

現実世界を理解するAI ～AI開発を加速させる「世界モデル」の動向～

株式会社日本総合研究所 先端技術ラボ

2026年6月15日

[お問い合わせ]

執筆者：[山銅 康弘](#)（日本総合研究所 先端技術ラボ）

本レポートに関するお問い合わせにつきましては、当社ホームページの [お問い合わせフォーム](#) よりご連絡ください。

- 本資料は作成日時点で弊社が一般に信頼できると思われる資料に基づいて作成されたものですが、情報の正確性・完全性を保証するものではありません。本資料の内容は、経済情勢などの変化により変更されることがあります。本資料の情報に起因して閲覧者及び第三者に損害が生じた場合も、執筆者、取材先及び弊社は一切責任を負いかねます。
- 本資料の著作権は株式会社日本総合研究所に帰属します。本資料の一部または全部を、電子的または機械的手段を問わず、無断で複製または転送などを行うことを禁止しています。

■ 世界モデルとは何か

- 世界モデルとは「物理的な因果関係や法則といった現実世界を理解・予測するAI」と表現できるが、明確な定義はない状況。
- 世界モデルは現行のAIが抱える物理的および時空間的な一貫性の欠如という課題を乗り越えるものとして有望視されている。

■ 世界モデルの活用

- 世界モデルはフィジカルAIをはじめとして応用先が幅広い。
- フィジカルAIにおける活用方法としては、(1) AIモデルそのものに世界モデルを組み込む、(2) AIの学習・評価環境として利用することが考えられる。これにより、AIの精度向上のほか、学習・評価のスピードアップ、コストダウンが期待される。

■ 世界モデルの研究・開発例

- 世界モデルの初期の研究としてはハ&シュミットフーバーの「World Models」と、ルカンの「JEPAモデル」が有名。
- 日本においては自動運転やロボット領域での研究・開発があり、実用化に向けた取り組みも行われている。

■ 世界モデルの今後

- 世界モデルを活用した高品質な動画生成は、AIが学習で使用するデータが不足している問題の打開策にもなるため、多くのAI研究・開発を加速させうる。
- 自動運転ではAIが認識・判断・制御までを実施するアプローチであるE2E方式においてスタンダードになる可能性がある。しかしE2E方式の覇権獲得や、実用化の直結を意味しているわけではない。
- 世界モデルの定義は未確定で曖昧なため、技術評価や社会的インパクトの評価に際しては使われる文脈について考慮する必要がある。
- 技術的な課題としては長期的な一貫性の取得や、物理的な因果関係や法則を理解していることを示す指標・ベンチマークの整備が必要。

- 現行の生成AIは現実世界の物理的な因果関係や法則を正しく理解していないという課題がある。その課題を乗り越える技術として「**世界モデル**」が注目されており、動向をレポートとしてまとめる。
- **世界モデルには明確な定義があるわけではない**。何か特定の技術やアーキテクチャを指し示しているとは限らず、現実世界を正しく理解、および予測するための実現方法は未確定である。
- この状況を踏まえたうえで、世界モデルと一般に考えられる研究、サービス事例を整理したのち、今後の展望について触れる。
- **本レポートの目的は、概念が曖昧であるがゆえにバズワードとして広がりうる世界モデルという概念について、今後の社会的インパクトや技術評価を行う際の、客観的な視座を提供することである**。
- 本レポートの想定読者は、「世界モデル」について知りたい人のみならず、AI開発・研究を専門としていない一般読者である。前提知識としては、AIに関する用語（LLM、フィジカルAI、汎用人工知能など）についてはイメージを持っていることが望ましい。

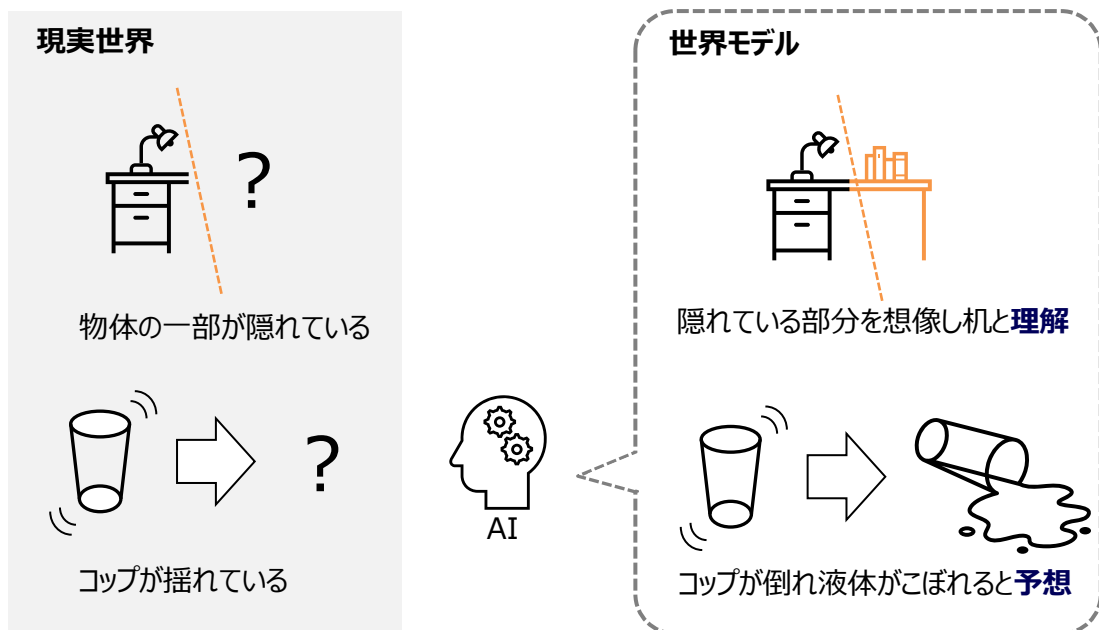
#	目次	頁
1	はじめに	3
2	世界モデルとは何か	4
3	なぜ世界モデルが注目されるのか	5
4	フィジカルAIへの活用	6
5	世界モデルをめぐる動向	7
6	世界モデルの初期の研究	8
7	世界モデルの広がり	9
8	世界モデルのサービス事例（海外）	10
9	世界モデルのサービス事例（日本）	12
10	考察・今後の課題	14
11	まとめ	15
12	参考文献	16

世界モデルとは

- 本レポートでは世界モデルを「**物理的な因果関係や法則といった現実世界を理解・予測するAI**」と定義する。
- 世界モデルの定義に関しては確定したものは現状存在しない。そのためある研究やサービスが果たして世界モデルなのか否かという議論が起こることがある。
- 世界モデルの今後の社会的インパクトや技術評価を行う際には、どのような意味で世界モデルという言葉が使われているのかを注意する必要がある。

世界モデルとは*1-2

- 「世界モデル」には確定した定義は存在しない。本レポートでは「物理的な因果関係や法則といった現実世界を理解・予測するAI」と定義する注。
- 世界モデルを備えたAIは、断片的な情報から想像（シミュレーション）し、人が行うような常識的な理解・予測を行う。



注：Nvidiaの定義*1を参考に日本総研作成。この定義は一般に世界モデルと呼ばれるものを含む包括的な定義となっている。しかしそれゆえに論者によっては世界モデルとみなされないものも含まれるが、まずは世界モデルの全体像（外延）を把握するためにも、この定義を採用する。

世界モデルの定義に関する議論

- OpenAIが開発した動画生成サービス「Sora」は、「**世界シミュレーター**」と位置づけられ、物理世界の人や動物、環境をシミュレートできると発表された。ただし内部的にはこれまでの生成AIと同様、過去の膨大な動画データから、統計的なパターンを再現し、ピクセルレベルでの動画生成を行っている。AI研究者の中では、これをもってAIが世界を理解していることにはならない、などの批判があり、**世界モデルとは結局のところ何なのか**という議論がある。



OpenAIが2024年に発表した動画生成AI「Sora」

ヤン・ルカン氏のXの投稿（2024/2/17のXの投稿）*3
「[Soraなど]プロンプトから現実のように見える動画を生成しているだけではシステムが物理世界を理解していることにはならない」



論者により世界モデルの定義は異なるため、あるサービスや研究が世界モデルなのか否か、という議論が起こる。

[画像出典]OpenAI, 2024/2/15, “Video generation models as world simulators”
(<https://openai.com/index/video-generation-models-as-world-simulators/>) 2026/4/6参照。

なぜ世界モデルが注目されるのか

- 世界モデルは、現行の生成AIが抱える「**物理的一貫性**」や「**時空間的一貫性**」が不十分という課題を乗り越えるものとして有望視されている。
- 特に自動運転やロボット等の“**フィジカル**”AIをはじめとして応用先が幅広いことから、世界モデルへの注目度が上がっている。

現行の生成AIの課題

- 現行の生成AIは、主にデータ間の統計的なパターンを学習し、それに基づきもっともらしい出力を生成する。そのためどこまで物理的な関係性を理解しているかは議論がある。
- 物理法則の理解や、状況に応じた対象への注意の制御が不十分となり、生成物において**時間的一貫性**や**物理的一貫性**が損なわれるという課題を抱えている。

物理的一貫性の損失



画像とテキスト「コップが倒れて割れて液体が出る」をもとに動画を生成

コップが割れずに中の液体が飛び出す
→AIが物理法則を理解していない

時空間的一貫性の損失

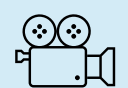
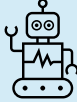
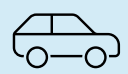





動画内である地点を描写し、視点が移動

視点が移動したときに物体（背後の人）が変容・消失している

世界モデルの応用先

- 世界モデルの利用は一貫性の課題を乗り越えるアプローチとして期待されており、応用先は幅広い。

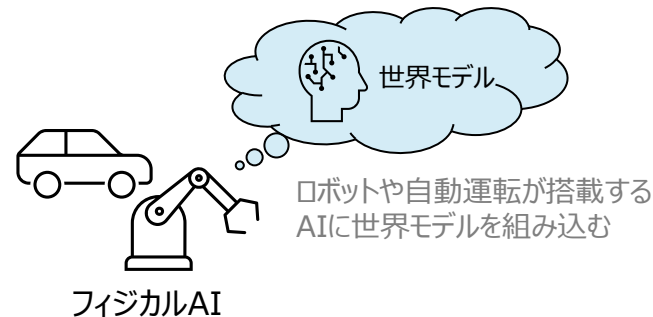
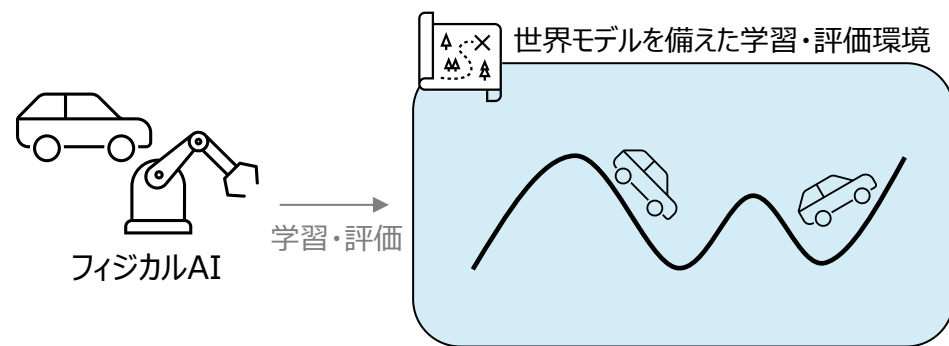
領域 ^注	世界モデル利用による影響	領域	世界モデル利用による影響
動画生成 	一貫性のある高品質な映像を生成。フィジカルAIの訓練用データ・シミュレータとして活用も可能。	自律型エージェント 	世界モデル内のシミュレーションを通じて学習することで、学習コストを削減、効率的な学習が可能。
自動運転 	一貫性のあるリアルな交通シナリオの生成と高精度な予測。	AI学習・評価環境 (AIが学習・評価するための身体的シミュレーション環境) 	動的かつ多様な仮想環境をリアルタイムに提供。仮想環境を利用するエージェントの適応力や汎化能力が向上。
ロボット 	環境の動態を予測し、未知の物理環境やタスクへの適応。および長期的な計画能力の向上。	社会シミュラクラ (エージェント同士の社会現象や人間行動の再現) 	現実社会のダイナミクスを仮想的に再現。社会現象の予測や複雑な人間関係のシミュレーションの精度が向上。

[画像出典]OpenAI, 2024/2/15, “Video generation models as world simulators”
(<https://openai.com/index/video-generation-models-as-world-simulators/>) 2026/4/6参照。

注：領域は参考文献[10]を参考。
[10]において「AI学習・評価環境」は「身体化環境（Embodied Environment）」と呼ばれている。

フィジカルAIへの活用

- 「物理的一貫性」や「時空間的一貫性」といった課題の解消は、現実世界で動作するフィジカルAI領域においてより重要となる。
- フィジカルAIにおける世界モデルの活用方法としては、AIモデルそのものに世界モデルを組み込むほか、AIの学習・評価環境として利用することが考えられる。

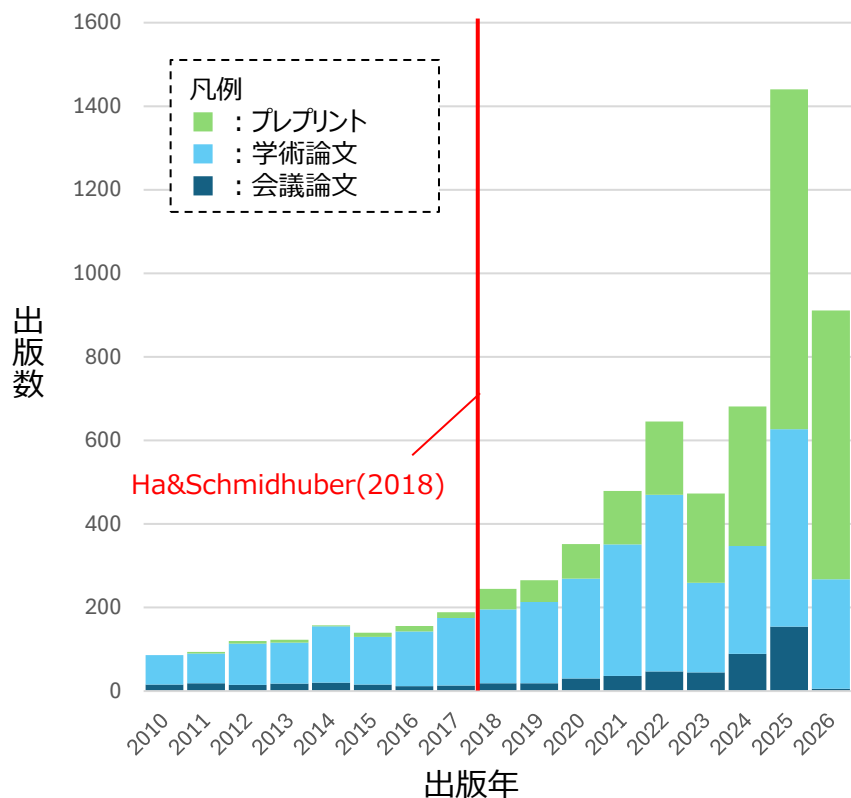
活用方法	概要	利点	具体的サービス (詳細は後述)
<p>AIモデルへの組み込み</p>	 <p>ロボットや自動運転が搭載するAIに世界モデルを組み込む</p>	<ul style="list-style-type: none"> 搭載するAIモデル自体が現実世界を理解。 →一貫性のある予測、精度の向上が可能。 	<ul style="list-style-type: none"> Dreamer 空間World Model (富士通) 人間系世界モデル (NEC) Terra (Turing)
<p>AIの学習・評価環境として利用</p>	 <p>世界モデルを備えた学習・評価環境</p> <p>学習・評価</p>	<ul style="list-style-type: none"> 現実世界に近い仮想環境や動画データを利用しAIを学習・評価。 →学習・評価のスピードアップ、コストダウン、精度の向上が可能。 <p>※現実の物理環境で学習・評価を行う場合、実機を用意して操作する時間やコストが発生。また事故を起こすシチュエーション等の稀なデータは得にくい。仮想環境上であれば様々なシチュエーションを何度も試すことができる。</p>	<ul style="list-style-type: none"> Cosmos (Nvidia) GAIA (Wayve) Genie(Google DeepMind) Terra (Turing)

世界モデルをめぐる動向

- 世界モデルの研究は、デビッド・ハ氏とユルゲン・シュミットフーバー氏が2018年に発表した論文「World models」を起点に研究数が増加。
- 近年は世界モデル開発の資金調達も活発化し、AIの思想的リーダーが世界モデルに注目していることから、期待感に拍車をかけている。

世界モデル研究の隆盛

- 2018年にプレプリント (arXiv) で公開された論文「World models」を起点に、世界モデルに関する研究が増加 (研究の詳細はp.8)。



Lens.orgで検索を行い日本総研作成。
 検索語: 'world model' OR 'world models' OR 'world foundation model' OR 'world foundation models'
 分野: Artificial Intelligence, Machine learning, Computer Science
 年: 2010/1/1-2026/3/31
 ドキュメントタイプ: journal article, preprint, conference proceeding article

世界モデルへの投資

- 世界モデルは社会的にも近年注目されており、世界モデル開発の資金調達が活発化。
- World Labsはフェイス・リ氏、Advanced Machine Intelligence (AMI) はヤン・ルカン氏が設立したスタートアップ企業であり、AIの思想的リーダーが世界モデルに注目していることも期待感に拍車をかけている。
- 世界モデルは応用先の幅広さから、世界モデル単体市場としてではなくAI開発全般を加速させる基盤技術として、今後も発展していくことが見込まれる。

会社名 (国)	調達額	主要調達先	発表日	サービス名: 詳細
World Labs (アメリカ) *4	10億ドル	AMD, Autodesk, Emerson Collective, Fidelity Management & Research Company, NVIDIA, Sea	2026/2/18	Marble: 画像やテキストから3D空間 (ガウスプラットフォーム) を生成する世界モデル
WAYVE (イギリス) *5	15億ドル	Microsoft, NVIDIA, Uber, Mercedes-Benz, Nissan, Stellantis	2026/2/25	GAIA: 自動運転向けの動画生成を行う世界モデル (詳細 p.9)
AMI (フランス) *6	10億ドル	Cathay Innovation, Greycroft, Hiro Capital, HV Capital, Bezos Expeditions	2026/3/9	JEPA: 動画等から物理現象や因果関係を理解・予測する世界モデル (詳細 p.7)

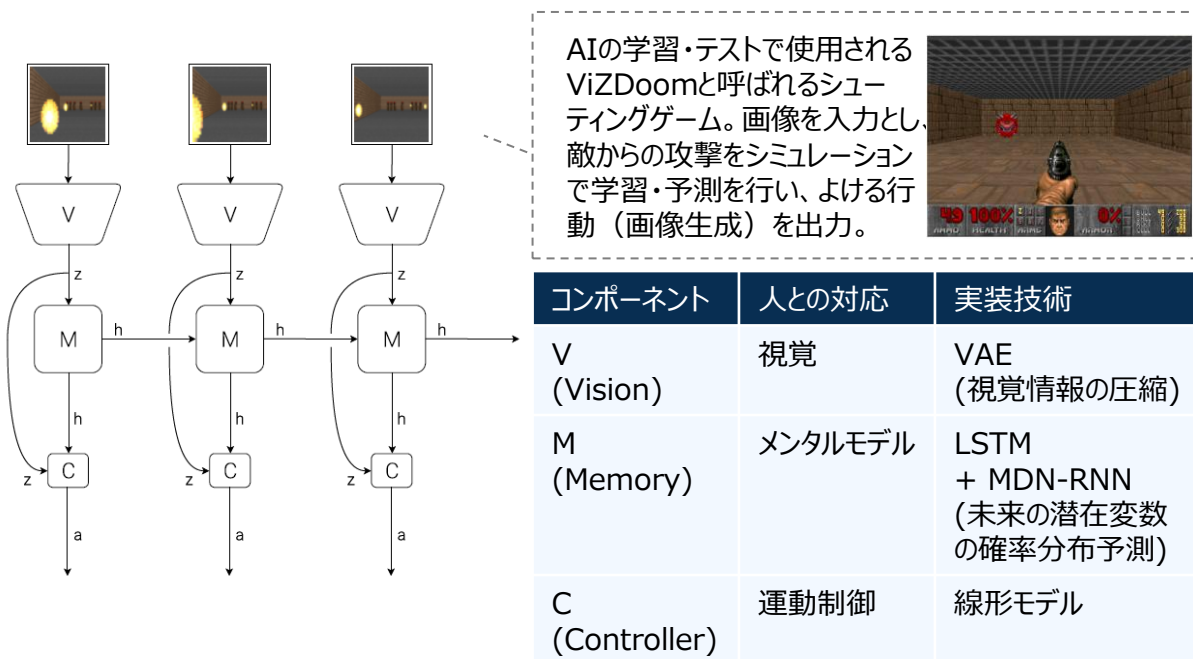
世界モデルの初期の研究

- 世界モデルは2018年に発表された論文「World Models」から広まったとされており、同様の思想を引き継いだAI研究の第一人者であるヤン・ルカン氏が提唱した**JEPAモデル**が世界モデルのアーキテクチャとして有名。
- 初期の研究はいかにヒトの脳に近いアーキテクチャを考えるかという、「**設計思想としての世界モデル**」が前面に出てきたといえる。

ハ & シュミットフーバーの「World Models」*7

自律型エージェント

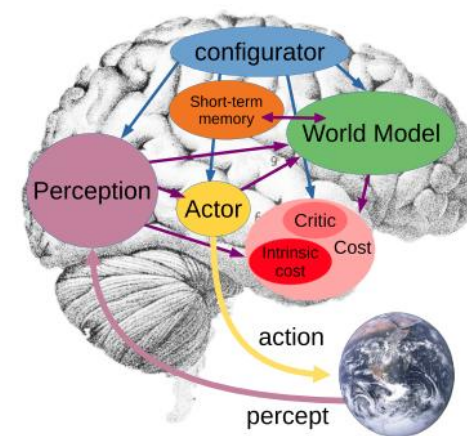
- ヒトは世界を理解・予測するための枠組み「メンタルモデル」を持つ。これにより少ない情報から頭の中でシミュレーションをすることができる。AIでも同様に、世界を理解・予測するための枠組みとして「世界モデル」を提唱。
- 具体的なアーキテクチャとして「VMCモデル」を考案し、世界モデルに該当するM (Memory) 上でシミュレーションを行い予測を行う。



ヤン・ルカンのJEPAモデル*8

自律型エージェント

- VMCモデルの思想を引き継ぎ、よりヒトの脳に近いモデルとして**Joint-Embedding Predictive Architecture (JEPA) モデル**を提案。
- 脳に近いアーキテクチャを考えた背景として、現状の生成AIは脳と異なり、見えたものすべてをピクセルレベルで生成している点を批判。生成を重視するAI研究の流れから脱し、具体的なタスクに即した予測に特化するようAI研究の方向性を変えることを主張している。
- より脳に近いアーキテクチャを用いることで、心理学におけるシステム1思考（直感的行動）とシステム2思考（熟慮的行動）の実現を目指す。システム2思考のためには世界モデルを実装することが必要不可欠と考えている。



コンポーネント	人との対応
Configurator	前頭前野 (制御・注意)
Perception	感覚野
World Model	前頭前野
Cost	基底核・扁桃体・前頭前野 (報酬)
Actor	前運動野
Short-term Memory	海馬

※実装技術については論文中では示唆に留まるため割愛

[画像出典]Ha & Schmidhuber, 2018, "World Models", arXiv (https://arxiv.org/abs/1803.10122) 2024/4/3参照。

[画像出典]Kepka et al., 2016, "ViZDoom: A Doom-based AI Research Platform for Visual Reinforcement Learning", arXiv (https://arxiv.org/abs/1605.02097) 2024/4/3参照。

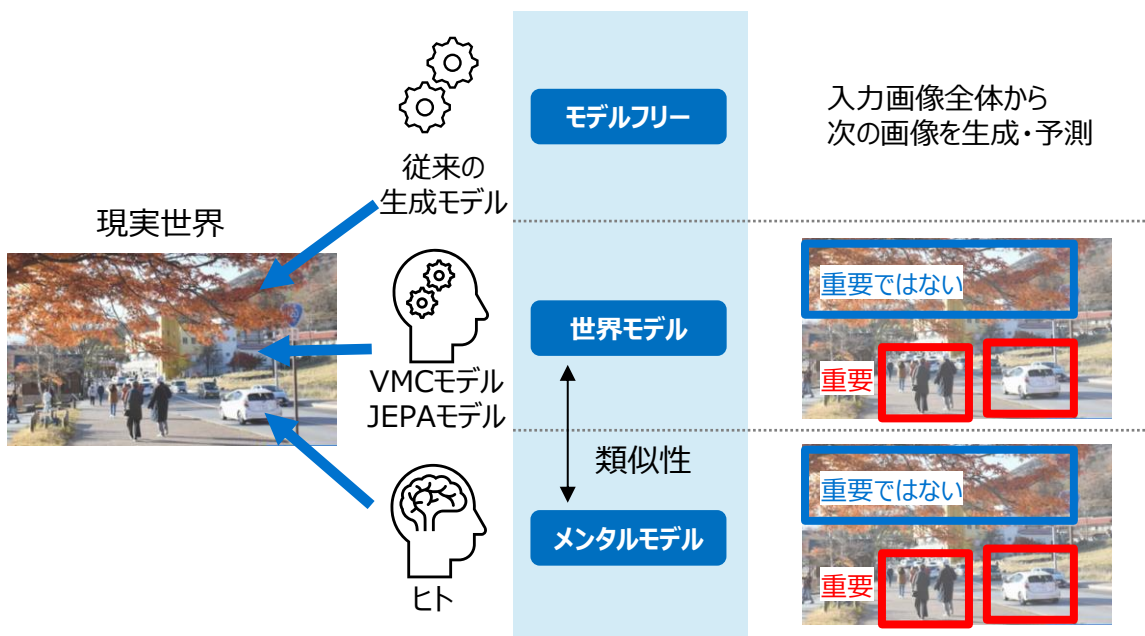
[画像出典]LeCun, 2022, "A Path Towards Autonomous Machine Intelligence Version 0.9.2, 2022-06-27" (https://openreview.net/pdf?id=BZ5a1r-kVsf) 2024/4/3参照。

世界モデルの広がり

- 初期の世界モデル研究はいかにヒトの脳に近いアーキテクチャを採用することで現行のAIの課題を乗り越えるかという点（設計思想としての世界モデル）に注力。
- その後、多くの世界モデルに関する研究やサービスが開発される中で、世界モデルの捉え方は多様化。
- さらには世界モデルはヒトの脳に近いアーキテクチャも提案されていることから、人に類似したAIという意味での汎用人工知能（AGI）の実現に近づくものという期待感も合わさる。しかしそれゆえ多くの期待が世界モデルと結び付けられ、**世界モデルが何か特定の技術やアーキテクチャを指すのか、課題を乗り越えた状態を指すのか等、使われ方が曖昧になっている。**

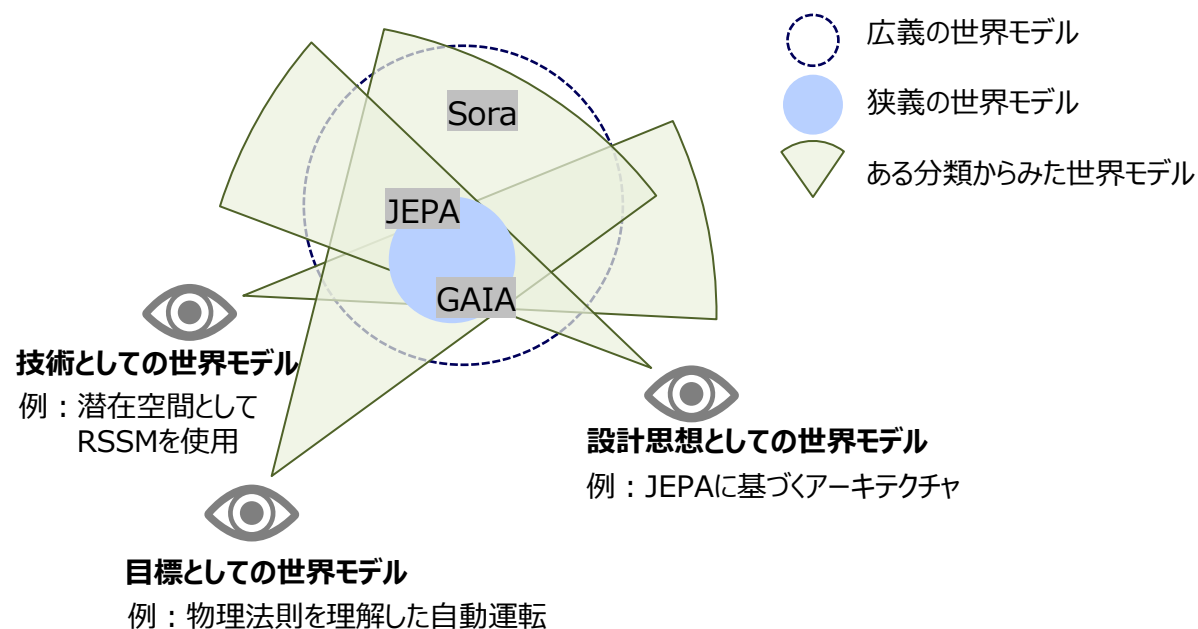
設計思想としての世界モデル

- 自動運転等のAIが生成・予測を行う上で、人物や車の動きの理解・予測は重要である一方、木や葉の細かい揺れは運転タスクにとって重要ではない。しかし従来の生成モデルではこれを同列に扱うため、計算コストが高く一貫性に欠ける。
- ヒトはメンタルモデルを通じて世界を理解・予測する。AIでもヒトと類似したアーキテクチャを目指し、ヒトにおけるメンタルモデルの代わりに世界モデルを通じて世界を理解・予測する。



世界モデルを分類する枠組みの多様化

- 世界モデルの研究・開発の進展に伴い、それらの分類が行われている^{*9-11}。しかし、**MECEな（漏れや重複のない）分類基準はまだ整備途中である。**
- 本レポートは、どの分類基準でも世界モデルに該当するサービスや研究事例（狭義の世界モデル）について説明する（pp.10~13）。



世界モデルのサービス事例（海外）

- 海外における世界モデルのサービス事例として、Nvidiaの**Cosmos**、Wayveの**GAIA**、Google DeepMindの**Genie**、**Dreamer**を取り上げる。
- 世界モデルの用途としては、AIシステム自体に世界モデルを組み込むほか、AIシステムの学習・シミュレータ環境として世界モデルを活用することが考えられている。

Cosmos (Nvidia) *12

ロボット注

- Nvidiaは2025年、「**世界基盤モデル**」と呼称する**Cosmos**シリーズを発表。
- CosmosはフィジカルAIモデルを構築するための開発基盤。物理的一貫性を得るための学習やテストを行ううえで仮想環境となり、これを用いることで現実環境における学習のリスクやコストを削減。

	名称	内容	入力	出力
	Cosmos Predict	世界生成モデル 入力データから予測を行い、30秒の動画を生成	テキスト 画像 動画	動画
	Cosmos Transfer	マルチコントロールモデル シミュレーション結果をもとに、リアルな環境における合成データを生成	シミュレーション動画	実写動画
	Cosmos Reason	視覚言語モデル (VLM) 入力データを物理的、事前知識の観点から理解・判断	画像 動画	言語

[画像出典] NVIDIA, n.d., “フィジカル AI NVIDIA Cosmos” (<https://www.nvidia.com/ja-jp/ai/cosmos/>) 2026/4/7参照。

GAIA (Wayve) *13-16

自動運転

- 自動運転のスタートアップ企業であるWayveは、自動運転向け動画生成世界モデル「**GAIA**」を開発。
- 2025年には最新のGAIA-3を発表。リアルな動画を生成するというだけの目的ではなく、生成された動画を自動運転AIが学習し、その**安全性を科学的に証明するための基盤**という位置づけになることを想定。

名称	リリース	パラメータ	学習データ	アーキテクチャ	特徴
GAIA-1	2023/3	91億	ロンドン	自己回帰トランスフォーマー 動画拡散モデル	視覚合成データ生成
GAIA-2	2025/3	87億	イギリス アメリカ ドイツ	動画トークナイザー 潜在拡散モデル	マルチカメラ生成による時空間的一貫性の向上
GAIA-3	2025/12	150億	3大陸 9か国 (日本含)	動画トークナイザー 潜在拡散モデル	視覚合成データ生成ツールから自動運転評価と安全性を検証するための基盤へ



複数視点の動画生成において

- ・歩行者
- ・自転車
- ・バイク
- ・ランドマーク（英国のビッグ・ベン）が破綻なく描写されている

[画像出典] WAYVE, 2025/12/2, “GAIA-3: Scaling World Models to Power Safety and Evaluation” (<https://wayve.ai/thinking/gaia-3/>) 2026/4/6参照。

世界モデルのサービス事例（海外）

- Google DeepMindのGenieはユーザのリアルタイムのインタラクションで世界が生成されることが特徴。
- Dreamerは少ないデータから世界モデル内でのシミュレーションによる学習ができることを示し、フィジカルAIの安全性や学習・検証におけるデータ効率の向上が見込まれる。

Genie (Google DeepMind) *17

AI学習・評価環境

- Google DeepMindは2025年に最新の**Genie3**を発表。テキストから3D世界を生成し、ユーザはキャラクターを操作可能。**キャラクターの動きに合わせて動的に周辺環境が生成**される。
- 環境はリアルな自然だけでなくフィクションを織り交ぜることも可能。
- 応用先としてはゲームの他、教育、自動運転の訓練等を想定しており、さらには汎用人工知能に向けた一歩だと述べる。
- 課題としては、行動範囲の制限、他のエージェントとの相互作用、完全な実在の地理の再現、テキストレンダリングの限界、数分間という継続時間を挙げている。

世界の物理的特性をモデル化しシミュレート

フィクションとの融合



ユーザのインタラクションに応じたリアルタイムの世界生成

数分間にわたる長期的な環境の一貫性

[画像出典] Google DeepMind, 2025/8/5, "Genie 3: A new frontier for world models"
(<https://deepmind.google/blog/genie-3-a-new-frontier-for-world-models/>) 2026/4/6参照。

Dreamer*18

自律型エージェント

- Danijar Hafner氏（現Google DeepMind所属）らによる一連の研究。2020年にDreamerV1が発表され、最新は2025年に発表された**DreamerV4**。ソースコードやアルゴリズムは誰でも利用可能。
- AIEージェントがオープンワールドのサンドボックスゲームMinecraft上でタスクを実行。実際にエージェントがプレイすることなく、学習データのみから想像によるトレーニングを行ったのち、実際にゲームをプレイしタスクをクリア。
- 応用先としては、物理ロボットの学習を用いた安全性とデータ効率の向上の他、インターネット動画を学習に利用することで、汎用的なシミュレータとしての世界モデルを作成することができる

Agent task: Gather wood



Agent task: Mine cobblestone



Agent task: Craft stone pickaxe



Agent task: Mine diamonds



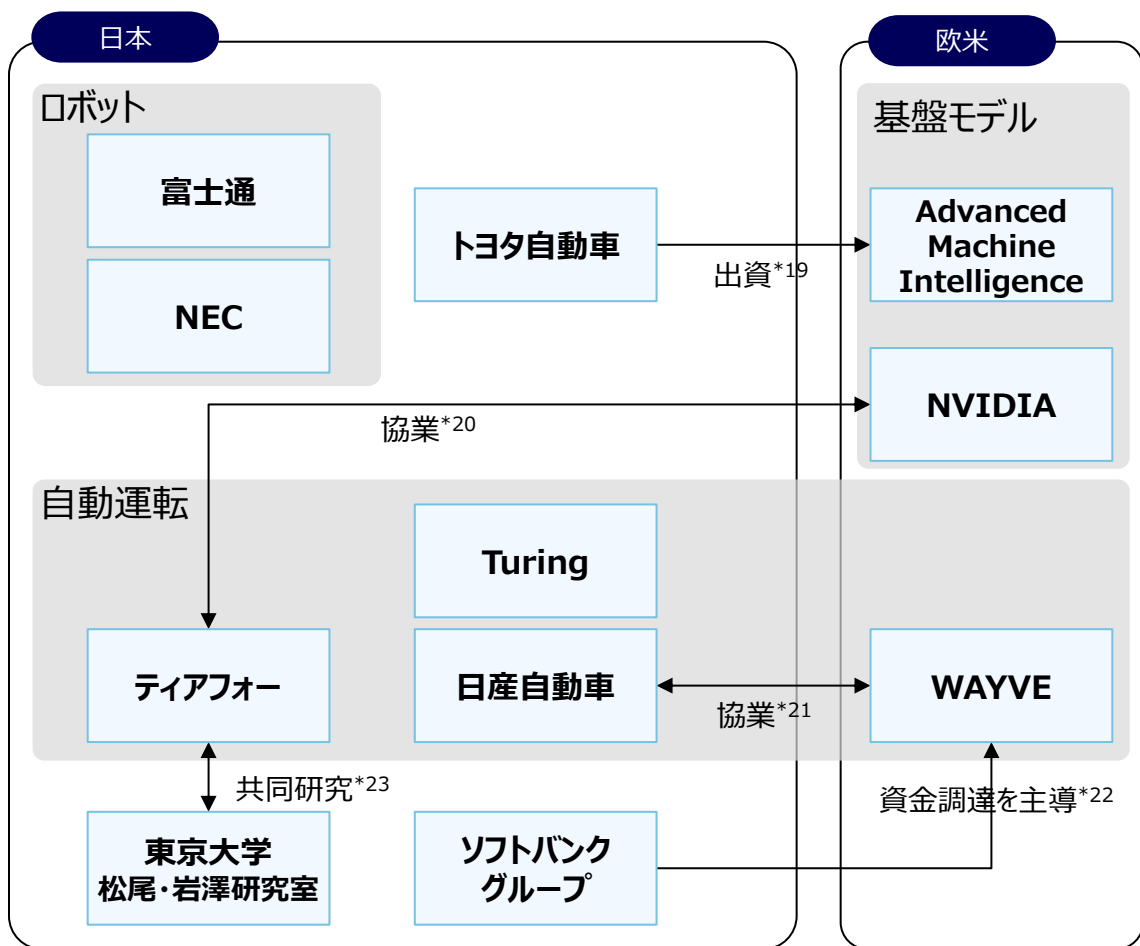
2541時間のプレイ動画データのうち、10~100時間のみアクション情報（何をしているのか明示されている情報）を付属し、残りは動画のみで学習。そこからアクション情報にない状況下であっても、Minecraftのプレイ空間にてタスク実行が可能。

[画像出典] Hafner et al., 2025, "Training Agents Inside of Scalable World Models", arXiv,
(<https://arxiv.org/abs/2509.24527>) 2026/4/7参照。

世界モデルのサービス事例（日本）

- 日本では自動運転やロボット等のフィジカルAIの分野において、応用にむけたアプリケーションレベルでの世界モデルの開発が進む。
- 自動運転領域では、E2E（AIが認識・判断・制御までを実施するアプローチ）方式を採用している企業において、世界モデルの活用の試みがみられる。

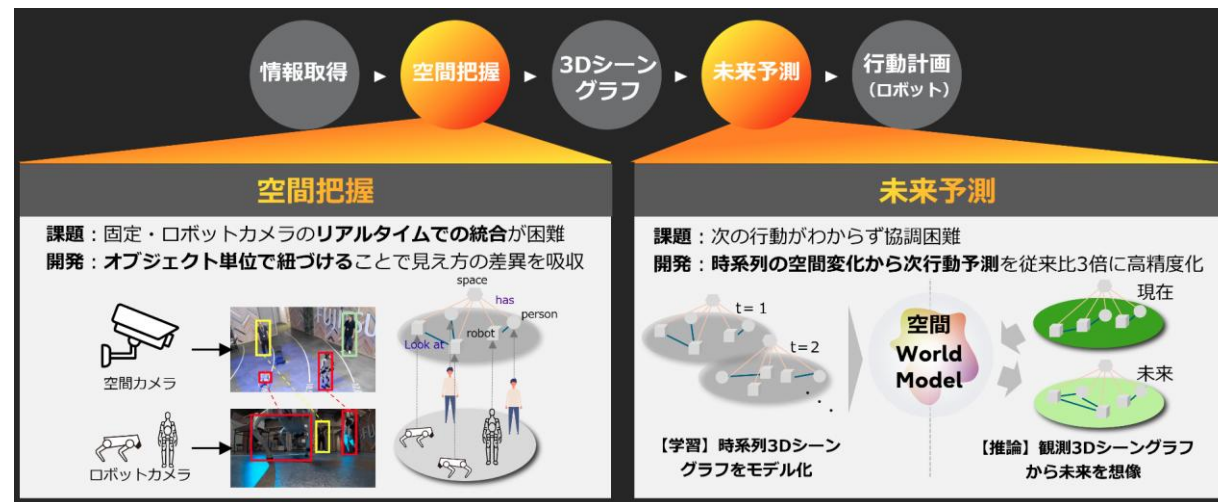
日本における主なプレイヤー



空間World Model（富士通）*24-25

ロボット

- 公共施設、商業施設における警備や接客などの対人向けサービスに使用するフィジカルAIモデルとして「空間World Model」を2025年に発表。
- 多くの世界モデルはロボットからの視覚画像のみを学習しているなか、空間World Modelは配備されているセンサーデータも学習に含めることで、空間全体の世界モデルを構築していることを技術的優位性としている。
- 空間World Modelの利用により、自律ロボットの高度化、人とロボットの協調、ロボット同士の協調が可能になる。2026年度中の技術実証を予定。



[画像出典]富士通, n.d., “空間World Model” (<https://documents.research.global.fujitsu.com/spatial-world-model/>) 2026/4/6参照。

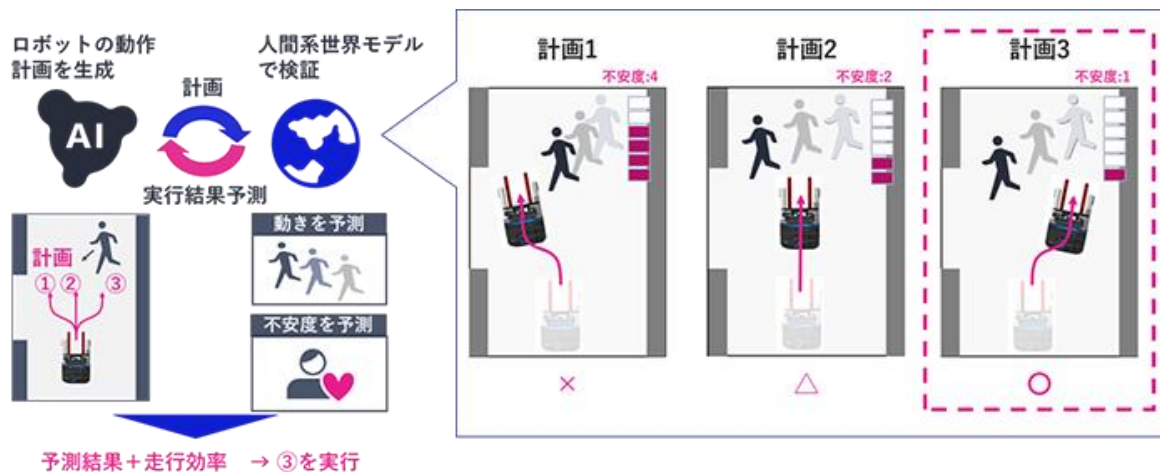
世界モデルのサービス事例（日本）

- ロボット領域では、富士通の他、NECが実用化に向けて世界モデルを開発。
- 自動運転領域ではTuringによるTerraが開発されているが、実用化には至っていない状況。

人間系世界モデル（NEC）*26

ロボット

- NECはロボット制御のフィジカルAIとして2026年に「人間系世界モデル」を発表。
- 人間系世界モデルの特徴として、以下の機能を実装
 - ✓ ロボットの挙動や物理的な周囲の状況によって変化する人の動きを予測
→人との衝突や接触を避けることによる身体的安全性の確保
 - ✓ ロボットの接近によって変化する、人の不安の程度をリアルタイムかつ定量的に推定
→人の不安を高めない行動の選択による心理的安全性の確保
- 労働力不足や危険作業への対応策になりうると考えており、2027年度中に実用化を目指す。

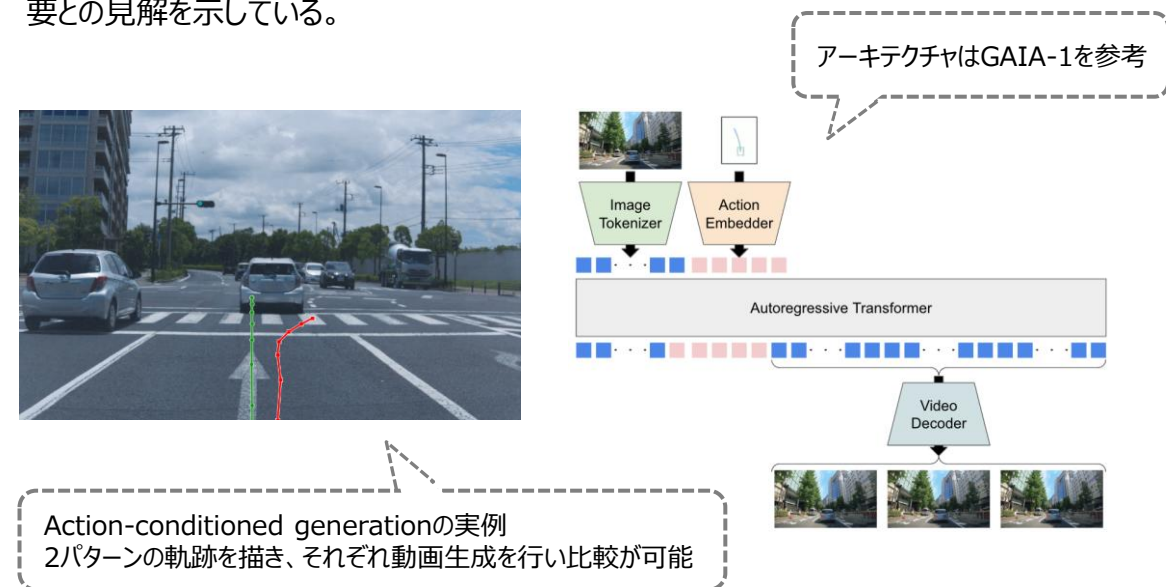


計画2でも人に当たらずに走行できるが、計画3のほうが人に不安を与えない

Terra (Turing) *27-29

自動運転

- Turingは日本初の自動運转向け生成世界モデルとして2024年に「Terra」を開発。
- Terraには、大きく以下の2つの機能が存在。
 - ✓ Video rollout : 動画を入力とし、そこから任意の後続フレームを出力
 - ✓ Action-conditioned generation : 動画内に軌跡を描き、それに沿って動画を生成
- 生成される動画の特徴として、学習データに入っていない事故映像も生成できている。
- 利用用途としてはシミュレータとして使うだけでなく、自動運転システムの一要素としての利用が可能。ただしTuring社は、現時点では実用化を目的としておらず、さらなる精度向上が必要との見解を示している。



[画像出典] NEC, 2026/3/12, “NEC、人の動きと心理状態を予測する世界モデルを活用し、先回りしてロボットを制御するフィジカルAIを世界で初めて開発” (https://jpn.nec.com/press/202603/20260312_01.html) 2026/4/6参照。

[画像出典] Turing, 2024/8/14, “運転版の“Sora”を作る: 動画生成の世界モデルTerraの開発背景” (https://zenn.dev/turing_motors/articles/6c0ddc10aae542) 2026/4/7参照。

考察・今後の課題

- 世界モデルを活用した高品質な動画生成は、AIが学習で使用するデータが不足している問題の打開策にもなるため、多くのAI研究・開発を加速させる。
- 自動運転ではE2E方式においてスタンダードになる可能性がある。しかしE2E方式の覇権獲得や、実用化の直結を意味しているわけではない。
- 世界モデルの定義は未確定で曖昧なため、技術評価や社会的インパクトの評価に際しては使われる文脈について考慮する必要がある。
- 技術的な課題としては長期的な一貫性の取得や、物理的な因果関係や法則を理解していることを示す指標・ベンチマークの整備が必要。

世界モデルは今後AI研究・開発の中でどう位置付けられるか

世界モデルという研究フロンティア

- 従来の生成AIが言語や動画の生成が目的になることが多い中、世界モデルの登場で「世界」を作り出すという目的のシフトにより研究・開発が活発化。

世界モデル利用用途の多様性・柔軟性

- 世界モデルの用途として、(1)システム自体に世界モデルを組み込む、(2)システムの学習・テストに世界モデルで構築した環境・データを用いることが可能。
- 用途の多様性、柔軟性から、既存のAI研究への活用がしやすく、多くの分野で今後言及、利用の検討が見込まれる。

良質な学習データ不足問題の打開策

- システムの学習・テストに世界モデルを使用するという観点で見たとき、良質な学習データの枯渇への打開策になると考えられる。既存のモデルの多くはデータを増やすことで性能を上げるというスケール則に頼っており、データの枯渇は性能向上の限界を意味する^{*30}。世界モデルで生成された一貫性のあるデータは、この課題への打開策として多くの分野で活用が見込まれる。

フィジカルAIへの影響

自動運転のE2E方式におけるスタンダード

- 自動運転領域において、E2E方式に関してはスタンダードになる可能性がある。
- 一方、世界モデルを採用しても、E2E方式自体の課題は残る（例えば、カメラからの入力による画像処理だけの場合、車や地面の反射による視界不良による誤認・誤操作が起こる問題^{*30}は残る）。
- そのため、世界モデルの採用が、自動運転におけるE2E方式の覇権獲得や、自動運転の実用化に直結するわけではない。

フィジカルAI開発の加速

- 世界モデルの活用による良質なシミュレーションデータをもとにしたフィジカルAIの学習が可能になる。これにより現実環境での学習・検証を減らすことにつながり、コスト・リスク・開発スピード面においてメリットがある。
- 事故映像等、現実では得にくいデータを世界モデルは潜在的に生み出すことができる。これにより現実のデータ以上のバリエーションがあり、AIモデルの質向上が見込まれる。

世界モデルの注意点・課題

定義の曖昧さによる誤用・乱用

- 世界モデルの実現 = AGIの実現、というわけではない。
- 世界モデルの定義が明確ではないため、評価の際にどの文脈、分類基準で使われているのか確認する必要がある。

「世界モデル」の 使われ方の分類	注意点
目標としての 世界モデル	結果的に（スケール則等で）一貫性が得られているだけで、世界モデル特有の技術・アーキテクチャが用いられているとは限らない。
技術としての 世界モデル	既存モデルより性能が上がっているとは限らない。
設計思想としての 世界モデル	既存モデルより性能が上がっているとは限らない。実装まで落とし込まれているとは限らない。

長期的な一貫性の取得

- 一貫性は向上しているが生成できる動画/環境は数分程度のため、長期的には不整合が生じる。

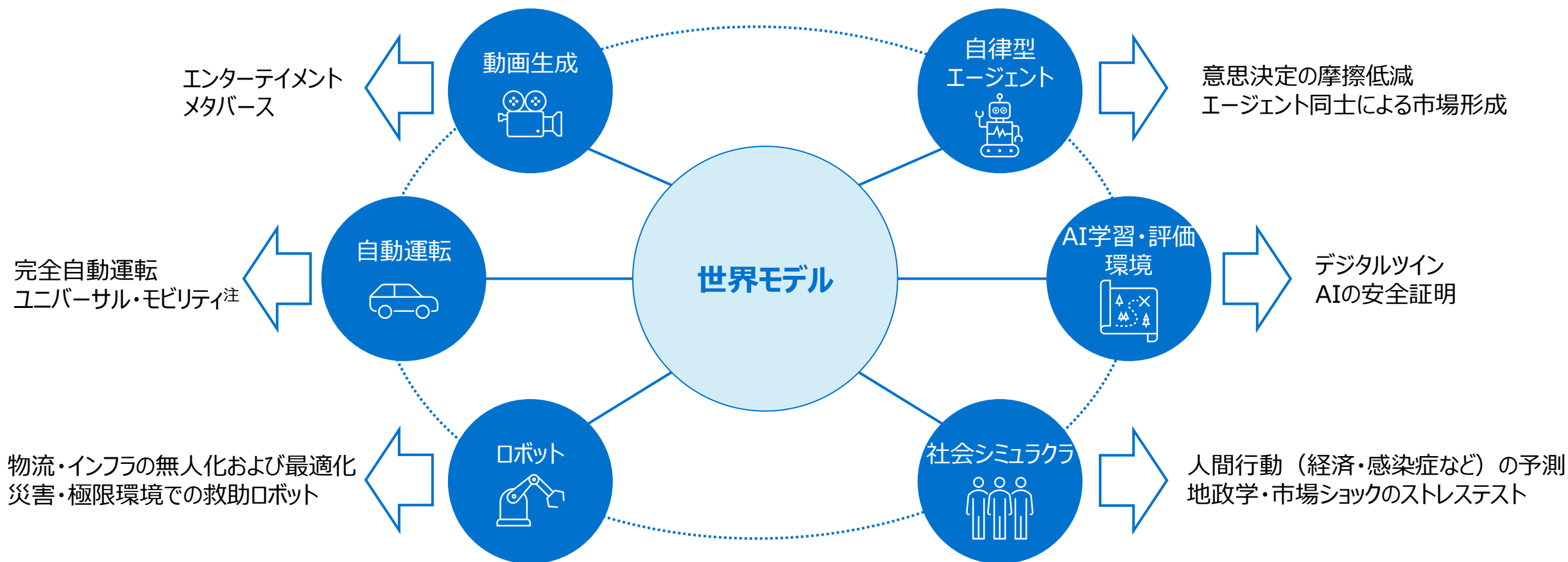
世界モデルの指標・ベンチマークの整備

- 世界モデルの共通的な評価を行うため指標・ベンチマークの整備が必要。特にNvidia CosmosのようなAIの評価基盤として位置付ける場合、仮想空間上での検証の妥当性を根拠づけるための、ガイドラインや法整備が必要。

まとめ

- 世界モデルは今後多くの領域において活用の可能性がある。動画生成の質向上により自動運転の精度が上がるなど、領域間を跨いだ活用も見込まれる。
- 世界モデルのような現実世界を理解するAIの発展は、AIの普及・実用化レベルにおいて求められている。そのため世界モデルの進展は、AIが社会に浸透するスピードを加速させる。
- ビジネスサイドにおける受け止め方としては、新奇な技術に振り回されることなく、社会のあらゆる方面でのAI普及・実用化が進むと一般に想定されてきた未来が確実性を帯びたという認識を持つ必要がある。今後、世界モデルを活用したAIの開発・実用化事例を踏まえ、自社のAIでも活用が可能か検討することが望ましい。

世界モデルの社会への波及効果



注：年齢や身体的状況に関わらず、「誰でも・いつでも・自由に」移動できる環境や手段のこと。



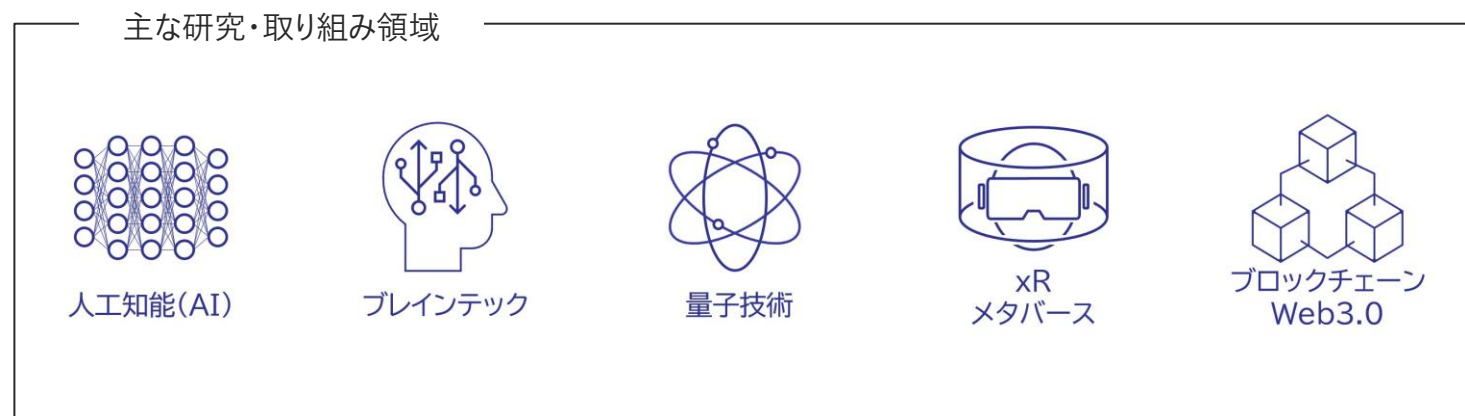
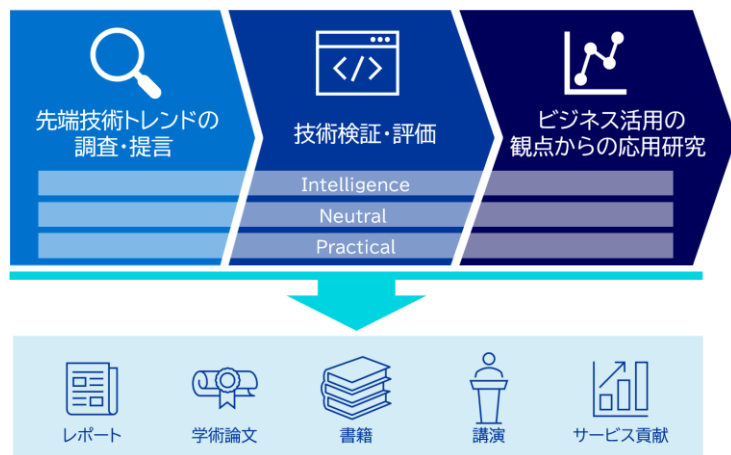
番号	文献名
*1	NVIDIA, n.d., “世界モデルとは? NVIDIA Glossary (用語集)” (https://www.nvidia.com/ja-jp/glossary/world-models/) 2026/4/16参照。
*2	松尾・岩澤研究室, 2022/11/30, “「世界モデル」とは何か? 知能の実現に向けて、松尾研が研究を推進する理由。” (https://weblab.t.u-tokyo.ac.jp/news/20221130/) 2026/4/16参照。
*3	LeCun, 2024/2/17, “X”(https://x.com/ylecun/status/1758740106955952191) 2026/4/3参照。
*4	World Labs, 2026/2/18, “World Labs Announces New Funding” (https://www.worldlabs.ai/blog/funding-2026) 2026/4/3参照。
*5	WAYVE, 206/2/25, “Wayve secures \$1.5B to deploy its global autonomy platform” (https://wayve.ai/press/series-d/) 2026/4/3参照。
*6	WIRED, 2026/3/12, “ヤン・ルカンのAIスタートアップ「AMI」が約1600億円を調達。世界モデル開発を加速” (https://wired.jp/article/yann-lecun-raises-dollar1-billion-to-build-ai-that-understands-the-physical-world/) 2026/4/3参照。
*7	Ha & Schmidhuber, 2018, “World Models”, <i>arXiv</i> (https://arxiv.org/abs/1803.10122) 2024/4/3参照。
*8	LeCun, 2022, “A Path Towards Autonomous Machine Intelligence Version 0.9.2, 2022-06-27” (https://openreview.net/pdf?id=BZ5a1r-kVsf) 2024/4/3参照。
*9	Zhu et al., 2024, “Is Sora a World Simulator? A Comprehensive Survey on General World Models and Beyond”, <i>arXiv</i> (https://arxiv.org/abs/2405.03520) 2026/4/3参照。
*10	Ding et al., 2025, “Understanding World or Predicting Future? A Comprehensive Survey of World Models”, <i>ACM Computing Surveys</i> , 58(3).
*11	Dickson, 2026, “Three ways AI is learning to understand the physical world” (https://venturebeat.com/technology/three-ways-ai-is-learning-to-understand-the-physical-world) 2026/4/3参照。
*12	NVIDIA, n.d., “フィジカル AI NVIDIA Cosmos” (https://www.nvidia.com/ja-jp/ai/cosmos/) 2026/4/7参照。
*13	WAYVE, 2023/10/3, “Scaling GAIA-1: 9-billion parameter generative world model for autonomous driving” (https://wayve.ai/thinking/scaling-gaia-1/) 2026/4/6参照。
*14	WAYVE, 2025/3/26, “GAIA-2: Pushing the Boundaries of Video Generative Models for Safer Assisted and Automated Driving” (https://wayve.ai/thinking/gaia-2/) 2026/4/6参照。
*15	Russel et al., 2025, “GAIA-2: A Controllable Multi-View Generative World Model for Autonomous Driving”, <i>arXiv</i> (https://arxiv.org/abs/2503.20523) 2026/4/6参照。
*16	WAYVE, 2025/12/2, “GAIA-3: Scaling World Models to Power Safety and Evaluation” (https://wayve.ai/thinking/gaia-3/) 2026/4/6参照。
*17	Google DeepMind, 2025/8/5, “Genie 3: A new frontier for world models” (https://deepmind.google/blog/genie-3-a-new-frontier-for-world-models/) 2026/4/6参照。
*18	Hafner et al., 2025, “Training Agents Inside of Scalable World Models”, <i>arXiv</i> , (https://arxiv.org/abs/2509.24527) 2026/4/7参照。
*19	日本経済新聞, 2026/3/10, “ヤン・ルカン氏、トヨタ・NVIDIAと新AI「世界モデル」 1630億円調達” (https://www.nikkei.com/article/DGXZQOGN101680Q6A310C2000000/) 2026/4/3参照。
*20	ティアフォー, 2026/3/19, “ティアフォー、NVIDIAのVLAモデルと世界基盤モデルを用いてAIベース型自動運転レベル4を加速” (https://tier4.jp/media/detail/?sys_id=4IPTZkyQIUoi2usopscMvj&category=NEWS) 2026/4/14参照。
*21	日産自動車, 2025/12/10, “日産とWayve、次世代運転支援技術の量産車への搭載に向けた協業契約を締結” (https://global.nissannews.com/ja-JP/releases/251210-01-j) 2026/4/3参照。
*22	WAYVE, 2024/5/7, “Wayve Raises Over \$1 Billion Led by SoftBank to Develop Embodied AI Products for Automated Driving” (https://wayve.ai/press/series-c/) 2026/4/3参照。
*23	ティアフォー, 2024/10/17, “2024年度 松尾研究所との共同研究プロジェクト：世界モデルの自動運転分野への応用” (https://tier4.jp/media/detail/?sys_id=4A8tHY41hoC9qs7maJRNJ1&category=BLOG) 2026/4/3参照。



番号	文献名
*24	富士通, n.d., “空間World Model” (https://documents.research.global.fujitsu.com/spatial-world-model/) 2026/4/6参照。
*25	富士通, 2026/12/2, “人とロボットが共存・協働する未来を拓く 空間World Model技術を開発” (https://global.fujitsu/ja-jp/pr/news/2025/12/02-02) 2026/4/8参照。
*26	NEC, 2026/3/12, “NEC、人の動きと心理状態を予測する世界モデルを活用し、先回りしてロボットを制御するフィジカルAIを世界で初めて開発” (https://jpn.nec.com/press/202603/20260312_01.html) 2026/4/6参照。
*27	Turing, 2024/8/14, “運転版の“Sora”を作る: 動画生成の世界モデルTerraの開発背景” (https://zenn.dev/turing_motors/articles/6c0ddc10aae542) 2026/4/7参照。
*28	Turing, 2024/8/14, “チューリング、日本初の自動運转向け生成世界モデル「Terra」を開発、リアルな運転シーンを自在に生成、より安全かつ効率的な自動運転開発を加速” (https://tur.ing/news/20240814/) 2026/4/3参照。
*29	Turing, 2024/10/1, “Turing TechTalk #5 自動運転のための世界モデル” (https://www.youtube.com/watch?v=qkIm3uW5Ats&list=PL757EZ4TpBICoefL5i5Ys6T2dJ0QsP3lh&index=33) 2026/4/3参照。
*30	Zia, 2025/12/28, “AI’s Next Scaling Law: Not More Data, but Better World Models” (https://www.unite.ai/ais-next-scaling-law-not-more-data-but-better-world-models/) 2026/4/24参照。
*31	Secci & Ceccarelli, 2020, "On failures of RGB cameras and their effects in autonomous driving applications," <i>2020 IEEE 31st International Symposium on Software Reliability Engineering (ISSRE)</i> , Coimbra, Portugal, pp. 13-24, doi: 10.1109/ISSRE5003.2020.00011.

先端技術ラボのご紹介

先端技術を活用したITサービスの創出に向けた技術の目利き役として、「先端技術トレンドの調査・提言」、「技術検証・評価」、「ビジネス活用の観点からの応用研究」に取り組んでいます。



当社ホームページの [特集サイト](#) では、IT分野における先端技術の調査レポート、及び所属する部員のプロフィール詳細がご覧いただけますので、ぜひご参照ください。

本レポート執筆者へのメディア取材や講演などに関するご相談につきましては、当社ホームページの [お問い合わせフォーム](#) よりご連絡ください。

株式会社日本総合研究所

日本総研は、シンクタンク・コンサルティング・ITソリューションの3つの機能を有するSMBCグループの総合情報サービス企業です。

東京本社 〒141-0022 東京都品川区東五反田2丁目18番1号 大崎フォレストビルディング

大阪本社 〒550-0001 大阪市西区土佐堀2丁目2番4号



日本総研

The Japan Research Institute, Limited