人工智能藝術如何展示女性的倫理問題

余創豪 CHONGHOYU@GMAIL.COM

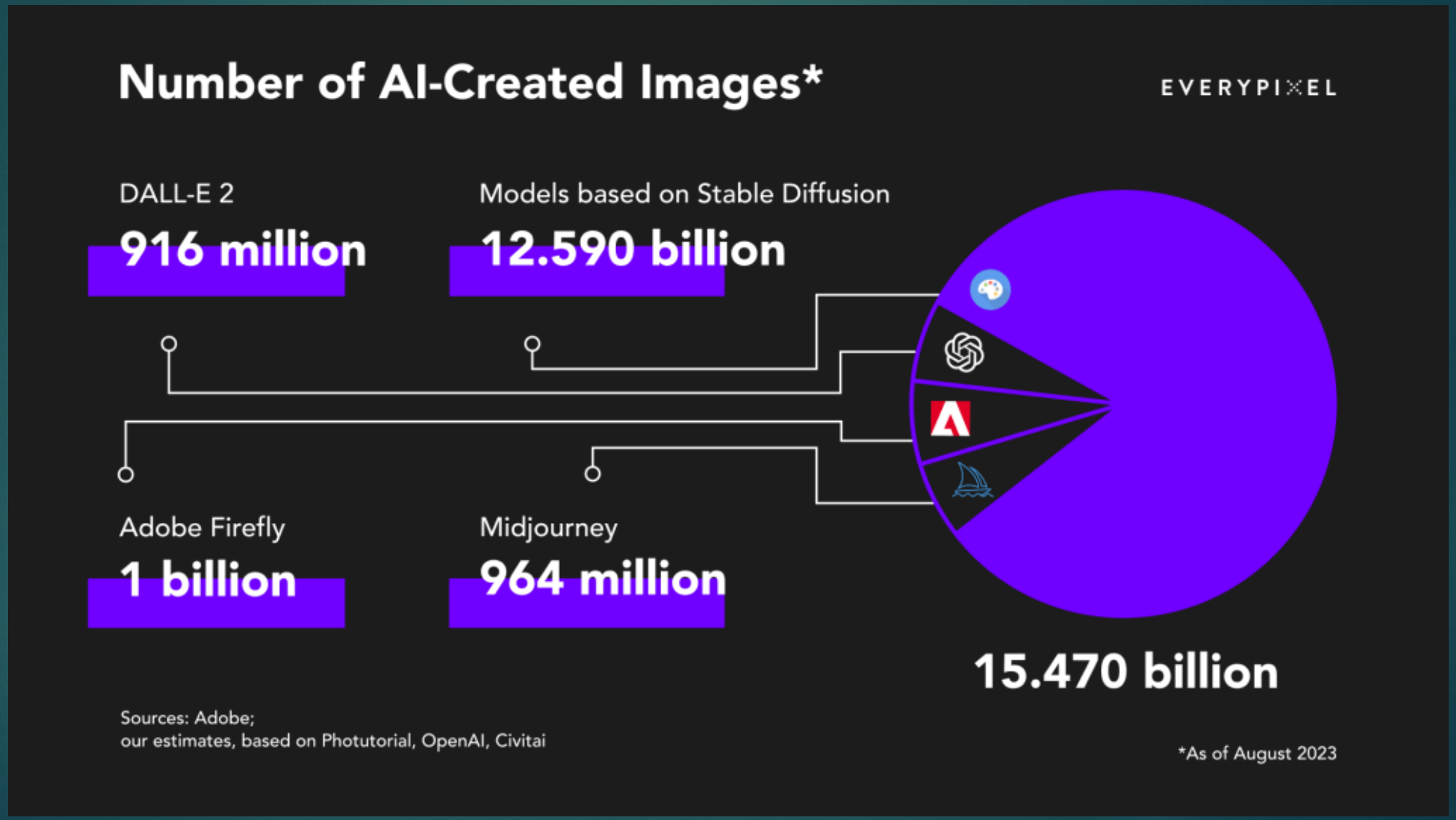
「AI(人工智能）：是祝福還是咒詛？」研討會（網上講座)

時間: 2023年12月9日 10:00 上午 香港特別行政區

人工智能藝術生成工具
AI Art Generator

- 人工智能藝術正在風靡世界，從推出到 2023 年 8 月，人工智能創作的圖像數量已經相當於攝影師過去一百五十年來拍攝的圖像數量總和。
- 著名的人工智能藝術系統包括了DALL-E3、Midjourney、Stable Diffusion、Adobe Firefly，後起之秀亦有不少，例如Leonardo AI、Tensor AI、Ideogram AI…… 等等。

人工智能藝術生成工具

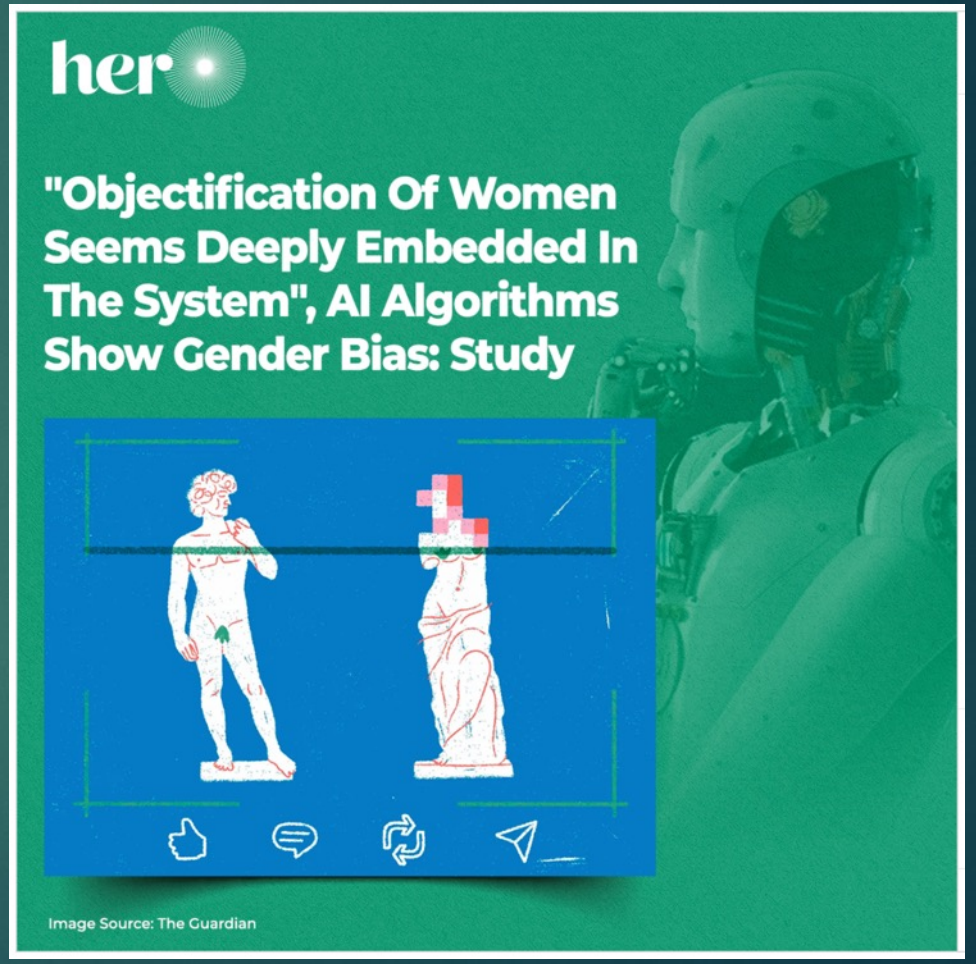

Number of AI-Created Images*
EVERYPIXEL
DALL-E 2
916 million
Models based on Stable Diffusion
12.590 billion
Adobe Firefly
1 billion
Midjourney
964 million
15.470 billion
Sources: Adobe;
our estimates, based on Photutorial, OpenAI, Civitai
*As of August 2023

人工智能壓制女性？

- 紐約大學教授 Hilke Schellmann 因其研究榮獲 2023 年普立茲獎。在她的研究中，她展示了社群媒體演算法的偏見如何壓制女性。
- 例如有幾百張女性穿着內衣進行日常活動（包括運動、進行醫學檢查」）的照片被人工智慧演算法標籤為具有性暗示。
- 微軟演算法認為一張顯示出孕婦腹部的照片具有性暗示的可能性高達90% 。

人工智能壓制女性？

- 又例如在 Instagram 上，一張穿著瑜伽褲並露出一點皮膚的女性的照片被「屏蔽」（shadowbanned），儘管該照片並未被刪除，但由於演算法將其定性為「 太火辣」（ too racy ），所以用戶難以搜尋出這圖片。
- 然而，赤裸上身的男人的圖像不會以同樣的方式受到評估。



我的回應

- 在大多數文化裏面，男人赤裸上身並沒有違反任何規範，但女性則不然，這並不算是雙重標準。
- 在很多文化中，公開穿着內衣做運動的照片亦不是常態。
- Instagam 、微軟和其他公司需要考慮到過於暴露的女性圖片可能會引來投訴，甚至法律訴訟，在 Hilke Schellmann 舉出的例子中，可能演算法可能過於敏感，出於自我保護，便會出現了有殺錯、無放過的現象，但這並不一定是出於歧視或者偏見。
- 反過來說，若果人工智能合成的女性圖像過於暴露，批評者又會說這是為了男性得到快感，將女性的身體商品化 。

人工智能壓制女性？

- 2022 年，《麻省理工學院技術評論》報道人工智慧的記者 Melissa Heikkilä 測試了一款名為 Lensa 的頭像生成應用程序，該應用程式使用 Stable Diffusion 將自拍照變成頭像。
- Heikkilä 得到的結果是一系列穿著暴露的卡通化身，這些化身看起來一點也不像她。

人工智能壓制女性？

- 相較之下，Heikkilä男同事的化身卻是衣冠楚楚，其形像是太空人、探險家、發明家。
- 而白人女同事的化身並沒有在她那般性感。
- 雙重歧視：性別、種族。

人工智能的倫理問題

- 屏蔽了可能具有性暗示的圖片是一個極端，但利用人工智能製造裸體或者性感圖像是另一個極端。
- 後者最極端的例子：DeepFake and DeepNude
- 「DeepFake」一詞始自 2017 年的 Reddit 論壇，用戶在該論壇上分享使用「深度學習 」（Deep learning ）創造或者修改過的色情影片。它是「深度學習」和「虛假」 兩個詞語的組合。
- Deeptrace Labs 2019 年發布的一份報告稱，在網路上發現的 15,000 個深度造假影片中，96% 是未經當事人同意而虛擬出來的色情內容，其中包含將受害人的臉孔接上裸露的身體。

Generative adversarial Network (GAN)

- DeepFake 建基於 GAN。
- 由曾經先後在Google和蘋果電腦從事人工智能研究的古德費洛（Ian Goodfellow）於 2014 年創建。
- Unsupervised learning
- GAN 創造出來的人面和真人無異 。

GAN

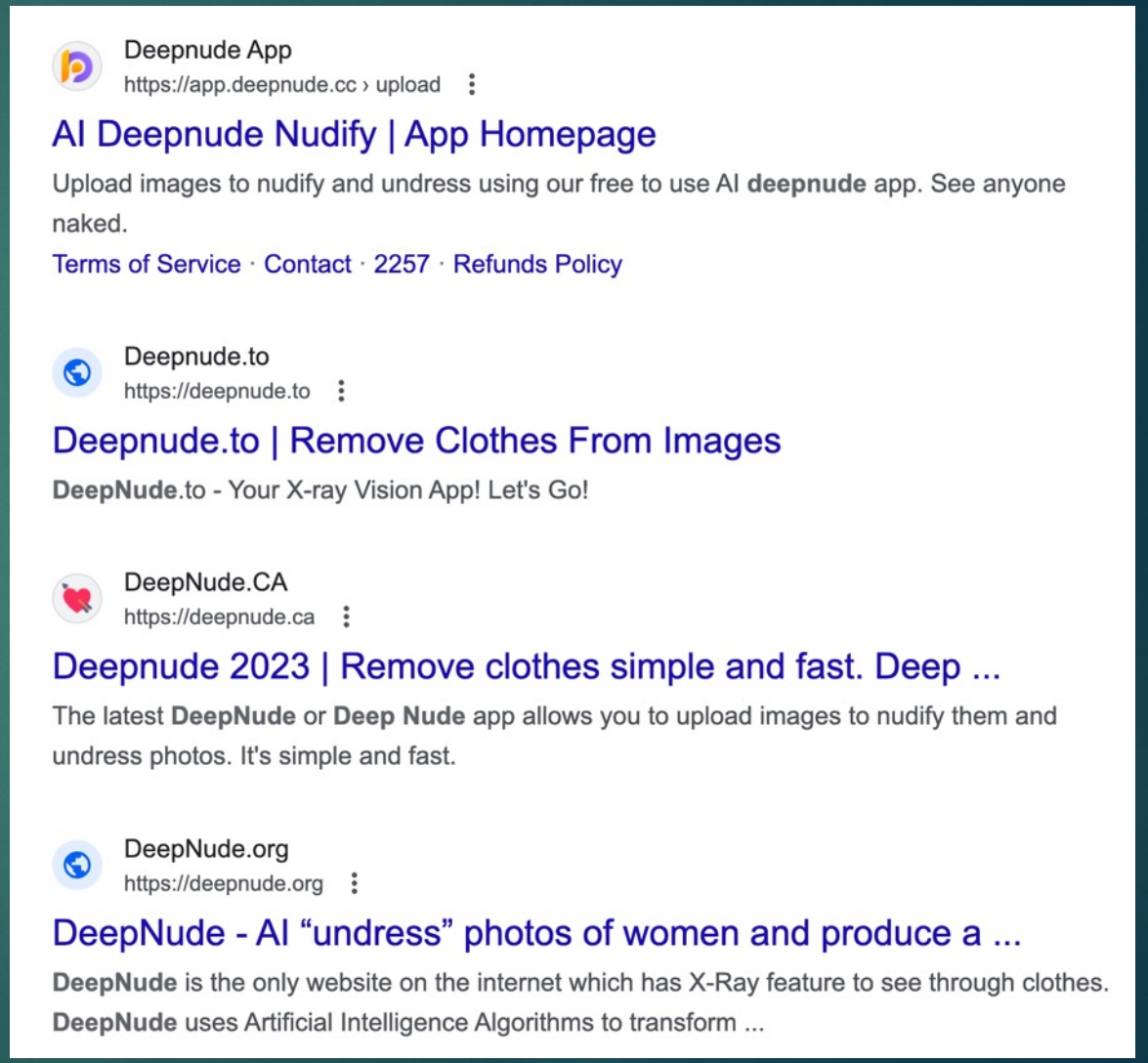
- Facebook 首席 AI 科學家楊立昆（ Yann LeCun ）盛讚：「這是過去 20 年來深度學習領域最酷的想法。 」
- 中國百度前首席科學家吳恩達（Andrew Ng） 說：GAN是「一項重大且根本性的進步」，激勵了全球不斷壯大的研究人員群體。
- 但這「最酷的想法」和「根本性的進步」卻變成了很多女性的惡夢。

DeepNude

- 2019 年 6 月，有人發布了一款名為 DeepNude 的電腦軟件，該應用程式通過 GAN，便可以轉瞬之間在圖像中將女性的衣服脫去。該應用程式有付費和非付費版本，付費版本售價是 50 美元。

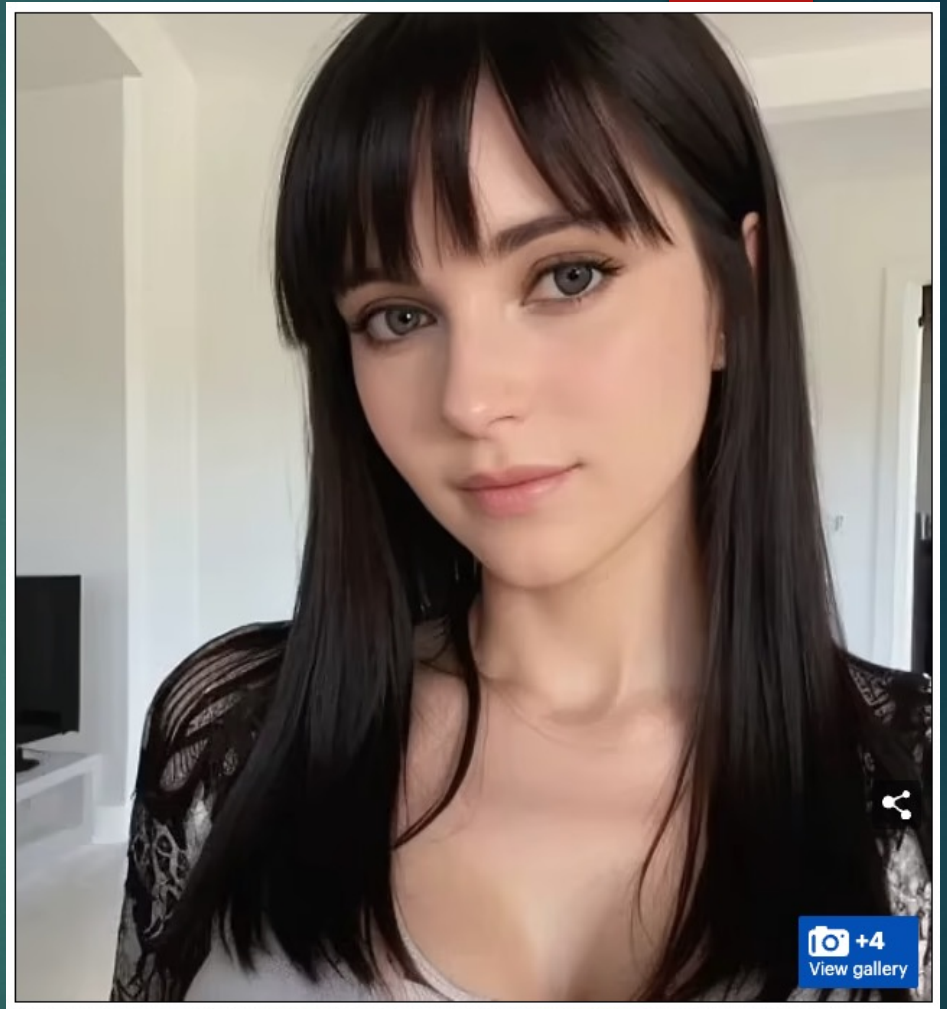
DeepNude

- 批評者認為該應用程式可能侵犯個人隱私，有時它被用於惡意攻擊和「報復性色情」（Revenge Porn），例如一對男女朋友分手之後，前男友利用DeepFake，將前女友的虛擬裸照在互聯網上散發。
- 面對輿論壓力，DeepNude 的創建者最終決定下架該應用程式，但 DeepNude 的其他變體仍然在互聯網上傳播。



人工智能的倫理問題

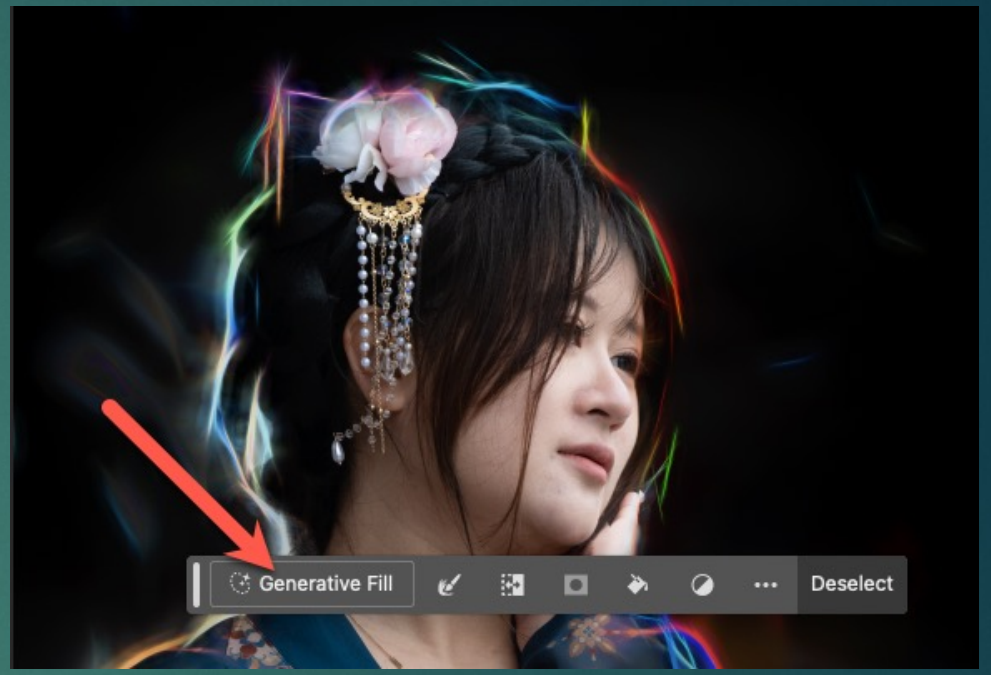
- 今年四月，一些 Reddit 用戶購買了人工智慧生成的裸體圖像， 那位虛擬的美女名為「克勞迪婭」（Claudia ），他們以為該圖像是真人。
- 這是一個實驗，兩名電腦系學生用人工智能創造了「克勞迪婭」， 他們想知道人們會否被人工智能欺騙，通過販賣裸照而得到$100之後，他們關閉了「克勞迪婭」 。
- 值得思考的是：若果人工智能虛構的色情照片能夠代替真人，這會不會減少了女性投身色情行業的機會呢？



One image of Claudia was a simple selfie captioned, 'Feeling pretty today.' The post had more than 200 comments, many praising her beauty. But one comment called out the image for being AI generated

人工智能生成的裸體圖像

- 人工智能應否製造裸體或者性感圖像呢？
- 為了避免潛在的濫用，DALL.E2/3 索性從數據庫中過濾掉了所有懷疑是色情的圖片，並且禁止可能製造出內容有問題的指令。
- 在Adobe，筆者曾經嘗試採用AI技術改造一個女性的面部圖像，但Adobe 誤會了那是身體的其他部份而拒絕執行指令。



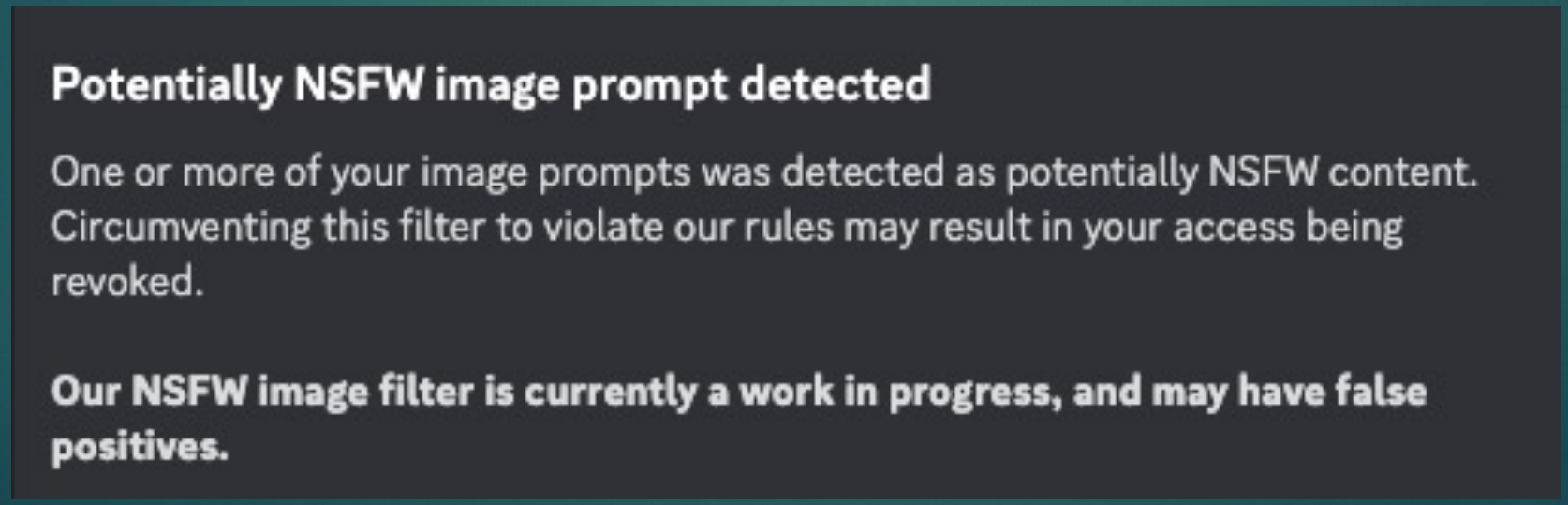
You need to fix some issues

- This prompt contains prohibited words: "Young girl dressed in steampunk style"

DALL-E 2 is subject to OpenAI's content policy. Names of public figures are not allowed.

人工智能生成的裸體圖像

- Midjourney也禁止用戶輸入一些敏感詞，例如sexy、nude、naked、bath， 若果輸入被懷疑是不當的詞語， Midjourney會發出這樣的警告：「你的一個圖像輸入被檢測為潛在的 NSFW 內容，規避此過濾器以違反我們的規則，可能會導致你的使用權限被撤銷。」NSFW是指「 對職場不安全」（Not safe for work ），通常是指裸體或者性感的圖像。



Stable Diffusion 對「照片」非常謹慎

- Stable Diffusion 沒有那麼嚴格，軟件只會將懷疑是NSFW的圖片弄成朦朧，或者隱蔽身體。
- 在大多數情況下，Stable Diffusion 對於呈現女性身體的照片非常謹慎。

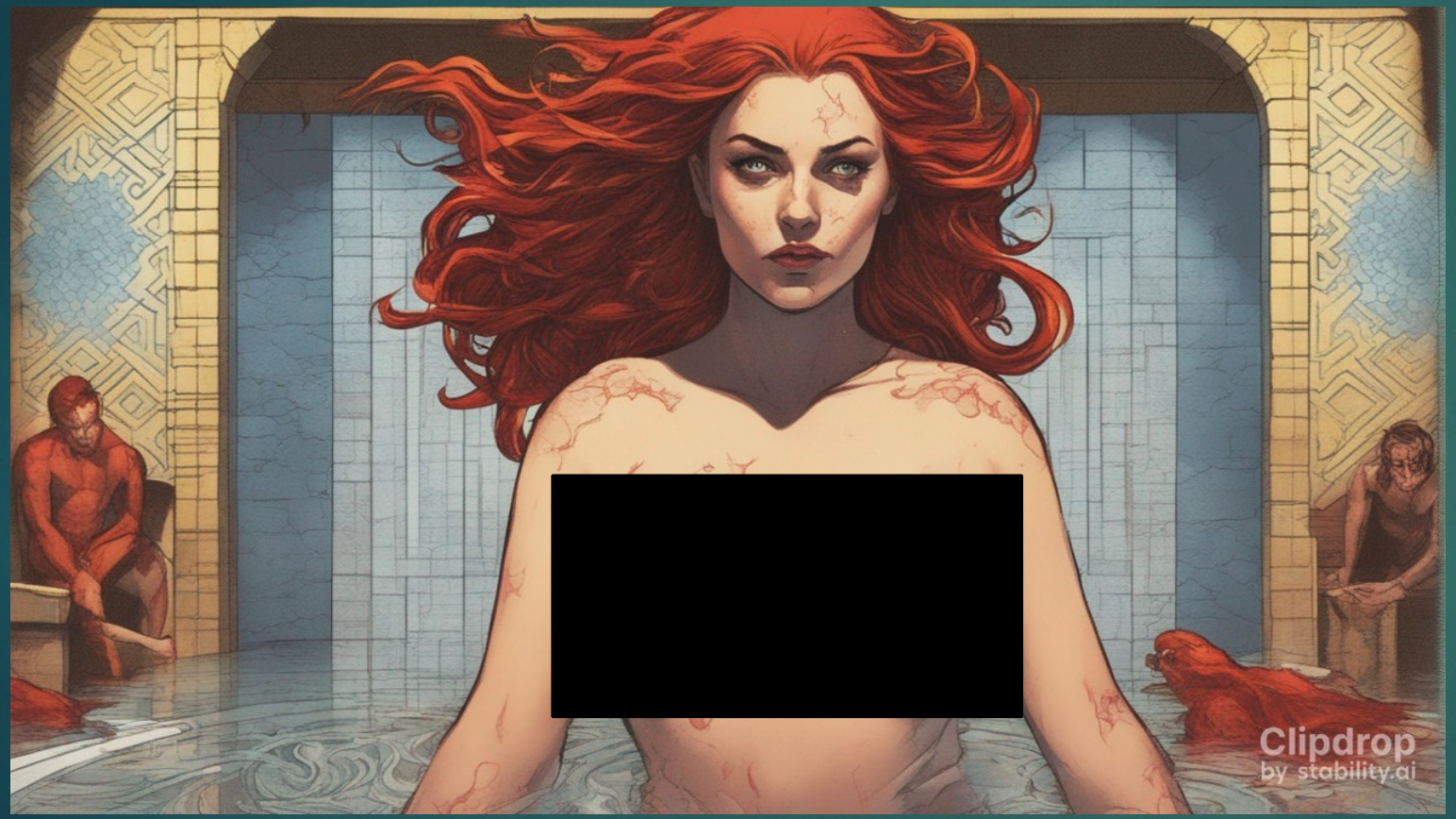
Stable Diffusion 對「照片」非常謹慎

例如在浴室裏面，女性模特兒用毛巾遮蔽整個身體。

Stable Diffusion 對「照片」非常謹慎

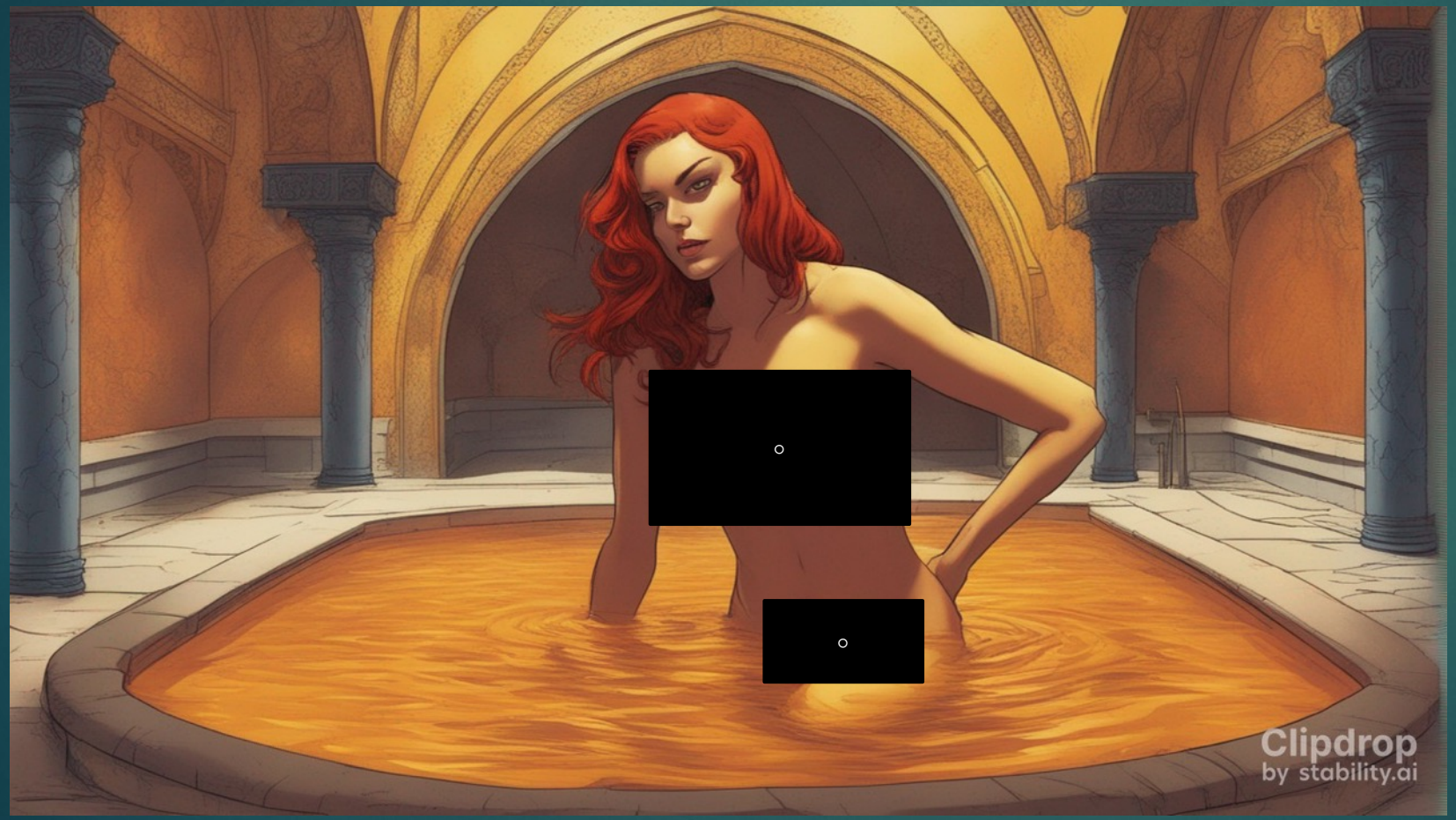
但有時候會有漏網之魚。

Stable Diffusion 不禁制裸體卡通



如果是繪畫形式的圖像，而不是好像攝影那般真實（photorealistic），Stable Diffusion 不會禁制。

Stable Diffusion 不禁制裸體卡通



這是基於以下的假設：並不接近真實的裸體圖像會有較低機會引起性幻想或者性衝動。

Stable Diffusion 不禁制裸體卡通

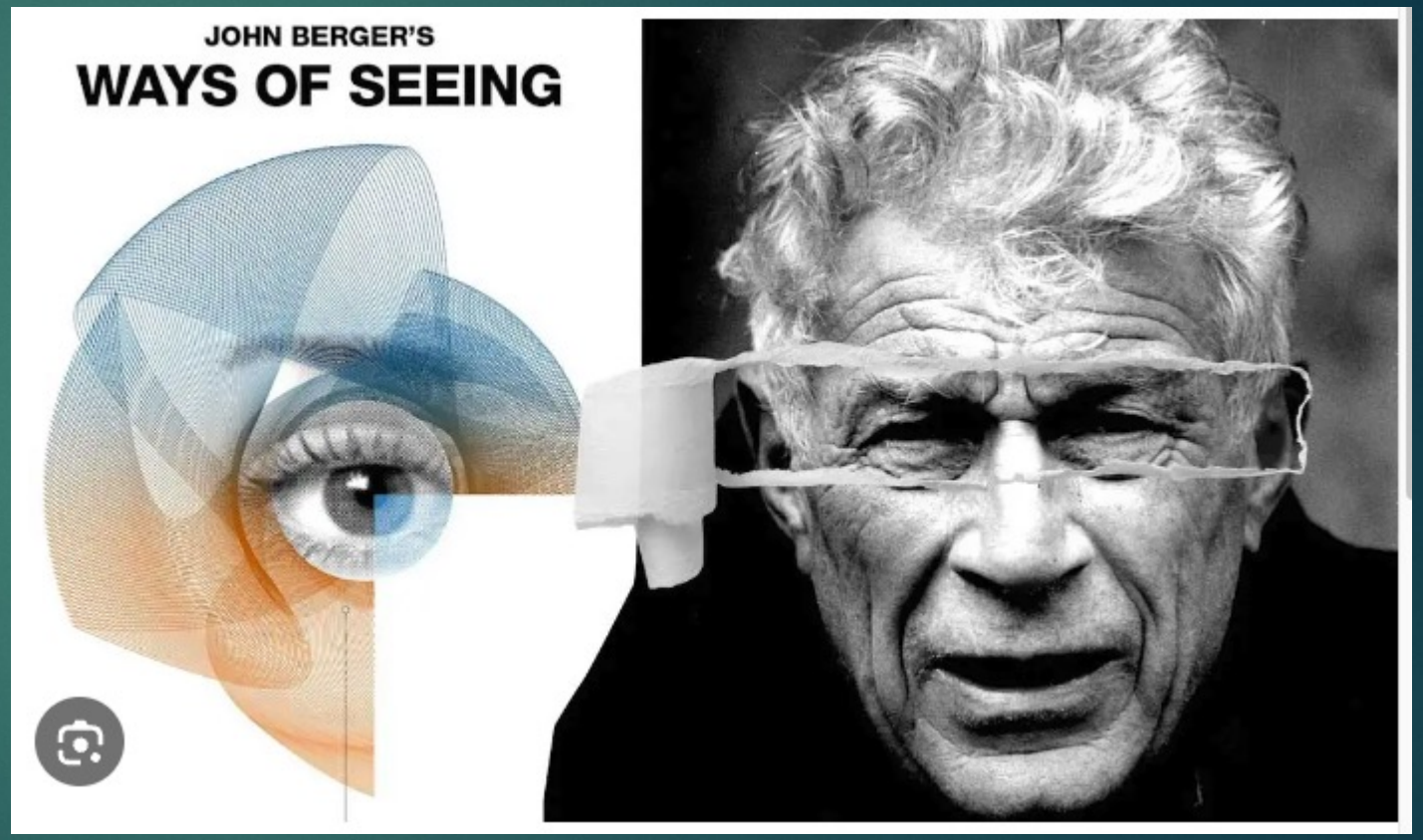
人對於視像的心理反應會隨着時代而改變，在未有照片和電影之前，中國已經有唐伯虎、趙子昂畫的春宮圖，但男性對春宮圖會有反應。

男性的凝視

- 筆者並不反對人體藝術，裸體圖像並不等同色情事實上，許多傳統的西方藝術都有裸體的形象，當中有些藝術品甚至是天主教會贊助的。
- 不過，一些人批評裸體或性感藝術將女性身體物質化和商品化，對女性的自尊和社會地位產生負面影響。
- 例如電影製片人勞拉·穆爾維（Laura Mulvey）在《視覺愉悅與敘事電影》中討論了「男性的凝視」（male gaze）的概念，她認為傳統電影經常將女性物化，將她們呈現為男性觀眾渴望和以「 金魚佬」 目光凝視的對象，這現象甚至延伸到藝術和廣告領域。

男性的凝視

- 藝術評論家約翰·伯格（John Berger）在其名著《不同的看法》中亦表達了類似的觀點，他認為，在藝術史裏面大部份作品都是為了滿足男性的凝視，女性是被動的觀察對象，而不是具有自主權的個體。



刻板形象？

- 不過，一些女性模特兒以自己的身材為驕傲，並且刻意在鏡頭前展示自己的身體，在自由社會裏面，每個人的選擇權利應該受到保障，那麼由其他人為她們設立道德標準又是否恰當呢？
- 一些女權主義者並不反對裸體或性感藝術，但她們批評許多傳統藝術和媒體傾向於表現理想化的女性體型，她們呼籲女性的身體形象應該更加多樣化，應該包括環肥燕瘦。
- 如果以這個標準來衡量，那麼無論是否裸體、性感、保守，大多數人工智能藝術軟件已經是犯了違反多元化和製造刻板形象的大忌，因為大多數輸出的女性形象都擁有天使面貌和魔鬼身材，而且樣貌多數是白種美女。

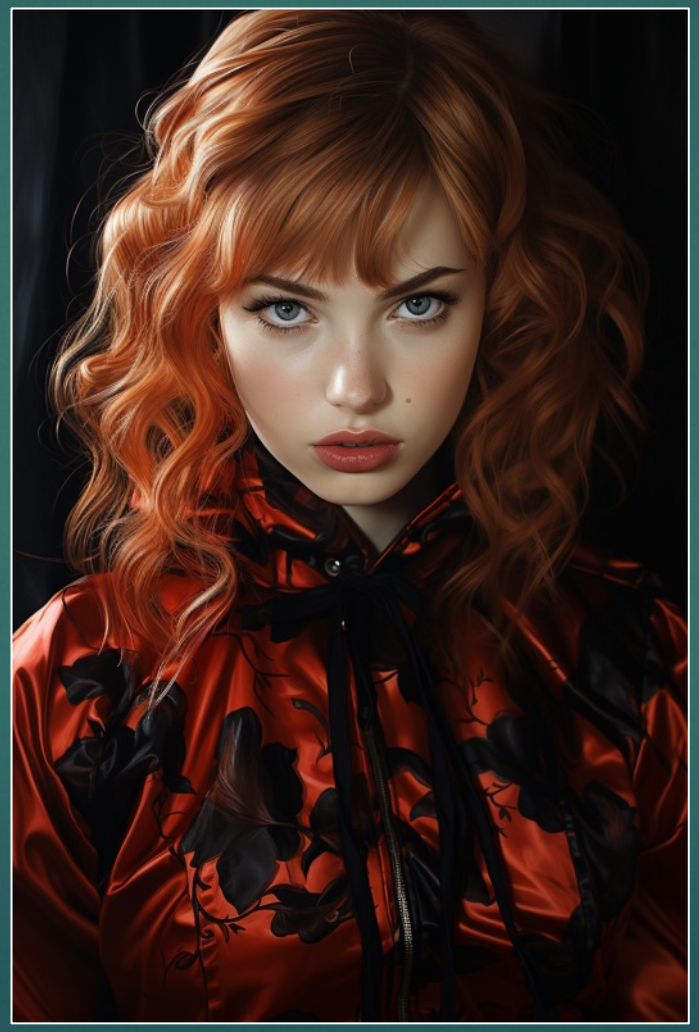
刻板形象？

- 人工智能採用了偏頗的樣本，基於缺乏代表性的樣本去學習世界的事物，人工智能便會生產出狹隘的刻板形象。一個著名的例子就是OpenAI早期的圖像生成系統DALL.E-2，當用戶要求系統製作出律師的圖像時，那麼全部顯示出來的人物都是中年白種男子。

刻板形象？

- 若果用戶要求空中小姐的照片，那麼輸出的結果都是年輕貌美的亞洲女子。

人工智能創造的美女多數是白種人



需要特別指明非白種人

如果我特別指明：「一名穿着和服的日本女子在日式建築」，那麼人工智能生成的圖像會準確顯示我輸入的內容。

參考熟悉的框架

- 這未必是歧視和偏見。
- 大多數創作者都會參考自己熟悉的資料和自身的經驗，而且，為了迎合本地市場，他們必須要提供顧客喜歡的東西。
- 美國人設計的人工智能藝術系統，當然會傾向於展示美國女性的圖像，問題是：美國女性有不同的年齡階層， 有高矮肥瘦，為什麼在沒有特別提示下，人工智能只會輸出美女？套用勞拉·穆爾維和約翰·伯格的理論，這是為了滿足「男性的凝視」。

滿足市場需要？

- 許多年前筆者主修傳播理論，教授曾經提出一個案例在堂上討論：克里斯汀・克拉夫特（Christine Craft）是一家電視台的 節目主持人，但 1981 年 8 月一個市場研究小組認定她太老了和太沒有吸引力，因此她被免去了主播職位，於是乎克拉夫特向電視台提出法律訴訟。
- 教授指出一個殘酷的現實：觀眾期望在電視螢光幕上看見一個年輕貌美的女直播。現在美國電視網絡主播的樣貌已經變成了高度多元化。

類似事件經常發生

- 年老色衰的女性新聞從業員被解僱

1999: Janet Peckinpaugh was awarded $3.79 million after suing a network for a demotion.

screencap

2002: Carol Kaplan filed a gender discrimination lawsuit against WGRZ-TV Buffalo.

Facebook fan page screencap

2002: Susan Hutchison filed a lawsuit against KIRO-TV in Seattle for age discrimination.

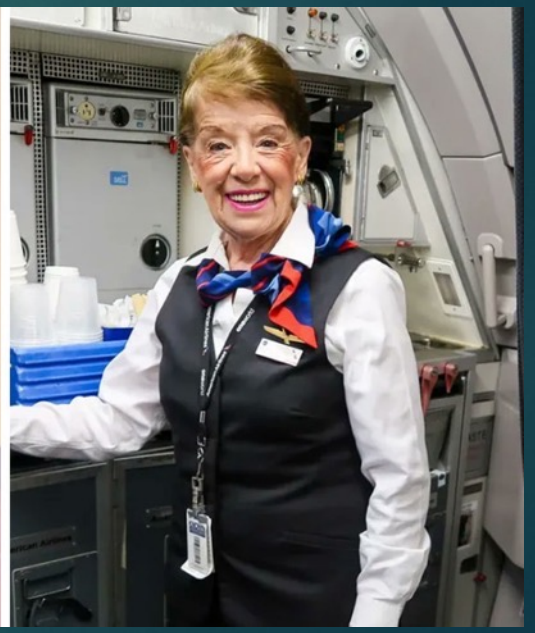
screencap

迎合市場需要？

- 過去美國航空公司視外表為招聘空姐最重要的因素，航空公司利用女性的性感形象去吸引乘客，因此，空中小姐的製服通常很貼身，而且她們需要穿戴高跟鞋；
- 空中小姐在這些情況下會被解僱：年齡達到三十多歲、體重不夠嬌小玲瓏、準備結婚……。

迎合市場需要？

- 二十世紀60年代美國人的平權意識逐漸高漲，這些外表要求逐漸被廢除，現在美國國際航空公司的空中服務員已經很多樣化。不過，亞洲航空公司仍然保留了聘請年輕貌美的女性為空服員這傳統。
- 不要忘記，是我和你這些顧客製造了市場的需要。
- DALLE2 只是反映現實？



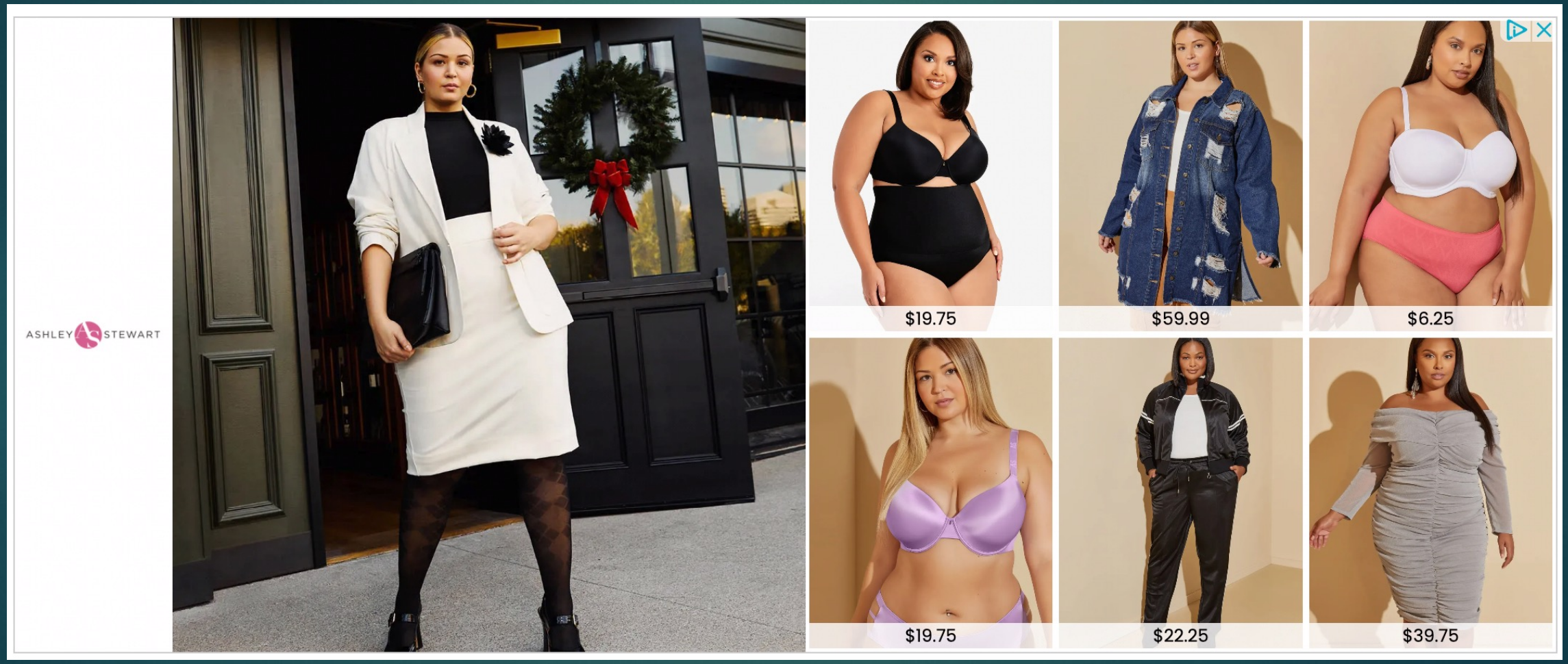
藝術多元化反映社會多元化

- 由真人藝術家創作的藝術已經克服了以上的問題，亦即是由過去的刻板形象、單一化進入到多元化。
- 例如哥倫比亞藝術家費爾南多·博特羅（Fernando Botero）畫筆下的人物都是身形粗壯。

廣告多元化反映社會多元化

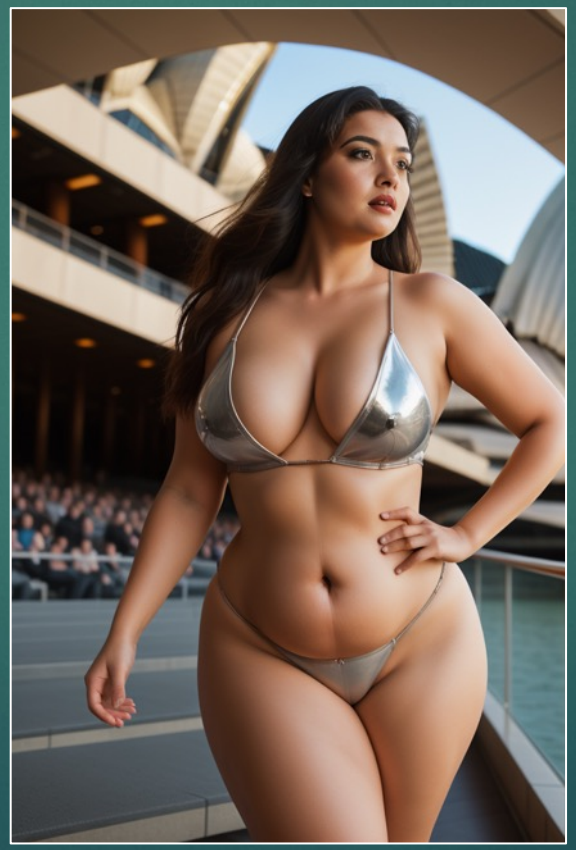
- 專門出售女性內衣的「維多利亞秘密」亦擺脫了以曲線玲瓏的女性圖像為廣告，而是展示非常豐滿的身形。
- Target 百貨公司的女裝部亦採用了同樣的廣告形式。

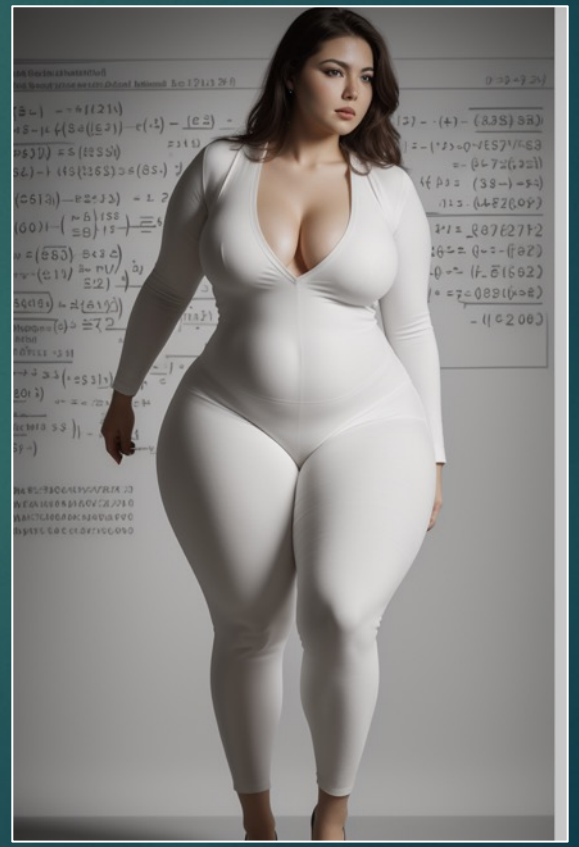
廣告展示非刻板的女性形象



Tensor Art 打破了刻板形象

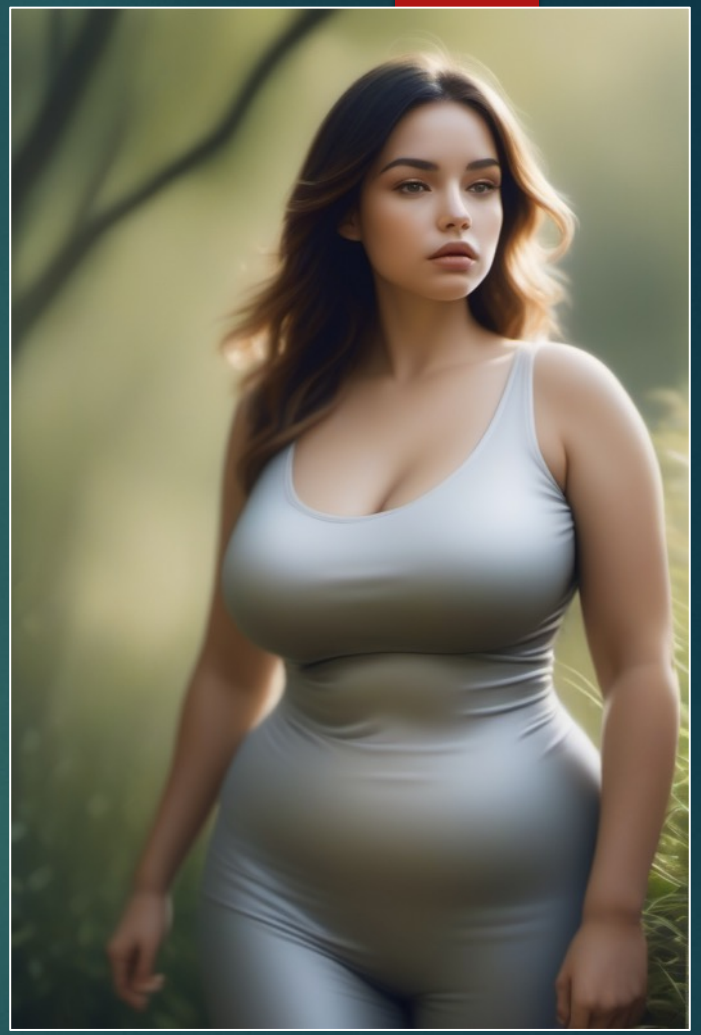
Tensor Art 衝破了女性一定要身材苗條的刻板形象，即使我在輸入指令中沒有說明高矮肥瘦，有時候創造出來的女性圖像擁有標準身材（左一），但有時候略肥（中間），有時候甚至過胖（右一）。





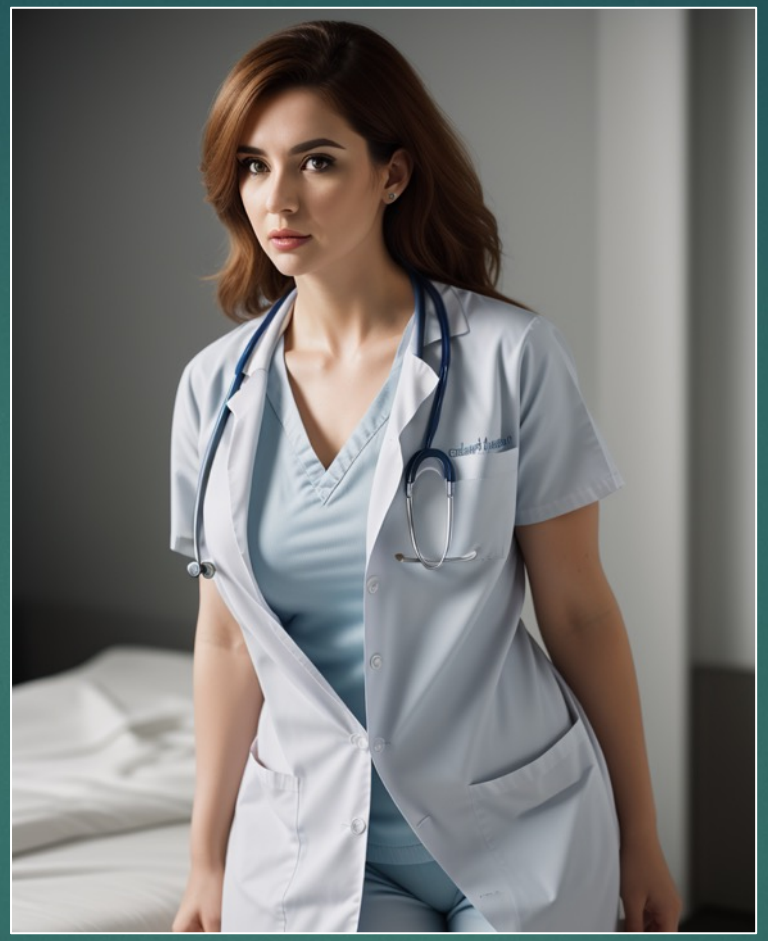
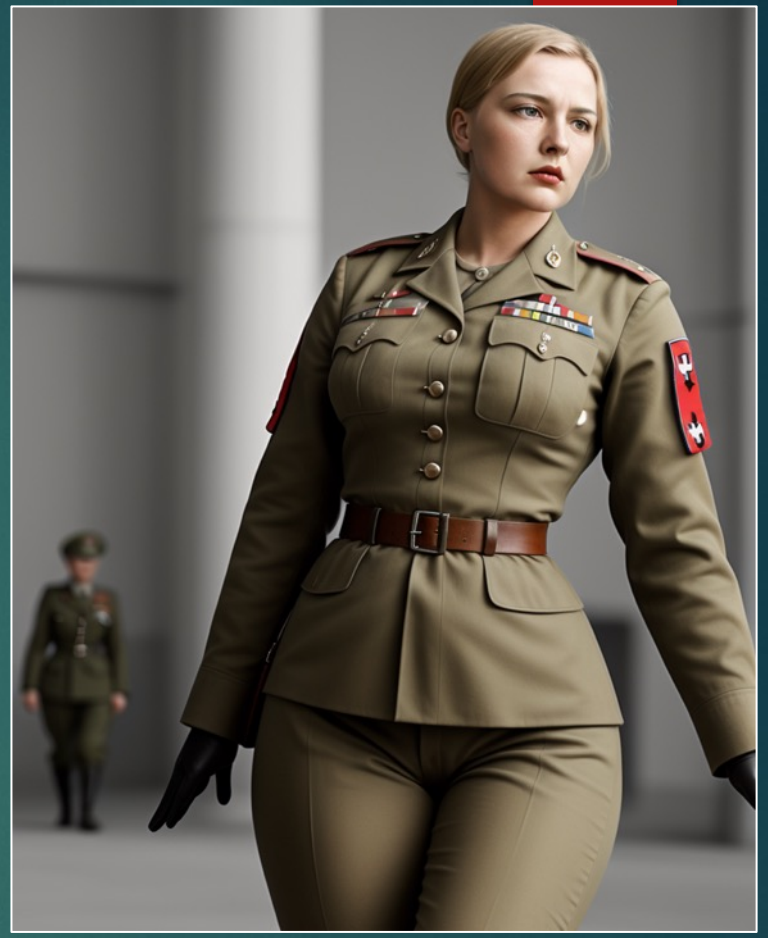
- 左面和右面的圖片都是採用同一指令製造出來的，但Tensor Art人工智能會創造出不同的身材。

更多例子



男性凝視的框框

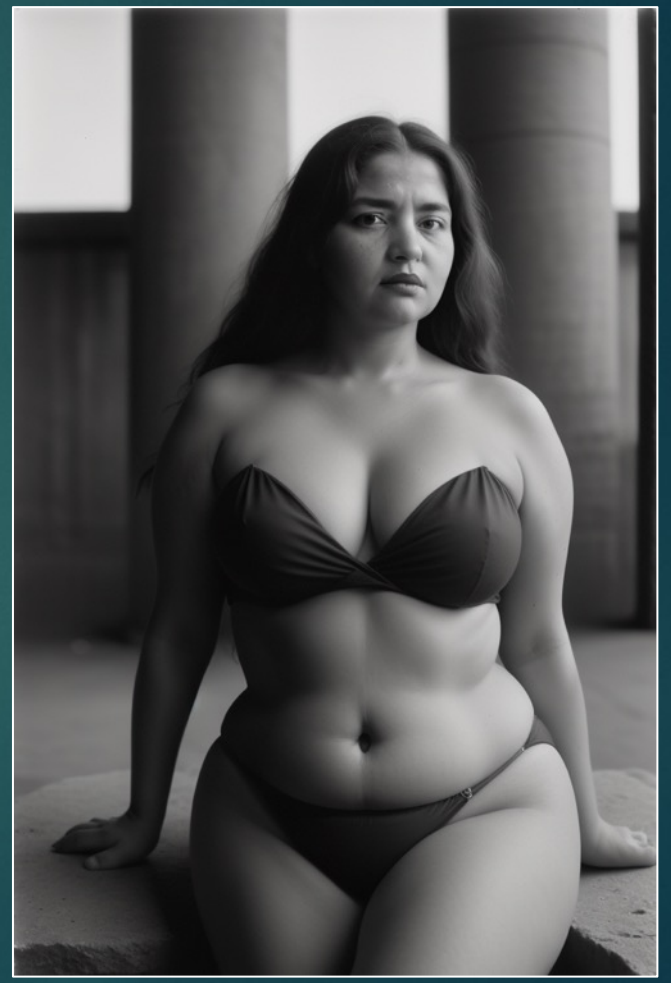
但 Tebsor Art 仍然無法衝出男性凝視的框框，如果沒有清楚要求圖像中的女性穿什麼衣服，絕大部份圖片中的女性都是穿泳衣或者緊身的性感衣服。

除非我需要清楚指明圖片中女性的角色，例如修女、護士、軍人，或者要穿什麼衣服，否則會呈現性感的衣着。

突破了樣貌的刻板形象

- 在以上的圖片中，雖然有些女性的身型粗壯，但樣子仍然是很甜美。
- 在很少數的情況下 ，即是我沒有說明對臉孔的要求，Tensor Art 製成的女子樣貌平庸。



總結

筆者討論了關於人工智能怎樣去呈現女性形象的倫理問題：

- 人工智能是否含有壓制女性的演算法？到底屏蔽了可能具有性暗示的圖片是貶損女性，抑或製造裸女或者性感女人圖像是貶損？
- 應否禁止裸體或性感的圖像，從而避免貶低女性、滿足男性凝視的結果？
- 人工智能是製作應否盡量多元化，而不是局限於製造身材苗條和面貌娟好的白種人美女？

禁止DeepNude 應該沒有爭議餘地，但其他問題並不能簡單地一錘定音，或者我們不一定要達到共識，而是應該用尊重多元化的態度去面對這些問題。