



プレスリリース

2024 年 4 月 24 日

アマゾン ウェブ サービス ジャパン合同会社

※本プレスリリースは、現地時間 2024 年 4 月 23 日に米国で発表された[プレスリリース](#)の抄訳版です。

AWS、安全な生成 AI アプリケーションを構築、拡張する基盤として 数万のお客様が活用する Amazon Bedrock の新機能を発表

**新機能 カスタムモデルインポートにより、
お客様自身のプロプライエタリーなモデルを Amazon Bedrock にインポートし、利用可能に**

**新機能 モデル評価により、RAG に最適化された Titan Embeddings V2 や
Cohere や Meta の最新モデルなど幅広いフルマネージドモデルを、より簡単かつ迅速に利用可能に**

**Guardrails for Amazon Bedrock により、お客様のアプリケーションの要件と
責任ある AI ポリシーに則した安全対策の効果的な展開を支援するテクノロジーを提供**

**addidas、ADP、Aha!、Amazon.com、Bridgewater Associates、Choice Hotels、Clariant、
Delta Aire Lines、Dentsu Group、FOX Corporation、GoDaddy、Hugging Face、Infor、Intuit、
Kone、KT Corporation、LexisNexis Legal & Professional、Lonely Planet、Netsmart、New
York Stock Exchange、Pfizer、PGA TOUR、Perplexity AI、Ricoh USA、Rocket Mortgage、
Ryanair、Salesforce、Siemens、Thomson Reuters、Toyota、Tui、United Airlines など、
数万のお客様やパートナーが Amazon Bedrock を活用して、生成 AI アプリケーションを構築・展開**

(シアトル — 2024 年 4 月 23 日) Amazon.com, Inc. (NASDAQ: AMZN) の関連会社である Amazon Web Services, Inc. (AWS) は 4 月 23 日、お客様が最も簡単、迅速、かつ、安全に高度な生成 AI アプリケーションを開発し、新たな体験を実現できるよう支援する Amazon Bedrock の新たなイノベーションを発表しました。Amazon Bedrock は、AI21 Labs、Anthropic、Cohere、Meta、Mistral AI、Stability AI、Amazon などが提供する幅広い主要な基盤モデルを、生成 AI アプリケーションを迅速に構築、展開するために必要な機能やエンタープライズグレードのセキュリティと合わせて提供するサービスで、数万に及ぶお客様の生成 AI 戦略の基盤として採用されています。Amazon Bedrock の強力な基盤モデルは、フルマネージドサービスとして提供されるため、お客様はインフラストラクチャを意識することなく、シームレスなデプロイ、拡張性、継続的な最適化を推進しながら、生成 AI アプリケーションを運用することができます。今回の発表は、お客様自身のフルマネージドモデルを Amazon Bedrock 上で稼働させ、自社の用途に最適なモデルを簡単に見つけ、生成 AI アプリケーションへの安全機構を導入できるよう支援し、さらに幅広い基盤モデルの選択肢を提供するものです。Amazon Bedrock の詳細については、aws.amazon.com/jp/bedrock をご覧ください。

世界で最も急速に成長しているスタートアップ企業から、高いセキュリティ要件を持つ企業や政府機関まで、あらゆる業界の組織が Amazon Bedrock を利用してイノベーションを加速し、生産性を向上させながら、新たなエンドユーザー体験を創出しています。ニューヨーク証券取引所 (NYSE) は、Amazon Bedrock が提供する幅広い基盤モデルの選択肢と最先端の生成 AI 機能を、数千ページもの規則類を処理し、わかりやすい言葉で回答するなどの複数のユースケースで活用しています。欧州最大の航空会社である Ryanair は、Amazon Bedrock を活用し、乗務員が国ごとの規制に関する質問への回答を迅速に見つけたり、マニュアルから要約を抽出できるようにし、乗客のスムーズな旅をサポートしていま



す。地域医療機関向け電子カルテの設計、構築、提供を専門とするテクノロジープロバイダーである Netsmart は、医療従事者の臨床文書作成体験の向上に取り組み、Amazon Bedrock 上に構築した生成 AI 自動化ツールを用いて、個人の健康記録の管理に費やす時間を最大 50%削減することを目指しています。これにより、Netsmart のサービスを利用する地域医療機関は、患者さんの診療報酬をより迅速に提出するとともに、患者さんのケアを向上させることができます。

AWS の AI・データ担当バイスプレジデントであるスワミ・シヴァスブラマニアン博士 (Swami Sivasubramanian) は、次のように述べています。「Amazon Bedrock は急速に成長しており、あらゆる業界・規模の数万もの組織が、他よりも迅速かつ簡単に検証段階から本番活用へと移行できることを評価し、生成 AI 戦略の基盤として活用しています。Amazon Bedrock はエンタープライズグレードのセキュリティとプライバシー、主要な基盤モデルの幅広い選択肢、生成 AI アプリケーションを構築する最も簡単な方法を提供するものとしてお客様に評価いただいています。私たちは今回の発表により、お客様に最も包括的な機能群と業界をけん引する主要な基盤モデルの選択肢を幅広く提供し、生成 AI によるイノベーションを大規模に民主化する取組をさらに強化することで、引き続きお客様のためのイノベーションを推進していきます」

新機能 カスタムモデルインポートにより、お客様自身のカスタマイズモデルが Amazon Bedrock で利用可能に。運用コストを削減し、アプリケーション開発の加速を支援

Amazon Bedrock では、AI21 Labs、Amazon、Anthropic、Cohere、Meta、Mistral AI、Stability AI などが提供する世界有数の強力な基盤モデルを活用できることに加え、医療や金融といった業界のお客様の間では、公開されたモデルを自社データを用いて専門領域用途向けにカスタマイズする例が増えています。お客様が自身の専有データを利用してこれらのモデルをカスタマイズする場合、優れた事前学習機能を提供する Amazon SageMaker を利用して、スクラッチでモデルの学習を行う、または、Llama、Mistral、Flan-T5 などの公開されたモデルの高度なカスタマイズを行うのが一般的です。Amazon SageMaker は 2017 年に発表され、これまでに一般公開されたなかでは最大規模のモデルである Falcon 180 などの世界中の高性能基盤モデルの構築や学習に活用されてきました。お客様はまた、あらゆる機能を自社で開発するのではなく、Amazon Bedrock で提供される Knowledge Bases、Guardrails for Amazon Bedrock、Agents for Amazon Bedrock、モデル評価 などの高度な内蔵の生成 AI ツールを自社のカスタマイズモデルでフルに活用したいと考えています。

Amazon Bedrock カスタムモデルインポートを利用することで、お客様は自社のカスタムモデルを Amazon Bedrock にインポートして、フルマネージドな API としてアクセスできるようになり、生成 AI アプリケーションの構築時にこれまでにない選択肢を得ることができます。Amazon SageMaker やその他のツールでカスタマイズしたモデルを、わずか数クリックで Amazon Bedrock に簡単に追加できます。自動での検証プロセスを経れば、Amazon Bedrock の他の基盤モデルと同様に自社のカスタムモデルにスムーズにアクセスでき、シームレスな拡張性、アプリケーションに対する強力な安全機構、責任ある AI の原則遵守、検索拡張生成 (retrieval augmented generation、RAG) によるモデルのナレッジベース拡張、複数ステップのタスクを完了するエージェントの簡単な作成、モデルの継続的な改良のためのファインチューニングの実行といった、現在と同様のメリットを、インフラストラクチャを管理する必要なくご利用いただけます。この新機能により、AWS はお客様が同じ API を介して Amazon Bedrock のモデルとお客様自身のカスタムモデルの組み合わせを簡単に選択できるよう支援します。現在、Amazon Bedrock カスタムモデルインポートはプレビュー版として提供しています。最も普及している Flan-T5、Llama、Mistral の 3 つのオープンモデルアーキテクチャーに対応しており、今後拡張していく予定です。

モデル評価により、お客様のアプリケーションに最適なモデルにアクセス、比較、採用できるよう支援

業界をけん引する基盤モデルを幅広く提供する Amazon Bedrock は、お客様の価格、性能、機能のあらゆる要件に対応し、モデルを単独、または他のモデルと組み合わせで稼働できるようにしています。



一方、特定の用途に最適なモデルを選択する際には、正確さと性能のバランスをとることが求められます。これまで、組織は新たなモデルが自社の用途に適しているのかを一つ一つ確認するために多くの時間を費やさなければならず、画期的な生成 AI 体験をエンドユーザーに迅速に提供するのは困難でした。今回、一般提供を開始したモデル評価により、Amazon Bedrock 上のモデルを迅速に分析、比較でき、モデルの評価にかかる時間を数週間から数時間に短縮します。これにより新たなアプリケーションや体験をより速やかに市場に投入できるようになります。事前に定義された評価基準（正確さ、堅牢性など）を選択して自社のデータセットやプロンプトライブラリーをアップロードしたり、組み込まれた公開リソースを選択することで、速やかなスタートが可能で、主観的な基準や微妙な判断が求められるコンテンツについては、ワークフローに人による判断を追加し、特定用途の指標（関連性、スタイル、ブランドボイスなど）に基づくモデルレスポンスを評価することが可能です。設定プロセスが完了すると、Amazon Bedrock は評価を実行してレポートを生成します。お客様は重要基準に対するモデルのパフォーマンスを把握し、用途に最適なモデルを迅速に選択することが可能となります。

Guardrails for Amazon Bedrock により、クラス最高のテクノロジーを利用して、個人情報や機密情報、冒涇表現や特定の単語を削除し、有害なコンテンツをブロックする保護機能を容易に実行

あらゆる業界で生成 AI が利用されるには、安全で、信頼でき、かつ責任ある方法で導入することが求められます。多くのモデルでは、あらかじめ組み込んだ制御機構を用いて、望ましくない有害なコンテンツのフィルタリングを行っています。しかし、大多数のお客様は、生成 AI アプリケーションをさらにカスタマイズすることで、関連性があり、企業ポリシーと合致しており、責任ある AI の原則に準拠した応答を返せるようにしたいと考えています。今回、一般提供を開始した Guardrails for Amazon Bedrock は、基盤モデルのネイティブ機能に加え、業界をリードする安全機構を提供することで、有害コンテンツを最大 85% ブロックすることを可能とし、現在提供されている他のモデルよりも優れたパフォーマンスを発揮します。Guardrails for Amazon Bedrock は、大手クラウドプロバイダーにより提供されている唯一のソリューションであり、組み込み型およびカスタムの安全機構を 1 つのサービスで提供するとともに、Amazon Bedrock のすべての大規模言語モデルやお客様がファインチューニングしたモデルと連動するようになっています。ガードレールを作成するには、お客様のアプリケーションのコンテキストにおいて許可されないトピックを、自然言語で定義すれば良いだけです。またヘイトスピーチや、人を傷つける言葉、性的表現、プロンプトインジェクション、脅しといったさまざまな領域をフィルタリングするためのしきい値のほか、個人情報や機密情報、冒涇表現、特定の禁止用語を削除するためのフィルターを設定することもできます。Guardrails for Amazon Bedrock は、一貫したユーザーエクスペリエンスを提供し、すべての生成 AI アプリケーションにおいて安全性とプライバシー統制を標準化することで、お客様が迅速かつ安全にイノベーションを推進できるよう支援します。

モデルの選択肢の拡大 : Amazon Titan Text Embeddings V2、Titan Image Generator の一般提供、Cohere と Meta の最新モデル

Amazon Bedrock のみでご利用いただける Amazon Titan モデルは、AWS が大規模で多様なデータセットでの事前学習を終えており、さまざまなユースケースに対応し、責任ある AI 利用のためのサポートが組み込まれています。現在、Amazon Bedrock では Amazon Titan ファミリーの拡充を続け、お客様にさらに多くの選択肢と柔軟性を提供しています。RAG での活用に最適化された Amazon Titan Text Embeddings V2 は、情報検索、質疑応答チャットボット、「おすすめ」のパーソナライズなど、さまざまなタスクに適しています。基盤モデルの応答を追加データによって強化するために、多くの組織は RAG を利用しています。RAG は、基盤モデルがナレッジソースに接続し、そのナレッジソースを参照することで応答を洗練させる一般的なモデルカスタマイズ手法です。しかし、これらの処理では、コンピューティングとストレージに多大な負荷がかかる可能性があります。来週から提供を開始する新しい Amazon Titan Text Embeddings V2 モデルは、精度を高めつつ、ストレージとコンピューティングのコストを削減します。お客様による柔軟な埋め込みが可能になることで、全体のストレージを最大 1/4 にまで減らし、運用コストを大幅に削減しながら、RAG ユースケースで 97% の精度を維持し、他の主要モデルよりも優れたパフォーマンスを発揮します。



また、今回一般提供を開始する Amazon Titan Image Generator により、広告、e コマース、メディア、エンターテインメントといった業界のお客様が、自然言語プロンプトを使って、スタジオ品質の画像を制作したり、既存の画像を低コストで強化・編集できるよう支援します。さらに、Amazon Titan は作成するすべての画像に目に見えない電子透かしを入れることで、AI で生成された画像の識別を助け、安全・確実で透明性のある AI 技術の開発を促進すると同時に、偽情報の拡散を抑制します。また、このモデルは電子透かしの有無をチェックすることもでき、Amazon Titan Image Generator で生成された画像かどうかを確認するのに役立ちます。

さらに本日より、基盤モデル Meta Llama 3 が Amazon Bedrock 上でご利用可能となりました。Cohere の Command R および Command R+モデルも近日、提供開始予定です。Llama 3 は、開発者、研究者、企業が生成 AI のアイデアを構築し、実験し、責任を持って拡張するために設計されています。Llama 3 モデルは、幅広いユースケースをサポートする、事前にトレーニングされ、きめ細かに調整された LLM です。特にテキストの要約や分類、センチメント分析、言語翻訳、コード生成に適しています。Cohere の Command R、および Command R+モデルは、お客様がエンタープライズグレードの生成 AI アプリケーションを構築する際にご利用いただける最先端の基盤モデルであり、10 の言語に対応し、グローバルなビジネス運営をサポートする高度な RAG 機能を備えています。

Amazon Bedrock をご利用のお客様とパートナー様の声

Amazon により構築された Rufus は、生成 AI を活用したエキスパートショッピングアシスタントです。Amazon の広範な製品カタログ、お客様によるレビュー、コミュニティ Q&A、および Web で収集された情報でトレーニングされており、さまざまなショッピングニーズや製品に関するお客様の質問に答え、比較を行い、会話の文脈に基づきおすすめを提案します。Amazon のストア基盤 AI 担当バイスプレジデント兼ディスティングイッシュドサイエンティストでもある Trishul Chilimbi 氏は、次のように述べています。「Amazon スタアのお客様に優れた会話型ショッピング体験をお届けするため、Rufus の開発に取り組み、これまでに Amazon が開発したなかでも特に先進的なモデルが完成しました。Rufus は当初想定した用途をはるかに超えた価値をお客様にもたらすことになるかと私たちは確信しています。Amazon Bedrock カスタムモデルインポートにより、Rufus の高度な基盤モデルを Amazon Bedrock 経由で Amazon 社内の開発者に提供できるようになりました。これにより、組織全体のさらに多くのビルダーが、フルマネージドの API として Rufus にアクセスすることができます。今回、ロジスティクスからスタジオまで多様なビジネスチームが Amazon Bedrock による効率的な開発体験のメリットを享受しながら、このモデルを使って開発を進められるようになり、Amazon のすべてのお客様に向けた新たな体験の創出が加速しています」

Dentsu Group は、マーケティングテクノロジーを提供する世界最大のプロバイダーの 1 社です。電通クリエイティブ テクノロジー担当グローバル責任者 James Thomas 氏は、次のように述べています。「この 3 か月間、プレビュー版の Amazon Titan Image Generator のモデルを活用し、自然言語によるプロンプトにより、特にプロダクト・プレイスメントやブランドイメージに沿った画像の生成を中心に、スタジオ品質のリアルな画像を大量に生成してきました。当社のクリエイティブチームは、Titan Image Generator から、世界各地で実施するプロダクトプレイスメントキャンペーンに役立つ魅力的な画像が生成されることに感銘を受けました。今後、当社の AI による生成コンテンツの透明性を高め、お客様とのさらなる信頼関係を構築に向けて、Titan Image Generator の新たな透かし検出機能を検証するのを楽しみにしています」

※抄訳版注釈

Aha!、Pearson、Salesforce、Accenture 各社のコメントに関しては、[本社発表プレスリリース](#)をご参照ください。



アマゾン ウェブ サービスについて

アマゾン ウェブ サービス(AWS)は、2006年に他社に先駆けてサービスを開始して以来、世界で最も包括的かつ幅広く採用されたクラウドサービスになっています。AWSは、コンピューティング、ストレージ、データベース、ネットワーク、分析、機械学習および人工知能(AI)、モノのインターネット(IoT)、モバイル、セキュリティ、ハイブリッド、メディア、ならびにアプリケーション開発、展開および管理に関する240種類以上のフル機能のサービスを提供しています。AWSのサービスは、33のリージョンにある105のアベイラビリティゾーン(AZ)でご利用いただけます。これに加え、マレーシア、メキシコ、ニュージーランド、サウジアラビア王国、タイ、ドイツの6つのリージョンにおける18のAZの開設計画を発表しています。AWSのサービスは、アジリティを高めながら同時にコストを削減できるインフラエンジンとして、急速に成長しているスタートアップや大手企業、有数の政府機関を含む数百万以上のアクティブなお客様から信頼を獲得しています。AWSの詳細については以下のURLをご参照ください。

<https://aws.amazon.com/>

Amazon について

Amazonは4つの理念を指針としています。お客様を起点にすること、創造への情熱、優れた運営へのこだわり、そして長期的な発想です。Amazonは、地球上で最もお客様を大切にしている企業、そして地球上で最高の雇用主となり、地球上で最も安全な職場を提供することを目指しています。カスタマーレビュー、1-Click注文、パーソナライズされたおすすめ商品機能、Amazonプライム、フルフィルメント by Amazon(FBA)、アマゾン ウェブ サービス(AWS)、Kindleダイレクト・パブリッシング、Kindle、Career Choice、Fireタブレット、Fire TV、Amazon Echo、Alexa、Just Walk Out technology、Amazon Studios、気候変動対策に関する誓約(The Climate Pledge)などは、Amazonが先駆けて提供している商品やサービス、取り組みです。Amazonについて詳しくはAmazon Newsroom (<https://amazon-press.jp>) および About Amazon (<https://www.aboutamazon.jp>) から。

報道関係からのお問い合わせ先

アマゾン ウェブ サービス ジャパン合同会社

広報部 awsjp-pr@amazon.com

広報代理店 株式会社プラップジャパン

担当：原口 (080-6656-1591)、中根 (080-6859-3639)

Email: aws_pr@prap.co.jp