

経済産業省が新設した「事業者における AI エージェントの安全な利活用に関するサブワーキンググループ」委員に代表の徳永が就任し AI エージェントのセキュリティガイドライン策定に参画 ～産官学・法曹の有識者とともに、国内の AI エージェント利活用の安全確保に貢献～

Polaris.AI 株式会社（本社：東京都港区、代表取締役 CEO：徳永優也）は、当社代表取締役 CEO の徳永優也が、経済産業省「産業サイバーセキュリティ研究会」のワーキンググループに新設される「事業者における AI エージェントの安全な利活用に関するサブワーキンググループ」の委員に就任したことをお知らせします。本サブワーキンググループは、AI エージェントを利用する企業が実務において参照できる「AI エージェントセキュリティガイドライン」の策定に向けて、専門的な検討・分析・助言を行う場になります。

背景：急拡大する AI エージェントとセキュリティの課題

企業における AI エージェントの活用が急速に広がり、国内市場は 2030 年までに数兆円規模へ拡大すると見込まれています。AI エージェントは、外部システムとの接続や業務の自律的な実行など、従来の生成 AI とは異なる形で実世界に影響を及ぼします。その自律性ゆえに、意図しない操作やデータ漏えいを引き起こす可能性があり、サイバー攻撃を受けた際の影響も大きいという固有のリスクを抱えています。こうしたリスクに対し、AI エージェントを利用する企業が実務において参照できるガイドラインは、現時点では十分に整備されていないのが実情です。

ガイドラインと本サブワーキンググループについて

経済産業省は、令和 8 年度「産業サイバーセキュリティ対策の強化に向けた環境整備事業」の一環として、AI エージェントの導入企業が実務で参照できるガイドラインの策定を予定しています。ガイドラインの検討方向性は以下のとおりです。

- ① AI エージェントの特徴やリスク
- ② 自律性と行動範囲に応じた統制レベル
- ③ AI ガバナンス・セキュリティ対策

「事業者における AI エージェントの安全な利活用に関するサブワーキンググループ」は、「産業サイバーセキュリティ研究会」のワーキンググループ内に設置され、学术界・法曹界・政府機関・産業界などの多様な有識者が委員として参画（※1）します。

（※1）委員名簿：

https://www.meti.go.jp/shingikai/mono_info_service/sangyo_cyber/wg_seido/wg_ai_agent/pdf/001_02_00.pdf

Polaris.AI 代表 徳永の参画：AI スタートアップの経営者として

様々なカテゴリーの有識者が名を連ねる本グループにおいて、Polaris.AI 代表の徳永は産業界（AI エージェントの開発・提供者）の一員として、また AI スタートアップの経営者として参画します。Polaris.AI はこれまで、AI エージェント・AI セキュリティ・ロボティクスを含む多数の AI 開発・社会実装を手がけてきました。AI の研究開発における最前線の知見と、実際に企業の現場で AI を実装してきた知見の双方をあわせ持つスタートアップの立場から、利用企業が実務で使えるガイドラインづくりに貢献します。

代表取締役 CEO 徳永優也 コメント



AI エージェントは、これまでの生成 AI のように「答える」だけでなく、企業のシステムやデータに接続し、一定の権限を持って実際の業務を動かす存在になりつつあります。その分、制御を誤れば、誤った判断がそのままシステム上の操作につながったり、情報流出や誤操作を引き起こしたりするリスクがあります。だからこそ、「何を任せるのか」「どこまで権限を与えるのか」「異常をどう検知し、止めるのか」まで含めた設計が重要です。また、リスクがあるから使わないのではなく、適切に制御しながら、安全に任せられる範囲を広げていくことも重要だと考えています。私たちは、実際に

AI エージェントを開発し、企業の業務に組み込んできた経験を生かし、「何を、どこまで AI に任せてよいのか」を企業が具体的に判断し、実装に落とし込めるガイドラインづくりに貢献していきます。

今後について

Polaris.AI は、本サブワーキンググループでの活動を通じて、AI エージェントの安全な利活用の推進に寄与するとともに、間違いの許されない Mission Critical 領域における AI の社会実装に、引き続き取り組んでまいります。

関連リンク

- ・ 経済産業省 産業サイバーセキュリティ研究会：
https://www.meti.go.jp/shingikai/mono_info_service/sangyo_cyber/index.html
- ・ 第 1 回 産業サイバーセキュリティ研究会 ワーキンググループ 1（事業者における AI エージェントの安全な利活用に関するサブワーキンググループ）：
https://www.meti.go.jp/shingikai/mono_info_service/sangyo_cyber/wg_seido/wg_ai_agent/001.html

Polaris.AI 株式会社について

Polaris.AI 株式会社は、東京大学松尾研究室発のスタートアップ（※2）です。「AI 時代の羅針盤（Polaris）になる」というビジョンのもと、製造業・インフラ・官公庁を中心に、間違いの許されない Mission Critical 領域における AI の社会実装を推進しています。

機械学習・アルゴリズム・ロボティクス・セキュリティ・大規模システム開発・産業ドメインにわたる専門性を持つ研究者・エンジニアが集結し、課題の見極めから技術選定・開発・本番運用まで、多数の AI プロジェクトを一気通貫で手がけてきました。主な事業は以下のとおりです。

- ・ **オーダーメイド AI 開発**
確立した解き方のない固有課題に対する、設計から現場定着までの AI 開発
- ・ **AI 導入・活用コンサルティング**
経営アジェンダとしての AI 戦略策定から、全社への定着・内製化まで
- ・ **研究開発支援／共同研究**
ロボット基盤モデル（VLA）、エッジ AI、AI セキュリティ等の先端領域における研究開発

私たちは、デモの場で動く AI ではなく、現場で使われ続ける AI を届けることを信条とし、お客様の競争力と社会の基盤となる成果の実現に取り組んでいます。

（※2）松尾研発スタートアップとは、松尾研出身者が創業または松尾研の支援を受け創業された企業の内、技術・事業力共に成長可能性が認められ、且つ松尾研の理念に共感し共に後進の育成に取り組む、選抜されたスタートアップ企業群です。（登録商標第 6667237 号、第 6710069 号）

社名：Polaris.AI 株式会社

所在地：東京都港区虎ノ門 2-2-1 住友不動産虎ノ門タワー13 階

代表者：徳永 優也

URL：<https://polarisai.co.jp/>

< 本件に関するお問い合わせ先 >

メール：biz@polarisai.co.jp

お問い合わせ用リンク：<https://polarisai.co.jp/contact>