

AIエージェント PoC 開始前チェックリスト

Feynma Technology

着手後に決めると、動いているものを止めてやり直しになる7項目

生成AIのPoCが本番に進まない理由の多くは、精度ではなく**着手前に決めていなかったこと**にあります。技術検証が終わった段階でセキュリティ部門や法務部門から質問が来て、そこで設計をやり直す。この手戻りが最も高くつきます。弊社がお客様とPoCを始める前に必ず整理している7項目です。**着手前に、この表を埋めてください**。空欄が残る項目は「決める人が決まっていない」というサインです。

7項目のうち、6番が最も高くつく

AIにどこまで操作を任せるかは、動くものができてから議論すると必ず揉めます。「取り消せない操作」「会社が責任を負う操作」「金額が確定する瞬間」を渡すかどうかは、着手前に決めてください。

1 ユーザーのアクセス経路

- ☐ VPN経由のみか
☐ 固定IPからのみ許可か
☐ インターネットからのアクセスを許可するか

なぜ先に決めるか／後から絞ろうとすると、動作確認済みの環境が動かなくなります。

2 利用規模

- ☐ PoC期間中の想定ユーザー数
☐ 本番展開時の想定ユーザー数

なぜ先に決めるか／1人で試すのか30人で試すのかで、選ぶべき構成が変わります。短期のPoCは結果重視で手法を問わなくてよい一方、複数人が使う中長期の運用ではフルマネージドサービスが正しい。**PoCの作り方と本番の作り方は別で考えるべきです**。

3 認証方式

- ☐ サービス側の認証機能をそのまま使うか
☐ 自社のID基盤（Entra ID等）と連携するか
☐ MFA（多要素認証）を必須にするか

なぜ先に決めるか／ID基盤との連携は情報システム部門の作業と承認が必要です。リードタイムが読めない項目です。

4 扱うデータの種類

- ☐ 公開情報のみで検証するか
☐ 顧客データ・機微情報を含めるか

なぜ先に決めるか／含めるなら3・5・6の要件が一段厳しくなります。**PoCは公開データで、という選択も有効です**。

5 データの処理・保存場所

- ☐ 東京リージョン限定とするか
☐ Cross-Region Inference の利用を許可するか

なぜ先に決めるか／リージョン制約は使えるモデルの選択肢を変えます。「あのモデルが使えない」と後から分かったら、精度の前提が崩れます。

6 AIエージェントの操作権限範囲 ★最重要

- ☐ エージェントが実行できる操作の範囲
☐ 重要な操作の前に人間の承認を挟むか。挟むならどの操作か
☐ プロンプトインジェクション（悪意ある指示の混入）への対策

7項目の中で、後から決めると最も高くつく項目です。「取り消せない操作」「社外に対して会社が責任を負う操作」「金額や契約条件が確定する瞬間」をAIに渡すかどうかは、動くものができてから議論すると必ず揉めます。

弊社の判断基準／取り消せない操作は人間に残す。AIには下書きの作成までを任せ、確定値はコード側から渡し、**AIには「どのボタンを押すか」だけを任せる**構成にしています。

7 PoCの可用性

- ☐ バックアップの要否
☐ リリース・更新のタイミング（業務時間中に止めてよいか）

なぜ先に決めるか／PoCとはいえ業務時間中に触られる前提なら、止められる時間が限られます。

この表の使い方

- ☐ PoCの企画段階で、空欄のまま着手しないことを確認する
☐ 埋まらない項目は「決める人が決まっていない」サイン。先に誰が決めるかを決める
☐ 6番だけは技術部門の判断で決める、業務部門・法務と合意する

8 どのセキュリティ基準に準拠していますか

クラウドのマネージドサービス経由で大規模言語モデルを利用する構成では、入力データが追加学習に使われない仕様のものがあります。どのサービスのどの仕様に基づくかを、契約・ドキュメントのレベルで確認して回答できるようにしておくことが重要です。

9 間違った回答をしたらどうするのですか

生成物への注意表示、多層のログ記録、人間の確認を前提とした運用設計。この3点をセットで説明できるようにしておきます。「間違えない」ではなく「間違いに気づける」を設計します。

利用するクラウド基盤が取得している第三者認証（ISO 27001、SOC 1/2/3 等）を根拠として説明できます。自社の対策と、基盤側の対策を分けて示すのが伝わりやすい形です。

✓ 先に線を引く。線の手前は徹底的に自動化する。線の向こうは人間が持つ。

線を引かずに始めたPoCは、たいてい本番に載りません。技術の問題ではなく、怖くて承認できないからです。