

●学会動向報告

特段ありませんでした

●研究発表

永井郁敏 会員：生成AI時代のサイバーリスク

－ 要 約 －

企業へのサイバー攻撃は急増し、そのコストは数兆円にも達している。その攻撃ターゲットは、グローバル企業から自衛手段を持たない中小企業や医療・教育機関と多岐にわたる。特に中小企業は自社を守る準備ができておらず、また復旧させるためのリソースも無いのが実情です。その被害は、ダウンタイム（システムが停止している時間）により、1分あたり数十万円、1時間あたり数千万円、1日あたり数億円の損失を被る可能性があるかと推定されています。つまり、いまサイバー攻撃を受けると、事業存続に関わる重大な損害につながるわけです。また企業連携の観点から、ネットワーク化された組織やサプライチェーンでは、その被害は一社一国にとどまりません。セキュリティの観点からも、事業体としての可用性を高め、これらを保証する必要があります。そして生成AIです！ その開発スピードは規制や基準の制定を凌駕しており、すべての面で後手に回っている。サイバー攻撃を仕掛ける側と防御する側双方で、AIを用いた攻防戦がはじまっている。今回は、事業継続性リスクの観点からサイバーリスクにフォーカスし、組織として求められるガバナンスを考えます。

●質問&意見交換（抜粋）

- ・ランサムウェアのワクチンについて、ランサムウェアごとに個別の対応が必要
- ・ランサムウェアのテロリストに対して、逆にマルウェアを送りつけることはできないのか？ 攻撃側が有利で防御側が不利であると感じる。
- ・ランサムウェア対策を国際的に対応することはしていないのか？ →国際的にも連携はとって対応はしているが、相手の所在地特定が難しい（インターネットであるため）
- ・生成AI等が発達しさらに悪事に用いられることが考えられる。国家セキュリティとして安全保障レベルでの枠組みを検討し対応する必要があるのではないか
- ・米国および欧州でAIの倫理的取り組みが始まっている
- ・取締役会の中で、日本の現実的な状況はどのようになっているのか？ →大手はCIOや外部の知見者を同席させているところはまだまだ少ない状況。そもそもセキュリティーエンジニア自体の絶対数が少ない。その具体的な方策をとっている企業も少ない状況
- ・インシデント開示におけるガイドラインor制度はどの程度確立されているのか？ →警察庁管轄なので、まずは連絡する必要があるが、情報を公開していない企業も多くあると推察される

- 生成AIのマルチモーダル機能を用い、ロゴマーク等も真偽の判断がつかないものが増えてきている
- 迷惑メール対策について。送り元のメールアドレスが常に変ったメールが来るが、その原理は？ →自動的にメールアドレスを生成して送り付けている。この部分の判定にAIを組み込み迷惑メール対策が可能になってきている。
- 個人レベルでのウィルス対策等、会員間で対策への経験を共有した
- アルトマン（オープンAI社）の件で、AIの安全性についての議論の中で、エシックスの解釈の違いがあること。また、一企業の中での顧客の業界等の違い等、AIからの判断を人間としてどれだけ判断ができるかが求められる。つまり人材育成が求められる。ローカルな言語（ex.方言など）で対応することで、それらが回避できないだろうか？ →判断する力、それをどのように高めるかが大切だが、難しいテーマ。
- ジェネラティブという言葉一つでも、クリエイティブとの違いを私たちはどう判断理解すれば良いか、「つくる」という理解で果たして良いのだろうか？ →考え方を考えると言った哲学的視点が求められるのではないか？
- 事業継続等のリスク対策を施しても、テロ行為による影響を企業として受けてしまった場合、その企業にどのような過失や責任があるのか？ 妥当な対応策はどんなレベルなのか？ →企業としてどこまで求められるのか？セキュリティレベルのチェックをする会社（評価会社）があるので、こう言った評価機関からの評価を受けることが一つ。また、取締役会でセキュリティ対策やリスクを評価し、その評価内容を議事録として残す。また、社員構成員全員に対して制約文書に署名させる等の対応が考えられます。
- 外部機関による評価を公表する必要があるか？ →公表は逆にハッカーによる標的になりかねないため、痛し痒しの部分がある。セキュリティー計画を明確に立て実施していくことが求められる。
- ID管理について。 →システムに対する管理者権限規定がまず求められます。システム側からの変更を要求するような対応も必要（当日言及はしていませんが、生体認証へ移行することで、パスワードを利用する必要がなくなります。このファイルをハッキングされるとアウトですが…Nagai）
- テロリストの本来の目的がわからない場合が多いため、対策も立てづらいのが実情だ

以上