デジラボビューティ 白書 No.3

バイブコーディング時代に人間が担う仕事

美容業界のための目的設定・検証・権限管理・運用責任

一般社団法人デジラボビューティ
解説 代表理事 鎌田麻央
2026 年 10 月 7 日 Version 1.4

AI に開発作業を任せるほど、人には、何を実現するかを決め、結果を確かめ、使い続ける条件を整える仕事が求められます。

デジラボビューティから皆さまへ

白書 No.1 は AI の透明性、No.2 は AI の安全性評価を扱いました。No.3 は、これらを開発・運用へ接続し、AI に任せる範囲と人・組織が担う判断を整理します。公的資料と企業の技術資料を先に確認し、その分析を踏まえて美容業務への適用例を示します。

本書では、原資料が示す内容と、当法人による分析・提言を区別します。政府向けの指針、検討段階の案、民間企業の技術設計は、それぞれ対象と位置付けが異なり、美容事業者へ一律に適用される基準ではありません。

経営・運用上の論点は、目的設定、検証、権限管理、監視、停止、復旧です。NVIDIA の Open Agent Safety Platform と、デジタル庁が 2026 年 9 月 29 日に公表した AI エージェント関連資料を含む一次情報を参照し、技術的な制御と組織的な責任分担を整理します。[G7][G8][N1][N2]

最初に知っておきたい 5 つのこと

1．AI による開発支援と、運用中に AI が自律的に操作する仕組みは、確認すべき対象が異なります。

2．公的資料では、人間中心、説明責任、公開前後の検証が重視されています。資料ごとの適用範囲を区別する必要があります。[G1][G2][G3]

3．AI エージェントの安全管理には、出力だけでなく、権限・接続先・中間動作・停止手段の確認が含まれます。[G7][G8][N1][N2]

4．デジタル庁の検討事項案と NVIDIA の技術設計は、一般の美容事業者に直接課される一律の義務ではありません。[G8][N1][N2]

5．本書の 7 つの仕事、30 日間の導入計画、18 項目のチェックリストは、原資料を踏まえた実務提案であり、公的認証基準や実証済みの効果ではありません。

情報基準日 2026 年 10 月 7 日 ／ Version 1.4
本書は当法人による解説・提言であり、政府の白書や公的認証基準ではありません。

今回参照した資料について

内閣府掲載の指針、総務省・経済産業省の AI 事業者ガイドライン、デジタル庁の生成 AI 調達・利活用ガイドラインを、判断・検証・運用管理の背景として参照します。[G1][G2][G3]

個人情報保護委員会の注意喚起と、経済産業省・IPA の人材・スキル資料を、データの取扱いと人材面の参考とします。[G4][G5][G6]

デジタル庁の 2026 年 9 月 29 日資料は AI エージェントの管理に関する検討を含みます。検討事項案を確定した制度と混同しません。[G7][G8][G9]

NVIDIA の公式資料は民間企業の技術設計です。安全管理の具体例として扱い、公的基準とは区別します。[N1][N2][N3]

民間教材は人間の 7 つの仕事を整理する出発点として参照します。美容業務向けの判断条件や成果物は当法人の提案です。[S1]

各資料の発行主体、日付、確認箇所、公式 URL は末尾に記載しています。本改訂は構成・表現の整理であり、外部資料の情報基準日は 2026 年 10 月 7 日のままです。

1 バイブコーディングをどう捉えるか

本書では、バイブコーディングを、自然な言葉で AI に意図を伝え、生成されたコードや画面を動かしながら修正する開発の進め方として扱います。この呼び方の範囲は一定していないため、特定の開発ツールや、コードを一切読まない方法だけに限定しません。

教材「バイブコーディングで人間が担う仕事」は、目的設定から運用責任までを 7 つの役割に整理しています。本書では、その整理を出発点に、美容業界で人が残すべき判断と、記録すべき成果物を具体化します。[S1]

人が責任を持つことと手作業で行うこと

要件の整理、設計案の比較、変更箇所の説明、テストの作成も、AI に支援を求められます。本書の分担は、AI の作業をコード生成だけに限定するものではありません。採用する条件、確認方法、公開の可否について、説明できる人・組織を決めるという提案です。

経営者が全行のコードを読む必要があるとは限りません。ただし、誰が技術的な確認を担当し、何を根拠に合格としたのかは把握する必要があります。判断に必要な知識が不足する部分は、エンジニアや分野の専門家に確認を依頼します。

美容の経験が開発に生きる場面

予約変更でお客様が迷う場所、スタッフが二重入力している項目、カウンセリングで説明が不足しやすい場面は、現場を知る人が見つけやすい課題です。その気づきを、入力項目、表示文、例外時の動きへ置き換えることが、開発における専門性になります。

人の仕事が減る可能性も含めて考える

この白書は、雇用や職種が必ず維持されるという予測ではありません。コード作成や定型作業に必要な時間が減れば、担当する仕事の量や配置が変わる可能性があります。一方で、品質確認や運用を担える人が不足すれば、作れる量が増えても実務への導入は進みません。まず自社の業務単位で、減る作業と増える確認を測ることを提案します。

本章の開発上の整理と美容現場への適用は当法人の解釈・提言です。人材面の背景資料は[G5][G6]を参照。

2 公的資料が示す方向性

政府資料は、バイブコーディングだけを対象とした統一基準ではありません。本書では、AI 利用の判断、人材育成、検証、個人情報の扱いに関係する記述を参照します。以下の右列は、原資料を踏まえた当法人の提案です。

資料	確認した内容	美容現場への提案
内閣府掲載の指針 2025 年 12 月 [G1]	AI を活用する範囲や条件は、人間が最終判断する。規模やリスクに応じて取り組む。	AI へ任せる操作と、担当者が承認する操作を決める。
総務省・経産省 第 1.2 版 概要 [G2]	開発・提供・利用の役割を整理し、人間中心や安全性、説明責任等を示す。	経営者、開発担当者、利用スタッフの担当を明らかにする。
デジタル庁 第 2.0 版 [G3]	政府の生成 AI システムについて、公開前のテストと、公開後の品質確認等を定める。	受入テストと運用記録の作り方の参考にする。
個人情報保護委員会 2023 年 6 月 [G4]	生成 AI への個人情報入力は、利用目的や提供者の取扱い等の確認が必要。	実データの投入前に、送信先・利用条件・保存を調べる。
経産省・IPA DSS ver.2.0 [G5]	DX に必要な役割・スキルを整理。2026 年 4 月改訂ではデータマネジメント等を強化。	作る人だけでなく、業務を設計しデータを管理する人を育てる。
経済産業省 人材・スキル 2024 [G6]	課題を問うこと、仮説の検証、成果物の評価・選択の重要性を示す。	研修では操作の速さに加え、判断の根拠を説明できるかを見る。

適用範囲を混同しない

[G1]は AI 法第 13 条に基づく自主的・能動的な取組を促す指針、[G2]は事業者の取組を支援するガイドラインです。[G3]は政府情報システム等を対象とする文書であり、一般の美容サロンに同じ手続きを直接義務付けるものではありません。[G4]が説明する個人情報保護法上の注意点は、各事業者のデータの扱いに即して確認します。

関連省庁の原資料 4 点を白書掲載ページで提供し、追加資料・IPA 資料等は公式 URL を案内します。確認箇所と位置付けは末尾の参考資料に記載しています。

3 デジタル庁の AI エージェント管理に関する検討

2026 年 9 月 29 日、デジタル庁は第 6 回先進的 AI 利活用アドバイザリーボードの資料を公表しました。その中では、ガバメント AI「源内」で試験提供される AI エージェントチャットや「源内工房」を前提に、AI エージェントのガバナンス、セキュリティ、セーフティが具体的に議論されています。[G7]

源内工房は、自然言語で指示し、AI エージェントがファイル操作、コマンド実行、Web 取得などを数十から数百ステップ連鎖させ、実際に動くプロトタイプを作れる環境として示されています。こ

れは、バイブコーディングが「コードを作る支援」から「複数の操作を自律的につなぐ仕組み」へ広がっていることを示す具体例です。[G7]

最終結果だけでなく、途中の動きを見る

AI エージェントは、最終的な出力が正しく見えても、その途中で想定外のファイルへアクセスしたり、外部サービスへ情報を送ったり、与えられた権限を越えようとしたりする可能性があります。デジタル庁のワーキング・グループでも、AI エージェントは接続先や動作を完全には予見しにくく、意図した通りに動いているかを確認するためには、中間動作と最終結果の両方を見る「プロセスアプローチ」が必要ではないか、という論点が示されています。[G7]

美容の現場へ置き換えると、予約処理の結果が正しかったとしても、顧客情報を不要な外部サービスへ送信していないか、SNS 投稿の途中で未承認の情報を取得していないか、会員管理で本来触れてはいけないデータへアクセスしていないかまで確認する必要があります。

MCP・API・ファイル・モデル更新を「接続の問題」として見る

デジタル庁の資料では、源内工房について、MCP や API 連携を無断で行えないよう外部送信を制限すること、アップロードファイルのマルウェア対策、危険性の高いリポジトリ等のダウンロード制御、基盤モデルの更新を自動化せず管理すること、他者へ共有するスキルの審査などが論点・対策として整理されています。[G7]

ここで重要なのは、「AI が何を答えるか」だけでなく、「AI が何につながっているか」を管理することです。AI エージェントが使えるツール、MCP サーバー、API、データ、外部サイトを一覧にし、必要なものだけを許可するという考え方は、企業の AI 導入でも基本になります。

AI にも ID とライフサイクル管理を

同じ 9 月 29 日の検討資料では、今後の政府ガイドライン充実に向けた案として、AI エージェントごとに一意の ID を割り振り、使用するツールや MCP サーバー等を一覧化したうえで、登録、承認、配布、公開、停止、廃止まで管理することが示されています。また、利用状況と管理責任者を定期的に確認し、脆弱性が判明した場合には、影響を受ける AI エージェントを迅速に特定し、停止または改修できる体制を求める案も示されています。[G8]

これは 2026 年 10 月 7 日時点では「検討事項案」であり、一般企業に直接適用される確定ルールではありません。ただし、AI エージェントが増える企業では、「誰が作ったか分からない」「どのツールにつながっているか分からない」「もう使っていないが動いている」といった状態を避けるための管理方法として参考になります。

人は毎回承認するのか、それとも監督するのか

デジタル庁の議論では、人が処理の途中で判断する Human-in-the-Loop に加え、AI の自律実行を人が監督し、必要なときに介入する Human-on-the-Loop の考え方も論点になっています。一方、行政処分など権利変動を伴う行為については、当面は原則として人の確認を必須とする方向が示されています。[G7]

企業でも、すべてを人が毎回承認すれば自動化の効果が失われます。反対に、重要な操作まで AI へ全面委任すれば、誤りが起きたときの影響が大きくなります。人の関与は一律ではなく、影響の大

きさ、元に戻せるか、個人情報や金銭を扱うか、外部へ公開するかによって決める必要があります。

当法人の分析：デジタル庁の検討と NVIDIA の設計の比較

観点	デジタル庁の検討から読み取れる方向	NVIDIA の技術設計での対応例
権限を絞る	AI が使えるツール・データ・接続先を限定する	OpenShell で実行範囲を制限する
ID と権限を確認する	AI エージェントを識別し、管理責任者・接続先を把握する	BlueField-4／DOCA で ID と委任権限を確認する
途中の動きを見る	中間動作と最終結果を確認する	Sentry が独立した位置から挙動を継続監視する
止められる状態にする	停止・廃止・改修までライフサイクル管理する	境界を越える動きに対して隔離・停止する
人の責任を残す	利用者・管理者・開発者の責任を分ける	共有責任モデルとして複数主体で安全を支える

政府資料と民間企業の技術資料は、適用範囲も目的も同じではありません。それでも両者を並べると、AI エージェントの安全を「AI 自身の判断力だけ」に依存させず、権限、ID、接続、監視、停止を AI とは別のシステムと組織の管理下に置く方向が共通して見えてきます。本書では、この共通点を美容業界の AI 運用へ置き換えて考えます。

デジラボビューティとして残していくもの

AI の技術や政策は短期間で更新されます。デジラボビューティでは、新しい公的資料や企業の一次情報が公表されたとき、過去の白書を消して書き換えるのではなく、情報基準日と版番号、更新履歴を残しながら積み上げます。美容業界で AI を導入するときに、「その時点で何が分かっていたか」「何が新しく追加されたか」を後から確認できるアーカイブにするためです。

デジタル庁の 9 月 29 日資料には、10 月上旬に「源内基準」を策定・公表する予定が記載されています。[G7] 本版は同日の公表資料を根拠とし、検討案と確定した指針を区別しています。2026 年 10 月 7 日の確認範囲では正式版の公表を確認できていないため、正式版が確認できた時点で差分を整理します。

4 NVIDIA の AI エージェント安全設計

2026 年 9 月 28 日、NVIDIA は「Open Agent Safety Platform」を公表しました。重要なのは、AI エージェントに「安全に動いてください」と指示するだけではなく、エージェントの外側に、実行制御・権限確認・監視・停止の仕組みを置くという設計思想です。[N1][N2]

NVIDIA は、長時間動作する AI エージェントが、曖昧な指示、利用可能なツール、エラーや制約への対応などをきっかけに、意図した仕事や運用条件から外れる「drift（逸脱）」が起こり得ると説明しています。そのため、エージェント自身に自分の安全管理を全面的に任せることはできない、という立場を示しています。[N1]

BlueField-4 と Sentry は「AI モデル」ではない

ここは名称を正確に区別しておきます。BlueField-4 は NVIDIA の DPU（Data Processing Unit）であり、AI モデルではありません。NVIDIA Sentry も生成 AI モデルではなく、BlueField-4 上で AI エージェントの挙動を外側から監視・制御するための「out-of-band watchdog（独立した見張り役）」として説明されています。[N1][N2]

したがって本書では、「BlueField-4 モデル」「Sentry モデル」ではなく、BlueField-4 DPU と NVIDIA Sentry という正式な役割で整理します。

実行環境の制限を担う OpenShell

OpenShell は、AI エージェントをサンドボックス化された実行環境で動かし、ファイル、ネットワーク、ツール、プロセス、認証情報など、何にアクセスできるかを運用者が定める仕組みです。制限は実行前だけでなく、エージェントが動いている間も適用されます。[N1][N3] 制御は AI エージェントのプロセスの外側で行われます。[N3]

美容の現場に置き換えると、「予約台帳は読めるが削除はできない」「SNS 投稿案は作れるが公開はできない」「顧客名簿へはアクセスしない」といった、AI が触れてよい範囲を先に決める考え方です。

独立した基盤で権限を確かめる BlueField-4 と DOCA

NVIDIA の説明では、BlueField-4 DPU をモデルへの経路上に置き、ホスト側の実行環境から独立した位置で、エージェントの通信や権限を確認します。DOCA ゲートウェイは、エージェントの ID と委任された権限を継続的に確認し、データ、ツール、API、サービスへのアクセスをゼロトラストの考え方で制御します。[N1][N2]

ここで重要なのは、AI が「自分にはこの権限があります」と主張するだけで通さないことです。AI の外側にある仕組みが、誰の権限で、何をしようとしているのかを確認する構造になっています。

継続監視を担う NVIDIA Sentry

NVIDIA Sentry は、BlueField-4 上でエージェントの挙動を継続監視し、ソフトウェア境界を越えようとする動きに対して隔離・停止を行うための独立した安全層として公表されています。NVIDIA は、Sentry がエージェントや攻撃者から見えにくい、ホストとは分離された信頼領域で動作すると説明しています。[N2]

NVIDIA の技術ブログでは、Sentry は OpenShell と組み合わせる「任意の追加安全層」と位置付けられています。したがって、すべての美容事業者が BlueField-4 を導入すべき、という意味ではありません。本書が注目するのは、規模にかかわらず応用できる「実行環境で制限し、独立した経路でも確かめる」という設計原則です。[N1]

当法人の分析：安全設計の美容業務への適用例

NVIDIA の構成	役割	美容業界での実装例
OpenShell	AI プロセスの外側で、ファイル・ネットワーク・ツール等のアクセス範囲を制限	AI に与える操作権限を最小化し、顧客 DB・予約・SNS・決済ごとに「読む／書く／公開する」を分ける
BlueField-4 + DOCA	エージェントの外側で ID・委任権限・通信経路を確認	API ゲートウェイ、サービスアカウント、期限付きトークン等で、AI とは別の仕組みが権限を確認する
NVIDIA Sentry	独立した監視・検知・隔離／停止	AI が変更できない監査ログ、異常通知、強制停止スイッチ、管理者へのエスカレーションを用意する
人間の運用責任	方針を決め、権限を付与し、必要時に止める	経営・業務・技術の担当を決め、重大操作は人が承認し、事故時の停止・復旧手順を持つ

デジラボビューティとしての結論

AI の安全性を高めるうえで、「AI が正しい判断をすること」だけに期待する設計は弱くなります。人が目的と許可範囲を決め、システムが権限を強制し、別の監視経路が実行を確かめ、必要なら止められる。この分離が重要です。

美容業界では、すべてのサロンがデータセンター級のハードウェアを導入する必要はありません。しかし、顧客情報、顔・肌画像、予約、SNS、決済を AI エージェントへ接続する場合は、「AI に何を頼むか」だけでなく、「AI が物理的・技術的に何までできる状態か」「AI 自身とは別に誰が確認し止められるか」を設計する必要があります。これは、バイブコーディング時代に人間が担う仕事そのものです。

注：本章は NVIDIA が公表した製品・リファレンス設計の解説と、デジラボビューティによる美容業界向けの実務的な翻訳です。NVIDIA の設計が法令・公的基準として美容事業者に義務付けられているわけではありません。

5 人間が担う 7 つの仕事

7 つの役割名は教材[S1]を参照し、下表の判断内容と成果物は当法人が美容業務向けに作成しました。AI は各段階の下書きや確認作業を支援します。人は、判断が必要な点を見極め、採用の根拠を残します。

人間の仕事	美容現場で決めること	残す成果物
目的設定	誰の、どの困りごとを解決するか。既存サービスで足りるか。	対象者と課題を記した企画メモ
要件判断	必須機能、扱うデータ、実装しない機能、完成の条件を選ぶ。	機能一覧と合格条件
設計選択	保存場所、閲覧権限、外部接続先、費用と保守の方法を選ぶ。	データの流れと権限表
差分確認	依頼した範囲の変更か。既存の料金・予約・権限に影響しないか。	変更一覧と実際の変更記録
テスト	通常の操作と失敗時の操作で、期待した結果になるかを確かめる。	入力例・期待結果・実測結果
公開判断	対象利用者、残る不具合、停止方法を確認し、公開範囲を決める。	公開可否と判断者の記録
運用責任	問い合わせに答え、障害を止め、更新時に再確認する体制を持つ。	運用手順と担当者・連絡先

判断を具体的な条件にする

例えば「予約しやすいアプリにして」だけでは、完成したかどうかを確かめにくくなります。「スタッフが確定するまで仮予約と表示する」「満席時に申込みができない」「送信に失敗した場合は未完了と表示する」のように、確認できる条件へ分解します。

一人の経営者が複数の役割を兼ねても構いません。ただし、確認できない技術領域まで一人で引き受けず、権限管理や決済などの重要部分は、担当できる専門家と分担します。

6 美容業務のアプリで人が決めること

以下は説明用の仮想例です。デジラボビューティや特定サロンの既存システムについて、実装・検証が済んでいることを示すものではありません。

予約の受付を助けるアプリ

人が決めるのは、仮予約と確定予約の違い、キャンセルの扱い、同じ枠への同時申込み、担当者が不在のときの対応です。画面を作る前に業務のルールを決めれば、予約が重なる条件もテストにできます。外部の予約サービスと連携する場合は、どちらの情報を正しい予約台帳とするかを決めます。

セミナーとアーカイブの会員アプリ

無料会員と有料会員で何を見られるか、契約終了後はいつ閲覧権限が変わるかを人が定めます。会員種別を画面に表示できても、動画の URL を直接開けば誰でも見られる状態では、権限管理ができたとはいえません。決済が失敗した場合や、通知が遅れた場合まで確認します。

カウンセリングの準備を助けるアプリ

本人の希望を整理し、担当者に伝えることを目的にするなら、AI が回答する範囲と、人へ引き継ぐ条件を定めます。身体の症状や疾病の判断に関わる機能は、通常の美容情報とは別に専門的な検討が必要です。画面に免責文を置くだけで、用途や機能の問題を解決できるとは考えません。

作り始める前に選択肢を比べる

選択肢	選びやすい条件	判断上の注意
既存サービスを使う	標準的な予約・決済・会員管理で足りる	継続費用、連携、データ移行条件
小さな補助機能を作る	既存業務の一部分だけ改善したい	既存システムへの影響を限定する
独自アプリを運用する	既存製品で満たせない価値がある	検証・保守・障害対応の担当と費用

最初の一歩としては、顧客情報や決済を扱わない内部用の試作品が取り組みやすい選択です。便利さを測り、運用できる見通しが立ってから、公開範囲を広げることを提案します。

7 動いたことと使えることを分けて確かめる

AI の「修正しました」という説明は、確認を始めるための情報です。実際の変更記録、操作結果、ログと突き合わせます。同じ AI に確認を依頼する場合も、合格条件を先に決めておき、生成した説明だけで合否を判断しません。以下は当法人による検証例です。

観点	確かめ方	確認する人
業務上の正しさ	料金・期限・定員・権限の条件を、決めた仕様と照合する。	業務責任者
失敗時の動作	未入力、二重送信、通信切断、戻る操作で結果を確かめる。	現場担当者と開発担当者
アクセス制御	他の会員のデータや管理機能へアクセスできないか、サーバー側を含めて検証する。	技術担当者
情報の送信と保存	端末、サーバー、AI 提供者、ログ、バックアップへのデータの流れを確認する。	技術担当者と情報管理担当者
既存機能への影響	変更前にできた予約・ログイン等が、変更後も機能するか確認する。	開発担当者
使いやすさ	スマートフォン、文字サイズ、キーボード操作、説明文の分かりやすさを確認する。	実際の利用者と現場担当者

写真を保存しないという説明の確かめ方

アプリの画面から写真が消えることと、関係するすべての場所に写真が残らないことは別の確認事項です。保存しない設計を採る場合は、通信先、処理中の一時保存、エラー記録、外部サービス側の保持条件まで調べ、実際の扱いに合う説明を表示します。[G4 を踏まえた提言]

公開の判断に必要な証拠

テストごとに、確認した版、操作条件、期待した結果、実際の結果、確認者を残します。第三者の情報が見える、誤って課金される、重要データが消えるといった重大な問題が残る場合は、その機能の公開を止めます。対象範囲を絞るか、修正後に同じ条件で再確認します。

[G3] 6.3.3 はリリース前の安定稼働・期待品質の確認を扱います。本ページの具体的テスト項目と担当分担は、当法人の実務提案です。

8 AI へ任せる範囲と人の権限を決める

人間中心の運用は、すべての操作を人が毎回承認することだけを意味しません。影響の小さい反復作業には、条件を定めた自動実行が有効です。影響の大きさ、元に戻せるか、誤りに気づけるかを基準に分担します。[G1][G2]を踏まえた当法人の提案です。 NVIDIA の最新設計も、エージェント自身の判断だけに安全管理を依存せず、実行環境と独立した監視層を組み合わせる考え方を示しています。[N1][N2]

任せ方	想定する作業	人が整える条件
事前に決めた範囲で自動化	内部用の文章案、テスト環境の変更、検証済みの定型処理	入力の範囲、利用権限、実行記録、異常時の停止
実行前に人が確認	顧客への案内、公開ページの料金変更、会員権限の変更	変更内容と対象者、確認者、実行後の確認方法
専門担当者の確認を組み込む	決済、認証、個人情報の移行、身体への影響に関わる機能	技術・法務等の確認範囲、責任分担、公開条件

判断できる情報と止める権限を渡す

担当者には、AI が何を変えるのか、誰に影響するのか、どこが未確認なのかを知らせます。確認する時間と、却下・保留・停止する権限も必要です。形式的に承認ボタンを押すだけの手続きでは、問題の発見にはつながりにくくなります。

組織として引き受ける仕事

業務責任者は目的と公開可否、技術担当者は実装と技術検証、運用担当者は監視と問い合わせを担当する形が考えられます。小規模組織では兼務し、難しい領域は外部支援を利用します。契約・法令に基づく責任の配分は個別事情で異なるため、「利用者一人がすべての責任を負う」と一律に整理しません。

公開判断の記録に残す項目

対象の版と公開先／変更目的／影響する利用者とデータ／テスト結果／残る課題／公開の条件／停止と復旧の担当者／判断者と日時。重要な機能を変えたときは、この記録を更新して再判断します。

AI でコードを作ることと、完成したアプリが運用中にも AI を使うことは別です。後者では、コードの変更に加え、AI モデルや参照データの更新による出力の変化も確認します。

9 現場で始める 30 日間の進め方

以下は導入計画の提案であり、実施済みの事業や効果を示すものではありません。新しい職種を増やす前に、現在の担当者が担う判断を明らかにし、小さな業務から学びます。

時期	実施内容	主担当	確認する成果物
1 週目	一つの業務を選び、困りごと、対象者、現在の所要時間を整理する。	業務責任者	企画メモと比較前の記録
2 週目	架空データで試作する。機能と完成条件を必要最小限に絞る。	開発担当者	試作品と合格条件
3 週目	現場担当者が試し、重要部分は専門担当者が検証する。	検証担当者	テスト結果と修正一覧
4 週目	利用範囲を限定して運用し、結果を基に継続・改修・中止を決める。	運用担当者と責任者	効果・負担・不具合の記録

研修では何を評価するか

「アプリが作れた」で研修を終えず、なぜ必要か、何を実装しなかったか、どのテストに失敗したか、いつ利用を止めるかを説明してもらいます。AI の提案に誤りを見つけ、根拠を示して修正できることも評価します。経産省が示す人材・スキルの方向性を、美容現場の課題に置き換える提案です。[G5][G6]

効果は確認と保守の時間を含めて測る

指標	測り方
業務時間	同じ業務量について、入力・確認・修正を含む時間を導入前後で比べる。
品質	入力漏れ、誤処理、やり直し、問い合わせを件数と内容で記録する。
継続費用	利用料、AI 利用料、保守時間、外部支援費用を月単位で集計する。
利用者の負担	迷った操作、途中離脱、人の支援が必要だった場面を確認する。

削減時間や売上増加の数値は、実測するまで約束しません。確認負担が大きく、既存サービスの方が目的を満たす場合は、独自開発を続けないという判断も適切です。

10 公開前に確認する 18 項目

デジラボビューティによる実務用チェックリストです。公的な適合認証ではありません。NVIDIA の「AI の外側で権限・監視・停止を持つ」という設計思想に加え、デジタル庁が示した AI エージェントの ID 管理、接続先の把握、中間動作の確認、停止・廃止までのライフサイクル管理の考え方を反映し、18 項目に更新しました。[G7][G8][N1][N2] 各項目に、確認者・確認日・記録の場所を添えて使います。該当しない項目には理由を残し、重要な未確認事項がある機能は公開範囲から外します。

□ 誰の何の課題を解決するか、目的を説明できる。

□ 実装する機能と実装しない機能、完成条件が決まっている。

□ 料金・期間・定員などの数値と表示内容を照合した。

□ 入力するデータの利用目的、送信先、保存・削除の条件を確認した。

□ 閲覧・編集・管理の権限を、画面とサーバー側の両方で確認した。

□ 通常の操作に加え、未入力・二重送信・通信失敗時を確認した。

□ 今回の変更が、既存のログイン・予約等を壊していないことを確かめた。

□ 画像・文章・コード等の利用条件と、表示の根拠を確認した。

□ 利用者が AI の役割と限界を理解でき、人に相談する方法がある。

□ 本番の公開範囲、公開する版、責任者の判断を記録した。

□ 問い合わせ、障害の停止、データ保全、復旧の担当と手順がある。

□ 更新時の再確認、費用確認、不要になった機能の終了方法が決まっている。

□ AI エージェントに与える権限を、業務に必要な最小範囲へ絞り、「読む／書く／削除する／公開する」を分けている。

□ AI とは独立した場所に監査ログや異常通知があり、AI 自身が記録を消したり監視を無効化したりできない設計を確認した。

□ 重大な誤動作が起きた場合に、AI の外部接続・API 権限・自動実行を人が速やかに停止できる。

□ AI エージェントごとに、管理責任者、利用目的、使用ツール・MCP・API、公開状態を把握している。

□ AI エージェントの中間動作と最終結果を確認できるログまたは記録がある。

□ 脆弱性や異常が見つかった場合に、対象 AI エージェントを停止・改修・廃止できる手順がある。

最初に作る一枚

まず、アプリ名、目的、利用者、扱うデータ、必須機能、合格条件、責任者、停止方法を一枚に書き出してください。この一枚を AI への依頼と、担当者同士の確認に使います。条件が変わったときは、変更理由も残します。

デジラボビューティの提言

当法人は、美容現場の知見を開発条件へ具体化し、検証記録に基づいて公開・運用を判断する体制を提案します。人が目的と許可範囲を定め、技術的な制御と継続的な確認を組み合わせることが、実務で AI を利用するための条件となります。

参考資料と添付原本 1

本文中の資料番号は以下に対応します。各資料の当初確認日は下記と更新履歴に示し、2026 年 10 月 7 日に主要な公的資料と NVIDIA の公式資料を再照合しました。関連省庁の原資料 4 点は白書掲載ページから取得できます。デジタル庁の追加資料および IPA 資料等は公式リンクを案内します。原資料と当法人の解釈を区別してお読みください。

[G1] 人工知能関連技術の研究開発及び活用の適正性確保に関する指針

人工知能戦略本部決定／内閣府掲載 ／ 2025 年 12 月 19 日

確認箇所：本文 p.2〜5（PDF 3〜6 ページ）、特に本文 p.3 の人間中心・アカウンタビリティ。AI 法第 13 条に基づく、自主的・能動的な取組を促す指針。

公式資料を開く

掲載ページの参考原資料：01_Cabinet_AI_Guidance_2025.pdf

[G2] AI 事業者ガイドライン 第 1.2 版 概要

総務省・経済産業省 ／ 2026 年 3 月 31 日

確認箇所：PDF 3、7〜8、15〜16、18 ページ。主体区分、策定の考え方、共通の指針、ガバナンス。添付は両省が作成した公式の概要資料であり、本編全文ではない。

公式資料を開く

掲載ページの参考原資料：02_MIC_METI_AI_Guidelines_v1.2_Summary.pdf

[G3] 行政の進化と革新のための生成 AI の調達・利活用に係るガイドライン 第 2.0 版

デジタル社会推進会議幹事会決定／デジタル庁掲載 ／ 2026 年 6 月 12 日

確認箇所：2.2、6.3.3、6.5、別紙 3・4。公開前検証、公開後の品質確認、調達・契約時の確認事項。政府情報システム等向けであり、本書では民間の実務設計の参考として扱う。

公式資料を開く

掲載ページの参考原資料：03_Digital_Agency_GenAI_Guidelines_v2.0.pdf

政府資料の掲載は、各府省庁による本書や当法人の推奨・承認・監修を意味しません。

参考資料と添付原本 2

[G4] 生成 AI サービスの利用に関する注意喚起等 別添 1

個人情報保護委員会 ／ 2023 年 6 月 2 日

確認箇所：PDF 2〜3 ページの（1）（3）。個人情報取扱事業者の入力時の注意点、一般利用者の留意点。利用目的、学習利用等の取扱い、規約の確認に参照。

公式資料を開く

掲載ページの参考原資料：04_PPC_Generative_AI_Notice_2023.pdf

[G5] デジタルスキル標準 ver.2.0

経済産業省・独立行政法人情報処理推進機構 IPA ／ 2026 年 4 月／公表 2026 年 4 月 16 日

参照箇所：改訂の考え方、DX リテラシー標準、DX 推進スキル標準。IPA は独立行政法人。本公開版では原本を再配布せず、公式 URL を案内します。

公式資料を開く

公式リンク参照（公開版への原本添付なし）

[G6] 生成 AI 時代の DX 推進に必要な人材・スキルの考え方 2024 公表ページ

経済産業省 ／ 2024 年 6 月 28 日

参照箇所：公表ページの専門人材に関する記載。課題設定・仮説検証・評価選択に関する説明を参照。2024 年の資料であり、最新のスキル標準は[G5]を参照。

公式資料を開く

公式リンク参照

[G7] 第 6 回先進的 AI 利活用アドバイザリーボード／第 3 回ワーキング・グループ実施報告

デジタル庁 ／ 2026 年 9 月 29 日

参照箇所：AI エージェントチャット、源内工房、MCP/API 連携制限、マルウェア対策、Human-in-the-Loop／Human-on-the-Loop、中間動作と結果の双方を確認するプロセスアプローチ、源内基準の策定予定。

公式資料：https://www.digital.go.jp/assets/contents/node/basic_page/field_ref_resources/162953b0-3dc0-46a6-a24c-0e3a335e0495/7545ff68/20260929_meeting_ai-advisory_outline03.pdf

[G8] 行政の進化と革新のための生成 AI の調達・利活用に係るガイドライン充実に向けた検討事項案

デジタル庁 ／ 2026 年 9 月 29 日

参照箇所：AI エージェント固有のガバナンス、AI エージェントごとの一意 ID、使用ツール・MCP サーバー等の一覧化、登録・承認・配布・公開・停止・廃止、脆弱性判明時の迅速な特定・停止・改修。2026 年 10 月 7 日時点では検討事項案。

公式資料：https://www.digital.go.jp/assets/contents/node/basic_page/field_ref_resources/162953b0-3dc0-46a6-a24c-0e3a335e0495/e8ccd70c/20260929_meeting_ai-advisory_outline06.pdf

[G9] ガバメント AI「源内」

デジタル庁 ／ 最終更新 2026 年 9 月 30 日

参照箇所：政府統一基準に準拠した生成 AI 利用環境、機密性 2 情報の取扱い、全府省庁約 18 万人を対象とする大規模実証、AI アプリ・共通データセット・運用基盤の展開。

公式資料：https://www.digital.go.jp/policies/genai

[S1] バイブコーディングで人間が担う仕事 A1 A-04

AI と作る Web アプリの教科書／AI Builder ／ 公表日表示なし／2026 年 9 月 29 日閲覧

民間の学習教材。本文を参照し、7 つの役割の整理を本書の出発点とした。原文の複製は添付していない。

公式資料を開く

公式リンク参照

[N1] NVIDIA Open Agent Safety Platform: A Reference for Continuous In-Silicon Agent Monitoring

NVIDIA Developer Technical Blog ／ 2026 年 9 月 28 日

参照箇所：OpenShell のゼロトラスト実行環境、5 つの原則、BlueField-4 上の Sentry、DOCA による ID・委任権限確認、Vera Rubin POD におけるモデル経路上の制御。

公式資料：https://developer.nvidia.com/blog/nvidia-open-agent-safety-platform-a-reference-for-continuous-in-silicon-agent-monitoring

[N2] NVIDIA Launches Open Agent Safety Platform to Secure Agents From Testing to Deployment

NVIDIA Newsroom ／ 2026 年 9 月 28 日

参照箇所：OpenShell、NVIDIA Sentry、BlueField-4 DPU、in-silicon enforcement、DOCA によるテレメトリ・ID 確認・ゼロトラストアクセス制御の説明。

公式資料：https://nvidianews.nvidia.com/news/open-agent-safety-platform

[N3] NVIDIA Open Agent Safety Platform / OpenShell

NVIDIA 公式製品・開発者情報 ／ 2026 年 10 月 4 日確認

参照箇所：OpenShell がエージェント実行環境の外側でポリシーを強制すること、OpenShell・Sentry・BlueField-4 の役割分担。

公式資料：https://www.nvidia.com/en-us/solutions/ai/agent-safety/

更新履歴

Version 1.0（2026 年 9 月 29 日）：初版。バイブコーディング時代の人間の役割、検証、運用責任を整理。

Version 1.1（2026 年 10 月 4 日）：NVIDIA Open Agent Safety Platform を追記。OpenShell、BlueField-4 DPU、DOCA、NVIDIA Sentry による「内側の制限＋外側の監視」という考え方を美容 AI へ翻訳し、公開前チェックを 15 項目へ更新。

Version 1.2 Draft（2026 年 10 月 4 日）：デジタル庁の 2026 年 9 月 29 日公表資料を追記。ガバメント AI「源内」における AI エージェントの MCP/API 連携制御、Human-in-the-Loop／Human-on-the-Loop、中間動作の確認、一意 ID、登録・承認・公開・停止・廃止までのライフサイクル管理を整理し、公開前チェックを 18 項目へ更新。源内基準正式版の公表後に差分を再確認する。

Version 1.3（2026 年 10 月 7 日）：公開版。主要出典を再照合し、OpenShell の制御位置を明確化。一般公開用の参考原資料と公式リンクを整理し、HP 本文と PDF を公開。検討案の位置付けは維持。

Version 1.4（2026 年 10 月 7 日）：No.1・No.2 との構成を統一。参照資料と 5 つの要点を冒頭に置き、公的資料・企業技術資料の解説、当法人の分析、実務提案の順へ再編。個人的な語りかけを整理。参照資料と情報基準日は維持。

本書の性格：公的資料、民間教材、企業の公式技術資料に基づく実務解説です。独自の実証調査や雇用予測ではありません。個別の法的適否や医療該当性を判定するものではなく、機能・データ・契約に応じた専門的確認が必要です。デジタル庁[G8]の内容には 2026 年 10 月 7 日時点で検討事項案が含まれ、一般企業への直接的な義務を示すものではありません。NVIDIA 資料は同社の製品・リファレンス設計に関する一次情報であり、政府基準、公的認証、またはデジラボビューティに対する推奨・承認を意味しません。デジラボビューティでは、今後も公的資料と一次情報の更新を確認し、情報基準日と更新履歴を残して継続改訂します。