Fab Academy 2026

AI + Design

Amira Abdel-Rahman

3D Generation

CAD + LLMs

Physics Simulation

Robotics

Recent Updates

3D Generation

We stopped hacking 2D image AI to make 3D. Now AI natively thinks in 3D like going from translating word-by-word to writing in the language directly.

CAD + AI

LLMs generate real parametric CAD (not just meshes) and can check their own work by rendering and visually inspecting the result.

Physics Simulation

LLMs can now act as "design agents" that iterate with physics simulators.

Robotics

Robots now see, understand language, and act physically. Using "VLA."

How Text-to-3D Used to Work (2022–2024)

"A fox in a lab coat"

↓

2D Image AI (Stable Diffusion)
Already knew how things should look

↓

Slowly sculpt a 3D shape
Rotate it, ask the 2D AI from every angle

↓

Often blurry, multi-face artifacts

~30+ minutes per object

NeRF (Neural Radiance Field)

A way to store a 3D scene inside a neural network. You feed it a camera position → it outputs what the camera would see. Not a mesh more like a magic camera.

SDS Loss (Score Distillation Sampling)

DreamFusion's trick: render the 3D object from a random angle, ask Stable Diffusion "does this look like the prompt?", use the answer to nudge the 3D shape. Repeat thousands of times.

3DGS (3D Gaussian Splatting)

A faster alternative to NeRFs. The scene is millions of tiny colored blobs ("splats"). Much faster to render, now used in games and VR.

How It Works Now

The new models learn 3D directly, no 2D workaround needed.

"A fox in a lab coat" or a photo

↓

3D Diffusion Transformer
Trained on millions of 3D objects, natively understands shape, depth, geometry

↓

Full 3D mesh + textures + PBR materials
Clean topology, production-quality, game-ready

10–25 seconds!

Diffusion Model

An AI that learns by destroying data (adding noise) then learning to undo it. At generation time: start from pure noise → gradually un-noise into a result. This is how Stable Diffusion and DALL-E work.

Flow-Based Diffusion Transformer

Combines: (1) a Transformer (the architecture behind ChatGPT, great at patterns) with (2) flow matching (a smoother, faster version of diffusion). Better quality, faster generation.

Models Examples

OPEN SOURCE ✦

Hunyuan3D 2.0–3.0

Tencent · 2025

Two-stage: geometry AI + texture AI. v3.0 "Omni" adds control via point clouds, skeletons, bounding boxes. 10–25 sec. Outperforms peers on benchmarks.

github.com/Tencent/Hunyuan3D-2

OPEN SOURCE

TRELLIS 2

Microsoft · 2025

4B-parameter transformer. 1536^3 resolution with 4K PBR textures. "Structured 3D Latents": a compressed 3D format.

github.com/microsoft/TRELLIS

COMMERCIAL

Meshy AI

Various · 2025

Bulk generation, auto-rigging for animation, PBR textures, multi-language prompts. The most production-ready consumer tool today.

meshy.ai

COMMERCIAL

Tripo / TripoSG

Tripo · 2025

Fast iteration, auto-rigging, animation generation. Good for rapid concept exploration. Free tier available.

tripo3d.ai

RESEARCH

Meta WorldGen

Meta · Nov 2025

Generates navigable 50×50m 3D worlds from text. Exports to Unity/Unreal. Walkable environments.

arxiv.org/abs/2411.11939

RESEARCH

Genie 3

DeepMind · Aug 2025

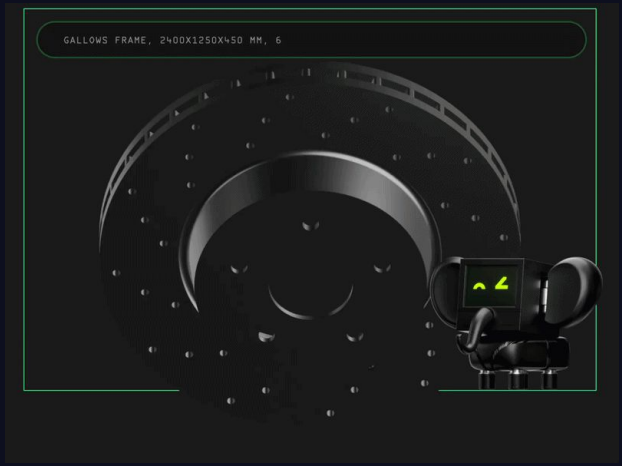
Playable 3D environments from a single image or text. Move around, interact with objects. Think: instant game level from a sketch.

deepmind.google

02 — CAD + AI

LLMs for CAD

Parametric, editable, constrainable CAD models.



CADCodeVerify's Self-Check Loop:

Generate → Render → Inspect → Refine

CAD-Llama

Fine-tunes LLM to output "Structured Parametric CAD Code": sketch → extrude → fillet sequences. Describes shapes hierarchically then generates operations.

arxiv.org/abs/2502.12345

CADCodeVerify

AI generates CAD → renders it → looks at its own output → asks "does this match?" → fixes errors. A self-correcting loop.

arxiv.org/abs/2502.06103

CADialogue

Talk to your CAD software: text, speech, images, or clicking geometry. Generates real executable macros within Fusion/Rhino.

doi.org/10.1016/j.cad.2025.103823

LLMto3D

Multi-agent LLM: Agent 1 breaks down description, Agent 2 writes geometry code, Agent 3 assembles + adds parametric sliders. 3D-printable output.

https://cdml-lab.github.io/LLMto3D/

AI + Physics: Validation

Can AI check if a design will actually stand up, not fall apart, or survive real forces?

Genesis

A universal physics simulation platform. Simulates rigid bodies, fluids, soft objects, cloth all differentiable (AI can learn from it). Up to 80× faster than real-time.

genesis-embodied-ai.github.io

Atlas3D

Generates 3D shapes that are self-supporting and stable. Adds physics loss during generation: "will this tip over? will this sag?" If yes, reshapes it automatically.

arxiv.org/abs/2403.15795

LLM-to-Phy3D

LLM + physics simulator in a loop. The LLM generates a 3D shape → simulator tests stability and aerodynamics → LLM refines the shape. Repeat until it passes.

arxiv.org/abs/2411.17209

Still hard: approximate physics, unstable gradients at collisions, very compute-heavy, limited material fidelity (warping, thermals).

Agentic Design Loops

LLMs act as "design agents" that iterate with physics simulators

1. LLM generates a 3D design
2. Physics simulator tests it (stress, gravity...)
3. Results sent back to LLM as feedback
4. LLM refines the design

Repeat until design passes

Differentiable Simulation

A physics sim where the AI can compute gradients ; meaning it can figure out "if I make the wall thicker, how much less will it bend?" automatically.

Robotics and VLAs

One model that sees, reads, and acts.

Old Approach (2022–2023)

Separate systems stitched together

Vision model
(sees the world)

↓

Language model
(plans what to do)

↓

Control policy
(moves the robot)

Slow, brittle, hard to coordinate

New: VLA (2024–2026)

One unified model does everything

Camera

"Pick up
the red cup"

↓ both at once ↓

VLA

Vision-Language-Action Model

One neural network: sees + understands + moves

↓ directly outputs

Motor commands
at 50 Hz (50 times/sec!)

Key VLA Models

FLAGSHIP

π0 / π0-FAST

Physical Intelligence

First large-scale generalist VLA. Trained on 8 robot types, one brain for arms, hands, bipeds. Uses flow matching for smooth 50Hz actions.

HUMANOID

Helix

Figure AI · Feb 2025

First VLA controlling a full humanoid body. Dual-brain: System 2 (understands scene) + System 1 ("muscle memory" for movement).

NVIDIA

GR00T N1

NVIDIA · Mar 2025

VLA for humanoids. Trains on robot data + human videos + synthetic data together. Benefits from NVIDIA's simulation ecosystem.

GOOGLE

Gemini Robotics

DeepMind · 2025

Built on Gemini 2.0. Can fold origami, play cards. On-Device version runs locally on the robot: no cloud, low latency.

OPEN SOURCE ✦

SmolVLA

Hugging Face · 2025

Only 450M parameters (tiny!) but matches much larger models. Fully open-source, trained on community data. Democratizes VLA research.

REASONING

CoT-VLA

CVPR 2025

Adds chain-of-thought to robots: imagines a subgoal image first ("what should this look like next?") then acts. Better at long multi-step tasks.

What's Still Hard

3D Generation

Output looks great but often not manufacturing-ready, non-watertight meshes, bad topology for CNC, wrong scale. Great for concepts; still needs cleanup. Multi-part assemblies are beyond reach.

CAD

Broken constraint histories, change one dimension and things break. Assemblies, mates, kinematic intent are extremely hard. Academic demos ≠ industrial complexity.

Physics Simulation

Trust & verification remain the bottleneck. AI toolpaths need expert review. Knowledge mostly limited to 3-axis milling. Multi-axis, tight tolerances underserved.

Robotics

Long-horizon tasks still break (20-step sequences unreliable). Sim-to-real transfer fragile. Data-hungry. Prompt injection can cause unsafe actions.

AI is a fast, creative starting point: not the final word. A very talented intern: always check their work.