OECD publishing

WHO DEVELOPS AI-RELATED INNOVATIONS, GOODS AND SERVICES?

A FIRM-LEVEL ANALYSIS

OECD SCIENCE, TECHNOLOGY AND INDUSTRY
POLICY PAPERS
September 2021 **No. 121**

Federal Ministry
of Labour and Social Affairs

OECD
BETTER POLICIES FOR BETTER LIVES

OECD Science, Technology and Innovation Policy Papers

This paper was approved and declassified by written procedure by the Committee for Industry, Innovation and Entrepreneurship (CIIE) on 16 June 2021 and prepared for publication by the OECD Secretariat.

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/.*

Note to Delegations:

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

DSTI/CIIE/WPIA(2021)5/FINAL

Who develops AI-related innovations, goods and services? A firm level analysis

Hélène Dernis, Laurent Moussiegt, Daisuke Nawa, and Mariagrazia Squicciarini

(OECD Directorate for Science, Technology and Innovation)

This study proposes an exploratory analysis of the characteristics of Artificial Intelligence (AI) "actors". It focuses on entities that deploy AI-related technologies or introduce AI-related goods and services on large international markets. It builds on the OECD Science, Technology and Innovation Micro-data Lab infrastructure, and, in particular, on Intellectual Property (IP) rights data (patents and trademarks) combined with company-level data. Statistics on AI-related patents and trademarks show that AI-related activities are strongly concentrated in some countries, sectors, and actors. Development of AI technologies and/or goods and services is mainly due to start-ups or large incumbents, located in the United States, Japan, Korea, or the People's Republic of China, and, to a lesser extent, in Europe. A majority of these actors operate in ICT-related sectors. The composition of the IP portfolio of the AI actors indicates that AI is frequently combined with a variety of sector-specific technologies, goods, or services.

Synthèse

Cette étude propose une analyse exploratoire des caractéristiques des « acteurs » de l'Intelligence Artificielle (IA). Elle se concentre sur les entités qui déploient des technologies liées à l'IA ou introduisent des biens et services liés à l'IA sur les grands marchés internationaux. Elle s'appuie sur l'infrastructure du Microdata Lab de la Direction de la science, de la technologie et de l'innovation de l'OCDE, et, en particulier, sur les données sur les droits de propriété intellectuelle (brevets et marques) combinées à des données au niveau des entreprises. Les statistiques sur les brevets et marques liés à l'IA montrent que les activités relatives à l'IA se concentrent fortement dans quelques pays, secteurs et acteurs. Le développement de technologies et/ou de biens et services d'IA est principalement dû à des start-ups ou de grands opérateurs historiques, situés soit aux États-Unis, au Japon, en Corée ou en République populaire de Chine et, dans une moindre mesure, en Europe. La majorité de ces acteurs opèrent dans des secteurs liés aux TIC. La composition du portefeuille de propriété intellectuelle des acteurs de l'IA montre que l'IA est fréquemment associée à une variété de technologies, de biens ou de services qui dépendent du secteur.

Kurzfassung

Diese Studie schlägt eine explorative Analyse der Eigenschaften von "Akteuren" der Künstlichen Intelligenz (KI) vor. Sie konzentriert sich auf Unternehmen, die KI-bezogene Technologien einsetzen oder KI-bezogene Waren und Dienstleistungen auf großen internationalen Märkten einführen. Sie benutzt die Infrastruktur des OECD Science, Technology and Innovation Micro-data Labs und stützt sich insbesondere auf Daten zu geistigen Eigentumsrechten (Patente und Marken) in Kombination mit Daten auf Unternehmensebene. Statistiken zu KI-bezogenen Patenten und Marken zeigen, dass KI-bezogene Aktivitäten in einigen Ländern, Sektoren und bei einigen Akteuren stark konzentriert sind. Die Entwicklung von KI-Technologien bzw. -Gütern und -Dienstleistungen findet man hauptsächlich bei Start-ups oder großen etablierten Unternehmen, die in den USA, Japan, Korea oder der Volksrepublik China und in geringerem Maße auch in Europa ansässig sind. Ein Großteil dieser Akteure ist in Sektoren der Informations-und Kommunikationstechnologie tätig. Die Zusammensetzung des Portfolios „geistiges Eigentum" der KI-Akteure zeigt, dass KI häufig mit einer Vielzahl von branchenspezifischen Technologien, Waren oder Dienstleistungen kombiniert wird.

Executive summary

This paper analyses the characteristics of economic agents that have filed for Intellectual Property (IP) rights to protect technologies related to Artificial Intelligence (AI) -through patents- or AI-related goods and services -through registered trademarks-, for the period 2014-18. It builds on recent experimental approaches to identify and characterise AI-related innovations (Baruffaldi et al., 2020[1]), as well as AI-related goods and services (Nakazato and Squicciarini, 2021[2]).

The work particularly aims to inform the policy discussion on the diffusion and adoption of AI. Gathering evidence about the way AI develops and permeates economies and societies is key, as AI has the potential to profoundly shape the way firms produce, perform and interact with suppliers, customers and other stakeholders.

The main findings of the analysis are as follows:

- The number of AI-related patents and trademarks filed worldwide increased importantly from the mid-2010s. During that period, the United States contributed the most AI-related patents and trademarks, followed by Japan, Korea and the People's Republic of China (hereafter "China").
- AI-related developments are highly concentrated in firms: the five companies with the largest IP portfolios related to AI collectively own 14% of the total IP portfolio related to AI, mainly with patents. The leading AI companies also belong to the world top R&D investors.
- Overall, AI actors rely more heavily on patents, in order to protect new AI-inventions internationally, than on trademarks, in order to protect AI-products. AI patents and AI trademarks are seldom bundled in the IP portfolios analysed.
- AI actors are polarised between large incumbents on one side and start-ups on the other side. Young and small entities tend to rely relatively more on AI trademarks, to signal AI-related activities to customers and competitors, and to position themselves onto the market.
- ICT-related sectors, notably "Computers & electronics" and "IT services", are responsible for the largest portfolios of IP to protect AI, while AI developments occur at a lower level across a variety of other sectors.
- Differences emerge in the way entities operating in certain economic sectors combine AI patents and trademarks in their IP portfolio. Companies in the services sector tend to have a larger number of trademarks in AI than the average of all sectors.
- AI is associated with application fields that depend on the sector in which the AI agents operate. Details of the composition of the IP portfolio of AI actors by industry show that agents are developing and protecting a broad range of technologies, goods and services in addition to AI.
- Overall, those technologies encompass ICT-related technologies, such as *Computer technology*, *Digital communication*, as well as *Electrical machinery*, *Semiconductors*, *Transport* and *Optics*, to name a few. In turn, other non-AI trademarks are highly connected to *ICT and audio-visual* goods and services.

Résumé

Ce document analyse les caractéristiques des agents économiques ayant déposé des droits de propriété intellectuelle (PI) pour protéger soit des technologies liées à l'intelligence artificielle (IA), par des brevets, soit des biens et services liés à l'IA, par des marques déposées, au cours de la période 2014-18. Il s'appuie sur de récentes approches expérimentales cherchant à identifier et caractériser des innovations en matière d'IA (Baruffaldi et al., 2020[1]), et de biens et services liés à l'IA (Nakazato and Squicciarini, 2021[2])

Le travail vise à éclairer le débat politique sur la diffusion et l'adoption de l'IA. L'IA va profondément façonner les processus de production des entreprises, leur fonctionnement, et leurs interactions avec les fournisseurs, clients et autres parties prenantes. Il est alors essentiel de collecter des preuves sur la manière dont l'IA se développe et imprègne les économies et la société.

Les principaux résultats de l'analyse sont les suivants :

- Les dépôts de brevets et marques liés à l'IA ont considérablement augmenté depuis le milieu des années 2010. Au cours de cette période, les États-Unis ont contribué le plus aux brevets et marques sur l'IA, suivis du Japon, de la Corée et de la République populaire de Chine (ci-après dénommée "Chine").
- Les développements liés à l'IA sont fortement concentrés au sein des entreprises : les cinq sociétés avec les plus gros portefeuilles de droits de PI liés à l'IA représentent 14 % du total des PI lié à l'IA, constitué principalement de brevets. Les principales sociétés actives dans l'IA comptent parmi les principaux investisseurs mondiaux en R&D.
- Dans l'ensemble, les acteurs de l'IA utilisent davantage les brevets, afin de protéger de nouvelles inventions en IA sur les marchés internationaux, que les marques, pour protéger les produits et services liés à l'IA. Les brevets et marques sur l'IA coexistent rarement au sein des portefeuilles de PI analysés.
- Les acteurs de l'IA se répartissent entre grandes entreprises d'un côté et start-ups de l'autre. Les jeunes et petites entreprises s'appuient davantage sur les marques, pour signaler des activités liées à l'IA à leurs clients et concurrents, et se positionner sur le marché.
- Les secteurs liés aux TIC, notamment « Informatique et électronique » et « Services informatiques », sont responsables des plus gros portefeuilles de droits de PI pour protéger de l'IA. Les développements en l'IA dans d'autres secteurs sont moindres.
- Les entités opérant dans certains secteurs économiques combinent inégalement brevets et marques liés à l'IA dans leur portefeuille de PI. Les entreprises du secteur des services semblent déposer un plus grand nombre de marques en IA que la moyenne de tous les secteurs.
- Le secteur dans lequel les agents opèrent conditionne les domaines d'application de l'IA. La décomposition du portefeuille de PI des secteurs indique que les agents développent et protègent un large éventail de technologies, de biens et de services en plus de l'IA.

- Dans l'ensemble, ces technologies couvrent des domaines des TIC, notamment l'*informatique*, la *communication numérique*, mais aussi les *machines électriques*, les *semi-conducteurs*, le *transport* ou l'*optique*. Par ailleurs, les marques hors IA se rapportent à des biens et services liés aux *TIC* et à l'*audiovisuel*.

Zusammenfassung

Diese Studie analysiert für den Zeitraum 2014-18 Eigenschaften von Wirtschaftsakteuren, die Rechte an geistigem Eigentum (intellectual property, IP) angemeldet haben, um Technologien im Zusammenhang mit Künstlicher Intelligenz (KI) durch Patente, oder KI-bezogene Waren und Dienstleistungen durch eingetragene Marken zu schützen. Sie beruht auf neuesten experimentellen Ansätzen zur Identifizierung und Charakterisierung von KI-bezogenen Innovationen (Baruffaldi et al., 2020[1]) sowie von KI-bezogenen Waren und Dienstleistungen (Nakazato and Squicciarini, 2021[2]).

Besonderes Ziel der Arbeit ist es, durch Information die politische Diskussion über die Verbreitung und Annahme von KI anzuregen. Das Sammeln von Erkenntnissen, wie KI sich entwickelt und Wirtschaft und Gesellschaft durchdringt, ist von zentraler Bedeutung; KI hat nämlich das Potenzial, die Art und Weise, wie Unternehmen produzieren, Leistungen erbringen und mit Zulieferern, Kunden und anderen Interessenvertretern interagieren, tiefgreifend zu verändern.

Die wichtigsten Ergebnisse der Analyse sind:

- Die Zahl der weltweit angemeldeten KI-bezogenen Patente und Marken ist seit Mitte der 2010er Jahre stark angestiegen. In diesem Zeitraum haben die Vereinigten Staaten die meisten KI-bezogenen Patente und Marken angemeldet, gefolgt von Japan, Korea und der Volksrepublik China (nachstehend “China”).
- KI-bezogene Entwicklungen sind in Unternehmen unterschiedlich stark konzentriert: Die fünf Unternehmen mit den größten IP-Portfolios im Zusammenhang mit KI besitzen zusammen 14% aller KI-bezogenen IP-Portfolios, und das hauptsächlich mit Patenten. Die führenden KI-Unternehmen gehören auch zu den weltweit führenden Investoren in Forschung und Entwicklung.
- Insgesamt setzen die KI-Akteure stärker auf Patente, um neue KI-Erfindungen international zu schützen, als auf Marken, um KI-Produkte zu schützen. KI-Patente und KI-Marken sind in den untersuchten IP-Portfolios selten gebündelt.
- KI-Akteure findet man hauptsächlich einerseits bei großen etablierten Unternehmen oder andererseits bei Start-ups. Junge und kleine Unternehmen verlassen sich tendenziell eher auf KI-Marken, um Kunden und Mitbewerbern KI-bezogene Aktivitäten zu signalisieren und sich auf dem Markt zu positionieren.
- Sektoren der Informations-und Kommunikationstechnologie, insbesondere "Computer & Elektronik" und "IT-Dienstleistungen", sind für die größten IP-Portfolios zum Schutz von KI verantwortlich, während KI-Entwicklungen niedrigerer Anzahl in vielen weiteren Sektoren auftreten.
- Unterschiede zeigen sich in der Art und Weise, wie Unternehmen in bestimmten Wirtschaftssektoren KI-Patente und -Marken in ihrem IP-Portfolio kombinieren. Unternehmen im Dienstleistungssektor haben tendenziell eine größere Anzahl von Marken im Bereich KI als der Durchschnitt aller Sektoren.
- KI wird mit Sektor abhängigen Anwendungsfeldern assoziiert, in denen KI-Akteure tätig sind. Details zur Zusammensetzung des IP-Portfolios von KI-Akteuren nach Branchen zeigen, dass die Akteure neben KI eine weitere große Palette von Technologien, Waren und Dienstleistungen entwickeln und schützen.

- Insgesamt umfassen diese informations-und kommunikationsverwandte Technologien - wie *Computertechnologie* und *digitale Kommunikation* - *elektrische Maschinen*, *Halbleiter*, *Transport und Optik*, um nur einige zu nennen. Andere Nicht-KI-Marken wiederum sind stark mit audiovisuellen, informations- und kommunikationstechnologischen Waren sowie Dienstleistungen verbunden.

Table of contents

Tables

Figures

Boxes

1. Introduction

Recent OECD work has proposed experimental approaches to identify and characterise Artificial Intelligence (AI[1])-related innovations, as well as AI-related goods and services, with the aim to inform the policy discussion on the diffusion and adoption of AI ((Baruffaldi et al., 2020[1]), and (Nakazato and Squicciarini, 2021[2])). Gathering evidence about the way AI develops and permeates economies and societies is key, as AI has the potential to profoundly shape the way firms produce, perform and interact with suppliers, customers and other stakeholders. Also, evidence shows that the development of technologies and products related to AI has accelerated over the last decade (OECD, 2021[3]).

This paper builds on (Baruffaldi et al., 2020[1]), and on (Nakazato and Squicciarini, 2021[2]) to shed further light on the actors driving this change, i.e. on the firms that are deploying AI-related technologies as well as those that are proposing AI-related goods and services on international markets. This can help inform the design of effective industrial policies aimed to foster the development and adoption of AI leading to more inclusive and sustainable economic growth paths.

The present analysis relies on data on Industrial Property (IP) collected and managed by the Micro-data Lab of the Directorate for Science, Technology and Innovation (STI). In particular, data on patents and trademarks have been linked to commercial data sources containing information from business registers, to identify and characterise AI innovating agents and shed light on recent trends in AI-inventions and in the type of AI-related goods and services that are protected by IP rights. The analysis further explores the characteristics of the companies protecting AI through patents and trademarks; the way in which different actors use and combine IP assets when protecting their AI-related assets; and the way applicants combine AI-related technological developments with developments in other technological areas, and with trademarks related to other types of goods or services.

Statistics on IP filings reveal a surge in the protection of AI developments since the mid-2010s. During that period, the United States contributed the most AI-related patents and trademarks, followed by Japan, Korea and China. Overall, AI actors tend to rely more on patents, to protect new AI-inventions internationally, than on trademarks to protect AI-related products, and seldom combine both assets in their IP portfolios.

Linked IP-to-company data suggest a strong polarisation of AI actors between large incumbents on one side and young start-up types of companies on the other side. Young and small entities tend to rely relatively more on AI trademarks to signal AI-related goods and services to customers and competitors, and to position themselves onto the market.

Actors operating in ICT-related sectors, notably “Computers & electronics” and “IT services”, had the largest portfolios of IP related to AI in 2014-18, mainly patents. Differences emerge in the way entities operating in a given economic sector combined AI patents and trademarks in their IP portfolio. Also, sector-dependent complementarities of AI-related developments with other fields emerge.

2. Protecting AI-related inventions, goods and services using IP rights

Administrative micro-data on IP constitute a rich source of information that can be leveraged upon to shed light on AI-related technological developments and on the supply of new AI-related goods and services onto the market. This can be done using patent and trademark (TM) records, respectively. The first, i.e. patent data, have been long acknowledged to represent a good proxy for inventive output (e.g. (Griliches, 1990[4])); the second, i.e. TM, have been found to constitute good proxies for non-technological innovations, especially services, and are used by firms and other economic actors to help customers to recognise the new products that they put on the market[2].

IP records notably include a description of what the IP intends to protect, as well as information on the owners of the inventions, goods and services protected by IP rights (i.e. name of the assignee(s) and location), as well as details about the technological scope or product categories of the IP. This allows identification of the specific technology domain(s) to which a patented invention pertains as well as the type of goods and services a trademark refers to.

With respect to AI-related developments in particular, recent work undertaken with experts from the Max Plank Institute (MPI) and from IP offices, in the framework of the OECD-led IP Statistics Task Force, proposes an operational definition of AI-related patents. This is done using combinations of keywords searches performed on the abstract of published patent documents and of codes of the International Patent Classification (IPC) and Cooperative Patent Classification (CPC) schemes[3].

A second stream of recent OECD work exploits information from registered trademark data to shed light on the diffusion of AI-related goods and services on different markets. It builds on the AI patents definition and recent work by a number of IP offices worldwide, and identifies trademarks relating to AI goods and services using combined keywords search on trademark descriptions (see (Nakazato and Squicciarini, 2021[2])).

Building on the two analytical approaches mentioned above, this paper exploits the rich datasets collected, managed and interlinked in the context of the STI Micro-data Lab infrastructure (see Box 2.1), to identify the companies responsible for AI-related technological developments, goods and services. It also sheds light on the broader technological and IP portfolios that AI-related companies own, to better understand technology and product complementarities with AI.

Box 2.1. The Intellectual Property (IP) Databases of the OECD STI Micro-data Lab

The OECD STI Micro-data Lab collects and links large-scale administrative and commercial micro-level datasets. It benefits from the support of the OECD-led IP Statistics Task Force, which gathers representatives from several IP offices. The data collected and managed in the context of the STI Micro-data Lab mainly encompass administrative data such as IP assets, including patents, trademarks and registered designs; scientific publications; online job postings; and information on companies from commercial providers.

With respect to IP data, the STI Micro-data Lab relies on data from the latest edition of the Worldwide Patent Statistical Database of the European Patent Office (EPO), also known as PATSTAT Global, Autumn 2020 edition. PATSTAT Global contains bibliographic records related to more than 100 million patent documents worldwide, covering more than 70 IP offices, as well as legal event data for more than 40 patent authorities. It also contains data related to registered trademarks and industrial designs from the Intellectual Property Office of the European Union (EUIPO), the Japan Patent Office (JPO) and the United States Patent and Trademark Office (USPTO).

These micro-data complement and enhance official statistics like macro-aggregated or survey-based data, and have the advantage of being granular in nature and comprehensive in time and geographical coverage. By providing detailed information about the behaviour of economic agents and the way science and technology develop and diffuse, these data help address policy-relevant questions, such as those related to the generation and diffusion of new technologies, the different ways in which firms innovate, science-industry links, researchers' mobility patterns and the role of knowledge-based assets in firms' economic performance.

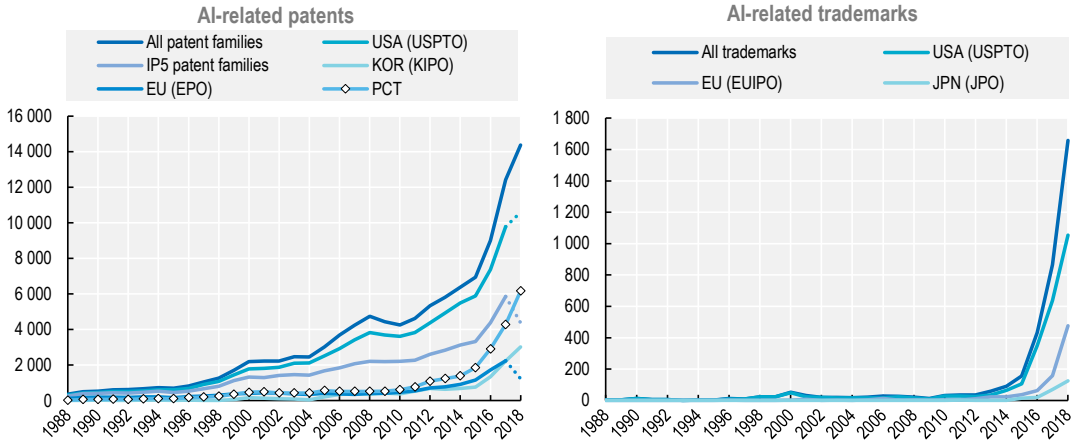
Unless otherwise specified, the patent analysis presented here refers to patent families filed within the Five IP offices (IP5 patent families), namely the EPO, JPO, Korean Intellectual Patent Office (KIPO), the China National Intellectual Property Administration (CNIPA) and USPTO (see (Dernis et al., 2015[5]), for the definition of IP5 patent families). While this approach is to some extent conservative, as it leaves outside of the analysis patents that have been protected only in one country, it makes data more comparable at the international level.

2.1. Filing IP to protect AI

AI-related IP assets increased significantly from the mid-2010s, as the statistics built on patents and trademarks depicted in Figure 2.1 show. The first panel of the figure mirrors the evolution of AI-related patents filed in different IP offices worldwide, while the second panel displays the latest trends in AI-related trademarks. Patent-based indicators rely on patent families.[4]

The patentability of software-based technologies, and hence AI technologies, varies across countries, depending on the rules and regulations of the Intellectual Property Office (IPO) where patent protection is sought. In the United States, software can be patented as such, whereas software becomes patentable at the European Patent Office (EPO) only if it is embedded in "Computer implemented inventions" (CII).[5]

Figure 2.1. Trends in AI developments, by market



Note: Patents counts are provided according to the earliest filing date. Figures for 2018 are underestimated because of data truncation.
Source: OECD, STI Micro-data Lab: Intellectual Property Database, http://oe.cd/ipstats, April 2021.

Overall, the number of patent families in AI nearly tripled since 2010. The US market featured most of the AI-patent filings between 1988 and 2018.[6] In 2017, more than 9 700 patents related to AI were filed at the United States Patent and Trademark Office (USPTO), four times more than the number of AI-patents protected in Korea. AI-related filings at the EPO follow a similar trend to that of the Korean Intellectual Property Office (KIPO). Latest data on patent applications filed under the Patent Cooperation Treaty (PCT) also report a steep acceleration in AI-related inventions since 2015.

Patent families group together patents filed at different offices to protect the same invention on several markets and typically mirror inventions of greater economic value (see (Squicciarini, Dernis and Criscuolo, 2013[6])). IP5 patent families, also presented in Figure 2.1 and used throughout the analysis, cover in particular inventions protected in at least two different jurisdictions worldwide, of which one of the five largest IP offices (see Box 2.1 and (Dernis et al., 2015[7]) for further details on the IP5 patent families definition). However, the surge of AI-related patents cannot be well captured by the IP5 patent families indicator at the end of the time period considered, as delays of at least 18 months exist between when a patent is filed and when it gets published in the different jurisdictions it has been applied. The slope of the trend line observed nevertheless points to an acceleration in developments occurring since 2015.

Statistics on trademarks reported in the right-hand panel of Figure 2.1 complement the picture and mirror the surge in protection of new goods or services related to AI. A steep increase in the number of AI-related trademarks was observed in the three offices considered in the analysis: the Intellectual Property Office of the European Union (EUIPO), Japan Patent Office (JPO) and USPTO. In 2018, more than one thousand AI trademarks were registered at the USPTO in 2018, ten times more than the level of 2015, indicating a marked increase of AI-related goods and services in the US market. The number of AI-trademarks filed at EUIPO is relatively lower, with about 470 trademarks filed in 2018.

Statistics compiled for the JPO are to be considered partial. The search strategy to identify AI-related patents and trademarks relies on English keywords: in the case of JPO, most patent abstracts or trademark descriptions are only available in Japanese characters. This issue has a strong impact on trademarks statistics derived from JPO, as almost no English translation is available for JPO trademarks. In the case of patents, the use of patent families

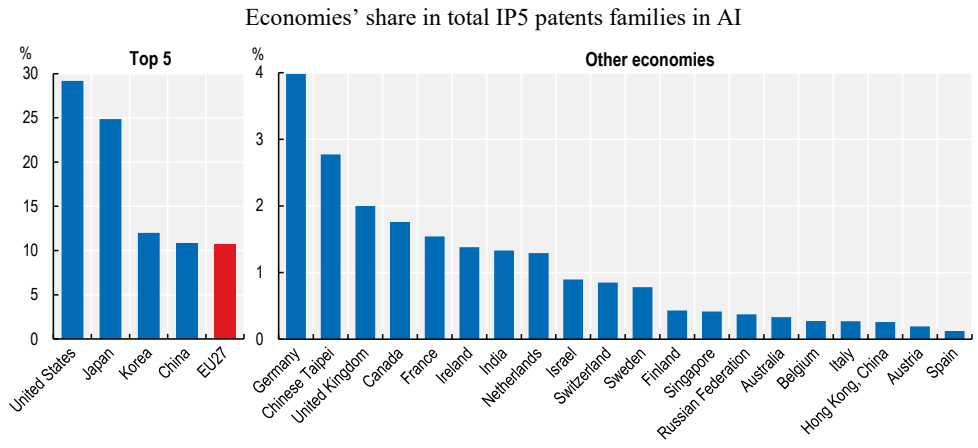
helped to overcome the problem, as families regroup patents filed in different offices worldwide, thus improving the likelihood to having an English translation of the abstract available. In the future, the AI-search strategy will be expanded to integrate non-English keywords and further enhance the coverage of AI-IP assets protected in Japan.

2.2. Top applicants protecting AI-related developments

Analysing the bibliographic information detailed in IP assets helps identify the actors involved in the development of technologies, goods and services related to AI. IP is typically owned by private companies, research institutes, academia, or individuals. As mentioned most of the analysis conducted here exclusively relies on IP5 patent families measures, to better reflect the inventive activity worldwide (see (Dernis et al., 2015[5]) and (Daiko Taro et al., 2017[8]) for further discussion about the use of IP5 families).

Figure 2.2 presents the distribution of IP5 patent families protecting inventions in AI according to the economies in which the patent applicants are located. In 2014-18, the United States (29%) and Japan (25%) were the top two owners of IP5 families protecting AI, contributing altogether to more than half of AI patents. Korea, China, and the EU27 area followed, each with a share of around 11-12% of patents.

Figure 2.2. Top 25 economies with AI patents, 2014-18

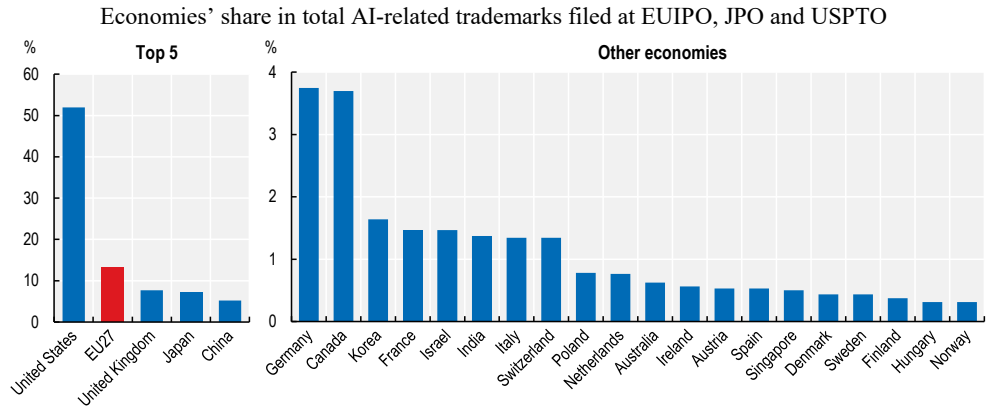


Note: IP5 patent families are presented according to the earliest filing date observed in the family and the location of applicants listed in the family, using fractional counts.
Source: OECD, STI Micro-data Lab: Intellectual Property Database, http://oe.cd/ipstats, April 2021.

Trademark-related statistics presented in Figure 2.3 rely on the set of AI trademarks registered at the EUIPO, JPO or USPTO. Contrary to patents, direct "families" of trademarks cannot be easily identified, as no "priority" filing exist for this IP asset. That is, we cannot clearly identify whether the same trademark has been registered in the three jurisdictions considered. While some international wordmarks provide links between trademarks protected on different markets, the majority of trademarks are likely to use country-specific names or distinctive signs that are difficult to compare. Hence, all trademark data presented here rely on pulling together all data published by the three offices. Furthermore, the identification of AI-related trademarks protected in Japan must be considered incomplete, because of translation issues (the description of goods and services protected by JPO trademarks being only available in Japanese characters).

As a result, counts of AI-related trademarks at the JPO are underestimated, as well as the number of trademarks filed by Japanese actors. Conversely, the share of AI-trademarks filed by US applicants is likely to be overestimated, due to the predominance of USPTO trademarks in the AI-trademark set and a well-documented home bias in national filings (see (Heckemeyer, Olligs and Overesch, 2018[9])). Therefore, the distribution of AI trademarks by economies, shown in Figure 2.1, cannot be considered as precisely reflecting AI-related trademarking activities worldwide, since the analysis only relies on two national IP authorities, i.e. USPTO and JPO, and an international one, the EUIPO. This caveat notwithstanding, we observe that the United States owned about 52% of AI-related trademarks filed at the three offices considered. The EU27 area followed, accounting for 13% of AI-trademarks, before the United Kingdom (8%) and Japan (7%).

Figure 2.3. Top 25 economies with AI trademarks, 2014-18



Note: Trademarks data refer to registrations at the EUIPO, JPO and USPTO, by filing date and location of the applicants, using fractional counts.
Source: OECD, STI Micro-data Lab: Intellectual Property Database, http://oe.cd/ipstats, April 2021.

As a next step, the sample of AI-related IP assets was curated and consolidated to help uncover the entities responsible for AI developments. This cleaning exercise was essential to compensate for the absence of entity unique identifiers in IP data (see Box 2.2 for further details on the data cleaning). Overall, 7 036 distinct applicants with IP assets related to AI were identified in the IP database for the period 2014-18. These comprise 5 770 AI patent owners and 1 436 owners of AI-related trademarks.

Leading economic actors in AI-IP - as measured by the size of their IP portfolio protecting AI related inventions, goods or services - are presented in Table 2.1. AI patenting activity is concentrated in a small number of companies: in 2014-18, the 10 companies with the most AI patents represented 20.5% of AI patents. The concentration of AI trademarks is lower, with 12% of AI trademarks held by the top 10 AI trademarking companies. Most companies featured in the top part of the AI-IP assets distribution also belong to leading world top R&D investors: the data are well in line with the findings published in the 2019 joint EC-JRC/OECD report on the IP bundle of the World Corporate Top R&D investors (Dernis et al., 2019[10]).[7]

Table 2.1. Top 50 applicants with IP assets in AI

Share of companies' IP portfolio in AI in total patents and trademarks in AI

IP applicant name	Location	Share in total IP in AI	Share in AI patents	Share in AI trademarks
Samsung Electronics Co Ltd	Korea	4.5	5.2	0.1
Canon Kabushiki Kaisha	Japan	2.6	3.0	0.0
Google Inc.	United States	2.0	2.2	0.7
Fujitsu Limited	Japan	1.8	2.1	0.1
Microsoft Technology Licensing LLC	United States	1.5	1.8	0.0
Intel Corporation	United States	1.3	1.5	0.4
Baidu Online Network Technology Beijing Co Ltd	China	1.2	1.2	1.3
Sony Corporation	Japan	1.1	1.3	0.1
International Business Machines Corporation	United States	1.1	1.1	1.1
Kabushiki Kaisha Toshiba	Japan	1.1	1.3	0.0
Nec Corporation	Japan	0.9	1.0	0.3
Fanuc Corporation	Japan	0.9	1.1	0.0
Huawei Technologies Co Ltd	China	0.9	0.8	1.6
Accenture Global Solutions Limited	Ireland	0.9	1.0	0.0
GM Global Technology Operations LLC	United States	0.9	1.0	0.0
Ford Global Technologies LLC	United States	0.9	1.0	0.0
Electronics And Telecommunications Research Institute	Korea	0.9	1.0	0.0
LG Electronics Inc.	Korea	0.8	0.8	0.9
Toyota Jidosha Kabushiki Kaisha	Japan	0.8	0.9	0.0
Qualcomm Incorporated	United States	0.7	0.9	0.0
Tata Consultancy Services Limited	India	0.7	0.8	0.2
Hitachi Ltd	Japan	0.7	0.8	0.0
Alibaba Group Holding Limited	Cayman Islands	0.7	0.8	0.0
Siemens Healthcare Gmbh	Germany	0.6	0.7	0.0
Tencent Technology Shenzhen Company Limited	China	0.6	0.7	0.0
Panasonic Intellectual Property Management Co Ltd	Japan	0.6	0.7	0.0
Koninklijke Philips Electronics NV	Netherlands	0.6	0.7	0.0
Robert Bosch Gmbh	Germany	0.6	0.7	0.0
Nvidia Corporation	United States	0.6	0.3	2.7
Fuji Xerox Co Ltd	Japan	0.6	0.6	0.0
Fujifilm Corporation	Japan	0.5	0.6	0.0
Baidu USA LLC	United States	0.5	0.6	0.0
Denso Corporation	Japan	0.5	0.6	0.0
Beijing Baidu Netcom Science and Technology Co Ltd	China	0.5	0.6	0.0
Olympus Corporation	Japan	0.5	0.6	0.0
General Electric Company	United States	0.5	0.5	0.1
Boe Technology Group Co Ltd	China	0.4	0.5	0.0
Konica Minolta Inc.	Japan	0.4	0.5	0.0
Deepmind Technologies Limited	United Kingdom	0.4	0.3	1.1
Siemens Aktiengesellschaft	Germany	0.4	0.5	0.1
Omron Corporation	Japan	0.4	0.5	0.0
Hyundai Motor Company	Korea	0.4	0.5	0.1
Mitsubishi Electric Corporation	Japan	0.4	0.5	0.0
Amazon Technologies Inc.	United States	0.3	0.3	0.6
Ricoh Company Ltd	Japan	0.3	0.4	0.0
The Boeing Company	United States	0.3	0.4	0.0
Guangdong Oppo Mobile Telecommunications Corp Ltd	China	0.3	0.4	0.0
Honda Motor Co Ltd	Japan	0.3	0.4	0.0
Casio Computer Co Ltd	Japan	0.3	0.3	0.0

Note: AI-related trademarks are mostly based on trademarks registered with the EUIPO and USPTO, because of issues identifying JPO trademarks: the share of AI trademarks owned by Japanese companies are underestimated.
Source: OECD, STI Micro-data Lab: Intellectual Property Database, http://oe.cd/ipstats, April 2021.

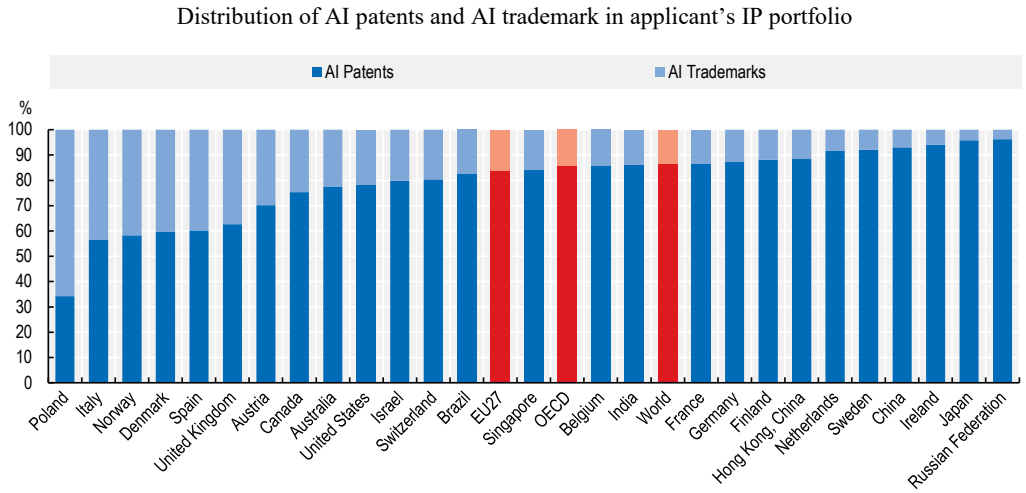
The Korean company, Samsung, ranked top, because of the sheer number of AI-related patents owned, i.e. 5.2% of all AI-related IP5 patent families filed in 2014-18. The Japanese company Canon followed, with about 3% of AI-related patents. In turn, Google Inc, in the US, came third in terms of AI-related IP, contributing respectively to 2.2% of AI patents and 0.7% of AI trademarks.

During the same period, a further three Japanese companies -Fujitsu, Sony and Toshiba – were among the top ten AI-IP users, representing together 4% of AI in IP (Table 2.1). Major ICT-related companies headquartered in the US, namely Microsoft, Intel and IBM also belonged to the top ten. Interestingly, IBM ranked high both in terms of AI patents and in terms of AI trademarks and the US based Nvidia Corporation topped the trademark list, with 2.7% of AI-related trademarks owned. The Chinese company, Baidu, contributed, in turn, to 1.2% of AI patents and 1.3% of trademarks. In Europe, Accenture Global Solutions (Ireland), Siemens Healthcare and Robert Bosch (Germany), Philips Electronics (Netherlands), followed by Deepmind (United Kingdom) appeared among the top companies with AI-related IP assets.

2.3. Combined use of IP to protect AI

Developers of AI-related technologies use different IP strategies to protect their inventions, goods and services. Overall, applicants tend to protect AI related inventions with multiple patents while they seem to use relatively fewer trademarks to signal their AI-related activities on the market. Figure 2.4 shows the proportion of AI-related patents and AI-related trademarks in the AI bundle of IP of applicants, broken down by geographical location. The basket of AI patents observed refers to IP5 patent families, while AI trademarks are those filed at EUIPO, JPO and USPTO.

Figure 2.4. Use of IP to protect AI, by location of applicants, 2014-18



Note: The patent portfolio of economies refers to IP5 patent families owned by economies, using fractional counts. Trademarks cover registrations at the EUIPO, the JPO and the USPTO, by applicant location using fractional counts. Only locations featuring more than 20 AI patent or trademarks in their portfolio are included.
Source: OECD, STI Micro-data Lab: Intellectual Property Database, http://oe.cd/ipstats, April 2021.

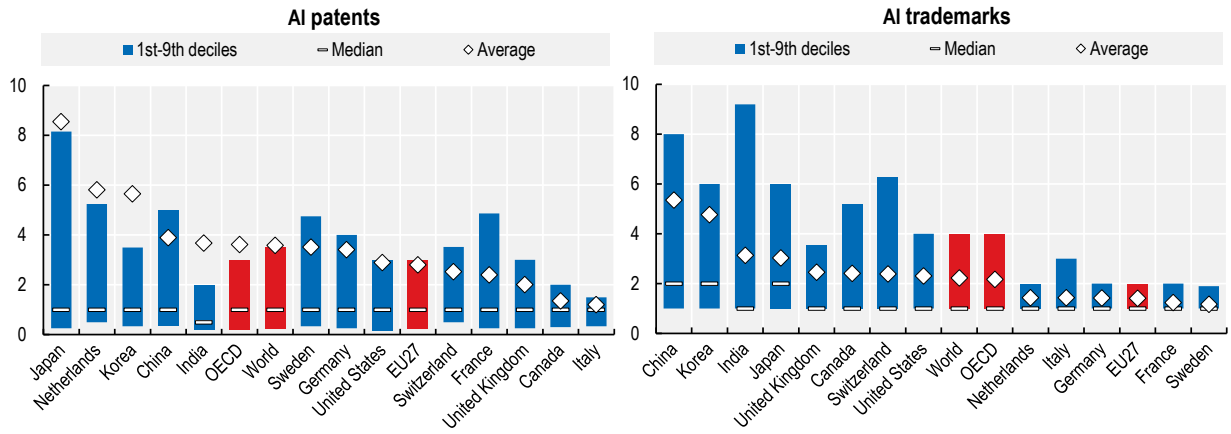
In 2014-18, patents were the predominant IP asset in the AI-IP portfolio of applicants, covering 87% of the IP bundle in AI, while selected economies reported a higher proportion of trademarks among their AI IP assets. Albeit its AI IP portfolio size was relatively smaller, Poland features the largest share of trademarks (66%). Data further suggest that

Italy (44%), Norway (42%), Denmark (40%), Spain (40%) and the United Kingdom (37%) tended to rely more on trademarks than other economies, when considering the distribution of patents and trademarks in their AI-related IP portfolio.. The relatively low proportion of AI trademarks in the IP bundle of Asian applicants, especially in the case of Japan, is conversely likely due to the data limitations, as underlined in the previous sections.

The way in which applicants rely on patents or on trademarks to protect AI-related technologies, goods or services tend to vary, depending on the economies in which economic actors are located. The upper part of Figure 2.5 displays the statistical distribution of AI patents in the applicants' IP portfolios, by economies, as well as that of AI trademarks in the applicants' portfolios in the same economy. On average, applicants filed for around 3.6 IP5 patent families to protect inventions in AI during 2014-18, more than the average number of AI trademarks (2.2) filed during the same period. The AI-patent distribution seem to be led by only a few companies, with 50% of the companies filing for only one AI-related patent during the period considered.

Figure 2.5. Distribution of AI-related patents and trademarks by applicants, 2014-18

Average number of patents and trademarks per applicant, median and deciles distribution

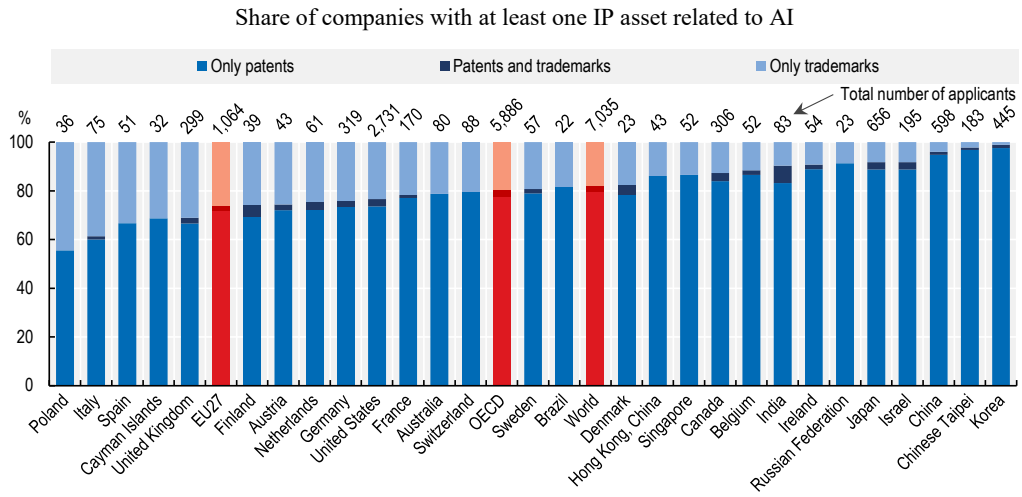


Note: Only locations with more than 40 applicants with AI patents and 10 applicants with AI trademarks are included.
Source: OECD, STI Micro-data Lab: Intellectual Property Database, http://oe.cd/ipstats, April 2021.

The most active AI applicants were located in Japan, with an average of 8.6 AI patents and 3 AI trademarks per applicant over the period considered. Japan also featured the largest distribution in terms of AI patents among all economies considered: the distribution for Japan being led by a few outliers that filed for a large number of AI-patent families. In China, applicants filed about 4 IP5 patents to protect AI inventions on average in 2014-18, and owned an average of more than 5 AI trademarks at the three offices considered.

A limited number of applicants tend to rely on both patents and trademarks to protect AI. Figure 2.6 features the distribution of applicants according to the composition of their IP portfolio, distinguishing those only using patents to protect AI, those only relying on trademarks, and applicants using combinations of the two. The low proportion of companies with both patents and trademarks related to AI is likely to be underestimated because of the automated data consolidation process: further manual curation of applicant names would be required to better capture the IP portfolio of each applicant (see Box 2.2).

Figure 2.6. IP use by applicants protecting AI, by location, 2014-18



Note: Data refer to applicants for which the AI-related IP portfolio is composed of patents only, trademarks only or a combination of both. Only locations with more than 20 companies in the sample are included. Locations are ranked by decreasing share of companies using only trademarks.
Source: OECD, STI Micro-data Lab: Intellectual Property Database, http://oe.cd/ipstats, April 2021.

Overall, 80% of applicants only used patents to protect AI, and 18% of them only owned AI related trademarks. About 2% of applicants in the sample combined both patents and trademarks. With respect to EU27 applicants, nearly 75% of their AI IP portfolios were seemingly composed only of patents. The combined use of patents and trademarks remained relatively marginal, whereas 33% to 44% of AI active entities located in Italy, Spain and Poland relied on AI trademarks only. Economic agents in Asian economies appeared to principally protect AI developments through patents, although this share is likely to be overestimated because of the data limitations explained in the previous paragraphs.

Box 2.2. IP data consolidation

Applicant names, as registered in IP documents, are not standardised, neither within nor across IP offices: a single applicant may be recorded as one or more entries depending on the way the name and address are written on the filing document (due to e.g. use of acronyms, spelling variations, etc.). A thorough data curation exercise is therefore required to be able to group similar names together and consolidate an applicant's IP portfolio.

In this work, the names of applicants were first harmonised using country-specific "dictionaries", aimed to deal with variations in legal entity denomination (e.g. *Limited* and *Ltd*), common names and expressions, as well as phonetic and linguistic rules that might affect the way in which enterprise names are written. Failing to account for such features of the data might mistakenly lead to double counting an applicant because different spellings of its name made it appear to be different entities. Names of JPO trademark applicants, only available in Japanese characters, were translated into Latin characters to allow for the matching algorithms to properly run.

In a second step, the names of patent applicants were matched to those of trademark applicants by means of comparing the strings of characters that compose an applicant name. This was done using token-based and string-metric-based algorithms, such as token frequency measures, Levenshtein (Levenshtein, 1965[5]) and Jaro-Winkler (Winkler, 1999[6]) distances. The precision of the match, which depended on minimising the number of false positive matches, was ensured through a selection of pairs of company names/ IP rights owners made on the basis of high-score thresholds imposed on the algorithms considered. Additional manual curation was needed to improve the recall and precision rates.

The harmonisation and matching procedures were carried out on a by-country basis using a series of algorithms contained in the *Imalinker* (Idener Multi Algorithm Linker) system developed by Idener (http://www.idener.es/). As a complement, IP applicant names were in turn linked to the company level data provided in the ORBIS© database, a private database maintained by Bureau van Dijk Electronic publishing. This additional matching exercise helped improve the linked patent to trademarks data, and enabled collection of additional data on the applicants (e.g. sectoral classification, age of companies, size, etc.). Further details on the matching steps are provided in Box 3.1.

Note that the semi-automated matching exercise performed between patents and trademarks applicants may underestimate the number of applicants featuring both patents and trademarks, as many name variations may remain. Additional manual curation would further improve the consolidation of the IP portfolio of applicants.

3. Characterisation of companies protecting AI with IP assets

While IP documents include a wealth of information about the inventions, goods or services that are protected, they do not feature any detail about the characteristics of the applicants, with the exception of their name and address. The sample of applicants with IP in AI was thus complemented by business register type data, to be able to shed light on the characteristics of AI applicants. The analysis here exploits data derived from the ORBIS© database available at the OECD (see Box 3.1 for further details).

3.1. General typology of applicants

The analysis of AI-IP applicants by companies' age or size reveals a polarisation of AI actors between large incumbent and young start-up types of companies. Available data on the age of companies indicate that, while 33% of AI applicants referred to companies of 20 years and above, 25% of the companies protecting AI-related IP were younger than 5 years of age (Figure 3.1). In turn, micro-firms, with less than 10 employees, represented 42% of the sample in 2018, and 30% of the sample were categorised as large companies.

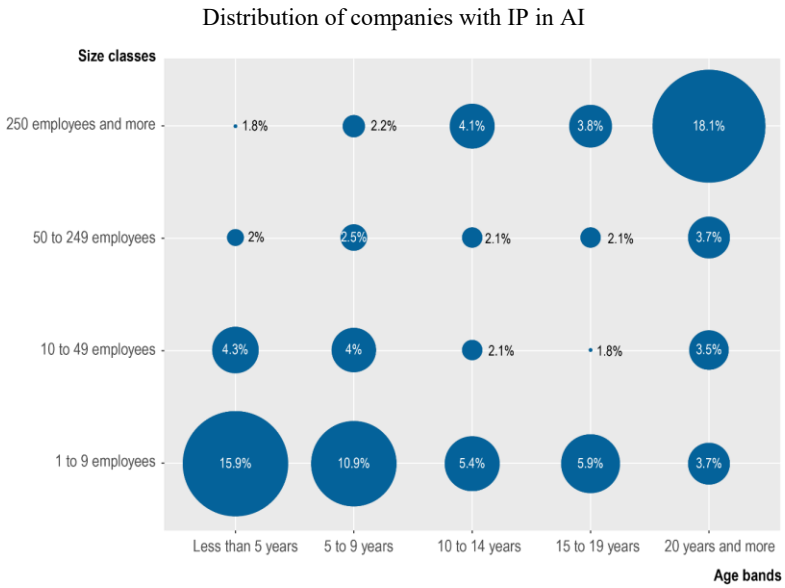
Figure 3.1. AI applicants by firm age or size classes, 2018

Note: The age of the applicants is calculated as the difference between 2018 and the incorporation date as provided in ORBIS© (data cover 74% of the sample). Size classes were constructed using the number of employees reported in ORBIS© for the year 2018, or latest available year down to 2015 (data cover 44% of the sample).
Source: OECD, STI Micro-data Lab: Intellectual Property Database, http://oe.cd/ipstats, and ORBIS, version 2.2020, Bureau van Dijk, April 2021.

The dichotomy between start-up and incumbent actors is emphasised when crossing both age and size data, as shown in Figure 3.2. However, note that the picture is likely to be distorted because of missing records on age and size in some countries, and because of the underrepresentation of small firms in the ORBIS© database (see more details on data availability by country in Box 3.1).

In 2018, nearly 27% of AI-IP companies were young start-ups, with less than 10 employees and less than 5 years old (16%) or between 5 and 10 years of age. On the other side of the spectrum, large and old companies represented in turn 18% of the sample.

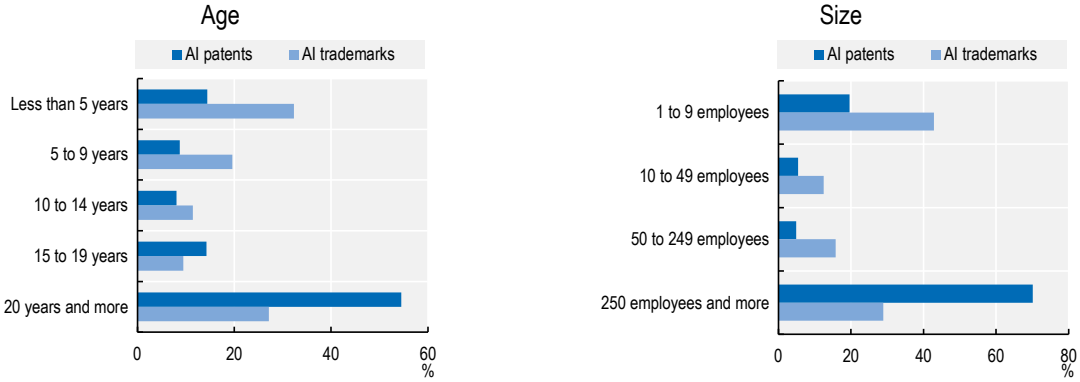
Figure 3.2. AI applicants by combinations of age and size classes, 2018



Note: Data refer to the sample of applicants matched to ORBIS© for which the incorporation date and the number of employees were available (43% of the sample). The age of the applicants is calculated as the difference between 2018 and the incorporation date as provided in ORBIS©. Size classes were constructed using the number of employees reported in ORBIS© for the year 2018, or latest available year down to 2015.
Source: OECD, STI Micro-data Lab: Intellectual Property Database, http://oe.cd/ipstats, and ORBIS, version 2.2020, Bureau van Dijk, April 2021.

The polarisation is less pronounced when looking at the distribution of the number of IP in AI by age and by size of firms (Figure 3.3). In 2014-18, most AI-patents were owned by old firms (55% of patents) and by larger firms (70%). However, younger and smaller firms relied more on trademarks to protect AI: companies aged below 5 years owned 32% of AI-trademarks; this share reached 43% for companies with less than 10 employees.

Figure 3.3. IP portfolio related to AI, by age and size, 2014-18



Note: The age of the applicants is calculated as the difference between 2018 and the incorporation date as provided in ORBIS©. Age information is missing for 16% of AI patents and 39% of AI trademarks. Size classes were constructed using the number of employees reported in ORBIS© for the year 2018, or latest available year down to 2015. Size classes are missing for 40% of AI patents and 49% of AI trademarks.
Source: OECD, STI Micro-data Lab: Intellectual Property Database, http://oe.cd/ipstats, and ORBIS, version 2.2020, Bureau van Dijk, April 2021.

Box 3.1. IP applicants linked to ORBIS© business register data

ORBIS© is a commercial database provided to the OECD by electronic publishing firm, Bureau Van Dijk. It is the largest cross-country firm-level database that is available and accessible for economic and financial research. The industry coverage reflects the non-farm business sector, i.e. all industries excluding agriculture and public services. However, the ORBIS© coverage varies by country, industry, over time and across variables within the data. A recent study by (Bajgar et al., 2020[11]) uncovered issues in the coverage and representativeness of the data, in particular with respect to the over-representation of large companies. The analysis here exploits the OECD ORBIS vintage from February 2020, the most recent available version at the time of the analysis.

Applicant names were linked to ORBIS© company names using the Imalinker system, following the steps described in Box 2.2. The outcomes of the matching exercise did not systematically result in unique pairs of names, as in many cases one applicant may be linked to several entries in ORBIS©. Multiple entries may also be found in relation to the same name due to the existence of branches or affiliates having the same company name. As a consequence, further filtering and consolidation are required, using the information contained in additional fields like the address of the company and the company's ownership structure (see (Squicciarini and Dernis, 2013[12]) for further details on the disambiguation rules).

Overall, 69% of applicants that used IP to protect AI-related inventions, goods or services could be linked to ORBIS©. Matching rates, compiled as the proportion of AI-related IP asset for which the applicant could be linked over the total number of AI-related IP filed in 2014-18, reached 86% of AI-patents and almost 80% of AI trademarks (see the table below). The availability of company-level variables is unevenly distributed across countries, and, on average, the date of incorporation is provided for 74% of companies in the sample, while recent data on the number of employees are missing for 56% of the sample. In turn, the industrial sector to which the applicants belong could been identified in 75% of cases.

Matching rates and coverage of ORBIS© variables, by country, IP in 2014-18

	Number of applicants with IP in AI	*of which matched to ORBIS©*	Matching rate (%)			Missing fields in ORBIS© (%)		
			AI patents	AI trademarks	AI-related IP	Age	Size	Sector
All countries	7,036	4,889	86.2	79.6	85.4	26.1	56.2	34.6
Australia	80	70	75.0	90.0	78.3	18.6	51.4	51.4
Belgium	52	43	80.0	80.0	80.0	14.0	39.5	20.9
China	598	351	78.2	89.2	78.8	16.8	57.8	24.8
France	170	176	94.3	92.0	94.0	2.3	66.5	11.9
Germany	319	222	89.4	68.3	86.9	16.2	44.1	27.9
Ireland	54	33	92.4	88.9	92.2	12.1	54.5	45.5
Israel	195	129	71.1	80.9	72.7	7.8	56.6	38.8
Italy	75	53	79.4	75.0	77.7	32.1	54.7	35.8
Japan	656	430	92.3	70.9	91.5	12.8	29.1	20.5
Korea	445	265	91.2	84.9	91.1	12.8	39.6	17.0
Netherlands	61	33	88.5	100.0	89.4	6.1	54.5	9.1
Sweden	57	54	96.0	92.9	95.8	13.0	38.9	25.9
Switzerland	88	73	88.7	86.0	88.3	6.8	17.8	12.3
United Kingdom	299	256	88.5	93.5	90.2	12.9	45.7	18.0
United States	2,731	2,233	87.9	85.3	87.4	39.2	67.1	46.0

Note: Top 15 countries, with more than 50 companies in AI. Matching rates refer to the number of AI patents for which applicants could be matched over the total number of AI patents (resp. trademarks or patents and trademarks), by economy.
Source: OECD, STI Micro-data Lab: Intellectual Property Database, http://oe.cd/ipstats, and ORBIS, version 2.2020, Bureau van Dijk, April 2021.

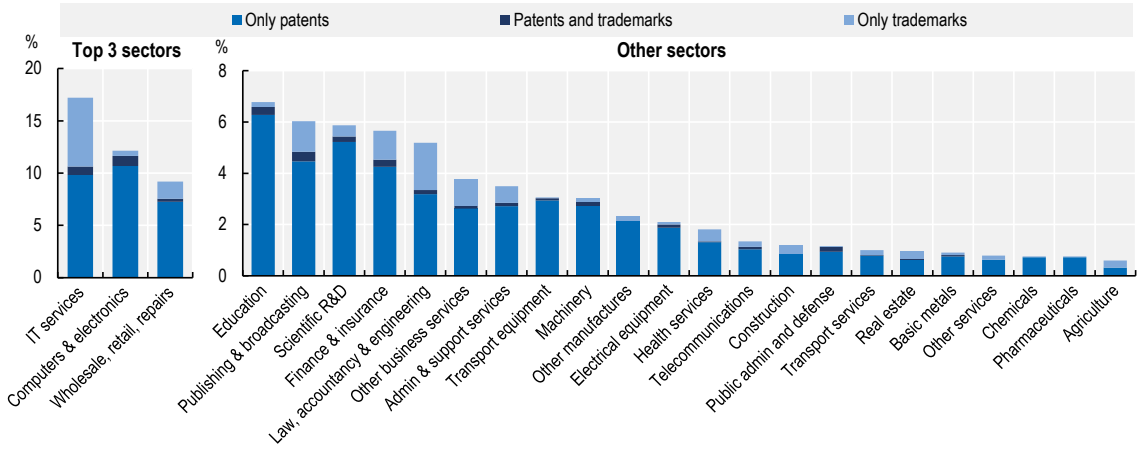
3.2. Top sectors protecting AI with patents or trademarks

The linked IP applicants to ORBIS© records were exploited to also investigate the sectors responsible for AI patenting or trademarking. Core economic sectors in which companies operate could be identified for 75% of the sample. Figure 3.4 displays the distribution of AI applicants by sector, according to the main groups of the International Standard Industrial Classification of All Economic Activities (ISIC, rev. 4), and by type of IP used.

A broad range of economic sectors appear to be protecting AI, confirming the transversal nature of AI and its applicability in many different domains, from health-related advances to marketing. In 2014-18, more than one third of AI-IP applicants came from Information and Communication Technologies (ICT)-related sectors: nearly 17% of applicants with AI-related IP assets belonged to the "IT services" sector, among which 10% exclusively relied on patents to protect AI. The "Computers & electronics" sector followed, representing 12% of AI-IP applicants, mainly patent-oriented. In turn, about 9% of applicants were operating in "Wholesale, retails, repairs", and 7% in the "Education" sector.

Figure 3.4. Top 25 economic sectors with IP in AI, 2014-18

Distribution of companies having filed for AI patents and/or AI trademarks by sector, and IP strategies

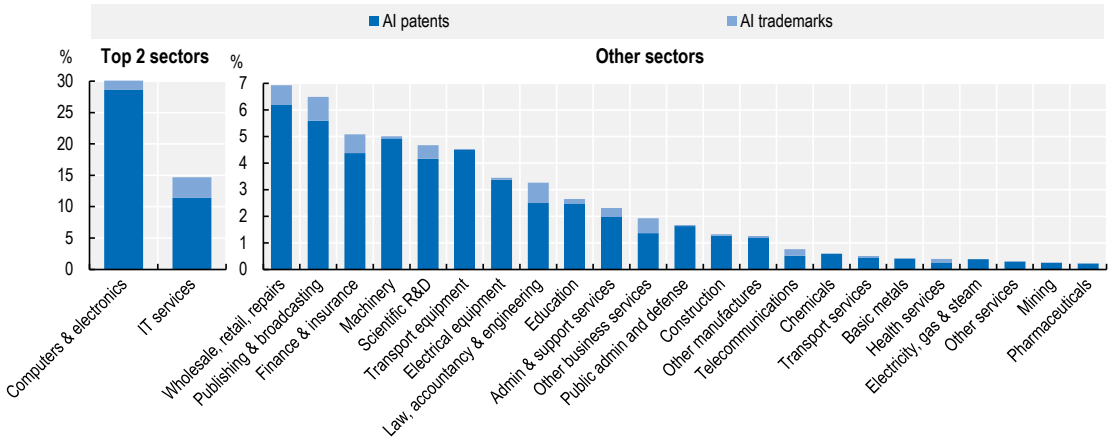


Note: Data refer to the share of applicants with a given IP portfolio combination out of the total number of applicants linked to ORBIS©.
Source: OECD, STI Micro-data Lab: Intellectual Property Database, http://oe.cd/ipstats, and ORBIS, version 2.2020, Bureau van Dijk, April 2021.

AI-IP activities appeared highly concentrated in ICT-sectors, notably in "Computers & electronics" that was responsible for 30% of IP filed to protect inventions, goods and services related to AI over the period considered (Figure 3.5). This sector accounted for 29% of AI patents and 14% of AI trademarks. Companies of the "IT services" sector, that represented 15% of IP in AI, were heavy users of trademarks to signal their offer of AI-services to competitors. This sector alone accounted for nearly one third of all AI-trademarks filed at the EUIPO, the JPO and the USPTO. The sectors of "Wholesale, retail, repairs", "Publishing & broadcasting" and "Finance & insurance" contributed each to about 7-8% of AI-related trademarks.

Figure 3.5. AI-related patents and trademarks by sector, 2014-18

Share of patents and trademarks owned by sectors in total patents and trademarks bundle



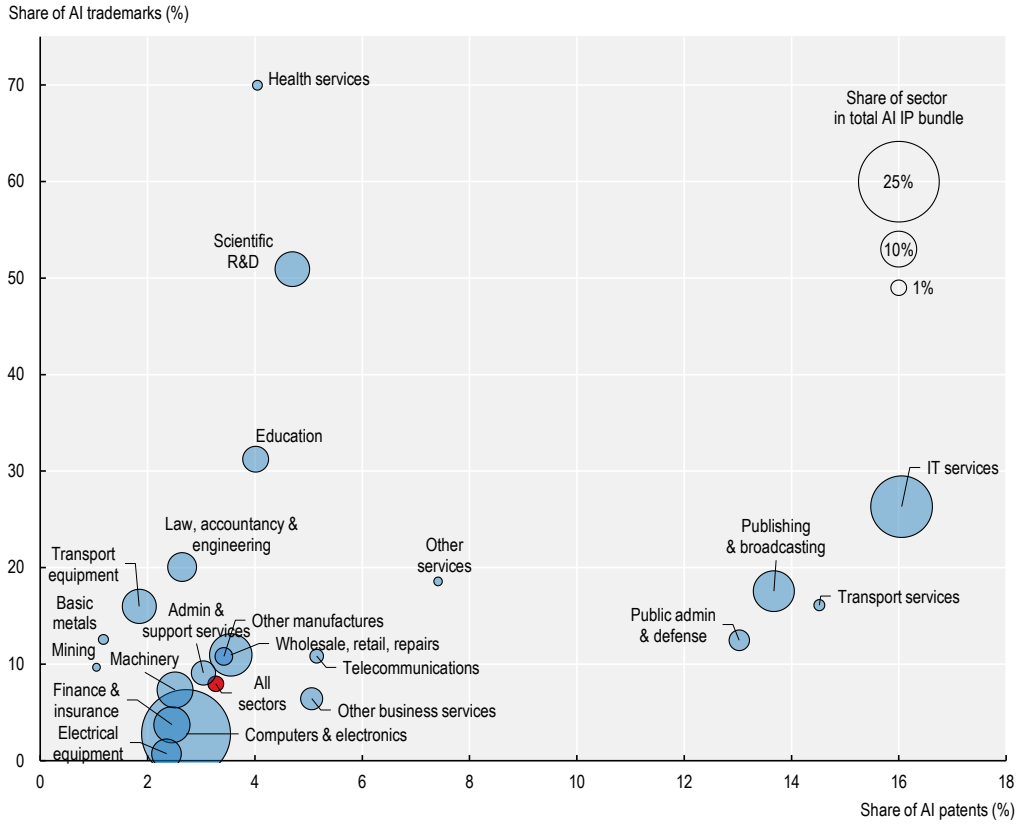
Note: Data refer to the share of applicants with a given IP portfolio combination out of the total number of applicants linked to ORBIS©.
Source: OECD, STI Micro-data Lab: Intellectual Property Database, http://oe.cd/ipstats, and ORBIS, version 2.2020, Bureau van Dijk, April 2021.

Figure 3.6 further explores the concentration of AI-related developments, measured through the IP prism, in the portfolios of companies. The x-axis displays the share of AI-related inventions in companies' patent portfolios, by sector, for companies in the AI-IP sample with at least one IP5 patent family in AI in 2014-18. The y-axis shows a similar proportion of AI-trademarks in companies' trademark portfolios. The size of the bubbles highlights the volume of IP in AI of the sectors (as also shown in Figure 3.5), to give an idea of where firms stand with respect to other firms in the same sector.

Over the period 2014-18, among all companies protecting AI through IP rights, those in "IT services" had the largest proportion of AI-related patents in their portfolio (16%), followed by applicants in "Transport services" (15%) and "Publishing and broadcasting" (14%), and "Public admin. & defense" (13%). The relatively lower concentration figures emerging when comparing patents and trademarks reveal a broader diversification of technological developments in companies' portfolios than in goods and services. In the case of AI-related trademarks, the concentration ratios point to strong specialisations in AI in several services sectors: 70% of trademarks filed by AI-IP applicants from "Health services" sectors relate to AI in 2014-18, and this share reached about 51% in "Scientific R&D" sectors, and 31% in "Education".

Figure 3.6. Concentration of AI in the IP portfolio of applicants, by sector, 2014-18

Share of AI patents (resp. AI trademarks) in AI-IP applicants' patent portfolio (resp. trademark portfolio), and share of economic sectors in the total AI IP bundle



Note: Only sectors featuring at least 10 applicants of IP in AI, having filed for a total of at least 20 patents and 20 trademarks in 2014-18, are included.
Source: OECD, STI Micro-data Lab: Intellectual Property Database, http://oe.cd/ipstats, and ORBIS, version 2.2020, Bureau van Dijk, April 2021.

4. Diversification of the IP portfolio of AI-active companies

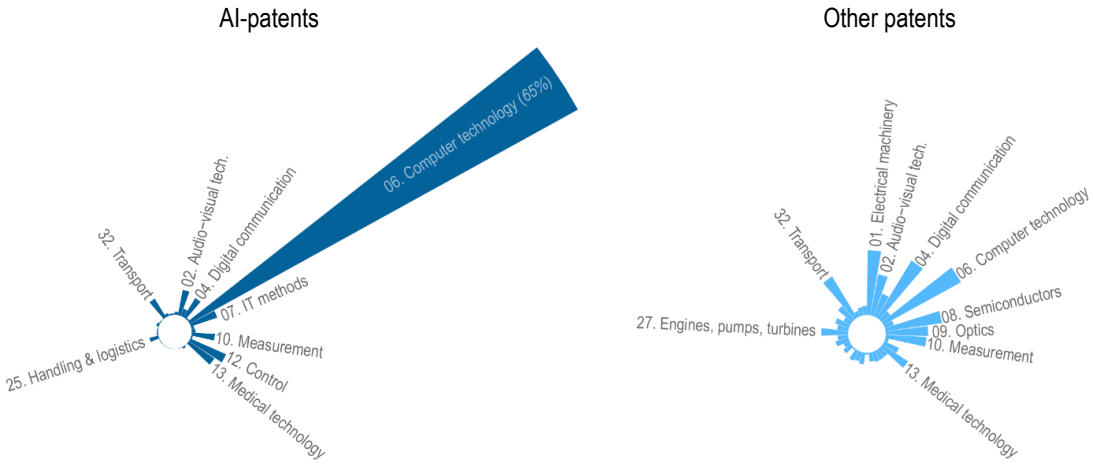
AI can be considered as a General Purpose Technology (GPT), which has many applications in different domains (see e.g. (Trajtenberg, 2018[13])). To see the extent to which AI developments are associated to developments in other technological domains, the IP portfolios of applicants active in AI was decomposed into different technology domains and different goods and services, using the classification schemes available in patent and trademark documents, respectively. This allows shedding light on the relative specialisation of companies in different fields, and their complementarity with AI developments.

4.1. Scope of IP rights protected by entities active in AI

During the examination process, several patent classes may be attributed to the examined patent, depending on the technological domains the invention relates to. This is done generally using the International Patent Classification (IPC) system, which features about 70 000 detailed technology classes. The different IPC codes allocated to patent documents can be aggregated into 35 technology fields using the IPC-technology concordance table[8] provided by the World Intellectual Property Organization (WIPO). Figure 4.1 compares the distribution of AI-related patents by technology domains (left panel) with that of non-AI patents that are owned by applicants in the AI-IP sample (right panel).

In 2014-18, AI-related patents, as defined in (Baruffaldi et al., 2020[1]), were highly concentrated in *Computer technology* domains (65%). Other AI-patents also embedded technologies related to *Control* (6%), *Medical technologies* (5%), *IT methods* or *Audio-visual technologies* (4% each). The portfolio of non-AI patents of the applicants in the sample appeared diversified in several technology domains: while 12% of patents also protected inventions in *Computer technology*, 9% relate to *Digital communication* (9%), *Electrical machinery* (9%), *Semiconductors* (8%), *Transport* (6%) and *Optics* (6%).

Figure 4.1. Technologies protected by AI-patent owners, 2014-18



Note: Patents are allocated to technology fields on the basis of their International Patent Classification (IPC) codes, following the concordance provided by WIPO, using fractional counts. Only the top technology domains are spelled out on the figures.
Source: OECD, STI Micro-data Lab: Intellectual Property Database, http://oe.cd/ipstats, April 2021.

The trademark portfolio of IP applicants protecting AI can conversely be analysed using the classes of the International Classification of Goods and Services (Nice) classification, which comprises 35 classes related to a wide range of goods and 11 classes related to services. The 46 Nice classes were further aggregated into 13 groups of product and services, as shown in Annex A.

In 2014-18, three-fourths of AI-related trademarks were filed to protect goods and services related to *R&D services* (39%, also including software) and *ICT and audio-visual* products (37%), as shown in the left panel of Figure 4.2, followed by trademarks related to *Advertising and business services* (13%). Non-AI trademarks filed by companies in the sample mostly protected *ICT and audio-visual* goods and services (46% of all non-AI trademarks), followed by *R&D services* (13%), *Advertising and business services* (12%) or *Leisure and education* (8%).

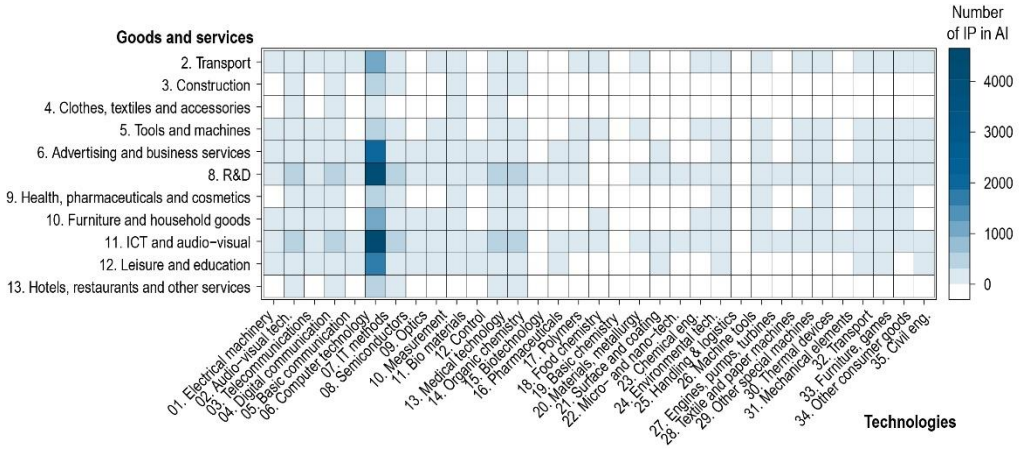
Figure 4.2. Goods and services protected by AI-trademark owners, 2014-18

Note: Trademarks are allocated to groups of goods and services on the basis of their Nice codes, following the concordance presented in Annex A, using fractional counts. Only the top technology domains are spelled out on the figures.
Source: OECD, STI Micro-data Lab: Intellectual Property Database, http://oe.cd/ipstats, April 2021.

The matrix presented in Figure 4.3 displays the way in which technology classes of AI-related patents included in the IP bundle of applicants were combined with the list of goods and services protected by AI-trademarks of the same applicants. As previously observed, most AI patents referred to inventions in *Computer technology* (Figure 4.3), and these were associated with trademarks in *R&D* services, including softwares, *ICT and audio-visual* products as well as *Advertising and business services* and *Leisure and education*.

Figure 4.3. AI-related inventions, goods and services combined in the IP bundle of AI applicants

Combinations of patented technologies and trademarked goods and services related to AI, 2014-18



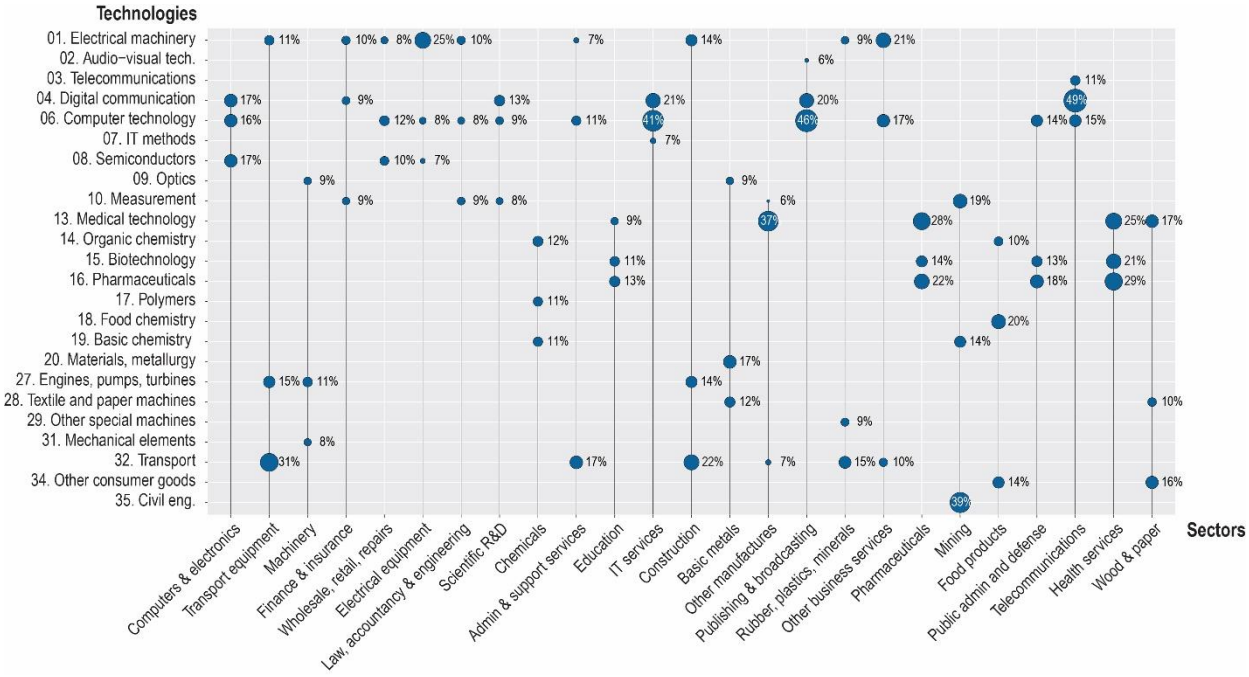
Note: AI-related patents are allocated to technology fields on the basis of their International Patent Classification (IPC) codes, following the concordance provided by WIPO; AI-related trademarks are allocated to groups of goods and services on the basis of their Nice codes, following the concordance presented in Annex A, using fractional counts.
Source: OECD, STI Micro-data Lab: Intellectual Property Database, http://oe.cd/ipstats, February 2021.

4.2. Non-AI innovations, goods or services protected by AI-IP companies

The analysis of the full IP portfolio of companies active in the AI space allows shedding light on the complementarities of AI developments with other technological domains, by sector of activity of the firms. The list of top technologies, goods and services protected through patents and trademarks by sector of activity is shown in Figure 4.4 and in Figure 4.5 below. Data relate to patents or trademarks that do not refer to AI, and both figures are ordered according to the size of the non-AI IP portfolio of each sector.

Figure 4.4. Non-AI technologies protected by AI applicants, by sector, 2014-18

Top 3 technologies in non-AI patents portfolio of sectors



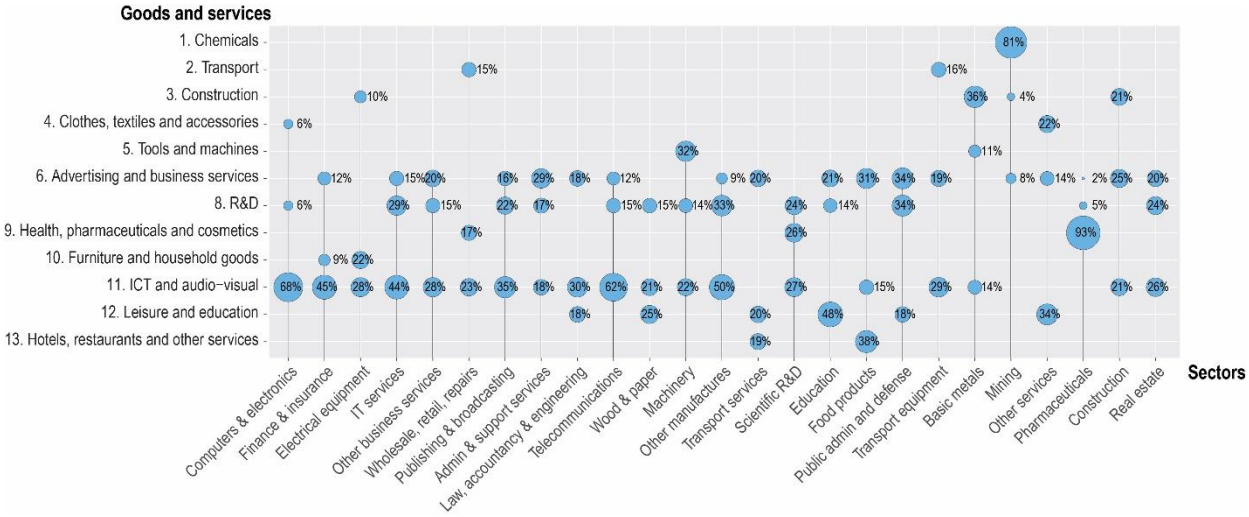
Note: Non-AI patents are allocated to technology fields on the basis of their International Patent Classification (IPC) codes, following the concordance provided by WIPO, using fractional counts. Sectors are ordered by descending number of the patent portfolio of AI patenting companies. Only the top 25 sectors are displayed.
Source: OECD, STI Micro-data Lab: Intellectual Property Database, http://oe.cd/ipstats, and ORBIS, version 2.2020, Bureau van Dijk, April 2021.

In 2014-18, half of technologies protected by "Computers & electronics" companies, which were leading in terms of AI-related patents, referred to *Digital communication*, *Semiconductors* and *Computer technology* (Figure 4.4). In the "Transport equipment" sector, companies protecting AI technologies also developed inventions related to *Transport* (31%), *Engine, pumps, turbines* (15%) or *Electrical machinery* (11%). "IT services" companies with AI patents in their patent portfolio were rather specialised in *Computer technology* (41%) and *Digital communication* (21%). Similar specialisations levels were observed in the case of "Publishing & broadcasting". "Chemical" industries co-developed AI-related techniques alongside with *Organic chemistry* (12%), *Polymers* or *Basic chemistry* (11% each). Applicants in the "Education" sector combined AI inventions with patents related to *Pharmaceuticals* (13%), *Biotechnology* (11%) or *Medical technologies* (9%) in their portfolio.

Similarly to patents, Figure 4.5 presents the distribution of the top 3 groups of non-AI goods and services that were the most frequently protected through trademarks, by sector of activity. During the 2014-18 period, almost all sectors that own AI trademarks were protecting *ICT and audio-visual* goods and services. The concentration in ICT and audio-visual was particularly strong for companies in the "Computers & electronics" sectors (68%) and "Telecommunications" (62%). *R&D* goods and services (of which software) ranked second in terms of trademark protection in a number of sectors, especially in "IT services" (29% of non-AI trademarks) and in "Publishing & broadcasting" sector (22%).

Figure 4.5. Non-AI goods and services protected by IA applicants, by sector, 2014-18

Top 3 goods and services in non-AI trademarks portfolio of sectors

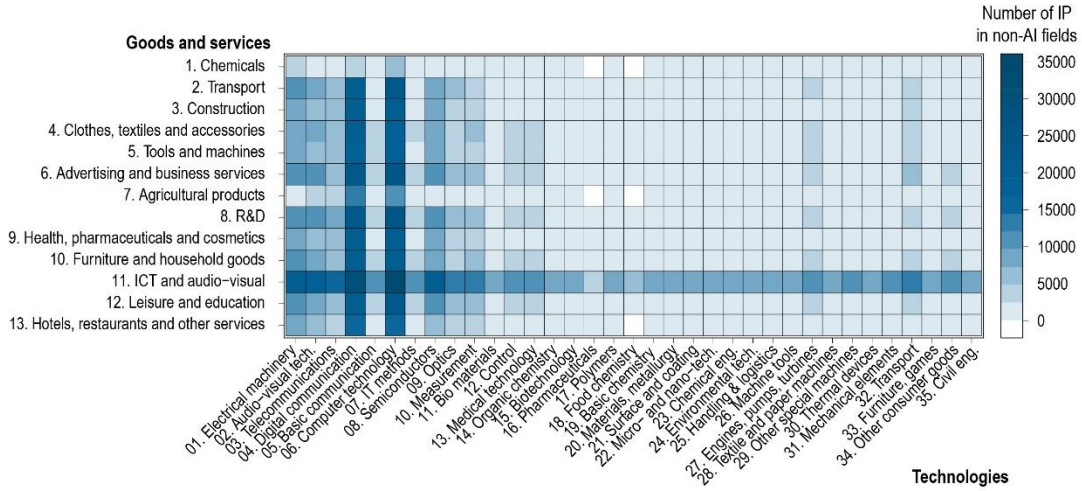


Note: AI-related trademarks are allocated to groups of goods and services on the basis of their Nice codes, following the concordance presented in Annex A, using fractional counts. Sectors are ordered by descending number of the trademark portfolio of AI trademarking companies. Only the top 25 sectors are displayed.
Source: OECD, STI Micro-data Lab: Intellectual Property Database, http://oe.cd/ipstats, and ORBIS, version 2.2020, Bureau van Dijk, April 2021.

The non-AI IP bundle of applicants in the sample covered a broad range of inventions, products and services. Figure 4.6 highlights the most frequent combinations of technologies and goods/services in companies' IP portfolio, beyond AI. While most patents referred to *Computer technology* (15%) and *Digital communication* (14%), these were typically associated to trademarks in a wide range of goods and services, which are generally sector dependent. Among the trademarking categories, *ICT and audio-visual* were products the most frequently protected by trademarks (23%).

Figure 4.6. Other inventions, goods and services combined in the IP bundle of AI applicants

Combinations of patented technologies and trademarked goods and services not related to AI



Note: Non-AI patents are allocated to technology fields on the basis of their International Patent Classification (IPC) codes, following the concordance provided by WIPO; AI-related trademarks are allocated to groups of goods and services on the basis of their Nice codes, following the concordance presented in Annex A, using fractional counts.
Source: OECD, STI Micro-data Lab: Intellectual Property Database, http://oe.cd/ipstats, April 2021.

5. Conclusion

This work, descriptive in nature, proposes an exploratory analysis of AI-actors' characteristics, focusing on entities that are deploying AI-related technologies as well as those that are proposing AI-related goods and services on large international markets. It builds on the STI Micro-data Lab infrastructure, and, in particular, on IP rights data (patents and trademarks) combined with company level data.

According to statistics derived from AI-related patents and trademarks, AI-related activities are highly concentrated in selected countries, sectors and actors. Most developments of technologies and products in the field of AI are due to start-up type of firms or to rather large incumbents. They are mainly located in a few economies (namely the United States, Japan, Korea or China). Companies protecting AI with IP rights mainly operate in ICT related sectors (especially in "Computer & electronics" or "IT services"), while AI developments also occur, albeit to a lesser extent, in other industries. Within companies' IP portfolio, AI is frequently combined with a variety of sector-specific technologies, goods or services.

The identification of AI actors presented in this paper provides a partial overview of AI developments. Other streams of work currently investigate into alternative sources of information on AI companies: using web-data (from GlassAI,[9] for a restricted sample of countries, namely Canada, Germany, the United Kingdom and the United States); and also using information contained in job postings in AI (from Burning Glass Technology).

Endnotes

[1] Artificial Intelligence (AI) refers to machines performing human-like cognitive functions, such as learning, understanding, reasoning and interacting, i.e. machine-based systems capable of influencing the environment by making recommendations, predictions, or decisions for a given set of objectives (OECD, 2019[14]).

[2] See e.g. (Millot, 2009[21]) and (Squicciarini, Millot and Dernis, 2012[20]) for a review of what trademark data help proxy.

[3] See (Baruffaldi et al., 2020[1])for more details. As all operational definitions, the identification strategy proposed may suffer from some biases, including: the need to update the taxonomy on a regular basis to be able to well track over time AI-related developments, which are unfolding at a fast pace; not all AI technologies are protected by patents and not all innovative firms rely on patents; AI-related patents include not only core AI technologies but also applications of AI. Despite these possible drawbacks, the taxonomy by (Baruffaldi et al., 2020[1])provides a solid analytical ground to characterise technological developments in AI.

[4] Patent families correspond to sets of patents filed within or across different offices to protect a similar invention. See (OECD, 2009[18]) for more details about the patent family concept.

[5] Computer implemented inventions are inventions "involving the use of a computer, computer network or other programmable apparatus, where one or more features are realised wholly or partly by means of a computer program". For more details, see the EPO guidelines for examination at https://www.epo.org/law-practice/legaltexts/html/guidelines/e/j.htm

[6] Because of timeliness issues due to delay of publication of patent documents (patent documents are typically published 18 months after filing), figures for 2018 are underestimated due to incomplete data.

[7] The IP bundle of the world top R&D investors relied on the consolidated IP portfolio within large conglomerates, including patents and trademarks filed by their affiliates.

[8] The latest version of IPC-technology concordance table (updated in August 2019), and the accompanying methodological paper, can be found on WIPO's web page at https://www.wipo.int/ipstats/en/

[9] GlassAI is a company that reads and interprets open web text at scale, exploiting any machine-readable information (e.g. sentences, paragraphs, etc.) contained on companies' websites. GlassAI was asked to search the web in order to identify any entity active in the AI space using the list of AI-related keywords described in (Baruffaldi et al., 2020[1]).

References

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://dx.doi.org/10.1787/c7bdaa03-en. [11]

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, https://doi.org/10.1787/18151965 (accessed on 10 July 2020). [1]

Daiko Taro et al. (2017), "World Top R&D Investors: Industrial Property Strategies in the Digital Economy", *Publications Office of the European Union*, http://dx.doi.org/10.2760/861062. [8]

Dernis, H. et al. (2015), "World Corporate Top R&D Investors: Innovation and IP bundles", *JRC Working Papers*, https://ideas.repec.org/p/ipt/iptwpa/jrc94932.html (accessed on 22 June 2020). [5]

Dernis, H. et al. (2015), *World Corporate Top R&D Investors: Innovation and IP bundles. A JRC and OECD Common Report.*, http://dx.doi.org/10.2791/741349. [7]

Dernis, H. et al. (2019), *World Corporate Top R&D investors: Shaping the Future of Technologies and of AI. A joint JRC and OECD report.*, http://dx.doi.org/10.2760/16575. [10]

Griliches, Z. (1990), "Patent Statistics as Economic Indicators: A Survey", *Journal of Economic Literature*, Vol. 28/4, pp. 1661-1707, https://www.jstor.org/stable/2727442 (accessed on 21 September 2020). [4]

Heckemeyer, J., P. Olligs and M. Overesch (2018), ""Home Sweet Home" versus international tax planning: where do multinational firms hold their U.S. trademarks?", *National Tax Journal*, Vol. 71/3, pp. 485-520, http://dx.doi.org/10.17310/ntj.2018.3.03. [9]

Levenshtein, V. (1965), "Binary codes capable of correcting deletions, insertions and reversals.", *Soviet Physics Doklady 10 (8): 707--710 (February 1966) translated from Doklady Akademii Nauk SSSR, V163 No4 845-848*, https://nymity.ch/sybilhunting/pdf/Levenshtein1966a.pdf (accessed on 23 July 2020). [15]

Millot, V. (2009), "Trademarks as an Indicator of Product and Marketing Innovations", *OECD Science, Technology and Industry Working Papers*, No. 2009/6, OECD Publishing, Paris, https://dx.doi.org/10.1787/224428874418. [21]

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://dx.doi.org/10.1787/2db2d7f4-en. [2]

OECD (2021), *OECD.AI trends & data overview*, https://oecd.ai/trends-and-data (accessed on 25 May 2021). [3]

OECD (2019), "Scoping the OECD AI principles: Deliberations of the Expert Group on Artificial Intelligence at the OECD (AIGO)", *OECD Digital Economy Papers*, No. 291, OECD Publishing, https://doi.org/10.1787/d62f618a-en (accessed on 10 July 2020). [14]

OECD (2009), *OECD Patent Statistics Manual*, OECD Publishing, Paris, https://dx.doi.org/10.1787/9789264056442-en. [18]

OECD (n.d.), *OECD.AI trends & data overview*, https://oecd.ai/trends-and-data (accessed on 25 May 2021). [19]

Squicciarini, M. and H. Dernis (2013), "A Cross-Country Characterisation of the Patenting Behaviour of Firms based on Matched Firm and Patent Data", *OECD Science, Technology and Industry Working Papers*, No. 2013/5, OECD Publishing, Paris, https://dx.doi.org/10.1787/5k40gxd4vh41-en. [12]

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://dx.doi.org/10.1787/5k4522wkw1r8-en. [6]

Squicciarini, M., V. Millot and H. Dernis (2012), "Universities' trademark patterns and possible determinants", *Economics of Innovation and New Technology*, Vol. 21/5-6, http://dx.doi.org/10.1080/10438599.2012.656526. [20]

Trajtenberg, M. (2018), *AI as the next GPT: a Political-Economy Perspective*, National Bureau of Economic Research, Cambridge, MA, http://dx.doi.org/10.3386/w24245. [13]

Winkler, W. (1999), "The State of Record Linkage and Current Research Problems", U.S. Bureau of the Census, https://www.census.gov/srd/papers/pdf/rr99-04.pdf (accessed on 23 July 2020). [17]

Winkler, W. and U. Bureau (n.d.), *The State of Record Linkage and Current Research Problems*. [16]

Annex A. Appendix

Table A A.1. Aggregation of NICE classes by fields

1. Chemicals
1. Chemical goods
2. Paints and colorants
4. Oils and fuels

2. Transport
12. Vehicles
39. Transport and packaging

3. Construction
6. Metals
17. Rubber and plastics
19. Building material
27. Carpets and floor covers
37. Building services

4. Clothes, textiles and accessories
14. Precious goods
18. Leather and complements
22. Fibrous products
23. Yarns and threads
24. Textiles
25. Clothing and footwear
26. Decorations

5. Tools and machines
7. Machineries
8. Hand tools

6. Advertising and business services
35. Business and advertising
36. Insurance and finance
45. Legal and personal services

7. Agricultural products
29. Food
30. Condiments and cereals
31. Animals and grains
32. Low and non alcohol drinks
33. Alcoholic drinks
34. Tobaccos.

8. R&D
42. R&D and software

9. Health, pharmaceuticals and cosmetics
3. Cleaning products
5. Pharmaceutical products
10. Medical instruments
44. Medical and hygiene services

10. Furniture and household goods
11. Lightening and heating
20. Furniture
21. House utensils

11. ICT and audio-visual
9. Instruments & computers
38. Telecommunications

12. Leisure and education
13. Firearms
15. Musical instruments
16. Papers and packaging
28. Games
41. Education and sport

13. Hotels, restaurants and other services
40. Treatment of materials
43. Food, drink and accommodation

Source: OECD, groupings based on WIPO, Nice classification, http://www.wipo.int/classifications/nice/en/