

AI時代のサイバー攻撃、度重なる情報漏洩からの守りを固めるための政府への7つの提言私案

個人情報の漏洩事案が過去に例をみないペースで公表されている。現在見えている事案はあくまで氷山の一角で、公表や検知すらされていない事案も数多く存在すると思われる。また、サイバー攻撃能力の高いAIの拡散によって今後さらに状況が悪くなる可能性が高い。ライフラインに関わる領域が攻撃されれば産業や人命にも関わる。そのような重大事案が発生する前に一刻も早く動かなければならない。一方で、具体的な打ち手に関する議論は国会の内外でまだまだ足りていないと感じる。そこで本記事では具体的に政府がとるべき打ち手を7つ提案したい。

今回の提言にあたり、多くの当事者、有識者に協力をいただいた。漏洩を経験したり不安を感じたりしている方々とは1,251件、セキュリティの専門家や事業者の方々とは278件のAIによる匿名インタビューセッションを実施した。また、セキュリティの専門家の方にオンライン/オフラインのインタビューも7件実施した。

下記に示す見解は党を代表するものではなく、安野個人としての提言である。ゆえに、積極的にまだ粗いアイデアも入れているしやや総花的である。本記事はオープンβ版のような立ち位置のものであり、是非、有識者の方にはツツコミをいただきたい。記事末尾にはコメント用のGoogle Docsへのリンクもある。引き続き党内外での議論を通じて磨き込んだ上で、可能であれば党として正式に諮った上で、関係各所に働きかけていこうと考えている。

状況認識

過去最悪ペースで情報漏洩事案が公表されているが、これはまだ氷山の一角である

国内におけるサイバー攻撃の被害は、2026年に入って急激に拡大している。民間の集計(humanhacker.ai)によれば、国内で公表されたサイバー事案は1月の29件から8月に67件、9月には75件へと増えた。別の集計(トレンドマイクロ)でも、企業や自治体が今年公表した不正アクセスなどの被害は9月末で600件に達し、2025年通年の593件をすでに上回っている。

漏洩の規模も大きい。東京商工リサーチの調査によると、100万件以上の個人情報漏洩は、2026年10月5日時点で10件に達した。2025年通年の6件をすでに上回り、1,000万件以上の漏洩に限っても、2012年から2025年までの14年間で3回だったものが、2026年だけですでに3回起きている。

守りが固い組織でも被害は相次いでいる。政府機関のデジタル庁、金融機関の大和証券、クラウド基盤を提供するIDCフロンティアなどが相次いで被害を受けた。IDCフロンティアの事例では、契約する495の企業・自治体に影響が及んだ。

ここまで挙げただけでも過去最悪のペースだが、重要なのは、公表されている漏洩が氷山の一角にすぎないということである。

行政への報告件数と、企業が自主的に公表する件数との間には、極めて大きな隔たりがある。英国政府のCyber Security Breaches Survey 2025/2026によれば、侵害を経験した企業のうち、最も影響の大きかった侵害を組織外へ報告したのは40%で、外部へ報告した企業の中でも警察へ届け出たのは6%にすぎない。米国では2023年1月にFBIがランサムウェア集団Hiveを摘発した際、長官は被害者のうち法執行機関に届け出たのは約20%だったと述べている。日本に

は公的統計が存在しないが、英米の報告率を踏まえれば、公表されている事案の背後により多くの侵害事案が起きている可能性が高い。

高いサイバー攻撃能力を持つAIが誰でも手に入るようになり、今後状況はさらに悪くなる可能性が高い

8月28日、中国のZ.aiは新しいモデル「GLM-5.3」をオープンウェイトとして一般公開した。Anthropicが公表した検証では、既知の脆弱性を対象にしたベンチマークにおいて、GLM-5.3は410回の試行中50回、動作する攻撃コードを完成させたという。2026年4月に提供された「Claude Mythos Preview」は、同じ評価で56回であり、GLM-5.3の性能はほぼMythos並みと言える。

GLM-5.3には悪用されないように安全装置がついているものの、比較的容易に回避が可能である。公開から数日のうちに、安全装置を外した改変版が複数の開発者によってインターネット上で配布された。Abliterationと呼ばれる手法を使うと、大規模なファインチューニングをしなくても、計算コスト70万円程度で、モデルの能力をほぼ保ったまま有害な要求への拒否を大幅に弱められる。

筆者は2026年4月15日の参議院デジタル社会・AI特別委員会で、米国企業以外も含む他社がMythos Preview相当の能力を持つモデルを開発できるまで早くて数か月、遅くても1年以内であり、AIを使った攻撃能力は早期に拡散する恐れがあるため対策を急ぐ必要があると申し上げていた。実際にこれが実現し、高い攻撃能力を持つオープンウェイトモデルがMythosの登場から5か月ほどで世界中に拡散してしまった形となる。

悪用から検知、公表までにかかるリードタイムを踏まえると、10月初旬までに公表されている漏洩の多くは、まだこのレベルのAIによるものではないと考えるべきだ。侵入から検知までにかかる時間の中央値は14日(Mandiant M-Trends 2026)、国内の不正アクセス事案で発生から公表までの中央値は56日、発覚から公表までは平均77日であり、技術の拡散から実際の公表までは時間差がある。これまでの事案でもAIは活用されているが、一世代以上前のモデルである可能性が高い。

したがって、Mythos級のモデルが拡散した影響が本格的に公表件数に現れ始めるのはおそらく26年の年末以降になると思われる。今見えている被害はまだ序の口である。誤解をおそれず言えば、最悪の場合、情報漏洩のような事業停止を伴わない被害ではなく、ランサムウェア被害のような事業停止を伴う被害が今のペースで起きるような事態も想定しうる。

状況認識については下記の動画等でも話をしたので興味がある方はご覧いただきたい。

AI時代のサイバー攻撃から守りを固める7つの提言

1) AIにはAIを: 人間が守るという前提を捨て、AIで継続的に守りを固める

AIによってさまざまな前提が変わった。脆弱性の公表から悪用までの時間は年々短くなり、脆弱性公表以前に悪用されるケースも増えた。つまり、パッチを素早く適用したとしてもシステムは守れない。既製ソフトウェアの欠陥だけでなく、個別企業が独自に構築したアプリケーション固有の

欠陥も、AI が見つけ出すようになっている。攻撃はますます安価になり、2026 年 9 月に報告された事例では、攻撃者が標的 1 社に費やした AI モデルの利用料は平均4,000円前後であったという。攻撃対象を慎重に選ぶのではなく、脆弱性があれば無作為に誰であれ攻撃されるようになった。この事態に対応するためには人間だけでは限界がある。だからこそ、防御側もAIを最大限活用しなければならない。

具体的には下記を行う。

- **AIによる継続的な脆弱性チェックを推奨する。** 専門家インタビューにおいても「資産の棚卸しやリスクアセスメント、監査は年1回が基準で、環境の変化に追いついていない」という意見があった。ISMSやプライバシーマークなどの認証においても年に一度のチェックが標準的である。これを見直し、年に一度ではなく、継続的に検証と修正を繰り返す運用を標準にする。米国FedRAMPという認証では、エクセルやPDFでチェックシートに回答する従来の形式ではなく、jsonで検証項目が連携され、AIエージェント向けの指示が Agents.mdによって整備され、数日に一度自動的に監査をする仕組みが実装されている。米Vanta社はAPIで企業のシステム群(クラウド、社内ツールなど)に直接つないで、セキュリティ監視とコンプライアンス対応を自動化するツールを提供している。人間の介在を前提としない(Human in the loopではない)チェック体制として参考にするべきである。
- 国内の防御側が継続的にフロンティアモデルを使えるよう、外交で確保する。米CISAは2026年6月にMythos Previewへのフルアクセスを得たと報じられた。Anthropicは10月6日に審査を通過した専門家・組織向けのCyber Verification Programを3段階に拡大し、最上位モデルへのアクセスを広げた。日本においてもYATA-Shieldなどの枠組みがあるが、政府は継続的に複数のフロンティアラボと交渉を続け、日本の機関、企業がアクセスを得られるようにすべきである。
- 国が高度なオープンウェイトモデルの検証環境を用意し、計算資源の支援をする。セキュリティ企業が今後はプライベートモデルだけではなく、オープンウェイトモデルを活用するケースが増える可能性がある。これに備え、ABCI等の既存の公的計算基盤に、自己ホスト型モデル(GLM-5.3などのオープンウェイトモデル)の検証環境を置き、審査制・利用ログ付きで認定事業者や企業に貸し出す。
- 企業に補助金ならぬ「補助トークン」を配る。セキュリティ検証用のAI推論のトークンを、用途・金額上限・成果指標を定めて配布し、より多くの企業が安価にシステムの守りを固められるようにする。既存のサイバーセキュリティお助け隊サービスや、デジタル化・AI導入補助金のセキュリティ対策推進枠を拡張する形を想定。補助金だけではなく補助トークンとして現物支援も行うことで、直接的にAIによる守りを届けることができる。
- 国産のサイバーセキュリティ専門モデルの開発を支援する。機密情報を国外に送らずに防御の検証を行えること、国内の業務システムに対応できること、海外事業者の提供方針が変わっても防御を継続できることを中長期で目指す。国産であることだけで安全と判断せず、モデルや運用環境の検証もあわせて行い、国内に開発・評価・運用の技術と人材を蓄積するべきである。
- 日本AISIのモデル評価能力を拡充する。日本のAIセーフティ・インスティテュート(AISI)は、関係機関と協力してAIのサイバー攻撃能力を継続的に評価し、変化があったときに各所に連携できるようにするべきである。

2) 被害を受けた個人に対してプッシュ型で支援を行う: 信用情報機関への報告、集団訴訟、損害賠償請求までの自動化を目指す

いくら対策をしても残念ながら被害は発生する。そこで、被害にあった方をサポートし、救済する仕組みづくりをすべきである。被害に遭われた方のインタビューでは「何をすればいいかわからな

かった」という声が非常に多かった。日本では集団訴訟のハードルは高く、弁護士を雇うコストも見合わないため、情報漏洩事案では泣き寝入りするしかない場合も多い。

だからこそ、被害を受けた方が能動的に動かなくても、「プッシュ型」で必要な支援が受けられることを目指すべきである。理想は情報漏洩の通知から、信用情報機関への報告、オプトアウト型集団訴訟、賠償金の請求までが自動で行われる状態と考える。

具体的には以下を実施する。

- 漏洩の有無を、本人が探しに行かなくても届く形にする。タイムズカーの事例では、退会者は専用フォームから自ら問い合わせる必要があり、本人が能動的に動かない限り漏洩の有無がわからなかった。退会者にもマイナポータル等の到達可能な経路で届けられる仕組みを作る。本人が次に何をすればよいか分かり、その場で手続きや相談につながることを目指す。
- ワンボタンで信用情報機関への報告が完了できるようにする。漏洩した本人確認書類でクレジットカードやローンを組まれないようにする必要がある。そのためには信用情報機関(CIC・JICC・全国銀行個人信用情報センター)の「本人申告」を行う必要があるが、現状は3機関に別々に500～1,000円を払って申し込む必要があり、与信を拘束する効力もない。米国の2018年の連邦法を参考に、信用凍結の無料化と、1機関に連絡すれば他の機関にも通知される仕組みを整えるべきである。
- オプトアウト型集団訴訟の枠組みを作り、被害者が自ら動かなくても賠償を受けられるようにする。韓国の制度を参考に、被害者が損害額を立証しなくても一定の賠償を受けられる「法定損害賠償制度」や、事前の意思表示を要しない「オプトアウト型集団訴訟」の導入を検討する。
- 仮に漏洩してもなるべく個人の不利益にならないようにする。例えば免許証を再発行したときの番号付番方法を変更し、再発行回数を推測できなくする。免許証を再発行すると免許証番号の末尾の数字が一つ繰り上がるようになっている。この数値を参照して利用者の信用力を推し量る事業者が存在している。たとえば一部のレンタカーの会社では免許証番号の末尾の数値に応じてサービス提供の可否が判断されてしまっており、これは情報漏洩の被害者の不利益を助長することになる。

ただし、プッシュ型の通知は本物と偽物の見分けがつきにくく、詐欺に悪用されやすい手法でもある点については慎重に検討が必要である。特に漏洩直後は、補償や被害確認を装った詐欺が起きやすい。これはマイナポータル等の信頼できる経路に限定する、仕組みの周知を徹底していく等の工夫により中期的には乗り越えられるのではないかと考える。

3) まじめな企業が馬鹿を見ない健全なインセンティブ設計をする

まじめに守った企業が報われ、守らなかった企業が相応の責任を負うべきである。企業の当事者のインタビューでも「まじめに公表すると叩かれる。隠そうとした事業者の方が得をする構造はおかしい」、「現場がリスクを訴えても、経営陣がセキュリティ対策をケチろうとする」という声があった。正直者が馬鹿を見ない健全なインセンティブ設計として、具体的には下記を行う。

ムチ: 責任と罰則を明確にする

- セキュリティ対策を取締役の責任として義務づけ、株主代表訴訟の対象とする。対策を怠って重大な漏洩を招いた場合に取締役個人の責任が問われることを明確にし、訴訟リスクという抑止力によって、セキュリティを現場任せではなく経営の課題として扱わせる。

- 上場基準として、**CISO(最高情報セキュリティ責任者)**や**セキュリティ担当取締役**の設置を努力義務とする。女性役員の登用に関する事例が参考になる。東京証券取引所は2023年、上場規程を改正し、プライム市場の上場企業に「2025年を目途に女性役員を1名以上選任するよう努める」ことを求めた。罰則のない努力義務だったが、女性役員がいない企業の割合は、改正時の18.7%から2025年には2.5%まで下がった。セキュリティも同じ手法で、経営の意思決定の場に責任者を置かせることができる。
- 個人情報保護のルールの中に、多要素認証(MFA)の義務化と脆弱性チェックの義務化を盛り込む。対象となる事業者の範囲と猶予期間を定めたうえで、最低限の対策を「推奨」から「義務」に引き上げる。
- 消費者裁判手続特例法を見直し、集団訴訟をしやすくする。1人数千円の慰謝料しか認められない構造を変え、集団的な救済を可能にして、企業に「守る方が安い」というインセンティブを作る。現状、情報漏洩の慰謝料として裁判所が認めた額は、1人あたり数千円から1万円程度であり、個人が単独で裁判を起こすのは、コストパフォーマンスが合わない。消費者裁判手続特例法を見直し、財産的請求を伴わない過失による情報漏洩の慰謝料請求も対象に含める。特定適格消費者団体の財政支援や、情報漏洩に対応できる団体の認定も進める。前述した法定損害賠償制度・オプトアウト型集団訴訟と組み合わせ、実効性を高める。
- 個人情報の安全管理措置義務違反に対する課徴金に、重過失を対象として加える。2026年改正個人情報保護法(2028年7月までに施行)で導入が決まった課徴金は、不正利用・不正取得・違法な第三者提供・統計特例違反の4類型に限られ、安全管理措置義務違反による漏洩は対象外とされた。重過失など一定の要件のもとで、算定方法と適切な防御を尽くした場合の扱いを定めたうえで課徴金の対象に加える。

アメ: 守る企業に報いる

- 税制優遇措置をつくり、セキュリティ投資する企業に報いる。SCS評価制度の取得や継続的検証の実施、古いシステムのリプレイスを税制・保険料・公共調達で優遇する。
- 流出した「個人情報」の中身を踏まえた対応を行う。定義上は同じ個人情報の漏洩であっても、ハッシュ化された識別子が漏洩することと、平文のパスワードや本人確認書類の画像が漏洩するのでは、実質的な意味合いはまったく異なる。先進的な企業では多層防御を採用しており、ハッシュ化された識別子よりも本人の名前や住所等の情報を重点的に防御している。しかし、報道等では漏れた情報の件数に着目されがちである。情報の種類と保護の状態に応じて責任の重さを明確に区別し、守っていた企業と守っていなかった企業が同じ扱いにならないようにする。
- 外部からの脆弱性の発見を、修正につなげる仕組みを強化する。事業者自身の検証に加え、外部の専門家が問題を発見し、安全に報告できる体制を整えるべきである。日本では、外部の専門家が脆弱性の存在を確かめようとすると、善意であっても不正アクセス禁止法などに抵触するおそれがあり、免責の仕組みもない。2003年のACCS事件で、脆弱性を指摘した研究者が有罪となって以降、指摘そのものをためらう萎縮が業界に残っている。一方で、海外では対応が進んでいる。米国司法省は2022年、害を最小限にとどめる善意のセキュリティ研究は訴追しないという方針を打ち出した。ベルギーは2023年に法改正を行い、速やかに国の窓口へ報告し、無断で公開しないなどの条件を守った発見者を刑事責任から免除している。日本でも、IPAなどの公的窓口への速やかな報告、必要最小限の確認、無断公開や情報の持ち出しをしないことといった条件を明確に定めたうえで、その範囲内の行為を処罰しないセーフハーバーを法律で設けるべきである。あわせて、運営者が応答しない場合でも取扱いを打ち切らず、所管省庁を通じて修正を促す仕組みを整える。

- 政府は、調査を許諾する対象・手法、報告窓口、受領・回答の期限、修正責任者を明示する「脆弱性開示プログラム」の普及を進める。政府システムなどから、発見者に報奨金を支払う「バグバウンティ」も導入する。既存制度では扱いにくい情報の公開状態や、報告後も対応が進まない案件については、IPA等が関係機関と連携して是正につなげる経路を整える。

4) 企業に対する注意喚起とパトロールの強化

特に中小企業においては危機意識が低く、サイバーセキュリティに対する投資が少ない。既存のセキュリティ補助金は制度はあるもののあまり使われていない。セキュリティ対策推進枠への申請があったのは、中小企業全体に対しては年0.02%に満たず、お助け隊サービスの累計導入も中小企業の0.3%程度である。

IPAの2024年度調査では、中小企業の62.6%が直近3期にセキュリティ投資をしておらず、その理由の最多は「必要性を感じていない」44.3%だった。補助や認証を増やしても、サイバーセキュリティの重要性が理解されておらず需要がそもそもない。認識を変えるための注意喚起や、異変を検知したら能動的に呼びかける必要がある。警察官が街をパトロールするのと同様に、国が民業を圧迫しすぎない範囲でサイバー空間のパトロールを行う。

具体的には以下を行う。

- 今まさに警戒が必要な動向についての注意喚起機能を強化する。流行し被害が出る攻撃の手口は共通していることが多い。特定の手法の注意喚起が被害の予防に有効である。日本では10/7～10/9に3日間で8機関がそれぞればらばらに注意喚起を行った（個人情報保護委員会、JPCERT/CC、警察庁、国家サイバー統括室、総務省、経済産業省、IPA、金融庁）。専門家へのインタビューではこれらの注意喚起が「網羅性がありすぎていて、言っていることもばらばらである」という声があった。米国ではFBI（捜査機関）とCISA（防御機関）が、捜査で得た最新の手口を、経営層向けの要約・具体的な対策・機械で読める痕跡情報をそろえて連名で公表し、状況の変化に応じて改訂している。日本においてもこのような取組を参考にしつつ注意喚起機能を強化するべきである。
- 外部露出しているIPアドレスに対し無償の脆弱性スキャンを全中小企業に提供する。英NCSCのEarly Warningは13,178組織が登録し年間31万個以上のIPアドレスに警報を送り、米CISAのCyber Hygieneは10,890組織に無償の脆弱性スキャンを提供している。公開されたIPアドレスのパトロールを行うことで特に弱い中小企業の守りを底上げする余地がある。

5) 事業者が持たなければならないデータをなるべく最小化する

- 本人確認はJPKI（マイナンバーカードの電子証明書）を標準とする。規制業種以外（カーシェア、不動産、ゲームの年齢確認、会員サービス）にもJPKI活用を広げるべきである。カードを持たない人向けの代替手段確保についても検討する。JPKIを規制業種以外に普及させるため、導入費用やAPI利用料の支援を行い、低コストでAPIを使えるようにする。
- JPKIで必要な情報のみを連携する「属性限定開示API」を実装する。本人確認（署名用電子証明書）を行うと、氏名・住所・生年月日・性別の基本4情報がまとめて事業者に渡り、必要な属性だけを選んで開示できない。「18歳以上か」といった真偽情報だけを返す「属性限定開示API」を整備し、使える要件も広げるべきである。運転資格の確認などの情報に対しても開示できるよう拡張する。

- 法令上の保存義務を総点検する。業種ごとの本人確認書類の保存義務を横断的に点検し、「法定記録(確認した事実・番号・日付等)」と「原画像」を区別した一覧と法的根拠を速やかに公表する。法的義務のない業種における原画像の長期保管に対しては、個人情報保護委員会が注意喚起を出し、スマートフォン撮影で本人確認を行う全事業者の実態総点検を求める。タイムズカーの漏洩事案では免許証の原画像データが長期にわたって保管されていたことが事態の悪化を招いていた。
- マイナンバーカードの取得は、中長期で義務化を検討する。インターネット社会で個人の身元を守るには、誰もが電子証明書で本人性を証明でき、画像のコピーでは本人確認が通らない状態をつくるのが社会の土台になる。

6) 人材・技術・窓口・統計などセキュリティを支えるリソースにより投資する

- 人材育成: 17万人の人材が不足している状況を改善する。国内のセキュリティ人材の需給ギャップは、ISC2の推計で2023年に約11万人、2024年には約17万人に広がった。国家資格「情報処理安全確保支援士」の登録者は「2020年までに3万人超」を掲げたものの2026年4月時点で26,453名に留まった。大学・高専のセキュリティ系定員を戦略的に増やす他、リスキリングを行うことで他職種の人(例えばソフトウェアの専門家)がセキュリティの専門家に転向できるような制度を作る。また、その際にはセキュリティ人材の解像度を高め、AIで検証する人、アプリを修正する人、ネットワーク・権限・復旧を設計する人、経営側で判断する人など、より必要な人材増や育成枠を精緻化する。
- 専門人材登用と人事制度改革を行い、デジタル庁、IPA、国家サイバー統括室(NCO)、消費者庁、個人情報保護委員会、AISIの専門人員を増強する。インタビューでは、「公的機関からの協力要請に対する報酬が市場相場の10分の1程度であった」「週に3日、4日の出社が必要なため協力が困難であった」という声があった。NCO以外においては人員は数年ごとにローテーションされるため、支援していても数年ごとに知見がリセットされてしまうという声もあり、セキュリティの分野においてはローテーションのあり方を見直す必要がある。
- 研究開発(1/3): 情報が漏洩しても「読めない」システムを作るため、秘密計算の研究開発を行う。秘密計算とはデータを暗号化したまま計算処理を可能にする暗号技術の総称である。他にもPETs(Privacy Enhancing Technologies)と呼ばれるプライバシーを担保するための技術群が存在する。現在はまだ利用可能な範囲や安全性については限定的だが、将来有望な技術の1つと考えている。
- 研究開発(2/3): 形式検証の研究開発を進める。ソフトウェアを数学的に検証する形式検証は、安全性を高める技術である。検証済みマイクロカーネルseL4は2026年にArm向けの機密性証明が完了し、検証済みコンパイラCompCertは航空機で認証を得ている。ナビエ・ストークス方程式等のミレニアム問題を解く能力をAIが獲得しつつある今、形式証明の応用可能性は広がっている。今後の技術開発によって、OSレベル、アプリレベルでの機密性証明が可能になれば、ソフトウェアの安全性をいっそう高めることが可能となる。
- 研究開発(3/3): メモリ安全言語への移行に関する研究開発を進める。メモリ安全言語とは、バッファオーバーフローや不正なメモリ参照などのメモリ関連のバグや脆弱性を防ぐ仕組みを備えたプログラミング言語であり、たとえばRustなどが該当する。米国ホワイトハウスは「ソフトウェアはメモリ安全になるべき」との声明を出している。日本においても政府情報システムの調達要件にメモリ安全への移行計画を中長期で組み込むべきである。
- 消費者ホットライン188を支える消費生活センター・国民生活センターの相談体制を増強し、漏洩相談に対応するAIチャット窓口を導入する。実施した漏洩当事者へのインタ

ビューでは「疑わしい程度だと門前払いされる印象」、「対応が合っているかすぐ確認できるチャットがほしい」という声があった。

- サイバー攻撃の実態の情報収集能力を強化する。本稿の冒頭で述べたとおり、国内の漏洩が実際に何件起き、どれだけ公表されているかは、現在の資料からは正確に分からない。被害を受けた多くの企業で攻撃手法が公開されることはない。マクニカ セキュリティ研究センターの調査によると2026年7月以降に公表された国内WEB侵害81件のうち、公表内容から原因を分類できなかったのは65件にものぼる。ゆえに、状況把握のためにも警察等と連携した企業への聞き取り機能の強化と情報の集約が必要である。

7) 情報漏洩以上の危機事態に備え、ライフラインの守りを固める

現在、情報漏洩事案が目立っているが、サイバー攻撃によってより深刻な事態が発生する可能性がある。社会に必須のライフラインが被害を受ければ、産業や人命に多大な影響がある。金融の勘定系システムや電力や水道の供給、工場、病院、インターネット接続などに対して今までになかったレベルでの損害が発生するような事態も想定すべきである。

ライフラインに影響した例は実際に存在する。2021年、米コロニアル・パイプラインでは、VPNアカウント経由の侵入で、東海岸の燃料輸送の半分弱を担うパイプラインが1週間近く停止した。国内でも2023年に名古屋港の統一ターミナルシステムが侵入され、コンテナの搬出入が約3日間止まった。大阪急性期・総合医療センターでは2022年の攻撃で電子カルテが約6週間停止した。

このような事案を想定し、震災に備えた防災訓練を実施するのと同様に、サイバー攻撃に備えた訓練や法整備を社会全体で行う。具体的には以下を実施する。

- 企業・自治体はサイバー攻撃による長期停止を想定して事業継続計画を作り、防災訓練のように訓練を実際に行う。攻撃時の一律停止がかえって危険になる設備もあるため、切り離す判断、安全に動かし続ける判断、停止する判断を事前に定めておく。政府は重要インフラや地域の医療機関、自治体に参加する訓練を支援する。訓練の際にはIT/OT資産の棚卸し等も実施する。
- サイバー版事故調査委員会を創設する。航空・船舶・鉄道などの分野では「なぜ起きたか、どうすれば再発を防げるか」を調べる独立機関として事故調査委員会が存在する。航空の安全性が劇的に上がったのは、一件一件の事故を徹底的に分析し、その結果を設計や運航の基準に反映し続けてきたからであると言われている。重大事案発生時に国が権限をもって調査できるようにすべきである。
- AIによる大規模サイバーインシデントにおいても緊急処理事態の認定をできるようにする。武力攻撃事態対処法第22条の緊急処理事態では、サイバー攻撃だけによる具体的な適用範囲について法律上明確な基準が示されていない。国民生活に甚大な影響が発生したとき、政府は法的根拠を持って省庁横断的な緊急対応を行えるようにすべきである。AIシステムの自律的な動作によって引き起こされた場合など、従前の考え方では攻撃主体が定義すらできない場合においても適切な対応をできるようにすべきである。

おわりに

本稿を書いていたときに思い出したのは『サイバーパンク2077』というSF作品である。この作品の中では、インターネットは危険な場所として描かれている。暴走したAIによって旧来のネットが汚染され、その全体が「ブラックウォール」と呼ばれる壁の向こうに封じられ使えなくなったという設定である。もはやこのような設定はフィクションだけの話ではないと考える。誰もが安心できる

情報化社会をAI時代においても維持するために、危機感を持って早急に対策を進めるべきである。

最初に述べたように、本記事は安野個人としての提言であり、荒削りのアイデアもなるべく多く入れ込んでいる。論点は多々あるため、是非、有識者の方にはフィードバックをいただきたいと思っている。

最後に、インタビューや調査に協力していただいた皆さんに心から感謝したい。調査に協力いただいたセキュリティ業界のエキスパートの皆様からいただいた声からは「危機感が伝わっていない」というもどかしさを強く感じた。実際に課題を解決できるよう、引き続き党内外での議論を通じて磨き込み、関係各所に働きかけていきたい。