

令和7年度補正予算 研究開発等計画

AIロボティクス分野等の安全性に係る事業実証・研究開発事業

令和8年1月
経済産業省

- 実施する重点課題（特に該当するものには◎、そのほかで該当するものがあれば○（複数可）を記載）

SIPや各省庁制度による研究開発成果の社会実装・市場開拓の加速化	他の戦略分野等との技術の融合による研究開発	スタートアップによるイノベーションの創出・促進	産学官を挙げた人材の育成・確保	グローバルな視点での連携強化
○	◎			

- 関連するSIP課題（該当するものに○を記載）

持続可能なフードチェーン	ヘルスケア	包摂的コミュニティ	学び方・働き方	海洋安全保障	スマートエネルギー	サーキュラーエコノミー	防災ネットワーク	インフラマネジメント	モビリティプラットフォーム	人協調型ロボティクス	バーチャルエコノミー	先進的量子技術基盤	マテリアル事業化・育成エコ
										○			

1. 施策の概要

「総合経済対策の策定について」において指示されたAIなどへの戦略分野の官民連携投資の強化においては、その前提として、**今後の産業利用が特に見込まれるロボティクス分野等のAI安全性評価に関する手法などを確立**させ、AIリスクの懸念を低減させることが必要である。そこで、以下の2つを実施する。

①AIの安全性評価等の中心機関として、**AIセーフティ・インスティテュート (AISI) の体制を強化**し、ロボティクス分野等の**事業実証WG**における民間事業者も参画した実証を通じて、**業種別のAIセーフティ評価に関するドキュメントを作成**する。

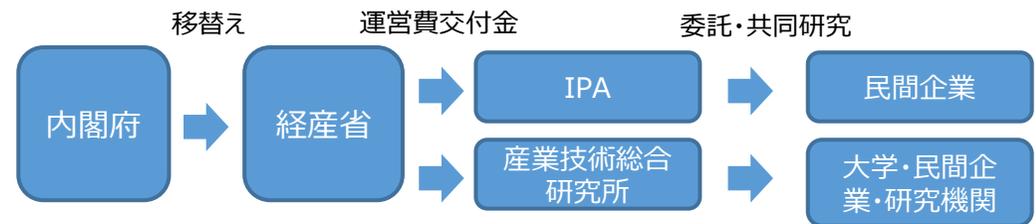
②**産業分野のAIセーフティ強化に関する研究開発事業を実施**する。今後の産業利用が特に見込まれる**協働ロボット分野等におけるAIセーフティに関する研究開発の実施**及び**AIセーフティ基準開発、適合評価に必要なベンチマークデータセットの構築、ISO/IEC JTC1/SC 42への打ち込み**等を行う。

本施策は、AIリスクの懸念を低減させるものであることからA I 基本計画（骨子）で示されたAIガバナンスの主導（「AIの信頼性を高める」）に資するものであり、AIセーフティ評価の普及により、各業種ごとのAI利活用を促していく（「AIを使う」）。

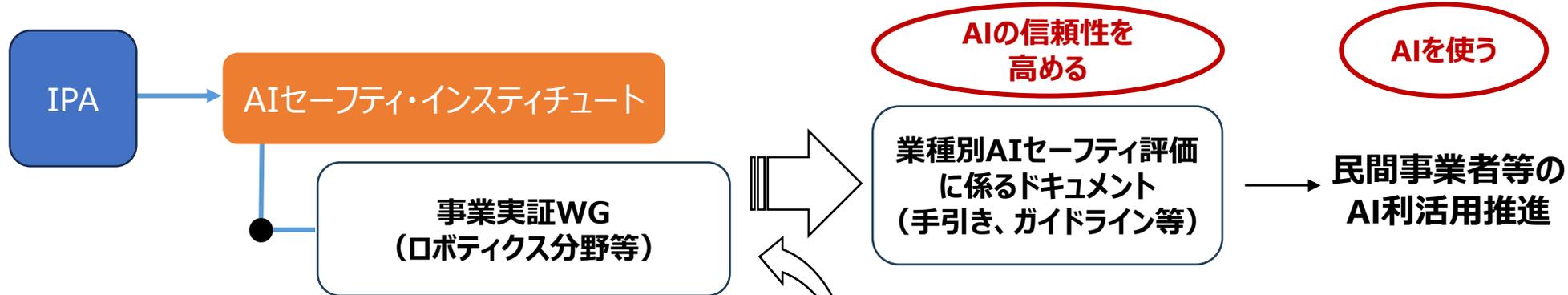
2. 施策の対象・成果イメージ

業種別のAIセーフティ評価に関するドキュメントや、適合評価に必要なベンチマークデータセットが作成されることで、各業種においてAIの利活用が促進される。

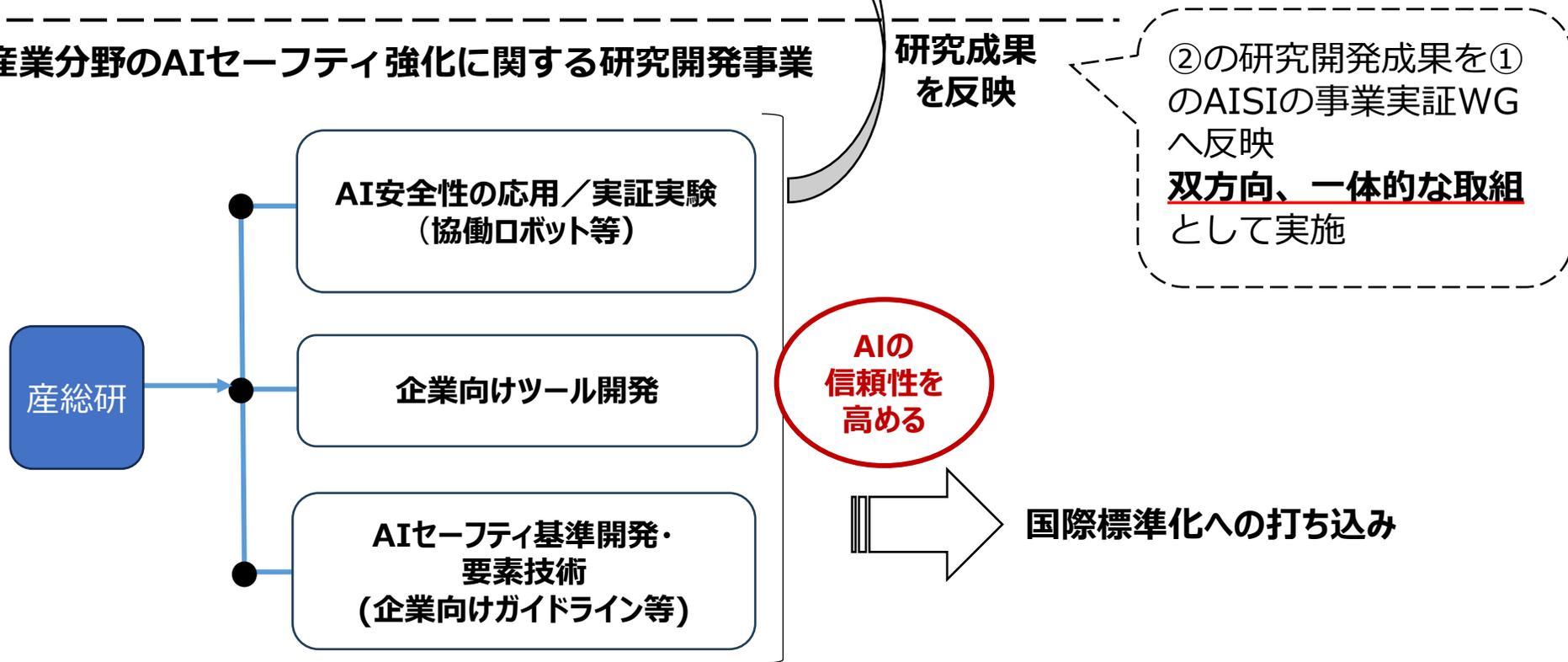
3. 資金の流れ



① AIセーフティ・インスティテュート (AISI) の体制強化・事業実証



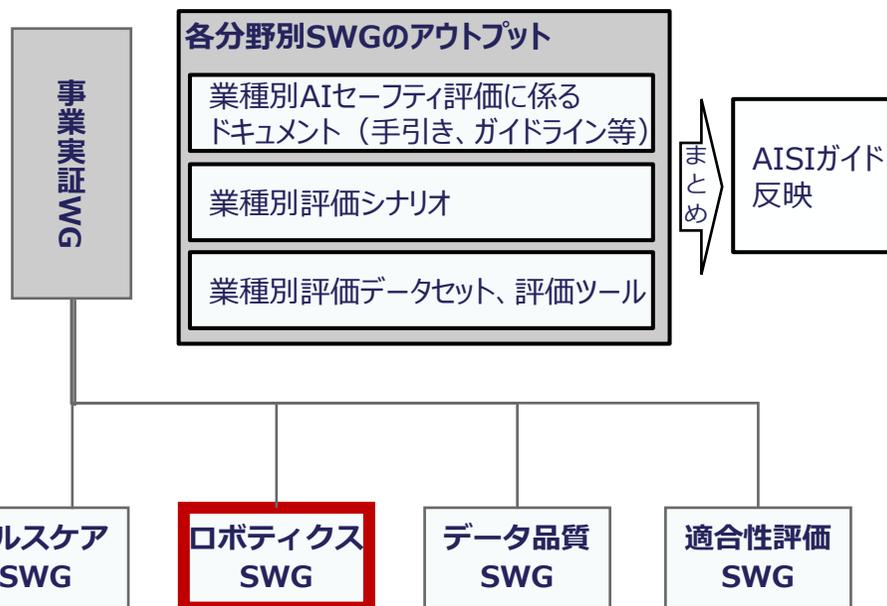
② 産業分野のAIセーフティ強化に関する研究開発事業



①AIセーフティ・インスティテュート（AISI）の体制強化・事業実証

- AIセーフティ評価に関するワーキンググループ（事業実証WG）を、AISI運営委員会の下のテーマ別小委員会として設置。民間事業者を中心に多様なステークホルダーが参画し、参画機関間の連携を図る場を提供、WG活動を推進する。
- AIセーフティ評価の活動を広く一般に普及させ、AIの利活用を促進させることを目的とし、民間企業を中心とした業界ごとの有識者ととともに、業界ごとのAIセーフティ評価に関する見解をまとめ、具体的な実証をする等のWG活動を推進し、業界ごとに特化されたガイドやデータを作り、その普及を図る。
- 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。

事業実証WGの全体像



ロボティクスSWGの活動

- AIセーフティ評価の対象ユースケースについて、「カフェ搬送」、「遠隔操作型小型車の自律移動（公道自律走行）」を選定。

（ユースケース1）カフェ搬送

- ✓ ロボットによる注文内容の把握や配膳迄の自律移動、注文者とのやり取りをシナリオとして評価

（ユースケース2）公道自律走行

- ✓ 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

- シミュレーション環境・実環境におけるロボットとの接触リスク、心理的リスク、衝突リスクなどの検証を想定。

- AIを導入するうえでの安全性の課題を構造化し、リスクに関する指標を体系化。



出所) 川崎重工株式会社
CO-CREATION PARK KAWARUBA



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」

②産業分野のAIセーフティ強化に関する研究開発事業

- ・マルチモーダルAIやフィジカルAIを対象としたAIセーフティ基準のコアとなる安全性評価・管理基盤技術の確立
- ・生活環境やロボットといったフィジカルに関する領域での安全性評価技術・ベンチマークデータセット構築・テスト環境構築技術の確立
- ・AIセーフティ基準の開発とガイドライン・企業向け実装解説の作成、ISO/IEC等における規格化に向けた標準化活動の実施

AIセーフティ基準・ガイダンスと標準化

国際標準化活動

ISO/IECでの
規格策定推進

整合性整理

AIセーフティ基準開発

機械学習品質
マネジメントガイドライン

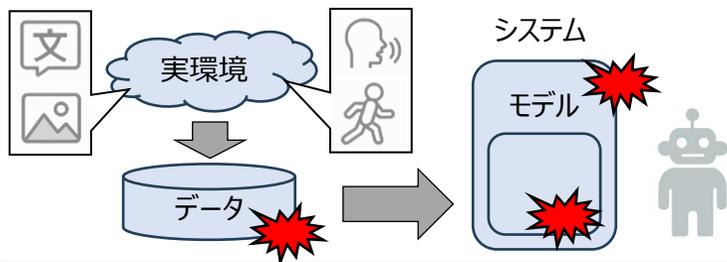
AIセーフティ
ガイドライン

社会実装・普及の促進

企業向け
実装解説

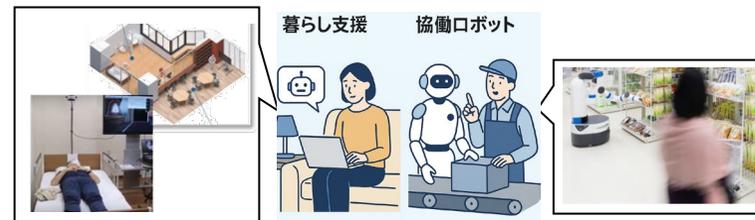
AI品質マネジメント
イニシアティブ

AIセーフティ評価・管理基盤技術



マルチモーダルAIやフィジカルAIに対する安全性評価技術は、各モダリティ単独の開発段階にとどまっており、モダリティ間連携や物理的インタラクションを含むリスク評価は体系化されていない。AIセーフティ基準策定に必須。

応用領域別AIセーフティ評価・実装技術



今後の産業利用が特に見込まれる暮らし支援および協働ロボットにおけるAIセーフティ評価・実装技術を開発し、実環境・仮想環境を活用した評価基盤を確立する。

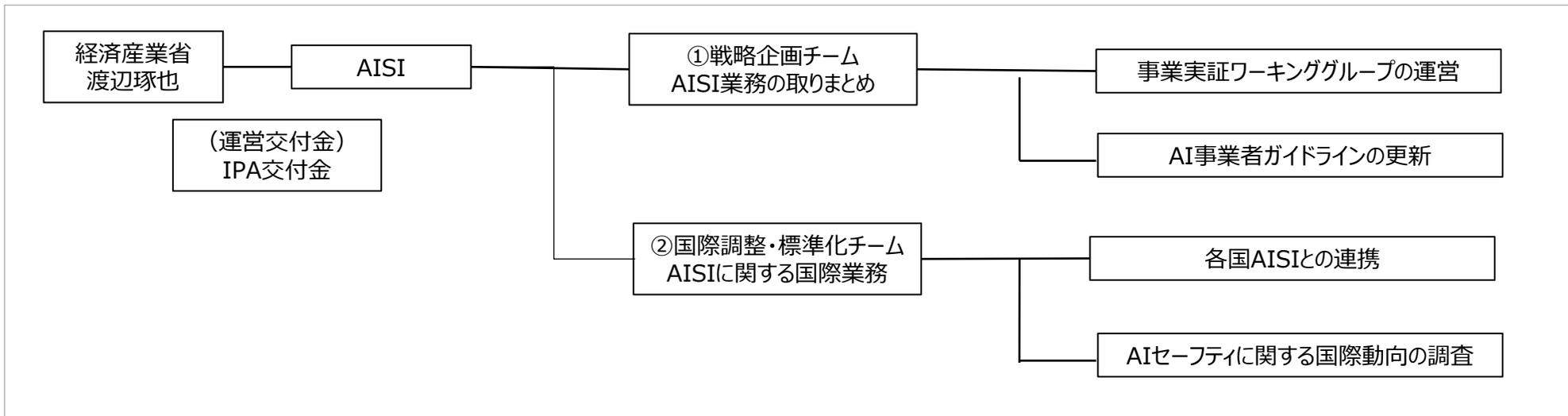
① AIセーフティ・インスティテュート（AISI）の体制強化・事業実証

テーマ名	令和7年度			令和8年度											
	1月	2月	3月	4月	5月	6月	7月	8月	9月	10月	11月	12月	1月	2月	3月
ASISコンソーシアムの運営等	事業実証ワーキンググループ（各サブワーキンググループ）の実施														
	開発コンソーシアムとの連携														
ガイドンス等の作成・更新等	AI事業者ガイドラインの更新														
	業種別AIセーフティ評価に係るドキュメントの作成														

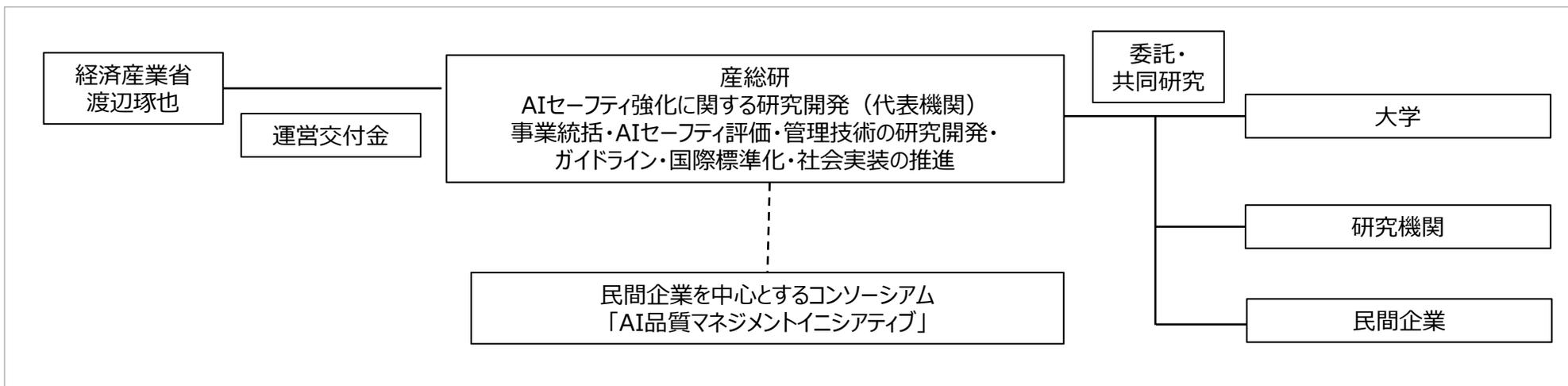
② 産業分野のAIセーフティ強化に関する研究開発事業

① AIセーフティ評価・管理基盤技術	個別の評価手法・ツールの研究開発										評価手法・ツールの実証実験			基準への反映論文執筆	
	モダリティ連携手法の研究開発														
	ベンチマークデータセット要件整理										試験的安全性評価				
② 応用領域別AIセーフティの評価・実装技術	ベンチマークデータセット・テスト環境構築技術の開発														
	評価シナリオ・手法の開発										評価シナリオ・手法の実証実験			基準への反映論文執筆	
③ AIセーフティ基準・ガイドンスと標準化	AIセーフティ基準開発・ガイドライン策定・企業向け実装解説の作成										ガイドライン改訂				
	コンソーシアムからのニーズ吸い上げ										コンソーシアムを通じた実証				
	機能安全規格等の策定推進										ISO/IEC/AISI技術体系との整合性整理				

① AIセーフティ・インスティテュート (AISI) の体制強化・事業実証



② 産業分野のAIセーフティ強化に関する研究開発事業



① AIセーフティ・インスティテュート (AISI) の体制強化・事業実証

テーマ名	実施内容の概要 到達目標 (KPI)
① AISIコンソーシアムの運営	事業実証ワーキンググループの下に設置されている各テーマごとのサブワーキンググループの実施および成果物（ガイドラインなど）の公表・更新。その過程で、産総研を始めとするパートナーシップ機関とも連携を行う。開発コンソーシアムや新規コミュニティとの連携 KPI：上半期中にサブワーキンググループの成果物（ガイドライン）を1件公表
② ガイダンス等の作成・更新等	AIリスクの増加に伴い、AI事業者ガイドライン（2024年公表）の更新を検討・実施する。日本のAI安全性の中心機関として、各国のAISIとの連携（意見交換やイベントへの参加）を行う。また、国際的な動きを踏まえ、AIセーフティに関する国際動向を調査し、必要に応じて国内のガイドラインなどに反映する。 KPI：年度末までにAI事業者ガイドラインを更新、各国のAISIとのイベントを1回開催

② 産業分野のAIセーフティ強化に関する研究開発事業

テーマ名	実施内容の概要 到達目標 (KPI)
AIセーフティ評価・管理基盤技術	マルチモーダルAIおよびフィジカルAIを対象に、AIセーフティ基準の詳細化を目的とした評価・管理基盤技術を確立する。データ、モデル、システムの要素技術の設計、評価ツールのプロトタイプ試作、ベンチマークデータセット要件整理、ベンチマークデータセットに基づく試験的安全性評価を行う。 KPI：2027年3月までに上記の活動を整理した報告書3件を作成し、AISIに提供
応用領域別AIセーフティの評価・実装技術	今後の産業利用が特に見込まれる「暮らし支援」および「協働ロボット」領域において、AIシステムの安全性評価・実装技術を開発し、安全性評価に用いる実環境と仮想環境のベンチマークデータセット、テスト環境構築技術を確立する。生活環境やロボットといったフィジカルに関するデータを活用したAIセーフティ評価シナリオ・手法を構築し、2027年3月までに「暮らし支援」および「協働ロボット」領域で各2件、合計4件のAI製品・サービス等を対象として評価シナリオ・手法の実証実験を行う。 KPI：2027年3月までに上記の活動を整理した報告書を作成し、AISIに提供
AIセーフティ基準・ガイダンスと標準化	AIセーフティ基準を開発するとともに、AIセーフティ基準の社会実装・普及の促進を目的として、AIセーフティに関する企業コンソーシアムを支援する。さらに、ISO/IEC等での標準化活動を行い、AI標準の国際的議論に貢献し、AISIに成果を提供する。 KPI：AIエージェント向けの基準をまとめたガイドラインを2027年3月までに公開・改訂 マルチモーダルAI等のセーフティに関する企業向け実装解説を2027年3月までに作成