COMPETITION IN ARTIFICIAL INTELLIGENCE INFRASTRUCTURE

OECD Roundtables on Competition Policy Papers, No. 330

OECD

Competition in artificial intelligence infrastructure

Disclaimers

This document was originally released on O.N.E. as an unclassified document under the reference code: DAF/COMP(2025)11.

This work is published under the responsibility of the Secretary-General of the OECD. The opinions expressed and arguments employed herein do not necessarily reflect the official views of the Member countries of the OECD.

This document, as well as any data and map included herein, are without prejudice to the status of or sovereignty over any territory, to the delimitation of international frontiers and boundaries and to the name of any territory, city or area.

Photo credits: © shironosov / iStock - Getty Images Plus

© OECD 2025.

Attribution 4.0 International (CC BY 4.0)

This work is made available under the Creative Commons Attribution 4.0 International licence. By using this work, you accept to be bound by the terms of this licence (https://creativecommons.org/licenses/by/4.0/).

Attribution – you must cite the work.

Translations – you must cite the original work, identify changes to the original and add the following text: *In the event of any discrepancy between the original work and the translation, only the text of the original work should be considered valid.*

Adaptations – you must cite the original work and add the following text: *This is an adaptation of an original work by the OECD. The opinions expressed and arguments employed in this adaptation should not be reported as representing the official views of the OECD or of its Member countries.*

Third-party material – the licence does not apply to third-party material in the work. If using such material, you are responsible for obtaining permission from the third party and for any claims of infringement.

You must not use the OECD logo, visual identity or cover image without express permission or suggest the OECD endorses your use of the work.

Any dispute arising under this licence shall be settled by arbitration in accordance with the Permanent Court of Arbitration (PCA) Arbitration Rules 2012. The seat of arbitration shall be Paris (France). The number of arbitrators shall be one.

Abstract

Artificial intelligence (AI) is a rapidly developing general purpose technology of global importance. This paper examines the potential competition issues arising from the physical infrastructure underpinning AI development, particularly computing resources such as advanced chips. It explores the structure of AI supply chains, identifies key players and highlights characteristics that make these markets susceptible to competition issues. The paper also considers the tools available to competition authorities and assesses their relevance and limitations in addressing emerging challenges. Finally, it discusses how effective competition in AI infrastructure markets relies on competition authorities closely monitoring market developments and responding with a balanced mix of enforcement and advocacy tools.

Key words: Competition, Industrial policy, Infrastructure, Innovation, Antitrust, Artificial Intelligence, AI, Semiconductors, Chips, Data centres.

JEL CODES: K21, L40, L41, L42, L63, O33, O34

This paper is part of the series "OECD Roundtables on Competition Policy Papers", https://doi.org/10.1787/20758677.

Acknowledgements

This note was prepared by Greg Jackson and Connor Hogg of the OECD Competition Division. Helpful comments and review were provided by Ori Schwartz, Antonio Capobianco, Alessandra Tonazzi and Richard May, also of the OECD Competition Division. This paper is part of the OECD Horizontal Project on ‘Thriving with AI: Empowering Economies, Societies and Citizens.’ It has benefited from contributions from colleagues in the OECD Directorate for Science Technology and Innovation, particularly Angela Attrey, Celine Caira, Lauren Crean, Alexia González Fanfalone, Jan-Peter Kleinhans and Filipe Silva. The note was prepared to serve as background material for discussions on “Competition and Artificial Intelligence” taking place at the December 2025 session of the OECD Competition Committee.

Table of contents

FIGURES

TABLES

BOXES

Executive summary

Artificial intelligence (AI) is rapidly emerging as a general-purpose technology with the potential to transform the economy. Its development depends heavily on advances in computing hardware (compute), which enables the training of increasingly complex foundation models and their deployment at scale.

The AI compute infrastructure supply chain is multi-layered and spans several interrelated markets. At its core are the chip designers and manufacturers, who produce the essential hardware powering AI data centres. These firms rely on specialised suppliers of materials and equipment, which are critical to the chip production process. The chips are then integrated into fully functioning server racks and deployed by data centre operators. These operators also rely on the energy and networking infrastructure required to power and interconnect the individual servers within a data centre, as well as link multiple data centres together. Once operational, cloud providers sell access to this compute power to AI developers and modelers, enabling them to train, and deploy increasingly sophisticated AI systems.

Despite covering a range of markets there are several common features which have important implications for competition:

- rapid innovation and dynamic market evolution with high levels of research and development
- high concentration and barriers to entry meaning at each level there are often only a few suppliers
- vertical and conglomerate integration is increasingly found across the AI compute supply chain
- increasing levels of crossholdings and investment partnerships
- high levels of state intervention including public investment and trade barriers
- high levels of demand, often outstripping supply.

Competition enforcement action across the OECD Members has been limited so far, even if some investigations are currently underway. Potential competition concerns in the AI compute sector are varied and evolving. For example, firms with significant market power may engage in exclusionary practices, such as bundling products in ways that disadvantage competitors. Market power may be also expanded through less traditional means, including acqui-hires and strategic partnerships that fall below merger notification thresholds. In addition, as certain layers of the AI stack mature and become commoditised, the risk of collusion may increase, particularly where fewer players dominate. Beyond conduct the strategic significance of AI to national competitiveness and security has led to increasing levels of government intervention across the supply chain. With industrial policy initiatives potentially shaping market dynamics in areas such as chip fabrication, data centre development and access to compute.

Competition authorities will likely need to respond using the wide range of tools available. This may start with building technical expertise across the AI compute stack through research and recruitment to enable timely and effective interventions. A particular challenge will be ensuring merger control regimes are flexible enough to capture acquisitions of nascent competitors, and scrutinising conglomerate mergers for expansion risks. This includes examining cross-layer partnerships to ensure they are not used to foreclose rivals or distort competition. Competition agencies may also consider advocating for pro-competitive policies, such as supporting public investment in open-source technologies, public compute resources, and infrastructure to help overcome economies of scale and enable competition to flourish.

1 Introduction

Artificial intelligence is a rapidly developing technology that has dominated public consciousness among consumers, firms and policymakers in recent years. AI is now being discussed as the next general-purpose technology with the transformative effects permeating throughout society being likened to developments like the steam engine and electricity (Calvino, Haerle and Liu, 2025[1]). [1]

Training and deploying generative AI services requires an enormous amount of computing resources (referred to as AI "compute" (OECD, 2023[2])), as well as a strong foundation of digital connectivity infrastructure and reliable energy sources. There has been little discussion to date at the intersection of competition policy and AI compute infrastructure. This paper furthers the discussion by focusing on competition in the physical infrastructure that underpin AI services, markets understood to be both highly complex and often highly concentrated (Gambacorta and Shreeti, 2025[3]). This includes AI accelerator chips and other compute hardware (i.e. the technologies needed to train, finetune and deploy AI models), as well as a more limited discussion on the power and cooling, networking and storage infrastructure. Other important non-physical inputs to AI including data provision, model development, finance and skills are outside the scope of this paper.

Given the economic significance and strong public interest in the AI sector, it is essential that competition policy and enforcement mechanisms are equipped to monitor and respond effectively to emerging risks in the AI supply chain. Ensuring that competition in these markets function well is critical not only for protecting consumer welfare, but also for sustaining innovation and long-term market dynamism. At the same time, governments are increasingly intervening in the AI supply chain to encourage investment and promote broader public interest objectives. This creates a need for competition authorities to engage with government and advocate to ensure such interventions consider competition policy. The focus of this background paper is on exploring the competition issues in AI infrastructure and the supply chains which underpin it. This paper is structured in three parts:

- Section 2 examines the supply chain, describing the relevant technologies and key players.
- Section 3 draws out some common features and highlights some of the potential implications for competition.
- Section 4 highlights current competition policy responses and tools as well as some of the challenges authorities face in responding to this rapidly evolving sector.

This paper is part of the OECD Horizontal Project on 'Thriving with AI: Empowering Economies, Societies and Citizens.' It builds on previous OECD policy papers in relation to AI and competition, namely those on AI, data and competition (OECD, 2024[4]) and Competition in the Provision of Cloud Computing services (OECD, 2025[5]). These OECD background papers on competition also complement a broad range of work ongoing in the OECD seeking to better understand the implications of AI. These include: 'The effects of generative AI on productivity, innovation and entrepreneurship' (Calvino, Reijerink and Samek, 2025[6]), 'The macroeconomic productivity gains from Artificial Intelligence in G7 economies' (Filippucci et al., 2025[7]) as well as papers on the value chain including: 'Mapping the semiconductor value chain' (OECD, 2025[8]) and 'Vulnerabilities in the semiconductor supply chain' (Haramboure et al., 2023[9]).

2 Overview of the AI supply chain

Artificial intelligence (AI) systems are defined in the OECD Recommendation on AI as (OECD, 2019[10]):

> *a machine-based system that, for explicit or implicit objectives, infers, from the input it receives, how to generate outputs such as predictions, content, recommendations, or decisions that can influence physical or virtual environments. Different AI systems vary in their levels of autonomy and adaptiveness after deployment.*

Generative AI systems are now commonly known and available on hundreds of millions of consumer devices at the touch of a button, but for most, the underlying technology remains a mystery. The training and deployment of an AI system relies on a highly complex and capital-intensive global supply chain. For the first half of 2025, capital expenditure relating to AI infrastructure in the United States was estimated to contribute around 1.1-1.2% to GDP growth, a greater portion than spending on Internet infrastructure during the adoption of the Internet of the late 1990s to early 2000s (Mims, 2025[11]; Kedrosky, 2025[12]; Aliaga, 2025[13]).

It is important for competition policymakers to recognise that while recent digital innovations may not on the surface resemble 19th century technologies such as railroads and telephones, digital innovations are nonetheless dependent on a supply chain of physical infrastructure to function (Rahman, 2018[14]).

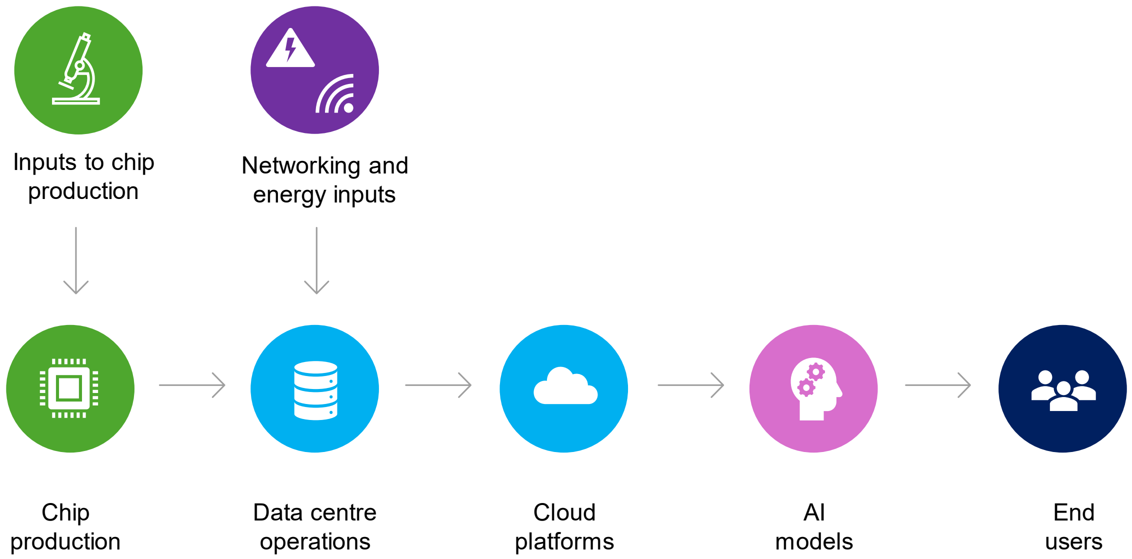
This section of the paper seeks to explain the layers of physical infrastructure that exist in the AI supply chain (in industry parlance, the technology “stack”) and outlines the current market features which may have implications for competition. This includes a strong focus on the supply chain for *chips* (the common name for microchips or integrated circuits) that are necessary to train and deploy AI systems. The chapter will also examine other digital infrastructure: namely the data centre and cloud computing layer of the AI supply chain, including the power and cooling systems needed to operate data centres, as well as the broadband network infrastructure that underpins the AI compute ecosystem.

2.1. Overview of the AI supply chain and the importance of compute

The AI applications which are found on consumer devices today rely on complex models which have been trained on vast databases. Just a few years ago these AI systems struggled with even basic tasks. Now they can solve many complex problems, write software or create realistic images and videos (Samborska, 2025[15]). A lot of the recent improvements in AI capabilities have come from scaling up AI systems.[2] These developments have required enormous improvements in computing power (Samborska, 2025[15]). The compute supply chain is therefore of crucial importance to the development of AI.

The provision of computing power to AI model developers and model users is typically provided by massive data centres and operated by cloud providers. These AI hardware operators rely on a range of key inputs including advanced compute hardware, networking (Mckinsey, 2025[16]), and energy (Chen, 2025[17]). Figure 1 shows a simplified supply chain. This chapter will begin by explaining the importance of integrated circuits (chips) to AI and explain the key parts of the chip supply chain (highlighted in green). The remainder of the chapter will discuss the other inputs involved in data centres which bring together the computing power to run AI models (purple and blue).

Figure 1. AI infrastructure supply chain diagram



Note: This diagram is a significant simplification of the key steps in the AI infrastructure supply chain with key inputs for competition identified by this paper. It does not survey the other inputs to AI systems, such as data, algorithms, models and know-how.
Source: Adapted from Pilz (2023[18]) An Assessment of Data Center Infrastructure's Role in AI Governance, https://www.konstantinpilz.com/data-centers/assessment.

2.2. Importance of Integrated circuits (chips)

Fundamentally, computer processing relies on integrated circuits (commonly referred to as microchips or chips) which contain transistors (switches) that enable a computer to make calculations (Intel, 2024[19]). Developments in AI are therefore inherently connected to innovations in chips (LeCun, Bengio and Hinton, 2015[20]).

The more transistors a chip contains, the greater its computational capacity. To increase performance, the chip industry has focused on packing more transistors onto each chip and shrinking their size. From 1950 to 2010, computing power roughly doubled every two years, a trend known as Moore's Law, named after Intel's founder who predicted this exponential growth (Schaller, 1997[21]; LeCun, Bengio and Hinton, 2015[20]). Since 2010, the growth in compute has accelerated dramatically, now doubling approximately every six months (Samborska, 2025[15]).

Improvements in chips helped AI move from simple rule-based systems to much more complex approaches (such as deep learning, inference and neural networks that underpin current AI developments). As AI has advanced to its current state, general purpose computing chips (such as Central Processing Units (CPUs)) which were central to previous computing developments became impractical. Instead, chips which could conduct simultaneous computations rose to the fore. Box 1 explains the key types of chips used in AI.

Box 1. Key types of chips used in AI explained

There are three main types of chips that are used in the training and deployment of AI systems, referred to as AI accelerator chips:

- **Graphical Processing Units (GPUs)** – Originally designed to process images for video game graphics, these are the primary chips used for AI training. By completing computing tasks in parallel rather than sequentially, GPUs are well optimised for the type of computations needed to train AI even though they are still designed for general purpose computing needs. They are by far the most widely available and most used chips in the AI sector.
- **Field-Programmable Gate Arrays (FPGAs)** – FPGAs are chips that can be programmed and reprogrammed to perform specific tasks after they have been manufactured, unlike GPUs which have fixed hardware structures. FPGAs are well-suited to AI computation because they can be tailored to the specific needs of AI applications. FPGAs have the advantage of being available "off the shelf" to be configured immediately but may be less cost effective and efficient in the long run compared to designing a chip for a specific need.
- **Application-Specific Integrated Circuits (ASICs)** – ASICs are chips that are custom designed to perform a single function with high efficiency and speed.* ASICs are typically hard wired to suit a specific AI algorithm. ASICs are very hard to commercialise given the very high upfront design and manufacture costs, and the lower demand given their ability to function for only a specific application. Currently only the largest tech firms active in AI development have deployed ASICs into their AI offerings, such as Google's Tensor Processing Units (TPUs), Google's own AI-specialised chips.

Note: * It is important to note that the distinction between ASICs and GPUs is often blurry as GPUs were themselves a type of ASIC when originally designed for processing graphics but are used for a wider range of tasks

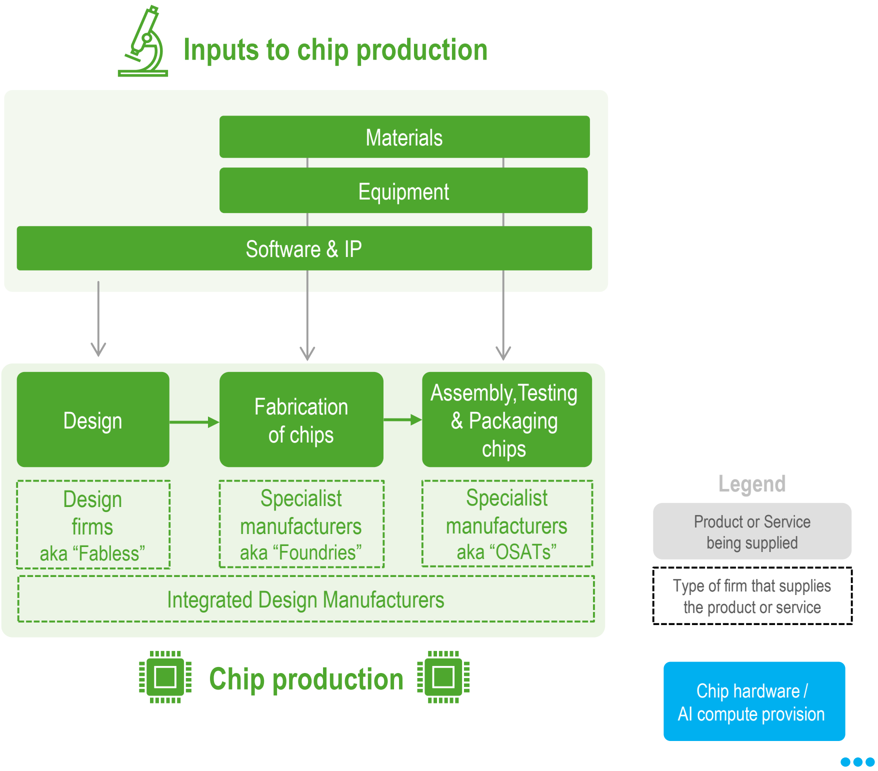
Source: Center for Security and Emerging Technology (2020[22]), AI Chips: What They Are and Why They Matter, Center for Security and Emerging Technology, https://doi.org/10.51593/20190014.

The following sections run through the key layers of the chip production process, starting with a brief overview of the value chain before discussing the different key firms operating at each level. There are multiple elements which come together in data centres to form the compute backbone for AI, this section is not exhaustive but seeks to highlight some of the key components in the value chain.

2.3. Overview of the chip supply chain

To produce the chips (shown in Box 1) which are key to AI's development and operation, there is a complex supply chain which relies on an array of different actors playing different roles. These steps are shown in Figure 2. A full discussion of the value chain can be found in Mapping the Semiconductor Value Chain (OECD, 2025[8]), which is summarised below. The process can be segmented into three major steps: Design, Fabrication, and Assembly Test and Packaging (ATP). Each of these process steps depends on distinct inputs including materials such as silicon but also equipment, often from highly specialised suppliers.

Figure 2. Layers of the chip supply chain



Note: This diagram is a significant simplification of the key steps in the AI infrastructure supply chain with key inputs for competition identified by this paper. It does not survey the other inputs to AI systems such as data, algorithms, models and know-how.
Source: Adapted from (OECD, 2025[8]) Mapping the semiconductor value chain https://doi.org/10.1787/4154cdbf-en.

Design refers to creating the architecture and layout of a chip. Chip design is highly complex interdisciplinary process, involving thousands of engineers, years of research and development, and hundreds of millions of dollars of investment (Semiconductor Industry Association, n.d.[23]). Chip design is a high-value activity that sits upstream of manufacturing and downstream of fundamental research. It adds significant intellectual property (IP) value and can shape the performance, cost and energy efficiency of AI systems. Chip design for AI accelerators goes through regular design cycles with the latest chips typically released every one or two years, which can quickly make previous models redundant given the exponential improvements in technology.[3] **Fabrication** refers to a key part of the manufacturing process in which highly specialised tools and processes are used to physically produce the chips. This process is extremely capital intensive with a high degree of economies of scale. After being fabricated, chips are cut and packaged individually onto circuit boards, placed into their frame and protective outer shell, and tested before being released for sale this is referred to as **Assembly /Testing and Packaging**.

At the different levels of the supply chain there are also different business models operating which we briefly define below.

- **Integrated Device Manufacturers (IDM)** design, produce and sell chips. Traditionally, IDMs handled all three process steps – design, front- and back-end manufacturing. However, IDMs have increasingly outsourced parts of their production, following the so-called "fab-lite" business model.
- **Fabless companies** design and sell chips, they outsource manufacturing to foundries (manufacturing facilities producing chips for other companies) and Outsourced Semiconductor Assembly and Test (OSAT) companies.

- **Foundries** operate fabrication plants (fabs) to provide manufacturing services to companies that design chips (fabless, Integrated Device Manufacturers (IDM) and system companies).
- **Systems companies** also design chips and rely on outsourcing for manufacturing however unlike fabless companies they design them for their own products and services**.**
- **Outsourced Semiconductor Assembly and Test (OSAT)** companies offer contract manufacturing services to external customers in relation to testing and packaging the chips
- **Semiconductor Intellectual Property** (IP) vendors design and sell functional blocks (also referred to as "IP cores") that are used by chip designers to shorten time to market and lower chip design costs. Examples include a block to enable the chip to connect to USB or ethernet port.

2.3.1. Suppliers of chips used in AI

As noted in Box 1, GPUs are currently the most used chips for running AI tasks in data centres. The GPU market is highly concentrated with high margins, frequent innovations and huge capital investment. Nvidia (a fabless company) has emerged as the market leader in the sector, with recent estimates suggesting that the firm has over 80% market share for GPU chips used for AI (Farooque, 2025[24]). Nvidia has gross margins of over 70% and has seen its revenues increase by 405% between 2023 and 2024 (NVIDIA, 2025[25]). In July 2025, Nvidia became the first public company to reach a USD 5 trillion market value, exceeding even major tech firms such as Microsoft and Apple (Mickle, 2025[26]; Montgomery and Robins-Early, 2025[27]).

AMD is Nvidia's most direct rival in the GPU chip design space, but the firm was slower to focus on GPUs and they have struggled to catch up. This is driven by Nvidia's first mover advantage, strong performance and the CUDA software that Nvidia built around its hardware (Pak, 2024[28]). Recently however AMD has announced some major supply arrangements including with OpenAI (AMD, 2025[29]).

Box 2. Role of software in GPU market

Software plays a critical role in unlocking the full potential of GPUs for AI workloads. While GPUs provide the raw computational power needed for training and inference, it is the software stack; comprising frameworks, libraries, drivers, and compilers; that orchestrates and optimises this hardware for efficient parallel processing. Innovations in software enable better memory management, faster data throughput, and support for increasingly complex models, making it possible to scale AI systems effectively. Without robust and adaptable software, even the most advanced GPUs would be underutilized in AI applications.

Nvidia was the leader in this area and created the CUDA software environment, which is designed to optimise communication between AI training and deployment tasks and Nvidia GPUs. Given Nvidias strong market position, CUDA has in essence become the industry standard and Nvidia's CEO Jensen Huang has described it as the ''operating system'' for AI. There have been several attempts to replicate this software environment using opensource standards, for example, Modular.

Source: Gambacorta, L. and V. Shreeti (2025[3]), *The AI supply chain*, https://www.bis.org/publ/bppdf/bispap154.htm Pak, A. (2024[28]), The CUDA Advantage: How NVIDIA Came to Dominate AI And The Role of GPU Memory in Large-Scale Model Training, https://medium.com/@aidanpak/the-cuda-advantage-how-nvidia-came-to-dominate-ai-and-the-role-of-gpu-memory-in-large-scale-model-e0cdb98a14a0 (accessed on August 2025); Bradshaw, T. (2024[30]), Nvidia's rivals take aim at its software dominance, https://www.ft.com/content/320f35de-9a6c-4dbf-b42f-9cdaf35e45bb (accessed on September 2025); Vipra, J. and S. Myers West (2023[31]), "Computational Power and AI", https://ainowinstitute.org/publications/compute-and-ai.

Intel, the IDM which for many years has dominated CPU sales (as the key chip for running the Windows operating system) has attempted to design a rival GPU but has only achieved a very small market share

(Cusumano, 2024[32]). Nvidia has recently invested in Intel and formed a partnership to produce customer CPUs that NVIDIA will integrate into its AI infrastructure platforms (Stokel-Walker, 2025[33]).

Outside of GPUs, major tech companies (including Amazon, Google, Microsoft and Meta) have also begun to design their own ASICs. However, these ASICs have typically only been available through each firm's own cloud services and are designed for a specific use case.[4] Some commentators suggest demands for ASICs may increase as demand for AI inference increases, as inference workloads are less complex and more focused on speed and cost less than the training of models (UncoverAlpha, 2024[34]). There are also several smaller startups trying to produce specialist chips (Nicol-Schwarz, 2025[35]), although to date none have been able to gain significant market share. Lastly, companies like Alibaba, Baidu and Huawei from the People's Republic of China (hereafter 'China') are also starting to produce their own AI accelerators (Gambacorta and Shreeti, 2025[3]). With recent restrictions on Nvidia's use in China, it may be that the Chinese market dislocates from other markets. As of 2025, Huawei was estimated to have 28% of the market share of AI accelerators in China compared to Nvidia's 54% (Olcott and Wu, 2025[36]).

As well as AI accelerator chips there are also a range of other chips important to AI compute. For example, memory chips are also an important and concentrated area with providers such as Micron, Sk Hynix and Samsung which have the majority of market share (Davies, 2025[37]). These providers are all IDMs. Two key types of memory chips are explained below (Davies, 2025[37]):

- **Dynamic Random-Access memory (DRAM)**, which have become more commodity-like in nature with designs following industry standards and differentiation mainly focusing on speed, capacity and power efficiency. To compete, suppliers must invest in large scale facilities and memory prices are highly cyclical.
- **HBM chips**, more recently have come to the fore with new innovations helping to alleviate the memory limitations in storing and retrieving data efficiency for AI. As the HBMs lifecycle has shrunk, it has been reported that standards are struggling to keep pace, meaning that at least temporarily, there is increased product differentiation in the market as customers such as Nvidia have used custom HBB solutions (trendforce, 2025[38]). SK Hynix has reported to have a market share of over 50% of the market (Davies, 2025[37]).

2.3.2. Manufacturers of chips used in AI

Referred to as *foundries* in industry jargon, they are responsible for the front-end manufacturing of chips. The most advanced foundries are some of the most complex and expensive production facilities on the planet and take years to build (OECD, 2025[8]). Building cutting edge fabs requires enormous capital investment and achieving a satisfactory return investment is possible only for firms with very large scale (Varas et al., 2021[39]). Currently, few firms have the technical capacity to compete for contracts to fabricate AI accelerator chips

Taiwan Semiconductor Manufacturing Corporation (TSMC) is the market leader among chip foundries sales to fabless companies. Data from 2024 estimates that TSMC controls a solid majority of worldwide chip manufacturing contracts, with a market share above 60% (Bowman, 2024[40]). TSMC's share further increases when looking at contracts for only the most advanced chips, reaching 90% market share (Bowman, 2024[40]). For AI accelerator chips, TSMC leadership in 2024 disclosed that 99% of the world's AI accelerators are made with TSMC technologies (Wu, 2024[41]). Samsung is currently the second largest firm in terms of foundry revenue but also manufactures its own chips. It has been able to develop capacity to manufacture cutting edge chips by focusing its strategy on attracting orders to manufacture ASICs (i.e. the custom chips often developed by large technology companies that are optimised to work for a specific AI workload, such as Google TPUs) (Bowman, 2024[40]). Intel has reportedly struggled to keep pace with technical developments, at the expense of its market share (OECD, 2025[8]; Aguirre, 2024[42]). However, it has also reportedly started large-scale production of the most advanced chips to be manufactured in the US after a recent USD 32 billion investment (Acton, 2025[43]).

After being fabricated, chips are cut and packaged, placed into their frame and protective outer shell, and tested before being released for sale. This *OSAT* step in the chip value chain was in recent times undertaken by firms that were not the chip fabricators. ASE Group, Amkor Technology, JCET and Tongfu Microelectronics as the leading firms specialising in these OSAT services (OECD, 2025[8]). For more advanced chips (including AI accelerators), TSMC, Samsung and Intel have developed greater internal capacity and are all seeking to outsource less to OSAT firms. While there is less publicly available information on relevant market shares, potential scenarios in the longer-term could see the leading AI accelerator chip foundries compete for OSAT contracts for the chips they are fabricating, with the standalone OSAT suppliers focusing on less advanced chips for other use cases (OECD, 2025[8]; Mordor Intelligence, n.d.[44]).

2.3.3. Suppliers of key of inputs into AI chip production

There are a broad range of inputs into chip production. Many of these are markets which are highly concentrated, but for the purposes of this paper, we briefly highlight the key suppliers in two of the higher value elements of the supply chain, silicon lithography and electronic design automation.

Silicon lithography

Silicon lithography is the process of making the intricate patterns of circuits that form chips. Innovations in the chip industry to squeeze more transistors into increasingly small chips require more advanced methods of silicon lithography to create the patterns at this nanoscopic scale. In recent years, ASML, a firm based in the Netherlands has established itself as the leading firm in lithography, providing the lithography machines for all AI accelerator chip manufacturers (Center for Security and Emerging Technology, 2020[22]; Aguirre, 2024[42]; Narechania and Sitaraman, 2023[45]).

No other firm has been able to commercialise an alternative to ASML's Extreme Ultraviolet (EUV) Lithography technology that is necessary to manufacture the latest generations of chips. The two other major firms active in lithography (Nikon and Canon) have not been able to operational EUV technology. They are unable to provide the machines vital for manufacturing AI accelerator chips as these require higher precision than they what is available from their machinery (Center for Security and Emerging Technology, 2020[22]; Aguirre, 2024[42]; Narechania and Sitaraman, 2023[45]).

ASML manages a vast and complex network of suppliers and on occasion has acquired or invested in several of their key input suppliers for these lithography machines. For example, buying the firms supplying light sources and error in light beam detection technology, as well as taking a 24.9% stake in a subsidiary of optics firm Carl Zeiss that manufactures optical parts for ASML (ASML, n.d.[46]; ASML, 2016[47]). More directly relevant to AI they have also recently they have also made a direct investment in the other end of the supply chain, taking a stake in French AI model developer Mistral (ASML, 2025[48]).

Electronic Design Automation

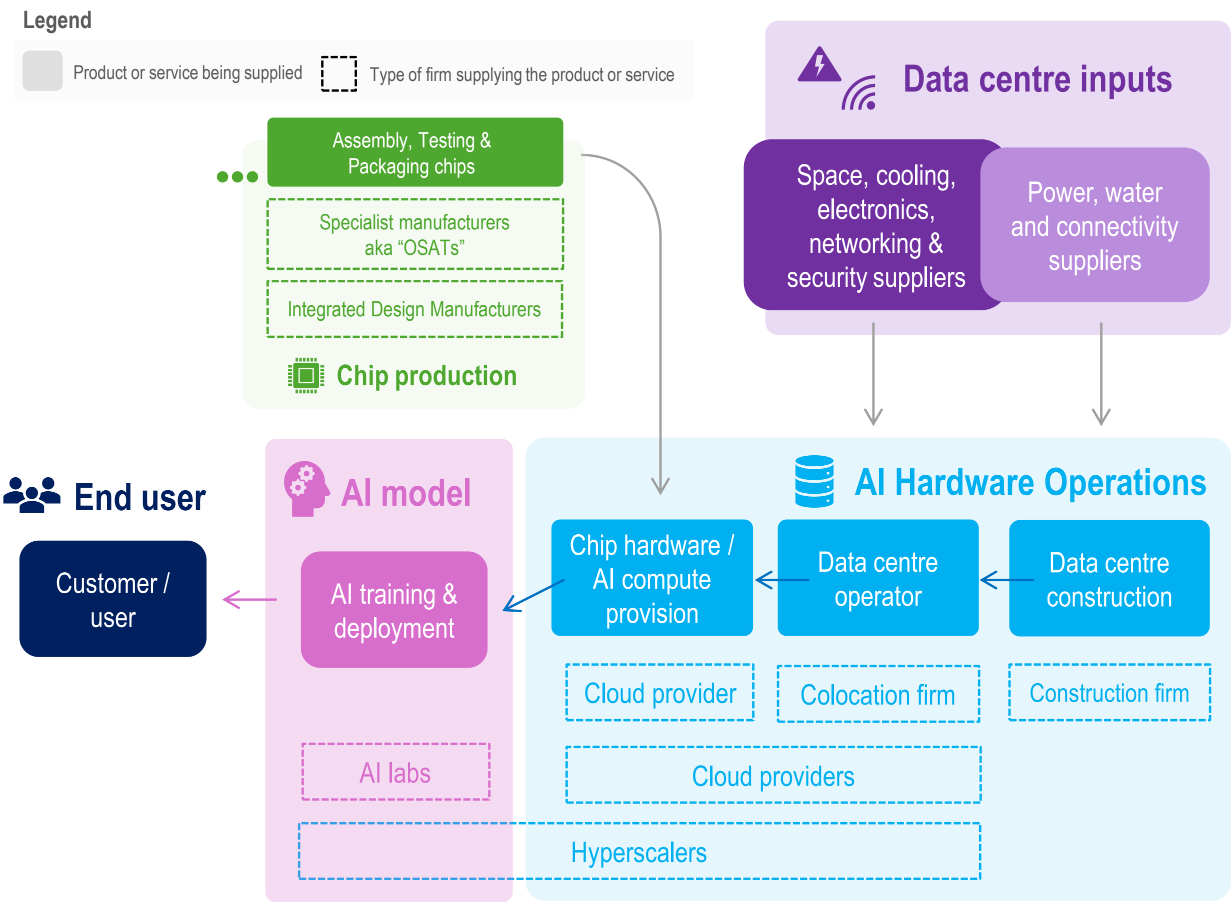
Electronic Design Automation (EDA) is "specialised software used by engineers to bring together semiconductor designs using IP cores and custom designs. It allows them to design, simulate and verify the design" (OECD, 2025[8]). EDA software is designed in collaboration with the development kits chip fabricators release to ensure designs can be manufactured in their foundries.

Regardless of who is designing and fabricating a chip, EDAs are key inputs given their ability to shorten time to market and lower chip design costs. Driven by demand for AI accelerator chips (as well as other chipsets like those for smartphones or automotive technologies), firms have seen high growth in recent years (Mordor Intelligence, n.d.[44]). Previous work from the (OECD, 2025[8]) has shown that three firms, Cadence, Synopsys and Arm "account for more than 60% of the global EDA market and 70% of the IP market". This concentration among just a small number of players means they have become indispensable for AI accelerator chip designers and manufacturers working at the cutting-edge.

2.4. Data centres, cloud computing and other inputs

While advanced chips are one of the core technologies powering the AI revolution, the technology must be operated together with other key inputs including power and cooling as highlighted in Figure 3 below. This section discusses the data centre operators which some in industry have started referring to as 'AI factories' (Harris, 2025[49]).

Figure 3. Diagram highlighting the different participants and inputs linked to data centres



Note: This diagram is a significant simplification of the key steps in the AI infrastructure supply chain with key inputs for competition identified by this paper. It does not survey the other inputs to AI systems such as data, algorithms, models and know-how.
Source: Adapted from Pilz (2023[18]), An Assessment of Data Center Infrastructure's Role in AI Governance, https://www.konstantinpilz.com/data-centers/assessment.

2.4.1. Data centres and cloud computing

Data centres serve as the backbone for AI development, providing the computational infrastructure required for both model training and inference. During training, vast datasets are processed through high-performance compute *clusters* equipped with the necessary GPUs or specialised AI accelerators. This phase is highly resource-intensive, demanding significant power, cooling and networking capacity to handle the parallel compute workloads. Once trained, models transition to inference, where they apply learned parameters to new data for predictions or decision-making. Inference still requires optimised hardware and low-latency environments (i.e. minimal delays in the time it takes for data to reach users) to deliver real-time responses at massive scale (OECD, 2025[5]).

Modern data centres are therefore evolving to balance these dual demands, integrating energy-efficient designs, access to clusters of AI accelerator chips, workload orchestration, and advanced networking to support the growing scale and complexity of AI applications. To reach the necessary scale of data centre compute and service provision, AI firms have typically relied on the data centres and services operated by the largest cloud computing suppliers (Narechania and Sitaraman, 2023[45]).

Cloud computing refers to a service model (OECD, 2014[50]):

> *for enabling on-demand network access to a shared pool of configurable computing resources (e.g. networks, servers, storage, applications and services) that can be rapidly provisioned and released with minimal management effort or service provider interaction (Mell and Grance, 2011[51]).*

There are broadly three categories of operators with regards to AI cloud providers (Lehdonvirta et al., 2025[52]):

- Government funded compute facilities – these are typically intended for academic or military use.
- Private compute clusters – these are owned by for-profit companies that build and use the compute for their own business purposes. They consist of large numbers of AI accelerators mounted into interconnected computers deployed in data centres. A private cluster can be used to power the company’s own AI development or rented out to another company.
- Public cloud providers that are also for-profit companies. They are called “public” not because of any government affiliation but because their services are in principle available on demand to the general public and thus shared by many customers.

Given the scale of data centre compute and service provision required to train and deploy AI services, the largest cloud providers are well placed to take significant market share in AI cloud provision. Google Cloud, Amazon Web Services and Microsoft Azure have been dubbed the *hyperscalers* and are subsidiaries of larger digital services companies (OECD, 2025[5]). As noted by the French (Autorité de la concurrence, 2023[53]),

> *All three belong to major digital companies that are among the world's largest market capitalisations. They already have a strong presence in digital services markets and have leveraged their considerable financial resources and internal needs to build up IT capacity worldwide and offer a large number of diverse cloud services, which have subsequently formed ecosystems.*

Compared to other areas of the AI infrastructure supply chain, competition authorities have already conducted far more in-depth research into cloud market dynamics. Across the market studies conducted to date, the combined market share of the hyperscalers across national and regional cloud computing services markets is consistently significant (OECD, 2025[5]).[5] The most recently reported market share estimates for the cloud market globally also suggest the largest three providers have over 60% share of the market (Richter, 2025[54]).

AI firms often use the hyperscaler cloud computing services, given the hyperscalers’ data centres ability to provide the necessary compute resources and networking capacity necessary for training and deploying AI (Stucke and Ezrachi, 2024[55]). Hyperscalers are also vertically integrated into firms that are active in AI development (OECD, 2025[5]). For example, Google operates its own data centres that power its own AI services (such as the Google Gemini AI assistant). Hyperscalers have also begun to strike partnerships or *strategic collaborations* with major AI firms. Typically, this involves the hyperscaler making a multi-billion dollar investment in the AI firm, with the AI firm agreeing to use the hyperscaler’s cloud services and data centres for training and deploying their AI models (OECD, 2025[5]). For example, in November 2024, Amazon invested an additional USD 4 billion into major AI firm Anthropic, with Anthropic announcing Amazon Web Services as their primary training partner (Amazon, 2024[56]).

However, there are other providers offering cloud compute capacity to AI. These include, for example, Oracle, which has recently signed a massive USD 300 billion deal to provide cloud computing power to OpenAI the model developer which developed the popular ChatGPT AI service (Tom's Hardware, 2025[57]). There are also other smaller providers such as such as CoreWeave, Crusoe, Nebius, and Lambda Labs, which apply a similar on-demand business model, but focus exclusively on providing AI compute services. CoreWeave has received investments from Nvidia, securing a multiple billion-dollar order that requires Nvidia to purchase any unsold CoreWeave capacity through to 2032 (Reuters, 2025[58]). In addition, some AI modellers such as Mistral have also sought to move up the chain and are beginning to look to develop compute capacity (Mistral AI, 2025[59])

2.4.2. Connectivity services and infrastructures, also referred to as "digital infrastructure"

As AI models grow in size and complexity, networking becomes a critical enabler of performance and scalability. Training large models like LLMs requires thousands of GPUs to work in parallel, exchanging data continuously at ultra-high speeds.

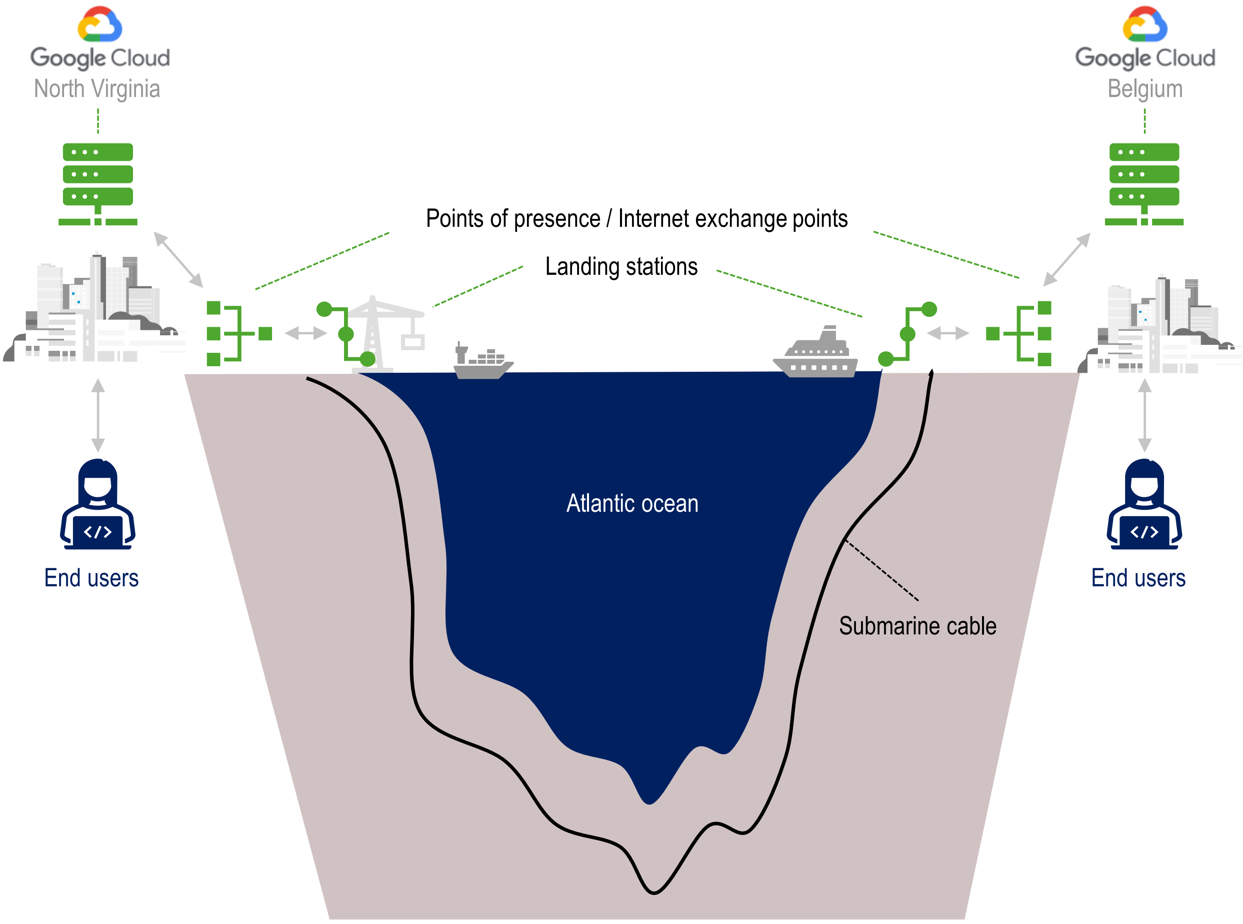
AI systems rely on distributed cloud computing environments, often spanning multiple data centres across regions. This creates an even greater need for robust digital infrastructure and global connectivity, which includes fibre networks, Internet Exchange Points (IXPs) and submarine cables (Filippucci et al., 2024[60]). In the past two decades, innovations in the digital services sectors have created far greater demand on this backbone infrastructure. For the largest technology companies, investing in and/or developing backbone infrastructure is necessary to ensure efficient and reliable data distribution. More recent developments such as the widespread adoption of cloud computing, and the need for a global network of data centres for AI compute has only increased demand for large technology firms who can afford to vertically integrate across several parts of backbone infrastructure (Filippucci et al., 2024[60]).

The OECD (2024[61]) has studied in detail the activities and investments of the five largest digital services firms by market capitalisation (i.e. Alphabet/Google, Amazon, Apple, Meta/Facebook and Microsoft). Since 2024 the level of capital investments made by these players has also continued to accelerate (Thomas, 2025[62]) . These firms have made substantial investments in digital infrastructure, including submarine cables, terrestrial cables (e.g. fibre), cloud infrastructure and data centres. Such investments aim to help them move data traffic between countries, as well as between their data centres and end users.

They have become particularly important investors of submarine cables, a critical component of backbone infrastructure. Prior to 2012, these firms accounted for less than 10% of total used capacity of submarine cables. By 2021, this figure had climbed to 69% (OECD, 2024[61]). Alongside investments through consortiums, technology companies independently financing submarine fibre cables have become more common in recent years. Nearly half of trans-Pacific cable investments scheduled to begin operation between 2023 and 2025 are backed or funded by these major technology companies (OECD, 2024[61]).

Figure 4 below shows a visual representation of the Internet backbone infrastructure underpinning Google's Dunant submarine cable connecting its US and European data centres and cloud computing networks.

Figure 4. Diagram of Google Durant submarine cable project and related infrastructure



Note: The diagram is a simplified version that does not explain more detailed aspects of submarine cables such as different cable sheaths at different depths, or the power and transmission equipment needed to send information along fibre optic cables.
Source: Adapted from Jayne Stowell (2018), "Delivering increased connectivity with our first private trans-Atlantic subsea cable", The Keyword (Google Blog), https://blog.google/products/google-cloud/delivering-increased-connectivity-with-our-first-private-trans-atlantic-subsea-cable/.

Large technology firms have also invested in ways to make sure data moves smoothly and reliably across the Internet from their data centres to users, including their AI compute-based products and services. They do this by making special agreements, called peering agreements, to connect their networks with others. They also work on systems that help store and deliver data faster, called content delivery networks (CDNs), often by partnering with communication companies or Internet providers. Some technology companies run their own CDNs to support their services and those of other businesses. In addition, firms have formed agreements with Internet service providers to place servers closer to users, which helps deliver content (including AI based) more quickly and efficiently (OECD, 2024[61]).

2.4.3. Access to energy and water for data centres

As discussed above AI's transformative potential is underpinned by its computational intensity. As more advanced AI systems are trained and deployed, their reliance on data centres and high-performance computing infrastructure has made energy supply and generation infrastructure a strategic input.

While estimates vary based on geography, the AI model being trained or queried, and the data centre infrastructure, it is clear that AI systems (namely the data centres where they are trained and deployed from) require a great deal of energy (Filippucci et al., 2024[60]; OECD, 2019[63]). Indeed, within the industry, data centre capacity is often measured in terms of power (megawatts) (Competition and Markets Authority, 2025[64]).

An illustrative example of the energy demand can be drawn from estimates relating to ChatGPT and other OpenAI products. The MIT Technology Review estimated that OpenAI's GPT-4 model (launched in 2023 and deployed to the public through products such as ChatGPT and Microsoft Copilot) used 50 gigawatt-hours of electricity to train (O'Donnell and Crownhart, 2025[65]). This equates to the same power usage of the population of San Francisco for three days. Analysis by investment firm Blackstone estimates compared AI outputs to a search engine query, with one ChatGPT query consuming roughly 10 times as much power, and a video generation request involves 10 000 times as much power (equivalent to charging a typical smartphone roughly 119 times) (Klimczak, 2024[66]).

In their market study into cloud computing, evidence gathered by the UK (Competition and Markets Authority, 2025[64]) showed that:[6]

- The energy supply to run and cool IT equipment in data centres is the largest portion of operational costs for running a data centre (compared to factors such as rent, maintenance, equipment, depreciation and labour costs).
- Access to energy may present challenges for establishing new data centre capacity, as continuous and reliable energy sources are not always present, especially in areas with existing high energy demand.

Against this backdrop, firms including Amazon, Microsoft and Google have acquired stakes or entered into long-term supply commitments with energy suppliers to meet the power requirements of their current and future data centres (Competition and Markets Authority, 2025[64]). This includes investing in reactivating nuclear energy plants and natural gas generators (da Silva, 2024[67]; Sherman, 2024[68]; Weise and Metz, 2025[69]).

Energy efficiency also creates additional incentives for the largest firms in the sector to vertically integrate across operating data centres, designing and manufacturing AI accelerator chips. Custom designing Application-Specific Integrated Circuits (ASICs) described above in Box 1 can increase the efficiency of AI compute, allowing firms to reduce their reliance on more heavy general AI chips such as the GPUs manufactured by Nvidia. For example, Google's development of its custom *Tensor* chips focuses on enhancing energy efficiency (Netherlands Authority for Consumers and Markets, 2022[70]); and Amazon's partnership with Annapurna Labs to manufacture chips for its AI data centres aims to reduce energy consumption (Weise and Metz, 2025[69]).

Operating compute in data centres, particularly those involving large-scale model training and inference, also generates significant heat. Data centres for AI compute also rely on water-intensive cooling systems (in addition to the indirect water usage relating to electricity generation). This is an often-overlooked aspect of the AI infrastructure supply chain, with far less data and academic literature available (OECD, 2022[71]). Estimates from (Mytton, 2021[72]) note that data centres are already among the top-10 water consuming industries in the United States, and "often cluster in similar geographic areas and many rely on scarce water supplies, particularly in the western United States" (Siddik, Shehabi and Marston, 2021[73]).

Access to abundant and reliable water sources can influence where data centres are located, creating regional advantages and investment flows for firms able to locate data centres in optimum locations in terms of water access. Access to sufficient water sources may also enable more innovative and efficient operation of AI data centres, by allowing firms to reduce their energy consumption. The French (Autorité de la concurrence, 2023[53]) market study highlighted firms such as OVHcloud and Scaleway, which are operating data centres with water based cooling systems that can save up to 40% in energy consumption compared to conventional air conditioner based cooling systems. Technology on cooling is developing rapidly, and as technologies such as closed loop systems develop access to large amounts water may become less important.

3 Market features in AI infrastructure

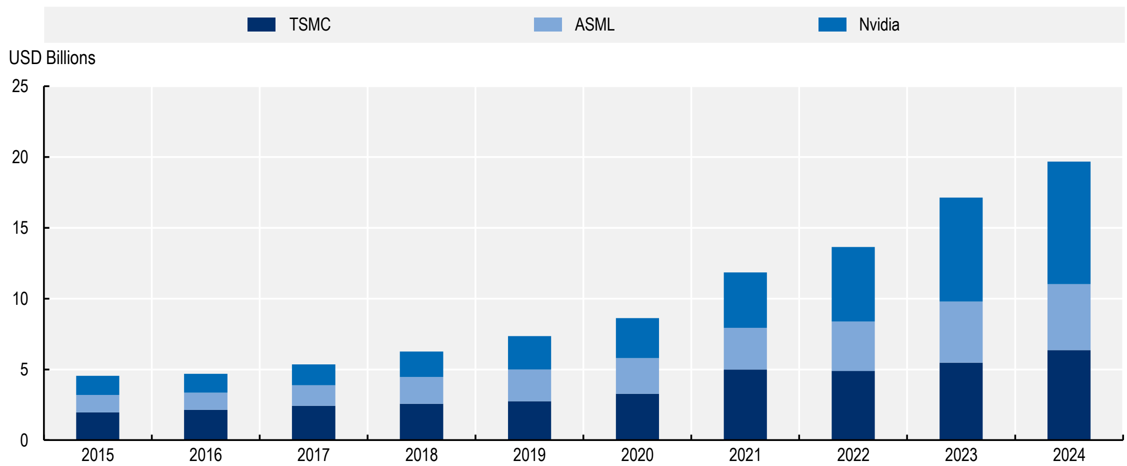
There are a range of different markets in the AI infrastructure chain, and each will have their own unique dynamics. While we are seeing constant innovation in both the AI systems being developed and launched and in the underpinning technology infrastructure, many layers of physical infrastructure underpinning AI have been relatively stable in relation to the fundamental organisation of the supply chain. This constancy makes it easier to assess potential competition issues and risks (Narechania and Sitaraman, 2023[45]).

Doing a full market-by-market analysis is beyond the scope of this paper. However, this chapter seeks to highlight some important features which appear to be present at multiple levels of the supply chain. These features may have important ramifications for competition and are therefore a good basis to start exploring potential competition policy responses in AI infrastructure. The intention is to survey market features that authorities and policy makers may wish to consider in more detail, noting that AI infrastructure covers a wide variety of different markets, with some common themes which this paper seeks to highlight.

3.1. High levels of innovation and intellectual property

The AI infrastructure supply chain is marked by a very high degree of research and development spend, innovation and intellectual property. At many layers of the supply chain such as chip design, lithography and fabrication firms must make exceptionally high investments in research and development. This is driven by the complexity and scale of the technologies involved. Competition in data centres may also increasingly focus on developing the optimal technologies for energy usage and cooling. Figure 5 highlights that firms involved in the chip ecosystem have significantly increased their research and development spending in the last ten years, with just three firms spending almost USD 20 billion on research and development in 2024.

Figure 5. Spend on research and development, example key firms in the chip ecosystem, 2015-2024



Notes: Converted to USD but not adjusted for inflation, R&D spending encompasses broader R&D investment beyond AI specific technology.
Source: OECD analysis based on annual reports.

In such sectors, intellectual property protections, including patents and trade secrets, can be essential to allow firms to recoup their R&D investments and maintain competitive advantages (Aguirre, 2024[42]). Licensing refers to the legal process by which the owner of IP rights grants permission to another party to use, produce, or commercialise the protected invention, work, or technology under agreed terms and conditions. Licensing is an important instrument for diffusing innovation, for allowing innovators to be rewarded for their efforts, and to promote co-operation and follow-on innovation during IP rights' period of exclusivity. However, licensing agreements can also have anticompetitive effects, such as anticompetitive foreclosure (OECD, 2019[74]).

Firms with leading technological positions do not necessarily have to licence their technology to other suppliers but may do so for strategic reasons, for example generating additional revenue without taking on manufacturing risk. In some situations, the refusal to licence IP can infringe competition law. This is typically only under exceptional circumstances where a provider is in a dominant position and the IP is an essential input in another market. For example, there was previously a case against Microsoft for refusing to licence the specifications required to ensure interoperability of Windows to manufacturers of competitor server operating systems (European Union, 2007[75]). More recently there have been reported investigations into the licencing terms of the Microsoft cloud (Godoy, 2024[76]) (Reuters, 2025[77]), as well as complaints reportedly being raised by Qualcomm (a chip maker) about ARM's (Chip IP developer) licensing model.[7]

There are broadly two types of licenses in the sector (Business Software Allliance, 2005[78]):

- proprietary or commercial licenses - which restrict usage to specific terms often involving fees and limited redistribution
- open-source licences – which aim to promote collaboration and innovation by allowing free use and modification.

The use of open-source licences has increasingly been used as a strategic tool by companies seeking to overcome the ecosystem advantages that a technology first mover has gained (see, for example, (Klotz, 2025[79])). This strategy encourages other players in the sector to contribute to systems and technologies around an open-source system which is not under proprietary control helping to reduce potential ecosystem barriers to entry and expansion.

Licensing can raise challenges for competition authorities in some circumstances, and they will need to carefully balance the need to protect intellectual property rights and ensure innovation and investment are not discouraged with the need to prevent dominant firms from leveraging the technology to exclude rivals from other markets. In doing so authorities may need to consider the likelihood to which rivals would be able to leverage open-source technologies to develop alternative solutions.

Beyond the complex interaction of intellectual property and competition, markets based on high levels of innovation may face other competition issues including high concentration and barriers to entry which are discussed further below.

3.2. High concentration and barriers to entry

While the innovation highlighted above is a positive outcome, in several layers of the supply chain there is also a high degree of concentration. Table 1 below illustrates the extent of concentration across several key segments of the AI infrastructure supply chain. In three segments, a single firm is reported to hold over 80% of the market, while in three others, the top three firms collectively hold over 60%.

Table 1. Examples of a concentrated supply chain

Key AI supply segment	Global share of largest provider reported over 80%	Global share of largest 3 players reported as having over 60% share
Advanced Lithography tools	**Yes** – ASML	
Advanced AI chip fabrication	**Yes** - TSMC	
GPUs for AI compute	**Yes** – Nvidia	
HBM chips		**Yes** – SK Hynix, Samsung and Micron
Cloud provision		**Yes** – AWS, Google and Microsoft
EDA software		**Yes** – Cadence, Synopsis and Siemens

Note: The OECD has not done a market definition exercise, and these shares may not therefore relate to economic markets but are commonly understood to be key parts of the AI supply chain.
Sources: Authorité de la concurrence (2023[53]), Opinion 23-A-08 of 29 June 2023 on competition in the cloud sector, https://www.autoritedelaconcurrence.fr/en/opinion/competition-cloud-sector; Netherlands Authority for Consumers and Markets (2022[70]) Market Study Cloud services; Ofcom (2023[80]) Cloud services market study - Final report; Japan Fair Trade Commission (2022[81]) Report on Fact-Finding Survey on Trade Practices by Digital Platform Operators Report on Trade Practices in Cloud Services Sector Contents, https://www.jftc.go.jp/en/pressreleases/yearly-2022/June/221102EN.pdf; Korea Fair Trade Commission (Korea Fair Trade Commission, 2022[82]), Announcement of the results of the cloud field survey, https://www.ftc.go.kr/www/selectBbsNttView.do?pageUnit=10&pageIndex=1&searchCnd=all&searchKrwd=%ED%81%B4%EB%9D%BC%EC%9A%B0%EB%93%9C&key=12&bordCd=3&searchCtgry=01,02&nttSn=42705 ; OECD (2025[8]); Mapping the semiconductor value chain: Working towards identifying dependencies and vulnerabilities. https://doi.org/10.1787/4154cdbf-en; Narechania and Sitaraman (Narechania and Sitaraman, 2023[45]), An Antimonopoly Approach to Governing Artificial Intelligence, https://dx.doi.org/10.2139/ssrn.4597080; Center for Security and Emerging Technology (2020[22]), AI Chips: What They Are and Why They Matter, https://cset.georgetown.edu/publication/ai-chips-what-they-are-and-why-they-matter/; Aguirre (2024[42]) On Labs and Fabs: Mapping How Alliances, Acquisitions, and Antitrust are Shaping the Frontier AI Industry, https://doi.org/10.48550/arXiv.2406.01722 and Narechania and Sitaraman (2023[45]), An Antimonopoly Approach to Governing Artificial Intelligence, Sitaraman, Ganesh and Narechania, Tejas N., An Antimonopoly Approach to Governing Artificial Intelligence (January 17, 2024). 43 Yale Law and Policy Review 95 (2024), Vanderbilt Law Research Paper No. 24-8, Available at SSRN: https://ssrn.com/abstract=4597080 or http://dx.doi.org/10.2139/ssrn.4597080.

The firms listed in Table 1 are examples, focussing on some of the highest value areas of the supply chain. However, it is important to note that high concentration is also found throughout the semiconductor supply chain where the technologies often involve specialised suppliers which have very high market shares (Burja, 2024[83]). Other elements such as fibre network infrastructure and energy networks by their nature may also be potentially natural monopolies or oligopolies at a local level.

It is important to note that market concentration is not inherently problematic. Given the high levels of research and development, the high concentration and profits in these markets can likely, at least initially, be viewed as a reward for their investments and innovation. This is often referred to as the escape effect (Arrow, 1962[84]; OECD, 2023[85]; Schumpeter, 1934[86]) and creates incentives for further innovation.

At some levels of the supply chain, the competition on innovation can even lead to a situation of 'competition for the market' where firms compete through innovations and investment to become the sole or dominant provider for the next decade (OECD, 2019[87]). In the context of AI infrastructure, such competition can drive rapid advancements in compute capabilities, data management and model deployment efficiency, as firms seek to establish themselves as indispensable suppliers in the AI ecosystem.

There are also barriers to entry which make it challenging for firms to compete which can contribute to the high level of concentration and weaken competition. These barriers include:

- Very high capital requirements – many levels of the supply chain are very capital intensive, for example the tools for manufacturing chips can cost hundreds of millions of US dollars alone (Shilov, 2024[88]). While entry into chip design is possible in partnership with a manufacturer such as TSMC (or more recently Intel), as discussed above the high levels of innovation also create a very high degree of investment needed in research and development.

- Economies of scale – given the very high level of fixed costs at many levels of the supply chain a large degree of scale is needed to cover the fixed costs. For example, at the cloud level it has been found that there are substantial economies of scale (OECD, 2025[5]).
- Long lead times – building and connecting data centres can take years and developing chip manufacturing plants even longer. Given the highly dynamic markets, these long lead-times create additional risks which may restrict entry.

In addition, some elements of the supply chain such as chip design also have proprietary software ecosystems which add additional barriers to develop. For example, the EU's case assessing the merger of Nvidia and OpenAI highlighted the high barriers to entry in discrete GPUs for use in datacentres including the need to have a well-developed software stack (European Commission, 2024[89]).

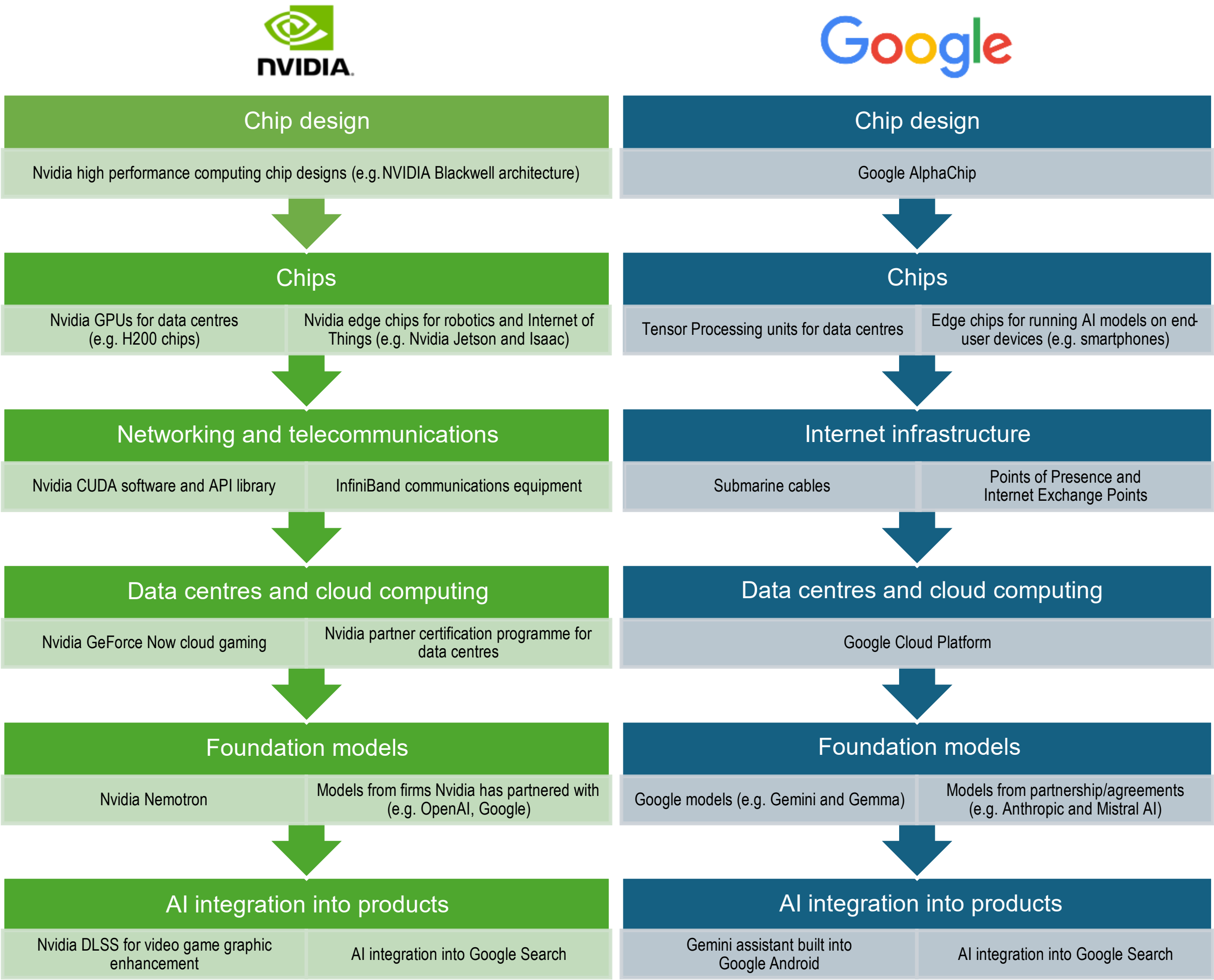
High concentration driven by innovation can still create a risk of anti-competitive conduct. Once a firm secures a dominant position, the incentives to innovate may shift, potentially leading firms seeking to protect their position. Potential behaviour includes the use of conditional rebates or other mechanisms to prevent rivals from reaching sufficient scale, or simply through targeting smaller innovative rivals through a series of *killer acquisitions*. Concentration can also increase the risk of collusion, particularly in markets with a small number of players with similar market shares and technology that has matured.

3.3. Vertical and conglomerate relations

At many levels of the AI infrastructure supply chain, the complexity historically led to a level of specialisation where no one company was able to produce everything in a fully vertically integrated way.[8] Today, though there is an increasing amount of vertical integration in the value chain, as well as firms expanding into offering multiple services which may be complementary (conglomerates). For example, the hyperscalers have begun developing AI accelerator chips and investing in downstream modellers (Google, 2024[90]; Amazon, 2024[56]; Meta, 2024[91]; Warnock, 2025[92]), and Nvidia has expanded into networking solutions through the acquisition of Mellanox (European Commission, 2019[93]).

Figure 6 below shows two case studies of vertical integration across the AI infrastructure supply chain, which extends down to consumer facing AI models and services utilising AI.

Figure 6. Case studies of vertical integration in AI infrastructure supply chain



Note: This figure is illustrative of examples of vertical integration and is not intended as an exhaustive list of products, market segments and supply chain steps. Similar vertical integration figures could be generated for numerous firms in the AI infrastructure supply chain (e.g. Amazon, Microsoft and Intel).

Source: https://www.gov.uk/government/publications/cma-ai-strategic-update/cma-ai-strategic-update; https://www.generativevalue.com/p/the-ai-semiconductor-landscape ; https://www.nvidia.com/en-us/ai/ ; https://www.nvidia.com/en-us/data-center/h200/ ; https://www.nvidia.com/en-us/data-center/; https://www.nvidia.com/en-us/data-center/colocation-partners/; https://developer.nvidia.com/ai-models/; https://www.nvidia.com/en-gb/geforce/technologies/dlss/ ; https://www.nvidia.com/en-us/geforce-now/; Authorité de la concurrence (2023[53]) Opinion 23-A-08 of 29 June 2023 on competition in the cloud sector , https://www.autoritedelaconcurrence.fr/en/opinion/competition-cloud-sector; Netherlands Authority for Consumers and Markets (2022[70]) Market Study Cloud services.

Vertical integration in the AI supply chain (as in other markets), has the potential for several benefits to consumer welfare. This includes enhancing investment incentives for producers, improving efficiency through enhanced co-ordination and removing double marginalisation (Beck and Scott Morton, 2021[94]). Given the very high barriers to entry, large providers moving to compete at other levels of the supply chain may also provide a route for potential competition which may otherwise be absent.

On the other hand, vertical integration involving firms with high market shares and power, can increase barriers to entry and raise the risk that dominant firms utilise their position at different parts of the value chain to tilt competition in their favour and make it harder for others to compete. In dynamic markets such as AI infrastructure, vertical foreclosure can be an optimal strategy because, although it can lead to short term profit loss, it affects future market structure, increasing the integrated firm's future profits (Fumagalli and Motta, 2020[95]). Such foreclosure could occur through different mechanisms including for example bundling/tying or even refusal to supply a key input.

These concerns may also extend beyond straightforward vertical integration between suppliers and producers to conglomerate integration at the same level where firms may employ similar tactics.

3.4. Partnerships and minority shareholdings

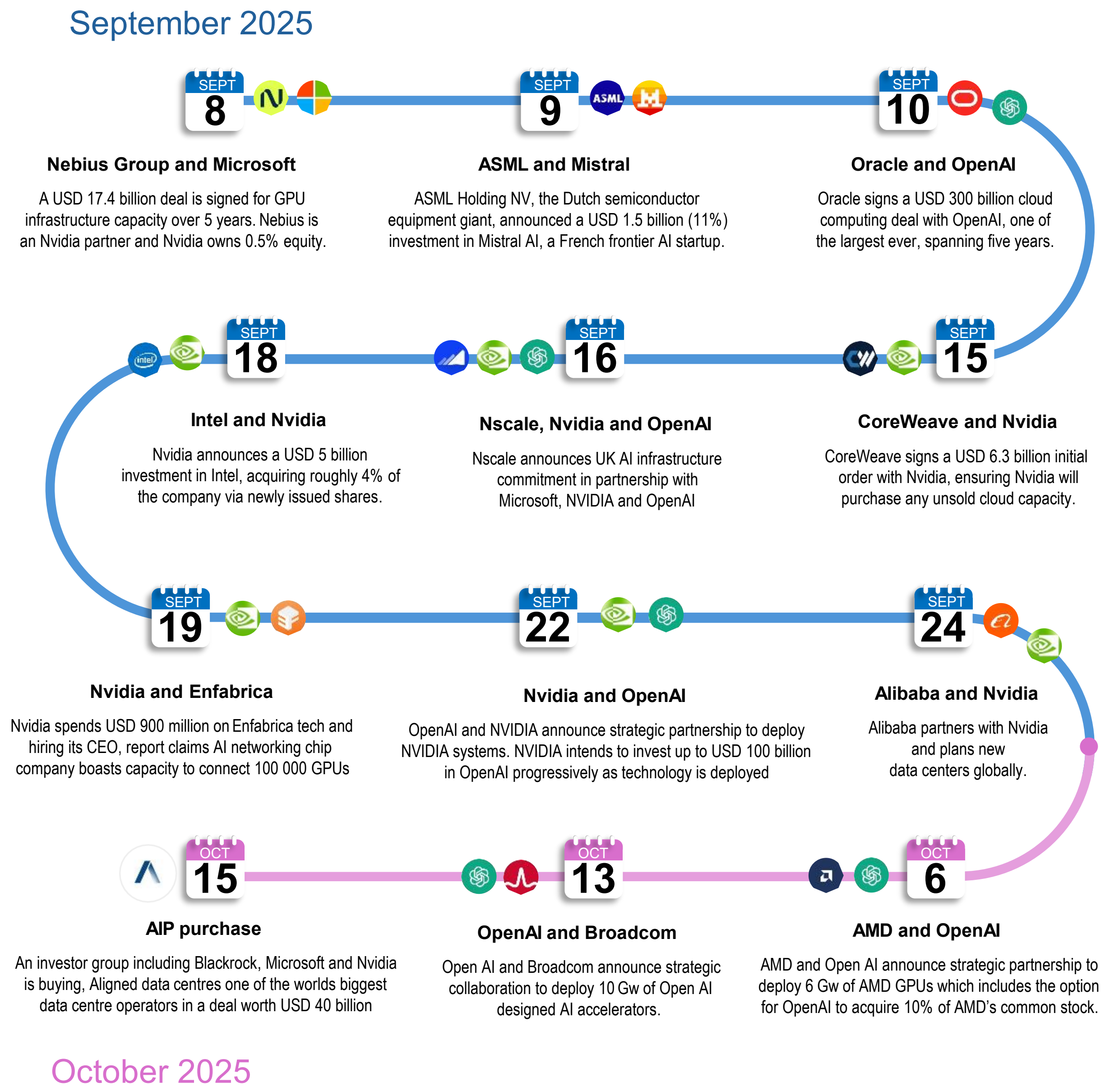
Agreements between firms can enhance efficiency, co-ordination and investment incentives. For instance, exclusive supply or distribution arrangements can ensure stable demand, enabling upstream firms to invest in specialised production or research and development. They can also help reduce transaction costs and improve logistical integration, which is important in sectors with high fixed costs and fast innovation cycles (OECD, 2021[96]). In the AI stack, partnerships can potentially help to accelerate innovation, optimise infrastructure use and grow the market (Ansari, 2025[97]). For example, they can help to ensure that companies developing AI models have access to the compute resources they need without undertaking huge capital investments (Auer and Zúñiga, 2025[98]).

It has become increasingly common for firms in the AI infrastructure supply chain to make investments and seek partnerships with other parties at different levels of the supply chain (Groza and Wierzbicka, 2024[99]). Often such arrangements appear to fall short of full mergers or acquisitions, will be for minority equity stakes and are structured together with a broader set of arrangements (Federal Trade Commission, 2025[100]).

Figure 7 highlights some of the recent deals which have been agreed in six weeks (September- October 2025). It demonstrates both the huge amounts of capital at stake, how quickly the industry is developing and the increasing number of cross holdings and partnerships. While some are simply long-term supply agreements others involve equity stakes and other arrangements. Others have been in place much longer such as the Microsoft - OpenAI partnership which has been in place since 2019 and recently been updated (OpenAI, 2025[101]).

While agreements can enhance efficiency and innovation, they can also pose potential risks to competition (Groza and Wierzbicka, 2024[99]) especially when involving competitors or potential competitors. As highlighted above, one major concern in relation to firms with significant market share in AI infrastructure is the risk that firms will seek to foreclose rivals. The partnership arrangements with firms at other levels of the supply chain may change the incentives or ability of firms in the market to carry out such foreclosure. For example, the CMA highlights the risk that partnerships could reinforce or extend existing positions of market power through the value chain (Competition and Markets Authority, 2024[102]).

Figure 7. Reported agreements in AI infrastructure between September and October 2025



Note: The above agreements are a selection of the most recent agreements at the time of writing, thy are demonstrative and not exhaustive.
Source: OECD summary of news based on: FT (2025) Microsoft taps Nebius to supply up to USD 20 billion of AI computing power, https://www.ft.com/content/a7aadbfe-1d6c-4159-890a-2f91d97a8d4b ; WSJ (2025) Exclusive | Oracle, OpenAI Sign USD 300 Billion Cloud Deal – WSJ https://www.wsj.com/business/openai-oracle-sign-300-billion-computing-deal-among-biggest-in-history-ff27c8fe ; ASML (2025) asml-mistral-ai-enter-strategic-partnership, https://www.asml.com/en/news/press-releases/2025/asml-mistral-ai-enter-strategic-partnership Reuters (2025), CoreWeave, Nvidia sign USD 6.3 billion cloud computing capacity order https://www.reuters.com/business/coreweave-nvidia-sign-63-billion-cloud-computing-capacity-order-2025-09-15/ Nvidia (2025) , NVIDIA and Intel to Develop AI Infrastructure and Personal Computing Products, https://nvidianews.nvidia.com/news/nvidia-and-intel-to-develop-ai-infrastructure-and-personal-computing-products TomsHardware (2025) 900 million on enfabrica tech, https://www.tomshardware.com/tech-industry/nvidia-drops-a-cool-usd900-million-on-enfabrica-tech-and-hiring-its-ceo-report-claims-ai-networking-chip-company-boasts-capacity-to-connect-100-000-gpus-together#xenforo-comments-3886527 Nvidia (2025) OpenAI and Nvidia announce strategic partnership, https://nvidianews.nvidia.com/news/openai-and-nvidia-announce-strategic-partnership-to-deploy-10gw-of-nvidia-systems ; Reuters (2025), Alibaba shares leap on Nvidia partnership, data center plans https://www.reuters.com/world/china/alibaba-launches-qwen3-max-ai-model-with-more-than-trillion-parameters-2025-09-24/.

Another aspect of the agreements which could cause competition risk is that even if equity stakes do not grant control communications facilitated by corporate governance rights can create opportunities for collusion between rivals (Ha and Law, 2025[103]). Recently for example the European Commission fined companies in the delivery market relating to the exchange of sensitive information (Fortuna, 2025[104]). As such minority shareholdings grow in AI infrastructure the risk of collusion may also grow.

The competition implications stemming from each agreement are difficult to assess and would need to carefully be examined. The possibility of examining under the merger regime and enforcement are discussed in section 4.

3.5. Switching barriers

Adding to entry barriers and limiting competition even between firms in the market there are several areas of the supply chain which have potentially high switching barriers. These barriers include:

- **Contracts and pricing:** Large customers often enter multi-year contracts with chip suppliers that include volume commitments, exclusivity clauses, or favoured pricing tiers to ensure their investments are worthwhile. Such commitments will often be driven by fears of scarcity but may mean that customers' ability to switch is limited in the short term. At the cloud level of the supply chain, concerns have been highlighted about the predicting contract spend, and potential barrier created by egress fees (charges for data leaving a cloud network) that may amount to penalising customers for switching providers. (OECD, 2025[5]).
- **Lack of standards and interoperability:** Given the fast pace of developments and the high levels of innovation, many areas of the supply chain do not have agreed standards and interoperable parts. This means, for example, that switching suppliers can often entail reconfiguring data centres, updating firmware, and modifying cooling and power systems. These physical and operational changes can be prohibitively expensive, especially enterprises with large-scale deployments. There has recently been an initiative launched by the Open Compute Project to seek to change this, with the support of Google, Meta and Microsoft (Open Compute Project, 2025[105]). Integration with existing cloud services may also be optimised for specific hardware, creating additional friction, and at the cloud level the ability to integrate services across multiple providers is limited by the use of proprietary interfaces and vendor specific solutions (OECD, 2025[5])
- **Feedback loops and ecosystems:** Like the previous waves of digital developments, key parts of the AI supply chain rely on integration between hardware and software. These integrated ecosystems can have benefits including seamless integration, enhanced user experience, convenience and efficiency, and enhanced security. Such ecosystems can also have strong network effects. These arise when the value of a system increases as more users, developers or data contributors participate (OECD, 2017[106]). For example, as more developers and enterprises adopt certain software ecosystems, the software become more valuable to other users as the number of tools and use cases increase. In AI infrastructure at different layers such software ecosystems or platforms have been made on a proprietary basis and tied to hardware, for example Nvidia's CUDA software. This then creates barriers to switch to alternative hardware, as a customer may need to make a complete ecosystem switch at significant expense. This is why challenger firms often seek to lead open-source initiatives to try and overcome the network effects by pooling resources with other challengers.

Switching barriers such as these are not necessarily be the result of illegal conduct but can add friction to competition by making it harder for companies to win on the merits and for customers to pick the optimal product or service for them. In section 4 there is a broader discussion of potential competition authority responses including through market studies and regulation which could potential seek to lower switching barriers.

3.6. Supply shortages/bottlenecks

The chip supply chain has in the past seen supply shortages (Jackson, 2024[107]). With the current forecast for demand for AI indicating massive demand, there could be a risk of supply shortages or bottlenecks emerging again, especially for cutting-edge chips and advanced lithography systems which are produced by a small number of companies in a limited number of locations.

The cost and scarcity of the specialised AI accelerator chips, such as GPUs and ASICs, may represent an important bottleneck for the sector (European Commission, 2024[108]). As chip designers begin to expand their offerings downstream into the cloud, and cloud providers begin to develop their own chips, a risk could grow that there are firms that could try to exclude rivals from upstream or downstream inputs or services. For example, a CMA market investigation highlighted that the supply of AI accelerator chips (such as GPUs) has been unable to keep pace with increasing demand associated with AI over the past three years. A delay in Nvidia chip delivery in 2024 had been a key factor in one cloud computing provider experiencing lower estimated revenue from AI compute for 2025 (Competition and Markets Authority, 2025[64]).

Shortages may also emerge in relation to local or national markets in relation to key energy, networking and water resources. With firms which are unable to gain access potentially being locked out of the market.

Shortages could both slow innovation in AI developments downstream and concentrate power, creating a risk of resource hoarding and preferential allocation. In such a context, firms with privileged access, often through exclusive agreements or via vertical integration, can accelerate their AI development while others are left behind. Control over who gets supplied becomes a strategic lever in this environment. Dominant cloud providers and chip manufacturers may, for example, prioritise internal projects or preferred partners, effectively shaping the competitive landscape.

On the other hand, demand outstripping supply can create the potential for alternative competitors to gain share, for example, in the operation of data centres several smaller providers appeared to have recently received large orders with AI modellers and cloud providers looking for as much supply as possible. The amount of capital being driven into the sector has recently raised concerns of a potential bubble, (Thornhill, 2025[109]) to the extent that there is any crash in confidence or demand for AI companies, the competition dynamics could change dramatically and make it harder for new entrants to raise funds.

3.7. Increasing state intervention

Given the importance of AI to all sectors and especially high technology industries, there is increasing government intervention in relation to the AI supply chain. Governments around the world are investing heavily in AI infrastructure because it is increasingly viewed as a strategic asset which is central to economic competitiveness, national security and technological sovereignty (Nakazawa and Pisa, 2025[110]).

Compute power is a key building block of modern AI systems, and control over these resources can shape innovation, industrial capabilities, and geopolitical influence. Public investment can also potentially help to address potential market failures, such as underinvestment in open-access compute. For example, the French (Autorité de la concurrence, 2024[111]) highlighted that certain technical and organisational developments, as well as public policies, may limit the barriers to entry including providing access to public supercomputers. Access to key infrastructure such as supercomputers can therefore be procompetitive as it can help to prevent potential exclusion or exploitation by dominant providers.

However, public investment itself can also raise competition policy challenges. If not carefully designed, public-private partnerships may entrench incumbents, distort markets or create dependencies on dominant platforms (OECD, 2022[112]). Preferential access to state-funded compute or data, for example, could reinforce existing power asymmetries. Moreover, governments may inadvertently favour national champions or larger firms, undermining cross-border competition and innovation.

Box 3 highlights some of the examples of government intervention from the first developments of the semiconductor to more recent government moves to develop AI infrastructure.

Box 3. Examples of OECD government investments and AI infrastructure partnerships

The origins of the semiconductor industry are intertwined with government programmes. In the 1950s and 1960s, the US Department of Defence and NASA were among the first major customers for integrated circuits, providing critical demand and funding that enabled firms like Fairchild Semiconductor and Texas Instruments to scale production. Defence contracts helped subsidise R&D, while government labs and procurement policies accelerated the commercialisation of transistor and chip technologies. This early public investment laid the foundation for Silicon Valley's rise and established a model of state-supported innovation in strategic technologies.

Governments in East Asia in the 1970s and 1980s were also highly involved in the development of their own semiconductor industries. For example, Japan provided extensive funding for research and development in semi-conductor manufacturing and South Korea also used targeted R&D subsidies, export incentives and infrastructure development. Chinese Taipei, meanwhile, leveraged its semiconductor leadership anchored by TSMC, with government-backed industrial parks and research institutes.

In recent years, governments have intensified efforts to secure strategic control over AI infrastructure for example:

- The US government launched a series of initiatives to bolster domestic AI infrastructure. The CHIPS and Science Act (2022) allocated USD 52 billion to support semiconductor manufacturing and R&D, with a focus on AI-relevant technologies. In 2025 the US government announced a deal with Intel to purchase 10% of its equity for a USD 10 billion investment to expand US manufacturing facilities.
- The European Union launched a public-private investment strategies in AI infrastructure, centred around the InvestAI initiative and the broader AI Continent Action Plan. InvestAI aims to mobilise EUR 200 billion in AI investment, including EUR 20 billion specifically earmarked for the development of AI gigafactories, large-scale computing facilities designed to support the training of frontier models. EU member states have also recently been investing in AI infrastructure. The EU Commission recently approved a state aid package from Germany to support Infineon in setting up a new semiconductor manufacturing facility. France has also invested in national compute capacity and supporting startups through Bpifrance and the Grand Défi initiative.
- The UK government's GBP 2 billion AI Opportunities Action Plan includes sovereign compute investments and regional AI Growth Zones.
- South Korea in 2025 announced a support package of over USD 23 billion to support its semiconductor industry and Japan was reported to have invested USD 27 billion in support of the semiconductor industry between 2000 and 2024 with plans for USD 65 billion or more by 2030.

Note: This box includes a sample of examples and is by no means exhaustive.
Sources: Miller (2022[113]) Chip war: the fight for the world's most critical technology; Simon & Schuster; OECD (2019[63]) Measuring distortions in international markets: The semiconductor value chain, https://doi.org/10.1787/8fe4491d-en; OECD (2025[8]), Mapping the semiconductor value chain: Working towards identifying dependencies and vulnerabilities, https://doi.org/10.1787/4154cdbf-en; European Commission (2025[114]) Commission approves EUR 920 million German State aid measure to support Infineon in setting up a new semiconductor manufacturing facility, https://ec.europa.eu/commission/presscorner/detail/en/ip_25_557; European Commission (2025[115]), EU launches InvestAI initiative to mobilise EUR 200 billion of investment in artificial intelligence,

https://ec.europa.eu/commission/presscorner/detail/en/ip_25_467 ; Intel Corp (2025[116]) Intel and Trump Administration Reach Historic Agreement to Accelerate American Technology and Manufacturing Leadership, https://www.intc.com/news-events/press-releases/detail/1748/intel-and-trump-administration-reach-historic-agreement-to; (UK Department for Science, Innovation & Technology (2025[117])AI Opportunities Action Plan,https://www.gov.uk/government/publications/ai-opportunities-action-plan/ai-opportunities-action-plan#lay-the-foundations; OECD (2025[118]) Policy Brief - Recent trends in semiconductor subsidies, https://www.oecd.org/en/publications/recent-trends-in-semiconductor-subsidies_5e91af33-en.html; European Commission (2022[119]), A Chip Act for Europe, https://www.decision.eu/wp-content/uploads/2022/09/A-Chips-Act-for-Europe-Commission-Staff-Working-Document.pdf; Reuters (2025[120]), South Korea unveils USD 23 billion support package for chips amid US tariff uncertainty, https://www.reuters.com/technology/south-korea-unveils-23-billion-support-package-chips-amid-us-tariff-uncertainty-2025-04-14/; Tomoshige (2022[121]) Japan's Semiconductor Industrial Policy from the 1970s to Today, https://www.csis.org/blogs/perspectives-innovation/japans-semiconductor-industrial-policy-1970s-today; Yamaguchi and Kihara (2024[122]) Japan unveils USD 65 billion plan to aid domestic chip industry, https://www.reuters.com/world/japan/japan-propose-65-billion-plan-aid-domestic-chip-industry-draft-shows-2024-11-11 /; The Economist (2025[123]) Japan storms back into the chip wars , https://www.economist.com/asia/2025/08/21/japan-storms-back-into-the-chip-wars; France (2025[124]) Make France an AI powerhouse, https://www.elysee.fr/en/emmanuel-macron/2025/02/11/make-france-an-ai-powerhouse.

In addition to government investments another form of intervention which has increasingly featured in markets are trade barriers including tariffs and export controls. For example, there has been an evolving set of restrictions in advanced chips and manufacturing equipment such as those imposed by the US on China.[9] Recently China has also placed restrictions on firms abilities to use Nvidia's chips in an effort to boost its domestic industry (Leng, Wu and Bradshaw, 2025[125]).

The interventions, while often justified by national security or industrial policy goals, can create trade barriers that can distort markets, fragment supply chains and entrench dominant players. The implications on competition are uncertain. By preventing companies from competing in certain jurisdictions, export restrictions can reduce global economies of scale and efficiency as well as reducing options in the affected jurisdiction. However, such actions are implemented by governments balancing other geopolitical considerations. And in the long-term in jurisdictions with limitations on imports, domestic industries may begin to develop in the market gap (if the affected market is sufficiently large). This could result in additional global competitors or simply distinct supply chains in separate geographic markets.

As well as direct investments and trade policies, government regulations and legislation more broadly will impact the way competition develops in the AI supply chain. This is likely to be particularly the case at the data centre level, as they require large volumes of land, construction, energy and water. these inputs may be subject to a range of local and national legislation for other policy reasons. Such regulations may impact the growth of data centres and the provision of accelerated compute, benefitting the companies which have already made investments in these assets over new entrants (Meyers and Bourreau, 2025[126]). In addition, complex regulation can provide other advantages to the biggest providers as they may be better able to manage the regulation due to their scale and involvement in the establishment of the regulations.

4 Potential competition policy responses in AI infrastructure

The economic and social impacts of AI are forecast to be widespread and significant, but they rely on AI infrastructure to also grow at pace. As highlighted in section 3, there are several features of the AI infrastructure value chain which make it susceptible to competition issues. It is therefore important that competition authorities consider the best responses to potential competition issues in these markets. This section discusses the potential application of competition tools to the sector, as well as the extent to which there has been activity so far.

4.1. Competition enforcement

Having effective competition enforcement is a key tool in ensuring that firm conduct is kept in check and competition is based on the merits of enterprise. The US Department of Justice (DOJ) Assistant Attorney General Gail Slater recently highlighted the importance of enforcement in markets such as AI infrastructure where technology is rapidly developing:

> *Antitrust enforcement focuses on creating a level playing field for businesses big and small and ensuring that market incumbents do not unfairly hinder newcomers and startups. This is always important, but it is crucial where the technology is still developing rapidly (US DOJ, 2025[127]).*

Recently several authorities have launched investigations into the chip designers developing the latest AI chips.[10] The supply chain at the chip design level has seen several significant global enforcement cases in the past prior to the development of AI. However, to date there has been no completed enforcement actions relating directly to the AI infrastructure supply chain.

This section highlights some of the potential areas that authorities may need to pursue potential antitrust cases, as well as highlighting some of the past cases as examples of risks which may arise again. We first discuss anti-competitive agreements before considering the risks of unilateral conduct.

4.1.1. Anticompetitive agreements

As highlighted in section 3, the AI infrastructure supply chain is characterised by an increasing number of agreements, often including investments in equity which fall below merger thresholds. These agreements have created an interrelated web of firms with crossholdings and relationships which could have the potential to weaken competition in the sector.

Vertical or conglomerate agreements are not per se illegal and in most jurisdictions they are subject to an economic effects analysis. However, given the importance of the sector, authorities should pay close attention to such agreements. So far, it appears that the agreements have received more attention through merger regimes (discussed under merger control below). Although, the US FTC has conducted an initial investigation into some of the arrangements between cloud providers and generative AI companies. A summary of the findings published in the US FTC staff report are included in Box 4.

Box 4. US FTC review of partnership agreements

The US Federal Trade Commission in January 2025 published staff findings on a study into partnerships involving generative AI companies and cloud providers. They sought to understand whether these relationships were being used to circumvent merger review or create anticompetitive advantages. The report highlighted three areas to watch regarding the potential implications of AI partnerships:

- The partnerships could affect access to certain inputs such as computing resources and engineering talent.
- The partnerships could increase contractual and technical switching costs for AI developer partners.
- The partnerships provide cloud service providers access to sensitive technical and business information unavailable to others.

Sources: US FTC (2025), Partnerships Between Cloud Service Providers and Ai Developers - FTC staff Report on AI Partnerships and Investments 6 (b) Study, https://www.ftc.gov/system/files/ftc_gov/pdf/p246201_aipartnerships6breport_redacted_0.pdf.

The agreements examined by the US FTC and other authorities primarily relate to agreements between generative AI providers and cloud providers, but such agreements are becoming increasingly prevalent at all levels of the AI infrastructure stack. Intervention may be warranted when agreements contain provisions which could lead to market foreclosure (such as exclusivity), for example preventing rivals from accessing key inputs (e.g. compute or reaching customers). This is especially critical in sectors like AI infrastructure, where access to advanced chips, can determine market viability.

Beyond vertical agreements, recently there have been arrangements between firms at the same level of the supply chain. Such horizontal arrangements and collaboration can raise the risk of information sharing or broader collusion (OECD, 2018[128]).

More generally the risk of collision may be relatively low at the leading edge of the AI infrastructure supply chain, given it is mostly operating with fast moving technology and high levels of innovation. At many levels of the supply chain has one firm with a leading market share. In markets with such asymmetric shares, there is typically lower risk of collusion (Motta, 2004[129]), especially when the market leaders can often make legal monopoly profits from their innovations. That said, there are other parts of the supply chain where technologies have matured, and the products may have become commoditised with oligopoly-like competition between a small number of players. In these areas the risk of collusion is likely to be higher (Asmat, 2019[130]). For example, Box 5 below highlights enforcement action from the 2000s relating to a computer memory chip cartel.

Box 5. The DRAM Cartel

Dynamic Random Access Memory (DRAM) is a type of semiconductor memory chip used in the primary memory systems of computers and other digital devices. As the DRAM market matured and commoditised, chips became interchangeable and industry standards were widely followed. This commoditisation in concentrated markets with high information sharing increased the risk of collusion.

The DRAM cartel was a significant international price-fixing cartel involving several manufacturers of DRAM chips between 1998 and 2002. The cartel included major DRAM producers including Samsung, Hynix, Infineon, Elpida, Micron Technology, NEC, Hitachi, Toshiba, Mitsubishi Electric, and Nanya Technology. These companies co-ordinated prices and restricted competition, artificially inflating DRAM prices sold to PC and server manufacturers.

A number of competition authorities took action and issued large sanctions, including:

- The US DOJ launched investigations in 2002 under the Sherman Antitrust Act. Between 2003 and 2006, several executives and companies pleaded guilty. For example: Infineon was fined USD 160 million in 2004. Hynix paid USD 185 million in 2005. Samsung paid USD 300 million in 2005. Executives at Samsung were also sentenced to prison.
- In 2010, the European Commission fined nine companies a total of EUR 331 million for their role in the cartel. Micron Technology received immunity for whistleblowing.

Sources: European Union (2011[131]) Final report of the Hearing Officer - COMP/38.511 - DRAMS, https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52011XX0621(02); United States District Court (2005[132]) United States of America v Samsung Electronics Company and Samsung Semicoductor Inc, https://www.justice.gov/atr/case-document/file/509296/dl; United States District Court (2005[133]) Unites States of America v. Infineon Technologies AG, https://www.justice.gov/atr/case-document/file/499551/dl; (2005[134]) United States of America v Hynic Semiconductor Inc, https://www.justice.gov/atr/case/us-v-hynix-semiconductor-inc; United States District Court (2006[135]) United States of America v Sun woo Lee, https://www.justice.gov/atr/case-document/file/501201/dl.

Lastly, the rapid growth in AI infrastructure may also have implications for the risk of collusion upstream in large-scale projects involving data centres, fibre networks, and energy-intensive compute facilities. These projects often require tenders for the largescale procurement of construction, including specialised engineering expertise. The high capital investment, and co-ordination across a limited number of capable firms, create conditions which increase the risk for bid rigging activity. The complexity and technical specificity of AI infrastructure tenders could obscure pricing benchmarks and make it easier for firms to align bids or rotate winners.[11] Additionally, the urgency to deploy infrastructure may lead to compressed timelines and reduced scrutiny in procurement processes, weakening safeguards against cartel formation. As governments and tech giants increasingly rely on public-private partnerships to build AI infrastructure, competition authorities should be alert to the structural vulnerabilities in upstream markets and continue advocating for enhanced procurement design and detection to actively deter collusion such as bid rigging (OECD, 2025[136]).

4.1.2. Anticompetitive unilateral conduct

As highlighted in section 3 there are several highly concentrated parts of the AI infrastructure supply chain. Firms with very high market share and positions protected by barriers to entry will likely have a high degree of market power (OECD, 2022[137]). This raises the risk of firms potentially abusing their market power.[12]

The starting point for authorities in assessing whether conduct is an abuse will be to seek to prove dominance (or monopoly power). Dominance is often defined as the ability to act to a substantial degree independently of competitors, customers and suppliers (i.e. the ability to exercise market power) (OECD, 2021[138]). This can be expressed in metrics such as the ability to earn high profits on a sustained basis which are above costs (including a reasonable return on capital). Some jurisdictions will have market share-based thresholds which allow competition authorities to establish a firm's dominance beyond a certain share of the market (OECD, 2024[139]). [13]

In practical terms when assessing dominance, whether structural presumptions apply or not, the approach typically starts with defining the relevant market. In several layers of the AI infrastructure value chain, meeting the threshold of dominance would seem likely as several providers along the value chain appear to have near monopoly market shares on certain technologies and are generating significant profit margins (Narechania and Sitaraman, 2023[45]). One key aspect competition authorities may need to consider is the dynamic nature of the markets, for example in conducting market share analysis by focusing on the share of forward orders rather than the most recent years data an analysis.

One of the features of AI supply chains highlighted earlier in the paper is that there are high levels of both vertical integration and concentration (and likely market power) at different levels of the supply chain. Such conditions create the risk that firms will use their position in one market to distort competition in other areas in their favour. This potential for exclusionary behaviour could take different forms including outright refusing to supply, tying/bundling, use of rebates or other pricing strategies

There have been numerous examples in previous rounds of computing developments of tying and bundling being used in a potentially anticompetitive manner. These cases have not always been successful, demonstrating the challenges for authorities in bringing cases. For example, IBM was accused of tying memory components to its CPU in the 1970s (although the court sided with IBM on the basis that that the integration was technologically superior) (Rowles, 2000[140]). There have however been successful cases, for example, the European Commission fined Microsoft EUR 497 million for bundling Windows Media Player with Windows OS (European Union, 2007[141]), foreclosing competition in media players. Google was fined EUR 4.34 billion for tying its search engine and Chrome browser to the Android OS, leveraging dominance in mobile OS to gain advantage in search and browser markets (European Union, 2018[142]).

In relation to AI infrastructure such bundling or tying could emerge as a risk at different levels for example, most AI models operate off cloud computing resources where compute access is rented from third parties and increasingly specialised for AI modelling and training. This advanced cloud computing is primarily offered by three hyperscalers, Google, Microsoft and Amazon (OECD, 2025[5]). Competition concerns may emerge from vertical relationships where cloud service providers are also involved in developing and deploying AI models and applications. If a company has significant market power in the upstream cloud services sector, by bundling its AI model downstream together with this infrastructure, it could make it in the future hard for model developers to compete independently. Another risk is that AI chips are bundled with other components. Furthermore, risk can result for the AI infrastructure supply chain is in the supply of complementary hardware products. As firms begin to offer multiple complementary hardware products in the supply of the data centre technology, there is a risk that if a supplier has a dominant position in one where there are high barriers to entry, it may use that position to gain greater market share in other complementary hardware products.[14]

Authorities should also be alert to the risk of pricing strategies, such as the use of conditional rebates, which may seek to encourage exclusivity through the incentives delivered (OECD, 2016[143]). Such cases have occurred in previous rounds of chip development for example, the EU has previously acted against Intel regarding the structure of their rebates, although the finding was later annulled on appeal (Court of Justice of the European Union, 2024[144]). A summary of the case is discussed in Box 6 highlighting how both pricing behaviour in relation to chips can be viewed as potentially exclusionary, and also that such behaviour may create efficiencies, requiring thorough economic analysis.

Box 6. Intel rebates case

Prior to the latest wave of AI computing, CPUs were the core technology of modern computing. Intel was the dominant manufacturer with AMD a smaller rival.

In 2009, the European Commission fined Intel EUR 1.06 billion under Article 102 TFEU for abusing its dominant position in the x86 CPU market. The Commission alleged that Intel granted loyalty rebates to computer manufacturers (e.g. Dell, HP, Lenovo) to induce exclusivity and foreclose competition, particularly from AMD.

The case was subject to multiple appeals with the CJEU ultimately annulling the case in 2024. The decision highlighted the need for thorough economic analysis of any such conduct and in particular to respond to any economic arguments put forward by the defendant sufficiently.

The core theory relating to such rebate cases is that there is typically a smaller contestable share of the market that competitors can compete on, and the conditional rebates can mean that the effective price competitors would need to compensate customers with is so low as to make effective entry/growth impossible.

Notes: The European Commission also found against Intel for abusing its dominant position by paying providers for delaying products. These findings of naked restrictions were note overturned on appeal.
Sources: Judgement of the Court (fifth Chamber) 2024, Commission vs Intel Corporation, https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:62022CJ0240; OECD (2016[143]), Fidelity Rebates, https://doi.org/10.1787/fd9e16da-en.

As highlighted above, previous technological waves have seen complaints and investigations into exclusionary behaviour, including tying and rebates. While such cases are no doubt complex and have in the past faced challenges in court. This does not mean that authorities should not continue assessing the impacts of potentially exclusionary pricing schemes, but authorities must ensure they have clear theories of harm. The challenge of bringing cases in such rapidly developing markets is the time they take, and therefore enforcement must be complemented with effective advocacy where necessary.

4.2. Merger control

The AI infrastructure supply chain is experiencing rapid technological innovation and unprecedented growth. In this environment, mergers & acquisitions (M&A) and strategic partnerships have become important tools for industry participants to secure and effectively manage their supply chains. This can drive efficiency and has benefits for startups who can secure investment for growth and R&D. However, they may also raise some significant competition concerns (OECD, 2020[145]). Effective merger control is therefore important to ensure anticompetitive arrangements are kept in check.[15]

As highlighted in section 3, the sector has high concentration and increasing levels of vertical integration, investments and partnerships. This creates a need for competition authorities to carefully scrutinise deals even where firms are not directly competing. There have already been several such merger investigations in the supply chain. Box 7 highlights two recent merger investigations into acquisitions at different levels of the supply chain where authorities considered vertical and conglomerate theories of harm.

Box 7. Recent examples of non-horizontal merger investigations in AI infrastructure

Mellanox/Nvidia merger

The Nvidia–Mellanox merger, valued at USD 6.9 billion and completed in 2020, was scrutinised by several competition authorities, focusing on conglomerate theories of harm. The transaction involved Nvidia, the leading provider of GPUs used in datacentres and Mellanox, a leading provider of network interconnects in datacentres.

The European Commission cleared the deal unconditionally in December 2019, finding no horizontal overlaps and limited vertical concerns. However, it examined conglomerate theories of harm ultimately concluding that in practice even if Nvidia had the incentive and ability to leverage Mellanox's products, it would only impact a small proportion of the GPU market and therefore not materially impact the competition from rivals.

The Chinese competition authority (SAMR) conducted a year-long Phase II review and ultimately granted conditional approval in April 2020. The SAMR identified conglomerate effects in neighbouring markets where the merged entity would hold dominant shares. To mitigate risks of tying, bundling, and foreclosure, SAMR imposed behavioural remedies, including commitments to supply on FRAND terms, maintain interoperability, and protect third-party confidential information. SAMR also required Nvidia to preserve organisational separation between GPU and interconnect teams and to maintain open-source commitments. Recently the SAMR has also reportedly opened an antimonopoly case in relation to potential breaches of the commitments agreed in 2020.

ARM/Nvidia merger

The proposed ARM–Nvidia merger, announced in 2020, was a USD 40 billion deal in which Nvidia sought to acquire ARM from SoftBank. ARM is a supplier of chip architecture used widely across the tech industry, including by Nvidia's competitors. The deal raised significant competition concerns globally, particularly around the risk that Nvidia could restrict access to ARM's technology or distort competition in downstream markets like data centres.

Concerns centred on the potential harm to Nvidia's rivals through foreclosure strategies, such as limiting access to ARM's CPU intellectual property and hindering interoperability between related products. The theory was that this may ultimately benefit Nvidia's downstream activities and increase its profits. Of particular relevance to AI infrastructure, vertical and conglomerate effects were identified as a risk which, could restrict access to CPUs, network interface controllers and GPUs, impacting data transfer efficiency and server performance.

After extensive investigations and concerns raised from multiple competition authorities including the CMA, US FTC and the EU Commission, Nvidia and SoftBank abandoned the deal in February 2022.

Sources: European Commission (2019), Case M.9424 – NVIDIA / MELLANOX, https://ec.europa.eu/competition/mergers/cases/decisions/m9424_778_3.pdf; Garrod et al. (2020) Nvidia/Mellanox: China's Close Scrutiny of Semiconductor Deals Continues, https://www.akingump.com/en/insights/alerts/nvidiamellanox-chinas-close-scrutiny-of-semiconductor-deals-continues; Perrone, H. (2025), "Chips in on a merger: The Arm-Nvidia case", International Journal of Industrial Organization, Vol. 98, p. 103130, https://doi.org/10.1016/j.ijindorg.2024.103130.

As well as conglomerate and vertical concerns, acquisitions of nascent competitors and potential killer acquisitions could also present challenges for merger control in AI infrastructure. Killer acquisitions refer to the strategic purchase of nascent or potentially competitive firms with the intent to neutralise future threats, often by discontinuing their innovation or integrating them into a broader ecosystem. As highlighted in section 3, AI infrastructure markets rely on a high degree of innovation, with leading edge innovations often gaining dominant positions in a market. This can lead to incumbents seeking to acquire small startups as a defensive measure, which can undermine dynamic competition and risk incumbents entrenching their positions in key layers of the AI stack. On the other hand, given the potential need for large amounts of capital to function in the AI sector, and potential efficiencies from transactions, authorities need to carefully consider what the appropriate counterfactual would be (OECD, 2020[145]). Given the highly dynamic nature of any such assessment authorities may need to rely more on qualitative evidence, including assessing internal documents of the acquirer to understand the extent of any perceived threat. In addition, analysis on the valuation of the transaction may also provide additional information (OECD, 2020[145]).

While this report focuses on AI infrastructure markets which is primarily about physical hardware, such infrastructure markets also rely on the knowledge and skills of key staff. Recently there has also been an increase in moves to acquire key staff such as researchers, engineers and leadership in AI infrastructure markets.[16] These arrangements where companies may seek to hire teams of key personnel can significantly reshape competitive dynamics. While usually hiring and firing is not the focus of merger frameworks, competition authorities have begun to assess such practices (Federle and de Amorin, 2024[146]), especially when the acquisition of staff leads to loss of competition or potential competition.

The CMA, and Bundeskartellmt for example, have previously examined the agreement between Microsoft, a cloud service provider, and Inflection, an AI company and model develop. The CMA found that hiring of key personnel and licensing IP constated a merger, however the CMA found that the transaction was unlikely to result in a realistic prospect that there would be a substantial lessening of competition (Competition and Markets Authority, 2025[147]). The Bundeskartellamt found that the merger did not meet the national thresholds for merger reviews (Bundeskartellamt, 2024[148]).

Such acquire-hires and other transactions involving innovative startups have exposed a potential gap in existing frameworks, where strategic investments, joint ventures, or exclusive supply arrangements may escape merger control scrutiny. Former European Commission Executive Vice President Margrethe Vestager noted that some competitively significant tech acquisitions fall below EU Merger Regulation thresholds yet still warrant scrutiny (European Commission, 2024[149]).[17] Several authorities therefore have begun to introduce call in powers, giving them discretion to examine mergers even when not meeting the mandatory thresholds (Bary, 2025[150]). However, the use of such powers should be balanced carefully against the potential legal uncertainty they can create if not applied consistently or utilised excessively.

4.3. Market studies

Market studies are a flexible tool for competition authorities to analyse whether there are competition problems in a sector outside of enforcement investigations or merger control (OECD, 2018[151]). Using these tools to explore the AI infrastructure supply chain has the benefit of increasing authority staff knowledge. In addition, proactive efforts during the early stages of development can uncover issues which may be resolved before market dynamics become intractable, a common critique of competition authority responses to digital platforms in the past two decades.

There have been several market studies in relation to AI and the cloud infrastructure market including in France, Japan, Korea, the Netherlands, Denmark, the US and the UK. These market studies so far have focused on the cloud segment of the market and have found issues including (OECD, 2025[5]):

- switching barriers including the use of egress fees (exit charges)
- restrictive licenses which impose markups for using key software on rivals' platforms
- issues with cloud credits potentially making prices so low smaller rivals can't compete
- issues with bundling and tying.

The competition issues in the cloud were discussed in more detail in the recent OECD policy paper “Competition in the Provision of Cloud Computing Services” (OECD, 2025[5]) and is not repeated in detail again here.

So far, no competition authority appears to have launched a dedicated market study solely on semiconductors for AI or the supply chain for compute. Nonetheless, the topic is increasingly embedded in broader inquiries into cloud infrastructure, AI partnerships, and digital ecosystems.

France has however an ongoing study exploring access to energy by AI players and there have also been studies undertaken into generative AI which have considered the potential competition issues upstream. These studies included reports from France (Autorité de la concurrence, 2024[111]), Japan (Japan Fair Trade Commission, 2025[152]), Korea (Korean Fair Trade Commission, 2024[153]), Portugal (Autoridade de Concorrencia, 2023[154]), Canada (Competition Bureau Canada, 2025[155]) and the UK (Competition and Markets Authority, 2024[102]).

Some of the findings from these studies include:

- Promoting competition in Generative AI markets is intrinsically tied to promoting competition in these upstream markets (Autoridade de Concorrencia, 2023[154]).
- Potential concerns include the risk of abuse by IT component providers which have dominant positions, the potential for lock-in by cloud platforms and a lack of transparency in relation to investments and partnerships (Autorité de la concurrence, 2024[111]).
- A risk that if large technology companies control key AI development resources, this control could lead to exclusionary practices that limit competition (Competition Bureau Canada, 2025[155]).
- Access to compute remains key to developing and deploying AI modelling, especially for training most of the cutting-edge models. UK stakeholders reported a limited supply of AI accelerator chips and the potential importance of new, smaller AI models which may have implications for compute access needs (Competition and Markets Authority, 2024[102]).
- That there may be different competitive dynamics in AI chips for the training and inference phases of AI. In the training phase NVidias GPUs continue to hold a significant share globally due to factors like development environments and the CUDA software ecosystem. In the training phase Nvidia is likely to maintain its status as market leader, but there may be more competition longer-term at the inference phase. However, there are large switching costs from switching ecosystems and that this will may switching of GPUs even if there are comparable competitor products (Japan Fair Trade Commission, 2025[152]).
- That there are significant economies of scale and scope, potential network effects and first mover advantages in the provision of cloud services and GPU provision. These can pose significant barriers to entry Korea (Korean Fair Trade Commission, 2024[153]).

Given the importance and global significance of the supply chain, market studies could be a valuable tool to ensure competition agencies are well informed and identify potential anticompetitive conduct which might warrant investigation. Such studies can also provide the basis for recommendations to governments for ex ante interventions. While in many jurisdictions market studies only provide the ability to make recommendations, several jurisdictions also have market investigation powers which provide the ability to intervene in markets and address competition issues, without needing to find breaches of the law (OECD, 2024[4]). A lighter touch option could be to gather information by monitoring developments. This could take different forms, from simply devoting time to tracking developments, and engaging with market operators (OECD, 2024[4]). Monitoring and information gathering could be a relatively light-touch first step in understanding whether the sector requires further scrutiny. Monitoring could be complemented with access to expertise, which could be provided by specialist staff (OECD, 2024[4]).

4.4. Advocacy and co-operation

AI infrastructure is both highly dynamic and attracts a great deal government and business attention. Advocacy can play an important role in ensuring markets are shaped in as pro-competitive a manner as possible. Such advocacy may cover several areas such as engagement with market participants as well as with governments.

4.4.1. Advocacy with market participants

One disadvantage of traditional antitrust enforcement tools is that they can be slow to achieve results (von Thun and Hanley, 2024[156]). In fast-evolving AI infrastructure markets, competition advocacy could play an important role in bridging the gap between enforcement and market realities before any anti-competitive conduct risk crystalises.

Effective advocacy can help build trust and awareness among market participants, encouraging whistleblowers, complainants and leniency applicants to come forward (OECD, 2023[157]). In AI infrastructure, where abuse may be subtle or opaque to authorities due to the highly technical nature of the market, insiders will often be the best source of actionable intelligence. By engaging with developers, startups, and infrastructure providers, authorities can better understand emerging risks and competitive bottlenecks, such as discriminatory access to compute or exclusionary bundling.

Finally, advocacy can have a deterrent effect. Public statements, guidance documents and soft law instruments can signal enforcement priorities and shape firm behaviour. In AI compute, where firms may be tempted to leverage infrastructure control for exclusionary purposes, visible advocacy can discourage such conduct before it occurs. However, advocacy alone is insufficient and to ensure deterrence agencies will need to effectively enforce against anticompetitive conduct if it is occurring.

4.4.2. Advocating with other parts of government

As discussed in section 2, governments and authorities will potentially be looking to implement laws and regulations affecting AI infrastructure through several policy angles. For example, AI infrastructure's environmental impacts through energy use and water consumption.

Competition authorities can play a role ensuring policy interventions are proportionate. Very rigid or prescriptive rules, especially those that favour large, vertically integrated firms can raise barriers for smaller players and new entrants in the AI compute ecosystem. Advocacy can help shape regulation that is risk-based, proportionate, and innovation-friendly, ensuring that other policy goals and competition are pursued in tandem. The OECD Competition Assessment Toolkit (OECD, 2019[158]) provides a framework competition agencies and governments can use to review proposed laws and regulations and promote more competition in their economies, leading to lower prices, greater choice and higher quality of goods and services.

As discussed in detail in section 2, various jurisdictions have passed laws to spur domestic manufacturing of chips or imposed export controls. Competition authorities also have a role to play in helping to ensure that state decisions on investments and trade policy are as pro-competitive as possible, seeking to integrate competitive neutrality in decisions where possible (OECD, 2024[159]). Authorities can also advocate to ensure governments making such large procurement and investment decisions are fully prepared for managing bid processes free from collusion (OECD, 2025[136]).

4.4.3. Co-operation

Almost all of the AI infrastructure chain is made up of global markets. The cross-border nature of the supply chain means that enforcement actions, merger reviews and policy interventions in one jurisdiction can have global implications. Accordingly, competition authorities around the world are recognising the need for closer international co-operation. For example, the EU, UK CMA, US DOJ and US FTC recently issued a joint statement on AI models and foundations. It noted risks to competition relevant to AI infrastructure, including the concentration of key inputs such as specialised chips and substantial compute. It also included principles for protecting competition in the AI ecosystem including fair dealing, interoperability and choice (EU, CMA, US FTC & US DOJ, 2024[160]). Similarly, the G7 issued a Digital Competition Communiqué which also highlighted the potential for competitive bottlenecks in availability and access to key inputs including compute infrastructure and specialised chips (G7 Italia, 2024[161]). It highlighted principles such as fair access to key inputs like AI chips and called for vigorous antitrust enforcement and enhanced cooperation (G7 Italia, 2024[161]).

As AI infrastructure markets evolve, deeper co-operation will be essential to ensure that merger remedies, abuse investigations, and regulatory interventions are effective and consistent across jurisdictions. Divergent approaches can raise compliance costs for business, stifle innovation and exacerbate the risk of competition issues persisting in some regions.

In addition to working to minimise unnecessary divergence, the consistency of the markets and complexity of the technologies provides opportunities for authorities to share their knowledge, analysis and expertise (OECD, 2024[162]).

Given the interdependencies between AI infrastructure and other regulated sectors, such as energy, telecommunications, and data governance, competition authorities may also need to increasingly co-ordinate with other domestic agencies. The development and operation of data centres, for example, relies heavily on access to stable and scalable energy supplies, while high-performance compute often depends on advanced connectivity and bandwidth provision regulated by telecommunication authorities. As AI infrastructure scales, issues such as grid access, spectrum allocation, and environmental impacts may intersect with competition concerns, particularly where infrastructure bottlenecks or preferential access arrangements arise. Co-ordinated oversight can help ensure that enforcement and policy interventions are coherent, avoid regulatory gaps and support a level playing field across sectors.

4.5. Potential regulation

As highlighted in section 3, the AI infrastructure supply chain is complex and dynamic, which can make enforcement challenging. In particular, abuse of dominance cases can be complex and time-consuming (OECD, 2021[163]). In fast-moving sectors like AI infrastructure which is evolving rapidly, such delays risk rendering enforcement ineffective, underscoring for some the potential need for more agile and anticipatory regulatory tools (Narechania and Sitaraman, 2023[45]).

In recent years several jurisdictions have moved to implement ex-ante regulatory regimes in digital markets.[18] Each jurisdiction’s legislation is drafted differently but given these regimes were typically designed in response to digital platforms rather than physical infrastructure or hardware. Therefore, many of the existing provisions will not cover AI infrastructure (beyond the potential partial coverage of cloud platforms) (OECD, 2025[5]).

There have already been calls from some for similar ex ante regulatory interventions in AI infrastructure. It has been suggested, for example, that competition could be improved by introducing access obligations on dominant infrastructure providers (such as chip producers) (Narechania and Sitaraman, 2023[45]). The idea being that such provisions would prevent firms which are supplying key inputs from favouring some customers over others, promoting fair access and competition downstream. Such regulations have typically been seen in other more traditional infrastructure sectors such as telecoms where monopoly providers have often been subject to Fair Reasonable and Non-Discriminatory (FRAND) provisions or in the creation of standards in industries with patents (Silva, 2025[164]). It is unclear whether such traditional regulatory approaches would be appropriate for AI infrastructure given the rapidity of technical developments. In addition, while such provisions extend the circumstances in which suppliers must not discriminate, enforcing such provisions may have potential challenges. For example, defining terms such as ‘fair’ and ‘reasonable’ in an industry with such large outlays in research and development.

Standardisation and interoperability have also been proposed as a way to encourage effective competition in the AI stack (von Thun and Hanley, 2024[156]). This has primarily been in the context of promoting interoperability in cloud computing. However, there are also issues of interoperability in AI infrastructure if there is a need to switch between different hardware components (e.g. switching chipsets). Moves towards standards and interoperable systems can have benefits for competition in relation to reduced switching costs, but have potential risks. For example, standards may simply further embed the systems supported by dominant providers (Perez, 2017[165]). Given these challenges and the continuing rapid market developments, authorities should consider exercising caution in pushing for standards early in the market’s development. An alternative approach could be to consider advocating for government support for open-source technologies to be developed which may provide alternatives to proprietary technology, serving as the potential base for future standards.

More broadly, while regulations can potentially spur competition and innovation, they also impose costs on market participants and may have unanticipated consequences. Some commentators suggest that early intervention may be unnecessary. Hagiu and Wright (2025[166]) for example argue that each large player with a strong position in one layer has a strong incentive to commoditise the other layers, especially the ones where their market position is weaker. Hagiu and Wright (2025[166]) suggest that there may end up being at least seven major players operating at multiple (possibly, all) levels of the AI stack which would be a significant improvement compared to other digital markets.

Before considering an intervention, authorities may need stronger evidence that the risk of firms preventing access or distorting competition is sufficiently high and that traditional enforcement tools are not sufficiently timely to resolve the concerns (Meyers and Bourreau, 2025[126]).

5 Conclusion

AI is an increasingly critical general-purpose technology with disruptive potential across a wide range of markets. The infrastructure supply chain that supports AI, including data centres and the compute resources they house, are attracting unprecedented levels of investment.

The AI infrastructure supply chain is characterised by both rapid innovation and several structural features with potential competition issues. On one hand, firms across the chain from semiconductor design to cloud-based compute services, are investing heavily in research, development, and deployment, driving technological progress at a fast pace. At the same time, the supply chain exhibits high barriers to entry at many levels due to the capital intensity, technical complexity and economies of scale involved. Several layers of the chain are highly concentrated, and there is a growing trend toward vertical and conglomerate integration, as firms seek to control multiple stages of the value chain. Strategic partnerships and exclusive arrangements are also becoming more common, potentially reinforcing existing positions and shaping access to critical inputs.

In addition to these structural features, the strategic significance of AI to national competitiveness and security has led to increasing levels of government intervention across the supply chain. Public investment, regulatory initiatives, and industrial policy measures are shaping market dynamics in areas such as chip fabrication, data centre development and access to compute. While such interventions may aim to support innovation and resilience, they also have the potential to influence competitive conditions, particularly where they intersect with existing market concentration or reinforce strong market positions.

To date, competition authorities have primarily focused their market studies on the cloud services and AI modelling layers of the value chain. However, there is scope for deeper analysis of upstream segments which would enhance understanding of how infrastructure influences competitive dynamics. Given the global nature of these markets, competition authorities should consider whether it is possible to co-ordinate such efforts. While enforcement activity in this area remains limited, several initial investigations have been launched. Authorities also face strategic choices in how to assess the growing number of partnerships, joint ventures and long-term agreements emerging across the supply chain, including whether these arrangements warrant scrutiny under merger control frameworks or as potential anti-competitive agreements.

Looking ahead, competition authorities may need to consider the full extent of their policy and enforcement powers as they engage with this rapidly evolving market. In the early stages, tools such as advocacy, deterrence and merger control may play a particularly important role in shaping market behaviour and guiding industry development. However, policymakers will also need to assess whether existing powers are sufficient to address potential abuses of dominance, particularly where exclusionary effects may arise. This could include evaluating the need for ex ante measures, such as interoperability requirements, to complement traditional enforcement. Moreover, while concerns around exclusion and market power may attract significant attention, authorities should remain vigilant in other areas. The scale of infrastructure construction underway suggests that risks such as bid rigging may become more salient, and as technologies mature and commoditise, the potential for price co-ordination or cartel behaviour may also increase.

References

Acton, M. (2025), *Inside Intel's big bet to save US chipmaking — and itself*, https://www.ft.com/content/e4b1fa29-6027-44d9-8375-b3368132dd1b. [43]

Aguirre, T. (2024), *On Labs and Fabs: Mapping How Alliances, Acquisitions, and Antitrust are Shaping the Frontier AI Industry*, https://doi.org/10.48550/arXiv.2406.01722. [42]

Aliaga, S. (2025), *Is AI already driving U.S. growth?*, https://am.jpmorgan.com/us/en/asset-management/adv/insights/market-insights/market-updates/on-the-minds-of-investors/is-ai-already-driving-us-growth/ (accessed on October 2025). [13]

Amazon (2024), *A city in the palm of your hand*, https://www.aboutamazon.com/stories/ai-chips-aws-trainium2-explain (accessed on September 2025). [178]

Amazon (2024), *Amazon and Anthropic deepen strategic collaboration*, https://www.aboutamazon.com/news/aws/amazon-invests-additional-4-billion-anthropic-ai (accessed on September 2025). [56]

AMD (2025), *AMD and OpenAI Announce Strategic Partnership to Deploy 6 Gigawatts of AMD GPUs*, https://ir.amd.com/news-events/press-releases/detail/1260/amd-and-openai-announce-strategic-partnership-to-deploy-6-gigawatts-of-amd-gpus (accessed on October 2025). [29]

Ansari, S. (2025), "Navigating AI partnerships between big and small tech firms", https://www.jbs.cam.ac.uk/2025/navigating-ai-partnerships-between-big-and-small-tech-firms/. [97]

Arrow, K. (1962), "Economic Welfare and the Allocation of", *National Bureau of Economic Research*, https://www.nber.org/system/files/chapters/c2144/c2144.pdf. [84]

Asmat, D. (2019), "Collusion Along the Learning Curve: Theory and Evidence from the Semiconductor Industry", *SSRN Electronic Journal*, https://doi.org/10.2139/ssrn.3470229. [130]

ASML (2025), *ASML, Mistral AI enter strategic partnership*, https://www.asml.com/en/news/press-releases/2025/asml-mistral-ai-enter-strategic-partnership (accessed on September 2025). [48]

ASML (2016), *ZEISS and ASML strengthen partnership for next generation of EUV Lithography due in early 2020s*, ASML, https://www.asml.com/en/news/press-releases/2016/zeiss-and-asml-strengthen-partnership-for-next-generation-of-euv-lithography (accessed on 11 August 2025). [47]

ASML (n.d.), *HMI*, ASML, https://www.asml.com/en/company/about-asml/hmi (accessed on 11 August 2025). [46]

Auer, D. and M. Zúñiga (2025), *AI Partnerships and Competition: Damned if You Buy, Damned if You Don't*, ICLE white paper, https://laweconcenter.org/resources/ai-partnerships-and-competition-damned-if-you-buy-damned-if-you-dont/. [98]

Autoridade de Concorrencia (2023), *Competition and Generative Artificial Intelligence*, [154] https://www.concorrencia.pt/sites/default/files/documentos/Issues%20Paper%20-%20Competition%20and%20Generative%20Artificial%20Intelligence.pdf.

Autorité de la concurrence (2024), *Generative artificial intelligence: the Autorité issues its opinion on the competitive functioning of the sector*, https://www.autoritedelaconcurrence.fr/en/press-release/generative-artificial-intelligence-autorite-issues-its-opinion-competitive (accessed on August 2025). [111]

Autorité de la concurrence (2023), *Opinion 23-A-08 of 29 June 2023 on competition in the cloud sector*, Autorité de la concurrence, https://www.autoritedelaconcurrence.fr/en/opinion/competition-cloud-sector (accessed on 22 January 2025). [53]

Bary, L. (2025), *Merger control and the rise of national call-in powers in the EU: what does it mean for M&A transactions?*, https://www.financierworldwide.com/merger-control-and-the-rise-of-national-call-in-powers-in-the-eu-what-does-it-mean-for-ma-transactions. [150]

Beck, M. and F. Scott Morton (2021), "Evaluating the Evidence on Vertical Mergers", *Review of Industrial Organization*, Vol. 59/2, pp. 273-302, https://doi.org/10.1007/s11151-021-09832-z. [94]

Bowman, J. (2024), *This 1 Number May Ensure TSMC's Market Dominance*, NASDAQ - The Motley Fool, https://www.fool.com/investing/2024/08/17/this-1-number-may-ensure-tsmcs-market-dominance/ (accessed on September 2025). [40]

Bradshaw, T. (2024), *Nvidia's rivals take aim at its software dominance*, https://www.ft.com/content/320f35de-9a6c-4dbf-b42f-9cdaf35e45bb (accessed on September 2025). [30]

Bundeskartellamt (2024), *Taking over employees may be subject to merger control in Germany – Bundeskartellamt not competent to review Microsoft/Inflection transaction as Inflection has no substantial operations in Germany*, https://www.bundeskartellamt.de/SharedDocs/Meldung/EN/Pressemitteilungen/2024/29_11_2024_Microsoft.html?nn=55030 (accessed on August 2025). [148]

Burja, S. (2024), *Tokyo Electron's Place in Global Semiconductor Manufacturing*, https://brief.bismarckanalysis.com/p/tokyo-electrons-place-in-global-semiconductor (accessed on October 2025). [83]

Business Software Allliance (2005), *Open source and Commerical Software*, https://www.wipo.int/edocs/mdocs/copyright/en/wipo_ip_cm_07/wipo_ip_cm_07_www_82575.pdf. [78]

Calvino, F., D. Haerle and S. Liu (2025), "Is generative AI a General Purpose Technology?: Implications for productivity and policy", *OECD Artificial Intelligence Papers*, No. 40, OECD Publishing, Paris, https://doi.org/10.1787/704e2d12-en. [1]

Calvino, F., J. Reijerink and L. Samek (2025), "The effects of generative AI on productivity, innovation and entrepreneurship", *OECD Artificial Intelligence Papers*, No. 39, OECD Publishing, Paris, https://doi.org/10.1787/b21df222-en. [6]

Center for Security and Emerging Technology (2020), *AI Chips: What They Are and Why They Matter*, Center for Security and Emerging Technology, https://cset.georgetown.edu/publication/ai-chips-what-they-are-and-why-they-matter/. [22]

Chen, S. (2025), *Data centres will use twice as much energy by 2030 — driven by AI*, https://www.nature.com/articles/d41586-025-01113-z (accessed on September 2025). [17]

CNBC (2025), *Nvidia just spent over $900 million to hire Enfabrica CEO, license AI startup's technology*, https://www.cnbc.com/2025/09/18/nvidia-spent-over-900-million-on-enfabrica-ceo-ai-startup-technology.html. [180]

Competition and Markets Authority (2025), *Cloud Infrastructure Services - Final decision report*, https://www.gov.uk/cma-cases/cloud-services-market-investigation (accessed on 11 August 2025). [64]

Competition and Markets Authority (2025), *Microsoft / Inflection inquiry*, https://www.gov.uk/cma-cases/microsoft-slash-inflection-ai-inquiry. [147]

Competition and Markets Authority (2024), *AI Foundation Models - update paper*, https://assets.publishing.service.gov.uk/media/661941a6c1d297c6ad1dfeed/Update_Paper__1_.pdf. [102]

Competition Bureau Canada (2025), *Consultation on Artificial Intelligence and Competition: What We Heard*, https://competition-bureau.canada.ca/en/how-we-foster-competition/education-and-outreach/consultation-artificial-intelligence-and-competition-what-we-heard. [155]

Court of Justice of the European Union (2024), *The Court of Justice upholds the annulment by the General Court of the Commission's decision finding an abuse of a dominant position on the part of Intel and imposing a fine of €1.06 billion on Intel*, https://curia.europa.eu/jcms/upload/docs/application/pdf/2024-10/cp240185en.pdf. [144]

Cusumano, M. (2024), *Intel's Fall from Grace*, https://cacm.acm.org/opinion/intels-fall-from-grace/ (accessed on September 2025). [32]

da Silva, J. (2024), "Google turns to nuclear to power AI data centres", *BBC News*, https://www.bbc.com/news/articles/c748gn94k95o (accessed on 11 August 2025). [67]

Davern, M. and M. Pinnuck (2025), *DeepSeek shatters beliefs about the cost of AI, leaving US tech giants reeling*, https://theconversation.com/deepseek-shatters-beliefs-about-the-cost-of-ai-leaving-us-tech-giants-reeling-248424 (accessed on September 2025). [172]

Davies, C. (2025), *Why memory chips are the new frontier of the AI revolution*, https://www.ft.com/content/f3ee292b-ba56-4e9f-944a-da26d5706583 (accessed on September 2025). [37]

EU, CMA, US FTC & US DOJ (2024), *Joint Statement on Competition in Generative AI Foundation Models and AI Products*, https://www.ftc.gov/system/files/ftc_gov/pdf/ai-joint-statement.pdf. [160]

European Commission (2025), *Commission approves €920 million German State aid measure to support Infineon in setting up a new semiconductor manufacturing facility*, https://ec.europa.eu/commission/presscorner/detail/en/ip_25_557. [114]

European Commission (2025), *EU launches InvestAI initiative to mobilise €200 billion of investment in artificial intelligence*, https://ec.europa.eu/commission/presscorner/detail/en/ip_25_467 (accessed on August 2025). [115]

European Commission (2024), *Case M.11766 - NVIDIA / RUN:AI*, https://ec.europa.eu/competition/mergers/cases1/202516/M_11766_10599589_2740_3.pdf. [89]

European Commission (2024), *Commission carries out unannounced antitrust inspections in the data centre construction sector*, https://ec.europa.eu/commission/presscorner/detail/en/ip_24_5926. [173]

European Commission (2024), *Competition Policy Brief - Competition in Generative AI and Virtual Worlds*, https://competition-policy.ec.europa.eu/document/download/c86d461f-062e-4dde-a662-15228d6ca385_en. [108]

European Commission (2024), *Speech by EVP Margrethe Vestager at the webinar on generative AI organised by the French 'Autorité de la Concurrence'*, https://ec.europa.eu/commission/presscorner/detail/en/speech_24_4966. [149]

European Commission (2022), *A Chip Act for Europe*, https://www.decision.eu/wp-content/uploads/2022/09/A-Chips-Act-for-Europe-Commission-Staff-Working-Document.pdf. [119]

European Commission (2019), *Case M.9424 - NVIDIA / MELLANOX*, https://ec.europa.eu/competition/mergers/cases/decisions/m9424_778_3.pdf. [93]

European Union (2025), *Case T-23/22: Order of the General Court of 13 January 2025 – Grail v Commission*, https://eur-lex.europa.eu/eli/C/2025/1769/oj/eng. [175]

European Union (2018), *Antitrust: Commission fines Google €4.34 billion for illegal practices regarding Android mobile devices to strengthen dominance of Google's search engine*, https://ec.europa.eu/commission/presscorner/detail/en/ip_18_4581 (accessed on September 2025). [142]

European Union (2011), "Final report of the Hearing Officer - COMP/38.511 - DRAMS", *Official Journal of the European Union*, https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52011XX0621(02). [131]

European Union (2007), *Microsoft Corp V Commission of The European Communities*, https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex:62004TJ0201. [75]

European Union (2007), *Microsoft Corp. v Commission of the European Communities.*, https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex:62004TJ0201. [141]

Farooque, M. (2025), *Nvidia Dominates AI Chips: Analyst Sees 85% Market Share Amid Fierce Competition*, https://finance.yahoo.com/news/nvidia-dominates-ai-chips-analyst-143411443.html?guccounter=1 (accessed on September 2025). [24]

Federal Trade Commission (2025), *Partnerships Between Cloud Service Providers and Ai Developers - FTC staff Report on AI Partnerships and Investments 6 (b) Study*, https://www.ftc.gov/system/files/ftc_gov/pdf/p246201_aipartnerships6breport_redacted_0.pdf. [100]

Federle, A. and A. de Amorin (2024), *EU: When does the hiring of another company's staff require merger control approval?*, https://competitionlawinsights.twobirds.com/post/102jp79/eu-when-does-the-hiring-of-another-companys-staff-require-merger-control-approv (accessed on August 2025). [146]

Filippucci, F. et al. (2024), “The impact of Artificial Intelligence on productivity, distribution and growth: Key mechanisms, initial evidence and policy challenges”, *OECD Artificial Intelligence Papers*, No. 15, OECD Publishing, Paris, https://doi.org/10.1787/8d900037-en. [60]

Filippucci, F. et al. (2025), “Macroeconomic productivity gains from Artificial Intelligence in G7 economies”, *OECD Artificial Intelligence Papers*, No. 41, OECD Publishing, Paris, https://doi.org/10.1787/a5319ab5-en. [7]

Fortuna, G. (2025), *EU antitrust fines food delivery giants in landmark cartel case*, https://www.euronews.com/my-europe/2025/06/02/eu-antitrust-fines-food-delivery-giants-in-landmark-cartel-case (accessed on September 2025). [104]

France (2025), *Make France an AI powerhouse.*, https://www.elysee.fr/en/emmanuel-macron/2025/02/11/make-france-an-ai-powerhouse (accessed on August 2025). [124]

Fumagalli, C. and M. Motta (2020), “Dynamic Vertical Foreclosure”, *The Journal of Law and Economics*, Vol. 63/4, pp. 763-812, https://doi.org/10.1086/709584. [95]

G7 Italia (2024), *G7 Competition Authorities and Policy Makers’Summit - Digital Competition Commmunique*, https://www.autoritedelaconcurrence.fr/sites/default/files/2024-10/G72024DigitalCompetitionCommunique%CC%81.pdf. [161]

Gambacorta, L. and V. Shreeti (2025), *The AI supply chain*, https://www.bis.org/publ/bppdf/bispap154.htm. [3]

Godoy, J. (2024), *Microsoft faces wide-ranging US antitrust probe*, https://www.reuters.com/technology/microsoft-faces-wide-ranging-us-antitrust-probe-2024-11-27/ (accessed on August 2025). [76]

Google (2024), *TPU transformation: A look back at 10 years of our AI-specialized chips*, https://cloud.google.com/transform/ai-specialized-chips-tpu-history-gen-ai (accessed on September 2025). [90]

Groza, T. and A. Wierzbicka (2024), “<div>”, *SSRN Electronic Journal*, https://doi.org/10.2139/ssrn.4967357. [99]

Hagiu, A. and J. Wright (2025), “Artificial intelligence and competition policy”, *International Journal of Industrial Organization*, https://doi.org/10.1016/j.ijindorg.2025.103134. [166]

Ha, H. and V. Law (2025), *New frontiers in antitrust enforcement: Minority stakes, labour markets and ESG coordination under scrutiny*, https://www.jsm.com/publications/2025/new-frontiers-in-antitrust-enforcement-minority-stakes-labour-markets-and-esg-coordination-under-scrutiny/ (accessed on September 2025). [103]

Haramboure, A. et al. (2023), “Vulnerabilities in the semiconductor supply chain”, *OECD Science, Technology and Industry Working Papers*, No. 2023/05, OECD Publishing, Paris, https://doi.org/10.1787/6bed616f-en. [9]

Harris, D. (2025), *AI Factories Are Redefining Data Centers and Enabling the Next Era of AI*, https://blogs.nvidia.com/blog/ai-factory/ (accessed on September 2025). [49]

Intel (2024), *The Transistor, Explained*, https://newsroom.intel.com/tech101/the-transistor-explained. [19]

Intel Corp (2025), *Intel and Trump Administration Reach Historic Agreement to Accelerate* [116]

American Technology and Manufacturing Leadership, https://www.intc.com/news-events/press-releases/detail/1748/intel-and-trump-administration-reach-historic-agreement-to (accessed on August 2025).

Jackson, F. (2024), *Global Chip Shortage: Everything You Need to Know*, https://www.techrepublic.com/article/global-chip-shortage-cheat-sheet/ (accessed on September 2025). [107]

Japan Fair Trade Commission (2025), *Report regarding Generative AI*, https://www.jftc.go.jp/file/250606.pdf. [152]

Japan Fair Trade Commission (2022), *Report on Fact-Finding Survey on Trade Practices by Digital Platform Operators Report on Trade Practices in Cloud Services Sector Contents*, https://www.jftc.go.jp/en/pressreleases/yearly-2022/June/221102EN.pdf. [81]

Kedrosky, P. (2025), *Honey, AI Capex is Eating the Economy*, Paul Kedrosky, https://paulkedrosky.com/honey-ai-capex-ate-the-economy/ (accessed on 11 August 2025). [12]

Kharpal, A. (2025), *China says Nvidia violated anti-monopoly law after preliminary probe*, https://www.cnbc.com/2025/09/15/china-nvidia-violated-anti-monopoly-law-will-continue-investigation.html (accessed on September 2025). [179]

King, I. and L. Nylen (2024), *Nvidia Gets DOJ Subpoena in Escalating Antitrust Probe*, https://www.bloomberg.com/news/articles/2024-09-03/nvidia-gets-doj-subpoena-in-escalating-antitrust-investigation?embedded-checkout=true (accessed on August 2025). [169]

Klimczak, S. (2024), *The Convergence of Data Centers and Power: A Generational Investment Opportunity | The Connection*, Blackstone, https://www.blackstone.com/insights/article/the-convergence-of-data-centers-and-power-a-generational-investment-opportunity/ (accessed on 11 August 2025). [66]

Klotz, A. (2025), *Huawei is making its Ascend AI GPU software toolkit open-source to better compete against CUDA*, https://www.tomshardware.com/tech-industry/artificial-intelligence/huawei-is-making-its-ascend-ai-gpu-software-toolkit-open-source-to-better-compete-against-cuda (accessed on August 2025). [79]

Korea Fair Trade Commission (2022), *Announcement of the results of the cloud field survey*, https://www.ftc.go.kr/www/selectBbsNttView.do?pageUnit=10&pageIndex=1&searchCnd=all&searchKrwd=%ED%81%B4%EB%9D%BC%EC%9A%B0%EB%93%9C&key=12&bordCd=3&searchCtgry=01,02&nttSn=42705. [82]

Korean Fair Trade Commission (2024), *Generative AI and Competition*, https://www.ftc.go.kr/eng/selectBbsNttView.do?key=564&bordCd=822&nttSn=13613. [153]

LeCun, Y., Y. Bengio and G. Hinton (2015), “Deep learning”, *Nature*, Vol. 521/7553, pp. 436-444, https://doi.org/10.1038/nature14539. [20]

Lehdonvirta, V. et al. (2025), “Measuring domestic public cloud compute availability for artificial intelligence”*, OECD Artificial Intelligence Papers*, No. 49, OECD Publishing, Paris, https://doi.org/10.1787/8602a322-en. [52]

Leng, C., Z. Wu and T. Bradshaw (2025), *China bans tech companies from buying Nvidia’s AI chips*, https://www.ft.com/content/12adf92d-3e34-428a-8d61-c9169511915c (accessed on September 2025). [125]

Mckinsey (2025), *AI infrastructure: A new growth avenue for telco operators*, https://www.mckinsey.com/industries/technology-media-and-telecommunications/our-insights/ai-infrastructure-a-new-growth-avenue-for-telco-operators (accessed on September 2025). [16]

Mell, P. and T. Grance (2011), *The NIST definition of cloud computing*, National Institute of Standards and Technology, Gaithersburg, MD, https://doi.org/10.6028/NIST.SP.800-145. [51]

Meta (2024), *Introducing Our Next Generation Infrastructure for AI*, https://about.fb.com/news/2024/04/introducing-our-next-generation-infrastructure-for-ai/ (accessed on September 2025). [91]

Meyers, Z. and M. Bourreau (2025), *A Competition Policy for Cloud and AI*, https://cerre.eu/wp-content/uploads/2025/06/A-Competition-Policy-for-Cloud-and-AI_FINAL.pdf. [126]

Mickle, T. (2025), "Nvidia Becomes First Public Company Worth $4 Trillion", *The New York Times*, https://www.nytimes.com/2025/07/10/technology/nvidia-4-trillion-market-value.html?smid=url-share (accessed on 11 August 2025). [26]

Miller, C. (2022), *Chip war: the fight for the world's most critical technology*, Simon & Schuster,. [113]

Mims, C. (2025), "Silicon Valley's New Strategy: Move Slow and Build Things", *The Wall Street Journal*, https://www.wsj.com/tech/ai/silicon-valley-ai-infrastructure-capex-cffe0431 (accessed on 11 August 2025). [11]

Mistral AI (2025), *Mistral Compute*, https://mistral.ai/news/mistral-compute. [59]

Montgomery, B. and N. Robins-Early (2025), *Nvidia becomes world's first $5tn company amid stock market and AI boom*, https://www.theguardian.com/technology/2025/oct/29/nvidia-first-company-5-trillion (accessed on 2025). [27]

Mordor Intelligence (n.d.), *Outsourced Semiconductor Assembly and Test (OSAT) Market Size & Share Analysis - Growth Trends & Forecasts (2025 - 2030)*, https://www.mordorintelligence.com/industry-reports/osat-market (accessed on 11 August 2025). [44]

Motta (2004), "Collusion and Horizontal Agreements", in *Competition Policy*, Cambridge University Press, https://doi.org/10.1017/cbo9780511804038.005. [129]

Mytton, D. (2021), "Data centre water consumption", *npj Clean Water*, Vol. 4/1, https://doi.org/10.1038/s41545-021-00101-w. [72]

Nakazawa, A. and M. Pisa (2025), *How governments are driving AI adoption for economic growth*, https://oecd.ai/en/wonk/how-governments-are-driving-ai-adoption-for-economic-growth (accessed on August 2025). [110]

Narechania, T. and G. Sitaraman (2023), "An Antimonopoly Approach to Governing Artificial Intelligence", *Yale Law & Policy Review*, Vol. 4/1, https://doi.org/10.2139/ssrn.4597080. [45]

Netherlands Authority for Consumers and Markets (2022), *Market Study Cloud services*, Netherlands Authority for Consumers and Markets, https://www.acm.nl/en/publications/market-study-cloud-services. [70]

Nicol-Schwarz, K. (2025), *Graphcore cofounder exits company one year on from SoftBank acquisition*, https://sifted.eu/articles/graphcore-cofounder-exits-company-one-year-on-from-softbank-acquisition (accessed on September 2025). [35]

NVIDIA (2025), *NVIDIA Announces Financial Results for Fourth Quarter and Fiscal 2025*, https://nvidianews.nvidia.com/news/nvidia-announces-financial-results-for-fourth-quarter-and-fiscal-2025 (accessed on September 2025). [25]

O'Donnell, J. and C. Crownhart (2025), *We did the math on AI's energy footprint. Here's the story you haven't heard.*, https://www.technologyreview.com/2025/05/20/1116327/ai-energy-usage-climate-footprint-big-tech/ (accessed on 11 August 2025). [65]

OECD (2025), "Competition in the provision of cloud computing services", *OECD Roundtables on Competition Policy Papers*, No. 323, OECD Publishing, Paris, https://doi.org/10.1787/595859c5-en. [5]

OECD (2025), "Mapping the semiconductor value chain: Working towards identifying dependencies and vulnerabilities", *OECD Science, Technology and Industry Policy Papers*, No. 182, OECD Publishing, Paris, https://doi.org/10.1787/4154cdbf-en. [8]

OECD (2025), *OECD Guidelines for Fighting Bid Rigging in Public Procurement (2025 Update)*, OECD Publishing, Paris, https://doi.org/10.1787/cbe05a56-en. [136]

OECD (2025), *Policy Brief - Recent trends in semiconductor subsidies*, https://www.oecd.org/en/publications/recent-trends-in-semiconductor-subsidies_5e91af33-en.html. [118]

OECD (2025), *Recommendation of the Council on Merger Review*, https://legalinstruments.oecd.org/public/doc/195/195.en.pdf. [174]

OECD (2024), "Artificial intelligence, data and competition", *OECD Artificial Intelligence Papers*, No. 18, OECD Publishing, Paris, https://doi.org/10.1787/e7e88884-en. [4]

OECD (2024), *Challenges and sources of divergence in cross-border merger review*, OECD Publishing, Paris, https://doi.org/10.1787/61972545-en. [162]

OECD (2024), *Competitive Neutrality Toolkit: Promoting a Level Playing Field*, OECD Publishing, Paris, https://doi.org/10.1787/3247ba44-en. [159]

OECD (2024), "Financing broadband networks of the future", *OECD Digital Economy Papers*, No. 365, OECD Publishing, Paris, https://doi.org/10.1787/eafc728b-en. [61]

OECD (2024), "The use of structural presumptions in antitrust", *OECD Roundtables on Competition Policy Papers*, No. 317, OECD Publishing, Paris, https://doi.org/10.1787/3b8c6885-en. [139]

OECD (2023), "A blueprint for building national compute capacity for artificial intelligence", *OECD Digital Economy Papers*, No. 350, OECD Publishing, Paris, https://doi.org/10.1787/876367e3-en. [2]

OECD (2023), *Communication by Competition Authorities: Objectives and Tools, OECD*, OECD Publishing, Paris, https://doi.org/10.1787/944bbdac-en. [157]

OECD (2023), "Competition and Innovation: A Theoretical Perspective", *OECD Roundtables on Competition Policy Papers*, No. 294, OECD Publishing, Paris, [85] https://doi.org/10.1787/4632227c-en.

OECD (2022), "Measuring the environmental impacts of artificial intelligence compute and applications: The AI footprint", *OECD Digital Economy Papers*, No. 341, OECD Publishing, Paris, https://doi.org/10.1787/7babf571-en. [71]

OECD (2022), "Subsidies, Competition and Trade", *OECD Roundtables on Competition Policy Papers*, No. 289, OECD Publishing, Paris, https://doi.org/10.1787/57a34bc6-en. [112]

OECD (2022), "The Evolving Concept of Market Power in the Digital Economy", *OECD Roundtables on Competition Policy Papers*, No. 278, OECD Publishing, Paris, https://doi.org/10.1787/2cfcb4a8-en. [137]

OECD (2021), "Abuse of dominance in Digital Markets", *OECD Roundtables on Competition Policy Papers*, No. 256, OECD Publishing, Paris, https://doi.org/10.1787/4c36b455-en. [163]

OECD (2021), "Economic Analysis and Evidence in Abuse Cases", *OECD Roundtables on Competition Policy Papers*, No. 269, OECD Publishing, Paris, https://doi.org/10.1787/63e6d5f0-en. [138]

OECD (2021), "Efficiency Analysis in Vertical Restraints", *OECD Roundtables on Competition Policy Papers*, No. 254, OECD Publishing, Paris, https://doi.org/10.1787/7108965c-en. [96]

OECD (2020), "Start-ups, Killer Acquisitions and Merger Control", *OECD Roundtables on Competition Policy Papers*, No. 248, OECD Publishing, Paris, https://doi.org/10.1787/dac52a99-en. [145]

OECD (2019), *Competition Assessment*, https://www.oecd.org/en/topics/sub-issues/competitive-and-fair-markets/competition-assessment.html. [158]

OECD (2019), "Competition For-the-market", *OECD Roundtables on Competition Policy Papers*, No. 238, OECD Publishing, Paris, https://doi.org/10.1787/f0a896a0-en. [87]

OECD (2019), "Licensing of IP Rights and Competition Law", *OECD Roundtables on Competition Policy Papers*, No. 230, OECD Publishing, Paris, https://doi.org/10.1787/6a74221e-en. [74]

OECD (2019), "Measuring distortions in international markets: The semiconductor value chain", *OECD Trade Policy Papers*, No. 234, OECD Publishing, Paris, https://doi.org/10.1787/8fe4491d-en. [63]

OECD (2019), "Recommendation of the Council on Artificial Intelligence", *Compendium of Legal Instruments of the OECD*, https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0449. [10]

OECD (2018), *Market Studies Guide for Competition Authorities*, OECD Publishing, Paris, https://doi.org/10.1787/7381b582-en. [151]

OECD (2018), *Summary of the workshop on cartel screening in the digital era*, https://one.oecd.org/document/DAF/COMP/M(2018)3/En/pdf. [128]

OECD (2017), *Network Effects and Efficiencies in Multisided Markets - Paper by H. Shelanski, S.Knox and A. Dhilla*, https://one.oecd.org/document/DAF/COMP/WD(2017)40/FINAL/en/pdf. [106]

OECD (2016), “Fidelity Rebates”, *OECD Roundtables on Competition Policy Papers*, No. 188, OECD Publishing, Paris, https://doi.org/10.1787/fd9e16da-en. [143]

OECD (2014), “Cloud Computing: The Concept, Impacts and the Role of Government Policy”, *OECD Digital Economy Papers*, No. 240, OECD Publishing, Paris, https://doi.org/10.1787/5jxzf4lcc7f5-en. [50]

OECD (2006), *Evidentiary Issues in Proving Dominance*, OECD Publishing, Paris, https://doi.org/10.1787/806588ad-en. [183]

OECD (1996), “Abuse of Dominance and Monopolisation : Key findings, summary and notes”, *OECD Roundtables on Competition Policy Papers*, No. 8, OECD Publishing, Paris, https://doi.org/10.1787/8f43c7ef-en. [167]

OECD (Forthcoming), *Artificial intelligence and competitive dynamics in downstream markets*. [182]

Ofcom (2023), *Cloud services market study - Final report*, Ofcom, https://www.ofcom.org.uk/siteassets/resources/documents/consultations/category-3-4-weeks/244808-cloud-services-market-study/associated-documents/cloud-services-market-study-final-report.pdf?v=330228. [80]

Olcott, E. and Z. Wu (2025), *How long can Nvidia stay ahead of Chinese competition?*, https://www.ft.com/content/c24a9b6c-1664-4e46-affb-0c0dc16e3c4a (accessed on September 2025). [36]

Open Compute Project (2025), *Realizing the Open Data Center Ecosystem Vision*, https://www.opencompute.org/blog/realizing-the-open-data-center-ecosystem-vision. [105]

OpenAI (2025), “The next chapter of the Microsoft–OpenAI partnership”, https://openai.com/index/next-chapter-of-microsoft-openai-partnership/. [101]

Pak, A. (2024), *The CUDA Advantage: How NVIDIA Came to Dominate AI And The Role of GPU Memory in Large-Scale Model Training*, https://medium.com/@aidanpak/the-cuda-advantage-how-nvidia-came-to-dominate-ai-and-the-role-of-gpu-memory-in-large-scale-model-e0cdb98a14a0 (accessed on August 2025). [28]

Perez, L. (2017), *Cloud Computing Standards*, https://ec.europa.eu/newsroom/document.cfm?doc_id=42968. [165]

Perrone, H. (2025), “Chips in on a merger: The Arm-Nvidia case”, *International Journal of Industrial Organization*, Vol. 98, p. 103130, https://doi.org/10.1016/j.ijindorg.2024.103130. [171]

Pilz, K. (2023), “An Assessment of Data Center Infrastructure’s Role in AI Governance”, https://www.konstantinpilz.com/data-centers/assessment (accessed on 11 August 2025). [18]

Rahman, K. (2018), “Regulating Informational Infrastructure: Internet Platforms as the New Public Utilities”, *Georgetown Law Technology Review* 2, https://ssrn.com/abstract=3220737 (accessed on 5 August 2025). [14]

Reuters (2025), *CoreWeave, Nvidia sign $6.3 billion cloud computing capacity order*, https://www.reuters.com/business/coreweave-nvidia-sign-63-billion-cloud-computing-capacity-order-2025-09-15/ (accessed on September 2025). [58]

Reuters (2025), *In latest trade warning to US, China says Nvidia violated anti-monopoly law*, https://www.reuters.com/sustainability/boards-policy-regulation/latest-trade-warning-us- [177]

china-says-nvidia-violated-anti-monopoly-law-2025-09-15/ (accessed on September 2025).

Reuters (2025), *Nvidia faces $5.5 billion charge as US restricts chip sales to China*, https://www.reuters.com/technology/nvidia-expects-up-55-billion-charge-first-quarter-2025-04-15/. [176]

Reuters (2025), *South Korea unveils $23 billion support package for chips amid US tariff uncertainty*, https://www.reuters.com/technology/south-korea-unveils-23-billion-support-package-chips-amid-us-tariff-uncertainty-2025-04-14/ (accessed on August 2025). [120]

Reuters (2025), *Trump's FTC advances broad antitrust probe of Microsoft, Bloomberg News reports*, https://www.reuters.com/technology/trumps-ftc-moves-ahead-with-broad-antitrust-probe-microsoft-bloomberg-news-2025-03-12/ (accessed on August 2025). [77]

Richter, F. (2025), *The Big Three Stay Ahead in Ever-Growing Cloud Market*, https://www.statista.com/chart/18819/worldwide-market-share-of-leading-cloud-infrastructure-service-providers/?srsltid=AfmBOop3pUNyYsIw5ZIt7QK8Uh5BqzTdLbayTPgWrGQEjHLhAbXNP7wZ (accessed on September 2025). [54]

Rowles, D. (2000), "Is a Tie-In or an Integration? U.S. v. Microsoft Weighs In", *6 B.U. J. SCI. & TECH. L. 12, 15*, https://www.bu.edu/law/journals-archive/scitech/volume6/rowles.pdf. [140]

Samborska, V. (2025), *Scaling up: how increasing inputs has made artificial intelligence more capable*, https://ourworldindata.org/scaling-up-ai. [15]

Schaller, R. (1997), "Moore's law: past, present and future", *IEEE Spectrum*, Vol. 34/6, pp. 52-59, https://doi.org/10.1109/6.591665. [21]

Schumpeter, J. (1934), "The Theory of Economic Development: An Inquiry into Profits, Capital, Credit, Interest, and the Business Cycle. Translated by Redvers Opie.", *Cambridge, MA: Harvard University Press.*. [86]

Semiconductor Industry Association (n.d.), *Chip design and R&D*, https://www.semiconductors.org/policies/chip-design/ (accessed on September 2025). [23]

Sherman, N. (2024), "Microsoft chooses infamous nuclear site for AI power", *BBC News*, https://www.bbc.com/news/articles/cx25v2d7zexo (accessed on 11 August 2025). [68]

Shilov, A. (2024), *ASML's High-NA chipmaking tool will cost $380 million — the company already has orders for '10 to 20' machines and is ramping up production*, https://www.tomshardware.com/tech-industry/manufacturing/asmls-high-na-chipmaking-tool-will-cost-dollar380-million-the-company-already-has-orders-for-10-to-20-machines-and-is-ramping-up-production (accessed on October 2025). [88]

Siddik, M., A. Shehabi and L. Marston (2021), "The environmental footprint of data centers in the United States", *Environmental Research Letters*, Vol. 16/6, p. 064017, https://doi.org/10.1088/1748-9326/abfba1. [73]

Silva, V. (2025), *FRAND-Licensing Litigation Across the Atlantic: A Comparative Assessment of US and UK Jurisprudence on Telecom Disputes*, https://laweconcenter.org/resources/frand-licensing-litigation-across-the-atlantic-a-comparative-assessment-of-us-and-uk-jurisprudence-on-telecom-disputes/. [164]

Stokel-Walker, C. (2025), *The state of Intel: What the company's $5 billion deal with Nvidia could mean for the floundering chipmaker — Investments, capex, and the beating heart of x86*, https://www.tomshardware.com/tech-industry/the-state-of-intel-what-the-companys-usd5-billion-deal-with-nvidia-could-mean-for-the-floundering-chipmaker-investments-capex-and-the-beating-heart-of-x86 (accessed on September 2025). [33]

Stucke, M. and A. Ezrachi (2024), "Antitrust & AI Supply Chains", *University of Tennessee Legal Studies Research Paper*, No. Research Paper #474, https://doi.org/10.2139/ssrn.4754655. [55]

Swain, G. (2025), *Qualcomm launches global antitrust offensive against Arm, accusing it of stifling competition*, Computer World, https://www.computerworld.com/article/3855432/qualcomm-launches-global-antitrust-offensive-against-arm-accusing-it-of-stifling-competition.html (accessed on August 2025). [168]

The Economist (2025), *Japan storms back into the chip wars*, https://www.economist.com/asia/2025/08/21/japan-storms-back-into-the-chip-wars (accessed on August 2025). [123]

TheGuardian (2024), *Nvidia shares fall as China launches antitrust investigation into company*, https://www.theguardian.com/technology/2024/dec/09/china-nvidia-antitrust-investigation (accessed on August 2025). [170]

Thomas, E. (2025), *We broke down the eye-popping AI spending for 4 Big Tech firms — and their plans to go even harder next year*, https://www.businessinsider.com/big-tech-spending-on-ai-capex-q3-2025-10. [62]

Thornhill, J. (2025), *AI's double bubble trouble*, https://www.ft.com/content/da16e2b1-4fc2-4868-8a37-17030b8c5498. [109]

Tomoshige, H. (2022), *Japan's Semiconductor Industrial Policy from the 1970s to Today*, https://www.csis.org/blogs/perspectives-innovation/japans-semiconductor-industrial-policy-1970s-today (accessed on August 2025). [121]

Tom's Hardware (2025), *OpenAI signs contract to buy $300 billion worth of Oracle computing power over the next five years —*, https://www.tomshardware.com/tech-industry/openai-signs-contract-to-buy-usd300-billion-worth-of-oracle-computing-power-over-the-next-five-years-company-needs-4-5-gigawatts-of-power-enough-to-power-four-million-homes (accessed on September 2025). [57]

trendforce (2025), *HBM Evolution Outpaces JEDEC Standards as Generation Cycles Reportedly Shrink to 2.5 Years*, https://www.trendforce.com/news/2025/09/01/hbm-evolution-outpaces-jedec-standards-as-generation-cycles-shrink-to-2-5-years/ (accessed on October 2025). [38]

UK Department for Science, Innovation & Technology (2025), *AI Opportunities Action Plan*, https://www.gov.uk/government/publications/ai-opportunities-action-plan/ai-opportunities-action-plan#lay-the-foundations (accessed on August 2025). [117]

UncoverAlpha (2024), *Inside the AI Chip Market: NVIDIA, TSMC, Custom ASICs*, https://www.uncoveralpha.com/p/inside-the-ai-chip-market-nvidia (accessed on September 2025). [34]

United States District Court (2006), *United States of America v. Sun woo Lee*, [135] https://www.justice.gov/atr/case-document/file/501201/dl.

United States District Court (2005), *United States of America v Hynic Semiconductor Inc*, https://www.justice.gov/atr/case/us-v-hynix-semiconductor-inc. [134]

United States District Court (2005), *United States of America v Samsung Electronics Company and Samsung Semicoductor Inc*, https://www.justice.gov/atr/case-document/file/509296/dl. [132]

United States District Court (2005), *Unites States of America v. Infineon Technologies AG*, https://www.justice.gov/atr/case-document/file/499551/dl. [133]

United States Securities and Exchange Commission (2025), *Annual report pursuant to Section 13 or 15(d) of the Securities Exchange Act of 1934*, https://www.sec.gov/Archives/edgar/data/1045810/000104581025000023/nvda-20250126.htm. [181]

US DOJ (2025), *Assistant Attorney General Gail Slater Delivers Keynote at Fordham Competition Law Institute's 52nd Annual Conference on International Antitrust Law and Policy*, https://www.justice.gov/opa/speech/assistant-attorney-general-gail-slater-delivers-remarks-52nd-annual-conference (accessed on September 2025). [127]

Varas, A. et al. (2021), *Strenghening the global semiconductor supply chain in an uncertain era*, https://www.semiconductors.org/wp-content/uploads/2021/05/BCG-x-SIA-Strengthening-the-Global-Semiconductor-Value-Chain-April-2021_1.pdf. [39]

Vipra, J. and S. Myers West (2023), "Computational Power and AI", https://ainowinstitute.org/publications/compute-and-ai. [31]

von Thun, M. and D. Hanley (2024), "Stopping Big Tech from Becoming Big AI: A Roadmap for Using Competition Policy to Keep Artificial Intelligence Open for All", *SSRN Electronic Journal*, https://doi.org/10.2139/ssrn.4990780. [156]

Warnock, S. (2025), *Microsoft's own AI chip delayed six months in major setback — in-house chip now reportedly expected in 2026, but won't hold a candle to Nvidia Blackwell*, https://www.tomshardware.com/tech-industry/semiconductors/microsofts-own-ai-chip-delayed-six-months-in-major-setback-in-house-chip-now-reportedly-expected-in-2026-but-wont-hold-a-candle-to-nvidia-blackwell (accessed on September 2025). [92]

Weise, K. and C. Metz (2025), "At Amazon's Biggest Data Center, Everything Is Supersized for AI", *The New York Times*, https://www.nytimes.com/2025/06/24/technology/amazon-ai-data-centers.html?smid=url-share (accessed on 11 August 2025). [69]

Wu, D. (2024), "TSMC's Dominance Is Starting to Worry More Than Just Rivals", *Bloomberg*, https://www.bloomberg.com/news/newsletters/2024-10-21/tsmc-s-dominance-is-starting-to-worry-more-than-just-rivals (accessed on 11 August 2025). [41]

Yamaguchi, T. and L. Kihara (2024), *Japan unveils $65 bln plan to aid domestic chip industry*, https://www.reuters.com/world/japan/japan-propose-65-bln-plan-aid-domestic-chip-industry-draft-shows-2024-11-11/ (accessed on August 2025). [122]

Notes

[1] AI’s effects on competition in other markets is being discussed in a forthcoming OECD Global Forum session, see OECD (Forthcoming[182]).

[2] While stock prices and investment plans of the compute industry suggest huge growth in compute needs, reports that Chinese model developer DeepSeek was able to produce an effective AI model with far less compute highlights that technological developments are inherently uncertain (Davern and Pinnuck, 2025[172]).

[3] Previously GPU releases were typically every 2-4 years, but Nvidia has recently committed to an annual release cycle.

[4] Recently there have been reports Google is looking to also offer its TPUs to other AI cloud providers. Eg see https://www.datacenterdynamics.com/en/news/google-offers-its-tpus-to-ai-cloud-providers-report/

[5] Market share of hyperscalers as reported in competition authority market studies: range from 40-62% for AWS, 10-35% for Azure and 5-10% for Google cloud (Autorité de la concurrence, 2023[53]; Netherlands Authority for Consumers and Markets, 2022[70]; Ofcom, 2023[80]; Japan Fair Trade Commission, 2022[81]; Korea Fair Trade Commission, 2022[82]).

[6] Similar findings were made in the cloud computing sector market studies conducted by the French Autorité de la concurrence (2023[53]) and the Netherlands Authority for Consumers and Markets (2022[70]).

[7] The reported dispute focuses on Arm’s instruction set architecture, the fundamental code enabling software communication with processors. The dispute is driven by Arm’s shift from an open licensing model, under which chipmakers like Qualcomm could develop custom processors based on Arm’s designs, to a more restrictive approach favouring its own chip products. Qualcomm argues that this move threatens competition in the semiconductor industry, which has relied on Arm’s technology for over two decades (Swain, 2025[168]).

[8] The extent of vertical integration is much more limited upstream production processes beyond chip design, with firms not tending to get into wafer fabrication or chip packaging. In addition, even for chip design the entry is typically conducted with established chip design companies to help them design AI accelerators, primarily Broadcom and Marvell.

[9] For example, Nvidia's AI chips have been a key focus of US export controls, with US officials working to keep the most advanced chips from being sold to China as the US tries to keep ahead in the AI race. After those controls were implemented, Nvidia began designing chips that would come as close as possible to US limits (Reuters, 2025[176]).

[10] For example, the French Authority conducted dawn raids in an investigation in the GPU sector. Its report on Generative AI indicated that the sector is being closely scrutinised by their Investigation Services department (Autorité de la concurrence, 2024[111]). The Chinese competition authority has also opened an investigation into Nvidia regarding potential breaches of Chinese Anti-monopoly law (Reuters, 2025[177]).

[11] For example, the EU Commission has recently carried out unannounced antitrust inspections in the data centre construction sector. The Commission stated it had concerns firms may have violated EU antitrust rules that prohibit cartels and restrictive business practices (European Commission, 2024[173]).

[12] Depending on the jurisdiction, the exploitation of single firm's market power is commonly referred to as 'abuse of dominance' or 'monopolisation' (OECD, 1996[167]).

[13] For example, EU case law presumes that firms are dominant with market share over 50% (OECD, 2006[183])

[14] There have been investigations recently reported into Nvidia the leading provider of GPUs, for example:

- The US Department of Justice has reportedly launched an investigation into Nvidia after complaints from competitors that it may have abused its market dominance in selling chips that power artificial intelligence. US DOJ investigators are reportedly looking at whether Nvidia is giving preferential supply and pricing to customers who buy its complete systems (King and Nylen, 2024[169]; TheGuardian, 2024[170]; United States Securities and Exchange Commission, 2025[181]).
- A statement from the State Administration for Market Regulation (SAMR) announcing a probe into Nvidia for potential violating China's anti-monopoly laws. The SAMR did not elaborate on how Nvidia might have violated China's anti-monopoly laws. It also said that the US chipmaker is suspected of violating commitments it made during its acquisition of Mellanox Technologies under terms outlined in conditional approval of that deal. As part of the merger clearance of Nvidia, Mellanox Nvidia was required for six years from the date of the decision, to supply its GPUs and Mellanox's high-speed network interconnection devices, relevant software and accessories to mainland China on fair, reasonable and non-discriminatory terms (Kharpal, 2025[179]).

[15] In 2025, the OECD adopted a new recommendation on merger review (OECD, 2025[174]) which: calls for a clear legal framework for merger review to be effective, efficient, and timely; provide clear principles applicable to merger notifications and review procedures; ensure that merger assessment is effective and transparent; provide clear guidance for the design, assessment, and adoption of remedies.

[16] For example, Nvidia recently spent over USD 900 million to hire key staff at Enfabrica and license the AI startup's technology (CNBC, 2025[180]).

[17] The recent Illumina–Grail case, decided by the Court of Justice of the European Union (CJEU) in September 2024, was highly significant for EU merger control, particularly regarding jurisdiction over below-threshold mergers and the use of Article 22 of the EU Merger Regulation (EUMR). The CJEU ruled that EU Member States cannot refer mergers to the European Commission under Article 22 unless they themselves have jurisdiction to review the transaction under national law (European Union, 2025[175]).

[18] The EU's Digital Markets Act (DMA) is one of the most well-known examples, imposing pre-emptive obligations on designated "gatekeepers" to ensure contestability and fair conduct in core platform services. Similar regimes have also been brought in in the UK, Korea and Japan, with several other jurisdictions also considering introducing such frameworks.

OECD