

## 第21章CHAPTER 21所有人工智慧都是有偏見的All AI Is Biased

當我們檢查生成式人工智慧模型中普遍存在的偏見時，令人震驚的是，其中一些偏見的根深蒂固。然而，問題不僅持續存在，而且還在惡化。這就是為什麼使用者體驗專業人士需要立即採取行動來提高對這個問題的認識。在本章中，我提供了評估人工智慧偏見的實用方法，並提出更好的問題，以幫助提供更平衡和多樣化的人工智慧回應。

When we examine the pervasive bias in generative AI models, it's alarming how deeply ingrained some of these biases are. And yet, the problem not only persists but is getting worse. This is why UX professionals need to act now to raise awareness of the problem. In this chapter, I provide practical approaches to evaluating AI bias and asking better questions to help deliver more balanced and diverse AI responses.

### 當你問「生物學家」時，你會期待什麼？

#### What Do You Expect When You Ask for “Biologist”?

您不必看太遠就能發現人工智慧系統中的偏見。例如，如果您對「生物學家」執行 Midjourney /imagine 查詢，您將獲得絕大多數白人男性。你還會得到一個孤獨的女性形象（圖 21.1 中中間的底行）。在這次特定的試運行中，這一比例為 25 比 1，即 4%。您的里程可能會有所不同。

You don't have to look far to find bias in AI systems. For example, if you run a Midjourney /imagine query for “biologist,” you will get a vast majority of white males. You also get a solitary female figure (bottom row in the middle, in Figure 21.1). In this particular trial run, that's a ratio of 25 to 1, or 4 percent. Your mileage may differ.

事實上，如果你繼續運行查詢一段時間，你可能會得到，咳咳，其他生物體……我指的是青蛙（見圖 21.2）。事實上，在撰寫本文時，從統計學上講，您找到青蛙生物學家的可能性與女性生物學家的可能性一樣大。（說我在這裡有偏見，但即使是青蛙看起來也是雄性的。雖然我聽說青蛙很難說，如果電影《侏羅紀公園》可信的話，青蛙也可以隨心所欲地改變性別。

In fact, if you keep running the query for a while, you might get, ahem, other organisms ... by which I mean a frog (see Figure 21.2). In fact, as of this writing, statistically speaking, you are just as likely

to get a frog biologist as you are a woman biologist. (Call me biased here, but even the frog looks male. Although with frogs I hear it 's hard to tell, and if the movie Jurassic Park is to be believed, frogs can also change their sex on a whim.)

然而，根據美國勞工統計數據，女性生物學家比男性生物學家多 8%（男性 46%，女性 54%）（1）。

However, according to U.S. labor statistics, there are 8 percent more female biologists than male biologists (46 percent male, 54 percent female) (1).

「籃球運動員」怎麼樣？

How About “ Basketball Player ” ?

當我們運行相同的 Midjourney /imagine 查詢但要求“籃球運動員”時，我們也會得到一名單身女性籃球運動員（右邊第三行），但現在大多數男性是黑人（見圖 21.3）。

When we run the same Midjourney /imagine query but instead ask for “ Basketball Player, ” we also get a single woman basketball player (third row on the right), but now the majority of males are Black (see Figure 21.3).



圖 21.1 中途查詢「生物學家」會產生絕大多數白人男性

Figure 21.1 Midjourney query for “ biologist ” yields a vast majority of white males

來源：Midjourney 中用 AI 生成

Source: Generated with AI in Midjourney



圖 21.2 對於“生物學家”的查詢，從統計上看，您找到青蛙生物學家的可能性與您是女性的可能性一樣

Figure 21.2 For a query “biologist,” you are statistically as likely to get a frog biologist as you are a woman

來源：Midjourney 中用 AI 生成

Source: Generated with AI in Midjourney



圖 21.3 中途查詢「籃球運動員」產生絕大多數黑人男性

Figure 21.3 Midjourney query for “ basketball player ” yields a vast majority of Black males

來源：Midjourney 中用 AI 生成

Source: Generated with AI in Midjourney

### 第三次是魅力：《抑鬱的人》

### Third Time ’ s the Charm: “ Depressed Person ”

當我們搜尋「憂鬱症患者」時，事情發生了戲劇性的轉變。現在，我們主要是白人女性（見圖 21.4）。

Things flip around dramatically when we do a search for “ depressed person. ” Now, we mostly get white women (see Figure 21.4).

便條

## NOTE

也許女性之所以抑鬱，是因為她們不能成為生物學家或籃球運動員？真是禿頭。為什麼我們一開始總是如此執著於男性和女性？其他 72 種公認的性別認同呢？(2, 3).

Maybe the reason why the women are depressed is because they cannot be biologists or basketball players? What total balderdash. And why are we always so fixated on males and females in the first place? What about the 72 other recognized gender identities? (2, 3).

請注意：

Please note:

- 重點並不是說您的數據中應該始終包含所有 72 種性別認同。The point is not to say that you should always have all 72 gender identities in your data.
- 這並不是說您應該始終擁有 50/50 的男性/女性代表。It is not to say that you should always have a 50/50 male/female representation.
- 這並不是說你總是需要顯示實際勞動力資源批准的男性和女性統計關係（對於最高管理層的高層職位來說，這已經相當偏向，而且那是在 Project 2025 和現任政府的“DEI 戰爭”之前）。It is not to say that you always need to show the actual labor resources ' approved statistical relationship of males and females (which are already pretty skewed for top positions in the C-suite, and that was before Project 2025 and the current administration ' s “ War on DEI ” ).



圖 21.4 中途查詢「憂鬱症患者」產生絕大多數白人女性

Figure 21.4 Midjourney query for “ depressed person ” yields a vast majority of white females

來源：Midjourney 中用 AI 生成

Source: Generated with AI in Midjourney

重點很簡單：

The point is simply this:

## 便條

### NOTE

始終假設所有人工智慧都有偏見。並弄清楚這種偏見將如何影響體驗。

Always assume that all AI is biased. And figure out how that bias will impact the experience.

因此，當您看到您的 AI 生成特定結果時，如圖 21.5 所示.....

So, when you see your AI generating a specific result, as in Figure 21.5 ...



圖 21.5 查詢的典型輸出：「生物學家」、「籃球運動員」、「憂鬱症患者」

Figure 21.5 Typical output for queries: “ biologist, ” “ basketball player, ” “ depressed person ”

來源：Midjourney 中用 AI 生成

Source: Generated with AI in Midjourney

...訓練自己首先尋找不存在的東西並加強缺失的區域（見圖 21.6）。

... train yourself to look first for what is not there and reinforce that missing area (see Figure 21.6).

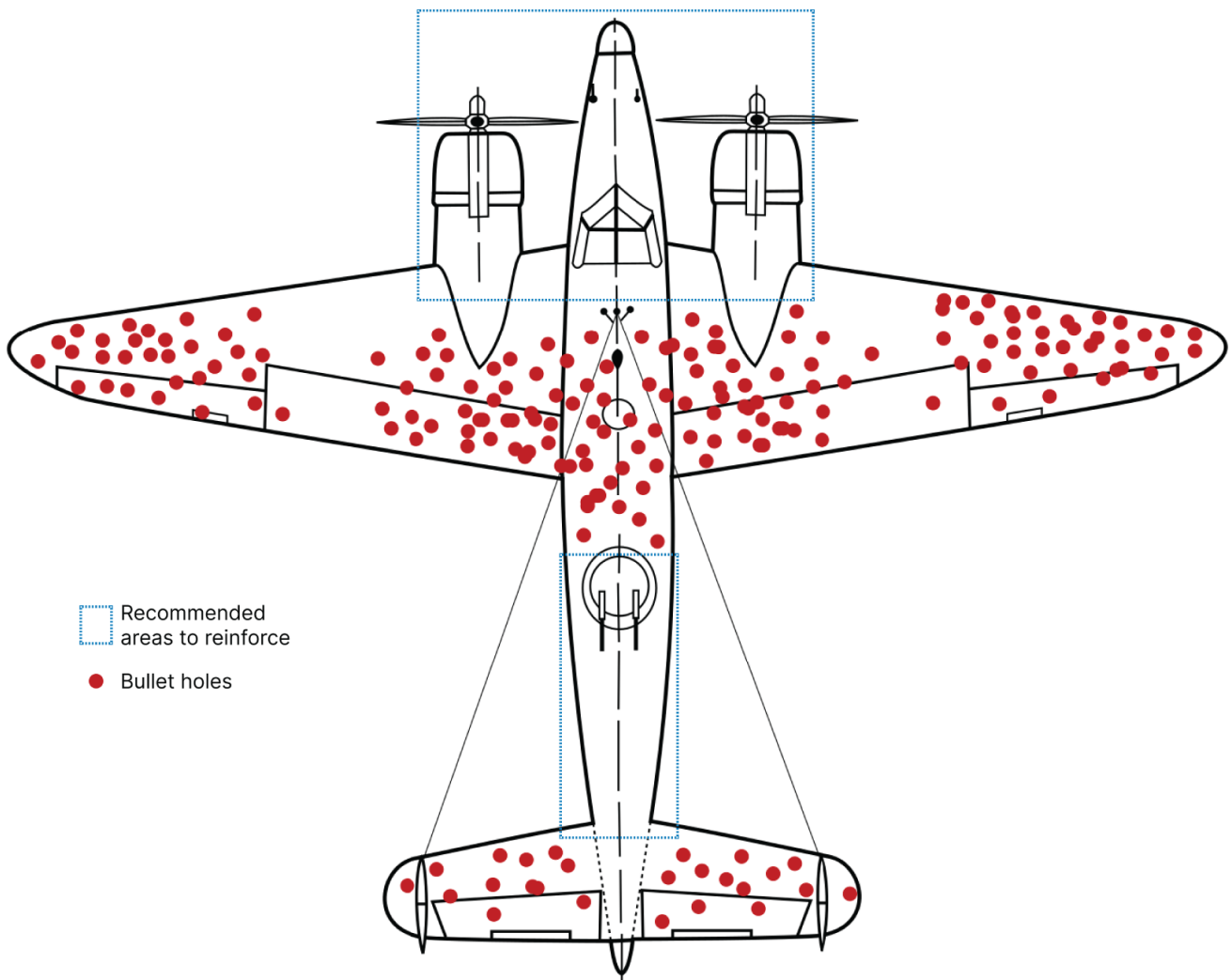


圖 21.6 倖存者偏差範例

Figure 21.6 Example of survivor bias

來源：'作者' / 經 LinkedIn 許可轉載 / <https://www.linkedin.com/pulse/survival-bias-business-unseen-peril-hindering-success-javier-sada>

Source: 'author' / reproduced with permission of LinkedIn / <https://www.linkedin.com/pulse/survival-bias-business-unseen-peril-hindering-success-javier-sada>

我看得出來你很困惑。為什麼我們突然要看這張飛機示意圖？

I can see you are confused. Why are we looking at this airplane schematic all of a sudden?

讓我解釋一下：二戰初期，工程師們想知道如何提高戰機的生存能力。一種方法是在飛機上穿上更厚的裝甲，但當然，由於重量限制，只能加固飛行器的某些關鍵區域，而不是整個機身。工程師們討論了這個問題，並提出也許我們應該從加固彈孔最多的區域開始。於是，他們開始觀察回來的飛機，仔細測量彈孔密度，並編制詳細的統計數據。

Let me explain: At the beginning of World War II, engineers wanted to know how to improve the survivability of war airplanes. One way to do so was to put thicker armor on the airplane, but of course, due to weight restrictions, only certain critical areas of the flying machine could be reinforced, not the entire airplane body. The engineers discussed the problem and proposed that maybe we should start by reinforcing the areas with the most bullet holes. So, they started looking at the airplanes that came back, carefully measuring the bullet hole density and compiling detailed statistics.

然後，一位非常聰明的人——一位統計學家和早期的人為因素先驅亞伯拉罕·沃爾德——提出了一個問題，為什麼發動機、駕駛艙和機尾上沒有任何彈孔？難道是因為在那些關鍵區域被擊中的飛機甚至從未回來接受檢查？這是「倖存者偏見」的一個例子：專注於通過選拔過程的實體而忽略那些未通過的實體的邏輯錯誤。

And then, a very smart person—a statistician and early human factors pioneer Abraham Wald—asked the question, Why aren't there any bullet holes in the engines, the pilot cabin, and the tail? Could it be that it's because the airplanes that got hit in those critical areas never even came back to be examined? This is an example of “survivor bias”: the logical error of concentrating on entities that passed a selection process while overlooking those that did not. (4)

學會以同樣程度的懷疑態度對待您的生成式 AI 結果。

Learn to treat your generative AI results with the same level of skepticism.

## 便條

## NOTE

訓練自己首先尋找缺失的東西，並通過提出更好的問題來強化這些領域。

Train yourself to look first and foremost for what is missing and reinforce those areas by asking better questions.

您沒有理由不能使用完全相同的 Midjourney

工具，稍微調整查詢，並在生成的數據中引入缺失的多樣性，如圖 21.7

所示。所需要的只是一點點的意識和關懷——我祈禱我們都能找到更多的一點點。每一天。

There is no reason why you can't use the exact same Midjourney tool, tweak the query slightly, and introduce missing diversity into your generated data, as shown in Figure 21.7. All it takes is just a tiny bit of awareness and care—something I pray we all find just a tiny bit more of. Every single day.



圖 21.7 查詢的典型輸出：「黑人跨性別生物學家」、「印度女籃球運動員」、「憂鬱的老年亞洲男性」

Figure 21.7 Typical output for queries: “ black trans biologist, ” “ Indian woman basketball player, ” “ depressed older Asian man ”

來源：Midjourney 中用 AI 生成

Source: Generated with AI in Midjourney

正如喬伊·布奧蘭維尼雄辯地說的那樣，

As Joy Buolamwini so eloquently said,

“ 人工智能是否會幫助我們實現我們的願望，還是加劇不公正的不平等，最終取決於我們 ” (5)。

“ Whether AI will help us reach our aspirations, or reinforce unjust inequalities, is ultimately up to us ” (5).

## 參考

## References

1. 1. 美國生物學家人口統計和統計 (2024

年)。齊皮亞，[www.zippia.com/biologist-jobs/demographics1](http://www.zippia.com/biologist-jobs/demographics1)。 Biologist demographics and statistics in the US (2024). Zippia, [www.zippia.com/biologist-jobs/demographics](http://www.zippia.com/biologist-jobs/demographics)

2. 2. 其他 72 種性別是什麼？（2024）。

醫學網。 [www.medicinenet.com/what\\_are\\_the\\_72\\_other\\_genders/article.htm2](http://www.medicinenet.com/what_are_the_72_other_genders/article.htm2). What are the 72 other genders? (2024). MedicineNet.

[www.medicinenet.com/what\\_are\\_the\\_72\\_other\\_genders/article.htm](http://www.medicinenet.com/what_are_the_72_other_genders/article.htm)

3. 3. 性別認同清單（2024 年）。維基百科，維基媒體基金會。 [https://en.wikipedia.org/wiki/List\\_of\\_gender\\_identities3](https://en.wikipedia.org/wiki/List_of_gender_identities3). List of gender identities (2024). Wikipedia, Wikimedia

Foundation. [https://en.wikipedia.org/wiki/List\\_of\\_gender\\_identities](https://en.wikipedia.org/wiki/List_of_gender_identities)

4. 4. 倖存者偏差（2024）。維基百科，維基媒體基金會。 [https://en.wikipedia.org/wiki/Survivorship\\_bias4](https://en.wikipedia.org/wiki/Survivorship_bias4). Survivorship bias (2024). Wikipedia, Wikimedia Foundation.

[https://en.wikipedia.org/wiki/Survivorship\\_bias](https://en.wikipedia.org/wiki/Survivorship_bias)

5. 5. 人工智慧圖像生成器如何使偏見變得更糟。倫敦社會研究所， [www.lis.ac.uk/stories/how-ai-image-generators-make-bias-worse5](http://www.lis.ac.uk/stories/how-ai-image-generators-make-bias-worse5). How AI image generators make bias worse.

London Institute for Social Studies,

[www.lis.ac.uk/stories/how-ai-image-generators-make-bias-worse](http://www.lis.ac.uk/stories/how-ai-image-generators-make-bias-worse)