OECD publishing

WHAT TECHNOLOGIES ARE AT THE CORE OF AI?

AN EXPLORATION BASED ON PATENT DATA

OECD ARTIFICIAL INTELLIGENCE PAPERS

November 2023 **No. 6**

Federal Ministry
of Labour and Social Affairs

OECD
BETTER POLICIES FOR BETTER LIVES

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

What technologies are at the core of AI? An exploration based on patent data

Flavio Calvino, Chiara Criscuolo, Hélène Dernis and Lea Samek (OECD)

This report outlines a new methodology and provides a first exploratory analysis of technologies and applications that are at the core of recent advances in AI. Using AI-related keywords and technology classes, the study identifies AI-related patents protected in the United States in 2000-18. Among those, “core” AI patents are selected based on their counts of AI-related forward citations. The analysis finds that, compared to other (AI and non-AI) patents, they are more original and general, and tend to be broader in technological scope. Technologies related to general AI, robotics, computer/image vision and recognition/detection are consistently listed among core AI patents, with autonomous driving and deep learning having recently become more prominent. Finally, core AI patents tend to spur innovation across AI-related domains, although some technologies – likely AI applications, such as autonomous driving or robotics – appear to increasingly contribute to developments in their own field.

Keywords: Artificial Intelligence, Patents, Innovation

JEL codes: C81, O31, O33, O34

Résumé

Ce rapport propose une première exploration des technologies et applications qui sont au cœur des avancées récentes de l'IA, en se fondant sur une méthodologie inédite. Les brevets liés à l'IA, déposés aux États-Unis entre 2000 et 2018, sont sélectionnés à l'aide de mots-clés et de classes technologiques liés à l'IA. Parmi ceux-ci, les brevets « au cœur » de l'IA sont identifiés suivant le nombre de citations reçues dans d'autres brevets liés à l'IA. Lorsque l'on compare ces brevets avec d'autres (IA et non-IA), les brevets au cœur de l'IA semblent porter sur des inventions à la fois plus originales et plus générales, avec une portée technologique plus large. De plus, les technologies liées à l'IA en général, à la robotique, à la vision par ordinateur et à la reconnaissance/détection automatique apparaissent systématiquement dans les brevets au cœur de l'IA, comme, plus récemment, la conduite autonome et l'apprentissage profond. Enfin, les brevets au cœur de l'IA ont tendance à stimuler les innovations dans l'ensemble des domaines de l'IA, même si certaines technologies – plus proches des applications de l'IA, comme la conduite autonome ou la robotique – semblent contribuer davantage à leurs domaines propres.

Kurzfassung

Diese Arbeit beschreibt eine neue Methodik und liefert eine erste explorative Analyse der Technologien und Anwendungen, die den Kern der jüngsten Fortschritte in der KI bilden. Unter Verwendung von KI-bezogenen Schlüsselwörtern und Technologieklassen werden in der Studie KI-bezogene Patente identifiziert, die in den Jahren 2000-18 in den Vereinigten Staaten geschützt wurden. Unter diesen Patenten werden "Kern"-KI-Patente auf der Grundlage ihrer Anzahl an KI-bezogenen Vorwärtszitaten ausgewählt. Die Analyse zeigt, dass diese Patente im Vergleich zu anderen (KI- und Nicht-KI-) Patenten origineller und allgemeiner sind und tendenziell einen breiteren technologischen Anwendungsbereich haben. Technologien im Zusammenhang mit allgemeiner KI, Robotik, Computer-/Bildverarbeitung und Erkennung/Detektion sind durchweg unter den wichtigsten KI-Patenten aufgeführt, wobei autonomes Fahren und Deep Learning in letzter Zeit an Bedeutung gewonnen haben. Schließlich treiben KI-Kernpatente in der Regel die Innovation in allen KI-verwandten Bereichen voran, auch wenn einige Technologien - wahrscheinlich KI-Anwendungen wie autonomes Fahren oder Robotik - zunehmend zu Entwicklungen in ihrem eigenen Bereich beizutragen scheinen.

Acknowledgements

The authors would like to thank Leonidas Aristodemou, Hanna-Mari Kilpelainen, Jens Lundsgaard, Jerry Sheehan and colleagues from the Directorate of Education and Skills for comments and suggestions, Marcio Carvalho and Shai Somek for excellent editorial support, participants to the 6th Annual IP Data and Research Conference organised by the Canadian Intellectual Property Office (CIPO) and the Centre for International Governance Innovation (CIGI), participants to the Workshop on Strategic Autonomy, Innovation, Digitalization and Artificial Intelligence at Rennes School of Business, and participants to the OECD Working Party on Industry Analysis. This work contributes to the OECD programme on AI-WIPS, supported by the German Federal Ministry of Labour and Social Affairs (BMAS).

Table of contents

Résumé 4

Kurzfassung 5

Acknowledgements 6

Executive Summary 9

Synthèse 10

Zusammenfassung 12

1 Introduction 14

2 Setting the scene 16

3 Data and methodology 18
- Identifying AI-related patents among all patents 18
- Identifying '*core*' AI patents 18
- Grouping core AI patents into different topics 19

4 Exploring the characteristics of core AI patents 22
- Measuring the quality of core AI patents: scope, originality, generality and radicalness 22
- Exploring core AI innovations thematically 24
- Exploring the role of core AI patents for subsequent innovations 25

5 Concluding remarks and next steps 32

Endnotes 34

References 36

Annex A. Additional information about data and methodology 41

Annex B. Additional results 44

FIGURES

Figure 3.1. AI-related patents and core AI patents, 2000-18 19
Figure 3.2. Pairwise correlation of AI keywords, core AI patents, 2000-18 20
Figure 4.1. Patent quality indicators, different patent groups, 2000-18 23
Figure 4.2. Word clouds of AI main topics, 2000-18 25
Figure 4.3. Core AI patents with AI citations, 2000-18 26
Figure 4.4. Algorithms-related AI patents and their citations, 2000-18 27
Figure 4.5. Recognition- or Detection-related AI patents and their citations, 2000-18 29
Figure 4.6. Autonomous Driving-related AI patents and their citations, 2000-18 30
Figure 4.7. Robotics-related AI patents and their citations, 2000-18 31

Figure A B.1. Patent quality indicators, core AI vs. top cited non-AI patents, 2000-18 44
Figure A B.2. Patent quality indicators, different patent groups, 2000-09 vs. 2010-18 45

TABLES

Table A A.1. List of AI keywords and corresponding AI topics 42

Executive Summary

Artificial intelligence (AI) is rapidly transforming economies and societies. Despite the centrality of AI in the current economic and policy debates, empirical analyses identifying and characterising the technological base underlying AI diffusion are still limited.

This report outlines a new methodology and provides a first exploratory analysis of AI-related core technologies, i.e., technologies and applications at the core of recent advances in AI, based on US patent data.

It first identifies AI-related patents registered at the United States Patent and Trademark Office (USPTO) between 2000 and 2018, i.e., before the more recent boom in generative AI, based on an established list of AI-related keywords as well as technology classes closely related to AI. It then defines the top cited ones as *"core"* AI patents, using counts of their AI-related forward citations.

Analysing the characteristics of core AI patents suggests that they are more original, i.e., rely on a broader range of technology fields, are more general, i.e., are cited by patents belonging to a wider range of fields, and tend to be broader in technological scope, compared to their less cited counterparts or other patents in non-AI technologies. Although AI patents are technologically more radical, i.e. they build upon technologies other than the ones classified in the patent itself, when compared to other non-AI patents, core AI patents do not significantly differ with respect to their less cited AI-related counterparts.

Zooming in on the content of core AI patents' abstracts reveals that some topics, such as *general AI*, *robotics*, *computer or image vision* and *recognition or detection*-related technologies, are consistently at the core of AI developments. Technologies that gained significant importance in more recent years include *autonomous driving* as well as *networks*, such as *deep learning.*

Exploring citation networks of AI patents citing core ones allows to further explore the direction of AI innovation and the existing interdependencies among different AI technologies. Core AI technologies tend to spur AI-related innovation beyond their own respective field. However, some of them – likely related to AI applications, such as *autonomous driving* or *robotics* – appear to increasingly contribute to developments in their own field.

This analysis is complementary to recent OECD work analysing AI innovation and diffusion using different sources of data, including firm-level surveys, online job postings, or data from internet websites. It provides evidence base that is relevant for implementing recommendations consistent with the OECD AI Principles and is also complementary to recent OECD work focusing on generative AI or on the role of AI for science.

Identifying and analysing AI core technologies is relevant not only to better understand the technological base underlying the diffusion of AI across the economy; it may also inspire future policy-relevant research in other domains. This could include the links between holding core AI patents and firm performance and innovation, or the possible interdependencies between core AI and other technologies, such as those related to the green transition.

Synthèse

L’intelligence artificielle (IA) opère une rapide transformation des économies et des sociétés. Malgré la place centrale qu’a pris l’IA dans les débats économiques et politiques actuels, les analyses empiriques identifiant et caractérisant les fondements technologiques sous-jacents à la diffusion de l’IA sont encore limitées.

En se fondant sur une méthodologie inédite utilisant les données sur les brevets américains, ce rapport propose une première exploration des technologies au « cœur » de l'IA, à savoir les technologies et leurs applications qui sont au cœur des avancées récentes de l'IA.

Les brevets liés à l'IA déposés auprès de l’office américain des brevets (USPTO) entre 2000 et 2018 sont sélectionnés sur la base de mots-clés et de classes technologiques liés à l’IA. Ces brevets ne couvrent pas les développements de l'IA générative, plus récents. Les brevets les plus cités sont catégorisés comme étant des brevets au cœur de l’IA, selon le nombre de citations reçues par des brevets liés à l’IA.

L’analyse des caractéristiques des brevets au cœur de l’IA suggère qu’ils sont plus originaux que leurs homologues moins cités ou que des brevets pris pour des technologies non liées à l’IA : ils s’appuient sur un plus large éventail de domaines technologiques. Ils semblent être plus généraux, étant cités par des brevets couvrant un plus large éventail de technologies, et ont une portée technologique plus vaste. Bien que les brevets sur l’IA portent sur des technologies plus radicales, c’est-à-dire qu’ils se fondent sur des technologies différentes de celles couvertes dans le brevet lui-même, les brevets au cœur de l’IA ne diffèrent pas significativement des autres brevets liés à l’IA moins fréquemment cités.

En regardant de plus près le résumé des brevets au cœur de l’IA, certains sujets, comme l’IA en général, la robotique, la vision par ordinateur, et les technologies liées à la reconnaissance ou à la détection automatique, se trouvent systématiquement au cœur des développements de l’IA. Parmi les technologies qui ont pris de l’ampleur ces dernières années, on trouve la conduite autonome ainsi que les algorithmes de réseaux, tels que l’apprentissage profond.

L’étude des réseaux de citations de brevets d’IA permet de discerner les orientations de l’innovation en matière d’IA, ainsi que les interdépendances entre technologies liées à l’IA. Les technologies au cœur de l’IA ont tendance à stimuler l’innovation liée à l’IA au-delà de leur propre domaine respectif. Cependant, certaines technologies – plus liées à l’application de l’IA, comme la conduite autonome ou la robotique – paraissent contribuer davantage à leurs domaines propres.

Cette analyse complète les travaux récents de l’OCDE portant sur l’innovation et la diffusion de l’IA, utilisant d’autres sources de données comme les enquêtes au niveau des entreprises, les offres d’emploi en ligne ou des données extraites de sites Internet. Elle fournit des preuves pertinentes pour la mise en œuvre de recommandations conformes aux Principes sur l’IA de l’OCDE et s’ajoute aux travaux récents de l’OCDE axés sur l’IA générative ou sur le rôle de l’IA pour la science.

L’identification et l’analyse des technologies au cœur de l’IA sont utiles non seulement pour mieux comprendre les fondements technologiques compris dans la diffusion de l’IA dans l’économie. Ce travail pourrait également servir de modèle pour des recherches dans d’autres domaines. Les liens entre la détention de brevets au cœur de l'IA et les performances et/ou l'innovation des entreprises, pourraient être étudiés, tout comme les interdépendances potentielles entre l'IA fondamentale et d'autres technologies, telles que celles liées à la transition verte.

Zusammenfassung

Künstliche Intelligenz (KI) verändert Wirtschaft und Gesellschaft rasant. Trotz der zentralen Bedeutung von KI in den aktuellen wirtschaftlichen und politischen Debatten gibt es nur wenige empirische Analysen, die die technologische Basis der KI-Verbreitung identifizieren und charakterisieren.

Dieser Bericht stellt eine neue Methodik vor und bietet eine erste explorative Analyse der KI-bezogenen Kerntechnologien, d. h. der Technologien und Anwendungen, die den Kern der jüngsten Fortschritte im Bereich der KI bilden, auf der Grundlage von US-Patentdaten.

Sie identifiziert zunächst KI-bezogene Patente, die zwischen 2000 und 2018, also vor dem jüngsten Boom der generativen KI, beim Patentamt der Vereinigten Staaten (United States Patent and Trademark Office, USPTO) angemeldet wurden, und zwar auf der Grundlage einer etablierten Liste von KI-bezogenen Schlüsselwörtern sowie von Technologieklassen, die eng mit KI verbunden sind. Anschließend werden die am häufigsten zitierten Patente anhand der Anzahl ihrer KI-bezogenen Vorwärtszitate als "Kern"-KI-Patente definiert.

Die Analyse der Merkmale von KI-Kernpatenten legt nahe, dass sie origineller sind, d. h. sich auf ein breiteres Spektrum von Technologiefeldern stützen, allgemeiner sind, d. h. von Patenten aus einem breiteren Spektrum von Bereichen zitiert werden, und tendenziell einen größeren technologischen Umfang haben als ihre weniger zitierten Gegenstücke oder andere Patente in Nicht-KI-Technologien. Obwohl KI-Patente technologisch radikaler sind, d. h. auf anderen als den im Patent selbst klassifizierten Technologien aufbauen, unterscheiden sich KI-Kernpatente im Vergleich zu anderen Nicht-KI-Patenten nicht wesentlich von ihren weniger zitierten KI-bezogenen Gegenstücken.

Wenn man den Inhalt der Zusammenfassungen der wichtigsten KI-Patente näher betrachtet, wird deutlich, dass einige Themen wie allgemeine KI, Robotik, Computer- oder Bildverarbeitung sowie Erkennungs- oder Detektionstechnologien stets im Mittelpunkt der KI-Entwicklungen stehen. Zu den Technologien, die in den letzten Jahren erheblich an Bedeutung gewonnen haben, gehören das autonome Fahren und Netzwerke wie Deep Learning.

Die Betrachtung von Zitationsnetzwerken von KI-Patenten, die Kerntechnologien zitieren, ermöglicht es, die Richtung der KI-Innovation und die bestehenden Interdependenzen zwischen verschiedenen KI-Technologien näher zu untersuchen. KI-Kerntechnologien sind in der Regel der Motor für KI-bezogene Innovationen, die über ihren eigenen Bereich hinausgehen. Allerdings scheinen einige von ihnen - wahrscheinlich im Zusammenhang mit KI-Anwendungen wie autonomes Fahren oder Robotik - zunehmend zu Entwicklungen in ihrem eigenen Bereich beizutragen.

Diese Analyse ergänzt die jüngsten OECD-Arbeiten zur Untersuchung von KI-Innovation und -Verbreitung unter Verwendung verschiedener Datenquellen, darunter Erhebungen auf Unternehmensebene, Online-Stellenausschreibungen oder Daten von Internet-Webseiten. Sie liefert eine Evidenzbasis, die für die Umsetzung von Empfehlungen im Einklang mit den KI-Grundsätzen der OECD relevant ist, und ergänzt auch die jüngsten Arbeiten der OECD, die sich auf generative KI oder die Rolle der KI für die Wissenschaft konzentrieren.

Die Identifizierung und Analyse von KI-Kerntechnologien ist nicht nur für ein besseres Verständnis der technologischen Basis, die der Verbreitung von KI in der Wirtschaft zugrunde liegt, von Bedeutung, sondern kann auch künftige politikrelevante Forschungsarbeiten in anderen Bereichen inspirieren. Dazu könnten die Zusammenhänge zwischen dem Besitz von KI-Kernpatenten und der Unternehmensleistung und -innovation oder die möglichen Interdependenzen zwischen KI-Kerntechnologien und anderen Technologien, z. B. im Zusammenhang mit dem grünen Wandel, gehören.

1 Introduction

Artificial intelligence (AI) is rapidly transforming economies and societies. AI is often referred to as a general-purpose technology that has the potential to improve economic outcomes, such as productivity, and tackle societal challenges, such as health or climate change (Cockburn, Henderson and Stern, 2018[1]; Brynjolfsson, Rock and Syverson, 2021[2]). Although AI presents significant opportunities for economic growth, it also poses risks for its inclusiveness, for example in terms of inequalities or democratic values (OECD, 2019[3]; 2023[4]; Manheim and Kaplan, 2019[5]; Li, Raymond and Bergman, 2020[6]; König and Wenzelburger, 2020[7]).

Despite the centrality of AI in the current economic and policy debates, empirical analyses identifying and characterising the technological base underlying AI diffusion are still limited.[1] In particular, although there have been recent attempts to delineate subject boundaries and provide an overall definition of what AI is and does, an empirical and operational identification and characterisation of technologies and applications at the core of recent advances in AI are still missing. A major challenge in this context is the fact that technological developments in AI become increasingly embedded in multiple technologies and sectors of the economy, a characteristic common to general-purpose technologies.

Although it is acknowledged that patent data only capture part of technological developments, they contribute to providing a solution to the challenge of defining AI-related technologies, allowing to distinguish core AI from other related developments as well as exploring the evolution of AI innovation. Indeed, they not only provide a quantitative and forward-looking indicator of technological developments, but also offer insights into the qualitative dimension of technological change, especially through their technological fields of detailed technology classification information (Jaffe and de Rassenfosse, 2017[8]; Hall, Jaffe and Trajtenberg, 2001[9]).

This report outlines a new methodology and provides a first exploratory analysis of AI-related core technologies, i.e., technologies and applications at the core of recent advances in AI, based on US patent data.

It first identifies AI-related patents registered at the United States Patent and Trademark Office (USPTO) between 2000 and 2018, i.e., before the more recent boom in generative AI, based on an established list of AI-related keywords as well as International Patent Classification (IPC) and Cooperative Patents Classification (CPC) codes closely related to AI (Baruffaldi et al., 2020[10]). [2] The top cited ones are defined as *"core"* AI patents using counts of their AI-related forward citations, and their characteristics, such as their generality, originality, scope and radicalness, explored and compared to other patent groups. In a next step, the most prevalent AI topics covered in core AI patents as well as their evolution over time are examined. Finally, leveraging a network-based approach, citation patterns of AI patents citing core ones are explored, focusing on different topics. This further allows characterising AI core technologies, hinting at a distinction between more general ones and likely applications, and provides information on their relevance for subsequent AI innovations and on existing technological interdependencies.

Identifying and analysing AI core technologies is relevant not only to better understand the technological base underlying the diffusion of AI across the economy, and its direction, but may also relevantly allow to carry out future policy-relevant research, oriented, for instance, at further exploring the links between holding core AI patents and firm performance or innovation, the diffusion of AI technologies across firm networks, or the possible interdependencies existing between core AI and other technologies, such as those related to the green transition.

The following section introduces the emerging literature relying on patents to study AI innovation while Section 3 presents the data and methodology used to identify and characterise AI-related core technologies. Section 4 provides insights based on those data, exploring the characteristics of core AI patents and their role for subsequent AI innovation, and Section 5 provides some concluding remarks, and points to possible next steps for future analysis.

2 Setting the scene

It is well argued by many scholars that AI is expected to trigger a cascade of complementary innovations across diverse applications and sectors and hence is considered a promising general-purpose technology (Cockburn, Henderson and Stern, 2018[1]; Trajtenberg, 2019[11]; Brynjolfsson, Rock and Syverson, 2021[2]; Crafts, 2021[12]). At the same time, AI is often described as an “invention of a method of invention” (IMI), which may act as a catalyst to generate new inventions that build upon previous knowledge (Griliches, 1957[13]; Cockburn, Henderson and Stern, 2018[1]).

Recent work by Hötte et al. (2023[14]) conceptualises three different views related to the trajectory of AI developments. The first one takes a long-range perspective where AI is considered an outcome of the long-term co-evolution of hardware, software, and networking technologies and hence the latest step towards technological automation (Aghion, Jones and Jones, 2017[15]; Bonaccorsi and Moggi, 2020[16]). Conversely, the mid-range view associates AI with modern forms of computing to perform tasks previously associated with human-like intelligence. Finally, the short-range perspective associates AI with the latest developments in machine learning, narrowed towards deep learning at the expense of other relatively unexplored domains.

To better identify and probe developments in AI technologies, a line of literature that relies on patent information to study AI innovation recently emerged. Hereby the landscape of AI innovation dynamics and its evolution is proxied by patents, which are a widely used, although not fully exhaustive, indicator to measure innovative activity in policy research and the innovation literature (OECD, 2009[17]).

Early analysis by Tseng and Ting (2013[18]) focuses on AI fields related to problem reasoning and solving, machine learning, network structure, and knowledge processing (of the single technological class “Data Processing: AI”) to examine their trends in the US through patent quantity and quality measures. Lately, the literature moved beyond the sole use of technological classes and started exploiting keywords sometimes in combination with more advanced machine learning techniques to further identify AI-related innovation, e.g., by parsing patent text (Damioli, Van Roy and Vertesy, 2021[19]; Giczy, Pairolero and Toole, 2021[20]). In fact, the use of keywords and text mining techniques are increasingly used to capture patents referring to technologies highly relevant to AI (De Prato and Cardona, 2019[21]; European Commission. Joint Research Centre, 2018[22]; Montobbio et al., 2022[23]). In particular, and in line with this work, recent analyses rely on a combination of pre-selected AI-related technological classes and on text mining techniques on IP data (Cockburn, Henderson and Stern, 2018[1]; WIPO, 2019[24]; Iori, Martinelli and Mina, 2021[25]; Baruffaldi et al., 2020[10]; Dernis et al., 2021[26]; Nakazato and Squicciarini, 2021[27]; Santarelli, Staccioli and Vivarelli, 2022[28]).

Regardless of the method employed, the aforementioned studies consistently reveal an acceleration of AI-related innovations – as proxied by patent data – particularly in more recent years. Furthermore, these suggest that AI innovations tend to be highly concentrated within a few firms, with implications for competition policy and regulation (Fujii and Managi, 2018[29]; Hötte et al., 2023[14]; Petit, 2017[30]).[3] However, with AI’s potential of being both a human-enhancing innovation, such as in the case of robot-assisted surgeries, and a human-replacing one, automating an increasingly high degree of tasks that were not previously automatable (OECD, 2023[4]), Trajtenberg (2019[11]) argues that the conventional emphasis of economic policy on the rate of innovation is worth redirecting towards its direction to better capture returns to technical change (see also Acemoglu and Johnson (2023[31]) among others).

When decomposing the determinants of AI technology inventions to help policymakers foster AI innovation, Fujii and Managi (2018[29]) observe a shift in the trend of AI technology developments globally over time. In particular, focusing on the IPC group ”Computer systems”, they find a shift from biological and knowledge-based approaches to patents granted for specific mathematical models and other AI technologies.[4] Ziao and Boschma (2022[32]) instead find that the CPC class “Instruments”, covering technologies related to recognition of data, digital data processing, and computational models, and Information and Communication Technologies (ICTs) are two major knowledge sources cited by AI patent applications, with ICTs (and in particular advanced digital technologies) becoming increasingly significant over time.

The importance of distinguishing between different types of AI innovations is also reflected in other recent empirical contributions. For instance, Grashof and Kopka (2022[33]) find that *AI techniques and methods*, such as algorithms (proxying its IMI characteristics), impact negatively while *AI applications* (proxying general-purpose technology characteristics) impact positively on radical knowledge generation (defined in terms of new combinations of CPCs) in European firms. These relationships vary with firm characteristics, with larger effects of AI techniques and methods in smaller firms and AI applications in larger firms.

This line of work also builds on the wider stream of research using network analyses, often to capture international industry (De Prato and Nepelski, 2012[34]) and academia-industry (Guan and Zhao, 2013[35]) collaboration patterns, or to understand technological development trajectories. For instance, recently Lee et al. (2016[36]) used a network approach to identify co-occurrences of robotics technologies in patents registered with the USPTO and the Korean Intellectual Property Office (KIPO) and found evidence of technological convergence in robotics. Iori et al. (2021[25]) identify AI-related patents granted by the USPTO by searching for CPC codes with a text-based search of keywords in titles and abstracts and exploit (un)directed citations and funding information to examine patents’ impact on follow-on innovations. Petruzzelli et al. (2023[37]) measure the technological impact of AI patents based on forward citations, excluding self-citations, and find that AI patents developed by university-industry collaborations exhibit a lower technological impact compared with those developed by other types of partnership, with AI patents developed by only companies having a higher technological impact than those developed by a universities and research institutes.

However, the empirical contributions distinguishing different AI technologies and looking at their characteristics and evolution over time are still limited, with the majority of studies considering AI’s general-purpose technology nature alone or examining its impact from a macroeconomic standpoint (Brynjolfsson, Rock and Syverson, 2021[2]; Crafts, 2021[12]) without taking into detailed account how it may influence the development process of AI solutions. This work aims at contributing to the existing literature twofold: first, it identifies and characterises highly cited innovations, which are likely to be most relevant for the direction of technological change, focusing on leading AI-related technologies and applications that are at the core of recent advances in AI. Second, it exploits network analyses to help reveal interdependencies between such core technologies and other AI-related innovations, hence helping better understand if and how such technologies may jointly contribute to shape a new knowledge base.

3 Data and methodology

The methodology consists of three main steps designed to: (1) identify AI-related patents, (2) select the sub-group of core AI patents and (3) group them into different topics based on their abstracts' content. These steps are described in detail in this section.

Identifying AI-related patents among all patents

The study focuses on patent applications filed between 2000 and 2018[5] at the USPTO, covering patents issued and pre-grant patent publications. Relative to other patent offices' available data, USPTO data provides the most comprehensive bibliographic data, particularly in terms of citations. Additionally, the USPTO's rules and regulations allow software to be patented while in other patent jurisdictions, such as the European Patent Office (EPO), software becomes patentable only in the context of "Computer implemented inventions" (CII).[6] Therefore, building the analysis on USPTO patents helps uncover core AI techniques from the broader spectrum of AI-related patents. Data derive from the Spring 2022 version of the Worldwide Patent Statistical Database of the EPO, also known as the PATSTAT database, that is further curated within the OECD STI Micro-data Lab infrastructure.

The set of AI-related patents are identified using the approach described in Baruffaldi et al. (2020[10]), which relies on a combination of AI keywords retrieved in the abstract or title of patent documents and on a selection of IPC (and CPC) codes that relate to AI. The list of AI keywords was derived from a text mining approach performed on scientific literature dedicated to AI. Both, the relevant classification codes and keywords are listed in Annex A. The time frame under analysis precedes the more recent boom in generative AI.

Identifying '*core'* AI patents

A patent citation-based approach is applied to distinguish core AI-related patents from remaining AI-related patents. According to the OECD Patent Statistics Manual (2009[17]), "patent and non-patent citations are the references provided in the search report which are used to assess an invention's patentability and help to define the legitimacy of the claims of the new patent applications. As they refer to prior art, they indicate the knowledge that preceded the invention and may also be cited to show the lack of novelty of the citing invention". Patent citations thus serve an important legal function by delimiting the scope of the property rights awarded by the patent. For instance, if patent B cites patent A, patent A represents a piece of existing knowledge upon which patent B builds and over which B cannot have a claim.

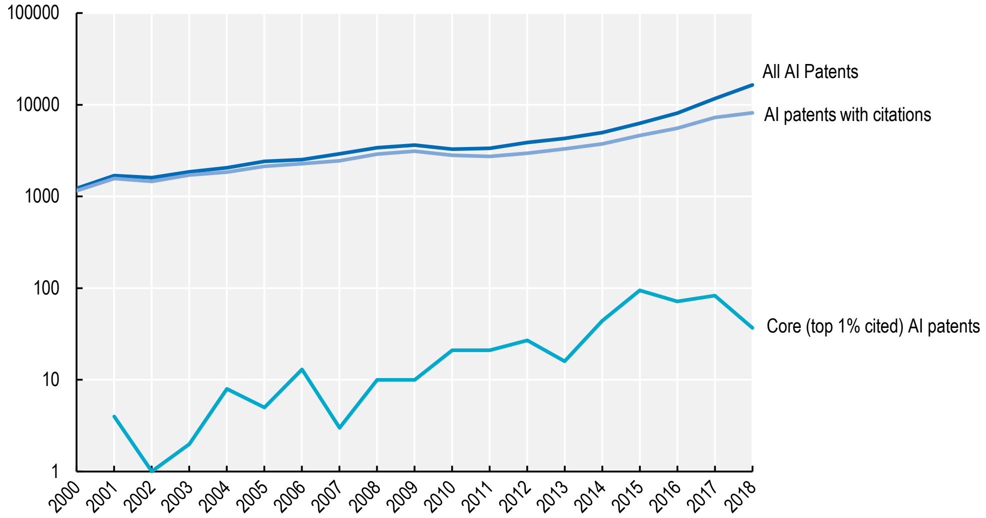
In this work, forward citation linkages are used to assess the technological impact a given AI innovation has for subsequent AI technological developments. Therefore, the analysis of the top highly cited AI patents is intended to help characterise different AI technologies and identify core ones. We build on the hypothesis that the more often an AI-related patent is cited by other AI-related patents, the more developments are built on it and, hence, the more likely it is to be at the core of AI developments, i.e., being relevant for a broad set of other AI-related technologies. The number of forward citations received by a focal patent are consolidated to account for all the citations received by its equivalent patent (e.g. a patent that protects the exact same invention in another country), as explained in Webb et al. (2005[38]).

In order to uncover the core of AI, which was used as seed for other AI-related developments, we identify the top 1 percent highly cited AI patents. In this exercise, no time window was applied to account for the number of citations received, so that important AI patents persistently cited are not ruled out. The number of forward citations an AI patent received is therefore accounted for over the whole period. To ensure comparability over time and across patents, citation counts are then normalised by the average number of forward citations received by patents from a reference cohort (i.e., patents in the same technology area – measured using the WIPO IPC-technology concordance [7] - and same filing year). Additionally, to control for the representativeness of individual inventions, continuations of USPTO patents are excluded from the sample. [8]

The top 1 percent set of patents that feature the largest normalised forward citation counts are flagged as "core AI" patents. Over the period 2000-18, more than 70% of AI-related patents have been cited by other patents. Among those, as shown in Figure 3.1, the methodology described above identifies 472 core AI patents between 2000 and 2018, with an acceleration of patent activity in recent years.

Figure 3.1. AI-related patents and core AI patents, 2000-18

Total AI-related patents, patents with forward citations, and core (top 1% cited) AI patents, logarithmic scale



Note: Data refer to patent applications filed at the USPTO (patents issued and pre-grant publications) that protect AI technologies. The sample of top 1% cited patents relies on normalised counts of forward citations.
Source: OECD, STI Micro-data Lab: Intellectual Property Database, http://oe.cd/ipstats, June 2023.

Grouping core AI patents into different topics

In order to better understand which AI technologies other developments build on, core AI patents are grouped into specific AI-related topics. This is done by identifying AI-related keywords that appear most frequently in abstracts of core AI patents. Patent abstracts contain key information on the invention, by summarising the content of the document in a concise manner. The list of AI-related keywords used follows the one developed by Baruffaldi et al. (2020[10]).

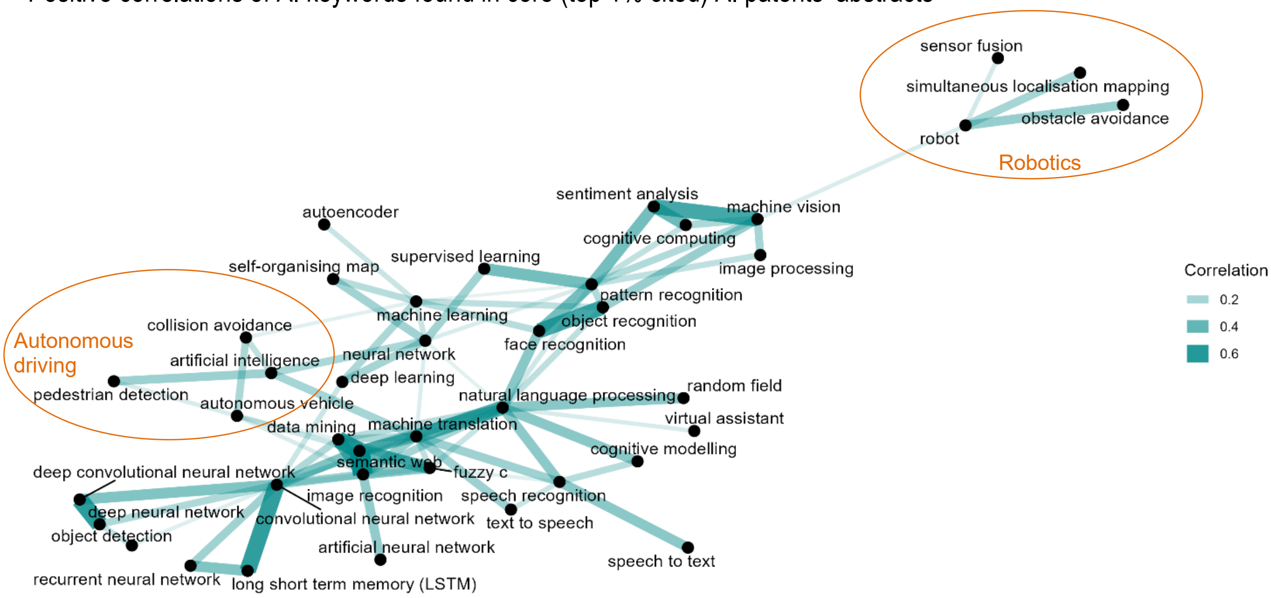
Those keywords are grouped into 11 representative AI topics relying on the support of AI experts in an iterative process and a pairwise correlation analysis. The latter reflects how often different keywords appear together in the abstract of each core AI patent relative to how often they appear separately. Figure 3.2 illustrates such correlations graphically in a network and highlights how each bundle of topic-specific keywords emerges. More specifically, each keyword is represented by a black *node* whereas the correlation between keyword pairs is reflected by the thickness of *edges* (or links) between the nodes. Therefore, the more frequently two keywords are mentioned together in an abstract of a core AI patent in relative terms, the thicker the edge is. The distance between nodes has no meaningful interpretation in this network.

AI topics identified include: *algorithms*, *autonomous driving*, *chatbot*, *computer or image vision*, *feature engineering*, *general AI,*[9] *natural language processing*, *networks* (such as *deep learning*), *recognition or detection*, *robotics*, and *speech*. A detailed list of the underlying keywords linked to each topic is provided in Annex A.

Although the purpose of this grouping exercise is mainly to ease visualisation of the networks which follows, the pairwise correlation of keywords shown in Figure 3.2 also provides first evidence that some topics, likely related to AI applications, e.g. autonomous driving and robotics, tend to appear self-contained within core AI patents.

Figure 3.2. Pairwise correlation of AI keywords, core AI patents, 2000-18

Positive correlations of AI keywords found in core (top 1% cited) AI patents' abstracts



Note: The figure displays the frequency with which AI-related keywords occur together rather than separately in an AI core patent's abstract. AI-related keywords are presented by nodes whereas the relative frequency is reflected by the thickness of the edges between the nodes. The distance between nodes has no meaningful interpretation. The sample refers to the top 1% cited patent applications (normalised counts of forward citations) filed at the USPTO (patents issued and pre-grant publications) that refer to AI-related technologies. It relies on a sub-sample (351 patents) of core AI patents, for which a pre-defined AI keyword was available. The remaining core AI patents (121 patents) were identified based on AI-related IPC/CPC codes alone and hence are not reported in this figure.
Source: OECD, STI Micro-data Lab: Intellectual Property Database, http://oe.cd/ipstats, June 2023.

To further analyse the content of each AI patent in more detail, a main topic out of all the identified AI topics was allocated to each patent, based on the frequency of topic-specific keywords. For example, if a patent's abstract contains three *speech*- and one *chatbot*-related AI keywords, this patent's main topic is set to *speech*. In the rare occasion in which no prevalent topic could be identified, the main topic is selected randomly out of all the topics mentioned. The direct allocation of an AI main topic, based on the set of pre-defined keywords, was possible for the great majority (351 out of the 472) of core AI patents.

The remaining 121 core AI patents were originally identified as AI-related solely based on their AI-related IPC/CPC codes, as none of the pre-defined AI-related keywords listed in Baruffaldi et al. (2020[10]) appear in their abstract. Therefore, they could not be directly allocated to a specific AI main topic. For those, main topics were imputed based on the similarity of their IPC codes combinations with those observed in patents that contain pre-defined AI keywords (and hence AI topics). More specifically, the main topic most often associated with a specific combination of four IPC codes was first used to impute AI main topics for patents without any keywords. The remaining list of patents without topics were then allocated to the topic that appears most frequently within combinations of three-IPC codes, and, in case the combination was not identified in other patent documents, the most frequent topic appearing in a combination of two IPC codes was used for the imputation.[10] As a result, all but nine (2%) core AI patents were allocated a specific AI main topic. The same exercise was also applied to AI patents that cite any of the identified core AI patents, resulting in 99% of citing AI patents being allocated a specific AI main topic.

4 Exploring the characteristics of core AI patents

Following the methodology discussed in the previous section, the analysis first characterises core AI patents according to their scope, originality, generality, and radicalness. This helps possibly distinguish between more general core technologies and likely applications. It then zooms in on the content of core AI patents' abstracts, exploring the main topics characterising such patents and their evolution over time. Finally, leveraging a network-based approach, it explores citation patterns of AI patents citing core ones focusing on different topics. This allows further characterising AI core technologies as well as provides information on their relevance for subsequent AI innovations and on existing technological interdependencies.

Measuring the quality of core AI patents: scope, originality, generality and radicalness

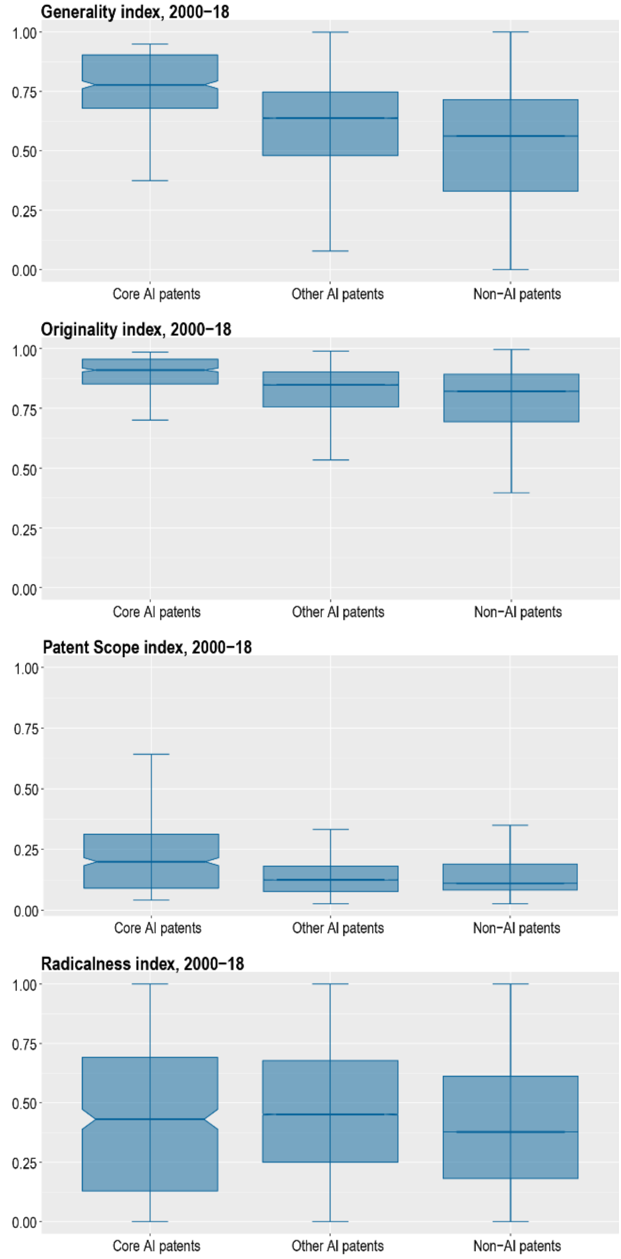
A first exploration relates core (top cited) AI patents to several patent quality measures reflecting their technological scope, originality, generality, and radicalness over the period 2000-18. A brief summary of each measure is provided in Box 4.1 below, for a more detailed description please see Squicciarini et al. (2013[39]).

Initial findings presented in Figure 4.1. reveal that, compared to their less cited counterparts or other patents in non-AI technologies, core AI patents *cite* patents belonging to a *wider* array of technology fields, i.e., they rely on a broader range of diverse knowledge sources, as indicated by the higher originality index. At the same time, they *are also cited* by patents belonging to a *wider* range of technology fields, as measured by the higher generality index. Core AI patents also have a broader technological scope. However, although AI patents compared to other non-AI patents are technologically more radical, there is no significant difference between the core AI and the less cited counterparts. Therefore, in this dimension core AI patents do not necessarily differ more from the predecessors they rely on than other AI patents do. These observed differences are overall evident even when comparing core AI patents to highly cited non-AI patents (presented in Annex Figure A B.1).[11]

When comparing the two periods 2000-09 and 2010-18, as shown in Annex Figure A B.2, it becomes evident that AI patents, and in particular core AI patents, became more original and more general over time. Therefore, core AI patents increasingly appear to rely on a larger number of former innovations and, at the same time, are increasingly relevant for a larger number of later innovations in different technology fields, providing first evidence in line with their general-purpose technology nature. Moreover, the figure also reveals that they became broader in scope over time. As technological breadth in a firm's portfolio is known to increase its valuation or venture capital funding (Lerner, 1994[40]; Caviggioli et al., 2020[41]), the higher scope levels further suggest that core AI patents also have a higher technological and likely higher market value, especially more recently. However, the seemingly higher level of radicalness of AI patents (with respect to non-AI ones) tends to be driven by earlier periods. AI patents, and in particular top AI patents, used to cite previous patents in classes other than the ones they were in more frequently. More recently, they are increasingly associated with lower radicalness, similar to the levels of non-AI patents.

Figure 4.1. Patent quality indicators, different patent groups, 2000-18

Technological generality, originality, scope and radicalness of core (top 1% cited) AI, other AI and non-AI patents



Note: The "core AI patents" sample (i.e. the core AI set covering up to 472 patents) refers to the top 1% cited patent applications (normalised counts of forward citations) filed at the USPTO (patents issued and pre-grant publications) that protect AI technologies. The "other AI patents" sample refers to the remaining AI patents identified (non-core AI patents). "Non-AI patents" relate to patents protecting non-AI technologies. Each indicator is constructed based on Squicciarini et al. (2013[39]), with more detailed description in Box 4.1; the patent scope index is normalised to the maximum value observed in patents belonging to the same cohort (filing date and WIPO technology). By construction, all indicators presented range between zero and one. Annex B presents each indicator but further distinguishing top cited non-AI patents, which are the top 1% highly cited patents in non-AI technologies, from the remaining non-AI patents.
Source: OECD, STI Micro-data Lab: Intellectual Property Database, http://oe.cd/ipstats, June 2023.

Box 4.1. Four indicators measuring the quality of core AI patents

The technological and/or economic value of a patent can be captured using a variety of measures, as described in Squicciarini et al. (2013[39]). The measures presented here rely on the information contained in the patent document, in particular the technological classes (here IPC classes) that are allocated to each patented invention, and combined, for three of them, with forward or backward citations. By construction, all indices have values between 0 and 1.

A *patent's scope* is defined as a normalised index based on the number of distinct 4-digit subclasses of IPCs the innovation is allocated to. For each patent document, the IPC counts are then normalised using the maximum value observed in patents belonging to the same cohort (i.e., patents in the same filing date and technology domain). Large patent scope index can be associated with higher technological and market value of a given patent.

The *patent generality index* relies on a modification of the Hirschman-Herfindahl Index and exploits, for a given patent, information on the number and distribution of the forward citations it received and the technology classes of these citing patents. This measure considers all IPC classes contained in the citing patent and observes their technological concentration. High indices are observed for patents that are cited by subsequent patents belonging to a wide range of technology fields, and hence, apply to a greater spectrum of technologies.

The *patent originality index* follows a similar logic to the one used to construct the generality indicator, but it relies on backward citations instead of forward ones. It observes the breadth of technologies and knowledge that were exploited to develop an invention. Inventions that combine a diversity of technologies are likely to be more original than patents that rely on a small set of technologies.

Finally, *radicalness* of a patent is observed by comparing the technological content of a patent with that of its citing patents. It is here defined following Shane (2001[42]): it is measured as a time invariant count of the number of IPC technology classes in which the patents cited by the given patent are, but in which the patent itself is not classified. The more a patent cites previous patents in fields other than the ones they rely on, the more the invention should be considered radical.

Exploring core AI innovations thematically

Zooming in on the content of core AI patents' abstracts reveals that some topics, such as *general AI*, *robotics*, *computer or image vision* and *recognition or detection*-related technologies, are consistently at the core of AI developments. These technologies are at the intersection of various disciplines, such as computer science, engineering, and mathematics, with a wide range of practical applications. For instance, the term *robotics* refers to a technology associated with the automation of tasks, which historically have been of a repetitive usually routine nature. More recently, robots are applied to broader domains, for instance they are used in highly specific hazardous environments or for human-enhancing tasks, such as robot-assisted surgeries, some of which require less and less human intervention.

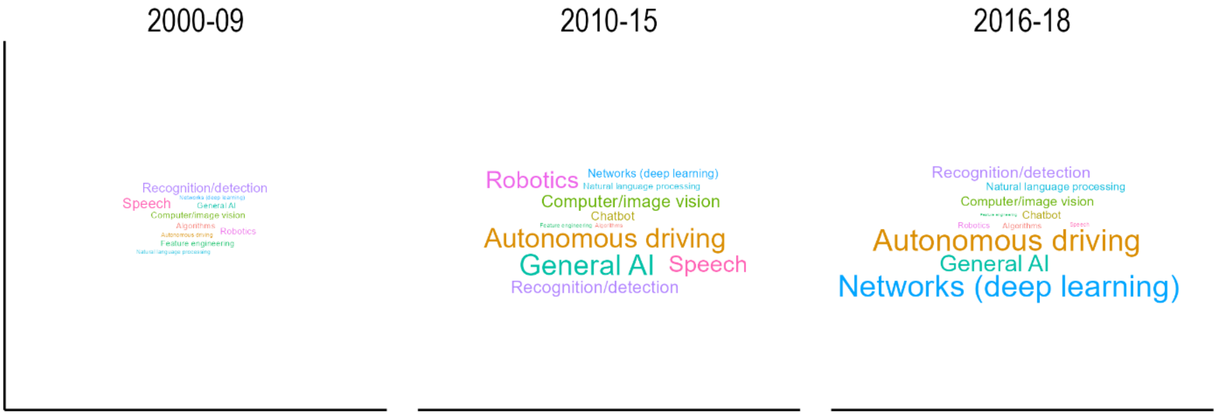
In comparison, technologies related to *feature engineering*, *chatbots*, *algorithms*, *natural language processing* and *speech* appear less prevalent over the period 2000 to 2018. This observation might seem surprising given the current AI landscape with the emergence of generative AI technologies, which is largely a result of substantial progress in the fields of *chatbots, natural language processing* and *speech*. However, these advancements may not be captured in this analysis because they are primarily based, among other factors, on transformer models which came to prominence just in 2017, i.e., the end of the period analysed here (Vaswani et al., 2017[43]).

Technologies that gained in significant importance in more recent years include *autonomous driving* as well as *networks*, such as *deep learning*, as shown in Figure 4.2. [12] The former can be partly attributed to recent advancements in sensor technologies and high-resolution cameras coupled with an increasing focus on safety, which spurred R&D and hence improved the perception capabilities of autonomous vehicles. The latter is likely a result of the rising availability of large-scale data and improvements in computing power, which enabled the acceleration in training deep learning models.

However, it is noteworthy that patents are just one aspect of the AI development ecosystem and hence less prevalent technology fields are not necessarily less developed. Instead, firms might prefer to maintain their competitive advantage by protecting these technologies as trade secrets without disclosing the details of their methods. Conversely, some advancements in chatbot and speech technologies come from open-source contributions.

Figure 4.2. Word clouds of AI main topics, 2000-18

AI main topics of core (top 1% cited) AI patents' abstracts, by sub-period



Note: The size of the words reflects the frequency with which each main topic appears in the sample, which refers to the top 1% cited patent applications (normalised counts of forward citations) filed at the USPTO (patents issued and pre-grant publications) that refer to AI technologies. The sample relies on a sub-sample (463 patents) of core AI patents (463 patents), for which an AI main topic was directly available or imputed.
Source: OECD, STI Micro-data Lab: Intellectual Property Database, http://oe.cd/ipstats, June 2023.

Exploring the role of core AI patents for subsequent innovations

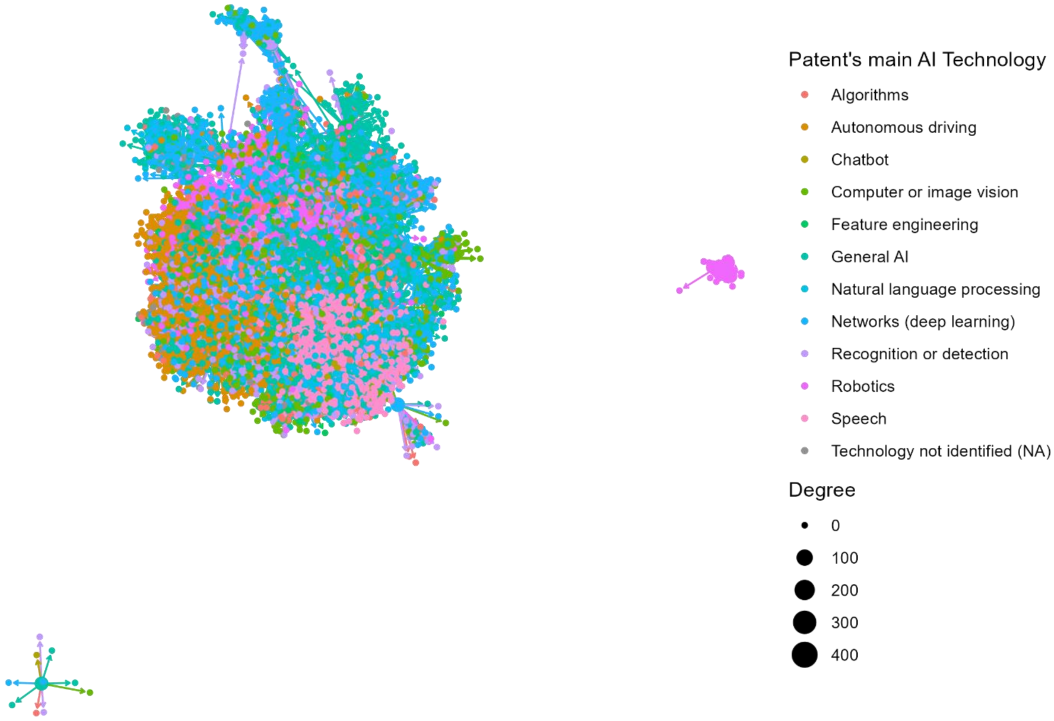
The previous section revealed the growing importance of some fields, such as *autonomous driving* and *networks*, at the expense of other domains over time. To gain deeper insights into how specific fields influence the trajectory of AI advancements over time, it is necessary to not only leverage the information provided in core AI patents but also to draw upon the insights supplied in their citing AI-related patents. Citations networks can help illustrate visually the evolution and flow of knowledge between these patents, pinpoint possible convergences and intersections of AI fields and offer additional support for the observations brought to light by the aforementioned indicators of patent quality.

Exploring citation networks, presented in Figure 4.3, shows that core AI technologies spur innovation in their own respective field, but more often tend to do so in other fields. In particular, the figure portrays the identified core AI patents as focal nodes with their citing AI patents as surrounding nodes in the network. The edges linking each focal node with their surrounding nodes are directed and reflect the flow of knowledge going from the core AI patent to the corresponding citing AI patent. In contrast to Figure 3.2, the edges here are all uniformly thin and hence neither the thickness nor the length of the edges (i.e. the distance between nodes) has any meaningful interpretation in this network. Furthermore, each node is colour-coded by the AI patent's main AI topic and the edges by the citing AI patent's main topic, providing first insights on the interdependencies existing between different topics.

According to Figure 4.3, core AI technologies appear relevant for a wide number of later AI innovations and, at the same time, also seem to rely on past core AI technologies. [13] At first glance, some technology clusters emerge, such as *autonomous driving* (coloured in orange), *speech* (pale pink) or *robotics* (bright pink), which seemingly contribute to developments in their own respective field. Conversely, other technologies related to, for instance, *networks* (bright blue), *recognition or detection* (purple), *general AI* (dark green) and *computer or image vision* (dark green) rather appear to build on technologies other than their own, possibly hinting at a wider-ranging potential for more diverse applications.

Figure 4.3. Core AI patents with AI citations, 2000-18

AI main topics of core (top 1% cited) AI patents and their AI citations

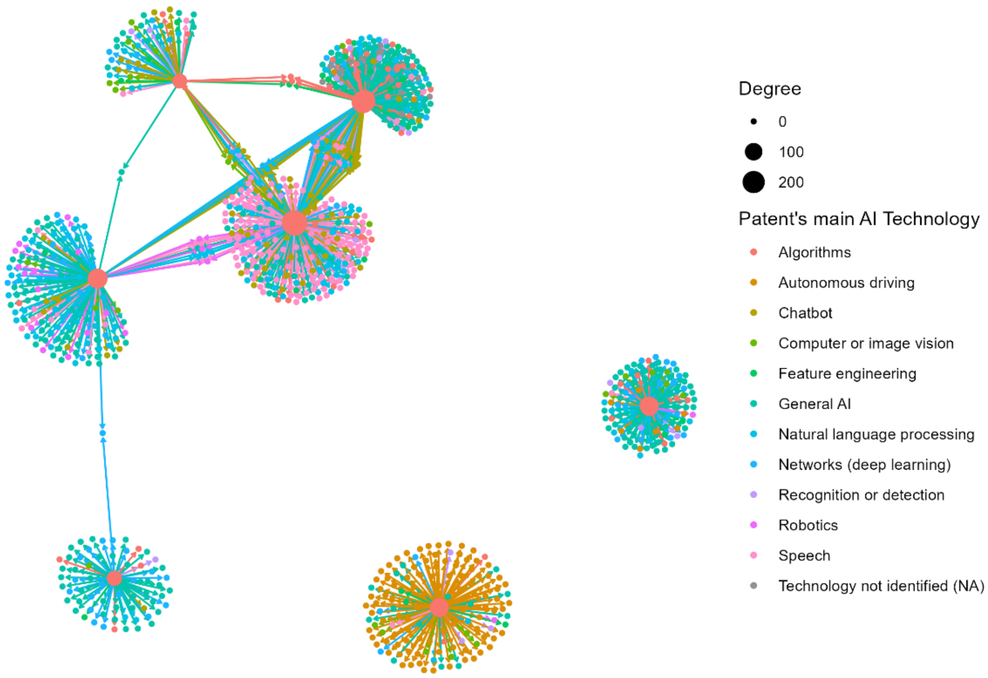


Note: The figure shows citation networks with the top 1% cited patent applications (normalised counts of forward citations) filed at the USPTO (patents issued and pre-grant publications) that refer to AI-related technologies as focal nodes. Surrounding nodes reflect other AI-related patents that are citing core AI patents. Neither the thickness nor the length of the edges has any meaningful interpretation. Nodes are colour-coded according to their AI main topics. AI topics have been imputed for patents, which were identified as AI-related based solely on IPC codes, based on patents with similar IPC characteristics. The few AI patents without any topics (even after imputation) are colour-coded grey.
Source: OECD, STI Micro-data Lab: Intellectual Property Database, http://oe.cd/ipstats, June 2023.

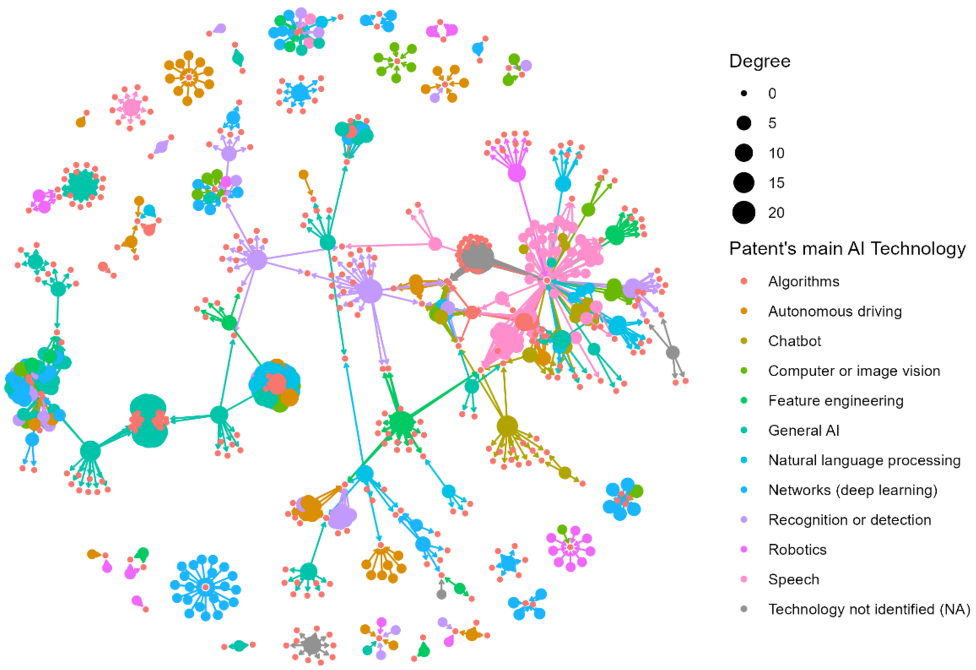
In what follows, the full network presented in Figure 4.3 is divided into separate layers, where each layer only contains core AI patents of one specific AI main topic together with its citing AI patents. This exercise provides further evidence that core AI technologies tend to be cited by AI patents belonging to AI technologies *other* than the one applied for in the core AI patent. For instance, the citation network of *algorithms*-related core AI patents with (any) AI citations in the top Panel of Figure 4.4 highlights their relevance for innovations associated with *autonomous driving, general AI, natural language processing, robotics, speech*, etc. This is further supported when looking at the reverse case and zooming in on any core AI patents and their *algorithm*-related AI citations, as shown in the bottom Panel of Figure 4.4. Given the different perspectives with which *algorithm*-related patents are viewed in this exercise, edges in the top Panel are colour-coded according to the AI main topic of the citing AI patent (the core AI patents are by default *algorithms*), while edges in the bottom Panel are colour-coded according to the AI main topic of the core AI patent (the citing AI patents are by default *algorithms*).

Figure 4.4. Algorithms-related AI patents and their citations, 2000-18

Algorithms-related core AI patents with AI citations



Core AI patents with Algorithms-related AI citations



Note: The figure shows citation networks with a subset of the top 1% cited patent applications (normalised counts of forward citations) filed at the USPTO (patents issued and pre-grant publications) that refer to algorithms as focal nodes (top Panel) and citing nodes (bottom Panel). Neither the thickness nor the length of the edges has any meaningful interpretation. Nodes are colour-coded according to their AI main topics. AI topics have been imputed for patents, which were identified as AI-related based solely on IPC codes, based on patents with similar IPC characteristics. The few AI patents without any topics (even after imputation) are colour-coded grey. Edges are colour-coded according to the AI main topic of the citing AI patent (top Panel) or the core AI patent (bottom Panel).
Source: OECD, STI Micro-data Lab: Intellectual Property Database, http://oe.cd/ipstats, June 2023.

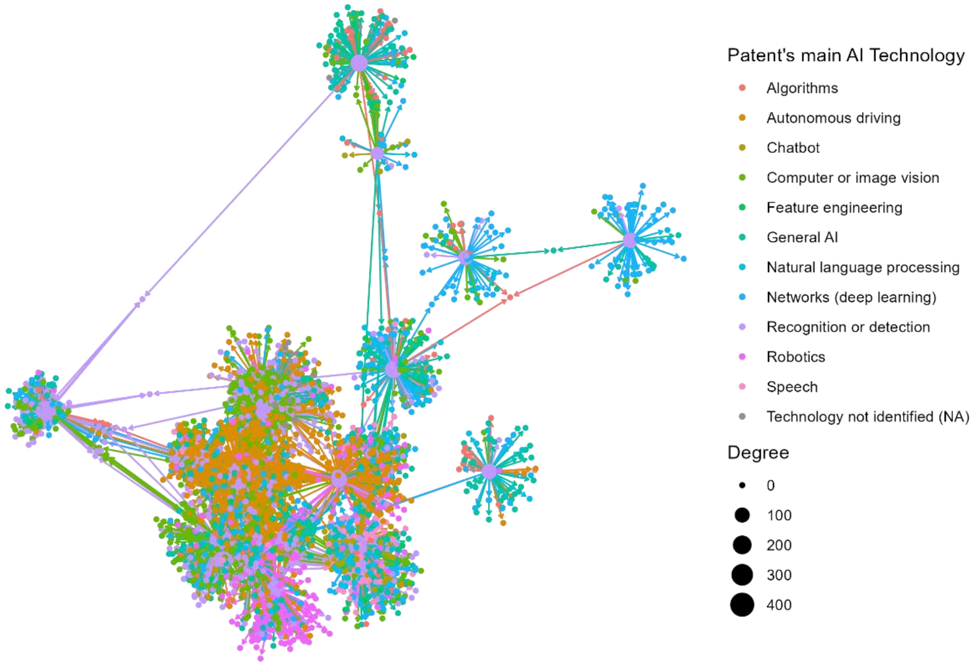
These initial findings are further confirmed when exploring another technology field separately. Figure 4.5 shows the citation network of core AI patents that have *recognition or detection* technologies as the main topic (top Panel). These technological fields also appear relevant for many innovations other than the ones associated with *recognition or detection* itself, and include *autonomous driving, robotics, computer or image vision* and *networks (deep learning).* Given that detection or recognition are crucial in perceiving and understanding the environment, it is not surprising that innovations related to autonomous driving and robotics, which both often rely on sensors to detect obstacles and interact with their surroundings intelligently, build on these technological fields. Equally deep learning networks, especially convolutional neural networks, have been revolutionary in image recognition tasks and are at the forefront of advancements in computer vision with substantial applications in autonomous driving and robotics.

When looking at the reverse case again and zooming in on any core AI patents and their *recognition or detection*-related AI citations, as shown in the bottom Panel of Figure 4.5, we find further evidence that *recognition or detection* technologies not only appear relevant for a wide range of later AI innovations (as discussed above) but also that they simultaneously seem to rely on past core AI technologies beyond their own field.

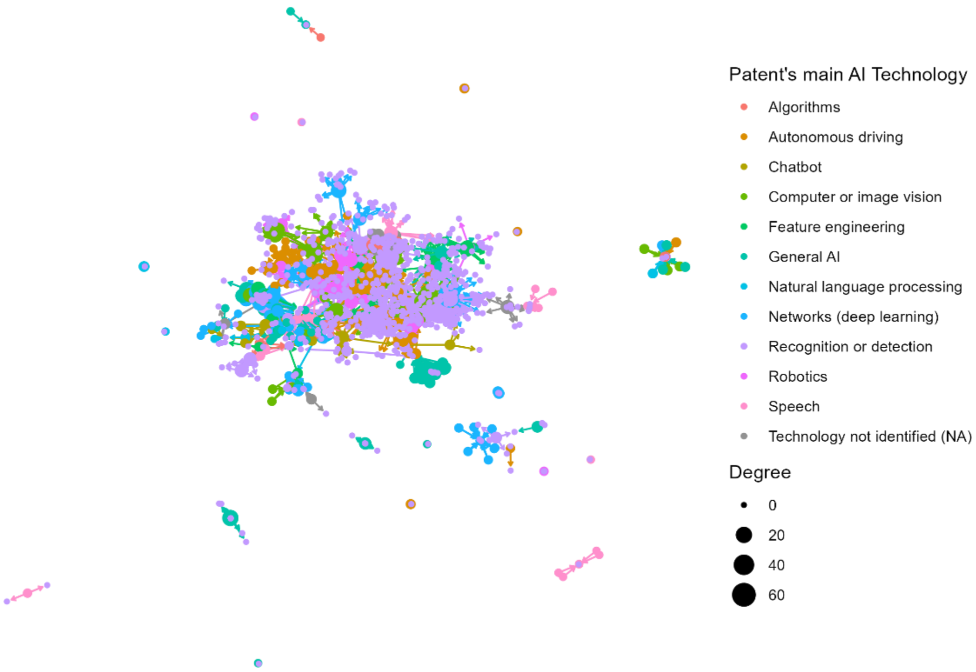
However, these observations are not applicable to all core AI technologies under analysis. In fact, when narrowing down on *autonomous driving* and *robotics*, the network analysis suggests that these technology fields might rather be considered AI applications, even if relying on other core AI technologies beyond their own field. Indeed, as shown by the top Panels of Figure 4.6 and Figure 4.7, core AI patents in *autonomous driving* and *robotics*, respectively, tend to rather contribute to future innovations in their own field, as they seemingly have their own relatively distinct citations clusters.

Figure 4.5. Recognition- or Detection-related AI patents and their citations, 2000-18

Recognition- or Detection-related core AI patents with AI citations



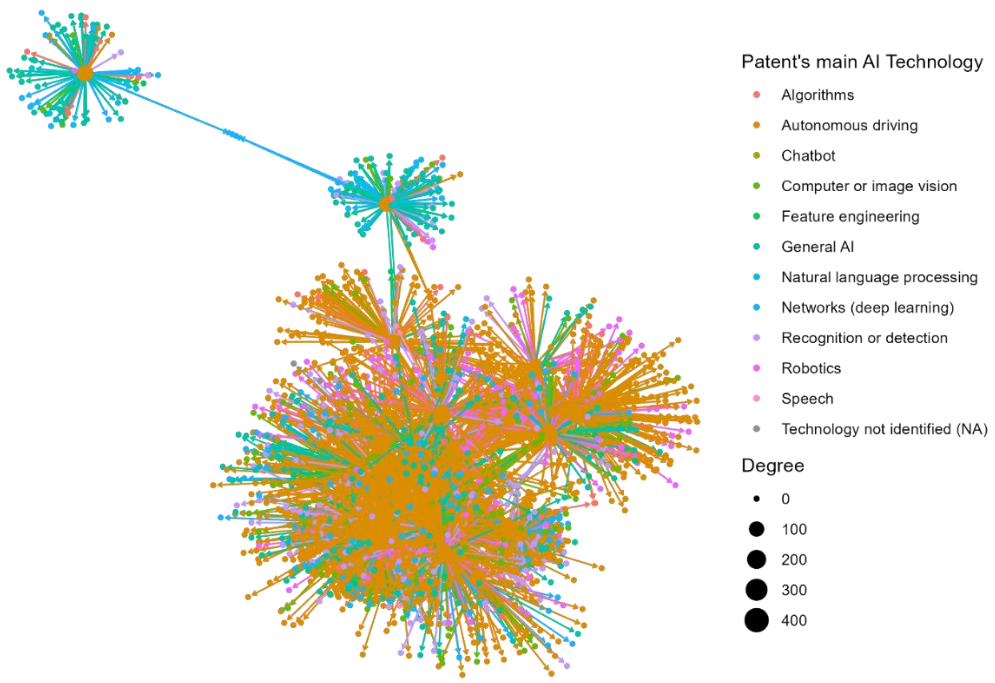
Core AI patents with Recognition- or Detection-related AI citations



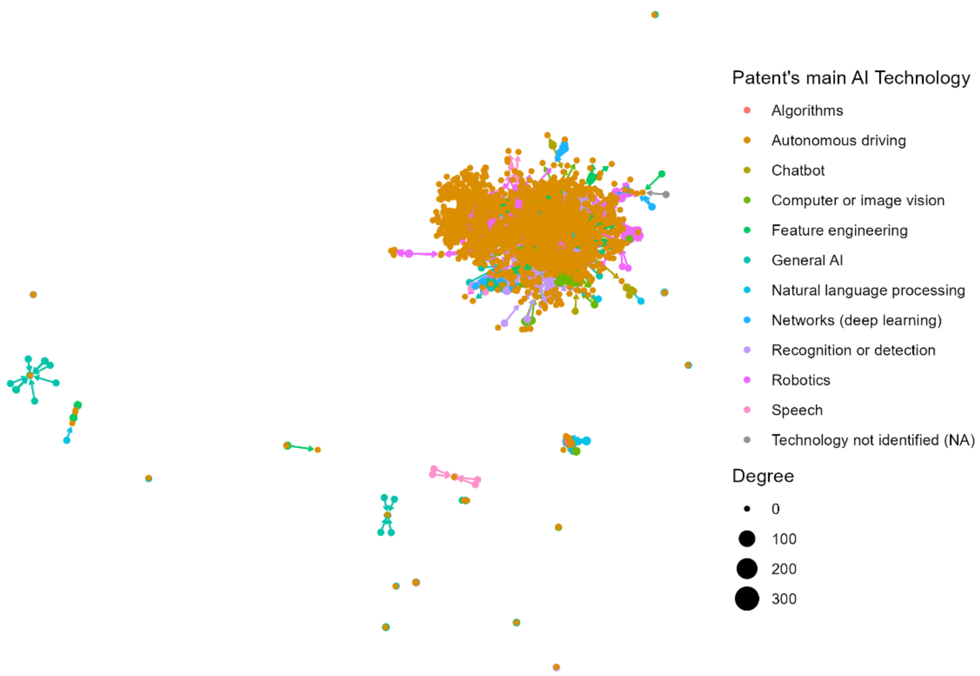
Note: The figure shows citation networks with a subset of the top 1% cited patent applications (normalised counts of forward citations) filed at the USPTO (patents issued and pre-grant publications) that refer to recognition or detection as focal nodes (top Panel) and citing nodes (bottom Panel). Neither the thickness nor the length of the edges has any meaningful interpretation. Nodes are colour-coded according to their AI main topics. AI topics have been imputed for patents, which were identified as AI-related based solely on IPC codes, based on patents with similar IPC characteristics. The few AI patents without any topics (even after imputation) are colour-coded grey. Edges are colour-coded according to the AI main topic of the citing AI patent (top Panel) or the core AI patent (bottom Panel).
Source: OECD, STI Micro-data Lab: Intellectual Property Database, http://oe.cd/ipstats, June 2023.

Figure 4.6. Autonomous Driving-related AI patents and their citations, 2000-18

Autonomous Driving-related core AI patents with AI citations



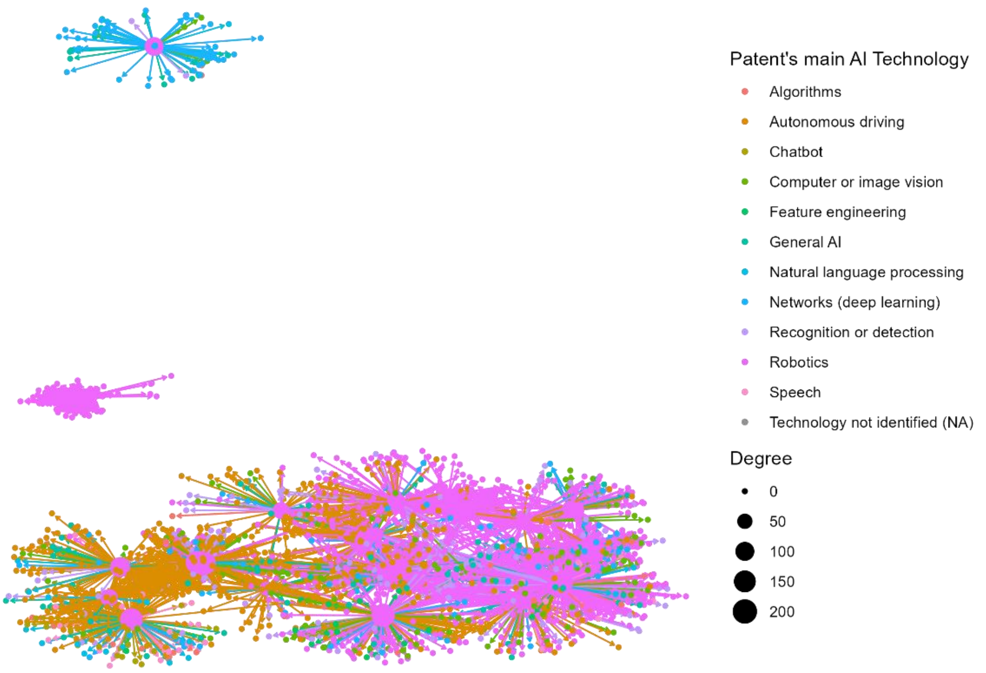
Core AI patents with Autonomous Driving-related AI citations



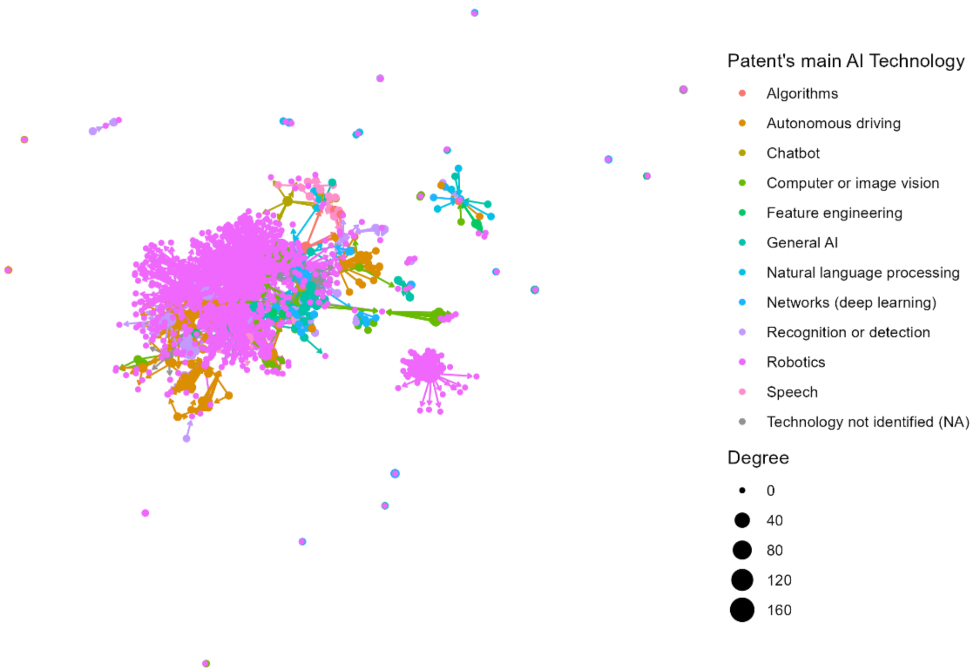
Note: The figure shows citation networks with a subset of the top 1% cited patent applications (normalised counts of forward citations) filed at the USPTO (patents issued and pre-grant publications) that refer to autonomous driving as focal nodes (top Panel) and citing nodes (bottom Panel). Neither the thickness nor the length of the edges has any meaningful interpretation. Nodes are colour-coded according to their AI main topics. AI topics have been imputed for patents, which were identified as AI-related based solely on IPC codes, based on patents with similar IPC characteristics. The few AI patents without any topics (even after imputation) are colour-coded grey. Edges are colour-coded according to the AI main topic of the citing AI patent (top Panel) or the core AI patent (bottom Panel).
Source: OECD, STI Micro-data Lab: Intellectual Property Database, http://oe.cd/ipstats, June 2023.

Figure 4.7. Robotics-related AI patents and their citations, 2000-18

Robotics-related core AI patents with AI citations



Core AI patents with Robotics-related AI citations



Note: The figure shows citation networks with a subset of the top 1% cited patent applications (normalised counts of forward citations) filed at the USPTO (patents issued and pre-grant publications) that refer to robotics as focal nodes (top Panel) and citing nodes (bottom Panel). Neither the thickness nor the length of the edges has any meaningful interpretation. Nodes are colour-coded according to their AI main topics. AI topics have been imputed for patents, which were identified as AI-related based solely on IPC codes, based on patents with similar IPC characteristics. The few AI patents without any topics (even after imputation) are colour-coded grey. Edges are colour-coded according to the AI main topic of the citing AI patent (top Panel) or the core AI patent (bottom Panel).
Source: OECD, STI Micro-data Lab: Intellectual Property Database, http://oe.cd/ipstats, June 2023.

5 Concluding remarks and next steps

The exploratory insights presented in this report appear relevant to better understand the technological base underlying recent stages of the digital transformation up to 2018. The analysis has outlined a methodology to define core AI patents and explored their characteristics, including their main topics and citation patterns.

Results indicate that core AI patents are more original, i.e., rely on a broader range of technology fields, more general, i.e., are cited by patents belonging to a wider range of fields, and tend to be broader in technological scope, suggesting that they also have a higher technological and likely higher market value, compared to their less cited counterparts or other patents in non-AI technologies. In fact, they became more original and general over time, especially relative to other non-AI patents, providing first evidence in line with the general-purpose technology nature of AI. Although AI patents are technologically more radical when compared to other non-AI patents, they do not significantly differ with respect to their less cited AI-related counterpart. Interestingly, they used to be more radical relative to other patents in the earlier stages of development, implying that they now cite previous patents in classes other than the ones they were in less frequently.

Zooming in on the content of core AI patents' abstracts reveals that some topics, such as *general AI*, *robotics*, *computer or image vision* and *recognition or detection*-related technologies, are consistently at the core of AI developments. Technologies that gained in significant importance in more recent years include *autonomous driving* as well as *networks*, such as *deep learning.*

Exploring citation networks of AI patents citing core ones has allowed to further explore the direction of AI innovation and the existing interdependencies among different AI technologies. Core AI technologies tend to spur innovation beyond their own respective field. However, some of them – likely related to AI applications, such as *autonomous driving* or *robotics* – appear to increasingly contribute to developments in their own field.

This analysis is complementary to recent and ongoing OECD work analysing AI innovation and diffusion using different sources of data, including official firm-level surveys (Calvino and Fontanelli, 2023[44]), surveys carried out by the OECD in collaboration with other organisations (e.g., the OECD/BCG/INSEAD survey, see OECD (2024 forthcoming[45]), or the OECD AI surveys of employers and workers, see also Lane et al. (2023[46])), online job postings (Squicciarini and Nachtigall, 2021[47]; Samek, Squicciarini and Cammeraat, 2021[48]; Green and Lamby, 2023[49]; Borgonovi et al., 2023[50]), data from internet websites (Dernis et al., 2023[51]), case studies (Milanez, 2023[52]), or a combination of sources (Calvino et al., 2022[53]). It provides evidence base that is relevant for implementing recommendations consistent with the OECD AI Principles and is also complementary to recent OECD work focusing on generative artificial intelligence (e.g., Lorenz, Perset and Berryhill (2023[54])) or on the role of AI for science (OECD, 2023[55]).

Considering its exploratory nature, based on and taking into account feedback from the OECD Committee on Industry, Innovation and Entrepreneurship and Working Party on Industry Analysis, this analysis could be extended into several directions. Future research may further investigate the extent to which some core AI technologies may differ from others, and how innovative activity – beyond AI patents – builds upon core AI technologies.

Future analysis may also be oriented at further exploring the links between core AI patenting and firm performance or innovation, or at assessing the possible interdependencies existing between core AI and other technologies, such as those related to the green transition.

Finally, future work may refine the methodology currently used to identify AI-related patents, expanding it beyond the use of keywords and technology classes, but combining those with a large-scale language processing approach, in which for instance a subset of labelled AI-related patents could be used to train a model that is able to classify the rest of the documents. This may also allow to further scale this line of research across countries using timely data, which may also cover periods after the recent boom in generative AI.

Endnotes

[1] Examples of recent literature using patent data to probe emerging AI technologies and their evolution are discussed in Section 2.

[2] The CPC is an extension of the IPC and is jointly managed by the European Patent Office (EPO) and the USPTO, with approximately 250,000 classifications.

[3] Hötte et al. (2023[14]) present an exception. When comparing different approaches in identifying AI innovations, they find different trends over time, with innovations found in scientific citations to AI research (reflecting the academic origins of AI) or USPTO patents (capturing the widespread use of AI in other inventions) highlighting a slow-down in recent years. This contrasts with the continued growth of AI innovations observed when relying on keywords, such as natural language processing, robotics, or neural networks (reflecting recent progress in dominant subject areas) or WIPO patents. While the USPTO classifies AI innovations using a trained ML classifier on patent text and citations (Giczy, Pairolero and Toole, 2021[20]), WIPO combines a set of keywords with computer-specific technological classification codes (WIPO, 2019[24]).

[4] One possible explanation put forward by the authors is the increasing complexity and range of AI technologies making it challenging to allocate patents into specific technology groups.

[5] The analysis is limited to the 2000-18 period to allow for a sufficient time span for citations to accumulate. Patents from 2019 onwards are partly available but are not considered for the analysis here due to truncation.

[6] Computer implemented inventions are defined as 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

[7] Patents are uniquely allocated to 35 technology fields as described in Schmoch (2008[56]), on the basis on the maximum number of IPC codes belonging to a given technology field in the patents. When a patent equally belongs to different technology fields, a random allocation is applied.

[8] Under the US patent law, a continuing patent application is a patent application that follows, and claims priority to, an earlier-filed patent application within the office.

[9] This category encompasses a wide range of concepts and technologies within the broader domain of AI, machine learning and intelligent systems, such as ambient or computational intelligence, cluster analysis, un-, semi- and supervised learning, or recommender system.

[10] Results are qualitatively very similar when restricting the sample to core AI patents solely identified through AI keywords.

[11] A t-test also confirms that the generality, originality and radicalness indices are significantly higher for the core AI than for other top cited non-AI patents. Concerning the patent scope indicator, although its median is higher for core AI patents, differences in averages are not statistically significant.

[12] As previously highlighted, results are qualitatively very similar when restricting the sample to core AI patents solely identified through AI keywords. This is the case when taking into account any AI topic mentioned in the abstract as well as when focusing on patents' AI main topic.

[13] Although not directly visual here due to the large number of patents displayed, when restricting the network to only core AI patents, it becomes evident that more recent core AI technologies also seem to rely on past ones.

References

Acemoglu, D. and S. Johnson (2023), *Power and Progress: Our Thousand-Year Struggle Over Technology and Prosperity*, Hachette Book Group. [31]

Aghion, P., B. Jones and C. Jones (2017), *Artificial Intelligence and Economic Growth*, National Bureau of Economic Research, Cambridge, MA, https://doi.org/10.3386/w23928. [15]

Agrawal, A., J. Gans and A. Goldfarb (eds.) (2019), *Artificial Intelligence as the Next GPT: A Political-Economy Perspective*, http://www.nber.org/chapters/c14025. [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, Paris, https://doi.org/10.1787/5f65ff7e-en. [10]

Bonaccorsi, P. and M. Moggi (2020), "The long wave of the Internet", *Laboratory of Economics and Management (LEM) Papers Series*, Vol. 2020/26,/Sant'Anna School of Advanced Studies, Pisa, Italy. [16]

Borgonovi, F. et al. (2023), "Emerging trends in AI skill demand across 14 OECD countries"*, OECD Artificial Intelligence Papers*, No. 2, OECD Publishing, Paris, https://doi.org/10.1787/7c691b9a-en. [50]

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. [2]

Calvino, F. and L. Fontanelli (2023), "A portrait of AI adopters across countries: Firm characteristics, assets' complementarities and productivity"*, OECD Science, Technology and Industry Working Papers*, No. 2023/02, OECD Publishing, Paris, https://doi.org/10.1787/0fb79bb9-en. [44]

Calvino, F. et al. (2022), "Identifying and characterising AI adopters: A novel approach based on big data"*, OECD Science, Technology and Industry Working Papers*, No. 2022/06, OECD Publishing, Paris, https://doi.org/10.1787/154981d7-en. [53]

Caviggioli, F. et al. (2020), "How venture capitalists evaluate young innovative company patent portfolios: empirical evidence from Europe", *International Journal of Entrepreneurial Behavior & Research*, Vol. 26/4, pp. 695-721, https://doi.org/10.1108/ijebr-10-2018-0692. [41]

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. [1]

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. [12]

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. [19]

De Prato, G. and M. Cardona (2019), "The AI Techno-Economic Segment Analysis", *Publications Office of the European Union* EUR 29952 EN. [21]

De Prato, G. and D. Nepelski (2012), "Global technological collaboration network: network analysis of international co-inventions", *The Journal of Technology Transfer*, https://doi.org/10.1007/s10961-012-9285-4. [34]

Dernis, H. et al. (2023), "Identifying artificial intelligence actors using online data", *OECD Science, Technology and Industry Working Papers*, No. 2023/01, OECD Publishing, Paris, https://doi.org/10.1787/1f5307e7-en. [51]

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. [26]

European Commission. Joint Research Centre (2018), *Artificial intelligence : a European perspective*, Publications Office, https://doi.org/10.2760/936974. [22]

Fujii, H. and S. Managi (2018), "Trends and priority shifts in artificial intelligence technology invention: A global patent analysis", *Economic Analysis and Policy*, Vol. 58, pp. 60-69, https://doi.org/10.1016/j.eap.2017.12.006. [29]

Gao, Z. (ed.) (2016), "Patent Network Analysis and Quadratic Assignment Procedures to Identify the Convergence of Robot Technologies", *PLOS ONE*, Vol. 11/10, p. e0165091, https://doi.org/10.1371/journal.pone.0165091. [36]

Giczy, A., N. Pairolero and A. Toole (2021), "Identifying artificial intelligence (AI) invention: a novel AI patent dataset", *The Journal of Technology Transfer*, Vol. 47/2, pp. 476-505, https://doi.org/10.1007/s10961-021-09900-2. [20]

Grashof, N. and A. Kopka (2022), "Artificial intelligence and radical innovation: an opportunity for all companies?", *Small Business Economics*, https://doi.org/10.1007/s11187-022-00698-3. [33]

Green, A. and L. Lamby (2023), "The supply, demand and characteristics of the AI workforce across OECD countries", *OECD Social, Employment and Migration Working Papers*, No. 287, OECD Publishing, Paris, https://doi.org/10.1787/bb17314a-en. [49]

Griliches, Z. (1957), "Hybrid Corn: An Exploration in the Economics of Technological Change", *Econometrica*, Vol. 25/4, p. 501, https://doi.org/10.2307/1905380. [13]

Guan, J. and Q. Zhao (2013), "The impact of university–industry collaboration networks on innovation in nanobiopharmaceuticals", *Technological Forecasting and Social Change*, Vol. 80/7, pp. 1271-1286, https://doi.org/10.1016/j.techfore.2012.11.013. [35]

Hall, B., A. Jaffe and M. Trajtenberg (2001), *The NBER Patent Citation Data File: Lessons, Insights and Methodological Tools*, National Bureau of Economic Research, Cambridge, MA, https://doi.org/10.3386/w8498. [9]

Hötte, K. et al. (2023), "AI Technological Trajectories in Patent Data: General Purpose Technology and Concentration of Actors", *INET Oxford Working Paper*, Vol. No. 2023-09. [14]

Iori, M., A. Martinelli and A. Mina (2021), "The direction of technical change in AI and the trajectory effects of government funding", *Laboratory of Economics and Management Working Paper Series* 2021/41. [25]

Jaffe, A. and G. de Rassenfosse (2017), "Patent citation data in social science research: Overview and best practices", *Journal of the Association for Information Science and Technology*, Vol. 68/6, pp. 1360-1374, https://doi.org/10.1002/asi.23731. [8]

König, P. and G. Wenzelburger (2020), "Opportunity for renewal or disruptive force? How artificial intelligence alters democratic politics", *Government Information Quarterly*, Vol. 37/3, p. 101489, https://doi.org/10.1016/j.giq.2020.101489. [7]

Lane, M., M. Williams and S. Broecke (2023), "The impact of AI on the workplace: Main findings from the OECD AI surveys of employers and workers", *OECD Social, Employment and Migration Working Papers*, No. 288, OECD Publishing, Paris, https://doi.org/10.1787/ea0a0fe1-en. [46]

Lerner, J. (1994), "The Importance of Patent Scope: An Empirical Analysis", *The RAND Journal of Economics*, Vol. 25/2, p. 319, https://doi.org/10.2307/2555833. [40]

Li, D., L. Raymond and P. Bergman (2020), *Hiring as Exploration*, National Bureau of Economic Research, Cambridge, MA, https://doi.org/10.3386/w27736. [6]

Lorenz, P., K. Perset and J. Berryhill (2023), "Initial policy considerations for generative artificial intelligence", *OECD Artificial Intelligence Papers*, No. 1, OECD Publishing, Paris, https://doi.org/10.1787/fae2d1e6-en. [54]

Manheim, K. and L. Kaplan (2019), "Artificial Intelligence: Risks to Privacy and Democracy", *21 Yale Journal of Law and Technology*, Vol. Loyola Law School, Los Angeles Legal Studies Research Paper No. 2018-37, http://Available at SSRN: h. [5]

Messeni Petruzzelli, A., G. Murgia and A. Parmentola (2023), "Opening the black box of artificial intelligence technologies: unveiling the influence exerted by type of organisations and collaborative dynamics", *Industry and Innovation*, pp. 1-31, https://doi.org/10.1080/13662716.2023.2213182. [37]

Milanez, A. (2023), "The impact of AI on the workplace: Evidence from OECD case studies of AI implementation", *OECD Social, Employment and Migration Working Papers*, No. 289, OECD Publishing, Paris, https://doi.org/10.1787/2247ce58-en. [52]

Montobbio, F. et al. (2022), "Robots and the origin of their labour-saving impact", *Technological Forecasting and Social Change*, Vol. 174, p. 121122, https://doi.org/10.1016/j.techfore.2021.121122. [23]

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. [27]

OECD (2023), *Artificial Intelligence in Science: Challenges, Opportunities and the Future of Research*, OECD Publishing, Paris, https://doi.org/10.1787/a8d820bd-en. [55]

OECD (2023), *OECD Employment Outlook 2023: Artificial Intelligence and the Labour Market*, OECD Publishing, Paris, https://doi.org/10.1787/08785bba-en. [4]

OECD (2019), *Artificial Intelligence in Society*, OECD Publishing, Paris, https://doi.org/10.1787/eedfee77-en. [3]

OECD (2009), *OECD Patent Statistics Manual*, OECD Publishing, Paris, https://doi.org/10.1787/9789264056442-en. [17]

OECD (2024 forthcoming), “The diffusion of artificial intelligence in firms: findings of a new international survey”, *OECD Publishing, Paris*. [45]

Petit, N. (2017), “Antitrust and Artificial Intelligence: A Research Agenda”, *Journal of European Competition Law & Practice*, Vol. 8/6, pp. 361-362, https://doi.org/10.1093/jeclap/lpx033. [30]

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. [48]

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. [28]

Schmoch, U. (2008), *Concept of a Technology Classification for Country Comparisons - Final Report to the World Intellectual Property Organisation (WIPO)*. [56]

Shane, S. (2001), “Technological Opportunities and New Firm Creation”, *Management Science*, Vol. 47/2, pp. 205–220, http://www.jstor.org/stable/2661570. [42]

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. [39]

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. [47]

Tseng, C. and P. Ting (2013), “Patent analysis for technology development of artificial intelligence: A country-level comparative study”, *Innovation*, Vol. 15/4, pp. 463-475, https://doi.org/10.5172/impp.2013.15.4.463. [18]

Vaswani, A. et al. (2017), “Attention Is All You Need”, *Advances in neural information processing systems*, Vol. 30, https://doi.org/10.48550/arXiv.1706.03762. [43]

Webb, C. et al. (2005), “Analysing European and International Patent Citations: A Set of EPO Patent Database Building Blocks”, *OECD Science, Technology and Industry Working Papers*, No. 2005/9, OECD Publishing, Paris, https://doi.org/10.1787/883002633010. [38]

WIPO (2019), *WIPO Technology Trends 2019: Artificial Intelligence*. [24]

Xiao, J. and R. Boschma (2022), “The emergence of artificial intelligence in European regions: the role of a local ICT base”, *The Annals of Regional Science*, https://doi.org/10.1007/s00168-022-01181-3. [32]

Annex A. Additional information about data and methodology

Annex A summarises additional information about the methodology applied to identify AI-related patents and provides the list of keywords and their corresponding topics.

Identifying AI-related patents

The following set of patents are tagged as AI-related and will be used for the aforementioned analysis. This search strategy enables to identify patents that are related to AI, either as they embed core AI techniques (such as specific algorithms or methods) and/or as they relate to technological applications of AI.

- classified in one of the IPC codes: G06N3, G06N5, G06N20, G06F15/18, G06T1/40, G16C20/70, G16B40/20 or G16B40/30
- classified in one of the IPC codes and featuring in their English abstract or claims at least one AI keyword: G01R31/367, G06F17/(20-28, 30), G06F19/24, G06K9/00, G06K9/(46-52, 60-82), G06N7, G06N10, G06N99, G06Q, G06T7/00-20, G10L15, G10L21, G16B40/(00-10), G16H50/20-70, H01M8/04992 or H04N21/466
- classified in one of the CPC codes and featuring in their English abstract or claims at least one AI keyword: A61B5/(7264,7267), B29C2945/76979, B29C66/965, B60G2600/(1876-1879), E21B2041/0028, F02D41/1405, F03D7/046, F05B2270/(707-709), F05D2270/(707-709), F16H2061/(0081-0084), G01N2201/1296, G01N29/4481, G01N33/0034, G01R31/367, G01S7/417, G05B13/(027-029), G05B2219/33002, G05D1/0088, G06F11/(1476,2257,2263), G06F15/18, G06F17/(20-28), G06F19/(34,707), G06F2207/4824, G06K7/1482, G06K9/00, G06K9/(46-52, 60-82), G06N3, G06N5, G06N7, G06N10, G06N20, G06N99, G06Q, G06T2207/(20081,20084), G06T3/4046, G06T7/(00-20), G06T9/002, G08B29/186, G10H2250/(151,311), G10K2210/(3024,3038), G10L15, G10L21, G10L25/30, G11B20/10518, G16B40, G16C20/70, G16H50/(20,70), H01J2237/30427, H01M8/04992, H02P21/0014, H02P23/0018, H03H2017/0208, H03H2222/04, H04L2012/5686, H04L2025/03464, H04L25/(0254,03165), H04L41/16, H04L45/08, H04N21/(4662-4666), H04Q2213/(054,13343,343), H04R25/507, Y10S128/(924-925) or Y10S706
- featuring at least three AI keywords in their English abstract or claims.

Table A A.1. List of AI keywords and corresponding AI topics

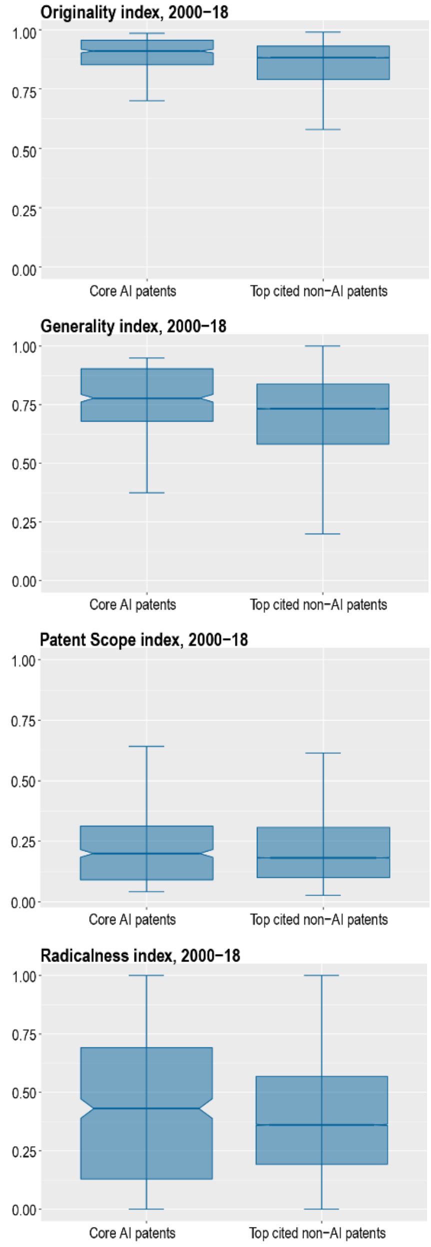
Algorithms	adaboost adaptive boosting ant colony ant colony optimisation artificial bee colony algorithm association rule backpropagation bayesian learning bayesian network bootstrap aggregation classification tree decision tree dictionary learning differential evolution algorithm dynamic time warping evolutionary algorithm evolutionary computation extreme learning machine fuzzy c fuzzy logic fuzzy number	fuzzy set gaussian mixture model gaussian process genetic algorithm genetic programming gradient boosting gradient tree boosting graphical model hebbian learning hidden Markov model hierarchical clustering k-means logitboost markovian memetic algorithm multi-objective evolutionary algorithm naive Bayes classifier natural gradient nearest neighbour algorithm particle swarm optimisation	q-learning random field random forest rankboost regression tree rough set rule learning self-organising map spectral clustering stacked generalization statistical relational learning stochastic gradient support vector machine support vector regression swarm optimisation temporal difference learning variational inference vector machine xgboost
Chatbot	chatbot	virtual assistant	
Autonomous driving	autonomous vehicle blind signal separation	collision avoidance motion planning	unmanned aerial vehicle pedestrian detection
Computer or image vision	computer vision face recognition image classification image processing	image recognition image retrieval image segmentation machine vision	obstacle avoidance visual servoing
Feature engineering	data mining dimensionality reduction factorisation machine feature engineering feature extraction feature learning	feature selection high-dimensional data high-dimensional feature high-dimensional input high-dimensional space independent component analysis	mapreduce non negative matrix factorisation simultaneous localisation mapping sparse representation
General AI	ambient intelligence artificial intelligence autonomic computing brain computer interface cluster analysis cognitive automation cognitive computing cognitive modelling collaborative filtering computational intelligence cyber physical system ensemble learning inductive monitoring	instance-based learning intelligence augmentation intelligent agent intelligent classifier intelligent infrastructure intelligent software agent Kernel learning learning automata machine intelligence machine learning meta learning multi task learning multi-agent system	multi-label classification multi-objective optimisation neuromorphic computing recommender system reinforcement learning relational learning rule-based learning semi-supervised learning similarity learning supervised learning swarm intelligence transfer learning unsupervised learning
Networks (deep learning)	adversarial network artificial neural network convolutional neural network deep belief network deep convolutional neural network	deep learning deep neural network generative adversarial network long short term memory (LSTM)	multi-layer perceptron neural network neural turing recurrent neural network

Natural language processing	autoencoder latent dirichlet allocation latent semantic analysis latent variable	machine translation natural language generation natural language processing natural language understanding	semantic web sentiment analysis text mining topic model
Recognition or detection	action recognition activity recognition community detection emotion recognition facial expression recognition	gesture recognition human activity recognition human action recognition link prediction object detection	object recognition pattern recognition trajectory planning trajectory tracking
Robotics	biped robot humanoid robot human-robot interaction industrial robot	legged robot multi-sensor fusion robot sensor data fusion	sensor fusion service robot social robot
Speech	speech recognition	speech to text	text to speech

Annex B. Additional results

Figure A B.1. Patent quality indicators, core AI vs. top cited non-AI patents, 2000-18

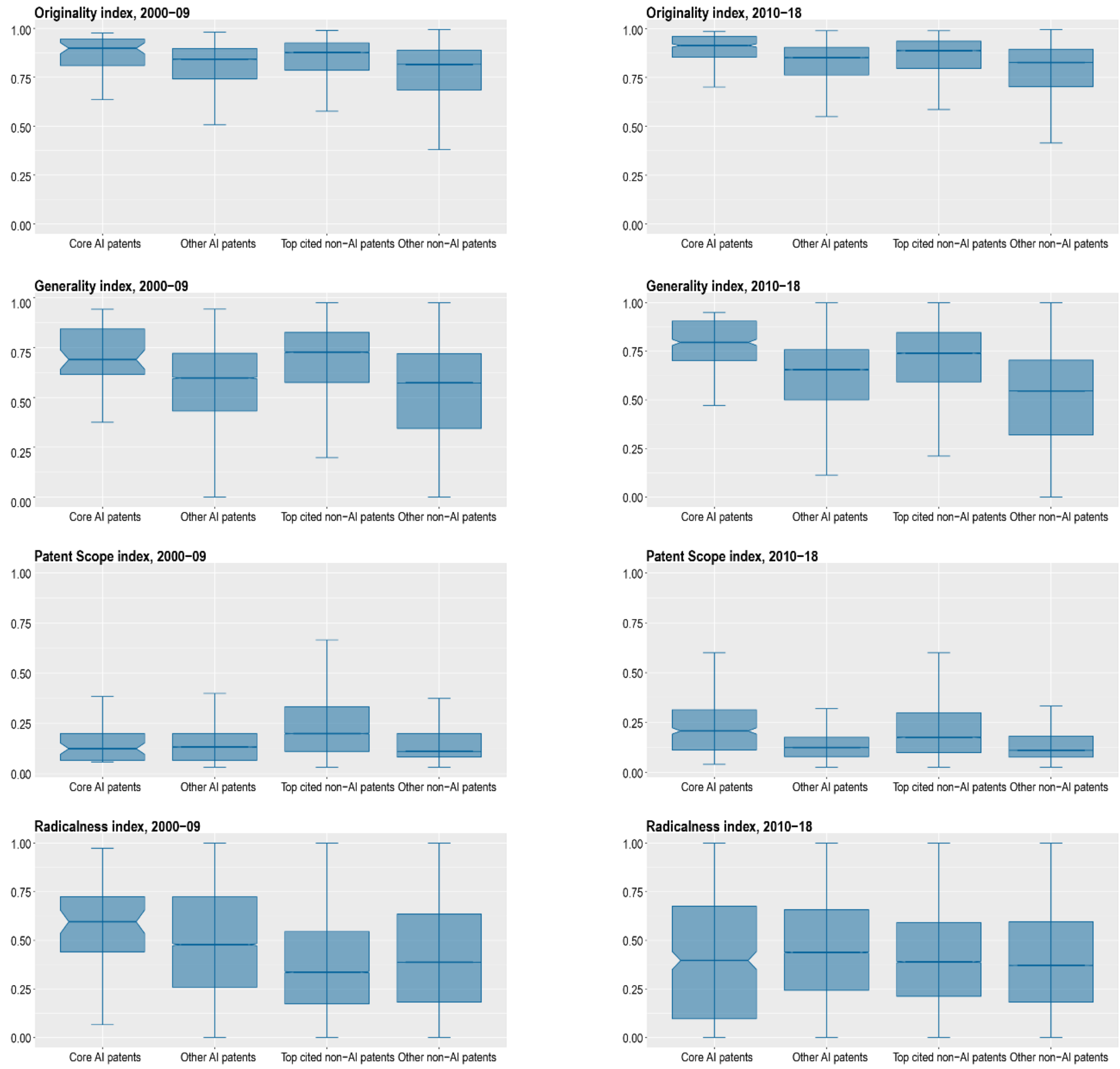
Technological originality, generality, scope and radicalness of core (top 1% cited) AI and top 1% cited non-AI patents



Note: The core AI patent sample (i.e. the core AI set covering up to 472 patents) refers to the top 1% cited patent applications (normalised counts of forward citations) filed at the USPTO (patents issued and pre-grant publications) that refer to AI technologies. The "Top cited non-AI patents" sample refers to patents in the top 1% highly cited patents in non-AI technologies. Each indicator is constructed based on Squicciarini et al. (2013[39]), with more detailed description in Box 4.1: the patent scope index is normalised to the maximum value observed in patents belonging to the same cohort (filing date and WIPO technology). By construction, all indicators presented range between zero and one.
Source: OECD, STI Micro-data Lab: Intellectual Property Database, http://oe.cd/ipstats, June 2023.

Figure A B.2. Patent quality indicators, different patent groups, 2000-09 vs. 2010-18

Technological originality, generality, scope and radicalness of core (top 1% cited) AI, other AI, top 1% cited non-AI and other non-AI patents



Note: The core AI patent sample (i.e. the core AI set covering up to 472 patents) refers to the top 1% cited patent applications (normalised counts of forward citations) filed at the USPTO (patents issued and pre-grant publications) that refer to AI technologies. The "other AI patents" sample refers to the remaining AI patents identified. The "Top cited non-AI patents" sample refers to patents in the top 1% highly cited patents in non-AI technologies. The "Other non-AI patents" sample refers to the remaining non-AI patents identified. Each indicator is constructed based on Squicciarini et al. (2013[39]) with more detailed description in Box 4.1: the patent scope index is normalised to the maximum value observed in patents belonging to the same cohort (filing date and WIPO technology). By construction, all indicators presented range between zero and one.
Source: OECD, STI Micro-data Lab: Intellectual Property Database, http://oe.cd/ipstats, June 2023.