「生成 AI を以って生成 AI を制す」 ― 新たな局面を迎える。サイバーセキュリティー対策とは

生成 AI は過去のどのテクノロジーとも異なっている。瞬く間にビジネスと社会を揺るがす存在になりつつあり、リーダーは想定や計画、早急な戦略の見直しを迫られる。

こうした急激な変化に CEO が適切に対処するための一助として、「IBM Institute for Business Value（IBM IBV）」は独自調査に基づいて、生成 AI に関する解説シリーズをまとめ、テーマごとに公表している。内容はデータ・サイバーセキュリティーからテクノロジー投資戦略、顧客体験にまで幅広く及ぶ。

今回は第七弾として、「サイバーセキュリティー」をお届けする。

生成 AI はサイバーセキュリティーの脅威を増大させるが、防御も高める。

生成 AI は新世代のサイバー脅威をもたらした。攻撃者がシステムの脆弱（ぜいじゃく）性につけ込む機会が増え、攻撃キャンペーン * を実行する方法も多様化している。

* 攻撃キャンペーンとは、標的を定めた上で長期にわたって行われる一連のサイバー攻撃を指す

だが幸いにして、逆の側面もある。生成 AI はサイバー攻撃に対する防御を高めることもできるのだ。例えばセキュリティー対応のプロセスは従来、多大な労力と時間を要するが、生成 AI によってその速度が速まることが期待される。また生成 AI は膨大なデータの分析やパターン認識を通じて異常を見つけ、脅威が出現すると迅速に検知することができる。

攻撃者が新たな手口を編み出すたびに、サイバーセキュリティー・チームは速やかな対応を迫られる。この“いたちごっこ”に追われる中でも、常に警戒を怠らないことが、脆弱性を解消し、一歩先んじた対策を講じる上で鍵となる。

IBV が考える、すべてのリーダーが知っておくべき 3 つのこと：

1. 生成 AI は新たなリスクと脅威にさらされる世界の到来を告げる。
2. 信頼できる生成 AI は、セキュアなデータなしには実現できない。
3. 生成 AI をサイバーセキュリティーに活用すれば、戦力増強につながる。

そして、すべてのリーダーが今すぐ実行すべき 3 つのこと：

1. 生成 AI はリスク管理上、極めて危うい状況にある。すぐにでもセキュリティーを確保すべきだ。
2. 「信頼できるデータ」を組織の根幹に据える。
3. 「スピードと規模」を軸にサイバーセキュリティー投資の方向性を見直す。

1. サイバーリスク + 生成 AI

リーダーが
知るべきこと

生成 AI は新たなリスクと脅威にさらされる世界の到来を告げる

今、サイバー攻撃者は生成 AI という全く新たな武器を手に入れ、機会をうかがっている。攻撃者は成り済ましメールを送ることに加え、声や顔、さらには人格までも模倣し、被害者をだまして罠にかけることが可能となった。

しかも、これはほんの始まりに過ぎない。

専門家は、今後半年～ 1 年間における生成 AI のさらなる普及に伴い、従来と比較して規模が大きく、迅速で、洗練され、高度な不正侵入が行われ、その後も過去に例のない脅威が頻繁に現れる可能性があると指摘する。リスク要因である発生可能性と潜在的影響から見て、AI による大規模な自律型攻撃が際立った脅威である。しかし、経営層がビジネス上、最も警戒するのは信頼するユーザーに成り済ます攻撃者だ。次に、マルウェアに代表されるような悪意のある不正なコードの生成をほぼ同程度の脅威として挙げている。

組織が生成 AI を導入することで、新たなリスクが発生する場合もある。実際、経営層の 47% は、生成 AI を業務に導入した場合、自社の AI モデルやデータ、サービスを標的にした新種の攻撃を呼び込む恐れがあるのではないかと危惧している。導入によって今後 3 年以内に組織内でセキュリティー侵害が発生する可能性が高いとする見方でほぼ全員（96%）が一致する。

データ侵害の平均コストは全世界で 445 万ドル（米国では 948 万ドル）に達していて、企業は新たなサイバーセキュリティー・リスクに対処するため、投資を拡大している。経営層によると、2023 年の AI サイバーセキュリティー予算は、21 年時に比べ 51% 増加し、25 年までにさらに 43% 増加すると予想されている。

経営層によると、
2023 年の AI サイバー
セキュリティー予算は、
21 年時に比べ
51% 増加している。

25 年までに
さらに 43% 増加すると
予想されている。

1. サイバーリスク + 生成 AI

リーダーが
実行すべきこと

生成 AI はリスク管理上、極めて危うい状況にある。すぐにでもセキュリティーを確保すべきだ

サイバーセキュリティーのリーダーに切迫感をもって行動するよう求め、生成 AI のリスク対策を今すぐ打ち出させる必要がある。その場しのぎや暫定対応では効果が低い。

自社がさらされているリスク環境について理解を共有する。

1. サイバーセキュリティーやテクノロジー、データ管理、運用のリーダーを招集し、リスクが高まっている状況について取締役レベルで議論することが必要である。
2. 議論においては、以下の論点を含める。
   （ア） 具体的にどのように生成 AI が悪用され、機密データの漏えいやシステムへの不正アクセスが起こるか。
   （イ） 新たな「敵対的」AI の最新情報に基づき、こうした AI が、どのようにしてほとんど気づかれずに重要なデータセットに変更を加え、どれほど有害な結果をもたらすか。

AI のトレーニングから導入後まで、全局面でセキュリティーを担保する。

1. AI モデルのトレーニング中はデータ保護に注意が必要である。例えば、使用するデータを暗号化しておくことが効果的である。
2. モデル開発中は、脆弱性の発生やマルウェア（悪意あるプログラム）の侵入、AI の悪用を防ぐような配慮が必要である。
3. モデル導入後も AI 特有の攻撃に対する監視を継続、強化することが必要である。具体的には、データ・ポイズニング（データの改ざん）やモデル盗用（モデル情報の収集・転送）などだ。

AI に特化した新たな防御策に投資する。

1. 企業は、新たな敵対的攻撃への防御策のための予算が、既存の AI システムを支えるインフラやデータの保護強化のためのセキュリティー管理や専門技術の予算枠では賄いきれない可能性を認識することが必要である。
2. AI モデルに対する敵対的攻撃を検知し、防御する対策のために、追加の予算措置も含め投資計画の見直しを進める。

2. データ + 生成 AI

リーダーが
知るべきこと

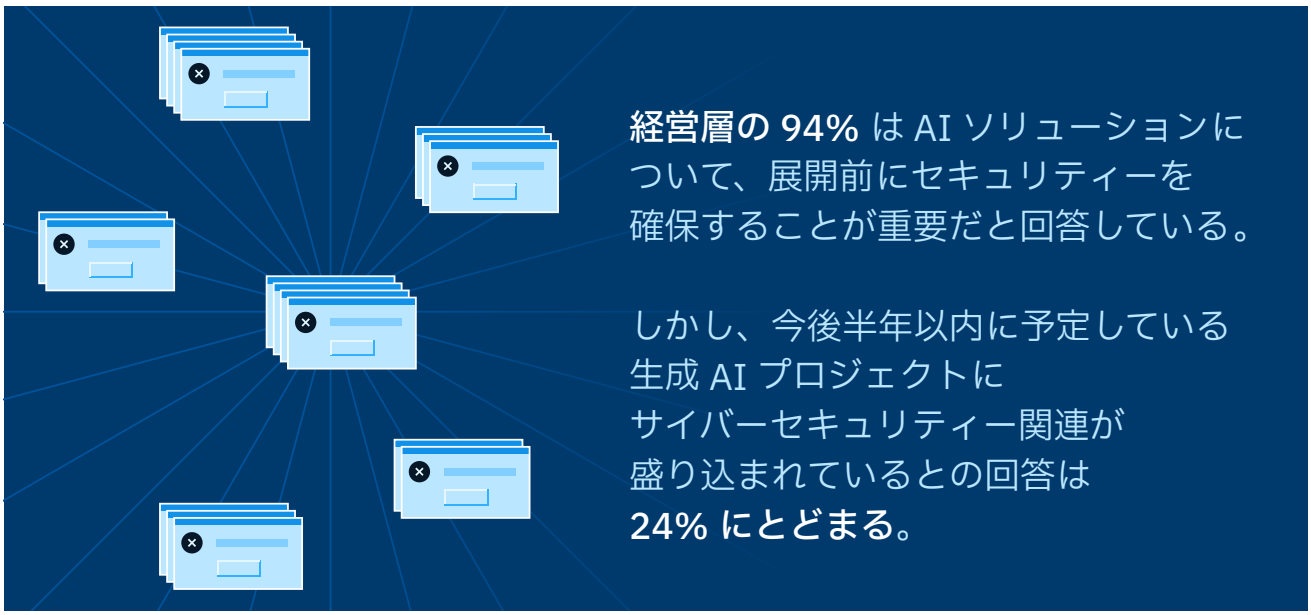
信頼できる生成 AI は、セキュアなデータなしには実現できない

データは生成 AI の生命線だ。どのモデルもデータに基づいて問いに答え、インサイト（洞察）を提供している。トレーニング・データがサイバー攻撃の標的になる理由はここにある。攻撃者はデータを盗んで最高値で売れるか画策し、データ侵入を行うことで不正な企てを実現する新手の攻撃を生み出している。組織の生成 AI モデルを駆動するデータに変更を加え、特定の目的で誤操作や誤情報を発生させれば、攻撃者はビジネスの意思決定を左右できるためだ。こうした脅威の高まりは、法律やセキュリティー、さらにはプライバシーに関する新たな懸念を次々に引き起こすため、CEO が全社規模で対策を講じる必要性が生じている。

経営層は問題の兆候に気づいており、生成 AI の導入を進める中で、さまざまなリスクが顕在化すると予想している。84% の経営層は、生成 AI の導入により生まれた新たな脆弱性が原因で、新たなサイバーセキュリティー攻撃が広範囲で、あるいは壊滅的な形で発生すると懸念している。3 人に 1 人は、こうしたリスクに対処するには、根本的に新しい形のガバナンスが必要であり、例えば、包括的な規制の枠組みや第三者による独立監査を導入すべきだと述べている。

AI ソリューションについては、展開前にセキュリティーを確保することが重要だとする回答が、経営層全体で 94% に達する。しかし、今後半年以内に予定している生成 AI プロジェクトにサイバーセキュリティー関連が盛り込まれているとの回答は 24% にとどまる。69% は生成 AI に対するサイバーセキュリティーより、イノベーションを優先して進めると答えている。

このことは、生成 AI のサイバーセキュリティーを巡って、必要性が認識されながらも、対応が大きく遅れている現実を示している。CEO はデータ保護へ投資し、データ・セキュリティーの現状やその根本的な原因の解明に正面から取り組む必要がある。さもなければ、望まぬ結果を招き、大きな代償を支払うことになるだろう。具体策としては、データ保護のために暗号化や匿名化に取り組むことや、データに対する攻撃の追跡・記録やモニタリング・システムを整備することなどが挙げられる。それによって、生成 AI モデルで使用するデータがセキュアであることを担保することが可能となる。



2. データ + 生成 AI

リーダーが
実行すべきこと →

「信頼できるデータ」を組織の根幹に据える

サイバーセキュリティーの取り組みを進化させ、複数の生成 AI モデルとデータ・サービス活用のための多岐にわたるセキュリティー要件を検討する。

AI 活用の中心に信頼とセキュリティーを据える。

1. 組織が AI 活用において信頼とセキュリティーを重視することは、組織全体にわたるバイアスの軽減、コンプライアンスに準拠したデータの取り扱い、恒常的なデータ品質向上の取り組みなどにつながり、データの信頼性確保の鍵となる。
2. キーとなるアクションは次のとおりである。
   (ア) 自社のデータ・ポリシーやコントロールにおいてセキュリティーやプライバシー、ガバナンス、さらにコンプライアンス（法令順守）が優先事項となるよう、策定や見直しを行う。
   (イ) AI によるバイアス（偏見や思い込み）やハルシネーション（もっともらしい虚偽の回答）などの懸念を防ぐために、透明性と説明責任がいかに重要であるかを全社に周知し、同時にリスク管理を徹底する。

AI の原動力であるデータを保護する。

1. 脅威モデリング（サイバーリスクを特定して対応を検討する手法）を拡張して、生成 AI 特有の脅威に対応する必要がある。
2. 具体的な脅威とは、データ・ポイズニングの恐れや、機密データおよび不適切コンテンツの出力結果への混入などである。
3. この対策としては、以下の手順を実施することが効果的である。
   (ア) 最高情報セキュリティー責任者（CISO）に指示を出して、AI のトレーニングや微調整に使われる機密データを特定して分類させる。
   (イ) データ損失防止技術を活用して、プロンプト（生成 AI に出す指示や質問）を介したデータ漏えいを防ぐよう求める。
   (ウ) 機械学習データセットに関わるアクセス・ポリシーとコントロールを実行する。

サイバーセキュリティーを「商品と一体のもの」として捉え、社員やパートナーなどの利害関係者を「顧客」と同様に重視する。

1. 企業はサイバーセキュリティーに関する認識・周知を高めるために、サイバーセキュリティーと製品とを最初から一体のものとして開発を進めたり、社員やパートナーなど利害関係者に対するセキュリティー対策を重視する必要がある。
2. キーとなるアクションは以下のとおりである。
   (ア) サイバーセキュリティーは生成 AI を活用した製品のリスク軽減に極めて重要な役割を担い、企業の収益に直結する、という認識を全社に根付かせる。
   (イ) 生成 AI を活用した製品につきものであるサイバーセキュリティーの脅威について従業員への教育を促進する。特に、従業員自らの行動を見直すことがデータおよびセキュリティーのハイジーン（リスク予防策）向上に極めて有効である点を強調する。
   (ウ) こうして企業のセキュリティーに関するマインドの転換を通じて、サイバーセキュリティーとビジネスの好循環へつなげていくことで、リスクを低減しつつ AI の導入促進が可能な状況を作り出す。

3. サイバーレジリエンス + 生成 AI

リーダーが
知るべきこと

生成 AI をサイバーセキュリティーに活用すれば、戦力増強につながる

生成 AI をサイバーセキュリティーに活用すれば、ビジネスを加速させる力となる。時間のかかる反復的なタスクが自動化され、従業員はさらに複雑かつ戦略的なセキュリティーの課題に集中することが可能になる。脅威を検出して調査し、過去のインシデント（事故）経験も踏まえて、組織のインシデント対応業務を迅速に行うこともできる。

生成 AI がもたらす成果への期待感から、迅速かつ広範な導入を求める声が CEO に対して高まっている。しかし、ビジネス・リーダーにとって不可欠なのは、成長に向けてひた走る組織が気づかぬうちに“砂上の楼閣”となってしまわないよう、生成 AI をレジリエンス強化にも活用することだ。そうすれば、経営層は生成 AI に伴うリスクを回避するだけでなく、強じんな組織をつくることが可能になる。

経営層の過半数（52%）は、生成 AI が社内リソースや機能、人材、さらにスキルの配分を改善する上で有効だと回答している。さらに、生成 AI がサイバーセキュリティーの人材に取って代わるわけではないが、生成 AI の導入により、彼らの能力を拡張・強化することができるとの回答は 92% に達する。

こうした最新テクノロジー・ツールは、従業員の複雑な作業を軽減し、最も重要なタスクに集中するために役立つ。だからこそ、経営層の 84% は、サイバーセキュリティー・ソリューションについて、従来型よりも生成 AI の活用を優先する計画だと答えたのだろう。

サイバーセキュリティーに生成 AI を活用すれば、企業のエコシステム全体に乗数効果を拡大することが期待できる。経営層の 84% は、オープン・イノベーションとエコシステムが、将来の成長戦略にとって重要であると回答した。今後 2 年間、エコシステム・パートナーを選ぶ上で、彼らの生成 AI 機能が判断に影響するとの見方が多数を占め、クラウドにおけるパートナーの場合 59%、ビジネス全体におけるパートナーでは 62% だった。

生成 AI が成熟していけば、リスクを軽減しつつ価値を実現する可能性はどんどん高まっていくだろう。リスク管理とレジリエンス強化の両面で幅広い能力を構築した企業は、この最新テクノロジーによって「より速く」「より遠くへ」進むことが可能になり、将来の成長を確保する足固めが整う。

経営層の **84%** は、
サイバーセキュリティー・
ソリューションについて、
従来型よりも生成 AI の活用を
優先する計画だと答えている。

3. サイバーレジリエンス + 生成 AI

リーダーが
実行すべきこと →

「スピードと規模」を軸にサイバーセキュリティー投資の方向性を見直す

セキュリティーを強化するために不可欠なツールとして AI を利用する。

また生産性の大幅な向上とビジネスの成長を実現する手段として、生成 AI と自動化をツールキットに組み込んで、セキュリティー・リスクとインシデント対応に迅速かつ広範に対処するよう、サイバーセキュリティー・リーダーに促す。

AI を活用し、セキュリティーの成果創出を加速させる。

1. 従業員が戦略的で重要なタスクに集中し、成果創出を加速・強化できるよう、定型タスクなどの自動化や効率化の推進が必要である。
2. キーとなるアクションは以下のとおりである。
   （ア） 人間の専門的な経験や判断を必要としない定型タスクを自動化する。
   （イ） 人間とテクノロジーが協業可能なタスクを洗い出し、生成 AI を活用することでそのタスクを効率化する。具体的には、セキュリティー・ポリシーの作成や、脅威ハンティング（ネットワーク、各ノード、データに潜む脅威の特定）、さらにインシデント対応などだ。

AI の導入によって新たな脅威の検知を迅速化する。

1. 今後、さらに脅威の加速や多様化が予想されており、従業員が攻撃者と同等のスピードや規模、精度、技術水準で対応できるようになることが求められている。
2. 生成 AI を使用してパターンや異常をより迅速に特定し、新たな脅威ベクトル（不正アクセス方法や攻撃箇所・攻撃ツリー）によってビジネスに支障が出る前に従業員が検知できるよう、企業のセキュリティー技術と知識を最新化する。

「協業の力」を強みにする。

1. 企業の包括的なビジネス成長のためには、セキュリティー対策においても信頼できるパートナー、エコシステムとの協業が重要となる。
2. キーとなるアクションは以下のとおりである。
   （ア） パートナーの協力を得て、自社の AI セキュリティーをどう完成させるかを明確化する。
   （イ） パートナーと協力して包括的な生成 AI 戦略を実行し、組織全体で価値創造を推進する。

日本語翻訳監修
大西克美

日本アイ・ビー・エム株式会社
IBM コンサルティング事業本部
サイバーセキュリティー・サービス
セキュリティー CTO　技術理事

IBM Institute for Business Value

CEO のための生成 AI 活用ガイド・シリーズ

サイバーセキュリティー

本レポートに記載されているインサイトは、IBM Institute for Business Value がオックスフォード・エコノミクス（Oxford Economics）社の協力を得て実施した 4 度の独自調査のほか、2023 年の「IBM Cost of a data breach report」および 22 年の「Open the door to open innovation」で得たデータに基づいている。第 1 回調査は、生成 AI がサイバーセキュリティーに与える影響について、米国企業の経営層 200 人を対象に 23 年 9 月～10 月に行った。第 2 回は、生成 AI がハイブリッドクラウドに与える影響について、同 414 人を対象に 23 年 5 月～ 6 月に実施。第 3 回は、生成 AI と AI 倫理について、同 200 人を対象に 23 年 8 月～ 9 月に行った。第 4 回は、生成 AI が労働に与える影響について、同 300 人を対象に 23 年 5 月に実施した。

ibm.co/ceo-generative-ai-cybersecurity

IBM Corporation
New Orchard Road
Armonk, NY 10504

Produced in the United States of America | October 2023

本書は英語版「The CEO's guide to generative AI: Cybersecurity - When it comes to cybersecurity, fight fire with fire」の日本語訳として提供されるものです。

DBMQB8RE-JPJA-00