The economic opportunities of open foundation models for Europe

—

Prepared for Meta

23 October 2023

Contents

Executive summary 3

1 Introduction 4
1.1 What are foundation models? 4
1.2 The distinction between open and closed foundation models 5
1.3 How this report contributes to the debate 7
1.4 Regulation of foundation models 7

2 How open foundation models unlock value 9
2.1 Overview 9
2.2 The benefits of open foundation models 9
2.3 Incentives 19
2.4 Conclusion 21

3 Mapping the benefits to the value chain of foundation models 22
3.1 The ecosystem surrounding open foundation models 22
3.2 How do users benefit? 24
3.3 Conclusion 31

A1 Technical Appendix—the importance of enabling competition 32
A1.1 The additional benefits provided by open models 32
A1.2 Estimates of benefits of generative AI 34

A2 Appendix—interviews 36

Figures and Tables

Figure 1.1 Overview of foundation model development and deployment 4
Figure 1.2 Spectrum of open and closed foundation models 6
Figure 3.1 The value chain of AI foundation models 22
Table A2.1 List of interviews conducted 36

Executive summary

Generative Artificial Intelligence (AI) models have recently been garnering the attention of the public and policymakers alike, and how best to regulate foundation models (and AI more generally) is subject of widespread debate.

Foundation models are 'any model that is trained on broad data (generally using self-supervision at scale) that can be adapted (e.g., fine-tuned) to a wide range of downstream tasks'.

Since AI and foundation models are expected by many to have a substantial positive impact on productivity over the coming years, it is vital for Europe to ensure that it makes the most of the opportunity that foundation models present. While the public debate remains ongoing and it is difficult to predict which risks may be more prevalent with open or closed foundation models, it is crucial for policymakers to bear in mind the considerable benefits uniquely attributable to open models.

This report prepared for Meta contributes to the debate regarding AI policy and regulation in two main ways. First, it provides an economic framework through which the benefits associated with open foundation models can be considered. Second, we have conducted interviews with industry participants in Europe, ranging from venture capital and AI innovators to banking firms and retail players. This direct access to European companies that are seeking to implement AI tools means this study provides additional evidence and a fresh perspective for policymakers to consider.

Open foundation models unlock the potential of community collaboration, allow greater customisation, and can enhance competition and decrease entry barriers especially for SMEs

Our research shows that AI has significant potential to accelerate the automation of activities and processes, potentially leading to growth gains in many sectors. Technology start-ups as well as key pillars of the economy—sectors like telecoms, financial services, the public sector, automotive and retail—are likely to benefit from rapid adoption of the many different use cases underpinned by foundation models.

In order to realise that potential, however, it is key that businesses of all sizes are able to unlock the benefits of AI, and harness them to create value to the benefit of businesses and society more broadly. Open foundation models have the potential to unlock significant additional productivity gains, which form the basis of improvements in incomes and welfare.

Longer-term incentives like independence, reduced costs, and data privacy in many cases favour open foundation models

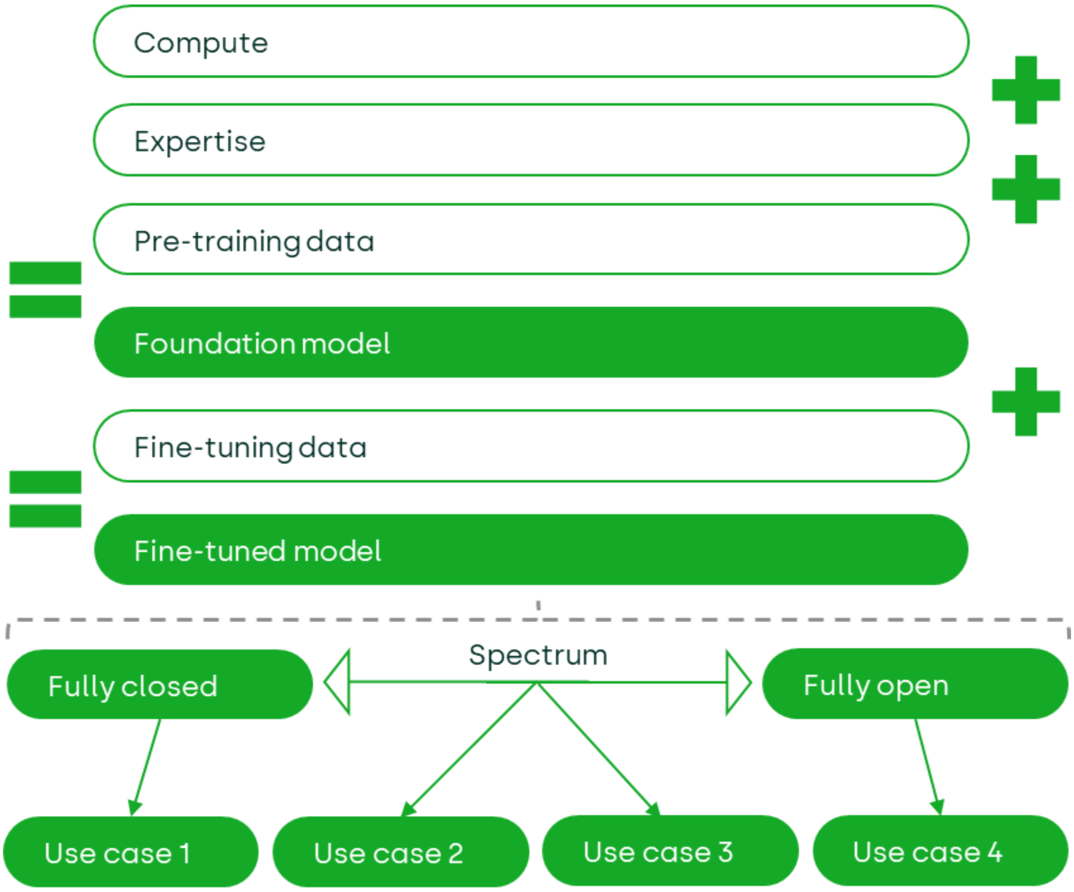
It is therefore essential for policymakers to clearly define what elements of foundation models are subject to regulation in order not to discourage—and where possible encourage and support—the uptake of AI.

1 Introduction

1.1 What are foundation models?

Foundation models can be defined as 'any model that is trained on broad data (generally using self-supervision at scale) that can be adapted (e.g., fine-tuned) to a wide range of downstream tasks'.[1] The development and deployment of foundation models is summarised in Figure 1.1.

Figure 1.1 Overview of foundation model development and deployment



Source: Oxera.

The first step is to pre-train the foundation model. This involves vast amounts of data being processed by the model, as well as computing

[1] Center for Research on Foundation Models (2022), 'On the Opportunities and Risks of Foundation Models', 12 July, p. 3, last accessed 2 October.

power and expertise. Pre-training is generally based on publicly available data, such as data available on the web or publicly available datasets, but pre-training can also be done on proprietary data.[2]

Pre-trained models can subsequently be fine-tuned to add specific capabilities—for example, specialising a pre-trained model to a particular domain or task that the data used to fine-tune the model relates to.[3]

Foundation models can power generative artificial intelligence (AI), which has recently been garnering the attention of the public and policymakers alike. Generative AI tends to be general purpose and, as described by the World Economic Forum, 'refers to a category of artificial intelligence (AI) algorithms that generate new outputs based on the data they have been trained on. Unlike traditional AI systems that are designed to recognize patterns and make predictions, generative AI creates new content in the form of images, text, audio, and more'.[4]

In this report, we will be using the Future of Life Institute's definition of general purpose AI: 'an AI system that – irrespective of how it is placed on the market or put into service, including as open-source software – can be used in, and adapted to, a wide range of distinct and downstream tasks, including some for which it was not intentionally and specifically designed'.[5]

1.2 The distinction between open and closed foundation models

The focus of our report is foundation models. There is a distinction between open and closed foundation models. The concept of open models is fluid and has been evolving since foundation models first became widely available—there is no one fixed definition.[6]

Open foundation models are those models that grant permissive IP rights to use the model to third parties and make the code and parameters publicly available, making them transparent and accessible.

The concept of 'open source' within AI foundation models differs from how open source is used more generally. It is broadly used in reference to a set of values—those that 'embrace and celebrate principles of open exchange, collaborative participation, rapid prototyping, transparency, meritocracy, and community-oriented development.'[7] Meanwhile, open

[2] CMA (2023), 'AI Foundation Models: Initial Report', 18 September, paras 3.5–3.7.
[3] CMA (2023), 'AI Foundation Models: Initial Report', 18 September, para. 2.12.
[4] World Economic Forum (2023), 'What is generative AI? An AI explains', 6 February, last accessed 2 October.
[5] Future of Life Institute (2023), 'FLI position paper: AI act Trilogue', July, last accessed 2 October.
[6] While some point to open source certification approved by Open Source Initiative (OSI), others argue that these licences were designed more to govern source code than the actual foundational models. The definition of 'open source' with respect to foundational models is under development by an OSI working group, but has not yet been released at the time of this report's publication.
[7] Opensource.com (2023), 'What is open source?', last accessed 13 September.

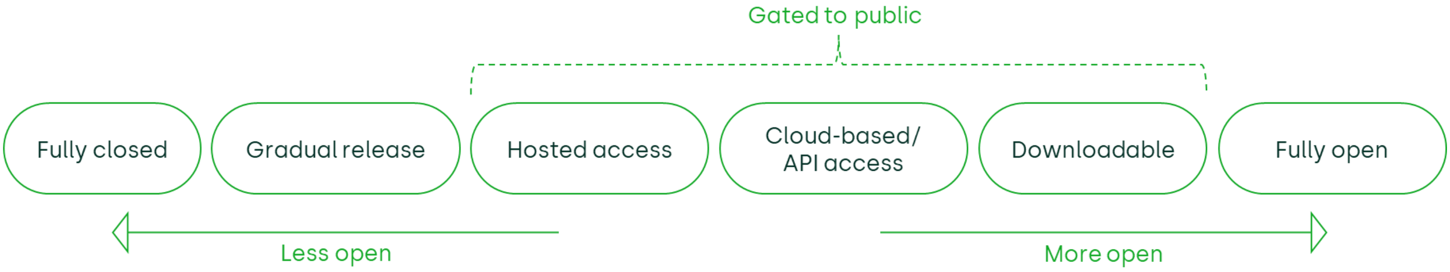
source within foundation models refers to the degree to which access to the model and the weights are provided to third parties.

For the purposes of this report, we define open foundation models ('open generative models', 'open generative AI models') as those models that grant permissive IP rights to use the model to third parties and make the code and parameters publicly available, making them transparent and accessible.[8] This is distinguished from closed source models, which do not grant these rights to external actors and may allow use of the model via an application programming interface (API) (or not at all).

Closed foundation models do not grant permissive IP rights to external actors and may allow use of the model via an API (or not at all).

At the same time, it is important to note that open and closed source models are not binary concepts. The degree of openness exists on a spectrum, as shown in the figure below, with some models offering more permissive IP rights than others and therefore being considered more 'open'. While some models may not be able to make everything available to third parties, they may still exist on the open end of this spectrum, and vice versa. For instance, despite the name, OpenAI's Chat-GPT 4 provides access to its model via an API and does not share its model and weights, features that are more usually attributed to closed models.

Figure 1.2 Spectrum of open and closed foundation models



Source: Oxera, based on Solaiman, I. (Hugging Face) (2023), ''The Gradient of Generative AI Release: Methods and Considerations', working paper, p. 4.

Fully closed systems do not allow use by those external to the developing organisation, but may in some cases provide gradual access to some groups, such as researchers. Those in the middle of the spectrum allow users outside the developer organisation to use them

[8] This definition aligns with the definition provided by the CMA. See CMA (2023), 'AI Foundation Models: Initial Report', 18 September, para. 3.49.

and provide feedback. However, users outside the organisation do not have access to the model itself or its training data. Those on the fully open end of the spectrum provide access to the parameters (in particular the weights and biases, explained to the right),[9] a description of the training data, the code itself, and do not have restrictions on who can access them (i.e. they are publicly available).[10]

Model weights and biases are the set of parameters that are iteratively calculated during the training of the model, based on the inputs and outputs provided in the training data.

1.3 How this report contributes to the debate

This report contributes to the debate regarding AI policy and regulation in two main ways. First, it provides an economic framework through which the benefits associated with open foundation models can be considered—to our knowledge there is no other analysis available that explores this topic. Second, we have conducted seven interviews with industry participants in Europe, ranging from venture capital and AI innovators to banking firms and retail players—this direct access to European companies that are seeking to implement AI tools means this study provides fresh evidence for policymakers to consider.

1.4 Regulation of foundation models

How best to regulate foundation models—and AI more generally—is the subject of widespread debate. The European Union (EU) is currently developing the AI Act, the first comprehensive AI law in the world, which will regulate the use of AI in the EU.[11]

The AI Act may also pave the way for AI regulation in other jurisdictions where work on developing AI regulation is also underway,[12] such as in the UK[13] and the USA. For example, in the USA, the White House recently published a document outlining three key principles which it regards as being fundamental to the future of AI—safety, security and trust—and some leading AI companies have made voluntary commitments to promote these principles.[14]

[9] Weights and biases, respectively, act as multipliers and additions to each of the calculations that convert inputs to outputs in foundation models. A foundation model is made up of a network of these fixed calculations. The number of parameters (i.e. weights and biases) of a model is also referred to as the size of the model. See CMA (2023), 'AI Foundation Models: Initial Report', 18 September, para. 2.6.
[10] Some may also consider models to be fully open only when the training data is also made public. Solaiman, I. (2023), 'Generative AI Systems Aren't Just Open or Closed Source', *Wired*, 24 May, last accessed 20 September.
[11] European Parliament (2023), 'EU AI Act: first regulation on artificial intelligence', 8 June, last accessed 20 September.
[12] Other countries are progressing AI regulation in the form of full acts, such as Brazil and Canada, or specific laws, such as China.
[13] Reuters (2023), 'Britain opts for 'adaptable' AI rules, with no single regulator', 29 March, last accessed 20 September.
[14] The White House (2023), 'Ensuring Safe, Secure, and Trustworthy AI', July, last accessed 3 October.

Bearing in mind the considerable benefits uniquely attributable to open foundation models that are identified throughout this report, we consider it essential for policymakers to clearly define what elements of foundation models are subject to regulation in order not to discourage—and where possible encourage and support—the uptake of AI.[15]

In addition, regulators should seek to ensure that key elements of open foundation models are not unfairly (or inadvertently) disadvantaged relative to closed foundation models. For example, collaborative development and public hosting components are crucial to harnessing the benefits of open foundation models, particularly through reducing barriers for smaller companies. This is especially the case since, at least in Europe, smaller companies exhibit greater positive impact in an open source context than what one would expect given their size, since they are more likely to create 'commits', i.e. contributions of code to the open source community.'[16]

Moreover, regulators will need to seek to differentiate between different use cases and types of foundation models (open vs closed) in order to set requirements appropriately for all approaches.[17] Under a risk-based approach, the focus is on ensuring that users of the models do not cause harm. One business we interviewed told us that with closed models such regulatory requirements can be met by the developer organisation, facilitating implementation, whereas with open models the burden is more likely to fall on smaller third-party firms using the model.

While this report offers general points for consideration in designing economically sound regulation for foundation models, the focus of this report is on assessing the economic benefits of open foundation models vis-à-vis closed models. Given that the public debate remains ongoing and that it is difficult to predict which risks may be more prevalent with open or closed foundation models, this report does not provide concrete principles of sound AI regulation more generally.

[15] See also Meyers, Z. and Springford, J. (2023), 'How Europe can make the most of AI', Centre for European Reform, September.
[16] See European Commission (2021), 'The impact of Open Source Software and Hardware on technological independence, competitiveness and innovation in the EU economy', final study report, May.
[17] See also Cepa (2023), 'An AI Challenge: Balancing Open and Closed Systems', 30 May.

2 How open foundation models unlock value

2.1 Overview

The main focus of this report is on assessing the benefits of open foundation models, particularly in relation to closed foundation models. However, a comprehensive assessment of the benefits of open foundation models requires considering those benefits in the context of how foundation models can be used more generally, and this report therefore also discusses some risks of open models and some benefits of closed foundation models.

Our research, informed by economic analysis as well as interviews with stakeholders active in the AI value chain, investigates the economic benefits of open foundation models.[18] The interviewees represent businesses across Europe active in a variety of sectors, such as technology, financial services, telecommunications, and retail.

The results of our research show that AI has significant potential to accelerate the automation of activities and processes, potentially leading to growth gains in many sectors. In order to realise that potential, however, it is key that businesses of all sizes are able to unlock the benefits of AI.

2.2 The benefits of open foundation models

This is where open foundation models have a significant role to play. Open models are uniquely positioned to unlock value to the eventual benefit of businesses and society more broadly.[19] While in the short run, especially for newer players lacking the necessary infrastructure, a closed model with a plug-in API can be more efficient, in the longer run there are many reasons why an open model provides more ownership, control, and even cost savings to business users.

This section discusses each of the following advantages of open foundation models.

[18] We have conducted interviews with employees of seven organisations, which have been helpful in enhancing our understanding of the potential for open foundation models. We acknowledge the fact that businesses may face a variety of incentives with respect to responses during interviews, and the conclusions expressed in this study are our own.

[19] For a broad overview of the policy considerations around the risks of closed and open generative AI models, see OECD (2023), 'Initial policy considerations for generative artificial intelligence', OECD Artificial Intelligence Papers, September.

Ecosystem of collaboration

With open models, businesses benefit from the contributions of a wide community of developers, researchers, and experts who can improve the model's performance, fix bugs, and identify security vulnerabilities, accelerating the development of applications and end uses (see section 2.2.1).

Reduced costs

Cost savings to businesses from open generative models relative to closed models come in a variety of forms, including direct cost savings relative to APIs, and faster implementation times (see section 2.2.2).

Democratisation of AI

Open foundation models can help democratise AI, and can allow smaller firms and start-ups to leverage machine learning in ways that would otherwise be prohibitively costly, thereby unlocking its potential for more businesses and consumers (see section 2.2.3).

Enhanced competition

Compared to closed foundation models, open foundation models are more likely to enhance competition, which has the potential to reduce costs and barriers to entry for both upstream providers of open foundation models and businesses active downstream (see section 2.2.4).

Interoperability and compatibility

Open models, unlike closed models, can more easily allow for interoperability and compatibility between many different inputs and components, such as data, algorithms, integration points, and programming languages, which allows for easier automated communication between platforms and organisations (see section 2.2.5).

Security and data privacy

Open foundation models also have the potential to enhance security and data privacy by, in contrast to closed models, allowing businesses the flexibility to in-house the models, and therefore keep custody of their data, as well as through open community stress testing of security measures (see section 2.2.6).

Transformative effects

Finally, by enabling community-based innovation, there is potential for transformative effects of open foundation models. This may be in the form of new use cases, leading to new services for end-users, or through new approaches to existing use cases (see section 2.2.7).

2.2.1 Ecosystem of collaboration

The potential development and testing by an open AI community of thousands of users has the potential to outweigh testing that can be undertaken by any individual company in isolation.

Community collaboration to improve open foundation models means that the foundation models themselves continue to become more productive for the developer organisation and third parties. Better-quality, more reliable and better-functioning generative models improve business productivity. As noted by the UK Competition and Markets Authority (CMA):

> The greater transparency of open-source models has several benefits. Users can have a better understanding of how the models work, which can help them to assess their accuracy and reliability. They can also modify the code of open-source models to improve them or add new features. Additionally, users can contribute to the development of open-source models by submitting bug fixes or new features. The development of open-source models can also be crowdsourced. Users can make suggestions to the original developers to improve the model.[20]

> 'We are only beginning to uncover the capabilities of Large Language Models. Experiments aimed at integrating LLMs with real-time data from third-party services and enabling actions on those services have yet to achieve significant adoption. However, open source models, due to their transparency and extensibility, present a massive opportunity to rapidly attract contributors who can build the necessary connectors to third-party services, potentially accelerating the practical utilisation and expansion of LLM capabilities.'
> **Sia Houchangnia, Partner, Seedcamp**

[20] CMA (2023), 'AI Foundation Models: Initial Report', 18 September, paras 3.50–3.51.

Open foundation models can further be used in other IT settings to aid in developing code (based on previous contributions of others) which can improve efficiency for software development. Given the lower costs of accessing these kinds of applications through open models compared to closed models, more businesses can reallocate resources to more productive/human-centric tasks. In other words, by allowing more businesses to automate tedious and repetitive tasks, open foundation models can allow for human capital to be better utilised for more productive/human-centric tasks such as innovating new uses or products.

2.2.2 Reduced costs

Cost savings to businesses from open foundation models relative to closed models come in a variety of forms. First, direct cost savings can be realised in comparison to commercial software available through API-based closed models, with the majority of respondents of a survey by the Linux Foundation estimating that costs would be two to four times greater than the cost of using open source software.[21] While this survey concerns open source software, as opposed to open AI foundation models specifically, most of the reasons cited by survey respondents relate to themes of accessibility, easier deployment, affordability and flexibility afforded by open source, which are common to foundation models and software.

> 'at present, those who offer cloud solutions have a clear advantage, as newer companies do not have the infrastructure or pipeline in place. So newer companies have less control, but at the moment are able to gain from the efficiencies of utilising the infrastructure of bigger players. But the retraining process will become cheaper and more efficient as users will focus less on retraining the full model, but rather fine-tuning it with specific data for very specific purposes.'
> **Soroosh Mashal, AudEERING**

Open source models often come with detailed research papers enabling knowledge sharing and allowing start-ups and small companies to ramp up operations quickly, thereby reducing training costs. In other words, it is far less resource-intensive to fine-tune existing open foundation models to specific tasks than develop models in-house to perform those same tasks, because some of the most labour-intensive elements of development can be automated. Moreover, the more fine-tuned a model to a specific purpose the 'smaller' it is considered. According to one firm interviewed, these smaller models are easier to manage and to harness community contributions. They are also better suited to some use cases where accuracy is paramount, for example for use in the interpretation of medical imaging.

By eliminating the need for stand-alone teams to develop many applications, businesses can allocate resources more efficiently. This can either result in direct cost reductions, or the ability to re-allocate

[21] Linux Foundation (2023), 'Measuring the Economic Value of Open Source', March, p. 15.

resources to building out bigger applications or customising and improving the models to directly align with their objectives. Under open models, because data does not need to leave the custody of the business user, businesses benefit from being able to use local servers and infrastructure in terms of improving speed for applications that require real-time responses, which in turn reduces costs and improves performance.

Cost savings can also be realised due to the faster implementation time of open generative models that are refined to specific use cases. This can often be more cost-effective for companies than using closed model APIs, especially if that company has some of the necessary infrastructure in place and is using the model fine-tuned to a very specific purpose. These potential cost savings in comparison to closed model APIs can either be absorbed by the business directly, or allow them to reallocate resources to more productive tasks.

As noted by Sia Houchangnia, Partner at Seedcamp:

'The open foundation model is not necessarily perfect in all applications, but it provides a very good baseline in a very wide range of applications. A single team could build a general model that has decent quality throughout, and companies can leverage that to reduce their costs. In particular this allows smaller companies to use machine learning, while bigger companies can build bigger applications. So foundation models in that way democratise the use of machine learning.'

Start-ups that wish to deploy AI can avoid the high costs of training and initial investment, as the foundation model developer organisation absorbs certain costs like the cost of the compute power associated with training the model. This allows more start-ups to enter at that stage of the value chain who would otherwise be unable to do so due to the prohibitively high costs of developing such models were closed ecosystems to predominate (since in a closed ecosystem they would have to develop their own foundation model to avoid those costs). These businesses are then in turn able to contribute to improving the efficiency of the open model, while passing on benefits to end-users through improved services.

2.2.3 Democratisation of AI

By providing a good-quality general model, open foundation models allow for more equal access to using machine learning for various applications, particularly by smaller companies. This generally means that more companies/services and more consumers will be able to benefit, including in terms of affordability and inclusivity, creating a 'democratisation' of machine learning to a certain extent.

Open source generative models allow smaller firms and start-ups to leverage machine learning in ways that would otherwise be prohibitively costly. One business interviewed noted that currently it is complicated and costly to run foundation models from an engineering perspective. Most companies do not have the resources to run these models, let alone improve them or re-train the model to specific uses. Allowing these companies to do so at zero cost through open models (which often provide a license that grants users the rights to use, reproduce, distribute, copy, create derivative works of, and make modifications to the models on a royalty free basis), is therefore a clear benefit of open models. In this way, open source models can be mutually beneficial to third parties as well as the developer organisation, by allowing the latter to leverage third parties to provide additions that improve their models.

Further, with the constant and rapid pace of innovation in AI algorithms, costs to run open source models are likely to decrease over time, making them more cost-effective in the longer run compared to closed models. Fundamentally, even with the current high resource costs of running generative models, open foundation models are likely to unlock access for start-ups that otherwise might not have been able to access these models at scale.

> [...] there's often a long-term effort dedicated to building on top of open-source models. This approach aims to eliminate dependency on the providers of closed-source models and to escape the potential walled garden of a closed-source ecosystem [...] As tech sovereignty becomes increasingly vital, the ability to leverage the power of LLMs without dependency on external providers is critical.'
> **Sia Houchangnia, Partner, Seedcamp**

2.2.4 Enhanced competition

Another key benefit of open foundation models relative to closed models is the increased likelihood of enhanced competition. In September 2023, the CMA published a report on AI foundation models, which highlights some key recommendations to minimise the risk of competition concerns with respect to foundation models and ensure positive outcomes, including:[22]

- ensuring that access to key inputs (e.g. data, computing power, capital and expertise) is on fair commercial terms without undue restrictions;
- ensuring that firms cannot use their market advantage to restrict access to competitor firms in those markets, or other developers;

[22] CMA (2023), 'AI Foundation Models: Initial Report', 18 September, para. 3.111.

- reducing barriers to new entrants to minimise any disproportionate incumbent/early-mover advantage;
- ensuring that a range of models—including open and closed foundation models—are available for firms.

While a range of models along the open to closed spectrum should be available, as per the fourth CMA recommendation above, in many ways, enabling open foundation models will allow the first three recommendations to be met.

On the first two points, open foundation models can decrease the cost of entry for the providers of such models by effectively providing zero-cost access to the algorithm, and since the wider community collaborates on improving the model and less needs to be done in-house. Relative to closed models, the likelihood of control by a handful of large firms is also lessened through this increased upstream competition, as models can be maintained and adjusted more easily by upstream providers, lessening any market power, and thereby reducing incentives to make foundation models available on uncompetitive terms.[23]

With respect to the third point, open foundation models can enhance competition in downstream markets relative to closed foundation models in several ways. For example, open foundation models can allow businesses to reduce potential risks around vendor lock-in where the provider of the foundation model is vertically integrated and may face an incentive to steer users towards ancillary services, such as a cloud computing service. Firms, even start-ups, have less dependency on any one particular vendor when open models are used, allowing the downstream company to have greater control. It can also allow for greater independence from certain technologies or geographies, enhancing European self-sufficiency.

Moreover, the avoided costs of paying for APIs by using open foundation models has the added benefit of reducing barriers to entry, particularly for smaller and mid-sized companies.

With access to the model and its weights, business can control the model's implementation and modifications, as well as fine-tune models to suit their specific purposes. They are also able to maintain independence as they are less reliant on upstream providers, reducing

[23] CMA (2023), 'AI Foundation Models: Initial Report', 18 September, para. 3.61.

the risk of vendor lock-in.[24] This has the potential to enhance competition, and it can drive innovation as firms experiment with different applications and business models in order to gain a competitive advantage.[25]

Enhanced competition due to open foundation models ultimately increases productivity and thereby leads to welfare gains for society more broadly, as evidenced in the established body of economic literature which links enhanced competition to welfare gains (which we explain and reference in more detail in Appendix A1). This is for two main reasons.

Enhanced competition due to open foundation models ultimately increases productivity and thereby leads to welfare gains for society more broadly

First, competition can act as a disciplining device by placing pressure on managers of firms to become more efficient. This is because under strong competition some less-efficient firms are unable to survive in the market, whereas under weak competition managers of those firms do not face the same risk of the firm being competed out of business.

Second, when competition works well more productive firms can gain market share at the expense of less productive ones, with less productive firms eventually exiting the market. This results in higher average productivity across firms.[26]

While this study does not quantify the impact of enhanced competition due to open foundation models, the welfare gains from AI generative models are generally expected to be large (as we explain in Appendix A1)—which means that any incremental benefit from enhanced competition resulting from open foundation models is likely to be substantial as well.

2.2.5 Interoperability and compatibility

Open models can more easily allow for interoperability and compatibility between many different inputs and components, such as data, algorithms, integration points, and programming languages, which allows for easier automated communication between platforms and organisations.

By removing the barriers to moving from one foundation model to another through better interoperability between AI platforms, this also allows for the easier substitution of one model for another. As noted by

[24] CMA (2023), 'AI Foundation Models: Initial Report', 18 September, para. 3.50.
[25] CMA (2023), 'AI Foundation Models: Initial Report', 18 September, para. 3.60.
[26] Arnold, J.M., Brys, B., Heady, C., Johansson, Å., Schwellnus, C. and Vartia, L. (2011), 'Tax policy for economic recovery and growth', *The Economic Journal*, **121**:550, pp. 59–80.

one interviewee, 'In the short term, most start-ups begin their exploration of LLMs by connecting to the APIs of closed models. However, many of them design their product architecture to be model-agnostic, ensuring a smooth transition between different models as needed.' This reduces inefficiencies that downstream users may face, allowing them more control to choose an alternative foundation model to better suit their requirements or if the performance of the initial open model is no longer sufficient. This ultimately allows for better- and higher-performing applications.

The interoperability offered by open foundation models can also facilitate productivity and innovation gains as some repetitive/tedious tasks may simply be easier and more effective under open generative models than closed models.[27] For example, while writing advertising copy can be done under both open and closed foundation models, advertisements generated under open foundation models may be easier to integrate across different platforms when open models offer better interoperability protocols.

2.2.6 Security and data privacy

Open models can allow businesses to better ensure that the model aligns with their security and data privacy standards by permitting more flexibility in fine-tuning the underlying model for their specific requirements. Companies with sensitive data, such as those in highly regulated sectors, can opt to bring open models in-house without the need for data to leave their custody. Businesses can also easily use their own local servers or infrastructure, enhancing data security.

A commonly cited risk is that open source models have the potential for use by bad actors, introducing biased or harmful material that is more difficult to remove by the developing organisation in an open source environment.[28] However, it is worth noting that open models can allow for more regular and in-depth monitoring to improve reliability, security issues, etc. One interviewee explained that the community—i.e. thousands of users—can pressure test for security vulnerabilities more thoroughly and consistently (for example, through prompt engineering) than any individual company would be able to, allowing for more testing

[27] Verdict (2021), 'Interoperability is essential for AI development', 22 January, last accessed 21 September, TS2 (2023), 'The role of open AI ecosystems in standards and interoperability', 8 May, last accessed 21 September.
[28] CMA (2023), 'AI Foundation Models: Initial Report', 18 September, para. 3.52.

of approaches to risk limitation.[29] Open models are also more accessible and transparent for regulators. At the same time, the extent to which open foundation models can prevent against use by malicious actors is difficult to predict and remains subject to debate. It is worth noting that there is always a risk of bad actors, as has been witnessed in the history of other technology, such as software. We understand from some interviewees that not only does the transparent and auditable nature of open AI models prevent against this, clear and strong regulation will also help mitigate this risk. As suggested by Soroosh Mashal, Principal AI Strategist at AudEERING, a company for audio AI applications:

'The biggest risk of open models is not being able to easily distinguish between what is AI generated or AI assisted and what is not. This may lead to AI retraining itself on AI generated data. Regulating safeguards must be in place to mitigate against this risk, for instance, through clear watermarks that distinguish AI generated material.'

In addition, the protection of internal data and privacy in open foundation models is often a cause for concern. The use or re-use of a pre-existing open foundation model to then make it available for the use of others within the community means that (insights derived based on) confidential data could be exposed. Firms may have a disincentive to participate if appropriate mitigations are not in place. However, relative to closed models, open models actually allow companies—particularly those in highly regulated industries, such as financial services—to develop and keep their work in-house, reducing the risk of data leakage to other third parties. In other words, open source models give users control since those who do not wish to share their information can bring the models in-house such that they can ensure that confidential data never leaves their custody.

2.2.7 Transformative effects

Finally, the potential transformative effects of open generative models should not be understated. As more businesses make use of open foundation models, and expand to new use cases, there is an increasing possibility of a whole new range of applications, or even a revolutionary approach to any particular use case.

The potential to use foundation models to aid scientific research is becoming increasingly apparent. This is particularly true for tasks that

[29] This type of community innovation may currently tend to work better for smaller models since they are more manageable from an engineering perspective. The positive impact of community innovation is likely to gradually expand to larger models as the tools and technology to manage them improves.

are crucial to scientific research, but that machines may be better at than humans: from mapping the number of proteins in a cell to simulating protein structures to RNA sequencing.[30] Put another way:[31]

'[i]f discovering new drugs is like searching for a needle in a haystack […] AI acts like a metal detector'.

Open foundation models can be particularly useful in scientific research because of the ability for many researchers to collaborate and share findings across generative models, and improve on each other's contributions. By facilitating the ability to collaborate in this way, open generative models may hold the potential to unlock progress in a way not previously envisaged, leading to transformative effects in the longer run.

While research and development are well known to have major positive impacts on productivity and innovation, it is worth noting that it takes a long time for basic research and new patents to translate into productivity benefits. For this reason, they are not included in our quantitative modelling in the following sections.

2.3 Incentives

The benefits of open foundation models vis-à-vis closed models are substantial and numerous. However, the monetary incentive to develop closed models is more direct than for open models. Thus for the benefits of open models to be accessed it is important that these incentive issues are solved by the participants in the open ecosystem.

The two key risks regarding incentives in open ecosystems are as follows.

- First, there is no clear party that is incentivised to improve the performance of open foundation models—because all members of the community benefit and the model is being given away for free. Conversely, as one interviewee noted, for closed models, the provider of the closed foundation model is clearly incentivised to make improvements to the model, since such improvements can be directly monetised by increasing sales

> 'When considering AI, it is helpful to draw a parallel with the history of the battery's evolution. The battery was invented in 1800, yet it took several decades before it was adapted into alternating and direct currents for electricity generation. Consequently, the late 19th century marked a surge in electricity applications. Similarly, keeping AI closed off and inaccessible will limit its potential. Open-source AI models grant access to this technology, enabling the development of innovative applications.'
> **Soroosh Mashal, AudEERING**

[30] Freakonomics Radio (2023), 'A.I. Is Changing Everything. Does That Include You?', interview with Dario Amodei, 6 September.
[31] The Economist (2023), 'How scientists are using artificial intelligence', 13 September, last accessed 22 September.

and/or prices, potentially making closed models better positioned for performance and faster deployment.[32]

- Second, without sufficient infrastructure and services to allow for easy and cost-effective deployment of open foundation models, there is a risk that open foundation models are currently not used to their full potential because downstream users consider implementation to be overly complex or expensive. The developers of open foundation models may—given that effectively the model is being given away for free—face a smaller incentive to facilitate downstream users in deployment.

There are, however, important incentives at play in the open environment. As discussed in section 2.2.5, upstream providers are incentivised to improve the performance of open models because the business case for open models relies on their widespread use. Through higher usage and the networks effects this can generate, the provider benefits from the community fixing bugs, catching security weaknesses, and generally improving the models' performance.

Over time this could allow them to monetise the open foundation model through add-ons, or apply the findings to closed API-based models. Thus, in some ways it is a symbiotic relationship between users and creators of open foundation models. As interoperability allows users to switch to alternatives if there are issues with a given open foundation model, the upstream provider loses users if it does not maintain and improve the foundation model. If users can 'walk away' because an open foundation model's performance is insufficient, upstream providers will have a strong incentive to ensure it works to improve its performance. As noted by Soroosh Mashal of AudEERING:

'it really depends on how the model is managed. If an open model is going to be successful the provider needs to manage it and establish clear guidelines for its use in order for the community to contribute to it. The open models I have seen be unsuccessful are those where the developing organisation has provided it without these guidelines or management.'

[32] At present, some, such as the UK CMA, consider that closed source models are larger and better-performing as they are trained on larger datasets and with more powerful hardware (the CMA also acknowledges that this gap in capabilities may not continue over time), see CMA (2023), 'AI Foundation Models: Initial Report', 18 September, paras 3.53–3.54. However, model performance remains an open topic of debate, with many factors affecting performance other than whether the model is open or closed, see for example Prompt Engineering Institute (2023), 'How Does Llama-2 Compare to GPT-4/3.5 and Other AI Language Models', last accessed 19 October.

Upstream providers are also incentivised to improve open models in other ways, for example through other monetisation opportunities. Companies may offer open foundation models because of their ability to be adopted more quickly and more widely. Open source models can also have the added benefit of attracting top talent interested in contributing to the field of open science.

In addition, commercial opportunity may arise because their open models allow for developers to 'build applications based around the model (or others created by the company) on top of its computing platform, and by building tailored solutions for customers.'[33] In other words, it may create knowledge for the upstream provider to build more customised solutions through add-ons, additional support, etc. under a form of 'freemium' model. Other monetisation opportunities include the development of end-to-end AI development platforms and multimodal AI systems that incorporate different types of data.[34]

2.4 Conclusion

It is worth noting that both open and closed foundation models serve different purposes within the AI ecosystem and come with various benefits and drawbacks.

Overall, open foundation models can unlock benefits to firms by allowing them to maximise their competitive advantage through retraining the models and developing specific, fine-tuned applications, versus the 'off-the-shelf' options offered by closed models. While closed foundation models can serve specific business interests and help acclimatise users to embedding AI in their business practices, competition benefits of open foundation models may better serve the longer-term interests of businesses as well as consumers.

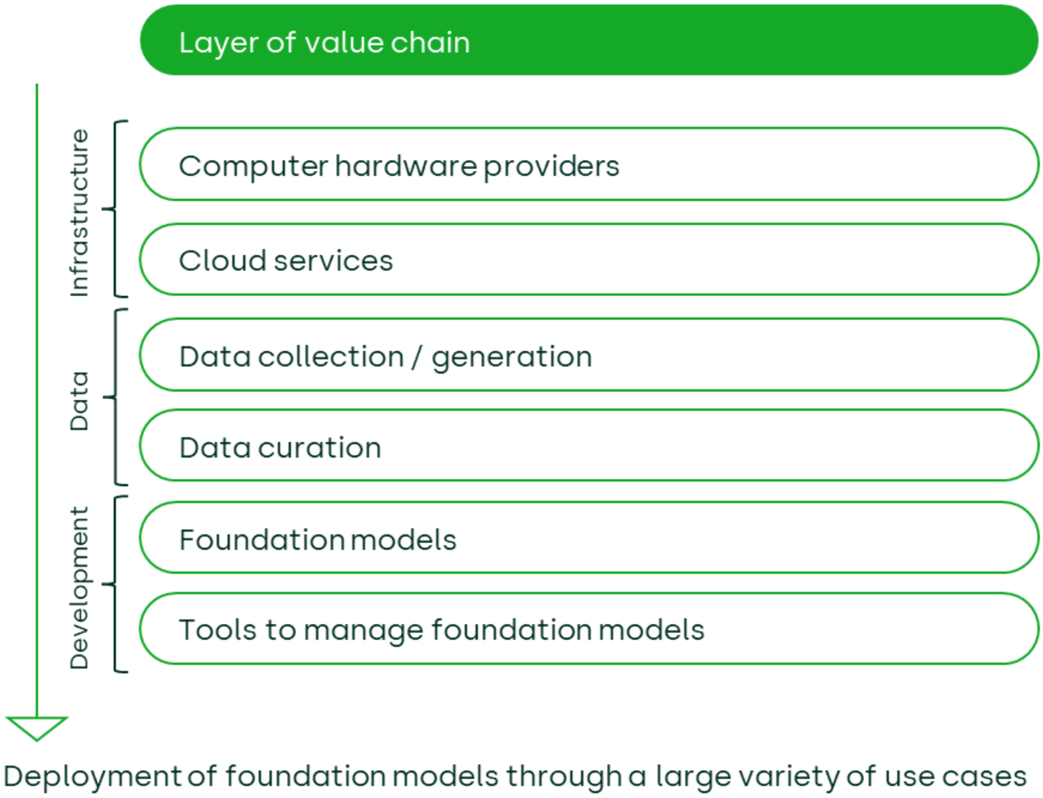
[33] The Economist (2023), 'Abu Dhabi throws a surprise challenger into the AI race, 21 September, last accessed 4 October.
[34] The Economist (2023), 'Abu Dhabi throws a surprise challenger into the AI race, 21 September, last accessed 4 October.

3 Mapping the benefits to the value chain of foundation models

3.1 The ecosystem surrounding open foundation models

Many steps are required before a foundation model can be deployed by a downstream user, and there are many distinct layers of providers that facilitate the various steps of that process. Together these layers form the value chain of AI foundation models, which can generally be summarised as follows in Figure 3.1, ranging from infrastructure to development. We discuss the final stage of foundation models—i.e. their various use cases and deployment to end-users—in section 3.2.

Figure 3.1 The value chain of AI foundation models



Source: Oxera.

Each of these layers of the value chain performs distinct tasks without which foundation models cannot be deployed, as outlined below. These layers jointly create the value that eventually accrues to end-users.

- **Computer hardware providers**—specialised chips that are used to train foundation models and infer the results from those models. These are produced through a complex value chain. Companies like Nvidia and Google (which is vertically integrated with its Google Cloud Platform) buy chips produced by foundries owned by businesses like TSMC (TSMC has recently entered into a joint venture with German and Dutch companies Bosch, Infineon and NXP to establish chip production facilities in Germany).[35] TSMC in turn uses equipment from multiple businesses in Europe, such as ASML in the Netherlands or Zeiss in Germany.[36]
- **Cloud services**—services that enable access to remote and scalable computing resources over a network, such as Amazon Web Services, Microsoft Azure, Google Cloud Platform and other (EU and non-EU) providers.
- **Data collection / generation**—data is a crucial input for training and fine-tuning foundation models. Typically the data used is collated from various sources, which can include the provider of the foundation model or the company using it, or scraped from the web. Data can also be acquired through external specialist data providers, like Surge AI or Prolific. Synthetically generated data can also be used and is predicted to become increasingly prevalent (as foundation models can also be used to generate synthetic data).[37]
- **Data curation**—before the data is used, it is generally cleaned and curated by the provider of the foundation model such that parts of the data that are harmful or do not benefit model performance are removed.
- **Foundation models**—a type of AI technology that is trained on vast amounts of data that can be adapted to a wide range of tasks and applications, such as GPT-4, LaMDA and Llama 2.[38]
- **Tools to manage foundation models**—tools like programming libraries or platforms for sharing, managing or deploying pre-trained models, such as PyTorch, TensorFlow or MLFlow.

[35] Reuters (2023), 'Bosch, Infineon, NXP, TSMC to establish JV for German wafer fab', 8 August, last accessed 3 October.
[36] CNBC (2021), 'Investors are going wild over a Dutch chip firm. And you've probably never heard of it', 24 November, last accessed 3 October.
[37] Gartner predicts 60% of data for AI will be synthetic to simulate reality, future scenarios and derisk AI, up from 1% in 2021. See Gartner (2023), 'Gartner Identifies Top Trends Shaping the Future of Data Science and Machine Learning', press release, 1 August, last accessed 22 September.
[38] Bommasani, R., Hudson, D.A. et al. (2021), 'On the opportunities and risks of foundation models', Working Paper.

At the deployment level, i.e. where service providers help businesses implement or enhance foundation models, foundation models also create opportunities for businesses. For example, the company Pinecone was founded in 2019 to provide an accessible packaged solution for the critical storage and retrieval infrastructure that is needed for running AI applications.[39] In addition, there are companies, such as Anthropic and Cohere, that specialise in the creation of foundation models in order to help the final users to apply them.[40] Such providers tend to rely on existing foundation models, and may or may not provide proprietary tools to manage them.

Another emerging area of business at the bottom of the value chain relates to AI governance, which can be defined as a legal framework aiming to ensure safe and responsible development of AI systems. As regulators are starting to make AI governance compulsory in order to guarantee a safer environment.[41] Daiki, an EU-based platform, helps organisations to build and use AI in a responsible manner by combining machine learning, legal and ethical expertise into the development of AI models.[42]

3.2 How do users benefit?

Downstream from the value chain are the end-users of the applications and services built on foundation models, i.e. where foundation models are eventually deployed. Due to the inherent flexibility of open foundation models, many types of end-users can benefit, including consumers as well as many different types of businesses.

The reasons why those end-users use foundation models also vary widely—ranging from enhancements to productivity software, support when writing code or when doing research, to healthcare applications like research on protein folding. Generally, foundation models are trained on a particular dataset that is geared towards the relevant use case, which can include, for example, data types such as text, images, or audio.

Throughout the rest of this section, we focus on the main use cases and benefits for five selected industries:

[39] See Pinecone (2023), 'Meet Pinecone', accessed 20 September 2023.
[40] See Google (2023), 'Building the most open and innovative AI ecosystem', 14 March, last accessed 3 October.
[41] European Commission (2023), 'Implementing AI Governance: from Framework to Practice', 17 July, last accessed 4 October.
[42] See Daiki website, last accessed 4 October.

- financial services (section 3.2.1);
- retail (section 3.2.2);
- telecoms (section 3.2.3);
- public sector (section 3.2.4);
- automotive (section 3.2.5).

Our research shows that, while some industries are better suited to certain use cases, each of these industries stands to benefit materially from easily accessible and open foundation models in the near term.

3.2.1 Financial services

Financial services firms, particularly banks and insurance companies, stand to gain from various use cases from the deployment of open foundation models. Many use cases allow these firms to benefit from cost savings through efficiencies. For example, one interviewee indicated that for banks, insurers, and other types of financial services firms, the use of virtual assistants simplifies the customer journey and reduces the communication points significantly when compared to human-led customer support. Another bank interviewed noted that 'in the medium-term, we see a big opportunity to explore new domains using generative AI models, such as providing very personalised advice and recommendations based on client data.'

> 'For mature corporates—such as big banks that hold and harness considerable amounts of customer data—open generative models can be more easily deployed and fine-tuned for specific purposes compared to closed models where upstream providers offer less flexible models via API.'
> **Europe-based financial services firm**

In addition, for products such as insurance or credit, a customer's risk profile can be much more easily assessed using data and analytics than through human decision-making. This makes the process more efficient, and lowers risk (both in the short and longer term) for the firm, while ensuring customers are directed to the correct product and priced at the right premium to meet their needs. Through this efficiency and better 'matching' of business interests and customer requirements, there are efficiency savings that can be realised by the business and passed on to customers.

According to one financial services firm interviewed, mature corporates, particularly big banks with large quantities of client data, also hold a substantial amount of analytics and insights on customer behaviour, which allows them to more easily monitor and prevent fraud. Since fraud liabilities in most jurisdictions often fall more on banks and financial institutions, utilising open generative models to predict whether a transaction aligns with a particular customer's behavioural data can allow firms to prevent fraud more easily and efficiently.

While it is possible to deploy these use cases under both open and closed models, open foundation models offer businesses more control and customisation. This is particularly beneficial for mature corporates—such as big banks that hold and harness considerable amounts of customer data. Open generative models can be more easily deployed and fine-tuned for specific purposes compared to closed models where upstream providers offer less flexible models via API. Some financial services firms benefit from hybrid open/closed models—for example, those that are sometimes referred to as freemium models, where the model and weights are provided under open source, but any additional professional services or support are paid for by the downstream user. One European bank interviewed indicated that for big banks that require specific applications, have considerable development power in-house, and hold vast amounts of customer data, the flexibility and customisation offered by open generative models or even freemium models can often be much better suited for their purposes than closed models.

Moreover, the use of open generative models can offer financial services firms, many of which have very high regulatory standards to meet with respect to data security and privacy, to in-house the models and retain custody of their customers' data. In closed models, customer data would in many cases fall under the custody of the upstream developer organisation. One financial institution interviewed indicated that:

'integration has been much easier with open AI models. It allows for deeper integration with our systems, because you have access to the model and weights and can fine-tune it more effectively. This allows us to control and safeguard the data using our security profiles. By being able to retain custody of the data, we are able to meet the very high standard of transparency and trust that banks are subject to.'

3.2.2 Retail

Similar to trends seen in the financial services sector described above, retail and e-commerce firms and their customers stand to gain from foundation models in areas like predictive analytics and personalised product recommendations.

Leveraging foundation models, one can analyse and forecast customer behaviour based on their usage patterns. For example, firms could identify those most inclined to like and purchase a certain product and subsequently recommend it to them, thereby streamlining the shopping process and increasing the effectiveness of marketing efforts. According to Yili Wu, CEO of SandStar, 'AI can also provide insights into the temporal patterns of consumer demand, including identifying seasonal item trends and estimating when these items will be in the highest demand.'[43]

The box to the right highlights the shopping assistant developed by Zalando based on ChatGPT, which falls more towards the 'closed' end of the spectrum of foundation models.[44] Retailers are deploying AI-based chatbots or assistants both on e-commerce platforms and in traditional brick-and-mortar shops. For example, 'chatbots can help customers navigate the store quickly and receive personalized product recommendations.'[45] This reduces the time a customer spends in a shop, resulting in time savings for the customer as well as leading to potentially higher sales conversion rates for the retailer. One e-commerce retailer interviewed observed that:

The use of AI is a fundamentally different way of interacting with computers. It's a different way of shopping than traditional online shopping, and in many ways it's closer to the experience of a physical shop. The difference of course is that instead of a person in the shop giving you advice on what to purchase, it's an AI system. This allows for better customisation because you can provide the AI your opinion and help direct it to something that suits you better.

Zalando, a German online lifestyle and fashion retailer, is launching a shopping assistant powered by a foundation model.

The assistant helps customers by allowing them to intuitively navigate through Zalando's assortment, and recommends products to customers based on their own fashion terminology and contextual cues.

In the future, this could be combined with customer preferences, such as preferred brands, to deliver more personalised selections of products.

[43] Wu, Y. (2023), 'Here's how artificial intelligence can benefit the retail sector', report for the World Economic Forum, 6 January, last accessed 4 October.
[44] See Zalando (2023), 'Zalando to launch a fashion assistant powered by ChatGPT', 19 April, last accessed 21 September.
[45] Wu, Y. (2023), 'Here's how artificial intelligence can benefit the retail sector', report for the World Economic Forum, 6 January, last accessed 4 October.

Retail firms are also increasingly using AI in their pricing strategies to 'spot patterns and relationships between multiple items' in order to estimate price elasticities for individual items and set prices optimally.[46]

For e-commerce retailers, prices change much more quickly than in traditional brick-and-mortar shops, making manual adjustment to price movements nearly impossible. For example, '[AI] can adjust prices based on multiple factors, even on a customer-by-customer basis', which means that when an online shopper adds an item to their cart but does not check out, 'AI can encourage them to complete the purchase by offering an attractive discount the very next day'.[47]

Across these use-cases, a business user of foundation models told us that being able to train, and in particular fine-tune, a foundation model on a company's own data is vital to differentiating a service relative to competitors—including with respect to personalised product recommendations or customer behaviour forecasts. While it is also possible to fine-tune some closed foundation models through APIs, using an open model can help to give businesses comfort that they retain full control of proprietary data, and it thereby facilitates differentiating and fine-tuning the model.

3.2.3 Telecoms

Telecoms firms can make use of foundation models in many ways, including by using them to provide improved customer service through chatbots based on foundation models. As a baseline, these chatbots could answer common questions and troubleshoot issues, and can include text-based as well as voice-based communication.

Especially when executed well, this has the potential to lead to substantial cost savings, since users would not feel the need for human assistance in finding what they are looking for. In addition, providing more intuitive online help functions can help businesses respond to a range of common customer queries, reducing the need for customer calls for one-on-one support.

In addition, a recent study found that the use of AI-based conversation assistants that offered suggestions for responses in real time to human customer service representatives in a call centre resulted in an increase

[46] The Economist (2022), 'How companies use AI to set prices', 26 March, last accessed 3 October.
[47] DLabs (2022), 'How AI Helps Retailers With Price Optimization', 21 November, last accessed 4 October.

in worker productivity of 14%.[48] This impact was particularly marked for newer and lower-skilled workers. The research also found that:

'AI assistance improves customer sentiment, reduces requests for managerial intervention, and improves employee retention.'[49]

Together, these effects improve productivity directly and indirectly, resulting in efficiencies and cost savings. While this example may be possible under both closed and open foundation models, open foundation models may allow for more customised integration of these conversation assistants, and also wider adoption of their use.

Finally, it is worth noting that there are many relevant use cases for foundation models in telecoms beyond customer service. For example, telecoms firms can integrate foundation models into network monitoring systems to more effectively detect or predict potential problems, propose troubleshooting procedures, and potentially automate specific tasks in order to minimise downtime and lower maintenance costs.[50]

3.2.4 Public sector

Public institutions stand to benefit from AI models in somewhat different ways than other sectors. Given that closed foundation models are offered under a more directly monetised business model, open foundation models may lend themselves better to more downstream 'social good' use cases.

Similar to other sectors, the use of open AI for public services can be better leveraged due to its relatively lower barriers to access. For public institutions on tight budgets, efficiency savings are as important as for firms in other sectors. These efficiency savings can be realised, for example, by utilising voice analytics for phone interactions for government services. The relatively lower implementation costs relative to closed models may allow for easier integration of open foundation models in government services.

Outside of government, public institutions such as museums and national archives can also utilise open foundation models in innovative ways. According to one public institution interviewed, because open

'open AI provides more of an incentive to create use cases for the social good. For instance, we can use open AI models to leverage our huge repository of audio, video and image data, including historical data, and make it available for the public'
Europe-based public sector organisation

[48] Brynjolfsson, E., Li, D. and Raymond, L.R. (2023), 'Generative AI at Work', NBER Working Paper 31161, April.
[49] Brynjolfsson, E., Li, D. and Raymond, L.R. (2023), 'Generative AI at Work', NBER Working Paper 31161, April, p. 2.
[50] See Master of Code (2023), 'Improving CX with Generative AI Chatbot in Telecom: Success Stories and Potential Use Cases', 14 July, last accessed 21 September.

foundation models allow for more customisation, they can be leveraged for use cases aimed at inclusivity, such as conversational search engines for people with different speech patterns (e.g. the elderly, children or non-native speakers of a language). This can allow for better accessibility to public services by a wide range of people.

3.2.5 Automotive

Automotive companies can be considered more manufacturing-intensive than the others analysed in this report and, for this reason, the ways in which foundation models help tends to be more visible, including designing vehicles or the components from which they are made. For instance, automotive companies are starting to use generative AI to improve in-vehicle voice assistants and self-autonomous driving.[51]

For example, in June 2023 Toyota Research Institute released a new technique to improve the efficiency of the design of vehicles through generative AI. Toyota designers can add engineering constraints in their sketches in order to reconcile both technical and design consideration into a single process, potentially having a positive impact in terms of the efficiency in designing and manufacturing vehicles.[52]

Similarly, Mercedes-Benz is incorporating ChatGPT into its MBUX voice assistant 'Hey Mercedes' in order to provide an improved user experience to its customers. In-vehicle voice assistants can sometimes be limited in the variety of actions they can perform, but with the implementation of a large language model, voice assistants can more easily sustain a conversation, thereby offering a better experience to drivers.[53]

In addition, Wayve, an Oxford based start-up founded in 2017, introduced in 2023 its own end-to-end (and therefore closed) generative AI model for autonomous driving named GAIA-1. The main function is to generate realistic driving videos through inputs such as video, text and actions. It has been trained on UK real-world driving data in order to understand and extract the concepts of real-world driving. This foundation model has been created with flexibility in order to be

[51] See Nvidia (2023), 'Generative AI Revs Up New Age in Auto Industry, From Design and Engineering to Production and Sales', 9 August, last accessed 4 October.
[52] Toyota (2023), 'Toyota Research Institute Unveils New Generative AI Technique for Vehicle Design', 20 June, last accessed 4 October.
[53] Mercedes-Benz Group (2023), 'Mercedes-Benz takes in-car voice control to a new level with ChatGPT', 16 June, last accessed 4 October.

implemented in future updates and is intended to improve autonomous driving performance and safety at the same time.[54]

The freedom and flexibility of open source allows vehicle manufacturers to enhance innovation by implementing their own technology into the baseline software. For example, Elektrobit, a German automotive software supplier, notes the transparency provided by open source that has been used for more than ten years to speed up the detection of issues and vulnerabilities.[55]

3.3 Conclusion

While the overall benefits of AI or foundational models (in general) are hard to predict at the current stage of adoption, the general consensus suggests these to be (very) high, with substantial positive impacts on productivity expected by the end of this decade (see section A1.2). This means it is vital for Europe to ensure that it makes the most of the opportunity that foundation models present.

As concluded by the European Parliament in a report, 'overall, it seems likely that, while AI has significant potential to boost productivity, the final effects will depend on the rate of AI diffusion across the economy and on investment in new technologies and relevant skills in the workforce.'[56]

[54] Wayve (2023), 'Introducing GAIA-1: A Cutting Edge Generative AI Model for Autonomy', 17 June, last accessed 4 October.
[55] Suse (2021), 'Elektrobit: Driving the Future of Automotive with Open Source and SUSE', last accessed 4 October.
[56] European Parliament (2019), 'Economic impacts of artificial intelligence', pp. 3–5.

A1 Technical Appendix—the importance of enabling competition

A1.1 The additional benefits provided by open models

A1.1.1 Why open models have the potential to enhance competition

As discussed in section 2.2.4, open foundation models can allow for a lower cost of market entry for potential providers of foundation models because the models are improved in part through community collaboration and hence can be maintained and adjusted at less cost. This means more organisations can develop and provide open foundation models, enhancing competition both upstream and downstream.

A1.1.2 How enhanced competition benefits productivity

Enhanced competition due to open foundation models has the potential to lead to more significant productivity gains. Productivity gains form the basis of improvements in incomes and welfare.[57] While productivity growth is affected by a range of factors, there is large body of evidence that shows that a well-functioning competitive process tends to enhance productivity, and that competition is an important determinant of the conditions under which productivity growth emerges.[58]

There are many academic studies that investigate the link between competition and productivity, and find results that indicate a substantial positive link.[59] The following studies are examples of such research.

- Holmes and Schmitz (2010) summarises the literature that examines industries experiencing dramatic changes in their competitive environment, and concludes that nearly all of the studies found that increases in competition had a positive impact on industry productivity.[60]
- Bourlès et al. (2013) investigates the impact of competition in intermediate goods markets on productivity downstream (i.e. where those intermediate goods are used) by analysing data on 20 industries in 15 OECD countries between 1985 and 2007. Their results suggest that increasing competition in upstream sectors

[57] Pilat, D. (1996), 'Competition, productivity and efficiency', *OECD Economic Studies*, **27**:2, pp. 107–46.
[58] Pilat, D. (1996), 'Competition, productivity and efficiency', *OECD Economic Studies*, **27**:2, pp. 107–46.
[59] CMA (2015), 'Productivity and competition: a summary of the evidence', 9 July.
[60] Holmes, T.J. and Schmitz Jr, J.A. (2010), 'Competition and productivity: a review of evidence', *Annual Review of Economics*, **2**:1, pp. 619–642.

could increase productivity growth by 1–1.5 percentage points per year.[61]

- Tang and Wang (2005) estimates the impact of product market competition on productivity using survey data. The authors find that medium-sized and large firms that state to perceive stronger product market competition tend to have higher productivity levels.[62]

The academic literature identifies two main ways in which stronger competition can lead to enhanced productivity.[63]

First, competition can act as a disciplining device by placing pressure on managers of firms to become more efficient. This is because under strong competition firms are unable to survive in the market, whereas under weak competition managers do not face the same risk of the firm being competed out of business. This leads managers to decrease 'x-inefficiency', which can be defined as the difference between the scenario in which the company is performing at its most efficient and its actual behaviour in the market. One example of an academic paper that studies this effect is Bloom and Van Reenen (2010), which finds that 'strong product market competition appears to boost average management practices through a combination of eliminating the tail of badly managed firms and pushing incumbents to improve their practices'.[64]

Second, when competition works well more productive firms can gain market share at the expense of less productive ones, with less productive firms eventually exiting the market. This results in higher average productivity across firms.[65] Baldwin and Gu (2006) studies such effects in the Canadian manufacturing industry over the period 1979–99, and finds that over 65% of productivity growth is directly related to

[61] The authors use anticompetitive upstream regulations as a proxy for the level of competitive pressure. See Bourlès et al. (2013), 'Do product market regulations in upstream sectors curb productivity growth? Panel data evidence for OECD countries', *Review of Economics and Statistics*, **95**:5, pp. 1750–1768.
[62] Tang, J. and Wang, W. (2005), 'Product market competition, skill shortages and productivity: Evidence from Canadian manufacturing firms', *Journal of Productivity Analysis*, **23**, pp. 317–339.
[63] CMA (2015), 'Productivity and competition: a summary of the evidence', 9 July. We do not discuss the impact of competition on innovation here since the relationship is complex. It is commonly referred to as taking the shape of an 'inverted-U', i.e. at low levels of competition an increase in competition tends to spur innovation, but beyond a certain point further increases in competition may have the opposite effect. See as a starting point Aghion, P., Bloom, N., Blundell, R., Griffith, R. and Howitt, P. (2005), 'Competition and innovation: An inverted-U relationship', *The Quarterly Journal of Economics*, **120**:2, pp. 701–728.
[64] Bloom, N. and Van Reenen, J. (2010), 'Why do management practices differ across firms and countries?', *Journal of Economic Perspectives*, **24**:1, pp. 203–224.
[65] Arnold, J.M., Brys, B., Heady, C., Johansson, Å., Schwellnus, C. and Vartia, L. (2011), 'Tax policy for economic recovery and growth', *The Economic Journal*, **121**:550, pp. 59–80.

higher-productivity firms gaining market share from less productive ones.[66]

A1.2 Estimates of benefits of generative AI

At this stage much remains unclear about how AI will develop, and as a result there is a large degree of uncertainty about the absolute level of benefits that are likely to arise.

This is also reflected in public estimates of the benefits of generative AI, which vary considerably between sources.

- A 2016 study by Accenture focused on 12 developed economies and estimated that AI had the potential to double annual global economic growth rates.[67]
- A 2016 study by Analysis Group estimated a range of between $1.49 trillion and $2.95 trillion in potential economic effects associated with AI by 2026.[68]
- A 2018 study by PwC estimated an up to 14% increase in global GDP by 2030 as a result of increased take-up and development of AI.[69]
- A 2023 study by McKinsey estimated that the yearly global benefits of generative AI are likely to range between $2.6 trillion and $4.4 trillion annually, which could roughly double if the impact of embedding generative AI into software that is currently used for other tasks is included.[70]
- A 2023 study by Goldman Sachs estimated that generative AI could account for a 7% increase in global GDP over ten years. This equates to about $7.1 trillion on a global basis, and $1.8 trillion (€1.7 trillion) in Europe and Central Asia alone.[71]
- A 2023 study by PublicFirst estimated that generative AI could increase gross value added (GVA)[72] in the EU by €1.2 trillion and save the average worker over 70 hours a year.[73]

[66] Baldwin, J.R. and Gu, W. (2006), 'Competition, Firm Turnover, and Productivity Growth, Economic Analysis (EA)', *Research Paper Series*, n. 042.
[67] Purdy, M. and Daugherty, P. (2016), 'Why Artificial Intelligence is the Future of Growth', Accenture p. 3.
[68] Chen, N., Christensen, L., Gallagher, K., Mate, R. and Rafert, G. (2016), 'Global Economic Impacts Associated with Artificial Intelligence', Analysis Group, p. 3.
[69] PwC (2018), 'The Macroeconomic impact of artificial intelligence', February, p. 3.
[70] McKinsey & Company (2023), 'The economic potential of generative AI: The next productivity frontier'.
[71] See World Bank Gross Domestic Product 2022. Goldman Sachs (2023), 'The potentially large effects of artificial intelligence on economic growth', 26 March.
[72] Gross Value Added is defined as the contribution made to the economy by an individual industry, sector, region, or producer. The sum of all GVA over all industries or sectors plus taxes on products minus subsidies on products sums to GDP. See Eurostat Statistics Explained.
[73] PublicFirst (2023), 'Google's impact in the EU', September, p. 7.

At the same time, there are sources that argue that AI may have a limited impact on growth. Aghion et al. (2017) argues that the ability of AI to improve productivity may be limited in certain sectors where retaining human labour is essential. This means that there is less opportunity for automation to improve growth, and thus the impact of AI on the overall economy may be limited.[74]

The overall benefits of AI or generative models are at present difficult to predict, but forecasts suggest these to be (very) high. Consequently, it is vital for Europe to ensure it makes the most of the opportunity that AI presents.

[74] Jones, C.I., Aghion, P. and Jones, B.F. (2017), 'Artificial Intelligence and Economic Growth', Stanford Institute for Economic Policy Research working paper No. 17-027, p. 46.

A2 Appendix—interviews

As part of this study Oxera conducted interviews with key stakeholders that are active in the AI space from a range of sectors. The general focus of the interviews was on the AI landscape, touching on topics like the expected impact of AI generally, the areas of society where the benefits of AI are likely to be felt most, and how these benefits might differ between open and closed approaches to foundation models. We also discussed how AI was specifically being applied in the interviewee's sector, including some of the use cases that are being developed or implemented in their organisation.

The discussions took place under the Chatham House rule, and many organisations felt more able to provide open commentary and insight without their identity being shared.[75] Where the organisation and interviewee are indicated we have received express permission to do so. Table A2.1 provides an overview of the interviews we conducted.

Table A2.1 List of interviews conducted

Organisation (if applicable)	Interviewee name and title (if applicable)	Sector	Description
Seedcamp	Sia Houchangnia, Partner	Financial Services—venture capital	European venture capital fund
Audeering	Soroosh Mashal, Principal AI Strategist	Tech	AI company based in Germany
Unattributed	N/A	Tech	AI coalition to promote AI adoption in a European country
Unattributed	N/A	Public Institution	European museum with vast multimedia archives
Unattributed	N/A	Financial Services—banking	Multinational bank and financial services company

[75] The Chatham House rule states that participants are free to use information from the discussion, but they are not allowed to reveal who made any particular comment. It is designed to increase openness of discussion. See Chatham House (2023), 'Chatham House Rule', last accessed 9 October.

Organisation (if applicable)	Interviewee name and title (if applicable)	Sector	Description
Unattributed	N/A	Financial Services—banking and insurance	Major European Bank and insurance group
Unattributed	N/A	Retail	E-commerce retailer with operations in Germany, Switzerland and Belgium (among others)

Contact

Robin Noble
Partner
+44 (0) 1865 253037
robin.noble@oxera.com

oxera.com