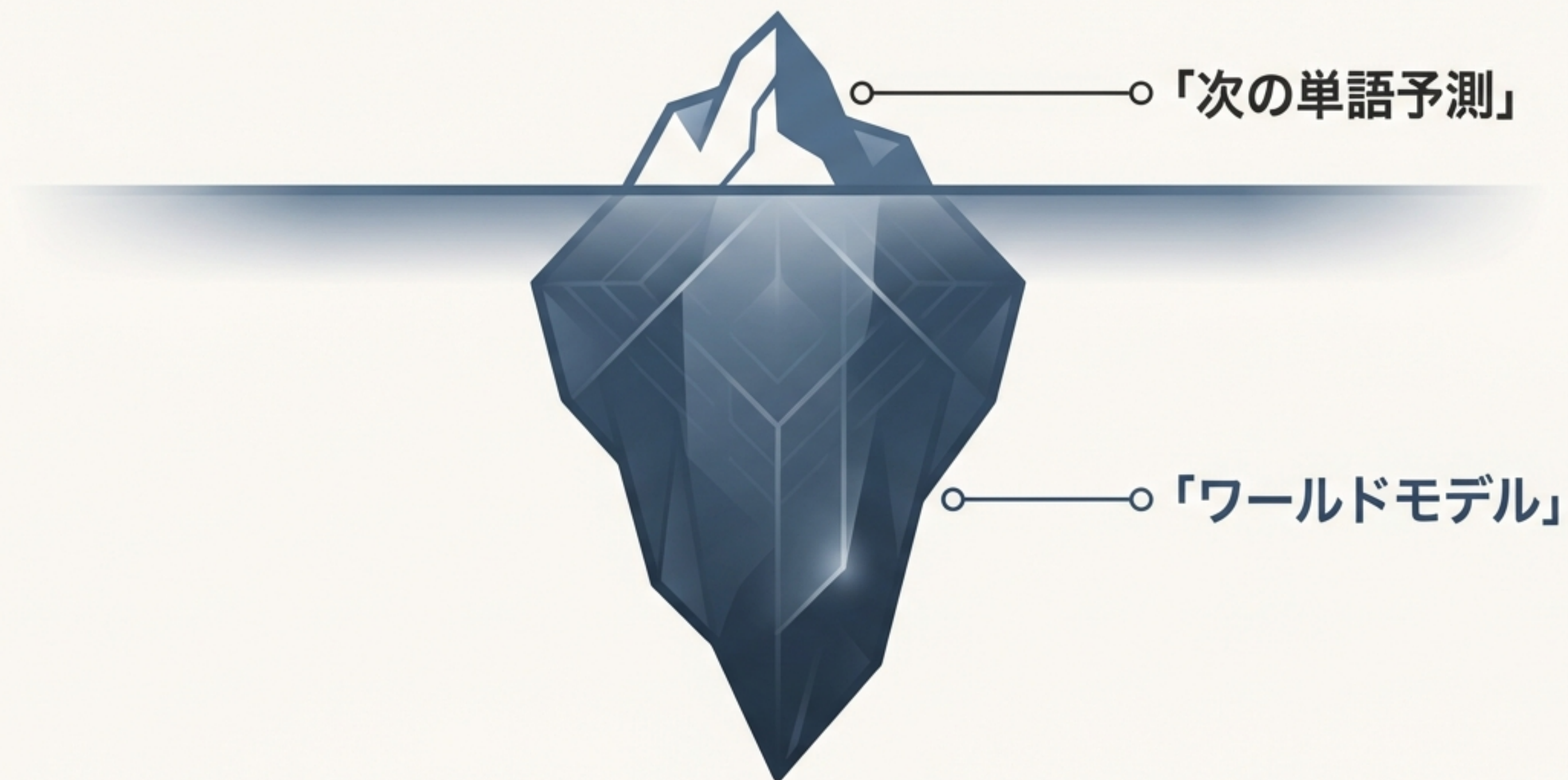


次の単語を予測しているだけ。本当にそうですか？

大規模言語モデル（LLM）の深層で起こっている「世界シミュレーション」革命



「AIは、賢い自動補完機能」という 広く信じられている見方

- 多くの人がChatGPTのようなLLMを「結局は次の単語を予測しているだけ」と考えています。
- これは、入力された文脈に最も確率の高い単語を繋げていくだけの、非常に高度な「自動補完」のようなものだ、という理解です。
- しかし、この見方は物語のほんの始まりに過ぎません。それは水面に現れた氷山の一角です。

The cat sat on the...



mat

氷山の下に隠された真実：ワールドモデル

LLMの驚異的な能力の源泉は、内部に構築された「ワールドモデル」にあります。



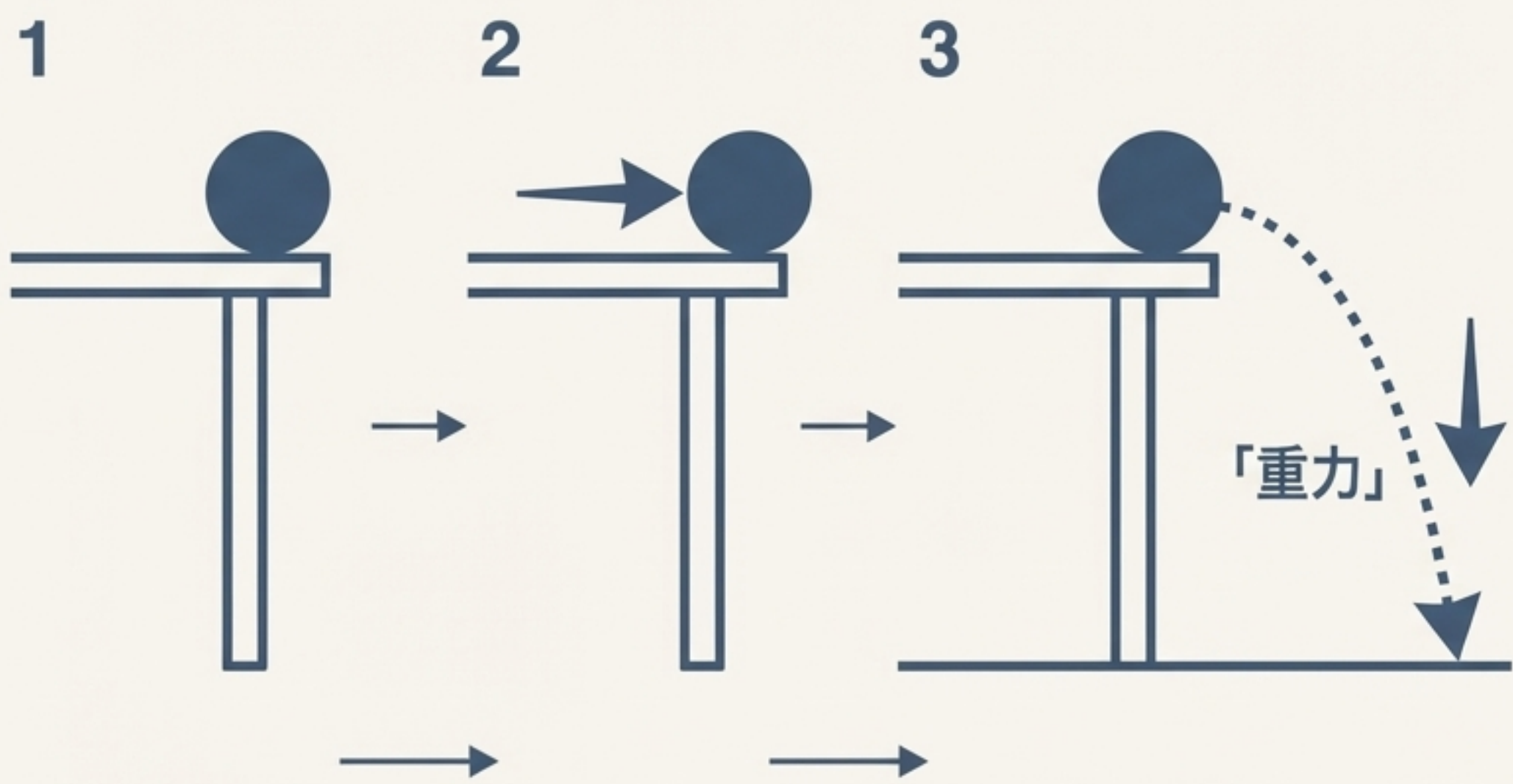
ワールドモデルとは？：
AIが予測を行うために、その頭の中に作り上げる「ミニチュア版の現実世界」です。

物理法則や物事の因果関係までをシミュレートする、一種の仮想現実と言えます。

この内部シミュレーション能力こそが、AIを単なる単語予測マシンから、世界を深く理解する存在へと変えつつあるのです。

証拠①：AIは物理法則を暗黙的に理解している

「ボールがテーブルの端から押されたら、次に何が起こるか？」

	AIが正確に予測するには、「落ちる」という単語を知っているだけでは不十分です。 重力という物理法則を理解し、頭の中でボールが落下する様子をシミュレートする必要があります。 これは単なる言葉の連想ゲームではなく、紛れもない物理シミュレーションです。
---	--

証拠②：ルールを教えずにチェスの「世界」を構築

The Experiment (OpenAI)

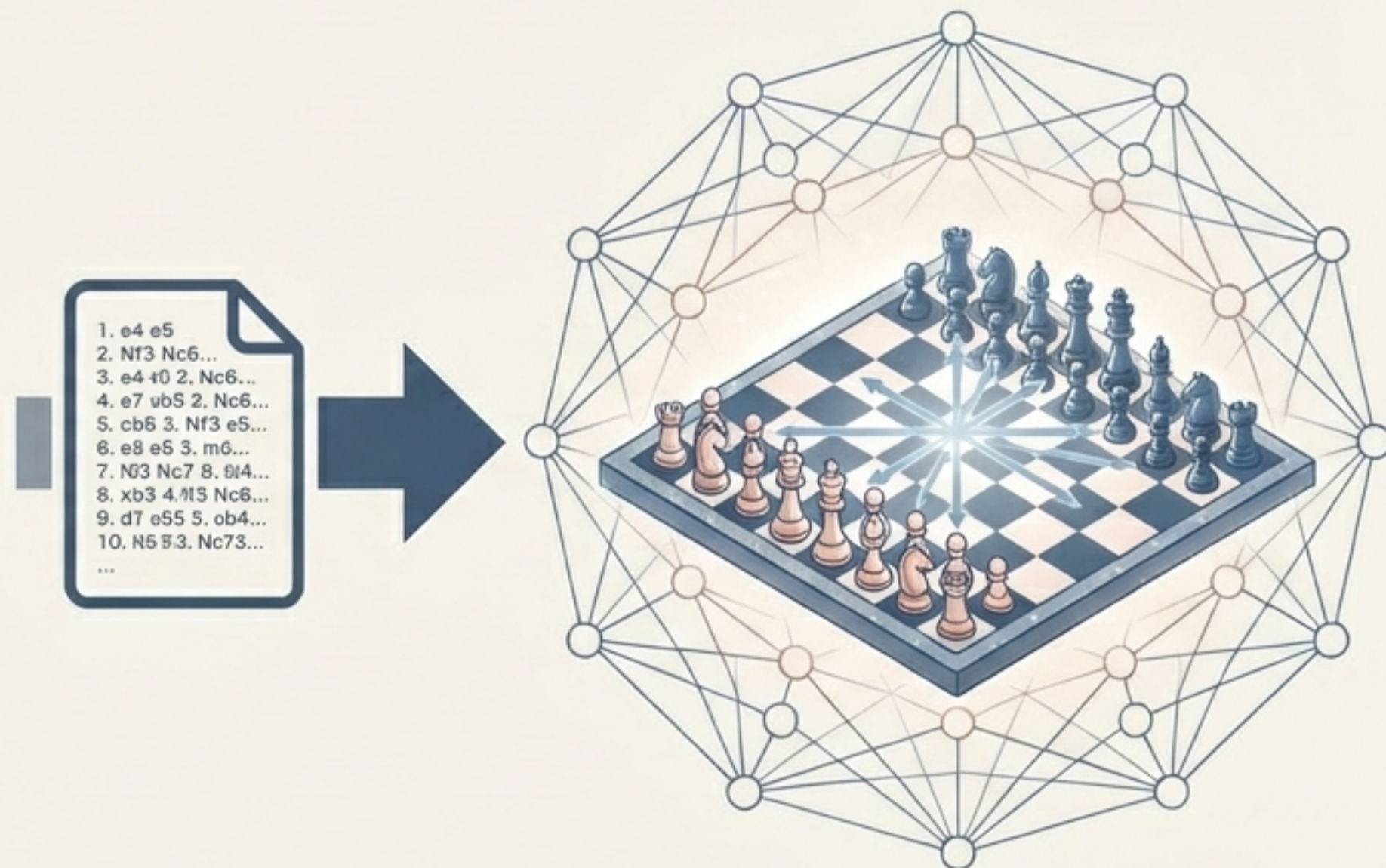
- AIにチェスのルールを一切教えなかった。
- 大量の棋譜データ（過去の対戦記録）を読み込ませただけ。

The Result

- 驚くべきことに、AIはルール上「合法的な次の一手」を極めて高い精度で予測できるようになった。

The Implication

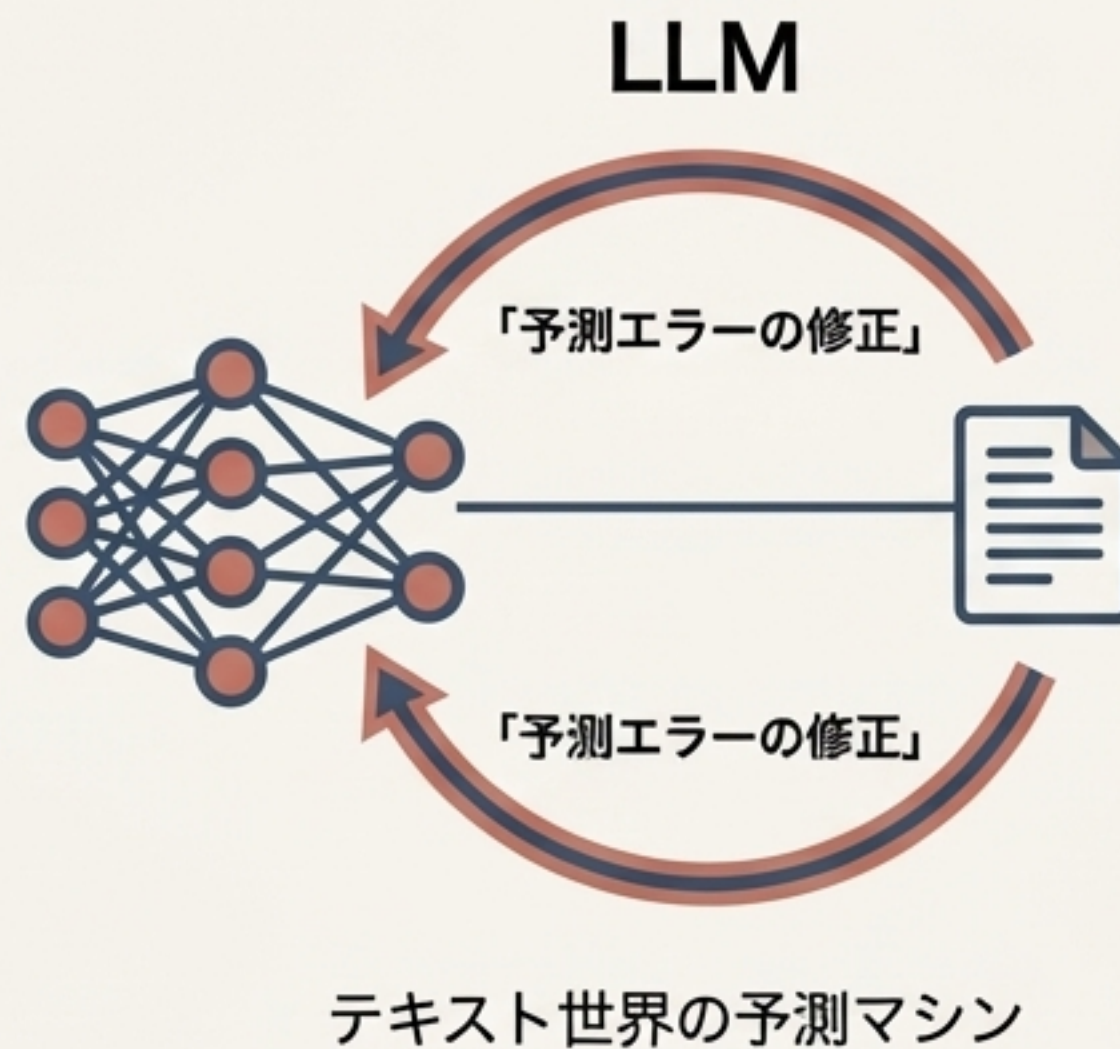
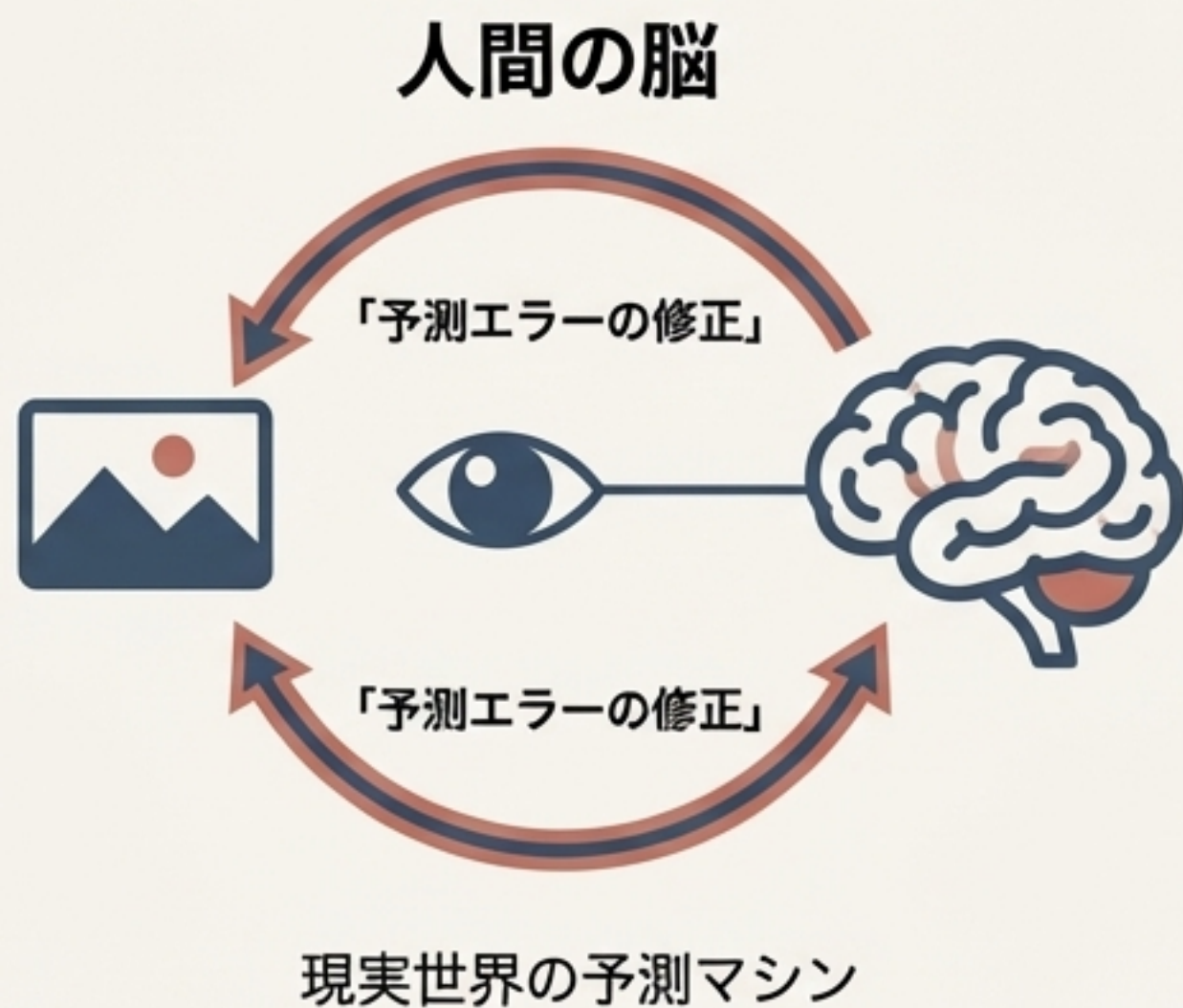
- これはAIの頭の中に、チェス盤や駒の動きを含む「チェスというゲームの世界そのものが構築されたことを示唆しています。



AIの学習プロセスは、人間の脳と不気味なほど似ている

神経科学の有力な理論では、人間の脳も「予測マシン」だとされています。

私たちは常に現実世界を予測し、その予測と現実のズレ（予測エラー）を修正することで世界を認識しています。



しかし、この「AI世界モデル」説に異議を唱える声がある

LLMが内部に世界をシミュレートしているという考えは、誰もが納得しているわけではありません。

AI界の「ゴッドファーザー」の一人とされる人物から、強烈な「待った」がかかっています。

彼の主張は、現在のLLMが持つワールドモデルは根本的に欠陥だらけで、脆弱“弱すぎる”というものです。

現実世界の予測マシン

テキスト世界の予測マシン

批判の核心： 「テキストだけの学習」 がもたらす致命的な弱点

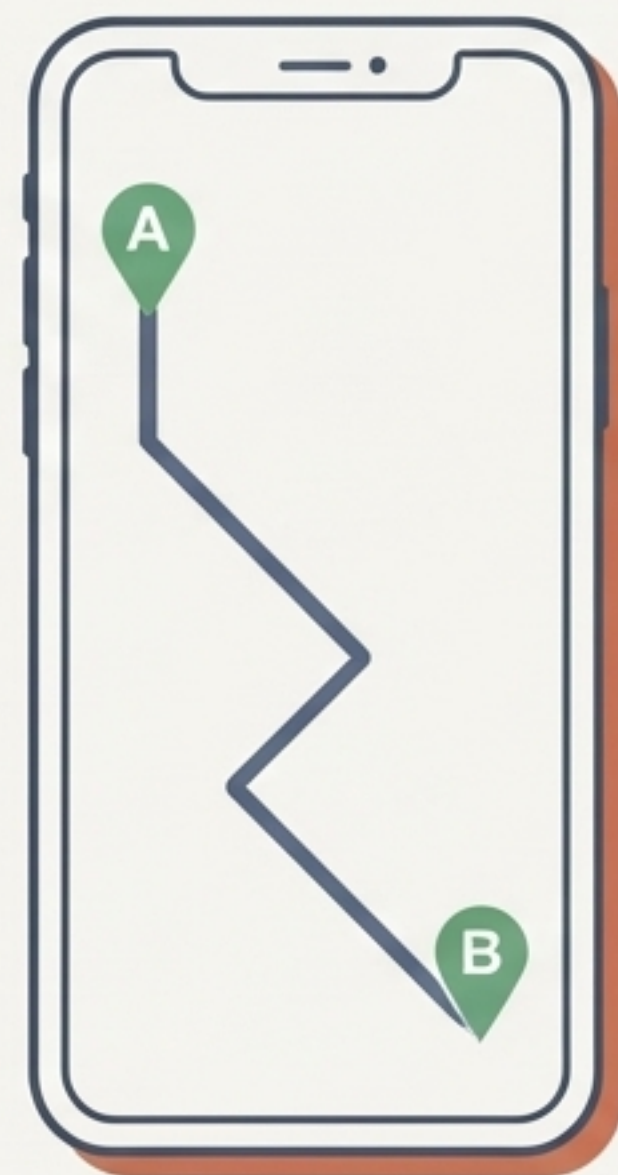
The Argument

- テキストデータのみで学習したモデルは、現実世界を本当に理解しているとは言えない。
- それらは脆弱で、少しでも状況が変わると簡単に破綻してしまう。

The Reason

動画や画像のような、よりリッチな情報から世界を学んでいないため、モデルの理解が表面的で脆いものになってしまう。

「地図アプリが、たった1本の通行止めでパニックに陥るように。」



▶ Before



▶ After

「LLMは行き止まりだ」

この考えに基づき、彼は現在のLLMは3～5年のうちに次世代の技術に取って代わられると予測している。

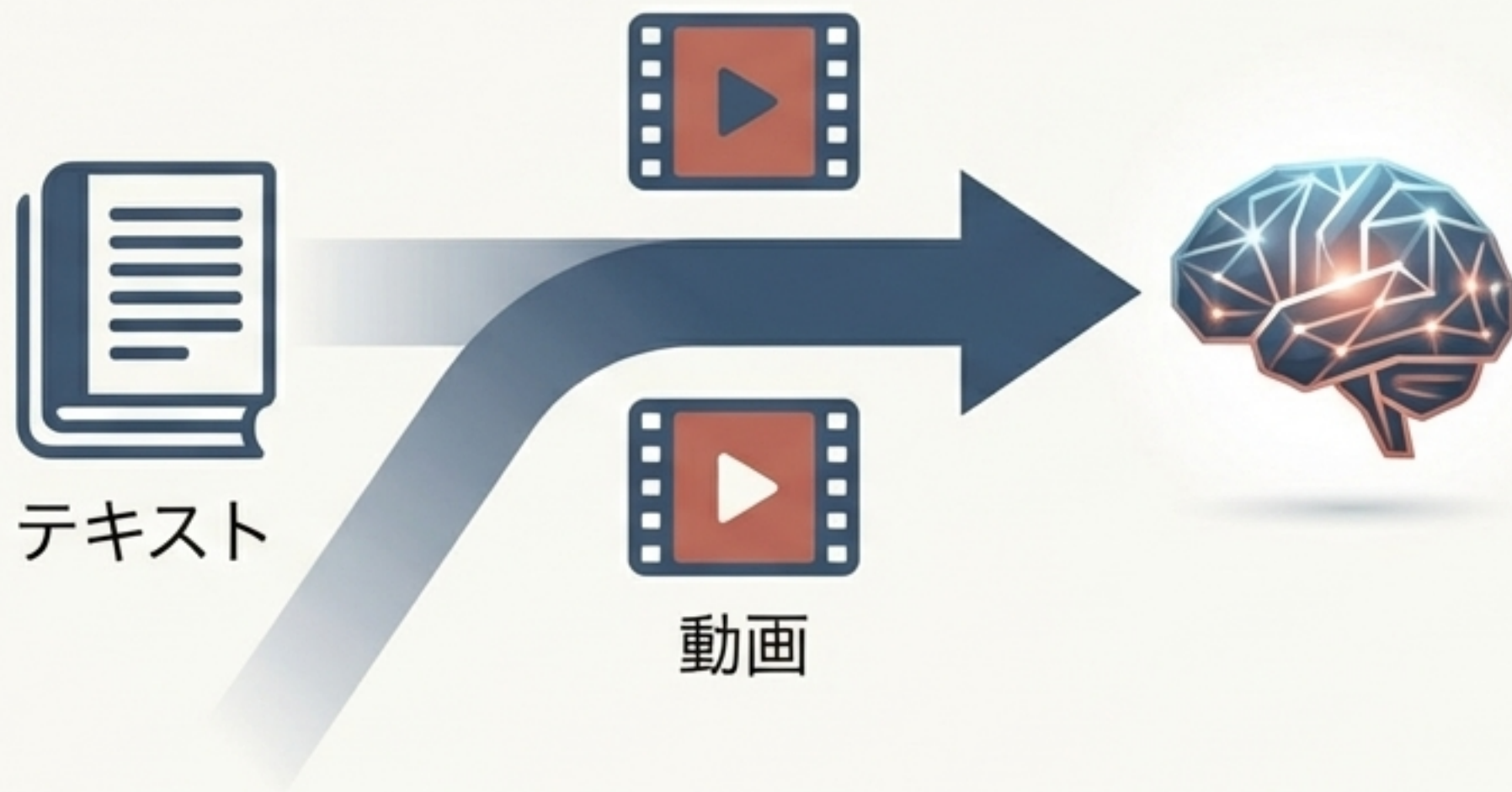
LLMが未来でないなら、 その先には何が待っているのか？

The Proposed Path

- テキスト中心の世界からの脱却。
- 動画のような視覚データから、直接世界の仕組みを学習するモデルへの移行。

The Goal

- 言語を介さず、観察によって世界の物理的・因果的ルールを理解するAIを構築すること。

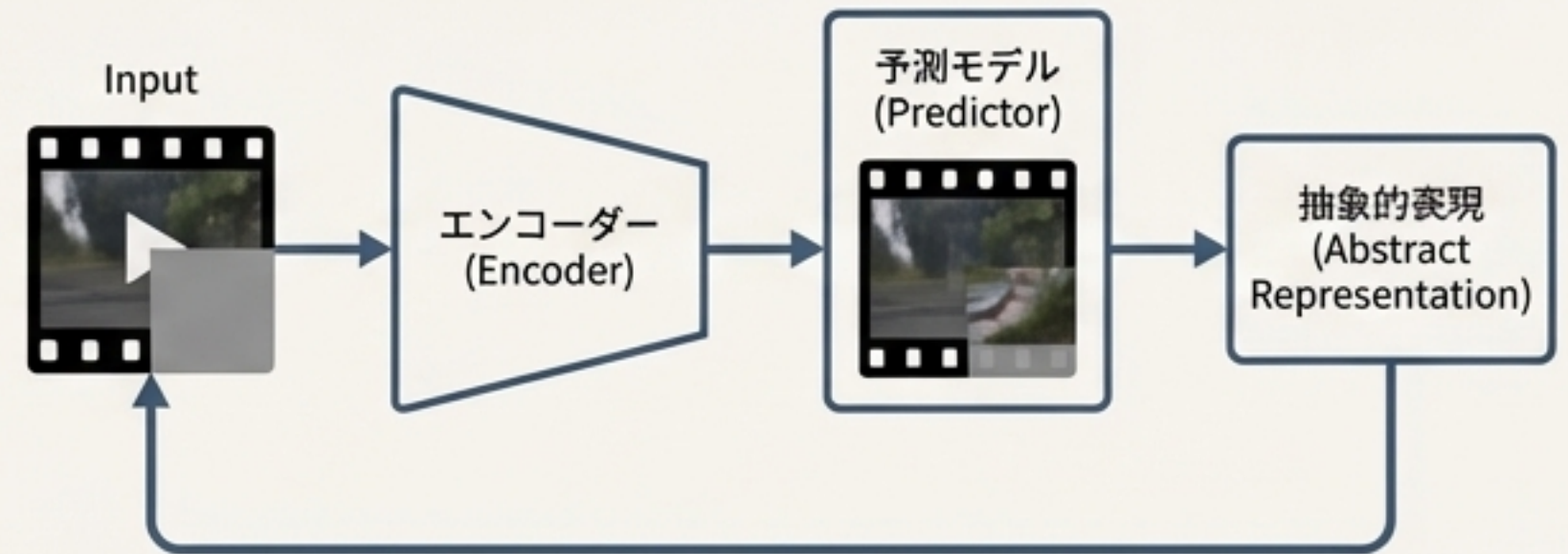


次世代アーキテクチャ「JEPA」が拓く真の世界モデル

批判者たちが提唱する未来のアーキテクチャの一つが「JEPA (Joint Embedding Predictive Architecture)」です。

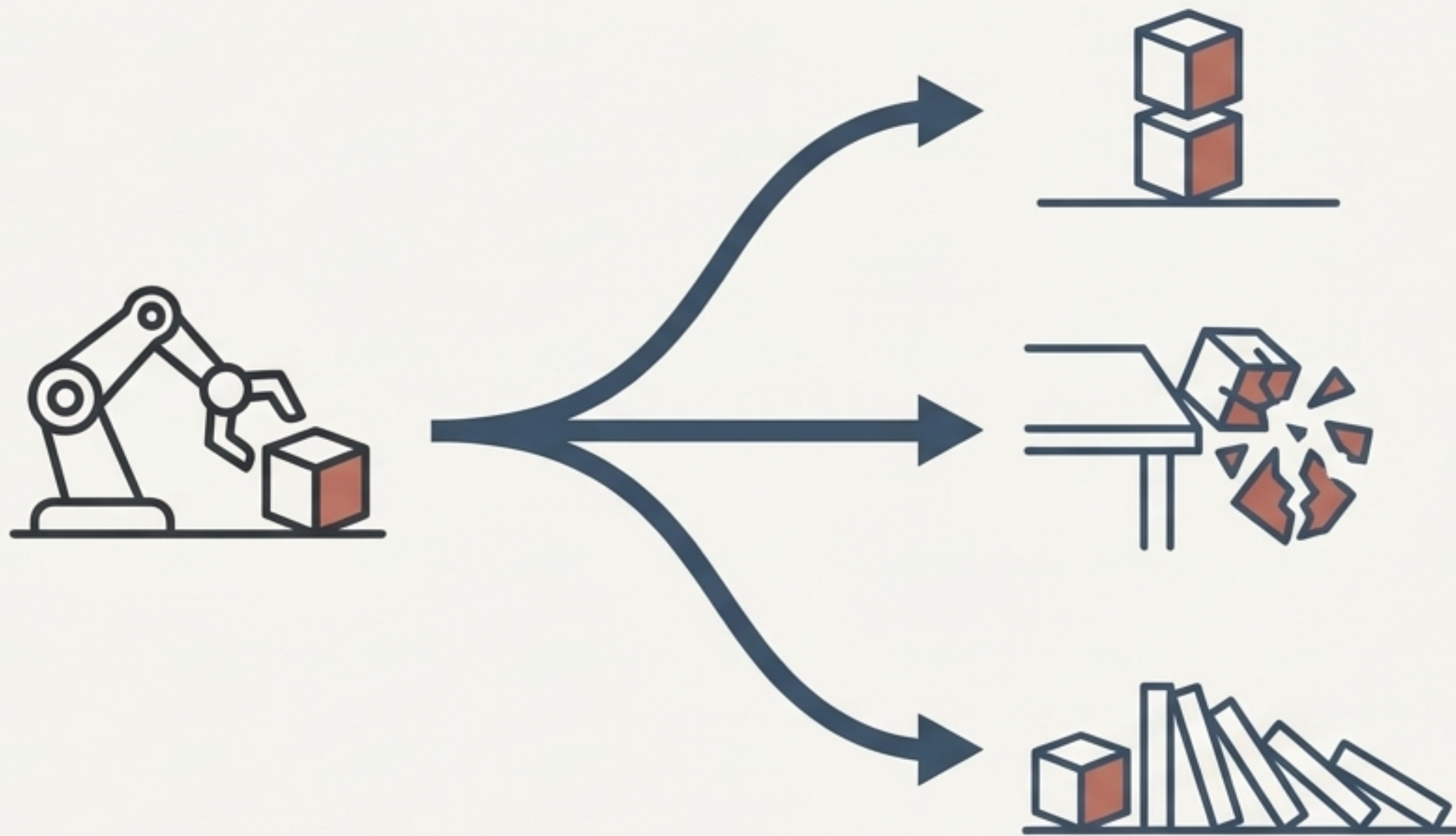
JEPAは、動画などの高次元データから、世界の抽象的な表現を学習することを目的としています。

これこそが、言語に依存しない「真の世界モデル」の始まりだと考えられています。



究極のビジョン：頭の中で「もしも」をシミュレーションするAI

人間のように、「もしこれをやったら、世界はどう反応するだろう？」という問いに答えられるAI。

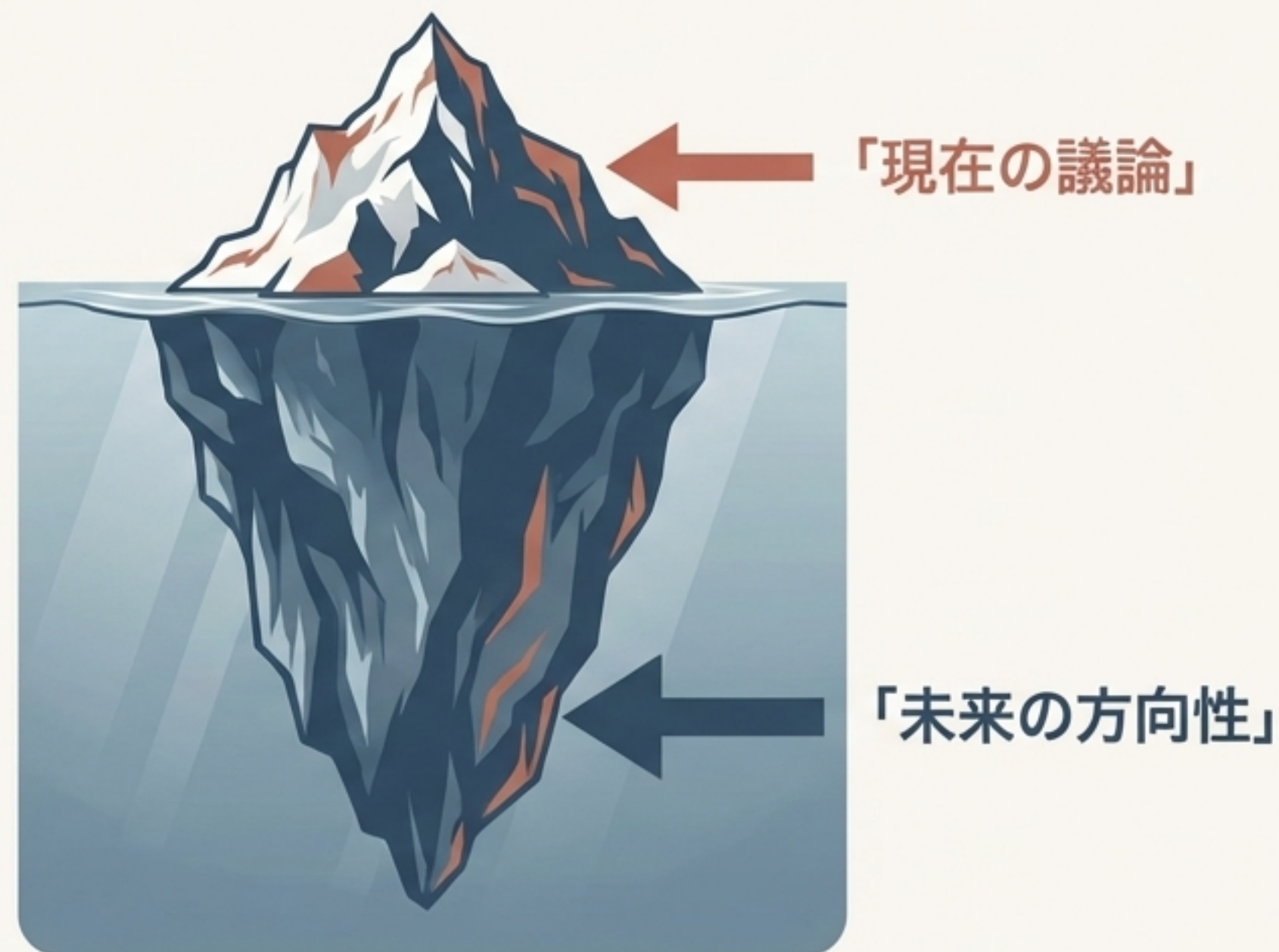


実際に行動する前に、頭の中で無数の「もしも」のシナリオをシミュレートし、最適な行動を計画する能力。

全ては水面下で起きている巨大な変化の兆候

AIを単に「次の単語を予測する賢い機械」と片付けるのは、氷山の一角しか見ていないのと同じです。

現在のLLMに関する議論そのものが、より強力な「ワールドモデル」を構築するという、AI開発の真の目的地へ向かう過程なのです。



私たちは、歴史の転換点に立っている

私たちは、テキストベースベースLLM時代の「頂点」にいますでしょうか？

それとも、これはワールドモデルという、さらに大きな時代の「ほんの幕開け」に過ぎないのでしょうか？

