

# Intro stable diffusion

# Gen AI



# Gen AI

## 생성 AI 서비스 조사 보고서 - Deloitte

- 마케팅 컨텐츠 작성
- 코딩 개발 지원
- 고객 응대 지원
- 제품 설계 지원
- 연구 보고서 조사 작성
- 합성 및 학습 데이터 생성
- 기업 내 데이터 검색, 정리, 분석 및 요약
- 게임 컨텐츠 설계 및 개발
- 언어 번역
- 시뮬레이션 및 추론
- 고도화된 개인 맞춤 교육 및 훈련
- 현장 업무 지원



Nitin Mittal  
Global Generative  
AI Leader  
Deloitte Consulting LLP



Costi Perricos  
Global Office of  
Generative AI Leader  
Deloitte UK



Lynne Sterrett  
US Generative AI Market  
Activation Leader  
Deloitte Consulting LLP

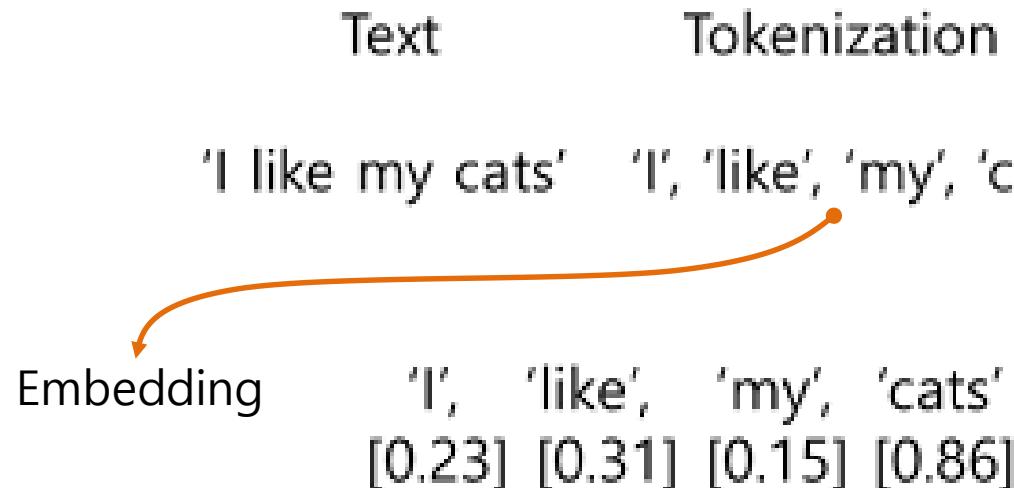


Deborshi Dutt  
US AI Strategic Growth  
Offering Leader  
Deloitte Consulting LLP



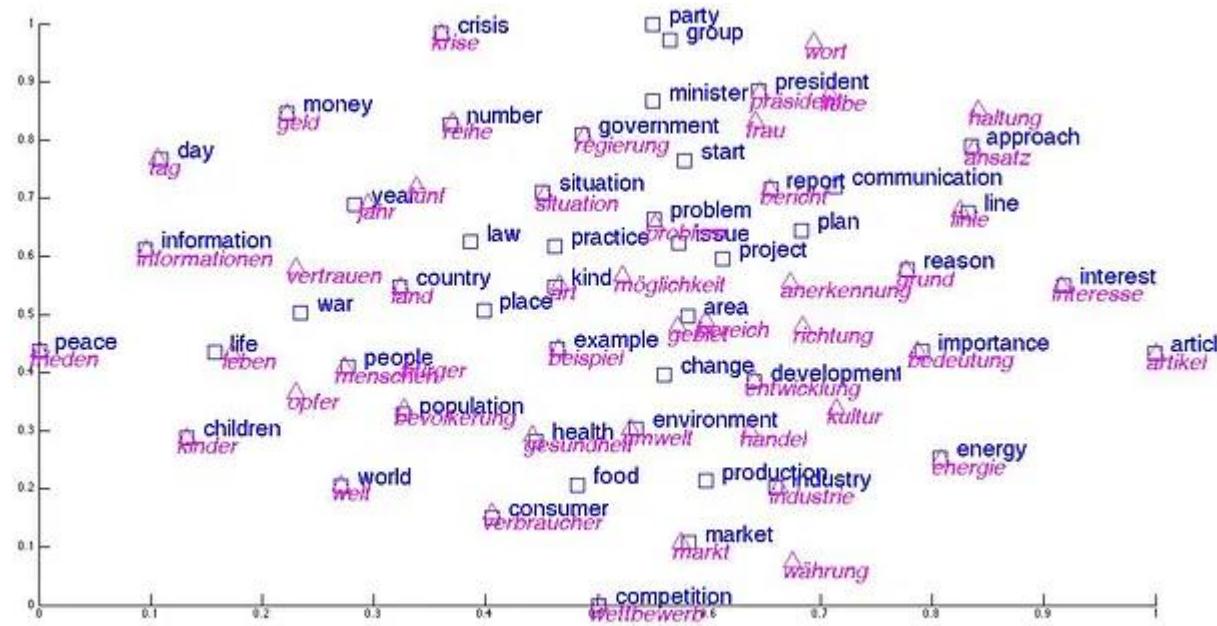
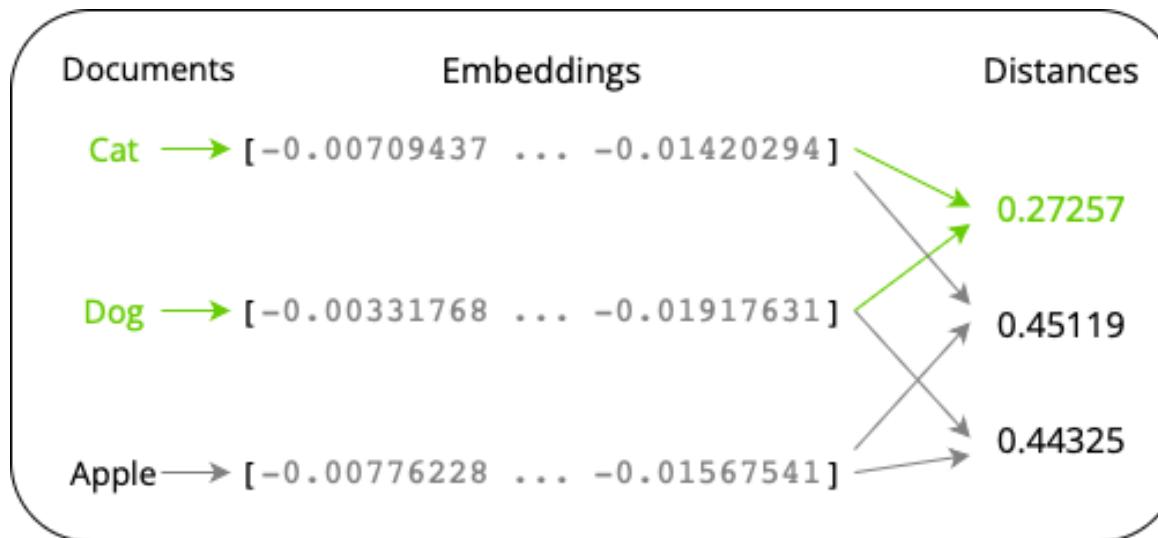
# Intro stable diffusion

# Token



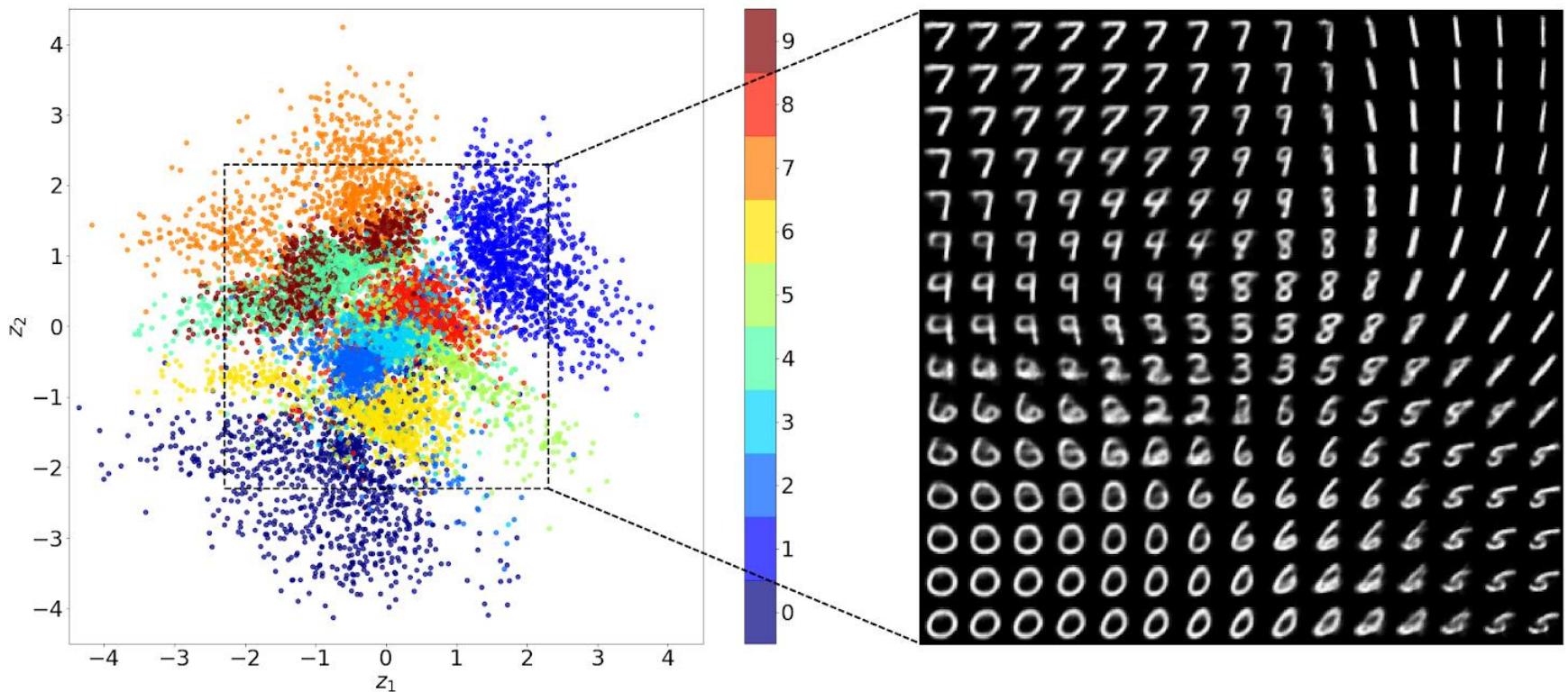
```
tokenizer.encode('123234234')
[10163, 24409, 24409]
tokenizer.encode('123')
[10163]
tokenizer.encode('234')
[24409]
tokenizer.encode('234')
[24409]
```

# Embedding



<https://daddynkidsmakers.blogspot.com/2023/12/blog-post.html>

# Latent space

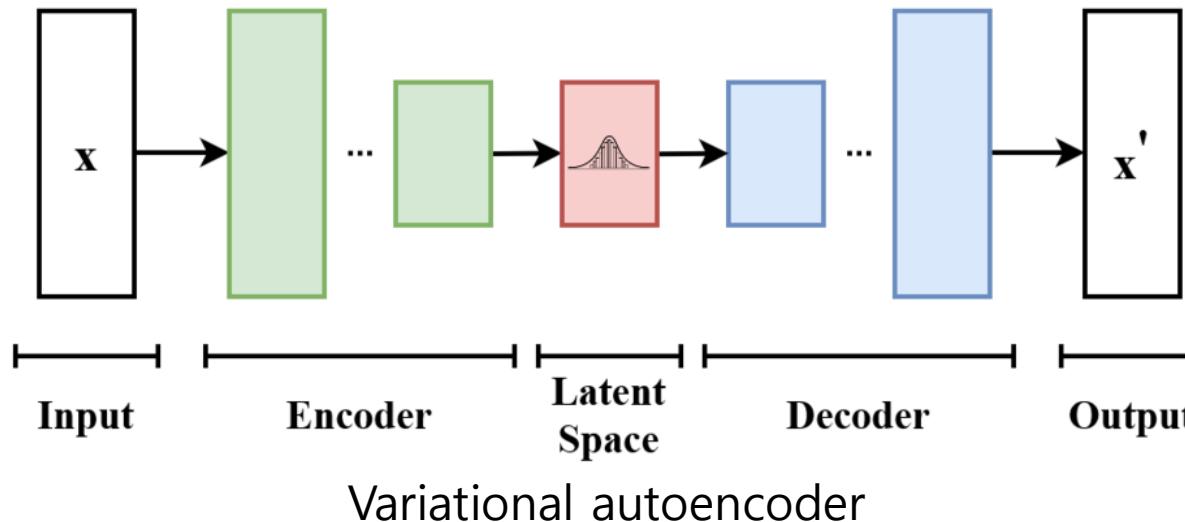


잠재공간에 맵핑된(인코딩된) 데이터(Alexej Klushyn, 2019.12, Learning Hierarchical Priors in VAEs)

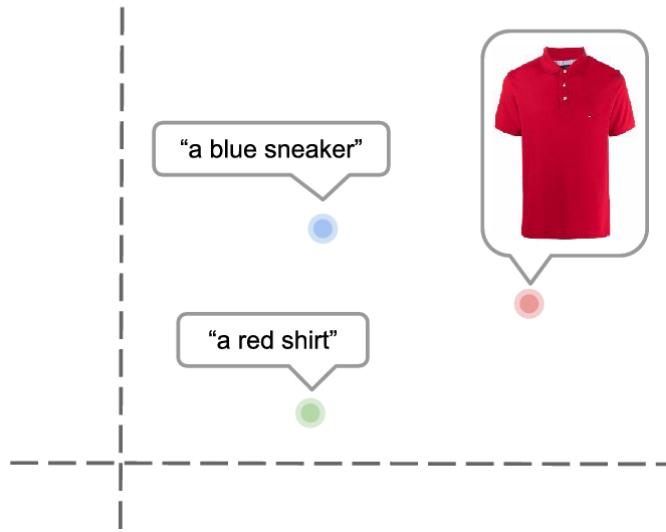
# Latent space



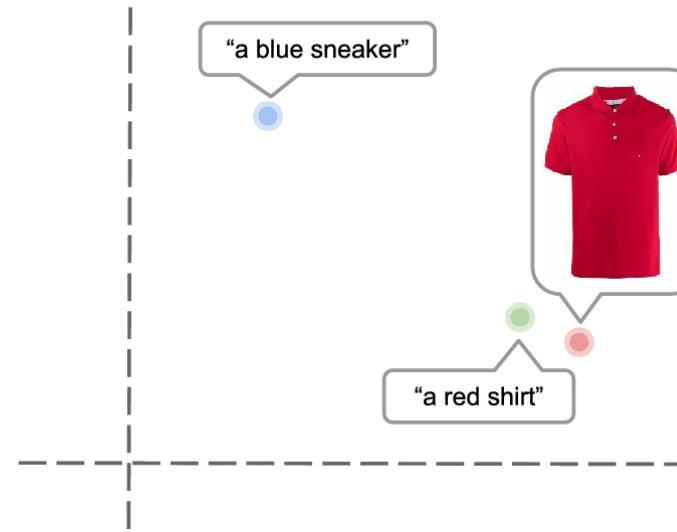
Villa savoye,  
1929, Le Corusier



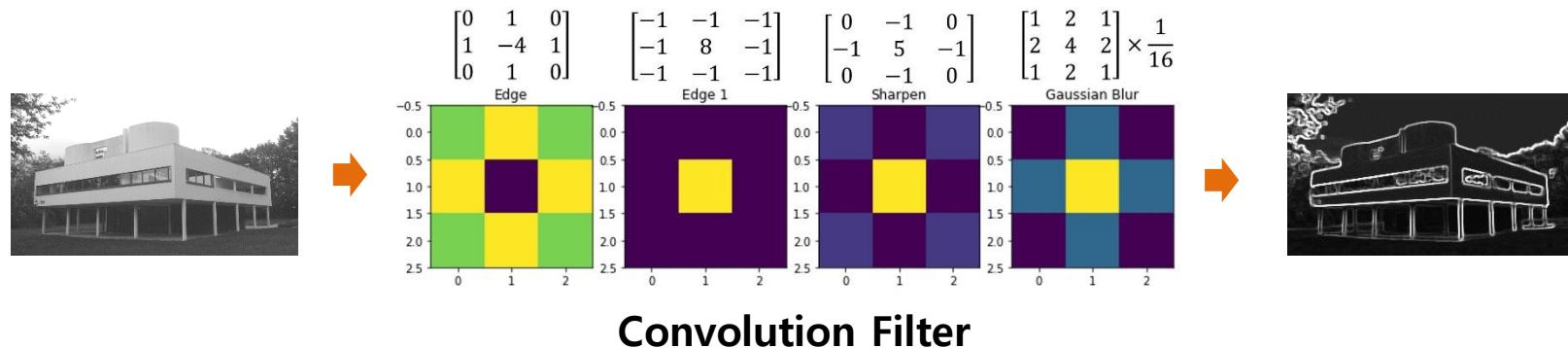
Latent Space Before Training



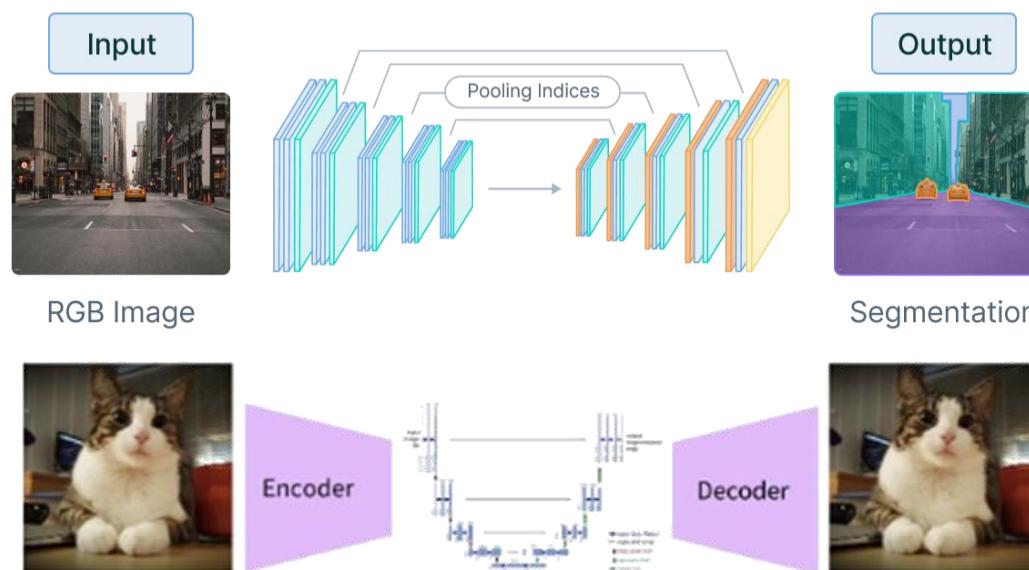
Latent Space After Training



# Convolution & Encoder-decoder U-Net



## Convolutional encoder-decoder



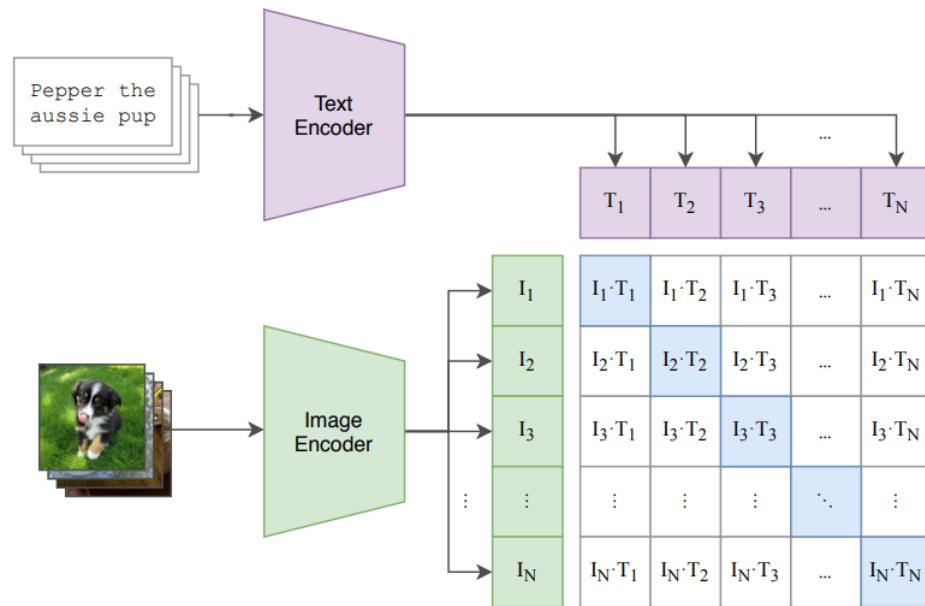
<https://www.vegaitglobal.com/media-center/knowledge-base/what-is-stable-diffusion-and-how-does-it-work>

# OpenAI CLIP multimodal 2021

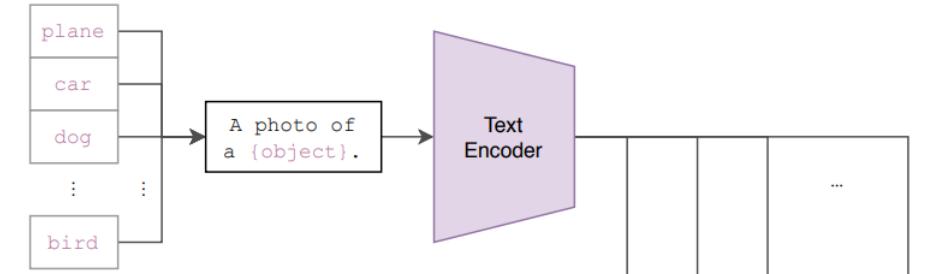


CLIP(Contrastive Language-Image Pre-Training. Open AI. 2021)

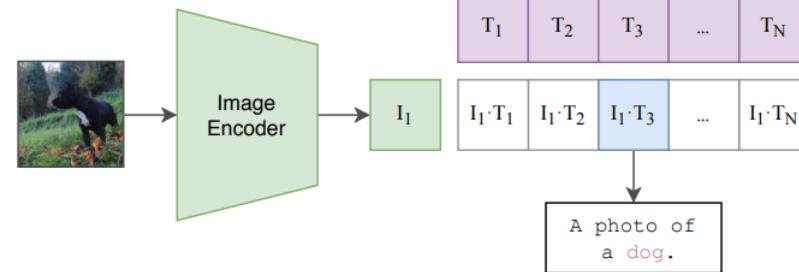
(1) Contrastive pre-training



(2) Create dataset classifier from label text



(3) Use for zero-shot prediction

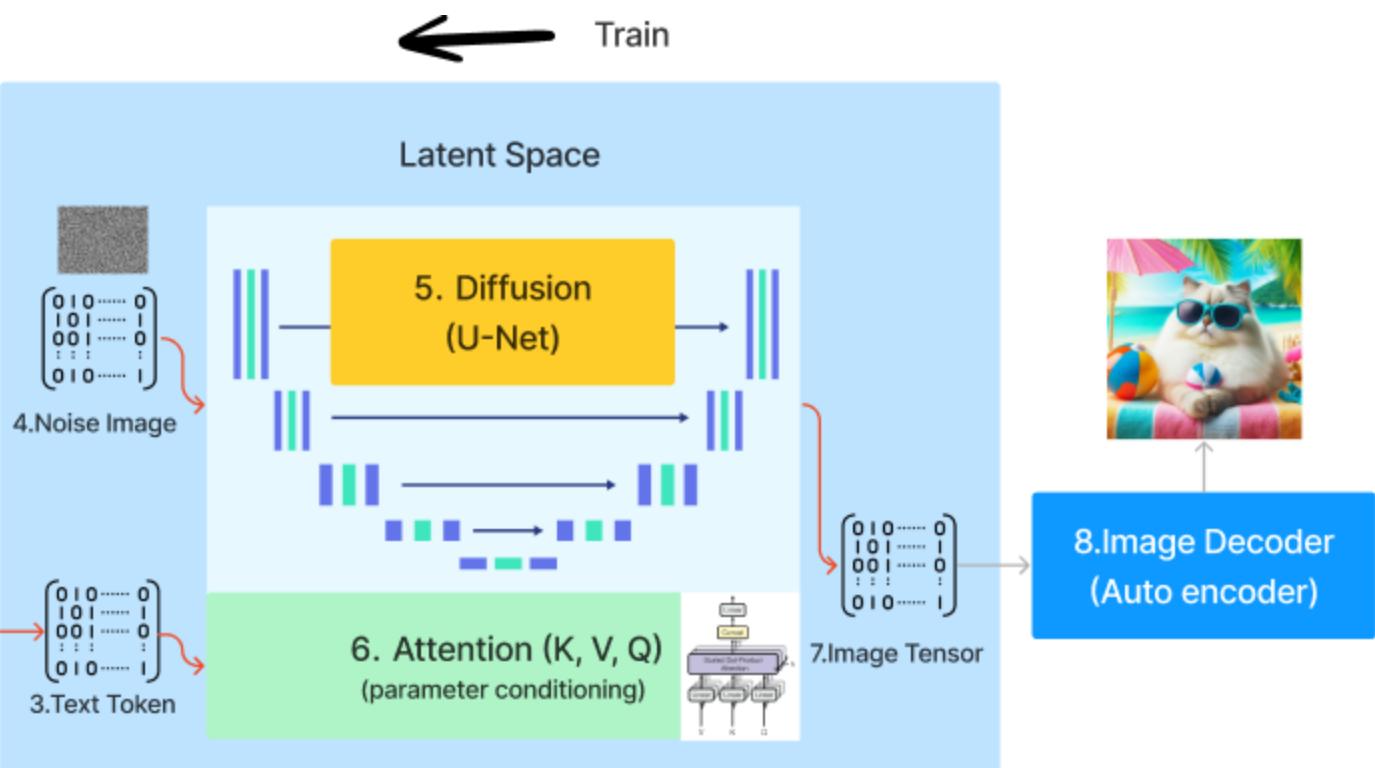


# CLIP

Stable Diffusion  
= CLIP + U-Net

1. Lovely cat in beach

2. Text Encoder (Embedding)



# Stable Diffusion 2022

## High-Resolution Image Synthesis with Latent Diffusion Models

Robin Rombach<sup>1,\*</sup> Andreas Blattmann<sup>1</sup> Dominik Lorenz<sup>2</sup> Patrick Esser<sup>3</sup> Björn Ommer<sup>1</sup>  
<sup>1</sup>Ludwig-Maximilians-Universität München & FZI, Heidelberg University, Germany      <sup>2</sup>Amazon ML  
<https://github.com/CompVis/latent-diffusion>

### Abstract

By decomposing the image formation process into a sequential application of denoising autoencoders, diffusion models (DMs) achieve state-of-the-art synthesis results with less computation than previous methods. However, these models are often trained in pixel space, which makes inference of powerful DMs often computationally expensive. To enable DM training on limited computational resources, we propose a novel approach to train them directly in latent space. We show that it is possible to apply them in the latent space of powerful pre-trained autoencoders. In contrast to previous work, training diffusion models in latent space does not require any modification to reach a near-optimal point between complexity reduction and denoising preservation, greatly boosting visual fidelity. By decomposing the image formation process into a sequence of texture, we turn diffusion models into powerful and flexible generators for general conditioning inputs such as text or house-based and high-resolution image inputs to generate images in a controllable manner. Our latent diffusion models (LDMs) achieve new state-of-the-art scores for image inpainting and class-conditional image synthesis and slightly outperform state-of-the-art methods for image-to-image, text-to-image synthesis, unconditional image generation, and super-resolution, while significantly reducing computational costs compared to pixel-based DMs.

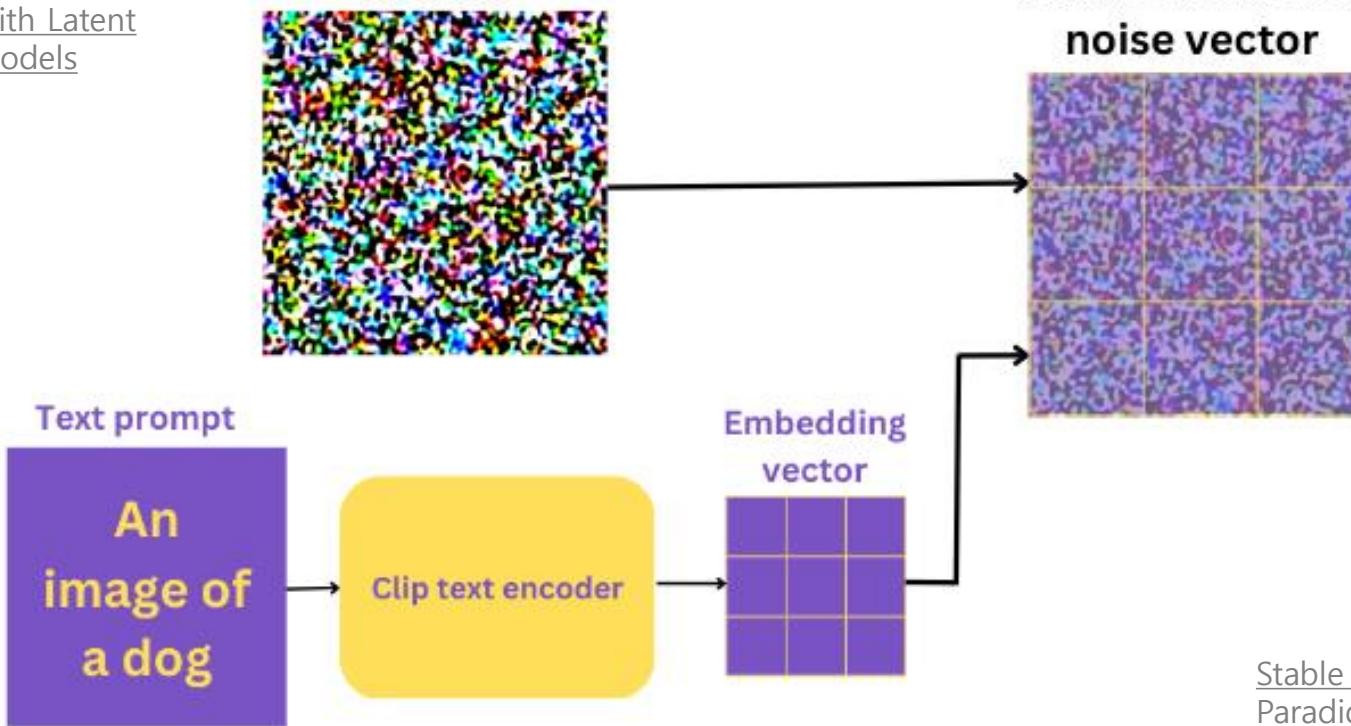
### 1. Introduction

Image synthesis is one of the computer vision fields with the most spectacular recent development, but also among



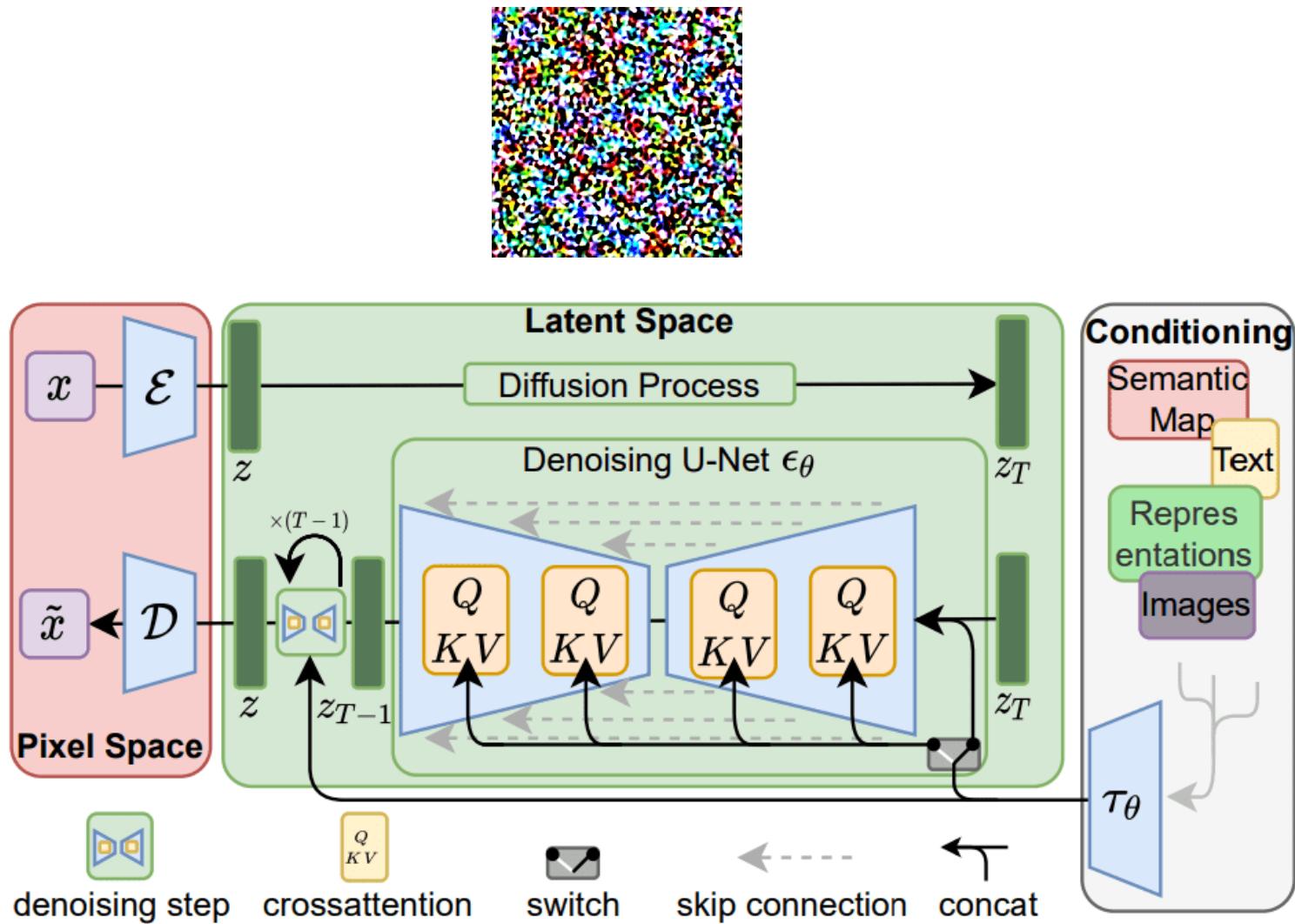
Figure 1: Boosting the upper bound on achievable quality with less aggressive downsampling. Left: a diffusion model trained on the ImageNet-1k dataset does not need the heavy spatial downampling of related generative models in pixel space, but can still greatly reduce the complexity of the data representation. Right: a diffusion model trained on the DIV2K [1] validation set, evaluated at 512 × 512. We denote the spatial downsampled version of the generated images as PSNR and PSNR are calculated on ImageNet-val [1]; see also Tab. X.

## High-Resolution Image Synthesis with Latent Diffusion Models



Stable Diffusion – A New Paradigm in Generative AI

# Stable Diffusion



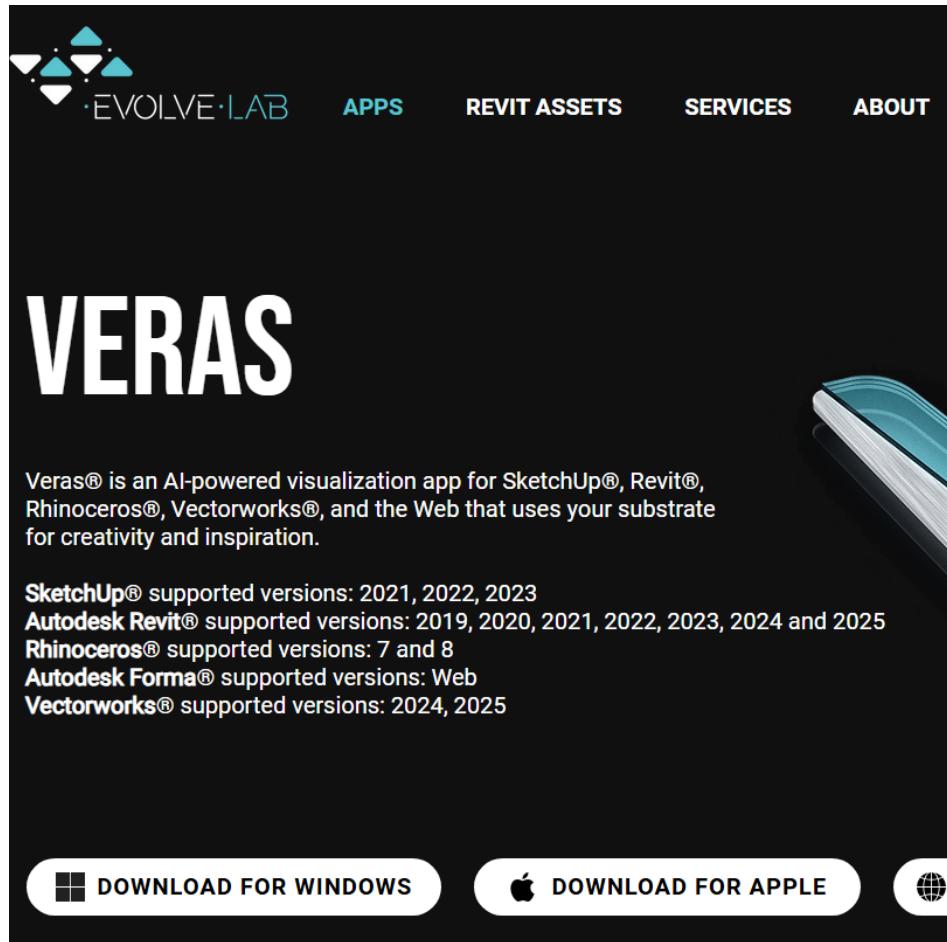
# Use cases

# Rendering



AI 기반 텍스트 렌더링 Revit 애드인 (VERAS, EnvolveLAB)

# Rendering

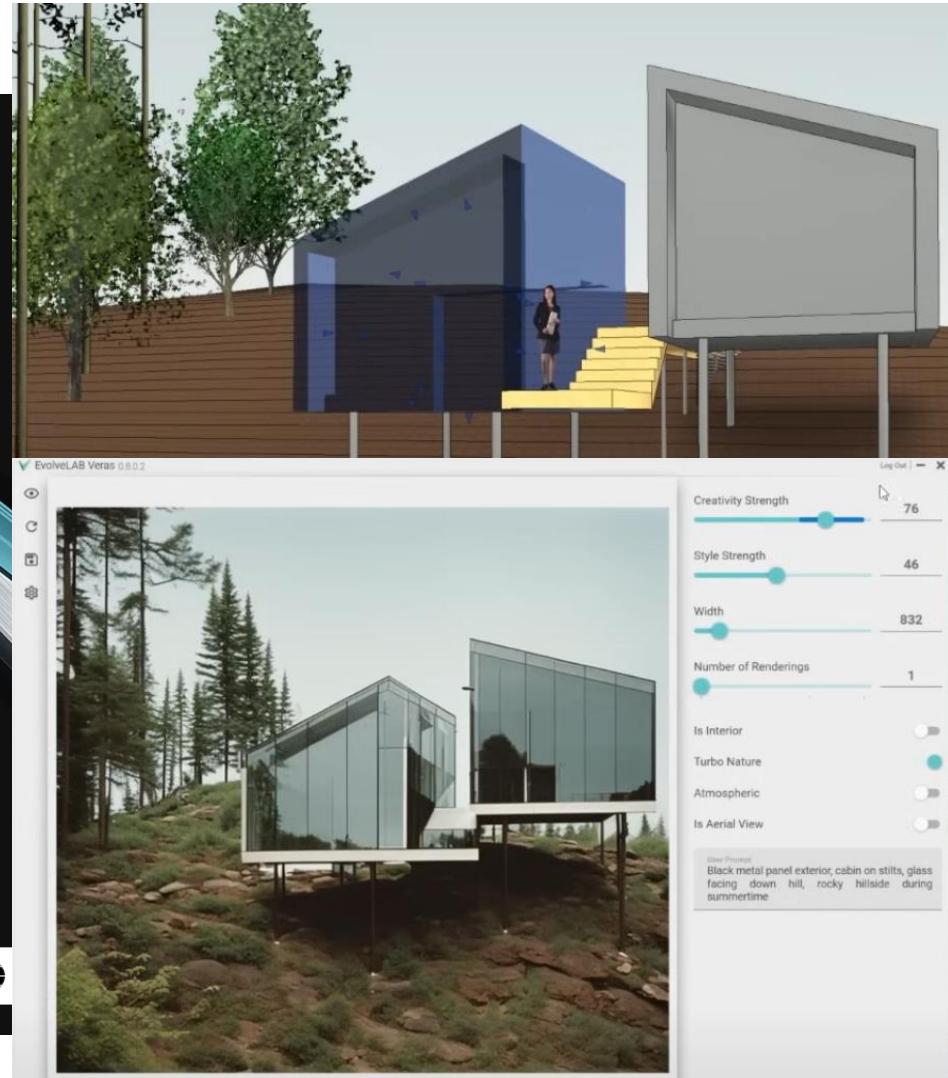


The screenshot shows the EvolveLAB website with a dark theme. At the top, there's a navigation bar with the logo (three blue triangles), followed by links for APPS, REVIT ASSETS, SERVICES, and ABOUT. Below this, a large white title "VERAS" is displayed. Underneath the title, a paragraph describes Veras as an AI-powered visualization app for various platforms. It lists supported versions for SketchUp, Revit, Rhinoceros, Vectorworks, and the Web. At the bottom, there are two download buttons: "DOWNLOAD FOR WINDOWS" and "DOWNLOAD FOR APPLE".

Veras® is an AI-powered visualization app for SketchUp®, Revit®, Rhinoceros®, Vectorworks®, and the Web that uses your substrate for creativity and inspiration.

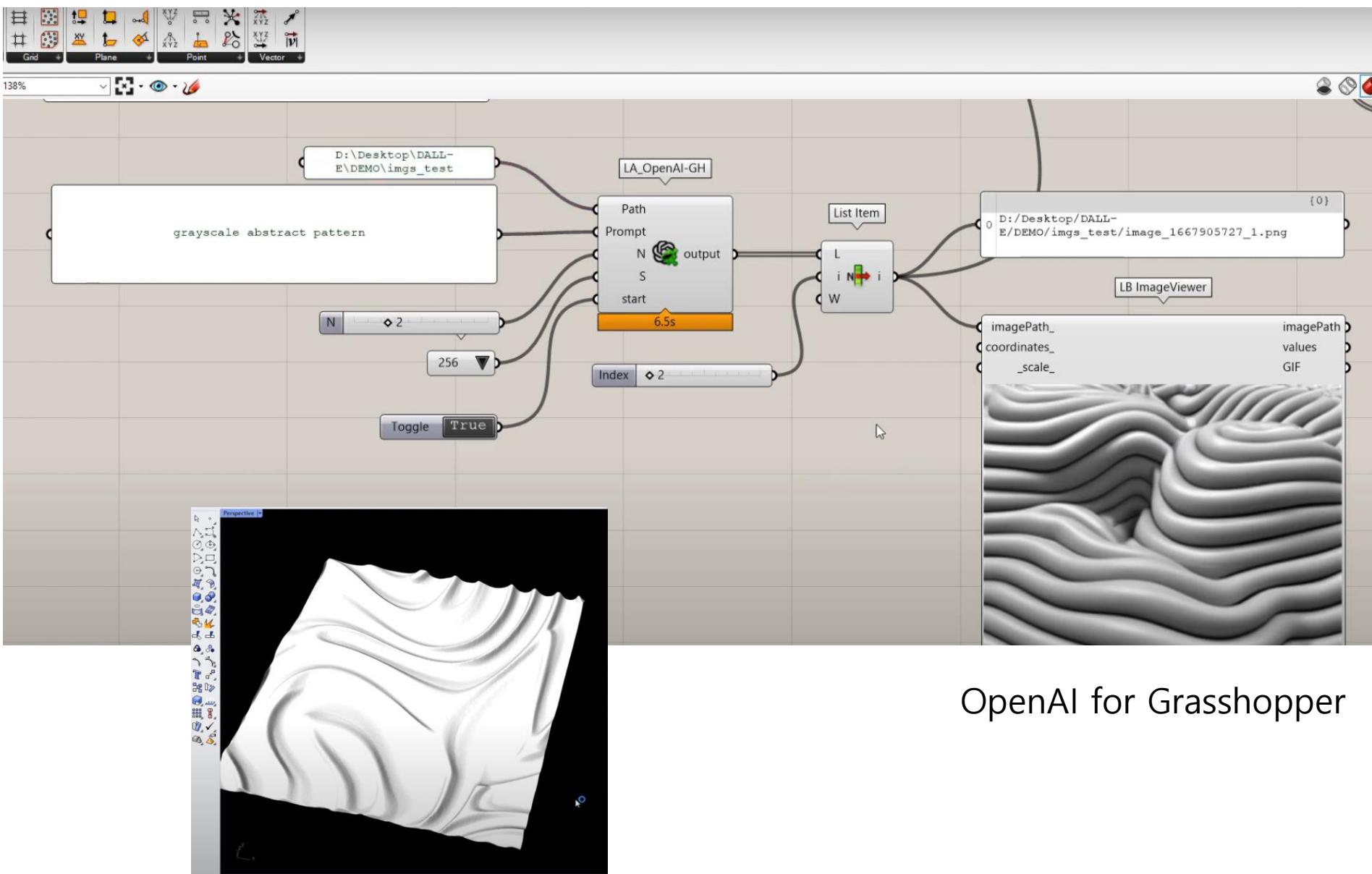
SketchUp® supported versions: 2021, 2022, 2023  
Autodesk Revit® supported versions: 2019, 2020, 2021, 2022, 2023, 2024 and 2025  
Rhinoceros® supported versions: 7 and 8  
Autodesk Forma® supported versions: Web  
Vectorworks® supported versions: 2024, 2025

[DOWNLOAD FOR WINDOWS](#) [DOWNLOAD FOR APPLE](#)



VERAS | EvolveLAB

# Modeling



# Modeling

2D IMG ► Reverse prompt ► 2D/3D Geo

► CAD

Digital Fabrication  
3D Printing  
CNC Milling ...

A modern, futuristic house with a unique organic and fluid architectural design, featuring smooth, flowing curves and perforated white surfaces that resemble waves or fabric in motion. The building has large glass walls and minimalist elements, with

I'll give you house design image. generate prompt from the image to use it in Midjourney text to image.

Sure! Please upload the house design image, and I'll generate a detailed prompt for you to use in Midjourney.



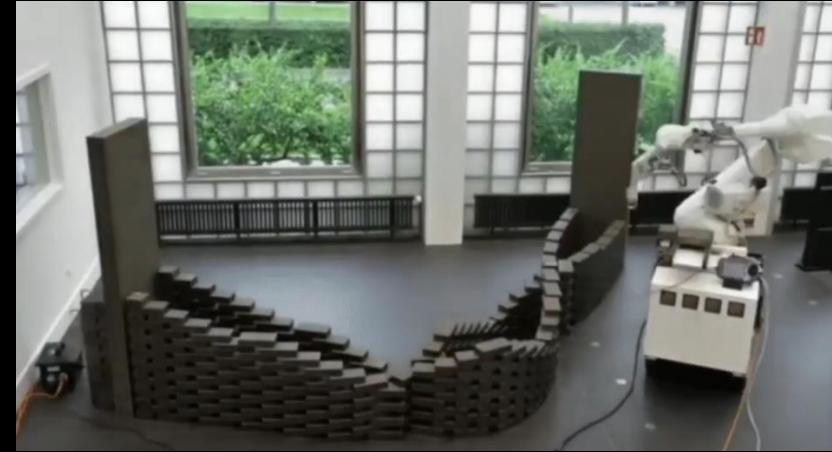
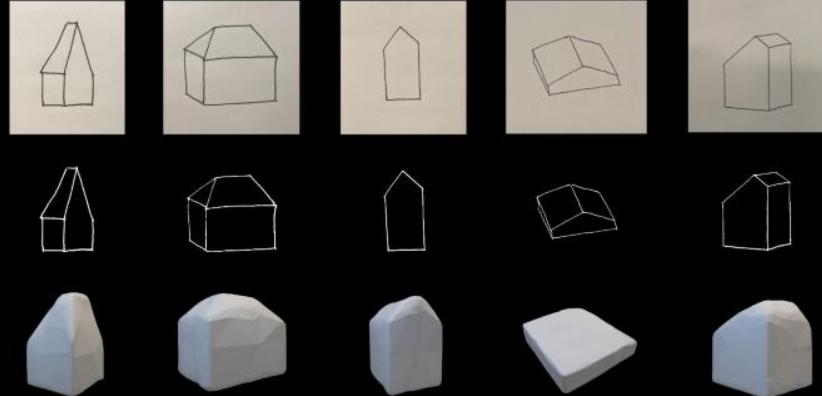
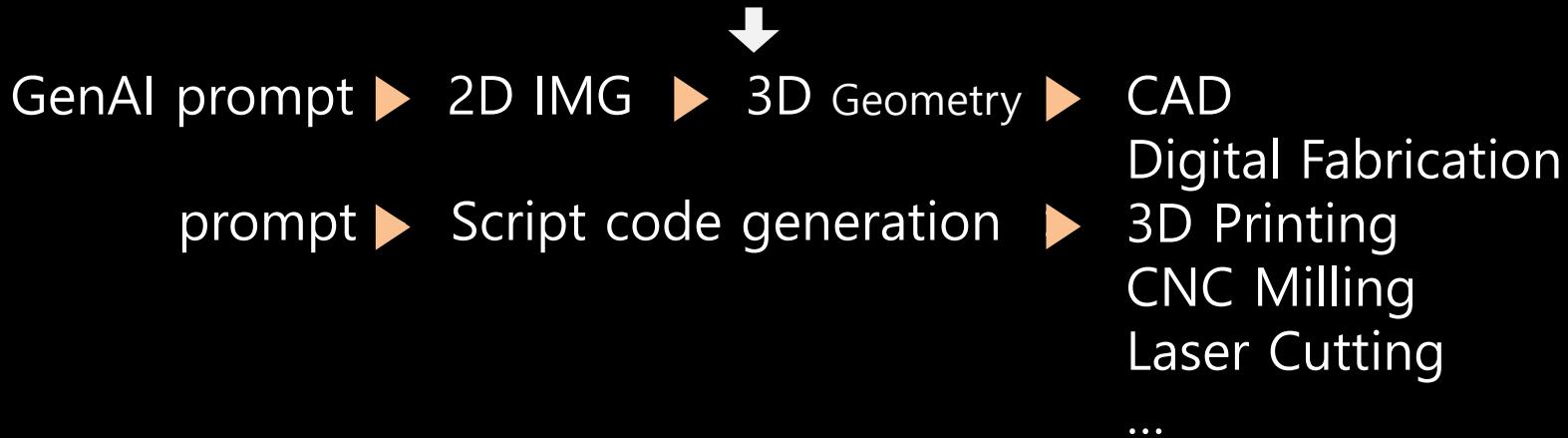
A modern, futuristic house with a unique organic and fluid architectural design, featuring smooth, flowing curves and perforated white surfaces that resemble waves or fabric in motion. The building has large glass walls and minimalist elements, with

Here's a prompt based on the design in the image for use in Midjourney:

"A modern, futuristic house with a unique organic and fluid architectural design, featuring smooth, flowing curves and perforated white surfaces that resemble waves or fabric in motion."



# Robotics



[Architectural Sketch to 3D Model: An Experiment on Simple-Form Houses | SpringerLink](#)

[Generative Design Workflows \(AI + Architecture\)](#)

# Environmental rendering using AI

[Create](#) [Remix](#) [Edit](#) [3D](#) [Create a new world by describing it. Sketch in guides if you like.](#) [Heart](#) [Download](#) [Share](#)

An open, futuristic interior with organic, flowing architecture. Smooth, curved white walls and ceilings echo the exterior, creating a fluid, airy space. Floor-to-ceiling glass walls allow natural light to flood in, illuminating sleek, minimalist furnishings in white and light gray. Polished marble floors add elegance, while built-in LED lighting highlights architectural curves. Floating sofas, suspended lights, and hidden storage create a clutter-free look, with subtle greenery adding warmth to the modern aesthetic.

522/600

Negative Text  Enhance Prompt

M3 UHD Render

GENERATE



Skybox AI

# Hand-drawn sketches to design

Discover & Remix   **Generation & History**   Assets

Image ②

Prompt

A modern living room space, inspired by Ukrainian design style, with a floor height of 5 meters, considering the relationship between space and people and the environment, and the lighting effect during the day

Reference Image ②

Tool Companions ②

Options

V1 V2

Style Scene

The screenshot displays the PromeAI interface. On the left, there's a sidebar with various icons (House, AI, 3D Model, etc.) and sections for 'Image' (with a preview), 'Prompt' (containing a detailed description of a modern living room), 'Reference Image' (with a preview), and 'Options' (with tabs for 'Style' and 'Scene'). The main area features a navigation bar with 'Discover & Remix', 'Generation & History' (which is highlighted in blue), and 'Assets'. Below the navigation is a large image of a modern living room interior. Underneath this image are several icons for editing and sharing. Further down are three smaller images of the same living room scene from different angles, labeled 'Tool Companions'. The overall theme is a modern, minimalist design.

[PromeAI - Free AI art generator : Bring creativity to life](#)

# AI for Sustainable Design

The screenshot shows a 3D architectural rendering of a city block with various buildings, roads, and green spaces. Overlaid on the scene are several blue and purple translucent layers representing wind flow and comfort zones. A large circular arrow indicates wind direction and speed. On the left, a sidebar titled 'Proposals' lists recent projects like 'Verlaine Today, 9:04 PM' and 'Tamira Oct 18'. Below it, the 'Layers' panel is open, showing categories such as Site limits, Buildings, Vegetation, Generic, Zones, Terrain, IGN Context, Roads, Property boundaries, and Railroads. On the right, there's a detailed 'Detailed wind' analysis section with a map, a 'Rapid wind' section with a small icon, and an 'Alerts' panel at the bottom.

**Real-time analysis informs your design impact**  
Leveraging machine learning models to give real-time results

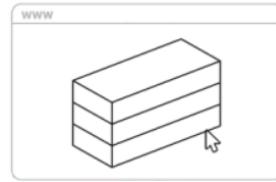
Patriarche, a multidisciplinary architecture firm, has transformed its early-stage urban planning by adopting Autodesk Forma. Traditional software was too slow for initial project phases, so Forma's AI-powered capabilities helped reduce the time for volume studies from two days to just one-and-a-half hours, allowing the team to focus on more complex design elements.

Autodesk Forma

# AI optimized design development with BIM

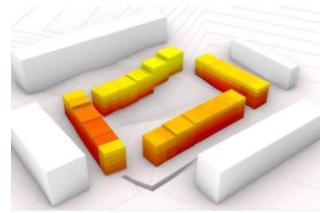
## AI Architecture Generator in Real-Time

INPUT



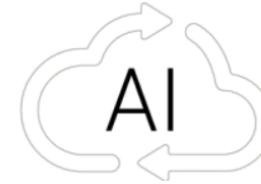
### I.-Introduction of Design Criteria and Modelling

The user inputs the design criteria that the project must meet and defines and models the solution online in an easy and intuitive way in 2D and 3D.



Environmental Responsiveness

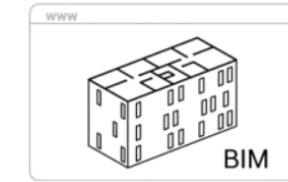
The system provides an extensive simulation of the building performance in real-time, allowing detailed knowledge of the passive behavior throughout the year. This data is used to provide the building and units layout with the best environmental performance that meet the user's design criteria.



### 2.-AI Optimized Design Development

Our cloud-based IA system generates in real-time the geometry that best fits the parameters entered for each user iteration.

OUTPUT



### 3.-Generation of the BIM solution and Project Data

The platform shows in real-time the resulting BIM solution and all its metrics to be later downloaded in XLSX, DXF, and IFC formats.



Architectures - AI  
Architecture Generator,  
Building Design

# Generating Design Alternatives



The Plan Generator allows users to specify room dimension & adjacency constraints, generate design options and export designs in DXF format. Maket's virtual assistant provides guidance on materials, costs and design possibilities.

The screenshot shows the Maket web application interface for generating floor plans. At the top, there is a navigation bar with the Maket logo, a progress bar indicating "1 of 3", and an "UPGRADE TO PREMIUM" button. Below the navigation bar, the project name "Test" is displayed. The main interface consists of a table where users can input room dimensions and adjacency constraints. The table has columns for "TYPE", "X (FEET)", "Y (FEET)", "FLOOR", and "NEXT TO (OPTIONAL)".

**Rooms Input:**

Type	X (Feet)	Y (Feet)	Floor	Next To (Optional)
Bedroom	12	12	1	Bedroom 2, Bathroom 1
Bedroom	10	10	1	Bedroom 1
Bathroom	8	8	1	Bedroom 1
Bathroom	10			

**Bottom Navigation Bar:**

- My Designs
- Floorplan Generator (selected)
- Visualizer
- Interior

**Help and Support:**

- A "Help" button with a speech bubble icon.
- A "Select" dropdown menu.

# AI for Sustainable Design

Products ▾ Support ▾ Learn ▾ Community ▾

Spacemaker

## Spacemaker: Early-stage planning, reimaged

Design better cities with outcome-based design

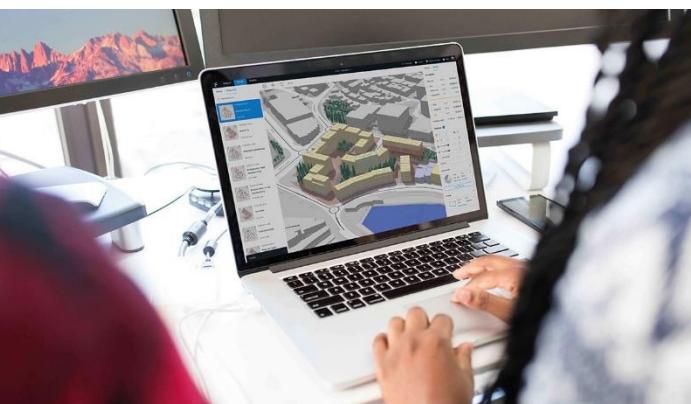
Already have an account? [Sign in](#)

[Start free trial](#) [Contact sales](#)

[Talk to our sales team. Financing options available. 1-855-664-8618](#)



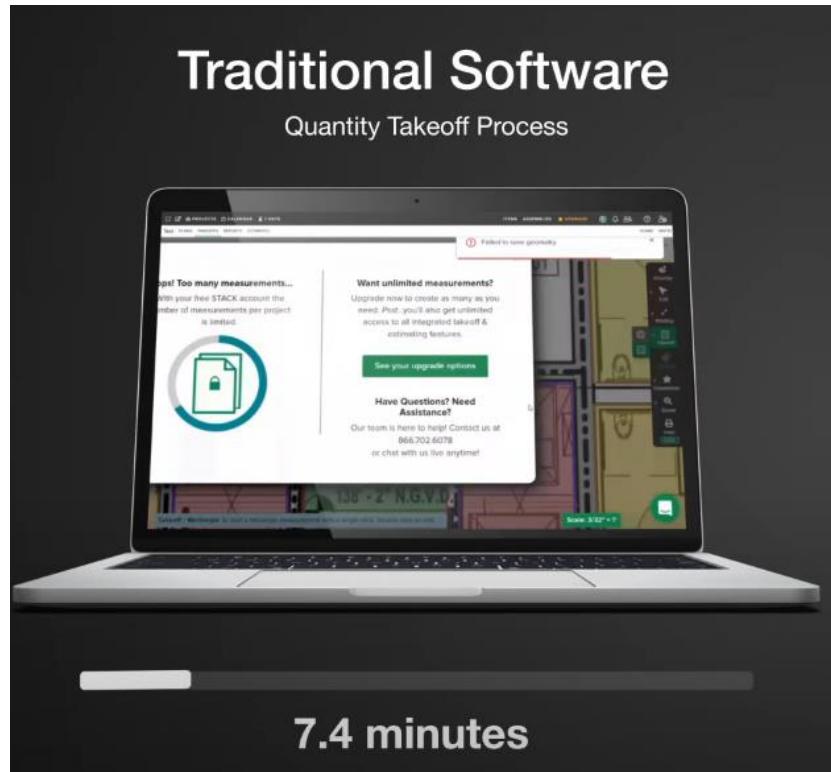
Overview Benefits Capabilities Customer Stories Webinars Resources FAQs



The image shows a detailed screenshot of the Autodesk Spacemaker software interface. The main view displays a 3D model of a city block with various buildings, roads, and green spaces. The buildings are color-coded in shades of purple, pink, and white. On the left side of the screen, there are several panels: one for 'Setup' showing a site plan and building footprints; another for 'Proposals' showing a timeline; and a third for 'Studies' showing a 'Heights' study. On the right side, there are more panels for 'Analysis' (including 'Site' and 'Building' sections), 'Results' (showing a grid of small site plans), and 'Design' (showing a 3D model with colored building footprints). A color wheel and a legend for 'Solar' and 'Wind' are also visible. At the bottom, there are tabs for 'Area' and 'Volume'.

Spacemaker | Autodesk

# AI-powered pre-construction software helping estimators



The Ultimate AI Companion for Estimators  
| Togal.AI

# LLM & Spreadsheet processing

```
1 import os
2 import pandas as pd
3 from pandasai import Agent
4
5 sales_by_country = pd.DataFrame({
6     "country": ["United States", "United Kingdom", "France", "Germany",
7                 "Italy", "Spain", "Canada", "Australia", "Japan", "China"],
8     "revenue": [5000, 3200, 2900, 4100, 2300, 2100, 2500, 2600, 4500, 7000]
9 })
10
11 os.environ["PANDASAI_API_KEY"] = " "
12
13 agent = Agent(sales_by_country)
14 answer = agent.chat('Which are the top 5 countries by sales?')
15 print(answer)
```



```
(venv_llm) PS F:\projects\pandas_ai> & C:/Users/MAC/.conda/envs/venv_llm/p
ython.exe f:/projects/pandas_ai/test.py
The top 5 countries by sales are: China, United States, Japan, Germany, Un
ited Kingdom
```

# Huggingface

Hugging Face

Search models, datasets, users...

Models

Datasets

Spaces

Posts

Docs

Solutions

Pricing



## Spaces

Discover amazing AI apps made by the community!

Create new Space

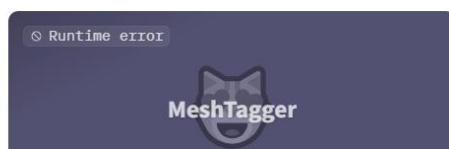
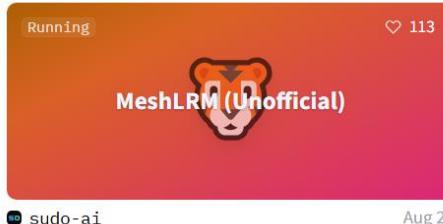
or [Learn more about Spaces](#)

mesh

Browse [ZeroGPU Spaces](#)

Full-text search

↑ Sort: Trending



[https://huggingface.co/new\\_space](https://huggingface.co/new_space)

### Create a new Space

Spaces are Git repositories that host application code for Machine Learning demos. You can build Spaces with Python libraries like Streamlit or Gradio, or using Docker images.

Owner	Space name
mac999	New Space name
Short description	Short Description
License	License

Select the Space SDK  
You can choose between Streamlit, Gradio and Static for your Space. Or pick Docker to host any other app.

Streamlit

Gradio 3 templates

Docker 15 templates

Static 3 templates

# Image to 3D mesh

Spaces

shariqfarooq/ZoeDepth

like 685

Running on A10G

⋮

Depth Prediction

Image to 3D

360 Panorama to 3D

## Image to 3D mesh

Convert a single 2D image to a 3D mesh



Keep occlusion edges

Submit



# Image to 3D mesh

Spaces

shariqfarooq/ZoeDepth

like 685

Running on A10G

≡

Depth Prediction

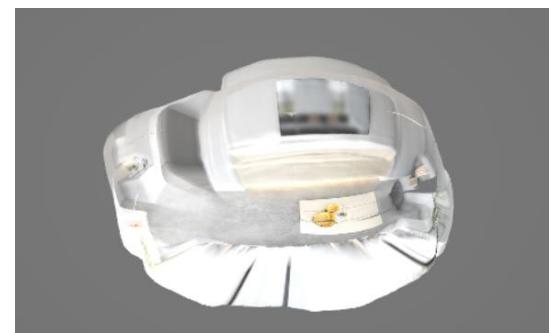
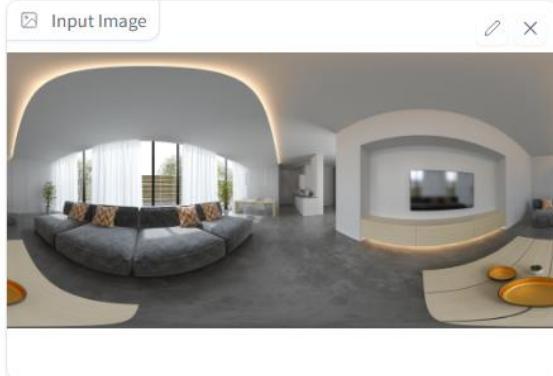
Image to 3D

360 Panorama to 3D

## Panorama to 3D mesh

Convert a 360 spherical panorama to a 3D mesh

ZoeDepth was not trained on panoramic images. It doesn't know anything about panoramas or spherical projection. Here, we just treat the estimated depth as radius and some projection errors are expected. Nonetheless, ZoeDepth still works surprisingly well on 360 reconstruction.



# Image to 3D mesh

Spaces · TencentARC / InstantMesh 1.15k like Running on ZERO

## InstantMesh: Efficient 3D Mesh Generation from a Single Image with Sparse-view Large Reconstruction Models

InstantMesh is a feed-forward framework for efficient 3D mesh generation from a single image based on the LRM/Instant3D architecture.

Code: [GitHub](#). Technical report: [ArXiv](#).

### Important Notes:

- Our demo can export a .obj mesh with vertex colors or a .glb mesh now. If you prefer to export a .obj mesh with a texture map, please refer to our [Github Repo](#).
- The 3D mesh generation results highly depend on the quality of generated multi-view images. Please try a different seed value if the result is unsatisfying (Default: 42).

The screenshot shows the InstantMesh web application interface. On the left, there is an "Input Image" panel displaying a photograph of a modern, white, organic-shaped building. Below it are several control sliders: "Remove Background" (checked), "Seed Value" (set to 42), and "Sample Steps" (set to 75). In the center, there is a "Processed Image" panel showing a semi-transparent 3D mesh reconstruction of the building. To the right, there is a "Generated Multi-views" panel displaying six smaller views of the same 3D mesh from different angles. At the bottom right, there is an "Output Model" section with tabs for "OBJ" and "GLB", and a preview image of the reconstructed mesh.

# Private LLM ollama & VLM



ollama run llama3.2-vision  
what is this figure? ./zaha\_ddp.jpg

```
>>> what is this figure? ./zaha_ddp.jpg
```

Added image './zaha\_ddp.jpg'

The image appears to be an interior view of a large, modern building with white architecture seems futuristic and minimalist, with clean lines and curves.

Upon closer inspection, it's possible that the image depicts a public space or building, such as a lobby, atrium, or hallway. The overall aesthetic suggests design, possibly intended to evoke feelings of calmness and serenity in those

이미지는 흰색 벽과 바닥이 있는 크고 현대적인 건물의 내부를 보여주는 것처럼 보입니다. 건축물은 미래적이고 미니멀하며, 선과 곡선이 깔끔합니다.

자세히 살펴보면 이미지가 로비, 아트리움 또는 복도와 같이 건물 내부의 공공 공간이나 상업 구역을 묘사한 것일 수 있습니다. 전반적인 미학은 현대적이고 세련된 디자인을 시사하는데, 아마도 공간에 들어오는 사람들에게 차분함과 평온함을 불러일으키기 위한 의도일 것입니다.

이미지는 흰색 벽과 바닥이 있는 크고 열린 공간을 보여줍니다. 천장도 흰색이고 천장 가장자리를 따라 여러 줄의 리세스 조명이 있습니다. 방 중앙에는 공간을 두 구역으로 나누는 길고 곡선형 벽이 있습니다. 벽에는 길이를 따라 일련의 창문이 있어 자연광이 방으로 들어올 수 있습니다.

곡선 벽 앞에는 반짝이는 흰색 바닥이 있는 넓고 열린 공간이 있습니다. 이곳은 건물 내의 어떤 종류의 공공 공간이나 상업 공간인 듯합니다. 이미지에는 사람이 보이지 않지만 사람들이 모이거나 모이는 장소일 가능성이 높습니다.

이미지의 전반적인 분위기는 현대적이고 매끈합니다. 깔끔한 선, 곡선, 미니멀리즘의 미학은 현대적인 디자인 스타일을 시사합니다. 기본 색상으로 흰색을 사용하면 밝기와 개방감이 생기며, 이는 공간에 들어오는 사람들에게 차분함과 평온함을 불러일으키려는 의도일 수 있습니다.

End