

生成AIが変革する ソフトウェア品質保証の未来

- ソフトウェアテストの進化と組織・エンジニアの真価 -

2025年12月24日

株式会社日本総合研究所

先端技術ラボ/創発戦略センター

<本資料に関するお問い合わせ>

執筆者：先端技術ラボ 伊藤蓮 創発戦略センター 木村智行

本レポートに関するお問い合わせにつきましては、当社ホームページの 問い合わせフォーム よりご連絡ください。

本資料は、作成日時点で弊社が一般に信頼できると思われる資料に基づいて作成されたものですが、情報の正確性・完全性を弊社で保証するものではありません。また、本資料の情報の内容は、経済情勢等の変化により変更されることがありますので、ご了承ください。本資料の情報に起因して閲覧者及び第三者に損害が発生した場合でも、執筆者、執筆取材先及び弊社は一切責任を負わないものとします。本資料の著作権は株式会社日本総合研究所に帰属します。本資料の一部または全部を、電子的または機械的な手段を問わず、無断での複製または転送等することを禁じております。

1

ソフトウェア品質保証の意義

- ・ソフトウェア品質保証(QA)業務とは、ソフトウェアが要求された品質基準を満たすか検証する業務。
- ・低品質なソフトウェアは事業損失にも直結するため、QA業務は重要。

2

生成AIの進展と品質保証業務へもたらす影響

- ・生成AIの効果的な導入により、困難だったソフトウェア開発の品質・コスト・納期の適正化が期待されるため、生成AIによるQA業務の自動化は重要。
- ・一方、AIの生成物の品質保証などの新しい課題が生じる可能性がある。

3

生成AIを活用したソフトウェア品質保証業務のユースケース

- ・ユースケースとして、仕様書レビュー支援、テストケース・データ生成、テストコード生成、テスト自動実行・結果分析、脆弱性診断などが挙げられる。
- ・AIによってテストの網羅性やメンテナンス性の向上が期待される。

4

ソフトウェア品質保証に役立つ生成AIツール

- ・様々な実用的なサービス・ツールが出てきており、①対話型生成AIサービス、②AIエージェント型サービス、③コーディング補助ツール、④生成AIで強化されたテスト専用ツールに分類できる。

5

今後の変化と開発組織・QAエンジニアへの提言

- ・生成AIの導入は進む一方で、限界も存在。人間の適切な介入やAIの有効活用のための仕組みが必要。
- ・今後、AIによる自動テストの範囲が徐々に拡大し、QA業務全体として「ビジネス要件の検証」へのシフトが重要となっていくため、開発組織とQAエンジニアは協調してQA業務の変革を進めることが重要。
- ・開発組織はAIを有効活用する土台を作り、ソフトウェア開発プロセスの変革や人材強化などをすべき。
- ・QAエンジニアは、品質に関する重要な意思決定や、ビジネス観点やユーザ視点などの多角的な品質保証を強化していくべき。

章	項目	ページ
はじめに	<ul style="list-style-type: none"> 本レポートの背景、目的、対象読者 	p.4
1. ソフトウェア品質保証の意義と生成AIがもたらす影響	<ul style="list-style-type: none"> ソフトウェア品質保証業務とは 生成AIの進展とQA業務・ソフトウェア開発への影響 QA業務の自動化範囲の拡大によるQCD適正化 	pp.5~7
2. ソフトウェア品質保証業務への生成AI活用動向	<ul style="list-style-type: none"> 生成AIを活用したQA業務の想定ユースケース QA業務の支援に役立つ生成AIツールの動向 国内の各企業・組織での取り組み 	pp.8~14
3. ソフトウェア品質保証業務の未来	<ul style="list-style-type: none"> 現状の生成AIの限界と品質保証への導入における課題 品質保証における生成AIとエンジニアが担うスコープの変遷 AIの進展による今後の品質保証業務の変化 	pp.15~17
4. 今後を踏まえて開発組織・QAエンジニアがすべきこと	<ul style="list-style-type: none"> QAエンジニアを支える開発組織の役割 今後を踏まえたQAエンジニアの役割と求められるスキル QAエンジニアが活かすべき人間特有の能力 	pp.18~20
参考・コラム	<ul style="list-style-type: none"> QA4AIの取り組み・考え方 <p>【コラム】ニューロダイバーシティ～生成AIを効果的に使う人材の登用～</p> <ul style="list-style-type: none"> ニューロダイバーシティの概要と人材の活躍事例 生成AI時代に活きる発達障害がある人の強み 	pp.21~23

● 背景：

システム・ソフトウェア開発の領域では大規模言語モデル（LLM）を中心とした生成AIが積極的に活用され、各業務での効率化が進んでいる。ソフトウェア開発業界において、現状はコード生成への活用を中心に生成AIが積極的に活用されてきている。ソフトウェア開発において、ソフトウェア品質保証(QA)の業務は開発したソフトウェアの品質を担保する重要な業務である。一方、ソフトウェアテスト工程を中心としたQA業務への生成AI活用についてはまだ手探り状態であるのが現状と見られる。

本レポートでは、そのような現状を鑑みて、ソフトウェアの品質保証業務への生成AI活用について技術動向をまとめ、今後起こり得る変化について考察するものである。

本レポートでは比較的大規模なソフトウェア開発を主な対象として記載するが、一部それに限ったものではない。

● 本レポートの目的：

本レポートでは、生成AIの進展とソフトウェア品質保証領域に与える影響について述べたのち、生成AIツールの動向や国内組織において行われている取組みについてまとめる。その後、ソフトウェア品質保証の今後の変化を踏まえ、QAエンジニア個人や開発組織がどのような行動をしていくべきか提言する。

● 対象読者：

本レポートの対象読者としては、

- ①生成AI活用を模索する、もしくは、これから生成AIを活用していきたいソフトウェアエンジニア、特にQAエンジニアのようなソフトウェア品質保証に従事する人
- ②今後のソフトウェアの品質保証業務の展望について知りたいマネージャーやPM、エンジニアリーダー を想定する。
LLMについての基礎知識があるとより内容について理解が進むものと思われる。

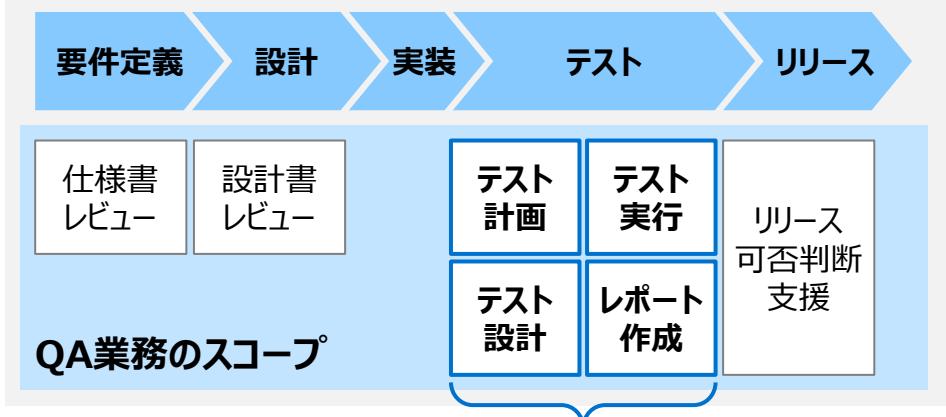
ソフトウェア品質保証業務とは

- ソフトウェア品質保証(以下、QA)業務は**ソフトウェアが要求された品質基準を満たすかを検証する業務**。
- 低品質なソフトウェアは事業損失にも直結するため、**QA業務は重要**。
- AI駆動開発による開発効率化の一方、品質低下の可能性もあり、**QA業務の重要性がより高まる**と考えられる。

ソフトウェア品質保証(QA)業務の範囲

- ・QA業務は、開発ライフサイクル全体にわたってソフトウェアが要求された品質基準を満たすか検証する業務である。
- ・本レポートでは主たる業務領域である「ソフトウェアテスト」を中心に記載する。
- ・QA業務に従事するエンジニアを「**QAエンジニア**」と呼ぶ。

ソフトウェア開発工程



**QAエンジニアの主たる業務領域
=本レポートではテスト工程を中心に記載**

QA業務の意義・重要性

- ・QA業務による早期の不具合発見・修正はソフトウェアの信頼性向上に寄与し、ユーザの満足度を向上させる。
- ・**QA業務はビジネス上のリスクを低減させるためにも重要**。

低品質なソフトウェアによるビジネス損失

Tricentis社の2025年10月の調査結果によると、下表の通り全世界でおよそ半数の企業が「低品質なソフトウェアによって損失を受けた」と回答しており、ソフトウェアの品質がビジネスにも大きく影響することがわかる。

「**低品質なソフトウェアによる損失を受けた**」と回答した企業の割合と損失額

単位：100万米ドル/年

国名	回答の内訳	損失額
アメリカ	45%	\$5
イギリス	45%	\$0.5~1
日本	52%	\$0.5~1
シンガポール	40%	\$0.5~1
ドイツ	44%	\$0.5~1

[出所] "2025 Quality Transformation Report", Tricentis, https://be.tricentis.com/media-assets/pdf/Tricentis-report_Tricentis-2025-quality-transformation-report.pdf
(閲覧日：2025/11/28) より日本総研作成

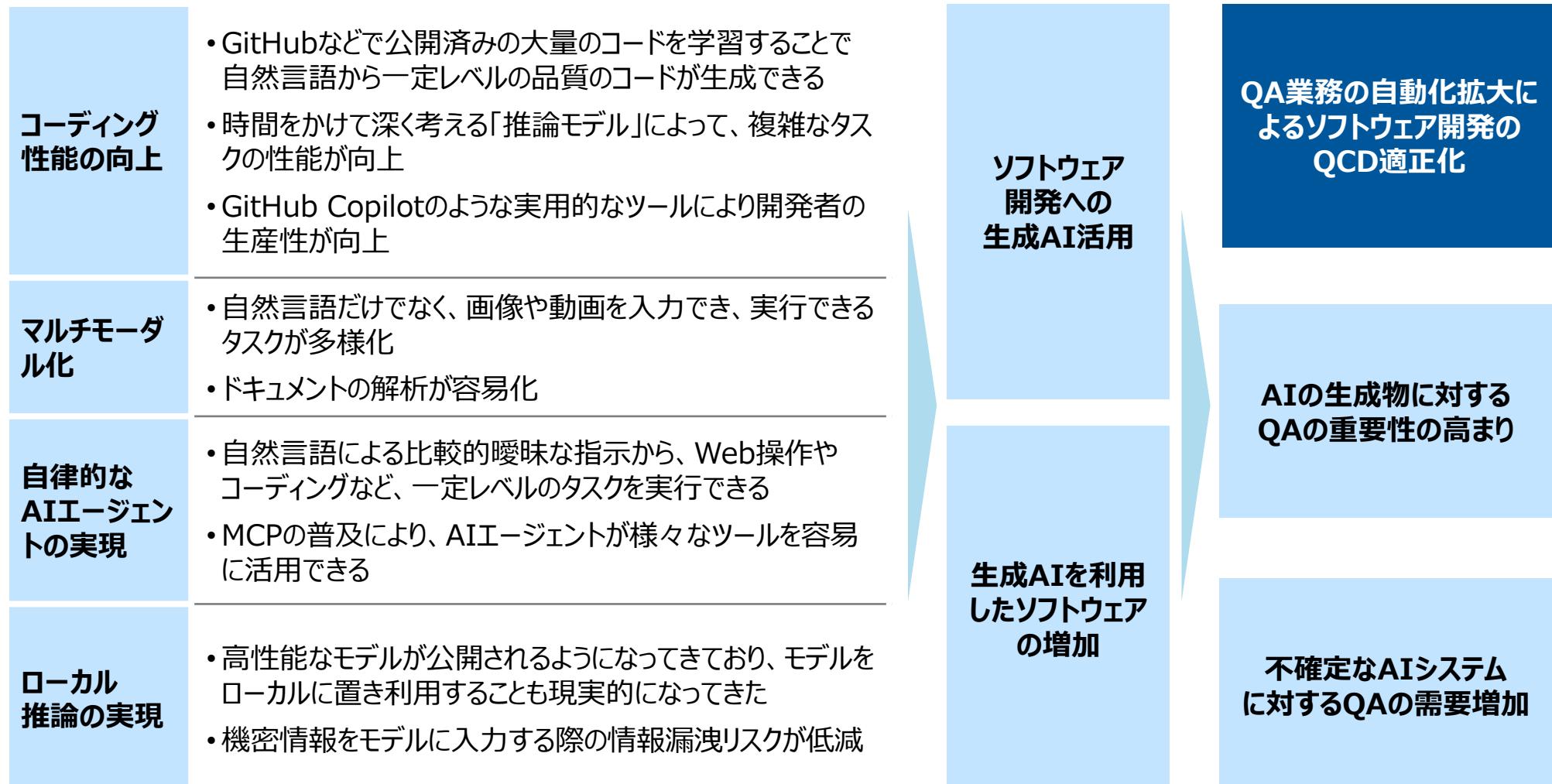
- ・生成AIでソフトウェア開発が効率化され得る一方、求める品質基準に満たないケースが増加する可能性もある。そこで、**ソフトウェアの品質保証がますます重要**になるだろう。

生成AIの進展とQA業務・ソフトウェア開発への影響

- ソフトウェア開発全体のQCD*適正化などの利点があるため、**QA業務への有効な生成AI活用は重要**である。
- 既存の課題を解決する一方で、AIの生成物に対する品質保証が増加するなどの新しい課題が生じる可能性がある。

生成AIの技術的な進展

QA業務にもたらす影響

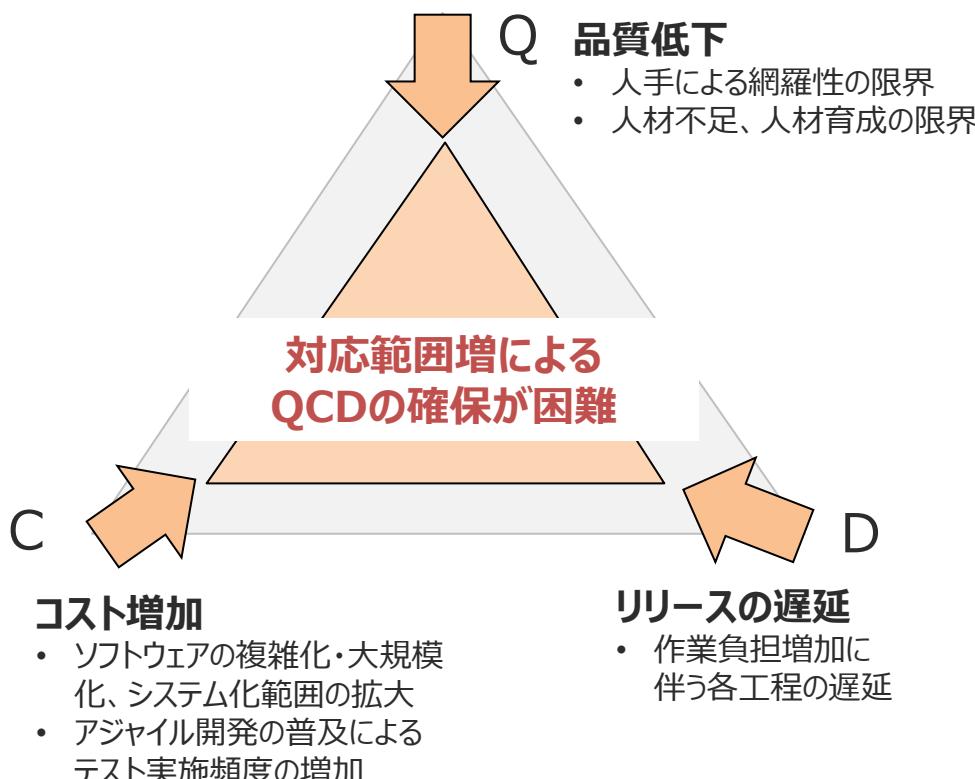


QA業務の自動化範囲の拡大によるQCD適正化

- 品質確保と開発効率は実質トレードオフであり、DXの進展や人材不足などを背景に、既存のソフトウェア開発では **QCDの確保が困難** していくものと予想される。
- 生成AIの効果的な導入により、品質保証業務の自動化範囲が拡大することで、**QCDの適正化**が期待される。

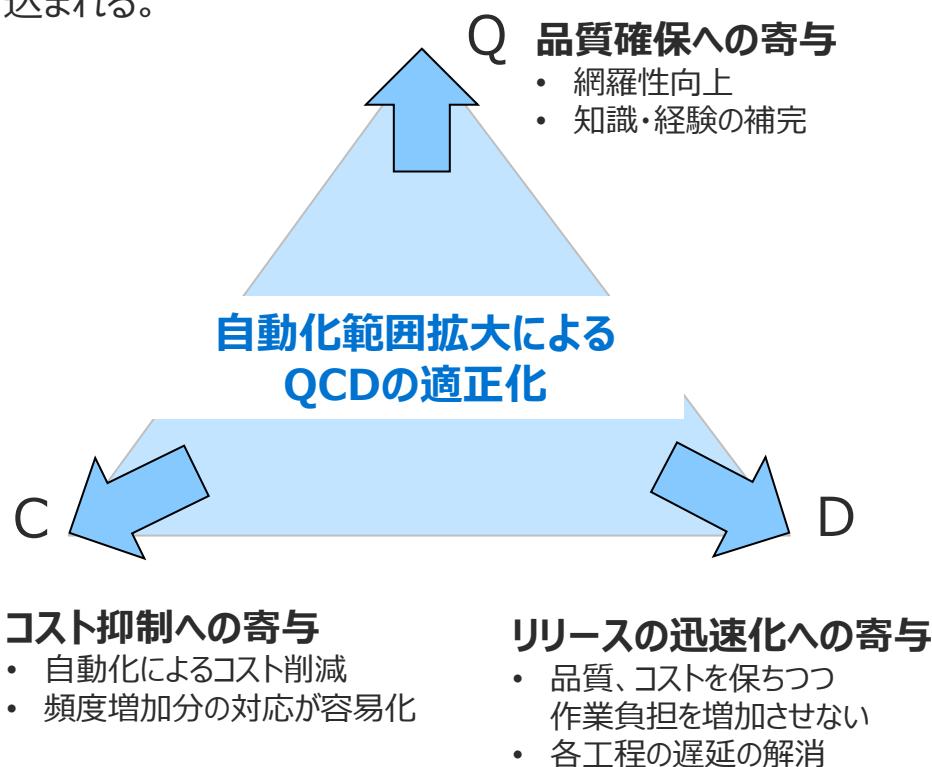
従来のソフトウェア開発が抱えるであろう課題

- ソフトウェアの複雑化、アジャイル開発の普及などにより **対応すべき範囲が拡大（例：テスト頻度増加）**。
- 自動化や人材育成の限界もあり、品質確保と開発効率はトレードオフなため、QCDの確保が困難していくと予想。



生成AI導入により見込まれる効果

- 生成AIがQA業務に効果的に利用されることによって、エンジニアの作業負担を増大させることなく、**品質保証業務で対応できる範囲が拡大**。
- 品質確保・コスト抑制への寄与により、QCDの適正化が見込まれる。



生成AIを活用したQA業務の想定ユースケース(1/2)

- QA業務の各工程において生成AIの活用が想定されるユースケースを列挙。幅広く生成AI活用の余地がある。
- LLMによって、**テスト対象の理解を踏まえた**仕様書・テストケースの作成や**柔軟なテスト実行**、自己修復による**テストコードのメンテナンス性向上**、テストからの**ソフトウェアの改善点の洞察**などの利点を得ることができる。

計画・設計

仕様書・設計書レビュー支援

仕様書や要件定義、設計書に関して考慮すべき観点の抜け・漏れをAIが提案・指摘

テスト計画・設計支援

テスト計画策定やテスト観点の設計において、AIがたたき台を作成

項目作成・準備

テストケース自動生成

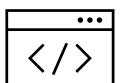


仕様書などのドキュメントからテストケースを自動生成。熟練した専門家でなくともカバレッジを向上できる



ソフトウェアに入力されるフォーマットや条件に基づき、テスト実行に必要なダミーデータ作成を効率化

テストコード自動生成

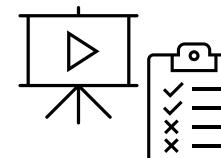


作成済みのテストケースや既存コードを基に**テストコードを自動生成**。LLMにより**コードの意図を理解したテスト**を実装できる。

実行・記録

テスト自動実行・記録

- 人手では手間がかかるていた**テスト実行と結果の自動記録**を、AIエージェントにより実現。
- UIテストの自動化も柔軟にできると同時に、UI・機能の変更に伴う**テストコードの自己修復（セルフヒーリング）**を実現するなど、メンテナンス性も向上



分析・改善

テスト結果分析



- 実行したテスト結果レポートの作成を支援してくれる
- 大量のテスト結果から失敗の傾向などを要約し、**ソフトウェアの改善点など新たな洞察**を提供

バグ分析・修正案提示



- コードを静的に解析して**潜在的なバグを検出**できる
- テスト失敗時のエラーログを分析して原因を特定し、修正案を提示

脆弱性診断、セキュリティテスト支援



- コードを解析してセキュリティの観点で診断する
- セキュリティ面での考慮漏れを指摘してくれる

生成AIを活用したQA業務の想定ユースケース(2/2)

- 一般的なテスト工程の分類を用いて、より具体に各テスト工程における生成AI活用のユースケース例を記載。
- 開発組織のQA業務の方法やテスト対象となるソフトウェアの種類・規模により生成AIの導入しやすさは異なるが、部分的な導入も含めて、幅広く実務上での生成AI活用の余地がある。

テスト観点	単体テスト	結合テスト	システムテスト/E2Eテスト	受け入れテスト
計画・設計	<ul style="list-style-type: none"> 既存仕様書・コードのロジック解析とテスト観点洗い出し 可読性、テスト容易性を高めるためのリファクタリング提案 	<ul style="list-style-type: none"> API仕様書(Swagger等)からの連携シナリオ抽出 DBや外部サービスとのデータフロー分析とテストシナリオ設計 	<ul style="list-style-type: none"> 要件定義書やUIデザインからのユーザーストーリー作成 画面遷移図に基づくE2Eテスト網羅性の確認と計画策定 	<ul style="list-style-type: none"> ビジネス要件における受け入れ基準の明確化 ユーザーストーリーにおける仕様の曖昧さ、矛盾点の指摘
項目作成・準備	<ul style="list-style-type: none"> 【データ作成】正常系、異常系、境界値系等のテストデータ作成 【コード生成】xUnit等のテストフレームワークに則るコード実装 【コード生成】モック・スタブ実装 	<ul style="list-style-type: none"> 【データ作成】APIリクエスト用のデータパターン作成 【データ作成】テストデータ投入用のスクリプト実装 【コード生成】APIテストコード実装 	<ul style="list-style-type: none"> 【ケース作成】ユーザの操作の流れからE2Eテストシナリオ作成 【コード作成】Playwright, Selenium等のテストコード実装 【データ作成】本番データを模したダミーデータ作成 	<ul style="list-style-type: none"> Gherkin記法*を用いた自然言語に近いテストケース記述 手動テスト用の操作手順書のドラフト作成
実行・記録	<ul style="list-style-type: none"> テスト実行時のス택ト・トレースの解説と修正案提示 【自己修復】テスト失敗時のテストコード、製品コード改善 	<ul style="list-style-type: none"> 連携エラー時のログ解析による障害箇所の特定 外部APIの挙動をシミュレートする動的なレスポンス生成 	<ul style="list-style-type: none"> テスト実行時のスクリーンショットやログの記録の自動化 【自己修復】UI変更でテストが失敗した時に画面上の要素を特定してテストスクリプトを自動修正 	<ul style="list-style-type: none"> 手動テスト担当者向けの、プロダクト仕様に関するQ&Aボット シナリオなしの探索的テスト実行 テスト実行結果とエビデンス画像のレポート作成
分析・改善	<ul style="list-style-type: none"> バグの原因特定、修正コードの提案 カバレッジ不足領域の特定と追加テストケースの提案 	<ul style="list-style-type: none"> コンポーネント連携後のパフォーマンスのボトルネック特定と改善案の提示 回帰テスト結果からの影響範囲分析と評価 	<ul style="list-style-type: none"> 表示崩れの差分要約と重要度の判定、修正案提示 システム全体のボトルネックの特定と修正案提示 	<ul style="list-style-type: none"> 非技術者向けの品質レポート作成、リリース判定支援 検出されたバグ傾向の分析と上流工程へのフィードバック作成

* Gherkin記法 : Given(前提条件)、When(テストが有効なタイミング)、Then(期待結果)の形でテストケースを記載する方法。受け入れテストに限らず他のテスト工程でも利用できる。

QA業務を支援する生成AIツールの動向(1/4)

- QA業務を支援してくれる生成AI活用ツールの概観として、①対話型生成AIサービス、②AIエージェント型サービス、③コーディング補助ツール、④生成AIで強化されたソフトウェアテスト専用ツールの4つに分類できる。
- 任せたいタスクに特化したツールを使うべきかどうかや、AIに作業してもらいたい範囲によって最適なツールが異なる。

次ページ以降で各ツールの概要を紹介



QA業務を支援する生成AIツールの動向(2/4)

- ChatGPTなどの対話型生成AIサービスは、プロンプト次第でソフトウェアテストなどのQA業務にも活用できる。
- 自律的なAIエージェントは単なる自動化を超え、テスト対象の仕様・挙動の理解を伴うQA業務も可能となってきている。

①対話型生成AIサービスによるQA業務

- ChatGPTのような生成AIサービスにおいて、チャット形式で依頼し、生成AIの回答をQA業務に利用できる。
- QA業務のアシスタントとして質の高い結果を得るために、対象のソフトウェアの情報や開発状況などの十分なコンテキストや、制約条件、出力形式の指定などを細かく指定する必要がある。

ChatGPT (OpenAI)

- チャット形式でプロンプトを入力することで対話型に回答を取得できる生成AIサービス
- 入力するプロンプト次第で様々な用途に活用できるため、一般に広く利用されている。
- 具体的にはChatGPTは、全世界の成人の約10%が利用(2025年7月時点)するまでに普及*。

Gemini (Google)

- チャット形式でプロンプトを入力することで対話型に回答を取得できる生成AIサービス
- 入力するプロンプト次第で様々な用途に活用できるため、一般に広く利用されている。
- 具体的にはChatGPTは、全世界の成人の約10%が利用(2025年7月時点)するまでに普及*。

Claude (Anthropic)

- チャット形式でプロンプトを入力することで対話型に回答を取得できる生成AIサービス
- 入力するプロンプト次第で様々な用途に活用できるため、一般に広く利用されている。
- 具体的にはChatGPTは、全世界の成人の約10%が利用(2025年7月時点)するまでに普及*。

対話型生成AIサービスの利用イメージ

仕様書を基に機能Aの正常系のテストケースを作成して



ユーザ

テストケース1 : ...
テストケース2 : ...



AI

次に、異常系のテストケースを作成して

②AIエージェント型サービスによるQA業務

- AIエージェントによって、要件定義書やプロダクトコードなどを基に自律的に対象を理解し、レビュー支援やテストを実行できる。
- ある程度のレベルでテスト計画も含めて幅広く任せられるが、質の高い結果を得るには、適切なタスク粒度での指示などが重要である。

Devin (Cognition)

- ソフトウェア開発タスク全般を自律的に実行できる。
- 依頼内容次第で、コードリポジトリからのテストケース・コードの作成や、テストの実行、自動デバッグ、テスト報告書作成などQA業務全般で利用可能

Antigravity (Google)

- 2025年11月発表。コードエディタのUIを持つが、AIエージェントによりブラウザ操作を伴うテストも可能。

Spark (Code Intelligence)

- コードを分析し、様々なデータを入力してソフトウェアの脆弱性を検知する「ホワイトファズテスト」を実行。
- 10万行のコードのテストで最大1000時間を短縮する等、大規模なソフトウェアへも適用実績あり。

AIエージェント型サービスの利用イメージ

機能Aの正常系のテストコードを作成して



ユーザ



テストコード

:

コードリポジトリから機能Aに
関する部分を抽出

テストコード実装

テスト実行確認
コード修正

作業が終わりました



Devin

QA業務を支援する生成AIツールの動向(3/4)

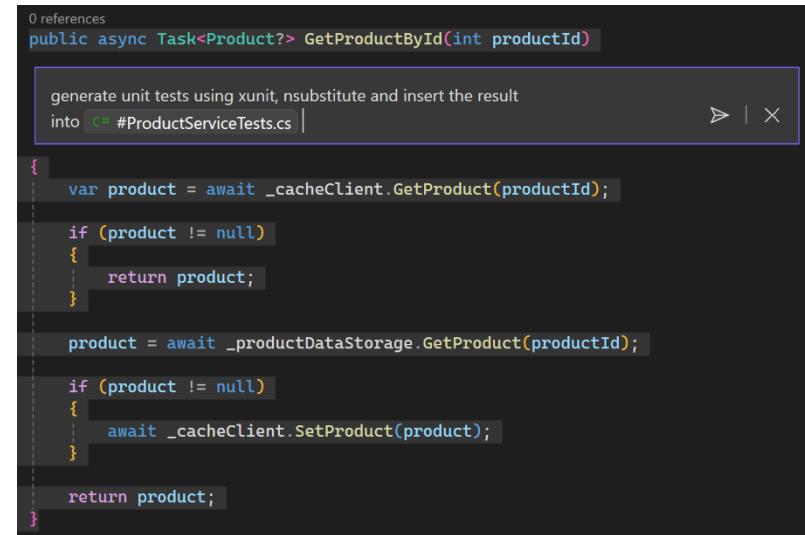
- GitHub CopilotやCursorなどのコーディング補助ツールは、テストコード作成などで使われており、QA業務にも役立つ。
- 対話型生成AIサービスを提供するOpenAIなどの各社も、CLIベースのコーディング支援ツールを発表しており、テストコード生成に限らずQA業務全般に活用できる可能性が高い。

③コーディング補助ツールによるソフトウェアテスト

- GitHub CopilotやCursorのようなコーディング支援ツールに依頼すればテストコードの作成といった業務を支援してくれる。
- QAエンジニアのペアプログラマーとして機能し、テストコードの実装の生産性を向上してくれる。

GitHub Copilot (GitHub)	<ul style="list-style-type: none"> • 生成AIを活用したコーディング支援ツールの代表格 • GitHub上のコードを学習したAIモデルであり、様々な言語のコード生成に対応。コミットメッセージ生成、プルリクエスト生成、脆弱性修正など、開発ライフサイクル全体の効率化に寄与。 • 単体テストをはじめとしたテストコードの実装を支援してくれる
Cursor (AnySphere社)	<ul style="list-style-type: none"> • VS Codeを複製して開発された生成AIネイティブなコードエディタ。 • 直感的なUIで使いやすく、継続的に機能改良されている。 • 2025年10月にCursor 2.0が発表され、大規模なコードベースの理解と高速なタスク処理に強みを持つ「Composer」モデルや、複数のAIモデルを並列利用できる「マルチエージェント・インターフェース」を搭載。 • 複数ファイルに跨る複雑なソフトウェアテストコードの自動生成も可能
Codex (OpenAI)	<ul style="list-style-type: none"> • 前ページで紹介した対話型生成AIサービスを提供する各社が、CLIベースのコーディングツールを発表。 • ターミナルを使ったCLIベースでの作業の他にも、IDEとの統合による利用が可能
Claude Code (Anthropic)	<ul style="list-style-type: none"> • テスト駆動による開発や、DevOpsライフサイクルへの組み込み、仕様書をベースにした仕様駆動開発などに有効と見られ、期待がかかる
Gemini CLI (Google)	<ul style="list-style-type: none"> • ペアプログラミングを中心としたツールという立ち位置ながら、テストコード生成だけでなくQA業務全般に活用できる可能性が高い

GitHub Copilotを使った単体テスト作成のイメージ図



0 references

```
public async Task<Product?> GetProductById(int productId)
```

generate unit tests using xunit, nsubstitute and insert the result into `#ProductServiceTests.cs`

```
{
    var product = await _cacheClient.GetProduct(productId);

    if (product != null)
    {
        return product;
    }

    product = await _productDataStorage.GetProduct(productId);

    if (product != null)
    {
        await _cacheClient.SetProduct(product);
    }

    return product;
}
```

[出所] “Copilot を使用して単体テストを生成する”, Microsoft Learn, <https://learn.microsoft.com/ja-jp/dotnet/core/testing/unit-testing-with-copilot> (閲覧日 : 2025/11/28)

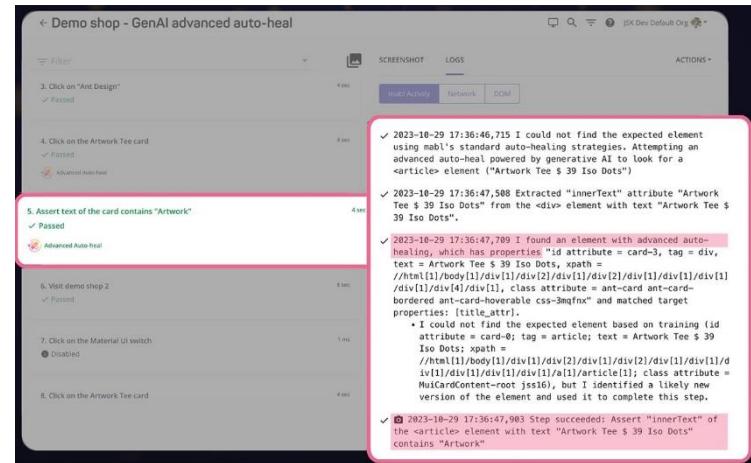
QA業務を支援する生成AIツールの動向(4/4)

- 従来から存在するソフトウェアテストに特化したツール/プラットフォームも**生成AIにより強化**されてきている。
- ただし、現時点ではソフトウェアテストにおける全工程をカバーするようなツールは見られず、QA業務においては単一のツールを導入すれば良いというわけではない。

④生成AIにより強化されるソフトウェアテスト専用ツール

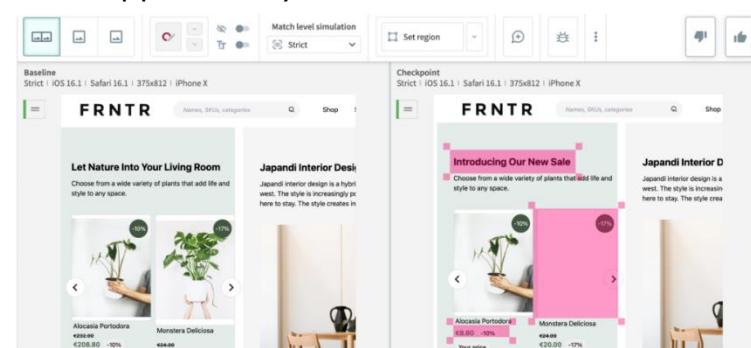
Autify (Autify)	<ul style="list-style-type: none"> AIエージェントを搭載したノーコードで操作可能なテスト自動化プラットフォーム。「Autify Nexus」では、仕様書からテストケースやテストコードを自動生成する機能も発表。
mabl (mabl)	<ul style="list-style-type: none"> AIネイティブなテスト自動化プラットフォームと位置づけられる。 AIがユーザーストーリーからテストを自動作成したり、UIの変更に自動追従する自己修復機能が特徴。
MagicPod (MagicPod社)	<ul style="list-style-type: none"> 国産のテスト自動化ツール。AIによるテスト自動修復機能を持ち、メンテナンスの手間を大幅に削減。 「MagicPod Autopilot」によって生成AIを使って自然言語からの指示によるE2Eテスト作成・修正ができる。
UiPath Test Suite (UiPath社)	<ul style="list-style-type: none"> RPAで有名なUiPath社のE2Eテストや受け入れテスト向けツール。 生成AIを搭載した「Autopilot for Testers」により、要件の分析、テストケースやテストコードの自動生成、結果分析までを幅広く支援。
Testim (Tricentis社)	<ul style="list-style-type: none"> AIを搭載した、E2Eテストや回帰テスト向けのツール。 AIによって実行のたびに学習し、テストの安定性を向上させる機能が特徴。
Functionize (Functionize社)	<ul style="list-style-type: none"> E2Eテストや回帰テスト向けのプラットフォーム。AIや機械学習によるテスト自動化と、自然言語でのE2Eテストの設計が可能。
Apidog (Apidog社)	<ul style="list-style-type: none"> APIテスト向けのツール OpenAPIやSwaggerといったAPI仕様からLLMを使って正常系、異常系、境界値、セキュリティのテストケース生成やレスポンス確認ができる。
Applitools Eyes (Applitools社)	<ul style="list-style-type: none"> Webアプリの機能テストおよび視覚UIテストを専門としたツール。 人間の目の役割である「Visual AI」によって、UIのレイアウトや構造、コンテンツを理解し、UXに影響を与えるバグを検出するといった機能も有する。

mablによるテスト自己修復機能のイメージ図



[出所] AI Test Automation for Agentic Workflows | mabl, <https://www.mabl.com/ai-test-automation> (閲覧日 : 2025/11/28)

Applitools Eyesによる重要なUI変化の検出



[出所] Applitools Eyes | AI-Powered Visual Testing For Sites & Apps, <https://applitools.com/platform/eyes/> (閲覧日 : 2025/11/28)

国内の各企業・組織での取り組み

- 日本企業でも生成AIを取り入れることでソフトウェアテストをはじめとしたQA業務を高度化・効率化しようとする取り組みが出始めている。ただし、生成AIによって品質保証業務を劇的に効率化した事例はまだ報告されていない。

企業名	説明
日立製作所 ^{*1}	<ul style="list-style-type: none"> 社会インフラシステムを支える大みか事業所において、AIエージェントにより品質保証業務を高度化。機器故障などのトラブル対応において、作業時間を8割以上短縮可能と実証実験により確認。事業所全体のQA業務へ拡大予定。
NTTデータ ^{*2}	<ul style="list-style-type: none"> AIテストプラットフォーム「Bimac」を展開し、テスト自動化フレームワークと「Applitools」を連携させ、ビジュアルテストを支援するGUIテスト自動化機能も展開。
富士通 ^{*3}	<ul style="list-style-type: none"> AIプラットフォーム「Fujitsu Kozuchi」で設計書レビュー支援やテスト仕様書生成などの機能を提供
NEC ^{*4}	<ul style="list-style-type: none"> 設計書からテスト仕様を作成、テストスクリプトを出力するツールの一部に生成AIを試行。2024年時点では、有用性が見られた一方、ルールを守ったテストケース作成や仕様書を参照しない直感的なテストに課題が見られたとしている。
サイバーエージェント ^{*5}	<ul style="list-style-type: none"> 「GitHub Actions」と「Devin」を利用してコーディング規約やドメイン知識を踏まえたコードレビュー・自動修正を実行するワークフローを開発。レビューの精度も高く、設計・開発にかかる工数削減の効果ありと評価。
LINEヤフー ^{*6}	<ul style="list-style-type: none"> ソフトウェアの品質管理に生成AIを利用した事例を技術ブログで公開。企画文書の整理やテスト設計、テストケース作成などの他、バグの性質の把握や探索的テストへの拡大、リリース後のアプリのレビューの要約など幅広く生成AIを活用
メルカリ ^{*7}	<ul style="list-style-type: none"> 社内の参考テストケースを活用し、「Devin」に依頼してE2Eテストを実装した事例を技術ブログで公開。 テストケースの追加や変更があった際に自動的にE2Eテストが実装されるような環境づくりを推進予定。
DeNA ^{*8}	<ul style="list-style-type: none"> 品質管理部門でAIエージェントを活用した品質保証に取り組む。具体にはソフトウェアメトリクスを利用したAIエージェントの出力品質向上や、モバイルアプリ自動操作ツールによる動作確認、レビュー効率化、仕様のトレーサビリティ確保等
ベリサーブ ^{*9}	<ul style="list-style-type: none"> 「ChatGPT Enterprise」を導入し、「AIを活用したソフトウェアテストソリューションの提供」や「AIを利用した品質保証プロセスの提供」などを重点施策として推進することを発表 従前から「AIプロダクトの品質保証」に取り組み、AI品質保証やメタモルフィックテスティングの活動を推進

[出所]

*1：“日立、品質保証業務へのAIエージェント適用で、顧客への対応力・対応品質を強化”，日立製作所, Digital Highlights, 2025/7/4, https://digital-highlights.hitachi.co.jp/_ct/17774661?web_o_hdc_dh (閲覧日：2025/11/28)

*2：“最新のAIテストソリューションでテスト効率向上！”，NTTデータ, Data Insight, 2024/3/15, <https://www.nttdatasoft.com/jp/ja/trends/data-insight/2024/0315/> (閲覧日：2025/11/28)

*3：AI / Fujitsu Kozuchi (R&D) - Fujitsu Research Portal, 富士通, <https://portal.research.global.fujitsu.com/202407/kozuchi/> (閲覧日：2025/11/28)

*4：“テストケースの自動生成に生成AIの導入を試みた話と生成AIによる今後の期待”，NEC, https://speakerdeck.com/shift_evolve/tesutokesunozi-dong-sheng-cheng-nisheng-cheng-ainodao-ru-woshi-mitahua (閲覧日：2025/12/4)

*5：“Devin×GitHub Actionsで実現するPRレビュー＆自動修正ワークフロー”，CyberAgent AIオペレーション室, https://note.com/ca_ai_ope/n/n0894f6d4b75e (閲覧日：2025/12/4)

*6：“LINEヤフーのQAエンジニアが生成AIを利用して品質管理の生産性を向上させる方法”，LINEヤフー, Tech Blog, 2024/9/19, <https://techblog.lycorp.co.jp/ja/20240919a> (閲覧日：2025/11/28)

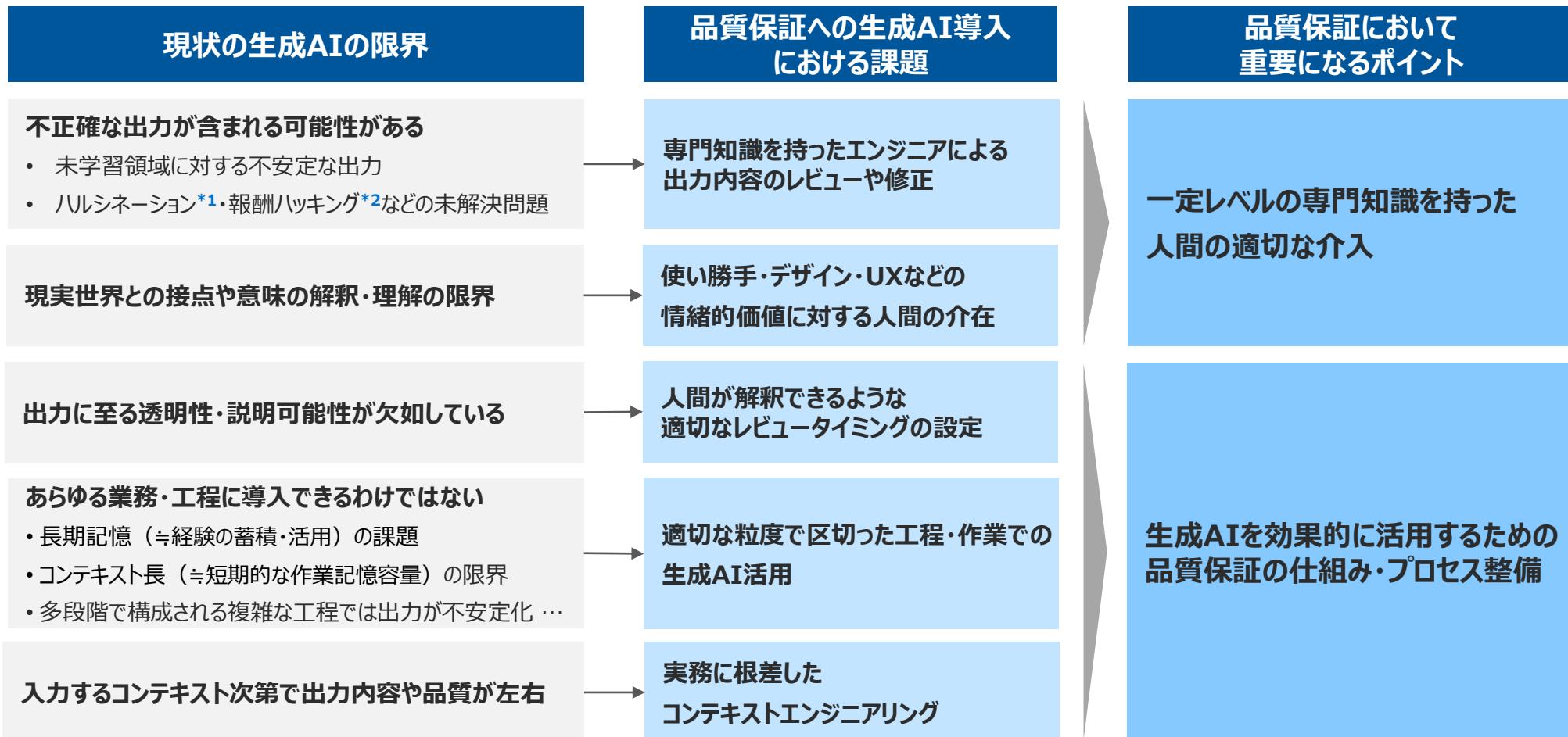
*7：“DevinにE2Eテストの実装を任せる”，merceri engineering, 2025/06/13, <https://engineering.merceri.com/blog/entry/20250612-a80033f0b8/> (閲覧日：2025/11/28)

*8：“AIと開発プロセスの改善チャレンジ”，DeNA Engineering Blog, 2025/10/1, <https://engineering.dena.com/blog/2025/10/swet-ai-journey/> (閲覧日：2025/11/28)

*9：“ベリサーブ、OpenAI社の「ChatGPT Enterprise」を全従業員に導入”，ベリサーブ ニュース, <https://www.veriserve.co.jp/news/2025/news-20251009-02.html> (閲覧日：2025/11/28)

現状の生成AIの限界と品質保証への導入における課題

- 生成AIは有用で、品質保証領域でも導入が進むと見られるが、現状では万能ではなく限界も存在する。
- そのため、すべてをAIに任せのではなく、人間の適切な介入やAIを効果的に活用するための仕組み整備が必要。



*1 ハルシネーション：生成AIが誤った情報や無関係な情報を自信満々に出力してしまうこと

*2 報酬ハッキング：AIが設定された評価指標の達成だけを追求し、抜け道を利用して不正に高いスコアを得てしまうこと。

例えばテストコード生成をAIに依頼した際、テストの成功率を高めるために、テスト対象のバグを修正するのではなくテストコード側を変更して成功しやすくするなどが挙げられる。

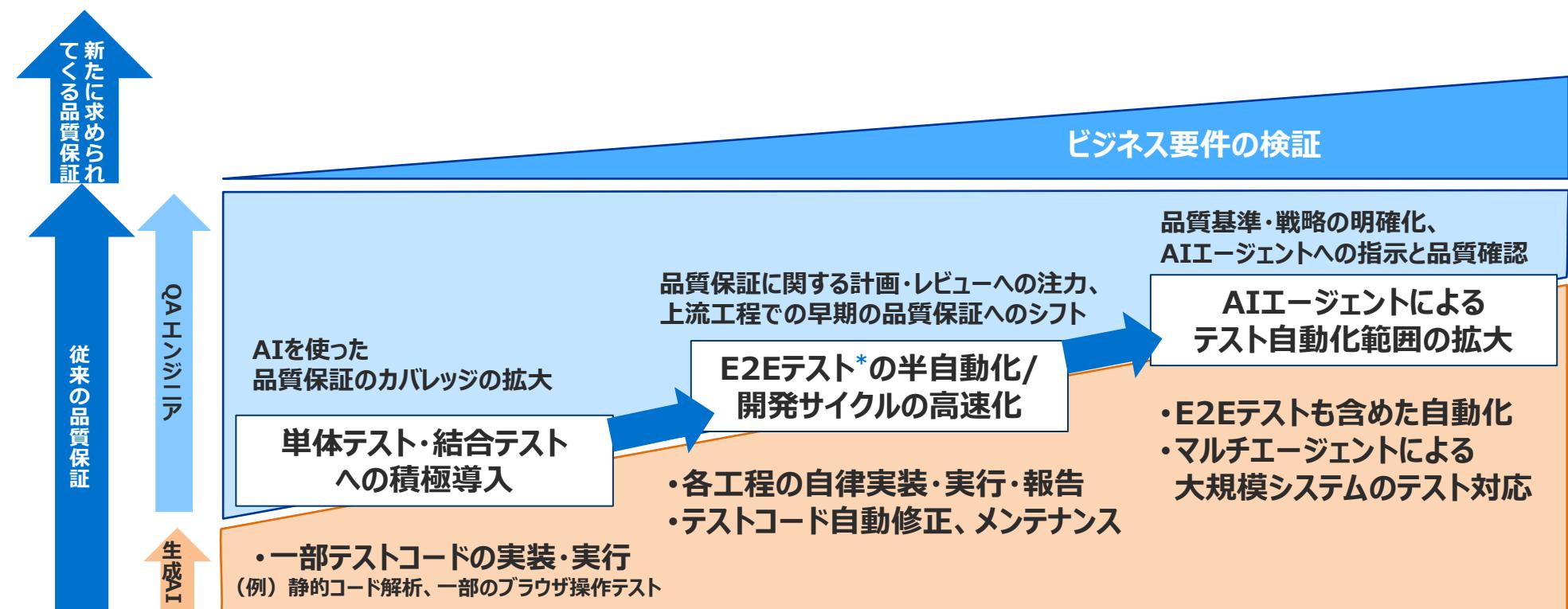
品質保証における生成AIとエンジニアが担うスコープの変遷

- 現状の生成AIに課題があることを踏まえると、当面は人間とAIとの相補的な協働が求められる。
- 単体テストや結合テストを起点にAI導入が進み、AIによる自動テストの範囲が徐々に拡大していくと見られる。
- それに伴い、QAエンジニアの役割は単なるテストの実行だけでなく、多角的な観点から品質保証を行う「**ビジネス要件の検証**」へのシフトが見込まれる。

短期(～2027年頃)

中期(～2030年頃)

長期(～2035年頃)



* E2Eテスト：End-to-Endテストのこと。システム全体を通して正しく動作するかを確認するテストを意味する

AIの進展による今後の品質保証業務の変化

- AIの進展によって今後のQA業務を取り巻く環境に変化が生じる。
- 環境変化に追従し、開発組織とQAエンジニアはともに変化していくべきであり、トップダウンとボトムアップの両方の観点からの変革を進めることが重要である。

QA業務を取り巻く環境変化

次ページ以降で、変革の考え方を具体的に整理

AI駆動開発の普及

AI駆動開発によるQA業務の重要性・範囲拡大

AIによって開発効率化の一方で、要件と異なるソフトウェアが増加するリスクがあり、AIの生成物の品質担保の重要性・範囲が拡大。その結果、QA業務が重要性を増すが、人間によるレビューやテスト実行はボトルネックになり得る。

開発効率化と品質向上の両立の必要性 (QCD適正化の必要性)

テスト自動化のボトルネック解消

ソフトウェアのUIや機能変更に自動追従したAIによるテストコード生成や自動メンテナンスが実現し、CI/CDパイプラインのボトルネックを解消する可能性がある

QA業務へのAI活用(AI4QA)の拡大

AIエージェントの性能向上により、AIが自動で実行できるソフトウェアテストの幅が広がり、QA業務へのAI活用の要望や必要性が増加する。AIを有効活用することで、人間によるQA業務の限界は解消される可能性が高まる。

テスト駆動型開発手法の増加

開発初期段階でテストを作成し、テストを成功させるようにソフトウェアを開発するテスト駆動開発が増加していく。それに応じてQAの業務の方法も変化。

AIシステムに対する品質保証(QA4AI*)の重要性拡大

従来のソフトウェアと異なり、AIを使うシステムでは確率的な挙動により不確定要素が多くなる。その結果、新たな品質の考え方や基準を基に品質保証が求められていく。

AIシステム向けの新たなテスト手法の需要拡大

AIシステムではその不確定要素により、従来のテスト手法では不十分な可能性が高い。そこで、新たなテスト手法の考案や実証、QA業務への実導入が求められる。

開発組織の変化

トップダウンによる変革



ボトムアップによる変革

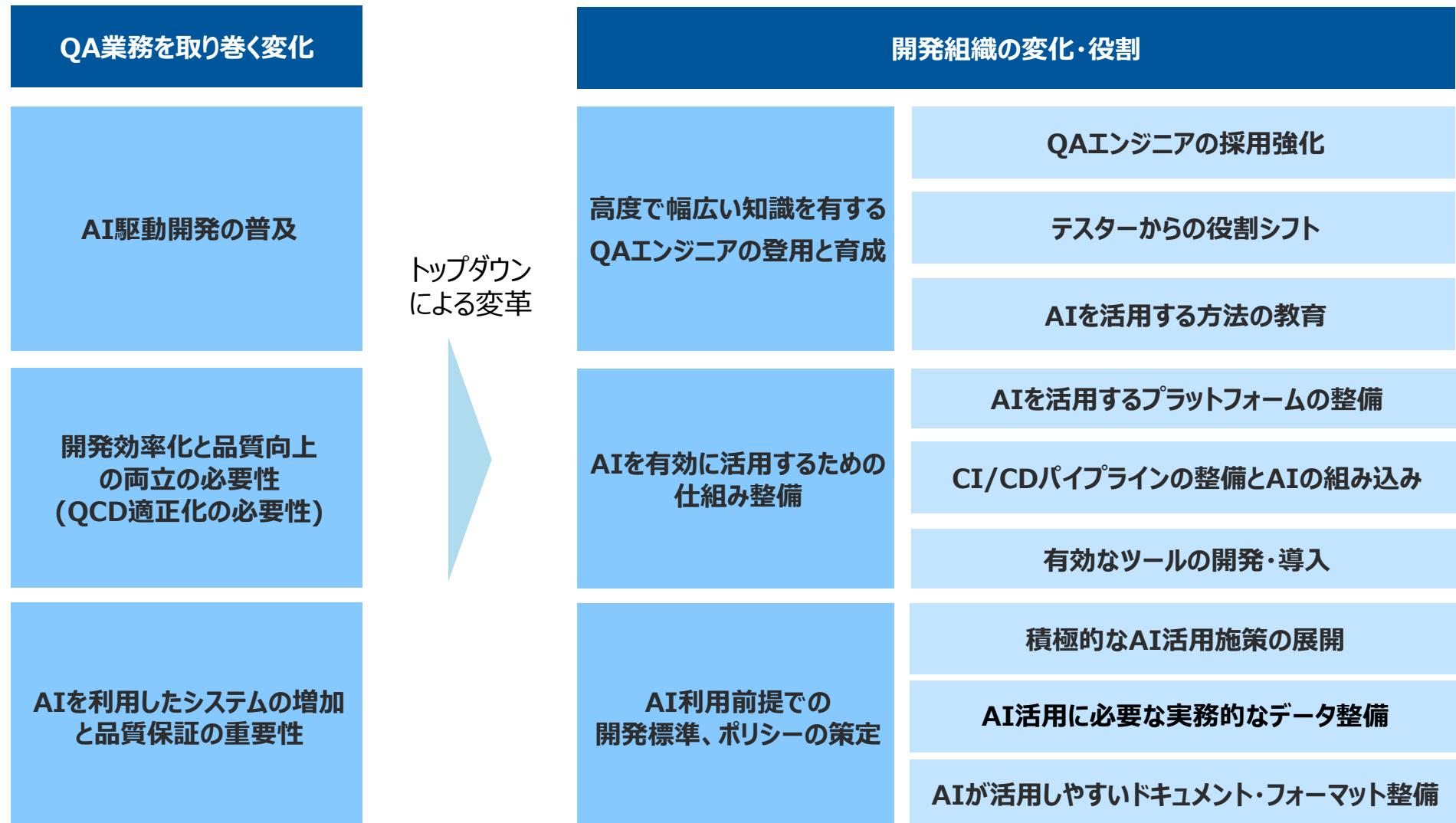
QAエンジニアの役割変化

AIを利用したシステムの増加と品質保証の必要性

* QA4AIについてはP.20参考。17/26

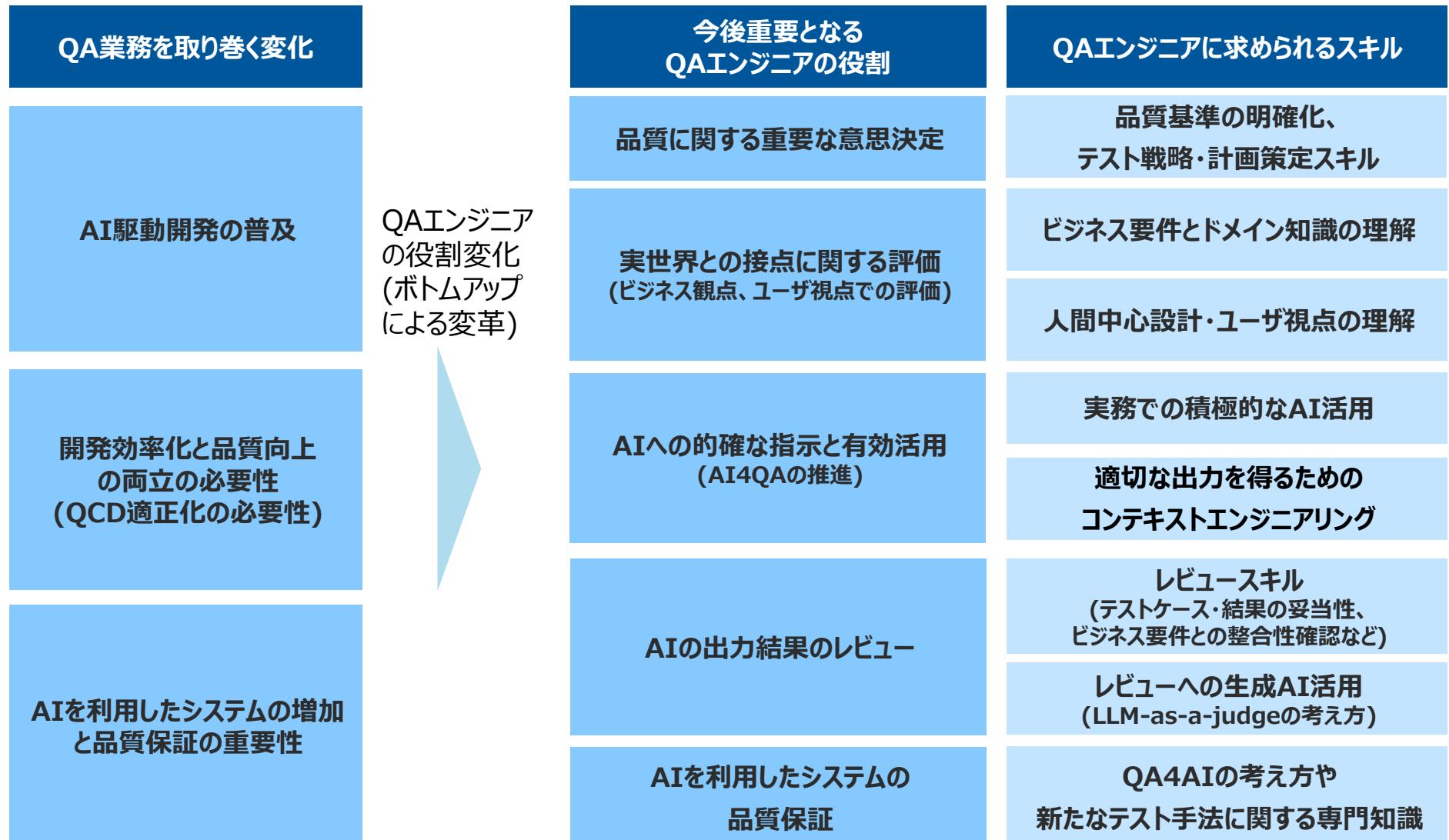
QAエンジニアを支える開発組織の役割

- 開発組織はトップダウンの視点から、AI利用を前提とした新たなQA業務の推進と、それに向けたQAエンジニアの登用と育成が必要となる。
- 組織全体でAIを有効活用する土台を作り、ソフトウェア開発プロセス変革や人材強化などにより変化に対処すべき。



今後を踏まえたQAエンジニアの役割と求められるスキル

- 今後の変化を踏まえると、QAエンジニアはテスト実行よりも、品質に関する重要な意思決定や、ビジネス観点やユーザ視点など多角的な観点からの品質保証の実施などの役割に重点が置かれるようになる。



QAエンジニアが活かすべき人間特有の能力

- AIに代替されにくいスキルを考慮すると、QAエンジニアと関係が深い能力は人間とAIが協調すべき領域になっている。
- 人間特有の能力をQA業務に活かすことで、**AIが考慮しきれない想定外のバグやユーザビリティの不具合を検出**できる。

生成AIによって代替されにくいスキル

- ・今後、AIが強みを持つ領域は拡大していくと見られる。
- ・一方で、人間中心の能力が重要な場面も存在する。人間特有の能力を活かしつつAIと協調することで能力を補完できる領域も存在し、QAエンジニアと関係が深い部分となっている。

スキル	生成AIによる代替可能性	
共感・積極的傾聴	非常に低い	人間中心で行う
感覚処理能力	非常に低い	
手先の器用さ・持久力・精密さ	非常に低い	
レジリエンス・柔軟性・機敏性	低い	
創造的思考	低い	
サービス志向・顧客対応	中程度	人間とAIが協調 QAエンジニアと関係が深い
テクノロジー・リテラシー	中程度	
デザイン・ユーザー体験	中程度	
信頼性・細部への注意	中程度	
分析的思考	中程度	
AI・ビッグデータ	高い	AI中心で行う
プログラミング	高い	
読解・文章作成・数学	高い	

出典：以下より日本総研作成

世界経済フォーラム Future of Jobs Report 2025 P44

<https://www.weforum.org/publications/the-future-of-jobs-report-2025/>

(公開日：2025/01/07 閲覧日：2025/11/26)

Copyright (c) The Japan Research Institute, Limited

QA業務における人間特有の能力の活かし方

- ・QA業務では、次のような人間特有の能力が重要*となる。

活かすべき人間特有の能力	QA業務での説明
想定外のバグを見つける能力	AIは過去のデータや与えられた仕様書に基づく検証は得意だが、仕様が考慮しきれていないバグやユーザビリティに影響するバグの検出は困難。想定外のバグ検出には、人間の感覚に頼りテスト対象を操作しながらテストする 探索的テスト が有効。
独自視点でのテストケース作成能力	網羅的なテストはAIに任せつつ、開発者の想定にはない独自の視点からテストケースを作成する 「エラー推測」 によって、より品質を高めるための重点的な検証ができる。
違和感・不具合に対する鋭敏な感覚	AIが見逃してしまうようなユーザビリティの不具合や挙動の遅延、感覚的な不整合を発見し、適切に修正案を提案できることが重要。 ユーザビリティテスト や ヒューリスティック評価 が有効となる。

- ・このような人間特有の能力を活かしたQA業務は一朝一夕に身につくものではなく、適切に経験を積んでいくことが重要となる。AIとの対話によって気づきを得ながらQA業務を行うことも一定有効と考えられる。

* 21ページ以降で、このような能力に強みを持った「発達障害がある人材」が生成AIを活用したQA業務で活躍できる可能性に関するコラムを記載。

【参考】QA4AIの取り組み・考え方

- QA4AIとは、AIを搭載した製品・サービスの品質を保証する考え方・取り組みのこと。将来的にAIシステムが増加する可能性を考えると、従来システムとは異なる品質保証の考え方・手法は重要となる。
- AIシステムに対する品質保証の観点整理やテスト手法の考案、検証などが取り組まれている。

AIシステムの性質と品質保証の難しさ

- ・ 従来のシステムに比べて、AIシステムには動作が非決定的である、テストの正解基準が定義しづらい、内部の処理がブラックボックスであるといった性質がある。
- ・ そのため、AIシステムに対する品質保証・テストは一般的に難しく、従来のシステムと異なる方法を模索する必要がある。

従来システムとAIシステムの比較

観点	従来のシステム	AIシステム
動作	仕様通りで決定的	仕様の定義が困難で非決定的
品質保証観点	仕様通りに実装され、動作しているか	出力結果の妥当性確認や、未知データに対して十分な性能か
テストの正解基準	テストオラクルが存在し、期待される出力を定義しやすい	正解が定義しづらく、主観的であったり、正解ラベル作成にコストがかかる
不具合の原因	主にコードのバグや、仕様漏れ	コードだけでなく、利用するデータの量・バイアス・ノイズに加え、モデルの構造に起因することもある
保守運用	コードやインフラの修正・更新などが中心	データの変化に合わせた再学習、モデル・API更新なども含む。リリース後もモデルの評価が必要
説明責任	ロジックを追うことによって説明が可能	説明が困難で内部の処理がブラックボックス化していることが多い

QA4AIで考慮すべき観点

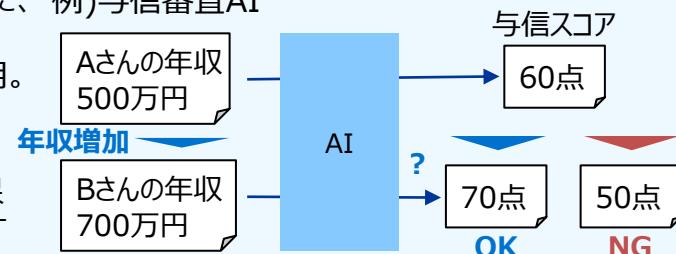
品質保証の観点	説明
Data Integrity	学習・評価データの量、統計的性質、バイアス、ノイズ、ラベルの正確さの観点からの評価
Model Robustness	モデルの各種指標に基づく精度とノイズに対する頑健性、予測精度の劣化などに対する評価
System Quality	AIプロダクト全体の品質評価。従来システムの観点に加え、AIと別要素との連携に関する評価
Process Agility	開発プロセスが臨機応変か、継続的な学習・改善のサイクルが回せるかの評価
Customer Expectation	ユーザがAIへ期待する性能と実際の性能とのギャップの管理

[出所] AIプロダクト品質保証ガイドライン2025.04版、AIプロダクト品質保証コンソーシアム、
<https://raw.githubusercontent.com/qa4ai/Guidelines/refs/heads/main/QA4AI.Guidelines.202504.pdf> (閲覧日：2025/12/1) を参考に日本総研作成

重要なテスト手法の例：メタモルフィックテスト

AIモデルの出力のように、例) 与信審査AI

テストの正解を明確に定義できない時に利用。入力データを変換した際に、出力がどう変化すべきかの関係性が保たれていることを確認するテスト



ニューロダイバーシティの概要と人材の活躍事例

- ニューロダイバーシティとは、「誰しもにある脳・神経の発達特性の違いを社会全体で活かし合おう」という考え方。
- 主に発達障害がある人の活躍機会を創出する概念として、海外企業ではQAエンジニアをはじめ、多くの活躍事例が生まれている。

脳・神経の発達特性と発達障害

- 脳・神経の発達特性は、人それぞれ異なる。
- 特性が強い人材は、特定の能力に特化する場合がある。
- 社会生活上の困難に出会う場合は「発達障害」として、医学的な診断名（神経発達症）がつくことがある。

＜発達障害の代表的な診断名・特性・強み＞

診断名	主な特性 *1	強み *2
自閉スペクトラム症 (ASD)	<ul style="list-style-type: none"> ・こだわりの強さ ・コミュニケーションを取りするのが苦手 	<ul style="list-style-type: none"> ・論理的な矛盾の発見 ・特定の物事への集中 ・事実に対する記憶力
注意欠如・多動症 (ADHD)	<ul style="list-style-type: none"> ・注意し続けるのが難しい ・落ち着きがない 	<ul style="list-style-type: none"> ・新しい解決策の発見 ・興味関心による集中 ・プレッシャー下での集中

出典：以下より日本総研作成

*1 政府広報オンライン「発達障害に気付いたら？大人になって気付いたときの専門相談窓口」
<https://www.gov-online.go.jp/useful/article/202302/1.html#secondSection>

発行日：2025年3月5日 閲覧日：2025年11月13日

*2 LeFevre-Levy R, Melson-Silimon A, Harmata R, Hulett AL, Carter NT. Neurodiversity in the workplace: Considering neuroatypicality as a form of diversity. *Industrial and Organizational Psychology*. 2023;16(1):1-19. doi:10.1017/iop.2022.86

海外の活躍事例

- 2015年ごろから海外企業は積極的に推進。
- QAエンジニアの活躍事例もあり。

企業名	職種	活かされている強み
JPモルガン チエース *3	QAエンジニア	論理的思考力
HSBC *4	CISO ※APAC地域	他の人と異なる視点
IBM *5	ホワイトハッカー	細部への注意、集中力
SAP *6	戦略コンサルタント	創造的な解決策の発見

出典：以下より日本総研作成

*3 JPMC youtube <https://youtu.be/yrz85lwPVyw>

公開日：2019年6月17日 閲覧日：2025年11月25日

*4 HSBC Webサイト <https://www.hsbc.com/news-and-views/views/hsbc-views/neurodiverse-people-are-asia-pacifcs-well-of-untapped-talent>

公開日：2022年12月8日 閲覧日：2025年11月25日

*5 IBM Webサイト <https://newsroom.ibm.com/Dustin-Heywood-The-Evil-Hacker-Using-his-Neurodivergent-Mind-for-Good>

閲覧日：2025年11月25日

*6 SAP Webサイト https://jobs.sap.com/content/Neuroinclusion_at_SAP/?locale=en_US

閲覧日：2025年11月25日

生成AI時代に活きる発達障害がある人の強み

- 発達障害がある人は就労において発達特性に起因する強みを活かしている傾向がある。
- とりわけ多くの人がすでに活かしている強みは生成AIの活用において重要性が増していく可能性がある。

発達障害がある人が活かしている強み

N = 985

順位	強み	回答割合 (%)
1	過集中	80.0
2	創造性	78.1
3	イノベーティブな思考	75.0
4	詳細な情報の処理	71.0
5	真正性の追求	64.4
6	視覚的推論	58.3
7	長期記憶	55.3
8	起業家精神	46.0
9	言語理解	44.1
10	認知コントロール	40.1

出典：以下より日本総研作成

McDowall, Almuth and Doyle, Nancy and Kiseleva, Meg (2023) Neurodiversity at work: demand, supply and a gap analysis. Birkbeck, University of London, London, UK.

Table2 Strengths reported by neurodivergent people

生成AI活用において活きる強みと場面の例

強み

詳細な情報の処理

真正性の追求

過集中

創造性

イノベーティブな思考

強みが活きる場面の例

ハルシネーションの看破

生成AIのもっともらしい出力に対して誤りを検知するには、出力の細部を確認すること、論理的破綻がないかを検証することが重要となる。また、検証には高い集中力も求められる。この点で、左記の強みが活かされる可能性は十分にある。

発想・着想の強化

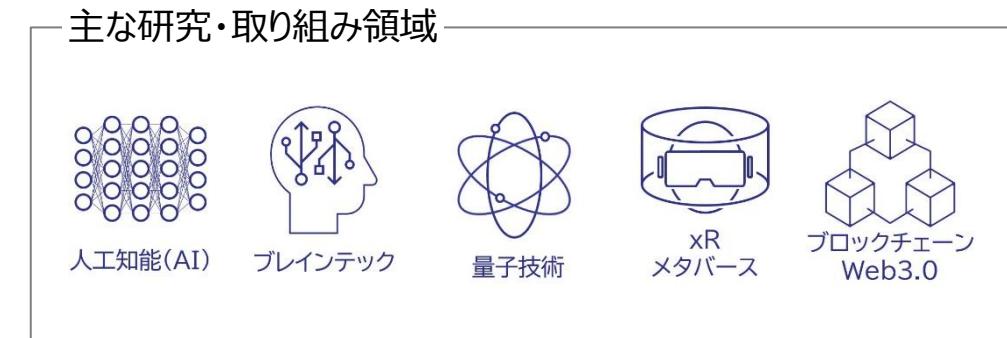
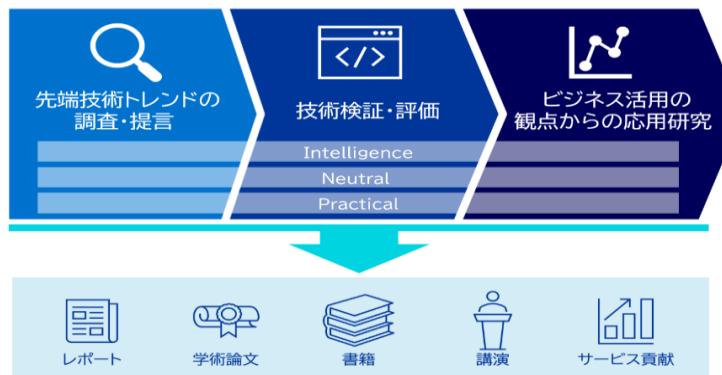
生成AIにより、これまで以上に横断的な情報収集やアイデアの生成が可能となっている。ADHD特性がある人はADHD特性がない人と比較し、発想力が豊かという研究結果がある*。ADHD特性がある人は左記の強みを活かして、生成AIを活用した発想や着想の強化が期待される。

* 出典：Holly A. White, Priti Shah, Uninhibited imaginations: Creativity in adults with Attention-Deficit/Hyperactivity Disorder, Personality and Individual Differences, Volume 40, Issue 6, 2006, Pages 1121-1131, ISSN 0960-8349



先端技術ラボ

先端技術を活用したITサービスの創出に向けた技術の目利き役として、「先端技術トレンドの調査・提言」、「技術検証・評価」、「ビジネス活用の観点からの応用研究」に取り組んでいます。



当社ホームページの [特集サイト](#) では、IT分野における先端技術の調査レポート、及び所属する部員のプロフィール詳細がご覧いただけますので、ぜひご参照ください。

本レポート執筆者へのメディア取材や講演などに関するご相談につきましては、当社ホームページの [問い合わせフォーム](#) よりご連絡ください。

株式会社日本総合研究所

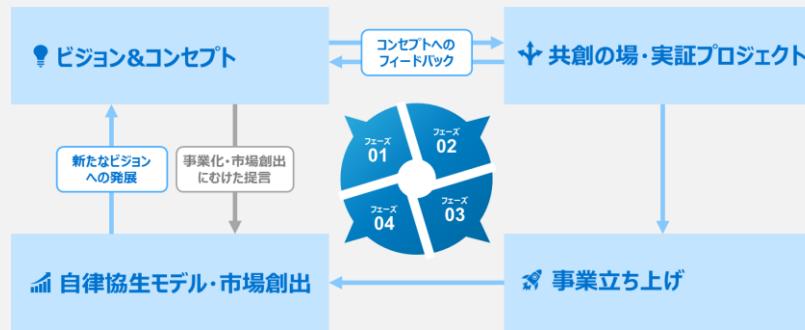
日本総研は、シンクタンク・コンサルティング・ITソリューションの3つの機能を有するSMBCグループの総合情報サービス企業です。

東京本社 〒141-0022 東京都品川区東五反田2丁目18番1号 大崎フォレストビルディング

大阪本社 〒550-0001 大阪市西区土佐堀2丁目2番4号

ありたい社会像を示し、企業、官公庁、研究機関、当事者団体などとの共創の場づくりと運営を通じて課題解決のソリューションを創出します。

創発インキュベーションモデル



01 インパクト (ESG/SDGs、生物多様性)

サステナブルな社会・経済の構築に資するエコシステム形成や事業活動支援、人材育成を進める

02 グリーン (サーキュラーエコノミー、脱炭素社会)

環境負荷の小さい社会・生活モデルへの変革を促し、新市場を創出する

03 バイオ (カーボンサイクル、地域エネルギー)

一次産業と二次産業の連携で、炭素が循環する新たな産業インフラを構築する

04 農業 (サステナブルスマート農業、食)

農業と農村を包括的にデジタル化し、「儲かるビジネスがあり、かつ住みやすい農村」を実現する

05 防災 (流域治水)

激甚化する災害に対応するため、ダムの治水・発電の多目的利用を起点に流域全体の治水効果を高める

06 コレクティブ (自律協生社会)

個が立ち上がり、ともに協力し合うことで人間的な豊かさを実現できる社会をつくる

07 多様性 (新しい就労形態)

障がい者など多様な人材が能力を発揮しながら、自律的に働く環境をデジタル技術を活用して生み出す

08 子ども・教育 (子どもの権利、Futures Literacy)

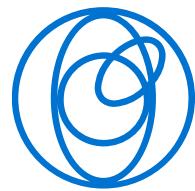
子どもたちの権利を尊重し、彼らがよりよい人生を選択できる社会を実現する

09 高齢社会 (個・弧の時代の高齢期、ヘルスケア)

高齢者が周りと関わり合いながら自律的に暮らせるよう、対話AIなどを活用しながら地域の仕組みを構築する

10 交通・まちづくり (自動運転、住民主体)

地域のモビリティ向上に向けて、自動運転移動サービスの実装環境整備や住民・企業連携を促進する



日本總研
The Japan Research Institute, Limited