City Political Economy Research Centre

Working Paper Series

Same End By Different Means: Google, Amazon, Microsoft and Meta's Strategies to Organize Their Frontier AI Innovation Systems

Cecilia Rikap

CITYPERC Working Paper No. 2023-03

City, University of London
Northampton Square
London EC1V 0HB
United Kingdom

Same end by different means: Google, Amazon, Microsoft and Meta's strategies to organize their frontier AI innovation systems

Cecilia Rikap (CITYPERC, City University of London)

Abstract

I combine quantitative methodologies and in-depth interviews to analyse United States Big Tech different strategies to organize and profit from their AI corporate innovation systems (CIS). I propose 1) "frenemies" for Microsoft, because even Chinese organizations and direct competitors integrate its CIS. 2) "University" for Google, whose AI strategy included leaving DeepMind autonomous to explore more fundamental AI but appropriation mechanisms are not translating into a clear business advantage. 3) "Secrecy" for Amazon, given its large concern with secrecy to profit from AI. 4) And "application-centred" for Facebook; its AI CIS is the narrowest, mostly attached to its platforms.

Highlights

- Analyses US Big Tech strategies to organize and profit from artificial intelligence
- Builds on the corporate innovation system concept to identify Big Tech AI strategies
- Combines network analysis, original data and in-depth interviews
- Strategies: Frenemies-Microsoft, University-Google, Secrecy-Amazon, Application-Facebook

Keywords: Big Tech; Artificial Intelligence; technological innovation systems; corporate innovation systems; Digital Capitalism.

1. Introduction

Soon after OpenAI released ChatGPT in November 2022, it became integrated into several Microsoft products. Microsoft was backing OpenAI since 2019 in exchange for privileged access to developments that were finally bearing fruit. Google hastily reacted by launching its own large language model "Bard", which made a factual mistake in its first demo, wiping USD 100bn off the market capitalization of its parent company Alphabet. By February 2023, Meta presented its own alternative, LLaMA (Large Language Model Meta AI)[1] and Amazon entered the race expanding its support to Hugging Face, a start-up whose artificial intelligence (AI) chatbot is offered as a service in Amazon Web Services (AWS) as well as

[1] https://gizmodo.com/facebook-chatgpt-google-ai-chatbot-google-bard-1850155514

Amazon Bedrock, a service for building and scaling generative AI applications also offered on AWS.[2]

The generic AI race illustrates these companies' technological -and market- convergence (Jacobides et al., 2021; Rikap & Lundvall, 2021). Beyond this shared evolution, scholars have largely scrutinized what these and other leading technology companies have in common coining terms like Big Tech and tech giants (see Section 2).

While the literature has focused on what Big Tech have in common, this paper inquiries about Alphabet (for simplicity for the reader hereon Google), Amazon, Microsoft and Meta (for simplicity hereon for the reader Facebook) different strategies to organize and profit the most from their respective AI corporate innovation system (CIS). CISs are innovation systems controlled by a dominant firm and constituted also by other organizations (firms, universities, public research organizations, etc.). The leading firm maximizes extracted rents from knowledge and innovation co-produced inside the CIS and defines its overall orientation (Lundvall & Rikap, 2022).

Hence, this paper explores the distinct ways in which the four corporations are seeking to develop, use and profit from frontier AI, notably frontier machine learning and big data (including data mining, data science and data analysis) technologies, which compose a technology cluster that is the only one likely to be a general-purpose technology; also conceived as an emerging general or general-purpose method of invention (Bianchini et al., 2022; Cockburn et al., 2018; Goldfarb et al., 2023; Rikap & Lundvall, 2021).

My main hypothesis is that although the four corporations share the same goal -i.e., to build the leading AI CIS, thus maximizing the opportunities to profit from AI-, each company has developed a different strategy to achieve it and exhibits different degrees of success. To identify and compare their AI CIS strategies, I combined quantitative research with in-depth semi-structured interviews.

CISs refer both to the co-production of knowledge and innovation and to how the leading corporation profits from the system's successes. Hence, I analysed Big Tech co-production of AI with other leading organizations. Moreover, I considered indicators of AI appropriation by analysing each giant's AI-related acquisitions, investments in AI start-ups, their place in the ranking of top AI patent assignees and the content of their AI patent portfolio. I also considered AI talent indicators since talent can be seen as a bridge between the co-production and appropriation of knowledge inside CISs. Among them, I looked at double-affiliations with universities, which provides evidence of sustained knowledge flows between each Big Tech and its CIS.

Results were validated with 15 semi-structured in-depth interviews that also enabled me to inquire about the AI strategies of each company, the differences among them and with other companies. I interviewed senior managers, scientists and engineers working for the chosen Big Tech (9 interviews representing 14 employee-company relations since 4 interviewees have worked for more than one Big Tech) and other leading corporations developing AI (6 additional interviews from Alibaba, Bosch AI, Globant, IBM Research and Mercado Libre).

---

[2] https://www.latimes.com/business/story/2023-02-21/amazons-aws-hugging-face-ai-deals-chatgpt; https://www.aboutamazon.com/news/aws/aws-amazon-bedrock-generative-ai-service?utm_source=amazonnewsletter&utm_medium=email&utm_campaign=041523&utm_term=generativeai

In a nutshell, my findings point to four different strategies to organize and profit from AI CISs. "Frenemies" can be used to define Microsoft's strategy. It has successfully integrated into its CIS the least expected actors, from Chinese Big Tech and other Chinese organizations to investing in AI start-ups that then sell services to competitors, with the paradigmatic case of OpenAI. Microsoft became the key bridge, thus the gatekeeper (Burt, 1995), that connects AI research in Western core countries and China. On the contrary, "university" defines Google's strategy because it acts almost like a leading university. It has the largest presence in AI conferences, both presenting papers and at their committees, has more employees with double-affiliations and, unlike the other Big Tech, still gives high priority to AI patenting and acquisitions. However, and unlike Microsoft and Amazon, it seems that these appropriation mechanisms are not so clearly translating into an AI business advantage. At the other end in terms of openness, compared to Google and Microsoft, Amazon has developed a frontier AI CIS, the most diverse in terms of functional applications, privileging secrecy. "Secrecy" defines its strategy, which is at the service of Amazon's goal to produce and apply AI only when it provides a straighforward benefit for customers, which translates into more profits and data, and for that the company prefers to limit disclosure. Unlike Amazon, by having more open CISs, Google and Microsoft are better positioned to influence the whole AI field and not only those directly connected to them. Finally, while the three companies have a strong foothold in frontier generic AI, Facebook's AI CIS is narrower and remains attached to its ongoing businesses, which is why it can be defined as "application-centred".

The rest of this paper is organized as follows. Section 2 elaborates on Big Tech literature, particularly regarding AI. The methodology is introduced in section 3, which is followed by a presentation of the results in section 4. Section 5 summarizes the findings and proposes a different strategy for each Big Tech. Identifying these differences, as the paper concludes in section 6, is not only relevant for academia but also for regulation and agency.

2. The Big Tech literature

2.1. Constructing the Big Tech concept

As large companies from the digital sector expanded their market capitalization, grew disproportionately and usually profited along those lines, the literature studying them blossomed. Terms like "Big Tech" and "tech giants" became vox populi. In a book that has the former in its title, Foroohar (2019, p. 5) provides a preliminary definition claiming that "everything in Big Tech goes big or it doesn't go at all—and the bigger it gets, the more likely it is to go bigger still." Others prefer to speak of "tech giants", widely used both by academics (Whittaker, 2019) and media (Hill, 2020), to emphasize the -generally pervasive- commonalities of large US (sometimes also including Chinese) large tech companies and how omnipresent they have become in everyone's lives. Acronyms have also proliferated and evolved: from FAANG (Facebook, Amazon, Apple, Netflix and Google) to recognizing that it was not Netflix but Microsoft that should be part of this club, thus GAFAM. And after Facebook became Meta and changing Google for its parent company Alphabet, MAMAA was introduced.[3]

However, “defining how big a digital tech company must be to be part of the ‘Big Tech’ club, and how exactly size should be measured” is a difficult and open discussion (Viera Magalhães & Couldry, 2021, p. 347). When focusing on US companies, the label usually includes a stable set of corporations including Google, Amazon, Facebook, Apple and Microsoft. Yet, depending on the topic, some may be left aside or the remaining ones may be complemented by other large companies. In the case of Viera Magalhães and Couldry (2021), who inscribed Big Tech “social good” initiatives within data colonialism, they dismissed Apple because it lacked such experiences and included IBM.

Often, scholars use both labels indistinctly (Li & Qi, 2022; Rikap & Lundvall, 2021; Van Dijck et al., 2018) or combine them. Li and Qi (2022) distinguish “Big Tech giants” from other platform companies in China by their differential profits, high for the former and low or even negative for the latter because the former control interconnected critical resources (user devices, operational systems, data and tools to gather them, data centers and AI technologies and payments systems and logistics services that connect the digital world with offline services). Among these resources, Rikap and Lundvall (2021) focus on data crunched with AI as a new method of invention that, according to their empirical evidence, risks being monopolized by Big Tech, while other authors focused on Big Tech infrastructural power (Blanke & Pybus, 2020; Hendrikse et al., 2022; Van Dijck et al., 2018). Among them, Van Dijck et al. (2018) argue that Big Tech power is based on their core infrastructure platforms used by all the other platforms to operate. This enables Big Tech to collect and combine diverse data flows and, according to Hendrikse et al. (2022), is resulting in a “Big Techification” of every aspect of social life.

Scholars have also mobilized the label “Big Tech” to refer to a shared feature among these companies. They referred to a Big Tech lending and financial intermediation model (Frost et al., 2019; Liu et al., 2022), a financialized “Big Tech model” (Klinge et al., 2022), a Big Tech acquisition strategy that negatively affects their competitors and venture capital (Affeldt & Kesler, 2021; Bourreau & de Streel, 2020; Glick & Ruetschlin, 2019; Kamepalli et al., 2020), Big Tech companies as data-driven intellectual monopolies (Rikap & Lundvall, 2020, 2021), and as a techno-economic configuration (Birch & Bronson, 2022). Topics of interest also included Big Tech lobby networks (Tarrant & Cowen, 2022) and their new marketization of philanthropy (Manning et al., 2020). Taxing them is also a topical theme within and beyond academia (Hendrikse et al., 2022; Mansell, 2021).

Other labels were also proposed to denote specific common features. Bratton (2016, Chapter 3) distinguishes “stack platforms” from mere platforms. The former operates with multiple interoperable layers -a stack-, thus not only the infrastructural ones. A piece published at the Harvard Business Review speaks of these companies as digital superpowers or hub firms (Iansiti & Lakhani, 2017) and the European Commission (2022) has suggested the term market gatekeepers. In relation to the latter, there are multiple contributions on Big Tech potential market power abuses and how to regulate them (Benghozi et al., 2020; Graef & Costa-Cabral, 2020; Hutchinson, 2022; Jacobides, 2020; Mazzucato et al., 2023).

[3] https://fortune.com/2021/10/29/faang-mamaa-jim-cramer-tech-facebook-meta/

Other authors have implicitly recognized that tech giants are not like other platforms when they focus on one of them, such as Kenney et al. (2021) observation that their study of Amazon is not representative of the typical platform company. Among Big Tech, Amazon has received particular attention, including comparisons with other tech giants or traditional retail (see for instance Baud & Durand, 2021; Coveri et al., 2022; Rikap, 2020; Wu & Gereffi, 2018).

Overall, "Big Tech" or "tech giants" monikers as well as other labels to bundle together large tech companies became so widely used that sometimes, like in Safadi and Watson's (2023) analysis of knowledge monopolies and their innovation divide, a definition of what they understand as Big Tech is missing, only providing non-exhaustive examples at different parts of their investigation. Pitelis (2022) uses one term to define the other. For him, Big Tech are platform-based technology giants. More positive stands also take the "Big Tech" identity as a given, such as Petit and Teece (2021) whose starting point is that these are diversified firms that compete with each other and with entrants and adjacent firms in oligopolistic markets.

Could we be missing something by bundling them together under different homogenizing labels? Categorizing is useful but never neutral, having effects on what is been classified or categorized (Bowker & Star, 2000), just like nowadays AI classification systems' performative effects raise concerns (Crawford, 2021). Precisely one area in which the construction of the Big Tech concept has received recent attention refers to their technological convergence around AI (Jacobides, 2020; Rikap & Lundvall, 2021).

2.2. Big tech and AI

Jacobides et al. (2021) studied the division of labour within AI, distinguishing between AI enablement (typically digital infrastructure), AI production and AI consumption. They found that Google, Amazon, Microsoft, Alibaba and Tencent are vertically integrated firms and classified them as AI Giants that produce AI for internal and external use building global AI ecosystems. The paper does not distinguish among these AI Giants but classifies Facebook as a different species, an AI-powered operator that produces internally part of the AI it uses but that also depends on AI Giants.

Furthermore, previous research found that since 2012 large technology firms are increasingly participating in major AI conferences favoured by a "compute divide" (Ahmed & Wahed, 2020). The latter was defined as uneven access to computing power which favours large technology firms and the elite universities collaborating with them. Similar results were found by Klinger et al. (2020) for AI research conducted by the private sector, in particular by tech giants that specialize in what the authors define as "data-hungry and computationally intensive deep learning methods". They also found that diversity of the AI research field has stagnated and pointed especially at the narrow thematical and methodological interests of the US most prestigious universities, Google, Microsoft, Facebook, Amazon and OpenAI. Among similar lines and judging by AI citations, Jurowetzki et al. (2021) found that Microsoft and Google are the most influential organizations in the AI field. Hence, if these companies narrow their AI research focus, it is highly likely that this will impact on the overall field.

Alternative methods that were marginalized include those that consider AI societal and ethical implications (Klinger et al., 2020). Nonetheless, it was also shown that Big Tech even shaped the AI ethics field since its infancy, among others, by funding almost every AI ethics researcher from leading universities (Abdalla & Abdalla, 2021; Ochigame, 2019).

A common feature between all these investigations and those of the previous section is that, even when focusing only on some Big Tech as Jacobides et al. (2021), the resulting set of large technology companies are studied as an overall homogeneous bundle. An exception is Heston and Zwetsloot (2020), who geolocalized Facebook, Google, IBM and Microsoft AI R&D (they excluded Amazon and Apple due to lack of data) and identified differences in the share of AI staff and AI labs across companies. However, they did not explore what these differences mean or why they take place. Moreover, the piece generally analyses findings for all the companies together, plunging in the concentration of AI labs in the San Francisco Bay Area and Seattle. Even those acknowledging differences, like Birch and Cochrane's (2022) assertions on Big Tech having heterogeneous techno-economic practices, did not explore claimed differences.

Overall, by constructing the Big Tech and other similar labels, the literature has certainly advanced our knowledge on digital capitalism, industrial organization and corporate power. However, this packaging risks losing sight of relevant heterogeneities. With the recent surge of deep neural networks within AI, in particular large language models powering generic AI, and their associated risks, achieving an in-depth understanding of each Big Tech AI strategy became all the more important. This is even more pressing considering several investigations pointing to the dangers of the concentration and even monopolization of AI cutting-edge research by Big Tech (see for instance Abdalla & Abdalla, 2021; Ahmed & Wahed, 2020; Jacobides et al., 2021; Rikap & Lundvall, 2021). Yet, to the best of my knowledge, there is still no comprehensive analysis of the different ways in which these companies are developing, shaping and capturing AI like the one conducted in the rest of this paper.

3. Methodology

3.1. Company selection

I have chosen to focus on Amazon, Facebook, Google and Microsoft. Previous research highlighted the place of these four companies, in particular of Google and Microsoft but also referring to Facebook and Amazon, both in AI conferences, AI papers and AI patents, outpacing their Chinese rivals (Ahmed & Wahed, 2020; Jurowetzki et al., 2021; Klinger et al., 2020; Rikap & Lundvall, 2021; World Intellectual Property Organization, 2019). I excluded IBM even though it is a major player both in terms of AI publishing and patenting because it is usually excluded from the Big Tech label, none of my interviewees mentioned it as an AI leader and even the IBM researcher that I interviewed confirmed that IBM is not attempting to compete with US Big Tech, recognizing that it is no longer a computing leader.

3.2. Data sources and methodology

To identify and compare chosen Big Tech companies' strategies to organize and profit from their AI CIS, I combined quantitative research with in-depth semi-structured interviews that were used to validate and delve into the quantitative results. Interviewees also pointed to additional indicators that were integrated into the analysis during the course of the investigation.

At the quantitative level, I used a set of indicators, methodologies and data sources to map the role of each company in the co-production of cutting-edge AI and identify differences in their chances to profit from it. Both dimensions deeply rely on AI talent, which can be considered a bridge between knowledge co-production and appropriation, thus I also considered related indicators. Table 1 provides a comprehensive summary.

Table 1. Summary of the quantitative methodological strategy

Dimensions of analysis		Proxy	Data source	Period of analysis
Co-production of AI	Positioning in the AI research field	Network analysis + Betweeness centrality	Top 14 AI Conferences bibliometric data extracted from Scopus	2012-2020
		Participation in conference committees	Conference websites	2023 (except for AAAI Conference that only had data for 2022)
	Content of AI research	Text mining and network analysis	Top 14 AI Conferences	2012-2020
Profiting from AI	AI-firms' acquisitions	Number and industries of AI acquisitions	Crunchbase	2012-2022
	Funding AI start-ups	Number of AI start-up firms' in which a Big Tech appears among the start-up's top 5 investors	Crunchbase	2021 (except for Facebook, data for 2023)
	AI granted patents	Ranking of top 30 AI granted patents assignees in 2022, comparison with WIPO's (2019) report for a previous period	Derwent Innovation	2022
	Content of AI patents	Text mining of the 30 most frequent multi-terms in abstracts and titles of each Big Tech AI patents	Derwent Innovation	2022
Bridge between production and profit	AI talent	Academic institutions with scholars that also work for a Big Tech (double affiliations)	Top 14 AI Conferences bibliometric data extracted from Scopus	2012-2020
		Open job posts in AI (absoute terms and in relation to total job posts)	Company career websites	April 2023

I proxied the frontier AI research network with a bibliometric sample of all the presentations at the top 14 AI conferences between 2012 and 2020 extracted from Scopus. Previous research has shown that the most influential AI research is presented at top AI conferences (Ahmed & Wahed, 2020). I chose AI conferences following previous research that used the lists provided by the Computer Science Rankings (www.csranking.org) (Ahmed & Wahed, 2020). I validated the extracted list with an AI computing scientist who suggested to include two smaller AI conferences (the "European Conference on Artificial Intelligence" and "Uncertainty in AI"). The final list of the most prestigious AI conferences used for this investigation is presented in Table 2. My resulting dataset contained 71,264 presentations.

2012 is an inflection point in AI, in particular for machine learning, including the achievement of a computer vision breakthrough after the introduction of the AlexNet convolutional neural network architecture (Ahmed & Wahed, 2020; Jurowetzki et al., 2021). 2020 is the end date because building this network was the first step of the investigation and at the time of retrieval, late 2021, it was the last year with complete information. Since I wanted to identify the evolution of this network, I split the sample into three sub-periods (2012-2014, 2015-2017 and 2018-2020).

Table 2. List of leading AI Conferences

Acronym	Conference Name
AAAI	Association for the Advancement of Artificial Intelligence
IJCAI	International Joint Conference on Artificial Intelligence
CVPR	Conference on Computer Vision and Pattern Recognition
ECCV	European Conference on Computer Vision
ICCV	International Conference on Computer Vision
ICML	International Conference on Machine Learning
KDD	Conference on Knowledge Discovery and Data Mining
NeurIPS	Conference on Neural Information Processing Systems
ACL	Association for Computational Linguistics
EMNLP	Empirical Methods in Natural Language Processing
NAACL	North American Chapter of the Association for Computational Linguistics
SIGIR	Annual International ACM SIGIR Conference on Research and Development in Information Retrieval
ECAI	European Conference on Artificial Intelligence
UAI	Uncertainty in AI

To proxy the AI frontier research network of organizations for each sub-period, I used network analyses combined with clustering, a technique that groups the closest entities forming communities within networks (Fortunato & Hric, 2016), to map the network of most frequent co-authoring organizations. Previous studies used network analysis for mapping relations between actors within a knowledge or innovation system (Cooke, 2006; Testoni et al., 2021; Trujillo & Long, 2018; Wasserman & Faust, 1994).

The data were processed using the CorText platform (Tancoigne et al., 2014), which allowed us to build network maps by using specific algorithms that associate entities according to their frequency of co-occurrence within a corpus (Barbier et al., 2012). To build the networks in this paper, the Louvain community detection algorithm was applied as cluster detection method (Blondel et al., 2008). To focus on the most influential actors for each sub-period, instead of mapping the whole network of co-authorships, I prioritized the 150 entities with the highest co-occurrence frequency. In the first period, only 147 organizations were above the minimum threshold to be included in the network. I used the chi-square proximity measure to determine nodes and edges to be considered in each network map. This is a direct local measure, meaning that it considers actual occurrences (co-authorships) between entities. To define the direct ties (edges), chi-square normalization prioritises links towards higher degree nodes; these are the most frequent co-occurrences (co-authorships) within the network. It thus privileges the strongest links for each organization, so edges can be interpreted as an indicator of closeness between organizations. I also calculated the betweenness centrality of each resulting node using Gephi to evaluate the position of the selected Big Tech. This is a standard measure for considering the intermediating role of each node in a network, defined as the sum of the ratio of the shortest paths between any two nodes in the network that pass through that node.

To retrieve affiliations, Scopus offers a field with authors' addresses that includes the name of their research institution. I used this field to proxy the overall AI frontier knowledge network of organizations. From a total of 59,907 addresses listed in the original sample, an in-depth cleaning process was conducted to identify affiliations resulting in a final list of 13,637 organizations. For this cleaning, patterns were identified in the portion of the address that referred to the name of the organization and, through an iterative process of modifying the original records using regular expressions, the list was purged of ambiguities. I followed Rikap (2019) and merged all the affiliations at the level of the corresponding parent organization.

The same procedure was used to build a network of organizations and privileged topics for the whole 9-year period. To identify the privileged topics within my sample, I text mined the 500 most frequent multi-terms appearing in titles, abstracts and keyword lists. Frequency was computed at the level of the document (a multi-term was counted only once per article even when it appeared more times in the same document) and the output list was cleaned to exclude words whose high frequency is explained by either their grammatical function (such as "and" and "or") or the level of grammaticalization within the scientific genre ("previous research", "proposed method", "results show", etc.). The final list consisted of 416 multi-terms. I built a network map that links the top 150 organizations and multi-terms to get a sense of the topics privileged by each Big Tech and other key organizations.

Then, I retrieved from each conference website the full list of committee members. These are the people that decide which papers will be accepted, thus, considering the importance of selected conferences, they occupy a crucial role in setting the AI frontier agenda. Since this was suggested by one of the interviewees and previous years' data was not always available, I retrieved information for 2023, except for AAAI for which data was available for 2022.

Additionally, I used Crunchbase to retrieve Big Tech acquisitions between 2012 and 2022 and the number of companies where they appeared among the top five investors by the end of 2021 to avoid the effects of the current global macroeconomic and tech sector distress. This information was not available for Meta, thus its information corresponds to 2023. Crunchbase provides the technological fields or industries (not distinguished in Crunchbase) where each firm operates, following a classification made by Crunchbase and firms themselves. I listed the frequency of appearance of all the technological fields or industries in each of the four companies' acquired firms.

I also retrieved all the AI granted patents in 2022 from Derwent Innovation. I applied the same methodology used by WIPO (2019) to identify AI patents to compare my results with those of that report. I also used text mining to extract the top 30 multi-terms appearing in each Big Tech AI granted patents for 2022.

Throughout the investigation and as identified in previous research (Gofman & Jin, 2022), I noticed that several scholars worked part-time at Big Tech companies. Therefore, I included an indicator of double-affiliations at the institution level by retrieving from my AI top conferences dataset all the academic institutions with scholars that also declared a Big Tech as their affiliation for the same article presentation.

Since interviewees also pointed out that internal talent was a source of differentiation among Big Tech, I also retrieved job posts information from each Big Tech career's website. Previous research has already used hiring data for labour market studies. Abis and Veldkamp (2020) used it to proxy data stocks by retrieving information on demand for data managers and data analysts. Future research could also map each company's AI workforce extracted from LinkedIn, which was not conducted here due to space limitations.

Finally, I conducted interviews to verify and expand on the quantitative results. I conducted semi-structured in-depth interviews with senior managers, researchers and engineers working for the chosen and other leading corporations with knowledge of Big Tech AI business. I interviewed nine employees from the four chosen giants working in the US, Germany and the United Kingdom (UK). Interestingly, four of them had worked for at least another Big Tech company before, and were asked the same questions for all the Big Tech they had worked at, providing in total 14 company-employee answers. I also interviewed six researchers and engineers from Alibaba, Bosch AI, Globant, IBM Reseearch and Mercado Libre. On top of asking concrete questions related to observed results -such as questions on Microsoft's AI research in China or the reduced importance of patents for some of the analysed Big Tech-, interviews inquired about Big Tech AI strategies more in general, the differences among them and with other companies. Interviews last between 30 minutes and one hour and were conducted between August-2022 and May-2023. All the interviewees required to remain anonymous.

These interviews cannot be considered as representative because of their reduced number and because most of them were secured by indirect connections with employees. Other indirect connections refused to be interviewed declaring that they did not have the authorisation or leverage to ask for permission. I also emailed all the Big Tech employees listed as members of AI Conferences committees and contacted people that was recommended by my interviewees without providing the names of those who had suggested me to contact them. I only received three responses and, in the end, only one agreed to be interviewed. Nevertheless, the consistency of the replies and their correspondence with my quantitative analysis justify their inclusion in this investigation. Another limitation of my qualitative investigation is that, for the same reasons, I was only able to interview one Chinese Big Tech employee.

4. Results

4.1. The co-production of AI

In this section, I focus on the differences among chosen Big Tech in relation to the co-production of frontier AI with other organizations, their current participation in top AI conferences committees and briefly discuss the main content of their AI conferences' presentations.

4.1.1. Big Tech positioning in the AI research field

Between 2012 and 2020, the position of the four companies in the AI top conferences network becomes more central (Figures 1, 2 and 3 in Appendix). Yet, underlying what seems to be a shared evolution, there are meaningful differences in terms of the place and type of privileged collaborations among Big Tech, especially when looking at the most recent period (Figure 3).

Microsoft, Google and Facebook were already plotted in the top AI conferences' network between 2012 and 2014 (Figure 1). Microsoft had the highest number of AI conference presentations and Google ranked fifth. However, the latter occupied a marginal position in the network, only directly connected to one institution ranking 116 in betweenness centrality (the lists of betweenness centrality are available in the online appendix). Microsoft was more connected but its direct links were mostly with four organizations from its same cluster and one from another one and it ranked 39th in betweennesses centrality. Facebook occupied a marginal position, ranked 65 in number of presented papers and last in betweennesses centrality.

Their relatively non-central position changes completely in the last period. Google and Microsoft, in that order, become the two organizations with the highest betweenness centrality of the network. They are also second and third in number of presentations. The Chinese Academy of Science, which has the highest frequency of presentations in this period, is 12ve in betweenness centrality, pointing to the detachment of China from the rest of the world. Precisely concerning the latter, Microsoft occupies a crucial bridging position connecting the West with China. Microsoft is part of a cluster mostly integrated by Chinese organizations (firms and universities) and is directly connected to four additional clusters. In total, Microsoft is directly linked to eleven universities from China, the US, Switzerland and the UK.

Microsoft opened a Chinese office in 1992 and its first R&D facility six years later. In 2010, it inaugurated its first major R&D campus outside the US, a high-tech industrial park, in Shanghai. Two interviewees that work or worked for Microsoft confirmed that, even though being in China is complicated because Microsoft will always be seen as a US corporation, the company is at the forefront of developing research and business there. According to one of the interviewees, Microsoft succeeded in China, among others, because it compromised the necessary level of investment, such as collaborations with Xiaomi on mix-reality.

By being both deeply related to several US and European universities and widely established in China, Microsoft plays the role of unifying the network, connecting what would otherwise be what Burt’s (1995) defines as a structural hole in a network. In other words, the globalization of AI cutting-edge research and the overall structure of this network relies crucially on Microsoft.

In turn, Google’s position also contributes to structuring the network even though it is not geopolitically as relevant as Microsoft. Google has the highest betweenness centrality and is directly linked to 19 other organizations from four clusters, including IBM, universities and public research organizations, all coming from core countries. Ten organizations are from the US, three from Canada and two from Israel. The other four are European organizations. Interestingly, an interviewee indicated that employees could go to two conferences per year even without presenting research. Presenting at conferences entails a quite simple internal approval process where researchers’ more relaxed view prevails over managers' attempts to establish more controls.

> “They barely look at the abstract, the strategy is the loosest part. (…) Managers push us to give more detailed feedback but we already review a lot for conferences, real reviews for all the conferences all year around. It is too much to do it for the internal paper. As long as there is no code that is exchanged, there is nothing in terms of secrecy or to protect.” (Google interviewee 1)

Facebook’s evolution in the network is also impressive. In the last period, it jumped to the 8th position in betweenness centrality and ranks eleventh in frequency of appearance. However, unlike Microsoft and Google, it seems more focused in terms of its place in the network, something that will become more apparent in the rest of this section.

Amazon joined the network of leading organizations in AI conferences in the second period. Although it progressively shows a more central position (its betweenness centrality moved from the 105th to the 48th position between the second and third periods), it remains far from the other Big Tech. In the last period, it was directly linked to 6 organizations from three clusters, all of them from the US except for the Max Planck. Nonetheless, interviewees stressed that the delay in developing a significant presence and the relatively non-central place in comparison to the other tech companies is not a sign of weakness or technological laggardness, but a top-down decision.

> “The founder of Amazon never really wanted publications to be a big thing because science is only useful for him if it is for customer benefits. It was done to be a more attractive employer and to validate what we do, (…). Amazon doesn’t publish to gain top leadership, well yes, but quality is more important (…). The number of good publications is the wrong metrics for selling products. A good metrics for Amazon would be how much of the customer retention and engagement is affected by science.” (Amazon interview 2).

Publications, the interviewee continued explaining, are not the best output because they are not written in a jargon that is easy to understand for engineers. What is needed is to have the inventor or researcher explaining and replicating the research for the engineer. The other interviewed Amazon employee also pointed out that Amazon is behind in terms of the culture of working towards external publication because its principles, which are constantly emphasized and include “learn and be curious”, do not include sharing information publicly.

Summing up, the four Big Tech seem to have different strategies when regarding the leading AI research network. Among them, Microsoft and Google occupy the most central positions, with the former occupying a bridging or gatekeeping role between the West and China by positioning itself in what would otherwise be a structural hole. Thus, their capacity to influence the field beyond direct collaborations is the maximum in the network.

4.1.2. The content of Big Tech research presented at leading AI conferences

Besides their common focus on machine learning, more precisely on deep neural networks, differences among chosen Big Tech are further stressed by looking at the more specific content within the privileged topics of their presentations at leading AI conferences. Figure 4 in Appendix presents a network that connects the most frequent multi-terms with the organizations whose presentations included them with the highest frequency. Table 3 lists the multi-terms directly connected to each of the four chosen Big Tech in Figure 4.

Table 3. Topics directly linked to Big Tech in Figure 3.

Google	Amazon	Microsoft	Facebook
Neural Networks	natural language	reinforcement learning	language model
reinforcement learning	natural language processing systems	natural language	machine translation
machine learning	transfer learning	language model	action recognition
language model	knowledge graphs	machine translation	
learning algorithms	word embeddings	data mining	
generative model	time series	neural machine translation	
machine translation	text classification	large amounts	
transfer learning	context information speech recognition		
gradient methods			
sample complexity			
data augmentation			
monte carlo methods			

Source: Author's analysis based on Scopus

Google's research is the most diverse and includes both general AI multi-terms and more specific ones. Intelligence chatbots, like ChatGPT and Google's Bard, are based on a "generative model" and trained with "reinforcement learning", which are terms directly connected to Google. Reinforced learning is a deep neural network technique originally developed by DeepMind that does not require a pre-set of labelled classifications to train the model. It is used for models where the function to be optimized is not fully accessible and inputs depend on previous actions (Alom et al., 2018). One of the Google employees that I interviewed defined it as "agentic" because the intelligence agent -the computer program- interacts with the environment and learns to act within it. It is a powerful tool because the AI model improves the more it is used, thus to some extent outsourcing the improvement of the model -hence part of the R&D- to the users or customers.

Amazon follows in terms of direct connections to multi-terms which are somehow skewed towards AI for language applications. Moreover, it includes the term transfer learning, which is also among Google's list. This is a technique in which algorithms transfer what they have learned from one or several datasets to another situation for which there is no sufficiently large dataset to train the model. This approach has been used for improving classifications in object recognition and text categorization databases using Amazon data (Zhuang et al., 2020).

Second, the term "time series" might seem outdated when thinking of AI. However, according to two Amazon employees that I interviewed, it speaks of the way in which Amazon approaches new technologies and R&D more generally. Amazon uses time series for long-term forecasting of demand and other aggregated variables for each of the countries or regions where it operates. One of the interviewees pointed out that it is still to be confirmed whether deep learning for such forecasting would provide additional accuracy. The interviewee added that "moving too much to AI puts restrictions and constraints on the insights we provide" (Amazon interview 1), which is why they still mostly use traditional statistical methods for long-term predictions. Meanwhile, forecasting at the item level and price setting are done with frontier deep neural network models.

Overall, the prevalent content of Amazon's AI conferences presentations speaks of the company's approach to technology, as explained by another Amazon employee interviewed.

> "I think that the good thing about Amazon's approach to AI is that it is agnostic and application focus, it doesn't matter to keep using an old method, it doesn't become a selection criterion for a project how new the proposed method is. My impression is that that is the selection criteria in other companies, just a simple random forest[4] can be useful and other big companies will less likely fund it rather than a state-of-the-art algorithm. We are technology agnostic at Amazon. Other companies will go for the more expensive things. ChatGPT is an example." (Amazon interview 2).

Microsoft, like Google, includes among its most frequent content the multi-term "reinforced learning". The rest of its directly linked terms refer to common aspects of Big Tech research, in particular AI functional applications for language, like Amazon. So we may say that Google and Microsoft are more focused on general frontier AI while Amazon develops frontier AI together with other forecasting techniques. In comparison with the other Big Tech, Facebook exhibits one exclusive multi-term, action recognition, which is a specific computer vision task used for recognizing and classifying human actions in videos or images further reinforcing the impression of this company as more focused on AI applied to its platforms.

4.1.3. AI conferences' committees

AI top conferences exhibit a significant presence of industry representatives in their committees (22%), mostly driven by US and, to a lesser extent, Chinese Big Tech (57 committee members) (Table 4). Big Tech influence can also be inferred by looking at

[4] Random Forest is a classification algorithm consisting of many decisions trees.

sponsorships. For instance, Amazon was a "platinum sponsor" of the AAAI Conference and Microsoft, Google and Baidu were silver sponsors. This is one of the conferences with the lowest industry participation, yet the AAAI 2023 organization team has only four members disclosed so far and 1 is at Microsoft. Hence, one may conclude that Big Tech have a strong foothold in defining what papers will be accepted and which will win prizes, which is a sign of their power to shape the AI field, as identified by an interviewee:

> “Most of the people leading the conference boards are in Big Tech, not all but at DeepMind we have a lot of those people. (…). They will say that they are independent and do it for the research but, are they? (…) Are they trying to steer the research and who gets the best paper? (...) I don’t know if it is significantly skewed, but do the members of the industry leave when they need to decide on papers from these companies? Someone told me that he tried to raise the alarm of conflict of interest (…) but they still stayed in the decisions.” (Google interviewee 1)

A concrete example of the entanglements between the organization of AI conferences and Big Tech interests are Women in Machine Learning sessions. This same interviewee explained that these spaces became a hiring forum, a place for Big Tech advertisement where PhD and postdocs that organize the sessions end up doing what the interviewee defined as “free labour for Big Tech hiring and legitimate a bad salary strategy because women are paid 30% less” (Google interviewee 1).

Table 4. Composition of leading AI Conferences committees*

Name of AI conference	Number of members in committee	From industry	Big Tech (US and Chinese)	Amazon	Google	Microsoft	Facebook	Share of industry participation
Association for the Advancement of Artificial Intelligence (AAAI)	45	3	1	0	1	0	0	7%
The International Joint Conference on Artificial Intelligence (IJCAI)	12	0	0	0	0	0	0	0%
Conference on Neural Information Processing Systems (NeurIPS)	39	20	13	1	9	1	2	33%
International Conference on Machine Learning (ICML)	25	6	2	0	1	0	0	24%
Conference on Knowledge Discovery and Data Mining (KDD)	58	15	8	1	3	2	0	26%
Association for Computational Linguistics (ACL)	42	10	7	2	2	1	1	24%
Empirical Methods in Natural Language Processing (EMNLP)	44	9	7	0	2	2	1	20%
Conference on Computer Vision and Pattern Recognition (CVPR)	38	8	3	1	2	0	0	21%
European Conference on Computer Vision (ECCV)	34	11	6	3	0	0	3	32%
International Conference on Computer Vision (ICCV)	34	11	5	1	1	0	3	32%
North American Chapter of the Association for Computational Linguistics	9	2	2	0	1	0	0	22%
International ACM SIGIR Conference on Research and Development in Information Retrieval	41	5	3	2	0	1	0	12%
Uncertainty in AI	25	1 (2)	(1)	0	0	0	0	4%
European Conference on Artificial Intelligence	20	3	0	0	0	0	0	5%
Totals	466	103	57	11	22	7	10	22%
Number of conferences with private presence				7	9	5	5	

Source: Author's analysis from AI Conferences websites. * All the committees are the ongoing ones except for AAAI where only the 2022 full committee was available. Uncertainty AI had as Sponsorship Chair a researcher from MILA that had worked for Microsoft, Google and Facebook.

Besides this common interest in the organization of top AI conferences, Google exhibits the largest presence, with a total of 22 members distributed in 9 of the 14 committees. Interestingly, it has nine of the 39 members of NeurIPS's committee, the main machine learning annual conference. In 2022, Google had the largest number of accepted papers in this conference (affiliations appear as Google, Google Research, Google Brain and DeepMind).[5] The other 3 Big Tech are also represented in this conference's committee, but only with one or two members. According to Abdalla and Abdalla (2021), at least two Big Tech have been sponsoring NeurIPS since 2015.

With half the number of members of Google, Amazon follows in terms of committees in which it has at least one representative (6) and total number of committee members. Conferences with Amazon employees in their committees range from overarching or general AI conferences to more specific events, including the Conference on Computer Vision and Pattern Recognition (CVPR), where only Amazon and Google integrate the committee and the

[5] https://github.com/sanagno/neurips_2022_statistics

Association for Computational Linguistics (ACL) whose committee is chaired by Dr. Yang Liu, affiliated to the University of Texas and Amazon. The relatively low representation, unlike that of Google in NeurIPS, speaks more of a strategy to establish a presence and get access to the internal discussions without necessarily having the sufficient leverage to steer -at least not without an alliance with others- the direction of the presented research.

Facebook's presence in these committees shows a more focused strategy, participating in only 5 conference committees, including those on computer vision, in line with the multi-term "action recognition that is prevalent in its AI papers (Table 3). Facebook has 3 out of 34 members in the European Conference on Computer Vision (ECCV) and the International Conference on Computer Vision (ICCV), respectively. In these committees, 11 members come from industry, so that Facebook has a 27% of the industry representation.

Like Facebook, Microsoft participates in only 5 committees and with a total of 7 members. Unlike the other companies, it does not participate in computer vision conferences' committees. Instead, on top of having a representative in NeurIPS and in another more general AI conference, it participates in committees of conferences on AI applied to language, which is certainly more aligned with its OpenAI partnership, to which I refer later (see section 4.2) and with the most frequent multi-terms of its AI conference presentations (Table 3).

4.2. Profiting from AI

Big Tech companies capacity to profit from AI is explored here by analysing their AI-related acquisitions and top investments, AI patents and where they stand in relation to secrecy as an appropriation mechanism.

4.2.1. Acquisitions and Top investors

Acquisitions and investments in other companies provide privilege access to technologies and skilled workforce. According to WIPO (2019), Google ranked first in AI-related acquisitions between 2009 and May 2018 with 18 acquisitions. Microsoft was third with nine, Amazon was fifth with 6 and Facebook eight with 5. Although these data evidence that the four have engaged in acquiring AI start-ups, their strategy in relation to what to privilege between acquisitions or investments and the industries from which they have acquired companies differ (Table 5).

Table 5. Big Tech AI acquisitions and investments in AI start-ups

	Microsoft	Amazon	Google	Meta
Industries appearing in more than one acquisition	Machine Learning	Machine Learning	Machine Learning	Machine Learning
	Software	Developer APIs	Analytics	Software
	Mobile	Apps	Software	Computer
	Developper Tools		Computer Vision	Mobile
	Natural Language Processing		Image Recognition	Computer Vision
	Information Technology		Natural Language Processing	Image Recognition
	iOS		Big Data	Developer APIs
	Developer Platform		Internet	Photography
Total number of industries	21	21	35	29
Total AI acquisitions since 2012	10	5	17	11
Cloud related acquisitions	1	1	0	2
Number of AI start-ups for which top 5 investors in 2021 (Meta info for 2023)	80	19	35	0

Source: author's analysis from Crunchbase

Google still leads in the number of AI-related acquisitions, which are also the most diversified both in the number of represented industries and in relation to the AI functional applications[6] found among them, including machine learning applied to images, language and analytics (Table 5). Google is also among the top 5 investors of many AI start-ups. However, it is widely outpaced by Microsoft. Microsoft acquires less but from sectors where it does not have a strong business (Mobile and iOS) and companies focused on strengthening its role inside its CIS as the owner of the tools and platforms that developers use to program specific solutions. This is reflected by the acquisition of companies working on Developer Tools and Developer Platforms.

The stories of DeepMind and OpenAI give testament of this different strategy. Google acquired the former, an AI forerunner UK start-up, in 2014. Until recently the company remained mostly independent but with the ultimate goal to generate valuable AI for Google. This started changing, as one of the interviewees explained, when DeepMind moved away from developing AI that played games because nobody cared about those applications, something that DeepMind top managers did not anticipate. The new strategy, the interviewee continued explaining, was to do things that people care about such as Alphafold, which developed an AI model that predicts protein structures. Originally, this project did not receive much attention internally until it became clear that AI for the sake of AI was not working for Google's overall strategy.

In the meantime, instead of acquiring or internally developing similar technologies, Microsoft reacted by investing USD 1 billion in 2019 in OpenAI.[7] This investment granted Microsoft an exclusive license to GPT-3, which underpins the first version of ChatGPT and was, by 2019, the most advanced language model with 175 billion parameters (Benaich & Hogarth, 2020).

To train its AI models, OpenAI needed supercomputers of a previously never seen scale, and Microsoft provided them in its cloud. At a Microsoft blog, Phil Waymouth, Microsoft senior

[6] Examples of AI functional applications based on deep learning and neural network approaches are speech processing, recognition and synthesis, natural language processing and images and video segmentation (World Intellectual Property Organization, 2019).

[7] https://thenextweb.com/artificial-intelligence/2019/07/23/openai-microsoft-azure-ai/

director of strategic partnerships concluded: "That shift from large-scale research happening in labs to the industrialization of AI allowed us to get the results we're starting to see today".[8] Put another way, Microsoft pushed OpenAI to move from research to applications, a change that led some OpenAI researchers to leave the company.[9] ChatGPT is an apparent result of this shift. Since then, Microsoft committed an additional USD 10 billion investment in OpenAI.[10] For Microsoft, investing instead of acquiring was a strategic move to assure that OpenAI applications are purchased even by rivals.

> "We know we have 49% of this company and the agreement has certain stipulations, privilege access to developments. OpenAI, for example, also works with Salesforce, which is one of our biggest competitors, but that is not a problem because if Salesforce uses OpenAI we still win because we earn revenue there. (…) In AI we didn't have to hit rock bottom, we are at the forefront. With Ballmer it was more of a "PC versus Mac" battle and there we were the lame ones, even though we never stopped making money and that's the goal, otherwise a company doesn't work. But Satya saw it coming and said 'let's do partnership with Open AI' and that mindset about how we can grow, be better all the time, brought us here" (Microsoft interviewee 1).

According to this interviewee, friends from Google had said that ChatGPT turned it upside down. Microsoft took the lead at a time when Google management was simultaneously communicating a slow-down in hiring and pushing its employees to be more "entrepreneurial", as a leaked internal memo from its CEO Sundar Pichai stated. [11]

In comparison to Microsoft and Google, Amazon acquires and invests less in AI start-ups, but its acquired firms tend to operate in multiple industries (Table 5). It seems that Amazon uses its cloud leadership as a vantage point to attract promising start-ups. I already mentioned that after OpenAI released the first version of ChatGPT, Amazon augmented its support to Hugging Face, a start-up developing another AI chatbot that is offered as a service in AWS.

Finally, and in line with the previous section's findings, Facebook acquired firms working on image and visual AI applications, which are more related to its relatively narrower business in comparison to the other Big Tech. Also, a major restructuring of Facebook's AI research took place by mid-2022, decentralizing its AI team to create AI Innovation Centers associated with each of its business units, both for its social networks and Metaverse. Facebook AI Research (FAIR) team became integrated into the company's Reality Labs Research.[12] This move seems to be further targeting AI to applications for Facebook's existing businesses. This narrower focus is also illustrated by its investment strategy in AI start-ups. We could not have access to the number of firms having Facebook among its top 5 investors in 2021 but in 2023 it was not among the top investors of any firm.

---

8 https://news.microsoft.com/source/features/ai/how-microsofts-bet-on-azure-unlocked-an-ai-revolution/?ocid=eml_pg394041_gdc_comm_mw&mkt_tok=MTU3LUdRRS0zODIAAAGKwmbrwlHO5mYvwKCSRwk2rcEO-79_q_J-nzO8jDiYkLCqxQDI3WXezvp1v-R1XS1chmfOLULFh7NnuL1mIejIT2WWNnZHWf1mc2zzg39WJ2aT7z8ppJQFXEi5

9 https://www.geekwire.com/2020/openai-renamed-closedai-reaction-microsofts-exclusive-license-openais-gpt-3/ and https://www.ft.com/content/8de92f3a-228e-4bb8-961f-96f2dce70ebb

10 https://blogs.microsoft.com/blog/2023/01/23/microsoftandopenaiextendpartnership/

11 https://www.theverge.com/2022/7/12/23206113/google-ceo-sundar-pichai-memo-hiring-slowdown-2022

12 https://ai.facebook.com/blog/building-with-ai-across-all-of-meta/

4.2.2. AI Patents and secrecy: complementary more than opposites

In 2019, WIPO (2019) published a report on Technological Trends in AI. It included the ranking of the top 30 patent applicants by number of patent families between 2013 and 2016, which was led by IBM (8,290) and Microsoft (5,930). The distance between the two forerunners narrowed when looking only at machine learning patents. Google ranked tenth and Amazon and Facebook were not listed.

Granted patents are a better indicator of possibilities to profit from AI than patent applications. So, using WIPO's (2019) definition of AI patents, I analysed AI granted patents for the top 30 patent grantees in 2022 (Table 6) and compared it with WIPO's (2019) findings. Although large companies use patents to create artificial barriers for rivals and usually do not profit from their whole portfolio, the indicator still provides valuable information for comparisons because these practices are shared among top patenting organizations in high-tech (Hall et al., 2013).

Table 6. Top 30 AI patent grantees in 2022

Organization	AI granted patents in 2022
Toyota	673
Samsung	538
Alphabet	452
Baidu	443
Honda Motor Co. Ltd.	367
IBM	295
Hyundai	280
Tencent	278
LG	254
Sensetime	211
Renault	206
Siemens AG	203
Sony Corporation	202
Ford	199
Bosch	175
Intel Corporation	172
University of Electronic Science and Technology of China	170
Huawei	168
HITACHI	164
General Motors LLC	160
Zhejiang University	158
Microsoft	149
NEC Corporation	147
Amazon	140
Mitsubishi	138
State Grid Corporation of China	137
Chinese Academy of Sciences	137
Canon Inc.	136
Tsinghua University	131
Fujifilm	122

Source: Author's analysis based on data extracted from Derwent Innovation

Compared to WIPO's (2019) findings, Microsoft seems not as focused on AI patents as before, not only judging by its place in the ranking (22nd) but also by the distance in the number of granted patents between those at the top and Microsoft. Besides generic multi-terms referring to machine learning, which are common for the four companies, the 30 most frequent multi-terms in Microsoft AI patents' titles and abstracts refer to virtual assistants and healthcare (Table A1 in appendix).

Another novelty regarding AI patents is that Amazon now integrates the top 30 ranking, with a similar number of granted patents in 2022 than Microsoft and with an AI portfolio that seems to be the most diverse of the four in AI functional applications, including image, audio, video and text. Here again, the multi-term "time series" pops up (Table A1 in appendix).

Facebook remains out of the ranking, occupying the 50th position. Its patents are connected to its existing platforms, with a focus on image and video and with multi-terms that can be easily connected to the Metaverse, such as “artificial reality environment”. Unlike the other Big Tech, terms referring to the cloud, natural language or AI for text are absent.

Among the four, Google seems to be the most focused on patenting AI, which jumped from the tenth to the third position in the ranking, placing itself as the first Big Tech company in AI granted patents in 2022. However, one of my interviewees argued that this was mostly a defensive strategy, to keep others from filling them and to prevent others from charging Google from using its own developments. A specificity of its AI patent portfolio are patents dealing with computer storage (possibly related to the cloud) and autonomous vehicles (Table A.1 in appendix).

All the interviewees agreed that scientific publications usually have an associated patent. To publish research that could generate profits, Big Tech first fill the patent before the paper can be submitted to review. At Google, inventors get a bonus (between USD 5,000 and 6,000) and the lawyers take care of everything, only asking the inventor to read the final draft, which according to one of my interviewees it is anyway quite time consuming.

Overall, it seems that even though patents still play a role in this field, they may not be the most relevant appropriation mechanism of frontier research, as was observed by the literature for ICT industry in general (Comino et al., 2019; Sampat, 2018). Interviewees agreed that secrecy and the speed of innovation are crucial for leading the AI field and that they are even complementary to publishing and patenting. In very simple terms, what makes a company be at the edge, will be kept secret, whereas complementary or not so cutting-edge developments are often published and/or patented for the reasons just mentioned.

> “Many ideas that are not commercialized are published and the partnerships with universities are very positive. Google has a lot of accelerators, it allows university researchers to use them and then we share ideas in common, but the core development of massive models is not going to be published, to keep the edge.” (Google interviewee 2)

An intangible asset that is always kept secret is internal data. A trade-off arises for researchers wanting to publish AI models that require large datasets to be trained. Using standardized or open datasets is easier if the aim is to publish. As the same interviewee explained, using instead internal data sources is more complicated due to compliance and privacy issues. To get access to an extract of an internal big data source, Google employees need to clearly state what they will use the data for and when they will start using it so that a copy of the required dataset is provided but only for a set period of time and then the copy is automatically deleted. Massive experiments, such as those underpinning Big Tech chatbots and other large language models, require massive scale data that are only available internally. These models are the frontier in AI, thus the requirement to use internal datasets kept secret is another reason why the AI field seems to be moving towards even more secrecy.

Another area where the importance of secrecy is apparent are Big Tech competitive research grants. Previous research highlighted that this was becoming a relevant funding source for academics and that Google was funding around 250 external research grants per year whereas Amazon, even though it has received 800 applications, only funded 50. A significant share of this funding is concentrated in a few AI leading institutions. Abdalla and Abdalla (2021) found that 52% (77 out of 149) of faculty based at Stanford, the MIT, the University of California Berkeley or the University of Toronto have received Big Tech funding. The figure grew to 84% after including in the calculation funding received by PhD students, internships and previous work experience. Big Tech employees avoid discussing new products with funded researchers and data are generally kept secret. One of the researchers interviewed by Popkin (2019, p. 666) said "They're giving you money to answer research questions, but it's still up to you to figure out how you're going to get access to data to answer the questions".

Secrecy is also an internal mechanism used to protect technological leadership. Interviewees doing AI research agreed that the edge in AI are small changes in configurations that most of the employees do not know about. Only the group that is programming those configurations will know. In general, things move so fast in these companies that non-compete agreements do not operate as in other industries. Sometimes, there is a period of paid leave between the time a scientist or engineer moves from one Big Tech to the other. This points to an underlying understanding that edge research moves so fast that it only takes months - between 3 and 6 according to interviewees- to move sufficiently ahead so that the former employee hired by another Big Tech will not be aware of the latest developments.

A final point to be made is that three Google interviewees observed a new trend further privileging secrecy. One particularly stated that

> "I see in the field a general push, led by OpenAI, to close down research, make it less open access, and offer the final product. I do not have internal information, but I suspect that Google and DeepMind are going to publish fewer academic papers in the future also as part of this process. (…) The reason why this did not happen before is because researchers want to publish, to advertise yourself, etc. But it is an incentive for the company to not disclose who is standing out, so the company has more control and keeps the talent for itself. In the publishing landscape you see the tension between workers and company and now the company is winning. (…) There is this marketing thing, you need good presence in the AI conferences but the point is that the focus of research is on massive multi-model and research on this side is going to be more sparse, kept more secret and I see that companies like mine will publish more reports with 80 pages and details of the model but written on the company terms and not complying with reviewers asking for more information to publish a paper." (Google interviewee 2)

As the previous quote underlies, the possibility of limiting this form of public disclosure raises employees' concerns, in particular for those that work part-time in academia, as I explore next.

4.3. AI talent: the bridge between co-production and appropriation

Here, talent is seen as a bridge between the co-production and appropriation of AI by Big Tech, thus connecting previous findings. As stated by one Amazon interviewee, having the most talented people and wanting them to stay is what matters the most in order to lead in the field. According to Big Tech statements and industry reports overviewed by Heston and Zwetsloot (2020), access to talent is also the main reason for setting up AI R&D laboratories outside the US. And, according to a BOSCH AI scientist, Big Tech companies are the AI forerunners precisely because they hire the most talented people.

> "They (Big Tech) have lots of money and the best way to lead is to hire AI researchers and ask them to investigate stuff. Google is doing it. They collect all the good researchers so that opponents in the market cannot. So, AI research is less accessible to competition and also the versatility of AI, it is the glue that fills in all the empty spaces nowadays." (BOSCH interview).

Highly skilled talent is hired full or part time and integrates other organizations from Big Tech CIS. Often, AI talent is drained from academia. By reconstructing the affiliation history of over 60,000 AI researchers, Jurowetzki et al. (2021) found that 8% had transitioned from academia to industry, with a sharp increase in transitions in the last decade, particularly of those working on machine learning and with higher citation rates. Gofman and Jin (2022) went deeper and found high and exponentially growing levels of brain drain of AI professors from US and Canadian universities into industry. The firms that hired the largest number of AI faculty were Google, Amazon and Microsoft. Facebook shared the 4th position with Uber and NVIDIA.

Part-time Big Tech employees are usually leading academics. In my sample of AI conference papers between 2012 and 2020, I could identify around 100 double affiliations between a Big Tech and a university or public research organization (Table 7).

Table 7. Institutions with AI scientists also working for Big Tech

Google	Microsoft	Facebook	Amazon
ASIT Japan	Aalto University	Georgia Tech	Caltech
Australian National University	Alan Turing Institute	Harvard	Carnegie Mellon University
Bar Ilan University	Carnegie Mellon University	ICREA	Heidelberg University
Brown University	China Sun Yat-Sen University	INRIA	Imperial College
Caltech	Chinese Academy of Sciences	Johns Hopkins University	Ohio State University
Carnegie Mellon University	ETH Zurich	McGill University	Rutgers University
CMU	Harbin Institute of Technology	New York University	University College London
Columbia University	Hebrew University	Sorbonne Universite	University of California
Cornell University	Hefei University of Technology Beijing	Stanford University	University of Edinburgh
ETH Zurich	Hong Kong Polytechnic University	Tel Aviv University	University of Southern California
Harvard	Indian Institute of Science	Texas A&M University	University of Texas
Hebrew University	MILA	Universite Le Mans	University of Washington
INRIA	MIT	University College London	University of Wisconsin-Madison
INSEE	Polytechnique Montreal	University of California	
Mila	Princeton University	University of Michigan	
Mines ParisTech	Shanghai Jiao Tong Unviersity	University of Texas	
MIT	South China University of Technology	University of Washington	
New York University	Stanford University		
Princeton University	Technion-Israel Institute of Technology		
Rutgers University	Tel Aviv University		
Stanford University	Tsinghua University		
Technion-Israel Institute of Technology	Universite de Montreal		
Tel Aviv University	University College London		
TTS Research	University of California		
University College London	University of Cambridge		
University of Alberta	University of Illinois		
University of California	University of Maryland		
University of Colorado	University of Massachusetts		
University of Edinburgh	University of Münster		
University of Michigan	University of Science and Technology of China		
University of Minnesota	University of Trento		
University of Oxford	University of Washington		
University of Texas	Weizmann Institute		
University of Warsaw			
University of Washington			
University of South California			

Source: Author's analysis based on the dataset of top 14 AI Conference presentations between 2012 and 2020.

As expected, given their higher presence in these academic events, Google and Microsoft have developed more of these collaborations both in terms of the number of academic institutions and countries involved. Microsoft's double affiliations are less concentrated in the US, mainly due to double affiliations with 8 Chinese institutions. The list also includes four Israelian, three British and three Canadian universities and a total of ten countries are represented in Microsoft's partnering organizations for double affiliations. Meanwhile, 20 of the 36 organizations with researchers also based at Google are US universities. Nonetheless, Google has researchers affiliated to other organizations in nine different countries, which is three times more countries than Amazon.

When inquired about the rationale for these double affiliations, one of Google's interviewees referred to Google's small office at the University of Alberta as a decision driven by the fact that this is one of the best places for reinforcement learning and added that "a guy form there is one of the fathers of the topic and he is at least part time in DeepMind." (Google interviewee 1).

Different interviewees mentioned that researchers with double affiliations typically push Big Tech companies to publish and present at AI conferences more. Moreover, Big Tech companies identify and capture talent by engaging in those conferences. One of the Google employees argued that conferences are used as a hiring forum “where a lot of lies are told about how it is inside, (…) they all want to have you for cheap.” (Google interviewee 1). Previous research has also showed that a precondition for receiving a competitive grant from Microsoft is to have had previous contacts with Microsoft employees such as networking at a conference (Popkin, 2019). Also, previous experience presenting at AI conferences was indicated as a necessary indicator of expertise at a Facebook job post for a postdoctoral position. A minimum requirement to apply for the job was: “Proven track record of achieving significant results as demonstrated by grants, fellowships, patents, or publications at leading workshops, journals or conferences in Machine Learning (NeurIPS, ICML, ICLR), Robotics (ICRA, IROS, RSS, CORL), or Computer Vision (CVPR, ICCV, ECCV)”

Table 7 seems to indicate that Amazon is the least engaged in fostering double affiliations. Interviewees confirmed that Amazon scholars, as the company calls academics that work part-time for Amazon, push the company to publish and present at AI conferences. This is an area of internal struggle at Amazon, which has the harshest policies in terms of never sharing confidential information when they present to others. This may explain why it relies less on this type of AI talent. Interviewees agreed that Amazon privileges internal presentations where Amazon scholars or senior academics hired as short-term consultants present what they are doing at their university and advice full-time employees, signing strict non-disclosure agreements unless they are engaged in projects where disclosure is not a problem.

> “They (Amazon scholars) are interested in publishing their research and this is why we sometimes go to conferences without sharing confidential data but yes part of the methodology. (…) We also have meetings where we present papers and get feedback specially on the science part. Amazon scholars give advice on methodologies, or suggest papers we should rely on, this culture of presenting and getting feedback on the methodologies is done internally across different teams and we have internal conferences that are larger than public conferences. Presenting there sometimes takes the same work (reviewers, specific format, the presentation, etc.) than public conferences. And obviously in external conferences you pass by a legal team (to assure you are not sharing confidential information) that can take a couple of months. It is a bit unpredictable and not that smooth, how many follow up questions they will have and how many things you will need to remove may require more work and it is a complicating factor”. (Amazon interview 1)

As with all the indicators considered so far, this result can be interpreted as part of a strategy to privilege secrecy while maximizing inflows of knowledge and information for the development of AI that is linked to Amazon’s business needs. Another example that fits into this interpretation is the AWS Cloud Credit for Research program. In 2018, last year with public data, it provided 387 credit grants to 216 different organizations, of which 49 went to the University of California, 32 to Harvard and only 9 organizations received 5 or more credit grants.[13]

---

[13] https://aws.amazon.com/government-education/research-and-technical-computing/cloud-credit-for-research/previous-recipients/

This program, as its name implies, offers AWS free credits to purchase cloud services.[14] In a context of an expanding compute divide (Ahmed & Wahed, 2020), one could expect that AWS credits would be in high demand, granting the company more leverage to define the terms of the grants and the opportunity to choose from more applications those that are the most aligned to its interests.

Credits represent an extremely low additional cost for Amazon. In the case of software as a service, they sell the use of the same lines of code sold to AWS clients. The same applies to datasets offered as a service to train AI models. In the case of processing power or storage capacity, having small projects consuming a very minor portion of its colossal infrastructure has very low opportunity costs for Amazon. Furthermore, this initiative has the potential to offer major gains to Amazon by being able to early identify, thus purchase or copy, projects that succeed. All the latter, without compromising or having to disclose any form of knowledge or data since cloud services are sold as black boxes.

Finally, there are differences in the AI talent hired as full-time employees by Big Tech. Table 8 presents figures and simple indicators of job postings for the four companies. I retrieved the number of searches that included the keywords "artificial intelligence", "machine learning" and "cloud" and compared them to the total number of job posts.

Table 8. Big Tech job postings referring to AI and related technologies

	Google	Microsoft	Amazon	Facebook
Jobs with AI or Machine Learning in the job description	251	120	321	105
Jobs with "cloud" in the job description	468	440	731	11
Total jobs posted	1044	627	3839	300
Share AI and/or ML in total jobs	24%	19%	8%	35%
Share of Cloud jobs in total jobs	45%	70%	19%	4%

Source: Author's analysis based on data extracted from Big Tech careers' websites on April 4th 2023. * To make figures more comparable among Big Tech, I did not count the following categories for calculating Amazon's total job posts: "fullfilment & operation management", "supply chain/transportation management", "business & merchant development" and "fulfillment associate".

In terms of absolute figures, it is quite telling that Amazon is the company that was hiring more AI talent and that was the 2nd company poaching AI professors according to Gofman and Jin's (2022) results. This is also the Big Tech with the fewest double affiliations (Table 7). Hiring more AI talent in is expected for a company whose strategy to remain as AI forerunner does not rely as much as the other Big Tech in building an open CIS.

Facebook ranks last in number of job posts, but interpretations should be cautious. Considering its recent weaker business and financial performance, hiring less than the other three giants shall not be a surprise. What is unexpected, nonetheless, is its high share of AI-related open positions in total job posts. An interviewee working at Facebook Reality Labs confirmed that the company was still hiring machine learning talent because it was the most crucial technology for the company. Finally, Table 8 also provides further evidence on the

[14] https://aws.amazon.com/awscredits/

difference between Google, Microsoft and Amazon, on the one hand, and Facebook, on the other, in relation to the relevance of the Cloud, which is aligned to Jacobides et al. (2021) observation of Facebook positioned differently in the AI division of labour.

5. Four strategies to build a leading AI corporate innovation system

Table 9 summarizes the previous section's findings and proposes four different strategies to organize and profit from a leading AI CIS. In one word, they could be summarized as: "frenemies" for Microsoft, "university" for Google, "secrecy" for Amazon and "application-centred" for Facebook.

Table 9. Four strategies to build a leading AI CIS

	Microsoft	Google	Amazon	Facebook
AI CIS strategy	**Frenemies**	**University**	**Secrecy**	**Application-centered**
AI Conference Presentations	+++	+++	+	+
Participation in AI conference committees	+	+++	++	++
Content of AI research	General topics with a focus on AI functional applications for language. Includes reinforcement learning	Maximum diversity with general and specific AI, including reinforcement learning	Highly diversified but skewed towards AI for language. Specific focus on time series and transfer learning	Very few direct links. Among them, "action recognition" is a specific computer vision task
Acquisitions	++	+++	+	++
Top investor	+++	++	+	-
AI patents (count)	+ (less important than in the past)	+++	+	-
AI patents (content)	Besides terms referring to more general machine learning, focus on virtual assistants and healthcare	Besides terms referring to more general machine learning, computer storage (possibly related to the cloud) and autonomous vehicles	The most diverse of the four in terms of AI functional applications	Connected to its existing platforms, with multi-terms that can be associated with the Metaverse
Double affiliations	+++ (less concentrated in the US - importance of China)	+++ (highly concentrated in the US)	+	+
Job posts	++	++	+++	+
AI CIS space	Central and global positionining, geopolitically strategic: connecting China with the West	Central and widely globalized but mainly outside Asia (China)	Core: limited to the leading AI organizations among those already doing frontier research	Narrow: it is the smallest of the four, driven by Facebook's narrower focus on AI connected to its applications/platforms
AI CIS scope	General, including research on generative AI and reinforcement learning. In terms of application fields, exhibits more focus than Amazon	General, including research on generative AI and reinforcement learning. In terms of application fields, exhibits more focus than Amazon	The most diverse in functional applications but without explicit indications of research on generative models or reinforced learning. Frontier AI is developed but only applied when there is a clear economic benefit	Focus on developing AI for its applications/platforms

Source: Author's analysis

"Frenemies" describes Microsoft's strategy. It has successfully integrated into its CIS the least expected actors, from Chinese organizations to competitors. By privileging investing in AI start-ups way more than other Big Tech, it enables those formally still separate companies to sell services to competitors, with the paradigmatic case of OpenAI while still at least partially controlling those other companies. In Microsoft's AI CIS, the development of generic AI is more inclined towards language applications, which is aligned to its investments in OpenAI. Additionally, its frenemies strategy resulted in Microsoft becoming the gatekeeper that connects AI research in Western core countries and China. This more open strategy, which does not endanger appropriation because of the speed of AI innovation and because the cutting-edge developments remain secret, is also compatible with Microsoft's lower levels of AI patenting.

The idea of frenemies was proposed by one Microsoft interviewee when I asked about the public Cloud, but I found that it also describes the company's strategy to set its AI CIS.

> "There is the question of the frenemies; and this happens a lot at Microsoft. It is a cultural shift that was brought by Satya when we moved from on premise to cloud, we had to adapt how we thought about partnerships." (Microsoft interviewee 1)

Among the four, Google excels in every indicator but, unlike Microsoft, it remains detached from China, building a CIS with top academic institutions from the rest of the world. Google's strategy for its AI CIS resembles a leading university. Among the four, it has the largest presence in AI conferences, both presenting papers and at their committees, has more employees with double-affiliations and, unlike the other Big Tech, still gives particular importance to AI patenting and acquisitions, even though it also relies on the management of internal and external knowledge flows, privileging secrecy for edge developments, just like the other giants. Google also has internal university-like features partly because many of the senior researchers have double affiliations.

> "The management style of my team is super academic, my manager is at the University of *XXX* half of the time, he is the big leader of the team and sees us as an army of postdocs" (Google, interview 1. The name of the university was removed to protect the anonymity of the interviewee)

Also, like leading universities these days, Google offers internal competitive grants for frontier projects that require more processing power to test and train AI models. To get the grant, the project needs to align with the goals of the company. Interviewees perceive Google as doing frontier AI and agree that merely incremental research that is not cutting-edge would not be funded.

It is not always straightforward how this "university" strategy turns into higher profits from AI for Google. In fact, it seems that Google's appropriation mechanisms are not translating into a clear AI business advantage to the point where it is hard to foresee where its technological lead is heading to.

> "The paper came out and we patented it but that is all (…) We wanted to have impact. With the paper we had a best paper award, and the recognition internally was really low. We made the front page of Hacker News and internally I got a bonus and that's it. We were not promoted or anything and my collaborators were frustrated. So, we tried to have impact and search for teams or Google products that would find our paper useful and tried many things, but I don't know how much of these attempts worked and eventually we moved on to new projects." (Google interview 1)

Judging by this example, the problem with the university mindset for Google looks close to those faced by leading universities that push their researchers to patent and find profitable applications for their research. Doing research takes time, researchers are not engineers or experts in translating knowledge into innovation, and applications are not just low hanging fruit once knowledge is developed. Thus, usually, universities barely profit from their patents (Popp Berman, 2011).

According to interviewees, the release of ChatGPT is perceived at Google as the crystallization of choosing the wrong strategy, too much focused on frontier AI research without sufficiently connecting it to business applications. In this context, one of Google's interviewees explained that the company is changing its focus to make AI more profitable. This change in strategy materialized by April 2023 when the company announced that DeepMind would merge with Google Brain. Just like in the university's realm, this poses greater challenges for those doing less applied research.

> "They are changing focus and people doing theory have a harder time to justify how they make profits. The big seniors, if things are going wrong, he can go to the university, they have an exit door." (Google interview 1)

On the other end, Amazon has also developed a frontier AI CIS, the most diverse in terms of functional applications, but by privileging secrecy and highly connected to its businesses. It is characterized by non-disclosure agreements, competitive grants that offer free AWS credits for secretly kept AI and an expansion of knowledge inflows from academia while discouraging outflows, such as AI conference presentations and publications. Nonetheless, it still developed a relatively prominent position in AI conferences' committees. Having such a panoptic view -which is also the case of Google- provides access to the latest AI and a space from which it can influence the field's agenda. Secrecy is at the service of Amazon's goal to produce and apply AI only when it provides a clear benefit for customers, which translates into more profits and data. As the company recently stated in response to the ChatGPT hype, for the last 25 years Amazon has been "infusing these capabilities into every business unit".[15]

A Google interviewee that had worked at Amazon acknowledged a difference in terms of secrecy and a clearer business mindset in Amazon in comparison to Google.

> "The organization of work at Amazon is much more top-down, with deadlines. What you deliver you measure it in money. You do some theoretical work, but if you can prove that it can make money." (Google interviewee 1).

Finally, while the three companies have a strong foothold in frontier generic AI, Facebook's AI CIS is narrower and remains close to applications for its platforms. Thus, it privileges participating in committees of computer vision conferences. The content of its AI patents and publications as well as the interviews reinforced this impression. This focus is coherent with the fact that Facebook does not participate in the cloud services market, which presents strong complementarities with AI and drives the other three giants to develop and offer as a service the most diverse and generic AI. As the Mercado Libre's data scientist that I interviewed stressed, ultimately, Big Tech are those that offer cloud services. When asked about the Cloud, one of Facebook's interviewees working in a strategic management position claimed that it was too late because it is a crowded market with clearly defined leaders. Moreover, this and another interviewee identified that Facebook's processes are too specific to its internal infrastructure, thus could not easily be offered as services to others.

The Metaverse could be interpreted as an attempt to expand not only Facebook's businesses but also its AI CIS, expanding it to virtual and augmented reality. However, this attempt remains narrow when compared to the AI CIS of the other three companies. As two interviewees pointed out, the Metaverse is a quite niche business.

---

[15] https://www.aboutamazon.com/news/aws/aws-amazon-bedrock-generative-ai-service?utm_source=amazonnewsletter&utm_medium=email&utm_campaign=041523&utm_term=generativeai

> "I am still not sold on the Metaverse. (…) I personally don't see the business in virtual reality, I don't use it. I barely see myself on Instagram anyway to begin with. I do think that the metaverse can be the niche of the gaming culture, but I don't really see it taking more space in the real world." (Facebook interviewee 1)
>
> "I think that it certainly helps for the revenue, but I personally don't think it would be anywhere near where the company projects it to be. The reality lab can become more mainstream and generate sizable revenue but not as big as the company wants it to be." (Facebook interviewee 2)

6. Final Remarks

The term Big Tech became vox populi and, within the tech-savvy from and beyond academia, there is ample agreement on the central role of these companies in the AI innovation system. From this starting point, the main contribution of this article has been to provide a first study of Google, Amazon, Microsoft and Facebook's different strategies to organize and extract profits from AI by organizing in distinct ways their respective AI corporate innovation systems.

I analysed multiple dimensions of their AI CISs with different quantitative techniques, from network analysis and clustering to proxy the leading AI field to text mining to get a synthetic idea of the content of these companies' AI frontier research and patents. I also considered indicators of their capacity to steer the direction of AI frontier R&D and ultimately profit from AI. This quantitative investigation was complemented with 15 in-depth semi-structured interviews. Nine were with senior managers, researchers and engineers from the four analysed companies and the remaining 6 with people working in similar positions at other digital sector companies.

The main findings point to Microsoft, Google and Amazon as building more generic AI CIS, while Facebook has a narrowed, more applied focus, concentrated on AI for its platforms. Within the former three, I also found multiple differences that can contribute to explaining their relative degrees of success. Google has opted for a sort of academic strategy. Emulating leading university traits, it has developed a cutting-edge AI CIS with an extremely central place in the non-Asian world. However, and as it often happens to universities, it is still unclear how Google will translate it AI CIS success into profitable initiatives. The remaining two companies differ in their strategy, but both seem to have better sorted out this question.

"Frenemies" describes Microsoft's strategy. It has successfully integrated into its CIS rivals from around the world. In the AI frontier research network, it is the bridging organization between Asia, in particular China, and the rest of the world. Microsoft controls not by acquiring but also by building a CIS with more organizations that are *de jure* independent but *de facto* controlled (and sometimes highly funded) by Microsoft. OpenAI is a case in point.

Finally, Amazon has developed what looks like the most diverse AI CIS in terms of functional applications, privileging secrecy and highly connected to its businesses. Secrecy is at the service of Amazon's goal to produce and apply frontier AI as long as it provides a clear benefit for customers, which translates into more profits and data.

Throughout this investigation, Amazon, Microsoft and Google cloud businesses popped up as interconnected with their AI CIS. This is an open area of research that should be tackled to better understand the shared and different strategies of these companies. Moreover, Microsoft's geopolitical role should be a subject of further investigation.

Given the results of this investigation, policy and agency should simultaneously address general and specific aspects of these companies' power over the global AI innovation system. The fact that the four Big Tech, with different degrees of success, have such a prominent role in the AI field raises several concerns considering the implications of AI for every dimension of life, from war and sovereignty to economic concentration and human rights. To change the playing field and regain public power to steer AI and better distribute its associated profits, the AI-labour market should be regulated globally. Institutions like the International Labour Organization could become the arena for discussing policies and regulating this market. For instance, regulations should prevent publicly funded academics from simultaneously work for a tech giant given the latter's larger capacity to exclusively profit from achieved results. Also, academic institutions need to be equipped with the latest digital infrastructure so that talented AI scholars can stay or be attracted to return. A survey published by Nature (2021) found that scientists working in industry are more satisfied and better remunerated than those in academia. This must be revised if the aim is to publicly redefine the purpose of AI and more evenly distribute its gains.

Public funding for AI conferences could include clauses that limit -or forbid- industry researchers in their committees. Industry researchers may well present their work in these events, but their participation in decision making spaces risks turning a public academic convening into a space controlled by private and for-profit interests. This is particularly the case of NeurIPS and Google.

In terms of specific policies, while less stringent intellectual property rights would impact Google, they will not have the same impact on the other companies, particularly for Amazon that limits knowledge outflows the most. In relation to the latter, a policy that prevents academics from signing non-disclosure agreements would limit unidirectional flows of knowledge from academic institutions to Amazon. Finally, the case of Microsoft invites to reflect on ownership structures. Corporate law should be rediscussed, and antitrust offices should not only look at the effects of mergers and acquisitions but also, as the case of Microsoft makes it clear, to major investments and preferential agreements between giants and small companies.

7. References

Abdalla, M., & Abdalla, M. (2021). The Grey Hoodie Project: Big tobacco, big tech, and the threat on academic integrity. *Proceedings of the 2021 AAAI/ACM Conference on AI, Ethics, and Society*, 287–297.

Abis, S., & Veldkamp, L. (2020). The changing economics of knowledge production. *Available at SSRN 3570130*.

Affeldt, P., & Kesler, R. (2021). *Competitors' Reactions to Big Tech Acquisitions: Evidence from Mobile Apps*.

Ahmed, N., & Wahed, M. (2020). The De-democratization of AI: Deep Learning and the Compute Divide in Artificial Intelligence Research. *ArXiv Preprint ArXiv:2010.15581*.

Alom, M. Z., Taha, T. M., Yakopcic, C., Westberg, S., Sidike, P., Nasrin, M. S., Van Esesn, B. C., Awwal, A. A. S., & Asari, V. K. (2018). The history began from alexnet: A comprehensive survey on deep learning approaches. *ArXiv Preprint ArXiv:1803.01164*.

Barbier, M., Bompart, M., Garandel-Batifol, V., & Mogoutov, A. (2012). Textual analysis and scientometric mapping of the dynamic knowledge in and around the IFSA community. In *Farming Systems Research into the 21st century: The new dynamic* (pp. 73–94). Springer.

Baud, C., & Durand, C. (2021). *Making profits by leading retailers in the digital transition: A comparative analysis of Carrefour, Amazon and Wal-Mart (1996-2019)*.

Benaich, N., & Hogarth, I. (2020). *State of AI Report*.

Benghozi, P.-J., Desmettre, S., & Gras, C. (2020). Faut-il démanteler les GAFA? *Regards Croisés Sur l'économie*, *25*(2), 235–243.

Bianchini, S., Müller, M., & Pelletier, P. (2022). Artificial intelligence in science: An emerging general method of invention. *Research Policy*, *51*(10), 104604.

Birch, K., & Bronson, K. (2022). Big tech. In *Science as Culture* (Vol. 31, Issue 1, pp. 1–14). Taylor & Francis.

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

Blanke, T., & Pybus, J. (2020). The material conditions of platforms: Monopolization through decentralization. *Social Media+ Society*, *6*(4), 2056305120971632.

Blondel, V. D., Guillaume, J.-L., Lambiotte, R., & Lefebvre, E. (2008). Fast unfolding of communities in large networks. *Journal of Statistical Mechanics: Theory and Experiment*, *2008*(10), P10008.

Bourreau, M., & de Streel, A. (2020). *Big Tech Acquisitions*. Centre on Regulation in Europe.

Bowker, G. C., & Star, S. L. (2000). *Sorting things out: Classification and its consequences*. MIT press.

Bratton, B. H. (2016). *The stack: On software and sovereignty*. MIT press.

Burt, R. S. (1995). *Structural holes: The social structure of competition*. Harvard university press.

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

Comino, S., Manenti, F. M., & Thumm, N. (2019). The Role Of Patents In Information And Communication Technologies: A Survey Of The Literature. *Journal of Economic Surveys*, *33*(2), 404–430.

Cooke, P. (2006). Global bioregional networks: A new economic geography of bioscientific knowledge. *European Planning Studies*, *14*(9), 1265–1285.

Coveri, A., Cozza, C., & Guarascio, D. (2022). Monopoly Capital in the time of digital platforms: A radical approach to the Amazon case. *Cambridge Journal of Economics*, beac044. https://doi.org/10.1093/cje/beac044

Crawford, K. (2021). *The atlas of AI: Power, politics, and the planetary costs of artificial intelligence*. Yale University Press.

European Commission. (2022). *Digital Markets Act*.

Foroohar, R. (2019). *Don't be Evil: How Big Tech Betrayed Its Founding Principles–and All of Us*. Broadway Business.

Fortunato, S., & Hric, D. (2016). Community detection in networks: A user guide. *Physics Reports*, *659*, 1–44. https://doi.org/10.1016/j.physrep.2016.09.002

Frost, J., Gambacorta, L., Huang, Y., Shin, H. S., & Zbinden, P. (2019). *BigTech and the changing structure of financial intermediation*. Bank of International Settlements.

Glick, M., & Ruetschlin, C. (2019). Big Tech Acquisitions and the Potential Competition Doctrine: The Case of Facebook. *Institute for New Economic Thinking Working Paper Series*, *104*.

Gofman, M., & Jin, Z. (2022). Artificial Intelligence, Education, and Entrepreneurship. *Journal of Finance, Forthcoming*.

Goldfarb, A., Taska, B., & Teodoridis, F. (2023). Could machine learning be a general purpose technology? A comparison of emerging technologies using data from online job postings. *Research Policy*, *52*(1), 104653.

Graef, I., & Costa-Cabral, F. (2020). *To regulate or not to regulate Big Tech*.

Hall, B. H., Helmers, C., Von Graevenitz, G., & Bondibene, C. R. (2013). A study of patent thickets. *Intellectual Property Office Research Paper*, *2013/26*.

Hendrikse, R., Adriaans, I., Klinge, T. J., & Fernandez, R. (2022). The big techification of everything. *Science as Culture*, *31*(1), 59–71.

Heston, R., & Zwetsloot, R. (2020). *Mapping US Multinationals' Global AI R&D Activity*.

Hill, K. (2020). I Tried to Live Without the Tech Giants. It Was Impossible. *New York Times*, *31*.

Hutchinson, C. S. (2022). Potential abuses of dominance by big tech through their use of Big Data and AI. *Journal of Antitrust Enforcement*, *10*(3), 443–468.

Iansiti, M., & Lakhani, K. R. (2017). Managing our hub economy. *Harv. Bus. Rev*, *10*, 117.

Jacobides, M. G. (2020). Regulating Big Tech in Europe: Why, so what, and how understanding their business models and ecosystems can make a difference. *Available at SSRN 3765324*.

Jacobides, M. G., Brusoni, S., & Candelon, F. (2021). The evolutionary dynamics of the artificial intelligence ecosystem. *Strategy Science*.

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 ArXiv:2102.01648*.

Kamepalli, S. K., Rajan, R., & Zingales, L. (2020). *Kill Zone*. National Bureau of Economic Research.

Kenney, M., Bearson, D., & Zysman, J. (2021). The platform economy matures: Measuring pervasiveness and exploring power. *Socio-Economic Review*, *19*(4), 1451–1483.

Klinge, T. J., Hendrikse, R., Fernandez, R., & Adriaans, I. (2022). Augmenting digital monopolies: A corporate financialization perspective on the rise of Big Tech. *Competition & Change*, 10245294221105572.

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

Li, Z., & Qi, H. (2022). Platform power: Monopolisation and financialisation in the era of big tech. *Cambridge Journal of Economics*.

Liu, L., Lu, G., & Xiong, W. (2022). *The big tech lending model*. National Bureau of Economic Research.

Lundvall, B.-Å., & Rikap, C. (2022). China's catching-up in artificial intelligence seen as a co-evolution of corporate and national innovation systems. *Research Policy*, *51*(1), 104395.

Manning, P., Baker, N., & Stokes, P. (2020). The ethical challenge of Big Tech's "disruptive philanthropy." *International Studies of Management & Organization*, *50*(3), 271–290. https://doi.org/10.1080/00208825.2020.1811522

Mansell, R. (2021). Adjusting to the digital: Societal outcomes and consequences. *Research Policy*, *50*(9), 104296.

Mazzucato, M., Strauss, I., O'Reilly, T., & Ryan-Collins, J. (2023). Regulating Big Tech: The role of enhanced disclosures. *Oxford Review of Economic Policy*, *39*(1), 47–69.

Nature. (2021). *Industry scores higher than academia for job satisfaction*. *600*(8). https://doi.org/10.1038/d41586-021-03567-3

Ochigame, R. (2019). The invention of 'ethical AI': How big tech manipulates academia to avoid regulation. *ECONOMIES OF VIRTUE*, 49.

Petit, N., & Teece, D. J. (2021). Innovating big tech firms and competition policy: Favoring dynamic over static competition. *Industrial and Corporate Change*, *30*(5), 1168–1198.

Pitelis, C. (2022). Big tech and platform-enabled multinational corporate capital (ism): The socialisation of capital, and the private appropriation of social value. *Cambridge Journal of Economics*.

Popkin, G. (2019). How Scientists can Team up with big Tech. *Nature*, *565*(7737), 665–668.

Popp Berman, E. (2011). *Creating the market university: How academic science became an economic engine*. Princeton University Press.

Rikap, C. (2019). Asymmetric Power of the Core: Technological Cooperation and Technological Competition in the Transnational Innovation Networks of Big Pharma.

Review of International Political Economy, *26*(5), 987–1021. https://doi.org/10.1080/09692290.2019.1620309

Rikap, C. (2020). Amazon: A story of accumulation through intellectual rentiership and predation. *Competition & Change*, *26*(3–4), 436–466. https://doi.org/10.1177/1024529420932418

Rikap, C., & Lundvall, B.-Å. (2020). Big Tech, knowledge predation and the implications for development. *Innovation and Development*, 1–28. https://doi.org/10.1080/2157930X.2020.1855825

Rikap, C., & Lundvall, B.-Å. (2021). *The Digital Innovation Race: Conceptualizing the Emerging New World Order*. Palgrave Macmillan.

Safadi, H., & Watson, R. T. (2023). Knowledge monopolies and the innovation divide: A governance perspective. *Information and Organization*, *33*(2), 100466.

Sampat, B. N. (2018). *A survey of empirical evidence on patents and innovation*. National Bureau of Economic Research.

Tancoigne, E., Barbier, M., Cointet, J.-P., & Richard, G. (2014). The place of agricultural sciences in the literature on ecosystem services. *Ecosystem Services*, *10*, 35–48.

Tarrant, A., & Cowen, T. (2022). Big Tech Lobbying in the EU. *The Political Quarterly*, *93*(2), 218–226.

Testoni, F. E., García Carrillo, M., Gagnon, M.-A., Rikap, C., & Blaustein, M. (2021). Whose shoulders is health research standing on? Determining the key actors and contents of the prevailing biomedical research agenda. *PloS One*, *16*(4), e0249661.

Trujillo, C. M., & Long, T. M. (2018). Document co-citation analysis to enhance transdisciplinary research. *Science Advances*, *4*(1), e1701130.

Van Dijck, J., Poell, T., & De Waal, M. (2018). *The platform society: Public values in a connective world*. Oxford University Press.

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.

Wasserman, S., & Faust, K. (1994). *Social network analysis: Methods and applications* (Vol. 8). Cambridge university press.

Whittaker, J. P. (2019). *Tech Giants, Artificial Intelligence and the Future of Journalism*. Taylor & Francis.

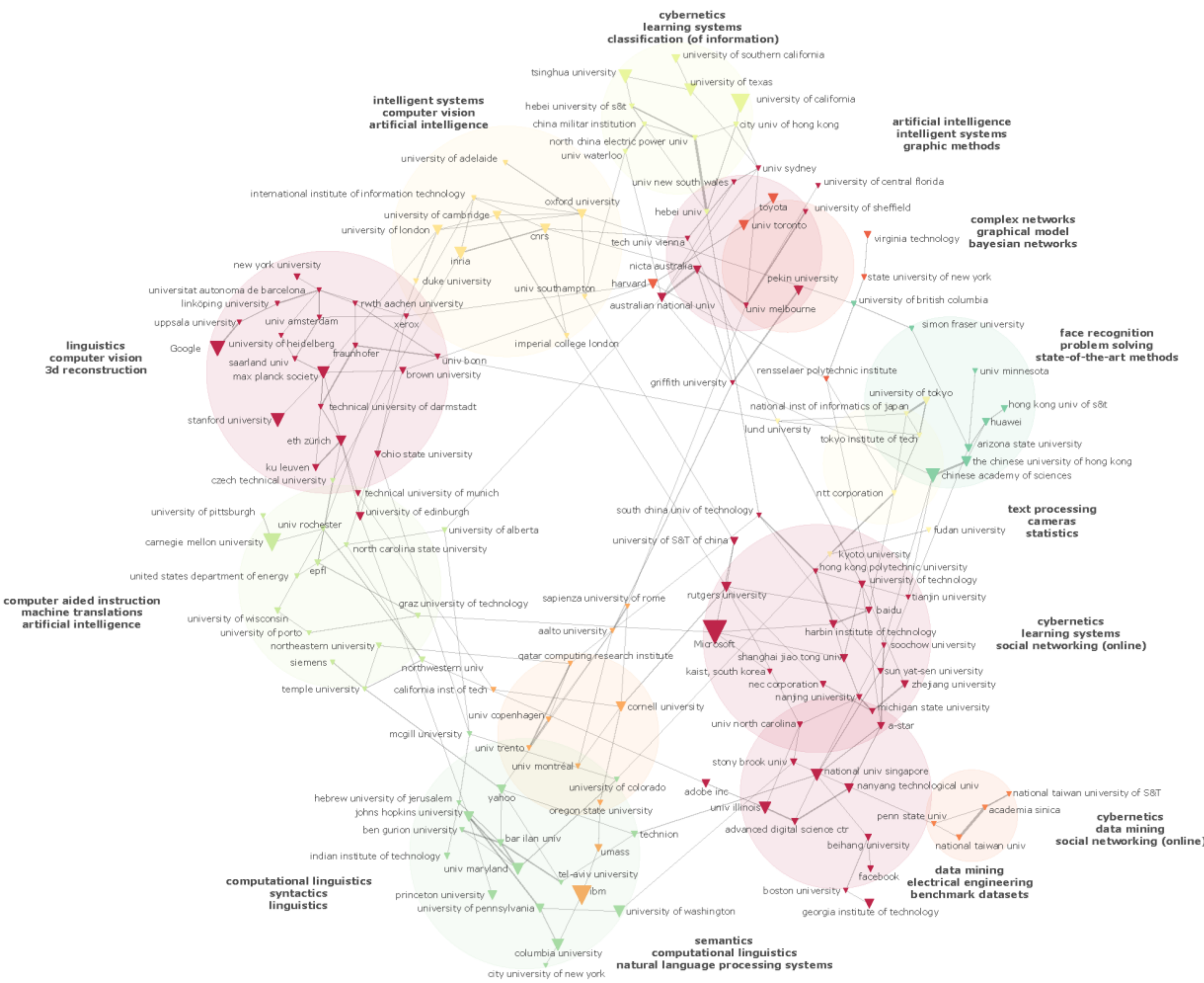
World Intellectual Property Organization. (2019). *WIPO Technology Trends 2019. Artificial Intelligence*. WIPO.

Wu, X., & Gereffi, G. (2018). Amazon and Alibaba: Internet Governance, Business Models, and Internationalization Strategies. In *International Business in the Information and Digital Age* (pp. 327–356). Emerald Publishing Limited.

Zhuang, F., Qi, Z., Duan, K., Xi, D., Zhu, Y., Zhu, H., Xiong, H., & He, Q. (2020). A comprehensive survey on transfer learning. *Proceedings of the IEEE*, *109*(1), 43–76.

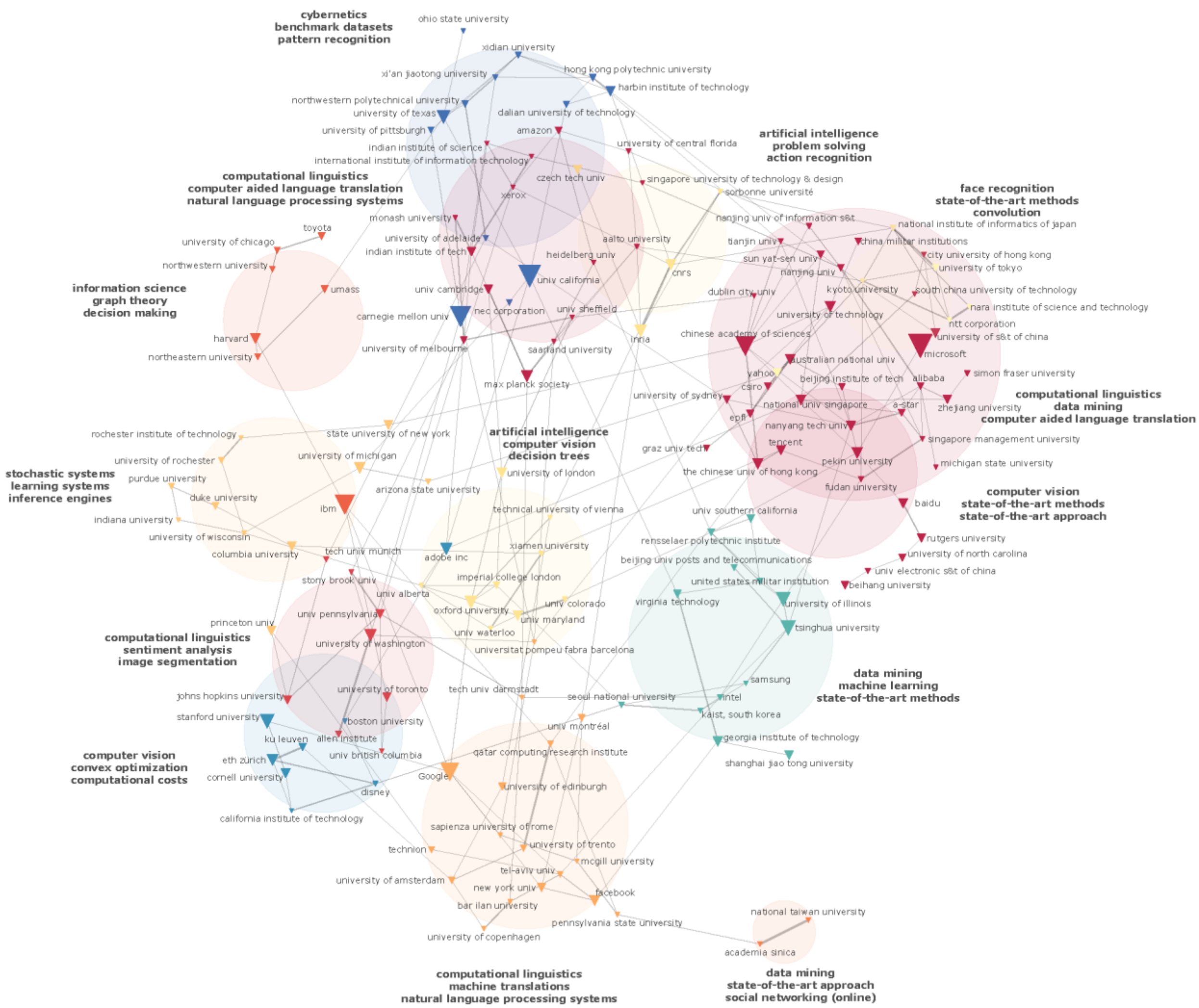
Appendix

Figure 1. AI leading research network (2012-2014). Source: Scopus.



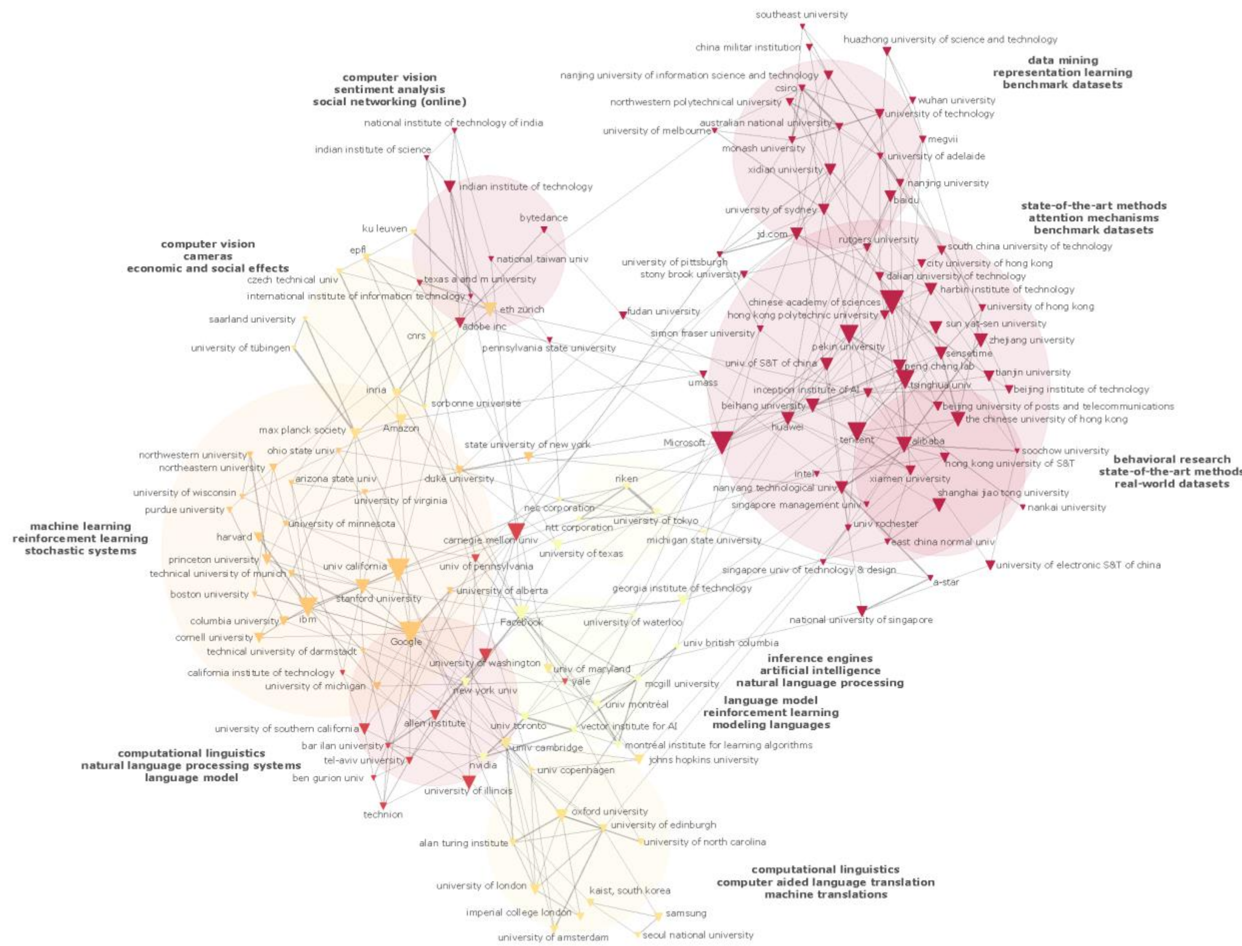
Source: Author's analysis based on a Scopus dataset

Figure 2. AI leading research network (2015-2017). Source: Scopus.



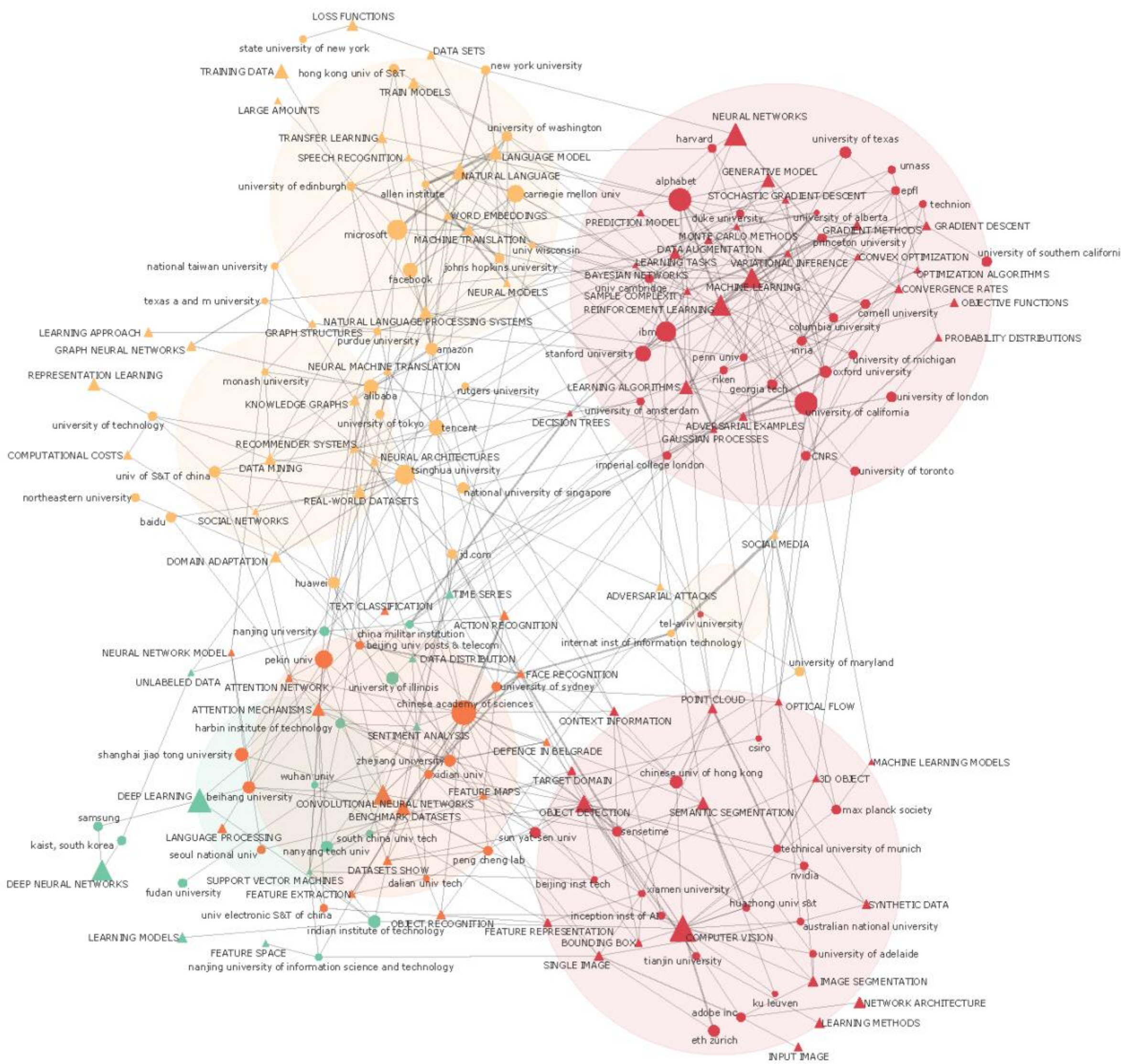
Source: Author's analysis based on a Scopus dataset

Figure 3. AI leading research network (2018-2020). Source: Scopus.



Source: Author's analysis based on a Scopus dataset

Figure 4. AI leading research network (2018-2020) and topics. Source: Scopus.



Source: Author’s analysis based on a Scopus dataset

Table A.1 Content of Big Tech AI patents in 2022

Microsoft	Google	Amazon	Facebook
computing system	including computer programs	image data	neural network
computer program product	computer storage medium	audio data	machine learning model
computing device	computer storage media	machine learning model	assistant systems
neural network	autonomous driving mode	neural network	computer-readable media
machine learning model	sensor data	autonomous mobile device	client system
machine learning	computing device	machine learning	processor system
clinical documentation	more processors	data representative	client system of a first user
training data	computer programs	input data	online system
user interface	computing system	video content	content item
input data	machine learning	other data	dot product
patient encounter	machine learning model	input image	matrix processor unit
sensor data	map information	physical space	computing system
encounter information	client computing device	motile device	artificial reality system
image data	autonomous mode	computer-readable media	result matrix
user input	data bus	using data	data matrix
object detection	computing unit	systolic array	video frames
user experience	input activation	time series	output image
machine train	plurality of cells	service provider	sending instructions
video game	process the image	communication devices	disclosed computer-implemented method
data samples	more computers	output data	multiplication results
conversation prints	encoder neural network	provider network	Mapping convolution
physical environment	cause the vehicle	Natural language	artificial reality environment
audio encounter information	processing unit	voice commands	hardware channel convolution processor unit
vision system	input activation value	automatic speech recognition	sensor data
context information	Example implementations	user device	head-mounted display
virtual assistant	perception system	various embodiments	calculation units
storage device	vehicle occupancy	process images	neural network model
more sensors	MAC operator	weight changes	corresponding channel convolution result matrix
computer program	first memory bank	items on the shelf	more computing systems
obtaining encounter information of a patient encounter	matrix multiplication	processing element	convolution weight matrices

Source: Author’s analysis based on data extracted from Derwent Innovation.