

Stable Diffusion

Text-zu-Bild-KI “Stable Diffusion” - Revolution der Bilder?



Das i

Ende der Slideshow
Quellen der Bilder im [Anhang](#)



Vorstellung

- Julian Egner
- seit >20 Jahren in der Softwareentwicklung
- seit 2016 bei der tarent
- Softwareentwickler
 - Java, Kotlin
 - Angular
 - Android

Ich bin *kein* AI-Experte



SDXL 1.0, prompt: spock vulcan salute

AGENDA

Einleitung

Funktionsweise

Prompting

Einsatzmöglichkeiten

Ökosystem

Soziale und Gesellschaftliche Auswirkungen

Fragen & Diskussion



echtes Photo



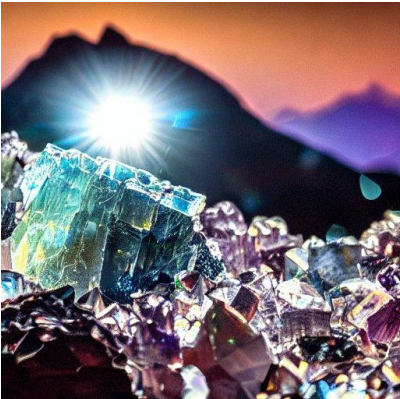


Einleitung - was ist Text-to-Image?

photograph of a (one massive colorful crystal:1.2)
growing out of the rocky mountain, (focus on
crystal:1.2), 4k, 8k, (highly detailed),
((landscape)),(translucent crystal:1.1), light going
through the crystal, bokeh, chromatic aberration,
mountain view



Die Software erstellt aus einem Text (dem Prompt) ein Image





Einleitung - was ist Image-to-Image?



In flames



Field of flowers



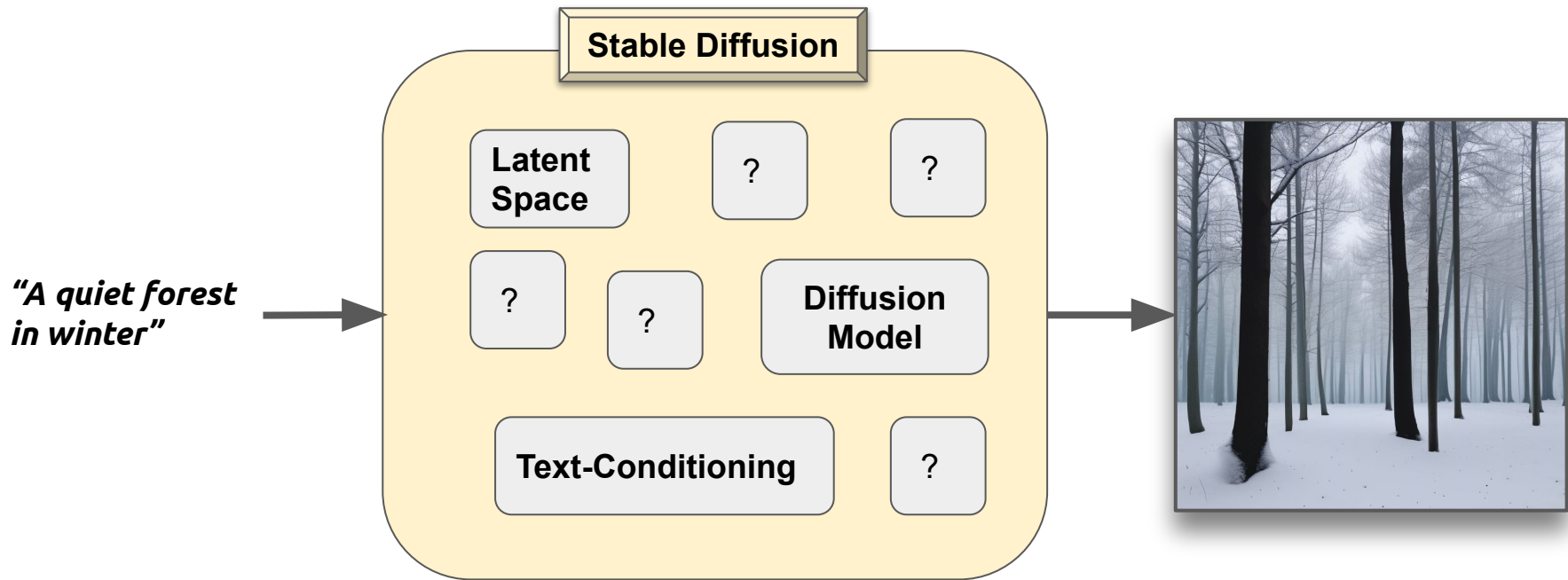
As a Monster from
monster inc movie



Funktionsweise

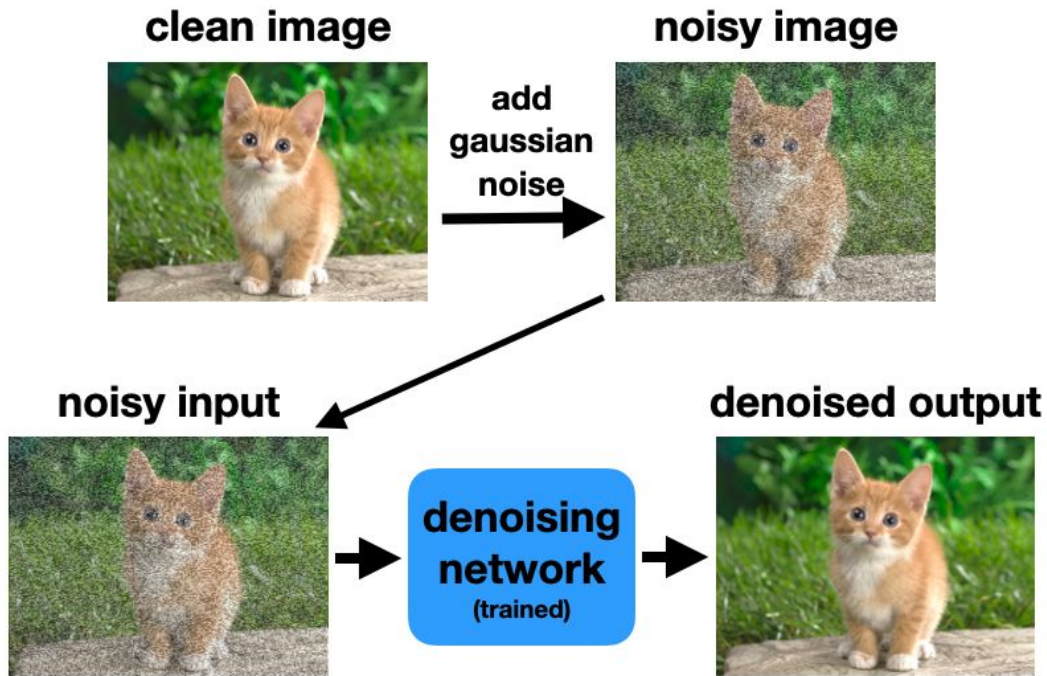


Funktionsweise: Überblick





Funktionsweise: Diffusion Models

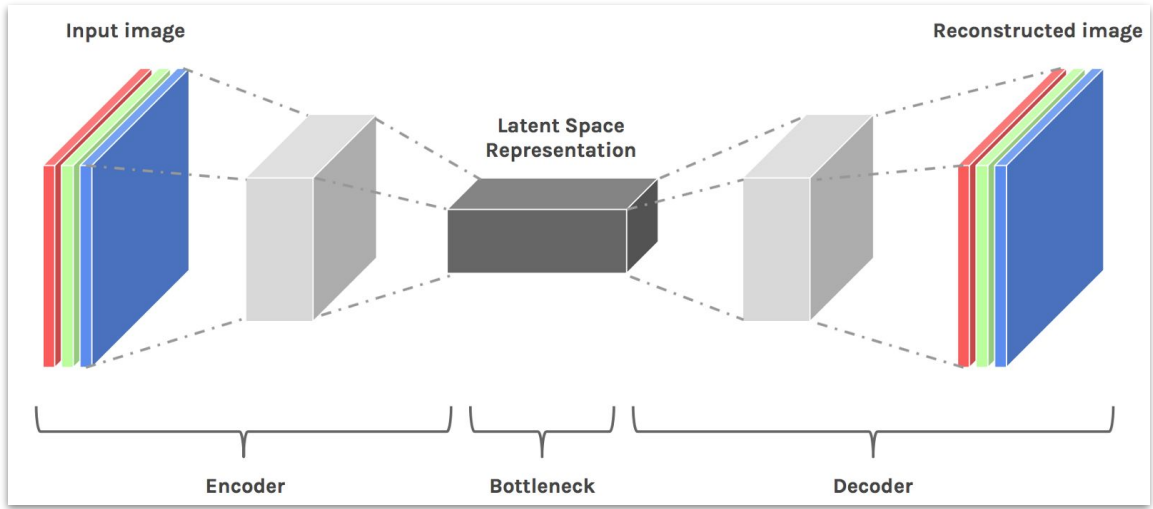




Funktionsweise: Latent Space

*"To enable DM [Diffusion Model] **training on limited computational resources** while retaining their quality and flexibility, we apply them in the **latent space** of powerful pretrained autoencoders."*

[<https://arxiv.org/abs/2112.10752>]



<https://hackernoon.com/latent-space-visualization-deep-learning-bits-2-bd09a46920df>



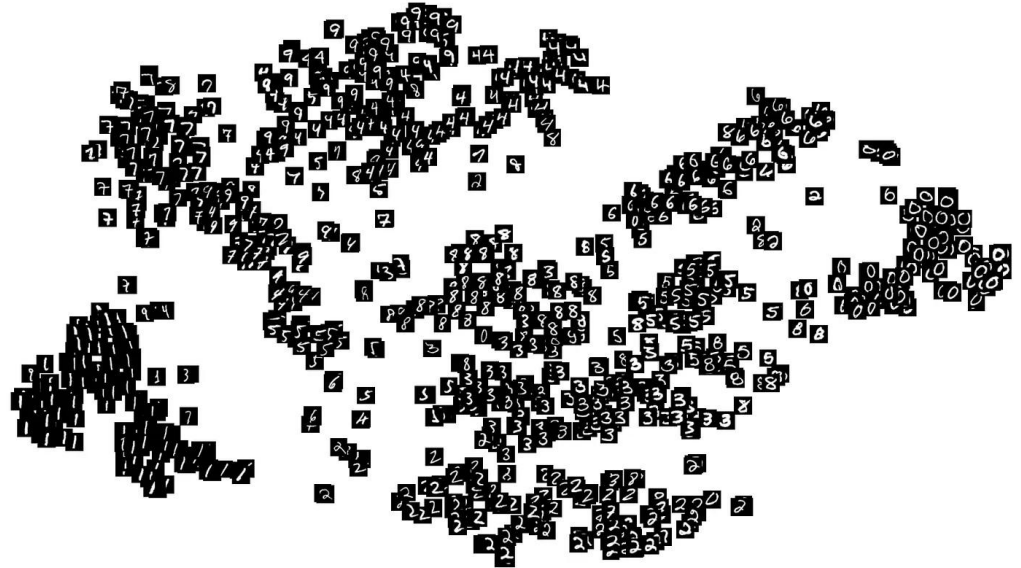
Funktionsweise: Latent Space

Stell dir einen grossen Datensatz von handgeschriebenen Zahlen (0-9) vor. Bilder der gleichen Zahl (z.B. 3) sind zueinander am ähnlichsten, verglichen mit Bilder anderer Zahlen.

Wenn wir ein Model darauf trainiert haben, Zahlen einzuordnen, dann ist es auch darauf trainiert, die “strukturellen Ähnlichkeiten” (‘structural similarities’) zwischen Bildern zu erkennen. Durch Lernen der Eigenschaften jeder Zahl kann das Model dann Zahlen erkennen.



Das passiert im Hintergrund, versteckt - “latent”.

Das Konzept des *latent space* ist der Kern des *deep learning* - Eigenschaften von Daten lernen und Datenstrukturen vereinfachen, um Muster zu finden.

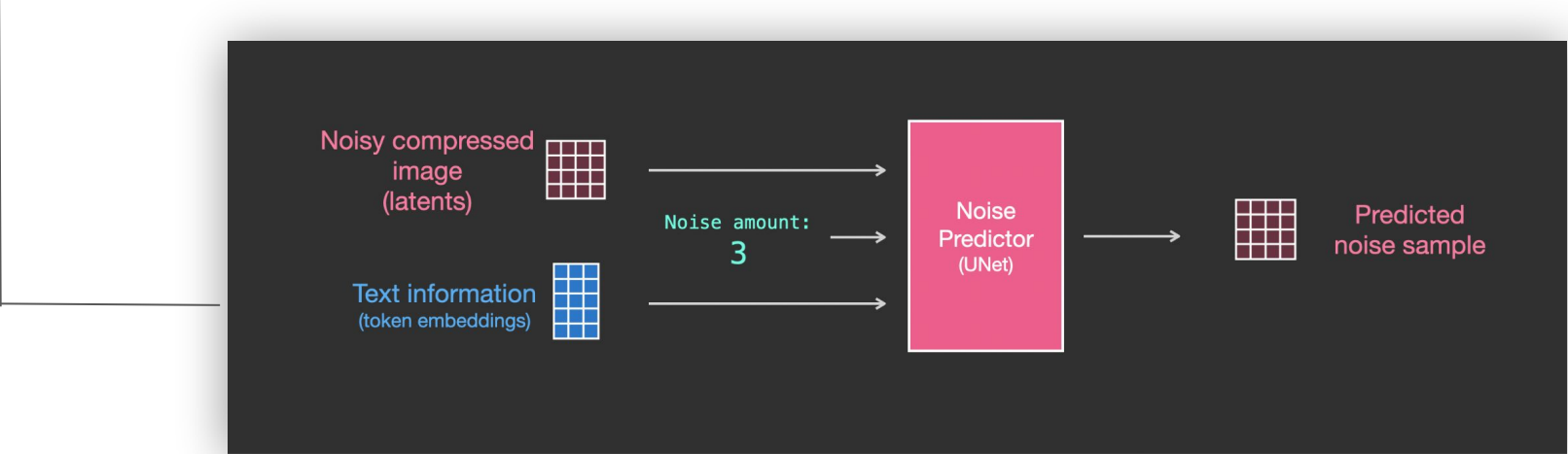




Funktionsweise: Text-Conditioning

IMAGE			
CAPTION	Photo pour Japanese pagoda and old house in Kyoto at twilight - image libre de droit	Soaring by Peter Eades	far cry 4 concept art is the reason why it 39 s a beautiful game vg247. Black Bedroom Furniture Sets. Home Design Ideas

<https://jalammar.github.io/illustrated-stable-diffusion/>





Funktionsweise: Trainings-Daten

- Dataset **LAION-Aesthetics**
 - ein Subset von **LAION 5B**
 - gefiltert danach, wie wahrscheinlich die Bilder als **“schön”** empfunden werden
 - Hunderte Millionen von Bildern wurden genutzt

Backend url:

Index:

Search:

Clip retrieval works by converting the text query to a CLIP embedding, then using that embedding to query a knn index of clip image embeddings

Display captions
 Display full captions
 Display similarities
 Safe mode
 Hide duplicate urls
 Hide (near) duplicate images
Search over
 Search with multilingual clip

French cat

How to tell if your feline is french. He wears a b...

Hilarious pics of funny cats! funnycatsgif.com

イケメン猫モデル「トキ・ナンタケツ」がかっこいい - NAVER まとめ

Hipster cat

cat in a suit Georgian sells tomatoes

French Bread Cat Loaf Metal Print

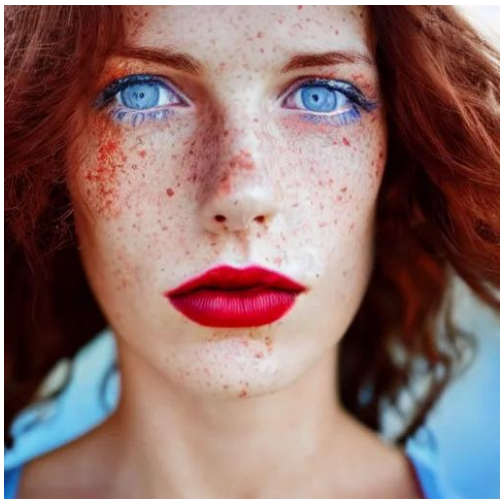
The Stable Diffusion model was trained on three subsets of LAION-5B: laion2B-en, laion-high-resolution, and laion-aesthetics v2 5+

<https://laion.ai/blog/laion-5b/>

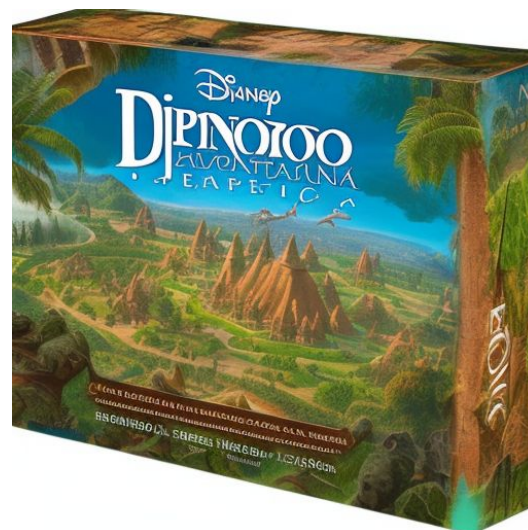


Funktionsweise: Schwächen

- Augen (nur 1.4, ab 1.5 behoben)
- Text (soll mit SDXL behoben sein)



<https://lexica.art/?prompt=ba8ad686-3b5a-49bf-9b98-d25f0225881d>



https://www.reddit.com/r/StableDiffusion/comments/z22mru/i_was_generating_some_random_dinotopia_got_way_to/



Funktionsweise: Schwächen

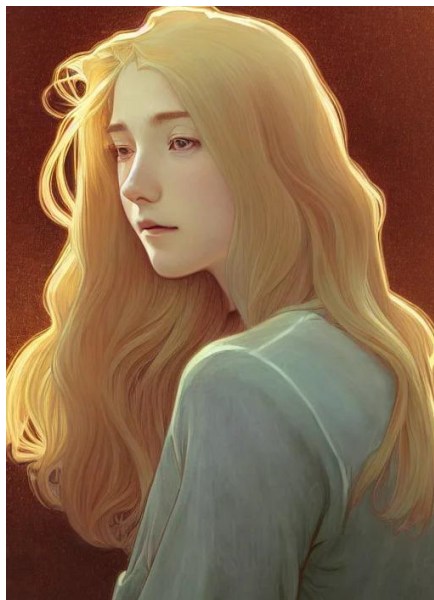
- Hände (ab 2.0 verbessert)
- Bias



"woman showing her hands"

https://www.reddit.com/r/StableDiffusion/comments/z3a4ye/prompt_woman_showing_her_hands_on_stable

Pretty young man with shoulder length shiny shimmering golden blond hair, head down, demure, shy, path traced, highly detailed, high quality, digital painting, by studio ghibli and alphonse mucha, leesha hannigan, disney



<https://lexica.art/prompt/f4d102bb-921f-4e78-909c-ac291631698a>

Prompting



Prompting: was sind Prompts?

Mit der Eingabe (Prompt) lenkt der Anwender die Software dahin, das zu generieren, was er möchte.

Häufig werden die Prompts sehr detailliert, aber auch ein kurzer Prompt kann interessante Ergebnisse liefern.

photograph of a (one massive colorful
**crystal:1.2) growing out of the rocky
mountain**, (focus on crystal:1.2), 4k, 8k,
(highly detailed),
((landscape)),(translucent crystal:1.1),
light going trough the crystal, bokeh,
chromatic aberration, mountain view

**crystal growing out of the rocky
mountain**



Prompting: Gewichtung

Gewichtung gibt einen Hinweis, welche Bildelemente wichtiger sind

photograph of a (one massive colorful crystal:**1.2**) growing out of the rocky mountain, (focus on crystal:**1.2**), 4k, 8k, (highly detailed), ((landscape)), (translucent crystal:**1.1**), light going trough the crystal, bokeh, chromatic aberration, mountain view

Mehrere Wörter können durch Klammern zu einem Bildelement zusammengefasst werden.

Mit **:1.2** wird dieses Bildelement um 20% höher gewichtet. Auch niedrigere Gewichtung ist mit **:0.8** möglich.

Mehrere Klammern wie ((landscape)) erhöhen die Gewichtung um jeweils 10%, zwei Klammern also um $1.1 * 1.1 = 1.21$

Eckige Klammern [landscape] senken die Gewichtung um 10%.

Empfohlen wird die Schreibweise (landscape:1.2), wegen der Lesbarkeit.

Ausserdem werden [drei Klammern](#) von seltsamen rechtsextremen in den USA verwendet, um antisemitische Kommentare zu machen. Deshalb kann es passieren, daß der Account gesperrt wird, wenn man Text in 3 Klammern postet



Prompting: Negative Prompts

negative Prompts geben der Software einen Hinweis, was man nicht haben möchte (Doppelte Köpfe, unscharfe Bilder etc)

negative Prompts sind bei Version 2.0 anscheinend besonders wichtig

Negative Prompt: make up, ugly, hands, blurry, low resolution, animated, cartoon, lowres, text, error, cropped, worst quality, low quality, jpeg artifacts, ugly, duplicate, morbid, mutilated, out of frame, extra fingers, mutated hands, poorly drawn hands, poorly drawn face, mutation, deformed, blurry, dehydrated, bad anatomy, bad proportions, extra limbs, cloned face, disfigured, gross proportions, malformed limbs, missing arms, missing legs, extra arms, extra legs, fused fingers, too many fingers, long neck



50 steps using Euler a, 768x768, CFG: 7.

Prompt: a beautiful ((closeup)) cinematic still image of a natural ((young)) woman as (medieval knight) standing in a forest near a ruined castle, sunset, headshot, textured skin, freckles, (((skin pores))), birthmark, auburn hair, blue eyes, dramatic lighting, Zeiss lens, 35mm film, f/2.8

Stable Diffusion 2.0



Generierung

<https://artificial-art.eu/> (kein login, viele Optionen und viele models)

powered by stable horde (crowdsourced distributed computing)

siehe [Anhang](#) und stablehorde.net

-
-
-
-

Eingabe

a beautiful ((closeup)) cinematic still image of a natural ((young)) woman as (medieval knight) standing in a forest near a ruined castle, sunset, headshot, textured skin, freckles, (((skin pores))), birthmark, auburn hair, blue eyes, dramatic lighting, Zeiss lens, 35mm film, f/2.8

Negative

Eingabe



make up, ugly, hands, blurry, low resolution, animated, cartoon, lowres, text, error, cropped, worst quality, low quality, jpeg artifacts, ugly, duplicate, morbid, mutilated, out of frame, extra fingers, mutated hands, poorly drawn hands, poorly drawn face, mutation, deformed, blurry, dehydrated, bad anatomy, bad proportions, extra limbs, cloned face, disfigured, gross proportions, malformed limbs, missing arms, missing legs, extra arms, extra legs, fused fingers, too many fingers, long neck

Seed

Legen Sie den Seed hier fest.

Sampler

k_euler

Anzahl

- 10 +

Schritte

- 30 +

Breite(1:1)

- 512 +

Höhe

- 512 +

Vorgabe

- 7 +

Clip Skip

- 1 +

Modell

[NSFW] Realistic Vision (6)

Postprozessoren

Select

Nahtlos

Karras

Reset Erstellen (10 Bilder) Abbrechen



Txt2Img

Img2Img

Copy Link

Rate Image



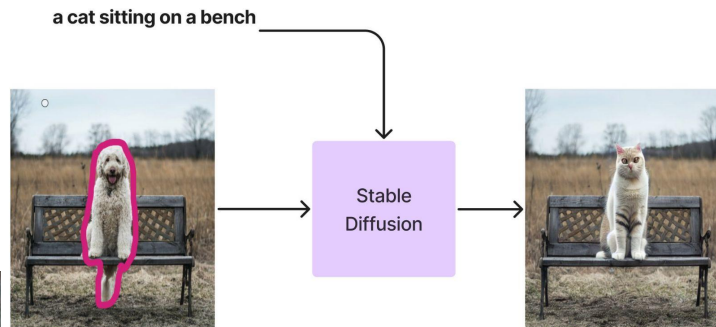
Prompting: mehr als Prompting

- Img2Img
- **Inpainting/ Outpainting**
- depth2Img (Stable Diffusion 2.0)

Outpainting: Erweitern eines Bilds



inpainting: Ersetzen eines Bildausschnittes





Prompting: mehr als Prompting

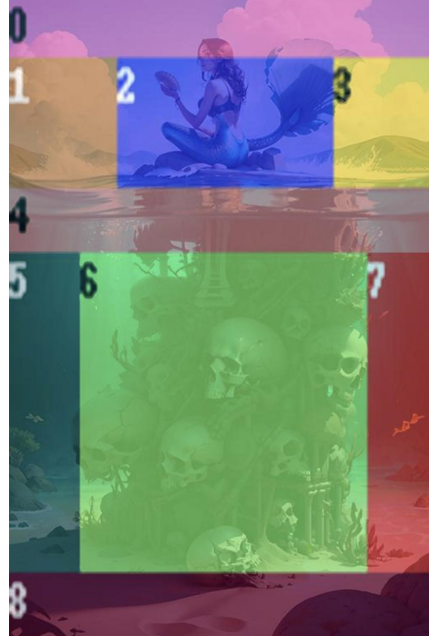
- Img2Img
- Inpainting/ Outpainting
- **depth2img** (Stable Diffusion 2.0)

The new depth-guided stable diffusion model, depth2img, extends the prior image-to-image feature from V1 with entirely new creative possibilities. Depth2img determines the depth of an input image (using an existing model) and then generates new images based on both the text and the depth information.





Prompting: region prompter extension



Wenn man versucht, einen Teil eines Bildes anzupassen (z.B. blaue Haare einer Person), kommt es vor, daß auch andere Bildelemente geändert werden, also z.B. ein Vorhang blau wird (prompt bleeding).

Völlig verschiedene Elemente in ein Bild zu bekommen, ist also ohne Nachbearbeitung nicht ganz einfach.

Diese Extension sorgt dafür, daß für die einzelnen Bereiche unterschiedliche prompts verwendet werden können.

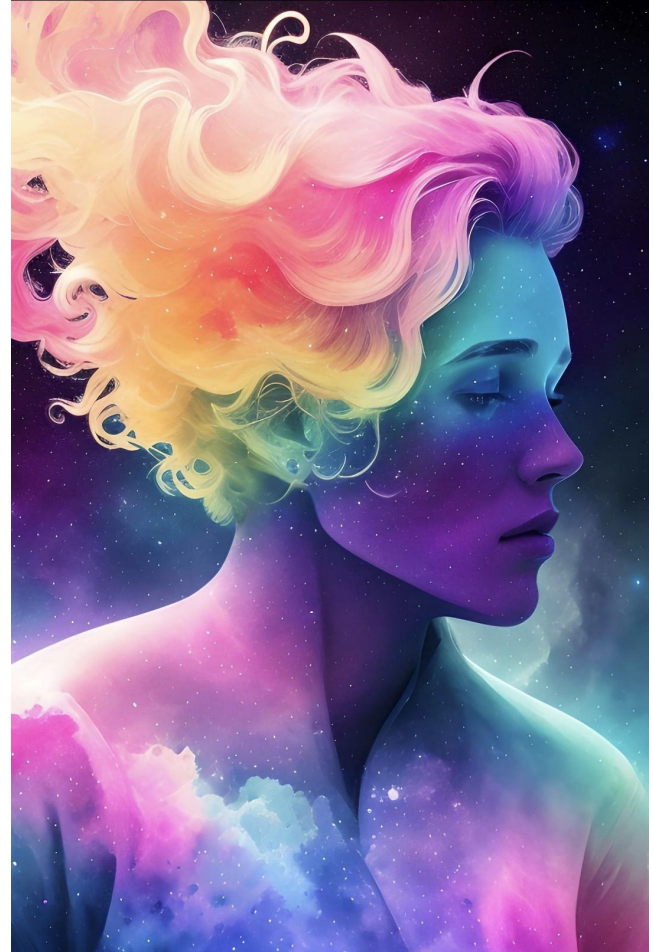
https://www.reddit.com/r/StableDiffusion/comments/13360c1/allure_of_the_lake_txt2img_region_prompter/

<https://github.com/hako-mikan/sd-webui-regional-prompter>

Einsatzmöglichkeiten

Einsatzmöglichkeiten: Kunst

https://www.reddit.com/r/StableDiffusion/comments/12fdexx/dreaming_of_the_universe





Einsatzmöglichkeiten: Musikalben



SDXL 1.0

Prompt: Cover of metal music album with evil dark duck and title text "Fear of the Duck"

<https://clipdrop.co/stable-diffusion>



Einsatzmöglichkeiten: Shirts

echtes Photo



Ada Lovelace



Einsatzmöglichkeiten: AI Actors



Quelle: https://www.reddit.com/r/StableDiffusion/comments/11qexu0/animate_your_stable_diffusion_portraits/



Einsatzmöglichkeiten: App Icon Generator



reddit.com/StableDiffusion



Einsatzmöglichkeiten: (Spiele-)entwicklung

Isometrische Teile bzw. Welten



https://www.reddit.com/r/StableDiffusion/comments/13cdl2j/my_ongoing_mission_to_create_the_perfect/



Einsatzmöglichkeiten: Comic Panels



[“Consistent comic character - test with SD Dreambooth trained on myself \(plus a couple of celeb co-stars\)”](#)



Einsatzmöglichkeiten: reverse Comic



https://www.reddit.com/r/StableDiffusion/comments/14okexp/simpsons_house



Einsatzmöglichkeiten: Image Compression



[Stable Diffusion Based Image Compression](#) by Matthias Bühlmann



Einsatzmöglichkeiten: Werbung

Generierung von Artikeln
wie z.B. Mahlzeiten für
eine Speisekarte



Lachs-Bild von:
https://www.reddit.com/r/StableDiffusion/comments/12afsuq/its_addicting_creating_food_in_sd_when_youre

Burger-Bild von: https://www.reddit.com/r/StableDiffusion/comments/z0rz2m2/we_really_enjoyed_having_lunch_at_this_restaurant/



Einsatzmöglichkeiten: Werbung (Mode)



https://www.reddit.com/r/StableDiffusion/comments/155k1ck/a_real_skirt_in_the_ai_fashion_model/

https://www.reddit.com/r/StableDiffusion/comments/157tm5w/generate_models_wearing_specific_clothes_using



Einsatzmöglichkeiten: Restauration



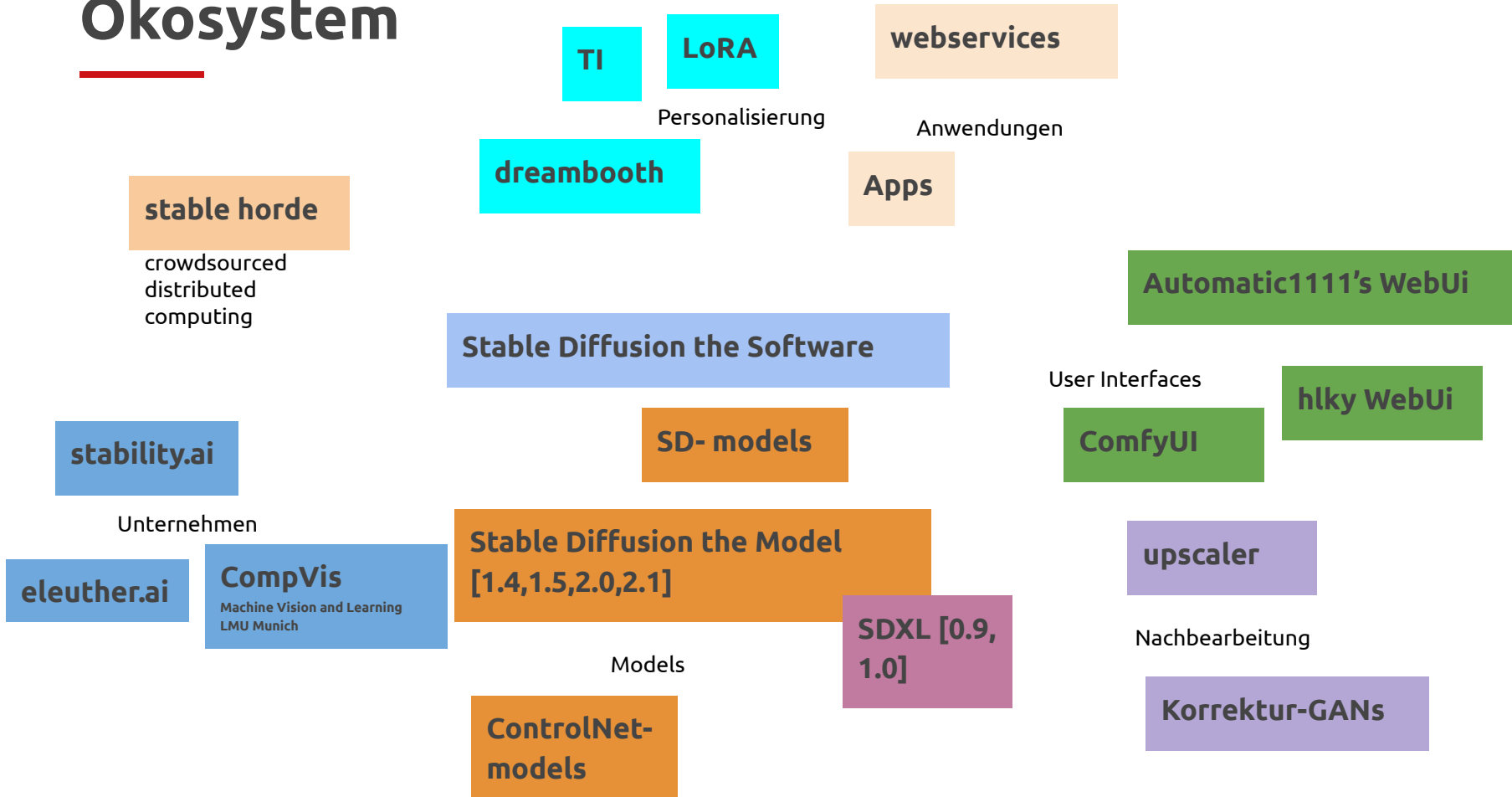
Soviet Soldier, 1946

https://www.reddit.com/r/StableDiffusion/comments/12fcqda/soviet_soldier_1946/

Ökosystem



Ökosystem



Ökosystem: models



Modern Disney

samdoesart



mehr models: siehe [Anhang](#)

Knollingcase



https://www.reddit.com/r/StableDiffusion/comments/z6eg5x/generating_porsches_with_the_knollingcase_model

https://www.reddit.com/r/StableDiffusion/comments/yhi8zo/modern_disney_lara_croft_prompt_settings_in

https://www.reddit.com/r/StableDiffusion/comments/vujief/samdoesarts_model_v2_huggingface_link_in_comments

Ökosystem: models

Model: dreamscapes & dragonfire

(Autor: DarkAgent)

<https://civitai.com/models/50294/dreamscapes-and-dragonfire-new-v20>

https://www.reddit.com/r/StableDiffusion/comments/144wsoe/goddess_of_creation/



Ökosystem: models



Model freedom

<https://civitai.com/models/87288/freedomredmond>

https://www.reddit.com/r/StableDiffusion/comments/146272w/freedom_is_here_the_generalist_21_768x_finetuned



Ökosystem: Models

Vergleich einiger Models. Mehr im [Anhang](#).

AbsoluteReality.safetensors AZovyaPhotoreal.safetensors CyberRealistic.safetensors EpicRealism.safetensors HenmixReal.safetensors ICBINP_SECO.safetensors Juggernaut.safetensors LRM.safetensors mujicMoRealistic.safetensors PatiensAbArte.safetensors Photomatix.safetensors Photon.safetensors RealisticVision.safetensors Serenity.safetensors

Seed:
2763607118



Seed:
2840926931



Seed:
3399183435



Quelle: https://www.reddit.com/r/StableDiffusion/comments/154nd8y/more_realistic_model_comparisons/



Ökosystem: ControlNet

Was ist ControlNet?

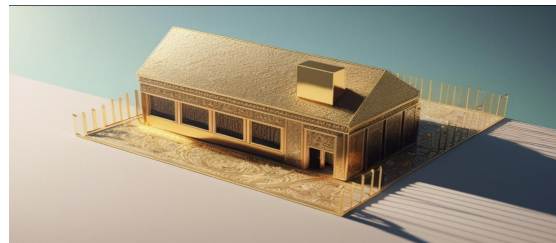
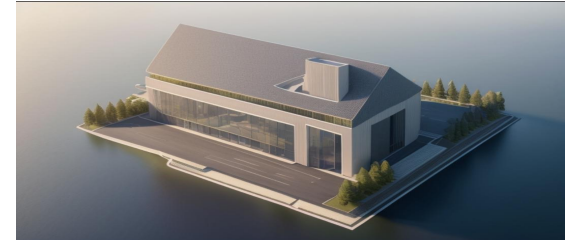
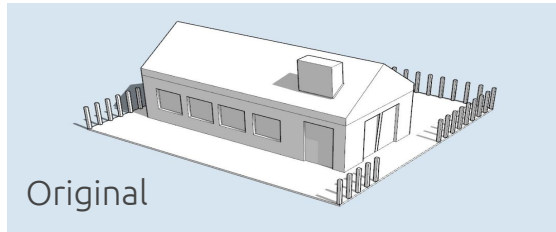
ControlNet ist ein neuronales Netzwerk, um Stable Diffusion Models genauer kontrollieren zu können.

Die einfachste Form, Models zu verwenden, ist text-to-image, wobei text prompts als conditioning (Steuerung) verwendet werden (wie vorhin gezeigt). ControlNet fügt ein conditioning hinzu.

Dieses zusätzlich conditioning kann verschiedene Ausprägungen haben, je nachdem, welches ControlNet-Model verwendet wird. ControlNet-Models werden zusammen mit dem üblichen Stable Diffusion-Model verwendet.



Ökosystem: ControlNet



Quelle: https://www.reddit.com/r/StableDiffusion/comments/12qu6u4/i_know_youre_probably_tired_of_seeing_this_but/



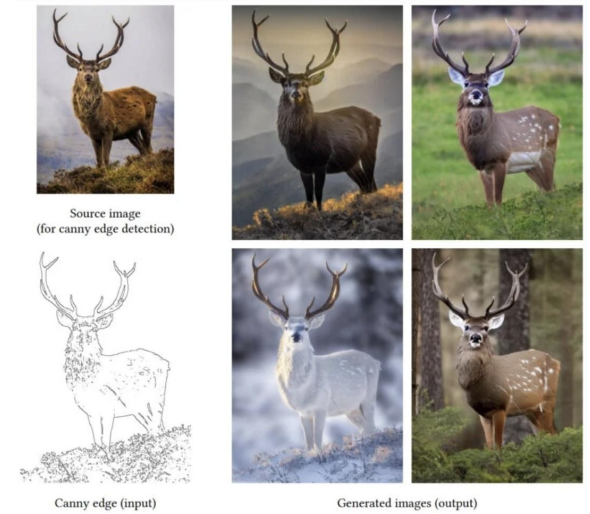
Ökosystem: ControlNet

ControlNet Models

Beispiele:

edge detection (hier: canny edge)

human pose detection (hier: openpose)

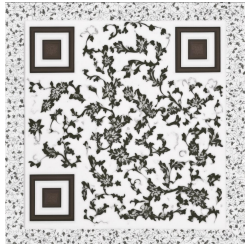
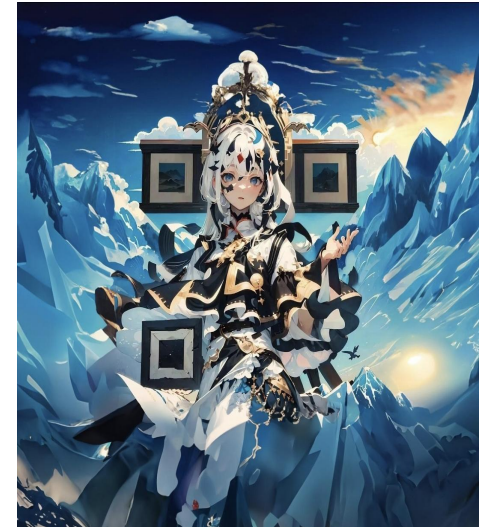




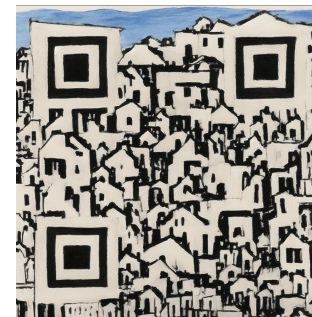
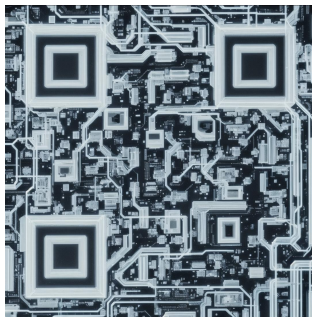
Ökosystem: ControlNet

ControlNet-Model: Controlnet for QR Code

- erstellt von <https://arbt.com/>
- noch nicht veröffentlicht
- Angekündigt:
- https://www.reddit.com/r/StableDiffusion/comments/141hg9x/controlnet_for_qr_code/



Ja, das sind funktionierende QR-Codes!





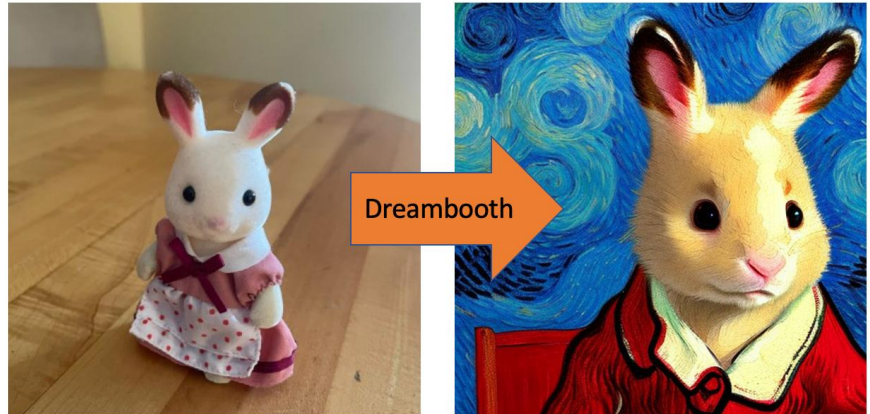
Ökosystem: Dreambooth, TI & LoRA

Stable Diffusion personalisieren

Wir können neue Konzepte hinzufügen und der KI etwas “beibringen”, ein neues Objekt oder einen neuen Stil.

- Dreambooth: mächtig, aber erzeugt neue Modelle (2-7GB)
- Textual Inversion (TI): sehr klein (< 1MB), kleine Änderungen, zusätzlich zum Model
- Low-Rank Adaptation (LoRA): mittel (20-700MB), zusätzlich zum Model

Details siehe [Anhang](#)





Ökosystem: stable diffusion XL (SDXL)

SDXL ist ein Model für stable diffusion (Also quasi Nachfolger von 1.5 und 2.1), keine "neue" Software

SDXL ist ein Zwischenschritt auf dem Weg zu stable diffusion 3

SDXL 1.0 wurde am 26.07.2023 [veröffentlicht](#).

Selber generieren:

<https://clipdrop.co/stable-diffusion>

Made with <https://clipdrop.co/stable-diffusion> (SDXL 0.9)
Prompt: a realistic blue frog with text 'FrosCon' on the back,
sitting in green grass
Type: Photo

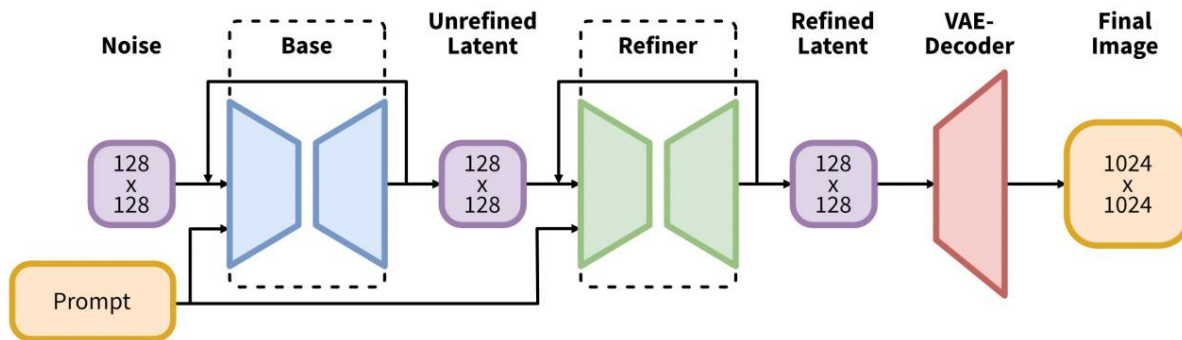




Ökosystem: SDXL

SDXL ist ein Zwischenschritt auf dem Weg zu Stable Diffusion 3.

Mehr Daten: Parameter SDXL: 6.6 Mrd., 1.5: 0.98 Mrd.



SDXL besteht aus **zwei** Modellen: *base* und *refiner* Modell.

Mehr Infos: siehe [Anhang](#)

Ökosystem: SDXL



SDXL 0.9 Base

SDXL 0.9 Refiner



SDXL 1 Base

SDXL 1 Refiner

SDXL besteht aus **zwei** Modells: *base* und *refiner* Model.

Mehr Infos: siehe [Anhang](#)

Ökosystem: SDXL



Made with <https://clipdrop.co/stable-diffusion> (SDXL 0.9)
Prompt: a duck with a sign saying Quack!
Type: No Style

Ökosystem: SDXL

Buffy the vampire slayer, shirt text reads slayer, metal festival with vampires



https://www.reddit.com/r/StableDiffusion/comments/14wog4r/sd-xl_buffy_the_vampire_slayer_shirt_text_reads

Ökosystem: SDXL

SDXL 1.0

Prompt:

Woman with black tshirt with iron maiden logo and text "fear of the duck" and a picture of an evil duck

Style: Photographic

<https://clipdrop.co/stable-diffusion>



Soziale und gesellschaftliche Auswirkungen



Copyright

Kann ich mit Stable Diffusion generierte Bilder sorgenfrei nutzen?

⚠ Wir sind keine Juristen!

- Stable Diffusion (Software + offizielles Model),
 - steht unter [“CreativeML Open RAIL-M“-Lizenz](#)
 - freie Nutzung von generierten Bildern
 - Einschränkung: “Bitte keine bösen Inhalte erstellen!”



Copyright

Viele Fragen noch offen:

- **Trainingsdaten**
 - [LAION](#) hält nur Referenzen zu öffentlich verfügbaren Bildern und Alt-Texten vor
 - Kein explizites Einverständnis der Rechteinhaber
- **Stile und Markenrecht** (“by Greg Rutkowski”, “modern/classic Disney”)
- **Prompts**

Grundsätzlich: egal, wie es erstellt wurde, man sollte besser nicht mit einem Bild “der Maus” öffentlich kommerziell Werben...



Auswirkungen

- Kommerziell
 - Wird es schwieriger, von Kunst zu leben?
 - Oder eröffnen sich ganz neue Geschäftsmodelle?
- Institutionell
 - Museen, Kunstpreise
- Ästhetisch
 - Inflation der Bilder
- Gesellschaftlich
 - *"Alles Fake!"*
 - Hyper-Individualitäts-Blasen

Fragen & Diskussion



Hinweis: Anhang

Wir haben **46 Seiten im Anhang** mit weiteren Informationen.

Diese Präsentation ist über <https://tinyurl.com/2ackd4ex> abrufbar

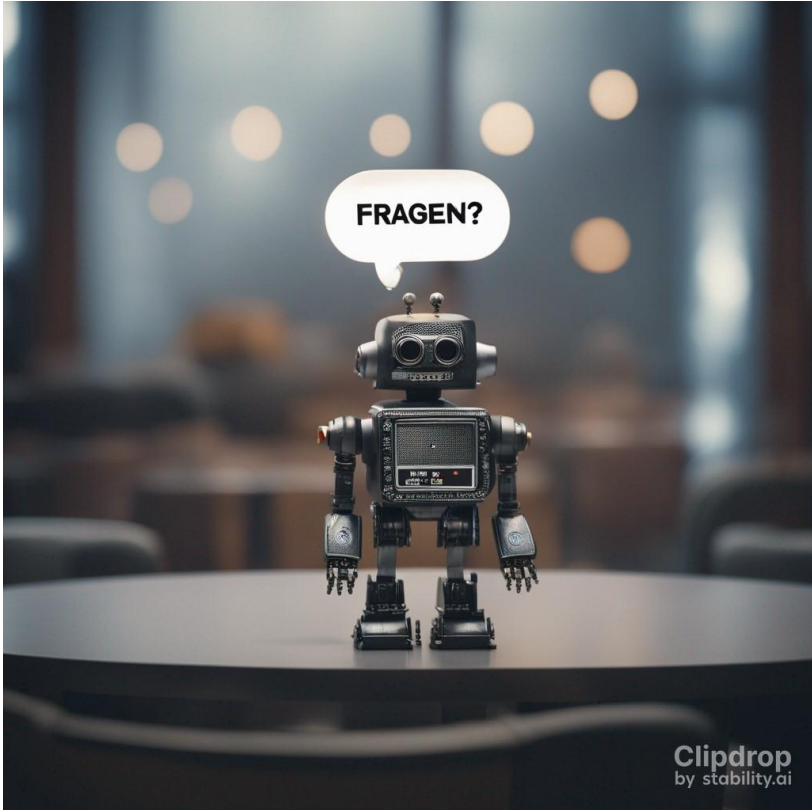


QR Code erstellt mit Taptok
<https://app.tt.social/tools/qr-generator-ai>

Prompt:
a white frog in blue water



Fragen & Diskussion



SDXL 1.0
Prompt: A robot with a speech bubble saying "Fragen?"
<https://clipdrop.co/stable-diffusion>



Vielen Dank

Wer später noch mit mir sprechen möchte, kann mich am Stand der tarent direkt am Eingang finden.

Erstellt mit SDXL 0.9 <https://clipdrop.co/stable-diffusion>

Prompt: crocodile with a sign that says See you later ...



weitere Informationen und Quellen

Anhang





Anhang: Bilder der Slideshow

Die Bilder der Slideshow am Anfang kommen von (in dieser Reihenfolge):

https://www.reddit.com/r/StableDiffusion/comments/z8jp2y/its_really_hard_developing_artist_styles_in_sd_20

https://www.reddit.com/r/StableDiffusion/comments/z9h9yo/is_this_detailed_enough_yet_sd_v2_768_a1111_fork

https://www.reddit.com/r/StableDiffusion/comments/z8gje0/hard_to_tell_these_are_not_real_women_damn

https://www.reddit.com/r/StableDiffusion/comments/z7ghbf/not_only_is_stable_diffusion_20_not_bad_but

https://www.reddit.com/r/StableDiffusion/comments/yokohg/thanks_to_nitrosocoke_for_the_fantastic

https://www.reddit.com/r/StableDiffusion/comments/yo5cla/pplease_dont_hurt_me_im_not_a_bad_slime_if_you

https://www.reddit.com/r/StableDiffusion/comments/yc17bg/ice_cream

https://www.reddit.com/r/StableDiffusion/comments/z7ko3m/using_takeon_hassanblend_and_knollingcase_to_make

https://www.reddit.com/r/StableDiffusion/comments/z97a0w/documenting_the_imaginary_architecture_of_havana

https://www.reddit.com/r/StableDiffusion/comments/yguwv6/prompt_to_create_silhouette_wallpapers_newest

https://www.reddit.com/r/StableDiffusion/comments/z9gf3x/knollingcase_is_a_great_model_for_models



Anhang: Bilder der Slideshow

neue Bilder (Juli 2023) der Slideshow am Anfang kommen von (in dieser Reihenfolge):

https://www.reddit.com/r/StableDiffusion/comments/1429jj2/its_important_to_have_good_advisors/

https://www.reddit.com/r/StableDiffusion/comments/13uerov/some_dnd_inspired_watercolor_style_portraits/

https://www.reddit.com/r/StableDiffusion/comments/136knxi/you_understand_that_this_is_not_a_photo_right/

https://www.reddit.com/r/StableDiffusion/comments/155iir2/most_realistic_image_by_accident/

https://www.reddit.com/r/StableDiffusion/comments/13b0tkk/i_know_people_like_their_waifus_but_here_are_some/

<https://www.reddit.com/r/StableDiffusion/comments/12x2m1t/midnight/>



Funktionsweise: Trainingskosten

⚠ Modelle brauchen sehr viele Ressourcen beim Training

*“We actually used **256 A100s** for this per the model card, **150k hours** in total so at market price **\$600k**”*

[\[Emad Mostaque, CEO stabiliy.ai\]](#)



Funktionsweise: Hardwareanforderungen

- **Intel/AMD Prozessor + Grafikkarte**
 - ⚠ GPU mit relativ viel VRAM (min. 8 GB für fertige Modelle, 30 GB für Modell-Verfeinerungen wie mit DreamBooth)
 - ⚠ Kann in der Cloud schwer zu bekommen sein (*I am looking at you, AWS!*)
- **Apple Silicon**
 - *It Just works*, sogar auf dem kleinen Macbook Air M1, aber ca. 60 Sekunden pro Bild
- **Mobil**
 - Hängt vom Gerät ab, moderne iPhones funktionieren

Hinweise siehe [Anhang](#)



Anhang: wie kann ich das nutzen?

Online: extrem viele Anbieter

Meine Empfehlung: stable horde (siehe nächste Seite)

- <https://huggingface.co/spaces/stabilityai/stable-diffusion>
kostenlos, lange Wartezeit, wenig Einstellmöglichkeiten
- <https://lexica.art/> Generate-Button, Anmeldung, kostenlos
- <http://beta.dreamstudio.ai/> 150 image credits free
- <https://getimg.ai/> 100 credits free (per month), with GFPGAN & Real-ESRGAN upscaling and inpainting/outpainting
- <https://openart.ai/> 100 free tokens
- <https://theartbutton.ai/> (login needed, tokens needed, refresh to 100 tokens every day)
- <https://visualise.ai/> 3 free runs, 100 tokens nach anmeldung
- <https://histre.com/integrations/generative/>
generierung online ohne login, aber die Ergebnisse werden direkt veröffentlicht



Anhang: wie kann ich das nutzen?

Online: stable horde (crowdsourced computing)

- <https://stablehorde.net/>
- <https://artificial-art.eu/> fork von dem UI, das bei aqualxx verwendet wird. no login, no tokens,..
- <https://aqualxx.github.io/stable-ui/> stable horde deployment: no login, no tokens, <60 sec, webp only..
- https://diffusionui.com/b/stable_horde
- <https://tinybots.net/artbot>
- Mastodon bot: https://sigmoid.social/@stablehorde_generator
Example: "[@stablehorde_generator](https://sigmoid.social/@stablehorde_generator) draw for me a beautiful night style: fantasy"

Meine Empfehlung: <https://artificial-art.eu/>



Anhang: wie kann ich das nutzen?

Selbst installieren: cloud

Install SD on GCP

<https://towardsdatascience.com/how-to-run-a-stable-diffusion-server-on-google-cloud-platform-gcp-c879357808bf>

Achtung: nicht einfach, an VMs mit GPU zu kommen (gilt auch für AWS)

Bei unserem Versuch hat es 2 Wochen gedauert, an GPUs zu kommen (AWS), dort waren dann aber zu wenige CPUs.

Ähnlich bei paperspace.com: Dort gibt es zwar spezielle auf Ubuntu basierende Images. Allerdings muss man größere RAM-Mengen mit Begründung anfordern:

Request Approval ×

This machine type has not been approved for this account yet. Tell us a bit more about your use case and we will prioritize your request. A member of our support team will approve your request shortly.

Note: Approvals can take 1-2 business days. Feel free to contact us for an update if it has been longer than expected.

Use case

ADD



Anhang: wie kann ich das nutzen?

Selbst installieren: lokal

GPU nötig, mindestens 4GB VRAM, besser mehr - nvidia oder apple

Run local with nvidia graphics card

<https://medium.com/geekculture/run-stable-diffusion-in-your-local-computer-heres-a-step-by-step-guide-af128397d424>

1-click install with GFPGAN & upscaling (min 4GB VRAM)

<https://github.com/cmdr2/stable-diffusion-ui>

Local, CPU only (no GPU needed)

https://www.reddit.com/r/MachineLearning/comments/x3pvqa/p_run_stable_diffusion_cpu_only_with_web/

local super-hi-res:

https://www.reddit.com/r/StableDiffusion/comments/x5wkj7/testing_even_higher_res_3200x1920_i_think_my_3090/

MacOS only

<https://diffusionbee.com/>



Anhang: wie kann ich das nutzen?

lokal auf smartphone (iOS)

maple diffusion

<https://twitter.com/amasad/status/1580772494230704128>

<https://github.com/madebyollin/maple-diffusion>

promptArt

<https://labml.ai/#promptArt>

[Draw Things: AI Generation \(ios App\)](#)



Anhang: wie kann ich das nutzen?

smartphone (Android)

Meine Empfehlung: stable horde über <https://artificial-art.eu/>

<https://www.arinteli.com/stable-diffusion-android-download/>

Berechnung läuft auf Servern

<https://play.google.com/store/apps/details?id=com.triceratop.aiapp>

anscheinend kostenlos, gibt aber IAP??

Berechnung läuft auf Servern

<https://play.google.com/store/apps/details?id=ai.pixelz.mobileApp>

tokens, kostenpflichtig

Berechnung läuft auf Servern



Prompting: Hints

<https://publicprompts.art/>

[How a Stable Diffusion prompt changes its output for the style of 1500 artists](#)

<https://mpost.io/best-100-stable-diffusion-prompts-the-most-beautiful-ai-text-to-image-prompts/>

[lexica.art](#) (wie immer)

<https://www.reddit.com/r/StableDiffusion/>

Prompting: Tricks

Delayed Keywords

https://www.reddit.com/r/StableDiffusion/comments/156s26v/delayed_keywords_is_a_nice_little_trick/

AND-Verknüpfungen im Prompt für mehrere Elemente

https://www.reddit.com/r/StableDiffusion/comments/155zidp/the_invention_of_the_razor_blade_c_5000_bc/





Prompting: Everything “by Greg Rutkowski”??

Betrifft Versionen bis 1.4 (die aber noch intensiv genutzt werden)

Um brauchbare Ergebnisse zu erzielen, fügen viele “by Greg Rutkowski” zu ihren Prompts hinzu (Ein lebender polnischer Künstler, der vor allem Fantasy bebildert, wie D&D und Magic).

Dadurch können seine eigenen Werke kaum noch gefunden werden, was er verständlicherweise nicht gut findet

“Well I guess soon I won't be able to find my own work on the internet cause it will be flooded with ai stuff.” <https://twitter.com/GrzegorzRutko14/status/1568294080756473858>

Eine mögliche Lösung wäre, einen (nicht existenten) Künstler zu nennen, dessen Nennung den gleichen Effekt hat: Sjampinjon Grzybski

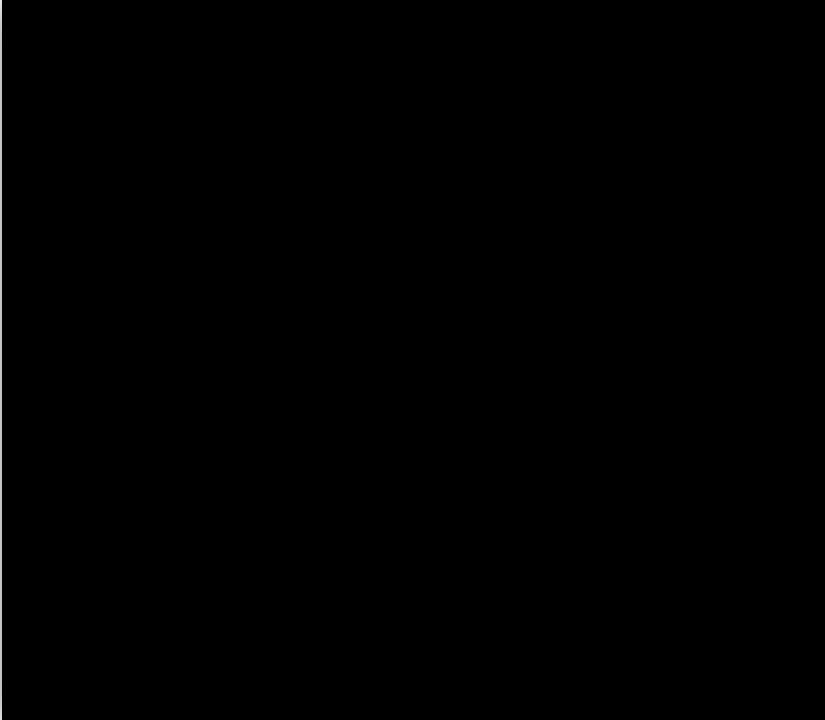
https://www.reddit.com/r/StableDiffusion/comments/xn4jnr/art_by_siampinjon_grzybski/

Oder man fügt eine Kunstrichtung statt eines Künstlers hinzu.

Hinweis: Bei stable diffusion 2.0 wurde die Verbindung zu Greg entfernt.



Einsatzmöglichkeiten: AI Actors



Quelle: [Google Colab - Thin-Plate-Spline-Motion-Model for SD.ipynb](#)



Quelle: [reddit](#)



Einsatzmöglichkeiten: Video Art

[Video “History of Earth, Life and Civilization – made with AI”](#) [1]



[1] [Wahrscheinlich](#) erstellt mit [DeForum Stable diffusion](#)



Anhang: Beispiel-Kosten

Anbieter	Produkt	Kosten
Paperspace.com	P5000	\$0.78/h
AWS	g4ad.xlarge	\$0.379/h
runpod.io	RTX A5000	\$0.490/h
vast.ai	RTX A4000	\$0.242/hr

Hinweis: diese Daten stammen aus Dezember 2022.



Anhang: Alternativen

DALL·E 2 (OpenAI)

<https://www.heise.de/hintergrund/KI-System-DALL-E-Ein-Alleskoenner-fuer-Kreative-7206468.html?hg=1&hgi=0&hgf=false>

Midjourney

Copyright: Alle von dir erstellen Bilder landen im Archiv.

Commercial Terms: <https://midjourney.gitbook.io/docs/billing#commercial-terms>

“Löscht Du Deinen Account, darfst Du die Bilder nicht mehr Dein nennen.”

<https://mizine.de/midjourney/einrichten/>

comic mit Midjourney erstellt

https://www.reddit.com/r/StableDiffusion/comments/yjqxhw/the_lesson_a_free_comic_download_made_with/

Imagen <https://imagen.research.google/>

Wombo dream <https://dream.ai>



Anhang: sozial-gesellschaftliche Auswirkungen

- <https://netzpolitik.org/2022/text-zu-bild-generierung-der-anfang-von-etwas-groessem>
- <https://arstechnica.com/information-technology/2022/09/with-stable-diffusion-you-may-never-believe-what-you-see-online-again/>
- [Is the AI spell-casting metaphor harmful or helpful?](#) [Simon Willison]
- Künstler, die generierte Bilder anpassen/aufhübschen
https://www.reddit.com/r/StableDiffusion/comments/y6lt1w/artist_with_10_years_of_experience_here_thinking/
- [Invasive Diffusion: How one unwilling illustrator found herself turned into an AI model](#) [waxy.org]



Ökosystem: 1.5 vs. 2.0

In 2.0 sind viele Bilder nicht mehr eingeflossen (NSFW-Filter) und die Zuordnung zu Künstlern wurde entfernt. Für Abbildungen von Menschen ist 2.0 nach ersten Auswertungen schlechter geeignet als 1.5, für andere Abbildungen aber besser.

Hinzugekommen sind aber ein Upscaler, depth2img, verbessertes Inpainting

Die Nutzbarkeit wurde mit 2.1 wieder verbessert, dennoch werden nach wie vor viele models eingesetzt, die auf 1.5 basieren.



“a professional photograph of a cat in a space suit” Generiert in Stable Diffusion 1.5 und 2.0



Anhang: Ökosystem: 1.5 vs. 2.0

Hinweis: theoretisch veraltet (aktuelle version ist 2.1), aber 1.5 wird immer noch intensiv verwendet (bzw. darauf basierende Models)

<https://stable-diffusion-art.com/how-to-run-stable-diffusion-2-0/>

<https://huggingface.co/spaces/fffiloni/prompt-converter>

https://www.reddit.com/r/StableDiffusion/comments/z6r79u/hmm_sd_20_is_actually_better/

https://www.reddit.com/r/StableDiffusion/comments/z6nyu0/sd_20_since_images_of_lovely_ladies_seem_to_get/

Anscheinend sind negative prompts bei 2.0 wichtiger

https://www.reddit.com/r/StableDiffusion/comments/z6ao2j/please_use_negative_prompts_with_stable_diffusion/

https://www.reddit.com/r/StableDiffusion/comments/z66wch/comparing_20_768_and_15_at_their_native/

https://www.reddit.com/r/StableDiffusion/comments/z5vqqu/beautiful_farmers_daughter_stable_diffusion_20/

https://www.reddit.com/r/StableDiffusion/comments/z637bb/sd_20/



Anhang: Ökosystem: 1.5 vs. 2.0

Hinweis: theoretisch veraltet (aktuelle version ist 2.1), aber 1.5 wird immer noch intensiv verwendet (bzw. darauf basierende Models)

https://www.reddit.com/r/StableDiffusion/comments/z5jmq0/stable_diffusion_20_has_surprisingly_good_lighting/

https://www.reddit.com/r/StableDiffusion/comments/z6apqi/how_to_generate_better_images_with_stable/

https://www.reddit.com/r/StableDiffusion/comments/z74iyj/sd_2_can_take_photorealistic_photos_good_tips_are

https://www.reddit.com/r/StableDiffusion/comments/z73jv6/living_room_stable_diffusion_20

https://www.reddit.com/r/StableDiffusion/comments/z76udu/sd_20_is_amazing_on_photorealism

https://www.reddit.com/r/StableDiffusion/comments/z7ghbf/not_only_is_stable_diffusion_20_not_bad_but

https://www.reddit.com/r/StableDiffusion/comments/z7k8nd/20_realistic_jewelry_withwithout_gems_prompt



Anhang: Ökosystem: Models

<https://huggingface.co/models?other=stable-diffusion> (> 4000 models)

<https://stablehorde.net/> (>200 Models, direkt nutzbar über die UIs)

Interessant sind z.B. “Asim Simpsons”, “realistic vision”, “mo-di-diffusion” (Pixar style), “dreamshaper”, “pixhell” (pixel art)

Dreamshaper 5 model

https://www.reddit.com/r/StableDiffusion/comments/12jt5y1/dreamshaper_5_is_here_sorry_it_took_me_a_while_i

Model pixhell https://www.reddit.com/r/StableDiffusion/comments/11v52ql/pixhell_21_sd_model

Model freedom

https://www.reddit.com/r/StableDiffusion/comments/146272w/freedom_is_here_the_generalist_21_768x_finetuned

https://www.reddit.com/r/StableDiffusion/comments/z6eq5x/generating_porsches_with_the_knollinqcase_model

<https://huggingface.co/nitrosococke/mo-di-diffusion>

https://www.reddit.com/r/StableDiffusion/comments/yh8zo/modern_disney_lara_croft_prompt_settings_in

https://www.reddit.com/r/StableDiffusion/comments/vuiejf/samdoesarts_model_v2_huggingface_link_in_comments

https://www.reddit.com/r/StableDiffusion/comments/yskhce/my_new_dd_model_trained_for_30000_steps_on_2500

https://www.reddit.com/r/StableDiffusion/comments/vxat2s/new_release_nitrodifffusion_multistyle_model_with

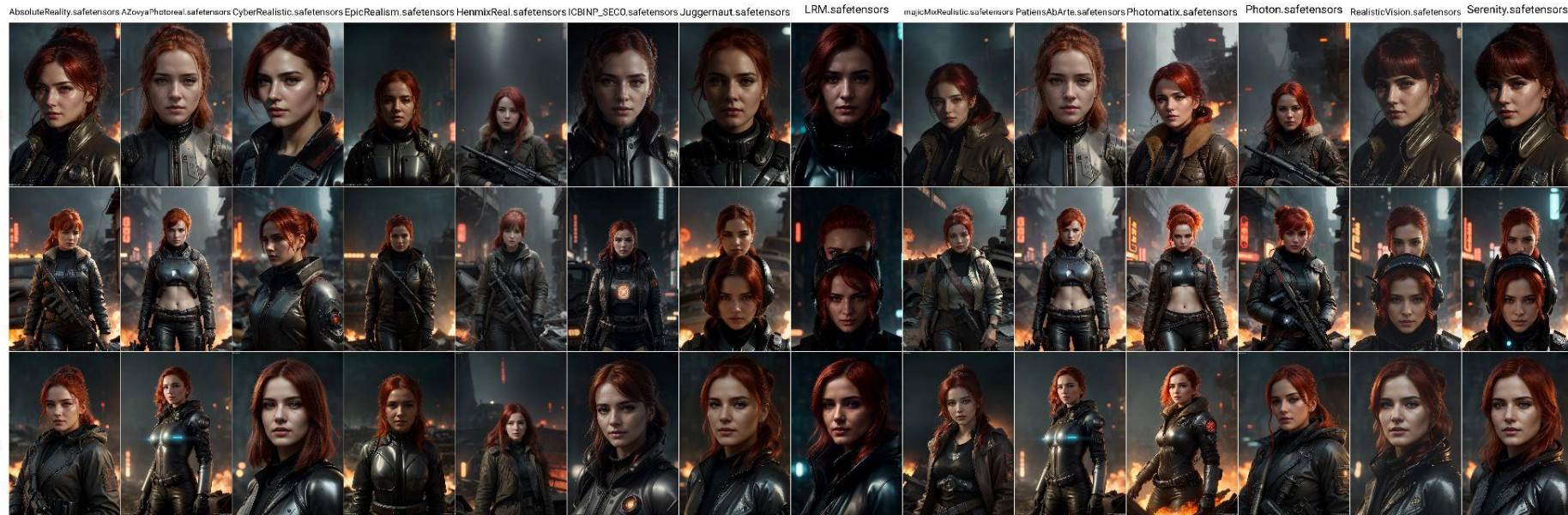
https://www.reddit.com/r/StableDiffusion/comments/z42idl/new_release_sd_20_dreambooth_model_futurediffusion

<https://huggingface.co/dallinmackay/Tron-Legacy-diffusion>

<https://huggingface.co/nitrosococke/classic-anime-diffusion>



Anhang: Ökosystem: Model-Vergleich



Quelle: https://www.reddit.com/r/StableDiffusion/comments/154nd8y/more_realistic_model_comparisons/



Anhang: Ökosystem: Model-Vergleich

AbsoluteReality.safetensors AzovyaPhotoreal.safetensors CyberRealistic.safetensors EpicRealism.safetensors HenmixReal.safetensors ICBINP_SECO.safetensors Juggernaut.safetensors LRM.safetensors mujicMixRealistic.safetensors PatiensAbArte.safetensors Photomatix.safetensors Photon.safetensors RealisticVision.safetensors Serenity.safetensors

Seed:
2493261843



Seed:
73679730



Seed:
3162725296



Quelle: https://www.reddit.com/r/StableDiffusion/comments/154nd8y/more_realistic_model_comparisons/



Anhang: Ökosystem: Model-Vergleich

AbsoluteReality.safetensors AzovyaPhotoreal.safetensors CyberRealistic.safetensors EpicRealism.safetensors HenmixReal.safetensors ICBINP_SECO.safetensors Juggernaut.safetensors LRM.safetensors majicMixRealistic.safetensors PatiensAbArte.safetensors Photomatix.safetensors Photon.safetensors RealisticVision.safetensors Serenity.safetensors

Seed:
349288523



Seed:
2543920889



Seed:
1569702187



Quelle: https://www.reddit.com/r/StableDiffusion/comments/154nd8y/more_realistic_model_comparisons/



Ökosystem: Dreambooth

[DreamBooth](#) is a method to personalize text-to-image models like Stable Diffusion given just a few (3-5) images of a subject. It allows the model to generate contextualized images of the subject in different scenes, poses, and views.



Input images



in the Acropolis



swimming



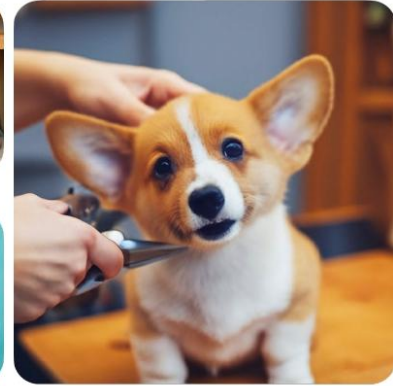
sleeping



in a doghouse



in a bucket



getting a haircut



Ökosystem: Dreambooth

Dreambooth ist die “mächtigste” Option, um ein Model zu tunen.

Aber: dadurch wird ein neues Model mit dem entsprechenden Platzbedarf geschaffen (2-7 GB)

Siehe auch:

<https://huggingface.co/docs/diffusers/training/dreambooth>

<https://stable-diffusion-art.com/dreambooth/>

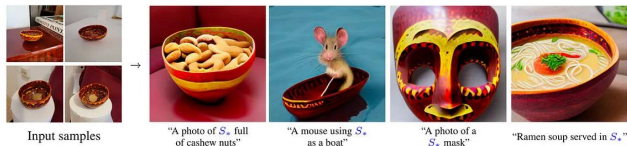




Ökosystem: Textual Inversion (TI)

Wir können neue Konzepte hinzufügen und der KI etwas “beibringen”, ein neues Objekt oder einen neuen Stil. Diesem wird ein Begriff zugeordnet, der bisher nicht verwendet wurde (“pseudo-word”).

Die Idee hinter Textual Inversion ist, dem text model mit wenigen repräsentativen Bildern ein neues Wort beizubringen.



Häufig wird TI auch für *style transfer* angewendet.

Textual Inversions sind sehr klein (< 1MB).

Siehe:

<https://medium.com/@onkarmishra/how-textual-inversion-works-and-its-applications-5e3fda4aa0bc>





Ökosystem: LoRA

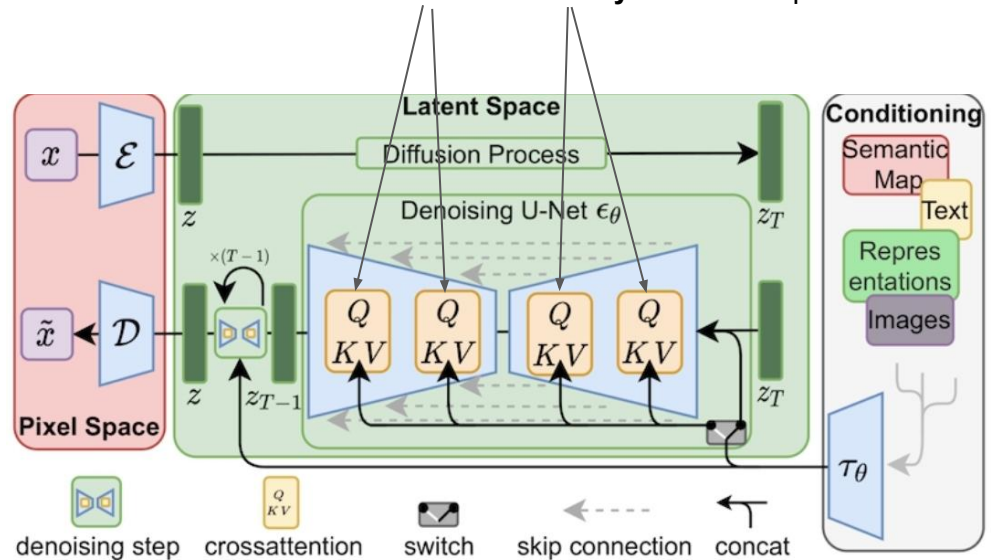
Low-Rank Adaptation (LoRA) models are small modifiers of checkpoint models

LoRA models are small Stable Diffusion models that apply tiny changes to standard checkpoint models. They are usually 10 to 100 times smaller than checkpoint models. That makes them very attractive to people having an extensive collection of models.

LoRA applies small changes to the most critical part of Stable Diffusion models: The **cross-attention layers**. It is the part of the model where [the image and the prompt meet](#).

Dreambooth is powerful but results in large model files (2-7 GBs). Textual inversions are tiny (about 100 KBs), but you can't do as much.

Siehe auch: [Anhang](#)





Anhang: Ökosystem: LoRA

Dreambooth is powerful but results in large model files (2-7 GBs). Textual inversions are tiny (about 100 KBs), but you can't do as much.

https://stable-diffusion-art.com/lora/#What_are_LoRA_models

<https://softwarekeep.com/help-center/how-to-use-stable-diffusion-lora-models>

<https://aituts.com/stable-diffusion-lora/>

wo finde ich LoRAs? <https://huggingface.co/models?other=stable-diffusion&search=lora>



Anhang: SDXL

Vergleich SDXL (base + refiner model) und 1.5 (base und finetuned models)

https://stable-diffusion-art.com/sd-xl-model/#Differences_between_SDXL_and_v15_models

<https://venturebeat.com/ai/stability-ai-announces-stable-diffusion-xl-beta-for-api-and-dreamstudio/>



Anhang: SDXL models



https://www.reddit.com/r/StableDiffusion/comments/155jx09/sdnext_kandinsky_v2_diffuser_model_not_just_sdxl

https://www.reddit.com/r/StableDiffusion/comments/1550f0s/dreamshaper_xl_alpha

https://www.reddit.com/r/StableDiffusion/comments/15694hh/sdxl_10_candidate_models_are_insane



Anhang: SDXL 1.0

Informationen und vor allem erste gute Ergebnisse mit SDXL 1.0

https://www.reddit.com/r/StableDiffusion/comments/15aapcb/sd-xl_1_0_is_out

https://www.reddit.com/r/StableDiffusion/comments/15azdjo/sd-xl_1_0_a1111_heron_lamp_designs

https://www.reddit.com/r/StableDiffusion/comments/15b67mp/sd-xl_1_0_an_ai_noob_fan_of_the_lord_of_the_rings

https://www.reddit.com/r/StableDiffusion/comments/15b7oqt/a_different_perspective_on_beauty_sd-xl_1_0

https://www.reddit.com/r/StableDiffusion/comments/15b5dxq/stable_diffusion_incorporate_text_in_image

https://www.reddit.com/r/StableDiffusion/comments/15ap2hg/sd-xl_1_0_artist_reference_sheet_with_rabbits

https://www.reddit.com/r/StableDiffusion/comments/15b2doz/sd-xl_base_refiner_1_0_gothic_cyberpunk_portraits

https://www.reddit.com/r/StableDiffusion/comments/15at3ko/first_results_sd-xl_1_0

https://www.reddit.com/r/StableDiffusion/comments/15be7jg/sd-xl_1_0_a1111_what_a_difference_refine_makes



Anhang: Ökosystem: Web UIs

- automatic's webUI
<https://github.com/AUTOMATIC1111/stable-diffusion-webui>
- hlky / stable-diffusion-webui
<https://github.com/Syqil-Dev/syqil-webui>

Hinweis: es gibt inzwischen mehr UIs. Am verbreitetsten ist nach wie vor das von automatic1111



Ökosystem: die Anderen

- [DALL-E 2](#)
 - proprietärer Cloud Service
 - OpenAI (Elon Musk, Sam Altman, Microsoft)
 - Diffusion Model
- [Imagen](#)
 - proprietärer Cloud Service
 - Google
 - Diffusion Model
- [MidJourney](#)
 - proprietärer Cloud Service
- [Wombo Dream](#)
 - proprietärer Cloud Service

Beispiel-Bild von Midjourney



https://www.reddit.com/r/StableDiffusion/comments/z2xhvq/scottish_landscapes/



Anhang: Ökosystem: DeepFloydIF

- DeepFloydIF
 - open source
 - stabilityAI (wie SD)
 - noch in Entwicklung
 - Diffusion Model
 - Text-fähigkeit



Anhang: stable horde

stable horde (crowdsourced distributed computing)

stablehorde.net

stable horde: Freiwillige stellen ihre Rechenkapazität zur Verfügung, die alle nutzen können. Kann ohne Anmeldung genutzt werden. Mit Anmeldung können "Kudos" verdient werden (vor allem durch Bereitstellung von Rechenkapazität), die für "teure" Generationen verwendet werden können (höhere Auflösung, viele Steps, besondere models...)

Seiten zum Generieren: <https://aqualxx.github.io/stable-ui/> <https://artificial-art.eu/>

Aufruf zur Unterstützung:

https://www.reddit.com/r/StableDiffusion/comments/12tultf/reminder_you_can_help_everyone_and_especially/

Zusammenarbeit stable Horde und LAION: <https://laion.ai/blog/laion-stable-horde/>



Anhang: Schwächen: Hände

2.0 kann Hände ordentlich darstellen

https://www.reddit.com/r/StableDiffusion/comments/z7salo/with_the_right_prompt_stable_diffusion_20_can_do

SDXL 0.9 bekommt das (ohne Erweiterungen und Nachbearbeitung) nicht besonders gut hin, für das Bild unten habe ich 200 Versuche gebraucht:



Generiert mit <https://clipdrop.co/stable-diffusion>
prompt: spock vulcan salute





Anhang

Prompt: "[Climate change report on wooden table](#), [readable text](#)"

Das ist teilweise überflutetes Land, links unten steht anscheinend "climate change" und es ist auf einem Tisch...

generiert mit [lexica.art](#) (November 2022), SD 1.5

<https://lexica.art/prompt/5c031846-5f8f-4d60-92c9-41db095bbed0>





Anhang: Funktionsweise: Anwendungssicht

Software + Model + Parameter + Prompt = Bild-Output



- Seed
- Sampler
- Bilddimensionen
- Steps
- Guidance

Bei **identischem Input** wird deterministisch immer **derselbe Output** generiert.

Hinweis: es kann sein, daß trotz gleichem input unterschiedliche Ergebnisse erzeugt werden, wenn sich die Hardware unterscheidet. Das konnte ich nicht verifizieren, daher habe ich diese Folie in den Anhang verschoben



Anhang depth2img vs. instructpix2pix

Zwei Modelle, die ermöglichen, die Komposition eines Bilder aufrechtzuerhalten und mit einem Prompt einzelne Elemente zu ersetzen

<https://stable-diffusion-art.com/depth-to-image/>

<https://stable-diffusion-art.com/instruct-pix2pix/>



Anhang: Workflow für Ada

Beginning with the workflow from here https://www.reddit.com/r/StableDiffusion/comments/10noz4f/sorry_had_to_post_this_a_good_prompt_a_good_model/

I fiddled around and after about 100 generations I finally got this.

Prompt: Vibrant neon colors comic, Portrait (woman) (30 year old woman), (((ada lovelace))), (long hair), Smirk, realistic shaded, elegant, award winning half body portrait of (((ada lovelace))) with flower in ((dark hair)), (rococo hair style), shaded flat illustration, comic sketch

Negative Prompt: highly detailed, fine detail, intricate, poorly drawn, crippled, crooked, broken, weird, odd, distorted, (big breasts), (big tits), erased, cut, mutilated, sloppy, hideous, ((ugly)), pixelated, ((bad hands)), aliasing, lowres, (monochrome), (black and white), ((b&w)), poorly drawn, sloppy, over exposed, over saturated, burnt image, sloppy, broken, fuzzy, aliasing, cheap, oldschool, poor quality, pixelated, sleepy, closed-eyes, lowres, pixelated, aliasing, old, granny, ugly, ((bad anatomy)), hideous, deformed, mutant, butchered, gore, sloppy, artifacts, mutilated, poorly drawn, smudged, pencil, glossy skin, doll, plastic, (signature), (watermark), (words), (letters), (logo), (username), ((disfigured)), ((close up))

Sampler: k_euler_a

Model: realistic vision

Postprocessor: RealESRGAN_4xplus

Size 512x512

Steps 30

Guidance 7

Karras enabled

Seed: 400424973

Made with <https://aqualxx.github.io/stable-ui/>



Anhang: Teaser des Vortrags

Titel: Text-zu-Bild-KI "Stable Diffusion" - Revolution der Bilder?

Short: "Stable Diffusion" generiert Bilder aus Texteingaben und ist seit Ende August 2022 für alle quelloffen verfügbar. Stehen wir vor einer Revolution der Bilder und dem „Ende der Kunst“?

Wir diskutieren Funktionsweise und Ökosystem, Copyright, soziale und gesellschaftliche Auswirkungen sowie die Kunst des "Prompting".

Beschreibung:

Der Vortrag soll einen Einstieg und die nötigen Hinweise geben, um selbst Bilder generieren zu können.

Themen sind Funktionsweise, Varianten wie Text-to-Image, Image-to-Image und ControlNet, Auswirkung von Modellen und Seeds sowie das Prompting. Es geht auch um das rasant wachsende Ökosystem und die möglichen Einsatzmöglichkeiten. Die aktuellen Entwicklungen und zu erwartenden Neuerungen sowie gesellschaftliche Auswirkungen sind weitere Themen.

Um direkt selbst einsteigen zu können, zeige ich die Nutzung von stable horde (crowdsourced distributed computing)

Der Vortrag ist eine aktualisierte Version des Vortrags, den wir im Dezember gehalten haben (<https://www.youtube.com/watch?v=ahvO05zEbf4>).



Denkanstoß

“This release is the culmination of many hours of collective effort to create a single file that compresses the visual information of humanity into a few gigabytes.”

[\[Stable Diffusion Public Release Announcement, stability.ai Blog\]](#)



Anhang: vorheriger Vortrag

Diesen Vortrag hatte ich zusammen mit meinem Kollegen Martin Lukas im Dezember 2022 in der inhouse-Konferenz "Tech Friday" bereits einmal gehalten.

Die damaligen Folien sind hier abrufbar: tinyurl.com/48w4vfm2

Das Video des damaligen Vortrags ist hier zu finden: <https://youtu.be/ahvO05zEbf4>



Ökosystem: Geschichte

Veraltet! Stand Dezember 2022.

Seitdem hat sich sehr viel getan (SD 2.1, Uls, SDXL, ControlNet ...)

22.08.2022 [Release Stable Diffusion](#) (1.4), Start Entwicklung [Automatic1111](#)

24.08.2022 [Lexica](#)

06.09.2022 Start Entwicklung [SD-Dreambooth](#)

20.10.2022 [Runway brings in/outpainting to level of Dall-E 2](#)

21.10.2022 [Stable Diffusion 1.5](#)

27.10.2022 Model "[Modern Disney](#)" (Pixar Style)

Explosionartige Vermehrung von neuen Models, Model-merges

24.11.2022 [Stable Diffusion 2.0](#)

Hier gibt es eine interessante Zusammenfassung der Geschichte von SD

https://www.reddit.com/r/StableDiffusion/comments/154p01c/before_sd1_new_era_starts_can_we_make_a_summary