knowbe4

May 2026

エージェンティックAIのリスクを「人間の強み」へ

なぜ今、セキュリティ文化なのか？―エージェンティックAI時代を生き抜く組織の条件

まえがき

常識は変わった。

ここ数十年の間、サイバーセキュリティは「人間という変数を抱えた、テクノロジーの問題」と捉えられがちでした。人を教育し、システムにパッチを適用する。その繰り返しでした。しかし、エージェンティックAI（自律型AI）の登場により、その常識は根底から書き換えられました。もはや、問題は「脅威からいかに人間を守るか」ではありません。「人間とAIエージェントがいかに協調し、信頼関係を築き、一刻を争う重大な意思決定をリアルタイムで行っていくか」ということです。そんな世界において、自律型AIは私たちの「チームメイト」であり、「意思決定者」であり、同時に悪用されてしまう「攻撃対象（アタックサーフェス）」でもあります。

また、多くの組織は、そこまで考えが及んでいません。しかし、本調査（研究）が示しているのは、この問題に対して正しい方向へと進んでいる組織には「ある重要な共通点」があるということです。それは、彼らがセキュリティを単なる「一機能（部署）」として扱うのをやめ、「組織の文化」として捉え始めたという点です。安全な行動は日々の業務の中に組み込まれており、従業員はただ規則に従う（遵守する）だけでなく、自ら進んでセキュリティに貢献しています。これらの組織は、目の前の脅威を追いかけるのをやめ、脅威を取り巻く「環境そのもの」をコントロールし始めているのです。

先進的な組織と、それ以外の組織との格差は、急速に縮まりつつあります。それだけに、対応を誤った組織（遅れた組織）が支払う代償は、急速に跳ね上がっています。

私たちKnowBe4は、「人間は最大の弱点（最弱のリンク）ではない。むしろ、組織が持つ最強の防御陣である」と断言してきました。この考えに今も疑いはありません。今やこの考えは人間の傍らで働くAIエージェントにまで及んでいます。つまり、私たちのミッションは、人間とエージェンティックAI（自律型AI）が共に手を取り合って強くなれるような、組織文化と信頼関係の構築を支援するということです。組織で働くすべての人が、その未来に向けて自律的かつ柔軟に取り組むことができる存在になれる。私たちは、それを確実に実現するためにあるのです。

次世代のサイバーセキュリティを定義していく組織は、より優れた脅威モデルが登場するのをただ待っているわけではありません。彼らは、より優れた人材を育て、より優れた文化を築き、そして隣で働くAIとの間に、より優れたパートナーシップを構築しているのです。本レポートは、その第一歩をどこから踏み出すべきかを示しています。

この調査は、未来を予測するものではありません。すでにその未来を生きている組織の姿を描き出したものです。重要なのは、あなたがその未来に到達できるかどうかではありません。「誰よりも先にそこに到達できるか」です。

エージェンティックAIがもたらすリスクを、人間の「勝利」へと変える未来を、共に築いていきましょう。

Perry Carpenter

ペリー・カーペンター
チーフデセプションストラテジスト

本書の構成

はじめに

サイバーセキュリティは、ツールを導入すれば解決されるという単なるテクノロジーの問題ではありません。それは常に、「人がテクノロジーをどう扱うか?」その交差点に存在しています。そしてこの交差する場所こそ、ツールだけでは解決できない形で、人間の行動がリスクを引き起こす舞台となるのです。AIはこれらの課題をさらに加速させ、日々の業務におけるリスクの現れ方を根底から変えつつあります。

サイバーリスクにおいて「人間の行動」が果たす役割の大きさは、セキュリティリーダーたちの間でも常に認識されてきました。それにもかかわらず、データ侵害がなくならないのは、この認識を「現実の複雑な状況」に当てはめて実することが難しいためです。プレッシャーや注意散漫、そして不確実性の中で下される日々の決断から、リスクは次々と生まれ続けています。

さらに、働く環境そのものが変化していることで、この課題は現在進行形で深刻化しています。従業員はもはや、一人で業務を行うわけではありません。コンテンツを生成し、自ら実行し、機密データと相互にやり取りを行う「エージェンティックAI」や「自律型エージェント」と共同で働いています。これにより、人間の行動(およびリスクに対する姿勢)と、エージェンティックAIの行動が複雑に交差する、全く新しい環境が生まれているのです。

本調査はこの「大きな転換」について深く掘り下げています。私たちは世界各国のサイバーセキュリティリーダー800名と、従業員3,200名を対象に調査を実施しました。人間とAIエージェントがもたらすリスク、そして次の時代への備えができているかについて、明らかになったことは以下の通りです。

58%
のセキュリティリーダーが、日々の「うっかりミス(日常的な過ち)」がサイバーセキュリティリスクの主な要因であると回答しています。

42%
のセキュリティリーダーが、今後の人的サイバーセキュリティリスクにおいて、「AIを悪用した攻撃」が最大の要因になると予測しています。

36%
過去12ヶ月の間に、人的リスクの管理が「大幅に改善された」と回答したセキュリティリーダーは、わずか36% にとどまっています。

97%
のセキュリティリーダーが、人的サイバーセキュリティリスクを追跡するための指標(メトリクス)を活用していますが、それが自社の取り組みを「完全に支えている」と答えたのは 61% にすぎません。

42%
セキュリティ意識の向上(トレーニングなど)「だけ」で、持続的な行動変容をもたらすことができると信じているセキュリティリーダーは 42% でした。

89%
従業員の 89% は、セキュリティ上のミスを報告する際、心理的安全性を感じていると回答しています。

19% 自社が「統合され、組織文化に深く根ざしたアプローチ」で人的サイバーセキュリティリスクを管理できていると報告したセキュリティリーダーは、わずか 19% でした。

▶ (リスク管理において)「統合され、組織文化に深く根ざしたアプローチ」を取り入れている組織は、そうでない組織に比べて、次のような傾向がより高くなっています。それは、セキュリティ意識向上のための活動が「ITチームと従業員との間の信頼関係を築くものである」と捉えていること、また「(ミスをした際などの)報告に関する明確なガイドラインを提供している」こと、そして「フィッシング詐欺の訓練を、ただのテストではなく、育成(コーチング)主導のアプローチで行っている」という点です。

セクション 1

統合型デジタルワークフォースにおけるリスクの現実

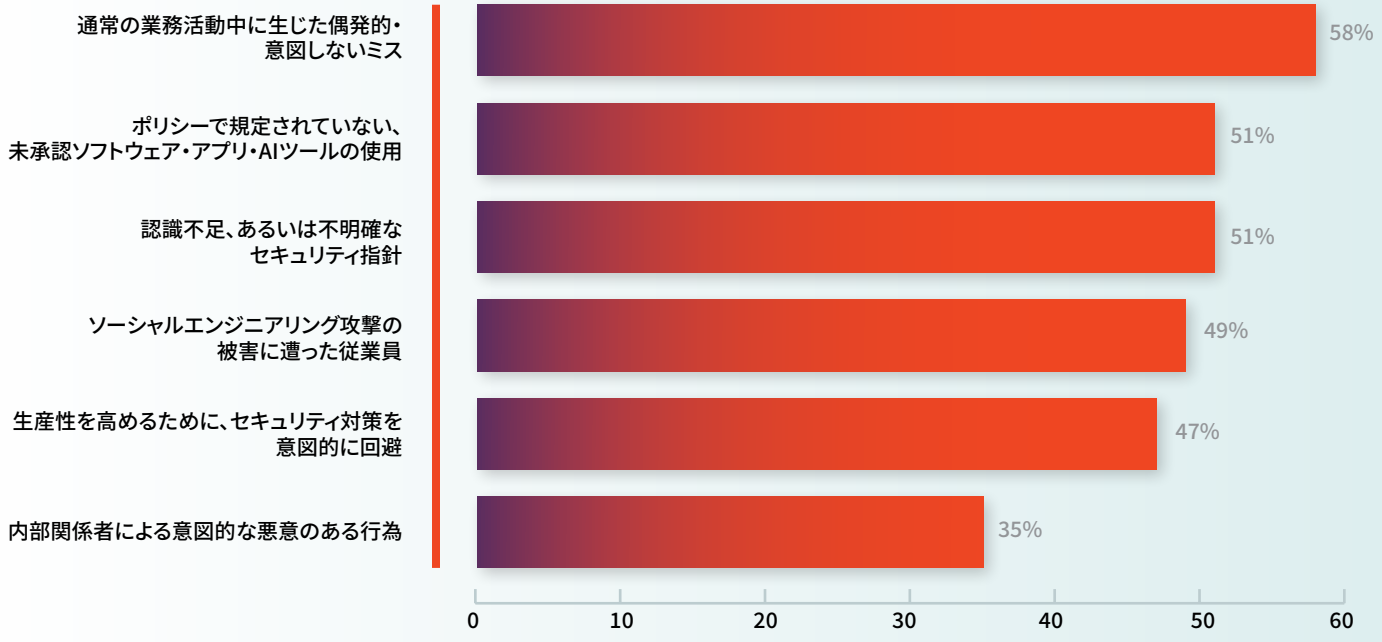
多くの組織が、従業員の安全確保や人的サイバーリスクへのアプローチを進化させていますが、リスクの性質そのものもまた変化しています。私たちの働き方はシフトしており、今日の従業員は、より素早い意思決定を迫られ、複雑なシステムを使いこなし、さらに意思決定や行動に影響を与える「エージェンティックAI」や「非エージェンティックAI」などのツールと共同で業務をこなさなければなりません。このような環境においては、人間とAIエージェントの双方によるミスが起こりやすく、かつ重大な結果を招きがちであり、日々の業務行動におけるリスクを高めています。

リスクは単独の大きなインシデントにあるわけではなく、日々の業務の中から始まっている

10人中およそ6人（58%）のサイバーセキュリティリーダーが、過去12ヶ月間において、自社のサイバーセキュリティに最も大きな影響を与えたのは「日々の業務におけるうっかりミス（日常的な過ち）」であったと報告しています。

> 「先手を打って警戒を怠らないこと、AIに依存しすぎないこと、そして常に『人間の目』で全体を監督し、二重チェックを行うことです。私たちの組織では、チェックのための強固なプロセスと手順を確立しています」
>
> シニアマネージャー
> （医療・医薬品業界、従業員数5,000名以上の組織）

過去12ヶ月間において、サイバーセキュリティに最も大きな影響を与えた「人的な行為」の種類



Q7. 過去12ヶ月間において、以下の「人的な行為」のうち、貴組織のサイバーセキュリティに最も大きな影響を与えたものはどれですか？
（対象：サイバーセキュリティ意思決定者、800名）

ミス（誤り）というものは、その本質からして悪意のあるものではありません。それは日々の業務における「現実」そのものです。ワークフローが複雑化し、生産性へのプレッシャーが高まるにつれ、従業員はこれまで以上に「スピードとセキュリティのバランス」を取ることを求められています。実際、セキュリティリーダーたちは、「セキュリティよりもスピードを優先せざるを得ないプレッシャー」を、今後12ヶ月間で人為的リスクが上昇する主要な要因の一つとして挙げています。これは、「ツールの複雑化」や「AIを悪用した攻撃の増加」に次ぐ上位の要因となっています。

55% の従業員が「安全対策は分かっていても、焦りや油断でミスをする」と回答。従業員の過半数が、時間的なプレッシャーや割り込み作業などによって、安全な行動が取れなくなる瞬間があると回答

Q4. ご自身の日々の業務についてお伺いします。以下の記述は、あなたにどの程度当てはまりますか？（同意の度合いをお答えください。）「たとえ安全な行動が頭では分かっていても、時間のプレッシャーや注意散漫によって、セキュリティ上の『ミス』を犯してしまうことがある」（対象：従業員、3,200名）

これは、より広範な課題を浮き彫りにしています。組織が業務効率を加速させるために、LLM（大規模言語モデル）やエージェンティックAIのツールを従業員に導入する一方で、それに見合った適切なガードレールの構築が常に追いついているとは限らないからです。こうした管理統制（コントロール）がないままでは、AIによって強化されたワークフォース（労働力）は、生産性を高めるどころか、かえってより大きなリスクを招く結果となってしまいます。

従業員の半数以上（55%）もまた、「たとえ安全な行動が頭では分かっていても、時間のプレッシャーや注意散漫によって、セキュリティ上のミスを犯してしまう可能性がある」と認めています。この事実は、2つの重大な現実を浮き彫りにしています。1つは、従業員がリスクを理解していたとしても、プレッシャーにさらされるといかに簡単に（安全への）意思が崩れ去ってしまうかという点。そしてもう1つは、「意識（アウェアネス）」と「実際の行動」との間にある深いギャップです。意識とは単なる精神状態にすぎませんが、最終的にセキュリティの成否を決定づけるのは日々の「行動」そのものだからです。このギャップを埋めるには、単なる注意喚起やコミュニケーションだけでは不十分です。組織が日々のワークフローにセキュリティを組み込み、安全な行動が「当たり前（デフォルト）」となるような文化を築くことで、従業員の行動を能動的に方向づけていく必要があります。そして、このギャップを埋める架け橋となれるのは、効果的な「ヒューマンリスク管理（HRM）」プログラムをおいて他にありません。

この「意識」と「行動」の間のギャップは、エージェンティックAIが日々のワークフローに統合されるにつれて、さらに顕著なものとなっています。しかも多くの場合、ユーザー自身がそれと意識しないうちに進行しているのです。従業員はメールの下書き、会議の要約、スプレッドシートの分析、ファイルの整理、レポートの作成などにAIを活用しており、開発者はソフトウェア開発を加速させるためにAI支援型のコーディングツールを使用しています。AIの活用は、もはや「選択肢の一つとしてのツール」から、業務の進め方を継続的に拡張する「目に見えない自動化された自律型レイヤー」へとシフトしているのです。そしてそれは、マシンのスピードとスケールで実行されています。

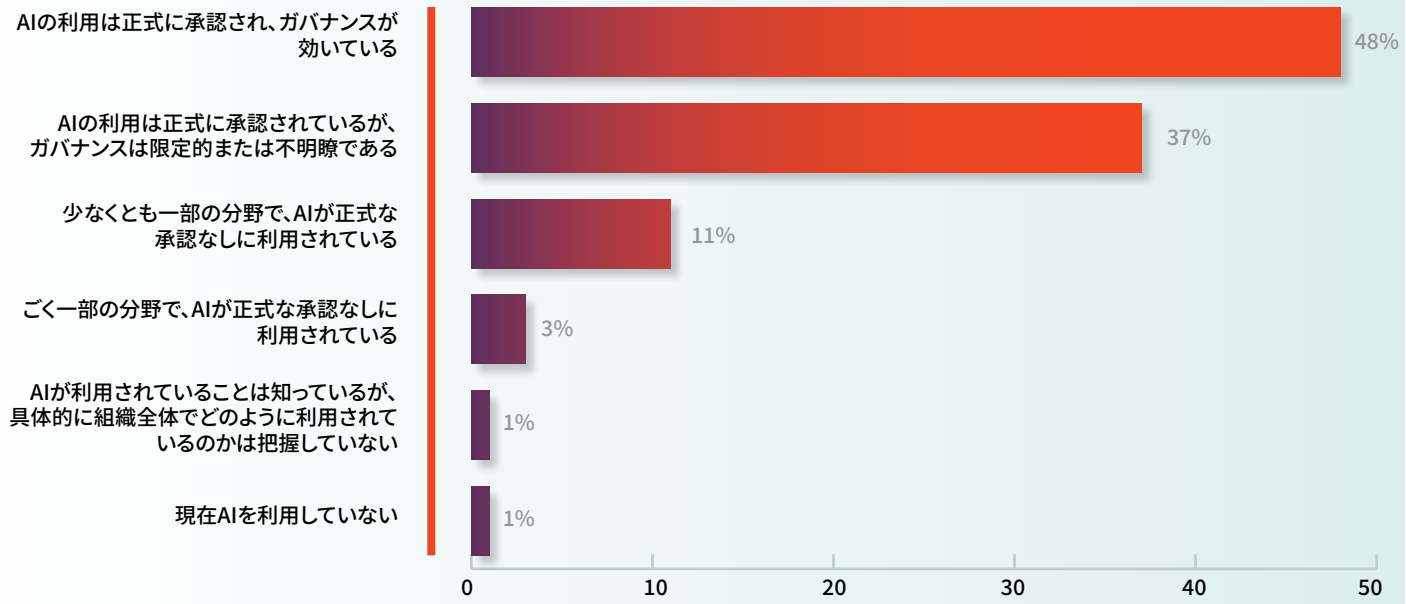
エージェンティックAIは管理統制が追いつかないほどのスピードで（攻撃対象として）露出を増やしている

エージェンティックAIは、今や日々の業務に広く組み込まれており、サイバーセキュリティリーダーの58%が「AIエージェントがすでに組織のワークフロー内で自律的に行動を実行している」と報告しています。その一方で、監視やガバナンス（統治）の欠如により、組織は危険にさらされた状態になっています。自社におけるAIの利用について**「正式に承認され、ガバナンスが効いている」と答えたリーダーはわずか48%にとどまり**、さらに37%は「正式に承認されてはいるものの、ガバナンスは限定的、あるいは不明瞭である」と回答しています。

58% の企業・組織で、AIツールやAIエージェントが複数の業務プロセスを自律型で実行中。そのうち 17% は、「人間の監視がほとんどない状態」AIを運用している

Q18. 今日の貴組織のワークフローにおいて、AIツールやAIエージェントの利用状況を最もよく表しているものはどれですか？（対象：組織でAIを利用しているサイバーセキュリティ意思決定者、795名）

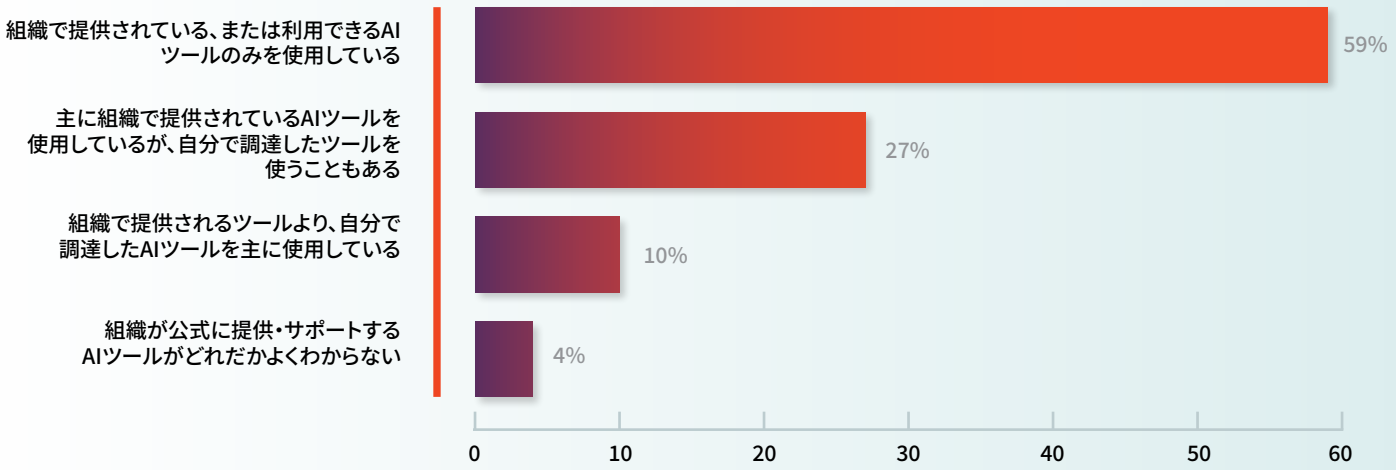
組織全体でのAIの利用・ガバナンス状況



Q17. 組織全体におけるAIツール（生成AIを含む）の利用状況について、貴組織の現状を最もよく表しているものはどれですか？（対象：サイバーセキュリティ意思決定者、800名）

従業員の3分の1以上（30%超）もまた、公式に提供されている選択肢がない場合や、会社の規制が厳しい場合には、自分自身でエージェンティックAIツールを調達して日常的に利用していると回答しています。

日々の業務における従業員のAIの利用状況



Q6. あなたの日々の業務におけるAIの利用状況について、最もよく当てはまるものはどれですか？（対象：職場でAIツールを利用していると回答した従業員、2,637名

また、サイバーセキュリティリーダーの半数以上（51%）が、「シャドーAI」やシャドーITツールの利用が、過去12ヶ月間の自社のサイバーセキュリティに大きな影響を与えたと考えています。この影響が浮き彫りにしているのは、これら未承認のツールが、もはや単に「管理外にあるソフトウェア」にとどまらないというパラダイムシフトです。それらは今や、監視の目が届かないところで自律的にタスクを実行する「シャドー従業員（見えない従業員）」として機能しているのです。

組織は、一貫したポリシーの適用や監視体制の強化に苦慮しており、半数近く（47%）が「AI（ツールおよびエージェントの双方）の安全な利用」を自組織の主要な課題として挙げています。その結果、リスクの所在は「チャットボットやエージェンティックAIが現場で実際にどう使われているか」へと移行しつつあります。これには、従業員がAIにどのような情報を共有（入力）しているか、AIの出力をどう解釈しているか、そしてAIの提案をしっかりと検証しているか、といった点が含まれます。エージェンティックAIの利用において適切な慣行や行動を促すことは、テクノロジーそのものの安全性を確保することと同じくらい重要なのです。

51% が「未承認のソフトウェア、アプリ、AIツールの使用が、過去12ヶ月間のセキュリティに甚大な影響を与えた」と回答

Q7. 過去12ヶ月間において、以下の「人的な行為」のうち、貴組織のサイバーセキュリティに最も大きな影響を与えたものはどれですか？（対象：サイバーセキュリティ意思決定者、800名）

AIツール Vs AIエージェント

- AIツールは一般的に「受動的（リアクティブ）」であり、ユーザーから指示（プロンプト）を受けた際に、テキストの生成、文書の要約、データの分析といった特定のタスクを実行します。AIツールは通常、人間の直接的なコントロールのもと、定義された限定的な範囲内で動作します。一方、AIエージェントはより「自律的」かつ「能動的（プロアクティブ）」です。彼らはユーザーに代わって計画を立て、意思決定を行い、何段階にもわたる一連の行動を実行することができます。その際、メール、カレンダー、ファイル、ビジネスアプリケーションといった連携されたシステムと相互にやり取りを行うことも少なくありません。AIエージェントは、状況を監視し、ワークフローを起動させ、時間の経過とともに自らの行動を適応させていくことができるため、生産性を劇的に向上させます。しかしその反面、適切に管理されなければ、より大きな複雑性とリスクをもたらすことになります。

人間とAIの相互作用（インタラクション）が、最大の脆弱性（弱点）になっている

従業員と自律型AIが協調して働くようになるにつれ、脅威のあり方（ランドスケープ）も変化しています。極めて精巧にパーソナライズされたフィッシング詐欺やディープフェイクなど、AIを悪用した攻撃は巧妙さを増し、見破ることが困難になっています。従業員の5人中4人以上（86%）が、「ディープフェイクコンテンツは今やリアルすぎて、何を信用すべきか判断するのが難しくなっている」と回答しています。その一方で、「自分もそうした攻撃にだまされてしまうかもしれない」と危機感を持っている人は64%にとどまりました。これは、残りの「自分はだまされない（リスクはない）」と考えている少なからぬ層が存在することを示しており、脅威が高度化する中で、根拠のない安心感（思い込み）が生じている可能性を浮き彫りにしています。

フィッシング詐欺は、依然としてリスクの主な要因であり続けています。従業員の10人中およそ6人（59%）が、フィッシングやなりすましメールを人的サイバーリスクの主要な原因として挙げており、さらに26%はそれを「最大の脅威」であると認識しています。

組織レベルにおいては、サイバーセキュリティリーダーの42%が、今後12ヶ月間で人的サイバーリスクが高まる主な要因として「AIを悪用した攻撃」を挙げています。それにもかかわらず、備えは限定的なのが現状です。多くの組織が、新たなリスクの管理全般に対しては「ある程度準備ができている」と感じているものの、予期せぬ、あるいは次々と現れる「AI関連の新たな脅威」に対して「非常に十分な準備ができている」と答えたリーダーは半数以下（48%）にとどまっています。

このような「リスクに対する認識」と「実際の備え」の間の乖離（ミスマッチ）は、組織にとって、先回りした（プロアクティブな）リスク管理戦略へとアプローチを進化させる好機でもあります。エージェンティックAIの導入が加速し、脅威がより巧妙化する中で、今後の焦点は「ガバナンスの強化」「行動の変容」「即応体制の向上」へと移行しつつあります。

リスクの本質は、常にITシステムとそれを扱う「ユーザー」の交差点に存在してきました。しかし、エージェンティックAIは今、この複雑なダイナミクスをさらに加速させています。だからこそ、「人間とAIの相互作用（インタラクション）のレイヤー」を保護することが、今や極めて重要な領域となっているのです。

86% の従業員が「ディープフェイクがリアルになりすぎて、何を信用すべきか判断するのが難しくなっている」と回答

64% の従業員が「AIを悪用した攻撃に、自分も騙される可能性がある」と回答

Q10. AIを悪用したサイバー脅威に関する以下の記述について、あなたにどの程度当てはまりますか？（同意の度合いをお答えください。）「ディープフェイクの音声や動画はリアルになりすぎて、今や何を信用すべきか判断するのが難しくなっている」「職場でディープフェイクの電話やメッセージなど、AIを悪用した詐欺や脅威にだまされてしまう可能性が自分にもある」（対象：従業員、3,200名）

42% のみが「今後12ヶ月間に発生し得る、想定外の『人的リスク』への備えが十分にできている」と回答

48% のみが「今後12ヶ月間に発生し得る、想定外の『AI関連リスク』への備えが十分にできている」と回答

Q24. 今後12ヶ月間にわたり、予期せぬ、あるいは次々と現れる「人間およびAIに関連したサイバーセキュリティリスク」を管理する上で、貴組織の準備はどの程度整っていると思いますか？（対象：サイバーセキュリティ意思決定者、800名）

セクション 2

人間および(AI)エージェントのリスクマネジメントにおける成熟度

日々の業務がエージェンティックAIによって形作られる割合が増すにつれ、各組織は従業員のセキュリティや人的サイバーリスクへのアプローチを変化させています。今やそれは、サイバーセキュリティ戦略の「中心的な柱」となりつつあります。しかし、その成熟度には依然としてばらつきがあり、自信（主観的な手応え）が実際の能力（客観的な実力）を上回っているケースが散見されます。

大半の組織が進歩を報告しており、10人中9人（90%）が「人的サイバーリスクを管理する能力が過去12ヶ月間で向上した」と回答しています。しかし、その向上を「劇的なものである（大幅に向上した）」と表現したリーダーはわずか36%にとどまりました。この事実は、多くの組織が進歩を感じてはいるものの、実際の「行動」や「文化」のレベルで本質的な変化を実感できている組織は一握りにすぎない、という現状を示唆しています。

36% のみが「過去12ヶ月間で、ヒューマンサイバーリスクを管理する能力が大幅に向上した」と回答

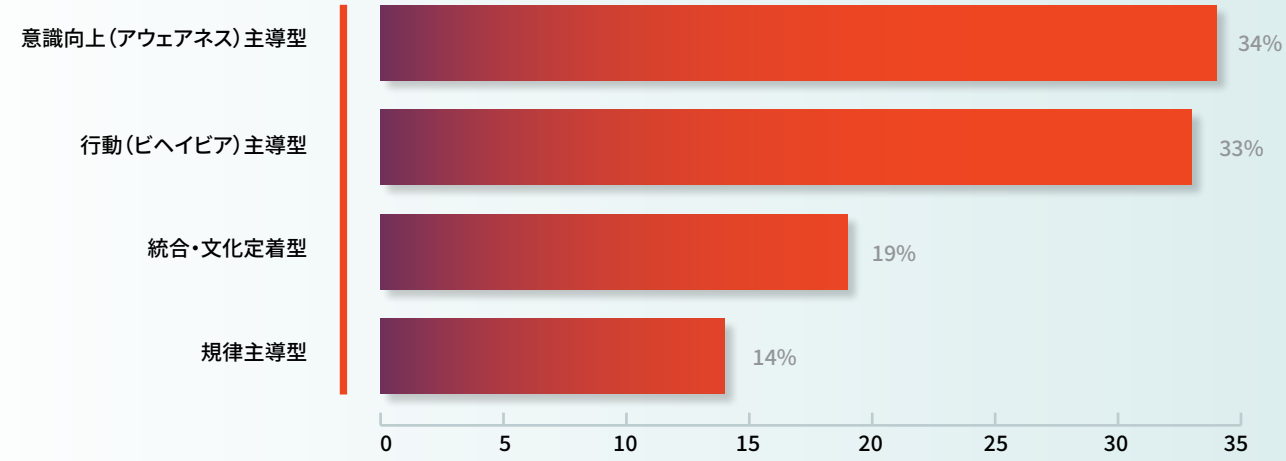
Q3. 過去12ヶ月間において、人的なサイバーセキュリティリスクを管理する貴組織の能力はどのように変化しましたか？（対象：サイバーセキュリティ意思決定者、800名）

人的リスクの評価について「指標や根拠に基づいている」と回答した割合は97%に達しているものの、それらの指標を「具体的な対策（実効性のあるインサイト）」へと落とし込むために必要なレベルの可視化ができている組織は、さらに少なくなります。測定自体は行われているものの、それらは断片的であったり、一貫性がなかったり、あるいは現実の行動と結びついていないケースが多いのです。

このギャップは、組織が実際にどのようにリスクを管理しているかを見ると、より鮮明になります。多くの組織はいまだに、従来の「セキュリティ意識（アウェアネス）向上」を目的としたアプローチにとどまっています。わずか3分の1強（34%）の組織が、「知識を提供すれば、それが行動の変容につながるだろう」と期待する、意識向上主導の戦略に頼っているのが現状です。

しかし、先ほどの調査結果が示していたように、55%の従業員が「たとえ安全な行動が頭では分かっていても、時間のプレッシャーや注意散漫によってセキュリティ上のミスを犯してしまうリスクがある」と認めています。このことは、意識を高めるだけでは不十分であるという事実を改めて裏付けています。特に、従業員がプレッシャーにさらされていたり、無意識のルーティンに頼ったりしている場合、意識が常に行動へと結びつくとは限りません。さらに、サイバーセキュリティリーダーのうち、「意識向上トレーニング単体で、従業員の行動を長期的に変容させることができる」と強く信じている人は、わずか42%にとどまっています。

組織における現在の人的サイバーセキュリティリスクの低減アプローチの種類と割合



Q5. 貴組織における現在の人的サイバーセキュリティリスクの低減アプローチについて、最もよく当てはまるものはどれですか？（対象：サイバーセキュリティ意思決定者、800名）

このように意識向上主導のアプローチに依存している現状は、従業員のセキュリティおよび人的サイバーリスクの核心にある「根本的なジレンマ」を浮き彫りにしています。組織は「行動が重要である」と理解しているものの、その多くはいまだに「意識を高めれば十分である」かのように対応しているのです。このマインドセットは、約4割の組織が「従業員の関与・当事者意識・納得感の醸成（42%）」や「安全な行動の促進（40%）」を、サイバーリスク管理における主要な課題として挙げている事実にも表れています。

現実には、行動というものは「置かれた状況（コンテキスト）」「従業員の動機（モチベーション）」「行動を起こす能力」の組み合わせによって形作られます。従業員は正しい行動を頭では理解していても、プレッシャーやスピード、あるいは認知の過負荷（キャパシティオーバー）に直面すると、往々にして異なる行動をとってしまうものです。つまり、リスクを低減させるために必要なのは、単に意識を高めることだけでなく、「意思決定が行われる環境そのものを変えること」なのです。

目指すべき「ゴールデンスタンダード（最高基準）」は、組織が強いリーダーシップ、明確な責任体制、そして効果的な測定手法に支えられながら、「意識」「行動」「コンプライアンス」を融合させる、組織の文化に完全に組み込まれた統合的なセキュリティアプローチです。しかし、このレベルの統合を実現できていると報告した組織は、わずか19%にすぎません。

このような、より成熟したアプローチを採用している組織は、他の組織とは根本的に異なります。彼らは単に意識向上（アウェアネス）だけに頼るのではなく、以下のような取り組みを行っています。

- セキュリティを日々のワークフローに組み込み、安全な行動が「当たり前」の選択肢となるようにする
- 単なるトレーニングにとどまらず、システムやリアルタイムのガイダンス（指示・誘導）を通じて適切な行動を強化する
- 「意識」「行動」「ガバナンス」を組み合わせ、一元化されたアプローチへと統合する

約半数（49%）の組織が「自社はワークフローに安全な行動を組み込めている」と考えています。しかし、これは明らかな「認識のギャップ」を浮き彫りにしています。大半の組織にとって、取り組みは進行中であるものの、一貫して文化として定着した（ワークフローに深く組み込まれた）成果を出せるレベルには、まだ達していないのが現状です。

従業員のセキュリティ（ワークフォース・セキュリティ）とは、本来、単なる「意識向上（アウェアネス）」や「コンプライアンス（法令・規律順守）」にとどまるものではありません。しかし現実には、多くの組織においてその理解が限定的であり、行動、文化、そして（各種施策の）統合といった他の重要な要素に本格的に取り組むことよりも、トレーニングやポリシーの策定ばかりに焦点を当ててしまっています。その結果、多くの組織は人的サイバーリスクに対処するための「明確に定義されたアプローチ」を確立するにいたっておらず、実際の実行プロセスが（リスクを低減させたいという）本来の意図と必ずしも一致していないのが現状です。

セクション 3

信頼、組織文化、そして行動を動かす要因

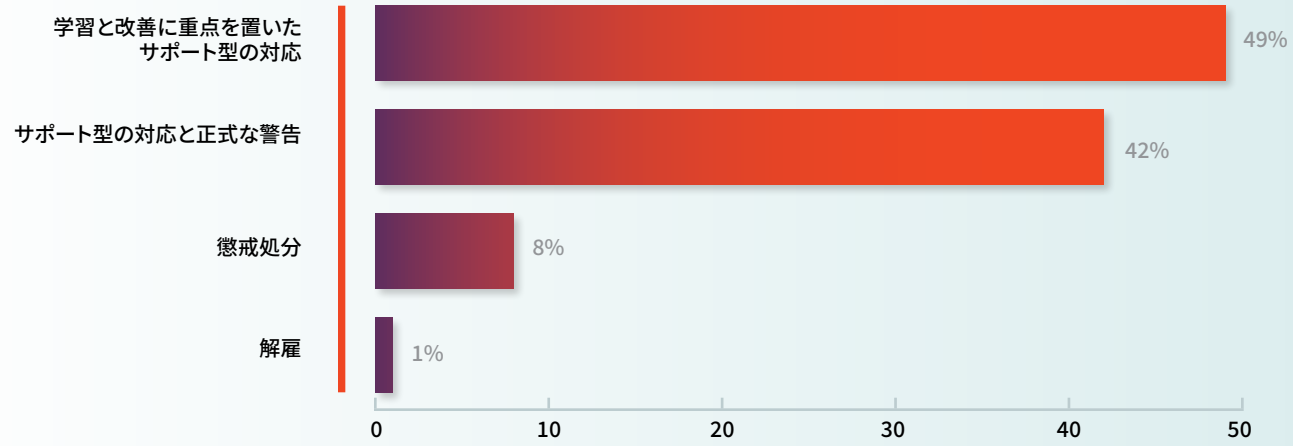
組織は、安全なセキュリティ成果を導き出す上で、「信頼」と「心理的安全性」がいかに重要であるかをますます認識するようになっています。大半の回答者が「従業員はITチームにミスを報告することに不安を感じていない（心理的安全性がある）」と報告しており、サイバーセキュリティリーダーの93%、従業員の89%がこの認識で一致しています。

かつてサイバーセキュリティは、ミスに対して懲戒処分を下すような「処罰」や「コンプライアンス（規律）重視」のアプローチに頼りがちでした。この方法はルールを強制する際に役には立つものの、透明性（実態のオープンな共有）を損なう原因にもなり得ます。だからこそ組織は、従業員がミスや問題を安心して報告でき、自分の行動が尊重され配慮をもって扱われると信じられるような、「信頼と確信に満ちた組織文化（カルチャー・オブ・コンフィデンス）」を築き上げる必要があるのです。

「強制（エンフォースメント）」から「コーチング主導」のセキュリティアプローチへ

組織は、より協調的で「コーチング主導」のサポート型アプローチへと移行しつつあります。現在、従業員がセキュリティ上の「ミス」を犯した際にこのような対応をとる組織は半数近く（49%）にのぼります。また、63%の組織が「教育やコーチング」を目的とした、学習主導型のフィッシング模擬訓練を導入しています。これは従業員の不手際を「摘発する」ためではなく、気づきを与えるために設計されたものです（ただし、24%の組織では、依然として公式な処分やペナルティが科されています）。

従業員が不注意によるサイバーセキュリティ上のミスを犯してしまった場合の主な対応



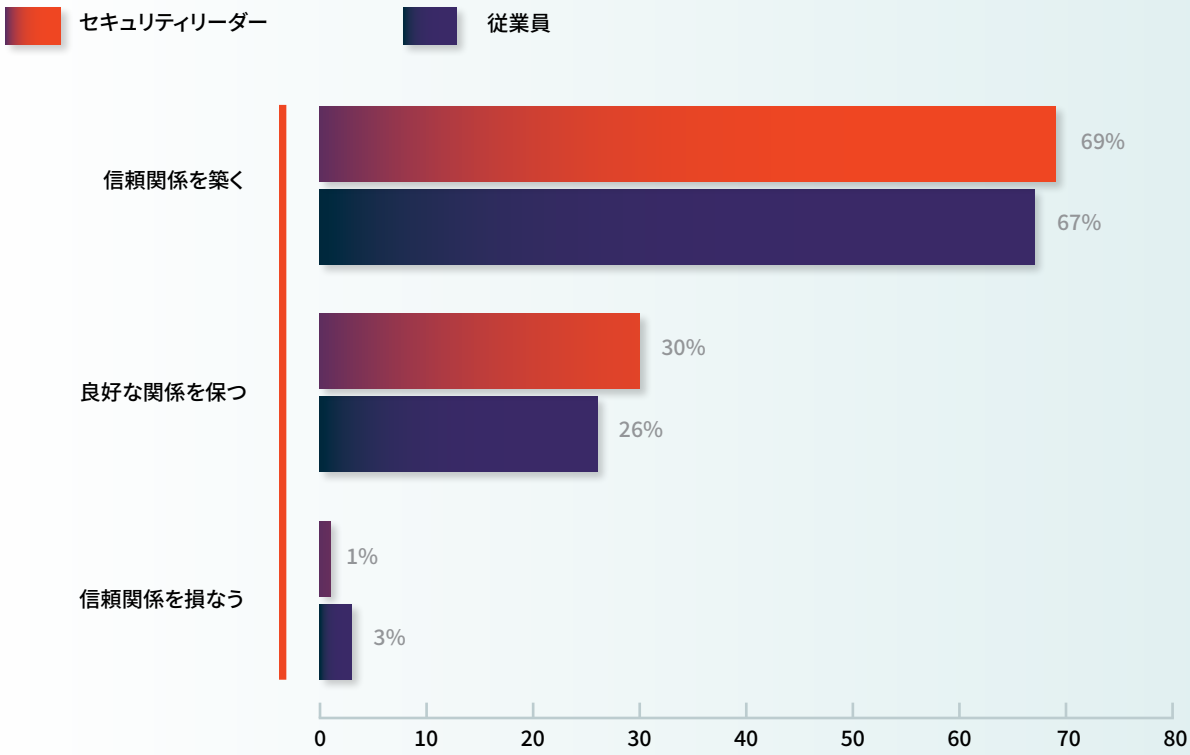
Q12. 従業員が不注意によるサイバーセキュリティ上のミスを犯してしまった際、貴組織では通常どのように対応していますか？（対象：サイバーセキュリティ意思決定者、800名）

組織は従業員にとって安全な（報告しやすい）環境を構築してきましたが、その浸透度には地域ごとにばらつきがある

従業員がセキュリティチームを信頼し、サポートされていると感じているとき、彼らは防御の「能動的な参加者」として行動する可能性が高くなります。このことは、従業員とセキュリティ部門との間のより広い関係性にも表れており、従業員の91%がIT・セキュリティチームを「支えてくれるパートナー（味方）」であると捉えています。さらに、従業員およびサイバーセキュリティリーダーの約3分の2（それぞれ69%と67%）が、セキュリティ意識の向上活動（アウェアネス活動）は、ITチームと従業員との間の信頼関係を築くのに役立っていると信じています。

しかしながら、この進展は地域によって一様ではありません。例えば、APJ（アジア太平洋・日本）地域では、ミスの報告に対する安心感や、セキュリティ文化への信頼度が、他の地域に比べて著しく低いことが判明しています。

サイバーセキュリティ意識向上活動がIT・セキュリティチームと従業員との関係性に与える影響



Q6. 貴組織のサイバーセキュリティ意識向上活動（アウェアネス活動）は、IT・セキュリティチームと従業員との関係性にどのような影響を与えていると思いますか？（対象：サイバーセキュリティ意思決定者、800名）/ Q3. 貴組織のセキュリティ意識向上活動（アウェアネス活動）は、従業員とIT・セキュリティチームとの関係性にどのような影響を与えていると思いますか？（対象：従業員、3,200名）

セキュリティ上のミスを犯した際に「安心して報告できる」と回答した従業員は、米州（南北アメリカ）の54%、EMEA（欧州・中東・アフリカ）の42%に対し、APJ地域ではわずか29%（3割未満）にとどまっています。この傾向はAPJ域内でもさらに分かれており、安心感が最も高いのはオーストラリア・ニュージーランド（43%）であるのに対し、シンガポール（27%）、日本（21%）では低くなっています。

APJの従業員が自社のITチームをどう捉えているかについても、同様のパターンが見られます。ITチームを「支えてくれるパートナーである」と強く信じている従業員は、米州の55%、EMEAの48%に対し、APJではわずか29%でした。ここでも域内で明確な差が出ており、やはりシンガポール（28%）や日本（24%）で信頼度が低いという結果が出ています。

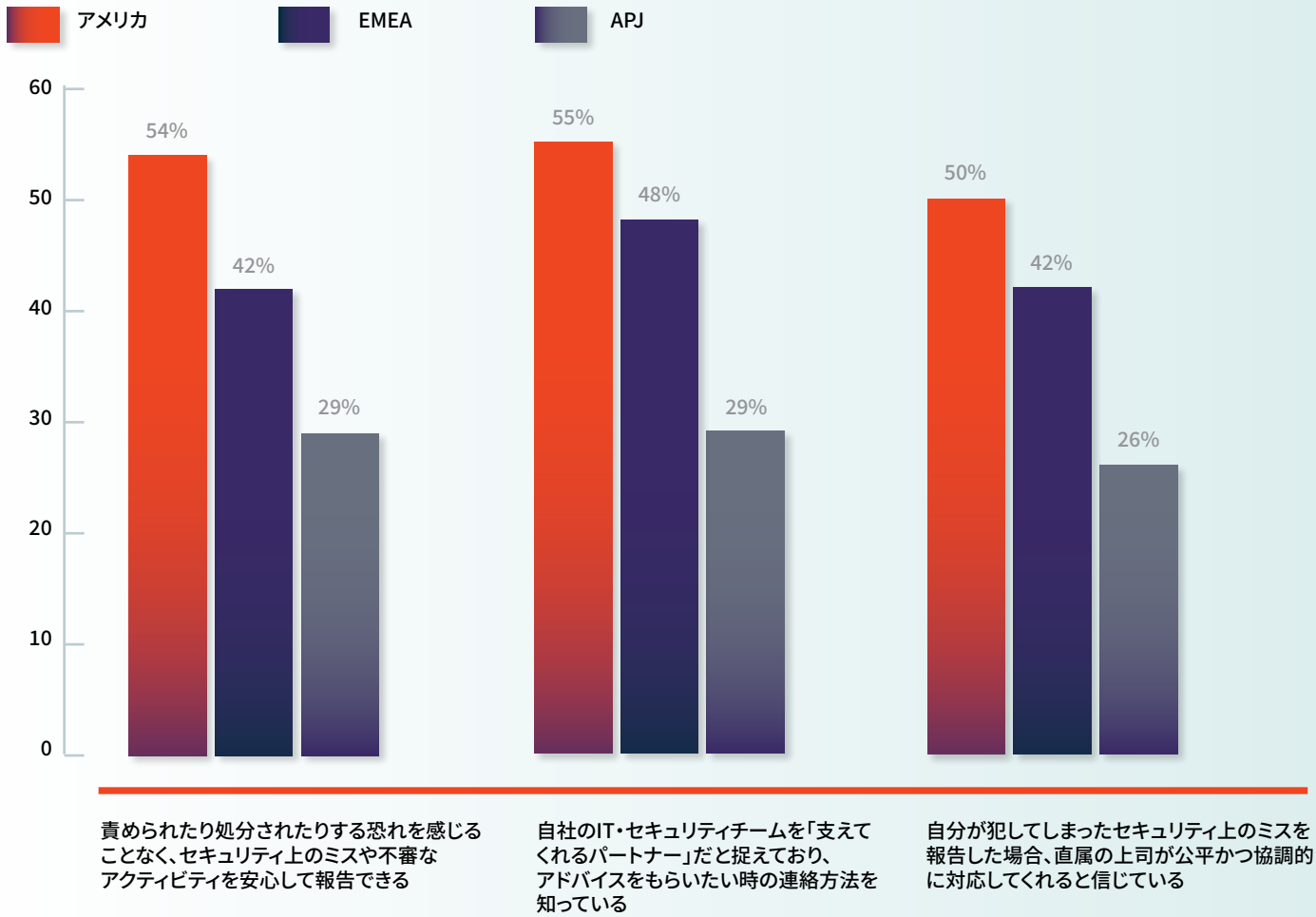
直属の上司（マネージャー）に対する信頼も、これとまったく同じ傾向を示しています。

APJ全体：自分がセキュリティ上のミスを報告した際、上司が「公平かつ協調的（サポート的）に対応してくれる」と信じている従業員は、わずか4分の1（26%）。日本：この割合はさらに下がり、そう答えた人はわずか2割（20%）にとどまりました。

このような地域間の差は、組織の文化が従業員のセキュリティ行動をいかに大きく左右するかを物語っています。従業員が声を上げることに不安や恐れを抱いている環境では、リスクが早期に報告される可能性は極めて低くなります。このような状況下では、どれほど優れたツールを導入していたとしても、その効果は限定的なものになってしまいます。なぜなら、ツールの価値を発揮するためには「問題の即座な報告」という、人の自発的な行動が前提となるからです。

したがって、今回の調査は「セキュリティとは、単なるポリシー（規律）やツールの問題ではない」という事実を改めて浮き彫りにしています。セキュリティの成否は、組織内の「信頼関係」や「文化」、そして「日々の業務の中で、一人ひとりが実際にどのように行動するか」にかかっているのです。

ミスの報告に対する安心感とセキュリティカルチャーへの信頼度に関する地域差（「強く同意する」と回答した割合）



Q4. 日々の業務を振り返って、以下の各項目にどの程度同意しますか？【回答選択項目】責められたり処分されたりする恐れを感じることなく、セキュリティ上のミスや不審なアクティビティを安心して報告できる / 自社のIT・セキュリティチームを「支えてくれるパートナー」だと捉えており、アドバイスをもらいたい時の連絡方法（アプローチの仕方）を知っている / 自分が犯してしまったセキュリティ上のミスを報告した場合、直属の上司が公平かつ協調的（サポート的）に対応してくれると信じている（対象：各項目に「強く同意する」と回答した従業員、3,200名）

すべてのアプローチが同じ成果をもたらすわけではない

信頼関係を育み、前向きな行動を促し、そして「安全に報告できる環境」を作り上げている組織ほど、意識向上（アウェアネス）を実際の行動へと結びつけやすい優位なポジションにあります。
このことは、より成熟し、日々の業務に**統合され、組織文化として深く根づいている**サイバーセキュリティアプローチのあり方にも明確に反映されています。

「統合・文化定着型」のアプローチを採用している組織は、強いリーダーシップ、明確な責任体制、そして効果的な測定手法に支えられながら、「意識」「行動」「コンプライアンス（規律）」を融合させています。こうした組織は、いくつかの重要な領域において他の組織を圧倒しています。例えば、このアプローチをとる組織の10社中8社以上（83%）が「セキュリティ意識向上活動（アウェアネス活動）がITチームと従業員との間の信頼関係を築いている」と回答しているほか、67%が「報告のための明確なガイダンス（手順）」を提供しており、51%が「コーチング主導型のフィッシング模擬訓練」を実施しています。

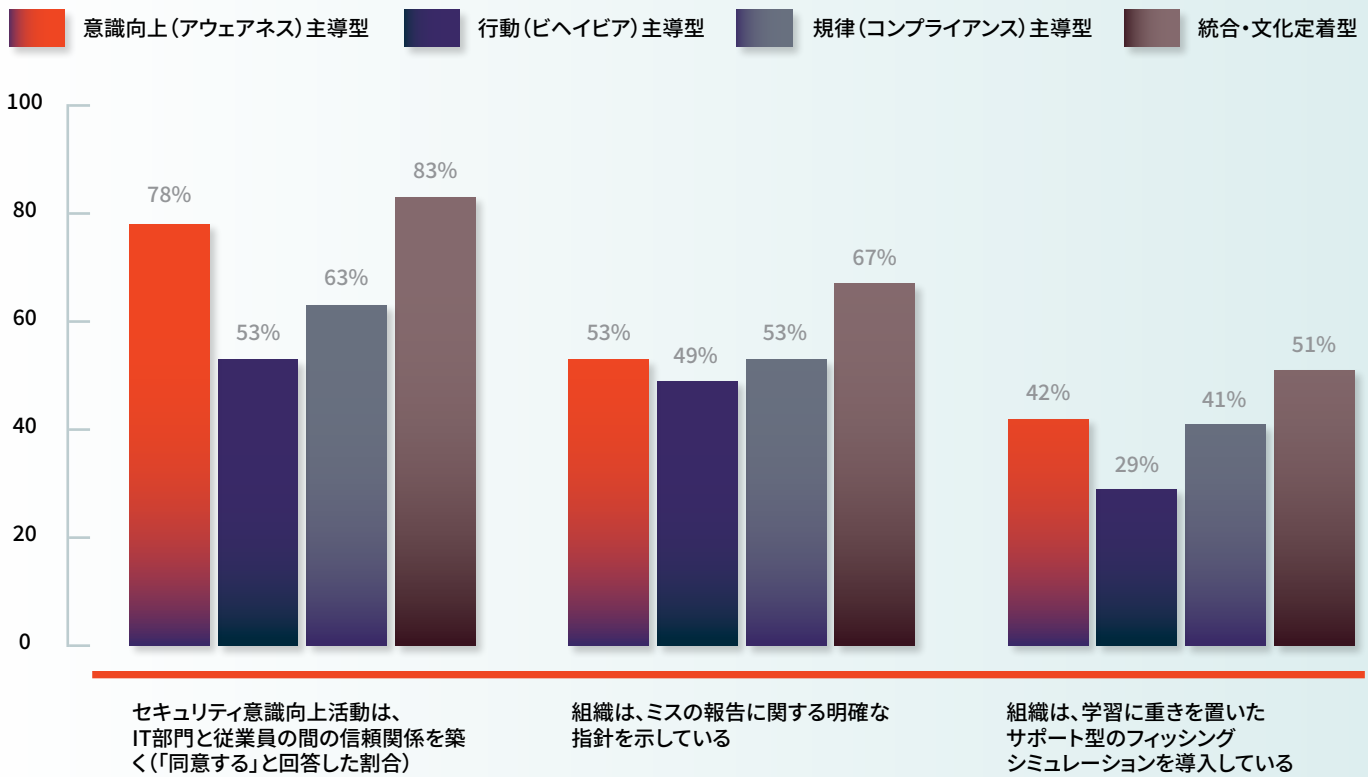
一方で、**「意識向上主導型」「コンプライアンス主導型」「行動主導型」**のいずれか単一のアプローチにとどまっている組織では、得られる成果にばらつきがあり、**統合・文化定着型**を採用している組織に比べて効果が劣るという結果が出ています。このことは、組織内に定着させるセキュリティプログラムやセキュリティ文化の「実効性」を高める上で、統合的なアプローチがいかに重要であるかを明確に物語っています。

この結果は、組織における人的リスク管理のアプローチに対して何を意味しているのでしょうか？

得られる成果におけるこのような差は、組織が「人的リスク」にどうアプローチするかという、より大きなパラダイムシフト（転換）を示しています。

成熟度の高い組織は、単に「ミスを防ぐこと」だけに焦点を当てるのではなく、従業員が主体的に行動し、報告し、改善していけるような環境づくりに取り組んでいます。その結果、従業員はただの「リスク要因」ではなく、ビジネス全体のセキュリティを支え、高めてくれる存在へと進化しているのです。

それぞれのアプローチが、セキュリティカルチャーに関する指標に与える影響



Q6. 貴組織のサイバーセキュリティ意識向上活動（アウェアネス活動）は、IT・セキュリティチームと従業員との関係性にどのような影響を与えていると思いますか？ / Q14. 従業員による不審なサイバーセキュリティアクティビティや人的ミスの報告を促すために、貴組織で導入しているもの（制度や仕組み）はありますか？ / Q13. 従業員向けのフィッシング模擬訓練やセキュリティテストに対する貴組織のアプローチとして、最もよく当てはまるものはどれですか？【クロス集計軸】Q5. 貴組織における現在の人為的サイバーセキュリティリスクの低減アプローチについて、最もよく当てはまるものはどれですか？（対象：サイバーセキュリティ意思決定者 800名、内訳：* 意識向上（アウェアネス）主導型：268名 * 行動（ビヘイビア）主導型：264名 * コンプライアンス（規律）主導型：115名* 統合・文化定着型：152名）

結論：
リスクの『管理』から、組織の『レジリエンス』の向上へ

ゴールは、最初から「ミスをゼロにすること」ではありませんでした。本当に目指すべきものは、常に「レジリエンス（しなやかな強さと回復力）」だったのです。現在、多くの組織がその実現に向けて動き出しています。先見の明がある組織は、適切なツール、エージェンティックAI、そして「何かが起きたときに従業員が安心して声を上げられる文化」の融合に取り組んでいます。テクノロジーは重要です。しかし、組織の文化はそれ以上に重要なのです。この両者が揃って初めて、人間とAIがもたらすリスクを「人間の強み」へと変える環境が整うのです。

テクノロジー、特にエージェンティックAIは、人間が適応できる以上のスピードで進化を続けています。そのため、現在のセキュリティ環境には、以下に示す3つの重大なギャップが存在しています。

1 確信と実態のギャップ（Confidence and capability）

組織は「状況は改善している」と信じていますが、実際の測定手法や可視化（現状の把握）は依然として限定的なレベルにとどまっています。

2 意識と行動のギャップ（Awareness and behavior）

組織は従業員のセキュリティ意識を高める取り組みを行っていますが、それが実際の業務において従業員が適切に行動することの保証にはつながっていません。

3 エージェンティックAIの普及と統制のギャップ（Agentic AI adoption and control）

AIはすでに日々のワークフローに深く組み込まれていますが、ガバナンス（統治）や監視（オーバーサイト）の体制構築は未だその後手に回っています。

これこそが、「Human Wins（人間の勝利／人間による成果）」という概念が最も重要になる理由です。従業員のセキュリティ対策において卓越した成果を上げる組織とは、以下を実践できる組織にほかなりません。

- 単に情報を伝えるだけでなく、望ましい行動へと導くシステムを設計する
- ミスの報告とそこからの学習を促す、協調的でサポート体制のある組織文化を築く
- 「失敗の追跡」から「前向きな行動の強化（称賛）」へとシフトする
- 「セキュリティ第一（セキュリティ・ファースト）」のマインドセットを、人間とAIエージェントの双方へと拡張する

この新たな現実において、従業員とAIエージェントの双方で構成される「デジタル・ワークフォース」は、サイバーリスクから組織を守るための極めて重要な防御層（レイヤー）となります。働き方が変化し続けるなかで、意識、行動、そして組織文化を融合させ、リスクの実態を常にクリアに把握できている組織こそが、脅威への露出（エクスポージャー）を減らし、持続可能なレジリエンスを築き上げることができるのです。私たちの働く環境は変わりました。今や従業員とAIエージェントは、「ひとつの統合された防御層」として機能しています。この変化に伴い、「優れたセキュリティ」の定義そのものも変わったのです。

リスクの全体像を正しく把握し、それに合わせて「意識」「行動」「文化」を連動させている組織は、単にリスクを管理するだけにとどまりません。彼らは、攻撃者にとっての「強固で付け入る隙のない標的（ハードターゲット）」を作り出しているのです。脅威が消え去ることはありません。しかしそれらの脅威は、はるかに高度な備えを固めた組織によって、ことごとく阻まれることになるでしょう。

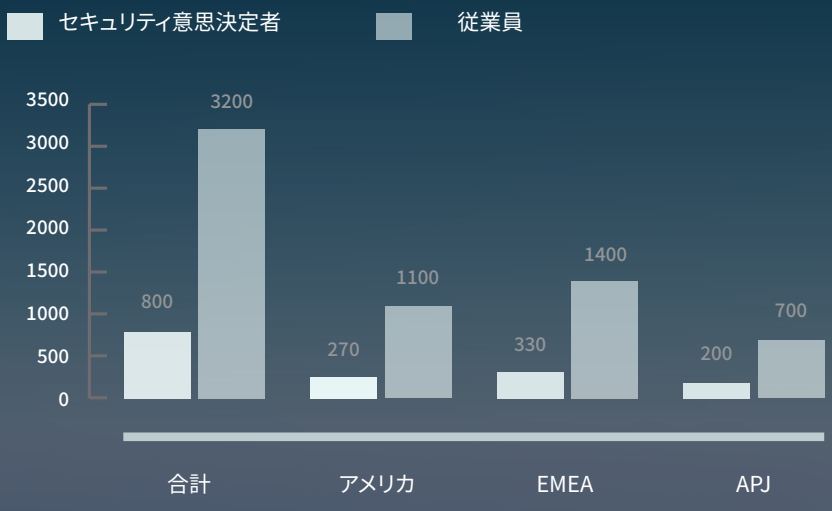
問い合わせる

調査方法

本調査は、南北アメリカ地域、EMEA（欧州・中東・アフリカ）、およびAPJ（アジア太平洋・日本）地域における、800名のセキュリティ意思決定者と3,200名の従業員を含む、計4,000名の専門職を対象に実施されたグローバル調査に基づいています。

回答者は従業員数250名以上の組織に所属しており、情報技術（IT）、医療・ヘルスケア、消費者サービスなど、幅広い業界の民間企業および公共部門（公的機関）を網羅しています。

対象者の内訳：地域別



回答者の内訳：業種別（セキュリティ意思決定者、計800名）

回答者の内訳：業種別（従業員、計3200名）

KnowBe4について

KnowBe4は、現代のワークフォース（従業員およびAI）が日々、よりスマートなセキュリティ判断を下せるよう支援しています。世界中で7万社以上の組織から信頼されているKnowBe4は、「デジタル・ワークフォース・セキュリティ」のパイオニアであり、AIエージェントと人間の双方の安全を確保しています。KnowBe4のプラットフォームは、攻撃シミュレーション（模擬訓練）やトレーニング、コラボレーションセキュリティ、そして独自のリスクスコア（Risk Score）と「AIDA（Artificial Intelligence Defense Agents）」を搭載したエージェントセキュリティを提供します。当プラットフォームは、15年間にわたり蓄積された「行動データ」を活用し、ソーシャルエンジニアリング、プロンプトインジェクション、シャドーAI（組織に無断で利用されるAI）といった高度な脅威に対抗します。人間とAIエージェントの双方を保護することにより、KnowBe4はワークフォースの信頼と防御において業界をリードしています。

詳細については、knowbe4.jpをご覧ください。

Follow us on Linkedin　Follow us on X

Vanson Bourne（ヴァンソン・バーン）について

Vanson Bourneは、テクノロジーセクターに特化した独立系の市場調査専門会社です。確実性と信頼性の高い調査ベースの分析において定評があり、その名声は厳格な調査原則と、あらゆるビジネスセクターおよび主要市場における技術・ビジネス部門の経営幹部（シニアデシジョンメーカー）から意見を募る卓越した調査能力に基づいています。詳細については、www.vansonbourne.com をご覧ください。

KnowBe4 Japan合同会社 | 〒107-0052 東京都港区赤坂 9−7−1 ミッドタウン・タワー 18F
03-4586-4540 | www.knowbe4.com/ja | info@KnowBe4.jp