WORKING PAPER | ISSUE 14/2023 | 18 JUNE 2023

COMPETITION IN GENERATIVE ARTIFICIAL INTELLIGENCE FOUNDATION MODELS

CHRISTOPHE CARUGATI

Foundation models (FMs) are the origin of breakthrough innovations in generative artificial intelligence (AI) applications, such as ChatGPT. Only responsible developments in competitive markets can help ensure that FMs deliver their full benefits at minimum risk.

FM developers require language models (LMs), data and computing power to generate natural language output, such as texts, from language input. Thus, the FM value chain is composed of three main elements: LMs, data and computing resources.

These markets are currently competitive, with multiple providers and degrees of openness thanks to several closed- and open-source models, open-source and proprietary data, and vigorous competition between firms at the computing-resources level, despite high degrees of concentration in some of these markets. These market characteristics ensure that FM developers face low or surmountable entry barriers.

Still, potential competition issues are likely to arise in the future. Dominant firms could leverage their dominant positions, refuse to give access to their LMs, scrape data, refuse to grant access to data, impose undue barriers to switching and lock their users into their ecosystems. Firms could also use LMs to achieve an anticompetitive agreement through algorithmic collusion.

Competition authorities should focus their efforts on short-term risks. They should also remain vigilant in terms of ensuring the competitive process between open- and closed-source models works and that open-source developers and public authorities do not impose undue restrictions to mitigate the risks of open-source models, which would deter their development in a way that would favour closed-source models. Finally, at this development stage, studies are lacking. Researchers and competition authorities should investigate the impact of fms on content providers and the digital advertising industry, the role of FMs in digital ecosystems, and cooperation mechanisms between competent authorities across regulatory fields and countries.

Christophe Carugati (christophe.carugati@imf.org) is a Research Fellow at Bruegel.

bruegel

Recommended citation:
Carugati, C. (2023) 'Competition in generative artificial intelligence foundation models', *Working Paper* 14/2023, Bruegel

1 Introduction

The release of ChatGPT by Open AI in November 2022 highlighted the role of language models (LMs), also known as foundation models (FMs), at the origin of breakthrough innovations in generative artificial intelligence (AI) applications.

Generative AI applications, which generate an output in response to a prompt in natural language – in other words, provide an answer to a question – have the potential to impact the economy dramatically. Some studies have found that LMs could impact 80 percent of workers in the United States (Eloundou *et al*, 2023) and could increase annual global GDP by 7 percent over ten years (Hatzius *et al*, 2023). LMs could also pose considerable risks to society. Malicious use of AI by humans could result in increasing discrimination, misinformation and disinformation. Other risks include overreliance on AI-generated content, divulging of sensitive information and environmental harm because of the energy required to use LMs (OECD, 2023b).

In this context, competition authorities can help ensure that the full benefits of FMs are delivered responsibly. Currently, a number of firms are innovating and competing to develop LMs and AI-powered applications, such as conversational chatbots. While these innovations occur quickly, there have been warnings that LMs might end up in the hands of only a few large technology firms with dominant positions in the digital sector. Large firms like Google and Microsoft have significant financial resources and key infrastructure such as cloud computing services[1]. So far, competition authorities have responded to competition concerns in several ways. In the US, the Federal Trade Commission (FTC) has said there is a need for regulatory intervention to avoid market concentration[2]. In the United Kingdom, the Competition and Markets Authority (CMA) has launched a review of AI FMs to understand how the market works[3]. In particular, the CMA examines FM entry barriers and their impact on competition in other markets. In France, the French *Autorité de la concurrence* (Adlc) has said that LMs will drive growth in demand for cloud services, thus impacting how the cloud market works[4].

[1] Diane Coyle, 'Preempting a Generative AI Monopoly', *Project Syndicate*, 2 February 2023, https://www.project-syndicate.org/commentary/preventing-tech-giants-from-monopolizing-artificial-intelligence-chatbots-by-diane-coyle-2023-02.
[2] Lina Khan, 'We Must Regulate A.I. Here's How', *The New York Times*, 3 May 2023, https://www.nytimes.com/2023/05/03/opinion/ai-lina-khan-ftc-technology.html. See also FTC (2023)
[3] See CMA press release of 4 May 2023, https://www.gov.uk/government/news/cma-launches-initial-review-of-artificial-intelligence-models.
[4] Autorité de la concurrence, Summary of Opinion 23-A-08 Of 29 June 2023 on Competition in the Cloud Sector.

Accordingly, this policy contribution examines how competition in FMs works. First, it defines FMs and entry barriers. It then assesses the potential competition issues. Finally, it outlines policy considerations, with some recommendations and future research questions.

2 Understanding foundation models

2.1 Definition

FMs are language models that use a large volume of input data to derive outputs, including texts, images, songs and videos. LMs are indispensable for running natural language processing (NLP) tasks, which automate functions such as summarising information or question-and-answer (OECD, 2023b). FMs rely on three essential resources: language models, data and computing resources.

2.1.1 Language models

LMs learn to perform tasks by analysing training examples from datasets. The models identify patterns from what they learn during the training phase in a machine learning (ML) process. For instance, models trained on data about cats learn to recognise and automatically identify cats. Most current LMs, such as generative pretrained transformer (GPT) models, predict the next input, such as a word, in sequence. In doing so, they enable faster training of larger datasets, which are broken down into pieces of words known as 'tokens'. This is essential for LMs because they are pretrained on large datasets to perform their tasks.

Several firms and research institutions, including Amazon, Anthropic, Baidu, Deepmind, Google, Hugging Face, Meta, OpenAI, Beijing Academy of AI (BAAI) and Yandex, have developed large language models (LLMs). These models are either proprietary closed-source or open-source. Developers of closed-source models, such as Open AI-owned GPT, licence their models to third parties for a fee. In return, third parties develop commercial applications built with the models but cannot modify the underlying models.

By contrast, developers of open-source models allow third parties to modify them, helping improve the models and correcting errors. However, malicious third parties can also exploit the models' vulnerabilities, posing risks for end-users (OECD, 2023b). Open-source models do limit their permitted uses by third parties. While some open-source providers allow use of their models for commercial applications, some impose usage restrictions[5]. Others only allow the use of their models for noncommercial research.

In addition to their characteristics as open- or closed-source, the performance of LLMs and their costs depend on the number of trainable parameters (Rae *et al*, 2022). The more trainable parameters, the more the models can learn from datasets. However, more parameters require more data and computing power, thus increasing the model's cost. Researchers are already developing small language models (SLMs) that rely on fewer trainable parameters, in order to reduce financial and environmental costs while achieving the same performance as some LLMs (Schick and Schütze, 2021)[6]. Researchers also propose to develop models by fine-tuning existing models with the Low-Rank Adaptation (LoRA) technique that enables fine-tuning on significantly less trainable parameters and computing power while achieving similar performance as the fine-tuning of LLMs on a higher number of parameters (Hu *et al*, 2021).

FMs have two main costs. The first occurs during the training phase when the models learn patterns from the datasets. During this phase, the models mobilise intense computing power for several weeks. In addition, models generally require fine-tuning on specific datasets to perform their desired tasks, thereby increasing computing power costs. The second occurs during the inference phase when the models generate an output following a prompt from the user. The inference costs depend on the number of generated tokens, which only require computing power during inference (Hoffmann *et al*, 2022). In addition, developers incur costs to store datasets in data centres.

Table 1 presents a non-exhaustive list of notable LLMs to compare valuable characteristics, including model type (closed-source/open-source), model size and permitted uses.

[5] For instance, Bloom prohibits use of its model in legal work, such as for writing legal contracts. See Bloom license: https://huggingface.co/spaces/bigscience/license.
[6] According to OECD (2023b), a non-exhaustive list of SLMs includes Google (ALBERT), Facebook (BART), Kakao (KoGPT), Google-owned DeepMind (RETRO), Baidu (ERNIE 3.0) and BigScience (T0).

Table 1: Overview of notable LLMs

Model	Release date	Developer	Type	Model size	Permitted use
Bloom	2022	BigScience	Open-source	176 billion (B)	Both commercial and non-commercial with restrictions
GPT-4	2023	OpenAI (investor included Microsoft)	Closed-source	1 trillion (Tr) (estimated)	Non-commercial and commercial
PaLM	2022	Google	Closed-source	540B	Non-commercial and commercial
LLaMA	2023	Meta	Open-source	65B	Non-commercial
ERNIE 3.0 Titan	2021	Baidu	Open-source	260B	Non-commercial and commercial
Wu Dao 2.0	2021	BAAI	Open-source	1.5Tr	Non-commercial and commercial
YaLM	2022	Yandex	Open-source	100B	Non-commercial and commercial
Claude	2023	Anthropic (investor included Google)	Closed-source	52B	Non-commercial and commercial
Amazon Titan FMs	2023	Amazon	Closed-source	N/S	Non-commercial and commercial
Jurassic-2	2023	AI21 labs	Closed-source	N/S	Non-commercial and commercial

Source: Bruegel and OECD (2023b). Note: N/S means non-specified. See https://huggingface.co/bigscience/bloom, https://openai.com/product/gpt-4, https://ai.googleblog.com/2022/04/pathways-language-model-palm-scaling-to.html, https://ai.facebook.com/blog/large-language-model-llama-meta-ai/, http://research.baidu.com/Blog/index-view?id=165, https://www.baai.ac.cn/english.html#Research, https://yandex.com/company/press_center/press_releases/2022/2022-23-06, https://www.anthropic.com/product, https://aws.amazon.com/bedrock/titan/, https://docs.ai21.com/docs/jurassic-2-models.

2.1.2 Data

Data includes information containing texts, images, songs, videos and computer code. In data-driven markets, research has found that data is a competitive asset characterised by '4Vs': volume (size of the dataset), variety (different information contained in the dataset), velocity (freshness or speed at which data is collected), and value (economic relevance) (Stucke and Grunes, 2016)[7]. Market studies by competition authorities in the online advertising sector have also found that data quality is important for competition (Autorité de la concurrence, 2018). Data plays a vital role in FMs, as models

[7] For an application of the 4Vs in a competition case, see the assessment of Apple's acquisition of Shazam in 2018 by the European Commission; M.8788 *Apple / Shazam*, 6 September 2018.

are pre-trained and fine-tuned on data. In this context, the abovementioned data characteristics are also relevant when assessing data in FMs.

First, LMs rely on large volumes and varieties of data, with better performance the more and more varied the data is (Kaplan *et al*, 2020). Developers pre-train their models by building datasets from proprietary and public data from the web, including, for instance, books, news, Wikipedia and GitHub. These datasets are either closed sources, such as the DeepMind-owned MassiveText dataset (Rae *et al*, 2022), or open source from community websites, such as the HuggingChat Open Assistant OASST1 dataset[8] or the Amazon Massive dataset[9].

Second, research suggests that data quality matters as much as volume and variety. For instance, developers of the Koala model used the open-source model LLaMA and fine-tuned it on high-quality open-source datasets containing dialogue data from the web, including dialogue with LLMs, such as ChatGPT. They found that the model performed similarly to closed-source LLMs trained on much larger volumes of presumably lower-quality data. Maximising data quality over quantity has the advantage of dramatically reducing the need for computing power and, thus, model training costs. Developers trained the Koala model for only six hours, costing less than $100 (Geng *et al*, 2023).

Third, the velocity or freshness of data is a major obstacle for most models, as they are pre-trained on data collected up to a certain point. This implies that these models do not contain up-to-date data and are thus unable to complete their tasks using the most recent information. To overcome this, some developers, including Microsoft, Google, Open AI and Neeva[10], run their models on up-to-date datasets of content discovered after crawling the web (index data) or click-and query (search data)[11].

Finally, data has significant economic value, especially because developers improve their models by training them on user dialogue data[12]. Data-driven network effects arise when user utility changes with improved learning from data, creating value for users (Gregory *et al*, 2021). This value relies on the improvement in the ability of the models to perform tasks for which user dialogue was relevant, and not on their ability to perform uncommon tasks for which they could not learn from user queries. Models can satisfactorily perform their tasks on any query based on the instructions describing the

[8] The dataset is available at https://huggingface.co/datasets/OpenAssistant/oasst1.
[9] The dataset is available at https://github.com/alexa/massive.
[10] Neeva stopped offering its service in June 2023.
[11] For instance, Microsoft Bing relies on its own index data. Jordi Ribas, 'Building the New Bing', *Microsoft Bing Blogs*, 21 February 2023, https://blogs.bing.com/search-quality-insights/february-2023/Building-the-New-Bing.
[12] For instance, Open AI collects data from conversations with ChatGPT to train and improve its models. See 'Data Controls FAQ', *Open AI*, https://help.openai.com/en/articles/7730893-data-controls-faq.

task, known as 'zero-shot learning' (Kojima *et al*, 2023). In other words, the models do not need to learn the appropriate answer to each query. They require only a description of the task. In contrast, most data-driven services, such as general search engines, require several prompts for uncommon queries to provide a satisfactory response to a query[13].

2.1.3 Computing resources

Computing resources – processing hardware, servers, supercomputers, and networking equipment – are essential components for running LMs. In practice, FM developers rent these components from cloud providers[14].

LMs, particularly LLMs, require significant computing power to train and run. Processing hardware includes central processing units (CPUs) to interpret and execute computations, graphics processing units (GPUs) to perform several computations simultaneously, random access memory (RAM) to store intermediate computations while running the models, and tensor processing units (TPUs) or natural processing units (NPUs) to accelerate machine-learning workloads. Table 2 outlines the global market share of the primary providers according to the processing hardware category. The table shows various concentration levels in these markets, with some critical markets dominated by only one or two players, such as Intel in the CPU market or Nvidia in the GPU market. It is worth mentioning that some firms, such as Meta, have announced the development of in-house chips for AI applications[15].

---

[13] AT.39740 *Google Search (Shopping)*, 17 June 2017, paras. 287 and 288.
[14] Other components include power supplies to provide electricity to the hardware, and cooling systems to cool off the heat generated by the hardware, avoiding damage and under-performance.
[15] Santosh Janardhan, 'Reimagining Our Infrastructure for the AI Age', *Meta Blog*, 18 May 2023, https://about.fb.com/news/2023/05/metas-infrastructure-for-ai/.

Table 2: Global market share of the primary providers of computing hardware by category

Provider	CPU	GPU	RAM	TPU/NPU
Intel	68.7% (2022)[16]	9% (2022)[17]	N/A	N/A
Advanced Micro Devices (AMD)	31.3% (2022)[18]	9% (2022)[19]	N/A	N/A
Nvidia	N/S	82% (2022)[20]	N/A	N/A
Micron Technology	N/A	N/A	26.4% (2022)[21]	N/A
Samsung Electronics	N/A	N/A	40.7% (2022)[22]	N/A
SK Hynix	N/A	N/A	28.8% (2022)[23]	N/A
Google	N/A	N/A	N/A	N/S

Source: Bruegel (see footnotes). Note: N/A = non-applicable. N/S = non-specified.

Servers are computers that deploy LMs to users. Developers train and run their LMs from servers that rely on the abovementioned processing hardware. Servers also store data and model parameters. AI servers are offered by firms including Inspur, Dell and Hewlett Packard Enterprise (HPE)[24].

Supercomputers are specialised computers built with thousands of processing hardware units, particularly CPUs and GPUs, to train and run LLMs at scale. Some firms, including Google, Microsoft, Meta and Nvidia, have developed supercomputing technologies by investing substantial financial resources. For instance, Microsoft partnered with Open AI in 2019 by investing $1 billion in developing its supercomputer, which has 285,000 CPUs and 10,000 GPUs[25], on its Microsoft Azure cloud service[26]. In 2023, Microsoft reinforced its investment with an additional $10 billion[27].

[16] Stephen Nellis, 'AMD Wins Nearly A Third of Processor Market, Arm's Climb Slows: Analyst Report', *Reuters*, 9 February 2023, https://www.reuters.com/technology/amd-wins-nearly-third-processor-market-arms-climb-slows-analyst-report-2023-02-09/.
[17] Data is for discrete GPUs (GPU is separate from the processor). See Wallstreetzen, 'Nvidia Corp Statistics & Facts', https://www.wallstreetzen.com/stocks/us/nasdaq/nvda/statistics.
[18] Stephen Nellis, 'AMD Wins Nearly A Third of Processor Market, Arm's Climb Slows: Analyst Report', *Reuters*, 9 February 2023.
[19] Wallstreetzen, 'Nvidia Corp Statistics & Facts'; see footnote 17.
[20] Ibid.
[21] Data is for Dynamic Random Access Memory (DRAM), 'DRAM Manufacturers Revenue Share Worldwide From 2011 to 2022, by Quarter', *Statista*, https://www.statista.com/statistics/271726/global-market-share-held-by-dram-chip-vendors-since-2010/.
[22] Ibid.
[23] Ibid.
[24] 'Artificial Intelligence (AI) Server Vendor Market Share Worldwide in the First Half of 2020', *Statista*, https://www.statista.com/statistics/1227556/ai-server-vendor-market-share/.
[25] Doug Black, 'Microsoft's Massive AI Supercomputer on Azure: 285k CPU Cores, 10k GPUs', *HPC Wire*, 20 May 2020, https://www.hpcwire.com/2020/05/20/microsofts-ai-supercomputer-on-azure-combinations-of-perceptual-domains/.
[26] Greg Brockman, 'Microsoft Invests in and Partners with Openai to Support Us Building Beneficial AGI', *Open AI*, 22 July 2019, https://openai.com/blog/microsoft-invests-in-and-partners-with-openai.
[27] Dina Bass, 'OpenAI Needs Billions to Keep ChatGPT Running. Enter Microsoft', *Bloomberg*, 26 January 2023, https://www.bloomberg.com/news/articles/2023-01-26/microsoft-openai-investment-will-help-keep-chatgpt-online.

Networking equipment is hardware that transfers data quickly between machines. It enables the training of LMs across multiple machines using distributed computing. In other words, the networking equipment connects multiple servers.

Cloud providers offer computing capabilities by renting hardware resources to their users, including processing hardware, storage, servers and supercomputing technologies over the internet. Cloud customers can easily use the hardware resources they need by scaling them up or down without investing in the infrastructure. Several providers, including Google (Google Cloud Platform), Amazon (Amazon Web Services) and Microsoft (Microsoft Azure), offer cloud capabilities to run LMs[28]. Some firms, including Google, Microsoft, Amazon and Nvidia, also provide cloud services to deploy and customise LMs[29]. The market is growing and competitive. Data for Q1 2023 shows that the cloud market is growing rapidly, with an annual growth rate of 20 percent compared to Q1 2022, and that Google (10 percent), Amazon (32-34 percent), and Microsoft (23 percent) account for 65 percent of the global market share[30]. The trend towards an oligopolistic cloud sector in the hands of Amazon, Google and Microsoft, has led competition authorities, including those in the UK, US, France, Netherlands, Japan and South Korea, to investigate competition issues in the cloud sector[31].

2.2 Entry barriers

In data-driven markets, entry barriers faced by firms typically include data advantage, network effects and economies of scale and scope (Cremer *et al*, 2019). In FM markets, barriers exist at the LM, data and computing resource levels.

2.2.1 At the LM level

The FM market is competitive and highly dynamic, with frequent releases of LMs of different sizes and degrees of openness from multiple providers. This prevents the emergence of a dominant incumbent

---

[28] Other cloud providers include Alibaba Cloud, Oracle Cloud, IBM Cloud (Kyndryl), Tencent Cloud, OVHcloud, DigitalOcean and Linode. Mary Zhang, 'Top 10 Cloud Service Providers Globally in 2023', *Dgitl Infra*, 1 January 2023, https://dgtlinfra.com/top-10-cloud-service-providers-2022/.

[29] See for instance the Nvidia NeMo framework, *Nvidia*, https://www.nvidia.com/en-us/ai-data-science/generative-ai/nemo-framework/.

[30] Synergy Research Group, 'Q1 Cloud Spending Grows by Over $10 Billion From 2022; the Big Three Account for 65% of the Total', 27 April 2023, https://www.srgresearch.com/articles/q1-cloud-spending-grows-by-over-10-billion-from-2022-the-big-three-account-for-65-of-the-total.

[31] See, for the US, Federal Trade Commission, 'An Inquiry into Cloud Computing Business Practices: The Federal Trade Commission is Seeking Public Comments', 22 March 2023, https://www.ftc.gov/policy/advocacy-research/tech-at-ftc/2023/03/inquiry-cloud-computing-business-practices-federal-trade-commission-seeking-public-comments. The press release mentions the links to the other investigations. In France, the French competition authority released its final report on 29 June 2023. French competition authority, Opinion 23-A-08 of 29 June 2023 on competition in the cloud sector.

player. Moreover, as the demand for LMs increases, new entrants and LMs will likely appear frequently. Developing LMs, especially LLMs, requires substantial financial resources to train and run models, and human resources to create models. However, two main factors could reduce the need for financial and human resources substantially in the future.

First, technical developments in SLMs and fine-tuning of LLMs could substantially reduce the financial resources required to develop models that perform as well as LLMs. Moreover, as the Koala model shows, FM developers can develop their models without investing in the infrastructure because they can train and run their models from cloud providers by scaling up or down their needs without significant investment (less than $100 to train the Koala model). These techniques are significant as they might mean that FM developers do not need to develop an LLM, especially when fine-tuning a pre-existing model for specific tasks.

Second, the availability of a community of developers and open-access research can substantially lower the human resources needed. While developers require in-house staff, valuable resources are readily available from community websites, such as HugginFace or GitHub, and open-source academic repositories, such as arXiv. Developers can thus find talents externally without the need to recruit them internally. In other words, unless developers want to create LLMs in their hardware infrastructure, barriers to entry can be relatively low.

2.2.2 At the data level

The need to collect a large volume and variety of data is a significant entry barrier. Some developers, such as Google, benefit from the data advantage of creating models based on their vast proprietary data. However, developers can also develop models from open-source data available on community websites and other websites that offer open-source data, eliminating the need to collect proprietary data[32]. In addition, the Koala model shows that developers can build their models by fine-tuning pre-trained models on high-quality open-source data, while achieving performance similar to that of LLMs trained on much larger volumes of proprietary data.

However, some barriers to the freshness of the data remain. Only a few providers, including Google and Microsoft, provide access to up-to-date web index and search data. In Europe, developers of search engines seeking access to search data can rely on the Digital Markets Act (DMA, Regulation (EU)

[32] For instance, around 37,000 open-source datasets are available on Hugging Face; see https://huggingface.co/datasets?sort=downloads.

2022/1925), which requires large search engine providers that fall within its scope, known as 'gatekeepers', to share search data, including ranking, query, click and view data, with competing search engine providers (Article 6(11) DMA). However, not all LMs require real-time data from the web. Those who do not need it can ensure the freshness of their data via an application programming interface (API), allowing the models to run on fresh proprietary or open-source data.

Therefore, the availability of open-source data, the ability to fine-tune pre-trained models on proprietary or open-source data, and the ability to run models on fresh data lower entry barriers significantly at the data level.

2.2.3 At the computing resources level

FM developers rely on a few firms to access computing power via CPUs, GPUs and TPUs. Barriers to entry into processing hardware markets are high, including significant levels of investment in research and development, economies of scale and existing partnerships with FM developers.

FM developers also rely on cloud providers, as they can train and run their models with the required computing power without investing in the hardware. The cloud sector is competitive and highly dynamic, frequently releasing new offerings to run and customise LMs. While the largest cloud providers might have a competitive advantage owing to economies of scale, access to large financial resources and the provision of their own LMs exclusively on their cloud services, both large and small players compete vigorously to attract users. The largest players have released new offerings to run and develop LMs, as Amazon does with Amazon Bedrock, Microsoft with Azure Machine Learning and Microsoft Azure Open AI service, and Google with Google Cloud. Smaller players also compete with them by offering cloud capabilities for AI[33]. Nonetheless, as discussed in section 2.3.1, cloud reports by competition authorities found barriers to switching that might prevent some cloud users from migrating from one cloud provider to another, thus preventing competition once the users adopt a cloud provider. Competition laws and regulations, such as the proposed European Data Act, can address barriers to switching. However, if users do not switch because they are satisfied with the cloud provider, it is doubtful that such tools will foster switching between cloud providers.

[33] For example, OVHcloud; OVHcloud, AI & machine learning, https://www.ovhcloud.com/en/public-cloud/ai-machine-learning/.

3 Potential competition issues

3.1 At the LM level

A firm with a dominant position in one market might abuse its position by leveraging one of its products or services from that market in order to promote a product or service in a market in which it is not dominant, known as "leveraging". A dominant firm in a given market can also restrict or refuse access to its LM. Dominant and non-dominant firms can also use LMs to achieve an anticompetitive agreement, known as 'algorithmic collusion'.

3.1.1 Leveraging

An FM is useful only if used with other products and services that are outside the FM's value chain, such as software, search engines and cloud services, to ensure the dissemination and uptake of the FM. In this context, some FM developers can leverage their dominant position in one market, which is outside of the FM's value chain, to promote their LM or AI-powered applications. There are three main types of leveraging practice. First, a firm can promote its own services over its competitors, known as self-preferencing. Second, a firm can make the acquisition of one product or service (the tying product) conditional on the purchase of another (the tied product), known as tying. Third, a firm can offer different products or services together, known as bundling (OECD, 2020). Leveraging breaches EU competition law only if it has a potential anticompetitive effect that outweighs any procompetitive effects by foreclosing rivals, thus excluding them from the market.

While these commercial practices use different means, they all leverage a product or service that has a dominant position (the dominant product) to increase sales of a product or service that is subject to competition (the competitive product), or to prevent the competitive product from competing with the dominant product. In the digital economy, leveraging is very common. Digital services often interact as they share inputs, such as data, and have overlapping consumers, allowing digital firms to propose an ecosystem of services.

Box 1 shows that over the last 20 years, competition authorities have found that several digital firms had the ability and incentive to leverage their dominant positions to deny competitors sufficient network effects (eg the European Commission *Microsoft/LinkedIn* merger remedies), prevent competitors from entering the competitive market (eg the *Microsoft Windows Media Player* case), protect the dominant market (eg the *US Microsoft Internet Explorer* case), and enter and foreclose competitors in a competitive market (eg the *Google Search (Shopping)* case).

Box 1: Notable digital leveraging cases

In 2016, the European Commission cleared Microsoft's acquisition of the professional social networking service (PSN) LinkedIn, conditional on behavioural remedies. The remedies aimed to address the competition concerns that Microsoft had the ability and incentive to leverage its dominant position in the PC Operating System (OS) and productivity software markets with its Windows OS and Office Suites to promote LinkedIn by preinstalling LinkedIn in its OS and imposing interoperability restrictions on competing PSNs in its Office Suites. This practice would have prevented competing PSNs from gathering sufficient user bases and network effects to compete with LinkedIn, excluding them from the market[34].

In 2004, the Commission found that Microsoft breached EU competition law by leveraging its dominant Windows OS (the tying market) by tying it with its media player Windows Media Player (WMP) (the tied market) by preinstalling WMP on Windows OS. This practice had the effect of foreclosing rivals from entering the tied market[35]. In Europe, the Digital Markets Act now prevents this practice by requiring gatekeepers to allow end-users to uninstall any software applications, including preinstalled software, and to easily change the default settings on the OS of products or services provided by the gatekeepers (Article 6(3) DMA). It also requires gatekeepers to allow the installation of third-party software applications or software application stores within the OS (Article 6(4) DMA).

In 2009, the Commission agreed with Microsoft to solve the alleged leveraging of its Windows OS (the tying market) by tying it with its web browser Microsoft Internet Explorer (IE) by preinstalling Microsoft IE on Windows OS. As in the *Microsoft WMP* case, the practice would have foreclosed rivals' web browsers from entering the tied market while strongly pushing content providers and software developers to develop their offerings primarily for Microsoft IE[36]. In the US, the Department of Justice (DOJ) in 1998 filed a lawsuit for similar tying practices. However, the DOJ considered that the tying practice would have protected Windows OS in the tying market by hindering the growth of Netscape Navigator, a competing web browser in the tied market. The aim would have been to limit the risk of the emergence of competing operating systems by restricting software developers from using Netscape's programming language, which allowed developers to design software for other operating systems[37].

[34] M.8124 *Microsoft / LinkedIn*, 6 December 2016, paras. 278-352.
[35] COMP/C-3/37.792 Microsoft, 24 March 2004. The other competition concern related to a refusal to supply interoperability information to competing providers of work group server.
[36] COMP/C-3/39.530 Microsoft Commitment, 1 December 2009.
[37] *United States v. Microsoft Corporation*, 253 F.3d 34, 2001.

In 2017, the European Commission found that Google abused its dominant position in general search engine services to promote its comparison shopping service (CSS) Google Shopping, by displaying and positioning its service more prominently in its search results and demoting rivals by applying adjustment algorithms that did not apply to Google's own CSS. This practice allowed Google to enter the CSS market and foreclose its competitors[38]. It is worth noting that the EU General Court upheld the Commission's decision in 2021 in the first instance, but the case is still pending before the EU Court of Justice[39]. In Europe, the DMA bans this practice by preventing gatekeepers from promoting their services over a similar third-party's products or services in crawling, indexing and ranking (Article 6(5) DMA)[40].

FM developers may have the ability and incentive to leverage their dominant positions in one market to promote through self-preferencing or tying products or services, including their own FMs or AI-powered applications, in a competitive market in which they are not dominant. In hypothetical scenarios, like in the *Google Search (Shopping)* case, they could promote their own generative AI-powered answer engine over rivals via their search engine service, foreclosing competing answer engine providers such as HuggingChat or ChatGPT[41]. Like in the *Microsoft Windows Media Player* case, they could tie their own AI-powered applications by preinstalling them on their OSs to make it more difficult for competitors to compete. Finally, they could tie the development of third-party AI-powered applications to their dominant software or OS by requiring software developers to develop them with their own FMs to prevent the growth of competing FMs.

3.1.2 Refusal to grant access to LMs

Some FM developers allow access to their closed-source models, enabling third parties to build applications on top. Developers have the ability and a strong incentive to allow access with broad conditions to ensure the widespread adoption of their models and the creation of an ecosystem of applications. However, some FM developers might also want to impose restrictive access conditions on developers that design competing functionalities to their own AI-powered applications to protect their applications from competing services. In other words, the restrictions imposed by the FM developer would block access to its FM to rivals that want to use the FM to develop competing functionalities. Such a scenario would be similar to the alleged anticompetitive practice in the 2020 US

[38] AT.39740 *Google Search (Shopping)*, 27 June 2017.
[39] T-612/17 *Google and Alphabet v Commission (Google Shopping)*, ECLI:EU:T:2021:763, 10 November 2021. C-48/22 P *Google and Alphabet v Commission (Google Shopping)* (pending).
[40] For an analysis of the provision, see Carugati (2022a).
[41] For an analysis of the competion in answer engines, see Carugati (2023).

FTC *Facebook* case, in which the FTC accused Facebook of infringing US antitrust law by refusing access to key APIs to developers designing competing functionalities[42].

In Europe, such a case would breach EU competition law only if the practice meets four cumulative conditions: the input is indispensable in a way that there is no actual or potential substitute for this input for technical, legal or economic reasons; the refusal is likely to eliminate all competition in another market; the refusal is not objectively justified[43]; and, if Intellectual Property Rights (IPRs) protect the input, the refusal prevents the emergence of a new product for which there is potential consumer demand (Graef *et al*, 2019)[44]. In practice, an FM (the input) is unlikely to be indispensable. Developers of AI-powered applications can use other open-source and closed-source FMs to develop their applications.

However, the FM can be indispensable when a dominant firm in a given market requires the use of its own FM to develop AI-powered applications for its dominant product or service. In this case, developers of AI-powered applications must use the FM of the FM developer if they want to develop applications for its dominant product. This would be the case, for instance, if Microsoft requires developers of AI-powered applications to use GPT to develop applications for Microsoft Windows or Microsoft Office. In this scenario, the use of GPT will be indispensable to develop applications for Microsoft Windows or Microsoft Office. In this circumstance, the condition of indispensability would be met. Then, the condition that the refusal is likely to eliminate all competition in another market can hold, as the refusal of access to the FM could prevent all competition in a competing AI-powered application market. Finally, the refusal would unlikely be objectively justified, as the FM developer would allow access to its FM to developers of non-competing functionalities.

Lastly, even if the practice does not meet the conditions for a refusal to deal, it can still breach EU competition law for discriminatory abuse if discrimination is not objectively justified and has the effect of excluding competitors[45].

[42] Federal Trade Commission, 'FTC Sues Facebook for Illegal Monopolization', 9 December 2020, https://www.ftc.gov/news-events/news/press-releases/2020/12/ftc-sues-facebook-illegal-monopolization. The case is still pending before the Court.
[43] C-7/97 *Bronner*, ECLI:EU:C:1998:569, 26 November 1998, para. 41.
[44] C-418/01 *IMS Health*, ECLI:EU:C:2004:257, 29 April 2004, para. 38 and para. 49.
[45] In France, the French competition authority found that Cegedim abused its dominant position through discriminatory access by preventing access to its database to customers using a management software while allowing access to customers using competing software. 8 July 2014: Health / Medical Information Databases, *Autorité de la concurrence*, 10 July 2014, https://www.autoritedelaconcurrence.fr/en/communiques-de-presse/8-july-2014-health-medical-information-databases.

3.1.3 Algorithmic collusion

Firms can use algorithms to implement anticompetitive collusive agreements, such as on price fixing. Collusion could be explicit between two or more firms, or 'hub-and-spoke', involving communication via the same algorithm. In the UK, in an example of explicit collusion, the CMA found that two online sellers of posters and frames used algorithms to implement and monitor their anticompetitive price agreement[46]. In the US, in an example of so-called hub-and-spoke collusion, an ongoing class action alleges that casino hotels in Las Vegas use the same algorithms to set supracompetitive prices[47]. Finally, algorithms can autonomously reach an anticompetitive agreement without explicit communication, known as 'algorithmic tacit collusion'. Although competition authorities have not yet found cases of this, the economic literature has found that machine-learning pricing algorithms can learn through trial and error to achieve a profit-maximising outcome by coordinating with other algorithms (Calvano *et al*, 2020; OECD, 2023a).

Competition laws can tackle such collusion in most cases. However, algorithmic tacit collusion poses several legal challenges related to its legality, detection and liability (OECD, 2017). While these challenges are crucial from a legal standpoint, there are beyond the scope of this paper, which seeks only to establish whether LMs can serve to achieve an anticompetitive agreement.

Moreover, if developers teach their LMs to achieve an anticompetitive outcome and train them on confidential data, LMs could serve as instruments for explicit or hub-and-spoke collusions. Even without these assumptions, LMs can learn to achieve the best strategy and generate output through text or computer code, which can lead to an anticompetitive tacit agreement. LMs have also reportedly reused confidential business information in public conversations, potentially leading to tacit collusions[48]. However, there are no economic studies yet on whether these scenarios can effectively lead to tacit collusion, and whether two LMs can communicate.

[46] Case 50223 *Online Sales of Posters and Frames*, 12 August 2016.
[47] *Richard Gibson, And Heriberto Valiente V.S MGM Resorts International et al*, 2:23-cv-00140, 2023.
[48] Mark Gurman, 'Samsung Bans Staff's AI Use After Spotting ChatGPT Data Leak', *Bloomberg*, 2 May 2023, https://www.bloomberg.com/news/articles/2023-05-02/samsung-bans-chatgpt-and-other-generative-ai-use-by-staff-after-leak#xj4y7vzkg.

3.2 At the data level

FM developers can train their models using public and proprietary data. Competition issues might arise when a dominant firm collects data from websites, known as 'data scraping', and when it refuses to provide access to data relevant to competition, known as 'refusal to access data'.

3.2.1 Data scraping

Data scraping involves collecting data from a source, such as the web. In some countries, including the US and the UK, some content creators, including publishers, have complained about potential IPR infringements for using their copyrighted content without permission[49]. In Europe, data scraping also poses a competition issue to the extent that this practice has a potential anticompetitive effect on content creators, irrespective of whether scraped data is IPR-protected[50]. Indeed, by generating content from content creators, FMs have the potential to reduce traffic to the original content creators because users may not consult the source, potentially excluding them from the market because of lower advertising revenue. Box 2 shows that US and EU competition authorities have already challenged Google over this practice.

[49] Blake Brittain, 'Lawsuits Accuse AI Content Creators of Misusing Copyrighted Work', *Reuters*, 17 January 2023, https://www.reuters.com/legal/transactional/lawsuits-accuse-ai-content-creators- misusing-copyrighted-work-2023-01-17/.

[50] C-457/10 *AstraZeneca AB and AstraZeneca plc v. European Commission*, ECLI: EU: C:2012:770, 6 December 2012, para. 112 and 132.

Box 2: Cases of data scraping against Google

In 2012, Google released vertical search services, including shopping and maps, through its Google general search engine service. Several competing vertical search services alleged that Google scraped their content without their consent, used their content in its own vertical search services and threatened to delist content providers who protested the practice.

In the US, the FTC investigated whether the practice could diminish the incentive for rivals to invest in new and innovative content and whether it could reduce Google's own incentive to innovate by providing its own content[51]. While the FTC had strong concerns that the practice could infringe US antitrust law, Google offered a five-year commitment in 2013 to refrain from scraping data from vertical websites without their consent on an opt-out basis for its own vertical services[52].

In Europe, the European Commission had the same competition concerns in its *Google Search (Shopping)* investigation[53]. However, the Commission left the issue open as it neither sent a statement of objections nor closed it publicly[54].

3.2.2 Refusal to access data

Some FM developers may need access to data relevant to competition to train and run their models. The data owner can accept or refuse access when developers request access to proprietary data.

However, as discussed in section 2.1.2, a dominant firm's refusal to permit access to an input, such as data, breaches competition law under certain conditions on a case-by-case basis. For instance, Microsoft uses its index data with GPT-4 to provide up-to-date generated content in the answer engines Bing Chat and ChatGPT. However, Microsoft reportedly threatened rivals with restriction of access to its index data if competitors did not stop using it to develop their own answer engines because, according to Microsoft, the practice violates the terms and conditions of use of its index[55]. In this case, the first condition, that refusal to grant access is likely to eliminate all competition in another market, is

[51] Federal Trade Commission, 'Statement of the Federal Trade Commission Regarding Google's Search Practices in the Matter of Google Inc., FTC File Number 111-0163', 3 January 2013.
[52] Google, Google Inc., File No. 111-0163.
[53] Joaquín Almunia, 'Policy Statement of VP Almunia on the Google Antitrust Investigation', 21 May 2012, https://ec.europa.eu/commission/presscorner/detail/en/SPEECH_12_372.
[54] European Commission memo of 15 April 2015, 'Antitrust: Commission Sends Statement of Objections to Google on Comparison Shopping Service', https://ec.europa.eu/commission/presscorner/detail/en/MEMO_15_4781.
[55] Leah Nylen and Dina Bass, 'Microsoft Threatens Data Restrictions in Rival AI Search', *Bloomberg*, 25 March 2023, https://www.bloomberg.com/news/articles/2023-03-25/microsoft-threatens-to-restrict-bing-data-from-rival-ai-search-tools#xj4y7vzkg.

unlikely to be met. Indeed, rivals like Google and, formerly, Neeva, compete with Microsoft by providing their own answer engines without relying on Microsoft index data. Therefore, the refusal is unlikely to eliminate all competition in the answer engine market. Thus, it is unlikely that Microsoft's threat constitutes a refusal of access.

3.3 At the computing resources level

Computing hardware level plays a significant downstream role and has various market players. While acknowledging that competition issues due to the provision of rebates by dominant providers might arise to attract customers along the value chain[56], this section focuses on competition issues arising from the provision of cloud services. Most FMs run from the cloud and competition issues are likely to arise when a dominant cloud provider prevents developers of FMs from switching their models, data and applications from one cloud provider to another, known as 'barriers to switching'. In addition, some FM developers are present in multiple markets, including cloud and software. They might have the ability and incentive to create an ecosystem around their services and lock developers of applications into this ecosystem, known as the 'ecosystem lock-in'.

3.3.1 Barriers to switching

Cloud customers use cloud services by scaling their needs up or down. Cloud providers attract customers by offering them cloud credit through free trials and support programmes. In practice, most cloud credits represent a monetary sum to be spent, aiming at retaining customers. The practice by a dominant cloud provider might have procompetitive effects, such as lower price, and anticompetitive effects, such as customer lock-in, which require a case-by-case analysis, depending on the details of the cloud credit[57].

In other words, cloud providers make it attractive for customers to access and use their services. However, they may make it difficult for customers to exit and switch from one provider to another, because of commercial or technical barriers.

At the commercial level, some cloud providers impose data transfer fees, known as 'egress fees'. This practice might have anticompetitive effects as it might lock cloud customers into their services, thus exploiting them and excluding rival cloud providers. However, in competitive law, the practice's legality

[56] See for instance in Europe, the Intel case about rebates. Comp/37.990 *Intel*, 13 May 2009.
[57] Avis n° 23-A-05 du 20 Avril 2023 Concernant le Projet de Loi Visant à Sécuriser et Réguler l'espace Numérique, *Autorité de la concurrence*, 20 April 2023.

depends on the details of each egress fee and objective justifications provided by the dominant cloud provider[58].

At the technical level, cloud customers might have difficulties transferring their data, known as 'data portability', and communicating with other cloud providers, known as 'interoperability'[59]. The assessment of portability and interoperability issues is beyond the scope of this paper, as there are cases specific to the cloud sector. However, it is worth noting that some EU and national legislation, such as the European Data Act or the French Law to Secure and Regulate the Digital Space, are still in the legislative process to target these issues by imposing obligations on cloud providers[60]. In relation to FMs, there are no studies yet on whether portability issues exist when transferring an FM and the applications developed on top of it from one cloud provider to another.

3.1.2 Ecosystem lock-in

Some FM developers, including Microsoft and Google, also provide associated services such as cloud and software services. As FM customers need access to cloud services to train and run their models and create software applications, developers offering FMs, cloud and software services have a strong incentive to leverage their overlapping customer bases and complementary services to create an ecosystem around their services and lock-in customers (Jacobides and Lianos, 2021). For instance, in an extreme hypothetical scenario, an FM developer could require or incentivise customers to use its cloud service to develop applications for its software applications or OS[61]. If the firm has a dominant position in one of these markets (eg the OS market), it will enable it to grow its market share in the associated non-dominated markets (eg the cloud market), as customers would have to use the latter associated services to access the dominant services. This practice might have anticompetitive effects, as the lock-in of customers might exploit them and exclude rivals. At the same time, the practice might have procompetitive effects, as it might create efficiencies in the form of greater compatibility between LMs and applications or operating systems. This practice may also be justified objectively to recover the investment cost in FMs, or for security reasons. In addition, the DMA counters such strategies,

[58] Ibid.
[59] Ibid.
[60] Ibid.
[61] In Germany, the German competition authority, the Bundeskartellamt, is investigating this issue as part of its investigation into whether to designate Microsoft as a firm subject to the German digital competition legislation of Section 19a German Competition Act. See 'Examination of Microsoft's significance for competition across markets', *Bundeskartellamt*, 28 March 2023, https://www.bundeskartellamt.de/SharedDocs/Meldung/EN/Pressemitteilungen/2023/28_03_2023_Microsoft.html.

because it requires gatekeepers to allow third-party hardware or service providers to interoperate with their OS for free (Article 6(7) DMA).

4 Policy considerations

4.1 Assessment

At the time of writing, FMs are still in the development phase, with various companies at several levels of the value chain competing to attract customers and supply the demand for FMs from developers of applications and end-users. This occurs rapidly when FM developers innovate frequently with new models and business offerings.

FMs can potentially disrupt several markets in the short term, including entrenched markets such as search engines displaced by answer engines (Carugati, 2023). However, the long-term impacts are still unknown but are likely to be disruptive, as FMs will force firms of all sizes in all sectors to rethink how they do business with their customers. Against this background, competition authorities should focus on short-term risks.

4.2 Recommendations

In this early development phase, competition authorities should ensure that users of FMs have a choice in terms of which FMs to use, and that dominant firms do not indulge in practices, such as those mentioned in section 2, which would deter entry or exclude rivals.

In addition, competition authorities should ensure that the competitive process between open- and closed-source models works. Because open-source models could be misused in a manner that harms users, developers of open-source models and public authorities are likely to intervene in the market by imposing restrictions on them.

Regarding the imposition of restrictions by open-source developers, competition authorities should ensure that the restrictions do not harm competition. For instance, in the EU *Google Android* case, the Commission found in 2018 that Google prevented manufacturers of smartphones from running a modified version not approved by Google of its open-source OS Google Android, known as 'Android forks', if they wanted to preinstall some flagship Google apps on their smartphones[62]. This restriction prevented the development of Android forks and its ability to exert competitive pressure on Google.

[62] AT.40099 *Google Android*, 18 July 2018.

The EU General Court confirmed the Commission's findings, but the case is still pending before the EU Court of Justice[63].

In terms of the imposition of restrictions by public authorities, competition authorities should work with policymakers during the legislative process, and with competent authorities during the enforcement phase, to ensure that they do not impose undue restrictions that would deter the development of open-source models in a way that would favour the use of closed-source models.

4.3 Future research questions

FMs raise competition issues that could have a wide impact, but have been few studies on these issues. Competition authorities and researchers should focus on three fundamental questions:

First, the data-scraping issue might reduce traffic to the original content providers, resulting in a loss of advertising revenue and the exclusion of some content providers. Thus, it might profoundly impact content providers and the advertising industry, with unknown consequences for the diversity of content and concentration in the digital advertising sector[64]. Competition authorities should monitor and eventually launch a market investigation into the impact of FMs on content providers and the advertising industry.

Second, ecosystem issues are complex and require further investigation. Several large digital ecosystems, including Microsoft, Google and Amazon, use FMs and generative AI as business drivers for their services[65]. While researchers and competition authorities are already working on digital ecosystems, they should assess the role of FMs in digital ecosystems (see for example ACCC, 2023).

Finally, several current and forthcoming laws, such as the European Artificial Intelligence Act (AI Act) on AI, will likely impact FM development. In addition, as FMs raise similar issues worldwide, their impact will be global. Competition authorities should then collaborate with competent authorities across regulatory fields and countries[66].

[63] T-604/18 *Google and Alphabet v Commission (Google Android)*, ECLI:EU:T:2022:541, 14 September 2022. C-738/22 P *Google and Alphabet v Commission* (pending).
[64] Several competition authorities worldwide did a market investigation into the digital advertising industry. See, for example, for the UK, CMA (2020).
[65] Tom Dotan and Miles Kruppa, 'Microsoft, Google, Amazon Look to Generative AI to Lift Cloud Businesses', *The Wall Street Journal*, 27 March 2023, https://www.wsj.com/articles/microsoft-google-amazon-look-to-generative-ai-to-lift-cloud-businesses-7159a43f.
[66] For an extensive discussion in the context of the DMA, see Carugati (2022b).

References

ACCC (2023) *Digital Platform Services Inquiry – September 2023 Report on the expanding ecosystems of digital platform service providers*, Issues Paper, Australian Competition and Consumer Commission

Autorité de la concurrence (2018) 'Opinion no. 18-A-03 of 6 March 2018 on data processing in the online advertising sector'

Calvano, E., G. Calzolari, V. Denicolò and S. Pastorello (2020) 'Artificial intelligence, algorithmic pricing and collusion', *American Economic Review* 110(10)

Carugati, C. (2022a) 'How to Implement the Self-Preferencing Ban in the European Union's Digital Markets Act', *Policy Contribution* 22/2022, Bruegel

Carugati (2022b) 'A Practical Arrangement for Cooperation Between Digital Economy Regulators', *Bruegel Blog*, 13 June

Carugati, C. (2023) 'Antitrust Issues Raised by Answer Engines', *Working Paper* 07/2023, Bruegel

CMA (2020) *Online platforms and digital advertising*, Market study final report, Competition and Markets Authority, available at https://assets.publishing.service.gov.uk/media/5fa557668fa8f5788db46efc/Final_report_Digital_ALT_TEXT.pdf

Crémer, J., Y.-A. De Montjoye and H. Schweitzer (2019) *Competition policy for the digital era*, European Commission, Directorate-General for Competition, available at https://data.europa.eu/doi/10.2763/407537

Eloundou, T., S. Manning, P. Mishkin and D. Rock (2023) 'GPTs are GPTs: An Early Look at the Labor Market Impact Potential of Large Language Models', mimeo, available at http://arxiv.org/abs/2303.10130

FTC (2023) 'Generative AI Raises Competition Concerns', *Technology Blog*, 29 June, Federal Trade Commission, available at https://www.ftc.gov/policy/advocacy-research/tech-at-ftc/2023/06/generative-ai-raises-competition-concerns

Geng, X., A. Gudiband, H. Liu, E. Wallace, P. Abbeel, S. Levine and D. Song (2023) 'Koala: A Dialogue Model for Academic Research', Berkeley Artificial Intelligence Research, 3 April, available at https://bair.berkeley.edu/blog/2023/04/03/koala/

Graef, I., T. Tombal and A. de Streel (2019) 'Limits and Enablers of Data Sharing: An Analytical Framework for EU Competition, Data Protection and Consumer Law', TILEC *Discussion Paper* No. DP 2019-024, Tilburg Law and Economics Center, available at https://doi.org/10.2139/ssrn.3494212

Gregory, R.W., O. Henfridsson, E. Kaganer and H. Kyriakou (2021) 'The Role of Artificial Intelligence and Data Network Effect for Creating User Value', *Academy of Management Review*, 46(3): 534-51, available at https://doi.org/10.5465/amr.2019.0178

Hatzius, J., J. Briggs, D. Kodnani and G. Pierdomenico (2023) 'Global Economics Analyst The Potentially Large Effects of Artificial Intelligence on Economic Growth', *Global Economics Analyst*, 26 March, Goldman Sachs Economics Research

Hoffmann, J., S. Borgeaud, A. Mensch, E. Buchatskaya, T. Cai, E. Rutherford … L. Sifre (2022) 'Training Compute-Optimal Large Language Models', mimeo, available at http://arxiv.org/abs/2203.15556

Hu, E.J., Y. Shen, P. Wallis, Z. Allen-Zhu, Y. Li, S. Wang, L. Wang and W. Chen (2021) 'LoRA: Low-Rank Adaptation of Large Language Models', mimeo, available at http://arxiv.org/abs/2106.09685

Jacobides, M.G. and I. Lianos (2021) 'Ecosystems and Competition Law in Theory and Practice', *Industrial and Corporate Change* 30(5): 1199–1229, available at https://doi.org/10.1093/icc/dtab061

Kaplan, J., S. McCandlish, T. Henighan, T. Brown, B. Chess, R. Child … D. Amodei (2020) 'Scaling Laws for Neural Language Models', mimeo, available at https://arxiv.org/abs/2001.08361

Kojima, T., S.S. Gu, M. Reid, Y. Matsuo and Y. Iwasawa (2023) 'Large Language Models are Zero-Shot Reasoners', mimeo, available at http://arxiv.org/abs/2205.11916

Lippmann, R. P. (1987) 'An Introduction'to Computing with Neural Nets', *IEEE ASSP Magazine* 4(2)

OECD (2017) *Algorithms and collusion Competition policy in the digital age*, Organisation for Economic Co-operation and Development

OECD (2020) *Abuse of Dominance in Digital Markets*, Organisation for Economic Co-operation and Development

OECD (2023a) *Algorithmic Competition*, OECD Competition Policy Roundtable Background Note, Organisation for Economic Co-operation and Development

OECD (2023b) 'AI Language Models: Technological, Socio-Economic and Policy Considerations', *OECD Digital Economy Papers* 352, Organisation for Economic Co-operation and Development available at https://doi.org/10.1787/13d38f92-en

Rae, J.W., S. Borgeaud, T. Cai, K. Millican, J. Hoffmann, F. Song … G. Irving (2021) 'Scaling Language Models: Methods, Analysis & Insights from Training Gopher', mimeo, available at http://arxiv.org/abs/2112.11446

Schick, T. and H. Schu ̈tze (2021) 'It's Not Just Size That Matters: Small Language Models Are Also Few-Shot Learners', Proceedings of the 2021 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, available at https://doi.org/10.18653/v1/2021.naacl-main.185

Stucke, M.E. and A.P. Grunes (2016) *Big Data and Competition Policy*, Oxford University Press

Vaswani, A., N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. Gomez, L. Kaiser and I. Polosukhin (2017) 'Attention Is All You Need', mimeo, available at https://arxiv.org/abs/1706.03762

Bruegel, Rue de la Charité 33, B-1210 Brussels
(+32) 2 227 4210
info@bruegel.org
www.bruegel.org