

Microsoft 認定 Agentic AI ビジネス ソリュー ション アーキテクト エキスパート(AB-100) 学 習用テキスト

予備知識ゼロから合格ラインへ —— 読んで理解する1冊目

無料サンプル(第1章まで)

AB-100 出題範囲(Study guide・Skills measured as of November 16, 2026)準拠
2026年10月版

ご利用にあたって

- このテキストは「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。
- 無料サンプルも著作物です。転載・改変しての再配布・二次販売はできません。
- AI モデルの学習データとしてのご利用(学習目的の入力・データセット化を含む)はできません。
- 製品版(学習用テキスト・頻出度別ノート・暗記用ノートの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。

SAMPLE

目次

この試験の範囲 [overview]

第1章 Plan AI-powered business solutions

[1.1.1/1.1.2/1.1.3/1.2.1/1.2.2/1.2.3/1.2.4/1.2.5/1.2.6/1.2.7/1.2.8/1.2.9/1.2.10/1.2.11/1.2.12/1.2.13/1.3.1/1.3

タスクの自動化・データ分析・意思決定でのエージェントの評価 [1.1.1]

接地に使うデータの確認 [1.1.2]

他の AI システムから使えるデータの整理 [1.1.3]

Azure 用クラウド導入フレームワークの AI 導入プロセス [1.2.1]

AI とエージェントを作る戦略 [1.2.2]

マルチエージェント ソリューションの設計 [1.2.3]

事前構築エージェントのユース ケース [1.2.4]

Copilot Studio・Foundry・Foundry Tools の決まりと制約 [1.2.5]

Copilot Studio での生成 AI とナレッジ ソースの使い方 [1.2.6]

Microsoft 365 Copilot の拡張とカスタム エージェントの選び方 [1.2.7]

カスタム AI モデルを作る判断 [1.2.8]

プロンプト ライブラリの作り方の指針 [1.2.9]

小さな言語モデル (SLM) のユース ケース [1.2.10]

プロンプト エンジニアリングの指針と技法 [1.2.11]

Microsoft AI Center of Excellence の要素 [1.2.12]

複数の Dynamics 365 アプリを使う AI ソリューション [1.2.13]

ROI の基準と総保有コスト [1.3.1]

業務プロセスの ROI 分析 [1.3.2]

構築・購入・拡張の判断 [1.3.3]

モデル ルーター [1.3.4]

サンプルはここまでです

このテキストの位置づけ: 3 点セットの本編。これだけで問題が解けるようになることが目標で、判断の型を合格ラインの深さで扱う。

進め方: ① 学習用を読む → ② 頻出度別で解き方を掴む → ③ 問題演習で腕試し → ④ 頻出度別に戻って再確認(試験直前もここ)。暗記用は空き時間や用語の再確認に使う。

SAMPLE

この試験の範囲 [overview]

AB-100(Microsoft Certified: Agentic AI Business Solutions Architect Expert の必須試験)は、AI を使った業務ソリューションを計画し、設計し、展開する立場の判断を問う Expert の試験です。合格には 1000 点満点の 700 点以上が必要です。日本語で受験できます。

出題範囲は 3 つの分野に分かれます。この教材も、同じ順に章を立てます(下の表)。配点は、学習ガイドの「Skills at a glance」にある割合です。

分野	配点	細目の数
1 Plan AI-powered business solutions	35～40%	20
2 Design AI-powered business solutions	30～35%	30
3 Deploy AI-powered business solutions	20～25%	24

学習ガイドの細目は 74 個。節の見出しの後ろにある [1.1.1] のような印が、その細目の番号。準拠する版は、2026 年 11 月 16 日時点の「Skills measured」。

第1章 Plan AI-powered business solutions

[1.1.1/1.1.2/1.1.3/1.2.1/1.2.2/1.2.3/1.2.4/1.2.5/1.2.6/1.2.7/1.2.8/1.2.9/1.10]

AI を使った業務ソリューションを、作る前に決める章です。何を任せるか、どのデータを使うか、どの製品で作るか、いくらかかって何が返るかを順に扱います。この章は、細目 1.1.1 から 1.3.4 までの20を扱う、試験の分野1です (配点 35〜40%)。[^c1-1] 製品の置き場は、先にこの表で頭に入れてください。

置き場	型	作り方	向く用途
Microsoft 365 Copilot の宣言型エージェント	拡張	指示・知識・アクションを設定する。モデルとオーケストレーターは Copilot のもの	個人の作業、Microsoft 365 アプリの中
カスタム エンジン エージェント	プロコード	独自のモデルとオーケストレーションを持ち込む。ホスティングは別	完全に制御したい複雑な処理
Copilot Studio	SaaS・ローコード	画面で作る。マネージドの基盤	業務プロセスの変革、中程度のカスタマイズ
Microsoft Foundry	PaaS	ローコードでもプロコードでも作る。モデルを選ぶ	戦略的な変革、深い統合
GPU とコンテナ	IaaS	モデルとランタイムを自分で持つ	高度なカスタム、特殊なコンプライアンス
製品内エージェント (Dynamics 365 など)	SaaS	製品に組み込み済み	その製品の業務

タスクの自動化・データ分析・意思決定でのエージェントの評価 [1.1.1]

エージェントは、目標に向けて手順とツールをモデルが自分で選びながら進める仕組みです。決まった台本を順に実行するプログラムとは違い、各ステップで次の一手を決めます。構成要素は次の5つです。

- 生成 AI モデル (推論エンジン)
- 指示
- 検索 (接地データ)
- アクション (関数 / API)
- メモリ (会話履歴と状態)

RAG (検索拡張生成) は、決まった検索手順で答えます。エージェントは各ステップで使う知識とツールをモデルが選ぶので、多段の推論ができます。その代わり動きは非決定的になります。[^c1-2]

使わない場面。 手順が決まっていて繰り返せる作業は、通常のコードか非生成 AI のほうが速く、安く、確実です。固定文書への質問応答や要約だけなら、エージェントでなく古典的な RAG で足ります (FAQ ボットなど)。[^c1-3]

効く場面。 次の3条件のどれかに当たるとエージェントが効きます。例はチケットの振り分け、経費処理、顧客対応です。

- 複数ステップの判断が要る

- 多数のツールやシステムを柔軟な順で呼ぶ
- 入力が不完全で、意図の解釈が要る

生成 AI は同じ入力でも出力が変わるため、非構造の入力や手順が固定でない場面に向きます。非生成 (決定的) の AI は、構造化された入力で結果が毎回同じであるべき予測や異常検知に向きます。[^c1-4]

ユース ケースは2つに分けて考えます。1つは個人の作業改善で、文章作成や会議準備が当たり、Microsoft 365 Copilot などと相性がよい領域です。もう1つは業務の自動化で、顧客のルーティングや需要予測が当たり、他システムとの連携が要ることが多く、複数の AI 種別を組み合わせることもあります。

エージェントが生む価値は4領域に整理されます。

向き	価値領域
内部	業務プロセスの再構築、従業員体験の向上
顧客向け	顧客エンゲージメントの再発明、イノベーションの加速

優先順位の採点。ユース ケースごとに、ビジネス効果、技術の実現性、ユーザーの望ましさを 1〜5 点で評価します。効果は戦略との整合、業務価値、変更管理にかかる期間で見ます。実現性はリスク、安全策、技術の適合で見ます。[^c1-5] この教材では、各案を3項目で採点し、案ごとに合計を出し、合計の高い順に並べて、最も高い案から着手する、の順で決めます。たとえば3案が (5,3,4)、(4,5,4)、(3,2,5) なら合計は 12、13、10 で、2番目の案から着手する並びになります (重み付けなど、合計以外の出し方もあります)。

成功指標の置き方。手順は次の順です。

1. 業務の現状をベースラインとして記録する
2. エージェントが改善する KPI を少数に絞る
3. パイロットの結果を go/no-go のゲートに使う
4. 本番後も計画的に実績を確認し、改善・廃止・他の機会への転用を決める

分析と意思決定を支える事前構築の例。

エージェント	役割	注意
Researcher	Web と仕事のコンテンツ (ファイル・メール・チャット) をまたぐ多段の調査をし、出典付きレポートを返す	処理は長め。軽い要約や短い返信は標準の Copilot チャット向き[^c1-6]
Dynamics 365 Sales の Opportunity Agent	商談を自律的に調べ、リスクと有望案件を示す	意思決定の支援
Qualification Agent	リードを調査して適格かを判定し、接触メールも送る	外部への送信まで自律で動く[^c1-7]

Copilot Studio の新規エージェントは、既定が生成オーケストレーションです。説明文をもとにトピック・ツール・他エージェント・ナレッジをモデルが選び、足りない入力は自動で質問します。従来型はトリガー フレーズの一致で動きます。[^c1-8] Foundry のエージェントは種別が3つです。[^c1-9]

- Prompt agent: 設定だけで作る。コード不要

- Voice-based prompt agent: 音声で応答する
- Hosted agent: 自前のコードをコンテナで実行する

コード無しで管理を最小にしたいなら Prompt agent です。

自動化の度合いと人の関与。度合いは、次の3段に分けて考えると決めやすくなります(公式が定めた区分ではなく、この教材の目安です)。

度合い	動き	当てる業務の目安
提案だけ	結果を示し、実行は人が行う	調査レポート、リスクの指摘。取り消せない操作を含む業務の初期段階
人の承認つき	下書きや候補を出し、承認で実行する	顧客への送信、金額の動く処理
自律で実行	判断から実行まで任せる	影響が小さく取り消せる、件数の多い定型

決め方の軸は、操作を取り消せるか、組織の外へ出るか、影響の大きさです。取り消せない操作は、最初から自律には任せません。迷う時は提案だけから始め、パイロットの実績を go／no-go に使って段階的に広げます。Customer Service の Rollout manager は、自律エージェントを段階的に管理しながら導入するための仕組みです。

接地に使うデータの確認 [1.1.2]

エージェントは情報を作らず、元のデータを統合して答えます。精度は元ソースの品質とアクセス性にそのまま左右され、分断された不完全なデータでは信頼できる答えになりません。[^c1-10] AI の導入計画では、まずデータ資産を棚卸しします。データソース・形式・品質・アクセス性を文書にして、準備作業の優先順位を決めます。[^c1-11] 下の表の5つの観点は学習ガイドの細目の語 (accuracy・relevance・timeliness・cleanliness・availability) で、「足りない時の出方」と「直し方」はこの教材の整理で、公式の文言ではありません。

観点	足りない時の出方	直し方
正確さ	誤った数字や古い規程をそのまま答える	元の文書を正しい版に直す
関連性	無関係な文書が混ざり、答えがぼやける	ナレッジに入れる対象を目的の文書に絞る
鮮度	改訂前の古い文書で答える	更新と廃止の運用を決め、古い版を外す
きれいさ	重複した文書が食い違う答えを返す	重複を統合し、正を1つに決める
使える状態	見せてはいけない文書が答えに出る、または見せるべき文書が出ない(権限の過不足)	元のデータ側の権限を直す

Microsoft 365 のデータは Microsoft Graph 経由で取得され、既存の権限が守られます。対象は SharePoint・OneDrive・Exchange・Teams です。そのため、答えに出る文書の過不足は AI 側の設定でなく、元の権限で直します。共同作業、参照、組織知識のコンテンツは Microsoft 365 に置く判断になります。[^c1-10]

ゴールデンデータセット。モデル評価の基準になる信頼できる標準データです。一貫性と品質のポリシーを定め、バージョンや環境をまたぐ性能の劣化を検出するのに使います。[^c1-12] 機密データと公開データを分けて扱うポリシーも定めます。分類(機密度)が変わったら、AI 側も追従させます。[^c1-13]

宣言型エージェントの接地の限界。 接地するレコードは 50 件まで、プラグインの応答は 25 件までです。接地と外部ツールの呼び出しは逐次で、連鎖や反復ループは作れません。Web 検索をサイトの URL で絞ると、元のサイトを直接検索せず Bing のインデックスを使います。動的に生成されるコンテンツは、不完全か古い結果になることがあります。[^c1-14]

Copilot Studio のナレッジでの確認点。

- 認証: SharePoint・Dataverse・コネクタは Microsoft Entra ID の認証で、質問したユーザーが見られるコンテンツだけを返します。公開 Web とアップロード文書は認証なしです。[^c1-15]
- 根拠のない回答の許可 (Allow ungrounded responses): オフにすると、ナレッジやツールを使わなかったターンの回答をブロックします。生成オーケストレーションが前提です。オフでも引用の無い正しい回答が隠れることがあるので、指示で出典の引用を求めます。手順は、生成オーケストレーションの設定を確かめてからオフにし、オフにした後で隠れる回答があれば、指示で出典の引用を求めます。[^c1-16]
- コンテンツ モデレーション: エージェント、生成回答ノード、プロンプト ツールの3段で設定でき、実行時はトピック (ノード) の設定が優先されます。既定は High です。Lowest は回答が増える反面、有害な出力のリスクが上がります。[^c1-17]
- テナント グラフ接地: Microsoft 365 Copilot のライセンスを持つユーザーが使う場合は課金なしです。それ以外では 1 メッセージ 10 Copilot Credits の高品質な RAG になり、コネクタ経由で Graph に同期した外部データも対象です。[^c1-18]

他の AI システムから使えるデータの整理 [1.1.3]

AI のためにデータ戦略を一から作り直す必要はありません。必要なデータ機能は、採用する AI ソリューションとユースケースで決まります。各 AI が必要とする組織データを特定し、安全にアクセスでき、適切に統制できる状態にします。[^c1-19] 進め方は、必要なデータ機能を決める→各 AI が要する組織データを特定する→データ領域ごとに取得方式 (検索・API・その両方) を分類して記録する、の順です。

データの置き場。 統合の基盤は Microsoft Fabric の OneLake です。中心のデータレイクとして、ドメインごとに管理されたデータ プロダクトを作り、エージェントはそれを AI の主要な入力にします。Fabric は組織データを分析・AI 向けに準備して提供し、Azure の AI ワークロードは承認された統合パターンで使います。ワークロードのチームごとに別の統合を作らせません。[^c1-20] 表のデータ (表形式の業務データ) は Dataverse か Microsoft Fabric の OneLake、文書は SharePoint やアップロード、他システムのデータは Microsoft Graph コネクタで Microsoft Search の索引に取り込む、という分担です。文書は索引づけて検索で引き、Azure AI Search や OneLake 検索インデクサーがその役を担います。

検索方式の文書化。 データ領域ごとに、検索 (ナレッジ インデックス)、API、その両方のどれかを分類して記録します。たとえば製品 FAQ は検索、注文の追跡は API です。どのデータソースが認定済み (certified) か、どの API に認証が要るか、どの権限が適用されるかも、統制の情報として同じ記録に書きます。ナレッジにする単位を、権限と更新の担当がはっきりする粒度 (サイト・ライブラリ・テーブルなど) にそろえるのは、この教材が勧める運用の目安です。[^c1-21]

組み込みの検索オプション。 まず Foundry と Copilot Studio の組み込み機能から始めます。

- Foundry IQ

- Fabric IQ
- Fabric データ エージェント
- MCP サーバー
- SharePoint 接続
- Azure AI Search
- OneLake 検索インデクサー

エージェント型検索 (agentic retrieval) は、固定の1回検索でなく、探し方を計画して実行するループです。Foundry ベースのエージェントでは Foundry IQ を既定の取得方式にすることが勧められています。[^c1-22]

MCP を使う場面。在庫のような最新値の照会や、チケット作成のような操作が要る時です。リモート サーバーのツールを共通の方法で呼べます。ID・ポリシー・監査を伴うため、すべてのツール呼び出しに認証を求め、詳細に監査します。Foundry が対応する認証方式は次の4つです。[^c1-23]

- キー ベース
- Microsoft Entra Agent Identity
- プロジェクトのマネージド ID
- OAuth ID パススルー

他エージェントとの接続方式の使い分け。[^c1-24]

相手	使う方式
API や簡単な HTTP	カスタム コネクタ、HTTP ツール
MCP のツール・リソース	MCP サーバー
A2A に対応したエージェント	A2A
Microsoft 365 Agents SDK 製のエージェント	Activity Protocol

データ境界。Microsoft Purview で、機密情報の種類やトレーニング分類子により機密度を分類します。AI ワークロードが対象者に適切なデータだけに届くよう境界を定め、DLP で出力経由の漏えいも防ぎます。[^c1-25] 検索データは、遅延が重要なら消費する AI サービスの近くに置きます。ただし、データ所在地の要件を満たすことが先で、遅延を理由に外すことはできません。リージョン間のフェイルオーバーでは、モデルのデプロイと検索データも二次側に用意します。[^c1-26]

Copilot Studio が取り込めるデータは、Dataverse・SharePoint・Web・アップロード文書・コネクタ (Microsoft Search の索引) です。生成回答ノードのナレッジ ソースは Bing Custom Search・Azure OpenAI・Custom Data に対応していません。これらは、同じノードのプロパティの Classic データ オプションで使います。[^c1-15]

Azure 用クラウド導入フレームワークの AI 導入プロセス [1.2.1]

Cloud Adoption Framework (CAF) の AI 導入は、Strategy (戦略)、Plan (計画)、Ready (準備) の順に進み、そのあと Govern (統制)、Manage (管理)、Secure (保護) を継続的に回します。Govern・Manage・Secure は必要に応じて前の段へ戻ります。段の名前は、AI 戦略、AI の計画、AI の準備、AI の統制、AI の管理、AI の保護です。[^c1-27]

段	最初にすること
Strategy	リーダーが AI 憲章または戦略を示す。AI CoE などの中央のガバナンス機能が Responsible AI の要件を決める。承認済みの AI への従業員のアクセスを、研修とセットで広げる
Plan	ユース ケース特定、スキル評価、スキル獲得、リソースとライセンスの確保、優先順位づけ、PoC、責任ある AI の実装の順
Ready	既存のランディング ゾーンを拡張する
Govern	NIST AI RMF に沿った4手順で運用する
Manage	AI CoE、MLOps／GenAIOps、デプロイ権限、モデル運用、コスト、データ、事業継続
Secure	リスクの発見、AI 資産の台帳、マネージド ID、Purview での分類と DLP、AI 向けのインシデント対応

ソリューション選びの原則。要件を満たす中で、運用責任が最も小さいものを承認します。カスタマイズが大きいほど、必要な専門スキルと運用責任が増えます。検討の順は SaaS、PaaS (Azure)、インフラです。管理型プラットフォームで足りるなら、Azure のインフラ上の AI を既定にしません。独自モデルやランタイムの制御、特殊な性能やコンプライアンスの要件がある時だけインフラを選びます。[^c1-19]

Plan の中身。まず、スキルとデータの準備状況进行评估します。スキルは AI 成熟度レベルで測ります。

レベル	できること
L1	基本の理解。任意の Copilot の利用
L2	モデルの選定とデプロイの経験。ファインチューニングや、RAG を使わない Foundry アプリ
L3	プロンプト設計とチャック化に習熟し、RAG アプリを作る
L4	高度な ML とインフラ管理。コンテナ上の大規模な学習

PoC は、優先度の高い社内向けの案件（顧客向けでない）から始めます。使う道具は用途で分かります。

- 生成 AI: Foundry
- Copilot: Copilot Studio と Microsoft 365 Copilot の拡張
- ML: Azure Machine Learning
- 分析 AI: Foundry Tools

PoC の結果から工数を見積もり、20～30% の余裕を足します。見込みの期間は、Microsoft Copilot が ROI まで最短（日～週）で、カスタム開発ほど長くなります。Plan の次の分岐は、Azure のカスタム AI なら Ready へ進み、Copilot の導入なら Ready を飛ばして Govern へ進む形です。[^c1-28]

Ready. AI 専用のランディング ゾーンやガバナンスは、通常は作りません。既存のベースラインと管理グループの階層を拡張します。インターネットに公開するものは online、社内向けは Corp の管理グループの配下に置きます (Azure ランディング ゾーンの区分)。ランディング ゾーン アクセラレータの基準ポリシーに、承認した AI サービス (Foundry モデル、Foundry Tools、Azure AI Search、VM など) の Azure Policy を足します。公開する AI ワークロー

ドには Azure DDoS Protection を使います。オンプレミスとの接続は、予測可能な性能が要るなら ExpressRoute、要件が緩ければ VPN Gateway です。[^c1-29]

Govern. NIST AI RMF に沿って、次の4手順を踏みます。

- 1. AI の組織リスクを評価する
- 2. ガバナンス ポリシーを文書にする
- 3. ポリシーを強制する (Azure Policy と Purview で自動化し、人手も足す)
- 4. リスクを監視する (高リスクは四半期、低リスクは年次で評価)

リスク評価に使う Responsible AI の6原則は、プライバシーとセキュリティ、信頼性と安全性、公平性、包括性、透明性、説明責任です。[^c1-13]

Manage と Secure. Manage の柱は、AI CoE の設置、MLOps (従来の ML) か GenAIOps (生成 AI) かの選択、ワークロード チームへのデプロイ権限の付与 (ポリシーで枠をはめる)、モデルの再学習・昇格ゲート・廃止日の追跡、コスト管理、データ、事業継続です。[^c1-30] Secure の柱は、プロンプト インジェクションやモデル反転などのテストによるリスクの発見、AI 資産の台帳、マネージド ID、Purview での分類と DLP、モデルとデータの保護、AI 向けのインシデント対応です。[^c1-27]

AI とエージェントを作る戦略 [1.2.2]

使う・拡張する・作るの選び方は、最初の問いで決まります。問いは、SaaS のエージェントで機能要件を満たせるかです。満たせるなら既製品を採用し、満たせなければカスタム エージェントを作ります (buy first, then build)。満たすか分からない時は、候補ごとに 1〜2 週間でプロトタイプを作り、開発速度、機能の網羅、統合の複雑さを比べます。ローコード (Copilot Studio) が機能とセキュリティの要件を満たすならそれを採用し、満たさなければプロコード (Foundry) へ進みます。この分担が、ローコードとプロコードの分け方です。[^c1-31]

作る経路は3つです。

経路	型	向く用途
Microsoft Foundry	PaaS	戦略的な変革、深い統合
Copilot Studio	SaaS・ローコード	プロセスの変革
GPU とコンテナ	IaaS	コンプライアンス、高度なカスタム

Copilot Studio のローコードに、Foundry の高度なモデルを組み合わせるハイブリッドも取れ、ビジネス部門が SaaS の信頼性のまま高度なユース ケースを扱えます。Copilot Studio は 60 日間の無料トライアルで、本番の前に検証できます。

モデルのホスト先は次のとおりです。

- エージェント自体: CPU 上
- 小さな言語モデル: Azure Container Apps (サーバーレス GPU)
- 大規模モデル: AKS (大規模 GPU クラスター)

Foundry のセットアップには2種類あります。

- basic: プロトタイプ向け。ネットワーク分離なし
- standard: データ・セキュリティ・ネットワークをきめ細かく制御できる

Foundry のメモリは管理型ですが、データ主権が厳しい場合は自前 (BYO) のメモリストアを選べます。[^c1-31]

単一が複数か。 範囲が単純で、速度を優先し、コストを抑えたいなら単一のエージェントから始めます。まず単一で検証し、必要になってから複数にします。明確な役割分担があるというだけでは、複数にする根拠になりません。最初から複数にする条件は3つです。

- セキュリティやコンプライアンスの境界をまたぐ
- 複数のチームが関わる
- 将来の拡張が見込まれる

ロードマップと採用プロセス。 エージェントの計画は、ビジネス計画、技術計画、組織の準備、データ アーキテクチャの順に立てます。採用プロセスはエージェントの計画、統制とセキュリティ、構築、管理の4区分です。ロードマップには優先度の高いユース ケースを選び、案件ごとに成功指標・期間・必要なリソースを定めます。案件ごとに成功指標を定めてから、単一のエージェントで検証し、必要になってから複数に広がります。[^c1-32]

マルチエージェント ソリューションの設計 [1.2.3]

マルチエージェントの型は、親と子のエージェントと、接続されたエージェントの2つです。子エージェント (インライン) は、親の内側の軽量なサブルーチンで、設定・認証・公開は親と共通です。接続エージェントは、独自のオーケストレーション、ツール、知識、モデル設定を持つ別のエージェントです。親が各エージェントへ仕事を振る役がオーケストレーターで、Copilot Studio では親のエージェントが担います。[^c1-33] 返答を親だけにする設計では、利用者の依頼を親が受けて内容に合う子エージェントか接続エージェントへ振り、振られたエージェントが自分のツールと知識で処理した結果も親が受け取って返答にまとめるので、子や接続エージェントが利用者へ直接は返しません。

判断	子エージェントで足りる	別の接続エージェントに分ける
目的	単一目的	専用のツールや知識を持つ専門領域
管理	小さな1チームが管理する	異なるガバナンスやアクセス制御が要る
展開	別設定・別デプロイ・別公開が不要	多数の親から再利用できる
再利用	他のエージェントで再利用しない	再利用する

分割の目安は数です。1つのエージェントの選択肢 (ツール、トピック、他のエージェント) が 30〜40 を超えると、ツールの識別精度が落ち始め、説明文が似ていると少数でも劣化します。マルチエージェントの代償は、オーケストレーションの往復によるレイテンシの増加と、テスト・管理・ガバナンスの対象の拡大です。別のエージェントは別の文字起こしを持つため、親子の相関付けが要ります。[^c1-33]

接続できる外部エージェント。 次の4種類です。

- A2A プロトコル対応のエージェント
- Foundry エージェント (プレビュー)
- Fabric データ エージェント (プレビュー)

- Microsoft 365 Agents SDK 製のエージェント(プレビュー)

A2A はエージェント間通信のオープン標準で、タスクの送信、豊富なメタデータ、予測可能な応答形式を備えます。HTTP コネクタと違って複数ターンに対応し、API を呼ぶのではなくタスクを委任できます。接続の手順は、対象エージェントの Agents ページで Add agent、A2A agent を選び、エージェント カードでなく通信エンドポイントの URL を入力します。カードが標準の .well-known にあれば、名前と説明が自動で入ります。[^{^c1-34}]

接続エージェントの制限とセキュリティ。Fabric データ エージェントは、Redirect ノードと指示内の参照が使えず、Microsoft Copilot へ展開した親でも動きません。接続エージェントを持つ親は、別の親の接続エージェントになれません。引用が親へ戻る際に失われることがあります。接続先が親に無い権限を持つことがあるので、親が禁止された操作を接続先経由で実行しない設計にします。会話履歴を渡すかどうかは、接続エージェントのコンテキスト設定で決まり、自動では渡りません。[^{^c1-33}]

指示の設計。返答を親だけにする設計では、サブエージェントの指示に、自分がサブエージェントだと明記します(子や接続エージェントが直接答える設計も、意図して選ぶなら取れます)。命令は強い形で書き、サブエージェントごとに、重複しない知識ソースと区別できる説明文を割り当てます。[^{^c1-35}]

プラットフォームをまたぐ構成。Microsoft 365 Copilot・Copilot Studio・Microsoft Foundry をまたぐ構成では、A2A が別の基盤のエージェントへタスクを委任する標準の経路になります。Foundry のマルチエージェントは、ワークフローで逐次の処理、条件分岐、状態管理を行います。A2A と Activity Protocol に対応し、Microsoft 365 や Agent 365 へ公開できます。複数エージェントを手動で連鎖させると壊れやすく、ワークフローが並列・逐次・条件の調整、状態管理、人への引き継ぎを含む分岐、透明性(監査)を受け持ちます。Copilot Studio の親からは A2A や Foundry のエージェントへ接続でき、Foundry で作ったエージェントは Microsoft 365 へ公開できます(方法は 1.2.7)。[^{^c1-31}]

事前構築エージェントのユース ケース [1.2.4]

事前構築(ファーストパーティ)のエージェントは、使いどころを見極めて、作る前に先に検討する対象です。AI の決定木では、個人の効率化は Microsoft Copilot(同じ頁の決定木の説明文では Microsoft 365 Copilot と表記され、製品名の書き方が割れています)、製品内の機能は In-product Copilot、役割別は Role-based Copilot を先に検討し、足りない時に Copilot Studio や Foundry で作ります。[^{^c1-19}]

置き場	例
Microsoft 365 Copilot (SaaS)	App Builder agent、Workflows agent、Researcher。導入の枠組みは Agent Success Kit[^{^c1-31}]
製品内の SaaS エージェント	GitHub Copilot (コード)、Fabric データ エージェント、Azure Copilot エージェント、Dynamics 365 エージェント、Security Copilot エージェント[^{^c1-31}]
Power Platform (モデル駆動アプリ、プレビュー)	事前構築ではなく、自作のエージェントを載せる場所。自律エージェント (Copilot Studio 製。Power Apps MCP サーバー経由で Agent feed にタスクを出し、人が確認して介入する)と、アプリ アシスタント エージェント(アプリ内のトピックとナレッジ) [^c1-36]
Dynamics 365 Sales	Qualification・Opportunity・Close・Research・Recommended Actions の各エージェント。Microsoft 365 Copilot 内の Sales agent (CRM は Dynamics 365 Sales と Salesforce) [^c1-37]

置き場	例
Dynamics 365 Customer Service	Case Management Agent、Customer Knowledge Management Agent、Quality Evaluation Agent、Customer Intent Agent。展開管理は Rollout manager、監視は Agent insights dashboard[^c1-37]
他の Dynamics 365	Field Service の Scheduling Operations Agent (プレビュー)、Finance の Account Reconciliation Agent (production ready preview)、Supply Chain の Supplier Communications Agent、Project Operations の Time Entry・Expense・Approvals Agent[^c1-37]

足りる場面と足りない場面。標準の Copilot チャットは短い要約・返信・軽いブレインストーミングに足ります。複数のソースを横断する深い推論と、共有できる出典付きのレポートが要るなら Researcher です (処理時間は長くなります)。^[^c1-6] カスタマイズの幅が足りない、固有のシステムへの接続が要る、といった場合に Copilot Studio や宣言型エージェントで作ります。ただし宣言型エージェントを作る前に、承認済みの既存エージェントを再利用・複製・設定・拡張できないかを確認し、指示・知識・接続・サポートの所有者を決めます。既存で足りるなら新規作成は不要です。^[^c1-38]

ライセンスの前提と配布。Researcher や Analyst など一部の Microsoft 製エージェントは、事前にインストールされ固定表示され、割り当ては管理センターの Copilot controls で管理します。利用者は Agent Store からエージェントをインストールできます。管理者が Agent Store にエージェントを取り込む前提は、組織に割り当てた Microsoft Copilot のライセンスと、Microsoft 365 管理センターでの Global Admin か Teams Admin の権限です。^[^c1-39] Agent Store に取り込む経路は4つです。^[^c1-40]

- 1. 事前構築エージェントを管理者が展開する
- 2. Copilot Studio で作成して公開する
- 3. 外部プラットフォームのエージェントを持ち込む
- 4. Agent 365 SDK で統合する

Copilot Studio・Foundry・Foundry Tools の決まりと制約 [1.2.5]

Copilot Studio のナレッジの数と種類の上限。

ナレッジの種類	従来型 (会話ブースト)	生成オーケストレーション
公開 Web サイト	Web URL 4	25
SharePoint	SharePoint URL 4	25
Dataverse	2 ソース (各最大 15 テーブル)	無制限 (Quotas and limits はモードを分けずに 1 エージェント 2 ソース・各最大 15 テーブルと書き、文書で値が違う)
コネクタ (Microsoft Search の索引)	エージェントごとに 2 (会話ブーストの上限界表には載っていない)	無制限
アップロード ファイル	無制限	すべて

ナレッジの種類	従来型(会話ブースト)	生成オーケストレーション
Azure OpenAI 接続	5	ナレッジソースとしては使えない(生成回答ノードの Classic data で使う)
Bing Custom Search	2	直接は使えない
Custom data	3	直接は使えない

生成オーケストレーションでは、ナレッジが 25 を超えると内部の GPT モデルで絞り込みます。アップロード ファイルは数えません。Custom data と Bing Custom Search は直接使えず、トピックの生成回答ノードの中に置くしかありません。Teams の引用は、1 回答で最大 20 件、タイトルは約 80 文字、スニペットは約 480 文字で切り詰められるため、カスタム メッセージを作る場合は引用の表示を自前で実装します。[^c1-15] 環境の管理者が生成オーケストレーションを無効にすると、その環境のエージェントは従来型だけになります。事前構築エージェントから作る場合は、その設定が使うオーケストレーションを決めます。[^c1-8] 子・接続エージェントは親の Use general knowledge 設定に従いますが、基盤モデルの知識が質問やメッセージの生成で使われることはあります。[^c1-33]

認証・チャネル・DLP の制約。新しいエージェントは、Authenticate with Microsoft (Microsoft Entra ID の認証) が既定でオンです。DLP は Power Platform のデータ ポリシーで、コネクタを Business・Non-business・Blocked のグループに分けて管理します。同じポリシーの中で、データを受け渡すコネクタは同じグループに入れる必要があります。認証なしの公開を防ぐには Chat without Microsoft Entra ID authentication のコネクタを、特定のチャネルへの公開を止めるにはそのチャネルのコネクタをブロックします。[^c1-41] データ ポリシーは、名前を付ける→コネクタを分類する→範囲を決める→環境を選ぶ→確認して作成する、の順に作るので、コネクタの分類より先に環境を選ぶことはありません。

Foundry と Foundry Tools の決まり。ポリシーでコンテンツ フィルターの設定を強制し、許可していないモデルの利用を禁止します。定期監査も行います。[^c1-30] Foundry のエージェントは標準で、Microsoft Entra ID、RBAC、コンテンツ フィルター、仮想ネットワークの分離を備え、トレース・評価・Application Insights で可観測性を確保します。[^c1-9]

モデル ルーターの制約。有効なコンテキスト ウィンドウは、一番小さい配下のモデルで決まります。大きな入力には、モデル サブセットで対応するモデルに絞るか、要約・切り詰め・検索で入力を縮めます。画像入力は受けますが、ルーティングの判断はテキストだけで行います。音声入力は処理しません。Claude のモデルは、モデル カタログから別にデプロイしておく必要があります。クォータは 1.3.4 の表のとおり、使用量の階層で決まります。[^c1-42]

ファインチューニングの前提。学習データの用意と、学習・ホスティングの追加コストが発生します。最新情報の検索や、アプリ側の安全制御の代わりにはなりません。効果は保証されず、保持したデータでの評価が要ります。[^c1-43]

宣言型エージェントの設計上の制約。オーケストレーターとモデルは開発者が制御できません。オンプレミスの API、OpenAPI 以外の仕様、ストリーミング API には不向きです。Microsoft 365 のセキュリティとコンプライアンスをそのまま継承するため、別の姿勢が必要な組織には向きません。ホスティングは、宣言型が Copilot のオーケストレーターとモデルを使うので追加が要りません。カスタム エンジン、Microsoft 365 の外 (Azure など) に追加費用でホストが要る場合があります。ただし Copilot Studio で作るカスタム エンジン、Microsoft 365 側でホストされるので、どちらになるかは作り方で決まります。[^c1-14]

Copilot Studio での生成 AI とナレッジ ソースの使い方 [1.2.6]

生成 AI の使い方は、オーケストレーションの方式で決まります。

観点	生成オーケストレーション	従来型
トピックの選択	説明文をもとにモデルが選ぶ	トリガー フレーズで単一のトピック
ツール	自動で呼び出す	トピックの中から明示して呼ぶ
ナレッジ	先回りで検索し、複数ソースを併用する	フォールバックで検索する
回答	生成型の回答を自動で作る	ノードで組む

切り替えは、エージェントの設定の「生成 AI」にあるオーケストレーションで、生成 AI オーケストレーションを使うかどうかを はい／いいえ で選びます。[^c1-8] 生成オーケストレーションでは、子・接続エージェント、トピック、ツール、ナレッジの説明文が、選択を左右する最大の要因で、トピックのトリガーは「エージェントが選択」に変わり、説明を書きます。トピックの最終応答は、メッセージ ノードでなく出力変数として返すと、エージェントが文脈に応じて柔軟な回答にまとめられます。[^c1-8]

ナレッジ ソースの設定場所と特徴。 ナレッジはエージェント レベル (ナレッジ ページ) と、トピック内の生成回答 ノードの2通りで設定します。公開済みのエージェントは、設定したナレッジで接地します。[^c1-15]

ソース	特徴
公開 Web サイト	Bing で、指定したサイトだけを検索する
ファイル	Dataverse へアップロードする
SharePoint	Microsoft Graph の検索を使う
Dataverse	RAG で検索する
Microsoft Graph コネクタ	組織データを横断し、Microsoft Search の索引から返す

概要 ページの Use information from the web (Web Search) をオンにすると、Grounding with Bing Search で、Bing が索引する公開サイト全体を検索します。追加した公開 Web ナレッジとは並行して検索されます。ナレッジから返る引用は、他のツールやアクションの入力に使えません。引用が付かない回答は、ungrounded をオフにしていると隠されることがあります。[^c1-15]

プロンプト ツールは、エージェントの中に組み込み、文書や画像の処理、特定のタスク向けの出力をモデルに指示するものです。ベーシック・スタンダード・プレミアムの3階層があり、課金率が違います (指示の書き方は 1.2.11)。[^c1-44]

Microsoft 365 Copilot の拡張とカスタム エージェントの選び方 [1.2.7]

Microsoft 365 Copilot を拡張するか、カスタム エージェントを作るかは、モデルとオーケストレーターを誰が持つかで決まります。拡張は、宣言型エージェントに、知識を足すコネクタと、アクションを足すプラグイン (API や MCP) を組み合わせる形です。入口のページは plugin を中心にした構成に変わり、宣言型エージェントのライフサイクル用に Work IQ Dev Tools (wiqd、プレビュー) も案内されています。[^c1-45]

観点	宣言型エージェント[^c1-46]	カスタム エンジン エージェント
モデルとオーケストレーター	Copilot のものを使う	独自に持ち込む(基盤・小規模・ファインチューニング・業界特化のモデル)
ホスティング	追加不要	作り方で決まる。Copilot Studio なら Microsoft 365 側、自前の構成なら Azure などに追加費用でホスト
動作	Microsoft 365 Copilot と Microsoft 365 アプリ (Teams・Word・Excel・Outlook)。ユーザー主導のみ。個人利用向け	プロアクティブな自動実行、個人とグループの利用、外部アプリへの展開
選ぶ場面	特定の対象者と成果に向けた、名前付きの会話体験。一貫した指示と制約、厳選した知識、アクション	オーケストレーション、モデル、ホスティング、UI を完全に制御したい複雑なワークフローや高度な統合

宣言型は、完全な制御が要る場合に不適です。既存の承認済みのエージェントや機能で足りるなら、どちらも作りません。[^c1-38] 宣言型のモデルは変更できないため、独自のモデルが要る場合はカスタム エンジン エージェントになります。

開発方式。

- ローコード: Copilot Studio (完全管理の SaaS)
- プロコード: Visual Studio か VS Code と Microsoft 365 Agents Toolkit
- SDK: Microsoft 365 Agents SDK と Teams SDK

道具[^c1-47]	型	対象	特徴
Copilot Studio	ローコード	個人向け	チャンネルが多い
Teams SDK	プロコード	グループ向け	Copilot と Teams。Teams 特化で、アクション プランナーを内蔵
Microsoft 365 Agents SDK	プロコード	ISV・ストア可	10 以上のチャンネル。Semantic Kernel や LangChain も使える
Foundry	ローまたはプロコード		Copilot と Teams

Foundry のエージェントを Microsoft 365 へ公開する方法は2つです。Foundry ポータルから直接公開する (Azure Bot Service と Microsoft Entra ID が自動でプロビジョニングされる) か、Agents Toolkit のプロキシ アプリを経由し、こちらは SSO、独自ロジック、マルチ環境に向きます。[^c1-47]

政府系クラウド。 宣言型エージェントは、GCC が完全対応、GCCH が Agents Toolkit 製のプロコードと Agent Builder、DoD が Agents Toolkit 製のプロコードだけ (Agent Builder は不可) です。[^c1-38]

拡張前の記録。 作るか再利用するかを決めながら、次の6項目を記録します。決めてから開発ツールを選び、作り始めます。

- 対象ユーザーと成果
- 指示と知識
- 必要なコネクタ・アクション・MCP

- ID・権限・同意・データ境界
- 可用性・言語・ライセンス
- 所有者・サポート・成功指標

費用の観点では、カスタム エンジン を自前でホストすると、追加のホスティング費用がかかります。Copilot Studio は Copilot Credits で課金され、Microsoft 365 Copilot のライセンスを持つユーザーによる社内エージェントの利用は、一部が課金の対象外です。[^c1-38]

カスタム AI モデルを作る判断 [1.2.8]

カスタム モデルを作る前に、既製のモデルとプロンプトで足りるかを確かめます。次に、モデル ルーターが複数モデルの比較と切り替えを1つのデプロイで代わるので、先に検討します。既存のベースラインとの比較評価も勧められています。[^c1-42] 独自データでの学習が要るのは、次のような場合です。

- 高品質なタスク固有の学習データがある
- 精度を上げたい
- プロンプトを短くしてコストと遅延を下げたい
- 形式や口調を統一したい
- ツール選択を改善したい
- 小型モデルを特化させたい

微調整 (ファインチューニング) と RAG とプロンプトの使い分け。

手段	向くこと	向かないこと
プロンプト	形式、口調、手順の指定。まず試す	組織の最新の知識を持たせること
RAG	最新情報や社内の文書を答えに入れる	口調や形式の癖を学習させること
微調整 (ファインチューニング)	形式・口調の統一、プロンプト短縮、小型モデルの特化	最新情報の検索や、アプリ側の安全制御の代わり

カスタマイズ手法。手法は3つです。

- SFT: 正解の入出力の例で学習する
- DPO: 好ましい応答と好ましくない応答の対で学習する
- RFT: 採点器つきの強化学習。明確な正解と採点ができる領域向けで、ML の専門性が要る

迷ったら SFT から始めます。データ量の目安は、最初のテストが 50〜100 件の高品質な例、本番用モデルは 500 件以上です。[^c1-48]

方式と種別。学習の方式は2つあります。

- 管理型ファインチューニング: データと設定を渡す。初心者向け
- 対話型トレーニング (プレビュー): Python でループを書く。承認制

どちらも学習用の GPU の用意は不要です。学習のタイプは次のとおりです。

タイプ	特徴
Standard	リージョン内で、データ所在地を保証する
Data zone	現在は米国のみ
Global	安価だが、データとウェイトが別リージョンへ出る
Developer	アイドル容量を使い、保証なし。実験向け

デプロイの選択肢と種別は次のとおりです。[[^]c1-48]

- 選択肢: Serverless、Managed compute (プレビュー)、Fireworks (プレビュー)
- 種別: Standard、Data Zone Standard、Global Standard、Developer (24 時間、SLA なし)、Provisioned Throughput など

成熟度と運用。 L2 でファインチューニングや RAG を使わない Foundry アプリ、L3 で RAG アプリや ML モデルの学習とデプロイ、L4 で大規模モデルの学習と実行 (VM、AKS、Container Apps) が対象です。従来のファインチューニングは、専門のデータサイエンティストと専用のパイプラインが要り、運用が複雑なため、エージェント化や簡易なカスタマイズで足りるなら、そのほうが運用を単純にできます。事前学習モデルの提供終了日を監視し、本番への昇格には品質ゲートを置きます。[[^]c1-30] 旧画面の Foundry (classic) の記事は新しいポータルでは使えず、現行は Fine-tuning in Microsoft Foundry です。[[^]c1-43]

プロンプト ライブラリの作り方の指針 [1.2.9]

Microsoft 365 には、プロンプトを集めて共有する場所が2つあります。Copilot Prompt Gallery は、Microsoft 作成のプロンプト (Suggested) と、ユーザーやチームが保存・共有するプロンプトのカタログです。Microsoft Copilot か Copilot Chat の中で使います。保存先はテナントの境界内にある Microsoft 365 Substrate のデータストアで、未認証のユーザーには Microsoft 製のプロンプトしか見えません。管理者は、特定ユーザーが保存・いいね・共有したプロンプトや、チームへ共有されたプロンプトを分析し、エクスポートして追跡できます。[[^]c1-49]

組織のプロンプトは、管理者が作ります。手順は Microsoft 365 管理センターの Copilot、Prompts、Create prompt で入力し、Publish です。必要なロールは AI Administrator か Search Editor です。[[^]c1-50]

項目	内容
公開できる数	テナントで最大 1,000 件
反映	プロンプト ラボへの表示まで通常約 3 時間
一括管理	作成・編集・削除に加え、全件のエクスポート、インポート、ピン留め
多言語	言語フィールドで登録できる。自動翻訳はなく、言語ごとに別のプロンプトを作る

ライブラリの運用では、次の4つを決めておくと回しやすくなります (公式のガイドラインの文言ではなく、Prompt Gallery と組織プロンプトの機能から組んだ目安です)。

- 目的別の分類: 業務ごとに探せるように分けて作る
- 共有の範囲: 個人、チーム、組織のどこまでにするか決めて公開する

- 版の管理: 改訂のたびにエクスポートで控えを残す
- 保守の責任者: 置いて、古くなったものを見直して削除する

この教材では、作る→共有の範囲を決めて公開する→改訂のたびに控えを残す→古いものを見直して削除する、の順に回します。

小さな言語モデル (SLM) のユース ケース [1.2.10]

小さな言語モデルは、端末内で動かせる軽さ、低い待ち時間、低いコストが売りで、特定の仕事に絞って使います。分類、感情分析、コンテンツ モデレーションのように絞ったタスクは、小さなモデルで始めます。教師モデルが生成した例で小型モデルを特化させると、より低コスト・低遅延で要件を満たせます。[^c1-48]

Foundry で提供される Microsoft 製の Phi モデルの仕様は次のとおりです。[^c1-51]

モデル	入力	出力	備考
Phi-4-mini-instruct	テキスト 131,072 トークン	4,096 トークン	日本語を含む多言語。ツール呼び出しなし。出力はテキスト
Phi-4-multimodal-instruct	テキスト・画像・音声 (131,072 トークン)	テキスト 4,096 トークン	日本語を含む多言語。ツール呼び出しなし
Phi-4	16,384 トークン	16,384 トークン	日本語を含む多言語。ツール呼び出しなし
Phi-4-reasoning	32,768 トークン	32,768 トークン	英語のみ。ツール呼び出しなし
Phi-4-mini-reasoning	128,000 トークン	128,000 トークン	英語のみ。ツール呼び出しなし

SLM を業務向けに調整する。 Foundry の SFT に対応するモデルに Phi 4 と Phi-4-mini-instruct があり、Phi 4 は SFT だけに対応します。Phi 4 は単純なタスク向けの低コストの選択肢と説明されています。DPO は GPT-4o / 4.1 系、RFT は o4-mini が対象です (classic の文書)。蒸留 (distillation) は、高性能モデルの知識を小型モデルへ移す用途です。classic の文書は GPT-4.1-mini から始め、反復を速くコストを抑えることを勧めています。効果は保証されないで、保持したデータで評価し、ベースラインと比べます。[^c1-43]

大きなモデルとの比べ方と併用。 比べる軸は、コスト、待ち時間、コンテキスト長、対応する入出力です。ホスト先は、小さな言語モデルなら Azure Container Apps のサーバーレス GPU (自動スケールで運用負荷が小さい)、大規模モデルなら AKS です。[^c1-31] カスタム エンジン エージェントなら、基盤 LLM のほかに小さな言語モデル、ファインチューニング済み、業界特化のモデルを選べます。宣言型では選べません。[^c1-46] モデル ルーターは、複雑さに応じて小さく安いモデルを使い、複雑な推論の時だけ大きなモデルを使います。プールには gpt-5-nano / mini、gpt-4.1-nano / mini、Claude の haiku 系などが含まれます。[^c1-42]

プロンプト エンジニアリングの指針と技法 [1.2.11]

プロンプトは次の5つの部品でできています。

- 指示
- 主要コンテンツ
- 例 (one-shot / few-shot)

- キュー
- 補助コンテンツ

Chat Completions では、指示がシステム メッセージになり、few-shot の例はユーザーとアシスタントの対で表します。[^c1-52] 技法は次のとおりです。

技法	中身
役割と文脈	システム メッセージにモデルの役割、対象、制約を書く
例示	few-shot の例を添えて、出力の型を見せる
先にタスク	作業を冒頭で指示してから文脈を渡す。GPT-4 系は順序の影響が小さいが、推奨は続く
末尾で繰り返す	新しい情報の影響が強い (recency bias) ので、最後に指示を繰り返すと効くことがある
出力のプライミング	最後に書き出しの語句 (箇条書きの先頭など) を置いて、形式を誘導する
段階の分割	先に事実を抽出し、その後で検索クエリを作る、のように手順へ分ける
出力構造の指定と引用	「事実だけ書け」より、出典の引用や構造 (entity1, relationship, entity2 など) の指定のほうが接地しやすい
接地コンテキスト	回答の元データをプロンプトに与える。元資料が最終回答に近いほど、モデルの作業が減る
Affordance	検索などの外部手段を呼ばせ、結果を戻して回答させる。作り話が減り、最新情報が得られる

Chain of thought (段階的な思考過程を出させる手法) は、非推論モデルだけに適用します。推論モデルは対象外で、推論過程の抽出は reasoning summary パラメーター以外では対応していません。Temperature は 0~2 で、高い値 (例 0.7) は多様に、低い値 (例 0.2) は集中して具体的にになります。Temperature と Top_p は、片方だけ調整することが勧められます。[^c1-52]

Copilot Studio の指示 (システム メッセージ) の書き方。 エージェントの振る舞い、してはいけない事、根拠の示し方を書きます。命令は MUST や NEVER のような強い形にします。丁寧な表現は、競合した時に優先度が下がります。出典の in-text 引用を常に求め、引用を省く指示は書きません。JSON のみといった厳格な出力形式の指定は、引用を妨げることがあります。簡潔に書きます。[^c1-35] [^c1-15]

Microsoft AI Center of Excellence の要素 [1.2.12]

AI CoE は、AI の成果を出すための社内の専門家チームです。分断された無統制な AI 導入を防ぎ、事業と技術の相談に乗ります。既存の Cloud CoE があれば、そこへ AI の実践を統合します。既存のチームで足りない場合などに独立させます。[^c1-53]

チームの構築は5手順です。

1. 経営層のスポンサーを確保する (運営委員会、月次レビュー)
2. リーダーを任命する (単一の窓口)

3. 多職種のチームを編成する
4. 組織内の配置を決める
5. 運用モデルを定義する

8つの責務。AI 戦略の定義、スキル開発、パイロットの主導、標準の策定と強制、受付と優先順位づけの流れ、再利用資産の作成、成果の測定と報告、AI サービスの運用管理(任意)です。測る KPI は、採用率、コンプライアンス水準、プロジェクトのサイクル タイムなどで、定期的に経営層へ報告します。再利用資産は、コンプライアンスのチェックリストなどを社内プラットフォームに公開するもので、テンプレートやコードのリポジトリのライブラリの維持は、AI サービスの運用管理(任意)の中身です。スキル開発では、現在のスキルを評価し、学習パスとハンズオンの実験の機会を提供します。[^c1-53]

CoE の要素。8つの責務は、5つの要素に並べ直せます(公式の分類ではなく、この教材の並べ直しです)。

要素	責務
戦略	AI 戦略の定義、パイロットの主導
人材とスキル	スキル開発
統制	標準の策定と強制、受付と優先順位づけの流れ
技術	再利用資産の作成、AI サービスの運用管理(任意)
価値の測定	成果の測定と報告

CoE の運営。スポンサーと役割は上の5手順で決め、利用の推進は再利用資産とスキル開発が担い、成功の物差しは KPI です。

運用モデルの進化。初期は集中型の CoE に専門性を集約し、導入を加速します。成熟したら、承認ゲートの門番から、ガードレールを示す助言型へ移り、AI の配信はプラットフォーム チームに組み込みます。移る時期は、集中管理が採用の妨げになっている兆候(承認待ちなど)を監視して見極めるので、変曲点を逃しません。[^c1-53] 戦略の面では、AI CoE のような中央のガバナンス機能が Responsible AI の要件と共通の境界を定め、事業部門はその範囲で機会を追求します。Manage では AI デプロイ全体の戦略監督と技術指導を担い、ガバナンスと革新を両立するデプロイ標準を作ります。Responsible AI のポリシーでは、CoE は標準を定めて相談に乗る役で、ゲートキーパーではありません。[^c1-19] [^c1-54]

Power Platform CoE キットとの関係。CoE キットは Power Platform の採用を支える参照実装で、Admin・Govern・Nurture の領域があります。これは Power Platform 向けの道具で、AI CoE の組織そのものではありません。CoE には人、コミュニケーション、プロセスも要ります。現在このキットは積極的には保守されておらず、中核の機能は Power Platform 管理センターに入っています。[^c1-55]

複数の Dynamics 365 アプリを使う AI ソリューション [1.2.13]

Sales・Customer Service・Customer Insights などは、Power Apps のモデル駆動アプリで Dataverse の上で動きます。そのため Copilot や AI の機能の多くを共有します。ただし、アプリによって有効化の方法や表示が違います。共通の機能は次のとおりです。[^c1-37]

- Microsoft Copilot (プレビュー)

- Agent feed (人とエージェントの協働の活動フィード、プレビュー)
- Power Apps MCP サーバー (プレビュー)
- フォームの入力支援
- ビューでのデータ検索

各アプリのエージェントは 1.2.4 の表のとおりで、Sales のエージェントは MCP ツールでほかのエージェントやアプリからも使えます。[^c1-7]

アプリをまたぐデータとエージェント。 Sales・Customer Service の側は Dataverse を共通の土台にします。Finance と Supply Chain Management の側とのつながり方は、次の表のとおりです。

つながり方	中身
共通の Dataverse	同じ Dataverse 上のアプリは、データとエージェントを共有しやすい
二重書き込み (dual-write)	Finance・Supply Chain Management と Dataverse を、近いリアルタイムで双方向に同期する。データは両側に持つ。[^c1-56]
仮想テーブル	Finance・Supply Chain のデータを、コピーせず Dataverse のテーブルとして見せる。データは Finance 側に置いたまま、Dataverse から作成・更新・削除もできる。生きた運用データを使い、Copilot Studio のエージェントのナレッジにも追加できる (ただし財務と運用アプリの生成ヘルプ=アプリ内ヘルプのナレッジとしては、現在は非対応と書く文書もある)。[^c1-57]
共有のナレッジ	Dataverse のテーブルや SharePoint を、複数のエージェントが同じナレッジとして使う

アプリ間の連携の例は3つです。[^c1-37]

- Customer Insights - Journeys、Contact Center、Copilot Studio 製の AI エージェントを組み合わせ、会話型のジャーニーで大規模な顧客体験を出す
- Project Operations: Time and Expense Agent が、Copilot Studio のエージェントで入力された時間・経費・資料を初期レビューし、Approvals Agent が承認フローを効率化する
- Microsoft 365 Copilot の側: Sales と Service が Dynamics 365 のデータを Outlook などを使い、メールの要約、返信の下書き、商談サマリーを出す

Dynamics 365 のエージェントは、AI の決定木では「製品内の SaaS エージェント」に当たります。既製品として先に検討し、足りない時に Copilot Studio や Foundry で作ります。[^c1-31] 段階的に導入するなら、Customer Service の Rollout manager で、AI 管理者や監督者が自律エージェントを管理しながら広げます。進め方は、既製品で足りるかを確かめる→足りない部分を作る→使う機能の提供状態を確かめる→段階的に広げる、の順なので、提供状態を確かめる前に段階的な展開へは進みません。Microsoft Copilot、Agent feed、Power Apps MCP、Scheduling Operations Agent、Account Reconciliation Agent (production ready preview) などはプレビュー表記なので、本番に載せる前に提供状態を確認します。[^c1-37]

ROI の基準と総保有コスト [1.3.1]

まず、ROI の物差しを、コストの低下、業務の速さ、よりよい意思決定、顧客体験の向上のような業務価値から選びます。次に、選んだ物差しに照らして案を採点します。採点の軸は、ビジネス効果（戦略との整合、業務価値、変更管理の期間）、技術的実現性、ユーザーの受容です。そのうえで成功指標のベースラインと目標 KPI を決めます。[^c1-5]

総保有コスト(TCO)の内訳。この教材では、次の6つの項目に分けて考えます(公式の区分ではありません)。

- ライセンス
- 従量の利用料
- 開発
- 運用
- 統制
- 教育

製品別のコスト要因

製品[^c1-19]	主なコスト要因
Microsoft Copilot	ライセンス
Copilot Studio	Copilot Credits
Foundry	モデルの利用とツールの利用
Fabric	容量、ストレージ、複製
Azure ML・インフラ	計算、ストレージ、転送

カスタム化の費用も TCO に入れます。カスタム エンジン エージェントを自前でホストするなら Microsoft 365 の外 (Azure など) のホスティング費が、ファインチューニングは学習費とホスティング費が加わります。[^c1-46] マルチエージェントでは、エージェント間の呼び出しごとに設計・エラー処理・状態の同期・監視が要り、重複したコンテキスト処理で費用が増えます。単一のエージェントはトークンと API の呼び出しを抑えやすいので、コスト優先なら単一から始めます。[^c1-32]

Copilot Studio の支払い方法。3つあります。事前購入の Copilot Credit パック (テナントで共有)、従量課金 (Azure サブスクリプションに環境を課金ポリシーで接続)、1 年の Copilot Credits 事前購入プラン (CCCU) です。購入した容量は毎月適用され、未使用の Credits は翌月に繰り越されません。エージェント フローは、容量が尽きると新規の実行がブロックされます。[^c1-58]

Copilot Credit の課金率。

項目[^c1-44]	Credits
従来型の回答	1
生成回答	2
アクション	5
テナント グラフ接地	10

項目[^c1-44]	Credits
エージェントフロー	100 アクションあたり 13
テキスト／生成 AI ツール (ベーシック／スタンダード／プレミアム)	1,000 トークンあたり 0.1／1.5／10

計算の例です。顧客向けのエージェントが 1 日 900 人を受け、1 人あたり生成回答 4 回とアクション 1 回なら、 $(4 \times 2 + 1 \times 5) \times 900 = 11,700$ Credits です。月の消費は、1 日の値に日数を掛けて出し、毎月適用される購入した容量と比べます。見積もりには、種類・トラフィック・オーケストレーション・ナレッジ・ツールを入れて予測する Microsoft Copilot Studio agent usage estimator を使います。Microsoft 365 Copilot のライセンスを持つユーザーが社内で使う Copilot Studio エージェント (Copilot Chat・Teams・SharePoint の従来型・生成回答・Graph 接地) は、メッセージバックや従量メーターを消費しません。Microsoft 365 の Agent Builder で作り、テナント グラフ接地を使わない生成回答も課金されません。[^c1-58]

Microsoft Copilot の提供形態は、Microsoft 365 E7 に含まれるか、対象の商用・政府・教育プランへのアドオンです。Business 向けは最大 300 ユーザー、年間契約です。ライセンスの文書は「Microsoft Copilot」の表記で、試験範囲の「Microsoft 365 Copilot」と製品名が一致しない箇所があります。[^c1-59] Foundry には専用の価格計算ページがなく、構成する Azure サービス (Speech・Language・AI Search など) を Azure 料金計算ツールで合算します。[^c1-60]

業務プロセスの ROI 分析 [1.3.2]

分析の入口は、業務の現状パフォーマンスをベースラインとして置くことです。エージェントが改善する業務の KPI を少数選び、パイロットの結果を go/no-go のゲートに使います。導入後は実績を目標 KPI と定期的に比べ、足りなければ設計の改善、廃止、他の機会への転用を決めます。PoC では開発時間、テスト サイクル、デブロイの複雑さを記録し、実測で工数を見積もって、計画には 20～30% のバッファを足します。[^c1-61] 例の数字は、すべて仮のものです (単位は万円)。

項目	値
年間のコスト	ライセンス 2,400、利用料 600、運用 700
初期のコスト	開発 1,800、教育 500
年間の効果	300 人×週 1.5 時間×46 週×時給 4,000 円＝8,280
ROI	$(8,280 - 6,000) \div 6,000 = 38\%$
回収期間	初期 2,300 ÷ 月の純効果 381.7＝約 6.0 か月
感度	効果が 70% なら 5,796 で ROI は -3.4%、50% なら -31%

6,000 は、年間のコストと初期のコストの合計です。感度の行は、期待の効果が下振れしても投資が回収できるかを見るための行です。

効果の測り方と追い方。 Copilot Analytics は、Copilot の利用データ、エージェントの利用データ、組織の Microsoft Graph データを入力にします。SAP・Salesforce・Workday など他社製品の組織指標もアップロードできます。測定は3つの観点で、準備と導入 (Readiness and adoption)、生産性への影響 (Productivity impact)、ビジネス価値と ROI です。これは Copilot controls の3本柱のうち「測定とレポート」の柱の中身です。[^c1-62]

見る場所	中身
Copilot Dashboard (Viva Insights)	Adoption タブと Impact タブ。Impact は会議・メール・Teams のチャットに費やす時間の変化など、働き方の変化を示す。Pulse と Glint のサーベイや業界ベンチマークとの比較もできる
Microsoft 365 管理センターのレポート	利用の指標
Copilot Studio の Analytics タブ	エージェントの利用
Viva Insights の高度な分析	組織指標と照合する

Copilot Business Impact Report は、組織の主要指標を取り込み、Copilot とエージェントのビジネス価値を測る上級レポートで、その学びを展開計画に反映します。[^c1-62] 運用コストは、リソース使用の監視、コスト超過の自動検出、Cost Management での CPU・GPU・メモリ・ストレージの追跡、予算アラートで見ます。[^c1-13] Foundry は、計算ツールで見積もり、小さなテストワークロードを動かし、Cost Management でリソース別・メーター別の実績と比べ、予算ベースラインを調整します。コストと使用量の記録は反映が遅れることがあり、分単位の照合には向きません。期間の傾向で見ます。[^c1-60] エージェントの価値は、効率（反復作業の自動化）、速度、拡張性で語り、測れる成果として運用コストの低下、顧客満足度の向上、イノベーションの加速を挙げます。[^c1-63]

構築・購入・拡張の判断 [1.3.3]

作る・買う・拡張するの判断は、基本が buy first, then build です。SaaS のエージェントが機能要件を満たすなら採用し、満たさなければ構築します。満たすか分からない時は、Copilot Studio と Foundry を時間枠つきで比べます。[^c1-31] 判断の流れは次のとおりです。

- 1. 手順が固定なら、コードや非生成 AI を選ぶ。静的な知識の検索や要約だけなら、古典的な RAG (Foundry で構築可) で足りる[^c1-3]
- 2. 既製品を探す。個人の効率化は Microsoft Copilot (ライセンス)、製品内の機能は基本製品の契約に含まれるか追加ライセンス、役割別は Security Copilot や Sales・Service・Finance のエージェント
- 3. 拡張する (extend)。Microsoft 365 Copilot を使う組織が専門知識とエージェントでカスタマイズするなら、宣言型エージェントによる拡張で、主なコスト要因はライセンスと従量課金。独自のモデルやオーケストレーションが要るならカスタム エンジンで、ホスティング費が加わる[^c1-19] [^c1-46]
- 4. 構築する。管理型を優先し、要件を満たす範囲で運用責任が最小のものを選ぶ

構築先は、Copilot Studio が SaaS・ローコード・プロセス変革・中程度のカスタマイズです。Foundry は PaaS で、プロコードも使い、戦略的変革、深い統合、モデルの選択に向きます。モデルを自分で持つ必要があるなら、GPU とコンテナーです。Foundry のエージェントは、管理を最小にするなら Prompt agent、完全に制御するなら Hosted agent、ほかで動くエージェントは Responses API で Foundry のモデルとツールだけを使います。[^c1-9]

差を決める要素は、差別化、時間、コスト、保守の負担、データの持ち場所です。差別化の必要が小さく、早く ROI に届けたく、保守の負担を避けたいなら既製品へ、差別化やデータの持ち場所の管理が要るなら拡張や構築へ寄ります。時間は、Microsoft Copilot が日～週で ROI に届く最短で、カスタム開発ほど長くなります。スキルと成熟度も効き、L1 は任意の Copilot、L2 以降でカスタムの分析 AI やファインチューニング、L3 で RAG アプリ、L4 で大規模な学習です。ライセンスの要件は、Copilot Studio が単体のライセンスかアドオン、Microsoft Copilot が Microsoft 365 のライセンスに Copilot のライセンスを加えたもの、Azure は Azure のアカウントです (再利用の確認は 1.2.4)。[^c1-4]

モデル ルーター [1.3.4]

モデル ルーターは、プロンプトを複雑さ、推論の要否、タスクの種別などで分析し、リアルタイムに最適な LLM へ振り分ける、学習済みの言語モデルです。ほかの Foundry モデルと同じようにデブロイして使い、プロンプトは保存しません。使う理由は、小さく安いモデルで足りる時にコストと遅延を下げ、複雑な推論が要る時だけ大きなモデルや推論モデルを使い、品質を保ったまま1つのデブロイにまとめられることです。従来は、モデルを個別に比べ、ルーティングのロジックを自作して、品質・コスト・遅延の釣り合いを探しました。ルーターは要求ごとに適したモデルへ振り分け、品質とコストの釣り合いを管理します。ベースラインと比較評価し、効果を確認めます。[^c1-42]

ルーティングの方針(モード)。

モード	選び方
Balanced (既定)	最高品質から 1〜2% 以内の候補で最安を選ぶ
Cost	5〜6% の許容幅の中で最安を選ぶ
Quality	コストを無視して最高品質のモデルを選ぶ

デブロイと設定(対象のモデルの選び方)。 Quick deploy か Custom deploy を選びます。Custom deploy ではルーティングモードとモデル サブセットを指定でき、サブセットで使うモデルを限定して、コンテキストや要件を満たします。現行のバージョンは 2025-11-18 (最新、随時更新) です。2025-08-07 と 2025-05-19 は凍結され、新しいモデルは加わりません。Auto-update を選ぶと新しいバージョンへ自動更新され、配下のモデルの変更で性能とコストが変わりうる点に注意します。対応するデブロイ種別は Global Standard と Data Zone Standard で、データゾーンの境界を守り、アクセス権とデブロイ種別で使えるモデルにだけルーティングします。配下のモデル プールには、OpenAI (gpt-5.x / gpt-4.1 / gpt-oss など)、Anthropic の Claude 系、xAI の grok、DeepSeek、Meta Llama、Fireworks の各モデルが含まれ、Claude 以外は別のデブロイが不要です。Claude 系をサブセットに入れるなら、ルーターより先に Claude をモデル カタログからデブロイしておきます。[^c1-42]

クォータ。 上限は使用量の階層で広がります。

階層	Global Standard	Data Zone Standard
Tier 1	1,000 RPM / 1,000,000 TPM	300 RPM / 300,000 TPM
Tier 2	2,000 RPM / 2,000,000 TPM	670 RPM / 670,000 TPM
Tier 3	4,000 RPM / 4,000,000 TPM	1,000 RPM / 1,000,000 TPM
Tier 4	7,000 RPM / 7,000,000 TPM	2,000 RPM / 2,000,000 TPM
Tier 5	10,000 RPM / 10,000,000 TPM	3,000 RPM / 3,000,000 TPM
Tier 6	15,000 RPM / 15,000,000 TPM	4,000 RPM / 4,000,000 TPM

実際に流せる要求数は、RPM と、TPM をトークン数で割った値のうち小さいほうです。Tier 1 の Global Standard で 1 要求が平均 2,500 トークンなら、 $1,000,000 \div 2,500 = 400$ 件 / 分が上限で、TPM が先に尽きます。平均 500 トークンなら RPM の 1,000 件 / 分が先に当たります (TPM 換算では 2,000 件 / 分)。Tier 1 の Data Zone Standard で 2,500 トークンなら、 $300,000 \div 2,500 = 120$ 件 / 分までです。[^c1-64]

課金と使いどころ。入力プロンプトの料金が、価格ページのレートで課金されます。Azure ポータルでデプロイのコストを監視できます。Foundry Agent Service では、適格な OpenAI・OSS・Anthropic のモデルで、ツールの利用を伴うエージェントのシナリオに対応します。[^c1-42]

トラブルシューティング。デプロイが失敗したらリージョンを確認します。Claude が使われないなら、先に Claude をデプロイします。想定外のモデルが選ばれたら、モード設定とサブセットを見直します。[^c1-42]

SAMPLE

サンプルはここまでです

続き(第2章～巻末)は、AB-100 問題集のご購入で、頻出度別ノート・暗記用ノートと合わせて3点すべてダウンロードできます。

SAMPLE

Microsoft 認定 Agentic AI ビジネス ソリューション アーキテクト エキスパート(AB-100) 学習用テキスト(無料サンプル)

版

2026年10月版(AB-100 出題範囲(Study guide・Skills measured as of November 16, 2026)準拠)

発行

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

お問い合わせ

info@shikaku-doujou.com

Microsoft、Microsoft Copilot、Copilot Studio、Microsoft Foundry、Dataverse、Power Platform、Power Automate、Azure および Microsoft の製品名・サービス名は、Microsoft Corporation の米国およびその他の国における商標または登録商標です。記載されているその他の会社名・製品名・サービス名は、各社の商標または登録商標です。

本教材は公式シラバスの内容を基に作成したもので、Microsoft Corporation との提携・承認関係はありません。

本文・図表・設問はすべてオリジナルの著作物です。無断での転載・複製・再配布・二次販売はできません。お気づきの点は info@shikaku-doujou.com までご連絡ください。