Business Software Alliance

サイバーセキュリティ専門人材を最優先に

フロンティアAIサイバーセキュリティ・システムの戦略的かつ段階的な展開に向けた自主的な官民パートナーシップの発足によるデジタルエコシステムの強化とレジリエンスの向上

フロンティアAI（人工知能）を活用したサイバーセキュリティ・システムは、脆弱性の発見と修正を改善する重要な機会を生み出します。一方で、その能力が防御者による活用に先んじて広く利用可能になれば、新たなリスクが生じることにもなります。

こうした技術のメリットを最大化すると同時にリスクを管理するには、政府、産業界、およびサイバーセキュリティ・コミュニティが、これらのシステムの戦略的な段階的展開を支援する自主的な官民パートナーシップを構築すべきです。これによって信頼できるサイバーセキュリティ専門人材へのタイムリーなアクセスを確保するとともに、セキュリティ・プラットフォームを通じて広範な展開を促進することが可能となります。体系的で透明性が高く、国際的に整合した予見可能なフレームワーク（枠組み）の下でこれらの取り組みを調整することにより、管理、能力、影響の実績を持つステークホルダーは、より広範なデジタルエコシステムのセキュリティとレジリエンス強化に貢献することができます。

フロンティアAIサイバーセキュリティ・システムが生み出す機会とリスク

GPT-6 AstraやMythosのようなフロンティアAIサイバーセキュリティ・システムの最近の進歩は、サイバーセキュリティの状況を急速に変化させています。これらのシステムは、従来なら高度なスキルを持つ専門人材のチームを必要としたであろうスピードと規模で、自律的に脆弱性を発見し、悪用することができます。

防御者にとって、こうした能力は大きな機会を生み出します。フロンティアAIサイバーセキュリティ・システムは、脆弱性の発見と修正の迅速化、ソフトウェアセキュリティの向上、インシデント対応の強化を実現するとともに、攻撃者が脆弱性を悪用する前に脆弱性を特定するのに役立ちます。これらのシステムとそれらを使用するプラットフォームの採用が徐々に広がることで、デジタルエコシステムのセキュリティとレジリエンスが大幅に向上します。

その一方で、これらのシステムは新たなリスクも生み出します。防御者が脆弱性を発見・修正できるようにするのと同じ機能が、悪意のある行為者によって脆弱性の発見と悪用に利用される可能性もあります。当面の最も重要なサイバーセキュリティ目標の一つは、攻撃者に先んじて、信頼できるサイバーセキュリティ専門人材がこれらの機能を活用できるようにすることです。

これまで、最も高性能なフロンティアAIサイバーセキュリティ・システムへのアクセスは、主に個々の開発者による場当たり的な決定によって管理されてきました。このアプローチは、急速なイノベーションの現実を反映しており、開発者が確信を持って迅速に行動することを可能にしてきました。しかしながら、ステークホルダーにとって理解しにくく、関連するすべての利害を十分に考慮できない可能性があります。また、能力が進化し続ける中で必要とされる透明性、予測可能性、調整を確保することも難しくなります。

そのため、より体系的なアプローチが必要です。すなわち、個別断片的な決定に代えて、自主的な官民パートナーシップを構築すべきです。このような自主的な官民パートナーシップは、信頼できるステークホルダーに対し、これらのシステムを戦略的かつ段階的に展開するための、体系的で透明性が高く、国際的に整合した予見可能なフレームワークを提供することができます。その結果、デジタルエコシステムのセキュリティとレジリエンスの強化の実現につながります。

自主的な官民パートナーシップの目標

自主的な官民パートナーシップは、フロンティアAIサイバーセキュリティ・システムの利点を活用しながら、関連するリスクを管理するために、以下の二つの主要な目的を追求すべきです。

1. 管理、能力、影響の実績を有する組織が脆弱性を発見、優先順位付け、修正できるように、フロンティアAIサイバーセキュリティ・システムにタイムリーにアクセスできるようにする。

2. 組織が悪意のある行為者からの防御に既に利用しているセキュリティ・プラットフォームを通じて、フロンティアAIサイバーセキュリティ機能の広範な展開を促進する。

体系的で透明性が高く、国際的な整合性を備え、予見可能なフレームワークの下でこれらの取り組みを調整することで、管理、能力、影響の実績を有するステークホルダーは、より広範なデジタルエコシステムのセキュリティとレジリエンスの強化を実現することができます。

これらの目標を同時に達成することで、フロンティアAIサイバーセキュリティ・システムのメリットを可能な限り迅速かつ広範に実現すると同時に、防衛側が脆弱性に対処する前に悪意のある行為者が脆弱性を悪用する機会を減らすことができます。

これらの目的を達成するには、パートナーシップとその決定は以下のようであるべきです。

» 体系的で、プロセス、役割、基準が明確に定義されている

» 透明性があり、ステークホルダーが決定の経緯を理解できる

» 国際的に整合し、国境を越えた調整が可能である

» 予見可能であり、プロセスと結果を明確に見通せる

自主的な官民パートナーシップの活動

このパートナーシップの事前準備として、どのフロンティアAIサイバーセキュリティ・システムおよびプラットフォームに段階的展開のフレームワークを適用すべきか判断すべきです。大半のAIサイバーセキュリティ・システムは、サイバーセキュリティ環境を大幅に変更することなく、既存のワークフローを改善するため、段階的展開は必要ありません。しかし、一部のフロンティア・システムは、自律的に脆弱性の発見、優先順位付け、または悪用を大規模に行うことで、現在利用可能な機能を大幅に超える可能性があります。そのようなシステムが広範に利用可能になると、サイバーセキュリティ面での大きな機会をもたらす一方で、システミックリスクを増大させる可能性もあります。

したがって、それぞれの機能と潜在的な影響に基づいて、システムとプラットフォームの対象範囲を特定すべきです。対象を必要最小限に限定するアプローチは、段階的導入の仕組みが必要な場合にのみ適用されるようにしつつ、イノベーションを維持し、サイバーセキュリティを強化する技術の広範な導入を可能にします。

また重要なこととして、フロンティアであることは本質的に一時的な状態であることも認識すべきです。機能が進化し、同等かそれに近いシステムが利用可能になると、サイバーセキュリティの成果を最大化するための段階的展開がもはや必要なくなることもあります。したがって、段階的展開を必要とする機会やリスクがなくなり次第、システムをその対象範囲から速やかに除外する明確な仕組みを確立すべきです。この動的なアプローチにより、イノベーションに対する不必要な制約やサイバーセキュリティ技術の広範な普及の遅延を回避しながら、パートナーシップが真のフロンティアAIサイバーセキュリティ・システムに焦点を当て続けることを可能とします。

特別の配慮を必要とするシステムを明確に特定することで、官民パートナーシップはイノベーションに対する不必要な制約を回避しながら、サイバーセキュリティとレジリエンスに最も大きな効果をもたらす領域に取り組みを集中させることができます。

戦略的な段階的展開を必要とするフロンティアAIサイバーセキュリティ・システムについては、パートナーシップは、組織の管理、能力、潜在的影響の包括的な評価に基づいて、その組織にアクセスを認めるかどうか、認める場合はそのタイミングを決定すべきです。この評価では以下の点を考慮すべきです。

» 管理要因（所有や支配関係における信頼性や透明性など）

» 能力要因（技術的能力、組織的能力、協調能力など）

» 影響要因（使命、影響範囲、公益など）

これらの要因を総合的に検討することで、これらのシステムを責任を持って利用し、その効果的な利用に必要な能力を有し、サイバーセキュリティとレジリエンスを有意に強化できる立場にあると信頼できる人材に、早期アクセスを確実に提供することが可能になります。

フロンティアAIサイバーセキュリティ・システムの戦略的かつ段階的展開を実現するパートナーシップは万人に恩恵をもたらす

自主的な官民パートナーシップへの移行により、ステークホルダーはこれらのシステムの展開に関連するリスクを責任を持って管理しながら、利益が広く享受されるようにすることができます。また、アクセスに関する決定をサイバーセキュリティと公益目的の両方の確実な前進に役立てることができます。

このパートナーシップは、これらの強力なシステムをその活用に最適な立場にある人々の手に届けることで、以下の実現につながります。

» 明確で一貫性があり、説明可能なアクセス基準と意思決定プロセスを通じて信頼を構築する

» 脆弱性の発見、優先順位付け、修正の迅速化により防御能力を強化する

» 脆弱性が広く悪用される前に特定と対処を可能にすることでシステミックリスクを低減する

» 開発者に予見可能なガバナンスのフレームワークを提供し、リスク管理を明確化する

» 志を同じくするパートナー国の国際的な整合性を促進し、各国のアプローチの分断を縮小する

フロンティアAIサイバーセキュリティ・システムは、脆弱性の発見と修正の迅速化、ソフトウェアセキュリティの向上、インシデント対応の強化を実現するとともに、攻撃者が脆弱性を悪用する前に脆弱性を特定するのに役立ちます。

この官民パートナーシップを構築し、サイバーセキュリティを強化するにはステークホルダー間の協力が不可欠

機会は膨大です。フロンティアAIサイバーセキュリティ・システムを効果的に展開すれば、脆弱性の発見と修正の大幅な改善と、最終的にセキュリティとレジリエンス全体を強化することができます。しかし、これらの成果が保証されているわけではありません。最も有能な防御者が最も強力なツールに優先的にアクセスできることを確実にすることは、政府と産業界の責務です。

自主的な官民パートナーシップの発足には、政府、AI開発者、サイバーセキュリティ・コミュニティの協力が必要になります。単独でリスクを管理できるステークホルダーは存在しません。しかし、すべてのステークホルダーが協力すれば、フロンティアAIサイバーセキュリティ・システムのメリットを実現できます。

フロンティアAIサイバーセキュリティ・システムの戦略的かつ段階的展開を管理するための自主的な官民パートナーシップに関するBSAフレームワーク

BSAは、パートナーシップの決定を純粋な定量的公式に要約できないことを認識しています。それには、関連するさまざまな基準を慎重に検討し、個々のケースの事情を総合的に勘案する、包括的かつ多面的な評価が求められます。とはいえ、パートナーシップの運営方法、組織にアクセスを提供すべきかどうか、提供する場合はそのタイミングに関して明確な基準を特定することは、より良い成果の促進と制度への信頼構築につながるはずです。

プログラム設計（PD）：自主的プログラムは予見可能な形で運営されるべきである。

» **目的（PD.PU）**：フロンティアAIサイバーセキュリティ・システムの段階的リリースに関するパートナーシップの目的および優先事項を設定し、理解を得ること。

- **PD.PU-1**：パートナーシップは、これらのシステムの段階的リリースを可能にすべきである。その際、サイバーセキュリティとレジリエンスの面で最大の利益をもたらす可能性が高い組織や分野へのアクセスを優先すべきである。
- **PD.PU-2**：パートナーシップは最終的に、自らのサイバー防御を強化する組織と顧客の防御を強化する組織の両方のアクセスを可能にすべきである。

» **国際的整合性 (PD.GA)**：パートナーシップは、フロンティアAIサイバーセキュリティ・システムへの一貫したアクセスと、法域を越えた協調的アプローチを促進するために、国際的な整合性を確保すべきである。

- **PD.GA-1**：パートナーシップは、志を同じくする法域に本社を置き、適用される基準を満たした組織に対して開かれたものであるべきである。
- **PD.GA-2**：パートナーシップは、フロンティアAIサイバーセキュリティ・システムの段階的リリースに関するアプローチにおいて、参加国政府間の調整を促進すべきである。

» **透明性 (PD.TR)**：パートナーシップのプロセス、基準、決定、スケジュールは、ステークホルダーにとって明確で、理解しやすいものであるべきである。

- **PD.TR-1**：パートナーシップは、組織の参加方法とリリース前のシステムへのアクセスに関する申請方法について明確な情報を提供すべきである。
- **PD.TR-2**：パートーナシップは、意思決定の際に適用するプロセスと基準について明確に説明すべきである。
- **PD.TR-3**：パートナーシップは、アクセスを認めない理由を明確に説明すべきである。
- **PD.TR-4**：パートナーシップは、新しいシステムのリリース前の段階について、どのくらいの期間が想定されているか明確な情報を提供すべきである。

» **アクセス可能性 (PD.AC)**：パートナーシップは、フロンティアAIサイバーセキュリティ・システムのメリットを、資格要件を満たした組織が、そのニーズ、潜在的影響、サイバーセキュリティの改善能力に基づいて享受できるようにすべきである。

- **PD.AC-1**：パートナーシップは、資格要件を満たした組織によるフロンティアAIサイバーセキュリティ・システムの利用を妨げる、コストをはじめとする障壁を取り除く方法を検討すべきである。
- **PD.AC-2**：パートナーシップは、オープンソースソフトウェアを保護する取り組みや、脆弱性の発見と修正の調整と重複排除など、影響の大きいサイバーセキュリティ用途を支援すべきである。

管理 (ST)：アクセスを認められた組織は、フロンティアAIサイバーセキュリティ・システムを責任を持って利用できる信頼性を備えていることを示すこと。

» **信頼性 (ST.TU)**：組織は、サイバー防御の強化を目的にフロンティアAIサイバーセキュリティ・システムを利用するという確信を与えるために、責任ある行動と実践の記録を示すこと。

- **ST.TU-1**：組織は、特にデータ保護、サイバーセキュリティ、AIガバナンスに関連する、責任ある行動と適用される法律の遵守の実績を有しているか？
- **ST.TU-2**：組織は、脆弱性の開示や修正を含め、ベストプラクティスや国際的に認められた標準を採用しているか？

- **ST.TU-3**：組織は国際的に認められた適切な認証を有しているか？
- **ST.TU-4**：組織は、フロンティアAIサイバーセキュリティ・システムへのアクセスに関連するリスクを生み出す可能性のある懸念国との関係、事業活動、所有関係、またはその他のつながりを有しているか？
- **ST.TU-5**：組織は、これらのシステムへの意図しないアクセスを制限できること、およびそれらが本来の目的で使用されるという高い信頼性を備えているか？

» **透明性（ST.TA）**：組織の所有、支配、実質的な影響に関し、アクセスの決定に必要な情報を得るのに十分な透明性があること。

- **ST.TA-1**：組織は、当該組織を所有または支配する、あるいは当該組織に実質的な影響を与える主体をパートナーシップが特定するために、所有や支配について十分な透明性を提供しているか？

能力（CA）：アクセスが認められる組織は、フロンティアAIサイバーセキュリティ・システムを効果的に利用してサイバー防御を強化するために必要な技術的能力、組織的能力、協調能力を有すること。

» **技術的能力（CA-TC）**：組織は、フロンティアAIサイバーセキュリティ・システムの効果的な利用に必要な技術的専門知識と経験を有すること。

- **CA.TC-1**：組織は、高度なサイバーセキュリティ・ツールやテクノロジーを利用してサイバー防御を強化する実証済みの専門知識と経験を有しているか？

» **組織的能力（CA.OC）**：組織は、フロンティアAIサイバーセキュリティ・システムの効果的な展開に必要なリソースと運用能力を有すること。

- **CA-OC-1**：組織は、フロンティアAIサイバーセキュリティ・システムを利用してサイバー防御を有意に強化するために必要な規模、リソース、ガバナンス、運用成熟度を備えているか？

» **協調能力（CA.CC）**：組織は、脆弱性の開示や修正の調整に関して、関連するステークホルダーとの信頼関係と実証済みの経験を有すること。

- **CA.CC-1**：組織は、ソフトウェア開発者、オープンソースソフトウェアコミュニティ、脆弱性研究者、顧客、およびその他のステークホルダーと協力して、組織や国の枠を超えて脆弱性の開示や修正を調整した実績を有しているか？

影響（IM）：アクセスを認められる組織は、フロンティアAIサイバーセキュリティ・システムを利用して、サイバーセキュリティとレジリエンスを有意に強化できる立場にあること。

» **使命（IM.MI）**：組織の使命と活動がサイバー防御の強化に貢献すること。

— **IM-MI-1**：組織は、製品、サービス、システム、またはデータの保護、サイバーセキュリティ・サービスの提供、脆弱性調査の実施、ソフトウェアセキュリティの改善、インシデント対応、重要インフラの保護、あるいはその他の活動を含め、サイバーセキュリティ防御の成果に貢献した実証済みの実績を有しているか？

» **影響範囲（IM.SR）**：組織は、フロンティアAIサイバーセキュリティ・システムを利用して、サイバーセキュリティを大幅に改善できる立場にあること。

— **IM.SR-1**：組織は、重要インフラ、広く使用されているテクノロジー、またはサイバーセキュリティに重大な影響を与えるその他のシステムを保護しているか？

— **IM-SR-2**：相当数またはタイプの個人、企業、重要インフラ事業者、その他の組織が組織の製品やサービスを利用しているか？

» **公益（IM.PB）**：組織がフロンティアAIサイバーセキュリティ・システムを利用することで、当該組織自体を超えてサイバーセキュリティ上のメリットが生まれる可能性が高いこと。

— **IM.PB-1**：組織の立場を考えた場合、当該組織がこれらのシステムにアクセスすることで、脆弱性の開示や修正、オープンソースソフトウェアの保護、またはその他の形での広範なデジタルエコシステムのセキュリティとレジリエンスの強化などを通じて、より広範なサイバーセキュリティ上のメリットを生み出すことになる可能性が高いか？