

大規模AIモデル運用に対応するMLOps基盤のリリース管理機能

MLOps Platform with AI Model Release Management Function for Large-Scale AI Models

石田 友美 ISHIDA Tomomi 合田 浩二 GODA Koji

MLOps^(注1)の重要性が高まる中、AIモデルの数が増加することで、各モデルの状態やリリース先の把握が複雑化し、運用負荷の増大が顕在化している。

これに対応するため、東芝は、大規模AIモデルの運用に対応したリリース管理機能をMLOps基盤に実装した。この機能により、再学習からリリースまでのプロセスの作業時間を短縮することが可能となった。また、承認プロセスを含む一連の運用フローを支援することで、モデルのリリースに掛かる作業負荷を軽減する。評価の結果、リリース管理機能の有効性が確認されるとともに、更なる最適化の余地も明らかになった。今後は、機能の拡充と運用性の向上を図り、一層広範な事業領域への展開を目指す。

Machine learning operations (MLOps) play a critical role in all processes from the development to the operation of artificial intelligence (AI) models. However, an increase in AI models tends to result in complicated update states and deployment destinations of individual models, leading to heavy operational workloads.

To rectify this situation, Toshiba Corporation has implemented a new AI model release management function that supports large-scale AI model operation in its MLOps platform. It helps shorten the time required for processes from relearning to release and reduces the workload for model releases by supporting a series of operational workflows including approval processes. User evaluation tests show that there is still room for optimization while confirming function effectiveness. We are working to further expand AI model release management function to a wide range of business activities by enhancing functionality and improving operability.

1. まえがき

社会やビジネスのあらゆる場面でデジタル化が進む中、実世界から得られる膨大なデータを活用し、価値ある情報へと変換するAI技術の重要性が高まっている。特に、AIを迅速かつ確実に業務へ適用するためには、データの収集・分析を支える基盤の整備と、AIモデルの運用体制の確立が不可欠である。

東芝は、設備の状態監視、予兆保全、需給予測、最適制御、品質検査など、様々な分野でAIの活用を進めており、現場の意思決定の高度化や製造プロセスの効率化を図っている。しかし、AIモデルは一度開発して終わりではなく、運用中にも外部環境の変化などにより精度が低下するため、継続的な更新が求められる。

このような背景から、AIモデルの学習・評価・展開・監視を一貫して管理できる仕組みが必要となる。また、モデル開発者や、システム開発者、運用担当者など、複数の専門家が連携しやすい環境の整備も重要である。これらを実現

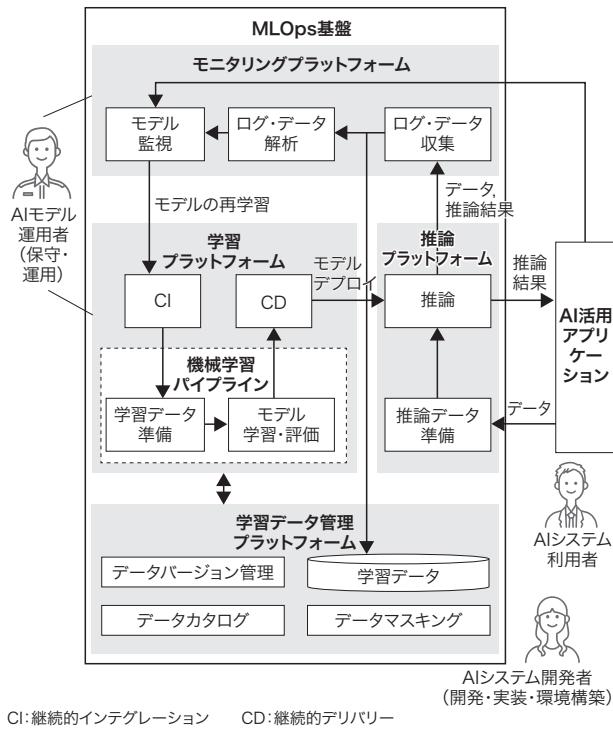
するための仕組みがMLOpsであり、AI運用の効率化と持続性を支える中核的な役割を果たしている。

当社は、AIモデルの高品質かつ迅速な運用を支える共通プラットフォームとしてMLOps基盤の開発と社内展開を進めている^{(1), (2)}。このMLOps基盤は、モデルの学習と評価を担う学習プラットフォーム、学習データの管理と再利用を支える学習データ管理プラットフォーム、学習済みモデルの推論を実行する推論プラットフォーム、及びモデルの精度や動作を監視するモニタリングプラットフォームの四つのプラットフォームから成る(図1)。

これらのプラットフォームを組み合わせることで、学習・デプロイ・推論・監視・メタデータ管理などの機能をワンストップで提供し、各専門家がこの基盤上で一貫して連携・自動化・管理できる環境を整えている。更に、より多くの事業部門で活用できるようにするために、利用者からの評価や運用上の課題を収集し、全社共通の基盤として改善を重ねている。

今回、従来のMLOps基盤を活用したAIモデル運用において、特にリリース管理に課題があることが判明した。ここでは、それを解決するために新たに開発したリリース管理機

(注1) Machine Learning Operationsの略で、AIモデルの開発から、運用環境への配置、運用までのライフサイクル全般を管理する一連のプロセス。



能について述べる。

2. リリース管理機能

2.1 背景

AIモデルの活用が進む中で、AIモデルの開発から運用までの一連の流れ(ライフサイクル)を効率的に管理することが重要になっている。AIモデルのライフサイクルは、データセットの作成・AIモデルの学習・評価・テスト環境へのデプロイ・テスト試行・リリース・運用・精度の監視・AIモデル更新の判断といった複数ステップで構成される

このライフサイクルの中では、様々な種類のデータやプログラムが使用され、状況に応じて追加・更新が発生する。また、設備の異常予兆検知などの用途では、数百規模の設備が対象になるため、同規模のAIモデルを個別に管理する必要がある。これは、設備ごとに動作環境やセンサー構成、稼働条件、故障パターンが異なるため、汎用的なAIモデルでは精度が出ず、設備固有のデータに基づいた個別最適化が不可欠であることが要因である。更に、異常の兆候は設備ごとに微妙に異なるため、一括管理では見逃しや誤検知のリスクが高まり、現場での信頼性を損なう可能性がある。

MLOps基盤は、各プラットフォームにAIモデルのライフサイクルに関する情報を保存している。しかし、運用中のAIモデルを更新する際には、必要な情報が複数の場所に分散して管理されており、一度に取得することが難しい。

例えば、運用中のAIモデルとリリース候補のAIモデルの性能を比較するには、推論プラットフォームから運用中のAIモデルID(識別情報)を調べ、そのIDを使って学習プラットフォームから学習時の記録を検索して必要な情報を取得し、その情報をを使って比較機能で比較する。この際、必要な情報を複数のプラットフォームから人手で集める必要があり、AIモデルの数が増えるほど管理が困難になり、業務のスケーラビリティが損なわれる。

更に、AIモデルのリリース目線での管理が十分に行われていない点も、運用上の問題である。運用中のAIモデルを更新する際には、リリース先を起点に必要な情報を取得し、再学習を行う必要がある。しかし、MLOps基盤では、プラットフォームごとに異なるOSS(オープンソースソフトウェア)が使用されており、情報が分散しているため、統合的に管理・参照できる仕組みが整っていない。このような状況では、AIモデルの数が数百件規模に達すると、手動での追跡や更新が困難で、情報の整合性や最新性を維持することが難しい。加えて、手動作業に起因するミスの発生も避けられず、結果として管理品質の低下が懸念される。AIモデル更新のたびに個別対応が求められるため、運用負荷やリスクが増大し、持続可能なAIモデル運用の障壁になっている。

このように、MLOps基盤には開発・運用に必要な機能は搭載されているが、運用フェーズで重要なリリース目線での業務を行うには、プロセスが複雑で作業ステップが多くなるという問題がある。

このような背景から、AIモデルのライフサイクルを一元的に管理し、リリース業務の効率化を図るためのリリース管理機能を開発した。

2.2 設計方針

リリース管理機能の設計にあたっては、まず再学習からAIモデルリリースに至る業務フローの定義を実施し、次に各工程で必要なトレーサビリティ情報を抽出した。トレーサビリティ情報とは、AIモデル更新に関連するリリース履歴、学習条件、評価結果、精度監視情報などを指す。これらを一元的に管理・活用できる構成を目指した。

リリース管理機能は、API(Application Programming Interface)クライアントとGUI(Graphical User Interface)の二つの主要コンポーネントで構成される(図2)。これは、事業部門やプロジェクトごとに業務フローが異なることを踏まえ、GUIをAPIクライアントを介して柔軟に構成できるよう

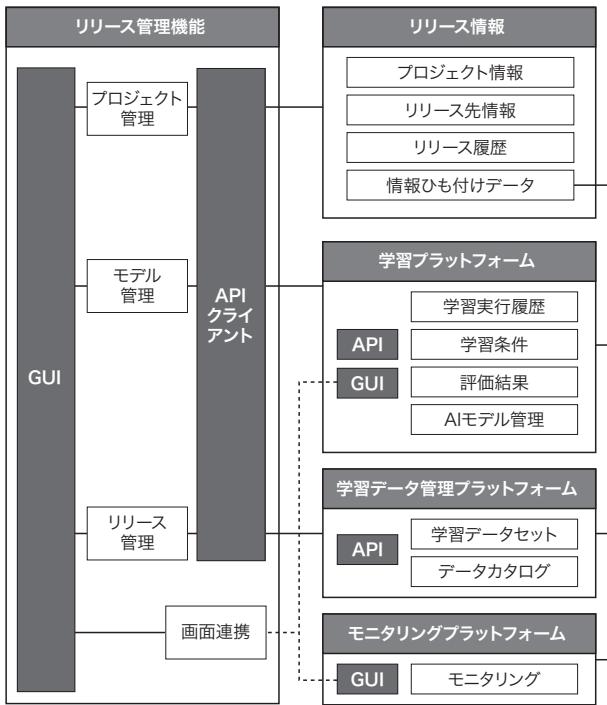


図2. リリース管理機能の構成

APIクライアントとGUIで構成し、トレーサビリティー情報を一元管理することで、モデル比較・承認・再学習・リリース業務の効率化を実現する。

AI model release management function architecture

にするためである。この構成により、分散して存在する情報をバックエンドで一元的に収集・管理し、モデルリリースに必要な承認プロセスを含めた業務の効率化を実現した。

2.2.1 APIクライアント

APIクライアントは、学習プラットフォームと学習データ管理プラットフォームからトレーサビリティー情報を取得し、統合的に管理するためのインターフェース群である。

学習プラットフォームのトレーサビリティー情報は、APIを活用して学習実行履歴、学習条件、評価結果、及びAIモデル管理の情報を取得可能な構成とした。これらの情報は、AIモデルの品質保証及びリリース判断に不可欠であり、業務フロー全体のトレーサビリティーを担保する上で重要な役割を果たす。

学習データ管理プラットフォームは、APIを活用することで、学習・評価に使用するデータセット情報とデータカタログのトレーサビリティー情報を取得可能な構成とした。これにより、学習に用いられたデータの出自や、加工履歴、属性情報などを体系的に把握できる。データカタログ情報は、AIモデルの再学習や精度検証において、使用データの妥当性を確認するための根拠となるものであり、AIモデルの品質保証及びリリース判断に不可欠である。

取得した情報は、プロジェクト情報、リリース先情報、及びリリース履歴と論理的にひも付けて、リリース情報として管理される。これにより、AIモデルの生成から運用、再学習、再リリースに至るまでのライフサイクル全体を通じたトレーサビリティーが確保できる。

2.2.2 GUI

GUIは、APIクライアントを活用し、業務フローに即した操作を可能とするユーザーインターフェースである。提供しているGUIは、汎用的な業務シナリオを基盤とした設計により、利用者が直感的に操作できる(図3)。

このGUIは、以下の機能を備えることで、多様な業務要件に対応する。

- (1) リリース先のグルーピング機能 プロジェクト単位でリリース対象を分類・整理することで、複数AIモデルの同時管理を効率化する。
- (2) 承認プロセスの統合 AIモデルリリースに必要な承認操作をGUI上で完結可能とし、リリース手順の簡素化と作業時間の短縮を実現する。
- (3) 既存GUIとの連携設計 学習プラットフォームが提供するAIモデル比較機能や、モニタリングプラットフォームの監視に適したGUIなど、既存のインターフェースが有効な場合にはAPIクライアントを介さず、直接表示する構成とすることで、操作性と視認性の向上を図っている。

これらの設計により、GUIは共通性と柔軟性を両立し、多様な利用者ニーズに対応できるインターフェースとして機能する。

2.3 評価

開発したリリース管理機能について、実際の利用者による評価を実施した。評価は、数百件規模のAIモデルの継続的なリリース・運用業務が発生することが見込まれる利用者に依頼し、実運用に即した視点からのフィードバックが得られた。

特に、学習結果の比較、及びリリース先を起点とした情報管理は、操作が容易であり、実務上の有用性が高いとの評価を得た。従来は、複数AIモデルの性能を比較する際に複数の画面や手順を経る必要があったが、リリース管理機能により、リリース先ごとに関連する学習結果を比較する作業の手順が大幅に簡素化された。これにより、AIモデル選定やリリース判断に掛かる作業負荷が軽減され、業務効率の向上に寄与できる。また、顧客向けAIモデルの管理にもリリース管理機能を活用したいとの意見が寄せられ、業務展開の可能性が示唆された。

一方、利用者からは、改善に向けた具体的な要望も寄せ

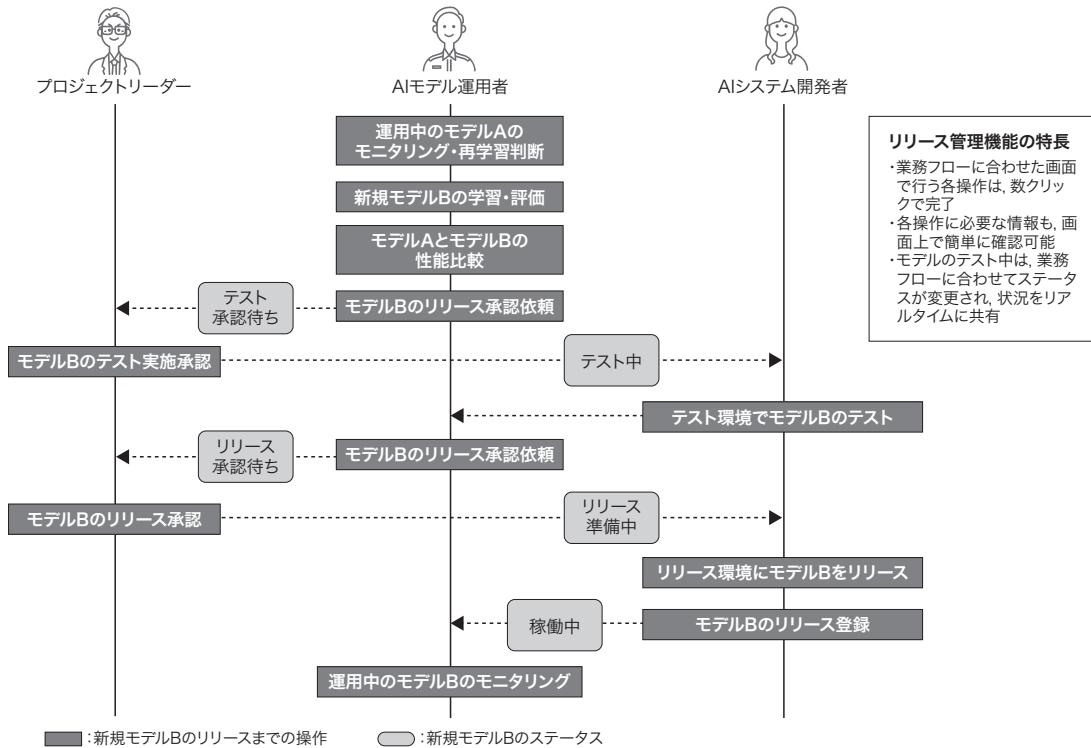


図3. リリース管理機能を使った汎用的な業務シナリオ

リリース管理機能の活用により、承認プロセスを含む汎用業務シナリオを、一層効率化できる。

Generic use case of AI model release management function

られた。中でも、リリース済みAIモデルが他のリリース先で再利用されているかを把握可能な機能の追加が、強く求められたため、これに対応する機能を新たに実装した。修正後のリリース管理機能は既に提供済みであり、今後の運用を通じて更なる改善を継続的に進めていく。

3. あとがき

当社におけるAIモデル運用の高度化を目的に、リリース管理機能を新たにMLOps基盤へ実装した。分散していたトレーサビリティ情報の一元管理することで、AIモデル品質の保証とリリース業務の効率化を実現した。

特に、学習プラットフォームとデータ管理プラットフォームの連携によるAPIクライアント設計、及び業務フローに即したGUI構築により、再学習・比較・承認・リリースといった業務を統合的に管理可能な環境を整備した。

利用者による評価では高い有用性が認められたが、今後は更なる機能拡充と運用性向上を目的として、より多くのフィードバックを収集し、継続的な改善を図る。

文 献

- 小畠知也、山田正隆. AIを適用したIIoTサービスの継続的運用を可能にするMLOps基盤. 東芝レビュー. 2023, 78, 4, p.29-32. <<https://www.global.toshiba/content/dam/toshiba/jp/technology/corporate/review/2023/04/f02.pdf>>, (参照 2025-08-22).
- 東芝. “信頼できるAIを支える東芝のAIガバナンス”. <<https://www.global.toshiba/jp/technology/corporate/ai/governance.html>>, (参照 2025-08-22).



石田 友美 ISHIDA Tomomi

総合研究所
デジタルイノベーション技術センター 実行基盤技術部
Platform Technology Dept.



合田 浩二 GODA Koji

総合研究所
AIデジタルR&Dセンター AI基盤技術開発部
AI Platform Technology Dept.