

〈1〉 フィジカル AI とは何か？ —その安全保障への含意—

防衛省 防衛研究所 理論研究部 研究員 清岡 克吉

1. はじめに
2. フィジカル AI とは何か？
3. フィジカル AI が変えるもの—安全保障の観点から
4. フィジカル AI の安全保障分野での課題
5. おわりに

1. はじめに

今日、人工知能（AI）は、日々のニュースで耳にしない日はなく、先進的な現代社会を構成する不可欠な要素の一つとなっているといえよう。AIは、民生分野に限らず、安全保障分野においてもその利用についても様々な議論や検討、実証がなされており¹、その技術は民生と軍事の両用性を持っているとも言える。そのようななかで、2025年1月にNVIDIA社の社長兼最高経営責任者を務めるジェンソン・ファン（Jen-Hsun “Jensen” Huang）が、全米民生技術協会が主催する電子製品見本市の「コン

シューマー・エレクトロニクス・ショー（CES）2025」での基調講演で「フィジカル AI（Physical AI）」という概念を提唱した²。フィジカル AI は、後述するように、同氏が提唱した AI の世代分けの議論の中で紹介された、次なる AI の世代に関する概念である。フィジカル AI は、誤解を恐れず、極めて単純に言うならば、現在の「言葉を生成する AI」からさらに移行した「行動を実行する AI」のことを指す。本稿では、このフィジカル AI とは何かについて概観したのち、この概念に基づく AI の安全保障分野への実行がどのような含意を持つのかを検討する。

2. フィジカル AI とは何か？

2.1. 概要と構成要素

フィジカル AI とは、先述の通り、NVIDIA のジェンソン・ファンが提唱した AI の世代を分類し、「次のトレンドとなる AI」の世代にレッテルを貼ったものであるといえる³。フィジカル AI とは、ファンの言

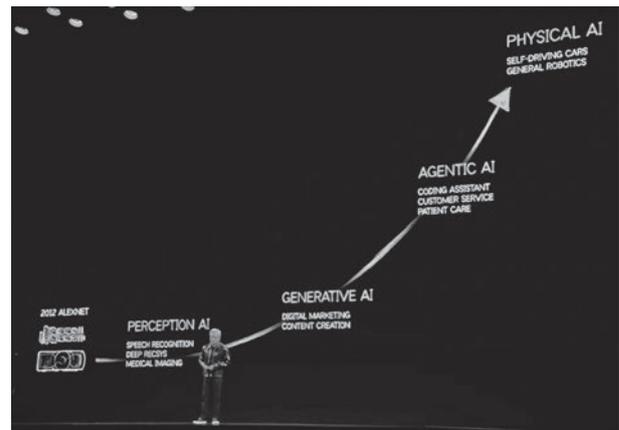
¹ Horowitz, M. C., Artificial Intelligence and International Security, Centre for New American Security, July 2018.; Congressional Research Service, Artificial Intelligence and National Security, August 2020. 等を参照。

² ブライアン・コールフィールド「CES 2025 : AI は「驚異的なペース」で進歩していると NVIDIA の CEO が語る」NVIDIA、2025年1月16日（最終閲覧日 2025年12月8日：<https://blogs.nvidia.co.jp/blog/ces-2025-jensen-huang/>）。

³ ファンが定義した「フィジカル AI」に相当する AI やその概念、技術に関する研究は同氏が「フィジカル AI」と名付けるよりも前から存在し、「身体性ある知能（physical intelligence）」や「身体性 AI（embodied AI）」などの呼称がなされていた。詳しくは Salehi, V., Fundamentals of Physical AI, Institute of Engineering Design of Mechatronic Systems, University of California, Berkeley, 2025. 等を参照。

葉を要約すれば、「現実世界における物理法則、環境との相互作用を学習し、理解し、自律的な出力として行動するAI」というような定義になる⁴。すなわち、手を動かす、体を持ったAI、もしくは生成AIとロボティクスの融合とでもいえるものになる。フィジカルAIは、この過程で、自然言語処理を行う生成AIで世界を認知し推論を行うため、人間が与えた指示に基づいて柔軟な行動を行うことができる。ロボットや自動運転車などが最も身近な事例であろう。このように物理的な出力が可能なAIは、製造業をはじめとする多様な産業用途に大きな変革をもたらす。その影響の大きさとして、2033年までにフィジカルAIの世界市場規模は約497億ドル（年平均成長率32.53%）に達するとの試算もある⁵。

ファンが提示したAIの世代分類によれば、AIはまず画像や音声を認識する「パーセプションAI」として発展し、その後、文章や画像、コードなどを生成する「生成AI」へと進化した。近年ではさらに、目標を理解し、計画を立て、複数の行動を組み合わせて実行する「エージェントAI」が登場している。フィジカルAIは、この流れの延長線上に位置づけられ、AIが推論や意思決定にとどまらず、物理世界において直接的な行動主体となる段階を示している。従来の生成AIが主にテキストや画像といった抽象的な情報を扱ってきたのに対し、フィジカルAIは、現実空間に存在する物体や環境そのものを対象とし、そこに働きかける能力を持つ点で質的に異なる。また、フィジカルAIは、NVIDIA社が名付けた概念であり、CES2025での講演では、同社のH100などの製品を使ったソリューションが多く含まれたが、その概念の汎用性と適応範囲は、より一般化が可能なものである。



(写真)CES2025において「フィジカルAI」について講演するジェイスン・ファン(出所:Omdia)

上述の通り、フィジカルAIが扱う入力情報は、自然言語だけではない。カメラ映像、LiDAR、レーダー、温度センサー、加速度センサー、触覚センサーなど、現実世界を構成する多様な物理情報が統合的に処理される。フィジカルAIは、これらの異なる種類の情報を同時に理解し、現在の環境がどのような状態にあるのかをリアルタイムで把握する。その上で、状況に応じて適切な行動を選択し、ロボットの腕を動かす、自律車両を走行させる、あるいは機械を制御するといった物理的な出力を行う。つまり、フィジカルAIとは「認識・推論・行動」が分断されていた従来のAIシステムを統合し、知覚から行動までを一貫したプロセスとして扱うAIであると整理できる。

フィジカルAIが注目されるようになった背景には、AI技術の発展はもちろんのこと、主として先進国での労働人口の減少に伴う課題の解決という、「実世界での触覚的・物理的な仕事をこなす能力」が本質的に求められているという社会的・産業的ニーズの変化がある。例えば、製造現場のロボットが荷物を搬送したり、組立作業を行ったりするためには、環境の3次元的な構造を理解し、物体の位置・形状・動きを把握して最適な運動制御を実行する必要がある。また、現在普及しつつある自動運転車の実装には、歩行者や障害物を認識し、交通の状況や天候の変化に応じて安全な走行計画を立案しなければなら

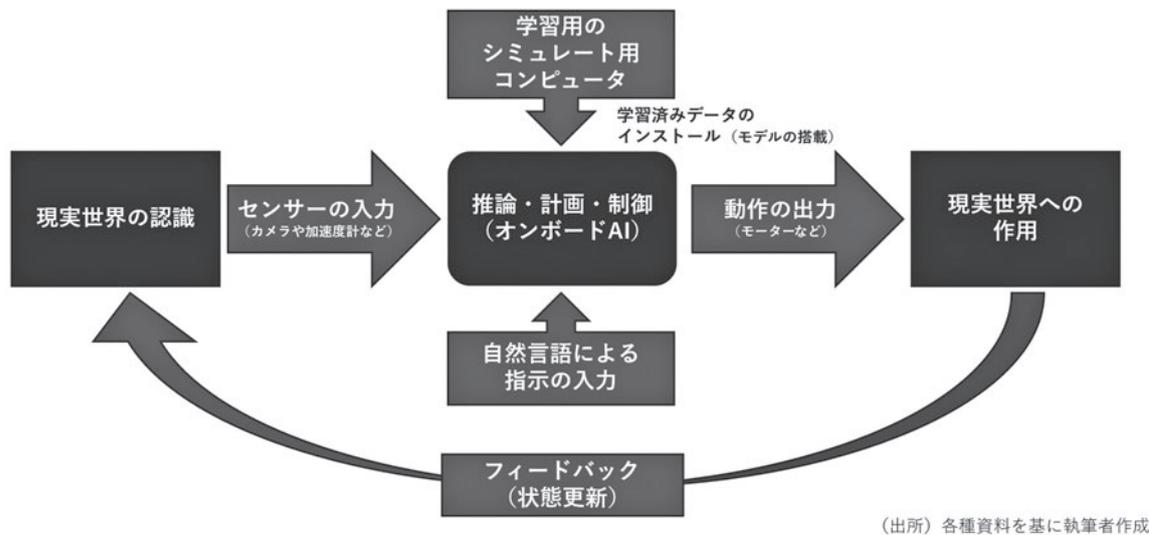
⁴ Huang, J., NVIDIA CEO Jensen Huang Keynote at CES 2025, NVIDIA, January 16, 2025, (<https://www.nvidia.com/en-us/on-demand/session/ces25-keynote/>).

⁵ S&S Insider pvt ltd, Physical AI Market to Hit USD 49.73 Billion by 2033, Owing to Advances in Robotics and Real-World Intelligent Systems, *Yahoo Finance*, December 5, 2025.

ないといった極めて複雑な状況が現実世界に存在する。このような実世界の不確実性を処理するには、単に過去のデータを分析してもっとも起こりうるシナリオを推定するだけでなく、今日の目の前にある

現実の環境を理解し、変化する条件にリアルタイムで適応する能力が必要であり、フィジカル AI はそのための認知と行動の統合的な知能を実現するものであると整理できる。

フィジカルAIの基本構造



(出所) 各種資料を基に執筆者作成

【図1】 フィジカル AI の基本構造 (出所：各種資料を基に執筆者作成)

上述の点で、フィジカル AI は従来の産業用自動化やロボット技術とも異なる。従来の産業ロボットは、あらかじめ決められた環境と作業手順を前提に、「if-then 型 (=条件指定型)⁶」の制御ロジックや事前プログラムに基づいて動作してきた。その強みは高い再現性と精度にあったが、製造ラインの組み換えや製造する製品の仕様変更などの環境変化や、想定外の事象への素早い適応力は限定的であった。これに対し、フィジカル AI は、経験から学習しながら振る舞いを更新する点に本質的な違いがある。フィジカル AI を用いたロボットでは、状況を認識して、事前に学習された大量のデータから自律的に適切な行動を出力するため、急な仕様変更があっても、再プログラムといった従来の産業用機械に必要な手順を省き、より汎用的な作業が可能になる。要するに、プログラム設計段階で想定できなかった部分を、物理演算と大量の学習からくる演繹的なデータ活用で補完するという理解ができるであろう。

フィジカル AI を技術的に成立させている要素は、大きく三つに整理できる。第一は、マルチモーダルな知覚と推論である。視覚、言語、音声、触覚、位置情報など、性質の異なる情報を統合し、環境の状態を総合的に理解する能力が不可欠となる。近年では、視覚と言語、行動を結びつける「Vision-Language-Action (VLA) モデル」や、世界の状態変化そのものを予測する「世界基盤モデル (World Foundation Models)」といった研究が進展しており、フィジカル AI の理論的基盤を形成している。

第二に、学習と検証の方法論の転換が挙げられる。物理世界における学習は、現実空間のみで行うにはコストや安全性の制約が大きい。この課題に対し、「デジタルツイン」などの現実世界を高い忠実度で再現した物理シミュレーションを用いて仮想環境で学習・検証を行い、その成果を実世界へ移行する「Sim-to-Real」アプローチが中核となっている。デジタルツインとは、工場、都市、車両、ロボット、あるいは

⁶ 例えば、ライン上に部品が流れてきたら掴む、5秒に一度部品を叩くなどの事前プログラミングされた動作。