

オープンソースAIにおける グローバルな協業のための 戦略的方向性の策定

GOSIM Open Source AI Strategy Forum
の主要なポイント

Cailean Osborne, PhD, *The Linux Foundation*

序文

Nicolas Flores-Herr, PhD, *Fraunhofer IAIS*

2025年7月

Sponsored by



オープンソース AI におけるグローバルなコラボレーションの戦略的方向性の策定

開発者は、オープンソース AI においてオープンソースの4つの自由を尊重するオープンな開発実践とライセンスを推進すべきです。



政府は、AI分野におけるグローバルな協力と競争を促進するための戦略的ツールとして、オープンソースAIを積極的に採用すべきです。



小型モデルは、大型モデルに匹敵する性能を発揮しつつ、手頃な価格、カスタマイズ性、およびプライバシーの利点を提供しています。



AIの研究開発におけるオープンサイエンスのアプローチは、再現性、知識の共有、学習、および潜在的なバイアスやセキュリティ脆弱性の検証を促進します。

オープンソースAIの成功は、モデル、コード、データセットの共有だけでなく、活気あるグローバルコミュニティの構築と維持にも支えられています。



中規模および低リソース言語向けのオープンデータセットの開発は、LLMにおける文化的・言語的バイアスに対処するための重要な優先事項です。

オープンソースの開発者は、他の人が使用、学習、および発展させることができるツールやドキュメントを共有することで、責任あるAIの実践に影響を与えています。



多様な機能と環境に対応したオープンソースの評価フレームワークとベンチマークは、オープンソースAIの信頼性と安全性を向上させるための鍵となります。

公共のコンピューティングインフラへのアクセスは、オープンソース AI への研究者や草の根の取り組みの参加における大きな障害となっています。



オープンソースAIスタートアップの成長と持続可能性を支援するためには、革新的な資金調達メカニズムとビジネスモデルが不可欠です。



中立的なFoundationは、AIスタック全体のオープンソースツールやオープンスタンダードに関するグローバル企業間のコラボレーションを促進することができます。



AI エージェント用の相互運用可能なオープンソースツールを開発し、クローズドシステムへのロックインを防ぐ可能性は、限られたチャンスしか残されていません。



目次

序文.....	4
概要.....	6
はじめに	8
オープンソース AI におけるグローバルなコラボレーションの促進	10
オープンソース AI がデジタル主権と自律性にもたらす可能性.....	14
オープンソース AI を通じて AI 研究と再現性を推進する	20
企業におけるオープンソース AI ソリューションの採用を阻害する要因と促進する要因.....	22
オープンソース AI における責任ある開発手法の推進	24
結論.....	27
謝辞.....	29
著者について	29
GOSIM について	29
巻末脚注.....	30



オープンソース AI がクローズドソース AI の性能に追いつくにつれ、オープンソースは低コスト、より良いプライバシー、より高いコントロール性、そしてはるかに柔軟なカスタマイズオプションを提供することで、AI 競争で優位に立つ立場にあります。しかし、現在の「オープンソース」という用語の使い方は曖昧です。多くのモデルは、ソフトウェアの世界でコンパイル済みバイナリに相当する重みのみを公開している一方で、特定のグループによる使用を禁止する制限的なライセンスを課しているモデルもあります。これらはすべて、「オープンソース」の精神と定義に反しています。AI モデルにおける「オープンソース」とは何なのか、なぜそれが必要なのか、そしてモデル開発者、ユーザー、そして社会全体にどのようなメリットをもたらすのかを明確に定義する、このタイムリーなレポートはまさに必要とされているものです。

- Michael Yuan, PhD, Founder of GOSIM and WasmEdge

序文

人工知能 (AI) の急速な進化、特にオープンソース AI の登場は、技術の進歩における重要な転換点となっています。このレポートは、2025 GOSIM Open Source AI Strategy Forum で、AI の研究と再現性、公平性、安全性、イノベーション、競争、そして主権を促進するためにオープンソースを活用することの戦略的必要性を強調した、一流の専門家たちによる洞察に満ちた議論と提言を簡潔にまとめたものです。

Fraunhofer IAIS の Foundation Models and Generative AI Systems チーム、および OpenGPT-X、TrustLLM、Eurolingua などのプロジェクトでは、パワフルでオープン、かつ信頼性の高い AI 言語モデル

を提供することで、ヨーロッパのデジタル主権の強化を目指しています。当社の主力オープンモデルは Teuken-7B で、EU の 24 言語すべてに対応し、Apache 2.0 オープンソースライセンスでリリースされています。Teuken 以外にも、当社のチームは、大規模な言語モデルをゼロから構築し、エンタープライズチャットボットから知識発見ツールまで、さまざまな業界で使用される高度なカスタマイズされた生成型 AI ソリューションを開発しています。この取り組みは、欧州のオープンソース AI コミュニティにおける幅広い取り組みと相まって、本レポートで取り上げているテーマと完全に一致しています。

Fraunhofer IAIS の活動と視点において、特に共感を呼ぶ3つの提言があります。

まず、ヨーロッパの観点からは、オープンソース AI は単なる可能性ではなく、国家、ヨーロッパ、そして国際的なコラボレーションの核となる必要不可欠な要素です。競争力のあるモデルを構築するには、高度なスキルを持つチームと膨大なコンピューティングリソースが求められる分野において、コラボレーションはヨーロッパ、そしてアフリカなどの他の大陸にも、進歩を加速し、競争優位性を確立するための現実的な道筋を提供します。資金援助や草の根の研究取り組みを通じて促進されるコラボレーションにより、堅牢で世界的に影響のある AI システムの開発に欠かせない専門知識、リソース、多様な視点の共有が可能になります。このようなコラボレーションにより、インフラストラクチャからアプリケーションに至る AI スタック全体にわたるオープンスタンダード、データセット、ツールが作成され、ベンダーロックインを防ぎ、相互運用可能なテクノロジーエコシステムが育まれます。この連携アプローチは、デジタル主権とオープンイノベーションに焦点を当てることで、グローバルなテクノロジー大手企業の市場力を補完するものです。

第二に、Apache 2.0 や MIT などの寛容なライセンスで AI モデルがリリースされることで、産業界や中小企業 (SME) による AI の採用が大幅に促進されます。これらのライセンスは、企業に厳しい制限なしにソフトウェアの使用、調査、変更、配布の自由を提供し、参入障壁を下げ、イノベーションと商用化のダイナミックなエコシステムを促進します。寛容なライセンスが提供する柔軟性により、企業は AI モデルを自社製品やサービスに安心して統合することができ、開発サイクルを加速し、法的複雑さを軽減することができます。これは、プロプライエタリなライセンスモデルでは高すぎる、あるいは複雑すぎて導入できない最先端の AI テクノロジーへのアクセスを可能にするため、中小企業にとって特に

有益です。しかし、長期的な安定性と信頼を確保するためには、これらのモデルライセンスの不変性を保証できる組織、理想的には非営利の NGO が将来的に不可欠となるでしょう。なぜなら、現在、企業はいつでも一方的にライセンス条件をより制限的なものに変更する権限を独占しているからです。

第三に、AI を持続可能な目標と倫理的な利用の方向へ導くためには、モデルをゼロから構築する能力が不可欠です。開発者が基盤となるアーキテクチャ、トレーニングデータ、開発プロセスに対して完全な制御と理解を持っている場合、バイアスを特定し軽減する能力、透明性を確保する能力、堅牢なセキュリティ対策を実装する能力が向上します。この深い基礎知識は、責任ある持続可能な AI の実践を促進し、社会や環境の価値観、EU AI 法などの規制の枠組みに沿ったシステムの構築を可能にします。オープンソース開発は、その性質上、このレベルの精査とコラボレーションを促進し、グローバルコミュニティが AI モデルとその倫理的影響の継続的な改善に貢献することを可能にします。

このレポートは、パリで開催された GOSIM 会議で私が得たインスピレーションをまとめたものであり、産業界、学界、政府、市民社会のステークホルダーが現在の課題に対処し、AI の民主化を共同で支援するための貴重なロードマップとなるものです。

ここで概説した原則と推奨事項を実践することで、誰もが恩恵を受ける活気ある、協働的で倫理的に健全なオープンソース AI エコシステムを築くことができます。

Dr. Nicolas Flores-Herr

*Team Lead Foundation Models & Gen AI Systems,
Fraunhofer IAIS*

概要

Eclipse Foundation と GOSIM が共催した GOSIM Open Source AI Strategy Forum は、2025 年 5 月 5 日、パリで、産業界、学界、市民社会、オープンソースコミュニティ、およびオープンソース Foundation のグローバルな専門家たちを集め、オープンソース人工知能 (AI) におけるグローバルな連携が直面する重要な課題と今後の道筋について議論しました。このレポートは、フォーラムの主なハイライトと要点をまとめたものです。

オープンソース AI エコシステムが直面する課題

フォーラムの参加者は、オープンソース AI エコシステムが直面している現在の課題について議論し、以下の点を含む内容を共有しました。

「オープンソース AI」に関するコンセンサスの欠如: AI コミュニティでは、「オープンソース AI」の定義について依然としてコンセンサスが形成されておらず、ほとんどの場合、「オープン」な AI モデルのリリースやライセンスは、オープンソースの 4 つの自由 (使用、研究、変更、再配布) を十分に満たしていません。

地政学的・規制上のリスク: デジタル主権や戦略的自律性に関するアジェンダと、AI 規制の相違は、グローバルな協力を阻害する障壁を生み出すだけでなく、グローバルなオープンソースエコシステムの地域的な分断を引き起こすリスクがあります。

オープンソース AI の経済的持続可能性: AI モデルの開発とデプロイメントの膨大なコストは、オープンソース AI の文脈において伝統的なオープンソースの資金調達モデルやビジネスモデルの持続可能性に疑問を投げかけています。

企業におけるオープンモデルの採用: 信頼性、安全性、信頼性は、企業におけるオープンモデルの採用における主要な障害となっています。制御された研究環境でのベンチマークでのモデル性能と、現実世界の規制された環境での性能との間には依然としてギャップが存在します。

倫理的で責任ある AI: オープンソース AI における責任ある実践の確保は、AI の普及のスピードと規模を考えると、AI 研究者、開発者、政策立案者にとって大きな課題となっています。さらに、EU AI 法が施行されたことで、オープンソース開発者は、この新しい規制環境が自分たちにどのような意味を持つかを早急に理解する必要があります。

オープンソース AI エコシステムに関する提言

フォーラムの参加者は、オープンソース AI エコシステムの今後の道筋を示す、以下の推奨されるアクションを提示しました。

オープン性を推進する: オープンソース AI 開発者は、オープンソース AI におけるオープンソースの自由を擁護するオープンな開発手法およびライセンスを推進し、AI モデル、コード、データ、およびドキュメントの使用、調査、変更、および共有を可能にする必要があります。

AI におけるグローバルなコラボレーションの促進: ベンダー中立の Foundation は、インフラストラクチャからアプリケーションに至る AI スタック全体におけるオープンソースツールやオープンスタンダードに関する、グローバルな開発者や企業間の連携を促進する上で重要な役割を果たしています。

オープンソース AI がデジタル主権に与える可能性: 政府は、デジタル主権を強化し、AI 分野における研究、イノベーション、およびセキュリティを促進しつつ、グローバルなオープンソースエコシステムの分断を緩和するため、オープンソース AI を戦略的ツールとして採用すべきです。

再現性と研究の実現：AI 研究者は、AI の再現性に関するオープンスタンダードを確立する必要があります。これには、AI トレーニングプロセスとデータの出所の透明性を確保するための評価フレームワークも含まれます。透明性が向上することで、AI へのアクセスが民主化されるだけでなく、公共の利益となる分野での研究や応用が可能になります。

オープンソース AI の企業導入を促進する：規制産業を含む多様なタスクと現実世界のコンテキストにおいて、モデル性能と安全性を評価・監視するためのオープンソース評価フレームワークとベンチマークスイートの拡大は、最優先課題です。

責任ある AI のプラクティスの推進：オープンソースの評価フレームワーク、ベンチマーク、およびドキュメントの開発と共有により、ベストプラクティスの学習、採用、およびイノベーションが促進されると同時に、開発者が規制を順守するのに役立つツールも提供されます。

全体として、フォーラムの参加者が議論した優先事項は、産業界、学界、政府、市民社会のステークホルダーが現在の課題に対処し、AI の民主化を支援するためのロードマップとなっています。

各パネルディスカッションからの提言の完全なリストは、結論のセクションに記載されています。



はじめに

「ソフトウェアが世界を飲み込んでいる」という表現は、ベンチャーキャピタリストの Marc Andreessen が 2011 年に考案した言葉で、ソフトウェアシステムが世界経済の伝統的な産業を破壊している状況を説明するために用いられています。¹ 2022 年、オープンコア企業とクローズドコア企業の市場シェア拡大率の加速を観察したベンチャーキャピタリストの Joseph Jacks は、Andreessen の比喩を拡張し、「オープンソースはソフトウェアが世界を飲み込むよりも速く飲み込んでいる」と指摘しました。² 今日では、オープンソースの AI ソリューションが、大規模言語モデル (LLM) から AI エージェントに至るまで、プロプライエタリな AI ソリューションの機能に追いつき、オープンサイエンスとイノベーションの原則に根ざした AI 開発とガバナンスの代替アプローチを提供していることから、AI についても同様の主張をしなくなるかもしれません。

オープンソースの AI エコシステムは、近年驚異的な成長を遂げています。Hugging Face Hub には現在 150 万を超えるモデルが登録されており、そのうち一部は数億回のダウンロードを記録しています。³ これらのモデルは、プロプライエタリなモデルに急速に追いつきつつあります。例えば、1 月に DeepSeek は、OpenAI のモデルと同等の性能を持つモデルをリリースし、その開発プロセスを詳細に記録した技術論文を公開することで、世界中の AI コミュニティに衝撃を与えました。MongoDB の Matt Asay は、これを「AI の Linux モーメント」と表現しました。⁴

その後、さまざまな研究機関がオープンソースモデルをリリースし、その中には小型モデルも含まれており、業界をリードするプロプライエタリモデルとの性能差を縮めるのに貢献してきました。⁵ 例えば、AI2 の Nathan Lambert は、LLM 「OLMo 2 32B」と共に、すべてのトレーニングコード、データ、ドキュメントをリリースした際、次のようにコメントしました：「長年、ChatGPT の真のオープンソース版が求められて

きましたが、ついにそれが実現しました。」⁶ これらの機能の向上により、企業はオープンモデルを採用する動きが加速しており、プロプライエタリな代替案と比較してコストの低さと高いカスタマイズ性が特に魅力的だと評価しています。⁷

同時に、オープンソースの AI エコシステムは、地政学的、経済的、規制上の課題に直面しています。最近の動向、例えば「DeepSeek 事件」は、AI 分野におけるグローバルな競争に関する懸念を浮き彫りにしており、各国政府はデジタル主権と AI 分野における能力強化に向けた戦略を検討しています。

AI の規制は地政学的境界線に沿って分岐しており、グローバルなオープンソースエコシステムの分断リスクが高まっています。EU の AI 法は世界初の AI 規制として施行され、オープンソース AI 開発者に対する規制義務を盛り込んでいますが、他の地域が同様のアプローチを採用するかどうかは不明です。一方、AI トレーニングデータに関する著作権をめぐる戦いは激化しており、米国の大手 AI 企業は「フェアユース」の広範な解釈を主張し、権利者はより厳格な規制を求めています。

「オープンソース」の真の定義は依然として激しい議論的となっており、多くの場合、「オープン」モデルの透明性は限定的なままです。さらに、企業はオープンモデルの主要な利点としてコストの低さとカスタマイズ性を挙げていますが、セキュリティ、リスク、および管理に関する懸念は、プロプライエタリモデルを選択する主な理由として依然として一般的です。⁸

このような背景から、Eclipse Foundation と GOSIM が共催する GOSIM Open Source AI Strategy Forum が、2025 年 5 月 5 日、フランスのパリで開催され、産業界、学界、市民社会、オープンソースコミュニティ、Foundation のグローバルな専門家たちが一堂に会し、オープンソース AI におけるグローバルな連携が直面する課題と今後の道筋について議論しました。チャタムハウスルールに基づき、フォーラムの参加者は、専用のパネルで以下の 5 つのトピックについて検討しました。

オープンソース AI におけるグローバルなコラボレーション: 地政学的緊張の高まりと AI ガバナンスに関する地域間の格差が拡大する中、国際社会はオープンソース AI の研究開発においてどのように連携できるのでしょうか? パネルでは、オープンソース AI の定義、規制上の課題、オープンソース AI エコシステムが直面する断片化のリスクについて議論しました。

オープンソース AI とデジタル主権: 各国政府は、グローバルなオープンソースのエコシステムを分断することなく、デジタル主権を強化するにはどうすればよいか? パネルでは、デジタル主権を強化し、研究、競争、セキュリティ、公共の価値を促進するための、双方にメリットのある戦略として、AI 機能の構築とグローバルなオープンソースのイノベーションの活用について議論しました。

オープンソース AI 研究と再現可能性: オープンソースは科学的発見と再現可能な研究においてどのような役割を果たすのか? パネリストたちは、AI 研究におけるオープンソースの潜在的な可能性と、研究者が直面する課題について議論しました。具体的には、計算リソースへのアクセ

ス確保から、データやトレーニングプロセスの透明性不足による AI モデルの再現困難さまで、多岐にわたる課題が挙げられました。

企業におけるオープンソース AI の導入課題: 企業によるオープンソース AI ソリューションの導入を妨げる要因とは何であり、その導入を促進するには何が必要か? パネルでは、統合の課題、エージェント型システムへの移行、およびオープンモデルを評価し、本番環境に導入するために企業が必要とする実用的なツールやベンチマークについて検討しました。

オープンソース AI における責任ある取り組みの実現: オープンソース AI のイノベーションにおいて、倫理的で責任ある取り組みをどのように確保すべきか? この会話では、規制の枠組み、人間中心の設計原則、デジタルコモンズについて話し合われました。

ライトニングトークでは、ドイツテレコム (DT) の Language Models Operating System (LMOS) プラットフォーム、Alibaba の最先端の Qwen3 モデル、Generative AI Commons の Model Openness Framework、Linux Foundation のオープンモデル向けの新しい寛容な OpenMDW ライセンスなど、オープンソース AI の推進と開発のためのツールが紹介されました。

このレポートでは、フォーラムでの議論から得られた重要な洞察と要点を紹介し、結論では推奨されるアクションのリストをまとめています。



オープンソース AI におけるグローバルなコラボレーションの促進

フォーラムは、オープンソース AI におけるグローバルなコラボレーションが直面する課題と今後の道筋を探るパネルディスカッションから始まりました。パネリストたちは、規制の断片化、地政学的緊張、オープンソース AI の定義に関する議論など、連携と知識の共有に課題をもたらしている要因を明らかにしました。こうした課題に直面する中、パネリストたちは、オープンソース AI におけるグローバルな連携を促進するためのアプローチについて議論しました。

議論では、以下の具体的な課題が浮き彫りになり、推奨されるアクションが提案されました。

オープンソース AI の定義に関する国際的な合意形成

パネリストたちは、「オープンソース AI」の定義について合意を形成することの難しさについて議論しました。AI モデルのオープン性およびライセンスに関する取り組みはさまざまであり、ほとんどの「オープン」モデルはオープンソースの 4 つの自由（使用、研究、改変、再配布）を遵守していないことを踏まえ、この用語に関する共通認識を確立することが不可欠であると主張しました。

この議論では、オープンソース イニシアチブ (OSI) の「オープンソース AI 定義 (OSAID) バージョン 1.0」⁹ が重要な参考資料となりました。OSAID は、AI システムがオープンソースとみなされるためには、そのシステムおよび構成要素が、個人が許可なく使用、研究、改変、再配布できる 4 つの自由を認める条件の下で提供されている必要があることを重視しています。特に、OSAID は、オープンソースの定義 (#2) の要件「ソースコードは、プログラマーがプログラムを変更する際に好ましい形式である必要がある」を AI システムに適用しています。この要件を満たすため、モデルパラメーター（事前学習された重みとバイアス）、トレーニングデータの前処理コード、AI システムのトレーニングと実行のためのコード、およびトレーニングデータ自体（それが不可能な場合、十分な詳細な情報）は、すべてオープンソースライセンスの下で公開される必要があります。

OSAID は、オープンソース AI コミュニティから賛同と批判の両方を浴びています。批判は、企業が開示できないデータで AI システムやモデルを訓練することを受け入れられない「マキシマリスト」的な立場（例：すべての訓練データが最初から公開されていることを要求する）から、コードやデータのアートファクトをオープンソース化することが過重な負担となるという懸念（例：プロプライエタリなコードの削除が必要だったり、プロプライエタリなデータの機密性）まで多岐にわたります。結果として、オープンソース AI は依然として議論的となっています。パネリストは、OSAID はバージョン 1.0 に過ぎず、100 人を超える貢献者が参加したグローバルな共創プロセスから生まれたように、グローバルなオープンソース AI コミュニティの議論に基づいて継続的に開発されていくと強調しました。あるパネリストは「これは OSAID v1.0 であり、パッチはコミュニティから提供されるべきだ」と述べました。

2 人目のパネリストも、オープンソース AI の定義に取り組み、その定義をより厳格にすることで得られるメリットを訴えていく人材がさらに必要であることに同意しました。彼らは、オープンソースの 4 つの自由のうち、使用と再配布の自由は定義で表現しやすいが、研究と改変の自由はより複雑であると指摘しました。

パネリストはさらに、公共部門は AI の導入がまだ初期段階にあるものの、政府機関の開発者は、モデルを十分に調査・修正するために必要なものを定義する必要が生じるため、よりオープン化を求める立場に置かれるようになるかもしれない、と指摘しました。

パネリストは2つの例を挙げました。1つ目は、政府機関の開発者がAIシステムの偏りや脆弱性を監査し、必要に応じて修正できるべきである、というものです。2つ目は、政府機関の開発者が、使用するソフトウェアやモデルのセキュリティアップデートを作成できるべきである、というものです。ただし、この例は、一部の政府機関がITセキュリティに対する考え方を変える必要があるため、実現はより困難です。パネリストは、「AIシステムの偏見を検出して除去し、セキュリティの問題を監査して修正するには、政府は、私たちの言葉で言えば、研究と改変を行う能力を持つ必要がある」と主張しました。私たちと同じように、政府もAIシステムが十分に研究され、改変できる状態をどのように定義すべきか、その方法を探っているのかもしれませんが。おそらく、私たちが協力することで、この複雑な問題に取り組む人々を増やすことができるでしょう。

OSAIID以外にも、あるパネリストがGenerative AI Commonsが提唱する「Model Openness Framework (MOF)」を、AIモデルのオープンネスと完全性を促進するための実用的なツールとして紹介しました。特に、MOFは機械学習モデルの開発ライフサイクルから、オープンモデル配布で公開される16のコード、データ、ドキュメントの成果物を特定し、適切なオープンライセンス下での成果物の公開に基づいて、モデルの完全性とオープンネスを評価するものです。あるパネリストは、OSAIIDとMOFは互いに補完的な関係にあると説明しました。MOFは、オープンライセンスで公開されたモデル成果物の可用性を評価するためのフレームワークを提供し、一方OSAIIDはオープンソースAIのオープン性を評価するためのバイナリ基準を提示しています。

オープンソース AI コラボレーションにおける地政学的な課題と規制上の課題

パネリストたちは、オープンソースAIにおけるグローバルな連携の妨げとなる地政学的障害について、率直に議論しました。また、AIモデルの輸出規制、規制の断片化とコンプライアンスの負担、およびより広範な政治的緊張が、慎重な検討と現実的な解決策を必要とする重大な課題であるとの認識を示しました。

パネリストたちは、さまざまな国や地域におけるAI規制の相違が拡大し、規制の断片化によってAIモデル、ソフトウェア、データセット、知識の共有が妨げられるリスクについて懸念を表明しました。EUのAI法は、世界初の包括的なAI規制として施行されましたが、AI規制について地域ごとに異なるアプローチが採用された場合、コンプライアンス業務が増えるリスクがあります。異なる規制を遵守する必要性により、特定の管轄区域に限定されたソフトウェアやモデルの流通が生じる可能性もあります。こうしたリスクは、AI規制に関する国際協力の必要性を際立たせています。あるパネリストは、最も困難な課題の一つはデータにあると主張し、著作権法がトレーニング用データセットのオープンな共有と利用に法的問題を引き起こしていると指摘しました。そのパネリストは、次のように述べています。「オープンソースは、著作権を利用してソフトウェアの自由な流通を保証する画期的な手法です。LLMやデータについても、同様のことを実現する方法を考え出す必要がある」と主張しました。

AI規制の指針となるアプローチについて尋ねられた際、あるパネリストはEUのAI法とサイバーレジリエンス法(CRA)を、オープンソースに対する一定の例外を盛り込み、他の管轄区域における類似の規制のモデルとなる初めての例として挙げました。パネリストの一人は、これらは最初の事例であるため完璧ではないかもしれないことを認め、法律の専門家はAIのトレーニングや著作権に関する議論など、新たな課題に取り組んでいると述べた。しかし、オープンソースAIにおけるイノベーションと投資を促進できることを示すことができれば、他の国や地域でも同様の法律が制定され、規制の断片化やグローバルなオープンソースエコシステムへの影響のリスクが軽減される可能性があるため、熱意を持ってコンプライアンスに取り組むことが私たちの共通の目標であるべきである。

オープンソース AI のグローバルな コラボレーションのためのフレームワーク

パネルディスカッションは、オープンソース AI におけるグローバルなコラボレーションとガバナンスの枠組みについての議論に移りました。あるパネリストは、ソフトウェア開発と AI 開発の違いを明らかにし、オープンソース AI におけるグローバルなコラボレーションには、専用のハードウェア、コンピューティング、データなど、AI 開発特有の特性やリソース要件に適した新しいコラボレーションの枠組みが必要であると指摘しました。さらに、「いくつかのガバナンスの原則や取り組みはすでに存在しますが、イノベーションのペースには大きく遅れをとっています」と付け加えました。オープンソースによるコラボレーションは、AI のイノベーションのペースに追いつくための手段であり、「未知の未知」に対処し、「既知の未知」が発生した場合にそれを管理するのに役立つ、オープンな開発プロセスとガバナンスのフレームワークを提供します。

パネルディスカッションでは、オープンソース AI における国際的なコラボレーションの成功事例が共有されました。パネリストの一人は、オープンソース AI のコラボレーションがデータ、アルゴリズム、モデル、ベンチマーク、ツールなど複数の分野で展開されていると説明しました。オープンソースにおけるグローバルなコラボレーションを促進するため、彼らは、CORE の価値観、すなわち、コラボレーション、オープン性、尊重、国家間の平等を推進することの重要性を明確にした。彼らは、その成功例として、中国と ASEAN の AI オープンソースアライアンスを挙げています。これは、中国と東南アジアの複数の国の研究者が AI の研究開発で協力し、各々がホストするオープンソースツールを共同開発して、アライアンス内に単一のコントロールポイントが存在しないようにするパートナーシップです。

また、グローバルコミュニティ間で最先端の可能性に関する対話を促進し、2025 年 1 月の「DeepSeek 事件」のような予期せぬ事態を回避するための手段として、オープンソースによる連携が挙げられました。特に、「DeepSeek 事件」は、地域間の視点の違いを浮き彫りにしました。欧米のパネリストは、このニュースに驚きを表明しましたが、中国の 2 人のパネリストは、中国以外の反応は、中国の AI コミュニティに対する理解が不十分であることから生じているとの見解を示しました。パネリストたちは、地域や言語の壁が、グローバルな AI の研究開発 (R&D) において相互の盲点を生み出していることを明らかにし、オープンソースは共同開発だけでなく、グローバルな対話のための手段でもあることを指摘しました。

オープンソース AI イノベーションの経済的課題

パネリストたちは、学術研究者、草の根の取り組み、企業など、さまざまなステークホルダーグループにとって、オープンソース AI 開発における経済的な課題について議論しました。企業について、あるパネリストは、オープンソース AI の持続可能性を確保するためには、解決すべき経済的な緊張関係があることが、彼らの調査で明らかになったと説明しました。オープンソース AI の競争は、価格の下落とモデルの品質向上をもたらしますが、企業がオープンモデルから得ることができる直接的な価値も低下させます。パネリストは、企業がソフトウェアをオープンソース化する場合、通常は事業の別の部門がそれを補填していると指摘しました。しかし、AI 開発は伝統的なソフトウェア開発に比べて大幅にコストが高く、オープンソースソフトウェア企業で機能するオープンコアビジネスモデルが、同様にオープンソース AI 企業でスケールし持続可能かどうかは、依然として未解決の課題です。

推奨されるアクション：

- コミュニティ主導の対話と共同設計を継続し、オープンソース AI コミュニティにおいてオープンソース AI の定義に関するコンセンサスを構築する。
- オープンソース AI において、オープンソースの 4 つの自由（使用、研究、改変、再配布）を尊重するオープンな開発およびライセンス付与の取組みを推進する。
- オープン AI アセット（モデルには OpenMDW ライセンス、コードには Apache v2 または MIT ライセンス、データには CC-BY または ODC-BY ライセンスなど）に対して、寛容的なライセンスの使用を促進する。
- オープンソース AI コミュニティにおいて、再現性、透明性、利便性におけるオープン性のメリットに焦点を当て、オープンサイエンスの価値観と実践を推進する。
- AI モデルおよびシステムの研究と改変におけるオープンソースの価値について政策立案者に働きかけ、バイアスやセキュリティ上の問題の監査と修正を可能にする。
- AI 規制に関する世界的な整合性について国際的な対話を促進し、規制の断片化とグローバルなオープンソースエコシステムへの悪影響を防止する。
- AI のインフラストラクチャおよびリソース要件を考慮し、オープンソース AI 企業を拡大・維持できる資金調達メカニズムおよびビジネスモデルを模索する。

オープンソース AI がデジタル主権と自律性にもたらす可能性

第2のパネルでは、オープンソース AI のイノベーションと、デジタル主権や自律性に関する政治的課題との関連性について考察しました。パネリストたちは、AI スタックにおける依存関係の影響について議論し、グローバルなオープンソースエコシステムの断片化を回避しつつ、デジタル主権や自律性を強化するために企業が取るべき行動について話し合いました。彼らは、オープンソースは、各国政府がデジタル主権を強化すると同時に AI におけるグローバルなコラボレーションを促進するための手段であり、また、変化し続ける地政学的状況において AI 研究者や開発者間の外交手段でもあることを主張しました。

議論では、以下の具体的な課題が明らかにされ、対応策が提案されました。

現在のデジタル主権の状況

パネリストの一人は、厳しい現実を指摘しました：「現在、デジタルブラックアウトは SF のシナリオではないことが明らかになりました。ブラックアウト状況下で自立して対応する能力がなければ、深刻な問題に直面するでしょう。」別のパネリストは、欧州は根本的にデジタル主権を欠如しており、チップからクラウドまでの AI の全領域で依存状態にあると主張しました。欧州ではクラウドインフラへの投資が不十分であり、競争力のある代替インフラを構築するには数十億ドルと数年を要するでしょう。パネリストは「デジタル主権」を単純な二元論に矮小化すべきではないと指摘し、絶対的な主権よりも絶対的な依存を回避することが重要であり、主権は「陰陽」のような概念として理解すべきだと主張しました。彼らはさらに、ドイツ、フランス、シンガポールのような中規模経済国にとって、AI 能力とインフラへの投資を組み合わせつつ、グローバルシステムにおける相互依存性を認識するバランスの取れたデジタル主権のアプローチが不可欠であると主張しました。

今後の対応方向：共同インフラ投資と「政策としてのオープンネス」

国家や地域のデジタル主権の状況を踏まえ、パネリストたちは、デジタル主権の課題に関する今後の道筋と、オープンソース AI が果たす役割について議論しました。2人目のパネリストは、現在のヨーロッパの状況は、モバイル時代以前のアフリカに似ており、AI およびグローバルなテクノロジーエコシステムにおいて競争力を持つためには、段階的なキャッチアップではなく、飛躍的な戦略が必要であると指摘しました。この目標を達成するため、彼らは AI 技術とデジタルインフラへの大規模な投資が不可欠であり、政府と企業が共同でこれらの投資を行う方法を模索すべきだと主張しました。具体例として、EU の支援を受けてドイツの SAP、DT、Siemens、Ionos、Schwarz Group が連携して AI ギガファクトリーを建設する官民連携の事例を挙げました。

同時に、パネリストは、AI におけるグローバル競争力を高めるためには、政府は「政策としてのオープンネス」を掲げるべきだと主張しました。これには、オープンソース AI の5つの相互に関連する柱、すなわちオープンサイエンス、オープンスタード、オープンソース、オープンデータ、オープンウェイトを戦略的ツールとして活用し、AI の研究開発と商業化の公平な競争条件を整え、相互運用可能なオープンなテクノロジーエコシステムを構築し、最終的にはグローバルテクノロジー大手企業の市場支配力に対する集団的な対抗勢力を形成することが必要となります。パネリストは、一方では国境を閉鎖したり、他方ではテクノロジー大手企業に対して永続的な「テナント」の地位を受け入れるのではなく、「政策としてのオープンネス」戦略は、オープンソースを逆手に取るのではなく、まさにそれを活用して、政府がデジタル主権を強化するための現実的な方法であると主張しました。

別のパネリストは、デジタル主権政策（ソフトウェアやモデルの配布制限など）により、グローバルなオープンソースエコシステムの分断化リスクを警告し、グローバルなオープンソースエコシステムの地域的な分断化を緩和することが、各国政府の共通の目標であるべきだと主張しました。彼らは、オープンソースエコシステムを分断することなくデジタル主権を強化できる、コンピューティングインフラへの地域投資などの生産的な介入と、50 件の scikit-learn や 50 件の PyTorch を構築するなど、努力の重複や資源の浪費のリスクがある配布制限などの逆効果な介入とを区別しました。彼らは、AI 能力とインフラストラクチャ、およびオープンソースへの同時投資は、AI におけるデジタル主権とグローバルなデジタル連携の両方を強化する、双方にメリットのある戦略であるという 2 人目のパネリストの見解に同意しました。3 人目のパネリストもこれに賛同し、「オープンソースは、この変化する地政学的体制における新しい外交手段である」と主張しました。

オープンソースがソブリン AI 能力の構築における役割

別のパネリストは、欧州全域の 1,100 人以上の研究者やエンジニアが、EU の 24 言語すべてで高性能なオープンなマルチモーダルな基盤モデルの開発に協力しているコミュニティ主導の取り組み「OpenLLM Europe」を紹介しました。OpenLLM Europe は、「真にオープンソースの AI」の開発に注力しており、すべてのモデル重み、トレーニングコード、データをオープンライセンスで公開するとともに、開発プロセスの透明性を確保するドキュメントも公開しています。彼らは、倫理的、法的、環境的な考慮を最初から最後まで組み込んだモデル開発に重点を置いています。具体的には、倫理原則や規制要件への準拠から、計算リソースを少なく済ませるため（一般に 240 億パラメーター未満）の小型で専門的なモデルの開発に注力しています。これにより、トレーニングとデプロイメント時のエネルギー消費を削減できます。パネルディスカッションの参加者は、モデルを伝統的なハードウェアで訓練・デプロイ可能に設計し、公共の研究インフラでモデルをトレーニングすると説明しました。例えば、OpenLLM Europe はフラッグシップの Lucie-7B 言語モデルをフランスの Jean-Z スーパーコンピュータでトレーニングしました。

同じパネリストは、小規模モデルの開発とオープンな共有のトレンドが、AI における協業の促進とデジタル主権の強化において特に有望な分野であることを指摘しました。開発者は、計算リソースとエネルギー消費が少なく済む小規模モデルを容易に開発、カスタマイズ、展開でき、これらのモデルの性能も大規模モデルに迫る水準に達しています。例えば、パラメーター数が 20 億未満の小規模モデルが、今年の最先端のプロプライエタリモデルを既に上回る性能を発揮していることが挙げられました。

パネルディスカッションでは、主権的なオープンソース AI の構築におけるもう一つの重要な点として、地域的な文脈、価値観、言語に特化したモデルのカスタマイズと評価が議論されました。言語モデルが文化的価値観や視点を反映していることを認識した上で、言語は単なるコミュニケーションツールではなく、文化的コミュニティを象徴するものであると指摘するパネリストもいました。別のパネリストもこれに賛同し、言語や文化の多様性に焦点を当てたプロジェクトは、AI が特定の言語や文化に偏らないことを確保する上で重要な役割を果たすことを指摘しました。彼らは、低・中リソース言語向けのオープントレーニングおよび評価データセットを共同で開発するオープンソースの取り組みをリードする研究者や開発者を称賛しました。

推奨されるアクション：

- オープンソースに投資し、それを活用してグローバルな AI 分野での協力と競争の両方を推進する、オープン性を重視した AI 産業戦略を策定する。
- AI 機能とデジタルインフラストラクチャへの共同投資の方法を模索する。
- オープンデータ、パブリックコンピューティング、オープンな開発およびガバナンスプロセスに裏打ちされた、真にオープンソースの AI モデルの開発と維持に資金を提供する。
- 公共の利益分野におけるオープンデータの研究開発と維持管理を支援する。これには、低リソースまたは中リソース言語向けのトレーニングおよび評価データセットが含まれる。
- 専門的な下流用途やアプリケーション向けにカスタマイズしやすい小型モデルの開発を優先する。



ライトニングトーク 1: LMOS by Deutsche Telekom

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

Eclipse LMOS は、クラウドまたはオンプレミスでエンタープライズグレードの AI エージェントシステムを大規模に構築および実行するための、オープンソースのベンダー中立のプラットフォームです。現在 Eclipse Foundation がホストしている LMOS は、DT のヨーロッパ全域で販売およびサービスエージェントを構築していた DT のエンジニアの小さなグループによる草の根の取り組みとして始まりました。既存のエンジニアリングチームとインフラを使用して構築されており、大規模なツールの入れ替えは必要ありません。現在、LMOS は本番環境で、複数の国における何百万ものエージェントによる会話を処理し、アクションを実行する AI エージェントの 30% 程度を人間のエージェントに引き継いでいます。これにより、企業はエージェントを独自に管理、進化させることができ、持続可能でスケーラブルな AI 運用への移行が進んでいます。LMOS は、その設計上、独立性が高く、移植性にも優れています。Kubernetes で動作し、マルチテナントに対応、既存のシステムと統合可能、モデルの柔軟性にも優れているため、組織は単一のエコシステムに縛られることなく AI の取り組みを拡張することができます。DT は、オープンなコラボレーションと進歩の共有を支援するため、LMOS をオープンソース化し、Eclipse Foundation に寄贈しました。LMOS は、草の根のエンジニアリングとオープンな実験が、企業がこれまでに築き上げたものを無駄にすることなく、具体的な成果につながることを示す生きた例です。

ライトニングトーク 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>

AI のオープン化を促進するため、Linux Foundation の Generative AI Commons は 2024 年に MOF を開発しました。MOF は、機械学習モデルのオープン性と完全性を評価するための 3 層分類システムであり、モデルの開発ライフサイクルから 16 のコード、データ、およびドキュメントの成果物を特定し、これらの成果物が適切なオープンライセンスでリリースされているかどうかに基づいて、そのオープン性と完全性を分類します。3 つの階層には、オープンモデルからオープンサイエンスモデルまでが含まれます。たとえば、オープンモデルとして分類するには、モデルの重み、モデルアーキテクチャ、ドキュメント（モデルカード、データカード、技術レポートなど）、評価結果を公開する必要があります。MOF を基に、2025 年、Linux Foundation は、機械学習モデル用に特別に設計された最初の寛容なライセンスとして、モデル（アーキテクチャおよび重み）だけでなく、モデルとともにリリースされる可能性のあるコード、データ、およびドキュメントも対象とする OpenMDW License Agreement v1.0 を策定しました。OpenMDW は、著作権法、特許法、データベース法、営業秘密法に基づく権利を付与し、特許訴訟保護と帰属表示を含み、モデルの出力を制限しません。これは、同じ配布物または別々に公開される成果物に適用される他のライセンス（例：オープンデータセット用の CC-BY または ODC-BY、オープンソースソフトウェア用の Apache 2.0 または MIT）と互換性のあるグローバルライセンスです。

ライトニングトーク 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 の Qwen3 は、最先端のオープン LLM ファミリーで、0.6B から 32B のパラメーターを持つ 6 つの高密度モデルと、30B と 235B のパラメーターを持つ 2 つの Mixture-of-Experts (MoE) モデルを含む。これらのモデルは、モバイルデバイス、スマートグラス、ロボティクスなど、多様なアプリケーション向けに最適化されている。Qwen3 は、ハイブリッド推論アプローチを先駆けて採用しており、複雑な推論タスク用の「思考」モードと、より迅速な回答のための「非思考」モードを切り替えることができます。36 兆個のトークンでトレーニングされた Qwen3 モデルは、多言語タスク (119 言語および方言に対応)、推論、ツール使用、人間と調和した対話において、最先端のパフォーマンスを発揮します。Qwen3 モデルは、Model Context Protocol と関数呼び出しサポートを通じて高度なエージェント機能を統合しています。Qwen3 モデルは、Hugging Face Hub、GitHub、ModelScope などのプラットフォームで Apache 2.0 ライセンスの下で利用可能です。

オープンソース AI を通じて AI 研究と再現性を推進する

パネルディスカッションの3つ目は、オープンソース AI と研究をテーマに、オープンソースが研究の推進と再現可能性に果たす役割に焦点を当てました。パネリストたちは、AI 研究における再現性の主な課題について議論しました。その中には、AI 開発プロセスやデータの出所の透明性の欠如、研究インフラへのアクセス制限などが挙げられました。AI 研究者が、草の根のオープンソースのコラボレーションを組織し、LLM の開発や評価に使用したコードやデータを公開し、ドキュメントや教訓を他者が精査、学習、活用できるように共有することで、これらの課題に対処している実態を明らかにしました。

議論では、以下の具体的な課題が明らかにされ、対応策が提案されました。

オープンソースは、AI 研究における再現性の基盤です

パネリストたちは、再現性があり、協調的な AI 研究の基盤となるオープン性の役割について議論しました。あるパネリストは、現在、ほとんどの LLM およびその基礎となるトレーニングプロセスやデータは非公開であり、科学者がモデルのパフォーマンスに関する主張を独自に検証したり、潜在的なバイアスを調査したり、既存の研究を厳密に発展させたりする可能性が著しく制限されていることを明らかにしました。彼らは、真の再現性を実現するには、データセットの構成からモデルのトレーニング、評価に至る AI モデルのパイプライン全体をオープンにする必要があると主張しました。しかし、モデルの重みを除けば、現在、AI 研究者がモデル開発パイプラインの重要な成果物(トレーニングデータやコードなど)を共有することは一般的ではありません。

LLM の透明性の欠如は、Eleuther AI、BigScience Workshop、LAION などのオープン LLM の開発を目的とした草の根イニシアチブを生み出しました。パネリストは、これらのコミュニティは、ドキュメント(技術レポートや研究論文など)を公開し、LLM の開発に使用したコードやデータをモデルチェックポイントとともに公開することで、LLM 開発パイプライン全体の透明性と再現性を確保することを目指していると説明しました。さらに、Ai2 や BAAI などの研究機関の研究者もオープンサイエンスの取組みを採用し、トレーニングコード、トレーニングデータ、ドキュメントをモデルチェックポイントと共にオープンソースライセンスの下で共有しています。

パネリストは、このオープンサイエンスアプローチによる AI 研究開発が、研究者が知識を共有し、互いに学び合い、過去の研究成果を検証・発展させることを可能にし、これにより信頼が強化され、科学的な進歩が加速されると説明しました。オープンソースの AI コラボレーションは、研究者がそれぞれの専門知識とリソースを共有できるため、AI 開発への障壁を低くします。彼らは、オープンソースは品質管理と健全性チェック」を提供すると主張し、オープンソースコミュニティの厳格な検証は、時には厳しくても、研究の品質を向上させ、根拠のない主張を防ぐと説明しました。

しかし、パネリストは、LLM のトレーニングや再現に必要な計算リソースへのアクセスが、研究者にとって大きなボトルネックとなっていることを明らかにしました。さらに、「オープンソースだけでは決して十分ではない」と主張し、共有だけでなく、コラボレーションのコミュニティと文化を構築することが重要であると強調しました。

オープン性は、科学的な発見における AI の普及の鍵

パネリストたちは、科学研究におけるオープンソース AI の役割と、さまざまな分野におけるオープンモデルの採用に伴う課題について議論しました。あるパネリストは、基礎モデルが医学、化学、生物学などの分野における科学的発見を促進する可能性がある一方で、そのブラックボックス的な性質とオープン性の欠如が、研究者が研究に基礎モデルを適用することを妨げる要因となっているという研究結果について、その見解を述べました。また、さまざまな科学分野の研究者がさまざまなオープンソース AI リソースを利用することで生まれた「心のスケールリング法則」というフレームワークを紹介しました。この発見に基づき、研究者は、実際の研究データと合成研究データを収集・共有し、科学者の AI スキルを育成し、基礎モデルの採用が科学的発見に有益な研究課題を戦略的に特定するという好循環を確立するよう研究コミュニティに呼びかけました。

エッジと組み込み AI におけるオープンソースアプローチ

パネルは、エッジ AI および組み込み AI におけるオープンソースのメリットについて議論に移りました。あるパネリストは、モデルやコードだけでなく、エッジで収集されたデータや基盤となるハードウェアアーキテクチャも網羅する、エッジ AI スタック全体のオープン化を促進するためのフレームワークを紹介しました。その後、パネリストたちは、組み込みシステム向け AI の開発に伴う課題と可能性について議論しました。あるパネリストは、組み込み AI 開発におけるオープンソースが最適なアプローチであるとの強い主張を展開し、この専門分野において最も効果的なコラボレーションを実現するため、オープンソースチャネルを通じた協業がコード、ツール、ベストプラクティスの共有を可能にする点を強調しました。

推奨されるアクション

- TRUE 原則（透明性、再現性、理解可能性、実行可能性）に基づく、透明で再現可能な AI の研究開発のためのオープンスタンダードを策定する。
- オープンサイエンスの実践（モデル、コード、データ、ドキュメントの公開を含む）を促進または、可能な場合は義務付ける、公的研究助成金を設計する。
- 公共の利益分野に対する補助金やクレジットなど、研究者が利用可能な国家または国際的な公共の研究インフラストラクチャの資金調達方法を検討する。
- 科学研究コミュニティにおける AI リテラシーを、オープンモデルを科学的な応用において使用し最適化する方法に関する教育資料や認定プログラムを通じて育成する。

企業におけるオープンソース AI ソリューションの採用を阻害する要因と促進する要因

パネルディスカッションの4つ目は、企業がオープンソース AI ソリューションを採用する上で共通する障害と促進要因について議論しました。パネリストたちは、管理された研究環境でのベンチマークにおけるモデルのパフォーマンスと実環境での導入とのギャップ、AI を既存のワークフローに統合することの難しさ、信頼性、信用、安全性を確保するための透明性の重要性など、企業が採用する際に直面するいくつかの課題について言及しました。パネリストたちは、オープンモデルを企業ソリューションにうまく統合するための実践的なアプローチ、教訓、および重要な指標について共有しました。また、エージェントシステム向けのオープンソースツールとオープンスタンダードの共同開発の可能性についても議論されました。

議論では、以下の具体的な課題が明らかにされ、対応策が提案されました。

企業におけるオープンソース AI ソリューションの採用における課題

パネリストたちは、管理された研究環境で達成されたベンチマークにおけるモデルのパフォーマンスと、本番環境での実際の導入準備状況との間には大きなギャップが存在することを指摘しました。あるパネリストは、オープンソースの AI 開発者は、オープンベンチマークでのパフォーマンスの向上やモデルランキングの上位入りを重視する傾向がある一方、企業は実環境でのモデル導入において異なる課題に直面すると指摘しました。このギャップは、企業がオープンモデルを広く採用する上で最大の障壁のひとつであると彼らは主張しました。なぜなら、企業は採用前に、信頼性および安全性の基準、規制要件を確実に満たす必要があるからです。

パネリストたちは、AI ソリューションの信頼性を確保し、オープンモデルの導入における信頼構築が企業における採用の重要な課題である点で一致しました。彼らは、信頼を実現するためのさまざまな戦略につい

て議論し、モデルの評価と検証の厳格化、モデル開発プロセスの透明性向上、堅牢なモニタリングとフィードバックメカニズムの実現などを挙げました。別のパネリストは、組織は、情報に基づいた意思決定を行うために、明確で定量化可能な費用対効果分析ツールと迅速な比較機能が必要であると主張しました。彼らは、自社では、コストと本番環境への導入までのスピードという2つの重要な指標を優先する実用的なアプローチを採用していると説明しました。さらに、AI モデルを企業で活用するための統合スタックを用意し、オープンモデルの精度とコストを「ボタンをクリックするだけで」簡単に評価・比較できるツールが、企業での導入には不可欠であると付け加えました。

パネリストたちは、AI ソリューションの成功的な導入は、オープンソースかクラウドソースかに関わらず、既存のワークフローへの統合の容易さとユーザーエクスペリエンスの質に依存すると一致しました。顧客が実際に信頼する AI ソリューションを構築する際、あるパネリストは、顧客の信頼を獲得するためには、高性能な技術力を構築するだけでなく、統合の品質と優れたユーザーエクスペリエンスを確保することが必要だと主張しました。AI テクノロジーは、付加価値のあるサービスや製品を生み出すためのツールであることを忘れてはならないと彼らは主張しました。別のパネリストは、顧客の信頼を得るためには、モデル統合の透明性が重要であることを指摘し、DeepSeek モデルを使用する際に信頼性の高い統合手法を実証しなければならなかった B2C 企業 Perplexity の例を挙げました。これには、モデルから政治的偏見を排除し、モデルを製品に統合する方法を顧客に透明性をもって説明することが含まれていました。

別のパネリストは、現在の AI の研究開発 (R&D) の進路の財政的持続可能性について懸念を表明し、大規模な AI モデルのトレーニングとデプロイメントのコストが多くの企業にとって財政的に持続可能ではないと指摘しました。

オープンソースがエージェント型システム構築における役割

モデルを超えて、パネルは、イノベーションと差別化の両方の可能性を生み出す新しいトレンドとしての AI エージェントについて議論しました。あるパネリストは、エコシステムがプロプライエタリなシステムに固執してしまう前に、世界中の企業が AI エージェントのための相互運用可能なオープンソースツールとオープンスタンダードの開発に積極的に協力すべきだと主張しました。彼らは、クラウドまたはオンプレミスでエンタープライズグレードの AI エージェントシステムを大規模に構築および実行するための、オープンソースのベンダー中立のプラットフォームである Eclipse LMOS の例を挙げました。彼らは次のように述べました。「独

立したオープンソース AI エージェントプラットフォームを構築するための時間は限られています。LMOS は、オープンなエージェントシステムがビジネス価値を最大化しつつ、長期的なロックインリスクを軽減できることを証明するために生まれました。」オープンソース AI エージェントプラットフォームの核心的な考慮点は、直感的な使いやすさ、スケーラビリティ、統合機能、カスタマイズ性、およびエンタープライズグレードのセキュリティ対策に焦点を当てるべきです。

オープンソース Foundation は、オープンソース AI ソリューションの開発を促進することができます

少数のテクノロジー大手企業が AI スタックを支配している現状に代わる選択肢として、パネリストは、オープンソース財団が、オープンソース AI ソリューションの共同開発を促進する上で重要な役割を果たすことができると主張しました。オープンソース Foundation のオープンなガバナンスにより、大小を問わず、あらゆるグローバル企業がオープンソース AI のイノベーションに公平に参加できるからです。

推奨されるアクション:

- トレーニングデータ、開発プロセス、評価に関する文書化要件など、AI モデルの透明性に関するオープンスタンダードを確立し、企業がデューデリジェンスを実施し、十分な情報に基づいて導入の決定を行えるよう支援する。
- 商用利用や改変を制限なく許可する許諾条件の緩やかなライセンスの活用を推進する。
- 規制対象の業界や分野における特定の機能やコンプライアンス要件に対応したオープンモデルの性能をテスト・監視するための、オープンソースの評価フレームワークと業界固有のベンチマークスイートを開発する。
- ベンダー中立の Foundation を活用し、AI スタック全体におけるオープンソースツールとオープンスタンダードの共同開発を促進する。
- 草の根のエンジニアリングを強力に支援し、実証実験を促進し、クローズドな AI システムへのロックインを防ぐ、相互運用可能なオープンソースツールの開発を優先する。

オープンソース AI における責任ある開発手法の推進

最後のパネルは、責任ある AI をテーマとし、AI のイノベーションと普及が追いついていないスピードで進む中、オープンソース AI コミュニティにおいて責任あるプラクティスを促進する上での重要な課題について議論しました。パネリストたちは、倫理原則や規制の認識から、他者が学び、発展させることができるベストプラクティスを実証・文書化するオープンソースプロジェクトまで、責任ある AI の実践を促進するためのさまざまな方法について議論しました。

議論では、以下の具体的な課題が明らかにされ、対応策が提案されました。

倫理原則を実践に反映させる

パネルでは、オープンソースの AI 研究開発において倫理原則を実践に反映させるという課題について議論しました。あるパネリストは、特にイノベーションと普及のペースが追いついていないほど速いことを考えると、これは決して簡単な課題ではないと指摘し、OpenAI が ChatGPT のユーザー数を 100 万人に達するまでにわずか 2 ヶ月半しかかからなかったことを紹介しました。彼らは、オープンソースの利点は AI 開発プロセスの透明性を高めることであり、ソフトウェアやモデルなどの AI テクノロジーのオープン性により、潜在的な偏りや脆弱性について分散型の監査が可能になり、その結果、これらのテクノロジーに対する信頼と安全性が向上すると指摘しました。そのため、AI におけるオープン性のプラクティスを開発、促進、奨励することが重要であると彼らは主張しました。

責任あるオープンソース AI 開発を促進する規制の役割

パネルでは、責任ある AI イノベーションの促進における規制の役割について議論が行われました。パネリストたちは、規制とイノベーションは対立するものではなく、規制の目的は安全で責任あるイノベーションの形成することであるべきだと述べました。あるパネリストは、市場参加者が安全やセキュリティに自発的に資源を投資しない分野では、規制が重要な役割を果たすことを指摘しました。このような介入により、たとえそれが当面の商業的インセンティブと一致しない場合でも、公共の利益が AI の開発に確実に反映されるようになります。別のパネリストは、EU の AI 法について言及しました。この法律は、汎用 AI モデルの提供者に、著作権遵守ポリシーやトレーニングデータの概要などの義務を課すことで、信頼性と安全性の向上を図るとともに、オープンソース AI 開発者のコンプライアンス負担を軽減するためのオープンソースの適用除外規定を設けています。

責任ある AI のためのオープンソースツールとオープンベンチマークの構築

パネルでは、責任ある信頼性の高い AI の開発と採用を促進するための実用的なツールとして、オープンソースの評価フレームワークとオープンベンチマークについて議論が行われました。あるパネリストは、「信頼性の高い AI の産業化」をモットーに、産業における信頼性の高い AI システムの開発と採用を可能にするオープンソースの方法論とツールを構築している European Trustworthy AI Foundation について概要を説明しました。これらのツールは、フランスの「Confiance IA」（AI に対する信頼）研究プログラムの結果を参考にし、メーカー、研究者、スタートアップ、研究所など、同 Foundation のマルチステークホルダーコミュニティによって開発されています。さらに、EleutherAI の LM Evaluation Harness、Hugging Face の lighteval、AI Verify

Foundation の Project Moonshot、英国 AI Security Institute の Inspect などのオープンソースの評価フレームワーク、およびさまざまな研究や業界固有のアプリケーション向けのオープンベンチマークのエコシステムが拡大しており、研究者や開発者に AI モデルの性能と安全性を体系的に評価するための不可欠なツールを提供しています。先述のパネルディスカッションのパネリストが述べたように、これらのオープンソースツールを維持・拡大し、より多くの現実世界のタスクやコンテキストをカバーできるようにすることは、規制産業において企業が安全かつ信頼できる方法でオープンモデルを採用するための鍵となるでしょう。

デモとドキュメンテーション：オープンソースコミュニティにおける模範的なリーダーシップ

パネルはさらに、オープンソース開発者が、他者が学び、それを基に開発を進めることができるような形で、責任ある開発手法を文書化することで、模範を示すことができる方法について議論しました。

パネリストは、オープンソースコードの最大のアーカイブを管理する非営利団体 Software Heritage が、そのアーカイブに保存されているコードを LLM のトレーニングデータとして責任を持って利用することを促進する取り組みを紹介しました。2023 年 10 月、Software Heritage は「Software Heritage Statement on LLMs for Code」を発表し、Software Heritage アーカイブのコードを用いて LLM をトレーニングする場合、開発者が遵守すべき 3 つの原則をまとめました。この原則では、1) モデルはオープンライセンスで、その使用に必要な文書やツ

ルとともに共有されること、2) トレーニングデータは SWHID 識別子によって完全に識別可能であり、バイアス調査、トレーニングデータに関心のあるコードが含まれているかどうかの確認、生成されたコードがトレーニングデータと類似している場合の適切な帰属表示が可能であること、3) 作成者がトレーニングデータセットから自分のコードをアウトアウトするための仕組みが確立されていること、が求められています。

声明の発表以来、Software Heritage の原則はコード向けのオープン LLM の開発に影響を与えてきました。例えば、Hugging Face、ServiceNow、Nvidia が参画する BigCode プロジェクトが開発したコーディング向けオープン LLM「StarCoder2」は、これらの原則に従って企業によって開発された最初のコーディング向け LLM です。しかし、BigCode プロジェクトはこれにとどまりませんでした。BigCode の開発者は、Software Heritage のデータを原則に従ってフィルタリングした方法を記載した研究論文と、他の開発者が学習、利用、開発に活用できるフィルタリング済みデータセット「The Stack v2」を公開しました。この例は、オープンソース AI 開発者が、倫理原則の実践的な実施内容を文書化し、倫理的に開発されたトレーニングデータセットやモデルを公開することで、より広範なエコシステムにおける責任ある AI 開発手法に効果的に影響を与えることができることを示しています。

推奨されるアクション:

- グローバルな研究者や開発者がモデル、コード、データの分散監査を実施できるようにすることで、セキュリティを強化する手段として、オープン性と透明性を促進する。
- オープンソース AI の研究者や開発者向けに、責任ある AI の実践と規制順守に関する無料の教育リソースおよび認定プログラムを作成する。
- 安全な AI 開発と規制順守の障壁を低減するための、オープンソースの評価フレームワークおよびベンチマークスイートを開発する。
- 事例 1: 責任ある開発手法を文書化することで、オープンソース開発者は、他の開発者が参考にし、学び、開発することのできるテンプレートを提供することができます。
- 事例 2: 倫理的に開発されたデータセットやモデルを公開することで、オープンソース開発者は、他の開発者が責任ある AI 開発手法を採用し、その発展に貢献できるようにします。



結論

GOSIM Open Source AI Strategy Forum は、オープンソース AI におけるグローバルな連携の課題と可能性の両方に焦点を当てました。オープンモデルは、性能面でプロプライエタリな代替品に急速に追いつき、コスト、カスタマイズ、透明性において大きな優位性を発揮していますが、エコシステムは、オープン性に関する標準や実施方法の相違、法規制の断片化、研究ベンチマークと実際の企業導入ニーズとのギャップなど、重大な障壁に直面しています。

フォーラムでの議論では、さまざまなステークホルダーグループにとっての可能性と優先課題が明らかになりました。例えば、政府はオープン化推進戦略を通じてデジタル主権を強化し、企業は AI スタック全体で相互運用可能なオープンソースのフレームワークとオープンスタンダードを共同開発して、プロプライエタリなソリューションへのロックインを防止し、オープンソース開発者は責任ある AI 開発手法を率先して推進すべきである、といったことが挙げられました。

パネルディスカッションでは、これらの課題に対処し、AI の民主化を共同で推進するために、産業界、学界、政府、市民社会、オープンソース Foundation、コミュニティのステークホルダーに対して、以下の提言が行われました。

オープンソース AI におけるグローバル コラボレーションの促進

- コミュニティ主導の対話と共同設計を継続的に実施し、オープンソース AI コミュニティにおいて「オープンソース AI」の定義に関するコンセンサスを構築する。
- オープンソース AI において、オープンソースの 4 つの自由(使用、研究、改変、再配布)を尊重するオープンな開発およびライセンス付与の実践を促進する。
- モデルには OpenMDW ライセンス、コードには Apache v2 または

MIT、データには CC-BY または ODC-BY などの寛容なライセンスを、オープン AI 成果物に使用することを促進する。

- 再現性、透明性、有用性におけるオープン性のメリットに重点を置いて、オープンソース AI コミュニティにおけるオープンサイエンスの価値観と実践を推進する。
- AI モデルおよびシステムの調査と変更におけるオープンソースの価値について政策立案者に理解を深めてもらい、バイアスやセキュリティ問題の監査と修正を可能にする。
- AI 規制に関する世界的な整合性について国際的な対話を促進し、規制の断片化とグローバルなオープンソースエコシステムへの悪影響を防止する。
- AI のインフラストラクチャおよびリソース要件を考慮し、オープンソース AI 企業を拡大および維持できる資金調達メカニズムおよびビジネスモデルを模索する。

オープンソース AI がデジタル主権と 自律性にもたらす可能性

- AI におけるグローバルな連携と競争のために、オープンソースへの投資と活用を推進する、オープン性を重視した AI 産業戦略を策定する。
- AI 能力とデジタルインフラへの共同投資の方法を模索する。
- オープンデータ、パブリックコンピューティング、オープンな開発およびガバナンスプロセスに裏打ちされた、真にオープンソースの AI モデルの開発と維持に資金を提供する。

- 公共の利益に資する分野におけるオープンデータの開発と維持を支援する。これには、低リソースまたは中リソース言語向けのトレーニングデータセットと評価データセットの開発も含まれる。
- 開発者が専門的な下流のユースケースやアプリケーション向けに容易にカスタマイズ可能な小型モデルの開発を優先する。

オープンソース AI を通じた研究の促進と再現性の向上

- TRUE 原則（透明性、再現性、理解可能性、実行可能性）に基づく、透明で再現可能な AI の研究開発のためのオープンスタンダードを開発する。
- モデル、コード、データ、文書などの公開を含むオープンサイエンスの実践を奨励、あるいは可能であれば義務化する、公的研究助成金を制度化する。
- 研究者が利用可能な国家または国際的な公共研究インフラの資金調達方法を検討し、公共の利益分野における補助金やクレジットを含む支援策を模索する。
- 科学研究コミュニティにおける AI リテラシーの向上やオープンモデルを科学的な応用において使用し最適化する方法に関する教育資料や認定プログラムを通じて促進する。

企業におけるオープンソース AI ソリューションの採用を阻害する要因と促進する要因

- トレーニングデータ、開発プロセス、評価に関する文書化要件など、AI モデルの透明性に関するオープンスタンダードを確立し、企業がデューデリジェンスを実施し、十分な情報に基づいて導入の決定を行えるようにする。
- オープンモデルの商品利用や改変を制限なく許可する寛容なライセンスの活用を推進する。
- オープンソースの評価フレームワークと業界固有のベンチマークス

イートを開発し、規制対象の業界や分野において、特定の機能やコンプライアンス要件を満たすオープンモデルの性能をテストし監視する。

- ベンダー中立の Foundation を活用して、AI スタック全体におけるオープンソースツールとオープンスタンダードの共同開発を促進します。
- 相互運用可能なオープンソースツールの開発を優先し、草の根のエンジニアリングを支援し、実験を促進し、閉鎖的な AI システムへのロックインを防止する。

オープンソース AI における責任ある開発手法の促進

- グローバルな研究者や開発者がモデル、コード、データの分散監査を実施できるようにすることで、セキュリティを強化する手段としてのオープン性と透明性を促進する。
- オープンソース AI の研究者や開発者向けに、責任ある AI の実践と規制順守に関する無料の教育リソースおよび認定プログラムを作成する。
- オープンソースの評価フレームワークとベンチマークスイートを開発し、安全な AI 開発と規制遵守の障壁を低減する。
- 事例 1: 責任ある開発手法を文書化することで、オープンソース開発者は、他の開発者が参考にし、学び、発展させることのできるテンプレートを提供することができます。
- 事例 2: 倫理的に開発されたデータセットやモデルを公開することで、オープンソース開発者は、他の人が責任ある AI 開発手法を採用し、発展させることを可能にします。

謝辞

著者は、GOSIM および Eclipse Foundation が GOSIM Open Source AI Strategy Forum を共催してくれたこと、また、Open Source Initiative、Generative AI Commons、Software Heritage、OpenWallet Foundation を含むすべての支援組織、そしてフォーラムの参加者およびボランティアの皆さんに感謝いたします。また、レポートの寄稿者である Chris Xie 氏、Anni Lai 氏、Gaël Blondelle 氏、Michael Yuan 氏、Stefano Maffuli 氏、Ciarán O' Riordan 氏、Arun Joseph 氏、Roberto Di Cosmo 氏、Xavier Tao 氏、Yumei He 氏、Hilary Carter 氏、Anna Hermansen 氏、Paula Grzegorzewska 氏にも感謝いたします。また、レポートの作成を担当した Linux Foundation Creative Services にも感謝いたします。

著者について

Cailean Osborne (博士) は、Linux Foundation のシニアリサーチャーとして、AI のオープン化を促進するための戦略的研究とアドボカシー活動を行っています。オックスフォード大学で社会データサイエンスの博士号を取得し、オープンソース AI エコシステムにおけるコラボレーションのダイナミクスについて論文を執筆しました。博士課程在学中は、北京大学のオープンソースソフトウェアデータ分析研究所で客員研究員を務めました。それ以前は、英国政府のデータ倫理・イノベーションセンターで国際政策責任者を務め、英国の AI 戦略を共同執筆し、OECD や欧州評議会などの国際的な AI ガバナンスの取り組みに英国代表として参加しました。現在はドイツのベルリンを拠点とし、余暇は語学学習を楽しんでいます。

GOSIM について

このイベントは、GOSIM AI Paris 2025 の併催イベントとして開催されました。GOSIM は、革新的なオープンソースプロジェクトが輝き、連携し、進化するための舞台を提供することを目的とした、グローバルなオープンソースコミュニティによって、そのコミュニティのために組織されたボランティアによる取り組みです。今後の GOSIM イベントの詳細については、www.gosim.org をご覧ください。

この日本語文書は、[Charting Strategic Directions for Global Collaboration in Open Source AI](#) の参考訳として The Linux Foundation Japan が提供するものです。

翻訳協力：吉田行男

卷末脚注

- 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



2021年に設立された **Linux Foundation Research** は、拡大を続けるオープンソースのコラボレーションを調査し、新たなテクノロジーのトレンド、ベストプラクティス、オープンソースプロジェクトの世界的な影響に関する洞察を提供しています。プロジェクトデータベースとネットワークを活用し、量的および定性的手法におけるベストプラクティスにコミットすることで、**Linux Foundation Research** は、世界中の組織に役立つオープンソースに関する洞察のライブラリを構築しています。

 x.com/linuxfoundation

 facebook.com/TheLinuxFoundation

 linkedin.com/company/the-linux-foundation

 youtube.com/user/TheLinuxFoundation

 github.com/LF-Engineering



Copyright © 2025 **The Linux Foundation**

本レポートは **Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International Public License** の下でライセンスされています。

この報告書を参照する場合は、以下のとおり引用してください。Caillean 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.

