AIのメリットとリスク

- ものづくり、金融、医療、教育、行政など様々な分野で、生成AI利用による飛躍的革新の可能性。
- 一方で、AIに関するリスクへの対応が官民ともに重要。
- 日本のAI関連産業の競争力強化も必要。

AIのメリット

労働力不足の解消

・問合せ対応、点検・監視など、業務の一部をAIで代替し、人は創造的な業務に注力

事務作業の効率化

・文献調査、要約
・添削・翻訳
・会議録等の作成
・資料の原案作成

サービスの質の向上

・病気の早期発見
・生徒・児童一人一人に合った教材の提供

イノベーションの創出

・新たな素材の開発
・新たな薬の、治療法の開発

地球規模の課題解決

・災害予測、対策
・農業の高収益化
・パンデミック対策
・安全保障

AIのリスク

- 機密情報漏洩、個人情報の不適正利用
- 犯罪の巧妙化・容易化
- 偽情報、誤情報等による社会不安
- サイバー攻撃の巧妙化
- 教育への影響
- 著作権侵害の懸念
- 雇用への影響

「AI戦略会議」における議論

これまでの基本戦略・理念 「AI戦略2022」「人間中心のAI社会原則」

生成AIなどの技術の変化

自然な対話が可能、精巧な画像生成が容易など
大きな便益・イノベーション、Society 5.0に寄与
一方で、AIに関するリスクはより切迫したものに

国際的な議論

G7広島サミットにおいて合意された**共通のビジョンと目標**「我々が共有する民主的価値に沿った、信頼できるAI」**の実現に向けて**G7にて議論を開始**（広島AIプロセス）**

新たに有識者会議である「AI戦略会議」を立ち上げ議論

■ AIに関する暫定的な論点整理（2023年5月26日 AI戦略会議）と政府の取組み

国際的な議論とリスクへの対応
- 広島AIプロセスなど国際的議論を主導
- 生成AIに関する懸念やリスクへの対応

AIの最適な利用
- 中央省庁による生成AIの段階的利用
- 幅広い世代のスキル・リテラシー教育

AI開発力の強化
- 計算資源の確保、データ整備
- 研究力向上、スタートアップ支援

取組み状況

- 関係省庁による生成AIの業務利用に関する申合せ（デジタル庁）
- 生成 AI サービスの利用に関する注意喚起等（個人情報保護委員会）
- 初等中等教育における生成AI利用に関する暫定的ガイドライン（文部科学省）
- 生成AIの利用と開発力強化に向けた事業拡充（全関係府省）
- 広島AIプロセス包括的政策枠組みの合意（総務省・外務省）
- AIセーフティ・インスティテュートの設立（内閣府等の関係府省庁、関係研究機関）
- 著作権等を含む知的財産権に関する中間とりまとめ等（知的財産戦略推進事務局・文化庁）
- AI事業者ガイドライン案とりまとめ（総務省・経済産業省）

AISIの設立について

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

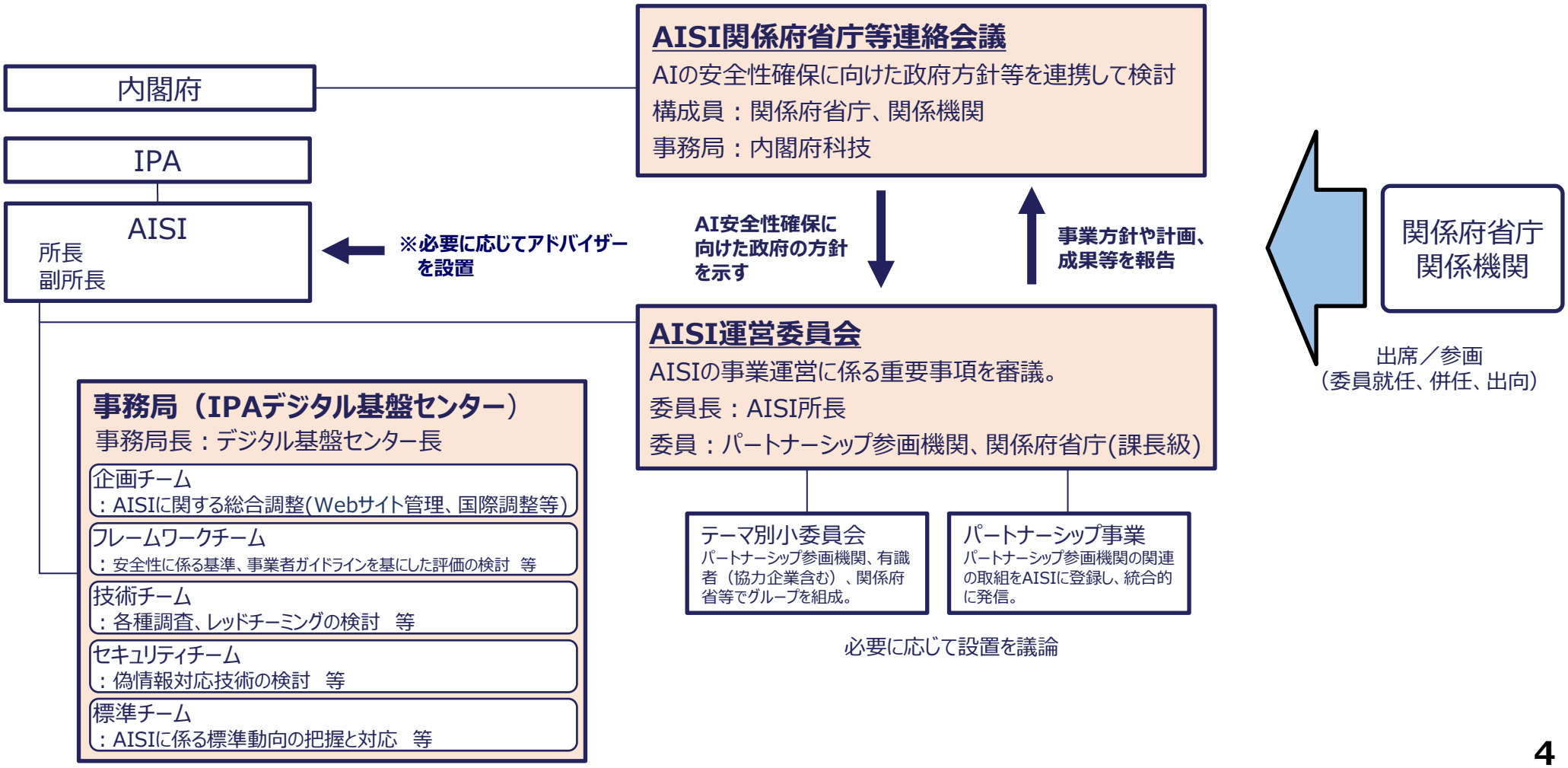
関係省庁・関係機関

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

AISIの体制整備

- 内閣府を事務局とする「AISI関係府省庁等連絡会議」を設置し、重要事項を審議（年間２～３回の開催を予定）。
- AISIの中に、AISI所長を委員長とする「AISI運営委員会」を設置（月１回の開催を予定）。運営委員会の下に、必要に応じて、「テーマ別小委員会」や「パートナーシップ事業」（研究機関等の関連の取組みをAISI事業として発信）を設置。
- AISIの事務局として、IPAデジタル基盤センターの中で５つのチームを編成。



各国のAIセーフティ・インスティテュート

	米国	英国	日本
概要	**大統領令（2023.10.31）** ガイドラインの整備から、専門人材の獲得、研究開発、政府によるAIの活用に至るまで**AIに関する幅広い内容**を含む連邦政府機関への指示	**AI安全性サミット（2023.11/1-2）** AI安全性サミットでは、AIがもたらす**便益を対外的に示し**つつ、**最先端のAI（フロンティアAI）のもたらすリスクについて、安全性に焦点を当てて共通の理解**を促進	**AI戦略会議 総理発言（2023.12.21）** AIをめぐる安全性に対する国際的認識が高まっています。日本としても、海外機関と連携し、AIの安全性の評価手法の研究や規格作成などを行う機関が必要との考えに立ち、「AIセーフティー・インスティテュート」を設立する。
安全性評価			
主体	**AIセーフティ・インスティテュート（US AISI）** NIST（国立標準・技術研究所）内に設置	**AIセーフティ・インスティテュート（UK AISI）** 首相直轄のフロンティアAIタスクフォースを発展改組	**AIセーフティ・インスティテュート（J-AISI）** 10府省庁、関係国立研究機関が参画
対象	**国家安全保障、経済安全保障、健康・安全に深刻なリスクをもたらすAI**	**デュアルユース、サイバー攻撃に利用されうるAI**	**生成AI、大規模汎用モデル等に限定しない【P】**
手法	● 開発者におけるAIの市場導入前の**安全性評価のためのガイドライン、ツール等をNISTが整備** ● 対象事業者に対して、ツールを用いた安全性評価の結果および対策の**報告義務**を課す	● 開発者におけるAIの市場導入前の**安全性評価のためのガイドライン、ツール等を英AISIが整備** ● 英AISIが対象事業者の**技術評価を実施**	● 安全性評価に係る調査、基準等の検討 ● 安全性評価の実施手法に関する検討 ● 他国の関係機関（英米のAI Safety Institute等）との国際連携