LUISS

Institute for European
Analysis and Policy

Jean Monnet Centre of Excellence
on EU Inclusive Open Strategic Autonomy

European Sovereignty in Artificial Intelligence:
A Competence-Based Perspective

Ludovic Dibiaggio, Lionel Nesta, Simone Vannuccini

Working Paper 20/2024

December 17, 2024

This paper was first published as part of the GREDEG Working Paper Series and the SSRN Working Paper Series.

European Sovereignty in Artificial Intelligence: A Competence-Based Perspective*

Ludovic Dibiaggio[†], Lionel Nesta[‡], Simone Vannuccini[§]

Abstract

We present a first-of-its-kind empirical study of technological sovereignty in artificial intelligence, adopting a competence-based perspective. We use patents and publication data to map competencies across AI techniques, functions and applications, and develop a novel measure of integration based on relative specializations and complementarities. We argue that our measure approximates technological sovereignty by capturing local capabilities to innovate in AI. We use our novel measure to explain AI innovation, and unpack integration determinants. Our focus is on the European Union, given its lagging position yet key role in a global landscape increasingly characterized by growing rivalries and fragmentation.

Keywords: artificial intelligence; competences; technological sovereignty; integration; European Union.

JEL codes: O31; O33; 034; D80; L86.

*This paper builds on a report of the same title that has been prepared for the Digital, Governance and Sovereignty Chair of Sciences Po Paris. The authors wish to especially thank Mohamed Keita for his technical support during the elaboration of the report. We also thank Sarah Guillou (OFCE SciencesPO) for her remarks.

†SKEMA Business School (France). Email: ludovic.dibiaggio@skema.edu.

‡Université Côte d'Azur, CNRS, GREDEG (France) and SKEMA Business School (France) and Sciences Po Paris, OFCE (France). Email: lionel.nesta@univ-cotedazur.fr.

§Université Côte d'Azur, CNRS, GREDEG (France). Email: simone.vannuccini@univ-cotedazur.fr.

1 Introduction

In this paper, we develop a novel analytical perspective to examine the relationship between developments in Artificial Intelligence (AI) and technological sovereignty. We define technological sovereignty in the context of AI as the ability of countries to mobilize local (i.e., domestic) competencies and integrate them across a stylized AI value chain, spanning from scientific developments to industrial implementation. Our central premise is that the integration of domestic competencies can be used as a meaningful measure of technological sovereignty.

In a world increasing characterized by geo-political and geo-economic fragmentation, the idea that countries should strive to enhance autonomy and resilience in the production of key, economy-impacting technologies is rapidly gaining traction. The ongoing discussion revolves around the notion of technological sovereignty because a handful of key, enabling and breakthrough technologies are considered strategic assets to lead in the global context. AI is one of such breakthrough technology and is increasingly seen as a strategic, dual-use asset on the geopolitical stage. For this reason, it has become one of the top policy priorities for countries across the world (Kak & West 2024, Bryson & Malikova 2021).

We focus on the European context, as the European Union (EU) is a major global player. However, as highlighted in a series of recent key European policy reports (Draghi 2024, Fuest et al. 2024, Aghion et al. 2024), the EU lags behind on the international stage, trailing the rapid advances at the global technological and competitiveness frontier, predominantly led by the United States and China. By adopting a continental perspective, we address the following questions: is the EU *capable* of achieving technological sovereignty in a fundamental technology like AI, which is increasingly considered pivotal and strategic globally? Does the domestic integration of complementary competencies in AI foster subsequent AI innovation? If so, what factors contribute to increasing within-country integration?

Our contribution is threefold. First, we conduct a set of empirical analyses based on patent and publication data for the period 1990-2021. Inspired by Corea (2019) and firstly implemented by WIPO (2019), we map a stylized value chain relevant to the production and innovation in AI, comprising *techniques* (T), *functions* (F), and *applications* (A) –what we label the TFA value chain. Second, we develop country-specific measures of integration based on relative specializations and complementarities along this stylized AI value chain. We interpret this metric as a proxy for technological sovereignty in AI, and apply the measure to the geopolitical blocks of United States, China, and Europe, together with the set of countries most active in AI. Last, we provide evidence of integration enhancing innovation, suggesting that integration is a strategic policy lever to increase technological sovereignty.

The paper proceeds as follows. In Section 2, we outline our theoretical framework that describes AI through the TFA model and justifies the conceptualization of technological sovereignty as the integration of domestic competences. As we focus on the EU, in Section 3 we describe the "political economy" of AI in Europe, which partially results from the EU gap of investments in the digital economy. Section 4 describes our data and our operationalization of integration into a measure. Throughout Section 5, we first offer an econometric analysis relating integration, as technological sovereignty, with innovation, and then reveal the determinants of integration. Section 6 concludes.

2 Setting the scene

2.1 Innovation as a process of integration

Advancing the concept of technological sovereignty requires a theoretical understanding of innovation as a highly uncertain and cumulative process. The dominant, and probably most realistic, view of innovation is that the links between science, technology, and commercial applications do not follow linear and sequential relationships (Kline & Rosenberg 1986, Fagerberg et al. 2006). After all, innovation is driven by complex interactions among heterogeneous agents with diverse, and sometimes conflicting, objectives (Freeman 1987, Aghion & Griffith 2005). But because of the cumulative nature of technological knowledge, innovation builds on previous ideas, essentially going from upstream knowledge to more downstream, localized and concrete knowledge applications. In other words, innovation has a directional flow, although feedback loops and complex interactions between upstream and downstream knowledge do occur (Acemoglu et al. 2016).[1]

In such a complex system, integration capabilities are critical for at least three reasons. First, integration entails shaping, selecting, and combining techniques dedicated to specific functions to deliver a given service.(Jacobides et al. 2009). This is what gives rise to the system's coherence. Therefore, integration requires the ability to coordinate diverse yet complementary competences held by different stakeholders. Second, the exploration and exploitation of these various technologies require complementary investments, organizational adaptations, and skill development along the value chain to fully realize their transformative potential (Bresnahan & Trajtenberg 1995). Third, integration emerges from a learning process driven by experimentation and adjustments along the value chain. Actors must explore and establish the most promising complementarities. As a result, integration capabilities—the process of combining dispersed technologies into functional applications—become a highly ambiguous, risky, and time-consuming process.

This is particularly true in the emerging phases of technology development. As Rotolo et al. (2015) explain, emerging technologies generate both uncertainty and ambiguity. Uncertainty stems from the competition between technological trajectories combined with the limited predictability of technological development and its potential applications, and the associated lack of factors such as technological skills, financial support, standards, and production costs. Ambiguity arises because

[1] In this respect, the role of basic research on technological development is well documented (Narin et al. 1997, Mansfield 1991, Veugelers & Wang 2019, Ahmadpoor & Jones 2017).

applications remain malleable and unstable, often due to the different or even contradictory meanings ascribed by various agents along the innovation value chain from scientists to end-users.

The combination of uncertainty and ambiguity makes the ability to integrate and reconfigure technologies along the industry value chain a critical source of long-term innovative performance (Pavitt 1984, Nelson 1993, Franco et al. 2009). Breschi et al. (2005) show how technological specialization at the sector level creates networks and linkages across firms that develop complementary capabilities and influence the rate and direction of technological change. At the same time, excessive integration can be potentially dangerous, as it may lock public and private organizations into suboptimal technological trajectories, ultimately jeopardizing their survival.

Integration capability is also the product of country-specific characteristics (Lundvall et al. 1992, Archibugio & Pianta 1992, Malerba & Orsenigo 1993, Nelson 1993). Public research programs, private R&D investments, the provision of human capital, and effective industrial policies are all key ingredients shaping the path towards technological leadership (Nelson 1993, Malerba & Orsenigo 1993, Furman et al. 2002). Building on this line of reasoning, Furman et al. (2002) show that cross-country differences in inventive productivity are determined by national characteristics, such as strong intellectual property rights, economic openness and a complex web of underlying infrastructures such as investments in higher education, the funding of basic research, the presence of industrial clusters, amongs other elements. These factors appear to be critical for predicting the countries' innovative performance.[2]

The view that national characteristics condition integration in innovative activities is also pervasive in the more user-centric approach of national innovation systems (Lundvall et al. 1992, Jensen et al. 2007). Innovation is conceptualized as a learning process based on the interactions between producers and users who engage in col-

[2] The authors further show that in the absence of strong linking mechanisms, upstream scientific and technical activity may spill over to other countries more quickly than opportunities can be exploited by domestic industries.

laborative technological combination search. Since users' needs and preferences for products that do not exist are often ambiguous, product development is an interactive problem-solving process based on experimentation, trial-and-error (Rosenberg 1982). The direction and the effectiveness of innovation, therefore, depend on the ability of producers and users to share accepted standards and codes, as well as create a trustful environment to develop a common understanding of the product.

The argument is that the national context shapes institutions defined as "common habits, routines, established practices, rules, or laws that regulate the relations and interactions between individuals and groups" (Edquist & Johnson 1996, p. 46). Comparative studies of innovation conditions across countries reveal significant differences in both formal and informal relationships between agents, including firms' governance, the organization of work within and between firms, social values (e.g., egalitarianism vs. large social disparities), cooperative behaviours, and worker mobility (e.g. Edquist & Lundvall 1993, OECD 1999). In other words, institutional conditions stimulate the formation of social capital and interactive learning, which are essential strategic capabilities for combining and integrating upstream and downstream knowledge along the value chain.

Altogether, we argue that domestic integration capabilities are key for the innovative performance of countries. While expertise in upstream knowledge is a critical source of innovation in new functionalities, downstream applications are essential for capturing value and creating incentives to invest in upstream R&D. Higher levels of complementarity along the innovation value chain increases the firm's propensity to innovate, spurring private research investments and thereby enhancing innovative capabilities within and across industrial sectors. (Fronzetti Colladon et al. 2025, Pichler et al. 2020).

2.2 The value chain of innovation in Artificial Intelligence

Our object of analysis is Artificial Intelligence (AI). AI is a breakthrough and a system technology (Dibiaggio et al. 2022, Sheikh et al. 2023, Vannuccini & Prytkova 2024) which has "prediction machines" (Agrawal et al. 2022) at its core. As a system, AI consists of a collection of complementary hardware and software components, plus data and talent. The AI system evolves thanks to feedback mechanisms and dynamic complementarities forming among its components: advancements in AI techniques prompt further inventive opportunities in complementary technologies or applications, thereby increasing the incentives for their adoption (Bresnahan 2003, Aghion et al. 2009). Subsequent spillovers generate positive feedback between technical inventions and the co-invention of functions that create opportunities for further innovations. New functionalities shape the design of products and services, reinforcing complementarities along the value chain (Rosenberg 1982, Mowery 1992, Nelson & Rosenberg 1993). At the same time, the uncertainty and ambiguity inherent in the dynamics of the structure of complementarities along the value chain, coupled with path-dependent investment trajectories, can generate dynamic coordination failures; the risk of prematurely committing to an inferior design or being locked into sub-optimal options is inherently high, underscoring the need to preserve diversity in the technological environment (David 1985, Arthur 1994, Aghion et al. 2009).

If the production of and innovation in AI are systemic efforts, a useful way capture their "system-ness" is by mapping the competencies required to innovate in AI across a stylized series of steps. We adopt WIPO (2019)'s Technique-Function-Application (TFA) model, or value chain. The key tenet of this framework is that different *AI techniques* give rise to specific *AI functions*; these functions are employed in different *AI applications*, which approximate techno-economic activities. As a result, we have a many-to-many relational structure. The TFA model is not a traditional value chain, or a full-stack representation of how value is accrued in the production of AI systems stage by stage. Rather, it is a simplified picture of how AI innovation is developed, from the more generic software developments to their adoption into specific function and practical applications. The TFA model is particularly useful, as it is tailored on the features of the patent and publication data we use. While a direct line from techniques to applications would trivialize the many non-linear circuits driving innovation in AI, the stylized value chain has the benefit of being able to capture the idea that an actor in AI innovation can *specialize* in one or more stages, and that — as we hypothesize — integrating competences along all stages may result in higher rates of innovation.

The rationale for working with a stylized TFA value chain of AI is grounded in the idea that solving specific AI-related problems (often approximated by functions) involves developing algorithms and methods that build on specific approaches or AI paradigms, such as symbolic or sub-symbolic AI. Each paradigm can be based on an array of techniques with specific properties, which can be more or less adapted to address certain types of problems. For instance, within the class of deep learning techniques, convolutional neural networks (CNNs) are used in computer vision tasks, while generative adversarial networks (GANs) have been used extensively to produce images. Reinforcement learning approaches have been successful in the AI-in-science context (e.g., in tasks related to addressing the protein folding problem), while language models have been pivotal in dealing with prediction tasks involving text embeddings. Currently, economic actors have bet on language models becoming the dominant foundational multi-modal design underlying all AI commercial applications; however, a large variety of techniques continues to exist in the AI world, which are subject to inventive activities.

2.3 Integration as technological sovereignty

Our premise is that the TFA model can be operationalized to gain insights into the degree of autonomy an actor possesses to produce and innovate in AI. This is particularly important, as the question of autonomy (or dependence) in the production of strategic technologies is emerging as a key political priority in a global landscape increasingly marked by rivalries. In this context, countries are striving to achieve technological sovereignty (Edler et al. 2023).

Technological sovereignty involves the capacity to make autonomous decisions about technology development, deployment, and regulation without undue influence or dependence on external entities. As a matter of principle, technological sovereignty should not be viewed as a static, nationalistic, defensive concept focused on erecting legal protection barriers. Instead, it should be understood as a dynamic concept centered on building the capability to develop adaptive capacities (Edler et al. 2023). The concept combines the ability to develop the competences and resources necessary to deliver technologies pivotal for competitiveness and growth, with the capacity to source and access the complementary technologies and assets required to produce industrial applications.

From an economic theory perspective, it is important to stress how the increasing consensus around technological sovereignty poses a direct challenge to the classical economic theories that advocate for specialization in international trade (Krugman 1979). In fact, the notion of technological sovereignty emphasizes building domestic capabilities and infrastructure to develop technologies independently, regardless of factor endowments or global efficiency arguments. This focus arises because economic considerations, such as the benefits of trade and specialization, take a backseat to the political concerns like national security, economic independence, or strategic interests.

As explained earlier, integration requires specific competencies to ensure coordination across the evolving boundaries of technological specializations with circular, interlocking, and often time-delayed relationships. We relate the concepts of integration, and technological sovereignty in AI (and in more broadly) as follows: an actor (a country, or the EU) can exhibit specialization in one, two, all, or none of the layers of the TFA value chain of AI depending on the activities carried out by firms in its territory. If the actor exhibits a relative advantage in AI innovation within one of the TFA domains, we consider it to have a comparatively high level of competence in generating new AI knowledge in that domain. The greater the number of domains this actor specializes, the more transversal its competencies become. If these domains are also complementary, they contribute to greater autonomy in producing all elements of AI innovation, from techniques to industrial implementations. A lack of integration, missing competencies, and reliance on the expertise of foreign actors prevent a country from fully capitalizing on its investments. Since advances in techniques can generate upstream and downstream innovation opportunities, the absence of downstream application producers may hinder a country from capturing the value of these opportunities. A clear illustration of this dynamic is when scientific discoveries made by public universitiesare exploited abroad by foreign corporations.

In summary, within a competence-based framework, technological sovereignty in AI can be defined as the ability to mobilize and integrate technological competencies domestically acrossthe entire AI innovation value chain ranging from the development of new or improved algorithms (techniques), the creation of new AI-based functions to the practical embodiment of these techniques and functions in new applications.

3 The political economy of AI in the European Union

3.1 Europe's follower position in the digital economy

To map European strengths and weaknesses in AI competences and innovation, a good starting point is to take a broader perspective that encompasses digital technology. In fact, according to Carlota Perez, AI "is better understood as a key development within the still-evolving information-communications-technology (ICT) revolution."[3]. Understanding the EU's global position in ICT can shed light on the roots of potential deficiencies and gaps related to AI. As argued by Bock et al. (2024) as well in the recent reports by Draghi (2024) and Fuest et al. (2024), the EU lags behind other actors, particularly the US, in terms of competitiveness, despite being the world's largest single market. Moreover this gap has been widening over time. This disparity began well before the Covid-19 pandemic, reflecting a gradual decoupling in overall economic performance. For instance, Bock et al. (2024)report that the gap in private investments in ICT between between the Eurozone and the US stood at approximately 150 billion euros in 2000, rising to a concerning 600 billion euros in 2019. Importantly, the gap holds for all types of private ICT investments (equipment, services, research and development). Following Bock et al. (2024), it appears that the primary contributors to United States' lead in R&D investments over other international actors are the ICT services sectors. This dominance can reasonably be attributed to the "GAFAMs effect".[4] In fact, even within the realm of AI, the returns from the exploitation of large datasets (particularly due to cross-domain network externalities), along with market dominance in cloud services, digital advertising, application markets, and the anticipated gains from the rise of AI models, are prompting tech giants to invest massively in R&D –and in upstream R&D in

[3]https://www.project-syndicate.org/magazine/ai-is-part-of-larger-technological-revolution-by-carlota-perez-1-2024-03 — last access August 2024).

[4]The companies included in this acronym are Google (now Alphabet), Amazon, Facebook (now Meta), Apple, and Microsoft.

particular– in addition to capital expenditure.

The European position is concerning for two reasons. First, the lack of investment in ICT services results in a slow pace of economic digitization. Second, the absence of European champions in digital services leads to reduced investments in R&D and ICT equipment, both of which are prerequisites for AI development. This discussion on Europe's investment gap, in particular in ICT — the foundational substrate of AI — highlights the challenge facing the EU if its ambmition to lead and achieve autonomy in the production of advanced digital technologies, with AI at the forefront. As Fuest et al. (2024) point out, the EU has been trapped in a "middle technology trap" for two decades, focuse on automotive manufacturing and lacking scale and R&D expenditures beyond that sector. Escaping this trap will require massive financial efforts, which open the door to discussions on (industrial) policy, resourcing, and investments at the continental level (Fontana & Vannuccini 2024). Given this context, we can now return to our key question: is the EU *capable* of producing a breakthrough system technology such as AI?

3.2 The State of European AI and Technological Sovereignty

The question of whether the EU can overcome its laggard position in the digital economy and take a global leadership in the development of emerging breakthrough technologies — and AI in particular — ultimately hinges on whether it possesses the necessary competencies. This can be reframed in terms of technological sovereignty: can the EU can the technology autonomously, to a certain extant?

This question is subject to intense debate within the EU. In European trade policy circles, technological sovereignty is closely tied to the concept of *open strategic autonomy*(Timmers 2018). Open strategic autonomy can be summarized as prioritizing autonomy while allowing for cooperation when feasible. It is defined as 'the ability to shape the new system of global economic governance and develop mutually beneficial bilateral relations, while protecting the EU from unfair and abusive practices, including to diversify and solidify global supply chains to enhance resilience to future crises".[5] Even in the open strategic autonomy format, sovereignty over technology production and innovation is a matter of ability and, thus, of competencies.

Focusing on AI, rapid advances in the technology its global scope of application — as well as its potential dual-use nature in the domain of defense — have placed it at the forefront of the discussion on technological sovereignty. While the real impact of AI on productivity might turn out to be rather modest (Acemoglu 2024), the industry forming around it and the shared narrative around AI's transformative impact have intensified the focus on AI, which is considered a "strategic asset" (Ding & Dafoe 2021). Next to potential trade dependencies, one must consider the increasing power of large tech companies ("Big Tech") to shape the AI technological landscape as well as different markets. The anti-competitive and innovation-harming role of Big Tech in AI and beyond is being increasingly placed under the spotlight, as these "intellectual monopolies" shape a novel technological regime around their objectives and appropriate most of the returns on global innovation (Rikap 2023). Big Tech's agenda may not align well with technological sovereignty priorities in Europe. Their cumulative advantage has generated increasing polarization in capabilities (both in resources and competencies), leading to brain-drain from academia to the private sector and to reduction in the diversity within AI research (Klinger et al. 2020, Ahmed et al. 2023). Furthermore, Big Tech's capital investments in AI-related hardware such as Nvidia's Graphics Processing Units (GPUs) are draining the supply of the key inputs of AI systems, which are then allocated exclusively to commercial uses rather than to pursue goals that favor the public interest.[6] Finally, their control of bottlenecks across the whole AI stack is a direct challenge to European autonomy in the development of (and innovation in) the technology.

For the EU, the increasing attention to technological sovereignty reflects the

[5] See the 2021 European Commission Staff Working Document — Strategic dependencies and capacities (last access: July 2024).

[6] See, for instance, the distribution of compute across private and public actors as provided by https://www.stateof.ai/compute (last access: December 2024).

block's concerns about losing the ability to act autonomously in a global technological system that is increasingly fragmented and in which trade and industrial policies are weaponized for geopolitical ends. In this context, AI policy is a matter of science, technology, and industrial policy, and a consensus on that is emerging at the institutional level and among civil society, that increasingly advocates for a reorientation of AI industrial policy from competitiveness to public interest priorities (Kak & West 2024). However, in terms of legislation, the dominant approach of the EU towards AI and more generally digital technology, platforms, and marketplaces has been that of protecting citizens and favoring market contestability. These principles inspired the most important European horizontal regulation exercises in the field: the General Data Protection Regulation (GDPR), the Digital Markets Act (DMA), and the Digital Services Act (DSA). The AI Act, just entered into effect, is another piece of the same puzzle. In this respect, the EU's political economy of AI has been one geared ensuring rights, product regulation, and addressing the dominance of Big Tech as oligopolistic platform business models.

While EU horizontal regulations in the digital realm have generally been a success and boosted the so-called "Brussels effect", with the rest of the world following and imitating European legislation, less has been done on the investments side. Orchestrating initiatives aimed at addressing the lack of continental champions in the hardware and services layers of ICT have not produced successes yet; for instance, the slow-moving Gaia-X project of a federated European cloud infrastructure[7] testifies to the difficulty of building alternatives to the early American hyperscalers, who enjoy path-dependent gains from their head start in the market and now control the upstream layers of the AI technology stack.

At the same time, we are witnessing an acceleration of industrial policy activism, especially after the introduction of the Inflation Reduction Act (IRA) and the CHIPS and Science Act in the US (Kleimann et al. 2023). Among the few tools specifically

[7]See, for example, https://www.politico.eu/article/chaos-and-infighting-are-killing-europes-grand-cloud-project/ (Last accessed: July 2024)

focused on AI, there is pilot of the National AI Research Resource (NAIRR) launched by the Biden Administration, and aimed at sharing computing resources among AI actors to lower the entry costs into the field.[8] The European Commission has been working in a similar direction with the proposal of creating "AI factories". [9] In line with the actions on the American side of the Atlantic, the EU initiative consists mostly of sharing high performance computational capacity — a key input into the production of AI systems — re-orienting the existing allocations of the European budget rather than providing additional resources to increase competitiveness and competences with regard to AI.

4 Measuring specialization and integration in AI

4.1 Data

Our analysis is based on patent and publication data. We consider the granularity that this type of data offers as most appropriate to address our research question, given our focus on AI *technology* and *competences*. For the part of our analysis related to patents, we exploit the PATSTAT dataset (Autumn 2023 edition). Despite their limits (Mezzanotti & Simcoe 2023), patents still provide the broadest coverage of inventive activities in all technological domains, and remain a good measure of innovation output when focusing on large companies and R&D firms. It must be stressed that AI-related inventions do not readily lend themselves to patenting. This is due largely to the fact that AI algorithms are software technology, and software can be patented only when embedded in a tangible (hardware) solution using AI. In this sense, our data might not cover some of the most recent advances in AI software. However, it will capture hardware-embedded technology that is pivotal to

[8]https://nairrpilot.org/. Some have pointed out how the design of this type of policy initiative, which builds on public-private partnerships and licensing agreements, risks favoring Big Tech rather than leveling the playing field: https://foreignpolicy.com/2024/02/12/ai-public-private-partnerships-task-force-nairr/.

[9]See https://digital-strategy.ec.europa.eu/en/policies/ai-factories.

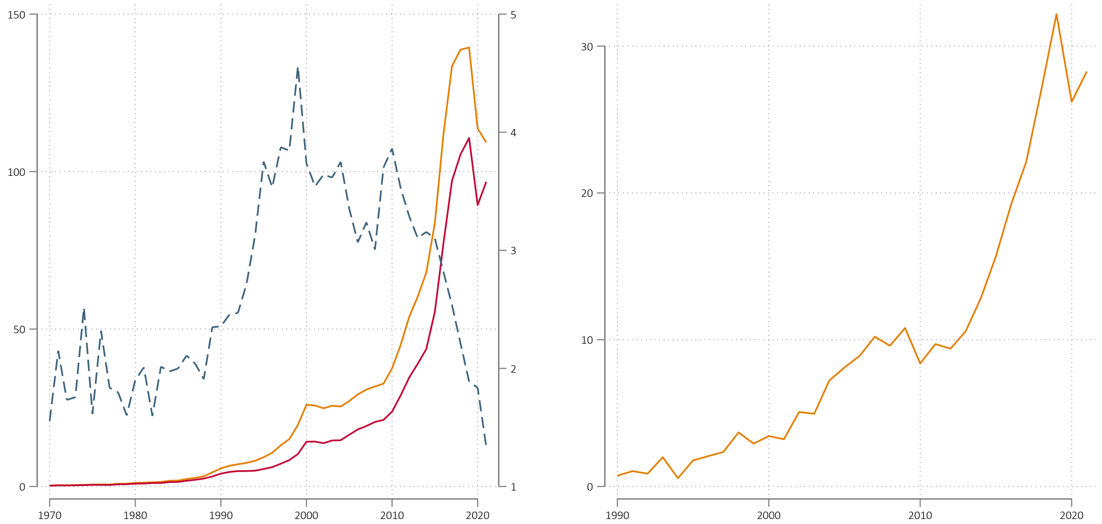
countries' physical production, products and services and competence formation.

We use AI patents for the period 1990-2021. We opted for measuring AI innovation and competencies over an extensive time span because building competencies is a cumulative, time-demanding process characterized by path dependence. Understanding gaps in European technological sovereignty requires a structural, long-term view, that can be achieved only by factoring in AI developments along a decades-long trajectory. Overall, by concatenating the different sources of information, we obtained 1,415,828 patents.

In addition to patents, our study uses scientific publications to track the evolution of AI-related scientific discoveries. We retrieved all papers in the Elsevier Scopus database (2023 edition) that were presented at international AI conferences in the period 1989 to 2023. To select publications concerned exclusively with AI, we relied on conferences identified as the main AI conferences by Baruffaldi et al. (2020). From Scopus, we retrieved all available proceedings of these conferences. We obtained 330,362 publications from conference proceedings from 1989 to 2023.

To identify AI patents and publications, we develop a protocol built on WIPO (2019) that combines two sources of information: patent classifications (technology classes) and keywords-based search. After having being identified, patents and publication are assigned to the TFA categories, and the location of invention (in terms of country) is determined. The latter step is key to our goal to locate and map AI-related competences. The fact that a patent is being developed by inventors from a particular country implies that complementary investments, in terms of infrastructure, researchers, engineers, national innovation system, networks, the underlying education and professional training system, etc., have been made in the first place. Details of the selection, assignment and location procedure can be found in Appendix A, including the full list of techniques and functions used to classify patents and publications in the Technique-Function-Application framework in Table B1.

Figure 1: The dynamics of AI-related patents and publications



(a) Patents (b) Publications

Sources: EPO PATSTAT (Ed. Autumn 2023) for patent data. SCOPUS for publication data. Authors' own calculations. The left axes in Figure 1a are in thousands of patents (orange line), of families (dark red line), and of publications (orange line, Figure 1b). The dotted blue line depicts the average family size (Figure 1a, right axis).

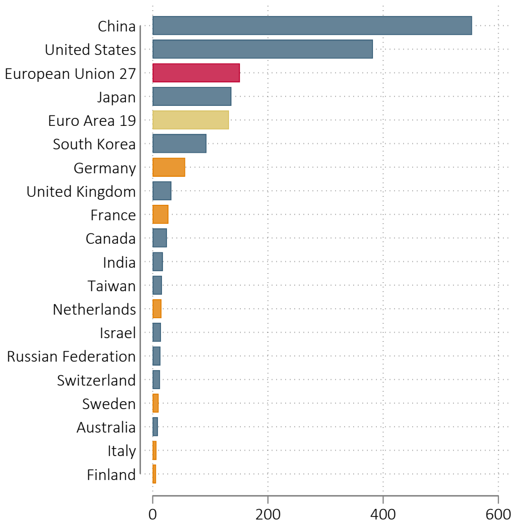
Figure 1 displays the evolution of the number of AI patents filed since 1970 (Panel 4a) and publications (Panel 4b). The two panels exhibit a similar exponential trend (the most recent decline being a product of time-lags characteristic of patent and publication data), with an acceleration around 2010, which is usually linked to the beginning of the so-called "Deep Learning era" (Sevilla et al. 2022), which marks the resurgence of interest in the so-called connectionist (that is, simply put, neural-network based) AI after previous "AI Winters" (Vannuccini & Prytkova 2024).

As our analysis is focused on Europe, we considered the European Union and the Eurozone (EZ) as countries.[10] Figure 2 displays the distribution of AI-related competences across countries. The top panel ranks countries according to their contribution in terms of frequencies. The bottom panel normalizes the figures by providing the number of patents and publications per million inhabitants. Figure 2 shows that the EU ranks third in terms of patents when looking at patents frequencies, yet with a significant gap: the number of EU27 patents is almost a third of

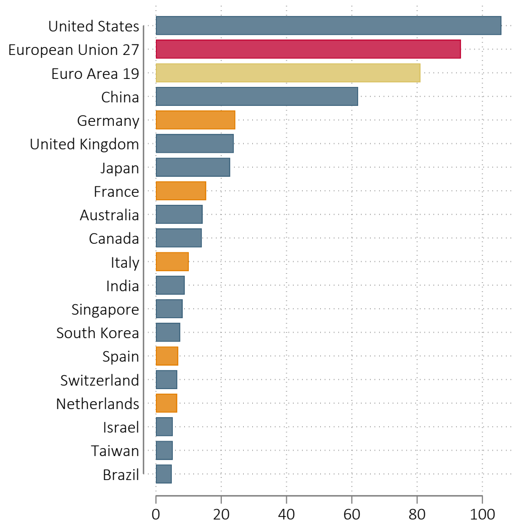
[10]We constructed the statistics on patents and publication in the EU and the EZ by aggregating the information about the individual member states (27 for the EU; 19 for the EZ). We excluded the United Kingdom from all EU statistics.

the number of US patents. Turning to publication counts in AI, EU27 publications match the number of US publications, and greatly exceed Chinese publications.

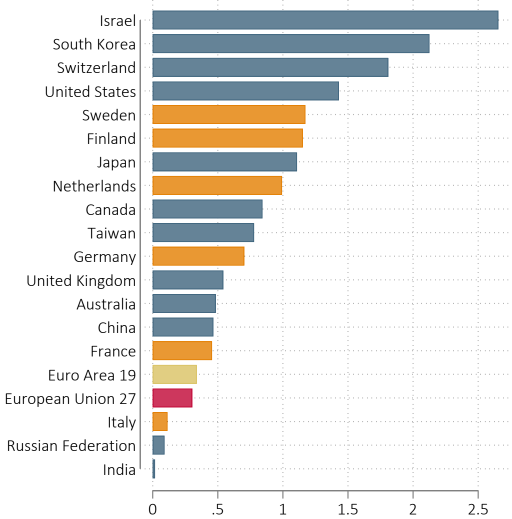
Figure 2: Country frequencies in AI-related patents and publications



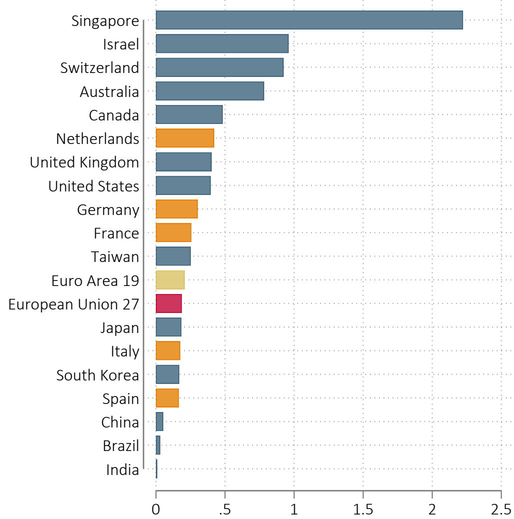
(a) Patents



(b) Publications



(c) Patents per capita



(d) Publications per capita

Sources: EPO PATSTAT (Autumn 2023 edition) and Scopus (2023 edition). The number of inhabitants per country is derived from the Penn World Tables version 10 (Feenstra et al. 2015). Number of patents and of publications per million inhabitants. Authors' own calculations.

The evidence can be read through the critical take of Dosi et al. (2006) on the European paradox. Traditionally, the paradox describes the gap between European frontier science and its sub-optimal industrial application. The term is often used to highlight failures in technology transfer and commercialization when compared to the US. However, our results suggest that in the case of AI the paradox may be more severe than originally conceived: the EU under-performs — relative to the US — both in patent and publication production. The take-home message is that the EU gap with the frontier is both science and innovation-based, rather than only innovation-based. Hence, the quest for improving AI competences is a transversal matter encompassing science, technology, and industrial policies.

Another element to consider is the long-term impact of this gap. As knowledge is for a large share cumulative, a lower accumulation of inventions compared to other areas of the world might turn into a persistent disadvantage. If a critical mass of knowledge production is needed to improve competitiveness and catch-up with the frontier, Europe might never be able to fill the gap formed over the decades.

4.2 Specialization in AI TFA

As we are interested in mapping competences to innovate in AI in a global context increasingly characterized by geopolitical rivalries and fragmentation, we compute the degree of specialization in AI TFA across the three major competing blocks: the EU, the US, and China. Specialization indicates the presence of a (relative) particular expertise and, hence, it is a measure of the decision to allocate resources among a portfolio of possible destinations. Specialization in sciences or in techniques has been the focus of attention of scholars in the literature on technical change, whether in firms (Cantwell 1989, Dibiaggio & Nesta 2005) or countries (Nesta & Patel 2004). One measure which has been used extensively is the so-called relative specialization index (RSA). In our context, the metric measures the share of a domain's AI patents in all AI patents for a given country relatively to the same share for the rest of the world.

The measure is constructed as follows: let $P_{c,d}$ be the number of patents held by country c in AI domain d, representing alternatively the technique t, the function f, or the application a domains. For a given year, the relative specialization advantage RSA index is defined by:

$$RSA_{c,d} = \frac{P_{c,d}/\Sigma_d P_{c,d}}{\Sigma_{b\neq c} P_{c,d}/\Sigma_{b\neq c}\Sigma_d P_{c,d}}, \tag{1}$$

where $d \in \{t, f, a\}$, $t = \{1, 2, \cdots, t, \cdots, T\}$, $f \in \{1, 2, \cdots, f, \cdots, F\}$, and $a = \{1, 2, \cdots, a, \cdots, A\}$.

The RSA index belongs to the zero-infinity interval (i.e. $nRSA \in [0\ ; +\infty[$), and its pivotal value is unity. Without altering the interpretation of the indicator, and in order to facilitate the visualization of the results, we normalized the index as follows:

$$nRSA_{c,d} = \frac{RSA_{c,d} - 1}{RSA_{c,d} + 1}, \tag{2}$$

where $nRSA_{c,d} \in [-1\ ; +1[$, with a threshold value of nullity indicating whether a country enjoys a relative specialization advantage ($nRSA > 0$) or disadvantage ($nRSA < 0$). It is important to stress that the RSA is a relative indicator, and therefore it does not provide information about absolute amounts: a country can be highly specialized without holding a large number of patents.

Figure 3 compares the specialization profile of the three macro blocks, respectively for techniques, functions, and applications. The broad insight that can be derived is that the EU does not display any specific specialization across the whole stylized AI value chain relatively to US and China. Impressive enough, China is characterized by high specialization in many, different techniques, functions and applications. Focusing on applications, China outperforms the other macro blocks on Telecommunications, Industry and Manufacturing, and Agriculture, while the US are relatively more specialized in Personal Devices, Computing and HCI, and Cybersecurity. The lack of European specialization is the result of individual EU countries not exhibiting clear patterns of specialization. In other words, there is no

Ricardian specialization in AI innovation across European member states. This fact can provide a policy opportunity: through coordination and support, the EU as a whole has a great deal of room for action to steer the direction of AI development towards specific areas.

4.3 Complementarity and integration across AI TFA

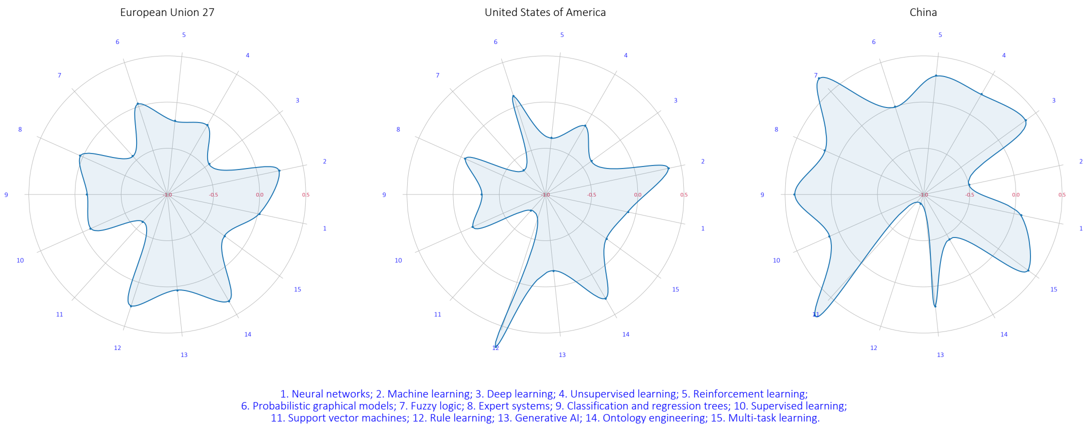
Our conception of technological sovereignty is rooted in the idea that the portfolio of competences of countries must be complementary to one another in order to yield services that cannot be reduced to their independent use. Applied to the AI TFA framework, sovereignty can be measured as the aggregate level of complementarity between the various AI domains of expertise in AI techniques, functions, and applications.

We exploit the fact that a single patent can be jointly assigned to techniques, functions, and applications. For example, a patent might use the techniques of "Probabilistic graphical models" to produce "Computer vision" functions for the application "Transportation", and doing so it signals a consistent, coherent (Nesta & Saviotti 2005) value chain.[11]

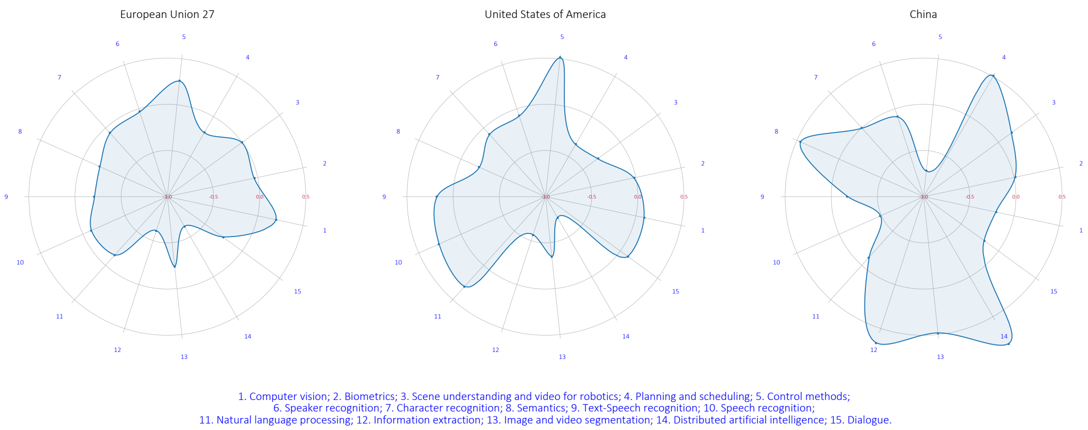
Our goal is to develop a statistical measure of complementarity that exploits joint frequencies. We assume that combinations of techniques and functions, and of functions and applications that are more productive are more complementary. We also maintain that they will occur more frequently than less productive ones. Hence, we first counted the frequency of joint occurrences of techniques and functions (what we call TF co-occurrences), and of functions and applications (what we call FA co-occurrences). We then compared the observed TF and FA joint frequencies with their expected ones, should such joint frequencies occur randomly.

[11] More specifically, if we combined the techniques, functions, and applications randomly, the number of possible combinations to be analyzed would be extensive (23 techniques combined with the 27 functions might yield 22 applications, giving rise to more than 13,000 possible value chains). Therefore, we claim that the actual combinations are meaningful and suggest an order led by synergies.

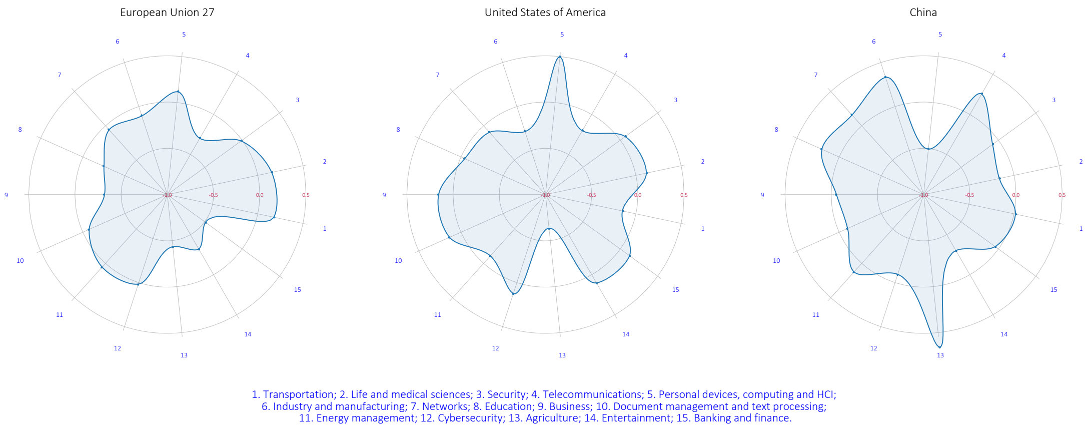
Figure 3: Normalized RSA in AI techniques, functions and applications



(a) techniques



(b) Functions



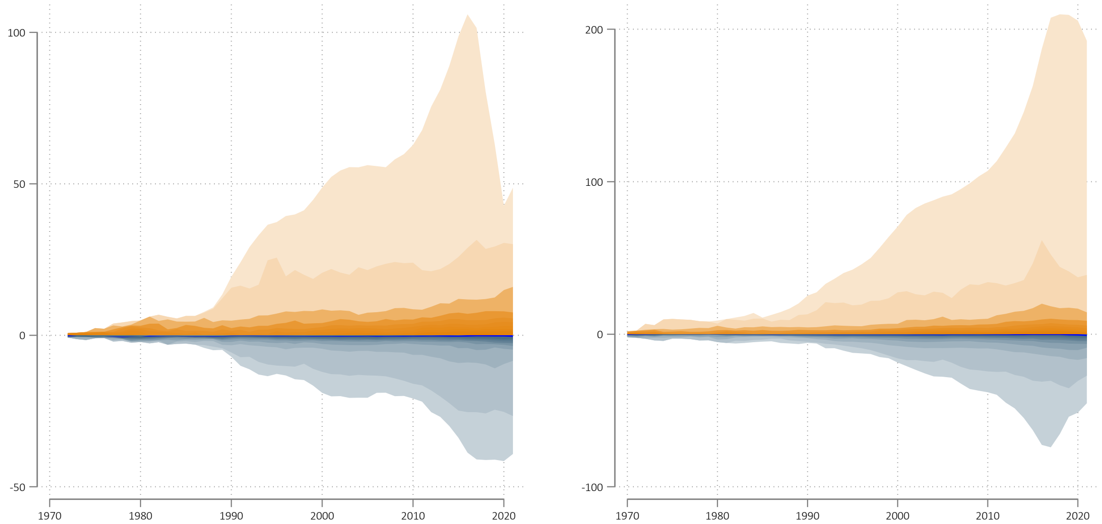
(c) Applications

An observed number of joint frequencies greater than their expected value reveals a positive association, a "mutual" attraction, or, one could say, a complementarity between techniques and functions (alternatively, between functions and applications).

Conversely, should the expected frequencies exceed the observed ones, we would conclude that the two AI domains exclude one another, and hence are not complementary to one another. We called the resulting measure τ_{ij}, where $ij = \{TF, FA\}$. We explain the derivation of the measure in details in Appendix C.

Figure 4 displays the evolution of the distribution of τ_{tf} (panel a) and τ_{fa} (Panel b) over the period of 1970 to 2020. Overall, the two panels display a strikingly similar pattern in which, initially, complementarities, and their lack thereof, are poorly defined. This pattern spans more than two decades, from 1970 to 1990. This initial period corresponds to the phase in which ICT technologies became increasingly pervasive in productive activities.

Figure 4: The dynamics of complementarity



(a) Techniques-Functions (TF) τ (b) Functions-Applications (FA) τ

These graphs display the distribution of complementarity measures τ over time, from the minimum to the maximum measures, and by darkening each every fifth percentile towards the median. Orange (resp. blue) colors depict positive (resp. negative) complementarity measures τ. Source: EPO PATSTAT (Ed. Autumn 2023). Authors' own calculations.

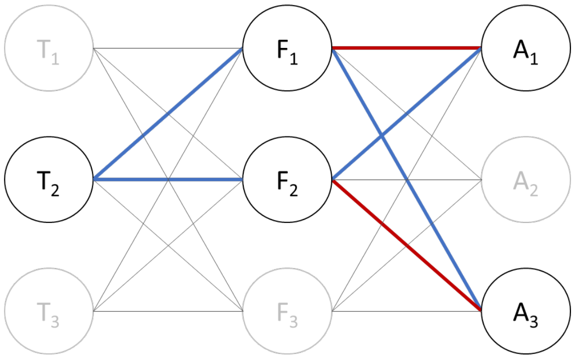
From the early nineties to the late 2010s, we observe an increase in the variability of complementarities. Significant positive ones grow and, as in a mirror, negative ones become clearer. This process exemplifies the fact that complementarities between AI techniques and AI functions, and between AI functions and AI applications, become gradually identified, and others are ruled out as a result of experimentation. This second phase matches the systematization of AI developments in scientific and applied fields. During this period AI continues to advance despite experiencing one of its "Winters". Starting in the early 2010s, the last phase corresponds to the rise of deep learning-driven AI as a well-bounded technology. It builds on the access to larger datasets and better computational capabilities. These are the two conditions needed for AI algorithms to be trained and expanded for a variety of potential uses. Thus, it not surprising to witness an increase in the range of the distribution, reaching very high positive and negative values. These results indicate that the TFA landscape is consolidating around better-identified sets of techniques, functions, and applications, and determinations about how to combine them in a way that yields services that cannot be reduced to their independent usage.

The information contained in patents and their breakdown into AI techniques, functions, and applications can be used to characterize the degree of integration of the AI innovation value chain. As we interpret technological sovereignty in AI as the capacity to mobilize local AI-related competences to develop AI-related innovations: an actor will exhibit a degree of integration when it masters the competences that appear to be complementary in the AI value chain. There are two ideas in this intuitive definition. First, actors must exhibit specialization in some AI-related areas, whether technical, functional, or application-related. Second, these exhibited levels of specialization between techniques and functions, and between functions and applications, must be complementary. Given this setting, and abstracting from the country and time indexes, we measured the overall TFA integration for a single application domain Γ_{TFA} as in the following:

$$\Gamma_{TFA,a} = \sum_{t \in T} \tau_{tf} \times \alpha_t \times \xi_t + \sum_{f \in F} \tau_{fa} \times \alpha_f \times \xi_f \tag{3}$$

where τ_{tf} and τ_{fa} are defined as in Equations C3 and C6, respectively. Variables α_t and α_f represent shares of techniques and functions in overall patents, i.e. $\alpha_t = P_t / \sum_t P_t$ and $\alpha_f = P_f / \sum_f P_f$, respectively. Last, variables ξ_t and ξ_f represent indicator variables, taking value 1 if the normalized value of RSA in technique t and function f are positive, 0 otherwise, i.e. $\xi_t = \mathbb{1}(RSA_t > 0)$ and $\xi_f = \mathbb{1}(RSA_f > 0)$, respectively.

Figure 5: AI innovation value chain for a fictitious country



To better understand the spirit of the measure, Figure 5 represents the AI innovation value chain with three AI techniques, three AI functions, and three AI applications.[12] Now let us imagine a country specializing in technique T2, functions F1 and F2, and applications A1 and A3, implying that $\xi_{T1} = 1$, $\xi_{F1}, \xi_{F2} = 1$, and last $\xi_{A1} = 1$ and $\xi_{A3} = 1$. The edges between the vertices represent the degree of complementarity between the techniques, functions, and applications (τ_{tf} and τ_{fa}). Edges in bold represent complementarities that are relevant for this country because they correspond to the revealed areas of specialization. As Figure 5 indicates, there is a positive association between technique T2 and functions F1 and F2 (blue edges). In addition, there is a negative association between function F1 and application A1 (red edges), but a positive association with function A3 (blue edges), unlike function F2. Overall, the degree of integration is the sum of the observed complementarities (the bold edges) linking the vertices corresponding to areas of AI specialization. This degree of integration can be either positive or negative, depending on whether countries specialize in areas that complement or exclude one another. We interpret this measure as indicating the complementarity between the TFA domains. In other

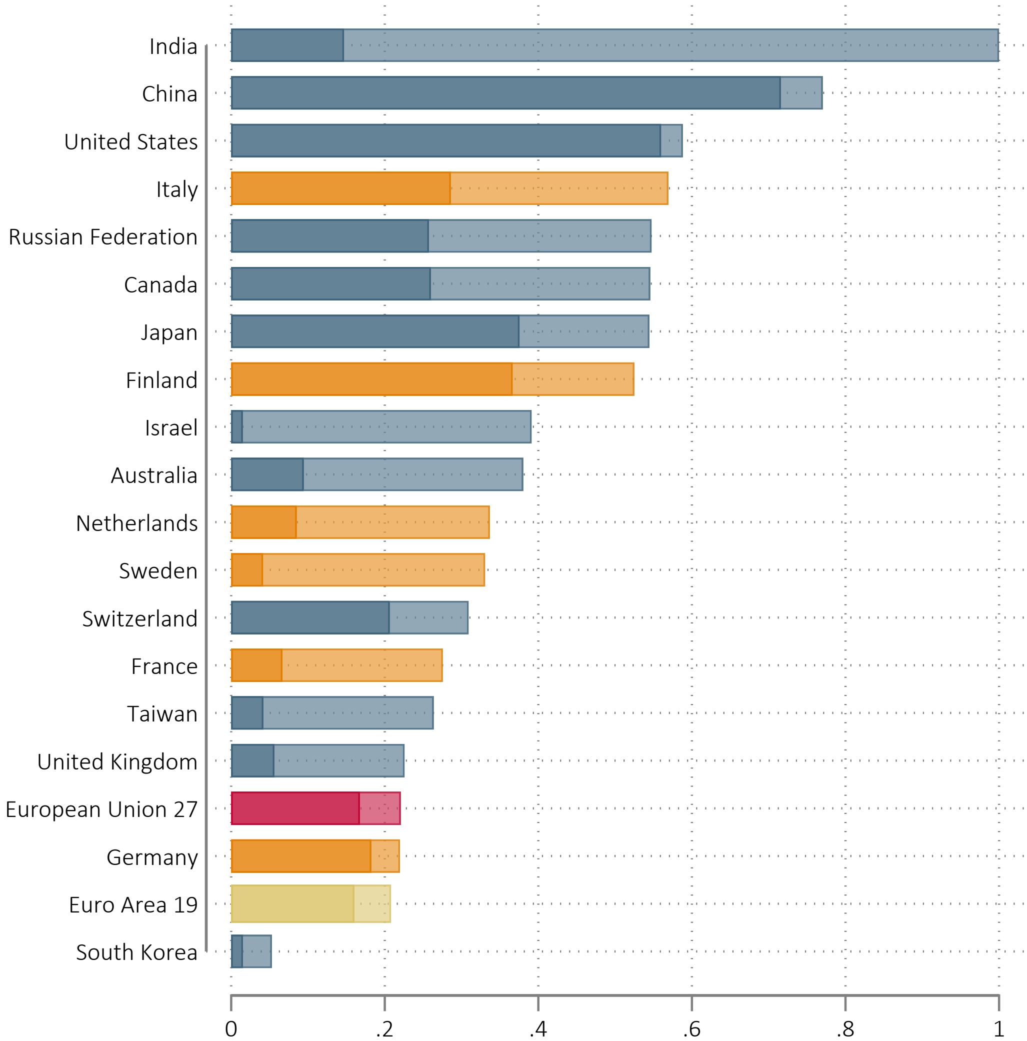
[12] The result is 27 possible technique-function-application chains with the total number of possible chains amounting to 1,000.

words, the degree of integration is simply the sum of the degrees of complementarity observed (between techniques and functions, then between functions and applications) along the possible chains. Observe that our measure of integration can be decomposed into two parts such that $\Gamma_{TFA} = \Gamma_{TF} + \Gamma_{FA}$. Doing so improves our ability to determine whether the locus of integration is located more in upstream integrations (Γ_{TF}) or downstream integrations (Γ_{FA}).

Figure 6 displays the levels of integration among the top patenting countries over the TFA AI value chain, do distinguish between TF integration and FA integration. As the figure indicates, Europe, whether the EU or the EZ, exhibits one of the lowest levels of integration. The United States and China belong to the other end of the spectrum, with values of integration reaching 60% and 80% of the highest value in the dataset belonging to India. Within Europe, Italy displays the highest level of integration, together with countries such as Finland, Sweden, and the Netherlands. France and Germany have low levels of integration and drive the overall poor performance of Europe regarding integration. Contrary to the United States and China, Europe exhibits low levels of integration. Within Europe, Italy displays the highest level of integration, together with countries such as Finland, Sweden, and the Netherlands. France and Germany have low levels of integration.

Another appealing feature of our measure of integration is that it can be decomposed by AI application fields, offering a diagnosis on whether a given AI application domain rests upon an integrated value chain. Table 1 provides integration scores by top AI application specializations, by geo-political blocks. We have three main observations. First, all integration scores are positive. This result implies that all value chains display positive complementarities on average, although some connections throughout the AI value chain may well be negative. Second, the locus of integration may vary a great deal, whether we consider upstream (Γ_{TF}) or downstream (Γ_{FA}) integration. For example concerning "Transportation" in Europe, the locus of integration is clearly located in the functions to applications complementarities,

Figure 6: Mean TFA Integration across countries (1990-2021)



The overall bar represents the value of integration over the Technique-Function-Application value chain (TFA integration). The darkened left hand side of the bar represents the value of integration over techniques and functions - TF integration). The light-color right-hand-side of the bar represents the value of integration over functions and applications (FA integration). Mean values are normalized to 1 for the frontier value (India). Source: EPO PATSTAT (Ed. Autumn 2023). Authors' calculations.

as previously observed. In contrast, the complementarities between techniques and functions are very poor. A similar pattern is evident concerning "Energy management" in Europe, "Personal devices, computing and HCI" in the US, and "Industry and Manufacturing" in China. Conversely, the locus of integration is located upstream in "Cybersecurity" in the US, and to a lesser extent in "Agriculture" in China. All other areas have a somewhat more balanced pattern in which there is integration throughout the entire value chain. Third, as exemplified by "Personal devices, computing and HCI" or by "Cybersecurity", two countries with significant specializations can exhibit different integration patterns. For example, with regard to " Cybersecurity", integration in Europe is more balanced throughout the TFA value chain than in the US, where integration is essentially upstream (TF integration). In a similar fashion, with regard to "Personal devices, computing and HCI", whereas integration in Europe is more balanced, that of the US leans more towards downstream complementarities (FA integration). Our interpretation is that this heterogeneity conceals local systems of innovation throughout the AI value chain involving specific public and private actors and specific sets of collaborations and interactions.

Table 1: Integration score by top AI application specializations, by geographic area

AI application	Γ_{TFA}	Γ_{TF}	Γ_{FA}
		Europe	
Transportation	0.499	0.005	0.494
Life and medical sciences	0.334	0.067	0.267
Personal devices, computing and HCI	0.287	0.135	0.152
Energy management	0.266	0.067	0.199
Cybersecurity	0.295	0.125	0.169
		United States of America	
Personal devices, computing and HCI	0.330	0.064	0.265
Business	0.296	0.148	0.148
Document management and text processing	0.313	0.137	0.176
Banking and finance	0.296	0.157	0.140
Cybersecurity	0.302	0.214	0.088
		China	
Agriculture	0.327	0.216	0.111
Industry and manufacturing	0.370	0.069	0.301
Education	0.303	0.124	0.179
Networks	0.332	0.177	0.155
Telecommunications	0.306	0.136	0.170

Period 2011-2021. See equation 4 for details about the Γ index. TFA: techniques-functions-application integration; TF: techniques-functions integration; TF: functions-application integration. Source: PATSTAT Autumn 2023 Edition. Calculations of the Authors.

5 Technological sovereignty and innovation

5.1 Is integration a source of innovation?

What remains to be tested is whether and to what extent technological sovereignty in AI, proxied by TFA complementarity and integration, does matter for AI innovation. We estimate a Cobb-Douglas patent production function whereby new patents in a given area of AI applications a stem from the relative specialization in AI application a ($nRSA_a$), the existing stock of AI-related knowledge stock in patents and publication, measures of the concentration of patents across techniques, functions, and applications, and of course, integration as measured in Equation 3.

Abstracting from subscript c accounting for country c, the model reads as follows:

$$k_{a,1} = AK_0^{\beta_K} S_0^{\beta_S} \mathbf{C_0}^{\mathbf{B_C}} \exp(\mathbf{B_Z Z_0} + \upsilon_{a,1}), \tag{4}$$

where k_a represent innovation in AI application a, and K and S represent overall patent and publication stocks (irrespective of the application domain a). Subscripts 0 and 1 indicate the timing of innovation, whereby additional patents in 1 come from existing stocks at the beginning of the period (hence, period 0). We forward the dependent variable one year to avoid any spurious correlation between the dependent variable and the vector of explanatory variables. We decompose the disturbance term $\upsilon_{a,1}$ into a year specific effect controlling for common shocks across countries, a country-application fixed effect to control for unobserved but stable differences between country-domains of application, and an *iid* disturbance term such that, respectively: $\upsilon_{a,1} = \kappa_y + \iota_c + \varepsilon$.

Knowledge stocks, whether using patents or publications, are measured using the permanent inventory method whereby new patents feed an existing stock of past patents given a rate of obsolescence ϱ – set to 15% – such that $K_t = (1-\varrho)K_{t-1} + k_t$, where k_t are new patents (when computing the patent stock) or new publications (when computing the publication stock).

Vector Z includes the variables of interest: the level of expertise E and the level of integration Γ, both being specific to application a, so that $\mathbf{B_Z Z} = \beta_E E_{a,0} + \beta_\Gamma \Gamma_{a,0}$. What we call the level of expertise E is the relative specialization advantage RSA in application a. Integration is measured as in Equation 3, and reflects the complementarity of the value chain between the various domains of techniques and functions with application a. Finally, vector $\mathbf{C}$ represents a vector of controls, namely, population and GDP per capita to control for both country size and wealth. We augment vector $\mathbf{C}$ with the various measures of concentration HHI to control for the concentration of expertise across techniques, functions, and applications. We also include a variable "Openness" to control for international interactions between the national innovation system and other countries.[13]

Taking the log-transformation of Equation 4 allows us to estimate the coefficient using least squares estimation methods.

Our intuition is that countries endowed with more integrated AI TFA value chains will have AI competences across the board, and therefore will be better equipped to produce AI innovations; in other words integration should support the production of new patents in AI applications. Given that knowledge creation draws on knowledge stocks, we expect the coefficient associated with patent stocks and publication stocks to be positive. We also expect the coefficient for the degree of expertise to be positive, implying that specialization in a given domain of applications has a positive effect on the creation of future innovations.

Table 2 provides the results of specification 4. We introduce the variables of interest sequentially, with the results appearing in Columns (1) to (4). In Column (1), we introduce the main control variables of knowledge stocks (patent and publication stocks), together with the normalized specialization index ($nRSA$). An important

[13]This variable is computed as the share of co-patents with foreign institutions over the overall number of patents for country c, relative to (i.e. divided by) the same share pertaining to all other countries.

difference between $nRSA$ and the knowledge stock variables is that the latter do not pertain to AI applications specifically. Hence, their parameter estimates can be interpreted as the effects of knowledge capital in science and technology in general on the generation of AI-related patents. Instead, specialization is AI application specific, so that its parameter estimate must be interpreted as the effect of expertise in the given application on the generation of future innovations.

Not surprisingly, all parameters are positive and significant, implying that the level of expertise in AI applications and overall knowledge stocks are key ingredients of future innovation in AI-related patents. Regarding specialization ($nRSA_a$), a 1% increase in $nRSA$ leads to a 46% increase in patent generation. By the same token, a 1% increase in overall patent stocks leads to a 0.46% in AI-related patents in specific applications. The significance of publication stocks in patent generation corroborates the idea that innovation in AI is science-based. Hence, a 1% increase in publication stocks leads to a 0.1% increase in AI-related patents. Finally, the parameter estimates of the patent stocks is more than three times as large as that of the publication stocks. This result is in line with the idea that experience in patenting matters for future patent generation. Beyond experience, publications and patents do not necessarily come from the same institutions. For instance, universities and public research institutions may focus their effort on publishing much more than patenting. Similarly, while the number of private companies involved in scientific research is limited, the number of private firms involved in patenting is much larger. This factor might affect the relationship between publication stocks and patenting output. These aspects cannot be accounted for by looking at scientific capabilities only, as evidenced in the publication stocks. These conclusions hold for all models (Columns 1-4) displayed in Table 2, given the stability of the parameter estimates.

A major finding that corroborates our competence perspective on technological sovereignty is the significant and positive effect of integration on patenting. There are a number of reasons for this result. First, as suggested earlier, when the neces-

Table 2: TFA local integration and the production of quality-weighted innovation

	(1)	(2)	(3)	(4)
$nRSA_a$	0.464***	0.452***	0.442***	0.445***
	(0.056)	(0.056)	(0.056)	(0.056)
Patent Stock (ln)	0.359***	0.360***	0.291***	0.294***
	(0.031)	(0.031)	(0.031)	(0.031)
Publication Stock (ln)	0.107***	0.106***	0.105***	0.102***
	(0.032)	(0.032)	(0.032)	(0.032)
TFA Integration (Γ_{TFA})		0.022**	0.024**	0.074***
		(0.010)	(0.009)	(0.021)
Openness			-0.262***	-0.260***
			(0.035)	(0.035)
TFA Integration $\times$ Openness				-0.023***
				(0.008)
T Herfindahl (patents)	-1.431***	-1.430***	-1.210***	-1.188***
	(0.442)	(0.443)	(0.423)	(0.425)
F Herfindahl (patents)	2.128***	2.181***	1.449***	1.538***
	(0.461)	(0.463)	(0.461)	(0.463)
T Herfindahl (publications)	0.750**	0.745**	0.892***	0.884***
	(0.303)	(0.302)	(0.290)	(0.289)
F Herfindahl (publications)	0.758***	0.750***	0.708***	0.706***
	(0.163)	(0.163)	(0.160)	(0.159)
Population (ln)	0.037	0.052	0.596**	0.556**
	(0.269)	(0.269)	(0.272)	(0.272)
GDP per capita (ln)	1.425***	1.419***	1.439***	1.434***
	(0.099)	(0.099)	(0.095)	(0.095)
R-squared	0.857	0.857	0.859	0.859
Within R-squared	0.197	0.197	0.206	0.207
Log Likelihood	-7,840	-7,837	-7,791	-7,787
LR test	-	5.44**	91.44***	8.78***

$N = 8,268$. Dependent variable: Quality-weighted number of innovation (number of patents). Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. All regressions include a full vector of unreported year fixed effects and country-field of application fixed effects. Constant is omitted for the sake of clarity. The LR test is carried out comparing the unrestricted model (m) with the restricted model ($m-1$).

sary expertise throughout the value chain is developed domestically, organizations find it easier and cheaper to identify and coordinate their activities rather than searching for similar competences abroad. Second, the result also suggests the existence of local clusters where the sharing of information, of human capital, and their associated positive externalities act as positive ingredients for future innovation (Romer 1990). Another important element is that value chain integration reduces uncertainty, allowing for further investments (Amendola & Gaffard 1998).

Observe that the previous remark holds irrespective of the diversity of the countries' portfolio throughout the AI value chain. More specifically, all models control for the diversity of the countries in terms of AI techniques and functions, whether stemming from technologies (as assessed by patents) or from science (as measured in publication data). It is noteworthy that all concentration measures display positive coefficients, implying that more concentrated investments around key techniques and functions matter. The only exception to this contention relates to techniques as measured in the patent data. Conversely, the diversification of AI techniques in the patent data acts as a positive input for future innovation. These results show that efforts to diversify investments into more AI techniques would contrast lock-in into some AI techniques and be beneficial for innovation in AI applications.

Finally, in Columns (3-4), we introduce Openness, a variable measuring the propensity of the country to engage in international collaborations in patent activity. In Column (3), Openness figures negatively, implying that more international collaborations generate fewer patents pertaining to the country. This result does not mean that international collaborations are detrimental to innovation in AI. Rather, we interpret this result as a confirmation of the importance of local innovation systems. A strong propensity to collaborate with foreign partners reveals a lack of equivalent expertise locally. As explained, searching for partners abroad is costly and less stable relative to relying on local networks of partners. Furthermore, openness also results in a loss of innovative opportunities for local partners throughout the supply chain. Knowledge spillovers and innovation options generated by the collaboration may benefit customers and suppliers in the partner's country.

A key feature of our measure of integration is that it allows us to identify the distinctive role of upstream versus downstream integration. We can rewrite Vector $\mathbf{Z}$ as $\beta_E E_{a,0} + \beta_{\Gamma_{TF}} \Gamma_{TF,a,0} + \beta_{\Gamma_{FA}} \Gamma_{FA,a,0}$. Coefficients $\beta_{\Gamma_{TF}}$ and $\beta_{\Gamma_{FA}}$ and their difference will provide information about the locus of integration as a source of future innovation. Table 3 re-runs the analysis exploiting the possibility of separating the upstream and the downstream integration effects. The signs, magnitude, and significance of the coefficients for the key variables tested ($nRSA$, patent and publication stocks, Openness) are stable compared to the results of the regressions with aggregate TFA integration. In this new setup, both the TF and FA integrations are positive and significant, indicating that AI inventions are enabled both by the alignment of competences between techniques and functions and between functions and applications. Models (7) and (8) consider TF integration based on publications rather than patents. Our goal is to capture the more science-based competences embodied in the technique-functions pairs, and possibly to identify the different actors involved. The coefficient related to TF integration based on publications loses its statistical significance, but re-acquires it when the interaction term with Openness is introduced in specification (8). One way to interpret the result is that TF integration feeds innovation, but only when the competences are developed domestically rather than by sources far from the local context. FA integration based on patents reveals similar insights. Openness affects innovation negatively both overall and when interacted with the integration terms. This result suggests that sourcing knowledge outside the local innovation systems reduces invention incentives and weakens the power of integration to produce new knowledge. All in all, this evidence supports the idea that technological sovereignty can enhance innovative performance.

5.2 The determinants of integration as sovereignty in AI

If integration — and, thus, technological sovereignty — favors innovation in AI, what factors favor integration? With our data, we can explore the organizational origins of integration at a granular level. Doing so illustrates the key modalities through which integration is built, and may also represent an actionable policy lever. In Table 4, we relate TFA integration (Model (9)), TF integration (Model (10)), and

Table 3: Partitioning upstream TF and downstream FA integration and the production of quality-weighted innovation

	(5)	(6)	(7)	(8)
$nRSA_a$	0.441***	0.445***	0.443***	0.444***
	(0.056)	(0.056)	(0.056)	(0.056)
Patent Stock (ln)	0.269***	0.272***	0.290***	0.282***
	(0.031)	(0.031)	(0.031)	(0.031)
Publication Stock (ln)	0.097***	0.093***	0.111***	0.135***
	(0.032)	(0.033)	(0.033)	(0.034)
TF Integration (Γ_{TF})	0.062***	0.069***		
	(0.013)	(0.019)		
FA Integration (Γ_{FA})	0.016*	0.064***	0.017*	0.069***
	(0.009)	(0.020)	(0.009)	(0.021)
Openness	-0.281***	-0.279***	-0.259***	-0.246***
	(0.035)	(0.036)	(0.035)	(0.035)
TF Integration × Openness		-0.004		
		(0.008)		
FA Integration × Openness		-0.022***		-0.024***
		(0.008)		(0.008)
TF Integration (publications)			0.010	0.080***
			(0.010)	(0.018)
TF Integration (pub.) × Openness				-0.031***
				(0.008)
R-squared	0.859	0.859	0.859	0.859
Within R-squared	0.209	0.210	0.206	0.209
Log Likelihood	-7,779	-7,775	-7,792	-7,777
Model Comparison	(5) vs. (3)	(6) vs. (4)	None	(8) vs. (7)
LR test	24.82***	24.84***	-	30.21***

$N = 8,268$. Dependent variable: Quality-weighted number of innovations (number of patents). Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. All regressions include a full vector of unreported year fixed effects and country-field of application fixed effects. The constant is omitted for the sake of clarity. The vector of control variables includes the series of Herfindahl indexes of AI techniques and AI functions derived from patents and publications, and population and GDP per capita entered in logs. Models (7) and (8) use relatedness measures and shares of techniques and functions derived from publication data.

FA integration (Model (11)). We can distinguish between private and public actors and combinations thereof (private-private, public-public, and public-private collaborations). Indicators of Openness and knowledge stocks (patents and publications) are included as well. With regard to TFA integration (Model (9)), collaborations among private actors are related to more integration. Private actors seem to play a positive role in enhancing TF integration, with public assignees being characterized by a significant but negative coefficient. FA integration is positively related to the presence of public actors. The stock of patents is significantly related to integration in all specifications, negatively for TFA and FA and positively for TF. These results suggest that prior knowledge is important for connecting complementary competences at the more technological level, while it hinders integration at the more market-proximate layer of the value chain. Interestingly, Openness has a positive and significant effect on TFA, TF, and FA integration. A possible interpretation of this result, especially when compared to the innovation analysis, is that an "open first, closed then" strategy might be at work. Local actors can develop or diversify their competences in AI by interacting with international partners. Once the competences are formed, local integration favors the production of new knowledge. In a nutshell, AI innovators trade openness at the competence-development stage for less openness at the innovation stage.

6 Conclusion

We have focused on artificial intelligence as one of the technologies driving a global "arms' race" between countries and geopolitical blocks to achieve technological sovereignty. As the international landscape becomes increasingly fragmented and multi-polar, the question of whether an actor is *capable* to innovate in AI is pivotal to understanding future dynamics of growth and competitiveness. This is particularly important for the European Union –a global player in terms of market size but a laggard in investments in digital and emerging technologies.

We posit that addressing the competitiveness challenges posed by global rivalries does not require dismantling globalization, but rather improving domestic capabilities and resilience. For this reason, we have focused on the competence dimension of technological sovereignty. Based on patent and publication data, we developed a

Table 4: The determinants of integration as sovereignty in AI

	(9)	(10)	(11)
Private assignee (ln)	0.038	0.280***	0.004
	(0.049)	(0.039)	(0.049)
Public assignee (ln)	0.066**	-0.093***	0.080**
	(0.033)	(0.025)	(0.033)
Private-Private coll. (ln)	0.153***	0.388***	0.105***
	(0.034)	(0.025)	(0.034)
Public-Public coll. (ln)	-0.008	-0.083***	0.002
	(0.018)	(0.016)	(0.018)
Public-Private coll. (ln)	-0.028	0.033*	-0.035
	(0.022)	(0.017)	(0.022)
Openness	0.160***	0.340***	0.120***
	(0.040)	(0.030)	(0.040)
Patent Stock (ln)	-0.192***	0.113***	-0.204***
	(0.041)	(0.038)	(0.041)
Publication Stock (ln)	0.038	-0.007	0.043
	(0.036)	(0.028)	(0.036)
R-squared	0.451	0.632	0.441
Within R-squared	0.010	0.103	0.080
Log Likelihood	-9,263	-7,608	-9,335

$N = 8,268$. Dependent variable: TFA integration in model (9). TF integration in model (10). FA integration in model (11). Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. All regressions include a full vector of unreported year fixed effects and country-field of application fixed effects. The vector of control variables includes patent stocks, publication stock, population and GDP per capita, all entered in logs. Constant is omitted for the sake of clarity.

measure of integration that we argue serves as a proxy for technological sovereignty. This measure captures the integration of countries' AI innovation across a stylized value chain comprising techniques, functions, and applications, and is then applied in a series of econometric analyses. Through this approach, we have linked integration to the innovative performance of countries in AI and further unpacked the integration measure into more fundamental determinants. Our results show that integration is key for innovation in AI.

As the EU reveals the least integrated among global blocks, its gap with the international technological and competitiveness frontier risks widening further and becoming permanent. At the same time, the scope for improvement is substantial, and the lack of integration in European AI should be seen as a call to policy action. We argue that European policies to enhance technological sovereignty should pursue two non-exclusive avenues. On the one hand, significant public investment programs, as recommended by Aghion et al. (2024) in the context of France and by Draghi (2024) for the European Union as a whole, are essential to addressing the investment gap. The true challenge lies in stimulating private investments in AI, a task particularly difficult for countries with limited representation among global AI leaders. On the other hand, efforts in developing a common understanding of the directionality of investments, for instance by allocating scientific and technological funding to areas with high returns (Fuest et al. 2024), represents another fundamental challenge. Increasing integration in EU AI will also depend on some form of continental coordination between European actors to facilitate the formation of a fully-fledged, Europe-wide AI industry.

References

Acemoglu, D. (2024), ‘The simple macroeconomics of AI’, *NBER Working Paper* (w32487).

Acemoglu, D., Akcigit, U. & Kerr, W. R. (2016), ‘Innovation network’, *Proceedings of the National Academy of Sciences* **113**(41), 11483–11488.

Aghion, P., Bouverot, A., Amabile, A., Cyprien, C. et al. (2024), Ia: Notre ambition pour la france, Technical report, Commission de l’Intelligence Artificielle, Government of France.

Aghion, P., David, P. A. & Foray, D. (2009), ‘Science, technology and innovation for economic growth: linking policy research and practice in ‘stig systems”, *Research policy* **38**(4), 681–693.

Aghion, P. & Griffith, R. (2005), *Competition and growth: Reconciling theory and evidence*, MIT Press.

Agrawal, A., Gans, J. & Goldfarb, A. (2022), *Prediction Machines, Updated and Expanded: The Simple Economics of Artificial Intelligence*, Harvard Business Press.

Ahmadpoor, M. & Jones, B. F. (2017), ‘The dual frontier: Patented inventions and prior scientific advance’, *Science* **357**(6351), 583–587.

Ahmed, N., Wahed, M. & Thompson, N. C. (2023), ‘The growing influence of industry in ai research’, *Science* **379**(6635), 884–886.

Amendola, M. & Gaffard, J.-L. (1998), *Out of Equilibrium*, OUP Catalogue, Oxford University Press.

Archibugio, D. & Pianta, M. (1992), *The Technological Specialization of Advanced Countries: A Report to the EEC on International Science and Technology Activities*, Springer Science & Business Media.

Arthur, W. B. (1994), *Increasing returns and path dependence in the economy*, University of Michigan Press.

Baruffaldi, S., van Beuzekom, B., Dernis, H., Harhoff, D., Rao, N., Rosenfeld, D. & Squicciarini, M. (2020), 'Identifying and measuring developments in artificial intelligence'.

Bock, S., Elewa, A., Guillou, S., Napoletano, M., Nesta, L., Salies, E. & Treibich, T. (2024), 'Le décrochage européen en question', *OFCE Policy Brief* (128), 2–22.
URL: *https://sciencespo.hal.science/hal-04585245*

Breschi, S., Malerba, F. et al. (2005), 'Clusters, networks, and innovation: Research results and new directions', *Clusters, networks, and innovation* pp. 1–26.

Bresnahan, T. F. (2003), 'The mechanisms of information technology's contribution to economic growth', *Institutions, innovation and growth: selected economic papers* p. 116.

Bresnahan, T. F. & Trajtenberg, M. (1995), 'General purpose technologies 'engines of growth'?', *Journal of Econometrics* **65**(1), 83–108.

Bryson, J. J. & Malikova, H. (2021), 'Is there an AI cold war?', *Global Perspectives* **2**(1), 24803.

Cantwell, J. (1989), *Technological innovation and multinational corporations*, Blackwell.

Cockburn, I. M., Henderson, R. & Stern, S. (2018), *The impact of artificial intelligence on innovation*, Vol. 24449, National bureau of economic research Cambridge, MA, USA.

Corea, F. (2019), *Applied artificial intelligence: Where AI can be used in business*, Vol. 1, Springer.

David, P. A. (1985), 'Clio and the economics of qwerty', *The American Economic Review* **75**(2), 332–337.

Dibiaggio, L., Keita, M. & Nesta, L. (2022), 'Artificial intelligence: Technologies and key players', *Sophia Antipolis, France: OTESIA Reports, SKEMA Business School* .

Dibiaggio, L. & Nesta, L. (2005), 'Patents statistics, knowledge specialisation and the organisation of competencies', *Revue d'économie industrielle* **110**(1), 103–126.

Ding, J. & Dafoe, A. (2021), 'The logic of strategic assets: From oil to AI', *Security Studies* **30**(2), 182–212.

Dosi, G., Llerena, P. & Labini, M. S. (2006), 'The relationships between science, technologies and their industrial exploitation: An illustration through the myths and realities of the so-called 'european paradox'', *Research Policy* **35**(10), 1450–1464. Triple helix Indicators of Knowledge-Based Innovation Systems.

Draghi, M. (2024), The future of european competitiveness, Technical report, European Commission.

Edler, J., Blind, K., Kroll, H. & Schubert, T. (2023), 'Technology sovereignty as an emerging frame for innovation policy. defining rationales, ends and means', *Research Policy* **52**(6), 104765.

Edquist, C. & Johnson, B. (1996), *Institutions and organizations in systems of innovation*, Univ.

Edquist, C. & Lundvall, B.-A. (1993), 'Comparing the danish and swedish systems of innovation', *National innovation systems: A comparative analysis* pp. 265–298.

Fagerberg, J., Mowery, D. C. & Nelson, R. R., eds (2006), *The Oxford Handbook of Innovation*, Oxford University Press, Oxford, UK.

Feenstra, R. C., Inklaar, R. & Timmer, M. P. (2015), 'The next generation of the penn world table', *American Economic Review* **105**(10), 3150–3182.

Fontana, O. & Vannuccini, S. (2024), How to institutionalise european industrial policy (for strategic autonomy and the green transition), Technical report, Research Paper of the Centro Studi sul Federalismo, Turin.

Franco, A. M., Sarkar, M., Agarwal, R. & Echambadi, R. (2009), 'Swift and smart: the moderating effects of technological capabilities on the market pioneering–firm survival relationship', *Management Science* **55**(11), 1842–1860.

Freeman, C. (1987), 'Technology policy and economic policy: Lessons from japan', *London: Pinter* .

Fronzetti Colladon, A., Guardabascio, B. & Venturini, F. (2025), 'A new mapping of technological interdependence', *Research Policy* **54**(1), 105126.
URL: *https://doi.org/10.1016/j.respol.2024.105126*

Fuest, C., Gros, D., Mengel, P.-L., Presidente, G. & Tirole, J. J. (2024), Eu innovation policy: How to escape the middle technology trap, Technical report, ifo Institute-Leibniz Institute for Economic Research at the University of University of Munich.

Furman, J. L., Porter, M. E. & Stern, S. (2002), 'The determinants of national innovative capacity', *Research policy* **31**(6), 899–933.

Jacobides, M. G., Brusoni, S. & Prencipe, A. (2009), 'Strategic dynamics in industry architectures and the challenges of knowledge integration', *European Management Review* **6**(4), 209–216.

Jensen, M. B., Johnson, B., Lorenz, E., Lundvall, B.-Å. & Lundvall, B. A. (2007), 'Forms of knowledge and modes of innovation', *The learning economy and the economics of hope* **155**.

Kak, A. & West, S. M. (2024), AI nationalism(s): Global industrial policy approaches to AI, Technical report, AI Now Institute.

Kleimann, D., Poitiers, N., Sapir, A., Tagliapietra, S., Véron, N., Veugelers, R. & Zettelmeyer, J. (2023), 'Green tech race? the us inflation reduction act and the eu net zero industry act', *The World Economy* **46**(12), 3420–3434.

Kline, S. J. & Rosenberg, N. (1986), An overview of innovation, *in* R. Landau & N. Rosenberg, eds, 'The Positive Sum Strategy: Harnessing Technology for Economic Growth', National Academy Press, Washington, DC, pp. 275–305.
URL: *https://nap.nationalacademies.org/read/612/chapter/18*

Klinger, J., Mateos-Garcia, J. & Stathoulopoulos, K. (2020), 'A narrowing of AI research?', *arXiv preprint arXiv:2009.10385* .

Krugman, P. R. (1979), 'Increasing returns, monopolistic competition, and international trade', *Journal of International Economics* **9**(4), 469–479.

Lundvall, B.-A. et al. (1992), *National systems of innovation: towards a theory of innovation and interactive learning*, Vol. 242, London Pinter.

Malerba, F. & Orsenigo, L. (1993), 'Technological regimes and firm bebavior', *Industrial and corporate change* **2**(1), 45–71.

Mansfield, E. (1991), 'Academic research and industrial innovation', *Research Policy* **20**(1), 1–12.

Maraut, S., Dernis, H., Webb, C., Spiezia, V. & Guellec, D. (2008), 'The oecd regpat database: a presentation'.

Mezzanotti, F. & Simcoe, T. (2023), 'Innovation and appropriability: Revisiting the role of intellectual property', *NBER Working Paper* (w31428).

Mowery, D. C. (1992), 'The us national innovation system: origins and prospects for change', *Research Policy* **21**(2), 125–144.

Narin, F., Hamilton, K. S. & Olivastro, D. (1997), 'The increasing linkage between u.s. technology and public science', *Research Policy* **26**(3), 317–330.

Nelson, R. R. (1993), *National innovation systems: A comparative analysis*, Oxford University Press.

Nelson, R. R. & Rosenberg, N. (1993), 'Technical innovation and national systems', *National innovation systems: A comparative analysis* **1**, 3–21.

Nesta, L. (2008), 'Knowledge and productivity in the world's largest manufacturing corporations', *Journal of Economic Behavior & Organization* **67**(3), 886–902.

Nesta, L. & Patel, P. (2004), *National patterns of technology accumulation: use of patent statistics*, Springer, pp. 531–551.

Nesta, L. & Saviotti, P. P. (2005), 'Coherence of the knowledge base and the firms' innovative performance: evidence from the US pharmaceutical industry', *The Journal of Industrial Economics* **53**(1), 123–142.

OECD (1999), *Managing National Innovation Systems*, OECD Publishing, Paris. **URL:** *https://doi.org/10.1787/9789264189416-en*

Ott, I. & Vannuccini, S. (2023), 'Invention in times of global challenges: A text-based study of remote sensing and global public goods', *Economies* **11**(8), 207.

Pavitt, K. (1984), 'Sectoral patterns of technical change: Towards a taxonomy and a theory', *Research Policy* **13**(6), 343–373.

Pichler, A., Lafond, F. & Farmer, J. D. (2020), 'Technological interdependencies predict innovation dynamics', *INET Oxford Working Paper No. 2020-04* . **URL:** *https://www.inet.ox.ac.uk/publications/no-2020-04-technological-interdependencies-predict-innovation-dynamics*

Rikap, C. (2023), 'Intellectual monopolies as a new pattern of innovation and technological regime', *Industrial and Corporate Change* p. dtad077.

Romer, P. M. (1990), 'Endogenous Technological Change', *Journal of Political Economy* **98**(5), 71–102.

Rosenberg, N. (1982), *Inside the black box: technology and economics*, cambridge university press.

Rotolo, D., Hicks, D. & Martin, B. R. (2015), 'What is an emerging technology?', *Research policy* **44**(10), 1827–1843.

Sevilla, J., Heim, L., Ho, A., Besiroglu, T., Hobbhahn, M. & Villalobos, P. (2022), Compute trends across three eras of machine learning, *in* '2022 International Joint Conference on Neural Networks (IJCNN)', IEEE, pp. 1–8.

Sheikh, H., Prins, C. & Schrijvers, E. (2023), Artificial intelligence: definition and background, *in* 'Mission AI: The new system technology', Springer, pp. 15–41.

Timmers, P. (2018), 'Cybersecurity is forcing a rethink of strategic autonomy'.

Vannuccini, S. & Prytkova, E. (2024), 'Artificial intelligence's new clothes? a system technology perspective', *Journal of Information Technology* **39**(2), 317–338.

Veugelers, R. & Wang, J. (2019), 'Scientific novelty and technological impact', *Research Policy* **48**(6), 1362–1372.

WIPO (2019), Wipo technology trends 2019 – artificial intelligence, Technical report.

Appendix A. Identification of patents and publications

In order to identify relevant AI patents, we combined two approaches. On the one hand, we used information from three patent classification systems: the International Patent Classification (IPC), the Cooperative Patent Classification (CPC), and the File Index / File forming terms (FI/F). On the other, we exploited keywords to search for AI technologies in patents' titles and summaries. Keyword-based approaches are increasingly used in order to navigate data in a more explorative and unstructured manner (Ott & Vannuccini 2023, Cockburn et al. 2018) and to go beyond standard classifications.

The protocol for our selection of AI patents is inspired by the methodology developed by WIPO (WIPO 2019), to which we added an additional step. The WIPO methodology consists of three building blocks of data from different selection strategies. Each block builds upon the previous one.

- Step 1: List of CPC codes specific to AI technologies/functions/applications
- Step 2: Specific list of keywords in the titles and summaries of the patents
- Step 3: Specific lists of CPC codes, IPC codes, and FI/F terms controlled by another specific list of keywords

The combination of the three datasets obtained through these steps results in a sample that represents all patents considered potential AI patents. Steps 2 and 3 are based on a search in the abstracts and titles of the patents of the keywords that the WIPO proposed. We also added a number of terms related to development in AI, such as generative AI, that have emerged more recently. Although the majority of patent titles and abstracts are written in English (approximately 80% for titles and approximately 90% for abstracts), some are written in other languages. Given that the keywords in the WIPO list are in English, it is difficult to search through texts written in other languages. Of the 36 languages used, we selected those which, according to the WIPO report, are spoken in countries that play a relatively important role in the development of AI (WIPO 2019). We translated the keywords into the following 11 languages: French, German, Spanish, Portuguese, Italian, Russian, Chinese, Japanese, Korean and Dutch. The last step was to apply Steps 2 and 3 to Japanese patents that do not use a patent classification system based on the CPC or IPC codes. To do so, we first retrieved the AI patents using the Japanese FI/F classification terms. We then performed a full join on the patent IDs in order to retrieve the corresponding IPC and CPC codes.

This procedure allowed us to identify 96% of the patents from the Japanese patent classification. In Step 2 we used the list of keywords from Step 1 to select the patents. The third step was to select a list of patents by IPC, CPC, and FI/F terms, and then filter them using the keyword list in Step 2. Finally, we built our own patent databases by categorizing the patents into the AI TFA categories using an algorithm we developed. We began by classifying a patent into a category/subcategory if the CPC/IPC code allowed it through the WIPO classification. If not, we searched for a series of keywords related to the category and subcategory in the abstract and/or title of the patent. In this way, we built three databases of patents that corresponded to the three categories of AI we considered: techniques, functions, and applications.

We also used this method to classify scientific publications about AI into the three categories of TFA. However, there were three major differences in our approach. First, given that our publications came from conferences devoted to AI, we did not have to determine which publications were relevant. Second, unlike patents, publications are not classified in technology classes (IPC, CPC, FI/F classes). Hence, we relied exclusively on our search for AI-related keywords in the titles and summaries of the publications. Last, we did not consider publications associated with application domains (the "A" in the TFA representation). We made this choice because publications usually focus primarily on advancing knowledge — in our case, introducing new (or advanced) techniques and functions — rather than specific production issues. Therefore, we assumed that scientific publications would be concerned with the development of techniques and functions.

An important issue for the purpose of this paper is to determine the location of the invention (in terms of country). As our interest is to locate and map AI-related competences, we use the country of residence of the inventors identified by their personal address referenced in PATSTAT rather than the country of the IP office.

The issue with using the inventors' country of location is that that information is missing in around 50% of AI patents. To correct that, we proceeded as follows.

1. *PATSTAT*. With the variable "psn_country_code", PATSTAT provides information on the inventor's location. Of 1,580,115 patents, we located the inventor's residency for 783,556 patents.

2. *OECD REGPAT database. January 2024.* REGPAT is an OECD database that provides the location of nearly 19 million patents from PATSTAT (Maraut et al. 2008). The ultimate goal of REGPAT is to link patents to NUTS3 regions, and therefore countries.

3. To complement this approach, we considered patents with only one inventor and a family size of 1 (only one IP office). We also assigned the country of the IP office as the country of invention.

By concatenating the different sources of information, we obtained 1,415,828 patents (93%) with a geographic location. It should be noted that a patent can have several inventors, therefore several locations.

Concerning publications, the information contained in Scopus allowed us to determine the location of the scientists more straightforwardly, using the address of the affiliation of the authors. We obtained 330,362 publications from conference proceedings from 1989 to 2023. As in the case of patents, we did not use weights to allocate publications to countries. If a publication was written by authors from different countries, we counted the publication as many times as there were countries rather than allocating weights to the countries.

Appendix B. List and frequencies of TFA domains.

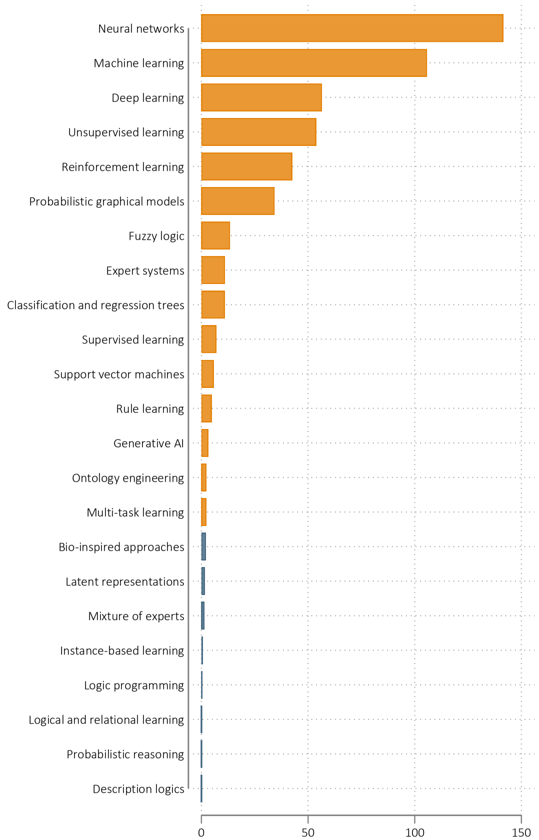
Table B1 presents the full list of techniques and functions used to classify patents and publications in the Technique-Function-Application framework. Figures B1 and B2 display the frequencies of the patent documents and publications, respectively. Not surprisingly, both figures B1 and B2 indicate over-dispersed distributions in the number of patents and publications dedicated to techniques and functions. Figures B1 also show that all applications do not use AI with the same intensity. The fields of transportation, life and medical sciences, security, and telecommunications are clearly dominant in their use of AI in our period of analysis. These differences may affect the estimation and interpretation of specialization in each domain. Becoming an expert in deep learning requires much more investment and resources than acquiring a specialization in fuzzy logic; at the same time, investing in the least crowded technical or functional domains might be a positioning strategy for taking the lead in niche areas, if any returns (scientific or economic) are to be expected in these areas.

Table B1: List of AI techniques, functions, and applications used in this report

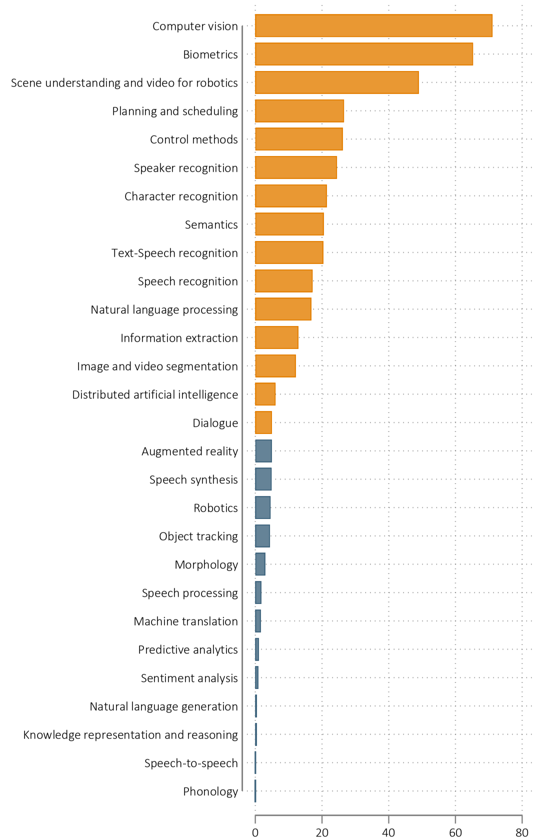
AI Techniques	AI Functions	AI Applications
Bio-inspired approaches	Augmented reality	Agriculture
Classification and regression trees	Biometrics	Arts and humanities
Deep learning	Character recognition	Banking and finance
Description logics	Computer vision	Business
Expert systems	Control methods	Cartography
Fuzzy logic	Dialogue	Computer in government
Generative AI	Distributed artificial intelligence	Cybersecurity
Instance-based learning	Image and video segmentation	Document management and text processing
Latent representations	Information extraction	Education
Logic programming	Knowledge representation and reasoning	Energy management
Logical and relational learning	Machine translation	Entertainment
Machine learning	Morphology	Industry and manufacturing
Mixture of experts	Natural language generation	Law, social and behavioral sciences
Multi-task learning	Natural language processing	Life and medical sciences
Neural networks	Object tracking	Military
Ontology engineering	Phonology	Networks
Probabilistic graphical models	Planning and scheduling	Personal devices, computing and HCI
Probabilistic reasoning	Predictive analytics	Physical sciences and engineering
Reinforcement learning	Robotics	Publishing
Rule learning	Scene understanding and video for robotics	Security
Supervized learning	Semantics	Telecommunications
Support vector machines	Sentiment analysis	Transportation
Unsupervized learning	Speaker recognition	
	Speech processing	
	Speech recognition	
	Speech synthesis	
	Speech-to-speech	
	Text-Speech recognition	

WIPO and author's own elaboration.

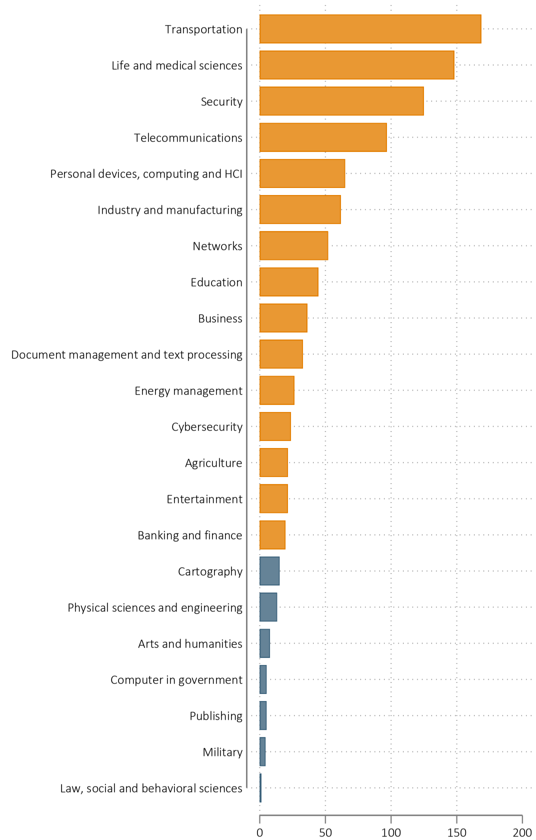
Figure B1: Top patent frequencies in AI techniques, functions, and applications



(a) AI techniques



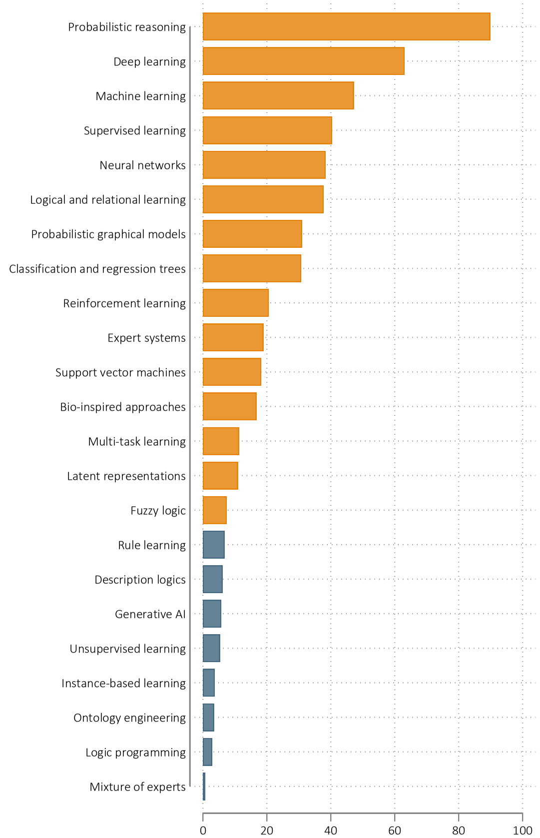
(b) AI functions



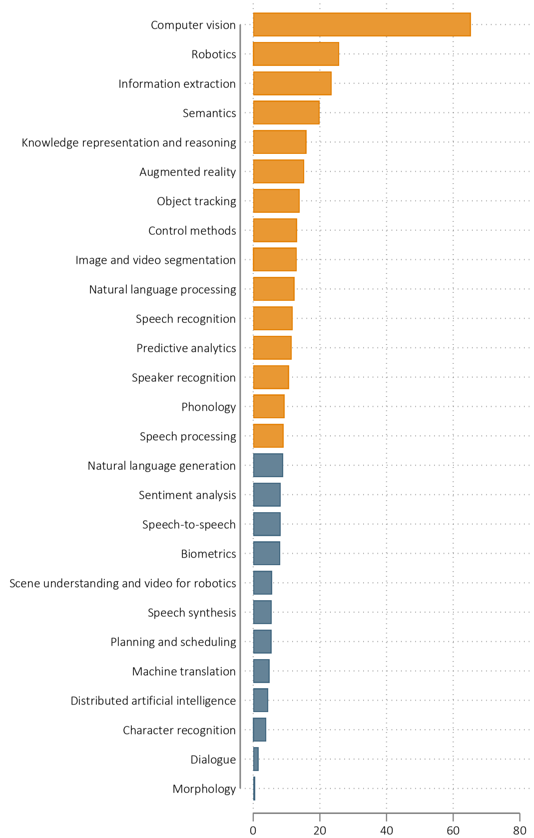
(c) AI applications

Source: EPO PATSTAT (Ed. Autumn 2023). Authors' own calculations. The top 15 frequencies appear in orange. The unit of the horizontal axis is in thousand of patents.

Figure B2: Top publication frequencies in AI techniques and functions



(a) AI techniques



(b) AI functions

Source: SCOPUS data. Authors' own calculations. The top 15 frequencies appear in orange. The unit of the horizontal axis is in thousand of publication.

Appendix C. Measuring Complementarity

To show our approach in details, we follow Nesta (2008) and quantify complementarity between any AI technique $t \in \{1, 2, \cdots, t, \cdots, T\}$ and any AI function $f \in \{1, 2, \cdots, f, \cdots, F\}$. Let the technological universe consist of K patent applications. Let $P_{tk} = 1$ if patent k is assigned to AI technique t, and 0 otherwise. The total number of patents assigned to technique t is thus $O_t = \sum_k P_{tk}$. In the same vein, let $P_{fk} = 1$ if patent k is assigned to AI function f, and 0 otherwise. The total number of patents assigned to function f is thus $O_f = \sum_k P_{fk}$. The number O_{tf} of observed joint occurrences of AI technique t with AI function f is $\sum_k P_{tk} P_{fk}$.

Given this setting, let us now define a random variable X_{tf} as the number of patents assigned to both technique t and function f under the assumption of random joint occurrence. Then, X_{tf} can be considered a hypergeometric random variable of mean μ_{tf} and variance σ^2_{tf} as follows (population K, number of successes O_t and sample size O_f):

$$\mu_{tf} = E(X_{tf} = x) = \frac{O_t O_f}{K} \tag{C1}$$

$$\sigma^2_{tf} = \mu_{tf} \left(\frac{K - O_t}{K} \right) \left(\frac{K - O_f}{K - 1} \right) \tag{C2}$$

If the actual number O_{tf} of co-occurrences observed between AI technique t and AI function f greatly exceeds the expected value σ^2_{tf} of random joint occurrences, then technique t and function f are highly complementary. Inversely, when $O_{tf} \leq \mu_{tf}$, AI technique t and AI function f are deemed as excluding one another, meaning they do not complement one another. Thus, complementary τ is defined as follows:

$$\tau_{tf} = \frac{O_{tf} - \mu_{tf}}{\sigma_{tf}} \tag{C3}$$

Typically, τ_{tf} is a real number that can be positive or negative and may be thought of as the degree of complementarity between couples of techniques and functions. The same logic can be applied to quantify how AI functions apply to specific AI-related applications. Define AI applications a such that $a \in \{1, 2, \cdots, a, \cdots, A\}$. Now let $P_{ak} = 1$ if patent k is assigned to AI application domain a, and 0 otherwise. The total number of patents assigned to AI-related application a is thus $O_a = \sum_k P_{ak}$. We then define μ_{fa}, σ^2_{fa} and τ_{fa} as, respectively:

$$\mu_{fa} = E(X_{fa} = x) = \frac{O_f O_a}{K} \tag{C4}$$

$$\sigma^2_{fa} = \mu_{fa}\left(\frac{K - O_f}{K}\right)\left(\frac{K - O_a}{K - 1}\right) \tag{C5}$$

$$\tau_{fa} = \frac{O_{fa} - \mu_{fa}}{\sigma_{fa}} \tag{C6}$$

Again, τ_{fa} is a real number that can be positive or negative and may be thought of as the degree of complementarity between couples of functions and applications.