SciencesPo
CHAIR DIGITAL, GOVERNANCE AND SOVEREIGNTY

Measuring the Openness of AI Foundation Models: Competition and Policy Implications

Thibault Schrepel

Associate Professor of Law at the Vrije Universiteit Amsterdam, Netherlands;

Faculty Affiliate at Stanford University CodeX Center, United States.

Email address: t.schrepel@vu.nl

Jason Potts

Distinguished Professor of Economics at RMIT University

Email address: jason.potts@rmit.edu.au

January 2025

Table of Content

Executive Summary

This paper provides the first comprehensive evaluation of AI foundation model licenses as drivers of innovation commons.

We introduce our analysis by outlining how AI licenses regulate access privileges to the fundamental inputs of AI innovation commons.

We show that AI licenses operate as a bottleneck, as their level of openness directly influences the flow of knowledge and information into the commons.

We then introduce a new methodology for evaluating the openness of AI foundation models. Our methodology extends beyond purely technical considerations to more accurately reflect AI licenses' contribution to innovation commons.

We proceed to apply it to today's most prominent models—including OpenAI's GPT-4, Meta's Llama 3, Google's Gemini, Mistral's 8x7B, and MidJourney's V6—and find significant differences from existing AI openness rankings.

We conclude by proposing concrete policy recommendations for regulatory and competition agencies interested in fostering AI commons based on our findings.

Introduction

Open source technology has become the backbone of modern digital infrastructures. Major web servers such as Apache HTTP Server, Nginx, and Lighttpd, as well as fundamental database systems like MySQL and PostgreSQL, all run on Linux Operating System. Open source libraries and frameworks, such as Django for Python, Laravel for PHP, and Node.js for JavaScript, are used to create dynamic and responsive websites. WordPress, which powers more than 40% of all websites, is also an open-source platform, while popular web browsers like Mozilla Firefox and Chromium are built on open-source projects. Additionally, the Android operating system incorporates open-source code, and Git is central to modern software development. These examples show that open source technology plays a vital role in stimulating innovation commons — i.e., a collaborative environment where knowledge, resources, and tools are openly shared among a network to foster collective innovation and problem-solving (**1.**).

Given that AI foundation models could become a key infrastructure for tomorrow's economy by powering generative AI ("GenAI") applications, one might want to measure the openness of these models. This is because open foundation models foster dynamic competition by contributing to the innovation commons in the long term, while they can be scrutinized, forked and not easily turned against ecosystem players in the short term. In a sense, open AI foundation models inherently foster innovation while addressing most antitrust concerns. The distinction between open and closed AI foundation models should then be taken seriously.

Current methodologies used to measure the openness of online systems and foundation models often fail to consider legal, economic, and social dynamics. These methodologies focus solely on technical aspects, which overlook important elements. In this paper, we introduce a new, comprehensive methodology for measuring the openness of AI foundation models (**2.**). We apply it to the most common AI foundation models, and we derive concrete policy implications based on our findings (**3.**).

1. The Promises of Innovation Commons in AI

AI foundation models are novel and curious economic objects. They are produced across economic institutions, including quasi-public research organizations and for-profit firms, and also, the commons. However, most public policy consideration (whether about safety, access, or concern with monopoly power) focuses on the public or private ownership and governance models, and mostly overlooks the role of the commons. This pattern of oversight is a hardy perennial of previous technology epochs too.[1]

This is a problem for two reasons. First, the development and use of AI is having a significant impact on the structure of competition, something that is accelerating with models that input and output natural language.[2] And second, generative AI makes extensive use of the commons as a crucial part of its innovation ecosystem, and so the way in which this technology gets into the commons will shape dynamic competition and ongoing value creation.

To address this shortcoming, this first section proceeds from the most general to the most specific. It first provides some simple economic models of AI foundations to highlight their economic and competitive dynamics (**1.1.**). It then narrows down to discuss the role of innovation commons in driving AI developments (**1.2**.), and examines how AI licenses can shape the commons **(1.3.**).

1.1. Simple Economics of AI Foundation Models

Research into artificial intelligence began in the 1950s, but recently the frontiers of machine learning have been significantly advanced by a series of practical breakthroughs in architecture and training that led to the deep learning

---

[1] See e.g., Eric von Hippel, *Democratizing Innovation* (MIT Press 2006); Eric von Hippel, *Free Innovation* (MIT Press 2017).

[2] Nicolas Petit, 'Antitrust and Artificial Intelligence: A Research Agenda' (2017) 8(6) *Journal of European Competition Law & Practice* 361-362; Jason Potts, Andrew Torrance, Dietmar Harhoff, and Eric von Hippel, 'Profiting from Data Commons: Theory, Evidence, and Strategy Implications' (2023) *9(1) Strategy Science* 1-17.

revolution.[3] The so-called *foundation models*[4] includes the class of Large Language Models, or LLMs, such as OpenAI's ChatGPT. GPT stands for Generative Pretrained Transformer, which is the novel and surprisingly successful training architecture developed by OpenAI researchers. But GPTs are also what economists call General Purpose Technologies.[5] In the short time since these breakthroughs were made, an extremely large amount of liquid capital (including public research funding, venture finance, stock market raises, debt finance, etc.) has flowed into existing frontier technology companies and new startups around the world to drive innovation in this technology.[6] This paper seeks to contribute to public policy governance of this epochal new industry and general purpose technology in support of its competitive development.

Consider the simple economics of AI innovation.[7] Fundamentally, AI is a type of automation that massively lowers the cost associated with predictive capabilities—foreseeing the subsequent actions of a skilled human based on given inputs. The predictive function is central to both language models, which process sequences of textual inputs, and autonomous vehicles, which interpret arrays of visual data. As such, AI emerges as a novel form of capital with extensive and versatile applications across various sectors. Compared to traditional modes of production and creation, AI is characterized by several distinct advantages, including reduced relative costs, enhanced speed, and considerable modularity that facilitates the integration of AI into both existing and novel processes, thereby extending its utility.

The economic implications of foundation models and generative AI hinge on their operational dynamics—whether they function as complements, thereby augmenting the productivity of particular resources, or as substitutes, potentially displacing existing economic rents.[8] The resultant market dynamics influence the valuation of resources, leading to fluctuations in their market prices.

[3] Yann LeCun, Yoshua Bengio, and Geoffrey Hinton, 'Deep Learning' (2015) 521(7553) *Nature* 436-444.
[4] Rishi Bommasani and others, 'On the Opportunities and Risks of Foundation Models' (2021) *arXiv preprint* arXiv:2108.07258v3.
[5] Tyna Eloundou and others, 'GPTs are GPTs: An Early Look at the Labor Market Impact Potential of Large Language Models' (2021) *arXiv preprint* arXiv:2303.10130v5.
[6] Agrawal, A., Gans, J., Goldfarb, A. (2023) 'Artificial intelligence adoption and system-wide change.' *Journal of Economics & Management Strategy.*
[7] Agrawal, A., Gans, J., Goldfarb, A., *Prediction Machines: The Simple Economics of Artificial Intelligence* (Harvard Business Press 2022).
[8] Aghion P, Jones BF and Jones CI, *Artificial Intelligence and Economic Growth* (NBER Working Paper No w23928, 2017).

Consequently, these innovation dynamics of competition entail disruptive and redistributive impacts on the owners of these resources, affecting the sustainability of firms and the returns to specific job categories.[9] In turn, foundation models and generative AI exemplify Schumpeterian innovation as they trigger 'creative destruction' across global industries that then causes the reallocation and revaluation of capital and labor as these shifts are uncovered, disclosed, and implemented.

Moving to AI institutional dimensions, a crucial aspect of the economic institutions governing the production and use of foundation models and generative AI is their reliance on the commons, as the form of institution that governs critical resources for innovation. This reliance stems from the technology's nature (software, mathematics), its production methods (extensive training sets), and the manner in which value is discovered (through distributed users).

A significant portion of the code used in developing generative AI is shared as open-source software within the software commons. Similarly, the tacit expert knowledge and insights regarding the design and deployment of foundation models circulate within what effectively are knowledge-sharing commons, often regulated by communities of practice.[10] Furthermore, much of the extensive training data used originates from the commons and is accessed through the commons. While special capabilities and knowledge may be safeguarded through contracts and intellectual property within private firms, generative AI is particularly distinguished as an economic good by its production, adoption, diffusion, and innovation predominantly occurring within the commons. This unique attribute is notable even among rapidly advancing and highly valuable frontier technologies. It follows that the economic theory of the commons, as developed by Elinor Ostrom and colleagues, along with the specialized theory of innovation commons, elucidates why such an arrangement can be designed or naturally evolve as an efficient institutional outcome.

Foundation models are also notably complex both economically and legally, posing significant challenges for public policy. These models are exceptionally capital-intensive due to substantial fixed costs associated with

[9] Mollick E and Euchner J, 'The Transformative Potential of Generative AI' (2023) 66(4) *Research-Technology Management*.
[10] Hugging Face has recorded over 1,2 billion downloads of AI foundation models in 2 years, *see* https://perma.cc/9VNZ-J4C5. The platform is now recording nearly 2 million downloads a day.

computing resources and the benefits of increasing scale. Foundation models exhibit characteristics of *labor* (in their AI capabilities which are fluid), *capital* (requiring significant investments for development and operation), and *technology* (comprising code that can be replicated). The ambiguity surrounding economic classification of foundation models (are they labor, capital or technology?) complicates the applicability of existing legal frameworks. Unlike conventional databases or software tools such as search engines, foundation models are technically embeddings. In machine learning, an embedding is a low-dimensional mapping of a discrete variable (such as a word) onto a continuous variable (such as a numerical vector).[11] AI does not represent 'artificial' intelligence as a distinct form of intelligence, akin to traditional machines, but rather facilitates the creation of 'pooled intelligence' through these embeddings. As such, AI foundation models function as mathematical embeddings. They operate as knowledge pools and belong to a class of commons. They utilize resources derived from the commons (training data) and generate value through interaction, such as training or prompting. The value of these models is correlated with the quantity and quality of input data and user interactions, increasing in value with more usage.

These curious economic, legal and mathematical properties of AI foundation models make them challenging objects for competition policy. On one hand, they have properties of market failure, including non-rivalry (the capital good is a digital matrix) and quasi-non-excludability (in training data, and in tacit skills to build and use effectively). They are difficult to protect with intellectual property or capital embodiment, yet they benefit from increasing returns.[12] They provide the technical infrastructure for building applications (the "GenAI" layer), which encourages developers to cluster around one (or, at most, several) foundation models to benefit from network effects and compatibility (i.e., ecosystem dynamic).[13] AI foundation models are also expensive to develop and operate due to massive computing costs. Venture funding typically requires a return that depends on business models that can exploit some kind of exclusivity

[11] Modern LLMs, in fact, are *embeddings of embeddings*, because while the model is an embedding (in code) into a large parameter set of a vast corpus of words, or any cultural product, those words are also an embedding (in language, in human minds) of words of a vast corpus of individually and socially processed sense impressions and experiences. See Potts, J., 'Embeddings' *Cultural Science Journal*, 14(1).
[12] Schrepel T and Pentland AS, 'Competition Between AI Foundation Models: Dynamics and Policy Recommendations' (2024) *Industrial and Corporate Change*
[13] Schrepel T, 'Toward a Working Theory of Ecosystems in Antitrust Law: The Role of Complexity Science' (2024) *Network Law Review* https://perma.cc/72E2-CWZR.

in access and use. Strong economic forces push towards monopoly or highly imperfect competition (where the price per unit of use is much higher than marginal cost due to very high fixed R&D costs).[14] On the other hand, much of the production and innovation here has occurred in the commons, and will likely continue to. So we need to understand how the innovation commons works.

1.2. Innovation Commons in Generative AI

The economic theory of the commons was developed by Elinor Ostrom and colleagues over many decades and hundreds of projects, forming a vast scientific literature, with a predominant focus on natural resource commons (e.g., forests, fisheries, watersheds).[15] The core idea is often expressed as a set of design principles on private-order rules to govern a community of resource creators and users.[16] These particular institutional rules, developed and enforced by the community itself, work because they make better use of local knowledge and understanding of conditions in use, and can deal with uncertainty and tacit knowledge. Commons can have lower transaction and governance costs and, under certain conditions, can shape incentives to more effectively and efficiently create and use a valuable resource than alternative institutional arrangements.[17] This argument applies not only to the governance of resources for consumption and production, but also, critically, to resources for innovation, often called 'knowledge commons'.[18]

An important class of knowledge commons are *innovation commons*.[19] The basic question with innovation commons, as with natural resource commons, concerns the economic logic of why a particular resource (in this case a new technology, as foundation models) is most efficiently created and used in the

[14] Shapiro, K. 'Competition and innovation: Did Arrow hit the bull's eye?' *The Rate and Direction of Innovation Revisited*. (University of Chicago Press 2012).
[15] Ostrom E, *Governing the Commons: The Evolution of Institutions for Collective Action* (Cambridge University Press 1990).
[16] Ostrom E, 'Beyond Markets and States: Polycentric Governance of Complex Economic Systems' (2010) 100(3) *American Economic Review* 641.
[17] Ostrom E, *Governing the Commons: The Evolution of Institutions for Collective Action* (Cambridge University Press 1990).
[18] Frischmann B, Madison M and Strandburg K, *Governing Knowledge Commons* (Oxford University Press 2014).
[19] Allen D and Potts J, 'How the Innovation Commons Contribute to Discovering and Developing New Technologies' (2016) 10(2) *International Journal of the Commons* 1035; Potts J, 'Governing the Innovation Commons' (2018) 14(6) *Journal of Institutional Economics* 1025; Potts J, *Innovation Commons: The Origin of Economic Growth* (Oxford University Press 2019).

commons as compared to alternative forms of institutional governance, such as private property (e.g, proprietary software in a firm) or as a regulated or nationalized public utility. The answer lies in knowledge sharing.

In the theory of innovation commons the key resource that shapes innovation is not shared equipment or capital goods (e.g., machines, compute infrastructure) but information and knowledge to reveal the entrepreneurial opportunity. Because information and knowledge are naturally distributed (e.g., experiments, prices, user demand, sources of supply of key resources, understanding of regulations and other constraints, etc.), innovation is unlikely to be carried out by a single firm, however large. Instead, innovating firms rely on innovation commons to play a central role in pooling information and knowledge.[20] Innovation commons pool distributed resources in order to help alleviate uncertainty around nascent technology and allow for "user innovations"[21]. These user innovations remain in the commons where costs of protection exceed the expected benefits protection could bring, and become commercial innovations where the costs are lower.

This view of innovation commons contrasts sharply with the standard economics of innovation, which focuses on how market structure affects firms' incentives and ability to innovate. The canonical Arrow (1962) model of market failure in fixed costs of R&D due to non-rivalry and uncertainty predicts that competition should decrease innovation.[22] Modern versions show how innovation shapes competition at the level of capabilities, as for instance Aghion et al. (2005) describe an inverted U-shape model in which "competition discourages laggard firms from innovating but encourages neck-and-neck firms to innovate."[23] Yet when applied to foundation models, the fundamental resources that need to be created and deployed are somewhat different, including tacit understanding of how to make a model architecture work, code libraries, sources of graphics chips and other specialized inputs, cleaned training sets and their setup, understanding of legal and regulatory barriers, how to harness communities of early users,

[20] Innovation commons provide firms with a mechanism for group cooperation that turns a market failure problem (the standard definition of the innovation problem) into a collective action problem.
[21] Von Hippel E, *Free Innovation* (MIT Press 2016).
[22] Arrow KJ, 'Economic Welfare and the Allocation of Resources for Invention' in RR Nelson (ed), *The Rate and Direction of Inventive Activity: Economic and Social Factors* (Princeton University Press 1962) 609.
[23] Aghion P, Bloom N, Blundell R, Griffith R and Howitt P, 'Competition and Innovation: An Inverted-U Relationship' (2005) 120(2) *Quarterly Journal of Economics* 701.

prompt libraries and discoveries of uses, and so on. These resources are inputs to innovation, yet are not always efficiently or effectively built within firms, due to the benefits of drawing on a broad set of inputs and the high relative cost of using specialized resources. Innovation commons thus play a crucial role in fostering innovation in generative AI. Our empirical thesis is that the robustness of innovation commons in AI can be observed where organizations and the broader community of users and stakeholders formally intersect in the licenses that govern access and use of AI foundation models.

1.3. The Role of Generative AI Licenses in Shaping Innovation Commons

AI foundation model licenses, comprising terms of use and related documentation, form the constitutional layer of innovation in generative AI. There are two competing explanations for this: one based on the strategic innovation commons model, the other on the standard IO-type economic model of competition. We contend that the strategic innovation commons model does a better job of capturing the dynamics of innovation in generative AI.[24]

Innovation commons theory suggests various strategic motivations for generative AI companies to engage with the commons as a *mode of competition*. Recognizing and delineating these motivations is crucial for aligning competition policy — and broader public policies such as taxation, industry regulation, data rights, and immigration — with the support of innovation and social benefits that comes with this new class of technology. Consider several specific ways in which the licenses of AI foundation models shape competition by impacting innovation commons.

First, licenses can be used to control access to applications, thereby controlling market entry. Licenses are a direct rent, and a way to gate or control competition through their issuance. Additionally, the terms of these licenses influence subsequent related markets, such as product embeddings, thereby also governing the development of business ecosystems, typically by restricting their growth or creating bottlenecks.[25] A critical aspect of using licenses to manage rents and ecosystems relates to the discovery of value and the property rights

---

[24] Potts J, 'Sources of Innovation in Generative AI' (2023) *Network Law Review* https://perma.cc/BD59-4WNW.

[25] Jacobides M and Tae C, 'Kingpins, Bottlenecks, and Value Dynamics Along a Sector' (2015) 26(3) *Organization Science* 889; Jacobides M, Brusoni S and Candelon F, 'The Evolutionary Dynamics of the Artificial Intelligence Ecosystem' (2021) 6(4) *Strategy Science* 412.

that accrue from opportunities identified by third parties or users. The implication of a right, akin to an option, within the license aligns with *real options theory*.[26] Effectively, the license determines where the financial value of these strategic real options is capitalized, whether in the licensor or the licensee.

Second, licenses can be used to protect copying of the model, i.e., the parameter embedding weights, which are very costly to produce (as a consequence of enormous compute cycles) but very easy to copy – the model is just a very large matrix. The license is drawing on copyright protection, but can specify tort consequences of copying to enable discrimination about particular groups copying versus others or other dimensions of merit. The license enables the model to enable what would otherwise be achieved with price discrimination, which is economically efficient, *in the absence of prices*.[27]

Third, the license enhances the liquidity of the underlying corporate asset by clearly defining the exact boundaries of rights associated with the model. The inclusion of precise and detailed descriptions (i.e., thick-description of property rights) in the licensing agreement reduces uncertainties regarding counterparty claims or ambiguities about joint ownership. This clarity facilitates the estimation of future claims and risks, thereby improving the ability to value capital and enhancing capital liquidity. In equilibrium, this will lower the cost of capital for the firm, improving its competitive position by enabling larger capital values to be effectively utilized for financing new rounds of R&D or the acquisition of complementary strategic assets, such as computing resources. Consequently, restrictive use licenses can strengthen an incumbent's competitive position.

Fourth, the license that functions to place resources specifically in the commons can work as a way to sharpen a firm's own competitive advantage by strategically placing resources that if they were proprietary in another firm would amount to a competitive threat. A firm can eliminate or weaken other firms competitive advantage, and business rents, by placing that resource in the commons to push the high-value parts of the industrial ecology to the strategic resources they control.[28] This strategy to place resources specifically in the

[26] Baldwin C, 'Optimal Sequential Investment When Capital Is Not Readily Reversible' (1982) 37 *Journal of Finance* 763; Abel AB, Dixit A, Eberly JC and Pindyck RS, 'Options, the Value of Capital, and Investment' (1996) 111 *Quarterly Journal of Economics* 753.
[27] Bergemann D, Brooks B and Morris S, 'The Limits of Price Discrimination' (2015) 105(3) *American Economic Review* 921.
[28] Jacobides M and Lianos I, 'Ecosystems and Competition Law in Theory and Practice' (2021) 30(5) *Industrial and Corporate Change* 1199; Jacobides M and Tae C, 'Kingpins,

commons to sharpen competitive advantage is usually called 'commoditize the complements'. This is the standard way in which open source software is used as a competitive strategy by firms.[29] A more cooperative approach to placing resources in the commons for profit maximization involves strategically reducing costs for downstream firms or users to incentivize the development of complementary assets or facilitate market discovery.[30] This commoditization strategy is particularly viable when the company placing the model in the commons can still monetize it within its existing ecosystem by providing users with attractive, new features.

In the standard model of competition, in contrast, the strategic competitive purpose of AI foundation model licenses is to control knowledge spillovers through contracting to secure property rights in the use of the product. The implicit economic logic is that: (a) the stronger the protection afforded or set out in the license the stronger the producers' property rights in the product, *it follows that* (b) the more controlled are any spillovers from use of that product by users (consumers, clients, third parties, etc), *it follows that* (c) the stronger the *ex ante* and *ex post* incentive to private capital investment in original and subsequent development of the product, *it follows that* the stronger are the private incentives to investment, (d) the more robust will be investment-driven competition. This applies to both Marshallian market entry-driven contestability and Schumpeterian technological innovation-driven market competition.

While *prima facie* reasonable in some industrial contexts, historical evidence and economic theory both find that licensing to limit spillovers can also have a negative effect on innovation and harm competition.[31] This is because incentives to invest involve a range of other factors too, specifically other inputs to innovation, many of which come from or are efficiently and best provided by the commons. The naive model, in which user license benefits innovation by protecting proprietary technology and capital investments from spillover, may not hold in practice where a substantial amount of innovation, and therefore dynamic

Bottlenecks, and Value Dynamics Along a Sector' (2015) 26(3) *Organization Science* 889.

[29] Schrepel T and Wolf T, 'Open-Source AI Scaling Laws' (2024) *Scaling Theory* https://perma.cc/63KJ-JNKE.

[30] Gambardella A and von Hippel E, 'Open Sourcing as a Profit-Maximizing Strategy for Downstream Firms' (2019) 4(1) *Strategy Science* 41.

[31] Boldrin M and Levine D, *Against Intellectual Monopoly* (Cambridge University Press 2008); Potts J, *Innovation Commons: The Origin of Economic Growth* (Oxford University Press 2019).

competition, issues from the innovation commons. The strategic innovation commons model thus offers a better explanation of the key mechanisms of dynamic competition in generative AI. For this reason, we ground our systematic content analysis (**2.**) on the basis of innovation commons theory rather than standard IO economics. We contend that our approach better captures the impact of AI foundation model licenses on innovation.

2. A Systematic Analysis of AI Foundation Models Licenses

We propose a holistic approach to document AI foundation model openness that considers technical, economic, legal, and social constraints (**2.1.**). Mindful of the need to make our methodology easily accessible, we apply it to the most prevalent AI foundation models currently in use (**2.2.**). We draw conclusions about which models are likely to foster innovation commons and rank them accordingly (**2.3.**).

2.1. Ranking Methodology

AI foundation model licenses regulate access privileges to the fundamental inputs of AI innovation commons. Operating as a bottleneck, their level of openness dictates the flow of knowledge and information into the commons, thus underscoring our emphasis on openness.

With this article, we propose a new methodology for ranking AI foundation model licenses openness by the extent to which they support innovation commons. Previous attempts at measuring their openness focused on technical aspects, following the definition of "open source" provided by the Open Source Initiative.[32] Examples include Liesenfeld, Lopez, & Dingemanse, and others which provide valuable insights on the technical aspects underpinning AI foundation models licenses.[33]

---

[32] Open Source Initiative, https://perma.cc/6MDS-DAZC. For a discussion of the Open Source Initiative's approach, *see* Benhamou Y, 'Open Source AI – Definition and Selected Legal Challenges' (2024) *Kluwer Copyright Blog* https://perma.cc/CPC7-6UMY.

[33] Liesenfeld A, Lopez A and Dingemanse M, 'Opening Up ChatGPT: Tracking Openness, Transparency, and Accountability in Instruction-Tuned Text Generators' (July 2023) in *Proceedings of the 5th International Conference on Conversational User Interfaces* 1 https://perma.cc/4YNT-UL5B; Basdevant A, François C, Storchan V, Bankston K, Bdeir A, Behlendorf B, Debbah M, Kapoor S, LeCun Y, Surman M, King-Turvey H, Lambert N, Maffulli S, Marda N, Shivkumar G and Tunney J, 'Towards a Framework for Openness in Foundation Models: Proceedings from the Columbia Convening on Openness in Artificial Intelligence' (2024) arXiv preprint arXiv:2405.15802; Alliance DPG, "Exploring a Gradient Approach to the Openness of AI System Components - Digital Public Goods Alliance" (Digital Public Goods Alliance - Promoting digital public goods to create a more equitable world, October 27, 2023) https://perma.cc/Q2VL-HGH3; Bommasani R, Klyman K, Kapoor S, Longpre S, Xiong B, Maslej N and Liang P, 'The Foundation Model Transparency Index v1.1: May 2024' (2024) arXiv preprint arXiv:2407.12929; Bommasani R, Kapoor S, Klyman K, Longpre S, Ramaswami A, Zhang D, Schaake M, Ho DE, Narayanan A and Liang P, 'Considerations for Governing Open Foundation Models' (December 2023) HAI Policy & Society Issue Brief (Stanford University, Princeton University, and RegLab); Open Source Initiative,

The European AI Act embraces that technical methodology to define open-source foundation models. Article 53(2) calls them open source when they "are released under a free and open-source license that allows for the access, usage, modification, and distribution of the model, and whose parameters, including the weights, the information on the model architecture, and the information on model usage, are made publicly available."[34] Recital 102 of the AI Act is slightly more detailed. It calls AI models "free and open source" when users can "run, copy, distribute, study, change and improve software and data, including models under the condition that the original provider of the model is credited, the identical or comparable terms of distribution are respected."[35]

This technical approach misses a view of the economics of the commons.[36] AI foundation model licenses are all between a *producer* of a generative AI model (including OpenAI, Anthropic, Midjourney, Meta, etc.) contracting with the *users* of that model (consumers, citizens, clients, other organizations, including firms, non-profits, trusts, governments, etc.), in which the producer writes the contract and, under most circumstances, the user either accepts or rejects as a whole. The purpose of the license is to cover the legitimate uses of the model and to specify the nature of the economic goods that are jointly produced by the user and the model, and in particular, the economic rights that attach to those products. An AI license, therefore, is a device for establishing property rights in the context of co-production.[37] The strong form of the assertion is that: (1) the model and the co-produced output are nonrivalrous goods; (2) efficient and effective use of the model draws extensively on local knowledge; (3)

'The Open Source AI Definition – 1.0' https://perma.cc/F3GE-WKF2; D.G. Widder, M. Whittaker, and S.M. West, 'Why "open" AI systems are actually closed, and why this matters' (2024) 635 Nature 827.

[34] Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence (Artificial Intelligence Act) [2024] OJ L 1689/1, art 53(2) https://perma.cc/95BB-JK6P

[35] Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence (Artificial Intelligence Act) [2024] OJ L 1689/1, Recital 102 https://perma.cc/95BB-JK6P

[36] Hess C and Ostrom E, 'Introduction: An Overview of the Knowledge Commons' (2007); Frischmann B, Madison M and Strandburg K (eds), *Governing Knowledge Commons* (Oxford University Press 2014); Potts J, 'Governing the Innovation Commons' (2018) 14(6) *Journal of Institutional Economics* 1025.

[37] Demsetz H, 'Toward a Theory of Property Rights' (1967) 57(2) *American Economic Review* 347; North D, *Structure and Change in Economic History* (Cambridge University Press 1981); Hodgson G, 'Much of the "Economics of Property Rights" Devalues Property and Legal Rights' (2015) 11(4) *Journal of Institutional Economics* 683.

benefits from discovery and use of applications will often dominate the costs of monopoly protection of investments.[38] The competition thesis underpinning our index measure is that the more a license works to create economic rights in the innovation commons, the more that license supports and facilitates dynamic competition.

To operationalize this theoretical understanding of how innovation and competition work in practice in foundation models and generative AI into a measure, we have identified a range of 18 variables that we can observe in AI licenses that both individually and collectively contribute to support the innovation commons. These 18 variables are all measures of openness, but unlike purely technical definitions of openness that are predominant in much contemporary analysis of AI, our approach is based on institutional effectiveness of the licenses, i.e., we focus on the specific problems they actually address and provide restitution for. In our view, there are three major economic problems that need to be solved in order for an AI innovation commons to be effective, and thereby to support dynamic competition. It must solve: (1) the *knowledge problem*; (2) the *implicit contracting problem*; and (3) the *collective action problem*.

A strong and effective innovation commons needs to solve all three problems. To some extent, solutions within each problem can substitute for each other, which is the logic of these clusters, and which also enables inference from our estimates. With this in mind, we construct an index X_i (X_i = measure of license i = f(knowledge problem + implicit contracting problem + collective action problem) for each of the 11 license types $\boldsymbol{i}$ ($\boldsymbol{i}$ = distinct generative AI producer entities) that *estimates the openness of the model based on its contribution to innovation commons* and resistance to monopolistic control.

2.1.1. Knowledge problem

In the standard economics of innovation, as originally formulated by Kenneth Arrow, perfect competition is the enemy of innovation.[39] Stated as a social welfare argument and policy program, monopoly (e.g., intellectual property rights) is a necessary price to create rents to incentivize innovation. The trade-off between social costs of imperfect competition and social benefits of innovation is a major

[38] Von Hippel E, *Democratizing Innovation* (MIT Press 2006).
[39] Arrow KJ, 'Economic Welfare and the Allocation of Resources for Invention' in RR Nelson (ed), *The Rate and Direction of Inventive Activity: Economic and Social Factors* (Princeton University Press 1962) 609.

focus of modern innovation economics. But in the dynamic competition approach (based on complexity, evolutionary and institutional economics),[40] innovation is primarily a ‘knowledge problem’ of information discovery and coordination.

Institutions solve knowledge problems by creating incentives to aggregate distributed information under uncertainty in different ways.[41] Coordinating distributed, uncertain information via the price mechanism often faces challenges due to absent property rights, within firms due to scalability issues, and by governments due to misaligned incentives and rent-seeking. However, a commons approach can effectively overcome these obstacles in the production of innovation.

The knowledge problem for AI licenses is to encourage the pooling of any and all information, data, and knowledge useful for innovation. This goes beyond just model weights and parameters to include broader information and rules for effective model usage, as well as an understanding of trade-offs, risks, liabilities (technical, computational, behavioral, legal, etc.), performance characteristics and other costs or constraints. The crucial questions are the breadth and depth of information the license encourages to pool and the strength of these incentives — or, practically speaking, how restrictive the disincentives are. To answer these questions, we document the six following variables.

[40] Sidak J and Teece D, ‘Dynamic Competition in Antitrust Law’ (2009) 5(4) *Journal of Competition Law and Economics* 581.
[41] Potts J, *Innovation Commons* (Oxford University Press 2019).

Variables:

- Knowledge accessibility: The extent to which the license facilitates access to the underlying knowledge, including training data, code, weights, and model architecture. Includes the ease of access for different user groups, such as researchers, developers, and the public.
- Documentation and support: Quality, comprehensiveness, and availability of documentation and support provided with the AI model. Includes tutorials, forums, FAQs, and other resources that can help users understand and utilize the technology effectively.
- Transparency: Degree of openness regarding the AI model's development process, including the methodologies, data sources, and algorithms used.
- Collaboration platforms: Support for or provision of platforms that facilitate collaboration, knowledge sharing, and problem-solving among users. Includes wikis, code repositories, forums, and other collaborative tools.
- Engagement and feedback: Provisions and mechanisms encouraging or facilitating community engagement and feedback to continuously refine the model and its use, report issues, and contribute ideas.
- Language and localization: Support for multiple languages — including the technical documentation *about* the model, and the languages supported *by* the model — and localization efforts (i.e., data training from diverse cultural backgrounds).

2.1.2. Implicit contracting problem

The implicit contracting problem, as researched by Williamson and others, is the problem of opportunism in preventing one party to a jointly contributed product (in this case, the innovation) from trying to renegotiate or capture the value of its contribution *ex post* if it turns out to be critical or exploitable in some way.[42]

The license must provide safeguards against the consequences of others reneging on an agreement to share and pool resources in the commons, neglecting or appropriating the contributions of others. To evaluate a license's effectiveness in mitigating opportunism—thereby safeguarding open cooperation—we assess the assurance it provides that contributions will remain

[42] Williamson O, 'The Economics of Organization: The Transaction Cost Approach' (1981) 87(3) *American Journal of Sociology* 548; Hart O and Moore J, 'Incomplete Contracts and Renegotiation' (1988) *Econometrica* 755.

protected from subsequent opportunistic behaviors by others. Our analysis is supported by six key variables.

Variables:

- Contribution policies: Clauses regarding contributions to the model, including *how* contributions are vetted and integrated. Includes code, data, algorithms, and other intellectual property.
- Credit and revenue sharing: Clear articulation of rights and obligations of all parties involved, including any requirements for attribution or crediting the original creators, but also what contributors can expect in return for their contributions (e.g., revenues or compensation).
- Specific anti-opportunism clauses: Provisions that explicitly guard against opportunistic behaviors by any party. Includes clauses that prevent parties from unilaterally altering the terms of use or distribution of contributed content that would disadvantage others.
- Exit rights: Clear terms defining the rights of contributors to withdraw their contributions and the conditions under which this is permissible in order to prevent exploitative uses.
- Amendment and termination clauses: Terms outlining clearly how the license can be amended and under what circumstances it can be terminated, including how these changes affect ongoing projects and contributions.
- Derivative works: Management of intellectual property rights related to derivative works, including if the license permits derivative works to be open source, features attribution requirements, requires derivative works to be shared under the same or similar terms as the original work, or allows for proprietary derivatives that can be commercialized separately.

2.1.3. Collective action governance problem

The collective action problem seeks to solve the problem of designing good rules for the commons, i.e., how to manage commons in a sustainable and equitable way, avoiding the pitfalls commonly described in the tragedy of the commons. Elinor Ostrom discovered that most successful commons shared eight design rules, and our conjecture is that this will also be true of AI innovation commons.[43] An effective licensing regime, we conjecture, will share these same qualities. Translated to generative AI, we identify six key variables.

[43] Ostrom E, *Governing the Commons* (Cambridge University Press 1990).

Variables:

- Access and use rights: Establish clear definitions of who can access and use the AI foundation models, including any limitations on usage, geographic restrictions, or prohibitions on certain types of modifications.
- Participatory governance: Facilitate mechanisms within the license that allow users and contributors to participate in the decision-making processes regarding modifications to the license terms or the governance of the model itself. Ensure it includes representatives from a diverse range of stakeholders, including contributors, users, and possibly affected groups.
- Costs of maintenance: Document mechanisms for users to contribute financially, e.g., crowdfunding campaigns or donation platforms and any paying costs of the maintenance of support for community engagement.[44]
- Accessible dispute resolution mechanisms: Identify clear, accessible, and low-cost mechanisms for dispute resolution mechanisms provided in the license.
- Non-competes clauses: Document the presence of non-competes and specific restricted activities, such as reverse engineering the foundation model, scraping data, identifying weights, or, more generally, developing similar technologies, working for competitors, or starting a competing project. Also, document the geographical scope and duration of non-competes.
- Interoperability and compatibility: Document the conditions for accessing the model and/or its API, including the price (if any), and general accessibility. Identify clauses that explicitly permit integration of the AI model with other software, hardware, or systems. Document statements or guarantees on the compatibility of the AI model with common industry standards, platforms, or environments, and support for data exchange standards that enable the AI model to share and receive data seamlessly with other systems.

We have selected a set ***i***, and for each *i* we manually inspected and analyzed the licensing agreement, scoring each license over the unit interval according to the criteria above for each variable. This gave us an unweighted composite index of 18 variables, clustered into three groupings. The index $\boldsymbol{X_i}$ is measured between 0-2. 2 is a perfect innovation commons. 0 is perfect secrecy or closed source.

All variables have equal weight, which represents a potential limitation. Certain variables, such as Knowledge Accessibility and Access and Usage

[44] Eghbal N, *Working in Public: The Making and Maintenance of Open Source Software* (Stripe Press 2020).

Rights, may exert a greater influence on the robustness of innovation commons compared to others like Credit and Revenue Sharing and Cost of Maintenance. Nevertheless, licenses can only support innovation commons where they score high in all three clusters we have created, which mitigate concerns about the centrality of individual variables. We also justify assigning equal weight to all variables based on our ambition to provide a neutral measure of AI openness that is not influenced by our subjective perception of which variables are more important, as their relative importance cannot be objectively measured. Nonetheless, our analysis provides readers with all the information needed to rank AI foundation models based on non-equally weighted variables. Lastly, note that past studies ranking foundation model openness based on technical variables have followed a similar approach by equally weighting variables, which is a common practice in the dedicated literature.[45]

2.2. Selection Methodology

In the absence of public studies documenting the number of users per foundation model, our selection of AI models is based on the Stanford HAI Artificial Intelligence Index Report 2024. Specifically, we have selected 11 foundation models that include state-of-the-art models, what the report calls “significant model releases,” and models that are frequently discussed in benchmarks. In total, we document 198 variables. Note that as we do not conduct statistics, the size of our sample and possible selection bias do not affect the validity of our result. We aim to present a diverse set of commonly used foundation models, including frontier models, smaller models, models commonly considered open or closed, and widely adopted models. We analyze the following models for this purpose: OpenAI’s GPT-4, Google’s Gemini Ultra, Meta’s Llama 3, MidJourney’s V6, Anthropic’s Claude 3, X’s xAI, Mistral’s 8x7B, BigScience’s Bloom, Cohere’s Aya, Cohere’s Command R, and TII’s Falcon 180B. Our Appendix #1 contains permanent URLs to the documentation we used for our analysis. A reference to these URLs is also included in our table along with the citations.

[45] Liesenfeld A, Lopez A and Dingemanse M, ‘Opening Up ChatGPT: Tracking Openness, Transparency, and Accountability in Instruction-Tuned Text Generators’ (July 2023) in Proceedings of the 5th International Conference on Conversational User Interfaces 1 https://perma.cc/4YNT-UL5B; Bommasani R, Klyman K, Kapoor S, Longpre S, Xiong B, Maslej N and Liang P, ‘The Foundation Model Transparency Index v1.1: May 2024’ (2024) arXiv preprint arXiv:2407.12929.

2.3. Results

The following table shows the 198 variables in the 11 foundation models we tested. It includes detailed information and references for each of these variables. We are making this resource open access to encourage other researchers to extract new insights from it.[46]

[46] Schrepel T and Potts J, *Documenting the Openness of AI Foundation Models* (2024) https://docs.google.com/spreadsheets/d/1XkyHq1oGu00F6qULQLQ5GrRou9zYEh6-fLzV9y7unpE/edit?usp=sharing

Knowledge problem

	Knowledge Accessibility	Documentation and Support	Transparency	Collaboration Platforms	Engagement and Feedback	Language and Localization
	KAC	**DAS**	**TRA**	**CPF**	**EAF**	**LAL**
Cohere (Aya)	2 (Model weights, da	2 (https://perma.cc/2/	2 (https://perma.cc/S	2 (https://perma.cc/8	1 (https://perma.cc	1 (Documentation onl
Bloom (560m)	2 (Model weights, da	2 (https://perma.cc/M	2 (https://perma.cc/Z	2 (https://perma.cc/V	0 (Nothing done to	1 (The technical docu
Mistral (8x7B)	1 (Only the weights a	2 (https://perma.cc/K	1 (https://perma.cc/C	2 (https://perma.cc/V	0 (Nothing done to	1 (The model is "nativ
TII (Falcon 180B)	1 (Model weights ava	2 (https://perma.cc/B	0 (White paper not p	2 (https://perma.cc/6	0 (Nothing done to	1 (Documentation onl
X (xAI)	1 (Only the weights a	2 (https://perma.cc/F	0 (Nothing said)	2 (https://perma.cc/5	0 (Nothing done to At the GenAI layer	0 (The technical docu
Meta (Llama 3)	1 (The model weights	2 (https://perma.cc/V	0 (Little details abou	2 (https://perma.cc/V	0 (Nothing done to	0 (Documentation ava
OpenAI (GPT-4)	0 (The code, training,	2 (https://perma.cc/M	0 (Nothing said exce	2 (https://perma.cc/4	0 (Nothing done to	1 (The technical docu
Google (Gemini)	0 (Nothing said)	2 (https://perma.cc/B	0 (No preprint identi	2 (https://perma.cc/6	0 (Nothing done to	2 (Documentation ava
Midjourney (V6)	0 (Nothing said)	2 (https://perma.cc/C	0 (Nothing said)	2 (https://perma.cc/7	0 (Nothing done to	0 (Documentation ava
Cohere (Command R	0 (The code, training,	2 (https://perma.cc/4(	0 (Very little said ab	2 (https://perma.cc/3	0 (Nothing done to	1 (The technical docu
Anthropic (Claude 3)	0 (Nothing said)	2 (https://perma.cc/6(	0 (Nothing said)	0 (Nothing found)	0 (Nothing done to	0 (General support ava
Score/Variables	**8**	**22**	**5**	**20**	**1**	**8**

Implicit Contracting Problem

	Contribution Policies	Credit and Revenue Sharing	Specific Anti-Opportunism Clauses	Exit Rights	Amendment and Termination Clauses	Derivative works
	COP	**CRS**	**SAO**	**EXR**	**ATC**	**DRW**
Cohere (Aya)	1 (No disclosure on hc "Unless You explicitly	1 (Credit mandatory w	1 (Guarding against op	0 (Nothing said)	1 (Nothing said, which	1 (Permissive use, but
Bloom (560m)	0 (Nothing said on how	0 (Nothing said on how	1 (Guarding against op	0 (Nothing said)	1 (Nothing said, which	2 (Overall permissive Use-based restrictions You must give any Thi
Mistral (8x7B)	1 (Nothing said on how	0 (Nothing said on how	1 (Guarding against op At then GenAI layer: I	0 (Contributions are "i At the GenAI layer, ac	0 (Nothing said at the At the GenAI layer: Te	2 (Permissive use at th When it comes to the (
TII (Falcon 180B)	1 (No disclosure on hc "Unless You explicitly	1 (Credit mandatory w	1 (Guarding against op	0 (Nothing said)	1 (Nothing said, which	1 (Permissive use, but
X (xAI)	0 (Nothing said on how	0 (Nothing said on how	1 (Rules at the founda	0 (Contributions are "i At the GenAI layer, nc	0 (Nothing said at the At the GenAI layer, te	2 (Permissive use at th When it comes to the (
Meta (Llama 3)	1 (Contributor License	0 (Nothing said on how	1 (Guarding against op	0 (No right to withdra Grant of Patent Licens	1 (Termination possibl	2 (Permissive use: "a.
OpenAI (GPT-4)	0 (Nothing said on how	0 (As far as Business T	0 (Nothing said as the	0 (Nothing said as the	1 (As far as Business T As far as the Terms of	2 (Permissive use as fa As far as the Terms of
Google (Gemini)	0 (Nothing said on how	0 (Nothing said on how	0 (Nothing said as the	0 (Nothing said)	0 (Terms are vague an	2 (Permissive use: "Sc
Midjourney (V6)	0 (Nothing said on how	0 (Nothing said on how	0 (Nothing said as the	0 (Nothing said as the	0 (The terms are vague	2 (Permissive use: "Yo Your ownership is sub If you are a company (
Cohere (Command R	0 (No disclosure on hc No attribution or rewa	0 (Nothing said on how	1 (Guarding against op	0 (Nothing said)	0 (For users, reasonabl However, the Terms of The terms are seeming	0 (Limited use: For us reproduce and Share t produce, reproduce, a
Anthropic (Claude 3)	0 (No disclosure on hc Anthropic has a "Resp	0 (Nothing said on how	0 (Nothing said as the	0 (Nothing said)	0 (Termination for "an Unless we specifically Changes to these term	2 (Nothing said for us
Score/Variables	**4**	**2**	**7**	**0**	**5**	**18**

Collective action governance problem

	Access and Use Rights	Participatory Governance	Costs of maintenance	Accessible Dispute Resolution Mechanisms	Non-competes clauses	Interoperability and Compatibility	**TOTAL SCORE**
	AUR	**PGV**	**COM**	**ADR**	**NCC**	**IOC**	
Cohere (Aya)	2 (Basic restrictions in	0 (Nothing said about	0 (Nothing said)	0 (Nothing said)	2 (Nothing said, which	1 (Nothing said about	**20**
Bloom (560m)	1 (Basic restrictions or "5. Use-based restricti	1 (Remains vague, but With regards to the res ("7. Updates and Runt	0 (Nothing said)	0 (Nothing said)	2 (Nothing said, which	1 (Nothing said about	**18**
Mistral (8x7B)	2 (Basic restrictions or When it comes to the (	0 (Nothing said when At the GenAI layer: no	0 (Nothing said)	1 (Nothing said at the Terms of Use: "Indem	2 (Nothing said, which	1 (Access to the mode	**17**
TII (Falcon 180B)	2 (Basic restrictions in	0 (Nothing said about	0 (Nothing said)	0 (Nothing said)	2 (Nothing said, which	1 (Access to the API/n 3. Subject to Section 9 4. Where TII grants pe	**15**
X (xAI)	2 (Basic restrictions or When it comes to the (	0 (Nothing said when At the GenAI layer, ar	0 (Nothing said)	1 (Nothing said at the At the GenAI layer, m	2 (Nothing said, which	1 (Nothing said, but ac	**14**
Meta (Llama 3)	1 (Basic restrictions ar	0 (Nothing said about	0 (Nothing said)	0 (Nothing said, but w	0 (Strict non-compete:	1 (Nothing said about	**12**
OpenAI (GPT-4)	0 (Restrictive use case As far as getting acces	0 (Nothing said as the	0 (Nothing said as the	2 (As far as Business T As far as the Terms of	0 (Strict non-compete. As far as the Terms of	0 (Access not free and	**10**
Google (Gemini)	1 (Basic restrictions, a	0 (Nothing said about	0 (nothing said as the	0 (No dispute resoluti	0 (Strict non-compete: "You may not use the	0 (Nothing said about	**9**
Midjourney (V6)	0 (Basic restrictions or	0 (Nothing said as the	0 (Nothing said as the	2 (Clear dispute resolu Each party will bear it	0 (Strict non-compete:	0 (Nothing said about	**8**
Cohere (Command R	0 (Access is strictly lir The CohereSolution a The Software and all	0 (Governance is not p The same goes for con	0 (Nothing said as the	0 (No dispute resoluti The same goes for nor	0 (Strict non-compete Commercial Saas Agr	1 (Documentation ava The Saas licence ment	**7**
Anthropic (Claude 3)	1 (Basic restrictions (a As far as commercial	0 (Non-participatory: "	0 (Nothing said as the	1 (No access to disput Governing law and ex Access to dispute resol	0 (Strict non-compete:	0 (Nothing said about	**6**
Score/Variables	**12**	**1**	**0**	**7**	**10**	**7**	

Table #1: "Documenting the openness of AI foundation models"
Access: Click here to access it.

Legend: The variables have been scored as follows: red (0 points) when the license strictly restricts innovation commons, yellow (1 point) when the license has a mixed impact on innovation commons, and green (2 points) when the license maximizes innovation commons. The provisions corresponding to each variable have been extracted from the licenses for verification purposes.

Four insights emerge from the table. The *first* outcome is that our results differ significantly from previous attempts to rank the openness of AI foundations from a technical perspective. The ranking order is different, and the magnitude of the gaps varies. In terms of ranking order, OpenAI does not rank last in our study as it does in Liesenfeld et al. (2023).[47] Bloom does not rank first in our ranking, whereas it does in Liesenfeld et al. (2023). Our results also differ from those of Bommasani et al. (2024), who, relying on technical variables, find that Llama 2 is more open than Mistral 7B and that Claude 3 is more open than Google's Gemini (2024).[48] Similarly, Liu et al. (2023) found that Llama-2 was more open than Falcon 7b while we find the opposite.[49]

When it comes to the size of the gap, our analysis shows that the distinction between so-called "open" and "closed" foundation models is not as clear-cut as purely technical analyses suggest or as the AI Act—which relies on a binary dichotomy between open and closed systems to establish exemptions—portrays.[50] For example, Meta's Llama 3 and OpenAI's GPT-4 are separated by only 2 points out of a possible 36. They score 12 and 10 points out of 36, respectively. Even the most open model we tested scores only 20 points out of 36. The fact that the main models occupy the middle of the spectrum is absent from technical rankings of AI model openness. Liesenfeld et al. (2023), for instance, ranked OpenAI's GPT-4 openness entirely with negative

---

47 Liesenfeld A, Lopez A and Dingemanse M, 'Opening Up ChatGPT: Tracking Openness, Transparency, and Accountability in Instruction-Tuned Text Generators' (July 2023) in *Proceedings of the 5th International Conference on Conversational User Interfaces* 1.

48 Bommasani R, Klyman K, Kapoor S, Longpre S, Xiong B, Maslej N and Liang P, 'The Foundation Model Transparency Index v1.1: May 2024' (2024) arXiv preprint arXiv:2407.12929

49 Liu Z, Qiao A, Neiswanger W, Wang H, Tan B, Tao T, Li J, Wang Y, Sun S, Pangarkar O, Fan R, Gu Y, Miller V, Zhuang Y, He G, Li H, Koto F, Tang L, Ranjan N, Shen Z, Ren X, Iriondo R, Mu C, Hu Z, Schulze M, Nakov P, Baldwin T and Xing EP, 'LLM360: Towards Fully Transparent Open-Source LLMs' (2023) arXiv preprint arXiv:2312.06550.

50 Maffulli S, "Meta's LLaMa 2 License Is Not Open Source" (Open Source Initiative, June 20, 2023) https://perma.cc/8UKD-G6VE; Biderman S, Schoelkopf H, Anthony Q, Bradley H, O'Brien K, Hallahan E, Khan MA, Purohit S, Prashanth USVSN, Raff E, Skowron A, Sutawika L and van der Wal O, 'Pythia: A Suite for Analyzing Large Language Models Across Training and Scaling' (2023) arXiv preprint arXiv:2304.01373; White M, Haddad I, Osborne C, Yang X-Y, Liu L, Abdelmonsef A, Varghese S and Le Hors A, 'The Model Openness Framework: Promoting Completeness and Openness for Reproducibility, Transparency, and Usability in Artificial Intelligence' (2024) arXiv preprint arXiv:2403.13784 ("Openness has always been a binary decision in the open-source movement; software is either open-source or not, with no in-between")

points except for one variable, while, in contrast, they scored one of BigScience's models as entirely open except for one parameter.[51] We therefore caution against characterizing openness as a binary concept, as the AI Act does, or loosely labeling models as "closed" or "open source." Most AI foundation models fall within the middle to lower end of the openness spectrum, with none occupying the extremes. Using such binary or loose approaches risks misclassifying AI foundation models, which is particularly problematic in regulations like the AI Act, where exemptions depend on being designated as "open."

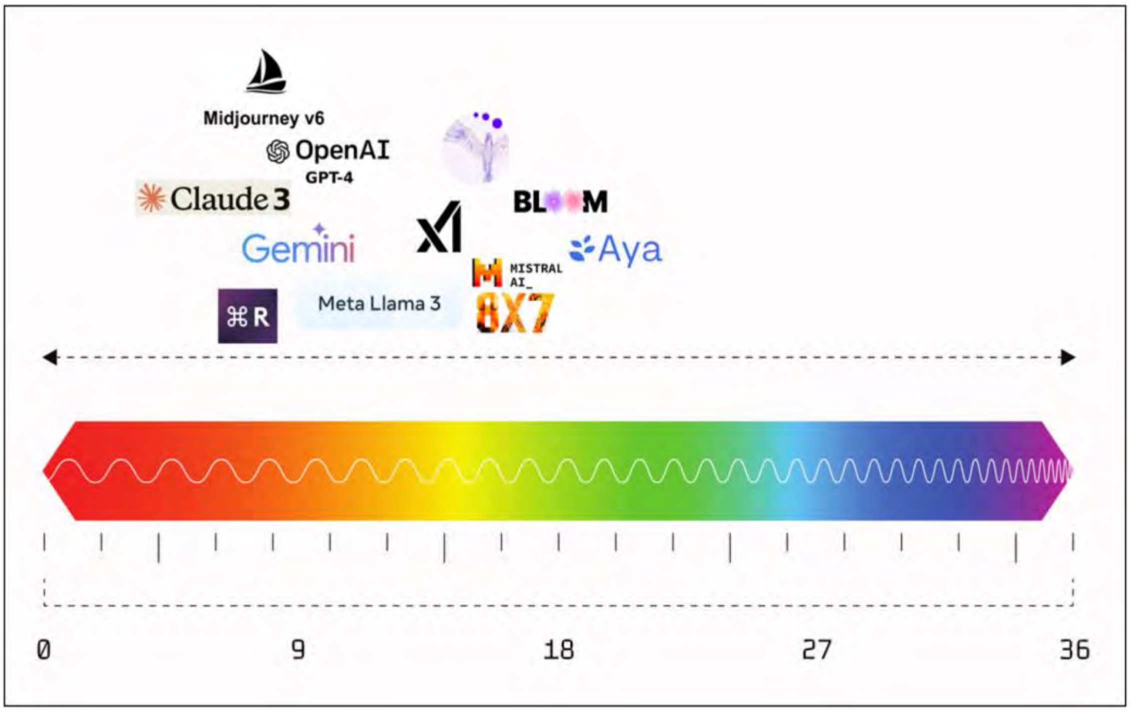

Figure #1: Ranking the openness of AI foundation models

Second, we see some trends among the AI foundation models. In general, they score 64 out of 132 points when it comes to addressing the "knowledge problem." They score more than twice as low when it comes to addressing the "implicit contracting problem" and the "collective action

[51] Liesenfeld A, Lopez A and Dingemanse M, 'Opening Up ChatGPT: Tracking Openness, Transparency, and Accountability in Instruction-Tuned Text Generators' (July 2023) in *Proceedings of the 5th International Conference on Conversational User Interfaces* 1.

governance problem," with a total of 35 out of 132 points for each. Part of this may be a simple consequence of our unweighted approach, suggesting that the knowledge problem simply looms larger in concern with openness than opportunism (implicit contracting problem) and governance (collective action problem). But we suspect that this is more likely a consequence of the very early stage of the new industry, and that as it develops these issues of strategic coordination will become increasingly significant for innovation and competition. Companies and organizations that want to open up their model have much to gain by improving their scores in these two areas, which are inherently less easily addressed by technical fixes and solutions.

Specifically, AI foundation models score very high on "documentation and support" (with a total of 22 out of 22 possible points), "collaboration platforms" (with a total of 20 out of 22 possible points), and "derivative works," (with a total of 18 out of 22 possible points). We observe a similar trend whether the models are at the open-source or closed-source end of the spectrum. Conversely, AI foundation models score very low on "exit rights" (with a total of 0 out of 22 possible points), "costs of maintenance" (with a total of 0 out of 22 possible points), "engagement and feedback" (with a total of 1 out of 22 possible points), "participatory governance" (with a total of 1 out of 22 possible points), and "credit and revenue sharing" (with a total of 2 out of 22 possible points). We also observe a similar trend whether the models are at the open source or closed source end of the spectrum. As above, the degree of openness from documentation and support and collaboration platforms is a happy consequence of the prior cultures and tools of open source software, as is the zero total scoring on maintenance costs.[52] But exit rights and credit and revenue sharing are both looming problems yet to be solved, and it is perhaps unsurprising to see clear signals of what could be costly commitments made explicit at this early stage. We suggest these could be areas for competition regulators to pay increasing attention in the near future.

Third, we find consistent <u>similarities</u> and <u>differences</u> between models at the "closed source" end of the spectrum and models at the "open source" end of it. In terms of similarities, the general openness of these models has little to no effect on the downstream use of the output. The more "open source" models impose restrictions on the usage of derivative works, for example, they prohibit

[52] Eghbal N, *Working in Public: The Making and Maintenance of Open Source Software* (Stripe Press 2020).

the use of their model for criminal activity. As for the more "closed source" models, they typically do not capture the copyrights of derivative works.

In terms of differences, the distinction between the more open and the more closed models has a clear impact on knowledge accessibility and access to the model (upstream). The more open a model is in general, the more free and easy the access is. Cohere's Aya and Bloom's 560M — the two models rank first and second overall — provide access to code, model weights, and datasets. They are the only two models to score 2 points out of 2 on "knowledge accessibility." The bottom five models in the overall ranking, OpenAI's GPT-4, Google's Gemini, Midjourney's V6, Cohere's Command R, and Anthropic's Claude 3, each score 0 points out of 2 on "knowledge accessibility" and "interoperability and compatibility," with the exception of Cohere's Command R, which scores 1 point on "interoperability and compatibility".

Fourth, the divide between big tech companies and others that has been used and abused over the past 20 years is of little use when it comes to GenAI. X, Meta, Google, and OpenAI (with a Microsoft partnership) are in the middle of our openness ranking. This means that the AI foundation models at the top and bottom of our ranking are provided by companies that are not typically described as 'big tech.' Given the importance of openness in the field for the reasons outlined above, the sensational divide between big tech and others should be abandoned.

3. Competition and Policy Implications

Our systematic evaluation of AI Foundation Models licenses provides valuable insights for policymakers and enforcers active in the AI space, highlighting the economic incentives and legal challenges of promoting licensing arrangements that help the innovation commons and technological progress flourish. On this basis, this section formulates actionable policy proposals to promote the AI foundation models that contribute most to the innovation commons (**3.1.**). The analysis continues with the development of enforcement strategies (**3.2.**). It provides a holistic view of regulation and enforcement, addressing issues of competition law, (intellectual) property law, data protection, public international law, and more.

3.1. Implications For Policymakers

Policymakers may consider two distinct levels of protection for the innovation commons benefiting generative AI. We first lay down the *minimal* requirements essential for the sustenance of innovation commons in the generative AI ecosystem. We then describe a *maximal* regulatory framework that would enable innovation commons to flourish. Policymakers will tailor their strategies according to political preferences and various constraints.

At *minimum*, we suggest that policymakers list all existing regulations affecting the 18 variables we have listed and ask whether these regulations are neutral or complicate the efforts of organizations to improve their openness scores. If the regulations fall into the latter category, we recommend conducting a new cost-benefit analysis to reassess their desirability in light of the benefits of generative AI.[53] Questions to ask about existing regulations impacting the variables we have used include:

1. Whether data protection makes it difficult to share datasets;
2. Whether intellectual property may have a similar effect on the sharing of training datasets;
3. Whether data protection hinder the development and adoption of open-source data processing and analytics tools by imposing strict data

[53] We recognize that fostering innovation commons is not the sole objective of policymakers. This objective should be balanced with others, which is why we suggest a cost-benefit analysis. For example, exploring the need to mitigate risks and how they relate to the openness of models, *see* Sayash Kapoor and others, 'On the Societal Impact of Open Foundation Models' (27 February 2024) arXiv:2403.07918.

handling requirements practices that are more difficult for decentralized, networked projects to meet;
4. Whether liability rules could lead companies to further restrict the use of derivative works;
5. Whether trade secrecy rules push AI training data and algorithms into leave the open;
6. Whether competition law makes it hard for companies to share data, with the risk of creating a cartel;
7. Whether corporate governance laws push for centralized control of foundation models in order to manage liability;
8. Whether standards and interoperability regulations inadvertently disadvantage open-source software by favoring proprietary technologies or closed ecosystems;
9. Whether cybersecurity regulations impose compliance burdens that are disproportionately difficult for open-source software projects to meet, hindering their adoption in critical infrastructure and sensitive environments;
10. Whether patent laws create barriers to open-source projects by allowing proprietary software companies to patent basic algorithms or software functionality, limiting their availability for open collaboration.

We also suggest that policymakers undertake analogous assessments for prospective regulations. Competition agencies, contingent upon their institutional framework, can assist in this cost-benefit assessment or undertake autonomously. Recent historical trends underscore the paramount importance of freeing ecosystems from regulatory capture and barriers to ensure openness worldwide.[54]

If policymakers want to take further action to foster innovation commons (i.e., implement a *maximalist* approach),[55] we suggest that they promote openness by adopting the following three-pillar agenda. The agenda, here again, builds on our systematic content analysis of AI foundation model licenses.

[54] As of January 2025, we have not identified a single competition agency publicly addressing regulatory capture or documenting the impact of regulation on innovation and competition in the space, *see* Schrepel T, Yerebakan A, and Baladima N, 'A Database of Antitrust Initiatives Targeting Generative AI' (2023–2025) *Network Law Review* (Winter 2023–2024) https://perma.cc/U99C-D3CE.
[55] E.g., Potts J, Torrance A, Harhoff D and von Hippel E, 'Profiting from Data Commons: Theory, Evidence, and Strategy Implications' (2024) 9(1) *Strategy Science* 1.

First, policymakers might consider introducing legal exemptions into new rules and standards for the more open system, as defined by our methodology. The AI Act has begun to address this issue, yet it falls short because it defines "open" solely in technical terms and, remarkably, does not provide exemptions for open systems that have scaled. This situation is problematic because the AI Act does not effectively discern between genuinely open or closed foundation models, focusing instead on mere technicalities. Additionally, the burden of compliance tends to be disproportionately higher for open source systems, even those that have scaled. Using Article 11 of the AI Act as an example, it is easy to see how organizations that release open source AI foundation models which can be modified freely by any user will have a harder time than companies that develop closed source models in-house to provide "[t]he technical documentation of a high-risk AI system (...) before that system is placed on the market or put into service" and to keep it "up-to date."[56]

In addition, open source models tend to have more inconsistencies in the understanding and application of legal standards, especially when contributors come from different jurisdictions with varying legal environments. This problem is compounded by the fact that open source projects typically do not have a centralized authority responsible for ensuring compliance. The decentralized nature of these projects means that while anyone can contribute, there might not be a dedicated team or individual tasked with compliance oversight. This lack of systematic legal analysis can lead to potential gaps in meeting legal requirements. Last, many open source projects operate with limited financial resources. Unlike commercial entities that develop closed source software, open source projects may not have the funding to hire legal experts or consultants to ensure ongoing compliance with new and existing regulations. This can be particularly challenging as laws and standards evolve.

Some of these concerns are mitigated by the involvement of corporations in driving some 'open' AI foundation models, rather than decentralized networks. Nevertheless, compliance costs remain higher for the more 'open' AI models. Companies running open models must review each user contribution, and often face greater legal variability than those managing closed systems with in-house developers who are informed about legal issues. For this reason, legal

[56] Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence (Artificial Intelligence Act) [2024] OJ L 1689/1, art 11 https://perma.cc/95BB-JK6P

exemptions appear to be more than just a nice feature for the development of open AI models.

Second, legal exemptions could be coupled with ambitious economic measures to help tilt the economic balance toward open source models. Several policies could help, including:

1. Tax deductions for organizations that develop AI models and choose to release them under open-source licenses;
2. Public funding for projects that commit to open-sourcing their results;
3. Preferential treatment given by governments to open-source AI systems in their procurement processes;
4. Economic support for open source AI foundations, similar to the role of the Linux Foundation in supporting open source software.
5. Partnerships between academia, industry, and government to advance open source AI projects.

Third, policymakers could provide technical support to open projects. To be clear, we are not saying that policymakers should engage in development of AI foundation models, but we are saying that they could provide developers support in the following ways:

1. Impose transparency requirements requiring companies to disclose the code, weight, and/or training data of their AI models under certain conditions, for example, when the models are developed with public funds.
2. Promote the creation and maintenance of open data repositories that can be used safely and ethically for training AI models. This would reduce barriers to entry to train AI and ensure a more level playing field.
3. Develop and enforce standards to ensure that AI systems are interoperable, which could naturally lead to more open structures.

These three pillars would break with the principle of “technological neutrality” in favor of opening up the generative AI ecosystem. They would not, however, reduce the incentives to develop closed-source models. Closed-source models play an important role in fostering competitive dynamics and innovation in the space. They tend to be Schumpeterian in nature and provide a clear and direct revenue stream. The closed-source financial model can fund further research and development, leading to innovative products and services. Closed source systems are thus an important competitive force in the generative

AI ecosystem, but, as we explained, so is open source. Given that open source is more vulnerable due to the disproportionate legal burden placed on it, the more fragile economic incentives, and the reliance on a few players to provide state-of-the-art technical advances, we believe that policymakers should take a proactive approach to protecting and promoting open source AI foundation models.[57]

[57] Evaluating open source foundation models chances of survival from the perspective of complexity economics, *see* Schrepel T and Pentland AS, 'Competition Between AI Foundation Models: Dynamics and Policy Recommendations' (2024) *Industrial and Corporate Change.*

3.2. Implications For Enforcers

As evidenced by the 18 variables we documented, the more open licenses, the more innovation commons benefit. Given that innovation drives competition as much as competition drives innovation in digital ecosystems, innovation commons are a central element of tomorrow's competition.[58] This already suggests why antitrust agencies might want to focus on closed models. But beyond that, the more open licenses are, the lower the risk of anticompetitive behavior. It follows that our systematic content analysis provides clear targets for antitrust enforcers, whereas an analysis limited to the technical aspects of openness does not. As highlighted by the Federal Trade Commission, models that are technically open can still be subject to licenses that enable anti-competitive strategies.[59] Considering non-technical aspects when defining the openness of AI licenses offers a more practical approach to antitrust enforcement. There are three main reasons for this:

1. Open foundation models, such as we define them, are more transparent than closed ones. They can be more easily scrutinized by antitrust agencies. For example, the number and identity of contributors, the terms of access, the provenance of data, and other parameters could be relevant to antitrust analysis. Our table makes this clear: our overall ranking of openness is strongly correlated with how many points models score in terms of "knowledge accessibility." Given that the more knowledge is accessible, the more competition there is, agencies should naturally target closed-source models.
2. Open-source foundation models can be forked, meaning that they can be freely duplicated and adapted. The more open foundation models are typically forked thousands of times, creating diversity and thus competition through innovation.[60] Our table highlights this market reality: the top 5 more open foundation models we have identified score 2 points out of 2 in "non-compete," meaning they do not include technical or legal

[58] Schrepel T, 'A Systematic Content Analysis of Innovation in European Competition Law' (2023) *Amsterdam Law & Technology Institute (ALTI) Working Paper* 2-2023; Petit N and Teece DJ, 'Innovating Big Tech Firms and Competition Policy: Favoring Dynamic Over Static Competition' (2021) 30(5) *Industrial and Corporate Change* 1168.
[59] Staff in the Office of Technology, "On Open-Weights Foundation Models" (*Federal Trade Commission*, July 10, 2024) https://perma.cc/F252-L5V3.
[60] For example, in less than a month after its release, Meta's Llama 3 has been forked over 1,900 times on GitHub, https://perma.cc/NZ62-8U8U.

restrictions on forking. This alone justifies why agencies may want to focus on the more closed-source models.

3. The companies and organizations offering the more open models do not have the typical power to leverage their market position that we see with closed, proprietary models, software, and services. The top 5 more open AI models listed in our table score a total of 9 points out of 10 points when it comes to "access and use rights." This means that the AI models are freely accessible, with no barriers to entry. They also score better on "interoperability and compatibility" than the bottom 5 models, which score 0. As a result, their code and/or API are directly accessible by competitors and downstream companies, leading to a thriving ecosystem without barriers to access the model. The same goes for "amendment and termination clauses." Although the top 5 open models score only 3 points out of 10, the bottom 5 score only 1 point, leading to greater concerns about refusal to deal. Finally, the top 5 open models score 5 points on "Specific Anti-Opportunism Clauses" against 1 point for the bottom 5, meaning that they better protect against technical capture of ecosystems. It follows that companies and organizations behind the more open models cannot, as easily as those operating closed models, rely on the market power gained through their foundation models to foreclose the ecosystem. Market power can only be leveraged through restrictions to openness.[61] The openness of AI models also prevents most vertical restraints and agreements between rivals, as companies forgo the technical, economic, and legal power to restrict competition with their open models. The more open their model is, the less leveraging power they retain, eventually reaching the point of having none.[62]

For these reasons, and given that many recent antitrust cases (in the EU, US, and elsewhere) and market regulations (such as the Digital Markets Act)

[61] The Portuguese Competition Authority highlights some of these strategies, *see* Portuguese Competition Authority, 'Competition and Generative AI: Opening AI Models' (AdC Short Papers, December 2024). Notably, all these strategies rely on restrictions to AI openness, undermining any claim that the related models are simply "open." Models that enable anti-competitive strategies are, at best, only partially open.

[62] To provide an example, the European Commission sanctioned Google for Android-related practices in 2018, *see* European Commission, Case AT.40099 Google Android C(2018) 4761 final. Android is closer to the open-source model than closed-source, so this case might seem surprising at first glance. However, the European Commission sanctioned Google for its Android anti-fragmentation policy, which imposes a restriction on open-source principles. A fully open-source Android would not have included such a policy and, therefore, would not have raised these competition law concerns.

cover leveraging practices, antitrust agencies should be sensitive to AI openness (such as we define it) as an antidote to this type of anti-competitive concern. This does not mean that the more open AI models have no anti-competitive risk. But it does mean that the more closed models are more likely to lead to anti-competitive practices, making them a clear target for antitrust agencies.

4. Conclusion

The deep-learning revolution in neural networks, barely a decade old, has ushered into the world a transformative new industrial technology called foundation models that are the basis of powerful applications known as generative AI. This paper has sought to contribute to public policy discussions and interventions about how antitrust or competition policy might be guided to shape the development of this new technology. The challenge is that standard economic policy models used to evaluate competitive dynamics and social welfare (with a focus on how market structure affects firms' incentives and ability to innovate) misses the mark due to the digital and general-purpose nature of the technology. Instead, our primary assertion is that progress in foundation models and generative AI is significantly driven by open-source development within the innovation commons. This dynamic not only stimulates competition but also promotes the discovery of value and enhances productivity across multiple markets and industries.

To assist competition agencies and regulators in supporting the innovation commons, we developed an index measure that we applied to 11 leading foundation models to estimate their degree of openness from an economic standpoint. Our analysis focused on foundation model licenses and public documents that we scarred over 18 variables, and which we broadly organized into three clusters in the way they solved the institutional problems of knowledge sharing, opportunism and governance. We found that, contrary to the technical definitions of openness, most models are still relatively closed, with the best scores of 17 to 20 out of a possible 36, going to smaller, specialized projects (Aya, 560m, Mistral), and with much of those scores loading heavily onto knowledge sharing. Large corporate models, like Meta's Llama and Google's Gemini, achieved mid-range scores in our index (12/36 and 9/36, respectively), due to more restrictive conditions.

There are obvious limitations to our analysis, given that it covered a fraction of the models in the public domain and the rapidly evolving nature of these models, such as changes in training sets. These limitations do not impact the validity of our study. Our analysis primarily measured the institutional openness resulting from organizational decisions and investments, which are critical as they fundamentally determine the economic costs of innovation and the reality of competition in this industry.

Finally, we derived actionable policy insights and enforcement targets from our systematic content analysis of AI models. Given that open-source AI models help foster competition and innovation without raising the anti-competitive concerns associated with closed-source models, we argue that policymakers should seek to ensure a regulatory environment that is at least neutral and at most favorable to the more open AI foundation models, and especially so given the inherently higher costs that openness imposes on an organization. Meanwhile, enforcers should target closed-source models as a priority, as they pose much greater anticompetitive risks. The openness of AI models provides a clear guide for defining antitrust targets. The more proprietary the model, the greater the risk. This, we hope, will help antitrust agencies define where to invest resources, rather than focusing on the identity or size of the companies offering these models.

Table #1: Variable grouping and list

	Variable name	Code
Knowledge problem	Knowledge Accessibility	KAC
	Documentation and Support	DAS
	Transparency	TRA
	Collaboration Platforms	CPF
	Engagement and Feedback	EAF
	Language and Localization	LAL
Implicit Contracting Problem	Contribution Policies	COP
	Credit and Revenue Sharing	CRS
	Specific Anti-Opportunism Clauses	SAO
	Exit Rights	EXR
	Amendment and Termination Clauses	ATC
	Derivative works	DRW
Collective action governance problem	Access and Use Rights	AUR
	Participatory Governance	PGV
	Costs of maintenance	COM
	Accessible Dispute Resolution Mechanisms	ADR
	Non-competes clauses	NCC
	Interoperability and Compatibility	IOC

Appendix #1: Permanent URLs to AI Foundation Models Documentation

Alphabetical order

Anthropic' Claude 3:

- Anthropic Consumer Terms (Perma.cc)
- Anthropic Commercial Terms (Perma.cc)
- Anthropic Privacy Policy (Perma.cc)
- Anthropic Acceptable Use Policy (AUP) (Perma.cc)
- Anthropic Responsible Disclosure Policy (Perma.cc)
- Getting Started with the API (perma.cc)
- Anthropic Support (English) (perma.cc)
- Anthropic Support (French) (perma.cc)
- Can I Use Claude in Different Languages? (perma.cc)
- Claude 3 Family News (perma.cc)
- Consumer Terms (perma.cc)
- Commercial Terms (perma.cc)

BigScience's Bloom:

- BigScience RAIL License (Perma.cc)
- License on Hugging Face (Perma.cc)
- BigScience Model Repository (perma.cc)
- Bloomz-560M Model (perma.cc)
- BigScience Ethical Charter (perma.cc)
- BigScience Workshop GitHub (perma.cc)
- BigScience Blog: Bloom (perma.cc)
- BigScience License Space (perma.cc)

Cohere's Command R:

- Terms of Use (Perma.cc)
- Privacy (Perma.cc)
- SaaS Agreement (Perma.cc)
- Service Level Objective (Perma.cc)
- Responsibility (Perma.cc)
- Security (Perma.cc)

- Data Usage Policy (Perma.cc)
- CC-BY-NC 4.0 License with Addendum (Perma.cc)
- Cohere Platform Documentation (perma.cc)
- Cohere LLMU Documentation (perma.cc)
- Cohere Community Discord (perma.cc)
- Cohere Command-R Documentation (perma.cc)
- Cohere C4AI CC BY-NC License (perma.cc)
- Cohere Integrations Documentation (perma.cc)

Cohere's Aya:

- Aya on Hugging Face (Perma.cc)
- AYA Collection Dataset (perma.cc)
- AYA Model Paper (perma.cc)
- AYA-101 Discussions (perma.cc)
- AYA Website (perma.cc)
- AYA Research (perma.cc)
- AYA Annotations UI GitHub (perma.cc)

Google's Gemini:

- Google Policies (Perma.cc)
- Google Privacy Policy (Perma.cc)
- Google Terms of Service (Perma.cc)
- Google Terms for Generative AI (Perma.cc)
- Google AI Terms (Perma.cc)
- Google AI Terms Preview (Perma.cc)
- Google Gemini API Terms Preview (Perma.cc)
- Google Gemini API Docs (Perma.cc)
- Google Gemini API AI Studio Quickstart (Perma.cc)
- Google Gemini API Get Started (Perma.cc)
- Gemini API Documentation (perma.cc)
- AI Studio Quickstart Guide (perma.cc)
- Google AI Discussion Forum (perma.cc)
- Gemini FAQ (perma.cc)
- Google Generative AI Terms (perma.cc)

Meta's Llama 3:

- Llama License (Perma.cc)
- Llama Acceptable Use Policy (Perma.cc)
- Responsible Use Guide (Perma.cc)
- Llama 3 GitHub Repository (perma.cc)
- Llama Model on Hugging Face (perma.cc)
- Meta-LLAMA Blog Post (perma.cc)
- LLAMA GitHub Contribution Guidelines (perma.cc)
- Facebook Individual Contributor License Agreement (perma.cc)

MidJourney's V6:

- Terms of Service (Perma.cc)
- Privacy Policy (Perma.cc)
- Data Deletion and Privacy FAQ (Perma.cc)
- Community Guidelines (Perma.cc)
- Midjourney Documentation (Perma.cc)
- Midjourney Discord (Perma.cc)
- Midjourney Discord Documentation (Perma.cc)

Mistral's 8x7B:

- Terms of Use (Perma.cc)
- Privacy Policy (Perma.cc)
- Data Processing Agreement (Perma.cc)
- Mixtral 8x7B on Hugging Face (Perma.cc)
- Mistral Documentation: Getting Started with Models (perma.cc)
- Mistral Documentation (perma.cc)
- Mistral Deployment Platform Documentation (French) (perma.cc)
- Mistral Contribution Guides (perma.cc)
- Mistral GitHub Repository (perma.cc)
- Mistral Discord Community (perma.cc)
- Mistral News: Mistral Large (French) (perma.cc)
- Apache License 2.0 (perma.cc)
- Mistral Terms of Use (French) (perma.cc)

OpenAI's GPT-4:

- OpenAI Policies (Perma.cc)
- OpenAI Terms of Use (Perma.cc)
- OpenAI Privacy Policy (Perma.cc)
- OpenAI Service Terms (Perma.cc)
- OpenAI Data Processing Addendum (Perma.cc)
- OpenAI Plugin Terms (Perma.cc)
- OpenAI Service Credit Terms (Perma.cc)
- OpenAI Business Terms (Perma.cc)
- OpenAI Usage Policies (Perma.cc)
- OpenAI Enterprise Privacy (Perma.cc)
- OpenAI Sharing and Publication Policy (Perma.cc)
- OpenAI Coordinated Vulnerability Disclosure Policy (Perma.cc)
- OpenAI Brand Guidelines (Perma.cc)

TII's Falcon 180B:

- Terms and Conditions (Perma.cc)
- Acceptable Use Policy (Perma.cc)
- Falcon-180B Model (perma.cc)
- Falcon Blog Post (perma.cc)
- Falcon RefinedWeb Dataset (perma.cc)
- Falcon RefinedWeb Discussions (perma.cc)
- Falcon-180B News (perma.cc)
- Falcon Terms and Conditions (perma.cc)

X's xAI:

- Terms of Service (Perma.cc)
- Privacy Policy (Perma.cc)
- Apache 2.0 License (Perma.cc)
- Grok-1 Model (perma.cc)
- XAI.org Developers (perma.cc)
- XAI.org GitHub (perma.cc)
- Grok-1 Discussions (perma.cc)
- X.ai Terms of Service (perma.cc

About the authors:

Thibault Schrepel is Associate Professor of Law at the Vrije Universiteit Amsterdam, and Faculty Affiliate at Stanford University CodeX Center.

Jason Potts is a Distinguished Professor of Economics at RMIT University, and Co-director of the Blockchain Innovation Hub at RMIT.

About the Digital, Governance and Sovereignty Chair:

Sciences Po's Digital, Governance and Sovereignty Chair's mission is to foster a unique forum bringing together technical companies, academia, policymakers, civil societies stakeholders, public policy incubators as well as digital regulation experts.

Hosted by the **School of Public Affairs**, the Chair adopts a multidisciplinary and holistic approach to research and analyze the economic, legal, social and institutional transformations brought by digital innovation. The Digital, Governance and Sovereignty Chair is chaired by **Florence G'sell**, Professor of Law at the Université de Lorraine, lecturer at the Sciences Po School of Public Affairs, and visiting professor at the Cyber Policy Center of Stanford University.

The Chair's activities are supported by: