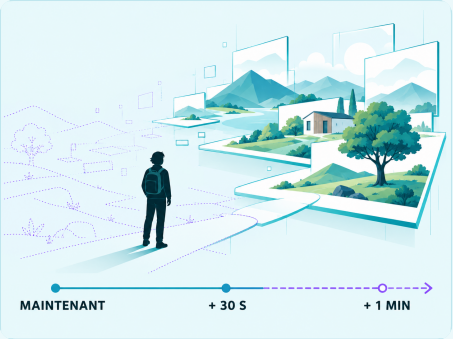
World Models
simuler un monde qui répond aux actions, entre usages réels et horizon 2030



GENIE 3, COSMOS 3, MARBLE, GAIA-4, JEPA : TROIS DÉFINITIONS, DES USAGES DÉJÀ RÉELS, UN HORIZON INCERTAIN (RELEVÉ AU 18 SEPTEMBRE 2026)

Les modèles de monde apprennent à prédire comment le monde réagit à une action. En 2026, ils servent déjà à tester des voitures autonomes et à fabriquer des données pour les robots, mais un monde généré ne tient qu'une à deux minutes de cohérence : l'horizon 2030 se joue sur la mémoire, la physique et la définition même du terme. Cette page distingue trois sens du mot, passe en revue les acteurs et les usages avérés, nomme les verrous, puis classe chaque promesse à venir selon son degré de certitude.

TEMPS RÉEL | SIMULATION | ROBOTIQUE & CONDUITE | RECHERCHE (JEPA)

» REPÈRES SEPTEMBRE 2026

~1 min
COHÉRENCE D'UN MONDE GÉNÉRÉ

Genie 3 garde en mémoire environ une minute de scène ; les sessions de Project Genie sont limitées à 60 s, en 720p et 24 images/s (selon la presse).

~200 M de miles
CONDUITE RÉELLE DE WAYMO, CONTRE DES MILLIARDS SIMULÉS

Chiffre de février 2026 ; le Waymo World Model, qui simule le reste, est bâti sur Genie 3.

~2 Md$
LEVÉS EN 2026 PAR DEUX START-UP DE MODÈLES DE MONDE

AMI Labs, 1,03 Md$ (mars) ; World Labs, 1 Md$ (annoncé en février).

4 Md
PARAMÈTRES DE COSMOS 3 EDGE, MODÈLE OUVERT EMBARQUÉ

Sorti le 20 juillet 2026 : actions de robot à 15 Hz sur une carte Jetson Thor ; poids et code publiés.

» SOMMAIRE

01 Un mot, trois paris techniques
02 Qui construit quoi, en septembre 2026
03 Ce qui sert déjà
04 Les verrous : durée, physique, preuve
05 L'horizon : de l'acquis au spéculatif
06 Que faire de tout cela en 2026

DÉFINITIONS / 01

01

Un mot, trois paris techniques

« World model » désigne des objets très différents : un monde que l'on regarde, un monde qu'un programme peut lire, un monde que l'on prédit pour décider. Les distinguer évite la plupart des malentendus.

L'idée n'est pas neuve : en 2018, Ha et Schmidhuber montrent qu'un agent peut apprendre dans un monde qu'il « rêve ». En juin 2026, World Labs a proposé une typologie reprise ici, **renderer, simulator, planner** (lue via la presse, l'essai d'origine n'a pas été relu). Un même produit peut relever de plusieurs sens.

01 RENDU

Le rendu : un monde que l'on regarde

Le modèle produit des images en réponse à des actions (déplacement, regard, commande) : chaque image dépend des précédentes et de l'action reçue.**Exemples :** Genie 3, Runway GWM-1, Odyssey-2.On obtient une vidéo interactive, pas une scène que l'on peut inspecter ni confier à un programme.

02 SIMULATEUR

Le simulateur : un monde qu'un programme peut lire

Le modèle produit un état du monde exploitable : une scène 3D persistante, ou des capteurs (caméras, lidar) cohérents entre eux.**Exemples :** Marble et Atlas (World Labs), Waymo World Model, GAIA-4 (Wayve).Le monde existe en dehors du regard : on peut le télécharger ou y rejouer un scénario.

03 PRÉDICTION

Le modèle prédictif : un monde pour décider

Le modèle prédit dans un espace abstrait (des représentations, pas des pixels) ce qui va se passer, pour choisir une action.**Exemples :** JEPA de Yann LeCun, V-JEPA 2 de Meta. Cosmos 3 (NVIDIA) prédit l'action et sa conséquence visuelle.Les modèles VLA, qui sortent directement des actions, en sont le voisin (page VLA).

» MODÈLE DE MONDE OU MODÈLE VIDÉO : LA DISTINCTION PÉDAGOGIQUE

Un modèle vidéo rend un clip entier en une passe ; un modèle de monde produit le moment suivant à partir de l'action reçue.

01

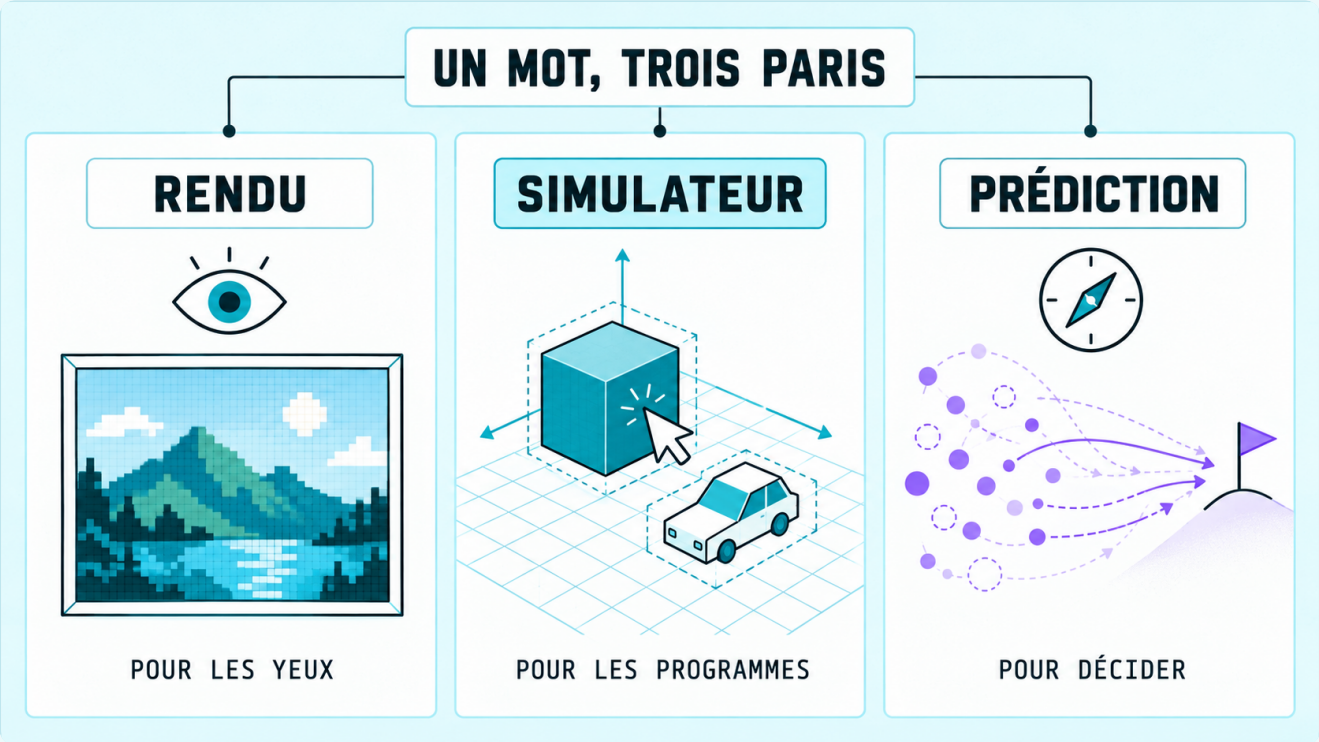
Modèle vidéo (Veo, Sora, Kling…)

- Un clip complet à partir d'un texte ou d'une image, rendu en une passe.
- **Forces :** haute fidélité, 4K, audio synchronisé, plusieurs plans.
- **Limites :** non interactif, aucune notion d'action ni de causalité contrôlée.

02

Modèle de monde (Genie 3, GWM-1, Odyssey-2…)

- Un monde généré image par image, en temps réel, qui répond aux actions.
- **Forces :** interactivité, causalité, utilisable pour entraîner un agent.
- **Limites :** 720p, 24 images/s, une à deux minutes de cohérence, calcul coûteux. La « physique » y est apprise statistiquement, pas simulée.



Un même mot, trois paris techniques : le rendu, le simulateur, le modèle prédictif.

ACTEURS / 02

02

Qui construit quoi, en septembre 2026

Neuf acteurs, rangés par pari technique, puis un zoom sur Project Genie, le monde interactif de Google ouvert aux abonnés.

01 RENDU

Google DeepMind · Genie 3 / Project Genie

Genie 3 (août 2025) : mondes navigables en 720p, 24 images/s, mémoire visuelle d'environ une minute. Project Genie, prototype Google Labs, s'ouvre le 29 janvier 2026 aux abonnés Google AI Ultra (environ 200 $ par mois), puis aux abonnés éligibles du monde entier depuis le 19 mai 2026.

02 RENDU

Runway · GWM-1

11 décembre 2025. Modèle autorégressif construit sur Gen-4.5 : image par image, en temps réel, 720p, jusqu'à environ deux minutes, piloté par des actions (caméra, commandes de robot, voix). Trois déclinaisons : Worlds, Robotics (SDK Python) et Characters. Accès anticipé.

03 RENDU

Odyssey · Odyssey-2

Vidéo interactive en temps réel (720p, 22 images/s), par API depuis janvier 2026 selon l'éditeur. Odyssey-2 Max (21 avril 2026, trois fois plus de paramètres) dépasse deux minutes de génération continue, en bêta privée.

04 SIMULATEUR

Decart · Oasis 3

10 juin 2026. Environnements de conduite photoréalistes, multi-caméras (une avant, deux latérales), en temps réel, vendus par API à 0,02 $ la seconde. TechCrunch relève les limites : la cohérence se dégrade et les véhicules se traversent.

05 SIMULATEUR

World Labs · Marble et Atlas

Marble (12 novembre 2025) génère un monde 3D persistant et téléchargeable (Gaussian splats, maillages), plutôt qu'un monde calculé à la volée ; offres payantes au lancement (prix non revérifiés). Atlas (1er septembre 2026) : jusqu'à une minute de vidéo en 1440p avec contrôle précis de la caméra, reconstruction 3D à partir de photos ; il doit alimenter les futures versions de Marble.

06 SIMULATEUR

NVIDIA · Cosmos 3

31 mai 2026 (GTC Taipei). Modèle ouvert pour la physical AI, licence OpenMDW 1.1 : Super (64 Md de paramètres), Nano (16 Md), Edge (4 Md, 20 juillet). Il génère des données synthétiques et prédit l'action et sa conséquence visuelle ; ce n'est pas un monde que l'utilisateur explore. Une coalition d'éditeurs (Runway, Black Forest Labs, Skild AI…) s'est ralliée.

07 SIMULATEUR

Waymo · World Model

6 février 2026. Bâti sur Genie 3, il génère caméras et lidar cohérents pour simuler des scénarios rares (tornade, inondation, obstacle inhabituel). Waymo comptait alors près de 200 millions de miles autonomes réels, contre des milliards en simulation. Usage interne, non commercialisé.

08 SIMULATEUR

Wayve · GAIA-4

3 août 2026. Simulation « en boucle fermée » : les images et le radar générés changent selon les décisions du conducteur IA, tandis que les autres usagers rejouent leur comportement réel. Sert à valider la conduite autonome.

09 PLANIFICATEUR

AMI Labs et Meta · JEPA

AMI Labs, fondée à Paris par Yann LeCun, a levé 1,03 Md$ en mars 2026 (3,5 Md$ de valorisation avant levée) pour des modèles de type JEPA, qui prédisent dans un espace abstrait plutôt que pixel par pixel ; la direction reconnaît que les produits prendront des années. Chez Meta, V-JEPA 2 (juin 2025, 1,2 Md de paramètres) pilote un bras robotique en zéro-shot après moins de 62 heures de vidéos de robots.

» PROJECT GENIE DE PRÈS

Correction importante : les « promptable events » (modifier le monde par une commande texte en cours d'exploration) sont une capacité du modèle Genie 3 présentée en 2025, absente du prototype public.

01

Ce que le prototype permet

- **World Sketching :** créer un monde à partir de texte et d'images.
- **World Exploration :** se déplacer, le chemin étant généré au fil des pas.
- **World Remixing :** remixer des mondes existants et télécharger des vidéos d'exploration.
- **Street View** (19 mai 2026) : partir d'un lieu réel et lui appliquer un style, pour les lieux situés aux États-Unis seulement.

02 ÉCART

Ce qu'il ne fait pas

Google l'écrit le 29 janvier 2026 : certaines capacités annoncées en août, comme les **promptable events-** , ne sont pas encore incluses dans le prototype. Aucune source consultée n'indique qu'elles aient été ajoutées depuis. Google parle toujours d'un « experimental research prototype ».

03

Limites reconnues par DeepMind

- Espace d'actions restreint ; interactions multi-agents non résolues.
- Lieux réels sans exactitude géographique.
- Texte lisible seulement s'il est fourni dans la description du monde.
- Pas de plusieurs heures d'interaction ; sessions de 60 s selon la presse.

USAGES / 03

Ce qui sert déjà

Quatre usages sont attestés par les éditeurs eux-mêmes. Aucun n'est encore un produit grand public : la valeur se trouve d'abord dans la simulation et les données synthétiques.

Un signal économique : au lendemain du lancement de Project Genie, les actions de plusieurs éditeurs de jeux vidéo et d'Unity ont reculé (Bloomberg, 30 janvier 2026, via Wikipedia) ; deux start-up de modèles de monde ont levé environ un milliard de dollars chacune. Les marchés parient sur la création de mondes, mais les usages avérés sont d'abord industriels.

01 EN PRODUCTION

Valider la conduite autonome

Waymo simule des scénarios rares ; Wayve valide en boucle fermée avec GAIA-4 ; Decart vend depuis juin 2026 un simulateur de conduite multi-caméras par API ; NVIDIA vise le même marché avec Cosmos 3 (partenaire : Li Auto).

02 EN PRODUCTION

Produire des données pour les robots

Cosmos 3 génère des données synthétiques et prédit action et conséquence ; NVIDIA affirme réduire les cycles d'entraînement et d'évaluation de plusieurs mois à quelques jours (affirmation de l'éditeur). Runway propose GWM Robotics avec un SDK Python.

03 RECHERCHE

Entraîner des agents

En novembre 2025, DeepMind a placé son agent SIMA 2 dans des mondes générés par Genie 3 qu'il n'avait jamais vus : l'agent s'y est orienté, a suivi des instructions et s'est amélioré seul. Première démonstration publique d'un « curriculum illimité » d'environnements.

04 PROTOTYPE

Créer des scènes 3D et des mondes

Marble et Atlas transforment texte, photos ou panoramas en scènes 3D éditables ; Project Genie permet d'explorer ses propres mondes. Les offres existent mais restent chères, limitées ou en accès anticipé.

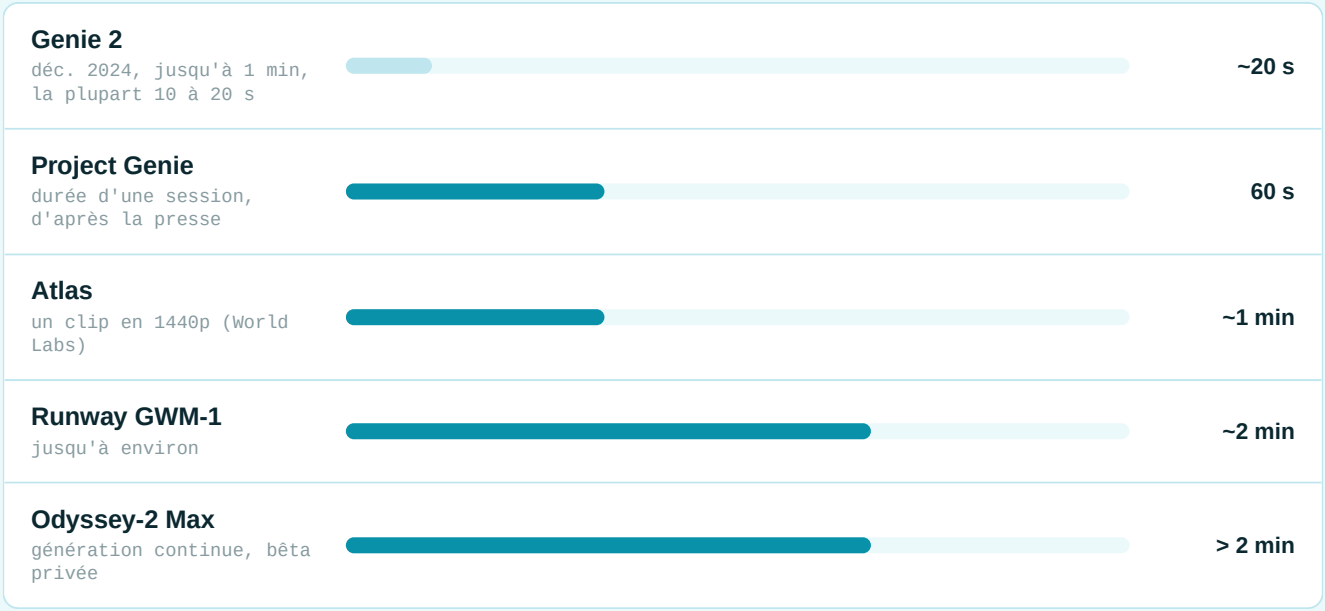
VERROUS / 04

Les verrous : durée, physique, preuve

Ce qui sépare une démonstration d'un outil fiable : la durée de cohérence, une physique apprise sans garantie, le coût de calcul et la difficulté de prouver qu'un simulateur est fidèle.

> **Une à deux minutes : c'est l'ordre de grandeur annoncé par les éditeurs pour la génération continue d'un monde interactif, loin des heures qu'exigerait une simulation de sécurité.**

» COMBIEN DE TEMPS TIENT UN MONDE GÉNÉRÉ ?



Durées annoncées par les éditeurs, non strictement comparables (session, génération continue, clip). Une heure représenterait vingt fois la largeur de ce graphique.

» SIX VERROUS

01
Cohérence dans la durée
Une minute de mémoire visuelle pour Genie 3, une à deux minutes de génération continue ailleurs. La cohérence se dégrade ensuite : objets qui changent, lieux qui dérivent.

02
Une physique apprise, non garantie
Meta a publié des tests de raisonnement physique (IntPhys 2, MVPBench) sur lesquels les modèles, V-JEPA 2 compris, restent proches du hasard pour détecter une violation de la physique. Avec Oasis 3, les véhicules se traversent.

03
Interactions multi-agents
Modéliser fidèlement plusieurs agents indépendants reste, selon DeepMind, un problème de recherche ouvert.

04
Le coût de calcul
Google décrit Genie 3 comme très coûteux à faire tourner. À 0,02 $ la seconde, la simulation de conduite d'Oasis 3 revient à environ 72 $ par heure, par flux.

05
Prouver qu'un simulateur est fidèle
Valider une voiture dans un monde généré suppose de prouver que ce monde ressemble au réel. La simulation en boucle fermée de GAIA-4 s'attaque à cette question ; elle n'est pas résolue.

06
Un accès encore restreint
Abonnement à environ 200 $ par mois pour Project Genie, Street View limité aux États-Unis, bêta privée pour Odyssey-2 Max, accès anticipé pour GWM-1 et Atlas.

HORIZON / 05

L'horizon : de l'acquis au spéculatif

Le chemin parcouru, puis les promesses de 2027 et 2030, chacune classée par degré de certitude.

> **Les jalons jusqu'à septembre 2026 sont des faits datés.** **Les cases 2027 et 2030 sont des anticipations : leur degré de certitude est une appréciation éditoriale de ce rapport, pas une prévision sourcée, sauf mention contraire. Aucune prévision chiffrée de marché fiable n'a été trouvée.**

» FRISE : DE 2018 À 2030, AVEC UN DEGRÉ DE CERTITUDE
Échelle : ACQUIS · PROBABLE · PLAUSIBLE · INCERTAIN · SPÉCULATIF (de moins en moins certain)

2018	**Ha et Schmidhuber, « World Models »** Un agent apprend dans un monde qu'il rêve : l'idée fondatrice.
Juin 2022	**LeCun propose JEPA** Prédire dans un espace abstrait plutôt que pixel par pixel.
2024	**Genie 1 (2D), puis Genie 2 (3D)** Quelques dizaines de secondes de monde jouable.
Juin 2025	**V-JEPA 2 (Meta)** Pilotage d'un bras robotique en zéro-shot.
Août 2025	**Genie 3** 720p, 24 images/s, en temps réel, environ une minute de mémoire.

Nov.-déc. 2025	**Marble, SIMA 2 dans Genie 3, Runway GWM-1** Le 3D persistant, l'agent qui s'entraîne dans un monde généré et le rendu temps réel de Runway.
Janv.-mars 2026	**Project Genie, Waymo World Model, levées** Ouverture aux abonnés Ultra (29 janvier), simulateur de Waymo (6 février), levées de World Labs et d'AMI Labs.
Mai-sept. 2026 AUJOURD'HUI	**Aujourd'hui** Genie + Street View mondial (19 mai), Cosmos 3 (31 mai) et Edge (20 juillet), Oasis 3 (10 juin), GAIA-4 (3 août), Atlas (1er septembre).
2027 PROBABLE	**Un outil standard de validation** Les modèles de monde deviennent courants pour valider la conduite autonome et la robotique : Waymo, Wayve, NVIDIA et Decart s'en servent déjà.
2027 PLAUSIBLE	**Des mondes plus longs et plus accessibles** Accès élargi au-delà de l'offre Ultra et sessions plus longues que 60 s pour le grand public. Google dit vouloir élargir l'accès, sans calendrier.
2027 PLAUSIBLE	**Les premières publications d'AMI Labs** Sous forme de recherche, pas encore de produit.
2030 INCERTAIN	**Les premiers produits d'AMI Labs** Horizon d'environ cinq ans évoqué en mars 2026 (selon le New York Times, cité par Wikipedia).
2030 PLAUSIBLE	**Modèles de monde et modèles d'action convergent** Le robot « imagine » avant d'agir : Cosmos 3 prédit déjà action et conséquence visuelle. Non daté (voir la page VLA).
2030 SPÉCULATIF	**Des heures de cohérence, une simulation acceptée comme preuve de sécurité** Aucune source ne le date : c'est un souhait, pas une trajectoire.

POUR ALLER PLUS LOIN **VLA : de l'image à l'action du robot ›**

REPÈRES / 06

06

Que faire de tout cela en 2026

Quatre réflexes pour garder la tête froide devant des démonstrations spectaculaires.

» À RETENIR

Quatre réflexes devant une démonstration

01 Lire la durée de session, la résolution et le mode d'accès (abonnement Ultra, bêta privée, accès anticipé) avant la vidéo de lancement : ils disent ce que l'outil fait vraiment.

02 Chercher l'usage là où il existe déjà : validation de conduite, données synthétiques pour la robotique. Pour la création grand public, traiter ces outils comme des prototypes.

03 Demander quel sens de « world model » le fournisseur emploie, rendu, simulateur ou modèle prédictif : les trois ne se remplacent pas.

04 Suivre la convergence avec les modèles d'action (page VLA) : c'est là que se jouera le passage de l'imaginé à l'agi.

RÉFÉRENCES / 07

07

Sources et ressources

[01] Google DeepMind : Genie 3, a new frontier for world models (août 2025)
[02] Google : Project Genie (29 janv. 2026, limites du prototype)
[03] Google : Project Genie et Street View (19 mai 2026)
[04] TechCrunch : Genie et Street View (19 mai 2026)
[05] NVIDIA : lancement de Cosmos 3 (31 mai 2026)
[06] NVIDIA Blog : Cosmos 3, licence OpenMDW 1.1
[07] Hugging Face / NVIDIA : Cosmos 3 Edge (20 juil. 2026)
[08] TechCrunch : World Labs lance Marble (12 nov. 2025)
[09] The AI Insider : World Labs lève 1 Md$ (févr. 2026)
[10] World Labs : Atlas (1er sept. 2026)
[11] Waymo : The Waymo World Model (6 févr. 2026)
[12] TechCrunch : le modèle de monde de Decart, Oasis 3 (10 juin 2026)
[13] Decart : site officiel
[14] Wayve : GAIA-4 (3 août 2026)
[15] Runway : GWM-1 (11 déc. 2025)
[16] Odyssey : Odyssey-2 Max (21 avr. 2026)
[17] TechCrunch : AMI Labs lève 1,03 Md$ (9 mars 2026)
[18] Wikipedia : Advanced Machine Intelligence Labs
[19] Meta AI : V-JEPA 2 et tests de raisonnement physique (juin 2025)
[20] arXiv : V-JEPA 2 (2506.09985)
[21] LeCun : A Path Towards Autonomous Machine Intelligence (2022)
[22] Ha et Schmidhuber : World Models (arXiv 1803.10122, 2018)
[23] TechTimes : la typologie renderer / simulator / planner de World Labs (juin 2026)
[24] Google DeepMind : SIMA 2 (13 nov. 2025)
[25] Wikipedia : Genie (world model)