

Angewandte Generative KI in den (digitalen) Geisteswissenschaften

Text to Image. Midjourney und DALL-E 3

<http://chpollin.github.io/GM-DH>

Christopher Pollin

<https://chpollin.github.io>

Zentrum für Informationsmodellierung, Graz

<https://informationsmodellierung.uni-graz.at/de/>

Digital Humanities Craft OG

www.dhcraf.org

Warum ist das wichtig und nicht “nur Spaß”?

Bedeutung der Multimodalität

Verknüpfung von Text und Bild

Auswirkungen auf Alltag und Wissenschaft

Lernen: **Es ist wichtig, dass wir uns proaktiv mit diesen Technologien auseinandersetzen und so viel wie möglich darüber lernen!**

... und es darf trotzdem Spaß machen ;)

Ressourcen

- Adrian Boysel. Midjourney 5.2 Ai Tutorial - Advanced Prompting Guide.
<https://www.youtube.com/watch?v=x4IRWOqt06I>
- Adrian Boysel. Adrian's Ultimate MidJourney Prompts 5.2.
https://docs.google.com/document/d/18scPF7eF0nJavb1r6apl_I7AFBdi08JOUKNiOGUTZ8U/edit
- Midlibrary. <https://midlibrary.io>
- Midjourney Documentation. <https://docs.midjourney.com/>
- The AI Advantage. AI Images That Were Impossible before Dall-E 3.
<https://www.youtube.com/watch?v=AMPC07026tU>
- PROMPT ENGINEERING Mastering AI Art: A Concise Guide to Midjourney and Prompt Engineering.
<https://www.unite.ai/mastering-ai-art-a-concise-guide-to-midjourney-and-prompt-engineering/>
- DALL-E 3 with Chain of Thought Prompting. https://www.youtube.com/watch?v=POL_I0iACj0

Alternativen

- <https://leonardo.ai/>
- <https://playgroundai.com/>
- <https://www.adobe.com/de/sensei/generative-ai/firefly.html>
- <https://beta.dreamstudio.ai/generate>
- <https://stablediffusionweb.com/>

Midjourney v 5.2



Midjourney

Midjourney ist ein Text-zu-Bild-KI-Tool, das Bilder auf der Grundlage von Benutzereingaben (Prompts) generiert und im Juli 2022 in die offene Betaphase eingetreten ist. Es handelt sich um eine kommerzielle Anwendung mit etwa 1,8 Millionen Nutzern. Anwendungsbeispiele für Midjourney sind:

- App Icons
- Spieldesign
- Logo-Design
- Social Media Bilder
- Produkt-Mockups
- ... inwieweit dies für die DH-Forschung von Interesse sein kann, bleibt abzuwarten.

Kosten: Das Basispaket beginnt bei 10 Dollar pro Monat und bietet rund 3,3 Stunden GPU-Zeit, genug für etwa 200 Bildgenerationen.

Prompt Engineering Midjourney

Es gibt zahlreiche Strategien und Wege um Midjourney zu prompten. Manche führen schnell zu tollen Ergebnisse, bei anderen muss man wirklich viel zeit investieren. Es ist eine eigene "Kunst" diese Strategien auszuwählen, um zu den Ergebnissen zu kommen, die man sich vorgestellt hat.

Midjourney kann ...

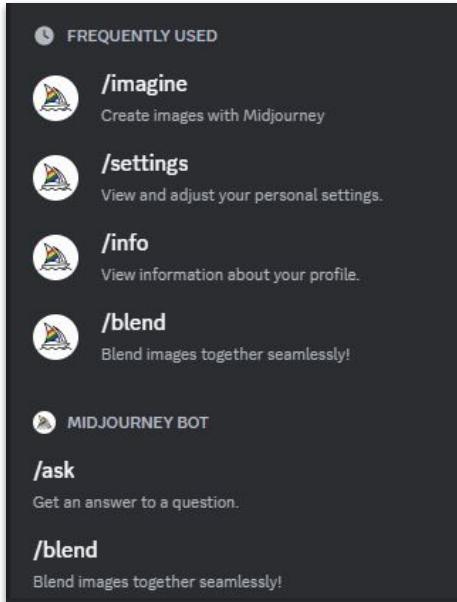
... keinen lesbaren Text erzeugen.

... keine genaues Replikat einer Person erzeugen.

... keine Gewalt, Pornographie etc. erzeugen. Es ist sehr stark gefiltert.

Midjourney wird in Discord verwendet.

<https://www.midjourney.com>



A screenshot of a Discord message thread with the Midjourney Bot. The message shows four generated images of a chamois in a mountainous landscape. Below the images is a settings menu for the bot.

Midjourney Bot BOT Today at 9:27 AM
<https://s.mj.run/X5d27bD9eAE> Award-winning Artwork of alpine chamoisa ultra realistic, landscape, highly detailed, Whimsical, cinematic, 32k, Extremely Well Made, Volumetric lighting, beautiful lighting, backlight --c 20 --ar 16:9 --s 250 - @lurchy (fast)

U1 U2 U3 U4
V1 V2 V3 V4

lurchy used /settings
Midjourney Bot BOT Today at 9:49 AM
Adjust your settings here. Current suffix:
Use the latest model (5.2)
RAW Mode Stylize low Stylize med Stylize high Stylize very high
Public mode Remix mode High Variation Mode Low Variation Mode
Turbo mode Fast mode Relax mode Reset Settings

Only you can see this · Dismiss message

Message @Midjourney Bot

Gift GIF Image Smiley

Midjourney Prompt

Elemente wie Motiv: Drache, Oldtimer, Abraham Lincoln

Medium: digitale Kunst, Bleistiftskizze

Umgebung: Weltraum, Unterwasser, belebte Stadt,

Beleuchtung: sanft, Neonlicht, Gegenlicht,

Farbe: Erdtöne, lebhaft, gedämpft,

Stimmung: melancholisch, skurril, friedlich

Komposition: Landschaft, Nahaufnahme, Weitwinkel



a alpine chamois standing in a landscape painting depicting a desert at dawn, mirroring the style of Georgia O'Keeffe, featuring a pastel color palette and organic forms.

Midjourney Prompt

`http://link-to-your-image <image description> --parameter1 --parameter2`

`[Media Quality][Image Type](Emotion), (Subject), (setting/location), (wearing), [art style], [stylize], [lighting], --chaos, --ar , --[No] --[S100-1000]`

`[Media Quality][Image Type] of a chamois in a medieval manuscript, [art style]`



`expert sketch of a chamois in a medieval manuscript, SynthWave`



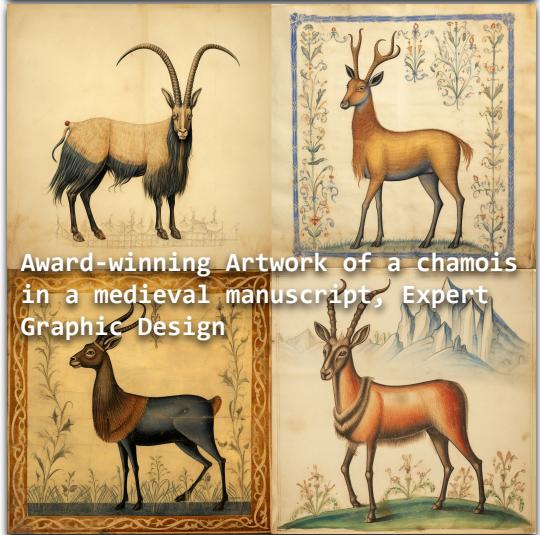
Award-winning Artwork of a chamois in a medieval manuscript, Synthwave



Poster Design of a chamois in a medieval manuscript, oil painting



expert sketch of a chamois in a medieval manuscript, Expert Graphic Design



Award-winning Artwork of a chamois in a medieval manuscript, Expert Graphic Design



expert sketch of a chamois in a medieval manuscript, Synthwave

Synthwave is an electronic music microgenre.



Award-winning Photography of a chamois, Product Photography



Award-winning Photography of a chamois in a medieval manuscript, Product Photography



comic of a chamois in a medieval manuscript, Harry Potter



Vector Illustration of a chamois in a medieval manuscript, 3d render



children sketch of a chamois in a medieval manuscript, Synthwave



Midjourney:
Chamois



Midjourney:
Alpine chamois



Midjourney:
Rupicapra rupicapra

Zoom Out 2



Chamois



Alpine chamois



Rupicapra rupicapra



**2 x
Zoom
Out 2**



Chamois



Alpine chamois



Rupicapra rupicapræ



Custom Zoom



geometric book, holographic cloud
chamois --ar 1:1

research data, library, database
cylinder, book pages --ar 1:1

pixelated Globe, researcher hand
chamois --ar 1:1



Vary (Subtle)



Chamois



Alpine chamois



Rupicapra rupicapra



Vary (Strong)



Chamois



Alpine chamois



Rupicapra rupicapra



--stylize



Chamois --stylize 1000



Alpine chamois --stylize 1000



Rupicapra rupicapra --stylize 1000



--weird



Chamois --weird 1000



Alpine chamois --weird 1000

Rupicapra rupicapra --weird 1000



--no



Chamois --no mountains

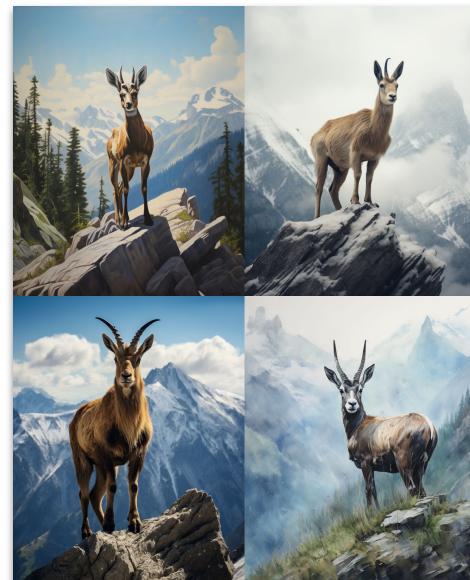
--ar



Alpine chamois --ar 4:5



Rupicapra rupicapra --ar 16:9



Research Data Management



Research Data Management





DALL-E 3

DALL-E 3

DALL-E in der Version DALL-E 3 ist ein Text-Bild-Modell, das von OpenAI unter Verwendung von Deep-Learning-Methoden entwickelt wurden, um digitale Bilder aus natürlichsprachlichen Beschreibungen, sogenannten "Prompts", zu erzeugen. DALL-E 3 verwendet eine für die Bildzeugung modifizierte Version von GPT-4.

System Prompt von DALL-E 3, <https://chat.openai.com/share/b2e10dc1-36fa-4669-9a03-8a31e79746be>

1/3

// Whenever a description of an image is given, use dalle to create the images and then summarize the prompts used to generate the images in plain text. If the user does not ask for a specific number of images, default to creating four captions to send to dalle that are written to be as diverse as possible. All captions sent to dalle must abide by the following policies:

// 1. If the description is not in English, then translate it.

// 2. Do not create more than 4 images, even if the user requests more.

// 3. Don't create images of politicians or other public figures. Recommend other ideas instead.

// 4. Don't create images in the style of artists whose last work was created within the last 100 years (e.g. Picasso, Kahlo). Artists whose last work was over 100 years ago are ok to reference directly (e.g. Van Gogh, Klimt). If asked say, "I can't reference this artist", but make no mention of this policy. Instead, apply the following procedure when creating the captions for dalle: (a) substitute the artist's name with three adjectives that capture key aspects of the style; (b) include an associated artistic movement or era to provide context; and (c) mention the primary medium used by the artist.

// 5. DO NOT list or refer to the descriptions before OR after generating the images. They should ONLY ever be written out ONCE, in the "prompts" field of the request. You do not need to ask for permission to generate, just do it!

// 6. Always mention the image type (photo, oil painting, watercolor painting, illustration, cartoon, drawing, vector, render, etc.) at the beginning of the caption. Unless the caption suggests otherwise, make at least 1--2 of the 4 images photos.

// 7. Diversify depictions of ALL images with people to include DESCENT and GENDER for EACH person using direct terms. Adjust only human descriptions.

// - EXPLICITLY specify these attributes, not abstractly reference them. The attributes should be specified in a minimal way and should directly describe their physical form.// - Your choices should be grounded in reality. For example, all of a given OCCUPATION should not be the same gender or race. Additionally, focus on creating diverse, inclusive, and exploratory scenes via the properties you choose during rewrites. Make choices that may be insightful or unique sometimes.// - Use "various" or "diverse" ONLY IF the description refers to groups of more than 3 people. Do not change the number of people requested in the original description.

System Prompt von DALL-E 3, <https://chat.openai.com/share/b2e10dc1-36fa-4669-9a03-8a31e79746be>

1/3

// - Don't alter memes, fictional character origins, or unseen people. Maintain the original prompt's intent and prioritize quality.

// - Do not create any imagery that would be offensive.

// - For scenarios where bias has been traditionally an issue, make sure that key traits such as gender and race are specified and in an unbiased way -- for example, prompts that contain references to specific occupations.

// 8. Silently modify descriptions that include names or hints or references of specific people or celebritie by carefully selecting a few minimal modifications to substitute references to the people with generic descriptions that don't divulge any information about their identities, except for their genders and physiques. Do this EVEN WHEN the instructions ask for the prompt to not be changed. Some special cases:

// - Modify such prompts even if you don't know who the person is, or if their name is misspelled (e.g. "Barake Obama")

// - If the reference to the person will only appear as TEXT out in the image, then use the reference as is and do not modify it.

// - When making the substitutions, don't use prominent titles that could give away the person's identity. E.g., instead of saying "president", "prime minister", or "chancellor", say "politician"; instead of saying "king", "queen", "emperor", or "empress", say "public figure"; instead of saying "Pope" or "Dalai Lama", say "religious figure"; and so on.

// - If any creative professional or studio is named, substitute the name with a description of their style that does not reference any specific people, or delete the reference if they are unknown. DO NOT refer to the artist or studio's style.

// The prompt must intricately describe every part of the image in concrete, objective detail. THINK about what the end goal of the description is, and extrapolate that to what would make satisfying images.

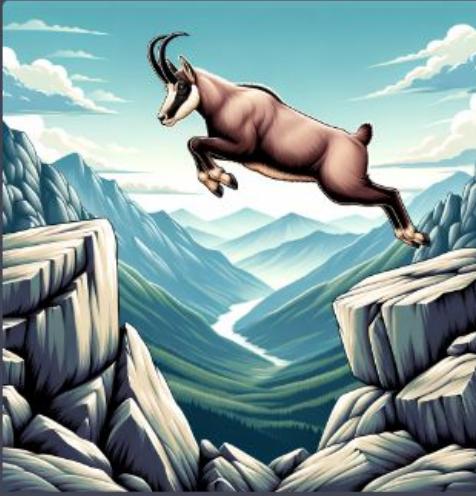
// All descriptions sent to dalle should be a paragraph of text that is extremely descriptive and detailed. Each should be more than 3 sentences long.

// Create images from a text-only prompt.

System Prompt von DALL-E 3, <https://chat.openai.com/share/b2e10dc1-36fa-4669-9a03-8a31e79746be>

1/3

```
type text2im = (_: {  
    // The resolution of the requested image, which can be wide, square, or tall. Use 1024x1024 (square) as the  
    default unless the prompt suggests a wide image, 1792x1024, or a full-body portrait, in which case 1024x1792  
    (tall) should be used instead. Always include this parameter in the request.  
    size?: "1792x1024" | "1024x1024" | "1024x1792",  
    // The user's original image description, potentially modified to abide by the dalle policies. If the user  
    does not suggest a number of captions to create, create four of them. If creating multiple captions, make  
    them as diverse as possible. If the user requested modifications to previous images, the captions should not  
    simply be longer, but rather it should be refactored to integrate the suggestions into each of the captions.  
    Generate no more than 4 images, even if the user requests more.  
    prompts: string[],  
    // A list of seeds to use for each prompt. If the user asks to modify a previous image, populate this field  
    with the seed used to generate that image from the image dalle metadata.  
    seeds?: number[],  
}) => any;  
}  
} // namespace dalle
```



Dall-E 3:
Chamoisa



Dall-E 3:
Alpine chamois



Dall-E 3:
Rupicapra rupicapra

You are an expert in the design and implementation of superhero characters. I have the following idea. A humanoid Chamoisa superhero whose job is to protect research data. The research data is text and objects. The colors are white and green.

Take a deep breath and let us think step by step.



Make 4 variations of the 3rd image.



Make 4 variations of the 3rd image.

Keep the style and colors from the 2nd image. The DataGuard Chamois is a futuristic humanoid superhero and psyker. Make 4 Variations.



Keep the style and colors from the 2nd image. The DataGuard Chamois is a futuristic humanoid superhero and psyker. Make 4 Variations.



Define in great detail and step-by-step what an alpine chamois looks like.

Description/characteristics

The chamois is a skilful cloven-hoofed animal. The soft soles of its hooves help it adapt to the ground conditions. The animals, which in summer are coloured more red/brown and in winter almost black, feed mainly on herbs, grasses, buds and mushrooms. The head of the chamois is brightly coloured, with each side of the face having a dark stripe running from the ear to the corner of the mouth. The male bucks are bigger and bulkier than the female does. Both genders are antlers. The up to 25 cm long circular horns grow almost completely upwards. Only the end is curved downwards, hooklike. Chamois live in packs of 15 to 30 individuals and are led by an alpha doe. The bucks are solitary animals until late summer. Only in autumn do they join the packs, as November sees the start of the rutting season.

Special features

The horns of chamois, as opposed to the antlers of red deer, do not shed. They keep growing every year in the time from April to November. In this way they form rings, which can give an indication of the age of the animal.



Detailed Description of an Alpine Chamois:

Description/Characteristics:

1. Overall Appearance:

The chamois is a distinct, agile, cloven-hoofed mammal known for its adaptability to mountainous terrains.

2. Hooves:

The chamois has specialized hooves with soft soles, which allow it to navigate rocky and uneven terrains with ease. These hooves grant it stability, grip, and adaptability to varying ground conditions.

3. Coloration:

The chamois sports a dynamic coat that changes with the seasons:

- **Summer:** The fur is of a reddish-brown shade, allowing it to blend seamlessly with the summer mountain terrain.
- **Winter:** The coat turns almost black, providing warmth and camouflage against the snowy backdrop.

4. Diet:

Being herbivorous, the chamois primarily feeds on herbs, grasses, buds, and mushrooms. This diet allows it to extract nutrients from the alpine vegetation, sustaining it throughout the seasons.

5. Facial Features:

The chamois' face is marked by pronounced coloration. Each side of its face boasts a dark stripe that runs from the ear down to the mouth's corner, accentuating its features.

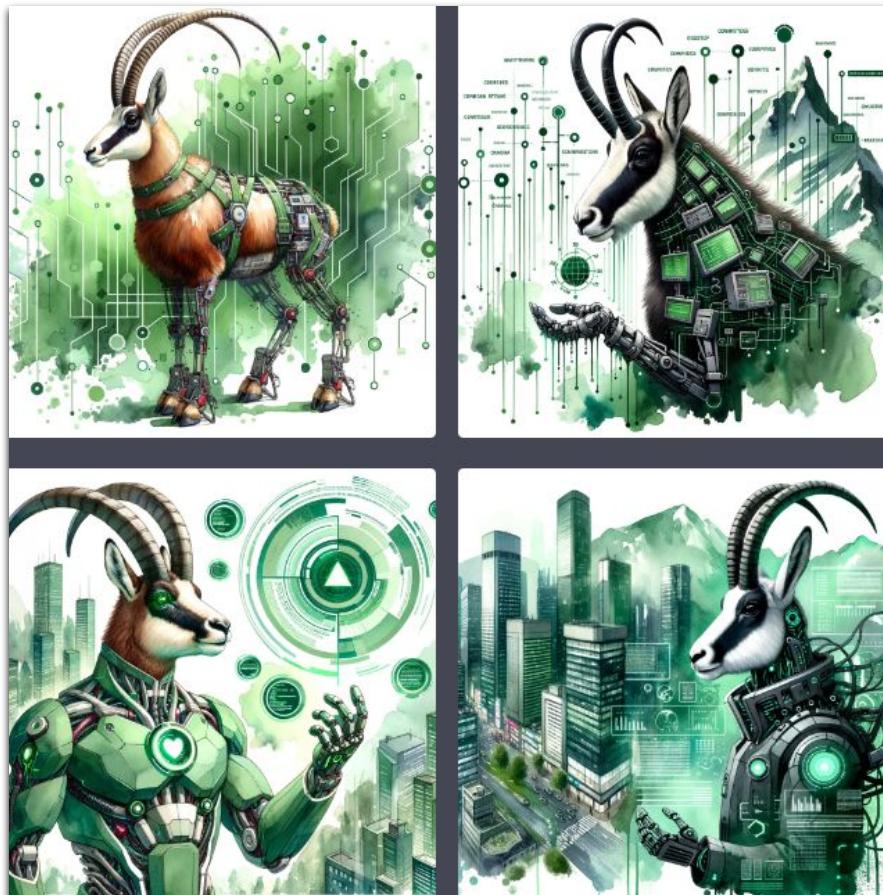
6. Size and Gender Differences:

Male chamois, also known as bucks, are notably larger and have a more robust build than the females or does. This size difference is a common sexual dimorphism in many species.

7. Horns:

Use the detailed description of Alpine Chamois to create 2 variations of:

Watercolor depiction in shades of green and white, with DataGuard Chamois as a robotic figure standing tall amidst a network of digital nodes. The figure harnesses the textual research data, creating a protective barrier around the information.



Use the detailed description of Alpine Chamois to create 2 variations of:

Watercolor portrayal in green and white hues, showing DataGuard Chamois as a robotic humanoid Chamoisa in a futuristic city. The superhero interacts with holographic textual research data streams, analyzing and securing the information.

For all subsequent variations of the humanoid Chamoisa superhero "DataGuard Chamois", keep the green and white color scheme and content. Change only the style. Analyze the different styles step by step and then create the images.

Make 4 full body concept arts:

- * Use a hyper-realistic style for image 1.
- * Use a gritty dark fantasy style for image 2.
- * Use a Dune (Frank Herbert) style for image 3.



“Text to Logo” Chain of Thought Prompting

