



Siri，為什麼妳不聰明點？

語音辨識軟體棒呆了，但用在電話上卻不是這麼回事。

蘋果去年發表 iPhone 4S，這款新手機看起來和前一代幾乎一模一樣。相機有好一點，晶片有快一點，但真正的新功能只有一樣：Siri。現在大家都知道，Siri 是能夠接受語音指令的軟體助理，不需要事先訓練：只要按下 Home 鍵，然後正常講話就好。

Siri 在文化世界掀起一股熱潮。你可以在 Youtube 上看到諧擬影片、在網路上找到基本指南，Android 手機上也有模仿的應用程式。現在人們不跟人講電話時，也會對著手機講話了，所以專家針對在公共場合使用手機提出了新的禮儀規範。語音辨識風靡一時，突然間，電視機也有語音辨識了，當然，競爭廠牌的手機就更不用說了。在這浪頭上，我們和電子產品的互動方式似乎永遠改變了。

但接下來出現的——卻是挫折。

「Siri 讓蘋果失信於人。」這是科技網站 Gizmodo 的頭條。人們抱怨有時說了一整段話，手機會思考一下，然後秀出一片空白。現在已經有人提出集體訴訟，宣稱蘋果的廣告不實。（根據蘋果說法，Siri 還在 β 測試階段。）

發生了什麼事？為什麼 Siri 這個電子產品的救世主變成了泡沫？其實大家忽略了一件事，做為虛擬助理的 Siri 和語音辨識引擎的 Siri 是有差別的。現實的情況是，這兩種不同的功能在成功率上有截然不同的成績。

Siri 做為助理的這一半，是源自於蘋果收購的 Siri 公司。（這間公司是從研究機構 SRI 參與的軍事人工智慧計畫衍生出來的。懂了吧？）但 Siri 的語音功能，也就是文字轉聲音這部份，是由 Nuance 這家公司提供，該公司也提供 Dragon NaturallySpeaking 這類軟體。

當你對著手機講話時，機器會產生一個聲音檔，傳送給 Nuance 的伺服器分析語音內容，再將文字傳回手機。這也是為什麼當網路訊號微弱或者手機上網塞車時，Siri 會有點短路（當你用 Wi-Fi 時，她表現會好很多）。資料得在手機和伺服器間傳遞，是 Siri 語音功能正確性不高的關鍵。



此外，背景雜音、風切聲，還有嘴巴到麥克風的距離不一，都會讓錄製過程不夠完美，而且結果的正確率會比個人電腦上的語音軟體要低。使用 Siri，可能每講一段話就要修正幾個錯誤。

個人電腦的語音軟體表現就好很多，正確率接近 100%，因為它不必面對這些特別的挑戰。而且在個人電腦上，你訓練軟體辨識的聲音只有一個：你自己。手機上則沒有這樣的訓練，計算工作因此變得格外艱難。

批評 Siri 的人說，我們已經習慣每次都順利運作的消費性科技了：e-mail、GPS、數位相機，但是仰賴網路的語音技術卻做不到。在這個年代，這樣的表現會讓人不滿。

但別因噎廢食。Siri 的「虛擬助理」部份，不管是設鬧鐘、打電話、傳簡訊或是行事曆，她都表現得很好。即使你只使用基本指令，像是「幾點叫醒我」、「打電話給某人」、「傳簡訊」和「提醒我」，也的確省下時間而不再手忙腳亂。不設限的手機語音功能技術尚未成熟，但它讓語音功能的未來，幾乎像一年前 Siri 所宣示的那樣光明。

且讓我們期待她走出 β 版吧。（鍾樹人 譯）

SA

波哥 (David Pogue) 是美國《紐約時報》科技專欄作家，並擔任美國公共電視「新星」節目「現在科學」單元的主持人。