# 7割が失敗するAIの研究開発・運用の現状と活用のための課題
## ―AI研究開発・運用の方向性に関する一考察―

川岸　卓司

目　次



## はじめに

近年「AI(Artificial Intelligence)」「人工知能」と呼ばれる技術が様々な分野で着目されており、陸上自衛隊においても様々な分野への導入が検討されている。これは、宇宙・サイバー・電磁領域などへの新たな脅威に加え、従来の活動領域への自律性致死兵器の登場により、従来通りの対応では必ずしも有効に対処しえないことが原因として挙げられる。一方、AIの専門家ではない陸上自衛官がAI技術を適用した装備品を防衛産業と議論する際、前提知識や導入課題の認識が異なっていることから、具体的な議論が進みづらい。また、様々な研究会や企業展示においてもAIのメリットが多く語られ、AIに関連するプロジェクトの開発・運用の7割近くが失敗している現状や問題点が語られていないように感じられる。このような中、防衛装備に関するAIの課題や活用について述べられているものは少ない。本論文では、AIの仕組みを明らかにし、AI技術を適応した装備品の研究開発・運用に係る課題を明らかにし、防衛産業を含む防衛省としての研究開発・運用体制の方向性について考察する。

## 1　AIの活用現状とデータの関係性

### (1)　本章の目的

近年、仕事に慣れた熟練者が実施していたような仕事も、AIにより一部の分野では人間以上の性能となっている。例えば、AIは画像を見て対象が何かを識別する実

験では人間の識別率を上回っており[1]、空港などの莫大な監視カメラの映像から個人を特定する人間でも困難なことができる。一方、近年の AI 活用議論の中、自衛隊でどのように活用できるのかについて明確に答えられる人は少ないと思われる。AI による軍事機能の代替え可能性について、小野圭二[2]がフレイ（Carl Benedikt Frey）とオズボーン（Michael A Osbone）の報告書[3]による民間での分類を軍事機能に適応し、軍の各機能で代替えできる可能性を論じている。論文によれば、幕僚組織では総務、情報、法務、副官業務（スケジュール調整）が可能性であり、戦闘部隊では警戒警備の一部、支援部隊では整備や輸送が AI による代替えの可能性があるとしている。一方、国内において AI による防衛分野への適応に関する課題や体制について述べている論文は小野の論文の他確認することができなかった。本論文では、AI の代替可能性に関する議論を具体的な開発・運用に言及するため、今実現可能な事柄について、AI の普及が進んでいる民間での事例や学術研究を元に自衛官が想像容易な 1 系（人事）、2 系（情報収集・分析）、3 系（作戦のための見積）及び 4 系（整備）の機能別で示す。また、次章以降の AI の仕組みや課題に関連する実現に必要な要素である「どのようなデータ」があれば AI 技術の適用が実現できるのかについて述べる。

### (2) 事例の紹介と必要なデータ

#### ア 1系（人事）

##### (ｱ) 事 例

HR テック（Human Resources Technology）と呼ばれる技術は、AI やビッグデータ解析などを駆使して採用、人材育成、評価、配置などの効率化と高品質を提供するシステムであり、多くの企業で導入が進められている[4]。採用では、採用率の高い時期や場所へ求人広告を出し、面接者の能力や希望を組み合わせ最適な職場を提案できる。また、将来必要な人材を育成するため、各職員に合わせたタイミングで次の配置や案件を提案するとともに、不満を抱き退職を考える前に適切な配置やケアが実施できる。また、関わったプロジェクトや利益、契約数からその職員を定量評価して給与システムと連携させ評価へ反映することもできる。このようなシステ

---

[1] Florian Schroff, Dmitry Kalenichenko, James Philbin,"A Unified Embedding for Face Recognition and Clustering",2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)DOI: 10.1109/CVPR.2015.7298682.

[2] 小野圭二『人工知能（AI）による軍の知的労働の代替』防衛研究紀要第 21 巻第 2 号、2019 年 3 月、13-14 項。

[3] Carl Benedikt Frey and Michael A. Osbone,"The Future of Employment; How susceptible are jobs to computerisation?"Oxford Martin School Working Paper, Univ. of Oxford(Sep.,13)

[4] Paul R. Daugherty and H. James Wilson,`HUMAN＋MACHINE`保科学世訳、東洋経済新報社、2018、70-74 項。

ムにより、人事担当者は紙と PC を見る時間に割いていた時間を削減し、人と人との直接的なコミュニケーションに当てることで、離職率を下げることやなども期待されている[5]。

### (イ) データの種類や取得方法

HR テックを運用するためには、採用時から現在までの各人の経歴や、能力、評価に加え、資格の有無や外国語の履修度合い、年齢や家族の有無や持ち家の有無などを入力したデータセットが必要である。また、転属や配置転換の度にアップデートし蓄積していくことが必要である。また、このような基礎状況に加え、心情や意識などを把握する要素（やりがい、不安事項や満足度など）が必要になると思われる。つまり、現在の人事担当者が資料や調査で使用しているものがあれば、それをデータベースとして適切な形で保存されていることが必要である。これらのデータに加え現在の人事担当者が経験的に考え、実施している事項を手順化することにより導入できる可能性がある[6]。

## イ 2系（情報収集・分析）

### (ア) 事 例

現在、インターネットや携帯電話の普及により、様々な情報が全世界中で発信されており情報が溢れている。このような現状で国際情勢やその影響を分析する情報収集するひとつの手段である公開情報（Open Source Intelligence：OSINT）収集のため、全世界の新聞やディジタル記事、他国の声明や議議決案に加え安全保障に関連する法律や経済状況など様々な情報を収集する必要がある。またこれらの情報は、その国の特性や過去の歴史を理解した人物がこれらの情報を分析することで、将来の予測や推察が可能であると考える。類似する事例として、ゴールドマンサックスの予測 AI を紹介する。株式や債券、通貨のブローカーから保険や投資銀行業務を行うゴールドマンサックスは、世界中の株価が与える影響を日々分析し、一瞬の判断が利益に直結する状況判断や長期的なリスクなどを含め日々膨大な情報を分析したレポートを作成し、意思決定に活用している[7]。このレポートを自動で作成する AI が検討されている。100 万件のアナリストレポートを AI が解析して作成し、人が作成していたものと遜色なく完成していたという。また、今までレポートを作

---

[5] Dianna L. Stone, Diana L. Deadrick, Kimberly M. Lukaszewski, Richard Johnsond,"Human Resource Management Review"Human Resource Management Review 25 (2015) 216–231.

[6] Daugherty and Wilson, HUMAN＋MACHINE, Pages 68-69.

[7] Nathaniel Popper, "The Robots Are Coming for Wall Street", New York Times, Feb. 25, 2016.（2019 年 12 月 10 日アクセス）
https://www.nytimes.com/2016/02/28/magazine/the-robots-are-coming-for-wall-street.html

成していた担当がAIのレポートを評価することで、AIが学習し次のレポートの完成度が高まっていく仕組みとなっている。

#### (イ) データの種類や取得方法

AIの作成したレポートは、人が参照していた様々なデータや類似のデータなどの情報をもとに100万件にも及ぶ情報から作成されている。このため、人が従来収集していた資料や、ネット上での検索キーワード、過去に状況判断に使用した数値などAIが「何を学習するのか」について多大な調査と研究をしていると思われる。さらに、実際の運用シーンではAIの作成したレポートを専門家が添削することで精度と信頼性を高めていると考えられる[8]。

### ウ 3系（作戦のための見積）

#### (ア) 事 例

Google Deepmindが実施するオンライン戦略ゲームでの研究事例を紹介する[9]。このゲームは、２つの部隊に分かれ敵を殲滅することで勝利を得るゲームである。この際、敵情の偵察をしつつ、自らの部隊の育成、資源の確保及び建物や障害の構築に関する命令をユーザが同時並行的に実施する必要がある。さらに、敵と接触した場所での攻撃命令などを入力する必要があり、長期的な戦いで必要な資源確保や建物への人的資源の全体最適化と局所戦闘での状況判断を適切に実施することが求められる。これらをAIに実施させる場合、AIが長期的な計画を作成しつつ、状況変化に対応して各種（人材育成、経済、築城等）計画修正を行いながら、局所戦闘での戦い方を計画し、命令することが求められる。このような長期計画と計画の修正の作業は、囲碁や将棋と異なりルール化が難しくAIが人間に勝つことが難しいと考えられていた。しかしながら、研究結果によると、このゲームの人間のプロはAIに勝つことはできなかった。

#### (イ) データの種類や取得方法

本事例では、このゲームのプロゲーマの操作データを当初の学習に使用した。この際のAIが観測したデータは、全体的な視点（人間のユーザも偵察した領域は全体的な視点を見ることができ、偵察していない場所は見られない、AIは人と同様の条件）、経済性（人材育成及び資源のステータス）及び各戦闘での戦い方（位置情報、障害、射撃等の状況だと推察）である。ある程度学習した後は、パラメータを少し変えAIが敵と味方に分かれ互いにゲームを実施して、より強いAIを決めてい

---

[8] Daniel Russo, "Get the AI advantage: Start your insight engine for business" IBM Watson blog, Jan. 30,2018.(2019年12月10日アクセス)
https://www.ibm.com/blogs/watson/2018/01/start-your-insight-engine-for-business/

[9] Google DeepMind" AlphaStar: Mastering the Real-Time Strategy Game StarCraft II " 24 Jan 2019.（2019年12月10日アクセス）
https://deepmind.com/blog/article/alphastar-mastering-real-time-strategy-game-starcraft-ii

く手法で自ら学習をしていった。勝った AI の内部をアップデートし、パラメータを少し変えた AI を作り次々とゲームを繰り返し、AI 自身を強くする強化学習と呼ばれる手法である。この繰り返し作業を並列分散処理によって 175 万時間分のゲームを実施[10]し、人に勝る状況判断が可能な AI を作成した。

#### エ　4 系（整備）

##### (ｱ)　事　例

故障前のわずかな異常兆候を発見して整備を促す GE(General Electric Company)の異常検知 AI について紹介する[11]。GE では、ボルトや発電機など様々な情報をディジタル情報で保管し、これを実物のボルトや発電機のセンサデータと組み合わせシミュレーション可能なディジタルツインと呼ばれる技術で管理している。これにより、今までの定期的な交換（△年あるいは○km 動いたら交換）ではなく、実際のボルトや発電機の状態を記録した莫大なセンサデータから長期的なシミュレーションを行い、個々の発電機特性を分析することで、それぞれの発電機に最適なタイミングで交換時期を予測できる。

##### (ｲ)　データの種類や取得方法

ディジタルツインは、コンピュータの仮想空間上に CAD などの 3 次元情報で部品を保持し、現実世界の部品に取り付けたセンサ情報をこの仮想空間に入力することで実際の部品への応力や温度を仮想空間で把握することができる。また、これらの部品を組み合わせた機器でのシミュレーションのみならず、システム全体（事例では、GE の発電機 1 万機）でのシミュレーションによって整備だけでなく、機器への負荷が最も少なく最も発電効率の高いプロペラ風向の最適化などができることを示している。さらに、実環境での故障やその原因を推定し、上記のシミュレーションへ導入することで、さらに予測精度を向上することができると考えられる[12]。

## 2　AI の仕組みとデータの関係性

### (1)　本章の目的

本章は、AI にさせたいことが明確になった後に「それを行う AI が作成可能なのか」を考察するため AI の仕組みとデータの関連性について解説する。AI の仕組みでは、機械学習やディープラーニングといった AI 作成に用いる手法そのものの数式的な解説ではなく、AI を作成する上で必要となる“AI の使用者が提供すべきデ

---

[10] 175 万時間（24 時間×365 日×200 年）のゲーム時間は、実際の年月ではなく、同時並行での累積時間かつ時間を数倍(例えば 1 秒を 0.1 秒)にして実施された結果、人間と同じ時間尺度で実施した場合の時間を表記している。

[11] Daugherty and Wilson, HUMAN+MACHINE, Pages 36-37.

[12] Ibid, Pages 41-46.

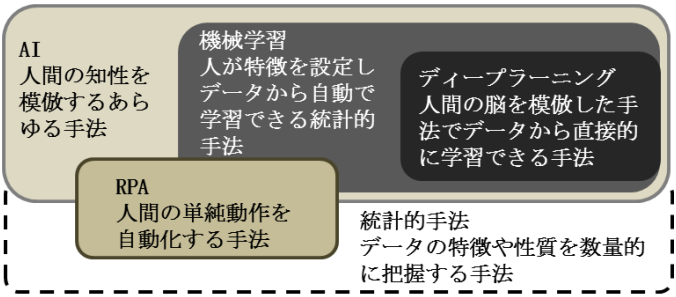
ータ”（以下、必要なデータ）について述べる。また、人間の作業効率化のためで導入されることの多い RPA(Robotic Process Automation)[13]とビッグデータとの関連について整理する。

### (2) AI の目的に応じた必要なデータ（AI の仕組み）

#### ア AI の変遷と RPA

AI で何かをさせるための必要なデータを考察するには、AI 作りのプロセスを理解することが重要となる。AI と呼ばれる「人の知的活動を補助または部分的に代替えする技術」[15]を作成する手法としてデータを AI に学習させる機械学習が用いられている。第 1 次 AI ブームはプログラムによる論理型、第 2 次 AI ブームは専門家が回答を入力しておき、聞けば答えてくれるようなエキスパートシステムがメインであり、共に演繹的な手法が用いられていた。近年の第 3 次ブームは 2010 年頃から現在（2019）を指し、従来とは異なり大量のデータを学習に用い帰納的に計算するものとなった[16]。一方、紙媒体の資料をパソコンに入力する作業や、表計算ソフトのデータをまとめたりするような単純作業を自動化する RPA と呼ばれる分野も AI と呼称することがあり、人の知的活動が広範囲にわたることから様々な定義が存在している[17]。（図 1）

図 1 AI に関連する用語の定義と範囲[14]




13 ソフトウエアロボットが業務プロセスを自動化すること

14 MathWorks,” Introduction to Deep Learning: Machine Learning vs. Deep Learning”,Mar. 24,2017.を参考に著者が作成（2019 年 12 月 10 日アクセス）
https://jp.mathworks.com/videos/introduction-to-deep-learning-machine-learning-vs-deep-learning-1489503513018.html

15 人工知能と人間社会に関する懇談会の定義を引用：日本が目指す Society5.0 を実現するための AI の研究開発及び活用のための社会的な課題や期待事項を検討するために 2016 年に内閣府に設置された懇談会

16 I/O 編集部『ビッグデータ＆人工知能ガイドブック』工学社、2017 年、25-38 項。

17 藤本浩司、柴原一友『AI にできること、できないこと』、日本評論社、2019 年、2-3 項。

### イ　ビッグデータと AI の関係性

情報化社会の進展により、個人の保有の携帯端末から発せられる位置情報から金融情報、SNS、WEB 履歴に加え、物がすべてネットに接続され様々な情報を送信する IoT(Internet of Things)が進展している。これにより保存されたテキストデータや数値データ、画像データなどの莫大なデータはビッグデータと呼ばれている。ビッグデータの AI による活用事例を示す。コンビニでの年齢、時間、場所などが入力された購買履歴データが数年間分ある場合、「どのような商品がいつ売れそうかを予測する AI」というのは実現できそうなイメージを持つ印象を受ける。しかし、顔写真のみから購買予測はできていないように AI の目的に応じ適切なデータがビッグデータに含まれていなければ、データの取得方法から考察する必要がある。現在、人間を越えるような分野で活用される AI は目的に応じた適切なデータから学習しており、「適切さ」にはデータの量や質が影響する。例えばアンケート結果や画像データ、センサデータなどから得られるデータはまず数値に変換される。アンケートのようにある時点で質問に答えデータを作成するには数量化や尺度化[18]といった考えに基づき、アンケートで評価したい特性や特徴に数字を割り当て数値データに変換している。温度センサなどが時間変化することで目的に影響する場合、温度センサの時間単位における時系列的な要素が重要となる場合がある。このようなデータの数量化や尺度、時系列に求められる質[19]については「数字が特徴や現象のどの側面をどの程度まで表現しているのか」[20]が重要となる。陸上自衛隊の演習場における部隊の活動を分析するならば、彼我の位置情報、地図データは等高線や植生、気温や天候、その当時の命令や通信の可否、分析目的に合わせた適切な質と量のデータが必要だと思われる。しかし AI の目的によっては上記の地図縮尺では不十分であり、木の密度や車両の燃料、隊員の疲労度など、まったく別のデータがないと部隊の活動を表現できない可能性もある。このようにビッグデータがあるからなんでもできるわけではなく、目的と対象を表現するデータの質と量が AI の性能のカギとなる。

## 3　AI の目的に合わせたデータと機械学習手法

### (1)　本章の目的

AI 仕組みから、質と量が十分なデータが重要であることが示された。一方、大量のデータから帰納的に出力する現在の AI はどのような目的に対して使用できるの

---

18 岩淵千明『データの処理と解析』、福村出版、1997 年、18-19 項。

19 時間分解能は msec か min、温度変化の記録幅は Δ100℃か Δ0.001℃など目的に応じて分解能＝質が異なること

20 岩淵、データの処理と解析、16 項。

表 1 AI の目的に応じたデータと機械学習手法の関係

| 目的別区分 | 例 | データ | | 手法の参考例 |
|---|---|---|---|---|
| | | 種類・量の例 | 留意事項 | |
| ①提案・予測系 | ・気象予測<br>・地域見積<br>・故障予測<br>・行動提案<br>・OSINT | ・観測データ<br>・時系列データ<br>・センサデータ<br>・行動ルールの定式化<br>（大量データ） | ・提案・予測結果の正解の有無<br>・シミュレーション実施の可否<br>・失敗のリスク許容・対策 | ・線形回帰<br>・SVM<br>・ロジスティック回帰<br>・NN　・決定木<br>・ベイジアン |
| ②画像・レーダ系 | ・探知・識別<br>・監　視<br>・画像改善<br>・SIGINT, IMINT | ・ラベル付データ<br>・観測環境データ<br>・実測データ<br>（大量データ） | ・実環境での「正解データ」の有無<br>・実レーダやカメラのなど使用機材の有無 | ・RNN　・CNN<br>・DNN<br>・クラスタリング<br>・LSTM |
| ③言語系 | ・音声認識<br>・文字認識 | ・収集作成データのアノテーション<br>・ラベル付データ<br>（大量データ） | ・実環境での「正解データ」の有無<br>・失敗のリスク許容・対策 | ・SVM<br>・word2vec<br>・LSA　・DNN<br>・線形回帰 |
| ④サイバーセキュリティ系 | ・ペネトレーション<br>・マルウェア対処<br>（監視、検知、処置） | ・マルウェア等の大量データ<br>・通信・実行ログ等攻撃時の振舞いデータ | ・システムの可用性<br>・サンドボックスやテスト環境の有無<br>・失敗のリスク許容・対策 | ・SVM<br>・word2vec<br>・LSA<br>・ロジスティック回帰　・NN |
| ⑤最適化・ゲーム系 | ・UAV最適設計<br>・兵站経路<br>・チェス、碁、将棋等<br>・指揮所訓練統裁支援システムの仮想敵作成 | ・各種定式化 CAD<br>・3D地図や諸元データ<br>・明確なルール及び教師データ | ・明確なルールとシュミレーションでの反復可能性<br>・実環境での有効性の検討 | ・RNN<br>・CNN<br>・DNN<br>（強化学習） |

か。本章では、AI 技術適用装備の参考になると考えられる①提案・予測系、②画像・レーダ系、③言語系、④サイバーセキュリティ系、⑤最適化・ゲーム系と整理[21]し、各区分に応じた実例やデータと機械学習手法について解説する。（表 1）

### (2)　目的に応じた区分とデータの関係性

#### ア　①提案・予測系

1 章で紹介した金融機関レポートの提案や異状兆候の検知などは、知識または経験がある熟練者が実施できていたことを、AI に代替えしていた。この際 AI が提案や予測の根拠としているデータは、様々な熟練者が用いているデータやセンサ情報などの大量の関連データから計算されていた。このように、AI が人間の経験や知識などプログラムで表現できないような思考を大量のデータから推測し、提案された結果と熟練者による正誤判定を繰り返すことで大量のデータから推測するための表現方法（モデル）を AI の内部で作り出している。機器の故障予測は、上記の熟練者が目や肌で感じるような温度や歪みを温度センサや歪ゲージなどを用い逐次

[21] AI 技術を防衛省の活用可能性から目的別に整理した事例がなく、著者が独自に区分したため学会や AI 白書の区分とは異なる。
AI 白書、独立行政法人情報処理推進機構、AI 白書編集委員会偏 2019、 323-331 項。

保存していき、このような時系列データから「正常動作」の状態を学習させる。そして、正常動作の状態から逸脱するような兆候があれば故障を予測することができる。上記のような提案・予測の AI を陸上自衛隊で検討した場合、OSINT の情報処理の一部など、現在の熟練した専門者がいれば、データ収集に日頃使っているサイトや検索用語を AI で学び、処理させレポートを作成させる。このレポートの添削を熟練者が実施する際、AI にチューニングをすることで適用できる可能性がある。また地域見積のような分析は、幕僚諸元のような見本となるルールと地図データを組み合わせ熟練者が正誤判定を繰り返すことで人間と同様に生成することができると考えられる。

一方、演習場や実際の行動を伴わない地図上のみで得られたデータは、実際の環境での「正解」となりえず、用途によっては演習場や図演でのみしか使用できない可能性を含んでいる。このような場合、ディジタルツインのような実空間を模す高精度な仮想空間でのデータの活用が必須になると考えられる。また、どんなに実戦的な訓練を実施したとしても、有事の際に本当にそのような状況なのかを誰も予測できないように訓練のデータから AI による実戦で適用可能な範囲については緻密に見積もることが必要であると考えられる。

### イ　②画像・レーダ判別系

この分野の AI は、すでに画像認識率が人間に勝っている例から、警戒監視ならびに映像情報(Imagery Intelligence：IMINT)など、多くの分野で適用可能な範囲が多い分野であると考える。一方、どのような AI も基本的にはラベル付データ[22]と呼ばれる画像と名前が一致しているデータが大量に必要である。この分野で精度が良いと呼ばれるディープラーニングを適応する場合、精度や目的に応じ一概には言えないが数万以上のラベル付きデータを学習させる必要があると言われている。可視画像から「戦車」を判別させたい場合、あらゆる角度から様々な戦車とわかる画像を入力する。インターネットで簡単に集められそうに感じるが、ここでデータの質が問題となる。インターネットには携帯や高価なカメラ、監視カメラなど様々な画像がある中で、「目的とする監視機材」での写り方が異なっていた場合、監視機材では使えずインターネット上の戦車画像を探してくることに特化した AI となる可能性が高い。また、雨や霧、山や市街地など実際の環境に準じた背景などの環境影響を網羅しても雪の環境が抜けていれば雪環境では使用できない可能性がある。また国内で日本の戦車を用いてすべての環境を網羅したとして、実際に偽装する敵の戦車を判別できるか確かめることは難しい。

---

22　ラベル付きデータ：AI に学習させる際にそのデータが何なのかを区別されたデータである。画像認識では、入力する画像に「戦車」のラベルをつけることで「戦車」とはどんな画像なのか学習できる。

ただし、沿岸監視やレーダなど今現在熟練した隊員が日夜判別しているような状況であれば、データを累積していくことで熟練した隊員と同等以上の性能で判別できるような装備は実現可能であると考えられる。この際、熟練隊員がどのような情報で総合判断しているかに留意したデータ累積が重要である。沿岸監視レーダであればレーダ波長、出力だけでなく潮や波、雨風や気温、季節などで判断していた場合は判断した時点のこれらの情報が重要となる。

### ウ　③言語系

言語系は文章解読や記入、音声認識などを対象としている。文章解読では災害時の Twitter 情報を水害や停電、避難数などを地図上に分類して表示するような AI はすでに実用化されている[23]。この例では、インターフェースを工夫し分類をユーザにタグ付けしてもらい、位置情報も付属しているため、見本データなどを学習する必要はない。一方、これを監視カメラ情報のみで実施する場合は②の技術が必要であると考える。 また作戦計画で記載された文章を読み取り、地図上に表示するような場合、文章内の地名と地図の地名が 1 対 1 で対応しており、地名呼称の表現が均一（○○一帯と呼称した際に周辺数百 m を指すのか、数十 km なのかが均一に定まっている）であれば可能だが 、状況に応じて異なるような場合は、それぞれの形容詞などを入力して学習する必要があると考えられる。

無線の声をテキスト化するようなものは、自衛隊の通常通話を学習することはもとより、無線の特性である微かに聞こえるような場合や偵察時の囁くような声から攻撃時の激動息切れの声、様々な背景雑音への対応から途切れ途切れで、なんとか内容が聞き取れる場合など多くのデータと学習が必須となることに留意が必要である。また、現在の状況を理解した上で無線の内容を予測しているような場合については、人間以上の精度は期待できない。

### エ　④サイバーセキュリティ系

サイバーセキュリティの分野では、悪意ある組織や個人が特定組織を狙った標的型攻撃により安全保障上の脅威の一部となっており、国家安全保障戦略や国際政治、市場利益、知的財産、安全保障及び軍事作戦などの各分野にまたがる問題の側面をもっている[24]。目に見えず、同時多発的に発生し兼ねないサーバー攻撃への取り組みに AI を活用する手法が検討されている。マルウェアでの攻撃にはシステムの脆弱性を攻撃するエクスプロイトと、目的のシステムに侵入した後に悪意のある動作

---

23　独立行政法人 情報通信研究機構、対災害 SNS 情報分析システム「DISAANA」（ディサーナ）を Web 上に試験公開（2019 年 12 月 10 日アクセス）
https://www.nict.go.jp/press/2014/11/05-1.html

24　“『標的型メール攻撃』対策に向けたシステム設計ガイド”．情報処理推進機構（2013 年 8 月）、2 項。

をするペイロードに区分される[25]。2016 年の国防高等研究計画局(Defense Advanced Research Projects Agency：DARPA)のプロジェクト Cyber Grand Challenge でサイバー攻撃及び防御を自動的に実施する AI の競技会が実施された際には、相手の脆弱性を攻撃しつつ自らオフラインにしてペイロードにも対処するものであったようだ。このように、攻撃と防御を AI で自動化することで人間には発見が難しく、脆弱性を見つけ対処するような抗堪性のあるシステムを構築することが期待されている[26]が、セキュリティカンファレンスの DEF CON 2016 で熟練者と AI の戦いの結果は AI の完敗であったことから、今後も人による様々な攻撃パターンの収集と対処の双方のデータベースの構築が重要であると考えられる。

### オ ⑤最適化・ゲーム系

最適化の例では、デザイナーが椅子を設計する際に強度や材料、座る面積などの制約条件を入力することで AI が自動的にこれらの条件を満たした椅子を設計している[27]。この際、AI には様々な形状の椅子のデザインを学習させておき、有限要素法、時間領域差分法、モーメント法のような力等の伝わりを仮想空間上で制限条件を満たすシミュレーションによって人間が思いもつかないような形状を作成することができると考えられる。このような分野は、小型軽量化などの部品の設計にも活用されており設計者が AI をチューニングすることで精度が向上していく。ドローン設計に応用した場合、状況に応じたペイロードや滞空時間、耐風性、通信距離など必要なパラメータを制約条件として入力することで、作戦に最適なドローン形状を AI がデザインできる。さらに、3D プリンタへ量産時間や残材料などの条件をかければ、人間が考察するより早く状況に適合し任務に最も適合したドローンを制作ができることになった[28]。ロボットアームの学習事例では、工場のロボットアームに新しい動作をプログラミングさせるには従来大変な労力がかかっていたが、人間が手動で一度入力し機械がその作業を覚え、仮想空間で実時間より早く何百万回もトライ&エラーを繰り返す強化学習[29]と呼ばれる手法により人より正確かつ迅速

---

[25] McAfee, ASCII.jp（2019 年 12 月 10 日アクセス）Bloghttps://ascii.jp/elem/000/001/403/1403066/

[26] NEC 技報、Vol.70（2017 年）、No.2（10 月）ビジネスの安全・安心を支えるサイバーセキュリティ特集 ～Futureproof Security 安心の先へ。～AI（人工知能）を活用した未知のサイバー攻撃対策（2019 年 12 月 10 日アクセス）
https://jpn.nec.com/techrep/journal/g17/n02/170215.html

[27] Daugherty and Wilson, HUMAN+MACHINE, Pages 184-188.

[28] 3D PRINT.COM," US Army and Marines Apply 3D Printing to Drone and Flight Applications", Dec. 19, 2017. （2019 年 12 月 10 日アクセス）
https://3dprint.com/197780/army-marines-3d-printing-drones/

[29] 強化学習： AI の学習の際、AI に目標を与え活動させ、その結果に対し報酬又は罰を与える。より報酬が与えられるよう繰り返すことで、より良い報酬を得た AI を作成することができる手法。

なロボットアームが実働している[30]。囲碁や将棋、チェスなどもルールと棋譜などを入力することで同様にAIが自ら何百万ものシミュレーションを行うことにより、勝ち筋を学んでいく。1章の3系の作戦を見積もるAIもこの分類に該当する。ルールが明白なゲームでは人間を超える成果を発揮している。一方、複雑な環境での学習のためには、仮想空間上でより精密な動作を再現するような仕組み及びこの環境での明確なルール（例えば視界、装備性能だけでなく住民避難や広報への影響を数値化）などを定義できなければ、戦闘のシミュレーションは難しく[31]、このような手法をそのまま戦術や戦闘予測に適用することは現段階では難しいと考えられる。

**カ　その他の分野**

上記分類に該当しないような群制御や自律ロボット分野などの他区分も多数存在しておりすべてを網羅することはできないため、ドローンなどの群制御については松野文俊の解説論文[32]、自律ロボットについては谷口忠大の著書[33]を確認して頂きたい。前者の論文は、アリなどの群で行動する昆虫を例に群制御ロボットの特性を解説し、2014年から数千台規模の群ロボットの研究が実施され、環境に応じた自立的な群れ構成や適用できる自己組織化群ロボットや系譜を丁寧に紹介している。後者は目的に応じた①から⑤の組み合わせであり、ロボットが自律で目的を達成するため経路探索、ゲーム理論、強化学習や自己位置推定、パターン認識や言語認識が丁寧に説明されており一通り必要となるデータと機械学習手法を学ぶことができる。なお、本記事で述べる手法やデータについては代表的な事例を示すのみであり、必ずしもその手法を適用できるわけではない。目的に合わせ企業担当者や専門家であるデータサイエンティストがデータと機械学習手法を決定する必要がある。手法の決定やAIの課題については、次章以降で細部を論じる。

---

30 三菱電機、AIを活用したロボットの力覚制御の高速化技術を開発、ニュースリリース、2017年11月（2019年12月10日アクセス）
http://www.mitsubishielectric.co.jp/news/2017/1121.html

31 I/O 編集部『ビッグデータ＆人工知能ガイドブック』、119-120項。

32 松野文俊、群行動の理解と群ロボット研究、日本ロボット学会誌, Vol. 35, No. 6, 2017、428-432項。

33 谷口忠大、人工知能概論、講談社、2014、15-199項。

## 4　AI の研究開発における課題

### (1)　本章の目的

前章までにおいて、AI の目的に応じたデータの必要性及び関連する学習手法について述べた。このように目的に特化して開発される AI を「弱い AI」と呼び、AI が自ら目的やデータの特徴設計が可能な汎用型の AI は「強い AI」と呼ばれている。シンギュラリティなど AI が人間を超えるような話は、後者が実用化された場合の影響を述べているが 2019 年現在では「強い AI」の基礎研究の段階であり、当面の実用化は難しい。このため現在は目的特化の「弱い AI」の開発が主となっており、前章までに紹介した事例も「弱い AI」に該当する。一方、2018 年の世界 6 か国の先進国において 500 社を調査した結果によると 98%の会社で AI による改善を実施しているものの、活用ができていると回答した会社はわずか 2%であった[34]。また AI 開発がスタートした企業のうち Google ベンチャーズの投資企業の 85%が開発失敗、AI　MIRAI（電通）でも AI が形になったものは 33%程度となっている[35]。このように前章までに紹介した AI 活用の成功例だけでなく失敗事例に着目して課題を検討する。また開発時の課題を明らかにすることによって自衛隊での AI 技術適用装備実現のための課題について述べる。

### (2)　AI の研究開発における課題

失敗事例として上述の記事では AI を用いる目的の誤りによるものから、目的に応じた適切なデータが存在しないような場合が述べられている。また、そもそも AI に学習させる上で「正解」となるようなラベル付きデータが存在しないようなもの、人が正しい判断かどうかを検証できないようなものなどがある。さらに、AI のアルゴリズムは成功したもののシステム全体を考慮しなかったために実装できなかったものまで多岐にわたる。ここで制作段階においても AI の基礎となる機械学習自体が失敗してしまう事例を述べる[36]。

#### ア　問題の設定項目の誤り

企業間の取引額が適切かを AI で判定する目的を設定したとする。人が監視して判断していたデータを用いた場合、人が以前から見逃していたことは判断できない。

---

34 Daugherty and Wilson, HUMAN+MACHINE, Page 154.

35 児玉拓也、『AI プロジェクトの 1/3 は失敗する。失敗例から導く AI 活用の勘所』人工知能特化型メディア 2019、（2019 年 12 月 10 日アクセス）https://ledge.AI/theAI-3rd-dentsu/

36 Alberto Artasanchez,"9 Reasons why your machine learning project will fAIl", KDnuggets, 2018.7（2019 年 12 月 10 日アクセス）
https://www.kdnuggets.com/2018/07/why-machine-learning-project-fAIl.html；
尾崎隆 機械学習プロジェクトが失敗する 9 つの理由 2018.8.3（2019 年 12 月 10 日アクセス）
https://tjo.hatenablog.com/entry/2018/08/03/080000#f-f5f08ebe
*参考資料は、9 つの点を述べているが、本記事の構成上著者が 8 つに整理した。また各事例については項目を除き著者が追記・編集した。

本来は、「人が判断したデータで AI 作成」ではなく、「取引額データそのものの値が異常値か正常値の範囲か判定する AI」とするような問題設定をしなければ、用いるべきデータすら誤ってしまう事例である。

### イ　間違った問題を解決するために AI を用いる

作成した AI に「価値があるのか」の根本的な問題である。例えば都内でゴミ出し量が予測できず収集車がオーバーフローすることを解決する例を考えると、ゴミステーションに IoT センサを付けて、画像からゴミ量を推定する AI と複数のゴミ収集車の回収残量と経路設定ができる AI を開発することは一見よい例だと思える。ただ、これらのシステム作成に 1 億かかり、そもそも収集車と作業員をピーク時期のみリースを活用したら 1000 万ですむような場合、費用対効果はマイナスになる。このような問題の解決方法は誤っているとされる。

### ウ　十分なデータがない・適切なデータがない

1 章や 2 章で解説した通り、目的に対応したデータがないような問題には適用は難しい。データサイエンス分野では”Garbage in, garbage out”と呼ばれ、ゴミデータからはゴミ AI しかできない。画像の学習において、ある画像=戦車のようにラベル付けを人がしたとしても、このラベル付けが 1 割間違っているデータを用いれば 99.9%精度の AI でも現実世界では 1 割が誤っていることに留意が必要である。さらに、AI のモデル構築の 90%がアルゴリズム自体の開発ではなくデータの準備と特徴量の加工や推定に用いられていることから適切なデータの質と量の有無は最も重要な課題であると考えられる[37]。

### エ　データが多すぎる

前項と対照的ではあるが、ビッグデータから目的に沿った学習を実施するためには、目的に適合した特徴あるデータを見つけ出しそれ以外のデータを削減する必要がある。データが多いほどその作業は煩雑になり、どのデータを削減するか判断する手法はあるものの、この判断は開発する人に委ねられている。

### オ　誤った人材配置

データ収集、データから特徴あるデータの推察、モデリングや実際の AI の動作だけでなく、通信インフラや実行環境のハードの制約解決者など本来は専門者を多く必要とする。最初は小さいチームでデータ収集・前処理や特徴量のエンジニアリング、モデリングやデプロイまで可能な少数精鋭を集め、組織が大きくなると全体設計や通信インフラ、運用解析などそれぞれのエキスパートが必要となる。ここに、必要とは異なるスキルの人の配置やスキルを保有しない人材を配置した場合、失敗

[37] Daugherty and Wilson, HUMAN+MACHINE, Page 241.

する。参照した記事では、各分野のエキスパートとこれらの人を束ねるデータサイエンティスト[38]がいれば良いと結論付けている。

**カ　間違ったツールの使用**

データから機械学習でAIを製作する際に、データ基盤ではMySQL、MongoDB、クラウドDWH等を選択する。ツールではPython、R、Jupyter、Keras、クラウド付属のツール等を使用していくが、これらのツールは時間とともにアップデートなどで使用できなくなる可能性がある。このため、それぞれの特性や実際の運用環境に適合したツールを選択しなければならない。

**キ　適切なモデルを用いていない**

データや目的からモデルが線形問題にもかかわらず非線形手法やディープラーニング、YOLO などを適用した場合、適切な出力にならない場合がある。AI の基礎となる機械学習のモデルも適切に選択しなければならない。

**ク　適切な評価方法を用いていない**

AIの学習段階で用いたデータをそのままテストに用いると100%精度の評価となるように、テストデータと評価方法を適切に設定しなければ AI の性能評価ができない。

### (3)　AI 技術適用装備の実現のための研究開発上の課題

これらの課題から、今後防衛省での AI 技術の適用装備を検討する上での研究開発上の課題として下記のように整理できる。

**ア　AI を用いる適切な目的の設定**

AI をそもそも用いるべき項目か。AI に適している問題か。

**イ　データの質量、データ構築可能性**

AI に学習させるデータの質・量は適切か。データの継続的に収集するデータベースの構築は可能か。

**ウ　AI モデル作成・システム全体の設計**

モデルは目的の結果を出力するのに本当に適切か。また、システム全体でその AI は適切な時期に適切な出力は可能か。

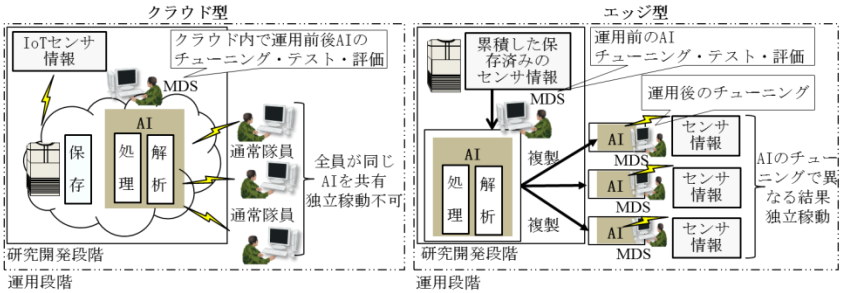
**エ　AI の評価方法の検討**

その AI は正しく設計され、求められる精度を出力できるか。

---

[38] データサイエンティスト（DS; Data Scientist）：データサイエンティスト協会では以下の3つのスキルをもった人材と定義している。 1. ビジネス力：課題背景を理解した上で、ビジネス 課題を整理し、解決する力 2. データサイエンス 力：情報処理、人工知能、統計学などの情報 科学系の知恵を理解し、使う力 3. データエンジニアリング 力：データサイエンスを意味のある形に使えるようにし、実装、運用できる能力をもつ。（2019年12月10日アクセス）http://www.datascientist.or.jp/files/news/2014-12-10.pdf

図３　AI 開発から運用のイメージ図




### オ　バイアスの有無、信頼性確認要領の検討

特定のデータやモデルで、結果に偏りが生じることはないか。セーフティとなるプログラムはシステムに組まれ、運用時に案出プロセスを可視化できるような形で設計されているか。

## 5　AI の運用における課題

### (1)　本章の目的

前章は開発時の課題について言及した。一方で、失敗事例の中では AI を開発したものの、運用時に失敗する事例もある。この課題を検討するにあたり、まず自衛隊での開発と運用イメージを述べ、現在日本で問題となっている AI の社会への適応上の課題及び倫理上的な課題の細部を明らかにすることにより、自衛隊での AI 技術適用装備の運用課題について明らかにする。

### (2)　AI 開発と運用のための過程

図 3 に AI 開発から運用のイメージ図について示す。自衛隊においての活用時には大きく 2 通りの状況での運用が考えられ、センサなどのデータ取得のためのセンサがネットワークで繋がったクラウド型と、ネットワークから切り離されても稼働するエッジ型である。前者はリアルタイムに様々なデータを取得し、クラウド上で保存されたデータを同一クラウド内の AI により解析し、結果を端末へ提示する。メリットは運用者のニーズに合わせた随時のアップデートが可能であり、適切なチューニングができれば、どの端末からでも精度向上した AI を用いることができる。一方でデメリットとして各端末はクラウドに接続しない限り AI の結果を受け取ることができない。後者は、一度データから AI を作成しテストした後は独立して運用可能であり、光学画像や熱画像、レーダなどのソフトウエアの一部に搭載し、これらのセンサ情報から敵を早期に発見できる。メリットは、クラウド接続の必要がなく電磁的・サイバー的に抗堪性があり、開発時に熟練した隊員以上の精度を達成

していれば全端末で同様の性能が発揮できる。デメリットとして、現場の特徴に合わせチューニングを繰り返した場合は、他の環境（北海道から沖縄へ移動など）で動作保証ができない可能性があるため、不用意に現場でのアップデートができなくなる。このような運用環境で生じ得る問題として、センサから出力された内容が本当に正しいものなのか、AI の提案したものに対してどのように責任を取るのかが問題となる。

### (3) AI の社会への適応や倫理的課題

AI 開発が終了し運用段階に進むと、AI にどこまで判断させるのか倫理的な課題や AI の判断過程がわからず何故その判断となったのか人間に理解できない問題が累積している。日本は、AI の社会への適応を前提として AI 戦略[39]を策定し、人材育成、社会への適応、データ戦略など幅広い戦略を策定し、AI 社会原則を以下のように定義した。

1 人間中心:AI は人間の能力拡張、判断は人
2 教育リテラシー:教育機会拡充
3 公正競争確保
4 公平性、説明責任及び透明性
5 プライバシー確保
6 セキュリティ確保
7 イノベーション

防衛省に関連した研究開発・運用等の記載はなく、社会への適応上の共通の原則が定義されており、人間が AI の結果の判断、AI 人材とデータの占有におる公正競争、AI の結果を説明できる説明責任や透明性と基盤となる個人情報の管理や保護、開発原則を掲げている。また、米ペンタゴンでは軍事利用を前提とした AI 運用時の倫理目標として以下の 5 つの原則[40]が必要としている。

1 責任：人間が AI の適切なレベルの判断を行い、この責任を負う。
2 公平性：バイアスがなく公平性を担保できる。
3 説明性：提案結果の案出プロセスを確認できる。
4 信頼性：明確なテストにより定義されたもので信頼性を担保できる。
5 自己回避：意図しない危害や混乱を検出してソフトウエア自身で気付き、自身で停止できる。

---

39 AI 戦略 2019～人・産業・地域・政府全てに AI～（内閣府　統合イノベーション会議決定）2019 年 6 月 11 日

40 Patrick Tucker, "The Pentagon’s AI Ethics Draft Is Actually Pretty Good", Defense One, Oct. 31, 2019. (2019 年 12 月 10 日アクセス)
https://www.defenseone.com/technology/2019/10/pentagons-AI-ethics-draft-actually-pretty-good/161005/?oref=defense_one_breaking_nl

図４　ディープラーニングでの誤判断事例[41]


1 は AI でさせるべき事項の決定と問題設定、2 はデータ特性、4 は評価方法、5 はプログラム上の組み込みの話であると考えられる。3 は、AI の判断を人間が理解することが難しいことを述べている。例えば画像・レーダ系で最も精度が良いとされるディープラーニングは脳の神経を模したニューラルネットを数式で表現しており、画像を画素単位や様々な方法で分解して次の式に渡す仕組みである。分解された単位をいくつもの階層を得て、出力層と呼ばれる画像が何であるのか判定している。この階層でのやり取りは人が見て理解できず、図 4 の戦車の画像に少しノイズが加わっただけで誤った出力となる。同様に止まれ標識に少しのモザイクで結果が変わってしまっても理由がわからない。

このため、このような階層を人間が理解できる形に説明する EAI(Explainable Artificial Intelligence)[42]などの研究も実施されている。このように、目的設定から AI の基礎の作成での課題に加え、社会原則や軍事利用への原則にあるよう、クリアすべき多くの課題がある。このため AI は、とにかくデータを貯めただけのビッグデータがあれば何かできるような物ではなく、従来の装備品と同様に技術的課題（特にデータ依存的）が多く存在している。さらに従来と異なりデータ依存のため装備の評価方法や AI 案出結果のプロセスが確認できないなど万能なものではない。適切な目的、問題設定やシステム全体の設計、学習データから運用時のデータ入出力、モデルの検証やシステムの評価方法など AI の仕組みや特性を把握し、自衛隊の運用要点を理解して、AI 活用を検討・実現できる組織や人材がいなければ十分な活用は難しいと思われる。

このため、AI の目的によってはデータや評価を装備化後の運用段階で稼働させながら AI のチューニングをしつつ実施するような開発体制が必要となると考えられ

---

41　Ian Goodfellow, Nicolas Papernot, Sandy Huang, Rocky Duan, Pieter Abbeel&Jack Clark,”　Attacking Machine Learning with Adversarial Examples”,Feb. 24,2017.を参考に著者が作成。左図の stop 表記は Evan Ackerman,”　Slight Street Sign Modifications Can Completely Fool Machine Learning Algorithms”,Aug. 4,2017.を引用。

42　"ExplAInable Artificial Intelligence", Program Information, DARAPA；AI 白書、346 項。XAI は DARAPA の研究プログラムの中核として説明能力を持つ AI の実現を目指すものとなっている。日本の AI 白書にも社会への適応のために同様の機能をもつ AI の必要性が謳われている。

る。この際、実際に使用する運用者と開発担当者、企業でチームを組み精度向上させていく仕組みが必要となると考えられる。ただし、運用段階において演習場でチューニングする際に取得されたデータが本来の目的（演習場以外の場所での使用）で使用可能であるかについては高度な専門的な見地が必要であると考える。

### (4) AI 技術適用装備の実現のための運用上の課題

これらの課題から、今後防衛省での AI 技術の適用装備を検討する上での運用上の課題として下記のように整理できる。

#### ア AI の判断に対する人間の責任範囲の明確化

人間のチェック機能は判断プロセスに組み込まれ運用されているか、AI の判断結果に組織は責任を負えるような運用になっているか。

#### イ AI 提案結果の案出プロセスの確認手法の確立

AI の案出プロセスの可視化を実施し、運用の現場で確認が可能か。

#### ウ 運用時のチューニング要領と評価方法の確立

運用時にモデルの修正やパラメータのチューニングを実施する際に、運用者自身が出力の妥当性を検証できるような評価法があるか。

## 6 防衛省としての AI 研究開発・運用体制の考察

### (1) 本章の目的

前章の課題に対する AI 研究開発・運用体制の方向性として、AI 開発と運用の過程を示し、人材、データ構築及び開発・運用体制から今後の方向性を検討する。まず、AI の基盤となる機械学習のためデータの有無や整備が必要なことから、これらを支える基盤となる人材、データ構築及び研究開発・運用体制について考察する（表 2）

表 2 課題と AI 研究開発・運用体制の関係

| | 課題 ＼ 研究開発・運用体制 | 人材 | | | データ構築 | 研究開発体制 |
|---|---|---|---|---|---|---|
| | | 運用系 | MDS* | 企業 | | |
| AI開発時 | 1 AIを用いる適切な目的の設定 | ○ | ○ | | | |
| | 2 データの質量、データ構築可能性 | | ○ | ○ | ○ | ○ |
| | 3 AIモデル作成・システム全体の設計 | | ○ | ○ | | ○ |
| | 4 AIの評価方法の検討 | ○ | ○ | ○ | | ○ |
| | 5 バイアスの有無、信頼性確認要領の検討 | | ○ | ○ | ○ | |
| AI運用時 | 1 AIの判断に対する人間の責任範囲の明確化 | ○ | ○ | | | ○ |
| | 2 AI提案結果の案出プロセスの確認手法の確立 | | ○ | ○ | | ○ |
| | 3 運用時のチューニング要領と評価方法の確立 | ○ | ○ | ○ | | ○ |

MDS: Military Data Scientist

### (2) 人材育成（Military Data Scientist の必要性）

このような開発・運用段階で必要な人材は、研究・開発時から運用まで一貫して AI に携わる必要がある。各課題の対策において必要な人材は、「運用系（指揮官・幕僚）」、「技術系（Military Data Scientist：MDS）」及び「企業」に区分できる。まず、運用系は、最も大切な AI に何をさせるかを運用面、技術面及び倫理面で検討できる防衛省側の人材であり、前章までの目的とデータの関係性を理解し、目的を幅広く提案する。次に、同じく防衛省側の人材である MDS は、データサイエンティストの資格を保有し、運用系が提案した目的と AI の関係性や、提案した目的の重要度、実行の可能度等を検討する。彼らは、上級クラスになるほどデータ構築体制への技術課題、コスト問題、統合運用等の高度な視点を保有しなければならない。一方、初級から中堅クラスは、運用時のチューニングや AI の案出プロセスの可視化などに特化した能力を保有し、開発全般への意見具申等を求められる。そして企業は、これらの防衛省の MDS と共にシステム全体の設計や AI 評価方法、信頼性の具現化について具体的な検討を協同して実施し、最新のハードとソフトの組み合わせを活用し、運用時に最適な出力ができるように設計を行っていく。企業とともに作成に携わった MDS は、そのプロジェクトの運用時の課題対策の第一人者として、省内への様々な意見や評価の提出を実施しつつ企業からチューニングや課題解決の支援をうけ戦力化していくようなループが必要となる。3 章 1 で述べたように、AI の開発・活用の 7 割近くが失敗に終わる原因の多くは、AI の活用を試みる組織自体に専門家が不在しており、依頼した企業に自社の改善すべきような側面（データ構築要素等）を委託し、本来 AI にさせたい目的や根本的な運用に適合できないことからも、MDS の役割は重要であると言える。

一方で運用者が AI への判断と責任を許容するためには、機械（AI）の作成した物に対する不審的な心情も克服する必要がある。これは小野が「意思決定のように知的労働の根幹部分を AI に依存することに対する人間の心理的な抵抗は否定できず、軍への AI 導入の障害となる」[43]と問題提起をしているように人間は、例え実際には AI による提案が人間よりも優れていたものと理解していても、人間の提案を好む傾向がある[44]。この不審的な心情の克服には、運用者が AI を実際に使用して AI のプロセス・チェック機能が十分に活用でき、出力結果を納得するまで時間がかかると思われる[45]。

---

[43] 小野、人工知能（AI）による軍の知的労働の代替、20-21 項。
[44] Daugherty and Wilson, HUMAN＋MACHINE, Pages 230-231.
[45] 藤本浩司、柴原一友『AI にできること、できないこと』、223-245 項。

### (3) データ構築

「十分なデータがない・適切なデータがない」で解説したとおり、AI のモデル構築の 90%がアルゴリズム自体の開発ではなくデータの準備と特徴量の加工や推定に用いられていることから、目的に対応したデータを保有またはデータ構築ができるような環境になければ、AI の基礎となる機械学習の適応は不可能である。防衛省で様々なデータを蓄積する取り組みが進んでいるが、すべてのデータが機械学習で使用できる構造[46]とはなっていない。データ蓄積の際には、長期的なデータ活用のヴィジョンを基に AI 開発に必要な質と量を備えた収集体制が必要である。例えば、人事活用であれば現在使用している文章化または数値化されていないようなデータの整備と蓄積要領の検討が必要である。また、作戦行動の提案を目指すのであれば FTC(Fuji Training Center)など[47]で収集できるデータに加え、指揮所訓練統裁支援システムで実行動と同様の質で、データの累積と分析方法の検討が必要になると考えられる。AI の目的に応じた分析方法をある程度具体化していかなければ、無用なデータを収集し必要なデータが無い状況が生起するため、蓄積と平行した分析作業で「何」のデータをどのような「質」で累積していくのかを決定する長期的なデータ活用ヴィジョンは必須であると考える。

また、自然災害など緊急性を用いる情報と関係のない過去の活動のデータのクラウド保存は区別する必要がある。IT の専門家達は緊急性の高い有用なデータは「ホット」なデータとして優先的な地位を与え即座に使えるパフォーマンスの高いシステムに置き、そうでないデータを「コールド」とすることで高性能サーバーとそれ以外のサーバーに保管する仕組みを組み込み活用できるようにすることがシステム全体設計に必要となる[48]。

### (4) 体　制

3 章での課題に対し、防衛省として官民でどのように長期的に取り組んでいくべきかのロードマップを構築できるような政策チームと、今ある課題で AI が適用可能な画像・レーダ系の分野に少数精鋭の防衛省のデータサイエンティスト（将来の MDS）と企業の開発チームを立ち上げ、課題である目的からデータ取得、分析、モデル化、評価及び運用時の評価・チューニングを一貫したチームで実施して成功例を蓄積していくことが最も重要である。実際に AI 技術適用装備の成功例までに至

---

46 構造・非構造データ：構造データはエクセル数値表のような決まった項目に決まった値が収められている状態だが、非構造データは SNS の文字をはじめ、画像や表記項目が一律でないデータをいい、データ分析では変換を必要として扱いづらいデータとされる。

47 FTC：陸上自衛隊の部隊訓練評価隊であり、陸上自衛隊の部隊の練度を模擬的な実践環境下で人員や車両へのセンサ、自己位置及び火力発揮の状況など様々なデータから定量的に評価し、部隊の精強化を図っている。

48 藤本浩司、柴原一友『AI にできること、できないこと』、224-245 項。

る開発制度上の問題点や AI の判断に対する法令[49]に関する改善が必要な事項を政策チームへ提言し、今後の AI 開発の人材、データ構築、研究開発体制及び倫理的課題解決の方向性を含め防衛省としての AI 技術適用装備の研究・開発・運用体制を明確にしていく必要がある。このような基準がなければ、66%の AI 活用が失敗している事例のように、従来装備と異なる特性をもつ AI 技術適用装備に対し、従来の技術試験や運用試験の仕組みが合わないことから運用時のチューニングや案出プロセスの明確化、評価判定など規定できず装備化ができない可能性もある。

さらに AI 技術適用装備の開発・運用で重要となるのが、AI 技術適用装備の提案結果に対する人間の責任や判断に対する説明をどのように組み込むのかを明確にすることである。AI の提案した事項に対し採り得る対策として 2 つの事項が考えられる[50]。

1　判断プロセス内に人間のチェック機能を付加

AI がミスしてもカバーできるように人間がチェックできるプロセスミスしても人間が責任を負うことのできる判断プロセスを構成

2　間違いの許容

人間では処理できないくらい作業量が多く、リスクが許容できるような場合、運用者が間違いを織り込んで使用

このような対策を踏まえた上で、現在の運用上の課題や長期的な技術進展から考察し得る将来の AI 装備をプロジェクト設計指針に定め、各プロジェクトの目的と課題を明確に決めることが重要である。この際、共通する事項をまとめ、対象とするプロジェクト AI 適用範囲を限定できると、検討すべき選択肢も限定的になり AI が性能を発揮しやすくなると考えられる。この際、DARAPA の例で紹介した XAI のような基礎研究や機械学習を進展させる可能性のあるエッジ AI や量子コンピュータ、AI の自己回避などに加え、データ収集に用いるセンシング分野の研究に対し、安全保障技術研究推進制度を活用した研究投資は必須であると考えられる。

さらに、データ整備や研究開発体制が整い、AI 技術適用装備の本格運用が始まった後、各目的別で機能発揮する AI 技術適用装備がそれぞれクラウドやエッジで状況判断を支援している状態となる。ここでは、システムとしてサイバー攻撃や電磁攻撃の抗堪性だけでなく、図 4 のようなディープラーニングの弱点を逆手にとったカウンターAI 攻撃のような対策として人間のチェック機能のあり方及びシステムの自己停止プログラムなど運用しつつアップデートさせていく必要があると考えられる。さらに、防衛省のみならず沿岸監視や警戒監視に必要な衛星や他省庁との

[49] 福岡真之介（編集）、『AI の法律と論点』、商事法務、2018 年、150-163 項。
[50]藤本浩司、柴原一友『AI にできること、できないこと』、162 項。

大量の情報のやり取りなどにもアップデートで必要となる機能だと考えられる。この際、AI 技術適用装備に提案結果や情報に対し、数値的や内部の動作など客観的な動作から、AI の判断に異を唱えられるような人材と運用体制になれなければ AI 装備を活用することができないと考えられる。

## おわりに

本論文では AI の民間事例から、データと AI の関係を解説し特性を明らかにするとともに、AI の開発時及び運用時の課題を分析することで防衛省としての AI 研究開発及び運用体制の方向性について論じた。防衛省や関連企業との AI 技術適用装備に関する議論はまだ端緒についたばかりである。そのため、AI の現状や課題について知見のないまたは誤った知識を有した者が携わった結果、目的に適していない価値のない AI 開発や体制の不備による不適切な運用により国民の信頼を揺るがすような事態は絶対にあってはならない。一方、今後の少子高齢化で隊員の確保がより難しくなる将来において、AI 技術の適用装備により様々な領域監視の一部を補佐していかなければ常続的な安全の確保は厳しくなっていく。このような現状に対し、本論文で語るべき内容が不足しており、多くの批判やご意見を頂くこととなると推察するが、本論文の否定や修正をもって AI 活用の課題や研究開発・運用体制の議論がより活発になることを期待する。