Telecommunications Policy 49 (2025) 102908

Contents lists available at ScienceDirect

Telecommunications Policy

journal homepage: www.elsevier.com/locate/telpol

The knowledge base of Big Tech: Research as a source of informational leadership by the dominant US digital platforms

Sandro Mendonça [a,b], Eduardo Silva [c], Bruno Damásio [c,*]

[a] Business Research Unit (BRU-IUL), ISCTE – Lisbon University Institute, Lisbon, Portugal
[b] REM-UECE, ISEG/ULisboa, Faculty of Business, City University of Macau, Macau
[c] Nova Information Management School (NOVA IMS), Universidade Nova de Lisboa, Campus de Campolide, 1070-312, Lisboa, Portugal

ARTICLE INFO

Keywords:
Big Tech
Digital platforms
Research
Publications
Dynamic capabilities
Technological diversification

ABSTRACT

Investing in research is an essential element in any business model in the digital age. At the forefront of this approach to innovation are digital platforms, which have resources and development strategies that enable them to enter and transform many markets outside their core activities. This work assembles a wealth of evidence to show that the omnipresence of Big Tech has now reached the science & technology sector in a significant way. We analyse the publication growth, influence, themes and partnership profiles of Big Tech (Amazon, Apple, Facebook, Google, and Microsoft). In a process that came alive in the 2000s, it is by the 2010s that these giant US-based ICT-services companies move fast and publish things. They show strong dynamic capabilities in "Physical Sciences" and "Computer Science". Co-authorships reveal links to American universities, mostly in California. Internationally, there is a strong preference for research collaboration with China, with the UK coming second. Results provide a better appreciation of Big Tech's multidimensional footprint and how the science & technology ecosystem is evolving and causing new policy pressures in the 21st century.

1. Introduction

Scientific research can play a significant role in corporate innovation and industrial leadership, strengthening, dynamising and diversifying a firm's business strategy (Mowery & Nelson, 1999; Rotolo et al., 2022). In this paper, we map and measure the growing research output of the largest digital platforms through their authored scientific publications. These companies display unparalleled data collection and analytics capacities, providing them with massive economies of scale and scope (Nuccio & Guerzoni, 2019). We may call "informational leadership" the outcome of these data science capabilities. The consequences are visible in terms of the structural market dominance based on corporates' tangible and intangible complementary assets, ranging from compute capacity, submarine cables, brand recognition, and lobbying resourcefulness (see, e.g., Mendonça et al., 2024; Rikap, 2024, Teece, 2023). Digital platforms can effectively compete with as well as leverage academia's scientific knowledge production, especially given the current landscape of university-industry ties (Lam, 2011).

Digital platforms emerged in the wake of the "Information Revolution", capitalising on the decentralised networking power of computers wired up through telecom infrastructure and protocols (Castells, 1997; Freeman & Louçã, 2001; Shapiro & Varian, 1999). In

This article is part of a special issue entitled: ITS Europe 2023 published in Telecommunications Policy.
* Corresponding author.
E-mail address: bdamasio@novaims.unl.pt (B. Damásio).

https://doi.org/10.1016/j.telpol.2025.102908
Received 3 August 2024; Received in revised form 21 November 2024; Accepted 18 January 2025
Available online 23 July 2025
0308-5961/

contrast to traditional industrial champions (Big Oil, Big Pharma, etc.), Big Tech operate on the basis of data-driven digital strategies and resources (Nuccio & Guerzoni, 2019). They re-intermediate information flows and build ecosystems of third parties to carry out business and co-create value in their private digital spaces (Cabral et al., 2021; Hein et al., 2020), resulting in strong network effects (Thomason, 2021). This ecosystem architecture grants the gathering of unique intelligence regarding users and potential partners, which in turn allows an expansion into ever more diversified business areas (Jacobides & Lianos, 2021; Klinge et al., 2023). Although considerable policy attention has been devoted to Big Tech, there continues to be a gap as far as future markets are concerned. Effectively, Big Tech engages in dynamic modes of competition by also investing in R&D and creating new markets, thus spreading its influence and presence forward over time. Light can be shed on the future pathways of Big Tech by focusing on the scientific activities undertaken by platforms, including those developed through the interchanges with other players, including non-profit institutions such as universities. Indeed, this research and innovation perspective is underexploited in the literature and can contribute to addressing the growing concerns over Big Tech, especially concerning the scientific underpinning of "techno-economic power" (Birch & Cochrane, 2022).

The activity of digital platforms in scientific research has been addressed by literature. Rikap and Lundvall (2021) theorised that this practice translates into data-driven "intellectual monopolies", by which each platform effectively orchestrates its own corporate innovation system. Others have noted an intensive research commitment by platforms in disruptive fields like artificial intelligence (AI), often publishing with academic researchers (Klinger et al., 2020) and presenting at reputable conferences (Ahmed & Wahed, 2020). Rikap (2023a, 2023b, 2024) also identified preferred research topics within the AI field, recurrent collaborations, mostly with American universities, and overall different research strategies between platforms as sources of rent. In turn, our work stresses how Big Tech's research profiles have been expanding in their scientific and technical capabilities in ample interchange with outside parties (Ma, 2022). Moreover, this steady but disruptive process also goes beyond simple competition for market share (a fight in the market) or for standard setting (a battle for the market) but points to a new kind of struggle for the very definition of future markets themselves (a fight for the next markets) (Geroski, 2003; de Streel & Larouche, 2004). As the new information and communications technologies (ICTs) motivated antitrust emphasis to move from *competition for how* to *competition for what* likewise the rise of *competition for why* may lead to new legal and regulatory challenges. Leverage over science content indicates a distinct domain in which discursive power and essential facilities are brought to bear on a complex economy by Big Tech (see Budzinski & Mendelsohn, 2021; Lubinki, 2004; Phoa & Gerbrandy, 2024; Shelanski, 2013; Silva et al., 2024). We suggest that this phenomenon in high-tech markets may invite, perhaps explicitly for the first time, the synchronous attention of science policy and competition authorities.

The primary motivation for this paper is to carry out a comprehensive assessment of Big Tech's knowledge base through scientific publications. So far, this perspective has chiefly focused on narrow industries, mainly in the fields of biotech and chemicals (Rotolo et al., 2022). To the extent that Big Tech has been explored, our study encompasses differences regarding extant work. For instance, we complement the pioneering inroads of Rikap and Lundvall (2021) by covering the period before 2014 and after 2019, by including conference papers, and above all, by considering the five largest digital platforms (Amazon, Apple, Facebook/Meta, Google/Alphabet, Microsoft), i.e. those huge US-based companies billed variously as Big Tech, Gatekeepers, Digital Platforms, FAANG or GAFAM. Comparatively, this enables us to nearly double the database. Note that the comparison between these five platforms is justified, given that they are the highest corporate R&D spenders in the world (Dyvik, 2022; Mendonça et al., 2024). Another difference to previous works is the inclusion of metrics such as citations, subject areas, co-authorships between Big Tech themselves, and a probe into specialisation and diversification issues. We complement and advance existing work in this scholarly niche by developing the most compressive body of evidence to date while also going into detailed quality controls for the purposes of robustness.

To understand Big Tech's contribution to the science and technology frontier, we deployed conventional techniques that uncover the trends and turns of knowledge production, collaborations and outcomes. Our paper reviews over 50,000 research publications penned by at least one author affiliated with the five platforms and subsidiaries. The methodology draws from the bibliometric and textmetric toolboxes of quantitative studies of science (Glänzel et al., 2019; Mingers & Leydesdorff, 2015), and we interpret the evidence through the prism of the theoretical compact known as evolutionary economics and innovation studies, with its emphasis on dynamic capabilities and knowledge-based resources (Fagerberg & E, 2013; Nelson & Winter 1982, 2018). Several cross-checks are carried out for the sake of buttressing the analysis, including comparing Big Tech's outputs and outcomes with aggregate dynamics and other global big business players.

Our data reveals a process of significant knowledge restructuring occurring in Big Tech over the long run. We confirm that Big Tech is active in the science domain and show that its academic publications are on the rise in terms of volume, variety and quality, revealing its distinctive profiles and how performance has changed over the years. Moreover, we assess which disciplines resonate more with Big Tech's publication efforts and list the academic institutions with which it mainly associates itself. A picture of influential "intellectual conglomerates" emerges, evocative of "technological diversification" processes that precede major redirection in knowledge accumulation trajectories. If multisided platforms require multisided accountability and regulatory systems, an implication is that explicit attention and evaluation must be assigned to them by the conventional structures of scientific governance (see Rahman, Karunakaran, & Cameron, 2024). Even if such companies are "intellectual monopolies", they are not "intellectual monoliths"; appreciating their specificities and eventual classes is instrumental to keeping policy analysis sensitive to the realities of actual corporate practice (Birch & Cochrane, 2022). Policy and regulatory concerns can thus benefit from realising the extent to which Big Tech now penetrate the academic sector and the innovation system at large.

This article is divided into the following sections. Section 2 presents a literature review framing the scientific research process environment (publication, funding, collaborations), as well as Big Tech's scientific activity and broad research patterns (themes, frequent partnerships). Section 3 states the research design, data treatment, and methodology procedures, respectively. Sections 4 and 5 present the major trends from Big Tech's published research profile, while Section 6 discusses them. Finally, Section 7 proposes conclusions and implications.

2. Research framework

2.1. The roll-out, renewal and re-upping of the digital age

Around the year 2000, the idea was prevalent that with the onset of e-commerce came lower prices (and their dispersion) while also allowing for faster adjustments to new conditions and information (Brynjolfsson & Smith, 2000). The "dotcoms" were creating "frictionless" and "weightless" markets, and in this way, they were the transmission mechanism of the Internet's potential to increase the overall efficiency of the economy (Oliner & Sichel, 2000). Indeed, information and communication technologies (ICTs) were then also visibly altering products and organisations themselves, something which was seen as being behind an unanticipated revival of productivity and the comeback of the American economy (Jorgenson, 2001). These changes gave economists and other observers a new agenda of research on the impact of computers and connectivity on growth and development (Gómez-Barroso & Marbán-Flores, 2020; Vu et al., 2020).

By 2010, the new buzzwords were "web 2.0" or "social networks", while the terms "access", "algorithms", "big data", "crowd", "gig", "monetisation", "sharing", and "wireless" were making headway (e.g. Belk, 2010). At the same time, branded "virtual real estate" was becoming increasingly recognised as an emergent template for value creation (e.g. Carroll & Romano, 2010). After the first Internet bubble burst, business models re-oriented from selling to sharing (Snowden, 2019, p. 4).

By the year 2020, the coalescing characteristics built around online intermediation and market multisidedness were becoming known as the "ecosystem economy" (Jacobides, 2019) or "platform capitalism" (e.g. Frenken et al., 2020; Marciano et al., 2020). Indeed, something on the Internet had shifted. The digital landscape was now akin to a privatised environment where communities of complementary and competing actors interacted. More and more business areas were punctuated by the presence of keystone players adopting the platform business model while also establishing the interfaces and the rules of the game (Teece, 2014, 2018). The consequences in terms of antitrust problems, privacy violations, fake news and hate speech were only then being grasped and articulated (see, e.g., Kahn, 2017; Petit, 2020).

2.2. Informational leaders in the data space

Rather than viewing the contemporary economy as an "online invisible hand", we approach platforms as the possessors and sponsors of the raw materials and means of production that shape a digital playing field that increasingly engulfs more and more activities of economic and social life. Taking the corporate perspective is helpful to understanding the agency and transformative power of platforms while positioning this type of organisational form as a mutation of the large traditional industrial leaders that prevailed until the 20th century (Frenken & Fuenfschilling, 2021). The analysis of corporate capabilities allows us to assess profiles and diversity, continuity versus disruption of business organisations in the context of a new phase of the current ICT techno-economic paradigm (Louçã & Mendonça, 2002). In fact, they can be thought to incarnate the distinctive business model of the current phase of informational capitalism (Freeman & Louçã, 2001; von Tunzelmann, 2003). Effectively, rising corporate concentration was industrially-led before the 1970s but informationally-led henceforth, with R&D and ICT investment indicating increasing economies of scale and scope (Know et al., 2024).

Big Tech companies are examples of platform-based enterprises (Xue et al., 2020). We take the hypothesis that these companies form the kernel of the contemporary digital economy (Klinge et al., 2023; Manganelli & Nicita, 2022). On the one hand, Big Tech, as platform owners, have the privilege of engaging in innovation within the platform environment but simultaneously apply terms and supervise conditions inside their digital realms as if they are legal frameworks while acting with discretion in harvesting big data, adjusting algorithms and policing content (Frenken & Fuenfschilling, 2021). That is to say, they control a corporate innovation system that is open (in a self-regulated way) to the participation of many other agents, such as end-users or business users (Rikap, 2023a, 2023b, 2024). On the other hand, their expansive ability to leverage common intangible and tangible infrastructure is a powerful driver of competitive advantage in a widening range of services and creates substantial network externalities (Ehrentraud et al., 2022). This process leads to potential excessive concentration and digital rentiership in the provision of technology-based products (Birch & Ward, 2023; Ma, 2022; Rikap & Lundvall, 2021). Such profitable large superstar firms that arose from new technologies but also from their ability to extract value from new learning-by-doing possibilities exert new pressures on institutional frameworks and raise political economy questions (Arora et al., 2023; Barkai & Benzell, 2024; Martínez & Bunkanwanicha, 2024; Ma et al., 2024; Tirole, 2024).

2.3. From antitrust to regulatory to political economy concerns

While achieving market leadership, platforms' exponential growth since the turn of the century has raised increasing concerns. Antitrust radically came to the fore until the early 1890s with the industrial age; in turn, the challenges of digital disruption reignited the debate in the late 2010s with Big Tech developments causing apprehension and anxieties while also appearing on the radar of several jurisdictions (e.g. Australian Commission A. C. and C., 2019; OECD, 2023). Issues have been many, such as market definition, measurement of dominance, identifiability of consumer harm, self-preferencing, explainability of algorithms, killer acquisitions, the dynamic effects of digital conglomerates, etc. (Allen et al., 2020; Bourreau & de Streel, 2020; Cabral et al., 2021; Crémer et al., 2019; Scott Morton et al., 2019).

Beyond competition policy, that is, from a consumer protection perspective, Big Tech has also attracted criticism because of ethical issues related to abuses over the usage of private personal information and preference engineering. These are consequences of market power by sprawling digital titans and do not conform to antitrust notions of short-term price effects; there has been analysis on extractive behaviour as these companies succeed in tilting the Internet toward a gated centralised control ecosystem (Bork & Sidak, 2012; Creser, 2020). Those practices can be the hidden collection of data through tracking (social plug-ins), manipulative engagement (built-in biases and stereotypes), and data brokering (selling to third parties), which can, in principle, be regulated (Roosendaal, 2011; Wu, 2017; de Freitas & de Moura Filho, 2022).

What is more, the heightened concerns about Big Tech have come to raise the prospects of the entrenchment of "Silicon Valley" power in society and the world at large. Ascribing political influence to these multinational corporations evokes historical parallels with Big Oil, Big Tobacco and Big Pharma for expanding influence in academia (Monsees et al., 2023), as even an article by the *Washington Post* (which became part of the broader Amazon emporium) intitled "Big Tech funds the very people who are supposed to hold it accountable" suggested (3 December 2023). In the case of Big Tech, however, their ability to commodify reality (through virtualisation) and connections (through datafication) means that they can be a much more powerful agent in fundamentally altering social relations and somehow interfering with the public interest in a profound sense (Foster & McChesney, 2014; Zuboff, 2019).

2.4. Crossing business lines into the scientific arena

The impact of the giant Internet gatekeepers has prompted several responses at the international level. In particular, reactions by China, the EU and the US have tried to be comprehensive and consistent in a number of areas, from media to finance, by combining *ex-ante* entity-based regulation and *ex-post* antitrust measures while also incorporating aspects of greater governance outreach (Crisanto et al., 2021). The collision of the growing digital ecosystems with existing markets not only brought about new logic to the business models but also created hitherto unfathomable strains on existing practices and incumbent institutions (Kenney & Zysman, 2019; Lianos, 2022). Platformisation allows for new frontiers in scale and scope, which may not be captured with standard indicators (see Hoberg & Phillips, 2022). Big Tech indeed destabilised every industry it penetrated, even those previously thought to be completely unrelated, and the public powers were slow in developing frameworks and imagining guardrails around its risks.

Much less studied, however, is the entry of Big Tech entities in non-commercial fields. In particular, the extended innovation ecosystem of profit-seeking digital platforms is intimately linked to science and technology. In specific fields, the footprint of large ICT-based corporations has been clear for the last decades, and there are signs of a recent increase (Abdalla et al., 2023). Surely enough, digital platforms have a potential for massive information gathering that would interest the scientific community, and the same could be said about their computational power and data storage capabilities (see, e.g. Nentwich & König, 2014). We can also easily assume that because science itself can generate continuous streams of data (from astrophysics to economics), large informational players would themselves also be interested in tapping into such learning and algorithm training possibilities as sources of increasing returns (see Santesteban & Longpre, 2020).

Therefore, the same distinctive platform characteristics, such as data analytics or the ability to improve a given industry efficiency and outcomes (through the provision of diagnostics or services enhancement), that allowed platforms to enter new sectors (content production, pharmaceuticals), can also be applied to scientific research (Ozalp et al., 2022; Thomason, 2021). Having established their own innovation hubs and research units (Berger & Brem, 2016; Webster, 2023) and invested significantly in R&D (Georgescu, 2022), Big Tech proceeded to partner with universities in scientific ventures (Jurowetzki et al., 2021).

2.5. Big tech as leading knowledge-based actors in the informational era

The dynamics of the interaction of Big Tech operators with the sectoral scientific systems thus pose peculiar questions. Some authors point to a cluster perspective, i.e. a collaborative environment composed of universities, startups, government, R&D centres and venture capitalists, where large companies like Big Tech spearhead the process by engaging actively in innovation and research (Engel, 2015). Naturally, this prompts inquiries on the capabilities that large digital players display when developing activities in the realm of science (Abdalla & Abdalla, 2021; Verhoef et al., 2021) and how they nurture liaisons with other entities while assuming duties compatible with the ethos of science such as openness or commons norms (Dolata, 2017; Paredes-Frigolett & Pyka, 2023).

Although the overall literature on Big Tech scientific activity is scarce, there are some studies conducted on their patent activity (Rikap & Lundvall, 2022; Viera Magalhães & Couldry, 2021) and regarding their upstream scientific publication activity (Rikap, 2023b; Rikap & Lundvall, 2022), which indicate a growing engagement with universities. Big Tech publications tend to receive a significant number of citations, even when compared with other more specialised players, particularly in the AI field (Giziński et al., 2023; Klinger et al., 2020). In fact, the number of Big Tech publications in deep learning and machine learning has increased to such an extent in recent years that it directly influences and affects the fields' direction (Ahmed & Wahed, 2020; Birhane et al., 2022). A key motivation for the increasing engagement of Big Tech with universities' capabilities in fundamental research can be the rising importance of AI technology for the companies' services and R&D strategies (Mucha & Seppala, 2020; Schuhmacher et al., 2021).

A key dimension of an innovation-intensive economy is the reliance on intellectual resources (data, research tools, spending), learning routines (let us call these dynamic capabilities), and knowledge networks (Paredes-Frigolett & Pyka, 2023; Powell & Grodal, 2006; Powell & Snellman, 2004). Big Tech has been known to hire top scientists, mostly in AI (Webster, 2023), usually by acquiring the small companies where they worked (Hellman, 2022; Varian, 2021). This action translates into in-house researchers having more robust collaborative networks and more chances of acceptance of submitted papers (Sabatier & Chollet, 2017). Therefore, formally published papers in scholarly journals will allow us a window into the perimeter and composition of Big Tech's knowledge base and a reflection of evolutionary trends.

3. Methodology

3.1. Published papers as a corporate innovation indicator

The effort by productive agents to directly improve their ability to innovate is a key realisation of evolutionary economics (Nelson & Winter 1982). In the field of strategic management, the company's role as a knowledge creator is usually framed in a resource-based or dynamic capabilities view of the firm (see Nelson et al., 2018). Empirically, the knowledge-base is acknowledged to partially surface through a number of indicators, either input (like R&D) or output (such as patents). Two notes are in order. First, protection and appropriation, rather than sharing and dissemination, are choices by learning agents. Second, exploration seldom occurs alone, and research takes place in an interactive context of many voluntary or involuntary interchanges with a variety of other economic and non-economic actors.

This paper examines how the five Big Tech corporations publish articles in highly-held academic journals. The observation that companies do publish, and quite extensively so, has been around the field of innovation studies for several years. Namely, Tijssen and Van Raan (1994) and Hicks (1995) showed that large R&D-active companies release their knowledge in the form of papers. When industry actors do this, they are managing the question of keeping insights to themselves or pursuing a more open policy while also making decisions regarding the structure of their internal innovation systems and their external collaboration patterns. Through publishing, companies reveal the existence of tacit know-how and other uncodified intangible assets, thus allowing for inferences regarding directional intentionality, selection strategies and the nurture of cumulative trajectories over time (for recent discussion, see Camerani, Grassano, & Rotolo, 2023; Yang et al., 2023).

In their comprehensive work on publishing in the corporate sector, Rotolo et al. (2022) note that analytical interest in the corporate publishing phenomenon has grown since the year 2000. They show that studies on this phenomenon primarily focus on single industry sector studies; these have tended to analyse publishing science-based industries such as pharmaceuticals, biotech, and chemicals. Overall, the rationale for publishing seems to fall under a five-category list: (i) accessing external knowledge and resources; (ii) acquiring and retaining talent; (iii) supporting appropriation strategies; (iv) building reputation; and (v) supporting commercialisation initiatives.

3.2. Research design

Journal articles are usually considered the currency of the scientific enterprise and are the prime type of data analysed in this contribution. The increase in the number of journals and subsequently in publication numbers (Farrukh et al., 2022), at par with the emergence of new analytical techniques (Berger & Baker, 2014), have made bibliometrics a highly sought means to study data patterns. However, there are other sources. Conference papers can be considered a type of complementary evidence; these are generally shorter than articles and serve as harbingers or precursors to journal papers (Drott, 1995), hence not conveying the same information. Conferences are vehicles for rapid and low-threshold dissemination of knowledge in formation, sometimes a venue for more technologically applied work. Their dynamic nature and flexibility for including papers on emerging areas make them a source sensitive to research trends, even if not at the same peer-review level as journals. Applied fields such as computer science and software engineering are areas where conference proceedings have a significant amount of influence (Lisée et al., 2008). Pre-prints are other forms of scientific communication, and they are fast, open-access, non-peer-reviewed. They may also anticipate papers that will be formally published in journals (Cabanac et al., 2001). The advantage of articles as a unit of analysis is their higher degree of selectivity and how they reveal a higher commitment to go through the conventional academic peer-evaluation process.

The study of published output has contributed to an awareness of how the scientific and engineering community develops and exchanges ideas. For example, bibliometric studies report that multi-authorship in publications makes them more prone to being cited (Katz & Ronda-Pupo, 2019) and that international collaborations have above average citation rates (Marginson, 2022).

To the best of our knowledge, few studies or references profile Big Tech's published research, discussing its quantity, content, quality, partnerships, or influence. The work by Rikap and Lundvall (2021, 2022) stands as the principal incursion into this topic. The authors describe the scientific activity developed by some Big Tech companies (Amazon, Google, Microsoft, Tencent and Alibaba) in detail, report findings on the number and content of publications and patents developed by them until 2019, as well as about the recurrent co-authors. Note that in Rikap (2023b), there is also the presentation of Big Tech (Amazon, Google, Facebook, Microsoft) research publications presented at AI conferences and a discussion of each company's research strategy. Besides these pioneering explorations, other works have not explored the topic as thoroughly or used bibliometrics to the same effect, focusing instead, for example, on the assessment of a sole platform – Facebook – published research (Gupta et al., 2015).

3.3. Perimeter of analysis

A set of observations on scientific publications by Big Tech was compiled for the sake of systematic analysis. The focus is the literature output of the five large US digital platform companies published in formal academic journals and conference proceedings. The implementation of the paper search targeted every journal article and conference paper whose authorship is wholly or partially attributed to Big Tech employing at least a co-author with a Big Tech affiliation. Journal articles and conference proceedings were found to be correlated, and after establishing the central pattern, attention was concentrated on journal articles since the available information is more detailed and the standards more comparable. Indeed, reliable thematic classifiers of conference proceedings are not accessible. Given our goal to understand Big Tech through their activities in more conventional publication circles and in order to have a clearer definition of the boundaries of the dataset, pre-prints and other forms of publication were not considered. This confine is admittedly a conservative restriction, especially in faster-paced areas like computer science and AI, where tools and results are released in repositories like arXiv, but, if anything, it underestimates patterns that are already quite expressive in our findings. All available information was subject to standard bibliometric checks, namely author disambiguation.

Data were obtained through Scopus using an API key, and the search query was made in the affiliations' name field. In addition, several iterations of the names (and affiliates of) the Big Tech companies studied: Amazon, Apple, Facebook (Meta), Google (Alphabet), and Microsoft were searched. Scopus, like Web of Science, is a scholarly reference analytics system. Its journal indexing allows the individual paper to be tagged to research areas using the journal disciplinary fields (Waltman & Van Eck, 2012; Wang & Waltman, 2016).

3.4. Data protocol

Fig. 1 details the workflow. After attaining the initial set, data were filtered to include only articles and conference papers in English. Then, only publications after the foundation dates of the respective companies and up to 2023 were considered; hence, the sample period goes from 1977 onwards. The dataset was also checked for false positives and duplicates. This aspect required thorough analysis, first to be able to detect all author affiliation entities with Big Tech such as names and publications, and second to exclude all publications completely unrelated to Big Tech. For instance, Amazon Forest-themed papers were detected in the Amazon datafile or companies with the word "Apple" in their name. Finally, only publications with DOI (digital object identifier) were considered in the analysis. This final criterion was applied given that it attests scientific quality to a publication and is often used in bibliometrics (Khurana, Ganesan, Kumar, & Sharma, 2022; Liu, 2021).

All the entries with similar names or variations of the five Big Tech companies were changed to the short best-known names ("Apple", "Amazon", "Facebook", "Google", and "Microsoft") to simplify the database. Exceptions were made for Big Tech affiliates with a significant number of publications attached, such as "Microsoft Research" and "Google Research", considering them different entities.

As in Rikap and Lundvall (2021), most publications examined in this study are not of Big Tech's full and exclusive authorship; instead, they result from partnerships. Hence, for the sake of interpretative expediency, we are taking the liberty to refer to all publications with one author affiliated with Big Tech as part of the latter's research output. Furthermore, because of pervasive co-authorship, the findings presented in the paper should not be taken as solely representing the position of the authors affiliated with Big Tech.

To achieve discernible results, we grouped the authors under the umbrella of university affiliations. For example, authors that are affiliated with these universities: UC Berkeley, UC Davis, UC Irvine, UC Los Angeles - UCLA, UC Merced, UC Riverside, UC San Diego, University of California San Francisco, UC Santa Barbara, and UC Santa Cruz, are all affiliated to the "parent" university – University of California.

Finally, for the purpose of robustness verification, we brought network analysis to bear on corporate research output, which can be challenging due to some authors having multiple affiliations (Yegros-Yegros et al., 2021). We obtained the network maps of the publications by the most prolific authors affiliated with each platform through VOS Viewer.

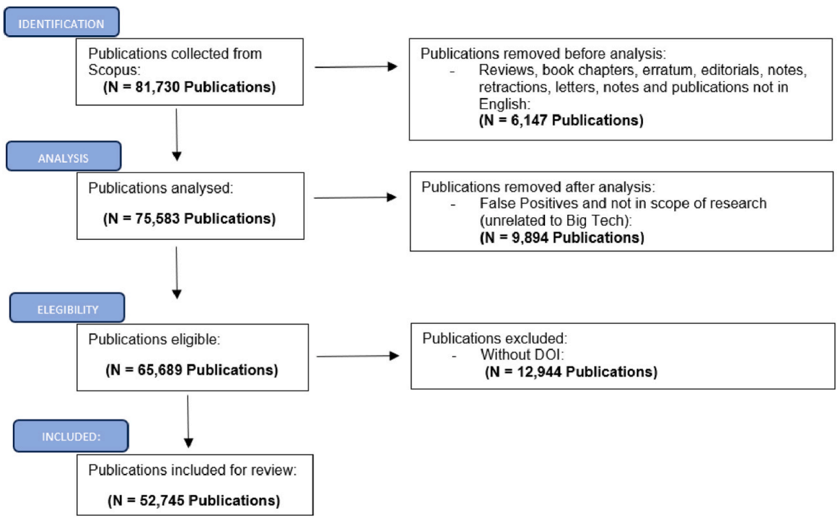

Fig. 1. Database building procedures.

3.5. Analytical orientation

The bibliometric approach followed here is a performance analysis (Donthu et al., 2021) focusing on volume, authorship and influence. Research networks are studied through the authors' affiliations, allowing a broad view of the links formed at institutional and international levels. Although not the whole picture, citations offer a glimpse of a given paper's innovative contribution and scientific repercussions.

Journals were labelled according to their subject area (e.g., Computer Science, Engineering) through the Scopus classification to study the thematic range of Big Tech's research. When said classification was unavailable, like in most conference papers, journals were labelled N/A (not available), constituting a limitation in the methodology. Therefore, given that all journal articles have an identified subject area (often more than one), only this type of publication was considered when performing a knowledge domain assessment.

Keywords and abstract content were also analysed. Text mining techniques were employed, removing non-alphabetic characters and stop words or examining the singular and plural forms of the same word as the sole one. Further analyses make use of approaches like concentration and specialisation indexes and leverage information from journal rankings.

4. General results

4.1. Basic stylised facts

In the six decades spanning our data sample (1977–2023), we detected 52,745 relevant publications. More than one-third of the total output is journal articles (19,739); the remainder are conference papers (33,006 items). The volume of total papers trended upward and accelerated in more recent times. A significant jump seems to have occurred around 2005. On average, the firms published the first paper about 3.6 years after being founded. By 2008, all the firms had contributed to the growing Big Tech publication pool.

In Fig. 2, we can discern three different subperiods in the growth published output. Until 2005, there is a modest volume. Until then, there was a meagre number of publications except for Microsoft, and almost residual for the others, also due to some companies being very recent (Facebook). From 2005 to 2014, Microsoft and Google experienced a tremendous increase in their level of publications. The year 2005 is also when proceeding papers decouple from journal papers, perhaps indicating a speedier phase of dynamic competition in the digital platform business. From 2016 onwards, we witnessed another burst of publications, which can be read as a moment of major turn-around in terms of learning and innovation priorities. This aspect temporally coincides with these companies' significant increase in R&D (Mendonça et al., 2024). Fig. 3 presents results by firms adopting this temporal breakdown, demonstrating that the last lap of this timeline shows Big Tech becoming more active, especially Amazon and Facebook. Apple consistently remains the company least interested in putting its discoveries and innovations into print.

Concerning citations, the general trend for both document types is one of growth throughout the timespan, as seen in Fig. 4. However, there were some fluctuations, especially after 2000, and there began to be a gap between the two document types. In the period between 2013 and 2018, conference papers recorded a significant number of citations, reaching the 250,000 citations threshold in 2016, while article citations peaked at almost 100,000 citations in 2017. Additionally, after the sizeable numbers recorded, both types showed a downward trend up to 2023, which is a typical bibliometric pattern as we get to more recent years.

Individually, Microsoft showed the highest number of citations in the whole period, followed by Google, although Google had more

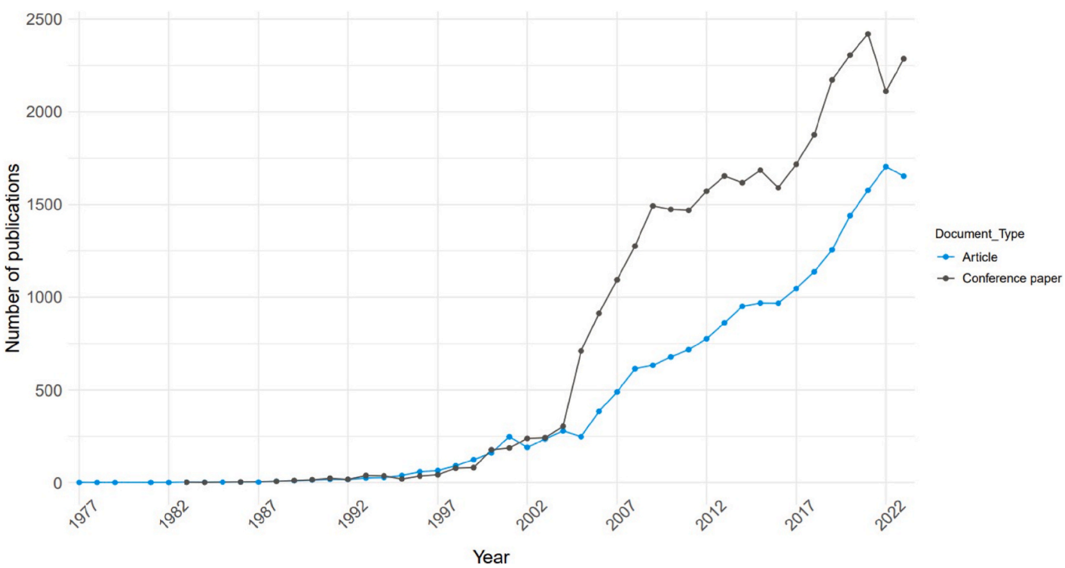

Fig. 2. Conference papers and articles by big tech.

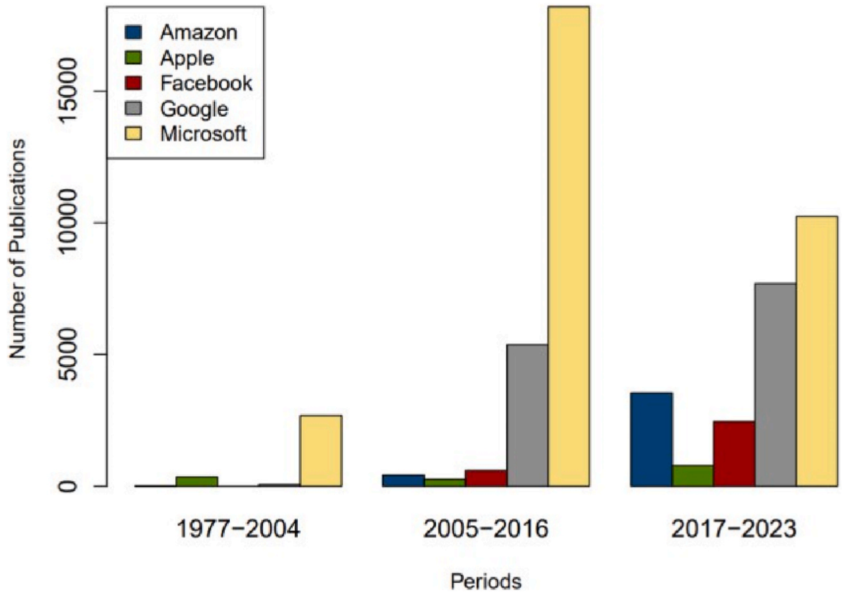

Fig. 3. Both types of publications by Big Tech (individually) over time.

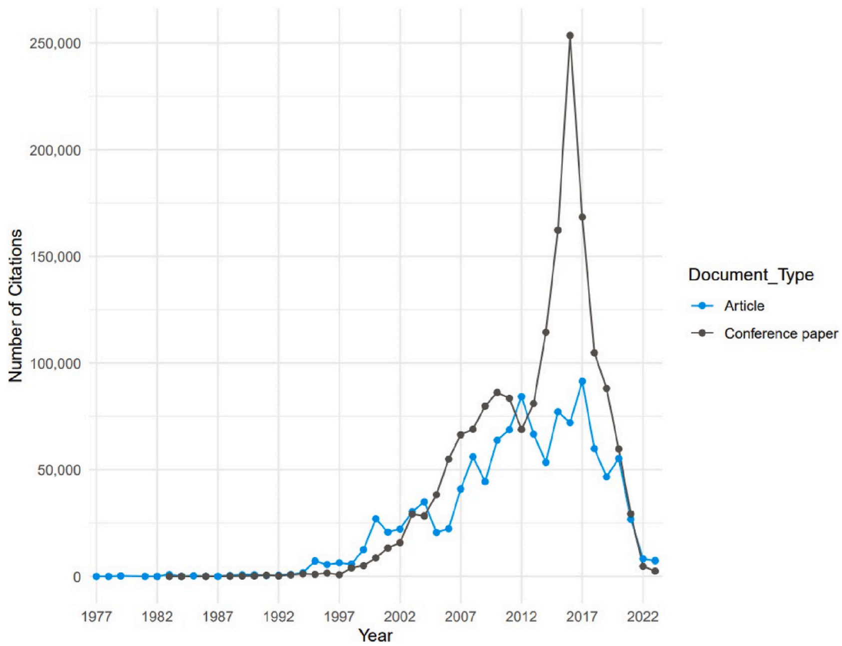

Fig. 4. Citations received by Big Tech publications.

citations in the later stages. Facebook has a robust number of citations, particularly after 2015, which is especially remarkable considering its lower number of publications. Apple and Amazon show residual values. After 2018, Google took the lead, followed by Microsoft and Facebook. However, no platform presents an unambiguous linear growth path so far, not allowing us to clearly assert how the scientific community is considering publications at this level of aggregation.

The performance of Big Tech should be seen in perspective. Fig. 5 sets their publication performance (in index numbers) against the grand total of all published scientific articles in the database. Between the early 2010s and the early 2020s, Big Tech's output grew at an average yearly rate of 7.8 %, which is higher than the general growth of 5.6 %; this represents a speed almost 40 % higher than the benchmark.

The five Big Tech companies studied are among the biggest spenders on R&D, having surpassed Big Pharma (e.g. Johnson & Johnson, Novartis, Roche, Pfizer) or multinational car manufacturers (Daimler, Toyota, Volkswagen) over the past decade and

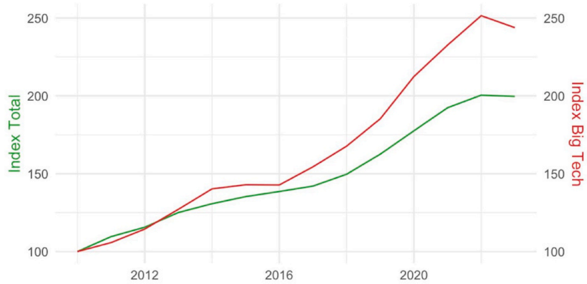

Fig. 5. Publication performance of Big Tech vs total Scopus publications performance.

boasting amounts in R&D only comparable with fellow ICT-equipment companies (e.g. Huawei, Intel, Nvidia, Samsung) (Dyvik, 2022). This investment increase was particularly felt in the period between 2015 and 2020, coinciding with peak growth in scientific publications (see Mendonça et al., 2024; see also Cano & Lind, 1991).

A comparison of Big Tech authored publications and citations received with the other top R&D spenders (Pharma, Automotive) can be inadequate, given that most of the latter firms have been established in the market for a long time, while Amazon, Google or Facebook were founded in the 1990s and 2000s. Furthermore, Big Pharma or Big Auto are more well known for their applied research efforts, which often became more visible in the form of patents rather than scientific publications (Dosi, Marengo, Staccioli, & Virgillito, 2023), and there are signs that Big Tech is following this trend (Merges, 2020). Nevertheless, Scopus searches into other companies' publications and, by applying the same criteria, yields that most have a comparably larger scientific output than Big Tech but a proportionally (and in some cases total) lower level of citations. Therefore, it is of great importance to stress the non-linear correlation between the increase in R&D spending and the number of scientific publications by Big Tech companies. It goes without saying that it is beyond the scope of this paper to relate their research with their financial performance, but the findings are suggestive (Table 1). Big Tech displayed robust growth in scientific output in the 2010–2023 period. It was only surpassed by the other "New Economy" sector, namely that of the ICT-equipment firms ("Hard Tech"). The "old economy" sectors exhibit a distinct pattern over time, especially Big Pharma, which has a contraction of publications. A revealing and complementary picture is revealed by the analysis of the research outcomes, i.e. here to be understood as citations of the published papers. Taking all the citations from the 16 companies as a whole (6,078,852 citations), we see that Big Tech commands the most influence. Hence, studying Big Tech's scientific output can be justified by its strong research growth at a corporate level, which is unmatched in terms of recognition.

4.2. Knowledge domains and standing

Regarding the study of the subject areas of Big Tech's published output, the results reported here only refer to journal articles. Fig. 6 captures the breadth of research fields by showing the number of Scopus categories in which the firms' aggregate has been published. The major finding is the steady increase in the variety of fields in which Big Tech is engaged.

There is a clear predominance of "Physical Sciences" and "Computer Science" as Fig. 7 illustrates. Areas like "Engineering", "Physics and Astronomy", and "Mathematics" are also important. Thus, the main subject areas of journal articles are in STEM (science, technology, engineering, and mathematics). Additional analysis for proceedings papers is convergent with this observation: they indicate the popularity of conferences related to neural systems, AI, computer vision and signal processing. It is also worth noting that "Social Sciences", "Business", and "Medicine" have non-trivial quantities of journal articles.

The dynamics over time can be grasped in Fig. 8, which takes snapshots at the turning points of the decades. Fundamental areas are on the increase, like "Physics and Astronomy". In the last few years, it has been notable how fields of application have become closer to final users' lives (like "Social Sciences").

The contributions to the key subjects have become more widespread among firms. Fig. 9 displays the five top areas in 2010 (by this year, all five Big Tech companies had already been publishing regularly) and computes a well-known measure of concentration (the Herfindahl- Hirschmann Index, the sum of the squared shares of the subject categories, or HHI). Results show a sliding down tendency, which points to capabilities in these fields to have become more even among the firms. Additional analysis for the aggregate article publications (i.e. the total and the individual contributions) shows that from 2018 there is evidence of cross-the-board deconcentrating

Table 1
Variation rate of Publications and weight in total citations of the different groups (2000–2023).

Group	Variation rate of publications	Weight in total citations
Big Tech	83.0 %	34.4 %
Big Pharma	−4.3 %	31.3 %
Big Auto	29.3 %	5.6 %
Hard Tech	101.4 %	28.7 %

Notes: Big Pharma - Johnson & Johnson, Novartis, Roche, Pfizer; Big Auto - Daimler, Toyota, and Volkswagen; Hard Tech - Huawei, Intel, Nvidia, and Samsung; no effort of consolation was made for the business groups.

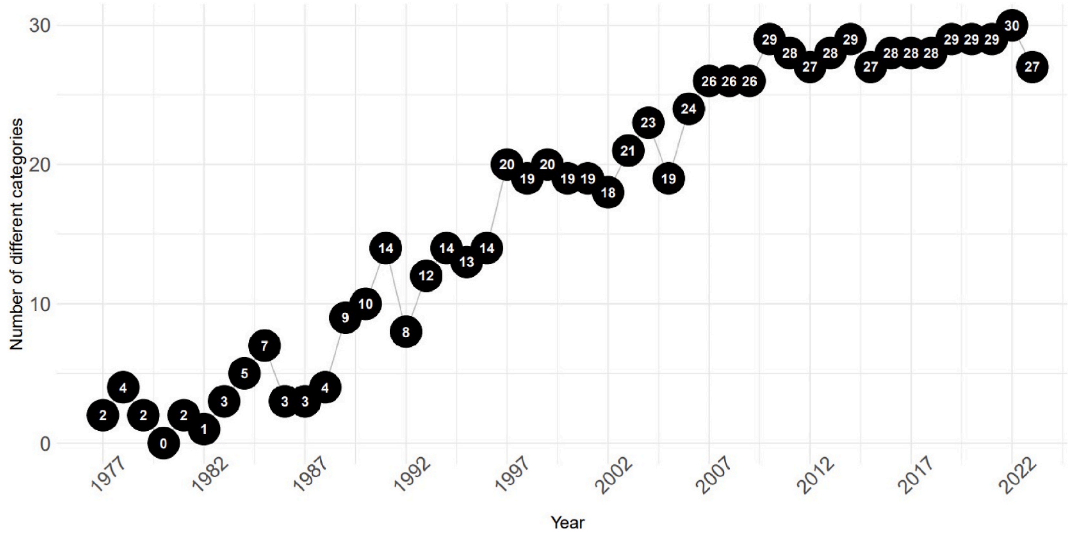

Fig. 6. Number of subject areas over time as revealed by journal articles.

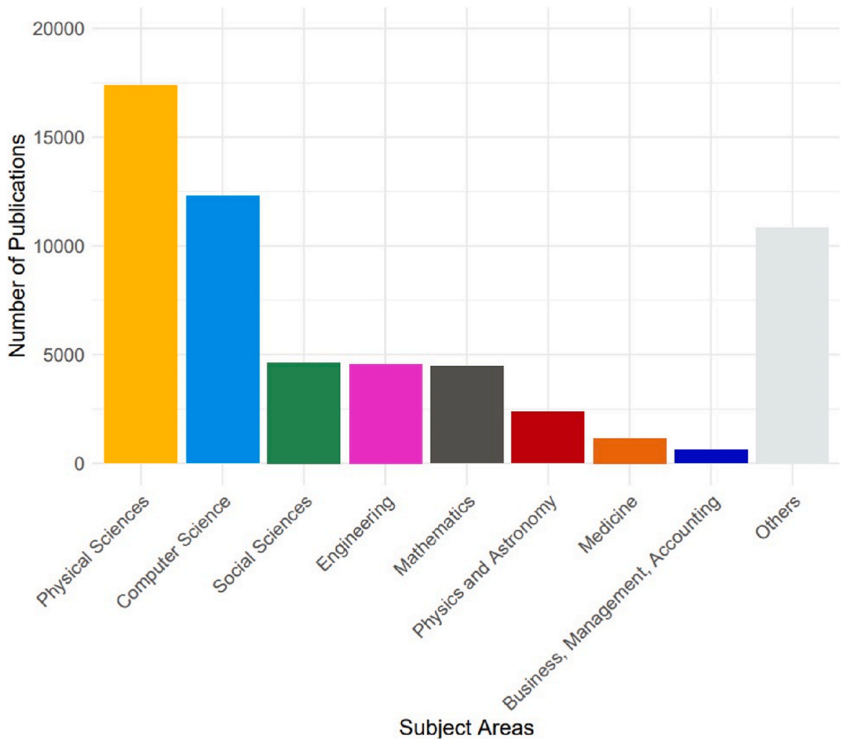

Fig. 7. Subject areas of journal articles published, 1977–2023.

publication output.

The trend for a multidisciplinary research agenda with a strong component of computational science is not exclusive to Big Tech. It can be seen as the longer-term trend of ICT-biased technological diversification (Mendonça, 2004, 2006) with other large companies that were traditionally located in different corners of the technological space, pharma, having also broadened and re-directed their research scope (Rafols et al., 2014). However, in the case of tech companies, this burgeoning interest is more pronounced due to their market footprint in different business areas (Stallkamp & Schotter, 2021). The case of healthcare is particularly noteworthy, given the small but increasing number of publications in the subject areas of "Medicine", "Neuroscience", or "Health Sciences". The investment in digital health and diagnosis solutions by Big Tech could potentially place them as significant players in this field (Schuhmacher et al., 2023). The data war chests and computational power of Big Tech (Thomason, 2021), as well as their heterogeneous business

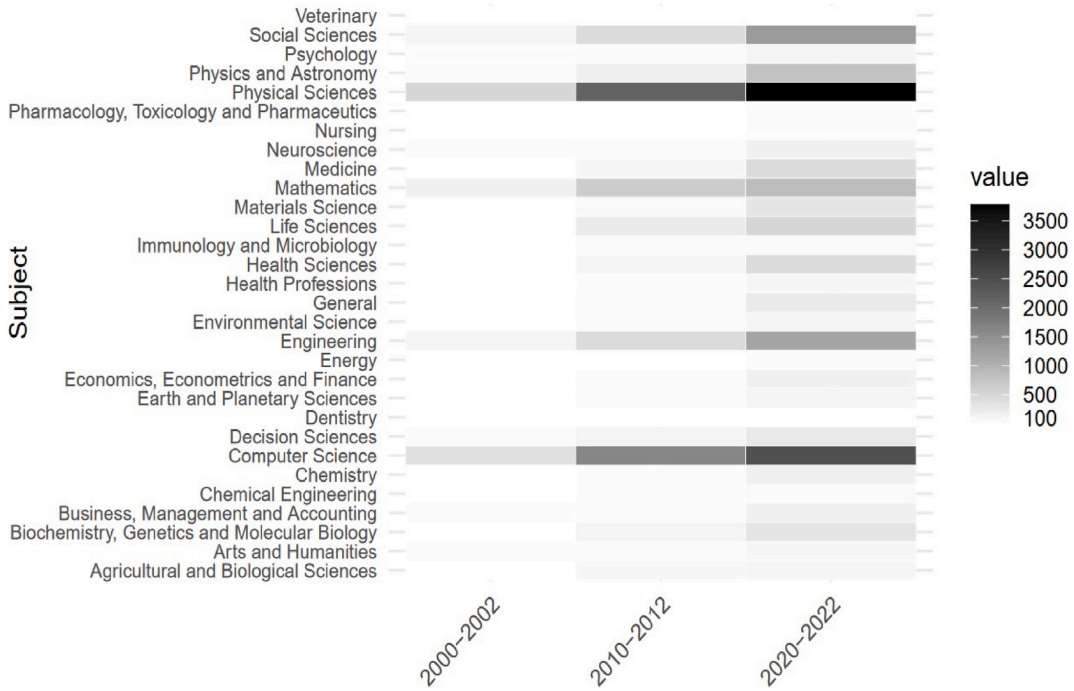

Fig. 8. Subject areas of journal articles published, 2000–2022. Note: The aggregates for the years 2000–2002, 2010–2012 and 2020-22 are computed.

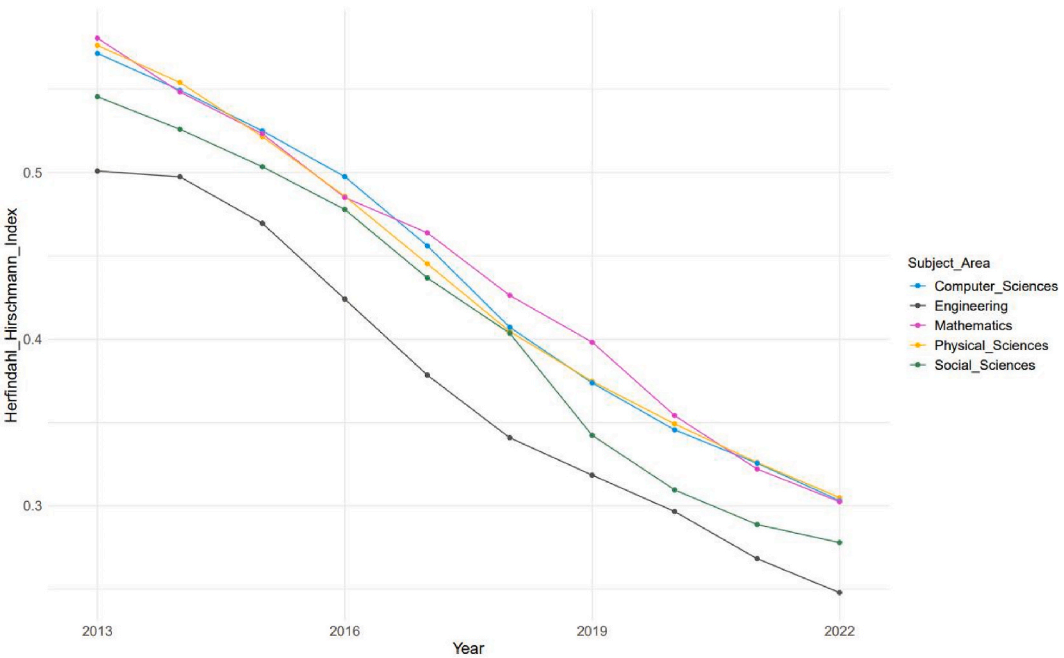

Fig. 9. Herfindahl-Hirschmann Index 2010–2022 in key subjects (moving average of 3 years).

ventures, suggest a new corporate profile as "intellectual conglomerates".

This picture of the broadening strength of Big Tech in terms of *volume* and *variety* of their cutting-edge expertise can be better understood if another layer of evidence regarding the standing of the journals in which they publish is added. One way to assess *quality* is to examine the standing of the outlet and to rely on rankings while observing the proportion of those articles published in the top 25 % of the journals (first quartile publications, or Q1). Between 2010 and 2020, the proportion of all articles published in Q1 journals increased from 24.0 % to 53.4 %. Taking an individual perspective allows an appreciation of the substantial jump in high-standing research (Table 2). We see that the firm with the lowest increase (Apple) almost doubled its number of journal articles published in leading journals, while the most prolific firms (Google, Microsoft) at least trebled their publication record.

4.3. Knowledge specialisation and relative performance

Relative advantages refer to an entity's capacity to engage in a particular area more successfully than its peers. Differentials in profiles may be grasped from a statistic like the revealed comparative advantage (RCA), which is a number that, when higher than a neutral value, like 1, shows that an actor is specialised in a given field. By using RCA in some period t we compute it by checking if the share of articles in a given field k in the total article production of firm i is larger than the equivalent share of all other firms.

In Table 3, we compute the RCA for the years 2010–2015 and 2016–2023. Joining years is implemented to avoid slight observation bias, and we use similar time windows (6-year and 8-year periods). We realise that most companies end up betting on computing, mathematics, and physics over time. Google and Microsoft are moving toward more similar profiles. Amazon and Facebook appear to be quite distinct in their profile. Apple maintains the strongest competence in Engineering.

4.4. Research collaboration and interactions

Most papers by Big Tech are coauthored with academics or researchers at scientific laboratories, as will be presented in section 5. Our analysis also identifies a small number of papers with full Big Tech authorship and an even smaller number with authors not affiliated with academia or Big Tech. This facet can suggest a dependence on external knowledge, especially from the research sector.

Multiple authorship in the realm of digital platforms is pervasive and on the increase. Fig. 10 shows that all firms follow an upward trend. The average number of authors collaborating in papers with Big Tech rose from around three in the year 2000 to over seven in the year 2022.

For the most part, research collaboration between the Big Tech players is relatively small (Table 4). Some firms do not coincide much in the same papers, namely Amazon-Apple and Apple- Facebook. However, there are some expressive collaborations, namely Microsoft-Google and Google-Microsoft, which surpass the 500 publications level. However, these two platforms also happen to be the largest article producers, so this co-authorship pattern may not be so surprising. Nevertheless, this may signal a possible future paradigm of the platform's scientific research, namely whether these entities will start to regard science as a common goal to pursue through pre-market partnerships or if, instead, a competition environment is established through creative subversion of the other's technological trajectories.

5. Specific results

5.1. Amazon

Amazon's core business is retail through e-commerce while at the same time committing itself to cloud solutions as part of its overall strategy (Berryman, 2014). The company's scientific ventures remain something to be studied, as well as the scientific network potential of Amazon Web Services, a global cloud computing service (Mathew & Varia, 2014).

The total number of Amazon-authored publications in our database from 2000 (the year the first publications with DOI appear in our dataset) to 2023 was 3970 (1542 articles and 2428 conference papers). The general growth path is upwards, showing a continuous increase with a kink in 2016, with 2023 being the year with the most publications: 925 (see Fig. 11a).

Amazon, like other Big Tech in general, has a greater proclivity towards the "Physical Sciences" and "Computer Science" while also displaying considerable presence in the "Social Sciences" (Fig. 11b). Looking at the most recurrent keywords available in the papers' front (for bigrams drilled from the abstracts, see Appendix A), displayed in Fig. 11c, the preferred themes in Amazon's research are linked to computing and AI through terminology like "machine learning", "deep learning" or "speech recognition". This focus on AI research, as well as the booming number of publications, can be attributed to the goal of reaching the level of production set by direct rivals such as Google and Microsoft (Satariano & Metz, 2023).

As for the co-authorship of Amazon's research output, most of the researchers are affiliated with American entities. As seen in Fig. 11d, Amazon does not exhibit many strong co-authorship partnerships; however, we are able to identify US universities like the University of California or the University of Texas in co-author affiliations. Therefore, the bulk of knowledge produced is by American-based researchers affiliated with the academic sector.

Finally, to further study the partnership dynamics in the publications' authorship, we looked for the articles and conference papers by the top 5 most prolific authors (see appendix A). The results confirm the general work trend with academics while indicating other collaborators, which are not present in Fig. 11, such as the University of Haifa and Rutgers University.

Table 2
Number of articles published in Q1 journals, 2010 and 2020.

	Amazon	Apple	Facebook	Google	Microsoft
2010	7	12	5	61	78
2020	68	23	68	223	387

Table 3
RCA, 2010–2015 and 2016–2023.

2010–2015	Amazon	Apple	Facebook	Google	Microsoft
Physical Sciences	0.93	**1.04**	1.00	0.97	**1.01**
Computer Sciences	0.94	0.53	**1.08**	0.99	**1.01**
Engineering	**1.45**	**3.63**	0.92	0.96	0.92
Mathematics	0.85	0.45	0.92	**1.05**	1.00
Social Sciences	**1.23**	0.50	0.89	**1.11**	0.96
2016–2023	Amazon	Apple	Facebook	Google	Microsoft
Physical Sciences	0.88	**1.03**	0.93	**1.02**	**1.03**
Computer Sciences	0.90	0.67	**1.08**	**1.01**	**1.04**
Engineering	**1.19**	**2.33**	0.94	0.96	0.84
Mathematics	0.97	0.57	**1.06**	**1.05**	1.00
Social Sciences	**1.45**	0.65	**1.06**	0.89	0.96

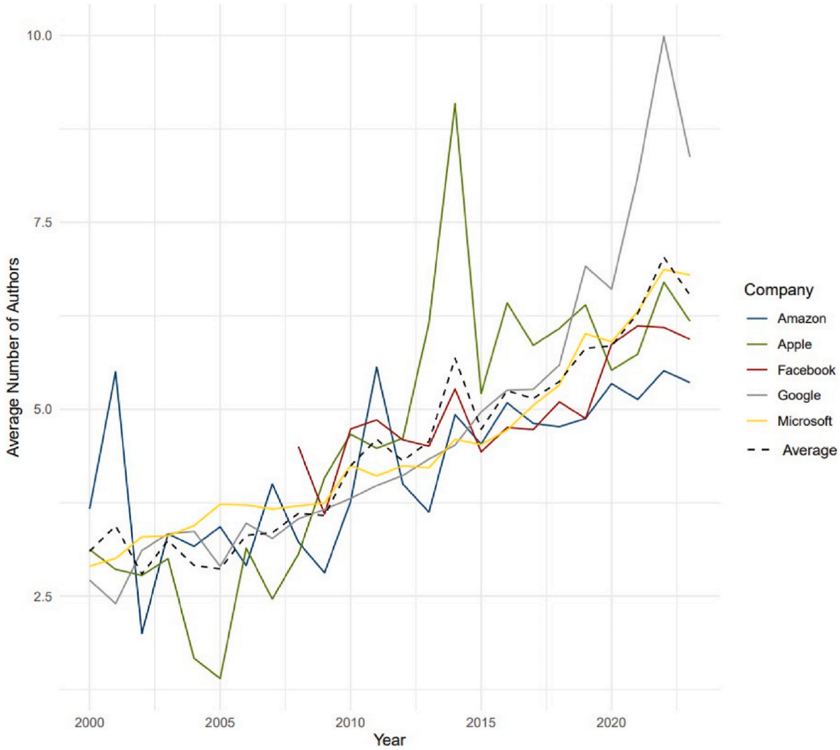

Fig. 10. Number of authors per paper, 2000–2023.

Table 4
Number of co-authored among the Big Tech firms, 2000–2023.

	Amazon	Apple	Facebook	Google	Microsoft
Amazon	–	13	54	92	113
Apple	14	–	14	43	22
Facebook	44	11	–	130	146
Google	103	31	145	–	537
Microsoft	101	17	148	501	–

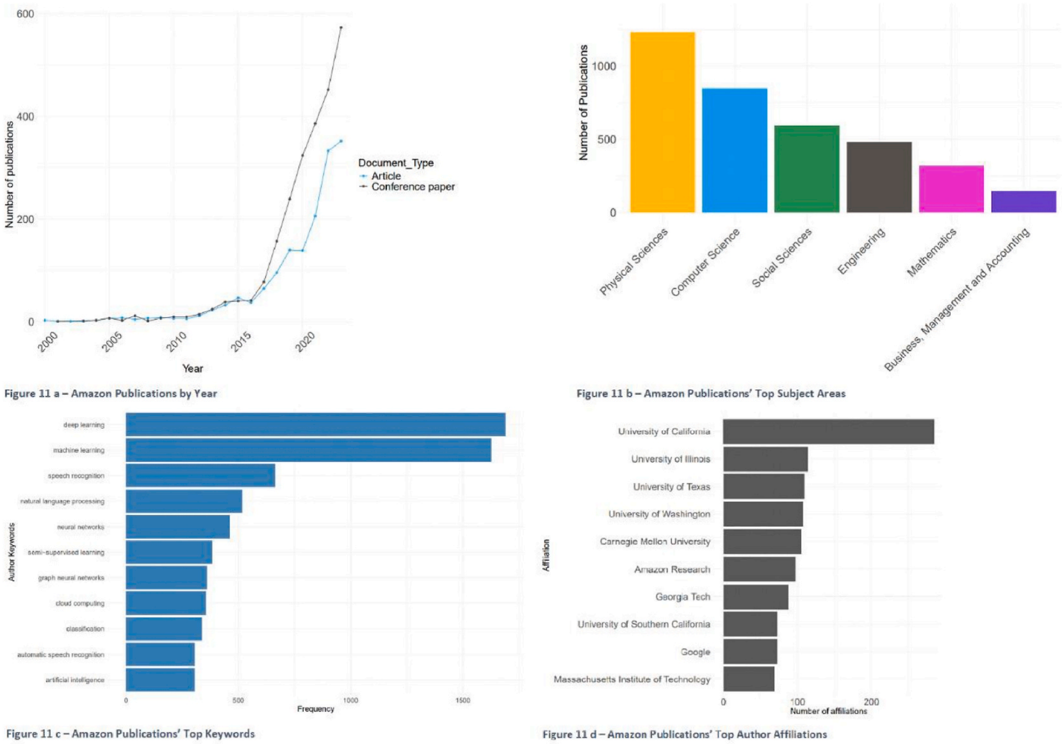

Fig. 11. Amazon publication profile.

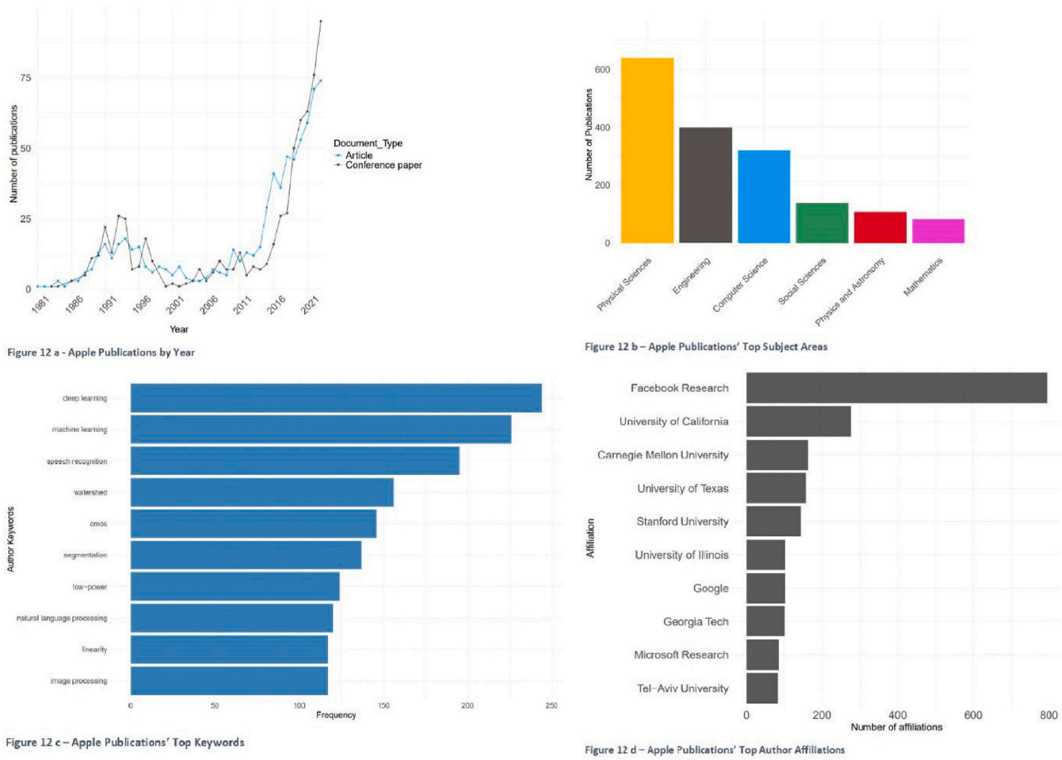

Fig. 12. Apple publication profile.

5.2. Apple

Like other Big Tech companies, Apple can be understood as a digital platform (through the Apple store), but its classic business is in software and hardware (Johnson et al., 2012; Zhang, 2017). Apple's publication record has two growth spurts. The first, which started in the mid-1980s, shows an increasing number of publications only to be interrupted in the mid-1990s (Fig. 12a). Then, there was a decrease to minimum levels at the turn of the century, always displaying a constant but small number of publications. From 2016 onwards, a rise is all too clear, with a maximum of 169 publications in 2023. Conference papers and journal articles share a strikingly similar time signature.

The firm's computer and consumer electronics product profile is reflected in the fields "Physical Sciences", "Computer Science", and "Engineering" (Fig. 12b). Keywords in the papers refer to topics in large volume data analysis like "deep learning" and "machine learning", as well as technology ones such as "speech recognition" and "CMOS" that play a role in Apple's services and products (Fig. 12c).

Regarding the authors' affiliations, a small proportion is associated with American Universities, with the "University of California" and "Stanford University" being the standouts (Fig. 11d), which can be part of a broader trend of Silicon Valley companies associating with top universities. Though the bulk of knowledge is co-produced with American academic researchers, the significant role authors affiliated with private companies play in Apple's research portfolio should be highlighted. Looking more closely at the most prolific authors' publications, we find the presence of authors of Cisco, Juniper Networks, Intel or Sony (Appendix B).

5.3. Facebook

As a digital platform with a business model oriented to connect people and circulate advertising, Facebook works primarily in relation to final users. Facebook can function as a support in many stages, from intellectual product development to distribution, and it can serve as a promotor of interactions between communities, including scientists (Nentwich & König, 2014; Schleyer et al., 2008). Abundant studies have been conducted on Facebook as a social network (González, Gasco, & Llopis, 2016; Niu, 2019), but very few on Facebook itself as a knowledge producer (Basak & Calisir, 2015; Coursaris & Van Osch, 2014; Gupta et al., 2015).

Facebook displays an uninterrupted growth in the number of publications from 2008 to 2021, going from a residual of publications to more than six hundred in total; in 2022, there is a marked drop to mid-decade levels (Fig. 13a). According to the business media reports, the first layoffs among Big Tech characterised 2022, and indeed, Facebook was the undertaking with the biggest workforce cull in this year; a conjecture may be that the 2022 publication drop may be related to reprioritisation taking place inside the firm.

The scientific journals with more Facebook-affiliated research are predominantly in "Physical Sciences" and "Computer Science".

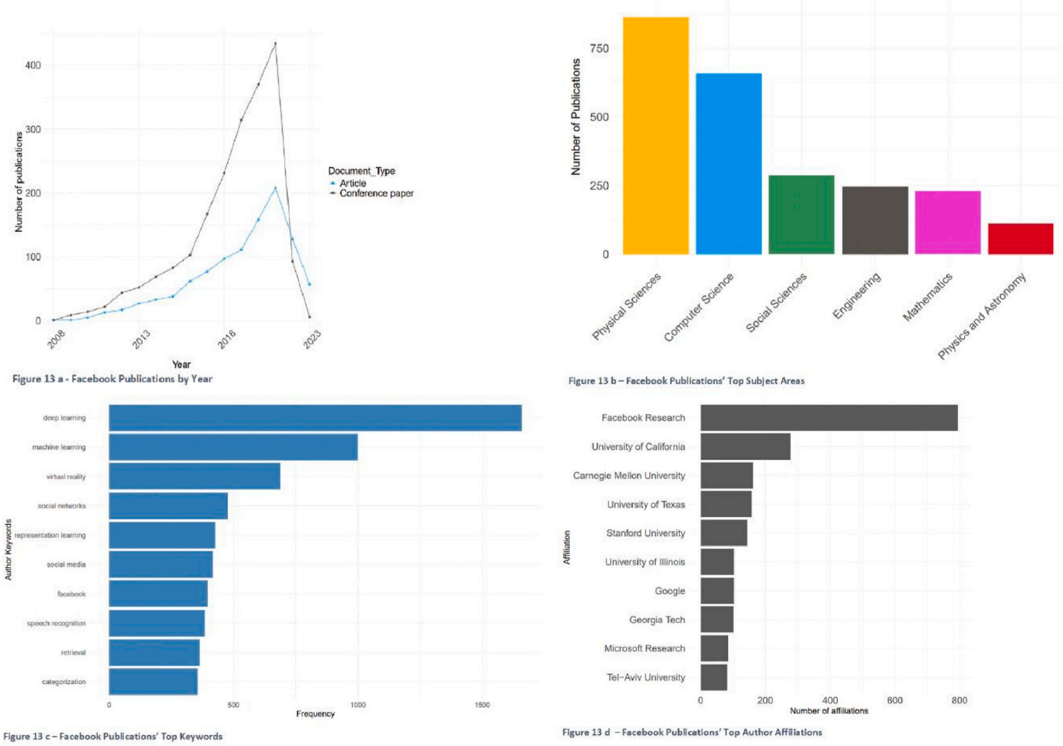

Fig. 13. – Facebook publication profile.

Some also cover the fields of "Social Sciences" and "Engineering" (Fig. 13b). There are also data science topics and AI ("deep learning", "machine" learning") and assorted technology ("speech recognition", "virtual reality") (Fig. 13c; see Appendix C).

A significant portion of the research was produced in-house with units such as "Facebook Research" (or "Meta Research") (Fig. 13d). However, as with other Big Tech firms, there is a significant bond with eminent American universities ("University of California", "Carnegie Mellon University", "Stanford University"). Naturally, most of the researchers' national affiliation is the US.

Looking further at the Meta partnerships by assessing the institutional background of its most prolific co-authors (Appendix C), we observe that it is firmly moved by the pull of the US-based universities, although Tel Aviv University and University College of London also stand out.

5.4. Google

Google is perhaps the Big Tech company with the greatest business diversity, being present in a search engine, advertising, AI, social media (YouTube) and other business areas. Not only did it emerge from a university background, but Google was also one of the first to put together scholarly knowledge (Jacsó, 2005) while providing bibliographic information with Google Scholar. Nonetheless, a study of Google's own research output could not be found.

Google's scientific and technological publications have an uninterrupted growth trajectory, totalising 13,129 publications (Fig. 14a). Google's publications rose steeply around 2005, presenting a steady growth, with the maximum volume recorded being 1323 publications in 2023.

Like other Big Tech companies, Google-affiliated researchers tend to publish more frequently in "Physical Sciences" and "Computer Sciences" themed journals (Fig. 14b). Google focuses on computer science and AI topics such as "machine learning", "deep learning", "speech recognition," or "neural networks", as seen in the publications' top keywords (Fig. 14c and Appendix D). These topics are often combined into a final product or application, like the incorporation of machine and deep learning technology in Google's speech recognition software (Lee et al., 2017).

Most of the publication's authorship affiliations are within Google itself, presenting some small recurring partnerships mainly with American universities ("University of California", "MIT.", "Stanford University") (Fig. 14d). As such, the authors' national affiliation pattern reveals a US prevalence (two-thirds of all publications' output), but China and the UK also figure as partners (see Appendixes). Additionally, when looking at the authorial network of the most prolific authors' publications (appendix D), we find a vast array of institutions ranging from astronomical observatories to European and Asian universities, implying that some Big Tech-affiliated researchers have a strong capability to associate with top scientists.

5.5. Microsoft

Microsoft is the oldest company in our dataset and displays most publications. Having had a global presence in the market from early on, Microsoft has delivered several product innovations that have had an economic impact (Acs et al., 2013). Additionally, like the other Big Tech companies, it is a technology trend-setter and seems to have turned its attention to data science recently (Kim et al., 2016, 2017).

As seen in Fig. 15a, Microsoft has recorded significant growth since the turn of the century. Articles keep going, but their presence in conference proceedings stabilised in the 2010s, unlike the other Big Tech companies, which could be credited to a certain stagnation phase of the firm (Prakash et al., 2021).

"Computer Sciences" and "Physical Sciences" are the more representative subject areas (Fig. 15b), although Microsoft also registers significant work in fields like "Social Sciences", hence displaying a diversified research portfolio. When examining the keywords (Fig. 15c) and bigrams (Appendix E), computing themes such as "machine learning", "deep learning", or "cloud computing" are prevalent.

Regarding Microsoft's co-author affiliations (Fig. 15d), like the other Big Tech companies, the overwhelming majority is affiliated with Microsoft, with some American and Chinese universitiesshowing a significant presence as partner institutions. In fact, the firm demonstrates a willingness to go cross-border in research. It has many works developed together with Chinese and UK entities, which distances Microsoft from the rest of the Big Tech lot, having a greater propensity to coauthor with non-US affiliated authors, particularly with Chinese ones. This latter trend is observable when referring to the publications by the most prolific authors (see Appendix E).

6. Discussion

The results obtained demonstrate vibrant academic activity and a robust increase in scientific production by researchers affiliated with Big Tech companies, with all these corporate outlets showing upward trajectories for most of the period in various subject areas. Microsoft and Google are the top-ranked publication hubs in terms of volume, but all display research agendas in high-tech fields like computing and physics. Therefore, there is a clear indication that Big Tech companies are private entities interested in producing cutting-edge knowledge in the realm of non-market research and innovation circles (see also Behrens & Gray, 2001). The interest seems to be higher than that revealed generally by corporate science (Camerani, Grassano, & Rotolo, 2023).

Big Tech seems to have become quite alive in publishing by the mid-2000s, and the late 2010s marked a rise and spread in overall performance (see Table 5). Significantly, evidence of a major reshuffling in terms of innovation and knowledge accumulation priorities occurs broadly in the same period where they became the world's foremost R&D corporate spenders (Mendonça et al., 2024). Note that

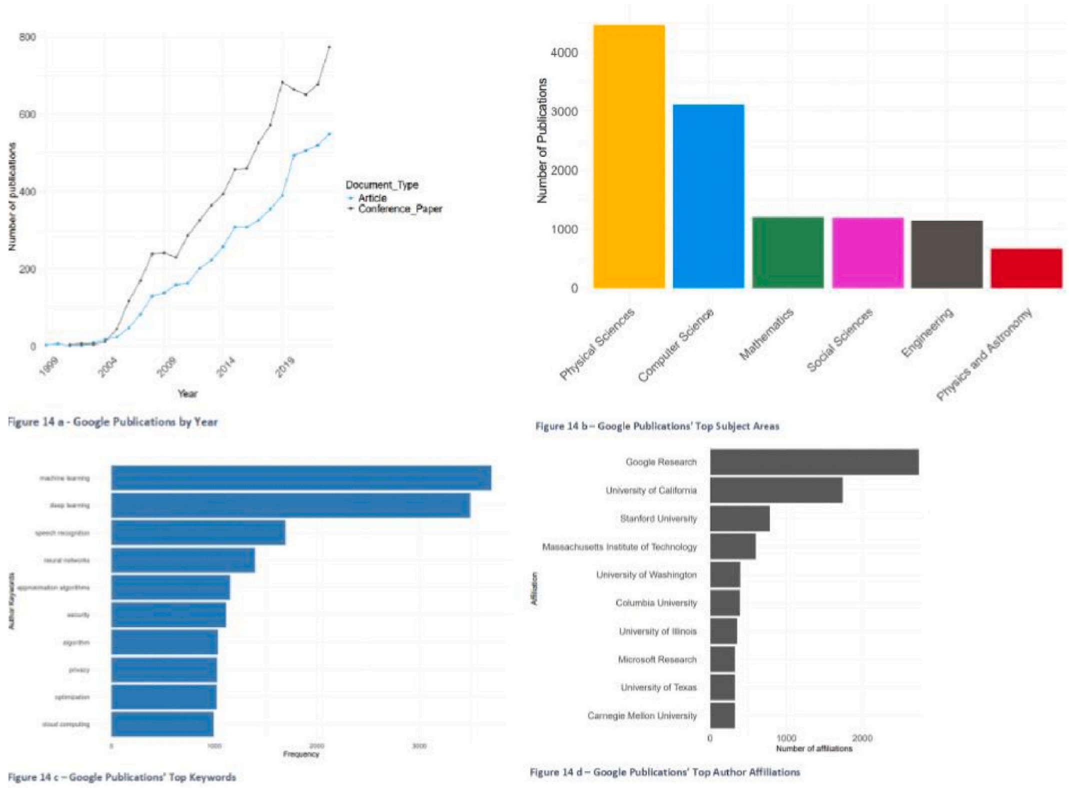

Fig. 14. – Google publication profile.

citation-wise, they display large levels proportionally, surpassing the thresholds of traditional R&D-focused companies like pharmaceuticals.

Big Tech's published output is hybrid, as it includes papers in conference proceedings and formal journals. We thus find evidence that Big Tech somehow make a point of attending conferences. Conference participation requires funding and sustained interest (Rowe, 2018). Conferences are also linked to a higher likelihood of publishing in a top-tier scientific journal (Gorodnichenko et al., 2021). Formal journal publishing requires heightened organisational commitment, and when considering the journals where they publish the most, we find evidence of top quality. Therefore, we argue that Big Tech has become fully committed scientific actors, presenting in the same vehicles (articles, conferences) as traditional research agents, even rising to a top position within this environment, given the financial leverage and ease with which they build a scientific network.

"Physical Sciences" and "Computer Sciences" are common and central and dynamic capabilities for Big Tech. Beyond these core fields, "Mathematics", "Social Sciences", and "Engineering" also represent important components of the knowledge base. Moreover, these companies nurture dynamic capabilities in a number of areas, and the scope of their learning efforts widened over time, thus suggesting an "intellectual conglomerate" profile. Such a pattern, recognised long ago as being related to the powers of ICT (Louçã & Mendonça, 2002), is associated with concepts like "technological diversification" and the "multi-technology corporation" (e.g. Mendonça, 2004, 2006). As for more precise individual research topics, "machine learning", "deep learning", or "speech recognition" are among the top keywords in all five tech companies' publications. This priority given to computer science and AI-related themes seems part of a common trend, with the automotive industry or the health sector titans also investing considerably in these subjects given their potential applications (Aggarwal et al., 2022; Awotunde et al., 2022; Sharma et al., 2021). Thus, research on general-purpose technologies like AI seems to have become a self-enforcing trend, given that it is the next step in platforms' new product development and novel business models (Rikap, 2023a; van der Aalst et al., 2019). That is to say, intelligentification building on platformisation, which had been built on cloudification. The combinational potential implicit in the giant informational leaders of this age thus warrants an "innovation-centric" view of economic regulation, industrial strategy and science & technology policy (see Cennamo, Kretschmer, Constantinides, Alaimo, & Santaló, 2023; de Streel & Larouche, 2004).

Regarding the authors' affiliations, there is a collaborative or "open innovation" environment, similar to the one developed by big pharma (Ma et al., 2022), in which informal ties with academia are prioritised when considering chief research collaborators. Though Big Tech does not have the same apparent motivations as pharmaceuticals, which relate to cost efficiency on research procedures (Rafols et al., 2014), instead tech companies look for ways to enrich their networks and complement their knowledge base with academics and external scientists (Nicoli & Iosifidis, 2023). Thus, the output is authored by the companies themselves with significant recurring collaborations, mostly renowned American universities like "Stanford University" or "Carnegie Mellon University"; the

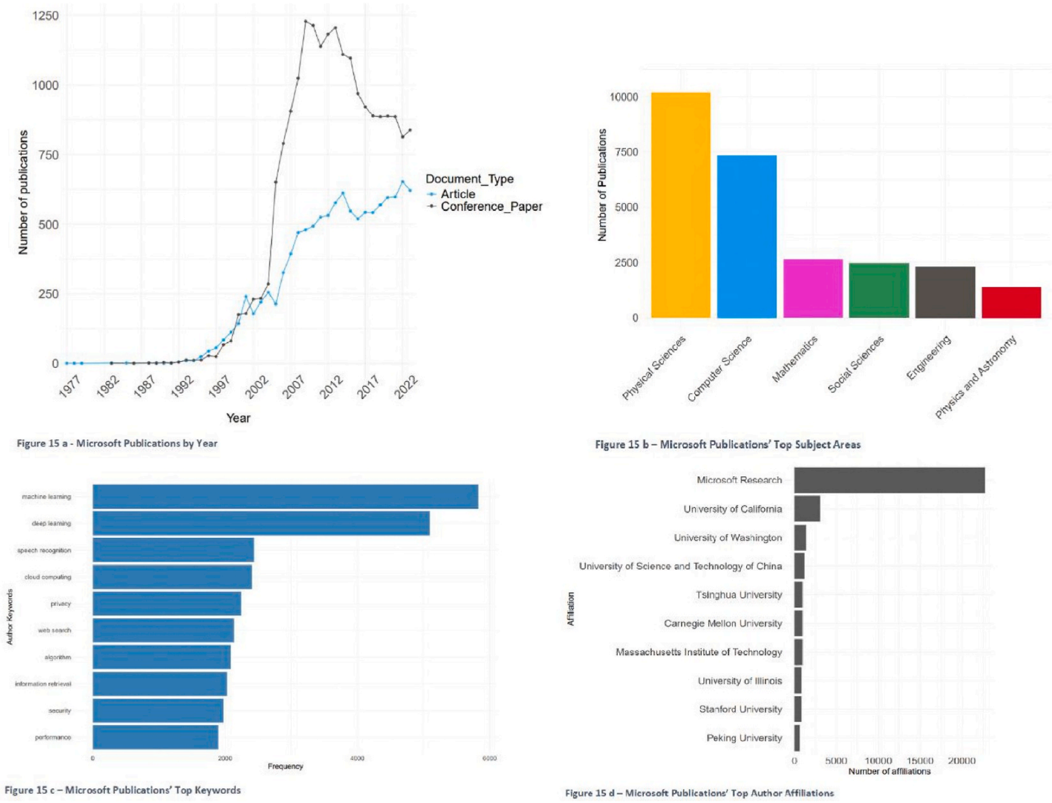

Fig. 15. Microsoft publication profile.

Table 5
Big Tech profiles, summary.

	Amazon	Apple	Facebook	Google	Microsoft
Rank in number of total publications	3	5	4	2	1
Publication take-off (and peak)	2017 (2023)	2015 (2023)	2017 (2021)	2008 (2023)	1999 (2013)
Top 3 subject areas	Physical Sciences Computer Sciences Social Sciences	Physical Sciences Engineering Computer Sciences	Physical Sciences Computer Sciences Social Sciences	Physical Sciences Computer Sciences Social Sciences	Physical Sciences Computer Sciences Mathematics
Number of subject areas in 2010 (and 2023)	9 (27)	15 (20)	5 (22)	23 (27)	28 (27)
Lead relative advantage (2016–2023)	Social Sciences	Engineering	Computer Sciences	Mathematics	Physical Sciences
Q1 article (% of total 2010–2023)	48 %	50 %	45 %	40 %	42 %
Outdoor ratio (% papers co-authored with external entities)	82 %	85 %	83 %	83 %	85 %
Top 3 collaborator institutions	Univ of California Univ of Illinois Univ of Texas	Univ of California Stanford University Carnegie Mellon Univ	Univ of California Carnegie Mellon Univ Univ of Texas	Univ of California Stanford University MIT	Univ of California Univ of Washington Univ of Science and Technology of China
Top 3 foreign country collaboration	China India Germany	United Kingdom China Japan	United Kingdom China Canada	United Kingdom China France	China United Kingdom India

pattern is not unlike the corporate giants of the industrial age (Durand, 2024; Freeman & Louçã, 2001). When going abroad, Big Tech engages more with Chinese and British institutions. All the big five, especially Microsoft, reveal this geography of collaboration in their top-tier international partnerships. In fact, China-based co-authors are present in 16 % of the publications for Big Tech as a whole (i.e.

journal and proceedings papers for the aggregate of the five companies), with the UK coming second with 12 %.

Considering all this information, we find that Big Tech's scientific output shares many traits with the most well-received papers in the computer science field. First, publishing in high quartile journals like "NeuroImage" and "IEEE Transactions On Medical Imaging", known for articles with the highest citation levels in the AI subfield (Hughes et al., 2023). Then, when considering the most cited publications in computer science, Big Tech has a matching profile with an increasing average number of authors correlating with increasing citation levels (Teh & Heard, 2021), as well as a strong partnership with high-ranking universities (Stanford, MIT, Carnegie Mellon). Finally, most of the academic authors co-authoring with Big Tech researchers are affiliated with highly cited institutions that perform strongly in the international arena (University of California, University of Texas, among others), resulting in abundant networking opportunities (Jiao et al., 2021).

These patterns suggest a pragmatic foray into research by Big Tech, select the best academic partners and develop work on the fields and subtopics more important to their business, like AI (machine learning, deep learning) or cloud computing (Rikap & Lundvall, 2022). Big Tech seems guided towards research as a consequence of their massive investment in R&D and establishment of a strategic program to relentlessly move on to the following markets (Berger & Brem, 2016; Webster, 2023), but also due to the reasons mentioned by Rotolo et al. (2022), namely building scientific reputation, access to external knowledge via partnerships with academia and talent recruitment. The proclivity to conduct research on AI and the partnerships with key researcher organisations in this area can raise concerns whether these corporations will take over universities as the main research agents on AI, a field already subject to barriers to entry (see Frank, Wang, & Cebrian, 2019) and amenable to disinformation, sorry, research … whether directly, in collaboration, by proxy or, importantly, for the State security apparatus (see also Abdalla & Abdalla, 2021; Thorp & Vinson, 2024; Valletti, 2024).

Major tech players have been awarded major contracts by the US military and intelligence agencies, including for AI and other frontier informational technologies González, 2024; Rikap, 2024). Sought or unsought the unwarranted influence of Big Tech seems to be on the rise, including in the formal political arena. Data collected from the US Federal Election Commission on corporate contributions exposes a number of facts related to large tech firms (Amazon, Apple, Facebook/Meta, Google/Alphabet, Microsoft, Netflix, and Twitter/X): from 2012 to 2024 tech companies their donation grew from USD 2,636,048 to USD 13,936,244, that is, 428.7 %; the greatest jump was from 2016 when all of them broke through the 1 million dollar threshold (OpenSecrets, 2024; Quiver Quantitative, 2024). More than ever, Big Tech became a political economy phenomenon.

7. Conclusions

Big Tech move fast and publish things. The omnipresence of Big Tech, active across many business areas, has now reached the research sector. Understanding how large digital platforms have become remarkably productive and influential knowledge players matters for understanding the evolving science and technology ecosystem in the 21st century.

This paper contributes with a systematic analysis of scientific and technological publications by Big Tech companies (see Rikap & Lundvall, 2021, 2022). It aimed to understand the basic stylised facts of these companies' research. This study focuses only on the large US-based ICT companies which have become the top worldwide corporate R&D spenders. These actors have unleashed global market consequences that have started to be addressed by competition policy (see Durand, 2024). However, now, the fact that they are actively engaged in future-generating activities like research and innovation also reveals them as knowledge powerhouses. Not considering their role in the science sector points to a gap in the multisidedness of the accountability systems required to monitor and evaluate their behaviour (see Rahman, Karunakaran, & Cameron, 2024).

Big Tech publishes extensively and in an increasing variety of fields over time. They have published more than average and have achieved a high influential standing compared to other global big business organisations. All companies are proficient in academic publishing and actively engage with university communities. Overall, we observe a steady increase in the volume, variety and quality of knowledge production, particularly when they became the world's largest corporate R&D spenders. Hence, we can conclude that there is a remarkable focus on publications as research output by Big Tech, not just patents. Google and Microsoft had the highest volumes individually, followed from a distance by Amazon, Facebook and Apple.

"Physical Sciences" and "Computer Sciences" are the dynamic capabilities most common across all Big Tech's output, with the themes of "machine learning", "deep learning", "speech recognition", or "neural networks" standing out, which are often part of their products and services development. Big Tech also has in common that it invests in a diversifying portfolio, having publications in areas such as "Social Sciences", "Mathematics" or "Physics and Astronomy", which can come as a result of their expansion to multiple sectors, as well as a trend by large multinationals to present a wide-ranging research agenda.

Big Tech seems not to show especially expressive partnerships with any particular entities. Notwithstanding, collaborations with universities are observed, primarily American ones. International collaboration mainly occurs with China-based institutions, with UK ones coming next. This bond with academia is expected to continue and even increase, as it happened with Big Pharma (Ma et al., 2022). Questions having to do with fair access to essential research tools and the disclosure obligations regarding strategic dependencies will require the development of new approaches at the intersection of science policy, sectoral regulation and competition policy.

This paper did not explore the motivations or gains that Big Tech accrues from these behaviours; indeed, this facet will be an outstanding issue to investigate in the future. Further probing can also be done on the specific subareas that have a practical end and are directly related to Big Tech's digital products, namely AI and user engagement. Evidence along these lines may matter for the sake of assuring transparency and accountability in an increasingly informational world in which research and innovation are crucial for dynamic competition and for dealing with societal challenges. Throughout, the scientific enterprise should remain contestable in this new era and be on guard against the acquisition of influence by the emerging military-informational complex.

CRediT authorship contribution statement

Sandro Mendonça: Writing – review & editing, Writing – original draft, Visualization, Supervision, Methodology, Investigation, Formal analysis, Conceptualization. **Eduardo Silva:** Writing – review & editing, Writing – original draft, Visualization, Software, Methodology. **Bruno Damásio:** Writing – review & editing, Writing – original draft, Visualization, Supervision, Methodology, Investigation, Formal analysis, Conceptualization.

Acknowledgments

This work was supported by national funds through FCT (Fundação para a Ciência e a Tecnologia), under the project - UIDB/04152/2020 - Centro de Investigação em Gestão de Informação (MagIC)/NOVA IMS (https://doi.org/10.54499/UIDB/04152/2020).

Appendices.

Appendix A – Amazon

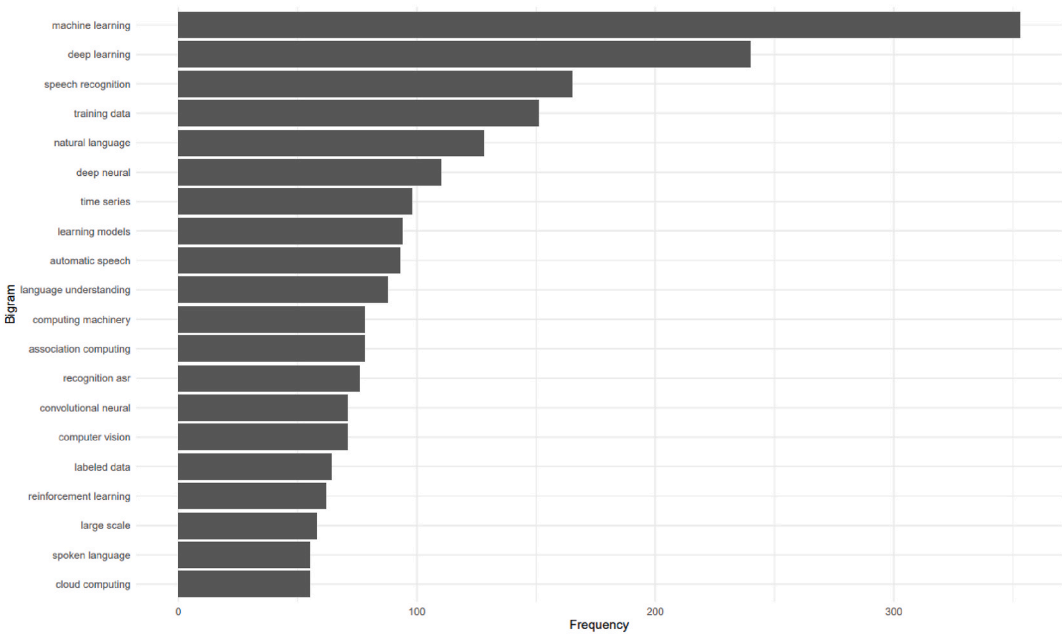

Fig. A1. Amazon Publications' Top 20 Bigrams.

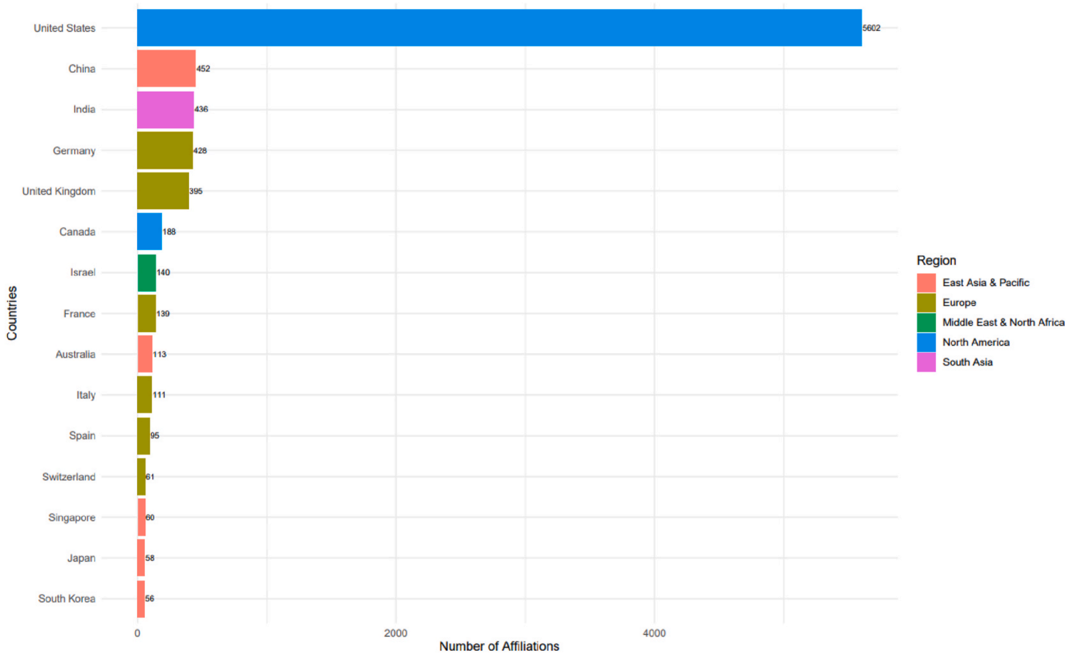

Fig. A2. Amazon publications authors' affiliations country of affiliation.

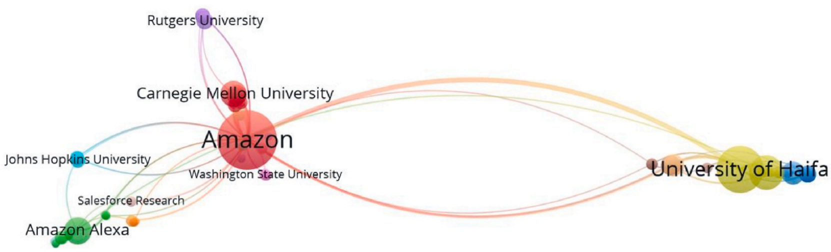

Fig. A3. Amazon network map considering publications of the top 5 most prolific authors.

Appendix B – Graphs

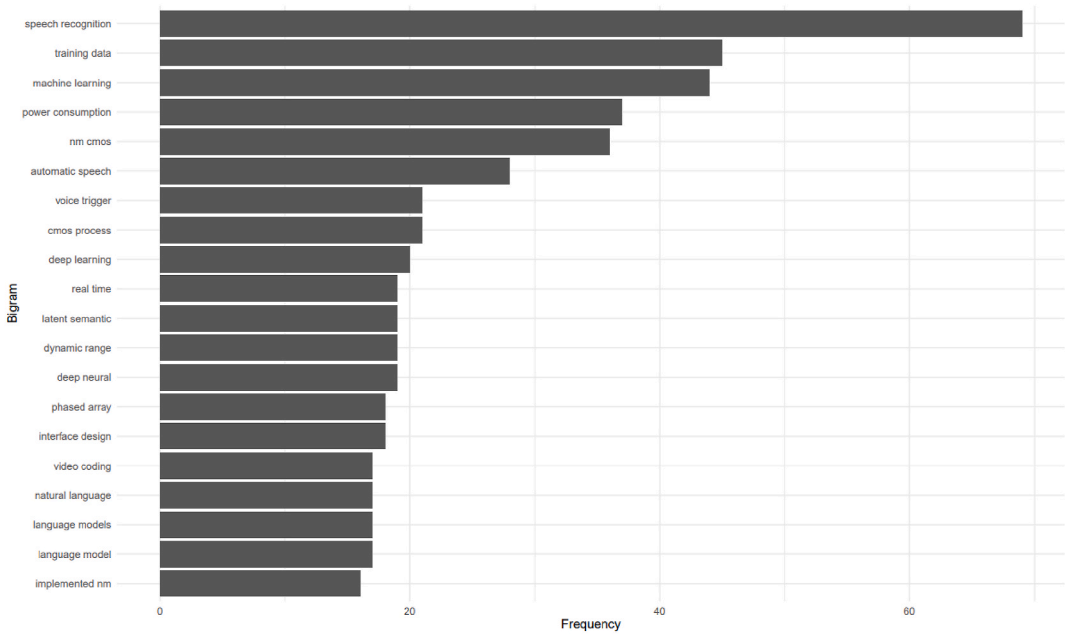

Fig. B1. Apple Publications' Top 20 Bigrams.

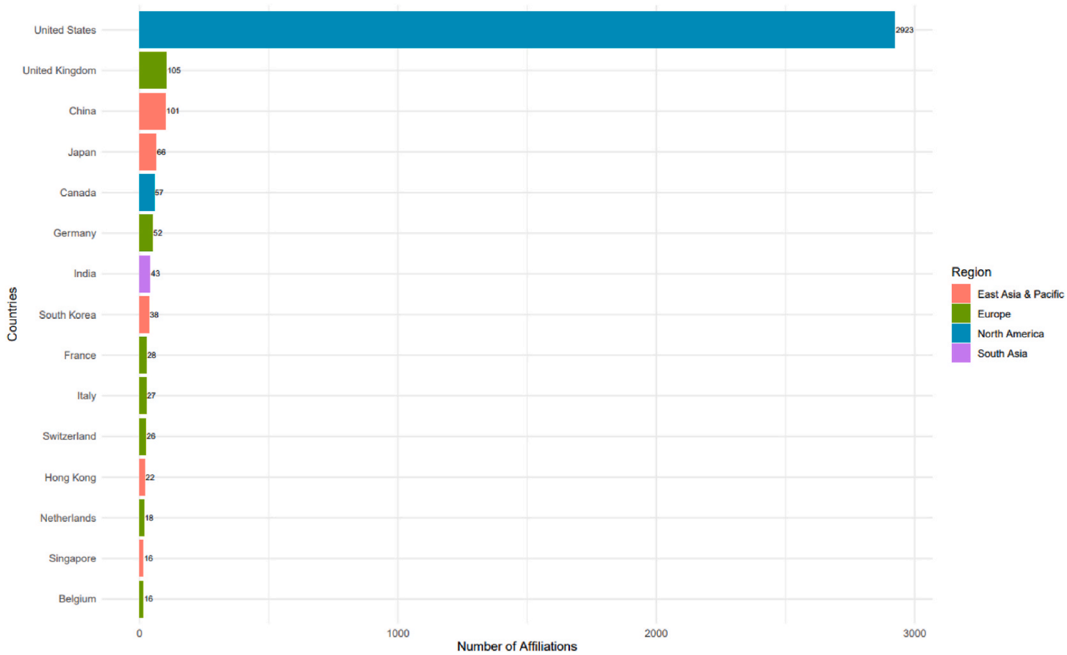

Fig. B2. Apple publications authors' country of affiliation.

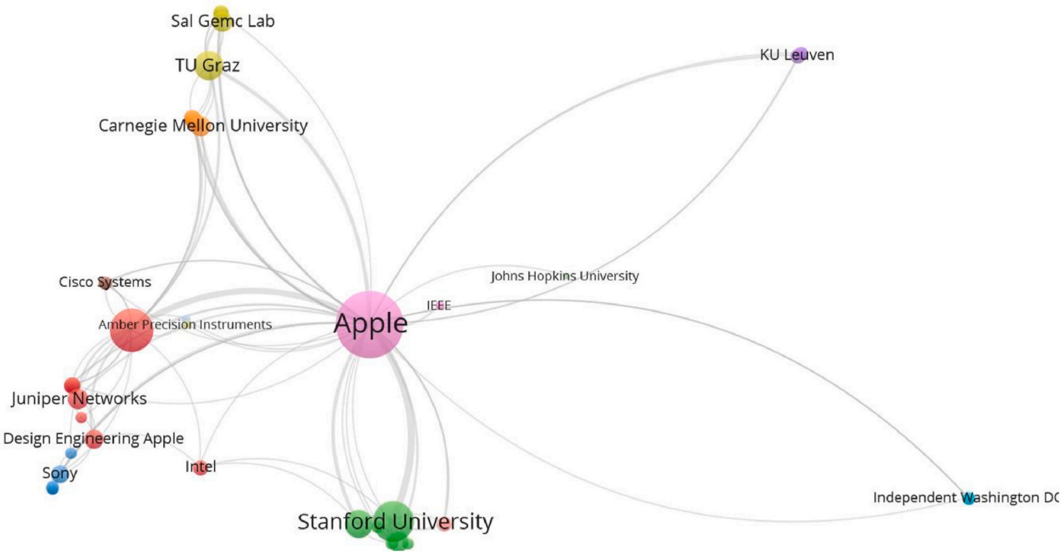

Fig. B3. Apple network map considering publications of the top 5 most prolific authors.

Appendix C – Facebook

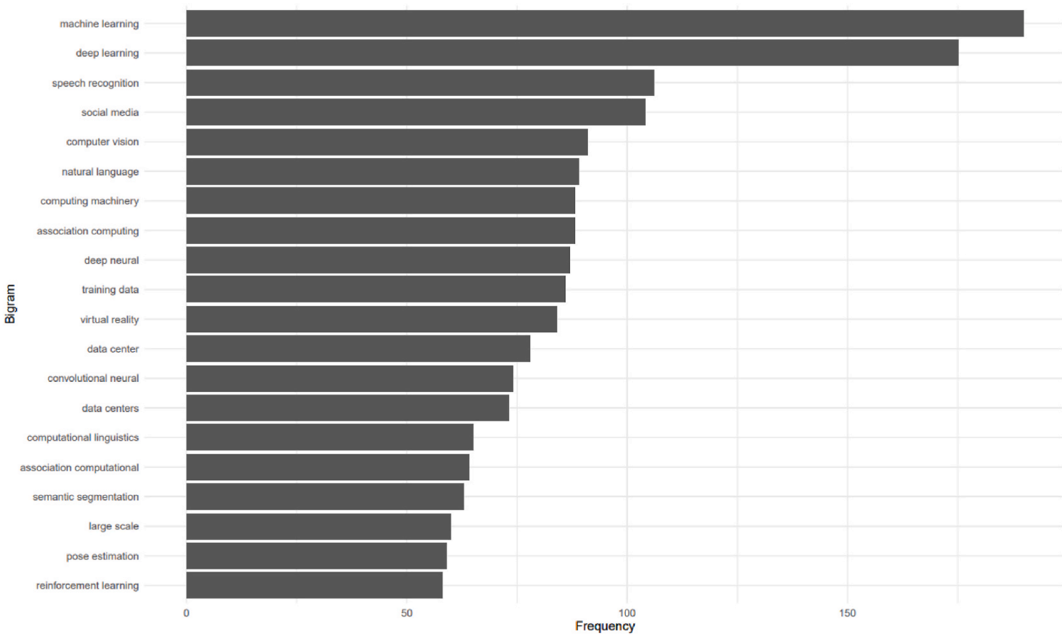

Fig. C1. Facebook Publications' Top 20 Bigrams.

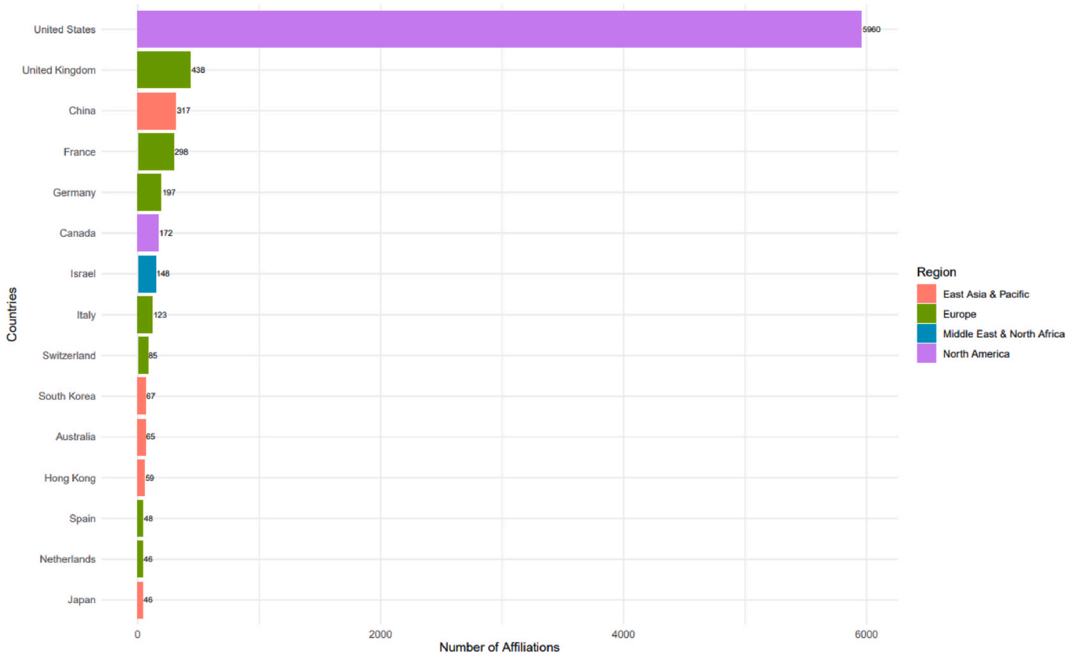

Fig. C2. Facebook publications authors' country of affiliation.

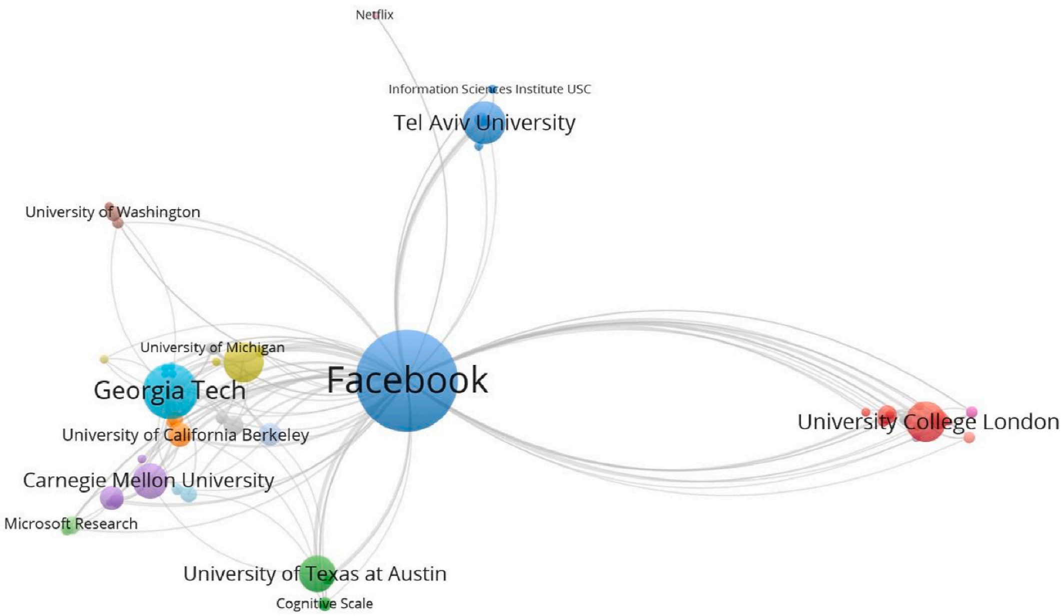

Fig. C3. Facebook network map considering publications of the top 5 most prolific authors.

Appendix D – Google

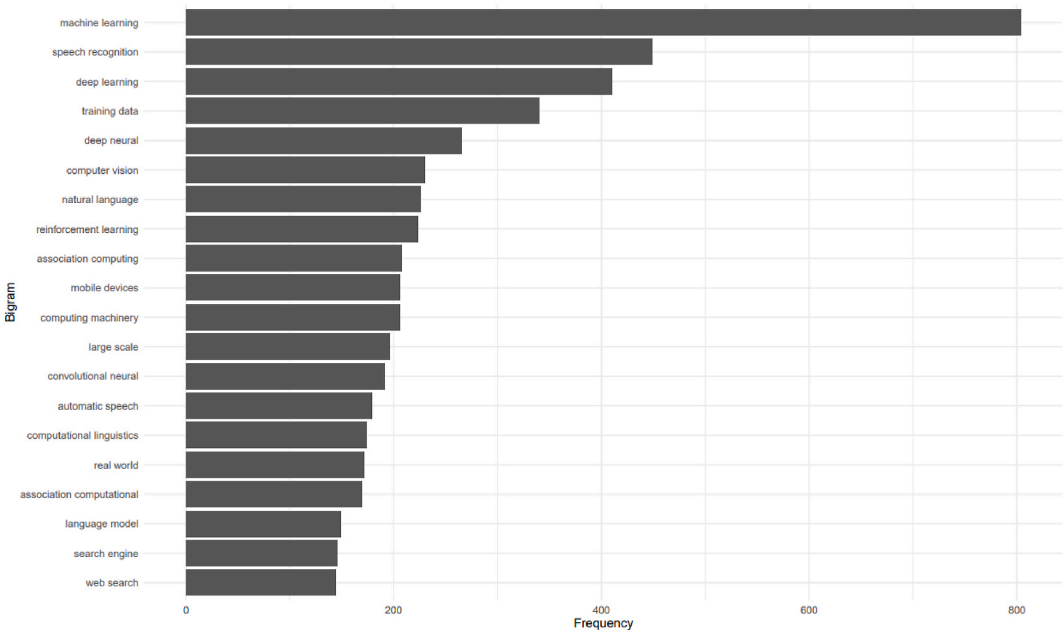

Fig. D1. Google Publications' Top 20 Bigrams.

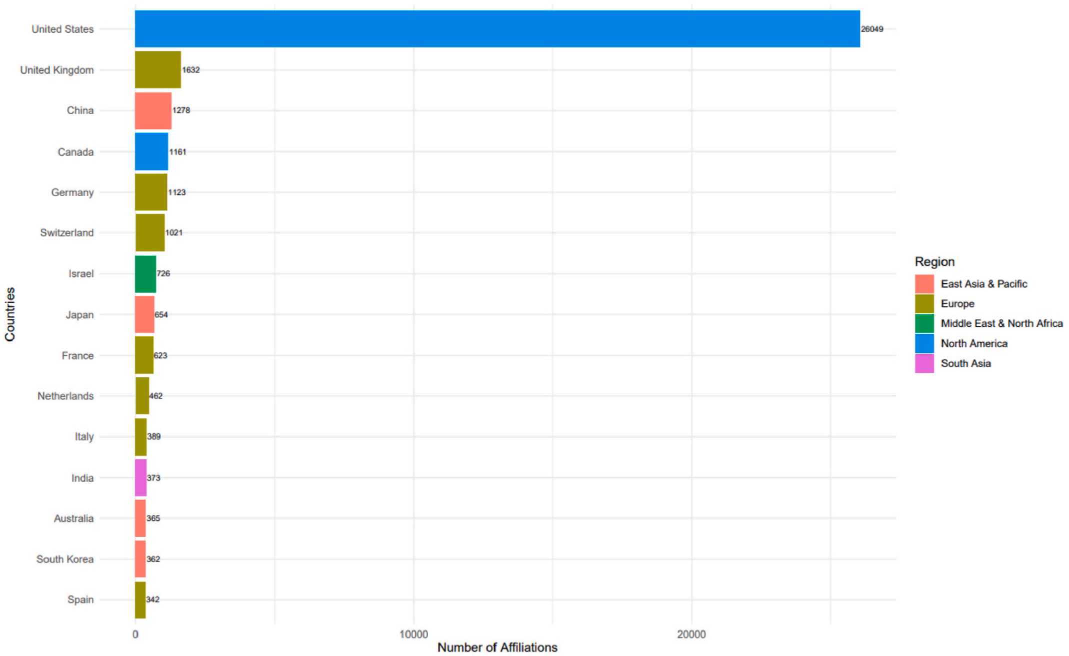

Fig. D2. Google publications authors' country of affiliation.

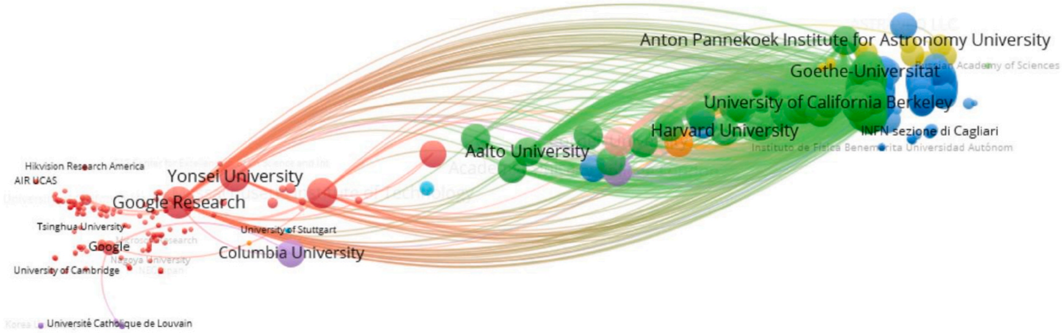

Fig. D3. Google network map considering publications of the top 5 most prolific authors.

Appendix E — Microsoft

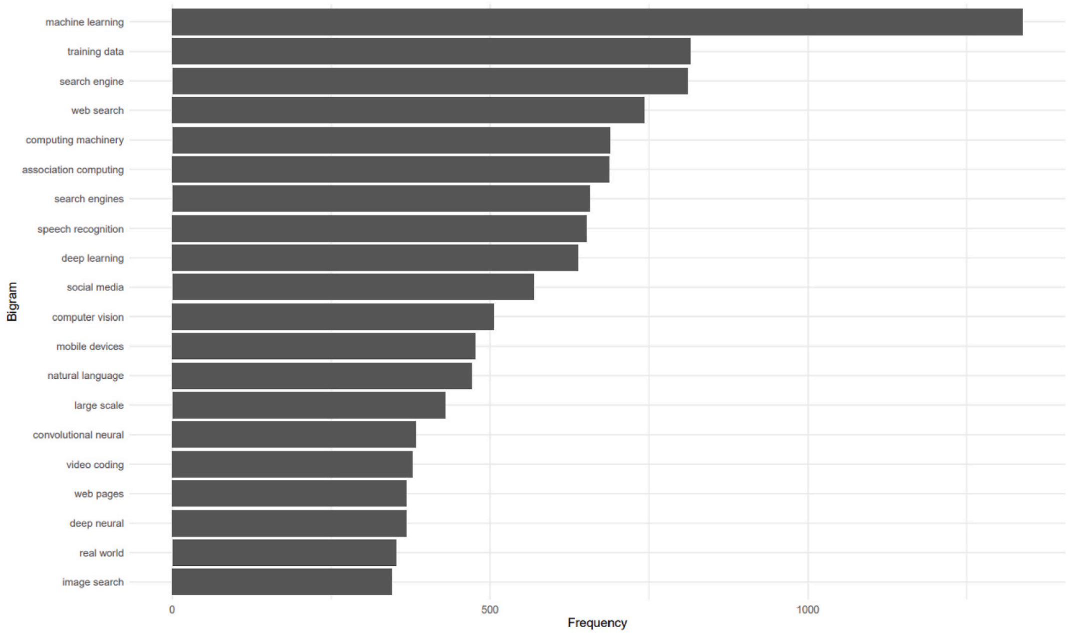

Fig. E1. Microsoft Publications' Top 20 Bigrams.

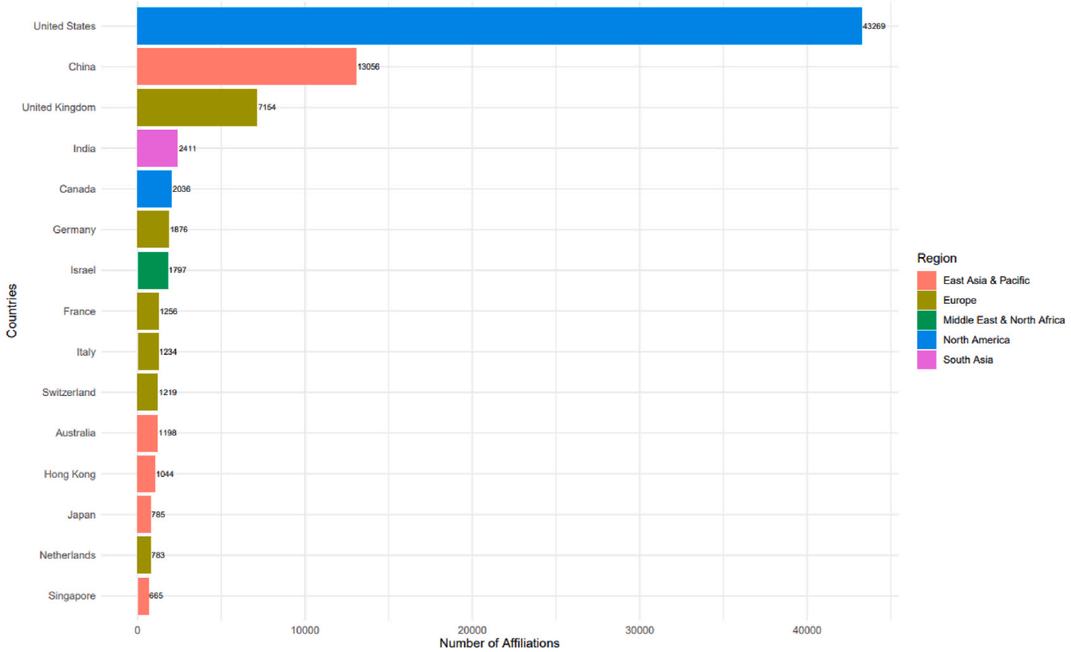

Fig. E2. Microsoft publications authors' country of affiliation.

Fig. E3. Microsoft network map considering publications of the top 5 most prolific authors.

Data availability

Data will be made available on request.

References

Abdalla, M., & Abdalla, M. (2021). The grey hoodie project: Big tobacco, big tech, and the threat on academic integrity. In *AIES 2021 - Proceedings of the 2021 AAAI/ ACM conference on AI, ethics, and society* (pp. 287–297). https://doi.org/10.1145/3461702.3462563

Abdalla, M., Wahle, J. P., Ruas, T., Névéol, A., Ducel, F., Mohammad, S. M., & Fort, K. (2023). The elephant in the room: Analyzing the presence of big tech in natural language processing research. *ArXiv Preprint*. https://doi.org/10.48550/arXiv.2305.02797

Acs, Z. J., Boardman, M. C., & McNeely, C. L. (2013). The social value of productive entrepreneurship. *Small Business Economics, 40*, 785–796. https://doi.org/10.1007/s11187-011-9396-6

Aggarwal, K., Mijwil, M. M., Al-Mistarehi, A. H., Alomari, S., Gök, M., Alaabdin, A. M. Z., & Abdulrhman, S. H. (2022). Has the future started? The current growth of artificial intelligence, machine learning, and deep learning. *Iraqi Journal for Computer Science and Mathematics, 3*(1), 115–123. https://doi.org/10.52866/ijcsm.2022.01.01.013

Ahmed, N., & Wahed, M. (2020). The De-democratization of AI: Deep learning and the compute divide in artificial intelligence research. *ArXiv Preprint*. https://doi.org/10.48550/arXiv.2010.15581

Allen, D. W., Berg, C., Markey-Towler, B., Novak, M., & Potts, J. (2020). Blockchain and the evolution of institutional technologies: Implications for innovation policy. *Research Policy, 49*(1), Article 103865. https://doi.org/10.1016/j.respol.2019.103865

Arora, A., Cohen, W., Lee, H., & Sebastian, D. (2023). Invention value, inventive capability and the large firm advantage. *Research Policy, 52*(1), Article 104650.

Awotunde, J. B., Oluwabukonla, S., Chakraborty, C., Bhoi, A. K., & Ajamu, G. J. (2022). Application of artificial intelligence and big data for fighting COVID-19 pandemic. In *Decision sciences for COVID-19: Learning through case studies* (pp. 3–26). https://doi.org/10.1007/978-3-030-87019-5_1

Barkai, S., & Benzell, S. G. (2024). 70 years of US corporate profits. *Journal of Corporate Finance, 87*, Article 102622.

Basak, E., & Calisir, F. (2015). Publication trends in Facebook: A scientometric study, 170–173 https://doi.org/10.15242/icehm.ed0315075.

Behrens, T. R., & Gray, D. O. (2001). Unintended consequences of cooperative research: Impact of industry sponsorship on climate for academic freedom and other graduate student outcome. *Research Policy, 30*(2), 179–199. https://doi.org/10.1016/S0048-7333(99)00112-2

Berger, J. M., & Baker, C. M. (2014). Bibliometrics: An overview. *RGUHS Journal of Pharmaceutical Sciences, 14*(4), 81–92. https://doi.org/10.5530/rjps.2014.3.2

Belk, R. (2010). Sharing. *Journal of Consumer Research, 36*(5), 715–734. https://doi.org/10.1086/612649

Berger, A., & Brem, A. (2016). Innovation hub how-to: Lessons from silicon valley. *Global Business and Organizational Excellence, 35*(5), 58–70. https://doi.org/10.1002/joe.21698

Berryman, R. M. (2014). *Amazon. com, Inc.: A case study analysis*. . (Accessed 11 November 2020).

Birch, K., & Cochrane, D. T. (2022). Big tech. Four emerging forms of digital rentiership. *Science as Culture, 31*(1), 44–58. https://doi.org/10.1080/09505431.2021.1932794.

Birch, K., & Ward, C. (2023). Introduction: Critical approaches to rentiership. *Environment and Planning A: Economy and Space, 55*(6), 1429–1437. https://doi.org/10.1177/0308518X231162363

Birhane, A., Kalluri, P., Card, D., Agnew, W., Dotan, R., & Bao, M. (2022). The values encoded in machine learning research. In *Proceedings of the 2022 ACM conference on fairness, accountability, and transparency* (pp. 173–184). https://doi.org/10.1145/3531146.3533083

Bork, R. H., & Sidak, J. G. (2012). What does the Chicago school teach about internet search and the antitrust treatment of Google? *Journal of Competition Law and Economics, 8*(4), 663–700. https://doi.org/10.1093/joclec/nhs031

Bourreau, M., & de Streel, A. (2020). *Big tech acquisitions: Competition & innovation effects and EU merger control*.

Brynjolfsson, E., & Smith, M. D. (2000). Frictionless commerce? A comparison of internet and conventional retailers. *Management Science, 46*(4), 563–585. https://doi.org/10.1287/mnsc.46.4.563.12061

Budzinski, O., & Mendelsohn, J. (2021). Regulating big tech: From competition policy to sector regulation? *ORDO, 72*(1), 215–255. https://doi.org/10.1515/ordo-2023-2015

Cabanac, G., Oikonomidi, T., & Boutron, I. (2001). Day-to-day discovery of preprint–publication links. *Scientometrics, 126*, 5285–5304. https://doi.org/10.1007/s11192-021-03900-7

Cabral, L., Haucap, J., Parker, G., Petropoulos, G., Valletti, T., & Alstyne, M. van (2021). In *The EU digital markets act: A report from a panel of economic experts*. Publications Office of the European Union. https://doi.org/10.2760/139337.

Camerani, R., Grassano, N., & Rotolo, D. (2023, April). Is corporate science growing or declining?. *27th International Conference on Science, Technology and Innovation Indicators (STI 2023)*. International Conference on Science, Technology and Innovation Indicators.

Cano, V., & Lind, N. C. (1991). Citation life cycles of ten citation classics. *Scientometrics, 22*, 297–312. https://doi.org/10.1007/BF02020003

Carroll, E., & Romano, J. (2010). *Your digital afterlife: When Facebook, Flickr and Twitter are your estate, what's your legacy? New Riders*.

Castells, M. (1997). An introduction to the information age. *City, 2*(7), 6–16. https://doi.org/10.1080/13604819708900050

Cennamo, C., Kretschmer, T., Constantinides, P., Alaimo, C., & Santaló, J. (2023). Digital platforms regulation: An innovation-centric view of the EU's Digital Markets Act. *Journal of European Competition Law & Practice, 14*(1), 44–51. https://doi.org/10.1093/jeclap/lpac043

Commission, A. C. and C.. (2019). *Digital platforms inquiry: Final report*.

Coursaris, C. K., & Van Osch, W. (2014). A scientometric analysis of social media research (2004– 2011). *Scientometrics, 101*(1), 357–380. https://doi.org/10.1007/s11192-014-1399-z

Crémer, J., de Montjoye, Y. A., & Schweitzer, H. (2019). *Competition policy*.

Creser, O. T. (2020). In antitrust we trust?: Big tech is not the problem-it's weak data privacy protections. *Federal Common Law Journal, 73*, 289.

Crisanto, J. C., Ehrentraud, J., Lawson, A., & Restoy, F. (2021). Big tech regulation: What is going on?. In *Bank for international settlements*. Financial Stability Institute.

de Freitas, L. C., & de Moura Filho, R. N. (2022). Aesthetic normalization of gender in the Instagram application: A portrait of the Brazilian woman. *Computer Law & Security Review, 47*, 105753. https://doi.org/10.1016/j.clsr.2022.105753

de Streel, A., & Larouche, P. (2004). Distruptive innovation and antitrust. *GW competition & innovation - Lab working paper series no. 2024/24*.

Dolata, U. (2017). Apple, Amazon, Google, Facebook, Microsoft: Market concentration - Competition - Innovation strategies. *Econstor, 33*. http://www.uni-stuttgart.de/soz/oi/publikationen/.

Donthu, N., Kumar, S., Mukherjee, D., Pandey, N., & Lim, W. M. (2021). How to conduct a bibliometric analysis: An overview and guidelines. *Journal of Business Research, 133*(March), 285–296. https://doi.org/10.1016/j.jbusres.2021.04.070

Dosi, G., Marengo, L., Staccioli, J., & Virgillito, M. E. (2023). Big pharma and monopoly capitalism: a long-term view. *Structural Change and Economic Dynamics, 65*, 15–35. https://doi.org/10.1016/j.strueco.2023.01.004

Drott, M. C. (1995). Reexamining the role of conference papers in scholarly communication. *Journal of the American Society for Information Science, 46*(4), 299–305. https://doi.org/10.1002/(SICI)1097-4571(199505)46:4<299::AID-ASI6>3.0.CO;2-0

Durand, C. (2024). *How silicon valley unleashed techno-feudalism: The making of the digital economy*. Verso.

Dyvik, E. H. (2022). *Ranking of the companies with the highest spending on research and development worldwide in 2022 (in billion U.S. dollars)*. Statista. https://www.statista.com/statistics/265645/ranking-of-the-20-companies-with-the-highest-spending-on-research-and-development/.

Ehrentraud, J., Evans, J. L., Monteil, A., & Restoy, F. (2022). *Big tech regulation: In search of a new framework*. *Financial Stability Institute,* Bank for International Settlements.

Engel, J. S. (2015). Global clusters of innovation: Lessons from Silicon Valley. *California Management Review, 57*(2), 36–65. https://doi.org/10.1525/cmr.2015.57.2.36

Fagerberg, J. B. R. M., & E, S. A. (2013). *Innovation studies: Evolution and future challenges*. Oxford University Press.

Farrukh, M., Javed, S., Raza, A., & Lee, J. W. C. (2022). Twenty years of green innovation research: Trends and way forward. *World Journal of Entrepreneurship, Management and Sustainable Development, 17*(3), 488–501.

Foster, J. B., & McChesney, R. (2014). Surveillance capitalism. *Monthly Review, 66*(3), 1–31.

Frank, M. R., Wang, D., Cebrian, M., et al. (2019). The evolution of citation graphs in artificial intelligence research. *Nature Machine Intelligence, 1*, 79–85. https://doi.org/10.1038/s42256-019-0024-5

Freeman, C., & Louçã, F. (2001). *As time goes by: From the industrial revolutions to the information revolution*. Oxford University Press.

Frenken, K., & Fuenfschilling, L. (2021). The rise of online platforms and the triumph of the corporation. *Sociologica, 14*(3), 101–113. https://doi.org/10.6092/issn.1971-8853/11715

Frenken, K., van Waes, A., Pelzer, P., Smink, M., & van Est, R. (2020). Safeguarding public interests in the platform economy. *Policy & Internet, 12*(3), 400–425. https://doi.org/10.1002/poi3.217

Georgescu, I. (2022). Bringing back the golden days of bell labs. *Nature Reviews Physics, 4*(2), 76–78. https://doi.org/10.1038/s42254-022-00433-7

Geroski, P. (2003). Competition in markets and competition for markets. *Journal of Industry, Competition and Trade, 3*(3), 51–166. https://doi.org/10.1023/A:1027457020332

Giziński, S., Kaczyńska, P., Ruczyński, H., Wiśnios, E., Pieliński, B., Biecek, P., & Sienkiewicz, J. (2023). Big Tech influence over AI research revisited: Memetic analysis of attribution of ideas to affiliation. *ArXiv Preprint*.

Glänzel, W., Moed, H. F., Schmoch, U., & Thelwall, M. (2019). In M. Glänzel, W. Moed, H. F. Schmoch, U, & Thelwall (Eds.), *Springer handbook of science and technology indicators*. Cham: Springer. https://doi.org/10.1007/978-3-030-02511-3.

Gómez-Barroso, J. L., & Marbán-Flores, R. (2020). Telecommunications and economic development–the 21st century: Making the evidence stronger. *Telecommunications Policy, 44*(2), Article 101905.

González, R. J. (2024). How big Tech and silicon valley are transforming the military-industrial complex. *Watson institute for international and public affairs*. Brown University.

González, M. R., Gasco, J., & Llopis, J. (2016). Facebook and academic performance: A positive outcome. *The Anthropologist, 23*(1–2), 59–67. https://doi.org/10.1080/09720073.2016.11891924

Gorodnichenko, Y., Pham, T., & Talavera, O. (2021). Conference presentations and academic publishing. *Economic Modelling, 95*, 228–254. https://doi.org/10.1016/j.econmod.2020.12.017

Gupta, B. M., Dhawan, S. M., Gupta, R., & Jalana, M. (2015). *Facebook research: A scientometric assessment of global publications, 2005-14, 1*. Library Philosophy and Practice.

Hein, A., Schreieck, M., Riasanow, T., Setzke, D. S., Wiesche, M., Böhm, M., & Krcmar, H. (2020). Digital platform ecosystems. *Electronic Markets, 30*(1), 87–98. https://doi.org/10.1007/s12525-019-00377-4

Hellman, J. (2022). Big tech's 'voracious appetite,'or entrepreneurs who dream of acquisition? Regulation and the interpenetration of corporate scales. *Science as Culture, 31*(1), 149–161. https://doi.org/10.1080/09505431.2021.2000597

Hicks, D. (1995). Published papers, tacit competencies and corporate management of the public/private character of knowledge. *Industrial and Corporate Change, 4*(2), 401–424. https://doi.org/10.1093/icc/4.2.401

Hoberg, G., & Phillips, G. M. (2022). *Scope, scale and concentration: The 21st century firm (No. w30672)*. National Bureau of Economic Research.

Hughes, H., O'Reilly, M., McVeigh, N., & Ryan, R. (2023). The top 100 most cited articles on artificial intelligence in radiology: A bibliometric analysis. *Clinical Radiology, 78*(2), 99–106. https://doi.org/10.1016/j.crad.2022.09.133

Jacobides, M. G. (2019). In the ecosystem economy, what's your strategy? *Harvard Business Review, 97*(5), 128–137.

Jacobides, M. G., & Lianos, I. (2021). Regulating platforms and ecosystems: An introduction. *Industrial and Corporate Change, 30*(5), 1131–1142. https://doi.org/10.1093/icc/dtab060

Jacsó, P. (2005). Google scholar: The pros and the cons. *Online Information Review, 29*(2), 208–214. https://doi.org/10.1108/14684520510598066

Jiao, F., Fang, J., Ci, Y., & Tu, W. (2021). Bibliometric analysis based on highly cited papers in computer science. *Journal of Physics: Conference Series, 1883*(2), Article 012053. https://doi.org/10.1088/1742-6596/1883/1/012053DownloadArticlePDF

Johnson, K., Li, Y., Phan, H., Singer, J., & Trinh, H. (2012). *The innovative success that is apple*. Inc.

Jorgenson, D. W. (2001). Information technology and the US economy. *The American Economic Review, 91*(1), 1–32. https://doi.org/10.1257/aer.91.1.1

Jurowetzki, R., Hain, D., Mateos-Garcia, J., & Stathoulopoulos, K. (2021). The privatization of AI research (-ers): Causes and potential Consequences–From university-industry interaction to public research brain-drain? *ArXiv Preprint*. https://doi.org/10.48550/arXiv.2102.01648

Kahn, R. (2017). Hate speech, democratic legitimacy and the age of trump. *International and Comparative Law Review, 17*(1), 239–253.

Katz, J. S., & Ronda-Pupo, G. A. (2019). Cooperation, scale-invariance and complex innovation systems: A generalization. *Scientometrics, 121*(1), 1045–1065. https://doi.org/10.1007/s11192-019-03215-8

Kenney, M., & Zysman, J. (2019). Work and value creation in the platform economy. *Work and Labor in the Digital Age*, 13–41.

Khurana, P., Ganesan, G., Kumar, G., & Sharma, K. (2022). A comparative analysis of unified informetrics with scopus and web of science. *Journal of Scientometric Research, 11*(2), 146–154. https://doi.org/10.5530/jscires.11.2.16

Kim, M., Zimmermann, T., DeLine, R., & Begel, A. (2016). The emerging role of data scientists on software development teams. In *Proceedings of the 38th international conference on software engineering* (pp. 96–107). https://doi.org/10.1145/2884781.2884783

Kim, M., Zimmermann, T., DeLine, R., & Begel, A. (2017). Data scientists in software teams: State of the art and challenges. *IEEE Transactions on Software Engineering, 44*(11), 1024–1038. https://doi.org/10.1109/TSE.2017.2754374

Klinge, T. J., Hendrikse, R., Fernandez, R., & Adriaans, I. (2023). Augmenting digital monopolies: A corporate financialization perspective on the rise of big tech. *Competition and Change, 27*(2), 332–353. https://doi.org/10.1177/10245294221105573

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

Lam, A. (2011). What motivates academic scientists to engage in research commercialization:'Gold','ribbon'or 'puzzle'? *Research Policy, 40*(10), 1354–1368. https://doi.org/10.1016/j.respol.2011.09.002

Lee, A., Taylor, P., Kalpathy-Cramer, J., & Tufail, A. (2017). Machine learning has arrived. *Ophthalmology, 124*(12), 1726–1728. https://doi.org/10.1016/j.ophtha.2017.08.046

Lianos, I. (2022). Value extraction and institutions in digital capitalism: Towards a law and political economy synthesis for competition law. *European Law Open, 1*(4), 852–890. https://doi.org/10.1017/elo.2023.2

Lisée, C., Larivière, V., & Archambault, É. (2008). Conference proceedings as a source of scientific information: A bibliometric analysis. *Journal of the American Society for Information Science and Technology, 59*(11), 1776–1784.

Liu, W. (2021). Caveats for the use of Web of Science Core Collection in old literature retrieval and historical bibliometric analysis. *Technological Forecasting and Social Change, 172*, 121023. https://doi.org/10.1016/j.techfore.2021.121023

Louçã, F., & Mendonça, S. (2002). Steady change: The 200 largest US manufacturing firms throughout the 20th century. *Industrial and Corporate Change, 11*(4), 817–845. https://doi.org/10.1093/icc/11.4.817

Lubinki, F. (2004). Simple propositions for competition law in the moder economy, competition law insight. https://www.competitionlawinsight.com/regulatory/complexity-continued–1.htm.

Ma, J. (2022). Emerging digital markets and regulation. In *Regulating data monopolies: A law and economics perspective* (pp. 17–54). Singapore: Springer Nature Singapore. https://doi.org/10.1007/978-981-16-8766-2_2.

Ma, Z., Augustijn, K., de Esch, I. J., & Bossink, B. (2022). Collaborative university–industry R&D practices supporting the pharmaceutical innovation process: Insights from a bibliometric review. *Drug Discovery Today, 27*(8), 2333–2341. https://doi.org/10.1016/j.drudis.2022.05.001

Ma, Y., Pugsley, B., & Zimmermann, B. (2024). *Superstar firms through the generations*. Mimeo.

Manganelli, A., & Nicita, A. (2022). *Regulating digital markets: The European approach. Palgrave macmillan Cham*. https://doi.org/10.1007/978-3-030-89388-0

Marciano, A., Nicita, A., & Ramello, G. B. (2020). Big data and big techs: Understanding the value of information in platform capitalism. *European Journal of Law and Economics, 50*, 345–358. https://doi.org/10.1007/s10657-020-09675-1

Marginson, S. (2022). Global science and national comparisons: beyond bibliometrics and scientometrics. *Comparative Education, 58*(2), 125–146. https://doi.org/10.1080/03050068.2021.1981725

Martínez, D. A., & Bunkanwanicha, P. (2024). Challenges with the rise of superstar firms. *ESCP impact paper No.2022-23-EN*.

Mathew, S., & Varia, J. (2014). Overview of Amazon web services. *Amazon Whitepapers, 105*, 1–22.

Mendonça, S. (2004). Large companies from all sectors patenting in ICT: Is there a link between corporate technological diversification and the information revolution? In J. Ö. Cantwell, A. Gambardella, & Ö. Granstrand (Eds.), *The economics and management of technological diversification* (pp. 227–257). Routledge.

Mendonça, S. (2006). The revolution within: ICT and the shifting knowledge base of the world's largest companies. *Economics of Innovation and New Technology, 15*(8), 777–799. https://doi.org/10.1080/10438590500510442

Mendonça, S., Archibugi, D., Gerbrandy, A., & Tsipouri, L. (2024). *Futures of big tech in Europe: Scenarios and policy implications*. foresight. https://doi.org/10.2777/93885

Merges, R. P. (2020). Patent markets and innovation in the era of big platform companies. *Berkeley Technology Law Journal, 35*(1), 53–112.

Mingers, J., & Leydesdorff, L. (2015). A review of theory and practice in scientometrics. *European Journal of Operational Research, 246*(1), 1–19. https://doi.org/10.1016/j.ejor.2015.04.002

Monsees, L., Liebetrau, T., Austin, J. L., Leander, A., & Srivastava, S. (2023). Transversal politics of big tech. *International Political Sociology, 17*(1). https://doi.org/10.1093/ips/olac020

Mowery, D. C., & Nelson, R. R. (Eds.). (1999). *Sources of industrial leadership: Studies of seven industries*. Cambridge: Cambridge University Press.

Mucha, T., & Seppala, T. (2020). *Artificial intelligence Platforms–a new research agenda for digital platform economy, 76*.

Nelson, R. R., Dosi, G., Helfat, C. E., Pyka, A., & Sinter, S. (2018). *Modern evolutionary economics: An overview*.
Nelson, R. R., & Winter, S. G. (1982). The schumpeterian tradeoff revisited. *The American Economic Review, 72*(1), 114–132. https://www.jstor.org/stable/1808579.
Nentwich, M., & König, R. (2014). Academia goes Facebook? The potential of social network sites in the scholarly realm. In *Opening science: The evolving guide on how the internet is changing research* (pp. 107–124). collaboration and scholarly publishing. https://doi.org/10.1007/978-3-319-00026-8_7.
Nicoli, N., & Iosifidis, P. (2023). EU digital economy competition policy: From ex-post to ex- ante. The case of alphabet, Amazon, Apple, and Meta. *Global Media and China, 8*(1), 24–38. https://doi.org/10.1177/20594364231152673
Niu, L. (2019). Using Facebook for academic purposes: Current literature and directions for future research. *Journal of Educational Computing Research, 56*(8), 1384–1406. https://doi.org/10.1177/0735633117745161
Nuccio, M., & Guerzoni, M. (2019). Big data: Hell or heaven? Digital platforms and market power in the data-driven economy. *Competition and Change, 23*(3), 312–328. https://doi.org/10.1177/1024529418816525
OECD. (2023). G7 inventory of new rules for digital markets: Analytical note. https://www.oecd.org/competition/analytical-note-on-the-G7-inventory-of-new-rules-for-digital-markets%022023.pdf.
Oliner, S. D., & Sichel, D. E. (2000). The resurgence of growth in the late 1990s: Is information technology the story? *The Journal of Economic Perspectives, 14*(4), 3–22. https://doi.org/10.1257/jep.14.4.3
OpenSecrets. (2024). Who are the biggest organization donors?. https://www.opensecrets.org/elections-overview/top-organizations?cycle=2024.
Ozalp, H., Ozcan, P., Dinckol, D., Zachariadis, M., & Gawer, A. (2022). "Digital colonization" of highly regulated industries: an analysis of big tech platforms' entry into health care and education. *California Management Review, 64*(4), 78–107. https://doi.org/10.1177/000812562210943
Paredes-Frigolett, H., & Pyka, A. (2023). The global stakeholder capitalism model of digital platforms and its implications for strategy and innovation from a Schumpeterian perspective. *Journal of Evolutionary Economics, 32*(2), 463–500. https://doi.org/10.1007/s00191-022-00760-z
Petit, N. (2020). *Big Tech and the Digital Economy: The Moligopoly Scenario*. Oxford University Press. https://doi.org/10.1093/oso/9780198837701.001.0001.
Phoa, P., & Gerbrandy, A. (2024). Regulating the discursive power of Big Tech companies. *VerfBlog*, 2024/7–202415. https://doi.org/10.59704/bfcfcba054a11467
Powell, W. W., & Grodal, S. (2006). Networks of innovators. In R. R. Fagerberg, J. Mowery, & D. C. Nelson (Eds.), *The Oxford handbook of innovation*. Oxford University. https://doi.org/10.1093/oxfordhb/9780199286805.003.0003.
Powell, W. W., & Snellman, K. (2004). The knowledge economy. *Annual Review of Sociology, 30*, 199–220. https://www.jstor.org/stable/29737691.
Prakash, D., Bisla, M., & Rastogi, S. G. (2021). Understanding authentic leadership style: The Satya Nadella microsoft approach. *Open Journal of Leadership, 10*, 95–109. https://doi.org/10.4236/ojl.2021.102007
Quiver Quantitative. (2024). Corporate election contributions dashboard. https://www.quiverquant.com/election-contributions.
Rafols, I., Hopkins, M. M., Hoekman, J., Siepel, J., O'Hare, A., Perianes-Rodríguez, A., & Nightingale, P. (2014). Big pharma, little science?: A bibliometric perspective on big Pharma's R&D decline. *Technological Forecasting and Social Change, 81*, 22–38. https://doi.org/10.1016/j.techfore.2012.06.007
Rahman, H. A., Karunakaran, A., & Cameron, L. D. (2024). Taming platform power: Taking accountability into account in the management of platforms. *Academy of Management Annals, 18*(1), 251–294. https://doi.org/10.5465/annals.2022.0090
Rikap, C. (2023a). Capitalism as usual? Implications of digital intellectual monopolies. *New Left Review, 139*, 145–160.
Rikap, C. (2023b). *Same end by different means: Google, Amazon, Microsoft and Meta's strategies to organize their frontier AI innovation systems CITYPERC Working Paper no. 2023-03*.
Rikap, C. (2024). The US national security state and big tech: Frenemy relations and in-novation planning in turbulent times. *Review of Keynesian Economics, 12*(3), 348–364.
Rikap, C., & Lundvall, B.Å. (2021). *The digital innovation race: Conceptualizing the emerging new world order*. Palgrave MacMillan. https://doi.org/10.1007/978-3-030-89443-6_2
Rikap, C., & Lundvall, B.Å. (2022). Big tech, knowledge predation and the implications for development. *Innovation and Development, 1–28*. https://doi.org/10.1080/2157930X.2020.1855825
Roosendaal, A. (2011). Facebook tracks and traces everyone: Like this! *Tilburg Law School Legal Studies Research Paper Series, 3*.
Rotolo, D., Camerani, R., Grassano, N., & Martin, B. R. (2022). Why do firms publish? A systematic literature review and a conceptual framework. *Research Policy, 51*(10), Article 104606. https://doi.org/10.1016/j.respol.2022.104606
Rowe, N. (2018). 'When you get what you want, but not what you need': The motivations, affordances and shortcomings of attending academic/scientific conferences. *International Journal of Research in Education and Science, 4*(2), 714–729. https://doi.org/10.21890/ijres.438394
Sabatier, M., & Chollet, B. (2017). Is there a first mover advantage in science? Pioneering behavior and scientific production in nanotechnology. *Research Policy, 46*(2), 522–533. https://doi.org/10.1016/j.respol.2017.01.003
Santesteban, C., & Longpre, S. (2020). How big data confers market power to big tech: Leveraging the perspective of data science. *Antitrust Bulletin, 65*(3), 459–485. https://doi.org/10.1177/0003603X2093421
Satariano, A., & Metz, C. (2023). Amazon invests up to $4 billion in AI Start-Up. *The New York Times*, 26 Sept. 2023, B4 link.gale.com/apps/doc/A766589749/AONE?u=anoñ8a60ec3b&sid=googleScholar&xid= bfd97c2b.
Schleyer, T., Spallek, H., Butler, B. S., Subramanian, S., Weiss, D., Poythress, M. L., & Mueller, G. (2008). Facebook for scientists: Requirements and services for optimizing how scientific collaborations are established. *Journal of Medical Internet Research, 10*(3), e24.
Schuhmacher, A., Gatto, A., Kuss, M., Gassmann, O., & Hinder, M. (2021). Big techs and startups in pharmaceutical R&D–A 2020 perspective on artificial intelligence. *Drug Discovery Today, 26*(10), 2226–2231.
Schuhmacher, A., Haefner, N., Honsberg, K., Goldhahn, J., & Gassmann, O. (2023). The dominant logic of big tech in healthcare and pharma. *Drug Discovery Today, 28*(2), Article 103457. https://doi.org/10.1016/j.drudis.2022.103457
Scott Morton, F., Bouvier, P., Ezrachi, A., Jullien, B., Katz, R., Kimmelman, G., … Morgenstern, J. (2019). *Committee for the study of digital platforms: Market structure and antitrust subcommittee report*.
Shapiro, C., & Varian, H. R. (1999). *Information rules: A strategic guide to the network economy*. Harvard Business Press.
Sharma, N., Sharma, R., & Jindal, N. (2021). Machine learning and deep learning applications-a vision. In *Global transitions proceedings, 2* pp. 24–28). https://doi.org/10.1016/j.gltp.2021.01.004, 1.
Shelanski, H. A. (2013). Information, innovation, and competition policy for the internet. *University of Pennsylvania Law Review, 161*, 1663–1705.
Silva, V., Gellert, R., & Borgesius, F. Z. (2024). The state in the platform economy: A typology of alternative approaches. *SSRN*.
Snowden, E. (2019). *Permanent Record: A Memoir of a Reluctant Whistleblower*. Macmillan.
Stallkamp, M., & Schotter, A. P. (2021). Platforms without borders? The international strategies of digital platform firms. *Global Strategy Journal, 11*(1), 58–80.
Teece, D. J. (2014). A dynamic capabilities-based entrepreneurial theory of the multinational enterprise. *Journal of International Business Studies, 45*, 8–37. https://doi.org/10.1057/jibs.2013.54
Teece, D. J. (2018). Profiting from innovation in the digital economy: Enabling technologies, standards, and licensing models in the wireless world. *Research Policy, 47*(8), 1367–1387. https://doi.org/10.1016/j.respol.2017.01.015
Teece, D. J. (2023). Big tech and strategic management: How management scholars can inform competition policy. *Academy of Management Perspectives, 37*(1), 1–15. https://doi.org/10.5465/amp.2022.0013
Teh, P. L., & Heard, P. (2021). Five hundred most-cited papers in the computer sciences: Trends, relationships and common factors. *Trends and Applications in Information Systems and Technologies, 2*(9), 13–22. https://doi.org/10.1007/978-3-030-72651-5_2
Thomason, J. (2021). Big tech, big data and the new world of digital health. *Global Health Journal, 5*(4), 165–168. https://doi.org/10.1016/j.glohj.2021.11.003
Thorp, H. H., & Vinson, V. (2024). Context matters in social media. *Science, 385*(671), 139.
Tijssen, R. J., & Van Raan, A. F. (1994). Mapping changes in science and technology: Bibliometric co-occurrence analysis of the R&D literature. *Evaluation Review, 18*(1), 98–115. https://doi.org/10.1177/0193841X9401800110
Tirole, J. (2024). Competition and industrial policy in the 21st century. *Oxford Open Economics, 3*(Supplement_1), i983–i1001.

Valletti, T. (2024). Debate: 'doubt is their product'—The difference between research and academic lobbying. *Public Money & Management*, 1–3. https://doi.org/10.1080/09540962.2024.2404249
van der Aalst, W., Hinz, O., & Weinhardt, C. (2019). Big digital platforms: Growth, impact, and challenges. *Business and Information Systems Engineering, 61*(6), 645–648. https://doi.org/10.1007/s12599-019-00618-y
Varian, H. R. (2021). Seven deadly sins of tech? *Information Economics and Policy, 54*, Article 100893. https://doi.org/10.1016/j.infoecopol.2020.100893
Verhoef, P. C., Broekhuizen, T., Bart, Y., Bhattacharya, A., Qi Dong, J., Fabian, N., & Haenlein, M. (2021). Digital transformation: A multidisciplinary reflection and research agenda. *Journal of Business Research, 122*(November 2019), 889–901. https://doi.org/10.1016/j.jbusres.2019.09.022
Viera Magalhães, J., & Couldry, N. (2021). Giving by taking away: Big tech, data colonialism and the reconfiguration of social good. *International Journal of Communication, 15*, 343–362.
von Tunzelmann, N. (2003). Historical coevolution of governance and technology in the industrial revolutions. *Structural Change and Economic Dynamics, 14*(4), 365–384. https://doi.org/10.1016/S0954-349X(03)00029-8
Vu, K., Hanafizadeh, P., & Bohlin, E. (2020). ICT as a driver of economic growth: A survey of the literature and directions for future research. *Telecommunications Policy, 44*(2), Article 101922. https://doi.org/10.1016/j.telpol.2020.101922
Waltman, L., & Van Eck, N. J. (2012). A new methodology for constructing a publication-level classification system of science. *Journal of the American Society for Information Science and Technology, 63*(12), 2378–2392.
Wang, Q., & Waltman, L. (2016). Large-scale analysis of the accuracy of the journal classification systems of web of science and scopus. *Journal of Informetrics, 10*(2), 347–364. https://doi.org/10.1016/j.joi.2016.02.003
Webster, P. (2023). Big Tech companies invest billions in health research. *Nature Medicine, 29*(5), 1034–1037. https://doi.org/10.1038/s41591-023-02290-y
Wu, T. (2017). The attention merchants: The epic scramble to get inside our heads. *Vintage*.
Xue, C., Tian, W., & Zhao, X. (2020). The literature review of platform economy. *Scientific Programming*, 1–7. https://doi.org/10.1155/2020/8877128
Yang, R., Wu, Q., & Xie, Y. (2023). Are scientific articles involving corporations associated with higher citations and views? An analysis of the top journals in business research. *Scientometrics, 128*(10), 5659–5685. https://doi.org/10.1007/s11192-023-04808-0
Yegros-Yegros, A., Capponi, G., & Frenken, K. (2021). A spatial-institutional analysis of researchers with multiple affiliations. *PLoS One, 16*(6), Article e0253462. https://doi.org/10.1371/journal.pone.0253462
Zhang, Q. (2017). Research on Apple Inc's current developing conditions. *Open Journal of Business and Management, 6*(1), 39. https://doi.org/10.4236/ojbm.2018.61003
Zuboff, S. (2019). *The age of surveillance capitalism*. Profile Books.