AI は“学習”していない！？
― セキュリティ議論が崩壊する構造的理由 ―

摘要（Executive Summary）

本ホワイトペーパーは、生成 AI に関するセキュリティ議論において頻発する
「AI が勝手に学習している」「入力情報がモデルに蓄積される」という誤解が、
AI の内部構造を正しく分離して理解していないことに起因することを明らかにする。

AI を①Learning（学習）、②推論エンジン、③マンマシン I/F の三層構造として再定義することで、
入力情報がどこで使われ、どこでは使われないのかを構造的に説明する。
その結果、適切に設計された AI 活用は、構造上セキュリティリスクを伴わないことを示す。

2025 年 12 月

著者：前田稔（エムズスタイル LLC）

構成編集：ChatGPT（OpenAI GPT-5）

1. なぜ AI セキュリティ議論は混乱するのか

現在、多くの企業において
「AI を利用する場合は、社内固有の情報の入力を避けるようにする」
といった、セキュリティ基準（？）がされていることが多い。

しかしながら、「社内固有」などの定義が明確にされていないため、多くの利用者が混乱していると思われる。

なぜ、このようなことが起きているのか？

この理由は、「AI の機能内容を正確に理解していないこと」に起因していると思われる。

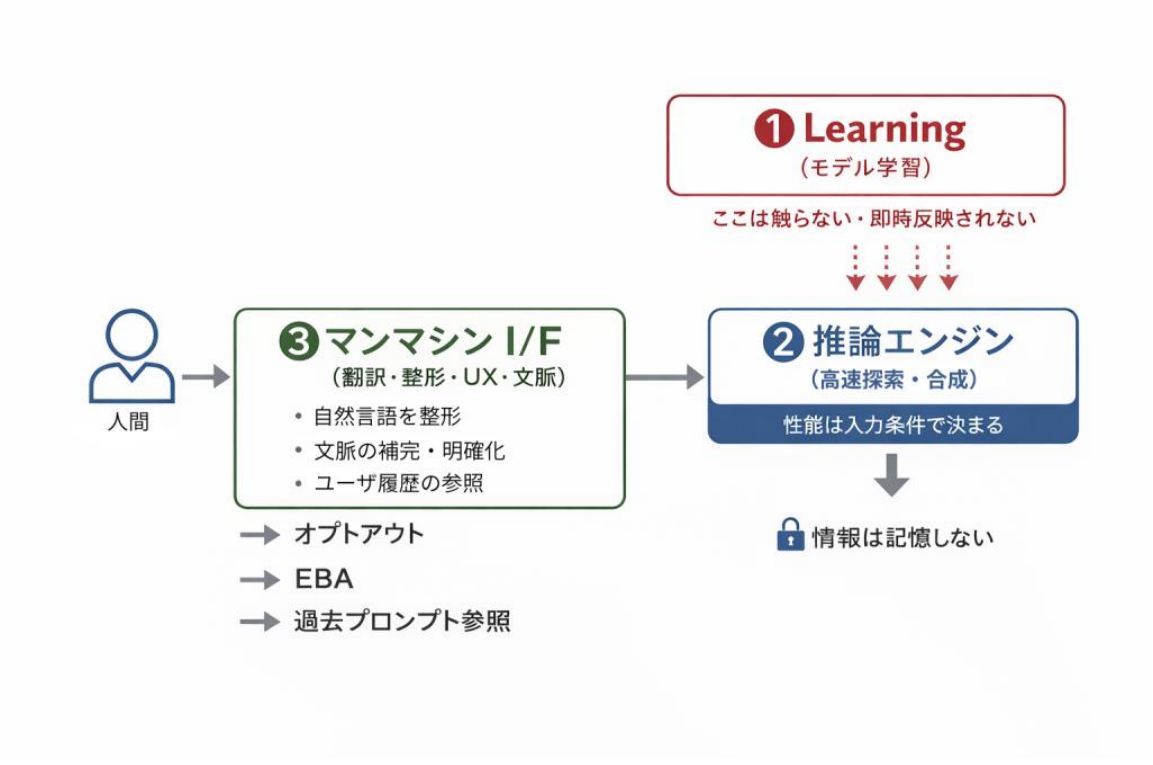
つまり、現在の AI セキュリティ議論では、「学習」「推論」「対話」が区別されないまま語られることが多い。
その結果、本来存在しないリスクが誇張され、逆に本質的な論点が見えなくなっている。

2. AI の三層構造

本ホワイトペーパーでは、AI を次の三層に分離する。

① Learning：モデルの重みを形成・更新する学習層

② 推論エンジン：学習済み構造を高速に探索・合成する層
③ マンマシン I/F：人間の言語を推論しやすい形に整形し、結果を人間向けに再表現する層



①の Learning が正に「モデル学習」の部分である。
この「モデル学習」ではインプット情報に対してその答え＝アウトプット情報を深層学習によって得られる部分となる。ここで学習されてしまうと確かに他の利用者からその学習結果を利用されることになる。

しかしながら、この学習は、インターネットなどで公になっている一般的な情報が用いられている。過去（2023 年より前）においては、入力データもこの学習に使われていたが、現在はどの AI も使っていない。

②の推論エンジンは、③から受け取ったインプット情報に対して、アウトプット情報を探索する機能である。ここには、一切可変的に学習する情報は存在しない。

③はマンマシン I/F。実は、人間のインプット情報はあまりにも曖昧なことが多い。そのまま②の推論エンジンに投入してしまうと推論エンジンが破綻し、答えを返せないことが多い。そのためマンマシン I/F が②の推論エンジンに分かるように入出力を修正している。最近の AI が忖度したりするのはまさにこの部分のせいである。

ここでも学習はされている。文脈の整形、最終的にユーザへ提示するときの文章などである。

ChatGPT などで学習をはずす「オプトアウト」機能があるが、それはこの部分の学習を指すので有り、Learning で言っている「モデル学習」を指しているのではない。

また、よく
「AI はユーザの癖を学習している」
「使っていくうちに賢くなる」
と言っている人がいるが、これも誤解である。
これは正確には「学習（Learning）」ではなく「（過去プロンプトの）参照（Referencing）」である。

※この構造は、コンピュータの構造に近い。つまり①メモリ、ストレージ、②CPU、③キーボード、ディスプレイと言ったものである。

3. セキュリティの本質的な判断基準

セキュリティ上の最大の論点は、「入力情報が①Learning で使われるかどうか」である。
構造上、入力情報が①に一切使われないのであれば、
モデルそのものが汚染されることはなく、学習リスクは存在しない。

③マンマシン I/F で行われるのは、UX 改善・対話最適化のための処理であり、
これは学習ではなくインターフェース改善である。

4. Enterprise 版の意味

Enterprise 版であれば問題ないという人もいるが、これも間違いである。

Enterprise 版だろうが、そうでなかろうが、この AI 構造に関してはまったく違いはない。

ただし、Enterprise 版においては

・オプトアウトがデフォルト

・ログの記録など責任分界点（説明）がされる。

ということである。

いずれにしろ、はっきりするのは、AI は入力やアップロードされたドキュメント情報を「モデル学習」には使っていないということである。

5. 欧米における対応

実務上は、欧米企業においても
「すべてを一律に禁止する」という運用は現実的ではないため、
入力情報を禁止するのではなく、
どのレイヤに影響を与えるかでリスクを切り分ける
という整理が一般的に採られている。

具体的には、
・入力情報がモデル学習に使われないこと
・推論結果が外部に再利用されないこと
・ログや責任分界が明確であること
を前提に、
業務文書や検討資料を用いた AI 活用が許容されているケースが多い。

AI 利用の是非とは無関係に、文書そのものに関する管理規定は、常に別途存在する。これは欧米でも日本でも共通である。
しかしながら、文書管理規定と AI 規定は「そもそもレイヤが異なる」

したがって、実務上の判断は、
「AI を使ってよいか」ではなく、
「当該文書が、既存の文書管理規定上、外部サービス利用が許容されているか」
を基準に行うのが合理的である。

分かりやすく言えば、
その文書がクラウド保存を許容されているのであれば、
同等の契約条件・責任分界を満たす AI 利用も、
原則として同じ扱いになると理解できる。

なお、これは無条件の利用を意味するものではなく、
既存の文書管理規定に定められた条件
（利用目的、契約条件、アクセス制御等）を
満たす場合に限られる。

6. 過去プロンプト参照の正体

ユーザの過去のプロンプトや行動履歴を参照する機能も、三層構造で見れば③に属する。
これは②に渡す前段階で文脈を整理し、推論の前提条件を安定させるための補助である。

結果として、②推論エンジンは論理齟齬のない入力を受け取り、
最大限のパフォーマンスを発揮できる。

7. EBA（EmzStyle Business Advisor）の位置づけ

EBA は①や②に手を加える仕組みではない。
③マンマシン I/F を構造的にアシストし、
人間の思考・前提・制約を整理した状態で②に渡すための仕組みである。

これはマニュアル車と最新のオートマ車の違いに近い。
特にポルシェの PDK が最適な回転数で最適なギアを選択するように、
EBA は推論エンジンが常に最適条件で動作する状態を作る。

8. なぜ正しい説明が出てこないのか

このような構造的説明が普及しない理由は単純である。
AI を実装・運用の構造として理解していない専門家が、
抽象語だけで語っているからである。

これは、車を知らない法律家が自動車事故を語るのと同じ構造的問題である。
三層構造を示せば、反論の余地はほぼ存在しない。

9. 結論

AI は「勝手に学習している」のではない。
学習が行われる場所は構造上明確に限定されており、
適切に設計された AI 活用はセキュリティ上、問題を生じない。

重要なのは恐怖ではなく、構造理解である。

本ホワイトペーパーについて

本ホワイトペーパーは、
前田 稔（エムズスタイル LLC）による独自の調査・分析および
構造知性フレームワークに基づき作成されています。

本資料は、特定の解決策や結論を提示するものではなく、
判断に必要な構造や視点を整理することを目的としています。

著作権・利用条件

本資料の利用条件は、以下に定める
「ホワイトペーパー利用規約」 に従うものとします。
🔗 https://emz-style.com/whitepaper-terms

利用区分の概要

- 無料版（要約・抜粋）
  社内共有・紹介目的での利用は可能です（改変・商用利用不可）
- 有料版（個人）
  個人学習目的に限り利用可能です（社内共有不可）
- 法人向けライセンス
  社内での配布・研修・教育用途での利用が可能です

※詳細は上記利用規約をご確認ください。

最後に

本資料をお読みになり、

- 判断に迷う点がある
- 自社の状況に当てはめると違和感がある
- このまま進めてよいのか確信が持てない

と感じられた場合は、
それ自体が重要なサインです。

ご相談・ご質問は、以下よりお気軽にお寄せください。
🔗 https://emz-style.com/contact
（※法人向けのご相談・講演・研修のご依頼もこちらから承っています）