Introduction

We, a group of civil society organisations active in the European Union, welcome the opportunity to contribute to the European Commission's consultation on competition and generative AI. The accelerating development and deployment of artificial intelligence (AI) is one of the greatest tests the next European Commission will face. Whether Europe's regulators meet the moment – or fail to – will shape our society and economy for many years.

Any sensible regulatory action in generative AI must begin with a hard truth: power in the sector is already highly concentrated. Decades of underenforcement of competition rules in digital markets have consolidated resources and capabilities in the hands of a few dominant firms, with high costs for our economies and societies.

The challenge thus resembles the history of digital markets over the past two decades, when regulators' hesitation allowed the tech behemoths to rise. But the risk of inaction today is greater still: because we begin with such a concentrated sector, the incumbents have swiftly moved to entrench their power and usurp potential competitors – including promising startups in Europe. This includes leveraging their existing positions of dominance – including search and operating systems – to capture new markets and technologies, and signing partnerships with smaller challengers to neutralise the competitive threat.

Current market structures heavily favour today's tech giants, which enjoy a degree of scale and vertical integration that is enabling them to secure a stranglehold over AI development and deployment. Their aim is to control the direction of innovation and ensure that AI does not threaten their dominance. This combination of extreme scale and narrow commercial self-interest is fuelling a race to the bottom for which society is paying the cost. Large-scale AI models are already being deployed in ways that violate fundamental rights and undermine the public interest, from violating privacy rights to creating and disseminating false and extreme material.

The development and rollout of generative AI systems also appears to be happening much faster than with previous technologies. This leaves little time to ruminate on the finer points of regulatory intervention; ex-post action will come too late and is unlikely to be effective. A change of stance is needed. Regulators, including in Europe, need to rise to the occasion swiftly with an approach that takes full account of the power conferred by market

concentration and digital ecosystems – not just in large-scale AI models but across the wider technology stack.

Early and robust intervention by competition authorities will be a key part of any successful strategy. Firm enforcement of competition rules will be essential in ensuring that generative AI, and AI in general, are deployed in ways that work for people and democracy, not against them. Failure to do so will result in dangerous outcomes for both, including untamed and unaccountable corporate power, the absence of human rights and public interest considerations from technological development, and the exclusion of European firms from the value chain in a concentrated market.

In the rest of this submission, we seek to draw attention to a number of specific developments and issues which we believe the European Commission should focus its efforts on. We look forward to dialogue with DG COMP on this topic in the months to come.

I: Bottlenecks and barriers to entry in the AI stack

There are bottlenecks and barriers to entry in every part of the AI stack. The firms driving large-scale AI development have infrastructural power, controlling access to the critical resources needed to build AI at scale – especially large datasets, cutting-edge semiconductors and cloud computing platforms – as well as the marketplaces, ecosystems and platforms through which generative AI technologies are distributed. In other words, they own the servers; they hold the data; they run the stores. Taken together, this gives incumbent firms powerful gatekeeper control over the trajectory of AI.

Any analysis of competition in AI must therefore look upstream. While much attention is directed toward the model layer, it is crucial to consider the broader ecosystem of dependencies that underpin it.

The main bottlenecks and barriers to entry in artificial intelligence can be conceptualised in terms of key inputs: **data, computing power, and talent.** These barriers can be identified in both the development and deployment of AI models.

Development of generative AI models

A whole host of competition issues arise when it comes to the training and development of AI models.

At the **data** level, while the core training material for large models often consists of publicly available datasets, firms with access to proprietary and curated datasets enjoy a competitive edge. Some datasets hold exceptional value due to their scarcity, specialised nature and/or exclusivity. For instance, generative AI models optimised for use in the financial sector rely on data available to only a select few companies.[1] Moreover, dominant firms can reinforce this data advantage through acquisitions (of data or of companies that possess data) and exclusive licensing agreements with data holders.[2]

[1] https://arxiv.org/abs/2303.17564
[2] https://openai.com/blog/axel-springer-partnership

Beyond access, the ability and capability to process **data** – sometimes referred to as "data work"[3] – represents a potential competitive barrier in terms of both the technical expertise required and the financial cost such work entails. Additionally, the power to address, deflect or evade legal and compliance requirements resulting from the unauthorised or illegitimate use of copyrighted material can serve as a novel source of competitive advantage. Notably, models that trained on copyrighted material prior to increased regulatory scrutiny may benefit from a kind of "first mover advantage". Finally, the importance of quality training data for large models[4] makes data labelling and curation companies central nodes and potential bottlenecks in the AI ecosystem.

The **computing power** needed to train large-scale models is dominated by the largest private cloud computing providers. As such, concentration in computational infrastructure functions as a gravitational field accelerating the centralisation of the generative AI ecosystem around these key providers, and indeed appears to be a central part of their commercial strategies[5] [6] [7] Recently, several dominant cloud providers have entered the chip market[8] and vice versa,[9] leading to further integration across the AI stack.

The **talent** needed to train, develop and build generative AI models is scarce. While high-level model architectures are often available in the public domain,[10] [11] the technical expertise needed to train and adapt these models is in short supply. This scarcity leads to high hiring costs, which solidifies the position of dominant companies with the most financial firepower.

Furthermore, training and validating a generative AI model or tool is an exceptionally labour intensive task that typically requires humans to manually review the material ingested or produced by a generative AI model. Given the sheer amount of data that needs to be reviewed, this can present a substantial cost that acts as yet another barrier for challenger firms while benefiting dominant players with deeper pockets.[12]

An additional relevant factor is that a large amount of such work is outsourced to lower-cost locations via online platforms or third-party contracting services. The largest online platform for this type of digital piece work (Mechanical Turk) is an Amazon owned service, with other major tech firms being major customers of outsourcing firms. This buyer power could enable large firms to negotiate lower prices (compared to smaller firms) for manual review work, or to establish exclusivity agreements with leading outsourcing firms.

---

[3] https://dl.acm.org/doi/10.1145/3411764.3445518
[4] https://arxiv.org/abs/2311.16867
[5] https://www.technologyreview.com/2023/12/05/1084393/make-no-mistake-ai-is-owned-by-big-tech
[6] https://www.cnbc.com/2024/01/30/microsoft-msft-q2-earnings-report-2024.html
[7] www.theinformation.com/articles/microsofts-ai-coding-product-becomes-weapon-in-battle-with-aws
[8] https://www.bloomberg.com/news/articles/2023-11-28/amazon-updates-homegrown-chips-even-as-it-tightens-nvidia-ties
[9] https://www.theinformation.com/articles/nvidia-muscles-into-cloud-services-rankling-aws?rc=qudiig
[10] https://arxiv.org/abs/1706.03762
[11] https://arxiv.org/pdf/2307.09288.pdf
[12] https://www.pymnts.com/cpi_posts/generating-concerns-exploring-antitrust-issues-in-the-generative-ai-sector

Distribution of generative AI models

While existing market concentration has the largest effect on the development of AI models and tools, there are also clear implications for the distribution of models.

At the level of **data**, firms with the most popular models and applications and the widest distribution channels also have access to the most information on how people use, abuse and interact with their products. This data can subsequently be used to improve these tools and distribution channels, in turn attracting more users and generating additional data which can in turn be used for retraining. These self-reinforcing data dynamics are similar to those seen in the platform economy, and can make it extremely difficult if not impossible for challengers to break through.

At the level of **compute,** while much attention has been paid to compute costs linked to the training of generative AI models, hosting these models incurs substantial costs as well. Hosting an AI model at reasonable scale is prohibitively expensive[13], requiring either a large capital buffer to sustain the burn rate or a preferential agreement with a cloud service provider (CSP). This further strengthens the gravitational pull of the leading compute providers over the nascent AI ecosystem. This means that even where governments create public computing infrastructure for training AI models – such as recent EU initiatives – this is unlikely to effectively promote competition unless alternative hosting infrastructures are established.

This problem is exacerbated by the emphasis large firms are placing on model marketplaces and "platformization" of the generative AI ecosystem. Microsoft Azure AI, Amazon Bedrock, Google Cloud and NVIDIA AI combine access to both computing capacity and various proprietary and open source models on the same platform, creating solutions for AI customers that tie them to existing cloud and enterprise software ecosystems. This increases the likelihood of customer lock-in and reinforces the role of CSPs as central nodes in the AI ecosystem, while strengthening their ability to extract economic rents at several critical points of the AI tech stack. Again, these dynamics mirror those found in the platform economy, and cut both ways. Not only do large incumbents have a distributional advantage, but challenger firms without access to these distribution channels will struggle to reach customers and grow.

Partnerships between dominant players and startups

As a consequence of this highly concentrated tech stack, promising startups are increasingly turning to incumbents for capital, data, computing power, and talent. In order to gain access to the resources they need to train and run their AI models, leading AI startups – including OpenAI, Mistral AI and Anthropic – have agreed to lopsided partnerships in which large tech firms provide financial investment and computing power in exchange for privileged or exclusive access to the startup's technology and explicit or implicit influence over its technical and corporate decision-making.[14]

---

[13] https://www.semianalysis.com/p/the-inference-cost-of-search-disruption
[14] https://www.techpolicy.press/monopoly-power-is-the-elephant-in-the-room-in-the-ai-debate/

While it is yet to be seen whether these partnerships meet the legal thresholds to be considered mergers, they appear likely to produce many of the same anti-competitive effects that such deals have produced in the past.[15] Generally speaking, one would expect these partnerships to blunt the partnered firms' incentives to compete against each other, given the financial and technological interdependencies involved. This applies especially to the smaller "partner", which risks losing access to critical financial and infrastructural support if it competes too aggressively against its senior partner. But it also applies to the larger firm, which will not want to undermine its investment by competing too aggressively against its investee.

Such partnerships give dominant firms an advantage over current rivals in the AI race, while also neutralising potential future challengers. The result is likely to be less choice and higher prices for consumers and businesses, less innovation, and greater competitive harm as market consolidation allows a few large platforms to exploit their dominance over those that depend on their services.

II: Actual and potential harms from market concentration

As the discussion above shows, today's AI ecosystem is already highly concentrated, and likely to become more so over time. If the evolution of digital markets over the past two decades is a guide, without timely and comprehensive competition enforcement, this market concentration is likely to result in a wide range of monopolistic or oligopolistic harms and anti-competitive practices.

Below we explore specific harmful practices that may emerge as the AI ecosystem matures over the coming months and years. While some of these harms remain hypothetical, they draw from an extensive track record of similar behaviour in digital markets, often by the same firms now set to dominate AI.

Self-preferencing, tying and leveraging

Today's dominant tech companies enjoy market power across a wide range of products, services and technologies, including online advertising, browsers, operating systems, app stores, cloud computing, productivity software and more. This "conglomerate"[16] or "ecosystem"[17] power enables these firms to exploit their dominance in one market to expand or entrench their hold over another, typically through a number of "leveraging" and "self-preferencing" behaviours. Previous (and in some cases ongoing) examples of such conduct include Microsoft tying its browser and media player to its operating system[18], and Google steering businesses and users towards its proprietary digital advertising and shopping services.[19]

---

[15] https://insights.som.yale.edu/insights/wave-of-acquisitions-may-have-shielded-big-tech-from-competition

[16] https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3350512

[17] https://www.promarket.org/2023/12/14/furthering-ecosystem-analysis-in-antitrust/

[18] https://ec.europa.eu/commission/presscorner/detail/en/IP_09_1941

[19] https://ec.europa.eu/commission/presscorner/detail/es/MEMO_17_1785

Generative AI provides new opportunities for digital gatekeepers, and the smaller companies they partner with, to replicate and build on such practices. Vertical integration across the AI stack – including semiconductors, cloud computing and foundation models – and integration into adjacent services – including search engines, browsers and app stores – will give dominant firms many opportunities to unfairly promote their own AI products over those offered by rivals.

Most obviously, cloud providers that also offer their own AI technologies – such as Google and Microsoft – could seek to steer customers towards using these in a variety of ways, including by preferentially ranking their own services or bundling services together in a way that is difficult for third-parties to compete with. A case in point is Microsoft, which is reportedly bundling certain security features for GitHub CoPilot with subscription to Azure.[20] Similarly, foundation model providers that also build applications – such as OpenAI – could give preferential treatment to their own applications compared to third-party offerings, for instance by giving them access to better model capabilities. Similar considerations could be applied to companies that develop both advanced semiconductors and AI models trained on them.

Exploitation of dominance

Firms in dominant positions at one or multiple levels of the AI stack will be in a position to exploit and abuse downstream actors that depend on their services and infrastructure. Such practices have been common on digital platforms in the past, particularly on platforms that serve as marketplaces between businesses and end users. For example, Apple and Google have attracted significant scrutiny for their conduct towards developers dependent on their app stores[21], as has Amazon for its treatment of third-party sellers on its marketplace.[22] This is reflected in a number of major antitrust and regulatory developments in Europe, including the recently enacted Digital Markets Act.

In fact, these dominant players are already profiting from the boom in AI applications by earning fees for in-app purchases made using their platforms. As the market develops, new opportunities may arise for firms which gain a position of power to exploit it. For example, OpenAI has launched an app store allowing developers to create customised versions of ChatGPT,[23] and other foundation model providers may follow suit. While this could provide much-needed competition to existing app store operators, it also could also place OpenAI or another foundation model provider in a similar position to exploit its power over application developers and other customers. This illustrates the need for basic rules of the game which prevent dominant players from exploiting dependencies and extracting rents.

The centrality of large cloud computing platforms in training and hosting AI models also creates opportunities for abuse of dominance. Already, today's dominant cloud platforms are

---

20 https://www.theinformation.com/articles/microsofts-ai-coding-product-becomes-weapon-in-battle-with-aws?utm_source=ti_app

21 https://www.gov.uk/cma-cases/mobile-ecosystems-market-study

22 https://ec.europa.eu/commission/presscorner/detail/en/ip_22_7777

23 https://openai.com/blog/introducing-the-gpt-store

facing intense antitrust scrutiny for anti-competitive practices, including imposing excessive switching fees and unfair licensing conditions on customers, and unnecessarily limiting interoperability.[24] As the AI ecosystem develops further, dominant cloud providers will have new opportunities to exploit their infrastructural position and capture a growing share of the value generated on their platforms.

While excessive and inconsistent fees are one manifestation of abusive dominance, there are other harmful practices to be on the alert for. Dominant providers of app stores, foundation models and cloud computing could also exploit their centrality to impose unfair terms and conditions on customers, force them to hand over valuable data, or suspend them without warning, mirroring many of the practices already seen in today's concentrated digital markets.

Limiting interoperability and restricting access to inputs

The infrastructural, gatekeeper role that a few dominant corporations look set to play in the emerging AI ecosystem will also give them the power and incentive to limit access to critical inputs and functionalities, including computing power, data, and technical gateways such as application programming interfaces (APIs). These behaviours are closely related to the self-preferencing and exploitative practices discussed above, all of which result from a single firm controlling access to multiple and overlapping platforms and ecosystems which businesses and consumers rely on.

Dominant cloud computing providers have the ability and incentive to degrade or shut off access in order to hurt potential rivals or punish customers that refuse to comply with unfair or exploitative terms and conditions. For example, a cloud provider that also offers its own AI solutions could degrade an AI developer's access to its high-performance computing platform if it felt that customer posed a competitive threat. Similar behaviour by dominant players has already been seen in other parts of the sector, such as Microsoft reportedly threatening data restrictions or increased prices for customers that develop competing generative AI search solutions.[25] Alternatively, if a dominant cloud provider wanted to prevent rivals from competing with it in hosting AI services, it could limit the ability of AI service providers to interoperate or carry over data to those rivals.

Similar risks arise in relation to foundation models themselves, should a small number of models end up playing a centralised, infrastructural role in the AI ecosystem. Mirroring the discussion of cloud computing above, vertical integration in foundation models and applications built on top of those models would give companies providing both the ability and incentive to degrade access for rival app developers. This might entail limiting developers to basic API queries, instead of giving them access to core model capabilities.

Last but not least, the role that wider digital ecosystems look set to play in the commercialisation of AI provides further opportunities for dominant platforms to limit access and interoperability. There are many potential ways such conduct could manifest, from a

---

[24] https://www.ofcom.org.uk/consultations-and-statements/category-2/cloud-services-market-study t

[25] https://www.bloomberg.com/news/articles/2023-03-25/microsoft-threatens-to-restrict-bing-data-from-rival-ai-search-tools

search engine provider (that also offers AI services) preventing potential challengers from training their AI tools on its search index data, to an operating system provider (that again also offers its own AI services) limiting the ability of third-party foundation models to access core functionalities.

III: Open versus closed systems

Several distinctions are necessary for a grounded discussion on “open” versus “closed” AI systems and the competition implications of making critical AI components widely available.

First, there is no binary distinction between open and closed models. Instead, there are varying degrees of openness, from fully closed models to hosted access to fully open.[26] It is therefore important to consider the full spectrum; indeed, not everybody claiming openness adopts complete openness. For instance, releasing only the trained model weights allows others to run, integrate, or adapt the model. It does not, however, enable reproducibility or deeper scrutiny of the artefact without, for example, access to training data, detailed documentation, or the source code used for the training process. Moreover, openness does not directly translate to specific policy goals or objectives – it is an underspecified category and a meaningful policy discussion requires greater precision about the actors and functionalities involved.

There is currently little visibility into the uptake of “open source” versus closed AI models and other components, as well as into their integration into consumer products and services. It is therefore hard to gauge how open components, particularly openly released models (i.e., model weights) compete against proprietary offerings in the market. However, the limited evidence we do have points towards the most advanced open models lagging behind their proprietary counterparts.[27]

In the long term, whether this gap will widen or close is contingent on a variety of factors, including the regulatory and competitive landscape as well as the viability of potential business models around open and closed offerings. Moreover, it should be noted that anti-competitive behaviour by large companies (including companies that claim to be open source) also presents significant threats to open source development through a number of mechanisms, including resource constraints, lobbying, capture of intellectual property and ecosystem capture.

Furthermore, there is a threshold of scale beyond which there are challenges in combining an open source approach with the business models that can finance the profitable running of the largest models. For instance, while Mistral provides its smaller models on an open source basis, the most capable Mistral “Large” model is not open sourced but only available through an API on the Microsoft Azure platform or Mistral’s own “Le Platforme”.[28]

While more openness in AI will not be sufficient in creating a competitive AI ecosystem, it can play an important function by providing building blocks across the AI stack for everyone

[26] https://arxiv.org/abs/2302.04844
[27] https://huggingface.co/spaces/lmsys/chatbot-arena-leaderboard
[28] https://mistral.ai/news/mistral-large/

to use, allowing for more experimentation, transparency, and ultimately innovation. As Kapoor and co-authors outline, it does so in three main ways: 1) through broader access, meaning increased diffusion of technological innovation; 2) greater customizability, meaning that open models are easier to customise and adapt to a variety of different use cases; and 3) local inference, meaning that open models make it easier to perform inference — e.g., for a model to answer a query — locally on a device without requiring data to be shared with a third party.[29]

An analysis limited to discrete markets, such as foundation models, would miss the point. The only way for open source models to compete against proprietary models is for there to be an ecosystem that enables this, which in turn will entail tackling concentration across the tech stack.

IV: Suggested changes in competition policy and rules

As argued throughout this submission, market concentration, barriers to entry, anti-competitive partnerships, and a range of existing and potential anti-competitive practices threaten to undermine innovation, openness and fairness in AI. The European Commission should use its existing competition powers as aggressively as necessary to tackle these problems, including the EU Merger Regulation, Articles 101 and 102 of the Treaty on the Functioning of the European Union (TFEU), and the Digital Markets Act. However, we also believe that responding to the emerging challenges posed by AI will require both adapting these existing tools and adopting new ones.

A new market investigation tool

Currently, the Commision is not able to take enforcement actions that target wider sectors, as opposed to investigating individual firms under Article 102 or colluding firms under Article 101 TFEU. We believe that a broader market investigation tool is needed, modelled on similar regimes elsewhere – including in the UK and Germany – as well as the “New Competition Tool” previously considered by the Commision.

These market investigation powers would build on the Commission’s existing sector inquiries and allow it to take a more in-depth and holistic view of important markets, including those relevant to AI, and to propose carefully crafted remedies in response to competition concerns.[30] This would ensure that the Commission was able to take a more expansive approach to addressing market concentration beyond the narrow confines of Articles 101 and 102 and the Sector Inquiry regime.

Upgrading the Digital Markets Act

The recently adopted Digital Markets Act promises to be a powerful tool for promoting fairness and contestability in digital markets. Unfortunately, in its current form, it does not appear suited to the task of achieving these objectives in relation to AI. Fortunately, a small number of changes would ensure that it is.

---

[29] https://crfm.stanford.edu/open-fms/paper.pdf

[30] https://op.europa.eu/en/publication-detail/-/publication/1851d6bb-14d8-11eb-b57e-01aa75ed71a1#

The first major gap is the lack of dominant cloud providers – such as Amazon Web Services and Microsoft Azure – among the core platform services initially designated in September 2023, despite cloud computing being listed as a "core platform service" under Article 2 of the DMA. This currently prevents the DMA from being used to prevent dominant cloud providers from exploiting their centrality in the tech stack to strengthen their market power in AI, for instance by forcing or nudging customers to use their AI models and services. This gap could be quickly rectified by launching a market investigation into specific core platform services under Article 17 of the DMA.

The second major gap is the absence of AI foundation models from the DMA's list of core platform services, which restricts the Commission's ability to tackle abuses inflicted on business and end users of such models (such as restricting model interoperability for third-party developers). Under Article 19 of the DMA, the Commission has the power to launch a separate market investigation examining whether new services should be added to the list of core platform services. We urge the Commission to make use of this power as soon as possible, given the already evident competition challenges in AI, and the fact that an investigation – and any subsequent legislation – will take time to achieve outcomes.

Investigating anti-competitive partnerships

Given the role of partnerships between dominant tech firms and startups in accelerating concentration in AI, we believe there is an urgent need to consider reforms to the EU's merger control regime, including both the EU Merger Regulation (EUMR) itself and relevant accompanying notices and guidelines.

Such reform is needed to facilitate investigations of partnerships and minority investments that potentially fall outside of the scope of the EUMR, but which can have a significant impact on competition in the market. Enforcement of the EU Merger Regulation is based on the "decisive influence" test, which currently sets a high bar for the existence of control by one firm over another, typically requiring the acquisition of majority ownership. However, acquisition of a non-controlling minority-shareholding can also be anti-competitive, as the Commission has itself acknowledged.[31]

Many countries, including EU Member States such as Austria and Germany and third countries such as the UK and US, already have the competence to review such investments. Therefore, the current formulation and/or interpretation of the "decisive influence" test should be adapted to ensure that the EU's merger regime is responsive to new market developments and able to review atypical arrangements that nonetheless present a Significant Impediment to Effective Competition (SIEC).

A legal presumption against acquisitions by dominant firms

We also believe there is a need for a broader structural change in how acquisitions by dominant tech firms are treated by the Commission. The current EU merger control regime

[31] https://eur-lex.europa.eu/legal-content/EN/ALL/?uri=CELEX%3A52014SC0221

requires the Commission to demonstrate that a merger is likely to be harmful to competition before intervening. However, predicting the future impact of a merger on competition, particularly in fast-moving digital markets, is a costly, time-consuming and highly uncertain exercise for regulators. Dominant tech firms are able to exploit these information and resource gaps and asymmetries to prevent acquisitions from being successfully challenged, including by commissioning large quantities of material to support their case.[32]

In part because of this system, many digital mergers now recognised as harmful were approved by the EU and other competition authorities, and there is a risk that this history now repeats itself in relation to AI.[33] To address this problem, we recommend the introduction of a legal presumption against acquisitions by dominant firms. This could be restricted to dominant firms in the tech sector, for example gatekeepers designated under the DMA, or it could apply to dominant firms in any industry. Acquisitions by this subset of firms would be presumed to be illegal under the EU's merger control regime, unless they were able to conclusively demonstrate that they would not create an SIEC.

Signatories to the submission:

Open Markets Institute (Europe)

AI Now Institute

Foxglove

ARTICLE 19

Mozilla Foundation

SOMO

Irish Council for Civil Liberties / Enforce

Rebalance Now

[32] https://academic.oup.com/antitrust/article-abstract/12/1/7/7128249?redirectedFrom=fulltext
[33] https://techcrunch.com/2024/02/01/eu-antitrust-new-world-order/