AIによって自律化が進む米国軍事技術の動向

The Impact of Artificial Intelligence on the Advancement of Autonomous Military Technologies in the U.S.

岡島　義憲
（情報統合技術研究合同会社）

山川 宏
（東京大学）

2023年11月24日

（注） 本本稿は、東京大学ムハンマド・ビン・サルマン未来科学技術センター（MbSC2030）の支援によるものである。

目次

1. 本調査について
2. Autonomy拡大のトレンド：Ubiquitous A.I. Integration
3. Autonomy実現に向けてのソリューション技術
   3.1 Digital Nervous System
   3.2 A.I.連携 と Human-Machine Collaboration
   3.3 人間中心主義
4. 考察
   4.1 Agent連携 と 人間中心主義
   4.2 A.I.リスク
5. まとめ と 所感

目次

1. 本調査について
2. Autonomy拡大のトレンド：Ubiquitous A.I. Integration
3. Autonomy実現に向けてのソリューション技術
   3.1 Digital Nervous System
   3.2 A.I.連携 と Human-Machine Collaboration
   3.3 人間中心主義
4. 考察
   4.1 Agent連携 と 人間中心主義
   4.2 A.I.リスク
5. まとめ と 所感

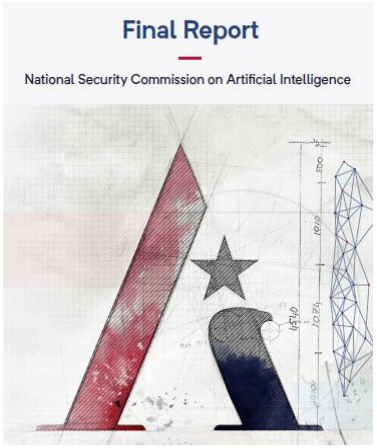
[1] National Security Commission on A.I. (NSCAI) ; " Final Report (2021)", March, 2021.
[2] Special Competitive Studies Project (SCSP) ; "Offset-X", May 2023.

調査目的 と 調査資料

目的 ： AI技術の導入により、進んでいると見られている兵器や軍事システムの自律化（Autonomy化）の技術トレンドを俯瞰することにより、今後、どういうリスクが問題となりうるかを探る。

調査資料 ： 以下、及び、関連資料 や ネット上の記事。

（1）米国 NSCAIのFinal Report [1]
（National Security Commission on A.I. ）



2021年

（2）米国 SCSP の Offset-X [2]
（Special Competitive Studies Project）

2023年

[1] National Security Commission on A.I. (NSCAI) ; " Final Report (2021)", March, 2021.
[2] Special Competitive Studies Project (SCSP) ; "Offset-X", May 2023.

用語について

- Teaming： 自律兵器や兵士の連携動作
  - > Human-Machine Teaming
  - > Machine-Machine Teaming
- Swarm ：自律兵器の 群れ(群隊)
  - > Swarms of Swarms
- Digital Nervous System
  - > 通信ネットワーク＆センサー＆サーバー
  - > 情報を集約、分析、指令するシステム
  - > 監視/管理/指示の機構
- Ubiquitous A.I. Integration
  - > Pipeline（データ処理プロセス）
- Digital Eco System
  - > A.I.に関する開発＆サポートの組織
  - > 開発環境/ツール開発、評価、Update

[1] NSCAI; "Foundations of Future Defense"、 Chapter-2, Final Report, 2021.
[2] NSCAI, "AI and Warfare", Chapter-3, Final-Report,Page-79-88, 2021.

軍事がA.I.技術に求める機能

要求1 ： 人間を超える(知的な)能力、敵よりも優れる能力

- 認知、協力：敵と味方と障害物を見極め、Team-Work良く見方と協力
- 計画、判断、評価：
  敵の攻撃を理解し、状況に対応して戦い方を変更し、連携を再構築する。
- 高速処理 ：収集した情報を元に、下記の一連の動作をPipeline処理

① Prepare
② Sense & Understand
③ Decide
④ Execute

要求2 ： 無人化（通信断絶時にも稼働可能）

目次

1. 本調査について
2. Autonomy拡大のトレンド：Ubiquitous A.I. Integration
3. Autonomy実現に向けてのソリューション技術
   3.1 Digital Nervous System
   3.2 A.I.連携 と Human-Machine Collaboration
   3.3 人間中心主義
4. 考察
   4.1 Agent連携 と 人間中心主義
   4.2 A.I.リスク
5. まとめ と 所感

[1] DARPA; “DARPA Testing Military Drones With Swarming Capability”.
[2] DATPA; ””OFFSET Swarms Take Flight in Final Field Experiment“, Dec.9, 2021.

Gremlins計画（DARPA、2019年）

- ネットワーク化された群れ（Swarm）が、共同して、敵をリアルタイムで把握し、応答しながら、任務遂行する「群れ制御技術」の開発計画[1]
（Machine-Machine Teaming）

図：DARPAのGremlins Military Dronesのコンセプト（出典）DARPA [1]

（※） Jon Harper; "'Air Force CCA drones could have greater range than crewed fighters: senior official", Aerospace Manufacturing, Nov., 15, 2023

Collaborative Combat Aircraft（CCA）計画

（協力的な戦闘ドローンについて、2023年11月）

- 4～5機の無人機が、有人戦闘機と、フォーメーションを組むというビジョン。
- 空軍は、CCA 1,000機構成される将来の軍隊を計画中
  （有人機の5倍から10倍の機数の導入を計画）
- 2028年に生産、2020年代末の導入を検討中

[1] Noah Robertson; “Pentagon unveils ‘Replicator’ drone program to compete with China”, DefenseNews, Aug.29, 2022.
[2] Thomas Newdick, Tyler Rogoway: “Replicator Is DoD’s Big Play To Build Thousands Of Autonomous Weapons In Just Two Years”, The Warzone,Aug. 28, 2023

自律無人システムをReplicator計画に集約

（2022年8月、 2023年 8月）

- 「小型＆安価な自律無人システム（陸・海・空用）の開発」を Replicator計画に集約し、生産規模拡大を進める[1]。
- （商業技術を活用し、損耗品のように） 航空機を量産し、迅速に納入することを目指す[1] 。
- 今後18～24か月以内に複数の分野で数千機の納入を目指す[2] 。
- ウクライナ軍が要求するUAVの数も「4桁」ではなく「5桁」に達している[2] 。

Tim McMillan; "Pentagon Secretly Working To Unleash Massive Swarms of Autonomous Multi-Domain Drones to Dominate Enemy Defenses", The Debrief, Feb. 6, 2023.

DARPAのAMASS program

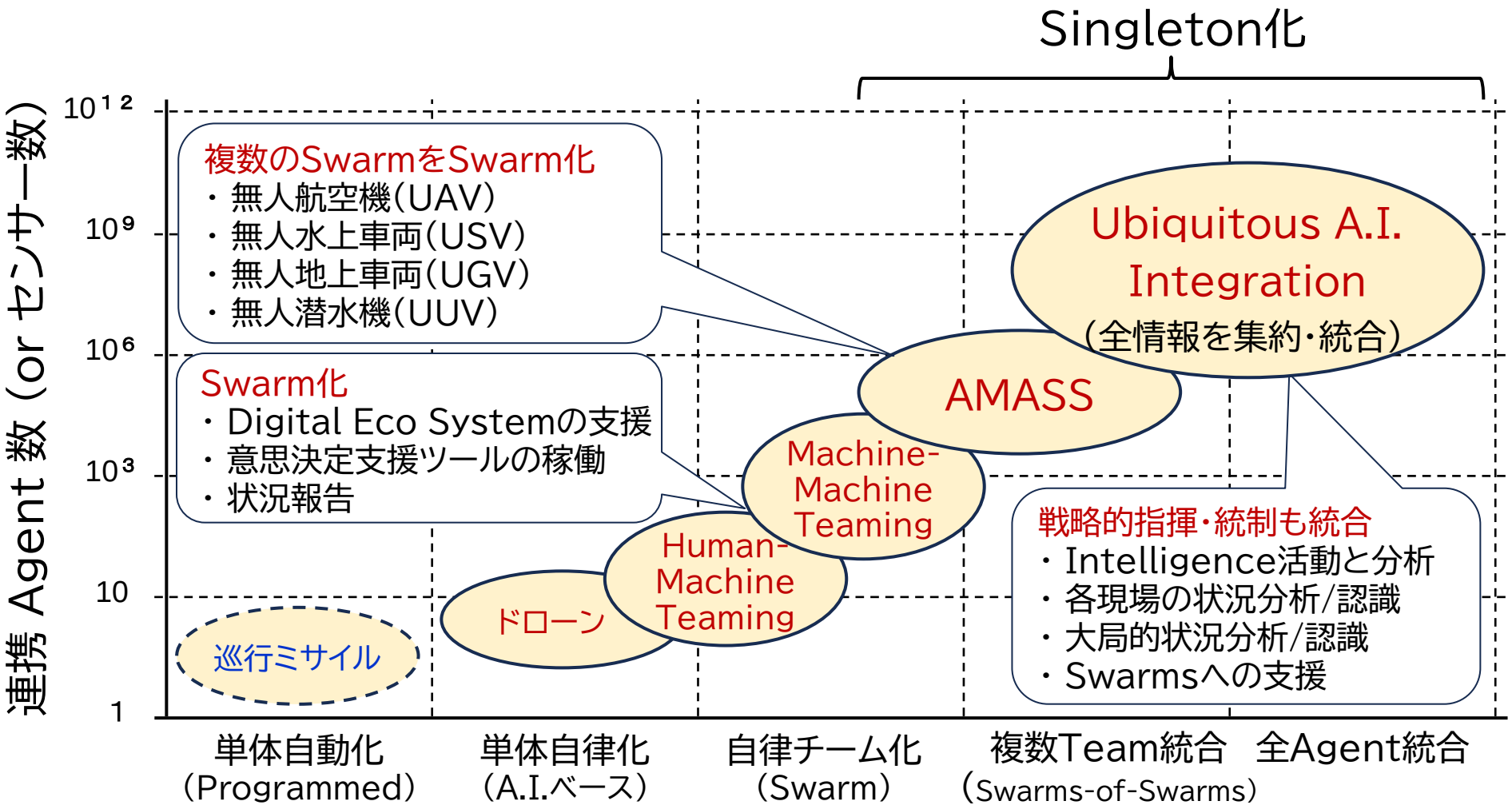
（2023年2月は企画段階）

（AMASS : Autonomous Multi-Domain Adaptive Swarms-of-Swarms）

- 無人航空機（UAV）、無人水上車両（USV）、無人地上車両（UGV）、無人潜水機（UUV）等の、多領域の自律車両の群れ（Swarm）を、更に、群隊化する。
- ミッションに基づき、間接射撃、精密兵器攻撃、諜報・監視・偵察を、数千～数万の機体の動作を自律的に計画し、制御/実行する。 注目
- 多様なセンサーと運動・非運動エフェクターを備えた低コストの群隊を、主に前方に事前配置。
- 開発の重点は、指揮統率システム。 新しい自律型プラットフォームを開発するのではなく、既存の小型で低コストのプラットフォームを活用する。 注目
- トリガーを引いたり、爆弾を落とすかを自律して決定する能力を持つロボットシステム

（出典）岡島の推定図

Autonomy進展の方向性：連携Agent数の拡大



NSCAI, “Emerging Threats in the AI era.”, Chapter-1, Final Report, pp.77-81, 2021.
[2] NSCAI, “AI and the Future of National Intelligence”, Chapter-5, Final Report, pp.108-111, 2021.

Ubiquitous A.I. Integration ：Autonomyの源泉

✓ 全A.I.-Systemsを、ネットワーク・ドメインを超えて統合し、自律的に連携させるシステム[1]

① 兵器 / 兵士
② 通信ネットワーク
③ データを管理する支援AI
④ 指揮・統制・命令組織
⑤ 情報組織

✓ 人間が定義したTeamにて連携し、与えられたミッションを実現すべく、自律的に戦闘パイプラインを進捗させる [2]

目次

1. 本調査について
2. Autonomy拡大のトレンド：Ubiquitous A.I. Integration
3. Autonomy実現に向けてのソリューション技術
   3.1 Digital Nervous System
   3.2 A.I.連携 と Human-Machine Collaboration
   3.3 人間中心主義
4. 考察
   4.1 Agent連携 と 人間中心主義
   4.2 A.I.リスク
5. まとめ と 所感

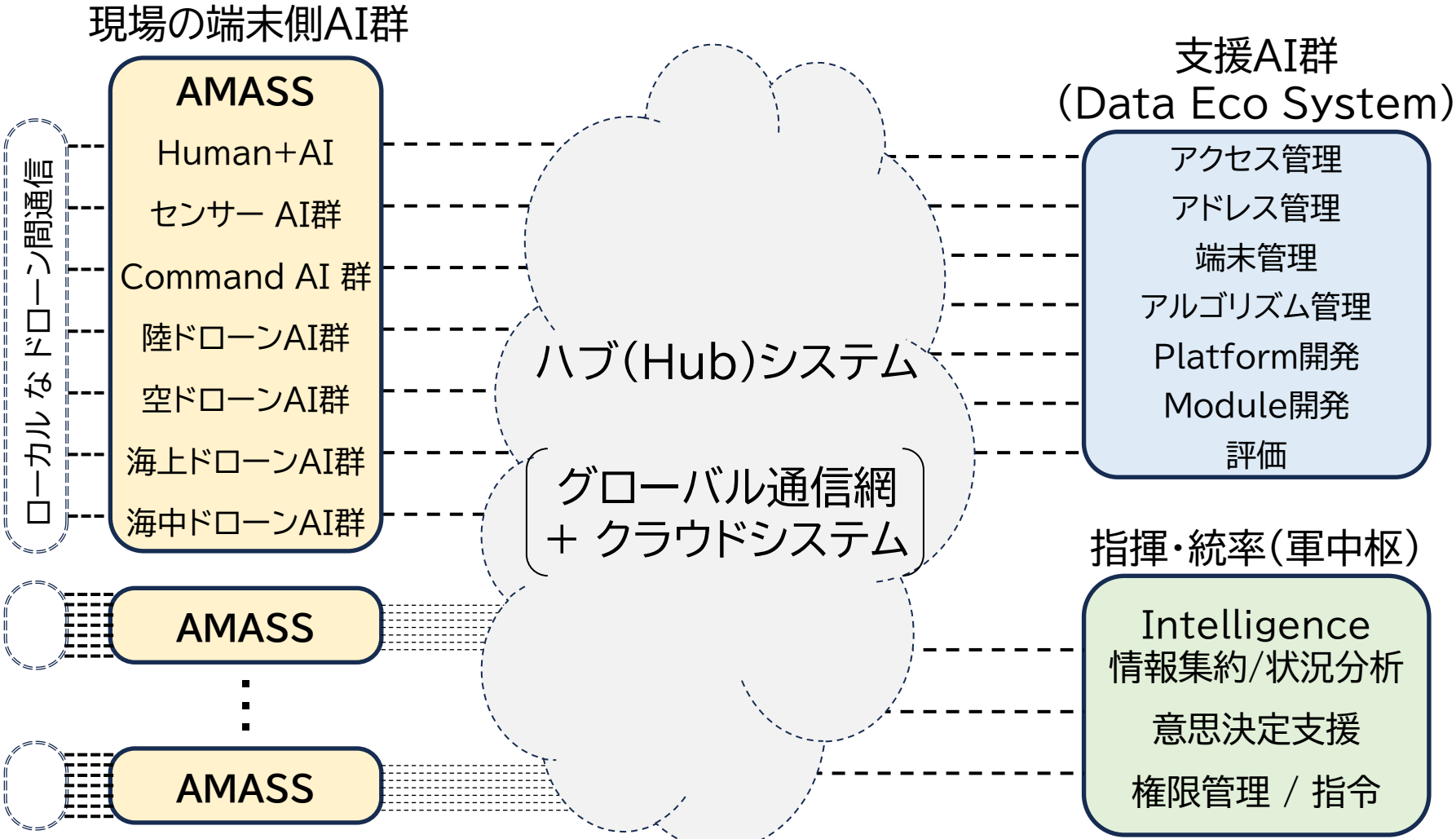
3.1. Ubiquitous A.I. Integration のための技術ソリューション

	（名称）	（関連技術）
求める機能	: Agent 連携	
アーキテクチャー	Digital Nervous System（含A.I.）+ Data Architecture	
構成要件		
開発/投資案件	: 次世代通信基盤	: 衛星コンステレーション、等
	: AI実装クラウドサーバー	
	: AI実装各種端末Platform	: Packaged AI 環境、等
	: AI実装センサー群	: データ・カタログ、等
	: AI実装Human-Machine-IF	
支援システム（クラウド）	: Data Eco System	: 共有AI資源（開発/管理）
	: アクセス制御	: ネットワーク・セキュリティ
	: 連合（分散型）リポジトリ	: Federation化クラウド
	: AI Marketplace	: AI能力開発/テスト/評価

軍用ITインフラ

3.2 Digital Nervous System

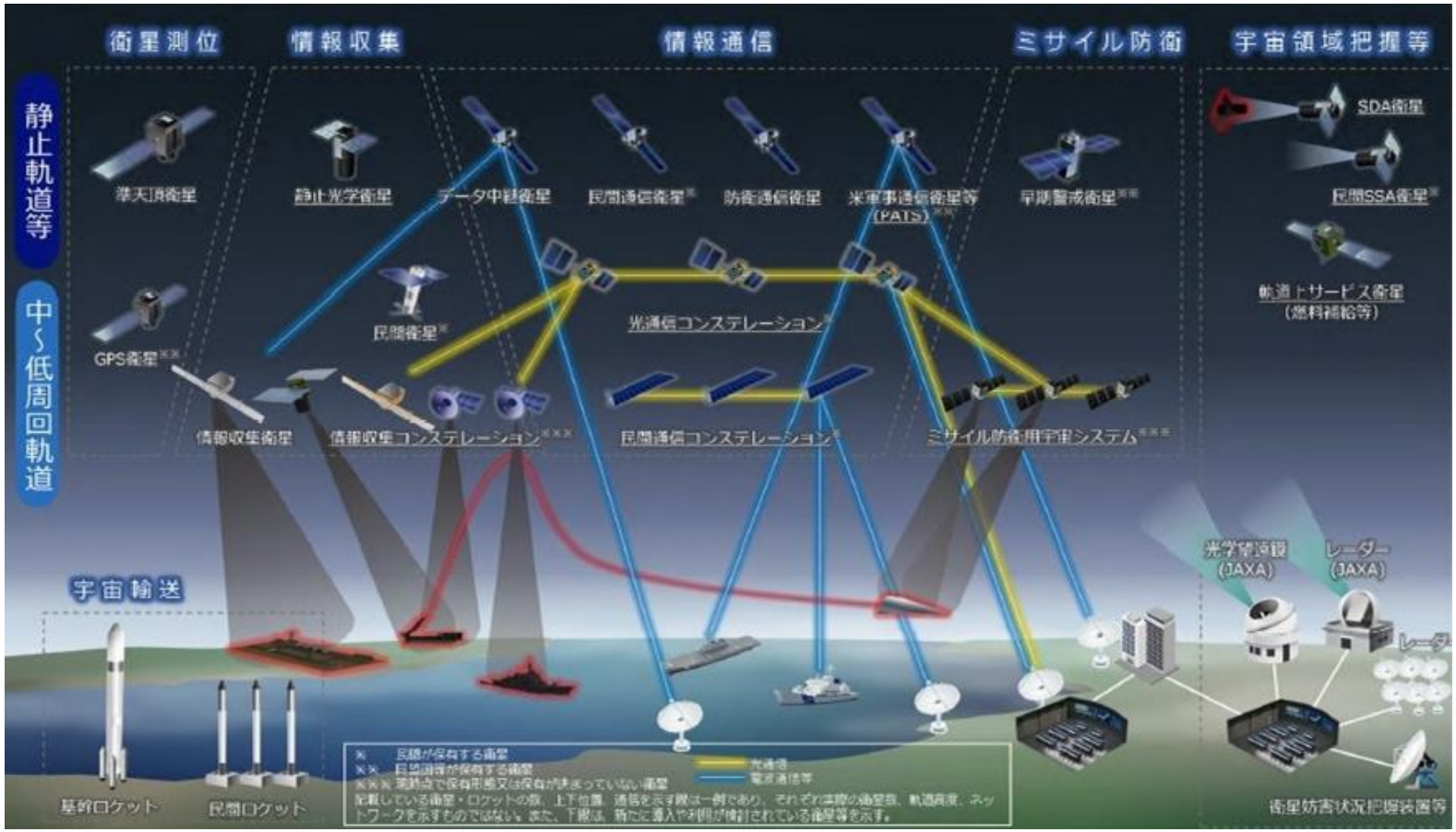
AMASS : Autonomous Multi-Domain Adaptive Swarms-of-Swarms)



出典：内閣府 宇宙開発戦略推進事務局；宇宙安全保障構想の概要(案),資料3-1, pp.5,令和5年

3.3 Digital Nervous Systemを支えるグローバル通信網

軍用のネットワークにては、静止衛星に加え、大量の小型衛星（衛星コンステレーション）のメッシュネットワークや、大気圏内の無人航空機を用いる。



[1] NSCAI, “Emerging Threats in the AI era.”, Chapter-1, Final Report, pp.45～51, 2021.

3.4 A.I.連携させるための 準備（その1）

✓ AI による分析が、企業と消費者の関係を変えたように、AIシステムに対するサイバー攻撃が成功すると、政府と個人、A.I.とA.I.の関係も変えてしまう。

✓ 敵のAI攻撃からの防御（共通のデータセキュリティ、サイバーセキュリティ）

- アクセスコントロール技術（認証技術）
- AIを用いたサイバー攻撃への防御
- 個人情報保護/データ保護の技術
- コンテンツの信頼性と出所を証明する技術
- 不正アクセスの追跡技術
- Red Team

AIセキュリティ・フレームワークの構成技術

- 外国投資審査
- Fact Check と サイバー攻撃への対応組織（外国悪影響対応センター）の設置
- サプライチェーンや重要インフラのインテリジェンスとリスク管理

3.5. A.I.連携させるための 準備（その2）： 共通言語

✓ 共通のData Architecture

- A.I.間のデータ通信で用いる共通のコード
  （「共通コード」には、Linkアドレスが付き、定義を参照可能とする）

✓ 共用データ の Federation管理

- ストレジやプラットフォームが扱うデータを共通のStrategyにて管理する。
  - ＞ トレーニングと評価に関するデータ群
  - ＞ パラメータ群
  - ＞ ツール群

3.6 人間中心主義

✓ AIは、軍の「任務指揮」の原則によって管理される（人間の指示に従う）

- AIは、意思決定プロセスの一部を自動化する。
  （軍人が任務の過程で知覚し、理解し、決定し、適応し、行動する方法を改善する）
- AIは、人間からの権限委譲後に自律動作可能とする。
- 人間中心の戦い方は、当分の間、標準のままとする。

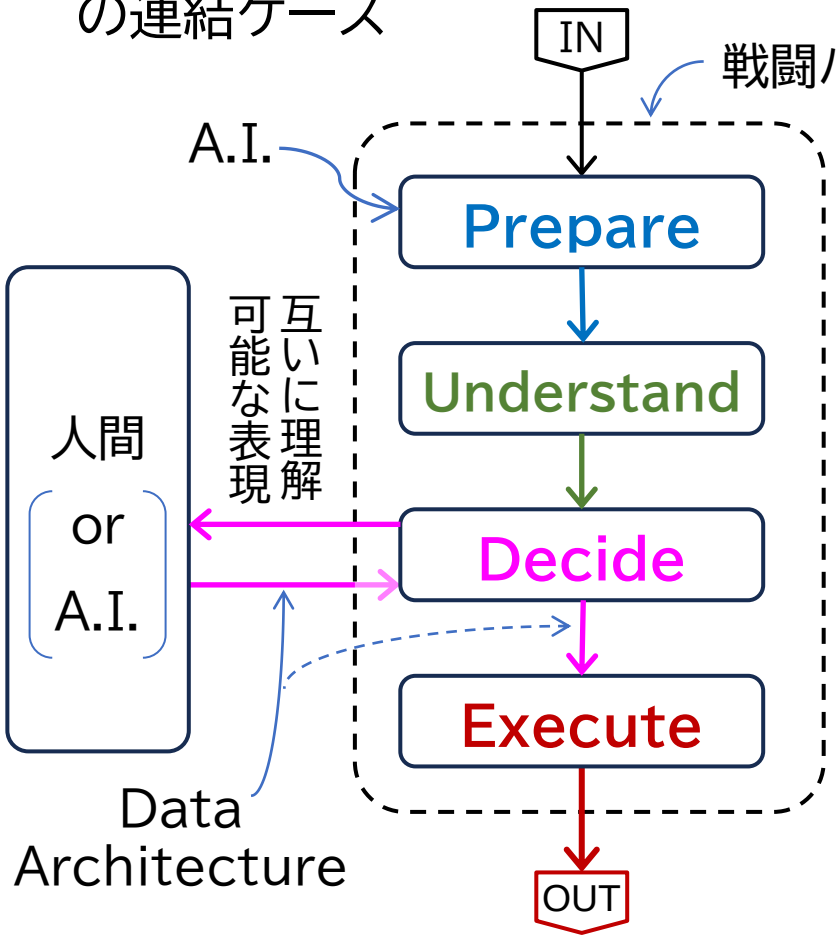
✓ 人間による承認と権限委譲前には、攻撃の実行Agentに、攻撃に関するデータ（標的データ等）を渡さない。・・・> 戦闘パイプラインの構造に影響

✓ AIは、予め、人間にとって有益なAIとなることを学習すること

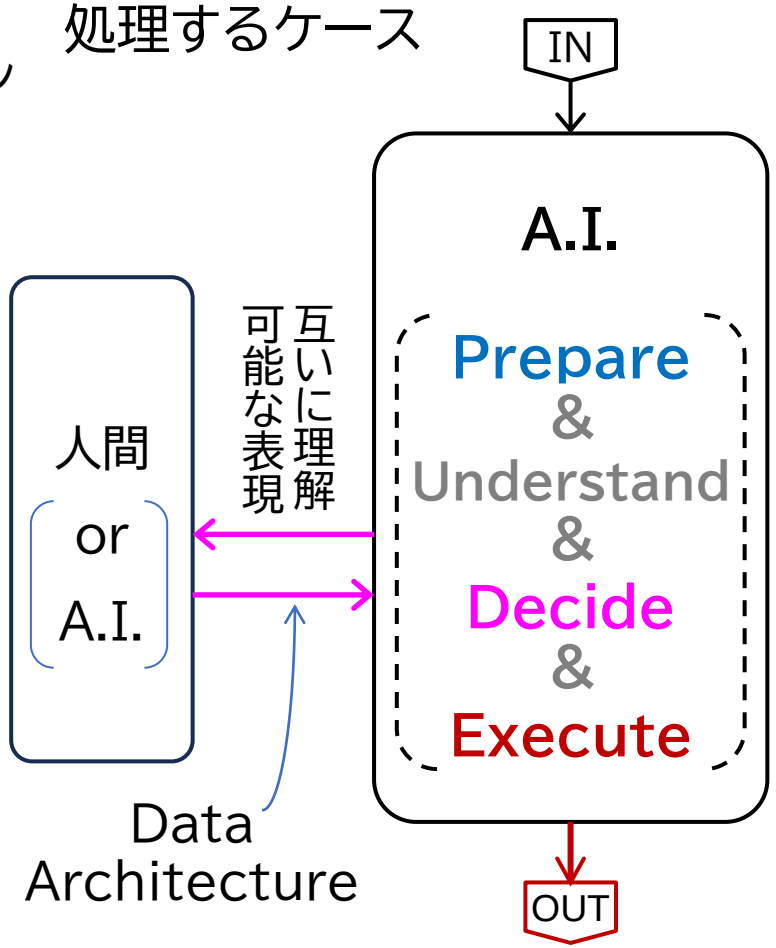
- 言語の理解、常識の理解、説明可能なAI、偏見を検出して緩和する

3.7 A.I.連携 に関する 疑問

（A） 複数エージェント
の連結ケース



（B） 単独A.Iが複数
処理するケース



目次

1. 本調査について
2. Autonomy拡大のトレンド：Ubiquitous A.I. Integration
3. Autonomy実現に向けてのソリューション技術
   3.1 Digital Nervous System
   3.2 A.I.連携 と Human-Machine Collaboration
   3.3 人間中心主義
4. 考察
   4.1 Agent連携 と 人間中心主義
   4.2 A.I.リスク
5. まとめ と 所感

（出典）岡島の推定図

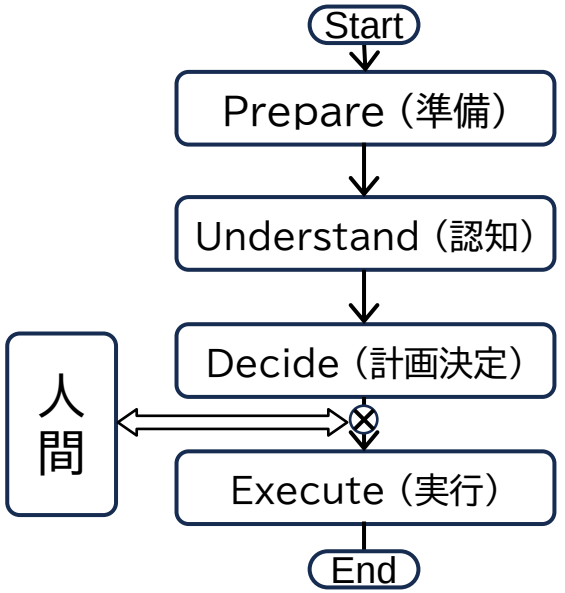
4.1 A.I.連携（無人化）と 人間中心主義（人間の介在）

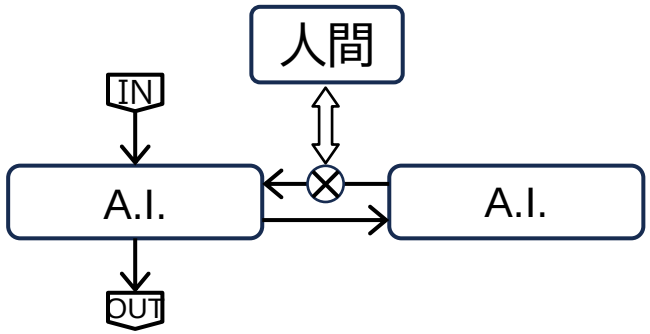
1） ドメイン内の連携（縦連携）

- 戦闘工程パイプラインのGating

2） ドメイン間の連携（横連携）

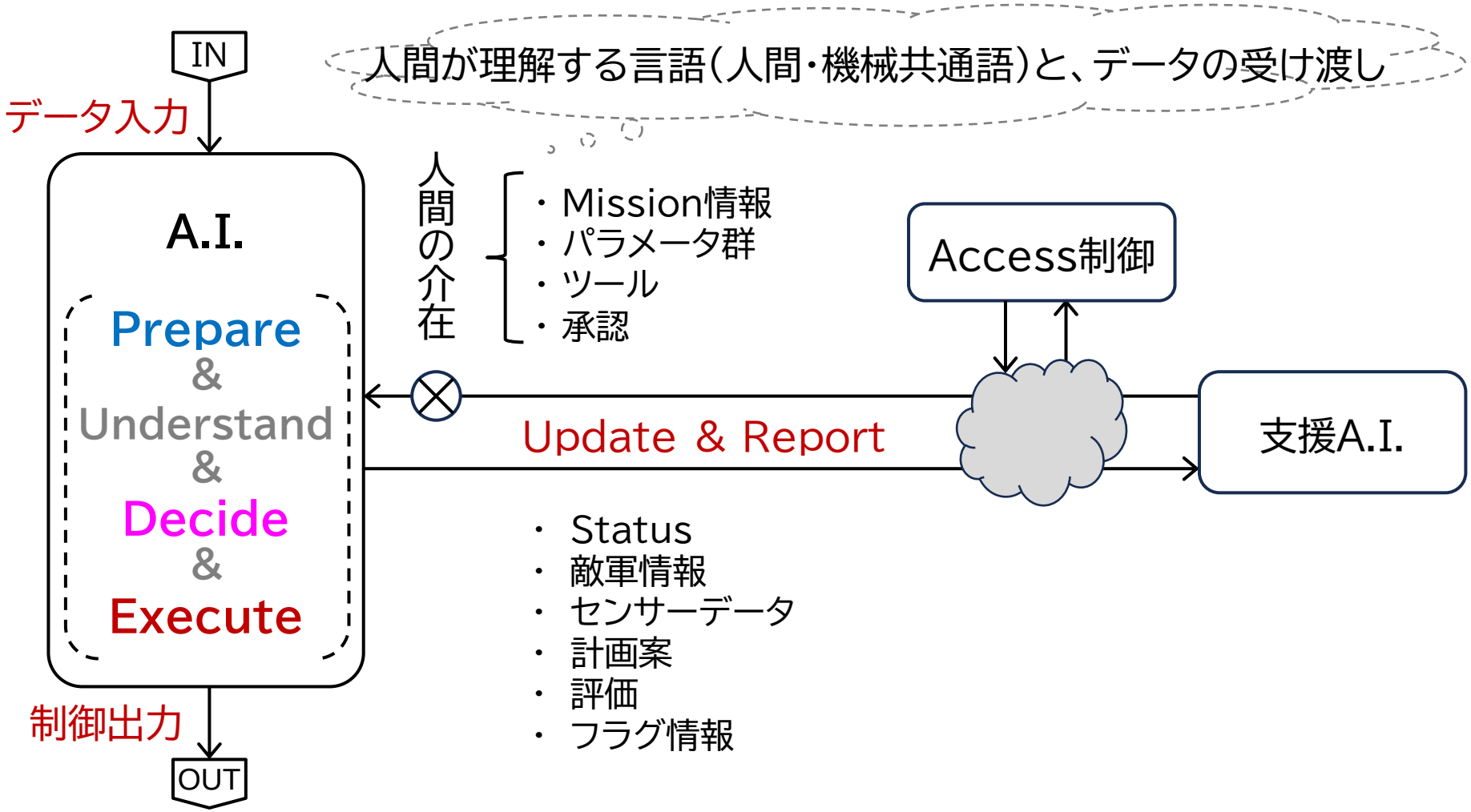
- 支援A.I.との連携
  - ＞ A.I.のUpdate
  - ＞ Missionの実装
  - ＞ 提案の承認
  - ＞ タスクの委任





（出典）岡島の推定図

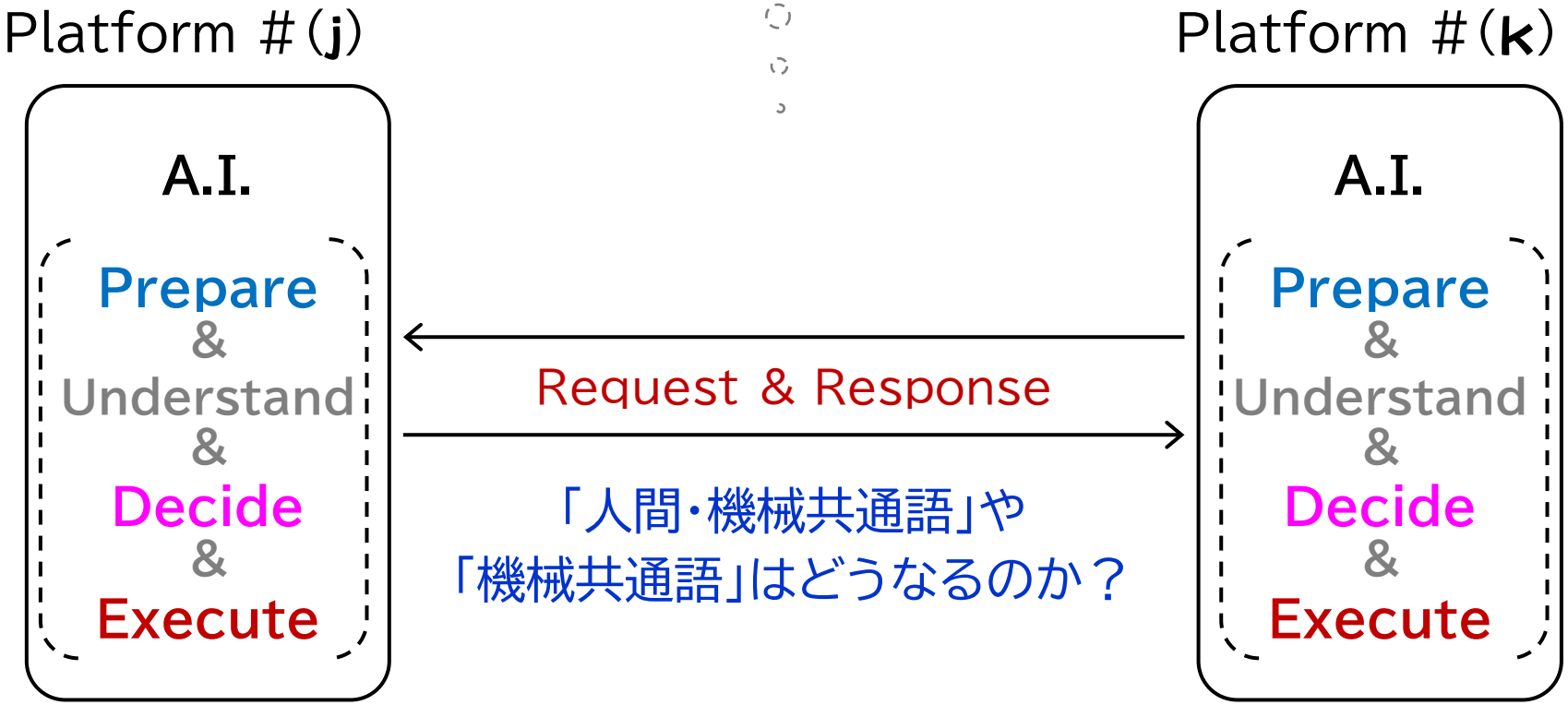
4. 2 Agent 連携 ① ：支援A.I.との連携



（出典）岡島の推定図

4. 2 Agent 連携 ② ：Platform間連携

互いに正確に理解可能な機械共通語によるデータの受け渡し



NSCAI; “AI and Warfare”, Chapter-3, Final Repoer, pp.75-88, 2021.
NSCAI, “AI and the Future of National Intelligence”, Chapter-5, Final Report, pp.108-111, 2021
NSCAI; “Autonomous Weapon Systems and Risks Associated with AI-Enabled Warfare”, Chapter-4, Final Report, pp.89-106, 2021.

4.3 A.I.統合 と AIリスク（その1）

✓ 人間中心主義」の遵守に関する懸念

1）「A.I.連携」は「人間中心主義」に反する？
人間を超える能力の追求競争は、人間の介在を減らす。

2）「指揮・統率」の自律化は、「人間中心主義」と相反する？
大局的判断を必要とする指揮統制機能を自律化するのは、異なるレベルのリスクを持ち込むつ可能性がある。

（そもそも、複雑な「指揮・統制の目標や価値」を人間は表現可能か？）

3）「学習十分」をどのように評価するのか？
トレーニングや評価が十分でないまま、実戦や演習にて使われると、予期しない動作を行う可能性がある。

NSCAI; “AI and Warfare”, Chapter-3, Final Repoer, pp.75-88, 2021.
NSCAI, “AI and the Future of National Intelligence”, Chapter-5, Final Report, pp.108-111, 2021
NSCAI; “Autonomous Weapon Systems and Risks Associated with AI-Enabled Warfare”, Chapter-4, Final Report, pp.89-106, 2021.

4.3 A.I.統合 と AIリスク（その2）

✓ 端末A.I. / 支援A.I. / クラウドインフラの管理に関する懸念

- 幾何級数的に増加するAgent群の連携動作は管理可能か？
- 巨大な支援A.I.のセキュリティは非常に重要
  （「A.I.が狂った」、「ハッキングされた」を、どのように検知するのか？ ）
- クラウド側のデータ管理、計算機構の管理、通信ネットワークの管理を民間企業に委託するのは妥当か？
  （民間企業が信頼可能と見做すエビデンスを何に求めるか？）

目次

1. 本調査について
2. Autonomy拡大のトレンド：Ubiquitous A.I. Integration
3. Autonomy実現に向けてのソリューション技術
   3.1 Digital Nervous System
   3.2 A.I.連携 と Human-Machine Collaboration
   3.3 人間中心主義
4. 考察
   4.1 Agent連携 と 人間中心主義
   4.2 A.I.リスク
5. まとめ と 所感

5.1 まとめ

・ 1章 ： 今回の 調査目的 （兵器の自律化トレンドの俯瞰と、A.I.リスク）と
調査方法 （米国政府への報告書）

・ 2章 ： 軍事のAutonomyは、全Agent統合（シングルトン化）を目指す。

・ 3章 ： 全Agent統合に向けての軍用ITインフラが刷新される

各種ハードウエア開発（Module開発、Platform開発、A.I.技術、
開発ツール環境、セキュリティ技術） の開発が進んでいる。

・4章 ： 全Agent統合を実現するには、
＞ 人間・機械共通語 や 機械共通語 が必要ではないか？
＞「自律性の向上」は、「人間中心主義」に反する方向となる可能
性ある。

5-2. 所感

「イングランドのボスワースの戦い、日本の関ケ原の戦い、インドのプラッシーの戦いといった決戦で結果を左右したのは、結局のところ裏切りだった。」

By S. Jaishankar（インドの外相）
“The India Way : Strategies for an Uncertain World”
（インド外交の流儀：先行き不透明な世界に向けた戦略）
白水社、pp.74、2022年12月

① 現代の認知戦は、「A.I.の裏切り」を工作する戦いとなる。

② 自律性の拡大は、人間中心主義と背反しうる。

③ この巨大な開発は、現在のITインフラを刷新する「将来の民間製品向け技術」に転用される。

ご清聴ありがとうございました。