

2025 年 11 月 10 至 13 日

韓國釜山

2025 國際檔案理事會東亞地區分會研討會

主題：歷史檔案管理的重大變革：

從莎草紙走向人工智能技術

中華人民共和國香港特別行政區政府

政府檔案處報告

為邁向人工智能奠定基礎：香港政府檔案處迎接檔案管理

轉型新時代的準備

（講者：政府檔案處檔案主任許崇德先生）

I. 引言

人工智能(AI)是當今時代最具變革性的力量之一。人工智能讓機器具有學習、推理和解決問題這些人類獨有的智慧能力。近年來，人工智能技術的急速發展與廣泛應用，不僅重塑了機構的運作模式，亦為檔案專業帶來機遇和挑戰。

2. 中華人民共和國香港特別行政區政府(「政府」)高度重視人工智能的巨大潛力，並將推動人工智能發展列為施政重點之一。2022年，創新科技及工業局公布《香港創新科技發展藍圖》，為香港人工智能發展勾勒出策略性方向和具體行動計劃。2023年，政府成立香港生成式人工智能研發中心(Hong Kong Generative AI Research and Development Centre [HKGAI])，進一步推進這方面的發展，專注研發開源基礎模型，包括香港自主研發的大語言模型及生成式人工智能文書輔助應用程式。自2024年中起，人工智能助理「港文通」(HKPilot)為政府人員在草擬文件、翻譯及撮要文件等方面提供支援。2024年7月，政府將當時的政府資訊科技總監辦公室與效率促進辦公室合併成為數字政策辦公室(數字辦)，體現對數碼發展的重視。2025年，政府再邁出重要一步，宣布撥款10億港元設立香港人工智能研發院，以推動人工智能的研發與產業應用。最近，《行政長官2025年施政報告》宣布成立「AI效能提升組」，負責統籌和指導政府部門於運作中有效應用人工智能技術。這些舉措均展現了政府推動在公共行政全面數碼轉型的決心。與此同時，新成立的數字辦對於協調人工智能在數據分析、客戶服務和文件處理方面的應

用，亦發揮重要作用，有助提升運作效率和服務質素。

3. 在此宏觀背景下，政府檔案處（檔案處）積極作準備，以把握人工智能技術帶來的機遇，從而優化歷史檔案的管理工作。為迎接新一波的轉型，我們構建數碼基礎設施，並開展試驗項目，以探索如何利用人工智能提升我們的工作效能。這些努力為歷史檔案管理的新篇章奠定堅實基礎。人工智能具備潛力，可協助我們更有效地保存歷史檔案，讓日後的使用者更易於取閱，從而為檔案管理開啟更美好、更創新的未來。

II. 打造數碼基礎，確保人工智能準備就緒

4. 紙本歷史檔案雖具歷史價值，但在取閱、搜尋和擴展性方面有其局限。例如受制於貯存載體，非結構化的格式亦耗費大量人手處理。故此，紙本檔案無法被機器讀取，也不能滿足人工智能對數據密集處理的需求，與人工智能驅動程序的兼容性較低。所以，從紙本環境過渡到數碼化檔案保管環境，是應用人工智能的先決條件。為此，檔案處優先發展核心數碼基礎設施，包括以電子檔案保管系統和數碼儲存庫，

作為人工智能應用的骨幹。同時，亦進行其他與紙本轉數碼化相關的準備及強化措施，例如制定政策指引、優化元數據管理，以及大規模數碼化實體檔案。

檔案處數碼基建設施(1)：電子檔案保管系統

5. 電子檔案保管系統是這個過渡期的重心。政府的目標是在 2025 年底前把這個系統推展至所有政府決策局／部門（局／部門），以優化原生數碼檔案的保存和管理政府檔案的效率。截至 2025 年 9 月底，所有局／部門均已推行電子檔案保管系統。預期在不久的將來，部分電子檔案將經由此系統平台移交至檔案處作進一步處理。這標誌著我們向構建數碼工作環境邁出重要一步，不僅促進政府檔案管理、移交和保存的效率，同時為未來應用人工智能締造有利條件。

6. 此外，各局／部門正收集電子檔案保管系統使用者的優化建議，以提升系統的用戶體驗。我們尤其歡迎各局／部門的電子檔案保管系統使用者就人工智能在電子檔案保管系統中的應用提出建議，這些建議將交給由數字辦和檔案處代表組成的電子資料管理計劃管理辦公室進一步審視。

檔案處數碼基建設施(2):數碼檔案庫

7. 數碼檔案本質上支援人工智能驅動工具，但亦有其脆弱性，尤其是在依賴特定的硬件和軟件開啟方面，這些硬件和軟件可能會過時，加上容易損毀、遺失及出現載體老化等問題。為此，檔案處在 2020 年利用商用套裝軟件建立數碼檔案庫，以確保數碼檔案未來可被人類和機器讀取和查閱。檔案處亦在探索將人工智能應用於數碼檔案以釋放它們的潛在價值。

8. 數碼檔案與先進技術和自動化流程兼容，有利人工智能應用。透過數碼檔案庫管理的數碼檔案，經政府雲端設施服務(Government Cloud Infrastructure Services [GCIS]) 等雲端運算平台支援，其可擴展的儲存空間和運算能力，成為了訓練及配置人工智能模型的必要條件。這種可擴展性，結合高速、低時延網路，能支援實時人工智能應用，實現高效的數據處理和分析。

9. 此外，數碼檔案庫提供管理完善的儲存空間，使非結構化數據（例如信件、會議記錄或照片）轉換成標準格式（例

如文字和影像檔案)的過程更方便。這種結構化促進數據分析、建模和自動化，因此對人工智能發展非常重要。數碼工作流程也透過把搜尋及分類等日常程序自動化，以及減少人手操作常見的錯誤和延誤，以推動人工智能的應用。數碼檔案庫的內置應用程式介面及以元數據驅動的工作流程，進一步簡化數據檢索及智能作業，為人工智能應用提供強而有力的後盾。因此，隨着檔案由紙本轉為數碼，並在數碼檔案庫創新工具的支援下，檔案處使用人工智能技術日益可行。

為長遠保存電子檔案制訂政策和指引

10. 為長遠保存電子檔案制訂政策和指引，是實現檔案由紙本轉數碼的重要一環，亦為應用人工智能作好準備，此舉尤其符合近年政府服務日益應用人工智能及政府推動公共服務數碼化的方針。這個由紙本走向數碼的轉變，使妥善管理和保存電子檔案成為日益迫切的要務，當中具長遠價值的檔案更需優先處理。對此，檔案處在 2023 年成立的長遠保存電子檔案專責小組已着手制訂兩個政策框架：一個用於規範檔案處內具歷史價值的電子檔案，另一個則涵蓋政府各局／部門內具長遠價值的電子檔案。這些政策旨在闡明保存檔案

處具歷史價值的電子檔案和各局／部門具長遠價值的電子檔案的主要原則和方法，包括已被鑑定為具長遠價值但尚未移交至檔案處的電子檔案。這些政策為檔案整個生命週期內（即由各局／部門開立至移交至檔案處作永久保存及在檔案處保管下作其他用途）的行動標準化提供依據，目的是透過遵行相關法律規定、國際標準和最佳做法，盡可能保護檔案處具歷史價值的電子檔案和各局／部門的電子檔案免受風險，並按需要一直確保這些檔案真確、完整、可靠、可供查閱和使用。當上述兩項政策獲通過後，便會制訂更詳細的指引和程序，以進一步支援檔案處和各局／部門長遠保存各自的電子檔案。

優化元數據管理以支援存檔、數碼保存及人工智能驅動的搜尋

11. 建立元數據管理是另一個重要基礎。如前所述，檔案處已成功建立數碼檔案庫。前文提及的一些現有功能，例如可擴展儲存空間和運算能力、高效可靠的政府雲端設施服務網絡、將非結構化檔案轉換成標準格式，以及內置應用程式介面和元數據驅動工作流程等，已證明檔案處在應用人工智能

優化存檔和數碼保存工作是可行的。

12. 由於檔案處的數碼檔案庫配備了內置功能和套裝應用程式介面，支援元數據的擷取與操作，預期透過自動收集數碼檔案中的元數據，可提升元數據處理能力，從而通過自動標記及檔案分類，提高人工智能輔助搜尋的可行性，讓使用者獲得更精準及符合語境的檢索結果。

為實體檔案進行大規模數碼化

13. 把實體檔案大規模數碼化，不但方便使用者透過電子平台更廣泛及更便捷地查閱檔案，更可加快檔案由紙本轉數碼的步伐，對應用人工智能至關重要。有鑑於此，檔案處在 2018 年開展為期十年的大規模數碼化計劃，旨在將大量以模擬格式儲存的檔案轉為數碼資料，共涉及 600 萬個影像。今年的製作目標約為 40 萬個數碼影像。截至 2025 年 9 月，已累計完成 480 萬個數碼影像，逐步邁向計劃最終目標。

III. 應用人工智能技術的試驗項目

14. 為進一步探索人工智能的應用潛力，檔案處亦與香港

一所大學及一家資訊科技公司合作開展試驗項目，藉此獲得應用人工智能技術的第一手經驗。

與大學協作

15. 檔案處明白，數碼人文和人工智能項目的複雜性講求跨學科技能、專門知識和豐富資源，所以協作對成功推行相關項目至關重要。為評估自身準備情況，檔案處自 2024 年起與香港一所大學合作推行數碼人文／人工智能計劃，涵蓋多種工作和技術，包括：在數碼人文學科舉辦使用歷史檔案的預備工作坊；將選定的二次大戰期間的土地記錄數碼化；以及在光學字符識別、從掃描影像提取文字及從歷史檔案生成結構化數據集等領域應用人工智能。檔案處透過這些工作，以建立更完善的檢索點，設置穩妥的索引和輔助檢索工具，最終完善檔案工作流程，使歷史檔案更便於檢索。

與資訊科技公司合作

16. 檔案處現正與一家資訊科技公司合作，利用人工智能技術生成新的搜尋標籤，以提升搜尋檔案處館藏照片的準確度，從而突破現時透過網上搜尋目錄檢索館藏（尤其是照片）

的局限。現時，公眾主要靠預設關鍵字搜尋藏品。以照片而言，標題無法充分描述它的內容。所以，以標題內的關鍵字作搜尋會有所局限。這個合作項目預期可以發展出一套創新的機器學習解決方案，利用影像分析技術生成新的搜尋標籤，自動識別數碼照片中的地點、地標、歷史人物及其他重要的歷史資訊。預期可以顯著提升公眾探索香港視覺歷史的能力，並開創與社區遺產互動的新方式。

IV. 應對人工智能時代的挑戰

17. 應用人工智能並非全無挑戰。在探索人工智能發展的同時，檔案處致力以各種方式妥善應對挑戰，包括加強數據安全保護、進一步規劃相關政策和指引、遵守檔案管理規定、作出周詳考慮，以及為員工提供更多培訓和專業發展機會。

致力確保數據安全

18. 數碼檔案庫是檔案處邁向以人工智能驅動檔案及歷史檔案管理的關鍵資產。雖然人工智能在提升資源管理效率、確保資源可靠和持續更新方面，具有顯著的效益，但同時亦

帶來挑戰，特別是為顧及數據安全，政府必須遵守嚴格的數據安全規程和災難管理方針。檔案處在繼續整合新興技術以探索人工智能的潛力之餘，亦會全力確保檔案和歷史檔案在數碼時代是高度安全、可靠及易於取閱的。

政策及指引規劃

19. 儘管人工智能潛力巨大，但在應用時須審慎。人工智能在歷史檔案管理工作的應用，例如檔案登錄、著錄、覆核查閱申請和數據探索等，仍處於試驗階段，將試驗結果轉化成綜合歷史檔案管理系統、做法和工作流程，需要投入大量研發資源，並需保持檔案可靠、可信和具問責性等。待取得更明確證據證明人工智能能勝任檔案工作後，檔案處將會適時制訂新政策及指引，並更新現行版本，向持份者闡明使用人工智能的原則，以及就負責任地使用人工智能提供清晰的指示。

遵守檔案管理規定

20. 遵守檔案管理規定是不容妥協的。檔案處將仔細評估

各局／部門電子檔案保管系統使用者就人工智能應用所提出的建議，以確保這些建議符合規定、可行並對使用者有益。數字辦、檔案處與各局／部門會合作至關重要，以確保人工智能工具發揮最大效用，優化而非擾亂健全的檔案管理業務。

在檔案保存及修復工作上應用人工智能技術

21. 作為檔案機構，檔案處亦負責保存及修復歷史檔案和圖書館資料，以延長其保存期。雖然人工智能具備應用在環境監察與文物修復等工作的潛力，但挑戰仍然存在。檔案處會通盤考量，進一步探討此議題，期望制訂出可兼顧各方面考慮的新方案。

培訓及專業發展

22. 員工的專業能力是可持續轉型的基石。人工智能技術日新月異，檔案管理人員要有效地運用人工智能工具，他們的技能和知識亦須與時並進。檔案處需迎接及適應環境不斷變化所帶來的挑戰，並積極為員工提供專業培訓機會。為協

助團隊緊貼科技的最新發展，檔案處為檔案主任職系人員制訂了全面的培訓及專業發展策略，重點涵蓋三個主要工作範疇：(1)歷史檔案及檔案管理的核心知識；(2)數碼技能；以及(3)跨領域知識和合作。我們的員工已參加由政府公務員學院及／或專業機構舉辦的培訓課程，內容涵蓋大數據、區塊鏈技術、數碼通訊及操守與合規性等主題。至於館長職系人員，檔案處亦鼓勵他們增進與保存及修復歷史檔案工作有關的人工智能技術知識。檔案處會繼續投放充足資源，鼓勵員工參與培訓及專業發展活動，使他們具備能力及作好準備，應對數碼及人工智能時代帶來的挑戰。

人工智能技術的局限

23. 儘管人工智能可帶來很多潛在的效益和優點，我們亦不能忽視其固有的局限。人工智能模型是基於數據訓練，若數據存在偏頗或謬誤，人工智能將延續甚至放大這些問題。檔案處目前正探索在特定領域融合人工智能的可能性，但無意把人工智能應用於所有運作。此外，必須注意的是，人工智能應作為專業判斷與檔案專業知識的補充，而非替代。人員的監督和管理仍然不可或缺。換言之，人員的專業知識應

與人工智能的效率互相配合，相輔相成。

V. 總結

24. 檔案處正為人工智能賦與未來歷史檔案管理工作奠定基礎。透過構建穩健的數碼基建設施、制訂全面的政策指引，以及推行創新的試驗項目，檔案處正逐步建立以負責任和可持續方式運用人工智能的能力。作為檔案管理人員，我們的使命始終不變：妥善守護受託管的文獻遺產，並確保後代能夠取閱。變化在於我們所能使用的工具與機遇。藉著人工智能，我們有能力改變的不只是保存過去的方式，還有如何將過去與未來相連。