EIGHTEEN HAIKUS

Background

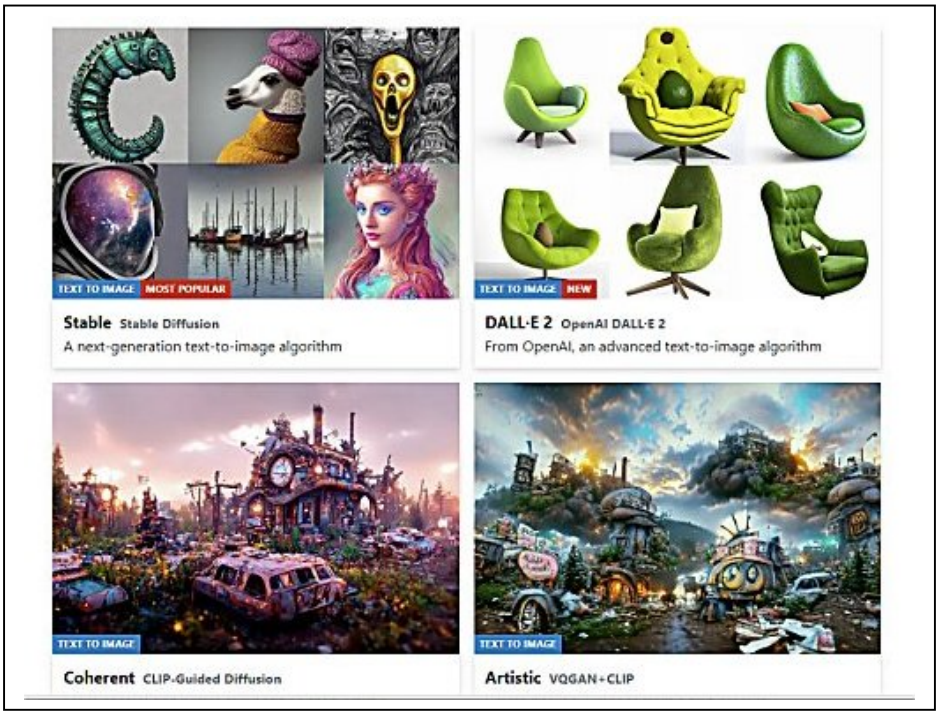
This started as an experiment into combining the modern techniques involved with AI generated art with the ancient structure of haiku.

I attempted to recapture the spirit of the traditional Japanese form within a more modern context. The AI art being based upon the words of the haiku itself. For this, the online website *NightCafe* was utilised.

NightCafe

In order to use *NightCafe* you will need to subscribe, which is free. Your email address is used to send you 5 free credits each day.

When you start creating, you will be presented with a range of **methods** from which to select:



The **Artistic** algorithm is the original text-to-image AI art algorithm. It is good at producing textures and scenery based on descriptive keywords), however the images often don't obey the laws of physics. For example: a building floating in the sky, or a repeating pattern throughout the image.

The **Coherent** algorithm is a newer addition and specialises in creating images that obey the laws of physics. However, it can be a little more hit-and-miss than the artistic algorithm.

The **Stable** algorithm is one of the latest additions and is even more coherent.

The **DALL·E 2** algorithm is an evolution of the original DALL-E by OpenAI that started the text-to-image scene.

Basically, play around with each. I have found some text is better suited to one method rather than another.

Once a method is selected you are presented with a range of **controls**:

Text Prompt: Where you enter words. Some words as restricted: child or naked.

Modifier: Where you can add preset modifier styles such as Oil Painting.

Start Image: Here you can add your own image to add to the process.

Seed: This affects the randomness of the image. You can accept that given or change. This is where you need to use trial and error.

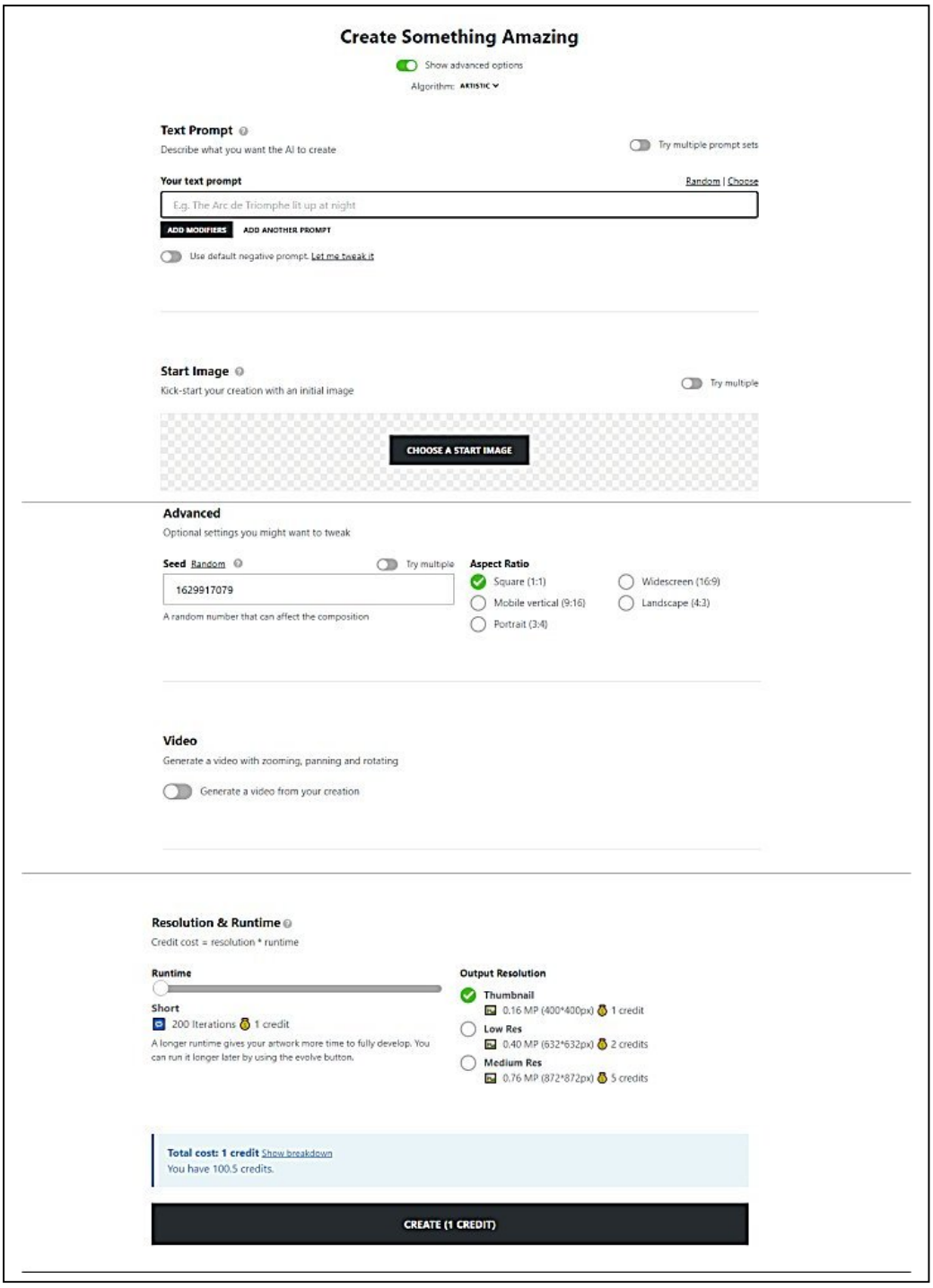
Aspects: Square, Widescreen etc.

Video: You can generate a video of your creation if you wish.

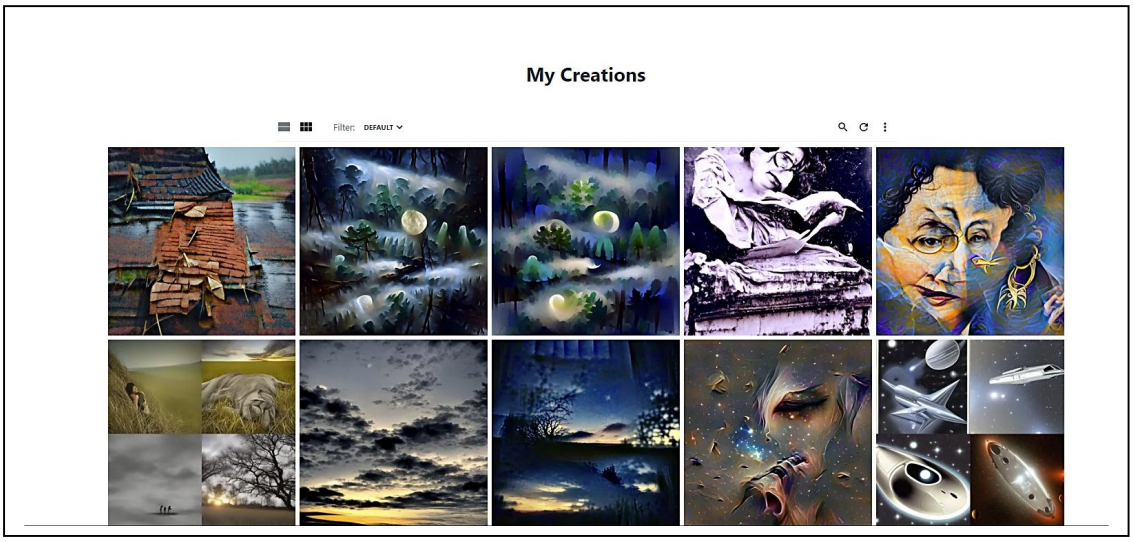
Runtime: Increase runtime for better quality, but this takes more credits.

Output Resolution: Thumbnail is for web use. Higher need for print. Can experiment with thumbnail and when happy upgrade the image to a higher resolution for a few additional credits.

Create: When happy select create.



You can download you creation either with or without your name at the bottom. It also gets saved on the **My Collection** page that only you can view.



With regard to **cost**: you are given free credits when you subscribe and an additional 5 credits are then sent to your email address each day. If you always collect these you can quickly build up enough credits for when you want to try the system out. I try to always have around 80 to 100 credits in the bank. Additional credits can be bought and you can pay to have the adverts removed from the web page, but so far I have never had the need to.

The Images

Old ponds and wells often occur in Traditional Japanese Haiku. For example this one from Basho (1644 – 1694):

Old pond.

A frog jumps in.

The sound of water.

So, this led me to my attempt:

The old well is deep.

But not as deep as sadness

that has no ending.

Putting these words into *NightCafe* resulted in the following image. Bear in mind you have to play around with the parameters and select what you consider the most appropriate result. It may also have to be refined in an image editor, such as *Photoshop*.

From a **copywrite** perspective, this remains with you, provided you do not include any images belonging to other people.

For example, I added the following image I had taken of a Rolls Royce engine from a Wellington bomber to my haiku:

Squat. Gunmetal grey.

Shrapnel from forgotten skies.

Tethered to the ground.

And produced the following end result:

Copywrite, in this case, remains with me as I took the photograph as well as producing the AI image using my haiku.

AI images are not merely dependent upon haiku or even poetry. Prose can be used as well, although there is a limit to the number of words you can employ. In fact you could create an image using a random selection of words selected to create

atmosphere. As an example I took the words: *Night*, *Forest*, *Mist*, *Moonlight* and this produced the following image:

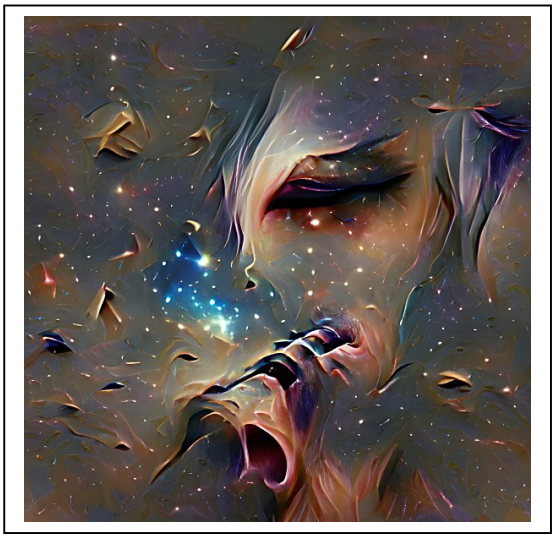
This is a new technique and needs to be explored and experimented with, but it offers exciting opportunities for writers to combine visual effects along with their words.

The following is an image produced from my short poem Silence:

Though stars may die and galaxies

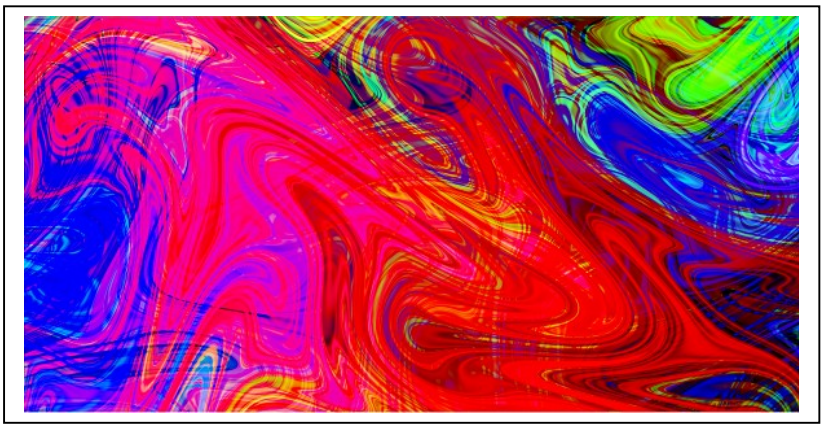
crumble to dust. Silence remains.

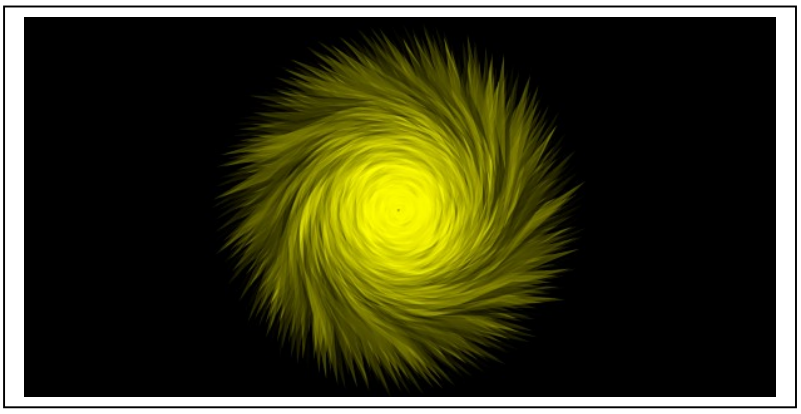
The long, unfilled pause; vast, timeless.

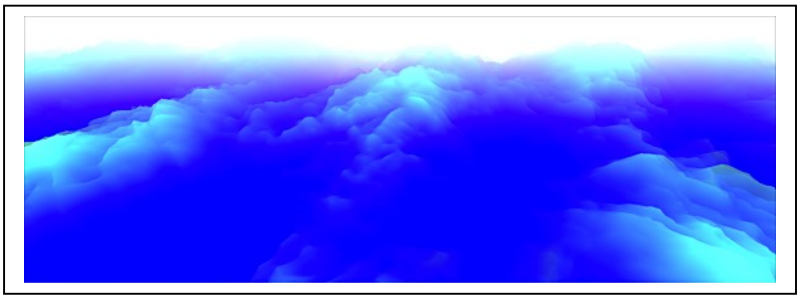


Like everything else it needs time spent experimenting in order to produce the type of images that you require to enhance your words. It won't just produce the desired result by magic. It is a technique to aid creativity and not a replacement for it.

Finally. I would like to mention another method that writers can consider utilising: *Cascading Style Sheets*. These are used mainly for website backgrounds and special effects. But, they can be tweaked to produce graphics to support poetry or prose or to produce striking book cover designs. Here are a few examples:







If all else fails, you could always try a free fractal generator to produce an image such as the following:

These *Cascading Style Sheets* and fractal images could then be added to your words in *NightCafe* to kick start the final image produced, just a thought. Have fun!

NightCafe's Copywrite Position

Where NightCafe is deemed the Copyright holder, as the AI-generation platform, we agree to transfer the Copyright ownership to you - the creator - to use however you like, provided that you didn't use any Copyright restricted images to create it.

Artists Class Action

Three Artists have filed a class action lawsuit targeted at Stability AI, Midjourney, and DeviantArt, who have released artificial intelligence powered image generators that can transform simple text prompts into convincingly rendered images. The artists claim that the companies have violated copyright laws by using their images, along with those of tens of thousands of other artists, to train their image generators and produce derivative works.

However, there are several issues with the lawsuit as constructed:

Firstly, only specific images, not styles, are protected by copyright. Meanwhile, collage is a protected medium under “Fair Use”.

Secondly, Whether or not the lawsuit accurately characterizes diffusion is also in question. According to Dr. Andres Guadamuz, a reader in intellectual property law at the University of Sussex: It is a common misconception that a machine-learning model is just a storage of images that then generates a collage. However, the technology is too complex to be generalized in this way. Stable Diffusion does not store copied works even in early training. Rather, it stores data on how space and colour relate to one another when representing certain objects, using what it learned from studying the training set. This complex process when explained by experts during trial might undermine the argument in the complaint. And then it will be up to a judge, or a jury, to decide if those complexities are meaningfully distant from our understanding of plagiarism or if these models truly constitute fair use of copyrighted materials.

The Technology

According to Dr. Andres Guadamuz: I like to classify what happens in AI generative tools in two stages, the input phase and the output phase. The input phase is comprised of the gathering of data to create a dataset, and this is used to train a model. In the case of Stable Diffusion, it uses a dataset called LAION, which has of over 5 billion entries consisting of the pairing of a hyperlink to a web image with its ALT

text description. This dataset then is used to train a model, I will not go into detail on models, suffice it to say that a model is a mathematical representation of a real-world process that is trained using a dataset, this can be used to make predictions or decisions without being explicitly programmed to perform the task. There are various types of models, but Stable Diffusion and Midjourney both use diffusion models. Diffusion models take an image, add noise to it, and then put it back together.

The current lawsuit uses the word collage repeatedly, but this is where another model comes in. This is known as CLIP. It is designed to improve the performance of AI models on a wide range of tasks involving both language and images. The model is trained using a large dataset of images and their corresponding text descriptions, and it learns to understand the relationship between language and images. This allows it to perform tasks such as image captioning and image classification with high accuracy. So, AI tools use a combination of a diffusion model trained on reconstructing images, as well as CLIP to understand words used to describe an image.

There is another very important element involved in generative models, and this is called latent space. In order to train a model with millions, and sometimes billions of single data points, it would be inefficient to treat every data point in the same way, there could be clusters of similar works. If we are thinking about images, you may not have to look at every single cat picture, it may suffice to cluster data that is similar. Imagine data as a room, you would put the cat pictures in the same space, the dog pictures in another space, etc. Latent space is the space of hidden or underlying factors that can explain the observed data, by clustering similar data, it is used in generative models where the goal is to learn a representation of the data that can be

used to generate new samples that are like the ones in the training set. This is very valuable because it helps to compress the inputs, there’s no need to copy all images of cats, as the model contains latent representations of cats.

The output phase is the generation of the image using all of the above models, and it is done using apps that can take a text prompt and generate a new image based on a combination of statistical data, language models, and latent space. In other words, this is not a collage.

Put bluntly, an AI image isn't collage, rather it is based upon the combination of three models:

1- Diffusion.

2- CLIP.

3- Latent Space.

Writers Perspective

Will AI affect writers? Definitely. There are already poetry generators online where you input key words and select a style: haiku, sonnet etc. I tried one as an experiment and it informed me that the words were AI generated but the structure owed much to Robert Frost.

To be honest, the results were pretty dire. But, so is much of the human generated poetry you come across online. Will it improve, undoubtedly? Watch this space. Here is a haiku I generated using a few keywords:

Forenoon leaping
How a fallen, bright tree blows
enjoying the grass

It doesn't quite meet the 5-7-5 criteria, but it sounds quite Zen!