

【参考】 AIセーフティ・インスティテュート（AISI）について



概要

AIの安全性に対する国際的な関心の高まりを踏まえ、AIの安全性の評価手法の検討等を行う機関として、米国や英国と同様に、日本においても、AIセーフティ・インスティテュートを2024年2月14日に設立した。

同機関は、内閣府をはじめ関係省庁、関係機関の協力の下、IPA（独立行政法人情報処理推進機構）に設置され、諸外国の機関とも連携して、AIの安全性評価に関する基準や手法の検討等を進めていく。

所長には、元日本IBMのAI研究者で、現在は損保ジャパンCDO（チーフ・デジタル・オフィサー）で京都大学防災研究所客員講師の村上明子氏が就任した。

業務内容（暫定）

1. 安全性評価に係る調査、基準等の検討
 - ①安全性に係る標準、チェックツール、偽情報対策技術、AIとサイバーセキュリティに関する調査
 - ②安全性に係る基準、ガイダンス等の検討
 - ③上記に関するAIのテスト環境の整備の検討
2. 安全性評価の実施手法に関する検討
3. 他国の関係機関（英米のAI Safety Institute等）との国際連携に関する業務

関係省庁・関係機関

関係省庁 内閣府（科学技術・イノベーション推進事務局）、国家安全保障局、内閣サイバーセキュリティセンター、警察庁、デジタル庁、総務省、外務省、文科省、経産省、防衛省

関係機関 情報通信研究機構、理化学研究所、国立情報学研究所、産業技術総合研究所

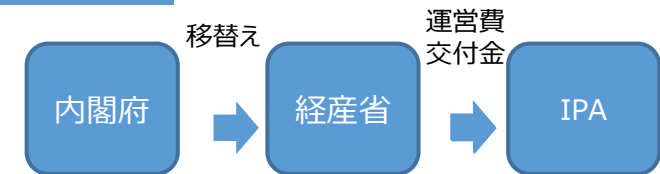
1. 施策の概要

- 各国AISIとのネットワークを活かし、国内外の最新技術を含め、AI安全性に関する動向を調査し、AI安全性に係る各国の政策、ツール開発、ルール形成等の状況把握、分析を行う。
- BRIDGEの各省事業の実施状況、成果などについて情報交換を行いながら、各事業におけるAI安全性に関する評価手法等を策定する。

2. 施策の対象・成果イメージ

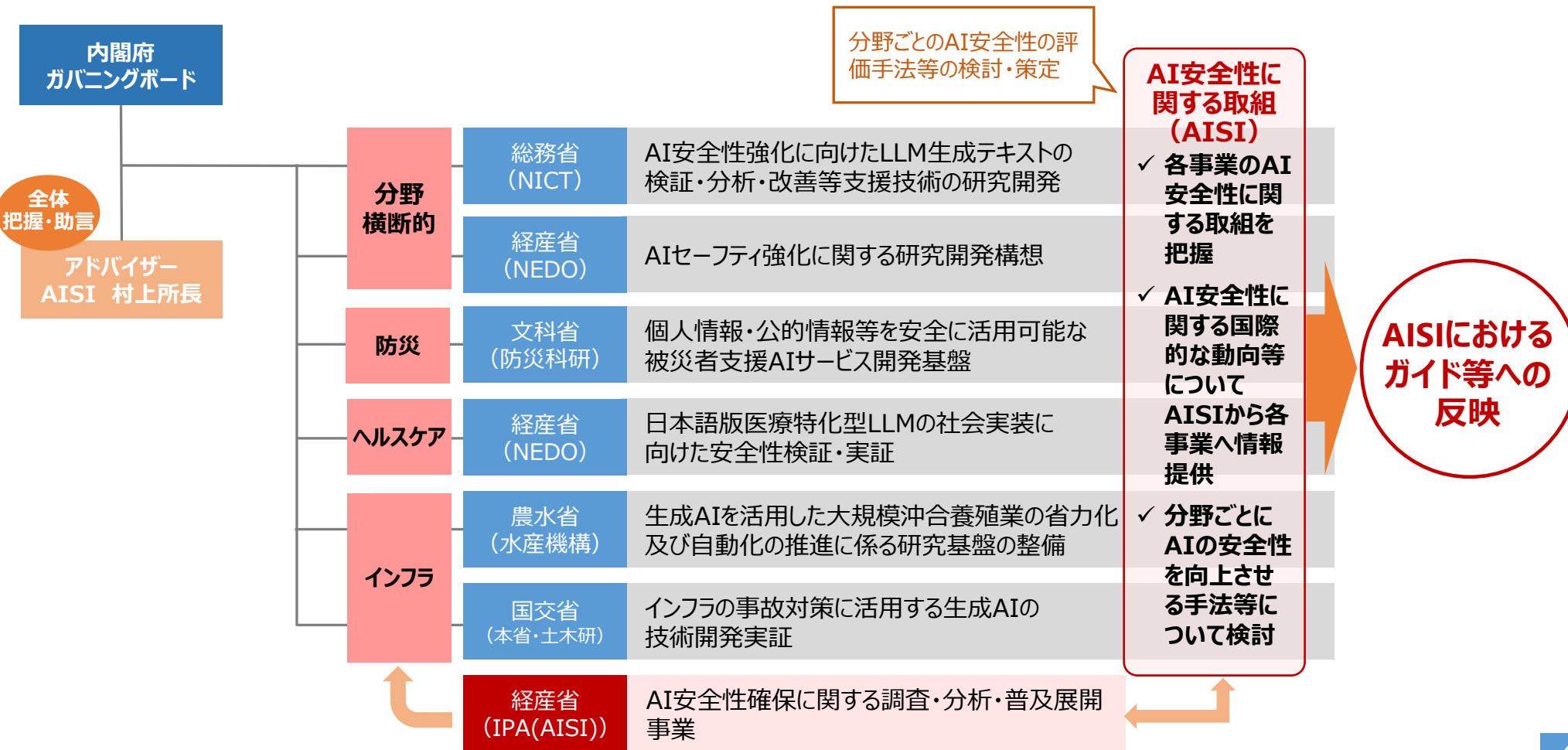
- AIの安全性に関する国内外の動向を把握しつつ、海外（英国、米国等）のAISIとの連携も深めることで、国内のAI安全性に関する技術基準等を国際的な場での議論においても示すことができ、我が国がAI安全性の国際標準等の議論においても主導していくことが期待できる。

3. 資金の流れ



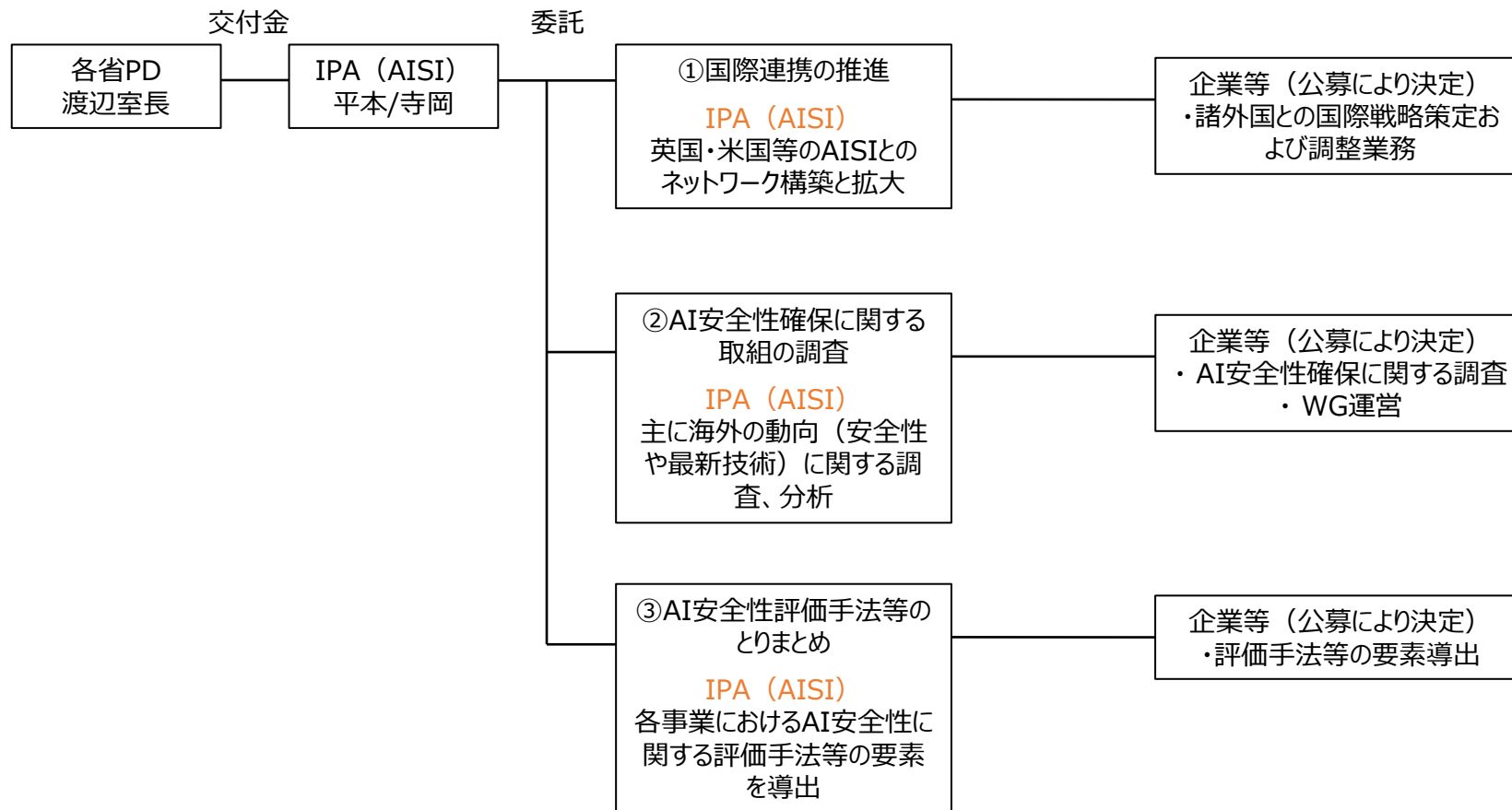
4. 取組内容（柱書き再掲）

- 各国AISIとのネットワークを活かし、国内外の最新技術を含め、AI安全性に関する動向を調査し、AI安全性に係る各国の政策、ツール開発、ルール形成等の状況把握、分析を行う。
- 以下の各省事業の実施状況、成果などについて情報交換を行いながら、各事業におけるAI安全性に関する評価手法等を策定する。



5. 取組スケジュール

内容	令和6年度			令和7年度												
	1	2	3	4	5	6	7	8	9	10	11	12	1	2	3	
国際連携の推進	英国・米国等のAISIとのAI安全性に関する基準等の調整															
AI安全性確保に関する取組の実態調査				<p>SIP事業への情報共有・展開</p> <p>実施状況・成果の共有</p> <p>海外のAI安全性全般（制度、基準等）に関する動向調査</p> <p>AI技術の最新動向調査</p> <p>分野ごとの国内外調査（分野横断、防災、インフラ、ヘルスケア）</p>												
AI安全性に関する基準等の方向性のとりまとめ																<p>SIP事業への情報共有・展開</p> <p>SIPへの共有・意見交換を通じて基準等へ反映</p> <p>AI安全性評価手法等のとりまとめ</p>



テーマ名	実施内容の概要 到達目標 (KPI)
①国際連携の推進	<ul style="list-style-type: none">● 国際AISIネットワークの構築と拡大を推進する。● AI安全性に係る各国の取組について、協調領域について議論するとともに、相互に貢献できる分野を把握する。
②AI安全性確保に関する取組の調査	<ul style="list-style-type: none">● 各国AISIとのネットワークを活かした調査等により、AI安全性に係る各国の政策、ツール開発、ルール形成等の状況把握、分析を実施する。● 各省事業に対して、上記の調査結果について情報提供および助言を行う。
③AI安全性評価手法等のとりまとめ	<ul style="list-style-type: none">● 分野ごとのAIの安全性に関する検討結果から、AIの安全性評価に関連する事項を抽出し、それらのうちさらに分野横断的に適応すべき（汎用化すべき）要素を導出するなど、AISIにおけるガイド等を改訂するための基礎検討を行う。