

利用 Midjourney 推測新興量子美學風格初探

Speculating on emerging quantum aesthetic styles with Midjourney : A preliminary study

陳筱媛¹ 劉辰岫^{2*}

¹國立陽明交通大學應用藝術研究所 碩士生

^{2*}國立陽明交通大學應用藝術研究所 助理教授

摘要

本研究利用 Midjourney 作為探索新興美學風格的工具，提出了「文圖文」(Text-to-Imageto-Text)的研究方法，將抽象的量子理論概念轉化為具象的視覺圖像，並初步獲得了「黑暗的」、「燈光」、「未來的」、「橘色」等形容量子風格的關鍵詞彙。此外，我們分析了四種提示語分隔符號(「,」、「::」、「space」及「no space」)在生成風格圖像時的影響，並歸納出最佳使用策略。本研究進一步透過系統化生成圖像，歸納出量子風格的視覺特徵，並推測「量子美學」的新興視覺語彙，為未來有意探索此議題的藝術家或設計師提供參考。研究結果顯示生成式 AI 在探索與定義新興視覺風格上的潛力，不僅擴展了生成式 AI 的應用方式，亦為理解量子理論提供了一種嶄新的視角，激發了大眾對量子理論的興趣與進一步探索的意願。

關鍵詞：量子美學、生成式 AI、美學風格、視覺傳達

Abstract

This study employs Midjourney as a tool to explore emerging aesthetic styles and introduces a "Text-to-Image-to-Text" research method, which translates abstract quantum theory concepts into concrete visual images. Through this process, we identified key descriptors of quantum style, such as "dark," "light," "futuristic," and "orange." Additionally, we analyzed the effects of four prompt delimiters (" , ", " :: ", "space," and "no space") on style generation and synthesized optimal usage strategies. The study further used a systematic approach to generate images and derived visual characteristics of the quantum style, proposing a novel visual language for "quantum aesthetics." This offers a reference for artists and designers interested in this subject. Our findings highlight the potential of generative AI in exploring and defining emerging visual styles, expanding its applications, and providing a new perspective for understanding quantum theory, inspiring greater public interest and engagement with the topic.

Keywords: quantum aesthetics, generative AI, aesthetic styles, visual communication

1. 前言

近年來，隨著量子計算技術的迅速發展，「量子」概念逐漸引起了廣泛的關注。然而，由於量子的本質極為抽象且高度複雜，使得該概念難以被大眾以直觀、易懂的方式理解。儘管有研

究建立量子理論與美學之間的聯繫，但主要著重於文學和心理學層面，尚未具體探討量子風格在視覺上的表現。Broekaert (2018) 通過量子理論解釋觀者在藝術作品體驗中的思考過程，使用疊加 (superposition)、坍塌 (collapse) 和去相干 (decoherence) 等量子物理狀態來形容觀者如何賦予作品多重意義和單一解釋。這些解釋在文學和心理學層面上呈現出富有詩意的表達方式。然而，對於尚未掌握具體量子概念的讀者而言，由於缺乏明確的量子圖像和描述性語言，仍難以想像「量子風格」的樣貌，更難以延伸成為創作的基石。

此外，文字生成圖像 (text-to-image，以下簡稱文生圖或 TTI) 模型的出現，例如：Midjourney，這些模型通過大量圖片進行訓練，並公開提供文字描述，神經網絡則利用數以百萬計的影像和文字標籤進行學習，從而能夠根據用戶的提示生成相似但每次結果皆不相同的圖像。最新的技術依賴於擴散模型 (diffusion models)，該模型不僅能實現逼真的畫質，還能生成從未出現在訓練數據中的圖像，因為模型已經用足夠多的例子來訓練提示 (prompt) 中的每個單獨元素 (Henriikka Vartiainen & Matti Tedre, 2023)。

在這個背景下，本研究將以 Midjourney 作為主要工具來探討「量子美學」，從視覺角度切入，研究生成式 AI 中已被訓練之大數據如何協助判別新興美學風格，挖掘大數據中既有的量子概念與視覺表現之間的潛在關聯。人們以語言與文字作為傳遞視覺想法的媒介，因此，除了視覺圖像之外，生成相應的風格描述文詞亦是探討美學風格的重要環節，能夠更有效地傳遞和詮釋風格的抽象意涵。具體而言，將利用 TTI 文生圖模型搜尋與「量子」相關的圖像，並嘗試透過系統化分析萃取出「量子」風格的視覺特徵，通過統整 Midjourney 對其生成之圖像進行描述的所有詞彙。本研究旨在構建「量子美學風格」的視覺特徵體系，該體系包含「視覺圖像」與「描述文詞」兩類成果，試圖為量子概念的可視化與普及提供一種新的詮釋途徑，同時深化設計與藝術領域對量子美學風格的探討與理解，並提供設計師和藝術家在創作上可參考和討論的依據。

2.TTI 生成式 AI 工具介紹與比較

目前在網路上有數十種知名 TTI (文生圖) 的生成式 AI 工具被開放使用，本研究從中選擇了四個使用門檻較低的工具進行應用，並對 Bing Image Creator、Leonardo.AI、Imagine、Midjourney 以上四種工具的使用模式與生成效果進行比較及分析，如表 1 所示，Bing Image Creator 和 Imagine 相較其他兩者並沒有特別突出的優勢；Leonardo.AI 除了可以免費使用以外，其使用者介面提供了參數調整拉桿等輔助功能，適合初學者生成具豐富變化的圖像；Midjourney (以下簡稱 MJ) 則在官網上詳細說明了 AI 使用的各項細節，並允許在提示詞 (prompt) 中使用官方提供的參數指令，以指定特定的生成風格，例如：在句尾加上“--style raw”指令，即可移除自動美化圖片風格功能，使生成的圖像更準確地反映提示詞的內容 (Midjourney, 2024.)，以圖 1 左 (未加 --style raw) 為例，其中的四張圖片皆由不同元素組成，下面兩張圖片在畫面中出現了人物，左上則出現幾何圖形之排列，而圖 1 右 (加上 --style raw) 中的四張圖片均生成相對單純且一致的畫面：圍繞圓心的光束、漩渦、絢爛色彩等，此結果有助於我們進一步觀察和識別標準的量子風格特徵，最終，本研究選擇使用“Midjourney”作為主要研究工具。

此外，我們進一步實驗並確認了在 MJ 中提示詞的使用方法，在使用文生圖模型時，提示詞的順序對生成結果有一定程度的影響。模型的注意力機制 (attention mechanisms) 通過分配不同的注意力權重來捕捉提示句中每個單詞之間的依賴關係，使得模型能夠更全面地理解與處

理輸入和輸出之間的語義關係 (Vaswani, A. et al., 2017)。在一篇探討 MJ 使用者在輸入提示句時所應用的策略及修正流程的質性研究中，一些參與者 (n=11) 提到單詞順序、重複和誇飾是強調與突出主題的有效方法。透過重新安排提示中的單詞，某位受訪者在採訪中所進行的實驗生成圖最終更加接近他想像的視覺品質 (Goloujeh, A. M. et al., 2024)。這些結果說明了通過設計提示詞順序，能夠更快速準確地引導模型生成符合預期之圖像。

表 1 四種文生圖 AI 工具優劣比較表

優缺點 \ AI 工具	Bing image creator	Leonardo.AI	Imagine	Midjourney
免費	O	O	O	X
調整參數指令	X	O	X	O
指定繪圖風格 (非文字)	X	O	X	O
權重分配	X	O	X	O
高解析度輸出	X	O	X	O
官方使用指南	O	O	O	O
官方社群支持	X	X	X	O
線上執行	O	O	O	O



圖 1

以不同提示詞所生成之圖，左：quantum。右：quantum --style raw。

3.研究方法

為了探討是否存在「量子風格」，以及「量子風格」究竟是什麼？本研究以日常物件作為風格載體，為了分析「量子風格」，經過前述對生成式 AI 工具的初步測試，提出「文圖文」(Text-to-Image-to-Text, 簡稱 TIT) 的探究方法。藉由生成式 AI 的文字提示 (text prompt) 功

能，生成大量含有「量子」關鍵字之「量子風格物件圖」，具象化 AI 模型裡已存在的「量子」美學元素，進而分析各圖像的共同特徵，以歸納出「量子」的視覺特徵。

在文圖文生成流程的第一階段「文字」部分，本研究在 MJ 中輸入含有不同「分隔符號」的文字提示，探討詞彙之間的連接方式對圖像生成的影響。為此，將分析各類分隔符號如何影響生成圖像的融合程度，進一步了解這些符號如何作用於圖像的整體呈現。同時，本研究亦著重於通過強化「量子」元素在畫面中的權重，以漸進式的方法辨別量子視覺風格的特徵。

3-1 「TIT」(Text to Image to Text) 的探究方法

本研究提出「文圖文」(TIT) 的探究方法 (圖 2)。為了分析「量子風格」，該方法透過生成式 AI 工具將含有「量子」的複合詞轉化為圖像，再利用解釋圖像的功能指令生成對應的文字描述，最終根據所有生成描述句中高度重複且共同出現的描述詞，確立「量子」風格的視覺元素，由此建立新興美學風格。



圖 2 文圖文 (TIT) 研究方法之流程圖

本研究選取三種物件類別作為實驗標的，包含「UI」、「cake」、「car」(表 2)。在第一階段「文生圖」(Text-to-Image) 的過程中，我們利用 Midjourney 輸入文字提示，生成「量子風格物件圖」，每個文字提示會產生一組共 4 張的生成圖，如圖 3 示意。在圖 3 的三張圖中，輸入物件的痕跡與模樣皆可被辨認，也看得出「量子」元素在畫面中與物件的結合表現，根據不同物件種類會有不同的結合型態，例如：「量子」與「UI」(使用者介面設計) 的結合圖 (圖 3 左) 中，由晶片、粒子、藍橘色的光等科技感視覺元素所組成；與「cake」的結合圖 (圖 3 中) 中，出現了轉盤、漩渦、繽紛色彩與裝飾等元素；與「car」的結合圖 (圖 3 右) 中，沒有出現太多發光視覺，以汽車造型為主要變化，通常為流線型。

第二階段「圖生文」(Image-to-Text) 過程中，以「quantum, cake」為例，再進一步探索「量子」與「蛋糕」的「量子物件圖」(圖 3 中)。在輸入描述圖片的指令後，根據圖 3 中所生成的文字敘述，如表 3 所示，每張圖會生成四個段落描述該圖片，其中「colorful」、「cube」、「glass」、「planets」、「box」、「bubbles」等詞彙頻繁出現在文字段落中，這些詞彙代表了此

「量子蛋糕」的視覺特徵，該結果有助於更加具體地描述生成圖像。因此，有必要生成一系列以蛋糕為主題的圖像，以擴充數據集，從而獲得更充分的支持來客觀描述這些圖像的視覺特徵。

將生成的描述句依據物件類別進行分類，並計算每個詞彙的重複次數，通過分析高度重複且在所有類別中均出現的詞彙，可以合理推測這些詞彙即為 AI 所認為的「量子風格」組合元素。由於人類以語言作為溝通的基礎，因此本研究利用此方法將風格轉化為文字，歸納出可描述該風格的詞彙。從圖 4 中可以觀察到，撇除介系詞及「style」一詞後，依據重複次數由高至低排序，「UI」使用者介面類的描述句中，「dark」、「design」、「futuristic」、「light」、「UI」、「orange」、「realistic」、「black」、「innovative」、「page」、「scientific」、「unreal」、「rendered」、「art」、「detailed」、「elements」、「interface」、「color」等詞彙重複出現最多次；「cake」蛋糕類的描述句中，頻繁出現「cake」、「light」、「dark」、「fluid」、「orange」、「white」、「plate」、「layered」、「chocolate」、「piece」、「black」、「mike」、「amber」、「detailed」、「cheesecake」、「hasselblad」、「campau」、「photorealistic」等詞彙；「car」車子類的描述句中，則以「car」、「light」、「futuristic」、「dark」、「realistic」、「interior」、「detailed」、「orange」、「renderings」、「illustrations」、「hyper-detailed」、「art」、「red」、「concept」、「photorealistic」、「design」、「realism」、「amber」、「american」等詞彙重複出現次數最高。根據這三類所產生的結果，重新檢視互相重複的詞彙，由於分類物件的描述詞會在此階段有極大機率不重疊，因此高度重複的詞彙將被視為描述「量子風格」的關鍵元素。從上述分析中獲得的前四名共同且高度重複的描述詞為「dark」（黑暗的）、「light」（燈光）、「orange」（橘色的）、「detailed」（細節的）。依照以上方法分析，此四個描述詞為最能用來描述「量子風格」視覺特徵的詞彙。

本研究目前僅以三類物件為主透過文圖文的方式產生描述量子風格的共同詞彙，理論上來說，使用的物件類別數目越多，共同描述詞彙會更趨向精準。

表 2 將物件歸為「UI」、「cake」、「car」三個類別，以及生成圖所使用的提示詞。

類別	編號	prompt
UI	UI-1	quantum, user interface
	UI-2	user interface, quantum
	UI-3	quantum, car UI
	UI-4	quantum, bus UI
	UI-5	quantum, mobile device UI
	UI-6	quantum, watch interface design
	UI-7	quantum, camera interface design
	UI-8	quantum, electrical appliances interface design
	UI-9	quantum, smart home devices interface design
	UI-10	quantum, game interface design
	UI-11	quantum, medical device interface design
cake	cake-1	quantum, cake
	cake-2	cake, quantum
	cake-3	quantum, restaurant
	cake-4	Tiramisu, quantum
	cake-5	quantum, cheese cake
	cake-6	quantum, mille-feuille
car	car-1	quantum, car
	car-2	car, quantum
	car-3	quantum, sedan
	car-4	quantum, roadster
	car-5	quantum, electric car
	car-6	quantum, car interior
	car-7	quantum, car garage
	car-8	quantum, car rear compartment



圖 3 以不同提示詞所生成之圖，左：quantum, UI。中：quantum, cake。右：quantum, car。表 3 「圖生文」所產生的描述句。將圖 3 中的左上圖透過 MJ 所產生的四個段落。

1	A colorful glass cube with swirling patterns and floating pearls, with an octane render of planets inside the box, in the style of Peter Gric.
2	3D rendering of an abstract glass cube with colorful swirls and bubbles, featuring floating space objects like planets or stars. The box is adorned with holographic stickers that give it a futuristic vibe. This artwork creates a sense of wonder by combining the ethereal beauty of crystalline surfaces with cosmic elements, inviting viewers to explore the depths of imagination within its walls.
3	A colorful and translucent cube with pearls and planets inside the box, bubbles on top of it, in the style of digital art, against a white background, with Octane rendering, and a glassy texture. The artwork is created by artist Heim Gao.
4	A multicolored glass cube with an open lid, inside which there is a planet of colorful pearls and rainbow bubbles in the style of digital art, with soft edges and blurred details, low resolution, white background, translucent resin waves, hyper-realistic watercolor illustrations, levitating organic forms, in dark pink and light green, an airbrush drawing, with a small space between it and the sphere on top, a sci-fi fantasy scene with a full-body composition.

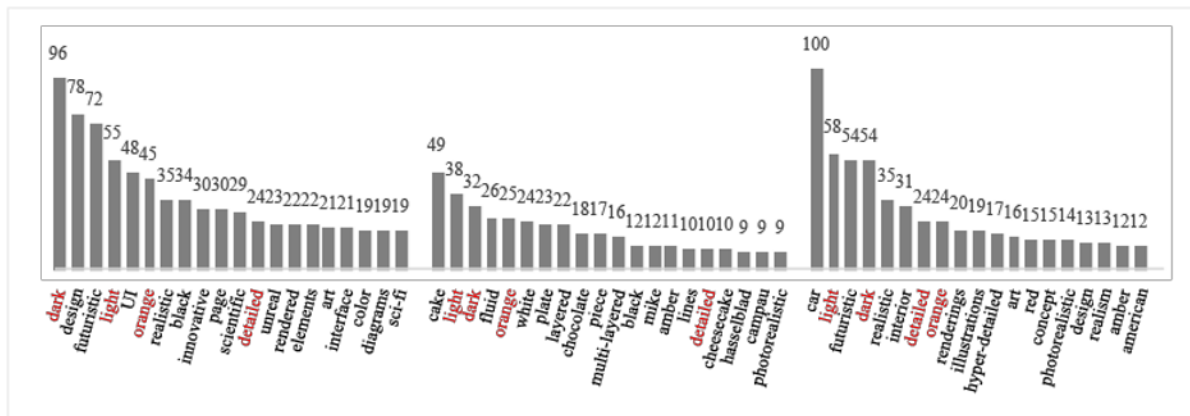


圖 4 以長條圖顯示各類別的高度重複詞，數字表示該詞出現次數。高度重複且共同出現在三個類別之詞以紅字標示。

3-2 分隔符號及改變風格權重之影響

在 Midjourney 中，提示詞順序、詞間分隔符號的選擇（表 4~6）、以及權重提示的使用（見表 7~9）等細微變化，均會對生成圖像產生影響。本研究旨在探討使用「,」、「::」、「space」（空格）及「no space」（無空格）作為詞間分隔符號，對圖像生成融合程度的影響。

於表 4~6 中，分別使用「,」、「::」、「space」及「no space」作為詞間分隔符號。然而，結果顯示各分隔符號之間並未導致生成圖像出現顯著差異，因而難以單獨識別所謂的量子風格。

因此，研究重點轉向使用「::」作為分隔符號，以期通過加強「量子」在畫面中的權重，達成量子視覺風格的辨識。

於表 7~9 中，在「量子」(quantum)一詞後添加「::」與數字，以表示該詞在畫面中生成時的權重大小（其中「::」等同於「::1」）。研究發現，於表 7 中，隨著「quantum::2」及「quantum::3」的提示，生成的圖像逐漸呈現出類似宇宙、微觀光束與粒子的形態，原始物體（如手錶）已不易辨識。表 8 中，當提示句「quantum::folding bike --style raw」逐步提升「量子」權重時，生成圖像逐漸變得抽象，在「quantum::2」時出現了騎單車的人物形象，且畫風從渲染圖拓展至插畫與攝影風格；在「quantum::3」時，圖像更明顯地強調量子視覺元素，如彩色、圓形等特徵。表 9 的結果也呈現類似的趨勢，生成圖像中出現了插畫、幾何設計、宇宙、彩色及光束等元素。由此可見，使用「::」分隔兩詞並增加「量子」一詞的權重，能夠顯著提升量子風格在生成圖像中的辨識度，並使其逐步融入至物體之中。

此外，另一現象也被觀察到：當使用「space」或「no space」作為分隔符號時，AI 可能將整串提示詞視為一個複合名詞（Compound Nouns），從而將最後一個名詞作為主體。例如，當使用「abluewhalea79yearsoldwoman」作為提示詞時，生成的圖像會以「a 79 years old woman」為主體，臉上有著海洋般深藍色的眼睛（見圖 5），而非按照先後順序將前後名詞依比例生成於圖像中。再以圖 6 中的 A、B、C 三張圖為例，提示詞中「apple」與「bed」之間分別以「,」、「space」及「no space」作為分隔符號，結果顯示，A 圖的分隔效果最佳，且「apple」依然是畫面中的主角；而在 B 與 C 圖中，兩者物體的融合度較高，且圖像的主體已轉變為「bed」，呈現出如蘋果造型的床、堆滿蘋果的床或蘋果樹屋內的床等情景。由此結果推論，當使用「space」或「no space」作為分隔符號時，Midjourney 在處理一串「名詞+名詞」結構的提示句時，可能將其視為一個複合名詞，生成具有「前者名詞風格的後者名詞」之圖像。不過，此推論對於不同提示詞內容的可靠度多高，則有待進一步驗證。

表 4 使用不同分隔符號區隔文字提示「quantum」與「watch interface design」之結果

prompt	quantum, watch interface design --style raw	quantum:: watch interface design --style raw	quantum watch interface design --style raw	quantumwatchinterfac edesign --style raw

表 5 使用不同分隔符號區隔文字提示「quantum」與「folding bike」之結果

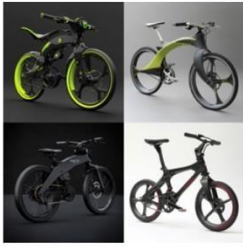
prompt	quantum, folding bike --style raw	quantum:: folding bike --style raw	quantum folding bike --style raw	quantumfoldingbike --style raw
				

表 6 使用不同分隔符號區隔文字提示「quantum」與「lounge chair」之結果

prompt	quantum, lounge chair --style raw	quantum:: lounge chair --style raw	quantum lounge chair --style raw	quantumloungechair --style raw
				

表 7 文字提示「quantum」與「watch interface design」加上不同強度權重指令的漸進生成圖

prompt	quantum, watch interface design --style raw	quantum:: watch interface design --style raw	quantum::2 watch interface design --style raw	quantum::3 watch interface design --style raw
				

表 8 文字提示「quantum」與「folding bike」加上不同強度權重指令的漸進生成圖

prompt	quantum, folding bike --style raw	quantum:: folding bike --style raw	quantum::2 folding bike --style raw	quantum::3 folding bike --style raw
				

表 9 文字提示「quantum」與「lounge chair」加上不同強度權重指令的漸進生成圖

prompt	quantum, lounge chair --style raw	quantum:: lounge chair --style raw	quantum::2 lounge chair --style raw	quantum::3 lounge chair --style raw
				



圖 5 輸入提示句「abluewhalea79yearsoldwoman」之生成圖

		Space	No Space
A	B	C	
			
apple, bed	apple bed	applebed	
D	E	F	
			
beer, belt	beer belt	beerbelt	
G	H	I	
			
nintendo, doughnut	nintendo doughnut	nintendodoughnut	

圖 6 三種分隔符號與三組提示詞之組合矩陣

4.結論與未來展望本研究利用 Midjourney 探索並試圖界定「量子風格」的視覺特徵，並生成兩大類型的成果，即代表性的「視覺圖像」與「描述文詞」。通過調整提示詞的順序、分隔符號的使用以及權重設置，生成了一系列圖像，並對其進行了初步的量子視覺特徵歸納。基於先前在文學和心理學層面上將量子理論與美學結合的研究，本研究進一步具象化量子概念，提供具有風格代表性的圖像和描述詞彙，從而擴大了「量子風格」的應用範圍。本文亦提出了一種創新的「文圖文」(TIT) 研究方法，為探索新興視覺風格提供了新的工具與途徑，亦為未來將新興風格應用於創作提供了理論基礎。透過此方法，本研究獲得描述量子風格的最佳四個描述詞，包含「dark」(黑暗的)、「light」(燈光)、「orange」(橘色的)、「detailed」(細節的)等。本研究目前僅以三類物件為主透過文圖文的方式產生描述量子風格的共同詞彙，理論上來說，使用的物件類別數目越多，共同描述詞彙會更趨向精準。

此外，研究推測 AI 在輸入不同分隔符號結構提示句時的內部處理邏輯，並展示不同提示句設計對生成結果的影響。過程發現使用「::」作為分隔符號能夠漸進式地調整量子風格的權重，幫助理解量子風格如何在物件中浮現，同時作為未來創作過程中可調變的自由度。未來的研究將進一步深入分析文中探討的四種分隔符號，確定其最佳應用策略，並生成更為廣泛的量子視覺元素體系，發展更為完整的量子美學。

參考文獻

Broekaert, J. (2018). The Tacit ‘Quantum’ of Meeting the Aesthetic Sign; Contextualize, Entangle, Superpose, Collapse or Decohere. *Foundations of Science*, 23, 255–266. <https://doi.org/10.1007/s10699-016-9519-2>

Goloujeh, A. M., Sullivan, A., & Magerko, B. (2024, May 11). Is It AI or Is It Me? Understanding Users’ Prompt Journey with Text-to-Image Generative AI Tools [Conference presentation]. CHI ’24: Proceedings of the CHI Conference on Human Factors in Computing Systems. <https://dl.acm.org/doi/10.1145/3613904.3642861>

Midjourney. (2024). Multi Prompts. Retrieved March 27, 2024, from <https://docs.midjourney.com/docs/multi-prompts>; The Effects of --Style Raw. Retrieved May 25, 2024, from <https://docs.midjourney.com/docs/style>

Thaeyne. (2023). *Midjourney v5 | From Colons to Parentheses | The Impact of Separator Choice in Prompts | 60 Examples* [Video file]. <https://youtu.be/EbPsNE-b73Q?si=yw0Ttfi2fvSctiTe>

Vartiainen, H., & Tedre, M. (2023). Using artificial intelligence in craft education: crafting with text-to-image generative models. *Digital Creativity*, 34(1), 1-21. <https://doi.org/10.1080/14626268.2023.2174557>

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., ... & Polosukhin, I. (2017). Attention is all you need. *Advances in neural information processing systems*, 30. <https://arxiv.org/abs/1706.03762>

Sprinkleofai. (2023, August 8). A Complete Guide to Writing Perfect Midjourney Prompts – With Examples! <https://sprinkleofai.com/writing-good-midjourney-prompts/>