



HAL
open science

The Governance of AI Companies: Reconciling Purpose with Profits

Paul Oudin, Teodora Groza

► **To cite this version:**

Paul Oudin, Teodora Groza. The Governance of AI Companies: Reconciling Purpose with Profits. Corporate Governance Law eJournal, 2024, pp.47. <10.2139/ssrn.4972751>. <hal-04932663>

HAL Id: hal-04932663

<https://sciencespo.hal.science/hal-04932663v1>

Submitted on 6 Feb 2025

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons CC BY-NC-ND 4.0 - Attribution - Non-commercial use - No Derivative Works - International License

The Governance of AI Companies: Reconciling Purpose with Profits

Paul Oudin¹

Teodora Groza²

Draft of November 24, 2024

Abstract

Artificial intelligence ('AI') is both a critical driver of economic change and a source of potentially extreme negative externalities. For this reason, two leading AI companies, OpenAI and Anthropic, implemented customised governance structures with the double aim of addressing these externalities while remaining financially attractive to their investors. Other AI companies across the world adopted milder governance safeguards for that purpose. This paper studies these innovative governance frameworks by providing what is to the best of our knowledge the most comprehensive review of AI companies' various governance structures available to date. It then shows that applicable rules are determinant in shaping companies' ability to tailor their governance structure to their specific needs and examines the limitations of corporate laws in three European jurisdictions – France, Germany, and Italy – and, to a smaller extent, the US – more specifically, Delaware and Nevada – in enabling flexible governance structures that balance profit motives with public benefit objectives. Finally, it proposes recommendations for creating a new corporate form in the European Union to better support the peculiar needs of AI and other innovative companies, in line with the European Commission's priorities for the next five years.

*

*

*

¹ Departmental Lecturer in Law & Finance, University of Oxford; External Professor, Economics Department of Ecole polytechnique (France).

² PhD Student & Lecturer, Sciences Po Paris Law School; Research Fellow, TrackAI (Sciences Po Law School & CodeX, The Stanford Center for Legal Informatics).

Introduction

Artificial intelligence (AI) is regularly presented as a critical driving force of progress. According to Bill Gates, its potential impact on our society can be compared to commonplace technologies such as ‘the microprocessor, the personal computer, the Internet, and the mobile phone’.³ Sundar Pichai, Alphabet’s CEO, went as far as to say that its effects ‘could be more profound than both fire and electricity’.⁴ While this level of optimism is not universally shared, researchers working on the productivity gains of new technologies have highlighted that the impact of AI ‘is even more pervasive, more rapidly improving, and [...] has been spawning even more complementary innovations than earlier technologies’.⁵

Whereas many of today’s digital infrastructures—and the internet itself—started as government-led projects, the race for AI happens behind the doors of private companies.⁶ In simple words, ‘AI is being developed by private firms, run by executives, supervised by boards of directors, and funded by investors’.⁷ This is consequential for several reasons.

First, as the website of Anthropic—a key contemporary AI player—points out, ‘AI may create unprecedentedly large externalities, ranging from national security risks, to large-scale economic disruption, to fundamental threats to humanity, to enormous benefits to human safety and health’.⁸ Today’s best-known technologies belong to the broader category of generative AI, ‘a type of artificial intelligence technology that can produce various types of content, including text, imagery, audio and synthetic data’.⁹ But there is also fierce competition between top AI companies for the development of ‘artificial general intelligence’ (AGI), defined as ‘software with human-like intelligence and the ability to self-teach’.¹⁰

³ Bill Gates, ‘The Age of AI has begun’ *GatesNotes* (21 March 2023) <<https://www.gatesnotes.com/The-Age-of-AI-Has-Begun>> accessed 21 August 2024.

⁴ Jonathan Vanian, ‘Sundar Pichai Says A.I. Will Be More Profound Than Fire or Electricity’ (*Fortune*, 17 April 2023) <https://fortune.com/2023/04/17/sundar-pichai-a-i-more-profound-than-fire-electricity/> accessed 31 August 2024.

⁵ Tej Parikh, ‘Erik Brynjolfsson: ‘This could be the best decade in history – or the worst’ (*Financial Times*, 20 April 2023) <https://www.ft.com/content/b71759fe-397b-4688-bc81-b082edb25f31> accessed 31 August 2024.

⁶ Mariana Mazzucato, *The Entrepreneurial State: Debunking Public vs. Private Sector Myths* (PublicAffairs, 2015).

⁷ Roberto Tallarita, ‘AI Is Testing the Limits of Corporate Governance’ (*Harvard Business Review*, December 2023) <https://hbr.org/2023/12/ai-is-testing-the-limits-of-corporate-governance> accessed 31 August 2024

⁸ Anthropic, ‘The Long-Term Benefit Trust’ (2023) Anthropic <<https://www.anthropic.com/news/the-long-term-benefit-trust>> accessed 20 November 2024.

⁹ See eg George Lawton, ‘What is generative AI? Everything you need to know’ *TechTarget* <<https://www.techtarget.com/searchenterpriseai/definition/generative-AI>> accessed 21 August 2024.

¹⁰ See eg AWS, ‘What is AGI (Artificial General Intelligence)?’ *Amazon* <<https://aws.amazon.com/what-is/artificial-general-intelligence/>> accessed 21 August 2024.

Whereas the timeline of the race to AGI is uncertain,¹¹ it is clear that AI technologies evolve and get deployed at remarkable speed.¹²

In this context, there is an inherent backlog between the evolution of AI technologies and the capacity of regulators to assess and tame their externalities. Risk-based AI regulations – such as the EU’ Artificial Intelligence Act (‘AI Act’) – are already a reality, but reliance on ex ante risk assessment is not enough. Whereas the riskiness of some applications is immediately visible – such as, for example, border control management systems or medical tools – the assessment of risks posed by instruments such as chatbots or personal assistants is not obvious. The picture is even more complicated when it comes to multi-purpose tools or AGI. Consequently, to ensure that applications whose riskiness is difficult or impossible to assess ex ante, regulation needs to be paired with flexible and adaptive levers. Corporate governance mechanisms can complement regulatory efforts by ensuring that companies’ decisions to train and deploy AI systems factor in the public interest regardless of their current or predicted risk classifications.

Second, the development of AI technologies requires access to tremendous resources, from computing power and accelerator chips to sizable databases and, most importantly, human talent.¹³ These inputs are extremely costly, and thus require massive amounts of public or private capital. As of now, no government has created an AI task force which would supply private initiatives with all required funds.¹⁴ Private donations have also so far proven insufficient.¹⁵ AI companies thus need to rely on funding by investors motivated by financial return. Consequently, they have the difficult task of reconciling two potentially conflicting objectives: being financially attractive to private investors, on one hand, and placing AI safety and the interests of humanity above profits, on the other hand. Indeed, profit maximisation may involve taking existential risks that can generate sizable externalities for humanity at large, whereas preventing these risks from occurring may, from time to time, hinder the company’s ability to generate profits – even if, as Anthropic claims, ‘public benefit is not at

¹¹ Christopher Mims, ‘This AI Pioneer Thinks AI Is Dumber Than a Cat’ (*The Wall Street Journal*, October 2024) <https://www.wsj.com/tech/ai/yann-lecun-ai-meta-aa59e2f5> accessed 20 November 2024.

¹² Statista, ‘Adoption Rate for Major Milestone Internet-of-Things Services and Technology in 2022, in Days’ (December 2022) <https://www.statista.com/statistics/1360613/adoption-rate-of-major-iot-tech/> accessed 1 May 2024.

¹³ Competition and Markets Authority, ‘AI Foundation Models: Update paper’ (16 April 2024) <https://www.gov.uk/government/publications/ai-foundation-models-update-paper> accessed 1 May 2024.

¹⁴ Nonetheless, several countries have made AI a key policy orientation. For example, the UAE has invested hundreds of millions of dollars into AI infrastructure purchases, including a \$900 million project to build a cluster of supercomputers. See AINow Institute, ‘AI Nationalisms: Global Industrial Policy Approaches to AI’ (AINow Institute, March 2024) <https://ainowinstitute.org/wp-content/uploads/2024/03/AI-Nationalisms-Global-Industrial-Policy-Approaches-to-AI-March-2024.pdf> accessed 31 August 2024.

¹⁵ OpenAI in particular failed to finance itself only based on donations, despite its initial endeavour to do so. On this topic, see Section 1.1 below.

odds with commercial success or stockholder returns, and [both] are often strongly synergistic'.¹⁶

The evolution of OpenAI epitomises the difficulties faced by AI companies seeking to secure profits while pursuing – or at least accounting for – the interests of humanity. The entity's founders initially wanted to obtain governmental funding but saw 'no clear path in the public sector'.¹⁷ Then, they structured the organisation as a tax-exempt nonprofit in order to be 'unencumbered by profit incentives'.¹⁸ With time, 'it became increasingly clear that donations alone would not scale with the cost of computational power and talent required to push core research forward'.¹⁹

Today, OpenAI feeds its research efforts from two avenues. On the one hand, the entity now subsumes a for-profit subsidiary to its public-interest nonprofit. The for-profit branch incentivizes investment by giving financial returns, albeit capped at a maximum rate.²⁰ On the other hand, OpenAI is enmeshed in a series of 'partnerships' with Microsoft and Apple, out of which the former gets access to capital, computing power, cloud, and easy access to customers.²¹ Both governance choices are attempts to solve the problem of getting access to capital and other resources whilst putting the public interest first. To structure this complex organisation, which led to several conflicts and scandals,²² OpenAI chose to create several legal entities, most of which are registered in the state of Delaware. Incorporating in Delaware, whose corporate law is reputed as being one of the most flexible in the world,²³ gave OpenAI substantial leeway in creating a tailor-made legal architecture. Nevertheless, as we shall see, even Delaware law does impose some constraints limiting companies' flexibility in tailoring their governance structure, especially compared to other US states – such as Nevada – which implemented the so-called 'Model Benefit Corporation Act' (MBCL). Overall, US companies enjoy substantial flexibility in this regard. Building on these flexible rules, several top AI companies in the US adopted 'purpose-led governance structures', whereby the objective and ability to make profit is reconciled with the mission to ensure the safety of their products and foster humanity's interests.

¹⁶ Anthropic, 'The Long-Term Benefit Trust' (19 September 2023) <<https://www.anthropic.com/news/the-long-term-benefit-trust>> accessed 31 August 2024.

¹⁷ OpenAI, 'Our Structure' <https://openai.com/our-structure/> accessed 31 August 2024.

¹⁸ Id.

¹⁹ Id.

²⁰ OpenAI, 'Our Structure' <https://openai.com/our-structure/> accessed 31 August 2024.

²¹ Id.

²² Cade Metz, Kate Conger, and Kevin Roose, 'OpenAI Ousts Sam Altman as Chief Executive' (*The New York Times*, 17 November 2023) <https://www.nytimes.com/2023/11/17/technology/openai-sam-altman-ousted.html> accessed 31 August 2024.

²³ See eg Lawrence A Hamermesh, 'The Policy Foundations of Delaware Corporate Law' (2006) 106 *Columbia Law Review* 1749, 1782; Lewis S Black, 'Why Corporations Choose Delaware' (2007) *Delaware Department of State Division of Corporations* <https://www.google.com/url?sa=t&source=web&rct=j&opi=89978449&url=https://corpfiles.delaware.gov/pdfs/whycorporations_english.pdf&ved=2ahUKewitxdeh9o2IAxW1Q6QEHfhUDtAQFnoECBQQAQ&usq=AOvVaw0Wb0EgQDRsRnt_dyKN8juj> accessed 24 August 2024.

On the European continent, companies do not enjoy the same room for manoeuvre. We show this by focusing on French, German, and Italian law, both because of their weight in the European economy and because these jurisdictions' legal traditions have exported considerable legal transplants both in other Member States' legal systems and at the EU level. As we shall see, these jurisdictions do not provide companies with similar flexibility when it comes to crafting legal structures. While Member States' national corporate laws allow (AI) companies to pursue charitable objectives alongside shareholder value or profit,²⁴ some of these jurisdictions either limit or altogether prevent the possibility of creating profit-distributing entities (that is, entities that can distribute profits to their investors) which pursue primarily a nonprofit objective. Even when purpose can be tailored, companies enjoy very limited flexibility in adjusting directors' fiduciary duties accordingly.

In other words, French, German and Italian law make it difficult, and sometimes impossible, to effectively structure profit-distributing entities with a nonprofit goal. These rigidities, which affect three of the EU's biggest economies and are likely to characterise other Member States' corporate laws, should be a source of concern at the European level. This is especially problematic given the current context where the European Union 'aims at making the EU a world-class hub for AI and ensuring that AI is human-centric and trustworthy'.²⁵ The momentum created by the Draghi Report, which puts in the spotlight the need to restore the EU's competitiveness, should not be wasted. The Draghi Report makes it clear that with 'the spread of AI, a window has opened for Europe to redress its failings and innovation and productivity'.²⁶

As it stands, the EU's regulatory landscape is not up for the task. Indeed, as a recent letter signed by high-profile AI researchers and CEOs highlights, the existing '[f]ragmented regulation means the EU risks missing out on the AI era'.²⁷ To avoid this risk, the EU should 'choose to reassert the principle of harmonisation [...] so that AI innovation happens here at the same speed as elsewhere'.²⁸ According to this open letter, failing to harmonise applicable

²⁴ This paper regularly refers to profit maximisation as traditional companies' baseline objective. A strong case can be made that companies should primarily pursue shareholder welfare in general, rather than value and profits in particular: Oliver Hart and Luigi Zingales, 'Companies Should Maximize Shareholder Welfare Not Market Value' (2017) ECGI - Finance Working Paper No. 521/2017 <https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3004794> accessed 31 August 2024. Nevertheless, this paper focuses on profits (which, arguably, is in the tremendous majority of companies shareholders' central concern) to better contrast traditional companies with the AI companies discussed in this paper which (claim to) give purpose equal or superior importance to that of profit.

²⁵ European Commission, 'European Approach to Artificial Intelligence' (Digital Strategy, 21 April 2021) <<https://digital-strategy.ec.europa.eu/en/policies/european-approach-artificial-intelligence>> accessed 31 August 2024.

²⁶ Mario Draghi, *The Future of European Competitiveness* (European Commission, 2023) <https://commission.europa.eu/topics/strengthening-european-competitiveness/eu-competitiveness-looking-ahead_en#paragraph_47059> accessed 20 September 2024.

²⁷ EU Needs AI, 'The EU's Digital Future Needs AI' (EU Needs AI, 2024) <https://euneedsai.com/> accessed 20 September 2024.

²⁸ Id.

rules would mean ‘reject[ing] progress, betray[ing] the ambitions of the single market and watch[ing] as the rest of the world build on technologies that Europeans will not have access to’.²⁹ One branch of the law in dire need of such harmonisation is corporate law, in a context where Member States’ corporate laws remain very fragmented and too rigid to respond to AI companies’ governance needs. This begs the question of what the EU can do to solve this issue.

To answer this question, this paper is structured as follows. In the first section, we answer the question of what makes AI special by showing the particularities of AI technologies that warrant the adoption of tailored corporate governance rules. In the second section, we illustrate this necessity by reviewing the governance of existing AI corporations. We dedicate special attention to OpenAI and Anthropic’s organisational structures, but we also look into the governance of AI entities that have chosen to rely on more traditional governance structures. Along the way, we discuss applicable corporate rules to show their importance in ensuring the effectiveness of these structures. In the third section, we introduce the law and economics literature on organizational forms, highlighting that the primary function of legal entities is to serve as cost-reducing devices and arguing that organisational law is flexible enough to accommodate new standard forms. This section also exposes what features of corporate law are necessary to allow purpose-led companies to adopt a governance structure allowing for the effective implementation of purpose. The fourth section builds on the two previous sections by showing that in three selected European jurisdictions – France, Germany and Italy – corporate rules are too rigid to allow AI companies to effectively adapt their governance to their nonprofit mission. Conversely, Delaware and Nevada law, covered in our analysis of OpenAI and Anthropic in section 2, offer much more flexibility to companies incorporated in these states. The fifth and final section shows how our recommendations could be implemented in the European Union, most notably through the creation of a new corporate form.

Before moving on, we need to clearly define the limits of our claim. We do not believe that the creation of a tailored corporate form is a panacea, either for solving the potential issues of AI misalignment or for fixing the EU’s innovation backlog. Nonetheless, looking at the behaviour of market actors, we notice that shareholders of listed companies have actively shown concern for the safe and ethical use and development of AI by filing shareholder proposals at these companies’ general meetings for this purpose. Moreover, several top AI companies in the US – not the least of which OpenAI and Anthropic – decided to tailor their governance structures in such a way as to wholly or partly insulate their directors from the pressure of profit-oriented investors, thereby suggesting that relying on shareholders’ goodwill is not enough to ensure AI safety. Building on these observations, we infer two things. First, these developments show the need for a government-supplied solution for aligning AI governance with the public interest. Second, the limited presence of experimental AI governance and the general scarcity of AI companies in Europe may partly stem from their inability to tailor purpose-led governance structures due to the rigidity of corporate laws in Member States.

²⁹ Id.

1. What Makes AI and Targeted Regulatory Efforts Special

The emergence of new technologies has always been followed by calls for regulatory reforms. Examples that come to mind are the evolution of tort law in response to industrialization or the adoption of data protection laws in response to widespread digitalization. Given this pre-existing dynamic, the emergence and increasing adoption of AI systems may give the impression that AI is only one more example that fits this narrative. Nonetheless, there are some characteristics of AI that render its legal implications unique, and the calls for finding adequate legal governance mechanisms particularly pressing.

1.1. AI as ‘The Most General Purpose Technology’

As opposed to most technological breakthroughs, AI belongs to the category of general purpose technologies (‘GPTs’), such as electricity or the steam engine.³⁰ GPTs have three particularities: ‘they are pervasive, affecting most sectors of the economy; they improve rapidly; and they spawn complementary innovations’. Even within the category of GPTs, AI is singled out by the fact that it amplifies these attributes to an unprecedented degree.

Before explaining why, it is important to clarify that AI does not refer to a homogenous category of technologies. The most relied-on AI technologies to date are generative AI (‘GenAI’) applications. GenAI refers to ‘a type of artificial intelligence technology that can produce various types of content, including text, imagery, audio and synthetic data’.³¹ Illustrations of GenAI tools are the world-famous chatbots ChatGPT and Claude, which are large language models (‘LLMs’) that rely on natural language processing to imitate humanlike conversations. Nonetheless, the ultimate goal of AI developers is to reach AGI, meaning the ability to produce machines that ‘possess the ability to to understand or learn any intellectual task that a human being can’.³² Because AGI seeks to automatize intelligence itself, we can easily label it as *the most* general purpose technology. However, even existing GenAI technologies exhibit characteristics that delineate them from previous technological breakthroughs.

First, the use of GenAI quickly became pervasive. GenAI applications have been integrated into our daily routines at unprecedented levels. These tools have also enjoyed product penetration rates that trump those of any previous success stories: ChatGPT needed only five days to reach the milestone of one million users, whereas Instagram took 76 days.³³ According

³⁰ Boyan Jovanovic and Peter L Rousseau, ‘General Purpose Technologies’ in Philippe Aghion and Steven N Durlauf (eds), *Handbook of Economic Growth*, vol 1B (Elsevier 2005) 1181.

³¹ ‘Generative AI’ *TechTarget* <https://www.techtarget.com/searchenterpriseai/definition/generative-AI> accessed 20 November 2024.

³² ‘What Is Artificial General Intelligence?’ *Google Cloud* <https://cloud.google.com/discover/what-is-artificial-general-intelligence?hl=en> accessed 20 November 2024.

³³ Statista, ‘Adoption Rate for Major Milestone Internet-of-Things Services and Technology in 2022, in Days’ (December 2022) <https://www.statista.com/statistics/1360613/adoption-rate-of-major-iot-tech/> accessed 1 May 2024.

to a 2023 study, the usage of LLMs – which are just one specific AI application – can already improve 15% of all worker tasks in the US. When these LLMs are augmented by other software tools, this figure goes up to 47 to 56%.³⁴

Second, GenAI improves particularly rapidly. To illustrate this, MIT researcher Andrew McAfee documents that ‘OpenAI’s GPT 3.5 system, released in late 2022, performed better on the U.S. bar exam than about 10% of the human test-takers. GPT 4, released in March 2023, performed better than 90% of those taking the bar exam’.³⁵ Furthermore, the prices of token inputs have been drastically reduced,³⁶ and the ability of LLMs to process information received from users has increased 40 times.³⁷

Third, the ability of GenAI tools to trigger complementary innovations is immediately visible. Whereas GenAI was initially limited to LLMs only, today there are countless applications that can produce and process images, video, and audio. Furthermore, these applications have been integrated into several platforms, thereby provoking downstream innovation. A 2024 Brookings report notes that ‘LLMs can now be accessed through web-based chatbots, real-time voice assistants, web-based experimentation platforms, and Application Programming Interfaces (APIs), offering varying levels of customization and integration’.³⁸ Furthermore, the rise of open-source models ‘offers new opportunities for transparency, innovation, and cost-effective research applications’.³⁹

Lastly, Erik Brynjolfsson highlights that ‘while most GPTs take decades to fully play out [...], for AI [...] the productivity effects on the economy [are] just a matter of a few months’.⁴⁰ Both existing data and predictions support Brynjolfsson’s point. A 2023 NBER study highlights that access to GenAI assistants boosts the hourly productivity rate of customer support agents by 14% on average and by 34% for novice and low-skilled workers.⁴¹ A late 2024 experiment ran

³⁴ Tyna Eloundou, Sam Manning, Pamela Mishkin, and Daniel Rock, ‘GPTs are GPTs: An Early Look at the Labor Market Impact Potential of Large Language Models’ (arXiv, 21 August 2023) <https://arxiv.org/abs/2303.10130> accessed 20 November 2024.

³⁵ ‘The Impact of Generative AI as a General-Purpose Technology’ *MIT Sloan School of Management* <https://mitsloan.mit.edu/ideas-made-to-matter/impact-generative-ai-a-general-purpose-technology> accessed 24 November 2024.

³⁶ ‘New Models and Developer Products Announced at DevDay’ *OpenAI* <https://openai.com/index/new-models-and-developer-products-announced-at-devday/> accessed 20 November 2024.

³⁷ Tyna Eloundou, Sam Manning, Pamela Mishkin, and Daniel Rock, ‘GPTs are GPTs: An Early Look at the Labor Market Impact Potential of Large Language Models’ (arXiv, 21 August 2023) <https://arxiv.org/abs/2303.10130> accessed 20 November 2024.

³⁸ ‘Large Language Models Level Up: Better, Faster, Cheaper’ *Brookings* <https://www.brookings.edu/articles/large-language-models-level-up-better-faster-cheaper/> accessed 24 November 2024.

³⁹ *Id.*

⁴⁰ Tej Parikh, ‘Erik Brynjolfsson: ‘This could be the best decade in history – or the worst’ (*Financial Times*, 20 April 2023) <https://www.ft.com/content/b71759fe-397b-4688-bc81-b082edb25f31> accessed 31 August 2024.

⁴¹ Erik Brynjolfsson, Danielle Li, and Lindsey R Raymond, *Generative AI at Work* (NBER Working Paper No 31161, April 2023, revised November 2023) <https://doi.org/10.3386/w31161>.

by Boston Consulting Group found that consultants that augmented their skills with GenAI were 49% more efficient when writing code for data cleaning.⁴² As a last example, EY report finds that ‘GenAI-driven productivity is set to provide a substantial lift to the economy, likely delivering a boost worth \$650b over the next decade and lifting real GDP by nearly 2.5% by 2033’.⁴³ All these data points take us to the same conclusion: the economic impact of AI will be unprecedented.

1.2. AI Risks and How to Govern Them

The characteristics of AI systems that render them uniquely revolutionary are also the cause of major concerns. Take the example of the ultimate use case of AI, AGI. Since it aims to mirror human thinking, its applications are literally limitless, meaning that it may soon become by far the greatest contribution of AI to human progress – or extinction. Top AI companies’ executives and employees themselves have been the first to voice concerns about the threat frontier AI technologies may pose to humanity.⁴⁴ Some recently went so far as to assert that ‘future systems could be as deadly as pandemics and nuclear weapons’ and bring about humanity’s extinction.⁴⁵ Geoffrey Hinton, the so-called ‘godfather of AI’ who won the 2024 Nobel Prize in physics for his work on AI, expressed concerns that AI may soon become an existential threat to humanity.⁴⁶ For this reason, AGI’s development⁴⁷ and deployment should be carried out in such a way as to preserve ‘AI safety’ or ‘AI alignment’ – that is, the alignment of AGI with humanity’s interests.⁴⁸

Of course, various AI systems that fall short of the sophistication of AGI can and already do pose palpable threats to our societies. From the notorious legal mistakes and hallucinations of

⁴² ‘Gen AI Increases Productivity and Expands Capabilities’ *Boston Consulting Group* <https://www.bcg.com/publications/2024/gen-ai-increases-productivity-and-expands-capabilities#:~:text=Participants%20who%20used%20GenAI%20achieved,over%20participants%20not%20using%20GenAI> accessed 20 November 2024.

⁴³ ‘Productivity Potential of Gen AI’ *EY* https://www.ey.com/en_us/insights/ai/productivity-potential-gen-ai accessed 22 November 2024.

⁴⁴ Silicon Valley’s ‘effective altruism’, in particular, gives a central place to safety in the development of AI technology: Lawrence J Trautman, ‘Sam Altman, OpenAI, and the Importance of Corporate Governance’ (2024) SSRN working paper 18 <<https://ssrn.com/abstract=4679613>> accessed 20 August 2024.

⁴⁵ Kevin Roose, ‘A.I. Poses ‘Risk of Extinction,’ Industry Leaders Warn’ *The New York Times* (30 May 2023) <<https://www.nytimes.com/2023/05/30/technology/ai-threat-warning.html>> accessed 21 August 2024.

⁴⁶ Nilima Marshall, ‘Geoffrey Hinton: AI pioneer who fears artificial intelligence could end humanity’ *Independent* (8 Oct 2024) <<https://www.independent.co.uk/news/science/nobel-prize-in-physics-british-university-of-toronto-nobel-california-b2625885.html>> accessed 23 Nov 2024.

⁴⁷ One author suggests for instance to grant AI rights aligning AGI’s interests with those of humans: Peter Salib and Simon Goldstein, ‘AI Rights for Human Safety’ (2024) SSRN working paper <https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4913167> accessed 20 Nov 2024.

⁴⁸ ‘What is AI Alignment?’ (IBM Research, 8 November 2023) <https://research.ibm.com/blog/what-is-alignment-ai> accessed 20 November 2024.

LLMs⁴⁹ to the potential public health threats of patients self-diagnosing with AI tools,⁵⁰ there is an entire spectrum of potential harms. One such tragic example is the suicide of a US teenager pursuant to a months-long exchange with an AI-generated character on the Character.AI app.⁵¹ As this story warns, even AI applications that are not obviously dangerous can pose existential threats to our society: amplifying existing biases and discrimination, eroding privacy, causing misinformation, displacing jobs, exacerbating mental health issues – the list goes on.

AI threats are difficult to cabin to specific policy areas, and even when this can be done, regulation might not be the best solution. Regulatory tools that are AI-specific and cut across several policy areas which operate on the basis of assessed riskiness, such as the EU's AI Act, can quickly become obsolete.⁵² Whereas the risks of some applications, such as facial recognition, is immediately obvious, applications such as Character.AI can appear harmless at first but prove able to cause severe mental health issues in the long run. This means that dangerous systems can remain un- or under-regulated, in an environment in which regulatory change is distinctly costly, lengthy, and liable to threaten innovation.

The alternative route taken by the UK is to provide for a series of cross-sectoral principles that need to be implemented by existing regulators within their subject-matters. The merits of this approach is that it is adaptive and gives considerable leeway to sectoral authorities, rendering it much more capable to metabolize technological developments. Nonetheless, this approach has been criticized for 'tak[ing] flexibility to the extreme by advocating for a non-statutory framework where regulators implement the principles without statutory duty.'⁵³

Whereas most of the discussion on AI governance to date has focused on what kinds of rules are best suited, a missing piece of the puzzle is the analysis of alternatives to regulation. We take the meaning of regulation as referring to norms that require or prohibit certain actions and activities, whereas governance rules refer to norms that tinker with the incentives of various actors. In the case of corporate governance – our subject-matter – the role of the rules is to provide a series of structural levers and constraints that ensure that companies are

⁴⁹ Stanford Human-Centered Artificial Intelligence, 'Hallucinating Law: Legal Mistakes in Large Language Models Are Pervasive' (Stanford HAI, 17 November 2023) <https://hai.stanford.edu/news/hallucinating-law-legal-mistakes-large-language-models-are-pervasive> accessed 20 November 2024.

⁵⁰ Andrew A, 'Potential Applications and Implications of Large Language Models in Primary Care' (2024) 12(Suppl 1) *Family Medicine and Community Health* e002602 <https://doi.org/10.1136/fmch-2023-002602> accessed 20 November 2024.

⁵¹ Independent, 'Character AI Suicide Lawsuit: Sewell Setzer III Death' *The Independent* (London, 2024) <https://independent.co.uk/news/world/americas/crime/character-ai-suicide-lawsuit-sewell-setzer-iii-death-b2634706.html> accessed 20 November 2024.

⁵² Marco Almada and Anca Radu, 'The Brussels Side-Effect: How the AI Act Can Reduce the Global Reach of EU Policy' (2024) 25(4) *German Law Journal* 646.

⁵³ 'New Study: UK Should Avoid EU-Style AI Regulation, but Its Current Regulatory Framework Falls Short' *Brunel University London* <https://www.brunel.ac.uk/news-and-events/news/articles/New-study-UK-should-avoid-EU-style-AI-regulation-but-its-current-regulatory-framework-falls-short> accessed 20 November 2024.

governed in ways that align with the public interest. The difference with regulation is that, in providing for governance structures or blueprints, the state lies down the contours of companies' decision-making processes, rather than policing their outcomes through establishing clear prohibitions via ex ante norms.

In a 2009 paper on corporate social responsibility ('CSR'), Bénabou and Tirole highlighted that the state has a comparative disadvantage in policing certain nuisances, and that private actors are sometimes in a better position to self-regulate through corporate governance practices. The authors argue that companies tend to 'substitute themselves for elected government' in the following situations: 1) regulatory capture; 2) jurisdictional limits to the enforcement of certain obligations; and 3) 'a combination of inefficiency, high transaction costs, poor information, and high delivery cost'.⁵⁴

We argue that the last scenario holds true for companies training and deploying AI systems. Because the environment evolves at an unprecedented pace and the potential use cases are limitless, regulatory limits are not enough to shelter society from AI misuse and misalignment. External checks need to be complemented by internal levers that steer the governance of these companies towards the public interest. This does not mean that the relationship between regulation and corporate governance is an either-or, but rather that both modules need to be leveraged to ensure adequate AI oversight.

2. The Governance of Existing AI Corporations

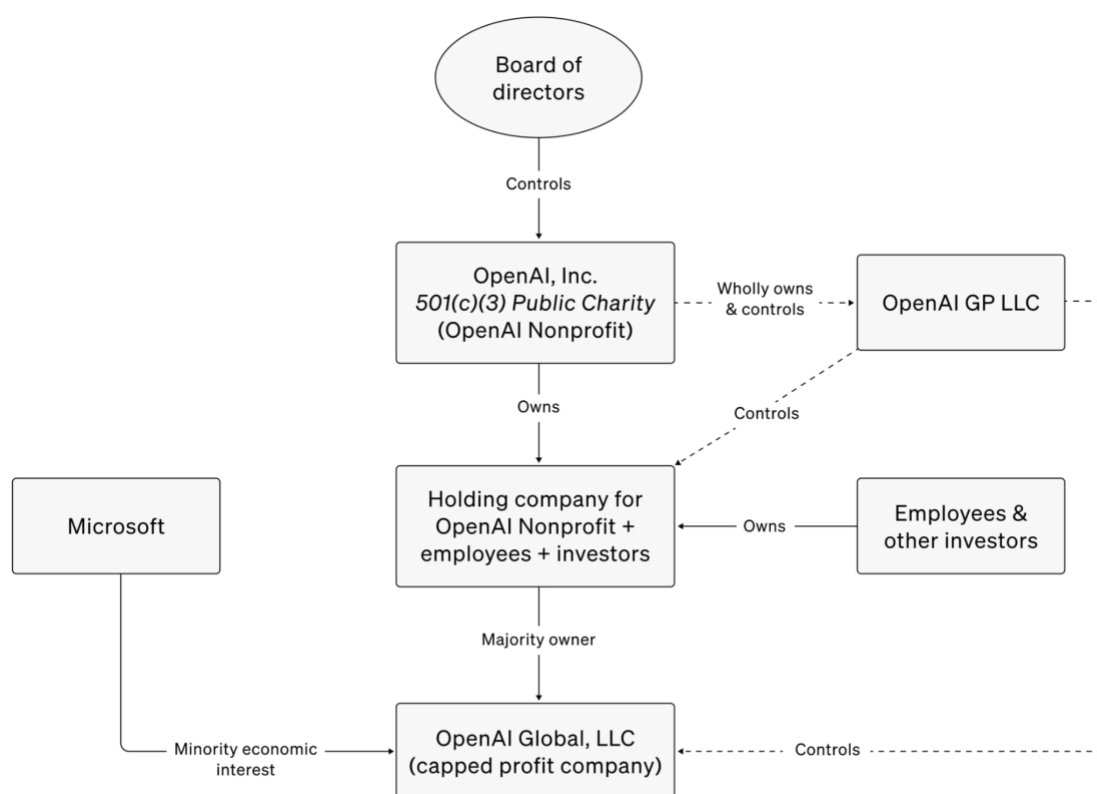
We start our journey by discussing the peculiar corporate governance of OpenAI, which relies on a unique combination of nonprofit and for-profit structures. This governance scheme gives its board, de jure if not de facto, the duty to place the interests of humanity before financial performance, while at the same time giving OpenAI's investors and employees a share in the company's profits. We then cover the governance mechanism implemented by Anthropic, another actor of reference of the AI industry. While Anthropic does not discard profit maximisation as a key objective of its activity, it concomitantly assigns itself the public mission of developing AI in the interest of humanity. Furthermore, according to its articles of association, Anthropic's powerful board of directors will eventually be composed of a majority of members designated by a trust which does not have any financial incentives. Thus, while OpenAI has the exclusive purpose of developing AI in the interest of humanity but can make profits, Anthropic has a 'bounded for-profit purpose', which ought to be fulfilled in the interest of humanity. Just as Anthropic, Inflection AI and xAI incorporated as benefit corporations, but without equivalent safeguards to those provided by Anthropic's trust. Finally, several other companies developing frontier AI technology recognize the importance of AI safety to various extents but did not implement corresponding corporate governance safeguards.

⁵⁴ Roland Bénabou and Jean Tirole, *Individual and Corporate Social Responsibility* (IZA Discussion Paper No 4570, Institute for the Study of Labor, 2009).

2.1. A Profit-Distributing Nonprofit: OpenAI

OpenAI, ChatGPT’s creator whose mission is ‘to build artificial general intelligence (AGI) that is safe and benefits all of humanity’,⁵⁵ is arguably the most famous and leading AI company worldwide. Far from being a single corporate entity, OpenAI is made of a web of closely intertwined legal structures, as explained below. Thus, any references to ‘OpenAI’ here and in subsequent developments should be understood as indistinctly referring to all or part of this group of structures, rather than to a specific legal entity. We specify the entity to which we refer when necessary.

OpenAI is structured as follows:⁵⁶



The official rationale for this group structure is that it seeks to reconcile two objectives: promoting the interests of humanity at large while securing the level of investment needed to develop AI systems and, ultimately, AGI. According to OpenAI, the company wants ‘to increase [its] ability to raise capital while still serving [its] mission, [but] no pre-existing legal structure [it knows] of strikes the right balance’.⁵⁷

⁵⁵ OpenAI, ‘Our structure’ <<https://openai.com/our-structure/>> accessed 31 August 2024.

⁵⁶ Ibid.

⁵⁷ OpenAI, ‘OpenAI LP’ (11 March 2019) <<https://openai.com/index/openai-lp/>> accessed 19 August 2024.

The entity's evolution shows that these two aims involve significant trade-offs. In the initial stages, OpenAI was a 'non-stock corporation' governed by the laws of Delaware⁵⁸ and qualifying as a public charity under US federal tax law.⁵⁹ This full nonprofit organisation was financed only by donations, not investments. The main drawback of this structure was that OpenAI could not distribute any benefits of any kind, and thus failed to attract sufficient funding. Indeed, according to OpenAI, 'it became increasingly clear that donations alone would not scale with the cost of computational power and talent required to push core research forward'.⁶⁰

An alternative possibility would have been to rely on a Delaware 'public benefit corporation' (PBC), as Anthropic and Inflection AI (discussed below) did. PBCs allow for the distribution of benefits while pursuing a public interest mission. However, the law defines these corporations as for-profit entities, and requires them to 'be managed in a manner that balances the stockholders' pecuniary interests' [and] the public benefit identified in its certificate of incorporation.⁶¹ By being incorporated as a Delaware PBC, OpenAI would thus have had to abandon the primacy of its nonprofit purpose to benefit humanity at large.

This is why OpenAI made the choice in 2019 of setting up a group of structures which, while managed as a nonprofit, gives investors and employees the possibility to get financial return⁶²—provided, however, that 'the company may never make a profit, and [...] is under no obligation to do so'.⁶³ While the nonprofit and for-profit dimensions of OpenAI's activity may sometimes lead OpenAI to forsake profits for the sake of humanity's interests, these two dimensions, far from being antagonistic, can in fact often be complementary. Indeed, OpenAI's business continuity—and, thus, its ability to develop safe AGI in the interest of humanity—requires funding by investors. In turn, OpenAI Nonprofit, while not required to do so, has an incentive to distribute part of its profits to investors to remain financially attractive in the future. In other words, OpenAI Nonprofit will not be able to pursue its grand mission without convincing investors that they will make profits by investing in its for-profit arm.

This corporate structure is designed as follows. The OpenAI Nonprofit is controlled by a board of directors whose fiduciary duties are towards 'humanity' and which is independent from OpenAI's investors. In turn, the OpenAI Nonprofit indirectly controls OpenAI Global, a

⁵⁸ OpenAI, Inc, Certificate of Incorporation executed on 8 Dec 2015 (Old Articles of Incorporation of OpenAI).

⁵⁹ Internal Revenue Code, Section 501(c)(3).

⁶⁰ OpenAI, 'Our structure' <<https://openai.com/our-structure/>> accessed 19 August 2024.

⁶¹ Delaware General Corporation Law, § 362.

⁶² On the advantages of setting up a for-profit subsidiary for a nonprofit entity, see David Avrum Levitt and Steven Richard Chiodini, 'Taking Care of Business: Use of a For-Profit Subsidiary by a Nonprofit Organization' *American Bar Association* (22 June 2014) <https://www.americanbar.org/groups/business_law/resources/business-law-today/2014-june/taking-care-of-business-use-of-a-for-profit-subsiary/> accessed 31 August 2024.

⁶³ *Ibid.*

‘capped profit company’ structured as an LLC.⁶⁴ Within the LLC, profits are capped in the sense that ‘returns for [the] first round of investors are capped at 100x their investment [and] this multiple [is expected] to be lower for future rounds’.⁶⁵ Thus, OpenAI Global is controlled by a nonprofit entity while being allowed to make profits.⁶⁶ To reflect this idea, we subsequently refer to OpenAI Global as a ‘profit-distributing nonprofit’, rather than a ‘for-profit entity’, as OpenAI does on its website. In OpenAI’s own words, the for-profit is ‘legally bound to pursue the Nonprofit’s mission, and carry out that mission’.⁶⁷ Furthermore, while OpenAI Global’s operating agreement is not public, this structure suggests that OpenAI Nonprofit’s directors are OpenAI Global’s managers.⁶⁸ For this reason, we subsequently refer to ‘OpenAI’s directors’ as both OpenAI Nonprofit’s directors and OpenAI Global’s managers, as they are the same persons.

This original structure warrants several observations. First, who actually controls OpenAI has been debated. For context, Microsoft owns a \$13 billion stake in OpenAI Global and reportedly played a key role in Sam Altman’s reintegration after he got fired by OpenAI Nonprofit’s board⁶⁹ – and before three board members were ‘forced out’ upon his return.⁷⁰ Elon Musk, one of OpenAI’s founders, went so far as to sue OpenAI and Sam Altman several times on the grounds that the company was now putting profits above its mission to benefit humanity.⁷¹

⁶⁴ For more explanations about OpenAI’s corporate structure, see Ellen P Aprill, Rose Chan Loui and Jill R Horwitz, ‘The Untold Nonprofit Story of OpenAI’ (*CLS Blue Sky Blog*, 5 March 2024) <<https://clsbluesky.law.columbia.edu/2024/03/05/the-untold-nonprofit-story-of-openai/>> accessed 19 August 2024; Ellen P Aprill, Rose Chan Loui and Jill R Horwitz, ‘Board Control of a Charity’s Subsidiaries: The Saga of OpenAI’ (2024) 182 *Tax Notes Federal* 289; Alexandra Andhov, ‘OpenAI’s Transformation: From a Non-profit to a 100 Billion Valuation’ (2024) SSRN working paper 8-14 <https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4750197> accessed 31 August 2024.

⁶⁵ OpenAI, ‘OpenAI LP’ (11 March 2019) <<https://openai.com/index/openai-lp/>> accessed 19 August 2024.

⁶⁶ On the governance implications of this structure, see Michael W Peregrine, Robert C Louthian III and Charles M Elson, ‘Corporate Governance Lessons from the OpenAI Controversy’ (*CLS Blue Sky Blog*, 26 Jan 2024) <<https://clsbluesky.law.columbia.edu/2024/01/26/corporate-governance-lessons-from-the-openai-controversy/>> accessed 19 August 2024.

⁶⁷ OpenAI, ‘Our structure’ <<https://openai.com/our-structure/>> accessed 31 August 2024.

⁶⁸ Technically, LLCs are managed by the members themselves (Delaware Limited Liability Company Act, § 18-402) or managers they appoint (Delaware Limited Liability Company Act, § 18-401) rather than directors.

⁶⁹ OpenAI, ‘OpenAI announces leadership transition’ (17 November 2023) <<https://openai.com/index/openai-announces-leadership-transition/>> accessed 25 August 2024.

⁷⁰ On this topic, see eg Daron Acemoglu and Simon Johnson, ‘Opinion: OpenAI’s drama marks a new and scary era in artificial intelligence’ *Los Angeles Times* (29 November 2023) <<https://www.latimes.com/opinion/story/2023-11-29/openai-sam-altman-firing-chatgpt-artificial-intelligence>> accessed 19 August 2024; James Broughel, ‘OpenAI Is Now Unambiguously Profit-Driven, And That’s A Good Thing’ *Forbes* (9 December 2023) <<https://www.forbes.com/sites/jamesbroughel/2023/12/09/openai-is-now-unambiguously-profit-driven-and-thats-a-good-thing/>> accessed 19 August 2024. On the departure of three board members, see Kari Paul, ‘OpenAI reinstates CEO Sam Altman to board after firing and rehiring’ *The Guardian* (9 March 2024) <<https://www.theguardian.com/technology/2024/mar/08/openai-sam-altman-reinstated>> accessed 31 August 2024.

⁷¹ The first lawsuit was filed on 29 February 2024 before being withdrawn a few months later: Jahnvi Nidumolu, Aditya Soni and Sheila Dang, ‘Elon Musk sues OpenAI for abandoning original mission for

This context led to sharp criticism of OpenAI's governance structure, with some going so far as to acknowledge its outright failure.⁷² According to two of its former board members, 'self-governance cannot reliably withstand the pressure of profit incentives'.⁷³ Others have labelled this structure as 'weird' or 'a failed experiment'.⁷⁴ Bracketing the debates on the merits of OpenAI's governance architecture, what is clear is that, given the several transformations it underwent, this bundle of entities has proven highly malleable, and countless future configurations are envisageable. In fact, OpenAI's original structure may get much closer to a traditional for-profit entity in the future. Indeed, OpenAI CEO Sam Altman recently announced that reliance on a for-profit benefit corporation may be considered, citing Anthropic and Elon Musk's xAI as examples of companies relying on this corporate structure.⁷⁵

This paper is not the place to discuss whether the board's decision to fire Sam Altman was well-founded. We simply observe that even if Altman's reintegration was pushed by Microsoft and motivated by OpenAI's ability to make profits, this decision did not necessarily conflict with OpenAI's mission. Indeed, as we saw above, OpenAI's ability to fulfil its mission in the interest of humanity in the future requires remaining attractive to investors, which in turn requires OpenAI being sufficiently profitable. Altman's failure to be 'consistently candid

profit' *Reuters* (1 March 2024) <<https://www.reuters.com/legal/elon-musk-sues-openai-ceo-sam-altman-breach-contract-2024-03-01/>> accessed 19 August 2024; Nick Robins-Early, 'Elon Musk abruptly withdraws lawsuit against Sam Altman and OpenAI' *The Guardian* (12 June 2024) <<https://www.theguardian.com/technology/article/2024/jun/11/elon-musk-withdraws-lawsuit-against-sam-altman-openai>> accessed 25 August 2024. Musk then sued OpenAI and Sam Altman again based on similar grounds on 5 August 2024: Nick Robins-Early, 'Elon Musk sues OpenAI again, alleging 'deceit of Shakespearean proportions'' *The Guardian* (5 August 2024) <<https://www.theguardian.com/technology/article/2024/aug/05/elon-musk-openai-lawsuit>> accessed 25 August 2024.

⁷² See eg 'Inside OpenAI's weird governance structure' *The Economist* (London, 21 November 2023) <<https://www.economist.com/business/2023/11/21/inside-openais-weird-governance-structure>> accessed 20 August 2024; Noam Wasserman, 'OpenAI's Failed Experiment in Governance' (*Harvard Business Review*, 30 Nov 2023) <<https://hbr.org/2023/11/openais-failed-experiment-in-governance>> accessed 20 August 2024; Kevin Roose, 'A.I. Belongs to the Capitalists Now' *The New York Times* (22 November 2023) <<https://www.nytimes.com/2023/11/22/technology/openai-board-capitalists.html?smid=li-share>> accessed 20 August 2024.

⁷³ Helen Toner and Tasha McCauley, 'AI firms mustn't govern themselves, say ex-members of OpenAI's board' *The Economist* (1 June 2024) <<https://www.economist.com/by-invitation/2024/05/26/ai-firms-mustnt-govern-themselves-say-ex-members-of-openais-board>> accessed 20 August 2024. These criticisms were strongly rejected by two current OpenAI board members: Bret Taylor and Larry Summers, 'OpenAI board members respond to a warning by former members' *The Economist* (30 May 2024) <<https://www.economist.com/by-invitation/2024/05/30/openai-board-members-respond-to-a-warning-by-former-members>> accessed 20 August 2024.

⁷⁴ Noam Wasserman, 'OpenAI's Failed Experiment in Governance' (*Harvard Business Review*, 7 November 2023) <https://hbr.org/2023/11/openais-failed-experiment-in-governance> accessed 20 September 2024.

⁷⁵ Aaron Holmes, Natasha Mascarenhas and Julia Hornstein, 'OpenAI CEO Says Company Could Become Benefit Corporation Akin to Rivals Anthropic, xAI' *The Information* (14 June 2024) <<https://www.theinformation.com/articles/openai-ceo-says-company-could-become-benefit-corporation-akin-to-rivals-anthropic-xai>> accessed 19 August 2024.

in his communications with the board⁷⁶ may not justify his ousting, should this discourage investors from investing in OpenAI. Declaring the failure of OpenAI's governance based on this single event might also be a stretch: Altman's reputation, network and unique experience as founder and CEO of OpenAI may make it difficult for any board of directors, however structured, to fire him as suddenly as OpenAI's former board did. Furthermore, the board's decision may very well have been misguided, at a time when the company's ability to grow and generate profits (which, presumably, was better ensured with its founder on board) arguably did not yet conflict with humanity's interests, thus potentially providing limited justification for the board's decision. Therefore, Altman's return does not necessarily mean that OpenAI's governance structure failed, but rather shows the limits of legal governance itself: for better or worse, there is only so much well-designed rules can do to tame a company's iconic leader, as suggested by OpenAI's former board members in the quote above. We can at best hope that, should OpenAI's board be faced with a clear-cut dilemma between maximising profits and safeguarding humanity's interests, it would be more inclined to protect the latter than the former than a traditional board.

Second, and relatedly, the fact that OpenAI Nonprofit's directors run both the nonprofit and the for-profit entity OpenAI Global required some legal engineering. On one hand, OpenAI's corporate structure clearly aims to prevent the profit-distributing entity from being run with financial performance in sight. This explains why OpenAI Global is managed by OpenAI Nonprofit's directors. On the other hand, OpenAI Global being a 'for-profit [...] legally bound to pursue the Nonprofit's mission',⁷⁷ its corporate officers' fiduciary duties should be tailored accordingly. As we shall see later in greater detail, OpenAI Global's structuration as an LLC gives it almost unlimited freedom to adapt or even exclude its managers' fiduciary duties. This means that whoever manages OpenAI Global may be completely immune from legal action by OpenAI's investors even if the latter's interests were blatantly disregarded by the managers. For now, we simply observe that this governance requirement certainly explains, at least in part, why OpenAI's profit arm has been structured as an LLC under Delaware law. We go back to the peculiarities of the LLC regime in Section 3.2.

2.2. The Bounded For-Profit Controlled by a Trust: Anthropic

Anthropic is a direct competitor of OpenAI. Besides operating AI assistant Claude, the company tries to develop AGI. Similarly to OpenAI, it claims to place safety at the forefront of its business considerations while aiming to become a leader in the AI industry: 'Anthropic is an AI safety and research company [which] build[s] reliable, interpretable, and steerable AI systems'.⁷⁸ In fact, the company was founded by former top employees of OpenAI

⁷⁶ OpenAI, 'OpenAI announces leadership transition' (17 Nov 2023) <<https://openai.com/index/openai-announces-leadership-transition/>> accessed 27 August 2024.

⁷⁷ OpenAI (n 44).

⁷⁸ Anthropic, 'Making AI systems you can rely on' <<https://www.anthropic.com/company>> accessed 21 August 2024.

(subsequently joined by other OpenAI executives⁷⁹) who reportedly claim to be more concerned for safety than OpenAI is.⁸⁰ This makes Anthropic an ideal candidate for a governance structure safeguarding the interest of humanity at large.

Despite their similar objectives, the governance structure of Anthropic significantly differs from that of OpenAI. First, it prima facie gives a larger place to profit-making. According to its bylaws, Anthropic pursues the mission of ‘responsibly develop[ing] and maintain[ing] advanced AI for the long term benefit of humanity’.⁸¹ However, it is structured as a public benefit corporation (PBC). As a result, it is a for-profit company by definition and directors necessarily owe fiduciary duties to the shareholders, as further discussed in Section 3.1.1. Furthermore, immediately after stating its public mission, Anthropic’s bylaws add that ‘in addition, the Corporation may engage in any lawful act or activity for which corporations may be organized under the Delaware General Corporation Law’.⁸² According to Anthropic, this structure ‘gives [the] board the legal latitude to weigh long- and short-term externalities of decisions—whether to deploy a particular AI system, for example—alongside the financial interests of [the] stockholders’.⁸³

One important feature that makes Anthropic different from traditional PBCs is the role played by one of its shareholders, the ‘Long-Term Benefit Trust’ (LTBT), which consists in ‘an independent body of five financially disinterested members’ and solely owns all of ‘Class T Common Stock’.⁸⁴ Class T Common Stock does not give rise to any dividend or liquidation right,⁸⁵ thereby depriving its holder of any financial incentive to maximise the company’s financial performance. However, the Class T common stockholder is entitled to elect one board member until 24 May 2026⁸⁶, then two board members until the company manages to raise at least \$5 billion of capital,⁸⁷ and, finally, three board members as soon as the company has managed to raise \$6 billion or, failing that, after 24 May 2027.⁸⁸ Anthropic’s board being

⁷⁹ Including Jan Leike and OpenAI co-founder John Schulman: Robert Hart, ‘OpenAI Exec Who Resigned Over Safety Concerns Joins Rival Firm—Here’s What We Know’ *Forbes* (29 May 2024) <<https://www.forbes.com/sites/roberthart/2024/05/29/openai-exec-who-resigned-over-safety-concerns-joins-rival-firm-heres-what-we-know/>> accessed 21 August 2024; Kimberley Kao, ‘OpenAI Co-Founder Leaves for Rival Anthropic’ *Wall Street Journal* (6 August 2024) <https://www.wsj.com/tech/ai/openai-co-founder-leaves-for-amazon-backed-rival-anthropic-4f34ebac?mod=itp_wsj> accessed 21 August 2024.

⁸⁰ Alex Kantrowitz, ‘Oh, Good, OpenAI’s Biggest Rival Has a Weird Structure Too’ *Slate* (2 December 2023) <<https://slate.com/technology/2023/12/anthropic-openai-board-trust-effective-altruism.html>> accessed 21 August 2024.

⁸¹ Anthropic, Amended and Restated Certificate of Incorporation executed on 20 May 2024 (Anthropic’s Articles of Incorporation), art III.

⁸² *Ibid.*

⁸³ Anthropic, ‘The Long-Term Benefit Trust’ (19 September 2023) <<https://www.anthropic.com/news/the-long-term-benefit-trust>> accessed 20 August 2024.

⁸⁴ *Ibid.*

⁸⁵ Anthropic’s Articles of Incorporation, art IV(D)(1-2).

⁸⁶ Anthropic’s Articles of Incorporation, art VIII(B)(2)(b) and VIII(B)(6)(v).

⁸⁷ Anthropic’s Articles of Incorporation, art VIII(B)(3)(b) and VIII(B)(6)(vi-vii).

⁸⁸ Anthropic’s Articles of Incorporation, art VIII(B)(4)(b) and VIII(B)(6)(iv).

composed of five members in total,⁸⁹ a majority of 60% of LTBT directors will eventually control the company.

Despite the peculiar nature and incentives of the LTBT, the company's directors still have fiduciary duties towards common shareholders, which is a necessary consequence of Anthropic being a corporation, as we further explain in Section 3.1.1. These duties may be violated if the directors were to completely disregard the company's for-profit nature while pursuing its public mission. Furthermore, the nonpublic nature of the trust agreement governing the LTBT makes it impossible to have a clear idea of LTBT board members' incentives to run the company in a specific way,⁹⁰ though some key features of the LTBT have been disclosed and seem to adequately align the company's public mission and the trust's purpose.⁹¹ Likewise, Anthropic's articles of incorporation do not say much about shareholders' voting powers, the crux of which is contained in the bylaws.⁹² Subject to that, shareholders may very well exercise their voting rights in a sense that privileges profit maximisation over the pursuit of humanity's interests – although the board may, at its sole initiative, modify the bylaws to constrain shareholders' voting powers to prevent this.⁹³

The role played by Anthropic's shareholders in the company's governance enables them to provide a counterpower to the board in case it violates its fiduciary duties – contrary to OpenAI's board, which de jure seems immune to revocation (though not de facto, as Sam Altman's comeback and OpenAI board members' subsequent departure suggests). This led commentators to consider Anthropic's governance structure as more resilient than that of OpenAI.⁹⁴ This role of shareholders is all the more likely to be central (but also criticised) given that companies such as Google and Amazon took significant stakes in Anthropic's capital,⁹⁵ though they reportedly do not hold voting rights.⁹⁶ At the same time, this makes it impossible, both de jure and de facto, to truly isolate board members' pursuit of humanity's interests from financial considerations. Indeed, as Anthropic rightly points out, 'all the key mechanisms of

⁸⁹ Anthropic's Articles of Incorporation, art VIII(B)(2)(a), VIII(B)(3)(a) and VIII(B)(4)(a).

⁹⁰ Zach Stein-Perlman, 'Anthropic's Certificate of Incorporation' (*LessWrong*, 12 June 2024) <<https://www.lesswrong.com/posts/ogXkDBLyxby3TXXKm/anthropic-s-certificate-of-incorporation>> accessed 20 August 2024.

⁹¹ For a short explanation of the purpose and functioning of the LTBT (by Anthropic's own outside counsels), see John Morley, David Berger and Amy Simmerman, 'Anthropic Long-Term Benefit Trust' (Harvard Law School Forum on Corporate Governance, 28 Oct 2023) <<https://corpgov.law.harvard.edu/2023/10/28/anthropic-long-term-benefit-trust/>> accessed 20 August 2024.

⁹² As can be inferred from Anthropic's Articles of Incorporation, art IV(B)(5) and IV(C)(5).

⁹³ Anthropic's Articles of Incorporation, art V.

⁹⁴ Ted Ladd, 'The Lesson Of OpenAI Is That Governance Matters' *Forbes* (15 Dec 2023) <<https://www.forbes.com/sites/teidladd/2023/12/15/the-lesson-of-openai-is-that-governance-matters/>> accessed 20 August 2024.

⁹⁵ Jackie Davalos and Brad Stone, 'OpenAI Rival Anthropic Defends Partnerships With Amazon, Google' *Bloomberg* (9 May 2024) <<https://www.bloomberg.com/news/articles/2024-05-09/openai-rival-anthropic-defends-partnerships-with-amazon-google>> accessed 20 August 2024.

⁹⁶ Billy Perrigo, 'How Anthropic Designed Itself to Avoid OpenAI's Mistakes' *Time* (San Francisco, 30 May 2024) <<https://time.com/6983420/anthropic-structure-openai-incentives/>> accessed 20 August 2024.

accountability in corporate law push directors to prioritize the financial interests of stockholders',⁹⁷ due to the specific nature of the fiduciary owed by the former to the latter.

We get back to the implications of corporations' fiduciary duty regime this in greater detail in Section 3.1.1. We simply need to stress here that, according to Anthropic, this prioritisation of stakeholders' interests is not a bug, but a feature: 'public benefit is not at odds with commercial success or stockholder returns, and if anything our experience has shown that the two are often strongly synergistic: our ability to do effective safety research depends on building frontier models (the resources for which are greatly aided by commercial success), and our ability to foster a "race to the top" depends on being a viable company in the ecosystem in both a technical sense and a commercial sense'. For this reason, Anthropic '[does] not expect the LTBT to intervene in these day-to-day decisions or in [its] ordinary commercial strategy'.⁹⁸ Instead, 'the need for fine-tuning of the governance structure ultimately derives from the potential for extreme events and the need to handle them with humanity's interests in mind, and we expect the LTBT to primarily concern itself with these long-range issues. For example, the LTBT can ensure that the organizational leadership is incentivized to carefully evaluate future models for catastrophic risks or ensure they have nation-state level security, rather than prioritizing being the first to market above all other objectives'.⁹⁹ Thus, the LTBT acts as a 'backstop' preventing Anthropic's leadership from taking catastrophic risks for the sake of making profits, in a similar fashion to what the risk management team of a bank would do.

Besides OpenAI and Anthropic, several US AI companies did not implement such sophisticated and original governance structures, even though they incorporated as public benefit corporations. We next discuss a few important examples.

2.3. The Vanilla Benefit Corporations: Inflection AI and xAI

Inflection AI, the creator of Pi (which stands for 'personal intelligence'¹⁰⁰), is another competitor of OpenAI and Anthropic. Inflection claims that 'safety is core to [its] culture and [its] top priority'.¹⁰¹ Like Anthropic, Inflection relies on a Delaware PBC.¹⁰² However, contrary to Anthropic, Inflection does not discuss the existential risk for humanity posed by AI on its webpage about safety, even though this risk is primarily what justifies relying on a governance structure that puts the interests of humanity above profits.¹⁰³

⁹⁷ Anthropic, 'The Long-Term Benefit Trust' (19 Sept 2023) <<https://www.anthropic.com/news/the-long-term-benefit-trust>> accessed 20 August 2024.

⁹⁸ Ibid.

⁹⁹ Ibid.

¹⁰⁰ Inflection, 'An Inflection Point' (8 March 2023) <<https://inflection.ai/an-inflection-point>> accessed 21 August 2024.

¹⁰¹ Inflection, 'FAQ' (2024) <<https://inflection.ai/>> accessed 21 August 2024.

¹⁰² Inflection, 'An Inflection Point' (8 March 2023) <<https://inflection.ai/an-inflection-point>> accessed 21 August 2024.

¹⁰³ See Inflection, 'Our policy on frontier safety' (30 Oct 2023) <<https://inflection.ai/frontier-safety>> accessed 21 August 2024. For the same observation, see nikola, 'Inflection.ai is a major AGI lab'

It is worth noting that Inflection’s CEO Mustafa Suleyman, a top AI figure who co-founded AI company DeepMind which was later acquired by Google, was recently hired by Microsoft together with former employees of Inflection. These hirings, which were considered as a merger by the UK Competition and Markets Authority (though not one worthy of competition concerns),¹⁰⁴ were part of a broader agreement between Microsoft and Inflection and followed a \$1.3 billion investment of Microsoft and Nvidia in Inflection, based on a \$4 billion valuation.¹⁰⁵ The acquisition by both of these companies of stakes in Inflection’s capital makes it very unlikely that Inflection will remain able or willing to resist these companies’ influence, should it conflict with the company’s alleged concern for safety.

Another interesting example is that of Elon Musk’s AI company xAI, incorporated as a Nevada benefit corporation (BC) and whose declared purpose is to ‘understand the universe’.¹⁰⁶ This company is subject to Nevada law, which as we shall see later contains a specific regime for ‘benefit enforcement proceedings’, allowing third parties to obtain injunctive measures in court if the corporation violates its purposes. From this perspective, xAI’s corporate regime widely differs from those of Anthropic and Inflection, despite all three companies being called ‘public benefit corporations’.

2.4. Safety without Governance: European Companies

Other companies in the AI sector also claim to be preoccupied by safety and humanity’s interests but did not implement any governance mechanisms requiring or incentivising their directors, in law or in fact, to pursue these objectives. German company Aleph Alpha, for instance, claims to ‘firmly believe in the power of AI to enhance the intelligence and productivity of the human workforce, while upholding ethical AI practices and responsible AI standards’.¹⁰⁷ The company presents financial performance and the upholding of ethical values as aligned: ‘trust is the cornerstone of any successful enterprise or government entity. With the unique explainability and control features included in the Aleph Alpha offering, [users] can rest assured that [their] AI-driven processes and decisions are based on transparent methods, traceable knowledge and aligned with [their] ethical values’.¹⁰⁸ Aleph Alpha does not specify whether any specific corporate governance safeguards are implemented to ensure

LessWrong (9 Aug 2023) <<https://www.lesswrong.com/posts/Wc5BYFfzuLzepQjCq/inflection-ai-is-a-major-agi-lab>> accessed 21 August 2024.

¹⁰⁴ Competition and Markets Authority, *Decision on Microsoft Corporation's Hiring of Certain Former Employees of Inflection and Its Entry Into Associated Arrangements* (4 September 2024) CMA/14/24.

¹⁰⁵ Krystal Hu and Harshita Mary Varghese, ‘Microsoft pays Inflection \$650 mln in licensing deal while poaching top talents, source says’ *Reuters* (22 March 2024) <<https://www.reuters.com/technology/microsoft-agreed-pay-inflection-650-mln-while-hiring-its-staff-information-2024-03-21/>> accessed 21 August 2024.

¹⁰⁶ xAI, ‘Our Mission – Understand the Universe’ <<https://x.ai/about>> accessed 21 August 2024.

¹⁰⁷ Aleph Alpha, ‘Join Our Mission’ <<https://aleph-alpha.com/join-our-mission/>> accessed 21 August 2024.

¹⁰⁸ Aleph Alpha, ‘The Augmented Intelligence Platform’ <<https://aleph-alpha.com/>> accessed 21 August 2024.

that ethical considerations are adequately taken into account. At any rate, as we shall see in the next section, German corporate law only provides companies with limited flexibility when it comes to tailoring their governance to the objectives pursued.

Several other AI companies claim to place safety at the centre of their preoccupations but do not explain how they guarantee it. UK-based AI company Stability.ai, for instance, claims to '[advance] safe and transparent AI with open models' by relying on six 'Safety Principles' but does not clarify how these principles are actually enforced, let alone at the top management level. Stability.ai's former CEO and controlling shareholder Emad Mostaque voluntarily stepped down on 22 March 2024 and stated that 'we should have more transparent & distributed governance in AI as it becomes more and more important. It's a hard problem, but I think we can fix it [...] The concentration of power in AI is bad for us all. I decided to step down to fix this at Stability & elsewhere'.¹⁰⁹ Similarly, Google's AI subsidiary DeepMind, co-founded by Inflection's CEO Mustafa Suleyman and which developed Bard, one of ChatGPT's main competitors, claims to be 'a team of scientists, engineers, ethicists and more, working to build the next generation of AI systems safely and responsibly'¹¹⁰ whose mission is to 'build AI responsibly to benefit humanity'. However, DeepMind does not seem to have implemented governance mechanisms guaranteeing that the benefit of humanity prevails over profit – even though its holding company Google is, itself, a 'full' for-profit company.

As another example, NXAI is an Austrian company engaged in the development of the European LLM technology xLSTM. The entity is structured as a joint venture company with three shareholders: two of them are existing technology companies PIERER Digital Holding, with a share capital of 37%, and Netural X, which holds another 37%.¹¹¹ The remaining 26% are held by NXAI's founder, AI researcher Sepp Hochreiter. Whilst the granularities of the deal are private, the significant shareholding of the company's founder indicates that Hochreiter wanted to ensure material influence over the entity's evolution. According to NXAI's website, each shareholder brings different advantages to the deal: 'NXAI is composed of a harmonious triad of elite AI research, digital product know-how, and industry expertise'.¹¹² While the company is clearly oriented towards 'unlock[ing] the vast potential of AI for industrial application', it is also committed to 'construct[ing] AI systems that obey

¹⁰⁹ PYMNTS, 'Stability AI CEO Resigns and Calls for 'Transparent' AI Governance' *PYMNTS* (24 March 2024) <<https://www.pymnts.com/artificial-intelligence-2/2024/stability-ai-ceo-resigns-and-calls-for-transparent-ai-governance/>> accessed 21 August 2024; Manish Singh, 'Stability AI CEO resigns because you're 'not going to beat centralized AI with more centralized AI'' *TechCrunch* (22 March 2024) <<https://techcrunch.com/2024/03/22/stability-ai-ceo-resigns-because-youre-not-going-to-beat-centralized-ai-with-more-centralized-ai/>> accessed 21 August 2024.

¹¹⁰ Google, 'Build AI responsibly to benefit humanity' *DeepMind* <<https://deepmind.google/about/>> accessed 21 August 2024.

¹¹¹ Bpv Hügel, 'bpv Hügel advises top researcher and AI pioneer Sepp Hochreiter on founding the AI company NXAI' (24 February 2024) <<https://www.bpv-huegel.com/wp-content/uploads/2024/02/240214-bpv-Huegel-advises-top-researcher-and-AI-pioneer-Sepp-Hochreiter-on-founding-the-AI-company-NXAI.pdf>> accessed 21 August 2024.

¹¹² NXAI, 'About Us' <<https://www.nx-ai.com/en/about>> accessed 21 August 2024.

European values and ethics standards'.¹¹³ However, the company has not publicly shared whether it implemented governance mechanisms aiming to further these goals.

2.5. Companies Without a Safety Statement

Last but not least, some AI companies do not express explicit concern over ethical or safety considerations. Notably, 'Europe's AI champion', the French start-up Mistral, doesn't make any reference to the public interest: '[o]ur mission is to make frontier AI ubiquitous [...]. This requires fierce independence, strong commitment to open, portable and customisable solutions.'¹¹⁴ The fact that this mission statement does not refer to safety or ethics can be attributed to Mistral's specific business model. Contrary to OpenAI and the likes, Mistral's products are 'open-weight models for everyone to customize and deploy where they want it'.¹¹⁵ The company doesn't commercialize applications for end users, but models that can be customized by developers and businesses. Given that Mistral sells tools rather than products, it most likely does not consider itself responsible or liable for how its models are customised and used. Many voices go so far as to argue that providers of open-source models should be entirely exempt from liability.¹¹⁶

Another tentative explanation for some AI companies not showing this kind of concern for safety, ethics or humanity's interests may be that OpenAI, Anthropic and xAI are much bigger companies than their competitors in the AI ecosystem, with valuations amounting to \$150bn, \$18bn and \$45bn, respectively.¹¹⁷ One may be tempted to argue that, because these very large players have purpose-led governance structures but smaller market actors do not, such structures are only useful for very large companies which develop the most successful generative AI products and can credibly participate in the race to AGI. However, we do not see any reason why this would be the case. First, as we saw in this section, smaller companies also show concern for AI safety, and may thus be candidates to purpose-led governance structures. Second, as explained in Section 1, a company need not develop such a GPT as ChatGPT or Claude to expose humanity in general, or their clients in particular, to considerable risk due to misuse or misalignment. This leaves the possibility that at least some AI companies may fail to adopt a purpose-led governance structure not because they do not

¹¹³ Ibid.

¹¹⁴ Mistral AI, 'Frontier AI in Your Hands' (Mistral AI) <<https://mistral.ai/>> accessed 24 November 2024.

¹¹⁵ Ibid.

¹¹⁶ John Speed Meyers and Paul Gibert, 'Questioning the Conventional Wisdom on Liability and Open Source Software' *Lawfare* (18 April 2024) <<https://www.lawfaremedia.org/article/questioning-the-conventional-wisdom-on-liability-and-open-source-software>> accessed 24 November 2024; Thomas Depierre, 'I Am Not a Supplier' *Software Maxims* (31 December 2022) <<https://www.softwaremaxims.com/blog/not-a-supplier>> accessed 24 November 2024; Nadia Eghbal, 'Roads and Bridges: The Unseen Labor Behind Our Digital Infrastructure' *Ford Foundation* (14 July 2016) <<https://www.fordfoundation.org/work/learning/research-reports/roads-and-bridges-the-unseen-labor-behind-our-digital-infrastructure/>> accessed 24 November 2024.

¹¹⁷ Based on their last funding rounds: Madhumita Murgia, 'Amazon pumps another \$4bn into AI start-up Anthropic' *Financial Times* (London, 22 Nov 2024) <<https://www.ft.com/content/01da4ed8-0371-44b0-a07d-63e07f32939e>> accessed 23 Nov 2024.

want to, but because their jurisdiction of incorporation does not leave them enough flexibility to do so. We discuss this question at a theoretical level in the next section before seeing in Section 4 the constraints several US and EU jurisdictions impose on purpose-led entities.

3. Legal Entities as Cost-Reducing Devices

As we showed in the previous section, AI companies are not just at the forefront of technological research; they are also pushing the boundaries of the organization of industry. AI players went great lengths to create tailor-made structures able to reconcile profits with the interests of humanity. We focus here on the role legal entities and rules applicable to them play in allowing or restricting these companies' ability to tailor their governance structure to their peculiar needs. Section 2.1 addresses this question at a high-level theoretical level. Section 2.2 then elicits the peculiar features legal entities must have for purpose-led entities to be effectively designed and implemented: a well-defined purpose, tailor-made fiduciary duties and legal actions, and an adequate governance structure.

3.1. Legal Entities as Governance Tools

By providing a template for structuring contractual relationships, legal entities 'serve as a nexus *for* contracting relationships'.¹¹⁸ There is a contentious debate in organisational law as to what is the role of standard legal entities and the extent to which these legal forms achieve any functionalities which could not be secured by contracting alone.¹¹⁹ While engaging with this debate is beyond the scope of our paper, we note that organisational law's utility does not (only) reside in surpassing the boundaries of contract law or in providing for features that contracts cannot achieve. Organisational law 'eliminates the need for [...] elaborate contracting and thereby avoids the transaction costs and moral hazard it involves'.¹²⁰ Thus, one of the key roles of organisational law in general, and of standard legal entities in particular, is reducing the costs of doing business.

The creation of bespoke legal structures, such as the one put in place by OpenAI, comes with hefty legal fees and considerable legal uncertainty as to how these structures will hold in court.¹²¹ Should there exist a standard legal form enabling companies to distribute profits while putting humanity first, these costs and the moral hazard involved therein would be

¹¹⁸ Michael C. Jensen and William H. Meckling, 'Theory of the Firm: Managerial Behavior, Agency Costs and Ownership Structure' (1976) 3(4) *Journal of Financial Economics* 305.

¹¹⁹ Hansmann and Kraakman argue that the only such feature is the ability of the corporate form to carve out a separate category of assets that belong to the corporate entity and cannot be reached by the creditors of the entity's owners and managers, i.e. entity shielding. Henry Hansmann and Reinier Kraakman, 'Organizational Law as Asset Partitioning' (2000) 44 *European Economic Review* 807.

¹²⁰ Henry Hansmann and Reinier Kraakman, 'The Essential Role of Organisational Law' (2000) 110 *Yale Law Journal* 387, 408.

¹²¹ Wassermann (n 50).

significantly diminished, compared to the alternative of setting up a complex web of contracts and legal entities.¹²²

Building on the vision that the role of organisational forms is to reduce the costs associated with doing business, the answer to the second question – whether and to what extent new standard forms should be created – is clearly in the affirmative. This is not a novel argument. In a paper dating back to the last decade, George S. Geis argues that technological advancements have led to the relocation of production organisationally towards hybrid institutional settings.¹²³ To stay true to its mission of reducing the costs of doing business, organisational law needs to modularise and adapt itself to changes in the organisation of industry. In the author's view, we are witnessing 'a transformation from discrete to continuous options for slicing ownership, control, commitment, and risk'.¹²⁴ This is precisely what we see with AI companies: the need for new organisational forms which separate profit rights from governance rights in order to ensure that the interests of humanity are safeguarded. Given these phenomena, the law should provide standardised forms allowing this separation.

If we go back in time even more, it becomes clear that legal forms, and the corporate form in particular, were time and again modularised to respond to the demands of doing business. The story of the liberalisation of the corporate form in the US illustrates this dynamic. Corporate charters were initially granted only by states and were highly restrictive: they 'specified a capital structure, a governance structure, and enumerated some of the privileges to which the corporation was entitled', and even 'imposed specific regulations on the operations of corporations', such as 'detailed specifications of the route to be followed for ships' or 'interest rate ceilings for banks'.¹²⁵ General incorporation laws were adopted in the US as early as 1811, but they still imposed relatively rigid templates on the firms whose creation they enabled.¹²⁶ From these restrictive charters to today's enabling regime, organisational law went a long way.

To showcase how corporate law was liberalised as a response to the evolution of the organisation of industry, take the example of New Jersey's adoption of the first holding company act in 1891.¹²⁷ Prior to its adoption, corporations were not allowed to own shares

¹²² This is due, inter alia, to the network effects of having various companies adopt the same kind of governance structures and the benefit of well-tested default rules. On this topic, see Michael Klausner, 'Corporations, Corporate Law, and Networks of Contracts' (1995) 81 *Virginia Law Review* 757; Michael Klausner, 'Fact and Fiction in Corporate Law and Governance' (2013) 65 *Stanford Law Review* 1325.

¹²³ George S. Geis, 'The Space Between Markets and Hierarchies' (2009) 95(1) *Virginia Law Review* 99

¹²⁴ *Id.* 103.

¹²⁵ Eric Hilt, 'Early American Corporations and the State' in Naomi R. Lamoreaux and William J. Novak, *Corporations and American Democracy* (Harvard University Press 2017) 37.

¹²⁶ Ronald E. Seavoy, 'Laws to Encourage Manufacturing: New York Policy and the 1811 General Incorporation Statute' (1972) 46(1) *The Business History Review* 85.

¹²⁷ In its first ever merger decision, the US Supreme Court held a New Jersey holding company liable under the Sherman Act's prohibition of cartels. *Northern Securities Company v. United States*, 193 U.S. 197 (1904).

and could thus not merge with each other or organise as groups. With rapid industrialisation, the dictates of economies of scale and scope required businesses to integrate upstream and downstream. To bypass this limitation, companies created the trust form, but with the rise of the antitrust movement, this savvy workaround was no longer available. As Ernst von Halle noted, the New Jersey act was ‘a transformation of the corporation law which facilitated the remodelling of the trusts, and their continued transaction of business in the state’.¹²⁸

As this story illustrates, organisational law is malleable to the demands of the organisation of industry and can be adjusted to enable the functioning of new entities. We next discuss how exactly this imperative translates in the context of purpose-led AI companies.

3.2. Legal Entities as Tools to Implement Purpose

Corporate purpose may be defined as ‘the reason a company exists’.¹²⁹ As this definition suggests, companies are not tied to the mere maximisation of profits. This is usually discussed through the prism of the so-called ‘stakeholder approach’ or ‘stakeholderism’, according to which companies should (be able to) pursue another purpose alongside profits. Whereas stakeholderism has been in the spotlight for decades, the emergence of AI gives this discussion a new dimension. The originality of AI companies is that oftentimes, they do not simply claim to pursue a public mission alongside profit maximisation. Instead, they implement concrete and sometimes sophisticated governance mechanisms to that end. OpenAI, in particular, not only declares to be first and foremost concerned with the interests of humanity, but also designed a governance structure where the company’s non-profit mission of fostering humanity’s interests entirely subsumes its ability to make profits. Profit is merely presented as a way to reach OpenAI’s non-profit objective, rather than an end in itself.

To be sure, OpenAI is an isolated example. To the best of our knowledge, no other company, Anthropic included,¹³⁰ qualifies as a profit-distributing nonprofit. However, more companies may very well be willing to use this type of structure if applicable corporate rules made it easier. In turn, companies’ ability to define their corporate purpose is a necessary condition for these companies to impose the primacy of a nonprofit goal on their directors, such as AI safety and the interests of humanity. But it is by no means a sufficient one. Merely stating that a specific purpose is pursued is very far from guaranteeing that the company’s managers and

¹²⁸ Ernst von Halle, *Trusts, or, Industrial Combinations and Coalitions in the United States*, (Macmillan and Co. 1896) 97.

¹²⁹ <https://www.unpri.org/corporate-purpose/corporate-purpose-what-investors-need-to-know/10729.article>

¹³⁰ To be more specific, Anthropic sees financial performance and humanity’s interests as aligned in normal times, and claims that humanity’s interests should prevail over financial performance in abnormal times (that is, when a ‘catastrophic risk’ is at stake). Therefore, Anthropic implicitly puts humanity’s interests above profits. Nevertheless, neither its governance structure nor its mission statement explicitly establish such a hierarchy.

directors will actually walk the talk.¹³¹ Taking purpose seriously requires creating a ‘purpose ecology’¹³² incentivising directors and managers to fulfil the company’s public mission.

From a theoretical perspective, two categories of legal devices can be used to tailor the governance structure to the company’s purpose. First, companies can implement ‘structural’ governance mechanisms whereby decision-making power is allocated to the company’s stakeholders according to the company’s purpose. This is what OpenAI and Anthropic did, by giving the power to appoint all or part of their respective directors to another constituency than their investors. As we shall see, US and EU laws usually provide companies with enough flexibility to adopt original structural mechanisms allocating voting and decisional power within the company to specific stakeholders with adequate incentives.

Second, companies and their investors should be able to align directors’ incentives with the interests pursued by the company by tailoring the scope and enforcement mechanisms of directors’ fiduciary duties. In most jurisdictions, directors owe the company and/or shareholders, and sometimes third parties (at least on paper), some variant of the duties of care and of loyalty that are prevalent in the US. In particular, ‘duty-of-loyalty doctrines encompass a variety of labels across jurisdictions, such as the duty of entire fairness, the prohibition against “wrongful profiting from position,” or the crimes of “abuse of corporate assets” (in France), and breach of trust (in Germany)’.¹³³ Furthermore, ‘[the] bite [of the duty of loyalty] crucially depends on how often and how stringently courts enforce it’.¹³⁴ While courts cannot substitute themselves to the directors in designing corporate strategy, fiduciary duties can be triggered whenever directors egregiously violate their duty to act in the best interest of a given constituency. For instance, directors of an AI company with fiduciary duties towards humanity at large could be considered as violating their fiduciary duties if they took a course of action that clearly privileges profit maximisation over humanity’s interests.

While structuring an effective fiduciary duty regime for mainstream companies is not an easy task, adapting it to AI companies with a nonprofit mission poses two additional related challenges. On the one hand, directors’ fiduciary duties should be designed in such a way as to insulate them from the threat of being sued by shareholders when privileging purpose over profits., On the other hand, because these duties often involve protecting other interests than those of the shareholders, the pool of potential claimants should extend beyond them. Indeed, discarding directors’ liability to further shareholders’ interest without concomitantly imposing on them an actionable duty to further the company’s purpose would decrease the

¹³¹ For a recent empirical study of the US market, see eg Lucian A Bebchuk and Roberto Tallarita, ‘Will Corporations Deliver Value to All Stakeholders?’ (2022) 75 *Vanderbilt Law Review* 1031. See also Leo E Strine, Jr, ‘The Dangers of Denial: The Need for a Clear-Eyed Understanding of the Power and Accountability Structure Established by the Delaware General Corporation Law’ (2015) 50 *Wake Forest Law Review* 761.

¹³² David Kershaw and Edmond Schuster, ‘The Purposive Transformation of Corporate Law’ (2022) 69 *The American Journal of Comparative Law* 478, 482.

¹³³ Kraakman et al, *The Anatomy of Corporate Law: A Comparative and Functional Approach* (3rd edn, OUP 2017) 161.

¹³⁴ *Ibid* 164.

scope of their liability with no guarantee that they will actually seek to promote that purpose (rather than, say, further their own interests or simply shirk).¹³⁵ This pool could include anything from a few public authorities or qualified associations to a specific subset of the population to humanity at large. A well-balanced enforcement regime should both incentivise potential claimants to enforce the company's purpose and protect the company and its directors from frivolous litigation.¹³⁶ Logically, the larger the pool of potential claimants, the more difficult it is to strike the right balance.

Third parties may be entitled to sue the company's directors either as personal claimants or derivatively, on behalf of the company. Personal claims may have reasonable grounds to succeed subject to directors' failure to pursue their company's nonprofit mission causing observable harm to the third parties bringing the action against the directors. This will typically not be the case when the harm caused by the company's activities is too diffuse and indirect for claimants to prove the existence of a personal prejudice. Moreover, the harm associated with such activities may not have materialised yet. For instance, directors' failure to, say, implement necessary safety processes within the company to avoid the development of a dangerous technology is very unlikely to be a valid ground for third parties' claims. Finally, aggrieved third parties may find it difficult to effectively coordinate to protect their interests, unless they can count on intermediaries, such as consumer associations, to do so on their behalf (which, in turn, may come with their own set of agency problems).

Derivative claims are usually open only to the company's shareholders (or, under certain circumstances, its creditors). However, opening these actions to third parties may make sense, at least to the extent that they would allow them to obtain something else than damages. For instance, the law may entitle third parties to ask the court for injunctive relief on behalf of the company in case directors adopt a course of action that egregiously departs from the company's purpose.

These are only illustrations of what could be done to make corporate law amenable to the enforcement of purpose: many other legal devices may be imagined, and the most effective and realistic ways to allow third parties to obtain the compliance of directors and the company itself with its purpose will necessarily vary from jurisdiction to jurisdiction. As a last remark, the fact that a director may, in theory, be found liable for pursuing a nonprofit mission at the expense of profit does not necessarily entail that this liability is likely to be triggered in practice. In many cases, the pursuit of nonprofit interests (such as humanity's betterment) can be framed as contributing to the long-term generation of profits and thus the company's interests.¹³⁷ Furthermore, investors may be unwilling to sue directors due to reputational

¹³⁵ Oliver Hart, 'An Economist's View of Fiduciary Duty' (1993) 43 *The University of Toronto Law Journal* 299, 303.

¹³⁶ In this sense, see *Zapata Co. v Maldonado*, 430 A2d 779, 786-789 (Del 1981).

¹³⁷ Einer Elhauge, 'Sacrificing Corporate Profits in the Public Interest' (2005) 80 *New York University Law Review* 733 (n X). This is also Anthropic's argument for being incorporated as a for-profit company, already exposed in Section 1.2.

concerns,¹³⁸ and courts may be reluctant to condemn them based on the pursuit of a charitable purpose which the company itself advocates for. Nevertheless, inadequate fiduciary duties may dissuade directors from taking decisions enhancing the company's nonprofit mission.

4. The Corporate Law of Purpose and its Implementation

Section 1 showed that most AI companies seek to pursue a certain ideal of the greater good, or at the very least express concern about AI safety. However, none of them is structured as a purely altruistic endeavour. To the contrary, all of these companies give substantial importance to their ability to make profits. OpenAI's experience showed that, even backed by generous donors, top AI companies are unlikely to be able to develop tomorrow's AI technology without either relying on substantial profit-motivated investments or being backed by the public sector. This requires designing tailor-made governance structures adequately incentivising management to address whatever public mission the company assigns to itself, from preventing excessive risk-taking to putting the interests of humanity at the heart of the company's mission.

AI companies based in the US could adopt governance structures tailored to their public mission due to the flexibility of US state laws. As we will see in this section, Delaware and Nevada corporate laws, though not perfect in every respect, leave quite significant flexibility to AI companies to tailor their governance structure to their specific purpose. The reason why we focus on these two states is, first, that Delaware is by far the privileged jurisdiction of incorporation for large US companies.¹³⁹ In particular, Delaware is where most of the biggest AI companies in the world, including OpenAI, Anthropic and Inflection AI, are incorporated. Second, and contrary to Delaware, Nevada reportedly is one of several US states to have modelled its benefit corporation law after the 'Model Benefit Corporation Legislation' (MBCL), which allows third parties to obtain injunctive measures as part of 'benefit enforcement proceedings' in case directors violate the company's purpose,¹⁴⁰ as discussed

¹³⁸ Gad Weiss, 'Aligned Structuring of AI Startups' (2024) SSRN working paper 50-51 <https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4937578> accessed 28 August 2024.

¹³⁹ See eg Government of Delaware, 'Delaware Division of Corporations: 2023 Annual Report' <<https://corp.delaware.gov/stats/>> accessed 17 Nov 2024.

¹⁴⁰ We found this information in Jaime Lee, 'Benefit Corporations: A Proposal for Assessing Liability in Benefit Enforcement Proceedings' (2017) 103 Cornell Law Review 1075, 1081 which reports that 'thirty-three states have passed statutes recognizing the benefit corporation as a new business form' and that 'most of the statutes were modeled after the Model Benefit Corporation Legislation'. This information is corroborated by various sources on the internet, and many states have similar benefit corporation laws to that of Nevada. Nevertheless, and despite the major importance of the MBCL, we have been unable to find anywhere the official version of the MBCL or a reliable confirmation of this figure, be it on B Lab's website (the reported creator of the MBCL: <<https://usca.bcorporation.net/benefit-corporation/>> accessed 17 November 2024) or anywhere else. The links contained in the paper cited above are outdated. One paper reports that at least twenty-seven states have adopted benefit enforcement proceedings (Summer Brown, 'Benefit Corporations: A Case Study in the Issues of Implementation and Adoption of the Fastest Growing Business Form in the United States' (2016) 35 Business & Professional Ethics Journal 199, 211). The Nevada State Bar also explicitly indicates that Nevada relied on the MBCL for its benefit corporation law: State Bar of Nevada, 'Benefit Corporation

below. Nevada is also the jurisdiction of incorporation of xAI. Delaware law and the MBCL are sometimes presented as the two competing legislative models followed by US states.¹⁴¹

This brings us to ask whether European law would offer as much flexibility to AI companies incorporated in the European Union. To answer this question, we cover the corporate laws from three European jurisdictions – France, Germany and Italy – and show that they offer unequal flexibility when it comes to defining the company’s purpose and tailoring its govern. As already explained in the introduction, we focus on these jurisdictions both because of their weight in the European economy and because of the influence of these jurisdictions’ legal traditions on other European countries.

This section shows to what extent Delaware, Nevada, French, German and Italian corporate laws allow or prevent companies from tailoring them to their needs. The company’s purpose and its directors’ fiduciary duties being conceptually strongly interrelated, we start in Section 4.1 with a discussion of ‘functional’ governance mechanisms, that is, the definition of purpose and its enforcement through fiduciary duties. We then discuss ‘structural’ governance safeguards – that is, the repartition of decisional power within the company – in Section 4.2.

4.1. Functional Governance Mechanisms: Purpose and Fiduciary Duties

There should exist a strong link between a company’s purpose and the fiduciary duties of its directors. The directors (and, as the case may be, the company’s managers) being in charge of the management and supervision of the company, they should be entrusted with the task of furthering its purpose. In turn, the scope of directors’ duties should depend on that purpose. While Delaware and Nevada laws establish a clear link between the company’s purpose and directors’ fiduciary duties, this link is much more tenuous in our selected European jurisdictions. We show this disparity by discussing the US and European corporate laws in turn.

4.1.1. Functional Governance Mechanisms in the United States

We start our discussion with an analysis of Delaware law before discussing a few peculiar features of Nevada law (which, as already discussed, are common to all states whose benefit corporation legislations are based on the MBCL).

4.1.1.1. Functional Governance Mechanisms under Delaware Law

Legislation: Safeguards for Businesses Giving Back’
<<https://nvbar.org/nvlawyermagazine/february-2013/>> accessed 17 November 2024.

¹⁴¹ See eg Summer Brown, ‘Benefit Corporations: A Case Study in the Issues of Implementation and Adoption of the Fastest Growing Business Form in the United States’ (2016) 35 Business & Professional Ethics Journal 199, 201 (also reporting that ‘Washington, California, and Florida have a slight mix of the Delaware and MBCL legislation’: *ibid* 213).

Delaware law mainly offers two possible legal forms for profit-making entities which aim to pursue another purpose than profit: PBCs (Anthropic's corporate form) and LLCs (the form chosen by OpenAI for the for-profit arm). We discuss each of them in turn, as each leaves investors a very different degree of flexibility when it comes to defining the company's purpose and directors' duties.

Admittedly, Delaware law grants corporations substantial flexibility in tailoring their governance structure to their specific needs, for instance by allowing the issuance of non-voting shares.¹⁴² Nevertheless, the law is more constraining when it comes to shaping directors' fiduciary duties. First, Delaware law offers corporations little flexibility in restricting or excluding directors' fiduciary duties, which goes to the crux of OpenAI's needs.¹⁴³ In corporations, directors owe a duty of loyalty to all 'stockholders in the aggregate in their capacity as residual claimants, [...] without regard to any special rights'. Furthermore, they are required 'to promote the value of the corporation for the benefit of its stockholders'.¹⁴⁴ Consequently, directors must 'treat stockholder welfare as the only end, considering other interests only to the extent that doing so is rationally related to stockholder welfare'.¹⁴⁵ Whenever different classes of shareholders exist, the reference class for assessing directors' compliance with their fiduciary duties is that of the common shareholders, and favouring the interests of specific classes can lead to a breach of the fiduciary duty of loyalty irrespective of who controls the company.¹⁴⁶

Second, even public benefit corporations benefit from limited flexibility in solving the tension between their public benefit purpose and the pursuit of profit. Under Delaware law, a corporation can qualify as a PBC subject to '[being] intended to produce a public benefit or public benefits and to operate in a responsible and sustainable manner'.¹⁴⁷ While Delaware law does not address the question of how purpose and profit can be reconciled directly, it does so indirectly by clarifying that directors' fiduciary duties is to balance 'the pecuniary interests of the stockholders, the best interests of those materially affected by the corporation's conduct, and the specific public benefit or public benefits identified'.¹⁴⁸ In accordance with the 'business judgement rule', a director will not breach his fiduciary duties 'if such director's decision is both informed and disinterested and not such that no person of ordinary, sound

¹⁴² Delaware General Corporation Law, § 151(a).

¹⁴³ Delaware General Corporation Law, § 102(b)(7). On the mandatory scope of the duty of loyalty held by board members of a Delaware corporation to the shareholders, see Gad Weiss, 'The Venture Corporation' (2023) SSRN working paper 27-36 <https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4338030> accessed 28 August 2024; Gad Weiss, 'Aligned Structuring of AI Startups' (2024) SSRN working paper 44-45 <https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4937578> accessed 28 August 2024.

¹⁴⁴ *Frederick Hsu Living Tr v ODN Holding Corp*, No CV 12108-VCL, 2017 WL 1437308, (Del Ch Apr 14, 2017), § 38.

¹⁴⁵ *Ibid* § 37 See also *Revlon, Inc v MacAndrews & Forbes Holdings, Inc*, 506 A2d 173 (Del 1986).

¹⁴⁶ *In re Trados Inc Shareholder Litig*, 73 A3d 17 (Del Ch 2013), §§ 41-42 and 62.

¹⁴⁷ Delaware General Corporation Law, § 362.

¹⁴⁸ Delaware General Corporation Law, § 365(a).

judgement would approve'¹⁴⁹—which means that the presumption of validity of their business decisions is very hard to overcome.¹⁵⁰ This provision must certainly be interpreted such that directors of a PBC may very well take decisions that imperil the company's public mission (because they represent a risk for humanity at large) for the sake of maximising profits without violating their fiduciary duties.¹⁵¹

In line with these observations, one author recently argued that 'the benefit corporation statutes that have been enacted are rife with vague, unenforceable language and toothless accountability mechanisms [...] unlikely to be upheld by courts while saddling directors with nebulous duties [that] increase liability risk and weaken shareholder rights'.¹⁵² As a result, the adoption by Anthropic and Inflection AI of the PBC form arguably provides, in itself, very limited guarantees that their directors will act upon the declared purpose of their respective companies.

Delaware corporate law does not allow directors to forsake the pursuit of profits for pursuing the company's public mission, let alone define the way in which profit maximisation and the public mission shall be reconciled.¹⁵³ This may be (part of) the reason why Anthropic's articles of incorporation do not provide that the company's directors shall never take any decision that may lead to a catastrophic risk for humanity, even though such a provision would have made sense given Anthropic's mission and governance structure. Relatedly, as Anthropic itself notes, directors' accountability mechanisms incentivise them to prioritise shareholders' interests.

¹⁴⁹ Delaware General Corporation Law, § 365(b).

¹⁵⁰ See Michael B Dorff, James Hicks and Steven Davidoff Solomon 'The Future or Fancy? An Empirical Study of Public Benefit Corporations' (2020) ECGI Law working paper <https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3433772> accessed 21 August 2024 and case law cited.

¹⁵¹ To be sure, managers and directors may be able to take decisions that harm the company's business based on its public mission, or even abstract ESG considerations when the company is not a PBC: see Michael B Dorff, James Hicks and Steven Davidoff Solomon, 'The Future or Fancy? An Empirical Study of Public Benefit Corporations' (2020) ECGI Law working paper 63 <https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3433772> accessed 21 August 2024 (n X) and case law cited; Brian Galle, 'Social Enterprise: Who Needs It?' (2013) 54 Boston College Law Review 2025. Still, based on existing (case) law, it seems that directors cannot be held liable for taking a decision going against the company's public mission so long as this decision can be justified based on financial grounds.

¹⁵² Cheng-Chi (Kirin) Chang, 'Substance over Symbolism: Do We Need Benefit Corporation Laws?' *Minnesota Law Review* 33, 34.

¹⁵³ Delaware case law further suggests that Delaware courts may be particularly deferential to the board's decisions when it comes to deciding how to balance profitability and public missions such as safety. In its famous case *Trinity Wall Street v Wal-Mart Stores, Inc.*, the Delaware Chancery Court ruled that Wal-Mart's board could lawfully exclude from its proxy materials a shareholder proposal aiming to forbid the sale of firearms with high-capacity magazines in supermarkets, on the grounds that this was an 'ordinary business decision' which was within the board's competence: *Trinity Wall Street v Wal-Mart Stores, Inc.*, 792 F 3d 323 (3d Cir 2015). While this decision was handed down in the specific context of shareholder proposals, it gives an idea of how Delaware courts may handle shareholder litigation asserting directors' failure to adequately balance purpose and profits.

To benefit from greater flexibility, Delaware law allows companies to opt for the LLC form. As compared to corporations (Delaware's flagship legal entity), LLCs have a much more flexible regime for fiduciary duties. As the Delaware LLC Act declares, 'it is the policy of [the LLC] to give maximum effect to the principle of freedom of contract'.¹⁵⁴ The law provides that fiduciary 'duties may be expanded or restricted or eliminated' in the LLC's operating agreement, with the only limitation being that 'the implied contractual covenant of good faith and fair dealing' cannot be contracted out.¹⁵⁵ This is echoed by the fact that LLCs can be structured both as for-profits or nonprofits¹⁵⁶ (including, as illustrated by OpenAI, profit-distributing nonprofits). However, OpenAI Nonprofit's directors are likely to play the role of OpenAI Global's managers, as already discussed. In this context, their fiduciary duties have probably been defined as being towards humanity at large and as not requiring them to maximise profits. This is all the more likely given that, according to OpenAI's website, OpenAI Global's activity is subject to OpenAI Nonprofit's mission to create 'safe AGI that is broadly beneficial', and that 'the Nonprofit's principal beneficiary is humanity, not OpenAI investors'.¹⁵⁷

However, these are only suppositions, as the granularities of OpenAI Global's agreement are unknown. Indeed, one of the key advantages of LLCs is that their agreements are private, meaning that 'that nobody who isn't aware of the operating agreement and any relevant collateral agreements can necessarily know the internal economic or governance structure of the entity'.¹⁵⁸ This means that, somewhat ironically given OpenAI's claim to act in humanity's interest, third parties have little visibility into the granularities of OpenAI Global's governance.¹⁵⁹

This rigidity probably explains in large part why OpenAI Global has been structured as an LLC and not a corporation. Indeed, corporations offer some technical advantages over LLCs,¹⁶⁰ such as its different tax regime.¹⁶¹ Furthermore, and perhaps more importantly, Delaware corporate case law is very developed and well-known by industry players, which limits legal uncertainty.¹⁶² This, in itself, has substantial value, especially for companies such

¹⁵⁴ Delaware Limited Liability Company Act, § 18-1101(b).

¹⁵⁵ Delaware Limited Liability Company Act, § 18-1101(c). In the absence of specific provisions to this effect in the LLC agreement, traditional fiduciary duties apply to LLCs under Delaware law: Kelly v Blum, Del Ch No 4516-VCP, Feb 24, 2010.

¹⁵⁶ Delaware Limited Liability Company Act, § 18-106(a).

¹⁵⁷ OpenAI, 'OpenAI Charter <<https://openai.com/charter/>> accessed 20 August 2024.

¹⁵⁸ Shawn Bayern, *A Research Agenda for Organizational Law* (Edward Elgar 2024) 29.

¹⁵⁹ Tech founders frequently refer to the period when OpenAI was a pure non-profit as 'when OpenAI was still open' to highlight the opacity of the contemporary governance regime.

¹⁶⁰ For a more comprehensive review of these relative advantages, see Gad Weiss, 'Aligned Structuring of AI Startups' (2024) SSRN working paper 22 <https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4937578> accessed 28 August 2024.

¹⁶¹ On the tax implications of the choice between a corporation and an LLC, see Victor Fleischer, 'The Rational Exuberance of Structuring Venture Capital Startups' (2003) SSRN working paper <https://papers.ssrn.com/sol3/papers.cfm?abstract_id=432840> accessed 31 August 2024.

¹⁶² Gad Weiss, 'The Venture Corporation' (2023) SSRN working paper 39 <https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4338030> accessed 28 August 2024.

as OpenAI whose innovative governance structures already expose them to substantial uncertainty.

Last but not least, it is worth stressing that there is uncertainty as to whether LLCs can provide in their agreement, that third parties can enforce managers' fiduciary duties, if any. Not only would such rights be obviously ineffective if the agreement remained private, but even if it were made public, one may wonder to what extent courts would be willing to accede to such requests from third parties. This is a major issue: extending directors' liability for pursuing the company's purpose rather than profit without extending their fiduciary duties only diminishes their accountability towards the shareholders, without correlatively increasing their incentives to pursue the interests of humanity. As we next see, other state laws, such as that of Nevada, have interesting features which may, from this perspective, make the enforcement of the company's purpose more effective.

4.1.1.2. Functional Governance Mechanisms under Nevada Law

Nevada corporate law, where xAI incorporated, clearly contrasts with Delaware law. Pursuant to Nevada law, a Nevada benefit corporation (Nevada BC) 'shall have the purpose of creating general public benefit',¹⁶³ defined as 'a material positive impact on society and the environment',¹⁶⁴ which 'is in addition to, and may be a limitation on',¹⁶⁵ traditional corporations' purpose to carry out 'the transaction of any lawful business, or to promote or conduct any legitimate object or purpose'.¹⁶⁶ Second, and more importantly, Nevada law requires directors of the benefit corporation to 'consider the impacts of any action or inaction upon' a long list of stakeholders and interests (shareholders, employees, customers, community and societal factors, the local and global environment, the short-term and long-term interest of the company and the public benefit purpose(s) assigned to it).¹⁶⁷ Furthermore, 'in discharging [its] duties, the board of directors [is] not required to give priority to any particular factor or the interests of any particular person or group [...] over any other factor or the interests of any other person or group.'¹⁶⁸

This long list of interests to be 'considered' is very likely to be toothless, both because effectively pursuing all of these interests at once is arguably impossible and because virtually any decision may always be justified by the pursuit of one interest even if it leads to sacrificing the others.¹⁶⁹ However, 'the benefit corporation [can state] its intention to give priority to a specific public benefit purpose identified in the articles of incorporation of the benefit

¹⁶³ NRS 78B.140(1).

¹⁶⁴ NRS 78B.040.

¹⁶⁵ NRS 78B.140(1).

¹⁶⁶ NRS 78.030(1).

¹⁶⁷ NRS 78B.150(1).

¹⁶⁸ NRS 78B.150(3).

¹⁶⁹ On this argument, see Lucian Bebchuk and Roberto Tallarita, 'The Illusory Promise of Stakeholder Governance' (2020) 106 *Cornell Law Review* 91, 164-165.

corporation',¹⁷⁰ in which case directors need to give priority to this purpose. The company may thus very well consider AI safety and/or the pursuit of humanity's interests as a specific public benefit whose pursuit shall prevail over any other objective.

What makes Nevada law striking is its provision that failure to pursue a specific public benefit may be subject to a 'benefit enforcement proceeding', defined as 'a claim or action for [inter alia] failure of a benefit corporation to pursue the general public benefit purpose of the benefit corporation or any specific public benefit purpose set forth in its articles of incorporation'.¹⁷¹ The great originality of this proceeding is that, besides shareholders and persons holding equity interests in companies of which the Nevada BC is a subsidiary, it may be triggered by 'any other person authorized in the articles of incorporation or bylaws of the benefit corporation to commence and maintain a benefit enforcement proceeding'.¹⁷² These persons may very well be third parties to the company, such as public authorities or members of the civil society, who have an interest in having the company's specific public benefit be enforced, even though they do not have any financial interest in the company.

Nevada law thus provides a great template for implementing effective governance mechanisms incentivising directors to pursue the company's public mission. Unfortunately, however, neither directors nor officers can be held liable for monetary damages for 'failure of the benefit corporation to pursue or create general public benefit or any specific public benefit identified in the articles of incorporation'.¹⁷³ What, then, is the point of a benefit enforcement proceeding? The law does not say anything about its purpose or outcome.¹⁷⁴ The State Bar of Nevada affirms that 'the remedies for such a suit are limited to injunctive rather than monetary relief'.¹⁷⁵ While the impossibility of bringing monetary claims with benefit enforcement proceedings may limit directors' incentives to pursue the company's public benefit rather than profit maximisation, these proceedings still are an interesting tool at third parties' hands to prevent the company from causing excessive harm. As a last remark, the Nevada corporation, just as Delaware's LLC, comes with its lot of legal uncertainty, as the Nevada corporation has not been tested in court as much as Delaware's corporate forms. Nevertheless, the original features of Nevada law offer one additional option to US companies willing to tailor their governance structure to their specific purpose.

4.1.2. Functional Governance Safeguards in the European Union

¹⁷⁰ NRS 78B.150(3).

¹⁷¹ NRS 78B.030.

¹⁷² NRS 78B.190(4).

¹⁷³ For directors, see NRS 78B.150(6); for officers, see NRS 78B.160(3). A similar provision shields the company against liability for monetary damages for failing to create specific public benefits, but does not say anything about its liability for failing to pursue such benefits (as opposed to creating them): NRS 78B.190(3). However, nothing in the law implies that such a liability can be triggered.

¹⁷⁴ All the law says is that the plaintiff can obtain reimbursement for its reasonable expenses if 'the court in a benefit enforcement proceeding finds that a failure to comply with this chapter was without justification': NRS 78B.190(4).

¹⁷⁵ State Bar of Nevada, 'Benefit Corporation Legislation: Safeguards for Businesses Giving Back' <<https://nvbar.org/nvlawyermagazine/february-2013/>> accessed 29 August 2024.

As we already hinted, French, German and Italian laws do not establish a clear link between the company's purpose and the fiduciary duties of its directors. Both are defined separately and independently from each other. For this reason, we start with a discussion of companies' ability to define their purpose in these jurisdictions before covering fiduciary duties separately.

4.1.2.1. Corporate Purpose in the European Union

A comparison of French, German and Italian corporate laws shows that, in the European Union, companies' ability to pursue a nonprofit mission as their primary purpose varies considerably.¹⁷⁶ In France, companies can be structured as '*sociétés à mission*'¹⁷⁷ (equivalent of the US benefit corporation), are free to state the corporate purpose of their choice (as long as it is lawful),¹⁷⁸ can 'specify a *raison d'être*, consisting of the principles that the company has adopted and for which it intends to allocate resources in order to carry out its business activities',¹⁷⁹ and ought to be managed 'taking into account the social and environmental challenges of [their] activity'.¹⁸⁰ These provisions seem to give some room for manoeuvre to companies wishing to pursue their activity for a nonprofit purpose. However, French law also provides that companies are 'constituted in shareholders' common interest' and must be managed in accordance with the company's corporate interest (*intérêt social*).¹⁸¹ This casts serious doubt as to the effectiveness of providing in the company's bylaws that directors ought to pursue an interest superseding that of the shareholders and/or the company itself. Italian corporate law is similarly rigid when it comes to defining the company's purpose: the best companies can do to promote a nonprofit objective is to incorporate as a benefit company (*società benefit*), which can only pursue a nonprofit mission 'in addition to the purpose of sharing profits', not instead.¹⁸² German law, on the other hand, grants more flexibility to shareholders: limited liability companies (*GmbH*)¹⁸³ and stock corporations (*AG*)¹⁸⁴ can both be formed for any legal purpose, including a non-profit one.¹⁸⁵ Non-profit GmbHs (known as *gGmbH*), in particular, enjoy favourable tax treatment.¹⁸⁶

¹⁷⁶ For a comparative analysis of corporate purpose in France, Germany, the UK and the US, see Holger Fleischer, 'Corporate Purpose: A Management Concept and its Implications for Company Law' (2022) 19 *European Company and Financial Law Review* 162, 170.

¹⁷⁷ French Code de commerce, art L210-10.

¹⁷⁸ French Code civil, art 1833.

¹⁷⁹ French Code civil, art 1835.

¹⁸⁰ French Code civil, art 1833.

¹⁸¹ French Code civil, art 1833.

¹⁸² Italian *Legge 28 dicembre 2015, n. 208 (Legge di Stabilità 2016)*, *Gazzetta Ufficiale* No 302, 30 December 2015, §376.

¹⁸³ *GmbHG*, § 1.

¹⁸⁴ *AktG*, § 23(3)(2).

¹⁸⁵ In this sense, see Mathias Habersack, 'Corporate Purpose' in Gregor Bachmann, Stefan Grundmann, Anja Mengel and Kaspar Krolop (eds), *Festschrift für Christine Windbichler zum 70. Geburtstag am 8. Dezember 2020* (De Gruyter 2021) 707.

¹⁸⁶ *Abgabenordnung*, § 51-68.

4.1.2.2. *Fiduciary Duties in the European Union*

As we already discussed, corporate purpose has very limited value insofar as its implementation is not guaranteed by a set of rules effectively incentivising the company's directors to act according to that purpose. Nevertheless, France, Germany and Italy do not establish a clear link between purpose and fiduciary duties, in the sense that the scope of directors' fiduciary duties does not depend on that purpose. This is an oddity that makes purpose rather ineffective.

French, German and Italian corporate laws leave companies limited flexibility to shape directors' fiduciary duties. Most jurisdictions impose mandatory fiduciary duties on the company's directors and managers, and the possibility to adjust the scope of these duties through the company's articles of incorporation is unclear at best. French law, for instance, provides that 'the directors and the chief executive officer are liable [...] to the company or to third parties, either for breaches of the laws or regulations applicable to [...] companies, or for breaches of the articles of association, or for misconduct in their management'.¹⁸⁷ While it imposes on companies' directors and managers a general duty to act in the company's interest,¹⁸⁸ which the company's articles of association may seek to define (directly and indirectly through the definition of the company's corporate purpose), the violation of this duty alone has, to the best of our knowledge, never been found in the past to trigger directors' liability, and may at best serve as a defence against it. Whether the violation of the company's *raison d'être* has any implication in terms of directors' and managers' liability is also doubtful at best.¹⁸⁹

German law provides for companies structured as AGs that 'no breach of duty [towards the company¹⁹⁰] shall be found if a management board member [or supervisory board member¹⁹¹] could reasonably assume, when making a business decision, that they were acting in the best interests of the company on the basis of appropriate information'.¹⁹² There is doubt as to whether the 'company's interest' can encompass abstract objectives (such as AI safety) or the

¹⁸⁷ French Code de commerce, art L223-22 (SARL), L225-251 (SA) and L227-8 (SAS); French Code civil, art 1850 (civil companies).

¹⁸⁸ French Code civil, art 1833.

¹⁸⁹ The impact study of the law creating the concept of *raison d'être* stressed itself that its legal consequences are 'hard to anticipate': French Assemblée Nationale, *Étude d'impact n°1088 sur le projet de loi relatif à la croissance et la transformation des entreprises* (15th Legislature, 2019) 548. For an overall study of the French civil liability regime of corporate officers, see Deen Gibirila, 'Fasc. 1053 : DIRIGEANTS SOCIAUX. – Responsabilité civile' (2019) Lexis360 Jurisclasseur Commercial (as updated by Hélène Azarian on 20 April 2023).

¹⁹⁰ The conditions of liability towards other persons than the company itself, such as the shareholders and creditors, are even more restrictive: Aktiengesetz, § 93 and 116.

¹⁹¹ Aktiengesetz, § 116.

¹⁹² Aktiengesetz, § 93 (Sorgfaltspflicht und Verantwortlichkeit der Vorstandsmitglieder). On the assessment of the German 'business judgement rule', see Bundesgerichtshof (BGH), ARAG/Garmenbeck (1997) II ZR 175/95, BGHZ 135, 244.

interest of third parties (including humanity at large). Regarding GmbHs, the law provides that 'the managing directors must exercise the diligence of a prudent businessman in the affairs of the company',¹⁹³ bringing similar questions as to whether directors' fiduciary duties can be tailored to the company's nonprofit mission.

Lastly, directors of Italian limited liability companies (*SRL*) can be held liable for 'failure to comply with the duties imposed on them by [...] the articles of association for the administration of the company', and can be exonerated by the shareholders from their liability under restrictive conditions.¹⁹⁴ Given the limited flexibility given by Italian law to companies in defining a nonprofit purpose and articulating it with profit, there is significant doubt as to the extent to which directors can be insulated from liability whenever they make their nonprofit mission prevail over profit.

Whereas preventing directors from being sued by the shareholders for violation of default fiduciary duties may avoid deterring them from pursuing the company's purpose, the effective implementation of corporate purpose and directors' incentives to act accordingly requires further extending the scope of directors' fiduciary duties to the benefit of third parties, as we saw in Section 2.2. As it stands, third parties can neither bring derivative claims against the directors nor ask for court injunctions when they depart from the company's purpose under French, German or Italian law. Conversely, directors of French,¹⁹⁵ German¹⁹⁶ and Italian companies¹⁹⁷ can in theory be found liable for having caused damage to third parties as a result of a breach of their fiduciary duties. However, we saw that in these jurisdictions, fiduciary duties are largely defined in isolation from the company's purpose, meaning that an extension of the company's purpose is unlikely to affect much the scope of directors' duties towards third parties. Furthermore, the standard of directors' liability towards third parties is in some jurisdictions much higher than for derivative suits or personal legal action brought by the shareholders against directors. In France, for instance, case law subjects civil liability actions of third parties against directors to a much higher liability threshold than actions carried out by shareholders or the company itself (as part of a derivative suit) by requiring a 'separable fault', defined as 'intentional misconduct of a particularly serious nature that is incompatible with the normal exercise of corporate duties'.¹⁹⁸

Another possibility would be to allow third parties to bring derivative suits on behalf of the company against directors, either to trigger their civil liability towards the company (which

¹⁹³ GmbHG, § 43.

¹⁹⁴ Codice civile, art 2476.

¹⁹⁵ French Code de commerce, art L223-22 (*SARL*), L225-251 (*SA*) and L227-8 (*SAS*); French Code civil, art 1850 (civil companies).

¹⁹⁶ German BGB, § 823. While this provision governs tort liability in general without explicit reference to companies' directors, it can validly serve as a basis for triggering their liability towards third parties: [•].

¹⁹⁷ Italian Codice civile, art 2395.

¹⁹⁸ See eg Cass com, 10 Feb 2009, n° 07-20.445.

would be of limited benefit to third parties, to whom damages would not accrue even indirectly) or to obtain injunctive reliefs from competent courts in case directors violate the company's purpose. Such injunctions could act as a prophylactic measure in case, for instance, the company takes excessive risk with the safety of its products. Third parties may ask the court to enjoin the directors to take any corrective action that the court deems necessary to prevent or terminate the harmful behaviour at hand. To the best of our knowledge, however, the perimeter of derivative suits cannot be tailored in either France, Germany or Italy to entitle specific third parties to bring legal action on behalf of the company against its directors. Together with the difficulties mentioned in the previous section which third parties' personal civil liability actions face, mandatory rules applicable to derivative suits make it virtually impossible, for third parties, to trigger directors' civil liability whenever they violate the company's purpose.

4.2. Structural Governance Mechanisms: Constituency Directors and Dual-Class Shares

As we saw, OpenAI's and Anthropic's governance structures feature independent directors appointed by stakeholders with no financial incentive to increase either company's profitability. Whereas all of OpenAI's directors follow this model, Anthropic chose to have only sixty percent of its board appointed by the LTBT. More generally, a great variety of structural governance mechanisms may be adopted and tailored to the specific needs of the company.

We next elaborate on the model chosen by OpenAI and Anthropic, that of constituency directors, before discussing the various ways in which shareholders' voting rights can be tailored to the company's purpose through so-called dual-class share structures. As all jurisdictions covered in this paper offer a great deal of flexibility to companies' shareholders when it comes to implementing structural governance mechanisms, we cover Delaware law together with French, German and Italian law. We drop the analysis of Nevada law, as its originality is for our purpose confined to purpose and fiduciary duties.

4.2.1. Constituency Directors

Constituency directors are entrusted with protecting a certain group of stakeholders (typically represented by a shareholder with adequate incentives, such as Anthropic's LTBT) and are usually appointed by this constituency. In the case of AI companies, constituency directors could be appointed with the mission of representing humanity at large or ensuring the safety of AI products. They could also be attributed specific powers to, say, publish reports about or veto decisions that they consider excessively risky. Allowing this kind of two-tiered governance structure may not only help AI companies reach a better equilibrium between

profitability and stakeholder governance, but also contribute to board diversity in a way that ultimately benefits the company's financial performance.¹⁹⁹

There are several ways of ensuring constituency directors' commitment to play their role. One of them consists in insulating them from the pressure of other directors and shareholders by ensuring their structural independence, for instance through the allocation of a specific budget and by entrusting a specific person with their nomination and revocation. This is, in essence, what OpenAI and Anthropic did for their respective boards.

French, German and Italian laws all leave sufficient flexibility to implement this kind of mechanism in bylaws, at least for some corporate forms. In Germany, the articles of incorporation can provide that managing directors (*Geschäftsführern*) of GmbHs,²⁰⁰ and up to a third of the supervisory board (*Aufsichtsrat*) in AGs,²⁰¹ can be appointed by specific shareholders. In France, companies incorporated as joint stock companies (*SAS*) are also free to design nomination and revocation mechanisms as they see fit,²⁰² just as Italian SRLs²⁰³ and SpAs.²⁰⁴ Likewise, Delaware law is extremely permissive when it comes to appointing constituency directors, as illustrated by Anthropic's and OpenAI's structures. For instance, the LLC regime offers considerable flexibility when it comes to designing governance rights. In particular, OpenAI Global's operating agreement may have transferred all voting rights and managerial powers to the OpenAI Nonprofit and its board of directors, thereby preventing profit-motivated members of the LLC from influencing its governance.²⁰⁵

While these safeguards limit the risk of shareholder pressure on directors, they do not actively incentivise them to fulfil their mission adequately. This is why, in addition to structural independence guarantees, constituency directors should be liable towards the constituency they represent, typically by owing fiduciary duties to them specifically rather than shareholders in general.²⁰⁶ As we saw, none of the jurisdictions we cover allows companies to provide that constituency directors shall have specific fiduciary duties in their articles of association – or, at least, none of them allows it explicitly.

¹⁹⁹ On the importance for AI companies of having the right persons in the boardroom, see Roberto Tallarita, 'AI Is Testing the Limits of Corporate Governance' (*Harvard Business Review*, 5 Dec 2023) <<https://hbr.org/2023/12/ai-is-testing-the-limits-of-corporate-governance>> accessed 19 August 2024. See also Louis Lehot, Lyman Thai, Patrick D Daugherty and Stephen Moore, 'Navigating Board Turbulence: Lessons From OpenAI' (Foley & Lardner LLP, 5 January 2024) <<https://www.foley.com/insights/publications/2024/01/navigating-board-turbulence-lessons-openai>> accessed 30 August 2024.

²⁰⁰ German GmbHG, § 45.

²⁰¹ AktG, § 101.

²⁰² French Code de commerce, art L227-6 and L227-9.

²⁰³ Italian Codice civile, art 2468(3).

²⁰⁴ Italian Codice civile, art 2351.

²⁰⁵ See esp Delaware Limited Liability Company Act, § 18-302 and § 18-404.

²⁰⁶ Simone M Sepe, 'Intruders in the Boardroom: The Case of Constituency Directors' (2013) 91 *Washington University Law Review* 309.

4.2.2. Dual-Class Shares

Structural governance mechanisms can also take the form of a class of shareholders taking on stronger control rights or even quasi-board member rights. As we have seen with Anthropic, preferred stockholders can be granted veto rights but no profit rights. Granting additional powers to special classes of shareholders limits the problems of fiduciary constraints, as the governance rights of these shareholders are not impacted by fiduciary duties owed to other classes of shareholders. We could imagine, for instance, granting the holders of a special category of shares the right to veto decisions which would affect the interests of humanity at large.

Such structures can freely be implemented by private companies,²⁰⁷ at least so long as the company is not publicly traded. While the European Union is about to adopt a directive requiring Member States to allow listed companies to issue multiple voting rights under certain conditions,²⁰⁸ the possibility for listed issuers to issue preferred shares with specific (veto) rights attached to them, rather than multiple (vanilla) voting rights, is not covered.

Finally, under Delaware corporate law, shareholder voting is required for certain life-and-death decisions, such as adopting or amending bylaws (unless the certificate of incorporation provides otherwise),²⁰⁹ mergers,²¹⁰ or dissolution.²¹¹ However, nothing in the law prevents corporations from requiring special classes of shareholders to vote on additional issues, such as business decisions which are liable to affect the interests of humanity. Section §151 DGCL provides that corporations can issue as many classes of stock as they please, and have the freedom to divorce control rights from profit rights.²¹²

5. Making Corporate Law Fit for AI in the EU

Our study of three European jurisdictions – France, Germany and Italy – showed that the corporate law landscape is likely to prevent rather than encourage AI companies to implement governance mechanisms adapted to their nonprofit concerns. These impediments, together with the fragmentation of Member States’ laws, run against the EU’s ambition, expressed *inter alia* in the Artificial Intelligence (AI) Act adopted on 13 June 2024, ‘to improve the functioning of the internal market by laying down a uniform legal framework in particular for the

²⁰⁷ See n 178, 179 and 180.

²⁰⁸ European Commission, ‘Proposal for a Directive of the European Parliament and of the Council on multiple-vote share structures in companies that seek the admission to trading of their shares on an SME growth market’ COM/2022/761 final.

²⁰⁹ §109 DGCL.

²¹⁰ §251(c) DGCL.

²¹¹ §275(b) DGCL.

²¹² This is what the article says: the corporation can issue as many classes of shares as it wants, ‘which classes or series may have such voting powers, full or limited, or no voting powers, and such designations, preferences and relative, participating, optional or other special rights, and qualifications, limitations or restrictions’.

development, the placing on the market, the putting into service and the use of artificial intelligence systems (AI systems) in the Union [...] while boosting innovation and employment and making the Union a leader in the uptake of trustworthy AI'.²¹³ This brings us to ask how European Union law could intervene to make the European ecosystem more attractive to companies in the AI sector wishing to tailor their governance structure to their specific purpose.

It is worth recalling here that Delaware and Nevada laws are also far from being perfect when it comes to giving AI companies the flexibility to tailor their corporate purpose and their governance structure. The reason why we formulate recommendations destined to the European Union specifically is that there is, at the moment, considerable political momentum in Europe in favour of the creation of a new corporate form that meets the needs of innovative companies. This being said, the recommendations of this section may very well be transposed to the US as well. Nonetheless, history shows that corporate law reforms at the federal level are a particularly thorny subject,²¹⁴ and the political momentum for such reform is not there – at least not yet.²¹⁵

We next discuss what substantive rules should be introduced into European corporate law to allow purpose-led AI companies in Europe to effectively tailor their governance structure. We then see how these recommendations could be implemented concretely, through the creation of new corporate form at the EU level. The two last subsections argue that, while AI companies' activity may have considerable externalities, no strong case can be made, as of yet, that these companies should be subject to mandatory rules requiring them to adopt purpose-led governance structures. However, European authorities could adopt a 'carrot-and-stick' approach whereby purpose-driven entities would have preferential access to EU public funding.

5.1. Increasing the Flexibility of Substantive Rules

Reconciling purpose with profits requires two main legal building blocks: leeway in defining the scope and effects of corporate purpose, and tailor-made fiduciary duties.

The first and most obvious way of increasing companies' ability to tailor their governance to their needs consists in giving companies more leeway in defining the scope and effects of their corporate purpose and its relationship with the pursuit of profits. As we saw in Section 3,

²¹³ Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence and amending Regulations (EC) No 300/2008, (EU) No 167/2013, (EU) No 168/2013, (EU) 2018/858, (EU) 2018/1139 and (EU) 2019/2144 and Directives 2014/90/EU, (EU) 2016/797 and (EU) 2020/1828 (Artificial Intelligence Act) [2024] OJ L217/1, recitals 1 and 2.

²¹⁴ Daniel Crane, *The Dissociation of Incorporation and Regulation in the Progressive Era and the New Deal* in Naomi Lamoreaux and William Novak (eds), *Corporations and American Democracy* (Harvard University Press 2017), 112.

²¹⁵ S 3348 – Accountable Capitalism Act, 115th Congress (2017–2018). (Elizabeth Warren's proposal for the Accountable Capitalism Act, which has never materialized).

there is considerable disparity between Member States when it comes to companies' ability to define the perimeter of their corporate purpose. Furthermore, the legal implications of that purpose, especially in terms of directors' liability, are usually far from certain. To entitle AI companies to tailor their governance structure to their purpose, corporate law should both make it clear that companies enjoy the greatest flexibility in defining that purpose and clarify its legal implications. In turn, corporate law should allow companies to define the perimeter of directors' fiduciary duties, clarify what can trigger these duties, to whom they are owed, who can sue the directors based on their violation (whether directly or derivatively on behalf of the corporation) and under what procedures. While some jurisdictions do not explicitly forbid tailor-made fiduciary duty mechanisms, the impact of these mechanisms remains largely untested and thus risky and unattractive for corporate actors. AI companies incorporated in Europe would benefit a lot from enabling rules that explain clearly to what extent they can customise directors' fiduciary duties through their bylaws. This would require establishing a clear link between corporate purpose and fiduciary duties, which has up until now been lacking in France, Germany and Italy.

Finally, allowing third parties to sue directors for violations of their fiduciary duties would also considerably enhance the efficacy of purpose-led governance structures. Admittedly, companies and public authorities alike may be worried that opening this type of action to the public at large may prove dangerous and overexpose them to frivolous action. One possibility would be to entrust qualified claimants to sue directors on behalf of the public – such as licensed associations, the 'national market surveillance authorities' entrusted by the AI Act to monitor AI markets²¹⁶ or the Commission itself through its AI Office – based on violations of some specific duties, such as their duty to ensure AI safety. While allowing these qualified claimants to trigger directors' civil liability on behalf of the public for a violation of these duties would represent a major conceptual innovation, we believe that its concrete implementation would not represent any major difficulties, as it mainly involves extending a pre-existing mechanism to additional claimants.

5.2. Modifying the European Company Regime or Creating a New Corporate Form

There is at the moment very strong momentum in favour of the creation of a new corporate form at the European level that would be more in line with the needs of innovative companies. In its Political Guidelines for the Next European Commission 2024-2029, for instance, President Ursula von der Leyen announced that she will propose 'a new EU-wide legal status to help innovative companies grow. This will take the form of a so-called 28th regime to allow companies to benefit from a simpler, harmonised set of rules in certain areas'.²¹⁷ The headline-grabbing Draghi report similarly mentions the need to streamline the regulatory landscape of

²¹⁶ AI Act, art 74.

²¹⁷ Ursula von der Leyen, 'Political Guidelines for the Next European Commission 2024-2029' (Strasbourg, 18 July 2024) 7.

corporate law at the EU level with the goal of enhancing the EU's competitiveness.²¹⁸ Furthermore, a petition signed by more than 600 VC funds and 9000 startups proposes the creation of a 'pan-European entity for Europe's startups'.²¹⁹

In line with these initiatives, an effective way of implementing our recommendations would be to create a new corporate form at the European level which would enjoy the greatest flexibility when it comes to defining the perimeter of their corporate purpose as well as that of their directors' fiduciary duties. From a regulatory perspective, this mechanism could follow the logic of the '*Societa Europaea*' ('European Company' or 'SE'). An SE can be created as a result of the merger of two pre-existing companies from two different EU Member States and whose corporate forms are listed in the European Company Regulation (the 'SE Regulation').²²⁰ The SE Regulation regulates the SE on some specific aspects but relies on national corporate laws for most rules applicable to it. Indeed, 'an SE shall be regarded as a public limited-liability company governed by the law of the Member State in which it has its registered office'.²²¹ Furthermore, 'the formation of an SE shall be governed by the law applicable to public limited-liability companies in the Member State in which the SE establishes its registered office'.²²² For instance, French companies that incorporate as an SE are also subject to the rules applicable to the *Société Anonyme*, which are by far more important and numerous than the few rules provided by the SE Regulation.

In the same vein as the SE, public interest companies may be subject by default to all of the rules applicable to a corporate form available in the Member State of incorporation (pre-existing or created for that purpose) and benefit on top of that from the additional features of the public interest company.

This corporate form, however, should not repeat the mistakes of the SE. The SE Regulation claimed to 'permit the creation and management of companies with a European dimension, free from the obstacles arising from the disparity and the limited territorial application of national company law'.²²³ However, the SE has never been particularly successful, and is far from leaving enough legal flexibility to encourage its broad adoption.²²⁴ For instance, a SE may only be constituted as a result of the merger of two companies from two different Member States²²⁵ or upon the transformation of a public limited-liability company which has

²¹⁸ Mario Draghi, 'The Future of European Competitiveness' (September 2024) https://commission.europa.eu/topics/strengthening-european-competitiveness/eu-competitiveness-looking-ahead_en#paragraph_47059.

²¹⁹ *EU Inc.* <https://www.eu-inc.org/> accessed 20 November 2024.

²²⁰ Council Regulation (EC) No 2157/2001 of 8 October 2001 on the Statute for a European company (SE) [2001] OJ L294/1 (SE Regulation), art 2(1) and (2).

²²¹ SE Regulation, art 3(1).

²²² SE Regulation, art 15(1).

²²³ European Company Regulation

²²⁴ In this sense, see Casimiro A Nigro and Alperen A Gözlügöl, 'Building an EU Venture Capital Market: What About Corporate Law?' in *Money Law, Capital, and the Changing Identity of the European Union* (2022) 165, 175.

²²⁵ SE Regulation, art 2(1) and (2).

a subsidiary in another Member State.²²⁶ Furthermore, all SEs are public limited-liability companies²²⁷ (a generic term referring to various corporate forms under Member States' national laws listed in the SE Regulation,²²⁸ such as the *Aktiengesellschaft* in Germany and the *Société Anonyme* in France) which tend to be heavily regulated. In particular, these companies are all subject to the relatively heavy rules of the Company Directive and its rules of transposition.²²⁹

To ensure the effectiveness of our proposals, we suggest creating a 'Simplified European Company' (SEC) which would have the following features. First, the national corporate rules applicable to this corporate form for matters not dealt with by the EU legislative instrument consecrating the SEC (be it a directive or a regulation²³⁰) should be those applicable to the most flexible corporate form available in each Member States (such as the SAS in France, the GmbH in Germany and the S.r.l. in Italy). This is because purpose-led companies need as much flexibility as possible to tailor their governance structures to their needs. Second, no additional requirements to those provided for by national laws should be imposed on these companies to allow them to benefit from the SEC regime. Finally, relying on the SEC regime should give shareholders the greatest flexibility in tailoring their corporate structure to their needs without putting any additional burden on them. In particular, the legislative tool implementing the SEC should provide that they should be free to articulate purpose and profits as they wish, and to define fiduciary duties accordingly, without requiring, for instance, that purpose be put above profit or that directors be liable towards a specific constituency.

Whatever the mechanism chosen, we see no particular reason why this corporate form would be reserved to AI companies. Any company that is serious about pursuing an alternative corporate purpose than profits and willing to implement governance mechanisms accordingly should be able to access it. Nevertheless, this corporate form would primarily be designed with AI companies in mind, given that they are those for which this need expressed itself the most evidently and due to the importance given to the European AI strategy.²³¹

5.3. Mandatory Rules

²²⁶ See Regulation, art 2(4).

²²⁷ SE Regulation, art 15(1).

²²⁸ SE Regulation, Annex 1.

²²⁹ SE Regulation, Annex 1 and Directive (EU) 2017/1132 of the European Parliament and of the Council of 14 June 2017 relating to certain aspects of company law [2017] OJ L 169/46 (Company Law Directive), Annex 1.

²³⁰ While a regulation would ensure maximal harmonisation, a directive may be a more effective way to adapt our proposals to the specificities of Member States' national laws. For instance, the disparity of fiduciary duty regimes in the EU and the need to cope with their peculiarities may justify the enactment of a directive requiring Member States to attain a specific objective (tailoring the perimeter of directors' fiduciary duties to the company's purpose) rather than define the perimeter of fiduciary duties directly. This may make the implementation of our recommendations easier, both technically and politically, while preserving their effectiveness.

²³¹ Commission, 'Artificial Intelligence for Europe' COM (2018) 237 final.

The implementation of governance mechanisms aiming to protect various stakeholders' interests, for instance by prioritising AI safety over profits, may lead companies to internalise the cost of (preventing) externalities linked to their activities. For this reason, companies may be willing to adopt less constraining rules than what would be socially optimal.²³² In turn, one may argue that AI companies should be subject to governance rules encouraging the ethical and responsible development of AI technology,²³³ for instance by requiring them to appoint at least one director in charge of AI safety. These governance mechanisms could also be complemented with penalties imposed on the company's directors for egregious violations of the AI Act. This may be an effective way of incentivising directors to account for stakeholders' interests (such as users of the AI product or humanity at large) without having to rely on the enforcement of fiduciary duties by third parties (a possibility which, as we saw, may prove excessively difficult to implement). A parallel can be made here with the systemic risks caused by the activities of systemically important financial institutions, which these institutions do not fully internalise. This led to various regulatory interventions aiming to mitigate these risks through, inter alia, dedicated governance principles.²³⁴

This paper does not purport to provide any definitive answer in this regard. The opportunity of mandating directors to pursue a greater purpose than profits is uncertain at best.²³⁵ Public benefit corporations and their European counterparts have also appeared relatively recently, with the corporate practice in terms of governance being still in its infancy. Adopting mandatory rules in this area may bridle the spontaneous emergence of good practices through trial and error by purpose-driven companies.²³⁶ Furthermore, any additional set of mandatory rules governing AI may have the counterproductive effect of deterring the incorporation of AI companies in the European Union for unclear benefits. At the same time, given the potential risks posed by AI systems and companies, this option should in our view be seriously considered. Adequate corporate governance rules incentivising companies' officers and directors to seriously consider AI safety may complement the various obligations and safety processes that the AI Act already imposes on companies developing and providing

²³² Oliver Hart, 'An Economist's View of Fiduciary Duty' (1993) 43 *The University of Toronto Law Journal* 299, 310.

²³³ Alan J Dignam, 'Artificial Intelligence: The Very Human Dangers of Dysfunctional Design and Autocratic Corporate Governance' (2019) Queen Mary School of Law Legal Studies Research Paper No. 314/2019.

²³⁴ On governance in particular, see Basel Committee on Banking Supervision, 'Corporate governance principles for banks' (2015) <https://www.bis.org/fsi/fsisummaries/corp_gov_principles.htm> accessed 31 August 2024.

²³⁵ For a recent illustration of this debate, see Paul Davies, 'Shareholder Voice and Corporate Purpose – The Purposelessness of Mandatory Corporate Purpose Statements' in Luca Enriques and Giovanni Strampelli, *Board-Shareholder Dialogue* (Cambridge University Press 2024) and Colin Mayer, 'The Purpose of Corporate Purpose Statements – A Response to "Shareholder Voice and Corporate Purpose: The Purposelessness of Mandatory Corporate Purpose Statements" by Paul A. Davies' in Luca Enriques and Giovanni Strampelli, *Board-Shareholder Dialogue* (Cambridge University Press 2024).

²³⁶ On this argument, see Jens Dammann, 'Publicly Traded Public Benefit Corporations: An Empirical Investigation' (2024) *Stanford Journal of Law, Business & Finance*, Forthcoming <https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4766678> accessed 17 Nov 2024.

‘high-risk AI systems’.²³⁷ Further research is certainly needed before rushing to the conclusion that the benefits of mandatory rules would balance the additional regulatory burden they may impose on European companies. The claim of this paper is merely that purpose-driven European companies should be given the necessary flexibility to tailor their governance to their specific purpose if they want to, as this would have the double benefit of making European law more attractive and, arguably, European companies’ activities less risky.

An intermediate solution would be to sanction not only companies, but also their authorised representatives for violations of the AI Act,²³⁸ and in particular its Chapter V on General Purpose AI Models²³⁹ (which include most frontier AI technologies). Directors’ exposure to such sanctions could, to a large extent, have the same incentive effect as allowing third parties to trigger their civil liability based on their failure to implement adequate safety measures. One caveat is that the efficacy of such a system depends on the AI Act’s own effectiveness in preventing excessive risk. As we have argued, *ex ante* risk assessments can quickly become inaccurate in the case of multi-purpose technologies, particularly when the end users are vulnerable citizens.

5.4. Incentivising Purpose-Led Governance

While OpenAI and Anthropic opted for strong governance mechanisms in favour of their nonprofit missions, such mechanisms may not be as adequate or attractive for startups which are first and foremost concerned with raising sufficient funding.²⁴⁰ Even milder versions, such as the appointment of minority directors entrusted with a nonprofit mission, may prove too constraining for these companies. After all, while AI companies’ founders may have the best of intentions, implementing governance mechanisms to benefit society at the expense of profit amounts to spontaneously internalising a negative externality, which is not an obvious choice.

A possibility would be to create a ‘carrot-and-stick’ corporate form at the EU level whereby AI companies would commit to a package of governance mechanisms while receiving regulatory and financial perks from the EU in exchange. The EU could create a corporate form for ‘public interest companies’ that would come with the following possible features:

- Requirement to pursue a public interest mission alongside profits or superseding profits;
- Possibility to impose on directors fiduciary duties to pursue a specific mission in the general interest (such as AI safety), whose articulation with profits would be tailored to the company’s specific needs; and

²³⁷ AI Act, art 6-49 (see in particular art 14 on human oversight-related requirements).

²³⁸ On the AI Act’s sanctions, see AI Act, art 99-101.

²³⁹ AI Act, art 51-56.

²⁴⁰ On this argument, see Gad Weiss, ‘Aligned Structuring of AI Startups’ (2024) SSRN working paper 22 <https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4937578> accessed 28 August 2024.

- Possibility to elect constituency directors by specific shareholders or even by other stakeholders who would have tailor-made fiduciary duties depending on their role.

Public interest companies adhering to governance features deemed to align them with the EU's AI strategy could receive preferential access to funding by EU institutions, notably as part of the Horizon Europe,²⁴¹ InvestEU²⁴² and Digital Europe²⁴³ funding programmes. These programmes allow companies to obtain funding and financial guarantees for projects in the AI sector. As we saw in the introduction, the reason why OpenAI did not remain a full nonprofit and had to attract profit-motivated investors was precisely that the US government did not provide it with the financing it needed to keep growing without relying on private markets. By combining public funding and guarantees with private funding and governance mechanisms, the EU could boost European AI companies' ability to raise funding while ensuring that they will pursue nonprofit goals, such as AI safety, in line with the EU's objective of supporting innovation 'while ensuring a high level of protection of health, safety, fundamental rights'.²⁴⁴ Funding by public authorities may also come with their involvement in AI companies' governance, with for instance one director appointed by the public funder (be it the EU itself or implementing partners, such as the European Investment Bank). This may mitigate the risk that a powerful shareholder or CEO uses its influence to put pressure on directors to pursue profit maximisation rather than the company's nonprofit mission.

Conclusion

The risks posed by AI have been first and foremost outlined by industry players, and their mitigation through tailored governance undertaken by AI companies themselves. From this perspective, allowing AI companies to place purpose above profits, and to adequately incentivise their directors to follow suit, would both foster market actors' ability to optimise their governance through private ordering and mitigate the externalities of AI technologies. Regulation aiming to foster AI safety need not put an additional regulatory burden on AI companies. Instead, implementing the policy recommendations formulated in this paper may both make EU Member States more attractive for AI companies and further encourage the adoption of good governance practices satisfying the objectives pursued by the European Union in its AI strategy.

²⁴¹ Regulation (EU) 2021/695 of the European Parliament and of the Council of 28 April 2021 establishing Horizon Europe – the Framework Programme for Research and Innovation, laying down its rules for participation and dissemination, and repealing Regulations (EU) No 1290/2013 and (EU) No 1291/2013 [2021] OJ L 170/1.

²⁴² Regulation (EU) 2021/523 of the European Parliament and of the Council of 24 March 2021 establishing the InvestEU Programme and amending Regulation (EU) 2015/1017 [2021] OJ L 107/30.

²⁴³ Regulation (EU) 2021/694 of the European Parliament and of the Council of 29 April 2021 establishing the Digital Europe Programme and repealing Decision (EU) 2015/2240 [2021] OJ L 166/1 (Digital Europe Regulation).

²⁴⁴ AI Act, recital 1.