

AWS Certified Generative AI Developer - Professional (AIP-C01) 試験ガイド

はじめに

AWS Certified Generative AI Developer - Professional (AIP-C01) 試験は、GenAI デベロッパーの役割を担う方を対象としています。この試験では、受験者が基盤モデル (FM) をアプリケーションとビジネスワークフローに効果的に統合できるかどうかを検証されます。この認定は、AWS テクノロジーを使用して GenAI ソリューションを本番環境に実装する方法に関する実践的な知識を証明するものです。

また、以下のタスクについての受験者の能力も検証します。

- ベクトルストア、検索拡張生成 (RAG)、ナレッジベース、その他の GenAI アーキテクチャを使用してソリューションを設計および実装する。
- FM をアプリケーションやビジネスワークフローに統合する。
- プロンプトエンジニアリングと管理手法を適用する。
- エージェント AI ソリューションを実装する。
- GenAI アプリケーションのコスト、パフォーマンス、ビジネス価値を最適化する。
- セキュリティ、ガバナンス、責任ある AI のプラクティスを実装する。
- GenAI アプリケーションをトラブルシューティング、モニタリング、最適化する。
- FM の品質と責任を評価する。

受験対象者について

受験対象者は、AWS またはオープンソーステクノロジーでのプロダクショングレードのアプリケーション構築について 2 年以上の経験と、一般的な AI/ML またはデータエンジニアリングの経験、GenAI ソリューションの実装について 1 年の実務経験を有している必要があります。

推奨される AWS の知識

受験対象者は、以下の AWS に関する知識を有している必要があります。

- AWS のコンピューティング、ストレージ、ネットワークサービスの経験

- AWS のセキュリティのベストプラクティスとアイデンティティ管理についての知識
- AWS のデプロイと **Infrastructure as Code (IaC)** ツールの経験
- AWS のモニタリングおよびオブザーバビリティサービスに精通していること
- AWS のコスト最適化の原則についての知識

受験対象者にとって試験対象外となるジョブタスク

受験対象者が実施できることが想定されていないジョブタスクは、以下のリストのとおりです。このリストはすべてを網羅しているわけではありません。以下のタスクは、本試験の対象外です。

- モデルの開発とトレーニング
- 高度な ML 手法
- データエンジニアリングと特徴量エンジニアリング

試験に出題される可能性のあるテクノロジーと概念のリスト、試験対象の **AWS** のサービスと機能のリスト、試験対象外の **AWS** のサービスと機能のリストについては、「付録」を参照してください。

試験内容

設問の種類

試験には、以下の出題形式が 1 つ以上含まれています。

- **択一選択問題:** 正しい選択肢が 1 つ、誤った選択肢 (不正解) が 3 つ提示される。
- **複数選択問題:** 5 つ以上の選択肢のうち、正解が 2 つ以上ある。設問に対する点数を得るには、正解をすべて選択する必要がある。
- **並べ替え:** 指定されたタスクを完了することを目的とした 3~5 つの答えのリストが提示される。設問に対する点数を得るには、正解を選択し、正しい順序に並べる必要がある。
- **内容一致:** 3~7 つのプロンプトのリストと一致する答えのリストが提示される。設問に対する点数を得るには、すべてのペアを正しく一致させる必要がある。

未解答の設問は不正解とみなされます。推測による解答にペナルティはありません。試験には、スコアに影響する設問が 65 問含まれています。¹

採点対象外の設問

試験には、スコアに影響しない採点対象外の設問が 10 問含まれています。AWS では、こういった採点対象外の設問でのパフォーマンス情報を収集し、今後採点対象の設問として使用できるかどうかを評価します。試験では、どの設問が採点対象外かは受験者にわからないようになっています。

試験の結果

AWS Certified Generative AI Developer - Professional (AIP-C01) 試験は、合否判定方式の試験です。試験は、認定業界のベストプラクティスおよびガイドラインに基づき、AWS の専門家が定めた最低基準に照らして採点されます。

試験の結果は、100～1,000 の換算スコアとして報告されます。合格スコアは 750 です。このスコアにより、試験全体の成績と合否がわかります。複数の試験間で難易度がわずかに異なる可能性があるため、スコアを均等化するために換算スコアが使用されます。

スコアレポートには、各セクションのパフォーマンスを示す分類表が含まれる場合があります。試験には補整スコアリングモデルが使用されるため、セクションごとに合否ラインは設定されておらず、試験全体のスコアで合否が判定されます。

試験の各セクションには特定の重みが設定されているため、各セクションに割り当てられる設問数が異なる場合があります。分類表には、受験者の得意分野と不得意分野を示す全般的な情報が含まれます。セクションごとのフィードバックを解釈する際は注意してください。

¹試験のベータ版は該当しません。ベータ試験全般の詳細については、[AWS 認定のウェブサイト](#)をご覧ください。

試験内容の概要

この試験ガイドには、試験に設定された重み、コンテンツ分野、タスク、スキルについての説明が含まれています。本ガイドは、試験内容の包括的なリストを提供するものではありません。

本試験のコンテンツ分野と重み設定は、以下のとおりです。

- コンテンツ分野 1: 基盤モデルの統合、データ管理、コンプライアンス (採点対象コンテンツの 31%)
- コンテンツ分野 2: 実装と統合 (採点対象コンテンツの 26%)
- コンテンツ分野 3: AI の安全性、セキュリティ、ガバナンス (採点対象コンテンツの 20%)
- コンテンツ分野 4: GenAI アプリケーションの運用効率と最適化 (採点対象コンテンツの 12%)
- コンテンツ分野 5: テスト、検証、トラブルシューティング (採点対象コンテンツの 11%)

コンテンツ分野 1: 基盤モデルの統合、データ管理、コンプライアンス

タスク 1.1: 要件を分析し、GenAI ソリューションを設計する。

スキル 1.1.1: 特定のビジネスニーズと技術的制約に沿った包括的なアーキテクチャ設計を作成する (適切な FM、統合パターン、デプロイ戦略の使用など)。

スキル 1.1.2: 本格的なデプロイに進む前に概念実証の技術的実装を開発し、実現可能性、パフォーマンス特性、ビジネス価値を検証する (Amazon Bedrock の使用など)。

スキル 1.1.3: 標準化された技術コンポーネントを作成し、複数のデプロイシナリオにわたって一貫した実装を確保する (AWS Well-Architected フレームワーク、AWS WA Tool Generative AI Lens の使用など)。

タスク 1.2: FM を選定して設定する。

スキル 1.2.1: FM を評価して選定し、特定のビジネスユースケースや技術要件に沿って最適に調整する (パフォーマンスベンチマーク、能力分析、制限評価の使用など)。

スキル 1.2.2: 柔軟なアーキテクチャパターンを作成し、コードを変更せずに動的なモデル選択とプロバイダーの切り替えができるようにする (AWS Lambda、Amazon API Gateway、AWS AppConfig の使用など)。

スキル 1.2.3: 耐障害性の高い AI システムを設計し、サービス中断中でも継続的に運用できるようにする (AWS Step Functions サーキットブレーカーパターン、リージョンの可用性が限られているモデルに対する Amazon Bedrock クロスリージョン推論、クロスリージョンモデルのデプロイ、グレースフルデグラデーション戦略の使用など)。

スキル 1.2.4: FM カスタマイズデプロイとライフサイクル管理を実装する [Amazon SageMaker AI を使用したドメイン固有のファインチューニングされたモデルのデプロイ、モデルデプロイのための低ランク適応 (LoRA) やアダプターなどパラメータ効率の高い適応手法、SageMaker Model Registry を使用したバージョンングとカスタマイズ済みモデルのデプロイ、自動デプロイパイプラインを使用したモデルの更新、デプロイが失敗した場合のロールバック戦略、モデルの廃止と交換のためのライフサイクル管理の使用など]。

タスク 1.3: FM 消費のためのデータ検証と処理パイプラインを実装する。

スキル 1.3.1: 包括的なデータ検証ワークフローを作成して、データが FM 消費の品質基準を満たしていることを確認する (AWS Glue Data Quality、SageMaker Data Wrangler、カスタム Lambda 関数、Amazon CloudWatch メトリクスの使用など)。

スキル 1.3.2: テキスト、画像、音声、表形式データなどの複雑なデータタイプを、FM 消費に特化した処理要件に沿って処理するデータ処理ワークフローを作成する (Amazon Bedrock マルチモーダルモデル、SageMaker Processing、AWS Transcribe、高度なマルチモーダルパイプラインアーキテクチャの使用など)。

スキル 1.3.3: モデル固有の要件に従って FM 推論用の入力データをフォーマットする (Amazon Bedrock API リクエストには JSON フォーマット、SageMaker AI エンドポイントには構造化データ前処理、ダイアログベースのアプリケーションには会話フォーマットの使用など)。

スキル 1.3.4: 入力データの品質を高めて、FM の応答の品質と一貫性を向上させる (Amazon Bedrock を使用したテキストの再フォーマット、Amazon Comprehend を使用したエンティティの抽出、Lambda 関数を使用したデータの正規化など)。

タスク 1.4: ベクトルストアソリューションを設計して実装する。

スキル 1.4.1: FM 拡張に特化した高度なベクトルデータベースアーキテクチャを作成し、従来の検索機能を超える効率的なセマンティック検索を有効にする (Amazon Bedrock ナレッジベースを使用した階層組織、Amazon OpenSearch Service と Amazon Bedrock 統合用 Neural プラグインを使用したトピックベースのセグメンテーション、Amazon RDS と Amazon S3 ドキュメントリポジトリ、Amazon DynamoDB とベクトルデータベースを使用したメタデータと埋め込みなど)。

スキル 1.4.2: 包括的なメタデータフレームワークを開発して、検索の適合率と FM インタラクションのコンテキスト認識を向上させる (S3 オブジェクトメタデータを使用したドキュメントのタイムスタンプ、カスタム属性を使用した著者情報、タグシステムを使用したドメイン分類など)。

スキル 1.4.3: 高パフォーマンスのベクトルデータベースアーキテクチャを実装して、FM 検索のセマンティック検索パフォーマンスを大規模に最適化する (OpenSearch シャーディング戦略、ドメインに特化したマルチインデックスアプローチ、階層型インデックス作成手法の使用など)。

スキル 1.4.4: AWS のサービスを使用して、リソースに接続するための統合コンポーネントを作成する (ドキュメント管理システム、ナレッジベース、GenAI アプリケーションの包括的なデータ統合のための社内 Wiki など)。

スキル 1.4.5: データ管理システムを設計してデプロイし、ベクトルストアに FM 拡張のための正確な最新情報が含まれるようにする (増分更新メカニズム、リアルタイム変更検出システム、自動同期ワークフロー、定期更新パイプラインの使用など)。

タスク 1.5: FM 拡張のための検索メカニズムを設計する。

スキル 1.5.1: 効果的なドキュメントセグメンテーションアプローチを開発し、FM コンテキスト拡張の検索パフォーマンスを最適化する (Amazon Bedrock チャンク機能、固定サイズのチャンクを実装する Lambda 関数、コンテンツ構造に基づく階層型チャンク用のカスタム処理の使用など)。

スキル 1.5.2: 最適な埋め込みソリューションを選択して設定し、セマンティック検索のための効率的なベクトル表現を作成する (次元とドメインの適合性に基づく Amazon Titan 埋め込みの使用、Amazon Bedrock 埋め込みモデルのパフォーマンス特性の評価、埋め込みをバッチ生成する Lambda 関数の使用など)。

スキル 1.5.3: ベクトル検索ソリューションをデプロイして設定し、FM 拡張のセマンティック検索機能を有効にする (ベクトル検索機能を備えた **OpenSearch Service**、pgvector 拡張を備えた **Amazon Aurora**、マネージドベクトルストア機能を備えた **Amazon Bedrock ナレッジベース**の使用など)。

スキル 1.5.4: 高度な検索アーキテクチャを作成し、FM コンテキストに対して取得された情報の関連性と正解率を向上させる (**OpenSearch** を使用したセマンティック検索、キーワードとベクトルを組み合わせたハイブリッド検索、**Amazon Bedrock reranker** モデルの使用など)。

スキル 1.5.5: 高度なクエリ処理システムを開発し、FM 拡張の検索効率と結果品質を向上させる (**Amazon Bedrock** を使用したクエリ拡張、**Lambda** 関数を使用したクエリ分解、**Step Functions** を使用したクエリ変換など)。

スキル 1.5.6: 一貫性のあるアクセスメカニズムを作成し、FM とのシームレスな統合を有効にする [関数呼び出しインターフェイスを使用したベクトル検索、モデルコンテキストプロトコル (**MCP**) クライアントを使用したベクトルクエリ、標準化された **API** パターンを使用した検索拡張など]。

タスク 1.6: FM インタラクションのためのプロンプトエンジニアリング戦略とガバナンスを実施する。

スキル 1.6.1: 効果的なモデル指示フレームワークを作成し、FM の挙動と出力を制御する (**Amazon Bedrock Prompt Management** を使用したロール定義の適用、**Amazon Bedrock** ガードレールを使用した責任ある AI ガイドラインの適用、テンプレート設定を使用した応答のフォーマットなど)。

スキル 1.6.2: インタラクティブな AI システムを構築し、コンテキストを維持して FM とのユーザーインタラクションを改善する (**Step Functions** を使用した明確化ワークフロー、**Amazon Comprehend** を使用した意図認識、**DynamoDB** を使用した会話履歴ストレージなど)。

スキル 1.6.3: 包括的なプロンプトマネジメントとガバナンスシステムを実装し、FM 運用の一貫性と監視を確保する (**Amazon Bedrock Prompt Management** を使用したパラメータ化テンプレートと承認ワークフローの作成、**Amazon S3** を使用したテンプレートリポジトリの保存、**AWS CloudTrail** を使用した利用状況の追跡、**Amazon CloudWatch Logs** を使用したアクセスのログ記録など)。

スキル 1.6.4: 品質保証システムを開発し、FM のプロンプトの有効性と信頼性を確保する (**Lambda** 関数を使用した期待される出力の検証、**Step Functions** を使用したエッジケースのテスト、**CloudWatch** を使用したプロンプト回帰のテストなど)。

スキル 1.6.5: FM のパフォーマンスを強化して、基本のプロンプト手法を超えてプロンプトを繰り返し改良し、応答品質を改善する (構造化された入力コンポーネント、出力フォーマットの仕様、思考連鎖指示パターン、フィードバックループの使用など)。

スキル 1.6.6: FM で高度なタスクを処理する複雑なプロンプトシステムを設計する (Amazon Bedrock Prompt Flows を使用した順次プロンプトチェーン、モデル応答に基づく条件分岐、再利用可能なプロンプトコンポーネント、統合された前処理と後処理のステップなど)。

コンテンツ分野 2: 実装と統合

タスク 2.1: エージェント AI ソリューションとツール統合を実装する。

スキル 2.1.1: 適切なメモリとステート管理機能を備えたインテリジェントな自律システムを開発する (Strands Agents と AWS Agent Squad を使用したマルチエージェントシステム、MCP を使用したエージェントとツールのインタラクションなど)。

スキル 2.1.2: 構造化された推論ステップに従って FM が複雑な問題を分解して解決できるようにする高度な問題解決システムを作成する (Step Functions を使用した ReAct パターンや思考連鎖推論アプローチの実装など)。

スキル 2.1.3: FM の挙動の制御を確実にするために、保護された AI ワークフローを開発する (Step Functions を使用した停止条件の実装、Lambda 関数を使用したタイムアウトメカニズムの実装、IAM ポリシーを使用したリソース境界の適用、回路ブレーカーを使用した障害の軽減など)。

スキル 2.1.4: 複数の機能にわたってパフォーマンスを最適化する高度なモデル調整システムを作成する (特化型 FM を使用した複雑なタスクの実行、モデルアンサンブルのためのカスタム集計ロジック、モデル選択フレームワークの使用など)。

スキル 2.1.5: 人間の専門知識を活用して FM 機能を強化するコラボレーション AI システムを開発する (Step Functions を使用したレビューと承認プロセスのオーケストレーション、API Gateway を使用したフィードバック収集メカニズムの実装、人間の拡張パターンの使用など)。

スキル 2.1.6: FM 機能を拡張し、信頼性の高いツール運用を実現するためのインテリジェントツール統合を実装する (Strands API を使用したカスタム挙動の実装、標準化された関数定義、Lambda 関数を使用したエラー処理とパラメータ検証の実装など)。

スキル 2.1.7: FM 機能を強化するモデル拡張フレームワークを開発する (Lambda 関数を使用した、軽量なツールアクセスを提供するステートレス MCP サーバーの実装、Amazon ECS を使用した、複雑なツールを提供する MCP サーバーの実装、MCP クライアントライブラリを使用した一貫性のあるアクセスパターンの確保など)。

タスク 2.2: モデルデプロイ戦略を実施する。

スキル 2.2.1: 特定のアプリケーションニーズとパフォーマンス要件に基づいて FM をデプロイする (Lambda 関数を使用したオンデマンド呼び出し、Amazon Bedrock プロビジョンドスループット設定、SageMaker AI エンドポイントを使用したハイブリッドソリューションの実装など)。

スキル 2.2.2: 従来の ML デプロイとは異なる大規模言語モデル (LLM) 特有の課題に対処することで FM ソリューションをデプロイする (メモリ要件、GPU 使用率、トークン処理容量について最適化されたコンテナベースのデプロイパターンの実装、特化型のモデルロード戦略の遵守など)。

スキル 2.2.3: GenAI ワークロードのパフォーマンスとリソース要件のバランスを取るように最適化された FM デプロイアプローチを策定する (適切なモデルの選択、特定のタスク向けに事前トレーニングされた小規模なモデルの使用、ルーチンクエリ実行のための API ベースのモデルカスケードの使用など)。

タスク 2.3: エンタープライズ統合アーキテクチャを設計し実装する。

スキル 2.3.1: FM 機能を既存のエンタープライズ環境にシームレスに組み込むエンタープライズ接続ソリューションを作成する (レガシーシステムとの API ベースの統合、疎結合を実装するイベント駆動型アーキテクチャ、データ同期パターンの使用など)。

スキル 2.3.2: 既存のアプリケーションを GenAI 機能で強化する統合 AI 機能を開発する (API Gateway を使用したマイクロサービス統合の実装、Lambda 関数を使用したウェブフックハンドラー、Amazon EventBridge を使用したイベント駆動型統合の実装など)。

スキル 2.3.3: 適切なセキュリティコントロールを確保するセキュアなアクセスフレームワークを作成する (FM サービスとエンタープライズシステム間の ID フェデレーション、モデルおよびデータアクセスに対するロールベースのアクセスコントロール、FM への最小権限の API アクセスの使用など)。

スキル 2.3.4: FM へのアクセスを有効にするとともに管轄区域を横断したデータコンプライアンスを確保する環境横断型 AI ソリューションを開発する (AWS Outposts を使用したオンプレミスデータ統合、AWS Wavelength を使用したエッジデプロイの実行、クラウドとオンプレミスのリソース間のセキュアなルーティングの使用など)。

スキル 2.3.5: エンタープライズ環境にセキュアで規制に準拠した消費パターンを実装するための CI/CD パイプラインと GenAI ゲートウェイアーキテクチャを実装する (AWS CodePipeline、AWS CodeBuild、GenAI コンポーネントの継続的デプロイとテストのためのセキュリティスキャンとロールバックをサポートする自動テストフレームワーク、一元化された抽象化レイヤー、オブザーバビリティと制御メカニズムの使用など)。

タスク 2.4: FM API 統合を実装する。

スキル 2.4.1: 柔軟なモデルインタラクションシステムを作成する (Amazon Bedrock API を使用した、さまざまなコンピューティング環境からの同期リクエストの管理、言語固有の AWS SDK と Amazon SQS を使用した非同期処理、API Gateway を使用したカスタム API クライアントへのリクエスト検証の提供など)。

スキル 2.4.2: FM から即時フィードバックを提供するリアルタイム AI インタラクションシステムを開発する (Amazon Bedrock ストリーミング API を使用した増分応答配信、WebSocket またはサーバー送信イベントを使用したテキストのリアルタイム生成、API Gateway を使用したチャンク転送エンコーディングの実装など)。

スキル 2.4.3: 信頼性の高い運用を確保するための耐障害性の高い FM システムを作成する (AWS SDK を使用したエクスポネンシャルバックオフ、API Gateway を使用したレート制限の管理、グレースフルデグラデーションのためのフォールバックメカニズム、AWS X-Ray を使用したサービス境界をまたぐオブザーバビリティの提供など)。

スキル 2.4.4: モデル選択を最適化するインテリジェントなモデルルーティングシステムを開発する (アプリケーションコードを使用した静的ルーティング構成の実装、Step Functions を使用した特化型 FM への動的コンテンツベースのルーティング、メトリクスに基づいたインテリジェントなモデルルーティング、API Gateway とリクエスト変換を使用したルーティングロジック)。

タスク 2.5: アプリケーション統合パターンと開発ツールを実装する。

スキル 2.5.1: GenAI ワークロードの特定の要件に対応する **FM API** インターフェイスを作成する (**API Gateway** を使用したストリーミング応答への対処、トークンの制限管理、再試行戦略を使用したモデルタイムアウトへの対処など)。

スキル 2.5.2: FM の採用と統合を促進するために、アクセシブルな **AI** インターフェイスを開発する (**AWS Amplify** を使用した宣言型 **UI** コンポーネントの開発、**OpenAPI** 仕様を使用した **API** ファースト開発アプローチ、**Amazon Bedrock Prompt Flows** を使用したノーコードワークフロービルダーなど)。

スキル 2.5.3: ビジネスシステムの機能強化を作成する [**Lambda** 関数を使用した顧客関係管理 (**CRM**) の機能強化の実装、**Step Functions** を使用したドキュメント処理システムのオーケストレーション、**Amazon Q Business** データソースを使用した社内ナレッジツールの提供、**Amazon Bedrock Data Automation** を使用した自動データ処理ワークフローの管理など]。

スキル 2.5.4: GenAI アプリケーションの開発ワークフローを加速させるためにデベロッパーの生産性を高める (**Amazon Q Developer** を使用したコードの生成とリファクタリング、**API** アシスタンスのコード提案、**AI** コンポーネントテスト、パフォーマンスの最適化の使用など)。

スキル 2.5.5: 高度な **AI** 機能を実装する高度な GenAI アプリケーションを開発する (**Strands Agents** と **AWS Agent Squad** を使用した **AWS** ネイティブオーケストレーション、**Step Functions** を使用したエージェント設計パターンのオーケストレーション、**Amazon Bedrock** を使用したプロンプトチェーンパターンの管理など)。

スキル 2.5.6: FM アプリケーションのトラブルシューティング効率を向上させる (**CloudWatch Logs Insights** を使用したプロンプトと応答の分析、**X-Ray** を使用した **FM API** コールのトレース、**Amazon Q Developer** を使用した GenAI 固有のエラーパターン認識の実装など)。

コンテンツ分野 3: AI の安全性、セキュリティ、ガバナンス

タスク 3.1: 入力と出力の安全コントロールを実装する。

スキル 3.1.1: FM への有害なユーザー入力を防ぐ包括的なコンテンツ安全システムを開発する (**Amazon Bedrock** ガードレールを使用したコンテンツのフィルタリング、**Step Functions** と **Lambda** 関数を使用したカスタムモデレーションワークフローの実装、リアルタイム検証メカニズムの使用など)。

スキル 3.1.2: 有害な出力を防ぐコンテンツ安全フレームワークを作成する (**Amazon Bedrock** ガードレールを使用した応答のフィルタリング、コンテンツモデレーション

と毒性検出に特化した FM 評価、Text-to-SQL 変換を使用した決定論的な結果の確保など)。

スキル 3.1.3: FM 応答のハルシネーションを低減するために、正解率検証システムを開発する (Amazon Bedrock ナレッジベースを使用した根拠のある応答の生成とファクトチェックの実行、信頼度スコアリングとセマンティック類似性検索、JSON スキーマを使用した構造化出力の提供など)。

スキル 3.1.4: FM の誤用に対する包括的な保護を提供する多層防御安全システムを作成する (Amazon Comprehend を使用した前処理フィルターの開発、Amazon Bedrock を使用したモデルベースのガードレールの実装、Lambda 関数を使用した後処理検証の実行、API Gateway を使用した API 応答フィルタリングの実装など)。

スキル 3.1.5: 敵対的な入力やセキュリティの脆弱性から保護する高度な脅威検出を実装する (プロンプトインジェクションとジェイルブレイクの検出メカニズム、入力サニタイズとコンテンツフィルター、安全分類ツール、自動化された敵対的テストワークフローの使用など)。

タスク 3.2: データセキュリティとプライバシーコントロールを実装する。

スキル 3.2.1: FM デプロイの包括的なセキュリティを確保する、保護された AI 環境を開発する (VPC エンドポイントを使用したネットワーク分離、IAM ポリシーを使用したセキュアなデータアクセスパターンの適用、AWS Lake Formation を使用したきめ細かいデータアクセスの提供、CloudWatch を使用したデータアクセスのモニタリングなど)。

スキル 3.2.2: FM インタラクション中に機密情報を保護するプライバシー保護システムを開発する [Amazon Comprehend と Amazon Macie を使用した個人を特定できる情報 (PII) の検出、Amazon Bedrock のネイティブデータプライバシー機能の使用、Amazon Bedrock ガードレールを使用した出力のフィルタリング、Amazon S3 ライフサイクル設定を使用したデータ保持ポリシーの実装など]。

スキル 3.2.3: FM の有用性と有効性を維持しながらユーザーのプライバシーを保護する、プライバシー保護に重点を置いた AI システムを作成する (データマスキング手法、Amazon Comprehend の PII 検出、機密情報の匿名化戦略、Amazon Bedrock ガードレールの使用など)。

タスク 3.3: AI ガバナンスとコンプライアンスのメカニズムを実装する。

スキル 3.3.1: FM デプロイの規制コンプライアンスを確保するコンプライアンスフレームワークを開発する (SageMaker AI を使用したプログラムによるモデルカード開発、AWS Glue を使用したデータリネージュの自動追跡、体系的なデータ

ソース帰属のためのメタデータのタグ付け、**CloudWatch Logs** を使用した包括的な意思決定ログの収集など)。

スキル 3.3.2: **GenAI** アプリケーションのトレーサビリティを維持するデータソース追跡を実装する (**AWS Glue Data Catalog** を使用したデータソース登録、**FM** 生成コンテンツのソース帰属のためのメタデータのタグ付け、**CloudTrail** を使用した監査ロギングなど)。

スキル 3.3.3: **FM** の実装を一貫して監視するために、組織ガバナンスシステムを構築する (組織のポリシー、規制要件、責任ある **AI** の原則に沿った包括的なフレームワークの使用など)。

スキル 3.3.4: 安全監査と規制への準備をサポートする継続的なモニタリングと高度なガバナンスコントロールを実装する (誤用、ドリフト、ポリシー違反の自動検出、バイアスドリフトモニタリング、自動アラートと修復ワークフロー、トークンレベルのリダクション、応答ロギング、**AI** 出力ポリシーフィルターの使用など)。

タスク 3.4: 責任ある **AI** の原則を実装する。

スキル 3.4.1: **FM** 出力での透明性の高い **AI** システムを開発する (推論表示を使用したユーザー向け説明の提供、**CloudWatch** を使用した信頼度メトリクスの収集と不確実性の数値化、ソース帰属のエビデンス提示、**Amazon Bedrock** エージェントトレースを使用した推論トレースの提供など)。

スキル 3.4.2: バイアスのない **FM** 出力を確保する公平性評価を適用する (**CloudWatch**、**Amazon Bedrock Prompt Management**、**Amazon Bedrock Prompt Flows** で事前定義された公平性メトリクスを使用した体系的な **A/B** テストの実行、**Amazon Bedrock** で **LLM-as-a-judge** ソリューションを使用した自動モデル評価の実行など)。

スキル 3.4.3: 責任ある **AI** プラクティスを遵守するためのポリシー準拠 **AI** システムを開発する (ポリシー要件に基づく **Amazon Bedrock** ガードレールの使用、モデルカードを使用した **FM** の制限のドキュメント化、**Lambda** 関数を使用した自動コンプライアンスチェックの実行など)。

コンテンツ分野 4: **GenAI** アプリケーションの運用効率と最適化

タスク 4.1: コスト最適化とリソース効率化の戦略を実施する。

スキル 4.1.1: 有効性を維持しながら **FM** コストを削減するトークン効率システムを開発する (トークンの推定と追跡、コンテキストウィンドウの最適化、応答サイズ

のコントロール、プロンプト圧縮、コンテキストプルーニング、応答制限の使用など)。

スキル 4.1.2: コスト効率の高いモデル選択フレームワークを作成する (コスト能力のトレードオフ評価、クエリの複雑さに基づく階層化された **FM**、応答品質に対する推論コストバランシング、価格対性能比の測定、効率的な推論パターンの使用など)。

スキル 4.1.3: **GenAI** ワークロードのリソース利用とスループットを最大化する高パフォーマンス **FM** システムを開発する (バッチ戦略、容量プランニング、使用率モニタリング、**Auto Scaling** 構成、プロビジョンドスループット最適化の使用など)。

スキル 4.1.4: 不必要な **FM** 呼び出しを回避することでコストを削減して応答時間を改善するインテリジェントなキャッシュシステムを作成する (セマンティックキャッシュ、結果フィンガープリンティング、エッジキャッシュ、決定論的リクエストハッシュ、プロンプトキャッシュの使用など)。

タスク 4.2: アプリケーションのパフォーマンスを最適化する。

スキル 4.2.1: レイテンシーとコストのトレードオフに対処し **FM** に関するユーザーエクスペリエンスを向上させる即応型 **AI** システムを作成する (事前計算を使用した予測可能なクエリの実行、時間的制約のあるアプリケーション向けのレイテンシーが最適化された **Amazon Bedrock** モデル、複雑なワークフローに対する並列リクエスト、応答ストリーミング、パフォーマンスベンチマークの使用など)。

スキル 4.2.2: **FM** コンテキスト拡張の検索対象情報の関連性と処理速度を向上させるように検索パフォーマンスを強化する (インデックス最適化、クエリの前処理、カスタムスコアリングによるハイブリッド検索実装の使用など)。

スキル 4.2.3: **GenAI** ワークロード固有のスループットの課題に対処するように **FM** スループット最適化を実装する (トークン処理最適化、バッチ推論戦略、同時モデル呼び出し管理の使用など)。

スキル 4.2.4: 特定の **GenAI** ユースケースで最適な結果が得られるように **FM** パフォーマンスを強化する (モデル固有のパラメータ構成の使用、**A/B** テストを使用した改善点の評価、要件に基づく適切な温度と **top-k/top-p** 選択の使用など)。

スキル 4.2.5: FM ワークロードに特化した効率的なリソース割り当てシステムを作成する (トークン処理要件に対する容量プランニング、プロンプトパターンと完成パターンに対する使用率モニタリング、GenAI トラフィックパターン向けに最適化された **Auto Scaling** 構成の使用など)。

スキル 4.2.6: GenAI ワークフロー向けに FM システムのパフォーマンスを最適化する (API コールプロファイリングを使用したプロンプト-完成パターン、ベクトルデータベースクエリ最適化を使用した検索拡張、LLM 推論固有のレイテンシー削減手法、効率的なサービスコミュニケーションパターンの使用など)。

タスク 4.3: GenAI アプリケーションのモニタリングシステムを実装する。

スキル 4.3.1: FM アプリケーションのパフォーマンスを完全に可視化する包括的なオブザーバビリティシステムを作成する (運用メトリクス、パフォーマンストレース、FM インタラクショントレース、カスタムダッシュボードでのビジネスインパクトメトリクスの使用など)。

スキル 4.3.2: 問題を事前に特定し FM 実装に固有の重要業績評価指標を評価するための包括的な GenAI モニタリングシステムを実装する (CloudWatch を使用したトークンの使用状況、プロンプトの有効性、ハルシネーション率、応答品質・トークンのバーストパターンと応答ドリフトの異常検出・Amazon Bedrock のモデル呼び出しログを使用した詳細なリクエストおよび応答分析の実行・パフォーマンスベンチマーク・コスト異常検出の使用など)。

スキル 4.3.3: FM アプリケーションに実用的なインサイトを提供する統合オブザーバビリティソリューションを開発する (運用メトリクスダッシュボード、ビジネスインパクト視覚化、コンプライアンスモニタリング、フォレンジックトレーサビリティと監査ロギング、ユーザーインタラクション追跡、モデル挙動パターン追跡の使用など)。

スキル 4.3.4: FM に最適なツール運用と利用を確保するために、ツールパフォーマンスフレームワークを作成する (コールパターン追跡、パフォーマンスメトリクス収集、ツールコールのオブザーバビリティ、マルチエージェントコーディネーション追跡、異常検知のための利用状況ベースラインの使用など)。

スキル 4.3.5: ベクトルストアの最適な運用と FM 拡張の信頼性を確保するベクトルストア運用管理システムを作成する (ベクトルデータベースのパフォーマンスモニタリング、自動インデックス最適化ルーチン、データ品質検証プロセスの使用など)。

スキル 4.3.6: 従来の ML システムにない **GenAI** 固有の障害モードを特定する **FM** 固有のトラブルシューティングフレームワークを開発する (ゴールデンデータセットを使用したハルシネーションの検出、出力差分手法を使用した応答一貫性分析の実施、推論パストレースを使用した論理エラーの特定、特化型のオブザーバビリティパイプラインの使用など)。

コンテンツ分野 5: テスト、検証、トラブルシューティング

タスク 5.1: **GenAI** の評価システムを実装する。

スキル 5.1.1: 従来の ML 評価アプローチを超えて **FM** 出力の品質と有効性を評価する包括的な評価フレームワークを開発する (関連性、事実に関する正解率、一貫性、流暢さのメトリクスの使用など)。

スキル 5.1.2: 最適な構成を特定するための体系的なモデル評価システムを作成する (**Amazon Bedrock** モデル評価、**FM** の **A/B** テストと **Canary** テスト、マルチモデル評価、トークンの効率、レイテンシーと品質の比率、ビジネス成果を測定するためのコストパフォーマンス分析の使用など)。

スキル 5.1.3: ユーザーエクスペリエンスに基づいて **FM** パフォーマンスを継続的に改善するためのユーザー中心の評価メカニズムを開発する (フィードバックインターフェイス、モデル出力の評価システム、応答品質を評価するアノテーションワークフローの使用など)。

スキル 5.1.4: **FM** の一貫したパフォーマンス基準を維持するための体系的な品質保証プロセスを作成する (継続的評価のワークフロー、モデル出力の回帰テスト、デプロイの自動品質ゲートの使用など)。

スキル 5.1.5: **FM** 出力をさまざまな視点から徹底的に評価する包括的な評価システムを開発する (**RAG** 評価、**LLM-as-a-judge** 手法による自動品質評価、人間のフィードバック収集インターフェイスの使用など)。

スキル 5.1.6: **FM** 拡張のための情報検索コンポーネントを評価および最適化するための検索品質テストを実装する (関連性スコアリング、コンテキストマッチング検証、検索レイテンシー測定の使用など)。

スキル 5.1.7: エージェントがタスクを正確かつ効率的に実行できるようにするためのエージェントパフォーマンスフレームワークを開発する (タスク完了率の測定、ツールの使用効率の評価、**Amazon Bedrock** エージェントの評価、マルチステップワークフローにおける推論品質評価の使用など)。

スキル 5.1.8: FM 実装のパフォーマンスメトリクスとインサイトをステークホルダーに効果的に伝えるための包括的なレポートシステムを作成する (視覚化ツール、自動レポートメカニズム、モデル比較の視覚化の使用など)。

スキル 5.1.9: FM 更新中に信頼性を維持するためのデプロイ検証システムを作成する (合成ユーザーワークフロー、ハルシネーション率とセマンティックドリフトに関する AI 固有の出力の検証、応答の一貫性を確保するための自動品質チェックの使用など)。

タスク 5.2: GenAI アプリケーションをトラブルシューティングする。

スキル 5.2.1: 必要な情報が FM インタラクションで完全に処理されるようにするためにコンテンツ処理の問題を解決する (コンテキストウィンドウのオーバーフロー診断、動的チャンク戦略、プロンプト設計の最適化、切り捨て関連のエラー分析の使用など)。

スキル 5.2.2: FM 統合の問題を診断して解決し、GenAI サービスに固有の API 統合の問題を特定して修正する (エラーロギング、リクエスト検証、応答分析の使用など)。

スキル 5.2.3: 基本的なプロンプト調整を超えて FM 応答の品質と一貫性を向上させるために、プロンプトエンジニアリングの問題をトラブルシューティングする (プロンプトテストフレームワーク、バージョン比較、体系的な改良の使用など)。

スキル 5.2.4: FM 拡張のための情報検索の有効性に影響する問題を特定して解決するために、検索システムの問題をトラブルシューティングする (モデル応答関連性分析、埋め込み品質診断、ドリフトモニタリング、ベクトル化問題の解決、チャンク化と前処理の修復、ベクトル検索パフォーマンスの最適化の使用など)。

スキル 5.2.5: FM インタラクションのパフォーマンスを継続的に改善するために、プロンプトメンテナンスの問題をトラブルシューティングする (テンプレートテストと CloudWatch Logs を使用したプロンプト混乱の診断、X-Ray を使用したプロンプトオブザーバビリティパイプラインの実装、スキーマ検証を使用したフォーマットの不一致の検出、体系的なプロンプト改良ワークフローの使用など)。

付録

試験に出題される可能性のあるテクノロジーと概念

以下は、試験に出題される可能性のあるテクノロジーと概念のリストです。このリストはすべてを網羅しているわけではなく、また、変更される場合があります。このリストにおける項目の掲載順序や配置は、その項目の相対的な重みや試験における重要性を示すものではありません。

- 検索拡張生成 (RAG)
- ベクトルデータベースと埋め込み
- プロンプトエンジニアリングと管理
- 基盤モデル (FM) の統合
- エージェント AI システム
- 責任ある AI のプラクティス
- コンテンツの安全性とモデレーション
- モデル評価と検証
- AI ワークロードのコスト最適化
- AI アプリケーションのパフォーマンスチューニング
- AI システムのモニタリングとオプザーバビリティ
- AI アプリケーションのセキュリティとガバナンス
- API 設計および統合パターン
- イベント駆動型アーキテクチャ
- サーバーレスコンピューティング
- コンテナオーケストレーション
- Infrastructure as Code (IaC)
- AI アプリケーションの CI/CD
- ハイブリッドクラウドアーキテクチャ
- エンタープライズシステム統合

試験での AWS サービスへの言及

AWS 認定では、略語や括弧付きの情報を含む周知の AWS サービス名に正式な短縮名を使用して、この試験を容易に読めるようにしています。例えば、試験では、**Amazon Simple Notification Service (Amazon SNS)** が **Amazon SNS** と記載されます。

試験のヘルプ機能 (すべての問題に用意されています) には、**AWS** サービスの短縮名と、それに対応する正式名のリストが含まれています。

試験に短縮名で表示されるサービスのリストについては、**AWS** 認定ウェブサイトの [AWS サービス名](#) を参照してください。リストに含まれているサービスのうち、試験対象外のものは試験に記載されていません。

注: 略語には、試験に正式名が記述されていないものや、ヘルプ機能に含まれていないものもあります。一部の **AWS** サービスの正式名 (**Amazon API Gateway**、**Amazon EMR** など) には、常に略語のまま記載される略語が含まれています。また、試験には、受験者が知っていると想定される他の略語が含まれる場合もあります。

試験対象の **AWS** のサービスと機能

以下に、試験対象の **AWS** のサービスと機能のリストを示します。このリストはすべてを網羅しているわけではなく、また、変更される場合があります。各 **AWS** のサービスは、サービスの主な機能に応じたカテゴリに分けられています。

分析:

- Amazon Athena
- Amazon EMR
- AWS Glue
- Amazon Kinesis
- Amazon OpenSearch Service
- Amazon QuickSight
- Amazon Managed Streaming for Apache Kafka (Amazon MSK)

アプリケーション統合:

- Amazon AppFlow
- AWS AppConfig
- Amazon EventBridge
- Amazon SNS
- Amazon SQS
- AWS Step Functions

コンピューティング:

- AWS App Runner
- Amazon EC2
- AWS Lambda
- AWS Lambda@Edge
- AWS Outposts
- AWS Wavelength

コンテナ:

- Amazon ECR
- Amazon ECS
- Amazon EKS
- AWS Fargate

カスタマーエンゲージメント:

- Amazon Connect

データベース:

- Amazon Aurora
- Amazon DocumentDB
- Amazon DynamoDB
- Amazon DynamoDB Streams
- Amazon ElastiCache
- Amazon Neptune
- Amazon RDS

デベロッパーツール:

- AWS Amplify
- AWS CDK
- AWS CLI
- AWS CloudFormation
- AWS CodeArtifact
- AWS CodeBuild
- AWS CodeDeploy

- AWS CodePipeline
- AWS のツールと SDK
- AWS X-Ray

機械学習:

- Amazon Augmented AI
- Amazon Bedrock
- Amazon Bedrock AgentCore
- Amazon Bedrock ナレッジベース
- Amazon Bedrock Prompt Management
- Amazon Bedrock Prompt Flows
- Amazon Comprehend
- Amazon Kendra
- Amazon Lex
- Amazon Q Business
- Amazon Q Business Apps
- Amazon Q Developer
- Amazon Rekognition
- Amazon SageMaker AI
- Amazon SageMaker Clarify
- Amazon SageMaker Data Wrangler
- Amazon SageMaker Ground Truth
- Amazon SageMaker JumpStart
- Amazon SageMaker Model Monitor
- Amazon SageMaker Model Registry
- Amazon SageMaker Neo
- Amazon SageMaker Processing
- Amazon SageMaker Unified Studio
- Amazon Textract
- Amazon Titan
- Amazon Transcribe

マネジメントとガバナンス:

- AWS Auto Scaling
- AWS Chatbot

- AWS CloudTrail
- Amazon CloudWatch
- Amazon CloudWatch Logs
- Amazon CloudWatch Synthetics
- AWS コスト異常検出
- AWS Cost Explorer
- Amazon Managed Grafana
- AWS Service Catalog
- AWS Systems Manager
- AWS Well-Architected Tool

移行と転送:

- AWS DataSync
- AWS Transfer Family

ネットワークとコンテンツ配信:

- Amazon API Gateway
- AWS AppSync
- Amazon CloudFront
- Elastic Load Balancing (ELB)
- AWS Global Accelerator
- AWS PrivateLink
- Amazon Route 53
- Amazon VPC

セキュリティ、アイデンティティ、コンプライアンス:

- Amazon Cognito
- AWS Encryption SDK
- IAM
- IAM Access Analyzer
- IAM アイデンティティセンター
- AWS KMS
- Amazon Macie
- AWS Secrets Manager
- AWS WAF

ストレージ:

- Amazon EBS
- Amazon EFS
- Amazon S3
- Amazon S3 Intelligent-Tiering
- Amazon S3 ライフサイクルポリシー
- Amazon S3 クロスリージョンレプリケーション

試験対象外の **AWS** のサービスと機能

以下に、試験対象外の **AWS** のサービスと機能のリストを示します。このリストはすべてを網羅しているわけではなく、また、変更される場合があります。試験の対象となる職務内容にまったく関係のない **AWS** のサービスは、このリストから除外されています。

アプリケーション統合:

- Amazon MQ

分析:

- AWS Clean Rooms
- AWS Data Exchange
- Amazon DataZone
- Amazon FinSpace

ブロックチェーン:

- Amazon Managed Blockchain (AMB)

ビジネスアプリケーション:

- Alexa for Business
- Amazon Chime
- AWS Wickr
- Amazon WorkDocs
- Amazon WorkMail

クラウド財務管理:

- AWS Budgets
- AWS Cost and Usage Report
- リザーブドインスタンスレポート
- AWS Savings Plans

コンピューティング:

- AWS Batch
- Amazon EC2 Image Builder
- Amazon ECS Anywhere
- Amazon EKS Anywhere
- AWS Elastic Beanstalk
- Amazon Lightsail
- AWS Local Zones
- AWS Serverless Application Repository

コンテナ:

- AWS App2Container
- AWS Copilot
- Red Hat OpenShift Service on AWS (ROSA)

カスタマーエンゲージメント:

- Amazon SES

データベース:

- Amazon Keyspaces
- Amazon Quantum Ledger Database (Amazon QLDB)
- Amazon Redshift
- Amazon Timestream

デベロッパーツール:

- AWS Cloud9
- AWS CloudShell
- Amazon CodeGuru

- AWS CodeStar
- Amazon Corretto

エンドユーザーコンピューティング:

- Amazon AppStream 2.0
- Amazon WorkLink
- Amazon WorkSpaces
- Amazon WorkSpaces Web

フロントエンドのウェブとモバイル:

- AWS Device Farm
- Amazon Location Service
- Amazon Pinpoint

ゲーム開発:

- Amazon GameLift
- Amazon Lumberyard

IoT:

- AWS IoT 1-Click
- AWS IoT Analytics
- AWS IoT ボタン
- AWS IoT Core
- AWS IoT Device Defender
- AWS IoT Device Management
- AWS IoT Events
- AWS IoT FleetWise
- AWS IoT Greengrass
- AWS IoT SiteWise
- AWS IoT TwinMaker

マネジメントとガバナンス:

- AWS コンソールモバイルアプリケーション
- AWS Health Dashboard
- AWS License Manager

- AWS Proton
- AWS Trusted Advisor

機械学習:

- AWS DeepComposer
- AWS DeepRacer
- Amazon DevOps Guru
- Amazon Forecast
- Amazon Fraud Detector
- Amazon HealthLake
- Amazon Lookout for Equipment
- Amazon Lookout for Metrics
- Amazon Lookout for Vision
- Amazon Monitron
- AWS Panorama

メディアサービス:

- Amazon Elastic Transcoder
- AWS Elemental MediaConnect
- AWS Elemental MediaConvert
- AWS Elemental MediaLive
- AWS Elemental MediaPackage
- AWS Elemental MediaStore
- AWS Elemental MediaTailor
- Amazon Interactive Video Service
- Amazon Kinesis Video Streams
- Amazon Nimble Studio

移行と転送:

- AWS Application Discovery Service
- AWS Application Migration Service
- CloudEndure Migration
- AWS Migration Hub
- AWS Snow Family

ネットワークとコンテンツ配信:

- AWS App Mesh
- AWS Cloud Map
- AWS Direct Connect
- AWS Private 5G
- AWS Transit Gateway
- AWS VPN

量子テクノロジー:

- Amazon Braket

ロボティクス:

- AWS RoboMaker

人工衛星:

- AWS Ground Station

セキュリティ、アイデンティティ、コンプライアンス:

- AWS Artifact
- AWS Audit Manager
- AWS Certificate Manager
- AWS CloudHSM
- Amazon Detective
- AWS Directory Service
- AWS Firewall Manager
- Amazon GuardDuty
- AWS Network Firewall

- AWS Private CA
- AWS Resource Access Manager (AWS RAM)
- AWS Security Hub
- AWS Shield
- Amazon Verified Permissions

ストレージ:

- AWS Backup
- Amazon FSx
- Amazon FSx for Lustre
- Amazon FSx for NetApp ONTAP
- Amazon FSx for OpenZFS
- Amazon FSx for Windows File Server
- Amazon S3 Glacier
- AWS Snow Family
- AWS Storage Gateway

アンケート

この試験ガイドはどの程度役に立ちましたか？ [アンケートへの回答](#)にご協力ください。