

# CHAPTER07

## 生成式AI工具應用



# INTERNET

# 本章摘要

## 7-1 生成式AI基本概念

[7-1-1 人工智慧](#)

[7-1-2 生成式AI](#)

INTERNET

3

## 7-2 AI聊天機器人

[7-2-1 ChatGPT](#)

[7-2-2 Jasper與Copy.ai](#)

[7-2-3 Google Bard](#)

INTERNET

18

## 7-3 AI生成圖片工具

[7-3-1 Midjourney](#)

[7-3-2 DALL-E](#)

[7-3-3 Image Creator](#)

[7-3-4 DreamStudio](#)

INTERNET

30

## 7-4 AI生成影音工具

[7-4-1 Make-A-Video](#)

[7-4-2 Pictory](#)

[7-4-3 RunwayML](#)

[7-4-4 音樂生成工具](#)

INTERNET

45

# 7-1 生成式AI基本概念

---

7-1-1 人工智慧

7-1-2 生成式AI



INTERNET

## 7-1-1 人工智慧

- 人工智慧(Artificial Intelligence, AI)是透過機器來模擬人類認知能力的技術，是以電腦程式來讓機器有智慧或讓電腦會思考。
- 它所涉及的範圍很廣，涵蓋了感知、學習、預測、決策等方面的能力。
- 人工智慧主要是希望機器能有像人類一樣的行為反應及機器能做出合理的推理與決策。

# 7-1-1 人工智慧

## 1956年

- 約翰·麥卡錫(John McCarthy)及其他當時數一數二的人工智慧專家在美國達特茅斯學院(Dartmouth College)舉辦了一次研討會，會議中正式提出人工智慧的定義，宣告人工智慧作為一門學科的誕生。

## 1993年

- 人工智慧的發展有了重大的突破，科學家想到與其告訴機器每個對應的指令，那就讓機器學會如何識字，自己判斷，找出規則，有能力自我學習，這方法讓人工智慧有了重大突破，這也是機器學習(Machine Learning)的開端。

# 7-1-1 人工智慧

1990年

- 神經網路(Neural Network)技術的發展，讓人們對AI開始有認知，慢慢進入了平穩發展時期。

1997年

- IBM的深藍(Deep Blue)戰勝國際象棋世界冠軍Garry Kasparov，這也成為AI的重要里程碑。

2013年

- 深度學習(DeepLearning)在語音和視覺識別有了突破性進展。

# 7-1-1 人工智慧

## 弱AI

- 是指經過訓練的AI，著重在執行特定作業，例如：達文西手術機器人、Apple的Siri。

## 強AI

- 是指期望電腦除了具有認知能力之外，還能夠推理、自學、溝通甚至擁有自我意識。

## 超級AI

- 是指將超越人類大腦的智慧與能力，擁有「自我」的概念，像人類一樣辨識到自己的存在，並會產生想要生存的意向。

# 7-1-1 人工智慧

## 機器學習

- 機器學習指的是讓機器具備自我改進能力及自動學習能力，是一門設計與開發演算法的學科；讓電腦可以根據經驗演化它的行為，自動最佳化下一次結果。
- 例如：**Gmail**中的垃圾郵件過濾為什麼可以那麼準確，甚至還可以根據每個人的特殊需求慢慢學習改進，這就是機器學習的成果。
- 機器學習可廣泛應用在機器視覺、大數據資料分析、資料探勘、語言與語音辨識、手寫辨識、生物特徵辨識、環境辨識、醫學診斷、詐欺檢測、證券市場分析等，同時也是人工智慧的核心技術。



# 7-1-1 人工智慧

## 神經網路與深度學習

- 神經網路(或稱類神經網路)就是利用電子科技模擬神經組織的運作而組成的。
- 神經網路中最基本的單元是神經元(Neuron)，一般稱作節點或單元。
- 神經網路是由很多層的神經元所建構而成的深度學習網路。

## 7-1-1 人工智慧

- 深度學習源於神經網路，神經網路是深度學習演算法的核心，讓電腦像是長了神經網路般，可進行複雜的運算，展現擬人的判斷及行為。
- 讓電腦進行深度學習主要有三個步驟：設定好神經網路架構、訂出學習目標、開始學習。

## 7-1-2 生成式AI

- 生成式AI是透過學習模型研究歷史數據的模式，運用大規模的語言模型，從網路上或其他大型資料庫取得海量資料進行訓練，讓使用者可以用自然語言與AI互動。
- AI則根據使用者所提出的問題或指令，自動生成新的數位內容，如文字、語音、圖像、視訊、商品、場景等，而這些生成的資料與訓練資料會維持相似，但不是複製。

## 7-1-2 生成式AI

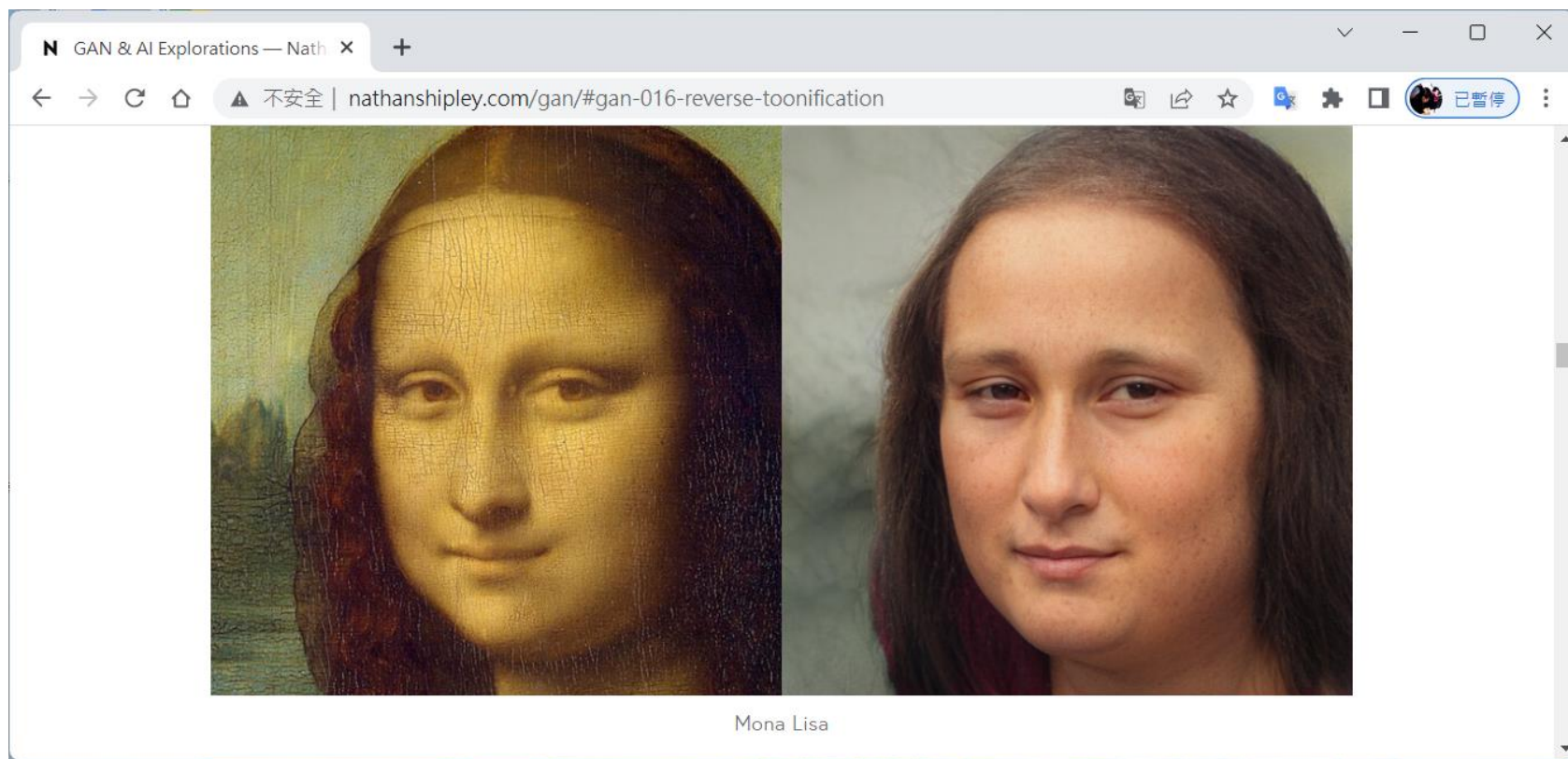
- 生成式AI主要依賴於深度學習技術，其中最常見的是生成對抗網路 (Generative Adversarial Network, GAN)及Transformer模型等。
- 從大量資料中透過GAN手法生成擬真資料，使用現有多模內容來建立新內容，目前生成式AI可以寫文章、編故事、虛擬人物、影音創作、數位設計、資料擴增、程式設計等，但也會被濫用於詐騙、偽造身分、政治造謠等。

## 7-1-2 生成式AI

### 生成對抗網路

- 是2014年蒙特婁大學博士生Ian Goodfellow提出的。
- 主要功能是模仿，讓電腦產生出以假亂真的圖片、影片、文字或是知名畫作。
- 例如：美國AI藝術家Nathan Shipley透過AI StyleGAN技術，將蒙娜麗莎、莎士比亞等畫像真人化，讓歷史人物來到現代。

# 7-1-2 生成式AI



## 7-1-2 生成式AI

- 生成對抗網路是由生成網路 (Generator Network) 及鑑別網路 (Discriminator Network) 所組成，透過讓兩個神經網路相互博弈的方式進行學習。
- 生成網路器主要是製造假圖片，由輸入數值來產生偽造圖片，鑑別網路可以同時觀察真實和偽造的圖片，判斷這個圖片到底是不是真的。

## 7-1-2 生成式AI

### Transformer模型

- 是非常強大的神經網路模型，藉由追蹤序列資料中的關係，學習上下文之間的脈絡及意義，就如同句子中的每一個字，它可以用於自然語言處理、圖像處理、音訊處理等各種生成式任務。
- 主要的特點就是它使用了一種被稱為注意力(Attention)或自我注意力(Self-Attention)機制的技術，這種技術可以幫助模型自動從數據中學習知識，並生成適當的回答。



## 7-1-2 生成式AI

### GPT

- **GPT (Generative Pre-trained Transformer, 生成式預訓練模型)**是一個基於Transformer架構的語言模型，是OpenAI開發的，透過機器學習技術進行自然語言處理，例如：回覆問題、生成文章和程式碼，或者翻譯文章內容等。
- **GPT開發出了許多版本，如：GPT-1、GPT-2、GPT-3、GPT-3.5、GPT-4等**，其中**GPT-4**不只能處理2.5萬字長篇內容，正確度高出40%，以整理和搜尋網路上的資訊為主，還支援視覺輸入及圖像辨識。

# 7-2 AI聊天機器人

---

7-2-1 ChatGPT

7-2-2 Jasper與Copy.ai

7-2-3 Google Bard



# INTERNET

## 7-2-1 ChatGPT

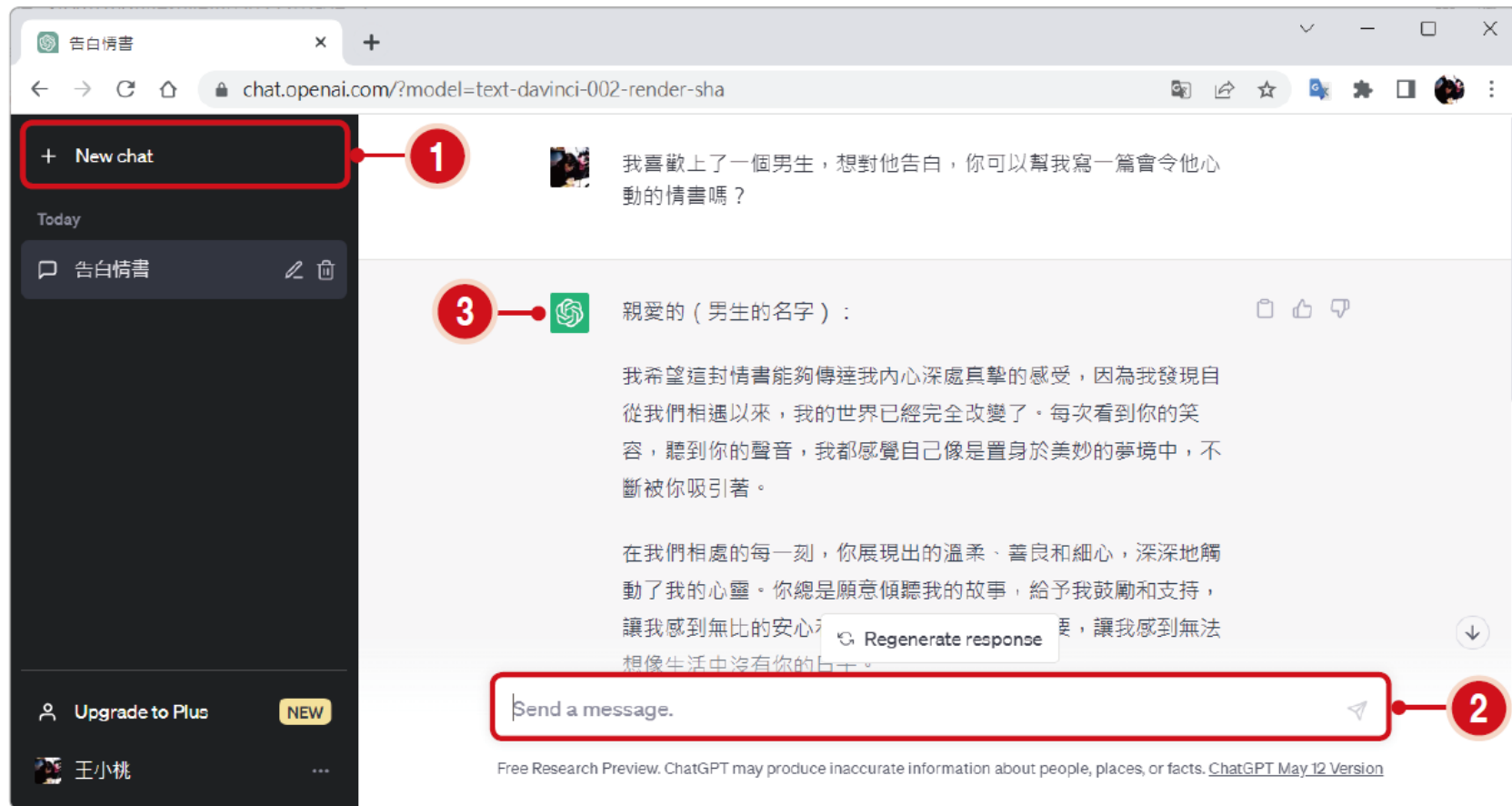
- ChatGPT能用各種語言回答各種問題，還能寫論文、算數學、寫詩、寫歌詞、寫程式等，被視為是AI的大突破。
- ChatGPT是使用基於GPT-3.5、GPT-4架構的大型語言模型(Large Language Model, LLM)，透過機器學習中的強化學習進行訓練和互動及人類回饋增強學習(Reinforcement Learning with Human Feedback, RLHF)，完成複雜的自然語言處理，因此讓對話的過程很有真實感，就像是在與朋友對話一樣。

# 7-2-1 ChatGPT

## 與ChatGPT對話

- 在與**ChatGPT**進行對話時，建議使用具體的指令，還要有明確的目的，避免太過廣泛或是開放式的問題，這樣比較能得到完整回覆，例如：要請**ChatGPT**撰寫情書、文章或E-mail時，將撰寫的目的、主題、對象清楚寫出，即可得到最佳的答案。

# 7-2-1 ChatGPT

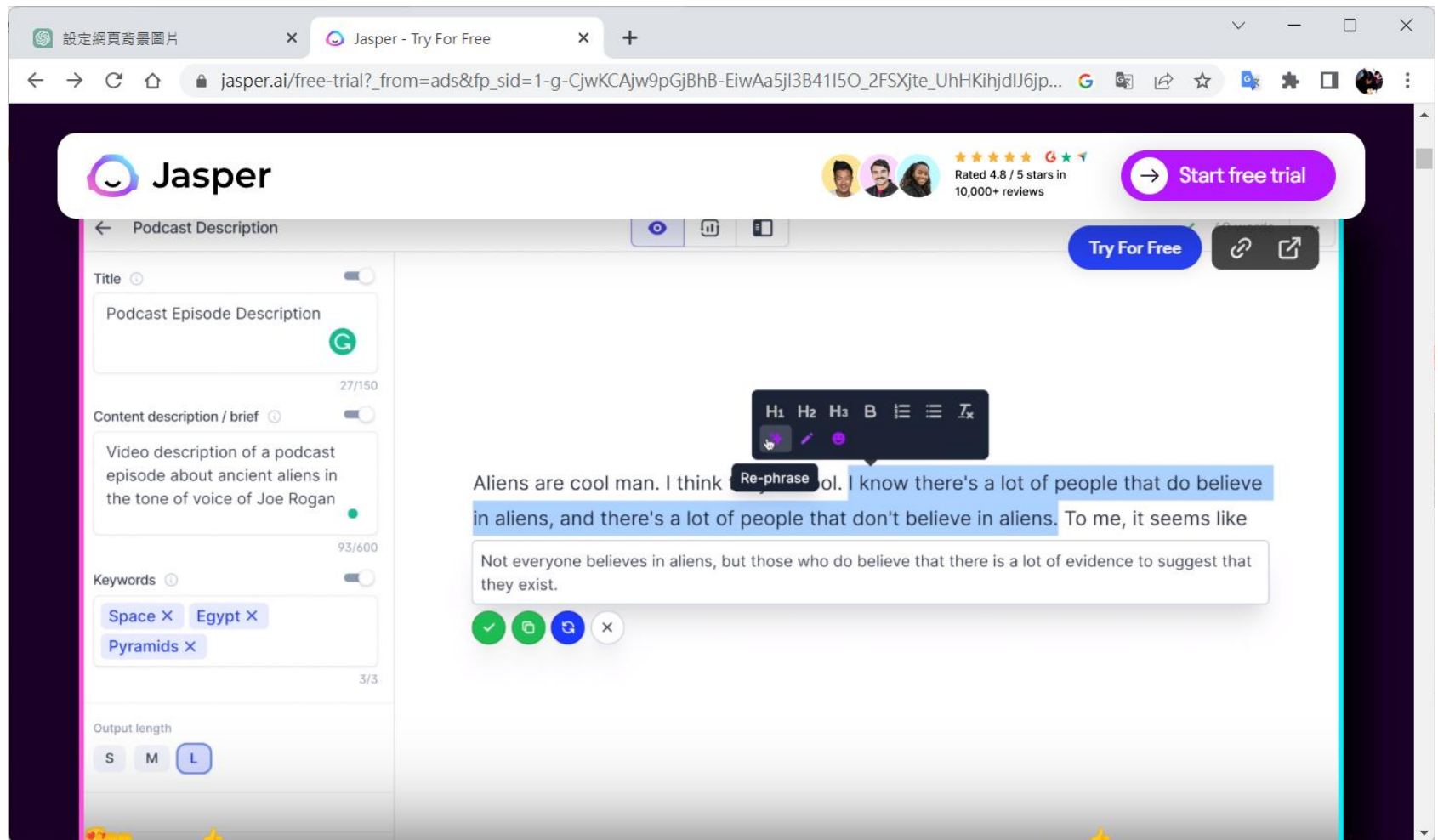


## 7-2-2 Jasper與Copy.ai

### Jasper

- 適合用來產生行銷文案、部落格文章、社群媒體貼文、電子郵件等文字，還能協助檢查文法並避免抄襲，而所產出的內容**100%**是原創長文，還提供常見的內容範本供使用者選擇使用，是商務與行銷人員的最佳工具，不過，它是需要付費使用的。

# 7-2-2 Jasper與Copy.ai



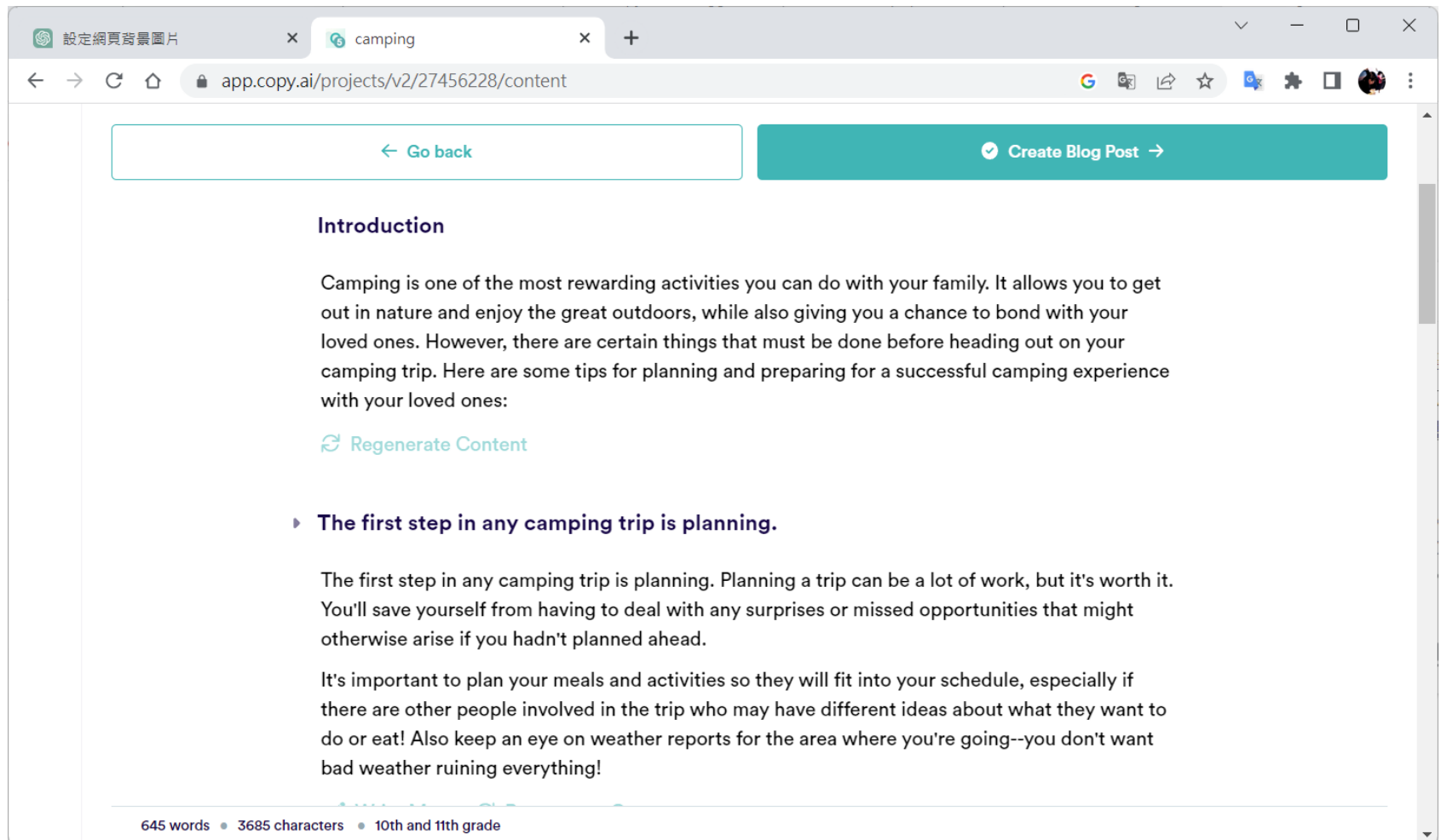
## 7-2-2 Jasper與Copy.ai

### Copy.ai

- 也是產生文案、文章的生成工具，只要輸入標題跟幾個關鍵詞，然後選擇想要敘述的寫作風格，就可以快速生成一段文字或者一篇文章，如：文章、廣告文案、產品描述、標題、標語、社群貼文等，並且還可以根據你的需求進行修改和調整。



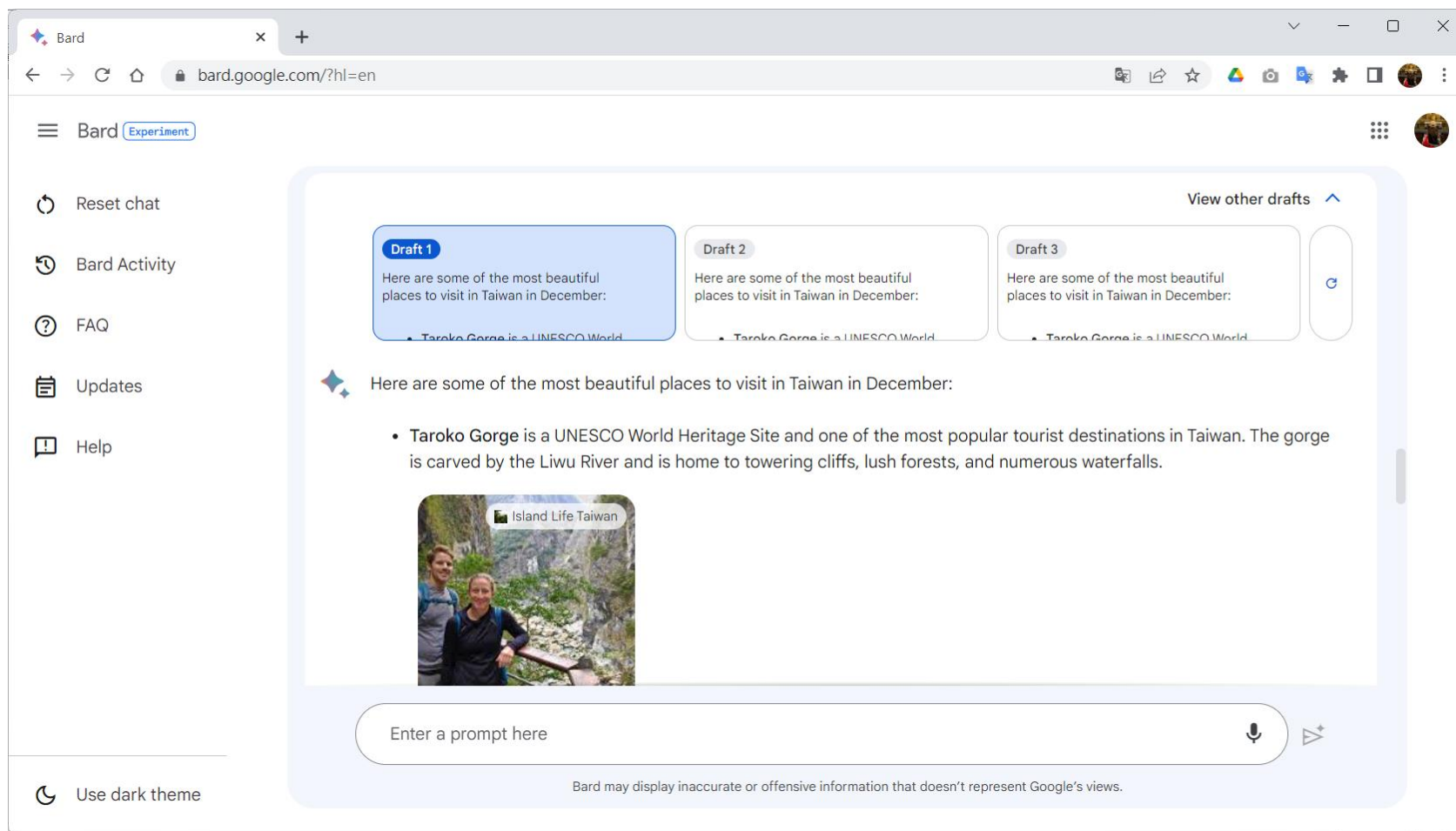
# 7-2-2 Jasper與Copy.ai



## 7-2-3 Google Bard

- **Google Bard**是**Google**推出的**AI**聊天機器人，使用**PaLM 2 (Parallel Learning Machines 2)**語言模型，增強邏輯、數學、因果關係的理解能力，增強程式語言能力，支援超過**20**種程式語言，可以生成也可以協助除錯。
- **Bard**的回應會提供三種答案讓使用者選擇，還有圖片作為輔助答案，還可以將回應結果輸出成**Docs**文件，或是存在**Gmail**當作草稿。

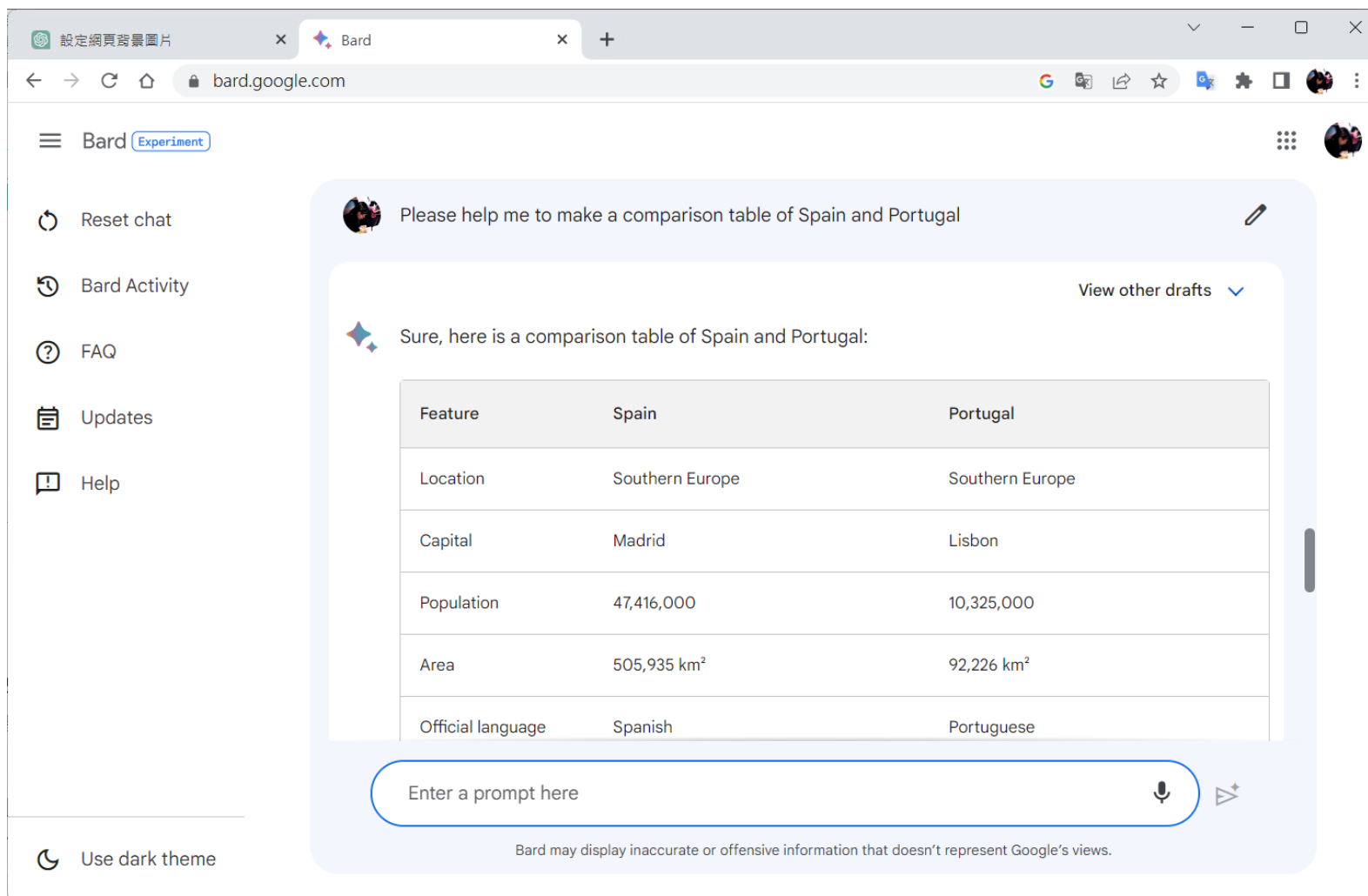
# 7-2-3 Google Bard



## 7-2-3 Google Bard

- Bard也將智慧鏡頭功能加入，例如：提供兩隻貓的照片，就可以請Bard生成相關的文案，Bard會自行理解貓的品種，並提供幾個相關文案供用戶參考選擇。
- Bard能做的事非常多，你可以請它幫你規劃簡報大綱、撰寫電子郵件、可撰寫程式碼及程式碼除錯、製作表格等。除此之外，Bard與Adobe Firefly合作，導入生成圖片的功能，並將Bard整合到Google文件、Google雲端硬碟、Gmail、Google地圖等服務中。

# 7-2-3 Google Bard



The screenshot shows the Google Bard web interface. The browser address bar displays 'bard.google.com'. On the left sidebar, there are links for 'Reset chat', 'Bard Activity', 'FAQ', 'Updates', and 'Help'. The main chat area shows a user prompt: 'Please help me to make a comparison table of Spain and Portugal'. Bard's response is: 'Sure, here is a comparison table of Spain and Portugal:'. Below the text is a table comparing Spain and Portugal across six features. At the bottom, there is a text input field with the placeholder 'Enter a prompt here' and a microphone icon. A disclaimer at the very bottom states: 'Bard may display inaccurate or offensive information that doesn't represent Google's views.'

Feature	Spain	Portugal
Location	Southern Europe	Southern Europe
Capital	Madrid	Lisbon
Population	47,416,000	10,325,000
Area	505,935 km <sup>2</sup>	92,226 km <sup>2</sup>
Official language	Spanish	Portuguese

## 7-3 AI生成圖片工具

---

7-3-1 Midjourney

7-3-2 DALL-E

7-3-3 Image Creator

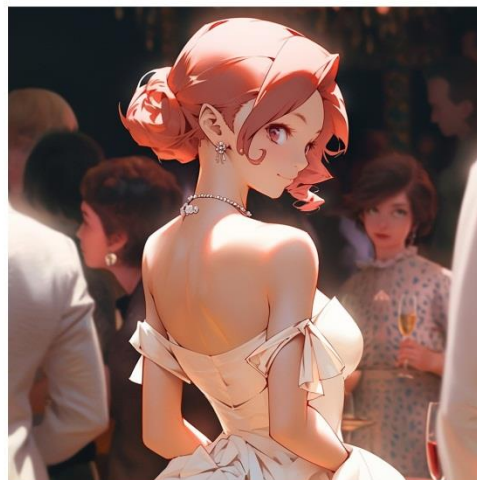
7-3-4 DreamStudio

INTERNET

## 7-3-1 Midjourney

- Midjourney 掀起了「人人都是藝術家」的風潮，還有創作者的作品在美國的科羅拉多州博覽會美術大賽中，奪得了「數位藝術類」獎項。
- Midjourney 功能強大，已成為最主流的AI生成圖片工具之一，該工具只要輸入關鍵字，系統會搜尋資料庫中相關的圖片，加以解構重組後，生成最適合的繪畫風格，創作出令人讚嘆的作品。

# 7-3-1 Midjourney

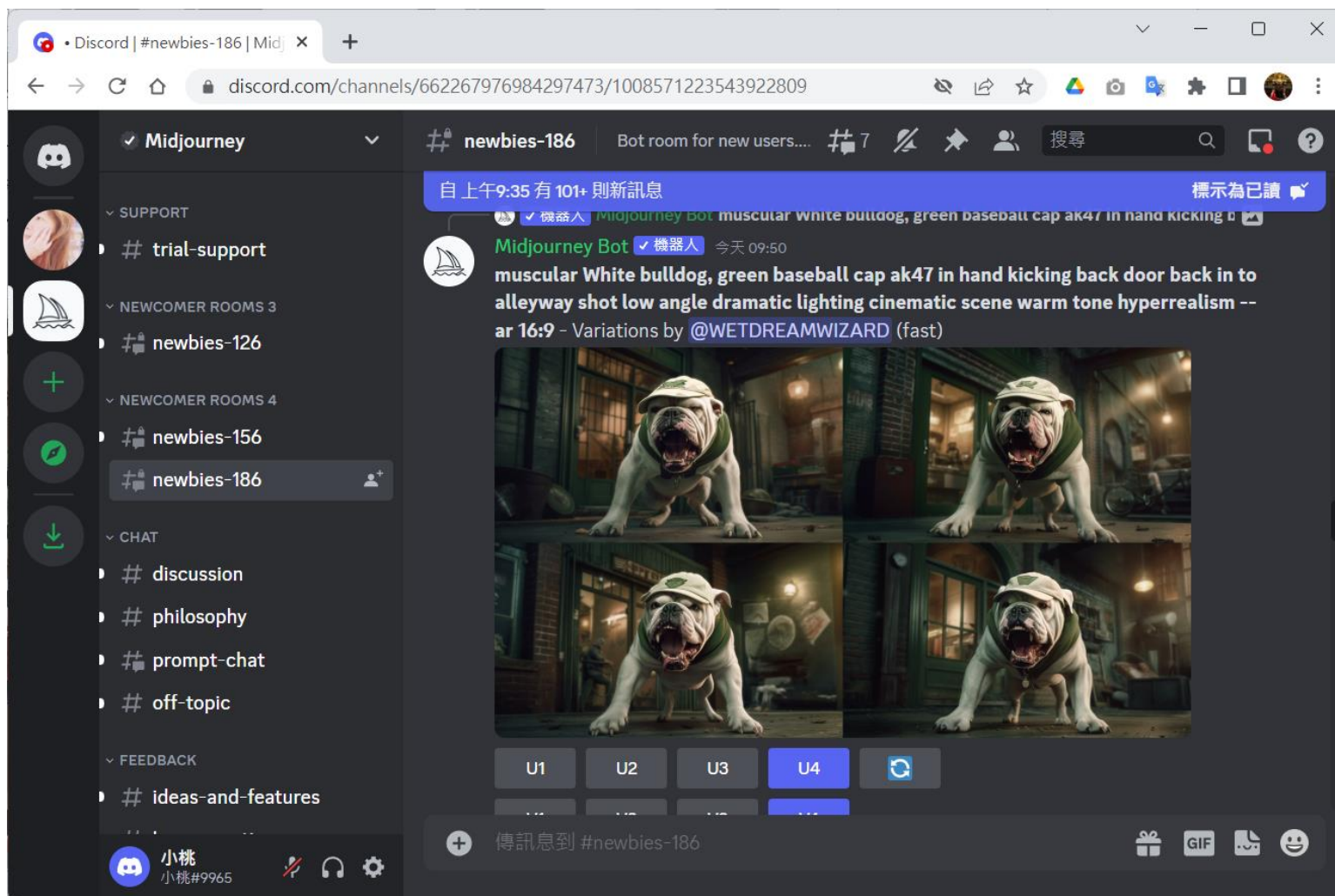




## 7-3-1 Midjourney

- **Midjourney**目前最新版本為**V5.1**，能夠處理更大、更複雜的語言數據集，能夠生成更真實、更細緻、更少錯誤、風格更加奔放、無縫紋理、更寬的縱橫比的圖片。
- **Midjourney**依附在**Discord**中，若已有**Discord**帳號可直接登入，若無則須先進行註冊及驗證。
- 進入**Discord**後，在左邊欄位中可以看到許多聊天室和對話串，還有歡迎新手加入的聊天群組，除此之外，還有一些作品可以觀看。

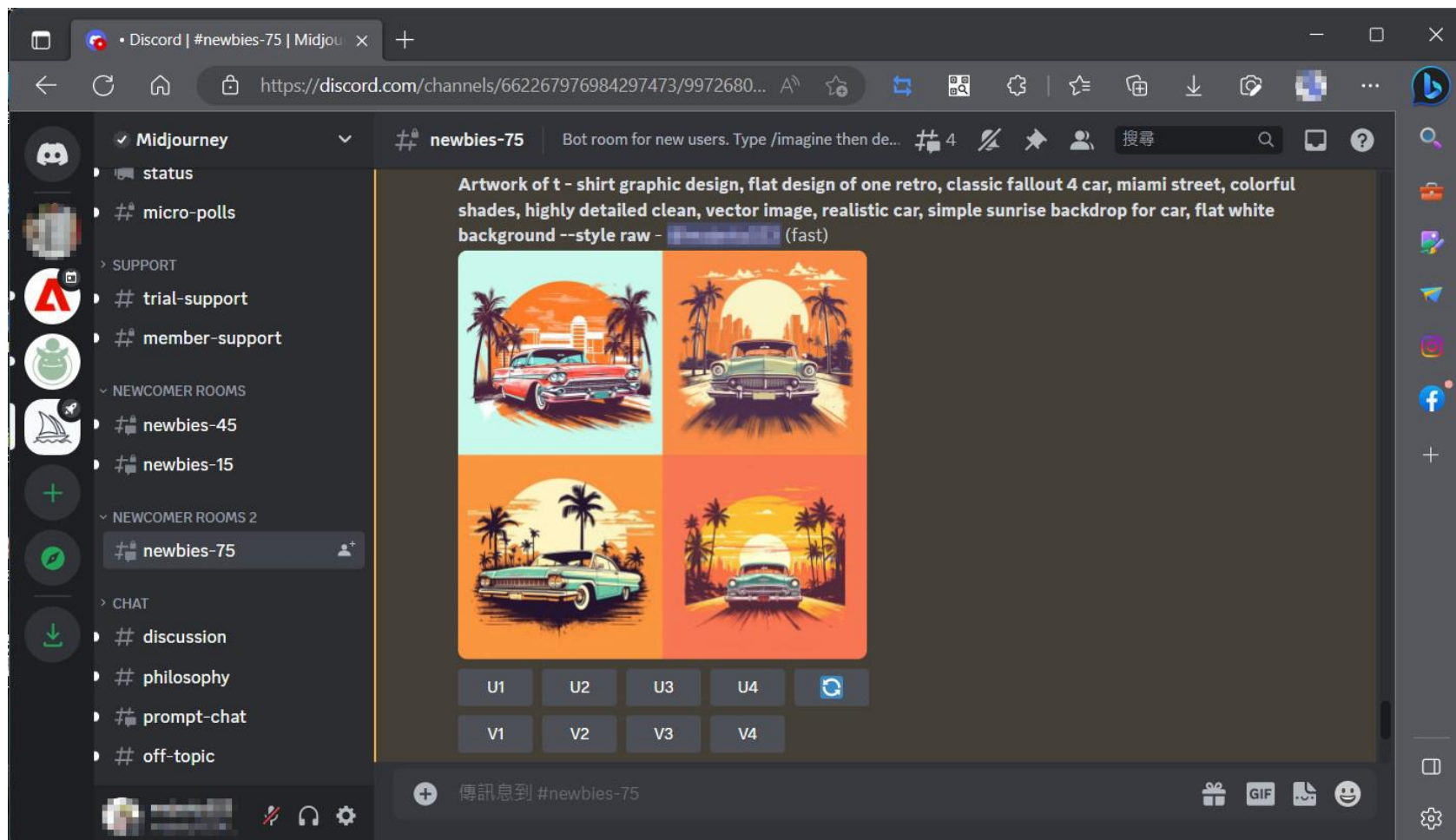
# 7-3-1 Midjourney



## 7-3-1 Midjourney

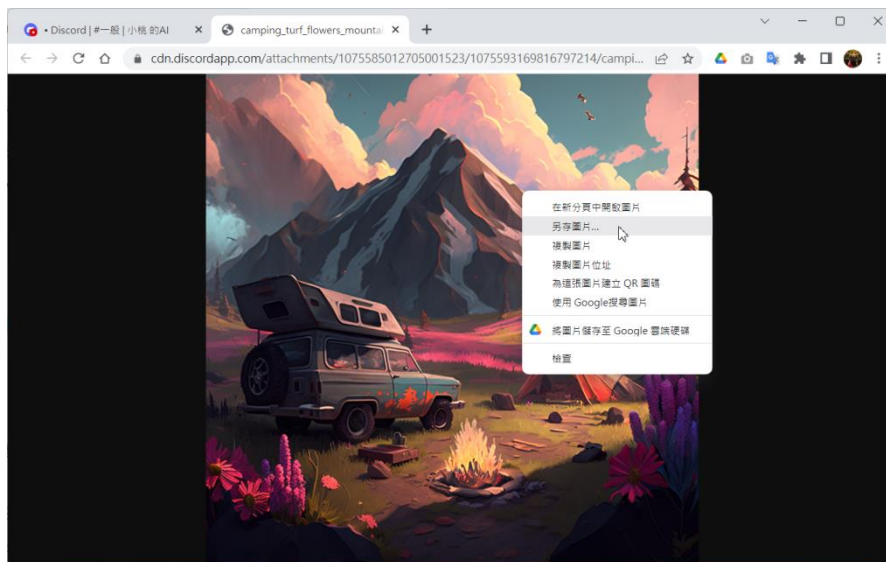
- Midjourney使用時像是和機器人對話的模式，輸入產生圖片的描述提示詞(prompt，或者稱為詠唱、咒語)及相關參數(--ar(建立長寬大於2:1圖片)、--hd(高畫質)、--niji(漫畫風格)、--test(藝術風格)、--tile(無縫重複圖像或填充空白))後就會快速產生四張圖片。
- 在下提示詞時掌握「人事物+風格+細節設定+相關參數」原則。

# 7-3-1 Midjourney



## 7-3-1 Midjourney

- 在Midjourney生成的圖片，若要下載時，先開啟該圖片，再按下「在瀏覽器開啟」按鈕，圖片就會顯示在另一個視窗中，在圖片上按下滑鼠右鍵，於選單中點選「另存圖片」選項，即可將圖片儲存起來。



## 7-3-2 DALL-E

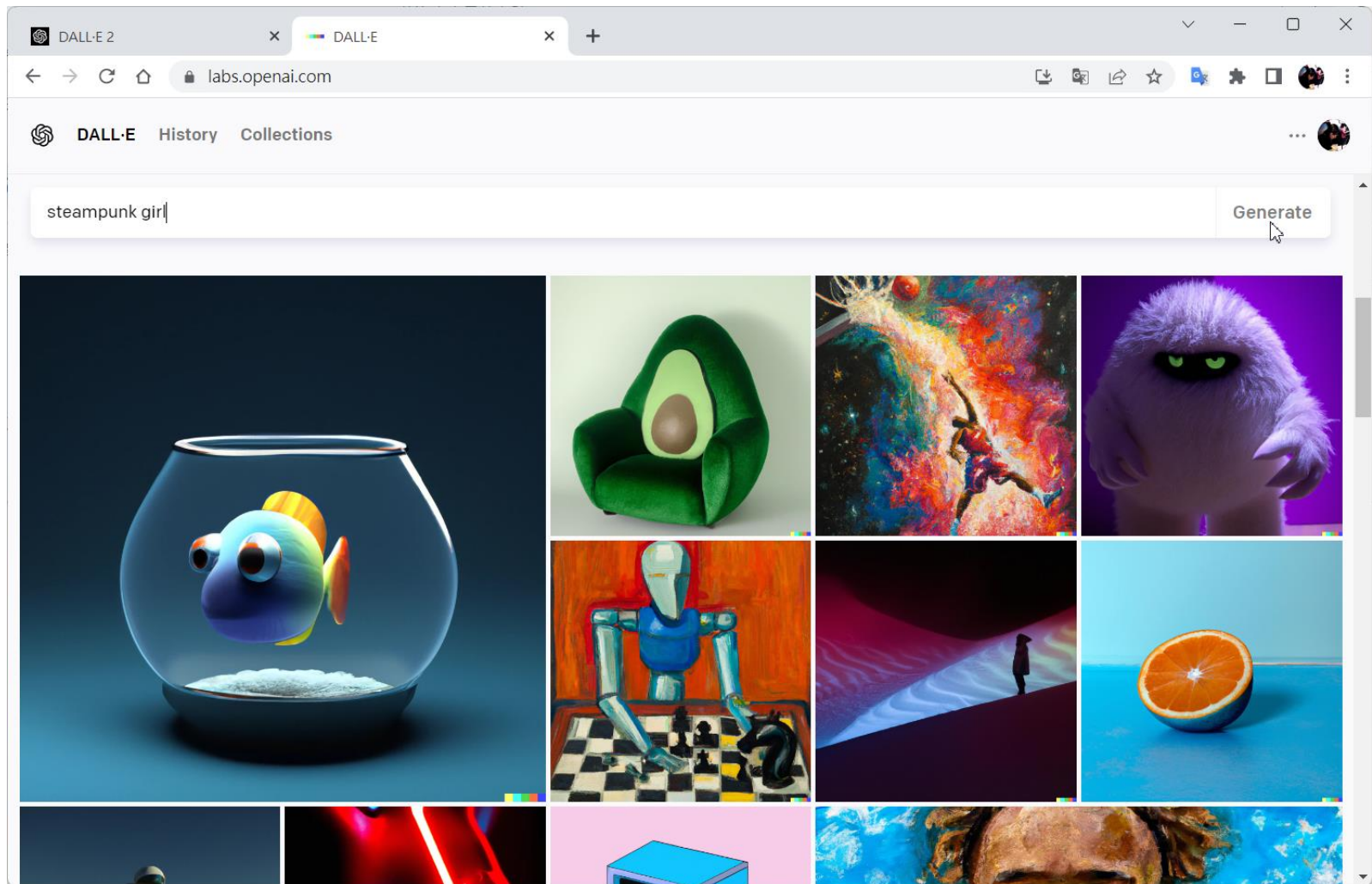
- **DALL-E是OpenAI所發布的，可以依據使用者輸入的語意描述，來產生一組圖片，也可上傳照片讓DALL-E以圖繪圖。**
- **DALL-E最新版本為DALL-E 2，能生成出更精緻、更逼真及能編輯的圖片，使用者可以使用文字敘述新增、替代或是移除圖片上的物品，甚至融合兩張現有圖片。而圖片解析度也從原本的256 x 256像素提升至1,024 x 1,024像素。**

## 7-3-2 DALL-E

- **DALL-E 2**是需要付費使用的，使用介面相當簡潔，只要透過網頁介面輸入描述就可以生成圖片，輸入時，官方規定嚴禁涉及仇恨、騷擾、暴力、自殘、裸露及非法活動等主題，以及假新聞、政局、醫療乃至疾病相關圖，只要在不違反OpenAI的政策，是允許商業用途的。



# 7-3-2 DALL-E

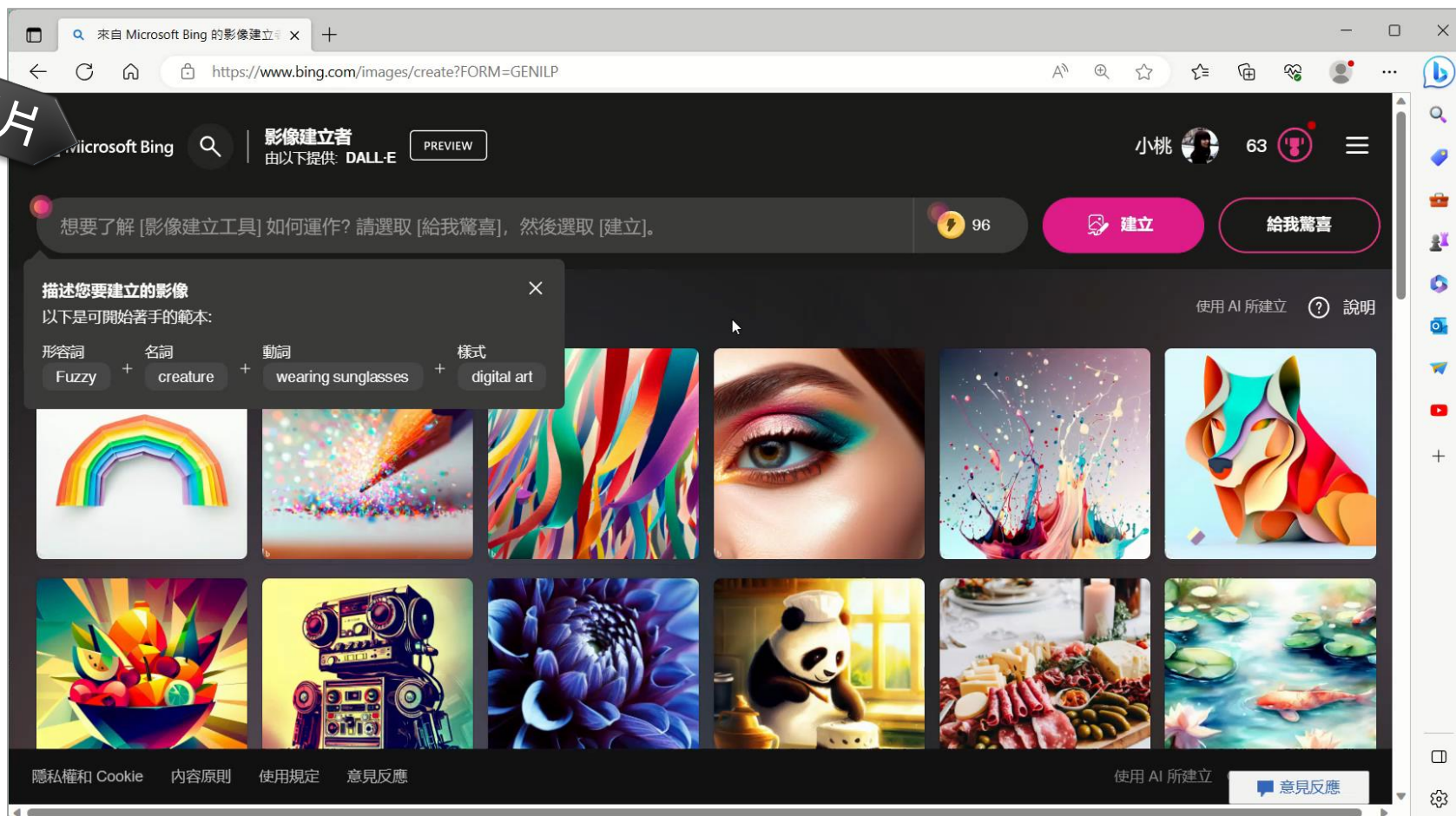




## 7-3-3 Image Creator

- **Image Creator (影像建立者)**是微軟開發的，使用OpenAI的**DALL-E**圖片生成技術，只要輸入圖片關鍵描述提示詞，就能夠在幾秒內立即生成圖片，該工具也整合到新版的**Bing**和**Edge**瀏覽器中。
- **Image Creator**只要使用微軟帳號登入，就可以免費使用，可透過電腦、手機或平板用任何一款瀏覽器都能夠直接開啟。

# 7-3-3 Image Creator

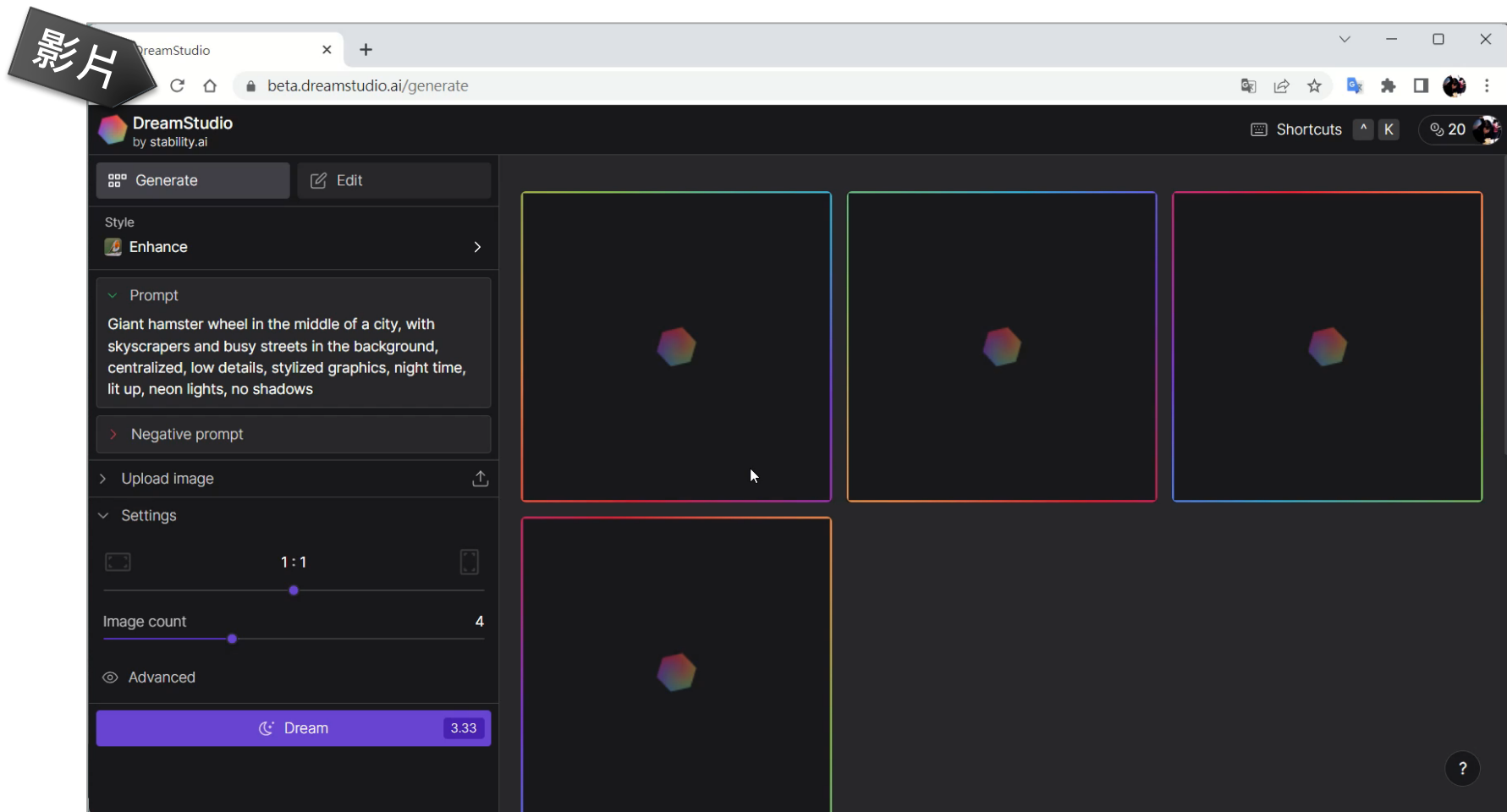


影片如無法播放，可安裝QuickTime，  
或開啟影片MP4檔

## 7-3-4 DreamStudio

- **DreamStudio**是由**Stability AI**公司所開發的，與**Midjourney**相比更加簡單、快速，是一款非常方便的生成工具，且圖片可允許在商業用途上(CC0)。
- 使用時，進入**DreamStudio**網頁後，可以使用**Google**、**Discord**帳號進行登入。
- 有提供免費的**25 credits** (約125張圖)。

# 7-3-4 DreamStudio



# 7-4 AI生成影音工具

---

7-4-1 Make-A-Video

7-4-2 Pictory

7-4-3 RunwayML

7-4-4 音樂生成工具



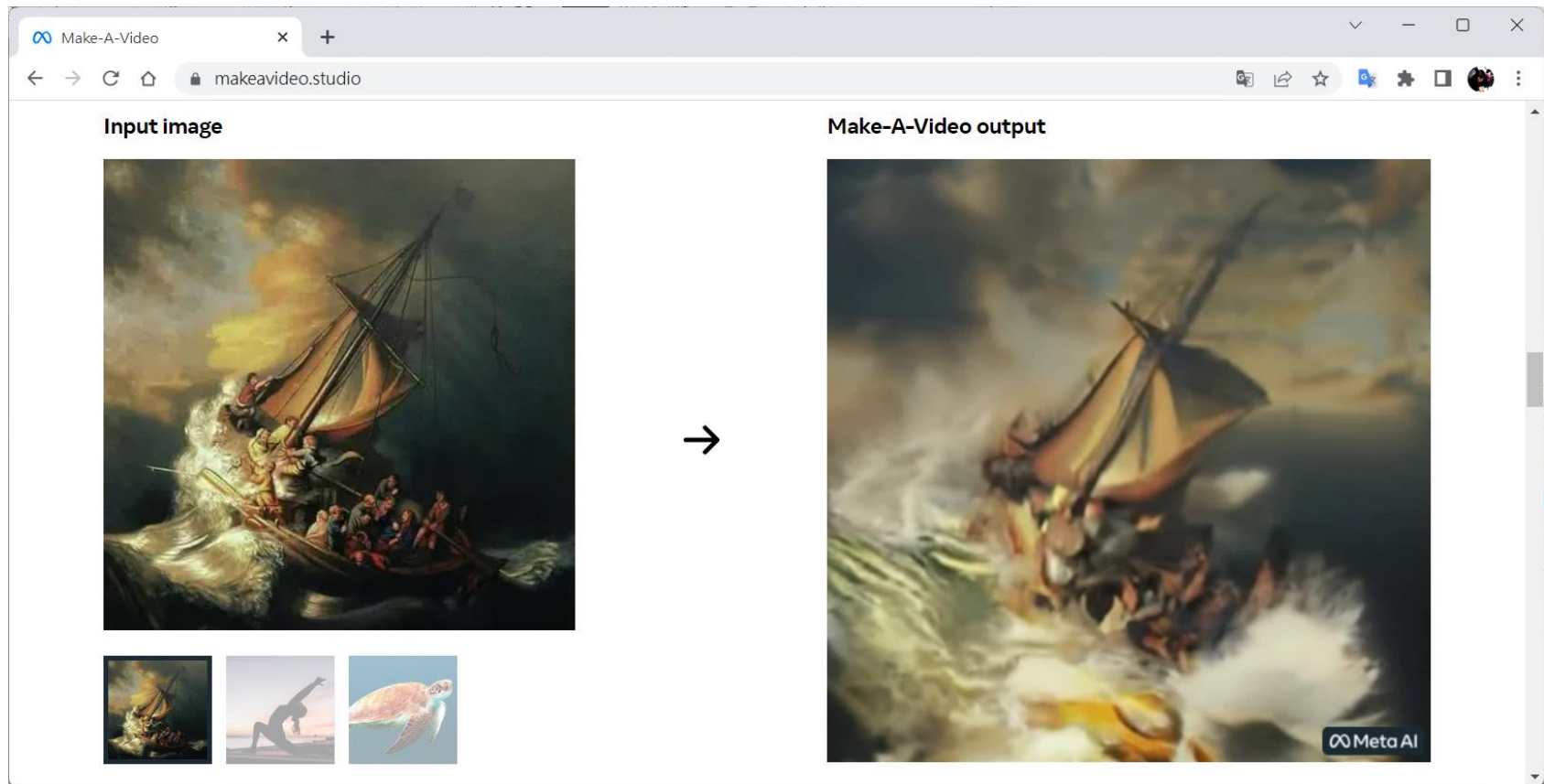
## 7-4-1 Make-A-Video

- **Make-A-Video**是Meta推出的AI影片生成器，使用者只要輸入文字描述，再加上單張或多張圖片，即可透過進行訓練的人工智慧模型運作，在短時間內生成自然生動且獨特的影片內容。
- **Make-A-Video**是藉由圖片、圖片的描述，以及未被標註的影片來訓練AI，訓練的素材來自**WebVid-10M**和**HD-VILA-100M**資料庫，總計數百萬部影片，其中也包括來自**Shutterstock**等圖庫的內容。

## 7-4-1 Make-A-Video

- **Make-A-Video**提供了現實、寫實及風格化三種影片類型，目前所有生成影片皆會有浮水印，確保觀眾知道該影片是透過**AI**生成，而不是真實拍攝。
- **Make-A-Video**會透過**GitHub**公開相關技術資源，並且提供開發社群研究使用。

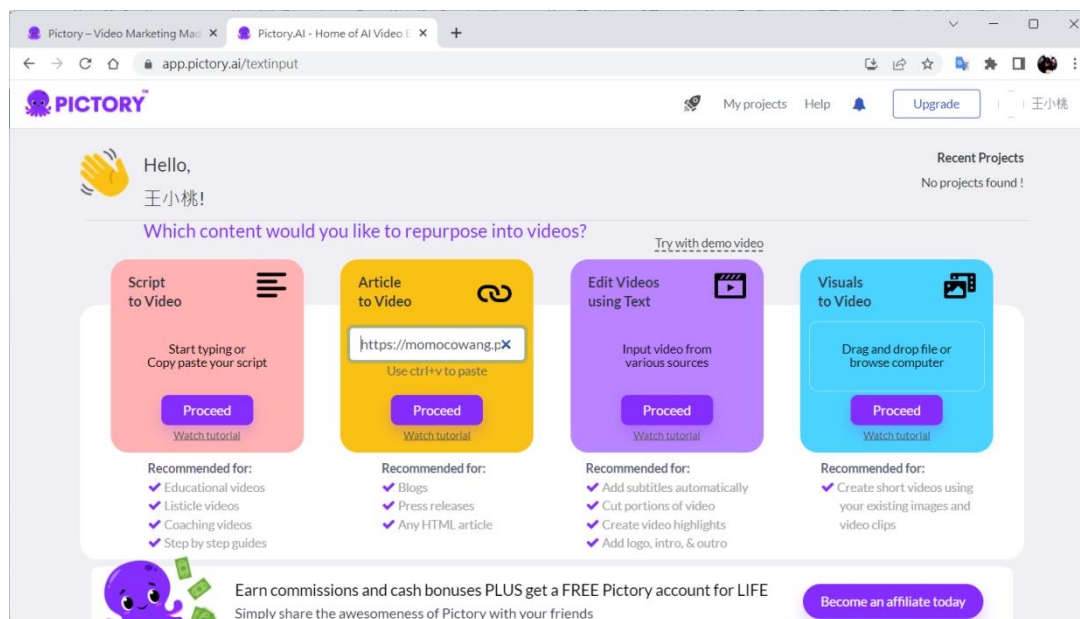
# 7-4-1 Make-A-Video





## 7-4-2 Pictory

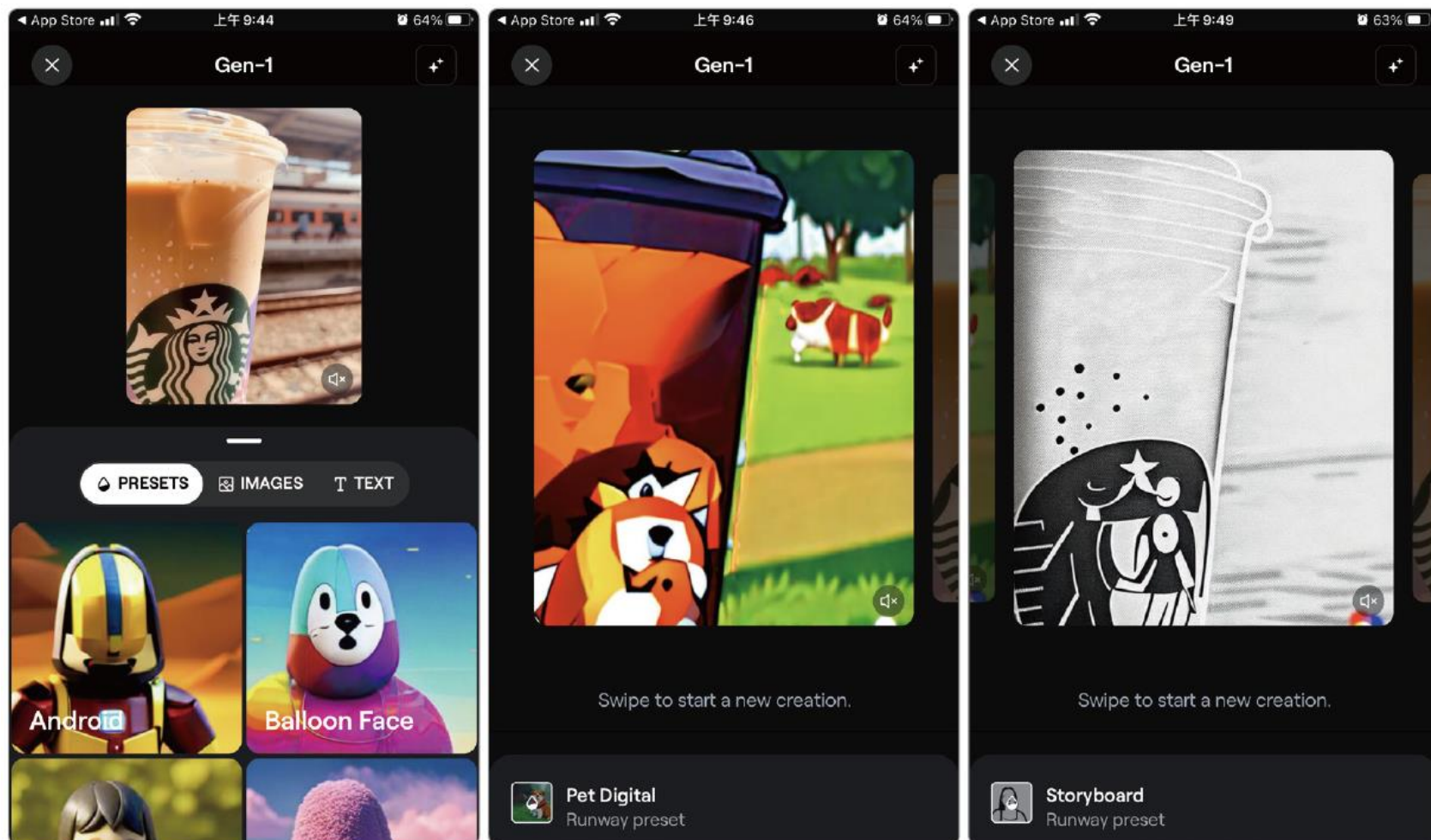
- **Pictory**可以將文字或網站連結轉換成影片，會自動分析文字的內容、選擇適合的圖片、音樂和配音，生成為影片。可免費使用，而且沒有浮水印或廣告。



## 7-4-3 RunwayML

- RunwayML是Runway在iOS推出的App，只要上傳一段影片、圖片或文字，就能生成相應的魔幻效果影片。RunwayML會識別出視訊中的物品，並把所有物品都變成統一的風格。
- RunwayML預設許多影片風格，包括雲朵、摺紙、水彩、紙墨、素描、泥塑等，使用者只要拍攝影片或上傳一段影片，再選擇喜歡的風格，即可生成相對應的AI效果，例如：選擇素描風格時，影片就會變成素描世界。

# 7-4-3 RunwayML



## 7-4-4 音樂生成工具

### MusicLM

- 是**Google**所研發的，可以根據使用者的想法或描述來生成獨一無二的歌曲，就會自動生成符合描述情境的歌曲音樂，也可以自定類型、音樂氛圍、情緒和情感等。
- **MusicLM**是由**280,000**小時的音樂數據集訓練而成，不過其生成音樂中有**1%**是直接從有版權作品複製而來，所以目前還在測試階段，並做持續的調整和修改。

## 7-4-4 音樂生成工具

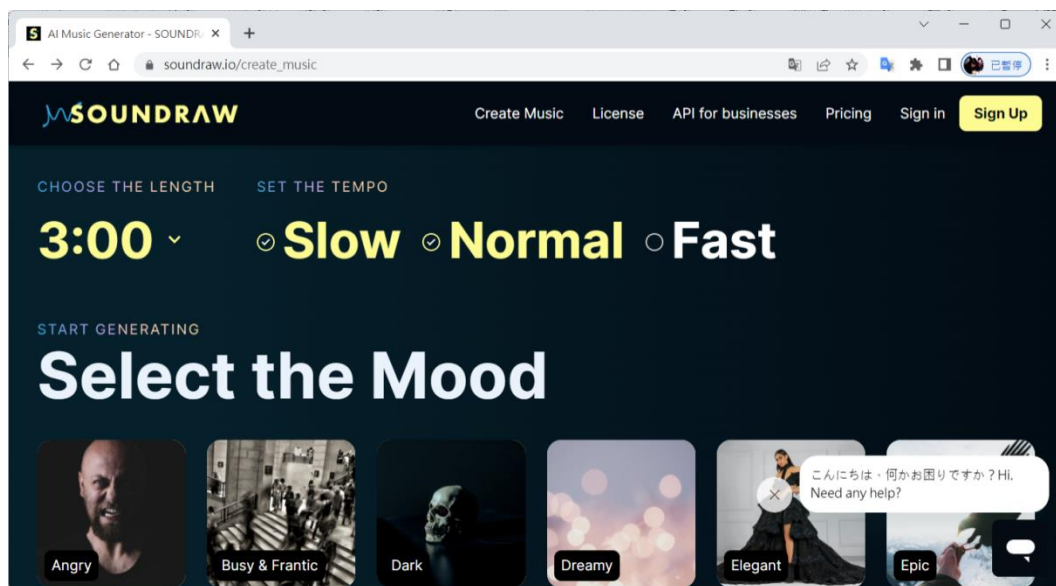
### MuseNet

- 是OpenAI開發的，它能透過深度神經網路獲得的資訊進行反覆訓練，並模仿出如莫札特等作曲家的作品，可以用**10**種不同的樂器生成**4**分鐘的音樂作品。

## 7-4-4 音樂生成工具

### Soundraw

- 可以輕鬆客製免版權的歌曲與音樂，使用者只需要設定音樂長度、類型、樂器、心情等，便可即興創作並調整音樂。



## 7-4-4 音樂生成工具

### AIVA

- 是世界上第一個正式註冊的AI作曲家，還發表了許多專輯，使用者可免費使用並提供下載，但限於個人非營利使用，若要將音樂商用或營利，則要使用付費訂閱服務，才能獲得完整音樂授權。

## 7-4-4 音樂生成工具

- AI所生成出的音樂仍有侵權疑慮，例如：**Spotify**因為環球音樂集團的投訴，刪除了來自新創公司**Boomy**的數萬首AI生成歌曲。
- 不過，在防堵侵權外，也有創作者利用AI激發自我潛能，帶來超越過往的作品，例如：陳珊妮發布的單曲《教我如何做你的愛人》，從歌聲到封面設計皆由AI生成，成果受到好評。