

# Discussion Paper Series

IZA DP No. 18434

March 2026

## The Diffusion of Artificial Intelligence Across Firms: Evidence from Europe

**Julio Garbers**

Luxembourg Institute of Socio-Economic Research (LISER)  
and University of Luxembourg

**Terry Gregory**

Luxembourg Institute of Socio-Economic Research (LISER),  
ZEW and IZA@LISER

The IZA Discussion Paper Series (ISSN: 2365-9793) ("Series") is the primary platform for disseminating research produced within the framework of the IZA Network, an unincorporated international network of labour economists coordinated by the Luxembourg Institute of Socio-Economic Research (LISER). The Series is operated by LISER, a Luxembourg public establishment (établissement public) registered with the Luxembourg Business Registers under number J57, with its registered office at 11, Porte des Sciences, 4366 Esch-sur-Alzette, Grand Duchy of Luxembourg.

Any opinions expressed in this Series are solely those of the author(s). LISER accepts no responsibility or liability for the content of the contributions published herein. LISER adheres to the European Code of Conduct for Research Integrity. Contributions published in this Series present preliminary work intended to foster academic debate. They may be revised, are not definitive, and should be cited accordingly. Copyright remains with the author(s) unless otherwise indicated.



# The Diffusion of Artificial Intelligence Across Firms: Evidence from Europe\*

## Abstract

We develop a novel firm-level indicator of Artificial Intelligence adoption in Europe (MAP-AI) by extracting information on AI usage from more than three million firm websites from Belgium, France, Germany, and Luxembourg (2016–2024) using a Large Language Model. The indicator captures realized AI adoption as signaled on their website rather than potential exposure. Our method allows to detect not only whether firms adopt AI, but also their role in the AI ecosystem and the type of AI technology they employ. Validation against human-coded benchmarks and external referenecs confirms high accuracy and external validity. We find that the share of AI-active firms grew from 1% in 2016 to 12% in 2024, with acceleration after 2022. We document a structural transformation in the AI ecosystem, as expanding AI adoption increases the share of adopters in overall AI activity, signaling widespread diffusion and more integrated AI use, including generative AI. While adoption is concentrated among larger, younger, knowledge-intensive firms in urban innovation clusters, workforce skills emerge as a key factor associated with AI adoption. Our skill-level analysis shows that foundational data skills form a necessary base for adoption, while a small set of specialized AI skills – such as machine learning and natural language processing – act as strong complements, highlighting human capital as a central driver of AI diffusion across firms.

## JEL classification

O33, C81, L25

## Keywords

Artificial Intelligence, firm-level data, Large Language Models, AI diffusion, human capital, skills

## Corresponding author

Julio Garbers

[julio.garbers@liser.lu](mailto:julio.garbers@liser.lu)

---

\* Julio Garbers gratefully acknowledges financial support from Luxembourg National Research Fund (FNR) under the ACROSS Doctoral Training Unit (grant no. PRIDE19/14302992). This project received financial support from the LISER Innovation call grant. We are grateful to Christina Gathmann, Albrecht Glitz, Skerdilajda Zanaj, Jens Ruhose, and Frédéric Docquier for many helpful comments and suggestions. We thank Etienne Bacher for his outstanding research assistance.

---

# 1 Introduction

Artificial Intelligence (AI) is becoming increasingly widespread across the economy, thanks in part to advances in generative AI over the past few years. It remains unclear how this will affect the economy. While some studies find hardly any significant effects on employment (Hampole et al., 2025), others point to initial signs of negative effects (Brynjolfsson et al., 2025). The theory indicates that it will depend on how firms integrate AI into their operations, products, and services (Agrawal et al., 2022, Acemoglu and Johnson, 2023). AI adoption is also likely to be highly uneven: while some firms and sectors may quickly integrate these tools, others may delay adoption – for instance, because of a lack of workforce skills – potentially widening productivity and competitiveness gaps. However, despite rapid technological progress, empirical evidence on AI’s economic effects remains limited, largely because AI adoption itself is difficult to observe at scale, especially at the firm level (Seamans and Raj, 2019). Most existing measures capture potential exposure to AI rather than firms’ realized use of AI technologies. This measurement gap limits our ability to study diffusion, heterogeneity, and the economic consequences of AI adoption, especially in Europe.

This paper develops a new firm-level indicator of AI adoption in Europe based on large-scale analysis of firms’ web content using a Large Language Model (LLM). We introduce MAP-AI (Mapping Artificial Intelligence in Firms), a novel dataset that captures not only whether and when firms adopt AI, but also their role in the AI ecosystem as well as the specific AI technologies they use. By leveraging unstructured web data, the measure overcomes key limitations of survey-based and exposure-based approaches, providing a scalable and up-to-date proxy for realized AI adoption at the firm level. The indicator thus lays the foundation for future causal and structural analyses of AI’s economic effects that have so far been constrained by data limitations. The MAP-AI data supporting the findings of this study are publicly available for use by researchers, policy-makers, and practitioners.<sup>1</sup>

Our analysis proceeds in three steps that leverage the new AI adoption measure to study diffusion, heterogeneity, and the determinants of AI adoption across European firms, with a special focus on skills.

In a *first* step, to build our novel firm-level AI adoption measure (MAP-AI), we exploit a dataset of firm websites collected by ISTARL.ai. Firm websites offer a signal of realized technology use. In particular, firms rely on them to describe products, services, and production processes they actively deploy. Exploiting this insight, we assemble a unique linked panel of more than three million firm websites from Belgium, France, Germany,

---

<sup>1</sup>The MAP-AI data underlying this study – including firm-level indicators of AI adoption, AI ecosystem roles, and AI technology types for European firms between 2016 and 2024 – are publicly available via GitHub: <https://github.com/MAP-AI-data/data>.

and Luxembourg between 2016 and 2024, using historical websites from Common Crawl. This large-scale digital footprint allows us to track the diffusion of AI over time and to observe dimensions of adoption that are largely missing in conventional firm-level data.

Based on this rich, unstructured website data, we leverage a Large Language Model (LLM) within a prompt-based classification to determine which firms use AI and which ones do not. We follow the definition of AI proposed by the High-Level Expert Group on Artificial Intelligence established by the European Commission, which defines AI as “systems that display intelligent behaviour by analysing their environment and taking actions – with some degree of autonomy – to achieve specific goals.”<sup>2</sup> Beyond detecting adoption, the framework classifies AI-active firms along meaningful dimensions including their role in the AI ecosystem (e.g., core technology provider, application and solution developer, adopter or innovative user) and the type of AI technology employed (e.g., predictive and analytical applications, generative AI and conversational systems, autonomous systems and robotics).<sup>3</sup> Distinguishing firms by AI role and technology type is essential because AI adoption is not uniform. Firms differ in whether they develop AI, build AI-based applications, or adopt AI for internal use, and these activities involve different capabilities and skill requirements. Moreover, differentiating AI types captures meaningful heterogeneity in AI usage that allows us to detect qualitative shifts and the emergence of AI technologies over time.

Extensive validation against human-coded benchmarks and complementary external references including exposure-based measures such as the AI Industry Exposure index of Felten et al. (2021), and adoption-based statistics from Eurostat, confirms that the indicator captures realized AI adoption with high accuracy and minimal misclassification. This procedure results in time-varying web-based measures of AI adoption across European firms.

Using the MAP-AI indicators, we provide descriptive evidence on the landscape of firm-level AI adoption across Europe. Specifically, we examine how the share of AI-using firms evolved between 2016 and 2024, how firms’ roles within the AI ecosystem and the types of AI technologies they employ have changed over time, and how AI activity is distributed across space and sectors. Overall, we find that AI adoption increased considerably – from about 1% of firms in 2016 to nearly 12% in 2024 – with a notable acceleration after 2022, likely driven by the diffusion of generative AI and LLMs. This increasing trend coincides with a qualitative shift toward more advanced and integrated AI applications, as well as a structural transition in the European AI ecosystem driven by rapid growth in AI adoption across firms, changing the composition of AI activity toward broader AI adoption, accompanied by different national and regional specializations.

---

<sup>2</sup>See Section D.1 for the full definition of AI we use, including a discussion of alternative definitions.

<sup>3</sup>See Sections D.2 and D.3 for the full definition of AI role and type categories.

In a *second* step, we shed light on the determinants of firm-level AI adoption. For this, we further link the firm-level MAP-AI indicators of AI activity to firm-level information from the *Orbis Europe* database and to workforce skill data obtained from *Revelio Labs*, which provides detailed information on workers' skills within each firm derived from LinkedIn users. The linkage is based on firms' website URLs, enabling a direct match between firm-level information from Orbis, our AI indicators, and workforce skills. This procedure results in a linked dataset that combines our detailed, time-varying web-based measures of AI adoption with rich firm-level and workforce skill information, providing an unique basis for studying the diffusion and determinants of AI across European firms.

Based on the linked data, we regress a binary indicator of AI adoption on three groups of explanatory variables: (i) firm characteristics (firm size, sector, age), (ii) location and environment (urban status, country, regional AI startup density, local competition), and (iii) firm-level workforce skill composition, distinguishing between AI-related and seven further skills categories. The analysis reveals that firm size, sector, and location are important predictors of AI adoption, with larger and younger firms, as well as those operating in ICT and knowledge-intensive sectors and innovative urban areas, being more likely to use AI technologies. These patterns point to the importance of both firm capabilities and environmental conditions in shaping AI adoption. We also find strong evidence that workforce composition plays a critical role. In particular, firms with higher shares of employees possessing AI- and data-related skills are strongly associated with AI adoption, even after controlling for other firm characteristics and regional factors. Notably, skills are highly consistent predictors across sectors, suggesting workforce skill composition as a central driver of AI adoption.

In a *third* step, we move beyond broad skill categories and directly link firm-level AI adoption to *specific workforce skills*, providing micro-level evidence into why AI diffusion remains uneven across firms, sectors, and regions. To do so, we employ a machine learning approach based on the Least Absolute Shrinkage and Selection Operator (LASSO), which allows us to identify the most relevant skills from a high-dimensional set of 2,709 candidate skills. The results indicate that AI adoption is associated with a bundle of complementary technical skills: while advanced data and programming skills form an important foundation (e.g., *Big Data*, *SQL*, *Python*), a small set of AI-specific skills – *Machine Learning* (along with *Algorithms* and *Natural Language Processing*) – stands out in distinguishing AI adopters from non-adopters. This pattern highlights that differences in AI adoption across firms are closely linked to the composition of their technical workforce, rather than the accumulation of more broader AI-complementary skills.

We contribute to at least two strands of the literature. First, we contribute to the literature on measuring AI. Much of this literature relies on exposure measures at aggregate occupational levels, for instance, estimating the share of tasks susceptible to automation, either at occupational or job levels (Frey and Osborne, 2017, Arntz et al., 2017). More re-

cent contributions refine these approaches specifically to AI exposure (Brynjolfsson et al., 2018, Felten et al., 2018, Webb, 2020, Tolan et al., 2021, Engberg et al., 2024, Marguerit, 2024). In contrast to these aggregate AI exposure measures, there is still relatively limited evidence on realized AI adoption at the firm level. Among the few exceptions are studies that directly measure adoption within firms for the US (Hampole et al., 2025, Acemoglu et al., 2024, Zolas et al., 2020, McElheran et al., 2024) and Germany (Licht and Wohlrabe, 2024), as well as recent work on frontier technologies more broadly, which include AI as one component (Arntz et al., 2024, 2025). Compared to this literature, our contribution lies in measuring firm-level AI adoption using web data (see, e.g. Dahlke et al., 2025), which has the advantage of capturing realized AI adoption – and not potential future exposure – from a large number of firms and across a longer time period. A few recent studies also use web data to measure firm-level innovation and AI adoption (see, e.g., Kinne and Lenz, 2021; Dahlke et al., 2024). Compared to this literature, our contribution lies in measuring firm-level AI adoption using web data, which allows us to capture realized AI use – rather than potential exposure – for a large number of firms over a long time horizon. We advance recent web-based approaches by (i) using LLMs to classify firm-level web content more accurately, (ii) constructing a longitudinal panel based on historical websites (taken from Common Crawl) to capture adoption dynamics over time, and (iii) distinguishing between the role and type of AI technology they use. Together, these features provide a more detailed and scalable measure of firm-level AI adoption than existing approaches.

Second, we contribute to the literature on the determinants of technology adoption, with a particular focus on AI. Prior research on advanced digital technologies more broadly has shown that adoption depends on the availability of specific skills and competencies within firms (Tambe and Hitt, 2014, Harrigan et al., 2021), and that firm size, sector, and workforce composition shape adoption decisions (Zolas et al., 2020). In contrast, a smaller but growing strand of work looks specifically at AI adoption. Seamans and Raj (2019) emphasize the need for detailed firm-level data to study AI determinants, while Cockburn et al. (2018) highlight the importance of complementary innovation capabilities and access to data for successful AI uptake. Similarly, Bessen et al. (2022) show that access to large data sets is a critical determinant of AI adoption and growth, particularly for startups. Using firm-level evidence, Yang (2022) document that AI adoption, as measured by AI-related patents, is associated with productivity and employment effects, underscoring the economic relevance of understanding the determinants of AI adoption. More recently, McElheran et al. (2024) provide large-scale U.S. evidence showing that AI adoption is shaped by firm size, industry, location, digital maturity, and workforce skills. Our study adds to this literature by providing large-scale, cross-country evidence on the determinants of AI adoption among European firms over almost a decade. By relying on a measure that captures realized AI use – rather than AI-related patent activity proxied by patents

or self-reported adoption – we are able to study adoption decisions across a much broader set of firms and sectors. Moreover, by linking firm-level AI adoption to detailed workforce skill composition, we identify the specific skills that AI adoption depends on and show that these relationships are highly consistent across sectors. From a policy perspective, MAP-AI offers a new empirical basis for evaluating innovation policy, skill formation strategies, and regional technology diffusion in Europe.

We proceed as follows. Section 2 outlines the main data source we use for the analysis, describes the AI indicators and presents multiple validation tests, and presents the differences between AI exposure and realized AI adoption. Section 3 outlines further data sources we use for the analysis, describes the mapping of firm websites to Orbis and Revelio Labs data and discusses the representativeness of our linked samples. Section 4 reports the main results on the AI landscape in Europe. We present results on how AI adoption is evolving over time, as well as the evolution and prevalence of AI roles and types. Further, we present results on how these trends vary by sector and on how AI role specializations vary across countries. Section 5 outlines our skill classification and reports the main results on the determinants of AI adoption, such as firm characteristics, location and environment, and skills. Section 6 zooms in on the role of skills as a foundation for AI adoption. Finally, Section 7 concludes.

## 2 Measuring Firm-Level AI Adoption

### 2.1 Firm-Level AI Adoption from Firm Websites

The primary data source for our firm-level AI adoption measure (MAP-AI) is ISTARAI.ai, a startup specialized in large-scale web scraping (see, e.g., [Dahlke et al., 2025](#)). The data contains web content collected from firms across Belgium, France, Germany, and Luxembourg. ISTARAI.ai uses the Common Crawl infrastructure – an open, publicly available web archive that regularly crawls billions of web pages across the Internet – to collect firm website content over time. This results in a rich panel of web-based content for more than four million firm websites from 2016 to 2024 – with 319,996 Belgian, 2,571,616 German, 1,176,519 French, and 20,671 Luxembourgish firms. The scraping process targets up to 35 pages per website per year. This gives us an average of 6.1 pages per firm website per year, with each page containing roughly 513 words. Importantly, the firm’s homepage (i.e., the primary domain, such as “firm.com”) is available for 94.7% of firms. Due to firm entry, exit, and changes in website availability, the dataset forms an unbalanced panel.

We use four key variables from the Istarai data: the website URL (i.e., the domain of a firm website), which serves as our primary firm identifier in the Istarai data; the page URL of each scraped page of a website (i.e., the address of a page within a website); the full HTML content of each scraped page; and the corresponding year of the HTML content.

The data processing consists of several steps to ensure high quality. First, we drop observations lacking HTML content or containing duplicate HTML content (within the same firm website page URL and year). Second, we extract the text from the HTML content of each page.<sup>4</sup> Third, we exclude pages with fewer than 100 or more than 100,000 characters or line breaks. This reduces both noise and ‘token counts’ for later LLM classification. Finally, we identify the main language in the text extracted from each page URL.<sup>5</sup> This way, we can remove text that is not in any of the languages understood by the LLM that we use to identify firm-level AI adoption. Overall, only between 1.1 and 2.5% of firm websites – depending on the country – are excluded from our sample after all the restrictions listed above.

**AI Adoption.** To identify firm-level AI adoption, we use a structured prompt-based approach applied to website text extracted from the firm’s website pages. Our method leverages `Magistral Small 2506` (see [Mistral-AI et al., 2025](#), for more details), a state-of-the-art LLM to classify whether a firm uses AI in a given year, providing us a measure of whether a firm uses AI and when firms adopt AI. Moreover, we apply the same pipeline to classify firms’ roles in the AI ecosystem and the types of AI technology they use.

The classification proceeds in a multi-step approach (for detailed prompts, see Section E.1). In the first step, we assess whether a firm is using AI at all. This is done by prompting the LLM with the text of each of the firm’s website pages, along with the year of the page, the detected language, and a definition of AI.<sup>6</sup> The model is asked to output a binary indicator for AI usage, returning 1 if the firm engages in any AI-related activity and 0 otherwise, followed by a short explanation justifying the answer.<sup>7</sup> The definition of AI we give the model comes from the High-Level Expert Group on Artificial Intelligence of the European Commission (see Section D.1 for the full definition of AI).<sup>8</sup> To ensure robustness, widely used web security, content delivery, and user verification services (e.g., CAPTCHA or reCAPTCHA, Cloudflare) are explicitly excluded from being classified as AI. See Section E.3 for two examples of firm website text: one classified as using AI and the other as not using AI.

**Firm Role in the AI Ecosystem.** Conditional on a firm being classified as AI-active, we categorize its role within the AI ecosystem based on the nature of its AI-related activ-

---

<sup>4</sup>The text is extracted from the HTML code using the `html2text` Python library, resulting in a clean Markdown-formatted text.

<sup>5</sup>We use `fastText`, a Python library provided by Meta that is able to recognize a total of 176 languages (see [Joulin et al., 2016b,a](#)).

<sup>6</sup>We process the text with the LLM page by page, as putting all the available text of a website at once, and passing it together, would require a model with a larger context window.

<sup>7</sup>We ask for a short explanation, as the accuracy of the model improves, when it has to think about a justification.

<sup>8</sup><https://digital-strategy.ec.europa.eu/library/definition-artificial-intelligence-main-capabilities-and-scientific-disciplines>.

ities (see Section D.2 for full definitions).<sup>9</sup> The classification distinguishes between firms that (i) develop foundational AI technologies (*Core Technology Providers*), (ii) build AI-based applications and solutions on top of existing technologies (*Application and Solution Developers*), (iii) primarily adopt and integrate AI into their own products, services, or production processes (*Adopters or Innovative Users*), (iv) engage in AI-related research, consulting, or support activities (*Research or Support Organizations*), (v) finance, incubate, or accelerate AI-related ventures (*Investors or Incubators*), and (vi) provide data infrastructure that enables AI deployment (*Data Infrastructure Providers*). Firms whose AI-related activities do not fit these categories are assigned to a residual *Other Role*.

For each role, the LLM is prompted with tailored instructions and returns a binary indicator (1 if the role applies, 0 otherwise), together with a brief textual justification. When the available website content does not provide sufficient information to infer a firm’s role, no role category is assigned (see Section E.1 for details on the prompts).

**Firm Type of AI Technology.** In a final step, conditional on a firm being classified as AI-active, we identify the specific type of AI technology firms use by adapting and extending Eurostat’s definition of AI technologies to our setting (see Section D.3 for the full definition of the AI technology types). The classification framework distinguishes between ten AI technology types: (1) *Generative AI and Conversational Systems*, (2) *Text and Language (Non-Generative)*, (3) *Image and Video (Non-Generative)*, (4) *Speech, Audio and Music*, (5) *Predictive and Analytical Applications*, (6) *Recommendation and Personalization*, (7) *Autonomous Systems and Robotics*, (8) *Security and Fraud Detection*, (9) *Foundational AI/ML Technologies*, and (10) *Other Type*. As in the role classification, the LLM is prompted with tailored instructions for each category and returns a binary indicator (1 if the technology applies, 0 otherwise), together with a brief textual justification. Firms for which none of the predefined types apply are assigned to the residual *Other Type*. When the available website content does not provide sufficient information to infer the AI technology used, no type category is assigned (see Section E.1 for details on the prompts).

Note that a potential concern with website-based measures is that firms may overstate or strategically signal AI use for marketing purposes rather than reflecting realized adoption. Two considerations mitigate this concern. First, firm websites foremost represent an observable communication channel for realized technology use: firms typically highlight technologies that are embedded in their products, services, or production processes, particularly when addressing customers, investors, or business partners. Second, we conduct several validation checks, including comparisons with human-coded assessments and external adoption-based statistics. Any purely symbolic or inflated claims would weaken,

---

<sup>9</sup>The role taxonomy is consistent with classifications used by innovation agencies such as Luxinnovation, but is adapted here to emphasize economically meaningful differences in firms’ AI-related activities.

rather than strengthen, the empirical alignment with these benchmarks. The results suggest that our indicator captures economically meaningful AI use rather than generic marketing language. From a policy perspective, this feature is important because website-based signals reflect technologies that firms actively implement and communicate to the outside, rather than potential AI use.

## 2.2 Validation of the AI Adoption Indicator

**Human-Based Benchmark.** We begin by validating the AI adoption indicator using a human-coded benchmark. We draw a random sample of 967 firm website pages from the Orbis–Istari dataset (see Section 3.3 for a discussion of the samples’ representativeness and Section 3.1 for further details on the Orbis-Istari dataset). The sample is stratified such that half of the firms are classified by the LLM Magistral as AI adopters, and the other half as non-adopters. This ensures balanced representation despite the relatively small share of AI-active firms in the overall population. Two human coders independently reviewed each firm’s website text, the LLM’s predicted label, and the model’s reasoning. They then indicated whether they agreed with the LLM’s classification or not. This human assessment serves as our ground truth for evaluating the model’s performance. Figure A.1 shows the resulting confusion matrix comparing human evaluations with Magistral’s predictions.

The results demonstrate strong alignment between human evaluations and model predictions. Out of the 967 firms, 480 were correctly classified as not using AI, and 428 as using AI. Notably, only 59 firms were misclassified as AI-active by the LLM but classified as non-AI firms by human coders. Importantly, no firms were incorrectly identified by the LLM as using AI when human coders classified them as non-AI. This indicates that the model performs well with *perfect recall* and minimal false positives. The corresponding performance metrics – *accuracy* of 0.94, *precision* of 0.88, recall of 1.00, and *F1 score* of 0.94 – demonstrate that the model reliably identifies genuine AI activity with near-perfect accuracy.

**Eurostat Comparison (External Validity).** Next, we assess the external validity of the indicator by comparing it to official statistics on AI use from Eurostat. Specifically, we use data from the Eurostat ICT usage and e-commerce in enterprises survey, pooling the available survey waves from 2021, 2023, and 2024 (the survey was not conducted in 2022). The Eurostat data report the share of firms that self-report using at least one AI technology, disaggregated by NACE industry. In our web-based data, we compute the average share of AI-using firms by NACE industry for the same years. The Eurostat survey does not report the share of firms using AI for all countries across every industry and year; in such cases, the mean is computed over the subset of countries with non-missing

values. We then compare these industry-level averages to the corresponding adoption rates reported by Eurostat. Figure A.2 plots these two measures against each other. The strong positive correlation between them (Pearson’s  $r = 0.94$ ) indicates close alignment with survey-based measures of *realized* AI adoption. Moreover, deviations occur in both directions rather than being concentrated in above the fitted line, which is inconsistent with systematic overstatement – for marketing reasons – of AI use on firm websites.

**Additional Validation Exercises.** To further assess potential false negatives, we conduct two additional validation exercises. First, we examine firms whose domain names end with ‘.ai’, a domain extension highly likely to be associated with AI-related activity. Among 385 firms with ‘.ai’ domains, 384 are correctly identified as using AI by our indicator, corresponding to an accuracy rate of 99.7 percent. Second, we use data from the *European AI Startup Landscape* to obtain an independent list of AI startups in Germany and France (data for Belgium and Luxembourg are not available).<sup>10</sup> Matching startups via their domain URLs yields 528 matches, of which 525 are classified as AI-active by our indicator, corresponding to an accuracy of 99.4 percent.

Taken together, these validation exercises show that the AI adoption indicator performs well across multiple independent benchmarks.

## 2.3 AI Exposure vs. Realized AI Adoption.

A key advantage of our indicator is that it measures firms’ realized AI adoption as revealed on their websites. Much of the existing literature – most notably Felten et al. (2021) – focuses on occupational or industry-level measures of AI exposure, which capture the potential impact of AI on tasks rather than realized use. While informative as forward-looking indicators, such measures do not necessarily reflect whether firms have begun adopting AI in practice. Our website-based indicator addresses this limitation by directly capturing firms’ publicly signaled AI use, providing an adoption-based perspective on AI diffusion. As a result, industries with high exposure may exhibit low adoption due to regulatory, skill, or organizational constraints, while others with more moderate exposure may display comparatively high uptake.

Figure 1 illustrates this distinction. While we find a positive correlation between the AI Industry Exposure (AIIE) and our AI adoption measure on the industry level, the relationship is far from one-to-one.<sup>11</sup> Outliers highlight precisely where exposure-based expectations and adoption-based realities diverge. For instance, Accounting & Payroll

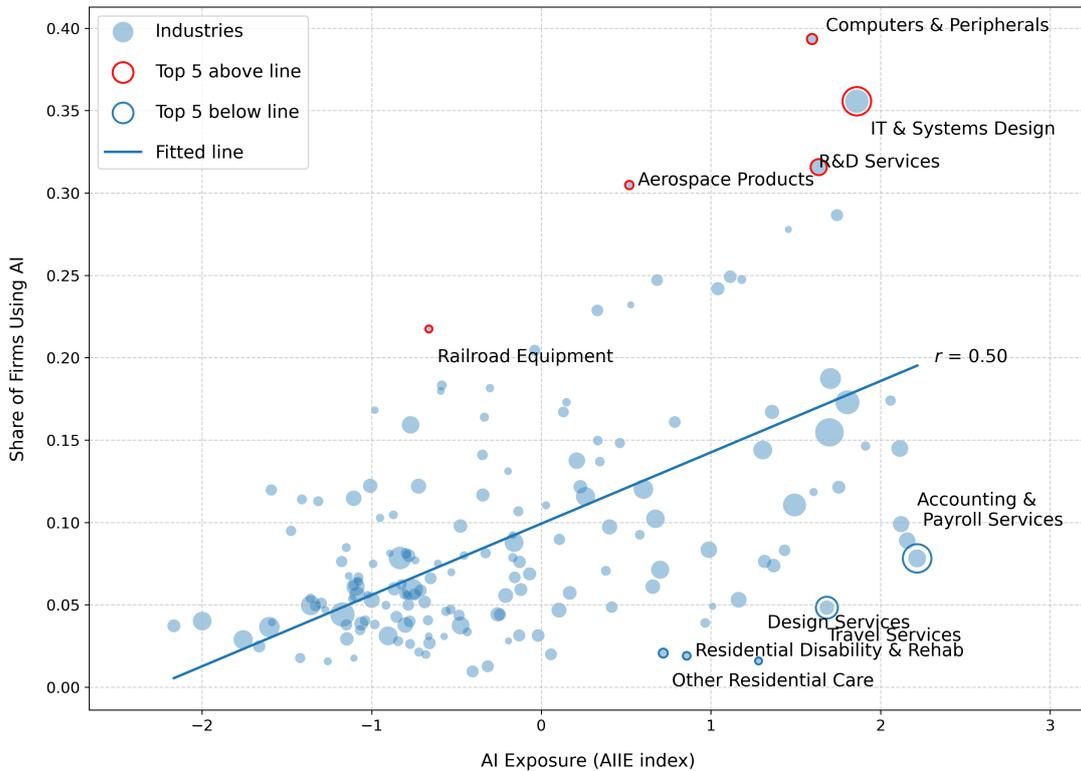
---

<sup>10</sup>The *European AI Startup Landscape*, initiated in 2020 by the applied AI Institute for Europe, Hub France IA, Ignite Sweden, AI Sweden, RISE Research Institutes of Sweden, and the Netherlands AI Coalition (NL AIC), compiles information on European AI startups (<https://www.ai-startups-europe.eu/>).

<sup>11</sup>We weight by the number of firms and drop industries with less than 100 firms.

Services and Legal Services have a very high AI exposure score, but our AI measure of AI adoption indicates that very few firms in these industries actually use or report using AI. On the other hand, according to our measure, the AI exposure score for firms in the industries Computer & Peripherals and IT & Systems Design, for example, should be considerably higher. This underscores the added value of our website-based indicator: it allows us to move beyond exposure (what AI *could* affect) to realized adoption (where AI *is* being implemented). In doing so, our AI adoption measure provides a more precise empirical basis for understanding the current diffusion of AI across firms and industries.

**Figure 1:** Industry-Level AI Exposure vs Share of Firms Using AI



*Note:* The AIIE index is taken from Felten et al. (2021), while the *share of firms using AI* is based on our MAP-AI indicator. Both measures are reported at the 4-digit NAICS industry level. Industry titles are abbreviated for readability. Bubble size is proportional to the number of firms in each industry, and industries with fewer than 100 firms are excluded.  $r$  is the Pearson correlation coefficient.

### 3 Further Data Sources

#### 3.1 Firm Characteristics from Orbis

This section describes the additional data sources used to analyze firm-level AI adoption based on our MAP-AI indicator. The Orbis Europe database, maintained by Bureau van Dijk (BvD, part of Moody's Analytics), is a cross-country firm-level dataset that collects

financial information of public and private firms in Europe. The database is continuously updated and provides standardized balance-sheet and income-statement variables – such as revenues, profits, and value added – along with key firm characteristics including industry classification, location, size, and ownership structure. The database covers the period from the early 2000s to the present, although coverage and data completeness improve substantially after 2010, reflecting the gradual inclusion of national business registers (see [Gopinath et al., 2017](#); [Kalemli-Özcan et al., 2024](#) for detailed discussions of the dataset and its coverage). Importantly, Orbis also contains firm-level website URLs, which we use to link firms to the Istari website data described in Section 2.1. We match firms using a two-step URL-based procedure that prioritizes unique domain identifiers and supplements them with additional website URLs, yielding a clean firm-level linkage (see Appendix Section C for details).<sup>12</sup>

Specifically, we merge the firm identifier (BvD ID), name, address, country, industry classification (NACE Rev. 2), legal form, firm size class (small, medium, large, very large), and year of foundation to the Istari data using the URL.<sup>13</sup> We conduct several country-specific and harmonized data-cleaning steps to ensure that each observation corresponds to a unique firm rather than a branch or an establishment (see Appendix Section C). We further aggregate industries into 5 broader sectors, following [Genz et al. \(2025\)](#).<sup>14</sup> These variables form the basis for assessing the representativeness of our matched Orbis-Istari sample and for analyzing the determinants of AI adoption.

The resulting linked Orbis–Istari dataset covers 3,030,365 unique firms between 2016 and 2024. With an average of 6.1 pages per firm website per year and approximately 513 words per page, the data provide an unprecedented large-scale view of European firms’ public digital footprints over nearly a decade.

## 3.2 Revelio Labs Data on Workforce Skills

We use workforce skill data from Revelio Labs to characterize firms’ human capital and to study the determinants of AI adoption. Revelio Labs constructs its dataset by collecting and harmonizing information from LinkedIn. The data includes detailed information on, for example, employment histories, employers, and self-reported skills for LinkedIn users worldwide, with coverage spanning the period from 2008 to 2024.<sup>15</sup>

The skills data consist of skills that users list on their profiles, which reflect both formal qualifications and on-the-job competencies. The data contains 8,704 unique skill

---

<sup>12</sup>Firms can have multiple website URLs in Orbis. For 4,308,098 firms (27.5%), at least one is available.

<sup>13</sup>The Orbis data we use was downloaded in August 2025.

<sup>14</sup>As an advantage, the classification distinguishes between knowledge-intensive and non-knowledge-intensive activities within both production and services, and includes a separate category for ICT.

<sup>15</sup>The data contains the full reported employment history of users; not only between 2008 and 2024. However, individuals who deleted their LinkedIn accounts before 2008 or created accounts after 2024 are not included in the data.

titles, which are clustered by Revelio Labs using their skill taxonomy.<sup>16</sup> After translating non-English entries, standardizing spelling, and consolidating close synonyms, we retain 8,604 distinct skills.

To link skills to firms, we use workers’ reported employment histories and assign skills to firms based on employment spells. The Revelio data is linked to the Orbis–Istari data using the company URL. This allows us to construct yearly firm-level measures of workforce composition, such as the share of employees with AI-related, data, or decision-making skills (see Section 5.1 for details of our classification). In total, we are able to match skill shares for 1,266,347 firms.

There are two aspects of the data that are worth noting: first, skills lack acquisition dates and therefore cannot be precisely timed. This essentially means that the variation of skill shares comes from the year-to-year changes in the workforce composition of a firm, and that there is no time variation in the skill set of a worker. For example, suppose a worker employed at a firm in 2016 reports Python as one of their skills. While we cannot determine whether they already acquired Python skills in 2016 – they might have learned it only later, say in 2020 – the information indicates that they were capable of acquiring Python skills at some point in their career. Second, LinkedIn restricted access to public skill information in 2021, so that all skills added to their profile before 2021 will be reported, while skills added after 2021 will not be included.

### 3.3 Representativeness of the Linked Samples

To assess how representative our firm website samples are, we compare Orbis firms successfully linked to website (Istari) and skill (Revelio) data with the broader Orbis population of firms in Belgium, France, Germany, and Luxembourg. Although Orbis does not constitute a full census, it provides a close-to-complete population reference, making it an appropriate benchmark for assessing the representativeness of the linked datasets.<sup>17</sup> Table 1 reports summary statistics and mean differences across three datasets: all firms in Orbis (column 1), all firms in Orbis linked to website data (column 2), and those further linked to employee skill data (column 3). Columns (4) and (5) display the respective differences with the reference population.

Overall, the results indicate that the linked Orbis–Istari and Orbis–Istari–Revelio samples remain broadly representative of the underlying firm population in Orbis. Firms successfully linked to website and skill information are somewhat more likely to be medium-sized or large, located in urban areas, active in knowledge-intensive or digitally visible

---

<sup>16</sup>For more details, see <https://www.data-dictionary.reveliolabs.com/methodology.html#skill-clustering>. The data also contains imputed skills, which we do not use.

<sup>17</sup>Bersch et al. (2014) compare the active stock of German firms in the Mannheim Enterprise Panel — a similar database as Orbis — with the Business Register of the Federal Statistical Office, indicating that Orbis gives a close to representative picture of the corporate landscape in Germany.

industries such as ICT and professional services, and more likely to be German rather than French. Despite these minor differences, both samples capture the overall structure of the firm population well, providing a representative and reliable foundation for the subsequent analyses of firm-level AI adoption.

Representativeness also holds at a fine level of industry disaggregation. Appendix Figure A.3 shows a strong correlation between the share of firms by 4-digit NACE industry in the Orbis–Istari sample and the full Orbis population. Taken together, these results indicate that the linked datasets provide a reliable and broadly representative basis for the analysis of firm-level AI adoption – a clear advantage to surveys, which are in many cases not representative for all 1-digit NACE industries.

**Table 1:** Balance Test: Representativeness of Firm Website Data

	Reference	Sample		Difference	
	Orbis (1)	Orbis-Istari (2)	Orbis-Istari- Revelio (3)	(2)-(1) (4)	(3)-(1) (5)
Year of creation	2007.1 (14.0)	2006.7 (14.2)	2007.2 (14.5)	-0.4	0.1
<b>Firm size (%)</b>					
Small	85.6	80.1	75.3	-5.6	-10.4
Medium-sized	11.3	15.7	17.9	4.3	6.6
Large	2.4	3.4	5.3	1.0	2.9
Very large	0.6	0.9	1.5	0.3	0.9
<b>Sector (%)</b>					
ICT	3.6	4.0	5.0	0.5	1.4
KI prod.	1.3	1.7	2.0	0.4	0.8
NKI prod.	19.5	18.2	15.1	-1.2	-4.4
KI service	22.3	22.7	25.8	0.4	3.5
NKI service	53.3	53.4	52.1	0.0	-1.2
<b>Legal form (stand.) (%)</b>					
Branches	3.2	2.5	2.7	-0.7	-0.6
Foreign companies	1.0	1.1	1.6	0.1	0.6
Other legal forms	5.6	2.7	3.1	-2.8	-2.5
Partnerships	18.4	15.6	15.9	-2.8	-2.5
Private limited companies	69.4	75.5	73.5	6.1	4.1
Public limited companies	2.4	2.5	3.2	0.1	0.8
<b>Location (%)</b>					
Rural	54.6	54.2	50.4	-0.4	-4.1
Urban	45.4	45.8	49.6	0.4	4.1
<b>Country (%)</b>					
Belgium	11.6	10.9	11.4	-0.7	-0.2
France	58.8	51.6	54.6	-7.2	-4.2
Germany	28.4	36.7	33.1	8.3	4.7
Luxembourg	1.2	0.8	0.9	-0.4	-0.3
Number of obs.	34,547,428	21,895,147	10,009,925		
Number of firms	15,682,646	3,030,365	1,266,347		

*Notes:* The table compares all firms in the Orbis database with those successfully matched to website and skill data across key observable characteristics. Differences are reported relative to the full Orbis population. The reported number of observations corresponds to the maximum sample size (i.e., variables with no missing values). Missing values are typically below 2–4% for industry, location, and year of incorporation, and therefore, do not affect the comparisons.

## 4 The AI Landscape in Europe

This section provides a descriptive overview of AI adoption across firms in Belgium, France, Germany, and Luxembourg, based on our MAP-AI indicator, as described in Section 2. In particular, we analyze how AI activity varies across time, sectors, and space, distinguishing between the firm’s role in the AI ecosystem, and by type of AI technology used. This way, we are able to describe the diffusion of AI using large-scale, unstructured textual data and to document emerging patterns in firms’ innovation activity in Europe.

### 4.1 Trends in AI Activity

**Overall.** Figure 2 shows the evolution of the share of firms using AI between 2016–2024, according to our website-based indicator. The indicator is constructed cumulatively, so that once a firm reports AI usage, it remains classified as AI-active in all subsequent years.<sup>18</sup> The figure reveals a clear upward trend of AI adoption: while only a small share of firms (1%) were involved in AI activity in 2016, the share has increased steadily over time, with slightly accelerated growth after 2022, presumably driven by the wider availability of LLMs. In 2024, 12% of firms already adopted AI, according to their website.

Country patterns reveal some heterogeneity in the speed of adoption. Luxembourg and Germany exhibit the highest shares of AI-related firms by 2024, suggesting faster and broader diffusion relative to their neighbors. France follows a similar trajectory but at a slightly lower level, while Belgium consistently lags behind, with about 8% of firms classified as AI-active in 2024. Despite these differences in magnitude, all countries display a common trend of steadily rising adoption throughout the period.

Overall, the figures highlight the gradual diffusion process: adoption begins with a small group of early innovators, then expands as more firms adopt AI across their business models, products, or services. By 2024, AI technologies are no longer limited to a small group of early adopters, but have spread across the economy.

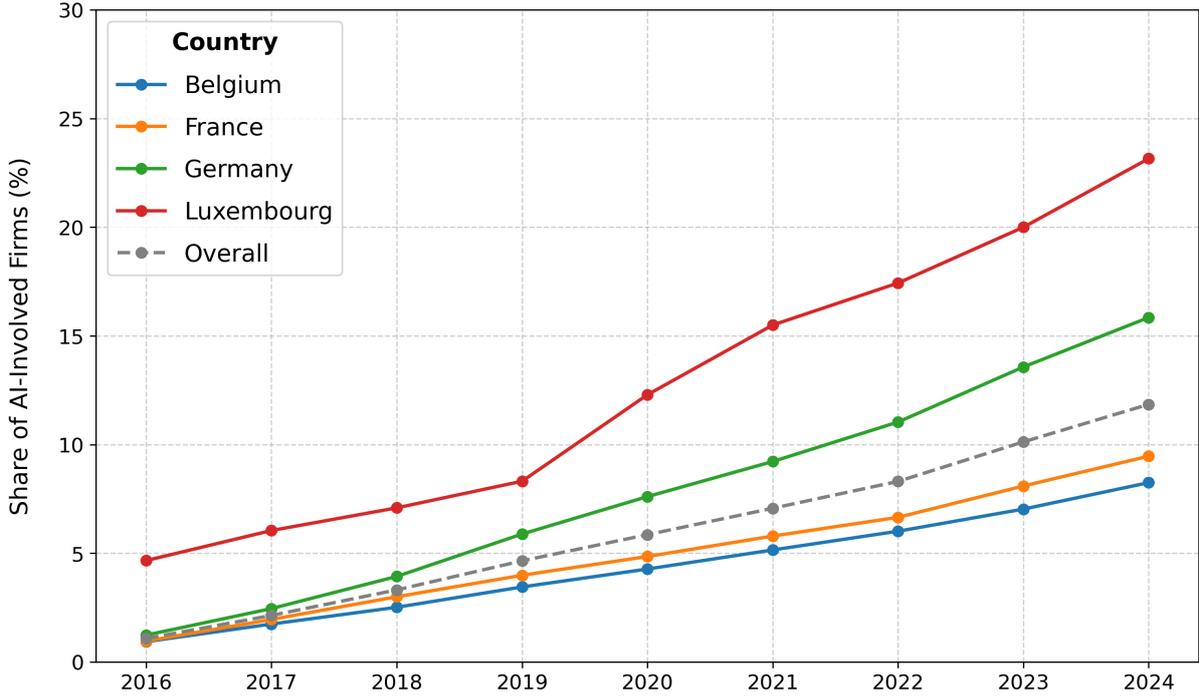
**Firms’ Role in the AI Ecosystem.** Figure 3 compares absolute and relative developments in firms’ roles within the AI ecosystem, revealing important structural shifts (see Section D.2 for the full definition of the firm role categories).<sup>19</sup> In absolute terms, the largest group is *Adopter or Innovative User*, representing about 46.6% of AI-active firms in 2024. They are followed by the *Application and Solution Developer* category (41.2%) and by the residual *Other Role* (23.1%). *Research or Support Organization* accounts

---

<sup>18</sup>We constructed the indicator in this way for two main reasons: first, it is unlikely that a firm that adopted AI goes back on its decision; and second, since we do not observe the entire website of firms, but rather a selection of pages, we want to reduce the year-to-year noise in AI adoption.

<sup>19</sup>Note that the role categories are not mutually exclusive (i.e., firms can have multiple roles), so that the shares do not add up to 1.

**Figure 2:** Share of AI-Involved Firms by Year and Country



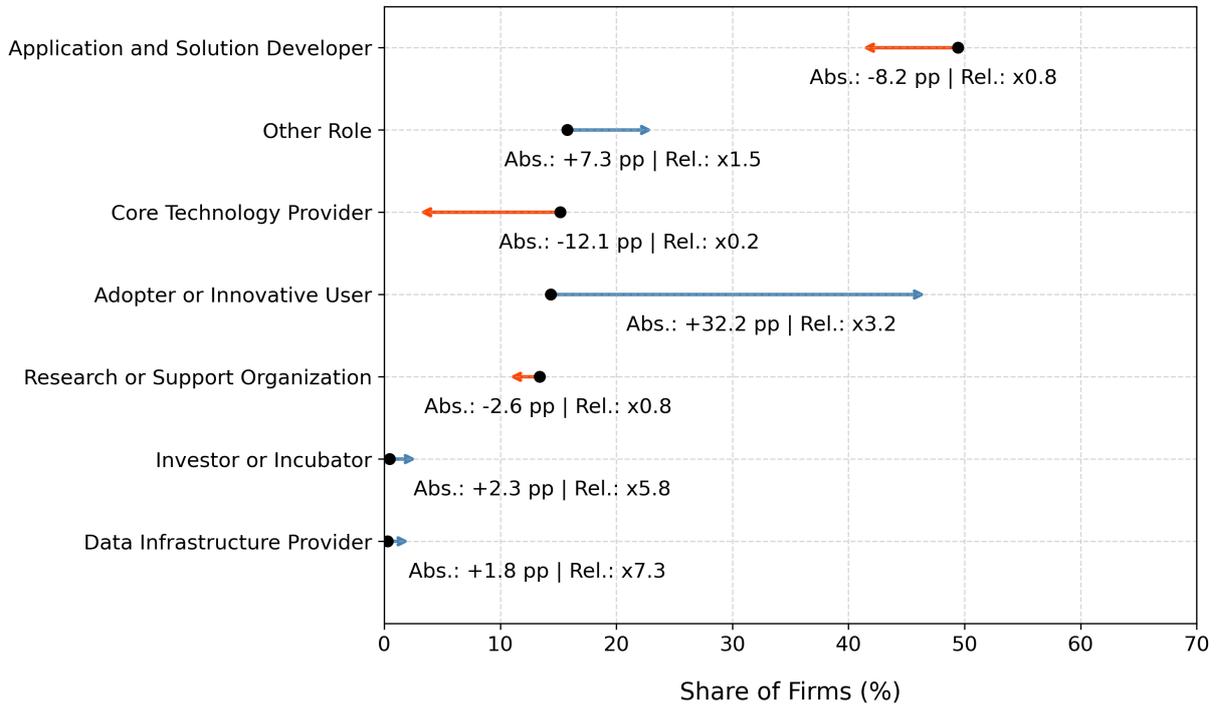
*Note:* The figure reports the cumulative share of firms mentioning AI on their websites between 2016 and 2024 based on the MAP-AI indicator. Once a firm is classified as using AI, it remains classified as using AI in all subsequent years, provided the website remains active.

for 10.8%, *Core Technology Provider* for 3.1%, *Investor or Incubator* for 2.7% and the smallest group is *Data Infrastructure Provider* at 2.1%.

The strongest growth is observed among *Adopter and Innovative User* (+32.2 pp), followed by the *Other Role* category (+7.3 pp). The increase in *Other Role* might indicate that there are emerging roles in the AI ecosystem that are not assigned an independent role, indicating a diversification of the AI ecosystem. In fact, BERTtopic model suggests that the “Other Role” category mainly captures emerging activities such as AI content publishing, infrastructure provision, education, and AI-enabled services (see Appendix Figure A.4). Moreover, *Investor or Incubator* and *Data Infrastructure Provider*, though still small in absolute size, increased almost sixfold (+2.3 pp) and more than sevenfold (+1.8 pp), respectively. By contrast, *Core Technology Provider*, *Application and Solution Developer*, and *Research or Support Organization* contract sharply (−12.1, 8.2, and 2.6 pp).

Taken together, these results suggest a shift away from a supply-driven AI ecosystem dominated by specialized technology providers toward a demand-driven landscape in which adoption and complementary roles play a central part. This decline reflects a relative change in shares driven by the rapid growth of AI adopters, rather than an absolute reduction in development activity; firms may simultaneously take multiple roles.

**Figure 3:** Share of Firms Using AI by Role in the AI Ecosystem (2016-2024)



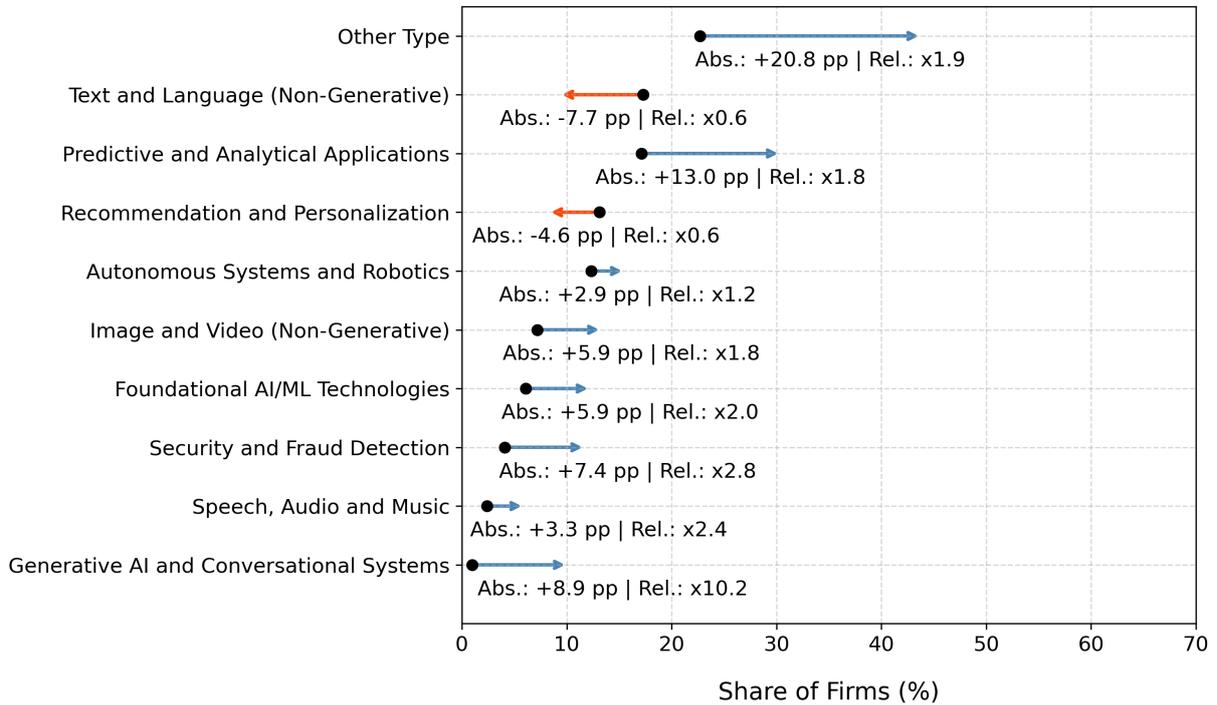
*Note:* Shares denote the proportion of AI-active firms by their role in the AI ecosystem based on the MAP-AI indicator. Firms may assume more than one AI role in the AI ecosystem, so the classified role categories are not mutually exclusive (i.e., shares do not add up to 1). This figure shows the absolute and relative changes between the levels of 2016 (black dots) and 2024 (tip of the arrows). About 12.9% and 14.9% of firm websites do not have enough information to infer the role in 2016 and 2024, respectively. See Section D.2 for definitions of categories.

The development indicates a maturing ecosystem in which foundational AI development is increasingly accompanied by broader diffusion and integration across firms and sectors.

**Type of AI Technology.** Comparing absolute levels with relative growth reveals a clear shift across AI technology types, as shown in Figure 4. In 2024, the most prevalent technologies are *Other Type* (43.6%), *Predictive and Analytical Applications* (30.2%), and *Autonomous Systems and Robotics* (15.3%).

The most dynamic type of AI technology category is *Other Type*, which expanded by 20.8 percentage points between 2016 and 2024, most probably reflecting the rapid emergence of diverse, application-specific AI use cases that do not align with standard technology taxonomies. As shown by the BERTopic results (Appendix Figure A.5), this category is largely composed of domain-oriented applications such as industrial optimization, health, and creative systems, indicating an increasing fragmentation and contextualization of AI technologies as adoption spreads across sectors. This expansion is followed by strong growth in *Predictive and Analytical Applications* and *Generative AI and Conversational Systems*, which expanded by 13 and 8.9 percentage points, respectively. The

**Figure 4:** Share of AI-Involved Firms by Type of AI Technology (2016-2024)



*Note:* Shares denote the proportion of AI-active firms whose websites mention an AI technology type based on the MAP-AI indicator. Firms may assume more than one AI technology type, so the classified AI type of technology categories are not mutually exclusive (i.e., shares do not add up to 1). The figure compares absolute levels (black dots) with changes over time (arrows indicating percentage-point and relative growth) between 2016-2024. About 12% of firm websites do not have enough information to infer the type in 2016 and 2024. See Section D.3 for definitions of categories.

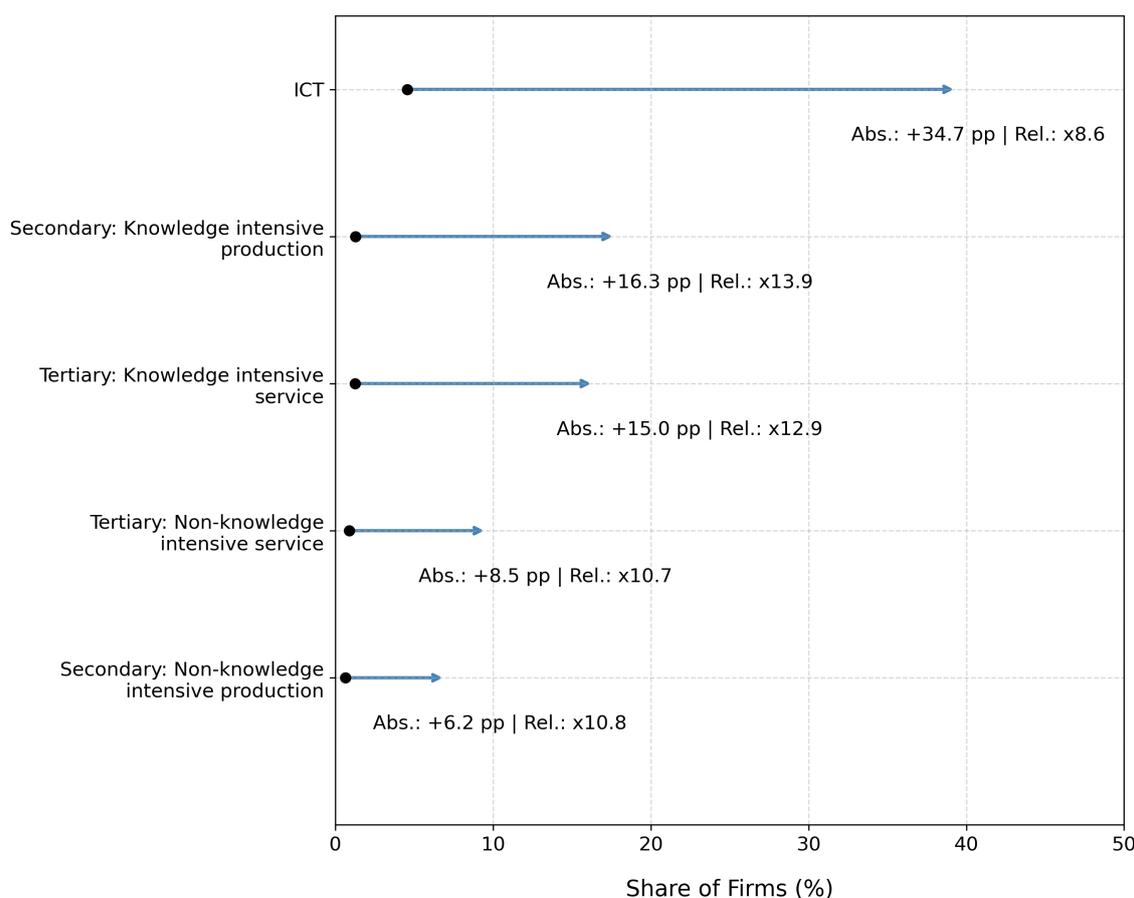
increase of the latest category most probably reflects the rise of AI copilots, autonomous agents, and LLM-powered chats in more recent years. Notably, it seems to partially replace more classic text and language analysis technologies, as the category *Text and Language (Non-Generative)* contracts (-7.7 pp).

Overall, these patterns point to a qualitative shift in AI use toward more advanced and integrated applications, driven by recent breakthroughs in foundation models. Rather than focusing on isolated analytical tasks, firms increasingly deploy AI for prediction, interaction, and decision support, enabling more autonomous and adaptive systems. This evolution reflects a transition from narrow, task-specific AI toward technologies that are more deeply embedded in core business processes and organizational workflows. In this sense, while early diffusion patterns in Europe resemble the highly concentrated, early-stage adoption documented for the United States by [McElheran et al. \(2024\)](#), the post-2022 acceleration suggests that European firms have entered a more advanced diffusion phase characterized by broader firm-level integration.

## 4.2 Trends in AI Adoption by Sector

Figure 5 shows that AI adoption has developed unevenly across sectors over time. At the aggregate level, the ICT sector is by far the most AI-intensive, with almost 48% of firms using AI in 2024. Knowledge-intensive industries – both in services and in production – also exhibit strong growth, with increases of around 15–16 percentage points. By contrast, non-knowledge-intensive sectors show more modest absolute gains (6–8 pp), though their relative increases remain substantial.

**Figure 5:** Share of AI-Involved Firms by Sector (2016-2024)



*Note:* The figure reports the share of AI-involved firms based on the MAP-AI indicator by sector classification used by [Genz et al. \(2025\)](#). Only industries with at least 200 firms are included.

Looking at the detailed NACE one-digit breakdown (see Figure A.6), the pattern becomes even clearer. Information and communication leads by a wide margin (+31.0 pp), followed by financial services (+17.7 pp), professional, scientific and technical activities and utilities (electricity, gas, steam) (+14.6 pp), and administrative and support service activities (+12.2 pp). Education also displays a substantial growth (+10.9 pp), while public administration and defence record the strongest relative increase (over thirtyfold) despite a low starting point.

Overall, the patterns suggest that AI adoption is accelerating across all sectors, with the industries in the ICT and knowledge-intensive sectors driving Europe’s AI diffusion. However, AI use is no longer confined to ICT and is spreading into traditional and low-tech industries, suggesting technology spillovers from ICT to other sectors through applications and digital transformation.

### 4.3 Spatial Distribution of AI Activity

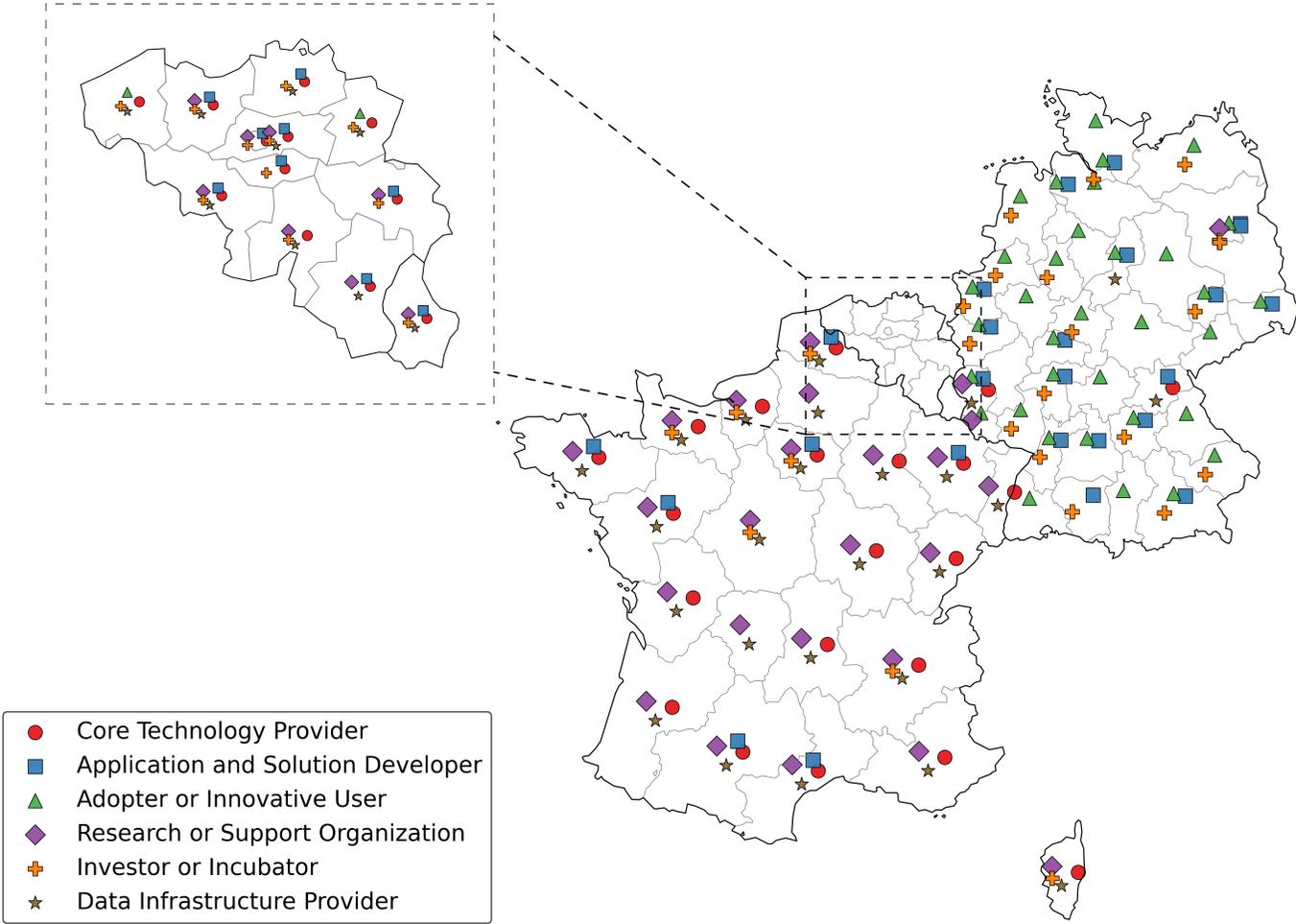
While the previous sections documented how AI adoption differs across roles and technology types, these patterns may also vary systematically across space. We therefore now turn to the spatial distribution of AI activity by role. Figure 6 maps AI activity by role in 2024 – represented by symbols – for NUTS-2 regions across our four countries. To highlight relative specializations, a symbol is only displayed for a region if the share of firms with a given role is above the median across all regions. The resulting map therefore identifies areas where specific roles are disproportionately represented within the AI ecosystem across the four countries.

The map reveals major differences across the four countries. In France, the AI landscape is dominated by core technology providers (red circles) and research or support organizations (purple diamonds). These roles are widely distributed across the country, pointing to a strong emphasis on fundamental AI development. By contrast, the relative presence of AI adopters and application developers is limited, suggesting that France’s strengths lie more in AI innovation and knowledge generation than in large-scale applied AI deployment.

Germany shows the opposite pattern. Here, AI application and solution developers (blue squares) and AI adopters or innovative users (green triangles) are widely distributed and appear in a majority of regions. This reflects the country’s strong industrial base and its orientation toward implementing AI into products, services, and production processes. While AI core technology providers and research or support organizations are present, their relative weight is lower than in France. Other roles, such as investors or incubators (orange crosses) and firms categorized under other role (brown stars), appear in both countries but are less systematically concentrated.

Luxembourg and Belgium also show a notable presence of AI core technology providers, but these firms are more closely clustered with research, application, and investment roles. This reflects a less specialized yet intensive AI ecosystem, where fundamental and applied AI development evolve side by side.

Figure 6: Spatial Distribution of AI-Involved Firms by Role



Note: Symbols are displayed if the regional share (NUTS-2 level) of firms using AI – based on the MAP-AI indicator – within a given role is above the median of the distribution across the four countries. See Section D.3 for definitions of categories.

Overall, the spatial distribution of AI activity reveals complementary national specializations. France emphasizes foundational AI research, Germany focuses on applications, and Belgium and Luxembourg host a less specialized but more diverse AI ecosystem that links research and deployment. These descriptive patterns motivate the regression analysis in Section 5, where we use the MAP-AI adoption indicator to study the firm-level determinants of AI use.

## 5 Determinants of AI Adoption

This section examines the determinants of AI adoption among firms in our four European countries, measured using the MAP-AI indicator introduced in Section 2, and exploits whether there is heterogeneity across firms’ roles in the AI ecosystem and sectors. As potential determinants, we exploit basic firm characteristics from the Orbis data (Section 3.1) as well as firms’ share of skills by category, detailed in the next subsection.

### 5.1 Classification of skills

We construct firms’ share of skills by category based on the Revelio Labs data discussed in Section 3.2. In particular, we classify the unique 8,604 skills into eight broad categories using a LLM. We use the LLM called `Qwen3-235B-A22B-Instruct-2507` (see Yang et al., 2025, for more details) for skill categorization due to its stronger reasoning capabilities, despite more resource demanding, relative to Magistral.<sup>20</sup> We categorize skills into seven types – AI, Data, Prediction, Judgment, Decision-Making, Social, Leadership –, and an “Other” group for residual skills, following the skill taxonomy of Feimi et al. (2025). Table B.1 shows the definitions together with examples. This taxonomy bridges economic theory (Agrawal et al., 2022) with empirical evidence by distinguishing between AI-related technical skills and their human complements. AI, Data, and Prediction skills capture the technological core of AI implementation, while Judgment, Decision-Making, Social, and Leadership reflect the complementary cognitive and interpersonal abilities that remain crucial as AI reduces prediction costs (i.e., assuming that AI in its current stage mainly implies better and cheaper prediction).

We proceed with the categorization of skills in two steps: first, for each skill, we elicit a structured description (label, short definition, core subskills, common tools, typical roles, related terms); and second, using the skill name and its description, we assign exactly one category  $\kappa(s)$  per skill (no multi-labeling). For detailed prompts, see Section E.2.

---

<sup>20</sup>We do not use Qwen for the website-based AI indicator classification because classifying millions of long web pages with such a large model would take several months on a high-performance computing system). Additionally, as shown in the validation of Magistral in Section 2.1, it performs exceptionally well in detecting AI activity on firm websites.

**Firm–year skill shares.** To construct the firm-year level skill shares for our eight broad skill categories, we proceed as follows. Let  $f$  denote a firm in the Revelio data,  $y \in \{2016, \dots, 2024\}$  a calendar year, and  $c \in \{\text{AI, Data, Prediction, Judgment, Decision-Making, Social, Leadership, Other}\}$  a skill category. We define the set of unique workers attached to  $f$  in year  $y$  as:

$$U_{f,y} := \{u : \text{start}(u, f) \leq y \leq \text{end}(u, f)\}$$

where ‘start’ indicates the year the worker started working at the firm, and ‘end’ indicates the year the worker stopped working for the firm, or 2024 if the worker was still working at the firm in 2024.<sup>21</sup> For each worker  $u \in U_{f,y}$ , let  $S_{u,f,y}$  be the set of that worker’s unique skills associated with positions at firm  $f$  during year  $y$  after mapping each skill to a category  $\kappa(s)$ . We compute total and category-specific counts:

$$N_{f,y} = \sum_{u \in U_{f,y}} |S_{u,f,y}|, \quad N_{f,y,c} = \sum_{u \in U_{f,y}} \sum_{s \in S_{u,f,y}} \mathbf{1}\{\kappa(s) = c\},$$

and define the firm–year skill share:

$$\text{share}_{f,y,c} = \begin{cases} \frac{N_{f,y,c}}{N_{f,y}}, & N_{f,y} > 0, \\ 0, & N_{f,y} = 0. \end{cases}$$

where worker-skill pairs contribute only once to any firm–year observation (i.e., we do not double-count internal position changes).

For example, consider a firm in 2020 with three observed employees: Employee 1 lists Machine Learning, OpenCV, Excel, and Customer Service (mapped to AI, AI, Data, and Social, respectively); Employee 2 lists Leadership, Communication, and Project Management (mapped to Leadership, Social, and Leadership, respectively); Employee 3 lists Stata, R and Data Visualization (mapped to Data, Data, and Data, respectively). The firm has  $N_{f,2020} = 10$  total skill mentions. The AI category share is  $\text{share}_{f,2020,\text{AI}} = 2/10 = 0.20$ , while Data is  $4/10 = 0.40$ .

Our firm-specific skill share measure counts total skill mentions rather than unique employees with at least one skill in a category. This approach has two key advantages given our data structure. First, because Revelio captures only LinkedIn users – a subset of each firm’s workforce – the number of observed employees varies substantially across firms and over time. Our skill shares aggregate more observations (individual skills) than employee counts, providing more stable measures when workforce coverage is limited. Second, and more fundamentally, our skill shares capture expertise rather than diffusion

---

<sup>21</sup>We exclude positions with a missing start date and set missing end dates to 2024.

across employees: a firm with one employee listing ten AI-related skills signals greater AI capability than a firm with ten employees each listing a single AI skill.

## 5.2 Empirical Specification

The literature on corporate technology investments identifies several key determinants of AI adoption (see Introduction). Building on these insights, we examine how firm-level characteristics, location and environment, and workforce skills shape the likelihood of adopting AI technologies. Leveraging our novel, web-based measure of firm-level AI adoption, the empirical analysis is descriptive and aims to document systematic associations between firm characteristics, workforce skills, and AI use, rather than to identify causal effects. Specifically, we regress a binary indicator of AI adoption on three groups of explanatory variables: (i) *firm characteristics*, including firm size (four categories), sector (five categories), and firm age; (ii) *location and environment*, accounting for contextual influences such as proximity to innovation clusters and regional competition, proxied by urban or rural status, the number of AI startups in the region – defined as the number of AI adopters founded within the last 3 years – and local competition measured by the reversed Herfindahl index, where lower values mean higher concentration and therefore low competition. Country dummies are included to capture cross-national time-invariant differences in regulatory and institutional environments; and (iii) *workforce skills*, encompassing the capabilities of the workforce (eight skill categories).

Our baseline specification models the probability of AI adoption as a function of these determinants. Formally, we estimate:

$$\Pr(\text{AI}_{it} = 1) = F(\alpha + \beta' \mathbf{FirmChar}_{it} + \phi' \mathbf{LocEnv}_{it} + \gamma' \mathbf{WorkfSkills}_{it} + \delta_t), \quad (1)$$

where  $F(\cdot)$  denotes a logit model,  $\text{AI}_{it}$  denotes the MAP-AI adoption indicator, equal to one if firm  $i$  uses AI in year  $t$ . The term  $\mathbf{FirmChar}_{it}$  represents firm characteristics, including firm size, sector, and age.  $\mathbf{LocEnv}_{it}$  represents location and environmental factors, including urban or rural status, regional AI startup density, and local competition, while  $\mathbf{WorkfSkills}_{it}$  measures skills and organizational capabilities, proxied by the firm’s share of workforce skills. Year fixed effects  $\delta_t$  absorb aggregate shocks, and standard errors are clustered at the firm level. Additionally, all continuous variables – firm age, regional AI startup density, competition, and workforce skill shares – are standardized to facilitate direct comparison of effect magnitudes.

In Appendix D, we further explore the heterogeneity in the determinants of AI adoption by estimating equation (1) separately for firms of different sectors. Among others, the extension sheds light on which determinants are consistent across sectors.

### 5.3 Results

**AI Adoption.** Table 2 presents the results of the baseline logit regressions examining the firm-level determinants of AI adoption. Column (1) uses the full Orbis–Istari sample, column (2) restricts the analysis to firms also linked to Revelio data, and column (3) extends the specification with firms’ workforce skill shares. All results should be interpreted as associations with web-revealed AI adoption as captured by MAP-AI, rather than potential AI exposure.

Firm size is among the strongest predictors of AI adoption. Compared with small firms, the likelihood of using AI increases sharply with firm size: medium-, large-, and very large-sized firms display progressively higher coefficients. This pattern aligns with expectations from the AI and broader technology diffusion literature (see, e.g., [McElheran et al., 2024](#), [Genz et al., 2025](#)), which holds that larger firms benefit from greater capability to adopt and integrate frontier technologies. The inclusion of workforce skills in column (3) slightly reduces the magnitude of these effects relative to column (2), suggesting that part of the size premium operates through firms’ internal skill composition rather than scale alone. Moreover, firm age is consistently and negatively associated with AI adoption. This suggests that organizational inertia continues to constrain digital transformation. Turning to sectoral differences, firms in ICT and knowledge-intensive sectors – both in production and services – display higher propensities to adopt AI than firms in the non-knowledge-intensive production sector (i.e., the reference category).

Regarding location and environment, firms in urban areas are more likely to adopt AI than rural firms. The association with regional AI startup density is strong and robust across specifications. Competitive intensity, proxied by the reversed Herfindahl index, also has a slight positive association, though of smaller magnitude and negative when accounting for skills. Country effects display clearer differentiation once workforce skills are included. In column (3), firms in Germany show the highest relative adoption rates, followed by France, while Belgium shows less AI adoption in comparison to Luxembourg. These results suggest that cross-country differences are largely explained by sectoral specialization and workforce composition, particularly the higher technical-skill intensity in Germany and France.

In column (3), we augment the baseline specification with firm-level workforce skill composition, capturing firms’ organizational and human-capital capacities to implement AI. Data skills exhibit the largest association with AI adoption, followed by AI-specific skills, underscoring the central role of data-related capabilities and technical expertise in enabling AI use. Importantly, however, adoption is not driven by technical skills alone. Non-technical skill categories – such as Social, Leadership, Judgment, and Decision-making skills – also display positive and statistically significant associations relative to the reference category.

**Table 2: Determinants of AI Adoption**

	(1) Orbis-Istari (1)	(2) + Revelio sample (2)	(3) + Skills added (3)
<b>(i) Firm characteristics</b>			
<i>Firm size (ref: small)</i>			
Medium-sized	0.21*** (0.01)	-0.05*** (0.01)	-0.08*** (0.01)
Large	0.85*** (0.01)	0.41*** (0.01)	0.24*** (0.01)
Very large	1.37*** (0.02)	0.89*** (0.02)	0.64*** (0.02)
Firm age (std)	-0.19*** (0.00)	-0.15*** (0.00)	-0.09*** (0.00)
<i>Sector (ref: NKI production)</i>			
ICT	1.95*** (0.01)	1.69*** (0.01)	1.19*** (0.01)
Knowledge-int. production	0.83*** (0.02)	0.64*** (0.02)	0.49*** (0.02)
Knowledge-int. service	0.80*** (0.01)	0.59*** (0.01)	0.49*** (0.01)
Non-knowledge-int. service	0.33*** (0.01)	0.25*** (0.01)	0.22*** (0.01)
<b>(ii) Location and environment</b>			
Urban (=1)	0.19*** (0.01)	0.14*** (0.01)	0.12*** (0.01)
AI startups in region (std)	0.24*** (0.00)	0.24*** (0.00)	0.23*** (0.00)
Competition (HHI, reversed, std)	0.03*** (0.00)	0.00 (0.00)	-0.01* (0.00)
<i>Country FE (ref: Luxembourg)</i>			
Belgium	-0.26*** (0.02)	-0.22*** (0.03)	-0.20*** (0.03)
France	-0.06* (0.02)	-0.02 (0.03)	0.08** (0.03)
Germany	0.08*** (0.02)	0.22*** (0.03)	0.45*** (0.03)
<b>(iii) Skills and organizational capabilities</b>			
<i>Firm-level workforce shares of the following skill types (ref: other, std):</i>			
AI			0.38*** (0.01)
Prediction			0.15*** (0.00)
Judgment			0.12*** (0.00)
Data			0.49*** (0.00)
Decision-making			0.10*** (0.00)
Social			0.16*** (0.00)
Leadership			0.12*** (0.00)
<i>Fixed-effects</i>			
Country	Yes	Yes	Yes
Year	Yes	Yes	Yes
Observations	20,326,284	9,349,205	9,051,145
Pseudo R <sup>2</sup>	0.10739	0.10696	0.19015
# of unique firms	2,810,762	1,183,224	1,183,219

*Notes:* Logit models. The dependent variable is the MAP-AI adoption indicator, equal to one if a firm is classified as AI-active in a given year based on website content, and zero otherwise. Column (1) uses the Orbis-Istari sample. Column (2) restricts to firms also observed in Revelio. Column (3) adds workforce skill shares. All continuous regressors are standardized. Standard errors clustered at the firm level. Significance: \* p < 0.05, \*\* p < 0.01, \*\*\* p < 0.001.

Overall, the results indicate that AI adoption largely reflects underlying firm capabilities and the innovation environments in which firms operate. Adoption is significantly more likely among larger and younger firms operating in ICT and other knowledge-intensive sectors, particularly in dynamic regions with high AI startup activity. In this respect, our findings closely mirror those documented by (McElheran et al., 2024), who show that early AI adoption is highly skewed toward large, young, skill-intensive firms and concentrated in a limited set of technologically dynamic regions, pointing to a persistent and potentially widening AI divide. In addition, our results further show that workforce skills are the most consistent determinant across sectors: Appendix Table B.2 indicates that AI-, data-, and decision-related skill shares are strongly and positively associated with AI involvement in all sectoral groupings, with stable magnitudes. Interestingly, our results show that they matter the most in the non-knowledge-intensive service sector, where they might be scarce. Taken together, our results suggest that workforce skill composition is a key driver of AI adoption.

## 6 Skill Foundations of AI Adoption

Having established that firms’ workforce skill composition is a key determinant of AI adoption, we now examine which specific skills relate most to AI adoption at the firm level, as measured by the MAP-AI indicator. In particular, we follow a machine learning approach to identify which specific skills are most predictive of firms’ likelihood of adopting AI technologies. This data-driven framework allows us to select the most relevant skills from a high-dimensional set of potential predictors. The analysis offers initial insights into which skills or worker types may be required to adapt to an AI-driven business, and which reorganizational efforts firms may need to undertake to navigate the digital transformation.

### 6.1 Empirical Strategy

For the following analysis, we draw on the Orbis-Istari-Revelio sample as before in Column (3) of Table 2, and construct a firm–skill matrix that captures the share of unique skills, rather than the share of our eight aggregated skill categories used in the determinant analysis in Section 5.

To identify which unique skills are most predictive of firm-level AI adoption, we estimate a high-dimensional LASSO logistic regression model in which the dependent variable indicates whether firm  $i$  is classified as adopting AI according to the MAP-AI indicator, and the explanatory variables capture the skill composition of its workforce – both at any point in time between 2016 and 2024. We conduct the estimation based on the cross-section of the data, since we are not aware of corresponding methods feasible for

our particular panel structure. To reduce computational burden and focus on skills with sufficient variation, we restrict attention to skills present in at least 1% of firms, yielding 2,709 (out of 8,604) candidate skills. The estimation framework can be written as

$$\Pr(\text{AI}_i = 1) = \text{logit}^{-1} \left( \beta_0 + \sum_{s=1}^{2,709} \beta_s \text{Skill}_{i,s} + \boldsymbol{\gamma}' \mathbf{X}_i \right), \quad (2)$$

where  $\text{Skill}_{i,s}$  represents the number of skills of type  $s$  in firm  $i$  as a share of all skills in firm  $i$ .

The vector  $\mathbf{X}_i$  includes basic firm-level controls for country, industry (NACE 1-digit), firm size, regional location (NUTS-1), and year of firm foundation.

We employ LASSO logistic regression, which applies an  $\ell_1$  penalty to the skill coefficients to perform automatic variable selection. The objective function minimizes:

$$-\ell(\beta_0, \boldsymbol{\beta}, \boldsymbol{\gamma}) + \lambda \sum_{s=1}^{2,709} |\beta_s|, \quad (3)$$

where  $\ell(\cdot)$  denotes the log-likelihood and  $\lambda$  controls the regularization strength. The  $\ell_1$  penalty forces coefficients of less predictive skills to exactly zero, thereby isolating the most distinctive capabilities associated with AI adoption.

This approach is particularly well-suited to our setting for four reasons. First, it explicitly manages the multicollinearity inherent in our skill dataset, where related skills frequently co-occur within firms. Second, it remains interpretable by keeping only the most important skills and eliminating the rest. Third, most firms only have a small fraction of all possible skills, which means that our data mainly consists of zeros. LASSO works efficiently with this type of data and naturally produces models that also focus on just a few key skills rather than all 2,709. Fourth, it balances predictive accuracy with model parsimony (i.e., keeping the model simple and using as few variables as possible), avoiding overfitting in high-dimensional settings where the number of skills substantially exceeds the typical size of feature spaces.

We calibrate the model to balance predictive accuracy with parsimony, selecting a regularization strength that yields an interpretable set of skills while maintaining strong predictive performance. To do this, we systematically test 40 different parameter settings and use a 5-fold stratified cross-validation.<sup>22</sup> We evaluate performance using the area under the ROC curve (ROC-AUC), a standard metric that measures how well the model distinguishes AI adopters from non-adopters, which is particularly appropriate given that only 20% of firms in our sample adopt AI. The optimal model identifies 613 skills out of

---

<sup>22</sup>This means that we divide the data into five parts, train the model on four parts, and test it on the remaining part, rotating through all combinations.

2,709 candidates (22.6%) as most predictive of AI adoption, achieving a cross-validated ROC-AUC of 0.791 ( $\pm 0.001$ ), indicating strong predictive performance.<sup>23</sup>

While the coefficients from the LASSO logit model do not admit a causal interpretation, their magnitudes are informative about the relative importance of different skills for predicting AI adoption. Because all skill variables are measured on the same scale and estimated jointly, larger coefficients indicate skills that are more strongly associated with AI adoption, conditional on the full set of skills and firm characteristics. Substantially larger coefficients observed for a skill therefore signal that these capabilities are markedly more informative in distinguishing AI adopters from non-adopters than other skill categories. For the presentation of results, individual skill coefficients are mapped to the eight broader skill categories and summarized using category-level counts and mean standardized coefficients. Overall, our approach provides a transparent overview of which types of skills are most prevalent and most predictive of AI adoption.

## 6.2 Results

Before turning to the ranking of the top ten individual skills for each skill category, Figure 7 summarizes the distribution and predictive strength of LASSO-selected skills across our eight broader skill categories. Panel (a) reports the number of skills selected by the LASSO within each category, while Panel (b) displays the corresponding mean coefficient of skills within each category. Coefficient magnitudes should be interpreted as relative measures of predictive importance rather than causal effects, given the penalized nature of the estimator.

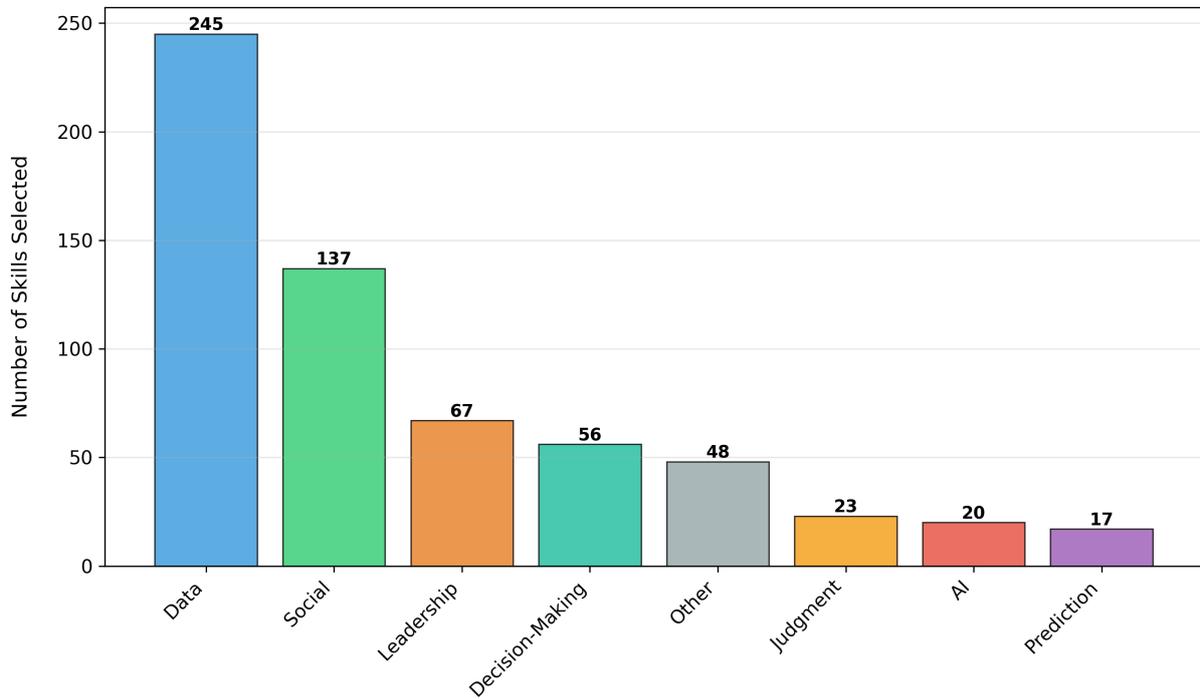
In total, 604 out of 2,709 skills were selected, suggesting that only a limited set of skills meaningfully distinguishes AI-adopting firms from non-adopters. The results reveal a pronounced contrast between prevalence and predictiveness. *Data*-related skills dominate in volume, with 245 skills selected—around 40% of all selected skills – followed by *Social* and *Leadership* skills. However, their mean coefficient is moderate. By contrast, directly *AI*-related skills as well as *Prediction* skills are comparatively rare, accounting for only 3.3% and 2.6% of selected skills. Despite their limited presence, directly *AI*-related skills stand out in Panel (b) with by far the largest average coefficient, exceeding those of all other categories by an order of magnitude.

This pattern illustrates that AI adoption is not primarily characterized by the broad presence of generally applicable capabilities, but rather by the presence of highly specific technological skills. While Data, Social, and Leadership skills are widespread across firms, their relatively modest average coefficients indicate that they are shared by both AI adopters and non-adopters and therefore only weakly distinguish between the two groups

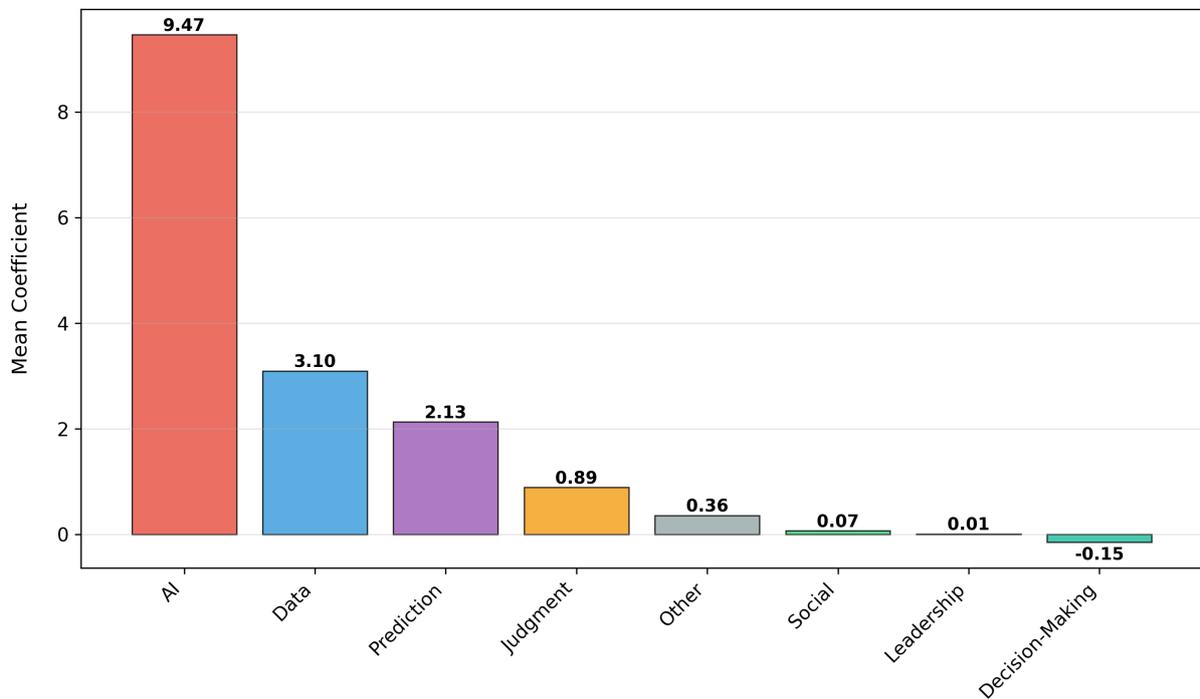
---

<sup>23</sup>An ROC-AUC of 0.791 means that if you randomly pick one AI-adopting firm and one non-adopting firm, the model using workforce skill information will correctly rank the adopter as more likely to adopt AI about 79% of the time.

**Figure 7:** Selected Skills by Category and Mean Coefficient Magnitude



**(a)** Number of Skills Selected per Category



**(b)** Mean Coefficient Magnitude by Category

*Note:* Panel (a) shows the count of skills with non-zero LASSO coefficients in each category. Panel (b) shows the average coefficient magnitude for selected skills within each category.

on average. These skills appear to reflect baseline competencies that are increasingly common across firms, rather than decisive factors that, on their own, drive AI adoption.

In contrast, directly AI-related skills – and, to a lesser extent, Prediction skills and more advanced Data skills – are rare but highly informative. Their larger average coefficients imply that, conditional on all other skills and firm characteristics, the presence of these capabilities are strong markers of AI adoption. At the same time, the widespread presence of Data skills alongside their positive, though smaller, coefficients suggests that Data and AI skills act as complements rather than substitutes: general data capabilities provide a necessary foundation, while AI-specific skills play a decisive role in enabling adoption.

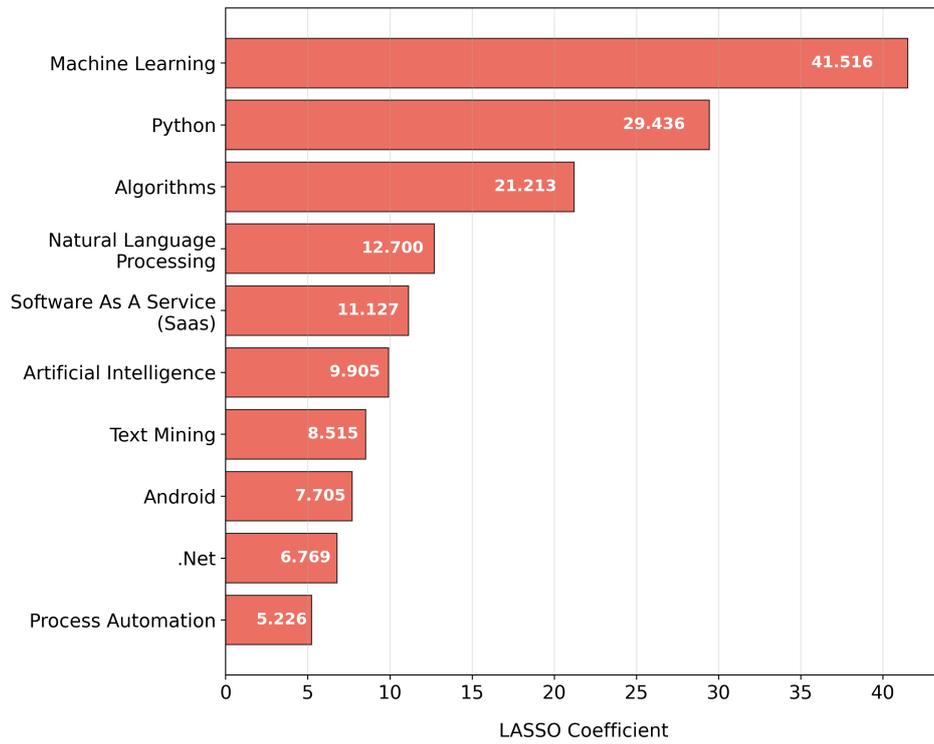
In the following, we focus on AI-related and Data skills and present the top ten individual skills within these two categories most strongly associated with AI adoption.

**Top 10 AI & Data Skills.** Figure 8 presents the top 10 LASSO-selected skills from both the *AI* and *Data* categories, ranked by coefficient magnitude. We focus on AI-related and Data skills, as these categories exhibit the strongest and most informative associations with AI adoption and capture the core technical capabilities underlying firms’ adoption decisions. The composition reveals important insights about the skill foundation of AI adoption.

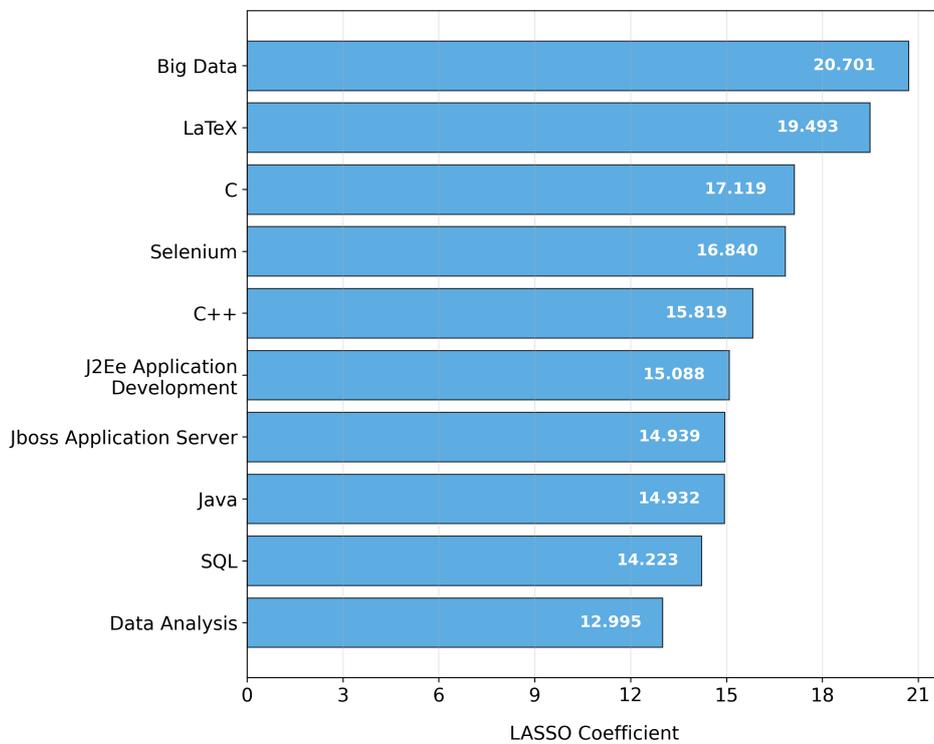
Regarding the top ten *AI* skills shown in Panel (a), *Machine Learning* (ML) clearly dominates, exhibiting by far the largest LASSO coefficient, followed by *Python* and *Algorithms*. This ordering is particularly instructive. ML – the most predictive *AI* skill – has a coefficient roughly twice as large as that of the highest-ranked *Data* skill (*Big Data*), underscoring the core role of ML expertise in distinguishing AI-adopting firms from non-adopters. The prominent position of *Python* skills further highlights its role as the primary programming language for implementing ML methods and integrating AI models into business applications. Together with the strong ranking of *Algorithms*, these results point to the importance of practical AI capabilities rather than general knowledge about AI concepts.

Turning to the top *Data* skills in Panel (b) reveals a complementary pattern. The leading Data skills – *Big Data*, *SQL*, and *Data Analysis* – reflect firms’ ability to process and extract value from large-scale datasets. At the same time, the presence of programming languages (*C*, *C++*, *Java*) and enterprise technologies (*J2EE Application Development*, *JBoss Application Server*) indicates that data capabilities in AI-adopting firms are related to data infrastructures and business-ready systems. This suggests that AI adoption relies not purely on descriptive analytics, but on data engineering and data mining capabilities that enable AI models to be trained and deployed at scale.

**Figure 8: Top 10 AI and Data Skills**



**(a) AI Skills**



**(b) Data Skills**

*Note:* The figure reports the top 10 AI (panel (a)) and Data (panel (b)) skills selected by LASSO, ranked by the mean coefficient.

Taken together, our results in this section indicate that AI adoption as measured by MAP-AI is not primarily associated with general data availability or organizational capacity alone, but rather with firms’ complementary technical skills. While advanced data and programming capabilities constitute an important foundation, AI-specific skills – most notably *Machine Learning* – stand out in differentiating AI adopters from non-adopters. Overall, the findings suggest that AI-adopting firms are characterized by the internal combination of specific technical skills, rather than by broad skill accumulation or general organizational skills.

## 7 Conclusion

This paper introduces MAP-AI (Mapping Artificial Intelligence in Firms), a novel firm-level indicator of realized AI adoption in Europe derived from large-scale analysis of firm websites leveraging recent advances in LLMs. By applying a state-of-the-art LLM to the text of more than four million firm websites, combined with firm characteristics and workforce skill information, we provide a cross-country, time-variant view of the evolving AI landscape in Belgium, France, Germany, and Luxembourg from 2016 to 2024. Our approach classifies not only whether firms use AI, but also their specific role in the AI ecosystem and the type of AI technology they employ, achieving 94% accuracy. This provides a large-scale and time-varying measure of realized AI adoption that complements traditional industry exposure-based approaches by capturing realized firm-level AI adoption instead of potential AI exposure.

Using MAP-AI, we document a rapidly expanding AI landscape. AI adoption grew from just 1% of firms in 2016 to nearly 12% in 2024, with acceleration after 2022 – coinciding with breakthroughs in generative AI and LLMs. Beyond this trend in AI adoption, we document a structural transformation: the European AI ecosystem is shifting from a supply-driven structure dominated by AI core technology providers toward a demand-driven landscape in which AI adopters and innovative users play increasingly central roles. The development indicates a maturing ecosystem in which foundational AI development increasingly coincides with broader diffusion and integration across firms and sectors.

AI adoption patterns are strongly shaped by firm characteristics, and location and environment factors. Larger, younger, and firms in ICT and knowledge-intensive sectors are more likely to integrate AI technologies. Luxembourg and Germany lead in overall adoption rates, while France shows particular strength in AI research and core technology development, and Germany dominates in application development and deployment. Regional innovation ecosystems matter significantly: firms in urban areas and with high AI startup density show higher adoption rates than comparable firms elsewhere, even after controlling for other factors.

Another very decisive factor distinguishing adopters from non-adopters is workforce skills. Our analysis suggests that Data skills form the necessary foundation, with firms possessing big data, and programming capabilities being highly associated with AI adoption. Specialized AI skills – while showing the strongest positive association when present – seem to provide value primarily as complements to these foundational Data capabilities rather than as substitutes. Data skills are most frequently selected by the LASSO model, but AI skills show the highest positive association, suggesting that firms have a higher likelihood of adopting AI when both Data and AI skills are present. Suggesting that a comprehensive technical capacity throughout the firm, combining foundational data competencies (e.g., Big Data and SQL) with specialized AI expertise (e.g., Machine Learning, Algorithms) and complementary capabilities (e.g., Judgment), is key for AI adoption. These patterns hold consistently across roles in the AI ecosystem and sectors, though they are particularly pronounced in less knowledge-intensive industries where technical skills provide the greatest comparative advantage.

Our paper demonstrates the strength of using LLMs to process web text data for research. Future research could build on this framework to study the economic consequences of AI adoption on productivity, innovation, employment composition, and wage structures. Extending the geographic coverage beyond the four countries to additional European countries or other regions would enable comparative analyses of how different institutional and policy environments shape adoption patterns. Moreover, as AI technologies continue to evolve rapidly, ongoing monitoring using our web-based measurement approach could provide near-real-time insights into emerging trends, new roles in the AI ecosystem, and the diffusion of new AI technologies.

## References

- Acemoglu, D., Anderson, G. W., Beede, D. N., Buffington, C., Childress, E. E., Dinlersoz, E., Foster, L. S., Goldschlag, N., Haltiwanger, J. C., Kroff, Z., Restrepo, P., and Zolas, N. (2024). Automation and the Workforce: A Firm-Level View from the 2019 Annual Business Survey. In *Technology, Productivity, and Economic Growth*. University of Chicago Press.
- Acemoglu, D. and Johnson, S. (2023). *Power and progress: Our thousand-year struggle over technology and prosperity*. Hachette UK.
- Agrawal, A., Gans, J., and Goldfarb, A. (2022). *Power and Prediction: The Disruptive Economics of Artificial Intelligence*. Harvard Business Review Press.
- Arntz, M., Böhm, M., Graetz, G., Gregory, T., Lehmer, F., and Lipowski, C. (2025). Firm-level technology adoption in times of crisis. IZA Discussion Paper No. 17846.
- Arntz, M., Genz, S., Gregory, T., Lehmer, F., and Zierahn-Weilage, U. (2024). Frontier technology adopters and the aggregate decline of routine jobs. IZA Discussion Paper No. 16740.
- Arntz, M., Gregory, T., and Zierahn, U. (2017). Revisiting the Risk of Automation. *Economics Letters*, 159:157–160.
- Barr, A. and Feigenbaum, E. A. (1981). *The Handbook of Artificial Intelligence, Volume 1*. Addison-Wesley, Reading, MA.
- Bersch, J., Gottschalk, S., Müller, B., and Niefert, M. (2014). The Mannheim Enterprise Panel (MUP) and firm statistics for Germany. *ZEW Discussion Paper*, No 14-104. Unpublished.
- Bessen, J., Impink, S. M., Reichensperger, L., and Seamans, R. (2022). The role of data for AI startup growth. *Research Policy*, 51(5):104513. Publisher: Elsevier.
- Brynjolfsson, E., Chandar, B., and Chen, R. (2025). Canaries in the Coal Mine? Six Facts about the Recent Employment Effects of Artificial Intelligence.
- Brynjolfsson, E., Mitchell, T., and Rock, D. (2018). What can machines learn, and what does it mean for occupations and the economy? In *AEA Papers and Proceedings*, volume 108, pages 43–47.
- Chollet, F. (2019). On the Measure of Intelligence. *arXiv preprint arXiv:1911.01547*.

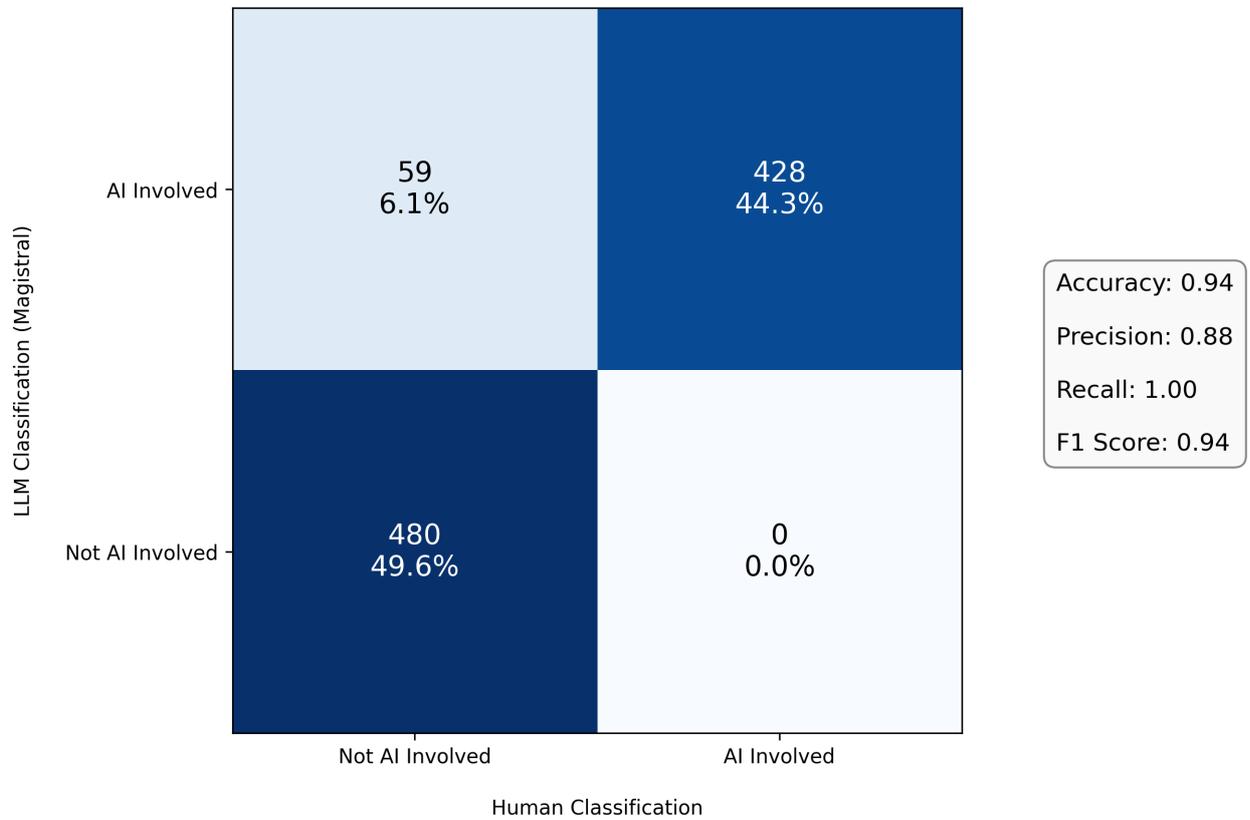
- Cockburn, I. M., Henderson, R., and Stern, S. (2018). The impact of artificial intelligence on innovation: An exploratory analysis. In *The economics of artificial intelligence: An agenda*, pages 115–146. University of Chicago Press.
- Dahlke, J., Beck, M., Kinne, J., Lenz, D., Dehghan, R., Wörter, M., and Ebersberger, B. (2024). Epidemic effects in the diffusion of emerging digital technologies: evidence from artificial intelligence adoption. *Research Policy*, 53(2):104917.
- Dahlke, J., Schmidt, S., Lenz, D., Kinne, J., Dehghan, R., Abbasiharofteh, M., Schütz, M., Kriesch, L., Hottenrott, H., Kanilmaz, U. N., et al. (2025). The webai paradigm of innovation research: Extracting insight from organizational web data through ai. ZEW Discussion Paper No. 25-019.
- Engberg, E., Hellsten, M., Javed, F., Lodefalk, M., Sabolová, R., Schroeder, S., and Tang, A. (2024). Artificial Intelligence, Hiring and Employment: Job Postings Evidence from Sweden. Technical report.
- Feimi, N., Gathmann, C., Gregory, T., and Marguerit, D. (2025). Skill Demand in the Age of AI: Evidence from Europe. *mimeo*.
- Felten, E., Raj, M., and Seamans, R. (2021). Occupational, industry, and geographic exposure to artificial intelligence: A novel dataset and its potential uses. *Strategic Management Journal*, 42(12):2195–2217. Publisher: Wiley Online Library.
- Felten, E. W., Raj, M., and Seamans, R. (2018). A method to link advances in artificial intelligence to occupational abilities. In *AEA Papers and Proceedings*, volume 108, pages 54–57.
- Frey, C. B. and Osborne, M. A. (2017). The future of employment: how susceptible are jobs to computerisation? *Technological Forecasting and Social Change*, 114:254–280. Publisher: Elsevier.
- Genz, S., Gregory, T., and Lehmer, F. (2025). The impact of firm-level technology adoption on workers. *mimeo*.
- Gopinath, G., Kalemli-Özcan, , Karabarbounis, L., and Villegas-Sanchez, C. (2017). Capital Allocation and Productivity in South Europe. *The Quarterly Journal of Economics*, 132(4):1915–1967.
- Hampole, M., Papanikolaou, D., Schmidt, L. D., and Seegmiller, B. (2025). Artificial intelligence and the labor market. Technical report, National Bureau of Economic Research.

- Harrigan, J., Reshef, A., and Toubal, F. (2021). The March of the Techies: Job Polarization Within and Between Firms. *Research Policy*, 50(7):104008. Publisher: Elsevier BV.
- High-Level Expert Group on Artificial Intelligence (2019). Ethics Guidelines for Trustworthy AI. Technical report, European Commission.
- Hofreiter, S. (2024). *Was kann künstliche Intelligenz? Wird sie Wohlstand für alle schaffen, Krebs heilen und das Klimaproblem lösen?* ecoWing Verlag, Linz, Austria.
- Joulin, A., Grave, E., Bojanowski, P., Douze, M., Jégou, H., and Mikolov, T. (2016a). Fasttext.zip: Compressing text classification models. *arXiv preprint arXiv:1612.03651*.
- Joulin, A., Grave, E., Bojanowski, P., and Mikolov, T. (2016b). Bag of tricks for efficient text classification. *arXiv preprint arXiv:1607.01759*.
- Kalemlı-Özcan, , Sørensen, B. E., Villegas-Sanchez, C., Volosovych, V., and Yeşiltaş, S. (2024). How to Construct Nationally Representative Firm-Level Data from the Orbis Global Database: New Facts on SMEs and Aggregate Implications for Industry Concentration. *American Economic Journal: Macroeconomics*, 16(2):353–374.
- Kinne, J. and Lenz, D. (2021). Predicting innovative firms using web mining and deep learning. *PLOS ONE*, 16(4):1–18.
- Licht, T. and Wohlrabe, K. (2024). Ai adoption among german firms. Technical report, CESifo Working Paper.
- Marguerit, D. (2024). Augmenting or Automating Labor? The Effect of AI Exposure on New Work, Employment, and Wages. *mimeo*.
- McElheran, K., Li, J. F., Brynjolfsson, E., Kroff, Z., Dinlersoz, E., Foster, L. S., and Zolas, N. (2024). AI Adoption in America: Who, What, and Where. *Journal of Economics & Management Strategy*, pages 1–24.
- Mistral-AI, :, Rastogi, A., Jiang, A. Q., Lo, A., Berrada, G., Lample, G., Rute, J., Barmentlo, J., Yadav, K., Khandelwal, K., Chandu, K. R., Blier, L., Saulnier, L., Dinot, M., Darrin, M., Gupta, N., Soletskyi, R., Vaze, S., Scao, T. L., Wang, Y., Yang, A., Liu, A. H., Sablayrolles, A., Héliou, A., Martin, A., Ehrenberg, A., Agarwal, A., Roux, A., Darcet, A., Mensch, A., Bout, B., Rozière, B., Monicault, B. D., Bamford, C., Wallenwein, C., Renaudin, C., Lanfranchi, C., Dabert, D., Mizelle, D., de las Casas, D., Chane-Sane, E., Fugier, E., Hanna, E. B., Delerce, G., Guinet, G., Novikov, G., Martin, G., Jaju, H., Ludziejewski, J., Chabran, J.-H., Delignon, J.-M., Studnia, J., Amar, J., Roberts, J. S., Denize, J., Saxena, K., Jain, K., Zhao, L., Martin, L., Gao,

- L., Lavaud, L. R., Pellat, M., Guillaumin, M., Felardos, M., Augustin, M., Seznec, M., Raghuraman, N., Duchenne, O., Wang, P., von Platen, P., Saffer, P., Jacob, P., Wambergue, P., Kurylowicz, P., Muddireddy, P. R., Chagniot, P., Stock, P., Agrawal, P., Sauvestre, R., Delacourt, R., Gandhi, S., Subramanian, S., Dalal, S., Gandhi, S., Ghosh, S., Mishra, S., Aithal, S., Antoniak, S., Schueller, T., Lavril, T., Robert, T., Wang, T., Lacroix, T., Nemychnikova, V., Paltz, V., Richard, V., Li, W.-D., Marshall, W., Zhang, X., and Tang, Y. (2025). *Magistral*.
- Seamans, R. and Raj, M. (2019). Artificial Intelligence, Labor, Productivity, and the Need for Firm-Level Data. In *The Economics of Artificial Intelligence: An Agenda*, pages 553–565. University of Chicago Press.
- Tambe, P. and Hitt, L. M. (2014). Job hopping, information technology spillovers, and productivity growth. *Management science*, 60(2):338–355. Publisher: INFORMS.
- Tolan, S., Pesole, A., Martínez-Plumed, F., Fernández-Macías, E., Hernández-Orallo, J., and Gómez, E. (2021). Measuring the occupational impact of AI: tasks, cognitive abilities and AI benchmarks. *Journal of Artificial Intelligence Research*, 71:191–236.
- Webb, M. (2020). The Impact of Artificial Intelligence on the Labor Market. Working Paper, Accessed at: [https://www.michaelwebb.co/webb\\_ai.pdf](https://www.michaelwebb.co/webb_ai.pdf).
- Yang, A., Li, A., Yang, B., Zhang, B., Hui, B., Zheng, B., Yu, B., Gao, C., Huang, C., Lv, C., Zheng, C., Liu, D., Zhou, F., Huang, F., Hu, F., Ge, H., Wei, H., Lin, H., Tang, J., Yang, J., Tu, J., Zhang, J., Yang, J., Yang, J., Zhou, J., Zhou, J., Lin, J., Dang, K., Bao, K., Yang, K., Yu, L., Deng, L., Li, M., Xue, M., Li, M., Zhang, P., Wang, P., Zhu, Q., Men, R., Gao, R., Liu, S., Luo, S., Li, T., Tang, T., Yin, W., Ren, X., Wang, X., Zhang, X., Ren, X., Fan, Y., Su, Y., Zhang, Y., Zhang, Y., Wan, Y., Liu, Y., Wang, Z., Cui, Z., Zhang, Z., Zhou, Z., and Qiu, Z. (2025). Qwen3 technical report.
- Yang, C.-H. (2022). How artificial intelligence technology affects productivity and employment: Firm-level evidence from Taiwan. *Research Policy*, 51(6):104536. Publisher: Elsevier.
- Zolas, N., Kroff, Z., Brynjolfsson, E., McElheran, K., Beede, D. N., Buffington, C., Goldschlag, N., Foster, L., and Dinlersoz, E. (2020). Advanced Technologies Adoption and Use by US Firms: Evidence From the Annual Business Survey. *National Bureau of Economic Research Working Paper*, No. 28290:1–93.

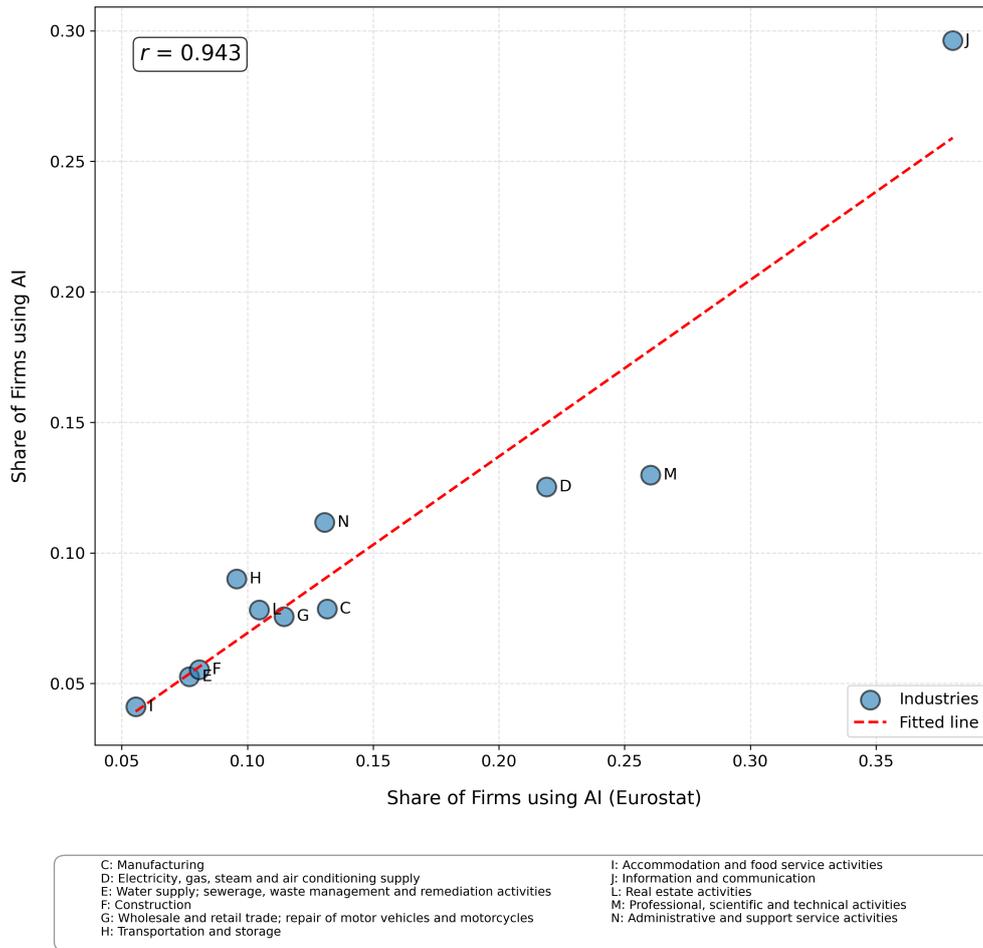
# A Figures

**Figure A.1:** Confusion Matrix



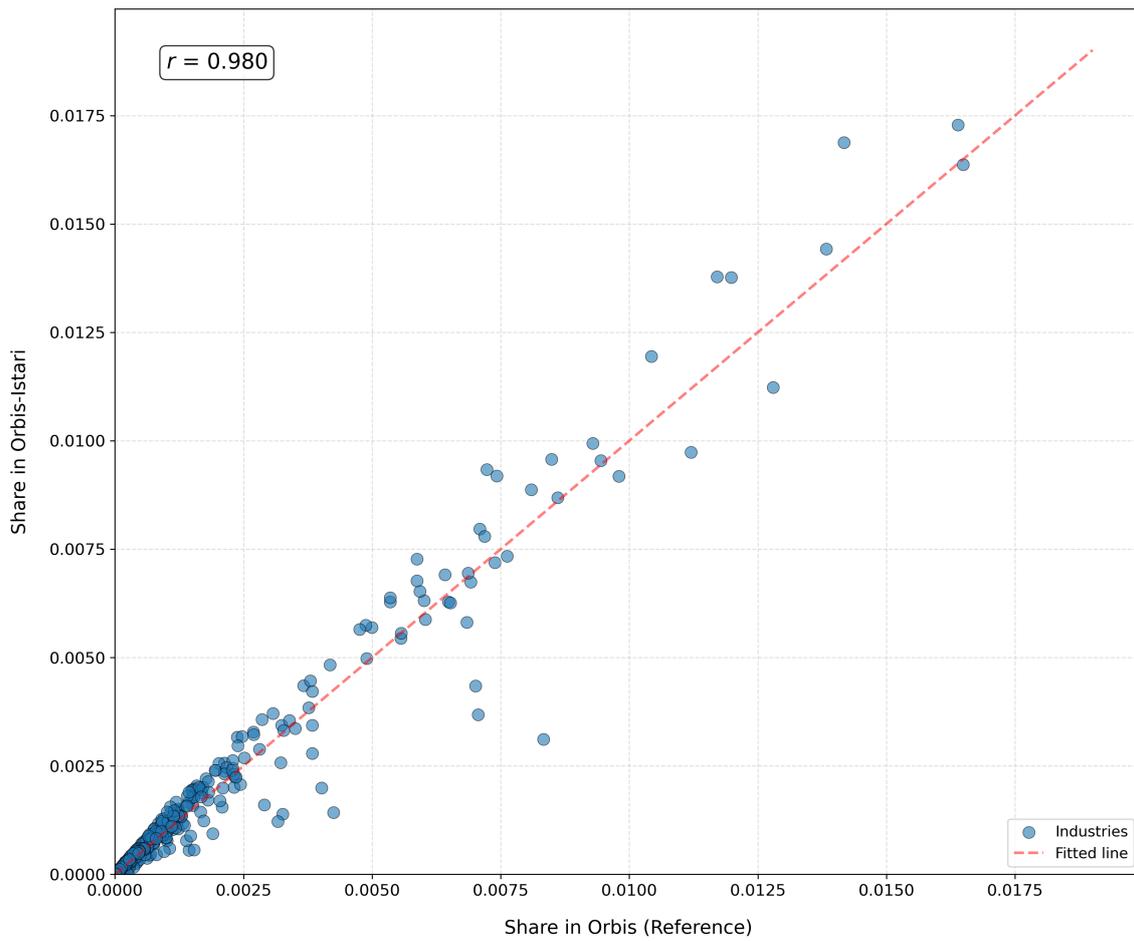
*Note:* The figure displays the confusion matrix comparing human-validated classifications (columns) and model-predicted classifications (rows). Each cell shows the number of firms and the corresponding percentage of the total sample. The accompanying box on the right reports standard performance metrics: accuracy, precision, recall, and F1 score.

**Figure A.2:** Industry-Level AI Adoption Shares: Web-Based Indicator vs. Eurostat



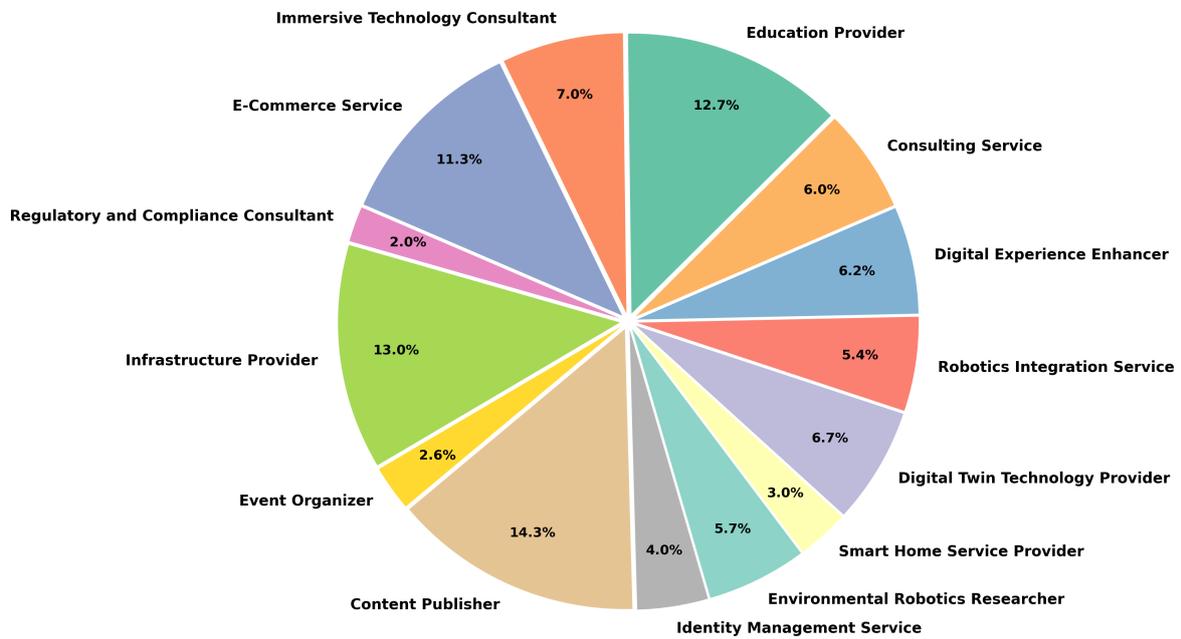
*Note:* The figure plots the average share of firms using AI by NACE industry based on our web-based AI adoption indicator for the years 2021, 2023, and 2024 (vertical axis) against the corresponding industry-level AI adoption rates reported by Eurostat (horizontal axis). Eurostat adoption rates are computed by pooling the ICT usage and e-commerce in enterprises survey waves from 2021, 2023, and 2024 (the survey was not conducted in 2022). Each point represents a NACE industry (letters correspond to industry codes listed below the figure). The dashed line shows the fitted linear relationship. The correlation between the two measures is high (Pearson's  $r = 0.943$ ), indicating close alignment with official survey-based statistics, while deviations from the fitted line reflect heterogeneity in realized AI adoption across industries. The Eurostat survey does not report the share of firms using AI for all countries across every industry and year; in such cases, the mean is computed over the subset of countries with non-missing values.

**Figure A.3:** Representativeness of Orbis-Istari at the 4-Digit Industry Level



*Note:* The figure plots, for each 4-digit NACE industry, the share of firms in the full Orbis population (horizontal axis) against the corresponding share in the Orbis-Istari sample (vertical axis). Each point represents a 4-digit NACE industry. The dashed line shows the fitted linear relationship. The strong correlation (Pearson's  $r = 0.943$ ) between the two measures indicates that the linked sample closely mirrors the full Orbis population even at a fine level of industry disaggregation.

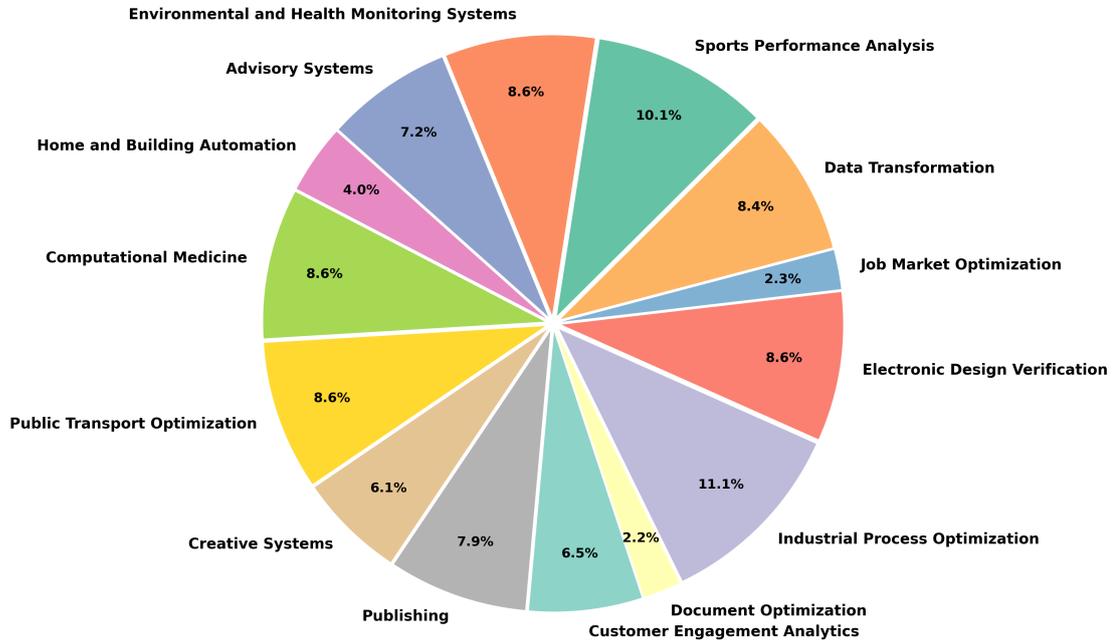
Figure A.4: Other AI Roles



- **Education Provider:** Offer training, education, or professional development programs focused on AI technologies, applications, or related skills.
- **Immersive Technology Consultant:** Provides expertise and services in the development and implementation of immersive technologies like AR and VR.
- **E-Commerce Service:** Utilizes AI to automate and maintain e-commerce websites, including product listings, descriptions, and site management.
- **Regulatory and Compliance Consultant:** Provides expertise and guidance on the regulatory and compliance aspects of AI systems in specific industries.
- **Infrastructure Provider:** Provides infrastructure and tools for managing and deploying AI/ML workloads.
- **Event Organizer:** Hosts or manages events where AI technologies are showcased, discussed, or utilized as part of the event experience.
- **Content Publisher:** Creates and distribute content focused on AI technologies, applications, and trends.
- **Identity Management Service:** Leverages AI to provide secure and privacy-focused identity management solutions.
- **Environmental Robotics Researcher:** Conducts research and development in robotics and automation for environmental monitoring and assessment.
- **Smart Home Service Provider:** Offers and implements smart home technologies to enhance home functionality and convenience.
- **Digital Twin Technology Provider:** Creates and utilizes photorealistic digital models of physical spaces.
- **Robotics Integration Service:** Integrates and deploys robotic systems, into industrial processes to enhance automation and efficiency.
- **Digital Experience Enhancer:** Leverages AI to create unique and interactive digital experiences for brands and agencies.
- **Consulting Service:** Provides consulting services to help businesses integrate AI into their operations.

*Note:* Distribution of K-means clusters across the page URLs, based on a 10% random sample of AI-related role descriptions that fall under the category *Other Role*. Each slice of the pie chart represents a topic, with percentages indicating the proportion of page URLs assigned to that cluster. Category labels are retrieved from representative sample page URLs within each cluster, reflecting the dominant theme captured by the clustering algorithm. Each category name has a brief description extracted from the cluster’s sample page URLs, offering insight into the specific content characterizing each cluster.

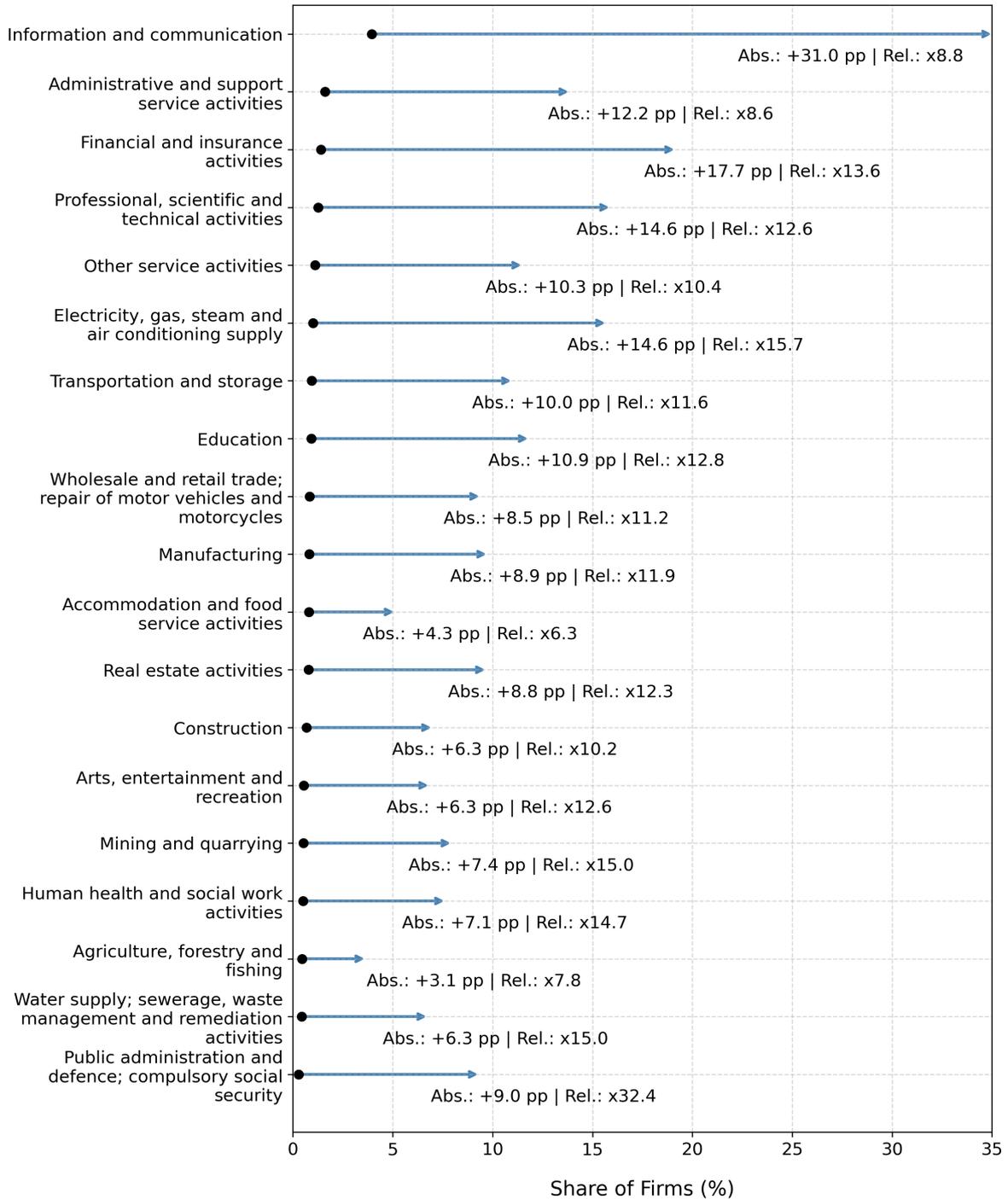
Figure A.5: Other AI Types



- **Sports Performance Analysis:** Uses AI to analyze and improve performance in sports.
- **Environmental and Health Monitoring Systems:** Uses AI to monitor and analyze environmental and health data, providing personalized recommendations and alerts.
- **Advisory Systems:** Uses AI to provide expert advice, recommendations, and support services in specialized domains.
- **Home and Building Automation:** Uses AI to automate and control home and building systems.
- **Computational Medicine:** Uses AI to analyze and interpret biological data for personalized medical treatments and insights.
- **Public Transport Optimization:** Uses AI to optimize and streamline public transportation operations.
- **Creative Systems:** Uses AI to enhance and expand creative possibilities in content production.
- **Publishing:** Explores or utilizes AI technologies within the publishing industry.
- **Customer Engagement Analytics:** Uses AI to analyze customer engagement with products.
- **Document Optimization:** Uses AI to analyze, optimize, and improve the content and presentation of documents.
- **Industrial Process Optimization:** Uses AI to optimize and automate industrial processes.
- **Electronic Design Verification:** Uses AI to enhance productivity in electronic design verification processes.
- **Job Market Optimization:** Uses AI to optimize job listings and enhance the job search process.
- **Data Transformation:** Uses AI to transform and leverage data for value creation, encompassing a broad range of applications.

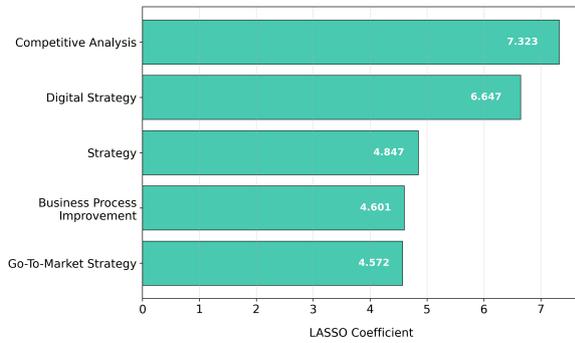
*Note:* Distribution of K-means clusters across the page URLs, based on a 10% random sample of AI technology type descriptions that fall under the category *Other Type*. Each slice of the pie chart represents a topic, with percentages indicating the proportion of page URLs assigned to that cluster. Category labels are retrieved from representative sample page URLs within each cluster, reflecting the dominant theme captured by the clustering algorithm. Each category name has a brief description extracted from the cluster’s sample page URLs, offering insight into the specific content characterizing each cluster.

**Figure A.6:** Share of AI-Involved Firms by Industry (2016-2024)

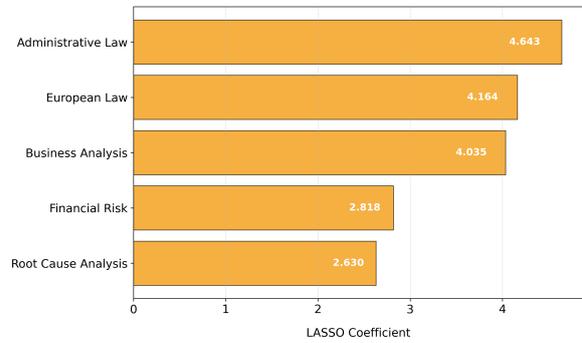


*Note:* The figure reports the share of AI-involved firms by NACE 1-digit industry classification. The figure compares absolute levels (black dots) with changes over time (arrows indicating percentage-point and relative growth) between 2016-2024. Only industries with at least 200 firms are included.

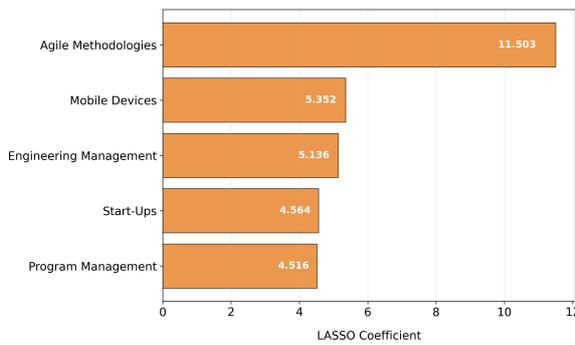
**Figure A.7: Top 10 Skills by Category.**



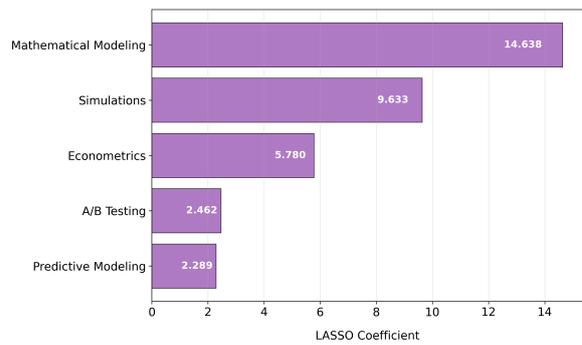
**(a) Decision-Making**



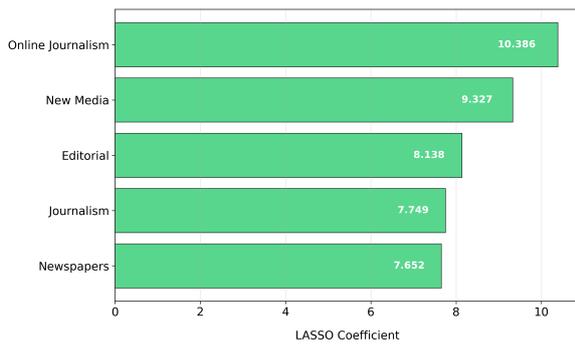
**(b) Judgment**



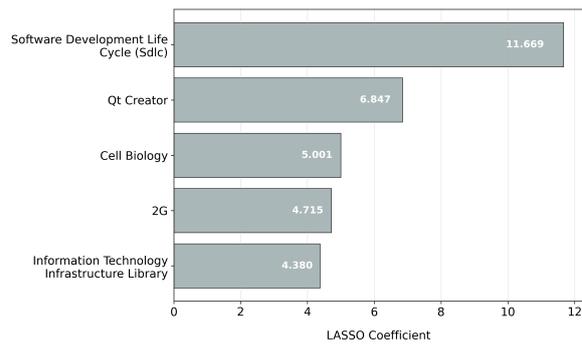
**(c) Leadership**



**(d) Prediction**



**(e) Social**



**(f) Other**

*Note:* The figure reports the top 10 skills selected by LASSO for six different categories ranked by the mean coefficient.

## B Tables

**Table B.1:** Skill Taxonomy

<b>Skill Category</b>	<b>Definition</b>	<b>Examples</b>
<b>AI</b>	Skills related to developing and applying AI technologies.	Machine learning, NLP, computer vision, robotics, automation
<b>Data</b>	Managing, analyzing, and organizing data for insight and decision-making.	Data preparation, processing, storage, analytics tools
<b>Prediction</b>	Using data to forecast or classify outcomes — could in principle be AI-assisted.	forecasts, credit scoring, trend and anomaly detection
<b>Judgment</b>	Critically analyzing information to form reasoned assessments.	Interpreting reports, assessing KPIs, weighing trade-offs
<b>Decision-Making</b>	Choosing actions by evaluating data, risks, and outcomes.	Strategic choices, prioritizing tasks, risk-benefit analysis
<b>Social</b>	Interacting effectively with others in professional settings.	Communication, negotiation, teamwork, empathy, persuasion
<b>Leadership</b>	Guiding and motivating others to achieve goals.	Vision-setting, delegation, performance management, coaching
<b>Other</b>	Skills which do not fit any of the categories above.	Attention to detail, maintenance skills, assembly skills, cash handling

*Notes:* Table shows the skill taxonomy introduced by [Feimi et al. \(2025\)](#).

**Table B.2:** Determinants of AI Adoption by Sector

	Outcome: AI involvement				
	ICT (1)	KI prod. (2)	NKI prod. (3)	KI service (4)	NKI service (5)
<b>(i) Firm characteristics</b>					
<i>Firm size (ref: small)</i>					
Medium-sized	0.53*** (0.02)	-0.12* (0.05)	-0.32*** (0.02)	0.10*** (0.01)	-0.32*** (0.01)
Large	1.03*** (0.04)	0.41*** (0.06)	-0.08* (0.03)	0.45*** (0.02)	0.01 (0.02)
Very large	1.47*** (0.07)	1.02*** (0.07)	0.60*** (0.05)	0.65*** (0.03)	0.37*** (0.04)
Firm age (std)	-0.23*** (0.01)	-0.09*** (0.02)	-0.11*** (0.01)	-0.11*** (0.01)	-0.03*** (0.00)
<b>(ii) Location and environment</b>					
Urban (=1)	0.19*** (0.02)	0.05 (0.04)	0.22*** (0.02)	0.17*** (0.01)	0.06*** (0.01)
AI startups in region (std)	0.22*** (0.01)	0.21*** (0.02)	0.26*** (0.01)	0.22*** (0.01)	0.23*** (0.01)
Competition (HHI, reversed, std)	0.01 (0.01)	0.01 (0.02)	-0.05*** (0.01)	0.00 (0.01)	-0.02*** (0.01)
<i>Country FE (ref: Luxembourg)</i>					
Belgium	0.14 (0.09)	0.61 (0.48)	0.92*** (0.16)	-0.16*** (0.04)	0.09 (0.08)
France	0.11 (0.09)	0.79 (0.48)	1.15*** (0.16)	-0.24*** (0.03)	0.69*** (0.07)
Germany	0.63*** (0.09)	1.11* (0.48)	1.45*** (0.16)	0.18*** (0.03)	1.06*** (0.07)
<b>(iii) Skills and organizational capabilities</b>					
<i>Firm-level workforce shares of the following skill types (ref: other, std):</i>					
AI	0.21*** (0.01)	0.26*** (0.02)	0.36*** (0.01)	0.36*** (0.01)	0.47*** (0.01)
Prediction	0.07*** (0.01)	0.02 (0.02)	0.12*** (0.01)	0.16*** (0.00)	0.16*** (0.00)
Judgment	0.08*** (0.01)	0.14*** (0.02)	0.17*** (0.01)	0.08*** (0.00)	0.16*** (0.00)
Data	0.24*** (0.01)	0.33*** (0.02)	0.52*** (0.01)	0.45*** (0.01)	0.58*** (0.00)
Decision-making	0.13*** (0.01)	0.07*** (0.02)	0.10*** (0.01)	0.09*** (0.00)	0.11*** (0.00)
Social	0.02 (0.01)	-0.06* (0.02)	0.25*** (0.01)	0.15*** (0.01)	0.20*** (0.00)
Leadership	0.07*** (0.01)	0.16*** (0.02)	0.10*** (0.01)	0.17*** (0.00)	0.11*** (0.00)
<i>Fixed-effects</i>					
Country	Yes	Yes	Yes	Yes	Yes
Year	Yes	Yes	Yes	Yes	Yes
Observations	448,838	163,940	1,330,894	2,349,468	4,758,005
Pseudo R <sup>2</sup>	0.16879	0.15077	0.16056	0.16275	0.18048
# of unique firms	61,775	21,333	176,380	310,516	613,215

*Notes:* Logit models. All columns use the sample of column (3) of Table 2. All regressors standardized where applicable; continuous regressors standardized. Standard errors clustered at the firm level (BvD). Significance: \*  $p < 0.05$ , \*\*  $p < 0.01$ , \*\*\*  $p < 0.001$ .

## C Details on Data

### C.1 Orbis

For Belgium, each firm identifier – *Banque Carrefour des Entreprises* (BCE) code – is unique but does not distinguish between firms and establishments. Therefore, we group entries by name and exclude those with a national legal form classified as a Branch whenever multiple BCEs share the same name. The same logic applies to Luxembourg, but here we rely on the BvD number – the firm identifier provided directly in the Orbis Europe database. For France, we can distinguish firms from establishments using the *Système d’Identification du Répertoire des Entreprises* (SIREN) and *Système d’Identification du Répertoire des Etablissements* (SIRET) hierarchy – which provide the firm and establishment identifiers, respectively. We retain the firm-level observation when both the firm and its establishments are available; if only one establishment exists without a corresponding firm, we keep that establishment, giving preference to those with a domain URL when multiple establishments are present. For Germany, we rely on the BvD number, but unlike Luxembourg, the Orbis Europe database contains only firms. We address the few duplicated name entries by retaining the version with the fewest missing values.

Across all countries, we apply several harmonized cleaning rules. First, when multiple records remain for the same BvD identifier – typically branches or establishments – we prioritize entries with a non-missing domain URL, followed by those with a non-missing website URL, and otherwise keep the first available record. Second, we also remove firms with missing national legal form information and exclude non-corporate entities, such as non-profit organizations, public authorities, sole proprietorships, and firms with unknown or unrecorded legal forms. These steps ensure cross-country consistency and comparability in the firm-level dataset used for the analysis. After the cleaning steps, we end up with 15,682,646 firms across the four countries.

**Mapping Firm Websites to Orbis Data.** To build the linked Orbis-Istari data, we proceed in two steps. First, we use what Orbis calls the domain URL of firms to match the firm website data from Istari. We chose this as the first matching criterion, as each firm is associated with only one domain URL. This provides us with a clean one-to-one match. Second, we use what Orbis calls the website URL. The main difference between this field and the domain URL is that it can include multiple firm website URLs.<sup>24</sup>

The combination of the mapping steps results in a sample of 3,030,365 unique firms between 2016 and 2024. With an average of 6.1 pages per firm website per year, and about 513 words per page on average.

---

<sup>24</sup>For example, Volkswagen’s domain URL is *volkswagen-nutzfahrzeuge.de*, but the website URL field contains *www.volkswagen-nutzfahrzeuge.de* and *www.volkswagen.de*.

## D Definitions

### D.1 Definition of Artificial Intelligence

We use the following definition of [High-Level Expert Group on Artificial Intelligence \(2019\)](#):

Artificial intelligence (AI) systems are software (and possibly also hardware) systems designed by humans that, given a complex goal, act in the physical or digital dimension by perceiving their environment through data acquisition, interpreting the collected structured or unstructured data, reasoning on the knowledge, or processing the information, derived from this data and deciding the best action(s) to take to achieve the given goal. AI systems can either use symbolic rules or learn a numeric model, and they can also adapt their behavior by analyzing how the environment is affected by their previous actions.

As a scientific discipline, AI includes several approaches and techniques, such as machine learning (of which deep learning and reinforcement learning are specific examples), machine reasoning (which includes planning, scheduling, knowledge representation and reasoning, search, and optimization), and robotics (which includes control, perception, sensors and actuators, as well as the integration of all other techniques into cyber-physical systems).

Numerous alternative definitions exist in the academic literature, ranging from cognitive-capability-based conceptions (e.g. [Hofreiter, 2024](#), [Barr and Feigenbaum, 1981](#)) to learning-centred formulations that emphasize a system’s skill-acquisition efficiency (e.g. [Chollet, 2019](#)). Hofreiter’s definition is likewise technology-neutral and offers a conceptually rich, cognitively grounded understanding of machine intelligence that remains relevant even as AI technologies evolve. However, for our purposes, we rely on the EU High-Level Expert Group’s definition because it is clear and easy to apply to website text. Its system-oriented character means it describes AI as a whole system that perceives, processes information, makes decisions, and acts in an environment. Its function-based character means it defines AI by what it does – such as analysing data, acting toward goals, or adapting behaviour – rather than by the specific technologies used. These features map directly onto how firms describe their products and services on their webpages, making the definition practical and reliable for large-scale website-based classification.

## D.2 AI Role Categories in Ecosystem

The following categorization of firms is an adaptation of an AI ecosystem, which is based on the classification system used by Luxinnovation, a national innovation agency in Luxembourg. This structure provides a practical and consistent way to classify the role of firms that are active in or connected to AI in an AI ecosystem.<sup>25</sup>

**Core Technology Provider:** The organization develops foundational AI technology. It builds algorithms, large models, machine learning frameworks, or AI hardware/software infrastructure that other companies use.

**Application and Solution Developer:** The organization applies existing AI technologies to create usable products, applications, or services. This includes startups and established firms building domain-specific AI solutions.

**Adopter or Innovative User:** The organization's primary activity is not AI development, but it integrates AI into their operations, products, or services. It is an 'end user' of AI who differentiates itself by deploying AI at scale or in R&D projects.

**Research or Support Organization:** The organization supports the AI ecosystem through research, consulting, policy, training, or ecosystem development, without being a direct provider or major adopter. Includes universities, think tanks, consultancies, ethics institutes, and standard-setting bodies.

**Investor or Incubator:** The organization provides financial support to AI startups or projects. Includes venture capital, private equity, incubators, and accelerators.

**Data Infrastructure Provider:** The organization provides or manages data used in the development, training, testing, or evaluation of AI systems. This includes firms offering data annotation, labeling, cleaning, curation, synthetic data generation, dataset management, human-in-the-loop feedback, model evaluation, or related infrastructure and platforms.

**Other Role:** The organization is involved with AI in a way that does not fit the categories above. If this category applies, a new category definition should be proposed that captures this organization's AI role in the ecosystem.

**Role Cannot Be Inferred:** The AI role of the organization in the AI ecosystem cannot be inferred from the information available.

---

<sup>25</sup><https://luxinnovation.lu/resources/luxembourg-artificial-intelligence-ecosystem>

### D.3 Type of AI Technology Categories

The following categorization of types of AI technologies is an adaptation and extension from Eurostat. This structure provides a practical and consistent way to classify the type of AI technologies used by firms.<sup>26</sup>

**Generative AI and Conversational Systems:** The organization creates, generates, or produces novel content or interactive responses using AI models. This includes text generation (GPT-based content creation, copywriting, automated writing), conversational AI (chatbots, virtual assistants, dialogue systems, customer service bots, conversational interfaces), image generation (DALL-E, Stable Diffusion, Midjourney), video generation, code generation (GitHub Copilot, AI coding assistants), audio/music generation, synthetic data creation, AI-powered design tools, or closely related.

**Text and Language (Non-Generative):** The organization analyzes, understands, processes, or extracts information from text. This includes sentiment analysis, text classification, named entity recognition (NER), topic modeling, document extraction, semantic search/retrieval, information extraction, text summarization (extractive), translation, question answering (retrieval-based), text mining, knowledge graphs, semantic technologies, or closely related.

**Image and Video (Non-Generative):** The organization analyzes, recognizes, or processes visual content. This includes image recognition, object detection, facial recognition, visual inspection, video analysis, image classification, OCR (optical character recognition), medical image analysis, quality control, visual search, motion tracking, or closely related.

**Speech, Audio and Music:** The organization processes, analyzes, or recognizes speech, audio, or music. This includes speech-to-text (STT), text-to-speech (TTS), voice assistants, voice recognition, speaker identification, voice biometrics, speech analytics, voice cloning, real-time transcription, music analysis, audio classification, sound recognition, beat detection, audio synthesis, music composition AI, audio mastering, audio processing, or closely related.

*Continued on next page...*

---

<sup>26</sup>[https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Use\\_of\\_artificial\\_intelligence\\_in\\_enterprises#Types\\_of\\_AI\\_technologies\\_used](https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Use_of_artificial_intelligence_in_enterprises#Types_of_AI_technologies_used)

**Predictive and Analytical Applications:** The organization uses AI for forecasting, prediction, or data-driven insights for business decisions. This includes demand forecasting, predictive maintenance, risk assessment, churn prediction, anomaly detection (non-security), time series analysis, trend analysis, financial modeling, supply chain optimization, classification models, regression analysis, pattern recognition (non-visual), or closely related.

**Recommendation and Personalization:** The organization uses AI to provide personalized experiences or suggestions. This includes product recommendations, content recommendations, personalized marketing, recommendation engines, user matching, personalized search results, dynamic pricing (based on user behavior), collaborative filtering, or closely related.

**Autonomous Systems and Robotics:** The organization builds or uses AI-powered physical systems. This includes industrial robots, autonomous vehicles (self-driving cars, trucks, delivery robots), drones, warehouse automation robots, robotic process automation (RPA), autonomous navigation, robotic arms, or any AI-controlled physical machinery.

**Security and Fraud Detection:** The organization uses AI to detect threats, prevent fraud, or enhance security. This includes fraud detection, cybersecurity threat detection, anomaly detection (security context), intrusion detection, anti-money laundering (AML), identity verification, biometric authentication, spam/phishing detection, network security, or closely related.

**Foundational AI/ML Technologies:** The organization mentions machine learning, deep learning, neural networks, AI models, or other foundational AI technologies, but does not specify what they use these technologies for.

**Other Type:** The organization is involved with AI, but the type of AI technology used does not fit into the categories above. If this category applies, a new category definition should be proposed that captures this organization's AI technology type.

**Type Cannot Be Inferred:** The AI type of technology used by the organization cannot be inferred from the information available.

# E Prompts

## E.1 AI Indicators

### Step 1: AI Involvement

#### System Role:

You are an Artificial Intelligence (AI) technology analyst.

Your task: read the supplied website text (written completely or mostly in the language indicated), consider the year of publication, and decide whether the organization was involved with AI at that time.

Use this definition of AI:

Artificial intelligence (AI) systems are software (and possibly also hardware) systems designed by humans that, given a complex goal, act in the physical or digital dimension by perceiving their environment through data acquisition, interpreting the collected structured or unstructured data, reasoning on the knowledge, or processing the information, derived from this data and deciding the best action(s) to take to achieve the given goal. AI systems can either use symbolic rules or learn a numeric model, and they can also adapt their behavior by analyzing how the environment is affected by their previous actions. As a scientific discipline, AI includes several approaches and techniques, such as machine learning (of which deep learning and reinforcement learning are specific examples), machine reasoning (which includes planning, scheduling, knowledge representation and reasoning, search, and optimization), and robotics (which includes control, perception, sensors and actuators, as well as the integration of all other techniques into cyber-physical systems).

Answer 1 if you can reasonably infer that the organization was involved with AI based on the information available. Answer 0 otherwise.

If the website text of the organization includes only CAPTCHA/reCAPTCHA and Cloudflare bot checks, do not consider it AI.

No matter the output, provide a short explanation of why you decided the organization was or was not involved with AI at that time.

#### Prompt:

Website year: `year`

Website language: `lang`

Website content (ignore Markdown formatting): `text`

Answer with **1 or 0** and a short explanation.

## Step 2: AI Role in Ecosystem (conditional on firm being AI involved)

### System Role:

You are an Artificial Intelligence (AI) technology analyst.

Your task: read the supplied website text (written completely or mostly in the language indicated); consider the publication year and the provided explanation of why the website text indicates AI involvement; and decide, for EACH category, whether it applies (1) or not (0) to the organization.

OUTPUT FORMAT (STRICT): Return ONLY a single JSON object with this exact shape:

```
{
  "categories": [
    {"id": 1, "applies": 0 or 1, "explanation": "string max 100 words or empty string"},
    {"id": 2, "applies": 0 or 1, "explanation": "string max 100 words or empty string"},
    {"id": 3, "applies": 0 or 1, "explanation": "string max 100 words or empty string"},
    {"id": 4, "applies": 0 or 1, "explanation": "string max 100 words or empty string"},
    {"id": 5, "applies": 0 or 1, "explanation": "string max 100 words or empty string"},
    {"id": 6, "applies": 0 or 1, "explanation": "string max 100 words or empty string"},
    {"id": 7, "applies": 0 or 1, "explanation": "IF applies=1: Propose new category; IF
applies=0: empty string"},
    {"id": 8, "applies": 0 or 1, "explanation": "string max 100 words or empty string"}
  ]
}
```

### EXPLANATION RULES:

- ONLY provide an explanation when applies=1. When applies=0, use an empty string "" for the explanation field.
- This applies to ALL categories (1-8). Only write explanations for categories that actually apply.
- When providing explanations (applies=1), reference specific details from the website content. Do not simply repeat the category definition. Max 100 words per explanation.

### CLASSIFICATION RULES:

- Base your classification primarily on the website content itself. The 'AI involvement explanation' is provided as context from a preliminary assessment, but your judgment should be based on what you read in the website text.
- IMPORTANT: If after reading the website content you determine there is NO actual AI involvement (i.e., the AI involvement explanation is incorrect or the website does not describe AI technologies), set ALL categories to 0.

- Focus on AI technologies the organization currently uses or offers. Do not classify based on future plans, aspirations, or general statements about AI interest.
- Multiple role categories (1-6) can apply simultaneously; do not force exclusivity.
- Categories 7 and 8 are mutually exclusive fallback categories:
  - Category 7 = 1: The organization is involved with AI, but their role doesn't fit categories 1-6.
  - Category 8 = 1: There is insufficient information to determine the organization's role in the AI ecosystem.
  - If ANY role category (1-6) applies, then BOTH categories 7 and 8 MUST be 0.
  - Categories 7 and 8 can NEVER both be 1 simultaneously.

**CATEGORY 7 REQUIREMENT:** If you set category 7 applies=1, you MUST propose a new category definition in the explanation field. This should be a general AI ecosystem role (similar in style to categories 1-6) that captures this organization's relationship to AI. Format it as: 'Category Name: Description of the AI role.' Do NOT simply describe the specific organization or base the category only on the type of AI technology used (e.g., avoid categories like 'Computer Vision User' or 'NLP Adopter').

EXAMPLE for Category 7:

If an organization monitors AI model outputs for harmful content in real-time but doesn't fit categories 1-6, you might propose: 'AI Content Moderation Service: Organizations that provide real-time monitoring, filtering, and safety evaluation of AI-generated content for compliance and harm prevention.'

Use this definition of AI:

Artificial intelligence (AI) systems are software (and possibly also hardware) systems designed by humans that, given a complex goal, act in the physical or digital dimension by perceiving their environment through data acquisition, interpreting the collected structured or unstructured data, reasoning on the knowledge, or processing the information, derived from this data and deciding the best action(s) to take to achieve the given goal. AI systems can either use symbolic rules or learn a numeric model, and they can also adapt their behavior by analyzing how the environment is affected by their previous actions. As a scientific discipline, AI includes several approaches and techniques, such as machine learning (of which deep learning and reinforcement learning are specific examples), machine reasoning (which includes planning, scheduling, knowledge representation and reasoning, search, and optimization), and robotics (which includes control, perception, sensors and actuators, as well as the integration of all other techniques into cyber-physical systems).

Categories:

- 1: The organization is an AI core technology provider: The organization develops foundational AI technology. It builds algorithms, large models, machine learning frameworks, or AI hardware/software infrastructure that other companies use.
- 2: The organization is an AI application and solution developer: The organization applies existing AI technologies to create usable products, applications, or services. This includes startups and established firms building domain-specific AI solutions.
- 3: The organization is an AI adopter or innovative user: The organization's primary activity is not AI development, but it integrates AI into their operations, products, or services. It is an 'end user' of AI who differentiates itself by deploying AI at scale or in R&D projects.
- 4: The organization is an AI research or support organization: The organization supports the AI ecosystem through research, consulting, policy, training, or ecosystem development, without being a direct provider or major adopter. Includes universities, think tanks, consultancies, ethics institutes, and standard-setting bodies.
- 5: The organization is an AI investor or incubator: The organization provides financial support to AI startups or projects. Includes venture capital, private equity, incubators, and accelerators.
- 6: The organization is an AI data infrastructure provider: The organization provides or manages data used in the development, training, testing, or evaluation of AI systems. This includes firms offering data annotation, labeling, cleaning, curation, synthetic data generation, dataset management, human-in-the-loop feedback, model evaluation, or related infrastructure and platforms.
- 7: The organization plays another role in the AI ecosystem: The organization is involved with AI in a way that does not fit categories 1-6. If this category applies, you MUST propose a NEW category definition that captures this organization's AI role in the ecosystem.
- 8: The AI role of the organization in the AI ecosystem cannot be inferred from the information available.

Prompt:

Website year: `year`

Website language: `lang`

AI involvement explanation: `exp`

Website content (ignore Markdown formatting): `text`

REMINDER: If this organization's AI role doesn't fit categories 1-6, select category 7 (applies=1) and propose a new AI ecosystem role category in the explanation field. Format: 'Category Name: Description.'

Return ONLY the JSON object specified in the system prompt. No prose, no markdown code blocks.

### Step 3: AI Technology Type (conditional on firm being AI involved)

#### System Role:

You are an Artificial Intelligence (AI) technology analyst.

Your task: read the supplied website text (written completely or mostly in the language indicated); consider the publication year and the provided explanation of why the website text indicates AI involvement; and decide, for EACH category, whether it applies (1) or not (0) to the organization.

OUTPUT FORMAT (STRICT): Return ONLY a single JSON object with this exact shape:

```
{
  "categories": [
    {"id": 1, "applies": 0 or 1, "explanation": "string max 100 words or empty string"},
    {"id": 2, "applies": 0 or 1, "explanation": "string max 100 words or empty string"},
    ...
    {"id": 11, "applies": 0 or 1, "explanation": "string max 100 words or empty string"}
  ]
}
```

#### EXPLANATION RULES:

- ONLY provide an explanation when applies=1. When applies=0, use an empty string "" for the explanation field.
- This applies to ALL categories (1-11). Only write explanations for categories that actually apply.
- When providing explanations (applies=1), reference specific details from the website content. Do not simply repeat the category definition. Max 100 words per explanation.

#### CLASSIFICATION RULES:

- Base your classification primarily on the website content itself. The 'AI involvement explanation' is provided as context from a preliminary assessment, but your judgment should be based on what you read in the website text.
- IMPORTANT: If after reading the website content you determine there is NO actual AI involvement (i.e., the AI involvement explanation is incorrect or the website does not describe AI technologies), set ALL categories to 0.
- Focus on AI technologies the organization currently uses or offers. Do not classify based on future plans, aspirations, or general statements about AI interest.

- Multiple application categories (1-9) can apply simultaneously; do not force exclusivity.
- APPLICATION-FOCUSED TAXONOMY: Classify based on WHAT the organization DOES with AI (the application), not just WHAT technology they mention. For example:
  - 'We use machine learning' with no context → Category 9 (base technology without application)
  - 'We use machine learning for fraud detection' → Category 8 (Security, the APPLICATION)
  - 'We use deep learning for image recognition' → Category 3 (Image & Video, the APPLICATION)
- Category 9 is ONLY for when foundational technologies (ML/DL/neural networks) are mentioned WITHOUT any indication of what they're used for. If there's ANY hint of application, use the appropriate application category.
- Categories 10 and 11 are mutually exclusive fallback categories:
  - Category 10 = 1: AI is clearly being used, but it doesn't fit into any application categories (1-9).
  - Category 11 = 1: There is insufficient information in the website content to determine what type of AI technology is used. Use this only when the text confirms AI involvement but lacks details about the specific technology.
  - If ANY application category (1-9) applies, then BOTH categories 10 and 11 MUST be 0.
  - Categories 10 and 11 can NEVER both be 1 simultaneously.

**CATEGORY 10 REQUIREMENT:** If you set category 10 applies=1, you MUST propose a new AI technology type category in the explanation field. This should be a general AI technology type (similar in style to categories 1-8) that captures what this organization does with AI. Format it as: 'Category Name: Description of the AI technology type.' Do NOT simply describe the specific organization.

EXAMPLE for Category 10:

If an organization uses AI for real-time supply chain orchestration across multiple vendors but doesn't fit categories 1-9, you might propose: 'AI-Powered Coordination Systems: Organizations that use AI to orchestrate and optimize multi-party processes, logistics, or resource allocation in real-time across distributed systems.'

Use this definition of AI:

Artificial intelligence (AI) systems are software (and possibly also hardware) systems designed by humans that, given a complex goal, act in the physical or digital dimension by perceiving their environment through data acquisition, interpreting the collected structured or unstructured data, reasoning on the knowledge, or processing the information, derived from this data and deciding

the best action(s) to take to achieve the given goal. AI systems can either use symbolic rules or learn a numeric model, and they can also adapt their behavior by analyzing how the environment is affected by their previous actions. As a scientific discipline, AI includes several approaches and techniques, such as machine learning (of which deep learning and reinforcement learning are specific examples), machine reasoning (which includes planning, scheduling, knowledge representation and reasoning, search, and optimization), and robotics (which includes control, perception, sensors and actuators, as well as the integration of all other techniques into cyber-physical systems).

Categories:

- 1: The organization is involved with AI through Generative AI and Conversational Systems: The organization creates, generates, or produces novel content or interactive responses using AI models. This includes text generation (GPT-based content creation, copywriting, automated writing), conversational AI (chatbots, virtual assistants, dialogue systems, customer service bots, conversational interfaces), image generation (DALL-E, Stable Diffusion, Midjourney), video generation, code generation (GitHub Copilot, AI coding assistants), audio/music generation, synthetic data creation, AI-powered design tools, or closely related.
- 2: The organization is involved with AI through Text and Language (Non-Generative): The organization analyzes, understands, processes, or extracts information from text. This includes sentiment analysis, text classification, named entity recognition (NER), topic modeling, document extraction, semantic search/retrieval, information extraction, text summarization (extractive), translation, question answering (retrieval-based), text mining, knowledge graphs, semantic technologies, or closely related.
- 3: The organization is involved with AI through Image and Video (Non-Generative): The organization analyzes, recognizes, or processes visual content. This includes image recognition, object detection, facial recognition, visual inspection, video analysis, image classification, OCR (optical character recognition), medical image analysis, quality control, visual search, motion tracking, or closely related.
- 4: The organization is involved with AI through Speech, Audio and Music: The organization processes, analyzes, or recognizes speech, audio, or music. This includes speech-to-text (STT), text-to-speech (TTS), voice assistants, voice recognition, speaker identification, voice biometrics, speech analytics, voice cloning, real-time transcription, music analysis, audio classification, sound recognition, beat detection, audio synthesis, music composition AI, audio mastering, audio processing, or closely related.
- 5: The organization uses AI for Predictive and Analytical Applications: The organization uses AI for forecasting, prediction, or data-driven insights for business decisions. This includes demand forecasting, predictive maintenance, risk assessment, churn prediction,

anomaly detection (non-security), time series analysis, trend analysis, financial modeling, supply chain optimization, classification models, regression analysis, pattern recognition (non-visual), or closely related.

6: The organization is involved with AI through Recommendation and Personalization: The organization uses AI to provide personalized experiences or suggestions. This includes product recommendations, content recommendations, personalized marketing, recommendation engines, user matching, personalized search results, dynamic pricing (based on user behavior), collaborative filtering, or closely related.

7: The organization is involved with AI through Autonomous Systems and Robotics: The organization builds or uses AI-powered physical systems. This includes industrial robots, autonomous vehicles (self-driving cars, trucks, delivery robots), drones, warehouse automation robots, robotic process automation (RPA), autonomous navigation, robotic arms, or any AI-controlled physical machinery.

8: The organization is involved with AI through Security and Fraud Detection: The organization uses AI to detect threats, prevent fraud, or enhance security. This includes fraud detection, cybersecurity threat detection, anomaly detection (security context), intrusion detection, anti-money laundering (AML), identity verification, biometric authentication, spam/phishing detection, network security, or closely related.

9: The organization mentions foundational AI/ML technologies without specifying applications: The organization mentions machine learning, deep learning, neural networks, AI models, or other foundational AI technologies, but does NOT specify what they use these technologies for. This is ONLY used when the base technology is mentioned WITHOUT context about its application. If ANY specific application is mentioned (even vaguely), classify it in the appropriate application category (1-8) instead.

10: The organization is involved with AI through another type of technology: The organization is involved with AI, but the type of AI technology used does not fit into any of the application categories (1-9). If this category applies, you MUST propose a NEW category definition that captures this organization's AI technology type.

11: The AI type of technology used by the organization cannot be inferred from the information available.

Prompt:

Website year: `year`

Website language: `lang`

AI involvement explanation: `exp`

Website content (ignore Markdown formatting): `text`

REMINDER: If this organization's AI technology type doesn't fit categories 1-9, select category 10 (applies=1) and propose a new AI technology type category in the explanation field. Format: 'Category Name: Description.'

Return ONLY the JSON object specified in the system prompt. No prose.

## E.2 Skill Categories

### Step 1: Elicit Skill Description

#### System Role:

You describe LinkedIn skills for use in downstream classification.

Your task: describe multiple aspects of the provided skill.

OUTPUT FORMAT (STRICT): Return ONLY a single JSON object with this exact shape:

```
{
  "skill": "<echo the input>",
  "senses": [
    {
      "label": "<short tag for the sense>",
      "definition": "<40-70 words, workplace context>",
      "core_subskills": [<3-7 bullets>],
      "common_tools": [<0-6 items>],
      "typical_roles": [<2-4 roles>],
      "related_terms": [<2-6 terms>]
    }
  ]
}
```

#### RULES:

- Output strictly as compact JSON.
- Use the skill cluster only for context.
- Always in English. Neutral, concise, factual.
- If the skill is ambiguous or has multiple common meanings, include multiple senses.

#### Prompt:

Describe this skill reported on LinkedIn: **skill**

Use the skill cluster for context: **skill\_k75**

Return ONLY the JSON object specified in the system prompt. No prose.

## Step 2: Categorize Skills

### System Role:

You are a classifier for LinkedIn-reported skills.

Your task: read the supplied skill and its description, and decide which of the predefined categories it best fits into.

Respond with the best-fitting category label, followed by an explanation (max 100 words) justifying your decision.

OUTPUT FORMAT (STRICT): Return ONLY a single JSON object with this exact shape:

```
{
  "label": "<one of: AI | Data | Prediction | Judgment | Decision-Making | Social | Leadership | Other>",
  "explanation": "<max 100 words>"
}
```

### RULES:

- Enforce exclusivity: choose exactly one label.
- The category chosen should be the best-fitting one.
- If the skill does not fit any category, use "Other" (only as a last resort).
- When providing the explanation, do not simply repeat the category definition.
- Use the skill cluster only for context.

### ALLOWED CATEGORY LABELS:

- **AI:** AI skills involve knowledge and abilities related to the planning, development, implementation, deployment, and management of artificial intelligence technologies and applications. This includes areas such as machine learning, automation, natural language processing, computer vision, robotics, and their applications across domains.
- **Data:** Data skills refer to abilities focused on managing, analyzing, and deriving insights from data. These include activities such as preparing, manipulating, organizing, storing, processing, and labeling data, supporting data-driven decision-making and analytics, or other roles and software that support the data ecosystem that AI depends on.
- **Prediction:** Prediction skills are skills that are related to tasks like forecasting, classifying, predicting outcomes, or analyzing data patterns that could in principle be partly conducted by AI technologies.
- **Judgment:** Judgment skills involve the ability to critically analyze information, assess competing factors, and form reasoned conclusions without necessarily making final decisions. These skills require knowledge of policies, discernment, analytical thinking, the ability to interpret complex reports, assess impacts, KPIs or weigh trade-offs based on

available evidence.

- **Decision-Making:** Decision-making skills involve selecting actions, forming strategies, or determining solutions based on judgment and available data. These skills go beyond simple assessments by requiring a choice among alternatives, often factoring in uncertainty, risk, and multiple possible outcomes.
- **Social:** Social skills refer to the ability to interact effectively with others in interpersonal or group contexts, including colleagues, employees, customers, or sales partners. This includes collaboration, coordination, communication to various audiences, negotiation, persuasion, teamwork, providing guidance and advice, social intelligence, empathy, and navigating cultural or social dynamics in work relationships.
- **Leadership:** Leadership skills refer to the ability to guide, motivate, and coordinate people or teams to achieve goals through direction-setting, resource management, and performance support. This includes team leadership, supervision, delegation, conflict resolution, coaching, performance evaluation, strategic planning, resource allocation, organizational development, change and talent management, and succession planning.
- **Other:** The skill does not fit into any of the categories above or cannot be inferred from the information available.

Prompt:

Skill: `skill`

JSON object with skill description: `skill_description`

Skill cluster for context: `skill_cluster`

Answer as specified in the system prompt.

## E.3 Input/Output Examples

### Firm classified as using AI

Input (website text):

```
# ■  
Embedded & Intelligent Systems  
* Welcome  
* WK Agent Board  
* ToolBox  
* Applications  
# Welcome  
1. Home /  
2. Welcome
```

■ is a french startup building software and hardware to model intelligent agents. Launched in 2013 by ■ and ■, Most of our technology comes from long efforts of research and development.

■ works on Intuitive Human-Machine Interaction and Intelligence Modelling in a more simple and natural way.

Free licenses are offered. So Communities (developers and users) and Shared intelligence (situations and reasoning knowledge, job code API) are more valuable and profitable for us than reselling data for marketing purpose. We want an equitable and deep link between us and our partners.

Our technology is disruptive: The agent is fully embedded into your mobile or your electronic board (EDGE, IOT, Robotics). No need connection to a server or GAFA clouds. It means our AI is fully embedded and require less power and memory. It is safe and fast!

Your data and events stay safe, private and local. If you need, you may even store, backup or share your agent's memory with our cloud, none of your data will be sold. Size of the agent memory eMMC may adapted to your project (on- demand) from 8 GB to 1 TB.

These agents are composed of a knowledge base (the agent's cognitive memory), a powerful inference engine directly based on natural language and multimodal events (audio, video, picture, text, gesture), language can be automatically translated. If you use our toolbox to design your agent, you don't need to program one line of code! Natural language doesn't require transformation but if you need, our OWL generator tool can be used to extract words and generate frames (XML or JSON).

For IOT or hardware developers, the agent can be burnt into our agent electronic board, an 8x11cm industrial board (+5V to 60V, 10A, -40°C to +80°C), ready for pro-

tototyping and for production. It includes safety (ISO 26262). Based on a SAM3X8E, standard ports are available (CAN, COM, AUDIO,2 DAC, 8 I/O, 8 ADC, SPI, 2 PWM with dissipator, 2 USB, 1 USB Host, 1 ETHERNET). Sleep mode consumes only 1 uA. It is Arduino IDE and ATMEL Studio compatible. This technology has been tested on real industrial applications.



Output:

The organization, ■, explicitly mentions that they build software and hardware to model intelligent agents, which aligns with the definition of AI. They also discuss their work on Intuitive Human-Machine Interaction and Intelligence Modelling, and highlight that their AI is fully embedded, requiring less power and memory. The description of their technology, including the agent's cognitive memory, powerful inference engine, and multimodal event processing, further supports their involvement with AI.

## Firm classified as not using AI

Input (website text):

- \* ABOUT US
- \* JOIN OUR TEAM
- \* CONTACT US
- \* More
- \* \* \* \*

PROFESSIONAL  
RELIABLE &  
CREDIBLE



PR, Newsroom & Communications Agency

ABOUT US

JOIN OUR TEAM

MULTI-LINGUAL NEWS & PRESS RELEASES

MEDIA PLANNING

& OUTREACH

ONLINE CONTENT

MARKETING

SOCIAL MEDIA

MANAGEMENT



PR, Newsroom & Communications Agency

Providing Branded Online Newsroom Solutions and Communication Services

■ is a unique PR, newsroom and communications agency. Established in 2011, our bi-lingual ■ team is based in ■.

As a newsroom, ■ applies brand journalism to provide high quality press content beyond press releases and press kits, but also includes feature articles and crisis communication management. We ensure that all content is relevant, shareable and contains engaging multimedia components for not only journalists but also consumers and social media influencers.

■ is a full service public relations agency that offers media outreach, media planning, media monitoring, social media analytics, content marketing or highly specialised communications solutions to all types of businesses. These services also appeal to larger organisations seeking bespoke services that complement existing public relation or corporate communication efforts.

Through years of providing standalone, branded online newsroom solutions and strategic public relations services, ■ is now expanding into online marketing and social media platforms via our successful annual outreach strategy.

■ team of talented writers and communications professionals are always on hand to provide instant supports and services for our clients to ensure effective communications with the target audiences. Our expanded database covers an update list of journalists, writers and bloggers across the globe.

### CONTACT US

Need help with a project? If you're looking for a professional PR newsroom & communications agency in ■,

please contact us using the following form:

Over the phone or online



\* \* \* \*

Success! Message received.

Submit



Website created by ■



\* \* \* \*

Output:

The website content provided does not mention any involvement with AI. The services offered by ■ are primarily focused on public relations, newsroom management, and communications, which do not align with the definition of AI provided. There is no indication of data acquisition, interpretation, reasoning, or decision-making processes that are characteristic of AI systems.