OECD Science, Technology and Industry Working Papers

This publication contributes to the OECD's Artificial Intelligence in Work, Innovation, Productivity and Skills (AI-WIPS) programme, which provides policymakers with new evidence and analysis to keep abreast of the fast-evolving changes in AI capabilities and diffusion and their implications for the world of work. The programme aims to help ensure that adoption of AI in the world of work is effective, beneficial to all, people-centred and accepted by the population at large. AI-WIPS is supported by the German Federal Ministry of Labour and Social Affairs (BMAS) and will complement the work of the German AI Observatory in the Ministry's Policy Lab Digital, Work & Society. For more information, visit *https://oecd.ai/work-innovation-productivity-skills* and *https://denkfabrik-bmas.de/*.

Federal Ministry
of Labour and Social Affairs

Note to Delegations:

This document is also available on O.N.E under the reference code:

DSTI/CIIE/WPIA(2022)7/FINAL

Table of contents

FIGURES

TABLES

Acknowledgments

The authors would like to thank Chiara Criscuolo for the guidance and useful comments throughout the project, Sara Calligaris, Hélène Dernis, Milenko Fadic, Luca Fontanelli, Andrea Greppi, Natia Mosiashvili, Dirk Pilat for their support providing relevant data, helpful discussions or suggestions, Francesca Borgonovi, Daniela Glocker, Marco Grazzi, Justine Herve, Hubert Plisiecki, Marina Romiti and Helke Seitz for valuable comments, participants to The Productivity Institute's 2022 annual conference, to the OECD Working Party on Industry Analysis, and to the second international AI in Work, Innovation, Productivity and Skills (AI-WIPS) conference. This work contributes to the OECD programme on AI-WIPS, supported by the German Federal Ministry of Labour and Social Affairs (BMAS).

Cody Morris contributed to this work while being an intern at the OECD Directorate for Science, Technology and Innovation.

Mariagrazia Squicciarini contributed to this work while serving as a Senior Economist and Head of Unit at the OECD Directorate for Science Technology and Innovation. She is now UNESCO's Chief of the Executive Office of the Social Human Sciences Sector.

Abstract

This work employs a novel approach to identify and characterise firms adopting Artificial Intelligence (AI), using different sources of large microdata. Focusing on the United Kingdom, the analysis combines data on Intellectual Property Rights, website information, online job postings, and firm-level financials for the first time. It shows that a significant share of AI adopters is active in Information and Communication Technologies and professional services, and is located in the South of the United Kingdom, particularly around London. Adopters tend to be highly productive and larger than other firms, while young adopters tend to hire AI workers more intensively. Human capital appears to play an important role, not only for AI adoption but also for firms' productivity returns. Significant differences in the characteristics of AI adopters emerge when distinguishing between firms carrying out AI innovation, those with an AI core business, and those searching for AI talent.

Keywords: Artificial Intelligence, Technology adoption, Productivity.

JEL codes: C81, J24, O33, O34

Résumé

Ce travail propose une nouvelle approche pour identifier et caractériser les entreprises adoptant l'intelligence artificielle (IA), en se fondant sur plusieurs sources conséquentes de microdonnées. L'analyse, centrée sur le Royaume-Uni, combine pour la première fois des données sur les droits de propriété intellectuelle, des informations extraites de sites internet, des offres d'emploi en ligne et des données financières d'entreprises. Elle montre qu'une part significative des firmes adoptant l'IA est active dans les technologies de l'information et de la communication et les services professionnels, et se situe au sud du Royaume-Uni, plus particulièrement autour de Londres. Les adoptants ont tendance à être très productifs et de plus grande taille que les autres entreprises, tandis que les jeunes adoptants semblent embaucher davantage de travailleurs de l'IA. Le capital humain parait jouer un rôle important, non seulement pour l'adoption de l'IA mais aussi pour le rendement de la productivité des entreprises. Des différences significatives dans les caractéristiques des adoptants de l'IA émergent lorsqu'on compare les entreprises innovantes dans l’IA, celles qui ont une activité principale liée à l'IA et celles qui recherchent des talents dans ce domaine.

Kurzfassung

In dieser Studie wird ein neuer Ansatz verwendet, um Unternehmen, die künstliche Intelligenz (KI) einsetzen, zu identifizieren und zu charakterisieren, indem verschiedene Quellen großer Mikrodaten genutzt werden. Die Analyse konzentriert sich auf das Vereinigte Königreich und kombiniert erstmals Daten zu geistigen Eigentumsrechten, Website-Informationen, Online-Stellenausschreibungen und Finanzdaten von Firmen. Sie zeigt, dass ein erheblicher Anteil der KI-Anwender in Informations- und Kommunikationstechnologien und professionellen Dienstleistungen tätig und im Süden des Vereinigten Königreichs, insbesondere um London herum, angesiedelt ist. Firmen, die KI einsetzen, sind in der Regel hochproduktiv und größer als andere Unternehmen, während junge KI-Anwender dazu neigen, verstärkt KI-Mitarbeiter einzustellen. Das Humankapital scheint eine wichtige Rolle zu spielen, nicht nur für die Einführung von KI, sondern auch für die Produktivitätsgewinne der Unternehmen. Signifikante Unterschiede in den Merkmalen der KI-Anwender ergeben sich, wenn man zwischen Unternehmen unterscheidet, die KI-Innovationen durchführen, die ein KI-Kerngeschäft haben und die nach KI-Talenten suchen.

Executive Summary

Artificial Intelligence (AI) is a fast-growing technology with the potential to drive growth, to transform the economic landscape and industries and to improve people's lives. However, evidence about its diffusion across firms is still relatively scant, especially beyond the United States. This work proposes a novel methodology based on big data allowing to identify AI adopters and analyses their characteristics, focusing here on data for the United Kingdom in 2019.

The approach combines different sources of large commercial and administrative microdata, including Intellectual Property Rights, website information, online job postings, and firm-level financials. Identification of AI adopters is based on a common set of AI-related keywords validated by the OECD. The data allow to focus on different margins of AI adoption, distinguishing between firms that carry out AI innovation; companies that have an AI core business; and other firms that require AI talent, and highlighting the importance of using complementary sources of information.

The analysis uncovers several stylised facts. AI adoption appears to some extent polarised, both in terms of industry activity – with most firms operating in Information and Communication Technology and professional services – and geographical presence – with AI adopters concentrated in the South of the United Kingdom. Hiring of AI talent appears however more widespread across sectors, providing evidence in favour of the general-purpose nature of AI technologies.

AI adopters tend to be leaders in terms of labour productivity in their industry. This may hint at possible implications for existing divides across groups of firms – such as divergences between leaders and laggards – as well as across regions.

A coexistence of young-small and old-large AI adopters also emerges. Among different types of AI adopters, firms that have AI at the core of their business are those with the largest share of young-small adopters.

Exploring the intensive margin of AI adoption – using a proxy of AI-hiring intensity – further highlights the role of AI professionals, for which a strong hiring demand is evident, and of young firms, which tend to have on average higher AI-hiring intensity.

When comparing AI adopters with other likely non-adopting firms, significant scale advantages emerge, with AI adopters being generally larger than other firms. AI adopters also tend to be more productive, although these productivity premia – especially evident in market services – do not necessarily imply a positive effect of AI on productivity. These can be rather related to self-selection of more productive firms into AI adoption.

Human capital appears to play an important role for AI adoption. AI-related departments in universities are found to increase the likelihood of using AI. Furthermore, among different occupational groups, managers – as well as professionals – likely help translate AI use into higher efficiency.

Synthèse

L'intelligence artificielle (IA) est une technologie en plein essor avec le potentiel de stimuler la croissance, de transformer le paysage économique et les industries et d'améliorer la vie des gens. Cependant, les indications de sa diffusion au sein des entreprises sont encore relativement rares, en particulier au delà des États-Unis. Ce travail propose une méthodologie inédite, fondée sur les données de gros volume, visant à identifier et caractériser les adoptants de l'IA, en se focalisant ici sur des données de 2019 pour le Royaume-Uni.

Cette approche combine plusieurs sources conséquentes de microdonnées commerciales et administratives, notamment sur les droits de propriété intellectuelle, les informations extraites de sites Internet, les offres d'emploi en ligne et les données financières d'entreprises. L'identification des adoptants de l'IA repose sur un socle de mots-clés liés à l'IA validés par l'OCDE. Grâce à la complémentarité des sources d'information, les données permettent de découper les phase d'adoption de l'IA, en distinguant les entreprises innovantes en IA, celles dont l'activité principale est liée à l'IA et les autres entreprises qui recherchent des spécialistes en matière d'IA.

Parmi les faits stylisés révélés par l'analyse, l'adoption de l'IA apparaît comme étant relativement polarisée, que ce soit en termes d'activité industrielle - la plupart des entreprises opérant dans le secteur des TIC et les services professionnels – comme en termes de présence géographique - les adoptants de l'IA étant plutôt concentrés dans le sud du Royaume-Uni. L'embauche de spécialistes en IA semble toutefois répartie dans tous les secteurs économiques, signe de la polyvalence des technologies liées à l'IA.

S'agissant de la productivité du travail, les entreprises adoptant l'IA ont tendance à être plus performantes dans leur secteur, ce qui pourrait avoir un impact sur les clivages entre groupes d'entreprises –entre leaders et retardataires – mais aussi entre régions.

L'analyse révèle également la coexistence de jeunes-petites et de vieilles-grandes entreprises parmi celles qui adoptent l'IA. Les entreprises avec l'IA comme activité principale comptent le plus grand nombre de jeunes-petits adoptants.

La marge intensive de l'adoption de l'IA, évaluée par une approximation de l'intensité d'embauche en IA, souligne le rôle des professionnels de l'IA, pour lesquels la forte demande d'embauche est évidente, et celui des jeunes entreprises, ayant une intensité moyenne d'embauche liée à l'IA plus élevée.

En compare les entreprises ayant adopté l'IA à d'autres entreprises susceptibles de ne pas l'avoir adoptée, on constate des effets d'échelle significatifs, les entreprises ayant adopté l'IA étant généralement plus grandes que les autres. Les adoptants de l'IA semblent également plus productifs, bien que ces primes de productivité - particulièrement prégnantes dans les services marchands – n'induisent pas nécessairement d'effet positif de l'IA sur la productivité.

Cela peut être lié au fait que les entreprises les plus productives auront davantage tendance à adopter l'IA.

Le capital humain paraît jouer un rôle important dans l'adoption de l'IA. On constate que la présence de départements dédiés à l'IA dans les universités augmente la propension à utiliser l'IA. En outre, entre les différents groupes professionnels, les cadres - comme les spécialistes - contribuent probablement à convertir l'utilisation de l'IA en efficacité accrue.

Zusammenfassung

Künstliche Intelligenz (KI) ist eine schnell wachsende Technologie, die das Potenzial hat, das Wachstum voranzutreiben, die Wirtschaftslandschaft und die Branchen zu verändern und das Leben der Menschen zu verbessern. Allerdings gibt es immer noch relativ wenige Belege für ihre Verbreitung in Unternehmen, insbesondere außerhalb der Vereinigten Staaten. In dieser Arbeit wird eine neuartige, auf großen Datenmengen basierende Methodik vorgeschlagen, die es ermöglicht, KI-Anwender zu identifizieren und ihre Merkmale zu analysieren, wobei der Schwerpunkt auf den Daten für das Vereinigte Königreich im Jahr 2019 liegt.

Der Ansatz kombiniert verschiedene Quellen großer kommerzieller und administrativer Mikrodaten, darunter Daten zu geistigen Eigentumsrechten, Website-Informationen, Online-Stellenausschreibungen und Finanzdaten von Firmen. Die Identifizierung von KI-Anwendern basiert auf KI-bezogene Schlüsselwörter, die von der OECD validiert wurden. Die Daten ermöglichen es, sich auf verschiedene Bereiche der KI-Einführung zu konzentrieren, wobei zwischen Unternehmen, die "KI-Innovationen" durchführen, Unternehmen, die ein "KI-Kerngeschäft" haben, und anderen Unternehmen, die "KI-Talente" benötigen, unterschieden wird, und zeigen, wie wichtig es ist, ergänzende Informationsquellen zu nutzen.

Die Analyse deckt mehrere stilisierte Fakten auf. Die Einführung von KI scheint zu einem gewissen Grad polarisiert zu sein, sowohl in Bezug auf die Branchenaktivität - die meisten Unternehmen sind in Informations- und Kommunikationstechnologien und professionellen Dienstleistungen tätig - als auch in Bezug auf die geografische Präsenz - die KI-Anwender findet man besonders konzentriert im Süden des Vereinigten Königreichs. Die Einstellung von KI-Talenten scheint jedoch branchenübergreifend verbreitet zu sein, was für die allgemeine Verwendbarkeit von KI-Technologien spricht.

KI-Anwender sind in der Regel führend in Bezug auf die Produktivität in ihrer Branche. Dies könnte auf mögliche Auswirkungen auf bestehende Unterschiede zwischen verschiedenen Unternehmensgruppen - wie etwa zwischen Spitzenreitern und Nachzüglern - sowie zwischen verschiedenen Regionen hindeuten.

Es zeigt sich auch eine Koexistenz von jungen, kleinen und alten, großen KI-Anwendern. Unter den verschiedenen Arten von KI-Anwendern sind die Unternehmen, bei denen KI den Kern ihres Geschäfts ausmacht, diejenigen mit dem größten Anteil an jungen und kleinen Anwendern.

Bei der Untersuchung der intensiven Marge der KI-Einführung - unter Verwendung eines Proxys für die KI-Einstellungsintensität - wird die Rolle von KI-Fachkräften, für die offensichtlich eine starke Einstellungsnachfrage besteht, und von jungen Unternehmen, die im Durchschnitt eine höhere KI-Einstellungsintensität aufweisen, noch deutlicher.

Beim Vergleich von KI-Firmen mit anderen Unternehmen, die wahrscheinlich keine KI einführen, ergeben sich erhebliche Größenvorteile, da KI-Anwender im Allgemeinen größer sind als andere Unternehmen. KI-Anwender sind auch tendenziell produktiver, obwohl diese Produktivitätsvorteile - insbesondere bei Marktdienstleistungen - nicht unbedingt eine positive Auswirkung der KI auf die Produktivität bedeuten. Sie können vielmehr auf die Selbstselektion produktiverer Unternehmen bei der Einführung von KI zurückzuführen sein.

Das Humankapital scheint eine wichtige Rolle bei der Einführung von KI zu spielen. Es hat sich gezeigt, dass KI-bezogene Abteilungen an Universitäten die Wahrscheinlichkeit des Einsatzes von KI später erhöhen. Darüber hinaus tragen unter den verschiedenen Berufsgruppen Führungskräfte - wie auch Fachkräfte - wahrscheinlich dazu bei, den Einsatz von KI in eine höhere Effizienz umzusetzen.

1 Introduction

Artificial Intelligence (AI) is rapidly reshaping economies and societies (OECD, 2019[1]). It has the potential to drive growth, to transform the economic landscape and industries, and to improve people's lives.

Relevantly, AI is often considered a general-purpose technology (GPT). As such, given improvements over time and potential breadth and depth of its penetration, AI may significantly affect productivity and growth over the next decades. Furthermore, it can importantly affect other economic outcomes. It is already changing the demand for jobs and skills, and is going to possibly help tackle societal challenges such as climate change. During the recent COVID-19 pandemic, AI systems have been crucial to monitor the evolution of the outbreak and guide policy responses.

Identifying which firms are adopting AI and studying their characteristics becomes therefore particularly relevant from an economic policy perspective. This may shed light on the factors enabling or preventing this recent stage of digital transformation, help characterise AI diffusion patterns, and better understand the implications of such diffusion for economic outcomes.

However, although very recently growing beyond the United States, studies analysing the diffusion of AI across firms are still in their infancy. This is generally due to the existence of significant challenges in data availability and measurement.

In this context, this work proposes and employs a novel approach to identify and characterise AI adopters, combining different sources of large commercial and administrative microdata.

The approach is applied to data from the United Kingdom, a country for which, at our knowledge, there is still limited empirical evidence[1] despite the importance of AI in the UK policy debate, highlighted for instance by the 2021 UK National AI Strategy.[2]

First, this work identifies AI adopters based on a common set of AI-related keywords, developed and validated by previous OECD work (Baruffaldi et al., 2020[2]), relying on and combining three different sources of data: i) AI-related Intellectual Property Rights (IPRs), which capture firms innovating in AI or embedding AI in their goods or services; ii) AI-related activities as stated on company websites; and iii) the demand for AI-related skills contained in online job postings by employers.[3]

Second, data on AI adopters are matched with firm-level financial statements that contain information on their characteristics, such as firm size, age, detailed sector of activity or location of the headquarters, and that allow computing proxies of productivity.

Combining these data sources provides a unique dataset with complementary information on AI adoption, allowing to focus on its different margins. In particular, it allows distinguishing three subgroups of AI adopters:[4] those that carry out AI innovation; non-innovators that mention AI-related keywords on their website, which we suppose are companies that have an AI core business; and other firms that require AI talent. Based on these data, the descriptive analysis provides interesting stylised facts about the characteristics of AI adopters, based on the most recent information.

A polarisation of AI adoption emerges, both in terms of industry activity – with most firms operating in Information and Communication Technology (ICT) and professional services – and geographical presence – with AI adopters concentrated in the South of the United Kingdom. The latter is not surprising given UK's geographic concentration of industrial activity. Hiring of AI talent appears however more widespread across sectors, providing evidence in favour of the general-purpose nature of AI technologies.

Furthermore, AI adopters tend to be leaders in their industry (in terms of labour productivity), with possible implications for existing divides across groups of firms (leaders vs. laggards) and regions. A coexistence of young-small and old-large AI adopters also emerges, and among firms that tend to have AI at the core of their business, most of those identified by the current analysis are young and small.

Exploring the intensive margin of AI adoption using a proxy of AI-hiring intensity further highlights the role of AI professionals, for which a strong demand is evident, and of young firms, which tend to have higher AI-hiring intensity.

When comparing AI adopters with other likely non-adopting firms, significant scale advantages emerge, with AI adopters generally being larger than other firms. AI adopters also tend to be more productive than other firms, although these productivity premia – especially evident in market services – do not necessarily imply a positive effect of AI on productivity, i.e. cannot be interpreted causally, but can also be related to self-selection of more productive firms into AI use.

Human capital appears to play an important role for AI adoption. Proximity to AI-related universities, i.e., universities that have AI-related keywords present on their website, seem to increase the likelihood of using AI. Among different occupational groups, managers as well as professionals appear to play an important role, likely helping translate AI use into higher efficiency.

Although largely exploratory, to our knowledge this is the most ambitious effort combining different commercial and administrative data sources to study AI adoption in firms, overcoming the limitations imposed by single data sources. The approach adopted builds upon the long-standing experience of the OECD in string and name matching and can be scaled to other countries beyond the United Kingdom.

The analysis builds upon and extends ongoing and previous OECD work analysing AI developments[5] and complements analysis based on – often confidential – ICT surveys (in particular, the ongoing cross-country OECD *AI diffuse* project[6]).

2 Existing evidence on AI adoption by firms

AI is a rapidly growing technology, with strong potential to affect productivity and other economic outcomes,[7] but it is conceptually different from other types of innovation. Indeed, given the depth and breadth of its possible penetration, it is often considered a GPT (see for instance discussions in Trajtenberg (2018[3]); Crafts (2021[4]); Brynjolfsson et al. (2021[5])). GPTs are characterised by pervasiveness, improvements over time, innovation spawning and spillover effects (Jovanovic and Rousseau, 2005[6]), but require complementary investments in intangible assets that may take time to materialise before translating technology use into productivity gains (Brynjolfsson, Rock and Syverson, 2021[5]).

However, despite the strong potential of AI and its relevance in economic and policy discussions, there is still scarce evidence about its diffusion across firms, and on the characteristics of firms using AI. This is mainly due to the paucity of AI firm-level data.

One recent strand of the literature has been analysing firm-level data that contain relevant information on AI use, in particular official national ICT surveys and custom data. Official national ICT surveys have recently started to integrate questions asked to firms concerning their use of AI (see for example Zolas et al. (2020[7]) focusing on the United States; Rammer et al. (2022[8])on Germany; Cho et al. (2022[9]) on Korea; Calvino and Fontanelli (2022[10]) on France).[8] Ongoing OECD work complementary to this paper is also exploiting microdata from ICT surveys in a cross-country perspective based on harmonised statistical code (*AI diffuse* project, see Calvino and Fontanelli (2022[11])). This literature has suggested that the use of AI technologies is still limited and heterogeneous across sectors and groups of firms. AI use appears prevalent among larger firms, in line with the existence of fixed costs or with the importance of complementary assets and economies of scale.

Custom data also contain information about selected categories of imported capital goods, and may allow to identify episodes of AI investments. In this context, Domini et al. (2022[12]) focus on the role of AI for labour demand and highlight that adoption events are not related to increases in wage inequality or gender wage gaps in France.

A different part of the literature has instead focused on analysing the process of AI diffusion based on keywords able to identify AI developments. In this context, Baruffaldi et al. (2020[2]) created a detailed three-pronged keyword-based approach to identify AI developments in science, algorithms and technologies. The list of keywords used to search for AI-related patents and scientific papers includes terms related to subject areas, such as automated recognition of patterns, neural networks, robotics, and autonomous vehicles (e.g. image or speech recognition, convolutional neural network, humanoid robot or unmanned aerial vehicle).[9] Their approach relies on established bibliometric and patent-based methods, complemented by machine learning (ML) implemented on open source software data, in order to produce an encompassing operational definition of AI.[10]

A related stream of work has focused more closely on analysing texts (or technology classes) of patent documents to identify AI-related inventions and thus AI-related firms. In fact, patents play an important role in the protection of innovations and they are increasingly being utilised to protect AI-related technological developments.[11] Using the United States Patent and Trademark Office (USPTO) patent data, Alderucci et al. (2020[13]) find a statistically significant and positive relationship between a firm patenting an AI innovation and higher employment, better productivity (higher value-added per employee) and fewer production workers. They also find that the strength of the relationship grows over the years following the initial innovation, which could be related to potential learning effects, adjustment costs, and the role of complementary investments.

Relatedly, combining USPTO patent applications published between 2002-19 with company accounts, Santarelli et al. (2022[14]) map the knowledge base centred on robotics and AI. Their analysis shows that such knowledge base is linked to the previous technological paradigm (see also Igna and Venturini (2022[15])), and that it is highly pervasive, highlighting that robotics and AI are strictly related. Damioli et al. (2021[16]) focus instead on 5,257 companies worldwide that filed at least one AI-related patent between 2000-16 and find AI patent applications to have a positive impact on firm labour productivity, especially on smaller firms and services industries.

Beyond patents, other IPRs may contain useful information to analyse AI activity. In particular, trademarks can shed light on the extent to which (new) companies and products appearing on the market rely on, exploit or propose AI-related goods and services. In this context, Nakazato and Squicciarini (2021[17]) exploit trademark applications filed at the European Union Intellectual Property Office (EUIPO), the Japan Patent Office (JPO) and the USPTO over the period 2009-18 and show that AI-related goods and services have expanded in consumer markets in recent years, especially protecting computer-related products or services, such as software or audio-visual devices. Combining AI-related patents and trademarks, Dernis et al. (2021[18]) highlight that developments of AI technologies, goods and services is mainly due to start-ups or large incumbents, located in selected countries, with a majority of actors operating in ICT-related sectors.

A different source of information to identify AI-active firms consist in their online presence. Some firms indeed highlight their AI activities by mentioning them on their websites. In this context, Dernis et al. (forthcoming[19]) focus on UK firms that have stated on their websites to be active in the AI space. These firms appear relatively small (less than 10 employees) and young (less than 5 years old). This could indicate that companies specialising in AI, are more likely to be smaller and younger start-ups compared to the older established firms that have AI-related patents.

A final source to identify AI activity by firms that has been more extensively exploited by recent work is related to their demand for AI-related skills. Indeed, in order to develop and adopt AI technologies firms need specialised human capital. Hence, looking at labour featuring AI-related skills may help shed light on the use of AI technologies (Tambe, 2013[20]). Using tools like ML algorithms and AI-related keywords, such as the ones found in Baruffaldi et al. (2020[2]), a number of papers including Alekseeva et al. (2020[21]; 2021[22]), Babina et al. (2020[23]), Squicciarini and Nachtigall (2021[24]) and Samek et al. (2021[25]) identify AI-related skills and jobs in online job postings data from Lightcast™ (formerly known as Burning Glass Technologies). Part of this analysis focuses on job postings rather than firms and highlights that, as AI permeates the economy, it moves into new sectors and AI skills are demanded in relation to a wider range of occupations, firms and industries.

More closely to this work, AI-related job postings have also been used to focus more directly on AI-active firms. In this context, Alekseeva et al. (2021[22]) use Lightcast™ data for the United States to determine the effects on a firm from adopting AI.[12] They document a positive association between AI adoption and a firm's growth in sales, capital expenditures and investments, driven by AI skills among managers. Using both Lightcast™ job posting and resume data from Cognism Inc., also Babina et al. (2020[23]) find that US firms investing more in AI (i.e., hiring AI-skilled labour) were more likely to see an increase in sales, employment, and market share. Similarly to Alekseeva et al. (2021[22]), they do not find evidence that investing in AI makes firms more productive, consistently with the need of complementary investments that may take time before fully materialising. Conversely, Bäck et al. (2022[26]) find a positive association between AI and productivity, observable for larger firms only and after a sufficient time lag, when combining online job postings data (from commercial job advertisement platform Oikotie Oy's) and firm-financials (from Orbis©) in Finland.

This work goes beyond the existing literature by exploiting more comprehensively the potential of methods to identify AI adopters based on keywords, taking a firm-level perspective and relevantly combining different sources of information that allow distinguishing different margins of AI adoption. Beyond the work of Baruffaldi et al. (2020[2]), which however focuses on AI developments in science, algorithms and technologies, and to some extent of Bloom et al. (2021[27]), which analyses the diffusion patterns of disruptive technologies in the United States, to our knowledge no contribution has combined information from several different sources to this extent.

Combining data on AI activity at the firm level with information on firm financials in order to study AI diffusion and the characteristics of AI adopters appears particularly promising also considering that official surveys contain only a limited set of questions on AI use, and those tend to be available only in selected years. In addition, surveys and custom data are often confidential. Furthermore, although informative, using IP-based methods alone may limit the identification of AI activity only to a subset of AI firms, which generally represent the most innovative ones.[13] Similarly, using only information on skill demand or web scraping may allow capturing only some aspects of AI diffusion, more related to new hires or web presence, while combining more data sources together would offer a unique and more comprehensive outlook about AI use by firms.

The proposed methodology, together with information about the different data sources, and the details concerning the matching procedure are discussed in the next section.

3 Methodology and data

This section presents a scalable multifaceted approach to identify different types of AI adopters. First, the rationale and detailed data sources are discussed. Second, the process of how the information from each data source is linked to characterise those actors is explained.

This work builds upon and extends the work by Dernis et al. (2021[18]), Samek et al. (2021[25]), Squicciarini and Nachtigall (2021[24]), Nakazato and Squicciarini (2021[17]) and Dernis et al. (forthcoming[19]) that have also used some of the sources of information but never combined all of them together.

Identifying and linking AI adopters from different data sources

This work identifies UK AI adopters based on a common set of AI-related keywords, developed and validated by previous OECD work, which include, for instance, machine learning, natural language processing or neural network (Baruffaldi et al., 2020[2]).[14] The approach relies on and combines three different sources of information: i) AI-related IPRs, which capture firms innovating in AI or embedding AI in their goods or services; ii) the demand for AI-related skills contained in online job postings by employers; and iii) AI-related activities as stated on company websites. It then links those data to company accounts in order to characterise the most recent patterns of AI adoption.

The following paragraphs provide a brief overview of the different data sources used to identify and characterise AI adopters, with further details available in Annex A.

First, AI-related IPRs provide key information to identify firms carrying out AI innovation. Given the very nature of patents, firms applying for AI-related patents can be considered as AI innovators since they aim at protecting AI-related technological advancements. Patents however do not capture all AI-related innovative activity, and patentability of AI-related technologies varies across countries.[15] Using trademark data relevantly complements the picture about AI-related innovators. Registering an AI-related trademark suggests that firms are relying on branding strategies to signal to their customers that they produce or sell goods and services embedding AI. A number of studies performed over the last decade show that trademarks are related to different types of innovations (to process innovation in particular), and are generally used to signal firms' innovativeness to current and prospective customers (see for instance Arora, Bei and Cohen (2016[28]); Castaldi, Block and Flikkema (2019[29])).

Data on AI-related IPRs are sourced from the Intellectual Property (IP) database of the OECD STI Microdata lab, using the same methodology and building upon the work of Dernis et al. (2021[18]), where additional information is also available. Data refer to AI-related patent applications (IP5 patent families) or trademarks registered at the EUIPO, JPO and the USPTO between the early 2000s and 2018. AI-related patents are identified using the AI-related keywords identified by Baruffaldi et al. (2020[2]) as well as the International Patent Classification (IPC) and Cooperative Patent Classification (CPC) classes, while AI-related trademarks are identified following Nakazato and Squicciarini (2021[17]), who build on Baruffaldi et al. (2020[2]) as well as other IP-based analyses.

Company websites are the second key source of information to identify firms' AI activity. In fact, AI-related keywords mentioned on company websites are relevant indicators of firm activity, and can suggest that AI may be central to the company business model. This work therefore exploits web-reading data provided by GlassAI - a UK software development company - to identify AI on websites, using the same approach used by Dernis et al. (forthcoming[19]). In particular, GlassAI was provided with the keywords described in Baruffaldi et al. (2020[2]) to identify AI activity. The web-reading exercise has been conducted in 2020. Where available, the information retrieved by GlassAI goes beyond entities' name and website and also covers location, company registration number, and sector.

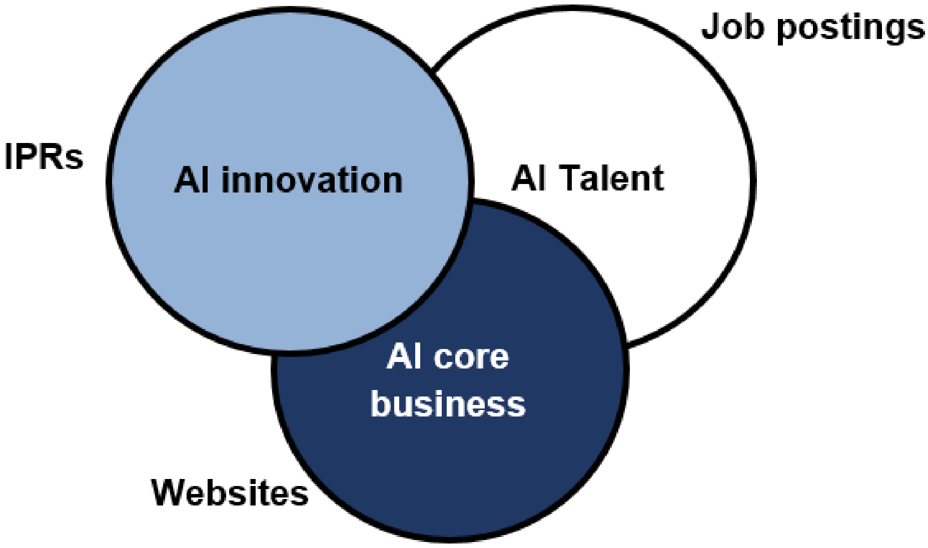
Third, AI-related online job postings are able to capture the extent to which firms are demanding AI talent (skills), responding to the changes in labour and skills demand brought by the diffusion of AI.[16] In order to identify the demand for AI-related skills, this work uses data provided by Lightcast™. Lightcast™ is a well-known data provider collecting online vacancy data through web-scraping. This has been widely used in the literature, as also highlighted in the previous section. Lightcast™ relies on over 40 000 distinct job boards and company websites and is active in many English-speaking and European countries, notably the United Kingdom. It claims to cover the near-universe of all online job postings and provides detailed information on labour and skill – including AI-skill – demand posted online. This information is then cleaned, structured and standardised. AI-related job postings have been identified according to the methodology described in Squicciarini and Nachtigall (2021[24]), which builds upon the work of Baruffaldi et al. (2020[2]), and then aggregated at the employer level to identify firms hiring AI-related human capital. Importantly, the data also contain information on location and international occupation classifications, which are used in the analysis. Data used in this work span over the period 2012-20.

Based on the information above, it is possible to identify a comprehensive set of AI adopters. In particular, a firm is tagged as an AI adopter if any of the following conditions applies: i) having applied for or registered an AI-related patent or trademark during the observed period; ii) having at least one AI-related job posting during the observed period; iii) mentioning AI-related keywords on their website at the time of web reading.[17]

Combining different data can not only broaden the scope of previous analyses that have considered one source of information about AI at the time, but importantly provides additional insights about different margins of AI adoption.

In particular, based on the abovementioned sources, three (mutually-exclusive) groups of AI adopters can be further identified: i) those that carry out AI innovation (i.e., those that apply for AI-related patents or register AI-related trademarks); ii) non-innovators that mention AI-related keywords in their website, which supposedly are companies that have an AI core business; and iii) other firms that demand AI talent, i.e. the residual group of firms posting AI-related jobs online (see also Figure 3.1 below for a graphical representation).[18]

Figure 3.1. Different groups of AI adopters



Notes: Data sources (IPRs, Job postings, and Websites) are reported outside, whereas group names of AI adopters (AI innovation, AI core business, and AI talent) identified in the respective data source are reported inside the Venn diagram.
Sources: Authors' elaboration.

In order to characterise AI adopters, the data are further combined with information on firms' financial accounts (see Annex A for further details). This is sourced from Bureau van Dijk's (BvD) Orbis©, a widely used commercial database that aggregates company accounts and financial reports across countries. It covers almost 400 million companies and entities worldwide of which around 41 million contain detailed financial information. Orbis© contains relevant information on firm characteristics, such as size, age, detailed sector of activity, and location of headquarters, and allows computing proxies of firm-level productivity. The same data source has been indeed used by several OECD and academic papers focusing on firm dynamics and productivity (e.g., Andrews et al. (2016[30]); see also Bajgar et al. (2020[31]) for a critical discussion about this data[19]). The latest available year with comprehensive coverage at the time of writing is 2019, which is the year for which the characterisation of AI adopters is carried out.

The current analysis, exploratory in nature, focuses on the United Kingdom, a country for which all sources of relevant data are available and generally of good quality and for which, to our knowledge, limited evidence on AI adopters and their characteristics exists, despite the importance of AI in the UK policy debate. In the future, the analysis could be further scaled to other countries for which relevant data are available or may become available. In fact, AI-related IPRs tend to be available across several countries; with patenting rates highest in the United States followed by Japan, Korea and the People's Republic of China (Dernis et al., 2021[18]). Information from company websites can, in principle, be retrieved beyond the United Kingdom (see, for instance, Kinne and Axenbeck's (2019[32]) study on innovation in firms from company websites) whereas information on AI talent from job postings is already available for a number of Anglo-Saxon and European countries (Squicciarini and Nachtigall, 2021[24]; Samek, Squicciarini and Cammeraat, 2021[25]; Pouliakas, 2021[33]; Acemoglu et al., 2022[34]). Orbis covers a wide range of countries, although the representativeness of the data differs significantly across them (see Bajgar et al. (2020[31]) for further discussion).

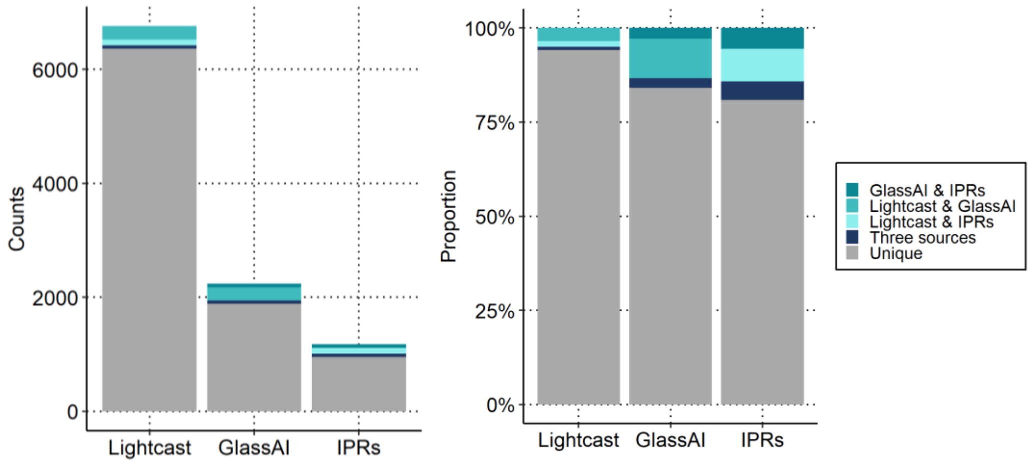
Matching the data in practice

After presenting the different data sources used in this work, this subsection briefly summarises the actual matching of those data as well as its quality. A detailed summary of the data cleaning and matching process as well as a description of the remaining challenges associated with the implemented approach are provided in Annex A.

Matching all data sources together is far from being a trivial exercise and involves significant conceptual, data and computational challenges. The matching exercise has built upon the long-standing experience of the OECD in string and name matching, also using software developed ad-hoc. This has been an iterative exercise that has involved different methods and software.

First, organisation names are harmonised and source data cleaned, within each different source. Then the harmonised names are compared using a series of string-matching algorithms – mainly token-based and string-metric-based, such as token frequency matching and Levenshtein (1966[35]) as well as Jaro-Winkler (1999[36]) distances.[20] Whenever possible and to address multiple matches, websites, company registration numbers, postcodes and/or sectoral information have also been exploited (in the case of GlassAI, registration numbers are indeed the primary matching criterion) in addition to names. The outcome of the matching is represented visually in Figure 3.2 and Figure 3.3.

Figure 3.2. AI adopters across different data sources



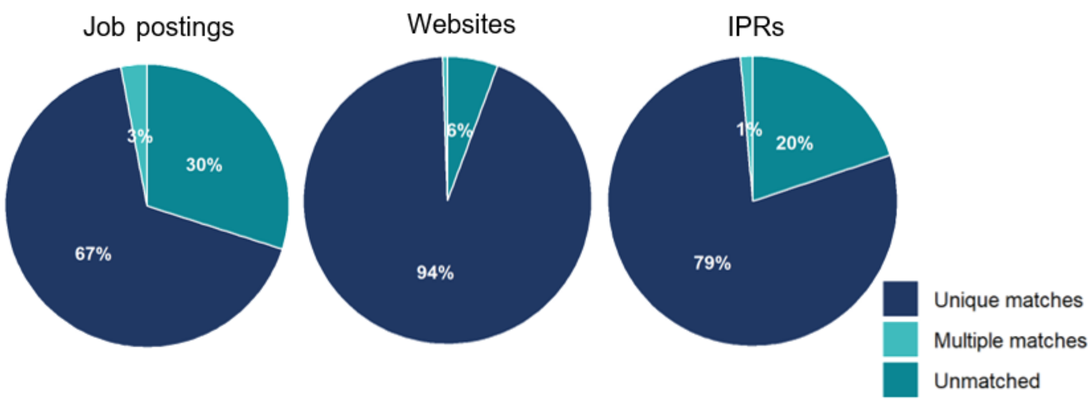
Notes: The sum reflects the number of firms that have been identified as AI adopters after the harmonisation and cleaning exercises (6,761 firms in Lightcast™, 2,245 in GlassAI and 1,178 in IPRs).
Sources: Authors' own compilation based on GlassAI (January 2021), Lightcast™ (February 2021), IPRs (STI Micro-data Lab: IntellectualPropertyDatabase, http://oe.cd/ipstats, July 2020) and Orbis© (version 2.2020, Bureau van Dijk, April 2021).

Figure 3.2 focuses on the composition and overlapping among the data sources identifying AI adopters, both in absolute numbers and percentages. It is evident from the figure that the different data sources on AI adoption turn out to be highly complementary, being likely related to different margins of AI adoption. This highlights the importance of using more than one source of information when shedding light on a phenomenon, like AI diffusion, that is dynamic and multi-faceted in nature.[21]

Figure 3.3 shows that matching rates with Orbis© are overall more than satisfactory, with always more than two thirds of firm names uniquely matched to Orbis©. Matching rates are close to 95% for GlassAI, also thanks to the availability of company registration numbers. In IPRs and online job postings, lower matching rates can also reflect the fact that – despite the cleaning efforts – not all private individuals (IPRs) and universities (job postings and IPRs) were removed before matching, mechanically increasing the shares of unmatched cases. Shares of multiple matches are very low, reflecting the extensive and iterative efforts made to try to solve those cases as much as possible.

Before getting to the final database that will be used for analysis, additional cleaning steps are implemented. Multiple matches are excluded from the AI adopters group, taking a conservative approach. As will be clear in the next sections, the analysis is however robust to including those among AI adopters. Finally, in line with existing studies based on Orbis© (e.g. Andrews et al. (2016[30])), a conventional sectoral filtering is implemented, focusing on the non-farm business economy (i.e., NACE Rev.2 division levels 5 to 82) in the rest of the analysis.

Figure 3.3. Matching rates with Orbis by source of information on AI activity



Notes: The sum reflects the number of firms that have been identified as AI adopters after the harmonisation and cleaning exercises (6,761 firms in Lightcast™, 2,245 in GlassAI and 1,178 in IPRs).
Sources: Authors' own compilation based on GlassAI (January 2021), Lightcast™ (February 2021), IPRs (STI Micro-data Lab: IntellectualPropertyDatabase, http://oe.cd/ipstats, July 2020) and Orbis© (version 2.2020, Bureau van Dijk, April 2021).

As already evident from the previous discussion, while the multi-perspective approach proposed provides comprehensive insights into the identification of AI actors and their characteristics, some challenges still remain, which mainly stem from data availability and from the characteristics of the databases we rely upon (see also Annex A).

In particular, it is clear that – despite all the efforts – it is not possible to claim that this methodology allows to identify each and every single AI adopter in the United Kingdom. In fact, not every single AI adopter may have applied for AI-related IPRs, posted AI-related jobs online, or mentioned AI-related keywords on its website. The set of AI adopters identified is however considerable and the proposed methodology allows to significantly extend the scope of analysis beyond existing work based on fewer data sources.

4 Exploring the characteristics of AI adopters

Having combined data from AI-related IPRs, online job postings, and information on company websites with firm-level financial accounts, this section now focuses on analysing the characteristics of AI adopters. The analysis focuses on AI adopters for which information on their characteristics is available in 2019, the latest available year with comprehensive coverage in Orbis©.[22] This is aimed at providing the most recent available picture about their characteristics.[23]

Some basic descriptive statistics are reported in Table 4.1 below. Those provide a starting point for the characterisation and already highlight significant heterogeneity in age and size across the different groups of AI adopters.[24] More specifically, the table already shows that firms with AI as their core business appear on average much smaller as well as younger, and have lower turnover than firms innovating in AI or searching for AI talent. Albeit substantial differences in terms of size, the latter two appear more similar in terms of age and turnover.

Table 4.1. Summary statistics about AI adopters

United Kingdom, 2019

	AI adopters (all)	AI innovation	AI core business	AI talent
Employment size – mean (std. dev.)	1,417 (11,514)	3,926 (27,205)	127 (2,026)	1,572 (8,031)
Firm age – mean (std. dev.)	14.8 (19.3)	19.0 (24.7)	7.6 (8.1)	18.8 (21.8)
Turnover [in GBP million] – mean (std. dev.)	1,498 (12,264)	2,027 (8,059)	69 (479)	1,558 (13,650)
N	2,796	335	1,007	1,434

Note: Firm age is calculated as the difference between 2019 and the incorporation date, employment refers to the number of employees and turnover to the operating revenues as reported in Orbis© for the year 2019. Implausible turnover or employee counts are dropped. N refers to the highest count of non-missing observations considering employment, firm age and turnover. As discussed in Annex A, shares of missing values are significantly higher for the turnover variable.
Source: Authors' own compilation based on GlassAI (January 2021), Lightcast™ (February 2021), IPRs (STI Micro-data Lab: IntellectualPropertyDatabase, http://oe.cd/ipstats, July 2020) and Orbis© (version 2.2020, Bureau van Dijk, April 2021).

The remainder of the section analyses the characteristics of AI adopters, focusing on their sector of activity, geographical location (proxied by the location of the headquarters), size, age, and position along the productivity distribution. The aim is to provide a set of stylised facts on AI adoption, based on the novel data collected and matched.

This is done not only focusing on all AI adopters, but also distinguishing among different types of AI adopters, using the three groups defined in the previous section ("AI innovation", "AI core business", and "AI talent").

While the core of the analysis focuses on the extensive margin of AI adoption, as identified in the previous section, two additional exploratory exercises are also carried out.

The first exercise consists in exploring the intensive margin of AI adoption for a subset of firms, for which this is possible. In particular, Lightcast™ job posting data can be used to build proxies of the extent to which firms are hiring more intensively AI talent, exploring further the characteristics of those firms.

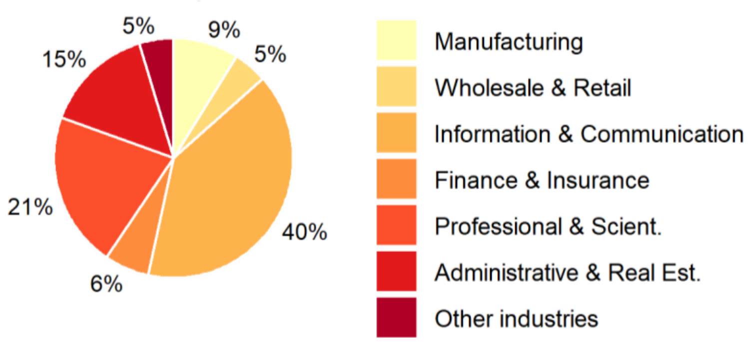
The second one focuses instead on analysing differences between AI adopters and other businesses, likely non-adopters. This is carried out exploiting – in a regression framework – other firms in Orbis© as a comparison group for the identified AI adopters. Further details are discussed in the following subsections.

Fact 1 - Most AI adopters are in ICT and professional services

When analysing the sectoral distribution of AI adopters, it is evident that most of them operate in ICT and professional services. Indeed, as shown in Figure 4.1, about 40% of all AI adopters are in ICT services sectors (ISIC Rev.4 58-63). While this sector accounts for a relatively small share of the economy (around 10%), AI adopters appear particularly concentrated in this sector.[25] This is followed by professional and scientific activities (ISIC Rev.4 69-75), which account for about 21% of AI adopters.

Figure 4.1. AI adopters by broad sector of activity

All AI adopters, United Kingdom, 2019



Notes: NACE classifications are constructed using the NACE2 (4 digit level) classification reported in Orbis© for the year 2019. The category "other industries" includes Mining and Quarrying, Utilities, Transport and Storage, Accommodation and Food, and Construction.
Sources: Authors' own compilation based on GlassAI (January 2021), Lightcast™ (February 2021), IPRs (STI Micro-data Lab: IntellectualPropertyDatabase, http://oe.cd/ipstats, July 2020) and Orbis© (version 2.2020, Bureau van Dijk, April 2021).

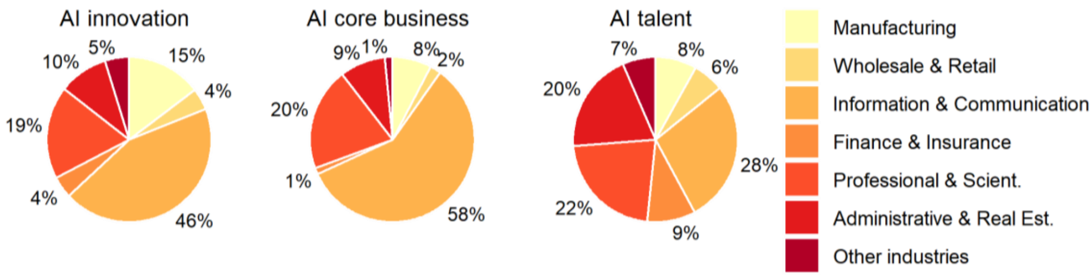
Focusing on different groups of adopters (Figure 4.2) provides additional interesting insights, which are also relevant to keep in mind when interpreting the rest of this characterisation section.

Unsurprisingly, adopters engaged in "AI core business" and "AI innovation" appear more concentrated in ICT sectors. Furthermore, a larger proportion of AI innovators is active in manufacturing if compared with other groups of AI adopters. This may be to some extent related to the very nature of patent data that, together with trademarks, is the main source of information for this group.

Adopters belonging to the group "AI talent" are instead more spread across sectors. Although this excludes AI innovators and AI-core businesses, which tend to be in ICT sectors, this seems to provide further support to consider AI as a GPT.

Figure 4.2. AI adopters by type and broad sector of activity

United Kingdom, 2019



Notes: NACE classifications are constructed using the NACE2 (4 digit level) classification reported in Orbis© for the year 2019. The category "other industries" includes Mining and Quarrying, Utilities, Transport and Storage, Accommodation and Food, and Construction.
Sources: Authors' own compilation based on GlassAI (January 2021), Lightcast™ (February 2021), IPRs (STI Micro-data Lab: IntellectualPropertyDatabase, http://oe.cd/ipstats, July 2020) and Orbis© (version 2.2020, Bureau van Dijk, April 2021).

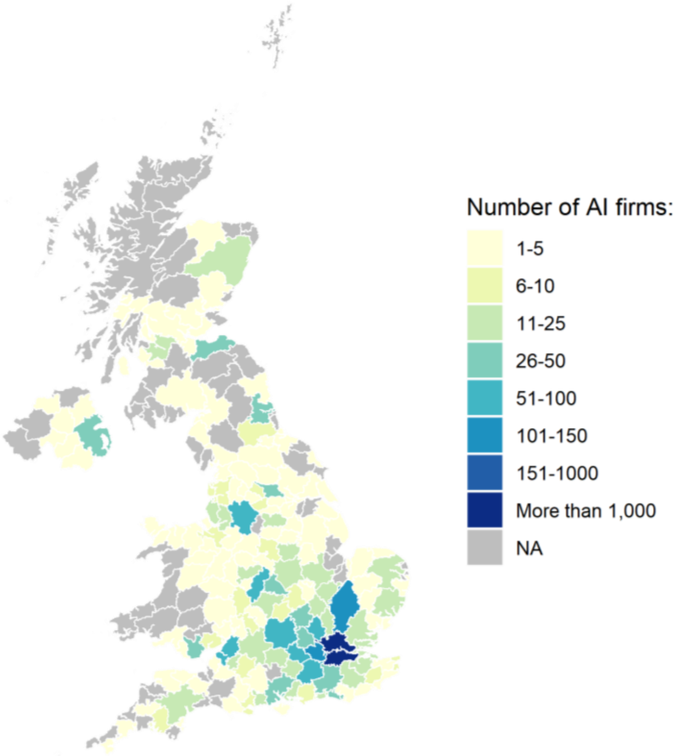
Fact 2 - AI adopters are concentrated in the South of the UK, especially around London

Focusing on the location of AI adopters highlights that they tend to be located in the South of the United Kingdom. In fact, as evident in Figure 4.3, which shows the number of AI adopters by travel-to-work area (TTWA),[26] the large majority of AI adopters are found in and around London (e.g., Luton, Heathrow or Reading). Other noteworthy AI hubs appear to be in Cambridge and Oxford, possibly also due to proximity to universities and hence research activities.

This does not seem to be exclusively due to the fact that the location information used refers to firm headquarters, which may be typically located around the capital. Indeed, when replicating the same heat map but using job postings' location information available for the subset of Lightcast™ firms, patterns remain qualitatively similar.[27] This is also in line with previous findings relying solely on job postings data (Samek, Squicciarini and Cammeraat, 2021[25]) and is not unexpected, considering United Kingdom's concentration of economic activity in those identified regions, in particular in ICT and services (OECD, 2022[37]) and its unequal distribution of skills across different geographical areas (OECD, 2020[38]).

Figure 4.3. Location of AI adopters

All AI adopters, United Kingdom, 2019

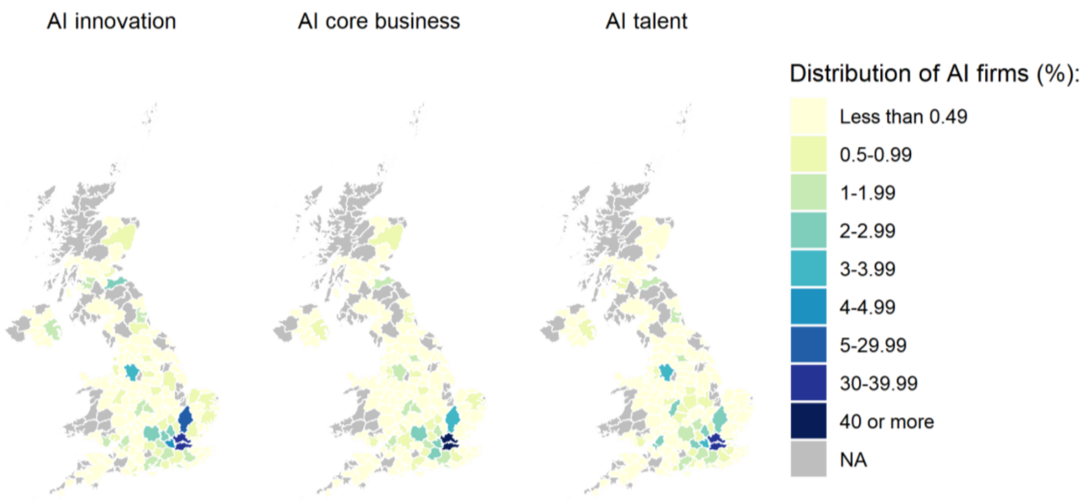


Notes: Location refers to the TTWA of headquarters, as reported in Orbis©.
Sources: Authors' own compilation based on GlassAI (January 2021), Lightcast™ (February 2021), IPR (STI Micro-data Lab: IntellectualPropertyDatabase, http://oe.cd/ipstats, July 2020) and Orbis© (version 2.2020, Bureau van Dijk, April 2021).

Figure 4.4 plots instead, for each group of AI adopters, their geographical distribution across the United Kingdom. While all groups remain geographically concentrated in the South of the United Kingdom and around London, a closer look at the figures shows that 43% of firms with "AI core business" are in London, compared to only 33% of those engaging in "AI innovation"; and 36% of the ones looking for "AI talent".

Figure 4.4. Location of AI adopters by type of adopter

United Kingdom, 2019



Notes: Share is defined as the number of firms carrying out "AI innovation", have an "AI core business" or demand "AI talent" in each TTWA divided by the total number of AI firms active in the respective field. Location refers to the TTWA of headquarters, as reported in Orbis©.
Sources: Authors' own compilation based on GlassAI (January 2021), Lightcast™ (February 2021), IPR (STI Micro-data Lab: IntellectualPropertyDatabase, http://oe.cd/ipstats, July 2020) and Orbis© (version 2.2020, Bureau van Dijk, April 2021).

Fact 3 – There is significant heterogeneity in AI adopters' age and employment size

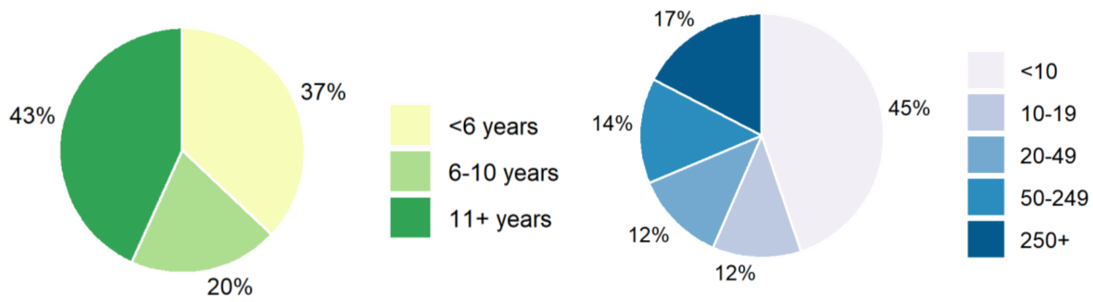
Analysing the age and size distributions of AI adopters highlights significant heterogeneity across firms. This is already evident when focusing on all AI adopters in Figure 4.5, which shows a substantial co-existence of young and old adopters, as well as small and larger ones.[28]

This is even clearer when distinguishing among different groups of adopters. In particular, Figure 4.6 shows that firms in the group "AI core business" are typically young and small, while there exists a significant polarisation between old-large AI adopters and young-small firms for the groups "AI innovation" and "AI talent".[29] These findings are in line with previous and ongoing work by Dernis et al. (2021[18]) and Dernis et al. (forthcoming[19]).

Figure 4.5. Shares of AI adopters by firm age and size class

All AI adopters, United Kingdom, 2019

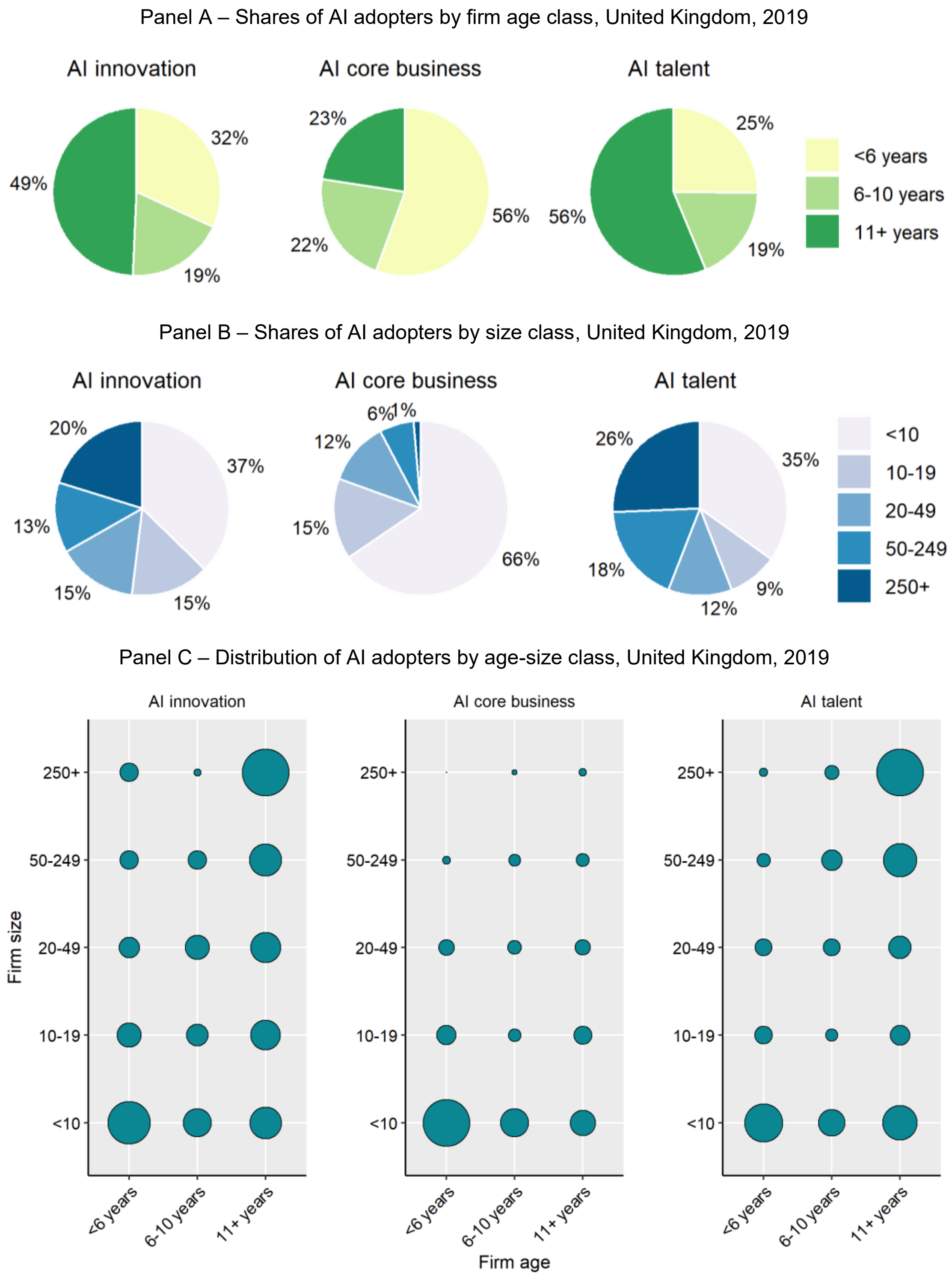
Panel A – Shares of AI adopters by firm age class Panel B – Shares of AI adopters by size class



Notes: Firm age is calculated as the difference between 2019 and the incorporation date, and employment size classes are constructed using the number of employees as reported in Orbis© for the year 2019.
Sources: Authors' own compilation based on GlassAI (January 2021), Lightcast™ (February 2021), IPRs (STI Micro-data Lab: IntellectualPropertyDatabase, http://oe.cd/ipstats, July 2020) and Orbis© (version 2.2020, Bureau van Dijk, April 2021).

Part of the differences across groups of adopters reflect differences in their sectoral composition (see Figure A B.2 and A B.3 in Annex B, which focus on manufacturing and services, respectively). Still, a considerable share of young and small AI adopters remains evident for the "AI core business" group even when separately focusing on the manufacturing sector. Large and old firms in the groups "AI innovation" and "AI talent" tend to operate in the manufacturing sector, while shares of small and young AI adopters are in general larger in services.[30]

Figure 4.6. AI adopters by type of adopter, firm age and size class



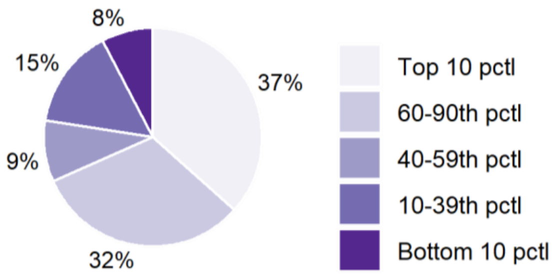
Notes: Firm age is calculated as the difference between 2019 and the incorporation date and size classes are constructed using the number of employees as reported in Orbis© for the year 2019. The size of circles is proportional to firm counts.
Sources: Authors' own compilation based on GlassAI (January 2021), Lightcast™ (February 2021), IPRs (STI Micro-data Lab: IntellectualPropertyDatabase, http://oe.cd/ipstats, July 2020) and Orbis© (version 2.2020, Bureau van Dijk, April 2021).

Fact 4 – AI adopters tend to be concentrated at the top of the sectoral productivity distribution

Studying the relative performance of AI adopters in terms of labour productivity suggests that they tend to be more prevalent among firms at the top of the labour productivity distribution. Indeed, Figure 4.7 plots the share of firms belonging to different percentiles of sector-specific distributions of (operating revenues) turnover over employment, a simple proxy of labour productivity used in this exploratory work. The figure shows that more than one third of all AI adopters is in the top 10% of the productivity distribution, and about another third is in the 60th-90th percentile.[31] This is also true when considering manufacturing and services separately (see Figure A B.4 in Annex B).

Figure 4.7. Shares of AI adopters by percentiles of the productivity distribution

All AI adopters, United Kingdom, 2019

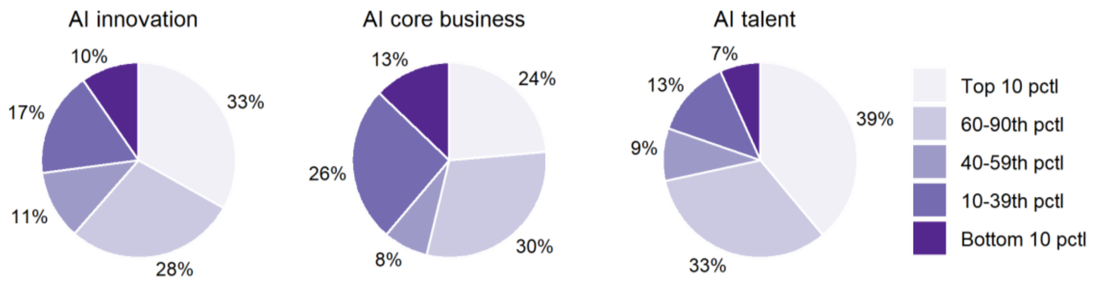


Notes: Productivity classes are constructed within SNA A38 sectors using a proxy of labour productivity (operating revenue turnover over number of employees). All data is collected from Orbis© for the year 2019 and implausible turnover or employee counts are dropped prior to calculating productivity.
Sources: Authors' own compilation based on GlassAI (January 2021), Lightcast™ (February 2021), IPRs (STI Micro-data Lab: IntellectualPropertyDatabase, http://oe.cd/ipstats, July 2020) and Orbis© (version 2.2020, Bureau van Dijk, April 2021).

As shown in Figure 4.8, AI firms confirm to be productivity leaders in their sectors also when focusing on different groups of AI adopters. This is particularly true for firms engaged in "AI innovation" and in search of "AI talent", while less so for firms in the "AI core business" group. Indeed, the previous discussion also suggests the latter group is mainly composed of young-small firms in the ICT sector that may still have to realise their full potential in terms of productivity.

Figure 4.8. Shares of AI adopters by type of adopter and percentiles of the productivity distribution

United Kingdom, 2019



Notes: Productivity classes are constructed within SNA A38 sectors using a proxy of labour productivity (operating revenue turnover over number of employees). All data is collected from Orbis© for the year 2019 and implausible turnover or employee counts are dropped prior to calculating productivity.
Sources: Authors' own compilation based on GlassAI (January 2021), Lightcast™ (February 2021), IPRs (STI Micro-data Lab: IntellectualPropertyDatabase, http://oe.cd/ipstats, July 2020) and Orbis© (version 2.2020, Bureau van Dijk, April 2021).

The relatively higher presence of AI adopters at the top of productivity distributions does not necessarily mean that AI use has a positive effect on productivity. These findings are purely descriptive and may hint at the existence of selection dynamics, in which already most productive firms select into AI use. Additional analysis on AI adoption and productivity is also carried out in the following subsections.

These findings may have relevant implications for the evolution of productivity divides between most productive firms and the rest, which have been already growing across many OECD countries in the past decades (Andrews, Criscuolo and Gal, 2016[30]). Furthermore, given the geographical distribution of AI adoption outlined above and their concentration around London, these findings may also have implications for regional disparities within the United Kingdom.

Further exploring the intensive margin of AI adoption

So far the analysis has focused on the characteristics of AI adopters, which were comprehensively identified using different data sources. A related but yet underexplored question is whether there exists some heterogeneity in the patterns of AI adoption across firms, and whether AI adopters differ in the intensity with which they use AI. A first consideration of this question is possible with a subset of the database built for the current analysis.

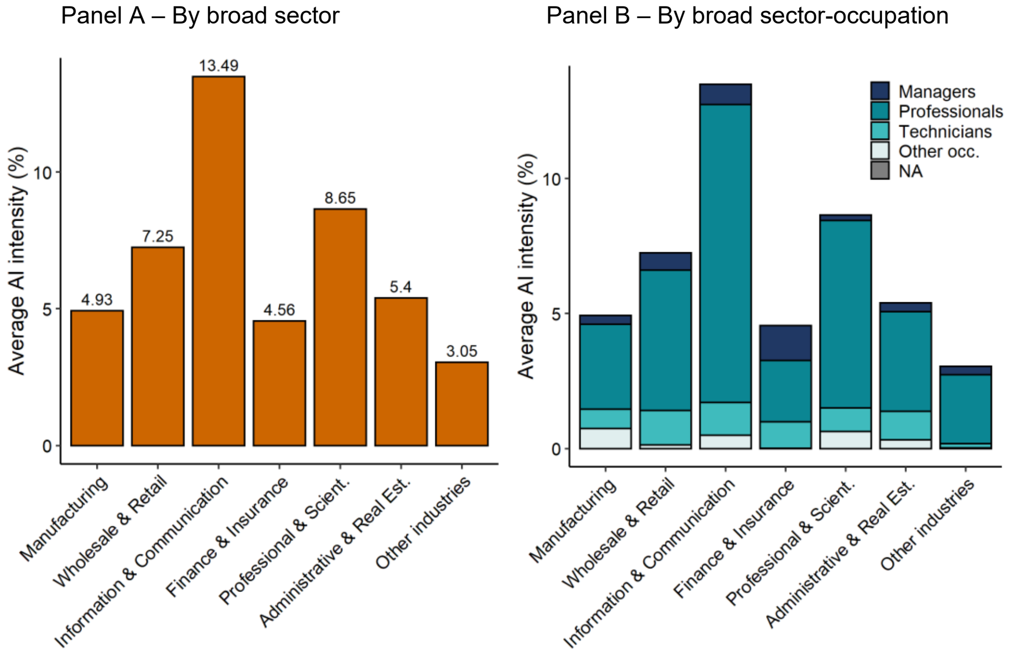
In particular, online job postings (Lightcast™) allow observing the extent to which different firms are demanding AI talent, which can be a proxy of increasing specialisation in AI, possibly hinting at higher AI intensity in the future, or longer-term commitment to AI.

In this context, a proxy of "AI-hiring intensity" is constructed for the subset of AI adopters identified through Lightcast™ data.[32] This is defined as the ratio between their number of AI job postings divided by their total job postings.[33] Exploiting the match with firm-level financials (Orbis©), it is then possible to explore some of the heterogeneity in the recent AI hiring patterns according to firm characteristics.

As observed in Figure 4.9 (Panel A) and to some extent unsurprisingly, firms in ICT and professional services sectors tend to have higher average AI-hiring intensity. More interestingly, among all AI jobs, AI professionals appear the most demanded, especially in ICT and professional services sectors (Figure 4.9, Panel B).[34]

Figure 4.9. Average AI-hiring intensity across sectors of activity

United Kingdom, 2019



Notes: The sample only contains AI adopters identified in Lightcast™ that were successfully linked to Orbis©. AI intensity is measured as the share of AI job postings out of all vacancies, on average for firms in each sector. AI intensity is assumed to be zero for the years in which firms do not advertise any AI jobs that is captured in Lightcast™. The category "other industries" includes Mining and Quarrying, Utilities, Transport and Storage, Accommodation and Food, and Construction.
Sources: Authors' own compilation based Lightcast™ (February 2021) and Orbis© (version 2.2020, Bureau van Dijk, April 2021).

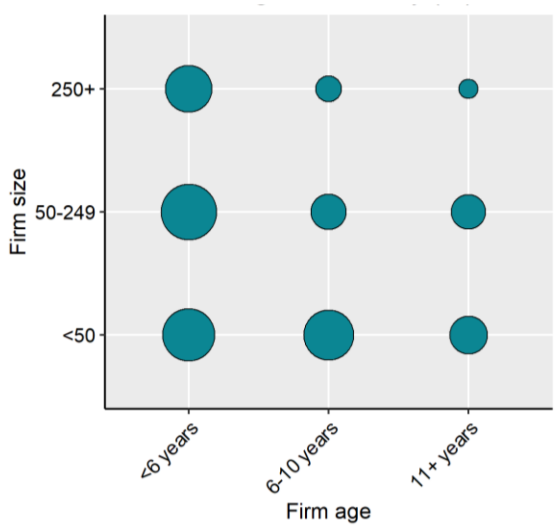
In line with the previous discussion, the South of the United Kingdom is also characterised by higher average AI-hiring intensity with respect to the North (Figure A B.5 in Annex B). However, patterns of AI-hiring intensity appear more spread out with respect to what can be observed in Figure 4.3.

When exploring the heterogeneity of AI-hiring intensity across age and size groups (Figure 4.10), young firms (especially those with less than 6 years of age) tend to have higher average AI-hiring intensity. This may reflect the fact that, for given size classes, younger firms may be, to some extent, better positioned to leverage the potential of the latest developments of AI. Alternatively, they may focus all the efforts in AI-related hiring, contrary to larger firms, which need to keep a more diverse hiring portfolio.

Further analysis may extend these initial findings, which represent a first attempt to tackle a question that has received little attention in the literature so far, mainly due to constraints in data availability.

Figure 4.10. Average AI-hiring intensity by firm age and size

United Kingdom, 2019



Notes: The sample only contains AI adopters identified in Lightcast™ that were successfully linked to Orbis©. AI intensity is measured as the share of AI job postings out of all vacancies, on average for firms in each sector. AI intensity is assumed to be zero for the years in which firms do not advertise any AI jobs that is captured in Lightcast™.
Sources: Authors' own compilation based Lightcast™ (February 2021) and Orbis© (version 2.2020, Bureau van Dijk, April 2021).

Towards a comparison of AI adopters with other firms

As a final exploratory exercise, the analysis compares – still focusing on 2019 – the identified AI adopters with other firms in Orbis©, in order to explore differences between AI adopters and other likely non-adopting businesses, both in terms of their characteristics and in terms of (a proxy of labour) productivity.

This is complementary to ongoing OECD work exploiting different data sources, such as administrative ICT surveys, to investigate similar issues. However, few additional caveats need to be taken into account when carrying out this analysis and interpreting its results.

As previously mentioned, Orbis©' representativeness may be more limited for small and less productive firms, and selection issues may likely arise (Bajgar et al., 2020[31]). In particular, when it comes to smaller firms, the most productive ones may be better covered in the data. Although this appears less problematic for AI adopters, which turn out to be at the top of productivity distributions, it may be a more relevant issue when comparing those with other firms in the economy.

Furthermore, despite the efforts carried out and the use of several data sources to identify AI use, the comparison group of other firms in Orbis© may still contain some AI adopters.[35] The productivity analysis will be therefore carried out imposing increasingly higher size thresholds, in order to assess how results may change and to attempt limiting those issues. Additional robustness checks are also carried out, such as including multiple matches in the AI adopters group, which corroborate the results presented in this section and are elaborated on later.

As a starting point of this analysis, descriptive statistics are reported in Table 4.2 below.

Table 4.2. Summary statistics about AI adopters and other firms

United Kingdom, 2019

	AI adopters (all)	Other firms
Employment size – mean (std. dev.)	1,417 (11,514)	19 (1,026)
Firm age – mean (std. dev.)	14.8 (19.3)	9.2 (11.3)
Turnover [in GBP million] – mean (std. dev.)	1,498 (12,264)	10 (410)
N	2,796	1,441,773

Note: Firm age is calculated as the difference between 2019 and the incorporation date, employment refers to the number of employees and turnover to the operating revenues as reported in Orbis© for the year 2019. Implausible turnover or employee counts are dropped. N refers to the highest count of non-missing observations considering employment, firm age and turnover. As discussed above, shares of missing values are significantly higher for the turnover variable.
Source: Authors' own compilation based on GlassAI (January 2021), Lightcast™ (February 2021), IPRs (STI Micro-data Lab: IntellectualPropertyDatabase, http://oe.cd/ipstats, July 2020) and Orbis© (version 2.2020, Bureau van Dijk, April 2021).

These basic descriptive statistics already show that AI adopters appear on average older and larger, vis-à-vis other firms in the database.[36] However, these do not take into account several confounding factors and compositional effects that may drive those differences.

The analysis therefore explores differences between AI adopters and other firms in a regression framework, in order to take into further account relevant confounding factors, such as sectoral composition and location.

In this context, Equation 1 below is estimated using a simple linear probability model

$$AI\ adoption_i = \alpha + age\ class_i{}'\beta_1 + size\ class_i{}'\beta_2 + \gamma_{sect} + \rho_{reg} + \epsilon_i\ ,$$

Equation 1

where *AI adoption* indicates a dummy variable that equals one for firm *i* tagged as an AI adopter and zero if tagged as a non-adopter; α is a constant; *age class* represents age class dummies (reference category: < 6 years old); $size\ class$ represents size class dummies (reference category: < 10 employees); γ_{sect} represents 2-digit-sector fixed effects; ρ_{reg} indicates TL2-regions fixed effects; and ϵ is the usual error term.

Table 4.3. AI adoption and firm characteristics

United Kingdom, 2019

	AI adoption dummy
Firm age class 6-10 years	-0.00004
	(0.00009)
Firm age class 11 or more years	0.00030***
	(0.00009)
Employment size class 10-49 employees	0.00357***
	(0.00018)
Employment size class 50-249 employees	0.01145***
	(0.00067)
Employment size class 250 or more employees	0.05757***
	(0.00274)
Observations	1,923,222
Industry dummies	YES
Regional dummies	YES
Adjusted R-squared	0.0140

Notes: Coefficients for the constant and missing firm age and employment size classes are not reported. Robust standard errors in parentheses (*** p<0.01, ** p<0.05, * p<0.1). Baseline category for firm age is <6 years and for employment size <10 employees.
Source: Authors' own compilation based on GlassAI (January 2021), Lightcast™ (February 2021), IPRs (STI Micro-data Lab: IntellectualPropertyDatabase, http://oe.cd/ipstats, July 2020) and Orbis© (version 2.2020, Bureau van Dijk, April 2021).

Results, presented in Table 4.3 above, suggest that there appear to be significant scale advantages in AI adoption, even after accounting for regional and sectoral heterogeneity. This is evident when focusing on the coefficients of the size class dummies, which increase in magnitude with size.

Furthermore, the age coefficients suggest that, to some extent and after taking into account other confounding factors, on average all AI adopters tend to be older than other firms, although differences appear rather small in magnitude.[37]

Although the data at hand offer limited potential to focus on the characteristics external to the firms that may affect AI adoption, one of the factors for which information is available and that appears interesting to explore further is the role of AI universities. AI universities have indeed been identified by the web-reading exercise carried out by GlassAI in the same way AI-related companies have, i.e., by focusing on AI-related keywords present on their websites. This allows exploring the role that such universities may play for the probability of AI adoption, as proxied by the *AI adoption* dummy reported in Equation 1.

In this context, Equation 1 is re-estimated adding as an additional explanatory variable a dummy that equals one if AI-related universities are present in the same TTWA, and zero otherwise. Results, available in Table A B.1 in Annex B, suggest that the presence of AI universities is positively related to AI adoption.[38] These are in line with the idea that tertiary education plays an important role to boost the availability of skilled human capital needed for AI adoption.

A number of robustness checks are carried out and qualitatively confirm the findings reported above. These include using different age and size bands, increasing the firm size threshold, and using logit or cloglog regressions instead of linear probability models.[39]

Beyond the role of firm characteristics for AI adoption, as discussed in the literature section, the recent debate has been focusing more and more on the role of AI for productivity. In this context, the database constructed can help assess the productivity characteristics of AI adopters, as a first step to better understand the link between AI adoption and productivity. Equation 2 is therefore estimated using ordinary least squares. The model reads as follows:

$$y_{i,t} = \alpha + \beta_1 \, AI\ adoption_{i,t} + \beta_2 \, size_{i,t-1} + \beta_3 \, age_{i,t} + \gamma_{sect} + \rho_{reg} + \epsilon_{i,t} \,,$$

Equation 2

where *y* is (the logarithm of) turnover over employment, the simple proxy used so far for labour productivity; *AI adoption* is the same dummy variable used in Equation 1; *size* corresponds to the (lagged logarithm of) number of employees; *age* corresponds to firm age; γ_{sect} represents 2-digit-sector fixed effects; ρ_{reg} indicates TL2-regions fixed effects; and ϵ is the usual error term

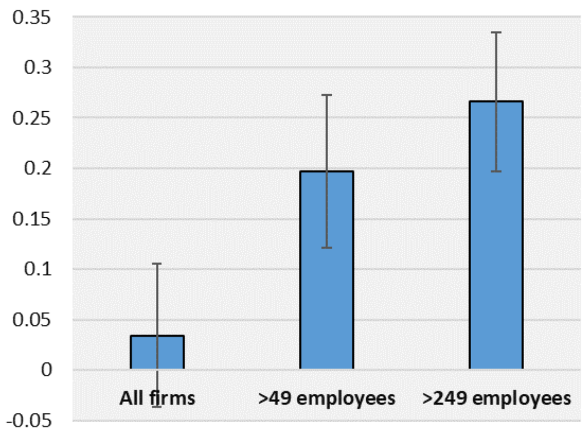
Equation 2 is fist estimated for different size groups and then for manufacturing and services separately.[40]

Results shown graphically in Figure 4.11 (see also Table A B.2 in Annex B) highlight that, especially when focusing on larger firms (particularly those with more than 49 employees, for which Orbis© is more representative), significant productivity premia emerge for AI adopters.[41] When separating out manufacturing and services (see Table A B.2 of Annex B), results show that productivity premia appear stronger in market services.

Figure 4.11. AI adoption and labour productivity

Regression coefficients of *AI adoption* dummies across different specifications

United Kingdom, 2019



Notes: The graph shows the coefficients (β_1) and error bars related to the *AI adoption* dummies from Equation 2, focusing on all firms, firms with more than 49 employees, and firms with more than 249 employees.
Source: Authors' own compilation based on GlassAI (January 2021), Lightcast™ (February 2021), IPRs (STI Micro-data Lab: IntellectualPropertyDatabase, http://oe.cd/ipstats, July 2020) and Orbis© (version 2.2020, Bureau van Dijk, April 2021).

These findings suggest that AI adopters may be more efficient than other firms, but do not imply that AI adoption causes an increase in firm productivity. In fact, results reported indicate simple conditional correlations. These may be likely driven by self-selection into AI adoption by firms that are already more productive or that have already invested in complementary assets required to effectively use AI, such as skills, other complementary technologies or intangible capital.[42]

Additional robustness checks have been carried out, which qualitatively confirm the main takeaways reported above. Those include replacing *y* with (contemporaneous) turnover and adding as an additional control (lagged) employment in a specification similar to the one adopted by Czarnitzki et al. (2022[39]), or changing the structure of fixed effects used as controls.[43]

Equation 2 is then estimated focusing on different groups of AI adopters using *AI adoption* dummies that equal one for firms belonging to any of the three groups, and zero for non-adopters. Zooming in on larger firms, for which productivity premia are evident and financial data are more representative (Table A B.3 in Annex B), suggests that productivity premia tend to be higher for the groups “AI innovation” and “AI talent”, while firms with “AI core businesses” tend to be less productive than other (non-AI) firms. These firms are relatively smaller and younger, and may be still in the initial phase of the “J curve” (Brynjolfsson, Rock and Syverson, 2021[5]), investing in complementary assets in view of future returns.[44]

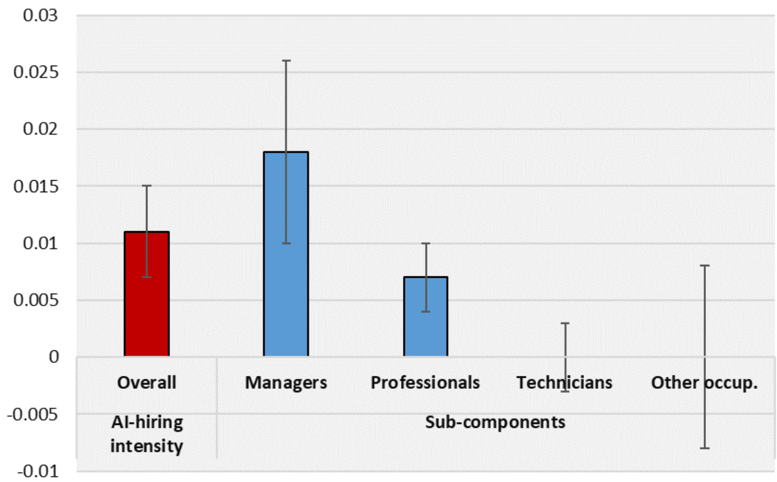
Finally, the role of the intensive margin of AI adoption and its link with productivity is explored. In particular, the measure of AI-hiring intensity discussed in the previous section – and available for adopters identified via online job postings – is used instead of the *AI adoption* dummy in Equation 2.[45]

Results, reported graphically in Figure 4.12 focusing on firms with more than 49 employees (see also Table A B.4 in Annex B for additional specifications using different firm size thresholds) suggest that productivity premia of (larger) AI adopters emerge also when focusing on the intensive margin, as captured by the proxy used.

Figure 4.12. AI-hiring intensity and labour productivity

Regression coefficients of AI-hiring intensity, firms with more than 49 employees

United Kingdom, 2019



Note: The graph shows the coefficients (β_1) and errors bars related to the AI-hiring intensity variable(s). They are based on Equation 2, but using AI-hiring intensity (or its four components) in place of the *AI adoption* dummy. The focus is on firms with more than 49 employees.
Source: Authors' own compilation based on Lightcast™ (February 2021) and Orbis© (version 2.2020, Bureau van Dijk, April 2021).

Furthermore, when focusing on the role of different AI-related occupations, i.e. decomposing AI-hiring intensity into four components (focusing on managers, professionals, technicians and other occupations), an important role of managers and – somewhat to a lesser extent – professionals also emerges.[46]

Additional robustness have been also carried out. Those include considering multiple matches in the group of AI adopters; using different sectoral filterings; prioritising consolidated accounts rather than unconsolidated ones; imputing labour productivity (by relying on employment and/or turnover from the previous year when missing in 2019). Those exercises qualitatively confirm the results reported.

5 Concluding remarks and next steps

This work has proposed a novel scalable approach to identify and characterise AI adopters, combining different sources of large commercial and administrative microdata, and overcoming significant challenges.

First, this work has identified AI adopters based on a common set of AI-related keywords, relying on and combining three different sources of data: i) AI-related IPRs, which capture firms innovating in AI or embedding AI in their goods or services; ii) AI-related activities as stated on company websites; and iii) the demand for AI-related skills contained in online job postings by employers. These data sources turn out to be highly complementary, highlighting the relevance of combining them.

Second, data on AI adopters have been matched with firm-level financial statements that contain information on their characteristics, such as firm size, age, detailed sector of activity, location of the headquarters, and that allow computing proxies of productivity. This has allowed to provide a picture of AI adoption based on the most recent available data, importantly distinguishing among different margins of AI adopters (firms carrying out AI innovation, those with an AI core business, and those searching for AI talent).

A number of stylised facts have been uncovered. In particular, a polarisation of AI adoption has emerged, both in terms of industry activity – with most firms operating in ICT and professional services – and geographical presence – with AI adopters concentrated in the South of the United Kingdom. Hiring of AI talent appears however more widespread across sectors, providing evidence in favour of the general-purpose nature of AI technologies.

Furthermore, AI adopters tend to be leaders in their industry (in terms of labour productivity), with possible implications for existing divides across groups of firms (leaders vs. laggards) and regions. A coexistence of young-small and old-large AI adopters has also emerged, and among firms that tend to have AI at the core of their business, most of those identified by the current analysis are young and small.

Exploring the intensive margin of AI adoption using a proxy of AI-hiring intensity has further highlighted the role of AI professionals, for which a strong demand is evident, and of young firms, which tend to have higher AI-hiring intensity.

When comparing AI adopters with other likely non-adopting firms, significant scale advantages have emerged, with AI adopters generally being larger than other firms. AI adopters also tend to be more productive, although these productivity premia – especially evident in market services – do not necessarily imply a positive effect of AI on productivity.

Human capital appears to play an important role for AI adoption. Proximity to AI-related universities seem to increase the likelihood of using AI. Among different occupational groups, managers as well as professionals appear to play an important role, likely helping translate AI use into higher efficiency.

Although largely exploratory, to our knowledge this is the most ambitious effort combining different commercial and administrative data sources to study AI adoption in firms. This is complementary to OECD work using other data sources, such as ICT surveys, to study AI diffusion.

The analysis can be extended in a number of directions. Further work could focus on the role of human capital or complementary assets in explaining differences between AI adopters and other firms. In particular, additional analysis could further focus on the role of management, or on the different set of skills required by AI adopters vis-à-vis other firms, or on collaborations between universities and firms (rather than their physical proximity). Furthermore, while the current analysis focuses on the United Kingdom, a country for which underlying data are generally of good quality and for which to our knowledge very limited evidence exists, the approach adopted is also scalable to other countries. Finally, future work could further analyse the different implications of AI adoption on other outcomes, also depending on the specific margin of AI adoption considered.

References

Acemoglu, D. et al. (2022), “Artificial Intelligence and Jobs: Evidence from Online Vacancies”, *Journal of Labor Economics*, Vol. 40/S1, pp. S293-S340, https://doi.org/10.1086/718327. [34]

Agrawal, A., J. Gans and A. Goldfarb (2018), *Prediction Machines: The Simple Economics of Artificial Intelligence*, Harvard Business Review Press. Cambridge, MA. [53]

Alderucci, D. et al. (2020), “Quantifying the Impact of AI on Productivity and Labor Demand: Evidence from U.S. Census Microdata”, https://www.aeaweb.org/conference/2020/preliminary/paper/Tz2HdRna. [13]

Alekseeva, L. et al. (2021), “The demand for AI skills in the labor market”, *Labour Economics*, Vol. 71, p. 102002, https://doi.org/10.1016/j.labeco.2021.102002. [22]

Alekseeva, L. et al. (2020), “AI Adoption and Firm Performance: Management versus IT”, *SSRN Electronic Journal*, https://doi.org/10.2139/ssrn.3677237. [21]

Amoroso, S. et al. (2021), *World Corporate Top R&D investors: Paving the way for climate neutrality*, https://doi.org/10.2760/49552. [40]

Andrews, D., C. Criscuolo and P. Gal (2016), “The Best versus the Rest: The Global Productivity Slowdown, Divergence across Firms and the Role of Public Policy”, *OECD Productivity Working Papers*, No. 5, OECD Publishing, Paris, https://doi.org/10.1787/63629cc9-en. [30]

Arora, A., X. Bei and W. Cohen (2016), “Why Firms Trademark (or not): Evidence from the US Trademark Data”, *Academy of Management Proceedings*, Vol. 2016/1, p. 17249, https://doi.org/10.5465/ambpp.2016.17249abstract. [28]

Babina, T. et al. (2020), “Artificial Intelligence, Firm Growth, and Industry Concentration”, *SSRN Electronic Journal*, https://doi.org/10.2139/ssrn.3651052. [23]

Bäck, A. et al. (2022), “Return of the Solow-paradox in AI? AI-adoption and firm productivity”, *Papers in Innovation Studies*, Vol. no. 2022/01. [26]

Bajgar, M. et al. (2020), "Coverage and representativeness of Orbis data", *OECD Science, Technology and Industry Working Papers*, No. 2020/06, OECD Publishing, Paris, https://doi.org/10.1787/c7bdaa03-en. [31]

Baruffaldi, S. et al. (2020), "Identifying and measuring developments in artificial intelligence: Making the impossible possible", *OECD Science, Technology and Industry Working Papers*, No. 2020/05, OECD Publishing, Paris, https://doi.org/10.1787/5f65ff7e-en. [2]

Bessen, J. (2019), "Automation and jobs: when technology boosts employment*", *Economic Policy*, Vol. 34/100, pp. 589-626, https://doi.org/10.1093/epolic/eiaa001. [51]

Bloom, N. et al. (2021), *The Diffusion of Disruptive Technologies*, National Bureau of Economic Research, Cambridge, MA, https://doi.org/10.3386/w28999. [27]

Borowiecki, M. et al. (2021), "The impact of digitalisation on productivity: Firm-level evidence from the Netherlands", *OECD Economics Department Working Papers*, No. 1680, OECD Publishing, Paris, https://doi.org/10.1787/e800ee1d-en. [57]

Brynjolfsson, E., W. Jin and K. McElheran (2021), "The power of prediction: predictive analytics, workplace complements, and business performance", *Business Economics*, Vol. 56/4, pp. 217-239. [55]

Brynjolfsson, E., D. Rock and C. Syverson (2021), "The Productivity J-Curve: How Intangibles Complement General Purpose Technologies", *American Economic Journal: Macroeconomics*, Vol. 13/1, pp. 333-372, https://doi.org/10.1257/mac.20180386. [5]

Calvino, F. et al. (2018), "A taxonomy of digital intensive sectors", *OECD Science, Technology and Industry Working Papers*, No. 2018/14, OECD Publishing, Paris, https://doi.org/10.1787/f404736a-en. [59]

Calvino, F. and L. Fontanelli (2022), "A portrait of AI adopters across countries: firm characteristics, assets' complementarity, and productivity", *OECD internal document*. [11]

Calvino, F. and L. Fontanelli (2022), *A portrait of AI adopters: evidence from France*, https://www.nber.org/conferences/si-2022-it-and-digitization. [10]

Calvino, F., D. Giachini and M. Guerini (2021), "The age distribution of business firms", *Journal of Evolutionary Economics*, Vol. 32/1, pp. 205-245, https://doi.org/10.1007/s00191-021-00747-2. [54]

Cammeraat, E. and M. Squicciarini (2021), "Burning Glass Technologies' data use in policy-relevant analysis: An occupation-level assessment", *OECD Science, Technology and Industry Working Papers*, No. 2021/05, OECD Publishing, Paris, https://doi.org/10.1787/cd75c3e7-en. [50]

Carnevale, A., T. Jayasundera and D. Repnikov (2014), *Understanding online job ads data*. [48]

Castaldi, C., J. Block and M. Flikkema (2019), "Editorial: why and when do firms trademark? Bridging perspectives from industrial organisation, innovation and entrepreneurship", *Industry and Innovation*, Vol. 27/1-2, pp. 1-10, https://doi.org/10.1080/13662716.2019.1685376. [29]

Center for Security and Emerging Technology (2020), *U.S. Demand for AI-Related Talent*, Center for Security and Emerging Technology, https://doi.org/10.51593/20200027. [47]

Chetty, R. et al. (2020), *The Economic Impacts of COVID-19: Evidence from a New Public Database Built Using Private Sector Data*, National Bureau of Economic Research, Cambridge, MA, https://doi.org/10.3386/w27431. [49]

Cho, J. et al. (2022), "What's driving the diffusion of next-generation digital technologies?", *Technovation*, p. 102477, https://doi.org/10.1016/j.technovation.2022.102477. [9]

Cockburn, I., R. Henderson and S. Stern (2018), *The Impact of Artificial Intelligence on Innovation*, National Bureau of Economic Research, Cambridge, MA, https://doi.org/10.3386/w24449. [52]

Crafts, N. (2021), "Artificial intelligence as a general-purpose technology: an historical perspective", *Oxford Review of Economic Policy*, Vol. 37/3, pp. 521-536, https://doi.org/10.1093/oxrep/grab012. [4]

Criscuolo, C., P. Gal and C. Menon (2014), "The Dynamics of Employment Growth: New Evidence from 18 Countries", *OECD Science, Technology and Industry Policy Papers*, No. 14, OECD Publishing, Paris, https://doi.org/10.1787/5jz417hj6hg6-en. [60]

Czarnitzki, D., G. Fernández and C. Rammer (2022), "Artificial Intelligence and Firm-Level Productivity", *SSRN Electronic Journal*, https://doi.org/10.2139/ssrn.4049824. [39]

Damioli, G., V. Van Roy and D. Vertesy (2021), "The impact of artificial intelligence on labor productivity", *Eurasian Business Review*, Vol. 11/1, pp. 1-25, https://doi.org/10.1007/s40821-020-00172-8. [16]

Dernis, H. et al. (forthcoming), "Identifying Artificial Intelligence actors using on-line data", *OECD Science, Technology and Industry*. [19]

Dernis, H. et al. (2021), "Who develops AI-related innovations, goods and services? : A firm-level analysis", *OECD Science, Technology and Industry Policy Papers*, No. 121, OECD Publishing, Paris, https://doi.org/10.1787/3e4aedd4-en. [18]

Domini, G. et al. (2022), “For whom the bell tolls: The firm-level effects of automation on wage and gender inequality”, *Research Policy*, Vol. 51/7, p. 104533, https://doi.org/10.1016/j.respol.2022.104533. [12]

Georgieff, A. and R. Hyee (2021), “Artificial intelligence and employment : New cross-country evidence”, *OECD Social, Employment and Migration Working Papers*, No. 265, OECD Publishing, Paris, https://doi.org/10.1787/c2c1d276-en. [56]

Gierten, D. et al. (2021), “Firms going digital: Tapping into the potential of data for innovation”, *OECD Digital Economy Papers*, No. 320, OECD Publishing, Paris, https://doi.org/10.1787/ee8340c1-en. [58]

HM Government (2021), *National AI Strategy*, https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/1020402/National_AI_Strategy_-_PDF_version.pdf. [61]

Igna, I. and F. Venturini (2022), “The determinants of AI innovation across European firms”, *Papers in Innovation Studies*, Lund University, CIRCLE - Centre for Innovation Research. [15]

Inaba, T. and M. Squicciarini (2017), “ICT: A new taxonomy based on the international patent classification”, *OECD Science, Technology and Industry Working Papers*, No. 2017/1, OECD Publishing, Paris, https://doi.org/10.1787/ab16c396-en. [41]

Jovanovic, B. and P. Rousseau (2005), *General Purpose Technologies*, National Bureau of Economic Research, Cambridge, MA, https://doi.org/10.3386/w11093. [6]

Kinne, J. and J. Axenbeck (2019), “Web Mining of Firm Websites: A Framework for Web Scraping and a Pilot Study for Germany”, *ZEW Discussion Paper*, Vol. No. 18-033, https://ftp.zew.de/pub/zew-docs/dp/dp18033.pdf. [32]

Levenstein, V. (1966), “Binary Codes Capable of Correcting Deletions, Insertions and Reversals”, *Soviet Physics Doklady*, Vol. 10/8, pp. 707-710. [35]

Nakazato, S. and M. Squicciarini (2021), “Artificial intelligence companies, goods and services: A trademark-based analysis”, *OECD Science, Technology and Industry Working Papers*, No. 2021/06, OECD Publishing, Paris, https://doi.org/10.1787/2db2d7f4-en. [17]

Nathan, M. and A. Rosso (2022), “Innovative events: product launches, innovation and firm performance”, *Research Policy*, Vol. 51/1, p. 104373, https://doi.org/10.1016/j.respol.2021.104373. [44]

Nathan, M. and A. Rosso (2015), “Mapping digital businesses with big data: Some early findings from the UK”, *Research Policy*, Vol. 44/9, pp. 1714-1733, https://doi.org/10.1016/j.respol.2015.01.008. [43]

OECD (2022), *OECD Economic Surveys: United Kingdom 2022*, OECD Publishing, Paris, https://doi.org/10.1787/7c0f1268-en. [37]

OECD (2020), *Enhancing Productivity in UK Core Cities: Connecting Local and Regional Growth*, OECD Urban Policy Reviews, OECD Publishing, Paris, https://doi.org/10.1787/9ef55ff7-en. [38]

OECD (2019), *Artificial Intelligence in Society*, OECD Publishing, Paris, https://doi.org/10.1787/eedfee77-en. [1]

ONS (2016), *Travel to Work Areas*, https://webarchive.nationalarchives.gov.uk/20160106004211/http://www.ons.gov.uk/ons/guide-method/geography/beginner-s-guide/other/travel-to-work-areas/index.html. [62]

Pouliakas, K. (2021), *Artificial intelligence and job automation: an EU analysis using online job vacancy data*, http://data.europa.eu/doi/10.2801/305373. [33]

Ragoussis, A. and J. Timmis (2022), "Global Transition Online", *World Bank Policy Research Working Paper* No. 9951, http://hdl.handle.net/10986/37103. [45]

Rammer, C. and N. Es-Sadki (2022), "Using Big Data for Generating Firm-Level Innovation Indicators – A Literature Review", *SSRN Electronic Journal*, https://doi.org/10.2139/ssrn.4072590. [46]

Rammer, C., G. Fernández and D. Czarnitzki (2022), "Artificial intelligence and industrial innovation: Evidence from German firm-level data", *Research Policy*, Vol. 51/7, p. 104555, https://doi.org/10.1016/j.respol.2022.104555. [8]

Samek, L., M. Squicciarini and E. Cammeraat (2021), "The human capital behind AI: Jobs and skills demand from online job postings", *OECD Science, Technology and Industry Policy Papers*, No. 120, OECD Publishing, Paris, https://doi.org/10.1787/2e278150-en. [25]

Santarelli, E., J. Staccioli and M. Vivarelli (2022), "Automation and related technologies: a mapping of the new knowledge base", *The Journal of Technology Transfer*, https://doi.org/10.1007/s10961-021-09914-w. [14]

Squicciarini, M., H. Dernis and C. Criscuolo (2013), "Measuring Patent Quality: Indicators of Technological and Economic Value", *OECD Science, Technology and Industry Working Papers*, No. 2013/3, OECD Publishing, Paris, https://doi.org/10.1787/5k4522wkw1r8-en. [42]

Squicciarini, M. and H. Nachtigall (2021), "Demand for AI skills in jobs: Evidence from online job postings", *OECD Science, Technology and Industry Working Papers*, No. 2021/03, OECD Publishing, Paris, https://doi.org/10.1787/3ed32d94-en. [24]

Tambe, P. (2013), "Big Data Investment, Skills, and Firm Value", *SSRN Electronic Journal*, https://doi.org/10.2139/ssrn.2294077. [20]

Trajtenberg, M. (2018), *AI as the next GPT: a Political-Economy Perspective*, National Bureau of Economic Research, Cambridge, MA, https://doi.org/10.3386/w24245. [3]

Winkler, W. (1999), “The State of Record Linkage and Current Research Problems”, *U. S. Bureau of the Census Working Paper* RR99-04. [36]

Zolas, N. et al. (2020), *Advanced Technologies Adoption and Use by U.S. Firms: Evidence from the Annual Business Survey*, National Bureau of Economic Research, Cambridge, MA, https://doi.org/10.3386/w28290. [7]

Annex A. Additional information about data and methodology

This Annex summarises and provides additional information about the data sources used in a complementary way in this work and the matching methodology employed to combine those.

Patent and trademarks

The OECD Science, Technology and Innovation (STI) Micro-data Lab collects administrative micro-data on Intellectual Property Rights (IPRs) records (patents and trademarks) containing information on the owners of inventions, goods and services protected by IPRs, as well as on their technological scope and product categories. The STI Micro-data Lab has collected, processed and analysed such information over years, which resulted in several influential publications, such as a joint publication of the European Commission's Joint Research Centre, the European Commission's science and knowledge service and the OECD on the world corporate top R&D investors (Amoroso et al., 2021[40]), a definition of Information and Communication Technology (ICT) based on the technology classes of the International Patent Classification (IPC) (Inaba and Squicciarini, 2017[41]) or a paper on defining and measuring patent quality (Squicciarini, Dernis and Criscuolo, 2013[42]).

Although IPRs are not the only mechanism to protect innovation and patentability of software-based technologies across countries, combining information about AI-related patents and trademarks already represents a significant improvement in the study of AI innovators. In fact, the existing literature has often used only one source of IPRs data to identify AI adopters.

Patent data used in this work are obtained from the European Patent Office (EPO)'s Worldwide Patent Statistical Database (also known as PATSTAT), spring 2020 edition, as available in the STI Micro-data Lab infrastructure. AI-related patents are identified using the AI-related keywords identified by Baruffaldi et al.'s (2020[2]) as well as the IPC and Cooperative Patent Classification (CPC) classes.

Trademarks instead relate to those registered at the following three offices: the EU Intellectual Property Office (EUIPO), the Japan Patent Office (JPO) and the US Patent and Trademark Office (USPTO). AI-related trademarks are identified following (Nakazato and Squicciarini, 2021[17]), who built on Baruffaldi et al.'s (2020[2]) as well as other IP-based analyses.

Although initially more than 3,500 applicants are identified in the United Kingdom, Table A A.1 shows that this number is significantly reduced once different spellings are accounted for and, importantly, individuals as well as universities are dropped as much as possible. This, to some extent, reflects the concentration of innovative activities among relatively few firms.

GlassAI

Information available on company websites can provide relevant indicators of firm activity. So far these have been exploited in a yet limited way, but offer a great potential for understanding the most recent trends in innovation and firm dynamics (see for instance Nathan and Rosso (2015[43]; 2022[44]); Kinne et al. (2019[32]); Ragoussis and Timmis (2022[45]); Dernis et al. (forthcoming[19]); see also Rammer et al. (2022[46]) for a survey).

The analysis uses information provided by GlassAI, a UK software development web scale intelligence company. GlassAI collected and provided experimental data on entities active in the AI space for the OECD by reading and interpreting open web text (e.g., sentences and paragraphs) contained on websites in 2020.

Using Baruffaldi et al.'s (2020[2]) AI-related keywords, GlassAI identified more than 2,000 firms (see Table A A.1) and almost 200 universities active in the AI space in the United Kingdom. Even if not all firms active in the AI space may refer to AI-related keywords on their website, at least most of those that have AI at the core of their business would supposedly do so.

Where available, the information retrieved by GlassAI goes beyond the entities' names and websites and covers their location, registered company number, and sector, among others. While industry and, by definition, name as well as website are available for all firms, the postal address and the company registration number are missing for almost 4% and around 9% of firms, respectively.

Lightcast™

Lightcast™ collects online vacancy data (e.g. employer name, location, occupation, industry, skills, education and experience levels) by web-scraping over 40,000 distinct job boards and company websites in many English-speaking and European countries.

Lightcast™ data have been extensively used in several high-level academic and policy publications, also focusing on the role of AI in firms (Center for Security and Emerging Technology, 2020[47]; Squicciarini and Nachtigall, 2021[24]; Samek, Squicciarini and Cammeraat, 2021[25]; Alekseeva et al., 2021[22]). Indeed, there is ample evidence suggesting that Lightcast™ is well aligned with government survey-based statistics on vacancies and exhibits good statistical properties, making it a useful source of timely information about labour market demand (Carnevale, Jayasundera and Repnikov, 2014[48]; Chetty et al., 2020[49]; Cammeraat and Squicciarini, 2021[50]). Nevertheless, it is noteworthy that vacancies provide information on the augmentation rather than the existing stock of (AI-related) human capital and it is not possible to disentangle whether such augmentation involves the replacement of existing labour and/or its complementation.

In this analysis, AI actors are identified based on AI-related skills (using Baruffaldi et al.'s (2020[2]) AI keywords) that are listed in online job postings.[47] The share of those identified jobs has been increasing consistently and accounted for more than 0.5% in 2020, equalling more than 33 thousand AI-related jobs in the United Kingdom that year. When translating job posting counts into organisation counts that are recruiting for AI-related positions over the period 2012-20 we find around 7,400 organisations. Although this number is considerable, this implicitly disregards a significant proportion of AI-related job postings for which organisation names are not available. Still, most firms advertise multiple vacancies over time and hence it is likely that the AI-active status will be captured by other AI postings for which an employer name is provided.

Once names are cleaned and harmonised, and universities are excluded, a little more than 6,700 AI adopters searching for AI talent remain (see Table A A.1). Although not all firms would post jobs online, high-skilled jobs (such as AI-related ones) have a relatively higher chance to be posted online (see Cammeraat and Squicciarini (2021[50]), for a survey) since AI arguably "lives" and thrives in the digital space.

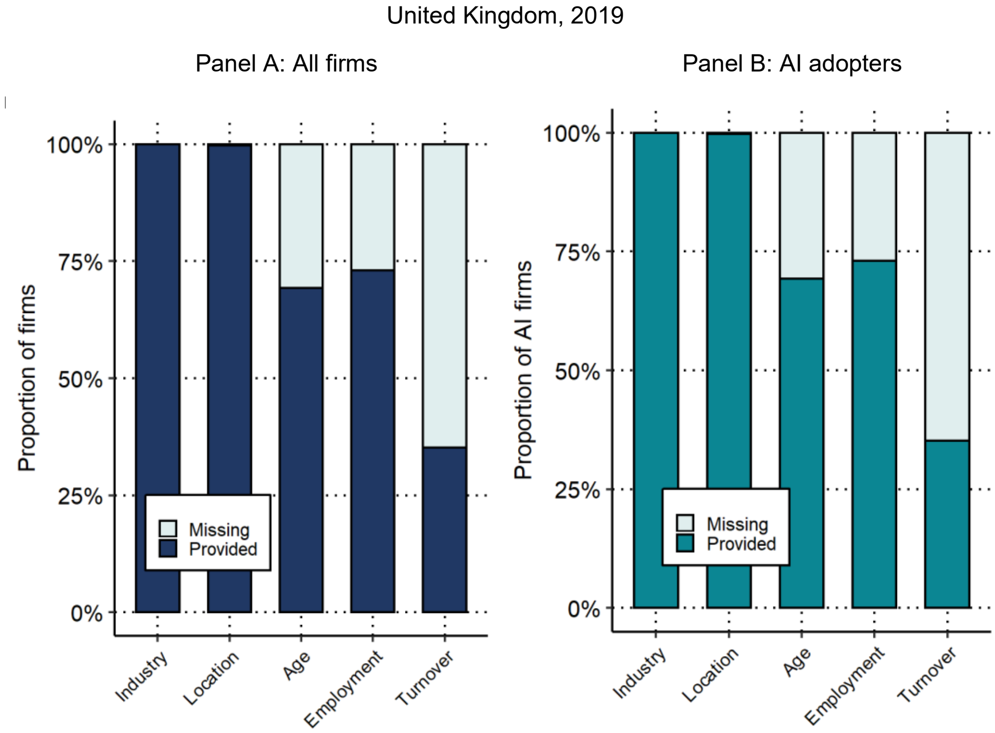
Orbis©

Orbis© by Bureau van Dijk is a well-known commercial database containing firm-level information collected from company accounts. It covers almost 400 million companies and entities worldwide of which around 41 million contain detailed financial information. It is the largest cross-country firm-level database that is available and accessible for economic and financial research (Bajgar et al., 2020[31]). The analysis relies on data provided in the Orbis© 2.2020 vintage.

Bajgar et al. (2020[31]) provide extensive discussion about Orbis© and its features, as well as its representativeness. While Orbis© is known to focus relatively more on larger and more productive firms, Bajgar et al. (2020[31]), based on a previous vintage, show that, when comparing Orbis© with STAN data, its representativeness for the United Kingdom is one of the best across countries, especially when focusing on employment. Still, smaller and younger firms may be to some extent under-represented in the analysis. Moreover, despite the relatively good coverage, not all information about firm characteristics is available in Orbis©.

This can be seen in Figure AA1, which focuses on 2019 – the year used for the characterisation. Coverage is very comprehensive for industry and location (headquarters) information, and very good for age and employment, especially for AI adopters. However, a considerable share of missing values is evident for turnover, which is used in the computation of the productivity proxy in the analysis.

Figure A A.1. Orbis variable coverage



Notes: The figure shows the share of firms for which the respective variable contains information. The left panel shows the share among all firms in Orbis©, while the right panel focuses on AI adopters (identified through the other three sources) only.
Source: Orbis© (version 2.2020, Bureau van Dijk, April 2021).

Matching methodology

The core of the matching has been performed using the Imalinker system (Idener Multi Algorithm Linker developed by IDENER), which relies on country-specific 'dictionaries' to first harmonise organisations' names and to then match those names from two different data sources. The harmonisation process helps dealing with legal entity denomination (e.g. 'Limited' and 'Ltd'), common names and expressions, as well as phonetic and linguistic rules that can affect the way in which names are written. Such harmonisation is based on applicant names for IPRs (patents and trademarks), organisation names for web-reading data (GlassAI), and employer names for online job postings (Lightcast™). Concerning Lightcast™ data, job postings with missing employer name were excluded at this stage.[48]

For financials (Orbis©), company names have been harmonised using the aforementioned country-specific dictionaries, and some basic cleaning has been implemented. In particular, as Orbis© has multiple accounts of different types available for some firms, we followed Bajgar et al. (2020[31]) in removing such duplicate accounts by prioritising unconsolidated ones.

The harmonised names are compared using a series of string-matching algorithms – mainly token-based and string-metric-based, such as token frequency matching and Levenshtein (1966[35]) as well as Jaro-Winkler (1999[36]) distances. Imalinker calculates matching accuracy scores for each pair matched by the different algorithms which can then be used to impose high score thresholds and minimise the number of false positive matches. Whenever possible, websites and company registration numbers are also exploited (in the case of GlassAI, registration numbers are indeed the primary matching criterion) in addition to names.

As Imalinker solely relies on names to link data sources, multiple matches are frequently identified and hence additional selection criteria have to be imposed. Therefore, linked firms are prioritised if their postcodes as well as their sectoral information match, and they have some firm-level information provided in Orbis©. Next, linked pairs among multiple matches are prioritised if either postcodes or sectors match, and Orbis© contains at least some firm-level information.

To confirm and possibly extend the list of identified matches, the exercise is replicated using R studio's *tidystringdist* and *data.table* libraries. First, some basic manual string cleaning is applied to the firm names, such as removing all punctuation, changing all characters to lowercase, and removing words or abbreviations from the names (this is a built-in function in the Imalinker system). Then, a table is created containing all possible combinations of the different firm names from two datasets. Finally, similar to the Imalinker, matching algorithms are applied to calculate the string distances and matches are confirmed based on imposed thresholds while multiple matches are narrowed down utilising postcode and sectoral information.

To identify and sift out universities[49] (from the job postings and the IPRs data) in either of the two matching instruments, a keyword list containing terms, such as "university", "college" and "school" is employed. Universities known to be active in the AI space are also separately searched for and tagged universities are spot-checked.

Individuals applying for patents or trademarks (in the IPR data) are also excluded from the analysis as they cannot be matched to any firm level information. To ensure that these entities are not false negatives, i.e., firms rather than individuals that failed to be matched to Orbis©, the string, i.e., their name, is compared to the typical naming pattern of "surname comma first name". Again, spot checks are carried out to confirm tagged individuals. The final counts of AI adopters identified in each data source are reported below in Table A A.1.

Table A A.1. Counts of identified AI adopters by source of data

Data source	Number of AI adopters
IPRs (Patents and trademarks)	1,178
GlassAI (Company websites)	2,245
Lightcast™ (Online job postings)	6,761

Notes: The table represents counts of AI adopters identified in each data source.
Source: authors' own compilation based on GlassAI (January 2021), Lightcast™ (February 2021), and IPRs (STI Micro-data Lab: IntellectualPropertyDatabase, http://oe.cd/ipstats, July 2020)

Remaining challenges

While the multi-perspective approach proposed provides comprehensive insights into the identification of AI actors and their characteristics, some challenges still remain, which mainly stem from data availability and from the characteristics of the databases we rely upon.

In particular, even after matching AI adopters with financials, not all information about firm characteristics is available in Orbis©. Relevantly, a considerable share of missing values is evident for turnover (and hence for the productivity proxy used in the characterisation, see Figure AA1).

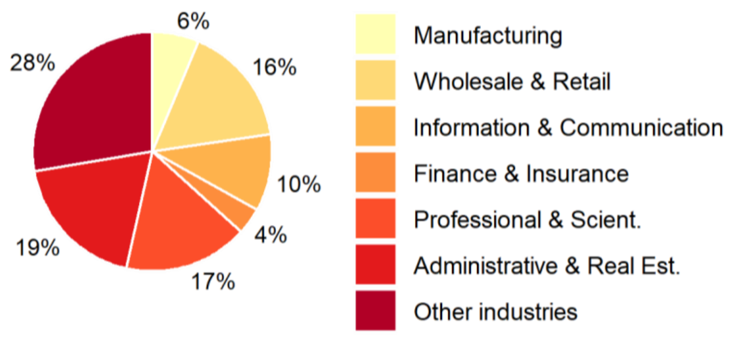
Furthermore, there are slight timing mismatches among the different data sources, due to data features and availability constraints. In particular, online vacancy data used refer to the period 2012-20, web scraping of internet websites was carried out in 2020, and IPRs data are available for longer time periods but only until 2018, while the last available year in Orbis©, for which the characterisation is carried out, is 2019.

Finally, it is clear that – despite all the efforts – it is not possible to claim that this methodology allows to identify each and every single AI adopter in the United Kingdom. In fact, not every single AI adopter may have applied for AI-related IPRs, posted AI-related jobs online, or mentioned AI-related keywords on its website. The set of AI adopters identified is however considerable and the proposed methodology allows to significantly extend the scope of analysis beyond existing work based on fewer data sources.

Annex B. Additional tables and figures

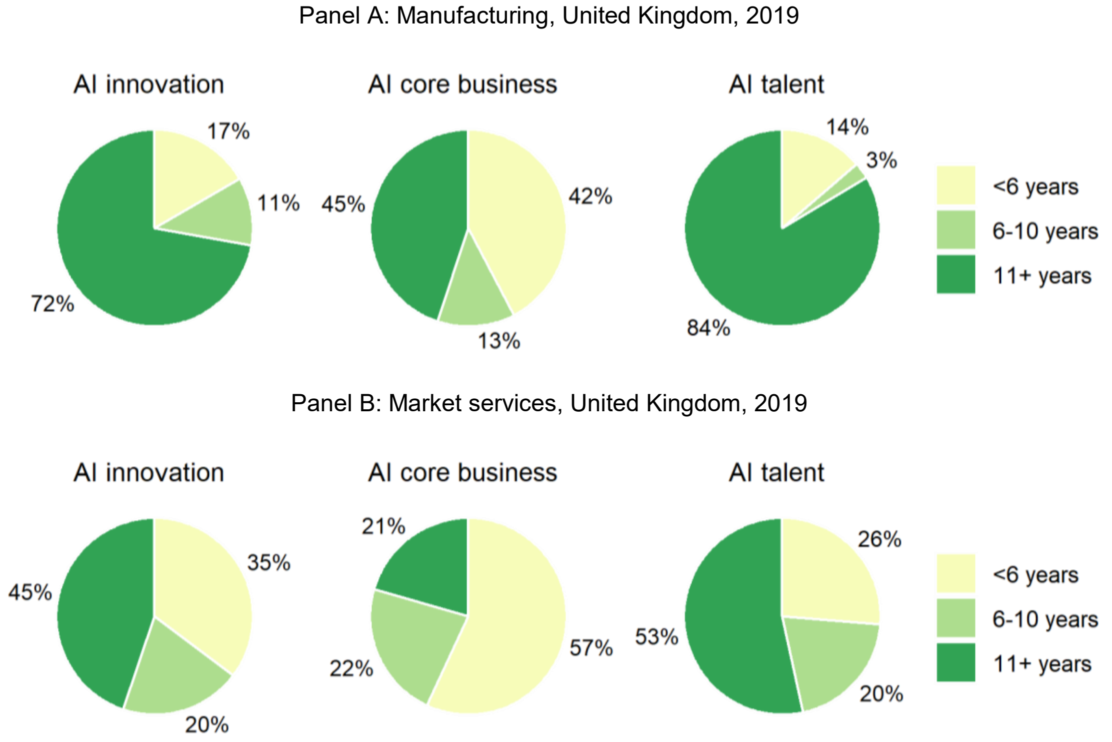
Figure A B.1. Firms in Orbis by broad sector of activity

United Kingdom, 2019



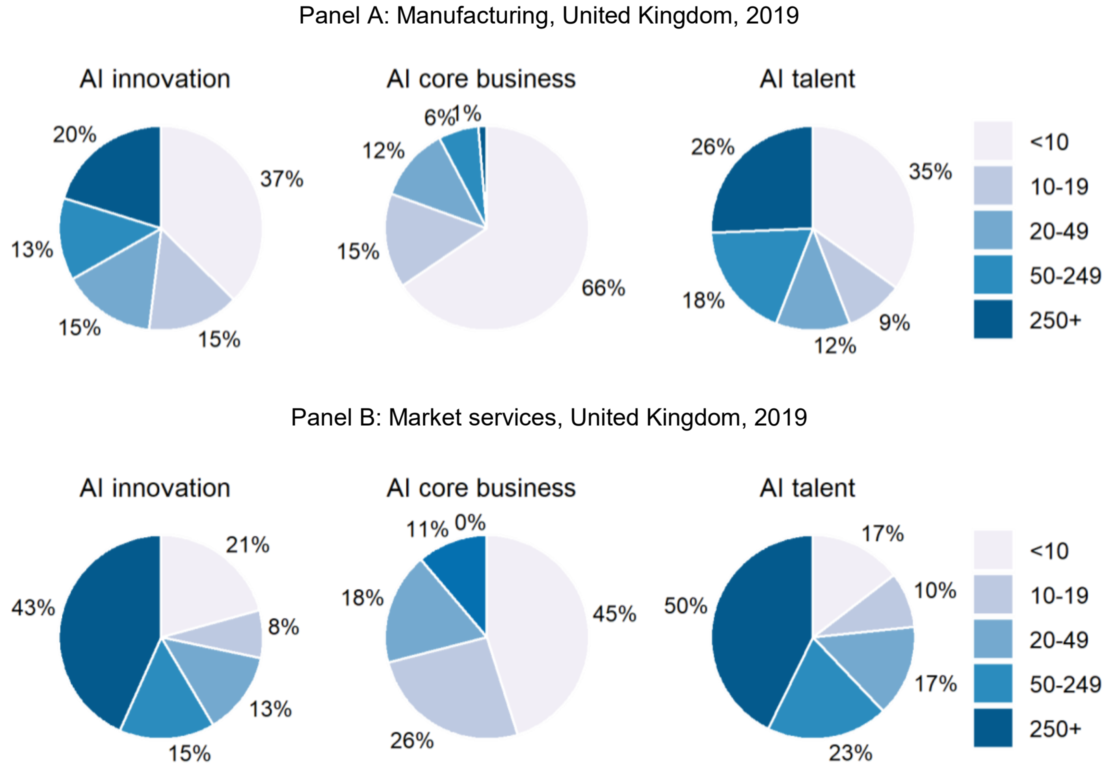
Notes: NACE classifications are constructed using the NACE2 (4 digit level) classification reported in Orbis© for the year 2019. The category "other industries" includes Mining and Quarrying, Utilities, Transport and Storage, Accommodation and Food, and Construction.
Sources: Authors' own compilation based on Orbis© (version 2.2020, Bureau van Dijk, April 2021).

Figure A B.2. Shares of AI adopters by firm age class and macro sector



Notes: Firm age is calculated as the difference between 2019 and the incorporation date as reported in Orbis© for the year 2019.
Sources: Authors' own compilation based on GlassAI (January 2021), Lightcast™ (February 2021), IPRs (STI Micro-data Lab: IntellectualPropertyDatabase, http://oe.cd/ipstats, July 2020) and Orbis© (version 2.2020, Bureau van Dijk, April 2021).

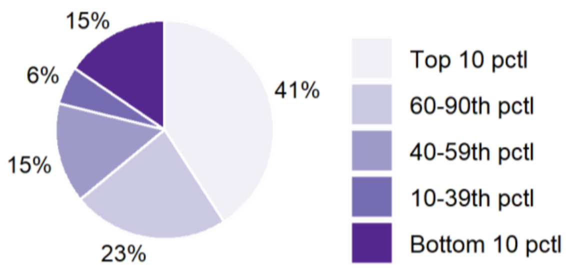
Figure A B.3. Shares of AI adopters by size class and macro sector



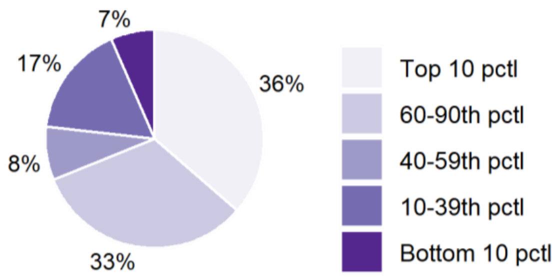
Notes: Size classes are constructed using the number of employees as reported in Orbis© for the year 2019.
Sources: Authors' own compilation based on GlassAI (January 2021), Lightcast™ (February 2021), IPRs (STI Micro-data Lab: IntellectualPropertyDatabase, http://oe.cd/ipstats, July 2020) and Orbis© (version 2.2020, Bureau van Dijk, April 2021).

Figure A B.4. Shares of all AI adopters by percentiles of the productivity distribution and macro sector

Panel A: Manufacturing, United Kingdom, 2019



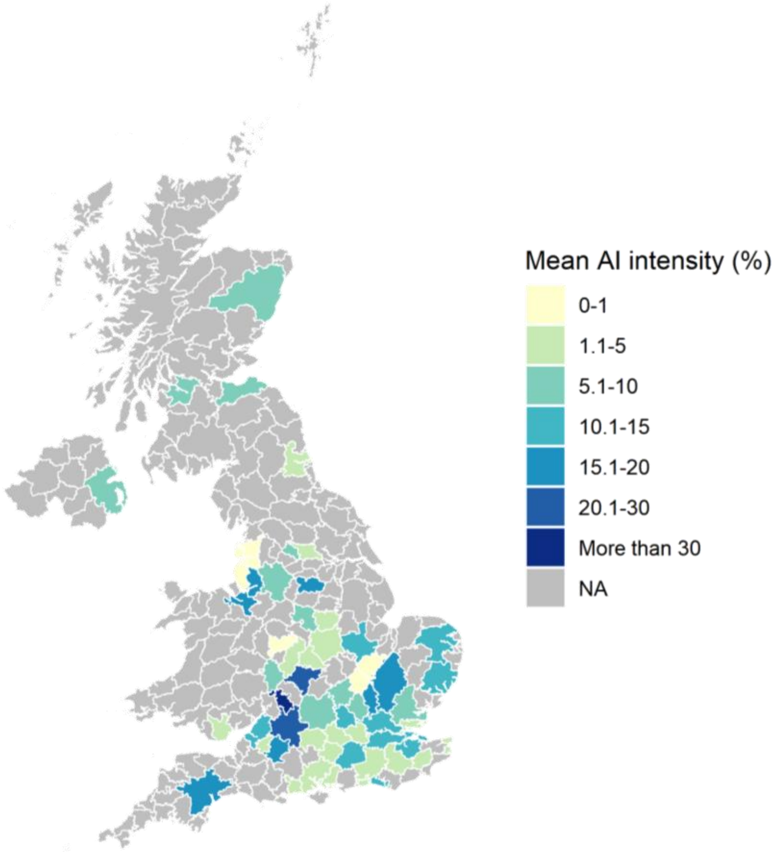
Panel B: Market services, United Kingdom, 2019



Notes: Productivity classes are constructed within SNA A38 sectors using a proxy of labour productivity (operating revenue turnover over number of employees). All data is collected from Orbis© for the year 2019 and implausible turnover or employee counts are dropped prior to calculating productivity.
Sources: Authors' own compilation based on GlassAI (January 2021), Lightcast™ (February 2021), IPRs (STI Micro-data Lab: IntellectualPropertyDatabase, http://oe.cd/ipstats, July 2020) and Orbis© (version 2.2020, Bureau van Dijk, April 2021).

Figure A B.5. Average AI-hiring intensity by TTWA

United Kingdom, 2019



Notes: Location refers to the TTWA of headquarters, as reported in Orbis©. Sample only contains AI adopters identified in Lightcast™ that were successfully linked to Orbis©. AI intensity is measured as the share of AI job postings out of all vacancies, on average for firms in each sector. AI intensity is assumed to be zero for the years in which firms do not advertise any AI jobs that is captured in Lightcast™.
Sources: Authors' own compilation based Lightcast™ (February 2021) and Orbis© (version 2.2020, Bureau van Dijk, April 2021).

Table A B.1. AI adoption, firm characteristics and AI universities

United Kingdom, 2019

	AI adoption dummy
Firm age class 6-10 years	-0.00004
	(0.00009)
Firm age class 11 or more years	0.00031***
	(0.00009)
Employment size class 10-49 employees	0.00356***
	(0.00018)
Employment size class 50-249 employees	0.01143***
	(0.00067)
Employment size class 250 or more employees	0.05754***
	(0.00274)
AI university within same TTWA	0.00036***
	(0.00005)
Observations	1,923,222
Industry dummies	YES
Regional dummies	YES
Adjusted R-squared	0.014

Notes: Coefficients for the constant and missing firm age and employment size classes are not reported. Robust standard errors in parentheses (*** p<0.01, ** p<0.05, * p<0.1). Baseline category for firm age is <6 years and for employment size is <10 employees.
Source: authors' own compilation based on GlassAI (January 2021), Lightcast™ (February 2021), IPR (STI Micro-data Lab: IntellectualPropertyDatabase, http://oe.cd/ipstats, July 2020) and Orbis© (version 2.2020, Bureau van Dijk, April 2021).

Table A B.2. AI adoption and labour productivity

United Kingdom, 2019

	All sectors			Manufacturing			Market services		
	(1) All firms	(2) >49 empl.	(3) >249 empl.	(4) All firms	(5) >49 empl.	(6) >249 empl.	(7) All firms	(8) >49 empl.	(9) >249 empl.
AI adoption	0.034	0.197**	0.266***	0.068	0.152*	0.088	0.091	0.201**	0.323***
	(0.071)	(0.076)	(0.069)	(0.113)	(0.075)	(0.132)	(0.080)	(0.095)	(0.086)
Firm age	0.004***	0.002***	0.003***	0.002***	0.001*	-0.000	0.005***	0.003**	0.004**
	(0.001)	(0.001)	(0.001)	(0.000)	(0.000)	(0.001)	(0.002)	(0.001)	(0.001)
Lagged log number of empl.	0.102***	-0.070***	-0.040***	0.042	-0.015	-0.032	0.095***	-0.083***	-0.041**
	(0.022)	(0.014)	(0.014)	(0.025)	(0.026)	(0.039)	(0.026)	(0.017)	(0.017)
Observations	39,385	16,495	4,984	5,507	3,696	963	29,211	11,087	3,564
Industry dummies	YES	YES	YES	YES	YES	YES	YES	YES	YES
Regional dummies	YES	YES	YES	YES	YES	YES	YES	YES	YES
R-squared	0.136	0.200	0.261	0.080	0.105	0.169	0.150	0.218	0.277

Notes: Robust standard errors in parentheses (*** p<0.01, ** p<0.05, * p<0.1). The constant has been estimated but is not reported.
Source: Authors' own compilation based on GlassAI (January 2021), Lightcast™ (February 2021), IPR (STI Micro-data Lab: IntellectualPropertyDatabase, http://oe.cd/ipstats, July 2020) and Orbis© (version 2.2020, Bureau van Dijk, April 2021).

Table A B.3. Type of AI adoption and labour productivity

United Kingdom, 2019

	(1) All firms	(2) >49 empl.	(3) >249 empl.
AI talent	0.060	0.254***	0.277***
	(0.081)	(0.077)	(0.074)
AI innovation	0.073	0.213*	0.304***
	(0.118)	(0.122)	(0.109)
AI core business	-0.219**	-0.444***	-0.339
	(0.097)	(0.080)	(0.240)
Firm age	0.004***	0.002***	0.003***
	(0.001)	(0.001)	(0.001)
Lagged log number of empl.	0.102***	-0.071***	-0.041***
	(0.022)	(0.014)	(0.014)
Observations	39,385	16,495	4,984
Industry dummies	YES	YES	YES
Regional dummies	YES	YES	YES
Adjusted R-squared	0.138	0.205	0.274

Notes: Robust standard errors in parentheses (*** p<0.01, ** p<0.05, * p<0.1). Baseline category for firm age is <6 years and for employment size <10 employees. The constant has been estimated but is not reported.
Source: authors' own compilation based on GlassAI (January 2021), Lightcast™ (February 2021), IPR (STI Micro-data Lab: IntellectualPropertyDatabase, http://oe.cd/ipstats, July 2020) and Orbis© (version 2.2020, Bureau van Dijk, April 2021).

Table A B.4. AI-hiring intensity and labour productivity

United Kingdom, 2019

	Overall AI-hiring intensity			Sub-components of AI-hiring intensity		
	(1) All firms	(2) >49 empl.	(3) >249 empl.	(4) All firms	(5) >49 empl.	(6) >249 empl.
Overall AI intensity	0.004	0.011**	0.036*			
	(0.007)	(0.004)	(0.019)			
AI intensity: Managers				0.018***	0.018**	0.008*
				(0.006)	(0.008)	(0.005)
AI intensity: Professionals				0.000	0.007**	0.022**
				(0.005)	(0.003)	(0.010)
AI intensity: Technicians				-0.000	-0.001	-0.015***
				(0.004)	(0.003)	(0.003)
AI intensity: Other occupations				0.007*	0.002	0.036*
				(0.004)	(0.008)	(0.018)
Firm age	0.004***	0.002***	0.003***	0.004***	0.002***	0.003***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Lagged log number of empl.	0.103***	-0.068***	-0.035***	0.103***	-0.069***	-0.036***
	(0.022)	(0.013)	(0.013)	(0.022)	(0.013)	(0.013)
Observations	39,191	16,356	4,914	39,191	16,356	4,914
Industry dummies	YES	YES	YES	YES	YES	YES
Regional dummies	YES	YES	YES	YES	YES	YES
Adjusted R-squared	0.138	0.205	0.274	0.138	0.205	0.275

Notes: AI intensity is measured as the share of AI job postings out of all vacancies, on average for firms in each sector. AI intensity is assumed to be zero for the years in which firms do not advertise any AI jobs, or for firms that are tagged as non-adopters. Firms tagged as AI-adopters based on web-reading or IPRs, but for which no information on AI-hiring intensity is available, are dropped from this analysis. Robust standard errors in parentheses (*** p<0.01, ** p<0.05, * p<0.1). Baseline category for firm age is <6 years and for employment size <10 employees. The constant has been estimated but is not reported.
Source: Authors' own compilation based on GlassAI (January 2021), Lightcast™ (February 2021), IPR (STI Micro-data Lab: IntellectualPropertyDatabase, http://oe.cd/ipstats, July 2020) and Orbis© (version 2.2020, Bureau van Dijk, April 2021).

Endnotes

[1] A relevant exception is a recent evidence analysis and primary market research to assess the extent of data foundations and AI adoption conducted by Ernst & Young, which were appointed by the governmental Department for Digital, Culture, Media and Sport (https://www.gov.uk/government/publications/data-foundations-and-ai-adoption-in-the-uk-private-and-third-sectors). The analysis is based on responses from about 400 organisations.

[2] In fact, AI was already at the core of the 2017 UK Industrial Strategy. In 2021, the United Kingdom published its National AI Strategy highlighting objectives related to investing in the long-term needs of the AI ecosystem, supporting the transition to an AI-enabled economy, and regulating AI technologies (HM Government, 2021[61]).

[3] The information on jobs is solely used for the purpose of identifying AI actors; investigating the labour market implications of AI adoption is beyond the scope of this paper.

[4] In this paper, the term “adopters” (and relatedly AI adopters) is used in a broad sense, in reference to the different groups of firms outlined in this paragraph, based on the data sources described in section 3.

[5] These include Dernis et al. (2021[18]); Samek et al. (2021[25]); Squicciarini and Nachtigall (2021[24]); Nakazato and Squicciarini (2021[17]).

[6] *AI diffuse* is a novel distributed microdata project aimed at studying the characteristics of adopters of digital technologies, notably including AI, the role of complementary assets, and the links between technology use and productivity, based on ICT surveys. The analysis is also complementary to the ongoing OECD-BCG-Insead survey on AI in business.

[7] The ongoing debate discusses more and more AI’s role in changing labour and skills demand (Bessen, 2019[51]; Samek, Squicciarini and Cammeraat, 2021[25]; Georgieff and Hyee, 2021[56]; Acemoglu et al., 2022[34]), its impact on facilitating prediction and innovation (Agrawal, Gans and Goldfarb, 2018[53]; Cockburn, Henderson and Stern, 2018[52]; Gierten et al., 2021[58]), and its potential to tackle societal challenges such as climate change. These issues are beyond the scope of the present analysis.

[8] A related part of the literature has instead employed previous waves of firm-level surveys that contain information on the use of big data or predictive analytics, particularly focusing on their relation with firm productivity (see for instance Brynjolfsson, Jin and McElheran (2021[55]) for the United States or Borowiecki et al. (2021[57]) for the Netherlands).

[9] The full list of keywords in Baruffaldi et al. (2020[2]), which previous OECD work and this study built upon, is available in their paper’s Annex Table C.1.

[10] Focusing instead on a broad set of technologies, Bloom et al. (2021[27]) identify and document the patterns of diffusion of 29 disruptive technologies in the United States based on textual analysis of patents, job postings, and earning calls. They document the initial geographical concentration of disruptive technologies, which then spread across space over time, and the concentration of initial hiring among higher-skilled jobs, with a gradual broadening of the types of jobs adopting a given technology. They show that pioneer locations have long-lasting benefits and those appear more likely to be located around areas with universities and high-skilled labour.

[11] Baruffaldi et al. (2020[2]) find that in 2017, AI-related patents made up more than 2.3% of the total share of IP5 patent families, twice the level observed in 2010 (~1.1%).

[12] This is measured by the hiring of an increasing share of AI-skills of the total number of skills the firm demands through these job postings.

[13] Existing studies to our knowledge do not use copyright data, which may be an additional relevant source of information.

[14] See also the discussion in the previous section and Annex Table C.1. in Baruffaldi et al. (2020[2]) for additional information.

[15] Patents are only one mechanism to protect innovation, and in general tend to be skewed towards manufacturing. Furthermore, the patentability of software-based technologies, and hence AI technologies, varies across countries. For instance, in the United States, software can be patented as such, while at the European Patent Office (EPO) it is only patentable if embedded in computer implemented inventions. See Dernis et al. (2021[18]) for further details.

[16] Job postings data proxy the demand for AI talent, as posted online and captured by the data provider, beyond initial stocks. Some relevant challenges to keep in mind when using this data source are further discussed in Annex A.

[17] Table A A.1 in Annex A shows the number of AI adopters identified based on each data source.

[18] Groups have been defined as mutually exclusive in order to ease the presentation of results and discussion. This also reflects the high complementarity of the different data sources, further discussed in the following section.

[19] In particular, Orbis© data tend to have better coverage for larger and more productive firms, possibly inducing selection bias. This is further discussed in the following section and in Annex A.

[20] The core of the matching has been performed using the Imalinker system (Idener Multi Algorithm Linker developed by IDENER), which has been complemented by additional iterations using R studio’s *tidystringdist* library.

[21] The fact that a larger number of AI adopters are identified through the AI talent group and that the proportion of AI jobs among all job postings in Lightcast™ has increased significantly over time (Samek, Squicciarini and Cammeraat, 2021[25]) may indicate that firms prepare and search for the necessary human capital to enter the AI space.

[22] Future work in the context of the OECD *AI diffuse* project may allow further comparison between the current findings and those based on data available in official ICT surveys.

Those data are more suitable for focusing on overall adoption rates, which are not the main focus of the current analysis.

[23] This implies that not all firms identified as AI adopters in the previous section will contribute to such characterisation. For instance, a firm posting AI-related jobs online in one year, which was linked to Orbis© but exited the market prior to 2019, does not contribute to the figures reported. This approach limits possible double-counting issues, maximises the accuracy of the information used for characterising adopters, and the accuracy of the comparison across different groups of firms.

[24] The distributions of employment, age and turnover are right-skewed for all three groups, with firms looking for AI talent exhibiting the highest median employee count, age and turnover.

[25] This is evident when comparing Figure 4.1 with Figure A B.1 in Annex B, which shows the sectoral distribution of firms in the overall economy when using Orbis©. Findings are also similar when relying on the OECD's Structural and Demographic Business Statistics.

[26] Due to diffuse commuting patterns, it is not possible to divide the United Kingdom into entirely self-contained labour market areas (where all commuting occurs within the boundary of that area). Therefore, TTWAs have been constructed, so that of the resident economically active population, "at least 75% of an area's resident workforce work in the area and at least 75% of the people who work in the area also live in the area" (ONS, 2016[62]). In other words, it is developed so that relatively few commuters cross a TTWA boundary on their way to work.

[27] Findings also remain qualitatively similar when showing the number of AI adopters as a share of all firms by TTWA. These are not reported for brevity but the figures are available upon request from the authors.

[28] Keeping in mind that Orbis© is not ideal to study the full support of age and size distributions, these figures identify a considerable number of small and young AI adopters. Firm size and age distributions are however typically more skewed (see Criscuolo, Gal and Menon (2014[60]); Calvino et al. (2021[54])), with e.g., more than 75% of micro (<10 employees) firms in most OECD countries. A tentative comparison in terms of age and size with likely non-adopters is carried out in one of the following subsections.

[29] Additional unreported analysis also suggests that – especially among those two groups – being part of a foreign-owned multinational is positively associated with AI adoption. Within those groups multinationals represent between 25% and 28% of AI adopters.

[30] Additional unreported analysis focusing on age and size heterogeneity among different groups of adopters further controlling for sectoral composition also qualitatively confirms the main insights discussed above. Focusing on AI adopters in digital-intensive sectors (as defined by the top quartile of the taxonomy proposed by Calvino et al. (2018[59])) also provides results that are in line with Figure 4.6 (Panel A and B).

[31] Around 4% and 14% of those AI adopters in the top 10% and the 60th-90th percentiles of the productivity distribution, respectively are also top R&D investors (see Amoroso et al. (2021[40]) for details). Top R&D investors active in the AI space in the United Kingdom operate mainly in manufacturing, and are large as well as mature firms.

[32] This corresponds to the whole "AI talent" group of AI adopters, plus the firms in the groups "AI innovation" and "AI core business" for which AI-related job postings were identified during the observed period.

[33] The measure is calculated at the firm level. Figures presented in this subsection refer to AI hiring intensity in 2019, for consistency with the rest of the paper. Additional unreported analysis confirms that increasing the length of the time window considered, to, for instance, four years from 2016 to 2019, provides qualitatively similar patterns.

[34] Unreported analysis shows that AI professionals are also more AI skills-intensive compared to managers and technicians, as reflected by the largest proportion of AI skills in their job postings.

[35] This indeed appears to be particularly the case when looking at the shares of AI adopters in different size bins, which are especially low for small- and medium-sized firms.

[36] This is also broadly confirmed when comparing whole age, employment, and turnover distributions, with the first quartiles, medians, and third quartiles being higher for AI adopters than for other firms.

[37] Unreported analysis focusing on different groups of AI adopters suggests that age coefficients' patterns for "AI innovation" and "AI talent" are in line with those reported in Table 4.3, while firms in the group "AI core business" exhibit small negative coefficients for older age bands, in line with evidence reported in the previous subsections.

[38] This is true across the different groups of AI adopters, and after controlling for (broad) regional heterogeneity.

[39] Those are not reported for the sake of brevity, but are available upon request from the authors.

[40] The different size groups for which Equation 2 is estimated are: all firms, firms with more than 49 employees, and firms with more than 249 employees.

[41] Additional unreported analysis suggests that, when considering firms with more than 49 employees, these premia tend to be highest among foreign-owned AI adopters.

[42] Suggestive evidence indeed seems to confirm that this may be the case since, when adding lagged productivity as additional explanatory variable in Equation 2, productivity premia tend to disappear. This is in line with ongoing analysis based on ICT surveys (e.g., Calvino and Fontanelli (2022[10])).

[43] Additional analysis has focused on average wages rather than labour productivity, in the same framework proposed in Equation 2. Results suggest that AI adopters tend to pay higher wages, especially largest ones, both in manufacturing and services. Wage data are however often missing in Orbis©. Results are exploratory and available upon request from the authors.

[44] A positive productivity coefficient is also evident for innovators in manufacturing, when considering all firms, possibly suggesting that those innovators may be significantly different from other firms, even within the more selected subsample of smaller firms in Orbis©.

[45] Firms tagged as non-adopters have therefore zero AI-hiring intensity. Firms tagged as AI-adopters based on web-reading or IPRs, but for which no information on AI-hiring intensity is available, are dropped from this analysis.

[46] Unreported regressions focusing on manufacturing and market services separately show that the positive coefficient for the overall AI intensity is driven by market services, while the positive coefficient for managers is evident in both macro sectors.

[47] The number of false positives, i.e. jobs and hence firms that are falsely identified as AI actors, could be a concern especially among the AI talent group if employers advertise non-AI jobs using AI-related keywords, e.g., to enhance the vacancy's appeal. However, by following Squicciarini and Nachtigall (2021[24]), we employ a conservative approach that significantly reduces such risk of false positives. For more details and sensitivity checks, please see Squicciarini and Nachtigall (2021[24]).

[48] Although this share is considerable (about half of all AI-related UK job postings), most firms tend to advertise multiple vacancies and hence it is likely that at least some of the vacancies with missing employer name are already captured by other AI-related jobs for which an employer name is provided.

[49] Information on AI-related universities from GlassAI is further exploited in the analysis to explore the role of tertiary education for AI adoption by firms, which are at the centre of the current analysis and for which data coverage in Orbis© is significantly more comprehensive.