

# 黃仁勳、小扎巔峰對談：萬字長文揭秘 Meta 的 AI 未來

之二



發佈內容這樣的人，就能夠擁有智能體來完成。同樣地，我認爲人們將會創造自己的智能體，用於各種不同的用途。有些會是定制化的實用工具，他們希望對其進行精細調整和訓練。有些則會是娛樂性質的。人們創造的一些東西只是有趣，你知道的，它有些可能有點傻或者對事物有一種幽默的態度，我們可能無法預料。

用戶可能不會將 MetaAI 構建爲助手，但我認爲人們對看到和與之互動相當感興趣。然後我們看到的其中一個有趣的應用場景是人們使用這些智能體進行支持。這一點是讓我有點驚訝的，MetaAI 最受歡迎的用途之一是人們基本上用它來模擬他們即將面臨的複雜社交情境。無論是職業場合，比如說，"我想問我的經理，我怎樣才能得到晉陞或加薪"，或者"我和朋友發生了爭執"，或者"我和女朋友遇到了困難"，基本上在一個完全沒有偏見的空間里，你可以模擬這種情況並看看會如何發展，如何進行對話並獲得反饋。

很多人，他們不僅僅想與同類型的智能體交互，無論是 MetaAI 還是 ChatGPT，或者其他大家都在使用的工具，他們想要創造自己的東西。所以這就是我們大致上在 AI 工作室所要做的方向。但這些都是我們更大的觀點的一部分，我們認爲不應該只有一種大型 AI 供人們交互。我們只是認爲，如果存在多種不同的 AI，世界會變得更好、更有趣。

**黃仁勳：**我認爲這非常酷，如果你是一位藝術家並且有自己的風格，你可以採用你的風格，你所有的作品集，你可以微調你的一個模型，現在這變成了一個 AI 模型，你可以來使用它並給出提示。你可以讓我，比如說，創作一些與我所擁有的藝術風格類似的作品，你甚至可以給我一幅作爲繪圖的藝術品，一個草圖作爲靈感，我就能爲你生成一些東西。然後，其他用戶會來到我的 AI 那里(獲得這些能力)。每一個餐廳，每一個網站在未來可能都會有這些 AI。

**扎克伯格：**是的，我有點認爲在未來，就像每個企業都有一個電子郵件地址、一個網站和一個或多個社交媒體賬戶，我認爲在未來，每個企業都將擁有一個與客戶交互的 AI 智能體。

**黃仁勳：**對。  
**扎克伯格：**這些事情我認爲是歷史上一直很難做到的。比如你考慮任何公司，客戶支持部門與銷售部門是兩個完全獨立的組織。而作爲 CEO，你並不希望它們這樣運作。只是，好吧，它們所需的技能確實有所不同。當你是 CEO 時，你得處理所有這些事情。

我的意思是，一般來說，當你搭建組織的時候，組織是分開的。因爲它們各自優化於不同的事務。但我認爲理想化的狀態是它應該是一種統一的存在，對吧？作爲消費者，你不想在購買東西時走一條路線，而在遇到已購商品問題時又走另一條路線，你只想有一個地方可以去，能解答你的問題，並且能夠以不同的方式與企業互動。我認爲這對創作者也同樣適用。

**黃仁勳：**我認爲這種個人消費者方面的互動，尤其是他們的投訴，將會使你的公司變得更好。

**扎克伯格：**是的，完全正確。  
**黃仁勳：**所有這些 AI 與客戶的互動都會讓 AI 捕捉到那些機構知識，而這些又都可以進入分析系統，這會促進後續人工智能。

**扎克伯格：**是的。你說的這個版本我認爲是整合度更高的版本，我們仍處於相當早期的

alpha 階段。但 AI 工作室讓人們能夠創建自己的 UGC 智能體和其他內容，開始進入這個讓創作者創作的飛輪。我對這個挺興奮的。

**黃仁勳：**那麼，我能用 AI 工作室來微調我的圖片，我的圖片集嗎？

**扎克伯格：**是的，我們正朝那個方向發展。

**黃仁勳：**好的。然後，我能把我寫的所有東西都加載進去，讓它作爲我的 RAG (檢索增強生成) 使用嗎？

**扎克伯格：**是的，基本上可以。

**黃仁勳：**好的，是的。每次我回來，它都會重新加載記憶。所以它記得上次離開的地方，我們繼續對話，就像什麼都沒發生過一樣。

**扎克伯格：**是的。而且我的意思是，就像任何會隨着時間改進的產品一樣，用於訓練它的工具也會變得更好。這不僅僅是關於你希望它說什麼。我認爲通常創作者和企業也有他們想要避開的主題。

**黃仁勳：**對。所以在這方面不斷進步。

**扎克伯格：**是的。你知道，我認爲理想的版本不僅僅是文本，對吧？這和我們隨着時間在做的一些編解碼虛擬形象的工作有些交集。你肯定希望能夠和那個智能體進行像視頻聊天一樣的交互，我認爲我們隨着時間會達到那一步的。我不認爲這些東西離我們太遠，這個飛輪正在非常快速地旋轉。所以這是令人興奮的。有很多新東西需要構建。

我認爲即使基礎模型的進展現在停滯不前，儘管我不認爲它會停滯，我想我們還會有五年的時間來讓行業基本上弄清楚如何最有效地利用迄今爲止已經構建的所有東西。但實際上，我認爲，基礎模型和基礎研究方面的進展正在加速。所以，這是一個非常狂野的時期。

**黃仁勳：**你的願景是這一切，你某種程度上促成了這一切。所以，謝謝你。你知道，我們是 CEO，我們是脆弱的花朵。我們需要很多支持。(笑)

**扎克伯格：**是的。我們現在已經相當堅韌了。我認爲我們是行業中兩位任職時間最長的創始人。對吧？

**黃仁勳：**我的意思是，這是真的。這是真的。  
**扎克伯格：**而且你的頭髮變灰了。我的只是變長了。(笑)

**黃仁勳：**我的頭髮是變灰了。你的變卷了。這是發生了什麼？

**扎克伯格：**它一直都很卷。所以我一直保持短髮。

**黃仁勳：**是啊。我只是，如果我知道成功需要這麼長時間……

**扎克伯格：**你可能一開始就不會開始。(笑)

**黃仁勳：**不，我可能會像你一樣從大學退學。早點開始。(笑)

**扎克伯格：**嗯，這是我們性格的一個好差別。(笑)

**黃仁勳：**你已經領先了 12 年。那很不錯。

**扎克伯格：**你知道，你做得相當不錯。

**黃仁勳：**我會繼續前進的。(笑)

所以，我喜歡的，關於您對每個人都能擁有 AI，每個企業都能在我們公司擁有 AI 的願景。我希望每位工程師和每位軟件開發者都能擁有一個 AI。

**扎克伯格：**是的。我們有很多 AI。

**黃仁勳：**我欣賞您的願景的一點是，您也相信每個人都和每個公司都應該能夠創造自己的 AI。所以您實際上是開源了。當您開源了 Llama 時，我認爲那很棒。Llama 2.1，順便說一下，我認爲 Llama 2 可能是去年 AI 領域最大的事件。

**扎克伯格：**我原本以爲(去年 AI 領域最大的事件)是 H100。(笑)

**黃仁勳：**這是個先有雞還是先有蛋的問題。是的。所以這兩者哪個先出現的？

**扎克伯格：**H100。 嗯，Llama 2 實際上不是 H100。(笑)

**黃仁勳：**是的，它是 A100。(笑)

**扎克伯格：**開源的未來：不受限制的構建，恢復開放黃金時代

**黃仁勳：**是的，謝謝。因此，但我說這是最大事件的原因在於，當 Llama 2.1 發佈時，它激活了每一家公司、每一個企業和每一個行業。突然之間，每家醫療公司都在開發 AI，每家公司，無論大小，初創企業都在構建 AI。它讓每一位研究者都能重新投入到 AI 領域，因爲他們有了一個起點去做些什麼。而現在，3.1 版本已經發佈，你知道的，我們共同努力部署 Llama 3.1，將其推向全球企業。這種興奮之情簡直難以言表。我認爲它將推動各種應用的發展。

但請談談你的開源理念。這一理念源自何處？你開源了 PyTorch，現在它已成爲人工智能開發的框架。而現在，你們已經開源了 Llama 3.1。圍繞它建立了一個完整的生態系統。所以，我認爲這非常棒。但這一切是從哪里來的呢？

**扎克伯格：**是的，這里面有很多歷史。我的意思是，隨着時間的推移，我們做了很多開源工作。我認爲部分原因，坦白說，是我們在其他一些科技公司之後才開始構建分布式計算基礎設施和數據中心這類東西。而且，正因爲如此，當我們構建這些東西時，它已經不是競爭優勢了。所以我們想，那好吧，我們不如乾脆將其開源。然後我們將從圍繞它的生態系統中受益。所以我們有很多這樣的項目。

我認爲最大的一個例子可能是開放計算項目(Open Compute)，我們公開了我們的服務器設計、網絡設計，最終還包括數據中心設計，並發佈了所有這些內容。通過使其成爲某種行業標準，整個供應鏈基本上都圍繞它進行了組織，這帶來

我有點希望在下一波計算浪潮中，我們將重現一個開放生態系統勝出並再次成爲主導的領域。總是會有封閉的和開放的兩種選擇。我知道兩者都有其存在的理由。兩者都有其好處。我並不是在這方面極端的人。我的意思是，我們也有閉源的產品。我們發佈的並非所有內容都是開源的。

但我認爲，對於整個行業所依賴的計算平台而言，尤其是軟件如果開放，會有很大的價值。這確實塑造了我在這方面的哲學。無論是與 Llama 合作的 AI 項目，還是我們在 AR 和 VR 領域的工作，我們基本上是在爲混合現實構建一個開放的操作系統，類似於 Android 或 Windows，旨在與衆多不同的硬件公司合作，製造各種類型的設備，我們基本上只是希望將生態系統恢復到那個開放的水平。我對下一代的發展相當樂觀，開放的系統將會勝出。

對我們來說，我只想確保我們自己能夠接入(AI)。這么說有點自私，但我的意思是，在創建這家公司一段時間後，我未來 10 到 15 年的目標之一就是，我只是想確保我們能夠構建基礎技術，因爲這將是我們構建社交體驗的基礎，因爲之前我嘗試構建的許多東西都受到了限制。然後你被告知，不，你無法真正構建那個，這是平台提供商在某一層面的說法。我只想說，不，去他的。對於下一代，我們將一路構建下去。

**黃仁勳：**好的，我們的直播要被掐了(因爲扎克伯格說了 Fword)

**扎克伯格：**是的，抱歉。抱歉。(笑)

**黃仁勳：**我當時就想說，嗶。(笑)

**扎克伯格：**是的。你知道，我們前 20 分鐘做得還不錯，但一談及封閉平台，我就生氣。(笑)

**黃仁勳：**有人致力於構建儘可能最好的 AI，無論他們如何構建，並將其作爲服務提供給世界，我認爲這才是一個很棒的世界。然後，如果你想自己構建 AI，你仍然可以自己構建 AI，這都很棒。但在使用 AI 的能力這件事上，你知道，有很多東西(我更喜歡開盒即用)。我更傾向於不自己製作這件夾克。我更傾向於讓別人爲我製作這件夾克。你明白我的意思嗎？

**扎克伯格：**是的，是的。

**黃仁勳：**皮革可以開源對我來說不是一個有用的概念。但我認爲你可以擁有偉大的服務，令人難以置信的服務，以及開放的服務，開放的。

**扎克伯格：**是的。

**黃仁勳：**那麼我們基本上涵蓋了整個範圍。但你用 Llama 3.1 所做的事真的很棒，你有 405B 的版本，你有 70B 的版本，你有 8B 的版本，你可以用它來生成合成數據，用更大的模型來教小模型。雖然更大的模型會更通用，它也不那麼脆弱，但你仍然可以構建一個更小的模型，適合於任何你希望的操作領域或運營成本。無論你希望的操作領域或運營成本是什麼，你都可以構建一個更小的模型來適應。

你創建了一個守衛，我想是叫 Llama 守衛(Llama Guard)，Llama 守衛用於護欄，非常棒。所以現在你構建模型的方法，它是以一种透明的方式構建的。你擁有一個世界級的完全團隊，世界級的倫理團隊。你可以以這樣一種方式構建它，讓每個人都知道它是正確構建的。我真的很喜歡這部分。

**扎克伯格：**是的，我先說完在我之前被岔開話題之前想說的話。你知道，我確實認爲我們有這種一致性，我們構建(開源 AI)是因爲我們希望它存在，並且我們不希望被某些封閉的模型所隔離。而這不僅僅像是開發一款軟件那麼簡單，AI 需要一個生態系統的支撐。因此，幾乎可以說，如果我們不將其開源，它甚至可能無法很好地運作。

我們這樣做並非因爲我們是什麼無私的人，儘管我認爲這對生態系統是有益的。我們這樣做是因爲我們認爲，通過建立一個強大的生態系統，可以使我們正在構建的東西達到最佳狀態，看看有多少人參與到了 PyTorch 生態系統的建設中。

**黃仁勳：**(這需要)大量的工程工作。我的意思是，僅視頻處理方面，我們可能就有數百人專門致力於讓 PyTorch 更優秀、更具擴展性，你知道的，性能更佳等等。

**扎克伯格：**是的，而且當某個技術成爲行業

(未完接下期)



了爲每個人節省資金的好處。因此，通過公開和開放這些設計，我們基本上通過這樣做節省了數十億美元。

**黃仁勳：** 嗯，開放計算項目也使得 NVIDIA HGX 成爲可能，我們爲一個數據中心設計的東西突然間也能在其他數據中心工作。

**扎克伯格：**確實如此。太棒了。所以我們有了這次非常棒的經歷。之後，我們還將其應用於我們的一系列基礎設施工具，比如 React, PyTorch 等。因此，可以說到了 Llama 出現的時候，我們對做這件事有一種積極傾向。

特別是對於 AI 模型，我想我有幾種看待這個問題的方式。首先，在過去 20 年里，在公司里構建東西中最困難的事情之一就是不得不應對這樣一個事實：我們通過競爭對手的移動平台發佈我們的應用。一方面，移動平台對整個行業產生了巨大的推動作用，這非常棒。另一方面，不得不通過競爭對手來交付你的產品是具有挑戰性的，對吧？

而且，在我成長的時代，Facebook 的第一個版本是在網絡上，那是開放的。然後轉向移動端，其好處是，現在每個人口袋里都有一台電腦。那真是太好了。但不利的一面是，我們在能做的事情上受到了更多限制。因此，當你審視這些計算時代的更迭，你會發現存在一個巨大的近因偏差，人們現在只關注移動設備時代，並認爲封閉的生態系統(會帶來更好的結果)，因爲蘋果基本上贏得了這場勝利並設定了規則。是的，我知道從技術上講，有更多的安卓手機存在，但蘋果基本上掌控了整個市場，而且(贏得了)所有的利潤，基本上安卓在發展方面也在跟隨蘋果的腳步。所以我認爲蘋果在這一代中明顯勝出。

但情況並非總是如此。

如果你回顧上一代，蘋果在搞他們那種封閉的模式(MacOS)。但微軟，它顯然不是一個完全開放的公司，但相較於蘋果，Windows 在各種不同的 OEM 廠商和不同軟件、硬件上運行，是一個更爲開放的生態系統。而 Windows(那時)是領先的生態系統。基本上在個人電腦這一代產品中，開放的生態系統勝出了。

