

繁體中文 AI 書齋版

人工智慧 翻譯動物的語言

傾聽鯨魚、海豚、鳥與蜜蜂之聲的
人工智慧故事

金京鎮律師

kimkj.com • AI 書齋 • 2026

人工智慧翻譯動物的語言

傾聽鯨魚、海豚、鳥與蜜蜂之聲的人工智慧故事

金京鎮律師

本書從陪伴巴哈馬海豚 40 年的研究者、在抹香鯨喀嗒聲中找出像母音般規律的語言學家，以及與座頭鯨 Twain 相互發出聲音交流 20 分鐘的科學家故事開始。全書整理成十二章，內容涵蓋人工智慧像畫地圖一樣描繪動物聲音並找出規則的方法、解讀鳥類、蜜蜂與猴子信號的新工具、尊重動物感覺世界的態度、用假聲音欺騙動物的危險與防範規則，以及賦予鯨魚法律人格的討論。為了讓國小高年級學生也能獨自閱讀，全書以簡短好懂的句子寫成，每個小章節都附上確認過的佐證資料，並且是經過兩位人工智慧審查員檢驗修訂後的版本。

公開研究資料庫

這本書是由人工智慧整理製作而成。內容收集並統整了全世界主要語言的資料。無論資料是用哪種語言寫成，NotebookLM 都不會限制提問的語言。讀者可以用自己的語言，向準備這本書時使用的同一個研究資料庫提問。即使部分資料是用韓文、英文或其他語言寫成，也是一樣的。只要在 NotebookLM 中打開設定，將輸出語言指定為想要的語言，回答與生成的內容就會以該語言呈現。

準備這本書時所使用的公開 NotebookLM 資料庫（人工智慧翻譯地球的生命）：

<https://notebook.google.com/notebook/bb152971-e02c-4008-b56e-4076e1180e53>

這本書的韓文原版在 kimkj.com AI 書庫：<https://kimkj.com/ai-library/?mod=document&uid=12655>

目錄

點選各章標題即可前往該章內文。若想從頭開始閱讀，請選擇導言。

導言：開啟地球共同體的聲音 • 4

第1章 Wild Dolphin Project與斑點海豚的世界 • 9

第2章 抹香鯨的音素字母與 Project CETI • 15

第3章 與座頭鯨 Twain 的相遇 • 20

第4章 深海交響樂團 • 25

第5章 大地的語言 • 30

第6章 天空的交響曲與看不見的舞蹈 • 35

第7章 從動物的大腦到深度學習畫布 • 40

第8章 動物的環境界與符號學邊界 • 45

第9章 欺瞞與保育的十字路口 • 50

第10章 非人類人格體與機器輔助權利 • 55

尾聲：跨物種網際網路與共存的地球 • 60

導言：開啟地球共同體的聲音

0.1 哈伯望遠鏡為我們打開的新雙眼

1995年12月，科學家把太空望遠鏡轉向天空。那座望遠鏡的名字叫做哈伯太空望遠鏡(Hubble Space Telescope)。望遠鏡對準的地方，是天空中最黑暗、最空曠的角落。許多學者以為那裡什麼都沒有。望遠鏡連續幾天在那片黑暗中收集微光。接著出現了意想不到的結果。黑暗中竟然藏著數千個星系。星系是由許多星星聚集而成的巨大群體。那裡不是空無一物，而是滿滿當當的。這件事給科學界留下了重大的啟示。只要用新的工具照亮，看似空曠的地方其實也藏著很多東西 [1]。

如今，人工智慧也正在做類似的事情。這次對準的目標不是天空，而是生命之間互相傳遞的聲音世界。這項新工具被稱為數位生物聲學(Digital Bioacoustics)。這是一門把生物發出的聲音錄下來，並用電腦進行研究的學問。很久以前，西方科學並不重視動物的聲音，只當成是本能發出的噪音。可是當人工智慧開始分析這些聲音時，大家的想法改變了。森林與海洋，原來是充滿豐富資訊的對話舞台 [2]。

人類的眼睛和耳朵，只能感受到世界的一小部分。我們很容易誤以為自己感受不到的東西就不存在。有一種叫做黑眼燈草鴉(dark-eyed junco)的小鳥。這種鳥的叫聲，在人類聽起來很平淡單調。可是鳥類耳朵的時間感比人類更敏捷。對鳥類來說，那其實是由兩層旋律交織而成的精緻歌曲。大象會用非常低沉的聲音傳向遠方來交流。這種聲音介於14到35赫茲(Hz)之間。其中較低的部分是人類聽不見的聲音。這種聲音可以傳遞到幾公里之外。如果環境條件良好，還能傳得更遠。蝙蝠會用人類聽不見的高音互相爭吵。連沒有耳朵的珊瑚幼蟲，也會靠聲音來尋找落腳的地方 [3]。

人類能聽見的聲音，上下都有一道牆。太高而聽不見的聲音稱為超音波(ultrasonic)。「超」是有超越的意思。這是超越了人類耳朵上方那道牆的聲音。太低而聽不見的聲音則稱為超低頻波(infrasonic)。這是超越了人類耳朵下方那道牆的聲音。蝙蝠的聲音屬於超音波。大象聲音中低沉的部分屬於超低頻波。人類在很長一段時間裡，都忽略了這兩種聲音。隨著錄音設備與分析技術進步，我們終於能夠逐一研究這些聲音。能聽見聲音的，並不只有長著耳朵的動物。也有研究發現，沒有耳朵的植物也會對聲音產生反應。

科學家在水中也能捕捉到這些聲音。放入水中的麥克風叫做水下聽音器(hydrophone)。深度學習(deep learning)會分析這樣收集到的聲音。深度學習是一種能從大量資料中自己找出規則的技術。人工智慧就這樣打破了人類感官的框架。通往地球漫長對話的大門正在緩緩開啟。

相關佐證資料

- [1] Bakker, K. (2022). *The Sounds of Life: How Digital Technology Is Bringing Us Closer to the Worlds of Animals and Plants*. Princeton University Press.
- [2] Earth Species Project. (2024-11-11). Introducing NatureLM-audio: An audio-language foundation model for bioacoustics. Earth Species Project. <https://www.earthspecies.org/blog/introducing-naturelm-audio-an-audio-language-foundation-model-for-bioacoustics>
- [3] Poole, J. H., Payne, K., Langbauer, W. R., & Moss, C. J. (1988). The social contexts of some very low frequency calls of African elephants. *Behavioral Ecology and Sociobiology*, 22(6), 385-392. <https://doi.org/10.1007/BF00294975>

0.2 把語言變成地圖的人工智慧

人工智慧是如何處理人類語言的呢？人工智慧會把話語變成一種地圖。這張地圖被稱為潛在空間(Latent Space)。這是一張看不見的地圖，把意思相近的詞聚集在一起。如果把這張地圖想像成布滿星星的夜空，就很容易理解了。每一個詞彙都變成一顆星星。意思相近的詞會聚攏在附近。意思相差很遠的詞則彼此隔得很遠 [1]。

這張地圖的威力不僅僅來自距離。詞彙之間的關係也可以用方向和距離來表示。例如，有一條箭頭連接著「國王」與「男人」。連接「女王」與「女人」的箭頭也和它非常相似。因此，如果從「女王」減去「女人」，再加上「男人」，就會非常接近「國王」。在某些詞彙地圖中，這種計算相當準確。「狗」這個詞也不會孤零零地落單。它會和「人類」、「吠叫」、「毛」這類的詞相連並找到自己的位置。

電腦無法像人類那樣體會詞彙背後的深層含意。它也不懂得悲傷或喜悅之類的情感。電腦所做的，只是計算彼此之間的關係而已。它只推算詞彙之間是如何相互連結的。這件事就像在解一道巨大的填空題一樣。就像一步一步填入空格裡的字詞一樣，人工智慧也逐一拼湊出詞彙的位置。語言精緻的樣貌就這樣被建構出來了 [2]。

聲音也可以用這種方式轉換成地圖。人工智慧會把聲音切成細小的片段，轉換成數字。再把這些數字像地圖上的星星一樣排列開來。相似的聲音會聚在一起。不同的聲音則隔得很遠。無論是人類的話語還是動物的聲音，在地圖上排列的方法都很相似。只不過，這種關聯是否能跨越物種並傳達真正的含意，目前仍在研究中 [3]。

語言就像是世界所投下的影子。人工智慧逆著追尋那個影子的足跡。進而把世界隱藏的結構繪製成地圖。有了這張地圖，接下來的故事才有可能展開。因為如果能把人類的語言畫成地圖，動物的聲音也同樣能繪製成地圖。

相關佐證資料

- [1] Mikolov, T., Yih, W., & Zweig, G. (2013). Linguistic regularities in continuous space word representations. Proceedings of NAACL-HLT 2013, 746-751.
- [2] Bakker, K. (2022). The Sounds of Life: How Digital Technology Is Bringing Us Closer to the Worlds of Animals and Plants. Princeton University Press.
- [3] Earth Species Project. (2026). What we do: Technology. Earth Species Project.
<https://www.earthspecies.org/what-we-do/technology>

0.3 沒有字典的翻譯與 Earth Species Project

2017年左右，機器翻譯領域發生了一件大事。那是在完全沒有連接兩種語言的字典的情況下。在過去，必須要有羅塞塔石碑(Rosetta Stone)。羅塞塔石碑是一座刻有同一段文字、多種語言的古老石碑。多虧了這座石碑，人們才得以解開古埃及文字的秘密。然而，人工智慧在沒有這種字典的情況下，就把兩種語言的詞彙配對了起來。這是透過能夠自主學習的機器學習所做到的 [1]。

方法就是把兩種語言的地圖疊合在一起看。人工智慧分別製作了英語的星圖，也單獨製作了西班牙語的星圖。接著，小心翼翼地將這兩張地圖疊加在一起。許多學者原本認為這兩張地圖不會吻合，因為每種語言看待世界的方式都不同。可是，這兩張地圖卻很平順地重疊在一起了。代表「狗」的那顆星星，幾乎落在了同一個位置上。這種將形狀不同的兩張地圖縮放並對齊重疊的方法，被稱為流形對齊(manifold alignment)。

這項成果不僅僅出現在歐洲的語言之中。即使是文法差異極大的語言，也呈現出類似的形狀。人類創造的人造語言世界語中也是如此。人類的多種語言，似乎共同擁有一個巨大的形狀。學者認為，這是因為「人類」同屬一個物種，感受著同一個地球的緣故 [2]。

Earth Species Project 的 Aza Raskin 在這裡提出了一個疑問。如果人類的語言可以這樣重疊，那動物的聲音是不是也能用同樣的方法呢？這個組織是2017年創立的非營利研究機構。他們收集白鯨、虎鯨以及紅毛猩猩的聲音。他們為每種聲音分別建立星圖。再將那些地圖小心地疊加在人類語言的地圖上。如此一來，人類就能用全新的雙眼去觀察鄰近生命的聲音。他們最近研發的工具是 NatureLM-audio。這是一款能聆聽動物聲音，辨識出物種和叫聲種類的人工智慧。2025年，他們公開了人人都能試用的示範版本。只要上傳聲音檔案並寫下問題，就會給出答案。到了2026年，它的功能變得更加強大。2026年7月，一個供多個研究機構共用資料的資料庫 alp-data 也正式上線了。這樣的趨勢，與「不同的人工智慧最終都會得出形狀相似的地圖」這一想法緊密相連。這種想法被稱為柏拉圖式表徵假說(Platonic Representation Hypothesis) [3][4][5]。

相關佐證資料

- [1] Lample, G., Conneau, A., Denoyer, L., & Ranzato, M. (2018). Unsupervised machine translation using monolingual corpora only. International Conference on Learning Representations (ICLR) 2018. arXiv:1711.00043.
- [2] Robinson, D., Miron, M., Hagiwara, M., et al. (2024/2025). NatureLM-audio: An audio-language foundation model for bioacoustics. International Conference on Learning Representations (ICLR) 2025. arXiv:2411.07186.
- [3] Earth Species Project. (2025-09). Hands-on with NatureLM-audio: A no-code demo. Earth Species Project.
<https://earthspecies.org/blog/hands-on-with-naturelm-audio-a-no-code-demo>

- [4] Earth Species Project. (2026-07-22). Introducing alp-data: A shared data layer for animal language processing. Earth Species Project.
<https://earthspecies.org/2026/07/22/introducing-alp-data-a-shared-data-layer-for-animal-language-processing/>
- [5] Huh, M., Cheung, B., Wang, T., & Isola, P. (2024). Position: The Platonic Representation Hypothesis. Proceedings of the 41st International Conference on Machine Learning (ICML). arXiv:2405.07987.

0.4 尋找聲音中隱藏的規則

動物的聲音是毫無意義的噪音嗎？還是像人類的話語一樣有規則呢？科學家透過統計聲音的數據來研究這個問題。一個眾所皆知的規則是：越常使用的詞往往越簡短。這個規則被稱為齊夫縮短定律(Zipf's law of abbreviation)。它是以語言學家 George Kingsley Zipf 的名字命名的。人類為了節省力氣，會把常用的話變短 [1]。

語言學家 Stuart Semple 和 Ramon Ferrer-i-Cancho 的研究團隊在許多動物身上檢驗了這個規則。他們收集了包含猴子在內的多種哺乳動物的聲音進行比對。電腦將聲音細細切分並加以計算。結果完全一致。越常發出的聲音，長度就越短。和人類語言相似的規則，也同樣存在於動物的聲音中。同時也出現了整體聲音越長、內部的組成片段就越短的規則。這個規則的名字叫做門澤拉特-奧爾特曼定律(Menzerath-Altmann law)。

另外還有一個數值，可以用來衡量下一個聲音有多難預測。它測量某個聲音之後會出現什麼聲音的預測難度。這個數值被稱為夏農熵(Shannon entropy)。它是以創立資訊理論的 Claude Shannon 之名命名的。如果聲音是毫無意義的尖叫，這個數值就會偏向某一極端。如果是有規則的話語，就會落在中間的數值 [4]。

Whale-SETI 研究團隊使用這個數值研究了鯨魚的聲音。團隊成員包括 Brenda McCowan、Fred Sharpe 和 Laurance Doyle。他們仔細分析了座頭鯨(humpback whale)的歌聲。鯨魚的歌聲恰好落在規則與自由之間的平衡點上。歌聲中也展現出一種結構：常出現的聲音較短，罕見的聲音較長。這樣一來，就能用更少的力氣傳達相同的意義。在電腦中，也有一種將常用資料記短、罕見資料記長來縮小體積的壓縮方法。這個方法的名字叫做霍夫曼編碼(Huffman coding) [2]。

研究團隊甚至進行了與活生生的鯨魚進行聲音互動的實驗。他們在阿拉斯加海域遇到了一隻名叫 Twain 的座頭鯨。研究團隊透過水下揚聲器，播放預先錄製好的鯨魚問候聲。這並不是人工智慧即時創造出的對話。這種問候聲被稱為 whup call。鯨魚就像人類一樣遵守輪流發言的順序，給予回應。牠等待聲音結束的空檔，再填補那個空白。這樣的來回互動持續了20分鐘。就像在吃午餐時和朋友輪流說話一樣，互相傳遞著聲音 [3]。

相關佐證資料

- [1] Semple, S., Ferrer-i-Cancho, R., & Gustison, M. L. (2022). Linguistic laws in biology. Trends in Ecology & Evolution, 37(1), 53-66. <https://doi.org/10.1016/j.tree.2021.08.012>
- [2] McCowan, B., Hubbard, J., Walker, L., Sharpe, F., Frediani, J., & Doyle, L. (2023). Interactive bioacoustic playback as a tool for detecting and exploring nonhuman intelligence: "conversing" with an Alaskan humpback whale. PeerJ, 11, e16349. <https://doi.org/10.7717/peerj.16349>
- [3] Earth.com. (2024). AI helps humans have a 20-minute 'conversation' with a humpback whale named Twain. Earth.com. <https://www.earth.com/animals/whale-seti-ai-helps-humans-have-20-minute-conversation-with-humpback-whale-twain/>
- [4] Bakker, K. (2022). The Sounds of Life: How Digital Technology Is Bringing Us Closer to the Worlds of Animals and Plants. Princeton University Press.

0.5 在抹香鯨的喀噠聲中找到母音

這個方法大放異彩的地方是在深海之中。那裡的主角是抹香鯨。抹香鯨的體長可達20公尺。牠的大腦重約9公斤。是人類大腦重量的6倍以上。這種鯨魚在漆黑一片的深海中捕食。因此牠們不用眼睛，而是用聲音來看世界。牠們發出短促的喀噠聲來探測周遭環境。這一串串喀噠聲被稱為 coda [1]。

長久以來，學者們都把 coda 視為沒有意義的信號。以為只不過是幾種固定的警示音。2020年，Project CETI 改變了這個看法。這是一個匯集了多所大學的人工智慧學者與鯨魚學者的團隊。他們在多米尼克海域收集錄音資料。他們花費10年以上收集的 coda 超過了8,700個。確切來說是8,719個。人工智慧分析了這些聲音。找出了遠遠多

於過去所知的21種模式的組合。還發現鯨魚會改變聲音的速度和彼此之間的間隔。牠們甚至會在聲音末尾加上像句子句號一樣的喀噠聲。也會微幅調整 coda 的長度。這種細微的節奏變化被稱為彈性速度(rubato)。這就像歌唱家把節拍稍微放慢或加快一樣。這是把微小的聲音片段組合起來，創造出豐富意義的方式。這種人類語言中也存在的方法，在動物身上得到了清楚的證實。抹香鯨還會根據傾聽的情境，使用不同的 coda。這種配合對方挑選說話方式的能力，被稱為聽眾設計(audience design)。

2025年，一項新研究帶來了更大的發現。這是美國柏克萊大學與 Project CETI 合作進行的研究，由 Gašper Beguš 教授帶領。研究團隊縮小了鯨魚喀噠聲之間的間隔，以比原本快上數倍的速度播放來聽。結果，原本分開的聲音連成了一整串。那個聲音聽起來就像人類的母音。研究團隊分辨出了「阿」的聲音和「伊」的聲音，也找到了從一個母音流動到另一個母音的雙母音。研究團隊還使用了能探查人工智慧內部的工具，觀察人工智慧在判斷鯨魚聲音時會對哪個部分產生反應。結果發現，聲音的共鳴峰非常重要。這種共鳴峰被稱為共振峰 (formant)，是依照嘴巴或頭部共鳴腔形狀而顯著突出的聲音波峰。這也是人類分辨母音時所依據的數值。2026年，接續這項成果的後續研究出爐了，這是一項將鯨魚聲音的結構與人類音韻學相互比對、進行更深入探討的研究 [2][5][3]。

2026年3月，Project CETI 公布了另一項研究。這項研究記錄了多米尼克海域抹香鯨的生產過程。它收錄了超過六小時的聲音以及從空中拍攝的影像。鯨群中的其他鯨魚協助了生產。這是第一次用數據證實非靈長類動物具有這種合作行為。這項研究再次展現了抹香鯨的社會是透過聲音緊密相連的 [4]。

抹香鯨沒有像人類一樣的聲帶。牠們改為使用頭部前方巨大的器官以及鼻腔內的組織。透過震動這些部位來製造聲音的共鳴。走過與人類截然不同道路的生命，自己創造出了與人類相似的母音。

相關佐證資料

- [1] Sharma, P., Gero, S., Rus, D., Goldwasser, S., & Andreas, J. (2024). Contextual and combinatorial structure in sperm whale vocalisations. *Nature Communications*, 15, 3617. <https://doi.org/10.1038/s41467-024-47221-8>
- [2] Beguš, G., Sprouse, R. L., Leban, A., Silva, M., & Gero, S. (2025). Vowel- and diphthong-like spectral patterns in sperm whale codas. *Open Mind*, 9. <https://doi.org/10.1162/OPMI.a.252>
- [3] Beguš, G., Sprouse, R. L., Leban, A., & Gero, S. (2026). The phonology of sperm whale coda vowels. *Proceedings of the Royal Society B*, 293. <https://doi.org/10.1098/rspb.2025.2994>
- [4] Aluma, A., Gero, S., et al. (2026). Description of a collaborative sperm whale birth and shifts in coda vocal styles during key events. *Scientific Reports*, 16, 9206. <https://doi.org/10.1038/s41598-025-27438-3>
- [5] University of California. (2025-11-13). UC Berkeley and Project CETI study shows sperm whales communicate in ways similar to humans. University of California. <https://www.universityofcalifornia.edu/news/uc-berkeley-and-project-ceti-study-shows-sperm-whales-communicate-ways-similar-humans>

0.6 不是支配，而是傾聽

在人工智慧研究中，有一個歷史悠久的教訓。那就是電腦科學家 Richard Sutton 所說的苦澀教訓 (The Bitter Lesson)。它的意思是：龐大的資料與強大的運算能力，勝過精巧聰明的設計。解讀動物聲音的工具也是這樣發展起來的。其中之一就是 DolphinGemma。這是由 Google、喬治亞理工學院與 Wild Dolphin Project 共同打造的。Wild Dolphin Project 是一項持續了40多年的海豚研究。DolphinGemma 是一個用來判斷聲音、內部數值約有4億個的小型人工智慧。因為體積很小，連拿在手裡的 Pixel 手機也能順暢運行。DolphinGemma 會用一種稱為 SoundStream 的方法把聲音切成細小的碎片。透過觀察這些碎片，就能預測接下來會出現什麼聲音。這和幫人類接續寫出下一個字詞的人工智慧運作方式相同。這項工具現在也正在大海中跟著海豚進行測試。Google 計畫將這項工具公開，讓任何人都能自由使用 [3][1]。

然而，這項飛速發展的技術也有令人擔憂的一面。Aza Raskin 反覆強調了這一點。人工智慧在理解意思之前，就能先模仿出聲音。它能做出和鯨魚一模一樣的聲音來持續對話。這就等於在不知道含義的情況下，將假聲音傾倒進大海中。抹香鯨透過聲音與行為來分享彼此的文化。從前當捕鯨船出現時，抹香鯨彼此學習避開危險的方法。捕鯨船出現僅兩年半的時間，被捕獲的鯨魚數量就減少了58%。這被認為是鯨魚互相通報危險的結果。牠們的溝通交流，比我們想像的連結得更為寬廣。座頭鯨的歌聲也是跨越大海廣泛傳播的文化。人類製造的假聲音可能會打亂這種連結。在完全明白含義之前就先製造聲音，是一件危險的事 [2][4][5]。

因此，這趟旅程的方向非常重要。這項技術絕不能成為隨意使喚動物的工具。它必須是用來如實承認並聆聽動物內心想法的過程。這被稱為跨物種理解（Interspecies Understanding）。人類不是在對別人滔滔不絕說話時改變，而是在放低聲音、接納對方時才會有所改變。

過去的人們曾相信地球是宇宙的中心。望遠鏡證明了那個想法是錯誤的。如今，聲音的地圖再次提醒了我們。人類並不是生命世界的中心。我們只不過是各自擁有聲音的巨大圓圈中的一部分。這份覺察，就是開啟地球聲音的第一步。

相關佐證資料

- [1] Google. (2025). DolphinGemma: How Google AI is helping decode dolphin communication. Google Blog. <https://blog.google/innovation-and-ai/products/dolphingemma/>
- [2] Robinson, D., Miron, M., Hagiwara, M., et al. (2024/2025). NatureLM-audio: An audio-language foundation model for bioacoustics. International Conference on Learning Representations (ICLR) 2025. arXiv:2411.07186.
- [3] Sutton, R. S. (2019-03-13). The bitter lesson. incompleteideas.net. <http://www.incompleteideas.net/InIdeas/BitterLesson.html>
- [4] Whitehead, H., Smith, T. D., & Rendell, L. (2021). Adaptation of sperm whales to open-boat whalers: Rapid social learning on a large scale? Biology Letters, 17(3), 20210030. <https://doi.org/10.1098/rsbl.2021.0030>
- [5] Bakker, K. (2022). The Sounds of Life: How Digital Technology Is Bringing Us Closer to the Worlds of Animals and Plants. Princeton University Press.

第1章 Wild Dolphin Project與斑點海豚的世界

1.1 翡翠綠巴哈馬的長年觀察

大海中生活著人類還不完全了解的高智慧生命。海豚就是其中之一。想了解海豚的心思該怎麼做呢？有一種方法，是把海豚關在水槽裡觀察。但是，有人選擇了一條不同的路。她就是 Denise Herzing 博士。Herzing 博士親自前往海豚生活的大海 [1]。

Herzing 博士立下了一項原則。那就是「在牠們的世界裡，用牠們的方式」。這代表不隨人類的心意去指使海豚。這也代表在不傷害海豚的情況下觀察牠們。這種方式稱為非侵入性觀察。非侵入是指不戳刺身體、也不加以捕捉的方式。

如果把海豚關在水槽裡，很難看見牠們原本的模樣。因為被關起來的海豚，生活方式會改變。因此，親自前往大海觀察的方法格外珍貴。

Herzing 博士領導著「Wild Dolphin Project」。這項研究持續了將近40年。40年大約是小學生年齡的四倍，是一段很長的時間。研究地點是巴哈馬海域。這裡的海水就像游泳池的水一樣清澈透明。所以很適合同時記錄海豚的動作與聲音 [2][3]。

這項研究始於1985年。在水中追蹤海豚的研究當中，這是世界上歷時最長的一項。他們一年也不間斷地探訪同一個海豚群體。多虧如此，研究人員得以完整見證一隻海豚從出生、長大到生下寶寶的整個生命歷程。這樣累積下來的記錄，後來成為訓練人工智慧的重要基礎 [4]。

這片海域生活著大西洋斑點海豚。學名是 *Stenella frontalis*。學名是學者們在全世界通用的正式名稱。進行研究的地方是水淺的沙洲海域。因為水深很淺，陽光可以直接照到海底。因此就算從水面上方，也能清楚看見海豚的動作。

這種海豚的體型比瓶鼻海豚小。但牠們是非常聰明的動物。有一個數值，可以用來衡量大腦相較於體型有多大。這個數值稱為大腦化指數。它是把大腦大小與體型相比較後得出的數值。海豚家族的這個數值普遍偏高。在哺乳類動物中，僅次於人類，屬於數值最高的一群 [5]。

研究團隊成功分辨出每一隻海豚。因為牠們鰭上的傷疤各自不同。身上長出的斑點花紋也各有不同。斑點海豚在成長過程中，身上的斑點會增加。小時候沒有斑點，隨著年紀變大會越來越多。因此只要看斑點，就能大致推估年齡。研究團隊就像記住班上同學的長相一樣，記住了每一隻海豚。多虧如此，他們得以追蹤由祖母、母親到寶寶相連的三個世代。

長年的觀察之所以珍貴，是有原因的。海豚的習性會由母親傳承給寶寶。海豚寶寶會在身邊學習母親覓食的方法。像這樣學習到的行為由母親傳給後代的過程，稱為文化傳承。如果只觀察短暫的時間，無法察覺這種連結。唯有長期觀察同一個群體，才能真正看見。

海豚的社會就像人類社會一樣複雜。雄性海豚會結交一輩子的朋友，一起行動。雌性海豚則會長久照顧寶寶，並教導牠們規矩。要建立如此複雜的社會，彼此之間就必須有溝通的管道。這就是我們必須解開海豚聲音之謎的原因。

相關佐證資料

- [1] Georgia Tech Research. (2025-04-14). DolphinGemma: How Google AI is helping decode dolphin communication. Georgia Tech Research News. <https://research.gatech.edu/dolphingemma-how-google-ai-helping-decode-dolphin-communication>
- [2] Herzing, D. L., Pack, A. A., Delfour, F., Starner, T., Mason, C., Gilliland, S., Ramsey, C., & Kohlsdorf, D. (2024). Imitation of computer-generated sounds by wild Atlantic spotted dolphins (*Stenella frontalis*). *Animal Behavior and Cognition*, 11(2), 136–166. <https://doi.org/10.26451/abc.11.02.02.2024>
- [3] Google. (2025-04-14). DolphinGemma: How Google AI is helping decode dolphin communication. The Keyword. <https://blog.google/innovation-and-ai/products/dolphingemma/>

[4] Millman, R. (2025-04-15). Google launches AI that talks to dolphins. Newsweek.

<https://www.newsweek.com/dolphingemma-google-ai-dolphins-2059955>

[5] Marino, L. (1998). A comparison of encephalization between odontocete cetaceans and anthropoid primates. *Brain, Behavior and Evolution*, 51(4), 230–238. <https://doi.org/10.1159/000006540>

1.2 嚴酷的大海與 CHAT 系統

要收錄海豚的聲音並不容易。因為海豚的聲音中，有許多是人類聽不見的高音。人類的耳朵大約能聽到20千赫（kHz）。千赫是測量聲音音調高低的單位。海豚甚至能發出接近150千赫的高音。這遠遠高於人類能聽見的聲音[1]。

海豚會使用很多種聲音。最具代表性的大致有三種。一種是音調高低起伏的哨聲。另一種是喀噠作響的咔嗒聲（click）。最後一種是短促爆發的脈衝聲（burst pulse）。脈衝聲是細密連續爆發的聲音。海豚在打架或近距離交談時經常使用。咔嗒聲碰到物體後會反彈回來。海豚透過回音，就能知道眼前有什麼。這項能力稱為回聲定位。回聲定位就是靠聲音來判斷物體位置的能力。

當時出現了一個問題。只把一支麥克風放進水裡並沒有用。因為無法知道在那麼多隻海豚中，究竟是誰發出了聲音。於是，Herzing 博士去拜訪了一位工程技術專家。他就是 Georgia Tech 的 Thad Starner 博士。Starner 博士是一位研發將小型電腦佩戴在身上技術的專家。

Starner 博士原本以為這項工作很快就能完成。但是大海是個棘手的對手。光是讓儀器穩定運作，就花了整整4年。最大的敵人是海水。鹽水從縫隙滲進去，燒壞了機器 [2]。

裝著機器的鋁製外盒也帶來了麻煩。它吸收了強烈的陽光，導致內部溫度飆升。這樣一來電腦就會當機。鋁還會阻擋無線訊號。研究團隊將外盒背面換成透明塑膠。電線與電路也用特殊樹脂完全包覆。直到這時，機器才能在水中順暢運作。

把機器長時間放在水裡也是一項考驗。因為鹽分會一點一點侵蝕金屬。研究團隊每次使用後，都會清洗並仔細檢查機器。只要有一點微小的縫隙，就可能讓整台機器報銷。

過程中也有許多危險。潛水員佩戴的大型電池一旦進水，可能會爆炸。機器漏出的微弱電流，也可能把鯊魚吸引過來。因為鯊魚擁有能感知微弱電訊號的感官。實際上，在測試過程中真的曾有鯊魚出現過。

鯊魚之所以能感覺到微弱電流，是因為有一種特殊的感官。這個器官稱為羅倫氏壺腹。它在鯊魚吻端周圍形成許多小孔。因此，連獵物發出的微弱電訊號牠都能感知到。研究團隊正是擔心水中的儀器會刺激到這種感官。

歷經千辛萬苦研發出來的設備，就是 CHAT 系統。CHAT 這個名字背後有其含義。翻譯過來就是「鯨豚聽覺增強遙測裝置」。簡單來說，它是一台可以在水中傳送、接收並辨識聲音的穿戴式電腦。鯨豚類是海豚與鯨魚的總稱 [3]。

CHAT 是一台在水下聆聽並記錄海豚聲音的水中電腦。這台機器記錄聲音非常細密。它一秒鐘分割接收聲音超過19萬次。這相當於比人類的耳朵聽得更加細密。只有聽得夠細密，才不會漏掉超高音。但相對地，儲存空間也會很快被占滿。因此必須搭配大容量的儲存裝置一起使用。

潛水員將這台機器背在背上，手腕上戴著小鍵盤。聲音則是透過骨傳導耳機來聽。骨傳導是一種不經由耳朵，而是振動頭骨來傳遞聲音的方式。潛水員按下鍵盤，約定好的聲音就會傳入水中。這種聲音不是海豚原本發出的聲音。而是人類全新創造出來的聲音。為了不跟海豚原本的聲音搞混，才特地挑選了全新的聲音。另一位潛水員的機器能辨認出這個聲音。只要符合預先設定的聲音，就會播報該聲音所對應的物品名稱。

傳遞聲音時，系統會用說話聲來提示。潛水員送出聲音時，會聽到男生的聲音。而在接收到聲音的那一端，則會有女生的聲音讀出含義。透過區分兩種聲音，讓人不會搞混究竟是誰在說話。

相關佐證資料

- [1] Georgia Tech Research. (2025-04-14). DolphinGemma: How Google AI is helping decode dolphin communication. Georgia Tech Research News. <https://research.gatech.edu/dolphingemma-how-google-ai-helping-decode-dolphin-communication>
- [2] Herzing, D. L., Pack, A. A., Delfour, F., Starner, T., Mason, C., Gilliland, S., Ramsey, C., & Kohlsdorf, D. (2024). Imitation of computer-generated sounds by wild Atlantic spotted dolphins (*Stenella frontalis*). *Animal Behavior and Cognition*, 11(2), 136–166. <https://doi.org/10.26451/abc.11.02.02.2024>
- [3] Kohlsdorf, D., Gilliland, S., Presti, P., Starner, T., & Herzing, D. (2013). An underwater wearable computer for two way human-dolphin communication experimentation. *Proceedings of the 2013 International Symposium on Wearable Computers (ISWC)*, 147–148. <https://doi.org/10.1145/2493988.2494346>

1.3 一起玩耍中學習的溝通遊戲

就算把機器全做好了，依然面臨一個難題。那就是該如何教導海豚這些聲音。他們決定不採用強迫訓練的方式。相反地，選擇了讓海豚自己主動學習的方法 [1]。

Herzing 博士挑選的方法是有名字的。那就是「社會競爭示範模式」。這種方法是向海豚展示兩個人互相問答的模樣。Irene Pepperberg 博士在教導鸚鵡時也曾用過這套方法。只要展現兩個人互動競爭的樣子，動物就會在旁邊跟著學習。做法是這樣的：兩個人在海豚面前示範玩遊戲。一個人拿著東西，另一個人發出代表那個東西的聲音。發出聲音的人接過東西，兩人便一起玩耍。

實際的遊戲過程是這樣的。一名潛水員手裡拿著黃色圍巾。另一名潛水員用手腕上的鍵盤發出圍巾的聲音。接著，拿著圍巾的潛水員就把圍巾遞給發出聲音的潛水員。兩個人就在水中拉著圍巾開心地玩耍。

和海豚一起玩的玩具有好幾種。除了黃色圍巾，還會用漂在海裡的馬尾藻來玩。馬尾藻是一種漂浮在海面上的褐色海藻。有時牠們也會拉著長繩玩耍。牠們也很喜歡在船頭激起的浪花中乘浪玩耍。研究人員為每種玩具都設定了不同的聲音來進行互動。

海豚在一旁靜靜看著這場遊戲。牠們察覺到了「發出聲音，玩具就會互相傳遞」的規則。不久之後，海豚開始跟著模仿那個聲音。牠們惟妙惟肖地模仿了人類製作的哨聲起伏。像這樣跟著發出相同聲音的行為稱為聲音模仿。學者們將其稱為聲學模仿。這些遊戲觀察，後來成為人工智慧研究的重要基石 [2][3]。

海豚多次主動模仿了這個聲音。有的聽了一次就立刻跟著模仿。有的則是在稍後才模仿。這證明了牠們能靈活模仿全新的聲音。不過，有一點需要注意。牠們究竟是否懂得那個聲音代表哪樣物品，目前尚未獲得證實。聲音模仿成功了。但還沒有確切的證據能證明牠們把聲音當成了物品的名字來使用。

海豚為什麼會主動發出聲音呢？這不只是為了獲得食物，而是因為想和人類一起玩耍。只要遊戲很有趣，海豚就會主動游過來。可以說是快樂帶動了彼此的溝通。

這項實驗揭示了一個重要的事實。那就是海豚能夠學習並模仿全新的聲音。這也說明我們不必強迫牠們學習人類的語言。可以說是海豚主動靠近，與人類互相交流聲音。

相關佐證資料

- [1] Georgia Tech Research. (2025-04-14). DolphinGemma: How Google AI is helping decode dolphin communication. Georgia Tech Research News. <https://research.gatech.edu/dolphingemma-how-google-ai-helping-decode-dolphin-communication>
- [2] Herzing, D. L., Pack, A. A., Delfour, F., Starner, T., Mason, C., Gilliland, S., Ramsey, C., & Kohlsdorf, D. (2024). Imitation of computer-generated sounds by wild Atlantic spotted dolphins (*Stenella frontalis*). *Animal Behavior and Cognition*, 11(2), 136–166. <https://doi.org/10.26451/abc.11.02.02.2024>
- [3] Google. (2025-04-14). DolphinGemma: How Google AI is helping decode dolphin communication. The Keyword. <https://blog.google/innovation-and-ai/products/dolphingemma/>

1.4 DolphinGemma 繪製的聲音地圖

人類所設定的聲音規則有其極限。因為海豚在群體中真正交流的聲音要複雜得多。要解開這種複雜的聲音，需要強大的工具。那個工具就是人工智慧 [1]。

2025年傳出了一個重大消息。Google DeepMind 與 Wild Dolphin Project 共同推出了一款人工智慧。Google DeepMind 是 Google 的人工智慧研究機構。這款人工智慧的名字叫做「DolphinGemma」。DolphinGemma 是以 Google 開發的名為 Gemma 的小型人工智慧為基礎。研發團隊將40年來累積的海豚聲音和影像輸入進去進行訓練 [2]。

以前的設備是背在背上的沉重背包電腦。潛水員必須背著沉重的儀器進入大海。現在只要手掌大小的智慧型手機就能勝任這項工作。裝備大幅縮小了。將 DolphinGemma 打造得如此小巧是有原因的。這款人工智慧在學習時使用的調整數值大約有4億個。這些調整數值稱為參數。就大型人工智慧而言，這算是相當精巧的規模。因此在手掌中的智慧型手機上也能運作。研究團隊將這款人工智慧放進 Google Pixel 手機中。不用把笨重的電腦留在船上，就能直接在海上使用 [4]。

DolphinGemma 這個名字也有其意義。它是在代表海豚的 Dolphin 後面加上了 Gemma。這代表以 Gemma 人工智慧為基礎，加入了海豚的聲音所打造出來。

最困難的事情是把聲音轉換成數字。水中的聲音是連續不斷流動的。人工智慧則比較喜歡切成一小段一小段的碎片資料。因此研究團隊把聲音細細切開。就像把文字拆成子音和母音一樣，把聲音切成細碎的小單位。這種聲音碎片稱為聲學標記（acoustic token）。意思就是把聲音細分而成的小碎片。

分割聲音時也有專門的工具。那就是 Google 名為 SoundStream 的聲音切割技術。這項技術能把水中的聲音細細切開。DolphinGemma 接著學習這些切片。

利用這些聲音碎片，DolphinGemma 練習了預測下一個聲音。也就是看著前面的聲音，來推測後面會接什麼聲音的練習。結果，DolphinGemma 變得能精準猜中接下來要出現的聲音。這就像人類懂得語言的規則後，就能猜出下一句話一樣。

能猜中下一個聲音，代表掌握了其中的規則。人工智慧就是從海豚的聲音中找到了這樣的規則。

DolphinGemma 還會自己把相似的聲音聚集在一起。人工智慧把聲音轉換成點，像地圖一樣排列開來。相似的聲音排得比較近，不同的聲音則放得比較遠。研究團隊長年收集了各種情況下的聲音。例如海豚媽媽和小海豚互相呼喚的簽名哨聲。簽名哨聲是每隻海豚都不同、就像名字一樣的聲音。也有打架時發出的爆發脈衝聲。還有尋找伴侶或追趕鯊魚時發出的咔嗒聲。把這些聲音與當時的情況配對起來有好處。這能幫助人工智慧更快找到聲音的線索。

Herzing 博士這樣解釋這項工作。她說如果由人類親手檢查完所有的資料，需要花上150年。她接著補充，人工智慧能以快得多的速度完成這項工作。150年比一個人活完兩輩子還要長 [3]。

相關佐證資料

[1] Google. (2025-04-14). DolphinGemma: How Google AI is helping decode dolphin communication. The Keyword. <https://blog.google/innovation-and-ai/products/dolphingemma/>

[2] Google DeepMind. (2025). DolphinGemma. Google DeepMind. <https://deepmind.google/models/gemma/dolphingemma/>

[3] Scientific American. (2025). DolphinGemma could enable AI communication with dolphins [Podcast]. Scientific American. <https://www.scientificamerican.com/podcast/episode/dolphingemma-could-enable-ai-communication-with-dolphins/>

[4] Lanz, J. A. (2025-04-14). Google launches AI to decode dolphin talk, runs on Pixel phones. Decrypt. <https://decrypt.co/314766/google-launches-ai-decode-dolphin-talk>

1.5 在大海展開的真正對話

2025年夏天展開了全新的一幕。Wild Dolphin Project 開始在海洋現場使用 DolphinGemma。在這之前，它只在電腦裡進行學習。現在潛水員直接把這款人工智慧帶進了水中。研究團隊使用了最新推出的 Google Pixel 手機 [1][9]。

在大海中使用的方法是這樣的。潛水員拿著裝在防水盒裡的 Pixel 手機潛入海中。人工智慧會在水中直接分析聲音。接著找出重複出現的聲音模式，並告訴研究團隊。不必等回到船上才檢查資料，在水裡就能立刻確認。

這次的海洋研究有一個特別之處。DolphinGemma 自己創造出了像海豚一樣的全新聲音。那是人類過去使用的程式很難製造出來的聲音。研究團隊計劃未來把這種聲音用在與海豚互相交流的研究上。也就是想試著用全新的聲音向海豚搭話 [2]。

這項技術使用得非常謹慎。因為如果人類隨意發出聲音，可能會擾亂海豚的社會。長久以來，海豚群體都是靠著聲音緊密連結在一起。因此研究團隊努力不去干擾海豚原有的聲音。教給海豚的聲音，也特地挑選了大自然中不存在的全新哨聲。這是為了不讓牠們跟原本的聲音搞混。研究團隊也努力不打擾海豚互相交流聲音的節奏。他們會等待海豚停下說話、安靜下來的空檔。這就像在排隊領營養午餐時依序輪流一樣，是一種禮貌。

並不是只有海豚接受這樣的研究。還有一個名為 Earth Species Project 的組織。這個組織利用人工智慧來分析各種動物的聲音。鳥類、鯨魚和大象等截然不同的動物都是研究對象。因為每種動物發出聲音的原因都不一樣。比較多種動物之後，聲音的規則就會看得更加清楚。2026年，他們還開放了一個能共享動物聲音資料的全新資料庫。名字叫做 alp-data。這是取自代表「動物語言處理」的英文首字母縮寫。建立它是為了讓許多研究人員能夠共享資料 [3][4]。

共享資料有好處。許多國家的研究人員把動物的聲音收集在同一個地方。收集的資料越多，人工智慧就學得越好。這個資料庫成為了許多動物聲音研究者共同使用的基礎。

也可以用同一個人工智慧同時處理多種動物。因為處理聲音的方法在不同動物之間很相似。從某種動物身上學到的方法，可以用在另一種動物身上。因此各項研究之間能夠互相提供幫助。

解開抹香鯨之謎的研究也在向前邁進。名為 Project CETI 的組織正在進行這項工作。抹香鯨會把好幾個喀噠作響的咔嗒聲組合在一起發出。這一組聲音稱為 coda。研究團隊在 coda 當中找到了聲音的規則。這個規則被稱為抹香鯨的聲音字母表。字母表這個詞是一種比喻。它指的不是人類的文字，而是把咔嗒聲組合的規則比喻成了字母表 [5]。

2026年有了一項全新發現。研究人員在抹香鯨的 coda 中，發現了與人類語言母音相似的規律。人類的語言中有像「啊」、「依」這樣的母音。抹香鯨的 coda 也依照聲音的性質分成了兩種類型。coda 的長度與排列順序中也存在著規則。這代表抹香鯨也擁有屬於自己的聲音架構 [6][7]。

這項研究未來也會謹慎地繼續進行。即使了解了動物的語言，也不代表懂得牠們的一切。因為人類與動物生活在不同的世界裡。儘管如此，傾聽牠們的聲音依然非常有意義。只要了解彼此，就更容易生活在一起 [8]。

這所有的研究都指向同一個方向。這不是強迫動物學習人類的語言。而是如實聆聽動物原原本本的聲音。有些人認為，這項研究對動物的權利也能有所幫助。因為只要知道動物在想什麼，就能更好地保護牠們。這場從大海展開的對話，正在成為連結人類與大自然的新橋樑。

相關佐證資料

[1] Google. (2025-04-14). DolphinGemma: How Google AI is helping decode dolphin communication. The Keyword.

<https://blog.google/innovation-and-ai/products/dolphingemma/>

[2] Scientific American. (2025). DolphinGemma could enable AI communication with dolphins [Podcast]. Scientific

American. <https://www.scientificamerican.com/podcast/episode/dolphingemma-could-enable-ai-communication-with-dolphins/>

[3] Earth Species Project. (2026-04-02). Animal language processing: An AI convergence in animal communication. Earth Species Project.

<https://earthspecies.org/2026/04/02/animal-language-processing-an-ai-convergence-in-animal-communication/>

- [4] Earth Species Project. (2026-07-22). Introducing alp-data: A shared data layer for animal language processing. Earth Species Project.
<https://earthspecies.org/2026/07/22/introducing-alp-data-a-shared-data-layer-for-animal-language-processing/>
- [5] Project CETI. (2024). Sperm whale phonetic alphabet proposed for the first time. Project CETI Blog.
<https://www.projectceti.org/blog-posts/sperm-whale-phonetic-alphabet-proposed-for-the-first-time>
- [6] Beguš, G., Dąbkowski, M., Sprouse, R. L., Gruber, D. F., & Gero, S. (2026). The phonology of sperm whale coda vowels. Proceedings of the Royal Society B, 293(2069), 20252994. <https://doi.org/10.1098/rspb.2025.2994>
- [7] National Geographic. (2026). We're one step closer to understanding the sperm whale 'alphabet'. National Geographic.
<https://www.nationalgeographic.com/animals/article/sperm-whale-vowel-patterns-communication-discovery>
- [8] Boing Boing. (2026-06-11). Decoding whale language could unlock new legal rights for them. Boing Boing.
<https://boingboing.net/2026/06/11/decoding-whale-language-could-unlock-new-legal-rights-for-them.html>
- [9] Schwartz, E. H. (2025-04-15). Google's new AI model could someday let you understand and talk to dolphins. TechRadar. <https://www.techradar.com/computing/artificial-intelligence/googles-new-ai-model-could-someday-let-you-understand-and-talk-to-dolphins>

第2章 抹香鯨的音素字母與 Project CETI

2.1 破解密碼的人與深海巨人相遇

地球上大腦最大的動物不是人類。而是抹香鯨。抹香鯨的學名是 *Physeter macrocephalus*。牠是生活在深海裡非常巨大的鯨魚。體長最長可以達到 20 公尺。這相當於把兩輛大公車連在一起的長度。體重可達 50 公噸。差不多和十頭大象加起來一樣重 [1][2]。

這種鯨魚的大腦大約有 8 到 9 公斤。人類的大腦大約只有 1.4 公斤。所以抹香鯨的大腦大約是人類大腦的 6 倍大。這個巨大的大腦，是用來處理群體生活與情感的。抹香鯨一生中有 80% 的時間，都在又深又暗的海底度過。在那麼深的地方，陽光照不進去。因此抹香鯨用聲音代替眼睛來看世界。

抹香鯨會發出聲音，並聆聽反射回來的回音。這是聲音撞到物體後，像回聲一樣傳回來。牠們藉由聽這個回聲，就能知道物體的位置和距離。這種方法叫做回聲定位（echolocation）。意思是利用聲音來觀察周圍環境的方法。這時使用的喀答聲，是用來尋找方向和測量獵物距離的。但是，抹香鯨在對話時會使用不同的聲音。牠們會按照固定的節拍，把簡短的喀答聲組合好幾次一起發出。這一組聲音就叫做 coda。coda 是鯨魚彼此交談時所使用的聲音片段 [3][5]。

長期以來，學者們並沒有把 coda 當作很特別的東西。只認為那是一種簡短的信號。然而在 2017 年，美國哈佛大學拉德克利夫高等研究院萌生了一個新想法。研究海洋生物的 David Gruber 博士錄下了鯨魚的聲音。正好有人偶然聽到了那個聲音。那就是美國 MIT 的密碼學研究者 Shafi Goldwasser 教授。密碼是為了隱藏秘密而設計、具有規律的信號。Goldwasser 教授在鯨魚的聲音中感受到了規律。那聽起來很像人類製造的密碼。於是她提議用電腦來解開那個規律 [4]。

這個想法與一位長期觀察鯨魚的學者連結了起來。他就是 Shane Gero 博士。Gero 博士在加勒比海多米尼克沿海，追蹤鯨魚家族超過了 13 年。到了 2020 年，Project CETI (Cetacean Translation Initiative) 正式成立。中文的意思是鯨魚語言翻譯計畫。來自許多國家的學者齊聚一堂。他們在多米尼克海域建立了研究基地。在深達海下 1,000 公尺處安裝了水下麥克風。還在鯨魚身上安裝了測量聲音與位置的小型儀器。這台儀器是由 Project CETI 的工程師們製造的 [6][7][8]。

這台小巧的儀器非常特別。裡面裝了三個麥克風。因為有多個麥克風，所以在吵雜的深海中也能辨別聲音。即使許多鯨魚一起發出叫聲，也能挑出其中一隻的聲音。這台儀器不只記錄聲音。還會同時測量鯨魚所在的深度與動作。水溫和位置也會一起記錄下來。因此可以完全知道牠們是在什麼時候、什麼地方發出了什麼樣的聲音。研究團隊用電腦分析了這樣收集到的 coda。分析的 coda 數量高達 8,719 個。這些聲音組合，成了破解鯨魚語言的關鍵鑰匙 [9]。

相關佐證資料

- [1] Whitehead, H. (2003). *Sperm Whales: Social Evolution in the Ocean*. University of Chicago Press.
- [2] Project CETI. (2024). About Sperm Whales. Project CETI. <https://www.projectceti.org>
- [3] Madsen, P. T., Wahlberg, M., & Møhl, B. (2002). Male sperm whale acoustics in a high-latitude habitat: implications for echolocation and communication. *Behavioral Ecology and Sociobiology*, 53(1), 31-41. <https://doi.org/10.1007/s00265-002-0548-1>
- [4] Radcliffe Institute. (2024). The CETI Moment. Harvard Radcliffe Institute. <https://www.radcliffe.harvard.edu/radcliffe-25/ceti-moment>
- [5] Andrews, R. (2024-05-07). Scientists Discover a 'Phonetic Alphabet' Used by Sperm Whales. *Smithsonian Magazine*. <https://www.smithsonianmag.com/smart-news/scientists-discover-a-phonetic-alphabet-used-by-sperm-whales-moving-one-step-closer-to-decoding-their-chatter-180984326/>
- [6] Gero, S., Whitehead, H., & Rendell, L. (2016). Individual, unit and vocal clan level identity cues in sperm whale codas. *Royal Society Open Science*, 3(1), 150372. <https://doi.org/10.1098/rsos.150372>
- [7] Project CETI. (2020). Cetacean Translation Initiative. Project CETI. <https://www.projectceti.org>
- [8] National Geographic Society. (2024). Project CETI. National Geographic Society. <https://www.nationalgeographic.org/explore/programs/project-ceti>

[9] Sharma, P., Gero, S., Payne, R., Gruber, D. F., Rus, D., Torralba, A., & Andreas, J. (2024). Contextual and combinatorial structure in sperm whale vocalisations. *Nature Communications*, 15, 3617.
<https://doi.org/10.1038/s41467-024-47221-8>

2.2 超越紅綠燈的聲音：coda 的四個要素

過去的教科書對抹香鯨聲音的看法很狹隘。上面寫著加勒比海的鯨魚只使用 21 種 coda。認為一個 coda 只傳達一種意思。就像交通紅綠燈一樣。例如紅燈停、綠燈行這種方式。然而，當 Project CETI 的人工智慧仔細檢視這些聲音時，這個想法被打破了 [1]。

人工智慧是人類製作的電腦程式。它能在大量資料中自己找出隱藏的規律。這套人工智慧在鯨魚的對話中，找到了 156 種 coda。這比過去所知道的 21 種多了 7 倍以上。而且，每一個 coda 裡面還隱藏著好幾個片段。鯨魚會變換四種要素，發出不同的聲音。

第一種是速度（tempo）。指的是喀答聲發出的速度有多快。人類心急時說話也會變快。鯨魚也會根據情況改變聲音的速度。第二種是節奏（rhythm）。指的是喀答聲之間的間隔。聲音密集或稀疏，代表的信號就有所不同。第三種是裝飾音（ornamentation）。這是在一組聲音的最後，多加上一個喀答聲。就像在句子末尾加上問號一樣，能稍微改變感覺。第四種是彈性速度（rubato）。就像在音樂課上把拍子稍微放慢或加快一樣，微調聲音的長度。鯨魚會根據對話的流向來調節這個長度 [3][4]。

這四種要素結合在一起，組成了 156 種 coda。這些 coda 依照時間順序連接，形成了更長的信號。單獨一個小小的聲音片段並沒有意義。但是把那些片段用各種方式組合起來，就會產生無數種信號。人類的語言也是這樣構成的。本身沒有意義的子音與母音聚在一起，就變成了詞彙與句子。研究團隊將構成 coda 的這些小小聲音片段，比喻為人類的子音和母音，稱之為音素字母。在語言學中，這種特徵被稱為雙重分節性（duality of patterning）。過去人們一直認為這種特徵只存在於人類的語言中。然而研究發現，鯨魚的語言中可能也有類似的結構 [2]。

鯨魚似乎還會根據說話的對象來改變聲音。對幼鯨發出的聲音，和對同伴發出的聲音並不相同。研究人員推測，鯨魚會考慮聆聽的對象來挑選聲音。雖然這一點尚未完全證實。鯨魚有時還會等到對方的聲音結束為止。等到出現空檔時，才接著發出聲音。這就像在教室裡，等同學說完話之後自己才開口一樣。這種輪流遵守次序來回發聲的方式，叫做輪流發聲（turn-taking） [6]。

每個鯨魚群體使用的 coda 都略有不同。就像不同地區的人說話腔調不同一樣。學者把這種群體稱為氏族（clan）。氏族是共同使用相同 coda 的大型群體。同一氏族的鯨魚會發出彼此相似的聲音。幼鯨在成長過程中會學習群體的 coda。這就像人類向父母學習語言一樣。因此，鯨魚的 coda 不是與生俱來的，而是後天學習而來的 [5]。

相關佐證資料

- [1] Sharma, P., Gero, S., Payne, R., Gruber, D. F., Rus, D., Torralba, A., & Andreas, J. (2024). Contextual and combinatorial structure in sperm whale vocalisations. *Nature Communications*, 15, 3617.
<https://doi.org/10.1038/s41467-024-47221-8>
- [2] Project CETI. (2024). Sperm Whale Phonetic Alphabet Proposed for the First Time. Project CETI Blog.
<https://www.projectceti.org/blog-posts/sperm-whale-phonetic-alphabet-proposed-for-the-first-time>
- [3] Andrews, R. (2024-05-07). Scientists Discover a 'Phonetic Alphabet' Used by Sperm Whales. *Smithsonian Magazine*. <https://www.smithsonianmag.com/smart-news/scientists-discover-a-phonetic-alphabet-used-by-sperm-whales-moving-one-step-closer-to-decoding-their-chatter-180984326/>
- [4] Al Jazeera. (2024-05-08). Scientists discover sperm whale 'phonetic alphabet'. Al Jazeera.
<https://www.aljazeera.com/news/2024/5/8/scientists-discover-sperm-whale-phonetic-alphabet>
- [5] Gero, S., Whitehead, H., & Rendell, L. (2016). Individual, unit and vocal clan level identity cues in sperm whale codas. *Royal Society Open Science*, 3(1), 150372. <https://doi.org/10.1098/rsos.150372>
- [6] Schulz, T. M., Whitehead, H., Gero, S., & Rendell, L. (2008). Overlapping and matching of codas in vocal interactions between sperm whales: insights into communication function. *Animal Behaviour*, 76(6), 1977-1988.
<https://doi.org/10.1016/j.anbehav.2008.07.032>

2.3 在深海找到的母音：Beguš 教授的發現

長期以來，鯨魚聲音的研究只測量了一件事。那就是喀答聲之間的時間間隔。卻沒能觀察聲音的高低。有一位突破了這道障礙的人。那就是美國加州大學柏克萊分校（UC Berkeley）的 Gašper Beguš 教授。Beguš 教授把目光轉向了聲音的高低。聲音是由許多不同高低混合而成的。這種混合狀態稱為頻譜（spectrum）。陽光照在玻璃上時，會散開成彩虹的顏色。聲音也是像這樣按照高低展開呈現的 [1][3]。

Beguš 教授觀察了人工智慧模仿製作鯨魚聲音的過程。人工智慧不僅重視時間間隔，也非常看重聲音的高低。因此，教授把鯨魚的喀答聲放慢了幾十倍。轉換成適合人耳聆聽的速度。於是，在喀答聲中看見了全新的東西。那是一種與人類母音非常相似的聲音 [2][4]。

人類在發出母音時，某些高低的聲音會產生強烈的共鳴。這種強烈共鳴的高低特徵稱為共振峰（formant）。這很像對著空瓶子吹氣時，會發出嗡嗡共鳴聲一樣。Beguš 教授從鯨魚的聲音中區分出了兩種母音。一種是 a coda（a-coda）。它表現出與人類 a 母音相似的高低結構。另一種是 i coda（i-coda）。它表現出與人類 i 母音相似的高低結構。

鯨魚並沒有把這兩種母音模糊地混在一起。而是將 a 和 i 清晰分明地發出來。甚至還有從一個母音滑向另一個母音的聲音。對人類來說，這就像是雙母音。就像 ai 那樣由兩個聲音連續發出的聲音。在人耳聽來，那原本只像是敲擊木頭的聲音。但是把播放速度放慢之後，就變成了互相來回傳遞 a 和 i 的對話。

這項發現的背後，有一套特別的人工智慧。這是一套人類能夠檢視它為什麼會得出這種答案的人工智慧。當這套人工智慧在處理鯨魚聲音的高低時，內部的一個部分只對某種特定的聲音產生格外強烈的反應。就像只有某個特定聲音出現時，開關才會被打開一樣。人類長期忽略了一件事。那就是鯨魚聲音中隱藏著類似母音的高低變化。人工智慧搶先發現了這個事實。

人工智慧透過自己模仿製作鯨魚的聲音來進行學習。即使人類沒有告訴它正確答案，它也能自己掌握。這種自己學習的方式稱為非監督式學習。人工智慧在製作聲音的過程中，自己挑選了哪一部分最重要。它所選出來的，正是聲音的高低。2026 年 4 月，這項研究發表在英國皇家學會的學術期刊上。內容指出 a-coda 與 i-coda 就像人類母音一樣遵循著規律。

鯨魚的母音中，還有更多與人類語言相似的地方。來看看加在一組聲音結尾的喀答聲。這個聲音會變得與旁邊緊接著的母音相似。人類在說話時，前後的發音也會互相影響。就像韓語裡的「국물」（湯）實際上會唸成「궁물」一樣。這是因為前後的聲音會變得互相相似。鯨魚的聲音中也發生了同樣的事情。

相關佐證資料

- [1] Beguš, G., Dąbkowski, M., Sprouse, R. L., Gruber, D. F., & Gero, S. (2026). The phonology of sperm whale coda vowels. *Proceedings of the Royal Society B*, 293(2069), 20252994. <https://doi.org/10.1098/rspb.2025.2994>
- [2] Beguš, G., Sprouse, R. L., Leban, A., Silva, M., & Gero, S. (2025). Vowel- and Diphthong-Like Spectral Patterns in Sperm Whale Codas. *Open Mind*, 9, 1849-1874. <https://doi.org/10.1162/OPMI.a.252>
- [3] University of California, Berkeley. (2024). Gašper Beguš, Speech and Computation Lab. UC Berkeley Department of Linguistics. <https://linguistics.berkeley.edu>
- [4] Zimmer, K. (2026-04-01). Sperm whale clicks follow similar rules to human speech. *Phys.org*. <https://phys.org/news/2026-04-sperm-whale-clicks-similar-human.html>

2.4 在鼻道製造母音的身體

人類與抹香鯨走過了不同的道路。在漫長的歲月中各自獨立演化。然而兩者都能發出像 a 和 i 這樣的母音。這種現象稱為趨同演化（convergent evolution）。這是指不同的生物面對相同的問題時，找到了相同的答案。就像鳥類和蝙蝠雖然是各自演化，卻都擁有翅膀一樣。因為生存環境的物理法則相同，所以找到的答案也變得相似 [1][2]。

人類為了發出母音，會使用喉嚨的聲帶。聲帶是喉嚨內部最初產生聲音的皺褶。人類為了發出母音，讓喉部位置下降。喉部是位於喉嚨中央的共鳴箱，聲帶就在裡面。人類的嘴唇和舌頭也演化成能進行細微的動作。但是抹香

鯨並不像人類那樣藉由振動喉嚨的聲帶來發聲。相反地，鯨魚在頭部的鼻道中，發展出了自己獨特的發聲器官。因為牠們必須頂住深海巨大的水壓來發出聲音。

抹香鯨的發聲器官由兩個部分成對搭配而成。第一部分是聲唇（phonic lips）。它是位於鼻道內部的皺褶結構。作用和人類的聲帶相同。當空氣通過時會產生振動，發出最初的聲音。這就像用手捏住氣球吹口放氣時，會發出吱吱聲一樣。第二部分是油脂囊與氣囊。這是填滿鯨魚巨大額頭的器官。最初的聲音通過這些囊袋時，會經過多次碰撞反彈。在這過程中，某些高低的聲音會產生強烈共鳴。這就和人類口腔修飾聲音的作用一樣 [3]。

這兩個部分共同創造出聲音。最初發出聲音的地方稱為聲源（source）。修飾聲音的通道稱為濾波器（filter）。撥動吉他弦發出聲音的地方就是聲源。讓那個聲音產生優美共鳴的琴箱就是濾波器。這兩者相遇，決定了聲音的音質。這個原理稱為聲源-濾波器理論（source-filter theory）。抹香鯨即使不用聲帶，也靠自己實現了這個原理 [4]。

這個發聲器官還隱藏著另一副面貌。正是鯨魚用來發出柔和母音的那個器官。但同一個器官也能發出極其強烈的聲音。測量出來的強度大約有 230 分貝。在水下測量的動物聲音中，屬於最強大的一類。不過，水中的分貝和空氣中的分貝標準並不相同。因此不能直接拿這兩個數值來比較。關於這種聲音會傷害人類的說法，目前尚未得到證實。鯨魚不會把這種強大的力量對準彼此。而是用溫柔的聲音互相包容。牠們就這樣在漫長的歲月裡，守護著海洋中的友情 [5]。

相關佐證資料

- [1] Beguš, G., Dąbkowski, M., Sprouse, R. L., Gruber, D. F., & Gero, S. (2026). The phonology of sperm whale coda vowels. *Proceedings of the Royal Society B*, 293(2069), 20252994. <https://doi.org/10.1098/rspb.2025.2994>
- [2] Madsen, P. T., Wahlberg, M., & Møhl, B. (2002). Male sperm whale acoustics in a high-latitude habitat: implications for echolocation and communication. *Behavioral Ecology and Sociobiology*, 53(1), 31-41. <https://doi.org/10.1007/s00265-002-0548-1>
- [3] Cranford, T. W., Amundin, M., & Norris, K. S. (1996). Functional morphology and homology in the odontocete nasal complex. *Journal of Morphology*, 228(3), 223-285. [https://doi.org/10.1002/\(SICI\)1097-4687\(199606\)228:3<223::AID-JMOR1>3.0.CO;2-3](https://doi.org/10.1002/(SICI)1097-4687(199606)228:3<223::AID-JMOR1>3.0.CO;2-3)
- [4] Fant, G. (1960). *Acoustic Theory of Speech Production*. Mouton.
- [5] Møhl, B., Wahlberg, M., Madsen, P. T., Heerfordt, A., & Lund, A. (2003). The monopulsed nature of sperm whale clicks. *Journal of the Acoustical Society of America*, 114(2), 1143-1154. <https://doi.org/10.1121/1.1586258>

2.5 三千萬年的海洋圖書館與鯨魚的權利

Project CETI 與 Beguš 教授的研究，留給了我們深刻的思考。抹香鯨並不是只憑本能胡亂吼叫的群體。牠們是透過聲音緊密連結在海洋各處的群體。是共享著各自的文化與家族羈絆的存在 [1]。

展現這種緊密連結的研究在2026年3月發表了。這是Project CETI詳細記錄抹香鯨生產過程的研究。在生下鯨魚寶寶的現場，其他母鯨也聚在一起。牠們在身旁幫助鯨魚媽媽。連毫無血緣關係的陌生鯨魚也來幫忙。這是非靈長類動物協助生產、並用數據證實的罕見案例。這時候，鯨魚的聲音也變了。每個重要時刻，聲音的質感都會改變。其中還包含類似母音的聲音。這項研究再次證明，鯨魚是會互相幫助的生物 [2]。

加勒比海的鯨魚有一種常常使用的聲音。學者們將它稱為「1加1加3 coda」。牠由喀嗒兩聲、短暫停頓，接著快速的三聲所組成。研究團隊謹慎地解讀這個聲音的意思。他們認為這可能代表「我在這裡」的意思。這種聲音是加勒比海鯨群共同使用的文化方言。每個地區群體使用的方言都不一樣。另一方面，抹香鯨族群數千萬年來一直在適應海洋環境中生活。在非常漫長的歲月裡，牠們在漆黑的海底靠著聲音確認彼此的存在 [3]。

鯨魚的聲音也是通報危險的網路。1800年代，人類搭船出海捕鯨。當時倖存的幾隻鯨魚傳播了消息。牠們用聲音告訴彼此躲避船隻的方法。結果，捕鯨的成功率大幅下降。僅僅兩年半的時間，就減少了大約58%。這是一份證明鯨魚聲音是龐大資訊網路的歷史紀錄 [4]。

人工智慧揭開的鯨魚語言，帶來了新的問題。那就是我們該如何像對待人類一樣對待鯨魚。因此，主張賦予鯨魚法律人格的行動正日益擴大。法律人格是指像人一樣受到法律保護的資格。這代表將牠們視為享有權利的對象，

而不是物品。2024年3月，紐西蘭與太平洋島國的原住民領袖發表了一項宣言。這是一份承認鯨魚擁有法律人格的宣言。名稱是 He Whakaputanga Moana。不過，這份宣言並非具有法律約束力的強制規則。這項行動與 Project CETI 攜手合作。美國紐約大學法學院的 MOTH 計畫也一同參與。MOTH 是代表超越人類之生命的名字 [5][6]。

另一方面，人類製造的噪音阻礙了鯨魚的聲音。那是船隻螺旋槳的聲音與海洋探勘的噪音。這些噪音掩蓋了鯨魚彼此交流的聲音。這樣一來，小鯨魚就很難學會鯨群的語言。文化也就無法傳承給下一代。Earth Species Project 也為這類研究貢獻心力。他們開發出了能同時處理多種動物聲音的人工智慧。它的名字是 NatureLM-audio [7][8]。

我們該做的事很明確。我們不該隨意向鯨魚發射人造聲音，擾亂牠們的文化。相反地，我們應該傾聽牠們漫長歲月以來在海洋中累積的聲音。那是延續了數千萬年的海洋圖書館。靜靜地傾聽那段旋律，才是最優先的事。這才是人類與鯨魚並肩共存的道路。

相關佐證資料

- [1] Whitehead, H. (2003). *Sperm Whales: Social Evolution in the Ocean*. University of Chicago Press.
- [2] EurekAlert!. (2026-03). Studies documenting rare sperm whale birth and ancient cooperative care released. EurekAlert! / Project CETI. <https://www.eurekalert.org/news-releases/1120823>
- [3] National Geographic. (2026). Whales could one day be heard in court and in their own words. National Geographic. <https://www.nationalgeographic.com/animals/article/whale-communication-legal-personhood>
- [4] Whitehead, H., Smith, T. D., & Rendell, L. (2021). Adaptation of sperm whales to open-boat whalers: rapid social learning on a large scale? *Biology Letters*, 17(3), 20210030. <https://doi.org/10.1098/rsbl.2021.0030>
- [5] Rodríguez-Garavito, C. (2026). Animal Communication, AI, and the Future of Nonhuman Rights. NYU Law MOTH Program. <https://www.law.nyu.edu/news/ideas/cesar-rodriquez-garavito-animal-communication-ai>
- [6] Te Ao Māori News. (2024-03-28). Māori king and other indigenous Pacific leaders sign up to granting whales legal personhood. Te Ao Māori News. <https://www.teaonews.co.nz/2024/03/28/maori-king-and-other-indigenous-pacific-leaders-sign-up-to-granting-whales-legal-personhood/>
- [7] Erbe, C., Marley, S. A., Schoeman, R. P., Smith, J. N., Trigg, L. E., & Embling, C. B. (2019). The Effects of Ship Noise on Marine Mammals—A Review. *Frontiers in Marine Science*, 6, 606. <https://doi.org/10.3389/fmars.2019.00606>
- [8] Earth Species Project. (2025). Using AI to Understand Animal Communication. Earth Species Project. <https://www.earthspecies.org/blog/using-ai-to-understand-animal-communication>

第3章 與座頭鯨 Twain 的相遇

3.1 地球邁向外星智慧的預演

想像一下，有一個訊號從遙遠的宇宙傳來。那個訊號只是雜音，還是有人傳來的話語呢？雜音是沒有任何意義的混亂聲音。我們沒有字典可以解讀那些話。這等於是完全沒有能解開含義的鑰匙。

為了解開這個問題，有些人沒有看向天空，而是把目光轉向了大海。有一個研究所專門尋找來自宇宙的智慧訊號。它的名字是搜尋地外文明研究所(SETI Institute)。這家研究所和美國加州大學戴維斯分校(UC Davis)的研究團隊攜手合作。長期研究鯨魚聲音的學者們也加入了進來。他們一起成立了名為「Whale-SETI」的研究團隊 [1]。

帶領這項研究的是 Brenda McCowan 教授。SETI 研究所的 Laurance Doyle 博士也一起參與。座頭鯨專家 Fred Sharpe 博士也加入了團隊。他們抱持著同一個想法：在遇見宇宙中陌生的存在之前，先來了解地球上的鄰居。

那個鄰居就是鬚鯨的一種——座頭鯨(Megaptera novaeangliae)。座頭鯨是會唱長長歌曲的大型海洋動物。這些歌主要由雄性來唱。同一個群體裡的雄鯨會互相學習唱歌。歌曲就這樣傳播開來，並且一點一點地改變。座頭鯨每年穿過廣闊的大海，也會用聲音互相聯絡 [2]。

研究團隊制訂的計畫是這樣的。在阿拉斯加海域放下一艘小船。把船的引擎完全關掉。把麥克風和水下喇叭放入水中。然後播放一次鯨魚的聲音。有一種方法是重新播放錄好的聲音，觀察動物會有什麼反應。這種方法被稱為回放實驗(playback) [3]。

播放的聲音是「接觸呼叫聲」。那是鯨魚在群體中告知自己位置時發出的聲音。這聲音被稱為「whup call」。這不是警告危險的聲音，而是和平友善的聲音。換成人類的話，大概就像「我在這裡」的意思。

這個聲音是前一天才在那片海域錄下的真正鯨魚聲音。那個群體裡有一隻38歲的雌性座頭鯨。牠的名字叫「Twain」。Twain 是研究團隊觀察了很久的一隻鯨魚。

為什麼研究團隊會選擇鯨魚呢？鯨魚和人類一樣壽命很長。牠們成群生活在一起。靠著聲音互相聯絡生活。還會把學到的歌曲傳給其他鯨魚。這些特點都和人類很像。因此研究人員認為，鯨魚非常適合當作理解陌生智慧的練習對象。

解讀宇宙訊號是一件非常困難的事。因為發送訊號的一方和接收訊號的一方，事先沒有任何約定。如果有把兩種語言並列排在一起的字典，翻譯就會很容易。但是外星訊號根本沒有這樣的書。所以我們必須從零開始找出其中的規則。

實驗開始前，船上一片安靜。研究團隊屏住呼吸，注視著海面。誰也不知道 Twain 會不會回答。

相關佐證資料

- [1] SETI Institute. (2023-12-12). Whale-SETI: Groundbreaking encounter with humpback whales reveals potential for non-human intelligence communication. SETI Institute. <https://www.seti.org/news/whale-seti-groundbreaking-encounter-with-humpback-whales-reveals-potential-for-non-human-intelligence-communication/>
- [2] McCowan, B., Hubbard, J., Walker, L., Sharpe, F., Frediani, J., & Doyle, L. (2023). Interactive bioacoustic playback as a tool for detecting and exploring nonhuman intelligence: "Conversing" with an Alaskan humpback whale. PeerJ, 11, e16349. <https://doi.org/10.7717/peerj.16349>
- [3] Ralls, E. (2026-04-29). AI helps humans have a 20-minute 'conversation' with a humpback whale named Twain. Earth.com. <https://www.earth.com/animals/whale-seti-ai-helps-humans-have-20-minute-conversation-with-humpback-whale-twain/>

3.2 與 Twain 的20分鐘

水下喇叭傳出了第一個聲音。沒過多久，Twain 出現了。Twain 悄悄地朝著船靠近。牠在船的周圍繞了一大圈。接著朝著水下喇叭發出了自己的聲音。

Twain 沒有叫了一聲就離開。人類播放聲音，Twain 就用聲音回答。人類停下來，Twain 也跟著等待。當聲音再次響起時，牠又會再次回應。這樣的互動持續了20分鐘。20分鐘差不多是一部短篇卡通的長度。

研究團隊在20分鐘內播放了38次 whup call。Twain 則用 whup call 回應了36次。其中有33次是緊接著人類播放的聲音所做出的即時回應。剩下的3次則是人類停止播放聲音後發出的。人類和野生鯨魚像這樣長時間相互交流聲音，還是史上第一次。

Twain 連聲音的間隔時間都配合得恰到好處。如果人類把聲音的間隔拉長，Twain 也會把間隔拉長。如果間隔很短，Twain 也會很快做出回應。牠完全配合了對方的節奏。

帶領這項研究的 McCowan 博士表示：這可以看作是人類與鯨魚之間一次深具意義的聲音互動。只不過，這並不代表人類翻譯出了鯨魚語言的含義。研究團隊解讀，這是彼此遵守順序、互相分享聲音的過程。

Twain 並不是20分鐘內都開開心心地發出聲音。一開始，牠很積極地參與聲音的互動。但是當研究團隊反覆播放完全相同的 whup call 時，牠的反應變了。Twain 看起來越來越焦躁。牠呼吸變得頻繁，甚至拍打水面。最後，牠終於離開了那裡。牠展現了參與、焦躁和離開這三個階段。

這次相遇有著明顯的局限。實驗對象只有 Twain 這一隻鯨魚。而且實驗也是碰巧迎來的唯一一次機會。因此，不能單憑這個結果就斷定鯨魚是在與人對話。研究團隊也謹慎地指出了這一點。

儘管如此，這次相遇展現了回放實驗的力量。這是一種播放預先錄下的聲音並觀察其反應的方法。透過這種方法，可以近距離觀察野生動物的反應。研究團隊稱這是一種全新的研究工具。

相關佐證資料

- [1] SETI Institute. (2023-12-12). Whale-SETI: Groundbreaking encounter with humpback whales reveals potential for non-human intelligence communication. SETI Institute. <https://www.seti.org/news/whale-seti-groundbreaking-encounter-with-humpback-whales-reveals-potential-for-non-human-intelligence-communication/>
- [2] McCowan, B., Hubbard, J., Walker, L., Sharpe, F., Frediani, J., & Doyle, L. (2023). Interactive bioacoustic playback as a tool for detecting and exploring nonhuman intelligence: "Conversing" with an Alaskan humpback whale. PeerJ, 11, e16349. <https://doi.org/10.7717/peerj.16349>
- [3] Frediani, J. (2024). "Conversing" with a humpback whale named Twain. The Safina Center. <https://www.safinacenter.org/blog/conversing-with-a-humpback-whale-named-twain>

3.3 遵守輪流順序的對話

人們在相互交談時會遵守規則。大家會等對方把話說完。等對方停下來，自己才開始說話。這種輪流發聲的規則被稱為輪替發聲(turn-taking)。這就像大家排隊領營養午餐時遵守順序一樣。

當對方說完話時，會有片刻安靜的瞬間。那段短暫的空隙被稱為對話停頓(conversational pause)。有了這樣的停頓，彼此的話才不會重疊在一起。座頭鯨在野外也會遵守這個規則。

Twain 在這20分鐘的互動中從未打破這個規則。在人類的聲音播放期間，牠沒有插嘴。牠耐心等待聲音完全結束。牠挑選了安靜下來的那個瞬間。然後在那個空檔填入自己的聲音。

研究團隊測量了 Twain 回應的時間。觀察聲音結束後過了幾秒才出現回答。這項工作並沒有使用人工智慧。研究人員是利用錄下的聲音和統計數據，一秒一秒仔細計算的。那個時間間隔與人類的對話規則非常吻合。

這個結果告訴了我們一件事。那就是 Twain 並不是因為機器聲音太煩而亂叫。Twain 能分清對方的輪次和自己的輪次，再發出聲音。研究團隊認為，Twain 是在主動參與這場聲音的互動。

人類從嬰兒時期就開始學習這個規則。看著媽媽的眼睛，和媽媽相互應答發聲。在學會說話之前，就先學會了這個順序。這是構成對話基礎的核心規則。

對話中的停頓非常短暫。有時甚至比眨一下眼睛還要快。Twain 沒有錯過這短暫的空隙。牠在最準確的時刻發出了聲音。

這種精準度如果沒有平時的練習是很難做到的。Twain 在野外平時就是這樣與同伴交流聲音的。牠把這種熟悉的溝通方式，原封不動地用在了人類傳來的聲音上。即使對方不是真正的鯨魚，牠也遵守了這個規則。

相關佐證資料

- [1] McCowan, B., Hubbard, J., Walker, L., Sharpe, F., Frediani, J., & Doyle, L. (2023). Interactive bioacoustic playback as a tool for detecting and exploring nonhuman intelligence: "Conversing" with an Alaskan humpback whale. *PeerJ*, 11, e16349. <https://doi.org/10.7717/peerj.16349>
- [2] SETI Institute. (2023-12-12). Whale-SETI: Groundbreaking encounter with humpback whales reveals potential for non-human intelligence communication. SETI Institute. <https://www.seti.org/news/whale-seti-groundbreaking-encounter-with-humpback-whales-reveals-potential-for-non-human-intelligence-communication/>
- [3] Ralls, E. (2024-01-06). Scientists have a 20-minute 'conversation' with a humpback whale named Twain. *Earth.com*. <https://www.earth.com/news/scientists-have-20-minute-conversation-with-a-humpback-whale-named-twain/>

3.4 數學揭開的鯨魚語法

Doyle 博士嘗試用數字來解析鯨魚的歌聲。因為他認為鯨魚的歌聲絕不是隨便發出的聲音。因此他使用了名為資訊理論(information theory)的工具。資訊理論是一種用數字來衡量聲音或文字中蘊含多少規則的方法。

研究團隊嘗試將齊夫定律(Zipf's law)套用在鯨魚的歌聲上。齊夫定律揭示了語言中隱藏的一種規律。在人類的語言中，常用詞彙和罕見詞彙的數量會呈現一定的比例。例如韓語中的「은、는、이、가」出現得非常頻繁，而難懂的詞彙則偶爾才出現一次。在出現次數多與少之間，總是保持著一定的比例。

2025年，一項在鯨魚歌聲中發現這個規律的研究出爐了。Inbal Arnon 研究團隊長期觀察了座頭鯨的歌聲。那是花了8年在南太平洋收集到的歌聲。那些歌聲與齊夫定律非常吻合。這是在人類語言之外，首次明確發現這種規律。

這個結果代表什麼呢？這可能表示鯨魚的歌聲中蘊含著規則。但光憑這一點，並不能證明鯨魚的歌聲具有意義或文法。研究團隊也沒有說鯨魚的歌聲與人類的語言相同。他們反而認為，這可能更接近人類的音樂。

研究團隊也檢視了夏農熵(Shannon entropy)。夏農熵是一種用數字來衡量「下一個聲音有多難預測」的方法。如果聲音完全雜亂無章，這個數值就會很高。如果只是不斷重複相同的聲音，這個數值就會接近0。鯨魚的歌聲正好介於這兩個極端之間。它既遵守規則，又不會千篇一律。

這種結構在其他鯨魚的研究中也能看到。這裡我們不妨舉另一種鯨魚為例。有一個名為 Project CETI 的研究團隊。他們專門研究抹香鯨。抹香鯨和座頭鯨是不同的物種。抹香鯨會把短促的喀噠聲組合在一起發出來。這一串串的聲音組合被稱為 coda。

2024年，Project CETI 研究團隊揭開了 coda 的結構。抹香鯨會混合節奏、速度、彈性速度(rubato)和裝飾音這四種要素來發出聲音。彈性速度是指將節拍稍微拉長或縮短。裝飾音則是在聲音中多加一個喀噠聲。牠們靈活組合這四種要素，創造出非常多樣的聲音。原本以為只有人類才會使用這種組合方式，沒想到抹香鯨也是如此 [4]。

相關佐證資料

- [1] Doyle, L. R., & McCowan, B. (2011). Information theory, animal communication, and the search for extraterrestrial intelligence. *Acta Astronautica*, 68(3-4), 406-417. <https://www.sciencedirect.com/science/article/abs/pii/S0094576509005682>
- [2] Arnon, I., et al. (2025). Whale song shows language-like statistical structure. *Science*, 387. <https://doi.org/10.1126/science.adq7055>
- [3] Sharma, P., et al. (2024). Contextual and combinatorial structure in sperm whale vocalisations. *Nature Communications*, 15, 3617. <https://doi.org/10.1038/s41467-024-47221-8>
- [4] Project CETI. (2024-05-07). Sperm whale phonetic alphabet proposed for the first time. Project CETI. <https://www.projectceti.org/blog-posts/sperm-whale-phonetic-alphabet-proposed-for-the-first-time>

3.5 寄往宇宙的信

Whale-SETI 研究還有一個最終目標。那不僅僅是為了了解鯨魚。而是要打造出一套工具，能夠辨識未來可能從宇宙傳來的陌生訊號。因為我們需要一把尺，用來區分單純的雜音和蘊含意義的訊號。

有一些用數學來解析鯨魚歌聲的方法。例如齊夫定律或夏農熵等方法。未來可以把這些方法應用在宇宙訊號上。不過，目前還沒有直接應用在宇宙訊號上。研究人員正在探討這些方法在未來是否值得一試。

許多學者指出了人類的傲慢。那就是我們連生活在同一個地球上的鯨魚聲音都聽不懂。他們反問：既然如此，我們又怎麼能聽懂宇宙另一端傳來的訊號呢？這個問題讓人們變得謙虛起來。

很久以前，人們曾相信地球是宇宙的中心。但事實證明那個信念是錯誤的。認為「只有人類才能溝通心意」的想法，或許也是如此。在大海深處，也許正住著用聲音相互交流的鄰居。

長期以來，人類都只把鯨魚的歌聲當作雜音。甚至捕殺了無數的鯨魚。還用船隻和機器讓大海變得喧鬧不堪。即便如此，鯨魚也沒有停止相互交流聲音。從非常久遠的過去開始，牠們就一直向彼此傳遞聲音。

因此，與 Twain 的這20分鐘是一場預演。那是為了在遇見陌生智慧時，提前練習可以使用的方法。在向宇宙發送話語之前，先傾聽身邊的聲音。懂得好好傾聽的能力，就是翻譯的起點。

相關佐證資料

- [1] SETI Institute. (2023-12-12). Whale-SETI: Groundbreaking encounter with humpback whales reveals potential for non-human intelligence communication. SETI Institute. <https://www.seti.org/news/whale-seti-groundbreaking-encounter-with-humpback-whales-reveals-potential-for-non-human-intelligence-communication/>
- [2] McCowan, B., Hubbard, J., Walker, L., Sharpe, F., Frediani, J., & Doyle, L. (2023). Interactive bioacoustic playback as a tool for detecting and exploring nonhuman intelligence: "Conversing" with an Alaskan humpback whale. PeerJ, 11, e16349. <https://doi.org/10.7717/peerj.16349>
- [3] Ralls, E. (2024-01-06). Scientists have a 20-minute 'conversation' with a humpback whale named Twain. Earth.com. <https://www.earth.com/news/scientists-have-20-minute-conversation-with-a-humpback-whale-named-twain/>

3.6 人工智慧開啟的新大門

在與 Twain 相遇之後，相關研究正在迅速發展。因為人工智慧能讓我們更深入地分析鯨魚的聲音。光靠人類自己需要耗費數十年的工作，人工智慧能夠快速給予協助。

有一個名為 Earth Species Project 的組織。他們致力於開發能聽懂動物聲音的人工智慧。那個人工智慧的名字是 NatureLM-audio。這個工具只要聽聲音就能分辨出是哪種動物。也能區分聲音的種類。像這樣的工作，它現在已經做得很棒了。

這個人工智慧也具備一些實驗階段的能力。例如數出錄音當中有幾隻動物。或是辨識從未學習過的動物聲音。這些工作目前還不夠完美。大概只是比胡亂猜測稍微準確一點的水準。

在人工智慧中，有一種作為基礎的大型人工智慧。這種大型人工智慧被稱為基礎模型(foundation model)。這是一種單憑一個模型就能廣泛處理多種任務的人工智慧。Google 開發的 Perch 就是這樣的一個例子。

人工智慧協助鯨魚研究的故事，可以依照時間順序來看。首先是 2018 年的事。Google 與美國國家海洋暨大氣總署 (NOAA) 攜手合作。國家海洋暨大氣總署是研究海洋與天氣的美國機構。雙方一起檢視了 14 年間收集到的海底聲音。其長度超過了 187,000 小時。這是人類就算不休息一直聽，也大約需要 21 年的份量。人工智慧從這筆龐大的資料中，挑出了座頭鯨的歌聲。

接著是 2025 年的事。Google 重新改良了 Perch，推出了 Perch 2.0。Perch 最初是用鳥叫聲學習的人工智慧。但是它用在海底聲音上也十分合適。在辨識鯨魚聲音的工作上拿出了優異的成績。學習鳥鳴的人工智慧，也為鯨魚研究帶來了幫助。

人工智慧強大的力量在於速度。人類需要聽上幾個月的聲音，人工智慧只要幾個小時就能瀏覽完畢。研究因此變得更快。它還能找出人類錯過的微弱聲音。多虧了這一點，舊資料中浮現出了新的事實。

這裡有一點需要注意。人工智慧能夠辨識鯨魚的聲音，並不代表已經解開了鯨魚的語言。雖然找到了聲音裡的規則，但還沒有理解其中的意思。目前還沒有確認任何證據，能證明鯨魚語言的意思已被完全破解。人類與鯨魚之間的字典，才剛剛翻開第一頁而已。

Earth Species Project 不只研究鯨魚。它也研究烏鴉的叫聲。也研究大象的叫聲。也研究白鯨的聲音。它用同一種工具處理多種動物的聲音。未來還計畫向動物播放人類製作的聲音。這是為了觀察動物會如何回應那些聲音。與 Twain 的聲音互動就是其中的第一步。

最近也出現了揭示鯨魚歌聲結構的研究。這是 2026 年美國雪城大學 (Syracuse University) Julia Zeh 研究團隊的研究。他們在唱歌的雄性座頭鯨身上安裝了小型記錄裝置。這是用吸盤短暫附著的錄音機兼動作記錄器。這項研究與 Twain 的聲音互動是不同的主題。它的內容指出鯨魚在水中上下移動的動作與歌聲互相配合。游向深處時和浮上水面時所唱的歌並不一樣。歌聲原來是配合身體動作編排而成的 [5]。

因此，研究至今仍在繼續。人類與鯨魚之間的字典，正一頁一頁地被填補。Twain 的那 20 分鐘，開啟了這部字典的第一個句子。

相關佐證資料

- [1] Earth Species Project. (2024-11-11). Introducing NatureLM-audio: An audio-language foundation model for bioacoustics. Earth Species Project. <https://www.earthspecies.org/blog/introducing-naturelm-audio-an-audio-language-foundation-model-for-bioacoustics>
- [2] Robinson, D., Miron, M., Hagiwara, M., Weck, B., Keen, S., Alizadeh, M., Narula, G., Geist, M., & Pietquin, O. (2025). NatureLM-audio: An audio-language foundation model for bioacoustics. International Conference on Learning Representations (ICLR) 2025. <https://arxiv.org/abs/2411.07186>
- [3] Google Research. (2018-10). Acoustic detection of humpback whales using a convolutional neural network. Google Research Blog. <https://research.google/blog/acoustic-detection-of-humpback-whales-using-a-convolutional-neural-network/>
- [4] Google Research. (2026-02-09). How AI trained on birds is surfacing underwater mysteries. Google Research Blog. <https://research.google/blog/how-ai-trained-on-birds-is-surfacing-underwater-mysteries/>
- [5] Syracuse University. (2026-08-31). How a whale's dive shapes its song. Syracuse University News. <https://news.syr.edu/2026/08/31/how-a-whales-dive-shapes-its-song/>

第4章 深海交響樂團

4.1 在喧鬧的大海中找出聲音

大海並不是安靜的地方。水中是一個充滿聲音的廣闊空間。鯨魚的歌聲、海豚的哨聲和海浪聲不停地交織在一起。在水裡，聲音傳播的速度比在空氣中快得多。比在空氣中快了大約4.5倍。因此，鯨魚用耳朵比用眼睛更能了解世界。在遠洋中尋找彼此時，牠們會使用聲音。追捕獵物和躲避危險時，也同樣使用聲音 [1][2]。

聲音在水裡每秒可以傳播大約1,500公尺。比聲音在陸地上傳播的距離遠得多。所以鯨魚連非常遙遠的朋友的聲音也能聽見。

問題在於人類製造的噪音。大船的螺旋槳聲音震盪著水中。探測海底的聲納（Sonar）聲音也很大。還有探測海底資源時引爆的爆炸聲。這些噪音可以傳播到幾千公里之外。如今，輪船的聲音是海洋中傳播最廣的人類噪音。噪音太大時，鯨魚的聲音就會被掩蓋在其中。學者們把這種現象稱為聲音被遮蔽。對於依靠聲音生活的鯨魚來說，這是一大威脅。

噪音也會讓鯨魚的身體生病。2001年美國發生重大事故後，船隻航行暫時大幅減少。當時大海變得安靜了。Rosalind Rolland 研究團隊抓住了這個機會進行觀察。牠們收集並調查了生活在加拿大芬地灣的露脊鯨糞便。當大海變安靜時，鯨魚體內的壓力物質減少了。這就像人類在吵鬧的工地旁會心跳加快一樣。這是噪音會使鯨魚疲憊不堪的第一個證據 [3]。

在各種聲音交雜的地方，要單獨挑出一隻鯨魚的聲音非常困難。學者們把這種難題稱為「雞尾酒會問題」。這就像在擠滿人的房間裡，只單獨聽某一位朋友說話一樣。就像學校餐廳雖然很吵，但旁邊同學說的話依然聽得見一樣。以前的機器不太擅長處理這件事。當船隻的金屬噪音和鯨魚的聲音重疊時，機器常常會搞混。

人工智慧用全新的方式解決了這個問題。這是一項能從混雜的聲音中挑出目標聲音的技術。機器首先將聲音轉換為數字地圖。在這張地圖上，輪船的聲音和鯨魚的聲音會落在不同的位置。這就像五顏六色的彈珠混在一起，再按照顏色重新分類放好一樣。機器區分這些位置，只把鯨魚的聲音重新聚集起來。Earth Species Project 研究團隊開發出一種即使沒有乾淨的標準答案聲音，也能消除噪音的方法。多虧了這項技術，人類得以更清晰地再次聽見鯨魚的聲音 [4]。

相關佐證資料

- [1] Bakker, K. (2022). The sounds of life: How digital technology is bringing us closer to the worlds of animals and plants. Princeton University Press.
- [2] NOAA Fisheries. (2025). A whale's world of sound. National Oceanic and Atmospheric Administration. <https://www.fisheries.noaa.gov/feature-story/whales-world-sound>
- [3] Rolland, R. M., Parks, S. E., Hunt, K. E., et al. (2012). Evidence that ship noise increases stress in right whales. Proceedings of the Royal Society B, 279(1737), 2363-2368. <https://doi.org/10.1098/rspb.2011.2429>
- [4] Miron, M., Keen, S., Hoffman, B., Hagiwara, M., et al. (2024). Biodenoising: Animal vocalization denoising without access to clean data. Earth Species Project. arXiv:2410.03427. <https://arxiv.org/abs/2410.03427>

4.2 虎鯨家族的方言

排除了吵鬧的噪音後，學者們開始解讀藏在鯨魚聲音裡的秘密。那個秘密就是每個家族都各自不同的「方言」。加拿大西海岸有一個研究虎鯨的基地。名字叫做 OrcaLab。這裡的 Helena Symonds 和 Paul Spong 博士數十年來一直在記錄野生虎鯨的聲音。牠們不捕捉也不打擾虎鯨，只用水下麥克風悄悄收集聲音 [1]。

虎鯨是以母親為中心生活的動物。幼鯨一輩子都不會離開母親的群體。因此，每個家族都會形成代代相傳的獨特聲音。學者們將這種現象稱為母系家族共同使用的方言。這就像我們班和隔壁班講話的口氣有點不一樣。虎鯨憑藉這種方言，在廣闊的大海中認出自己的家人。這也有助於防止近親交配。

OrcaLab 收集聲音的時間非常長。光是錄音就接近2萬小時。人類很難把這些聲音全部聽完。於是學者們把這項工作交給了人工智慧。德國與加拿大研究團隊開發的程式名稱是 ORCA-SPOT。這個程式能從漫長的錄音中自動

找出虎鯨的聲音。與人類標註的標準答案相比，100次中答對了93次。原本需要花費好幾年的工作，機器只用了八天就完成了 [5]。

人工智慧在分類聲音時，會把重點放在關鍵部分。這就像人類在長篇文字中給重要的詞語畫底線一樣。機器就是這樣連細微的語調差異也能捕捉到。虎鯨寶寶聽著母親和群體的聲音學習方言。這就像人類小嬰兒模仿父母說話一樣。學者們看著人工智慧繪製的聲音地圖，還發現了新的事實。那就是同一個家族的聲音歷經世代更迭，也在一點一點改變。這就像長期使用的語言會隨著時代產生微小的改變一樣。這代表鯨魚的聲音不只是天生的本能。這是像人類語言一樣，經過學習、傳承並不斷演變的文化證據。

了解方言對於保護鯨魚也派得上用場。因為透過聲音就能知道哪一個家族生活在哪裡。可以觀察船隻頻繁往來的海域中有哪些家族。用聲音計算家族數量，就能迅速掌握數量是否減少。這類研究至今仍在持續。2026年，能自動搜尋並分類虎鯨聲音的新程式 orcaAI 問世了。這個程式學習的不是加拿大，而是冰島海域的虎鯨聲音。那裡的虎鯨以捕食鯡魚為生。即使人類不需要長時間聆聽，機器也能代為檢查漫長的錄音 [6]。

相關佐證資料

- [1] Bakker, K. (2022). The sounds of life: How digital technology is bringing us closer to the worlds of animals and plants. Princeton University Press.
- [5] Bergler, C., Schröter, H., Cheng, R. X., et al. (2019). ORCA-SPOT: An automatic killer whale sound detection toolkit using deep learning. Scientific Reports, 9, 10997. <https://doi.org/10.1038/s41598-019-47335-w>
- [6] Bonhoeffer, S., Selbmann, A., et al. (2026). orcaAI: A machine learning tool to detect and classify acoustic signals of killer whales in audio recordings. Marine Mammal Science, 42(1), e70083. <https://doi.org/10.1111/mms.70083>

4.3 重現冰層下白鯨的聲音

如果說虎鯨是在清澈的沿岸孕育方言，那麼也有生活在冰冷混濁水域的鯨魚。那就是生活在北極和加拿大聖羅倫斯河（St. Lawrence River）的白鯨（Beluga Whale）。因為全身雪白，牠們被稱為「北極的天使」。由於擅長發出像小鳥一樣清脆高亢的聲音，牠們也有「海洋金絲雀」的外號。聖羅倫斯河的白鯨數量目前正在大幅減少。這是工廠噪音與水質污染累積造成的。即使劃定了保護區，數量依舊沒有明顯增加，令人非常擔憂 [1][7]。

加拿大溫莎大學（University of Windsor）的 Jaclyn Aubin 研究員參與了這項工作。她決定解開白鯨聲音的秘密。這項工作並不容易。白鯨是喜歡群居且善於社交的動物。幾十隻聚在一起同時發出聲音。牠們會同時傾瀉出哨聲和急促的喀喀聲。各種聲音混雜在一起，人類很難分辨哪裡是起點、哪裡是終點。

研究團隊把這些聲音交給了人工智慧。首先，搜尋聲音的機器從長段錄音中只剪出白鯨的聲音。接著，另一台機器將那些聲音按種類進行分類。這時也使用了 Earth Species Project 開發的程式 AVES。AVES 是一個學習動物聲音的基礎模型。所謂基礎模型，是指預先學習大量資料，並廣泛應用於各種任務的人工智慧。即使人類不給予標準答案，AVES 也能自主學習聲音的結構 [8]。

將人工智慧分類的聲音與人類分類的結果進行對比，兩者幾乎完全相同。即使沒有人類插手，機器也能像專家一樣分類白鯨的聲音。這樣做有一個好處。人類需要聽上好幾天的漫長錄音，機器能迅速檢查完畢。可以持續記錄生活在河裡的白鯨何時何地發出叫聲。這些記錄成為保護日益減少的白鯨的重要基礎。

冰層破裂和碰撞的聲音是一大干擾。為了消除這種噪音，研究團隊搭配使用了專門降噪的人工智慧。這台機器能從嘈雜的聲音中只還原出白鯨的聲音。多虧了它，差點消失的白鯨聲音得以重新被聽見。就這樣，人工智慧守護了埋在冰層下的聲音。

相關佐證資料

- [1] Bakker, K. (2022). The sounds of life: How digital technology is bringing us closer to the worlds of animals and plants. Princeton University Press.
- [7] Cotillard, T., Aubin, J., Vergara, V., et al. (2024). Automatic detection and classification of beluga whale calls in the St. Lawrence estuary. The Journal of the Acoustical Society of America, 156(6), 3723-3740. <https://doi.org/10.1121/10.0030472>

[8] Hagiwara, M. (2023). AVES: Animal vocalization encoder based on self-supervision. ICASSP 2023 - 2023 IEEE International Conference on Acoustics, Speech and Signal Processing. <https://doi.org/10.1109/ICASSP49357.2023.10095642>

4.4 學習鳥鳴的 AI 聆聽海洋

明明在聊大海，為什麼突然提到鳥呢？這是有原因的。因為學習過鳥鳴的人工智慧，在破解海洋動物聲音時發揮了巨大的力量。這項研究最初源自 Google。是由電腦科學家 Tom Denton 博士帶領的研究團隊所推動的「Perch」計畫。這個團隊的夢想很遠大。那就是將地球上發出的各種大自然聲音統整在一起 [10]。

Perch 團隊用鳥鳴來訓練人工智慧。讓它學習世界上超過1萬種鳥類的歌聲。光是學習的時間就長達數十萬小時。這樣打造出的人工智慧會將聲音轉換為數字座標來整理。可以想像成是為每種聲音安排位置，畫出一張寬廣的地圖。

然而，意想不到的事情發生了。學習鳥鳴的人工智慧，竟然也能很好地分類水下動物的聲音。這代表從一種動物身上學到的規則，也能適用於其他動物的聲音。鳥在空氣中鳴叫，而鯨魚在水裡發聲。生活環境和聲音都截然不同。即使如此，從鳥鳴學會的地圖竟然也完全符合鯨魚的聲音。

研究團隊用青蛙驗證了這種能力。當時使用的資料是 AnuraSet。這是收集了42種巴西青蛙叫聲的資料。錄音地點不是亞馬遜，而是在塞拉多和大西洋森林這兩處森林。這份資料不是 Google 製作的，而是由 Cañas 研究團隊建立的。Google 借用這份公開資料測試了人工智慧。雖然只展示了幾種青蛙叫聲，人工智慧的答對率卻超過了95%。這多虧了只看少數範例就能學會新任務的方法。這種能力並不局限於青蛙。面對鯨魚的喀嗒聲和海豚的哨聲，也同樣行得通 [9]。

這項研究還在持續發展。2025年8月，Google 推出了 Perch 的新版本（Perch 2.0）。這個人工智慧不僅學習鳥類，也學習了多種陸地動物的聲音。學習的動物超過1萬4千種。除了鳥類，還學習了哺乳動物、青蛙和昆蟲的聲音。但是，牠幾乎沒有學習過海底的聲音。儘管如此，2026年2月 Google 宣布，這個人工智慧在鯨魚聲音研究中取得了優異的成績。在多項鯨魚聲音測試中，這個人工智慧的表現不是第一名就是第二名。Google 還專門為海豚打造了人工智慧 DolphinGemma。這個程式體積很小，在智慧型手機上也能運行。牠學習了 Wild Dolphin Project 長期收集的聲音。這並不代表已經解開了海豚語言的涵義。它是用來尋找聲音結構與規則的研究工具 [11]。

DolphinGemma 不僅能分類聲音。牠還能模仿海豚的聲音，自主產生全新的聲音。將這些聲音在海裡播放給海豚聽的實驗也在持續進行。DolphinGemma 學習的聲音是經過長年收集而來的。Wild Dolphin Project 從1985年就開始收集聲音。這是生活在巴哈馬海域的大西洋斑點海豚的聲音。牠們將交配或爭鬥等行為與聲音一起記錄了下來。Google 將這個程式公開，讓任何人都能使用。

相關佐證資料

[9] Cañas, J. S., Toledo, L. F., et al. (2023). AnuraSet: A dataset for benchmarking Neotropical anuran calls identification in passive acoustic monitoring. PeerJ Computer Science, 9, e1575. <https://doi.org/10.7717/peerj-cs.1575>

[10] Google Research. (2026, February 9). How AI trained on birds is surfacing underwater mysteries. Google Research Blog. <https://research.google/blog/how-ai-trained-on-birds-is-surfacing-underwater-mysteries/>

[11] Google. (2025, April 14). DolphinGemma: How Google AI is helping decode dolphin communication. Google Blog. <https://blog.google/innovation-and-ai/products/dolphingemma/>

4.5 守護泥水中粉紅河豚的小型 AI

水中人工智慧不僅在廣闊的大海中運作，也在河流中工作。那就是在肉眼什麼也看不見的亞馬遜混濁泥水裡。這裡生活著擁有粉紅色身軀的淡水豚 Boto。巴西的 Mamirauá Institute 和巴西國家亞馬遜研究所（INPA）帶領著保護 Boto 的工作 [12]。

Boto 的數量目前正在迅速減少。牠們會被人類丟棄在河裡的漁網纏住。也會受到採金礦時流出的水銀污染。還飽受開發噪音的折磨。亞馬遜河因為泥土和樹葉很多，水質總是混濁不清。因此用相機或無人機根本找不到 Boto。唯有聲音才是保護 Boto 的關鍵 [13]。

研究團隊在河水裡放入了能夠聆聽聲音的水下麥克風。這個麥克風長時間不停歇地記錄聲音。人類很難把這麼長的錄音全部聽完。於是人工智慧替人進行檢查。這裡使用的人工智慧是一種透過觀察聲音圖案來進行判斷的機器。它會過濾掉叢林中的風雨聲和樹枝折斷聲。然後找出 Boto 發出的喀嗒聲。

這個人工智慧以95%的準確率找到了 Boto 的聲音。也能以92%的準確率分辨出船隻經過的聲音。在視線完全看不見的泥水深處，機器僅憑聲音守護著 Boto 的棲息地。生活在河裡的另一種海豚——土庫海豚的聲音也一起被分辨了出來。就這樣，人工智慧用聲音尋找看不見的海豚。

這些紀錄立刻用來保護 Boto。研究團隊用聲音描繪出 Boto 經常出沒的路徑。也查出牠們在什麼季節會聚集到哪裡。到了水位上漲的季節，Boto 會從大河游進被水淹沒的森林。透過分析聲音，清楚看見了這個移動過程。這張地圖有助於改變船隻航行路線與下網捕魚的位置。聲音成了守護 Boto 生命的告示牌。即使在沒有大型電腦的叢林深處，尋找能長時間聆聽聲音並守護牠們的方法，相關研究仍在持續進行 [2]。

相關佐證資料

[2] NOAA Fisheries. (2025). A whale's world of sound. National Oceanic and Atmospheric Administration.

<https://www.fisheries.noaa.gov/feature-story/whales-world-sound>

[12] Erbs, F., Gaona, M., van der Schaar, M., Zaugg, S., & André, M. (2023). Towards automated long-term acoustic monitoring of endangered river dolphins: A case study in the Brazilian Amazon floodplains. Scientific Reports, 13, 10801. <https://doi.org/10.1038/s41598-023-36518-1>

[13] Kane, S. (2023, August 1). AI unlocks secrets of Amazon river dolphins' behavior, no tagging required. Mongabay. <https://news.mongabay.com/2023/08/ai-unlocks-secrets-of-amazon-river-dolphins-behavior-no-tagging-required/>

4.6 傾聽海洋

在第4章看到的人工智慧水下研究，改變了一項長久以來的觀念。人類長久以來，都把其他動物當作只順從本能生活的存在 [1]。但最近的研究展現出了不同的樣貌 [1]。虎鯨的方言展現了世代相傳的文化 [1]。白鯨在河流中互相傳遞聲音的模樣，正是牠們形成緊密相連社會的證據 [1]。

關於鯨魚話語的新發現也接連出現。這是 Project CETI 與多所大學研究團隊的研究成果。2024年，他們揭開了抹香鯨聲音的結構。抹香鯨會把多個喀嗒聲連在一起組成訊號。這一組聲音被稱為 coda。研究團隊發現 coda 裡有著速度和節奏等規則。將這些規則組合起來，就會產生無數種不同的聲音。這代表那些訊號比想像中更有條理。研究團隊還發表了整理這些聲音碎片的「聲音字母表」。2026年，研究更進了一步。他們在抹香鯨的聲音中，找到了兩種與人類語言母音相似的聲音形式。那是就像人類的「啊」聲和「依」聲那樣區分的形式。不過，這並不代表已經破解了抹香鯨話語的意思 [14][15]。

協助人們用人類語言詢問並解答動物聲音的人工智慧也在不斷成長。這就是 Earth Species Project 打造的 NatureLM-audio。這是專為動物聲音設計的第一個聲音與語言基礎模型。當人類用平常的話發問時，人工智慧就會分析聲音並回答。它能區分物種、找出聲音，甚至能數出有幾隻動物在鳴叫 [16]。

解讀鯨魚語言的技術，也對法律提出了新的問題。2025年10月，法學家與 Project CETI 研究團隊共同發表了論文。他們認為鯨魚是能夠分享文化並相互溝通的存在。因此提議也應賦予鯨魚不遭受殘酷對待的權利。紐約大學法學院的「超越人類的生命（MOTH）」計畫團隊帶領著這項工作。人工智慧開啟的新耳朵，就這樣延伸到了法律與倫理的提問 [17]。

這段時間以來，我們向大海傾倒了太多噪音。大型船隻的聲響與探勘爆炸聲，切斷了鯨魚彼此溝通的管道。如今，人工智慧成了小心翼翼解讀海洋聲音的聽診器。這個工具提醒了人類一件事。那就是放低我們的聲音，如實傾聽大自然原本的聲音。當我們放下想要支配海洋的心態時，才能真正聽清鯨魚悠久的歌聲。深海的交響樂團至今尚未停止演奏。只要我們用心傾聽，那些聲音就會再次回到我們身邊 [2][1]。

相關佐證資料

[1] Bakker, K. (2022). The sounds of life: How digital technology is bringing us closer to the worlds of animals and plants. Princeton University Press.

- [2] NOAA Fisheries. (2025). A whale's world of sound. National Oceanic and Atmospheric Administration. <https://www.fisheries.noaa.gov/feature-story/whales-world-sound>
- [14] Sharma, P., Gero, S., Payne, R., et al. (2024). Contextual and combinatorial structure in sperm whale vocalisations. *Nature Communications*, 15, 3617. <https://doi.org/10.1038/s41467-024-47221-8>
- [15] Beguš, G., Dąbkowski, M., Sprouse, R. L., Gruber, D. F., & Gero, S. (2026). The phonology of sperm whale coda vowels. *Proceedings of the Royal Society B*, 293(2069). <https://doi.org/10.1098/rspb.2025.2994>
- [16] Robinson, D., Miron, M., Hagiwara, M., Pietquin, O., et al. (2024). NatureLM-audio: An audio-language foundation model for bioacoustics. arXiv:2411.07186. <https://arxiv.org/abs/2411.07186>
- [17] Rodríguez-Garavito, C., Gruber, D. F., Nemeth, A. O., & Beguš, G. (2025, October). What if we understood what animals are saying? The legal impact of AI-assisted studies of animal communication. *Ecology Law Quarterly*, 52(1). <https://www.ecologylawquarterly.org/print/what-if-we-understood-what-animals-are-saying-the-legal-impact-of-ai-assisted-studies-of-animal-communication/>

第5章 大地的語言

5.1 黑猩猩的聲音有規則

人類的話語是把幾種聲音組合連接起來的。用少數幾種聲音，就能創造出無窮無盡的意思。長久以來，人們一直認為這種本領只有人類才有。然而，長期觀察動物聲音的學者們，開始重新思考這個想法。

舞台是西非象牙海岸的塔伊國家公園。這裡生活著一群野生黑猩猩。德國馬克斯普朗克演化人類學研究所與法國國家科學研究中心的學者們，長期觀察著牠們。Cédric Girard-Buttoz 博士與 Roman Wittig 博士帶頭展開了這項研究 [1]。

這項研究並非一兩年就結束。而是歷經多年觀察同一個群體所得到的結果。學者們收集了 46 隻成年黑猩猩的聲音。錄下來的聲音多達 4,826 個。過去人們只把這些聲音當作表達情緒的呻吟聲。但是一個一個仔細清點後，發現裡面其實井然有序。

學者們長期聆聽這麼多聲音，並將它們分類。像用尺測量一樣，仔細區分聲音的高低與長短。結果發現黑猩猩使用的基本聲音可以分為 12 種。例如哼唧聲、吠叫聲和尖叫聲等。這就像人類語言的子音和母音一樣，是最微小的聲音碎片。

黑猩猩並不是隨隨便便發出這 12 種聲音。牠們會按照一定的順序連接組合。這樣組合出來的聲音組合超過了 390 個。有些順序經常出現，有些順序則幾乎不出現。這代表聲音的連接是有固定結構的。

現在，分析這麼多聲音時，稱為機器學習的人工智慧會協助學者。機器學習能從無數聲音中，快速找出重複出現的結構。人類獨自需要花費數年才能完成的計算，它幾天之內就能完成。找出規則的是學者，人工智慧只是協助這項工作的工具。

聲音組合呈現出層次結構。由兩個聲音相連的組合，有時也會包含在三個聲音的組合之中。人類的語言也是這樣構成層次的。就像詞彙組合成短語，短語組合成句子一樣。

學者們指出了人類語言的一大特徵。那就是把沒有意義的微小聲音碎片連接起來，創造出新意思的能力。單一聲音碎片本身沒有意義。連接組合的順序才真正創造出意義。這種能力被稱為聲音的雙重結構。不過，黑猩猩的 12 種聲音與人類的子音、母音並不完全相同。學者們目前仍在謹慎地進行研究 [2]。

隨著時間推移，新的研究接踵而來。2025 年，學者們發現黑猩猩會將兩種聲音配對，從而改變意思。仔細研究這些雙聲音組合，發現一共有 16 種。有些組合直接把兩種聲音的意思加在一起。有些組合則轉變成了全新的意思。

例如，某種聲音後面接上另一種聲音，意思就會改變。有時會變成在樹上築巢睡覺的訊號。或者把通報危險的聲音與呼叫援助的聲音連在一起，傳達要小心蛇的訊息。只有單一聲音和兩個聲音連在一起時，具有彼此不同的意思 [3]。

這些發現是告訴我們人類語言是如何誕生的足跡。黑猩猩用這些聲音組合來通報危險，並協調彼此的行動。這代表編織組合聲音的本領，並非人類所獨有。黑猩猩的聲音規則，展現了從遙遠過去延續至今的溝通痕跡 [4]。

相關佐證資料

[1] Girard-Buttoz, C., Wittig, R. M., et al. (2022). Chimpanzees produce diverse vocal sequences with ordered and recombinatorial properties. *Communications Biology*, 5, 410. <https://doi.org/10.1038/s42003-022-03350-8>

[2] Girard-Buttoz, C., et al. (2025). Versatile use of chimpanzee call combinations promotes meaning expansion. *Science Advances*, 11(19), eadq2879. <https://doi.org/10.1126/sciadv.adq2879>

[3] Leroux, M., et al. (2023). Call combinations and compositional processing in wild chimpanzees. *Nature Communications*, 14, 2225. <https://doi.org/10.1038/s41467-023-37816-y>

[4] Sun, W. (2025, May 9). Origins of language: Wild chimps mirror linguistic structures in human language. *Phys.org*. <https://phys.org/news/2025-05-language-wild-chimps-mirror-linguistic.html>

5.2 獅尾狒神似人類的說話節奏

如果說黑猩猩在森林裡展現了聲音規則，那麼另一種猴子發出的聲音則與人類更為相似。那就是生活在衣索比亞高山上的獅尾狒（*Theropithecus gelada*）。牠們在險峻的岩石懸崖上組成龐大群體生活。牠們彼此交談的聲音，聽起來就像人類在輕聲聊天一樣 [1][2]。

獅尾狒的聲音在人類耳朵聽起來格外親切。彷彿好幾個人用低沉的嗓音在喃喃細語。那種聲音整天在懸崖間迴盪。這就是學者們長久以來深深被這種聲音吸引的原因。

美國密西根大學的 Thore Bergman 教授長期研究了這種聲音。獅尾狒在友好相處時會快速抖動嘴唇。這種聲音被稱為抖唇聲。牠們還會疊加上震動喉嚨發出的粗厚聲音。這種聲音被稱為 wobble 聲。

Bergman 教授測量了這種聲音的節奏。人類說話時，嘴巴每秒大約開合三到八次。獅尾狒的 wobble 聲則是每秒震動六到九次。這兩種節奏彼此重疊。計算每秒重複次數的單位稱為赫茲（Hz）。

這項發現於 2013 年公諸於世。人類與獅尾狒生活在彼此相距遙遠的地方。即使如此，發出聲音的節奏卻非常相似。就像事先約定好了一樣。

獅尾狒不只是抖動嘴唇而已。牠們會重疊兩種聲音，藉此容納更多的意思。學者們將獅尾狒的聲音與豚尾狒的聲音進行了比較。一起聚居生活的群體越大，聲音的種類也越多。這就像可以交流的鄰居越多，能分享的話語也就越多一樣。

如今，這類聲音資料已經累積得非常豐富。因此學者們開始將人工智慧作為工具來使用。人工智慧能快速將無數聲音分類。將過去只能靠耳朵推測的節奏，用數字清晰地呈現出來。揭開真相的是學者，人工智慧則是協助這項工作的手 [3]。

這裡產生了一個疑問。人類與獅尾狒的祖先不同，為什麼節奏會如此相似呢？外表不同的動物朝著同一個方向演化，這種現象稱為趨同演化。就像兩個不同的班級各自設計的加油口號，卻帶有相似的節奏一樣。過著相似的生活，自然就會具備相似的本領。

相關佐證資料

[1] Bergman, T. J. (2013). Speech-like vocalized lip-smacking in geladas. *Current Biology*, 23(7), R268-R269.
<https://doi.org/10.1016/j.cub.2013.02.038>

[2] Gustison, M. L., le Roux, A., & Bergman, T. J. (2012). Derived vocalizations of geladas (*Theropithecus gelada*) and the evolution of vocal complexity in primates. *Philosophical Transactions of the Royal Society B*, 367(1597), 1847-1859.
<https://doi.org/10.1098/rstb.2011.0218>

[3] Nature. (2025, September 17). AI is helping to decode animals' speech. *Nature News*.
<https://doi.org/10.1038/d41586-025-02917-9>

5.3 地底裸鼯鼠的方言

如果說猴子在明亮的地面上交流聲音，那麼在黑暗的地底也存在著聲音的社會。那是一點光線也沒有的泥土洞穴深處。這裡生活著裸鼯鼠（*Heterocephalus glaber*）。這種動物像螞蟻或蜜蜂一樣過著群居生活。由一隻女王負責繁衍後代，其餘成員則分工合作。像女王與工鼠分工角色的這種社會，稱為真社會性 [1]。

德國柏林 Max Delbrück Center 的研究團隊收集了牠們的聲音。Alison Barker 博士帶頭展開了這項研究。裸鼯鼠相遇時會發出溫柔的唧唧叫聲。研究團隊收集了 3 萬多個這類聲音，並逐一進行分析。像用尺測量一般，精確提取了聲音高低與長短等多種特徵 [2]。

研究揭開的事實是這樣的。每個群體都有各自不同的發聲習慣。對人類來說，這就像不同地區各自擁有的方言一樣。這被稱為地底群體方言。裸鼯鼠在漆黑的洞穴中無法用眼睛看。只能完全依靠聲音來區分自己人與外人。

洞穴裡總是一片漆黑。沒有光線，根本無法辨認臉孔。因此聲音就扮演了名牌的角色。說同一種方言就是同一個群體。如果聽到陌生的方言，就會被當作別的群體。單單一種聲音，就成了守護群體的守門人。

這種方言並非天生具備的。剛出生的小鼠會不斷模仿成年鼠的聲音。在這個過程中學習群體的發聲習慣。從其他群體新加入的裸鼴鼠也是如此。牠們透過模仿新群體的聲音來習得該方言。方言並非由血統決定，而是透過學習決定的。

這種學習過程與人類嬰兒牙牙學語的過程十分相似。就像嬰兒聽著大人的話慢慢模仿一樣，幼鼠也是這樣逐步學會聲音的 [3]。

最值得關注的地方是女王的力量。只要女王活著，群體的方言就能維持統一。然而一旦女王突然死亡，方言就會瓦解混亂。聲音的規則頓時全面動搖。這代表女王正是把群體的聲音凝聚在一起的支柱。

在分類這麼多聲音時，人工智慧提供了很大幫助。人工智慧迅速辨別出每個群體不同的發聲習慣。但是，揭開方言是透過學習習得、且由女王維持著這一事實的，是長期的觀察。人工智慧只是協助那項觀察的工具。

女王倒下，聲音也隨之瓦解，這個事實寓意深遠。要將聲音凝聚為一體，必須有一個中心。這就像我們班上共同訂立的班規一樣。讓大家遵守規則的中心一旦動搖，規則本身也會跟著動搖。

相關佐證資料

[1] Barker, A. J., et al. (2021). Cultural transmission of vocal dialect in the naked mole-rat. *Science*, 371(6528), 503-507. <https://doi.org/10.1126/science.abc6588>

[2] Max Delbrück Center for Molecular Medicine. (2021, January 28). Naked mole-rats 'chirp' in colony-specific, culturally transmitted dialects. *EurekAlert!* <https://www.eurekalert.org/news-releases/510214>

[3] Vaidyanathan, G. (2021, February). The queen sets the tone: Deciphering the dialects of naked mole-rats. *Mongabay News*. <https://news.mongabay.com/2021/02/the-queen-sets-the-tone-deciphering-the-dialects-of-naked-mole-rats/>

5.4 狐獴的警報聲與鬣狗的笑聲

陸地上的許多動物也會在聲音中融入豐富的意思。狐獴 (*Suricata suricatta*) 就是一個很好的例子。狐獴用聲音告知敵人是從天上來還是從地面來。老鷹來襲或蛇類逼近時，發出的聲音都不相同。連危險有多緊急，都會包含在聲音裡面 [1]。

這項事實是由瑞士的 Marta Manser 教授很久以前揭開的。Manser 教授在南非喀拉哈里沙漠長期觀察了狐獴。逐一聆聽並分類牠們的警報聲。危險越緊急，聲音就越粗糙嘈雜。通報危險種類的聲音與通報緊急程度的聲音各有不同。狐獴光靠一種聲音，就能帶動整個群體採取行動。

現在，人工智慧也能對這類警報聲進行分類。人工智慧能迅速辨識出人類耳朵難以察覺的細微聲音紋理。但是，聲音中蘊含著危險種類與緊急程度這件事，是由長期的現場觀察率先揭開的。

稱霸非洲草原黑夜的斑鬣狗 (*Crocuta crocuta*) 也是如此。鬣狗會發出像是在笑的聲音。這種咯咯笑聲中包含著許多資訊。法國的 Nicolas Mathevon 研究團隊分析了這種聲音。雖然聽起來像雜亂無章的噪音，但事實並非如此。裡面包含了年齡以及在群體中的位階順序。甚至還含有辨認每個不同個體的標記。這就像人類的聲紋一樣 [2]。

Mathevon 研究團隊錄下了美國柏克萊 17 隻鬣狗的聲音。並從笑聲中提取出兩種特質。一種是嘴部與喉嚨形狀所構成的聲音紋理。這被稱為共振峰 (formant)。另一種是聲音的高低。這被稱為音高 (pitch)。地位高的鬣狗與地位低的鬣狗，這兩種特質彼此不同。只要聽聲音，就能知道誰的力量更強大。

與人類親近的狗 (*Canis lupus familiaris*) 也是如此。美國密西根大學的研究團隊研究了狗的聲音。收集了吠叫聲、低吼聲和嗚咽聲。並且嘗試將分析人類語言的人工智慧應用在狗的叫聲上。這款人工智慧的名字是 Wav2Vec2。結果成功辨別出了叫聲是在玩鬧還是凶狠。甚至在一定程度上猜中了狗的品種、性別與年齡。準確率達到了 70%。處理人類語言的工具，在狗的叫聲上也同樣奏效 [3]。

這裡產生了一個疑問。前面提到獅尾狒與人類的節奏相似屬於趨同演化。雖然祖先不同，但因為過著相似的生活而變得相似。然而人類與狗發聲的基礎，卻是因為不同的原因而相似。人類與狗都是哺乳動物。在很久以前從同一個祖先分化而來。因此利用喉嚨與聲帶發出聲音的身體構造，從一開始就很相似。像這樣因為祖先相同而根源

相似的現象，稱為同源。就像人類的手與动物的前腳是由相同的骨骼分化而來一樣。相似性有著這兩種途徑。一種是分開生活卻變得相似的途徑，另一種則是從同一個根源分化出來的途徑。

相關佐證資料

- [1] Manser, M. B. (2001). The acoustic structure of suricates' alarm calls varies with predator type and the level of response urgency. *Proceedings of the Royal Society B*, 268(1483), 2315-2324. <https://doi.org/10.1098/rspb.2001.1773>
- [2] Mathevon, N., et al. (2010). What the hyena's laugh tells: Sex, age, dominance and individual signature in the giggling call of *Crocuta crocuta*. *BMC Ecology*, 10, 9. <https://doi.org/10.1186/1472-6785-10-9>
- [3] Abzaliev, A., Pérez Espinosa, H., & Mihalcea, R. (2024). Towards dog bark decoding: Leveraging human speech processing for automated bark classification. *Proceedings of LREC-COLING 2024*. <https://aclanthology.org/2024.lrec-main.1432/>

5.5 單一人工智慧學習多種動物的聲音

到目前為止，學者們都是針對每種動物分別研究其聲音。黑猩猩歸黑猩猩研究，獅尾狒歸獅尾狒研究。但最近，同一款人工智慧卻能同時學習多種動物的聲音 [1]。

Earth Species Project 走在這項領域的最前端。這個組織試圖用人工智慧來解開動物的聲音。他們打造了名為 NatureLM-audio 的模型。這是一款能聆聽動物聲音並回答人類提問的人工智慧。它於 2024 年 11 月首次亮相，並在 2025 年開放給所有人使用。

這個模型學習了數千種生物的聲音。光憑聲音就能猜出是哪種動物。還能分辨出是雌性還是雄性、是成體還是幼體。甚至還能數出錄音中有幾隻動物。它甚至自行展現出從未被教導過的能力。

該組織在 2024 年 10 月獲得了巨大助力。從多位贊助者手中獲得了 1,700 萬美元的研究經費。這筆錢超過 230 億韓元，是一筆龐大的資金。他們利用這筆資金大幅擴充了研究團隊。該組織認為，到 2030 年將能對動物的聲音取得重大突破 [2]。

2026年夏天，許多研究團隊共同使用的資料庫也開放了。這個資料庫的名字是動物語言處理資料 (alp-data)。它把分散在各處的動物聲音資料收集在同一個地方。現在正一起觀察烏鴉、白鯨與大象的聲音。資料收集得越多，人工智慧就學得越快 [5]。

推動這股趨勢的獎項也出現了。這就是柯勒杜立德獎 (Coller-Dolittle Prize)。這個獎項為解讀動物聲音的研究增添了力量。2026年的獲獎者是美國柏克萊的茱莉·艾莉博士。艾莉博士解讀了一種名為斑胸草雀 (zebra finch) 的小鳥聲音。她找出了這種鳥使用的11種基本聲音。裡面包含了肚子餓、害怕和打招呼等意思。艾莉博士也得到了名為機器學習的人工智慧協助 [3]。

人工智慧也在動物的聲音中，發現了和人類語言一樣的節省規則。人類會把常常說的話縮短。越常使用，聲音就越短。2025年，有一位學者研究了16種鯨魚的聲音。接著把這些聲音與人類51種各國語言做比較。在座頭鯨和藍鯨的歌聲中，也看得到這種節省現象。越常唱出的聲音就越短。懂得節省聲音的智慧，並不是人類所獨有的。不過並非所有鯨魚都這樣，學者們正在進一步研究 [4]。

讓我們想一想黑猩猩的聲音規則、獅尾狒的說話節奏，以及裸鼴鼠的方言。這些鄰居的聲音，帶給我們許多值得思考的事情。人類的語言並不是天上憑空掉下來的奇蹟。而是在許多生命共同站立的寬廣基礎上綻放出來的。

人工智慧會替我們找出聲音裡的規則。人類自己要做上好幾年的工作，它幾天之內就能完成。多虧了人工智慧，學者們得以觀察更多的聲音。但是，人工智慧並不是直接把動物的話翻譯過來。它只是協助將聲音分類並找出規則而已。

聽懂動物的聲音，對保護牠們也很有幫助。如果知道哪種聲音代表害怕，就能察覺動物難受的時刻。也能從聲音推測出是什麼在困擾著牠們。讀懂聲音的力量，就會變成照顧鄰居的力量。

一直以來，我們只把這些聲音當成樹叢裡的雜音。甚至還隨意挖開牠們生活的土地。然而，牠們長久以來一直用聲音傳承著群體的歷史。現在我們該做的，並不是先讓牠們聽人造的聲音。而是應該從謙虛地聆聽牠們的聲音開始。

人工智慧成了一雙好耳朵，幫助我們去聆聽。它能替我們聽出人類曾經錯過的微小細節。但是，要聽什麼、為什麼要聽，是由人類決定的。掌握好工具的同時，謙虛聆聽的心也必須同行。這樣一來，我們才能與大地的鄰居們長久並肩生活。

相關佐證資料

- [1] Robinson, D., Miron, M., Hagiwara, M., & Pietquin, O. (Earth Species Project). (2024). NatureLM-audio: An audio-language foundation model for bioacoustics. arXiv:2411.07186. <https://arxiv.org/abs/2411.07186>
- [2] Earth Species Project. (2024, October 17). Earth Species Project secures \$17M in grants to apply frontier AI to decoding animal communication. <https://earthspecies.org/2024/10/17/earth-species-project-secures-17m-in-grants-to-apply-frontier-ai-to-decoding-animal-communication/>
- [3] Jeremy Collier Foundation. (2026, July 3). Scientists win US\$100,000 in 2026 Collier Dolittle Prize. <https://www.jeremycollierfoundation.org/news-and-insights/press-releases/scientists-win-us100-000-in-2026-coller-dolittle-prize>
- [4] Youngblood, M. (2025). Language-like efficiency in whale communication. Science Advances, 11(6), eads6014. <https://doi.org/10.1126/sciadv.ads6014>
- [5] Earth Species Project. (2026, July 22). Introducing alp-data: A shared data layer for animal language processing. <https://earthspecies.org/2026/07/22/introducing-alp-data-a-shared-data-layer-for-animal-language-processing/>

第6章 天空的交響曲與看不見的舞蹈

6.1 鳥類與人工智慧互相進行的對話

人類的對話並不是預先寫好的講稿。我們會聆聽對方的說話內容與語氣。就連對方短暫停頓的片刻，我們也會仔細留意。我們會配合那個節奏接上自己的回答。鳥類的歌聲中，也包含著這樣一來一往的交流。

過去，科學家在研究鳥叫聲時，通常是播放錄音給牠們聽。這種重新播放聲音的方法，稱為回放（playback）。傳統的回放只是播放預先設定的簡短聲音或固定錄音。這種做法很難捕捉到鳥類長時間一問一答的對話節奏。因此，人們長期以來都以為鳥類的歌聲是出生時就已經注定的聲音 [1]。

為了跨越這道障礙，兩個研究團隊聯手合作。這兩個團隊分別是想利用人工智慧解開動物溝通奧秘的 Earth Species Project，以及加拿大的麥吉爾大學（McGill University）。研究由羅根·詹姆斯（Logan James）博士與莎拉·伍利（Sarah Woolley）教授帶領。他們收集了體型跟麻雀差不多大的雌性斑胸草雀（zebra finch）的聲音。彼此一來一往交流的聲音超過了 1,000 小時 [2]。

研究團隊用這些聲音打造出人工智慧模型 ZF-AIM。當鳥兒唱出一小節，這個模型就會立刻生成下一小節來回答。這就像智慧型手機在我們打字時，會預先顯示下一個詞彙一樣。像這樣聆聽對方的聲音並立刻接話的方式，稱為一問一答的對話。

研究團隊非常謹慎地設計了實驗。他們把雌性斑胸草雀分別放在不會漏音的隔音箱裡。並讓箱子裡的鳥透過麥克風和揚聲器互相傳遞聲音。當鳥發出聲音時，ZF-AIM 就會配合那個聲音生成下一段聲音播放給牠聽。這代表鳥類與人工智慧正在即時進行聲音的互動。

研究團隊比較了三種不同情況。分別是與活生生的鳥互動時、單純播放錄音時，以及與即時人工智慧互動時。只有在與活生生的鳥或人工智慧一問一答時，才會出現自然的特徵。單純播放錄音時，則看不到那樣的特徵。

實驗結果顛覆了長久以來的想法。過去教科書上寫著，雌性斑胸草雀的鳴叫聲是由遺傳決定、不會改變的。但是與人工智慧進行聲音互動的鳥兒卻不一樣。牠們會配合對方發出的聲音改變自己的聲音。這證明了鳥類的對話並不是單純重複固定好的聲音。

研究團隊為什麼會選擇斑胸草雀呢？因為斑胸草雀不會獨自生活。牠們會成群結隊，不停地互相呼應發聲。牠們一邊互相呼喚回答，一邊變換位置。這種熱絡的對話非常適合用來進行研究。人工智慧則原原本本地模仿了那種快速的一問一答。

研究團隊在鳥類的對話中發現了三個特徵。第一，兩隻鳥的聲音變得越來越像。第二，牠們能迅速轉換聲音。第三，牠們並不會對任何聲音都有反應。而是會挑選自己想回答的聲音來回應。這就像在鬧哄哄的教室裡，只有聽到自己的名字時才會回頭一樣。

研究團隊嘗試將模型的一部分一部分分別暫停。當暫停負責配合節奏的部分時，對話就亂了套。當暫停負責塑造聲音特徵的部分時，其他地方又出現了偏差。這表示這兩個部分各自承擔著不同的任務。

研究團隊認為，這種方法對理解人類的對話也很有幫助。人類在對話時，也會猜測對方下一句會說什麼。在鳥類與人工智慧的互動中，也看得到這種猜測。這種聲音一問一答的規則，說不定適用於許多動物。

研究團隊把兩個 ZF-AIM 模型面對面放著讓牠們對話。完全不靠人手干預，讓人工智慧彼此交流聲音。那場對話重現了真正鳥類之間對話的特徵。這代表模型確實學會了當中的規則。

這項研究指向了一個更遠大的目標。那就是用人工智慧生成的聲音主動向動物搭話。現在已經進展到鳥類能夠回應人工智慧聲音的階段。下一步則是人類發送帶有意義的聲音給動物。到了那天，人類與鳥類說不定就能進行真正的對話了 [3]。

相關佐證資料

[1] Earth Species Project. (2023, December 3). Using AI to understand animal communication. Earth Species Project. <https://www.earthspecies.org/blog/using-ai-to-understand-animal-communication>

- [2] James, L. S., Hoffman, B., Woolley, S. C., et al. (2026). Leveraging AI-powered interactive playbacks to decipher rules of communication in zebra finches. bioRxiv. <https://doi.org/10.64898/2026.02.12.705387>
- [3] Earth Species Project. (2026, April 2). Animal Language Processing: An AI convergence in animal communication. Earth Species Project. <https://earthspecies.org/2026/04/02/animal-language-processing-an-ai-convergence-in-animal-communication/>

6.2 尋找歌聲中隱藏的規則

在斑胸草雀與人工智慧互動聲音的同時，鳥的大腦中也正在進行學習。鳥類和人類一樣，小時候會模仿大人的聲音唱歌。這個時期稱為敏感期（sensitive period），也就是最適合學習聲音的時期。這種學習並不單純由遺傳決定。共同生活的群體氛圍也會帶來很大的影響 [1]。

幼小的斑胸草雀是聽著成年鳥的歌聲長大的。一開始模仿得有些笨拙。但漸漸地會越來越接近成年鳥的歌聲。這和人類小嬰兒學說話的樣子非常相似。在群體中聽得越頻繁，學得就越好。

鳥類的大腦裡有一個主導歌唱順序的部位。這裡被稱為 HVC，也可以說是歌唱指揮總部。這個部位負責安排要在什麼時候發出什麼聲音的順序。這就像指揮家引領節奏的模樣。

美國加州大學聖地牙哥分校（UC San Diego）的提摩西·根特納（Timothy Gentner）教授團隊觀察了歐洲椋鳥（European starling）。這種鳥能將數百個聲音片段串連成長長的歌曲來歌唱。單憑人類的耳朵去聽，很難找出其中的規律。於是研究團隊把未貼標籤的歌聲輸入到人工智慧中 [2]。

人工智慧把相似的聲音片段聚集在一起。就如同把歌聲像地圖一樣展開。在那張地圖上，聲音形成了一個個群體。人類先前沒能看出的結構這才映入眼簾。用這種方法，就能把各種動物的聲音放在一起互相比較。

研究團隊把這個方法應用在多種動物的聲音上。他們把鳥、老鼠、猴子、人類和鯨魚的聲音放在同一張地圖上。即使物種不同，也能用相同的方式進行比較。各種聲音是由誰發出來的，其中的特徵也清晰地呈現出來。

人類即使聽了很久的鳥鳴，也無法找齊所有的規則。因為聲音節奏很快，而且片段非常多。人工智慧卻能不知疲倦地分析數千小時的歌聲。把人類遺漏的順序一個一個找出來。

金絲雀（canary）的歌聲中也發現了隱藏的規律。美國波士頓大學（Boston University）的研究團隊分析了金絲雀悠長的歌聲。過去人們習慣使用只看前一個聲音來預測下一個聲音的方法。這種簡單的方法稱為馬可夫模型。然而金絲雀的歌聲並沒有那麼簡單 [3]。

研究團隊發現，歌聲前面的部分決定了很久之後才會出現的聲音。一開始唱出的聲音，會一路引領到很後面的聲音。這種特性稱為長期依賴性。這就像長句子裡的第一個詞和最後一個詞相互呼應一樣。看起來像是隨心所欲地歌唱，其實背後隱藏著嚴密的結構。

隨著這些研究不斷累積，一幅宏大的藍圖逐漸展現。Earth Species Project 打造出能同時學習多種動物聲音的人工智慧。這個模型的名字叫做 NatureLM-audio。這是一個同時學會了人類語言、音樂和大自然聲音的模型。這個模型只要聽聲音，就能分辨出是哪種動物。連從未學過的動物聲音也能推測出來。如果人類用文字寫下問題，它會聽取聲音並回答。只要寫下發出這個聲音的動物是什麼，它就會幫忙挑選出物種。連錄音中有幾隻動物在叫也能數得出來 [4]。

2026 年，Earth Species Project 為這一趨勢取了名字。那就是把動物聲音當作資料來處理的研究，也就是動物語言處理（Animal Language Processing）。這個名稱在 2024 年首次被使用。它是借用處理人類語言的自然語言處理而來的名稱。解開鳥鳴規則的工作也是在這一領域中蓬勃發展 [5]。

相關佐證資料

- [1] Earth Species Project. (2023, December 3). Using AI to understand animal communication. Earth Species Project. <https://www.earthspecies.org/blog/using-ai-to-understand-animal-communication>
- [2] Sainburg, T., Thielk, M., & Gentner, T. Q. (2020). Finding, visualizing, and quantifying latent structure across diverse animal vocal repertoires. PLOS Computational Biology, 16(10), e1008228. <https://doi.org/10.1371/journal.pcbi.1008228>

- [3] Markowitz, J. E., Ivie, E., Kligler, L., & Gardner, T. J. (2013). Long-range order in canary song. *PLOS Computational Biology*, 9(5), e1003052. <https://doi.org/10.1371/journal.pcbi.1003052>
- [4] Earth Species Project. (2024, November 11). Introducing NatureLM-audio: An audio-language foundation model for bioacoustics. Earth Species Project. <https://www.earthspecies.org/blog/introducing-naturelm-audio-an-audio-language-foundation-model-for-bioacoustics>
- [5] Earth Species Project. (2026, April 2). Animal Language Processing: An AI convergence in animal communication. Earth Species Project. <https://earthspecies.org/2026/04/02/animal-language-processing-an-ai-convergence-in-animal-communication/>

6.3 漆黑蜂巢裡的八字舞

天空的秩序並不只存在於大型鳥類身上。小蜜蜂的振翅之中也蘊含著這種秩序。人類首次揭開動物精準溝通奧秘的，正是蜜蜂的舞蹈。奧地利學者卡爾·馮·弗里希（Karl von Frisch）揭開了這個祕密，並因此獲得諾貝爾獎 [3]。

蜜蜂一旦找到有花蜜的地方，就會回到蜂巢。接著牠會搖擺身體，畫出數字 8 的形狀。這種舞蹈被稱為八字舞（waggle dance）。蜂巢內部漆黑一片，沒有任何光線。蜜蜂以蜂巢牆面的上方作為基準。所謂的上方，就是重力的反方向。

蜜蜂會筆直穿過八字舞的中間線。這時身體筆直前進的角度，就傳達了方向。以向上為基準傾斜了多少角度，就是花蜜所在的方向。搖擺身體的時間長短則代表距離。搖擺得越久，代表距離越遠。單靠這支舞，就能傳達數公里之外的地點資訊。

如果花蜜非常近，牠們就不跳八字舞，而是跳圓形舞。這時候牠們不會詳細告知方向。因為很近，不需要特別告知太多細節。隨著距離變遠，牠們就會改跳八字舞。

蜜蜂傳達的訊號不只有肢體動作。胸部的細微震動和翅膀的聲音也會同時發出。還有沿著蜂巢牆面傳遞的震動。像這樣同時使用多種訊號的方式，稱為多模態。也就是同時運用動作、聲音與震動的方式。就算蜂巢裡一片漆黑，同伴也能用身體感受到這些訊號。

這些訊號與人類的多種感官十分相似。我們和朋友說話時，不會只用聲音。還會同時配合眼神與手勢。蜜蜂也是同時運用動作、震動與聲音。

馮·弗里希長時間觀察這種舞蹈，解開了其中的規律。他用顏色標記每一隻蜜蜂。一隻隻比對哪隻蜜蜂去了哪裡。這是過去人類靠雙手花費漫長時間才能完成的工作。人工智慧卻能以快得多的速度完成這項工作 [1]。

德國柏林自由大學（Free University of Berlin）的提姆·蘭德格拉夫（Tim Landgraf）教授團隊研發出一種自動解讀這種舞蹈的方法。研究團隊在蜂巢上安裝了攝影機。這是一種不會傷害蜜蜂的方法。人工智慧能自行在影像中找出蜜蜂的舞蹈。

這種人工智慧迅速解讀了人類肉眼遺漏的動作。它解讀出蜜蜂筆直前進的方向。也測量了蜜蜂搖擺身體的時間。結合這兩點，它推測出了花蜜所在的位置。原本由人類手動完成的工作，現在全由機器代勞。

研究團隊並沒有就此停步。他們把解讀出來的舞蹈輸入到機器蜂中。這隻機器蜂是在蜂巢內部模仿八字舞的機器。它和哈佛大學在空中飛行的微型機器蜂不同。當機器人在蜂巢正中央跳起舞時，真正的工蜂便跟隨了那段舞蹈 [2]。

工蜂朝著機器人指示的方向改變了飛行方向。研究團隊花了數年時間改進機器人的動作與震動。還讓它散發出花香。就這樣讓蜜蜂更能聽從機器蜂的舞蹈引導。

這項實驗展現了兩件事。第一，人類確實正確解讀了蜜蜂的舞蹈。如果機器人發出錯誤的訊號，蜜蜂就不會照著做。第二，蜜蜂確實信任彼此的舞蹈並採取行動。小昆蟲的一個微小動作裡，竟然蘊含著方向與距離。

相關佐證資料

- [1] Wario, F., Wild, B., Rojas, R., & Landgraf, T. (2017). Automatic detection and decoding of honey bee waggle dances. *PLOS ONE*, 12(12), e0188626. <https://doi.org/10.1371/journal.pone.0188626>

[2] Landgraf, T., et al. (2018). Dancing honeybee robot elicits dance-following and recruits foragers. arXiv:1803.07126. <https://arxiv.org/abs/1803.07126>

[3] Frisch, K. von. (1967). The dance language and orientation of bees. Harvard University Press.

6.4 烏鴉在空中畫出的立體對話

在空中飛行的聰明動物中，烏鴉顯得與眾不同。烏鴉以聰明伶俐而聞名。鳥類的大腦雖然比人類小。但神經細胞卻密密麻麻地聚集在一起。因此人們稱烏鴉為「披著羽毛的類人猿」[1]。

西班牙萊昂大學（University of León）與 Earth Species Project 研究團隊觀察了烏鴉。研究對象是生活在西班牙北部的小嘴烏鴉（carrión crow）。研究團隊在鳥的尾羽上貼上了微型感測器。重量只有 12.5g，非常輕巧。並且設計成過一段時間就會自動脫落。這是一種不會傷害鳥類的裝置。

這個感測器裡裝有微型麥克風和動作感應器。麥克風記錄了鳥的聲音。感應器則記錄了翅膀與身體的擺動。何時發出什麼聲音、如何擺動，這些資料一起被記錄了下來。這就像人類用智慧手錶來計算步數的裝置一樣。

研究團隊連續三年安裝了這種感測器。他們在雛鳥孵化大約 10 到 15 天時收集聲音。人工智慧把收集到的聲音分成三類。分別是佩戴感測器的成年鳥叫聲、身旁其他成年鳥的叫聲，以及雛鳥的叫聲。以此分辨出是誰在什麼時候叫喚。

收集到的聲音全由人工智慧進行篩選與分辨。Earth Species Project 開發了一款名為 Voxaboxen 的工具。這是一個能從漫長錄音中逐一挑出叫聲的工具。這個工具找出了超過 12 萬個烏鴉叫聲。其中甚至有研究烏鴉超過 25 年的學者也從未聽過的聲音。那是因為聲音太小，人類耳朵容易漏聽的聲音。即使聲音相互重疊，這個工具也能一一拆分開來。即使在許多鳥同時鳴叫的錄音中，它也發揮了充分的作用 [4]。

研究團隊分別檢視了聲音和動作資料。把這兩種資料放在一起對照查看，是未來要做的事。目前是建立好可以一起對照查看的資料的階段。有了這些資料，打開了更深入了解烏鴉對話的大門。

人工智慧並沒有解開烏鴉的所有意思。人工智慧幫忙把細微的聲音分門別類。揭開烏鴉心思的，是很久以前的動物學家 [2]。

美國華盛頓大學(University of Washington)的 John Marzluff 教授團隊發現了烏鴉能記住人臉。研究團隊戴著橡膠面具捕捉烏鴉。之後，烏鴉對著戴著同樣面具的人大叫並撲上前去。這表示牠們記住了危險人物的臉。牠們還把這個記憶告訴了其他烏鴉。

同一所大學的 Kaeli Swift 和 Marzluff 教授觀察了另一種行為。烏鴉會聚集在死去的同伴身邊大叫。人們把這稱為葬禮。但這不只是因為悲傷。那是學習是什麼讓同伴死去的場合。牠們把危險告訴群體，並長久記住 [3]。

以前，人們只覺得烏鴉的叫聲很吵鬧。用人工智慧分類之後，發現並非如此。每一種聲音似乎都包含著不同的意思。提醒大家要小心誰的聲音、叫大家往哪裡去的聲音，似乎都夾雜在其中。鳥類的叫聲，看起來很接近人類的語言。

在其他鳥類研究中，又往前邁進了一步。2026年，Earth Species Project 發表了只靠聲音就能辨識個別動物的人工智慧。這項研究的對象不是烏鴉，而是斑胸草雀。這為分辨鳥群中每一隻鳥的聲音開闢了道路。現在人類可以回溯每一隻鳥在什麼時候說了什麼話。能觀察誰常常叫、誰比較安靜。這樣的工具也可以用在烏鴉研究上 [5]。

研究越是共享資料，進展就越快。2026年7月，Earth Species Project 推出了新工具。它的名字是 alp-data。這是把散落各處的動物聲音資料整合在一起的工具。它能用同一種方式，讀取超過35個聲音資料集。以前每個研究團隊都要各自整理資料。現在大家可以用同一個工具一起使用。烏鴉聲音的研究，在這種共享資料的基礎上也能進行得更快。在人工智慧揭開的天空背後，鳥兒們每天都在進行著人類從不知道的對話 [6]。

相關佐證資料

[1] Baglione, V., Canestrari, D., Cusimano, M., Hoffman, B., Moreno, V., & Trapote, E. (2025). Capturing vocal communication in a free-living corvid: high-resolution data from low-impact miniaturized tags. *Animal Cognition*, 28, 85. <https://doi.org/10.1007/s10071-025-02018-0>

- [2] Marzluff, J. M., Walls, J., Cornell, H. N., Withey, J. C., & Craig, D. P. (2010). Lasting recognition of threatening people by wild American crows. *Animal Behaviour*, 79(3), 699-707.
<https://www.sciencedirect.com/science/article/abs/pii/S0003347209005806>
- [3] Swift, K. N., & Marzluff, J. M. (2015). Wild American crows gather around their dead to learn about danger. *Animal Behaviour*, 109, 187-197. <https://www.sciencedirect.com/science/article/abs/pii/S0003347215003188>
- [4] Mahon, L., et al. (2025). Robust detection of overlapping bioacoustic sound events. arXiv:2503.02389.
<https://arxiv.org/abs/2503.02389>
- [5] Earth Species Project. (2026, May 5). From what to who? AI learns to recognize individual animals. Earth Species Project. <https://earthspecies.org/2026/05/05/from-what-to-who-ai-learns-to-recognize-individual-animals/>
- [6] Earth Species Project. (2026, July 22). Introducing alp-data: A shared data layer for Animal Language Processing. Earth Species Project.
<https://earthspecies.org/2026/07/22/introducing-alp-data-a-shared-data-layer-for-animal-language-processing/>

第7章 從動物的大腦到深度學習畫布

7.1 自己學習聲音的人工智慧

從前的科學家用耳朵聽動物的聲音，用雙手測量。他們把聲音的高低和長度一一記錄下來。這個方法在安靜的房間裡還算好用。但在海邊或森林裡就不太靈光了。因為風聲和波浪聲會混在一起。當時還有許多誰也不知道意思的動物聲音。光靠人力來做，這項工作實在太龐大了 [1]。

現在由稱為深度學習的人工智慧來接手這項工作。深度學習有許多種方式。有由人類告訴它正確答案的方式。也有不用答案、自己尋找規則的方式。處理動物聲音的人工智慧，大多使用不用答案自己學習的方式。它聽了無數的動物聲音，自己找出藏在裡面的規則。

有一種人工智慧，在廣泛學習之後，可以廣泛運用在各種任務上。這樣的人工智慧被稱為動物聲音基底模型。英文是生物聲學基礎模型（Bioacoustic Foundation Model）。Foundation 是「基礎、底子」的意思。只要把基底打得穩固，遇到新問題也能立刻派上用場 [2, 4]。

基底模型會把一次學好的基礎，分用到許多不同的工作上。可以用來區分聲音。也可以用來消除雜音。就算出現新問題，也不需要從頭重新學習。只要在學好的基礎上再加一點點內容就行了。像這樣在基礎之上進行的小任務稱為應用任務。英文稱為下游任務（downstream task）[2, 4]。

引領這個領域的早期人工智慧之一是 AVES。AVES 沒有靠人類標註的答案卡來學習。它聽了從各地收集來、大約 5,000 小時的聲音。這就算是每天 24 小時不休息地聽，也要聽超過 200 天。

AVES 學習的方法和填空遊戲很像。AVES 直接接收聲音原本的波形。它把這些波形切得細細的。然後故意抹掉其中的幾個地方。再根據前後的聲音，把抹掉的部分猜出來。它無數次重複這個遊戲。不把聲音轉換成圖片，而是直接用聲音本身來學習。

AVES 所掌握的是聲音的紋理。它觀察聲音多有規律地重複。觀察聲音變高變低的起伏流動。觀察聲音描繪出的形狀。只要懂得這種紋理，無論是什麼動物的聲音都能處理。因為就算物種不同，聲音的紋理也是相通的。

在 AVES 出現之前，答案標籤一直是一大阻礙。每種聲音都必須由人類貼上名稱，它才能學習。沒有名牌的聲音就學不會。AVES 開闢了一條不需要名牌也能學習的道路。因此它能把世上的各種聲音都當成材料。

經過這個過程，人工智慧會把聲音的特徵換成數字。這組數字稱為嵌入（embedding）。這相當於把聲音換成了電腦看得懂的數字座標。多虧了嵌入，無論哪種動物的聲音，都能用相同的方式來處理。

這個座標和教室的座位表很像。看了座位表，就知道誰坐在哪裡。看了嵌入，就知道哪種聲音落在什麼位置。相似的聲音會聚在相近的座位。相距很遠的聲音，就是彼此不同的聲音。

這種方式有一個很大的優點。人類不需要一個一個貼上正確答案。貼標籤是一件很花時間的事。要為一個聲音加上名稱，必須由專家親自去聽。人工智慧就算沒有答案，也能找出聲音的規則。因此，連至今還沒有人知道意思的聲音也能處理。這一點和過去的方法截然不同。

相關佐證資料

- [1] Hagiwara, M. (2023). AVES: Animal vocalization encoder based on self-supervision. ICASSP 2023 - IEEE International Conference on Acoustics, Speech and Signal Processing. arXiv:2210.14493. <https://arxiv.org/abs/2210.14493>
- [2] Robinson, D., Robinson, A., & Akrapongpisak, L. (2023). Transferable models for bioacoustics with human language supervision (BioLingual). arXiv:2308.04978. <https://arxiv.org/abs/2308.04978>
- [4] Robinson, D., Miron, M., Hagiwara, M., & Pietquin, O. (2025). NatureLM-audio: An audio-language foundation model for bioacoustics. ICLR 2025. arXiv:2411.07186. <https://arxiv.org/abs/2411.07186>

7.2 一次處理多種動物的聲音

要培養基底模型，需要大量的聲音資料。但是每種資料的格式都不同，非常費工夫。2026 年 7 月，Earth Species Project 推出了一個小幫手。它的名字是 alp-data。它把分散在 35 處以上的動物聲音資料整合在一起。將聲音的速度與標籤統一成單一格式。研究人員省去了整理資料的辛苦，能專注在學習工作上 [10]。

Earth Species Project 也一起推出了可以用眼睛觀看資料的工具。這是一個能一眼瀏覽多種動物聲音資料的工具。像這樣把資料集中在一起並加以挑選，正是培養人工智慧的第一步。優良的資料越多，人工智慧就會越聰明。

資料齊備之後，更強大的人工智慧誕生了。那就是 David Robinson 研究團隊開發的 BioLingual。這個人工智慧學習了陸地動物與海洋動物數百個物種的聲音。它學習的資料是聲音與說明文字成對的組合。這樣的配對超過了 100 萬對 [2]。

BioLingual 把聲音和文字連結起來。它把百科全書上的動物說明文字，和實際錄下的聲音配成一對。這時使用的方法稱為對比學習（contrastive learning）。這種方法是在電腦裡的地圖上，把意思相同的聲音和文字拉得更近。意思不同的聲音和文字則互相推遠。

將聲音和文字連結的想法，來自一個叫做 CLAP 的人工智慧。它是由卡內基美隆大學與微軟研究團隊共同開發的。CLAP 把人類發出的各種聲音和描述那些聲音的文字配成對。BioLingual 把這個想法運用到了動物聲音上 [3]。

多虧了這一點，人類可以用文字來尋找聲音。例如寫下「春天呼喚伴侶的知更鳥高亢鳴叫」。人工智慧就會在無數的錄音中找出那個聲音。而且不需要事先另外教它。像這樣把沒學過的事情立刻做好的能力，稱為零樣本（zero-shot）。也就是就算沒有特別學過，也能一次猜中的本領。

BioLingual 一身就能做很多事。它能分辨發出聲音的動物是哪種生物。也能在長長的錄音中找出有動物聲音的地方。只要用文字詢問，它就會找出相對應的聲音。如果再多學一點新資料，實力還會更加進步。要是換成人類需要好幾個人分工的事，人工智慧一個就能包辦。

BioLingual 學習的聲音來自許多資料庫。有收集鳥叫聲的 Xeno-canto。也有大家拍攝大自然並上傳的 iNaturalist。這些都是研究人員在漫長歲月裡一個個錄音累積而成的。人工智慧就是以這些資料庫為養分學會的。

這樣的人工智慧對研究人員大有幫助。過去為了找一個聲音，要連續聽好幾天的錄音。現在只要短短一行字就能找到。聲音研究的速度因此大幅加快了 [2, 3]。

相關佐證資料

[2] Robinson, D., Robinson, A., & Akrapongpisak, L. (2023). Transferable models for bioacoustics with human language supervision (BioLingual). arXiv:2308.04978. <https://arxiv.org/abs/2308.04978>

[3] Elizalde, B., Deshmukh, S., Al Ismail, M., & Wang, H. (2023). CLAP: Learning audio concepts from natural language supervision. ICASSP 2023. arXiv:2206.04769. <https://arxiv.org/abs/2206.04769>

[10] Earth Species Project. (2026-07-22). Introducing alp-data: A shared data layer for animal language processing. earthspecies.org. <https://earthspecies.org/2026/07/22/introducing-alp-data-a-shared-data-layer-for-animal-language-processing/>

7.3 聽取聲音並用語言解說的人工智慧

Earth Species Project 是一個想用人工智慧解開動物語言的研究團隊。這個團隊開發的人工智慧就是 NatureLM-audio。這個人工智慧廣泛處理鳥類、陸地動物和海洋動物的聲音 [4]。

NatureLM-audio 能夠聽取聲音並用語言回答。只要人類播放動物的聲音，它就會用文字說明那隻動物正在做什麼。它能猜出是哪種生物。也能數出錄音裡有幾隻動物。甚至能分辨發出聲音的動物是幼崽還是已經長大。這些事情就算沒有另外教，它也自己做到了。

這個人工智慧是聽聲音後寫出文字的人工智慧。而不是製造新聲音的人工智慧。過去人工智慧已經擁有聽懂人類語言並用文字回答的能力。如今這股能力拓展到了動物聲音上。這是專為動物打造的第一個大型聲音語言模型。

這個人工智慧的實力是透過測驗來衡量的。這個測驗的名字叫做 BEANS-Zero。這是一門只挑選沒學過的問題來考的測驗。在這個測驗中，該人工智慧成功完成了辨別物種的任務。也成功分類了叫聲的類型。甚至做到了聽聲音並用文字解釋發生了什麼事。

學習越廣泛的人工智慧，處理陌生任務時就越拿手。它的成績比只學過動物聲音的人工智慧更好。這是因為它連人類的語言和各種環境聲音都廣泛聽過的緣故。只學狹隘知識的小型人工智慧，面對新任務時就會迷失方向。唯有基礎夠寬廣，才能解開新的問題。

Earth Species Project 將這個人工智慧用在各種動物研究上。用來研究烏鴉的聲音，也一起觀察白鯨的聲音，還能處理大象的聲音。用同一個人工智慧同時研究多種動物。像這樣用同一種方法處理多個物種聲音的概念，稱為動物語言處理（animal language processing）[11]。

使用這個人工智慧的方法並不難。只要播放錄好的聲音給它聽，並提出問題。詢問發出這個聲音的動物是什麼。它就會用文字給出答案。這就如同用言語向懂得聲音的工具請教一樣 [4, 12]。

人工智慧聽聲音並用文字回答時，有一點必須小心。人工智慧有時會煞有介事地捏造不存在的事實。有時會給出毫不相干的物種名稱。有時甚至會寫出動物根本沒做過的動作。像這樣給出不實答案的現象，稱為幻覺（hallucination）。

有一個減少這種錯誤的方法。那就是讓人工智慧在回答之前，先去查閱實際的紀錄。只讓它寫出用真實動物紀錄確認過的內容。這種方法稱為檢索增強（retrieval-augmented）方式。這就像先在圖書館查書確認之後再開口說話一樣。

Earth Species Project 想要解開動物的語言來保護大自然。只要了解聲音，就能率先知道動物所面臨的困境。可以從聲音中讀出族群正在減少的訊號。人工智慧能比人類更快找出那個訊號 [4, 11]。

相關佐證資料

[4] Robinson, D., Miron, M., Hagiwara, M., & Pietquin, O. (2025). NatureLM-audio: An audio-language foundation model for bioacoustics. ICLR 2025. arXiv:2411.07186. <https://arxiv.org/abs/2411.07186>

[11] Earth Species Project. (2026-04-02). Animal language processing: An AI convergence in animal communication. earthspecies.org. <https://earthspecies.org/2026/04/02/animal-language-processing-an-ai-convergence-in-animal-communication/>

[12] Earth Species Project. (2024-11-06). Introducing NatureLM-audio: An audio-language foundation model for bioacoustics. earthspecies.org. <https://www.earthspecies.org/blog/introducing-naturelm-audio-an-audio-language-foundation-model-for-bioacoustics>

7.4 沒有字典時的翻譯方法

要翻譯兩種語言，通常需要字典。必須有人告訴我們英文的 dog 就是「狗」。但是動物的語言裡沒有字典。因為沒有人記錄過鯨魚用什麼聲音說了些什麼。沒有字典要怎麼翻譯意思呢 [5]？

2017 年，電腦科學家找到了一個方法。就算兩種語言之間連一個配對的例句都沒有，也能進行翻譯。這稱為非監督式機器翻譯（unsupervised machine translation）。意思是就算不告訴它答案，人工智慧也能自己完成翻譯。

要為鯨魚製作配對的例句幾乎是不可能的。因為沒有人知道鯨魚用什麼聲音在說什麼。想要配對，就必須先知道意思；但是要想知道意思，又必須先有翻譯。非監督式翻譯為突破這個死胡同帶來了線索。

這個方法的關鍵在於詞彙所形成的形狀。如果把某種語言的大量詞彙放進人工智慧，它就會畫出一張地圖。意思相似的詞彙會靠得很近。詞彙之間的距離和方向會產生規則。在地圖上，男人和女人的距離有多遠，國王和女王的距離也有多遠。像這樣由詞彙組成的地圖，稱為潛在空間（latent space）。這是一張把詞彙意思用位置表現出來的地圖。

在這裡發生了一件神奇的事。試著畫出不同語言的地圖，會發現形狀幾乎一模一樣。英文地圖和芬蘭文地圖長得很像。這是因為人類的大腦以相似的方式來理解世界。因此只要旋轉一張地圖並疊到另一張地圖上，詞彙就會重疊在一起。英文 dog 的位置上正好會疊上芬蘭文的 koiras。這等於是在沒有字典的情況下產生了翻譯字典。

研究人員正試著把這個方法搬到鯨魚身上。Project CETI 正在進行這項工作。他們分別繪製抹香鯨的聲音地圖與瓶鼻海豚的聲音地圖。接著研究如何旋轉並重疊這兩張地圖。運用能讓兩張地圖平順咬合旋轉的數學方法。目前還沒有完全解開所有意思。現在正處於尋找聲音結構、觀察要在哪裡重疊地圖的階段。

Project CETI 在抹香鯨的聲音中，發現了與人類母音相似的特徵。揭開這個特徵的人是語言學家 Gašper Beguš。他曾在加州大學柏克萊分校進行研究。抹香鯨的聲音片段 coda 分為兩種類型。coda 是鯨魚在對話時發出一連串短促喀嗒聲。一種是 a coda，另一種是 i coda。就像人類用嘴形區分「a」和「i」一樣，鯨魚則是用聲音的共鳴來區分。這項研究發表於 2026 年的英國皇家學會期刊 [6]。

Project CETI 先前也揭開了鯨魚聲音的結構。他們用人工智慧分析了抹香鯨的 8,700 個 coda。結果發現，coda 就像樂高積木一樣可以拼湊組合。速度和節奏等幾種特徵，以各種方式組合在一起。他們把這種結構比喻為人類語言的字母，稱之為聲音字母表。不過，這項研究還沒有解開每個聲音代表什麼意思。目前只是找出了結構規則的階段 [7]。

同年 3 月，Project CETI 也公布了關於抹香鯨的另一項發現。他們詳細記錄了幼鯨出生的瞬間。當時觀察到了好幾隻鯨魚一起照顧幼鯨的模樣。這項研究展現了鯨魚的合作育幼。這和找出聲音母音模式的研究是兩項不同的研究 [13]。

找出聲音的結構，為沒有字典的翻譯帶來了幫助。如果鯨魚聲音裡有和人類語言相似的結構，就更容易找到兩張地圖重疊的位置。聲音切分得越細，地圖就越細密。細密的地圖之間更容易重疊在一起 [5, 6, 7]。

相關佐證資料

- [5] Goldwasser, S., Gruber, D. F., Kalai, A. T., & Paradise, O. (2023). A theory of unsupervised translation motivated by understanding animal communication. NeurIPS 2023. arXiv:2211.11081. <https://arxiv.org/abs/2211.11081>
- [6] Beguš, G., Dąbkowski, M., Sprouse, R. L., Gruber, D. F., & Gero, S. (2026). The phonology of sperm whale coda vowels. Proceedings of the Royal Society B. <https://doi.org/10.1098/rspb.2025.2994>
- [7] Sharma, P., Gero, S., et al. (2024). Contextual and combinatorial structure in sperm whale vocalisations. Nature Communications, 15, 3617. <https://doi.org/10.1038/s41467-024-47221-8>
- [13] Project CETI / Harvard SEAS. (2026-03). Studies documenting rare sperm whale birth and ancient cooperative care released. EurekAlert! <https://www.eurekalert.org/news-releases/1120823>

7.5 跨越物種的人工智慧與無法跨越的高牆

只學習了一種動物聲音的人工智慧，有時也能辨認出其他動物的聲音。這被稱為跨越物種的適應力。英文是跨物種泛化（cross-species generalization）。就像學過鳥叫聲的人工智慧，連青蛙的叫聲也能區分一樣 [8, 9]。

最明顯的例子是 Google 的 Perch。2025 年推出的 Perch 2.0 主要透過鳥叫聲來學習。它聽了非常多以鳥類為主的陸地動物聲音。這裡面包含的動物超過 14,500 種。這個人工智慧幾乎沒有學過海洋中的聲音。即使如此，它還是能很好地分辨鯨魚的聲音。在只用少量資料建立的分類器中，它取得了優異的成績。學習了鳥類聲音的人工智慧，認出了鯨魚 [8, 9] [8]。

Perch 跨越的高牆不只有鯨魚。學習鳥叫聲的 Perch 也處理了青蛙的叫聲。巴西亞馬遜的青蛙叫聲資料集稱為 AnuraSet。Perch 將這些青蛙叫聲依照種類區分開來。它從沒學過青蛙的叫聲，卻成功做到了 [8, 9]。

這些結果是光明的一面。然而，人工智慧無法跨越的高牆也很大。

第一，資料太少了。學習人類語言的人工智慧，是吸收了幾十萬小時以上的錄音長大的。野生動物的聲音就算全部收集起來，也遠遠少於這個數量。而且過去的錄音設備只挑大聲的聲音錄了下來。鯨魚在近距離彼此交談的微弱耳語卻被漏掉了。這種微弱的聲音稱為 soft click。重要的聲音就這樣從資料中遺漏了 [6]。

soft click 是鯨魚彼此靠得很近時發出的微弱聲音。在人類看來就像悄悄話一樣。這微弱的聲音裡，可能蘊含著交流心意的意義。因為過去的錄音機漏掉了這種聲音，導致重要的對話被遺漏了。

第二，生活的世界太不一樣了。抹香鯨生活在連光線都沒有的深海裡。牠們用額頭的油囊發射超音波，用聲音來「看」周圍的環境。人類則是用眼睛來看世界。用聲音感知世界的鯨魚，與用眼睛看世界的人類，感受方式截然不同。因此，兩張地圖可能無法完全重疊在一起。

每個物種各自感受到的不同世界，稱為 Umwelt。意思是「周遭的世界」。提出這個詞的是生物學家 Jakob von Uexküll。意思是每種動物用自己的感官所感受到的世界都不同。哲學家 Ludwig Wittgenstein 也留下過類似的話。他說就算獅子會說話，人類也聽不懂。因為獅子和人類生活的世界是不同的。

第三，隨便使用的話很危險。人工智慧可以模仿雌鳥的叫聲來吸引雄鳥。這會擾亂鳥類尋找伴侶、撫育幼鳥的生活秩序。用假聲音欺騙野生動物的行為，稱為生態欺騙（ecological deception）[14]。

危險還不只停留在模仿。人工智慧可以看著動物的反應來修改聲音。當動物做出回應，就配合那個回應調整聲音。像這樣即時觀察動物的反應並改變聲音的實驗，稱為反應型聲音實驗。用得好，它會成為溝通的鑰匙。用得不好，它就會變成欺騙動物的工具。

鯨魚在數千萬年來，一直傳承著各自的聲音文化。假的聲音可能會污染這份悠久的文化。一旦聲音的秩序被打亂，就很難恢復。因此，在模仿牠們之前必須格外小心。

把龐大的人工智慧帶到森林和海洋時，也會遇到問題。龐大的人工智慧很笨重，而且需要消耗很多能量。森林裡的錄音機很難裝下這麼龐大的東西。因此，研究人員只把龐大人工智慧的核心精華轉移到小型設備上。這就像把大碗裡的精華，盛到小碗裡裝著一樣。

因此，研究人員主張要建立必須遵守的原則。那就是不隨意介入動物聲音的原則。這被稱為最先到要遵守的原則。我們在模仿動物的聲音之前，應該先學會如何安靜地聆聽。

解開鯨魚的語言，法律也可能會改變。2026 年，出現了賦予鯨魚法律權利的討論。如果能透過聲音了解動物感受到了什麼、經歷了什麼，人類就很難再隨意對待動物。解開聲音的工作，與保護動物是緊密相連的 [15]。

人工智慧開闢了一條通往動物語言的新道路。它把人類需要花上幾百年的工作縮短到了幾年。但是，道路被開啟並不代表已經到達終點。資料仍然很少，而且動物的世界與人類的世界不同。人工智慧只是一項工具。如何使用這項工具，取決於人類自己。謙虛聆聽的心態，必須走在技術的前面。

相關佐證資料

[6] Beguš, G., Dąbkowski, M., Sprouse, R. L., Gruber, D. F., & Gero, S. (2026). The phonology of sperm whale coda vowels. *Proceedings of the Royal Society B*. <https://doi.org/10.1098/rspb.2025.2994>

[8] Burns, A., et al. (2025). Perch 2.0: The bittern lesson for bioacoustics. arXiv:2512.03219. <https://arxiv.org/abs/2512.03219>

[9] Choi, C. Q. (2025). AI trained on birdsong can recognize whale calls. *IEEE Spectrum*. <https://spectrum.ieee.org/foundation-models-google-birds-whales>

[14] Beguš, G. (2024). New patterns in animal communication with AI: Ethical and legal implications. *Interspecies Internet Lectures*.

[15] Bolling, R. (2026-06-11). Decoding whale language could unlock new legal rights for them. *Boing Boing*. <https://boingboing.net/2026/06/11/decoding-whale-language-could-unlock-new-legal-rights-for-them.html>

第8章 動物的環境界與符號學邊界

8.1 各自不同的世界，環境界

世界不是只有一個。每個生命都生活在屬於自己的世界裡。有一個人最早清楚提出了這個想法。他是在愛沙尼亞出生的生物學家 Jakob von Uexküll。這位學者認為，每種動物對世界的感受都各不相同 [1]。

Jakob von Uexküll 把這種世界稱為環境界。德語叫作 Umwelt。環境界是指一個生命用自己的感官去感受並建立的專屬世界。人類用眼睛看見顏色，用耳朵聽見聲音。但不是所有動物的感受方式都和人類一樣。

「環境界」這個詞並不難懂。即使在同一個操場上，狗和人類感受到的東西也不一樣。狗靠氣味來認識世界，人類則靠眼睛。每個生命都用適合自己的感官來建立屬於自己的世界。

我們來看看抹香鯨（*Physeter macrocephalus*）。抹香鯨會潛入數百公尺深的深海。那裡是一片沒有任何光線的黑暗。在黑暗中，顏色沒有任何意義。代替視覺的是，抹香鯨會從額頭發出強烈的聲音。牠們聆聽回傳的聲音來辨識周圍的世界。抹香鯨的世界不是用顏色，而是用聲音描繪出來的世界 [2]。

人類在很長一段時間裡都有一個錯覺。以為自己感受不到的東西就是不存在的。人類把自己的感官當成了唯一的標準。這種態度被稱為人類中心主義。人類中心主義是指只把人類的感官與想法當成世界標準的態度。因為這種態度，人類過去只把動物的聲音當成單純的叫聲或本能反應。

有一門學問專門研究動物的聲音與肢體動作。這就是動物符號學（Zoosemiotics）。動物符號學是研究動物之間傳遞信號含義的學問。所謂信號，就是包含意義的聲音、氣味與肢體動作。美國學者 Thomas Sebeok 為這門學問命名。這門學問建立在 Jakob von Uexküll 的環境界思想之上。包含意義的信號並非只存在於人類的語言中。所有生命都在自己的世界裡互相傳遞信號 [3]。

Jakob von Uexküll 的想法引出了新的問題。義大利學者 Nicola Zengiaro 提出了生成動物符號學的想法。他指出，人工智慧光靠機器模仿動物的信號，是無法理解其中含義的。必須同時觀察動物生活的環境與肢體動作，才能正確理解其中的意義 [4]。

現在，人工智慧正在為這個問題開啟新的大門。人工智慧能透過運算，讀取比人類聽覺更寬廣的聲音範圍。在人類原本以為一片寂靜的地方，其實充滿了聲音。海洋、森林與地底下，都有信號在彼此傳遞。人工智慧幫助我們找出那些隱藏的規則。

這個問題也與人類社會息息相關。創立 Earth Species Project 的 Aza Raskin 表示，科技把人與人分開了。人類也與大自然漸行漸遠。用人工智慧聆聽動物的聲音，就是把這條斷掉的線重新連結起來 [5]。

相關佐證資料

- [1] Bakker, K. (2022). The sounds of life: How digital technology is bringing us closer to the worlds of animals and plants. Princeton University Press. <https://doi.org/10.2307/j.ctv2hnkcc3>
- [2] Bakker, K. (2022). The sounds of life: How digital technology is bringing us closer to the worlds of animals and plants. Princeton University Press. <https://doi.org/10.2307/j.ctv2hnkcc3>
- [3] Sebeok, T. A. (1972). Perspectives in zoosemiotics. Mouton.
- [4] Zengiaro, N. (2025). AI and animal communication: A generative zoosemiotics perspective. Linguistic Frontiers, 8(2). <https://doi.org/10.2478/lf-2025-0021>
- [5] Earth Species Project. (2026, April 2). Animal language processing: An AI convergence in animal communication. Earth Species Project. <https://earthspecies.org/2026/04/02/animal-language-processing-an-ai-convergence-in-animal-communication/>

8.2 超越人類感官的聲音與震動

人類的耳朵有極限。人類只能聽見 20 赫茲到 2 萬赫茲之間的聲音。赫茲是計算聲音在 1 秒內震動幾次的單位。比這個範圍更高或更低的聲音，人類是聽不見的。但那些聲音並沒有消失。人工智慧捕捉到這些聲音，並將它們呈現在人類眼前 [1]。

我們先來看看植物的聲音。長久以來，人們認為植物和聲音沒有關係。但以色列特拉維夫大學的 Yossi Yovel 教授與 Lilach Hadany 教授研究團隊改變了這個想法。缺水或莖部受傷的番茄和菸草發出了超音波。超音波是人類聽不見的高頻聲音。那些聲音以相當於人類交談的音量在空氣中傳播開來。

植物為什麼會發出聲音呢？當水分乾涸時，莖內細小的水分管道中會產生微小氣泡。當這些氣泡破裂時，就會發出喀喀聲。人類的耳朵聽不見，但機器能捕捉得到。

研究人員把這些聲音輸入深度學習人工智慧中。深度學習是能夠自己找出聲音特徵的人工智慧。結果，缺水的聲音和受傷的聲音被分成了不同的群組。這兩種聲音的波形有明顯的不同。波形就像水波起伏一樣，是聲音的形狀。像昆蟲那樣能聽見超音波的動物，就能聽見這種聲音。

植物其實也會聽聲音。Yovel 教授與 Hadany 教授研究團隊觀察了月見草。他們在花朵旁邊播放蜜蜂飛行時發出的聲音。結果，花朵在 3 分鐘內讓花蜜變得更甜了。面對其他雜音，花朵則沒有任何反應。這等於是花瓣扮演了收集聲音的耳朵角色 [2]。

美國密蘇里大學的 Heidi Appel 教授與 Rex Cocroft 教授研究團隊也進行了類似的實驗。他們讓一種名叫阿拉伯芥的小草聽了各種聲音。面對風雨聲，小草很平靜。但是當播放毛毛蟲啃食葉子的震動時，情況就不同了。小草立刻製造出驅趕昆蟲的物質。這表示它能把啃食葉子的震動與其他震動區分開來 [3]。

2026 年發表的一項研究又向前邁進了一步。Yovel 教授與 Hadany 教授研究團隊觀察了蛾類。雌蛾在挑選產卵地點時，會傾聽植物的聲音。蛾避開了缺水植物的聲音，選擇了健康的植物。這是動物對植物聲音產生反應的第一個證據 [4]。

地底與微小生命的世界也是如此。蜜蜂在黑暗的蜂巢裡跳 8 字舞。這支舞是傳達花朵方向與距離的信號。但這個信號不僅僅是用眼睛看的舞蹈。翅膀的震動與沿著蜂巢傳遞的震顫也融合在其中。Tim Landgraf 教授研究團隊利用人工智慧解讀了這支舞蹈。機器人與攝影機觀察蜂巢超過了一個月。人工智慧自動將影像中的舞蹈一一解讀出來 [5]。

老鼠也會用人類聽不見的聲音說話。老鼠在焦慮或心情好時會發出超音波。華盛頓大學的 Kevin Coffey 研究團隊開發了一款名為 DeepSqueak 的人工智慧。這個人工智慧把老鼠的聲音細細拆解並加以分析。結果成功從聲音中解讀出老鼠的心情狀態 [6]。

人類的耳朵聽不到老鼠的聲音。老鼠害怕時會發出 22 千赫茲左右的聲音。心情好時則會發出 50 千赫茲左右的高頻聲音。千赫茲代表 1000 赫茲。DeepSqueak 把這些聲音轉換成圖像，繪製出心情的地圖。

這些發現共同說明了一件事。人類聽不見，並不代表聲音不存在。大地、花草與微小的動物一直都在互相傳遞信號。人工智慧把這些隱藏的聲音展現在人類面前。

相關佐證資料

- [1] Khait, I., Lewin-Epstein, O., Yovel, Y., & Hadany, L. (2023). Sounds emitted by plants under stress are airborne and informative. *Cell*, 186(7), 1328-1336. <https://doi.org/10.1016/j.cell.2023.03.009>
- [2] Veits, M., et al. (2019). Flowers respond to pollinator sound within minutes by increasing nectar sugar concentration. *Ecology Letters*, 22(9), 1483-1492. <https://doi.org/10.1111/ele.13331>
- [3] Appel, H. M., & Cocroft, R. B. (2014). Plants respond to leaf vibrations caused by insect herbivore chewing. *Oecologia*, 175(4), 1257-1266. <https://doi.org/10.1007/s00442-014-2995-6>
- [4] Seltzer, R., Zer Eshel, G., Hadany, L., & Yovel, Y. (2026). Female moths incorporate plant acoustic emissions into their oviposition decision-making process. *eLife*, 14, e104700. <https://doi.org/10.7554/eLife.104700>
- [5] Landgraf, T., et al. (2024). Autonomous tracking of honey bee behaviors over long-term periods with cooperating robots. *Science Robotics*, 9(95). <https://doi.org/10.1126/scirobotics.adn6848>
- [6] Coffey, K. R., Marx, R. G., & Neumaier, J. F. (2019). DeepSqueak: A deep learning-based system for detection and analysis of ultrasonic vocalizations. *Neuropsychopharmacology*, 44(5), 859-868. <https://doi.org/10.1038/s41386-018-0303-6>

8.3 同時解讀多種感官的人工智慧

動物不會只靠單一聲音來說話。聲音、肢體動作與氣味結合在一起，共同產生了意義。像這樣同時使用多種管道的溝通方式，稱為多模態（multimodal）。這是不僅包含聲音，還融合了眼神、肢體動作與氣味的溝通方式 [1]。

如果只看聲音，很容易錯過真正的含義。即使是相同的聲音，隨著肢體動作不同，意思也會改變。人類笑著說同一句話和生氣時說，意思也不一樣。因此，必須同時觀察多種信號，才能真正理解其中的意思。

巴哈馬海域斑點海豚的研究就是一個很好的例子。Denise Herzing 博士與 Thad Starner 教授製作了一台名為 CHAT 的裝置。這台裝置能協助海豚與人類互相傳遞聲音。但光靠聲音是無法解開對話內容的。還必須同時記錄海豚與人眼神交會的瞬間，以及身體擺動的角度。必須同時觀察多種信號，才能明白真正的意思。

烏鴉的研究也證明了多種信號的重要性。美國華盛頓大學的 John Marzluff 教授研究團隊觀察了烏鴉。研究團隊戴上特製面具抓捕烏鴉後又放走牠們。之後，當戴著相同面具的人出現時，烏鴉大聲叫喚並群聚過來。烏鴉記住了欺負自己的人的面孔。當那個人走來時，牠們會發出警告叫聲，向同伴告知危險 [2]。

這份記憶維持了很久。烏鴉過了 2 年多都沒有忘記那張臉。光聽聲音很難發現這樣的事實。必須把烏鴉的聲音和行為放在一起看，才能解讀出其中的含義。

為了更廣泛地展開這類研究，Earth Species Project 出面採取了行動。這個團體是致力於用人工智慧解開動物語言的研究團隊。他們主張擺脫只看聲音的方法。因為貼在動物身上的感測器能收集到各種資料。感測器可以測量聲音、動作以及水深等數值 [3][4]。

貼在動物身上的感測器能測量很多東西。它會收集游動距離、水深以及心跳節奏。把這些資料加到聲音中，意義就會變得更加清楚。這樣就能分辨出那是肚子餓的叫聲，還是害怕時的叫聲。人工智慧會依照時間順序，同時解讀這些不同的資料。

Earth Species Project 在 2026 年的文章中介紹了一個新名詞。那就是動物語言處理（Animal Language Processing）。這個詞最早在 2024 年的發表中被使用。這種方法不會硬把人類語言的文法套在動物身上。無論是聲音還是肢體動作，它都把各種信號視為有規則的資料來進行分析。這種態度不以人類語言為標準，而是試圖如實解讀動物的世界。

協助這項研究的資料庫也誕生了。Earth Species Project 在 2026 年公開了名為 alp-data 的工具。他們把鳥類、鯨魚和昆蟲的 35 種以上聲音資料收集在同一個地方。目前以聲音資料為主，未來還計畫增加動作與影像資料。這些資料公開給所有人，讓大家都能用相同的資料進行研究。這樣一來，研究成果才能快速累積。

相關佐證資料

[1] Herzing, D. L., Starner, T., et al. (2025). CHAT Jr: A wearable bioacoustics system for two-way communication research between humans and cetaceans. NeurIPS 2025. <https://neurips.cc/virtual/2025/131501>

[2] Marzluff, J. M., Walls, J., Cornell, H. N., Withey, J. C., & Craig, D. P. (2010). Lasting recognition of threatening people by wild American crows. *Animal Behaviour*, 79(3), 699-707. <https://doi.org/10.1016/j.anbehav.2009.12.022>

[3] Earth Species Project. (2026, April 2). Animal language processing: An AI convergence in animal communication. Earth Species Project. <https://earthspecies.org/2026/04/02/animal-language-processing-an-ai-convergence-in-animal-communication/>

[4] Earth Species Project. (2026, July 22). Introducing alp-data: A shared data layer for animal language processing. Earth Species Project. <https://earthspecies.org/2026/07/22/introducing-alp-data-a-shared-data-layer-for-animal-language-processing/>

8.4 抹香鯨的聲音字母與人工智慧

抹香鯨用喀嗒聲來說話。許多短促的喀嗒聲聚集在一起，被稱為 coda。coda 是抹香鯨彼此交流的短促喀嗒聲組合。Project CETI 正在研究這種 coda。Project CETI 是一項試圖用人工智慧解開鯨魚語言的研究計畫 [1]。

2024 年，Project CETI 與美國麻省理工學院（MIT）研究團隊有了重大發現。他們用人工智慧分析了超過 8700 段鯨魚聲音資料。結果發現，coda 是一套具有規則的聲音系統。研究團隊把這套規則稱為抹香鯨聲音字母 [2]。

這套字母包含了四個要素。第一個是喀嗒聲之間維持固定拍子的節奏。第二個是整體的拍速。第三個是根據對話流動把速度放慢或加快的彈性速度。第四個是在末端再加上一聲喀嗒聲的裝飾音。將這些要素結合，就會形成各種不同的 coda。研究團隊光是經常出現的組合就找出了 143 種以上。這和人類語言中有子音與母音的結構非常相似。

2025 年的研究又更進一步。美國柏克萊大學與 Project CETI 的 Gašper Beguš 研究團隊仔細觀察了 coda。喀嗒聲中竟然包含著類似人類母音的聲音。研究團隊找到了兩種不同的鯨魚母音。就像人類發出的「啊」和「依」一樣，是彼此不同的聲音。這代表抹香鯨就像人類一樣，能將聲音進行細緻的劃分 [3]。

coda 的形式會隨著對話情境而改變。後面的鯨魚會配合前面鯨魚發出的聲音來改變自己的聲音。這就像人類互相搭腔、一來一往地交談一樣。透過龐大的資料與運算分析，這種模式更清楚地展現了出來。

Project CETI 在 2026 年也傳來了新消息。研究團隊詳細記錄了抹香鯨生產的畫面。他們收錄了超過 6 小時的水下聲音，同時也收集了從空中拍攝的影像。從這些資料中，展現出鯨群合力協助幼鯨誕生的景象。在人類與猴子這類群體之外的動物中，這是首次以數據證實這種合作照料行為的案例 [4]。

要把鯨魚的語言翻譯成人類的語言並不容易。因為沒有能對照這兩種語言的字典。Beguš 研究團隊正在尋找新方法。那就是在沒有標準答案對照表的情況下，由機器自己找出聲音的規則來推測含義的方法 [5]。

這樣的研究如果沒有人工智慧，很難進行。光靠人類的耳朵，很難找出那麼多喀嗒聲的規律。人工智慧能快速比對無數聲音，找出藏在裡面的規律。

有一件事絕不能忘記。就算找到了 coda 的規律，也不代表完全明白其中的意思。了解聲音的結構，和明白其中的意思並不相同。抹香鯨究竟在說些什麼，目前還沒有得到證實。這項研究才剛開啟大門。

相關佐證資料

- [1] Project CETI. (2024, May 7). Sperm whale phonetic alphabet proposed for the first time. Project CETI. <https://www.projectceti.org/blog-posts/sperm-whale-phonetic-alphabet-proposed-for-the-first-time>
- [2] Sharma, P., Gero, S., Payne, R., Gruber, D. F., Rus, D., Torralba, A., & Andreas, J. (2024). Contextual and combinatorial structure in sperm whale vocalisations. *Nature Communications*, 15, 3617. <https://doi.org/10.1038/s41467-024-47221-8>
- [3] Beguš, G., Sprouse, R. L., Leban, A., Silva, M., & Gero, S. (2025). Vowel- and diphthong-like spectral patterns in sperm whale codas. *Open Mind*, 9. <https://doi.org/10.1162/OPMI.a.252>
- [4] Project CETI. (2026, March). Studies documenting rare sperm whale birth and ancient cooperative care released. *EurekAlert!* <https://www.eurekalert.org/news-releases/1120823>
- [5] Beguš, G., Leban, A., & Gero, S. (2026). Approaching an unknown communication system by latent space exploration and causal inference. *Royal Society Open Science*, 13(8), 250829. <https://doi.org/10.1098/rsos.250829>

8.5 符號學界線與跨物種外交

人工智慧為人類打開了寬廣的感官世界。那麼，大海裡生命的聲音是用什麼收集的呢？人類會使用水下聽音器。水下聽音器就是捕捉水裡聲音的麥克風。人們也會在鯨魚或海豚身上貼上小型記錄裝置，並在海洋各處施放捕捉聲音的浮標感測網路。人工智慧便會讀取這些收集到的聲音 [1]。

針對未來即將出現的技術，也有人提出了一些想法。那就是把即將到來的第六代行動通訊網路(6G)當作感測器來使用。6G 是比現在快得多的無線通訊網路。這個想法認為，只要讀取通訊網路的電波碰撞到物體後的情形，就能知道物體的動作。不過，這目前還只是提案階段的想法。而且電波幾乎無法穿透海水。因此，這個方法主要打算用來觀察陸地上與空氣中的情況，並不適合用來收集大海裡的聲音 [2]。

在這裡出現了一條重要的界線。所謂的符號學界線，就是區分信號與意義究竟能解讀到哪裡。假設人工智慧收集了動物的所有信號，難道我們就可以把這些信號硬套進人類語言的框架裡嗎？例如，把鯨魚的聲音一一對應到人類的音樂上，這樣真的可以嗎？這種配對只是把人類的想法強加在動物身上而已 [3]。

讓我們來看一個錯誤配對的例子。有些人可能會說，10 隻海豚當中有 3 隻會使用名詞；也可能把鯨魚的喀答聲一一對應到人類的音樂。然而，這是硬把人類的框架套上去。動物的聲音，必須在動物的世界裡去解讀。

解讀聲音這件事伴隨著倫理問題。學者們指出，人類不能隨意模仿動物的語言。如果用錯了，可能會被用來欺騙或騷擾動物。因此，學者們主張應該先建立規則 [4]。

尤克斯庫爾的環境界想法，在今天依然是一盞指路明燈。動物和植物活著，並不是為了成為人類字典裡的翻譯詞。牠們的信號，是經歷漫長歲月流傳下來、屬於牠們自己的文化。現代研究人員繼承了尤克斯庫爾的精神，認為不能隨便用人工智慧去模仿動物。人工智慧的目標，並不是做出會說話的動物玩具。

有一種把動物聲音錄下來、再放給牠們聽的實驗，這被稱為再生(playback)。這種實驗是利用喇叭播放錄下來的聲音，觀察動物會有什麼反應。如果隨便使用這個工具，可能會打亂動物的世界。被回放聲音嚇到的動物，可能會分不清真正的信號和虛假的信號。因此，研究人員提醒大家一定要小心。

我們該走的道路不是征服，而是傾聽。也就是人類從地球唯一詮釋者的位置走下來，守護動物的聲音，讓這些聲音能在牠們的土地與海洋中長久延續。研究人員把這稱為種間外交(interspecies diplomacy)。種間外交是指人類與其他生命為了不互相傷害、和平相處所訂下的約定 [5]。

傾聽並不只是默默地聽，還延伸到守護動物所居住的海洋與森林。想要守護聲音，就必須一同守護發出聲音的地方。這就是種間外交的真正意義。

相關佐證資料

- [1] Bakker, K. (2022). The sounds of life: How digital technology is bringing us closer to the worlds of animals and plants. Princeton University Press. <https://doi.org/10.2307/j.ctv2hnkcc3>
- [2] Ericsson. (2025). Integrated sensing and communication (ISAC) explained. Ericsson. <https://www.ericsson.com/en/6g/isac>
- [3] Zengiaro, N. (2025). AI and animal communication: A generative zoosemiotics perspective. Linguistic Frontiers, 8(2). <https://doi.org/10.2478/lf-2025-0021>
- [4] Alcantara, L., & Andrews, K. (2026). Can we talk to the animals? The ethics of using machine learning to decode animal communication. Topoi. <https://doi.org/10.1007/s11245-026-10409-2>
- [5] Earth Species Project. (2026, April 2). Animal language processing: An AI convergence in animal communication. Earth Species Project. <https://earthspecies.org/2026/04/02/animal-language-processing-an-ai-convergence-in-animal-communication/>

第9章 欺瞞與保育的十字路口

9.1 假聲音引發的混亂

現在，人工智慧已經能非常逼真地模仿動物的聲音。不管是鳥叫聲、鯨魚歌聲還是大象叫聲，它都能製作出來。就像人類學習說話一樣，人工智慧也是聽著聲音來學習規則。學會之後，它就能自己創造出新的聲音。這項技術可以用來幫助動物。但如果用得不对，就會嚴重擾亂大自然 [1]。

在動物研究中，有一種叫做「回放 (playback)」的方法。這是一種把錄下來的聲音重新播給動物聽，並觀察牠們反應的實驗。以前，研究人員只會使用人類錄下的真實聲音。過去的錄音混雜著雜音，動物有時能察覺出來。現在，人工智慧能做出非常流暢逼真的假聲音。因此動物更容易受騙。這些假聲音會讓動物在求偶與照顧幼崽時陷入混亂 [7]。

許多國家的學者都對這個危險提出了警告。Christian Rutz 教授和他的同事們最具代表性。他們認為假聲音會改變野生動物的行為。用聲音欺騙動物的這種行為被稱為「生態欺瞞」。意思是在大自然中欺騙動物，進而改變牠們的行為。聽到了虛假的求偶聲，鳥兒就會徒勞地尋找伴侶。結果浪費了體力和時間，連照顧雛鳥的機會都被剝奪了。

Karen Bakker 教授也長期提醒大家注意這個危險。Bakker 教授是帶領大家利用數位技術聆聽動物聲音的學者。她在 2022 年出版的著作《生命之聲》(The Sounds of Life) 中表達了這個理念。Bakker 教授於 2023 年去世。她留下的警告至今依然延續著 [2]。

人類常把自己的心情投射到動物的聲音上。我們很容易把鯨魚的歌聲當作情歌。但目前還沒有人知道那些聲音真正的含義。隨意猜測就會犯錯。學者們提醒我們也要小心這種習慣。

長期研究大象的 Joyce Poole 博士分享過一個故事。這件事發生在肯亞的安波塞利 (Amboseli)。研究團隊把一隻已經過世的象媽媽叫聲播給象群聽。大象們聽到想念的聲音後，四處奔跑尋找。其中，過世象媽媽的女兒不斷呼喚著媽媽，等待著回應。可是聲音的主人卻在哪裡都找不到。大象甚至能記住死去家人的聲音。大象會用人類耳朵聽不太到的低沉聲音傳遞到很遠的地方。多虧了這點，牠們能分別記住幾十隻大象的聲音。一個聲音就這樣深深震撼了整個象群的心 [3][4]。

面對鯨魚也必須非常謹慎。座頭鯨是會唱歌的鯨魚。當一片海域出現新歌時，就會像流行歌曲一樣傳播開來。澳洲外海的歌聲甚至會傳播到南太平洋彼岸。這首歌曲會在幾個月到幾年的時間裡慢慢傳播。慢慢跨越約 6,000 公里的距離。這是一段相當於從首爾到非常遙遠國家的遙遠距離。小鯨魚也會聽著成年鯨魚的歌聲跟著學。這就跟流行歌透過智慧型手機傳播的樣子很像。除了人類之外，像這樣互相學習唱歌的現象非常罕見。如果把人工智慧製作的假鯨魚歌聲播到大海裡，就會產生嚴重的問題。延續已久的歌唱傳統可能會被擾亂。這等於是在我們解開動物語言之前，就先把牠們的文化破壞掉了 [5]。

因此研究人員在進行這項實驗時也非常小心。只在絕對必要時才進行短暫的實驗。實驗結束後，還會確認動物是否安好。同樣的技術，根據使用方式的不同，既可能成為幫助，也可能造成傷害。

相關佐證資料

- [1] Rutz, C., Bronstein, M., Raskin, A., Vernes, S. C., Zacarian, K., & Blasi, D. E. (2023). Using machine learning to decode animal communication. *Science*, 381(6654), 152-155. <https://doi.org/10.1126/science.adg7314>
- [2] Bakker, K. (2022). *The sounds of life: How digital technology is bringing us closer to the worlds of animals and plants*. Princeton University Press.
- [3] Poole, J. (1996). *Coming of age with elephants: A memoir*. Hyperion.
- [4] McComb, K., Moss, C., Sayialel, S., & Baker, L. (2000). Unusually extensive networks of vocal recognition in African elephants. *Animal Behaviour*, 59(6), 1103-1109. <https://doi.org/10.1006/anbe.2000.1406>
- [5] Garland, E. C., Goldizen, A. W., Rekdahl, M. L., et al. (2011). Dynamic horizontal cultural transmission of humpback whale song at the ocean basin scale. *Current Biology*, 21(8), 687-691. <https://doi.org/10.1016/j.cub.2011.03.019>

[7] Küçükuncular, A. (2025). Ethical implications of AI-mediated interspecies communication. *AI and Ethics*, 5, 6379-6391. <https://doi.org/10.1007/s43681-025-00828-z>

9.2 無所不在的監視網與動物的隱私

以前，深山密林和廣闊的大海是動物的避難所。人類的耳目觸及不到那裡。現在不一樣了。小巧聰明的聲音感測器被放置在森林與海洋的各個角落。這個裝置 24 小時不斷聆聽聲音。人工智慧會立即分析那些聲音。設置在森林與海洋中的這種自動錄音裝置，被稱為「被動聲學監測網」[6]。

現在，野生動物隨時都處於被注視的目光之下。就像是一座看不見的監視塔始終開啟著。共同創立 Earth Species Project 的 Aza Raskin 指出了這項轉變。他表示，多虧了人工智慧，我們才能不間斷地監測森林與海洋。與此同時，他也警告必須制定規則，免得這股力量變成束縛動物的嚴密監視[8]。

這種裝置也能用在好事情上。鯨魚經常因為撞上大型船隻而失去生命。利用聲音感測器找出鯨魚所在的位置並改變航道。這樣大型船隻就不會撞到鯨魚了。它還能捕捉到森林中偷偷砍伐樹木的電鋸聲。藉此保護瀕危動物並遏止盜獵行為。

但如果落入壞人手中，它就會變成武器。就像一把刀既能用來做菜，也能傷人一樣。有人可能會模仿虛假的求偶聲或覓食聲來誘騙動物。雖然目前這種情況尚未普遍發生，但學者們十分擔憂這個危險。原本為了幫助動物而開發的技術，竟可能反過來被用來捕捉動物[7]。

動物的隱私也是一個問題。隱私是不想讓別人看見的個人專屬領域。Karen Bakker 教授曾說，我們應該把鯨魚和海豚視為擁有權利的存在。有時人類甚至會暗中偷聽媽媽和寶寶之間輕聲細語的對話。還隨意收集那些聲音並公開給任何人使用。學者們認為這應該被視為侵犯隱私。這就像別人偷偷看了我的日記還上傳到網路上是一樣的道理[2]。

長期守護森林與海洋的原住民也有他們的權利。大公司有時會偷偷拿走來自他們土地上的聲音。並用這些聲音來訓練人工智慧。聲音資料也是那片土地的財產。未經主人同意是不能拿走的。當地人對產自自己土地的資料擁有自主決定權，這項權利被稱為「原住民數據主權」。

網路上還流傳著動物對抗監視的傳聞。有傳言說抹香鯨群曾同時停止發聲。也有傳言說烏鴉發出鳥叫聲來混淆人工智慧。但這些事情目前尚未得到科學證實。這只是網路影片中的故事，千萬不能當作事實來相信。

好的用途和壞的用途都源自同一項技術。因此，由誰抱持著什麼樣的心態去使用才是關鍵。動物也有屬於自己、需要被守護的專屬空間。

相關佐證資料

[2] Bakker, K. (2022). *The sounds of life: How digital technology is bringing us closer to the worlds of animals and plants*. Princeton University Press.

[6] Earth Species Project. (2024). Introducing NatureLM-audio: An audio-language foundation model for bioacoustics. <https://www.earthspecies.org/blog/introducing-naturelm-audio-an-audio-language-foundation-model-for-bioacoustics>

[7] Küçükuncular, A. (2025). Ethical implications of AI-mediated interspecies communication. *AI and Ethics*, 5, 6379-6391. <https://doi.org/10.1007/s43681-025-00828-z>

[8] Faunalytics (Hanes, A.). (2025, October 30). The ethics of using AI to talk to animals. Faunalytics. <https://faunalytics.org/the-ethics-of-using-ai-to-talk-to-animals/>

9.3 藏在聲音裡的人類標記：音訊浮水印

我們需要分辨假聲音的方法。學者和工程師們提出了一種解法。那就是在人工智慧製作的聲音中隱藏「標記」。這種方法被稱為「音訊浮水印 (audio watermarking)」[9]。

浮水印是在音訊檔案中嵌入看不見的標記的技術。這個標記是隱藏在音訊檔案裡的電腦訊號。人類很難察覺到這種訊號。但檢測程式卻能立即讀取該訊號。Google 開發的 SynthID 就是這種標記的一個例子。這種標記會分散嵌入在聲音的多個位置。因此很難被剪掉或抹除。就算調小音量或改變播放速度，標記依然會保留下來。

這個標記能告訴我們聲音的來源。機器能藉此確認某個聲音是不是人類製造的假聲音。這樣一來，我們就能將真實的動物聲音和假聲音分開處理。分辨真假是守護大自然的第一步。如果沒有標記，就很難知道什麼才是真實的 [1]。

這個標記還能用來保護研究資料。森林和大洋中安裝了數以千計收集聲音的裝置。如果裡面混進了假聲音，研究成果就會變得一團糟。把真的鳥鳴聲和機器製造的鳥鳴聲混在一起統計，答案就會出錯。有了標記，就能把假聲音過濾掉。如此一來，研究資料就能令人信賴並安心使用 [7]。

也有人提出可以將它用於防範盜獵。方法是在研究用聲音中加入標記，以確認聲音的來源。這樣就能區分出經認證的研究聲音與來路不明的假聲音。不過這種方法也有其限制。如果壞人製作了沒有標記的聲音，就無法把它過濾出來。如果使用抹去標記的聲音，也很難查出來。因此，單靠一項技術是無法完全阻止盜獵的 [8]。

這種方法目前還處於提議階段。雖然是個好點子，但不會憑空實現。只有當所有國家與企業都承諾加入標記時，它才能發揮力量。只要有一處流出沒有標記的假聲音，就會產生漏洞。這就像漁網只要破了一個洞，魚兒就會溜走一樣。因此光靠技術是不夠的。我們需要大家共同遵守的規則。

這項技術也能應用在人類製作的其他聲音上。也就是在人工智慧生成的聲音或歌曲中加入標記。這樣就能減少被假聲音欺騙的情況。原本為了保護動物而開發的方法，也同時保護了人類。最近網路上充斥著人工智慧製作的歌曲與人聲。因此這類標記在未來會變得更有用處。

相關佐證資料

- [1] Rutz, C., Bronstein, M., Raskin, A., Vernes, S. C., Zacarian, K., & Blasi, D. E. (2023). Using machine learning to decode animal communication. *Science*, 381(6654), 152-155. <https://doi.org/10.1126/science.adg7314>
- [7] Kùçükuncular, A. (2025). Ethical implications of AI-mediated interspecies communication. *AI and Ethics*, 5, 6379-6391. <https://doi.org/10.1007/s43681-025-00828-z>
- [8] Faunalytics (Hanes, A.). (2025, October 30). The ethics of using AI to talk to animals. Faunalytics. <https://faunalytics.org/the-ethics-of-using-ai-to-talk-to-animals/>
- [9] Google DeepMind. (2024). SynthID: Identifying AI-generated content. <https://deepmind.google/models/synthid/>

9.4 建立人類與動物之間的規則

用人工智慧處理動物的聲音需要規則。因為不能任由有權勢的人隨心所欲地撼動大自然。聲音會跨越國界傳播到遙遠的地方。因此只靠一個國家遵守是不夠的。學者們主張制定多個國家共同遵守的規則。意思是這是人類與動物之間必須遵守的約定 [10]。

美國紐約大學法學院有一個名為「超越人類之生命（More Than Human Life）」的研究團體。簡稱為 MOTH。這個團體成立於 2022 年。由研究法律的學者與科學家共同組成。這個團體主張，不該把動物當成單純的資料，而應將牠們視為擁有各自生命的存在。他們提出了一套包含四大承諾的守則。這四項承諾分別是：做好準備、共同研商、防止傷害以及切實守護。

第一，是事前審查的規則。有些實驗會讓人工智慧與野生動物進行聲音互動。在進行此類實驗之前，必須先評估其風險。確認是否會危害動物的求偶與照顧幼崽的文化。只有在確認安全之後才允許進行實驗。這就像在動工之前必須先通過安全檢查一樣。

第二，是防止濫用的規則。不允許任何人隨便使用模仿動物聲音的技術。他們主張徹底禁止將此技術用於打獵。只有通過審查的單位才能謹慎使用。這就像不能把危險工具隨處亂放一樣。同時也主張防止核心技術被隨意用於營利牟利。

第三，是防止誇大報導的規則。像「人工智慧已經做出能與動物自由對話的翻譯機」這類誇大不實的新聞時有所聞。學者們認為這種新聞是有害的。因為這會讓人們把動物當成玩具一樣對待。因此他們主張應坦率說明這項技術目前還無法做到的事 [11]。

規則並不是要阻礙研究。而是幫助研究走在正軌上的柵欄。就像操場上畫好界線，大家才能安心奔跑玩耍一樣。制定規則需要匯集許多人的聲音。不僅僅是科學家。懂法律的人、生活在森林與海洋的原住民、愛護動物的人都

要一起參與決定。透過多元的視角審視，疏漏之處就會減少。規則的目標只有一個。那就是尊重動物，而非支配動物 [7][8]。

相關佐證資料

- [7] Küçükuncular, A. (2025). Ethical implications of AI-mediated interspecies communication. *AI and Ethics*, 5, 6379-6391. <https://doi.org/10.1007/s43681-025-00828-z>
- [8] Faunalytics (Hanes, A.). (2025, October 30). The ethics of using AI to talk to animals. Faunalytics. <https://faunalytics.org/the-ethics-of-using-ai-to-talk-to-animals/>
- [10] Rodríguez-Garavito, C. (2025). Listening to the more-than-human world: Legal and ethical principles for nonhuman animal communication technologies. More-Than-Human Life (MOTH) Program, NYU School of Law. <https://www.law.nyu.edu/news/ideas/cesar-rodriguez-garavito-animal-communication-ai>
- [11] Kluskens, C. (2025). The responsible media representation of artificial intelligence for non-human animal communication. NeurIPS 2025 Workshop on AI for Non-Human Animal Communication.

9.5 最新消息與我們所處的十字路口

到了 2026 年，動物語言研究依然進展飛速。Earth Species Project 將這項研究統整為「動物語言處理（Animal Language Processing）」。這就像電腦理解人類語言一樣，透過大規模收集並分析動物的聲音。它擺脫了以往一次只觀察單一物種的方式。改為同時檢視許多種動物聲音的做法。這並不是他們第一次發明這個名詞，而是將過去分散的研究整合在一起。這項發表於 2026 年 4 月公布 [12]。

同年 7 月，他們還建立了一個全新的資料庫。名稱為「alp-data」。這是為了讓眾多研究人員能共享動物聲音資料而匯整的地方。當中收錄了超過 35 個資料集。共享資料能讓研究進展得更快。這就像班上同學互相分工共享用具，事情就能很快完成一樣 [13]。

這項研究的基礎是「NatureLM-audio」。它是專門為了動物聲音打造的大型人工智慧。這個人工智慧只要聽聲音，就能猜出是哪種動物。它還能數出聲音裡有幾隻動物。也能分辨是幼鳥還是成鳥等成長階段。它還展現出能認出從沒學過的動物聲音的本領。這樣的人工智慧能在幾分鐘內，完成人類需要花好幾天的工作。因此研究人員能更快地檢視聲音 [6]。

鯨魚的研究也有了進展。Project CETI 收集了抹香鯨的聲音片段 coda。coda 是一連串「喀、喀、喀」的咔嗒聲組合。研究團隊分析了超過 8,700 個 coda。結果找到了構成聲音的基本片段。研究人員把這些基本片段分成四種。研究團隊稱之為「抹香鯨的聲音字母」。這就像人類文字區分子音和母音一樣。鯨魚會改變聲音的速度與節拍來表達不同的意思。牠們也會加入更多咔嗒聲來改變意思。學者們還找到了像人類母音一樣組合聲音的規則。這並不是說鯨魚會像人類一樣發出「啊」或「一」的聲音。而是一種比喻，說明就像人類組合子音和母音來創造各種詞彙一樣，鯨魚也會按照規則來組合聲音 [14][15][16]。

人工智慧也用來保護逐漸消失的動物。夏威夷烏鴉就是一個例子。這種鳥被長期圈養後回到了野外。研究人員正在觀察牠們在這段期間，是否忘記了群體中的某些叫聲。他們利用人工智慧比對過去和現在的聲音。這是為了幫牠們找回忘記的叫聲。Earth Species Project 也同時研究烏鴉、白鯨與大象的聲音。當我們一起觀察多種動物時，就能看出聲音共通的規則。

我們現在正站在岔路口。一條路是欺騙。這是一條用假聲音欺騙並擾亂動物的路。另一條路是保育。這是一條靜靜傾聽並守護動物的路。這兩條路都建立在同一項技術之上。要走哪一條路，取決於人類的選擇。人工智慧只是一項工具。聆聽聲音的力量越大，守護這股力量的規則也必須越完備。力量變強了，就必須一起學習如何駕馭那股力量。當我們放下傲慢、敞開雙耳時，人類與動物就能並肩生活在一起 [10][1]。

相關佐證資料

- [1] Rutz, C., Bronstein, M., Raskin, A., Vernes, S. C., Zacarian, K., & Blasi, D. E. (2023). Using machine learning to decode animal communication. *Science*, 381(6654), 152-155. <https://doi.org/10.1126/science.adg7314>
- [6] Earth Species Project. (2024). Introducing NatureLM-audio: An audio-language foundation model for bioacoustics. <https://www.earthspecies.org/blog/introducing-naturelm-audio-an-audio-language-foundation-model-for-bioacoustics>

- [10] Rodríguez-Garavito, C. (2025). Listening to the more-than-human world: Legal and ethical principles for nonhuman animal communication technologies. More-Than-Human Life (MOTH) Program, NYU School of Law.
<https://www.law.nyu.edu/news/ideas/cesar-rodriguez-garavito-animal-communication-ai>
- [12] Earth Species Project. (2026, April 2). Animal language processing: An AI convergence in animal communication.
<https://earthspecies.org/2026/04/02/animal-language-processing-an-ai-convergence-in-animal-communication/>
- [13] Earth Species Project. (2026, July 22). Introducing alp-data: A shared data layer for animal language processing.
<https://earthspecies.org/2026/07/22/introducing-alp-data-a-shared-data-layer-for-animal-language-processing/>
- [14] Project CETI. (2024). Sperm whale phonetic alphabet proposed for the first time.
<https://www.projectceti.org/blog-posts/sperm-whale-phonetic-alphabet-proposed-for-the-first-time>
- [15] Sharma, P., Gero, S., Payne, R., et al. (2024). Contextual and combinatorial structure in sperm whale vocalisations. Nature Communications, 15, 3617. <https://doi.org/10.1038/s41467-024-47221-8>
- [16] National Geographic. (2024). We're one step closer to understanding the sperm whale 'alphabet'.
<https://www.nationalgeographic.com/animals/article/sperm-whale-vowel-patterns-communication-discovery>

第10章 非人類人格體與機器輔助權利

10.1 從沉默到作證

人類的法律在很長一段時間裡，只把權利給人類。由許多人聚在一起組成的團體，像是公司，也擁有權利。這樣的組織被稱為法人（法人）。法人是法律創造出來的虛擬人。公司或學校都屬於這一類。法人雖然不是真正的人，但在法律面前會受到像人一樣的對待 [3]。

過去的法律堅信一件事。那就是只有明白自己的義務、能夠許下約定的一方，才可以擁有權利。動物沒辦法用言語為自己辯護。因此，法律並不把動物看作權利的主人。而是把動物當成人類可以買賣的物品和財產。

這道古老的高牆現在動搖了。因為人工智慧與生物聲學攜手合作。生物聲學是研究動物發出聲音的學問。這兩項技術結合在一起，開始用數字呈現動物的心聲 [2]。

以前，人類只用耳朵聽動物的聲音。所以錯過了很多聲音。現在，人們在海裡廣泛布置了錄音設備。Project CETI (Cetacean Translation Initiative) 正在做這件事。翻譯成中文就是鯨魚語言翻譯計畫。這項計畫長期收集抹香鯨的聲音。人工智慧會逐一仔細分析這些收集來的聲音。電腦甚至能找出人類忽略的規律 [4]。

抹香鯨會互相傳送名為 coda 的聲音片段。coda 是把短促的喀嗒聲連續接在一起的聲音。裡面包含標示群體的 identity coda。這是同一個家族共同使用的標記般的聲音。抹香鯨用這種聲音來認出自己的同伴。很難說這和人類的個人名字完全一樣。它更像是全班一起喊的班級口號。

抹香鯨的聲音裡藏著規律。牠們把幾個聲音單位組合在一起，形成各種不同的形式。人類的語言也以相同的方式運作。我們把幾個子音和母音組合起來，創造出無數的詞彙。只不過，每一個 coda 究竟代表什麼意思，目前仍在研究中。聲音的結構雖然已經揭曉，但意義還沒有完全解開。

研究人類大腦的學者對語言提出了新的解釋。那就是麻省理工學院 (MIT) 的 Evelina Fedorenko 教授研究團隊。他們在 2024 年的學術期刊《Nature》發表了研究成果。他們認為，語言並不是產生深度思考的必要條件。這代表語言只是把內心想法傳達給別人的工具。這說明了思考和說話是兩回事。不能因為不會說話，就認為沒有思想 [1]。

這種想法動搖了法律長久以來的固執觀念。不能因為動物不會說人類的語言，就忽視牠們的智慧。動物其實也一直在用自己的方式，互相傳遞精密的聲音。人工智慧協助我們找出了人類過去沒能發現的聲音規律。

這裡必須把兩件事分開來看。知道聲音中存在規律，與修改法律是兩回事。即使科學發現了新事實，法律也不會立刻改變。必須有社會共識、長期的討論以及新的判例，法律才會有所行動。不過，一旦證實動物的聲音是具有意義的溝通，法律也會隨之改變。法律就難以繼續把動物單純當成物品了。

相關佐證資料

- [1] Fedorenko, E., Piantadosi, S. T., & Gibson, E. A. F. (2024). Language is primarily a tool for communication rather than thought. *Nature*, 630, 575-586. <https://doi.org/10.1038/s41586-024-07522-w>
- [2] Sharma, P., Gero, S., Payne, R., Gruber, D. F., Rus, D., Torralba, A., & Andreas, J. (2024). Contextual and combinatorial structure in sperm whale vocalisations. *Nature Communications*, 15, 3617. <https://doi.org/10.1038/s41467-024-47221-8>
- [3] Rodríguez-Garavito, C., Gruber, D. F., Nemeth, A. O., & Beguš, G. (2025). What if we understood what animals are saying? The legal impact of AI-assisted studies of animal communication. *Ecology Law Quarterly*, 52(1). <https://doi.org/10.15779/Z383X83N5Q>
- [4] Project CETI. (2024, May 7). Sperm whale phonetic alphabet proposed for the first time. Project CETI. <https://www.projectceti.org/blog-posts/sperm-whale-phonetic-alphabet-proposed-for-the-first-time>

10.2 NYU MOTH 計畫與機器輔助權利

有一個核心機構正努力把這種改變落實在法律上。那就是紐約大學 (NYU) 法學院的超越人類生命計畫。英文稱為 MOTH (More Than Human Life)。這項計畫由環境法學者 César Rodríguez-Garavito 教授領導 [1]。

Garavito 教授與多位學者一起進行研究。加州大學柏克萊分校（UC Berkeley）的 Gašper Beguš 博士也參與其中。他們提出了一個嶄新的想法。那就是把人工智慧解析出來的動物聲音，當作法庭上的證據。這被稱為機器輔助權利（Machine-Assisted Rights）。這還不是廣泛使用的詞彙，而是他們最新提出來的倡議。

用簡單的話來解釋機器輔助權利，就是這樣。人工智慧去解讀動物的溝通和生存環境。用數字來衡量動物因為人類而遭受的損害。然後代替動物把這些數字作為證據提交給法庭。這等於是扮演不會說話的動物的翻譯官兼代理人。這就像小朋友無法自己提出文件時，由大人代替協助一樣。

這種方式試圖繞開法律長久以來的限制條件。過去的法律只把提出訴訟的資格賦予人類。這種提出訴訟的資格，就叫做原告適格。動物沒辦法取得這個資格。因此學者們想把人工智慧測量出的損害當作證據。藉此替動物敲開法庭的大門。

Beguš 教授的研究團隊有了重大發現。他們在抹香鯨的聲音中找到了共振峰（formant）結構。共振峰是聲音中能量聚集、聽起來特別清晰響亮的頻段。這很像人類改變嘴型發出母音時產生的現象。研究團隊在 2025 年的發表中指出，抹香鯨的 coda 中存在 a coda 與 i coda。它們就像人類的母音一樣，是彼此不同的聲音 [2]。

然而，人工智慧重現的聲音地圖也揭露了令人難過的事實。大型船隻的螺旋槳聲和資源探勘的爆炸聲充斥著海洋。這種聲音被稱為人類製造的海洋噪音。2021 年發表於學術期刊《Science》的一項研究探討了這個問題。內容指出，工業發展之後，海洋變得非常吵鬧。也有研究指出，船隻噪音可能會影響抹香鯨的 coda 結構 [3]。

MOTH 計畫主張，可以把這種損害當作法庭上的證據。即使無法完美翻譯聲音的意思也沒關係。他們認為，只要能數字展現抹香鯨群體的溝通受到了多大破壞就可以了。也就是把聲音扭曲變形的圖表本身直接當作證據。他們認為，這樣一來法院就能命令船隻改變航道，或者阻止軍用聲納的使用。這種主張提到了抹香鯨的兩項權利。那就是免於痛苦的權利，以及維護文化的權利 [4]。

這種想法也能強化現有的法律。美國有保護瀕危動物的法律。也有專門保護海洋哺乳動物的法律。人工智慧測量出的受害數值，能成為讓這些法律切實執行的力量。這代表在建立新權利之前，我們就能更強而有力地運用現有法律。

在這裡同樣必須分開來看。知道聲音遭到破壞是一回事，改變法律又是另一回事。目前尚未確認是否有真正把這些證據提交給法庭的訴訟案件。現在還只是學者們提議可以這樣運用的階段。

相關佐證資料

- [1] Rodríguez-Garavito, C., Gruber, D. F., Nemeth, A. O., & Beguš, G. (2025). What if we understood what animals are saying? The legal impact of AI-assisted studies of animal communication. *Ecology Law Quarterly*, 52(1). <https://doi.org/10.15779/Z383X83N5Q>
- [2] Beguš, G., Sprous, C., Lebanidze, T., Gero, S., & Andreas, J. (2025). Vowel- and diphthong-like spectral patterns in sperm whale codas. *Open Mind*, 9, 1849-1874. <https://doi.org/10.1162/OPMI.a.252>
- [3] Duarte, C. M., Chapuis, L., Collin, S. P., et al. (2021). The soundscape of the Anthropocene ocean. *Science*, 371(6529), eaba4658. <https://doi.org/10.1126/science.aba4658>
- [4] Surma, K. (2025, October 29). AI is decoding whales' communications. Could that be a turning point in the push for their rights? *Inside Climate News*. <https://insideclimatenews.org/news/29102025/ai-sperm-whale-communications-legal-rights/>

10.3 加勒比海的證言與原住民的智慧

非人類人格體的討論，並不只停留在西方學者的電腦裡。非人類人格體是指雖然不是人類，但可以受到法律保護的存在。這項討論與守護海洋許久的原住民智慧相遇，獲得了長足的發展 [2]。

2024 年 3 月 28 日，太平洋庫克群島（Cook Islands）的拉羅湯加發生了一件具有歷史意義的大事。毛利族國王與庫克群島的領袖們齊聚一堂。大溪地、東加與夏威夷的領袖們也一同參與。他們承認鯨魚是擁有權利的獨立主體。宣布鯨魚是雖非人類但受法律保護的存在。這份共同宣言的名稱是 He Whakaputanga Moana，也就是大洋宣言。這份宣言於 2026 年 2 月在紐西蘭懷唐伊再次廣為宣傳 [1][4]。

這份宣言是第一份由原住民發起、承認鯨魚是有感知存在的宣言。宣言提到了鯨魚擁有自由游泳的權利，以及在健康海洋中生活的權利。

這份宣言植根於傳統生態知識。所謂傳統生態知識，是把大自然看作共同生活的親族的一種古老智慧。原住民代代相傳這份智慧。他們早已明白，鯨魚的歌聲並不是偶然發出的噪音。他們相信，那些歌聲承載著海洋的記憶與家族的歷史。

現代的人工智慧指出了與這種古老直覺相同的方向。Project CETI 長期觀察加勒比海多米尼克外海的鯨魚。研究團隊利用人工智慧仔細分析這些鯨魚的 coda。只要把幾種特徵組合起來，能夠區分的 coda 數量就會大幅增加。過去人們認為 coda 只有大約二十一種型態。但當解開組合方式之後，可區分的形式居然超過了一百種。這就像人類用少數文字拼出許多詞彙一樣。這也與原住民一直以來相信歌聲承載著意義的想法不謀而合 [3]。

庫克群島的領袖們很久以前就把鯨魚視為親族。把鯨魚當成海洋中的尊長來對待。這樣的心意早在科學出現很久之前就存在了。現在，人工智慧為這樣的心意增添了數字作為根據。

因此，原住民與科學家攜手創造了嶄新的方式。那就是把人工智慧的運算與原住民的海洋知識並列在一起的方法。在這裡，原住民不是輔助研究人員。而是從設計人工智慧開始就共同參與決策的主角。科學家與原住民平等攜手，守護海洋。

這樣的共識並不局限於單一國家之內。因為大洋橫跨了多個國家的海域。鯨魚不分國界，穿梭在寬廣的海洋中。所以單憑一個國家的法律，無法保護所有的鯨魚。大洋宣言正是多個國家與部落攜手合作的第一步。

相關佐證資料

- [1] He Whakaputanga Moana. (2024, March 28). Whale protection and legal personhood declaration. Rarotonga, Cook Islands. <https://www.whaledeclaration.com/>
- [2] Rodríguez-Garavito, C., Gruber, D. F., Nemeth, A. O., & Beguš, G. (2025). What if we understood what animals are saying? The legal impact of AI-assisted studies of animal communication. *Ecology Law Quarterly*, 52(1). <https://doi.org/10.15779/Z383X83N5Q>
- [3] Sharma, P., Gero, S., Payne, R., Gruber, D. F., Rus, D., Torralba, A., & Andreas, J. (2024). Contextual and combinatorial structure in sperm whale vocalisations. *Nature Communications*, 15, 3617. <https://doi.org/10.1038/s41467-024-47221-8>
- [4] Inside Climate News. (2026, February 5). How a 'powerful vision' on whales and oceans could change worldviews. Inside Climate News. <https://insideclimatenews.org/news/05022026/maori-scientist-daniel-hikuroa-legal-rights-of-whales/>

10.4 物種演化的權利

在人工智慧開闢的這條道路盡頭，有著更宏大的想法。那是一種超越命令船隻改道的想法。也就是非人類智慧自主走向未來的權利。部分學者將此稱為物種演化的權利。這項權利尚未寫入法律，目前還是學者們的提議 [1]。

人類的語言經過漫長歲月的發展。從幾種聲音開始，逐漸成熟為複雜的語法。動物的聲音文化也不是停滯不前的標本。它是經歷世代傳承，透過學習、模仿與改變而不斷成長的活資產。各地區的鯨魚歌聲略有不同，原因也正在於此。這就像人類的語言在不同地區會分化出方言一樣 [3]。

觀察抹香鯨媽媽就能明白這一點。鯨魚媽媽會教導剛出生的鯨魚寶寶群體的聲音。鯨魚寶寶會像牙牙學語一樣，用生疏的聲音模仿很久。這很像小嬰兒在學說話時笨拙模仿的樣子。這代表抹香鯨的聲音並不是出生時就銘刻在身上的本能。而是代代相傳的文化遺產。就像人類傳承書籍與歌曲一樣，鯨魚也傳承著聲音。

然而，人類傾倒入海洋的噪音阻礙了這種學習。工業噪音與日夜不息的軍事爆炸聲充斥著海洋。年幼的鯨魚無法聽清楚父母的聲音。因為學習聲音所需要的重要音高頻段被攪亂了。這就像教室太吵鬧時，聽不見老師說話一樣。這不只是傷害單一隻鯨魚。而是切斷了整個群體文化與記憶的脈絡 [2]。

因此，部分生態哲學家與法學者主張一件事。那就是守護動物安全學習並傳承聲音的權利。這時，動物所生活的世界，被稱為用各自感官去體會的世界。蝙蝠感受到的世界與鯨魚感受到的世界彼此不同。他們提議建立一項權利：讓動物不受人類噪音或虛假聲音干擾，能夠自主成長。這目前還不是法律，仍處於學者們討論的階段。

人工智慧製造出的虛假鯨魚聲音也必須小心提防。人類隨意播放的虛假聲音，會混淆幼鯨的學習。因此這項提議也包含了防範假聲音的責任。如果使用不當，再好的技術也會變成新的噪音。翻譯機應該是用來聆聽鯨魚世界的工具。絕不能成為擾亂鯨魚世界的工具。

海洋的噪音只要減少，很快就會平息。2021 年發表於《Science》的研究指出了這一點。因為噪音是在現場發出的，只要減少成因，海洋很快就會安靜下來。只要我們減少商業噪音，鯨魚就能再次清楚聽見彼此的聲音。

相關佐證資料

- [1] Rodríguez-Garavito, C., Gruber, D. F., Nemeth, A. O., & Beguš, G. (2025). What if we understood what animals are saying? The legal impact of AI-assisted studies of animal communication. *Ecology Law Quarterly*, 52(1). <https://doi.org/10.15779/Z383X83N5Q>
- [2] Duarte, C. M., Chapuis, L., Collin, S. P., et al. (2021). The soundscape of the Anthropocene ocean. *Science*, 371(6529), eaba4658. <https://doi.org/10.1126/science.aba4658>
- [3] Sharma, P., Gero, S., Payne, R., Gruber, D. F., Rus, D., Torralba, A., & Andreas, J. (2024). Contextual and combinatorial structure in sperm whale vocalisations. *Nature Communications*, 15, 3617. <https://doi.org/10.1038/s41467-024-47221-8>

10.5 新的證據與接下來的故事

這個領域至今仍在迅速成長。每年都有新的研究發表，新的工具問世。這些都是幾年前難以想像的事情 [1]。

柏克萊與 Project CETI 的研究就是一個好例子。正如前面所見，研究團隊在抹香鯨的 coda 中發現了類似母音的聲音。2026 年，後續研究發表在《英國皇家學會報告 B》（*Proceedings of the Royal Society B*）上。這項研究指出，抹香鯨類似母音的聲音就像人類的母音一樣遵循著規則。這讓抹香鯨像人類一樣發出聲音的證據變得更加確鑿了。

Earth Species Project 開闢了另一條道路。這個組織打造了名為 NatureLM-audio 的人工智慧。這是第一個專為動物聲音設計的大型聲音語言模型。這個模型只要聽聲音，就能猜出是哪種動物。它還能數出錄音裡有幾隻動物。甚至連從沒學過的動物聲音，在某種程度上也能辨認出來 [2]。

2026 年，這個組織擴大了研究方式。他們不再只深入研究單一物種，而是把多種動物的聲音收集在一起。他們藉此拓展了用電腦同時分析多種動物聲音的方法。2026 年 7 月，他們公開了為此建立的共用資料庫 alp-data。這是讓全世界的研究人員共享相同資料的基礎。這就像在圖書館裡，任何人都能借閱同一本書一樣 [3]。

2026 年 3 月，Project CETI 發表了兩篇新研究。其中一篇是詳細記錄抹香鯨生產過程的研究。這是第一次如此細緻地記錄下海底生產的過程。另一篇則揭示了鯨群協助生產的模樣。當鯨魚媽媽生產時，其他鯨魚會在身旁守護著牠。這項研究詳細展示了非靈長類動物身上的這種合作照護行為。看到這種行為，有些學者開始探討鯨魚的權利。他們認為，複雜的社會性與照護行為，可以作為討論權利時的參考資料。不過，我們很難直接將互相幫助的行為視為法律上的義務。鯨魚幫助朋友，和人類遵守法律，是完全不同的兩回事 [4]。

法律要將動物像人一樣看待，還有很長的路要走。到目前為止，許多國家的法院都不太承認動物具備訴訟資格。2022 年美國紐約州法院的判決就是一個例子。動物權利團體曾為大象 Happy 提起訴訟。那是一場要求將牠從狹窄圍欄中釋放出來的訴訟。法院並沒有接受這項訴求。理由是不能把為人類設計的制度直接套用在大象身上。就像這樣，了解事實與改變法律是兩回事 [5]。

這所有的證據都指向同一個方向。那就是動物的聲音不是噪音，而是含有意義的溝通。隨著證據不斷累積，法律越來越難把動物當成單純的物品。科學走在前方，法律則在後頭緩步跟隨。人工智慧正逐漸成為把動物聲音帶進法庭的口譯員。為地球上的生命進行翻譯，與確立牠們的權利緊密相連。

相關佐證資料

- [1] Beguš, G., Sprous, C., Lebanidze, T., Gero, S., & Andreas, J. (2026). The phonology of sperm whale coda vowels. *Proceedings of the Royal Society B*, 293(2044). <https://doi.org/10.1098/rspb.2025.2994>

- [2] Robinson, D., Miron, M., Hagiwara, M., Weck, B., Keen, S., Alizadeh, M., Narula, G., Geist, M., & Pietquin, O. (2025). NatureLM-audio: An audio-language foundation model for bioacoustics. International Conference on Learning Representations (ICLR) 2025. <https://arxiv.org/abs/2411.07186>
- [3] Earth Species Project. (2026, July 22). Introducing alp-data: A shared data layer for animal language processing. Earth Species Project. <https://earthspecies.org/2026/07/22/introducing-alp-data-a-shared-data-layer-for-animal-language-processing/>
- [4] Project CETI. (2026, March). Studies documenting rare sperm whale birth and ancient cooperative care released. EurekAlert! <https://www.eurekalert.org/news-releases/1120823>
- [5] Nonhuman Rights Project, Inc. v. Breheny. (2022). 38 N.Y.3d 555, New York Court of Appeals. <https://www.nycourts.gov/ctapps/Decisions/2022/Jun22/52opn22-Decision.pdf>

尾聲：跨物種網際網路與共存的地球

E.1 抹香鯨的喀噠聲與深海字母

大海深處總是漆黑一片。陽光照不到的數千公尺深處，就像漆黑的夜晚一樣。在那片黑暗中，抹香鯨用聲音來看世界。抹香鯨會從額頭發射出喀噠喀噠的聲音。這種簡短的單一聲音被稱為喀噠聲(click) [3]。

抹香鯨會連續發出好幾個喀噠聲。這樣連在一起的聲音組合被稱為 coda [2, 3]。喀噠聲是單一的聲音，而 coda 則是把好幾個喀噠聲串在一起的聲音束。一個 coda 大約由 3 到 40 個喀噠聲連續組成。在人類耳朵聽來，coda 聽起來只是一模一樣的喀噠聲。長久以來，人類都把這種聲音當作沒有意義的噪音 [2]。

但是人工智慧重新傾聽了這種聲音。一個名為 Project CETI 的研究團隊展開了行動。CETI 是一個由多國學者組成、想要解開鯨魚語言的團隊。該團隊的研究員 Pratyusha Sharma 與 Jacob Andreas 教授分析了這些聲音。長期觀察多米尼克抹香鯨的 Shane Gero 博士也參與其中 [1]。

研究團隊收集並仔細研究了大量的聲音。他們逐一檢視了 8,719 個抹香鯨的 coda。並在其中找到了構成 coda 的四個基本要素。這些要素包括聲音的節奏、速度，以及額外添加的微小喀噠聲等。只要將這些要素靈活變換，就能組合出不同的 coda。

人工智慧在看似完全相同的喀噠聲中找到了規律。抹香鯨會稍微拉長或縮短喀噠聲之間的間隔。也會改變聲音的速度。有時還會在喀噠聲組合的前後加上微小的聲音。這些微小的差異聚集在一起，形成了不同的聲音結構。

抹香鯨還會微調聲音的速度。牠們有時會把同一個 coda 稍微拉長發出。有時則會稍微縮短節奏、快速發出。這就像唱歌時把拍子放慢或拉快一樣。這種微小的拉長與縮短，也是改變聲音的關鍵。

新的研究在此更進一步。學者們仔細觀察了 coda 中哪種音高的共鳴特別強烈。人類在發出「a」和「i」等母音時，也有強烈共鳴的位置。抹香鯨的 coda 中也出現了這樣的位置。CETI 團隊找到了這兩種鯨魚的母音。它們與人類發出的「a」和「i」聲音非常相似 [4]。

抹香鯨是怎麼發出這種母音的呢？抹香鯨會在頭部內側的聲唇製造聲音。再透過頭部的氣囊來修飾調整這些聲音。這就像人類在聲帶發聲，並透過改變嘴型來發出母音一樣。也就是說，製造聲音的地方和修飾聲音的地方是分開的。

研究團隊指出，抹香鯨能發出約 156 種不同的 coda。這些 coda 並不是單個字母。而是將節奏、速度等幾種要素編織在一起的聲音組合。原來在深海之中，也存在著用少數材料組合出豐富聲音的本領。不過，每個 coda 代表什麼意思目前尚未揭曉。研究團隊目前只是找到了聲音的結構。

即使是相同的 coda，在不同群體之間也有些微差異。這與人類的方言很相似。在同一個群體內，牠們會重疊發出聲音，互相呼應節奏。當一隻鯨魚發出聲音，另一隻鯨魚就會立刻接上聲音。抹香鯨就是這樣共同建立起對話的流動。

有些 coda 扮演著類似名牌的角色。抹香鯨過著大家族群居的生活。每個群體都有各自常用的 coda。只要聽到這種 coda，就能知道是屬於哪個群體的鯨魚。這就像人類透過方言來推測對方的家鄉一樣。

重要的是，正是這些微小差異的累積構成了聲音。抹香鯨會將幾種聲音要素靈活組合。根據組合方式的不同，就會產生各式各樣的 coda。這就像人類用幾個字母組合出無數個單字一樣。原來在深海裡，也隱藏著這樣精巧的結構。

要捕捉這些聲音，需要優良的錄音設備。CETI 團隊研發了一種能吸附在鯨魚身上的小型錄音器。這個裝置不僅能錄下鯨魚的聲音，還能記錄水中的各種環境狀況。因為唯有同時了解牠們在什麼情境下發出什麼聲音，才能解開其中的含義。人類在獲得人工智慧這項新工具之後，才終於看見了深海的字母。

相關佐證資料

- [1] Sharma, P., Gero, S., Payne, R., Gruber, D. F., Rus, D., Torralba, A., & Andreas, J. (2024). Contextual and combinatorial structure in sperm whale vocalisations. *Nature Communications*, 15, 3617.
<https://doi.org/10.1038/s41467-024-47221-8>
- [2] Oceanographic. (2026). Scientists listen in on sperm whale communication with new tech. *Oceanographic Magazine*.
<https://oceanographicmagazine.com/news/scientists-listen-in-on-sperm-whale-communication-with-new-tech/>
- [3] Project CETI. (2026). Research. Project CETI. <https://www.projectceti.org/research/index>
- [4] Kemeny, R. (2026, April 14). We're one step closer to understanding the sperm whale 'alphabet'. *National Geographic*.
<https://www.nationalgeographic.com/animals/article/sperm-whale-vowel-patterns-communication-discovery>

E.2 跨越只有人類會說話的想法

長久以來，人類總認為自己非常特別。相信在地球上只有人類能夠說話並交流心意。與人類語言不同的動物聲音，往往只被視為噪音。這種認為只有人類最特別的想法持續了很長時間 [1]。

這種信念如同一道堅固的高牆。正因為這道牆，動物的聲音長久以來被埋沒在黑暗之中。人類不曾試著傾聽身邊其他生命在想些什麼。在這段歲月裡，究竟有多少精彩的故事悄悄溜走，我們不得而知。

這道牆源自於人類總用自己的說話方式去衡量動物的習慣。人類認為必須要有詞彙和句子才算是語言。只要不符合這個框架的聲音，就被排斥為毫無意義。然而，動物的聲音完全沒有理由去迎合人類的語言框架。因為那是源自於各自不同的身體與不同生活方式的聲音。

人工智慧正在一步步拆除這道高牆 [3, 4]。人工智慧把動物的聲音整理在思維地圖上。學者們將這張地圖稱為潛在空間(latent space)。在這張地圖上，如果聲音相似，點就會聚集在一起。如果聲音不同，點就會相距較遠 [4]。

學者們將人類的語言地圖與動物的聲音地圖並排放在一起。這兩張地圖並非毫無關聯的陌生人。當對齊地圖的形狀並重疊在一起時，出現了相似的位置。人類的某些詞彙甚至與動物的某些訊號落在了同一個位置上。

這裡面隱藏著一種非常有趣的方法。將兩張地圖重疊時，不需要事先製作好的翻譯對照表。學者們只憑兩張地圖的外形特徵，就能找出相互吻合的位置。這就像把兩張紙重疊貼在窗戶上比對花紋一樣。學者們目前正在測試這種對齊形狀重疊的方法。

有一些學者正在引領這樣的研究。美國語言學家 Gašper Beguš 正利用人工智慧尋找動物聲音的規律。他們不會把人類語言的規則強加在動物身上。而是等待規律從聲音本身的形貌中自然浮現出來。

這裡萌生出了新的想法。動物並不是等待著人類語言的對象。動物是各自感受著自己的世界，並用聲音描繪那個世界的夥伴。動物用自己的感官所體會到的世界，被稱為 Umwelt。蝙蝠用聲音看世界，狗用嗅覺解讀世界。鯨魚有鯨魚的世界，鳥類也有鳥類的世界。

這種重疊對齊的方法在艱難的情況下發揮了力量。人類與動物之間沒有事先編好的辭典。也沒有把詞彙一一配對的對照表。儘管如此，人工智慧依然只看兩張地圖的形狀來嘗試尋找配對。這有望成為連接彼此不同世界的一座新橋樑。

因此，我們絕不能把動物的聲音置於人類語言之下。我們需要以平等的心態去尊重動物的思想與節奏。那是一種把所有生命之聲都看得同等珍貴的心。人類若想成為地球大家庭真正的一份子，必須先具備這樣的心態 [2]。

相關佐證資料

- [1] Rodríguez-Garavito, C. (2026, February 19). AI-enabled decoding of whale communication could bolster animal rights. *NYU School of Law Ideas*. <https://www.law.nyu.edu/news/ideas/cesar-rodriquez-garavito-animal-communication-ai>
- [2] MOTH - More Than Human Life. (2026). MOTH x CETI: Exploring how understanding sperm whale communications can be a force for good. *mothrights.org*. <https://mothrights.org/news/moth-x-ceti-exploring-how-understanding-sperm-whale-communications-can-be-a-force-for-good/>
- [3] Bioneers. (2026). The quest to decode whale-speak. *Bioneers*.
<https://bioneers.org/the-quest-to-decode-whale-speak/>

[4] Rutz, C., Bronstein, M., Raskin, A., Vernes, S. C., Zacarian, K., & Blasi, D. E. (2023). Using machine learning to decode animal communication. *Science*, 381(6654), 152-155. <https://doi.org/10.1126/science.adg7314>

E.3 連結地球生命的跨物種網際網路

人類之間有網際網路。網際網路是讓相隔遙遠的人們能夠互通訊息的網路。但是，地球上並不是只有人類在生活。鯨魚和大象、鳥兒和蜜蜂，甚至是樹木都生活在一起。有一個夢想是將這些生命全部連結在一起。這個夢想被稱為跨物種網際網路(Interspecies Internet) [1]。

這個夢想在 2013 年首次被提出。網際網路的締造者 Vint Cerf 與音樂家 Peter Gabriel 率先發起。學者 Neil Gershenfeld 與 Diana Reiss 也一同加入。他們主張不要讓網際網路成為人類獨有的工具。想法是把這個網路也向動物敞開。

跨物種網際網路不僅僅是用應用程式來翻譯動物的聲音。它意味著一個讓地球所有生命共享彼此消息與危險的龐大網路。就像人類的神經連結全身各個角落一樣，這個網路將連結整個地球的生命。

其實，這種連結的片段早已存在於大自然中。許多動物會偷聽其他物種發出的警告聲。當狐獴發出危險警報時，身旁的其他動物聽懂後也會一起逃跑。科學家已在 70 多種動物身上證實了這種偷聽行為。這等於是不同的動物在共享同一份危險情報 [2]。

海豚也展現了這種分享的能力。當海豚群體在海中相遇時，牠們會互相發出哨聲。各自不同的哨聲就像名牌一樣。牠們透過聲音來確認彼此是誰。這條看不見的聲音之路，早已構成了一張網路 [3]。

這個網路不只連結聲音。樹木與花草也會發出訊號。當缺水時，番茄和菸草等植物會發出人類耳朵聽不到的聲音。一株承受壓力的植物，一小時內會發出大約 30 到 50 次喀噠聲。人工智慧可以收集這些聲音，解讀出植物的健康狀況 [4]。

人工智慧正在協助將這些零散的片段拼湊在一起。未來即將推出的新通訊網路也能助一臂之力。稱為 6G 的通訊網路能讓電波在物體上反彈，進而感測周遭的動靜。將通訊網路當作感測器的想法被稱為網路即感測器(network as a sensor)。不過，將這項技術連結到動物溝通網路，目前仍只是一種可能性 [5]。

善用這個網路就能守護大自然。假設有人開始在熱帶雨林中濫伐樹木。鳥類四散的動靜與植物發出的異常訊號，就會透過網路匯集到中央。人類能立刻收到消息並予以制止。跨物種網際網路可以不是監視的工具，而是共同生活的工具。它將成為地球所有生命分擔彼此危險、互相照顧的巨大神經網路。

相關佐證資料

[1] Interspecies Internet. (n.d.). About. Interspecies Internet. <https://www.interspecies.io/about>

[2] Magrath, R. D., Haff, T. M., Fallow, P. M., & Radford, A. N. (2015). Eavesdropping on heterospecific alarm calls: from mechanisms to consequences. *Biological Reviews*, 90(2), 560-586. <https://doi.org/10.1111/brv.12122>

[3] Quick, N. J., & Janik, V. M. (2012). Bottlenose dolphins exchange signature whistles when meeting at sea. *Proceedings of the Royal Society B*, 279(1738), 2539-2545. <https://doi.org/10.1098/rspb.2011.2537>

[4] Khait, I., Lewin-Epstein, O., Sharon, R., et al. (2023). Sounds emitted by plants under stress are airborne and informative. *Cell*, 186(7), 1328-1336. <https://doi.org/10.1016/j.cell.2023.03.009>

[5] Nokia Bell Labs. (n.d.). Network as a sensor. Nokia. <https://www.nokia.com/bell-labs/research/6g-networks/6g-technologies/network-as-a-sensor/>

E.4 新工具揭示的事物：學習大自然語言的人工智慧

人工智慧也朝著學習動物聲音的方向邁進。Earth Species Project 推出了一款新的人工智慧。它的名字是 NatureLM-audio。這是第一個專門為動物聲音打造的大型聲音人工智慧。它於 2024 年 11 月首次公開，並在 2025 年的國際學術研討會(ICLR)上正式發表 [1]。

這個人工智慧可以同時執行多項任務。它能聆聽聲音並猜出是哪種動物。還能分辨該動物是幼崽還是成年個體。它會為聲音分類，並附上聲音的說明。甚至還能計算錄音中有幾隻動物。

令人矚目的是，它做到了人類沒有教過的事情。通常人工智慧只能做學習過的事情。但這個人工智慧在學習大量聲音的過程中，自己掌握了新本領。計算有幾隻動物這件事，並不是一開始就學會的。而是聽多了之後自然而然學會的。

Earth Species Project 同時研究多種動物的語言。目前他們正在深入研究烏鴉、白鯨與大象的聲音。他們與世界各地的生物學家攜手合作收集資料。不只專注於單一物種，而是將多個物種並列對比研究，這正是該團隊的特點。

2026 年 7 月，又有一項新工具問世。名稱為 alp-data。這是一個將多種動物聲音資料匯集在同一處的工具。裡面包含了鳥鳴聲、鯨魚聲和昆蟲叫聲等 35 種資料集。現在，學者們可以從單一窗口取用多種動物的聲音 [2]。

Earth Species Project 將這一趨勢稱為動物語言處理(Animal Language Processing)。這是一種利用人工智慧來研究動物溝通的方法。他們不會將人類的語言規則直接套用在動物身上。而是首先承認每種動物的溝通方式可能各不相同。在此基礎上，讓人工智慧利用資料來尋找規律 [3]。

該團隊也詢問了人們的想法。他們在 2025 年調查了全球 67 個國家、共 1,057 人。許多人對能夠理解動物語言表示歡迎。同時也表達了絕不能隨意對待動物的擔憂。隨著技術的發展，謹慎以對的心態也隨之增長 [4]。

這些新工具改變了學者們的工作方式。過去，一位學者往往只長期深入研究一種動物。資料也以各自不同的形式分散在各處。現在，大家可以在同一個窗口一起查看多種動物的資料。原本分散的研究，正匯聚成一股巨大的潮流。

下一個目標是聲音播放實驗。這是一項用人工智慧製作的聲音，試著向動物搭話的實驗。這項計畫主要由 Project CETI 提出。Earth Species Project 則比起隨意播放聲音，把更多心力放在收集和檢視資料上。人類製作的聲音動物是否聽得懂，目前還沒有得到證實 [5]。

相關佐證資料

- [1] Earth Species Project. (2024, November 11). Introducing NatureLM-audio: An audio-language foundation model for bioacoustics. Earth Species Project. <https://earthspecies.org/2024/11/11/introducing-naturelm-audio-an-audio-language-foundation-model-for-bioacoustics/>
- [2] Earth Species Project. (2026, July 22). Introducing alp-data: A shared data layer for animal language processing. Earth Species Project. <https://earthspecies.org/2026/07/22/introducing-alp-data-a-shared-data-layer-for-animal-language-processing/>
- [3] Earth Species Project. (2026, April 2). Animal language processing: An AI convergence in animal communication. Earth Species Project. <https://earthspecies.org/2026/04/02/animal-language-processing-an-ai-convergence-in-animal-communication/>
- [4] Earth Species Project. (2025, September 29). What the world thinks about AI and animal communication: Findings from our first global survey. Earth Species Project. <https://earthspecies.org/2025/09/29/what-the-world-thinks-about-ai-and-animal-communication-findings-from-our-first-global-survey/>
- [5] Project CETI. (2026). Research. Project CETI. <https://www.projectceti.org/research/index>

E.5 謹慎聆聽：負責任的工具與共存的承諾

新工具的力量很強大。力量強大的工具如果用錯了，就會帶來害處。人工智慧現在已經能夠做出跟動物一模一樣的聲音。如果把這種力量用在壞的地方，甚至可能欺騙動物。試想一下用假聲音引誘野生動物的情況吧。盜獵者可能會利用假聲音進行獵捕 [1][2]。

因此，我們必須遵守三項承諾。

第一，是不隨意干涉的承諾。動物群體在漫長的歲月中，累積了屬於自己的聲音文化。人類不應該為了做生意或控制牠們，就混入假聲音來打亂這套秩序。人工智慧不是奪走動物聲音的武器，而應該是順應牠們的意願、守護動物文化的工具。

為什麼需要這樣的承諾呢？長久以來，人類總是只想戰勝自然。了解聲音的力量，一不小心就可能變成操控對方的力量。如果隨意干涉動物群體，牠們的生活就會被打亂。因此，懂得越多，就越要懂得適時退後一步。

第二，是在聲音中留下標記的承諾。學者們建議，在人工智慧製作的動物聲音裡植入看不見的標記。這個標記稱為音訊浮水印（audio watermark）。就像偽鈔上的防偽圖案一樣，人的耳朵聽不見，但機器能辨認出來。這樣一來，就能分辨哪些聲音是真的，哪些聲音是假的 [3]。

這種標記也能用來保護資料。學者們會把錄音機長時間掛在海洋和森林中。像這樣人類不介入、只收集聲音的工作，稱為被動聲學觀測。如果混入了假聲音，這些資料就會受到污染。只要在聲音中植入標記，就能過濾掉假聲音，保持資料的乾淨純粹。

第三，是共同制定的承諾。紐約大學法學院的 MOTH 計畫帶頭進行這項工作。MOTH 是一個探討非人類生命權利的組織。機器找到的資料，必須與長久守護海洋的原住民的古老智慧相結合。原住民經過長期觀察獲得的自然知識，稱為傳統生態知識。不能只憑人類的利益來制定規則。

這三項承諾源自同一種心態。那就是即使獲得了新的力量，也不隨意擺布對方的態度。長久以來，人類一直只把自然當成工具。現在是該反省這種習慣的時候了。人工智慧這項新力量並不是拿來誇耀的事物，而是一份沉重的責任。

如果懂得抹香鯨的語言，就能進一步思考新的權利。學者們主張鯨魚擁有不受騷擾的權利，同時也主張牠們擁有守護自身文化的權利。2024年，紐西蘭與庫克群島的原住民領袖宣告，應將鯨魚視為擁有與人類相同權利的存在。這意味著鯨魚不是任由人類隨意使用的物品。

這本書記錄了翻譯地球生命的故事。抹香鯨的喀嗒聲、鳥兒的歌唱、大象低沉的共鳴，人工智慧正在一步步解開這些謎團。翻譯不是為了戰勝對方，而是為了去理解對方。人工智慧是翻譯地球生命的橋樑。當我們小心翼翼地走過這座橋時，就能打造一個與鄰近生命共同生活的地球。翻譯並沒有在此結束，而是才剛剛開始。

相關佐證資料

- [1] Inside Climate News. (2025, October 29). AI is decoding whales' communications. Could that be a turning point in the push for their rights? Inside Climate News. <https://insideclimatenews.org/news/29102025/ai-sperm-whale-communications-legal-rights/>
- [2] MOTH - More Than Human Life. (2026). MOTH x CETI: Exploring how understanding sperm whale communications can be a force for good. mothrights.org. <https://mothrights.org/news/moth-x-ceti-exploring-how-understanding-sperm-whale-communications-can-be-a-force-for-good/>
- [3] Rutz, C., Bronstein, M., Raskin, A., Vernes, S. C., Zacarian, K., & Blasi, D. E. (2023). Using machine learning to decode animal communication. Science, 381(6654), 152-155. <https://doi.org/10.1126/science.adg7314>

支持這本 AI 書齋著作

如果這本書對您有幫助，歡迎透過 PayPal 支持後續的翻譯、公開版本以及開放的人工智慧知識計畫。

PayPal

<https://paypal.me/kimkjkj>



請掃描 QR 碼，或直接開啟上方連結。

金京鎮律師