cerre Centre on Regulation in Europe

WHAT POLICY INTERVENTIONS FOR A COMPETITIVE AI SECTOR?

REPORT

July 2025

Zach Meyers
Marc Bourreau

cerre

Report

What policy interventions for a competitive AI sector?

Zach Meyers
Marc Bourreau

July 2025

As provided for in CERRE's bylaws and procedural rules from its "Transparency & Independence Policy", all CERRE research projects and reports are completed in accordance with the strictest academic independence.

The project, within the framework of which this report has been prepared, received the support and/or input of the following CERRE member organisations: Amazon, Autorité de Régulation des Communications Électroniques, des Postes et de la Distribution de la Presse (Arcep), Belgisch Instituut voor Postdiensten en Telecommunicatie (BIPT – IBPT), and Microsoft. However, they bear no responsibility for the contents of this report. The views expressed in this CERRE report are attributable only to the authors in a personal capacity and not to any institution with which they are associated. In addition, they do not necessarily correspond either to those of CERRE, or of any sponsor or of members of CERRE.

info@cerre.eu – www.cerre.eu

Table of Contents

About CERRE

Providing high quality studies and dissemination activities, the Centre on Regulation in Europe (CERRE) is a not-for-profit think tank. It promotes robust and consistent regulation in Europe's network and digital industry and service sectors as well as in those impacted by the digital and energy transitions. CERRE's members are regulatory authorities and companies operating in these sectors, as well as universities.

CERRE's added value is based on:

- its original, multidisciplinary and cross-sector approach covering a variety of markets, e.g., energy, mobility, sustainability, tech, media, telecom, etc.;
- the widely acknowledged academic credentials and policy experience of its research team and associated staff members;
- its scientific independence and impartiality; and,
- the direct relevance and timeliness of its contributions to the policy and regulatory development process impacting network industry players and the markets for their goods and services.

CERRE's activities include contributions to the development of norms, standards, and policy recommendations related to the regulation of service providers, to the specification of market rules and to improvements in the management of infrastructure in a changing political, economic, technological, and social environment. CERRE's work also aims to clarify the respective roles of market operators, governments, and regulatory authorities, as well as contribute to the enhancement of those organisations' expertise in addressing regulatory issues of relevance to their activities.

About the Authors

Zach Meyers is Director of Research at CERRE, where he has a wide remit, including managing cross-sectoral programmes and projects.

Previously the assistant director of the Centre on European Reform, Zach Meyers has a recognised expertise in economic regulation and network industries such as telecoms, energy, payments, financial services and airports. In addition to advising in the private sector, with more than ten years' experience as a competition and regulatory lawyer, he has consulted to several governments, regulators and multilateral institutions on competition reforms in regulated sectors. He is also a regular contributor to media.

Zach holds a BA, LLB (with First Class Honours) and a Master of Public & International Law from the University of Melbourne.

Marc Bourreau is a CERRE Academic Co-Director and Professor of Economics at Telecom Paris, Institut Polytechnique de Paris, France, where he acts as the Director of the Chair for Innovation & Regulation.

He holds a master's degree in engineering from Telecom Paris and a Ph.D. in Economics from University Paris 2 Panthéon-Assas. He has published in leading journals in Economics. His current research interests concern the economics of digital platforms, the impact of competition and regulation on entry and investment in network industries, and licensing and trading of standard essential patents.

Executive Summary

To tackle the economic problems set out in Mario Draghi's report on EU competitiveness, European businesses must become far more productive and innovative than they are today, including by taking better advantage of new technologies like AI. In particular, by lowering barriers to entry and growth across many markets, AI has the potential to **boost competition and dynamism across the EU economy.**

In an era where Europe is increasingly concerned about excessive unilateral dependence on certain trading partners, AI could also offer new opportunities for Europe to grow the size of its technology sector and secure strategic footholds in the tech value chain. AI is a general-purpose technology which is creating new economic markets.

That means availability of diverse, high-quality and affordable AI systems is critical to Europe's economic growth. F**irms active in the AI sector therefore need sustainable and enduring competitive pressure to continue to invest, innovate, and make their services as widely available as possible**.

We reach a positive assessment of levels of competition in the AI sector. We find, in summary, that:

- **The AI sector is characterised by high levels of static competition** – with low barriers to entry and significant diverse businesses active in the sector. However, when examined through a dynamic lens – by looking at levels of radical innovation, the investments being made by large and small AI firms, including and private equity and venture capital investors, and the degree of uncertainty about which business models, user propositions and technological solutions will ultimately succeed – levels of competition appear even more significant.
- The AI sector relies on the provision of 'accelerated compute'. On a static analysis, a small number of players appear to have a very strong position in providing this service to the AI sector and one firm appears to be dominant in the design of AI accelerator chips. However, a more dynamic analysis suggests that **the largest players in this part of the value chain are subject to emerging competitive challenges** and that competition authorities should allow these challenges to play out. Challenger and incumbent companies are all attracting significant investment.

The concentrated nature of some parts of the AI value chain and the strong position of some vertically integrated companies mean authorities are right to keep a close eye on the sector, despite our findings. There remain certain factors that could constrain competition in the future. These include unilateral conduct (such as firms withholding access to essential inputs like data or accelerator chips); general characteristics of the market (such as difficulties AI developers may have in switching accelerated compute providers due to technical incompatibilities); and the impact of regulations on growth of the sector.

This report examines existing and proposed regulatory measures aimed at addressing these problems. Competition problems identified by authorities to date relate to how the market *could* develop – rather than conduct which has already taken place. Our overall conclusion is that, given the high levels of competitive pressure which currently exist, authorities ought to be cautious about intervening pre-emptively. We do not find compelling reasons for **competition authorities to assume market**

consolidation is likely or that 'winner takes all' outcomes will occur in the AI sector as they have in some digital markets. We recommend that any new regulatory obligations are tightly targeted and are designed to allow the positive and negative economic impacts of particular behaviour to be weighed. In particular, interventions should avoid unnecessarily dampening the current high levels of experimentation, innovation and investment.

We recommend that policy makers:

- **Consider carefully complaints that firms are withholding or imposing unreasonable terms of access to data for AI developers**. Policy makers should, however, ensure rules on mandatory data sharing, like the GDPR, the Data Act and the DMA, are applied in ways which protect firms' incentives to invest in collecting and innovating in the use of data.
- **Consider promoting an industry-led initiative to ensure users of AI services can 'port' their usage history from one model to a competing service**, in order to protect users' choices and help avoid creating 'lock in effects' or 'winner takes all' outcomes.
- **Scrutinise carefully business practices among firms providing chips for accelerated compute**. However, those practices should be reviewed on their merits – and the case for any broad and overarching intervention into the chip-design market at this point seems premature given the growing potential for competition in this space.
- **Keep a 'watching brief' on the possibility that developers of open AI models will seek to close off access to these models in future**, while keeping in mind that there is currently limited evidence of firms adopting an 'open early, close late' approach. If a larger AI model developer were to close access to a previously open model or resource, competition authorities should take into account whether there are legitimate business reasons for such conduct (such as protecting the AI developer's incentives to invest and innovate).
- **Carefully consider potential interventions to protect AI developers' ability to switch between providers of accelerated compute**. However, in addressing technical barriers to switching, authorities must balance the dangers that interventions will limit innovation and differentiation in the sector; take into account the possibility of private sector technical solutions to help overcome some of the switching barriers that exist; and consider the specific circumstances of AI developers, which are likely to be fairly sophisticated customers.
- **Apply the ban on egress fees set out in the Data Act, and continue monitoring the provision of discounts by providers of accelerated compute to avoid a 'customer lock-in' effect**, while keeping in mind that discount schemes often indicate a vibrant competitive market and offer significant benefits to innovative customers.
- **Monitor and scrutinise partnerships between major cloud service providers and AI startups.** However, these partnerships have strong efficiency effects and should not be blocked per se. Pay particular attention to exclusivity clauses in partnership agreements.
- **Apply caution when scrutinising practices associated with vertical integration** (such as self-preferencing, tying and leveraging) since these may often indicate pro-competitive conduct, such as helping consumers discover new services and helping high-cost, high-risk business ideas get scale. Competition authorities will need to review the full context of the conduct and its impacts to determine whether such conduct is anti-competitive, and ex ante regulation like

the Digital Markets Act might not be the best tool to tackle any perceived problems with vertically integrated players.

- **Ensure that regulatory barriers such as the AI Act are applied proportionately** in order to further growth and competition in the AI and cloud sectors, and continue the European Commission's current regulatory simplification exercise to help minimise regulatory costs which might dissuade entrepreneurs and smaller firms from entering the AI sector, or businesses across Europe generally from taking up AI.

1. Introduction

As the European Commission has acknowledged in its AI Continent Action Plan, availability of diverse, high-quality and affordable AI systems is critical to Europe's economic growth.[1] Europe faces enormous economic headwinds – among them, a shrinking working population, high energy prices impacting its manufacturing sectors, the loss of Europe's technological edge against its geopolitical competitors in many sectors, and several of its largest export markets turning to protectionism. To escape this crisis, European businesses will have to become far more productive and innovative than they are today, including by taking better advantage of new technologies.[2]

AI offers particular promise because it is a general-purpose technology. Its purpose is to lower the cost of making predictions based on prior data. This is a general task with countless applications across multiple economic sectors. Consequently, AI has the potential to lower barriers to entry across many markets, and to help many existing firms become more efficient. By lowering the cost of making predictions, AI can be a driver of even more innovation and transform markets by allowing completely new products and services to emerge. **AI is therefore a technology which has potential to boost competition across the EU economy. These benefits to competition and innovation will be fundamental drivers for European growth in the future**, given how market power is increasing across economic sectors[3] and the pace of innovation is slowing as innovations become harder to find.[4]

The EU should be especially optimistic about AI because it offers particular benefits for the services sector. The EU's traditional strength in export-oriented manufacturing has benefited from past waves of innovation, such as robotics and automation. But services represent a much higher proportion of the European economy – currently about 70% of total GDP – and an economic model based on ever-increasing exports of technically sophisticated goods looks harder to maintain, given growing trade barriers with some of the EU's largest trading partners. This means Europe's future will have to lie in high-end services like design of high-tech devices and software. Productivity growth in the services sector has been stagnant for many years, however. AI offers an opportunity to drive increased productivity and competitiveness in the services sector – one of the most exciting uses of AI, for example, is to help computer programmers write code much faster than they could previously and to help other office workers process data and produce text much more efficiently than they previously could. Despite the ICT revolution, European services firms have often struggled to use technology to boost their output and efficiency, however. The US saw a huge productivity boom in services from the ICT revolution of the 1990s but Europe missed out.[5]

In an era where Europe is increasingly concerned about excessive unilateral dependence on certain trading partners, AI could also offer new opportunities for Europe to grow the size of its technology sector and secure strategic footholds in the tech value chain. AI is a general-purpose technology which – as we explain below – is creating new economic markets. Many of these markets are still growing rapidly, with plenty of opportunity for challengers to grow their customer bases rapidly. Even the largest tech firms are facing uncertainty about which business models will succeed, which types

[1] European Commission, 'AI Continent Action Plan', 9 April 2025.
[2] Mario Draghi, 'The future of European competitiveness', September 2024; Enrico Letta, 'More than a Market', April 2024.
[3] Wenjie Chen et. al, 'Rising Corporate Market Power: Emerging Policy Issues', IMF, 15 March 2021.
[4] Robert J Gordon, 'The Rise and Fall of American Growth: The U.S. Standard of Living since the Civil War', Princeton University Press, 2016.
[5] Robert Gordon and Hassan Sayed, 'Transatlantic technologies: The role of ICT in the evolution of US and European productivity growth', National Bureau of Economic Research, 2020.

of AI-powered services will be valued by customers, and what tomorrow's centrifugal points in digital ecosystems will be. **Using AI will give Europe an opportunity to gain market leadership in a sector where market leadership is still up for grabs and market structures are far from settled.**

To maximise the economic benefits of AI and its opportunities to boost European competitiveness, **the AI sector needs to be able to enjoy rapid growth – and firms active in the sector need sustainable and enduring competitive pressure to continue to invest, innovate, and make their services as widely available as possible**. As we explain below (and have explored in more depth in our earlier issue paper[6]), AI is characterised by a diversity of business models – and a noteworthy degree of openness. Many of the most successful firms allow widespread use of their assets, including their computing resources, models, and platforms. Several even adopt a fully open-source approach. This suggests that the sector has a significant amount of growth ahead, and that a priority should be to reduce any artificial regulatory barriers to investment in AI and to encourage take-up, rather than to intervene too eagerly in a sector where competition is currently delivering good results.

This also has implications for how competition authorities approach the sector.[7] Competition authorities have been scrutinising the AI sector closely – understandably given its fundamental importance to future economic growth and its potential relevance to many different segments of the economy. In general, competition authorities appear now to have fewer concerns than they did a year or two ago. Our own analysis – as set out in our earlier issue paper – suggests that this more relaxed approach is indeed warranted. We found, in summary, that:

- The AI sector is characterised by high levels of static competition. However, when examined through a dynamic lens – by looking at levels of radical innovation, the investments being made by large and small AI firms and private equity and venture capital investors, and the degree of uncertainty about which business models, user propositions and technological solutions will ultimately succeed – levels of competition appear even more significant.
- The AI sector relies on the provision of 'accelerated compute'. On a static analysis, a small number of players appear to have a very strong position in providing this service to the AI sector. However, a more dynamic analysis suggests that the largest players in this part of the value chain are subject to emerging competitive challenges, such as new specialised entrants and technological changes which may weaken their existing positions.
- Despite the signs that competition in key parts of the AI sector may be strong and improving, there remain certain factors that could damage competition in the future. These include unilateral conduct (such as firms withholding access to essential inputs like data or accelerator chips); general characteristics of the market (such as difficulties AI developers may have in switching accelerated compute providers due to technical incompatibilities); and the impact of regulations which pose unreasonable barriers to entry to firms seeking to enter the AI sector or support it, for example by building data centres.

[6] Zach Meyers and Marc Bourreau, 'A competition policy for cloud and AI: Issue Paper', CERRE, [x] June 2025 ('Issue Paper').
[7] Bundeskartellamt and Autorite de la Concurrence, 'Working Paper - Algorithms and Competition', 6 November 2019; Autorite de la Concurrence, 'Generative artificial intelligence: the Autorité issues its opinion on the competitive functioning of the sector', 28 June 2024; Autoriteit Consument & Markt, 'Onderzoek naar toezicht op algoritmische toepassingen', 10 December 2020; Competition and Markets Authority, 'AI Foundation Models: Update paper', April 2024.

This report builds on the findings of our issue paper. It focuses on the factors which might constrain competition in the sector in future, and examines existing and proposed regulatory measures aimed at addressing these problems. Our overall conclusion is that, given the high levels of competitive pressure which currently exist, authorities ought to be cautious about intervening too pre-emptively in existing market dynamics: many interventions may have unintended consequences such as reducing incentives to innovate and some interventions risk dampening conduct (like discounted access to cloud computing) which has a strong positive impact on competition between AI providers. We therefore recommend that any new regulatory obligations are tightly targeted and, where they have the objective of furthering competition, are designed in ways which allow the positive and negative economic impacts of particular behaviour to be weighed.

Our report focuses on **competition between AI developers and at higher levels of the value chain, in particular the provision of accelerated compute, to ensure that AI services can be widely available on the market. We recognise that uses of AI may create competition concerns, for example when AI systems provide recommendations, affect what choices users are shown, or directly make decisions on behalf of users. Many of these concerns – and potential regulatory interventions to address them – are analysed in a separate CERRE issue paper and report on AI agents.**[8]

This paper is structured as follows:

- Section 2 summarises our earlier findings on current levels of competition and relationship between AI developers and cloud computing services and what may change in the future. We summarise the potential barriers to effective competition (including barriers imposed by economic actors, general market characteristics, and regulatory interventions) which have been postulated by competition authorities or which are revealed by our own analysis.
- Section 3 takes a deeper dive on particular barriers to effective competition and links these to regulatory initiatives which have been imposed, or which European authorities have proposed for the AI sector. This section evaluates how well these each address the potential barriers to effective competition.
- We conclude in Section 4 with concrete conclusions about how authorities should decide whether to make competition and regulatory interventions and the shape any such interventions should take.

[8] Friso Boeston and Jan Kramer, 'AI Agents and Ecosystems Contestability', CERRE, 5 November 2024.

2. Competition in the AI Value Chain Today

In our earlier issue paper, we described how competition works in key parts of the AI 'value chain': starting with inputs into AI models, competition between AI models, and then examining downstream activities.

2.1 Key Conclusions from Prior Work

Our main finding was that, in general, competition appears to be thriving. Both AI and the provision of accelerated compute are characterised by very high levels of investment by a variety of different players, including some of the largest tech firms and also those many times smaller. Innovation in these markets is occurring rapidly, with high levels of uncertainty about which business models will succeed, and the markets are growing in size with room for many different types of players to increase their customer bases. In markets where demand is relatively static and most customers are already served, competition tends to focus on price and quality, and innovations tend to be incremental; but that is not the case where firms are innovating and experimenting – and, in the case of hyperscalers,[9] facilitating other AI developers to innovate – in order to unlock latent customer demand.

Our key conclusions were as follows:

- In mature markets – where most customers are already served, competition tends to focus on price and quality, and innovations tend to be incremental – a traditional approach to competition policy which focuses on metrics such as market shares and the impact of price changes in determining the boundary of an economic market and whether a particular player has market power might be appropriate.[10] **In high-growth, high-potential sectors like cloud and AI, a more dynamic approach to competition policy would be more useful.** Such a policy would seek to identify and promote innovation – such as increasing competition 'for' markets, particularly when businesses are still trying to work out how to get the most use of a new technology – rather than simply increasing the number of providers offering exactly the same service. In digital markets, in particular, competition 'for' emerging markets may drive significant positive consumer outcomes. This requires closer examination of the way technologies and business models are evolving, the degree of uncertainty under which large players are operating, the pace of technological change, and prospects for disruptive innovation, rather than a static view of existing players in the market.

  It is widely accepted that competition authorities have not always taken a sufficiently dynamic approach to applying competition policy in the past.[11] In the context of the AI sector, assessing competition with 'traditional' metrics such as market share would be particularly inadequate, since the 'markets' themselves are still being determined, and since players in the AI sector are competing to position themselves against potential future competitors, and face

[9] We do not necessarily mean to suggest that any hyperscalers' activities take place in a separate market, nor do we take a stance on this question. This market may include other players besides hyperscalers.
[10] Zach Meyers, 'Competition policy must reflect Europe's reality, not its aspirations', Centre for European Reform, 23 October 2024.
[11] Wolfgang Kerber & Simonetta Vezzoso, 'Competition and innovation: Incorporating a more dynamic perspective into enforcement', 3 January 2025.

uncertainty about which AI businesses will succeed and whether the sector will create new focal points for digital ecosystems.

- The AI sector is characterised by high levels of static competition – but measures of dynamic competition indicate that competitive tension in the AI sector is likely even higher than a static assessment would assume. A spectrum of business models exist – at one end, closed-source proprietary models, some of which are only available to selected customers, often with usage limitations; more open models such as Meta's Llama model, where information such as the weights and design of the model are public and the model can be widely used, but the full training model and code is not always available; and on the other end of the spectrum fully open models (like Hugging Face's Bloom model) which can be freely used, adapted, and deployed to provide training to new models, and where the source code and training data are transparently disclosed. Even among the largest AI developers, including the hyperscalers, many of their smaller AI models are published openly with only the most advanced models being fully closed-source.[12]

  Many AI providers are also innovating by exploring different means of differentiation, for example by increasing their performance at certain specific tasks. This suggests that AI providers are still exploring how to fulfil different customers' needs. Innovations do not only come from the largest and most established or highest-performing models: DeepSeek demonstrates that quite radical innovations can come from unknown new entrants.

  AI providers generally appear to be operating their services at a loss. That is particularly significant now that the marginal costs of performing inference to answer queries appears to be increasing.[13] This can indicate high levels of dynamism because companies tend to use below-cost or zero pricing (sometimes supported by cross-subsidising profits from other services) in order to 'create' a market and encourage customers to try out an innovative service.

  This illustrates that the AI sector is characterised by high levels of dynamism, with both demand and supply growing rapidly, and firms not yet clear about the future shape of economic markets in this sector. This implies high levels of experimentation and risk-taking, as illustrated by high levels of investment by large and small players.

  Even on more static measures of competition, the AI sector appears healthy. Thousands of different foundational models exist. Many models offered by AI developer start-ups have an edge on models from the largest firms in certain respects: when start-up Anthropic launched its Claude 3 FM in March 2024, for example, it could beat Google's Gemini FM on both undergraduate-level expert knowledge and graduate-level expert reasoning.[14] AI providers are continuously leapfrogging each other in innovation and performance.[15]

- **AI developers need access to data** and there is the potential for this to become a bottleneck in some contexts. It does not appear currently to be a constraint on the development of large language models. But the sector appears to be moving towards greater investment in use-

[12] Autorite de la Concurrence, 'Opinion 24-A-05 on the competitive functioning of the generative artificial intelligence sector', 28 June 2024, paras 183-4.
[13] Bertin Martins, 'How DeepSeek has changed artificial intelligence and what it means for Europe', Bruegel policy brief, 12/25, March 2025.
[14] Angela Yang, 'Move over, ChatGPT: AI startup Anthropic unveils new models that challenge Big Tech', NBC, 4 March 2024.
[15] Competition and Markets Authority, 'Cloud Infrastructure Services: Provisional decision report', 28 January 2025, para 3.479.

case-specific models, which may help AI developers to have greater access to specialised datasets over which one firm, or a small number of firms, may have control. This raises the risk of firms which hold that data withholding, or imposing unreasonable terms, on access in order to protect themselves from competition or business model disruption.

- **The provision of accelerated computing power to AI firms requires a different competition analysis to the general cloud computing sector.** AI developers generally require access to computing power which can perform many different operations in parallel (known as 'accelerated compute'); this functionality can only be provided by specialised 'accelerator' processing chips. It is used by AI developers to train their models; for intermediary steps like fine-tuning a model with more curated datasets, so it can be adapted to deliver better results at particular types of tasks or uses; and to respond to an input provided to the model by a user (known as 'inference'). Many AI developers choose to use the largest cloud computing companies (Amazon Web Services, Microsoft and Google), known collectively as the 'hyperscalers', for accelerated compute. AI developers' access to compute remains one of the most significant inputs to AI development and a key input over which competition authorities have had concerns.

  The general cloud computing sector has been extensively studied by numerous competition authorities in Europe, many of which have identified problems with how markets are operating. However, the provision of accelerated compute to AI developers has different competitive characteristics – with more entrants, potentially lower barriers to entry, and specialised firms which offer a different set of advantages over the hyperscalers in serving AI developers. That has important implications for competition policy.

- **Competition between providers of accelerated compute appears to be intensifying.** Between the hyperscalers, Microsoft and Google trail Amazon in their share of the overall cloud computing sector, but this is not holding them back from winning a greater share of AI projects. There are also a growing number of both general cloud computing providers beyond the hyperscalers which are growing their AI capacity (such as Oracle, IBM, Alibaba Cloud, OVHCloud and Scaleway).[16] Barriers to entry for smaller accelerated compute providers are falling too: as AI models have started to train on each other's outputs, the costs of AI development have shifted away from initial training and towards fine-tuning and inference.[17] This avoids smaller accelerated compute providers needing to spend significant up-front capital on building their AI model training capacity – potentially allowing smaller providers to more easily compete with the hyperscalers. Smaller cloud computing providers which specialise in providing accelerated compute to AI developers, such as Lambda Labs, Denv-r, TensorWave, CoreWeave, Vultr and Nebius, are also growing in number and achieving significant success. By focusing specifically on AI (rather than repurposing existing cloud computing infrastructure) these specialised providers have some competitive advantages over the larger players.[18] Third, many AI developers are focused on 'multi-cloud' approach, using multiple providers (including in many cases a non-hyperscaler) and ensuring they can shift their compute volumes between their different vendors. This illustrates that many AI

[16] OpenAI, 'Announcing the Stargate Project', 21 January 2025, available at https://openai.com/index/announcing-the-stargate-project/.
[17] Tim Bradshaw, 'How 'inference' is driving competition to Nvidia's AI chip dominance', Financial Times, 11 March 2025.
[18] See, eg, CoreWeave S-1, 3 March 2025, available at https://www.sec.gov/Archives/edgar/data/1769628/000119312525044231/d899798ds1.htm?ref=runtime.news.

developers are keeping their options open and are trying to avoid 'lock in' to any one cloud provider, and should impose significant competitive pressure. Fourth, today only a few AI developers (such as Samsung and Meta) can self-provide accelerated compute: however, as the computing demands of some AI models decrease, many more AI developers and users may be able to self-supply in future.[19] Fifth, in Europe, publicly funded supercomputers and 'AI gigafactories' may also provide an alternative for some AI developers.

Several big tech players – such as Apple, Meta and NVIDIA – do not have large-scale cloud computing platforms of their own. **These large and influential players have the resources and incentives to avoid the provision of accelerated compute becoming too concentrated, and are actively pursuing business strategies that should help limit the economic leverage hyperscalers might be able to exert over the AI sector.** Meta, for example, is focused on developing AI models which are more open than many alternatives and can operate across different cloud environments. Apple is working on AI models which can – to a large extent – run inference on a user's own device, limiting the need to rely on cloud computing.[20] NVIDIA – the most significant designer of accelerated AI chips today – has an active strategy of supporting alternatives to the hyperscalers, for example by giving smaller cloud computing services early access to NVIDIA's latest accelerator chips and financially investing in alternative providers of accelerated compute such as CoreWeave. This business strategy ensures NVIDIA maintains a diverse customer base, but also helps to ensure that AI providers have a good range of access to different external providers of accelerated compute.

- **In general, partnerships between AI firms and the largest cloud computing providers have improved competition**. These partnerships often involve the cloud computing provider taking a financial stake in the AI developer in return for concessions about the AI developer's access to accelerated compute, along with an agreement to host the AI model on the hyperscaler's platforms, or to integrate the AI developer's services into the hyperscaler's existing services. They seem to benefit AI developers but are not necessarily required to successfully develop AI systems, as a number of AI developers have successful products without formal partnerships with hyperscalers. These partnerships differ greatly, and are not always transparent, but in at least some of these deals the AI developer has agreed to use one of the large cloud computing providers as their 'primary' or 'preferred' provider.

  Initially, many competition authorities were concerned that these partnerships 'locked in' AI firms to one computing provider, or would result in hyperscalers having excessive influence over AI firms. However, **as competition authorities have scrutinised many of these deals, they have tended to discover that most are non-exclusive and do not significantly impede AI firms' freedom to innovate**, including to compete with the hyperscaler's own AI services. Authorities have generally not found these partnerships to pose competitive problems that justified intervention.[21] Consequently, many of these agreements appear to have benefited

[19] https://www.google.com/url?sa=t&source=web&rct=j&opi=89978449&url=https://go.techinsights.com/l/1043171/2024-10-23/8lvv8x&ved=2ahUKEwjG5IiA2fGMAxUOSPEDHbcJCcYQFnoECBcQAQ&usg=AOvVaw022juzX3HqcXMsYl8hxBpe

[20] See, eg, Apple, 'Introducing Apple's On-Device and Server Foundation Models', 10 June 2024.

[21] The Microsoft/OpenAI partnership was examined by the European Commission, which concluded the partnership was not a deal which qualified for review under the EU's merger control regime.; The CMA launched investigations into the Amazon/Anthropic and Google/Anthropic partnerships. The Google/Anthropic deal did not meet the criteria for merger review.; The CMA also investigated Microsoft's partnership with Inflection AI, but found that it did not substantially reduce competition, and Amazon's partnership with Anthropic.

the AI firms – which usually retain significant autonomy. This was illustrated by OpenAI's decision to diversify away from Microsoft: despite having a much-publicised quasi-exclusive partnership with the firm, OpenAI recently negotiated a new partnership with Microsoft's cloud rival Oracle and agreed to a $11.9 billion deal with CoreWeave, a smaller and more specialised provider of accelerated compute.[22]

- Past experience shows that radical changes to the structure of digital markets can happen relatively quickly. Nevertheless, several factors – such as the decreasing costs of developing AI, the growing interdependence of AI models, and the prominence and success of relatively open AI models – suggest that **competition authorities should not assume market consolidation is likely or that 'winner takes all' outcomes will occur in the AI sector**. That means, in assessing whether any competition problems have emerged, and the risk of intervening too late to effectively correct these problems, authorities should give adequate weight to the risks of being overcautious.
- Even without a 'winner takes all outcome' being likely, the concentrated nature of some parts of the AI value chain and the strong position of the hyperscalers across multiple parts of the value chain, create the **possibility of future harm to competition**. However, authorities should be cautious about too quickly assuming that certain practices are harmful. For example, many practices which have the potential to limit static competition, in some contexts, can also be markers of (and necessary for) strong dynamic competition. For example, a firm may leverage one of its existing services to help promote an innovative new AI feature in order to help disseminate and encourage take-up of innovative AI services. The ability to leverage services can be necessary to help build a business case for a risky investment.

As some competition authorities have already observed, the **hyperscalers along with NVIDIA enjoy some advantages over other accelerate compute providers and AI developers** – for example from having access to expensive inputs (such as accelerator chips) and existing extremely popular services which can serve as "anchor tenants" for AI models, provide data to train that model, ensuring that their own accelerate compute enjoys significant demand.[23] Not all of these advantages can be replicated by smaller companies. **However, non-hyperscaler business models have their own advantages, not all of which hyperscalers have been able to match** – indeed, the level of investment, dynamism and innovation in the development of AI models and in independent accelerated compute companies like CoreWeave suggests that investors believe many companies in the AI sector have a bright future beyond the hyperscalers.

The idea that AI could become a new focal point for digital ecosystems has led to **suggestions from some competition authorities that independent AI developers should be 'protected' from the influence of large technology firms** – and concern that independent AI developers may be forced "to cooperate with big tech firms to get access to computing infrastructure and end users".[24] Presumably, these authorities' concern is that AI developers should feel freer to focus on radical and disruptive innovation rather than 'sustaining' innovation which complements rather than upends hyperscalers' existing market positions in tech markets. The CMA, for example, has set out 'principles' to guide how

[22] https://www.nextplatform.com/2025/03/11/what-a-tangled-openai-web-we-coreweave/
[23] Autorite de la Concurrence, 'Opinion 24-A-05 on the competitive functioning of the generative artificial intelligence sector', 28 June 2024.
[24] Bertin Martins, 'Why artificial intelligence is creating fundamental challenges for competition policy', Bruegel, policy brief, 18 July 2024.

the AI sector should look – emphasising high levels of interoperability and that hyperscalers should not exercise "undue influence" over independent AI firms.[25]

In general, we **do not find evidence to support a view that the influence of the hyperscalers would inevitably be negative for competition**. Competitive pressure by hyperscalers can provide incentives for smaller and independent firms to pursue diverse business models. If there is fierce competition between hyperscalers, they will have the incentives to pursue radical innovations in order to disrupt their competitors.[26] The hyperscalers may also have the ability to pursue some radical innovations which might not otherwise make it to market (and help radical innovations by independent AI developers succeed) due to their size, access to capital, and their ability to leverage their existing customer bases to encourage a critical mass of consumers to try an experimental new service. Chatbots like ChatGPT have attracted significant consumer interest, but they have not yet posed existential challenges to existing large digital platforms. It is more likely that the benefits of AI will be enjoyed by greater numbers of customers - and disruptions may happen - when they are well integrated into existing services, which will likely require hyperscalers to play a significant role in the AI ecosystem.

2.2 Maintaining a Competitive Market

Our issue paper identified a number of concerns that competition authorities might have about how to maintain a competitive and dynamic market. This includes being alert to several types of developments that could hinder effective competition in future.

Potential unilateral anti-competitive conduct

First, we identified **potential anti-competitive conduct arising from firms which could control the inputs to AI developers:**

- **One set of concerns arises from the risk that holders of data valuable for training AI models might withhold access to that data**. Competition authorities previous raised concerns about large technology firms controlling the inputs independent AI developers would need, such as data.[27] In practice, data has turned out to be widely accessible and difficult to monopolise, and AI technology is developing in ways which rely less on data (for example, only highly curated datasets may be required to fine-tune a model for a particular use case, and there is growing focus on the quality rather than the quantity of training data). However, in particular use cases, AI developers will require access to specific datasets for which no alternative is available. These datasets may be withheld in ways that might limit competition for using AI in particular use cases – in particular where they compete with or pose a commercial challenge to the data holder.[28] These strategies could be pursued by a hyperscaler because they each have a broader ecosystem of digital services to protect. However, this type of strategy could

[25] Competition and Markets Authority, 'CMA AI strategic update', 29 April 2024.
[26] Statista, 'Global search engine traffic market share of Bing from January 2018 to January 2025', available from https://www.statista.com/statistics/1219326/market-share-held-by-bing-worldwide/.
[27] Competition and Markets Authority, 'AI Foundation Models: Initial report', September 2023.
[28] For example, Google limited access to its search data and Microsoft reportedly threatened to cut access to its search index to customers using the data to build AI tools: see Leah Nylen and Dina Bass, 'Microsoft threatens data restrictions in rival AI search', Bloomberg, 25 March 2023. See also Autorite de la Concurrence, 'Opinion 24-A-05 on the competitive functioning of the generative artificial intelligence sector', 28 June 2024, para 261.

also be pursued by any other holder of data which fears its business model being disrupted by AI.

- The price for cutting-edge AI accelerator chips is increasing, which could pose constraints on the growth of the AI sector. **This raises potential questions about NVIDIA's strong position in the market and whether its pricing and commercial strategy reflects the significant costs and risks of its development efforts, or instead indicates that it is exercising a degree of market power.** These concerns also relate to the use of CUDA, the proprietary coding language owned by NVIDIA, which is the industry standard for running AI software on NVIDIA's accelerator chips and is the only coding language that is fully compatible with those chips.
- One notable development in the sector is that many smaller AI models are becoming more heavily reliant on larger models for example to produce high-quality training data, to "teach" a smaller AI model chains of reasoning, or to specialise in fine-tuning larger model, or with AI developers designing value-added services to work on top of existing models.[29] This means that an 'input' to AI models is increasingly the output of other models. **This raises potential concerns if 'open' models become closed once the shape of economic markets become clearer and firms start to focus more on increasing their share of a well-defined market rather than 'growing the pie'.**

Relationship between accelerated compute and AI

Second, we identified **a series of possible concerns relating to the relationship between providers of accelerated compute and AI developers**. These include:

- **Commercial barriers to switching.** Providers of accelerated compute may offer 'credits' or discounts to AI developers to attract them as customers. These practices are also prevalent for AI developers,[30] including in AI developer/hyperscaler partnerships.[31] Discounts are generally pro-competitive: they can illustrate strong competition between cloud computing companies to win customers, and can boost competition in the AI sector by lowering AI developers' costs. Anti-competitive effects from discounts seem unlikely in a market which is highly competitive. In general, since offering discounts simply requires a willingness to accept lower returns for a period, competing cloud computing companies should be able to offer their own discounts – Oracle, a large cloud computing provider which is still significantly smaller than the hyperscalers, offers significant credits to AI providers,[32] for example – and 'free trials' or similar schemes are common in many highly competitive parts of the economy. Although discounts make life difficult for competitors, they will usually indicate a fierce battle to win customers rather than to squeeze out rivals. **However, competition authorities may need to examine discounts if they have problematic characteristics, such as (contractually or in practice) imposing a high degree of exclusivity and/or limiting AI developers' choices, and are therefore likely to be a way to attract developers and then lock them in.** Authorities

[29] Autorite de la Concurrence, 'Opinion 24-A-05 on the competitive functioning of the generative artificial intelligence sector', 28 June 2024, p 43.
[30] Autorite de la Concurrence, 'Opinion 24-A-05 on the competitive functioning of the generative artificial intelligence sector', 28 June 2024, fns 146-147.
[31] Federal Trade Commission, 'Partnerships Between Cloud Service Providers and AI Developers FTC Staff Report on AI Partnerships & Investments 6(b) Study', January 2025, p 3.
[32] Aaron Holmes, 'AI Startups Find an Unlikely Friend: Oracle', The Information, 22 February 2023, https://theinformation.com/articles/ai-startups-find-an-unlikely-friend-oracle.

have also been investigating the impact of other commercial practices like charging egress fees for moving data, which some smaller firms argue can pose an unreasonable barrier to switching.[33] These may also pose challenges to some AI developers who might otherwise switch between accelerated compute providers.

- **Technical barriers to switching and 'lock in' effects.** Identifying whether these are anti-competitive poses challenges because technical barriers to switching can often arise as unintended side effects of introducing new features and service differentiation. These features and differences can benefit customers even if they (as a necessary consequence) make switching provider more difficult. Some lock-in effects may occur without any active decision by a market player to make switching more difficult. For example, consumers and businesses may use particular AI services, which over time may learn from that consumer or business's usage of the service, making it difficult to switch without losing that usage history.
- **The role of accelerated compute providers in providing 'channels to market' for AI developers.** Some providers of accelerated compute to AI developers also operate downstream platforms (such as Google's Model Garden, Amazon Bedrock and Microsoft Azure AI Model Catalogue) enabling AI developers to reach business users. The development of such 'platforms' from the hyperscalers is positive for competition and innovation, since they enable widespread and affordable accessibility of different foundation models to businesses. One concern, however, is that the operators of some of these platforms have developed their own foundation models. This has caused some competition authorities to question whether these platforms may either in future use their platforms to actively drive customers towards their own foundation models or influence the **direction of AI development in ways that complement rather than challenge the hyperscalers**. Vertically integrated players could do so through self-preferencing their own models on their AI model platforms for business users, tying services, or integrating and bundling their AI services with other 'must have' services, or giving only certain AI developers access to the most cutting-edge models.
- The final stage in the AI value chain is **the use of models for customer-facing applications**. In some cases, AI developers are choosing their own (or a chosen third party's) AI model for integration into their existing products and services. This raises concerns about self-preferencing, tying, integration and bundling. In particular, integration with an incumbent's services can give an AI developer an important 'anchor tenant', guaranteeing revenue and use of their AI model. **However, it may also create questions about what level of integration is beneficial – and when users should have a choice of AI models that will work with a particular digital service.** This may be more of a concern where it involves leveraging or tying of an AI-related service with a potentially dominant digital services in a more mature market. In these contexts, however, competition authorities will need to examine the impacts of the practice in the particular case. Competition authorities should be cautious about intervention, since the integration of new AI model functionality may represent an innovation and benefit consumers.

[33] Competition and Markets Authority, 'Cloud Services Market Investigation: Summary of provisional decision, 28 January 2025.

Regulatory interventions that may limit competition

Finally, we noted in passing that there are a number of regulatory interventions proposed or already passed which are aimed at achieving non-competition objectives, but which might have an impact on levels of competition in the AI sector. These include:

- **Regulatory interventions regarding IP rights**: policymakers are considering which rights AI developers should have to train models on data covered by intellectual property ('IP') laws without the active consent of the rights holders.[34] This will have potential competitive impacts between AI developers, since larger AI developers will be more likely to have the resources to negotiate access to content protected by IP compared to smaller firms.

- **Regulatory and commercial constraints on the rollout of data centres:** provision of accelerated compute requires sufficient access to data centres to host the computing equipment, and in turn those data centres require electricity grid connections, network connectivity, and water. However, the planning process for building new data centres, and the need for supporting infrastructure, often means the infrastructure to support provision of more accelerated compute is lacking. Regulatory or other barriers to constructing data centres could have the effect of both hindering the growth of AI and also ossifying the existing market structure, by advantaging firms which have already made investments in data centres over new entrants.

- **Laws such as the AI Act and the GDPR may pose significant barriers to small firms who wish to grow in the market, due to the significant one-off and ongoing costs that compliance with these laws requires**. One important concern among EU policy-makers is to ensure that AI in the EU is developed in ways which protect fundamental rights. This objective guided the AI Act, and justifies the application of large parts of the EU digital acquis to AI. While such laws may have sensible objectives, competition authorities should nevertheless remain conscious of the competitive impacts these laws could have due to costs and complexity of compliance. Since larger firms may tend to be more capable of ensuring compliance, the impact of such laws has the potential to disproportionately affect newer market entrants.

- **'Sovereignty' oriented interventions:** Another important concern in Europe is fear of overreliance on foreign cloud computing and AI services. This has long provoked European fears. Some EU member states and parts of civil society have for years raised concerns about US laws, which may in some circumstances require US cloud firms to hand over Europeans' data to law enforcement authorities.[35] Policy changes in Washington are now contributing to fears among some European policymakers that there are new risks to dependency on foreign services, prompting renewed efforts to support European cloud computing and AI investments. Such measures have some potential to distort competition, including if they create an unlevel playing field and/or provide strong incentives on firms or public authorities to use 'European' solutions instead of those that might offer the best value or most appropriate functionality. However, if well designed such interventions may also contribute to competition by enabling the sustainable entry or growth of new players in the sector.

[34] Ana Rački Marinković, 'Liability for AI-related IP infringements in the European Union', Journal of Intellectual Property Law & Practice, 19(10), October 2024.
[35] Zach Meyers, 'Can the EU afford to drive out American cloud services?', Centre for European Reform, 2 March 2023.

2.3 Summary of Potential Issues That May Limit Competition

In the table below we summarise the potential concerns.

Potential competition concern	Target of potential concern	Potential impacts on competition
Foreclosure of, or unreasonable provision of, access to data	Any data holder	Would limit the growth of AI in areas which compete with a data holder's own services. May have disproportionate impact on smaller AI developers which lack resources to negotiate with significant data holders
Foreclosure or unreasonable terms of access to accelerator chips	NVIDIA	Would limit the growth of, and potentially competition between, accelerated compute providers by raising barriers to entry
'Lock-in' or foreclosure of open AI models	Providers of more open AI models	Would limit the growth of AI firms which rely on previously more open AI models to build their own models and services
'Lock-in' effects from commercial practices of accelerated compute providers (e.g. discounts and egress fees) or technical barriers	Providers of accelerated compute	May limit AI developers' choices and ability to choose between providers of accelerated compute
Self-preferencing, tying, bundling or similar leveraging practices when providing AI model platforms and/or between an AI service and an existing end-user facing tech service	Providers of AI model platforms	May impact growth and diversity of AI sector by limiting some AI services' access to customers or free innovation by AI developers, and may impact competition in adjacent markets.
Regulatory / commercial constrains on rollout of data centres and other supporting infrastructure	Public authorities	May impact growth of data centres and provision of accelerated compute, which may turn benefit companies which have already made investments in these assets over new entrants.
Application of the AI Act and the existing digital acquis to AI	All market participants	Complex or burdensome regulation may have disproportionate impacts on smaller firms and new entrants.

3. Potential Harms and Regulatory Interventions

We set out below our detailed views regarding the potential future harms to competition. In relation to each concern, we also describe the regulatory interventions currently in place or proposed at the EU level for the AI and cloud sectors. We provide, in each case, an assessment of:

- Whether the intervention (and its scope) is justified based on the likelihood and severity of the potential concerns we have set out above materialising;
- Whether there are likely to be any costs or unintended consequences of these interventions on competition, innovation or investment; and
- Whether there are alternative regulatory approaches which could reflect a better trade-off.

3.1 Access to Data

Competition authorities are raising concerns about access to data for AI developers due to the risk that holders of valuable data for training or fine-tuning AI models might withhold access to it or impose unreasonable terms on that access.

In the issue paper, we argued that, in practice, data has proven to be widely accessible and difficult to monopolise, and that AI technology is advancing in ways that are less dependent on developers having direct access to ever-larger sources of data. However, for certain use cases, AI developers may still require access to specific datasets for which there is no alternative. These datasets may be withheld in a way that limits competition when it comes to using AI for specific purposes, particularly when they compete with or pose a commercial challenge to the data holder. Hyperscalers could pursue these strategies because they each have a broader ecosystem of digital services to protect. However, any other holder of data that fears its business model being disrupted by AI could also pursue this type of strategy.

The UK Competition and Markets Authority (CMA, 2023) distinguishes between two types of data: training data, which is used to develop the core capabilities of models, and alignment data, which is used to fine-tune their behaviour to reflect specific objectives. For example, this could involve aligning the model with the expectations of a human.

With regard to training data, the CMA (2023) notes that AI developers primarily rely on publicly available sources, including web-scraped content. However, it also notes a growing trend towards incorporating proprietary datasets. The CMA is particularly concerned about vertically integrated firms, especially major digital platforms, as they may have easier access to valuable proprietary data, thereby reinforcing their competitive advantage. With regard to alignment data, the CMA notes that such data tends to be proprietary, costly, and difficult to obtain. In particular, one concern might be that AI models improve in quality based on their ability to learn from how users respond to their outputs: meaning the more they are used, the more useful data the AI developer can collect to improve the model, in term prompting even more usage. At the moment, it is not obvious that such

'feedback loops' will necessarily be strong.[36] However, if they do become pronounced, they may create economies of scale which could eventually lead to 'winner takes all' dynamics.

From a general point of view, the CMA considers that firms controlling a critical input in the AI value chain can have the ability and incentive to restrict access to it in order to reduce competition. According to the CMA, data constitutes such a critical input (CMA, 2024).

The French Competition Authority (FCA, 2024) has identified several potential competition issues relating to access to data for AI developers:

- Firms that control key data could refuse competitors access to it. This assumes that they are active in the development of AI models or applications downstream, and that they may try to exploit their market power over key data to reduce competition downstream. Alternatively, they could offer discriminatory access by providing more favourable terms to themselves or their partners, thereby distorting competition.
- Exclusivity clauses can be implemented by firms that obtain data access to prevent their competitors from accessing the same data. This would harm competition to the extent that they are no alternatives to this type of data.
- In order to raise their rivals' costs, firms could purchase access to data at high prices, assuming that their rivals would have to pay similar prices. This could prevent smaller firms from accessing the data and also discourage potential new market entrants. In particular, the FCA cites the remuneration of content creators for access to their data, which could be set at high levels by dominant players to restrict competition. DG COMP (2024) raises a similar concern.
- Incumbents and vertically integrated firms could exploit legal rules (e.g.: those relating to the protection of personal data) or security concerns to restrict third-party access to their users' data and, consequently, restrict competition.

Although certain AI technological developments are reducing reliance on large-scale data inputs, access to data seems to remain a critical bottleneck, at least in some contexts.

In line with the analyses of the CMA and the FCA, it is therefore essential to ensure that where data is a genuine bottleneck, it cannot be withheld by the data holder in a way which precludes a potential competitor from entering a market or from exerting competitive pressure which would not otherwise exist. However, especially given the success of AI developers and the thriving levels of competition among them, it is also important that regulatory interventions do not undermine incentives for investment and innovation in the market – for example, by unfairly depriving firms of the opportunity to exploit data they have collected through their own investments or through undermining incentives to collect data in the first place.

There are a number of existing regulatory interventions in the EU which aim to improve access to data. Some of these have relatively low risks of negative consequences for innovation and investment. For example, the EU has passed some useful regulations which aim to help improve access to and the ability to trade data across the bloc, so that data can be obtained by firms best placed to exploit it. EU

[36] Klaus Kowalski, Cristina Volpin, and Zsolt Zombori, 'Competition in Generative AI and Virtual Worlds', Competition Policy Brief, September 2024

laws require public authorities to open up their datasets, for example;[37] the Union has removed barriers to intra-EU non-personal data transfers;[38] and the Data Governance Act (DGA) seeks to encourage voluntary data-sharing in the private sector.[39] Supporting voluntary data-sharing could indeed help to address some of the concerns about data access without the need for regulation.

However, other laws take a more directive approach by obliging certain data holders to share data they have collected. The GDPR and the Digital Markets Act (DMA) both include provisions which empower users to extract their data from a digital service themselves or to require the data holder to send it directly to another service provider.[40] The DMA also requires any regulated search engine to provide access to its 'click and query' data to a competing search engine.[41] The Data Act requires providers of 'connected devices' and associated services to make data collected by those devices and services available to third parties at the user's request.[42]

Rules empowering users to retrieve and direct how their own personal data is used and disseminated may be justified on the basis that this helps protect users' fundamental rights over their own personal data. However, obligations which mandate broader types of data-sharing can suffer from two flaws. First, they may not be comprehensive, because they are not able to apply to all circumstances where a data holder is in a position to throttle competition: for example, data may be 'hoarded' by many different types of firms, not just those regulated under the DMA as gatekeepers or providers of connected devices. Second, these rules may be overinclusive. For example, not all data is necessarily essential or non-replicable, even though identifying which data is which is probably difficult.

Designing ex ante regulation with broad overarching rules about how data should be shared, in order to prevent negative impacts on competition in the AI sector, is therefore likely to be rather complex. However, one context in which data portability obligations might be justified is to ensure that users of AI models can extract their usage history from a service and transfer it to another service. For example, as noted in our issue paper, as business users and consumers use an AI service, they will build up a history of queries and outputs, allowing the AI model to provide answers more tailored to the user's needs. Ensuring users can transfer this history from one AI model to another may be key to ensuring ongoing freedom of choice.[43] This may be considered analogous to existing data portability requirements in the GDPR. Where possible, to avoid imposing unnecessary uniformity and standardising services which are still innovating quickly, the development of portability tools to enable data sharing between innovative AI services should be by industry consensus. This might be more possible to achieve today – when the market is open and no player can be considered to have a dominant position – compared to in more consolidated markets where one or two players would have a lot to lose by promoting portability and the ability to switch.

In other cases, whether regulatory intervention to promote data mobility is necessary will depend on a detailed analysis of the facts: such as whether the data holder is in a position of dominance; whether the data in question is genuinely indispensable for competition; whether the data holder has

[37] https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex%3A32019L1024
[38] https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A32018R1807
[39] https://www.consilium.europa.eu/en/press/press-releases/2022/05/16/le-conseil-approuve-l-acte-sur-la-gouvernance-des-donnees/
[40] Eg, Digital Markets Act arts 6(9) and 6(10).
[41] DMA art 6(11).
[42] Data Act art 3.
[43] Chris Riley, 'The future of AI hinges on data portability and APIs', Data Transfer Initiative, 11 February 2025, available at https://dtinit.org/blog/2025/02/11/future-of-AI-portability; Chris Riley, 'Digging in on personal AI portability', Data Transfer Initiative, 4 June 2024, available at https://dtinit.org/blog/2024/06/04/digging-in-personal-AI.

legitimate reasons for withholding (or imposing certain conditions) on access; and whether the terms being imposed have anti-competitive effects. Given the complexity of the required analysis, access to data might be an area where competition authorities ought to carefully scrutinise potential abuses of dominance with a view to bringing enforcement proceedings on an ex post basis, rather than policy makers intervening with ex ante rules which might be poorly calibrated.

3.2 Terms of Access to Accelerator Chips

NVIDIA's strong position in the supply of accelerated AI chips has raised concerns from competition authorities, in particular due to the increasing prices for their cutting-edge AI accelerator chips and their scarcity.[44] The UK Competition and Markets Authority notes that "Nvidia continues to be the lead supplier" for accelerated computing (CMA, 2024), while the French Competition Authority recognises NVIDIA's "dominant position in the sector for the IT components needed to train foundation models" (FCA, 2024). While the CMA (2024) acknowledges that several competitors, including well-established chip manufacturers such as Intel and AMD, as well as major cloud service providers like Amazon, Microsoft and Google, have released or announced plans to release AI chips in the near future, the FCA (2024) also emphasises the ongoing reliance of the ecosystem on NVIDIA's proprietary CUDA programming environment. As of today, this remains the only environment that is fully compatible with NVIDIA's GPUs.

As a result, competition authorities in France, the UK, Europe, the United States and China are currently investigating NVIDIA's business practices.[45] NVIDIA's dominant position raises a range of potential competition concerns, including risks of excessive pricing, supply restrictions, the imposition of unfair contractual terms, and discriminatory practices (FCA, 2024). If AI developers cannot access accelerator chips easily and at competitive prices, there is also a broader concern that computational resources, a key input for AI innovation, could become increasingly concentrated in the hands of a few players (FCA, 2024).

A recent analysis of the accelerated computing market is provided in the European Commission's decision of 20 December 2024 regarding the acquisition of Run:ai, a supplier of GPU orchestration software, by NVIDIA.[46] In its decision, the EC concludes that "NVIDIA likely holds a dominant position in the global market for discrete GPUs for use in datacentres". The details of the Commission's market analysis are also of particular interest.[47]

First, the European Commission defines a separate market for the supply of discrete GPUs for use in datacentres. In other words, integrated GPUs, other accelerators, and offerings of accelerated compute capacity as a service do not belong to the same market as discrete GPUs.[48] Viewed from the angle of AI developers, however, this conclusion may not hold in the future, given the range of choices

[44] NVIDIA's H100 accelerator chip can cost up to $40,000 whereas a chip from two generations before (V100) cost only $10,000: TechInsights, AI Outlook Report 2025. In 2024, the waiting time to purchase an H100 chip (used for AI applications) from NVIDIA was also significant at around 3-4 months (DGComp, Competition in Generative AI and Virtual Worlds, Competition Policy Brief, September 2024).

[45] See, eg, Meaghan Tobin et al, 'China Opens Investigation Into Nvidia Over Potential Antitrust Violations', New York Times, 9 December 2024.

[46] See the press release, https://ec.europa.eu/commission/presscorner/detail/en/ip_24_6548.

[47] The European Commission cleared the acquisition of Run:ai by NVIDIA on the grounds that "the merged entity will not have the ability to foreclose competing suppliers of discrete GPUs for use in datacentres by leveraging its position in the market for the supply of GPU orchestration software."

[48] In particular, the EC excludes custom accelerators developed by companies such as Broadcom and Marvell, Amazon's Trainium and Inferentia accelerators, Microsoft's Maia 100 AI accelerator, and Broadcom's cooperation with Google on the tensor processing units ('TPUs').

that AI developers have. For example, while integrated GPUs may be important inputs for customers that want to run on-premises AI training, AI developers also have the option of procuring accelerated compute from cloud computing firms like Google which are deploying their own alternatives to NVIDIA's GPUs.

Second, in this market for the supply of discrete GPUs for use in datacentres, the Commission concludes that NVIDIA's holds a dominant position and significant market power. This assessment is supported by:

- **NVIDIA's very high market shares, which have remained stable over time:** The European Commission estimates that NVIDIA held an annual market share of over 70% in volume and over 90% in value in 2024. Furthermore, despite significant market growth in recent years — for instance, a threefold increase of market size between 2022 and 2023 — and strong customer demand for GPU supplier diversification, NVIDIA's market share has remained remarkably stable in both volume and value.
- **NVIDIA's substantial profit margins***:* The European Commission notes that "NVIDIA's margins are significant" and increasing: "NVIDIA's EBIT margins have been increasing over time suggesting decreasing competitive pressure from potential future entry on NVIDIA's pricing."
- **Significant barriers to entry and expansion that limit competitive pressure***:* The European Commission considers that there has been no meaningful entry in recent years in the market for discrete GPUs for use in datacentres. Furthermore, it considers that the barriers to entry are particularly high, due to the significant investment and lead time needed to design the hardware and software, and the need to provide a complete software stack to be competitive.
- **The lack of effective countervailing buyer power**: The market investigation conducted by the European Commission has shown that switching between hardware vendors is technically complex and takes time for users of discrete GPUs. This is due in particular to the strong complementarity between the hardware (GPUs) and the associated software (CUDA for NVIDIA). Indeed, switching hardware obliges customers to rewrite software, retrain staff, etc. The Commission considers that even hyperscalers cannot bypass NVIDIA's products by developing proprietary solutions due to their high and increasing needs for accelerate compute.

The European Commission's analysis of the discrete GPU market suggests that, in the short term at least, NVIDIA will continue hold significant market power. However, it is noteworthy that NVIDIA's business model appears to turn on making its chips widely available: as opposed to investment by other AI chip developers, some of whom are currently limiting direct access to their chips for in-house purposes only. Furthermore, as we explained in our earlier issue paper, NVIDIA's market position may be at risk of erosion because of a range of market developments. For example, numerous large tech firms (not limited to the hyperscalers but also including firms like IBM) are developing their own customised AI-optimised chips. OpenAI is also now designing its own chips.[49] Smaller players like SambaNova, Cerebras and Groq have also designed AI chips which have some advantages over NVIDIA's – such as on cost and power consumption.[50] Large tech firms are also providing or supporting software development kits like Amazon's Neuron, which can support switching between NVIDIA and

[49] Anna Tong, Max Cherney and Krystal Hu, 'OpenAI set to finalize first custom chip design this year', Reuters, 10 February 2025.
[50] TechInsights, AI Hardware Summit 2024.

third-party chips. Furthermore, the recent shift in emphasis away from compute-intensive initial training of models towards spending more compute resources on the inference stage has opened up new opportunities for firms to design chips which could be better than NVIDIA's at inference.[51]

Therefore, while scrutiny of business practices by providers of accelerated compute chips might be justified, those practices will need to be assessed on their merits. The case for any broad and overarching intervention into the chip-design market at this point seems premature given the growing competitive threats which NVIDIA faces. Competition authorities will need to keep a close eye on this part of the sector in order to examine whether these competitive threats – such as hyperscalers bypassing NVIDIA's offerings – do indeed effectively constrain NVIDIA's commercial behaviour.

3.3 Foreclosure of Access to Open AI Models

Today there is a significant degree of openness between different parts of the supply chain, including with many AI models being open to use by other model developers.

In part, the current market structure appears to have arisen because the largest AI developers do not seem able to prevent their models being used to develop other, competing models (and/or do not currently have incentives to do so, perhaps in order to maximise their influence and importance in the AI supply chain). This could change over time as patient investors in AI start to demand a return on investment: while some models are open today, the model developers could introduce technical or licensing use restrictions to prevent their use for creating new AI models. If this occurs, it may pose challenges to some AI providers, potentially reshaping the market. In particular, the FCA highlights the risk that AI model developers could prohibit or restrict the development of competing models.[52] However, if AI providers have incentives to keep their models open (or are unable to effectively restrict how their models are used) then openness may prove to be an enduring characteristic of the sector.

Over time, the best performing AI models may try to prevent smaller models from "free riding", for example by using the best performing model to prepare customised training data which a smaller model can use (for example to train the smaller model on chains-of-thought). Such restrictions could be developed either through technical means, enforceable contractual limitations, or new pricing models (for example higher prices for more extensive queries and outputs).

There remains an open question about whether these platforms may in the future have incentives to become closed: for example, when the market is more mature and the race for new customers ends, and firms may choose to focus on keeping existing users locked in rather than being attractive for new users. Such concerns have been expressed in particular by the CMA (AI 2023),[53] the FTC (2023),[54] and the FCA (AI 2024).[55]

[51] https://www.ft.com/content/d5c638ad-8d34-4884-a08c-a551588a9a28
[52] FCA, AI 2024.
[53] The CMA mentions the "potential concern that suppliers of FMs could initially use open-source models to develop their ecosystem of partners and developers, but later choose to transition away from open-source approach."
[54] Federal Trade Commission's blog of 29 June 2023 (Generative AI Raises Competition Concerns | Federal Trade Commission (ftc.gov)) states that "Experience has also shown how firms can use "open first, closed later" tactics in ways that undermine long-term competition. Firms that initially use open-source to draw business, establish steady streams of data, and accrue scale advantages can later close off their ecosystem to lock-in customers and lock-out competition.'"
[55] The FCA is concerned that "companies may initially adopt an open approach to generative AI to extend their market power by locking-in user companies."

There does not currently appear to be evidence of actively squeezing out or unduly influencing AI developers. On the contrary, the largest cloud computing platforms seem to be actively building their portfolios of third-party foundation models.[56] They see this as a way to improve their services and grow their market share in cloud computing, which is (unlike AI) currently a highly profitable business.[57] This makes sense given the degree of competition between AI platforms: no competition authority has suggested that any of the hyperscalers' AI platforms is approaching a position of dominance. Where there are formal partnerships between the hyperscalers and AI developers, these occasionally exclude or restrict how the AI model can be offered on other cloud platforms,[58] however, most are non-exclusive. Some – like the Microsoft/OpenAI partnership – have become less exclusive over time. As we have noted above, in many cases it appears that AI developers retain significant freedom and autonomy even when they have entered into partnerships with hyperscalers.

If a larger AI model developer were to impose restrictions in future, whether those restrictions constituted an abuse of dominance would need to be assessed at the time. It may be difficult to argue that such restrictions are anti-competitive if the market for AI remains dynamic, because that may imply no single model provider has a position of dominance. Furthermore, there might be legitimate business reasons – rather than anti-competitive reasons – for a model provider to limit the use of the model by its competitors. For example, since operating models is expensive, the larger provider may lose its incentive to continue investing and innovating if competitors are free to "piggyback" off its costs. If an open model becomes closed, this may have impacts on the market, including for AI developers who have not prepared for such a market development. Nevertheless, the purpose of competition law is not to protect smaller firms per se; and competition law does not generally demand that firms make their investments freely available to competitors to use.

In any event, the risk of large FMs becoming more closed over time appears to be encouraging many AI providers to rely more on genuinely open-source AI models. In particular, much of the AI sector is characterised by remarkably high levels of public sharing of resources such as know-how, data repositories, and coding, many shared through open-source facilities such as those of HuggingFace. The existence of tools like HuggingFace provides a degree of assurance that the fate of many smaller AI developers need not rely solely on a few of the largest cloud service providers, including the hyperscalers.

On this basis, our conclusion is that at this stage there is little case for authorities to intervene now to prevent "open early, close late" business strategies among AI developers – but that competition authorities should consider to closely monitor the sector to ensure that these strategies are not adopted in ways that are designed to reduce competition.

[56] Competition and Markets Authority, 'Cloud Infrastructure Services: Provisional decision report', 28 January 2025, para 3.454.
[57] See James Bessen, 'The New Goliaths: How Corporations Use Software to Dominate Industries, Kill Innovation, and Undermine Regulation', 2022.
[58] Competition and Markets Authority, 'Amazon.com Inc.'s partnership with Anthropic PBC: Decision on relevant merger situation', 27 September 2024.

3.4 Risks of Lock-In to Accelerated Compute Providers and Barriers to Switching

AI developers which use an accelerated compute provider may wish to move their operations to their own premises, switch providers or adopt a multi-cloud architecture (or some combination of these options). However, they may encounter various costs and constraints when doing so. If AI developers face unreasonable barriers to switching, then their accelerated compute providers may face less competitive pressure and have a degree of market power over their customers - as customers will only switch if their expected benefits from switching and/or multi-cloud exceed these costs, allowing the AI developer's incumbent provider to charge more than they otherwise could.[59] This may lead to dampened pressure to reduce costs and increase functionality and innovation in the future, if either (i) the accelerated compute provider is in a position to offer different terms to new customers than existing locked-in customers; or (ii) the market stops growing so that accelerate compute providers change their focus from attracting new customers to extracting maximum profitability from their existing customer base. However, addressing barriers to switching involves complex commercial and technical issues which, if inappropriately designed, may have the perverse consequences of reducing rather than improving competitive dynamics.

First, we cover technical barriers to switching. However, accelerated compute providers have also adopted commercial practices that may increase switching costs for customers. For example, they charge egress fees for transferring data outside of the cloud service. Consequently, overall switching costs tend to be high for customers of cloud services, raising concerns about commercial practices that could reinforce lock-in, such as discounts or credits. In this context, discounts or credits are seen as potentially problematic, as they could enable dominant service providers to attract customers with the intention of locking them in later.

3.4.1 Technical Barriers to Switching

Technical switching costs arise from the technology itself, rather than strategic decisions by providers of accelerated compute. In the previous issue paper, we argued that technical switching costs can be a by-product of introducing new features and differentiating services, which benefit customers even if they make switching difficult afterwards. The French Competition Authority gives the example of hyperscalers offering proprietary solutions, such as automated machine learning services, to companies wanting to create or fine-tune their models more easily. However, these solutions can only be deployed on the CSP's infrastructure (FCA, 2024). As one example, where an AI developer fine-tunes one model using another larger model as the 'base', AI developers do not necessarily have continued direct access to the fine-tuned model other than through the hyperscaler's platform. This could result in AI developers being unable to move their models to another accelerated compute provider for the purposes of inference. This could pose a potential constraint to some smaller cloud providers which specialise in supporting inference but not training and/or fine-tuning.

As we pointed out in the issue paper, lock-in may happen without any anti-competitive intent. For example, 'learning effects' – where a customer's staff become more familiar with a particular cloud environment – increase switching costs, because they mean that after a customer switches to a new

[59] As the industrial organization literature on switching costs explains, firms (in this case, accelerated compute providers) may compete fiercely to win over customers initially. However, this only mitigates, not eliminates, the market power conferred by switching costs.

cloud environment there will be a substantial period of 're-learning'.[60] For example, as noted in the issue paper, accelerated compute providers provide specialised and proprietary tools and software to train and fine-tune models, including Machine Learning Operations services ('MLOps') such as Google Vertex AI, Azure Machine Learning, and Amazon SageMaker, all of which streamline the AI lifecycle. Switching accelerated compute provider could result in significant costs and difficulties as employees learn how to use a different set of tools. At the same time, the fact that some hyperscalers are developing and using their own chips in competition with NVIDIA's may imply higher switching costs, since an AI developer may have to recode their software to work on a different type of chip.[61] These switching efforts may be substantive: for example it took even a relatively large and sophisticated AI firm like OpenAI at least 17 months to migrate to Microsoft Azure in December 2020.[62]

The UK's Competition and Markets Authority (CMA) considers that the technical barriers for customers who wish to switch or use multiple clouds are substantial (CMA cloud 2025). It mentions the following sources as contributing to these technical barriers to switching:

- Feature and interface differentiation in core services and ancillary services and tools, which makes services not easily substitutable and limits the ability to switch and/or multi-cloud;
- Latency, which may be a technical barrier (not necessarily intentional), especially in the context of multi-cloud;
- Skill gaps across clouds, which is related to the idea we explained above of technical switching costs existing due to learning effects.[63] IT engineers certified with one cloud provider (such as AWS) may take weeks or months to become certified with a different provider.[64]

The FCA also notes that setting up a cloud architecture with a specific provider can be time-consuming and costly. This generates switching costs when moving to a different provider (FCA, 2023). These costs may be difficult to anticipate due to a lack of transparency surrounding the technical switching process. The CMA (2025) notes that some mitigations (from cloud providers, third parties, and the customers themselves) exist but only help partially.

A number of regulatory interventions have already been passed to address technical barriers to switching:

- Article 30 of the Data Act (13 December 2023) contains various technical measures to facilitate switching. These obligations apply to all cloud service providers (DPSPs, or Data Processing Service Providers under the Data Act) as a symmetric regulation, regardless of their size or market dominance. Specifically, the Data Act introduces obligations to improve the interoperability of cloud services, thereby reducing technical switching costs. It makes a distinction between services that are "infrastructural elements" – where only basic data storage and compute is provided – and the remaining data processing services that typically

[60] This would be problematic to the extent that the customer finds the new cloud environment more efficient once re-learning has been achieved.
[61] https://www.ftc.gov/system/files/ftc_gov/pdf/p246201_aipartnerships6breport_redacted_0.pdf p 34
[62] https://www.ftc.gov/system/files/ftc_gov/pdf/p246201_aipartnerships6breport_redacted_0.pdf p 34
[63] According to the report 'Unlocking Europe's AI Potential in the Digital Decade 2025', which was conducted by Strand Partners and commissioned by AWS, 40% of European businesses cited the AI skills gap as a critical barrier to realising the benefits of AI. https://www.unlockingeuropesaipotential.com/
[64] https://assets.publishing.service.gov.uk/media/66a0afa2ab418ab055592bf2/Google_Cloud_s_response_to_the_CMA_s_technical_barriers_working_paper.pdf

include the provision of platforms and software as a service (also known as "PaaS" and "Saas"). For "infrastructural elements", the Data Act introduces an obligation of "functional equivalence", according to which cloud service providers should facilitate functional equivalence for the features that both the source and destination services offer independently (art 30.1). For the remaining services, cloud service providers should provide open interfaces to facilitate the switching process, which are publicly available and free of charge (art 30.2), based on common standards for interoperability (art 30.3). If no such standard exists (art 30.5), cloud service providers are required to offer structured, machine-readable data export at the request of the customer.

- The Data Act also includes obligations for data portability. Article 26 stipulates that cloud service providers have to:
- Inform users of the procedures available for switching and data porting. This should include information on the available porting methods and formats, as well as any known restrictions and technical limitations relating to the destination data processing services; and
- Provide users with a reference to an up-to-date online register, hosted by the data processing service provider, containing details of all the data structures and formats, as well as the relevant standards and open interoperability specifications, in which the exportable data will be available.
- The SREN Law of 21 May 2024 in France also includes data portability and interoperability obligations, stipulating that access to APIs for these purposes should be provided free of charge (Art. 28). The French law provides more specific details on how technical interfaces should be implemented than the Data Act. Specifically, it designates the French telecom regulator, ARCEP, to specify the interfaces for interoperability and data portability (Art. 29).

In general, addressing certain types of technical switching costs through ex ante regulatory intervention may be a better solution than ex post application of competition law. Technical barriers to switching appear to arise across numerous market players, and at least in part as a secondary consequence of firms innovating, rather being limited solely to firms which might have market power or with an anti-competitive intent. They are therefore not a problem best addressed through ex post competition law interventions.

However, regulatory interventions aimed at reducing technical switching costs need to be carefully calibrated and involve a trade-off: on the one hand, they may increase competitive pressure on accelerated computer providers, and therefore increase those providers' incentives to lower their prices or improve their service offerings; but on the other hand they may demand a degree of 'standardisation' across services which can reduce the potential for differentiation or innovation, or may impose costs and limitations on how an accelerate compute provider can develop a new service or functionality.

In sectors like telecoms, rules on switching were generally imposed only on markets which were already relatively standardised and/or where the perceived potential for innovation that would impact switching was relatively low. Neither of these factors is present in the provision of accelerated compute. This implies regulatory intervention would need to be carefully calibrated to reflect the right trade-offs.

Furthermore, regulatory interventions may yet prove unnecessary. While the provision of accelerated compute remains highly innovative and dynamic, there is significant potential for private sector solutions which would reduce some of the technical barriers to switching without regulatory intervention. For example, providers of accelerated compute (including the hyperscalers) have strong incentives to make switching to their services easy, and to develop and promote tools (such as Google's Kubernetes) to help customers switch or work across different cloud environments. Third party tools already exist which also help bridge cloud providers. And many providers of accelerated compute support a variety of open standards, file formats and protocols. For example, the three hyperscalers do tend to allow integration of third-party solutions, including open source ones, such as TensorFlow and PyTorch. This can allow AI developers to fine-tune open-source models in one cloud environment and then deploy the model using a different cloud provider.[65]That suggests the right interventions at this stage should focus on issues like transparency about available porting methods and formats (as in Data Act art 26) to enable developers to more easily create tools to make switching easier.

Furthermore, AI developers are likely to be relatively sophisticated customers of providers of accelerated compute and are therefore more likely than average cloud computing customers to take issues like technical barriers to switching into account when they decide on how to source accelerated compute. The case for broad regulatory intervention to reduce the technical barriers to switching between providers of accelerated compute is therefore not yet clear, even if such interventions might be justified in relation to the general cloud computing sector.

3.4.2 Egress Fees

We turn now to commercial barriers to switching. Egress fees, also known as bandwidth or data transfer fees, are charged by a cloud service provider for transferring data outside its network. This could be to send data to another provider, to the customer's own premises, or to use several providers simultaneously, in a setup known as multi-cloud.

Egress fees have been the subject of public debate and have become a concern for policymakers because (i) hyperscalers tend to pass on a significant proportion of their overall costs through these fees, despite the fact that (according to the French competition authority) most of these costs are "common to all the different services offered by a cloud service provider", and (ii) egress fees are proportionate to traffic and "customers are unable to anticipate future data traffic requirements in advance" (FCA Cloud, 2023).

For these reasons, in the previous issue paper, we argued that egress fees may increase switching costs and, therefore, lead to lock-in for AI developers. There does not appear to be agreement between the hyperscalers and competition authorities about whether egress fees tend to be cost-oriented (FCA Cloud, 2023). However, the French Competition Authority raised the following concerns (FCA cloud 2023):

- **Egress fees are difficult to anticipate; as a result, firms may not properly take them into account when comparing offers from different cloud service providers:** Egress fees are

[65] Felix Theisinger, 'Multi-Cloud: Minimizing lock-in risks for AI and Generative AI', Cognizant, December 2024, https://www.cognizant.com/de/de/insights/blog/articles/multi-cloud-ai-lock-in-risks.

proportionate to outgoing traffic. However, when signing up with a cloud service provider, it can be difficult for a company to anticipate how much data it will need to transfer when switching providers in the future. Consequently, when comparing cloud offers, companies may overlook egress fees, which will be incurred in the distant future and are difficult to predict. For this reason, competition in egress fees may not function effectively. However, it is unclear the extent to which this reasoning is likely to apply to AI developers, which are likely to be more sophisticated purchasers of accelerated computing services. More research is needed into whether or not AI developers are able to, and do, adequately account for egress fees when making purchasing decisions.

- **Charging (high) egress fees enable large cloud computing firms to offer low subscription prices and credits**. By allocating most of their costs to egress fees, large cloud computing firms are able to offer very low prices (for storage, computing, etc.) to attract customers, something that smaller CSPs cannot easily replicate. The FCA also notes that this effect is reinforced by the cloud credits granted by these large CSPs (see below).

Egress fees represent an endogenous switching cost for customers of cloud service providers. If an AI provider customer has to pay high egress fees to migrate to a different provider or to their own premises, they may eventually decide to remain with the original provider, even if a competing provider offers a higher quality service or one which better meets the AI provider's needs.

Note also that egress fees are proportionate to traffic. Therefore, if an AI provider's activity on the cloud increases over time and this leads to an increase in storage, the switching costs represented by egress fees for transferring stored data also increase. This could cause problems in particular when large CSPs attract customers with high development or innovation potential.

Egress fees make it costly to use several cloud service providers in an integrated way, if it implies recurring data transfers between them. Similarly, they can make it would be costly for a company to use tools that are not on the cloud if they require recurring data exchange with the cloud.

The CMA (cloud 2025) holds a similar view as the FCA regarding egress fees. It considers that egress fees "represent a barrier to switching and/or multi-cloud" and may "contribute to a degree of 'lock-in' where customers are less able to switch cloud provider, or use multiple cloud providers, once they have made their initial choice upon entering the market for cloud services."

The Data Act (13 December 2023) bans egress fees after an interim period (Article 29). Specifically, during an initial phase from 11 January 2024 to 12 January 2027, providers of data processing services may impose switching charges (e.g. egress fees) on customers for the switching process, provided they do not exceed the cost of the process itself. From 12 January 2027, a full ban on switching fees will be implemented: from this date onwards, providers of data processing services shall not impose any charges on customers for the switching process.

A similar obligation was introduced in Article 27 of the French SREN Law of 21 May 2024. The law stipulates that egress fees should be capped at their actual cost. Interestingly, unlike the Data Act, the French law specifies how these costs are to be determined. Specifically, it states that they are to be determined by ARCEP, the French telecoms regulator, which has considered that there are no

incremental costs of data transfer in case of switching.[66] The French law contains a sunset clause of 12 January 2027, after which the Data Act's full ban on switching fees will come into effect.

As Manganelli and Schnurr (2024) argue, reducing egress fees can decrease the financial switching costs for customers of cloud service providers. This can result in reduced lock-in and increased switching. This would benefit both AI providers and smaller providers of accelerated compute.

However, Manganelli and Schnurr (2024) also emphasise that eliminating egress fees could raise other prices (e.g.: for storage or computing), if cloud service providers cannot directly recover any of the costs of external data transfer. We could add that even a cost-oriented egress fee could increase storage and/or computing fees, if CSPs can no longer pass some of their common costs to the switching fee.

Moreover, Manganelli and Schnurr (2024) argue that a ban on egress fees could be especially challenging for small cloud service providers[67] as cross-subsidisation of data traffic costs is more difficult for them. Indeed, if small cloud service providers cannot replicate the business model of large CSPs (i.e.: offering low storage and computing prices alongside high egress fees), it is unclear how they will be able to replicate the new offers made by large CSPs after the elimination of egress fees.

Manganelli and Schnurr (2024) conclude that price regulation of egress fees could be preferable to a full ban (zero egress fees).

Gans et al. (2023)[68] discuss the impact of eliminating egress fees on consumer welfare and competition. They conclude that other fees could increase when egress fees are eliminated, which would harm customers, in particular those who transfer small amounts of data and thus would pay low egress fees. Moreover, the full elimination of all egress fess could lead to an inefficient transfer of data out of the cloud provider, increasing costs for the provider and as a result potentially its prices. Thus, this paper echoes the caveats raised by Manganelli and Schnurr (2024) that banning egress fees can inflate other prices. However, as Biglaiser et al. (2024)[69] note, their analysis relies on the assumption that the cloud service market is competitive.

Mantovani et al. (2024)[70] also study the impact of a ban on egress fees, but in a model with imperfect competition between cloud service providers. In particular, they consider a scenario in which a dominant cloud provider competes with a fringe provider. This is in line with the scenario envisaged by policymakers, in which a dominant provider could use egress fees to reinforce its position. They demonstrate that consumer surplus is greater under a ban on egress fees, but only if switching costs are not too high. Intuitively, there is also more switching with the ban, but in both directions, from the fringe to the dominant provider and vice versa. Total welfare decreases with the ban because the dominant provider is hurt more than consumers benefit. Therefore, this analysis supports the decision

[66] See also the results of Arcep's public consultation about these costs, https://www.arcep.fr/uploads/tx_gspublication/consultation-cloud-changement-fournisseur-services-architectures-tarifs-oct2024.pdf.
[67] Unlike the DMA for example, the Data Act is a symmetric regulation, which applies to all cloud service providers indifferently, regardless of their size or market share.
[68] Joshua Gans, Mikaël Hervé & Muath Masri (2023) Economic analysis of proposed regulations of cloud services in Europe, European Competition Journal, 19:3, 522-568
[69] Biglaiser, Gary, Crémer, Jacques and Andrea Mantovani (2024). The Economics of the Cloud, TSE Working Paper N° 1520.
[70] Biglaiser, Gary, de Cornière, Alexandre, Crémer, Jacques and Andrea Mantovani (2024). Should Egress Fees Be Eliminated? An Analysis of Cloud Services and Beyond, mimeo.

to ban egress fees in Europe as a means of stimulating switching to the benefit of customers of cloud service providers.

3.4.3 Cloud Credits

Cloud credits (also called rebates) are trial offers in the form of service allowances offered by a provider, granting free access for a defined period.

In the issue paper, we argued that credits should be seen generally as pro-competitive: they reflect the competition between cloud service providers to attract customers, and they benefit AI developers, in particular small ones, by lowering their costs. However, credits can be a concern if, for example, they are delivered with a requirement of exclusivity, with 'minimum spend' levels which are sufficiently high as to preclude important customers from providing other accelerated compute providers with sufficient scale to remain effective competitors, or if they are a way for providers to attract their customers in order to lock them in and later exploit them.

The FCA also considers that cloud credits can generate efficiency gains for customers of cloud service providers. They view them as particularly attractive to startups, enabling them free access to train, fine-tune, and deploy their solutions. However, the FCA also raises concerns about potential competition problems. To do so, it suggests distinguishing between two types of credits (FCA 2023):

- Free-trial (or promotional) credits, which are offered by most cloud computing service providers; and
- Support programme credits, which target in particular customers with high innovation potential, such as startups, researchers, or students, and are only offered by large CSPs (e.g. hyperscalers).

The FCA considers that free-trial credits do not pose a priori any competitive risks. However, it sees support programmes potentially more problematic, and this type of programme should therefore be analysed with particular attention.

In its 2024 report on AI, the FCA raises two main potential competition concerns (FCA 2024):

- *Cloud credits can act as a barrier to entry or expansion for smaller CSPs.*

The provision of substantial cloud credits by large cloud service providers (in particular, hyperscalers) may attract innovative customers (e.g., startups) and draw them away from smaller players that cannot replicate such credits. Indeed, the French Competition Authority considers these cloud credits to be unusually large compared to commercial practices observed in other industries (for example, trial offers), calling into question their profitability or rationality. In particular, the FCA considers that "[t]he ability of a competing provider "as efficient" as the companies offering cloud credits to enter or remain profitable in this (or a related) market is questionable, given the amount of cloud credits offered by hyperscalers and the relatively major investments required for smaller providers."

The FCA concludes that cloud credits may represent a barrier to entry or expansion for (as efficient) smaller cloud service providers. The mechanism that the FCA has in mind is one of demand reduction: by capturing the demand from innovative customers (e.g., startups), the large CSPs can reduce the demand that their competitors can address, thereby limiting entry and expansion prospects. This is

particularly problematic for smaller competitors as these startups could become major customers in the future.

However, one could argue that this concern is only valid to the extent that smaller CSPs cannot win back customers attracted by large CSPs' generous cloud credits. Therefore, this requires a lock-in effect, which is precisely the second concern of the FCA.

- *Cloud credits can be used as a customer lock-in strategy.*

The FCA is concerned that cloud credits may contribute to customer lock-in in due to the significant technical and financial barriers to migration (e.g., egress fees) that we have covered earlier. The FCA also notes that the amount of credits can be increased if the beneficiary start-ups use the CSP's proprietary products. This could potentially increase switching costs further and reinforce lock-in.

Although cloud credits are a cause for concern, there are no specific regulatory provisions regarding them at the European level. In France, Article 26 of the SREN Law (21 May 2024) states that 'cloud credits' must not exceed one year and may not include an exclusivity clause.

Cloud Computing Services (CCSs) are one of the Core Platforms Services listed in the DMA (Article 2(2)). Manganelli and Schnurr (2024) argue that designating a CCS provider as a gatekeeper under Article 3.8 of the DMA would be "very burdensome". The French Competition Authority further notes that even if it were the case, the DMA does not very well address the type of competition problems raised, that is, situations of lock-in through technical or commercial practices. Gatekeepers' obligations would therefore need to be updated, which would be a complex and long process (FCA 2023).

As discussed in the issue paper, and as Manganelli and Schnurr (2024) also argue, the regulation of cloud credits should be approached with caution. While they may raise competition concerns, they also indicate a vibrant competitive market and offer significant benefits to innovative customers. Manganelli and Schnurr (2024) also argue that there is a trade-off between the short-term benefit of cheap access to the cloud for customers with limited financial resources but strong innovative potential, and the long-term harm that could result from lock-in.

In conclusion, we recommend continuing to monitor the provision of discounts by accelerated compute providers to avoid a 'customer lock-in' effect, bearing in mind that discount schemes often indicate a vibrant competitive market and offer significant benefits to innovative customers.

3.5 Merger and Partnerships

Two types of operation may raise competition concerns: (i) mergers between AI firms, and (ii) partnerships between AI firms and major cloud service providers.

Regarding mergers between AI firms, currently, authorities seem right by not intervening, since AI firms seem to have choices and significant freedom. However, if the market becomes much more consolidated, deeper scrutiny could be justified.

Authorities will need to be careful when assessing consolidation: for example, allowing AI firms to consolidate and continue to provide value as part of a firm's broader digital ecosystem may be a better

option than leaving them with no option but to close shop. However, it is too early to assume that significant market consolidation – for example the case for accepting only one or two providers of FMs – would be inevitable. Authorities should therefore avoid jumping to the conclusion that they may need to step in quickly to avoid either market "tipping" towards one or two providers.

Regarding partnerships between AI firms and major cloud service providers, in general, partnerships between AI firms and the largest cloud computing providers have improved competition and benefited AI developers, in particular European ones. These partnerships often involve the cloud computing provider taking a financial stake in the AI developer in return for concessions about the AI developer's access to accelerated compute, along with an agreement to host the AI model on the hyperscaler's platforms, or to integrate the AI developer's services into the hyperscaler's existing services. These partnerships differ greatly, and are not always transparent, but in at least some of these deals the AI developer has agreed to use one of the large cloud computing providers as their 'primary' or 'preferred' provider.

The French Competition Authority (FCA, 2024) agrees that partnerships between AI firms and major cloud service providers are not problematic per se. They provide startups with access to the financial and technological resources of major companies, enabling faster development and spurring innovation. Conversely, the major companies gain access to innovative products and technologies to incorporate into their own offerings. Similarly, the CMA (2024) states that these partnerships may be "an essential ingredient for the success of independent developers."

However, the two competition authorities have also expressed concern that these partnerships could be exploited by dominant firms for anti-competitive purposes.

Specifically, the French Competition Authority has identified the following potential risks to competition (FCA, 2024):

- Partnerships can reduce competition if the target company and the acquirer are competitors;
- If partners operate at different stages of the AI value chain, this can have negative vertical effects, particularly if the partnerships include exclusivity clauses;
- Partnerships can increase market transparency, which can be harmful if the acquirer gains access to information about competitors using the target's technology;
- Risks to competition may arise if a major company invests in different targets that are competitors, or if competing major firms invest in the same company.

According to the CMA, the following factors may raise greater concern for a partnership:

- If a partner has market power upstream (over a critical input) or downstream (where models are deployed);
- If some of the models involved are "highly capable" and have significant future potential;
- If one party gains influence over another's development of models and/or deployment downstream.

The French Competition Authority argues that these concerns can be addressed ex ante through merger control if the partnership gives the acquirer de facto control of the target and the transaction

exceeds the EU and national notification thresholds. The concerns can also be assessed ex post through competition law, for example on the basis of an abuse of dominant position (FCA, 2024).

The terms of partnership agreements may also create specific concerns about competition. For example, under some AI developer/hyperscaler partnerships, the AI developers have agreed to relatively high expenditure on the cloud computing services of their hyperscaler partner. However, whether these create de facto exclusivity obligations needs to be considered by reference to the particular commercial deal. For example, Anthropic has agreed to spend $4 billion with Amazon Web Services over five years – but it also has a similar deal with Google.[71] Hyperscalers argue that a minimum spend commitment provides them with a certain degree of revenue predictability, which is crucial for infrastructure planning and covering their investment costs. However, minimum spend commitments may also have an impact on AI developers' ability to switch.

In conclusion, partnerships between AI firms and major cloud service providers have strong efficiency effects and should not be blocked per se. However, given the high stakes and the substantial market power of the major firms involved, the deals made between major CSPs and AI startups should be monitored and scrutinised.

In fact, Article 14 of the DMA requires gatekeepers to inform the European Commission of any intended merger involving other digital service providers or operators in the digital sector, even if it is not subject to notification under the EU Merger Regulation or national merger laws. Therefore, the NCAs have the ability to implement such close monitoring.

3.6 Self-Preferencing, Tying, Leveraging and Vertical Integration

As we have noted in the issue paper, a further potential set of concerns is associated with the vertical integration of some large players in the cloud and AI sectors – for example, some players may design their own accelerator chips; provide accelerated computing services; operate platforms through which AI models are accessible to end users; operate their own AI models; and provide downstream services such as search engines, operating systems, browsers, online marketplaces, and social media networks with which AI services can be integrated or accessed. Furthermore, at some points in the value chain these players may have dominant positions.

Although vertical integration can deliver significant efficiency gains, there is also a well-understood set of concerns which can arise from such vertical integration, including:

- Self-preferencing or discrimination – for example, where a vertically integrated player offers services to itself and to third parties, it may provide preferential access to its own (or to preferred third-party) services.[72] For example, its own AI model might receive prominence on its AI model platform, or it might only give certain preferred AI developers access to the vertically integrated provider's most cutting-edge models.

[71] https://www.ftc.gov/system/files/ftc_gov/pdf/p246201_aipartnerships6breport_redacted_0.pdf p 32

[72] See also the French law to secure digital space (Article 26 of LOI n° 2024-449 du 21 mai 2024 visant à sécuriser et à réguler l'espace numérique) which prohibits self-preferencing practices of a cloud provider offering also software services (i.e., when the cloud service provider applies more favourable software licensing terms when the software is used with its own cloud services compared to when the software is used with a third-party cloud services provider. https://www.legifrance.gouv.fr/jorf/article_jo/JORFARTI000049563594

- Tying and bundling – where two services can only be purchased together or one is cross-subsidised. This can be problematic where one service is a 'must have' service and is used to drive usage of a service in a competitive market, such as the vertically integrated firm's AI services.
- Product integration – where an AI service is integrated into the firm's existing services, without providing equal opportunities for third party AI services to be similarly integrated.

Caution is required, however, before jumping to the conclusion that any of these practices are inherently harmful to competition. Often, each of these practices will have a pro-competitive rationale, even if they damage the interests of smaller AI developers. For example:

- Self-preferencing is not usually harmful to competition where there exists a range of effective alternatives for the service in which the self-preferencing is occurring (e.g. the marketplace or AI model platform). In these cases, AI developers will still have a range of ways to make their service accessible to customers. Furthermore, where there is effective competition between platforms, this can drive openness – because platforms will compete to attract the largest number of AI models to their platform. This appears to be the dynamic we see in the market today, where each of the hyperscalers' AI model platforms offers a wide range of third-party models, and there are competing independent AI model platforms such as Hugging Face.
- Tying and bundling can be essential ways to persuade customers to try a new feature, and a particularly valuable way of pushing such a feature which is novel and not well understood by customers, so they have a chance to experiment with it and identify its value.
- Integration of AI into existing services is often pursued by small and large tech firms as a way of delivering new product improvements and convenience to users. In some cases, integration may be essential to help the hyperscaler innovate – for example, by building a strong enough customer base to justify making an otherwise too-risky investment. Furthermore, it will not always be clear whether there is even (commercially or technically) potential for third-party competition to provide an AI feature deeply embedded with a large firm's existing service: arguably such features are simply an extension of the existing service (as when smartphones began integrating cameras) rather than a standalone market (as when smartphones began allowing third-party apps). There remain many other ways for AI developers to be successful.

Competition authorities will need to assess such justifications and their rationale extremely carefully, but it is commonly accepted that practices like self-preferencing should not be assumed to be anti-competitive.[73] This is why – except in some narrow cases such as the Net Neutrality Regulation – regulation does not typically impose blanket requirements on firms to treat all customers equally. While business decisions may harm some independent AI developers, they will not necessarily harm levels of competition, provided there remains a thriving ecosystem of AI firms.

This suggests that concerns regarding vertical integration are best addressed through the ex post application of competition law, rather than through ex ante rules which do not take full regard of the circumstances.

[73] Still, the DMA considers self-preferencing as anticompetitive for gatekeepers providing some services (Art. 6(5)).

The Digital Markets Act is the primary EU-level regulatory measure which aims to address the potential concerns with vertical integration in the digital sector. The DMA's rules include, for example, requirements to allow the uninstallation of software applications and the ability to easily change defaults and enable switching. These may reduce the ability of firms to engage in tying and bundling. They also include a general prohibition on self-preferencing in ranking and non-discrimination in access to app stores, search engines and social networks.

The DMA anticipates that cloud services could be regulated but none currently are: even if they were, the law contains few provisions which would effectively address any potential concern regarding vertical integration between cloud and AI services. For example, the requirement to provide non-discriminatory access does not apply to a gatekeeper's cloud computing services. And it is not clear how the rule against self-preferencing would apply to AI model platforms which are usually used solely by businesses rather than directly intermediating between businesses and end-users.

The DMA also does not specifically address AI models, though it may regulate AI to the extent it is integrated into and forms part of an existing core platform service such as a browser or operating system. Furthermore, as discussed by Bostoen and Kramer, the DMA might also regulate certain uses of AI, such as AI agents, under the category of 'virtual assistants'.[74] This may prevent AI agents from unduly self-preferencing by pointing users to a gatekeeper's own service. Both ways in which the DMA apply could provide a partial response to concerns about vertical integration between AI models and other digital services.

France, Germany, and the Netherlands are preparing a joint proposal urging the European Commission to extend the DMA's scope to more comprehensively address cloud and AI. There may be merit in reconsidering how the DMA applies to cloud and AI. However, any broadening of the scope of the DMA ought to be done cautiously. In general, since the DMA's rules and prohibitions do not take the full set of circumstances into account, as competition law does, the efficiency of the DMA seems more relevant to digital markets where entrenchment is clear and both static and dynamic competition remains at suboptimal levels (for example, because the degree of innovation is relatively low). That does not appear to be the case for the provision of AI models or accelerated compute since in both cases levels of innovation and investment are relatively high and no clear "winner" has emerged. Any amendments to the DMA to more effectively regulate cloud and AI therefore should be considered primarily to address the risk that cloud and AI services are used to further entrench existing incumbents, rather than on the assumption that large providers of cloud or AI services necessarily warrant additional regulation.

3.7 Other Factors Impacting Competition

Having considered the way potential market dynamics or competition issues could impact whether Europe has access to low-cost, high-quality and innovative AI, our conclusions are largely that the market is working well and that where there are potential competition issues in many cases these may not materialise or may be addressed with only targeted interventions. Largely that is because – as we explained in the issue paper – other players in the AI and cloud sectors have significant incentives and ability to promote a thriving competition.

[74] https://cerre.eu/wp-content/uploads/2024/11/CERRE_Final-Issue-Paper_AI-Agents-and-Ecosystems-Contestability.pdf p 13.

That means one of the more significant constraints on the economic benefits the sector could bring to the EU will arise from regulatory barriers that constrain firms, especially smaller and less well-resourced ones, from growing in Europe. To ensure the successful growth of the sector in Europe, the essential inputs to AI need to be widely accessible and regulatory barriers preventing the growth of the sector must be tackled. These may include:

- **Access to data** – as we have explained above, the EU has made good progress in recent years in ensuring open accessibility of public data, and working to improve trade in the market for non-personal data in Europe. Policies governing the use by AI developers of IP-protected material should also take into account the significant economic benefits of the AI sector, and the possibility that significant restrictions on the use of IP-protected material would benefit only the largest players in the market, who are likely to be in a better position to negotiate terms of access. Targeted antitrust measures, or use of regulatory tools like the DMA, to address where data holders may be constraining access to critical data will also be important.
- **Access to data centres** – this infrastructure is essential to host the data storage and processing equipment needed for AI models to be trained and run inference. However, the planning process for building new data centres, and the need for electricity grid investments, often means the infrastructure to support provision of more accelerated compute is lacking. In 2024, about $465 billion was spent on data centre investments and that figure is expected to grow significantly. In practice, regulatory or other barriers to constructing data centres could have the effect of both hindering the growth of AI and also ossifying the existing market structure, by advantaging firms which have already made investments in data centres over new entrants.
- **Proportionate regulatory obligations**: laws like the AI Act, the GDPR, the DMA and the Digital Services Act together form a complex set of rules, some of which impact the AI sector directly and others indirectly. The complexity risks posing a particular barrier to smaller "challenger" firms which are less able than global monoliths to navigate the European regulatory environment – and to ensure that rules like the GDPR are not misapplied by firms with advantages in the market to exclude potential competitors. The Commission's current work programme which aims to ensure laws like the AI Act are implemented in a proportionate way, together with the simplification exercise aiming at ensuring consistency and coherence between the different parts of the EU digital acquis, will be important measures to increase competition.
- **Use of interventions to support AI and cloud "sovereignty":** We note that there is growing pressure in Europe to support and encourage more "sovereign" providers of cloud computing, and to enable European firms to have more of a role in the AI value chain. The Commission has, for example, mooted an EU Cloud and AI Development Act, which might set minimum standards for cloud computing services in Europe, potentially providing advantages to local European firms.[75] Regardless of whether or not initiatives to support digital sovereignty in Europe are necessary, it is important these initiatives promote rather than undermine competition. They may achieve positive competitive impacts if they foster and support the entry and growth of new local players in the sector, for example by ensuring customers value supply chain diversity. However, if they are primarily focused on excluding market

[75] European Commission, 'A Competitiveness Compass for the EU', 29 January 2025.

opportunities for existing players – for example by adopting an approach to competition policy which is inherently hostile to foreign firms or those which already have a presence in other digital markets – then there is a significant risk of decreasing competition.

4. Conclusion

European policy makers have identified AI as a technology which could help boost Europe's sluggish economic growth, and which offer Europe an opportunity to return to the forefront of technological leadership. To achieve this, AI must be widely available and with competitive pressure keeping prices low, quality high and forcing firms in the sector to continue to invest and innovate.

As we have explained in our earlier issue paper, the sector – along with the provision of accelerated compute on which AI developers rely – is performing well in Europe. Both AI and provision of accelerated compute are growing at a fast pace and a range of challenger and incumbent companies are all attracting significant investment. The AI sector looks competitive even on traditional metrics of market power and is even more competitive when metrics of dynamic competition are examined. While a few large firms are currently very prominent in the provision of accelerated compute, and one firm is very prominent in the design of AI accelerator chips, a dynamic competitive analysis suggests that both these markets are highly innovative and the firms that operate in these markets are under pressure to deliver good results for the AI sector as a whole.

The concentrated nature of some parts of the AI value chain, the strong position of the hyperscalers across multiple parts of the value chain, create the possibility of future harm to competition. However, authorities should be cautious about reacting to this possibility prematurely. One possible reason for anxiety to move quickly and pre-emptively is that competition authorities clearly moved too late, and acted too slowly, in some more mature digital markets where one or two players have now dominated for many years. However, several factors – such as the decreasing costs of developing AI, the growing interdependence of AI models, and the prominence and success of relatively open AI models – suggest that **competition authorities should not assume market consolidation is likely or that 'winner takes all' outcomes will occur in the AI sector**. As we explained in the issue paper, other players in the AI and cloud sectors have significant incentives and ability to promote a thriving competition.

As we have explained, most competition problems identified by authorities to date relate to how the market *could* develop – rather than conduct which has already taken place. However, since we foresee a stronger possibility that the sector will remain highly competitive than in other digital markets, in assessing whether to act pre-emptively, authorities should give adequate weight to the risk that intervention proves unnecessary, and the potential unintended consequences of such interventions. Where intervention is warranted, these factors suggest the interventions should be targeted and designed to avoid unnecessarily dampening the current high levels of experimentation, innovation and investment.

In particular we recommend that policy makers:

- **Consider carefully complaints that firms are withholding or imposing unreasonable terms of access to data for AI developers**. Policy makers should at the same time ensure rules on mandatory data sharing, like the GDPR, the Data Act and the DMA, are to the extent possible applied in ways which protect firms' incentives to invest in collecting and innovating in the use of data.

- **Consider promoting an industry-led initiative to ensure users of AI services can 'port' their usage history from one model to a competing service**, in order to protect users' choices and help avoid creating 'lock in effects' or 'winner takes all' outcomes.
- **Scrutinise carefully business practices among firms providing chips for accelerated compute**. However, those practices should be reviewed on their merits – and the case for any broad and overarching intervention into the chip-design market at this point seems premature given the growing potential for competition in this space.
- **Keep a 'watching brief' on the possibility that developers of open AI models will seek to close off access to these models in future**, while keeping in mind that there is currently limited evidence of firms adopting an 'open early, close late' approach. If a larger AI model developer were to close access to a previously open model or resource, competition authorities should take into account whether there are legitimate business reasons for such conduct (such as protecting the AI developer's incentives to invest and innovate).
- **Carefully consider potential interventions to protect AI developers' ability to switch between providers of accelerated compute**. However, in addressing technical barriers to switching, authorities must balance the dangers that interventions will limit innovation and differentiation in the sector; take into account the possibility of private sector technical solutions to help overcome some of the switching barriers that exist; and consider the specific circumstances of AI developers, which are likely to be fairly sophisticated customers.
- **Apply the ban on egress fees set out in the Data Act, and continue monitoring the provision of discounts by providers of accelerated compute to avoid a 'customer lock-in' effect**, while keeping in mind that discount schemes often indicate a vibrant competitive market and offer significant benefits to innovative customers.
- **Monitor and scrutinise partnerships between major cloud service providers and AI startups.** However, these partnerships have strong efficiency effects and should not be blocked per se. Policy makers should pay particular attention to exclusivity clauses in partnership agreements.
- **Apply caution when scrutinising practices associated with vertical integration** (such as self-preferencing, tying and leveraging) since these may often indicate pro-competitive conduct, such as helping consumers discover new services and helping high-cost, high-risk business ideas get scale. Competition authorities will need to review the full context of the conduct and its impacts to determine whether such conduct is anti-competitive, and ex ante regulation like the Digital Markets Act might not be the best tool to tackle any perceived problems with vertically integrated players.
- **Ensure that regulatory barriers such as the AI Act are applied proportionately** in order to further growth and competition in the AI and cloud sectors, and continue the European Commission's current regulatory simplification exercise to help minimise regulatory costs which might dissuade entrepreneurs and smaller firms from entering the AI sector, or businesses across Europe generally from taking up AI.

cerre Centre on Regulation in Europe

Avenue Louise 475 (box 10)
1050 Brussels, Belgium
+32 2 230 83 60
info@cerre.eu
www.cerre.eu

Centre on Regulation in Europe (CERRE)
CERRE Think Tank
CERRE Think Tank