Scientometrics (2026) 131:2333–2372
https://doi.org/10.1007/s11192-025-05468-y

Detecting AI adoption at scale: a web mining and LLM methodology

Ana Pastor-Merino[1] · **Xavier Martínez-Barbero**[1] · **Josep Domenech**[1]

Received: 2 May 2025 / Accepted: 24 October 2025 / Published online: 25 November 2025

Abstract
The recent emergence of Artificial Intelligence (AI) technologies has created multiple opportunities for companies. The growing interest in AI adoption has led to an increase in technological innovations and business applications. At the same time, web mining techniques have been used to monitor technological trends, often through basic natural language processing methods. However, advances in Large Language Models (LLMs) enable a more detailed and in-depth analysis of web content, necessitating the development of new methodologies. This paper proposes a methodology for integrating LLMs into the technology mining of corporate websites, with a focus on Spanish companies. The study uses data collected in 2023 and 2025, applying web scraping techniques to gather textual information. The analysis process involves automated AI detection, manual validation, and evaluation through standard classification metrics. The findings contribute to a better understanding of AI adoption trends and inform future research on tech mining applications for industry.

Keywords Artificial intelligence · Large language models · Web scraping · Technology mining · AI adoption

Introduction

The diffusion of Artificial Intelligence (AI) technologies across firms is reshaping how businesses operate, compete, and innovate. As AI applications become more integrated into products, services, and decision-making processes, understanding their adoption is essential for assessing the pace and nature of digital transformation. This is particularly relevant for evaluating firm-level productivity, sectoral dynamics, and broader economic performance (Makridakis, 2017; Tegmark, 2018; Cockburn et al., 2019). The extent and heterogeneity of AI adoption also have implications for policy design, labor market adaptation, and strategic investment (Stornelli et al., 2021).

Recent studies have documented early patterns of AI diffusion using traditional sources such as surveys and administrative data. For example, McElheran et al. (2024) use U.S.

✉ Ana Pastor-Merino
anpasme@doctor.upv.es

[1] Universitat Politècnica de València, Valencia, Spain

census survey data to show that AI adoption varies across firm sizes and geographies. Likewise, Czarnitzki et al. (2023) analyze German innovation survey data to show positive productivity effects linked to AI use. Other studies rely on case studies and interviews to uncover sector-specific drivers and barriers (Yang et al., 2024).

Traditional methods such as surveys, though valuable, often suffer from limitations such as restricted sample sizes, response biases, and reporting lags (Gök et al., 2015). In response to these limitations, recent studies have explored digital traces to measure AI adoption, employing keyword-based Natural Language Processing (NLP) techniques on job postings, scientific publications, or web content. Specifically, focusing on the United States, Alekseeva et al. (2021) analyzed online job vacancies to estimate the demand for AI-related skills, while Babina et al. (2024) tracked changes in workforce composition through resumes and hiring data to infer AI investments. Similarly, Hunt et al. (2024) combined scientific publication data and job ads to assess spatial patterns of AI diffusion.

In the European context, Dahlke et al. (2024) measured AI adoption based on the textual content from company websites in the DACH region. Their approach combines keyword-based feature selection with a supervised transformer-based classifier trained to differentiate between genuine, deep integration of AI and superficial or promotional mentions.

Recent advances in Large Language Models (LLMs) present a complementary path for detecting AI-related content. LLMs can interpret context, infer implicit meaning, and adapt to heterogeneous language use (Brown et al., 2020; Dahlke et al., 2025; Abbasiharofteh et al., 2024). However, the application of LLMs to systematic, large-scale measurement of firm-level AI adoption remains limited and requires methodological validation.

This paper presents a method to identify AI adoption using LLM-based analysis of the textual content of company websites. The approach is applied to a large sample of firms operating in Spain, allowing for scalable and context-aware identification of AI-related activity. Drawing on the innovation definitions outlined in the Oslo Manual (OECD & Eurostat, 2018), the method distinguishes between different forms of AI use and tracks adoption patterns over time. Specifically, we analyze data from 2 years, 2023 and 2025, to assess temporal changes in the prevalence and nature of AI adoption.

The main contributions of this study are fourfold: (i) it proposes a replicable method to detect firm-level AI adoption based on LLM processing of textual data from corporate websites; (ii) it applies the method to Spain, extending empirical research on AI adoption to a large multilingual European setting; (iii) it provides a temporal comparison of AI adoption across two years, capturing evolving firm behavior; and (iv) it contributes to methodological debates on digital footprints indicators by critically evaluating the advantages and limitations of LLM-based approaches.

The remainder of the paper is organized as follows: Sect. 2 presents a review of the relevant literature; Sect. 3 outlines the data and methodology used in the study; Sect. 4 discusses the evaluation of the proposed methodology; Sect. 5 presents the experimental results derived from the analysis; Sect. 6 discusses their implications, addresses limitations, and outlines directions for future research; finally, Sect. 7 summarizes the key findings and conclusions of the study.

Related work

This section reviews three strands of literature relevant to the study. First, it frames AI adoption as a form of innovation, drawing on established conceptual frameworks. Second, it examines the use of company websites as a source of firm-level information, highlighting their value as a digital trace of technological activity. Third, it discusses recent developments in natural language processing, focusing on the application of LLMs for automated classification tasks.

AI Adoption as Innovation

AI refers to the capability of machines to perform tasks requiring human intelligence, such as learning, reasoning, and decision-making (Russell & Norvig, 2016). We define AI adoption as the point at which a firm not only acquires an AI-enabled tool or service but also integrates it into at least one core business workflow and communicates that integration externally. In this work, such communication is operationalized as the discursive presence of AI within company website content: statements indicating that AI technologies are being used, embedded in offerings, or applied internally. This definition reflects the idea that adoption is not solely a technical implementation but also a social process, where technologies must be legitimized and publicly signalled as part of the organisation's identity and value proposition (Battisti & Stoneman, 2003; Ulucanlar et al., 2013; Dahlke et al., 2024).

The Oslo Manual provides a widely accepted framework for classifying innovation into two categories: product and process. Product innovations involve goods or services that are new or significantly improved in their characteristics or intended uses. Process innovations refer to new or significantly improved production or delivery methods (OECD & Eurostat, 2018). Under this framework, AI adoption qualifies as innovation in both dimensions. For instance, companies developing smart devices or autonomous vehicles are engaging in product innovation by embedding AI into their offerings. In contrast, the use of AI such as machine learning models for internal purposes, like automating document classification or enabling predictive maintenance illustrates process innovation. According to the Oslo Manual, the key areas of process innovation include: Production of Goods and Services, Distribution and Logistics, Marketing and Sales, Information and Communication Systems, Administration and Management, and Product Development and Business.

Firms often describe these innovations explicitly on their websites (Gök et al., 2015; Daas & Van Der Doef, 2020). For instance, by noting the integration of AI into platforms, the use of intelligent systems in operations, or the launch of new AI-based services. These statements provide concrete signals that AI has been embedded in the firm's activities and are thus suitable for detection via automated text analysis.

AI adoption often requires changes in the firm, like improving data systems, training staff, and reorganizing workflows (Lee et al., 2022, 2023; Chen & Tajdini, 2024). These shifts can lead to new or improved ways of delivering products or running processes, meeting the Oslo Manual's definition of innovation. AI can involve incremental or radical change (Dewar & Dutton, 1986): the former improves existing products or processes, while the latter introduces entirely new technologies or business models that may disrupt markets (Bresnahan, 2010). For instance, adding facial recognition to a regular surveillance camera is incremental, but creating a fully AI-based product or service is radical.

Given the intangible and fast-evolving nature of AI, measuring its adoption through digital traces becomes increasingly relevant. Website content, by capturing how firms describe their technological capabilities and market offerings, provides a timely and externally visible proxy for innovation (Kinne & Lenz, 2021; Rammer & Es-Sadki, 2023; Bottai et al., 2024). Focusing on these online signals allows for scalable, comparative, and context-aware identification of AI-related activities. The next sections examine how prior studies have mined web content for innovation signals and explore how advances in natural language processing can enhance the reliability and comparability of these measures.

Company website mining for firm characteristics

Company websites have become a convenient source of information for researchers seeking to understand different aspects of a company's profile, including strategic orientation, innovation capacity, and market behavior (Blazquez & Domenech, 2018). Unlike traditional data sources, websites offer greater timeliness, granularity, and coverage. They are frequently updated, providing real-time insights into business activities, technological adoption, and strategic changes that official databases and surveys often fail to capture due to reporting lags (Rammer & Es-Sadki, 2023; Bottai et al., 2024; Crosato et al., 2024). They also capture fine-grained data on small and medium-sized enterprises, which often have fewer reporting obligations and may not be captured adequately by traditional sources (Mazzoni et al., 2024). Consequently, website data complement established indicators such as surveys, patents, and official registries (Héroux-Vaillancourt et al., 2020)

Several studies have examined how company websites can be used to identify innovation-related activities, mainly through keyword-based methods and cross-sectional analyses. Early research focused on the frequency of specific terms: Gök et al. (2015) counted R&D-related terms on UK "green-goods" Small and Medium-sized Enterprise (SME) websites to infer research activity, showing that such lexicons capture downstream innovation not visible in patents or surveys, while Héroux-Vaillancourt et al. (2020) built formative indices for R&D, intellectual property, collaboration and financing by matching carefully curated word lists against Canadian nanotechnology pages. Using Dutch CIS data, Daas and Van Der Doef (2020) trained a text model that still relies on discriminative words and highlighted issues of concept drift and minimum word counts when scaling keyword approaches to the full business population. Subsequent work blended keywords with machine-learning: Axenbeck and Breithaupt (2021) showed that terms such as "develop" together with English-language share and page volume predict product and process innovation in German firms via Random Forests, and Kinne and Lenz (2021) used survey-labelled websites to teach a deep neural network that outperforms static lexicons yet still draws on salient word patterns to flag product innovators nationwide.

More recent studies move beyond pure keyword counts: Dahlke et al. (2024) used large-scale website scraping in the DACH region to identify AI adopters and measure spatial diffusion effects, initially using keywords but later complementing this with transformer-based models to overcome limitations in semantic inference. In contrast, Bottai et al. (2024) showed that structural website features such as HTML tag profiles, rather than textual content, could distinguish innovative from non-innovative Italian manufacturing SMEs, indicating that non-lexical signals may also be informative.

The use of website data for innovation studies can be situated within the broader field of tech mining, which aims to extract insights from large volumes of unstructured data using bibliometric and text mining tools (Zhang et al., 2017).

As NLP continues to evolve, its integration into tech mining is transforming how such data is extracted, structured, and interpreted. Unlike patent databases, which primarily reflect technological output, website data offers a richer view by incorporating information about market activities, external collaborations, and real-time corporate positioning (Kinne & Axenbeck, 2020). While bibliometrics offers a retrospective view of formal outputs, website data provide a forward-looking perspective on firms' public signaling of innovation.

Despite these benefits, website data present challenges. Content is unstructured and varies widely in style, requiring robust scraping and cleaning pipelines. The use of websites as a source may introduce biases: firms may use marketing language that exaggerates capabilities, leading to false positives, or omit internal innovations that they choose not to publicize, causing under-coverage (Gök et al., 2015). Technical barriers also exist. Dynamic content rendered through scripts can be difficult to extract, and access may be limited by server availability or restrictions imposed by the robots exclusion protocol (Koster et al., 2022). Multilingual websites add another layer of complexity, as variations in translation and local terminology hinder consistent keyword-based extraction (Bottai et al., 2024). Finally, the frequency of website updates differs across firms and sectors, meaning that websites may not reflect the actual timing of technology adoption.

The following section discusses how advances in natural language processing, and particularly through LLMs, can address some of the limitations associated with website mining and improve the reliability of AI adoption.

Large Language Models for automated classification

Recent developments highlight the transformative role of LLMs in mining unstructured content. These models enhance tasks such as information extraction, summarization, and semantic classification with unprecedented precision (Porter et al., 2024).

LLMs demonstrated capabilities in classification and labeling across fields such as education, healthcare, and business analytics, effectively processing complex inputs like essays or clinical notes with human-comparable accuracy (Savelka & Ashley, 2023; Huang et al., 2024). Their semantic understanding supports contextual tasks, including sentiment analysis, topic classification, and sensitive content detection (Chang et al., 2024). However, challenges persist. LLMs are prone to "hallucinations" and may generate inconsistent outputs, sometimes misinterpreting or overlooking important context (Liu et al., 2023; Huang et al., 2025).

Enhancing LLM-based labeling can be achieved by incorporating structured rubrics into prompt design. In education, rubrics define assessment criteria binary or multi-class to standardize evaluations and reduce subjectivity (Jonsson & Svingby, 2007; Panadero & Jonsson, 2013). Similarly, the "LLM-as-a-Judge" framework employs detailed instructions and constraints to guide models as consistent scoring agents across domains (Chiang et al., 2024; Gu et al., 2025).

These capabilities are now being leveraged in tech mining applications, especially when analyzing unstructured content from company websites and scientific literature. LLMs facilitate the discovery of hidden patterns and trends in textual data, supporting more informed decision-making in innovation management (Bouschery et al., 2023). This shift illustrates how tech mining is moving beyond classical bibliometric analysis and embracing AI-powered approaches that extend the boundaries of research intelligence.

Despite the advantages, maintaining model consistency under repeated tasks remains a challenge. An emerging solution involves complementing LLM-based classification with Retrieval-Augmented Generation (RAG) techniques (Lewis et al., 2020; Gao et al., 2024). This approach segments or chunks input text, providing the LLM with only the most relevant fragments, thus reducing ambiguity and improving the model's focus on critical information (Liu et al., 2023).

Thus, building on prior research in AI adoption, web content analysis, and the applicability of Large Language Models, this study proposes a scalable and replicable method for identifying and monitoring firm-level AI adoption based on the semantic analysis of company website content.

Methodology

This section describes the methodology used to detect AI adoption among Spanish firms through the analysis of their website content. First, we present the data sources, which include structured firm-level information and web-based text. Then, we explain the processing steps and analysis pipeline to identify and categorize AI use within companies.

Data sources

This study draws on firm-level data from two complementary sources to construct a balanced panel of Spanish companies observed in 2023 and 2025. The objective is to measure the prevalence and evolution of AI adoption by combining structured business information with website-based content.

The first source is the *Sistema de Análisis de Balances Ibéricos* (Bureau van Dijk, 2025, SABI), a commercial database widely used in economic and innovation studies. It provides structured firm-level information such as employment size, geographic location, sector classification, and company website. Firm size is classified according to the European Commission (2003) SME employment thresholds: firms with fewer than 10 employees are considered micro, those with 10 to 49 are small, those with 50 to 249 are medium, and those with 250 or more are large.[1] Economic activity is classified using the NACE Rev. 2 system at the section level (European Commission, 2008). All major sections are represented except three: Public administration and defense (section O), Activities of households as employers (section T), and Activities of extraterritorial organizations and bodies (section U). These sections were absent or virtually unrepresented in the initial frame and were therefore omitted. Geographic location is recorded at the province level, corresponding to the NUTS 3 level (Commission, 2024).

The second source consists of the contents of company websites, which were crawled in the first quarter of 2023 and 2025. They serve as digital footprints reflecting external communication of technological capabilities. The collection and processing of website content are described in detail in Sect. 3.2.

[1] Although the full SME definition also considers turnover and balance sheet totals, these variables were not used due to limited availability and inconsistency across firms. This simplification may affect the classification of capital-intensive firms with small workforces.

The sample frame comprises all Spanish firms listed in SABI for the year 2025 that report a website. From this universe, we applied a series of filters to construct a balanced 2-year panel:

(i) Firms must be active in both 2023 and 2025 according to SABI records.
(ii) A valid and publicly accessible website URL must be present in both years.
(iii) Size, sector and geography information at a firm-level must be available.

The application of these criteria resulted in a final panel of 118,605 unique firms observed in both years, for a total of 237,210 company–year observations. The stepwise filtering process is illustrated in Fig. 1. Firms that were newly created, exited the market, or had incomplete records across the 2 years were excluded from the final panel. As a result, the sample may overrepresent larger and more digitally established firms, and underrepresent microenterprises and those without a public online presence. The implications of this restriction for generalizability are discussed in Sect. 6.1.

Table 1 displays the final sample distribution of firms across size categories and economic sectors for each year analyzed in the study.

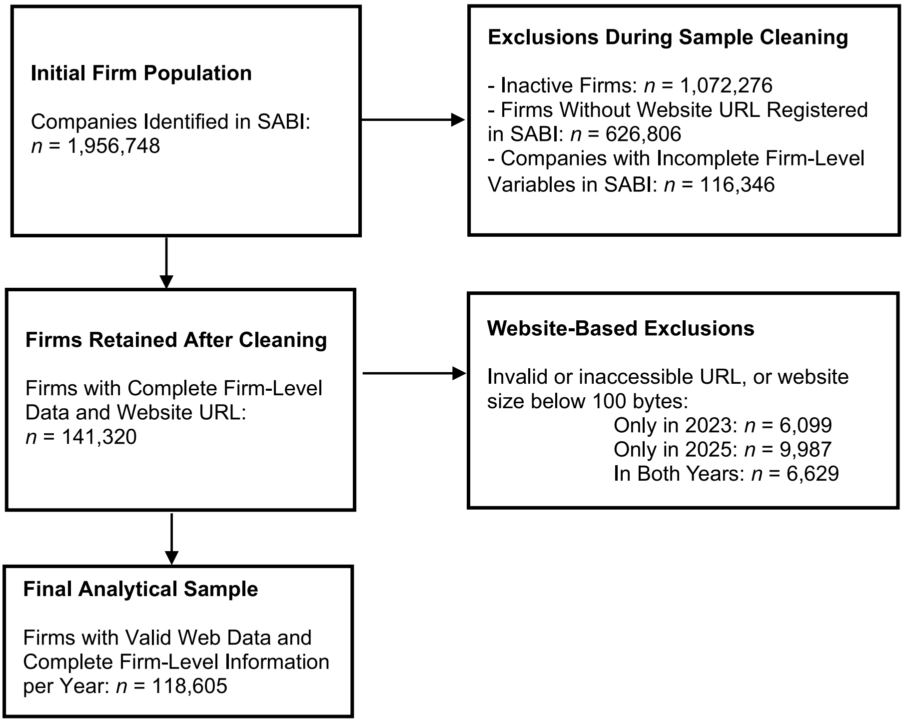


Fig. 1 Flowchart of the Sample Creation Process. The diagram illustrates the steps taken to construct the final sample for the study

Table 1 Sample composition by economic sector and firm size

	Large	Medium	Small	Micro	Total
Administrative and support services	35	429	1515	3454	5433
Agriculture, livestock, forestry, and fishing	4	89	414	636	1143
Artistic, recreational, and entertainment activities	5	93	385	1099	1582
Construction	7	506	4880	8364	13,757
Education	3	159	680	1633	2475
Electricity, gas, steam, and air conditioning supply	0	6	69	185	260
Extractive industries	0	8	128	105	241
Financial and insurance activities	2	22	232	1446	1702
Healthcare and social services	11	184	981	2190	3366
Hospitality	11	363	2552	2562	5488
Information and communications	6	241	1327	3766	5340
Manufacturing industry	23	1232	8838	10,598	20,691
Other services	1	46	343	992	1382
Professional, scientific, and technical activities	11	272	2762	11,367	14,412
Real estate activities	1	18	315	4274	4608
Transportation and storage	5	283	1856	1822	3966
Water supply, sanitation, waste management, and decontamination	1	34	216	223	474
Wholesale and retail trade; vehicle repair	17	849	9184	22,235	32,285
Total	143	4834	36,677	76,951	118,605

Website content processing pipeline

This section describes the procedure used to derive AI adoption indicators from firm website content, as outlined in Fig. 2. The methodology integrates web crawling, semantic filtering, and large language model classification to extract structured information from unstructured text. The resulting indicators capture the presence and type of AI use in firms' processes and products.

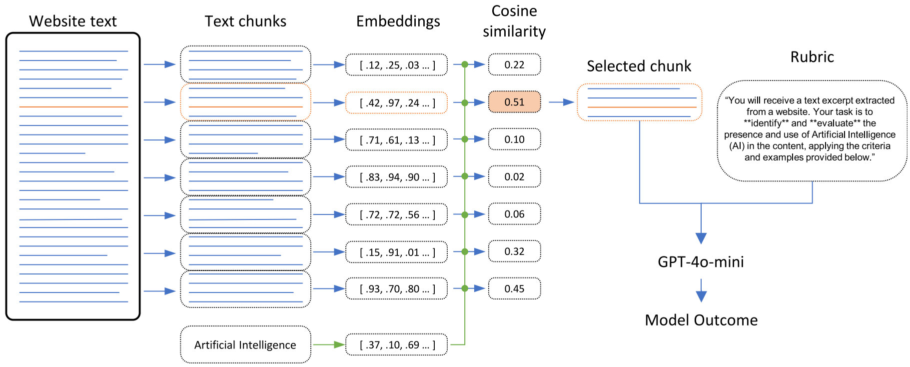


Fig. 2 Pipeline for assessing the use of AI by companies from website text

Web crawling and text preprocessing

Company website content was collected using a standardized and replicable crawling procedure applied consistently across both observation periods (2023 and 2025).

The GNU `wget` crawler was configured to operate within ethical and technical constraints, imposing a maximum download quota of 200 MB per website and a rate limit of 500 KB/s to minimize server load. Only HTML pages and PDF documents were retrieved, as these are the most common formats used for publicly accessible firm-level communication.

The crawler was set to follow internal links up to two levels deep. Internal links were identified as URLs pointing to the same domain or closely related variants, such as minor naming variations (e.g., removal of dashes) or identical second-level domains with different top-level domains.

After retrieval, HTML files were converted into plain text by removing all HTML tags and embedded scripts. PDF files were processed and converted to text using standard text extraction methods (`pdftotext`, included in standard Linux distributions).

No language-based filtering was applied, resulting in multilingual corpora primarily in Spanish, but also containing English and regional languages (e.g., Catalan). Text preprocessing involved basic cleaning, specifically HTML tag removal and whitespace normalization. No additional linguistic preprocessing steps, such as stopword removal or lemmatization, were implemented, as semantic analysis in subsequent stages relies on the original context-preserving structure of the text.

The resulting texts from each company's website, across all retrieved HTML pages and PDFs, were concatenated into a single corpus per firm without imposing any size constraints. The median size (in bytes) and number of subpages per company website are reported in Table 2.

Chunking and relevance filtering

The consolidated text of each website was segmented into 1000-byte non-overlapping fragments (or chunks), without relying on paragraph or sentence boundaries. Each chunk was then transformed into high-dimensional vectors using contextual embeddings generated by OpenAI's `text-embedding-3-small` model. This representation captures the semantic meaning of the text (Devlin et al., 2019; Peters et al., 2018).

To identify AI-relevant content, we encoded the phrase "Artificial Intelligence" using the same embedding model. This phrase served as a reference vector for semantic filtering, chosen for its generality and neutrality across business and technical contexts.

Cosine similarity was then computed between each chunk and the reference vector. Cosine similarity is a widely used metric in NLP that measures the cosine of the angle between two vectors, indicating their degree of semantic similarity (Mikolov et al., 2013). It is defined as:

Table 2 Median size and number of extracted web pages per firm

Year	Size (bytes)	Number of pages
2023	623,952	12
2025	975,384	13

Table 3 Description of variables used in the study

Source	Variable	Description	Type	Mean	
				2023 (%)	2025 (%)
SABI	Size	Size of the company	Categorical	–	–
	Sector	NACE section of the primary activity	Categorical	–	–
	Location	Province in which the company is registered	Categorical	–	–
Website	AI1	The company uses AI	Binary	3.52	4.80
	AI2	AI adopted in business processes	Binary	3.24	4.42
	$AI3_{PGS}$	AI used in Production of Goods & Services	Binary	1.40	1.76
	$AI3_{DL}$	AI used in Distribution & Logistics	Binary	0.40	0.45
	$AI3_{MS}$	AI used in Marketing & Sales	Binary	1.06	1.75
	$AI3_{ICS}$	AI used in Information & Communication Systems	Binary	0.84	1.20
	$AI3_{AM}$	AI used in Administration & Management	Binary	0.95	1.56
	$AI3_{PDB}$	AI used in Product Development & Business	Binary	0.94	1.39
	AI4	AI adopted in products and services	Binary	2.45	3.26
	AI5	AI is used in products radically	Binary	1.74	2.23

$$\text{Cosine Similarity} = \frac{A \cdot B}{\|A\|\,\|B\|}, \tag{1}$$

where $A \cdot B$ denotes the dot product of the two vectors, and $\|A\|\|B\|$ represents the product of their magnitudes. In this context, higher similarity values indicate chunks that are semantically closer to the concept of artificial intelligence.

Semantic similarity does not imply evidence of AI adoption, but rather highlights segments of text likely to contain relevant descriptions. From each firm's corpus, we selected the 10 chunks with the highest similarity scores, focusing the analysis on the most relevant content while excluding irrelevant or low-alignment material.

To improve the contextual grounding of each selected chunk, a symmetric 1.1KB context window was introduced around each of these top-selected chunks,[2] capturing adjacent text and sentence boundaries where possible. This enriched input allows the language model to better interpret references to AI by incorporating surrounding content. Moreover, the combination of chunking and contextual extension helps avoid scoring fragments that are merely headings or lists of SEO keywords, as the model is guided by the semantic depth of surrounding discourse.

LLM classification and variable construction

Each text chunk identified as potentially AI related through semantic filtering was analyzed using an LLM, specifically GPT-4o-mini. The LLM classification was guided by a structured evaluation rubric detailed in "Appendix 2". This rubric operationalizes the adoption of AI by distinguishing between its presence in company processes and products, aligned with the innovation categories described in Sect. 2.1. The prompt was developed through an iterative prompt engineering process, in which various formulations were tested and

[2] The choice of chunk size, number of selected chunks, and context window was determined empirically.

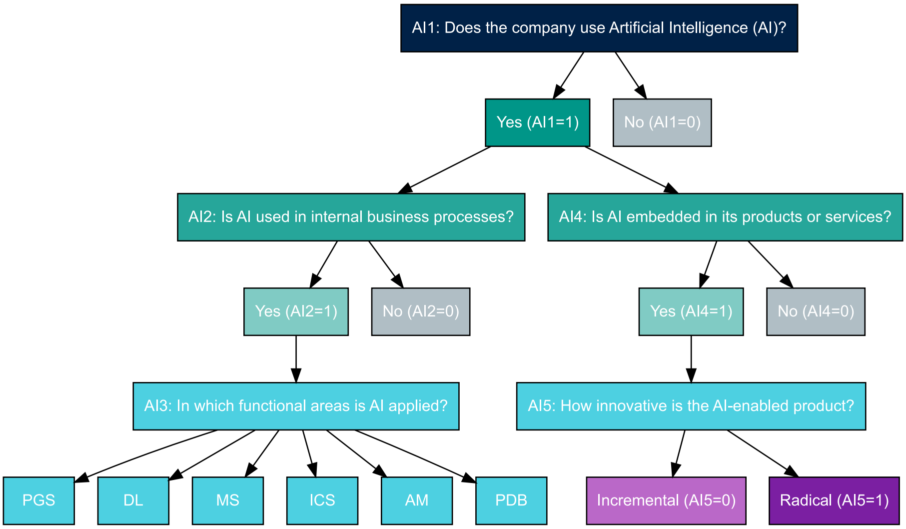


Fig. 3 Flowchart illustrating the evaluation strategy included in the rubric for detecting AI adoption in companies

refined through multiple interactions with the model. We compared the outputs of different prompt versions against manually labeled examples to assess clarity, consistency, and alignment with the intended classification goals. Rather than relying solely on keywords, the final prompt includes illustrative examples for each concept (e.g., types of AI use in production or marketing), to guide the model's interpretation and reduce ambiguity.

For each chunk, the model produced a structured output in JSON format, containing binary values for the predefined set of variables listed in Table 3. To assign firm-level labels, the outputs of all selected chunks for a company were aggregated.

A firm was marked as an AI adopter if at least one chunk was considered positive for the relevant variable by the LLM. That is, the firm-level variable was computed as the logical OR across all the classified chunks associated with the company. This rule applies to all variables derived from the LLM.

The classification procedure incorporates a conservative interpretation of clear evidence of implementation, aimed at distinguishing genuine AI adoption from superficial or promotional mentions. The rubric guides the LLM to identify only those instances where AI is functionally integrated into internal operations or embedded within products or services, based on specific and context-rich descriptions (e.g., predictive analytics in logistics, AI-driven recommendation systems). Abstract, vague, or aspirational references are excluded from positive classification. When the input lacks sufficient detail or is ambiguous, the model defaults to a negative response. This strict interpretive framework is further reinforced by the initial semantic filtering, which limits the input to highly relevant fragments, including adjacent context. Together, these design choices systematically reduce the likelihood of false positives resulting from AI washing.

This approach ensures that the presence of a single, well-formed statement about AI adoption is sufficient to mark a company as adopter, even if other fragments are silent or ambiguous. It also avoids diluting the signal from the relevant content by averaging or requiring multiple positive instances.

The evaluation strategy, represented in Fig. 3, starts by determining if AI is explicitly utilized within the company. If AI presence is confirmed, further analysis identifies its application in internal processes and external products or services. For internal applications, the model identifies specific business domains where AI is used, such as production, logistics, marketing, and others, offering a granular view of operational adoption. "Appendix 3" showcases paragraph examples categorized as either process- or product-oriented AI implementations.

When AI is integrated into products or services, its use is categorized as either enhancing existing functions (incremental innovation) or introducing fundamentally new capabilities (radical innovation). This distinction helps in understanding the depth of AI influence on the firm's offerings.

Evaluation of LLM-based classification method

This section evaluates the LLM-based method in terms of reliability and validity. Using a stratified sample of firm websites manually labeled, we assess whether the model produces reliable and valid classifications of AI adoption. This evaluation draws on emerging best practices for LLM use in research (Hajikhani & Cole, 2024), addressing concerns around bias and transparency through stratified validation and full procedural documentation (see "Appendix 5").

Evaluation design and methodology

Evaluation sample

To evaluate the classification model, we carried out an evaluation procedure using a human-labeled dataset. We employed a stratified sampling approach to ensure representation across all firm categories. Specifically, we stratified the sampling frame by company size and economic sector and selected an equal number of companies from each stratum, resulting in a validation sample of 228 firms.

This equal-allocation design guarantees that every business segment is represented, and it affords sufficient coverage of large firms, despite their smaller numbers, since they account for the highest share of AI adopters. A strictly proportional sampling strategy would have underrepresented large firms and risked biasing our accuracy assessment. The resulting labeled dataset serves as the ground truth against which we compare the LLM-based classifications. The manual labeling was conducted by two members of the research group with diverse backgrounds to ensure an independent and multidisciplinary evaluation. Both evaluators received the same prompt and classification rubric used by the LLM. In cases of disagreement, the annotation was jointly reviewed and resolved through discussion between the two evaluators.

Model reliability

LLMs are known to introduce variability in outputs even under fixed prompts and inputs. Reliability refers to the degree to which a measurement procedure yields consistent and reproducible results when it is applied repeatedly under identical conditions (Nunnally &

Bernstein, 1994). Applied to our setting, reliability refers to the stability of the LLM-based classifier across multiple runs that use the same input data and prompts.

To assess the model's stability across executions, we applied the classification pipeline ten times to the websites of the companies in the evaluation sample. The outputs from these runs were compared to assess the level of agreement in binary classification decisions. Two statistical measures were used to quantify reliability.

The Intraclass Correlation Coefficient (ICC) is a standard metric used to assess the reliability of repeated measurements when the same subjects are rated multiple times. In our context, the "subjects" are the company websites, and the "rater" is the LLM classification pipeline applied repeatedly under identical conditions. The formula is:

$$\mathrm{ICC} = \frac{\mathrm{MS_R} - \mathrm{MS_E}}{\mathrm{MS_R} + (k-1)\mathrm{MS_E}}, \tag{2}$$

where $\mathrm{MS_R}$ is the mean square for rows (subjects), $\mathrm{MS_E}$ is the mean square error, and k is the number of raters or repeated measurements.

High ICC values (close to 1) indicates that the classifier generates stable and repeatable labels (Qin et al., 2019), suggesting the model is reliable for large-scale automated annotation.

Cohen's κ is a complementary measure calculated to evaluate the level of agreement in categorical classifications across runs. Unlike simple percent agreement, κ adjusts for the likelihood of chance agreement, offering a more rigorous evaluation of reliability (Cohen, 1960).

$$\kappa = \frac{p_\mathrm{o} - p_\mathrm{e}}{1 - p_\mathrm{e}}, \tag{3}$$

$$p_\mathrm{e} = \sum_{i=1}^{k} p_i^{(1)} \cdot p_i^{(2)}, \tag{4}$$

where p_o is the observed agreement proportion, p_e is the expected agreement by chance, and $p_i^{(1)}$ and $p_i^{(2)}$ are the marginal proportions for category i from each run.

Cohen's κ values are interpreted as follows (Landis & Koch, 1977): $\kappa > 0.41$ indicates moderate agreement, $\kappa > 0.61$ indicates substantial agreement, while $\kappa > 0.81$ indicates almost perfect agreement.

Model validity

Validity refers to the extent to which a measurement procedure accurately captures the concept it is intended to assess (Messick, 1995). In our context, it concerns whether the LLM-based classification correctly identifies AI adoption by firms based on their website content. A classification method should produce labels that are not only reliable but also aligned with expert human judgment when applied to the same input data.

To assess the validity of the LLM in classification tasks, we employ widely recognized evaluation metrics, including accuracy, precision, recall, and the F_1Score (Powers, 2020). These metrics provide a comprehensive understanding of the model's effectiveness by measuring different aspects of its predictive capabilities and are calculated based on True Positives (TP), True Negatives (TN), False Positives (FP), and False Negatives (FN).

Accuracy represents the proportion of companies correctly identified regarding their use or non-use of AI. It reflects the model's overall effectiveness in distinguishing between companies that adopt AI solutions and those that do not:

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}}. \tag{5}$$

Precision assesses the performance of the model in correctly classifying a company as an AI user. It measures how many of the companies identified as AI users truly implement it, ensuring that AI implementation is not overestimated:

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}}. \tag{6}$$

Recall quantifies the model's ability to identify all companies that use AI. This is particularly relevant to capture the full extent of AI adoption across companies:

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}}. \tag{7}$$

The $F_1 Score$ provides a balanced measure between precision and recall, making it especially useful when there is an uneven distribution of AI-adopting and non-adopting companies. This metric ensures that both false positives and false negatives are considered in the overall evaluation:

$$F_1\text{Score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}. \tag{8}$$

To complement the conventional metrics, we also report the Matthews Correlation Coefficient (MCC). MCC is better suited than accuracy or F_1Score for evaluating models under class imbalance, as it incorporates all elements of the confusion matrix and does not inflate performance when one class dominates. It is computed using the following formula:

$$\text{MCC} = \frac{\text{TP} \times \text{TN} - \text{FP} \times \text{FN}}{\sqrt{(\text{TP} + \text{FP})(\text{TP} + \text{FN})(\text{TN} + \text{FP})(\text{TN} + \text{FN})}}. \tag{9}$$

The MCC value ranges from −1 to 1, where 1 indicates perfect classification, 0 represents random performance, and −1 signifies total disagreement between predictions and actual labels (Chicco & Jurman, 2020).

Evaluation results

Model reliability

Results of the ten times executions under identical conditions are presented in Table 4. The highest reliability, measured by ICC and Cohen's κ, is observed when classifying companies as AI users (AI1). High reliability is also found in classifying whether companies employ AI in business processes (AI2) and in products and services (AI4).

In contrast, the method exhibited relatively lower reliability in identifying the functional domains where AI is used (AI3 variants), likely reflecting the greater complexity involved in these distinctions. This was not the case for AI3_{MS}, suggesting that the detection of AI

Table 4 ICC and Cohen κ results for each AI response

	ICC	Cohen κ
AI1	0.89	0.88
AI2	0.87	0.87
$AI3_{PGS}$	0.67	0.66
$AI3_{DL}$	0.61	0.59
$AI3_{MS}$	0.78	0.78
$AI3_{ICS}$	0.45	0.44
$AI3_{AM}$	0.60	0.59
$AI3_{PDB}$	0.64	0.64
AI4	0.80	0.80
AI5	0.74	0.74

use in marketing and sales is more reliable. Finally, the classification of the product's innovation radicalness (AI5) also showed good reliability.

Model validity

Table 5 reports the model's predictive validity metrics averaged across ten repeated executions for each classification output defined in the framework. The high overall accuracy across models (mean ≈ 0.951) confirms the system's capacity to distinguish between firms that do or do not adopt AI (AI1), apply AI in internal processes (AI2), embed AI into products or services (AI4), and distinguish between incremental or radical innovation (AI5). These high-level decisions yield F_1Scores between 0.68 and 0.77, and Matthews Correlation Coefficients in the range of 0.66 to 0.74, indicating robust discriminative power and reliable signal detection even under potential class imbalance.

However, performance is more heterogeneous for AI3—the dimension that identifies the specific functional area in which AI is applied. While some categories such as $AI3_{MS}$ and $AI3_{PGS}$ exhibit strong accuracy (0.96 and 0.95) and acceptable F_1Scores (0.59 and 0.53), others like $AI3_{ICS}$ and $AI3_{DL}$ show more imbalanced trade-offs between precision and recall, resulting in lower F_1Scores (0.35 and 0.42) and MCC values under 0.50. This is connected with the lower reliability of these same classifications.

Table 5 Model performance metrics for each AI dimension

	Accuracy	Precision	Recall	F_1Score	MCC
AI1	0.94	0.69	0.80	0.74	0.71
AI2	0.95	0.72	0.82	0.77	0.74
$AI3_{PGS}$	0.95	0.64	0.46	0.53	0.52
$AI3_{DL}$	0.95	0.85	0.29	0.42	0.47
$AI3_{MS}$	0.96	0.61	0.58	0.59	0.58
$AI3_{ICS}$	0.94	0.45	0.30	0.35	0.34
$AI3_{AM}$	0.97	0.53	0.60	0.56	0.55
$AI3_{PDB}$	0.93	0.63	0.35	0.45	0.44
AI4	0.95	0.66	0.71	0.68	0.66
AI5	0.96	0.56	0.75	0.64	0.63

In summary, the model exhibits suitable validity across core classification dimensions (AI1, AI2, AI4, AI5), while showing moderately lower but still informative performance in the more granular classification of functional AI use (AI3).

Results

This section presents the key results of our analysis on AI adoption among Spanish firms, based on web content extracted from their official websites in 2023 and 2025. The results are organized into three main dimensions: adoption patterns across firm sizes and economic sectors, the geographic distribution of adopters at the provincial level, and the typology of AI use within firms—distinguishing between internal process integration and product-oriented applications.

AI Adoption segmented by size and sector

Table 6 presents AI adoption patterns by industry and firm size in 2025, revealing important differences in how companies integrate AI technologies depending on their characteristics. A clear gradient emerges across firm sizes: large companies consistently show higher adoption rates, followed by medium-sized firms, while small and micro-enterprises

Table 6 AI adoption rates across industries and firm sizes in 2025

Description	Large	Medium	Small	Micro	Total
Administrative and support services	0.143	0.065	0.066	0.041	0.051
Agriculture, livestock, forestry, and fishing	*	0.022	0.014	0.017	0.017
Artistic, recreational, and entertainment activities	*	0.043	0.044	0.020	0.027
Construction	*	0.032	0.021	0.019	0.020
Education	*	0.132	0.121	0.061	0.082
Electricity, gas, steam, and air conditioning supply	*	*	0.043	0.059	0.062
Extractive industries	*	*	0.008	0.010	0.008
Financial and insurance activities	*	0.045	0.121	0.076	0.082
Healthcare and social services	*	0.071	0.057	0.041	0.047
Hospitality	*	0.008	0.008	0.009	0.008
Information and communications	*	0.573	0.414	0.180	0.257
Manufacturing industry	0.130	0.062	0.026	0.015	0.023
Other services	*	0.152	0.050	0.034	0.042
Professional, scientific, and technical activities	*	0.360	0.148	0.060	0.083
Real estate activities	*	0.222	0.054	0.033	0.035
Transportation and storage	*	0.046	0.023	0.015	0.021
Water supply, sanitation, waste management, and decontamination	*	0.118	0.028	0.013	0.027
Wholesale and retail trade; vehicle repair	0.118	0.079	0.043	0.031	0.036
Total	0.154	0.103	0.057	0.040	0.048

Values represent the proportion of firms in each category using AI

*Indicates sectors with fewer than 15 firms present more modest levels of uptake. This gradient is consistent with prior evidence on scale advantages in digital technologies.

AI adoption is highest in the Information and Communications sector, driven by its digital maturity and reliance on data-intensive services. Professional, Scientific, and Technical Activities also show strong uptake, particularly among medium and large firms, reflecting AI's value in enhancing knowledge-based work. In contrast, traditional sectors like Agriculture, Hospitality, Construction, and Manufacturing exhibit limited adoption, with AI use often confined to support functions rather than core operations, especially among smaller firms. This contrast aligns with expectations of greater uptake in knowledge-intensive sectors compared to traditional industries.

Figure 4 illustrates the change in the proportion of companies adopting AI between 2023 and 2025, broken down by firm size. The results reveal a consistent increase in AI adoption across all size categories, with the most pronounced growth observed among larger firms. Medium-sized firms also saw notable growth, while small and micro-enterprises experienced only modest gains, indicating persistent resource and capability gaps.

Figure 5 presents the percentage change in AI adoption by economic sector between 2023 and 2025. The results confirm that the expansion of AI use is highly uneven across industries, with a few sectors driving most of the observed growth.

The Information and Communications sector leads in AI adoption growth, reinforcing its role in digital innovation. Significant increases are also observed in Energy and Education. Moderate gains in Professional, Scientific, Technical, and Financial sectors reflect continued integration of AI in service delivery. In contrast, sectors like Construction, Manufacturing, and Hospitality show limited progress, with Extractive Industries stagnant and Agriculture experiencing a slight decline.

AI Adoption segmented by Geography

Figure 6 illustrates the geographic distribution of AI usage across Spanish provinces in 2023 and 2025, respectively. The maps are segmented by quintiles, with darker shades

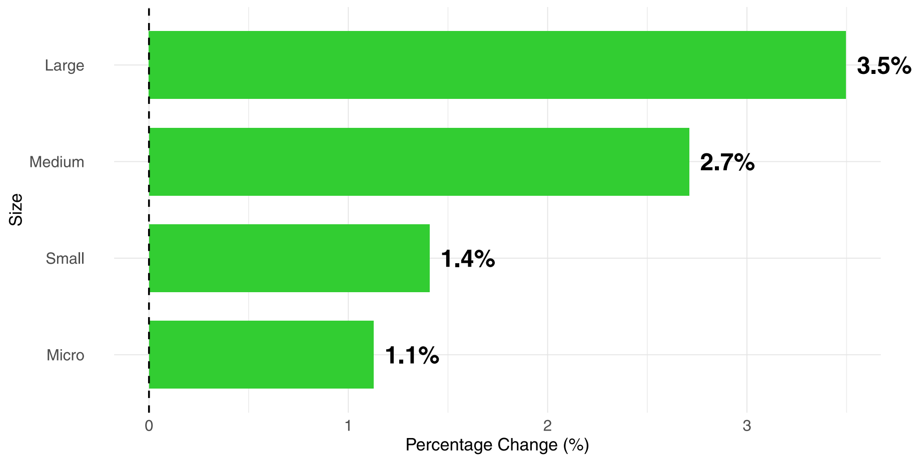


Fig. 4 Change in the proportion of AI-adopting companies by size between 2023 and 2025

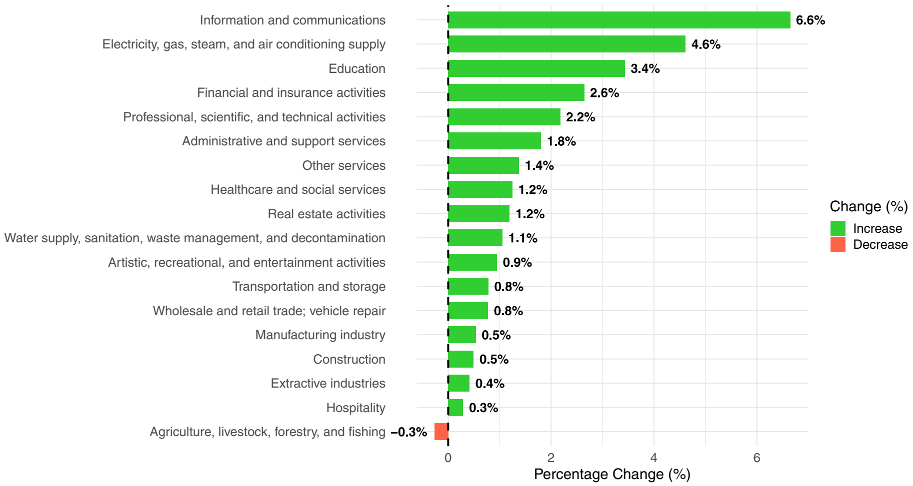
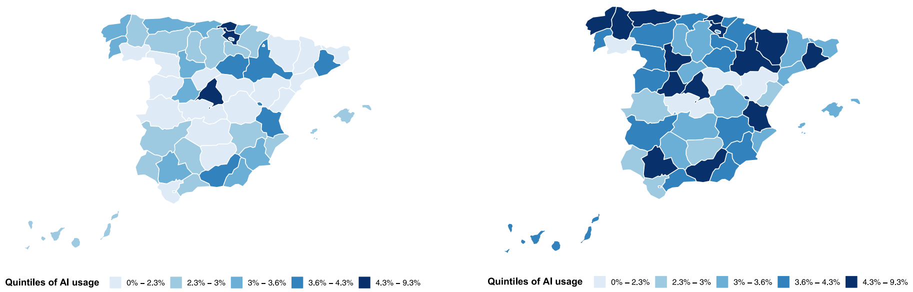


Fig. 5 Percentage change in AI adoption by sector from 2023 to 2025. Increases are shown in green and decreases in red. (Color figure online)



(a) AI use by firms in Spain, 2023

(b) AI use by firms in Spain, 2025

Fig. 6 AI adoption across Spanish provinces (NUTS-3 level) in 2023 and 2025. Provinces are grouped into quintiles by the proportion of firms adopting AI

indicating higher levels of AI adoption based on the proportion of firms mentioning the use of AI on their corporate websites.

To quantify these spatial dynamics, we computed dispersion statistics for the provincial distribution of AI adoption. The Herfindahl–Hirschman Index (HHI) decreased from 0.115 in 2023 to 0.106 in 2025, while the Gini coefficient fell from 0.691 to 0.669, and the Theil index from 1.054 to 1.002. These declines indicate a modest reduction in concentration across provinces, consistent with a gradual democratization of adoption patterns.

In 2023 (Fig. 6a), the spatial pattern of AI usage is relatively dispersed, with higher adoption concentrated in a few central and northern provinces. The majority of provinces fall within the lower to middle quintiles, denoting that AI adoption was still in its early diffusion phase, with limited regional concentration. Some industrial hubs and urbanized areas already displayed moderate adoption, but the national landscape was characterized by fragmentation and uneven development.

By 2025 (Fig. 6b), the landscape changes significantly. A clear intensification of AI usage is observed, with a larger number of locations moving into the upper quintiles. High AI adoption is no longer confined to a few regions; instead, it spreads across both northern and southern provinces, indicating a broader diffusion of AI technologies. Notably, several traditionally less digitized provinces have progressed into higher quintiles, suggesting that regional gaps are beginning to narrow—although some disparities persist.

This spatial pattern matches expectations. In 2023, higher adoption appears in Madrid, Catalonia, and the Basque Country, which are established innovation hubs. By 2025, adoption spreads to more provinces, showing a diffusion from core regions to others. This suggests that the indicator is capturing plausible regional dynamics of AI adoption.

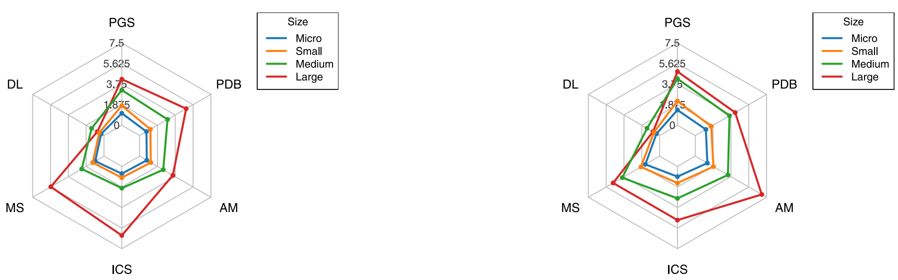
AI Adoption segmented by Process/Product Application

Based on firm website content, the model indicates that among companies identified as AI adopters in 2025, AI use tends to be more commonly associated with process-related activities (92.08%) than with product-related ones (67.92%). When AI is mentioned in relation to products, around 68.4% of those mentions are interpreted as referring to radical innovation.

Focusing on the process-oriented results, Fig. 7 illustrates how companies of different sizes applied AI across various business functions in 2023 and 2025, specifically in the context of process innovation. The radar charts clearly reveal a persistent size-based gradient: larger firms consistently demonstrate higher levels of AI integration in all functional areas compared to medium, small, and micro enterprises with the exception of Distribution and Logistics.

In 2023 (Fig. 7a), large firms concentrated their AI efforts in Marketing and Sales, Product and Business Development, and Administration and Management. Medium-sized companies showed moderate engagement across functions, with relatively stronger use in Production of Goods and Services and Distribution and Logistics. Micro and small firms reported minimal use of AI, with their limited activity primarily focused on accessible domains like Information and Communication Systems and Distribution and Logistics.

By 2025 (Fig. 7b), the functional reach of AI broadened for all firm sizes, but the disparity between large firms and the rest grew even more pronounced. Larger companies



(a) AI usage in 2023 (b) AI usage in 2025

Fig. 7 Comparison of AI usage across business functions by firm size in 2023 and 2025. The figures show the shift in AI adoption across different business functions and firm sizes over time exhibited intensified adoption in almost every area, especially in Administration and Management and Product and Business Development. Medium-sized firms increased their presence in Production, while smaller firms maintained a limited but slightly expanded use in technical areas like Information and Communication Systems and Distribution and Logistics.

These results are consistent with expectations. Firms more oriented toward products, such as those in digital services and communications, show stronger signals of AI in product innovation, while process-intensive industries like manufacturing display greater emphasis on internal applications. The predominance of process-related use across all sectors also reflects the common trajectory in which companies first apply AI to improve operations before embedding it in external offerings. This alignment suggests that the indicator captures plausible distinctions between product- and process-oriented adoption.

The distribution of the different business functions remains consistent across firm sizes and over time. This is visually reflected in the concentric shape of the radar plots, where firms of different sizes follow similar patterns of AI application within each year.

Overall, the charts highlight that firm size plays a key role in shaping not only whether AI is adopted, but also where within the business it is implemented. Larger firms appear more capable of leveraging AI across multiple functional domains.

Discussion

The integration of LLMs into the analysis of corporate web content represents a methodological advancement on innovation measurement and digital transformation. By applying an LLM-based pipeline to analyze the textual content of corporate websites, we introduce a novel and scalable approach for detecting firm-level innovation and technology adoption. This method improves upon traditional survey-based techniques, which, in addition to being more costly, often suffer from limited sample sizes and delayed reporting (Gök et al., 2015; McElheran et al., 2024). It also overcomes key limitations of previous web-based approaches that rely on keyword matching, which can miss implicit references to AI or generate false positives, particularly in multilingual contexts (Alekseeva et al., 2021). In contrast, our LLM-based method provides semantic understanding of textual content, enabling more accurate detection of AI adoption even when firms use varied or indirect language. Furthermore, this approach offers a dynamic alternative to static measurement methods by allowing continuous monitoring of technological adoption through publicly available digital content. As such, it enhances the methodological toolkit for innovation studies, providing timely and context-aware indicators suited to tracking fast-evolving technologies.

More recent studies, such as Dahlke et al. (2024), also address these limitations through the use of supervised language models trained to differentiate between substantive and superficial references to AI. Their method combines keyword-based filtering for feature selection with a supervised transformer model trained for classification. Our approach uses chunk selection based on cosine similarity to select relevant content, followed by classification through a rubric-based LLM. This design supports comparability across firms and sectors, reduces dependency on labeled data, and enables monitoring of AI adoption using open web content.

The findings from this study offer detailed evidence of how AI adoption is spreading among Spanish firms following the launch of ChatGPT in 2022. To support external validity, we compare our results with official data from the Spanish National Statistics Institute (INE), which reports a 5.75% AI usage rate among firms with more than 10 employees in 2023. This aligns well with our overall estimate of 4.8% across all firm sizes in 2025, and 10.3% among medium and large firms. Sectoral patterns also converge across datasets. In both our data and that of INE, the services sector emerges as the most AI-intensive. In contrast, more traditional sectors—including Agriculture, Hospitality, and Manufacturing—consistently report lower adoption rates.

Geographically, high adoption regions coincide. Madrid (INE: 12.82%) and the Basque Country (INE: 10.66%) lead in both datasets, while Extremadura, Ceuta, and Melilla consistently show lower levels. Our province-level data further reveal a gradual diffusion of AI into less urbanized areas between 2023 and 2025, supporting the notion of narrowing regional gaps. Unlike INE data, our approach also captures functional differences—highlighting that most adopters use AI in internal processes (92%) and, when used in products, it often represents radical innovation (68%).

Overall, we observe a substantial increase in the prevalence of AI use between 2023 and 2025, confirming that the adoption of AI technologies has accelerated markedly in recent years. While the overall AI adoption levels observed in this study are consistent with similar research (Calvino & Fontanelli, 2023; McElheran et al., 2024; Dahlke et al., 2024), our LLM-based methodology provides a more detailed perspective on adoption patterns. Specifically, this approach allows for differentiation of AI integration across firm sizes, economic sectors, and geographic areas, as well as identification of the functional domains in which AI technologies are employed. Consequently, it delivers a more granular understanding of the heterogeneity and evolution of AI adoption compared to previous methods relying on keyword-based approaches or aggregate survey data.

The results show that larger firms exhibit substantially higher AI uptake than small and mid-sized enterprises, consistent with prior research evidence that AI diffusion shows a "large firm bias". Cross-country data indicate that the share of AI-using firms is highest among those with over 250 employees, often roughly double the adoption rate of the next-largest size category (Calvino & Fontanelli, 2023). This results align and complement similar studies which focused on other geographies such us European countries or United States (Calvino & Fontanelli, 2023; Alekseeva et al., 2021; McElheran et al., 2024).

As presented in Fig. 5, this study shows that sectoral differences in AI adoption are similarly pronounced, with knowledge-intensive industries leading the way and traditional sectors lagging behind. Other studies also confirm that AI use is especially prevalent in high-tech and professional industries—for example, the Information and Communication sector and professional, scientific and technical services show the highest adoption rates (Calvino & Fontanelli, 2023; McElheran et al., 2024).

Beyond confirming well-known patterns such as the dominance of large firms and knowledge-intensive sectors in AI adoption, our study reveals a widening gap in AI adoption across both firm size and sectoral dimensions. Larger firms consistently exhibit higher growth rates in AI adoption compared to medium-sized, small, and micro-enterprises, suggesting increasing disparities linked to resource availability and technical capabilities. This divergence raises important questions about the potential for digital inequality, as smaller firms may face structural barriers preventing them from fully leveraging AI-driven productivity gains. Similarly, the sectoral analysis underscores that industries already positioned at the technological frontier, mainly Information and Communications, Education, and Financial and Insurance activities, continue to accelerate their adoption rates, whereas traditional sectors like Agriculture, Construction, Manufacturing, and Hospitality show limited progress. This growing sectoral divide highlights the uneven distribution of digital transformation benefits and suggests the need for targeted policy interventions aimed at reducing barriers to AI adoption in lagging sectors and smaller firms, to mitigate potential long-term competitive disadvantages.

A further dimension of uneven adoption highlighted by our study is geography. AI adoption by firms in Spain remains geographically concentrated in major urban and economic hubs. At the same time, dispersion statistics indicate modest spread from 2023 to 2025 (Gini 0.69 to 0.67; Theil 1.05 to 1.00; HHI 0.115 to 0.106), suggesting a gradual but limited narrowing of regional gaps. Firms headquartered in larger cities still show substantially higher rates of AI-related activity, reflecting their role as centers of technological innovation, talent, and digital entrepreneurship, while less populated regions continue to lag. This core-periphery pattern of technological diffusion is again consistent with prior research. McElheran et al. (2024) document that in the US, a handful of "superstar" cities (e.g., San Francisco, Seattle, New York) and emerging hubs led the early adoption of AI among startups, whereas other regions lagged, raising concerns about spatial inequalities. In Europe, recent analyses similarly point to regional hotspots driving AI uptake. For instance, Dahlke et al. (2024) identify strong geographic clustering of AI adoption in the DACH region (Germany, Austria, Switzerland), finding that firms embedded in knowledge-rich regions and networks are much more likely to adopt AI.

Our findings reveal that the majority of Spanish firms adopting AI tend to apply it within internal processes while more than half also use it in products. This aligns with previous studies suggesting that firms often integrate new technologies incrementally, using them first to enhance productivity (Alekseeva et al., 2021; Czarnitzki et al., 2023).

In terms of functional application, AI is most frequently used in Administration and Management, followed by Production of Goods and Services and Business Development. These areas show the highest levels of process-related adoption, especially among larger firms, indicating a focus on decision-making, coordination, and strategic planning. This pattern is consistent with recent studies highlighting the growing role of AI in management and innovation functions (Alekseeva et al., 2021; Rammer & Es-Sadki, 2023), in contrast to earlier expectations that emphasized its use in operational tasks (Hollenstein & Woerter, 2008).

Beyond process innovation, our findings reveal that when AI is embedded in products, most of these cases qualify as radical innovations. This supports the theoretical perspective that AI-driven product innovations have the potential to significantly reshape industries, particularly through technologies such as deep learning, autonomous systems, or advanced diagnostics (Bresnahan, 2010; Schumpeter & Swedberg, 2013). Unlike incremental process improvements, radical product innovations introduce fundamentally new functionalities or redefine existing offerings, often enabling firms to move into entirely new market spaces. Although fewer companies integrate AI into products compared to processes, those that do are often pursuing high-impact transformations. This observation resonates with the notion of AI as a general-purpose technology whose most disruptive potential lies in enabling entirely new forms of value creation (Cockburn et al., 2019).

Notably, the use of a rubric-guided LLM allows for classification of AI adoption by business function, offering a level of granularity not typically captured in survey-based or keyword-driven approaches. The results indicate that administrative and product development functions dominate among large firms, while early signs of AI use in logistics and information systems are emerging among smaller companies. These functional patterns provide a more nuanced picture of how AI is being integrated across organizational layers and firm sizes. This approach reflects recent developments in the "WebAI" paradigm for innovation research (Dahlke et al., 2025), which emphasizes the growing role of AI-driven tools in extracting structured insights from firms' digital self-representations. In this context, the present study contributes to ongoing debates about the use of web data for innovation measurement by demonstrating how LLMs—when paired with domain-specific rubrics and contextual filtering—can move beyond superficial mentions and toward meaningful indicators of technological adoption and functional integration.

Limitations

A number of limitations should be acknowledged in this study. First, not all active companies in Spain are included in the SABI database, which may limit the coverage of the sample. Second, there is not always a one-to-one correspondence between companies and website URLs—some firms may share a single URL (as in the case of franchises), while others may operate multiple digital domains, complicating the mapping process. Third, the selection of firms based on the availability of a corporate website introduces a coverage bias, as companies without an online presence are systematically excluded. Fourth, limiting the analysis to firms observed in both 2023 and 2025 may introduce survivorship bias, potentially underrepresenting newer firms or those that exited the market during the period and overrepresenting larger and more digitally established firms. Fifth, evaluation results were lower in some AI_3 subcategories, likely due to the greater complexity of more granular functional classifications beyond the product/process level. Sixth, the human-labeled sample was annotated by members of the research team rather than external experts, despite efforts to ensure multidisciplinary input.

Additionally, broader methodological challenges common to web-based data analysis must also be considered. These include the use of marketing language that may exaggerate capabilities, heterogeneous web structures and content quality, technical barriers to extracting dynamic or restricted pages, multilingual inconsistencies and local terminology that hinder keyword-based detection, and asynchronous website updates that may not reflect the actual timing of technology adoption (Gök et al., 2015; Koster et al., 2022; Bottai et al., 2024; Dahlke et al., 2025).

Implications for future research

While this work focused on AI adoption, the same approach can be extended to monitor the emergence and diffusion of new and frontier technologies across sectors and geographies. Future research could adapt the detection rubric and embedding strategies to track technologies such as quantum computing, synthetic biology, green hydrogen, or digital twins, providing early signals of innovation hotspots and market shifts.

In addition, this methodology opens avenues for analyzing the evolution of technological discourse over time. By comparing website content across years, researchers can detect changes in the prominence, framing, and co-occurrence of technological concepts, enabling a dynamic view of innovation cycles and hype curves. This approach can also help identify first movers and trendsetters by flagging firms that consistently reference cutting-edge technologies before they become mainstream.

From a policy perspective, the application of LLMs for technology mining supports evidence-based decision-making. Innovation agencies and governments can leverage this methodology to monitor real-time adoption of priority technologies, evaluate the impact of policy interventions, and detect sectoral or regional gaps in technological uptake. The ability to extract signals from open web data complements official statistics and enables anticipation of technological transitions ahead of traditional reporting mechanisms.

As an initial step in this direction, "Appendix 4" presents a bag-of-words analysis comparing product- and process-related AI use, highlighting distinct lexical patterns. Future work could expand on this by incorporating topic modeling techniques as a complementary validation layer, in line with recent advances in web-based innovation research.

Finally, further research could delve deeper into product-oriented classifications and explore how varying intensities and forms of innovation are reflected in web-based discourse. Combining this approach with alternative data sources—such as patent filings or firm-level surveys—would enhance the robustness and validity of innovation measurement frameworks.

Conclusion

This study presents a novel methodology for detecting AI adoption in companies through the integration of Large Language Models and web mining techniques. By analyzing the content of corporate websites for 118,605 Spanish firms, we provide a scalable and replicable approach to monitoring AI adoption in real-time. Our framework combines semantic embeddings, chunk-based retrieval, rubric-guided classification, and stratified manual validation, offering an alternative to traditional survey-based methods and enhancing our understanding of how firms communicate innovation externally.

The results reveal several important patterns. Firm size is an important determinant of AI adoption. Larger firms are substantially more likely to adopt and disclose AI technologies, both in 2023 and 2025. They also show broader functional integration, particularly in areas such as administration, product development, and marketing. Sectoral differences are pronounced. AI is most prevalent in digital- and knowledge-intensive industries such as Information and Communications, Professional Services, and Education. Conversely, traditional sectors like Agriculture, Construction, and Hospitality show lower and slower uptake. AI is more commonly used in internal processes than in products. This suggests that companies find it easier to integrate AI into existing operations rather than embed it into market-facing innovations, partly due to the sectoral nature of some firms and the lower implementation complexity of process AI. When AI is applied to products, it often represents a radical innovation. Over two-thirds of AI product cases are classified as radical, highlighting the transformative potential of AI when used in external offerings. Geographic distribution of AI adoption is becoming more balanced. Between 2023 and 2025, we observe greater diffusion across Spanish provinces, indicating that regional gaps in technological uptake may be narrowing.

Finally, this framework can contribute to technology foresight and strategic intelligence by identifying weak signals and mapping early-stage ecosystems. Combined with additional data sources—such as funding data, scientific publications, or startup activity—LLM-based tech mining can become a cornerstone tool for shaping forward-looking innovation policy and guiding public and private investment strategies.

Appendix 1: Complementary information

See Table 7.

Table 7 AI adoption rates across industries and firm sizes in 2023

	Large	Medium	Small	Micro	Total
Administrative and support services	0.057	0.033	0.046	0.027	0.033
Agriculture, livestock, forestry, and fishing	*	0.022	0.014	0.022	0.019
Artistic, recreational, and entertainment activities	*	0.032	0.023	0.015	0.018
Construction	*	0.034	0.019	0.013	0.016
Education	*	0.075	0.069	0.036	0.048
Electricity, gas, steam, and air conditioning supply	*	*	0.014	0.016	0.015
Extractive industries	*	*	0.000	0.010	0.004
Financial and insurance activities	*	0	0.069	0.054	0.056
Healthcare and social services	*	0.038	0.045	0.030	0.035
Hospitality	*	0.006	0.005	0.006	0.005
Information and communications	*	0.485	0.307	0.129	0.190
Manufacturing industry	0.087	0.045	0.021	0.011	0.017
Other services	*	0.130	0.029	0.023	0.028
Professional, scientific, and technical activities	*	0.265	0.109	0.044	0.061
Real estate activities	*	0.111	0.025	0.023	0.023
Transportation and storage	*	0.039	0.014	0.009	0.013
Water supply, sanitation, waste management, and decontamination	*	0.059	0.023	0.004	0.017
Wholesale and retail trade; vehicle repair	0.118	0.051	0.036	0.024	0.028
Total	0.119	0.076	0.043	0.029	0.035

Values represent the proportion of firms in each category using AI

*Indicates sectors with fewer than 15 firms

Appendix 2: LLM Prompt instructions

AI implementation evaluation on websites

You will receive a text fragment extracted from a website. Your task is to **identify** and **evaluate** the presence and use of Artificial Intelligence (AI) in the content, applying the criteria and examples outlined below.

AI detection in content

Determine if the content mentions or implies the use of AI, avoiding false positives. To do this, review the following aspects:

Direct References: Identify key terms such as ‘artificial intelligence’, ‘machine learning’, ‘deep learning’, ‘neural networks’, ‘automatic learning’, ‘cognitive computing’, ‘generative AI’, ‘generative models’, and ‘AI algorithms’.

Conversational Systems: Chatbots, virtual assistants, automatic response generation. Example: “Our chatbot answers queries in seconds with generative AI.”

AI News: Announcements of internal developments, investments in AI. Example: “Our company has developed a new AI model for financial analysis.”

Applications and Services: AI-based personalization, computer vision, predictive analytics, recommendation systems. Example: “We use AI to predict market demand.”

Technological Innovations: Use of LIDAR, GANs, RPA, NLP, continuous learning. Example: “We use GANs to generate visual content” or “We use NLP to process text information.”

Specific Models: Mentions of models like GPT-4, BERT, DALL·E, Stable Diffusion, etc. Example: “Our system uses GPT-4 to automatically analyze legal documents.”

Decision Automation: Use of AI in decision-making processes. Example: “We optimize banking credit with machine learning models.”

Avoid false positives: - Do not assume the company uses AI just because the concept is mentioned generically. - Look for clear evidence of implementation in products or processes.

If there is insufficient evidence, indicate that AI was not detected.

AI implementation evaluation

If AI presence is detected, proceed to classify it according to the following criteria:

Key definitions

To avoid ambiguity, the following definitions are established:

- **Product:** AI-based solutions designed to be marketed and used externally by clients or end-users. Examples: smart devices, autonomous vehicles, AI-as-a-service platforms.
- **Process:** AI applications implemented internally to improve efficiency, automate tasks, or optimize operations. Examples: internal chatbots, data analysis systems, or administrative process automation.

[AI1] Use of artificial intelligence in the company (1 or 0)

Is the use of AI technologies in products and/or processes evidenced by the information available on the website? - 1 (Yes): AI use is identified. - **0 (No):** No clear evidence of AI use. *Note* If AI mention is ambiguous, assign **0**. Furthermore, if AI1 = 0, the following metrics (AI2, AI3, AI4, AI5) must be set to **0**.

[AI2] Implementation level in processes (1 or 0) (if AI1 = 1)

Is the use of AI in internal company processes mentioned or inferred? Consider ‘processes’ to be applications aimed at improving efficiency, automating routine tasks, or optimizing operations. Example: internal chatbots, predictive analytics in logistics, or recommendation systems for administrative processes. - **1 (Yes):** AI integration in internal processes is evident. - **0 (No):** No conclusive evidence found. If information is ambiguous, assign **0**.

[AI3] Application area (if AI2 = 1)

If AI2 = 1, classify the AI implementation in the corresponding area using the following codes and examples:

PGS: Goods or Services Production Example: Industrial automation with robots and AI systems, predictive maintenance using sensor analysis, quality control with computer vision.

DL: Distribution and Logistics Example: Route optimization with AI, inventory management with machine learning, warehouse automation with robots and drones, real-time tracking.

MS: Marketing and Sales Example: Personalized advertising with AI, chatbots for customer support, trend analysis with machine learning, advertising campaign automation.

ICS: Information and Communication Systems Example: Cybersecurity with AI, virtual assistants for technical support, system performance optimization, data processing with AI.

AM: Administration and Management Example: Predictive financial analysis, accounting automation, AI-based recruitment, automatic document management, AI-based management assistants.

PDB: Product and Business Process Development Example: AI-assisted product design, virtual simulation and testing, business process optimization, AI-assisted software development.

For each column, indicate, in binary form, whether AI implementation is detected (1) or not (0) in the respective area. It is essential that, if AI2 = 1, at least one of these fields is marked as 1; otherwise, AI2 should be assigned 0.

[AI4] Use in products (1 or 0) (if AI1 = 1)

Is AI part of a product or service offered? (Consider 'product' to be AI-based solutions designed for clients or end-users). - **1 (Yes):** AI is an integral component of the product or service. Example: a smart device with facial recognition or an autonomous vehicle. - **0 (No):** No AI application in products or insufficient information. If information is ambiguous, assign **0**.

[AI5] Product type (if AI4 = 1)

If AI4 = 1, determine whether the product depends essentially on AI for its operation or if AI is just an added value:

0: Incremental: Gradual improvements that do not drastically change the market but maintain competitiveness. These products use AI as a complement, improving efficiency, functionality, or user experience without completely transforming the market. Example: a surveillance camera using AI for facial detection but still recording video without it. Other examples include: - Recommendation systems. - Interface improvements. - AI features that optimize existing tasks.

1: Radical: Disruptive changes that create new industries or significantly transform existing ones. Example: the development of autonomous vehicles that use AI to operate without human intervention, creating new forms of transportation and entirely new markets.

If AI4 = 0 or information does not allow conclusions, assign **0**.

Expected output

The response must be exclusively in table format with the following columns:

- **AI1:** 0/1: Is AI used in the company?
- **AI2:** 0/1: Is AI use detected in internal processes?
- **PGS:** 0/1: Is AI used in Goods or Services Production?
- **DL:** 0/1: Is AI used in Distribution and Logistics?
- **MS:** 0/1: Is AI used in Marketing and Sales?
- **ICS:** 0/1: Is AI used in Information and Communication Systems?
- **AM:** 0/1: Is AI used in Administration and Management?
- **PDB:** 0/1: Is AI used in Product and Business Process Development?
- **AI4:** 0/1: Is AI used in a product or service?
- **AI5:** 0/1: Product Category (0 if not applicable or incremental, 1 for radical)

Do not include any additional text outside of the table.

Appendix 3: Examples of chunks classified as process/product innovation

Category	Description
AI process	"Route planning is **dynamically optimized using maps and sales forecasts** to identify high-priority machines. This reduces unnecessary visits by up to 30%"
AI process	"Manual processes are prone to errors and consume valuable time that could be dedicated to strategic tasks. AI Solution: Robotic Process Automation (RPA) uses **AI-powered bots to handle repetitive tasks, such as data management or email dispatch**, allowing employees to focus on higher-value activities"
AI product	"Our solutions are built on deep knowledge of big data technologies. They are scalable, secure, and capable of handling complex datasets. From generative AI and natural language processing to machine learning and predictive analytics, **our AI-powered solutions are designed to unlock new opportunities and optimize operations**"
AI product	"Our integrated Data Fabric platform **unifies all your information and artificial intelligence** technology in one place. **Our mission is to help companies maximize the value of their data**, establishing ourselves as leaders in digital transformation"

The original texts were in Spanish and have been translated into English for illustration purposes

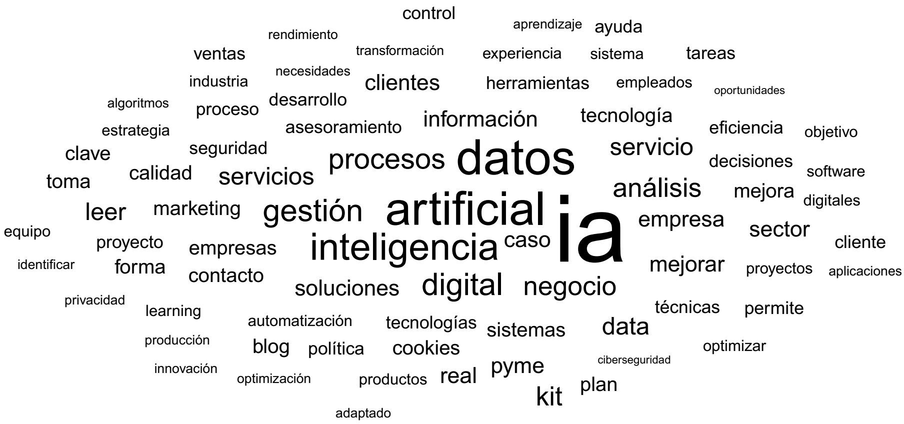


Fig. 8 Word cloud for process-related AI use

Fig. 9 Word cloud for product/services-related AI use

Appendix 4: Topic modeling exploration

This "Appendix" provides an exploratory linguistic and thematic analysis of how Spanish firms describe artificial intelligence on their corporate websites. The objective is to complement the quantitative classification results with qualitative and probabilistic text analysis, offering deeper insight into the narratives through which firms articulate AI adoption. We combine descriptive visualization and topic modeling techniques to examine how the discourse differs across two dimensions of adoption: (i) internal, process-oriented applications (AI2) and (ii) external, product- and service-oriented applications (AI4). Together, these approaches help to uncover the cognitive and communicative structures underpinning the diffusion of AI across organizational boundaries.

Table 8 Summary of model selection metrics for AI2 (internal-process AI adoption)

K	Coherence	ARI	Perplexity
4	0.128	0.688	2041.14
50	0.185	0.444	1148.95
100	0.174	0.458	1030.86
200	0.137	0.524	969.41

Table 9 Summary of model selection metrics for AI4 (product- and service-related AI adoption)

K	Coherence	ARI	Perplexity
4	0.048	0.216	835.16
50	0.184	0.673	495.37
100	0.168	0.650	476.09
200	−0.007	0.772	492.26

(A) Word Cloud

As a first step, we conducted a complementary exploratory analysis by generating comparative word clouds from AI-related fragments in processes (AI2) and products/services (AI4). The visualizations reveal a clear lexical divergence: process-related texts emphasize automation, efficiency, and workflow, pointing to internal optimization, while product-related texts highlight recommendation, personalization, and customer, reflecting user-facing innovation. This contrast reinforces our classification framework and provides a basis for deeper topic modeling (Figs. 8, 9).

(B) Topic Modeling using Latent Dirichlet Allocation (LDA)

To deepen the analysis of AI-related language on company websites, we applied Latent Dirichlet Allocation to the previously classified text fragments. LDA is an unsupervised probabilistic model that detects hidden thematic structures by treating each document as a combination of topics, and each topic as a distribution of words.

We evaluated multiple topic configurations and compared their performance using standard indicators of model quality and thematic stability for both internal-process AI adoption (AI2) and product- and service-related AI adoption (AI4). In both cases, models with $K = 50$ and $K = 100$ achieved the strongest overall balance between semantic coherence and clustering consistency, whereas smaller specifications such as $K = 4$ produced broader but more general thematic structures that captured dominant narratives with limited differentiation. Larger models (e.g., $K \geq 200$) yielded lower perplexity and higher cluster stability but led to thematic fragmentation and overlapping semantic fields that reduced interpretability. Based on this evidence, we retained $K = 100$ as the preferred extended configuration to explore sub-thematic variation with sufficient granularity, and the resulting topic structures are summarised in the tables below (Tables 8, 9).

The topic modeling over AI2 and AI4 reveals two complementary dimensions of AI-related discourse among Spanish firms. The most probable topics, such as *Solutions/Services* and *Data Intelligence*, suggest a stable and overarching narrative focused on technological capabilities and service provision. Within AI2, topics like *Datos IA*, *IA Datos*,

Table 10 Excerpt of topic model based on AI-related paragraphs

Topic	Prob	Prob_AI2	Prob_AI4	AutoTopic	TopTerms
18	0.0262	0.0262	0.0253	Solutions Services	solutions, services, technology, intelligence, project, projects, technologies, data, contact, artificial
72	0.0259	0.0278	0.0003	Datos IA	datos, ia, gestión, información, dato, artificial, técnicas, inteligencia, análisis, pyme
17	0.0211	0.0225	0.0009	IA Datos	ia, datos, inteligencia, patrones, artificial, aprendizaje, algoritmos, automático, análisis, sistemas
10	0.0209	0.0211	0.0184	Data Intelligence	data, intelligence, artificial, time, operational, systems, customer, processes, process, technology
50	0.0190	0.0198	0.0073	IA Marketing	ia, marketing, clientes, campañas, contenido, artificial, inteligencia, estrategias, personalización, marcas
64	0.0183	0.0195	0.0017	Negocio Estrategia	negocio, estrategia, inteligencia, rendimiento, mejorar, oportunidades, artificial, datos, actual, análisis
79	0.0179	0.0192	0.0007	Procesos Negocio	procesos, negocio, proceso, mejora, áreas, empresariales, producción, optimización, caso, asesoramiento
70	0.0174	0.0167	0.0263	Política Privacidad	política, privacidad, legal, aviso, cookies, contacto, acepto, derechos, servicios, teléfono
25	0.0167	0.0179	0.0016	Plan Artificial	plan, artificial, inteligencia, servicio, asesoramiento, pyme, inversión, básico, adaptado, objetivo
68	0.0151	0.0160	0.0031	Decisiones Datos	decisiones, datos, toma, información, tomar, modelos, análisis, informadas, inteligencia, artificial
24	0.0144	0.0145	0.0137	Proyecto Desarrollo	proyecto, desarrollo, objetivo, investigación, proyectos, innovación, empresas, programa, tecnologías, financiado
62	0.0139	0.0140	0.0124	Servicios Contacto	servicios, contacto, soluciones, blog, inicio, consultoría, cloud, casos, noticias, éxito
73	0.0138	0.0141	0.0092	Energía Gestión	energía, gestión, energética, eficiencia, consumo, sistemas, mantenimiento, inteligentes, sostenibilidad, renovables
69	0.0138	0.0127	0.0289	Amb IA	amb, ia, artificial, més, els, gestió, procesos, dels, dades, clients
29	0.0136	0.0142	0.0056	Visión Artificial	visión, artificial, calidad, sistemas, control, producción, industrial, automatización, industria, soluciones
4	0.0136	0.0139	0.0092	Cliente Experiencia	cliente, experiencia, clientes, necesidades, soluciones, servicio, información, ofrecer, gestión, permite
54	0.0135	0.0143	0.0025	Industria Artificial	industria, artificial, inteligencia, empresas, soluciones, procesos, tecnología, servicios, ia, proyectos
16	0.0128	0.0136	0.0024	Ventas Digitales	ventas, digitales, digital, empleados, pyme, técnicas, asesoramiento, servicio, área, utilizando
85	0.0128	0.0135	0.0032	Desafíos IA	desafíos, ia, tecnologías, soluciones, sistemas, tecnología, sostenibilidad, implementación, eficiencia, mercado
84	0.0127	0.0126	0.0138	Logística Papel	logística, papel, inteligencia, empresas, artificial, tecnología, crucial, sostenibilidad, eficiencia, optimizar
2	0.0125	0.0106	0.0383	IA Lenguaje	ia, lenguaje, natural, contenido, chatgpt, herramientas, procesamiento, generar, texto, imágenes
6	0.0124	0.0128	0.0060	Seguridad Ciberseguridad	seguridad, ciberseguridad, amenazas, inteligencia, artificial, sistemas, protección, datos, proteger, ataques
9	0.0122	0.0123	0.0106	Marketing Digital	marketing, digital, seo, google, online, social, diseño, media, sem, posicionamiento

Table 10 (continued)

Topic	Prob	Prob_AI2	Prob_AI4	AutoTopic	TopTerms
86	0.0121	0.0130	0.0008	Producción Procesos	producción, procesos, automatización, inteligencia, ia, artificial, precisión, mantenimiento, empresas, reduciendo
13	0.0121	0.0114	0.0225	Sistema Real	sistema, real, mantenimiento, sistemas, control, inteligente, permite, sensores, predictivo, detección
44	0.0120	0.0124	0.0072	IA Profesionales	ia, profesionales, inteligencia, artificial, mundo, máster, herramientas, sector, tecnología, formación
91	0.0118	0.0124	0.0046	Kit Digital	kit, digital, consulting, programa, ayudas, servicios, pymes, empresas, transformación, artificial
58	0.0117	0.0123	0.0032	Digital Transformación	digital, transformación, digitales, tecnologías, herramientas, digitalización, objetivos, optimizar, estrategia, desarrollo
97	0.0114	0.0115	0.0106	Leer IA	leer, ia, inteligencia, artificial, mundo, transformando, sector, artículo, empresas, blog
36	0.0113	0.0117	0.0048	Datos Análisis	datos, análisis, avanzado, aprendizaje, extracción, procesos, automático, sistema, mejora, inteligencia
99	0.0112	0.0118	0.0030	Asesoramiento Servicio	asesoramiento, servicio, ciberseguridad, básico, análisis, avanzado, artificial, inteligencia, datos, negocio
22	0.0112	0.0110	0.0136	IA Sector	ia, sector, inteligencia, artificial, evento, tecnología, empresas, soluciones, tecnológicas, organización
75	0.0109	0.0108	0.0126	Tareas Repetitivas	tareas, repetitivas, automatización, productividad, ia, procesos, recursos, humanos, costes, eficiencia
77	0.0109	0.0107	0.0130	Cookies Política	cookies, política, sitio, aceptar, privacidad, información, experiencia, navegación, utiliza, utilizamos
20	0.0108	0.0114	0.0018	Business Intelligence	business, intelligence, analítica, gestión, ia, procesos, asociada, clientes, inteligencia, categoría
96	0.0102	0.0108	0.0014	Data Gestión	data, gestión, automatización, procesos, rpa, cloud, automation, aplicaciones, datos, negocio
89	0.0102	0.0098	0.0149	Cookies Policy	cookies, policy, privacy, experience, reality, contact, accept, artificial, future, legal
83	0.0100	0.0101	0.0088	Diseño Proyectos	diseño, proyectos, construcción, arquitectura, impresión, desarrollo, ingeniería, inteligencia, artificial, materiales
19	0.0097	0.0092	0.0164	IoT Dispositivos	iot, dispositivos, internet, nube, datos, seguridad, tecnología, desarrollo, proceso, productos
45	0.0095	0.0096	0.0078	Talento RRHH	talento, rrhh, equipo, laboral, experiencia, crecimiento, constante, personas, gestión, recursos
67	0.0095	0.0098	0.0051	Calidad Desarrollo	calidad, desarrollo, información, control, trazabilidad, plataforma, algoritmos, manipulación, artificial, regional
33	0.0094	0.0097	0.0055	Industrial Automatización	industrial, automatización, industriales, tecnologías, procesos, consentimiento, cookies, robótica, navegación, datos
66	0.0093	0.0075	0.0338	Enero Noviembre	enero, noviembre, octubre, diciembre, junio, mayo, abril, read, artificial, febrero
8	0.0091	0.0086	0.0173	Imágenes IA	imágenes, ia, calidad, forma, búsqueda, artificial, persona, inteligencia, fuente, interacciones
23	0.0091	0.0089	0.0126	Inteligencia Artificial	inteligencia, artificial, ia, viajes, consultores, agencias, proceso, información, nivel, soc

Table 10 (continued)

Topic	Prob	Prob_AI2	Prob_AI4	AutoTopic	TopTerms
15	0.0090	0.0096	0.0017	Ciberseguridad Seguridad	ciberseguridad, seguridad, protección, frente, plan, básica, información, sistema, ia, iso
71	0.0088	0.0095	0.0004	Cadena Suministro	cadena, suministro, demanda, gestión, logística, inventario, previsión, aprendizaje, automático, real
78	0.0088	0.0082	0.0160	Equipo Clientes	equipo, clientes, artificial, expertos, soluciones, inteligencia, sector, servicios, utilizamos, compromiso
55	0.0087	0.0082	0.0155	Comunicación Inteligencia	comunicación, inteligencia, artificial, medios, noticias, comentarios, paso, ia, años, temporada
74	0.0087	0.0093	0.0004	IA Selección	ia, selección, artificial, candidatos, inteligencia, datos, proceso, empresas, amazon, candidato
63	0.0087	0.0077	0.0217	Artificial IA	artificial, ia, inteligencia, generativa, marketing, power, cámara, éxito, eventos, copilot
11	0.0086	0.0091	0.0026	Artificial Inteligencia	artificial, inteligencia, salud, tratamiento, circular, innovación, economía, ia, tratamientos, móstoles
37	0.0085	0.0078	0.0175	Seguridad Monitorización	seguridad, monitorización, inteligencia, artificial, control, sistemas, accesos, real, reconocimiento, detección
39	0.0084	0.0083	0.0100	Artificial Inteligencia	artificial, inteligencia, página, tecnología, director, información, objetos, sistema, seleccionar, espacios
60	0.0084	0.0079	0.0146	Realidad Virtual	realidad, virtual, aumentada, tecnología, inteligencia, artificial, experiencias, entorno, espacio, experiencia
94	0.0083	0.0084	0.0067	Innovación Inteligencia	innovación, inteligencia, tecnologías, tendencias, artificial, mercados, emergentes, empresas, desarrollo, estrategias
14	0.0083	0.0083	0.0080	Software Desarrollo	software, desarrollo, soluciones, informática, sector, erp, servicios, especialistas, medida, artificial
43	0.0083	0.0084	0.0071	Datos Azure	datos, azure, información, ia, base, optimizar, cantidad, fiscal, forma, mejorar
92	0.0082	0.0079	0.0120	LinkedIn Facebook	linkedin, facebook, twitter, instagram, blog, contacto, noticias, youtube, artificial, search
48	0.0081	0.0087	0.0006	Gestión Tareas	gestión, tareas, clientes, comercial, datos, ia, automatización, electrónicos, oportunidades, correos
56	0.0081	0.0046	0.0555	Modelo IA	modelo, ia, artificial, inteligencia, openai, tecnología, deepseek, chatgpt, investigación, millones
32	0.0079	0.0077	0.0112	Redes Sociales	redes, sociales, neuronales, inteligencia, artificial, educación, marcas, social, tendencias, contenido
87	0.0079	0.0061	0.0322	Curso Artificial	curso, artificial, inteligencia, programación, anual, cursos, desarrollo, introducción, online, unidad
93	0.0079	0.0078	0.0094	Salud Seguros	salud, seguros, servicios, ia, sector, atención, soluciones, gestión, productos, pacientes
51	0.0079	0.0081	0.0055	Transporte Servicios	transporte, servicios, control, rutas, información, aplicaciones, empresas, mercancías, flota, gestión
35	0.0079	0.0066	0.0251	IA Fundamental	ia, fundamental, vida, artificial, inteligencia, tecnología, audífonos, convertido, herramienta, logística
95	0.0078	0.0079	0.0066	Agricultura Cultivos	agricultura, cultivos, años, agua, información, sistema, drones, riego, ia, decisión
76	0.0077	0.0079	0.0048	IA Sistemas	ia, sistemas, riesgo, gestión, servicios, seguridad, defensa, integral, presenta, alto
5	0.0077	0.0079	0.0047	Gestión Artificial	gestión, artificial, soluciones, financiera, clientes, información, inteligencia, seguridad, neural, visión

Table 10 (continued)

Topic	Prob	Prob_AI2	Prob_AI4	AutoTopic	TopTerms
21	0.0076	0.0080	0.0022	Resultados Artificial	resultados, artificial, inteligencia, evaluación, ia, servicio, diagnóstico, inicial, técnica, necesidades
40	0.0074	0.0076	0.0048	Machine Learning	learning, machine, deep, desarrollo, equipos, artificial, técnicas, aplicaciones, modelos, blockchain
80	0.0074	0.0077	0.0036	Mercado Tendencias	mercado, tendencias, apoyo, inteligencia, material, robótica, competitiva, artificial, futuro, ventaja
61	0.0074	0.0076	0.0043	Gestión Negocio	gestión, negocio, ia, productos, retail, descubre, quieres, datos, wordpress, optimiza
46	0.0072	0.0068	0.0129	ERP Artificial	erp, artificial, administración, inteligencia, crm, integraciones, información, ia, ley, rpa
98	0.0071	0.0066	0.0144	Datos Chat	datos, chat, personales, tratamiento, dirección, responsable, online, bot, casos, usuarios
52	0.0071	0.0063	0.0173	Electrónico Correo	electrónico, correo, inteligencia, transporte, artificial, logística, nombre, empresas, comentario, comments
38	0.0070	0.0070	0.0081	IA Inteligencia	ia, inteligencia, llamadas, crisis, artificial, cliente, tecnología, agente, artificiales, actions
59	0.0070	0.0067	0.0111	Formación Inteligencia	formación, inteligencia, artificial, cursos, personal, derecho, selección, problema, e-learning, ia
90	0.0069	0.0071	0.0041	Segmento Empleados	segmento, empleados, ia, usuarios, duración, importe, empresas, humano, instructor, máximo
49	0.0069	0.0066	0.0113	Preferencias Gestionar	preferencias, gestionar, usuario, marketing, servicios, opciones, administrar, publicidad, consentimiento, guardar
41	0.0068	0.0069	0.0058	Personas Sitio	personas, sitio, usuario, accesibilidad, interfaz, aplicación, usuarios, utiliza, discapacidad, pantalla
31	0.0067	0.0068	0.0059	España IA	españa, ia, financiero, práctica, noticia, sector, artificial, marca, inteligencia, seguridad
28	0.0067	0.0066	0.0075	Solución Gestión	solución, gestión, procesos, funcionalidades, automatización, flujos, integración, plataformas, ia, deberá
42	0.0066	0.0054	0.0223	IA Inteligencia	ia, inteligencia, artificial, propiedad, formación, gestión, intelectual, diseño, técnico, tecnología
81	0.0065	0.0068	0.0031	Gestión Comercio	gestión, comercio, turismo, electrónico, link, oficina, cultura, electrónica, virtual, patrimonio
34	0.0065	0.0067	0.0025	Datos Almacenamiento	datos, almacenamiento, bases, paneles, creación, estructurados, permitirá, capacitaciones, visuales, exportación
88	0.0064	0.0063	0.0089	Smart Inteligencia	smart, inteligencia, artificial, división, tráfico, servicios, digital, digitalización, defensa, ia
53	0.0064	0.0053	0.0208	Artificial Cookies	artificial, cookies, inteligencia, guardar, activar, activa, visitantes, preferencias, cambios, información
1	0.0063	0.0057	0.0156	IA Admin	ia, admin, público, beneficios, artificial, inteligencia, herramientas, global, especialidades, alumbrado
57	0.0062	0.0057	0.0142	Center Assessoria	center, assessoria, contact, ia, ip, artificial, telefonía, soluciones, inteligencia, seguir
47	0.0059	0.0059	0.0059	Informe Móvil	informe, móvil, asistentes, company, experiencia, información, digital, fibra, panel, empresas
30	0.0059	0.0063	0.0006	Requisitos Oposiciones	requisitos, oposiciones, red, información, área, guía, palibex, sesión, entrega, ia
65	0.0058	0.0059	0.0042	Global Management	global, management, sector, artificial, operador, translation, customer, digital, mobile, innovation

Table 10 (continued)

Topic	Prob	Prob_AI2	Prob_AI4	AutoTopic	TopTerms
3	0.0057	0.0054	0.0106	Aplicada Artificial	aplicada, artificial, inteligencia, especialista, campo, universidad, ciberseguridad, data, consorcio, robot
12	0.0056	0.0053	0.0094	Español English	español, english, inglés, news, contact, spain, quote, menú, tecnología, jamones
100	0.0055	0.0058	0.0017	III IA	iii, ia, centro, forma, inteligencia, artificial, response, herramientas, detection, mantenimiento
26	0.0054	0.0054	0.0041	Madrid Barcelona	madrid, barcelona, españa, diciembre, contacto, horario, jueves, coches, automatic, lunes
27	0.0052	0.0056	0.0006	Vino Artificial	vino, artificial, inteligencia, ia, recursos, gestión, forma, procesos, clave, aribau
7	0.0045	0.0046	0.0026	Noticias Proyectos	noticias, proyectos, participa, ia, nutai, visión, detección, actualidad, perte, cobotvnc
82	0.0041	0.0039	0.0070	Inteligente Jugar	inteligente, jugar, ia, bote, potencia, mes, juega, acceder, lun, juegos

and *Procesos de Negocio* exhibit markedly higher probabilities than in AI4, highlighting a process-oriented view of AI centered on data management, decision-making, and internal efficiency. This pattern reflects a discourse grounded in organizational processes and managerial optimization rather than product communication.

Conversely, AI4 displays stronger associations with market-oriented and web-derived themes, including *IA Lenguaje* (language models), *Modelo IA* (e.g., OpenAI, DeepSeek, ChatGPT), and legal or compliance-related topics such as cookies and privacy policies. These reflect the external projection of AI as a product feature, technological attribute, or regulatory concern. Taken together, the results illustrate a clear distinction between functional adoption—where AI supports internal processes—and communicative adoption—where AI forms part of firms' product and marketing narratives. This duality underscores the multidimensional nature of AI diffusion across organizational boundaries.

The probability values reported in Table 10 correspond to the average topic weights estimated from the LDA model. Specifically, Prob represents each topic's overall prevalence across all relevant paragraphs, while Prob_AI2 and Prob_AI4 indicate its relative importance within the process-oriented and product-oriented subsets, respectively. This information helps to compare the semantic dominance of each topic across both dimensions of AI adoption. This approach resembles that presented in the y material of Dahlke et al. (2024).

Overall, the topic model reinforces the distinction between internal and external AI narratives, revealing how firms simultaneously frame AI as a tool for operational improvement and as a market-facing technological asset.

Appendix 5: LLM classification details

See Table 11.

Table 11 LLM classification configuration

Component	Description
Model	GPT-4o-mini
Platform	OpenAI API
Prompt structure	Custom rubric-based classification prompt (see "Appendix 2")
Task	Binary classification of AI adoption (IA = 1/0), with sublabels for process (AI2) and product (AI4)
Temperature	1.0
Top-p	1.0
Fragments per firm	Top 10 highest-ranked fragments by semantic similarity to AI queries
Chunking configuration	each chunk 1000 bytes; symmetric context window of 1.1 KB
Validation method	Stratified manual validation across firm size, sector, and region
Fine-tuning	None

Acknowledgements This work was supported by the Generalitat Valenciana under Grant CIAICO/2023/272; the Agencia Estatal de Investigación (MCIN/AEI/10.13039/501100011033) and ERDF/EU under Grant PID2023-152106OB-I00; and the Universitat Politècnica de València (PAID-01-24, PAID-12-25).

Declarations

Conflict of interest The authors have no competing interests to declare that are relevant to the content of this article.

References

Abbasiharofteh, M., & Kriesch, L. (2024). No. 2024/16 Not all twins are identical: The digital layer of "twin" transition market applications. Papers in Innovation Studies (2025-07-11). https://journals.lub.lu.se/piis/article/view/27650

Alekseeva, L., Azar, J., Giné, M., Samila, S., & Taska, B. (2021). The demand for AI skills in the labor market. *Labour Economics, 71*, 102002. https://doi.org/10.1016/j.labeco.2021.102002

Axenbeck, J., & Breithaupt, P. (2021). Innovation indicators based on firm websites—Which website characteristics predict firm-level innovation activity? *PLoS ONE, 16*(4), e0249583. https://doi.org/10.1371/journal.pone.0249583

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

Battisti, G., & Stoneman, P. (2003). Inter- and intra-firm effects in the diffusion of new process technology. *Research Policy, 32*(9), 1641–1655. https://doi.org/10.1016/S0048-7333(03)00055-6

Blazquez, D., & Domenech, J. (2018). Big Data sources and methods for social and economic analyses. *Technological Forecasting and Social Change, 130*, 99–113. https://doi.org/10.1016/j.techfore.2017.07.027

Bottai, C., Crosato, L., Domenech, J., Guerzoni, M., & Liberati, C. (2024). Scraping innovativeness from corporate websites: Empirical evidence on Italian manufacturing SMEs. *Technological Forecasting and Social Change, 207*, 123597. https://doi.org/10.1016/j.techfore.2024.123597

Bouschery, S.-G., Blazevic, V., & Piller, F.-T. (2023). Augmenting human innovation teams with artificial intelligence: Exploring transformer-based language models. *Journal of Product Innovation Management*. https://doi.org/10.1111/jpim.12656

Bresnahan, T. (2010). General purpose technologies. In *Handbook of the economics of innovation* (Vol. 2, pp. 761–791). Elsevier. https://doi.org/10.1016/S0169-7218(10)02002-2

Brown, T., Mann, B., Ryder, N., Subbiah, M., Kaplan, J.-D., Dhariwal, P., & Amodei, D. (2020). Language models are few-shot learners. In H. Larochelle, M. Ranzato, R. Hadsell, M. Balcan & H. Lin (Eds.), *Advances in neural information processing systems* (Vol. 33, pp. 1877–1901). Curran Associates, Inc. https://proceedings.neurips.cc/paperfiles/paper/2020/file/1457c0d6bfcb4967418bfb8ac142f64a-Paper.pdf

Bureau van Dijk. (2025). *SABI: Sistema de análisis de balances ibéricos*. https://sabi.bvdinfo.com/

Calvino, F., & Fontanelli, L. (2023). *A portrait of AI adopters across countries: Firm characteristics, assets' complementarities and productivity*. OECD Science, Technology and Industry Working Papers No. 2023/02. https://doi.org/10.1787/0fb79bb9-en

Chang, Y., Wang, X., Wang, J., Wu, Y., Yang, L., Zhu, K., & Xie, X. (2024). A survey on evaluation of large language models. *ACM Transactions on Intelligent Systems and Technology, 15*(3), 1–45. https://doi.org/10.1145/3641289

Chen, J., & Tajdini, S. (2024). A moderated model of artificial intelligence adoption in firms and its effects on their performance. *Information Technology and Management*. https://doi.org/10.1007/s10799-024-00422-5

Chiang, C.-H., Chen, W.-C., Kuan, C.-Y., Yang, C., & Lee, H.-Y. (2024). Large language model as an assignment evaluator: Insights, feedback, and challenges in a 1000+ student course. In *Proceedings of the 2024 conference on empirical methods in natural language processing*, 2924 (pp. 2489–2513). Association for Computational Linguistics. https://doi.org/10.18653/v1/2024.emnlp-main.146

Chicco, D., & Jurman, G. (2020). The advantages of the Matthews correlation coefficient (MCC) over F1 score and accuracy in binary classification evaluation. *BMC Genomics, 21*(1), 6. https://doi.org/10.1186/s12864-019-6413-7

Cockburn, I.-M., Henderson, R., & Stern, S. (2019). The impact of artificial intelligence on innovation: An exploratory analysis. In A. Agrawal, J. Gans, & A. Goldfarb (Eds.), *The economics of artificial intelligence* (pp. 115–148). University of Chicago Press.

Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement, 20*(1), 37–46. https://doi.org/10.1177/001316446002000104

Crosato, L., Domenech, J., & Liberati, C. (2024). Websites' data: A new asset for enhancing credit risk modeling. *Annals of Operations Research, 342*(3), 1671–1686. https://doi.org/10.1007/s10479-023-05306-5

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

Daas, P.-J., & Van Der Doef, S. (2020). Detecting innovative companies via their website. *Statistical Journal of the IAOS*. https://doi.org/10.3233/SJI-200627

Dahlke, J., Beck, M., Kinne, J., Lenz, D., Dehghan, R., Wörter, M., & Ebersberger, B. (2024). Epidemic effects in the diffusion of emerging digital technologies: Evidence from artificial intelligence adoption. *Research Policy, 53*(2), 104917. https://doi.org/10.1016/j.respol.2023.104917

Dahlke, J., Schmidt, S., Lenz, D., Kinne, J., Dehghan, R., Abbasiharofteh, M., Schütz, M., Kriesch, L., Hottenrott, H., Neftakanilmaz, U., Grashof, N., Hajikhani, A., Liu, L., Riccaboni, M., Balland, P.-A., Wörter, M., & Rammer, C. (2025). The WebAI paradigm of innovation research: Extracting insight from organizational web data through AI. Technical Report. ZEW Discussion Papers.

Devlin, J., Chang, M.-W., Lee, K., & Toutanova, K. (2019). BERT: Pre-training of deep bidirectional transformers for language understanding. https://doi.org/10.48550/arXiv.1810.04805.

Dewar, R.-D., & Dutton, J.-E. (1986). The adoption of radical and incremental innovations: An empirical analysis. *Management Science, 32*(11), 1422–1433. https://doi.org/10.1287/mnsc.32.11.1422

European Commission. (2003). Commission recommendation of 6 May 2003 concerning the definition of micro, small and medium-sized enterprises. *Official Journal of the European Union, L, 124*, 36–41. https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A32003H0361.

European Commission. (Ed.). (2008). *NACE Rev. 2: Statistical classification of economic activities in the European Community*. Publications Office.

European Commission. (2024). *Regions in the European Union: Nomenclature of territorial units for statistics (NUTS)*. Publications Office. https://doi.org/10.2785/714519

Gao, Y., Xiong, Y., Gao, X., Jia, K., Pan, J., Bi, Y., & Wang, H. (2024). Retrieval-augmented generation for large language models: A survey.https://doi.org/10.48550/arXiv.2312.10997

Gök, A., Waterworth, A., & Shapira, P. (2015). Use of web mining in studying innovation. *Scientometrics, 102*, 653–671. https://doi.org/10.1007/s11192-014-1434-0

Gu, J., Jiang, X., Shi, Z., Tan, H., Zhai, X., Xu, C., & Guo, J. (2025). A survey on LLM-as-a-judge. https://doi.org/10.48550/arXiv.2411.15594

Hajikhani, A., & Cole, C. (2024). A critical review of large language models: Sensitivity, bias, and the path toward specialized AI. *Quantitative Science Studies, 5*(3), 736–756. https://doi.org/10.1162/qss_a_00310

Héroux-Vaillancourt, M., Beaudry, C., & Rietsch, C. (2020). Using web content analysis to create innovation indicators—What do we really measure? *Quantitative Science Studies, 1*(4), 1601–1637. https://doi.org/10.1162/qssa00086

Hollenstein, H., & Woerter, M. (2008). Inter- and intra-firm diffusion of technology: The example of E-commerce. *Research Policy, 373*, 545–564. https://doi.org/10.1016/j.respol.2007.12.006

Huang, J., Yang, D.-M., Rong, R., Nezafati, K., Treager, C., Chi, Z., & Xie, Y. (2024). A critical assessment of using ChatGPT for extracting structured data from clinical notes. *NPJ Digital Medicine, 7*(1), 106. https://doi.org/10.1038/s41746-024-01079-8

Huang, L., Yu, W., Ma, W., Zhong, W., Feng, Z., Wang, H., & Liu, T. (2025). A survey on hallucination in large language models: Principles, taxonomy, challenges, and open questions. *ACM Transactions on Information Systems, 432*, 1–55. https://doi.org/10.1145/3703155

Hunt, J., Cockburn, I., & Bessen, J. (2024). *Is distance from innovation a barrier to the adoption of artificial intelligence?* Technical Report No. w33022 (2025-04-14). National Bureau of Economic Research. https://doi.org/10.3386/w33022

Jonsson, A., & Svingby, G. (2007). The use of scoring rubrics: Reliability, validity and educational consequences. *Educational Research Review, 22*, 130–144. https://doi.org/10.1016/j.edurev.2007.05.002

Kinne, J., & Axenbeck, J. (2020). Web mining for innovation ecosystem mapping: A framework and a large-scale pilot study. *Scientometrics, 125*(3), 2011–2041. https://doi.org/10.1007/s11192-020-03726-9

Kinne, J., & Lenz, D. (2021). Predicting innovative firms using web mining and deep learning. *PLoS ONE, 164*, e0249071. https://doi.org/10.1371/journal.pone.0249071

Koster, M., Illyes, G., Zeller, H., & Sassman, L. (2022). Robots exclusion protocol. RFC Editor. https://doi.org/10.17487/RFC9309

Landis, J.-R., & Koch, G.-G. (1977). The measurement of observer agreement for categorical data. *Biometrics, 331*, 159. https://doi.org/10.2307/2529310

Lee, M.-C., Scheepers, H., Lui, A.-K., & Ngai, E.-W. (2023). The implementation of artificial intelligence in organizations: A systematic literature review. *Information and Management, 605*, 103816. https://doi.org/10.1016/j.im.2023.103816

Lee, Y.-S., Kim, T., Choi, S., & Kim, W. (2022). When does AI pay off? AI-adoption intensity, complementary investments, and R &D strategy. *Technovation, 118*, 102590. https://doi.org/10.1016/j.technovation.2022.102590

Lewis, P., Perez, E., Piktus, A., Petroni, F., Karpukhin, V., Goyal, N., & Kiela, D. (2020). Retrieval-augmented generation for knowledge-intensive NLP tasks. In *34th Conference on neural information processing systems (NeurIPS 2020)*, Vancouver, Canada, 2020.

Liu, N.-F., Lin, K., Hewitt, J., Paranjape, A., Bevilacqua, M., Petroni, F., & Liang, P. (2023). Lost in the middle: How language models use long contexts. *Transactions of the Association for Computational Linguistics*. https://doi.org/10.48550/arXiv.2307.03172

Makridakis, S. (2017). The forthcoming Artificial Intelligence (AI) revolution: Its impact on society and firms. *Futures, 90*, 46–60. https://doi.org/10.1016/j.futures.2017.03.006

Mazzoni, L., Pinelli, F., & Riccaboni, M. (2024). Measuring corporate digital divide through websites: Insights from Italian firms. *EPJ Data Science, 131*, 51. https://doi.org/10.1140/epjds/s13688-024-00491-0

McElheran, K., Li, J.-F., Brynjolfsson, E., Kroff, Z., Dinlersoz, E., Foster, L., & Zolas, N. (2024). AI adoption in America: Who, what, and where. *Journal of Economics and Management Strategy, 332*, 375–415. https://doi.org/10.1111/jems.12576

Messick, S. (1995). Validity of psychological assessment: Validation of inferences from persons' responses and performances as scientific inquiry into score meaning. *American Psychologist, 509*, 741–749. https://doi.org/10.1037/0003-066X.50.9.741

Mikolov, T., Chen, K., Corrado, G., & Dean, J. (2013). Efficient estimation of word representations in vector space. https://doi.org/10.48550/arXiv.1301.3781

Nunnally, J.-C., & Bernstein, I.-H. (1994). *Psychometric theory* (3rd ed.). McGraw-Hill.

OECD, & Eurostat. (2018). *Oslo Manual 2018: Guidelines for collecting, reporting and using data on innovation* (4th ed.). OECD (2025-02-25). https://doi.org/10.1787/9789264304604-en

Panadero, E., & Jonsson, A. (2013). The use of scoring rubrics for formative assessment purposes revisited: A review. *Educational Research Review, 9*, 129–144. https://doi.org/10.1016/j.edurev.2013.01.002

Peters, M.-E., Neumann, M., Iyyer, M., Gardner, M., Clark, C., Lee, K., & Zettlemoyer, L. (2018). Deep contextualized word representations (2025-01-28). https://doi.org/10.48550/arXiv.1802.05365

Porter, A.-L., Zhang, Y., & Newman, N.-C. (2024). Tech mining: A revisit and navigation. *Frontiers in Research Metrics and Analytics, 9*, 1364053. https://doi.org/10.3389/frma.2024.1364053

Powers, D.-M.-W. (2020). Evaluation: from precision, recall and F-measure to ROC, informedness, markedness and correlation (2025-03-04). https://doi.org/10.48550/ARXIV.2010.16061

Qin, S., Nelson, L., McLeod, L., Eremenco, S., & Coons, S.-J. (2019). Assessing test–retest reliability of patient-reported outcome measures using intraclass correlation coefficients: Recommendations for selecting and documenting the analytical formula. *Quality of Life Research, 284*, 1029–1033. https://doi.org/10.1007/s11136-018-2076-0

Rammer, C., & Es-Sadki, N. (2023). Using big data for generating firm-level innovation indicators—A literature review. *Technological Forecasting and Social Change, 197*, 122874. https://doi.org/10.1016/j.techfore.2023.122874

Russell, S., & Norvig, P. (2016). *Artificial intelligence: A modern approach* (3rd ed.). Pearson.

Savelka, J., & Ashley, K.-D. (2023). The unreasonable effectiveness of large language models in zero-shot semantic annotation of legal texts. *Frontiers in Artificial Intelligence, 6*, 1279794. https://doi.org/10.3389/frai.2023.1279794

Schumpeter, J.-A., & Swedberg, R. (2013). *Capitalism, socialism and democracy*. Routledge.

Stornelli, A., Ozcan, S., & Simms, C. (2021). Advanced manufacturing technology adoption and innovation: A systematic literature review on barriers, enablers, and innovation types. *Research Policy, 50*(6), 104229. https://doi.org/10.1016/j.respol.2021.104229

Tegmark, M. (2018). *Life 3.0: Being human in the age of artificial intelligence* (First Vintage books edition). Vintage Books.

Ulucanlar, S., Faulkner, A., Peirce, S., & Elwyn, G. (2013). Technology identity: The role of sociotechnical representations in the adoption of medical devices. *Social Science and Medicine, 98*, 95–105. https://doi.org/10.1016/j.socscimed.2013.09.008

Yang, J., Blount, Y., & Amrollahi, A. (2024). Artificial intelligence adoption in a professional service industry: A multiple case study. *Technological Forecasting and Social Change, 201*, 123251. https://doi.org/10.1016/j.techfore.2024.123251

Zhang, Y., Porter, A.-L., & Chiavetta, D. (2017). Scientometrics for tech mining: An introduction. *Scientometrics, 1113*, 1875–1878. https://doi.org/10.1007/s11192-017-2344-8