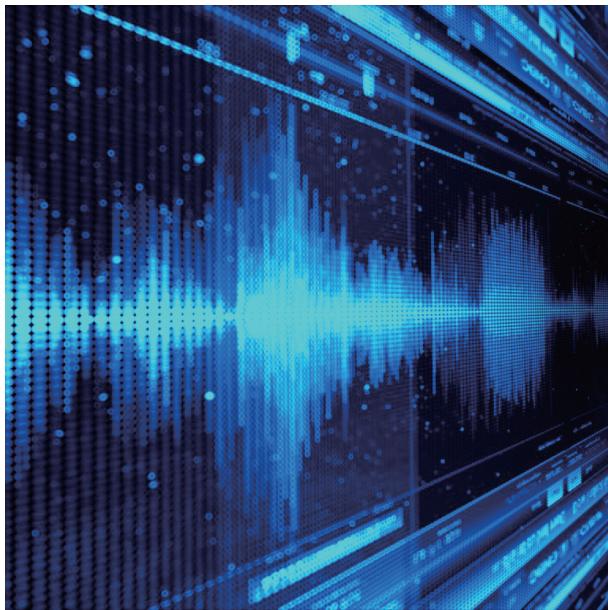


# 語音互動時代到來 對話助理已面世

■作者：Moshe Sheier/

Director of Strategic Marketing, CEVA

聲控虛擬助理已以極快的速度蔓延。蘋果的 Siri、亞馬遜的 Alexa、Google Home、微軟的 Cortana 以及許多其他產品正在相互競爭，助我們一臂之力。讓我們看看此次變革的促因、其發展過程中面臨的挑戰以及在其中發現的問題。



## 從語音辨識發展為語音理解

還是在幾年前，語音辨識是一個巨大的挑戰，並且是人們最為夢寐以求的無操作使用者介面之一。第一批版本的 Siri 誤聽我所說內容或反復回覆「抱歉，我沒聽清楚」的情景仍記憶猶新。如今，在成本降低和效率提高的處理能力的推動下機器學習方面已有重大突破，這使語音辨識變得非常普遍，以至於使它成為一種理所當然的技術。在最新發表的主題中，谷歌的高級研究員 Jeff Dean 聲稱在神

經網路套用語音辨識時詞語錯誤率將降低 30%。除了語音辨識的直接改進，噪音消除和語音增強也因神經網路而受益頗多。Cypher 技術為一個傑出典範，它使用深度神經網路隔離聲音。ASR 有針對性的消音改善了語音辨識引擎的原始資料，這使任務成功的可能性更大。這些因素促成了如今的局面，即語音辨識在多種裝置上成為可靠且實用的介面。

因此，該行業已轉為處理下一巨大挑戰：理解豐富的口語，包括並列複合句和根據上下文做出的判斷。簡單來說，機器是在逐漸獲得進行對話的能力。據專業從事語音理解的公司 Houndify 稱，您可以向其對話介面詢問以下問題：「告訴我位於舊金山於明天入住兩晚、價格介於 200 到 300 美元、可攜帶寵物、有健身房和泳池且只配有空調的 3 星或 4 星酒店。」除了複雜的問題以外，他們還宣稱他們的平台還支援語境和後續問題，這使語音互動與我們正常講話相似。這種複雜程度在不到十年前還屬於科幻小說的領域。在幾乎所有行業巨頭皆在各種平台發佈其機器學習架構之後，在實現完全語音理解方面的努力也從開源社區取得動力。

此加速發展背後的推動力是最近幾年機器學習推動下人工智慧的飛躍性進展。除了軟體方面的發展之外，硬體也在穩步提升，從而實現更加強大有效的計算。這種同步發展的範例就是谷歌的張量處理單元 (TPU)。這些皆為谷歌內部開發的加速器晶片，用於推動其機器學習專案，如 AlphaGo。量身定製適合深度神經網路算法需求的硬體，這樣可使硬體具有顯著競爭優勢。



在處理連線行動裝置時，電力和回應時間限制甚至更為重要。大小、重量和電池使用壽命方面的限制使挑戰更具難度，但這些毫無疑問皆會一一被戰勝。使低功耗便攜式裝置充分利用智慧優勢毫無疑問是本次變革的下一步措施。實現這一技術要求注意考慮達到所需功效所需的每一毫瓦和每一時鐘週期。

## 家庭領域的對話助理

在智慧型家庭案例中，亞馬遜開啓了 Echo 趨勢，將其對話助理 Alexa 推入到千家萬戶。這個總是在聆聽的成功無線擴音器最初不僅僅是一個麥克風，而且還是具有一些炫酷功能的遠距離聲控音樂系統。從此以後，Alexa 變得越來越智慧便利，引入了許多全新的功能和技能，如訂購披薩和在騎行時歡呼。亞馬遜還發佈了一個便攜式版本，名為 Echo

Tap。Tap 的主要不足之處正如其名稱一樣，您輕觸一下才可講話。一款更加有效的內嵌解決方案可支援與插入的 Echo 版本相同的總是聆聽的語音啓用功能，並且提供真正的無操作介面。

在最近召開的谷歌 I/O 會議中，谷歌宣稱他們發佈的智慧型助理的版本稱為 Google Home。該單元加入與 Android 裝置相同的谷歌助理，在呼出「OK Google」語句後啓用。此助理已擁有許多技能，您可以在此精美的詞彙表中瀏覽這些技能。在 I/O 主題會議中，谷歌執行長 Sundar Pichai 宣稱此助理將繼續不斷完善，支援語境、後續問題和複雜句，從而實現完全對話模式。

蘋果最近向第三方開發人員開放 Siri API 的決策，也為在戰略市場中展開競爭做好蓄勢待發之勢。這顯而易見是一個重要舉措，因為使用者將傾向於使用具有最豐富的生態系統、應用程式和功能的助



理。這些資訊時代的網際網路巨頭已清楚地認識到進入家庭是深入到消費者内心和錢包的關鍵入口點.....

一個真正給予幫助的虛擬助理是無處不在的，無時無刻都在為您提供所有公司服務。對話助理將不會限制在家庭領域，不久我們將會在許多其他領域(如汽車、行動裝置，甚至我們的耳朵內)發現這一技術。

## 記錄一切資訊：對話助理的隱私問題

多數情況下，原始語音資料上傳到雲端後，

這些助理的處理才會完成。這需要使用大量的伺服器，並且伺服器需擁有強大的計算能力方可完成處理。所有資料上傳後將會由谷歌、亞馬遜記錄和保留，或者由任何公司的服務使用。退一步說，這將引發一些隱私問題。基本上，這意味著我們正在向我們所提到的一切資訊由這些公司記錄保留的階段發展。繞開此問題的一種可行方法為擁有一個足夠強大的處理器，可在本地執行軟體，且無需上傳語音資料。在這種情況下，只有在需要取得資訊時才需存取雲端，如果天氣、方向或其他服務，但不提供具有敏感性的隱私資訊。另一方面，機器理解的改進和實現是巨量資料收集的成果。要繼續學習和改進，系統將不斷需要更多資料。也許解決方案將會以規範的形式產生。看到整個演變過程，也將是一件趣事。

## 查找更多資訊

CEVA一直活躍于戰勝將智慧引入微小裝置這些挑戰的最前線。我們的超低功耗 CEVA-TeakLite-4 處理器用於高效、總是聆聽 / 弱音解決方案，具有降噪、語音識別等更多功能，全部集於一個超薄套件中。此處理器最近已於聲控 Galaxy S7 中演示，由 CEVA 提供技術支援，採用客戶的 DSP Group 晶片。 [CTA](#)

## Xilinx 宣布 16 奈米 UltraScale+ 全系列提前量產

美商賽靈思 (Xilinx) 日前宣布旗下 16 奈米 UltraScale+ 系列提前達成重要的量產里程碑。自全系列元件首次出貨至今日宣布本季度開放量產元件訂單，用了僅不到一年的時間。

賽靈思 UltraScale+ 系列是 14 奈米或 16 奈米工藝製程上，業界唯一基於 FinFET+ 的可編程技術產品，其產品系列包含 Kintex、Virtex UltraScale+FPGAs 及 Zynq UltraScale+ MPSoCs 等元件，目前全系列已開放量產訂單。

