

Measuring artificial intelligence: a systematic assessment and implications for governance

Kerstin Hötte¹ , Taheya Tarannum², Vilhelm Verendel^{3,*}  and Lauren Bennett⁴

¹Kedge Business School, Department of Accounting, Finance and Economics, 40 Av. des Terroirs de France, 75012 Paris, France. e-mail: kerstin.hotte@kedgebs.com, ²Oxford Martin Programme on Technological and Economic Change, University of Oxford, 34 Broad Street, OX1 3BD Oxford, UK. e-mail: taheya.tarannum@oxfordmartin.ox.ac.uk, ³Department of Physics, Chalmers University of Technology, Kemigården 1, 412 96 Gothenburg, Sweden. e-mail: vive@chalmers.se and ⁴The Leverhulme Doctoral Training Programme for the Ecological Study of the Brain, University College London, Gower street, WC1E6BT London, UK. e-mail: lauren.bennett.21@ucl.ac.uk

*Main author for correspondence.

Governing artificial intelligence (AI) inventions is a major policy concern. Yet, definitions and measurement approaches remain contested. We compare four patent-based definitions reflecting distinct understandings of AI. Using US patents (1990–2019), we assess the degree to which each approach describes AI as a general-purpose technology (GPT) and examine patent concentration by a few dominant firms. We find that between 3% and 17% of all US patents in 2019 are classified by at least one of the approaches as AI patents. Yet, only 1.4% of all AI patents are simultaneously identified by all four approaches. All approaches are consistent with AI having GPT characteristics, with the Keyword-based patents exhibiting the highest growth and generality. GPT indicates public good characteristics, which could be used to justify public support. Across methods, AI patents are concentrated among a few firms, highlighting market power and regulatory challenges. The wide variation in the subsets and characteristics of AI patents identified by these approaches suggests that currently multiple classification methods should be considered to formulate robust, inclusive, and effective analyses for AI governance.

JEL Classification: O31, O33, O34

1. Introduction

Artificial intelligence (AI) governance is a significant policy concern, reflecting the aim to mitigate risks and maximize societal benefits (Jelinek *et al.*, 2021; Mazzucato *et al.*, 2022; Schmitt, 2022). Governance based on data, measurement, and tracking requires a clear identification of AI inventions, and so far a consensual definition is lacking (Krafft *et al.*, 2020). A wide audience is interested in understanding recent trends in AI development but the lack of clear identifying methods can hamper assessments of AI impacts and the effectiveness of regulatory policy (Dafoe, 2018). Ambiguities in defining and measuring AI contribute to empirical controversies concerning its impacts on labor markets, technological leadership, and productivity (Alderucci *et al.*, 2020; Brynjolfsson *et al.*, 2021; Bresnahan, 2023; Babina *et al.*, 2024; Cazzaniga *et al.*, 2024). In this paper, we use patents as a quantitative and qualitative record of AI inventions. Comparing differences across four different AI classification methods, we investigate if and how empirical conclusions about the characteristics and concentration of AI inventions can be sensitive to the method.

AI is widely seen as a general-purpose technology (GPT), expected to profoundly transform production processes, labor markets, and economic leadership at national and global levels (Valdes and Rudyk, 2017; Cockburn *et al.*, 2018; Agrawal *et al.*, 2019; Cockburn *et al.*, 2019; Webb, 2019; Alderucci *et al.*, 2020; Brynjolfsson *et al.*, 2021). Testing these predictions requires analysis based on empirical measurement; yet ex-ante measurement of radical novelty remains challenging (Schumpeter *et al.*, 2005). AI is still widely regarded as being in the early stages of development (Brynjolfsson *et al.*, 2021), and its long-run effects have not yet manifested and can be shaped by decisions undertaken today (Petit, 2017; Jacobides *et al.*, 2021; Fanti *et al.*, 2022).

This paper does not aim to identify the best AI definition, but rather examines agreements and differences across measurement methods and their implications for policy and research. We do so by comparing four samples of AI patents, each reflecting a different way to understand and define AI. The samples are identified by Keywords focusing on trends in neural networks robotics and natural language processing (NLP):

- (1) scientific citations reflecting the academic origins of AI;
- (2) the World Intellectual Property Organization (WIPO) classification method, accounting for both the hardware and software dimensions of AI; and
- (3) the United States Patent and Trademark Office (USPTO) approach capturing the widespread use of AI in other inventions.

In the union of all four samples, we identify 732 k as AI patents from 1990 to 2019. Yet, the individual samples vary greatly by scale: 54 k patents are captured by the Keyword, 178 k by the Science, 159 k by the WIPO, and 595 k by the USPTO approach. Strikingly, all four methods agree on only 1.37% of all identified AI patents, with pairwise overlaps of 10%–20% or less throughout the entire period. Further, the four approaches reflect disparate time trends, with the Science and USPTO sample showing an AI slowdown in recent years, while Keyword and WIPO patents exhibit accelerating growth since the 2000s.

We evaluate whether each of the four samples reflect the beliefs of AI being a GPT, by assessing three GPT characteristics established in the literature (Bresnahan and Trajtenberg, 1995; Hall and Trajtenberg, 2006; Petralia, 2020):

- (1) Growth: GPTs are engines of growth with continued technological improvements (Petralia, 2020). We measure this feature by the growth rates of each AI sample, and patents that rely on them, as indicated by citations.
- (2) Generality: GPTs can be used across a wide range of products and processes. We examine the technological diversity of patent citations. As GPTs often experience long delays before being widely taken up (Comin and Mestieri, 2018), we measure citation lags between AI and subsequent inventions (Hall and Trajtenberg, 2006).
- (3) Complementarity: GPTs complement technologies in many fields (Petralia, 2020), being reflected in a high technological diversity. We quantify the diversity of co-classifications of AI patents across technology groups.

Evaluating the “GPTness” of AI is relevant because it indicates the potential for continued macroeconomic growth with large social benefits in the long run (Lipsey *et al.*, 2005), and the public good characteristics of GPTs can justify public support (Bresnahan and Trajtenberg, 1995).

Whether AI is a GPT or not cannot be taken as empirically given: If AI inventions have little or no GPT characteristics, some arguments claiming that AI invention needs public support would be undermined. However, an observed lack of GPTness could also be a matter of classification or barriers to the realization of AI benefits, which do not necessarily arise by themselves: commercial actors may maximize the private rather than the social value of an invention, leading to a premature lock-in to an inferior pathway of AI development (Klinger *et al.*, 2020; Bresnahan, 2023). This consideration motivates an additional analysis of the key actors in AI development and the concentration of inventive activities.

We find that all methods characterize AI as a GPT, yet the AI inventions identified by Keywords and the WIPO method show the highest levels of GPTness. Inventive activities by firms appear least concentrated in technology areas identified by Keywords. Relying on GPTness and diversity

in the market of AI development, one may conclude that AI as captured by Keywords bears the greatest potential for public benefits, compared to the other methods. This could guide policy aimed at supporting AI inventions for the public good.

Our patent-based approach captures only a segment of the AI ecosystem and evaluates a narrow set of policy-relevant criteria (Jacobides *et al.*, 2021; Roche *et al.*, 2023). Still, within this scope, we find concentration levels among the Big Tech companies to be highest in electronics and computing, while invention in AI applications appears to be more widely spread. Attempts to empirically assess the impact of AI and policy should therefore currently rely on a variety of methods to identify AI before drawing conclusions.

Our systematic comparison can guide methodological choices in patent-based research by highlighting the implications and conceptual considerations when measuring AI. Further, our results inform discussions on AI definitions by showing consensual features of AI and quantifying nuanced differences in the types of industries, technology fields, and key players involved in AI development.

The relevance of this research ranges from researchers (who reflect on AI definitions or rely on classifications in applied research) to policymakers (who rely on AI definitions and empirical insights on AI impacts in everyday policy). On the longer run, more knowledge about AI definitions and their implementation may help clarify some of the empirical controversies about AI impacts on society.

The remainder of our paper is structured as follows: Section 2 introduces GPTs and AI in patent data; Section 3 describes the methods, followed by the results (Section 4). In Sections 5 and 6, we discuss the findings and limitations, and Section 7 concludes.

2. Background

Here, we review (i) the arguments for AI being a GPT (Section 2.1), (ii) the history of AI and its definitions (Section 2.2), and (iii) patents as a data source (Section 2.3).

2.1 What is a GPT, and does it matter for AI?

Bresnahan and Trajtenberg (1995) define GPTs as technologies that pervade the economy and spur inventions via complementarities. Well-established GPTs, such as electricity and Information and Communication Technologies (ICTs), are widely considered as long-term drivers of societal value, growth, and technological progress. Quantitative research identifies GPTs by three characteristics: rapid intrinsic improvement (growth), economic pervasiveness (generality), and productivity spillovers across sectors (complementarity) (Bresnahan and Trajtenberg, 1995; Lipsey *et al.*, 2005).

The first criterion refers to GPTs' inherent capacity to rapidly improve. If the technology is sufficiently mature to be valuable for many uses, this should be reflected in high growth rates. The second criterion refers to GPTs' ability to engender new methods of production or innovation. Due to their pervasiveness, GPTs inspire a wave of technological inventions as they embody new universal tools for production and research. Other impactful technologies, such as nuclear power and fMRI, lack the generality required to pervade a significant number of sectors (Agrawal *et al.*, 2018; Brynjolfsson *et al.*, 2019). The third criterion highlights how GPTs generate productivity spillovers by complementing existing products and processes. This often comes with creative destruction and technological discontinuities rendering established technologies and jobs obsolete (Lipsey *et al.*, 2005).

Together, these criteria require longer time for GPTs to evolve, diffuse, and realize their full economic impact (Lipsey *et al.*, 2005). Once the inventions have evolved and spread sufficiently, GPTs depend on complementary infrastructure and secondary innovations to transform organizations and reshape production processes across the economy (Bresnahan, 2023).

While much of the literature frames AI as a GPT (e.g. Valdes and Rudyk, 2017; Agrawal *et al.*, 2018; Cockburn *et al.*, 2018; Cockburn *et al.*, 2019; Webb, 2019; Alderucci *et al.*, 2020; Brynjolfsson *et al.*, 2021), some scholars argue that the GPT concept may be too simplistic for understanding AI. Recent work has raised concerns about conceptualizing AI as a GPT, particularly from a policy perspective (Jacobides *et al.*, 2021). Others caution that the realization

of GPT-related benefits cannot be taken for granted (Bresnahan, 2023). GPTs may generate market failures due to their public good characteristics, which can lead to suboptimal levels of private investment (Bresnahan and Trajtenberg, 1995). However, it remains unclear whether AI constitutes a single GPT or an amalgamation of multiple technologies that perform diverse functions, including data provision, prediction, classification, soft- and hardware integration, and AI edge applications (Jacobides *et al.*, 2021).

Certain AI ecosystem components are more essential than others, and countries and companies show different specializations in subareas of AI, but those that lack access to the critical components run the risk of being cut off from realizing the benefits of AI (Klinger *et al.*, 2020; Jacobides *et al.*, 2021; Bresnahan, 2023; Franco *et al.*, 2023; Babina *et al.*, 2024; Cazzaniga *et al.*, 2024).

Therefore, if AI is understood as a system of interdependent technologies and resources—potentially organized in a hierarchical structure—the study of AI should be broadened to include the key actors shaping its technological evolution. The growing concentration of AI development among a few dominant firms (Klinger *et al.*, 2020; Babina *et al.*, 2024) has raised concerns about the influence of Big Tech (Jacobides *et al.*, 2021) and geographic or institutional hubs (Klinger *et al.*, 2020). This concentration also casts doubt on the case for public funding of AI, as advocated from a GPT perspective (Bresnahan and Trajtenberg, 1995; Jacobides *et al.*, 2021). On the contrary, additional public support—if not accompanied by competition policy and regulatory oversight—may reinforce existing asymmetries in the AI ecosystem (Petit, 2017; Hennemann, 2020).

In this paper, we analyze whether the GPT framing is consistent with the four different AI patent classification approaches. Our analysis is informative along three main dimensions: (i) timescale and history: different definitions of AI affect the timescale on which it is viewed and the perceived history of its development; (ii) concentration and key actors: theoretical approaches differ in which key actors that are identified to shape AI innovation and the implications for market power; and (iii) GPT characteristics: the choice of definition may suggest different guidelines for the direction supporting AI technologies, depending on the degree to which AI satisfies the characteristics and related public goods of a GPT. Our analysis investigates the methodological basis of empirical research on AI and its impact, showing how empirical controversies can arise from different approaches to defining and measuring AI inventions in the economy.

2.2 AI history and different definitions

AI broadly refers to technologies that perform tasks requiring intelligence. Its conceptual roots trace at least back to Turing (1950) and work in the 1950s. Subsequent decades saw sustained R&D investment that led to advances in using computers as problem-solving machines. However, unmet high expectations, combined with limited computing power and reduced funding, resulted in stagnation periods (“AI winters”) in the 1970s and 1990s (Stuart and Norvig, 2003).

Modern AI—since at least the mid-2000s—has been largely driven by methods from machine learning (ML), which typically involve computational techniques for detecting and modeling patterns in diverse data sources (Mitchell, 1997). ML draws on foundations from computer science, statistics, logic, probability, and optimization. In many cases, ML software is integrated with hardware components such as sensors, actuators, and control systems to form intelligent systems. This combination of software and hardware is one established approach to building AI systems.

Recent work argues that modern AI research is becoming increasingly privatized and narrowly focused on deep learning—a specific form of ML—at the expense of other relatively underexplored areas, such as symbolic learning (Bianchini *et al.*, 2020; Klinger *et al.*, 2020; Jurowetzki *et al.*, 2021; Whittaker, 2021). Moreover, the transformative potential of AI appears to be relatively underexploited by adopting firms (Bresnahan, 2023). It remains uncertain whether this narrowing will influence how AI is defined in the future, potentially shaping which inventions are captured under the AI label. Future developments may also involve new subfields that are, at present, largely unexplored.

Taken together, these observations reveal both long-term technological trends and a recent concentration of inventions around a limited set of specific technologies. The choice of how broadly to define AI is central to both research and policy debates, yet this decision is far from

straightforward. Efforts to define AI in more technology-neutral terms are already underway. For instance, in EU legislation regulating AI,¹ an *AI system* is defined as:

“a machine-based system that is designed to operate with varying levels of autonomy and that may exhibit adaptiveness after deployment, and that, for explicit or implicit objectives, infers, from the input it receives, how to generate outputs such as predictions, content, recommendations, or decisions that can influence physical or virtual environments.”

and the specific connection to ML is further emphasized in the text as

“AI systems often have machine learning capacities that allow them to adapt and perform new tasks autonomously. Machine learning refers to the computational process of optimizing the parameters of a model from data, which is a mathematical construct generating an output based on input data. Machine learning approaches include, for instance, supervised, unsupervised and reinforcement learning, using a variety of methods including deep learning with neural networks.

[...]

Comparably simpler techniques such as knowledge-based approaches, Bayesian estimation or decision-trees may also lead to legal gaps that need to be addressed by this Regulation, in particular when they are used in combination with machine learning approaches in hybrid systems.

[...]

AI systems can be used as stand-alone software system, integrated into a physical product (embedded), used to serve the functionality of a physical product without being integrated therein (non-embedded) or used as an AI component of a larger system. If this larger system would not function without the AI component in question, then the entire larger system should be considered as one single AI system under this Regulation.”

These definitional efforts deliberately avoid being overly specific, highlighting the difficulty of drawing clear system boundaries. As technological capabilities advance, definitions of AI for policy and regulation are likewise evolving. For example, since 2023, the EU AI Act has been refined to emphasize the potential outputs of systems classified as AI, and more recently, it has incorporated the requirement that such systems possess the ability to adapt and generalize.

Whilst the United States currently lacks a unified definition of AI, California Assembly Bill 331—introduced in 2023 as part of ongoing efforts to regulate AI—diverges from the EU AI Act by not explicitly addressing the autonomous nature of such systems.² The UK similarly lacks a unifying definition of AI, with conceptualizations varying significantly across government departments. Whilst the Department for Science, Innovation & Technology defines AI systems narrowly, emphasizing only those with “adaptable” and “autonomous” functionality,³ the UK’s National Cyber Security Centre identifies AI based on outputs—specifically, its ability to perform tasks typically associated with human intelligence.⁴ Evidently, key definitions differ not only in the types of technologies considered to constitute AI, but also in whether AI should be defined by its outputs, internal composition, adaptability, or autonomy.

2.3 Patents as a data source

Attempts of defining AI in legislation are driven by a perceived need for regulation, drawing on empirical research on AI, including economic studies based on patent data (e.g. *Cockburn et al.*, 2018; *Fujii and Managi*, 2018; *Webb*, 2019; *WIPO*, 2019b; *USPTO*, 2020; *Verendel*, 2023). Patents provide a rich data source for research, offering detailed records of inventions and the embedded technological knowledge. Patent offices maintain millions of records, categorized into

¹ https://www.europarl.europa.eu/doceo/document/TA-9-2023-0236_EN.html.

² https://leginfo.ca.gov/faces/billTextClient.xhtml?bill_id=202320240AB331.

³ <https://www.gov.uk/government/publications/ai-regulation-a-pro-innovation-approach/white-paper>.

⁴ <https://www.ncsc.gov.uk/section/advice-guidance/>.

technological fields via hierarchical classification systems. A single patent may be assigned to one or more codes from hierarchical patent classification systems, each reflecting a distinct aspect of the underlying technology (Jaffe and De Rassenfosse, 2017). In AI measurement, these codes help describe the qualitative nature of a patent, for example, whether it is related to specific computing techniques or hardware for data transmission, or related to functions in another sector of the economy.

Patent documents also include citation links to other patents and academic articles, which help differentiate the patented invention from prior technologies. Inventors must demonstrate the novelty of their invention in relation to existing technologies. These citations are commonly used by innovation scholars, as they reveal the cumulative dependencies between technologies and can serve as an indicator of how much an invention builds on existing knowledge and technology (Jaffe and De Rassenfosse, 2017).

Several studies on AI have leveraged these characteristics to conduct empirical quantitative assessments of AI's impact, with some framing AI as a potential GPT (Cockburn *et al.*, 2018). However, it is important to note that there are many alternatives to patent data, including industry-, firm-, and technology-level non-patent data, both quantitative and qualitative, as well as conceptual frameworks (e.g. Klinger *et al.*, 2018; Trajtenberg, 2018; Bresnahan, 2023; Goldfarb *et al.*, 2023).

3. The four different AI classification approaches

In our analysis, we collect four sets of AI patents using the four distinct AI definitions applied to patent data, and compare them based on their GPTness and patterns of concentration. For GPTness, we assess the growth, generality, and complementarity of AI patents classified by (i) Keywords, (ii) science citations, (iii) the WIPO method, and (iv) the USPTO method, drawing from all USPTO patents granted between 1990 and 2019. To analyze the concentration of inventive activity, we measure the Herfindahl–Hirschman Index (HHI) and evaluate dominant firms. The four approaches are implemented as follows:

First, the Keyword method replicates Cockburn *et al.* (2018), using text search terms (e.g. NLP, robotics, neural networks).⁵ This approach reflects short-term perspectives on AI, attributing its progress to themes that became dominant since the mid-2000s.

Second, we use scientific citations, including gray literature and conference proceedings (Marx and Fuegi, 2020), as an AI identifier. This approach conceptualizes AI technologies as the result of scientific and academic research (Arthur, 2009; Jee and Sohn, 2023). As discussed below, this method has limitations, as citation practices vary across different technological fields.

Third, the WIPO method emphasizes AI's technical foundations, such as ML and applied computing. It combines a set of Keywords with technology-specific classification codes using Boolean conditions (WIPO, 2019a).

Fourth, the USPTO method adopts the broadest understanding of AI. It employs a trained ML classifier on patent text and citations to identify innovations related to knowledge processing, speech, hardware, evolutionary computation, NLP, ML, computer vision, and planning and control (Giczy *et al.*, 2021). This expansive conceptualization of AI is reflected in the large volume of AI patents identified through this method, including a significant share of downstream AI applications.

For further background on these definitions and details on their implementation, we refer the reader to Appendices A and B.

4. Results

We compare the four AI samples by their general characteristics (Appendix 4.1), GPTness (Appendix 4.2), and concentration (Appendix 4.3). Several results are described using standard

⁵ The keywords used in this paper focus on three subfields of AI: symbolic systems (e.g. natural language, image grammars, pattern recognition, symbolic reasoning), learning algorithms (e.g. machine learning, neural networks, Bayesian belief networks, unsupervised learning, deep learning), and robotic (e.g. computer vision, robot systems, humanoid robotics, sensor network, and systems and control theory) Also, see Appendix Table C1 for the full list of keywords.

Table 1. AI patents by four classification approaches

	Keyword	Science	WIPO	USPTO
A. Growth				
Growth rate (1990–99)	0.98	6.69	2.38	3.87
Growth rate (2000–09)	0.39	1.32	1.21	1.16
Growth rate (2010–19)	3.77	0.91	3.02	1.07
B. Inventor type (patent assignee)				
% Commercial	0.86	0.85	0.90	0.91
% Individual	0.06	0.03	0.05	0.05
% University	0.05	0.10	0.04	0.03
% Other non-profit	0.03	0.04	0.02	0.02
C. Industry affiliation (patent assignee)				
% Pharmaceuticals (manufacturing)	0.01	0.17	0.01	0.01
% Computer (manufacturing)	0.68	0.76	0.79	0.84
% Machinery & equipment (manufacturing)	0.49	0.28	0.54	0.22
% Other manufacturing	0.14	0.10	0.09	0.09
% Computer programming (service)	0.06	0.07	0.09	0.12
D. Country of origin (patent applicant)				
% USA	0.62	0.75	0.65	0.72
% Japan	0.14	0.08	0.15	0.10
% S. Korea	0.04	0.02	0.03	0.02
% Germany	0.05	0.03	0.03	0.03
% China	0.01	0.01	0.02	0.01
% Canada	0.02	0.02	0.02	0.02
E. Public support				
% Public support (1990–1999)	0.31	0.39	0.29	0.21
% Public support (2000–2009)	0.34	0.48	0.31	0.23
% Public support (2010–2017)	0.33	0.52	0.28	0.22
F. CPC 1-digit codes				
% Human necessities (A)	0.15	0.18	0.08	0.08
% Performing operations (B)	0.31	0.05	0.11	0.04
% Chemistry; metallurgy (C)	0.03	0.16	0.01	0.01
% Physics (G)	0.66	0.75	0.95	0.79
% Electricity (H)	0.25	0.26	0.27	0.33
% General/cross-sectional (Y)	0.05	0.04	0.03	0.04
Number of patents	54,145	178,004	158,652	595,047

This table compares the scale and scope of AI invention identified by each definition, disaggregated by inventor types (commercial, individual, non-profit, university), industry affiliation (based on NACE Rev. 2 classification), country of origin, reliance on public R&D support, and technological classification. Note that the data on public support ends in 2017.

patent classification data from the Cooperative Patent Classification (CPC) codes, that groups granted patents in different technology areas on different level of detail.

4.1 General characteristics of AI

Table 1 shows, for each method, growth rates (panel A), inventor types (panel B), main industries (panel C), countries of origin (panel D), public support (panel E), and main technology classes (panel F).

All methods indicate increasing diversification of AI inventions across industries and countries, suggesting economic diffusion. Furthermore, all approaches indicate that AI inventions are disproportionately generated by commercial enterprises (Panel B) and are predominantly associated with the computer and machinery manufacturing sectors (Panel C). Although the number of patents identified varies by orders of magnitude across methods, all consistently show a dominance of the United States in AI patenting. However, relying solely on USPTO data inherently over-represents US-based inventors.

We find differences in the growth rates of the Science and USPTO approaches, suggesting that these two groups have not experienced the same post-2000 acceleration as the other two approaches (Keywords and WIPO).

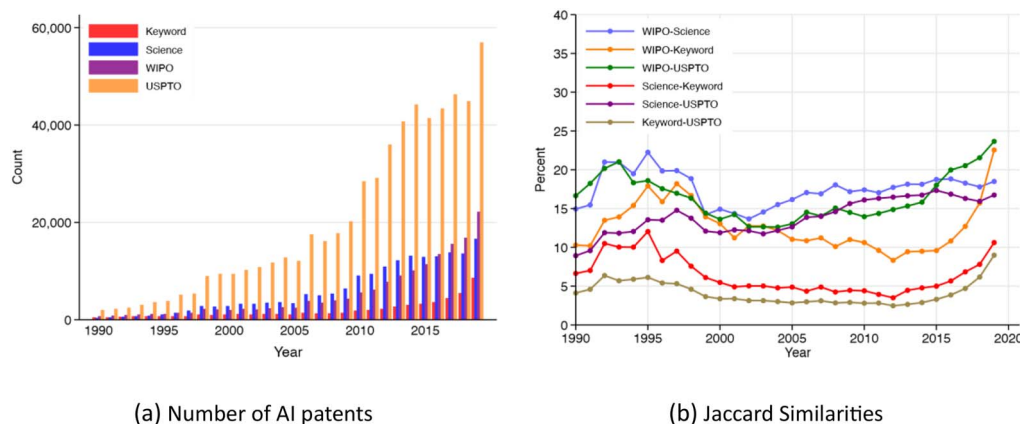


Figure 1. AI patents by year (1990–2019). *Note:* The right panel shows the number of AI patents over time as identified by the four approaches. The left panel shows the evolution of the Jaccard similarities computed for each year in our dataset.

When comparing how inventions are classified using CPC 1-digit codes, all approaches show a similar association with the “Electricity (H)” section. WIPO patents are most strongly concentrated in the “Physics (G)” section, while the Science approach most strongly identifies patents in “Chemistry (C).” The Keyword and Science approaches yield the most diverse range of AI-related patents across CPC categories. The high share (95%) of WIPO patents in “Physics (G)” is a direct result of the method’s design: the WIPO approach explicitly filters for patents classified under this CPC section (WIPO, 2019a).

Figure 1 illustrates the pace of AI invention. Expressed as a share of all granted US patents, the USPTO approach classifies approximately 16.6% of patents in 2019 as AI related (see Figure D10 in Appendix D3.1). Across all methods, this share has risen from 1%–2% in the 1990s to 3%–17% by 2019.

To quantify the overlap among the four AI patent samples, we compute pairwise Jaccard similarities and track their evolution over time, as shown in Figure 1b.⁶ All pairs of AI samples exhibit low overlaps, with Jaccard values at or below 20%. The WIPO approach shows the highest agreement with other methods. Notably, the WIPO—Keyword and Science—USPTO pairs show the most pronounced increases in similarity over time. The Keyword method has the lowest Jaccard similarity with the other approaches, likely due to the relatively small number of patents in this sample. Overall, only 10,062 patents—or 1.37% of all unique granted patents labeled as AI by any of the four methods—are simultaneously identified as AI patents by all four methods. This low overlap shows that quantitative assessments of the scale and diffusion of AI can be highly sensitive to the choice of classification method.

In the following section, we adopt a qualitative perspective to examine how these definitional differences influence the extent to which AI qualifies as a GPT.

4.2 GPT characteristics of AI

We study the GPTness of AI by comparing the four patent samples along three key dimensions: growth, generality, and complementarity. The results are summarized in Table 2 and discussed in detail below.

4.2.1 Growth

Panel A of Table 2 reports the average annual growth rates of AI patents, while Figure 2 illustrates how these rates have evolved over time. To show how overall trends have changed, we apply local regression (LOWESS) smoothing (Cleveland, 1979) to each time series.

⁶ The Jaccard similarity for two sets of patents is given by $J(A,B) = \frac{|\text{Patents in both A and B}|}{|\text{Patents in union of A and B}|} = \frac{|A \cap B|}{|A \cup B|}$ with $J(A,B) \in [0,1]$ where $J(A,B) = 0$ if both sets do not overlap and $J(A,B) = 1$ if both sets are identical. In other terms, the overlap between the sets can range from 0% to 100%.

Table 2. Measure of GPT characteristics of AI

	Keyword	Science	WIPO	USPTO
A. Growth				
Avg. growth rate	0.12	0.15	0.14	0.13
B. Generality				
Avg. generality index (1 digit)	0.81	0.77	0.76	0.73
Avg. generality index (3 digit)	0.94	0.90	0.90	0.87
Avg. generality index (4 digit)	0.97	0.96	0.95	0.95
C. Complementarity				
Avg. number of CPC (1 digit)	1.43	1.40	1.36	1.27
Avg. number of CPC (3 digit)	1.67	1.64	1.64	1.43
Avg. number of CPC (4 digit)	1.92	1.97	2.05	1.64

This table gives a comparison of the GPT-like characteristics of AI inventions classified by each distinct technique. Note that the generality index is defined as share of citations to patents in different CPC classes at different aggregation levels. Citations within the same class are excluded.

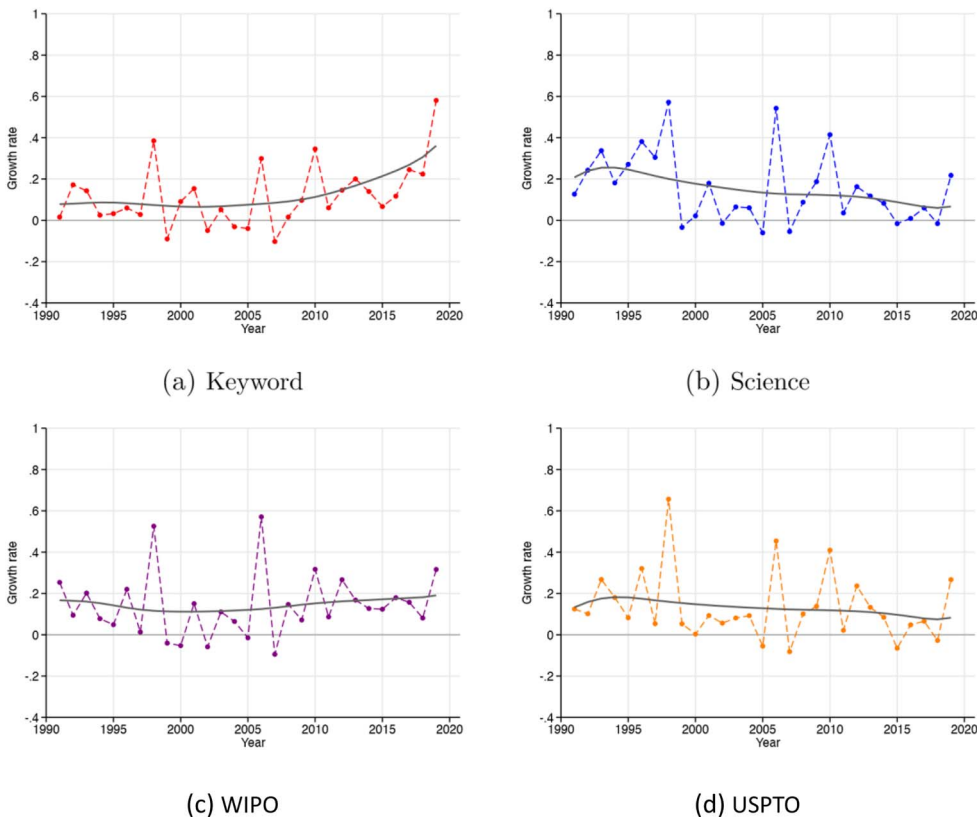


Figure 2. Growth of patents by year. *Note:* The four AI approaches (a)-(d) have different growth patterns over time. The averages for all are positive, but the Keyword (a) and WIPO (c) approaches both have increasing growth rates.

We observe the following: First, all methods exhibit positive growth rates, consistent with GPT expectations. Second, each series exhibits a dip in growth during the early 2000s, followed by renewed acceleration in more recent years.⁷ Third, the smaller AI samples identified by the Keyword and WIPO methods show a strong acceleration in the later years, in contrast to the Science and USPTO samples, which exhibit more moderate growth. Overall, the results indicate persistent growth across all methods, albeit with varying dynamics and strength depending on the classification approach.

⁷ This slowdown is not unique to AI; patenting activity declined across multiple sectors during this period (see Figure D4 in Appendix D1.1).

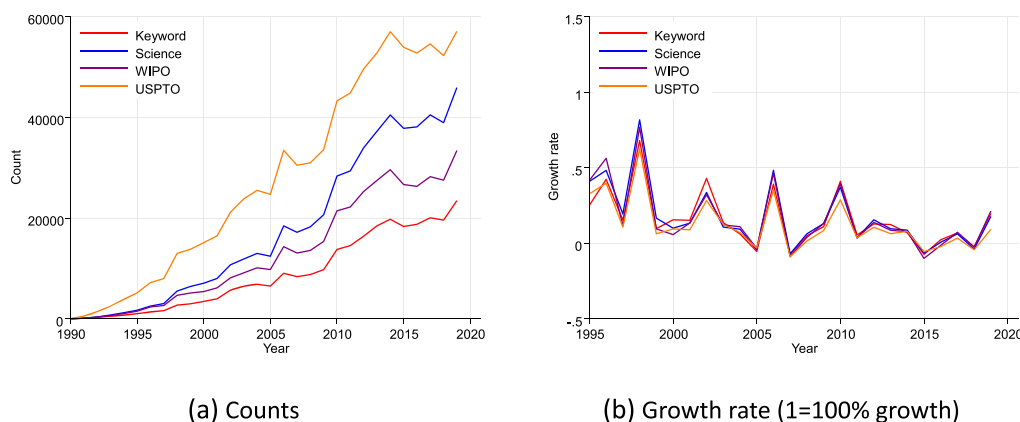


Figure 3. Patents citing AI. *Note:* Panel (a) shows the actual number of AI citing patents. Panel (b) shows growth rates plotted from 1995 and onwards.

In [Appendix D1](#), we compare our AI patent samples to benchmarks that have been discussed in the literature as potential GPTs, such as nanotechnology, climate technology, biochemistry, and general computing. Overall, the AI samples exhibit higher growth rates than most of these GPT candidates. A Wilcoxon test shows that AI patents grow significantly faster than average, though differences across AI types are mostly not significant (see [Appendix D2.1](#)).⁸ Among the AI classification methods, growth rate differences are generally not statistically significant, with the exception of the USPTO approach, which shows significantly lower growth compared to the others.

If AI functions as a GPT, it should enable follow-on inventions in non-AI sectors—reflected in the growth of patents that cite AI patents but are not themselves classified as AI. [Figure 3](#) shows rising numbers of AI-citing patents across all methods. The relative sizes of the groups are consistent with earlier patterns (see [Figure 2](#)). The positive downstream growth suggests that each approach generates a growing number of invention spillovers to non-AI sectors.

Significance tests in the appendix indicate that differences between the Keyword, Science, and WIPO approaches are not statistically significant—except for the Keyword sample, which showed slower uptake in the 1990s before accelerating. These three approaches score significantly higher than the USPTO sample.

Overall, these findings support the view that these classification methods capture AI technologies that have a persistent impact on the wider innovation landscape.

4.2.2 Generality

We use two indicators to evaluate the generality of AI inventions—that is, the extent to which AI patents are cited across diverse technological fields. First, we calculate a generality index ([Trajtenberg *et al.*, 1997](#); [Hall and Trajtenberg, 2006](#)), which reflects the dispersion of CPC 1-digit codes among patents citing AI inventions. This measure is conceptually similar to the HHI, but inverted to capture breadth rather than concentration.⁹ Technical details and additional results are provided in [Appendix A2](#) and [D3.2](#).

[Figure 4](#) shows how the generality index evolved over time. The Keyword sample consistently shows highest generality, with citations broadly distributed across technology fields. The WIPO and Science samples follow, with the Science-based approach scoring slightly higher than WIPO. The USPTO sample shows the lowest generality at all CPC aggregation levels—somewhat

⁸ It should be noted that these statistical tests are based on a relatively small number of observations, and the growth trajectories of the four AI samples vary substantially over the three-decade period.

⁹ The formula of the generality index GI_t is given by $GI_t = 1 - \sum_j^{N_j} \left(\frac{\#cites_{ij,t}}{\sum_{k=1}^{N_j} \#cites_{ik,t}} \right)^2$ where $\#cites_{ij}$ is the number of citations to patents labeled as AI by method i from CPC class j , using CPC codes at the 1-digit, 3-digit, and 4-digit level; $\#cites_{ij}$ excludes citations within the same class. N_j is the number of different CPC classes, and t is for a given time.

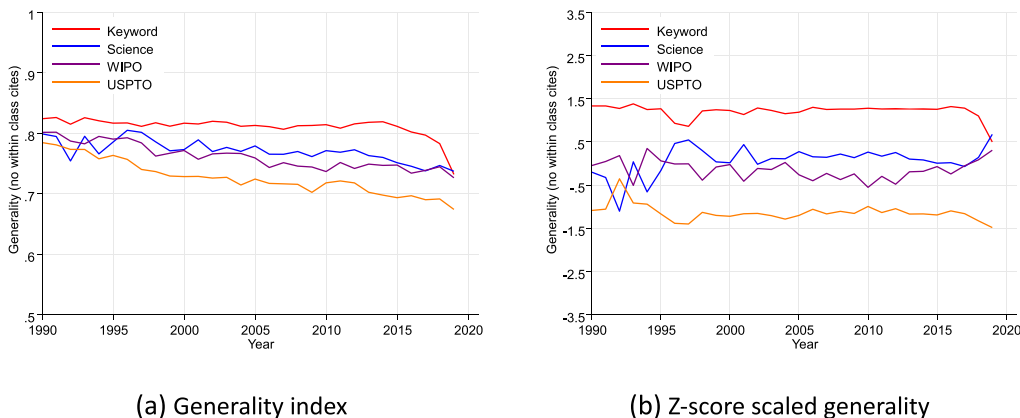


Figure 4. Generality index at the 1-digit CPC-section level (a). Note: The z-scored value equals the level of the generality index minus its average across the four approaches divided by the standard deviation for each year (b).

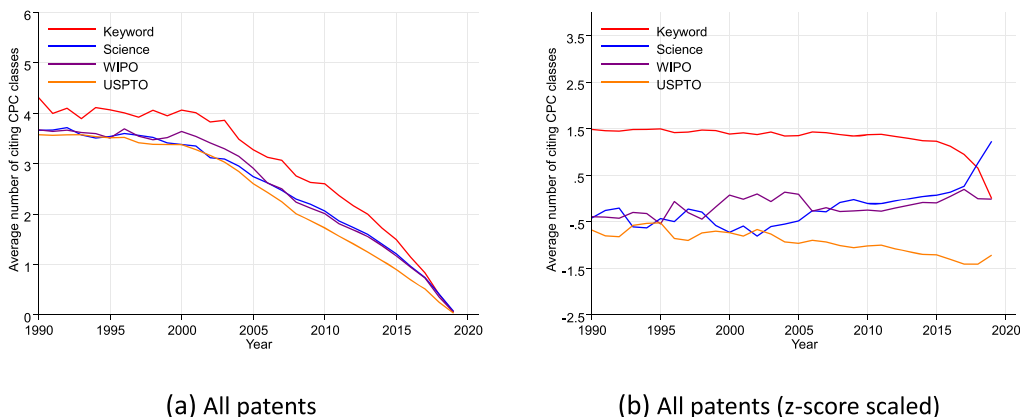


Figure 5. (a) Average number of classes citing AI. (b) The z-scored value equals the level of the generality index calculated at the 1-digit level minus its average across the four approaches divided by the standard deviation for each year.

surprisingly, given the large size of this sample. The decline in the generality of the Keyword and USPTO samples toward the end of the period should be interpreted with caution, as recently granted patents have had less time to be cited.

Second, we assess generality using the mean annual number of unique CPC classes that cite each AI patent. This measure accounts for the overall growth in patenting activity over time and avoids over-representing the generality of more recent patents. As before, we compute this metric at multiple CPC aggregation levels.

At all CPC aggregation levels, Keyword patents show highest generality (Table 2, Figure 5). At the 1-digit level, the Science sample ranks second, closely followed by WIPO. However, WIPO scores higher than Science at the more granular 3- and 4-digit levels. The USPTO sample consistently shows the lowest number of unique citing CPC classes across all levels. As before, the decline observed in more recent years can largely be attributed to citation time lags.

Significance tests (Appendix D2.2) confirm that the Keyword sample’s generality is statistically higher, and the USPTO sample’s lower generality is also significant. Differences between the Science and WIPO methods are negligible at the 1-digit level.

In Appendix D3.2, we additionally report generality based only on AI patents that received at least one citation (Table D39). This refines the analysis to include only relevant inventions, given that backward citations are often considered a proxy for patent quality (Barbieri et al., 2025). Here too, the Keyword sample shows the highest generality across all CPC levels. Science again

Table 3. Average citation lags by approach

Period	Keyword	Science	WIPO	USPTO
All periods	10.16	8.90	9.63	9.80
1990–1999	14.17	13.26	13.77	13.64
2000–2009	9.92	9.08	9.38	9.34
2010–2019	4.33	4.15	4.19	4.33

This table shows the average number of years taken until a patent in the sample is cited. The average number of years is lower in recent years, because the data ends in 2019 causing a truncation of the maximal time lag.

ranks slightly higher than WIPO at the 1-digit level, with WIPO scoring higher at the 3- and 4-digit levels. The USPTO sample remains the least general across all metrics.

As before, we benchmark our AI samples against all patents and other GPT candidates. [Table D2](#) shows that the generality index of all patents is higher than our AI samples, which is a natural feature, confirming the usefulness of the generality index in capturing the widespread use of patents. Among other groups, biochemistry/genetic engineering, nanotechnology, and climate inventions related patents have high generality scores. [Figure D5b](#) reveals that the generality index follows a relatively stable time series pattern, with some fluctuations toward the end of the period likely due to citation lags.

Using our second measure of generality—the mean number of unique citing CPC classes per AI patent—we find that all four AI samples exhibit higher generality than the overall patent population across all CPC aggregation levels ([Table D3](#)). Again, biochemistry/genetic engineering and climate-related technologies rank high by this measure. These results support the validity of our generality metrics, at least in relation to other established GPT candidates.

Additionally, we examine the generality of non-AI patents that cite AI inventions. While differences between the four AI samples are smaller for this group, the qualitative patterns remain consistent (see [Appendix D3.3](#)).

Patent citations to GPTs often occur with long time lags ([Hall and Trajtenberg, 2006](#)), as the diffusion of such technologies typically involves an early phase of “learning and destruction”—a period requiring organizational adjustments and complementary innovations before the GPT can be widely adopted ([Crafts, 2021](#); [Bresnahan, 2023](#)). In [Table 3](#), we report the average citation lags for each AI sample, calculated as the average number of years between a patent’s grant year and the grant years of citing patents. Keyword patents show the longest citation lags and once again score highest for GPTness. Taken together, the Keyword sample consistently ranks highest across all generality-related metrics, supporting its identification as the most GPT-like among the four classification approaches.

4.2.3 Complementarity

To measure complementarity, we examine the co-classification of AI by multiple CPC codes attached to patents. If AI complements a wide range of other technologies, then AI patents would be co-classified across a variety of technological fields.

[Figure 6](#) shows the share of 3- and 4-digit CPC classes covered by each AI patent set. The USPTO sample spans the broadest range of technology classes, covering between 70% and 90% of all possible CPC codes. This can be explained by the large number of AI patents in the USPTO approach, compared to the others ([Table 1](#)).

However, the smaller AI samples took off over time: beginning around 2010, the share of CPC codes associated with the Keyword, Science, and WIPO patents increased rapidly. Significance tests indicate that the differences among the Keyword, Science, and WIPO approaches are not statistically significant. However, all three score significantly lower than the USPTO sample (see [Tables D29](#) and [D31](#)).

To account for differences in sample sizes, we compute the annual average number of 1-, 3-, and 4-digit CPC codes per patent (see [Tables 2](#), [D46](#), and [D3.2](#)). Across all years, an average WIPO-classified patent is linked to 1.64 3-digit classes and 2.05 4-digit classes. Keyword and Science patents are associated with slightly fewer CPC classes on average, while USPTO patents exhibit the lowest degree of multidisciplinary by this measure. The superior performance of

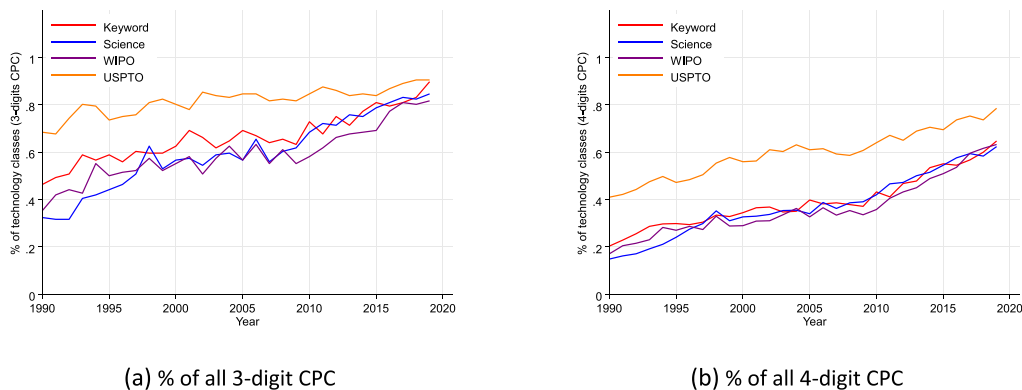


Figure 6. Diversity of AI—share of technology classes. *Note:* Panel (a) shows the percentage of 3-digit CPC codes and panel (b) shows the percentage of 4-digit CPC as a share of all codes in the respective category. Note that the total numbers of 3-digit and 4-digit CPC codes are 136 and 674, respectively according to the February 2022 version.

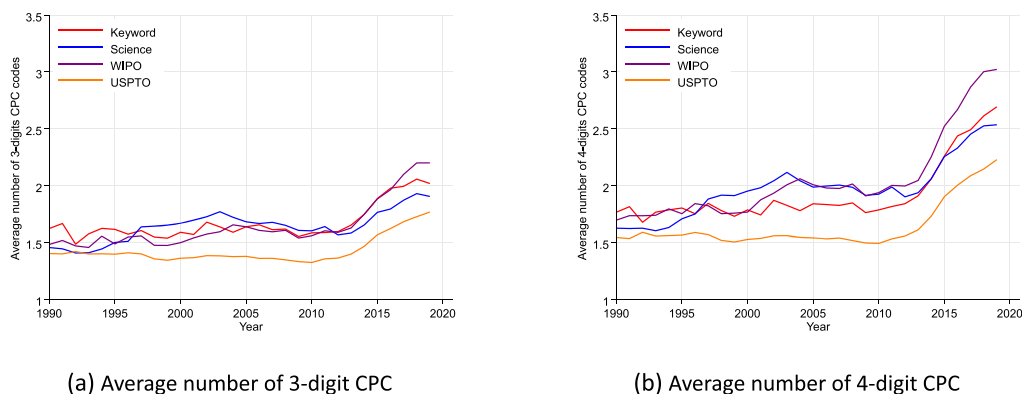


Figure 7. Patent-level diversity—average technology classes on different levels of granularity: (a) 3-digit CPC classification codes and (b) 4-digit CPC codes.

the WIPO method at the 4-digit level and the lower diversity of the USPTO sample are both statistically significant.

At the 1-digit CPC level, Keyword and Science AI patents exhibit similar values (1.39–1.40), and the difference between them is not statistically significant. By contrast, the other two approaches rank significantly lower, with the USPTO method displaying the least complementarity. Figure 7 illustrates the evolution of average co-classifications over time at the 3- and 4-digit levels. All panels indicate a trend toward increasing technological diversity in the second half of the last decade, with the most pronounced growth observed for WIPO patents.

In Appendix D1.1, we report results for other benchmark GPT candidates, confirming that our measures effectively reflect the broadening scope of these technologies, as suggested in the literature.¹⁰

Summing up, all AI samples exhibit increasing technological diversity over time. When accounting for differences in sample size, we find that the WIPO method captures the highest diversity at more disaggregated CPC levels, while the Science and Keyword approaches show greater diversity at the broader 1-digit CPC section level. The high diversity of WIPO patents at finer levels of classification likely reflects the method’s design, which targets patents assigned

¹⁰ Due to the heterogeneity of CPC classes, comparing diversity across different GPT candidates is challenging. Our selection of benchmarks is based entirely on CPC codes at the 3- and 4-digit levels. For instance, the Y02 codes used to identify climate technologies naturally span a broad range of technological areas, making them inherently more diverse.

to various definitions of computing, operational processes, and hardware components of AI (see [Appendix, Section B.4](#)). In contrast, the Science and Keyword approaches are not constrained by the CPC classes included.

4.3 Concentration in AI

Many policy debates around AI highlight concerns about increasing concentration in the market for AI technology, raising concerns about competition and equitable access to the benefits of AI ([Petit, 2017](#); [Jelinek *et al.*, 2021](#); [Mazzucato *et al.*, 2022](#); [Schmitt, 2022](#); [Babina *et al.*, 2024](#)). In the following, we provide empirical evidence on these concerns across different conceptions of AI, as captured by our four AI patent samples.

The Science approach captures the highest share of non-commercial AI inventions—those filed by individuals, non-profits, and universities. It also shows a strong concentration of AI patenting in the pharmaceutical sector, in contrast to the WIPO and Keyword approaches, which are more centered in the machinery and manufacturing industries. This reflects the Science sample's higher share of biotech-related patents, which are often filed by academic inventors and originate from university-based research.

Geographically ([Table 1, Panel D](#)), the Keyword and WIPO samples include a higher share of AI patents filed by foreign inventors, particularly from Japan. Notably, Chinese inventors are absent from the top ranks, despite China's increasingly prominent role in AI development in recent years ([Jacobides *et al.*, 2021](#)). This absence may be partly due to the time period under study, as China's emergence at the global frontier of high-quality patenting has primarily occurred in the most recent decade. However, it may also reflect a methodological bias stemming from the exclusive use of US patent data in this analysis.¹¹

Moreover, the four classification methods draw different pictures of the extent to which AI inventions rely on public R&D support ([Table 1, Panel E](#)). In the USPTO sample, only about one-fifth of AI patents received public funding, whereas in the Science sample, this share rises to nearly half in the most recent decade. This pattern aligns with differences in inventor types ([Table 1, Panel B](#)): AI patents identified by the USPTO method appear primarily commercially driven, while those in the Science sample are more often linked to academic inventors. This is consistent with the fact that public R&D funding is frequently channeled through universities and research institutions.

In [Table 4](#), we present the top 10 firms by their share of AI inventions for each of our AI samples. Additional results, including top assignees and concentration by technological field at the 1- and 3-digit CPC level, are reported in [Appendix D3.5](#). All methods consistently highlight AI patenting dominated by a small number of leading technology and communication firms. Notably, IBM, Microsoft, and Google occupy the top ranks across all methods. However, their top positions are concentrated within CPC sections G (Physics) and H (Electricity), regardless of the AI classification method. At a more granular level, their dominance is most pronounced in specific domains—namely, Electronic Communication (H04) and Computing (G06). For other technology fields, we observe a wider range of leading firms (see [Appendix D3.5](#)).

To assess concentration beyond the top-10 firms, we calculate standard concentration measures—the concentration ratio (CR) and the HHI—reported in [Table 5](#). The four-firm (eight-firm) CR captures the share of AI patents attributed to the top four (top eight) assignees, while the HHI is calculated as the sum of squared patent shares of all firms, providing a more comprehensive measure of market concentration. Across measures, Keyword patents exhibit lowest concentration (HHI and CR values). Further disaggregation by 1-digit and 3-digit CPC codes is presented in [Appendix D3.5](#), showing how concentration varies across technological fields.

Contrary to concerns raised in the literature ([Petit, 2017](#)), we do not find evidence that AI patenting has become increasingly concentrated among fewer firms over the past three decades. While we observe some fluctuations in concentration levels, there is no consistent upward trend across any of the four AI samples. It is important to emphasize that our analysis is limited to the distribution of patented AI inventions and does not capture other dimensions of concentration. In

¹¹ A discussion of the implications of this bias can be found in Section 6.

Table 4. Top AI-producing firms

Keyword		Science		WIPO		USPTO	
Company name	%	Company name	%	Company name	%	Company name	%
IBM Corp	0.05	IBM Corp	0.07	IBM Corp	0.07	IBM Corp	0.09
Microsoft	0.02	Microsoft	0.06	Microsoft	0.04	Microsoft	0.04
Samsung	0.02	Google	0.03	Google	0.03	Google	0.02
Fanuc Corp	0.02	Apple	0.02	Canon	0.02	Intel	0.02
Google	0.02	Sony	0.01	Samsung	0.02	Samsung	0.01
Siemens	0.02	Siemens	0.01	Sony	0.02	Hewlett Packard	0.01
Honda Motor	0.01	Hewlett Packard	0.01	Intel	0.01	AT&T	0.01
Amazon	0.01	Intel	0.01	Amazon	0.01	Sony	0.01
Intel	0.01	AT&T	0.01	Siemens	0.01	Amazon	0.01
Sony	0.01	Canon	0.01	Fujitsu	0.01	Canon	0.01

This table reports the top-10 AI producing firms for each AI definition. IBM, Microsoft, and Google are among the top-five AI patenting firms across all four groups. The column share reports the share of commercial patents accounted by a firm within each AI definition. For example, IBM accounts for 4%–7% of all AI patents produced by commercial firms.

Table 5. Concentration of firms innovating in AI

	Keyword	Science	WIPO	USPTO
A. Concentration ratio (CR)				
Four-firm CR	0.117	0.174	0.162	0.177
Eight-firm CR	0.175	0.219	0.221	0.228
B. Herfindahl–Hirschman Index (HHI)				
HHI (overall)	0.007	0.012	0.011	0.014
HHI (1990–1999)	0.008	0.012	0.013	0.014
HHI (2000–2009)	0.006	0.013	0.015	0.016
HHI (2010–2019)	0.009	0.013	0.011	0.014

This table shows the measures of concentration of AI-producing commercial firms. The concentration ratio (CR) measures the market share of top-four (or top-eight) firms. The Herfindahl–Hirschman Index (HHI) is calculated as the sum of squares of the shares of patent produced by each firm within each of the four AI definition patent samples.

particular, broader concerns about AI-driven concentration often relate to final goods markets—where productivity gains may accrue disproportionately to dominant firms (Babina *et al.*, 2024)—or to the control over strategic inputs such as data, computational infrastructure, and digital platforms (Jacobides *et al.*, 2021; Franco *et al.*, 2023), which fall outside the scope of our patent-based approach.

Furthermore, our analysis of field-specific dominance and concentration of AI inventions across 1- and 3-digit CPC classes reveals no substantial variation in concentration levels across technological fields or classification approaches. While aggregate indicators suggest that overall concentration remains limited, we do observe a set of dominant actors operating within particular niches of AI technology. This pattern is consistent with the conceptualization of AI as an ecosystem comprising interdependent technologies, applications, and complementary assets (Jacobides *et al.*, 2021). From the perspective of patented inventions, our findings do not provide evidence of a general trend toward rising concentration across the AI landscape.

5. Discussion

Defining AI remains challenging, as it is an emerging technology whose full impact is yet unknown (Schumpeter *et al.*, 2005; Krafft *et al.*, 2020; Barbieri *et al.*, 2025). The evolution of AI is shaped not only by technological progress but also by socio-technical dynamics and regulatory frameworks (Geels, 2005). Shaping AI has gained increasing attention in current discussions on AI governance, aimed at mitigating risks and maximizing benefits for all (Mazzucato *et al.*, 2022; Bresnahan, 2023). Effective AI policy-making relies on the ability to clearly delineate the domain

Table 6. Summary of findings: general-purpose technology, concentration, and sample size

	Keyword	USPTO	WIPO	Science	Metric	Based on
Growth	Strongest	Weakest	Middle	Middle	Growth rate	Counts
Generality	Strongest	Weakest	Middle	Middle	Generality index	Citations
Complementarity	Middle	Weakest	Strongest	Middle	Avg. # tech. classes	CPC codes
Concentration	0.009	0.014	0.011	0.013	HHI index	Commercial AI patents
Sample size	54 k	595 k	159 k	178 k	Counts	All AI patents

This brief summary of our results shows which patent group generates the strongest average estimate of each GPT characteristic over the last 10 years. Top three rows: Comparing the GPT characteristics of the different classification methods into either strongest, weakest, or in the middle (between strongest and weakest).

of AI and to identify areas where regulatory intervention or targeted support may be warranted (Krafft *et al.*, 2020).

Our analysis of GPTness and concentration provides insights relevant to two key dimensions of AI policy. First, policy efforts could prioritize forms of AI that generate the greatest public value—those characterized by broad technological spillovers and societal benefits (Mazzucato *et al.*, 2022; Bresnahan, 2023). Such AI inventions may be identified by their high levels of GPTness, which reflect their capacity to stimulate follow-on innovation across the economy (Lipsey *et al.*, 2005).

Second, AI policy could aim to mitigate the risks of power concentration that arise when core resources to develop and deploy AI are unequally distributed. Addressing concentration helps prevent the premature narrowing of innovation pathways toward areas with the greatest short-term private returns, while neglecting other areas with broader societal benefits (Klinger *et al.*, 2020; Jacobides *et al.*, 2021; Bresnahan, 2023).

5.1 AI as a general-purpose technology

All four AI classification methods show consistent trends: limited patenting in the 1990s (“AI winter”) followed by strong growth post-2000 (Stuart and Norvig, 2003; Klinger *et al.*, 2020). Each supports the view of AI as a GPT and associate AI with similar industries and core technologies. They also reveal comparable patterns of concentration among a similar set of key firms. These shared characteristics are consistent and informative for scholars, policy-makers, and analysts who are investigating what findings are clear across different definitions of AI patents, as a basis for research and regulatory action (Krafft *et al.*, 2020).

Quantitatively, we observe some differences across approaches: AI patenting identified by the Keyword and WIPO methods shows an acceleration in the 2010s, whereas the other approaches suggest a relative slowdown. Keywords and WIPO also identify patents that are most GPT-like and capture more from machinery manufacturing. This is reflected in the key firms emerging as top AI inventors, particularly within computing and electronic communication technologies. The Keyword method, while capturing the narrowest set of inventions (~54 k patents), shows the highest levels of GPTness (Table 6) and the lowest levels of concentration in inventive activity. Framing AI as a GPT using our patent-based metrics aligns with patterns observed in historical GPTs—such as computers, communication technologies, and electrical engineering—in patent data (Petralia, 2020).

5.2 Concentration in AI patenting

Concentration and GPTness interact, as GPTs have public good characteristics, with high innovation spillovers and lagged private returns for diverse actors, reducing the incentives for short-term profit maximization (Bresnahan and Trajtenberg, 1995). This may help explain the lower concentration levels in AI inventions that are most GPT-like. The low concentration of the Keyword-based AI sample aligns with their high GPTness ranking (Section 4.2), reinforcing the interpretation that they capture most GPT-like segments of AI. The Keywords many patents centered on ML, a technology also highlighted by others as having high GPT potential (Goldfarb *et al.*, 2023).

The full realization of GPT benefits is not guaranteed. Commercial actors may prioritize private returns over social value, potentially resulting in premature lock-in to suboptimal trajectories of AI development (Klinger *et al.*, 2020; Mazzucato *et al.*, 2022; Bresnahan, 2023). Our identification of core technological domains—particularly those captured by Keywords with the strongest GPTness and lowest concentration—may indicate areas in which public support could be justified. However, any such decisions should be informed by methodologically pluralist and up-to-date evidence, as each classification approach—including those used in the present study—has inherent limitations (see Section 6).

Given their GPTness, Keyword-identified technologies are strong candidates for targeted support if there is a preference for public good characteristics. However, when addressing innovation bottlenecks, the challenges may not lie primarily at the level of invention—as captured by patents—but rather at the diffusion stage, where business transformation and AI adoption take place (Bresnahan, 2023). This aligns with concerns about concentration, as the capacity to adopt and benefit from AI systems remains unevenly distributed across firms (Babina *et al.*, 2024).

We observed no evidence of increasing concentration at the invention stage, suggesting that entry barriers in AI innovation have not intensified—contrary to some earlier concerns (Agrawal *et al.*, 2019; Tambe *et al.*, 2020; Mazzucato *et al.*, 2022). This may reflect the nature of digital inventions, which are typically characterized by low capital intensity, short technological cycle time, and high modularity (Bresnahan, 2023). The modular structure of the AI ecosystem facilitates the entry of new players, particularly when computational resources and hardware components can be sourced externally at reasonable costs (Jacobides *et al.*, 2021; Bresnahan, 2023). Our analysis of concentration across technological subfields further reveals a diverse range of key actors, each specialized in particular fields. However, our patent-based analysis does not capture trends in vertical integration across the AI value chain, control over key network resources, or emerging patterns of polarized AI-driven productivity growth—all of which are relevant to competition and antitrust policy (Petit, 2017; Ducuing, 2020; Franco *et al.*, 2023; Rikap, 2024).

5.3 Research guidance on measuring AI patents

Our comparison highlights that empirical conclusions in patent-based AI research vary significantly with classification method. For instance, the USPTO and Science approaches suggest a recent slowdown in AI invention activity—a trend not observed with the Keyword or WIPO methods. These discrepancies can lead to diverging interpretations and explanations, such as whether a decline or rise reflects a narrowing and corporatization of AI research (Klinger *et al.*, 2020; Jee and Sohn, 2023). We also find minor but meaningful differences in the top-ranked firms identified as AI leaders and their nationalities. These distinctions matter when analyzing the performance of national innovation systems, given their influence on emergent pathways of AI specialization (Jacobides *et al.*, 2021; Rikap, 2024). For example, the Science approach includes Apple among the top-ten AI patentees but excludes Samsung, which consistently ranks among the top five in the other methods. Such variations can have material consequences when investors, banks, and policymakers rely on patent-based indicators to assess economic performance and inform funding decisions.

Our comparison may inform patent-based AI research in selecting classification methods aligned with specific research objectives. Our findings suggest that *AI as a GPT* might be best studied by using the Keyword or WIPO approaches. The USPTO method appears more appropriate for analyzing the broad diffusion of AI, while the Science approach may be best suited for studying the interactions between research and AI development.

The simple Keyword approach—which yields a relatively narrow set of key patents—appears well-suited for research on emerging GPTs. Smaller patent groups may offer a clearer distinction from the broader landscape (Kovács *et al.*, 2021) and may signal greater potential for future growth. This contrasts with the USPTO method, which classifies 16.6% of all US patents in 2019 as AI-related, potentially overstating its current breadth.¹² However, we cannot rule out that the

¹² During our analysis, we also discovered that the Keyword method reproduced from Cockburn *et al.* (2018) may be further simplified. We found that the majority of patents can be identified using a narrower set of four terms (ML, neural network, robot, pattern recognition), rather than the original list of over 40 words.

high levels of GPTness of the Keyword method reflect only a short-term trend of the ML uptake in commercially valuable inventions, as captured by patents, while the full scope of AI remains unleashed (Klinger *et al.*, 2020).

Measuring emerging technologies is inherently difficult as definitions clarify only after development, diffusion, and final use (Schumpeter *et al.*, 2005; Lafond and Kim, 2019; Barbieri *et al.*, 2025). This paper quantifies areas of consensus and nuanced differences in AI classification. While we can only speculate, this emerging consensus may signal the type of AI that is most likely to be advanced in the future.

6. Limitations

Our data ends in 2019 and reflects a view of AI shaped by pre-ChatGPT developments—prior to the rapid rise of large language models (LLMs) and generative AI. As such, it captures how AI was defined and measured before these transformative technologies started to shape discussion in the field. Assessing how recent trends affect our four classification approaches and GPT assessment is beyond the scope of this study, given its complexity. Patent classification systems coevolve with technology, especially rapidly changing areas of technology (Barbieri *et al.*, 2025). Non-systematic checks of recent amendments in the CPC system reveal a high revision intensity in CPC codes related to those used by the WIPO approach and used in our GPT assessment. Changes in the CPC codes maybe themselves an indicator of the evolving nature of AI.

Keyword based methods are also time-sensitive. Exploratory queries to recent LLMs (ChatGPT and Mistral) seeking Keywords to identify current AI patents generated lists still predominantly centered on ML, but also included terms associated with specific AI edge applications hardware components and integrated AI systems.¹³ This shift aligns with the observed diffusion of AI across a growing number of sectors. Consequently, we would also expect the range of scientific fields contributing to AI innovations to have expanded significantly.

Expanding our cross-sectional analysis of AI classification into a longitudinal dimension would be a valuable endeavor. Such an analysis could yield new insights into the evolution of AI trajectories, which may be relevant for policy. Moreover, reclassification dynamics of emerging technologies is relatively understudied, yet understanding this process could help in technological forecasting and policy implementation, with relevance beyond AI (Barbieri *et al.*, 2025).

Beyond these conceptual limitations, our research comes with other well-documented limitations and biases related to patent data. First, patents are utilized heterogeneously across industries, firms, and technologies, and some inventions are unsuitable for patent protection. Alternative mechanisms of intellectual property protection may predominate in specific sectors, leading to biases in patent-based analyses (Granstrand, 2009). Additionally, variations in the frequency of co-classifications and citations across technological fields (Jaffe and De Rassenfosse, 2017; Hötte *et al.*, 2021) may influence empirical measures of GPTness.

Second, patents capture technical inventions but do not reflect key dimensions of AI-driven technological change, such as diffusion, process innovations, input substitution, AI-related services, data resources, and other complementary assets (Jacobides *et al.*, 2021; Bresnahan, 2023; Goldfarb *et al.*, 2023). Nevertheless, our findings align with studies using alternative indicators—for example, Goldfarb *et al.* (2023), who draw on occupational skill requirements and conclude that ML qualifies as a GPT.

Third, our analysis is limited to USPTO patents. While US patents are often considered a good proxy for the global technological frontier, this may not hold for AI. National specialization in AI varies considerably, reflecting differing priorities and institutional contexts (Fujii and Managi, 2018; Jacobides *et al.*, 2021; Rikap, 2024). Although foreign inventors can and do file at the USPTO, they may encounter barriers or have reduced incentives compared to filing at home. In addition, differences in examination practices, citation behavior, and classification standards across patent offices complicate the generalization of our analyses to the international level (Jaffe and De Rassenfosse, 2017).

¹³ We conducted exploratory queries in July 2025 using various versions of ChatGPT (4.5, 4o) and LeChat from Mistral. Notably, Mistral highlighted a larger number of hardware-related terms.

Fourth and in addition to our discussion above, the observed levels of GPTness across the different approaches may be specific to the time period analyzed (1990–2019). The high GPTness found for the Keywords and WIPO may reflect the recent dominance of a narrow set of technology-related buzzwords (especially ML). These results might differ if reproduced in the future, as the boundaries and focus of AI continue to evolve. This limitation relates to the debate on “AI narrowing” (Klinger *et al.*, 2020), which illustrated how recent AI research has become less diverse, concentrated within a few large companies, and focused on a limited set of ML techniques. However, the future breadth of AI development may be shaped by political and regulatory decisions made today.

Lastly, our analysis is restricted to GPTness and concentration as criteria relevant for AI policy, while remaining silent on other objectives such as safety, societal impact, and privacy (Krafft *et al.*, 2020; Roche *et al.*, 2023). Efforts to empirically assess the impact of policy on the evolution of AI should draw on a range of AI measurement methods and incorporate evaluation criteria that are aligned with the goals set by policy.

7. Conclusion

We conduct a systematic analysis of four distinct approaches to identifying AI patents, each capturing a different set of patents, with only a small overlap. Our findings show that both quantitative and qualitative assessments of AI are sensitive to the choice of classification methods. In addition, we examine patterns of concentration in AI innovation. Despite methodological differences, we find consistent qualitative overlaps across key dimensions, which are strong compared to benchmark technologies. This suggests the emergence of a partial convergence in our understanding of what constitutes AI.

Our analysis offers guidance for policymakers and innovation scholars seeking to identify patented AI inventions. For example, researchers aiming to study the role of AI as a GPT may find that a simple Keyword-based method captures a narrow set of patents that best exhibit canonical GPT features: endogenous growth, broad usefulness, and technological complementarity.

Taken together, our results (i) provide robust empirical support for conceptualizing AI (particularly ML) as exhibiting key characteristics of a GPT, (ii) demonstrate the usefulness of patent data for tracking AI innovation, and (iii) underscore the importance of employing multiple classification methods to reduce methodological biases, especially when assessing emerging technologies and issues like market concentration.

Acknowledgments

The authors thankfully acknowledge valuable feedback received from the participants of the OMPTEC-FoW and INET complexity group meetings. K.H. and T.T. thankfully acknowledge support from OMPTEC and Citi. V.V. acknowledges support from Intelligence Advanced Research Projects Activity (IARPA), via contract no. 2019-1902010003, for developing capabilities that this work builds upon. L.B. gratefully acknowledges support from the General Sir John Monash Foundation.

Appendix A. Measuring GPTs

Currently, there exists a number of alternative metrics to capture GPT characteristics. Given the lack of consensus, many believe GPTs should be better identified as sophisticated networks of technologies sharing “underlying principles and mutual dependencies” (Petralia, 2020).

Historically, patent growth rates have been used to capture the endogenous elaboration of technologies similar to GPTs (Moser and Nicholas, 2004; Jovanovic and Rousseau, 2005; Petralia, 2020). Petralia (2020) uses patent growth rates, co-classifications, and a text-mining algorithm to successfully reproduce the canonical GPTs contained within the broad USPTO categories of electricity and computer communication. However, the author finds great heterogeneity within these pools of patents, which contain both dynamic and stagnant inventions. Moreover, the author notes that the identification of more diffuse and diverse GPTs, such as AI, may require “bottom-up” classification approaches using lower levels of aggregation that can scan multiple technological classes for common principles.

Hall and Trajtenberg (2006) attempt to capture GPTs by measuring the patent growth rates and unbiased generality measures for the most-cited US patents and the patents which cite them. The authors also find great heterogeneity between patents, which underscores the need for multiple metrics to satisfactorily capture GPTs.

In the next section, we motivate our selection of patent measures for GPT characteristics and connect each with empirical facts about AI’s dissemination and the three canonical GPT features.

A.1. Growth

For more than a decade, AI methods have become more powerful and complex as a result of new technical methods, increased data availability, and improved hardware. Consequently, AI invention has shifted away from specific application-based methods to more generalized learning-orientated systems (Cockburn *et al.*, 2019). With this refinement, the performance of many sub-fields of AI, such as image and text recognition, has seen remarkable improvements in performance (Brynjolfsson *et al.*, 2021). This is reflected in the exponential growth of patenting activity referencing terms such as ML and deep learning (see Appendix, Figure C2).

Based on these observations, we measure improvements in AI via the growth rates of each group of patents and changes to their share of all patents, from 1990 to 2020 (Hall and Trajtenberg, 2006; Petralia, 2020). We also look at the growth of the patents that cite such technologies: the “GPT hypothesis” in previous work has been that inventions that build on GPT-like technologies should spawn more new inventions (Hall and Trajtenberg, 2006).

Let $N_{i,t}$ denote the number of patents in a group $i \in \{\text{keyword,science,WIPO,USPTO}\}$ at time t , indexed by year. We compute the growth rate as

$$(N_{i,t} - N_{i,t-1})/N_{i,t-1} \quad (\text{A.0})$$

A.2. Generality

AI has already begun to pervade a myriad of industries because it expands beyond computer science into such diverse fields as structural biology, transport, and imaging (Cockburn *et al.*, 2019). In the early 1990s, AI methods remained largely confined to computer science. However, over the past decade, the majority of patents referencing these technologies have appeared in secondary domains (Cockburn *et al.*, 2019). Based on the work of Trajtenberg *et al.* (1997), we capture this stylized fact through the “generality” of patents, measuring the dissemination of AI across different technology fields.

To do so, we build on patent citation data and assume that a forward citation link entails information about the use of a patent in a subsequent invention (Jaffe and De Rassenfosse, 2017). To operationalize wide usefulness, we rely on a modified version of the generality metric by Trajtenberg *et al.* (1997) and Hall and Trajtenberg (2006) given by

$$1 - \sum_i^{N_j} \left(\frac{\#\text{cites}_{ij,t}}{\sum_{j=1}^{N_j} \#\text{cites}_{ij,t}} \right)^2 \quad (\text{A.0})$$

where $\#\text{cites}_{ij}$ is the sum of citations to patents labeled as AI by classification approach i from technology class j , whereby we use the CPC 1-digit level as class. The number of citations $\#\text{cites}_{ij}$ excludes citations within the same class: N_j is the number of different CPC classes, and t is the time. Our approach differs to that of Trajtenberg *et al.* (1997) as we apply the method to each group of AI patents belonging to a variety of CPC sections. For the main analysis, we focus on 1-digit CPC sections, as these are more technologically distant than 3-digit or 4-digit classes and subclasses, whose results we also report.

Our generality measure is calculated for the entire group of patents in i with N_i unique patents. To address concerns that this metric may be affected by differences between group sizes, we additionally

calculate patent-level metrics given by the average number of citing classes, i.e.

$$\frac{1}{N_{i,t}} \sum_{p=1}^{N_{i,t}} \sum_{j=1}^{N_d} 1(\#ncites_{p,j,t} \geq 0) \quad (\text{A.0})$$

where $1(\#ncites_{p,j,t} \geq 0) = 1$ if patent p in i is cited by at least one patent in technology class j out of the total number of classes N_d at level with $d \in \{1,3,4\}$ in the code. N_i is the number of patents in approach i , and t is the time. Again, we exclude within-class citations and present results at both the 1-digit CPC section level ($d = 1$) and higher orders of disaggregation ($d = 3$ or $d = 4$).

A.3. Complementarity

Thirdly, GPTs augment existing products and processes in a range of novel contexts to generate productive complementarities throughout the economy (Bresnahan and Trajtenberg, 1995; Petralia, 2020). AI technologies have been shown to complement and rely on secondary inventions, related to areas such as cloud computing and big data, which increase access to larger and more affordable data-sets (Brynjolfsson *et al.*, 2019). Furthermore, because diverse AI systems share similar underlying structures and can share information, advances in one application of ML, such as machine vision, can spur inventions in other fields, such as autonomous vehicles.

Following the approach of Petralia (2020), we measure the extent to which AI patents enhance and supplement other inventions through the diversity of their technology class co-occurrences. For our analysis, we calculate the share of 3- and 4-digit CPC codes ($d = 3,4$) assigned to the patents in each group of AI patents. Specifically, we calculate the following *diversity measure* over time is

$$\frac{\#CPCs_{i,d,t}}{N_d} \quad (\text{A.0})$$

where i denotes each of the four patent classification approaches, d is the classification level and t is the year. N_d refers to the number of CPC codes found in use for a particular group of patents, where the codes include d digits. Note that there are 136 and 674 CPC codes, respectively at the 3- and 4-digit level (according to the February 2022 version of CPC codes).

As the above measure could be biased by patent volume, we also calculate the average number of distinct 1-, 3-, and 4-digit CPC codes per patent per year. The *diversity per patent* over time is

$$\frac{1}{N_{i,t}} \sum_p \#CPS_{p,i,d,t} \quad (\text{A.0})$$

where d represents the technology class represented by 1-, 3-, or 4-digit CPC codes, and t is the time. The time series graphs for the latter measures depict how an average patent's complementarity across technology sections evolves over time.

B. Measuring AI

In our analysis, we compare four methodologically and conceptually distinct approaches to identifying AI inventions in patents based on (i) Keyword search, (ii) science citations, (iii) the WIPO, and the (iv) USPTO method. Here, we introduce these classification approaches in detail.

B.1. Data source

We apply our methods to all patents granted by the USPTO from 1990 to 2019. For the analysis, we create four groups of AI patents for each classification method and complement each with supplementary information.

From PATSTAT (Spring 2021 edition, [EPO 2021]) we sourced patent grant dates and from the USPTO we downloaded the Master classification file (April 2021 version) which contains CPC classifications of patents.¹⁴ We added further data on patent-to-patent citations and patent titles from GooglePats obtained in an earlier project (Hötte *et al.*, 2021). For our analysis, we

¹⁴ <https://bulkdata.uspto.gov/data/patent/classification/cpc/>

supplemented the citation data with citation year and the technology classes of both the citing and cited patent. In doing so, we obtained networks that represent citations from technology fields at different levels of aggregation to our four sets of AI patents. We also made use of the Reliance on Science database (Marx and Fuegi, 2020) for citation data between patents and science.

B.2. Keyword search

Our first classification technique is a straight-forward approach based on Keyword search, in which researchers use their discretion to develop a set of terms that reflect the very recent developments in AI. In this paper, we use the set of Keywords provided in the appendix of Cockburn *et al.*, 2018.¹⁵ The Keywords used in this paper focus on three subfields of AI: symbolic systems, learning algorithms, and robotics (see Table C1 for the full list of Keywords). According to the authors, the symbolic systems represent “complex concepts through logical manipulation of symbolic representations” and include “NLP” and “pattern recognition.” Learning algorithms include core analytic techniques such as neural networks, deep learning, and ML. The last category, robotics, is related to automation or applications of AI (e.g. computer vision and sensory networks).

We search for these Keywords in patent titles, abstracts, claims, and descriptions using USPTO data. We match the resulting list with patents granted by the USPTO between 1990 and 2019. The main advantage of the Keyword approach is its simplicity and ease of implementation. Moreover, carefully chosen Keywords can capture recent changes in the AI field. However, the success of this approach depends on the judgment and familiarity of the researcher to the field of AI. Missing important Keywords could lead to under-representation of a subfield. Our approach yields 67,187 patents.

B.3. Science citations

This classification approach harnesses the scientific basis of patents. In particular, we classify a patent as an AI patent if it makes at least one citation to a scientific paper in the scientific field of “Computer Science; Artificial Intelligence” (short, AI paper) as categorized by the Web of Science (WoS). Scientific citations are added to patent documents for multiple reasons such as describing the technological content of the invention or distinguishing the legal claim from other publicly available knowledge (see Narin *et al.*, 1995; Meyer, 2000; Tijssen, 2001; Ahmadpoor and Jones, 2017; Marx and Fuegi, 2019). A citation link to an AI paper indicates that the patent is technologically related to AI because it builds on scientific advancements in this field. A limitation of this approach is that it only identifies AI patents within the subset of patents that make citations to science.

For this method, we use data from the Reliance on Science (RoS) database (Marx and Fuegi, 2019, 2020), which comprises a mapping from patents to scientific articles indexed in Microsoft Academic Graph (Sinha *et al.*, 2015). Scientific articles are tagged by the WoS fields indicating the field of science into which an article is grouped.¹⁶

The citation links in the RoS database cover citations made by both the patent applicant and examiner, as well as citations indicated at both the front page and body of the patent document. Marx and Fuegi (2019) identified citations through a sequential probabilistic text recognition technique. Each citation link is tagged with a confidence score indicating the reliability of the matching approach. In the RoS data, roughly one third (34%) of all US patents granted in 2019 can be attributed with at least one citation to science.

In our study, we identified AI papers by their WoS categories and extracted all patents with at least one citation link to an AI paper. We kept only citation links with a reliability score greater than three, which corresponds to a precision rate of 99.5% and a recall of 93%. This approach yields 178,004 AI patents.

¹⁵ While we use the keywords from Cockburn *et al.* (2018), we do not fully replicate their approach. They use two subsets of patents: (i) patents classified by the USPC code 706 (Artificial Intelligence) and 901 (Robots); and (ii) patents identified by searching titles for the selected keywords. Here we use patents identified by keyword search only, but we extend our search to match keywords also from abstract, claims, and description. We do not use the USPC classification codes since the WIPO method takes a more comprehensive approach combining keywords with IPC or CPC classifications. Also, with our extensive keyword search, we miss only a few patents which are in the first group (i.e. USPC 706 and 901) but not in the second group of Cockburn *et al.*, 2018.

¹⁶ Note that this assignment was made at the paper level using a probabilistic mapping, which is different from the journal-based categorization of Clarivate Analytics (WoS).

B.4. World Intellectual Property Organization method

The WIPO methodology for classifying AI patents was published in 2019 and validated by a team of patent experts (WIPO, 2019a,b). The aim behind the methodology is to capture three aspects of AI invention: (i) core AI techniques (deep learning, other learning methods, various type of logic, clustering, etc.); (ii) functional applications of AI that can be used to simulate human-like cognitive capacities (such as vision, language, or decision-making); and (iii) end-user application fields (such as automation in business, health, or military).

This methodology is based on both a Keyword search of patent texts and the use of patent classification codes (CPC and IPC). In this technique, some patents are classified based on only a subset of the technological codes, or Keywords, whilst others are identified by a combination of both.

The list of Keywords used in this approach covers core AI methods as well as computing and mathematical concepts used in such technologies. These Keywords are matched to the text in the patent titles, abstracts, and claims.

This approach identifies 158,652 patents.

B.5. United States Patent and Trademark Office classification

The USPTO approach uses a supervised ML classifier to identify AI patents (see Giczynski et al., 2021). This ML model is trained to classify eight components of AI technologies, namely: ML, evolutionary computation, NLP, speech, vision, knowledge processing, planning/control, and AI hardware. The ML model is trained on the abstracts and claims of a seed (positive set) and an anti-seed (negative set). The seeds are chosen carefully for each respective component by taking an intersection of CPC, IPC, and USPC codes, as well as Derwent’s World Patents Index™. The seeds are expanded based on patent families, CPC codes, and citations to identify all patents linked to the seed set. The anti-seed set is selected randomly from all remaining patents. For training, each text is pre-processed and embedded via the Word2Vec algorithm. The ML models also encode backward and forward citations in a citation vector. The predictions from the ML model are further validated using a small subset of patents that are manually examined.

Published in August 2021, the resulting dataset contains 13.2 million USPTO patents and pre-grant publications issued or published between 1976 and 2020. For consistency with our other approaches, we only consider patents granted between 1990 and 2019 and exclude pre-grant publications. The remaining data yield 595,047 patents.

C. Keywords in detail

Table C1. List of Keywords from Cockburn et al. (2018)

Symbols	Learning	Robotics
Natural language processing	Machine learning	Computer vision
Image grammars	Neural networks	Robot
Pattern recognition	Reinforcement learning	Robots
Image matching	Logic theorist	Robot systems
Symbolic reasoning	Bayesian belief networks	Robotics
Symbolic error analysis	Unsupervised learning	Robotic
Pattern analysis	Deep learning	Collaborative systems
Symbol processing	Knowledge representation and reasoning	Humanoid robotics
Physical symbol system	Crowdsourcing and human computation	Sensor network
Natural languages	Neuromorphic computing	Sensor networks
Pattern analysis	Decision making	Sensor data fusion
Image alignment	Machine intelligence	Systems and control theory
Optimal search	Neural network	Layered control systems
Symbolic reasoning		
Symbolic error analysis		

C.1. Words used in the Keyword approach

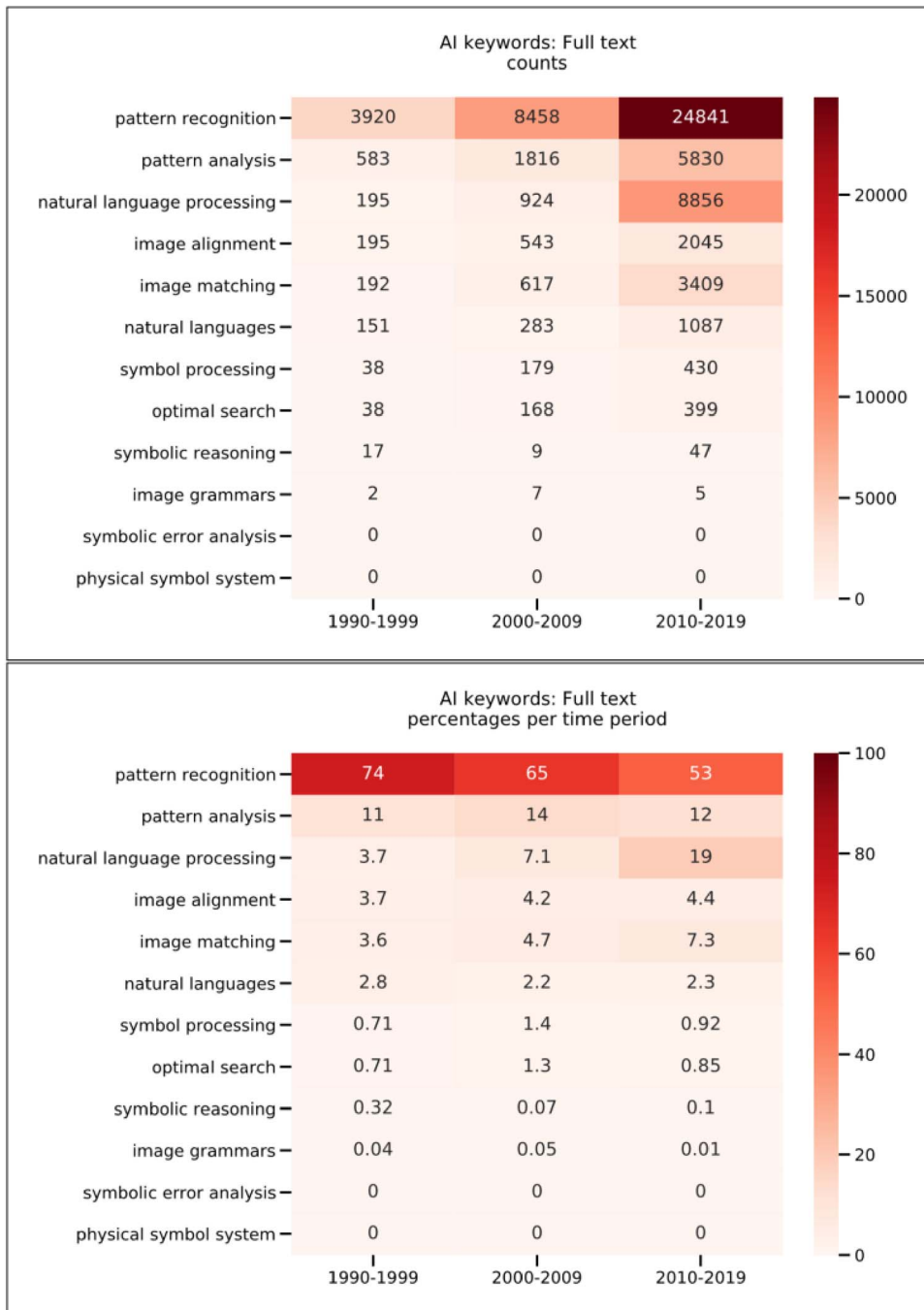


Figure C1. Symbolic Keywords: in full texts.

C.2. AI Keywords in patent texts

We split all the patent texts into three time periods (1990–1999, 2000–2009, and 2010–2019) and search through the texts for Keywords. Then, in each period (and for each



Figure C2. Learning Keywords: in full texts.

category) we count the unique number of matching documents and what percentages of the AI patents match according to this Keyword. [Figures C1–C3](#) below illustrate both counts and shares.

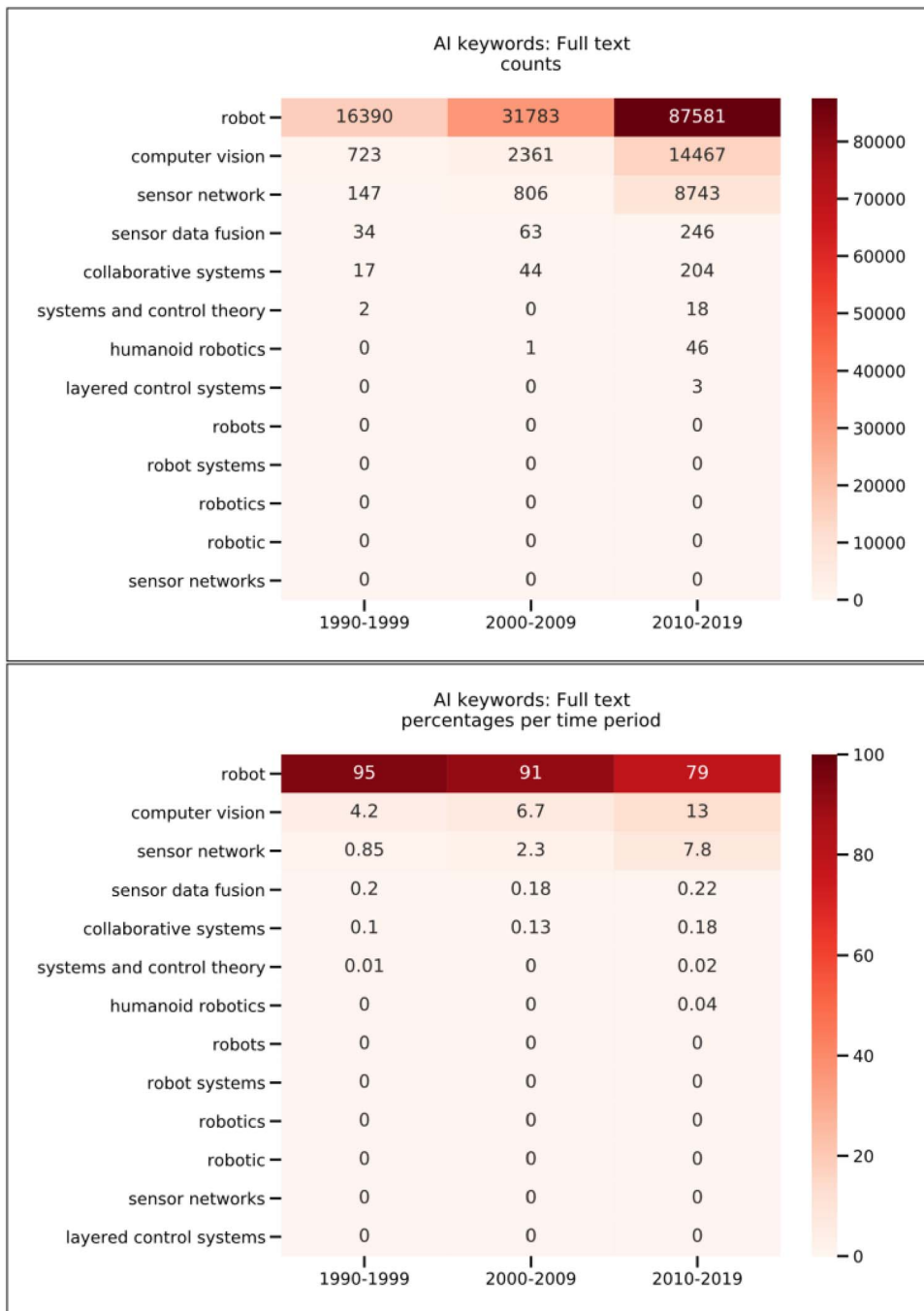


Figure C3. Robotics Keywords: in full texts.

D. Supplementary results

D.1. Comparison to benchmarks

The following figures reproduce time series of growth rates, counts, and shares for additional groups of patents. The benchmarks were identified in previous discussions of GPT technologies in the literature (nanotechnology, biochemistry, green technologies, computing). Climate patents were also included as a group of technologies where one can expect wide diversity, as climate inventions can be expected to cover many sectors of the economy.

D.1.1. Growth

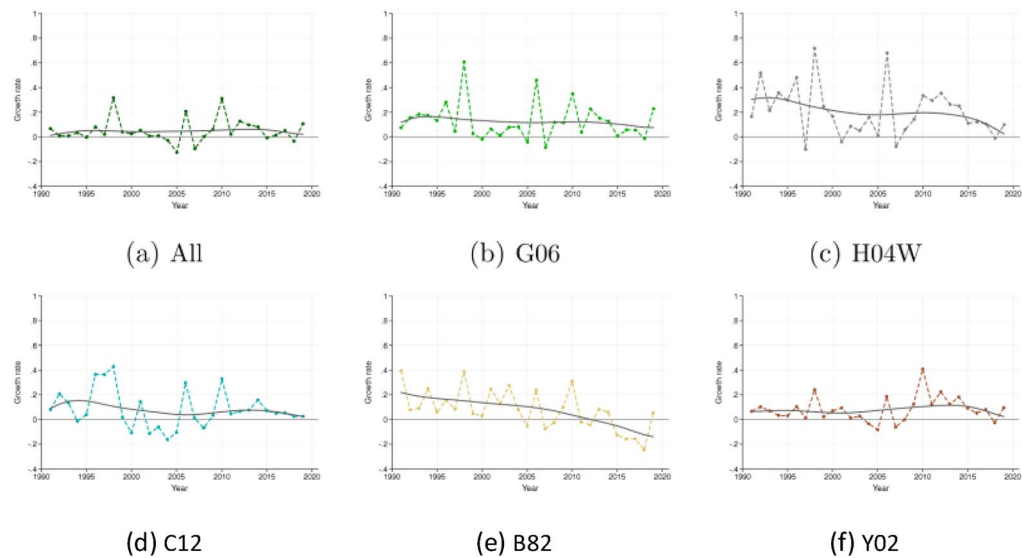


Figure D4. Growth rates of benchmark patents by year. *Note:* “All” refers to all patents, G06, H04W, B82, C12, and Y02 refers to computing, wireless communications, biochemistry/genetic engineering, nanotechnology, and climate invention-related patents, respectively.

D.1.2. Generality

Table D2. Average generality index (1990–2019): Benchmark categories

	All	G06	H04W	C12	B82	Y02
1 digit	0.82	0.62	0.62	0.74	0.79	0.85
3 digit	0.95	0.82	0.82	0.85	0.92	0.95
4 digit	0.82	0.62	0.62	0.74	0.79	0.85

The generality index is defined as share of citations to patents in different CPC classes at different aggregation levels (see [Appendix A2](#)). Citations within the same class are excluded. “All” refers to all patents, G06, H04W, B82, C12, and Y02 refers to computing, wireless communications, biochemistry/genetic engineering, nanotechnology, and climate invention-related patents, respectively.

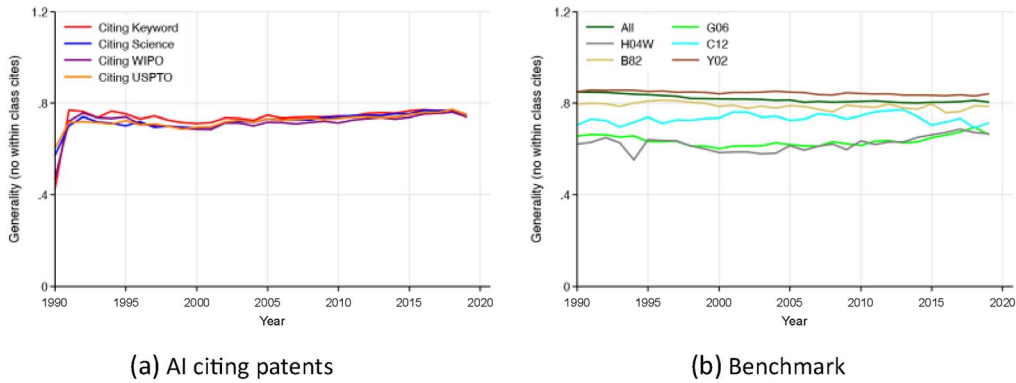


Figure D5. Generality index at the 1-digit CPC-section level.

Table D3. Average CPC classes making citations: Benchmark categories

	All	G06	H04W	C12	B82	Y02
1 digit	1.27	1.00	0.68	1.46	2.36	1.99
3 digit	2.48	1.99	1.19	2.52	4.39	3.32
4 digit	3.97	3.31	2.80	4.18	6.42	5.10

The table shows number of different CPC classes making a citation to an average patent of the respective group. Citations within the same class are excluded. “All” refers to all patents, G06, H04W, B82, C12, and Y02 refers to computing, wireless communications, biochemistry/genetic engineering, nanotechnology, and climate invention-related patents, respectively.

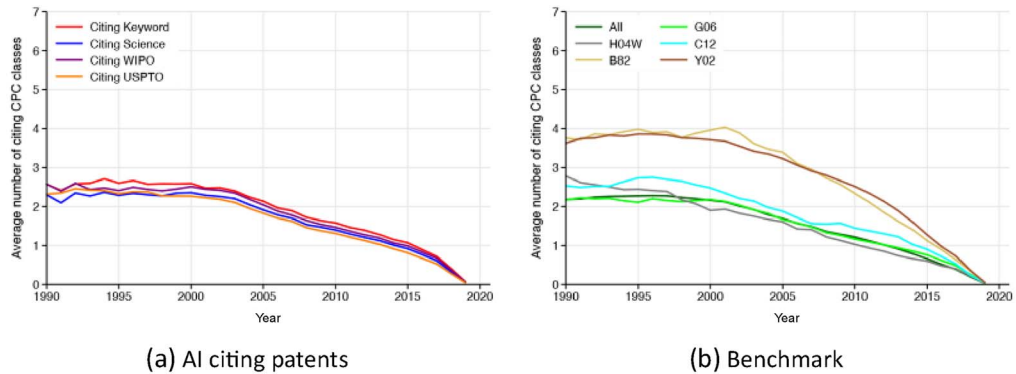
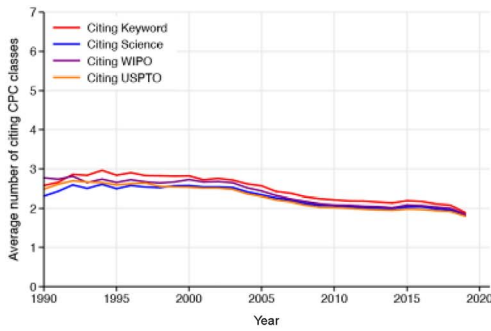


Figure D6. Average number of CPC classes citing AI. *Note:* “All” refers to all patents, G06, H04W, B82, C12, and Y02 refers to computing, wireless communications, biochemistry/genetic engineering, nanotechnology, and climate invention-related patents, respectively.

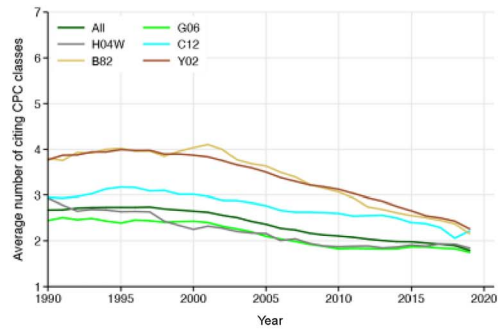
Table D4. Average number of citing CPC classes—cited patents

	All	G06	H04W	C12	B82	Y02
1 digit	2.39	2.03	2.00	2.78	3.35	3.29
3 digit	4.28	3.78	3.35	4.60	6.22	5.48
4 digit	6.44	5.96	5.63	7.47	9.08	8.42

The table reports numbers of different CPC classes making a citation to an average patent of the respective group that receives at least one citation. Citations within the same class are excluded. “All” refers to all patents, G06, H04W, B82, C12, and Y02 refers to computing, wireless communications, biochemistry/genetic engineering, nanotechnology, and climate invention-related patents, respectively.



(a) AI citing patents



(b) Benchmark

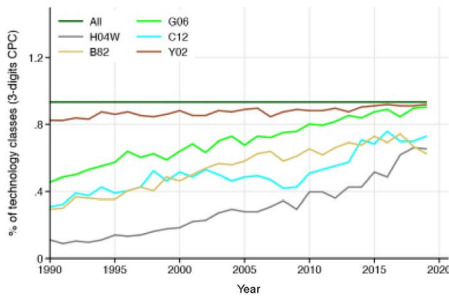
Figure D7. Average number of CPC classes citing AI: Subset of cited patents. *Note:* “All” refers to all patents, G06, H04W, B82, C12, and Y02 refers to computing, wireless communications, biochemistry/genetic engineering, nanotechnology, and climate invention-related patents, respectively.

Table D5. Average citation lags by patents in benchmark categories

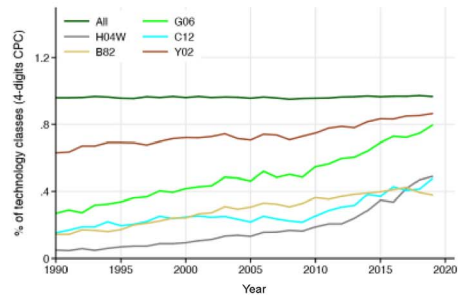
Period	All	G06	H04W	C12	B82	Y02
1990–1999	13.57	12.78	12.47	14.58	12.14	13.49
2000–2009	9.19	9.00	8.75	10.15	8.58	8.92
2010–2019	4.29	4.19	3.83	4.29	4.33	4.08

This table shows the average number of years it takes until a patent in the sample is cited. The average number of years is lower during the more recent decade as the maximal time lag is truncated since our data ends in 2019. “All” refers to all patents, G06, H04W, B82, C12, and Y02 refers to computing, wireless communications, biochemistry/genetic engineering, nanotechnology, and climate invention-related patents, respectively.

D.1.3. Complementarity



(a) % of all 3-digit CPC



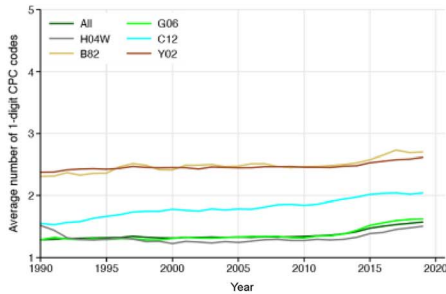
(b) % of all 4-digit CPC

Figure D8. Share of technology classes: Diversity of benchmark categories. *Note:* Panel (a) shows the percentage of 3-digit CPC codes and panel (b) shows the percentage of 4-digit CPC as a share of all codes in the respective category. Note that the total numbers of 3-digit and 4-digit CPC codes are 136 and 674, respectively. “All” refers to all patents, G06, H04W, B82, C12, and Y02 refers to computing, wireless communications, biochemistry/genetic engineering, nanotechnology, and climate invention-related patents, respectively.

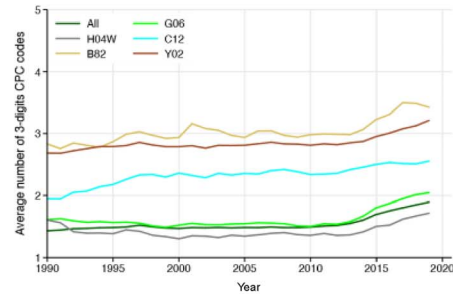
Table D6. Yearly average number of 3- and 4-digits CPC codes per patent

	All	G06	H04W	C12	B82	Y02
1 digit	1.36	1.36	1.32	1.80	2.48	2.47
3 digit	1.54	1.62	1.43	2.32	3.03	2.85
4 digit	1.80	1.81	2.26	2.93	3.50	3.39

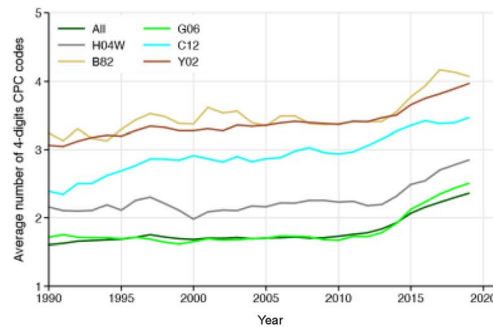
The table shows the average of annual average number of technology classes by 1-, 3-, or 4-digit CPC per patent. “All” refers to all patents, G06, H04W, B82, C12, and Y02 refers to computing, wireless communications, biochemistry/genetic engineering, nanotechnology, and climate invention-related patents, respectively.



(a) Average Number of 1-digit CPC



(b) Average Number of 3-digit CPC



(c) Average Number of 4-digit CPC

Figure D9. Average technology classes: Patent-level diversity of benchmark categories. *Note:* “All” refers to all patents, G06, H04W, B82, C12, and Y02 refers to computing, wireless communications, biochemistry/genetic engineering, nanotechnology, and climate-related patents, respectively.

D.2. Significance tests

Here, we provide the results of a series of pair-wise Wilcoxon signed rank tests showing whether the differences between the means reported in Tables 2, D40, D41, D39, D46 are significant.

D.2.1. Growth

Table D7. Growth rates

Period	Pair	Keyword	Science	WIPO	USPTO	All	G06	H04W	C12	B82
1990–2019	Science	1.00	–	–	–	–	–	–	–	–
1990–2019	WIPO	1.00	1.00	–	–	–	–	–	–	–
1990–2019	USPTO	1.00	1.00	1.00	–	–	–	–	–	–
1990–2019	All	0.02	0.00	0.00	0.00	–	–	–	–	–
1990–2019	G06	1.00	1.00	1.00	1.00	0.00	–	–	–	–
1990–2019	H04W	0.27	0.62	0.74	0.17	0.00	0.04	–	–	–
1990–2019	C12	1.00	0.08	0.40	0.76	1.00	0.75	0.01	–	–
1990–2019	B82	1.00	0.29	1.00	0.69	1.00	0.86	0.04	1.00	–
1990–2019	Y02	1.00	0.60	0.32	0.67	0.03	0.86	0.00	1.00	1.00
1990–1999	Science	0.18	–	–	–	–	–	–	–	–
1990–1999	WIPO	1.00	0.85	–	–	–	–	–	–	–
1990–1999	USPTO	0.62	1.00	1.00	–	–	–	–	–	–
1990–1999	All	1.00	0.45	1.00	0.18	–	–	–	–	–
1990–1999	G06	0.32	1.00	1.00	1.00	0.45	–	–	–	–
1990–1999	H04W	0.45	1.00	1.00	1.00	0.45	0.85	–	–	–
1990–1999	C12	1.00	1.00	1.00	1.00	1.00	1.00	1.00	–	–
1990–1999	B82	1.00	1.00	1.00	1.00	0.18	1.00	1.00	1.00	–
1990–1999	Y02	1.00	0.32	1.00	0.32	1.00	0.18	0.45	1.00	0.45
2000–2009	Science	1.00	–	–	–	–	–	–	–	–
2000–2009	WIPO	1.00	1.00	–	–	–	–	–	–	–
2000–2009	USPTO	1.00	1.00	1.00	–	–	–	–	–	–
2000–2009	All	1.00	0.43	1.00	0.26	–	–	–	–	–
2000–2009	G06	1.00	1.00	1.00	1.00	0.56	–	–	–	–
2000–2009	H04W	1.00	1.00	1.00	1.00	1.00	1.00	–	–	–
2000–2009	C12	1.00	0.43	1.00	0.78	1.00	1.00	1.00	–	–
2000–2009	B82	1.00	1.00	1.00	1.00	0.43	1.00	1.00	1.00	–
2000–2009	Y02	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
2010–2019	Science	0.55	–	–	–	–	–	–	–	–
2010–2019	WIPO	1.00	0.71	–	–	–	–	–	–	–
2010–2019	USPTO	1.00	1.00	0.85	–	–	–	–	–	–
2010–2019	All	0.09	0.71	0.09	1.00	–	–	–	–	–
2010–2019	G06	1.00	1.00	0.30	1.00	0.09	–	–	–	–
2010–2019	H04W	1.00	1.00	1.00	1.00	0.13	1.00	–	–	–
2010–2019	C12	0.55	1.00	0.30	1.00	1.00	1.00	0.19	–	–
2010–2019	B82	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.40	–
2010–2019	Y02	1.00	1.00	1.00	1.00	0.19	1.00	0.85	0.71	0.09

Entries show the *P*-value of a two-sided paired Wilcoxon signed rank test for the hypothesis that the compared pair ranks equal.

Table D8. Summary statistics for growth rates

	Keyword	Science	WIPO	USPTO	All	G06	H04W	C12	B82	Y02
Mean 1990–2019	0.12	0.15	0.14	0.13	0.05	0.13	0.21	0.08	0.08	0.08
Median 1990–2019	0.09	0.12	0.12	0.09	0.03	0.08	0.17	0.05	0.08	0.07
St. dev. 1990–2019	0.15	0.17	0.16	0.16	0.10	0.15	0.20	0.15	0.16	0.10
Mean 1990–1999	0.09	0.26	0.16	0.20	0.06	0.19	0.32	0.18	0.17	0.08
Median 1990–1999	0.03	0.27	0.09	0.12	0.03	0.16	0.30	0.13	0.09	0.07
St. dev. 1990–1999	0.14	0.17	0.17	0.19	0.10	0.18	0.23	0.17	0.14	0.07
Mean 2000–2009	0.05	0.10	0.09	0.09	0.01	0.08	0.12	–0.01	0.09	0.03
Median 2000–2009	0.03	0.06	0.07	0.09	0.01	0.07	0.07	–0.06	0.09	0.02
St. dev. 2000–2009	0.12	0.18	0.19	0.15	0.09	0.15	0.21	0.14	0.13	0.09
Mean 2010–2019	0.21	0.11	0.18	0.12	0.08	0.12	0.19	0.09	–0.02	0.14
Median 2010–2019	0.17	0.07	0.16	0.08	0.07	0.09	0.19	0.06	–0.03	0.11
St. dev. 2010–2019	0.16	0.13	0.09	0.15	0.10	0.12	0.12	0.09	0.16	0.12

Entries show the *P*-value of a two-sided paired Wilcoxon signed rank test for the hypothesis that the compared pair ranks equal.

Table D9. Growth rates (citing AI)

Period	Pair	Keyword	Science	WIPO	USPTO	G06	H04W	C12	B82
1990–2019	Science	1.00	–	–	–	–	–	–	–
1990–2019	WIPO	1.00	1.00	–	–	–	–	–	–
1990–2019	USPTO	0.02	0.00	0.01	–	–	–	–	–
1990–2019	G06	0.14	0.00	0.00	1.00	–	–	–	–
1990–2019	H04W	0.91	0.89	1.00	0.01	0.00	–	–	–
1990–2019	C12	0.89	0.62	0.96	1.00	1.00	0.10	–	–
1990–2019	B82	1.00	1.00	1.00	1.00	1.00	1.00	1.00	–
1990–2019	Y02	0.07	0.12	0.96	1.00	1.00	0.05	1.00	1.00
1990–1999	Science	1.00	–	–	–	–	–	–	–
1990–1999	WIPO	1.00	1.00	–	–	–	–	–	–
1990–1999	USPTO	1.00	0.14	0.26	–	–	–	–	–
1990–1999	G06	1.00	1.00	0.14	1.00	–	–	–	–
1990–1999	H04W	1.00	1.00	1.00	1.00	0.14	–	–	–
1990–1999	C12	1.00	1.00	1.00	1.00	1.00	0.88	–	–
1990–1999	B82	1.00	1.00	1.00	1.00	1.00	1.00	1.00	–
1990–1999	Y02	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
2000–2009	Science	1.00	–	–	–	–	–	–	–
2000–2009	WIPO	1.00	1.00	–	–	–	–	–	–
2000–2009	USPTO	0.68	0.68	1.00	–	–	–	–	–
2000–2009	G06	1.00	1.00	1.00	1.00	–	–	–	–
2000–2009	H04W	1.00	1.00	1.00	1.00	0.90	–	–	–
2000–2009	C12	1.00	1.00	1.00	1.00	1.00	1.00	–	–
2000–2009	B82	1.00	1.00	1.00	1.00	1.00	1.00	1.00	–
2000–2009	Y02	0.07	1.00	1.00	1.00	1.00	1.00	1.00	1.00
2010–2019	Science	1.00	–	–	–	–	–	–	–
2010–2019	WIPO	1.00	1.00	–	–	–	–	–	–
2010–2019	USPTO	0.33	0.62	1.00	–	–	–	–	–
2010–2019	G06	0.14	0.21	0.62	1.00	–	–	–	–
2010–2019	H04W	1.00	1.00	1.00	1.00	0.45	–	–	–
2010–2019	C12	1.00	1.00	1.00	1.00	1.00	1.00	–	–
2010–2019	B82	1.00	1.00	1.00	1.00	1.00	1.00	1.00	–
2010–2019	Y02	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00

Table excludes those patents that themselves are AI by the respective classification approach. Entries show the *P*-value of a two-sided paired Wilcoxon signed rank test for the hypothesis that the compared pair ranks equal.

Table D10. Summary statistics for growth rates (citing AI)

	Keyword	Science	WIPO	USPTO	G06	H04W	C12	B82	Y02
Mean 1990–2019	0.95	0.75	1.74	0.53	0.96	1.29	0.70	1.67	0.70
Median 1990–2019	0.14	0.13	0.12	0.09	0.11	0.19	0.08	0.12	0.10
St. dev. 1990–2019	3.80	2.56	8.01	1.71	4.08	5.34	2.58	7.77	2.65
Mean 1990–1999	2.81	2.15	5.38	1.52	2.91	3.89	2.07	5.14	2.07
Median 1990–1999	0.57	0.58	0.68	0.54	0.44	0.73	0.43	0.39	0.40
St. dev. 1990–1999	6.69	4.43	14.24	2.93	7.21	9.43	4.50	13.83	4.63
Mean 2000–2009	0.14	0.13	0.13	0.10	0.11	0.17	0.11	0.16	0.10
Median 2000–2009	0.12	0.10	0.12	0.09	0.09	0.12	0.04	0.13	0.09
St. dev. 2000–2009	0.17	0.16	0.16	0.13	0.13	0.19	0.22	0.17	0.15
Mean 2010–2019	0.10	0.09	0.09	0.06	0.05	0.09	0.07	0.07	0.07
Median 2010–2019	0.07	0.07	0.08	0.05	0.05	0.09	0.06	0.06	0.07
St. dev. 2010–2019	0.13	0.13	0.14	0.10	0.10	0.13	0.15	0.13	0.10

Table excludes those patents that themselves are AI by the respective classification approach.

D.2.2. Generality

Table D11. Generality index at 1-digit level

Period	1-digit	Keyword	Science	WIPO	USPTO	All	G06	H04W	C12	B82
1990–2019	Science	0.00	–	–	–	–	–	–	–	–
1990–2019	WIPO	0.00	0.00	–	–	–	–	–	–	–
1990–2019	USPTO	0.00	0.00	0.00	–	–	–	–	–	–
1990–2019	All	0.00	0.00	0.00	0.00	–	–	–	–	–
1990–2019	G06	0.00	0.00	0.00	0.00	0.00	–	–	–	–
1990–2019	H04W	0.00	0.00	0.00	0.00	0.00	0.00	–	–	–
1990–2019	C12	0.00	0.00	0.00	0.00	0.00	0.00	0.00	–	–
1990–2019	B82	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	–
1990–2019	Y02	0.66	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01
1990–1999	Science	0.09	–	–	–	–	–	–	–	–
1990–1999	WIPO	0.09	1.00	–	–	–	–	–	–	–
1990–1999	USPTO	0.09	0.09	0.09	–	–	–	–	–	–
1990–1999	All	0.09	0.09	0.09	0.09	–	–	–	–	–
1990–1999	G06	0.09	0.09	0.09	0.09	0.09	–	–	–	–
1990–1999	H04W	0.09	0.09	0.09	0.09	0.09	0.09	–	–	–
1990–1999	C12	0.09	1.00	1.00	0.09	0.09	0.09	0.09	–	–
1990–1999	B82	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09	–
1990–1999	Y02	0.11	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09
2000–2009	Science	0.09	–	–	–	–	–	–	–	–
2000–2009	WIPO	0.09	0.09	–	–	–	–	–	–	–
2000–2009	USPTO	0.09	0.09	0.09	–	–	–	–	–	–
2000–2009	All	0.09	0.09	0.09	0.09	–	–	–	–	–
2000–2009	G06	0.09	0.09	0.09	0.09	0.09	–	–	–	–
2000–2009	H04W	0.09	0.09	0.09	0.09	0.09	0.09	–	–	–
2000–2009	C12	0.09	0.09	0.09	0.09	0.09	0.09	0.09	–	–
2000–2009	B82	0.11	0.09	0.09	0.09	0.09	0.09	0.09	0.09	–
2000–2009	Y02	0.92	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.32
2010–2019	Science	0.09	–	–	–	–	–	–	–	–
2010–2019	WIPO	0.09	0.09	–	–	–	–	–	–	–
2010–2019	USPTO	0.09	0.09	0.09	–	–	–	–	–	–
2010–2019	All	0.09	0.09	0.09	0.09	–	–	–	–	–
2010–2019	G06	0.09	0.09	0.09	0.09	0.09	–	–	–	–
2010–2019	H04W	0.09	0.09	0.09	1.00	0.09	0.12	–	–	–
2010–2019	C12	0.67	0.09	0.09	0.09	0.67	0.09	0.09	–	–
2010–2019	B82	1.00	0.09	0.09	0.09	1.00	0.09	0.09	0.09	–
2010–2019	Y02	1.00	0.09	0.09	0.09	1.00	0.09	0.09	0.09	1.00

Entries show the *P*-value of a two-sided paired Wilcoxon signed rank test for the hypothesis that the compared pair ranks equal.

Table D12. Summary statistics for generality index at 1-digit level

Period	3-digit	Keyword	Science	WIPO	USPTO	All	G06	H04W	C12	B82
1990–2019	Science	0.00	–	–	–	–	–	–	–	–
1990–2019	WIPO	0.00	0.74	–	–	–	–	–	–	–
1990–2019	USPTO	0.00	0.00	0.00	–	–	–	–	–	–
1990–2019	All	0.00	0.00	0.00	0.00	–	–	–	–	–
1990–2019	G06	0.00	0.00	0.00	0.00	0.00	–	–	–	–
1990–2019	H04W	0.00	0.00	0.00	0.22	0.00	0.00	–	–	–
1990–2019	C12	0.00	0.00	0.00	0.05	0.00	0.00	0.01	–	–
1990–2019	B82	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	–
1990–2019	Y02	0.74	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
1990–1999	Science	0.09	–	–	–	–	–	–	–	–
1990–1999	WIPO	0.09	0.39	–	–	–	–	–	–	–
1990–1999	USPTO	0.09	0.29	0.09	–	–	–	–	–	–
1990–1999	All	0.09	0.09	0.09	0.09	–	–	–	–	–
1990–1999	G06	0.09	0.09	0.09	0.09	0.09	–	–	–	–
1990–1999	H04W	0.09	0.09	0.09	0.09	0.09	0.29	–	–	–
1990–1999	C12	0.09	0.09	0.09	0.26	0.09	0.29	0.39	–	–
1990–1999	B82	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09	–
1990–1999	Y02	0.39	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09
2000–2009	Science	0.09	–	–	–	–	–	–	–	–
2000–2009	WIPO	0.09	1.00	–	–	–	–	–	–	–
2000–2009	USPTO	0.09	0.09	0.09	–	–	–	–	–	–
2000–2009	All	0.09	0.09	0.09	0.09	–	–	–	–	–
2000–2009	G06	0.09	0.09	0.09	0.09	0.09	–	–	–	–
2000–2009	H04W	0.09	0.09	0.09	0.09	0.09	0.09	–	–	–
2000–2009	C12	0.09	0.09	0.09	0.09	0.09	0.09	0.09	–	–
2000–2009	B82	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09	–
2000–2009	Y02	1.00	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09
2010–2019	Science	0.09	–	–	–	–	–	–	–	–
2010–2019	WIPO	0.09	1.00	–	–	–	–	–	–	–
2010–2019	USPTO	0.09	0.09	0.09	–	–	–	–	–	–
2010–2019	All	0.09	0.09	0.09	0.09	–	–	–	–	–
2010–2019	G06	0.09	0.09	0.09	1.00	0.09	–	–	–	–
2010–2019	H04W	0.09	0.21	0.12	1.00	0.09	0.37	–	–	–
2010–2019	C12	0.09	0.09	0.09	0.09	0.09	0.09	1.00	–	–
2010–2019	B82	0.74	0.67	0.37	0.09	0.09	0.09	0.09	0.09	–
2010–2019	Y02	1.00	0.09	0.09	0.09	1.00	0.09	0.09	0.09	0.09

Entries show the *P*-value of a two-sided paired Wilcoxon signed rank test for the hypothesis that the compared pair ranks equal.

Table D13. Generality index at 3-digit level

	Keyword	Science	WIPO	USPTO	All	G06	H04W	C12	B82	Y02
Mean 1990–2019	0.81	0.77	0.76	0.73	0.83	0.7	0.68	0.78	0.82	0.81
Median 1990–2019	0.81	0.77	0.76	0.72	0.83	0.7	0.68	0.79	0.82	0.81
St. dev. 1990–2019	0.02	0.02	0.02	0.03	0.02	0.02	0.02	0.01	0.01	0.01
Mean 1990–1999	0.82	0.79	0.79	0.76	0.84	0.73	0.68	0.78	0.83	0.81
Median 1990–1999	0.82	0.79	0.79	0.76	0.84	0.73	0.68	0.78	0.83	0.81
St. dev. 1990–1999	0.01	0.02	0.01	0.02	0	0.01	0.01	0	0	0
Mean 2000–2009	0.81	0.77	0.76	0.72	0.83	0.7	0.66	0.79	0.82	0.81
Median 2000–2009	0.81	0.77	0.76	0.72	0.83	0.7	0.66	0.79	0.82	0.81
St. dev. 2000–2009	0	0.01	0.01	0.01	0	0.01	0.01	0.01	0	0.01
Mean 2010–2019	0.8	0.76	0.74	0.7	0.81	0.68	0.7	0.78	0.81	0.81
Median 2010–2019	0.81	0.76	0.74	0.7	0.82	0.68	0.7	0.78	0.81	0.81
St. dev. 2010–2019	0.03	0.01	0.01	0.01	0.03	0.01	0.02	0.02	0.02	0.01

Table D14. Summary statistics for generality index at 3-digit level

Period	4-digit	Keyword	Science	WIPO	USPTO	All	G06	H04W	C12	B82
1990–2019	Science	0.00	–	–	–	–	–	–	–	–
1990–2019	WIPO	0.00	0.05	–	–	–	–	–	–	–
1990–2019	USPTO	0.00	0.00	0.00	–	–	–	–	–	–
1990–2019	All	0.00	0.00	0.00	0.00	–	–	–	–	–
1990–2019	G06	0.00	0.00	0.00	0.00	0.00	–	–	–	–
1990–2019	H04W	0.00	0.00	0.00	0.00	0.00	0.00	–	–	–
1990–2019	C12	0.00	0.00	0.00	0.69	0.00	0.00	0.00	–	–
1990–2019	B82	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	–
1990–2019	Y02	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
1990–1999	Science	0.09	–	–	–	–	–	–	–	–
1990–1999	WIPO	0.09	1.00	–	–	–	–	–	–	–
1990–1999	USPTO	0.09	1.00	1.00	–	–	–	–	–	–
1990–1999	All	0.09	0.09	0.09	0.09	–	–	–	–	–
1990–1999	G06	0.09	0.09	0.09	0.09	0.09	–	–	–	–
1990–1999	H04W	0.09	0.09	0.09	0.09	0.09	0.09	–	–	–
1990–1999	C12	0.09	0.09	0.09	0.09	0.09	0.09	0.09	–	–
1990–1999	B82	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09	–
1990–1999	Y02	0.09	0.09	0.09	0.09	1.00	0.09	0.09	0.09	0.09
2000–2009	Science	0.09	–	–	–	–	–	–	–	–
2000–2009	WIPO	0.09	0.17	–	–	–	–	–	–	–
2000–2009	USPTO	0.09	0.09	0.09	–	–	–	–	–	–
2000–2009	All	0.09	0.09	0.09	0.09	–	–	–	–	–
2000–2009	G06	0.09	0.09	0.09	0.09	0.09	–	–	–	–
2000–2009	H04W	0.09	0.09	0.09	0.09	0.09	0.09	–	–	–
2000–2009	C12	0.09	0.09	0.17	0.09	0.09	0.09	0.09	–	–
2000–2009	B82	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09	–
2000–2009	Y02	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09
2010–2019	Science	0.63	–	–	–	–	–	–	–	–
2010–2019	WIPO	0.33	0.27	–	–	–	–	–	–	–
2010–2019	USPTO	0.09	0.09	0.09	–	–	–	–	–	–
2010–2019	All	0.09	0.09	0.09	0.09	–	–	–	–	–
2010–2019	G06	0.09	0.09	0.09	0.65	0.09	–	–	–	–
2010–2019	H04W	0.09	0.09	0.09	0.21	0.09	0.16	–	–	–
2010–2019	C12	0.09	0.09	0.09	1.00	0.09	1.00	0.65	–	–
2010–2019	B82	0.65	0.59	0.52	0.09	0.09	0.09	0.13	0.09	–
2010–2019	Y02	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09

Entries show the *P*-value of a two-sided paired Wilcoxon signed rank test for the hypothesis that the compared pair ranks equal.

Table D15. Generality index at 4-digit level

	Keyword	Science	WIPO	USPTO	All	G06	H04W	C12	B82	Y02
Mean 1990–2019	0.94	0.9	0.9	0.87	0.95	0.86	0.87	0.88	0.93	0.94
Median 1990–2019	0.94	0.9	0.9	0.87	0.95	0.86	0.86	0.88	0.93	0.94
St. dev. 1990–2019	0.01	0.01	0.01	0.02	0.01	0.01	0.02	0.01	0.02	0
Mean 1990–1999	0.94	0.91	0.91	0.89	0.96	0.87	0.88	0.88	0.93	0.94
Median 1990–1999	0.94	0.91	0.92	0.9	0.96	0.88	0.88	0.88	0.93	0.94
St. dev. 1990–1999	0	0.01	0.01	0.02	0	0.01	0.01	0.01	0	0
Mean 2000–2009	0.94	0.9	0.89	0.86	0.95	0.85	0.85	0.88	0.93	0.94
Median 2000–2009	0.94	0.9	0.89	0.86	0.95	0.85	0.85	0.88	0.93	0.94
St. dev. 2000–2009	0	0.01	0.01	0.01	0	0	0.01	0	0	0
Mean 2010–2019	0.93	0.89	0.9	0.86	0.94	0.86	0.87	0.88	0.92	0.94
Median 2010–2019	0.94	0.89	0.9	0.86	0.95	0.86	0.88	0.88	0.93	0.94
St. dev. 2010–2019	0.02	0	0.01	0.01	0.02	0	0.02	0.01	0.03	0

Entries show the *P*-value of a two-sided paired Wilcoxon signed rank test for the hypothesis that the compared pair ranks equal.

Table D16. Summary statistics for generality index at 4-digit level

Period	1-digit	Keyword	Science	WIPO	USPTO	All	G06	H04W	C12	B82
1990–2019	Science	0.00	–	–	–	–	–	–	–	–
1990–2019	WIPO	0.00	1.00	–	–	–	–	–	–	–
1990–2019	USPTO	0.00	0.00	0.00	–	–	–	–	–	–
1990–2019	All	0.00	0.00	0.00	0.00	–	–	–	–	–
1990–2019	G06	0.00	0.00	0.00	0.00	0.74	–	–	–	–
1990–2019	H04W	0.00	0.00	0.00	0.00	1.00	1.00	–	–	–
1990–2019	C12	0.00	0.00	0.00	0.00	1.00	1.00	1.00	–	–
1990–2019	B82	0.74	0.00	0.00	0.00	0.00	0.00	0.00	0.00	–
1990–2019	Y02	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
1990–1999	Science	0.09	–	–	–	–	–	–	–	–
1990–1999	WIPO	0.09	0.56	–	–	–	–	–	–	–
1990–1999	USPTO	0.09	0.09	0.09	–	–	–	–	–	–
1990–1999	All	0.09	0.09	0.09	0.09	–	–	–	–	–
1990–1999	G06	0.09	0.09	0.09	0.09	0.09	–	–	–	–
1990–1999	H04W	0.09	0.09	0.09	0.09	0.39	0.15	–	–	–
1990–1999	C12	0.09	0.09	0.09	0.09	0.09	0.09	0.39	–	–
1990–1999	B82	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09	–
1990–1999	Y02	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09
2000–2009	Science	0.09	–	–	–	–	–	–	–	–
2000–2009	WIPO	0.09	0.10	–	–	–	–	–	–	–
2000–2009	USPTO	0.09	0.09	0.09	–	–	–	–	–	–
2000–2009	All	0.09	0.09	0.09	0.09	–	–	–	–	–
2000–2009	G06	0.09	0.09	0.09	0.09	0.09	–	–	–	–
2000–2009	H04W	0.09	0.09	0.09	0.09	0.09	0.09	–	–	–
2000–2009	C12	0.09	0.09	0.09	0.09	0.09	0.09	0.13	–	–
2000–2009	B82	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09	–
2000–2009	Y02	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09
2010–2019	Science	0.19	–	–	–	–	–	–	–	–
2010–2019	WIPO	0.09	0.09	–	–	–	–	–	–	–
2010–2019	USPTO	0.09	0.09	0.09	–	–	–	–	–	–
2010–2019	All	0.09	0.09	0.09	0.09	–	–	–	–	–
2010–2019	G06	0.09	0.09	0.09	0.09	0.48	–	–	–	–
2010–2019	H04W	0.09	0.09	0.09	0.09	1.00	0.09	–	–	–
2010–2019	C12	0.09	0.09	0.09	0.19	0.84	1.00	1.00	–	–
2010–2019	B82	0.09	1.00	1.00	0.58	0.19	0.48	0.41	0.12	–
2010–2019	Y02	0.09	1.00	1.00	0.48	0.09	0.41	0.19	0.09	0.48

Entries show the *P*-value of a two-sided paired Wilcoxon signed rank test for the hypothesis that the compared pair ranks equal.

Table D17. Average number of citing classes (all) at 1-digit level

	Keyword	Science	WIPO	USPTO	All	G06	H04W	C12	B82	Y02
Mean 1990–2019	0.97	0.96	0.95	0.95	0.98	0.94	0.91	0.95	0.97	0.98
Median 1990–2019	0.97	0.96	0.95	0.94	0.98	0.94	0.9	0.95	0.97	0.99
St. dev. 1990–2019	0.01	0	0.01	0.01	0.01	0.01	0.02	0.01	0.01	0
Mean 1990–1999	0.97	0.96	0.96	0.96	0.99	0.94	0.91	0.95	0.97	0.99
Median 1990–1999	0.98	0.96	0.96	0.96	0.99	0.94	0.91	0.95	0.97	0.99
St. dev. 1990–1999	0	0.01	0.01	0.01	0	0.01	0.01	0	0	0
Mean 2000–2009	0.97	0.95	0.95	0.94	0.98	0.93	0.9	0.95	0.97	0.98
Median 2000–2009	0.97	0.95	0.95	0.94	0.98	0.93	0.9	0.95	0.97	0.98
St. dev. 2000–2009	0	0	0	0	0	0	0	0	0	0
Mean 2010–2019	0.97	0.96	0.95	0.94	0.98	0.94	0.92	0.94	0.96	0.98
Median 2010–2019	0.97	0.96	0.96	0.94	0.98	0.94	0.92	0.94	0.97	0.98
St. dev. 2010–2019	0.02	0	0.01	0	0.01	0	0.02	0.01	0.02	0

Entries show the *P*-value of a two-sided paired Wilcoxon signed rank test for the hypothesis that the compared pair ranks equal.

Table D18. Summary statistics for average number of citing classes (all) at 1-digit level

Period	3-digit	Keyword	Science	WIPO	USPTO	All	G06	H04W	C12	B82
1990–2019	Science	0.00	–	–	–	–	–	–	–	–
1990–2019	WIPO	0.00	0.00	–	–	–	–	–	–	–
1990–2019	USPTO	0.00	0.01	0.00	–	–	–	–	–	–
1990–2019	All	0.00	0.00	0.00	0.00	–	–	–	–	–
1990–2019	G06	0.00	0.00	0.00	0.00	0.00	–	–	–	–
1990–2019	H04W	0.00	0.00	0.00	0.00	0.65	0.91	–	–	–
1990–2019	C12	0.00	0.00	0.00	0.00	1.00	0.00	0.00	–	–
1990–2019	B82	0.99	0.00	0.00	0.00	0.00	0.00	0.00	0.00	–
1990–2019	Y02	0.00	0.02	1.00	0.00	0.00	0.00	0.00	0.00	0.00
1990–1999	Science	0.09	–	–	–	–	–	–	–	–
1990–1999	WIPO	0.09	0.96	–	–	–	–	–	–	–
1990–1999	USPTO	0.09	1.00	0.25	–	–	–	–	–	–
1990–1999	All	0.09	0.09	0.09	0.09	–	–	–	–	–
1990–1999	G06	0.09	0.09	0.09	0.09	0.09	–	–	–	–
1990–1999	H04W	0.09	0.09	0.09	0.09	0.09	0.59	–	–	–
1990–1999	C12	0.09	0.09	0.09	0.09	0.25	0.09	0.09	–	–
1990–1999	B82	1.00	0.09	0.09	0.09	0.09	0.09	0.09	0.09	–
1990–1999	Y02	0.09	1.00	1.00	1.00	0.09	0.09	0.10	0.09	0.09
2000–2009	Science	0.09	–	–	–	–	–	–	–	–
2000–2009	WIPO	0.09	0.09	–	–	–	–	–	–	–
2000–2009	USPTO	0.09	0.25	0.09	–	–	–	–	–	–
2000–2009	All	0.09	0.09	0.09	0.09	–	–	–	–	–
2000–2009	G06	0.09	0.09	0.09	0.09	0.09	–	–	–	–
2000–2009	H04W	0.09	0.09	0.09	0.09	0.64	0.09	–	–	–
2000–2009	C12	0.09	0.09	0.09	0.09	0.09	0.09	0.09	–	–
2000–2009	B82	0.15	0.09	0.09	0.09	0.09	0.09	0.09	0.09	–
2000–2009	Y02	0.09	0.09	0.64	0.09	0.09	0.09	0.09	0.09	0.09
2010–2019	Science	0.17	–	–	–	–	–	–	–	–
2010–2019	WIPO	0.10	1.00	–	–	–	–	–	–	–
2010–2019	USPTO	0.09	0.09	0.09	–	–	–	–	–	–
2010–2019	All	0.09	0.09	0.09	0.09	–	–	–	–	–
2010–2019	G06	0.09	0.09	0.09	0.10	0.17	–	–	–	–
2010–2019	H04W	0.10	0.09	0.09	0.10	1.00	0.10	–	–	–
2010–2019	C12	0.09	0.09	0.09	0.10	1.00	0.10	1.00	–	–
2010–2019	B82	0.09	1.00	1.00	0.58	0.11	0.49	0.38	0.10	–
2010–2019	Y02	0.09	1.00	1.00	0.38	0.09	0.38	0.17	0.09	0.41

Entries show the *P*-value of a two-sided paired Wilcoxon signed rank test for the hypothesis that the compared pair ranks equal.

Table D19. Average number of citing classes (all) at 3-digit level

	Keyword	Science	WIPO	USPTO	All	G06	H04W	C12	B82	Y02
Mean 1990–2019	2.98	2.53	2.57	2.38	2.14	2.17	2.14	2.15	3.08	2.77
Median 1990–2019	3.38	2.85	3.03	2.72	2.45	2.51	2.43	2.26	3.79	3.3
St. dev. 1990–2019	1.24	1.1	1.15	1.17	1.06	1.03	1.06	1.08	1.5	1.24
Mean 1990–1999	4.04	3.58	3.59	3.5	3.16	3.11	3.22	3.28	4.31	3.81
Median 1990–1999	4.03	3.57	3.61	3.53	3.17	3.11	3.26	3.29	4.33	3.81
St. dev. 1990–1999	0.12	0.09	0.08	0.08	0.04	0.04	0.15	0.1	0.09	0.07
Mean 2000–2009	3.41	2.82	2.94	2.68	2.43	2.49	2.36	2.28	3.73	3.24
Median 2000–2009	3.38	2.85	3.03	2.72	2.45	2.51	2.43	2.26	3.79	3.3
St. dev. 2000–2009	0.52	0.42	0.55	0.54	0.44	0.46	0.36	0.43	0.6	0.38
Mean 2010–2019	1.48	1.2	1.17	0.94	0.84	0.9	0.85	0.88	1.2	1.25
Median 2010–2019	1.61	1.31	1.27	0.99	0.85	0.97	0.88	0.95	1.18	1.27
St. dev. 2010–2019	0.86	0.65	0.64	0.57	0.58	0.55	0.54	0.58	0.89	0.89

Entries show the *P*-value of a two-sided paired Wilcoxon signed rank test for the hypothesis that the compared pair ranks equal.

Table D20. Summary statistics for average number of citing classes (all) at 3-digit level

Period	4-digit	Keyword	Science	WIPO	USPTO	All	G06	H04W	C12	B82
1990–2019	Science	0.00	–	–	–	–	–	–	–	–
1990–2019	WIPO	0.00	0.00	–	–	–	–	–	–	–
1990–2019	USPTO	0.00	0.10	0.00	–	–	–	–	–	–
1990–2019	All	0.00	0.00	0.00	0.00	–	–	–	–	–
1990–2019	G06	0.00	0.00	0.00	0.00	0.00	–	–	–	–
1990–2019	H04W	0.00	0.88	0.07	0.88	0.00	0.00	–	–	–
1990–2019	C12	0.00	0.00	0.00	0.00	0.02	0.00	0.00	–	–
1990–2019	B82	0.21	0.05	0.42	0.00	0.00	0.00	0.03	0.00	–
1990–2019	Y02	0.00	0.88	0.01	0.44	0.00	0.00	0.88	0.00	0.02
1990–1999	Science	0.17	–	–	–	–	–	–	–	–
1990–1999	WIPO	0.09	1.00	–	–	–	–	–	–	–
1990–1999	USPTO	0.09	1.00	1.00	–	–	–	–	–	–
1990–1999	All	0.09	0.09	0.09	0.09	–	–	–	–	–
1990–1999	G06	0.09	0.09	0.09	0.09	0.09	–	–	–	–
1990–1999	H04W	1.00	0.17	1.00	0.45	0.09	0.09	–	–	–
1990–1999	C12	0.09	0.09	0.09	0.09	0.09	1.00	0.09	–	–
1990–1999	B82	1.00	0.71	1.00	0.38	0.09	0.09	1.00	0.09	–
1990–1999	Y02	0.09	1.00	0.09	0.38	0.09	0.09	0.29	0.09	0.09
2000–2009	Science	0.09	–	–	–	–	–	–	–	–
2000–2009	WIPO	0.09	0.09	–	–	–	–	–	–	–
2000–2009	USPTO	0.09	0.64	0.09	–	–	–	–	–	–
2000–2009	All	0.09	0.09	0.09	0.09	–	–	–	–	–
2000–2009	G06	0.09	0.09	0.09	0.09	0.09	–	–	–	–
2000–2009	H04W	0.09	0.09	0.09	0.53	0.09	0.09	–	–	–
2000–2009	C12	0.09	0.09	0.09	0.09	0.64	0.09	0.09	–	–
2000–2009	B82	0.92	0.09	0.09	0.09	0.09	0.09	0.09	0.09	–
2000–2009	Y02	0.09	0.09	0.64	0.09	0.09	0.09	0.09	0.09	0.09
2010–2019	Science	0.16	–	–	–	–	–	–	–	–
2010–2019	WIPO	0.09	0.19	–	–	–	–	–	–	–
2010–2019	USPTO	0.09	0.09	0.09	–	–	–	–	–	–
2010–2019	All	0.09	0.09	0.09	0.09	–	–	–	–	–
2010–2019	G06	0.09	0.09	0.09	0.09	0.09	–	–	–	–
2010–2019	H04W	0.09	0.09	0.09	1.00	0.09	0.30	–	–	–
2010–2019	C12	0.09	0.09	0.09	0.09	0.11	0.74	0.09	–	–
2010–2019	B82	0.09	0.19	0.11	1.00	0.16	1.00	1.00	0.19	–
2010–2019	Y02	0.09	1.00	1.00	0.33	0.09	0.33	0.30	0.09	0.09

Entries show the *P*-value of a two-sided paired Wilcoxon signed rank test for the hypothesis that the compared pair ranks equal.

Table D21. Average number of citing classes (all) at 4-digit level

	Keyword	Science	WIPO	USPTO	All	G06	H04W	C12	B82	Y02
Mean 1990–2019	6	4.82	5.07	4.57	3.59	4	3.92	3.56	5.88	5.12
Median 1990–2019	6.73	5.31	5.92	4.95	4	4.41	4.11	3.64	7.29	6.06
St. dev. 1990–2019	2.9	2.47	2.62	2.67	1.95	2.22	2.4	2	3.23	2.59
Mean 1990–1999	8.65	7.37	7.56	7.36	5.54	6.24	6.65	5.75	8.54	7.49
Median 1990–1999	8.58	7.27	7.54	7.44	5.57	6.27	6.78	5.74	8.64	7.57
St. dev. 1990–1999	0.4	0.44	0.26	0.24	0.1	0.14	0.68	0.29	0.39	0.23
Mean 2000–2009	6.88	5.22	5.78	4.97	4.01	4.44	3.96	3.69	7.29	5.94
Median 2000–2009	6.73	5.31	5.92	4.95	4	4.41	4.11	3.64	7.29	6.06
St. dev. 2000–2009	1.4	1.05	1.47	1.42	0.91	1.19	0.9	0.86	1.74	0.99
Mean 2010–2019	2.46	1.86	1.88	1.38	1.21	1.32	1.16	1.24	1.82	1.91
Median 2010–2019	2.6	1.96	2.01	1.41	1.19	1.39	1.18	1.3	1.69	1.83
St. dev. 2010–2019	1.54	1.07	1.1	0.88	0.87	0.83	0.76	0.85	1.44	1.44

Entries show the *P*-value of a two-sided paired Wilcoxon signed rank test for the hypothesis that the compared pair ranks equal.

Table D22. Summary statistics for average number of citing classes (all) at 4-digit level

Period	1-digit	Keyword	Science	WIPO	USPTO	All	G06	H04W	C12	B82
1990–2019	Science	0.00	–	–	–	–	–	–	–	–
1990–2019	WIPO	0.00	0.26	–	–	–	–	–	–	–
1990–2019	USPTO	0.00	0.00	0.00	–	–	–	–	–	–
1990–2019	All	0.00	0.00	0.00	0.42	–	–	–	–	–
1990–2019	G06	0.00	0.00	0.00	0.00	0.00	–	–	–	–
1990–2019	H04W	0.00	0.00	0.00	0.00	0.00	0.00	–	–	–
1990–2019	C12	0.00	0.42	0.32	0.01	0.00	0.00	0.00	–	–
1990–2019	B82	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	–
1990–2019	Y02	0.42	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
1990–1999	Science	0.09	–	–	–	–	–	–	–	–
1990–1999	WIPO	0.09	1.00	–	–	–	–	–	–	–
1990–1999	USPTO	0.09	0.34	0.29	–	–	–	–	–	–
1990–1999	All	0.09	0.09	0.09	0.09	–	–	–	–	–
1990–1999	G06	0.09	0.09	0.09	0.09	0.09	–	–	–	–
1990–1999	H04W	0.09	0.09	0.09	0.09	0.09	1.00	–	–	–
1990–1999	C12	0.09	0.26	0.32	0.97	0.09	0.09	0.09	–	–
1990–1999	B82	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09	–
1990–1999	Y02	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09
2000–2009	Science	0.09	–	–	–	–	–	–	–	–
2000–2009	WIPO	0.09	1.00	–	–	–	–	–	–	–
2000–2009	USPTO	0.09	0.09	0.09	–	–	–	–	–	–
2000–2009	All	0.09	0.09	0.09	1.00	–	–	–	–	–
2000–2009	G06	0.09	0.09	0.09	0.09	0.09	–	–	–	–
2000–2009	H04W	0.09	0.09	0.09	0.09	0.09	0.09	–	–	–
2000–2009	C12	0.09	1.00	1.00	0.65	0.09	0.09	0.09	–	–
2000–2009	B82	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09	–
2000–2009	Y02	0.39	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09
2010–2019	Science	0.09	–	–	–	–	–	–	–	–
2010–2019	WIPO	0.09	0.09	–	–	–	–	–	–	–
2010–2019	USPTO	0.09	0.09	0.09	–	–	–	–	–	–
2010–2019	All	0.09	0.09	0.55	0.09	–	–	–	–	–
2010–2019	G06	0.09	0.09	0.09	0.53	0.09	–	–	–	–
2010–2019	H04W	0.09	0.09	0.22	0.53	0.26	0.53	–	–	–
2010–2019	C12	0.56	0.09	0.09	0.09	0.09	0.09	0.09	–	–
2010–2019	B82	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09	–
2010–2019	Y02	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.26

Entries show the *P*-value of a two-sided paired Wilcoxon signed rank test for the hypothesis that the compared pair ranks equal.

Table D23. Average number of citing classes (cited) at 1-digit level

	Keyword	Science	WIPO	USPTO	All	G06	H04W	C12	B82	Y02
Mean 1990–2019	9.26	7.82	8.42	7.5	5.26	6.15	7.62	5.72	8.81	7.87
Median 1990–2019	10.62	8.63	9.9	8.04	5.84	6.74	7.93	5.66	10.77	9.37
St. dev. 1990–2019	4.55	4.13	4.47	4.55	2.93	3.51	4.85	3.38	5.16	4.12
Mean 1990–1999	13.32	12.11	12.56	12.27	8.2	9.67	13.16	9.52	13.2	11.65
Median 1990–1999	13.44	11.9	12.6	12.44	8.24	9.67	13.47	9.58	13.29	11.8
St. dev. 1990–1999	0.57	0.68	0.41	0.31	0.2	0.16	1.33	0.59	0.89	0.5
Mean 2000–2009	10.78	8.48	9.72	8.15	5.87	6.85	7.64	5.78	10.84	9.21
Median 2000–2009	10.62	8.63	9.9	8.04	5.84	6.74	7.93	5.66	10.77	9.37
St. dev. 2000–2009	2.28	1.85	2.69	2.56	1.43	1.99	1.91	1.48	3.07	1.71
Mean 2010–2019	3.68	2.88	2.98	2.09	1.7	1.93	2.06	1.86	2.39	2.76
Median 2010–2019	3.81	3	3.17	2.12	1.67	2.04	2.07	1.92	2.18	2.56
St. dev. 2010–2019	2.4	1.7	1.78	1.37	1.24	1.23	1.39	1.3	1.93	2.14

Entries show the *P*-value of a two-sided paired Wilcoxon signed rank test for the hypothesis that the compared pair ranks equal.

Table D24. Summary statistics for average number of citing classes (cited) at 1-digit level

Period	3-digit	Keyword	Science	WIPO	USPTO	All	G06	H04W	C12	B82
1990–2019	Science	0.00	–	–	–	–	–	–	–	–
1990–2019	WIPO	0.00	0.08	–	–	–	–	–	–	–
1990–2019	USPTO	0.00	0.00	0.00	–	–	–	–	–	–
1990–2019	All	0.00	0.00	0.00	0.01	–	–	–	–	–
1990–2019	G06	0.00	0.00	0.00	0.00	0.08	–	–	–	–
1990–2019	H04W	0.00	0.00	0.00	0.00	0.54	0.08	–	–	–
1990–2019	C12	0.00	0.00	0.00	0.08	0.00	0.31	0.05	–	–
1990–2019	B82	0.54	0.00	0.00	0.00	0.00	0.00	0.00	0.00	–
1990–2019	Y02	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
1990–1999	Science	0.09	–	–	–	–	–	–	–	–
1990–1999	WIPO	0.09	1.00	–	–	–	–	–	–	–
1990–1999	USPTO	0.09	1.00	1.00	–	–	–	–	–	–
1990–1999	All	0.09	0.09	0.09	0.09	–	–	–	–	–
1990–1999	G06	0.09	0.09	0.09	0.09	0.09	–	–	–	–
1990–1999	H04W	0.09	0.09	0.09	0.09	0.15	1.00	–	–	–
1990–1999	C12	0.09	0.09	0.09	0.09	0.09	0.84	0.95	–	–
1990–1999	B82	1.00	0.09	0.09	0.09	0.09	0.09	0.09	0.09	–
1990–1999	Y02	0.09	1.00	1.00	1.00	0.09	0.09	0.09	0.09	0.09
2000–2009	Science	0.09	–	–	–	–	–	–	–	–
2000–2009	WIPO	0.09	0.32	–	–	–	–	–	–	–
2000–2009	USPTO	0.09	0.10	0.09	–	–	–	–	–	–
2000–2009	All	0.09	0.09	0.09	0.09	–	–	–	–	–
2000–2009	G06	0.09	0.09	0.09	0.09	0.34	–	–	–	–
2000–2009	H04W	0.09	0.09	0.09	0.09	0.09	0.09	–	–	–
2000–2009	C12	0.09	0.09	0.09	0.34	0.09	0.77	0.09	–	–
2000–2009	B82	0.10	0.09	0.09	0.09	0.09	0.09	0.09	0.09	–
2000–2009	Y02	0.09	0.09	0.34	0.09	0.09	0.09	0.09	0.09	0.09
2010–2019	Science	0.15	–	–	–	–	–	–	–	–
2010–2019	WIPO	0.11	1.00	–	–	–	–	–	–	–
2010–2019	USPTO	0.09	0.09	0.09	–	–	–	–	–	–
2010–2019	All	0.09	0.09	0.09	0.09	–	–	–	–	–
2010–2019	G06	0.09	0.09	0.09	1.00	0.15	–	–	–	–
2010–2019	H04W	0.15	0.11	0.15	0.79	0.52	0.74	–	–	–
2010–2019	C12	0.15	0.18	0.15	0.09	0.09	0.09	0.11	–	–
2010–2019	B82	1.00	0.09	0.09	0.09	0.09	0.09	0.09	0.09	–
2010–2019	Y02	1.00	0.09	0.09	0.09	0.09	0.09	0.09	0.09	1.00

Entries show the *P*-value of a two-sided paired Wilcoxon signed rank test for the hypothesis that the compared pair ranks equal.

Table D25. Average number of citing classes (cited) at 3-digit level

	Keyword	Science	WIPO	USPTO	All	G06	H04W	C12	B82	Y02
Mean 1990–2019	3.44	2.99	2.96	2.83	2.81	2.67	2.53	3.05	3.74	3.4
Median 1990–2019	3.67	3.17	3.19	2.96	2.92	2.79	2.54	3.07	4.05	3.57
St. dev. 1990–2019	0.78	0.67	0.71	0.74	0.6	0.58	0.6	0.5	0.76	0.56
Mean 1990–1999	4.14	3.66	3.66	3.62	3.42	3.25	3.24	3.59	4.38	3.95
Median 1990–1999	4.12	3.66	3.65	3.63	3.44	3.24	3.28	3.6	4.4	3.95
St. dev. 1990–1999	0.1	0.07	0.06	0.07	0.03	0.04	0.14	0.08	0.09	0.07
Mean 2000–2009	3.68	3.14	3.12	2.94	2.91	2.78	2.49	3.09	4.04	3.55
Median 2000–2009	3.67	3.17	3.19	2.96	2.92	2.79	2.54	3.07	4.05	3.57
St. dev. 2000–2009	0.41	0.35	0.45	0.42	0.3	0.35	0.27	0.2	0.39	0.26
Mean 2010–2019	2.5	2.16	2.11	1.95	2.08	1.97	1.86	2.48	2.78	2.71
Median 2010–2019	2.61	2.16	2.12	1.94	2.06	1.97	1.83	2.52	2.73	2.68
St. dev. 2010–2019	0.43	0.22	0.19	0.18	0.22	0.15	0.12	0.25	0.33	0.25

Entries show the *P*-value of a two-sided paired Wilcoxon signed rank test for the hypothesis that the compared pair ranks equal.

Table D26. Summary statistics for average number of citing classes (cited) at 3-digit level

Period	4-digit	Keyword	Science	WIPO	USPTO	All	G06	H04W	C12	B82
1990–2019	Science	0.00	–	–	–	–	–	–	–	–
1990–2019	WIPO	0.00	0.00	–	–	–	–	–	–	–
1990–2019	USPTO	0.00	0.04	0.00	–	–	–	–	–	–
1990–2019	All	0.00	0.00	0.00	0.00	–	–	–	–	–
1990–2019	G06	0.00	0.00	0.00	0.00	0.00	–	–	–	–
1990–2019	H04W	0.00	0.37	0.06	0.90	0.00	0.00	–	–	–
1990–2019	C12	0.00	0.00	0.00	0.07	0.00	0.01	0.19	–	–
1990–2019	B82	0.35	0.00	0.00	0.00	0.00	0.00	0.00	0.00	–
1990–2019	Y02	0.00	0.04	0.83	0.04	0.00	0.00	0.16	0.00	0.03
1990–1999	Science	0.11	–	–	–	–	–	–	–	–
1990–1999	WIPO	0.09	1.00	–	–	–	–	–	–	–
1990–1999	USPTO	0.09	1.00	1.00	–	–	–	–	–	–
1990–1999	All	0.09	0.09	0.09	0.09	–	–	–	–	–
1990–1999	G06	0.09	0.09	0.09	0.09	0.09	–	–	–	–
1990–1999	H04W	1.00	0.59	1.00	1.00	0.09	0.09	–	–	–
1990–1999	C12	0.09	0.09	0.09	0.09	0.09	1.00	0.09	–	–
1990–1999	B82	1.00	1.00	1.00	0.90	0.09	0.09	1.00	0.09	–
1990–1999	Y02	0.09	1.00	0.11	0.59	0.09	0.09	0.90	0.09	0.09
2000–2009	Science	0.09	–	–	–	–	–	–	–	–
2000–2009	WIPO	0.09	0.09	–	–	–	–	–	–	–
2000–2009	USPTO	0.09	0.22	0.09	–	–	–	–	–	–
2000–2009	All	0.09	0.09	0.09	0.09	–	–	–	–	–
2000–2009	G06	0.09	0.09	0.09	0.09	0.10	–	–	–	–
2000–2009	H04W	0.09	0.09	0.09	0.09	0.09	0.09	–	–	–
2000–2009	C12	0.09	0.09	0.09	0.22	0.09	1.00	1.00	–	–
2000–2009	B82	1.00	0.09	0.09	0.09	0.09	0.09	0.09	0.09	–
2000–2009	Y02	0.09	0.09	1.00	0.09	0.09	0.09	0.09	0.09	0.09
2010–2019	Science	0.30	–	–	–	–	–	–	–	–
2010–2019	WIPO	0.46	0.30	–	–	–	–	–	–	–
2010–2019	USPTO	0.09	0.09	0.09	–	–	–	–	–	–
2010–2019	All	0.10	0.09	0.09	1.00	–	–	–	–	–
2010–2019	G06	0.10	0.09	0.09	1.00	1.00	–	–	–	–
2010–2019	H04W	0.39	0.54	0.46	1.00	1.00	0.46	–	–	–
2010–2019	C12	0.22	1.00	0.46	0.09	0.09	0.09	1.00	–	–
2010–2019	B82	0.39	0.39	1.00	0.09	0.09	0.09	0.46	1.00	–
2010–2019	Y02	1.00	0.09	0.46	0.09	0.09	0.09	0.09	0.09	0.09

Entries show the *P*-value of a two-sided paired Wilcoxon signed rank test for the hypothesis that the compared pair ranks equal.

Table D27. Average number of citing classes (cited) at 4-digit level

	Keyword	Science	WIPO	USPTO	All	G06	H04W	C12	B82	Y02
Mean 1990–2019	6.76	5.54	5.71	5.27	4.57	4.76	4.47	4.9	6.85	6.1
Median 1990–2019	7.3	5.91	6.23	5.37	4.78	4.89	4.29	4.96	7.79	6.56
St. dev. 1990–2019	2.25	1.89	2	2.14	1.37	1.65	1.83	1.28	2.27	1.67
Mean 1990–1999	8.86	7.55	7.69	7.6	6.01	6.51	6.69	6.29	8.67	7.76
Median 1990–1999	8.82	7.43	7.67	7.65	6.04	6.52	6.81	6.24	8.76	7.82
St. dev. 1990–1999	0.36	0.39	0.25	0.23	0.09	0.14	0.67	0.26	0.4	0.23
Mean 2000–2009	7.4	5.78	6.12	5.42	4.79	4.94	4.18	4.98	7.88	6.49
Median 2000–2009	7.3	5.91	6.23	5.37	4.78	4.89	4.29	4.96	7.79	6.56
St. dev. 2000–2009	1.22	0.96	1.3	1.25	0.73	1.03	0.79	0.55	1.4	0.8
Mean 2010–2019	4.03	3.28	3.32	2.79	2.93	2.82	2.55	3.42	4.02	4.06
Median 2010–2019	4.22	3.24	3.35	2.76	2.87	2.84	2.49	3.44	3.89	3.98
St. dev. 2010–2019	1.04	0.49	0.49	0.43	0.43	0.34	0.26	0.53	0.85	0.62

Entries show the *P*-value of a two-sided paired Wilcoxon signed rank test for the hypothesis that the compared pair ranks equal.

Table D28. Summary statistics for average number of citing classes (cited) at 4-digit level

	Keyword	Science	WIPO	USPTO	All	G06	H04W	C12	B82	Y02
Mean 1990–2019	10.4	8.94	9.43	8.56	6.66	7.26	8.56	7.75	10.12	9.29
Median 1990–2019	11.53	9.61	10.42	8.72	6.96	7.49	8.29	7.7	11.51	10.14
St. dev. 1990–2019	3.62	3.28	3.53	3.79	2.15	2.71	3.89	2.38	3.93	2.87
Mean 1990–1999	13.64	12.4	12.79	12.66	8.89	10.1	13.25	10.42	13.4	12.07
Median 1990–1999	13.76	12.2	12.82	12.78	8.94	10.09	13.52	10.41	13.53	12.26
St. dev. 1990–1999	0.54	0.61	0.42	0.28	0.21	0.17	1.31	0.55	0.91	0.51
Mean 2000–2009	11.6	9.4	10.29	8.86	7	7.6	8.05	7.77	11.68	10.05
Median 2000–2009	11.53	9.61	10.42	8.72	6.96	7.49	8.29	7.7	11.51	10.14
St. dev. 2000–2009	2	1.71	2.43	2.31	1.19	1.77	1.7	1.04	2.61	1.43
Mean 2010–2019	5.97	5.03	5.21	4.17	4.1	4.09	4.4	5.05	5.28	5.75
Median 2010–2019	6.19	4.97	5.29	4.17	4.03	4.16	4.29	5.08	5.02	5.52
St. dev. 2010–2019	1.69	0.88	0.9	0.76	0.65	0.56	0.45	0.91	1.13	1.06

Entries show the *P*-value of a two-sided paired Wilcoxon signed rank test for the hypothesis that the compared pair ranks equal.

D.2.3. Complementarity

Table D29. Share of CPC classes at 3-digit level

Period	3-digit	Keyword	Science	WIPO	USPTO	All	G06	H04W	C12	B82
1990–2019	Science	0.29	–	–	–	–	–	–	–	–
1990–2019	WIPO	0.15	0.29	–	–	–	–	–	–	–
1990–2019	USPTO	0.00	0.00	0.00	–	–	–	–	–	–
1990–2019	All	0.00	0.00	0.00	0.00	–	–	–	–	–
1990–2019	G06	0.00	0.00	0.00	0.00	0.00	–	–	–	–
1990–2019	H04W	0.00	0.00	0.00	0.00	0.00	0.00	–	–	–
1990–2019	C12	0.00	0.00	0.00	0.00	0.00	0.00	0.00	–	–
1990–2019	B82	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.29	–
1990–2019	Y02	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
1990–1999	Science	0.32	–	–	–	–	–	–	–	–
1990–1999	WIPO	0.32	0.14	–	–	–	–	–	–	–
1990–1999	USPTO	0.09	0.09	0.13	–	–	–	–	–	–
1990–1999	All	0.13	0.13	0.13	0.09	–	–	–	–	–
1990–1999	G06	0.10	0.13	0.13	0.13	0.09	–	–	–	–
1990–1999	H04W	0.13	0.09	0.09	0.09	0.13	0.09	–	–	–
1990–1999	C12	0.10	0.32	0.09	0.09	0.13	0.09	0.13	–	–
1990–1999	B82	0.13	0.13	0.09	0.09	0.13	0.09	0.09	0.17	–
1990–1999	Y02	0.09	0.09	0.13	0.13	0.13	0.09	0.09	0.13	0.09
2000–2009	Science	1.00	–	–	–	–	–	–	–	–
2000–2009	WIPO	1.00	0.77	–	–	–	–	–	–	–
2000–2009	USPTO	0.18	0.09	0.18	–	–	–	–	–	–
2000–2009	All	0.18	0.18	0.18	0.18	–	–	–	–	–
2000–2009	G06	0.18	0.18	0.18	0.18	0.18	–	–	–	–
2000–2009	H04W	0.18	0.09	0.09	0.18	0.18	0.09	–	–	–
2000–2009	C12	0.09	0.09	0.12	0.09	0.18	0.09	0.09	–	–
2000–2009	B82	0.63	0.63	1.00	0.09	0.18	0.18	0.18	0.18	–
2000–2009	Y02	0.18	0.09	0.09	0.18	0.18	0.18	0.18	0.09	0.18
2010–2019	Science	0.88	–	–	–	–	–	–	–	–
2010–2019	WIPO	0.09	0.13	–	–	–	–	–	–	–
2010–2019	USPTO	0.13	0.09	0.09	–	–	–	–	–	–
2010–2019	All	0.13	0.09	0.09	0.13	–	–	–	–	–
2010–2019	G06	0.13	0.13	0.13	0.88	0.09	–	–	–	–
2010–2019	H04W	0.09	0.09	0.09	0.09	0.13	0.09	–	–	–
2010–2019	C12	0.09	0.09	0.13	0.09	0.13	0.13	0.09	–	–
2010–2019	B82	0.09	0.09	0.88	0.09	0.13	0.09	0.13	0.88	–
2010–2019	Y02	0.09	0.09	0.09	0.13	0.13	0.13	0.13	0.13	0.13

Entries show the *P*-value of a two-sided paired Wilcoxon signed rank test for the hypothesis that the compared pair ranks equal.

Table D30. Summary statistics for share of CPC classes at 3-digit level

	Keyword	Science	WIPO	USPTO	All	G06	H04W	C12	B82	Y02
Mean 1990–2019	0.61	0.6	0.59	0.82	0.93	0.7	0.3	0.51	0.54	0.87
Median 1990–2019	0.6	0.59	0.57	0.83	0.93	0.71	0.28	0.49	0.57	0.88
St. dev. 1990–2019	0.15	0.16	0.12	0.06	0	0.13	0.17	0.12	0.14	0.03
Mean 1990–1999	0.47	0.43	0.48	0.76	0.93	0.56	0.13	0.4	0.37	0.85
Median 1990–1999	0.45	0.43	0.51	0.75	0.93	0.56	0.12	0.4	0.36	0.85
St. dev. 1990–1999	0.08	0.1	0.07	0.05	0	0.06	0.03	0.06	0.06	0.02
Mean 2000–2009	0.59	0.59	0.58	0.82	0.93	0.7	0.27	0.48	0.57	0.87
Median 2000–2009	0.6	0.58	0.57	0.83	0.93	0.71	0.28	0.49	0.57	0.88
St. dev. 2000–2009	0.04	0.03	0.04	0.02	0	0.04	0.05	0.04	0.05	0.02
Mean 2010–2019	0.78	0.77	0.71	0.87	0.93	0.85	0.49	0.64	0.68	0.9
Median 2010–2019	0.79	0.77	0.69	0.86	0.93	0.85	0.46	0.69	0.67	0.91
St. dev. 2010–2019	0.08	0.06	0.08	0.03	0	0.04	0.11	0.09	0.04	0.02

Entries show the *P*-value of a two-sided paired Wilcoxon signed rank test for the hypothesis that the compared pair ranks equal.

Table D31. Share of CPC classes at 4-digit level

Period	4-digit	Keyword	Science	WIPO	USPTO	All	G06	H04W	C12	B82
1990–2019	Science	0.83	–	–	–	–	–	–	–	–
1990–2019	WIPO	0.83	0.33	–	–	–	–	–	–	–
1990–2019	USPTO	0.00	0.00	0.00	–	–	–	–	–	–
1990–2019	All	0.00	0.00	0.00	0.00	–	–	–	–	–
1990–2019	G06	0.00	0.00	0.00	0.00	0.00	–	–	–	–
1990–2019	H04W	0.00	0.00	0.00	0.00	0.00	0.00	–	–	–
1990–2019	C12	0.00	0.00	0.00	0.00	0.00	0.00	0.00	–	–
1990–2019	B82	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.31	–
1990–2019	Y02	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
1990–1999	Science	0.77	–	–	–	–	–	–	–	–
1990–1999	WIPO	0.48	0.42	–	–	–	–	–	–	–
1990–1999	USPTO	0.09	0.09	0.09	–	–	–	–	–	–
1990–1999	All	0.09	0.09	0.09	0.09	–	–	–	–	–
1990–1999	G06	0.09	0.09	0.09	0.09	0.09	–	–	–	–
1990–1999	H04W	0.09	0.09	0.09	0.09	0.09	0.09	–	–	–
1990–1999	C12	0.09	0.48	0.09	0.09	0.09	0.09	0.09	–	–
1990–1999	B82	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09	–
1990–1999	Y02	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09
2000–2009	Science	0.09	–	–	–	–	–	–	–	–
2000–2009	WIPO	0.92	0.09	–	–	–	–	–	–	–
2000–2009	USPTO	0.09	0.09	0.09	–	–	–	–	–	–
2000–2009	All	0.09	0.09	0.09	0.09	–	–	–	–	–
2000–2009	G06	0.09	0.09	0.09	0.09	0.09	–	–	–	–
2000–2009	H04W	0.09	0.09	0.09	0.09	0.09	0.09	–	–	–
2000–2009	C12	0.09	0.09	0.09	0.09	0.09	0.09	0.09	–	–
2000–2009	B82	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09	–
2000–2009	Y02	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09
2010–2019	Science	0.48	–	–	–	–	–	–	–	–
2010–2019	WIPO	0.09	0.22	–	–	–	–	–	–	–
2010–2019	USPTO	0.09	0.09	0.09	–	–	–	–	–	–
2010–2019	All	0.09	0.09	0.09	0.09	–	–	–	–	–
2010–2019	G06	0.09	0.09	0.09	0.22	0.09	–	–	–	–
2010–2019	H04W	0.09	0.09	0.09	0.09	0.09	0.09	–	–	–
2010–2019	C12	0.09	0.09	0.09	0.09	0.09	0.09	0.22	–	–
2010–2019	B82	0.09	0.09	0.09	0.09	0.09	0.09	0.25	0.46	–
2010–2019	Y02	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09

Entries show the *P*-value of a two-sided paired Wilcoxon signed rank test for the hypothesis that the compared pair ranks equal.

Table D32. Summary statistics for the share of CPC classes at 4-digit level

	Keyword	Science	WIPO	USPTO	All	G06	H04W	C12	B82	Y02
Mean 1990–2019	0.37	0.37	0.36	0.6	0.96	0.49	0.17	0.27	0.29	0.74
Median 1990–2019	0.33	0.35	0.33	0.6	0.96	0.48	0.14	0.25	0.31	0.72
St. dev. 1990–2019	0.14	0.13	0.12	0.1	0.01	0.15	0.13	0.08	0.09	0.06
Mean 1990–1999	0.24	0.24	0.25	0.48	0.96	0.33	0.06	0.2	0.18	0.68
Median 1990–1999	0.24	0.23	0.27	0.48	0.96	0.33	0.06	0.2	0.17	0.68
St. dev. 1990–1999	0.06	0.07	0.05	0.05	0	0.05	0.02	0.03	0.03	0.03
Mean 2000–2009	0.33	0.36	0.33	0.6	0.96	0.47	0.14	0.24	0.3	0.72
Median 2000–2009	0.33	0.35	0.33	0.6	0.96	0.48	0.14	0.24	0.31	0.72
St. dev. 2000–2009	0.02	0.02	0.02	0.02	0	0.03	0.03	0.01	0.03	0.01
Mean 2010–2019	0.54	0.53	0.5	0.71	0.97	0.66	0.32	0.36	0.39	0.81
Median 2010–2019	0.56	0.53	0.5	0.7	0.97	0.67	0.31	0.38	0.39	0.82
St. dev. 2010–2019	0.09	0.07	0.09	0.05	0.01	0.08	0.11	0.07	0.02	0.04

Entries show the *P*-value of a two-sided paired Wilcoxon signed rank test for the hypothesis that the compared pair ranks equal.

Table D33. Average diversity at 1-digit level

Period	1-digit	Keyword	Science	WIPO	USPTO	All	G06	H04W	C12	B82
1990–2019	Science	1.00	–	–	–	–	–	–	–	–
1990–2019	WIPO	0.02	0.05	–	–	–	–	–	–	–
1990–2019	USPTO	0.00	0.00	0.00	–	–	–	–	–	–
1990–2019	All	0.00	0.02	1.00	0.00	–	–	–	–	–
1990–2019	G06	0.00	0.06	1.00	0.00	1.00	–	–	–	–
1990–2019	H04W	0.00	0.01	0.03	0.00	0.00	0.00	–	–	–
1990–2019	C12	0.00	0.00	0.00	0.00	0.00	0.00	0.00	–	–
1990–2019	B82	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	–
1990–2019	Y02	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.82
1990–1999	Science	1.00	–	–	–	–	–	–	–	–
1990–1999	WIPO	0.09	1.00	–	–	–	–	–	–	–
1990–1999	USPTO	0.09	0.84	0.15	–	–	–	–	–	–
1990–1999	All	0.25	1.00	0.09	0.09	–	–	–	–	–
1990–1999	G06	0.09	1.00	0.15	0.09	0.45	–	–	–	–
1990–1999	H04W	1.00	1.00	0.84	0.09	1.00	1.00	–	–	–
1990–1999	C12	0.09	0.09	0.09	0.09	0.09	0.09	0.09	–	–
1990–1999	B82	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09	–
1990–1999	Y02	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.71
2000–2009	Science	0.09	–	–	–	–	–	–	–	–
2000–2009	WIPO	0.29	0.09	–	–	–	–	–	–	–
2000–2009	USPTO	0.09	0.09	0.09	–	–	–	–	–	–
2000–2009	All	0.09	0.09	0.83	0.09	–	–	–	–	–
2000–2009	G06	0.09	0.09	0.83	0.09	0.83	–	–	–	–
2000–2009	H04W	0.09	0.09	0.09	0.29	0.09	0.09	–	–	–
2000–2009	C12	0.09	0.09	0.09	0.09	0.09	0.09	0.09	–	–
2000–2009	B82	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09	–
2000–2009	Y02	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.29
2010–2019	Science	1.00	–	–	–	–	–	–	–	–
2010–2019	WIPO	1.00	1.00	–	–	–	–	–	–	–
2010–2019	USPTO	0.09	0.09	0.09	–	–	–	–	–	–
2010–2019	All	0.58	0.11	0.84	0.09	–	–	–	–	–
2010–2019	G06	1.00	1.00	1.00	0.09	0.84	–	–	–	–
2010–2019	H04W	0.09	0.09	0.09	0.27	0.09	0.09	–	–	–
2010–2019	C12	0.09	0.09	0.09	0.09	0.09	0.09	0.09	–	–
2010–2019	B82	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09	–
2010–2019	Y02	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09

Entries show the *P*-value of a two-sided paired Wilcoxon signed rank test for the hypothesis that the compared pair ranks equal.

Table D34. Summary statistics for average diversity at 1-digit level

	Keyword	Science	WIPO	USPTO	All	G06	H04W	C12	B82	Y02
Mean 1990–2019	1.39	1.4	1.36	1.27	1.36	1.36	1.32	1.8	2.48	2.47
Median 1990–2019	1.36	1.42	1.33	1.24	1.33	1.32	1.29	1.78	2.48	2.46
St. dev. 1990–2019	0.09	0.1	0.12	0.08	0.08	0.11	0.08	0.15	0.11	0.05
Mean 1990–1999	1.35	1.3	1.27	1.24	1.31	1.29	1.33	1.64	2.39	2.43
Median 1990–1999	1.35	1.28	1.27	1.24	1.32	1.29	1.29	1.65	2.37	2.43
St. dev. 1990–1999	0.03	0.08	0.02	0.01	0.02	0.01	0.08	0.08	0.08	0.03
Mean 2000–2009	1.35	1.43	1.33	1.24	1.33	1.32	1.26	1.79	2.48	2.46
Median 2000–2009	1.36	1.43	1.33	1.24	1.33	1.32	1.26	1.78	2.48	2.45
St. dev. 2000–2009	0.02	0.02	0.03	0.01	0.01	0.01	0.02	0.04	0.03	0.01
Mean 2010–2019	1.47	1.47	1.48	1.35	1.45	1.47	1.37	1.97	2.58	2.52
Median 2010–2019	1.48	1.47	1.47	1.35	1.44	1.48	1.36	1.99	2.55	2.51
St. dev. 2010–2019	0.11	0.08	0.14	0.11	0.09	0.12	0.09	0.08	0.1	0.06

Entries show the *P*-value of a two-sided paired Wilcoxon signed rank test for the hypothesis that the compared pair ranks equal.

Table D35. Average diversity at 3-digit level

Period	3-digit	Keyword	Science	WIPO	USPTO	All	G06	H04W	C12	B82
1990–2019	Science	0.65	–	–	–	–	–	–	–	–
1990–2019	WIPO	0.38	0.87	–	–	–	–	–	–	–
1990–2019	USPTO	0.00	0.00	0.00	–	–	–	–	–	–
1990–2019	All	0.00	0.00	0.00	0.00	–	–	–	–	–
1990–2019	G06	0.79	0.79	0.46	0.00	0.00	–	–	–	–
1990–2019	H04W	0.00	0.00	0.00	0.79	0.00	0.00	–	–	–
1990–2019	C12	0.00	0.00	0.00	0.00	0.00	0.00	0.00	–	–
1990–2019	B82	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	–
1990–2019	Y02	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
1990–1999	Science	1.00	–	–	–	–	–	–	–	–
1990–1999	WIPO	0.09	1.00	–	–	–	–	–	–	–
1990–1999	USPTO	0.09	0.09	0.09	–	–	–	–	–	–
1990–1999	All	0.09	1.00	0.76	0.09	–	–	–	–	–
1990–1999	G06	1.00	1.00	0.09	0.09	0.09	–	–	–	–
1990–1999	H04W	0.09	1.00	0.49	1.00	1.00	0.09	–	–	–
1990–1999	C12	0.09	0.09	0.09	0.09	0.09	0.09	0.09	–	–
1990–1999	B82	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09	–
1990–1999	Y02	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09
2000–2009	Science	0.09	–	–	–	–	–	–	–	–
2000–2009	WIPO	0.09	0.09	–	–	–	–	–	–	–
2000–2009	USPTO	0.09	0.09	0.09	–	–	–	–	–	–
2000–2009	All	0.09	0.09	0.09	0.09	–	–	–	–	–
2000–2009	G06	0.92	0.09	0.09	0.09	0.09	–	–	–	–
2000–2009	H04W	0.09	0.09	0.09	0.86	0.09	0.09	–	–	–
2000–2009	C12	0.09	0.09	0.09	0.09	0.09	0.09	0.09	–	–
2000–2009	B82	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09	–
2000–2009	Y02	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09
2010–2019	Science	1.00	–	–	–	–	–	–	–	–
2010–2019	WIPO	0.09	0.10	–	–	–	–	–	–	–
2010–2019	USPTO	0.09	0.09	0.09	–	–	–	–	–	–
2010–2019	All	0.09	0.09	0.09	0.09	–	–	–	–	–
2010–2019	G06	0.32	1.00	0.09	0.09	0.09	–	–	–	–
2010–2019	H04W	0.09	0.09	0.09	0.10	0.09	0.09	–	–	–
2010–2019	C12	0.09	0.09	0.09	0.09	0.09	0.09	0.09	–	–
2010–2019	B82	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09	–
2010–2019	Y02	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09

Entries show the *P*-value of a two-sided paired Wilcoxon signed rank test for the hypothesis that the compared pair ranks equal.

Table D36. Summary statistics for average diversity at 3-digit level

	Keyword	Science	WIPO	USPTO	All	G06	H04W	C12	B82	Y02
Mean 1990–2019	1.61	1.64	1.64	1.43	1.54	1.62	1.43	2.32	3.03	2.85
Median 1990–2019	1.55	1.65	1.58	1.39	1.48	1.56	1.39	2.35	2.99	2.81
St. dev. 1990–2019	0.15	0.14	0.21	0.12	0.12	0.15	0.11	0.16	0.19	0.12
Mean 1990–1999	1.55	1.51	1.5	1.39	1.48	1.57	1.43	2.16	2.88	2.77
Median 1990–1999	1.56	1.48	1.48	1.4	1.48	1.57	1.4	2.16	2.86	2.79
St. dev. 1990–1999	0.04	0.1	0.04	0.02	0.03	0.04	0.09	0.15	0.09	0.06
Mean 2000–2009	1.54	1.69	1.58	1.36	1.48	1.54	1.35	2.36	3.01	2.81
Median 2000–2009	1.54	1.68	1.59	1.36	1.48	1.54	1.35	2.36	3.01	2.81
St. dev. 2000–2009	0.02	0.05	0.05	0.02	0.01	0.02	0.03	0.04	0.07	0.03
Mean 2010–2019	1.74	1.73	1.85	1.53	1.67	1.75	1.49	2.46	3.2	2.96
Median 2010–2019	1.73	1.71	1.82	1.52	1.65	1.74	1.46	2.48	3.15	2.91
St. dev. 2010–2019	0.2	0.14	0.26	0.17	0.15	0.21	0.14	0.08	0.22	0.14

Entries show the *P*-value of a two-sided paired Wilcoxon signed rank test for the hypothesis that the compared pair ranks equal.

Table D37. Average diversity at 4-digit level

Period	4-digit	Keyword	Science	WIPO	USPTO	All	G06	H04W	C12	B82
1990–2019	Science	0.01	–	–	–	–	–	–	–	–
1990–2019	WIPO	0.00	0.10	–	–	–	–	–	–	–
1990–2019	USPTO	0.00	0.00	0.00	–	–	–	–	–	–
1990–2019	All	0.00	0.00	0.00	0.00	–	–	–	–	–
1990–2019	G06	0.00	0.00	0.00	0.00	0.52	–	–	–	–
1990–2019	H04W	0.00	0.00	0.00	0.00	0.00	0.00	–	–	–
1990–2019	C12	0.00	0.00	0.00	0.00	0.00	0.00	0.00	–	–
1990–2019	B82	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	–
1990–2019	Y02	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
1990–1999	Science	1.00	–	–	–	–	–	–	–	–
1990–1999	WIPO	1.00	1.00	–	–	–	–	–	–	–
1990–1999	USPTO	0.09	0.09	0.09	–	–	–	–	–	–
1990–1999	All	0.09	1.00	0.09	0.09	–	–	–	–	–
1990–1999	G06	0.09	1.00	0.09	0.09	1.00	–	–	–	–
1990–1999	H04W	0.09	0.09	0.09	0.09	0.09	0.09	–	–	–
1990–1999	C12	0.09	0.09	0.09	0.09	0.09	0.09	0.09	–	–
1990–1999	B82	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09	–
1990–1999	Y02	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.10
2000–2009	Science	0.09	–	–	–	–	–	–	–	–
2000–2009	WIPO	0.09	0.26	–	–	–	–	–	–	–
2000–2009	USPTO	0.09	0.09	0.09	–	–	–	–	–	–
2000–2009	All	0.09	0.09	0.09	0.09	–	–	–	–	–
2000–2009	G06	0.09	0.09	0.09	0.09	0.28	–	–	–	–
2000–2009	H04W	0.09	0.09	0.09	0.09	0.09	0.09	–	–	–
2000–2009	C12	0.09	0.09	0.09	0.09	0.09	0.09	0.09	–	–
2000–2009	B82	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09	–
2000–2009	Y02	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09
2010–2019	Science	0.25	–	–	–	–	–	–	–	–
2010–2019	WIPO	0.09	0.09	–	–	–	–	–	–	–
2010–2019	USPTO	0.09	0.09	0.09	–	–	–	–	–	–
2010–2019	All	0.09	0.09	0.09	0.09	–	–	–	–	–
2010–2019	G06	0.09	0.09	0.09	0.09	0.64	–	–	–	–
2010–2019	H04W	0.09	0.09	0.70	0.09	0.09	0.09	–	–	–
2010–2019	C12	0.09	0.09	0.09	0.09	0.09	0.09	0.09	–	–
2010–2019	B82	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09	–
2010–2019	Y02	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.15

Entries show the *P*-value of a two-sided paired Wilcoxon signed rank test for the hypothesis that the compared pair ranks equal.

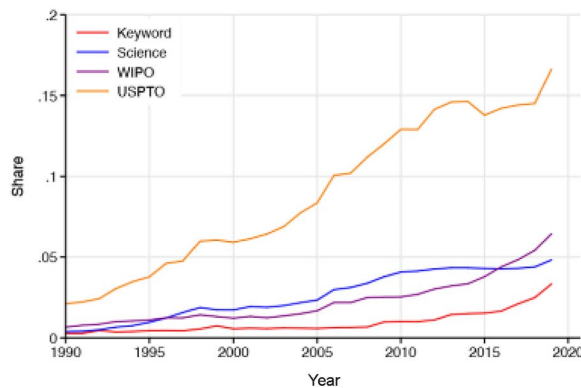
Table D38. Summary statistics for average diversity at 4-digit level

	Keyword	Science	WIPO	USPTO	All	G06	H04W	C12	B82	Y02
Mean 1990–2019	1.88	1.97	2.05	1.64	1.8	1.81	2.26	2.93	3.5	3.39
Median 1990–2019	1.77	1.97	1.96	1.56	1.71	1.71	2.2	2.89	3.41	3.36
St. dev. 1990–2019	0.26	0.25	0.38	0.21	0.21	0.25	0.21	0.3	0.27	0.23
Mean 1990–1999	1.77	1.73	1.76	1.55	1.68	1.69	2.16	2.64	3.31	3.2
Median 1990–1999	1.77	1.67	1.75	1.56	1.68	1.71	2.14	2.65	3.3	3.21
St. dev. 1990–1999	0.05	0.13	0.04	0.03	0.04	0.04	0.07	0.2	0.14	0.1
Mean 2000–2009	1.75	2	1.95	1.53	1.7	1.69	2.16	2.9	3.46	3.35
Median 2000–2009	1.75	1.99	1.98	1.54	1.7	1.69	2.17	2.89	3.44	3.36
St. dev. 2000–2009	0.03	0.06	0.09	0.02	0.01	0.03	0.09	0.07	0.1	0.05
Mean 2010–2019	2.12	2.19	2.43	1.83	2.01	2.05	2.45	3.24	3.72	3.63
Median 2010–2019	2.09	2.16	2.39	1.82	2	2.02	2.4	3.31	3.66	3.58
St. dev. 2010–2019	0.35	0.26	0.44	0.28	0.24	0.32	0.26	0.2	0.33	0.22

Entries show the *P*-value of a two-sided paired Wilcoxon signed rank test for the hypothesis that the compared pair ranks equal.

D.3. Additional results

D.3.1. Volume and time trends



(a) Share of AI patents

Figure D10. AI patents by year (1990–2019). *Note:* This figure shows the evolution of AI patents over time as identified by the four different approaches, as a share of all US patents granted in the same year.

D.3.2. Generality

Table D39. Average number of citing CPCs (1990–2019): Cited patents

	Keyword	Science	WIPO	USPTO
1 digit	2.71	2.42	2.41	2.26
3 digit	5.99	5.00	5.23	4.74
4 digit	9.92	8.54	9.05	8.21

This table shows the numbers of different CPC classes making a citation to an average patent of the respective group, conditional on the patent being cited at least once. Citations within the same class are excluded.

Table D40. Average generality index (1990–2019)

	Keyword	Science	WIPO	USPTO
1 digit	0.76	0.73	0.72	0.68
3 digit	0.91	0.87	0.87	0.84
4 digit	0.96	0.95	0.94	0.93

The generality index is defined as share of citations to patents in different CPC classes at different aggregation levels (see [Appendix A2](#)). Citations within the same class are excluded.

Table D41. Average number of citing CPCs (1990–2019)

	Keyword	Science	WIPO	USPTO
1 digit	2.15	1.83	1.82	1.68
3 digit	5.18	4.14	4.39	3.91
4 digit	8.90	7.41	8.04	7.13

This table shows the numbers of different CPC classes making a citation to an average patent of the respective group. Citations within the same class are excluded.

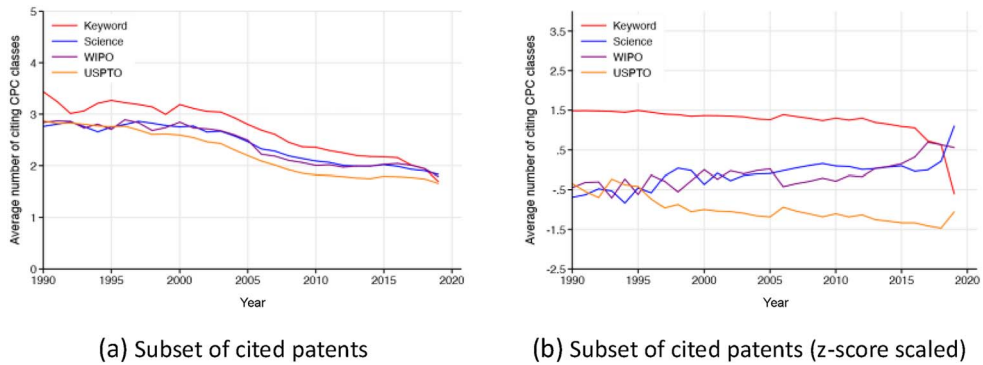


Figure D11. Average number of classes citing AI. *Note:* The z-scored value equals the level of the generality index minus its average across the four approaches divided by the standard deviation for each year.

D.3.3. Generality of AI descendants

In this section, we report additional results for the wide-ranging usefulness of technological descendants of AI, i.e. those patents that cite an AI patent but are not AI themselves. This serves as an additional indicator of the widespread of AI in a range of different products and processes. The results confirm the persistence of the ranking, indicating the highest generality of keyword patents across different indicators.

Table D42. Average generality index: AI descendants

	Citing keyword	Citing science	Citing WIPO	Citing USPTO
1 digit	0.74	0.73	0.72	0.72
3 digit	0.89	0.87	0.87	0.88
4 digit	0.96	0.95	0.95	0.95

Generality is measured as $G = 1 - \sum s^2$ with s as share of citations to patents in different CPC classes at different aggregation levels. Citations within the same class are excluded.

Table D43. Average number of citing CPCs: AI descendants

	Citing keyword	Citing science	Citing WIPO	Citing USPTO
1 digit	1.32	1.15	1.23	1.16
3 digit	2.75	2.27	2.57	2.33
4 digit	4.82	3.99	4.57	4.06

This table shows the number of different CPC classes making a citation to an average patent of the respective group. Citations within the same class are excluded.

Table D44. Average number of citing CPCs (cited): AI descendants

	Citing keyword	Citing science	Citing WIPO	Citing USPTO
1 digit	2.35	2.17	2.24	2.17
3 digit	4.60	4.04	4.39	4.10
4 digit	7.51	6.60	7.29	6.63

The table shows the numbers of different CPC classes making a citation to an average patent of the respective group that receives at least one citation. Citations within the same class are excluded.

Table D45. Average citation lags by group of AI citing patents

Period	Citing keyword	Citing science	Citing WIPO	Citing USPTO
1990–1999	12.79	12.59	12.66	12.47
2000–2009	8.95	9.03	9.01	8.84
2010–2019	4.30	4.29	4.22	4.28

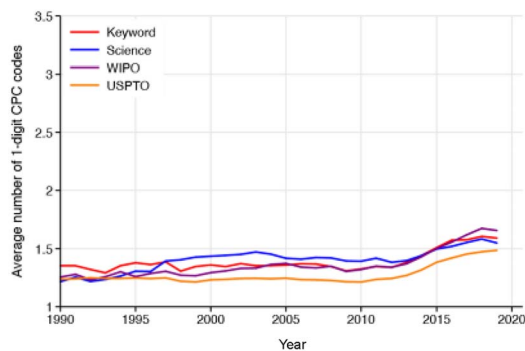
This table shows the average number of years it takes until a patent in the sample is cited. The average number of years is lower during the more recent decade as the maximal time lag is truncated since our data ends in 2019. Note that the group of AI citing includes all patents that cite AI but are not identified as AI by the respective approach.

D.3.4. Complementarity

Table D46. Average number of 1-, 3-, and 4-digit CPCs per patent

	Keyword	Science	WIPO	USPTO
1 digit	1.39	1.40	1.36	1.27
3 digit	1.61	1.64	1.64	1.43
4 digit	1.88	1.97	2.05	1.64

The table shows the average of annual average number of technology classes by 1-, 3-, or 4-digit CPC per patent.

**(a)** Average number of 1-digit CPC**Figure D12.** Patent-level diversity—average technology classes.

D.3.5. Concentration by technology field

Table D47. List of top-5 firms and HHI (by CPC section). IBM, Microsoft, and Google are among the top-five AI patenting firms across all four groups.

	CPC A	CPC B	CPC C	CPC G	CPC H	CPC Y
Keyword	Intuitive Surgical Inc iRobot Corp Siemens Samsung TECH HOLDINGS CORP	Fanuc Corp Honda Motor Seiko Epson Yaskawa Elect Corp iRobot Corp	Applied Materials Inc Hitachi SYMYX TECH INC EBARA CORP General Electric	IBM Corp Microsoft Google Samsung Fanuc Corp	IBM Corp Applied Materials Inc Samsung Microsoft Google	Fanuc Corp Samsung iRobot Corp IBM Corp Boeing
HHI (Keyword)	0.0139	0.0079	0.0102	0.0100	0.0079	0.0071
Science	GENENTECH INC HUMAN GENOME SCI- ENCES INC Siemens ZYMOGENETICS INC Microsoft	SILVERBROOK RE- SEARCH PTY LTD iRobot Corp AUTOMOTIVE TECH INT INC FORD GLOBAL TECH LLC MAGNA ELECT INC	GENENTECH INC HUMAN GENOME SCI- ENCES INC ZYMOGENETICS INC MILLENNIUM PHAR- MACEUTICALS INC EIDU PONT DE NEMOURS & CO	IBM Corp Microsoft Google Apple Sony	Microsoft IBM Corp Google AT&T Sony	IBM Corp SILVERBROOK RE- SEARCH PTY LTD Microsoft DEXCOM INC MONSANTO TECH LLC 0.0051
HHI (Science)	0.0035	0.0080	0.0044	0.0161	0.0108	0.0051
WIPO	Siemens Sony Samsung Microsoft IBM Corp	Toyota FORD GLOBAL TECH LLC Honda Motor Genral Motors	SILVERBROOK RE- SEARCH PTY LTD Applied Materials Inc UBIOME INC ROCHE MOLECULAR SYSTEMS INC NANOGEN INC	IBM Corp Microsoft Google Canon Samsung	IBM Corp Microsoft Google Sony Samsung	IBM Corp Toyota SILVERBROOK RE- SEARCH PTY LTD Genral Motors Google
HHI (WIPO)	0.0049	0.0088	0.0190	0.0117	0.0084	0.0069
USPTO	Siemens Philips General Electric MEDTRONIC INC Sony	Genral Motors FORD GLOBAL TECH LLC Toyota Robert Bosch Honda Motor	AFFYMETRIX INC Applied Materials Inc EIDU PONT DE NEMOURS & CO SYNGENTA PARTICIPA- TIONS AG MONSANTO TECH LLC	IBM Corp Microsoft Google Intel	IBM Corp Microsoft AT&T Google Samsung	IBM Corp Intel General Electric Genral Motors FORD GLOBAL TECH LLC 0.0097
HHI (USPTO)	0.0059	0.0056	0.0020	0.0172	0.0102	0.0097

Entries show the *P*-value of a two-sided paired Wilcoxon signed rank test for the hypothesis that the compared pair ranks equal.

Table D48. List of top-5 firms and HHI (by selected top 3-digit CPC). IBM, Microsoft, and Google are among the top-five AI patenting firms across all four groups.

	H04 (ELECTRIC COMMUNICATION)	G06 (COMPUTING)	A61 (MEDICAL & VETERINARY SCIENCE)	G01 (MEASURING; TESTING)	Y02 (CLIMATE TECHNOLOGIES)
Keyword	IBM Corp Microsoft Samsung Google Intel 0.01167	IBM Corp Microsoft Google Amazon Intel 0.01699	Intuitive Surgical Inc Siemens ETHICON LLC HANSEN MEDICAL INC Philips 0.02382	Samsung Siemens Boeing General Electric IBM Corp 0.00337	Fanuc Corp Samsung iRobot Corp Boeing IBM Corp 0.00727
HHI (Keyword)	0.01167	0.01699	0.02382	0.00337	0.00727
Science	Microsoft	IBM Corp	GENENTECH INC	Microsoft	SILVERBROOK RE-SEARCH PTY LTD IBM Corp
	IBM Corp	Microsoft	HUMAN GENOME SCI-ENCES INC	Siemens	IBM Corp
HHI (Science)	Google Sony AT&T 0.01207	Google Apple Sony 0.02173	Siemens ZYMOGENETICS INC Philips 0.00392	IBM Corp Google GENENTECH INC 0.00264	Microsoft DEXCOM INC MONSANTO TECH LLC 0.00522
WIPO	IBM Corp Microsoft Google	IBM Corp Microsoft Google	Siemens Philips IBM Corp	Google IBM Corp Toyota	IBM Corp Toyota SILVERBROOK RE-SEARCH PTY LTD General Motors
HHI (WIPO)	Sony AT&T 0.00928	Canon Sony 0.01345	CARDIAC PACEMAK-ERS INC Samsung 0.00575	Boeing Honeywell 0.00396	Google 0.00707
USPTO	IBM Corp Microsoft AT&T Google	IBM Corp Microsoft Google Intel	Siemens Philips General Electric MEDTRONIC INC	IBM Corp Siemens General Electric SCHLUMBERGER TECH CORP Hitachi	IBM Corp Intel General Motors General Electric
HHI (USPTO)	Samsung 0.01109	Hewlett Packard 0.02263	CARDIAC PACEMAK-ERS INC 0.00876	0.00409	FORD GLOBAL TECH LLC 0.00970

Entries show the *P*-value of a two-sided paired Wilcoxon signed rank test for the hypothesis that the compared pair ranks equal.

References

- Agrawal, AJAY, JOHN McHale, and ALEXANDER Oetl (2018), 'Finding needles in haystacks: Artificial intelligence and recombinant growth,' in A. Agrawal, J. Gans and A. Goldfarb (eds), *The Economics of Artificial Intelligence: An Agenda*. University of Chicago Press: Chicago, US, pp. 149–174. <https://doi.org/10.7208/chicago/9780226613475.003.0005>
- Agrawal, AJAY, JOSHUA Gans and AVI Goldfarb (2019), 'Economic policy for artificial intelligence,' *Innovation Policy and the Economy*, 19(1), 139–159. <https://doi.org/10.1086/699935>.
- Ahmadpoor, MOHAMMAD and BENJAMIN F. Jones (2017), 'The dual frontier: patented inventions and prior scientific advance,' *Science*, 357(6351), 583–587. <https://doi.org/10.1126/science.aam9527>.
- Alderucci, DEAN, LEE Branstetter, EDUARD Hovy, ANDREW Runge, and NIKOLAS Zolas (2020). 'Quantifying the impact of AI on productivity and labor demand: Evidence from US census microdata,' In: *Allied Social Science Associations - ASSA 2020 Annual Meeting*.
- Arthur, W. BRIAN (2009), *The Nature of Technology: What It Is and How It Evolves*. Simon and Schuster: New York, US.
- Babina, TANIA, ANASTASSIA Fedyk, ALEX He and JAMES Hodson (2024), 'Artificial intelligence, firm growth, and product innovation,' *Journal of Financial Economics*, 151, 103745. <https://doi.org/10.1016/j.jfineco.2023.103745>.
- Barbieri, NICOLO, KERSTIN Hötte and PETER Persoon (2025), 'The evolving boundary of green technology,' *Technical Report*. arXiv preprint arXiv:2503.21310. <https://doi.org/10.48550/arXiv.2503.21310>.
- Bianchini, STEFANO, MORITZ Müller, and PIERRE Pelletier (2020). 'Deep learning in science'. arXiv preprint arXiv:2009.01575. <https://doi.org/10.48550/arXiv.2009.01575>.
- Bresnahan, TIMOTHY F. (2023), 'What innovation paths for AI to become a GPT?' *Journal of Economics & Management Strategy*, 33(2), 305–316. <https://doi.org/10.1111/jems.12524>.
- Bresnahan, TIMOTHY F. and MANUEL Trajtenberg (1995), 'General purpose technologies 'engines of growth?'' *Journal of Econometrics*, 65(1), 83–108. [https://doi.org/10.1016/0304-4076\(94\)01598-T](https://doi.org/10.1016/0304-4076(94)01598-T).
- Brynjolfsson, ERIK, DANIEL Rock, and CHAD Syverson (2019). 'Artificial intelligence and the modern productivity paradox,' in A. Agrawal, J. Gans and A. Goldfarb (eds), *The Economics of Artificial Intelligence: An Agenda*. University of Chicago Press: Chicago, US, pp. 23–60. <https://doi.org/10.3386/w24001>.
- Brynjolfsson, ERIK, DANIEL Rock and CHAD Syverson (2021), 'The productivity J-curve: how intangibles complement general purpose technologies,' *American Economic Journal: Macroeconomics*, 13(1), 333–372. <https://doi.org/10.1257/mac.20180386>.
- Cazzaniga, MAURO, Ms FLORENCE Jaumotte, LONGJI Li, MR GIOVANNI Melina, AUGUSTUS J. Panton, CARLO Pizzinelli, EMMA J. Rockall and Ms MARINA MENDES TAVARES (2024), *Gen-AI: Artificial Intelligence and the Future of Work*. Staff Discussion Notes, International Monetary Fund. <https://www.imf.org/-/media/Files/Publications/SDN/2024/English/SDNEA2024001.ashx>.
- Cleveland, WILLIAM S. (1979), 'Robust locally weighted regression and smoothing scatterplots,' *Journal of the American Statistical Association*, 74(368), 829–836.
- Cockburn, I. M., Henderson, R., and Stern, S. (2019). 'The impact of artificial intelligence on innovation: an exploratory analysis,' in Agrawal, Gans, Goldberg (eds), *The Economics of Artificial Intelligence*. University of Chicago Press: Chicago, US. pp. 115–148. <https://doi.org/10.7208/9780226613475-006>.
- Cockburn, IAIN M., REBECCA Henderson and SCOTT Stern (2018), 'The impact of artificial intelligence on innovation,' *Working Paper. National Bureau of Economic Research*. Cambridge, Massachusetts, US, pp 115–146. <https://doi.org/doi:10.3386/w24449>.
- Comin, DIEGO and MARTI Mestieri (2018), 'If technology has arrived everywhere, why has income diverged?' *American Economic Journal: Macroeconomics*, 10(3), 137–178. <https://doi.org/10.1257/mac.20150175>.
- Crafts, NICHOLAS (2021), 'Artificial intelligence as a general-purpose technology: an historical perspective,' *Oxford Review of Economic Policy*, 37(3), 521–536. <https://doi.org/doi:10.1093/oxrep/grab012>.
- Dafoe, ALLAN (2018), *AI Governance: A Research Agenda*. Tech. Rep. Governance of AI Program. Future of Humanity Institute, University of Oxford: Oxford, UK, <http://www.fhi.ox.ac.uk/wp-content/uploads/Go vAI-Agenda.pdf>.
- Ducuing, CHARLOTTE (2020), 'Data as infrastructure? A study of data sharing legal regimes,' *Competition and Regulation in Network Industries*, 21(2), 124–142. <https://doi.org/10.1177/1783591719895390>.
- EPO (2021), 'PATSTAT online (version 2.7.2), patent information for experts [data set],' *Global European Patent Office*, <https://data.epo.org/expertservices/index.html>.
- Fanti, LUCREZIA, DARIO Guarascio and MASSIMO Moggi (2022), 'From heron of Alexandria to Amazon's Alexa: a stylized history of AI and its impact on business models, organization and work,' *Journal of Industrial and Business Economics*, 49(3), 409–440.
- Franco, SEBASTIÁN FERNÁNDEZ, JUAN M. Graña, DAVID Flacher and CECILIA Rikap (2023), 'Producing and using artificial intelligence: what can Europe learn from Siemens's experience?' *Competition & Change*, 27(2), 302–331. <https://doi.org/10.1177/10245294221097066>.

- Fujii, HIDEMICHI and SHUNSUKE Managi (2018), 'Trends and priority shifts in artificial intelligence technology invention: a global patent analysis,' *Economic Analysis and Policy*, 58, 60–69. <https://doi.org/10.1016/j.eap.2017.12.006>.
- Geels, FRANK W. (2005), *Technological transitions and system innovations: a co-evolutionary and socio-technical analysis*. Edward Elgar Publishing: UK.
- Giczy, ALEXANDER V., NICHOLAS A. Pairolo and ANDREW Toole (2021), 'Identifying artificial intelligence (AI) invention: a novel AI patent dataset.' USPTO economic working paper No. 2021-2,' *The Journal of Technology Transfer*, 47(2), 476–505. <https://doi.org/10.2139/ssrn.3866793>.
- Goldfarb, AVI, BLEDI Taska and FLORENTA Teodoridis (2023), 'Could machine learning be a general purpose technology? A comparison of emerging technologies using data from online job postings,' *Research Policy*, 52(1), 104653. <https://doi.org/10.1016/j.respol.2022.104653>.
- Granstrand, OVE (2009). 'Innovation and Intellectual Property Rights,' *The Oxford Handbook of Innovation*, in JAN Fagerber and DAVID Mowery (eds). Oxford University Press: Oxford, UK. <https://doi.org/10.1093/oxfordhb/9780199286805.003.0010>.
- Hall, BRONWYN H and MANUEL Trajtenberg (2006). 'Chapter 14: Uncovering general purpose technologies with patent data,' *New Frontiers in the Economics of Innovation and New Technology Essays in Honour of Paul A. David*, in CRISTIANO Antonelli, DOMINIQUE Foray, BRONWYN H. Hall, and W. EDWARD Steinmueller (eds), Cheltenham, UK: Edward Elgar Publishing. <https://doi.org/10.4337/9781845427924.00022>
- Hennemann, MORITZ (2020), 'Artificial intelligence and competition law,' *Regulating Artificial Intelligence*. Springer International Publishing: Cham, pp 361–388.
- Hötte, KERSTIN, ANTON Pichler and FRANÇOIS Lafond (2021), 'The rise of science in low-carbon energy technologies,' *Renewable and Sustainable Energy Reviews*, 139, 110654. <https://doi.org/10.1016/j.rser.2020.110654>.
- Jacobides, MICHAEL G., STEFANO Brusoni and FRANCOIS Candelon (2021), 'The evolutionary dynamics of the artificial intelligence ecosystem,' *Strategy Science*, 6(4), 412–435.
- Jaffe, ADAM B. and GAÉTAN De Rassenfosse (2017), 'Patent citation data in social science research: overview and best practices,' *Journal of the Association for Information Science and Technology*, 68(6), 1360–1374. <https://doi.org/10.1002/asi.23731>.
- Jelinek, THORSTEN, WENDELL Wallach and DANIL Kerimi (2021), 'Policy brief: the creation of a G20 coordinating committee for the governance of artificial intelligence,' *AI and Ethics*, 1(2), 141–150. <https://doi.org/10.1007/s43681-020-00019-y>.
- Jovanovic, BOYAN and PETER L. Rousseau (2005), 'General purpose technologies. Working paper,' *National Bureau of Economic Research*. Cambridge, Massachusetts, US. <https://doi.org/10.3386/w11093>.
- Jee, JUNG, SU and SO YOUNG Sohn (2023), 'A firm's creation of proprietary knowledge linked to the knowledge spilled over from its research publications: the case of artificial intelligence,' *Industrial and Corporate Change*, 32(4), 876–900. <https://doi.org/10.1093/icc/dtad002>.
- Jurowetzki, ROMAN, DANIEL Hain, JUAN Mateos-Garcia and KONSTANTINOS Stathoulopoulos (2021), 'The privatization of AI research (–ers): causes and potential consequences – from university-industry interaction to public research brain-drain?', arXiv preprint arXiv:2102.01648. <https://doi.org/10.48550/arxiv.2102.01648>.
- Klinger, JOEL, JUAN C. Mateos-Garcia and KONSTANTINOS Stathoulopoulos (2018), 'Deep learning, deep change? Mapping the development of the artificial intelligence general purpose technology,' arXiv preprint arXiv:1808.06355. <https://arxiv.org/abs/1808.06355>.
- Klinger, JOEL, JUAN Mateos-Garcia and KONSTANTINOS Stathoulopoulos (2020), 'A narrowing of AI research?', arXiv preprint arXiv:2009.10385. <https://arxiv.org/abs/2009.10385>.
- Kovács, BALÁZS, GIANLUCA Carnabuci and FILIPPO CARLO Wezel (2021), 'Categories, attention, and the impact of inventions,' *Strategic Management Journal*, 42(5), 992–1023. <https://doi.org/10.1002/smj.3271>.
- Krafft, PM, MEG Young, MICHAEL Katell, KAREN Huang, and GHISLAIN Buggingo (2020). 'Defining AI in policy versus practice,' *Proceedings of the AAAI/ACM Conference on AI, Ethics, and Society*. Association for Computing Machinery: New York, NY, United States, pp. 72–78. <https://doi.org/10.1145/3375627.3375835>.
- Lafond, FRANÇOIS and DANIEL Kim (2019), 'Long-run dynamics of the US patent classification system,' *Journal of Evolutionary Economics*, 29(2), 631–664. <https://doi.org/10.1007/s00191-018-0603-3>.
- Lipsey, RICHARD G., KENNETH I. Carlaw and CLIFFORD T. Bekar (2005), *Economic Transformations: General Purpose Technologies and Long-term Economic Growth*. Oxford University Press: Oxford, <https://EconPapers.repec.org/RePEc:oxp:obooks:9780199290895>.
- Marx, MATT and AARON Fuegi (2019). *Reliance on Science: Worldwide Front-Page Patent Citations to Scientific Articles*. Boston University Questrom School of Business Research Paper No. 3331686. November 15, 2019. <https://doi.org/10.2139/ssrn.3331686>.
- Marx, MATT and AARON Fuegi (2020), 'Reliance on science: worldwide front-page patent citations to scientific articles,' *Strategic Management Journal*, 41.9, 1572–1594. <https://doi.org/10.1002/smj.3145>. *Reliance on Science in Patenting (Version v30) [Data set]*. (2021). Zenodo, <https://doi.org/10.5281/zenodo.4850763>.

- Mazzucato, MARIANA, MARIETJE Schaake, SEB Krier, and JOSH Entsminger (2022). *Governing artificial intelligence in the public interest. Working Paper Series, IIPP WP 2020–12*. UCL Institute for Innovation and Public Purpose: London, UK. <https://www.ucl.ac.uk/bartlett/public-purpose/wp2022-12>.
- Meyer, MARTIN (2000), 'Does science push technology? Patents citing scientific literature,' *Research Policy*, 29,3, 409–434. [https://doi.org/10.1016/S0048-7333\(99\)00040-2](https://doi.org/10.1016/S0048-7333(99)00040-2).
- Mitchell, TOM MICHAEL (1997), *Machine Learning* Vol. 1, McGraw-Hill: New York.
- Moser, PETRA and TOM Nicholas (2004), 'Was electricity a general purpose technology? Evidence from historical patent citations,' *American Economic Review*, 94(2), 388–394. <https://doi.org/10.1257/0002828041301407>.
- Narin, FRANCIS, KIMBERLY S. Hamilton and DOMINIC Olivastro (1995), 'Linkage between agency-supported research and patented industrial technology,' *Research Evaluation*, 5(3), 183–187. <https://doi.org/10.1093/rev/5.3.183>.
- Petit, NICOLAS (2017), 'Antitrust and artificial intelligence: a research agenda,' *Journal of European Competition Law & Practice*, 8(6), 361–362. <https://doi.org/10.1093/jeclap/lpx033>.
- Petralia, SERGIO (2020), 'Mapping general purpose technologies with patent data,' *Research Policy*, 49(7), 104013. <https://doi.org/10.1016/j.respol.2020.104013>.
- Rikap, CECILIA (2024), 'Varieties of corporate innovation systems and their interplay with global and national systems: Amazon, Facebook, Google and Microsoft's strategies to produce and appropriate artificial intelligence,' *Review of International Political Economy*, 31(6), 1735–1763. <https://doi.org/10.1080/09692290.2024.2365757>.
- Roche, CATHY, P. J. Wall and DAVE Lewis (2023), 'Ethics and diversity in artificial intelligence policies, strategies and initiatives,' *AI and Ethics*, 3(4), 1095–1115. doi:10.1007/s43681-022-00218-9 1, 21.
- Schmitt, LEWIN (2022), 'Mapping global AI governance: a nascent regime in a fragmented landscape,' *AI and Ethics*, 2(2), 303–314. <https://doi.org/10.1007/s43681-021-00083-y>.
- Schumpeter, JOSEPH A., MARKUS C. Becker, HANS ULRICH Eßlinger, ULRICH Hedtke and THORBJØRN Knudsen (2005), 'Development,' *Journal of Economic Literature*, 43(1), 108–120 (post-hum published and introduced by Knudsen, Thorbjørn, pp. 108–120). <https://doi.org/10.1257/0022051053737825>.
- Sinha, ARNAB, ZHIHONG Shen, Yang SONG, HAO Ma, DARRIN Eide, Bo-June (Paul) HSU, and KUANSAN Wang (2015). *An overview of Microsoft Academic Service (MAS) and Applications. Proceedings of the 24th International Conference on World Wide Web (WWW 15 Companion)*. ACM, New York, NY, USA, 243–246. <https://doi.org/10.1145/2740908.2742839>.
- Stuart, RUSSELL and PETER Norvig (2003), *Artificial Intelligence: A Modern Approach*. Prentice Hall: Upper Saddle River, NJ, USA,
- Tambe, PRASANNA, LORIN Hitt, DANIEL Rock, and ERIK Brynjolfsson (2020). *Digital capital and superstar firms. NBER Working Paper 28285*. National Bureau of Economic Research: Cambridge, Massachusetts, US. <https://doi.org/10.3386/w28285>.
- Tijssen, ROBERT J. W. (2001), 'Global and domestic utilization of industrial relevant science: patent citation analysis of science–technology interactions and knowledge flows,' *Research Policy*, 30(1), 35–54. [https://doi.org/10.1016/S0048-7333\(99\)00080-3](https://doi.org/10.1016/S0048-7333(99)00080-3).
- Trajtenberg, MANUEL (2018), 'AI as the next GPT: a political-economy perspective,' NBER Working Paper 24245. *National Bureau of Economic Research*. Cambridge, Massachusetts, US. <https://doi.org/10.3386/w24245>.
- Trajtenberg, MANUEL, REBECCA Henderson and ADAM Jaffe (1997), 'University versus corporate patents: a window on the basicness of invention,' *Economics of Innovation and New Technology*, 5(1), 19–50. <https://doi.org/10.1080/10438599700000006>.
- Turing, ALAN (1950), 'Computing machinery and intelligence,' *Mind*, 59(236), 433.
- USPTO (2020), *Inventing AI: Tracing the diffusion of artificial intelligence with U.S. patents*. U.S. patent and trademark office, Office of the Chief Economist. <https://www.uspto.gov/sites/default/files/documents/OCE-DH-AI.pdf>.
- Valdes, Y. M. and ILJA J. Rudyk (2017), 'Patents and the fourth industrial revolution: The inventions behind digital transformation,' in *Technical report, European Patent Office, Handelsblatt Research Institute*. <https://www.epo.org/service-support/publications.html?pubid=163#tab3>.
- Verendel, VILHELM (2023), 'Tracking artificial intelligence in climate inventions with patent data,' *Nature Climate Change*, 13(1), 40–47. <https://doi.org/10.1038/s41558-022-01536-w>.
- Webb, MICHAEL (2019). *The Impact of Artificial Intelligence on the Labor Market*. Working Paper. National Bureau of Economic Research: Cambridge, Massachusetts, US. <https://doi.org/10.2139/ssrn.3482150>.
- Whittaker, MEREDITH (2021), 'The steep cost of capture,' *Interactions*, 28(6), 50–55. <https://doi.org/10.1145/3488666>.
- WIPO (2019a). *Technology Trends 2019: Artificial Intelligence: Data Collection Method and Clustering Scheme: Background Paper*. Geneva: World Intellectual Property Organization. https://www.wipo.int/export/sites/www/tech_trends/en/artificial_intelligence/docs/techtrends_ai_methodology.pdf.
- WIPO (2019b), *WIPO Technology Trends 2019: Artificial Intelligence*. World Intellectual Property Organization: Geneva. https://www.wipo.int/edocs/pubdocs/en/wipo_pub_1055.pdf.