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

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

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

player. Moreover, as the demand for LMs increases, new entrants and LMs will likely appear frequently.

Developing LMs

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. A dominant firm in a given market can also restrict or refuse access to its LM, known as "leveraging". 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.

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 browser384

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-

3.1.3

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,

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 inv

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

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 incen(o)-5 (v (nd 4.-2 (e)-5J0 Tc 0 Tw 3.9 (6 Td( )Tj-

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].

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-

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', *P RDj/T2_1 1f0.005w0.02 0 Tw [())3 ( )]TJEMC1P AM[22,pan <0 Tw 4 0 Td()Tj(,)Tj/T2_1 1f0.005w 2.38 0.37 0 i,i hric*

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

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',

© Bruegel 202All rights reserved. Short sections, not to exceed two paragraphs, may be quoted in the original language without explicit permission provided that the source is acknowledged. Opinions expres in this publication are those of the author(s) alone.