

AI に関するオープン性の機会と共通言語に関する G7 ビジョン

対象

本文書は、政府、規制当局、政府間機関、シンクタンク、非政府組織（NGO）、AI モデル提供者、財団、オープンソース・コミュニティ、市民社会を含む、G7 加盟国及びそれを越えた広範な AI エコシステムを対象としているものである。また、AI の提供者、零細・中小企業（MSME）を含む企業、イノベーション及びデジタルを推進する主体、研究者、科学者、研究資金提供者、公的機関にとっても参考資料となり得るものでもある。本文書の目的は、AI のオープン性を説明する用語の使用においてより明確さを求めることと、コミュニティと協力してこの問題に関する G7 の更なる取組を促進することにある。

本文書は、拘束力を持たない参考資料である。AI ガバナンス、サイバーセキュリティ、データ保護、知的財産権に関連するものを含め、適用される国内又は国際的な枠組みに影響を与えるものではない。

導入

1. 我々、G7 デジタル・技術担当大臣は、オープンソースソフトウェアが、国境を越えた技術アクセスへの障壁を低減することなどを通じて、国際的な研究協力や発見、協調的イノベーション、そしてデジタル分野での進歩にとって不可欠なものとなっていることを認識する。同様に、AI 分野においても、オープン性は G7 諸国及びそれを越えた経済圏に多大な利益をもたらしており、今後も更なる機会が期待される。G7 加盟国は、イノベーション、研究、経済成長への AI のオープン性の多大な利益を認識しつつも、デュアルユース（軍民両用）に関する懸念、セキュリティ上の脆弱性、技術スタックに関連するものを含むサプライチェーンのリスクなどの問題を踏まえて、オープン性を評価すべきであることも認識している。

過去 30 年間、ソフトウェアは G7 諸国の経済を変革し、直接的・間接的な利益をもたらすことで、前例のない成長の機会を可能にしてきた。その間、オープンソースソフトウェアは我々の経済の礎となり、ソフトウェアの開発コストを大幅に削減し、再現性を可能とし、研究を加速させ、世界中の研究員の参画障壁を低減することで発見の原動力となってきた。

人工知能は G7 諸国の経済を変革するものと見込まれており、AI のオープン性は、ソフトウェアのオープン性やその他のオープン性と同様の役割を果たす可能性を秘めている。しかし、AI は定義の明確さという点において新たな課題を提起している。すなわち、オープンウェイト AI やオープンソース AI の意味については依然として論争が巻き起こっており、オープンソースの従来の定義が一定の指針を示しているとはいえ、更なる明確化が必要である。

AI 分野におけるこうした明確さの欠如は、こうした技術のオープン性の度合いに疑念を抱かせる傾向があり、結果としてその恩恵を損なってしまう。したがって、我々 G7 のデジタル・技術担当大臣は、より広範なエコシステムに対し、AI のオープン性を記述する際に用いられる言葉のニュアンスを意識するよう促す。エコシステムがより共通の理解へと向かうことを支援するため、我々は、AI のオープン性の機会と共通言語に関する G7 ビジョンのため、本文書を提示する。これには、この新興分野の発展に寄与することを意図した AI のオープン性による経済的利益に関する声明と、より微妙なニュアンスを含むオープン性の類型化を含む一連の原則が含まれる。明確な区別は、ユーザー、政策立案者及び市場が、AI のオープン性の程度の違いがもたらす実際的な影響をよりよく理解するのに役立つ。

2. AI のオープン性は、イノベーションと協力を促進し、企業やコミュニティの技術へのアクセスを拡大することで、我々の経済に不可欠な貢献をしてきた。

AI のオープン性がもたらす便益は、物理的インフラの最適化から、AI を活用した先進的なビジネスソリューションの実現に至るまで、AI ライフサイクルのあらゆる段階に及ぶ。『AI のオープン化の便益』に関する OECD の報告書で強調されているように、オープンウェイト AI モデル及び関連するオープンソースソフトウェアは、国内総生産を含め、マイクロ及びマクロ経済にプラスの影響を与えうる。その恩恵には、特に大規模な導入において高い費用対効果が挙げられるほか、イノベーション、地域での価値創造、業界ごとの適応、そして戦略的なデジタル自律性の支援も含まれる。同時に、悪意のある利用、セキュリティ上の脆弱性、説明責任や監督に関する課題など、完全にオープンではない AI がもたらす潜在的なリスクを特定するための今後の取組が必要となる。

こうした文脈において、「オープン」と称される AI に関する、より明確かつ共有されたビジョンは、不確実性を低減し、イノベーションを支援し、経済成長に寄与し、オープンウォッシングのような慣行を回避することで、大きな付加価値をもたらすだろう。

原則

原則1 – AI のオープン性はコミュニティ主導である

オープンソースの規範、標準及び定義の策定は、営利・非営利を問わず、これらの技術を構築・利用する開発者、研究者、科学者、スタートアップ、実務家などのコミュニティによって、長年にわたり形作られてきた。G7 加盟国は、オープン性が実務上何を意味するかを形作る上で、これらのコミュニティが果たす中心的な役割を認識する。また、既存の取組を評価するとともに、オープンな AI に関する共通基準の策定において、公的機関とオープンソース・コミュニティとの継続的な連携を奨励する。

原則2 – AI のオープン性はスペクトラムである

AI のオープン性は二元的なものではない。それは、制限付きライセンスの下で重みのみを共有するモデルから、オープンライセンスの下ですべての要素を完全に公開するモデルに至るまで、スペクトラム上に存在する。G7 加盟国及びステークホルダーは、「オープン」という用語を無条件に使用するのではなく、特定の AI にどの程度のオープン性が適用されるかを認識し、言及することが推奨される。

オープン性のスペクトルの性質を考えると、厳密な定義を網羅的に策定することは困難である。しかし、技術的な類型を確立する最初の試みは、G7 加盟国がそれをより正確に記述するのに役立つ。以下でそれらをさらに展開する。

原則3 – AI には複数の要素がある

AI のオープン性は、以下に挙げる要素（用語集参照）の組み合わせによって決定される。これには以下が含まれるが、これらに限定されない：

- 利用制限
- モデルの重み
- デプロイコード
- 学習コード
- 学習データ

利用可能になる各追加コンポーネントは、オープン性の度合いに影響を与え、それとともに、イノベーション、再現性及び強靱性の可能性にも影響を与える。

原則4 – AI のオープン性は適切なラベルによって記述される

AI に適用される「オープン」という用語について、一般的に受け入れられた定義が存在しない中で、関係者は、特定のモデルのオープン性の度合いを正確かつ透明性を持って反映する用語を使用することが推奨される。少なくとも、AI を「オープン」と表現する際は、「オープン」という用語を包括的な特徴付けとして使用したり、実際に提供されているものよりも高い度合いのオープン性を暗示したりするのではなく、どの部分がアクセス可能か、制限の有無があるかを明確に述べるべきである。

分類

以下の階層型分類は、AI のオープン性の異なる度合いを、最もオープンなものからよりオープンでないものへと順に記述したものである。オープン性はスペクトラム的な性質を持つため、この分類は網羅的なものではなく、純粋な二分法を超えた共通の参照点を提供することを目的としている。G7 加盟国は、AI について言及する際の共通の技術的参照として、この分類を採用することが推奨される。G7 は、AI を「オープン」と表現する際、アクセス条件、ライセンス条項、ドキュメント及び利用制限に関する透明性の重要性を強調する。

- **オープンデータ付きオープンソース AI**
-

オープンデータ付きオープンソース AI とは、モデルの重み、デプロイコード、学習コード及び完全な学習データを含め、オープンソースライセンスの下で無償で公開される AI を指す。

- **オープンソース AI**

オープンソース AI とは、モデルの重み、デプロイコード、学習コード、場合によっては完全な学習データを含め、オープンソースライセンスの下で無償で公開される AI を指す。

ただし、法的又は技術的な理由により共有が不可能な場合、全トレーニングデータの提供には例外が適用されることがある。そのような場合、利用できないデータに関するデータ情報（用語集参照）を代わりに提供する必要がある。

- **オープンウェイト AI**

オープンウェイト AI とは、オープンソースライセンス（用語集参照）の下で、その重みとデプロイコードとともに無料で公開される AI を指す。

- **ウェイト利用可能 AI**

ウェイト利用可能 AI とは、商用、地理的、または用途に関する制限など、使われ方に制限を含むライセンスの下で、重みとデプロイコードと共に無料で公開される AI を指す。

用語集

AI モデル：本文書において、AI モデルとは、予測及び推論を通じて結果を生み出す AI システムの構成要素を指す。

AI システム：本文書において、AI システムとは、さまざまなレベルの自律性をもって動作するように設計され、導入後に適応性を示す可能性があり、明示的又は暗黙的な目的のために受け取った入力から、物理的又は仮想的な環境に影響を与える可能性のある予測、コンテンツ、推奨又は決定といった出力を生成する方法を推論する、機械ベースのシステムを指す。

トレーニングデータ情報：モデルのトレーニングに使用されたデータに関する十分に詳細な情報。これには以下を含める必要がある：出所、範囲、特性、収集及び選択方法、ラベリング手順並びに処理及びフィルタリング手法を網羅し、共有不可能なデータを含むすべてのトレーニングデータの完全な説明；一般に公開されているすべてのトレーニングデータの一覧及びその入手先；並びに、有料を含む第三者から入手可能なすべてのトレーニングデータの一覧及びその入手先。

デプロイコード：AI を実行するために又は標準的な推論フレームワークとの互換性を確保するために必要なコード。

オープンソースライセンス：制限なく使用、検査、改変、再配布を許可するライセンス。認定されたオープンソースライセンスの一覧は、<https://opensource.org/licenses> 及び <https://www.gnu.org/licenses/license-list.html> で確認できる。

学習コード：学習データの処理及び AI の学習に使用されるコード。これには、データ処理及びフィルタリングコード、すべての引数と設定を含む学習コード、検証及びテストコード、トークナイザーやハイパーパラメータ探索コードなどのサポートライブラリ並びにモデルアーキテクチャの仕様が含まれる。

学習データ：AI の学習に使用されるデータ。

使用制限：モデルの使用方法を制限するライセンスの条件。商用、地理的、用途別の制限などを含む。

重み：学習を通じて生み出された、AI モデルの学習済み数値パラメータ。