サイバーデセプション

― 攻撃者心理を逆手に取る能動的防御戦略 ―

サマリー版

ICSCoE 第９期中核人材育成プログラム 卒業プロジェクト

2026年6月

このホワイトペーパーで伝えたいこと

サイバーデセプションは、攻撃者の認知・判断・行動に働きかける防御である。

従来型の防御は、攻撃を「防ぐ・見つける」ことが中心だった。
しかし、攻撃者の侵入を完全に防ぐことは難しい。

そこで、防御側は発想を変える。
攻撃者心理を踏まえて環境や得られる情報を設計し、迷わせる・遅らせる・誘導する・観測する。
サイバーデセプションは、既存の防御を置き換えるものではない。
防御側の選択肢を広げる、追加的な防御レイヤーである。

ホワイトペーパーの作成背景

攻撃者優位の現状に対し、技術的防御に「人・攻撃者の認知」という視点を組み込み、
日本のサイバーレジリエンス向上に貢献したい。

■課題

- 日本における「サイバーデセプション」の**認知度が低い。**
- 戦略、設計・運用、既存防御策との連携などについて、実務で活用できる形で整理された**情報が限られている。**

■目的

ホワイトペーパーでは、サイバーデセプションについて以下の3点を整理する。

- 必要性や有効性を整理し、企業・組織における導入検討の**きっかけ**を提供する。
- 基本概念、類型、効果と限界を**体系的に整理**し、共通理解を形成する。
- 戦略的な導入方法、既存対策との連携、留意点を整理し、**実務に適用**できる形で提示する。

サイバーデセプションというアプローチ

既存対策は攻撃側に対して「受動的」なアプローチであり、**防御側が常に後手に回る不利な構造**にある。その結果、低い攻撃コストに対して、防御コストは高くなっている。この構造を変えるアプローチとして、サイバーデセプションが有効な選択肢となる。

■ 攻撃者優位の非対称性

境界防御の限界

攻撃を「100%防ぐ」ことは非常に困難である。

アラート過多の弊害

過剰な検知ノイズが運用を疲弊させている。

■ サイバーデセプションの活用

既存の多層防御基盤

【本番環境/業務システム】

- 境界防御（FW/IPS等）
- 端末監視（EDR等）

サイバーデセプション

【本番環境に埋め込む罠】

- 偽のパスワードやファイル

【独立・隔離された罠】

- おとりネットワーク

すり抜けた未知の脅威 → **罠による捕捉・遅延・誘導**

既存の多層防御を基盤としつつ、
攻撃者を欺く能動的な防御レイヤーの導入

デセプション(欺瞞)の概念と攻撃者にもたらす影響

デセプションとは**「相手の認知・判断・行動に影響を与え、自らに有利な結果を導くための意図的な働きかけ」**、
実体を隠し、虚像を見せることで、攻撃者の判断を誤らせる。

■デセプション技術を体系化したモデル「Bell-Whaleyフレームワーク」

実体を隠す技術		虚像を見せる技術	
Masking（隠蔽）	背景に溶け込ませて不可視化する。	Mimicking（模倣）	実在する物を模倣し、本物と思わせる。
Repacking（偽装）	別の無害なものに見せかける。	Inventing（創出）	本来は存在しない対象や情報を作り出す。
Dazzling（幻惑）	大量のノイズで、真実を見分けにくくする。	Decoying（誘引）	注意を引き、特定の方向やおとりに誘導する。

■デセプションが攻撃者にもたらす影響

- 人間は限られた時間と情報の中で判断を下す際に、すべての選択肢を論理的に検証できず、直感や経験則に頼ってしまう**「限定合理性」という制約**を持っている。
- デセプションは、攻撃者に意図的に実態を隠し、虚像を見せる（例えば、偽の情報や環境を提示する）ことで、**攻撃者の認知バイアスを利用し、誤った判断に導く。**

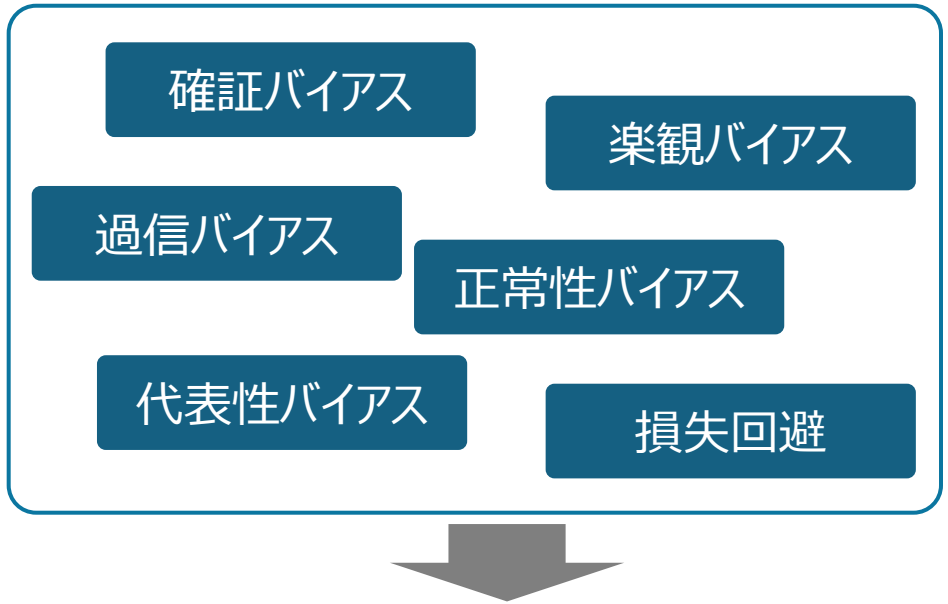
サイバーデセプションと戦略的な設計・構想

（第4章、第5章）

デセプションをセキュリティ戦略に組み込み、攻撃者に作用させる考え方を**攻撃者心理**から紐解く。
サイバーデセプションを戦略的な設計・運用に必要な**基本概念**について整理した。

■攻撃者心理に作用するサイバーデセプション

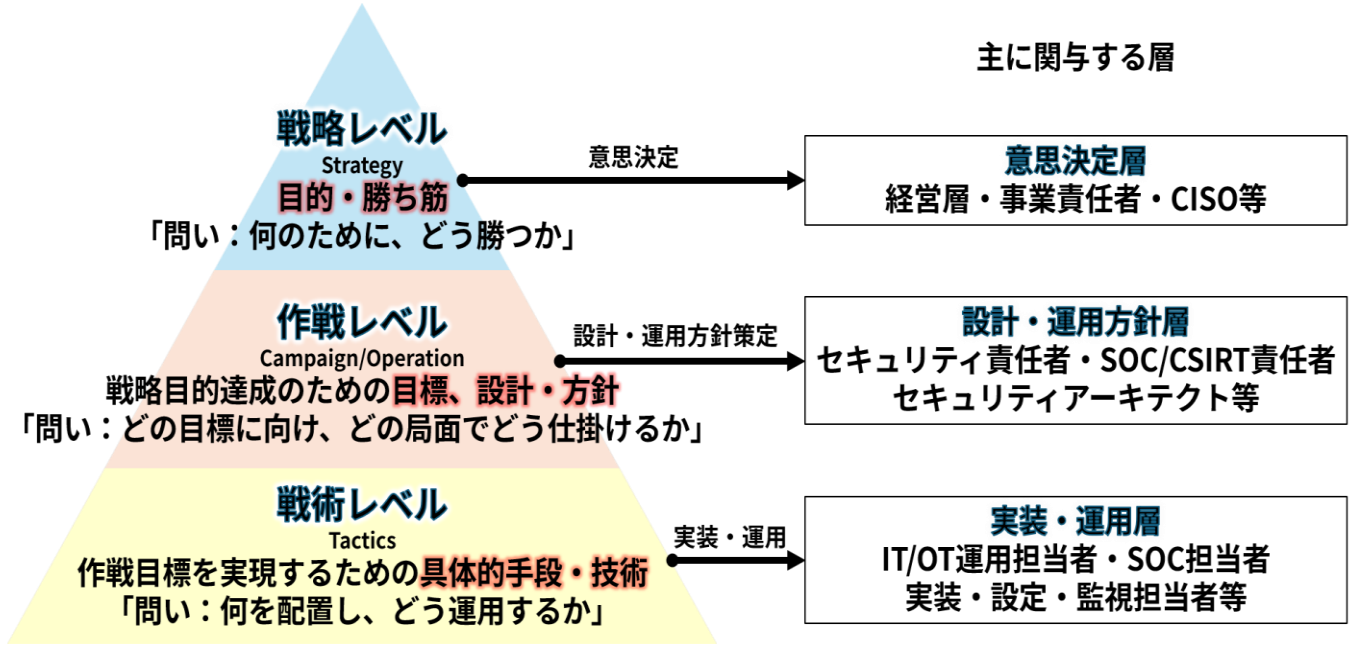
- 攻撃者の意思決定に対し、サイバーデセプションによる**認知バイアス**を作用させる。



- 攻撃者の判断を**防御側に有利な方向へ誘導**することが重要

■戦略的な設計・運用に必要な基本概念

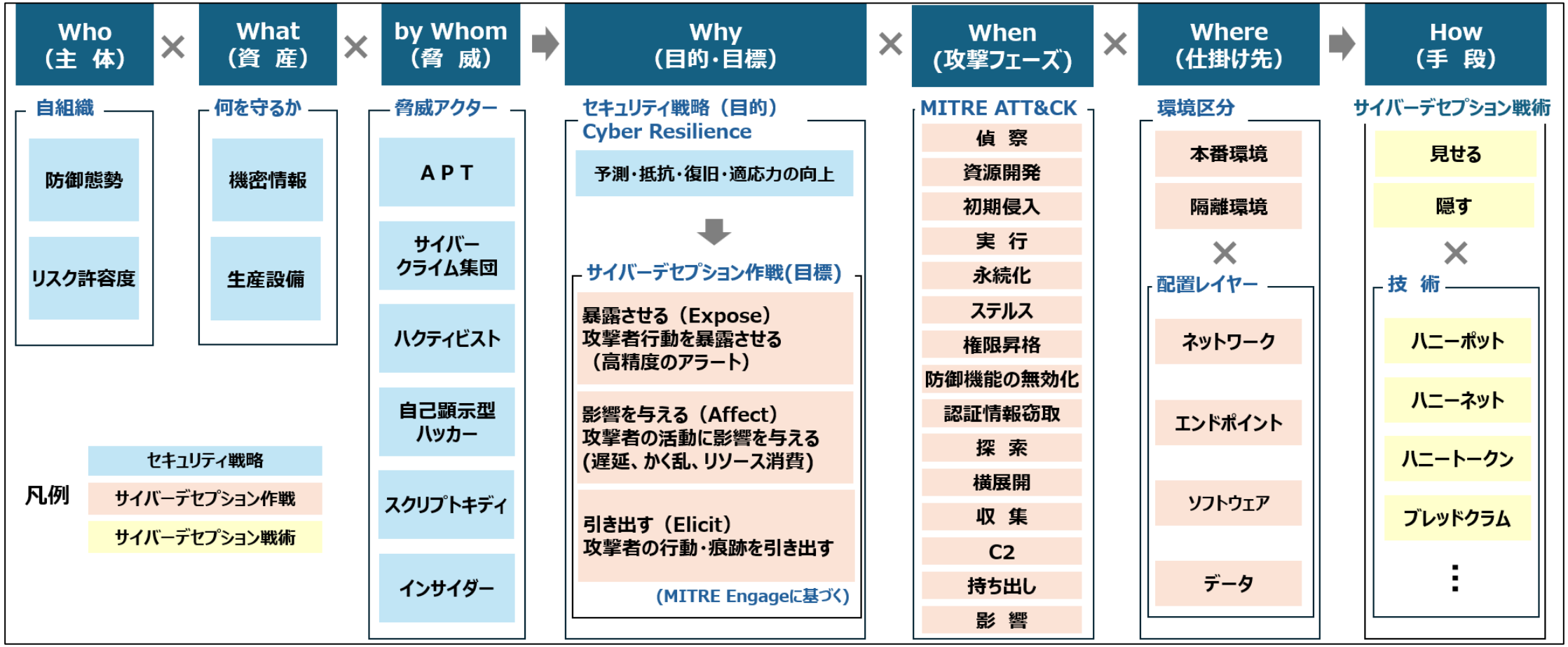
- サイバーデセプションを単なるおとり環境の配置ととらえると、効果は十分に発揮されない。
- **戦略・作戦・戦術**の3レベルで整理し、各レベルで関与する層と担う役割を明確化した。



サイバーデセプションを戦略的に組み込むための実践的フレームワーク

（第5章）

セキュリティ戦略に組み込むための実践的フレームワークとして**CyberDeception Design Framework(CDF)**を提唱する。「ハニーポットを導入する」といった技術起点ではなく、セキュリティ戦略（目的）から考えることが重要



「何を置くか」ではなく、「**何のために、誰から何を守るか**」から設計する。

CDFを実務に落とし込むための設計・運用プロセス

（第6,7章）

第5章で提示したCDFを起点に、サイバーデセプションを「構想」から「設計・運用・評価・改善」へ落とし込む。
重要なのは、技術を選ぶことではなく、**目的・配置・運用・評価を一貫して設計**することである。

■設計・運用プロセスの全体像

1　基本設計
- 作戦目標、配置方針、連携方針、運用方針を整理する。
- 攻撃フェーズ、仕掛け先、期待効果を明確化する。

2　詳細設計
- 技術要素ごとの役割、導入効果、構成パターンを具体化する。
- 攻撃者心理、判断傾向、配置上のポイントを整理する。

3　実装・運用
- 安全性、隔離、監視、ログ取得、SOC/CSIRT連携を考慮する。
- 正規業務への影響や法務・倫理面の留意点を確認する。

4　評価・改善
- 検知、誘導、遅延、観測などの効果を評価する。
- 運用負荷や安全性を点検し、継続的に見直す。

■実務導入時の着眼点

1 目的から逆算する
何を置くかではなく、何を達成したいかを先に決める。

2 攻撃者に自然に見えるか
配置場所、文脈、デコイ環境との整合性を確認する。

3 運用できる設計にする
アラート、ログ、監視、対応手順まで含めて設計する。

4 継続的に見直す
放置せず、効果・負荷・安全性を評価し改善する。

#	狙う効果	Why	When	Where	How
1	境界における偵察行為の情報収集	Expose（Collect）	【偵察】Nmap などによるポートスキャン/サービス探索	【ネットワーク層】インターネット、DMZ、外部公開サービス周辺	ハニーポット（低対話型、偽公開サービス）による偵察行為の観測
2	未知の侵害経路と初期侵入の早期検知	Expose（Detect/Collect）	【初期侵入】VPN/Web ログイン画面、公開アプリへの認証や脆弱性悪用試行	【ソフトウェア層】VPN ポータル、外部公開 Web、認証画面	ハニーポット（偽VPN/偽ログイン画面）による侵入試行の検知

（表）CDFに基づく設計項目の具体化例（一部抜粋）

CDFは、構想で終わらせず、**設計・運用・評価・改善に落とし込んで初めて機能**する。

検証項目・選定観点とその狙い

（第8章）

第6章の設計例から、**攻撃者の認知・判断・行動への作用**が表れやすい３つを選定し、検証環境にて確認した。
ここでは、各検証項目について選定観点とねらいを整理する。

検証項目	選定観点	ねらい
① 応答遅延を組み込んだ低対話型ハニーポットの検証	**探索コストの増大**	応答遅延により、攻撃者が偵察に要する時間を増大させられるかを定量的に確認する。
② メタデータを作り込んだハニーファイルの検証	**真正性判断への影響**	ファイル名に加えて、タイムスタンプやファイルサイズなどのメタデータを自然に作り込むことで、攻撃者の真正性判断を困難にできるかを確認する。
③ 配置文脈を変えたハニークレデンシャルの検証	**偽認証情報による次行動の誘発**	偽認証情報の配置方法や文脈の違いが、攻撃者の誘引効果やおとり環境への誘導に与える影響を確認する。

サイバーデセプションの今後の課題

（第9章）

サイバーデセプションは従来の防御手法を補完する**追加的な防御レイヤー**である。導入すれば自然と効果を発揮するものではなく、適切な目的設定や継続的な評価・改善を行い、実効性を確保していくべきである。

■今後の課題

1 検証結果の一般化と評価手法の高度化

検証結果をより信頼性の高いものとし、サイバーデセプションの効果を客観的に評価できる仕組みを確立する。

主な課題

- 検証環境やAIシミュレーションには限界がある
- 攻撃者像や環境による差異が大きい
- 効果を測定、評価する共通指標が不足している

2 実務適用に向けた設計・運用モデルの確立

効果を最大化しつつ、安全に運用し続けるための設計・運用モデルを確立し、組織に定着させる。

主な課題

- おとりを設置するだけでは効果が出ない
- 設計等が不十分な場合ののっとり等のリスクがある
- 運用負荷や投資対効果の判断が難しい

3 多層防御の一翼としての戦略的位置づけ

サイバーデセプションを多層防御の一部として位置づけ、AI時代の攻撃者にも対応できる戦略へと進化させる。

主な課題

- 個別技術として捉えられ、戦略的活用が進まない
- AIを活用する攻撃者への対応が不十分
- 動的、リアルなおとりの設計が必要

【参考】ホワイトペーパーにおけるコラムについて

ホワイトペーパーの中には、デセプションをより身近に捉えるため、各章の終わりにコラムを入れている。技術論だけでは見えにくい「人が何を信じ、なぜ行動するのか」という視点から、デセプションの世界に関心を持っていただきたい。

日本人は騙すことが嫌い？

「欺瞞」という言葉への抵抗感を、国民性という見方も手がかりにしながら、デセプションを“だます技術”ではなく、相手の認知や判断に働きかける防御の考え方として捉え直す。

ゲームで考えるサイバーデセプション

ダンジョン探索型ゲームを題材に、攻撃者が見ている環境そのものを設計し、判断を迷わせるという発想を直感的に理解する。

偽物の美学

贋作の世界を題材に、本物らしさを支える細部、来歴、文脈の作り込みから、サイバーデセプションに必要なリアリティを考える。

欺瞞の食卓

食品における“本物らしさ”や代替・加工の工夫を手がかりに、見た目や文脈が人の判断に与える影響を考える。

イギリス人はデセプションがお好き？

第二次世界大戦中の欺瞞作戦を題材に、デセプションを成功させるための細部へのこだわりと、相手に信じさせるための演出を考える。

「デセプション導入済」って、言っちゃう？

サイバーデセプションを公表することが、攻撃者の種類によって抑止にも挑発にもなり得るという、公表判断について考える。

なぜキズだらけのマグロが注目されるのか

豊洲市場での実体験をもとに、人が「他者の関心を示す痕跡」や「重要そうに見えるシグナル」から価値を推測する仕組みを考える。

自然界の欺瞞とサイバーデセプションのこれから

自然界に見られる欺瞞を手がかりに、サイバーデセプションが今後どのように発展し得るかを考える。

著作権及び利用条件等

著作権及び利用条件

本書に関する著作権及びその他すべての知的所有権は、「情報処理推進機構 産業サイバーセキュリティセンター 第９期中核人材育成プログラム　卒業プロジェクト　サイバーデセプション　―攻撃者心理を逆手に取る能動的防御戦略―」（以下、「本プロジェクト」という。）に帰属する。

本書は、企業、団体、行政機関、教育機関その他の組織または個人が、サイバーデセプションに関する理解促進、調査、検討、教育、実務上の参考資料として活用することを目的として提供するものである。

特に、権利者の承諾を得ることなく、無断で以下の行為を行うことを禁止する。

- 本書の全部または一部を、書籍、雑誌、Webメディアその他出版物などに転載、再配布または販売する行為
- 販売、宣伝、営業活動を目的とした教材、セミナー資料、研修コンテンツ、商品・サービス資料などへ転用する行為
- 本書の内容を改変し、本プロジェクトの成果物であることが判別できない形で利用する行為

本書を引用する場合は、引用部分を明確に区別したうえで、出典を適切に明記するものとする。本書に記載された第三者の著作物、商標、製品名、サービス名などに関する権利は、それぞれの権利者に帰属する。本書におけるこれらの記載は、説明上の便宜を目的とするものであり、本書が特定の製品、サービス、企業などを推奨、保証、または支持するものではない。

本書を利用する者は、以上の条件を理解し、これに同意したうえで利用するものとする。

免責事項

本書は、サイバーデセプションに関する調査研究の成果を情報提供の目的で整理したものであり、特定の製品、サービス、技術、対策の導入または利用を推奨・保証するものではない。

本書の内容は、作成時点における情報、文献調査、検証結果および本プロジェクトの考察に基づくものであり、サイバー攻撃の手法、技術動向、法制度、標準化動向、製品仕様、運用環境などの変化により、予告なく変更される場合がある。

本プロジェクトは、本書の内容について、正確性、完全性、有用性、最新性、商品性、特定目的への適合性その他いかなる事項についても、明示または黙示を問わず一切の保証を行わない。

本書の利用により生じた直接的、間接的、付随的、特別、派生的または結果的損害、ならびに利益の損失、事業上の損失、データの喪失、システム障害その他一切の損害について、本プロジェクト、本書の著者、関係者および監修者は、法的原因の如何を問わず一切の責任を負わない。

本書には、外部資料、Webサイト、ツール、製品、サービスなどへの参照またはリンクが含まれる場合があるが、本プロジェクトは、それらの内容、正確性、継続性、安全性、適法性その他一切について責任を負わない。

本書の有効期限は、発行日から2年間とする。ただし、有効期限内であっても、技術動向、脅威環境、法制度、標準化動向などの変化により、内容が実態に適合しなくなる可能性がある。本プロジェクトは、本書の内容を常に最新の状態に更新・修正する義務を負うものではなく、実態との乖離によって生じたいかなる損害についても責任を負わない。

注意事項

本書は、本プロジェクトの調査研究および検証結果に基づいて作成されたものであり、独立行政法人情報処理推進機構（IPA）ならびに本プロジェクトメンバーの所属組織の公式見解を示すものではない。

本プロジェクトにおける調査・検証の対象期間は、2026年4月から2026年6月末までである。

本書で取り上げる検証結果は、限られた期間、環境、条件のもとで実施したものであり、すべての組織、システム、攻撃者、運用環境にそのまま適用できるものではない。

実際にサイバーデセプション技術を導入・運用する場合は、各組織のシステム構成、保護すべき資産、想定する脅威、リスク許容度、法務・コンプライアンス要件、運用体制などを踏まえ、個別に検討する必要がある。

本書では、文章の構成検討、表現の推敲、図表・挿絵の作成補助に、ChatGPT等の生成AIサービスを一部利用している。生成AIの出力は、執筆者が確認・編集したうえで使用しており、本書の内容に関する最終的な責任は執筆者に帰属する。