AIシステムの品質を保証するためのプロセス・技術体系

Process and Technology Framework for AI System Quality Assurance

久連石　圭 KUREISHI Kei　　村田　由香里 MURATA Yukari　　仲　　義行 TSUZUKI Yoshiyuki

AIは高度で複雑な処理の実現が期待できるものの，出力に誤りが含まれることもあるなど品質を担保する上での問題も多く，AIを活用したシステム（以下，AIシステムと略記）の品質を保証することは難しい。

そこで東芝は，AIシステムの品質を保証するためのプロセス・技術体系を整備した。AIシステム特有の品質保証の観点をまとめたAI搭載システム品質保証ガイドラインをベースに，AI品質保証プロセスを定義し，品質評価に活用できる技術の開発を進めている。更に，品質保証活動の記録をAI品質カードに可視化する。これをAIシステム開発に適用することで，高品質なAIシステムを提供していく。

While AI is capable of advanced, complex processing, it poses many issues that make AI system quality assurance difficult, such as erroneous output.

With this in mind, Toshiba Corporation has developed a process and technology framework for AI system quality assurance. We have defined AI quality assurance processes based on AI system quality assurance guidelines that focus on their unique issues and are currently developing technology that can be used for quality evaluation. Further efforts include visualization of quality assurance activity records in the form of AI quality cards. Applying such methods and technologies to AI system development efforts allows us to provide high-quality AI systems.

1. まえがき

近年，AI技術の発展に伴い，AIをシステムに組み込んで活用することが多くなっている。AIシステムは，AIを利用していないシステムよりも高度で複雑な処理の実現が期待できる。一方，AIには不確実性などの特徴があり，品質保証は難しい。

ここでは，機械学習の仕組みを活用してモデルを構築したものをAIとする。AIは，以下に示す二つの特徴が，その品質保証を難しくしている。

(1) 構築されたAIモデルの性能は学習データに依存し，学習していない状況に対応できないことがある。そのため，未学習のデータが入力されると，出力を誤ることや，安全性・公平性を損なうおそれがある。

(2) 多くの場合，AIモデルは大量のパラメーターによって構成されるため，複雑性を持つ。そのため，AIモデルが判断・推論した結果に対して，その動作の過程や理由を正確に述べることが難しい。

これらを解決するため，AIシステム特有の品質保証の取り組みが国内外で行われている。我が国では，AIプロダクト品質保証コンソーシアムがAIプロダクト品質保証ガイドライン[(1)]を公表し，産業技術総合研究所が機械学習品質マネジメントガイドライン[(2)]を公開した。また，欧州ではAIを規制する法案[(3)]が審議されるなどの動きがある。

そこで東芝は，国内外のガイドラインを社内の開発プロセスに取り込んだ形で利用できるように，AIシステムの品質保証に必要なプロセス・技術体系を整備し，プロセスの整備や品質技術の開発を進めている。

ここでは，AIシステム品質保証のためのプロセス・技術体系の全体像について述べる。

2. AIシステム品質保証のためのプロセス・技術体系の概要

当社では，東芝グループ総合品質保証基本方針に基づき，製品の品質保証を行っている。AIシステムも品質保証を確実に遂行するため，AIシステムに適用するための品質保証プロセス・技術体系を**図1**のとおり整備した。

まず，品質保証の観点を定義し，品質保証として何をすべきかをまとめたAI搭載システム品質保証ガイドラインを定めた。これを基に，品質保証の手順を定めたAI品質保証プロセスの構築と，AIの品質を評価するためのAI品質技術・AIテスト技術の開発を行っている。更に，これらの取り組みや評価結果をまとめて，品質の可視化をすることで，AIシステムの品質を明らかに示していく。次章から，それぞれの項目について述べる。

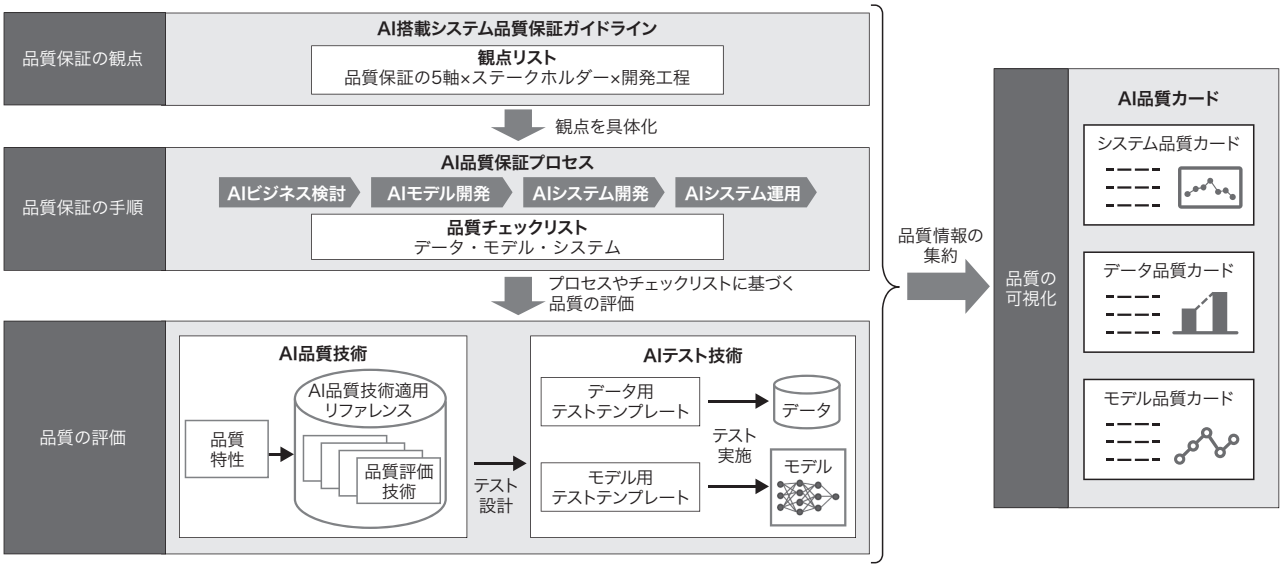

図1. AIシステム品質保証プロセス・技術体系

東芝は，品質保証に必要な観点，プロセス，及び技術を体系的に整備し，AIシステムの品質保証に取り組んでいる。

AI system quality assurance process and technology framework

3. AI搭載システム品質保証ガイドライン

AI搭載システム品質保証ガイドラインでは，**図2**に示すように品質保証の5軸，ステークホルダー，及び開発工程の三つの要素から，観点を網羅的に抽出している。

品質保証の5軸は，AIプロダクト品質保証コンソーシアムのAIプロダクト品質保証ガイドラインを参考に，AIシステムの品質保証に必要な視点として，次のとおり定義した。

- データ　　学習や評価に利用するデータに関する視点
- モデル　　学習したモデルの構成や評価に関する視点
- システム　　システム全体の品質や安定性に関する視点
- 開発プロセス　　開発や評価の進め方に関する視点
- 顧客　　顧客が持つ期待を明確にするための視点

ステークホルダーの分類では，AIシステムの開発・運用に携わる役割を明確にした。AIモデルの学習・評価を担当するAIアナリストチームやシステム開発を担うAIシステム開発チームだけでなく，営業や企画を担当するAIプランニングチームや運用を行うAIシステム運用チームもステークホルダーとして挙げている。品質保証チームは，AIシステムの品質保証に責任を持つステークホルダーであり，開発を担う各チームとは独立した役割としている。

開発工程は，AIシステム開発のライフサイクルとして，AIビジネス検討，AIモデル開発，AIシステム開発，及びAIシステム運用の四つに分けた。

AI搭載システム品質保証ガイドラインでは，上記三つの要素を組み合わせ，開発の中で何を気にするべきかを示す観点リストを定義した。これにより，誰が，どの工程で，何を確認すべきかを明確にできる。

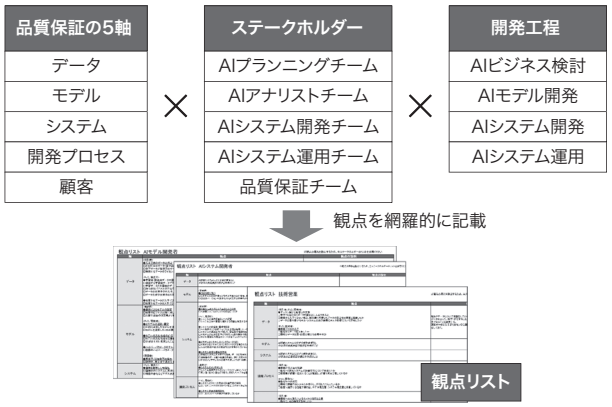

図2. AI搭載システム品質保証ガイドラインの観点

品質保証の5軸・ステークホルダー・開発工程の三つの要素を組み合わせて，品質保証に必要な観点リストとして整理した。

Perspectives on AI system quality assurance guidelines

4. AI品質保証プロセス

AIシステムの開発と品質保証の流れをまとめたものが，AI品質保証プロセス(**図3**)である。AI品質保証プロセスは，プロセス文書と品質チェックリストで構成している。

プロセス文書には，開発工程に対応付けた詳細な活動

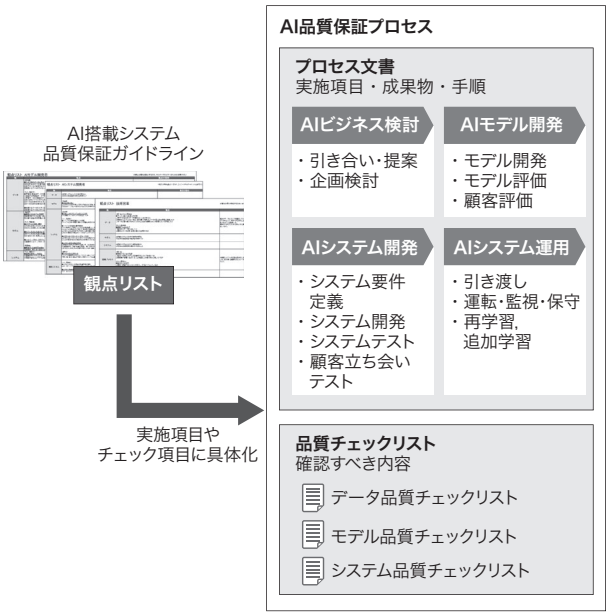

図3. AI品質保証プロセスの構成

AI搭載システム品質保証ガイドラインの観点を具体化したプロセス文書とチェックリストにより，AI品質保証プロセスを構成する。

AI quality assurance process configuration

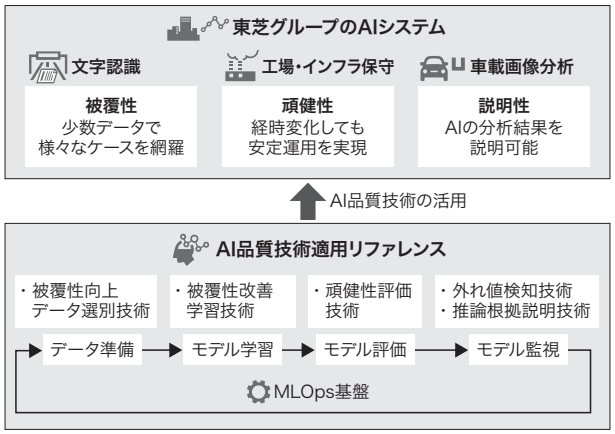

図4. AI品質技術の活用方法

MLOps基盤の各フェーズにAI品質技術を用意し，AIシステムの品質評価に適用する。

Usage of AI quality technologies

（アクティビティー・タスク）と成果物をまとめた。AI搭載システム品質保証ガイドラインで示す観点を基に，AIシステムの開発や品質保証に必要な活動を整理した。品質チェックリストは，AIシステムの開発や品質保証に必要なチェック項目を，学習データを対象にした“データ品質チェックリスト”，AIモデルを対象とした“モデル品質チェックリスト”，及びAIシステムの全体を対象とした“システム品質チェックリスト”の三つに分けて構成している。システム開発の各アクティビティーやタスクで，チェックリストを確認しながら開発を進め，リリース前にチェック結果を確認することで，品質保証の観点に漏れなく対応した開発ができたことを確認する。

AI品質保証プロセスでは，このように基本的なプロセスをまとめているが，開発プロセスは開発する組織やシステムで異なることが一般的である。そこで，AI品質保証プロセスを開発組織や対象システム向けにカスタマイズするためのテーラリングガイドも作成している。テーラリングガイドは，開発するAIシステムの特徴に応じたプロセスパターンを定義し，プロセスパターンに応じたプロセスやチェック項目の取捨選択方法をガイドする。

5. AI品質技術とAIテスト技術

4章で述べたプロセスやチェックリストによる品質保証活動を実現するため，AIシステムの品質特性を考慮したAI品質技術の開発・整備も進め，AI品質技術適用リファレンスやAIテスト技術としてまとめている。

AI品質技術適用リファレンスとは，AIシステムの品質評価や品質改善に活用できる技術を，品質特性に対応付けてカタログ化したものである。AI品質技術適用リファレンスでは，機械学習品質マネジメントガイドラインを参考に，当社が拡張した品質特性を定義している。品質特性ごとに技術を用意することで，AIシステムに期待される品質の確保を目指す。更に**図4**のように，MLOps（注1）の各フェーズに対応するAI品質技術を整備することで，学習・評価時だけではなく，データ準備やモデル監視においても品質向上にアプローチする。

これらのAI品質技術は，システムのテスト工程でも活用している。AIシステムに対するテストは，網羅的かつ効率的に実施することが求められるため，必要なテストをまとめたAIテストテンプレートを定義した。AIテストテンプレートでは，AIシステムのテストで重要なデータやモデルに関するテストの手順をまとめ，AI品質技術を活用した。これにより，様々なAIシステムに対して，必要なテストが漏れなく実施できるようになる。

6. AI品質カードによる品質の可視化

これまで述べた品質保証の観点に基づく品質チェックや品質技術を用いた品質評価の結果を，顧客が理解できる形に可視化する取り組みも進めている。

（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。

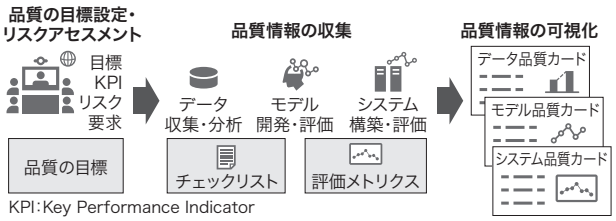

図5. AI品質カードによる品質の可視化

開発の初期に定義した品質の目標に対して，どの程度目標を達成できたか品質情報を可視化する。

Quality visualization using AI quality cards

可視化の流れは，**図5**に示すとおりである。開発の初期段階であるAIビジネス検討段階で，AIシステムの要求や考え得るリスクから，AIシステムが達成すべき品質の目標を定義する。その後，AIモデルの開発やシステムへの組み込みを行う中で，様々な評価を実施し，品質に関するデータを蓄積する。これらの蓄積したデータを，開発の節目やシステムのリリース前に，AI品質カードと呼ぶ報告書にまとめ，開発初期に設定した目標が達成されているかどうかを示す。

AI品質カードはシステム全体を示すシステム品質カードに加え，AIの学習に使ったデータを示すデータ品質カードと学習したAIモデルを示すモデル品質カードの3種類で構成される。AI品質カードには，顧客の要求やシステムにあるリスクに関連付いた項目を並べ，評価結果を記入する。これにより，システムはどのような性能を持っているか，どのようなデータで学習しているか，AIの弱点は何か，どのように運用すればよいかなど，AIシステム，及び利用しているデータやモデルの特徴や品質を，顧客に分かりやすく示すことができる。

7. 品質保証活動のAIシステムへの適用

これまで述べたAIシステムの品質保証ガイドラインや品質技術を，当社が開発するシステムに適用開始した。

一つの例として，開発組織や対象のシステムに適合したプロセスを整備する活動を進めている。AIシステムは，AIモデルの開発を担当する組織とシステム構築を担当する組織に分かれて開発することも多い。加えて，学習データの追加やAIモデルの再学習が頻繁に行われ，各作業を反復的に実施することがある。このような場合に，それぞれの組織が行うべき作業項目や連携の手順を明確にすることで，実施すべき事柄やチェックすべき項目の漏れを防ぐようにしている。

8. あとがき

これまで述べたように当社は，AI搭載システム品質保証ガイドラインで定義したAI品質保証の観点をベースに，AI品質保証プロセスやAI品質技術，AIテスト技術，可視化技術を構築している。これらによって，AIの不確実性などにより品質保証が難しいといった問題を解決する。

現在，このプロセス・技術体系に沿った教育プログラムを作成し，AI開発に携わるメンバーへの展開を進めている。品質技術も，AIシステムに適用できるように技術開発を進めている。

このような活動を通じて，顧客がAIシステムを安心して利用できるよう，更なるプロセスの改善や新たな技術の開発を進めていく。

文　献

(1) AIプロダクト品質保証コンソーシアム編．AIプロダクト品質保証ガイドライン 2022.07版．2022，288p．<http://qa4ai.jp/QA4AI.Guideline.202207.pdf>，(参照 2023-04-25)．

(2) 国立研究開発法人 産業技術総合研究所．機械学習品質マネジメントガイドライン 第3版．2023，256p．<https://www.digiarc.aist.go.jp/publication/aiqm/AIQuality-requirements-rev3.2.1.0079-signed.pdf>，(参照 2023-04-25)．

(3) European Commission. “Proposal for a Regulation laying down harmonised rules on artificial intelligence”. <https://digital-strategy.ec.europa.eu/en/library/proposal-regulation-laying-down-harmonised-rules-artificial-intelligence>, (accessed 2023-04-25).

久連石　圭　KUREISHI Kei
デジタルイノベーションテクノロジーセンター
先端ソフトウェア技術室 ソフトウェアエンジニアリング技術部
Software Engineering Technology Dept.

村田　由香里　MURATA Yukari
デジタルイノベーションテクノロジーセンター
先端ソフトウェア技術室 ソフトウェアエンジニアリング技術部
Software Engineering Technology Dept.

仲　　義行　TSUZUKI Yoshiyuki
研究開発センター
知能化システム技術センター AI応用推進部
AI Application Dept.