

【與三星研究院共築未來②】波蘭三星研發中心：打造 AI 驅動技術，開創便利新世界

於本系列報導中，三星新聞中心專訪來自三星全球研發中心的技術專家，深入了解其工作內容及創新如何造福消費者。



本系列報導的第二位受訪專家為來自波蘭三星研發中心（SRPOL）的 AI 團隊負責人 Lukasz Slabinski。Lukasz 於 2013 年加入 SRPOL 團隊，最初擔任高級工程師，歷經 8 年辛勤耕耘，目前晉升 SRPOLAI 的 AI 團隊掌舵人。請繼續閱讀以下報導，深入了解 SRPOL 及 AI 團隊的重大研究突破。



系列報導《Q&A》

問：相較於智慧視覺技術開發，據悉語音辨識設計解決方案的複雜度更高。您曾遇過哪些挑戰？又是如何克服？

- *Sergei Lytvynenko* (烏克蘭三星研發中心)



答：人類語言如口音、語調和方言均不斷演進，不僅極其主觀且變化多端。為克服重重挑戰，我們利用先進機器學習技術建構模型，而該模型能記憶由數據分析自動辨識的通用語言模式。

- *Lukasz Slabinski* (波蘭三星研發中心)

問：據悉語音辨識解決方案的設計極其複雜。研究語言相關技術時，曾遇過哪些挑戰？又是如何克服？

以我個人的觀點，語言相關技術遠比其他技術複雜。人類的溝通語言高達 7,000 多種且不斷演進，並可再細分為各種口音與方言。此外，人類語言的客觀性，遠不及能以數學公式描述的圖片。人們將其想法，以聲音或符號型態編碼成訊息，待由他人解碼並詮釋。由於此過程每一階段皆具個人性、創造性、非決定性，因此，基於語言的人類交流具高度複雜性，且十分模稜兩可，意即唯美詩篇、有趣笑話與溝通誤解只有一線之隔。

從事自然語言處理 (NLP) 的研發人員經常面臨人類與生俱來的侷限性。即便是工作同事，亦或是同住家人，彼此皆會產生溝通問題。而只精通 2 種語言的工程師，該如何設計、編碼一套涵蓋 40 種不同語言的機器翻譯系統？我們借助機器學習技術，解決此項難題。

所謂「訓練」過程，是根據數據集中的範例，自動提取通用模式，並以模型形式加以記憶。為建構機器翻譯系統，團隊訓練一套神經網路，可根據數百萬個縝密蒐集整理的範例，映射不同語言的句子。聽起來容易，但此階段具備三項基礎挑戰。

第一項挑戰為設計適合機器學習的模型架構，該架構需能記憶和歸納足夠的語言模式，用以解決特定問題，例如機器翻譯、情緒分析、文本摘要等。

第二個挑戰則是準備充足的訓練數據；因機器學習系統僅能辨識與記憶訓練數據集中的模式。

最後一項挑戰，是在專屬雲端或終端裝置平台上，部署已完成訓練的機器學習模型。

我們運用工程師豐富的专业知識及縝密的方法蒐集數據，並以最先進的機器學習架構進行無數次試驗，逐

—化解上述挑戰。

問：請簡單介紹波蘭三星研發中心（SRPOL）、AI 團隊，以及主要研發工作？

SRPOL 為波蘭規模最大的國際軟體研發中心，分別在波蘭首都華沙，以及科技重鎮克拉科夫設有據點，並與當地新創企業、大學和研究機構密切合作。

SRPOL AI 團隊的使命，即打造基於 AI 的功能、工具和服務，使人們生活更加便利與豐富。我們聚焦 NLP 和智慧音訊領域，同時具備包括推薦系統、室內定位、視覺分析和 AR 等跨領域專業知識。

問：自 2018 年起，您擔任波蘭研究中心 AI 團隊負責人，期間推動多項 NLP 等領域的專案。您與團隊目前正投入何種研究？

團隊持續進行 10 多年前啟動的 NLP 領域研究，包含機器翻譯、對話系統（涵蓋問題回覆和文本分析）等開發。同時兵分二路，分別研究可擴充規模、強大的雲端服務，以及可快速離線作業的裝置終端應用。

對我們而言，智慧音訊是較新的領域。幾年前，隨著該領域的重要性與日俱增，我們開始將其列為研究重點。目前已致力於聲音辨識、分離、強化和分析。研究過程會將音訊處理的所有層面納入考量，從聲學場景理解，到搭配資源有限的硬體設備（如無線耳機），針對嵌入式音訊演算法進行微調。



問：您聚焦的技術包括 **NLP**、文本和數據挖掘、智慧語音等。哪些研究直接影響三星產品或服務的開發？團隊如何提升用戶的生活便利性？

SRPOL 在 AI 技術商業化領域已具悠久歷史，但並非單打獨鬥。我們很自豪能成為大環境中重要一員，SRPOL 與其它三星研發中心密切合作，致力實現商業化目標。

例如，團隊為三星行動裝置開發多項智慧文字輸入功能，包括螢幕鍵盤、主題標籤、Samsung Note 標題推薦，以及智慧手錶的智慧文字回覆。

此外，我們亦對「三星應用商店」推薦系統貢獻良多，該系統可根據用戶喜好推薦最有趣的遊戲。

問：身為智慧音訊等新興 AI 領域的倡導者，您認為目前產業趨勢為何？此技術將如何影響人們日常生活？

我相信智慧音訊技術未來將改變消費性電子產品領域的遊戲規則。深耕音訊分析極其重要，因其是實現以人為導向的 AI 系統中，不可或缺的重要拼圖。

強大的 NLP 系統能透過文字和語音分析更了解用戶。相機和視覺內容輸入的背後，電腦視覺演算法必不可缺。對普羅大眾而言，很難想像在無導航情況下駕駛車輛、輸入訊息時少了拼字修正器，亦或無法使用網際網路搜尋資訊。但截至目前，除去部分專業應用，我們鮮少使用智慧音訊技術強化聽力。我認為此情況即將發生改變。

試著想像 - 假設現在存在一種普遍的技術，可讓人們選擇耳朵想聽到的內容以及聆聽方式。如在繁忙都市的公園中與朋友共進午餐時，可選擇只聽大自然和交談對象的聲音。或能直接於人類腦海中，建構身臨其境的 3D 音訊體驗、被稱為元宇宙 (Metaverse) 的一種先進 VR 或 AR 系統。光是上述兩項概念，即可組合數百種全新、可行的使用案例。而將眼光放遠，若能聽見人類目前聽不到的聲音，會是何種情景？目前人耳聽得到的音頻範圍相當窄。世界充滿了深具意義的聲音，但 AI 技術觸及的部份卻只有冰山一角。隨著智慧音訊技術持續發展，我相信這將影響人類生活。



▲波蘭三星研發中心研究人員，在無回音室利用人頭和軀幹模擬器（HATS），進行主動式降噪（ANC）技術開發工作。

問：您如何將當前趨勢融入波蘭三星研發中心的研究中？

除 NLP 和音訊之外，我們亦積極尋找並建構多模式系統最有效的途徑。為此，團隊從不同觀點切入研究，並分析使用案例。歸功由工程師、語言學家、數據科學家所組成的多元跨學科團隊，我們得以進行此類型分析。

問：截至目前為止，您在 SRPOL 最重要的成就為何？

為機器翻譯解決方案。此項解決方案已連續五年於各大競賽中拔得頭籌：2017 年至 2020 年國際口語機器翻譯評測比賽（IWSLT）；2020 年國際機器翻譯大賽（WMT）；以及 2021 年國際機器翻譯大賽（WAT）。上述皆為該領域最負盛名的國際競賽。

今年能在 WAT 獲得獎項肯定，是特別值得驕傲的一項里程碑，開發亞洲語言解決方案對身為波蘭工程師的我們而言，堪稱一項壯舉。最終成果證明三星技術的實力，遠遠超越展示成果。

另一項令我引以為傲的成就，是智慧音訊團隊技術開發的進展速度。我們從零開始，短短幾年即成績斐然，於 2019 和 2020 年連續二年站上 DCASE Challenge（聲音場景與事件偵測分類挑戰賽）頒獎台。我們亦發表數篇有關此領域的科學論文，並取得多項專利。我相信，三星將持續在此領域發光發熱。

【與三星研究院共築未來】 AI 專家訪問機器學習專家



機器學習 (ML) 為 AI 解決方案中的一項基礎技術，因此「AI」與「ML」經常被交互使用。開發新 ML 演算法時，您應該面臨不少挑戰。可否談談您的研究機構面臨哪些新挑戰？為了克服，您採取了哪些努力？

下一篇文章，將訪問來自中國北京三星研發中心的機器學習專家 - Bin Dai。