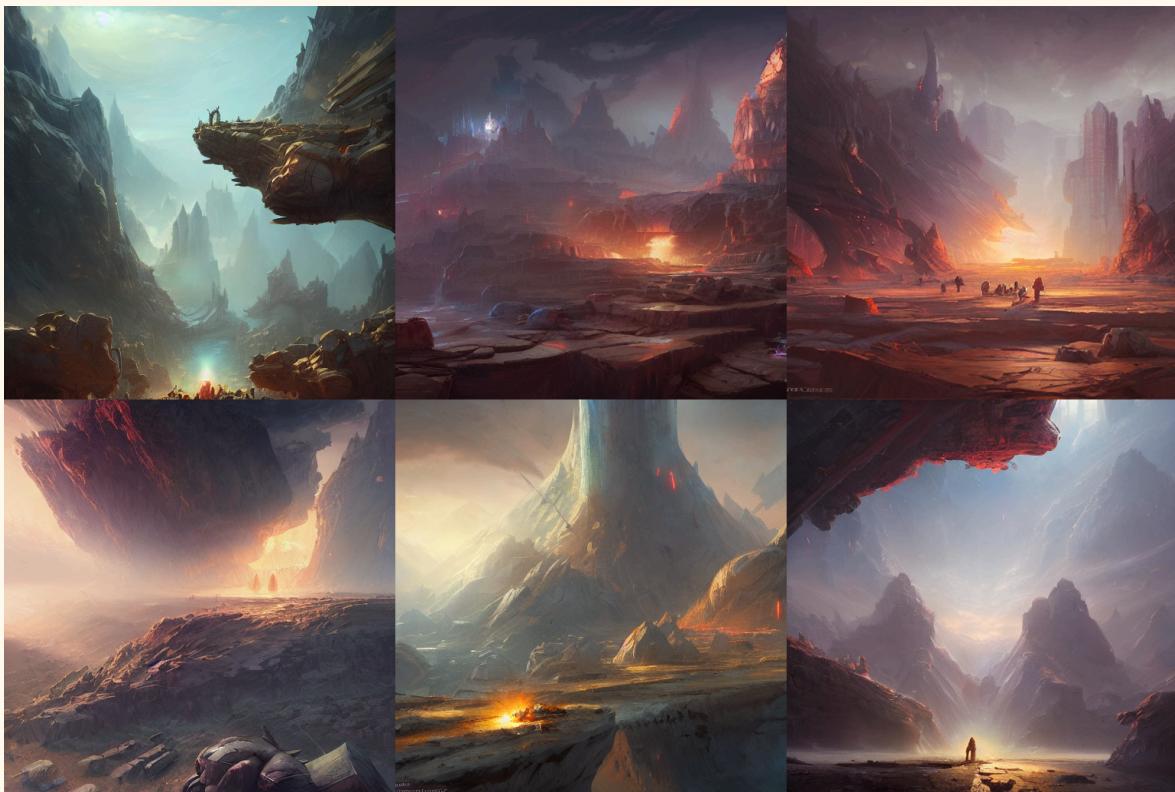


# Stable Diffusion

## Ultimate Beginner's Guide

---

By Arman Chaudhry (discord:Handi#2783, [https://twitter.com/Arman\\_point0](https://twitter.com/Arman_point0))



## INTRODUCTION

Stable Diffusion (SD) is a text to image model similar to others you might be familiar with such as Dalle-2, Midjourney, Craiyon. You can submit a text prompt to the AI and within seconds unique images, like the ones above<sup>1</sup>, will be generated just for you. This guide is for both people who have never used any image AI and veterans alike, and is intended to get you up to speed quickly and some tips and tricks to get the most out of SD.

---

<sup>1</sup> !dream "epic scene, concept art by Noah Bradley, Darek Zabrocki, Tyler edlin, Jordan Grimmer, Neil Blevins, James Paick, Natasha Tan, highly detailed, ultra detailed, ultra realistic, trending on artstation, masterpiece" -n 6 -g -s 75 Prompt by Arman Chaudhry and Morbuto -Seed 1667540758

## How do I use it or get access?

Access to the AI is through discord, which your beta invite should link you to. To generate an image first join any one of the “Dream” channels on discord. If you cannot type in that channel, just hold tight and you will get access soon (normally it takes a day or two). There’s no need to request access or fill out another form, if you are in the discord it will be taken care of and there will be a discord notification and announcement when you get access. Once in the Dream channel you can start prompting.

## Anatomy of a Prompt

A prompt is structured as following:

```
!dream "Your prompt here" -optional_modifiers_here
```

The modifiers are not necessary. If you want to get started it’s that simple, try it out now!

## Modifiers Overview

While a simple prompt can produce good results, getting familiar with the modifiers available will make it easier and more consistent to get what you’re looking for. Here is a quick overview of them and later sections will go into more detail. Note modifiers are case sensitive.

### 1. **-h or --help**

This will return a list of all the modifiers you can use, their default settings and options.

### 2. **-H or --height**

This chooses the height for your image, and takes inputs that are multiples of 64. 512x512 is the default. Changing height or width results in artifacts such as duplication of limbs or people so some trial and error might be needed when generating non-standard aspect ratio images. See the section on custom widths and heights before modifying. Example prompt is: !dream “Your prompt here” -H 768

### 3. -W or --width

Same as height, this takes multiples of 64 and default is 512. Changing it might increase the amount of images with artifacts but you can still generate great images with a non-standard aspect ratio. See the section on custom widths and heights before modifying. Example prompt is: !dream “Your prompt here” -W 768

### 4. -C or --cfg\_scale

Classifier-free guidance (CFG) scale is one of the most asked about and there is a section dedicated to later in this guide as well as a link to a visual comparison of CFG settings in the additional references section.. Basically though, this changes how strongly the AI will follow your prompt and default is set at 7. I recommend not changing it until you've read the section dedicated to CFG scale. Example prompt is: !dream “Your prompt here” -C 12.0

### 5. -n or --number

This sets how many images will be returned for your request. Default is 1 image and you can increase it up to a maximum of 9 images per request. Example prompt is: !dream “Your prompt here” -n 9

### 6. -i or --separate-images

Images used to come in a grid and this command was for specifying you wanted each image to be a separate file. Now this is the default mode so you do not need to specify it, but I included this incase you see -i in the bot's interpretation of your prompt or in other's prompts. Example prompt is: !dream “Your prompt here” -i

### 7. -g or --grid

The opposite of -i, this will return your generation as one image file, with all images put together in a grid. This is useful when you want to easily compare images or rapidly prototype prompts and settings. Example prompt is: !dream “Your prompt here” -g

### 8. -A or --sampler

This sampler is what the AI model uses to actually decide how to generate your image, but is a setting I don't recommend you change unless you are a power user. There is also almost no difference in the output image depending on which sampler you use with 2 notable exceptions (check the linked sampler study in the additional resources section.)

Default is k\_lms and options are k\_lms, ddim, plms, k\_euler, k\_euler\_ancestral, k\_huen, k\_dpm\_2, and k\_dpm\_2\_ancestral. Example prompt is: !dream “Your prompt here” -A k\_euler

## 9. -s or --steps

The AI model starts from random noise and then iteratively denoises the image until you get your final image. This modifier decides how many denoising steps it will go through. Default is 50, which is perfect for most scenarios. For reference, at around 10 steps you have generally a good idea of the composition and whether you will like that image or not, at around 20 it becomes very close to finished. If cfg\_scale and sampler are at default settings, then the difference 20 steps and 150 (the maximum) is often times hard to tell. So if you want to increase the speed at which your images are generated try lowering the steps. Increasing steps also often adds finer detail and fixes artifacts (often but not always). Example prompt is: !dream “Your prompt here” -s 20

## 10. -S or --seed

All images generated start with random noise and the seed determines what that random noise will be. By default the bot will randomly pick seeds for you, but if you happen to like an image you generated you can reuse the seed (with the same prompt and modifiers) and get the exact same result back. The usefulness of this comes when you want to make small variations to the modifiers you used or prompt, then reusing that seed will lead to the generated images being close to the original composition you liked, which help you hone in on your desired result. Example prompt is: !dream “Your prompt here” -S 12345



```
!dream
"environmental
portrait of princess
at a japanese
mountain onsen at
golden hour,
autumn colour oak
trees, Concept art
by Jorge jacinto,
masterpiece, highly
detailed and ultra
realistic, trending
on artstation" -W
704 -n 1 -s 125
Prompt by
Morbuto#8802 on
Discord, seed
2891151871
```

## 11. -t

-t will let you see the tokenized version of your prompt. What are tokens? They are how the AI model sees and interprets your inputs. Currently there is a 77 token limit, including the “startoftext” and “endoftext” token, so effectively you have 75 tokens of room for your prompt. This modifier is useful when you have long prompts and aren’t sure if it will all get seen and used by the AI. If your prompt is too long, you might get an error or warning stating your prompt was truncated- that means the parts after the 77 token limit were thrown away. Also don’t worry if the tokens seem to break your words into multiple parts or numbers get separated with spaces, that’s just how it’s always worked. An example prompt is !dream “Your prompt here” -t

## 12. -a and --ac

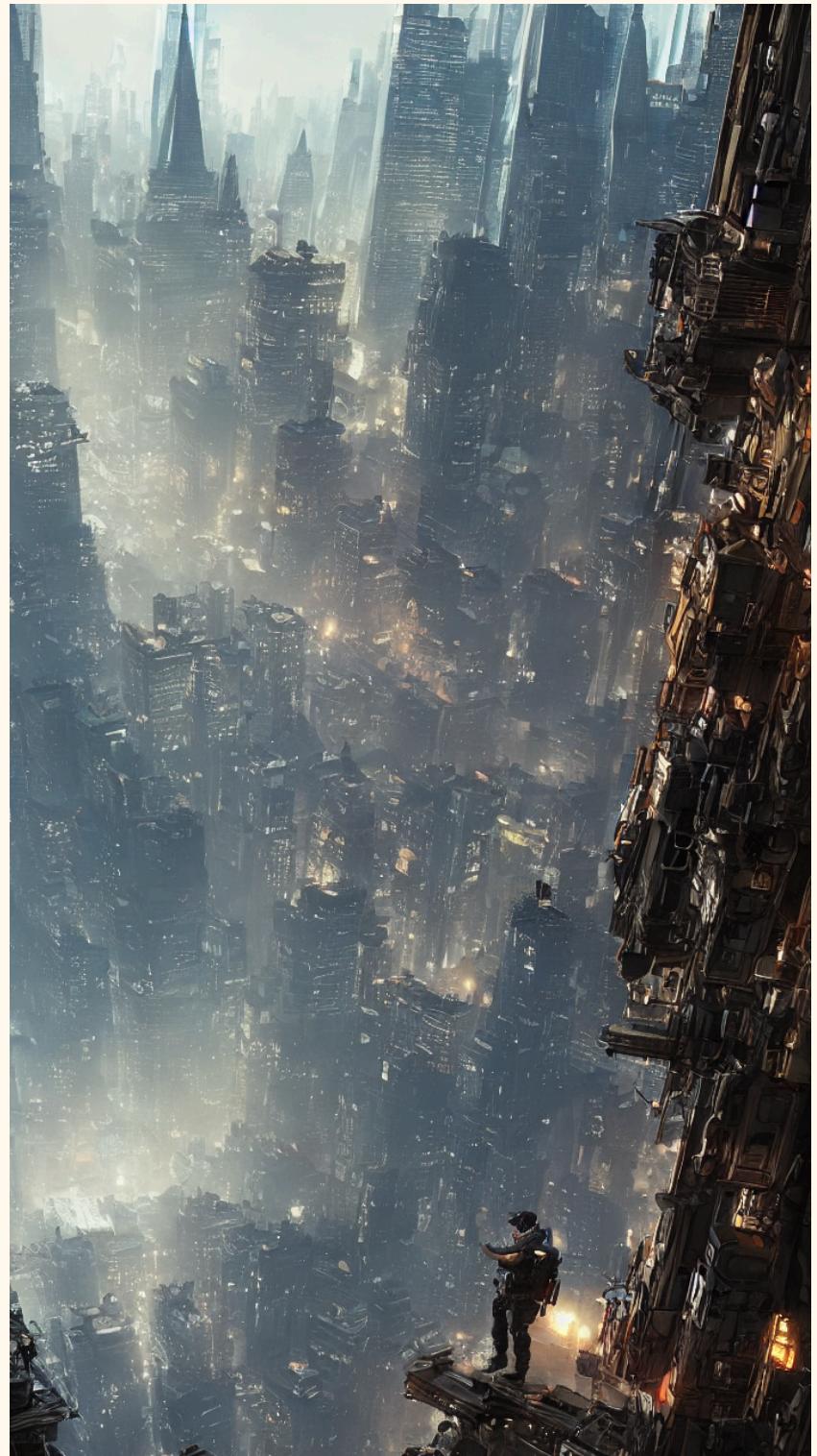
-a is the command to have your generated image be in the ASCII art format. If you enjoy leet hacker aesthetics you can try this out and get something to put in your terminal. You can also add the -ac modifier to control the number of columns for the ascii art (minimum 40, maximum 160). The result will be sent as a text file. Example prompts are !dream “Your prompt here” -a and !dream “Your prompt here” -a -ac 128

## Custom Width and Height Tips

The Stable Diffusion model currently natively works with 512x512 pixels, meaning that all custom width and height ratios are somewhat of a work-around. This means you will get a lot more artifacts and weird generations, but also some stunning images that can be made with non-standard aspect ratios like the image on the right<sup>2</sup>. I suggest generating many variations at a low step amount until you find a composition you like. Then use the seed for that composition to run at higher steps or with slight variations to further fine tune the image.

### Note on duplicates

Often you run into duplicated faces and bodies in these aspect ratios due to the context window of the AI. If your prompt strongly focuses on something (say the eyes of a person) then the SD might try to ensure each 512x512 patch contains eyes. This could lead to a request of a portrait of one person to contain two or more.




---

<sup>2</sup> !dream "A lone police officer overlooking a ledge towards the city below in Ghostpunk New York City | highly detailed | very intricate | cinematic lighting | by Asher Brown Durand and Eddie Mendoza | featured on ArtStation" -H 1280 -W 704 Prompt by Thicc Birb <https://twitter.com/WeirdStableAI> Seed 1921211903

Rewording/refocusing prompts can often fix duplication issues.

## Size and Time tradeoff

The larger sizes often associated with custom dimensions could make your generations take much longer. Important things to note are that there is a maximum of 1 megapixel resolution allowed (that works out to 1024x1024 for square images, but any single dimension can go over that). If you are getting errors with your prompt, you are likely being timed out for having either too large of a dimension, or too many steps for an image of that size. Lowering your image size or the amount of steps per image might fix it.

Even if the size is allowed, your custom size might still take a long time to generate, slowing down the generate, tweak, generate workflow, so consider lowering the steps to 10 or 20 to get fast generations that still give you a good idea of the composition. Then reuse the seed and prompt you like at 50 or more steps to get the final image.

## Pitfalls to avoid

Due to the context window limitations of the model, some prompts that are fine for a standard 512x512 image result in suboptimal results. Particularly bad are prompts that ask for a certain number of something, since the context window might at different points in the image ensure that locally there are the right number of things, but globally that results in duplication. Also another thing to note is specifying specific human anatomy, like how the eyes or nose look. The same context window issue could cause multiple pairs of eyes or noses to come about depending on how heavily you highlighted them in your prompt. That's not to say this will happen every time or with all images, you can still get good results ignoring these tips but following them might produce more consistent images.

## Classifier Free Guidance Scale Tips

The most common question is what is CFG and what should I do with it. The simple answer is don't touch it. The default of 7 is perfect for more use cases. If you are still new to Stable Diffusion you can easily generate 1,000 images (and believe me, some people have) without touching the setting and get good results. The effect of CFG is technically to make the bot follow your prompt more closely, but it has many unintended effects. For instance, increasing the CFG will increase the saturation of your image. Another effect is the creation of artifacts like extremely discolored areas. A solution to the artifacts is to increase the number of steps. A CFG of 12 with the normal default 50 steps might cause artifacts, but I find the artifacts might clear up with an additional 25-75 steps. I recommend only going high on the CFG scale if you have a tricky prompt that might be helped by forcing the AI to stick more closely to your prompt. Although smaller increases or decreases of CFG could be used strategically to change your image aesthetic. But first a note on the acceptable range. For best results at 50 steps, stick to CFG between 7 and 10!

### **Limiting Range**

You can technically set the CFG scale to 9999 (or higher probably) but for actual good useful images, I recommend limiting the CFG range to between 0 and 20. You can also go negative to get the "opposite" of your prompt, but that is normally not widely used. CFG of higher than 20 do not appear to have any noticeable benefit, only further degradation and introduction of artifacts. So I would encourage you to be conservative in your usage of CFG scale.

### **Artistic vs Photorealistic**

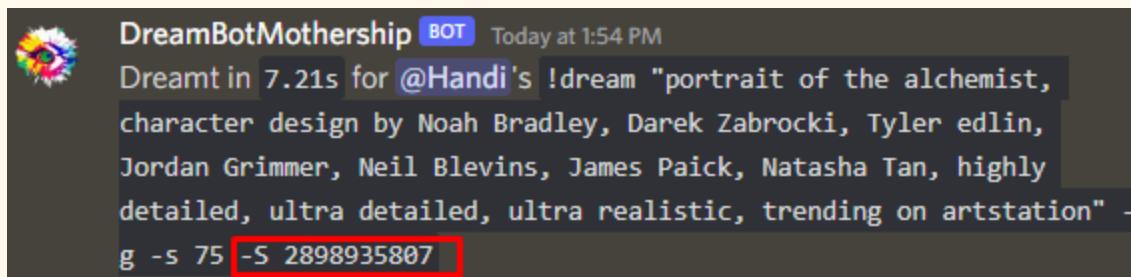
Now on to why someone might want to change from the default in the first place! Something people have noticed is that higher CFG scale will make stylized images become more stylized and painterly, so if you have some concept art or something you're creating but it lacks a certain something, maybe try turning up CFG scale to around 12 or as high as 15. Meanwhile lowering from the default to 5 or 3 ends up producing more realistic images. Now this influence is minuscule in comparison to your prompt, so don't think you need to change it for each prompt you make. However when you are refining a prompt to get it just right, tweaking these values could add that final touch.

## Seeds and Variations

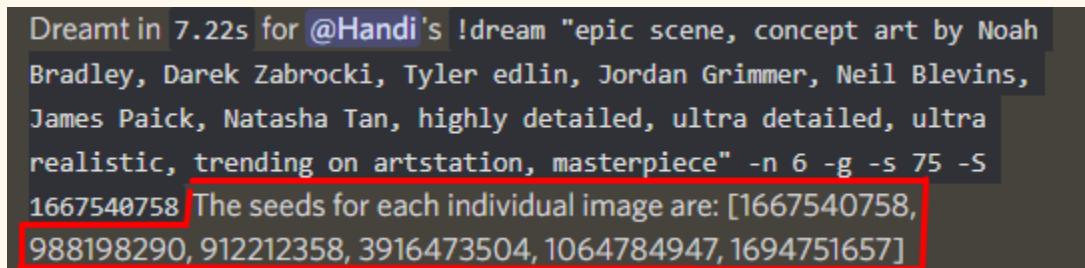
While variations are not officially supported in the current version of Stable Diffusion, you can still make something close to variations. SD will always return to you the random seed that was used to generate an image. If you generate multiple images, it will return a list of all of the seeds used. This is the key to creating these pseudo-variations. If you reuse a prompt with the same seed (as well as all other settings such as steps, cfg scale, etc), you will get exactly the same image. Where it gets interesting is that small tweaks to the settings or the prompt will now result in extremely similar images, just slightly modified. You could make small changes like hair color, age, ethnicity, etc. to the prompt and get generally the same composition, with just those specified things changed.

### Finding your seed

The seed is returned to you with each bot message. For single images it is normally the last thing on the prompt as shown below:



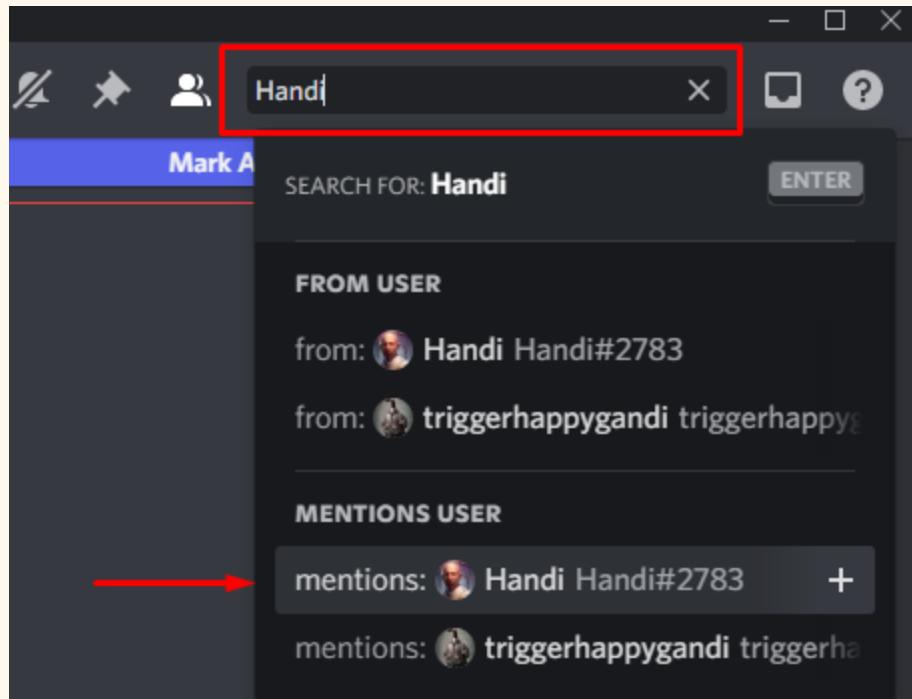
For prompts with multiple images the seeds are given to you in an array as follows:



## Finding your Images

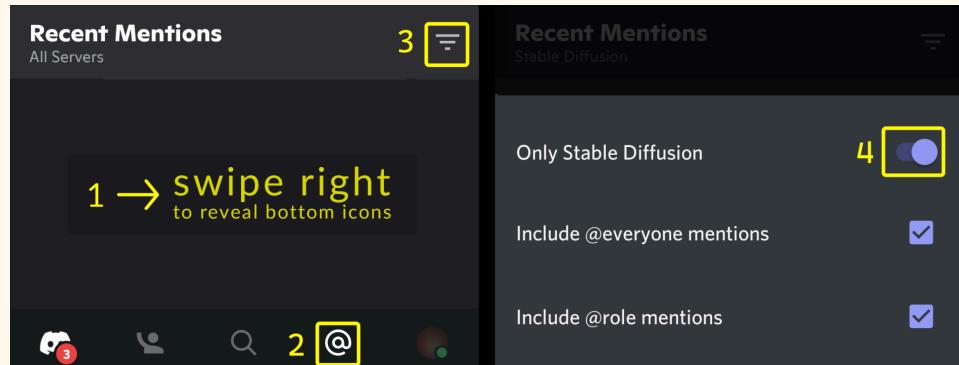
### Desktop

At the top right hand side of discord there is a search bar (marked in red on the screenshot to your right). In there simply type in your discord name and click the mentions option that corresponds to your name (like the red arrow is pointing to. This will open up all messages that mention you in the server which include, which includes every single bot message that has your images. You could also access mentions by clicking the inbox button to the right of the highlighted search box.



### Mobile

Simple swipe right on your Discord app and you should see icons appear at the bottom of your app. Click the "@" icon (highlighted in the yellow square marked 2) to see your mentions (this will include all mentions from the bottom with your images, but also messages from other discord servers. You can click the hamburger menu button on the top right (marked as yellow square 3 in the screenshot) to bring up settings and select "Only Stable Diffusion", like the yellow square marked 4 shows, to limit it to mentions from the discord.



## Additional References

While this guide is meant as a quick start resource and reference, there are *a lot* of studies and references done by some amazing people. Once you feel ready to be a more advanced user, follow these links and take your game to the next level.

[Proxima Centauri B's Stable Diffusion Artist Studies](#)

[Stable Diffusion CFG Scale Studies](#)

[Stable Diffusion Sampler Studies](#)

[Stable Diffusion Launch Day Presentation And Tutorial](#)

[Stable Diffusion - Prompt Weighting](#)

[promptoMANIA:: Prompt Building Tool](#)

[Stable Diffusion Akashic Records](#)

[Clip Search Tool by Rom1504](#)

[Clip Interrogator by @pharmapsychotic](#)