

“A renaissance painting of an old muslim scholar writing a book” — DALLE

This member-only story is on us. [Upgrade](#) to access all of Medium.

★ Member-only story

Writing Stories with Illustrations with GPT-3 and DALL-E — A Short

Exploration



Hafidz Zulkifli · Follow

5 min read · Jun 20, 2022



108



1



Introduction

The introduction of GPT-3's generative writing capability to the mass market meant that authors can now leverage AI to assist them in the ideating storylines or improving the writing or sentence structure.

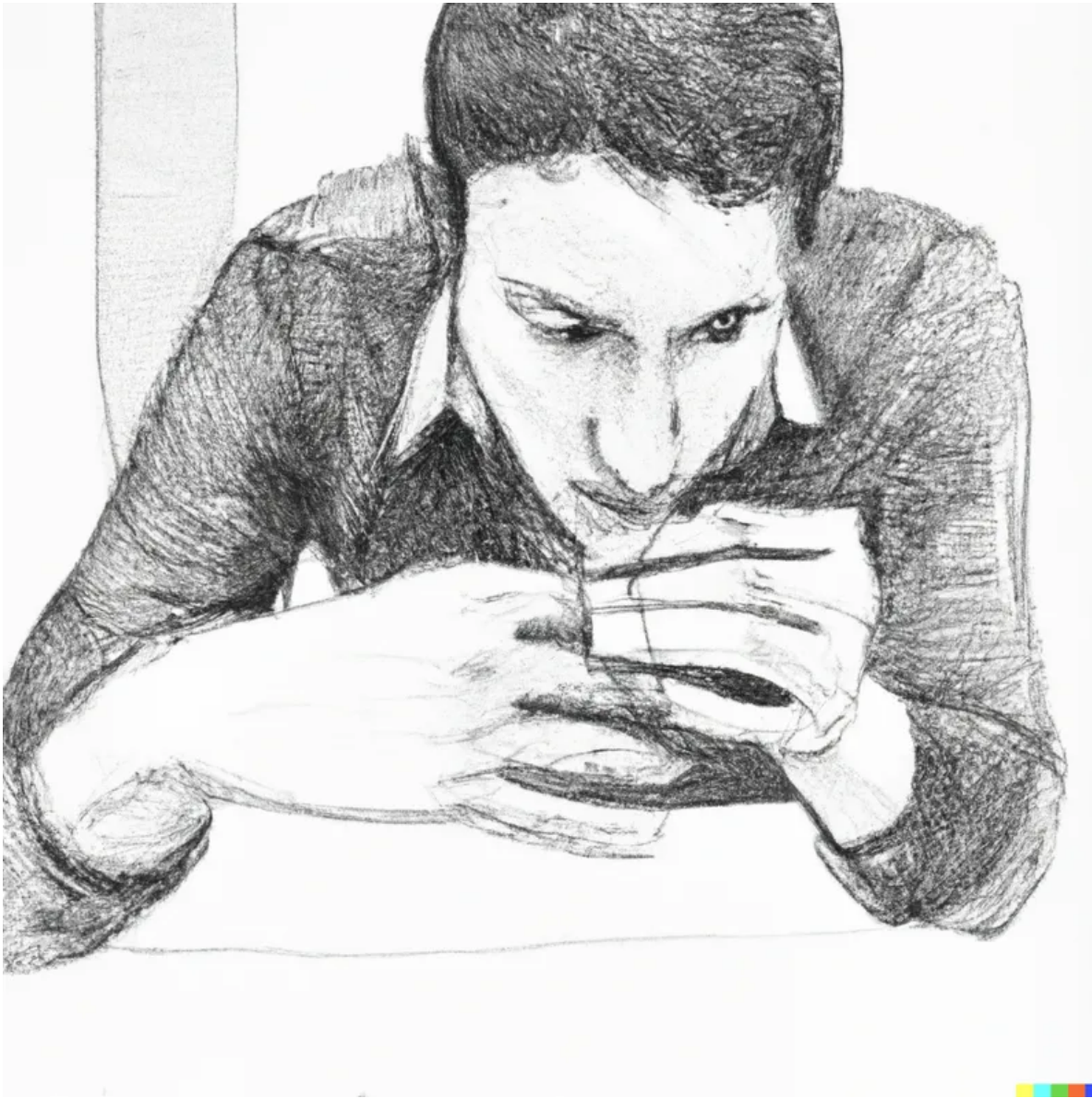
Hell, one guy even wrote an entire poetry book using GPT-3 and had it published within 24 hour.

Coupled with DALL-E, a story teller can now not only produce their own writing, but also enrich it with illustrations.

The intent of this article is not so much an exploration of how the AI behind GPT-3 or DALL-E works, but more about an imagining of what's it's like for a beginner to dive into the world of story writing and illustration with the assistance of these platforms.

The Short Story

Title: The Rapid Decline



He looks at the screen in disbelief as the market valuation of his startup literally falls off the cliff.

“Oh shit.”

After three months of hard work, he thought that they would be getting funded by the best VC in the world. But just when they were about to close the deal, the stock market goes into a tailspin. This never happened before. He knew that the VC would pull out of the deal. He was right. He gets a call the next morning. That’s it. The dream is over. All that hard work, gone.

He picks up his phone and speed dials, Alan — his CFO.



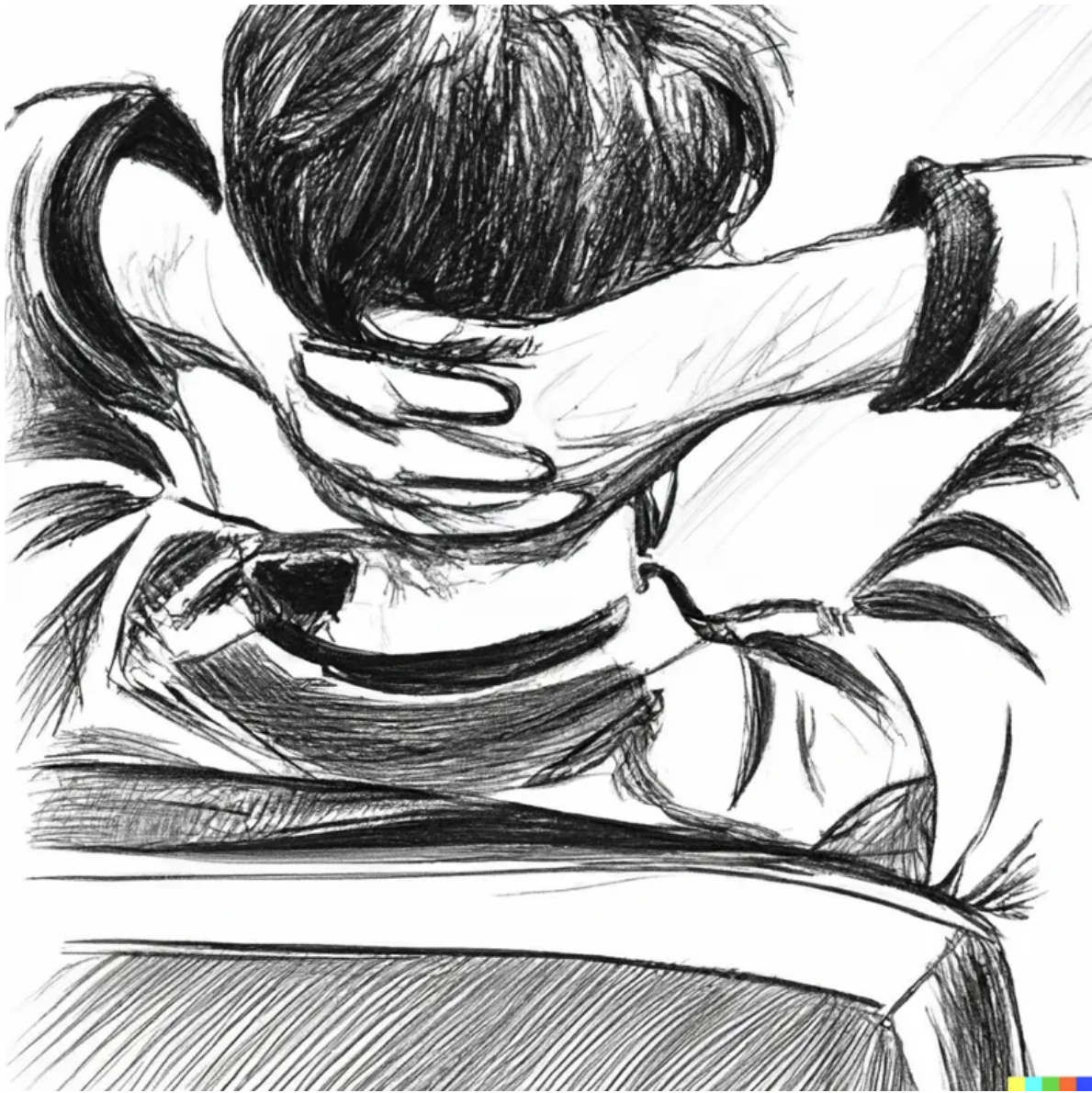
“Alan, have you seen the news?”

“Yes, I can’t believe it.”

“It’s all over isn’t it?”

“It seems so. We’ll have to call an urgent board meeting, I’m afraid.”

“Ok, I’ll arrange it.”

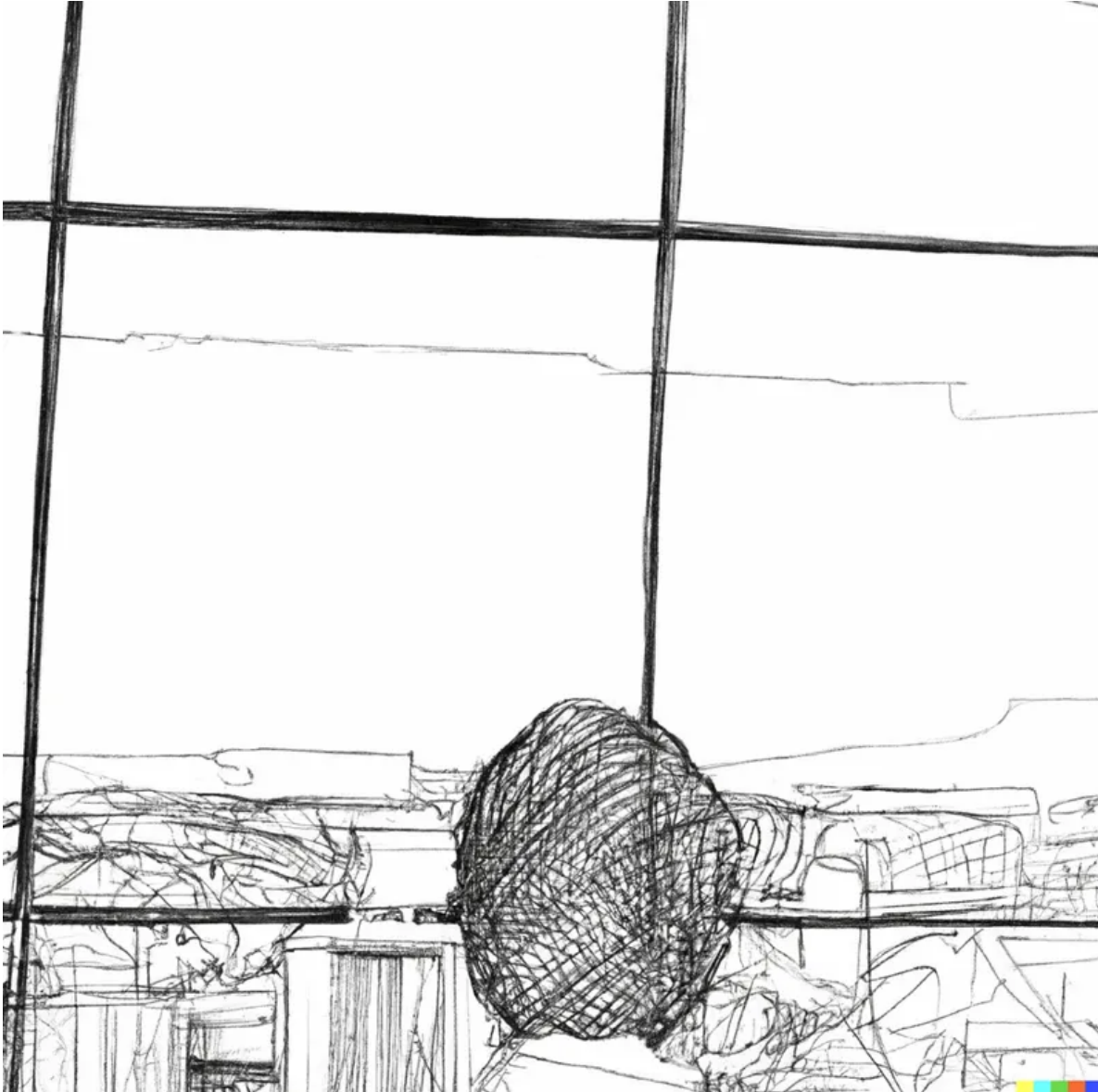


He hangs up the phone and lets out a deep sigh. He sits back in his chair and rocks back and forth, hands holding his head, fingers entwined. He wonders what he's going to say to his parents. They were so proud of him, they had supported him all through his life, through his studies, through his career, and into his business.

He had let them down.

Alan puts the phone down and walks over to his window. He looks out at the beautiful view of the city, high above the hustle and bustle of the streets, and

thinks about the hundreds of people who will soon be out of a job.



He thinks about the board meeting, and how much he dislikes the CEO. He wonders how he will break the news to his wife, who is at home with their son. Turning away from the window and walking back to his desk, he sits down and puts his head in his hands.

He knows it is over, there is no hope left, the business is finished.

Learnings

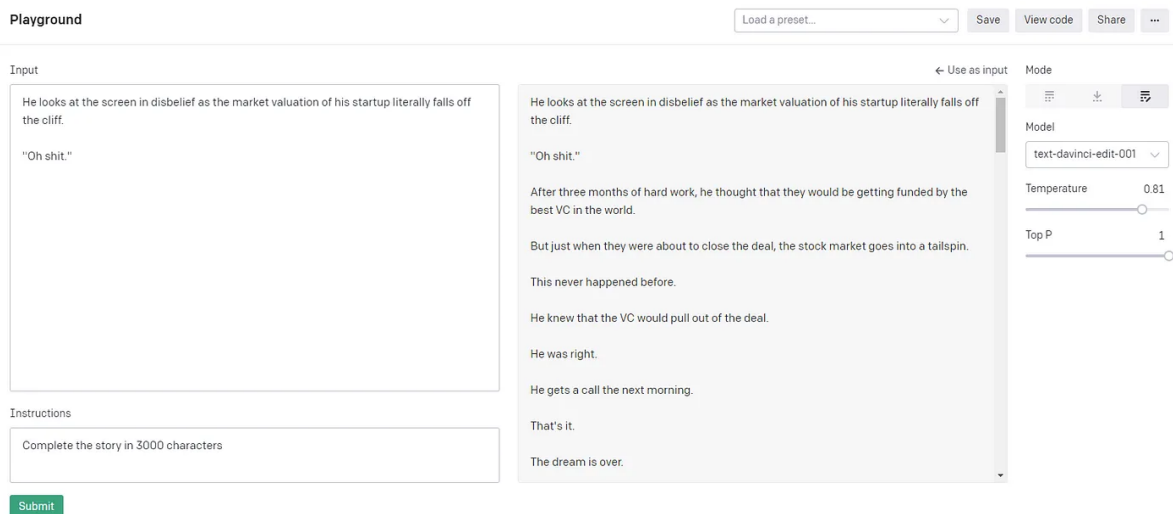
The story began with the following prompt,

He looks at the screen in disbelief as the market valuation of his startup literally falls off the cliff.

“Oh shit.”

From there, it became a series of fine tuning the difference levers that’s available in OpenAI’s GPT-3 Playground. Among them:

1. Changing the instructions.
2. Varying the temperatures to get different levels of creativity.
3. Piecing together different ideas that were generated in multiple prompt submissions.



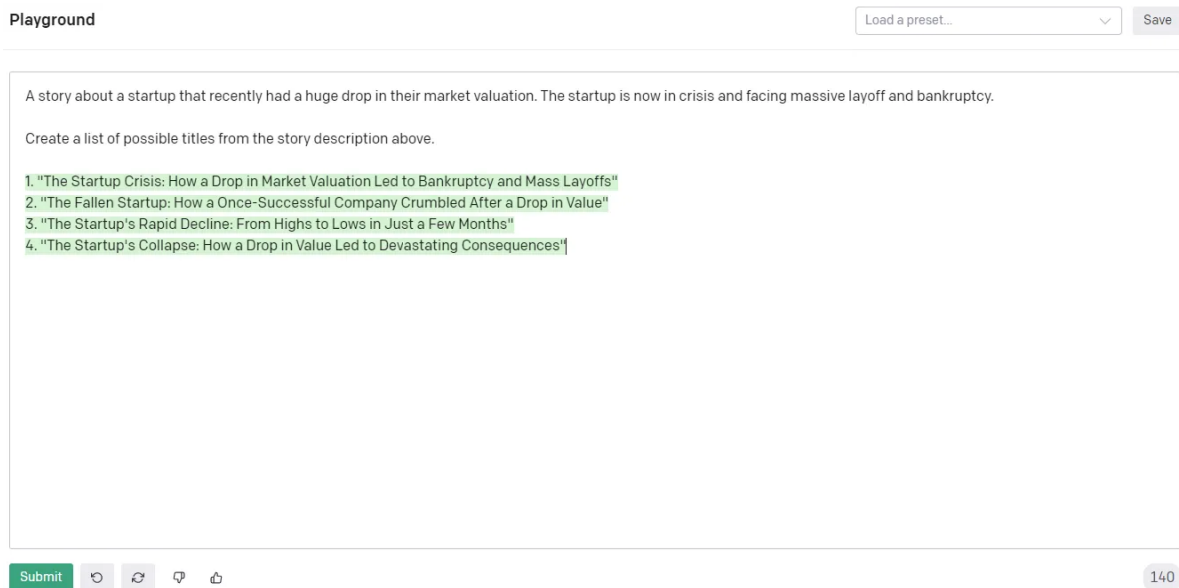
The prompt, instructions and generated text.

But it quickly became clear that a good story arc is something that the author would have to come up with themselves. Things like character building, scene change, plot twists and reveal etc; are probably something that the author will need to have planned at the onset if possible.

Each prompt submission generated quite a number of story plots that seemed interesting — but was quite nonsensical when properly read. Some

was due to the repetitive nature of a language model (in some setting, the generated text went in loops), but in others, it lacked the basic understanding of how certain business concepts actually worked — thus “world knowledge” or “ground truth” on the author’s part is important to establish believability in the story.

Dialogues was a lot of fun to generate, as compared to self monologue and reflection.



The screenshot shows a web interface for a language model playground. At the top left, it says "Playground". On the right, there is a dropdown menu labeled "Load a preset..." and a "Save" button. The main area contains a text box with the following content:

A story about a startup that recently had a huge drop in their market valuation. The startup is now in crisis and facing massive layoff and bankruptcy.

Create a list of possible titles from the story description above.

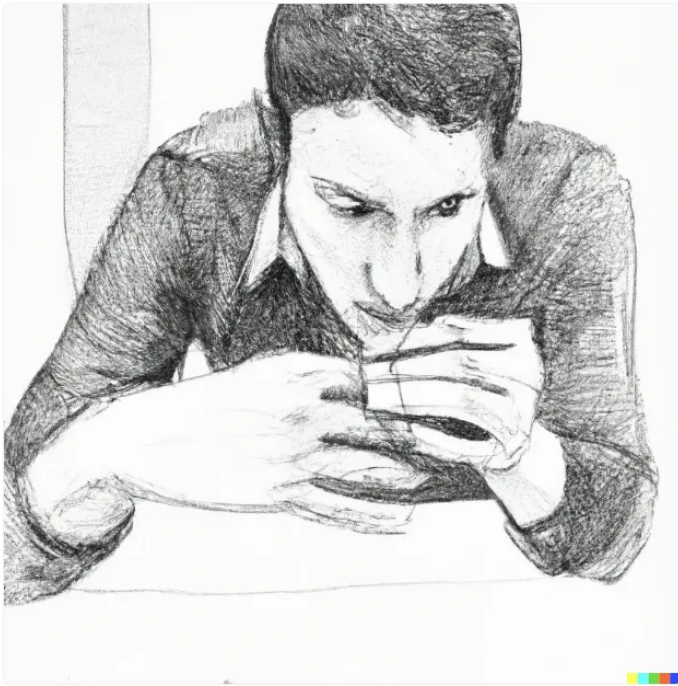
1. "The Startup Crisis: How a Drop in Market Valuation Led to Bankruptcy and Mass Layoffs"
2. "The Fallen Startup: How a Once-Successful Company Crumbled After a Drop in Value"
3. "The Startup's Rapid Decline: From Highs to Lows in Just a Few Months"
4. "The Startup's Collapse: How a Drop in Value Led to Devastating Consequences"


At the bottom left of the text box, there is a "Submit" button and several small icons (undo, redo, search, share). At the bottom right, there is a "140" character count indicator.

Looking for possible titles, using GPT-3

Coming up with illustrations were fun at first, but grew harder over time — due to the need to maintain congruency between the different pictures generated.

For example, here’s the description used to generate the first picture.



...  Share

“man with a worried look in an office, a beautiful pencil sketch”

 hafidz x DALL-E
Human & AI


Picture A: “man with a worried look in an office, a beautiful pencil sketch”

Producing the second picture was somewhat tricky, as the second guy seemed to be missing a lot of hair compared to the first. Also, he was not wearing the right shirt.



...  Share

“man calling someone on a phone with a worried look, a detailed pencil sketch”

 hafidz x DALL-E
Human & AI

Picture B: “man calling someone on a phone with a worried look , a detailed pencil sketch”

To remedy this, DALL-E provides a way to ‘erase’ parts of the generated picture that you don’t agree with and have it revised using a new description. After erasing the entire forehead area of the Picture B, I’ve some additional description that provides more information on what I’m actually expecting.



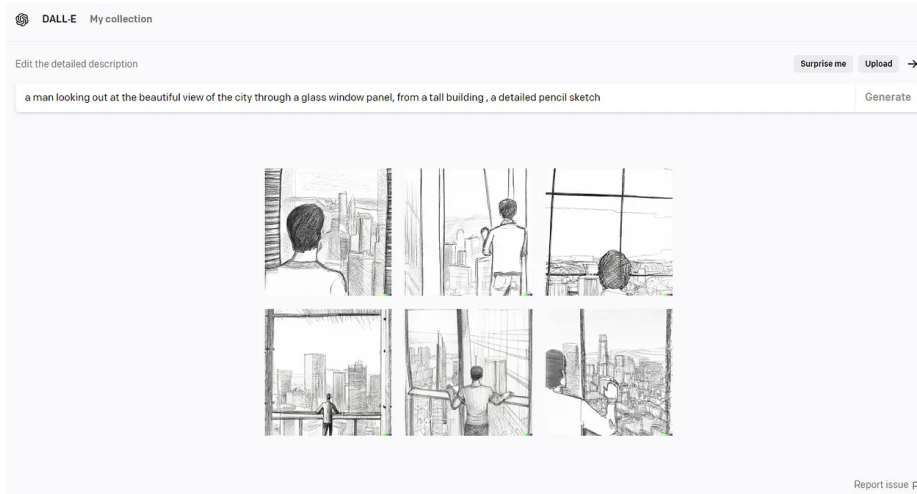
Picture C: “man with dark shirt with white collar, having full set of hair, calling someone on a phone with a worried look, a detailed pencil sketch”

Over time that ‘man’ (in Picture A) had to be described as ‘*man with dark shirt with white collar, having full set of hair*’ (in Picture C).

Such is the level of detail that one needs to apply when describing the scene to be generated.

Summary

As someone who’s never been trained to write stories professionally, and who’s also bad at art — these tools are a godsend.



Experimenting with art choices, in the style of pencil sketches (left) and renaissance paintings (right).

They won't however, replace the need for actual authors or illustrators. If anything, these tools should be able to improve the workflow of both professions; similar to how Github's Autopilot are currently allowing software developers to work faster by automatically generating lines of codes and functions based on prior context that it analyzes.

Exciting times indeed.

- Gpt 3
- Dalle 2
- Copywriting
- Illustration
- AI



Written by Hafidz Zulkifli

