

# 挑戰 GPT-4：GOOGLE 大型語言模型 GEMINI 來勢洶洶

Google 於 6 日推出了 AI 大型語言模型——Gemini 一代，建立在多模態的基礎上，Gemini 可以同時處理文本、圖像、影片、音頻和程式碼等資料，且根據從資料中心到行動端的運算能力差異，預計推出三個版本，分別是模型最大的 Ultra、泛用性最好的 Pro、和應用於手機端的 Nano。Google 表示，在提升人們生活品質的能力方面，Gemini 大幅超越了以往的 AI。



## 怎麼超越？

### GEMINI V.S. GPT-4

Google 稱 Gemini Ultra 在 32 項學術性的標準測試中，有 30 項表現優於 GPT-4，且在一個名為 MMLU 的多任務測試中，超越了人類專家，該測試涵蓋了數學、物理、法律、醫學和倫理等 57 個科目，而 GPT-4 的分數比人類專家還要低。

GPT-4 本身不具備圖像處理的能力，模型訓練資料目前截於2023年4月，而 Gemini 可以處理多種型態的資料，訓練資料也更新，另 Google 計劃將 Gemini AI 整合到其各種產品中，可能增強 Google Search、Gmail 和 Google Workspace 等服務的能力，不過還未透露是否有能在企業內部部署的服務。

## 實際測試 GEMINI：BARD

目前 Google 聊天機器人 Bard 已經開始使用 Gemini Pro 技術，開放了實驗版給大眾體驗，那麼 Bard 又會怎麼回答上面的問題呢？讓我們看看他的說法：「首先，Gemini 的訓練資料集比 GPT-4 大得多。Gemini 的訓練資料集包括文字和程式碼，而 GPT-4 的訓練資料集僅包括文字。這意味著 Gemini 能夠生成更準確和更有創意的文字，包括程式碼。」  
小編測試後覺得生成內容的幻想成分還是比 GPT-4 多，也沒有 GPT-4 穩定，但因為目前使用的還不是 Ultra，所以仍待觀察。

---

## 挑戰 GPT-4：GOOGLE 大型語言模型 GEMINI 來勢洶洶

---

不過結合 Google 強大的搜尋能力以及圖像處理，像搜尋一張風景圖在哪，Gemini 可以給出好非常多的答案，小編以前有上傳過一張陽明山擎天岡的照片到 Facebook，Bard 回答出了正確答案，而且還附註是在哪個 Facebook 頁面 不過對於沒有上傳過的圖片，辨識效果就沒有那麼好，像交大校區的交映樓就被認成臺大的某個建築物 而 GPT-4 對於上述問題都只能回答「識別資訊不足，需要更明顯的地標」。

### 如何應用？

Gemini 可以處理多種型式的輸入與輸出，比如辨識圖像中的物品，接著搜尋物品名稱、翻譯等，或再生成圖片，像是給兩種顏色的毛線球圖片，就能產出兩種顏色毛線玩偶的圖片。

除此之外，Gemini 可以理解並生成 Python、Java、C++ 等程式語言，在 Google 釋出的範例中，給 Gemini 一段粒子流動動畫，就可以生成相應的 JavaScript (p5.js) 程式。