

JaSST'26 Tokyo

テクノロジーセッション 14-2

“AI駆動開発に対するQAを考える会”を立ち上げました！

株式会社ベリサーブ クラウドサービス事業部

白川 亮太

2026年3月20日（金）

本日本日お伝えしたい内容

- AI駆動開発における品質について、定期的に情報交換できる場を作りました！
- これまでの活動内容を紹介し、どのような会なのかをお伝えします。

自己紹介

- 白川 亮太（しらかわ りょうた）
- 経歴
 - ▶ 2008年新卒でベリサーブに入社。
 - ▶ WebサービスやSIer案件を中心に、テスト業務、PMO、テストプロセス標準構築支援などに従事。
 - ▶ 2025年から顧客に対してAI駆動開発に対する品質標準作成支援を行っている。
- 趣味
 - ▶ ボードゲーム作り
 - ▶ ゲーム



AI駆動開発に対するQAを考える会

概要

- ▶ AI駆動開発に対する品質保証に関するリスクや問題、それらに対する解決策などの情報交換・発信を行い、業界に貢献する。
- ▶ 過去に2回開催済。次回は5月頃を予定している。

発足の背景

- ▶ AI駆動開発が急速に浸透していく昨今の開発現場において、プロセス品質やプロダクト品質にAIがどのような影響を及ぼしているか、世の中に出ている情報が少なかった。
- ▶ ベリサーブのオウンドメディアであるHQW!は、多くの企業、エンジニアの方にご登録いただいている。このネットワークを活用し、世の中のAI駆動開発に対する品質保証の情報を集められると考えた。

登壇者
ベリサーブ クラウドサービス事業部
クラウドサービス第三課 白川 亮太

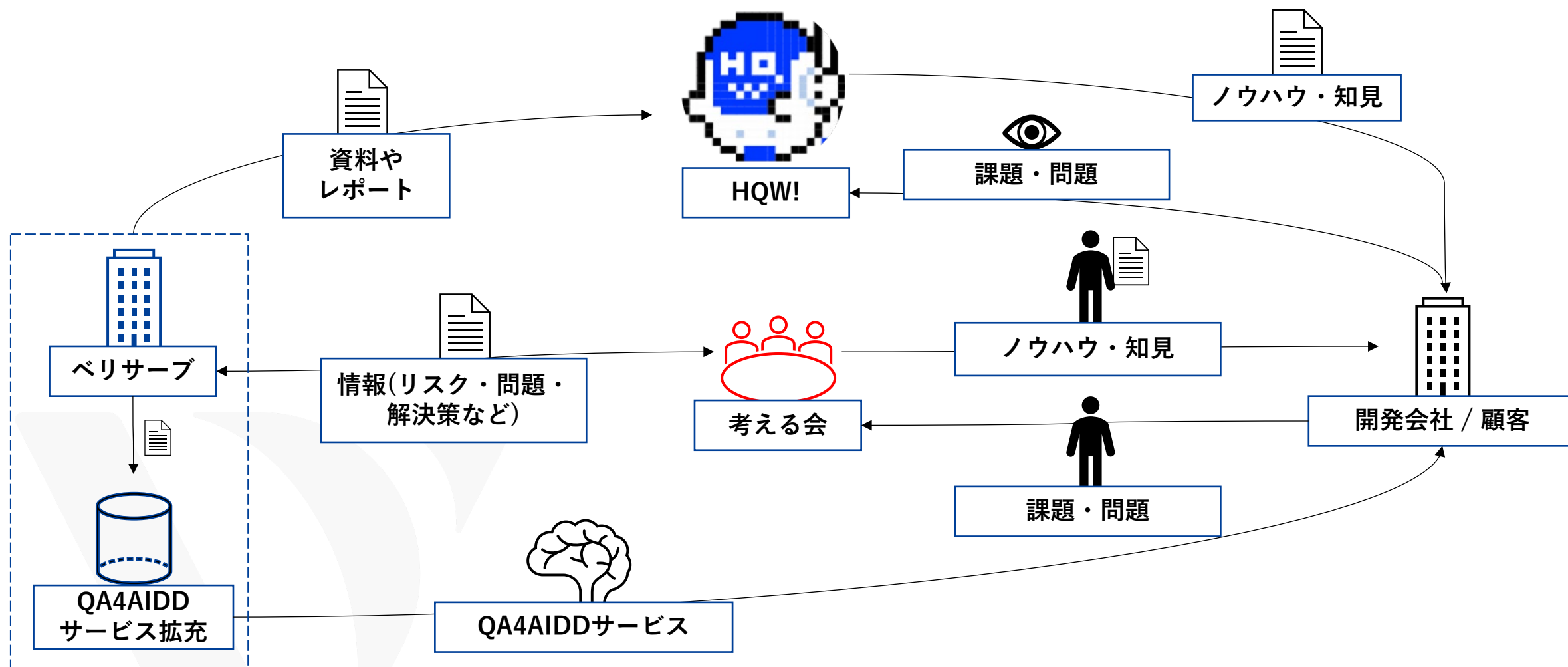
第2回 AI駆動開発における
リスクと対策の
抽出ワークショップ

2026年1月26日 月
19:00~21:00
—— 現地開催 ——

登壇者
ベリサーブ 研究開発部 上野 彩子

AI駆動開発に対するQAを考える会

・会のインプット・アウトプットサイクルイメージ



第1回 [講演型]

QA4AIDD (AI駆動開発に対するQA) とは？

AI駆動開発に対するQAを考える会

・第1回 【QA4AIDD（AI駆動開発に対するQA）とは？】

- ▶ ベリサーブ 研究開発部 AI技術戦略推進課 上野 彩子が登壇し、AI駆動開発時代のQAの在り方や課題、今後の展望について語る会。
- ▶ 講演後、1時間のフリーディスカッションタイムを設け、参加者同士で情報交換に花を咲かせた。



第1回の実績

- 開催日：2025/12/09（火）
- 参加者数：20名
- 講演資料は、HQW!にて公開中。

▶ URL：

<https://www.veriserve.co.jp/helloqualityworld/download/20260318001/>



AI駆動開発とは（本日の定義）

- 「AI駆動開発」とは、ソフトウェア開発プロセスにAIを活用し、要件定義、設計、実装、テストの各工程を支援・自動化することを目指すアプローチ。
- 自律型コーディングエージェントが進化するにつれ、AI駆動開発が一般的になっていく。

INNOVATION

Emerging Software Trends That Will Define Development By 2030



By Expert Panel®, Forbes Councils Member.

for Forbes Technology Council, COUNCIL POST | Membership (fee-based)

Published Oct 21, 2025, 01:15pm EDT

Forbes誌の「2030年までに標準となる20のソフトウェア開発手法」にAI活用とその周辺技術が数多くランクイン

Forbes. Emerging Software Trends That Will Define Development By 2030.

<https://www.forbes.com/councils/forbestechcouncil/2025/10/21/emerging-software-trends-that-will-define-development-by-2030/>

AI駆動開発の課題

資料抜粋

AI駆動開発は大幅な生産性向上をもたらす一方で、プロセス面、AIの出力の品質面において課題があることが知られている。

利点

- 生産性向上、開発サイクルの短縮
- 単純作業の自動化

課題

運用の難しさ

- 組織の開発プロセスの変更が必要

出力の品質

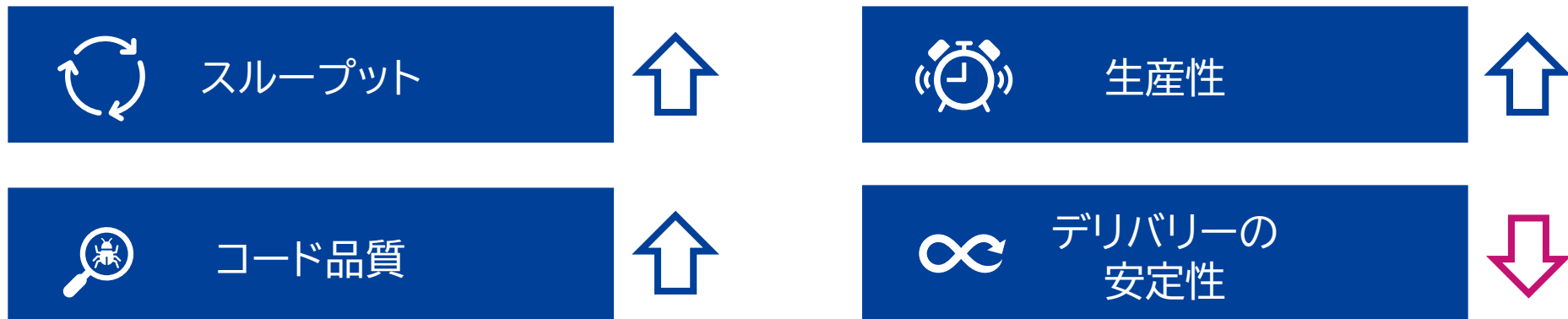
- 脆弱性のあるコードの出力
- 再現性のなさ
- 成果物同士の一貫性の欠如
- 人間の指示に反した出力

AI駆動開発の課題－運用の難しさ

資料抜粋

AI駆動開発はチームのスループットの向上、コード品質の向上、個人レベルでの生産性の向上に貢献する一方、デリバリーの不安定性を増加させているとの報告がある¹。

開発スピードの向上に対し、**管理プロセスがAI駆動開発に適応していない**ためであると考えられている。



¹ Google Cloud. 2025 DORA State of AI-assisted Software Development Report.
<https://cloud.google.com/resources/content/2025-dora-ai-assisted-software-development-report>.

AI駆動開発の課題ーコーディングエージェントの出力の品質

資料抜粋

AIはWeb上の多くの安全でないコードを学習しているため、コーディングエージェントの出力には次のようなリスクがあるといわれている。

脆弱性	<ul style="list-style-type: none">• LLMは人間の開発者に比べて脆弱なコードを生成する傾向 (Islam et al., 2024; Tihanyi et al., 2024)• メモリ安全違反やSQLインジェクションなどの脆弱性が含まれていた (Huynh et al., 2025; Sharma et al., 2025)
バイアス	<ul style="list-style-type: none">• GPT-4が生成したコードの38.92%にジェンダーバイアスが含まれていた (Huynh et al., 2025)
保守性	<ul style="list-style-type: none">• ChatGPTが生成した4,066本のプログラムのうち、1,930本のコードが保守性の問題を抱えていた (Liu et al., 2024)

AI駆動開発の品質保証①：AIに正しく指示を与える

資料抜粋

- 作業プロセスや出力に対するルール、仕様など開発のためのインプット情報をリポジトリ内に配置し、毎回参照させる。
- 指示は具体的に与える。



開発はTDDで行ってください。



Aという機能をTDDで開発したいです。
まず仕様を理解してください。

仕様に基づいて、最初に失敗するテストコードを書いてください。

先ほどのテストをパスする最小限の実装を生成してください。

全てのテストが通った後、コードのリファクタリング案を提案してください。

?



```
import pytest
from main import sum_even_numbers
```

```
def test_sum_even_numbers_basic():
    assert sum_even_numbers([1, 2, 3, 4, 5]) == 6 ...
```



QA4AIDDの実践例: AI出力の確認方法

資料抜粋

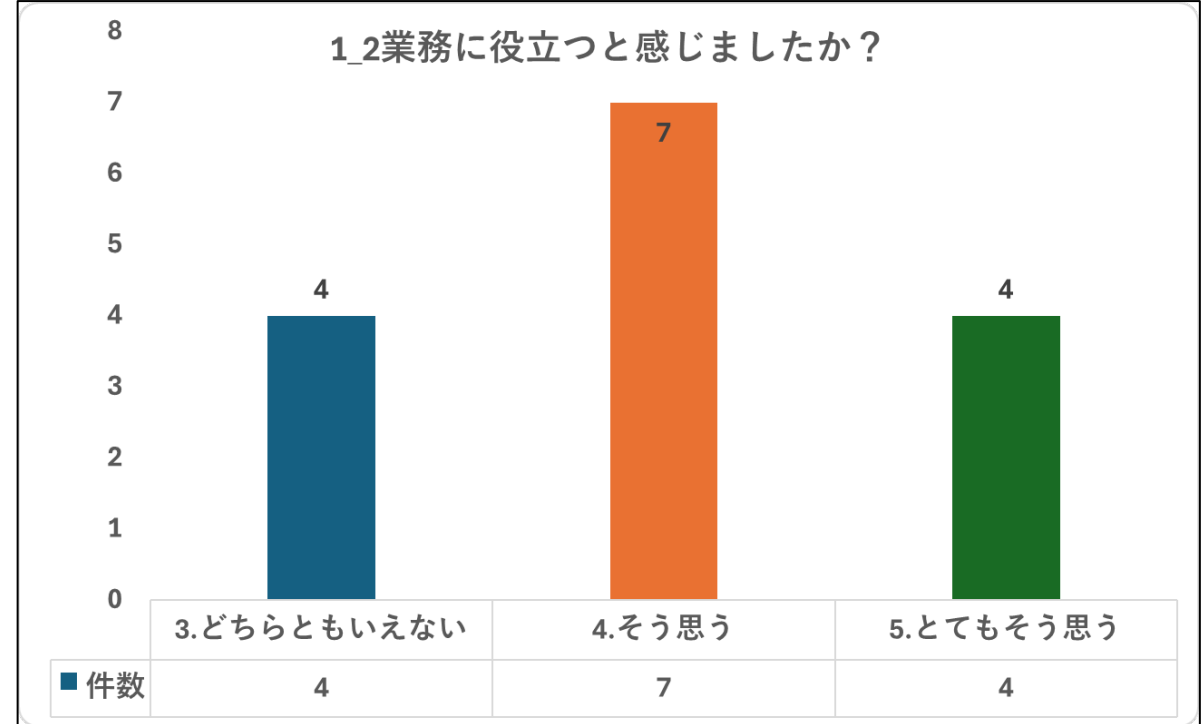
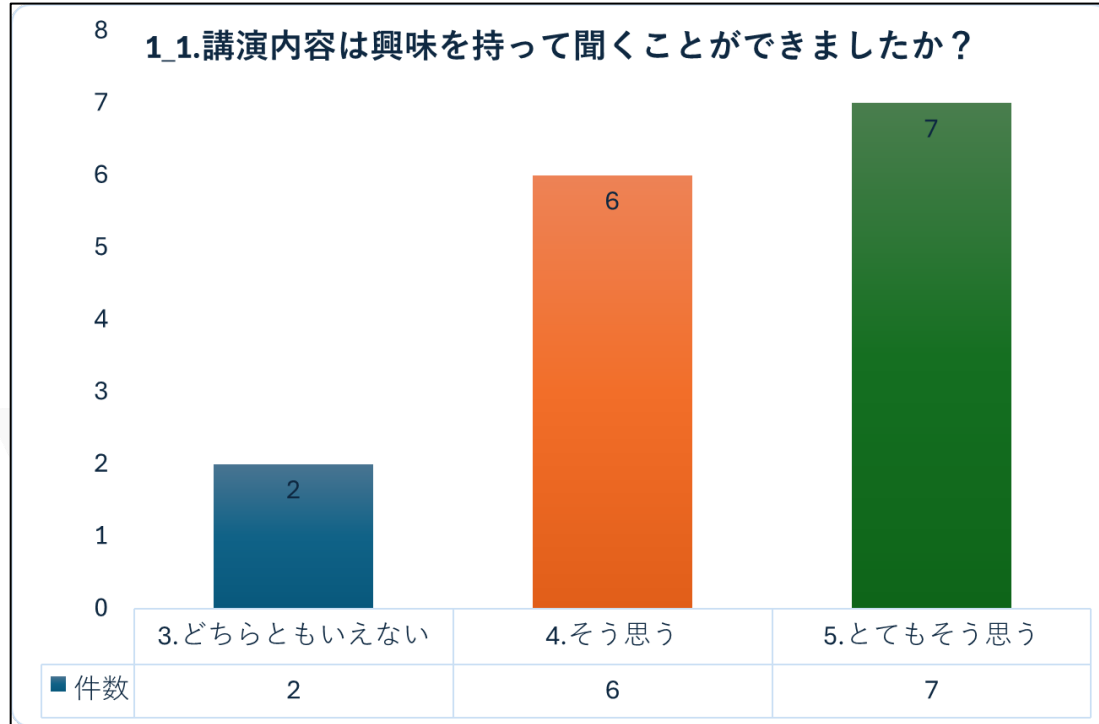
各開発工程で実施する主な活動は以下の通り。これらの活動を通し、望ましいAIの出力が得られるよう指示を継続的に改善する。

要件定義・設計	<ul style="list-style-type: none"> • プロンプトに例外ケースを列挙し、モデルに反例を探させる
コード生成	<ul style="list-style-type: none"> • 静的解析ツールによるコーディングルールの違反、潜在的なバグ、脆弱性の検出 • コードの保守性や可読性に対する指摘 • 脆弱性、保守性、可用性などのコードメトリクスを取得し、品質状況をスコア化する
コードレビュー	<ul style="list-style-type: none"> • AIに渡した観点に沿ったレビューができているか、AIエージェントの軌跡を確認する • 人が見るべきところを決めてレビューする
コンポーネントテスト	<ul style="list-style-type: none"> • テスト駆動開発のプラクティスを導入する • コンポーネントテストがない場合は、作成してからAI駆動開発を導入することを推奨
システムテスト	<ul style="list-style-type: none"> • 機能が仕様通りであるかの確認の他に、AI駆動開発特有の弱点であるセキュリティ、信頼性、保守性の観点を盛り込んだテストを行う

第1回の実績

アンケート結果（1/3）

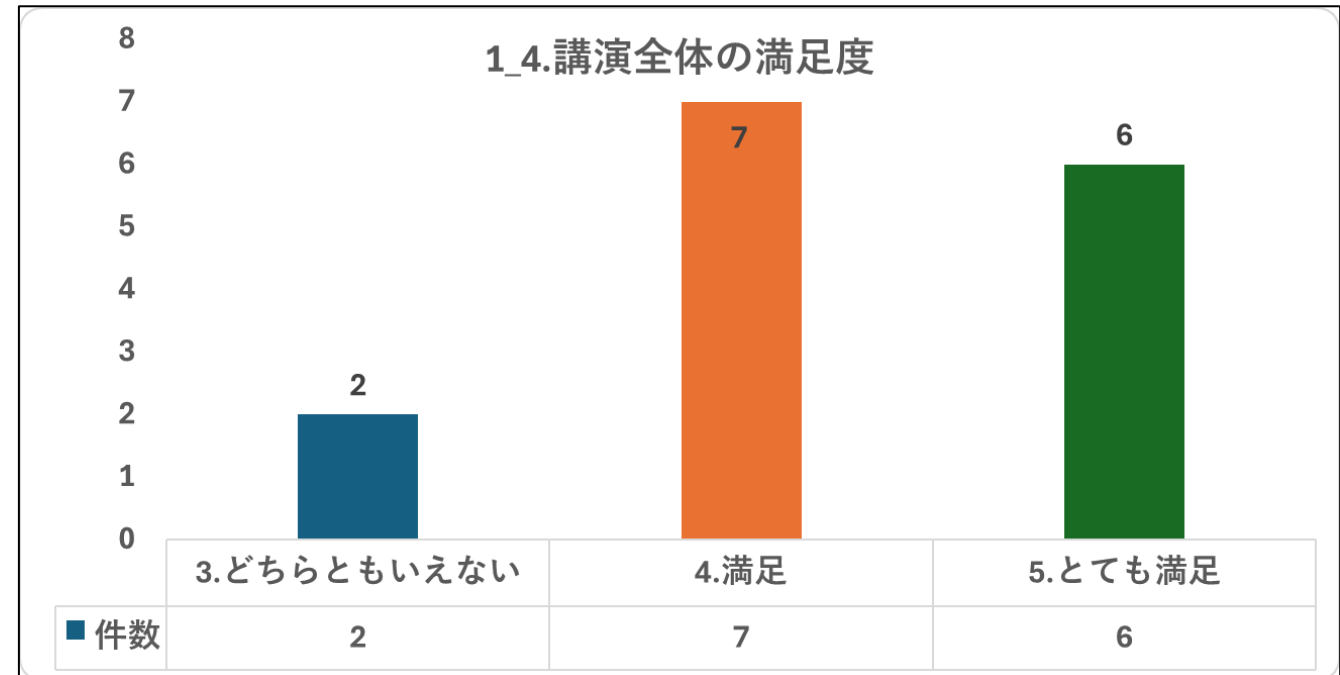
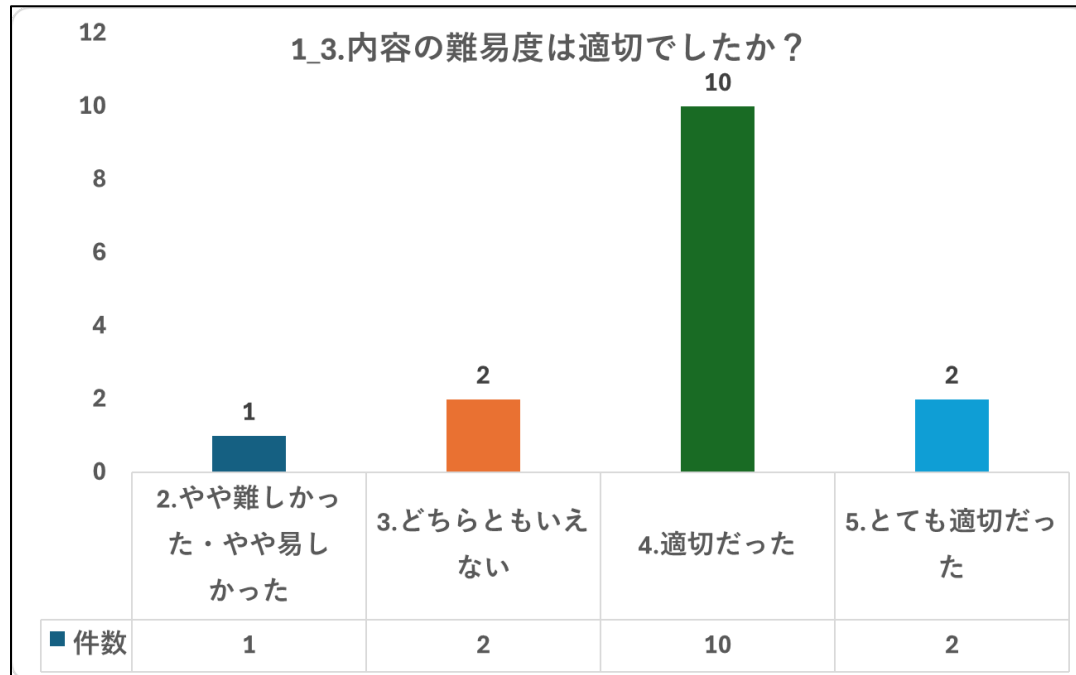
- アンケート回答率：75%（20名中15名）
- 参加者の半数以上が業務に役立つと感じており、初回としては上々の滑り出しである。



第1回の実績

アンケート結果 (2/3)

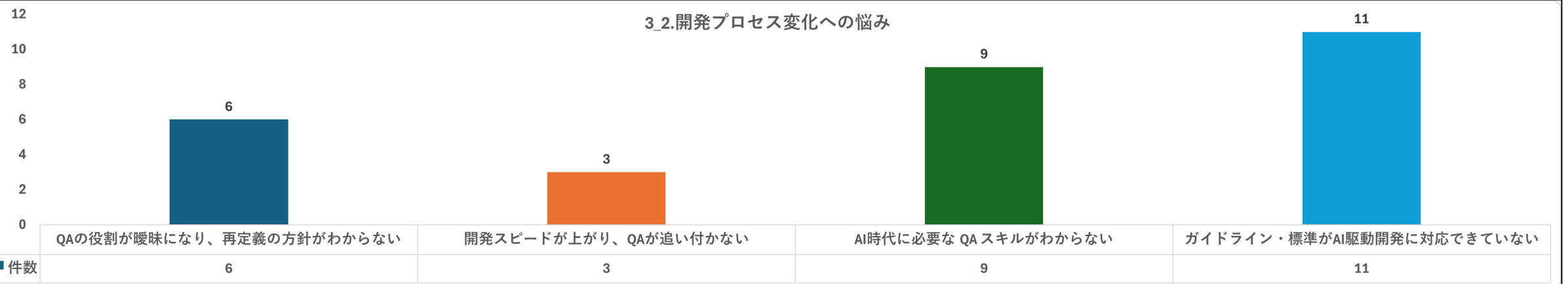
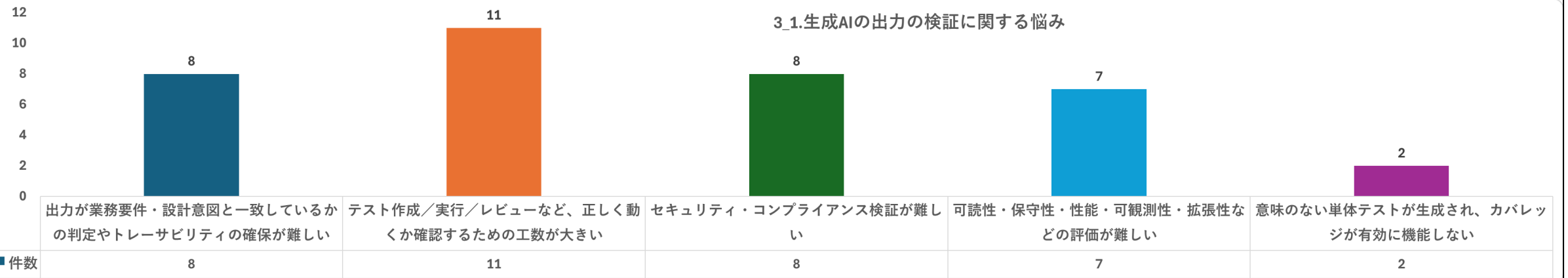
- 講演+フリーディスカッション方式となっており、情報交換が活発に行われていた。そのため、講演全体の満足度が高かったものと考えられる。



第1回の実績

アンケート結果 (3/3)

- 出力されたものの量や、質の不安定さから、レビューの難しさについて悩んでいる人が多い。
- プロセスについては、会社やPRJとしての標準がまだ対応できていないことが多い。

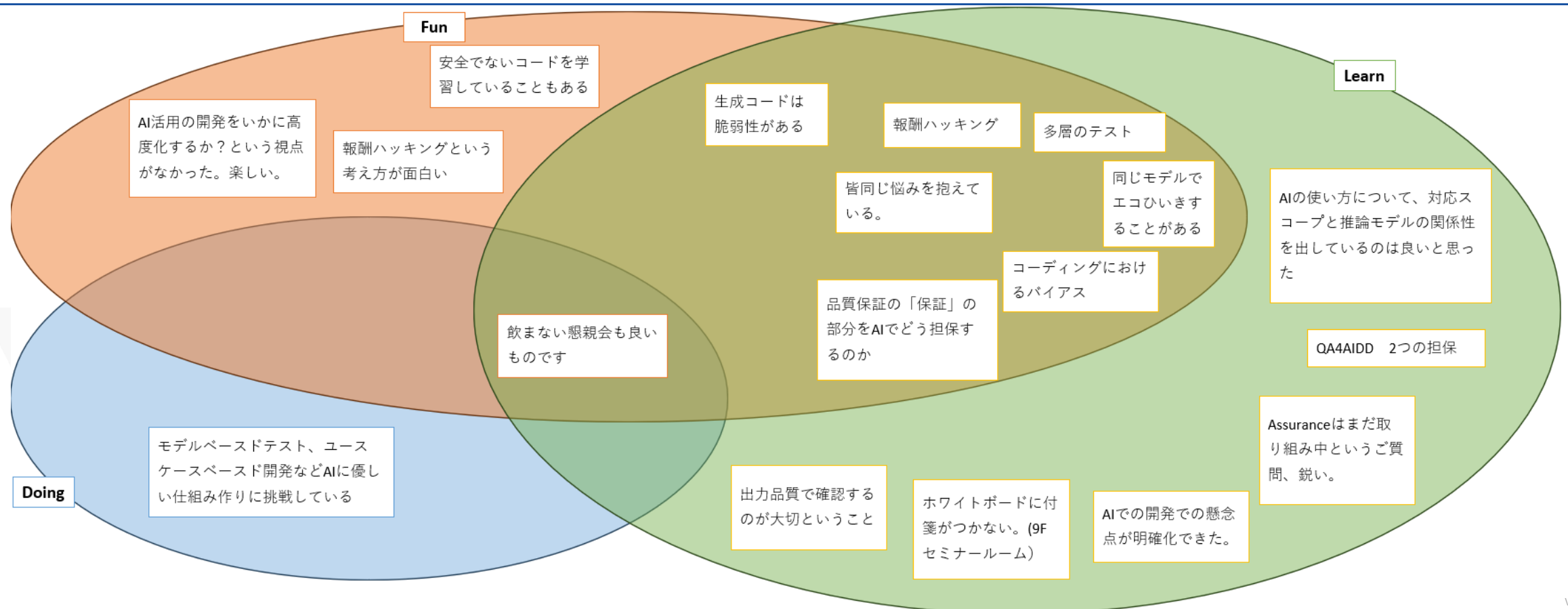


第1回の実績

• Fun Doing Learn

- Fun：面白かったこと、興味深かったこと
- Doing：自身のQA4AIDDに関する取り組み
- Learn：本日の学び

- Fun、Learnの数が圧倒的に多く、Doingはまだ少ない。
- 今後も定期的に本イベントを開催し、Doingが増えていくことを望んでいる。



第2回 [ワークショップ型] AI駆動開発におけるリスク抽出と 対策立案ワークショップ

第2回の実績

・ 第2回

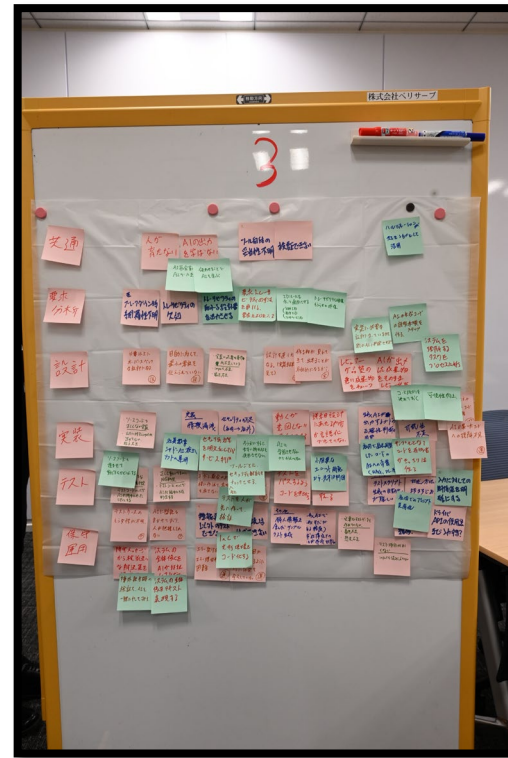
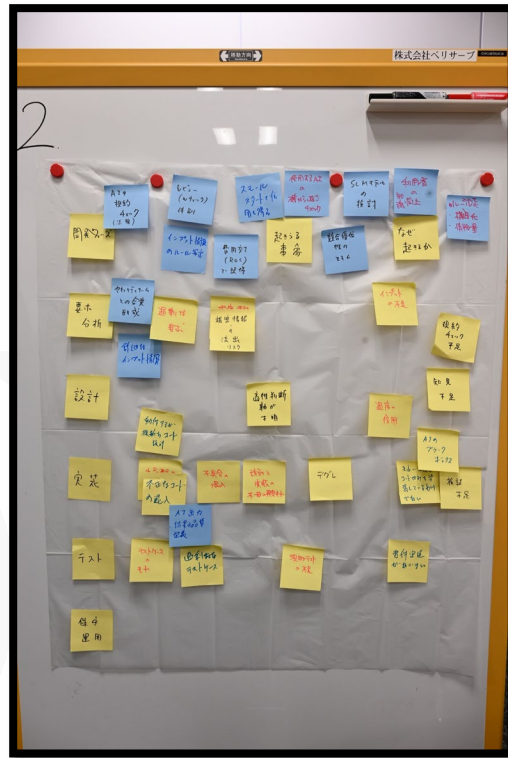
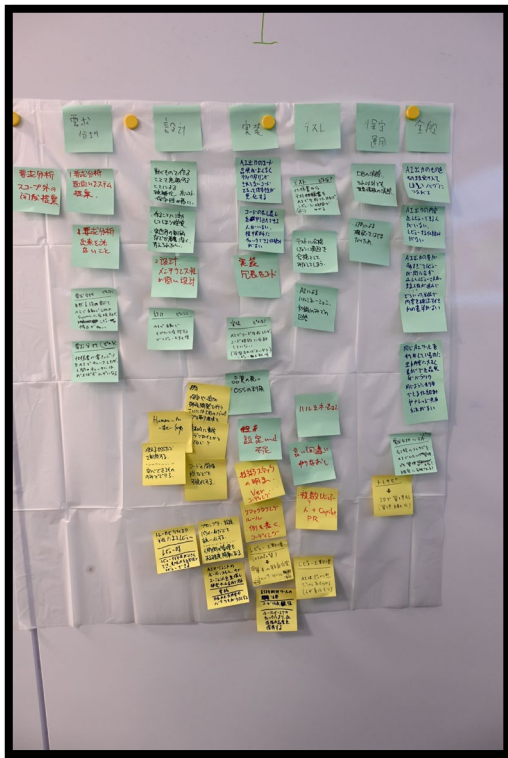
【AI駆動開発におけるリスクと対策の抽出ワークショップ】

- ▶ 参加者がグループに分かれて、AI駆動開発におけるリスクの抽出と、それに対する対策を立案するワークショップ。
- ▶ 工程毎におけるリスクを先に洗い出し、共有し合った上で、リスクに対する対策の検討を行った。



第2回の実績

- 開催日：2026/01/26（月）
 - 参加者数：16名
- ※4グループに分けてワークショップを実施



今回の題材：もし、あなたの組織でこんなことがあったら…

資料抜粋

あなたは新規プロダクト開発のプロダクトオーナーです
開発には、組織で初めてAIを活用すること（AI駆動開発）を考えています

AIを活用するに当たり上層部から大きな反対はなかったものの、
リスクが不明であるとして、リスク洗い出しおよび対策立案の依頼を受けました

どうすれば上層部に納得してもらえる説明ができるでしょうか

ワーク①： AI駆動開発におけるリスク洗い出し

リスクは以下の3点を必ずセットで考えます

どのフェーズで起きるか

- 要求分析
- 設計
- 実装
- テスト
- 保守運用

AIを使ったことでどのような起きてほしくないことが起きるか

- 開発上の具体的な事象

誰にどんな影響が出るか

- 品質
- コスト
- 責任(説明責任・法的・倫理)

ワーク②： リスクに対する対策立案

- 洗い出したリスクに対して、現実的な対策を検討する
 - ▶ ルールを策定する
 - ▶ プロセスで防ぐ
- 上層部に説明できるかを意識

第2回の実績

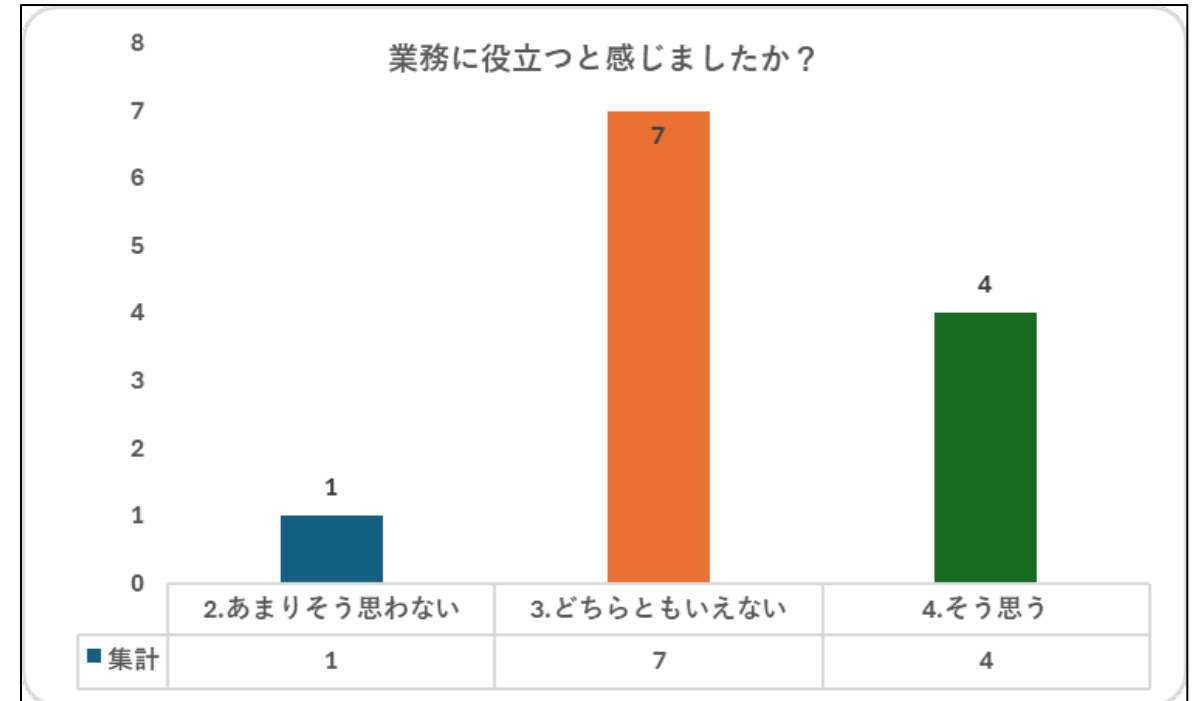
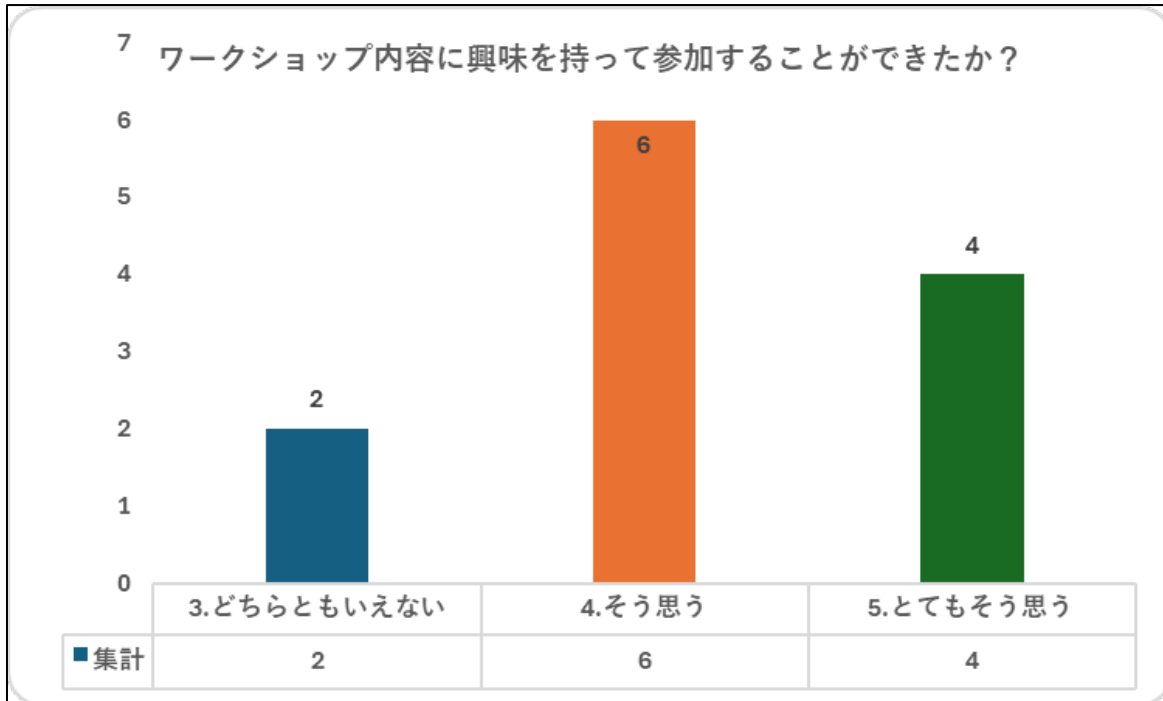
・ワークショップ結果の要約

フェーズ	主なリスク	リスクの具体像	主な対策	補足・運用ポイント
企画・要件定義	AI出力を前提条件として過信してしまう	要件・設計の妥当性をAIに委ね、人の検討が浅くなる	・ AIは補助とし、人が最終判断する原則の明確化	「AIが言っているから正しい」を防ぐ思想づくりが重要
			・ AI活用時の注意点を教育で共有	
設計	AI出力をレビューできない	設計案・設計書をAI生成したが、妥当性を判断できる人がいない	・ 設計レビュー観点の教育	AI生成物でもレビュー基準は従来と同じであることを明示
			・ レビュー責任者の明確化	
実装	AI出力をそのまま受け入れ、バグを混入	生成コードを理解せずにコミットしてしまう	・ 人によるコードレビューの徹底	レビュー省略を「効率化」と誤解しない運用が必要
			・ AIは「下書き」扱いとするルール化	
実装	出力量が多すぎてレビューが破綻	AIが大量コードを生成し、人のレビューが追いつかない	・ チェックツールの活用	人レビューを減らすのではなく「集中させる」目的で使う
			・ ルールベースでの自動検証	
テスト	AI生成テストの品質が保証できない	テスト観点漏れ・誤った期待値に気づけない	・ テスト観点の教育	AI生成テストも「テスト対象」と考える
			・ ルールベースチェック	
全体共通	AI活用スキル・体制が不足	レビューできる人、判断基準、仕組みが存在しない	・ 基本教育の充実	単発教育ではなく継続的な仕組み化が鍵
			・ AI利用ルール／ガイドライン整備	

第2回の実績

・アンケート結果（1/3）

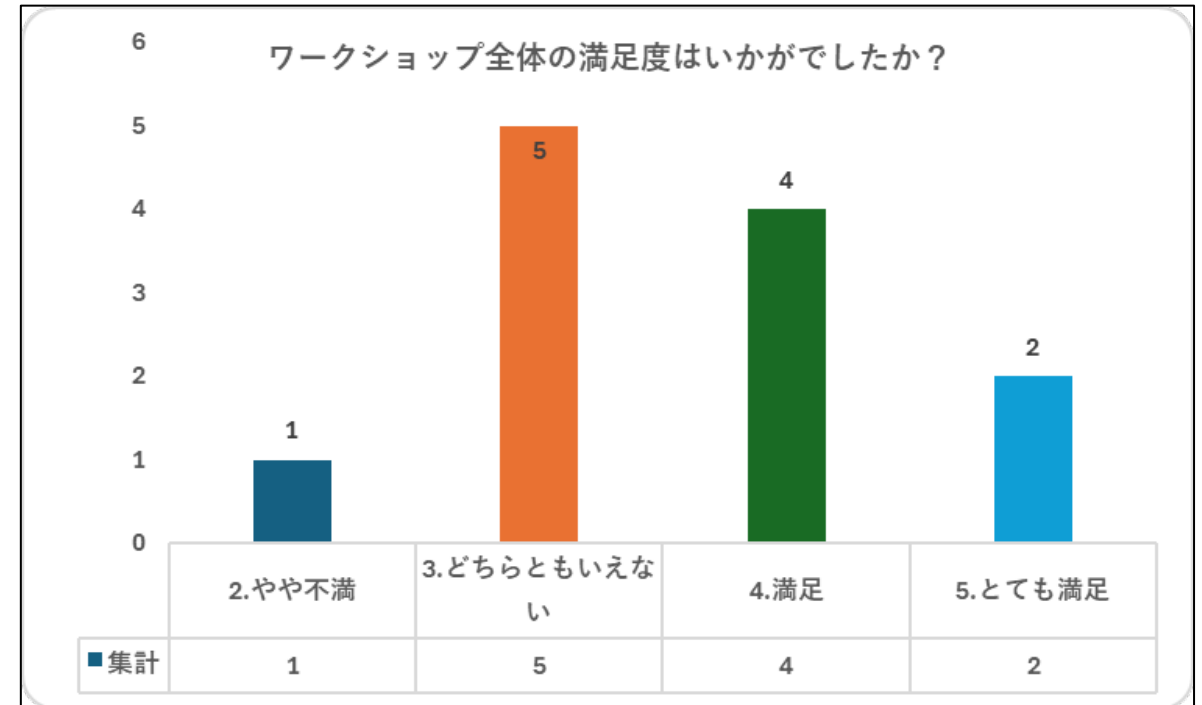
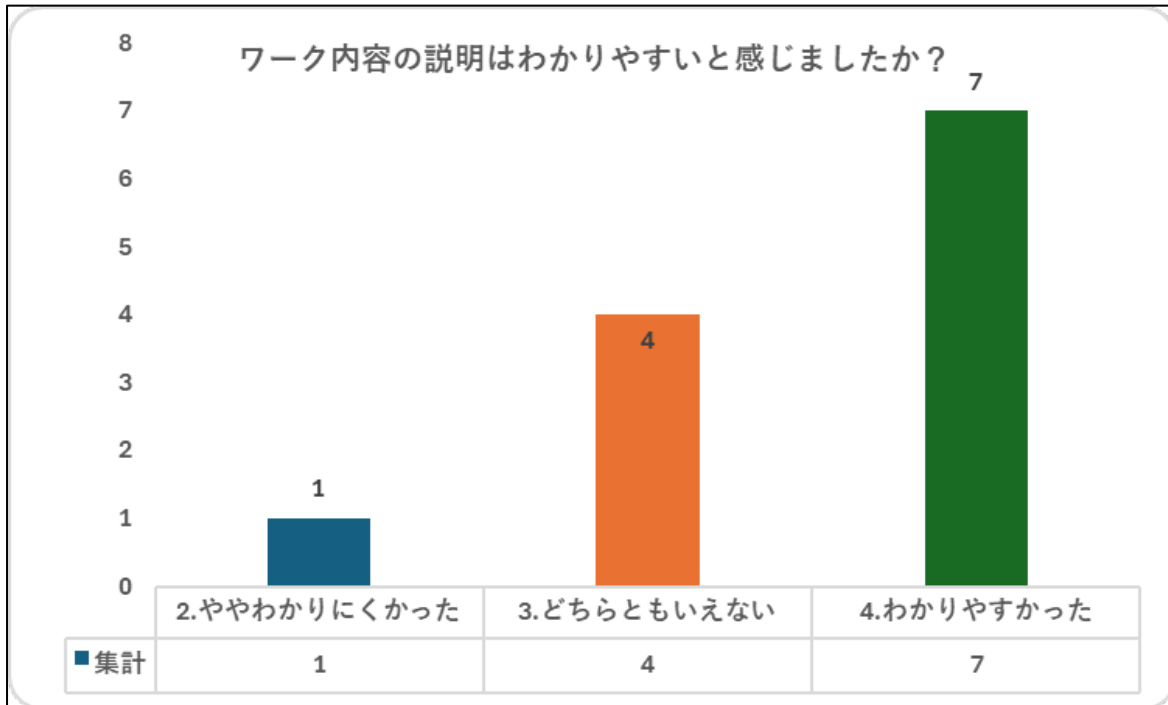
- アンケート回答率：75%（16名中12名）
- 第1回と比較し、業務に役立つと感じる人が少なかった。まだAI駆動開発をしていない方が多く、リスクと対策がイメージしづらかったと推察する。



第2回の実績

・アンケート結果（2/3）

- リスクと対策を挙げたものの、対策が見つからないリスクも多くあり、もやもやしたまま終わってしまい、満足度が高くなかったものと考察する。
- セキュリティ、膨大な出力結果に対するレビューをどうすればよいか、出力の不確実性に対する対策など、悩みが多く挙がっていた。

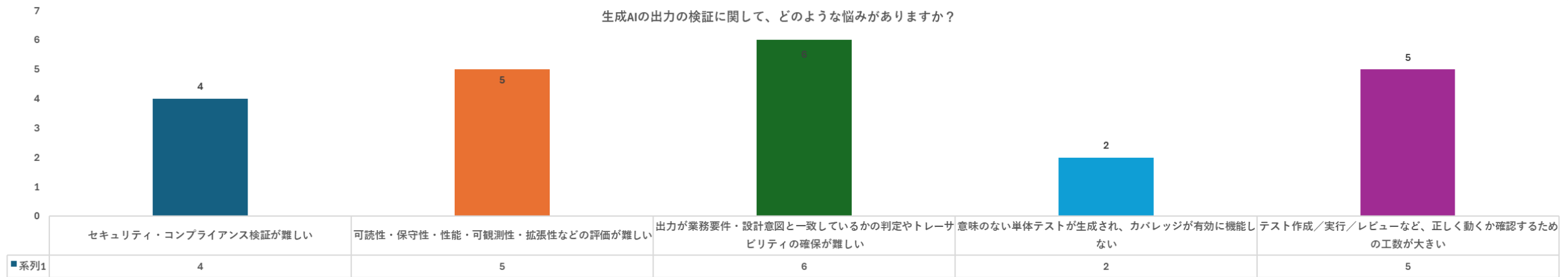


第2回の実績

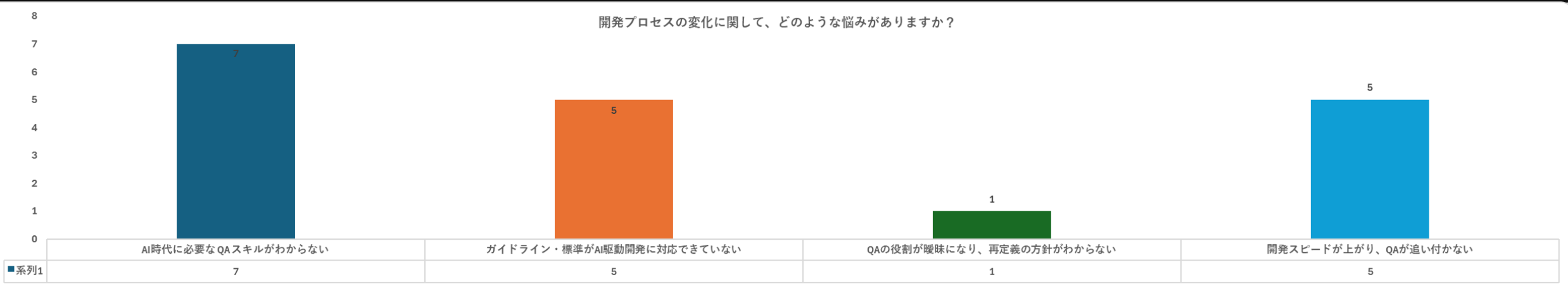
・アンケート結果 (3/3)

- 出力されたものの量や、質の不安定さから、レビューの難しさについて悩んでいる人が多い。
- プロセスについては、そもそも”AI時代に必要なQAスキルがわからない”が1位となった。

生成AIの出力の検証に関して、どのような悩みがありますか？



開発プロセスの変化に関して、どのような悩みがありますか？



今後の展望



今後の展望

・ 今後実施予定の内容 (1/2)

グループ	内容	想定形式
フレームワーク系	派生開発でのコンテキストエンジニアリング	講演、ディスカッション
	AIモデルの品質保証フレームワーク解説（NIST AI RMF、ISO/IEC 42001、EU AI Act動向など）	講演、読書会
	日本企業におけるフレームワークの適用事例の紹介	LT
事例とケーススタディ	SDD（仕様駆動開発）の事例共有	LT
	AIが出力したコードの効果的なレビュー事例	LT
	AI×各産業（自動車・金融・医療・製造）におけるQA事例	講演
	巨大なコードベースをAI駆動開発で直させた事例	講演、LT
	AI駆動開発がうまくいった事例・うまくいかなかった事例	講演、LT

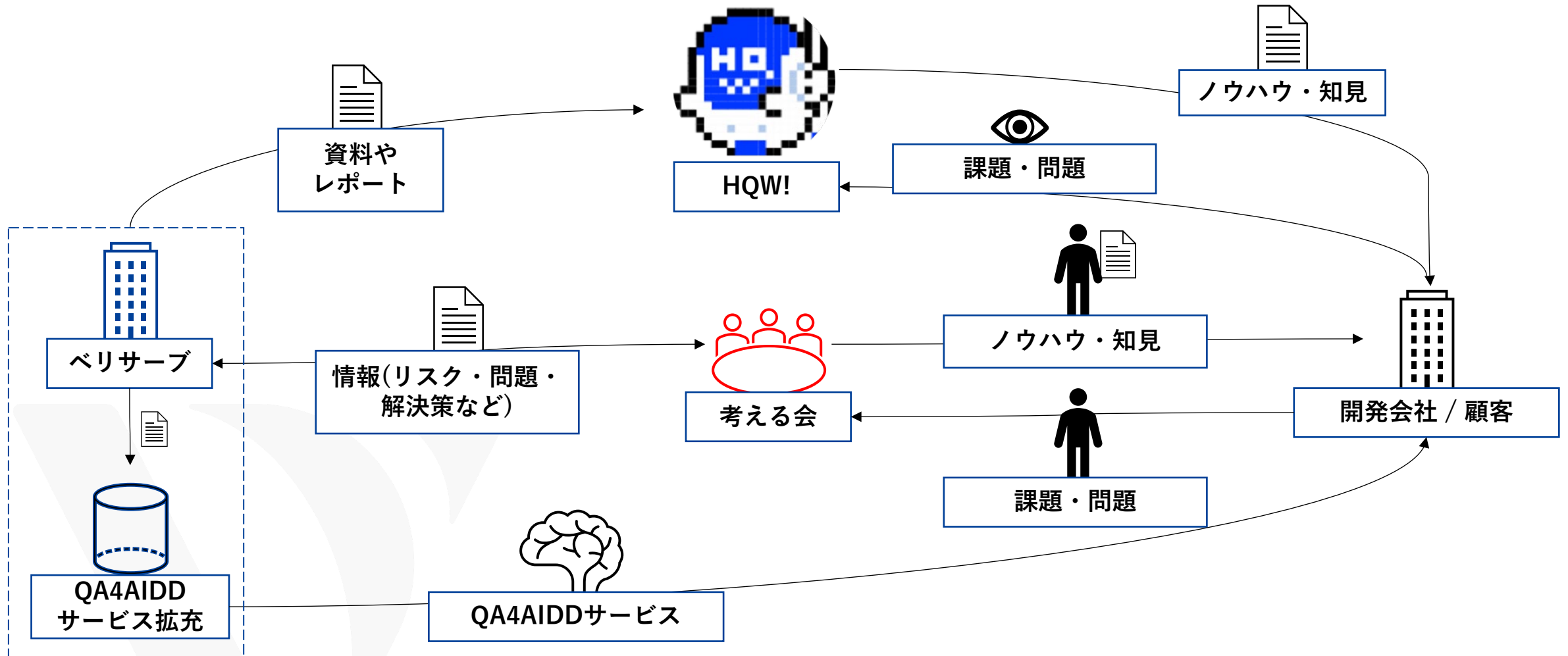
今後の展望

・ 今後実施予定の内容 (2/2)

グループ	内容	想定形式
ワークショップ・イベント	プロンプト改善ワークショップ	ワークショップ
	SDDハンズオン	ハンズオン
	AGENT.mdのベストプラクティスを持ち寄る会	研究会
	AI品質指標作成ハンズオン	ハンズオン
	テーマ別分科会の常設化	研究会
	AI品質に関するミニコンテスト	コンテスト
ガイドライン・プロセス	社内でAI駆動開発を広める工夫	LT
	AIプロダクトにおける“品質”とは？（UX観点の安全性・透明性・説明性）	講演、ワークショップ
	AI駆動型開発の標準プロセスをどう作るか？	講演、ワークショップ
	AI生成コードのレビュー手法・自動化ツールの紹介	LT
	AI駆動開発によるエンジニアのモチベーション低下、認知的オフロード問題をどうすればよいか	座談会
	AI駆動開発組織の立ち上げ方とガバナンス（プロセス品質の一種）	講演、ワークショップ

今後の展望

・会のインプット・アウトプットサイクルイメージ（再掲）



さいごに

- 次回は**5/15（金）**に開催予定です。テーマは「**派生開発におけるコンテキストエンジニアリング（仮）**」
- 次回以降の予定はHQW!、およびConnpassで発信いたします。



AI駆動開発に対するQAを考える会

さいごに

- 本日はご清聴いただきありがとうございました。
- “AI駆動開発に対するQAを考える会”は、今後も定期的に情報交換・情報発信できる場を提供していきます。
- 興味のあるタイトルがあれば、ぜひお気軽にご参加ください！



ご清聴ありがとうございました

株式会社ベリサーブ
クラウドサービス事業部
白川 亮太

26-001



イノベーションを加速させる
知恵と品質技術にアクセスする
テクノロジーライフメディア

www.veriserve.co.jp/helloqualityworld/