor, products, and other productivity mechanisms. A compact summary of the channels, their definitions, and interpretations is provided in Appendix C, Table C1.

Labor channels. The first group concerns how AI changes labor tasks by augmenting labor, automating tasks, or creating new tasks. In the framework, these channels correspond to changes in B_{pj}, Z_{pj}, and $\mathcal{T}_p$. This follows the task-based view of technological change, where new technologies affect productivity and labor demand by changing the allocation of tasks between workers and machines (e.g., Autor et al. 2003; Acemoglu and Restrepo 2020).

Labor augmentation occurs when AI helps workers perform existing tasks more effectively. Examples include forecasting, pricing, risk assessment, customer targeting, fraud detection, and decision support. In the framework, this corresponds to an increase in B_{pj}. Existing evidence suggests that such gains can be substantial, but uneven across tasks and workers (e.g., Noy and Zhang 2023; Dell'Acqua et al. 2024; Brynjolfsson et al. 2023).

Automation occurs when AI performs tasks previously carried out by workers, such as document processing, customer service, quality control, or fraud detection. In the framework, this corresponds to an increase in Z_{pj}. Automation may raise output per worker, but can also reduce labor demand in the tasks being automated (e.g., Autor et al. 2003; Acemoglu and Restrepo 2020).

New labor tasks arise when AI creates new activities for workers, such as AI oversight, model evaluation, data curation, prompt design, AI governance, or workflow integration. In the framework, this corresponds to an expansion of $\mathcal{T}_p$.

Product channels. The second group concerns how AI changes the products and services firms offer by improving existing product quality or enabling new products and services. In the framework, these channels correspond to changes in q_p and $\mathcal{P}$. They capture cases where AI raises productivity by changing the value or range of output, rather than only by changing how existing output is produced.

Existing product quality improves when AI makes current products or services more valuable to customers. Examples include personalization, better diagnostics, improved reliability, faster service, higher precision, or more consistent product quality. In the framework, this corresponds to an increase in q_p.

New products or services arise when AI enables offerings that the firm did not previously provide, such as AI-enabled digital services, data-driven products, automated advisory tools, or new customer-facing solutions. In the framework, this corresponds to an expansion of the product set $\mathscr{P}$.

Other productivity channels. The third group covers productivity channels that are not primarily about labor tasks or product characteristics. These channels concern scalability, non-labor input efficiency, and R&D effectiveness. In the framework, they correspond to changes in σ_p, E_p, and the rate at which firms improve A_p, q_p, or $\mathscr{P}$.

Scalability improves when AI makes it easier to expand output without a proportional increase in inputs. This may occur if AI lowers communication, coordination, monitoring, or knowledge-transfer costs. In the framework, this corresponds to an increase in σ_p. The channel is related to the literature on digital technologies, organizational change, and scale economies (e.g., Bresnahan et al. 2002; Brynjolfsson et al. 2017).

Non-labor input efficiency improves when AI makes capital, energy, materials, inventories, logistics, or purchased services more productive. Examples include predictive maintenance, energy optimization, materials savings, inventory optimization, or improved logistics planning. In the framework, this corresponds to an increase in E_p.

R&D effectiveness improves when AI makes research, experimentation, design, discovery, or product development faster or cheaper. This channel is dynamic. If AI permanently raises the rate at which firms generate useful ideas, it may affect productivity growth rather than only the productivity level (e.g., Filippucci et al. 2024a; Trammell and Korinek 2023). In principle, higher R&D effectiveness could accelerate improvements in any technology-relevant object in the framework, including A_p, q_p, $\mathscr{P}$, B_{pj}, Z_{pj}, E_p, or σ_p.

3 Data and text-based measurement

3.1 Sample and scope

Annual reports are obtained from the Danish Business Authority's digital reporting platform, *virk.dk*, which serves as the central repository for statutory corporate filings in Denmark. The use of this source ensures broad coverage and consistent reporting formats while relying on publicly available documents prepared for investors, regulators, and other stakeholders. Firms in the sample are drawn from a broad range of sectors, including manufacturing, trade and transport, finance, and other large private-sector activities, allowing for comparisons of AI narratives across industries.

The starting point for firm identification is the CVR number assigned by the Danish Business Authority. The CVR is a unique eight-digit identifier assigned to all registered businesses, associations, and authorities in Denmark, regardless of company form. This identifier is useful because it provides a consistent administrative link between annual reports and registered legal entities. However, CVR numbers do not necessarily correspond one-to-one to economically meaningful firms. A larger business group may contain multiple CVR-registered entities, while subsidiaries, branches, and related legal units may share ownership or business activities without being identical firms. To account for this distinction, we link the CVR-based reporting data to firm IDs using the database from Kaarsen and Sørensen (2026). The database groups CVR numbers into a single firm when they share the same nearest parent company that publishes consolidated annual accounts. The rationale is that, when a lower-level parent company could rely on consolidated reporting at a higher level but nevertheless prepares consolidated accounts for its own sub-group, this is an informative reporting choice. We interpret such consolidation as evidence that the parent presents the subsidiaries as part of the same economic unit, and therefore treats them as belonging to the same firm for reporting purposes. The resulting identifier, therefore, provides a meaningful firm-level unit of analysis while retaining the CVR number as the administrative foundation for matching annual reports with registry information.

The unit of analysis is the firm-year annual report. Each report is treated as a strategic communication artifact that reflects the firm's self-presentation, priorities, and expectations at a given point in time. The initial report universe is reduced through successive filtering steps. Starting with 1,248 firms in the broad sample, annual reports are downloaded and, where possible, converted into machine-readable text where possible. The keyword procedure then identifies 423 firms with reports that include AI-related content. Furthermore, sorting out false-positive AI keywords leaves us with 326 true AI keyword matches. Of those, 252 firms report AI usage, while 74 explicitly state non-AI usage. Figure A2 in Appendix A illustrates this sequential filtering process.

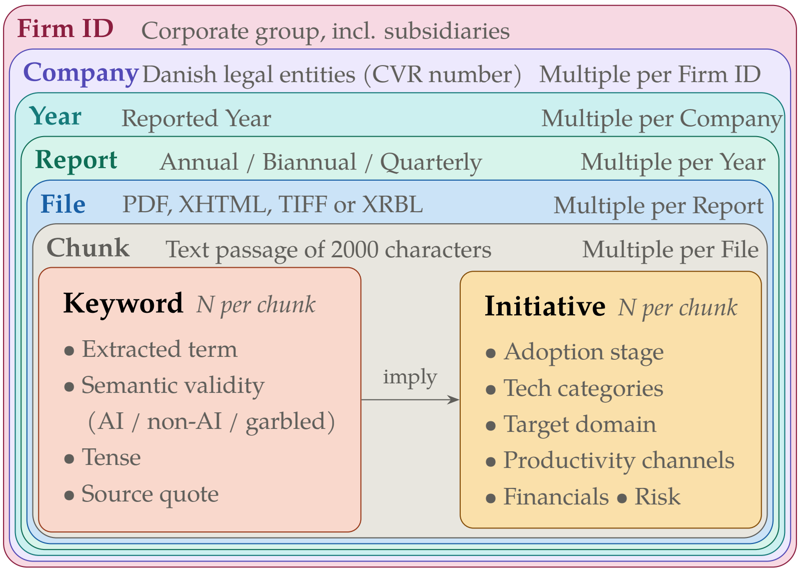
3.2 Text extraction and pre-processing

Annual reports filed on *virk.dk* are published in heterogeneous formats, including text-based PDFs, scanned PDFs, XBRL-related files, and mixed documents containing tables, images, signatures, and appendices.[1] To enable systematic comparison across firms and years, these documents must first be converted into a standardized machine-readable corpus. The purpose of the extraction pipeline is therefore twofold: to make heterogeneous reports comparable for text analysis, and to preserve enough of the original narrative structure to interpret AI-related statements in context.

Figure 1 summarizes the resulting data structure. The extraction procedure links each text passage to a hierarchy of identifiers: company, reporting year, report, source file, and text chunk. This structure allows the analysis to preserve the connection between AI-related statements and the firm-year report in which they appear, while also enabling classification at the more granular chunk level. The chunk is the main unit for keyword identification, semantic validation, and subsequent coding of AI-related initiatives.

[1] A full picture of different formats and their shares among the reports can be found in Appendix C, Figure C1.

Figure 1: From annual reports to validated AI-related text chunks



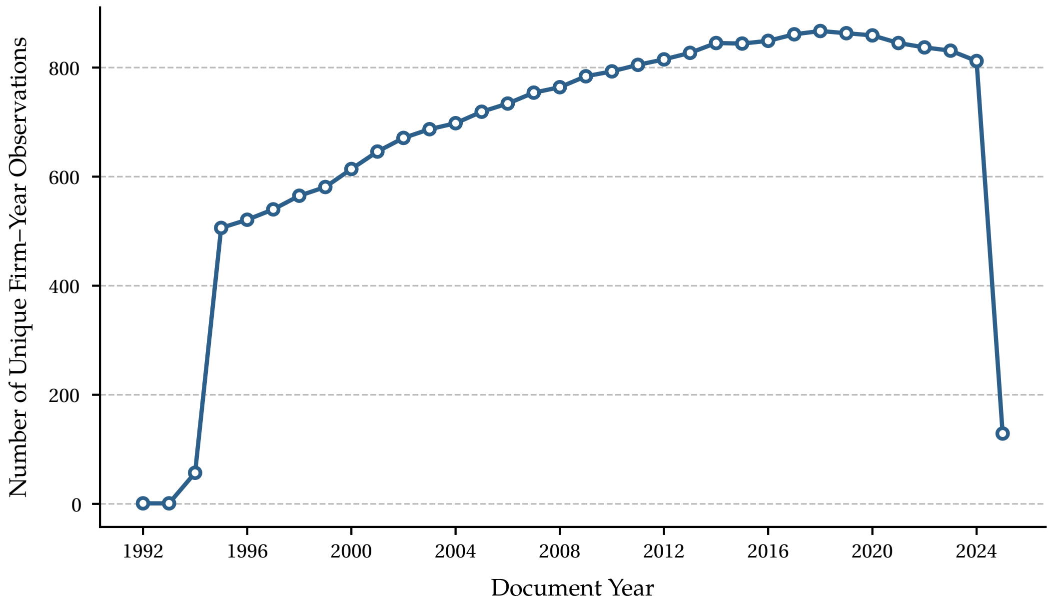
Note: This figure illustrates the hierarchical structure of the text-based dataset. Each firm identifier (Firm ID) represents a corporate group, which may include one or more subsidiaries, each identified by a distinct company identifier (CVR number). Within a given firm-subsidiary-year combination, annual reports are linked through the CVR number and reporting year. Each firm-year may contain one or more reports, and each report may consist of one or more files in different formats, such as PDF, XHTML, TIFF, or XBRL-related files. After text extraction and preprocessing, files are segmented into text chunks of approximately 2,000 characters. AI-related chunks are then identified through keyword matching and semantic validation. Validated chunks form the basis for the subsequent classification of AI initiatives, including adoption stage, technology category, target domain, value driver, and financial or risk-related content.

A more detailed overview of the processing workflow, including keyword detection, semantic validation, subject classification, and final annotation, is provided in Appendix A, Figure A1.

Reports are first downloaded in their original formats. Text-based PDFs are processed using standard PDF text extraction tools, while scanned documents and image-based pages, including TIFFs or embedded images, are converted using optical character recognition (OCR). This step makes textual content embedded in non-readable formats accessible for downstream analysis (Smith 2007). The extracted text is then cleaned and standardized. This includes removing headers and footers, correcting common OCR errors, normalizing character encoding, and excluding purely graphical elements that do not contribute to the narrative content.

The aggregated reports (full sample) over time are shown in Figure 2. The coverage is sparse before the mid-1990s, rises steadily, and is broadly stable from the mid-2010s onward, though 2025 remains incomplete.

Figure 2: Reports over time



Note: This figure shows the aggregate reports over time. Reports for 2025 are not fully available as of the time of data collection (May 2026).

3.3 Keyword filtering and semantic validation

The text corpus resulting from processing the annual reports is structured at the chunk level. Preserving chunk-level structure is important because AI-related language often appears in specific parts of annual reports, such as management reviews, strategy sections, risk discussions, sustainability statements, or governance sections. The preprocessing procedure is therefore designed not only to reduce noise, but also to ensure that AI-related statements can be evaluated within the surrounding narrative rather than as isolated keyword occurrences.

The identification of AI-related content follows a context-aware text-processing strategy. The whole process is visualized in Appendix A, Figure A1. Rather than scanning annual reports as undifferentiated text, each report is segmented based on keyword occurrences. The keyword represents the epicenter of chunk construction, where each chunk is processed as a coherent narrative unit. This approach is computationally more demanding than document-level keyword counts, but it improves interpretability by preserving the context in which AI is mentioned. For example, the same keyword may refer to ongoing implementation, planned investment, operational use, regulatory risk, deliberate non-use, or an unrelated false positive.

AI-related candidate chunks are first identified using keyword-based filtering. The keyword list includes terms such as "artificial intelligence", "machine learning", and "generative AI".[2] The keyword step is designed to maximize recall: it identifies text passages that may contain relevant AI-related content, but it does not itself define the final measure. Raw keyword hits are therefore treated only as candidate observations. Once a keyword is found, the text around that keyword is parsed. In the final implementation, candidate chunks are constructed using a character-based window of up to 2,000 characters around the relevant AI keyword or keyword cluster. When multiple AI keywords occur close to one another, the window is centered around the midpoint of the relevant keyword cluster rather than around a single keyword. Thus, multiple nearby keyword occurrences are grouped into a single candidate chunk, so that the number of candidate chunks can be smaller than the number of raw keyword occurrences. This procedure preserves the surrounding narrative context while keeping chunks sufficiently compact for the subsequent semantic validation step.[3]

Candidate chunks are subsequently screened using semantic validation. This validation step distinguishes substantive AI-technology references from false positives, ambiguous mentions, and cases where the keyword appears in another context. In particular, the validation procedure separates valid AI-related technology mentions from references to professional titles, organizational names, abbreviations, or other non-substantive uses of the same terms. It also identifies cases where firms explicitly state that AI is not used or not relevant, which are retained as meaningful observations rather than discarded as non-disclosed. The final measure is therefore based on validated narrative content, not on mechanical keyword frequency. This approach follows the broader text-as-data literature, which emphasized the importance of contextual interpretation when analyzing corporate disclosure (e.g., Gentzkow et al. 2019; Loughran and McDonald 2016).

3.4 Taxonomy and classification

Validated AI-related chunks form the basis for the classification of AI initiatives. We define an AI initiative as a firm-level statement in which the company has direct agency over an AI-related activity, decision, capability, or restriction. An initiative therefore requires more than

[2] The full list of keywords can be found in Appendix C, Table C2.

[3] Candidate chunks have a mean length of 1,990 characters, with a median of 2,000 characters. The minimum observed chunk is 975 characters, and the maximum is 2,000 characters. These statistics reflect the imposed upper bound of the chunking procedure. Chunks may contain more than one AI keyword if several keywords appear within the same local context.

the presence of an AI keyword. It must describe something the firm itself does, develops, tests, plans, expands, governs, or explicitly chooses not to do with respect to AI.

This distinction is important because annual reports often mention AI in different ways. Some passages describe the firm's own use of AI, for example in operations, customer interaction, product development, risk management, or internal processes. Other passages mention AI only as part of broader market trends, customer demand, regulatory developments, competitor activity, board expertise, or glossary definitions. The latter are informative about the context in which firms operate, but they are not treated as AI initiatives because they do not describe firm-level agency. By contrast, explicit statements that a firm does not use AI in a specific area are retained as initiatives, since they represent an active disclosure of a boundary around AI use rather than a simple absence of disclosure.

The classification therefore proceeds in two steps. First, validated chunks are evaluated to determine whether the company is the relevant subject of the AI-related statement. If the passage describes direct firm-level agency, it is classified as an AI initiative. If the passage only refers to external actors, general technological developments, or passive context, it is classified as a contextual mention and is not used as an initiative in the main taxonomy. Second, each extracted initiative is annotated with additional attributes that describe how AI is positioned in the annual report. These attributes include the adoption stage, the relevant technology category, the target domain within the firm, the stated productivity relation, and whether the passage contains financial or risk-related content.

The initiative is therefore the main analytical unit for the taxonomy-based part of the analysis. It differs from a keyword occurrence because it is based on semantic content rather than word frequency, and it differs from a chunk because a chunk is only the surrounding text window from which initiatives are extracted. A single validated chunk may contain no initiative, one initiative, or more than one initiative, depending on whether the text describes one or several distinct firm-level AI-related activities or restrictions. This structure allows the analysis to preserve the contextual richness of annual-report language while converting validated narrative content into comparable firm-level measures.[4]

[4]The LLM-based classification was implemented using a structured prompt that specified semantic validation rules, adoption-stage definitions, contextual-mention categories, non-use logic, evidence requirements, and field-level extraction rules. The full prompt is explained and shown in Appendix B. We assessed the robustness of the LLM-based semantic validation and classification procedure across several model specifications. In an initial implementation, candidate chunks were processed using Gemini 3 Flash Preview. We subsequently replicated the

Where possible, AI initiatives are also coded according to the productivity-relevant mechanism described in the underlying text. This coding is used in the later productivity analysis to connect firms' AI narratives to the conceptual framework introduced in Section 2. Importantly, this classification is separate from the adoption-stage taxonomy. The adoption-stage taxonomy describes how AI is positioned in the report, for example as active use, planning, capability building, or explicit non-use. The productivity-channel classification instead describes the mechanism through which the firm suggests AI may affect productivity (labor, product, or other productivity channels). The detailed definition of these channels was developed in Section 2.2, while they are interpreted in Section 5.3.

Descriptive statistics for the annual reports are presented in Table 1. The empirical analysis focuses on large firms operating in Denmark. The broader report universe includes 1,248 companies when public-sector entities are included. This broader universe accounts for 53.1 percent of Danish employment, corresponding to 1,358,673 employees out of 2,558,981. Public-sector entities are included at this stage only to illustrate the broad economic relevance of the report universe and the fact that large reporting entities account for a substantial share of employment in Denmark. However, the baseline analysis in this paper focuses on the private sector. This is mainly because private firms have an incentive to signal through their annual reports. Public firms are generally required to publish annual reports, but do not necessarily communicate their strategies through them to investors or other stakeholders. Excluding public-sector entities, the sample comprises 923 companies and accounts for 38.5 percent of private-sector employment, representing 617,423 employees out of 1,603,343.

validation and classification step using Claude Sonnet 4.6. Comparing model outputs against manual checks and internal consistency criteria indicated that the Claude-based specifications produced fewer false-positive classifications, in particular for ambiguous abbreviations, contextual mentions, and non-substantive uses of AI-related terminology. We therefore use Claude Sonnet 4.6 for the main classification.

Table 1: Descriptive Statistics

Statistic	Value
Time span	1992-2025
Industries (category 1)	21
Firms	
private-sector (baseline)	923
including public sector	1,248
AI keyword matches	326
AI usage	252
Employees	
private-sector baseline	617,423
incl. public sector	1,358,673
national (for reference)	2,558,981
AI related chunks	
candidates	7774
with valid keywords	6781
Firms with explicit "no-AI" statements	74

Note: This table shows the summary statistics of the used data. National employment statistics calculated based on Kaarsen and Sørensen (2026).

4 Trends in AI mentions and reported use

4.1 The rise of AI references over time

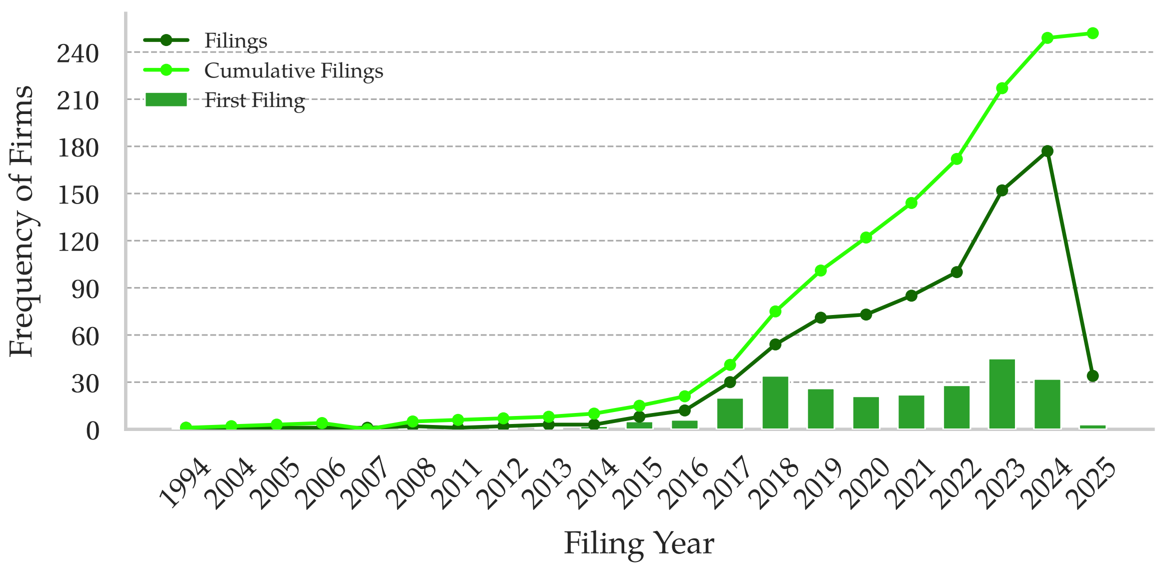
Figure 3 reports the evolution of firms' AI-related annual-report disclosures over time. The bars show the number of firms for which a given year represents the first observed annual-report disclosure of active AI use. The dark green line shows the annual number of firms with validated AI-related filings, while the light green line reports the cumulative number of firms that have mentioned AI up to that year.

The figure highlights two patterns. First, AI-related reporting was rare before the mid-2010s, with only isolated disclosures in earlier years. From 2017 onward, both annual AI filings and first-time active-use disclosures increased markedly. This indicates that AI has become increasingly visible in firms' public reporting narratives. Second, the cumulative number of AI-reporting firms rises especially sharply after 2021 and continues to grow through 2024 and 2025. This suggests that the diffusion of AI narratives is driven not only by repeated mentions among a small set of early adopters but also by a growing number of firms mentioning AI for the first time.

The annual filing series shows a particularly strong increase in 2023 and 2024, consistent with the broader diffusion of machine learning, data analytics, and generative AI technologies and with rising public attention (e.g., Allen 2026; OECD 2024; HAI 2025). The sharp increase in AI mentions in annual reports from 2023 onward coincides with OpenAI's public release of ChatGPT as a research preview in November 2022 (OpenAI 2022). This timing suggests that the growing public accessibility and visibility of generative AI may have contributed to its increasing salience in firms' reporting narratives.[5] The cumulative series is therefore more informative for long-run diffusion, while the annual filing series captures year-specific disclosure intensity.

Importantly, the figure measures the first observed disclosure of active AI use in annual reports, not the true first year of technological adoption. Firms may have implemented AI before publicly mentioning it, and some may use AI without considering it sufficiently material or strategic to disclose in their annual reports. The figure should therefore be interpreted as evidence on the timing and diffusion of AI visibility in corporate reporting narratives rather than as a direct measure of AI adoption.

Figure 3: Reported active use of AI



Note: This figure shows annual and cumulative AI-related disclosure in annual reports. Bars indicate the first year in which a firm is observed to disclose active AI use. A filing is defined as a validated AI-related filing of active use. The dark green line reports the annual number of firms with validated active-use AI-related filings, while the light green line reports the cumulative number of firms with at least one validated AI-related filing in active use up to that year. The measure captures the first observed disclosure in annual reports, not necessarily the true first year of AI adoption. Reports for 2025 are incomplete at the time of data collection (May 2026) and should therefore be interpreted cautiously.

[5] Note that the decline in annual filings in 2025 should be interpreted with caution, since 2025 annual reports are not yet fully available at the time of data collection (May 2026).

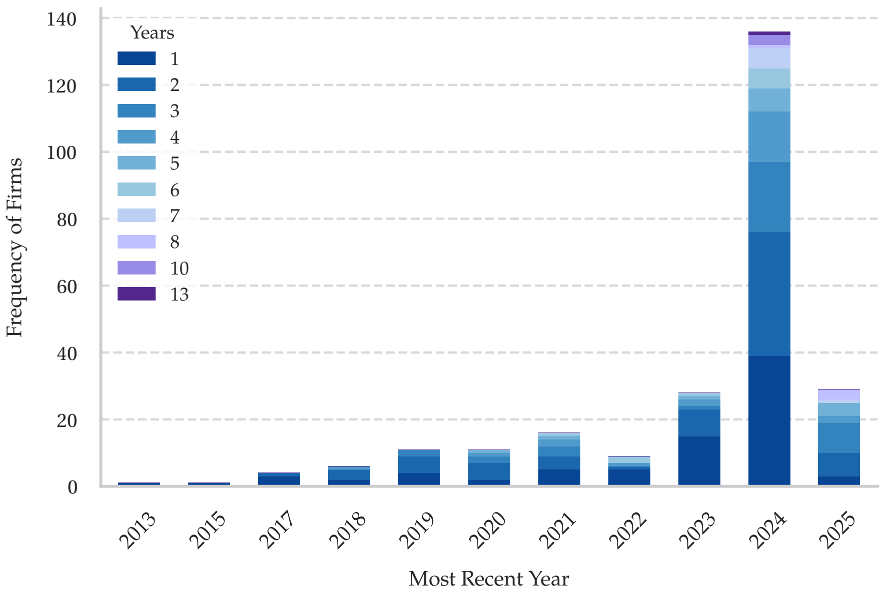
Figure 4 adds a second perspective on the diffusion of AI-related reporting by examining how long firms continue to mention AI once it appears in their annual reports. While Figure 3 documents the timing of AI visibility, Figure 4 distinguishes between firms that mention AI only once and firms that discuss AI repeatedly across multiple reporting years. This distinction is important because repeated disclosure may indicate that AI has become part of the firm's more stable strategic narrative, whereas one-off mentions may reflect experimentation, temporary market attention, or reporting salience in a specific year.

The figure shows that early AI-related reporting is mostly sparse and isolated. In the earlier years, firms with AI activity were few, and most appeared to have short reporting spans. By contrast, the large increase in firms whose most recent AI activity occurs in 2024 comprises not only one-year AI mentions but also firms with multi-year reporting histories. In particular, a substantial share of the 2024 observations have activity spans of two to three years or even four to seven years. This suggests that the recent rise in AI disclosure reflects both new entrants into AI-related reporting and firms that have continued to discuss AI over several annual reports.

At the same time, the results also show that long reporting spans remain less common than shorter ones. Most firms with AI-related disclosures are concentrated in the short and mid-term categories, especially one-year and two- to three-year. This is consistent with the relatively recent acceleration of AI visibility in corporate reporting: many firms only began mentioning AI in the last few reporting years, so long disclosure histories cannot yet be observed for them. The figure should therefore not be interpreted as evidence that AI use itself is short-lived. Rather, it shows that sustained public reporting on AI is still emerging.

Consequently, Figure 4 reinforces the earlier interpretation of annual-report AI disclosure as a narrative measure. The rise in AI references over time is not only an increase in the number of firms mentioning AI, but also an increase in the number of firms for which AI appears repeatedly in public reporting. This persistence suggests that, for at least part of the firm population, AI is moving from a temporary or isolated topic toward a recurring element of strategic communication.

Figure 4: Span of AI reporting across annual reports



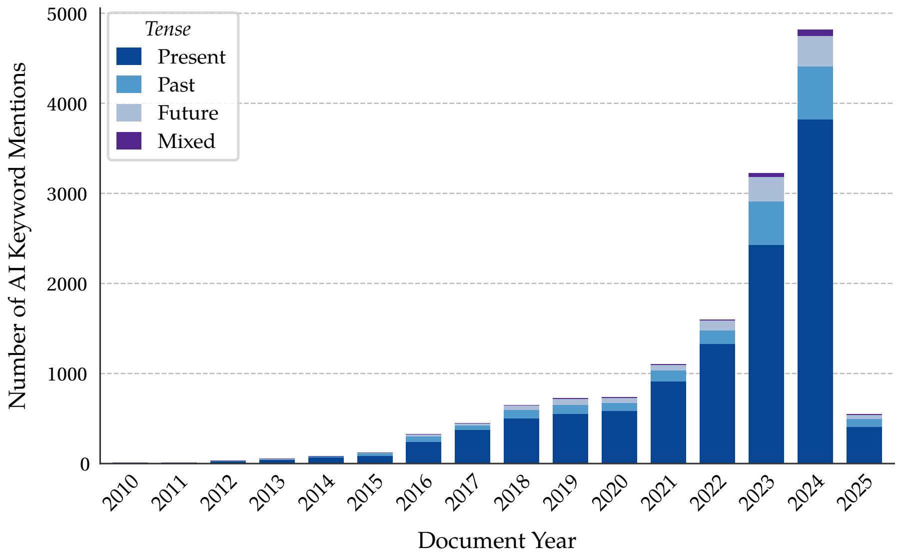
Note: The figure shows the span of consecutive mentions of AI-related context in annual reports. The measure proxies persistence in narrative reporting, not necessarily continuous operational use. Denominator: firms with at least one validated AI mention.

4.2 Temporal framing: current, past, and future AI use

Figure 5 describes the temporal language surrounding AI references. A central finding is that firms predominantly discuss AI in present or current terms. Some reports refer to past use or prior investments, while future-oriented references also appear, but the dominant pattern is current-oriented framing. This suggests that AI is often presented as part of existing activities, processes, or systems rather than only as a speculative future technology.

This result should not be interpreted as evidence on the intensity or maturity of AI deployment. Present-tense language indicates how firms frame AI in their annual reports, but it does not prove that AI is deeply embedded in operations or that it has measurable productivity effects. Nevertheless, the temporal framing is informative because it shows whether firms publicly associate AI with current operations, past experience, or future plans.

Figure 5: Temporal language in AI-related annual-report statements



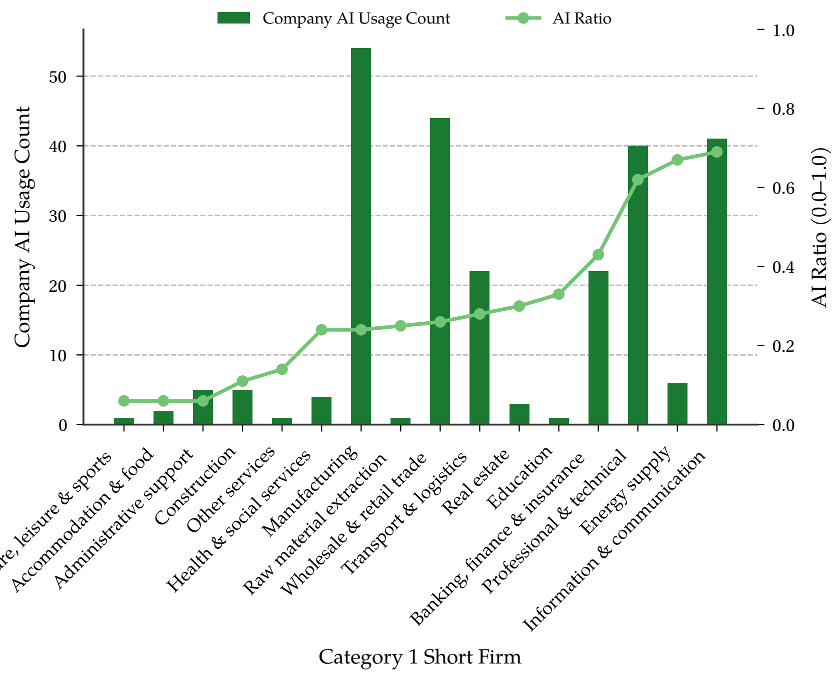
Note: The figure classifies validated AI-related statements according to whether the surrounding language refers primarily to present, past, or future use. Unit: validated AI-related chunks. Classification based on narrative tense and temporal framing in the immediate context of AI mentions.

4.3 Which industries talk about AI?

In the next step, we are interested in which industries? firms talk about AI. To do so, we categorize our sample by different firm characteristics. Figures 6 and 7 describe heterogeneity in AI-related disclosure across industries and firm-size groups. In both figures, the bars report the number of firms with AI-related disclosure, while the line reports the corresponding AI disclosure ratio, defined as the share of firms within the respective group that mention AI in their annual reports. This distinction is important because absolute counts and disclosure intensity capture different aspects of AI visibility. Counts show where most AI-reporting firms are located in the sample, whereas the ratio shows how widespread AI disclosure is within a given industry or size group.

Figure 6 shows that AI-related disclosure is concentrated in several large and data-intensive sectors. Information and communication, banking, finance and insurance, wholesale and retail trade, manufacturing, and professional and technical services account for many of the firms with AI-related annual report disclosures. This pattern is consistent with the expectation that AI is especially visible in industries where digital infrastructure, data processing, automation, analytics, or customer-facing services are central to business operations. At the same time, the disclosure ratio indicates that some sectors have high AI visibility relative to the number of firms in those sectors. Information and communication, professional, scientific and technical services, and banking, finance and insurance display relatively high within-industry AI ratios, suggesting that AI narratives are particularly widespread in knowledge-intensive and data-intensive activities.

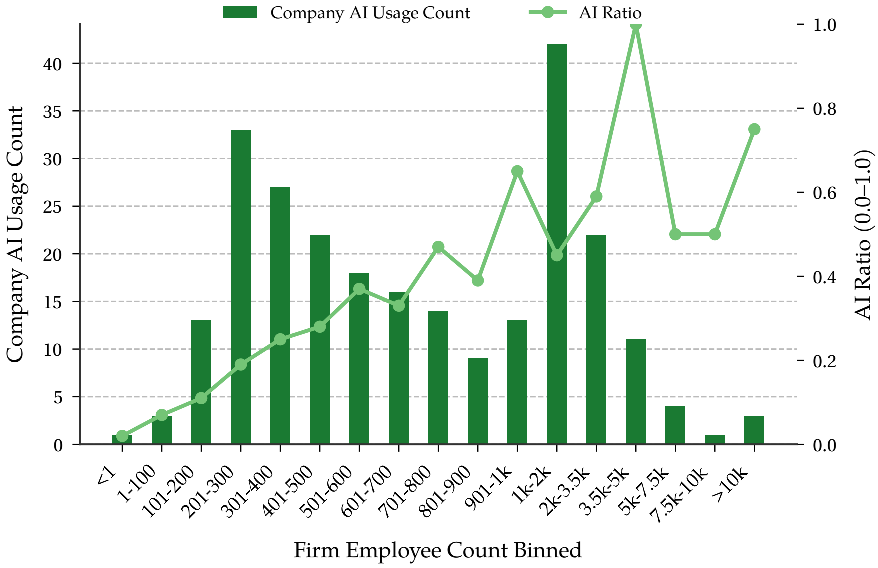
Figure 6: AI-related disclosure and intensity ratio by industry classification



Note: This figure shows AI-related disclosure by industry. Bars report the number of firms with validated AI-related disclosure in each industry. The line reports the AI disclosure ratio, defined as the share of firms within the industry with validated AI-related disclosure. The count measure captures the absolute concentration of AI-reporting firms, while the ratio captures within-industry disclosure intensity. Industry labels are shortened for readability as defined in Table C3.

Figure 7 shows the corresponding pattern by firm size. In absolute terms, AI-related disclosure is most common in several medium-sized and large firm groups, with particularly high counts among firms with 201–400 employees and firms with 1,000–2,000 employees. The disclosure ratio generally tends to be higher among larger firms, but the relationship is not monotonic. Some of the largest size groups contain relatively few firms, so their within-bin ratios can be sensitive to small denominators. The figure therefore provides clearer evidence that larger firms are prominent among AI-reporting firms than of a smooth size gradient in the probability of disclosure.

Figure 7: AI-related disclosure and intensity by firm size



Note: This figure shows AI-related disclosure by firm-size group. Bars report the number of firms with validated AI-related disclosure in each employment-size bin. The line reports the AI disclosure ratio, defined as the share of firms within the employment-size bin with validated AI-related disclosure. The count measure captures the absolute number of AI-reporting firms, while the ratio captures within-bin disclosure intensity. Ratios for size bins containing relatively few firms should be interpreted cautiously.

Both figures, individually and combined, suggest that AI-related annual-report disclosure is unevenly distributed across the firm population. AI narratives are especially visible in data- and knowledge-intensive industries and are also prominent among larger firms in absolute terms. However, the difference between counts and ratios indicates that AI visibility is shaped by both the size of the underlying group and the intensity with which firms in that group disclose AI. This reinforces the interpretation of annual-report AI mentions as selective narrative signals: firms mention AI when it is sufficiently relevant, strategic, or material to communicate publicly, and this relevance varies systematically across industries and firm-size groups.

5 Firm narratives about their AI usage

5.1 Tracking active use

The previous results showed that firms increasingly refer to AI, mostly in present terms, and that large, data-intensive sectors in particular do so. The next question is whether these references describe concrete firm-level activity or whether they remain closer to general market language, governance statements, or external context. To address this, we classify validated AI-related narratives along two dimensions. The first dimension captures the *adoption stage*:

whether AI is described as active use, pilot or testing, planning or exploration, continued research and expansion, talent and culture building, ecosystem and partnerships, awards and recognition, intellectual property, or explicit non-use.[6] The adoption-stage classification, therefore, helps avoid treating all AI mentions as equivalent. The second dimension captures the *target domain*: the part of the firm or business activity to which the AI statement refers, such as enterprise-wide strategy, research and product development, HR and workforce, IT and cybersecurity, operations and manufacturing, customer and marketing, legal and governance, personal data, or finance and accounting.[7]

We differentiate those two categories because AI mentions in annual reports do not all carry the same economic meaning. A statement that AI is already used in operations differs from a statement that the firm is exploring future applications, investing in employee capabilities, participating in an AI partnership, or restricting AI use for personal data. The classification, therefore, separates the maturity of the AI narrative from the organizational area in which AI is located. This helps interpret whether firms describe AI mainly as an operational tool, an experimental technology, a capability-building effort, a governance issue, or a broader strategic theme.

Figure 8 shows the distribution of AI-related initiatives by adoption stage. The dominant category is active use. This indicates that most classified AI initiatives in annual reports are not framed merely as general interest in AI, or distant future plans. Instead, firms most often describe AI as something that is already being used, deployed, implemented, or embedded in current activities. This result is important because it suggests that annual-report AI narratives increasingly refer to concrete firm-level activity rather than only to external technological trends. At the same time, the figure shows that active use is not the only way in which firms discuss AI. Continued research and expansion, talent and culture, ecosystem and partnerships, and initial planning or exploration also appear as meaningful categories. These categories point to the complementary activities surrounding AI disclosure. Firms do not only report the use of AI applications, but they also describe investments in capabilities, organizational learning, partnerships, and further development. This suggests that AI adoption disclosure is often embedded in a broader process of experimentation, capability building, and strategic positioning.

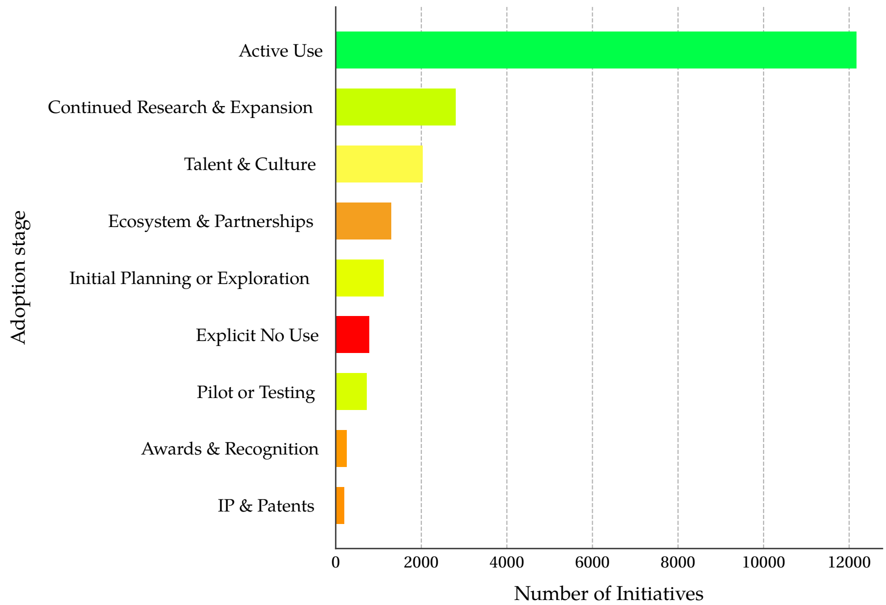
[6] Parts of these dimensions have been used before in a binary manner. For instance, active use has been, and explicit non-use will be, described in binary terms to discuss these target stages in more detail.

[7] Detailed descriptions of the definitions for the adoption stages and the narrative targets can be found in Appendix C, Table C6.

Interestingly, there is a relatively small number of pilot or testing narratives. One interpretation is that firms may be more likely to disclose AI once it has moved beyond a narrow testing stage and can be presented as active use, continued development, or part of broader organizational initiatives. Alternatively, pilot activities may be underreported in annual reports because they are not yet considered material enough for public disclosure. This reinforces the interpretation of annual reports as selective communication documents: they capture AI initiatives that firms consider relevant to disclose, not the full internal portfolio of experiments.

Additionally, we also identify a notable number of initiatives that explicitly state that AI is not used. These cases are distinct from simple silence: firms actively disclose that AI is not applied, not relevant, or excluded in particular contexts. Although explicit non-use is much less common than active use, its inclusion in the taxonomy is important. It shows that annual reports are used not only to communicate adoption, but also to define boundaries around AI use. We examine these statements in greater detail in Section 5.4.

Figure 8: AI initiatives and adoption stages



Note: This figure shows the number of classified AI-related initiatives by adoption stage. Adoption stages indicate whether AI is described as active use, testing, planning, continued development, capability building, partnerships, recognition, intellectual property, or explicit non-use. This graph relies on AI-related initiatives. The counts pool initiatives across firms, reporting years, and annual reports. Thus, a firm may therefore contribute multiple observations. Additionally, adoption-stage categories are not mutually exclusive, so an initiative may contribute to more than one category.

5.2 Where is AI used in the firm?

Having shown that active use is the dominant adoption stage, we next examine where firms locate this active use within the organization. Figure 9 classifies AI-related initiatives coded as active use by their target domain, that is, the business function or organizational area to which the AI application refers. The figure shows that active AI use is most frequently located in Research & Product. Operations & Manufacturing is the second-largest target domain, closely followed by Customer & Marketing. Firms therefore do not describe AI use as confined to a single organizational function. Instead, active use appears prominently in product development, operational processes, and customer-facing activities, alongside broader enterprise-wide applications. This pattern is consistent with the interpretation of AI as a general-purpose technology: firms present AI as relevant across several parts of the organization rather than as a tool restricted to one narrow function.

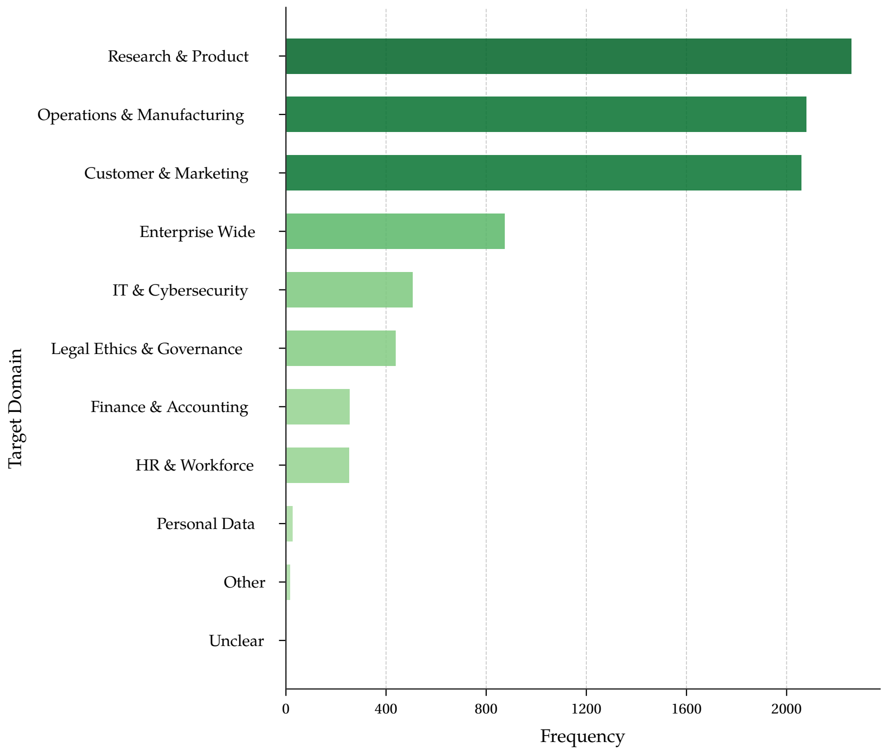
The prominence of Research & Product is particularly informative. It suggests that firms often describe AI in relation to the products and services they offer, or the processes through which these products and services are developed. This may reflect AI being embedded directly in new or improved products, but it may also capture AI supporting research, design, experimentation, software development, or product innovation. In either case, these narratives are consistent with product-related productivity mechanisms developed in Section 2.2: AI may raise the value of output by enabling new products, improving existing products, or increasing the efficiency with which such products are developed.

Operations & Manufacturing and Customer & Marketing represent the next largest categories. These domains point to more immediate business applications. In Customer & Marketing, AI may be used to improve customer interaction, personalization, sales processes, customer analytics, or service quality. In Operations & Manufacturing, AI is more likely to be linked to internal process improvements, logistics, production, supply chains, or operational efficiency. These categories therefore suggest that active AI use is not only product-oriented, but also tied to customer-facing and process-oriented parts of the firm.

These passages linked to enterprise-wide domains describe AI as a broader organizational capability rather than as an application tied to one specific function. Such narratives may reflect attempts to integrate AI into shared systems, workflows, decision-making processes, or employee tools across the firm. This is consistent with the idea that part of AI's productivity effect may come from organizational change and lower coordination costs, rather than from isolated applications in a single department.

The smaller categories should be interpreted cautiously. Active AI use is mentioned less frequently in IT & Cybersecurity, Legal, Ethics & Governance, HR & Workforce, and Finance & Accounting. This does not necessarily imply that AI is rarely used in these functions. Rather, it shows that annual-report narratives about active AI use less often locate AI in these supporting or governance-related domains. Some applications in these areas may be embedded in broader systems, described under enterprise-wide transformation, or simply not considered sufficiently material for annual-report disclosure.

Figure 9: Active AI use and target domain



Note: This figure shows active AI-use initiatives by target domain. Target domains indicate the organizational area or business function to which the AI-related statement refers. A single initiative may be assigned to more than one target domain where the annual-report passage refers to multiple areas. Unit: validated AI-related initiatives classified as active use as described in Section 3.4.

5.3 Productivity implications

5.3.1 Productivity channels in narratives of AI use

Having established whether AI is described as active use, experimentation, capability building, or explicit non-use, the analysis turns to the productivity mechanisms implied by these narratives. We classify AI-related initiatives according to the productivity channels introduced in Section 2. This classification asks through which mechanism firms suggest AI may affect productivity, rather than whether AI is currently used or where in the organization it is applied. These channels capture distinct routes through which AI may become productivity-relevant. Our framework distinguishes eight non-mutually-exclusive productivity channels, organized into three broader groups. The first group concerns labor-related mechanisms. AI may augment labor by supporting workers' or managers' decisions, automate existing tasks by reducing the need for human input, or create new labor tasks by changing the composition of work and the skills required inside the firm. The second group concerns product-related mechanisms. AI may improve the quality of existing products and services, or enable the development of new products, services, or business models. The third group captures other productivity mechanisms beyond the direct labor and product margins. AI may improve scalability by allowing firms to expand activities without proportional increases in organizational costs, reduce the use of non-labor inputs, or raise R&D effectiveness by making experimentation, discovery, design, or innovation processes faster or more productive.

This classification allows the analysis to connect firms' annual-report narratives to economically meaningful productivity mechanisms. A single AI initiative may be linked to more than one channel. For example, an AI-based customer service tool may automate routine interactions, augment employees handling complex cases, improve service quality, and support scalability. The classification should therefore not be interpreted as assigning each initiative to one exclusive productivity effect. Rather, it identifies the set of productivity mechanisms that firms themselves make visible in their AI narratives.

In addition to identifying the channel, we also distinguish the explicitness of the underlying statement. Some annual reports state a productivity mechanism directly, for example, by saying that AI reduces manual work, improves product quality, supports employees, or enables a new service. Other passages imply a productivity channel more indirectly, for example by describing an AI tool or application without explicitly spelling out its economic effect. This distinction is analytically useful because firms may be more willing to state some channels explicitly than others. For instance, firms may openly describe AI-enabled innovation or quality improvements, but be more cautious about labor substitution, staffing reductions, or cost-saving effects. We therefore interpret the productivity-channel counts together with the explicitness of the supporting statements.[8]

Figure 10 provides an overview of the productivity-channel classification. The bars show the number of firms for which a given productivity channel appears in annual report narratives. The line reports the average number of channel mentions per document-year-firm and serves as a proxy for narrative intensity. The figure therefore distinguishes between the extensive margin, that is, how many firms mention a given channel, and the intensive margin, that is, how strongly the channel appears within firm-year reporting.

Several results are shown in this Figure. The first result is that AI is linked to productivity through a broad set of mechanisms. The narratives are not concentrated in a single channel. Firms connect AI to labor-task augmentation, product quality, automation, new products or services, scalability, R&D and innovation, non-labor input efficiency, and the creation of new labor tasks. This supports the interpretation that firms do not frame AI solely as a labor-saving technology, as often discussed in the literature (e.g., Acemoglu and Restrepo 2018; Acemoglu and Restrepo 2019; Hampole et al. 2025) and in public debates (e.g., Filippucci et al. 2024b; Katz 2026; Lane and Saint-Martin 2021; OECD 2023). Instead, AI is described as affecting how work is performed, what firms produce, how innovation takes place, and how firms scale their operations.

The largest firm-count category is labor-task augmentation. This suggests that the most widespread productivity narrative is not full replacement of workers, but the use of AI to support existing tasks. Firms frequently present AI as a tool that helps employees perform their current activities better, faster, or more accurately. This pattern is consistent with the broader finding that annual reports often describe AI in current and operational terms: AI is framed as a complement to existing work processes rather than only as a substitute for labor.

[8]Examples of classified passages, including the model's reasoning and the assigned explicitness level, are reported in Appendix F, Table F2 - F9. These examples illustrate how the classification distinguishes direct statements of productivity mechanisms from more implicit references.

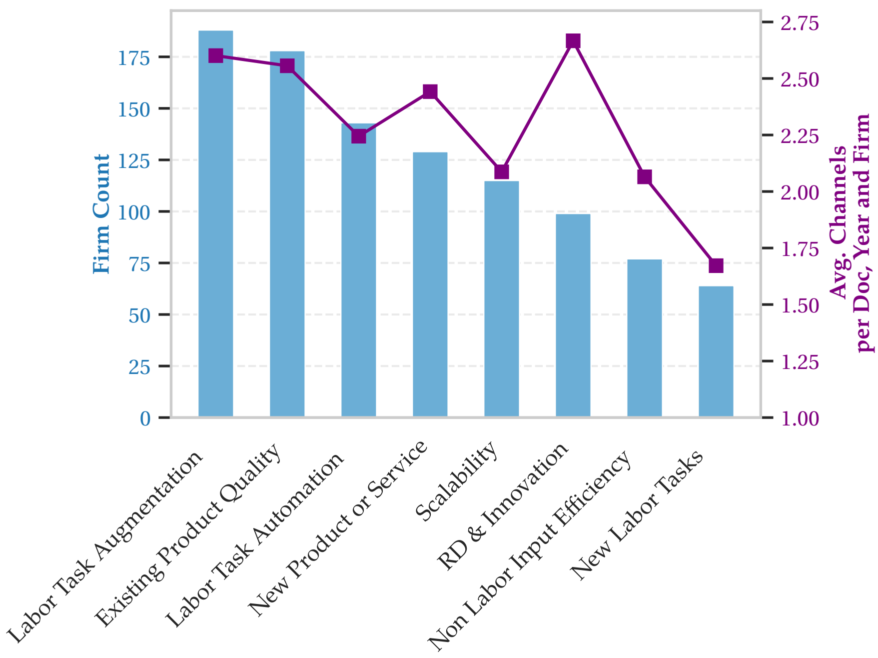
Second, the next-largest category is the improvement in the quality of existing products or services, followed by the automation of existing labor tasks. This ordering is important. Automation is clearly a central productivity channel, but it does not dominate the narratives. A large number of firms instead associate AI with improving the quality, reliability, personalization, speed, or functionality of products and services already offered. This suggests that firms expect AI-related productivity gains not only through lower labor input, but also through higher-quality or more valuable output. This is an important result because these product-side mechanisms have received less systematic empirical attention than labor-related channels in much of the recent AI-productivity literature.

A third result is that the creation of new products or services and increased scalability also appear frequently. These channels point to broader productivity mechanisms than immediate task-level efficiency. New product and service narratives suggest that AI may expand the set of outputs firms can offer. Additionally, scalability narratives suggest that AI may help firms grow without proportional increases in coordination, monitoring, or communication costs. These channels are particularly relevant for interpreting AI as a general-purpose technology, since they point beyond isolated process improvements toward changes in firms' product space and organizational capacity. These channels have received less systematic empirical attention than labor-related mechanisms.

Fourth, Figure 10 also shows a distinction between firm counts and narrative intensity. Labor-task augmentation has the highest firm count and one of the highest intensity values, indicating that it is both widespread and repeatedly emphasized when it appears. R&D and innovation, by contrast, has a lower firm count than several other channels but the highest intensity measure. This channel captures improvements in the effectiveness of the R&D process. Its high narrative intensity is notable because AI may affect productivity growth, rather than only productivity levels, if it increases the rate at which firms generate useful ideas, processes, products, or services (e.g., Cockburn et al. 2018; Aghion et al. 2019; Filippucci et al. 2024a; Trammell and Korinek 2023). The results therefore suggest that AI-related innovation effects may be less widespread across the firm population, but more central in the narratives of firms for which it is relevant.

At the lower end, non-labor-input efficiency and the creation of new labor tasks are less common. This does not imply that AI has no effect on capital, energy, materials, purchased services, or the creation of new tasks. Rather, it indicates that such channels are less frequently made explicit in annual reports. In particular, the low count for new labor tasks suggests that firms are less likely to describe AI as creating new roles, workflows, or human responsibilities than as augmenting existing work or improving existing outputs.

Figure 10: Productivity channels in AI-related firm narratives



Note: This figure shows AI-related productivity channels identified in annual-report narratives. Bars report the unweighted number of unique firms with at least one initiative assigned to the respective productivity channel. The line reports, among firm-year-documents in which the channel is observed, the average number of assignments to that channel per document-year-firm. Channels are not mutually exclusive, so a firm may be associated with more than one productivity channel. Note that the intensity measure (average mentions of respective productivity channel, purple) has a y-axis ranging between 1 and 3.

Table 2 reports the corresponding distribution at the initiative level. Because a single initiative may be associated with several productivity channels, the counts refer to channel assignments rather than mutually exclusive initiatives, and the shares sum to 100 percent across all assignments. Labor-task augmentation is the largest category, accounting for 24.0 percent of assignments, followed by improvements in existing product quality at 20.1 percent. New products or services account for 15.6 percent, while labor-task automation represents 13.5 percent. R&D and innovation, and scalability account for 10.4 and 8.8 percent, respectively. Non-labor input efficiency and new labor tasks are substantially less frequent. The initiative-level results therefore reinforce the firm-level conclusion that AI narratives cover several productivity mechanisms, while placing particular emphasis on labor augmentation, product quality, and the creation of new products and services.

Table 2: Productivity channels in AI initiatives

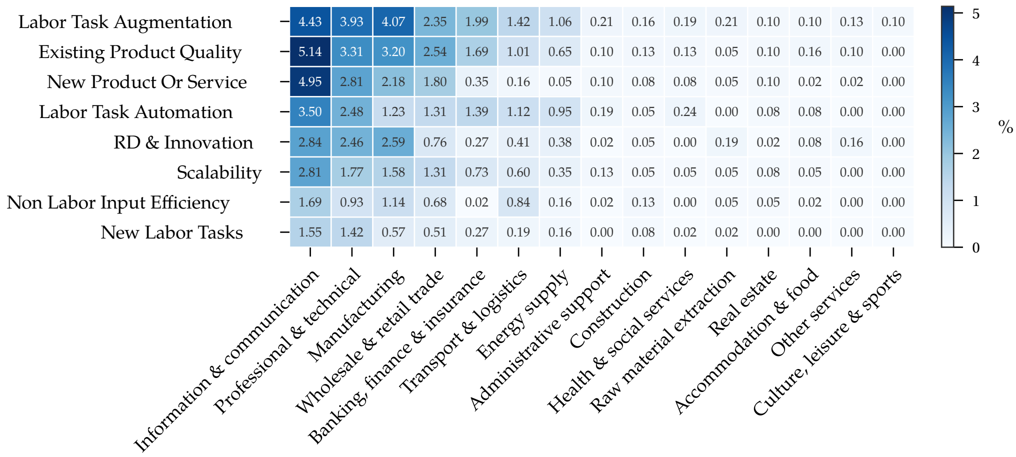
Productivity channel	Count	Share
Labor task augmentation	2567	24.03%
Labor task automation	1438	13.46%
New labor task	354	3.31%
Existing product quality	2142	20.05%
New product or service	1671	15.64%
Non-labor input	457	4.28%
R&D and Innovation	1110	10.39%
Scalability	943	8.83%

Note: This table shows the distribution of productivity channels at the initiative level. A single AI-related initiative may be associated with more than one productivity channel because the classification allows multiple labels. Counts therefore refer to channel assignments rather than mutually exclusive initiatives. The denominator used to calculate the shares is the total number of productivity-channel assignments. Table F1 in Appendix F shows the corresponding values on the firm-year level.

Figure 11 shows how industry–channel combinations contribute to the overall population of AI-reporting firms. Labor-task augmentation, existing product quality, new products or services, and automation appear across several of the industries in which AI reporting is most prevalent. Information and communication, professional and technical services, manufacturing, wholesale and retail trade, and banking, finance and insurance account for many of the largest cells.

Information and communication contributes particularly strongly to existing product quality and new products or services, while professional and technical services also contributes substantially across several channels. Manufacturing makes a sizable contribution to labor augmentation, product quality, R&D and innovation, and non-labor input efficiency. These patterns are consistent with differences in sectoral production processes and business models documented in Figure 10.

Figure 11: Productivity channels by industries



Note: This figure shows the distribution of productivity-channel narratives across industries. Each cell reports the share of all AI-reporting firms that belong to the respective industry and have at least one narrative assigned to the respective productivity channel. Firms may be associated with multiple productivity channels, so cells are not mutually exclusive and percentages do not sum to 100. Darker cells indicate a larger share. Industries are displayed on the horizontal axis and productivity channels on the vertical axis.

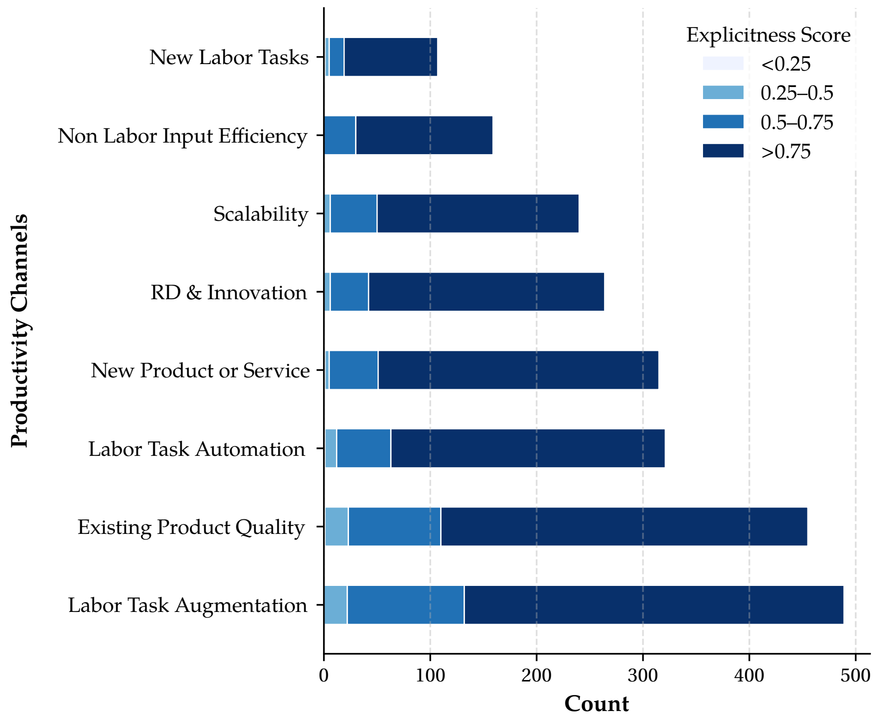
5.3.2 Explicitness of the statements

After documenting which productivity channels appear in firm narratives, we next examine the explicitness of these statements. This distinction matters because the observed frequency of a productivity channel may reflect not only its substantive relevance, but also how directly firms are willing to describe it in annual reports, as discussed before. We therefore complement the productivity-channel counts with an explicitness score that captures whether the supporting text states the productivity mechanism directly or only indirectly.

Figure 12 shows the distribution of explicitness scores across productivity channels using four equally sized intervals. The main result is that the classifications are supported predominantly by highly explicit statements. For every productivity channel, the largest share of observations receives an explicitness score above 0.75. Thus, the productivity-channel results are not primarily driven by passages from which the economic mechanism must be inferred indirectly. Rather, firms frequently state the relevant mechanism in comparatively direct terms.

Labor-task augmentation and improvements in existing product quality have the most highly explicit observations, reflecting their overall prominence in the data. Automation and new products or services also contain substantial numbers of highly explicit statements. Even the less frequent channels, including scalability, non-labor-input efficiency, and new labor tasks, are not concentrated primarily in low-explicitness classifications. Their lower aggregate importance therefore appears to reflect fewer relevant narratives rather than systematically weaker textual evidence.

Figure 12: Explicitness of productivity-channel statements



Note: This figure shows the explicitness score for all productivity channel statements. Thus, we evaluate every AI initiative related to a productivity channel. The explicitness score captures how directly the annual-report statement describes the productivity mechanism underlying the classification. Higher values indicate more explicit textual evidence. The exact definition of explicitness and implicitness is given in the prompt in Appendix B. Channels are not mutually exclusive, so a single AI-related initiative may be associated with multiple productivity channels.

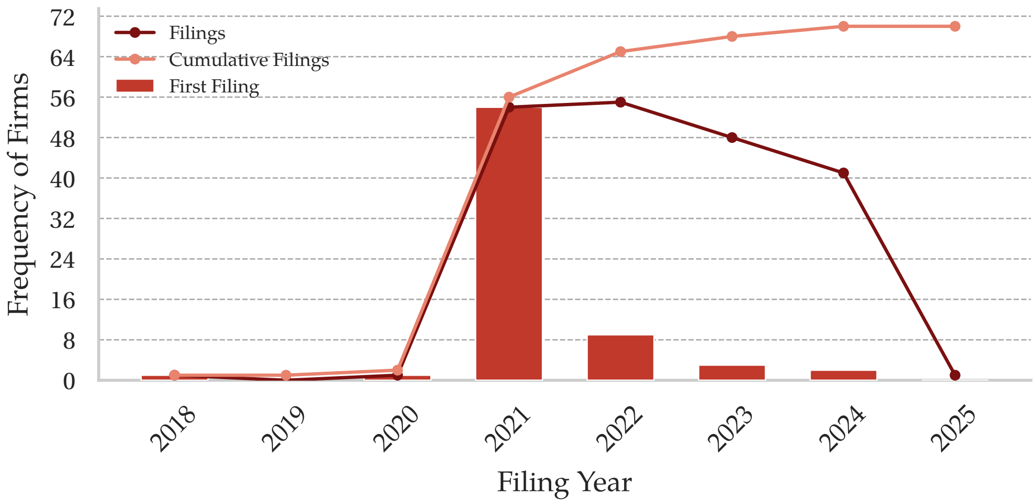
The productivity-channel and explicitness results do not support a narrow interpretation of AI as merely a labor-saving technology. Labor-task augmentation is the most widespread channel, and automation is clearly evident, but firms also frequently link AI to improvements in existing product quality, new products or services, R&D effectiveness, and scalability. Firms therefore frame AI as productivity-relevant through both input-side and output-side mechanisms. The explicitness analysis further shows that these classifications are not driven only by weak or indirect inference: for several of the most frequent channels, particularly labor-task augmentation and improvements in existing product quality, firms provide relatively direct textual evidence. At the same time, the observed distribution reflects both the mechanisms firms perceive and their willingness to articulate those mechanisms publicly. Annual reports therefore provide informative but selective evidence about how firms frame expected productivity effects, rather than evidence about the realized magnitude of those effects.

5.4 Explicit non-use of AI

A second empirical pattern is that some firms do not merely omit AI from their annual reports, but explicitly state that they do not use AI in specific contexts. Figure 13 shows the evolution of such explicit non-use statements over time. These cases are analytically distinct from reports that mention no AI at all. A missing AI reference may simply mean that AI is not considered material, not discussed in the relevant reporting sections, or not captured by the text pipeline. By contrast, explicit non-use indicates that the firm has chosen to communicate a boundary around AI use.

Figure 13 shows that explicit non-use is almost absent before 2021, with only isolated observations between 2018 and 2020. It then increases sharply in 2021, when both the number of filings containing explicit non-use statements and the number of firms making such statements for the first time rise substantially. Explicit non-use became salient as a reporting category consistent with the same time that AI-related governance, profiling, and data-processing concerns became more visible in corporate disclosure (e.g., Cao et al. 2023). After 2021, the number of new first-time non-use statements declines markedly, while the cumulative number of firms with at least one explicit non-use statement continues to increase more gradually. This pattern indicates that the main expansion in explicit non-use occurs early in the period, followed by fewer additional firms disclosing this position in later years.

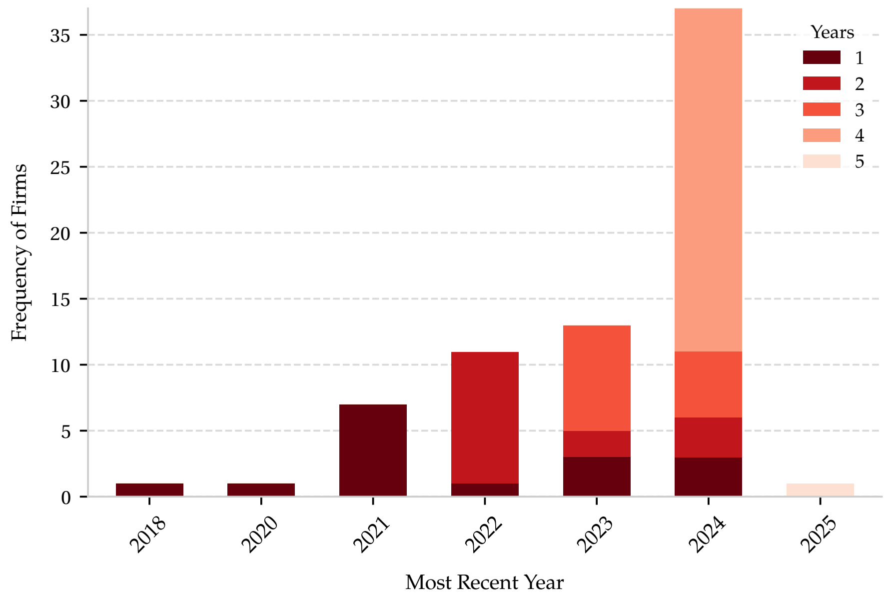
Figure 13: Explicit non-use of AI in annual reports



Note: This figure shows annual and cumulative explicit non-use statements in annual reports. Bars indicate the first year in which a firm is observed to explicitly state that AI is not used, not applied, not relevant, or excluded from specific activities. The dark red line reports the annual number of firms with explicit non-use filings in the given year, while the light red line reports the cumulative number of firms with at least one explicit non-use filing up to that year. The measure captures explicit disclosure of non-use in annual reports and is analytically distinct from the absence of AI-related content. Reports for 2025 are incomplete at the time of data collection (May 2026) and should therefore be interpreted cautiously.

The pattern in Figure 13 raises a further question: whether explicit non-use statements are isolated disclosures or whether they represent a more persistent reporting position. Figure 14 addresses this by showing the span of explicit non-use statements across annual reports. The figure indicates that explicit non-use is often repeated rather than reported only once. Among firms whose most recent non-use statement occurs in 2023 or 2024, multi-year reporting spans account for a substantial part of the observations. This suggests that explicit non-use frequently becomes a recurring disclosure position, although the figure captures persistence in reporting rather than the underlying duration of non-use itself. This persistence suggests that explicit non-use may reflect a stable narrative position rather than a temporary absence of AI activity or a mechanically repeated reporting phrase. In this sense, explicit non-use should be interpreted as an active disclosure category: firms are not only silent about AI, but in some cases repeatedly communicate that AI is not used, not applied, or excluded from specific activities.

Figure 14: Span of explicit non-use across annual reports



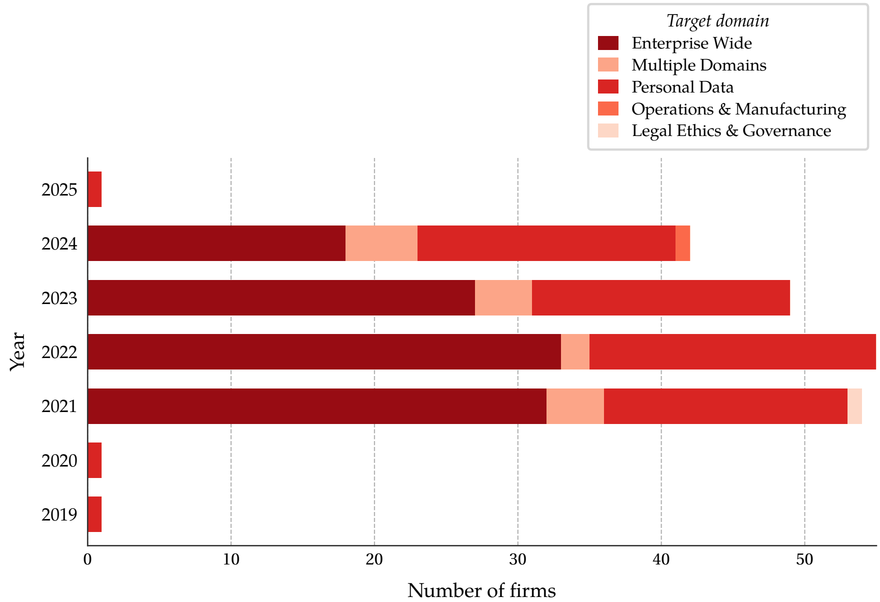
Note: The figure groups firms with at least one explicit AI non-use statement by the most recent year in which such a statement is observed. The stacked segments show the number of consecutive reporting years over which explicit non-use is observed for the firm. The measure captures persistence in explicit non-use reporting, not necessarily a complete absence of AI use across the firm. Reports for 2025 are incomplete at the time of data collection.

Explicit non-use can be further differentiated by the domain in which firms state that AI is not used. An explicit non-use mention does not necessarily imply that the firm lacks all algorithmic or data-driven systems. Rather, it shows that the firm publicly frames AI as unused, excluded, immaterial, or not applicable in a specific context. Additionally, a firm can state non-use in one area but use AI in another. To distinguish different areas, we use the categories of the target domain.

Figure 15 shows that explicit non-use statements are concentrated primarily in two domains: enterprise-wide statements and personal data. Enterprise-wide non-use is particularly prominent between 2021 and 2023, while personal-data-related non-use also accounts for a substantial share throughout the main period. The remaining categories, including legal, ethics and governance, operations and manufacturing, and statements spanning multiple domains, are considerably less frequent. The importance of enterprise-wide statements indicates that some firms formulate non-use as a relatively broad organizational position. These disclosures may state that the company does not use AI or related algorithmic technologies. Personal-data-related statements, on the other hand, are clearly delimited and commonly specify that AI, machine learning, automated decision-making, or similar technologies are not used for profiling, processing personal data, or assessing customers, employees, applicants, or other individuals.

The figure therefore reveals two distinct forms of explicit non-use. Some firms communicate a broad enterprise-level position, while others draw a narrower boundary around sensitive applications involving personal data and automated decisions. Therefore, Figure 15 provides an important contrast to the broader diffusion narrative: as AI becomes more visible in annual reports, firms also become more explicit about the boundaries of acceptable AI use.

Figure 15: Domains of explicit AI non-use

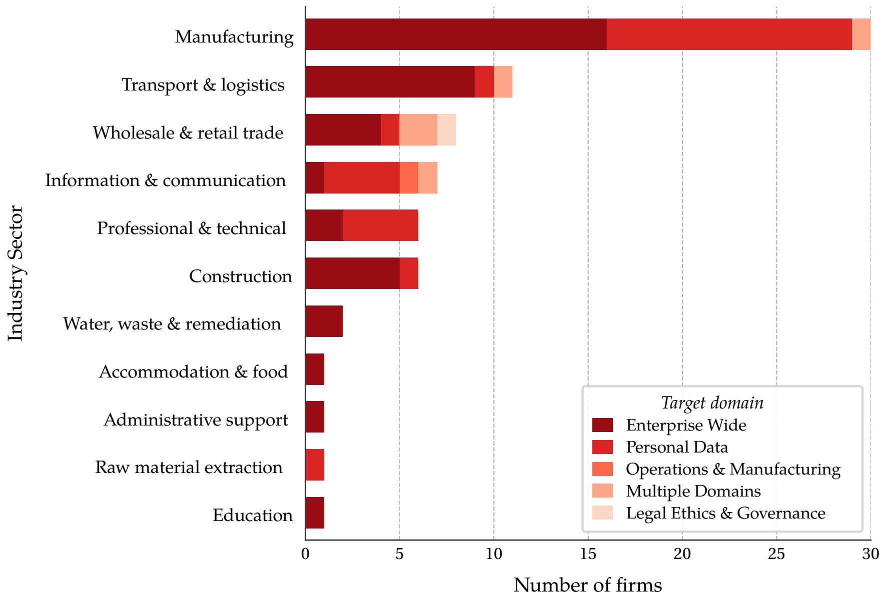


Note: This figure displays the target domains in which firms explicitly state that AI is not used, not applied, not relevant, or excluded. Target domain categories include enterprise-wide non-use, multiple domains (two or more domains), personal data, operations and manufacturing, and legal, ethics, and governance. The measure captures explicit non-use statements and is distinct from reports without validated AI mentions. Observation numbers for 2025 are low because not all reports were available at the time of data collection (May 2026).

Figure 16 adds an industry dimension to explicit AI non-use. In absolute terms, manufacturing contributes by far the largest number of firms with explicit non-use statements. Within manufacturing, the displayed observations are concentrated mainly in enterprise-wide and personal-data-related non-use. The prominence of manufacturing should be interpreted in light of the sector's large representation in the sample and does not by itself show that manufacturing firms have a higher propensity to report non-use.

Transport and logistics, wholesale and retail trade, information and communication, professional and technical services, and construction also contribute explicit non-use observations, but at substantially lower levels. Across these sectors, enterprise-wide and personal-data-related statements account for most of the displayed cases. Operations and manufacturing is rarely the target domain, even for firms in sectors where operational processes are central. This suggests that explicit non-use is more often framed as a broad organizational position or as a restriction concerning personal data than as a statement that AI is unsuitable for production or operational activities.

Figure 16: Target domains of explicit AI non-use by sector



Note: This figure shows explicit AI non-use statements by industry and target domain. Target domains indicate the organizational area or functional context in which firms state that AI is not used, not applied, not relevant, or excluded. A single firm may appear in more than one target domain if a non-use statement is found in more than one report over time. Data are generated in an earlier iteration with Gemini, instead of the main model Claude Sonnet 4.6. Because of that, its precise category counts should be interpreted as supplementary evidence and not compared mechanically with the Claude-based figures.

The explicit non-use results provide an important contrast to the diffusion narrative. If AI references simply increased over time, the evidence might be read as a straightforward story of technological spread. The presence and persistence of explicit non-use complicates that interpretation. Some firms do not merely fail to mention AI, but they actively state that AI is not used, not relevant, or excluded from specific activities. This suggests that non-use can itself be a meaningful part of firms' technology narratives.

Explicit non-use may reflect several different considerations. It may indicate that management does not view AI as material to the firm's operations, that AI use is limited by governance or compliance concerns, or that the firms want to draw a boundary around automated decision-making, profiling, customer interaction, or personal-data processing. From a disclosure perspective, such statements can be interpreted as part of firms' strategic narrative positioning rather than as neutral omissions (e.g., Merkl-Davies and Brennan 2007). These interpretations should be approached with caution because annual reports provide narrative explanations rather than direct evidence of internal constraints. Nevertheless, persistent explicit non-use is informative for understanding diffusion. It shows that AI is not only disclosed as adopted, tested, and discussed unevenly, but is also sometimes publicly bounded. This matters for productivity analysis because expected productivity effects depend not only on technological ability but also on firms' perceived relevance, risk assessment, and willingness to reorganize around the technology.

6 Robustness, sensitivity checks and discussion

6.1 Data-driven productivity drivers

The main productivity-channel analysis classifies AI-related narratives according to the theoretically motivated channels introduced in Section 2. This has the advantage of linking the empirical evidence directly to the economic mechanisms through which AI may affect productivity. At the same time, any theory-based classification raises the concern that the results may partly reflect the categories we impose. As a robustness check, we therefore complement the main analysis with a more data-driven classification of firms' stated AI value drivers.

This robustness check asks a related but distinct question. Instead of asking which theoretically defined productivity channel is present, it asks what firms themselves say AI is expected to contribute to the productivity of their business. The resulting categories capture the stated value proposition of AI, including efficiency and cost savings, customer experience, innovation and growth, risk, security and compliance, talent and workforce, and sustainability and ESG. These categories should not be interpreted as one-to-one equivalents of the theoretical productivity channels. Rather, they provide an empirical contrast: if firms' stated productivity drivers resemble the mechanisms identified in the theory-based productivity classification, this strengthens the interpretation that the main results are not an artifact of the imposed taxonomy.

There are two main reasons for this exercise. First, it allows the data to reveal the productivity rationale firms attach to AI without requiring each statement to be assigned directly to one of the theoretical productivity channels. Second, it provides a bridge between managerial language and economic interpretation. Firms usually do not write annual reports in the language of production functions, labor augmentation, or scale elasticities. Instead, they describe AI in terms of recognizable business objectives: improving processes, serving customers, innovating, managing risk, developing capabilities, or meeting ESG goals. The data-driven productivity-driver classification captures this managerial language and allows us to examine whether it points to productivity mechanisms similar to those in the main theoretical classification.

Table 3 summarizes the data-driven productivity drivers identified in firms' AI-related narratives. The categories capture the business objectives that firms themselves associate with AI, rather than the theoretical mechanism through which those objectives may affect productivity. Innovation is the largest category, accounting for 24.0 percent of all driver assignments, followed by efficiency and cost savings at 22.7 percent. Growth and customer experience each account for approximately 17 percent, while risk, security, and compliance represent 10.9 percent. Sustainability and ESG, and talent and workforce are considerably less frequent. Because individual passages may be associated with more than one driver, the counts represent driver assignments rather than mutually exclusive observations.

The distribution provides an initial indication that firms' stated rationales for AI extend beyond cost reduction alone. Innovation, growth, and customer experience together account for a substantial share of the classifications, suggesting that firms frequently describe AI in terms of improved or expanded output and market opportunities. Efficiency and cost savings nevertheless remain central, consistent with the labor-task automation, augmentation, and non-labor input channels identified in the main analysis. Risk, security, and compliance appear less frequently as a direct value proposition, but it may function as an important enabling or boundary condition for productive AI use. The relatively small shares of talent and workforce and sustainability and ESG suggest that these objectives are more specialized or less often made salient in annual reports.

Figure 17 reports the distribution of data-driven productivity drivers across industries. The left panel shows total counts of AI-related narratives by productivity driver and industry. This panel shows where most AI-related productivity-driver narratives are concentrated in absolute terms. The right panel reports the corresponding within-industry shares and therefore abstracts from differences in industry size. It shows whether firms within a sector tend to frame AI mainly around efficiency, innovation, customer experience, risk and compliance, workforce issues, or sustainability.

Table 3: Data-driven productivity drivers in AI initiatives

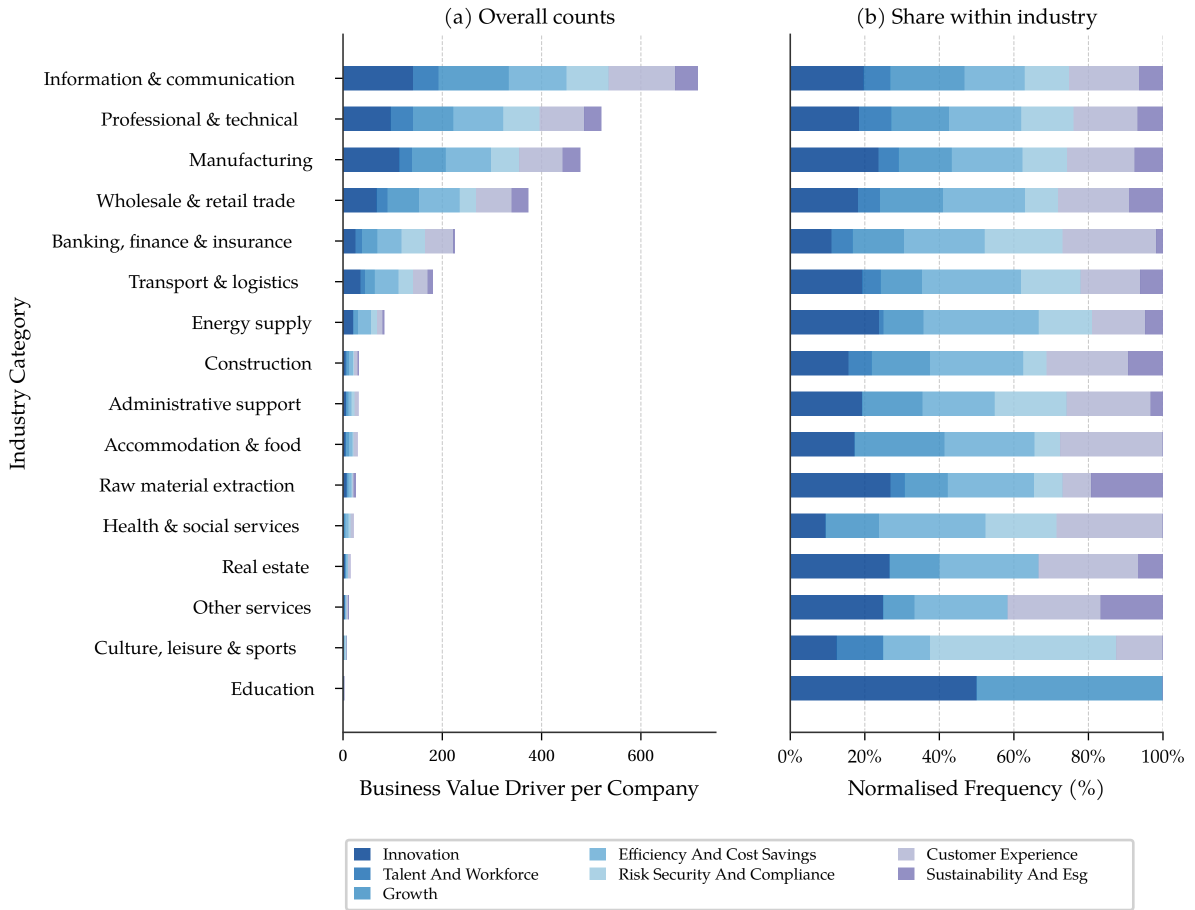
Productivity driver	Logic	Count	Share
Efficiency, cost savings	process improvement, cost reduction, error reduction, speed, workflow automation	3420	22.7%
Customer experience	personalization, customer support, service quality, engagement, recommendations, customer analytics	2613	17.3%
Innovation	new products and services, R&D, experimentation, discovery, product development, and data-driven business models	3618	24.0%
Growth	market expansion, revenue growth, scaling, new customer segments, competitive advantage	2652	17.6%
Risk, security, compliance	risk identification and monitoring, fraud detection, cybersecurity, legal compliance, responsible AI, governance, data protection, and regulatory reporting	1644	10.9%
Sustainability, ESG	sustainability reporting, emissions monitoring, resource optimization, ESG compliance, environmental risk assessment	535	3.6%
Talent, workforce	recruitment, workforce planning, training, employee support, HR analytics, changes in tasks and skills	584	3.9%

Note: This table shows the data-driven productivity drivers identified in firms' AI-related narratives, along with their logic, counts, and shares. A single AI-related passage may be associated with more than one productivity driver because the classification allows multiple labels. Counts therefore refer to driver assignments rather than mutually exclusive passages. The denominator used to calculate the shares is the total number of productivity-driver assignments. Table G1 in Appendix G shows the corresponding values on the firm-year level.

Thus, the first result is that AI narratives are attached to concrete business objectives rather than to abstract technological language alone. Across industries, firms describe AI in relation to efficiency and cost savings, innovation and growth, customer experience, risk, security and compliance, sustainability and ESG, and talent and workforce. This supports the main interpretation of the paper: firms do not merely mention AI as a fashionable technology, but connect it to specific organizational and economic objectives. In absolute terms, data-driven productivity-driver narratives are concentrated in the same sectors in which AI disclosure is generally most visible, especially manufacturing, information and communication, wholesale and retail trade, professional and technical services, and banking, finance and insurance. This pattern is consistent with the earlier heterogeneity results and suggests that the productivity-driver exercise is not identifying an unrelated set of firms or sectors. Rather, it recovers productivity-relevant narratives in the same parts of the economy where AI disclosure is already most salient.

The within-industry shares add a second layer. Although the largest sectors account for most observations in absolute terms, the relative composition of productivity drivers differs across industries. Some sectors place greater emphasis on efficiency and cost savings, while others attach relatively more weight to innovation and growth, customer experience, or risk and compliance. This shows that the data-driven categories capture differences in how firms explain AI. AI is not framed as the same business tool in all sectors, but its stated productivity depends on the sectoral context in which it is used.

Figure 17: Productivity drivers in AI narratives by industry



Note: The left panel displays total counts of AI-related narratives by productivity driver across sectors. The right panel shows the within-sector distribution as shares. Unit: validated AI-related chunks. Denominator: all chunks with classification in the private-sector baseline sample.

Figure G1 in Appendix G provides a more granular view by showing the industry-by-productivity-driver matrix. The heatmap confirms the patterns from Figure 17. Efficiency and cost savings, innovation and growth, and customer experience are widely represented across the largest AI-reporting sectors. Manufacturing, information and communication, wholesale and retail trade, and professional and technical services account for many of the largest cells, indicating that these sectors contribute substantially to the observed data-driven productivity narratives.

With this figure, we show that the data-driven categories do not collapse into a single dominant objective. Efficiency and cost savings are important, but they are not the only stated rationale for AI. Innovation and growth, customer experience, and risk, security, and compliance also appear prominently in several sectors. This mirrors the main productivity-channel results, where AI was not framed only as labor automation or input saving, but also as a source of product improvement, innovation, and organizational capability. At the same time, the figure highlights that some value drivers are more sector-specific. Sustainability & ESG and talent & workforce appear less frequently overall, but they are present in selected industries. This suggests that the data-driven productivity-driver classification captures both broad productivity motives and more specific organizational or regulatory objectives.

Consequently, the data-driven productivity-driver analysis supports the main productivity-channel results. When firms' AI narratives are classified according to their stated business objectives rather than according to the theoretical productivity taxonomy, the same broad interpretation emerges. AI is linked not only to efficiency and cost savings, but also to innovation, customer experience, workforce capabilities, sustainability, and governance. These categories do not map one-to-one into the theoretical channels, but they point to similar underlying mechanisms. This suggests that the main conclusion, that firms frame AI as productivity-relevant through multiple channels rather than through automation alone, is not driven solely by the pre-defined productivity-channel classification.

6.2 Confidence score as an alternative classification measure.

The main productivity channel analysis uses the explicitness score to capture how directly annual reports state the productivity effects. The explicitness score therefore captures a feature of the disclosure itself: how directly the text states the productivity mechanism.

As a robustness check, we repeat the analysis using the classification confidence score. The confidence score captures how certain the model is that a given productivity-channel label is the correct classification for the text. In the prompt, both scores are defined on a scale from 0 to 1, but they measure different concepts. The classification confidence score measures the certainty of the label choice: a value of 1 indicates that the text unambiguously maps to the assigned channel and that no other channel is plausible, while lower values indicate a weaker fit, greater indirectness, or that other channels could also apply. By contrast, the explicitness score measures how directly the text states the productivity mechanism: a value of 1 corresponds to direct statements or named quantitative effects, while lower values indicate that the mechanism is implied or requires inference across sentences. The prompt explicitly instructed the model to treat the two scores as distinct: explicitness concerns the text itself, whereas confidence concerns the label choice.

We choose this test for two reasons: First, it tests whether the productivity channel results are sensitive to using explicitness as the main measure of textual strength. Second, it helps distinguish disclosure behavior from classification reliability. A passage can be implicit but still classified with high confidence if the described use case clearly corresponds to a given productivity channel. Conversely, a passage can contain explicit productivity language but receive lower confidence if the statement could plausibly map to several channels. The confidence score should therefore not be interpreted as the economic importance or magnitude of a productivity channel, but rather that it measures only the certainty of the textual classification.

The results are very similar to the explicitness-based analysis. The same broad ranking of productivity channels emerges: labor-task augmentation, improvements in existing product or service quality, and labor-task automation remain among the most prominent channels. Innovation-related channels, including new products or services and R&D and innovation, also remain clearly visible, while non-labor input efficiency and the creation of new labor tasks continue to appear less frequently. This indicates that the main interpretation does not depend on using the explicitness score. Firms' AI narratives remain multidimensional and cannot be reduced to a narrow automation-or-labor-saving story. Moreover, the confidence score analysis supports the reliability of the underlying classification. The most frequent channels are not concentrated among low-confidence observations. Instead, the main productivity channels are supported by observations for which the model assigns relatively high classification confidence. This suggests that the observed patterns are not primarily driven by uncertain or borderline cases, but by passages that can be classified with reasonable confidence into the productivity-channel taxonomy.

Finally, we can show that the explicitness and confidence measures are strongly related. Appendix G, Figure G2 shows the classification confidence score against the explicitness score. The relationship is positive and strong, with an R^2 of 0.852. This is intuitive: passages that state the productivity mechanism more directly are usually easier to classify with high confidence. At the same time, the two measures are not identical. The scatter around the regression line shows that some passages receive high confidence despite lower explicitness, consistent with cases where the productivity channel is clearly implied by the use case even if the firm does not state the mechanism directly.

6.3 Limitations

Several limitations follow from the nature of the evidence and its interpretations. Annual reports identify expectations, framing, and perceived relevance, not realized adoption or productivity effects. Firms may use AI without reporting it, mention AI for strategic or reputational reasons, or describe pilot projects in ways that overstate their operational importance. Conversely, some firms may avoid detailed AI disclosure because projects are experimental, commercially sensitive, or not considered material for reporting purposes. The results should therefore not be interpreted as estimates of AI intensity or causal effects on productivity. Instead, the contribution of the analysis is to document how firms publicly narrate AI and to use these narratives as early signals of expected productivity channels. The evidence shows where firms themselves locate AI in relation to efficiency, decision-making, innovation, organizational change, and non-use. This provides a basis for interpreting AI diffusion before its effects are fully visible in standard productivity measures.

7 Conclusion

This paper studies how large firms in Denmark describe artificial intelligence in their annual reports and what these narratives reveal about the productivity mechanisms firms associate with AI. Rather than treating any AI mention as evidence of adoption, we distinguish between substantive firm-level initiatives, contextual references, and explicit statements of non-use. We then classify firms' AI narratives according to where AI is used, how mature the reported activity appears to be, and through which productivity channels firms suggest AI may matter.

The analysis is based on annual reports filed with the Danish Business Authority. After converting heterogeneous report formats into machine-readable text, we identify candidate AI-related passages using keyword searches and validate them semantically with an LLM-based classification procedure. This approach produces a narrative measure of AI use and positioning. It does not measure the intensity of adoption or realized productivity effects. Instead, it captures which AI-related activities firms consider sufficiently relevant, material, or strategic to communicate publicly.

Three main findings establish the usefulness of annual reports for studying firms' engagement with AI. First, AI-related disclosure was rare before the mid-2010s but has increased sharply in recent years, particularly after 2021. The increase reflects both repeated reporting by earlier mentions and a growing number of firms discussing AI for the first time. Second, firms predominantly describe AI in present-oriented terms and most often classify their initiatives as active use. This suggests that annual reports contain substantial information about current firm-level applications, rather than only general technological trends or distant plans. Third, explicit non-use is a distinct and persistent narrative category. Some firms repeatedly state that AI is not used in particular contexts, especially in relation to personal data, profiling, automated decision-making, and governance. These statements show that firms communicate not only where AI is used, but also where they deliberately draw boundaries around its use.

The productivity analysis shows that firms associate AI with a broad set of mechanisms. Labor-task augmentation is the most widespread channel: firms commonly describe AI as helping employees perform existing tasks more effectively. Labor-task automation is also important, but it does not dominate the narratives. Firms frequently connect AI to improvements in the quality of existing products and services and to the creation of new products or services. Scalability and greater R&D effectiveness also appear regularly, while improvements in non-labor input efficiency and the creation of new labor tasks are less common. These patterns suggest that firms do not frame AI primarily as a labor-replacing technology. Instead, they describe it as affecting both how output is produced and the quality, variety, and scalability of that output.

The explicitness analysis strengthens this interpretation. Across all productivity channels, most classified statements receive high explicitness scores. The results are therefore not driven mainly by vague AI references from which productivity effects must be inferred indirectly. Firms often state comparatively clearly that AI supports workers, automates tasks, improves products, enables new services, strengthens innovation, or facilitates scaling. At the same time, annual reports remain selective disclosure documents. The prominence of a channel reflects both its perceived relevance and firms' willingness to discuss it publicly. The results should therefore be interpreted as evidence on firms' stated and expected productivity mechanisms, not as estimates of their realized economic effects.

The findings have two broader implications. First, empirical and policy discussions of AI productivity should look beyond automation and labor displacement. The narratives indicate that product quality, new products and services, R&D effectiveness, and scalability are also central to how firms understand the economic value of AI. These channels may affect productivity through higher-value output and faster innovation, even when labor input does not immediately decline. Second, productive use of AI is likely to depend on complementary capabilities. The concentration of AI disclosure among larger and more knowledge-intensive firms is consistent with the importance of data infrastructure, specialized skills, managerial capacity, and organizational adaptation. Policies intended to support AI-related productivity should therefore focus not only on access to AI technologies, but also on firms' ability to integrate them effectively and responsibly.

Annual reports cannot provide a complete measure of AI adoption, nor do they reveal whether the expected productivity gains ultimately materialize. Their value lies elsewhere: Annual reports show how firms publicly interpret an emerging general-purpose technology, which uses they consider economically relevant, and which constraints they choose to emphasize. In this sense, firm narratives offer early evidence on the channels through which AI may shape productivity before these effects become fully visible in conventional firm-level performance measures.

References

Acemoglu, D. and Restrepo, P. (2018). Artificial intelligence, automation, and work. In *The economics of artificial intelligence: An agenda*, pages 197–236. University of Chicago Press.

Acemoglu, D. and Restrepo, P. (2019). Automation and new tasks: How technology displaces and reinstates labor. *Journal of Economic Perspectives*, 33(2):3–30.

Acemoglu, D. and Restrepo, P. (2020). The wrong kind of ai? artificial intelligence and the future of labour demand. *Journal of Economic Perspectives*, 34(1):3–32.

Aghion, P., Jones, B. F., and Jones, C. I. (2019). Artificial intelligence and economic growth. In Agrawal, A., Gans, J., and Goldfarb, A., editors, *The Economics of Artificial Intelligence: An Agenda*, pages 237–282. University of Chicago Press.

Aldasoro, I., Gambacorta, L., Pal, R., Revoltella, D., Weiss, C., and Wolski, M. (2026). Ai adoption, productivity and employment: evidence from european firms. Technical report, EIB Working Papers.

Allen, J. S. (2026). Monitoring ai adoption in the us economy. *FEDS Notes. Washington: Board of Governors of the Federal Reserve System*.

Autor, D. H., Levy, F., and Murnane, R. J. (2003). The skill content of recent technological change: An empirical exploration. *Quarterly Journal of Economics*, 118(4):1279–1333.

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

Baker, S. R., Bloom, N., and Davis, S. J. (2016). Measuring economic policy uncertainty. *Quarterly Journal of Economics*, 131(4):1593–1636.

Basnet, A., Elias, M., Salganik-Shoshan, G., Walker, T., and Zhao, Y. (2025). Analyzing the market's reaction to ai narratives in corporate filings. *International Review of Financial Analysis*, 105:104378.

Botosan, C. A. (1997). Disclosure level and the cost of equity capital. *Accounting Review*, 72(3):323–349.

Bresnahan, T. F., Brynjolfsson, E., and Hitt, L. M. (2002). Information technology, workplace organization, and the demand for skilled labor: Firm-level evidence. *Quarterly Journal of Economics*, 117(1):339–376.

Brynjolfsson, E., Li, D., and Raymond, L. R. (2023). Generative ai at work. NBER Working Paper 31161.

Brynjolfsson, E., Rock, D., and Syverson, C. (2017). Artificial intelligence and the modern productivity paradox: A clash of expectations and statistics. In *NBER Macroeconomics Annual*, volume 32, pages 85–135.

Calvino, F. and Fontanelli, L. (2024). A portrait of ai adopters across countries: Firm characteristics, assets' complementarities and productivity. Technical report, OECD.

Cao, S., Jiang, W., Yang, B., and Zhang, A. L. (2023). How to talk when a machine is listening: Corporate disclosure in the age of ai. *The Review of Financial Studies*, 36(9):3603–3642.

Christensen, H. B., Hail, L., and Leuz, C. (2013). Mandatory ifrs reporting and changes in enforcement. *Journal of Accounting and Economics*, 56(2–3):147–177.

Cockburn, I. M., Henderson, R., and Stern, S. (2018). The impact of artificial intelligence on innovation: An exploratory analysis. In *The economics of artificial intelligence: An agenda*, pages 115–146. University of Chicago Press.

Czarnitzki, D., Fern'andez, G. P., and Rammer, C. (2023). Artificial intelligence and firm-level productivity. *Journal of Economic Behavior and Organization*, 211:188–205.

Davis, A. K., Piger, J., and Sedor, L. M. (2007). Beyond the numbers: Measuring the information content of earnings press release language. *Journal of Accounting Research*, 45(4):1001–1033.

Dell'Acqua, F. et al. (2024). Navigating the jagged technological frontier: Field experimental evidence of the effects of ai on knowledge worker productivity and quality. Working paper.

Diamond, D. W. and Verrecchia, R. E. (1991). Disclosure, liquidity, and the cost of capital. *Journal of Finance*, 46(4):1325–1359.

European Commission (2017). Guidelines on non-financial reporting (methodology for reporting non-financial information). Technical report, Official Journal of the European Union.

Filippucci, F., Gal, P., Jona-Lasinio, C., Leandro, A., and Nicoletti, G. (2024a). The impact of artificial intelligence on productivity, distribution and growth: Key mechanisms, initial evidence and policy challenges. *OECD Artificial Intelligence Papers*, (15). Available online: `https://www.oecd.org/content/dam/oecd/en/publications/reports/2024/04/the-impact-of-artificial-intelligence-on-productivity-distribution-and-growth_d54e2842/8d900037-en.pdf`. Accessed 13 July 2026.

Filippucci, F., Gal, P., and Schief, M. (2024b). Miracle or myth? assessing the macroeconomic productivity gains from artificial intelligence. *OECD Artificial Intelligence Papers*, (29).

Gentzkow, M., Kelly, B., and Taddy, M. (2019). Text as data. *Journal of Economic Perspectives*, 33(3):3–24.

Grimmer, J., Roberts, M. E., and Stewart, B. M. (2022). *Text as Data: A New Framework for Machine Learning and the Social Sciences*. Princeton University Press.

Grimmer, J. and Stewart, B. M. (2013). Text as data: The promise and pitfalls of automatic content analysis methods for political texts. *Political Analysis*, 21(3):267–297.

HAI (2025). The 2025 AI Index Report. Technical report, Stanford Institute for Human-Centered Artificial Intelligence.

Hampole, M., Papanikolaou, D., Schmidt, L. D., and Seegmiller, B. (2025). Artificial intelligence and the labor market. *NBER Working Paper*, (33509).

Healy, P. M. and Palepu, K. G. (2001). Information asymmetry, corporate disclosure, and the capital markets: A review of the empirical disclosure literature. *Journal of Accounting and Economics*, 31(1–3):405–440.

Huang, X., Teoh, S. H., and Zhang, Y. (2014). Tone management. *The Accounting Review*, 89(3):1083–1113.

Humlum, A. and Vestergaard, E. (2025). Still waters, rapid currents: Early labor market transformation under generative ai. *NBER Working Paper*, (33777).

Kaarsen, N. and Sørensen, A. (2026). The rockwool foundation firm demographics database. Unpublished working paper.

Katz, D. (2026). Artificial intelligence and the economics of adjustment. *International Monetary Fund: F&D Magazine*. Available online: https://www.imf.org/en/publications/fandd/issues/series/analytical-series/straight-talk-artificial-intelligence-and-the-economic. Accessed 13 July 2026.

Lane, M. and Saint-Martin, A. (2021). The impact of artificial intelligence on the labour market: What do we know so far? *OECD Social, Employment and Migration Working Papers*, (256). Available online: https://www.oecd.org/en/publications/2021/01/the-impact-of-artificial-intelligence-on-the-labour-market_a4b9cac2.html. Accessed 13 July 2026.

Li, F. (2008). Annual report readability, current earnings, and earnings persistence. *Journal of Accounting and Economics*, 45(2–3):221–247.

Li, F. (2010). The information content of forward-looking statements in corporate filings—a naïve bayesian machine learning approach. *Journal of Accounting Research*, 48(5):1049–1102.

Loughran, T. and McDonald, B. (2011). When is a liability not a liability? textual analysis, dictionaries, and 10-ks. *Journal of Finance*, 66(1):35–65.

Loughran, T. and McDonald, B. (2016). Textual analysis in accounting and finance: A survey. *Journal of Accounting Research*, 54(4):1187–1230.

Merkl-Davies, D. M. and Brennan, N. M. (2007). Discretionary disclosure strategies in corporate narratives: Incremental information or impression management? *Journal of Accounting Literature*, 26:116–196.

Noy, S. and Zhang, W. (2023). Experimental evidence on the productivity effects of generative artificial intelligence. Working paper.

OECD (2019). *Artificial Intelligence in Society*. OECD Publishing.

OECD (2023). *OECD Employment Outlook 2023: Artificial Intelligence and the Labour Market*. OECD Publishing. Available online: https://www.oecd.org/en/publications/oecd-employment-outlook-2023_08785bba-en. Accessed 13 July 2026.

OECD (2024). *The Adoption of Artificial Intelligence in Firms*. OECD Publishing.

OpenAI (2022). Introducing chatgpt. Available online: https://openai.com/index/chatgpt/?utm_source=chatgpt.com. Accessed: 2026-07-14.

Smith, R. (2007). An overview of the tesseract ocr engine. In *Proceedings of the Ninth International Conference on Document Analysis and Recognition*, pages 629–633.

Statistics Denmark (2026). ITAV19: Enterprises use of robot technology and artificial intelligence (10+ employees) by use, activity (NACE REV.2) and enterprise size. StatBank Denmark. Accessed: 2026-05-21.

Syverson, C. (2017). Challenges to mismeasurement explanations for the u.s. productivity slowdown. *Journal of Economic Perspectives*, 31(2):165–186.

Trammell, P. and Korinek, A. (2023). Economic growth under transformative ai. *Annual Review of Economics*, 18.

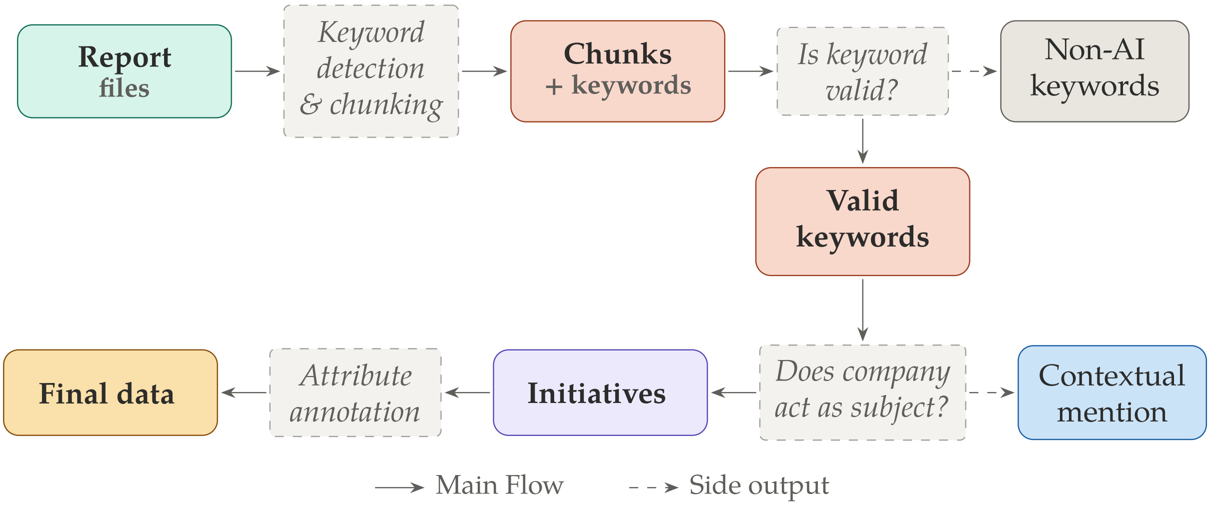
Verrecchia, R. E. (2001). Essays on disclosure. *Journal of Accounting and Economics*, 31(1–3):391–425.

Appendix

A Technical details on the methodology

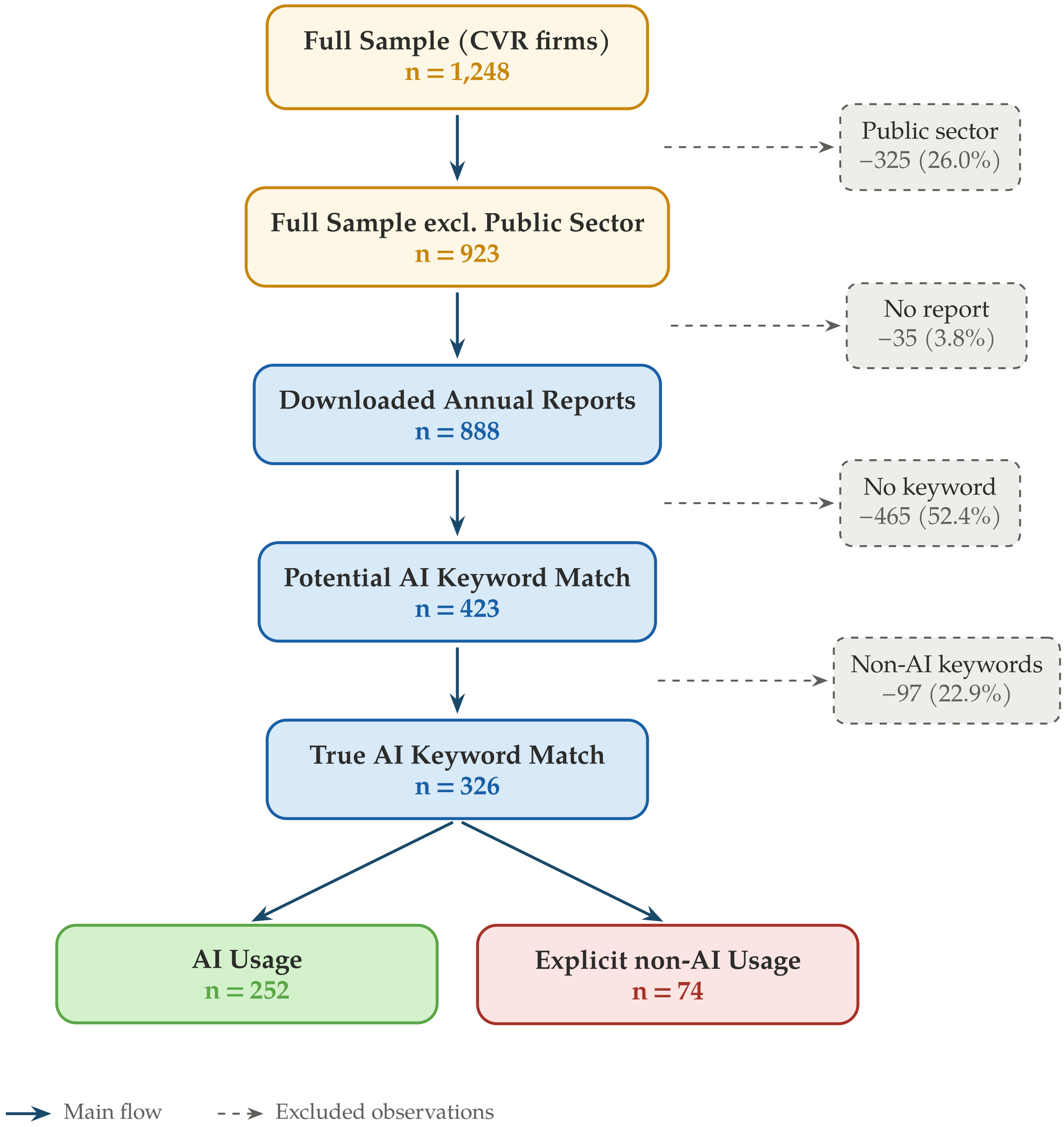
Figure A1 provides a detailed overview of the text-processing workflow. The procedure starts with the extracted annual report files and applies keyword detection to identify candidate passages. Because raw keyword matches may include false positives, ambiguous abbreviations, or references unrelated to artificial intelligence as a technology, the candidate chunks are not used directly as the final measure. Instead, each candidate's passage is screened for semantic validity. Valid AI-related mentions are then evaluated in context to determine whether the company itself is the relevant subject of the statement. This distinction separates firm-level AI narratives from contextual references to external actors, general market trends, competitors, or regulatory developments. The final step annotates subject-relevant AI-related passages as initiatives and assigns additional attributes used in the empirical analysis. These annotations include adoption stage, technology category, target domain, productivity driver, and financial or risk relevance. The resulting dataset is therefore based on validated and contextually classified AI-related text chunks, rather than simple document-level keyword counts.

Figure A1: Processing workflow for identifying and classifying AI-related report text



Note: This figure summarizes the text-processing workflow used to transform annual-report files into the final analytical dataset. Report files are first searched for AI-related keywords and segmented into text chunks around detected keyword occurrences. Candidate chunks are then screened to assess whether the keyword is used in a substantively AI-related sense. Invalid keyword matches, such as ambiguous terms, abbreviations, or non-technology references, are excluded as non-AI keywords. Valid keyword mentions are then evaluated in context to determine whether the company itself is acting as the subject of the AI-related statement or whether the mention refers only to external actors, market developments, or general context. Subject-relevant chunks are classified into AI initiatives and subsequently annotated with additional attributes, such as adoption stage, technology category, target domain, productivity driver, and financial or risk relevance. The final dataset, therefore, consists of validated and annotated AI-related text passages rather than raw keyword hits. Dashed arrows indicate side outputs or filtering steps, while solid arrows indicate the main processing flow.

Figure A2: Sample construction



Note: This figure illustrates the sequential filtering process used to construct the final sample of firms with validated AI-related content in their annual reports. We start with n = 1,248 largest firms in the Danish Business Authority registry (CVR). Those are reduced to 923 private-sector firms after removing the public-sector firms. Due to availability, annual reports have been downloaded only for 888 firms. The keyword screening identifies 423 firms with potential AI mentions, of which 326 firms pass semantic validity checks. Among firms with validated AI-related content, 252 report at least one firm-level AI-use initiative and 74 contain at least one explicit non-use statement. These groups are not mutually exclusive, since a firm may report AI use in one domain while explicitly ruling it out in another.

B Prompt

B.1 Description of the prompt used for LLM-based semantic validation and classification

To ensure transparency and reproducibility, this appendix reports the full prompt used for the LLM-based classification of AI-related candidate chunks. The prompt guides the model through a conservative, rule-based coding procedure. It instructs the model to extract only AI-related activities over which the firm has direct agency and to classify each identified initiative according to its adoption stage, organizational scope, deployment mode, artifact type, business-value drivers, and productivity channels. Particular attention is given to the treatment of explicit non-use statements, the distinction between internal and customer-facing applications, the separation of classification confidence from textual explicitness, and the requirement that all classifications be supported by specific evidence from the underlying text. The prompt also specifies the required JSON output structure and prohibits the inclusion of unsupported or obsolete fields.

The full prompt is reproduced below. Its substantive instructions, classification rules, thresholds, and output requirements are preserved; only the formatting has been adapted to improve readability.

B.2 Full prompt

```
SYSTEM_INSTRUCTION = f"""
```

You are a specialized Data Extraction AI. Your goal is to analyze text chunks from Annual Reports and extract structured data about Artificial Intelligence (AI) adoption. Your three **PRIMARY** focal dimensions are: Workforce Impact, Productivity Channels, and Business Value Drivers. All other fields exist to provide interpretive context for these three.

1. INITIATIVE IDENTIFICATION & ADOPTION STAGE

Only extract items where the company has **direct agency** over an AI activity.

Adoption Stage values:

> `ACTIVE_USE` · `PILOT_OR_TESTING` · `INITIAL_PLANNING_OR_EXPLORATION`
> `CONTINUED_RESEARCH_AND_EXPANSION` · `TALENT_AND_CULTURE` · `ECOSYSTEM_AND_PARTNERSHIPS`
> `AWARDS_AND_RECOGNITION` · `IP_AND_PATENTS` · `EXPLICIT_NO_USE` · `OTHER` · `UNCLEAR`

→ Always fill adoption_stage_confidence (float 0.0–1.0) rating your classification certainty:

- 1.0 = unambiguous signal: 'we deployed', 'is live in production', 'launched in Q3'
- 0.7 = clear signal but minor hedging: 'we are using', 'has been implemented'
- 0.5 = reasonable inference but language is hedged or another stage is plausible
- 0.2 = very indirect signal—an alternative stage is genuinely plausible

This score captures classification uncertainty only—it is independent of text clarity and evidence explicitness, which are captured in other fields.

2. THE DATA ETHICS PARADOX (CRITICAL)

- Scoped restriction: EXPLICIT_NO_USE + narrow target_domain.
- Global ban: EXPLICIT_NO_USE + ENTERPRISE_WIDE target_domain.
- 'Very limited' use → ACTIVE_USE (not no-use).
- 'Not yet but monitoring' → INITIAL_PLANNING_OR_EXPLORATION.

3. ANTI-HALLUCINATION RULES

- Do NOT extract AOI, API, RPA, or similar as AI unless the text explicitly defines them as Artificial Intelligence.
- {f'- Language: Input text may be {LANGUAGE} or English. ALWAYS output all fields in ENGLISH.' if LANGUAGE else ''}

4. INITIATIVE SCOPE & DEPLOYMENT MODE

Set both initiative_scope and ai_deployment_mode on every Initiative, then explain **BOTH** choices together in one scope_and_mode_reasoning field.

initiative_scope—WHO benefits?

- INTERNAL_OPERATIONS: AI improves the company's own processes/productivity.
  Signal: 'we use AI to...', 'our teams use...', 'internal tool', 'operational efficiency'.
- CUSTOMER_FACING: AI embedded in a product/service delivered to external customers.
  Signal: 'we offer customers...', 'our clients can...', 'AI-powered feature in our product'.
- BOTH: Text explicitly describes both beneficiaries.

ai_deployment_mode—HOW is AI packaged?

- PRODUCT_OR_FEATURE: AI is the core of—or a distinct named feature in—a deliverable.
- PROCESS_TOOL: AI supports or automates an internal business process; not the end-product.
- BOTH: Described as serving both roles.
- Note: INTERNAL_OPERATIONS + PRODUCT_OR_FEATURE is valid (e.g. an internal AI platform).

scope_and_mode_reasoning: Cite the specific words that reveal both who benefits and how AI is packaged. Write once—do not repeat the same passage for each field separately. If ambiguous on either axis, name what is unclear.

4b. AI ARTIFACT TYPE [NEW]

Classify what kind of AI artefact this initiative represents using `ai_artifact_type`:

- `NEW_PRODUCT_OR_SERVICE`: A net-new product/service brought to market. Did not exist before.
  Ask: 'Is the company launching something new?' → Yes → `NEW_PRODUCT_OR_SERVICE`.
- `EXISTING_PRODUCT_ENHANCEMENT`: AI added to an existing product to improve or extend it.
  Ask: 'Is AI being added to something that already existed?' → Yes → `EXISTING_PRODUCT_ENHANCEMENT`.
  Note: this maps closely to `GROWTH` in `business_value_drivers`.
- `INTERNAL_PROCESS_TOOL`: AI used inside the company for operational tasks; not customer-facing.
- `PLATFORM_OR_INFRASTRUCTURE`: Foundational AI layer enabling other initiatives.
  e.g. data platform, MLOps pipeline, AI governance framework.
- `NAMED_FEATURE`: A specific named AI capability within a broader product.
  e.g. 'Clara', 'Advisory GPT', 'Copilot for Finance'.

Also set `artifact_name` to the specific name if one is explicitly stated in the text. Set to null if the initiative is unnamed or refers to AI generically.

5. WORKFORCE IMPACT

```
## v7: workforce_impacts commented out âĂŤ not part of current
research focus.
## Do NOT populate workforce_impacts. Omit the field entirely from the
output.
```

6. BUSINESS VALUE DRIVERS

Populate `business_value_drivers` as a list of BusinessValueDriver objects.

- `driver`: `EFFICIENCY_AND_COST_SAVINGS` | `RISK_SECURITY_AND_COMPLIANCE` | `INNOVATION` | `GROWTH` | `CUSTOMER_EXPERIENCE` | `TALENT_AND_WORKFORCE` | `SUSTAINABILITY_AND_ESG` | `OTHER` | `UNCLEAR`
- CRITICAL DISTINCTION—INNOVATION vs. GROWTH:
  - `INNOVATION`: Something net-new—new product, capability, market, R&D breakthrough.
    Ask: 'Did this exist before?' → No → `INNOVATION`.
  - `GROWTH`: Scaling / enhancing what already exists.
    Ask: 'Is the company doing more of something they already do?' → Yes → `GROWTH`.
  - Both can coexist—use separate entries.
- `classification_confidence` (float 0.0–1.0): How confident is the model that this driver label is the correct classification for what the text describes?

1.0 = unambiguously maps to this driver, no other is plausible

0.7 = clear fit but another driver could also apply

0.4 = reasonable but text is vague or fit is indirect

0.1 = marginal—inferable but barely supported

Rate each driver INDEPENDENTLY—do NOT normalise across drivers.

– `explicitness_score` (float 0.0–1.0): How explicitly does the text state this driver?

1.0 = named figures, direct statements ('saved DKK 15m', 'reduced time by 40%')

0.7 = clear intent without quantification ('AI enables us to serve clients faster')

0.4 = implied by the described activity but not directly stated

0.1 = very indirect—requires inference across multiple sentences

Distinct from `classification_confidence`: explicitness is about the TEXT, confidence is about the LABEL CHOICE.

– `evidence_text`: Cite specific words or phrases. If weak or indirect, say so explicitly.

7. PRODUCTIVITY CHANNELS

Populate `productivity_channels` as a list of ProductivityChannel objects.

Each entry has three fields:

– `channel`: one of the enum values defined below—one value per entry

– `classification_confidence` (float 0.0–1.0): how certain you are this label is correct.

Rate INDEPENDENTLY per entry—do NOT normalise across channels.

– `explicitness_score` (float 0.0–1.0): how explicitly the text describes this mechanism.

1.0 = mechanism described with specifics ('processes 10,000 invoices/day')

0.7 = mechanism stated but without operational detail

0.4 = mechanism implied by the described activity

0.1 = very indirect—requires inference across sentences

– `evidence_text`: cite the specific words. If weak or indirect, say so explicitly.

CRITICAL: Do NOT code any channel from generic statements about AI, digitalisation, innovation, efficiency, or transformation. A specific use case must be described. This applies equally to all channels below.

Valid channel values:

```
LABOUR_TASK_AUTOMATION | LABOUR_TASK_AUGMENTATION | NEW_LABOUR_TASKS |
NON_LABOUR_INPUT_EFFICIENCY | EXISTING_PRODUCT_QUALITY |
NEW_PRODUCT_OR_SERVICE |
RD_AND_INNOVATION | SCALABILITY
```

Channels are not mutually exclusive—code all that apply with concrete evidence. For each channel, code it only if the initiative contains:

Explicit mention: the initiative directly states that AI affects this channel.

Implicit mention: the initiative describes a concrete AI use case that clearly implies this channel, even if the channel is not named.

CHANNEL DEFINITIONS

`LABOUR_TASK_AUTOMATION`

AI replaces or reduces the need for human work in tasks previously performed by employees, e.g. automated customer service, document processing, coding, quality inspection, reporting, or administrative work. Also includes cases where AI is linked to reductions in labour costs, staffing needs, hours worked, manual work, outsourcing needs, or hiring. Do not code from unspecified cost reductions or generic efficiency claims. "AI reduces costs" is insufficient. "AI handles first-line customer queries" or "AI processes invoices automatically" is sufficient—it describes a specific task previously requiring human labour.

`LABOUR_TASK_AUGMENTATION`

AI helps employees perform existing tasks better, faster, or with higher accuracy, while humans remain responsible for the task, e.g. decision support tools, recommendations, copilots, forecasting tools, or AI-assisted analysis. The key criterion is that a human remains in the loop and retains responsibility for the outcome. When unclear whether humans remain in the loop, default to this channel unless the report explicitly mentions reduced headcount, hours, or outsourcing as a result of AI. Note: some use cases (e.g. forecasting) may qualify for both `LABOUR_TASK_AUTOMATION` and `LABOUR_TASK_AUGMENTATION` depending on whether the report describes AI as replacing an analyst role or as a tool used by analysts.

`NEW_LABOUR_TASKS`

AI creates new tasks, roles, workflows, or capabilities for employees that did not previously exist, e.g. prompt engineering, AI monitoring, model training, data labelling, AI governance, or new human tasks made possible by AI systems. Code only when the report provides evidence that the firm is actively deploying people in these new roles—not merely mentioning them as general possibilities or future ambitions. Generic references to hiring data scientists or building AI teams are sufficient if described as currently underway.

`NON_LABOUR_INPUT_EFFICIENCY`

AI makes other inputs more productive, efficient, or reliable, e.g. capital equipment, energy, materials, or purchased services. Includes cases where AI reduces waste, downtime, or cost of these inputs. A common example is predictive maintenance, where AI reduces equipment downtime or extends asset lifetimes by anticipating failures before they occur. Do not code from generic operational efficiency claims unless the report describes a specific non-labour input that AI acts on.

`EXISTING_PRODUCT_QUALITY`

AI improves the quality, reliability, personalisation, speed, consistency, or functionality of products or services that the company already offered. If personalisation is described as a genuinely new capability the product previously lacked entirely, consider `NEW_PRODUCT_OR_SERVICE` instead. If AI improves an existing product along dimensions it already had—faster, more accurate, more consistent, better tailored—code `EXISTING_PRODUCT_QUALITY`. When ambiguous, default to this channel unless the report explicitly frames the offering as new.

`NEW_PRODUCT_OR_SERVICE`

AI enables a product or service that the company did not previously offer, including cases where AI is an essential feature of the new offering. The key distinction from `EXISTING_PRODUCT_QUALITY`: does AI improve something existing, or enable something new? If the product would not exist without AI, or if AI defines what the offering fundamentally is, code this channel. When ambiguous, default to `EXISTING_PRODUCT_QUALITY` unless the report explicitly frames the offering as new. Note: will often co-occur with `ai_artifact_type` = `NEW_PRODUCT_OR_SERVICE`.

`RD_AND_INNOVATION`

AI makes research, product development, experimentation, design, testing, or innovation processes faster, cheaper, more accurate, or more effective. This includes drug discovery, product design, simulations, software development, prototyping, testing, literature screening, hypothesis generation, or patent landscape analysis. Code even if the report does not yet mention a specific new product—investment in AI-enabled R&D capacity is sufficient. Do not code from generic references to innovation or digital transformation unless a specific R&D or development process is described.

`SCALABILITY`

AI makes it easier or cheaper to scale operations by lowering the costs of communication, monitoring, coordination, or internal knowledge transfer, e.g. standardising processes across subsidiaries, replicating best practices across units, or enabling firm-wide monitoring without proportional headcount increases. Do not code from generic growth or transformation language. “AI enables the company to grow” is insufficient. “AI allows the same compliance monitoring to be applied across all 12 subsidiaries without additional staff” is sufficient.

8. GENERAL RULES

– `target_domain`: select multiple values if applicable.

– All reasoning and `evidence_text` fields: cite specific words from the text.
  Do not introduce information absent from the chunk.
  Weak or indirect signals must be flagged, not papered over.
– Empty lists `[]` for missing list fields.
– Float scores `confidence_score` and `text_clarity_score` must be in [0.0, 1.0].

9. OUTPUT FORMAT (CRITICAL)

Respond with **valid JSON only**—no markdown fences, no preamble, no trailing text. The number of items in `analyses` must EXACTLY match the number of `<chunk>` elements in the input.

Required keys on every Initiative:

`name`, `description`, `adoption_stage`, `adoption_stage_confidence`,
`target_domain`, `ai_artifact_type`, `artifact_name`,
`initiative_scope`, `ai_deployment_mode`, `scope_and_mode_reasoning`,
`business_value_drivers`, `productivity_channels`

PROHIBITED fields—do NOT add any of these:

`document_metadata`, `ai_evidence_found`, `ai_evidence_items`, `chunk_id`,
`evidence_type`,
`text_summary`, `contextual_mentions`, `evidence`, `summary`,
`initiative_scope_reasoning`, `ai_deployment_mode_reasoning`
(the last two are merged into `scope_and_mode_reasoning`).
`intensity_score`, `evidence_explicitness`—renamed in v7 to
`classification_confidence` and `explicitness_score`.
`workforce_impacts`—commented out in v7.
`reasoning`—removed in v8 (ChunkAnalysis level).
`adoption_stage_reasoning`—removed in v8 (Initiative level).

Required skeleton—follow this structure EXACTLY. One ChunkAnalysis per chunk; one example initiative shown with all required fields:

```
{{
  \"chunk_index\": \"01/03\",
  \"chunk_has_initiatives\": true,
  \"text_clarity_score\": 0.75,
  \"confidence_score\": 0.8,
  \"initiatives\": [
    {{
```

```
      \"name\": \"Example Initiative Name or null\",
      \"description\": \"What the initiative does - drawn from the text.\",
      \"adoption_stage\": \"ACTIVE_USE\",
      \"adoption_stage_confidence\": 0.9,
      \"target_domain\": [\"IT_AND_CYBERSECURITY\"],
      \"ai_artifact_type\": \"INTERNAL_PROCESS_TOOL\",
      \"artifact_name\": null,
      \"initiative_scope\": \"INTERNAL_OPERATIONS\",
      \"ai_deployment_mode\": \"PROCESS_TOOL\",
      \"scope_and_mode_reasoning\": \"Cite the specific words here.\",
      \"business_value_drivers\": [
        {{
          \"driver\": \"EFFICIENCY_AND_COST_SAVINGS\",
          \"classification_confidence\": 0.85,
          \"explicitness_score\": 0.7,
          \"evidence_text\": \"Cite specific words from the text here.\"
        }}
      ],
      \"productivity_channels\": [
        {{
          \"channel\": \"LABOUR_TASK_AUTOMATION\",
          \"classification_confidence\": 0.8,
          \"explicitness_score\": 0.6,
          \"evidence_text\": \"Cite specific words from the text here.\"
        }}
      ]
    }}
  ]
}}

"""
```

Note: The formatting has been adapted for readability. The substantive instruction text and classification rules are preserved.

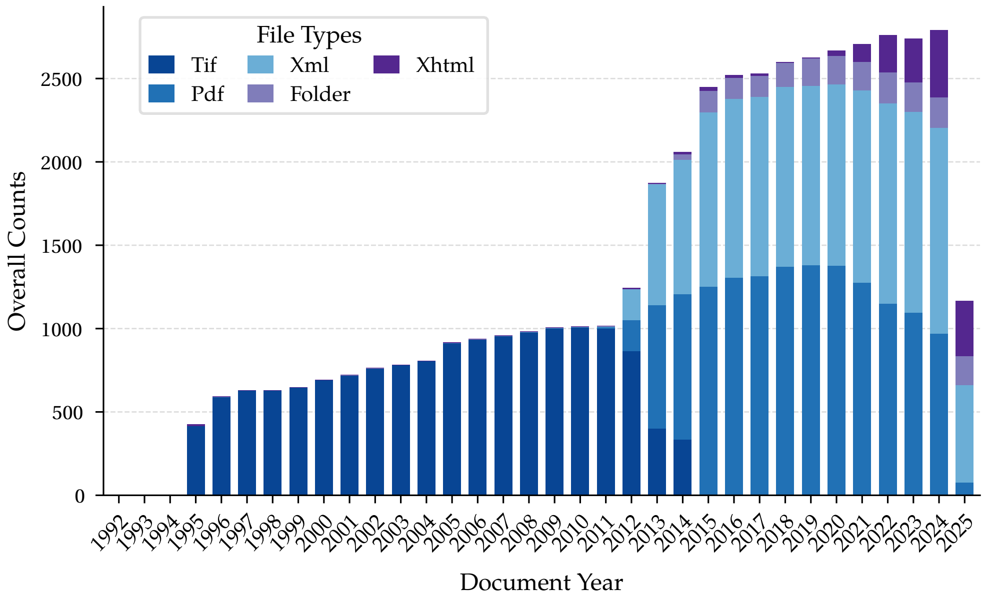
C Keywords, examples, and further descriptions

Table C1: Summary of productivity channels and through which logic they can be described

Category	Channel	Logic
Labor	labor augmentation	decision support, forecasting, planning, pricing, risk assessment, fraud detection, recommendations, analytics, employee assistance, improved accuracy or speed while workers remain involved
	automation	automated processing, workflow automation, document processing, automated customer service, reduced manual work, labor saving, error reduction, faster task completion
	new tasks	AI oversight, model monitoring or evaluation, data curation, prompt design, AI governance, workflow integration, employee training, creation of new AI-related roles or responsibilities
Product	quality of existing products	personalization, improved reliability, accuracy, functionality, speed, service quality, diagnostics, customer experience, consistency, or performance of an existing offering
	new products & services	new AI-enabled products, services, features, platforms, advisory tools, data-driven offerings, business models, or customer-facing solutions
Other productivity	scaling cost	expansion across customers, markets, locations, products, or organizational units; standardized systems; reusable infrastructure; centralized platforms; output growth without proportional increases in labor or other inputs
	non-labor input	predictive maintenance, energy optimization, materials savings, inventory optimization, logistics planning, reduced downtime, improved use of equipment, capital, energy, or purchased inputs
	R&D efficiency	faster or cheaper research, experimentation, discovery, design, testing, prototyping, drug development, product development, or generation and evaluation of new ideas

Note: This table summarizes the interpretation of the productivity-channel taxonomy and the logic behind each channel.

Figure C1: Different formats and their share over time



Notes: This figure shows the number of annual-report source files by file format and document year. The source material includes TIFF image files, PDF documents, XML files, folders containing multiple report components, and XHTML files. The changing composition reflects both developments in firms' filing practices and changes in the formats made available through the Danish Business Authority's reporting platform. Counts refer to source files rather than unique firm-year reports, and a single report may contain multiple files. Reports for 2025 are incomplete at the time of data collection (May 2026).

Table C2: Keywords and their respective counts

Keyword	Count
ai	17642
artificialintelligence	3534
machinelearning	1921
kunstigintelligens	1192
generativeai	1009
genai	959
ki	809
al	661
datascience	362
chatbots	171
chatbot	143
maskinlæring	125
datamining	117
cognitivecomputing	116
computervision	110
deeplearning	107
kl	79
naturallanguageprocessing	78
neuralnet	78
largelanguagemodels	70
deepfake	46
generativai	39
anomalydetection	28
gai	24
largelanguagemodel	18
billedgenkendelse	14
sprogmodeller	14
kognitivcomputing	12
reinforcementlearning	10
aietik	9
sprogmodel	5
artificialneuralnetwork	4
knowledgegraph	4
computeralgoritmer	3
datavidenskab	2
convolutionalneuralnetwork	1

Note: This table shows the full list of AI-related keywords and their counts in the full sample. While processing the text, all spaces have been removed.

Table C3: Industry category labels

Original category	Short label used in figures
Accommodation facilities and restaurant business	Accommodation & food
Administrative services and support services	Administrative support
Banking and financial services, insurance	Banking, finance & insurance
Building and construction business	Construction
Culture, amusements and sports	Culture, leisure & sports
Education	Education
Electricity, gas and district heating supply	Energy supply
Health care and social measures	Health & social services
Information and communication	Information & communication
Liberal, scientific and technical services	Professional, scientific & technical
Manufacturing business	Manufacturing
Other services	Other services
Raw material extraction	Raw material extraction
Real estate	Real estate
Water supply; sewage system, waste management and cleaning of soil and groundwater	Water, waste & remediation
Wholesale and retail trade; repair of motor vehicles and motorcycles	Wholesale & retail trade

Note: Short labels are used only for figure readability. The original industry category definitions remain unchanged.

Table C4: Examples of explicit non-use statements within annual reports

Quotes from the annual reports documenting AI-non-use for personal data processing
Special personal data that reveals racial or ethnic origin, political opinions, religious or philosophical beliefs, trade union membership, genetic data, biometric data, or data concerning health or revealing a person's sexual activity or orientation will in no event be subject to AI or automated decision-making.
Vi bruger heller ikke maskinlæring, AI eller lignende til at afdække forbrugsmønstre hos kunder, medarbejdere eller andre privatpersoner.
...nor do we use machine learning, AI or similar to profile customers, employees or other private individuals.
Colas Danmark A/S, with the exception of Matomo Analytics, which replaced Google Analytics in 2022, does not use algorithms for data analytical purposes, artificial intelligence or automated processes in respect of the rights and possibilities of individuals.
Furthermore, we do not use machine learning, artificial intelligence, or other similar technologies to forecast or assess the consumption behaviors of customers, employees, or other individuals
It does not use AI on personal data, including but not limited to data related to employees, contractors, job applicants, visitors or customers.
Pharma Nord anvender ikke overfor kunder eller medarbejdere algoritmer eller automatiske afgørelser (AI), som kan være dataetisk problematiske
Selskabet anvender ikke avancerede teknologier som kunstig intelligens eller 'machine learning' til projekter der behandler personlige eller i øvrigt følsomme informationer.
Vi anvender ikke AI og automatiseret beslutningstagning der involverer persondata.
Vi handler ikke med persondata og bruger ikke kunstig intelligens eller tilsvarende til at danne profiler af potentielle kunder, medarbejdere eller andre.

Note: This table shows example statements from the annual reports justifying in which area and why AI is explicitly not used.

Table C5: Adoption stages

Adoption stage	Definition
Active use	current use, deployment, embedment, applications in firm's activities, systems, products, or processes
Pilot or testing	tests, pilots, trials,evaluations through proof-of-concept, beta, or experimentation activities
Planning or exploration	future possibilities, planned investments, strategic exploration, areas under consideration, without clear evidence of current use
Research and expansion	ongoing research, development of new applications, expansion of existing capabilities, establishment of related R&D activities
Talent and culture	employee training, hiring, skills development, board or staff capabilities, development of an AI-oriented organizational culture
Ecosystem and partnerships	partnerships, collaborations, sponsorships, public initiatives, external networks, participation in broader AI ecosystems
Awards, recognition, and IP	awards, recognition, patents, intellectual property, other signals of technological achievement
Explicit non-use	explicit statements of not used, not applied, or excluded from a specific activity or domain

Note: This table defines the different disclosed adoption stages found in AI-related annual reports. Adoption stages describe the stage within a company at which an AI-related initiative or an explicit non-use statement is made. A single passage may be assigned to more than one adoption stage where the text refers to multiple areas.

Table C6: Target-domain categories

Target domain	Definition
Enterprise-wide	general company strategy, broad transformation, activities across the firm
Multiple domains	two or more categories have been associated with one initiative
Research & Product	R&D, product development, innovation, new product/service design, prototyping
Operations and manufacturing	internal operations, production, logistics, supply chains, manufacturing, process optimization
Customer & Marketing	customer service, sales, marketing, advertising, customer experience, branding, communications
IT & Cybersecurity	IT infrastructure, data security, network protection, threat detection, system administration
Legal, ethics, and governance	compliance, audit, corporate governance, ethical rules, responsible AI, legal constraints
Finance & Accounting	financial planning, accounting, budgeting, reporting, auditing, investment, risk management
HR & Workforce	recruitment, employee management, training, workforce planning, talent development
Personal data	processing, protection, restriction, non-use of personal data
Other	initiatives not fitting the above categories
Unclear	insufficient information to determine domain

Note: This table defines the different target domain categories found in AI-related annual reports. Target domains describe the organizational area or functional context to which an AI-related initiative or explicit non-use statement refers. A single passage may be assigned to more than one target domain where the text refers to multiple areas.

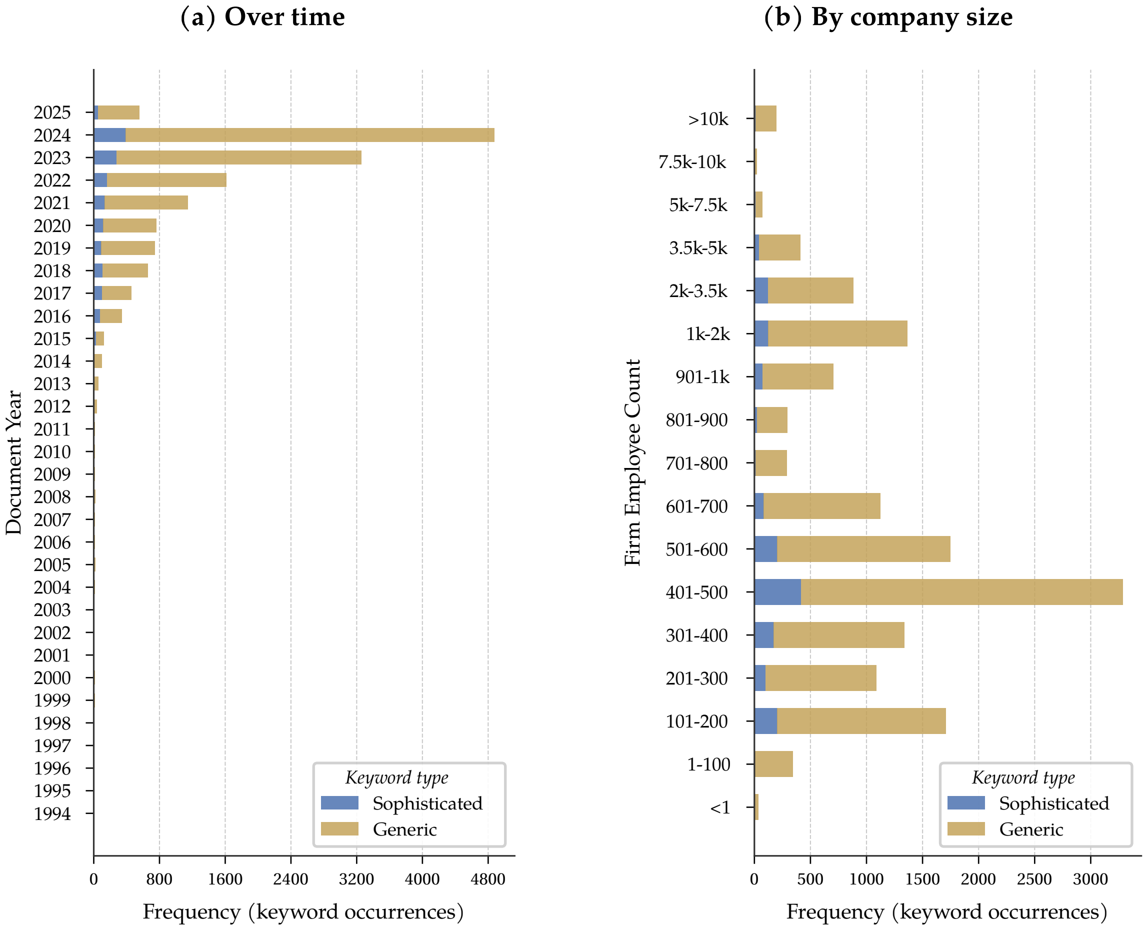
D Generic vs sophisticated keywords

We are analyzing a related question: *how* firms talk about AI once it appears in annual reports. Specifically, Figure D1 distinguishes between generic AI-related keywords (e.g., "AI", or "artificial intelligence") and more sophisticated keywords (e.g., "machine learning" or "generative AI").

Panel (a) in Figure D1 shows that the use of AI-related keywords increases sharply over time, especially from 2023 onward. However, this increase is driven primarily by generic rather than sophisticated terminology. Sophisticated keywords also become more common, but remain much less frequent than generic ones throughout the period. This suggests that firms increasingly discuss AI in annual reports, but typically do so in broad, non-technical terms. Panel (b) in Figure D1 shows a similar pattern across firm-size groups. Generic keywords dominate across all size categories, even though the total number of keyword mentions varies across groups. This indicates that the tendency to describe AI in broad terms is widespread rather than concentrated in a single part of the firm-size distribution.

Thus, Figure D1 suggests that AI-related annual-report narratives are expanding not only in frequency but also in breadth. Yet this expansion is characterized more by the spread of generic AI language than by the use of specific and technically differentiated terminology. This is consistent with the broader interpretation of annual reports as strategic communication documents in which AI is often framed at a relatively general level.

Figure D1: AI-related keywords: generic vs sophisticated



Note: This figure shows AI-related keywords distinguished by generic (e.g., “AI”, “KI”) and sophisticated (e.g., “machine learning”, “generative AI”) keywords. The left panel displays the prevalence of generic versus sophisticated keyword usage over time. The right panel shows the corresponding distribution across firm-size groups. The unit is validated AI-related chunks, while the denominator is all chunks containing keyword matches in the private-sector baseline sample.

E Comparison to other measures

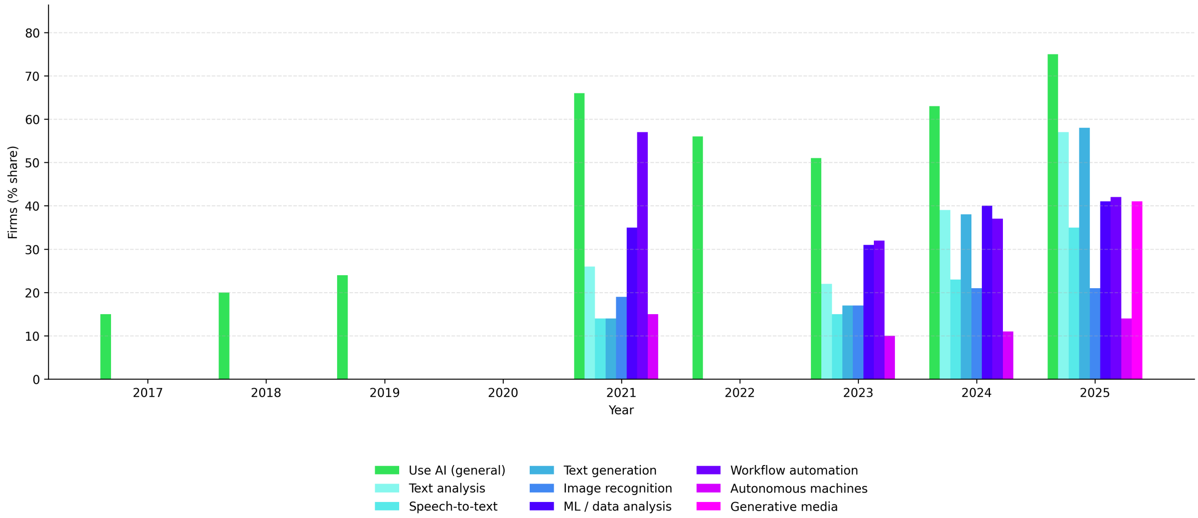
The measure developed in this paper captures AI as it appears in firms' annual report narratives. It is therefore different from standard measures of AI adoption. Existing empirical measures of AI use typically rely on surveys, technology investment data, job postings, patent data, web-scraped technology indicators, or case-study evidence. Recent AI-monitoring and adoption studies provide useful benchmarks for this type of measurement exercise, but they typically focus on observed or reported adoption rather than on the narrative framing of AI in corporate disclosures (e.g., Aldasoro et al. 2026; Allen 2026; OECD 2024). Each of these measures captures a different dimension of AI diffusion. For instance, survey data are well-suited to measuring whether firms report using AI and, in some cases, for which functions. Investment or technology-spending data can indicate whether firms allocate resources to digital or AI-related capital. Job postings may reveal demand for AI-related skills, while patent data captures inventive activity. However, these measures usually say less about how firms publicly frame AI, whether they connect it to strategy, and through which productivity-relevant mechanisms they expect AI to matter. Annual reports, therefore, provide a complementary measure rather than a substitute for existing adoption measures. They reveal which AI-related activities firms choose to make visible, how they describe them, and whether they connect them to operational efficiency, decision-making, innovation, or organizational transformation. This interpretation is consistent with the conceptual framework, which treats AI narratives as qualitative signals of expected productivity mechanisms rather than as direct evidence of productivity outcomes.

A useful benchmark for the annual-report measure is provided by Statistics Denmark's survey-based data on firms' use of artificial intelligence (Statistics Denmark 2026). These data have two advantages for the present purpose. First, they are based on a standardized survey instrument and therefore provide an external reference point for the level and development of AI adoption among Danish firms. Second, the survey distinguishes between different types of AI use and perceived areas of effect. This makes it possible to compare our narrative evidence from annual reports with a more conventional adoption-based measure. This comparison follows recent efforts to monitor AI diffusion through surveys and other adoption indicators, while emphasizing that such measures capture a different margin than annual-report narratives (e.g., Allen 2026; OECD 2024).

Figure E1 shows the share of Danish firms with at least 250 employees that report using AI separately by usage areas. The figure shows a clear increase in reported AI use over time. Overall, AI use is low in the earlier years but rises substantially after 2021, especially between 2023 and 2025. By 2025, the overall share of firms reporting AI use is above 75 percent. The specific usage areas also increase over time, but with noticeable differences across applications.[9] Text analysis, text generation, and speech-to-text are among the more widespread applications in recent years, while image recognition, machine learning/data analysis, workflow automation, autonomous machines, and generative media are less evenly distributed. The figure therefore suggests both broad diffusion and substantial heterogeneity in the types of AI applications firms disclose.

[9] It has to be noted that in 2022, the question about the company's use of AI is formulated differently compared to 2021, which has resulted in a lower share. Additionally, the media-generation subcategory was added to the survey only in 2025. For more information, see Statistics Denmark (2026).

Figure E1: Use of artificial intelligence, by usage area, large firms (≥ 250 employees)



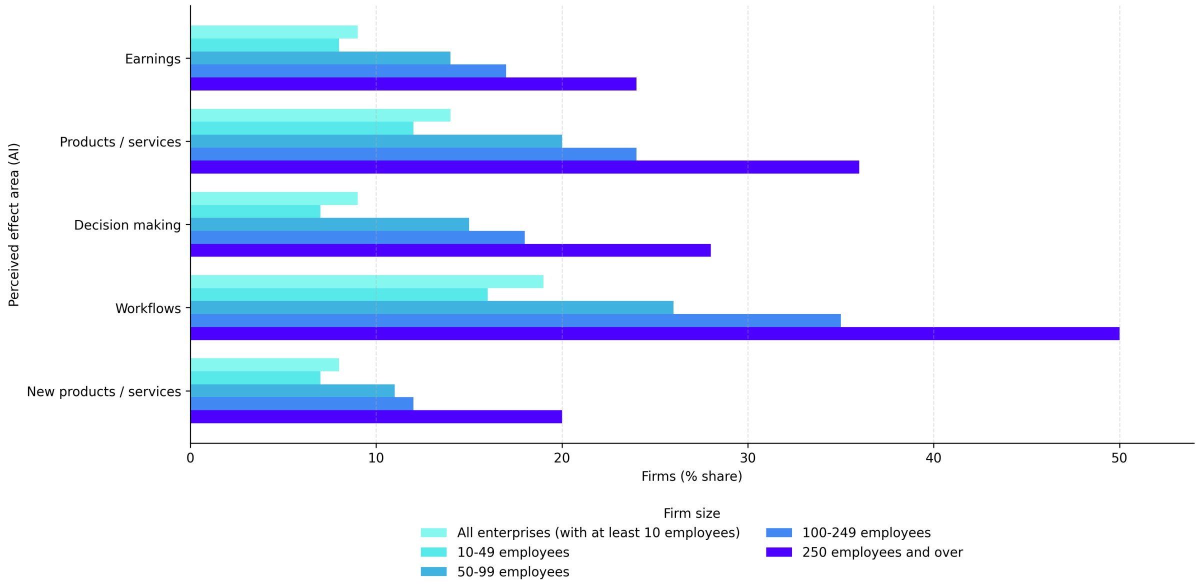
Note: This figure shows the share of firms using artificial intelligence by usage area for all firms with at least 250 employees. The underlying data is obtained from Statistics Denmark (2026).

This benchmark is useful for interpreting evidence in the annual report. If the annual-report measure increases over the same period as survey-based AI use, this supports the interpretation that the rise in AI narratives reflects a real increase in the salience and diffusion of AI among Danish firms. At the same time, the two measures should not be expected to coincide mechanically. The measure by Statistics Denmark (2026) reports on use among firms, whereas our annual-report measure captures whether AI is sufficiently relevant, strategic, or material to be included in public corporate disclosure. A firm may report AI use in a survey without discussing AI in its annual report. Conversely, a firm may discuss AI in an annual report as part of market context, planning, risk, or governance without reporting intensive operational use. The comparison should therefore be interpreted as external validation of broad trends rather than as a one-to-one validation of individual firm-level adoption.

Figure E2 provides a second and more conceptually important benchmark. It shows firms' perceived effects of AI in 2024, broken down by firm size and area of effect. The size gradient is pronounced: larger firms report perceived effects more frequently than smaller firms across all categories. For firms with 250 employees or more, the most frequently reported perceived effect concerns workflows, followed by products and services, decision-making, earnings, and new products or services. For all firms with at least 10 employees, the corresponding shares are substantially lower. This pattern is consistent with the view that AI use and expected AI-related gains are more common among larger firms. This is due to complementary investments in data infrastructure, skills, and organizational capabilities being more likely to be available.

The categories in Figure E2 are particularly useful because they closely correspond to the productivity channels used in the Section 2. Perceived effects on workflows map directly onto the channel of process automation and efficiency. Effects on decision-making correspond to the channel of decision-making and management. Effects on products and services, as well as new products and services, relate to innovation and product development. Earnings are not a productivity channel in the same narrow sense, but rather an outcome that may reflect the combined effects of efficiency, better decision-making, innovation, and scaling. The survey, therefore, provides external support for the idea that firms themselves understand AI through mechanisms that are close to our productivity-channel framework. Thus, the channels are not only theoretically motivated but also align with the way firms report perceived AI effects in survey data.

Figure E2: Perceived effect of AI in 2024, by productivity area and firm size



Note: This figure shows the perceived effects of AI by productivity area and firm size. The underlying data is obtained from Statistics Denmark (2026).

F Productivity channels

Table F1: Productivity channels at the firm-year level

Productivity channel	Count	Share
Labor task augmentation	497	20.71%
Labor task automation	324	13.50%
New labor task	114	4.75%
Existing product quality	465	19.38%
New product or service	322	13.42%
Non-labor input	165	6.88%
R&D and Innovation	268	11.17%
Scalability	245	10.21%

Note: This table shows the distribution of productivity channels at the firm and year level. A single firm and year pair may be associated with more than one productivity channel. Counts therefore refer to channel assignments rather than mutually exclusive initiatives. The denominator used to calculate the shares is the total number of productivity-channel assignments.

Table F2: **Examples and reasoning for the productivity channel "Labor task automation"**

Score	Example	Reasoning
0.3	"til de kunder, som er særligt udfordrede på de digitale platforme, eller hvor svaret ikke kan findes via vores selvbetjeningsløsninger, skal vi naturligvis fortsat kunne tilbyde kvalificeret personlig betjening"	suggests AI search reduces the volume of cases requiring human intervention, but this is indirect
0.4	"chatbot technology"	chatbots typically automate customer query handling previously requiring human agents, though the text does not explicitly state this outcome
0.5	"technologies, like chatbots or virtual agents, which users can talk to"	implies automation of customer interaction tasks previously requiring human agents
0.6	"enabled real-time fraud surveillance"	implies automated monitoring replacing or reducing manual surveillance tasks
0.7	"Fast-tracked automation is supporting enterprises in achieving empathy-led automation of their customer care modules"	AI automates customer care tasks previously requiring human agents
0.8	"ArtEngine removes the burden of minutiae involved in material creation work"	such as photoconversion to physically based rendered materials, resolution enhancement, deblur, seam removal, unwarping, and color matching
0.9	"machine learning solution developed by DXC to automate medical coding"	explicitly describes AI replacing human work in a specific task
1.0	"automated the review of all claims using Machine Learning... Natural Language Processing... and a scenario-based modelling approach', 'could automate 91% of claims' approval"	the solution directly replaces manual human adjudication for the majority of claims

Table F3: Examples and reasoning for the productivity channel "Existing product quality"

Score	Example	Reasoning
0.2	'Transforming Telecommunications with GenAI'	implies improvement to existing telecom services, but no specific mechanism is described
0.3	levere endnu mere vÃ ¦ rdi til vores klienter	framed as delivering more value through legaltech tools, suggesting improvement of existing legal service quality, though indirect
0.4	'as more intelligent industrial processes are developed, the focus has shifted to the level of information and material flows that depend increasingly on fuzzy logic and neural networks'	implies AI enhances the intelligence and capability of existing industrial process solutions, though the link is indirect
0.5	'kognitive procesomrÃ¥der, hvor kunstig intelligens spiller en central rolle'	AI enhances IBM's existing service offerings in automation, security, and multi-cloud
0.6	'streamlined their business processes, enabling them to reduce energy consumption and enhance indoor air quality with their innovative ventilation products'	AI (as part of Data & AI focus) improves the quality and sustainability performance of WindowMaster's existing ventilation products and business processes
0.68	Linux and other platforms that are enabled with enterprise AI and are hybrid cloud ready...	a security-rich, high-performance enterprise operating system
0.8	deep insights powered by machine learning	empower our customers to create, optimize and deliver personalized experiences for their end-users, increasing customer retention and engagement rates
0.9	'making arrival times 50% faster during peak travel periods'	AI improves speed and efficiency of existing elevator dispatching service

Table F4: Examples and reasoning for the productivity channel "Labor task augmentation"

Score	Example	Reasoning
0.1	Very indirect: 'anvender kun i meget begrænset omfang teknologier såsom artificial intelligence, machine learning eller algoritmer i forbindelse med markedsføringen'	implies some form of AI-assisted marketing activity, but no specific task or mechanism is described. Signal is weak
0.2	The Company uses Artificial Intelligence (AI) only on technical data, such as operational- or subsurface data from energy operations	Augmentation of technical analysis is implied but no specific task or human-in-the-loop detail is provided
0.3	utilises advanced technologies, such as AI and automation, supported by active oversight from the Executive Board and the Audit Committee	oversight language implies humans remain in the loop, but the specific tasks AI augments are not described
0.4	'predict optimal sailing speeds'	implies vessel operators or navigators use the predictions to adjust speed
0.5	expanding our capabilities beyond the lab... use data science and artificial intelligence to discover new targets and biomarkers"	implies AI assists researchers in discovery tasks while humans remain responsible for scientific decisions
0.6	'generative AI both as a business opportunity and a working tool, how we approach and apply GenAI in our work is important'	the tool is described as a working tool used by professionals, implying augmentation of existing tasks. Humans remain responsible ('ethical generative AI behaviour in client work')
0.7	'new GenAI-powered tools that greatly improve the productivity of our developers'	developers remain in the loop
0.8	kan herved gøre ledere proaktivt opmærksomme på medarbejdere der bør forfremmes	managers remain responsible for the promotion decision, with AI providing proactive recommendations
0.9	'The Quality Developer always makes their own assessment, but with support from the AI assistant, the process becomes both easier and more efficient'	humans remain in the loop and retain responsibility

Table F5: Examples and reasoning for the productivity channel "New labor tasks"

Score	Example	Reasoning
0.3	'established our Advisory Sounding Board'	implies new roles/workflows for employees, but no direct AI link is described. Signal is weak and indirect
0.4	'the use of AI is high on the agenda across the organisation and the efforts will continue through 2025 to increase AI literacy'	implies employees are being prepared for AI-related tasks, but no specific new roles or workflows are described in this chunk
0.5	We are developing skilled, knowledgeable teams with an understanding of the centrality of data to our business... there were 10,000 course enrollments through the WPP AI Academy	implies active deployment of people into AI-related capabilities, though specific new roles are not named
0.6	'we had the opportunity to attend the very first global generative AI learning month with sessions on harnessing the benefits whilst ensuring ethical use. These sessions were available to all our professionals.'	active deployment of people in new AI-related learning and governance tasks
0.7	'we are currently setting up a dedicated GenAI Board to approve use cases in the future'	this describes the active creation of a new governance role/body for AI oversight, representing new tasks and workflows for employees
0.8	'adding modules such as advanced AI training to help our people build essential skills in areas like prompt engineering and apply AI in practical, real-world scenarios'	actively deploying people in new AI-related tasks and roles

Table F6: Examples and reasoning for the productivity channel "New product or service"

Score	Example	Reasoning
0.2	AI is listed as a named service offering ('BI, AI, IoT, CRM, IT Development, SAAS Solutions') in the service catalogue	no description of what the AI service entails is provided, making this a very weak signal
0.3	'new-age ones like Cloud and GenAI' as service lines	implies GenAI is offered as a service to customers, but the passage is a governance biographical context with no operational detail about the GenAI offering itself
0.4	'SAP is involved in driving innovation in all fields of the digital economy, such as the Internet of Things, machine learning, and artificial intelligence'	implies AI-enabled offerings, but no specific new product is named in this chunk
0.5	'technology innovation in areas as diverse as creative AI, campaign optimization and market simulation for clients'	framed as new innovation delivered to clients, though specifics are limited
0.6	Digitization and AI are driving rapid innovation in the payments market, paving the way for new solutions	Automated cash management solutions and smart integrated payment solutions are becoming more common
0.7	'TCS' digital innovation lab has built several digital applications and solutions for multiple industries, which are being implemented by customers'	describes new AI/analytics-enabled products being deployed
0.8	'the first commercially available cognitive computing platform' delivered through 'cloud environments and as-a-Service models'	'the first commercially available cognitive computing platform' delivered through 'cloud environments and as-a-Service models'
0.9	It-jobbank har således lanceret en AI Apply Helper	a newly launched AI-powered product feature that did not previously exist

Table F7: Examples and reasoning for the productivity channel "Non-labor input efficiency"

Score	Example	Reasoning
0.3	'Shortening production lead times', 'Preventing social infrastructure accidents'	implies AI acts on physical/capital assets to improve their efficiency or reliability
0.4	'preventive maintenance with artificial intelligence'	implies AI acts on equipment/assets to prevent failures, though the mechanism is not explicitly detailed in this chunk
0.5	AI driven tech debt reduction	data resilient solutions across Multi/Hybrid IT systems
0.6	'spare pÃ¥ energiforbruget og mindske dyre nedbrud'	AI acts on energy and equipment inputs to reduce waste and downtime
0.7	'substantially reduce doses while maintaining high levels of yield' and 'optimize the volumes of plant protection products used'	AI directly reduces consumption of plant protection products (a non-labour input) for the end customer
0.8	leverages AI to simultaneously improve performance while minimizing cost across a client's hybrid cloud environment	leverages AI to simultaneously improve performance while minimizing cost across a client's hybrid cloud environment.

Table F8: Examples and reasoning for the productivity channel "R&D and innovation"

Score	Example	Reasoning
0.3	'new and inventive technology developed in whole or in part by relying on AI tools'	implies AI contributes to R&D or technology development processes, though the passage is a risk disclosure and does not describe the mechanism in detail
0.4	analytics ecosystem...integrate data, allowing for a fast and transparent decision making process"	in the context of monitoring preclinical programs and clinical trials
0.5	'data science and artificial intelligence (AI) Key to accelerating development'	describes AI as accelerating the drug development process, a core R&D activity
0.6	Genmab's vision to transform cancer treatment has driven its passionate, innovative and collaborative teams to invent next-generation antibody technology platforms and leverage translational research and data science, fueling multiple differentiated cancer treatments that make an impact on people's lives	leverage translational research and data science, fueling multiple differentiated cancer treatments
0.7	IBM Research continues to advance these technologies, for example integrating neural and symbolic techniques to build AI that can perform more complex tasks by understanding and reasoning more like humans	describing how IBM uses these technologies in its Cloud Paks, software solutions, and research initiatives.
0.8	'PhD research explores the application of artificial intelligence (AI) in enhancing site analysis for regenerative urban design'	AI is being applied to make a specific R&D and design analysis process more effective, conducted in partnership with Computational Design team and Ramboll Tech
0.9	our AI models are able to identify novel targets in just days, rather than years as is common for standard drug discovery methods... PIONEER identifies patient-specific tumor mutations, so called neoantigens	Evaxion Biotech explicitly names and describes three distinct proprietary AI models (PIONEER, ObsERV, and AI-DeeP), each tied to a specific application in immune-target discovery and drug development.

Table F9: Examples and reasoning for the productivity channel "Scalability"

Score	Example	Reasoning
0.3	'Centralizing Custodial Operations' implies standardization and consolidation of operations, which aligns with scalability through AI	though no specific AI mechanism is described
0.4	'Across the four offices on Copenhagen, Aarhus, Odense and Aalborg & top capabilities within Modern Workplace, Data & AI... are now delivered to customers'	implies replication of AI capabilities across multiple locations, though the mechanism is not described in operational detail
0.5	'the world's most powerful, secure and flexible foundation for AI and data-intensive applications and workloads'	implies infrastructure enabling scalable AI deployment across enterprises
0.6	cloud-baserede løsninger sikrer kunderne råderum og fleksibilitet, så de hurtigt kan justere i forhold til ændringer i deres behov uden at binde flere ressourcer end nødvendigt	it frames "kognitiv computing" as a response to the need for tools that can keep pace with continuously growing data volumes ("stadigt stigende datamængder"), directly tying AI-driven innovation to the company's ability to scale its analytical capacity as demand increases.
0.68	'Watsonx and our Consulting capabilities to deploy AI at-scale are a powerful combination'	explicit reference to deploying AI at scale across client organizations
0.8	'Since it's impossible to cover all the stores across the country'	explicitly acknowledges that AI/ML enables scale that human effort alone cannot achieve

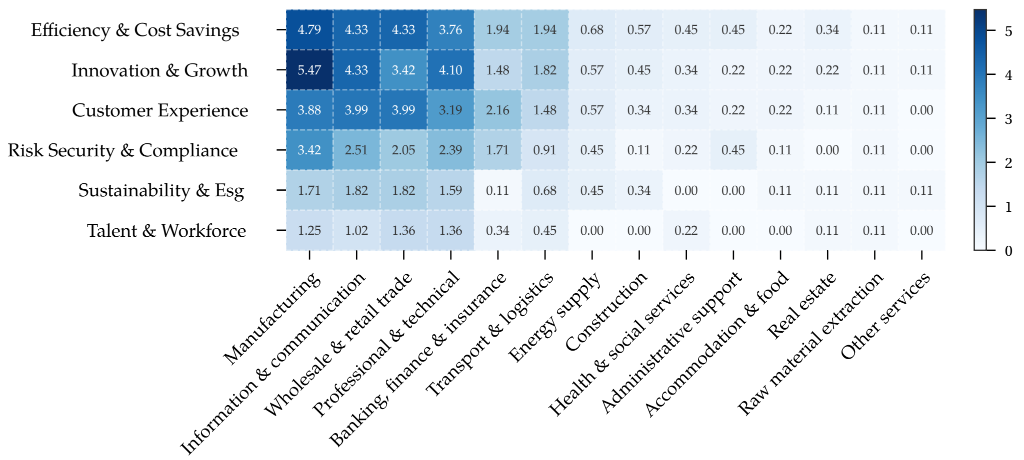
G Robustness checks

Table G1: Data-driven productivity drivers at the firm-year level

Productivity driver	Logic	Count	Share
Efficiency, cost savings	process improvement, cost reduction, error reduction, speed, workflow automation	554	20.1%
Customer experience	personalization, customer support, service quality, engagement, recommendations, customer analytics	518	18.8%
Innovation	new products and services, R&D, experimentation, discovery, product development, and data-driven business models	532	19.3%
Growth	market expansion, revenue growth, scaling, new customer segments, competitive advantage	441	16.0%
Risk, security, compliance	risk identification and monitoring, fraud detection, cybersecurity, legal compliance, responsible AI, governance, data protection, and regulatory reporting	356	12.9%
Sustainability, ESG	sustainability reporting, emissions monitoring, resource optimization, ESG compliance, environmental risk assessment	182	6.6%
Talent, workforce	recruitment, workforce planning, training, employee support, HR analytics, changes in tasks and skills	171	6.2%

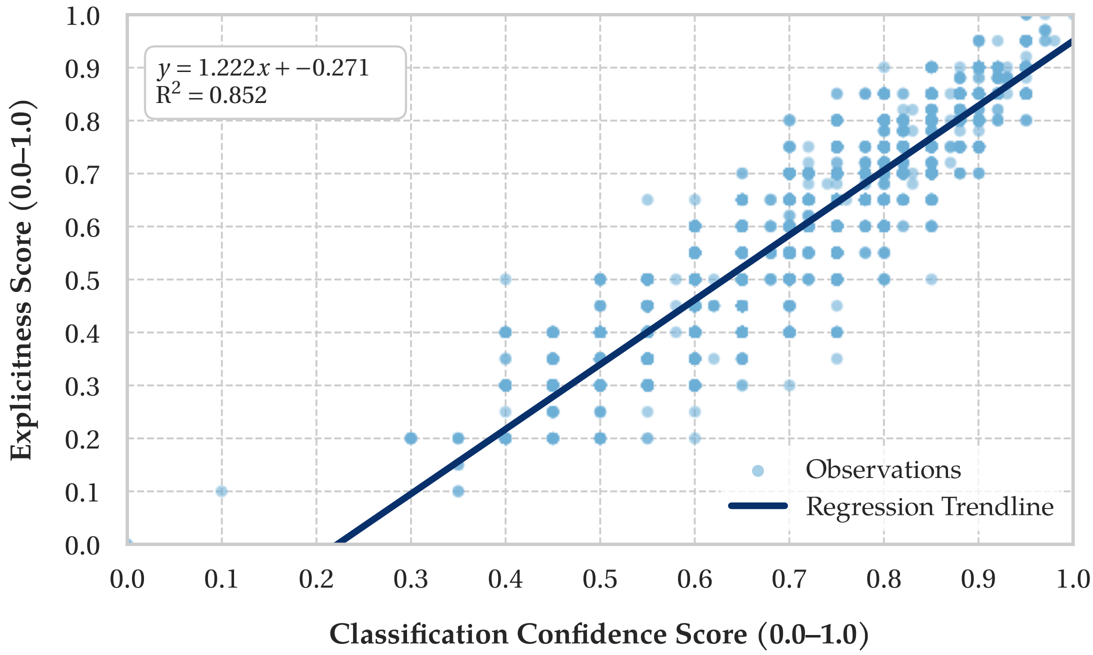
Note: This table shows the data-driven productivity drivers identified in firms' AI-related narratives, along with their logic, counts, and shares. A single AI-related passage may be associated with more than one productivity driver because the classification allows multiple labels. Counts therefore refer to driver assignments rather than mutually exclusive passages. The denominator used to calculate the shares is the total number of productivity-driver assignments.

Figure G1: Productivity drivers by industries



Note: This figure shows the share of AI-reporting firms by industry and data-driven productivity driver. Cell values indicate the percentage of all AI-reporting firms associated with each industry- productivity-driver combination. Darker colors indicate higher shares. Industries are on the x-axis, and productivity drivers on the y-axis. Categories are not mutually exclusive.

Figure G2: Confidence vs. explicitness score



Note: This figure plots the classification-confidence score against the explicitness score for productivity-channel assignments. Each point represents a classified productivity-channel observation. The confidence score captures the model's certainty that the assigned channel is appropriate, whereas the explicitness score captures how directly the underlying passage states the productivity mechanism. The solid line shows the fitted linear relationship. The positive association and an R^2 of 0.852 indicate that more explicit statements are generally classified with greater confidence, although the two measures remain conceptually distinct.