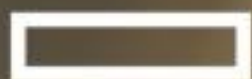


15. BECHTLE IT-FORUM THÜRINGEN

BECHTLE

2024

15. Mai 2024 • STEIGERWALD Stadion ^{Erfurt}



Hewlett Packard
Enterprise



HUAWEI

intel.



Smart Spaces in Zeiten von GenAI

Anwendungen und Beispiele

Volker Meschonat – Business Development Smart Spaces
Bechtle IT Forum Erfurt, 15. Mai 2024

OUR BODY OF WORK

NVIDIA pioneered accelerated computing to tackle challenges no one else can solve. We engineer technology for the da Vincis and Einsteins of our time. Our work in AI is transforming 100 trillion dollars worth of industries, from gaming to healthcare to transportation, and profoundly impacting society.

NVIDIA Works with Thousands of Customers and Millions of Developers

Scientists, researchers, developers, and creators are using NVIDIA to do amazing things. More than 3 million developers and 10,000 startups create thousands of applications for accelerated computing. We've shipped more than a billion CUDA[®]-based GPUs.



**3 MILLION
DEVELOPERS**

Pioneering Accelerated Computing

Accelerated computing requires full-stack optimization, from chip architecture, systems, and acceleration libraries, to refactoring the applications. The global NVIDIA ecosystem spans 4 million developers, 40,000 companies, and over 3,000 applications.

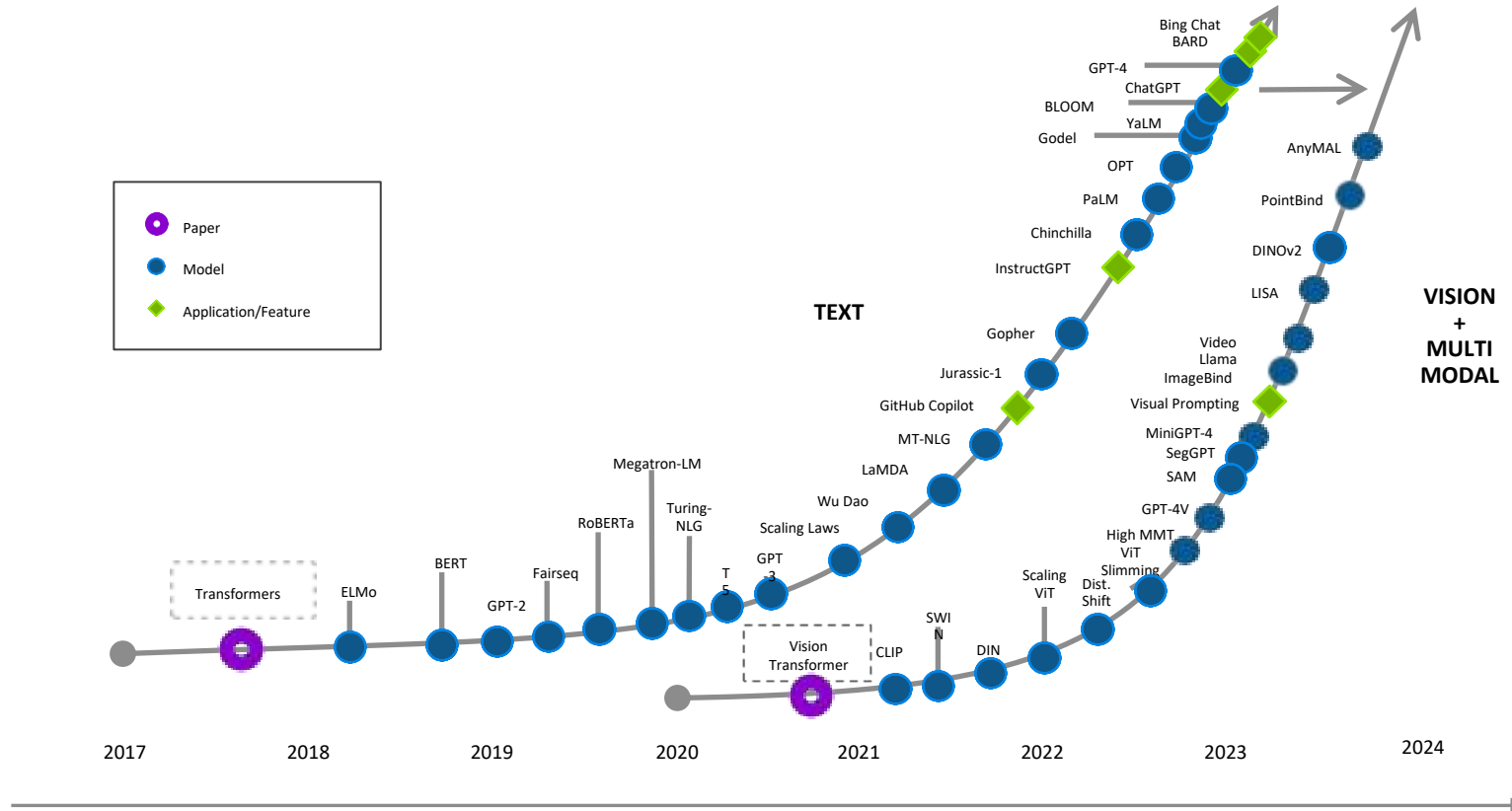




Analytik für smarte Räume:

Wandle Daten zu Einblick, Wert und Information!

Generative AI is Advancing at an Astonishing Pace



Credit: Andrew Ng

Revolutionizing How We Interact With Perception Apps

Legacy CNNs



Rigid / Rule-based
Requires tons of labeled data
Slow Development Cycle



Generative AI



Zero-shot Learning
Generalizable
Faster Development Cycle
Natural Language Prompts

TEXT GENERATION

Text generation Chat settings Parameters Model Training Interface mode

Create an itinerary for a 3-day trip to San Francisco Bay Area.

Sure

Input

TEXT & VISION

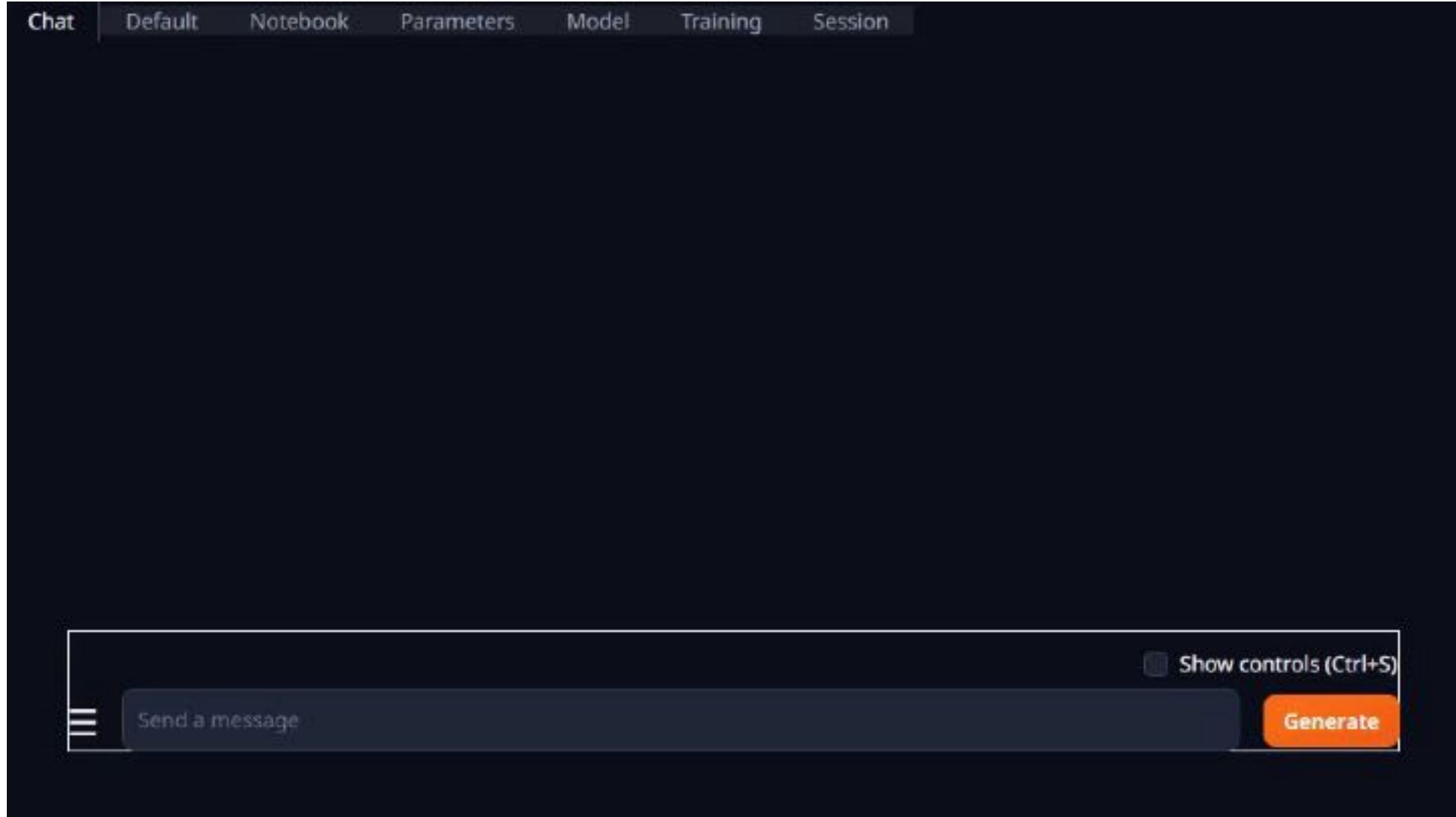


IMAGE GENERATION

The screenshot displays the web interface for image generation. At the top, the browser address bar shows "Not secure | 10.11.0.35:0.7860". The interface includes a navigation menu with options like "bt2img", "img2img", "Extras", "PNG info", "Checkpoint Merger", "Train", "Settings", and "Extensions".

The main prompt area contains the text: "Macro closeup photography of dewdrops on a spiderweb, with morning sunlight creating rainbows. Bright green gradation. Bokeh. Secular highlights. 4k." Below this is a field for a negative prompt with the instruction "Negative prompt (press Ctrl+Enter or Alt+Enter to generate)".

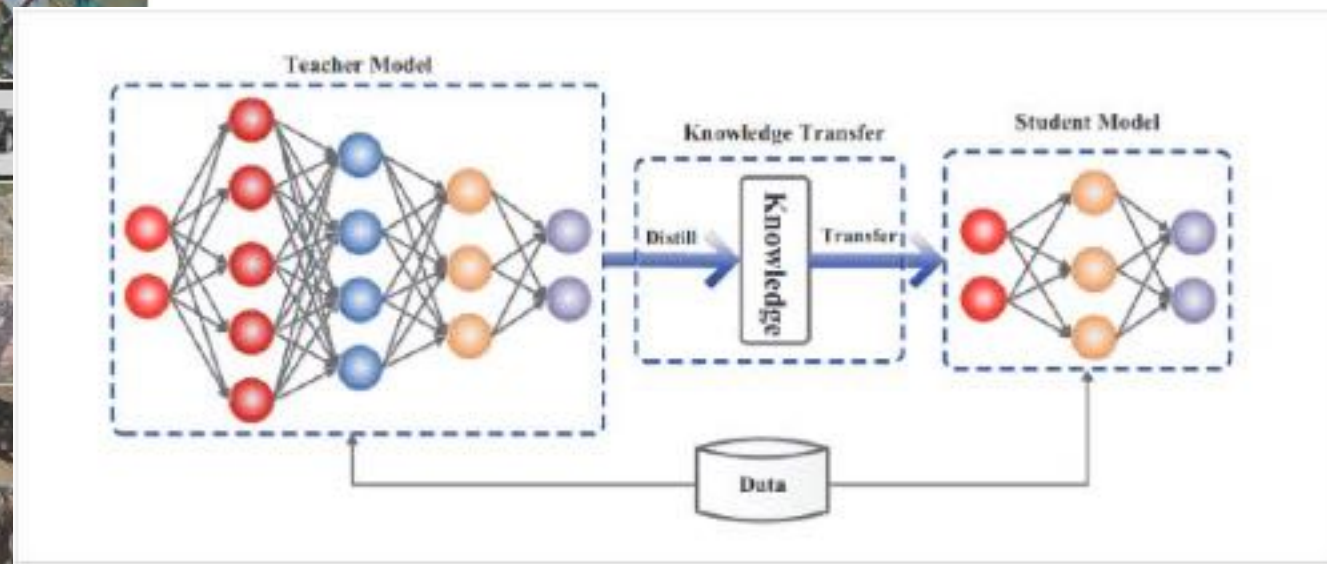
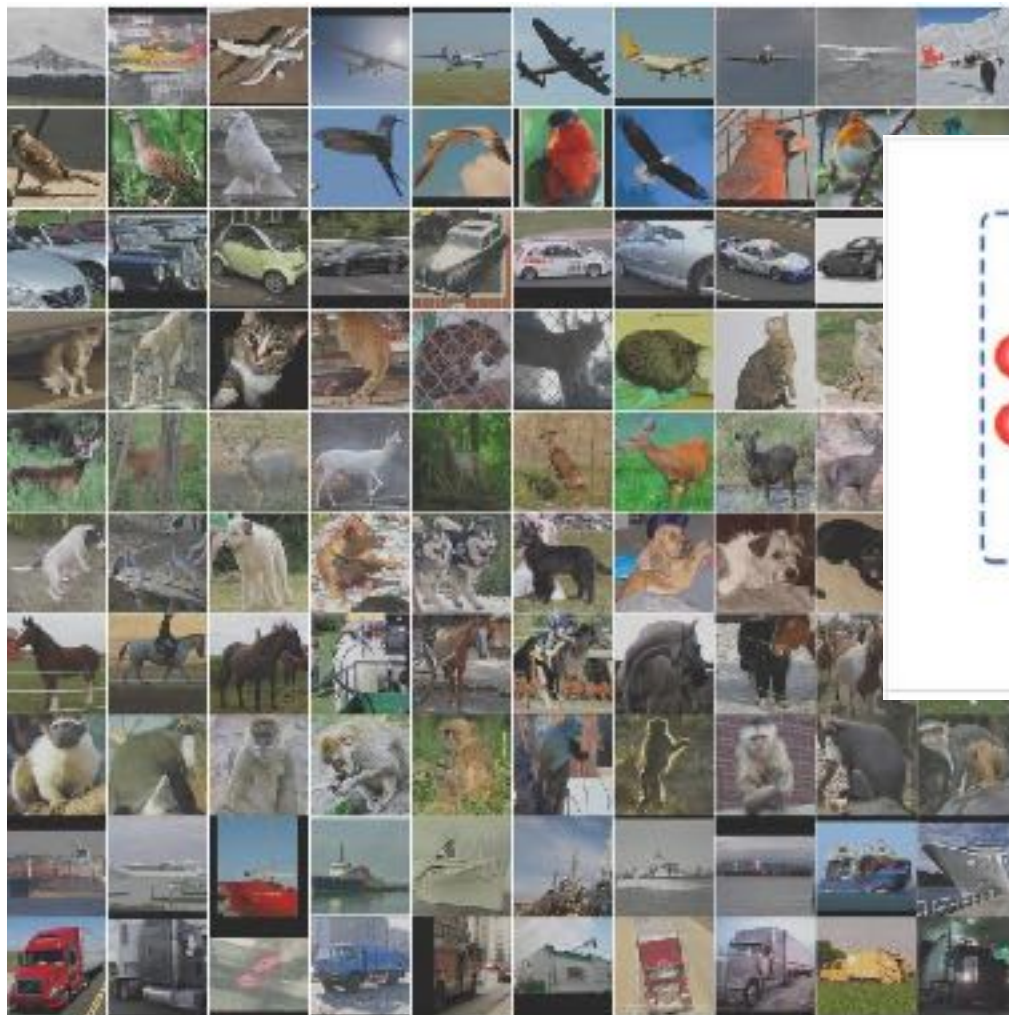
On the right side, there is a "Generate" button with a red circular icon, which is being clicked by a green hand cursor. Below the "Generate" button are several icons for actions like "Save", "Download", "Share", "Print", and "Fullscreen".

The settings section on the left includes:

- Sampling method: Euler a
- Sampling steps: 20
- Width: 512
- Height: 512
- Batch count: 1
- CFG scale: 7
- Seed: -1
- Script: None

At the bottom of the interface, there are buttons for "API", "Github", "Credits", "Startup profile", and "Reload UI".

MODEL DISTILLATION



VISION TRANSFORMERS



Search Query

Model: CLIP ViT-L/14@336px
Images: 279,958

Image

Drop Image Here

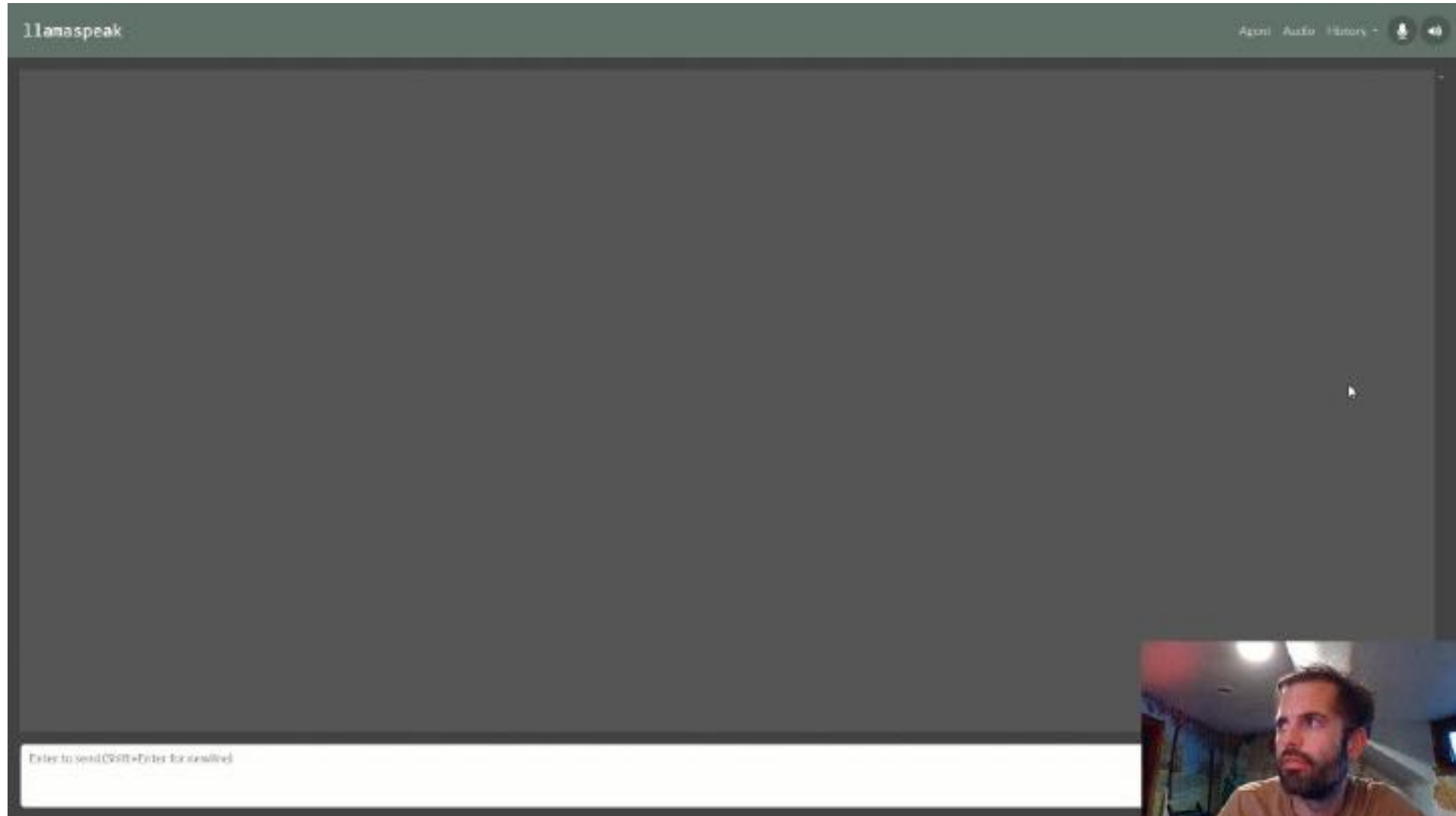
- or -

Click to Upload

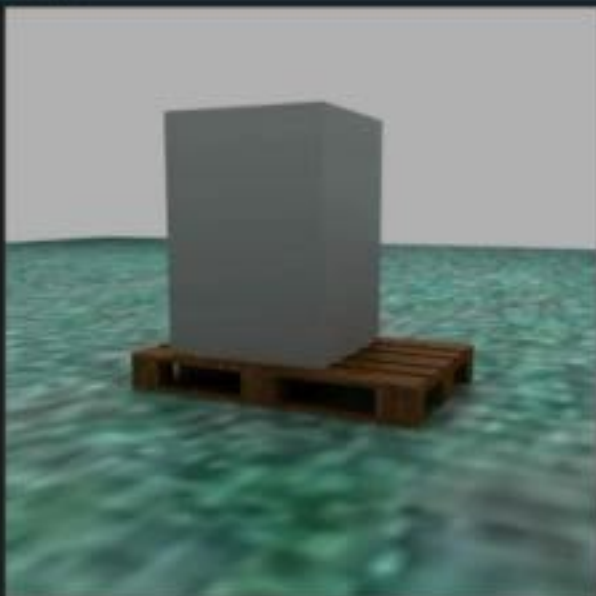
Gallery



TEXT TO SPEECH



```
image = PIL.Image.open("abfc2f6a-9417-11ed-9f97-38d547278c4d/rgb_image.png").resize((512, 512))
entity_id_image = PIL.Image.open("abfc2f6a-9417-11ed-9f97-38d547278c4d/entity_id_image.png").resize((512, 512))
image
```



```
def run(prompt, image, mask, image_for_depth):
```

```
    image_a = pipeline(
        prompt=prompt,
        image=image,
        strength=0.99,
        guidance_scale=9.0,
        negative_prompt="broken, ugly, disfigured, cartoony, nonrealistic, green, blue",
        mask=mask,
        mask_interpolation_mode="bilinear",
        image_for_depth=image
    ).images[0]
```

```
    plt.figure(figsize=(20, 20))
    plt.subplot(131)
    plt.imshow(image)
    plt.subplot(132)
    plt.imshow(mask)
    plt.subplot(133)
    plt.imshow(image_a)
    plt.show()
```


NVIDIA VIDEO AGENT

MODEL DESCRIPTION

This demo shows LMMs ability to generate responses based on video and text inputs.

Vid-LLaMA



Start Chatting

≡ SELECT A SAMPLE



Vid-LLaMA



Vid-LLaMA



Vid-LLaMA

MODEL INPUT INSTRUCTIONS

Upload a video and enter a prompt to receive an AI-generated response.

Type and press Enter

I

RESPONSE

processing | 8.1/2.1s

briefly describe what is happening in the video



...



[a person [a face (happy, sad, ...), ...], ...]



Ihr Kontakt

Volker Meschonat

+49 174 9248 732

volkerm@nvidia.com



Danke für Ihr Interesse und Ihre investierte Zeit!