

HY-World 2.0: Tencent's Open-Source Model Turns Text or a Photo Into a Walkable 3D World

Kabui, Charles

2026-04-20

[Read at ToKnow.ai](#)



HY-World 2.0: Text or a Photo Into a Walkable 3D World

Tencent's Open-Source Four-Stage Pipeline for 3D World Generation

- 4 Stages**
Text/image to 3D world pipeline
- 1.2B**
WorldMirror 2.0 parameters
- Marble**
Matches closed-source SOTA

April 20, 2026

ToKnow.ai

Tencent released [HY-World 2.0](#), an open-source framework that converts a text prompt, a single photo, or a video into a navigable 3D world. Unlike video world models like Genie 3 that produce flat clips, HY-World 2.0 outputs real 3D assets: meshes and [3D Gaussian Splats](#) (a

technique that represents scenes as collections of colored 3D blobs for fast, high-quality rendering) that import directly into Unity, Unreal Engine, or Blender. The [pipeline](#) runs four stages: panorama generation, trajectory planning with 3D scene understanding, world expansion with memory modules that enforce geometric consistency across views, and final composition via WorldMirror 2.0, a 1.2B-parameter reconstruction model. It ships with WorldLens, a rendering platform with collision detection and character support. On benchmarks, it matches the closed-source model [Marble](#) and leads all open-source alternatives.

A game designer can type “medieval village at sunset” and get a playable 3D environment in minutes instead of weeks of manual modeling. Generated assets are engine-compatible, slotting into existing workflows without conversion. The memory architecture is what makes this practical: earlier open-source attempts at 3D world generation drifted into incoherence as scenes grew. HY-World 2.0 maintains consistency across views, producing worlds stable enough to actually use.

This signals a move from video world models to 3D-native ones. Video models generate pixels you watch. 3D models generate geometry you interact with: zero-error camera control, real physics, and permanent assets that render in real time on consumer hardware. Open-sourcing the full pipeline lets the community improve each stage independently.

Read More: [World Labs raised \\$1B for similar spatial intelligence work, but kept its models closed.](#)

Sources:

- [HY-World 2.0 Paper \(arXiv\)](#)
- [HY-World 2.0 GitHub Repository](#)
- [HY-World 2.0 Project Page](#)
- [HuggingFace Daily Papers](#)

Disclaimer: For information only. Accuracy or completeness not guaranteed. Illegal use prohibited. Not professional advice or solicitation. Read more: [/terms-of-service](#)