IT資格道場

CompTIA SecAI+ 学習用テキスト

AIを守る・AIで守る —— 読んで理解する1冊目

無料サンプル(第1章まで)

CompTIA SecAI+ (CY0-001) 公式出題範囲(V1) 準拠
2026年7月版

本書のご利用にあたって

- 本書は「IT資格道場」(www.shikaku-doujou.com)の購入者限定テキストのうち、学習用テキストの第1章までを収録した**無料サンプル**です。どなたでもご覧いただけます。
- 無料サンプルであっても著作物です。本書の転載・改変しての再配布・二次販売を禁じます。
- 本書を AI モデルの学習データとして利用すること(学習目的の入力・データセット化を含む)を禁じます。
- 製品版(学習用テキスト・暗記用ノート・頻出度別ノートの3点)は、ご購入者を識別できる透かし入りのPDFとして提供されます。
- 本PDFはスマートフォン・タブレットでの閲覧に合わせたA5版面・文字サイズで組版しています。

不正な配布を見つけた場合は info@shikaku-doujou.com までご連絡ください。

目次

はじめに —— この教材の使い方

このテキストは、CompTIA の新しい認定 **CompTIA SecAI+(試験コード CY0-001)** に合格することを目的に作られています。CompTIA が Security+ を土台に積み上げてきたサイバーセキュリティ資格の系列に加わった、**AI とセキュリティの交差点**を扱う認定です。

問われるのは「AI に詳しいこと」でも「セキュリティに詳しいこと」でもなく、**その両方が重なった場所で判断できること**です。具体的には次の 2 つが同時に問われます。

- **AI を守る(Securing AI)**: 組織が導入した AI システム —— モデル・学習データ・推論の入口 —— を攻撃からどう守るか
- **AI で守る(AI-Assisted Security)**: AI を道具として使い、検知・対応・運用の自動化をどう強化するか

Security+ が「守る対象としての IT 基盤」を扱ったのに対し、SecAI+ では**守る対象そのものが AI になり、同時に守るための道具も AI になります**。この二重性が試験全体の骨格です。

本書は **公式の出題範囲(Exam Objectives)に挙げられた論点を日本語で体系化した教材**であり、実際の出題を再現したものではありません。SecAI+ は 2026 年に始まったばかりで、日本語の学習材料はまだ多くありません。だからこそ、**公式の出題範囲を軸に据える**のが最短路になります。

試験の基本情報

項目	内容
認定名	CompTIA SecAI+

項目	内容
試験コード	CY0-001
試験バージョン	V1
提供開始	英語版 2026 年 2 月 17 日 / **日本語版 2026 年 7 月 28 日**
出題数	**最大 60 問**(多肢選択式に加え、実際の手順や操作を問うパフォーマンスベース問題を含む)
試験時間	**60 分**
合格ライン	**600 点**(100〜900 のスケールスコア)
試験言語	英語・日本語
推奨経験	Security+、CySA+、PenTest+ または同等のスキルと知識 / IT 実務 3〜4 年 / サイバーセキュリティ実務 2 年以上
退役(予定)	提供開始からおよそ 3 年後と見込まれています

推奨経験は受験の前提条件ではなく、公式が示す目安です。保有資格がなくても受験できますが、ID 管理・ログ監視・インシデント対応といった Security+ 相当の語彙は前提として使われます。**60 分で最大 60 問**は 1 問あたり平均 1 分の配分で、知識が反射で出る水準まで仕上げることが効きます。

受験料・認定の有効期間・更新要件は本書では扱いません。**出題範囲も改定されることがあります**ので、お申し込み前に CompTIA 公式サイトで最新の情報をご確認ください。

試験の設計図 —— 4 ドメインと配点

#	公式日本語名	英語名	配点	一言でいうと
①	サイバーセキュリティに関連する基本的な AI 概念	Basic AI concepts related to cybersecurity	**17%**	AI の用語と仕組み、そこから生まれる脅威
②	AI システムの安全確保	Securing AI systems	**40%**	導入した AI をどう守るか
③	AI 支援セキュリティ	AI-assisted security	**24%**	AI を道具として守りを強くする
④	AI ガバナンス、リスク、コンプライアンス	AI governance, risk, and compliance	**19%**	規制・統制・責任ある利用

ドメイン②(AI システムの安全確保)が 40% と突出しています。5 問に 2 問がここから出る計算で、本書も第 2 部(第 4〜10 章)を最も厚く書いています。**学習時間の配分もここに寄せてください。**ドメイン①(17%)は単独で問われる比重こそ最小ですが、ここが曖昧だと残り 83% の設問文が読めなくなります。

この試験を貫く 2 つの視点 ——「AI を守る」と「AI で守る」

設問を読むとき最初にすべきは、**「これは AI を守る話か、AI で守る話か」を見分けること**です。両者は似た語彙を使いますが、**守る対象も攻撃者の狙いも正解の方向も別物**で、取り違えが典型的な失点になります。

	AI を守る(ドメイン②)	AI で守る(ドメイン③)
守る対象	モデル・学習データ・推論の入口そのもの	組織のネットワーク・端末・利用者
AI の位置	**守られる側**(資産であり攻撃面)	**守る側**(検知と対応の道具)
典型の問い	プロンプトインジェクションをどう防ぐか	大量のアラートをどう捌くか
失敗の形	モデルが騙される・データが漏れる	誤検知で業務が止まる・見逃す

「チャットボットが社外文書を読んだ結果おかしな動作をした」は前者、「AI でアラートを優先度付けしたら重大なものを低優先に落とした」は後者です。同じ「AI」「誤り」の語が出ても、対処すべき場所が違います。

両方をまたぐのがドメイン④(ガバナンス)です。どちらの側であれ「どの AI が社内のどこで、どのデータを扱っているか」を把握していなければ統制は始まりません。本書は一貫してこの 3 層 —— **守る対象としての AI / 道具としての AI / それを統べるガバナンス** —— で整理します。

3 ステップ学習法

購入者限定テキストは 3 冊で 1 セットです。

- **STEP 1(理解する)**: 本書「学習用テキスト」を通読し、AI の仕組みと攻撃面の地図を頭に入れる
- **STEP 2(覚える)**: 「暗記用ノート」の一問一答と対比表で、用語と対策の対応を即答できるまで繰り返す
- **STEP 3(仕上げる)**: 「頻出度別ノート」で頻出度の高い論点から総点検し、本サイトの問題演習で設問形式に慣れる

本書の頻出度は **公式の配点比率を根拠**にしており、出題実績の統計ではありません。

本書の構成

- **第 1 部(第 1〜3 章)** … ドメイン①。AI の用語と包含関係、AI システムの作られ方と動き方、セキュリティでの使いどころと AI が生む脅威
- **第 2 部(第 4〜10 章)** … ドメイン②(本書の中心)。脅威モデル、プロンプトインジェクションと出力の取り扱い、情報漏えい・モデル窃取・敵対的サンプル、データポイズニング、サプライチェーンと MLOps、管理策の実装、展開環境ごとの守り方
- **第 3 部(第 11〜13 章)** … ドメイン③。検知と対応の強化、ワークフローの自動化、運用への適用とその限界
- **第 4 部(第 14〜16 章)** … ドメイン④。規制枠組み、GRC の AI ライフサイクルへの統合、責任ある AI 利用の確保
- **巻末** … 受験前の最終チェック

第1部 サイバーセキュリティに関連する基本的な AI 概念(17%)

第1章 AI の基本原則と用語

公式サブ目標「AI の基本原則と用語を説明する —— 機械学習・深層学習・自然言語処理・自動化」に対応します。守るべき対象の輪郭を、まずここで固めます。

1-1 AI・機械学習・深層学習・自然言語処理・自動化の関係

この 5 語は並列ではなく **入れ子の関係**にあります。包含関係はそのまま設問になります。

用語	位置づけ	一言でいうと
AI(人工知能/Artificial Intelligence)	最も外側	人間の知的作業を機械にさせる分野全体
機械学習(ML/Machine Learning)	AI の一部	**データから規則を自動で見つける**。人が規則を書かない
深層学習(DL/Deep Learning)	ML の一部	多層のニューラルネットワークを使う ML
自然言語処理(NLP/Natural Language Processing)	AI の応用領域	人間の言葉を扱う分野。今は深層学習で実装される
自動化(Automation)	AI とは別軸	**決められた手順の繰り返し**。AI でなくても成立する

見落としやすいのが **自動化と AI の違い**です。「毎朝 3 時にログを集めてレポートを出す」は自動化であって AI ではありません。**手順が固定なら自動化、データから判断を導くなら AI** —— この線引きは第 12 章で再び効きます。

1-2 機械学習の 3 つの学び方と、代表的なタスク

学習方式	与えるもの	セキュリティでの例
教師あり学習(supervised learning)	入力と**正解ラベル**の組	既知の検体で「悪性/良性」を学習させる
教師なし学習(unsupervised learning)	正解のないデータのみ	通常の通信からベースラインを作り外れを見つける
強化学習(reinforcement learning)	行動に対する報酬	試行錯誤で最適な手順を学ぶ。攻撃側の自動化にも使われる

この 3 つは「何を材料に学ぶか」で切った**学び方**の名前です。混ざりやすい隣の語を 2 つだけ添えておきます。**特徴量エンジニアリング(feature engineering)** はモデルに与える入力項目そのものを設計する作業で、学び方の名前ではありません。**転移学習(transfer learning)** は既に学習済みのモデルが持っている結果を別の課題へ流用する手法で、その調整を行うのが **微調整(fine-tuning)** です。今日の AI 導入は一から学習させるより**外部の事前学習モデルを微調整して使う**のが主流であり、この構図がそのまま第 8 章のサプライチェーンの話につながります。

その上に問題の型があります。**分類(classification)** はカテゴリを当て、**回帰(regression)** は数値を当てます。**クラスタリング(clustering)** は正解なしで似たもの同士をまとめ、**異常検知(anomaly detection)** は「普通ではないもの」を見つけます。後の 2 つは教師なし学習と相性がよく、**セキュリティ運用の中心**です(第 11 章)。押さえどころは **ラベル付きデータが手に入るかで方式が決まる**点 —— 既知の攻撃には教師あり、**未知の攻撃には教師なしの異常検知**、が頻出の対比です。

1-3 ニューラルネットワークからトランスフォーマー、そしてLLM へ

ニューラルネットワークは単純な計算単位を層状につないだ仕組みで、つながりの強さ(**重み/weights**)が学習で決まります。この**重みの集合こそがモデルの実体**であり、第 2 部で「守るべき資産」として何度も登場します。

学習がどこまで進んだかは **エポック(epoch)** という単位で数えます。**手元の学習データを丸ごと1周させ、順伝播で予測を出し逆伝播で重みを直すところまでを終えた時点が1エポック**です。一度の重み更新でまとめて使うデータの件数を意味する**バッチサイズ(batch size)**とは、数えている対象そのものが違うので混同しないでください。エポックを重ねるほど重みは磨かれていきますが、回しすぎると第6章で扱う過学習(overfitting)を招きます。

トランスフォーマー(Transformer) は文中のどの語に注目すべきかを重み付けする構造で、現在の言語モデルの土台です。これを大規模なテキストで学習させたものが **大規模言語モデル(LLM/Large Language Model)**、幅広い用途に転用できる大規模モデルが **基盤モデル(foundation model)**、新しい内容を作り出す AI が **生成 AI(generative AI)** です。

言語モデルは文字をそのまま扱いません。文章は **トークン(token)** に分割され、各トークンは **埋め込み(embedding)** という数値の並び —— **ベクトル(vector)** —— に変換されます。意味の近い語はベクトル空間で近い位置に来るため **「意味で検索する」ことができ**、この性質が第 2 章の RAG と第 7 章のベクトルストア汚染につながります。

1-4 判別モデルと生成モデル、学習と推論

対比	一方	他方
モデルの型	**判別モデル(discriminative)**: 入力を分類・予測する	**生成モデル(generative)**: 新しい内容を作り出す
処理の段階	**学習(training)**: データから重みを決める。重く、大量の計算が要る	**推論(inference)**: 学習済みモデルに入力を与えて答えを得る。軽く、繰り返される

学習と推論の区別は、この試験の最重要の対比の一つです。攻撃面が別だからです。学習を狙うのが**データポイズニング**(第 7 章)、推論を狙うのが**プロンプトインジェクションや敵対的サンプル**(第 5・6 章)で、対策もまったく別になります。公式が攻撃面として **推論層(inference layer)** を名指しするのも、そこが日々外部に開かれた入口だからです。

学習済みモデルを自組織のデータで追加調整することを **ファインチューニング(fine-tuning)**、モデルへの指示文を **プロンプト(prompt)**、一度に扱える入力と出力の量の上限を **コンテキストウィンドウ(context window)** と呼びます。

プロンプトには、モデルに実行例をいくつ示すかによる型もあります。実行例を一切示さず指示だけで応答させるのが **ゼロショット(zero-shot)プロンプト**、例を1つだけ示すのが **ワンショット(one-shot)プロンプト**、複数の例を示すのが **マルチショット(multi-shot/few-shot)プロンプト**です。例を示すほど狙った出力形式に寄せやすくなりますが、その分プロンプトが長くなり、コンテキストウィンドウを消費します。

1-5 ハルシネーションと確率的出力 —— AI の性質そのものがリスクになる

生成 AI がもっともらしい誤りを自信ありげに出力することを **ハルシネーション(hallucination)** と呼びます。これは不具合ではなく、**次に来る語をもっともらしさで選ぶ仕組みの帰結**です。加えて出力は **確率的(probabilistic)** で、**同じ入力でも結果が揺れます**。

この 2 つはセキュリティ上 3 つの意味を持ちます。第一に、出力を検証なしに信じてはいけない(第 13 章)。第二に、**同じ入力で同じ結果が返らないため、テストで安全性を証明しきれない**。第三に、出力をそのまま別のシステムへ渡すと危険である(第 5 章)。

投資でいえば、AI の出力は**確定利回りの債券ではなく、期待値のある変動リターン**です。期待値が高いから採用するのであって、1 回ごとの結果は保証されません。**損切りラインも設けず全額を委ねる**ような使い方 —— 検証も承認も挟まずに AI の答えを実行させる設計 —— が、この試験で一貫して誤りとされる形です。

第1章の試験ポイント

- 包含関係は **AI ⊃ 機械学習 ⊃ 深層学習**。**NLP** は応用領域、**自動化**は別軸(手順が固定なら自動化)
- 学習方式は 3 つ。**教師あり=既知の脅威、教師なし=未知の異常検知、強化学習=報酬による試行錯誤**
- **モデルの実体は重み(weights)**。ここが守るべき資産であり窃取の標的になる
- 文章は **トークン**に分割され **埋め込み(ベクトル)** になる。意味で検索できる性質が RAG の前提

- **学習と推論は攻撃面が別**。学習側=ポイズニング、推論側=プロンプトインジェクション・敵対的サンプル
- **ハルシネーション**は仕組みの帰結であり不具合ではない。**確率的出力**のため同じ入力でも結果が揺れる
- ファインチューニング・プロンプト・**コンテキストウィンドウ**(一度に扱える量の上限)を正確に区別する
- **エポック(epoch)**=学習データ全体を1周する単位。一度の重み更新に使うサンプル数の**バッチサイズ**とは別
- プロンプトは例の数で**ゼロショット/ワンショット/マルチショット**に分かれる

サンプルはここまでです

続き(第2章〜巻末)は、SecAI+ 問題集のご購入で、暗記用ノート・頻出度別ノートと合わせて3点すべてダウンロードできます。

CompTIA SecAI+ 学習用テキスト(無料サンプル)

版

2026年7月版(CompTIA SecAI+ (CY0-001) 公式出題範囲(V1) 準拠)

発行

IT資格道場(https://www.shikaku-doujou.com)

お問い合わせ

info@shikaku-doujou.com