

# ロボティクスSWGの 取組状況と今後の進め方について

## ロボティクスSWG

国立研究開発法人 産業技術総合研究所  
ウェルビーイング実装研究センター  
副研究センター長 中坊 嘉宏

2025年10月2日

AISI事業実証WG 上期報告会

**AISI** Japan  
AI Safety Institute



## 中坊 嘉宏 Yoshihiro NAKABO

国立研究開発法人 産業技術総合研究所  
研究戦略本部 ウェルビーイング実装研究センター  
副研究センター長

- 2000年 東京大学計数工学専攻博士課程修了 博士（工学）
- 2005年 産総研入所。その後は介護ロボット、サービスロボット、産業用ロボット、AI搭載ロボット等の安全にかかる研究開発に従事。
- 2025年から ウェルビーイング実装研究センターにて働く人のウェルビーイングに資するロボットの研究を開始。
- ISO TC299 ロボティクス、IEC TC125 小型電動配送車ほかの標準化委員会WG主査、また国内の農機や公道走行配送車、人機械協調安全に関する委員会委員等を務める。

## AIロボティクスへの期待

- ◆ 近年、生成AIやマルチモーダルAIの発展により、ロボティクス分野においても、AIの利用が急速に拡大。
- ◆ AIを実装したロボット（AIロボット）の社会実装が進むことで、産業界では効率化や生産性向上が期待される。また、ロボットの対人インタラクション能力向上により、接客等サービス業においては高付加価値を提供できる可能性。
- ◆ さらにロボット基盤モデルやフィジカルAI、エンボディドAI等の、AIロボットに関する新たな研究開発も進展している。今後は、いっそう汎用性を増したロボットが登場し、応用範囲と利活用ケースはますます拡大発展していくと見込まれる。



出所) 三菱総研「AI・ロボティクスの現在地とこれから2024年10月号」

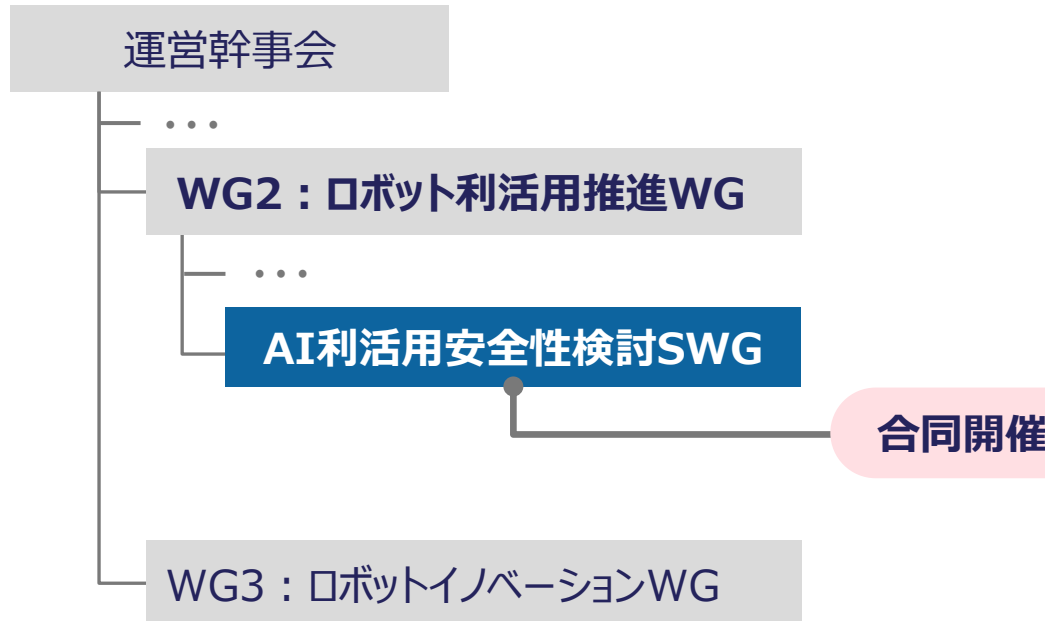
## リスク・課題

- ◆ 従来のロボットにおける衝突や接触等の物理的リスクに加えて、AIが実装されることで、コミュニケーションの際の威圧的発言等の心理的リスクや、偽誤情報拡散等の社会的リスクも想定される。
- ◆ 機械安全や機能安全ではカバーされないリスクに対しての、追加的な安全設計の指針が求められる。

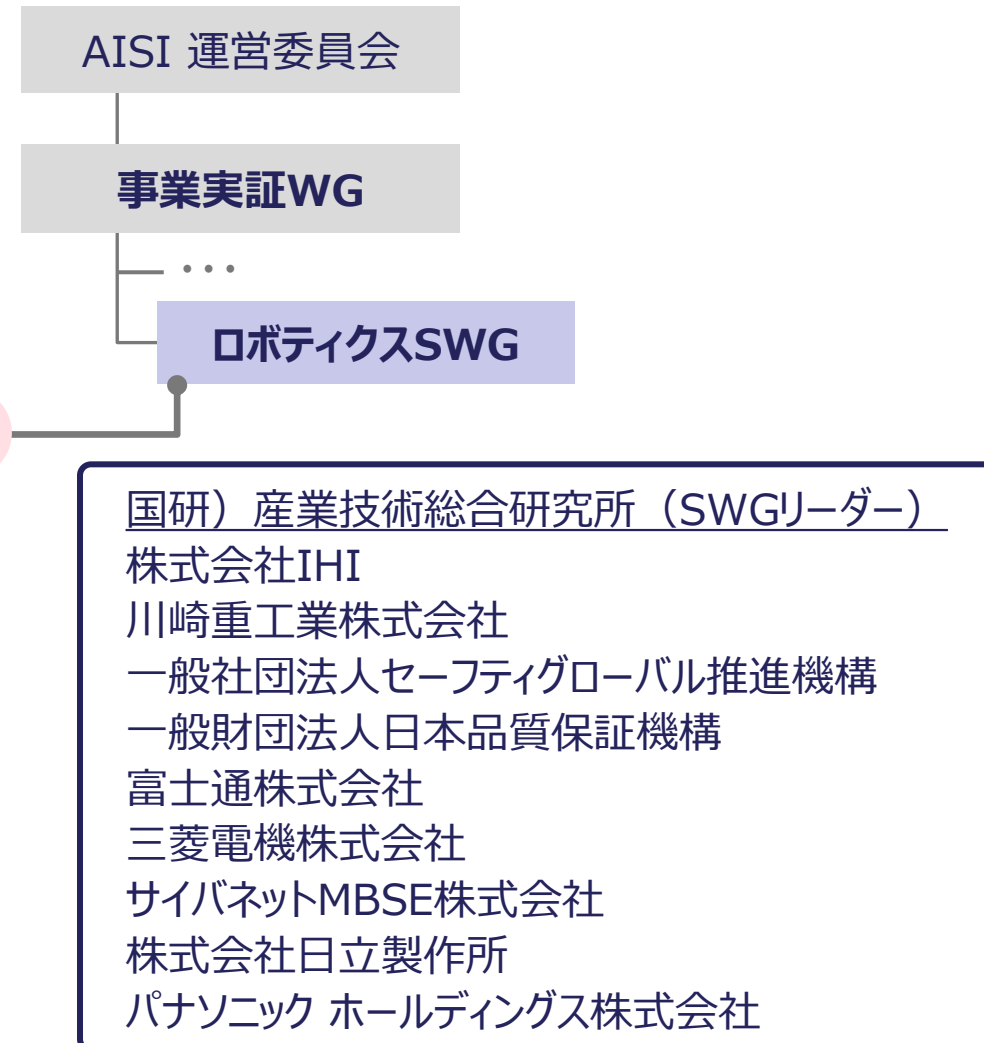
**AIとロボットの複合的な振る舞いが生む便益と、そこに内在するリスクのバランスを捉えた評価の枠組みの構築が重要。**

⇒ ロボティクスSWGでは、社会がAIロボットを安全かつ安心して利活用することを促進するため、開発メーカーやシステム提供者、研究機関等と連携して、より実用に近い応用例からAIセーフティ評価の模擬環境と仮想シナリオによる実証を通じたロボット類型ごとの多層的評価を進め、将来の標準的な枠組みの確立を目指す。

## ロボット革命・産業IoTイニシアティブ協議会（RRI）



## AIセーフティ・インスティテュート（AISI）





- 近年、対話/移動/操作型のサービスロボットや協働ロボットにおいて、認知・判断・行動の一連の知的処理を担うAIの役割が実用に近づいている。
- AIEージェントやエンボディドAI等が発展し、人とロボット/ロボット同士のコミュニケーションやコラボレーションが進む。
- AIにより新たに生じうるリスクに対して、AIロボットに係る安全の基準が明確でないため、利活用範囲の拡張に慎重。

|          |                                                                                                                                                                                  |                                                                                                                                                                                                                               |
|----------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| AI利活用の現状 | <ul style="list-style-type: none"> <li>・対人インターフェース領域におけるマルチモーダルAI導入による人間の五感的判断代替を模索。</li> <li>・人とロボットの協働／ロボット同士の制御・協働。</li> <li>・「Embodied AI」の進展による、状況理解と行動選択の高度な融合。</li> </ul> | AIセーフティ評価の現状 <ul style="list-style-type: none"> <li>・サービスロボットの安全基準は明確化。社会への導入が進む。</li> <li>・一方、人とロボットのコミュニケーションを通じてサービスを受けるようなシーンのリスク（不適切な誘導など）に懸念はある。</li> <li>・ロボティクスに特化したAIに係る基準が明確でないため、ロボットの活用範囲の拡張に慎重になっている。</li> </ul> |
|----------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|

|             | 対話型                                                                                                                                                                                                                                                                                                                                                                             | 移動型                                                                                                                                                                                                                                                                                                                                                                             | 操作型                                                                                                                                                                                                                                                                                                                                                                                              |
|-------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 特徴          | 音声認識、音声合成、感情認識などの技術を統合し、自然言語による双方向のコミュニケーション能力を持つ。利用者の意図や感情を理解し、適切な応答を生成することが可能。                                                                                                                                                                                                                                                                                                | センサやカメラを用いて周囲環境を認識し、自己位置推定と経路計画による自律移動能力を持つ。障害物回避や動的環境への適応が可能。                                                                                                                                                                                                                                                                                                                  | 人間と一定の距離を保ち、動作を監視しつつ、必要に応じて停止し、衝突を回避する。力加減や動作速度の調整による柔軟な作業が可能であり、人間との共同作業により、作業効率の向上が期待される。                                                                                                                                                                                                                                                                                                      |
| ユースケース分類（例） | <ul style="list-style-type: none"> <li>● 接客支援 <ul style="list-style-type: none"> <li>- 顧客への商品案内</li> <li>- 自然対話による接客 等</li> </ul> </li> <li>● 介護・メンタルケア <ul style="list-style-type: none"> <li>- 音声・表情解析による感情推定</li> <li>- 対話を通じた共感・励まし 等</li> </ul> </li> <li>● 観光案内 <ul style="list-style-type: none"> <li>- 外国人観光客向けの観光案内、経路案内</li> <li>- おすすめスポットの紹介</li> </ul> </li> </ul> | <ul style="list-style-type: none"> <li>● 自動配送 <ul style="list-style-type: none"> <li>- 動的に経路を作成</li> <li>- 渋滞等の環境変化を特定し、経路変更 等</li> </ul> </li> <li>● 自動清掃 <ul style="list-style-type: none"> <li>- 障害物回避</li> <li>- 汚物等を検知し、適切な清掃方法を特定 等</li> </ul> </li> <li>● 警備・見回り <ul style="list-style-type: none"> <li>- 施設内の見回り・正常確認</li> <li>- 異常検知した場合に駆け付け 等</li> </ul> </li> </ul> | <ul style="list-style-type: none"> <li>● 作業指示の理解 <ul style="list-style-type: none"> <li>- 人間からの口頭による指示を理解</li> <li>- 指示された内容を行動に変換 等</li> </ul> </li> <li>● 手順書・マニュアル作成 <ul style="list-style-type: none"> <li>- 作業内容のモニタリング</li> <li>- 手順書やマニュアルの作成 等</li> </ul> </li> <li>● 人手の動きの再現 <ul style="list-style-type: none"> <li>- 熟練作業者による製作方法を詳細に認識</li> <li>- 作業スタイルを再現 等</li> </ul> </li> </ul> |

- ロボティクス分野のAIセーフティ評価では、物理的リスクに加え、心理的・社会的リスクの抽出と制御が不可欠である。
- 想定されるリスクに対する受容可能な範囲を議論し、AIロボットの社会実装を進める上でステークホルダー間の責任分界やリスクアセスメントの実施方法等を検討することが重要。

## 想定されるリスク

物理的リスク（怪我、破損）  
衝突・接触、感電・火災 等

心理的リスク（不安、不信）  
恐怖・威圧的発言 等

社会的リスク（レピュテーションリスク）  
差別的扱い、偽誤情報の拡散 等



AIロボットがもたらす便益に対し、AIを利用することで**新たに発生しうる、どのようなリスクをどの程度まで許容**できるか

## 主な 論点

### スコープ

- ・ ロボットとAIを組み合わせた場合にどのようなリスクを想定するか。
- ・ 従来の機能安全評価ではなく、AIセーフティ評価として狙うべきスコープはどこか。

### 責任分界

- ・ AIの独自性が、製造者・提供者・利用者にどのような影響を及ぼすか。
- ・ AIが進化する性質を踏まえ、製品のライフサイクルにおいて、AIの安全性を誰がどこまで管理し、責任を持つか。

### リスク アセスメント

- ・ AIロボティクスの安全性評価について従来の機能安全のリスクアセスメントで全て説明できるのか。
- ・ 様々なユースケースが想定されるが、リスクアセスメントにおいて共通的な要素は何か。

### 評価環境

- ・ 人間とロボット、もしくは人間と環境の間をつなぐための言語モデルと、既存のシミュレーターをどう組み合わせるか。
- ・ 環境の認識から動作までを一つのモデルで扱える End-to-End モデルはどのようなデータセットや評価環境が必要か。

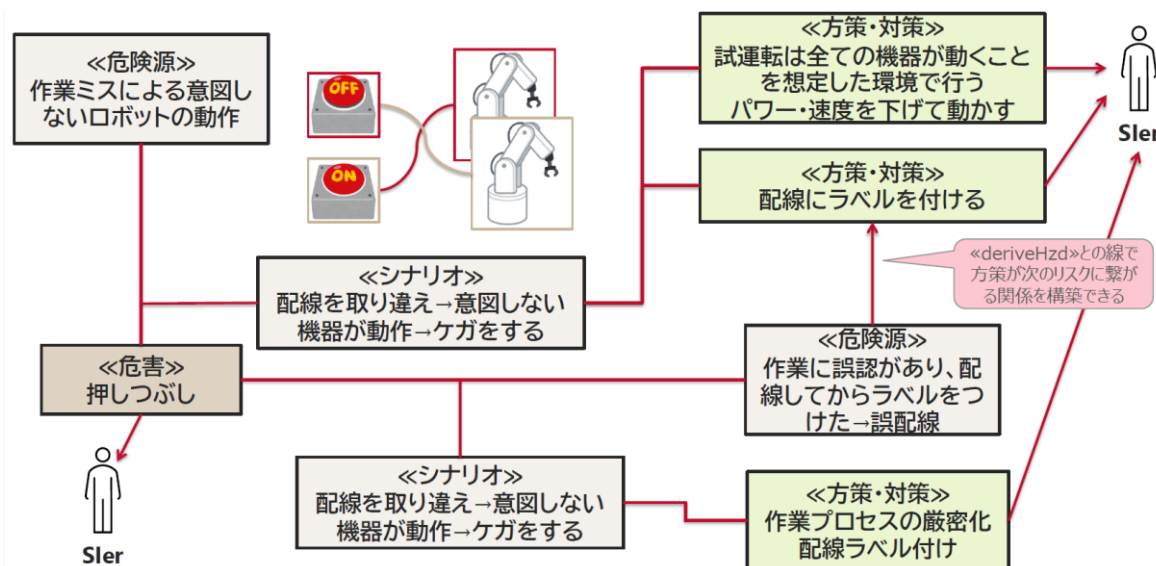
# 現在の取組み（1/3）

- AIRiskの観点から、SafeML\*を利用したリスクアセスメントや安全設計について検討。
  - ロボット開発の観点から、3フェーズの安全設計開発プロセスについても検討。
- ⇒ 今後、これらのスキームを利用したAIセーフティ評価手法について、実証も行いながら議論と検討を深めていく。

\* SafeML：システムの安全情報を記述するためのモデリング言語。RRI WG3にて検討中。

## SafeMLによるリスクアセスメント

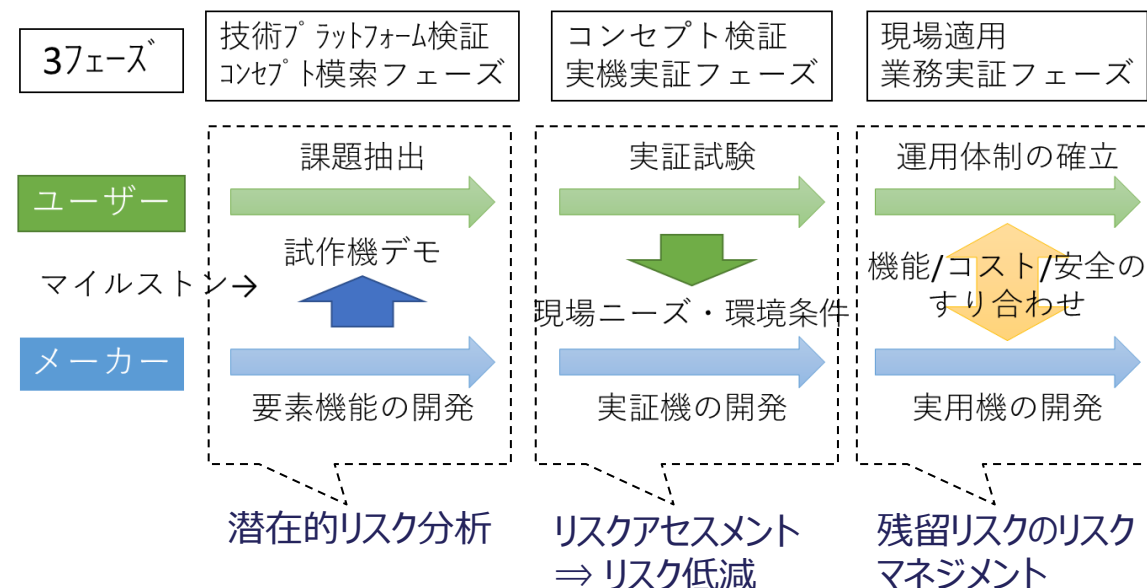
- ・ 実施したリスク対策が次の危険源へ繋がるおそれがある。
- ・ SafeMLを利用することにより、繰り返しを持つ複雑なリスクアセスメントも表現が可能となる。



出所）第2回ロボティクスSWG資料 サイバネットMBSE「安全設計におけるSE/MBSEの技術について」

## 3フェーズ開発プロセス

- ・ 時系列でモデル化した開発プロセスにおける3つのフェーズごとにマイルストーンを設定。
- ・ 潜在的なリスク分析から残留リスクのマネジメントまで対応



出所）第3回ロボティクスSWG資料 産総研「ロボットの安全設計開発」

# 現在の取組み（2/3）ユースケース①

- AIセーフティ評価の対象ユースケースについて、参加メンバーからの提案を踏まえ、「カフェ搬送」および「遠隔操作型小型車の自律移動」を選定。
- 「カフェ搬送」では、ロボットによる注文内容の把握（認知・判断）や配膳までの自律移動（行動の適切性や安全性）、注文者とのやり取り（インタラクションやプライバシー）をシナリオとして評価することを想定し検討中。

## シナリオ （案）

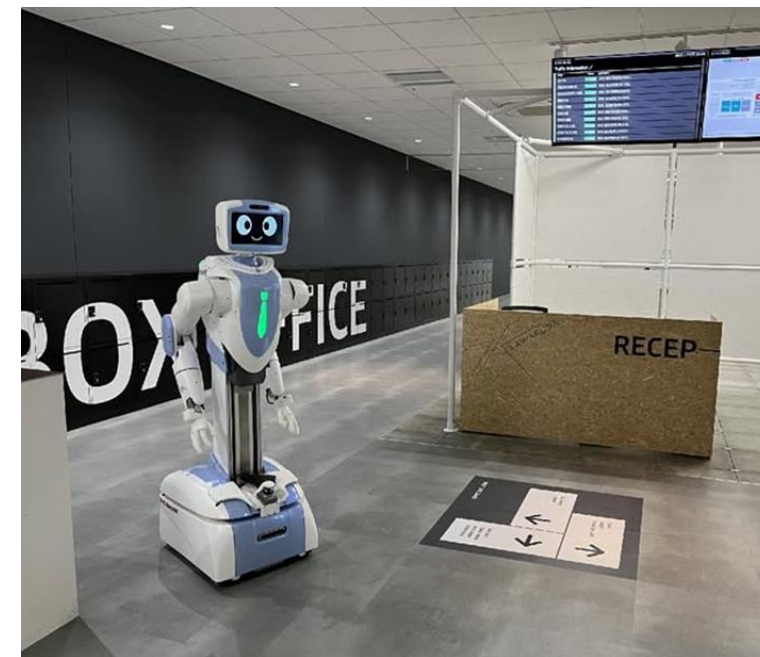
- 注文を受けたロボットがカフェで飲み物を受け取り、配膳
- 他のロボットや人とすれ違う、注文者が不在

## 評価観点 （検討中）

- 認知・判断：注文者の識別、受取場所と配達先の把握
- 行動の適切性：速度・経路制御、障害物回避
- インタラクション：注文者とのやり取り
- 安全性：安定走行、すれ違い時の安全確保
- プライバシー：注文者や注文内容の取扱いへの配慮

## 評価手法

- 川崎重工株式会社のソーシャルイノベーション共創拠点「KAWARUBA」の模擬環境（大田区）でシナリオを再現し実証を予定。
- ユーザインタビュー、ログ分析、チェックリスト評価 等



出所）川崎重工株式会社 CO-CREATION PARK KAWARUBA



# 現在の取組み（3/3）ユースケース②

- 「遠隔操作型小型車の自律移動」では、様々な自律移動ロボットを人間が遠隔から監視・操作する際の運用・効率性や安全性について評価する。
- 特に、人間とAIが連携して運用するケース（Human in the loop）において、AIによる支援の効果や人とAIとの責任分界点について検証を行い、人とAIロボットの協働における課題を分析する。

## シナリオ （案）

- 複数の自律移動ロボットが屋内外をシームレスに移動
- AI支援のもと、10台のロボットを人間が遠隔にて操作

## 評価観点 （検討中）

- 責任分界：人間（運行管理者）とAI（運行管理システム）の役割と責任範囲
- 安全性：多数のロボットの遠隔監視による衝突等の異常検知・回避
- 運用・効率性：自律移動による作業の効率性、遠隔操作・監視業務の運用面の課題

## 評価手法

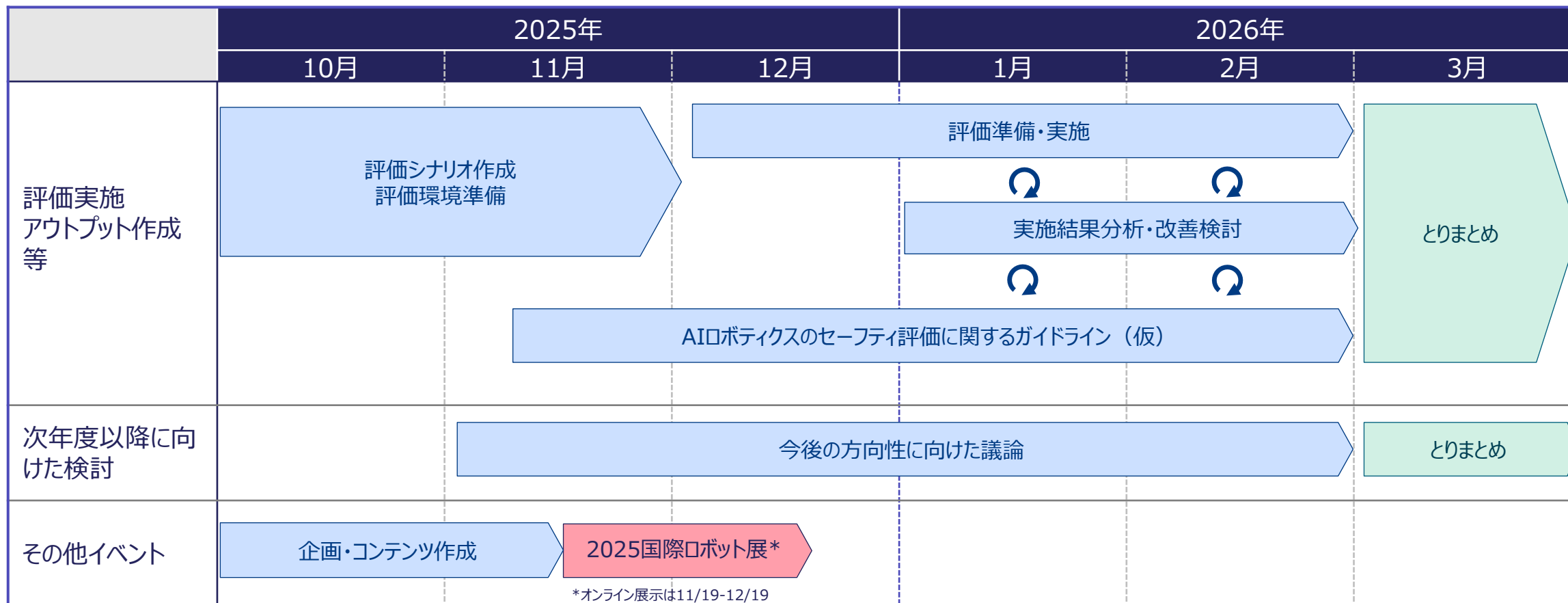
- 屋内外シームレスな環境での実証
- 遠隔操作者による監視の定性/定量評価、ログ分析 等



出所）パナソニックホールディングス株式会社「人とロボットの共創社会へ AI活用で導く“現場のリモート革命”」

# 今後の計画

- AIセーフティ評価を進め、AIロボティクスの開発・提供者に広く活用されるためのガイドラインの作成を目指す。
- 次年度以降に向け、ロボット基盤モデル（End to Endモデル）やフィジカルAI、ヒューマノイドロボット等の技術動向を踏まえ、AIロボティクスにおける、評価のあるべき姿や課題についても議論を開始。
- AIセーフティの普及啓発を目的に、「2025国際ロボット展」にて動画コンテンツのオンライン展示（11/19-12/19）とウェビナー（12/1）を実施予定。



# AISI

Japan AI Safety Institute