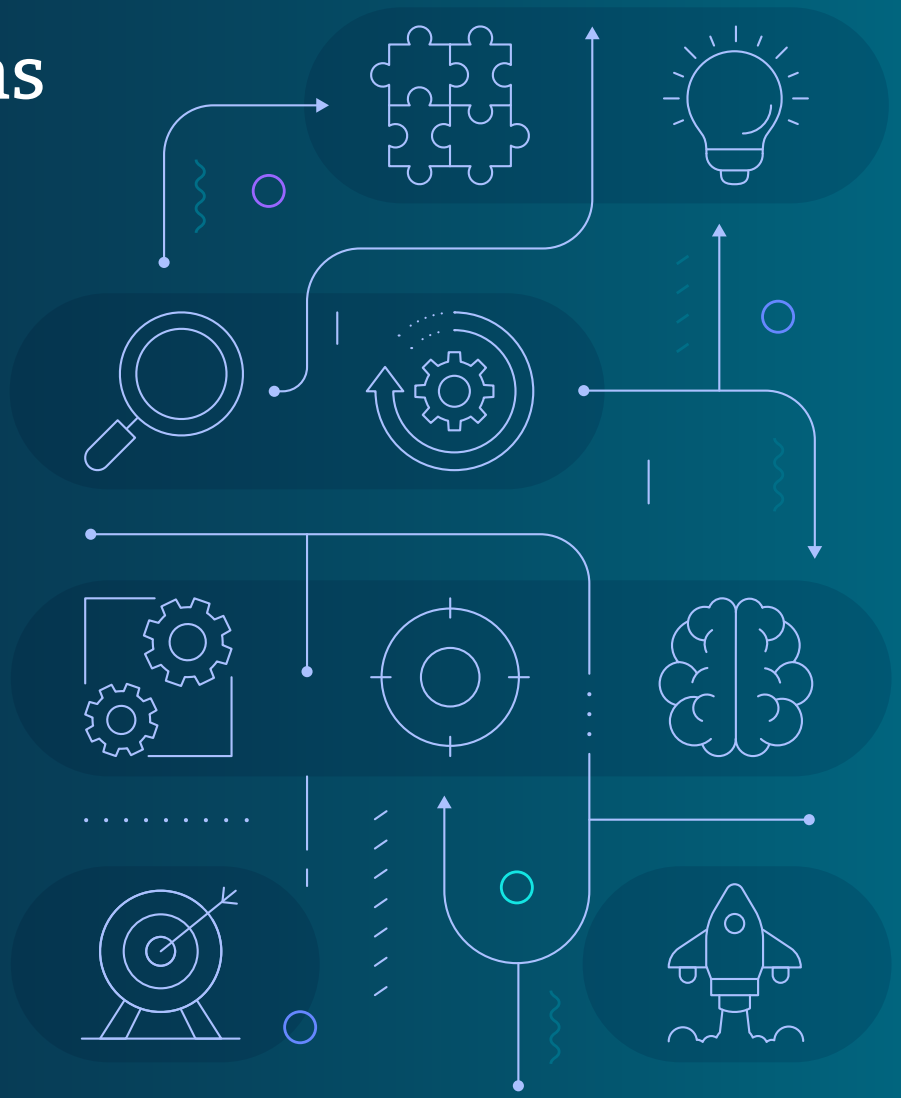
THE LINUX FOUNDATION | Research

GOSIM

ECLIPSE FOUNDATION

Charting Strategic Directions for Global Collaboration in Open Source AI

Key Takeaways from the GOSIM Open Source AI Strategy Forum



Cailean Osborne, PhD, *The Linux Foundation*

Foreword by
Nicolas Flores-Herr, PhD, *Fraunhofer IAIS*

July 2025

Sponsored by FUTUREWEI Technologies

Charting Strategic Directions for Global Collaboration in Open Source AI

Developers should champion open development practices and licenses that **uphold the four freedoms of open source in open source AI.**

Governments should **embrace open source AI as a strategic tool** to foster both global collaboration and competition in AI.

Small models increasingly rival the performance of large models while providing affordability, customizability, and privacy benefits.

The **open science approach to AI R&D** facilitates reproducibility, knowledge sharing, learning, and auditing of potential biases and security vulnerabilities.

The success of open source AI relies not only on sharing models, code, and datasets but also on **building and sustaining a vibrant global community.**

Developing **open datasets for medium- and low-resource languages** is a key priority for tackling cultural and linguistic biases in LLMs.

Open source developers influence responsible AI practices by sharing tools and documentation that others can use, learn from, and build on.

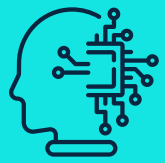
Open source evaluation frameworks and benchmarks for diverse capabilities and contexts are key to enhancing trust and safety in open source AI.

Access to public compute infrastructure is a major bottleneck for the participation of researchers and grassroots initiatives in open source AI.

Innovative funding mechanisms and business models are needed to support the growth and sustainability of open source AI startups.

Neutral foundations can foster collaboration between global enterprises on open source tools and open standards across the AI stack.

There is a narrow window of opportunity to develop interoperable open source tools for AI agents and prevent lock-in to closed systems.



Contents

As open source AI catches up to closed source AI in performance, open source is poised to win the AI race by offering lower cost, better privacy, more control, and far more flexible customization options. However, the current use of the term "open source" is murky. Many models have only published their weights, which is equivalent to compiled binaries in the software world, while others have restrictive licenses that prohibit certain groups of people from using them. They are all against both the spirit and definition of "open source." This timely report is much needed to clearly define what is "open source" when it comes to AI models, why we need it, and how it can benefit model developers, users, and society at large.

- Michael Yuan, PhD, Founder of GOSIM and WasmEdge

Foreword

The rapid evolution of artificial intelligence (AI), in particular the arrival of open source AI, marks a pivotal moment in technological advancement. This report succinctly captures the insightful discussions and recommendations from leading experts at the 2025 GOSIM Open Source AI Strategy Forum, which underscore the strategic imperative of leveraging open source for promoting research and reproducibility, fairness and safety, innovation and competition, as well as sovereignty in AI.

In our Foundation Models and Generative AI Systems team at Fraunhofer IAIS and in projects like OpenGPT-X, TrustLLM, and Eurolingua, it is our aim to bolster Europe's digital sovereignty by providing powerful, open, and trustworthy AI language models. Our flagship open model is Teuken-7B, which supports all 24 official EU languages and is released under the Apache 2.0 open source license. Beyond Teuken, our team has built large language models from scratch and advanced customized generative AI solutions used across various industries, from enterprise chatbots to knowledge discovery tools. This work, along with the broader efforts in the European open source AI community, aligns perfectly with the themes explored in this report.

Three recommendations particularly resonate with our work and perspective at Fraunhofer IAIS.

First, from a European perspective, open source AI is not just an opportunity; it is a necessity, forming a nucleus for national, European, and international collaborations. In a field demanding highly qualified teams and compute resources to build competitive models, collaboration offers Europe, and indeed other continents such as for example Africa, a viable pathway to accelerate progress and establish a competitive edge. Collaborative efforts, fostered through funded or even grassroots research initiatives, enable the pooling of expertise, resources, and diverse perspectives, which are vital for developing robust and globally impactful AI systems. Such collaborations allow for the creation of open standards, datasets and tools across the entire AI stack, from infrastructure to applications, preventing vendor lock-in and fostering an interoperable technological ecosystem. This collective approach complements the market power of global technology giants, by focusing on digital sovereignty and open innovation.

Second, the release of AI models with permissive licenses, such as Apache 2.0 or MIT, significantly boosts their adoption by industry and small to medium-sized enterprises (SMEs). These licenses provide businesses with the freedom to use, study, modify, and distribute the software without stringent restrictions, lowering entry barriers and fostering a dynamic ecosystem of innovation and commercialization. The flexibility offered by permissive licenses allows companies to integrate AI models into their products and services with confidence, accelerating development cycles and reducing legal complexities. For SMEs, this is particularly beneficial as it provides access to cutting-edge AI technology that might otherwise be prohibitively expensive or complex under proprietary licensing models.

However, to ensure long-term stability and trust, organizations, ideally nonprofit NGOs, that can guarantee the immutability of these model licenses will be essential in the future, as companies currently possess the unilateral power to alter their terms at any time to a more restricted license.

Third, the ability to build models from scratch is crucial for steering AI towards sustainable goals and ethical usage. When developers have complete control and understanding of the underlying architecture, training data, and development processes, they are better equipped to identify and mitigate biases, ensure transparency, and implement robust security measures. This deep foundational knowledge fosters responsible and sustainable AI practices, enabling the creation of systems that align with societal and environmental values and regulatory frameworks, such as the EU AI Act. Open source development, by its very nature, encourages this level of scrutiny and collaboration, allowing a global community to contribute to the ongoing improvement of AI models and their ethical implications.

The report captures the inspiration I gained at the GOSIM meeting in Paris and serves as a valuable roadmap for stakeholders across industry, academia, government, and civil society to navigate the current challenges and collectively support the democratization of AI.

By embracing the principles and recommendations outlined, we can foster a vibrant, collaborative, and ethically sound open source AI ecosystem that benefits everyone.

Dr. Nicolas Flores-Herr
Team Lead Foundation Models & Gen AI Systems,
Fraunhofer IAIS

Executive summary

The GOSIM Open Source AI Strategy Forum, co-organized by the Eclipse Foundation and GOSIM, convened global experts from industry, academia, civil society, and open source communities and foundations in Paris on May 5th, 2025, to discuss critical challenges facing global collaboration in open source artificial intelligence (AI) and pathways forward. This report summarizes the key highlights and takeaways from the forum.

Challenges facing the open source AI ecosystem

The forum participants discussed challenges currently facing the open source AI ecosystem, including the following.

Lack of consensus about "open source AI": The AI community still lacks consensus on what constitutes "open source AI," and in most cases, "open" AI model releases and licensing fall short of upholding the four freedoms of open source (use, study, modify, and redistribute).

Geopolitical and regulatory risks: Digital sovereignty or strategic autonomy agendas and divergent AI regulations risk creating barriers to global collaboration as well as causing the regional fragmentation of the global open source ecosystem.

Economic sustainability of open source AI: The steep costs of AI model development and deployment raise questions about the viability of traditional open source funding and business models in the open source AI context.

Enterprise adoption of open models: Trust, safety, and reliability remain key blockers to enterprise adoption of open models. A gap remains between how well models perform on benchmarks in controlled research settings versus real-world, regulated contexts.

Ethical and responsible AI: Ensuring responsible practice in open source AI represents a major challenge for AI researchers, developers, and policymakers, especially given the speed and scale of AI diffusion. In addition, with the EU AI Act now in effect, open source developers must urgently understand what this new regulatory environment means for them.

Recommendations for the open source AI ecosystem

The forum participants highlighted recommended actions that provide a path forward for the open source AI ecosystem, including the following:

Championing openness: Open source AI developers should champion open development practices and licenses that uphold the open source freedoms in open source AI, enabling the use, study, modification, and sharing of AI models, code, data, and documentation.

Fostering global collaboration in AI: Vendor-neutral foundations play a key role in facilitating collaboration among global developers and enterprises on open source tools and open standards across the entire AI stack, from infrastructure to applications.

The open source AI promise for digital sovereignty: Governments should embrace open source AI as a strategic tool to strengthen digital sovereignty and foster research, innovation, and security in AI while mitigating the fragmentation of the global open source ecosystem.

Enabling reproducibility and research: AI researchers should establish open standards for reproducibility in AI, including evaluation frameworks for the transparency of AI training processes and data provenance. Greater transparency not only democratizes access to AI but also enables research and applications in public interest domains.

Facilitating enterprise adoption of open source AI: It is a priority to expand open source evaluation frameworks and benchmark suites for testing and monitoring model performance and safety for diverse tasks and real-world contexts across regulated industries.

Promoting responsible AI practices: The development and sharing of open source evaluation frameworks, benchmarks, and documentation facilitate the learning, adoption, and innovation of best practices while also providing tools that can help developers comply with regulations.

Overall, the priorities that forum participants discussed provide a roadmap for stakeholders across industry, academia, government, and civil society to navigate current challenges and support the democratization of AI.

A complete list of recommendations from each panel discussion can be found in the Conclusion.

Introduction

It is a common saying that "software is eating the world," a phrase that venture capitalist Marc Andreessen coined in 2011 to describe how software systems are disrupting traditional industries across the global economy.[1] In 2022, upon observing the accelerating growth rate of market capture by open-core companies versus closed-core companies, venture capitalist Joseph Jacks extended Andreessen's metaphor, remarking that "open source is eating software faster than software is eating the world."[2] Nowadays, we may be tempted to make a similar argument about AI, with open source AI solutions catching up with the capabilities of proprietary AI solutions, from large language models (LLMs) to AI agents, and offering an alternative approach to AI development and governance rooted in the principles of open science and innovation.

The open source AI ecosystem has undergone meteoric growth in recent years. There are now over 1.5 million models on the Hugging Face Hub, and some have hundreds of millions of downloads.[3] Those models are steadily catching up with the capabilities of proprietary models. For example, in January, DeepSeek sent shockwaves through the global AI community for releasing models with capabilities on par with OpenAI's models, as well as a detailed technical paper documenting its process for others to learn from and build on. As Matt Asay of MongoDB put it, this was "the Linux moment for AI."[4]

Since then, various labs have released open models, including small models, that have helped close the gap with industry-leading proprietary models.[5] For example, upon releasing the LLM, OLMo 2 32B, along with all training code, data, and documentation, Nathan Lambert from Ai2 commented: "For a long time, people have asked for a truly open source version of ChatGPT, and we finally have it."[6] With these capability gains, enterprises are increasingly embracing open models, finding the lower costs and greater customizability compared to proprietary alternatives particularly attractive.[7]

At the same time, the open source AI ecosystem faces a number of geopolitical, economic, and regulatory challenges. Recent developments, such as "the DeepSeek moment," have raised concerns about global competition in AI, and governments are weighing strategies to strengthen their digital sovereignty and capabilities in AI.

Regulation for AI is also diverging across geopolitical lines, which risks fragmenting the global open source ecosystem. While the EU's AI Act has gone into effect as the world's first regulation for AI, which includes regulatory obligations for open source AI developers, it remains unclear whether other jurisdictions will follow a similar approach. Meanwhile, the copyright battle around AI training data is intensifying, with major AI companies pushing for broad "fair use" interpretations in the USA and rights holders demanding stricter controls.

The question of what truly constitutes "open source" remains a subject of heated debate, and in most cases, the transparency of "open" models remains limited. What is more, while enterprises cite lower costs and customizability as key benefits of open models, concerns about security, risks, and control remain common reasons for opting for proprietary models.[8]

Against this backdrop, the GOSIM Open Source AI Strategy Forum, co-organized by the Eclipse Foundation and GOSIM, convened global experts from industry, academia, civil society, and open source communities and foundations in Paris, France, on May 5th, 2025, to wrestle with challenges facing global collaboration in open source AI and pathways forward. Under the Chatham House rule, forum participants explored the following five topics in dedicated panels.

Global collaboration in open source AI: How can the international community work together on open source AI research and development despite growing geopolitical tensions and regional divergence on AI governance? The panel examined the open source AI definition, regulatory challenges, and the fragmentation risks facing the open source AI ecosystem.

Open source AI and digital sovereignty: How can governments strengthen their digital sovereignty without fragmenting the global open source ecosystem? The panel discussed building AI capabilities and leveraging global open source innovation as a win-win strategy to enhance digital sovereignty and foster research, competition, security, and public value.

Open source AI research and reproducibility: What role does open source play in scientific discovery and reproducible research? The panelists discussed the promise of open source for AI research and the challenges researchers face, from securing access to compute to the difficulties of reproducing AI models due to limited transparency about data or training processes.

Open source AI adoption challenges in enterprise settings: What are the barriers to enterprise adoption of open source AI solutions, and what could facilitate this adoption? The panel examined integration challenges, the transition to agentic systems, and the practical tools and benchmarks enterprises need to evaluate open models and deploy them in production.

Ensuring responsible practices in open source AI: How do we ensure ethical and responsible development practices in open source AI innovation? The conversation spanned regulatory frameworks, human-centric design principles, and the digital commons.

Lightning talks showcased tools for promoting and developing open source AI, from Deutsche Telekom's (DT's) Language Models Operating System (LMOS) platform and Alibaba's state-of-the-art Qwen3 models to the Generative AI Commons' Model Openness Framework and The Linux Foundation's new permissive OpenMDW license for open models.

In this report, you will find the key insights and takeaways from the forum's discussions, with a summary of a list of recommended actions in the conclusion.

Fostering global collaboration in open source AI

The forum began with a panel discussion exploring the challenges facing global collaboration in open source AI and pathways forward. The panelists highlighted how regulatory fragmentation, geopolitical tensions, and debates about the definition of open source AI, among other factors, are creating challenges for collaboration and knowledge sharing. Faced with these challenges, the panelists discussed approaches to facilitating global collaboration in open source AI.

The discussion highlighted the following specific challenges and recommended actions.

Building international consensus on the definition of open source AI

The panelists discussed the challenge of reaching a consensus on the definition of "open source AI." They argued that developing a common understanding of the term is imperative in light of the divergent practices in the openness and licensing of AI models, with most "open" models failing to uphold the four freedoms of open source (i.e., use, study, modify, and redistribute).

The Open Source Initiative's Open Source AI Definition (OSAID) version 1.0[9] was a key reference point in this discussion. The OSAID emphasizes that for an AI system to be considered open source, the system and its constituent parts must be made available under terms that grant individuals the four freedoms of using, studying, modifying, and sharing it without permission. In particular, the OSAID applies the requirement from the Open Source Definition (#2)—"the source code must be the preferred form in which a programmer would modify the program"—to AI systems. To satisfy this requirement, it requires that model parameters (i.e., pre-trained weights and biases), the code for preprocessing training data, the code for training and running the AI system, and the training data itself (if that's not possible, sufficiently detailed information on it), are all released under open source licenses.

The OSAID has received both endorsements and critiques from the open source AI community. The critiques range from "maximalist" views that find it unacceptable to train AI systems or models on data that companies cannot reshare further (e.g., requiring that all training data are open to begin with) to concerns that open-sourcing code and data artifacts is too burdensome (e.g., due to the need to remove proprietary code or the sensitivity of proprietary data). The result: open source AI remains a contested term. The panel underscored that the OSAID is only version 1.0, and, just as it emerged from a global co-design process involving over 100 contributors, it will continue to be developed based on discussions among the global open source AI community. As a panelist said, "It's the OSAID v1.0, and the patches should come from the community."

A second panelist agreed that we need more people working on the open source AI definition and championing the benefits of possibly making the definition more demanding. They underlined that of the four freedoms of open source, the freedoms to use and redistribute are easy to express in a definition, but the freedoms to study and modify are more complex.

The panelist went on to argue that while the public sector is still in the early stages of AI adoption, developers in government bodies may soon find that they have obligations that put them in a similar position of demanding more rather than less openness, as they may need to define what is necessary for them to be able to sufficiently study and modify a model.

The panelist offered two examples. One example is that developers in government bodies should be able to audit AI systems for biases or vulnerabilities and modify them if needed. The second example is that developers in government bodies should be able to produce security updates for the software or models they use. However, this example is more difficult because it would require changing how some governments think about IT security. The panelist argued that "To detect and remove bias and to audit and fix security issues in AI systems, governments need to be able to, in our words, study and modify. Like us, they might be trying to figure out how to define when an AI system can be sufficiently studied and modified. Maybe we can work together, and this could be one path to getting more people working on this complex issue."

Beyond the OSAID, a panelist mentioned the Model Openness Framework (MOF) by the Generative AI Commons as a practical tool for promoting the openness and completeness of AI models. In particular, the MOF identifies 16 code, data, and documentation artifacts from the development lifecycle of machine learning models that are released in an open model distribution and evaluates the completeness and openness of models based on the release of artifacts under appropriate open licenses. One panelist explained that the OSAID and the MOF are highly complementary: the MOF provides a framework for assessing the availability of openly licensed model artifacts, while the OSAID provides a binary yardstick for open source AI.

Geopolitical and regulatory challenges for open source AI collaboration

The panel did not shy away from discussing the geopolitical obstacles to global collaboration in open source AI. The panel acknowledged the risks of export controls on AI models, regulatory fragmentation and compliance burdens, and broader political tensions as significant challenges that require careful consideration and pragmatic solutions.

Panelists voiced concerns regarding the increasing divergence of AI regulations across different jurisdictions and the risk of regulatory fragmentation hindering the sharing of AI models, software, datasets, and knowledge. While the EU's AI Act has entered into force as the world's first comprehensive AI regulation, there is a risk that the compliance work will multiply if we end up with different regional approaches to AI regulation. The need to comply with divergent regulations could even result in software or model distributions that are limited to specific jurisdictions. These risks underline the need for international cooperation on AI regulation. One panelist argued that one of the hardest challenges stems from data, where copyright legislation is creating legal challenges for the open sharing and use of training datasets. They made the case that "Open source is a hack that uses copyright to guarantee the free circulation of software. We need to figure out how we can do something similar for LLMs and data."

When asked if there are any approaches that act as lighthouses for AI regulation, the panel mentioned the EU's AI Act and the Cyber Resilience Act (CRA) as the first examples of AI and software regulation that contain certain exemptions for open source and provide a template for similar regulation in other jurisdictions. A panelist admitted that they may not be perfect as the first of their kind, and legal experts are grappling with novel challenges, such as AI training and copyright debates. Nonetheless, enthusiastic compliance should be our shared goal because if we can show that it can foster innovation and investment in open source AI, we may see similar legislation developed in other jurisdictions, reducing the risk of regulatory fragmentation and the impacts on the global open source ecosystem.

Frameworks for global collaboration in open source AI

The panel turned to discussing global collaboration and governance frameworks in open source AI. One panelist highlighted the difference between software development and AI development, emphasizing that global collaboration in open source AI requires new collaboration frameworks that are suited to the unique characteristics and resource requirements of AI development, including specialized hardware, compute, and data. They added that "several governance principles and initiatives already exist, yet they lag far behind the pace of innovation." Collaboration via open source represents a means of keeping up with the pace of AI innovation, providing open development processes and governance frameworks that can help deal with "unknown unknowns" and manage "known unknowns" as they emerge.

Panelists shared examples of successful international collaboration in open source AI. A panelist explained that open source AI collaborations take place in several areas, including data, algorithms, models, benchmarks, and tools. To foster global collaboration in open source, they highlighted the importance of promoting CORE values: collaboration, openness, respect, and equality among nations. They offered the China–ASEAN AI Open Source Alliance as a successful example. This is a partnership between researchers in China and several Southeast Asian countries collaborating on AI research and development and jointly developing open source tools that each country hosts to ensure no single control point in the alliance.

The panel also highlighted open source collaboration as a means to foster dialogue about state-of-the-art capabilities among the global community and to avoid surprises such as "the DeepSeek moment" in January 2025. In particular, the DeepSeek story revealed different regional perspectives. While Western panelists confessed their surprise at the news, two Chinese panelists suggested the reaction outside of China reflected incomplete visibility into the Chinese AI community. The panelists highlighted how regional or linguistic barriers create mutual blind spots in global AI research and development (R&D), and open source offers not only a means for collaborative development but also for global dialogue.

The economic challenges of open source AI innovation

The panelists discussed the economic challenges in open source AI development for various stakeholder groups, including academic researchers, grassroots initiatives, and enterprises. Focusing on enterprises, a panelist explained that their research indicates that there is an economic tension that needs to be solved to ensure the sustainability of open source AI. While competition in open source AI drives down prices and improves model quality, it also reduces the direct value that enterprises can capture from their open models. The panelist noted that when enterprises open-source their software, another part of the business typically subsidizes it. However, AI development incurs significantly higher costs than traditional software development, and it remains an open question whether open-core business models that work for open source software companies can similarly scale and sustain open source AI enterprises.

Recommended actions:

- Build consensus on the definition of open source AI in the open source AI community through continued community-led dialogue and co-design.
- Promote open development and licensing practices that uphold the four freedoms of open source (use, study, modify, and redistribute) in open source AI.
- Promote the use of permissive licenses for open AI artifacts, such as the OpenMDW license for models, Apache v2 or MIT for code, and CC-BY or ODC-BY for data.
- Champion open science values and practices in the open source AI community with emphasis on the merits of openness for reproducibility, transparency, and usability.
- Educate policymakers about the value of open source for the study and modification of AI models and systems, enabling the auditing and fixing of biases or security issues.
- Foster international dialogue on global alignment on AI regulation to prevent regulatory fragmentation and adverse impacts on the global open source ecosystem.
- Explore funding mechanisms and business models that can scale and sustain open source AI enterprises, accounting for infrastructure and resource requirements in AI.

The promise of open source AI for digital sovereignty and autonomy

The second panel explored the relationship between open source AI innovation and political agendas concerning digital sovereignty or autonomy. The panelists confronted the implications of dependencies in the AI stack and discussed actions that companies should take to strengthen digital sovereignty or autonomy while avoiding the fragmentation of the global open source ecosystem. They underlined open source as a means for governments to simultaneously strengthen their digital sovereignty and foster global collaboration in AI, as well as a means of diplomacy between AI researchers and developers in the changing geopolitical landscape.

The discussion highlighted the following specific challenges and recommended actions.

The state of digital sovereignty today

A panelist stressed a sobering reality: "Now we know that a digital blackout is not a science fiction scenario. If we don't have the capabilities to manage ourselves independently in a blackout scenario, then we will have serious problems." A second panelist argued that Europe fundamentally lacks digital sovereignty and exists in a state of dependency across the AI stack, from chips to the cloud. In Europe, there has been insufficient investment in cloud infrastructure, and it will cost billions and several years to build up a competitive alternative. The panelist suggested we should not reduce "digital sovereignty" to a simple binary, arguing that it is crucial to accept that absolute sovereignty is less important than avoiding absolute dependency and that sovereignty can be understood as a "yin/yang" rather than a binary concept. They went on to argue that for medium-sized economies such as Germany, France, or Singapore, such a balanced approach to digital sovereignty will be necessary, which may combine making investments in AI capabilities and infrastructure while recognizing their interdependence in the global system.

Ways forward: Joint infrastructure investments and "openness as policy"

Given the state of the digital sovereignty of nation-states and regions, the panelists discussed pathways forward for digital sovereignty agendas and the role that open source AI can play. The second panelist suggested that Europe's current position resembles Africa's pre-mobile era, requiring a leapfrog strategy rather than incremental catch-up to become a competitive force in AI and the global technology ecosystem. To achieve this, they argued that major investments in AI capabilities and digital infrastructure were necessary and that governments and enterprises should explore ways to jointly make these investments, citing examples of public–private partnerships such as the coalition between German companies SAP, DT, Siemens, Ionos, and Schwarz Group in building an AI gigafactory with support from the EU.

At the same time, the panelist argued that governments should embrace "openness as policy" to increase their global competitiveness in AI, which entails leveraging five interconnected pillars of open source AI—open science, open standards, open source, open data, and open weights—as strategic tools to level the playing field in AI R&D and commercialization, build an interoperable, open technology ecosystem, and ultimately create a collective counterbalance to the market power of global technology giants. Rather than

closing borders on the one hand or accepting perpetual "tenant" status vis-à-vis technology giants on the other hand, the panelist argued that "openness as policy" strategies offer governments a pragmatic way of strengthening their digital sovereignty precisely through—and not in spite of—open source.

Another panelist warned about the risk of fragmentation of the global open source ecosystem as a result of digital sovereignty policies, such as restrictions on the distribution of software or models, arguing that mitigating the regional fragmentation of the global open source ecosystem should be a common goal of governments. They distinguished between productive interventions, such as regional investment in compute infrastructure, which can strengthen digital sovereignty without fragmenting the open source ecosystem, and counterproductive interventions, such as distribution restrictions, which risk the duplication of efforts and a waste of resources, such as the building of 50 scikit-learns or 50 PyTorchs. They agreed with the second panelist that simultaneous investments in AI capabilities and infrastructure and open source represent a win-win strategy that strengthens both digital sovereignty and global digital collaboration in AI. A third panelist agreed, arguing that "Open source is a new kind of diplomacy in this changing geopolitical regime."

The role of open source in building sovereign AI capabilities

Another panelist introduced OpenLLM Europe, a community-driven initiative with over 1,100 researchers and engineers across Europe who collaborate on the development of open multimodal foundation models that are performant in all 24 official EU languages. OpenLLM Europe focuses on developing "truly open source AI," which involves releasing all the model weights, training code, and data under open licenses, as well as documentation that provides transparency about their development process. They place emphasis on developing models with ethical, legal, and environmental considerations in mind from start to end, from complying with ethical principles and regulatory requirements to focusing on the development of smaller, specialized models (in general, less than 24 billion parameters) that require less computational power—and hence incur less energy consumption—to train and deploy. The panelist added that they design models that can be both trained and deployed on traditional hardware and train models on public research infrastructure; for example, OpenLLM Europe trained its flagship Lucie-7B language model on France's Jean-Z supercomputer.

The same panelist underlined that the trend of building and openly sharing small models represents a particularly promising area for facilitating collaboration in AI and strengthening digital sovereignty. Developers can easily build, customize, and deploy small models with less compute and energy consumption, and the performance of these models is also catching up with that of large models. For example, the panel mentioned that small models with less than 2 billion parameters are already outperforming cutting-edge proprietary models from last year.

The panelists discussed that another aspect of building sovereign open source AI is customizing and evaluating models specifically for regional contexts, values, and languages, recognizing that language models embody cultural values and perspectives. One panelist argued that languages are not just communication tools; they also represent cultural communities. Another panelist agreed, emphasizing that projects focusing on linguistic and cultural diversity serve crucial functions in ensuring that AI is not biased toward any language or culture. They commended researchers and developers who are leading open source initiatives that are collaboratively developing open training and evaluation datasets for low- and medium-resource languages.

Recommended actions:

- Design AI industrial strategies that embrace pro-openness postures that invest in and leverage open source as a way to both collaborate and compete globally in AI.
- Explore ways to jointly invest in AI capabilities and digital infrastructure.
- Fund the development and maintenance of truly open source AI models backed by open data, public compute, and open development and governance processes.
- Fund the development and maintenance of open data for public interest domains, including training and evaluation datasets for low- or medium-resource languages.
- Prioritize the development of small models that are easier to customize for specialized downstream use cases and applications.

Lightning Talk 1: LMOS by Deutsche Telekom

Link: https://eclipse.dev/lmos/docs/introduction/

Eclipse LMOS is an open source, vendor-neutral platform for building and running enterprise-grade AI agent systems at scale on cloud or on-premises. Now hosted by the Eclipse Foundation, LMOS began as a grassroots initiative that a small group of engineers at DT developed while building sales and service agents across DT's European footprint. It was built using existing engineering teams and infrastructure, without requiring massive tooling overhauls. In production today, LMOS powers millions of agent-handled conversations across multiple countries, achieving around 30% human agent handover for action-taking AI agents. The business is now able to manage and evolve agents independently, reflecting a shift toward sustainable, scalable AI operations. LMOS is sovereign and portable by design. It runs on Kubernetes, supports multi-tenancy, integrates with existing systems, and offers model flexibility, enabling organizations to scale AI efforts without being locked into a single ecosystem. DT open-sourced LMOS and donated it to the Eclipse Foundation to support open collaboration and shared progress. LMOS is a living example of how grassroots engineering and open experimentation can lead to tangible results without discarding what enterprises have already built.

Lightning Talk 2: Model Openness Framework and the OpenMDW License by the Linux Foundation

Link to MOF: https://arxiv.org/abs/2403.13784
Link to OpenMDW License: https://github.com/OpenMDW/OpenMDW

To promote openness in AI, the Generative AI Commons of the Linux Foundation developed the MOF in 2024. The MOF is a three-tiered classification system for evaluating the openness and completeness of machine learning models and identifies 16 code, data, and documentation artifacts from the development lifecycle of models and classifies its openness and completeness based on the release of these artifacts under appropriate open licenses. Its three tiers range from open models to open science models. For example, to classify it as an open model, the model weights, model architecture, documentation (e.g., model card, data card, and technical report), and evaluation results must be released. Building on the MOF, in 2025, the Linux Foundation developed the OpenMDW License Agreement v1.0 as the first permissive license designed specifically for machine learning models, covering models (i.e., architecture and weights) as well as code, data, and documentation that may be released with a model. OpenMDW grants rights under copyright, patent, database, and trade secret laws, includes patent litigation protection and attribution, and places no restrictions on model outputs. It is a global license that is compatible with other licenses used for artifacts released in the same distribution or separately (e.g., CC-BY or ODC-BY for open datasets and Apache 2.0 or MIT for open source software).

Lightning Talk 3: Qwen3 by Alibaba

Link: https://www.alibabacloud.com/blog/alibaba-introduces-qwen3-setting-new-benchmark-in-open-source-ai-with-hybrid-reasoning_602192

Alibaba's Qwen3 is a family of state-of-the-art open LLMs that includes six dense models (ranging from 0.6B to 32B parameters) and two Mixture-of-Experts (MoE) models (30B and 235B), all optimized for applications across mobile devices, smart glasses, and robotics, among others. Qwen3 pioneers a hybrid reasoning approach, enabling the model to switch between the "thinking" mode for complex reasoning tasks and the "non-thinking" mode for faster responses. Trained on 36 trillion tokens, Qwen3 models offer cutting-edge performance in multilingual tasks (supporting 119 languages and dialects), reasoning, tool use, and human-aligned dialogue. Qwen3 models also integrate advanced agent capabilities through the Model Context Protocol and function-calling support. Qwen3 models are available under Apache 2.0 on platforms such as the Hugging Face Hub, GitHub, and ModelScope.

Facilitating AI research and reproducibility through open source AI

The third panel addressed the topic of open source AI and research, with a focus on the role of open source in enabling research and reproducibility. The panelists discussed key obstacles to reproducibility in AI research, including limited transparency about AI development processes and data provenance, as well as limited access to research infrastructure. The panelists highlighted how AI researchers are navigating these challenges by organizing grassroots open source collaborations, openly releasing the code and data used to develop or evaluate LLMs, and sharing documentation and lessons learned for others to scrutinize, learn from, and build on.

The discussion highlighted the following specific challenges and recommended actions.

Open source is fundamental to reproducibility in AI research

The panelists discussed the role of openness as a cornerstone of reproducible and collaborative AI research. One panelist highlighted that today, most LLMs and their underlying training processes and data remain closed, which severely limits the possibilities for scientists to independently verify claims about model performance, investigate potential biases, or build upon existing research in a rigorous manner. They argued that the entire pipeline of AI models—from dataset composition to model training to evaluation—must be open to achieve true reproducibility. However, beyond model weights, currently, it is not common for AI researchers to share key artifacts from their model development pipeline, such as the training data or code.

The lack of transparency about LLMs sparked the creation of grassroots initiatives dedicated to developing open LLMs, including Eleuther AI, the BigScience Workshop, and LAION. The panelist explained that these communities aim to make their entire LLM development pipeline transparent and reproducible by publicly sharing documentation (e.g., technical reports and research papers), as well as openly releasing the code and data used to develop the LLMs alongside model checkpoints. In addition, researchers at research institutes such as Ai2 and BAAI have also employed open science practices, sharing training code, training data, and documentation along with model checkpoints under open source licenses.

The panelist explained that this open science approach to AI R&D enables researchers to share knowledge, learn from each other, and verify and build upon previous findings, which fosters greater trust and accelerates scientific progress. Open source AI collaborations lower the entry bar for researchers to participate in AI development, as researchers can pool their respective expertise and resources. They argued that open source also provides quality control and "sanity checks," explaining that, while sometimes harsh, the open source community's rigorous scrutiny improves research quality and prevents unfounded claims.

However, the panelist highlighted that access to the necessary compute resources to train or reproduce LLMs represents a major bottleneck for researchers. In addition, they argued that "open source alone is never sufficient," stressing that building a community and a culture of not only sharing but also collaborating is important.

Openness is key to AI adoption in scientific discovery

The panelists discussed the role of open source AI in scientific research and the challenges associated with the adoption of open models in various disciplines. A panelist shared insights into their research, which finds that while foundation models have the potential to facilitate scientific discovery in fields such as medicine, chemistry, and biology, their black-box nature and lack of openness often deter researchers from applying them in their research. The panelist introduced their "scaling law of mind" framework, which emerged from the use of different open-source AI resources by researchers in various scientific fields. Based on their findings, the panelist urged the research community to establish a virtuous cycle of collecting and sharing real and synthetic research data, cultivating AI skills among scientists, and strategically identifying research problems where the adoption of foundation models can benefit scientific discovery.

Open source approaches to edge and embedded AI

The panel turned to a discussion of the benefits of open source in edge and embedded AI. A panelist presented a framework for fostering more openness across the entire edge AI stack, encompassing not only models and code but also the data collected at the edge and the underlying hardware architectures. The panelists then discussed the challenges and opportunities associated with developing AI for embedded systems. A panelist made a strong case for open source as the optimal approach to embedded AI development, emphasizing that working through open source channels facilitates the most effective collaboration in this specialized field, enabling the sharing of code, tools, and best practices among researchers.

Recommended actions:

- Develop open standards for transparent and reproducible AI research and development based on the TRUE principles (transparent, reproducible, understandable, executable).
- Design public research grants to incentivize or, if possible, require open science practices, including the open release of models, code, data, and documentation.
- Explore ways to fund national or international public research infrastructure that is accessible to researchers, including subsidies or credits for public benefit domains.
- Cultivate AI literacy in scientific research communities via educational materials and certifications on how to use and fine-tune open models for scientific applications.

Blockers and enablers of enterprise adoption of open source AI solutions

The fourth panel unpacked common blockers and enablers of enterprise adoption of open source AI solutions. The panelists highlighted several challenges facing enterprise adoption, including the gap between model performance on benchmarks in controlled research settings and real-world deployments, the difficulties of integrating AI into existing workflows, and the importance of transparency for ensuring reliability, trust, and safety. The panelists shared practical approaches, lessons learned, and key metrics for the successful integration of open models in enterprise solutions. The discussion also addressed the opportunity for the collaborative development of open source tooling and open standards for agentic systems.

The discussion highlighted the following specific challenges and recommended actions.

The challenges of enterprise adoption of open source AI solutions

The panelists pointed out that a significant gap exists between the performance of models on benchmarks achieved in controlled research settings and their actual readiness to be deployed in production. One panelist pointed out that while open source AI developers often focus on performance improvements on open benchmarks and model leaderboard rankings, enterprises face different challenges when attempting to implement models in real-world production environments. They argued that this gap represents one of the most significant barriers to the widespread adoption of open models by enterprises, as they need to be able to confidently satisfy trust and safety standards and regulatory requirements before adoption.

The panelists agreed that building trust and ensuring the reliability of AI solutions remains a crucial challenge to enterprise adoption of open models. They discussed various strategies for achieving trust, including rigorous testing and validation of models, increasing transparency about model development processes, and implementing robust monitoring and feedback mechanisms. Another panelist made the case that organizations require clear, quantifiable cost/benefit analysis tools and rapid comparison capabilities to make informed decisions. They explained that in their enterprise, they employ a pragmatic approach by prioritizing two key metrics: cost and speed to production. They added that having an integration stack ready for enterprise use of AI models and tools to easily evaluate and compare the accuracy and cost of open models “at the click of a button” are key to their adoption in enterprise settings.

The panelists agreed that the successful adoption of AI solutions, regardless of whether they are open source or closed source models, hinges on the ease of integration into existing workflows, as well as the quality of the user experience. When considering how to build AI solutions that customers actually trust, a panelist argued that earning customer confidence requires not just building performant technical capabilities but also ensuring integration quality and an excellent user experience. They argued that we should not forget that AI technologies are tools that we use to create value-adding services or products. Another panelist highlighted the importance of the transparency of model integrations in earning customer trust, citing the example of Perplexity, a B2C company that had to demonstrate trustworthy integration practices when using DeepSeek models. This included de-censoring the models of political bias and transparently explaining to its customers how they were integrating the models into its product.

Another panelist expressed their concern about the financial sustainability of the current trajectory of AI R&D, noting that the costs of training and deploying large AI models are not financially viable for many enterprises.

The role of open source in building agentic systems

Looking beyond models, the panel discussed AI agents as a new trend that is creating opportunities for both innovation and differentiation. A panelist argued that enterprises across the world should actively collaborate on developing interoperable open source tools and open standards for AI agents before the ecosystem gets locked into proprietary systems. They highlighted the example of Eclipse LMOS, which is an open source, vendor-neutral platform for building and running enterprise-grade AI agent systems at scale on cloud or on-premises. They argued, "We have a narrow window to build sovereign, open source AI agent platforms. LMOS was born to prove that open agentic systems can unlock business value while mitigating long-term risks of enclosure." The core considerations for open source AI agent platforms should center on intuitive ease of use, scalability, integration capabilities, customizability, and enterprise-grade security measures.

Open source foundations can foster the development of open source AI solutions

As an alternative to the status quo, where a small handful of technology giants dominate the AI stack, a panelist made the case that open source foundations can play a crucial role in fostering the collaborative development of open source AI solutions, with their open governance enabling fair participation by global enterprises, large and small, in open source AI innovation.

Recommended actions:

- Establish open standards for AI model transparency, including documentation requirements for training data, development processes, and evaluations, to enable enterprises to conduct due diligence and make informed adoption decisions.
- Champion the use of permissive licenses that allow for the commercial use and modification of open models without restrictions.
- Develop open source evaluation frameworks and industry-specific benchmark suites that can be used to test and monitor the performance of open models for specific capabilities or compliance requirements across regulated sectors and industries.
- Leverage vendor-neutral foundations to facilitate the collaborative development of open source tools and open standards across the AI stack.
- Prioritize the development of interoperable, open source tools that empower grassroots engineering, encourage experimentation, and prevent lock-in to closed AI systems.

Promoting responsible development practices in open source AI

The final panel concerned the topic of responsible AI, addressing key challenges in fostering responsible practices in the open source AI community amidst the unprecedented pace of AI innovation and diffusion. The panelists discussed various ways to promote responsible AI practices, from ethical principles and regulatory awareness to open source projects that demonstrate and document best practices that others can learn from and build on.

The discussion highlighted the following specific challenges and recommended actions.

Translating ethical principles into practice

The panel grappled with the challenge of translating ethical principles into practice in open source AI research and development. A panelist highlighted that this is no trivial challenge, especially in light of the unprecedented pace of innovation and diffusion, noting that it took OpenAI just 2.5 months to reach one million users of ChatGPT. They highlighted that an advantage of open source is that it increases transparency about AI development processes, and the openness of AI technologies, such as software and models, enables distributed auditing for potential biases and vulnerabilities, in turn enhancing trust and safety in these technologies. For this reason, they argued, it is vital to develop, promote, and incentivize openness practices in AI.

The role of regulation in fostering responsible open source AI development

The panel discussed the role of regulation in fostering responsible AI innovation. The panelists mentioned how regulation and innovation are not in opposition and that the goal of regulation should be to shape safe and responsible innovation. One panelist highlighted that regulation plays a crucial role in areas where market players would not naturally invest resources in safety or security. Such interventions ensure that public interest considerations are incorporated into AI development, even when they might not align with immediate commercial incentives. Another panelist mentioned the EU's AI Act, which creates obligations for providers of general-purpose AI models—including copyright compliance policies and training data summaries—that aim to enhance trust and safety while providing open source exemptions that seek to reduce the compliance burden on open source AI developers.

Building open source tools and open benchmarks for responsible AI

The panel discussed open source evaluation frameworks and open benchmarks as practical tools to facilitate responsible and trustworthy AI development and adoption. One panelist provided an overview of the European Trustworthy AI Foundation, which, under the motto of "Industrializing Trustworthy AI," is building open source methodologies and tools that enable the development and adoption of trustworthy AI systems in industry. These tools draw on the outcomes of France's "Confiance IA" (confidence in AI) research program, and

are being built by the foundation's multi-stakeholder community of manufacturers, researchers, startups, and research labs. Beyond them, open source evaluation frameworks such as LM Evaluation Harness by EleutherAI, lighteval by Hugging Face, Project Moonshot by the AI Verify Foundation, and Inspect by the UK AI Security Institute, as well as a growing ecosystem of open benchmarks for various research and industry-specific applications, provide researchers and developers with essential tools for systematically evaluating the performance and safety of AI models. As a panelist in the previous panel mentioned, maintaining and expanding these open source tools so that they cover a greater number of real-world tasks and contexts will be key to enabling enterprises to adopt open models in regulated industries in a safe and trustworthy manner.

Demonstrate and document: Leading by example in the open source community

The panel went on to discuss how open source developers can lead by example by documenting their responsible practices in a way that enables others to learn from and build on.

A panelist presented the efforts of Software Heritage, a non-profit foundation that maintains the largest archive of open source code, to promote the responsible use of code hosted in its archive as training data for LLMs. In October 2023, Software Heritage published its "Software Heritage Statement on LLMs for Code," which outlines three principles that developers must adhere to if they train LLMs on code in the Software Heritage archive. The principles require that 1) models are openly licensed and shared together with the documentation and tooling needed to use them; 2) training data are fully identifiable through SWHID identifiers to enable bias studies, verify if any code of interest was included in the training data, and provide appropriate attribution when generated code resembles the training data; and 3) mechanisms be established for authors to opt out their code from training datasets.

Since the publication of the statement, Software Heritage's principles have influenced the development of open LLMs for code. For example, StarCoder2, an open LLM for code developed by the BigCode project involving Hugging Face, ServiceNow, and Nvidia, was the first LLM for code that a company developed in accordance with these principles. However, the BigCode project did not stop there. The BigCode developers released a research paper that documented how they filtered Software Heritage data in line with the principles, as well as a filtered dataset called "The Stack v2," for others to learn from, use, and build on. This example illustrates how open source AI developers can effectively influence responsible AI practices in the wider ecosystem by documenting their practical implementation of ethical principles and openly releasing their ethically developed training datasets and models.

Recommended actions:

- Promote openness and transparency as a means to enhance security by enabling global researchers and developers to conduct the distributed auditing of models, code, and data.
- Create free educational resources and certification programs on responsible AI practices and regulatory compliance for open source AI researchers and developers.
- Develop open source evaluation frameworks and benchmark suites in order to lower the barriers to safe AI development and regulatory compliance.
- Lead by example #1: By documenting responsible practices, open source developers can provide templates for others to follow, learn from, and build on.
- Lead by example #2: By releasing ethically developed datasets and models, open source developers enable others to adopt and advance responsible AI practices.

Conclusion

The GOSIM Open Source AI Strategy Forum highlighted both the challenges of and opportunities for global collaboration in open source AI. While open models are rapidly catching up to proprietary alternatives in performance and offer significant advantages in cost, customization, and transparency, the ecosystem faces critical barriers, including divergent standards and practices of openness, regulatory fragmentation across jurisdictions, and gaps between research benchmarks and real-world enterprise deployment needs.

The forum's discussions revealed opportunities and priorities for various stakeholder groups; for example, for governments to strengthen digital sovereignty through pro-openness strategies, for enterprises to collaboratively develop interoperable open source frameworks and open standards across the AI stack to prevent lock-in to proprietary solutions, and for open source developers to lead by example in championing responsible AI development practices.

The panel discussions made the following recommendations for stakeholders across industries, academia, governments, civil societies, and open source foundations and communities to navigate these challenges and collectively advance the democratization of AI.

Fostering global collaboration in open source AI

- Build consensus on the definition of "open source AI" in the open source AI community through continued community-led dialogue and co-design.
- Promote open development and licensing practices that uphold the four freedoms of open source (use, study, modify, and redistribute) in open source AI.
- Promote the use of permissive licenses for open AI artifacts, such as the OpenMDW license for models, Apache v2 or MIT for code, and CC-BY or ODC-BY for data.
- Champion open science values and practices in the open source AI community with an emphasis on the merits of openness for reproducibility, transparency, and usability.
- Educate policymakers about the value of open source for the study and modification of AI models and systems, enabling the auditing and fixing of biases or security issues.
- Foster international dialogue on global alignment on AI regulation to prevent regulatory fragmentation and adverse impacts on the global open source ecosystem.
- Explore funding mechanisms and business models that can scale and sustain open source AI enterprises, accounting for infrastructure and resource requirements in AI.

The promise of open source AI for digital sovereignty and autonomy

- Design AI industrial strategies that embrace pro-openness postures that invest in and leverage open source to both collaborate and compete globally in AI.
- Explore ways to jointly invest in AI capabilities and digital infrastructure.
- Fund the development and maintenance of truly open source AI models backed by open data, public compute, and open development and governance processes.

- Fund the development and maintenance of open data for public interest domains, including training and evaluation datasets for low- or medium-resource languages.
- Prioritize the development of small models that developers can more easily customize for specialized downstream use cases and applications.

Facilitating research and reproducibility through open source AI

- Develop open standards for transparent and reproducible AI research and development based on the TRUE principles (transparent, reproducible, understandable, executable)
- Design public research grants to incentivize or, if possible, require open science. practices, including the open release of models, code, data, and documentation.
- Explore ways to fund national or international public research infrastructure that is accessible to researchers, including subsidies or credits for public benefit domains.
- Cultivate AI literacy in scientific research communities via educational materials and certifications on how to use and fine-tune open models for scientific applications.

Blockers and enablers of enterprise adoption of open source AI solutions

- Establish open standards for AI model transparency, including documentation requirements for training data, development processes, and evaluations, which enable enterprises to conduct due diligence and make informed adoption decisions.
- Champion the use of permissive licenses that allow for the commercial use and modification of open models without restrictions.
- Develop open source evaluation frameworks and industry-specific benchmark suites to test and monitor the performance of open models for specific capabilities or compliance requirements across regulated sectors and industries.
- Leverage vendor-neutral foundations to facilitate the collaborative development of open source tools and open standards across the AI stack.
- Prioritize the development of interoperable, open source tools that empower grassroots engineering, encourage experimentation, and prevent lock-in to closed AI systems.

Promoting responsible development practices in open source AI

- Promote openness and transparency as a means to enhance security by enabling global researchers and developers to conduct the distributed auditing of models, code, and data.
- Create free educational resources and certification programs on responsible AI practices and regulatory compliance for open source AI researchers and developers.
- Develop open source evaluation frameworks and benchmark suites in order to lower the barriers to safe AI development and regulatory compliance.
- Lead by example #1: By documenting responsible practices, open source developers can provide templates for others to follow, learn from, and build on.
- Lead by example #2: By releasing ethically developed datasets and models, open source developers enable others to adopt and advance responsible AI practices.

Acknowledgments

The author would like to thank GOSIM and Eclipse Foundation for co-organizing the GOSIM Open Source AI Strategy Forum, as well as all the supporting organizations, including the Open Source Initiative, Generative AI Commons, Software Heritage, and OpenWallet Foundation, and all forum participants and volunteers. The author would also like to thank report contributors Chris Xie, Anni Lai, Gaël Blondelle, Michael Yuan, Stefano Maffuli, Ciarán O'Riordan, Arun Joseph, Roberto Di Cosmo, Xavier Tao, Yumei He, Hilary Carter, Anna Hermansen, and Paula Grzegorzewska. The author would also like to thank the Linux Foundation Creative Services for producing the report.

About the author

Cailean Osborne, PhD is a Senior Researcher at the Linux Foundation, where he conducts strategic research and advocacy for promoting openness in AI. He has a PhD in Social Data Science from the University of Oxford, where he wrote his thesis on collaboration dynamics in the open source AI ecosystem. During his PhD, he was a visiting researcher at the Open Source Software Data Analytics Lab at Peking University. Previously, he was the International Policy Lead at the UK government's Centre for Data Ethics and Innovation, where he co-authored the UK's AI Strategy and served as a UK delegate in international AI governance initiatives at the OECD and Council of Europe, among others. He is based in Berlin, Germany and in his free time he likes to learn languages.

About GOSIM

This event took place as a co-located event of GOSIM AI Paris 2025. GOSIM is a volunteer-based endeavor organized by and for the global open source community with the objective of providing a stage where innovative open source projects can shine, collaborate, and evolve. To learn more about future GOSIM events, please go to **www.gosim.org**.

Endnotes

1 https://www.wsj.com/articles/SB10001424053111903480904576512250915629460

2 https://www.coss.community/cossc/open-source-is-eating-software-faster-than-software-is-eating-the-world-3b01

3 https://huggingface-open-source-ai-year-in-review-2024.static.hf.space/index.html

4 https://www.infoworld.com/article/3960764/deepseeks-open-source-movement.html

5 https://huggingface.co/blog/evijit/smollm-deepseek-bias-eval

6 https://www.linkedin.com/posts/natolambert_a-very-exciting-day-for-open-source-ai-we-activity-7306014817247539201-Jmy5/?utm_source=share&utm_medium=member_desktop&rcm=ACoAABRNgu0BbMNu7gnqkFd65chSBUwqigx3Ktc

7 https://www.linuxfoundation.org/research/economic-impacts-of-open-source-ai?hsLang=en

8 https://www.mckinsey.com/capabilities/mckinsey-digital/our-insights/tech-forward/open-source-in-the-age-of-ai

9 https://opensource.org/ai/open-source-ai-definition

Sponsored by FUTUREWEI Technologies

THE LINUX FOUNDATION | Research

Founded in 2021, **Linux Foundation Research** explores the growing scale of open source collaboration, providing insight into emerging technology trends, best practices, and the global impact of open source projects. Through leveraging project databases and networks and a commitment to best practices in quantitative and qualitative methodologies, Linux Foundation Research is creating the go-to library for open source insights for the benefit of organizations the world over.

x.com/linuxfoundation

facebook.com/TheLinuxFoundation

linkedin.com/company/the-linux-foundation

youtube.com/user/TheLinuxFoundation

github.com/LF-Engineering

To reference this work, please cite as follows: Cailean Osborne, "Charting Strategic Directions for Global Collaboration in Open Source AI: Key Takeaways from the GOSIM Open Source AI Strategy Forum," foreword by Nicolas Flores-Herr, PhD, Fraunhofer IAIS, The Linux Foundation, July 2025.