Introduction to world model

世界模型（World Models）作為人工智慧領域的核心概念，正在重新定義機器如何理解、預測和與環境互動。本報告深入探討了世界模型的完整發展歷程，從1990年代的理論基礎到2025年的商業應用，為讀者提供這一前沿領域的全面視角。

歷史起源與理論基礎

Dyna算法：世界模型的奠基之作

世界模型的概念可以追溯到1990年Richard S. Sutton提出的Dyna算法。Sutton在GTE實驗室工作期間，開發了這一開創性的模型化強化學習（Model-Based Reinforcement Learning, MBRL）架構。Dyna算法的核心創新在於將三個關鍵組件整合為統一框架：

試錯學習（Trial-and-Error Learning）：智慧體通過與真實環境互動來學習最優策略。 這⼀過程基 於強化學習的基本原則，智慧體根據獲得的獎勵信號調整其⾏為。

世界模型學習（World Model Learning）：系統同時學習環境動態的內部模型， 該模型能夠預測在

給定狀態下執⾏特定動作的結果。這個預測模型成為後續規劃的基礎。

規劃機制（Planning）：利⽤學習到的世界模型， 智慧體可以在不與真實環境互動的情況下， 通過 內部模擬來評估不同動作序列的潛在結果，從⽽制定更優的策略。

Dyna算法的理論基礎建立在動態規劃（Dynamic Programming）理論之上，特別是與時序差分學習（Temporal-Difference Learning）和強化學習的緊密聯繫。這種整合方法使得智慧體能夠同時從真實經驗和模擬經驗中學習，大幅提高了樣本效率。

Sutton在其理論框架中提出了\*\*放鬆規劃（Relaxation Planning）\*\*的概念，這是一種增量式規劃方法，可以隨時中斷和恢復。與傳統的深度搜索不同，放鬆規劃通過一系列淺層搜索（通常只有一層深度）來逐步改進策略，最終達到與任意深度傳統搜索相同的效果。

早期發展與挑戰

在Dyna算法之後的十餘年間，模型化強化學習領域面臨著可擴展性問題。早期的方法主要依賴表格化表示（Tabular Representations）或簡單的函數近似，這限制了它們在高維狀態空間中的應用能力。儘管如此，這一時期的研究為後續的發展奠定了重要的理論基礎。

研究者們逐漸認識到，有效的世界模型需要能夠處理高維觀察（如圖像）、複雜動態系統，以及不確定性建模等挑戰。這些認識為後來的深度學習時代世界模型研究指明了方向。

深度學習時代的突破

Ha & Schmidhuber的開創性⼯作

2018年標誌著世界模型研究的重大轉折點。\*\*David Ha和Jürgen Schmidhuber發表的論文「World Models」\*\*不僅重新定義了世界模型的概念，更提供了在高維觀察環境中實現世界模型的具體技術路徑。

技術架構詳解

Ha和Schmidhuber的世界模型框架包含三個核心組件，這一設計深受人類認知系統的啟發：

視覺組件（Vision Component, V）：採⽤變分⾃動編碼器（Variational Autoencoder, VAE）架構， 負責將⾼維觀察（64×64×3 RGB圖像）壓縮為低維潛在表⽰（32維潛在向量z）。VAE的使⽤確保了潛在空間的連續性和可解釋性，同時引入的隨機性有助於提⾼模型的泛化能⼒。

記憶組件（ Memory Component, M）：使用混合密度網絡-循環神經網絡（Mixture Density Network-RNN, MDN-RNN）來建模環境的時序動態。該組件預測下一個時間步的潛在向量分布p(z\_{t+1}|z\_t, a\_t, h\_t)，其中h\_t是RNN的隱藏狀態。混合高斯分布的使用使模型能夠處理環境中的隨機事件和多模態結果。

控制器組件（Controller Component, C）：採用極簡的線性模型設計，僅包含867個參數。控制器接收來自視覺組件的當前觀察編碼和記憶組件的隱藏狀態，輸出相應的動作。這種設計哲學將模型的複雜性集中在世界模型部分，而保持策略組件的簡潔性。

關鍵創新與貢獻

夢境訓練（Dream Training）：Ha和Schmidhuber證明了智慧體可以完全在由其世界模型⽣成的「夢境環境」中進⾏訓練，並將學習到的策略成功遷 移到真實環境中。在VizDoom的「Take Cover」任務中，智慧體在虛擬環境中學習躲避火球，並在真實環境中表現出⾊。

溫度參數調節：通過調節MDN-RNN的溫度參數， 研究者可以控制⽣成環境的不確定性⽔平。 較⾼的溫度使虛擬環境更加困難和不確定， 從⽽防⽌智 慧體利⽤世界模型的缺陷。實驗顯⽰，在溫度為1.15的設定下訓練的智慧體在實際環境中達到了1092±556的最佳成績。

進化策略優化：控制器的訓練採⽤協⽅差矩陣適應進化策略（CMA-ES）⽽非梯度下降， 這種選擇使得系統可以避免傳統強化學習中的信⽤分配問 題，同時便於並⾏化實現。

PlaNet：向潛在動態的演進

2019年， Danijar Hafner等⼈提出的Deep Planning Network (PlaNet)進⼀步推進了世界模型的發展。 PlaNet的核⼼創新在於引入了潜在動態模型（Latent Dynamics Model）， 該模型結合了確定性和隨機性轉換組件。

PlaNet提出的潜在過度射擊（Latent Overshooting）⽬標函數是⼀個重要的技術創新。 這個多步驟變分推理⽬標確保 模型能夠準確預測多個時間步的未來狀態，這對於有效規劃⾄關重要。

與Ha和Schmidhuber的⼯作相比，PlaNet在以下幾個⽅⾯實現了顯著改進：

更精細的抽象表⽰：PlaNet將環境抽象為關鍵特徵，使得規劃過程更加⾼效

純模型化⽅法：完全依賴學習到的動態模型進⾏決策，不需要與真實環境的額外互動

連續控制優化：特別針對需要精確連續控制的任務進⾏了優化

Dreamer系列：世界模型的成熟化

Dreamer：通過潜在想像學習⾏為

2020年，Hafner團隊發布了Dreamer算法，這標誌著世界模型從概念驗證向實⽤系統的重要轉變。 Dreamer的核⼼思想是「通過潜在想像學習⾏為」（Learning Behaviors by Latent Imagination）。

Dreamer整合了PlaNet的潜在動態建模與強化學習， 使智慧體能夠在潜在空間中模擬詳細的軌跡並 基於想像的場景優化策略。 與傳統⽅法相比， Dreamer避免了在真實環境中的試錯學習， ⽽是依靠內部模擬來改進策略。

DreamerV2：離散世界模型的突破

2021年的DreamerV2實現了兩個關鍵技術突破， 使得世界模型⾸次在Atari環境中超越了Rainbow等 頂級模型化⾃由算法：

離散變分⾃動編碼器（ Discrete VAE）： 放 棄 了 傳 統 VAE中 單 峰 正 規 分 布 的 強 假 設 ， 改 ⽤ OneHotCategoricalDistribution。這種設計能夠更好地表⽰系統的隨機轉換，特別適⽤於具有離散狀態轉換的環境。

KL平衡（KL Balancing）：這⼀技術創新解決了變分學習中的訓練不穩定問題，確保了模型訓練的 收斂性。

DreamerV3：多樣化領域的掌控

2023年發布的DreamerV3代表了世界模型發展的重要⾥程碑， 實現了以下 突破性成就：

統⼀超參數框架：DreamerV3能夠在不同環境域中使⽤相同的超參數配置實現

異性能， 這證明了其作為通⽤算法的潛⼒。 實驗涵蓋了從連續控制到離散動 作、從低維輸入到⾼維圖像的各種環境。

Minecraft鑽⽯收集：DreamerV3成為⾸個能夠在Minecraft中從零開始收集鑽⽯

的算法，這⼀成就展⽰了世界模型在複雜、開放式環境中的能⼒。

Symlog預測：引入Symlog函數來處理不同規模的獎勵和價值函數， 解決了訓練中的數值不穩定問 題。 Symlog變換公式為：symlog(x) = sign(x)ln(|x|+1)， 其逆變換為symexp(x) = sign(x)(exp(|x|)-1)

模化效益：DreamerV3證明了模型參數規模與性能的正相關關係， 從8M到 200M參數的模型顯⽰了⼀致的性能提升趨勢。

商業化應⽤與突破

SORA：視頻⽣成作為世界模擬器

2024年，OpenAI發布的Sora代表了世界模型在視頻⽣成領域的重⼤突 破。 Sora的技術報告明確提出了「視頻⽣成模型作為世界模擬器」的 概念，這⼀觀點重新定義了世界模型的應⽤範疇。

技術架構創新：Sora採⽤Transformer架構操作視頻和圖像的時空補

丁， 能夠處理可變持續時間、 解析度和長寬比的視覺數據。 這種統⼀ 表⽰⽅法實現了⼤規模⽣成模型的⾼效訓練。

新興模擬能⼒：Sora展現了多項令⼈驚嘆的新興能⼒：

3D⼀致性：能夠⽣成具有動態攝像機運動的視頻，當攝像機移動和旋轉時，⼈物和場景元素在三 維空間中保持⼀致

長程連貫性和物體永續性：能夠在長視頻中維持時間⼀致性，即使物體被遮擋或離開畫⾯也能保 持其特性

世界互動性：能夠模擬影響世界狀態的動作，如畫家在畫布上留下持續的筆觸

數位世界模擬：能夠模擬⼈⼯過程， 包括視頻遊戲， 可同時控制 Minecraft中的玩家並⾼保真度渲染世界及其動態

Google DeepMind的Genie系列

Genie 2的發布標誌著互動式世界⽣成的重要進展。與Sora專注 於視頻品質不同，Genie 2引入了⽤⼾驅動的世界操作新維度。

Genie 2採⽤⾃回歸潜在擴散模型架構， 在⼤型視頻數據集上訓練，逐幀⽣成視頻。其處理流程包括：⾃動編碼器將視頻幀壓縮到潜在空間；基於Transformer的⾃回歸模型根據先前幀和智慧體動 作預測下⼀個潜在幀；潜在擴散過程精煉⽣成真實的視頻幀；解碼潜在表⽰為視覺幀。

2025年推出的Genie 3實現了實時互動能⼒， 同時在⼀致性和真實性⽅⾯相比Genie 2有顯著改進。 這標誌著世界模型從離線⽣成向實時互動應⽤的重要轉變。

NVIDIA Cosmos：物理AI的世界基礎模型

NVIDIA Cosmos代表了世界模型商業化應⽤的最新⾼度。作為專為物理AI開發構建的平台，Cosmos提供了最先進的⽣成世界基礎模型（World Foundation Models, WFMs）、 ⾼級分詞器、 護欄和加速數據處理管道。

技術規模與能⼒：Cosmos WFMs 基於9,000兆tokens和2,000萬⼩時的機器⼈和駕駛數據訓練， 涵蓋現實世界的幾乎所有場景。這種⼤規模訓練使模型能夠⽣成物理基礎的合成數據，⽤於訓練和評估現有模型。

開放模型授權：Cosmos採⽤開放模型授權⽅式，使5個⼈和⼩型初創企業也能夠利⽤⾼級物理AI技術。這 ⼀策略有望加速整個⽣態系統的發展。

產業合作夥伴：包括1X、 Agile Robots、 Agility、 Figure AI、 Foretellix、 Uber、 Waabi、 XPENG等在內的全球物理AI領導企業已成為Cosmos的⾸批採⽤者。

Meta的導航世界模型

Meta的導航世界模型（Navigation World Model, NWM）專注於視覺運動能⼒智慧體的導航任務。 NWM採⽤條件擴散Transformer（ Conditional Diﬀusion Transformer, CDiT）架構，在⼈類和機器⼈ 智慧體的⾃我中⼼視頻多樣化集合上訓練，規模達到10億參數。

NWM的核⼼優勢在於其動態約束整合能⼒。 與具有固定⾏為的監督導航策略不同， NWM能夠 在規劃過程中動態整合約束條件，這使其成為下⼀代導航系統的靈活⽽強⼤的⼯具。

在未知環境中， NWM利⽤其學習到的視覺先 驗從單⼀輸入圖像想像軌跡，展⽰了強⼤的泛 化能⼒。實驗證明其在從零開始規劃軌跡或對 外部策略採樣的軌跡進⾏排名⽅⾯都表現出⾊。

當前挑戰與限制

技術挑戰

不確定性建模與部分可觀察性：當前的世界模型在處理不確定性和部分可觀察環境⽅⾯仍⾯臨重⼤挑戰。 如Reddit討論中提到，建模不確定性和部分可觀察性是當前⾯臨的主要技術難題之⼀。

可遷移表⽰學習：創建能夠跨不同任務和環境遷移的表⽰是另⼀個關鍵挑戰。Ha和Schmidhuber在原始論⽂中已經指出， VAE可能編碼對任務不相關的觀察部分，這限制了模型在新任務上的重⽤能⼒。

真實性與抽象性的平衡：如何在內部模擬中實現真實性與抽象性的適當平衡是⼀個持續的研究問題。過於真實的模型可能導致計算複雜性過⾼， ⽽過於抽象的模型則可能遺漏關鍵資訊。

理解與解釋性問題

缺乏真正理解：如VE3分析指出，⼤型語⾔模型在世界建模⽅⾯⾯臨的⼀個根本問題是缺乏真正的理解。它們基於數據模式⽽非概念的真正理解來⽣成答案，這可能導致錯誤，特別是在需要深度理解或推理的場景中。

⿊盒性質：隨著模型變得越來越複雜，世界模型的可解釋性成為⼀個重要問題。如Verity AI指出，未來的AI世界模型可能變得無法解釋， 因為其決策過程將類似於⼈類直覺， ⽽我們也無法完全解 釋⼈類直覺。

靜態知識基礎：當前的世界模型主要依賴於訓練數據，無法像⼈類那樣持續學習和適應。這種限制使得模型難以構建強健且無偏⾒的世界模型，特別是在動態或不熟悉的環境中。

可擴展性與實⽤性限制

計算資源需求：現代世界模型，特別是像NVIDIA Cosmos這樣的⼤規模模型，需要巨⼤的計算資源進⾏訓練和部署。這限制了其在資源受限環境中的應⽤。

控制環境外的脆弱性：儘管世界模型在受控環境中表現出⾊，但在開放式、不可預測的真實世界環境中的表現仍有待驗證。

資料依賴性：世界模型的品質嚴重依賴於訓練資料的品質和多樣性。如果訓練資料缺乏多樣性或存在偏⾒，模型的世界理解將受到相應限制。

未來發展⽅向

邁向⼈⼯通⽤智慧

世界模型作為AGI的基⽯：Yann LeCun等研究者認為，世界模型是實現⼈⼯通⽤智慧的關鍵組件。LeCun提出了包括⾃監督學習、世界模型和認知架構在內的三⼤關鍵技術，這些技術的整合被視為AGI發展的必要條件。

體現智慧的發展：如體現AGI綜述所指出， 體現智慧（Embodied AI）與 世界模型的結合正在推動AI系統向更 類⼈的互動能⼒發展。 體現AGI被定 義為能夠在真實世界任務中表現出⼈ 類⽔準能⼒的體現AI形式。

技術融合與創新

多模態整合：未來的世界模型將更好地整合視覺、聽覺、觸覺等多種感知模式，形成更全⾯的世界表⽰。NVIDIA Cosmos已經在這個⽅向上邁出了重要步伐。

因果推理整合：如TuringPost分析指出， 世界模型未來發展的⼀個 重要⽅向是整合因果AI。這種整合對於實現真正的AGI⾄關重要。

層次化與混合表⽰：未來的世界模型可能採⽤層次化、 多層次和連續/離 散混合表⽰的架構，以更好地處理複 雜的真實世界場景。

應⽤領域擴展

⾃動駕駛⾰命：世界模型在⾃動駕駛領域的應⽤正在快速發展。 從 GAIA-1 到 DriveDreamer ， 再 到 MUVO等專⾨針對駕駛場景的世界 模型正在推動⾃動駕駛技術的進步。

機器⼈技術突破：如Towards AI分析所述，世界模型正在推動機器⼈技術向更智慧、 更⾃主的⽅向發展。機器⼈將能夠預測、規劃和適應複雜環境，實現真正的⾃主性。

科學發現與探索：世界模型在科學發現中的應⽤潛⼒巨⼤。通過模擬複雜系統，研究者可以探索新的科學現象並測試理論假說。

結論與展望

世界模型的發展歷程展現了⼈⼯智慧領域從理論探索到實際應⽤的完整演進過程。 從1990年Sutton 的Dyna算法奠定的理論基礎， 到2018年Ha和Schmidhuber的開創性框架， 再到近年來Dreamer系列 的技術突破和商業化應⽤的蓬勃發展，世界模型已經成為現代AI系統的核⼼組件。

當前的發展趨勢顯⽰， 世界模型正在朝著更⼤規模、 更多模態、 更⾼智慧的⽅向演進。 NVIDIA Cosmos、 OpenAI Sora、 Google Genie等商業化產品的成功展⽰了世界模型的巨⼤應⽤潜⼒。 同 時，導航世界模型、遊戲世界⽣成等專⾨化應⽤也證明了這⼀技術的廣泛適⽤性。

然⽽， 世界模型的發展仍⾯臨諸多挑戰。 不確定性建模、 真實世界泛化、 計算效率、 可解釋性等問 題需要持續的研究投入。 特別是在邁向⼈⼯通⽤智慧的路徑上， 世界模型需要與其他AI技術深度融 合，形成更加完整的智慧系統架構。

12 展望未來， 世界模型有望成為實現真正智慧機器的關鍵技術。通過持續的技術創新和跨領域合作， 我們正在逐步構建能夠真正理解和互動世界的AI系統。這不僅將推動⼈⼯智慧技術的⾰命性發展， 更將為⼈類社會帶來前所未有的變⾰機遇。

在這個充滿挑戰與機遇的時代，世界模型研究正站在歷史的轉折點上。從學術研究到商業應⽤， 從 理論探索到實際部署， 世界模型正在重新定義智慧機器與物理世界互動的⽅式。 這⼀發展趨勢預⽰ 著我們正在進入⼀個AI系統不僅能夠感知世界，更能夠真正理解和預測世界的新時代。

發展時間軸：

世界模型發展歷程時間軸

1990年：

* Dyna 演算法的誕生： Richard S. Sutton 於 GTE 實驗室工作期間，提出了 Dyna 演算法，奠定了世界模型（World Models）的理論基礎。Dyna 演算法整合了試錯學習、世界模型學習和規劃機制。Sutton 提出了「放鬆規劃（Relaxation Planning）」的概念。

1990年代至2000年代：

* 早期發展與挑戰： 模型化強化學習領域面臨可擴展性問題，主要依賴表格化表示或簡單函數近似，限制了其在高維狀態空間的應用。研究者逐漸意識到世界模型需要處理高維觀察、複雜動態系統和不確定性建模等挑戰。

2018年：

* 「World Models」論文發表： David Ha 和 Jürgen Schmidhuber 發表了開創性論文「World Models」，重新定義了世界模型概念，並提供了在高維觀察環境中實現世界模型的具體技術路徑。其框架包含視覺組件（VAE）、記憶組件（MDN-RNN）和控制器組件，並引入了「夢境訓練」和「進化策略優化」。

2019年：

* PlaNet 的提出： Danijar Hafner 等人提出了 Deep Planning Network (PlaNet)，引入了潛在動態模型，並提出了「潛在過度射擊（Latent Overshooting）」目標函數，改進了多時間步預測能力。

2020年：

* Dreamer 演算法發布： Hafner 團隊發布了 Dreamer 演算法，標誌著世界模型從概念驗證向實用系統的重要轉變。Dreamer 實現了「通過潛在想像學習行為」，在潛在空間中模擬軌跡並優化策略。

2021年：

* DreamerV2 的突破： DreamerV2 實現了兩項關鍵技術突破：採用離散變分自動編碼器（Discrete VAE）和引入 KL 平衡（KL Balancing），首次在 Atari 環境中超越了 Rainbow 等頂級模型化自由演算法。

2023年：

* DreamerV3 的發布： DreamerV3 實現了統一超參數框架，能在不同環境域中實現超高性能；成為首個能在 Minecraft 中從零開始收集鑽石的演算法，展現了在複雜開放式環境中的能力；引入 Symlog 預測來處理不同規模的獎勵和價值函數；並證明了模型參數規模與性能的正相關關係。

2024年：

* OpenAI Sora 發布： OpenAI 發布了 Sora，代表了世界模型在影片生成領域的重大突破，明確提出了「影片生成模型作為世界模擬器」的概念，展現了 3D 一致性、長程連貫性、物體永續性和世界互動性等新興模擬能力。

2025年（預計/報告涵蓋時間點）：

* Google DeepMind 的 Genie 2 & 3： Genie 2 發布，引入了用戶驅動的世界操作新維度，採用自回歸潛在擴散模型架構。Genie 3 實現了實時互動能力，並在一致性和真實性方面有顯著改進。
* NVIDIA Cosmos 發布： NVIDIA Cosmos 作為專為物理 AI 開發構建的平台，提供了最先進的生成世界基礎模型（WFMs），基於海量機器人與駕駛數據訓練，並採用開放模型授權。
* Meta 的導航世界模型 (NWM) 發展： Meta 專注於視覺運動能力智慧體的導航任務，採用條件擴散 Transformer (CDiT) 架構，並具備動態約束整合能力。

世界模型核心人物列表

1. Richard S. Sutton (理查德·S·薩頓)

* 簡介： 1990 年在 GTE 實驗室工作期間，提出了 Dyna 演算法，奠定了世界模型和模型化強化學習的理論基礎。他將試錯學習、世界模型學習和規劃機制整合到統一框架中，並引入了「放鬆規劃」概念。他是強化學習領域的開創性人物。

1. David Ha (大衛·哈)

* 簡介： 2018 年與 Jürgen Schmidhuber 共同發表了開創性論文「World Models」，重新定義了世界模型的概念，並提供了在高維觀察環境中實現世界模型的具體技術路徑。其框架包含視覺組件、記憶組件和控制器組件，並引入了「夢境訓練」等創新。

1. Jürgen Schmidhuber (尤爾根·施密德胡伯)

* 簡介： 2018 年與 David Ha 共同發表了關於「World Models」的開創性論文。他是深度學習領域的重要先驅，尤其在遞歸神經網路（RNN）和長短期記憶（LSTM）網路方面貢獻卓著，這些技術是 Ha & Schmidhuber 世界模型記憶組件的基礎。

1. Danijar Hafner (丹尼亞爾·哈夫納)

* 簡介： 2019 年提出了 PlaNet (Deep Planning Network)，進一步推進了世界模型的發展，引入了潛在動態模型。其團隊在 2020 年發布了 Dreamer 演算法，隨後在 2021 年和 2023 年相繼發布了 DreamerV2 和 DreamerV3，極大地推動了世界模型的成熟化和實用化，並在多個任務中取得了突破性成就。

1. Yann LeCun (楊·勒昆)

* 簡介： 報告中提到，Yann LeCun 等研究者認為世界模型是實現人工通用智慧 (AGI) 的關鍵組件。他是深度學習領域的奠基人之一，特別是在卷積神經網路方面貢獻巨大，目前擔任 Meta AI 的首席人工智慧科學家。他提出的 AGI 發展關鍵技術包括自監督學習、世界模型和認知架構。