フィジカルAIに対するサイバーセキュリティ対策：統合レポート

Librus株式会社
事業開発部

1. フィジカルAIとは何か

フィジカルAI(Physical AI)とは、AIがセンサーなどを通じて現実世界（物理空間）を理解し、ロボットなど物理的な実体を伴って自律的に行動する技術の総称である。従来のデジタル空間で動作する生成AIとは異なり、フィジカルAIは現実世界と直接相互作用する点が特徴であるNTT。自動運転車、人型ロボット、産業用ロボット、ドローンなど、様々な形態で実装され、製造、物流、医療、家庭サービスなど幅広い分野での活用が期待されている。

2. フィジカルAIの脆弱性と脅威の構造

2.1 外部起因の脆弱性（Exogenous Vulnerabilities）

最新の学術研究によれば、フィジカルAIシステムは外部環境と内部システムの両面から脆弱性に晒されている。arXivで公開された論文"Towards Robust and Secure Embodied AI"（2025年2月）は、フィジカルAIの脆弱性を体系的に分類している。

物理攻撃（**Physical Attacks**）: センサー欺瞞攻撃が代表的である。カメラ、LiDAR、レーダーなどの知覚システムに対する敵対的攻撃により、物体認識を誤らせることが可能である。例えば、特殊なパターンのステッカーを貼付することで自動運転車の画像認識システムを欺く攻撃が実証されている。

サイバーセキュリティ脅威: 無線通信の脆弱性が深刻な問題となっている。2025年9月に発見されたUnitree社製ロボット（Go2、G1、H1、B2）の脆弱性（CVE-2025-35027、CVE-2025-60017）は、Bluetooth Low Energy（BLE）とWi-Fi設定における重大な欠陥を露呈したIEEE Spectrum。これらの脆弱性により、コマンドインジェクション、ルートOSアク

セス、ハードコード化された暗号鍵の悪用が可能となり、攻撃者はロボットを完全に制御できる状態であった。

2.2 内部起因の脆弱性(**Endogenous Vulnerabilities**)

センサー障害とソフトウェア欠陥: 内部システムレベルの脆弱性として、センサーの故障、ソフトウェアのバグ、ファームウェアの欠陥が挙げられる。これらは環境認識の精度低下や制御システムの誤動作を引き起こす。

AIモデルへの攻撃: 大規模視覚言語モデル(LVLMs)や大規模言語モデル(LLMs)を組み込んだフィジカルAIシステムは、ジェイルブレイク攻撃や命令の誤解釈による脆弱性を抱えている。2026年現在、AI関連の脆弱性は最も急速に増加しているサイバーリスクであり、調査対象組織の87%がこれを認識している。

3. 具体的な脅威シナリオと実例

3.1 Unitree G1人型ロボットの脆弱性事例
2025年9月、VicOne Lab R7とセキュリティ研究者により、Unitree G1人型ロボットに3つの重大な無線脆弱性が発見された。この事例は商用人型ロボットプラットフォームの初の大規模公開エクスプロイトとして記録されている。
日経クロステック
https://xtech.nikkei.com/atcl/nxt/column/18/02801/101500027/

攻撃シナリオとして、以下が実証された:

- Bluetooth経由での不正アクセスによるロボット制御の乗っ取り
- センサーデータの窃取とプライバシー侵害
- 悪意のあるコマンドの実行による物理的な危害の可能性
- ロボット間のウイルス感染の連鎖(ロボット・ツー・ロボット感染)

さらに重大な問題として、収集されたデータが中国のサーバーに無断で送信されていた可能性が指摘され、データ主権とスパイ活動の懸念が浮上している。

3.2 AIを活用した自律的サイバー攻撃

AIによるサイバー攻撃の高度化も顕著である。2026年の予測では、AIエージェントを活用した攻撃が、従来の数日かかっていた攻撃面のマッピングを数分で完了し、自律的なエクスプロイテーションを実行できるようになっている。

Lazarus Alliance

https://lazarusalliance.com/the-biggest-cybersecurity-threats-of-2026/

4. 攻撃対象領域（Attack Surface）の拡大

フィジカルAIシステムの攻撃対象領域は、従来のPCやスマートフォンとは根本的に異なる。VicOne CEOの指摘によれば、意思決定を司るAIモデル、センサー入力、アクチュエータ制御、無線通信、クラウド接続など、多層的な攻撃ベクトルが存在する。

自動運転車を例にとると、以下の攻撃面が存在する：

- V2X（Vehicle-to-Everything）通信の脆弱性
- LiDARやカメラへの物理的攻撃
- GPSスプーフィング
- ECU（電子制御ユニット）への侵入
- OTA（Over-The-Air）アップデートの改ざん

5. サイバーセキュリティ対策の体系的アプローチ

5.1 設計段階のセキュリティ（**Security by Design**）

多層防御アーキテクチャ：VicOne Lab R7が提唱する統合検証プロセスは、システムレベルとAIロジックを一括で保護するアプローチであるVicOne。出荷前の包括的なセキュリティスキャンにより、システムとAIモデルの隠れた脆弱性を可視化し、優先的に対処すべきポイントを明確化する。

セキュアな通信プロトコル：認証と暗号化の強化が必須である。World Economic Forum Global Cybersecurity Outlook 2026によれば、64%の組織がAIツールをデプロイする前にセキュリティ評価プロセスを導入しており、これは前年の37%から大幅に増加している。

5.2 運用段階の防御

ランタイム防御と**R-SOC**：VicOne Lab R7のRthenaは、フィジカルAI専用に設計されたR-SOC（ロボティクス・セキュリティ・オペレーションセンター）を提供し、継続的な監視と迅速な対応を実現する。全方位型ランタイム防御エージェントにより、OTA（Over-The-Air）アップデートのなりすまし、モデル改ざん、センサー乗っ取りなどの脅威に対処する。

ゼロトラストアーキテクチャ：World Economic Forumの報告書"AI Agents in Action"では、AIエージェントの資格情報、権限、相互作用を人間ユーザーと同様に管理する必要性が強調されている。すべての相互作用をデフォルトで信頼しないゼロトラスト原則に基づく継続的な検証、監査証跡、堅牢なアカウンタビリティ構造が不可欠である。

5.3 AIモデルの完全性保護

データポイズニング対策：AIモデルのトレーニングデータへの攻撃を防ぐため、データの完全性検証と異常検出が重要である。87%の組織がAI関連の脆弱性を2025年に最も急速に成長したサイバーリスクと認識している。

モデルの検証とモニタリング：AIモデルの読み込み時および実行時における完全性保護が不可欠である。攻撃者は軽量かつ効率的に動作する特化型モデルを使用し、クラウドの計算資源を活用してより複雑な攻撃を仕掛けることが可能であるVicOne。

6. 規制とコンプライアンスの動向

6.1 国際標準とフレームワーク

ISO/IEC 42001:2023：世界初のAIマネジメントシステム国際標準であり、AIシステムのリスク管理、影響評価、ライフサイクル管理、サードパーティサプライヤー監視を規定している。

従来のロボット安全規格：数十年にわたり、ロボット工学の安全性の基盤はISO 13849-1（パフォーマンスレベル）とIEC 61508（SIL - 安全整合性レベル）であった。フィジカルAIシステムがこれらの規格に準拠できるかが重要な課題となっている。

6.2 地域別規制の差異

EU（欧州連合）：

- CRA（Cyber Resilience Act）：2027年12月11日全面適用予定。製品安全と長期的なソフトウェア責任を連動
- AI規則（AI Act）：高リスクAIに対するリスクベースの規制

米国：

- NISTフレームワーク、SBOM（Software Bill of Materials）開示、IoT Cyber Trust Markの任意ラベリングによる市場の透明性重視

中国：

- サイバーセキュリティ法（CSL）、データセキュリティ法（DSL）、個人情報保護法（PIPL）、MLPS 2.0によるデータ主権重視
- GB/T 45502-2025（サービスロボット向け基準、2025年10月1日施行）

VicOne Lab R7のプラットフォームは、CRA準拠からSBOMの自動化まで、AIロボットの出荷段階から安全性・準拠性・監査対応を支援している。

7. 組織的対策とベストプラクティス

7.1 リスクの可視化と優先順位付け
VicOne Lab R7は「リスクの見える化」から始めることを推奨している。可視化がなければ、影響度の小さい領域にリソースを割いてしまい、本当に業務を止めかねない脆弱性を見逃す可能性がある。ワンストップ型サイバーセキュリティスキャンプラットフォームにより、システムとAIの隠れた脆弱性を可視化し、優先的に対処すべきポイントを明確化する。

7.2 脅威インテリジェンスと情報共有
地政学的な不安定性がサイバーセキュリティを再定義している。World Economic Forum Global Cybersecurity Outlook 2026によれば、64%の組織が地政学的に動機付けられたサイバー攻撃を全体的なサイバーリスク軽減戦略に考慮している。

高レジリエンス組織のCEOの52%が国家アクターに関する脅威インテリジェンスを優先し、48%が政府機関や情報共有グループとの連携を強化している。これに対し、レジリエンスが不十分な組織のCEOではそれぞれ13%、6%にとどまっている。

7.3 サプライチェーンセキュリティ
高レジリエンス組織のCEOの78%がサプライチェーンとサードパーティの依存関係をレジリエンス強化の最大の課題と認識している。対策として、70%がセキュリティ機能を調達プロセスに統合し、59%がサプライヤーの成熟度評価を優先している。

8. 人材とスキルギャップの課題

8.1 サイバーセキュリティ人材不足
World Economic Forumの調査によれば、54%の組織がAIをサイバーセキュリティに活用するための知識・スキルの不足を実装の障壁として挙げているWEF。

地域別では、サハラ以南アフリカ（70%）とラテンアメリカ・カリブ海地域（69%）のCEOが、現在のサイバーセキュリティ目標を達成するためのスキルが不足していると認めている。

8.2 AIリテラシーの重要性
World Economic Forumの"The Future of Jobs Report 2025"によれば、「ネットワークとサイバーセキュリティ」は2030年に向けて最も急速に成長するスキルの上位3つに入っている（AI・ビッグデータ、テクノロジーリテラシーとともに）。AIは人間の専門知識を置き換えるのではなく、専門家が戦略的監督、ガバナンス、ポリシーに焦点を移し、日常的な運用タスクを自動化に委任することを可能にしている。

9. 業界別の対策動向
AIツールをサイバーセキュリティ能力の強化に採用する動きは業界によって異なる。エネルギーセクターは侵入・異常検知を重視（69%）、素材・インフラセクターはフィッシング保護を優先（80%）、製造・サプライチェーン・輸送セクターは自動化されたセキュリティ運用の利用が多い（59%）。

10. 今後の展望と提言

10.1 協調的アプローチの必要性

フィジカルAIのセキュリティは、単一の組織や国だけでは確保できない。VicOne Lab R7とDecloak Intelligenceの戦略的パートナーシップのように、ファームウェア、通信、AIモデルの完全性、センサーのプライバシー制御に至るまで、多層的かつ包括的なサイバーセキュリティを実現する業界横断的な協力が不可欠である。

10.2 プロアクティブな防御戦略

World Economic Forumは、サイバーセキュリティの未来は今日の選択に依存すると強調している。先見性、能力、イノベーションへの投資、業界、セクター、国境を越えた協力の強化により、ボラティリティを推進力に変え、より安全でレジリエントなデジタル未来を共に構築できる。

10.3 継続的なモニタリングと適応

セキュア・バイ・デザインは「出発点」であり、「ゴール」ではない。AIロボットが通信・センサー・学習モデルを備えたフィジカルAIへと進化する中で、出荷後に新たな攻撃対象領域が生まれる。運用時の継続的な監視、OTA署名の検証、モデルの完全性チェックなど、現場での安全を保つためには継続的なサイバーセキュリティ対策が不可欠である。

結論

フィジカルAIは、AIがサイバー空間から物理世界へと飛び出す歴史的な転換点を示している。しかし、この技術革新は同時に、前例のないサイバーセキュリティの課題をもたらしている。2025年のUnitree G1の脆弱性事例が示すように、わずかな設計上の見落としが数千台のロボットを危険に晒す可能性がある。

エビデンスベースの分析から明らかなのは、フィジカルAIのセキュリティには、設計段階からのセキュリティ組み込み、運用段階での継続的監視、AIモデルの完全性保護、サプライチェーンセキュリティ、人材育成、国際協力という多面的なアプローチが必要だということである。地政学的不安定性、AI脆弱性の急増、サイバー対応型詐欺の増加という2026年の脅威環境において、組織は技術的対策と戦略的ガバナンスを統合し、レジリエンスを構築する必要がある。

フィジカルAIの時代において、サイバーセキュリティはもはやバックオフィスの技術的機能ではなく、政府、企業、社会にとっての中核的な戦略的関心事である。私たちの選択が、より安全で信頼できるフィジカルAI社会の実現を左右するのである。

監修者:
鎌田光一郎:青山学院大学法学部卒業。SMBC日興証券株式会社にて証券営業、経営管理業務に従事したのちPwCコンサルティンク゛合同会社に転籍。金融機関に対するコンサルティンク゛業務に従事。その後、Libras株式会社を設立、代表取締役に就任。

お問い合わせ先
Librus株式会社(代表取締役 鎌田光一郎)
105-0004東京都港区新橋6丁目13-12 VORT新橋Ⅱ 4F
03-6772-8015
お問い合わせフォーム
https://librus.co.jp/contact