

1999 臺北市民當家熱線數位轉型與 個資保護對策之研究

受委託單位：達文西個資暨高科技法律事務所

計畫主持人：葉奇鑫律師

協同主持人：王慕民律師



臺北市政府研究發展考核委員會委託

中華民國 111 年 12 月 19 日

非 經 本 會 同 意
不 得 轉 載

摘要

隨著全球邁入數位化時代，落實數位轉型、構建數位社會已成全球趨勢。政府服務之數位化、智慧化是其中重要面向。臺北市政府（北市府）配合服務型智慧政府推動計畫，致力推動數位轉型，現於 1999 臺北市民當家熱線（1999 熱線）擬導入人工智慧優化市民服務。其於提升服務品質、優化行政資源分配、促進創新之同時，亦將涉及民眾隱私及個人資料保護、新興科技適當利用等議題，即有事前評估法規風險、權利保障或倫理疑義之必要。

據此，本研究即（1）蒐集國內外政府或民間團體數位轉型及導入人工智慧後，在政策、法規及制度修正之案例，梳理爭點與解決方案；（2）對 1999 熱線執行個人資料影響評估（DPIA），針對性檢視北市府現行對於 1999 熱線規劃導入之人工智慧應用方式可能面臨的法律風險；以及（3）於法規、案例分析及 DPIA 之基礎上，從法規面、制度面及流程面提供 1999 熱線導入人工智慧技術後，服務流程再設計之參考，並提出「臺北市政府開發及應用人工智慧指引」草案，供北市府評估作為日後導入人工智慧之作業依據。

關鍵字：數位轉型、人工智慧、政府服務、個人資料保護、制度再造

Abstract

As the world enters the digital era, the pursuit of digital transformation and building of digital society has become a global trend, to which the digitalization and intelligentization of government service is a key aspect. As a part of the action plan for building a service-oriented smart government, the Taipei City Government has been carrying forward its digital transformation, and is now planning to introduce artificial intelligence (AI) applications into the 1999 Taipei Citizen Hotline (the 1999 Hotline). While expected to enhance service quality, optimize administrative resources allocation and promote innovation, the planned upgrading could also give rise to concerns on citizens' privacy/data protection, and the appropriate use of new technologies. Therefore, it is necessary to evaluate legal risks and clear any issues regarding adverse impacts on rights and ethics.

Therefore, this research: (1) compares cases and examples from home and abroad and in both public and private sectors on the adjustment of policies, regulations and institutions, and analyzes any focal points and their solutions; (2) carries out a data protection impact assessment (DPIA) on the 1999 Hotline, and examines any potential legal risks of the planned introduction of AI applications by the Taipei City Government; and (3) on basis of the aforementioned analysis and DPIA, proposes regulatory, institutional and procedural recommendations for the reformation of the 1999 Hotline service process after the upgrading. This research also products a draft Guidelines of the Taipei City Government on the Development and Use of Artificial Intelligence, for future reference on further introduction of AI applications.

Keywords : Digital transformation, Artificial intelligence, Government services, Personal data protection, Systematic reform

目錄

第一章 研究主旨	1
第一節 研究主題	1
第二節 研究緣起與動機	1
第三節 研究目的	3
第四節 研究範圍與限制	4
第二章 研究方法	7
第一節 法規及文獻分析	7
第二節 案例彙整	7
第三節 個資保護影響評估	9
第四節 指引草案與焦點座談	11
第三章 蒐集之法規、文獻分析	13
第一節 歐盟：可信賴人工智慧倫理指引.....	13
一、 可信賴 AI 系統三特徵	13
二、 可信賴 AI 系統四倫理原則	13
三、 可信賴 AI 系統七遵守面向	14
第二節 歐盟：AI 法草案	17
一、 歐盟 AI 法草案提出背景	17
二、 AI 法草案之監理制度設計	19

三、	AI 法草案與歐盟 GDPR 之適用關係	23
四、	歐盟理事會對 AI 法草案之審修.....	24
第三節	英國 AI 與資料保護指引	31
一、	AI 系統之課責性	32
二、	AI 系統個資處理之合法性、公平性與透明性	34
三、	AI 系統之資料安全與資料最小化義務	37
四、	AI 系統與資料主體（當事人）權利	39
五、	AI 與資料保護工具組	40
第四節	美國人工智慧權利法案藍圖.....	52
一、	安全有效之系統	53
二、	演算法歧視之保護措施	55
三、	資料隱私	57
四、	告知與解釋	60
五、	人為替代、評估及備案（fallback）	61
第五節	日本：以人為本的人工智慧社會原則.....	65
一、	以人為本原則	65
二、	認知教育原則	65
三、	隱私保護原則	66
四、	確保安全原則	67

五、	公平競爭原則	67
六、	公平、課責與透明原則	67
七、	創新原則	68
第六節	新加坡：人工智慧治理模範框架.....	69
一、	內部治理架構與措施	69
二、	決定 AI 輔助決策中的人為參與程度.....	71
三、	作業管理	73
四、	利害關係人互動與溝通	75
第七節	香港－開發及使用人工智能道德標準指引.....	77
一、	人工智慧策略及治理	78
二、	風險評估及人為監督	80
三、	人工智慧模型的開發及人工智慧系統的管理.....	82
四、	與利害關係人的溝通及交流.....	83
五、	自我核對清單	83
第八節	法規及指引文獻比較分析	88
第四章	蒐集之案例彙整	93
第一節	官方專線服務導入 AI 案例	93
一、	比利時（佛拉蒙地區）Flemish Infoline 1700..	93
二、	愛爾蘭（稅務局）VDA（Virtual Digital Agent）	

三、	西班牙（中央政府）060 專線	95
四、	我國（勞動部）多元智能電服中心.....	95
五、	我國中國信託商業銀行 AI 語音客服.....	100
六、	我國長庚醫院 AI 智能客服	100
第二節	語音辨識 AI 案例	101
一、	愛沙尼亞（國家議會）Hans.....	101
二、	德國（聯邦訊息技術中心）C-19.....	101
三、	德國（聯邦移民及難民辦公室）難民語言偵測..	101
四、	我國法務部語音辨識系統	102
第三節	AI 線上聊天機器人案例	103
一、	英國（英國駕照及行照監理局）聊天機器人....	103
二、	奧地利（聯邦數位及經濟事務部）Mona.....	104
三、	丹麥（腓特烈堡自治市）Kiri.....	104
四、	愛沙尼亞（統計局）Iti	105
五、	芬蘭（奧盧市）OuloBot	105
六、	羅馬尼亞（克盧日－納波卡市）Antonia.....	105
七、	新加坡（微軟）對話式 AI 技術解決方案.....	106
八、	我國高雄市政府 1999 智能客服萬事通.....	108

第四節	案例比較分析	109
第五章	臺北市民當家熱線現況分析	111
第一節	1999 臺北市民當家熱線簡介	111
第二節	個資保護影響評估 (DPIA)	112
一、	評估程序	112
二、	背景說明與資料流	113
三、	法規遵循風險識別	125
四、	資安技術風險識別	161
五、	剩餘風險及因應建議	164
第三節	臺北市民當家熱線導入人工智慧應用評估小結..	169
第六章	指引草案與焦點座談	173
第一節	臺北市政府開發及應用人工智慧作業指引草案之提出	173
第二節	焦點座談	173
一、	焦點座談議程	174
二、	焦點座談討論議題大綱	175
三、	焦點座談家意見彙整紀錄	177
第三節	臺北市政府開發及應用人工智慧作業指引草案全文	181

第七章 結論與建議	207
第一節 立即可行之建議	207
一、 法規面	207
二、 流程面	208
三、 制度面	210
第二節 長期性建議	211
一、 法規面	211
二、 流程面	212
三、 制度面	213
第八章 參考文獻	215
附錄一：1999 熱線檢核實作結果	221
附錄二：啟動會議記錄.....	227
附錄三：期中報告審查會議紀錄.....	231
附錄四：期中審查意見回應說明.....	238
附錄五：期末報告審查會議紀錄.....	243
附錄六：期末報告審查意見回應說明.....	252

表目錄

表 1	AI 系統資料保護檢核表（英國 ICO）	40
表 2	人工智慧開發使用自我評估核對清單（香港 PCPD）	84
表 3	AI 法規及指引比較表	88
表 4	實務案例比較表	110

圖目錄

圖 1	AI 系統風險層級示意圖.....	18
圖 2	AI 決策對個人之損害矩陣模型（新加坡 IMDA & PDPC）....	72
圖 3	AI 開發與部署之三階段（新加坡 IMDA & PDPC）	73
圖 4	消費者與 AI 互動之資訊提供（新加坡 IMDA & PDPC）	76

第一章 研究主旨

第一節 研究主題

蒐集國內外政府或民間團體數位轉型及導入人工智慧後，在政策、法規及制度修正之案例，梳理爭點與解決方案，提供 1999 臺北市民當家熱線導入人工智慧技術後，服務流程再設計之參考(包含法規面、制度面及流程面)。

第二節 研究緣起與動機

隨著全球邁入數位化時代，落實數位轉型、構建數位社會已成全球趨勢。政府服務之數位化、智慧化是其中重要面向。我國行政院於 109 年核定國家發展委員會（國發會）「服務型智慧政府 2.0 推動計畫（110 年至 114 年）」，以「擁抱數位未來，打造開放與創新的智慧政府」為願景，旨在推動政府開放資料推動力道更強勁、數位服務更簡單好用、政府決策模式更快速精準發展，力求資料驅動政府服務變革，藉由數位治理創造公共服務價值，縮小民眾與政府資訊不對稱差距，提升施政透明與創新政府數位服務¹。該計畫現已進入第二年期程。秉持擁抱數位轉型的宗旨，服務型智慧政府之基本元素包括²：

¹ 國家發展委員會，服務型智慧政府 2.0 推動計畫(110 年至 114 年)，109 年 8 月，頁 38。

² 國家發展委員會，服務型智慧政府 2.0 推動計畫(110 年至 114 年)，109 年 8 月，頁 29。

一、資料治理

政府採用各項新興科技前，必須檢視內部資料治理的成熟度，從資料蒐集、處理、利用與刪除等生命週期中，建構資料管理措施，分別從資料標準、資料傳出、資料管理、資料應用、資料安全、資料隱私等角度剖析政府「資料治理」之作為。

二、以民為本

發展「由民需觸動服務變革」的數位轉型模式，政府必須從不同對象的角度來設計政府服務，並根據外部需求，跨部會整合既有數位服務。

三、服務優化

優化不只是一要節省政府行政成本或提升運作效率，而是解決民眾生活面臨的困難，聯合政府各部門資源，統合服務能量。

基於前述三項基本元素，服務型智慧政府 2.0 推動計畫策略四「深化新興科技應用之智慧服務」即強調「積極採用人工智慧應用加強為民服務模式」，提供民眾更好的服務體驗。

然而，政府數位轉型或導入人工智慧的過程，強調以「資料驅動」優化服務決策，勢將涉及民眾權利遭受干預之風險。在 2019 年由國

際隱私權專家協會（International Association of Privacy Professionals，IAPP）在新加坡舉辦之亞洲隱私論壇（IAPP Asia Privacy Forum 2019）中，新加坡個資主管機關代表 Lanx Goh 曾提出新加坡政府對待人工智慧之態度，希望能以不斷進步之立法，在個資保護與創新之間取得平衡，建立可信賴之環境。並提醒業界於利用人工智慧科技時，應特別注意「告知、同意、目的限制、正確性、保護性、保存期限、資料移轉限制、當事人近取權和更正權、透明性」等九個重點。

據此，臺北市政府（以下稱北市府）配合國發會服務型智慧政府 2.0 推動計畫，致力推動數位轉型，現於 1999 臺北市民當家熱線（以下稱 1999 熱線）擬導入人工智慧優化市民服務，必將涉及對於民眾個人資料於我國個人資料保護法中的法規遵循（包含蒐集、處理與利用資料之法律依據、目的限制原則、必要性原則、公開透明原則下的告知義務、知情同意、當事人權利如近用、刪除等權利保障、委外管理、安全維護等領域），即有事前評估法規風險、權利保障或倫理疑義之必要。

第三節 研究目的

自法規面、制度面與流程面執行研究，就相關議題提出因應建議及配套措施，作為北市府推動數位轉型規劃之參考依據。其中：

一、個資保護影響評估

本研究執行之「個資保護影響評估」，可評估北市府現行對於 1999 當家熱線規劃導入之人工智慧應用方式可能面臨的法律風險，並據此於本研究提出解決措施建議。

二、比較分析

本研究執行之「比較法規、文獻研究」可彙整他國經驗與實務指引，提供北市府 1999 當家熱線法規、制度及流程等服務流程再造之建議，進而提出「臺北市政府開發及應用人工智慧指引」草案，供北市府評估作為日後導入人工智慧之作業依據。

三、焦點座談

本研究執行之「焦點座談」，可期由專家學者就本研究提出之「個資保護影響評估」，及後續提出之因應措施，乃至於本研究於期末階段草擬之「臺北市政府開發及應用人工智慧指引」草案等產出，提供所屬領域之專業意見。

第四節 研究範圍與限制

本研究以歐盟、英國、日本、新加坡與香港為法規、文獻研究對象，除因初步研究所得資源較豐之外，該等人工智慧發展先進國家之

政策與法規亦對我國較具借鏡價值。

案例研究方面，本研究以 1999 熱線於市政服務中之功能定位為基礎，經初步查詢資料後，選定歐盟、英國、新加坡及我國之實際案例作為研究對象。惟因本研究對於所蒐集之大多數公部門案例，未能獲得所涉政府機關對於人工智慧應用的事前個資保護或倫理評估、法規調適之具體資訊（或因國際上尚未通過施行正式的人工智慧相關法律，各機關均回歸個人資料保護及其他法律適用，並遵守特定人工智慧指引，且與 1999 熱線相類似之人工智慧應用多具資訊提供、輔助政府作業之功能，並無實質決策之故）。

另本研究著重民眾權利保護或行政倫理，在為民服務業務中導入人工智慧技術後易生之爭議，據以提出法規面、制度面與流程面之適法建議，是故有關人工智慧應用之技術面實作、評估等議題，將不在本研究中處理（北市府於 109 年曾委託執行「1999 市民熱線導入 AI 人工智慧最適方案評估研究」，業已就 1999 市民熱線導入人工智慧執行技術評估）。

第二章 研究方法

第一節 法規及文獻分析

本研究蒐集他國公部門服務導入人工智慧應用之案例(著重與本案市民進線語音服務較有關聯者)，以及國際實務關於人工智慧開發與使用之法規、指引等文獻，作為本案比較分析標的。所研析資料彙整如下：

一、 法規：歐盟－人工智慧法（草案）

二、 指引文獻

（一）歐盟－可信賴人工智慧倫理指引

（二）英國－AI 與資料保護指引

（三）美國－人工智慧權利法案藍圖

（四）日本－以人為本的人工智慧社會原則

（五）新加坡－人工智慧治理模範框架

（六）香港－開發及使用人工智能道德標準指引

第二節 案例彙整

一、 服務專線服務導入 AI 案例

（一）比利時（佛拉蒙地區）Flemish Infoline

（二）愛爾蘭（稅務局）VDA（Virtual Digital Agent）

(三) 西班牙 (中央政府) 060 專線

(四) 我國 (勞動部) 多元智能電服中心

(五) 我國 (中國信託商業銀行) AI 語音客服

(六) 我國 (長庚醫院) AI 智能客服

二、 語音辨識 AI 案例

(一) 愛沙尼亞 (國家議會) Hans

(二) 德國 (聯邦訊息技術中心) C-19

(三) 德國 (聯邦移民及難民辦公室) 難民語言偵測

(四) 我國 (法務部) 語音辨識系統

三、 AI 聊天機器人案例

(一) 英國 (英國駕照及行照監理局) 聊天機器人

(二) 奧地利 (聯邦數位及經濟事務部) Mona

(三) 丹麥 (腓特烈堡自治市) Kiri

(四) 愛沙尼亞 (統計局) Iti

(五) 芬蘭 (奧盧市) OuloBot

(六) 羅馬尼亞 (克盧日-納波卡市) Antonia

(七) 新加坡 (微軟) 對話式 AI 技術解決方案

(八) 我國 (高雄市政府) 1999 智能客服萬事通

第三節 個資保護影響評估

為評估 1999 市民熱線導入人工智慧對民眾個資保護之衝擊，本研究對現行規劃中之「語音轉文字及文本內容關鍵字分析」應用，參考國際趨勢之「資料治理」措施，業已執行「個資保護影響評估（Data Protection Impact Assessment, DPIA）」（參第五章第一節與附件）。

DPIA 之執行過程略為「識別」、「檢視」、「評估」、「鑑別」、「因應」等 5 步驟：

1、識別蒐集、處理或利用個人資料之行為、範圍與流程，至少包含下列事項：

- (1) 所規劃者屬蒐集、處理或利用個人資料之行為，及該行為之背景、目的與涉及之個人資料。
- (2) 該行為涉及個人資料之保存方式與期限。
- (3) 該行為涉及將個人資料提供予第三人（包含受機關委託之受託者）者，應識別該第三人或其業別。

2、檢視蒐集、處理或利用個人資料之合目的性與合比例性，尤應檢視下列原則之遵循：

- (1) 該行為符合誠信原則。
- (2) 該行為之目的應特定、明確、合法。
- (3) 該行為僅蒐集、處理或利用對於該目的適當、相關且必要之個人資料。

3、評估個人資料保護法之遵循性，至少包含下列事項：

- (1) 依個人資料保護法第 6 條、第 15 條、第 16 條、第 19 條、第 20 條規定，評估蒐集、處理與利用個人資料之法律依據。
- (2) 如以同意為蒐集、處理或利用個人資料之法律依據者，依個人資料保護法第 7 條規定，評估同意之合法性。
- (3) 依個人資料保護法第 8 條及第 9 條規定，評估告知義務之存否與踐行方式。
- (4) 依個人資料保護法第 11 條規定，評估個人資料之正確性維持與保存期限或保存條件。
- (5) 依個人資料保護法第 3 條、第 10 條、第 11 條、第 13 條規定，評估當事人權利行使之原則、例外及行使方式。
- (6) 非公務機關為國際傳輸個人資料者，依個人資料保護法第 21 條規定，評估中央目的事業主管機關之限制。
- (7) 委託他人蒐集、處理或利用個人資料者，依個人資料保護法施行細則第 8 條規定，評估對受託者之監督。
- (8) 依個人資料保護法第 27 條規定，評估個人資料之安全維護措施，及依個人資料保護法第 12 條規定，評估個人資料事故通知之機制。

4、鑑別個人資料之風險，至少應包含下列事項：

- (1) 以當事人權益受影響程度鑑別風險。
- (2) 依據前三項之識別、檢視與評估結果，鑑別風險種類與來源，並宜徵詢機關內外關係人（例如員工、受託者、接收個人資料之第三人）之意見。
- (3) 鑑別風險發生之可能性，以及風險發生對當事人權益影響之嚴重性。

5、依據鑑別之風險，預先規劃足以避免或降低風險之因應措施。

依據個資保護影響評估結果，對該行為或涉及之系統，採取規劃之風險因應措施。並持續監控剩餘風險，即時改善風險因應措施之有效性。

第四節 指引草案與焦點座談

本研究提出「臺北市政府開發及應用人工智慧作業指引」草案，並於 111 年 11 月 29 日邀集相關產官學研代表，針對本研究之個資影響評估報告、前述指引與建議方向，就 1999 臺北市民當家熱線導入人工智慧應用之適法、倫理等疑義，就法規面、流程面與制度面提供意見。

第三章 蒐集之法規、文獻分析

第一節 歐盟：可信賴人工智慧倫理指引

歐盟將 AI 視作「形塑歐洲數位未來（shaping Europe's digital future）」之重要方向。歐盟執委會（European Commission, EC）於 2018 年將提升公部門與私部門對 AI 之投入、提升 AI 在歐盟的使用列作其關鍵目標之一。2019 年，歐盟執委會提出「可信賴人工智慧倫理指引（Ethics Guidelines for Trustworthy AI）」，針對以符合倫理、可信賴之方式設計及使用 AI 提供指引。

一、可信賴 AI 系統三特徵

- 1、合法性（lawful），即 AI 系統應遵循一切相關法令規範。
- 2、倫理性（ethical），即 AI 系統應遵守倫理原則與價值。
- 3、穩健性（robust），即 AI 系統在技術和社會面向，皆應留意防範造成意外傷害。

二、可信賴 AI 系統四倫理原則

- 1、尊重人類之自主權（respect for human autonomy）：與 AI 系統互動之個人，仍應保有充分且有效之自我決定能力。AI 系統不應脅迫、欺騙或操縱人類³，AI 之設計應以人類為中心，旨在輔助與

³ High-Level Expert Group on Artificial Intelligence, Ethics Guidelines for Trustworthy AI (2019), 12.

增強人類之社會文化技能與認知。此即說明應確保人類對 AI 系統之監督，AI 應在工作環境中對人類提供支持。

- 2、避免傷害 (prevention of harm)：AI 不應對人類造成不利影響，亦不應加劇既有之衝突或傷害。AI 系統及其運行環境應確保安全，具備穩健技術，且防範惡意濫用。此外，弱勢族群應於 AI 運用中受到更多關注，並被視為服務對象。
- 3、公平性 (fairness)：AI 系統之開發、部署與利用，必須具備公平性。這一方面要求確保平等與公正之利益分配與成本分擔，另一方面要求個人及群體免受偏見、歧視或污名化對待。AI 系統絕對不可以欺騙或其他不當方式損害人類之自由意志。人類並應有權質疑/挑戰 AI 自動化決策之結果。
- 4、可解釋性 (explicability)：AI 應儘量避免黑箱 (Black Box) 決策，其系統處理程序須公開透明，並儘可能使相關決策結果具備可解釋性，以利實現決策結果之可質疑/挑戰性。此外，特定情形下，可能需要以其他措施實現可解釋性，例如可溯性、可審核性、對系統能力的充分溝通等。

三、可信賴 AI 系統七遵守面向

1、人類自主性與監督 (human agency and oversight)：

- (1) 若某一 AI 系統可能對人權有不利影響，則應在其開發前執行

基本權影響評估。

- (2) 人類使用者必須能夠對 AI 系統之使用作出知情決定，並可挑戰該等系統。個人必須有權不受僅由自動化決策作出之重大決定之拘束。
- (3) 人類必須能夠（視應用領域及風險）監督 AI 系統。

2、技術穩健性和安全性(technical robustness and safety)：

- (1) AI 系統應以防範風險、促進穩定行為、使危害最小化為設計考量。
- (2) AI 系統除需具備準確性、可靠性和可重複性等技術特質，同時也需在出現問題前訂定完善的備援計畫。

3、隱私和資料治理（privacy and data governance）

- (1) AI 系統的整個生命週期應確保充分尊重隱私和資料保護，防範歧視。
- (2) AI 系統必須確保適當的資料治理機制，同時考慮到資料的品質和完整性，並確保合法近用資料為可行。

4、透明度（transparency）

- (1) 資料、系統和 AI 的商業模型應該是透明的、可追溯的。
- (2) 應以利害關係人能夠理解的方式解釋 AI 系統的邏輯及運作模式。人類參與者和使用者需要意識到他們正在與 AI 系統進

行互動，並且必須了解 AI 系統的功能和限制。

5、保持多樣性、不歧視和公平（diversity, non-discrimination and fairness）

- (1) AI 系統不公平的偏見可能會加劇對弱勢群體的偏見和歧視，導致邊緣化現象更為嚴重。為避免此種情況，AI 系統應該設計為所有人皆可以近用，達成使用者多樣性的目標。
- (2) 組織在 AI 系統整個生命週期內，應定期徵求可能受 AI 系統影響的利益相關方之意見。

6、社會和環境福祉（societal and environmental well-being）

AI 系統必須兼顧永續發展、環境友善，並能提供正向的社會影響。

7、課責性（accountability）

應建立機制以妥當處理 AI 所導致的結果的責任歸屬，演算法的可審計性（Auditability）為關鍵。此外，應確保補救措施為無障礙設計。

第二節 歐盟：AI 法草案

一、歐盟 AI 法草案提出背景

2021 年 4 月 21 日，歐盟執委會提出「人工智慧法（Artificial Intelligence Act）」草案。這部草案是世界首部針對人工智慧提出專門性立法草案，旨在構建歐盟關於人工智慧技術的統一規則，從而實現如下目標⁴：

- (1) 確保投入歐盟市場並使用之 AI 系統符合安全標準，且尊重現行法所保障之基本權利和歐盟價值。
- (2) 確保法律確定性，以促進對 AI 的投資與創新。
- (3) 強化現行法中適用於 AI 系統的基本權利和安全要求之治理和有效執行。
- (4) 促進合法、安全且可信賴 AI 應用之單一數位市場之發展，防範市場碎片化。

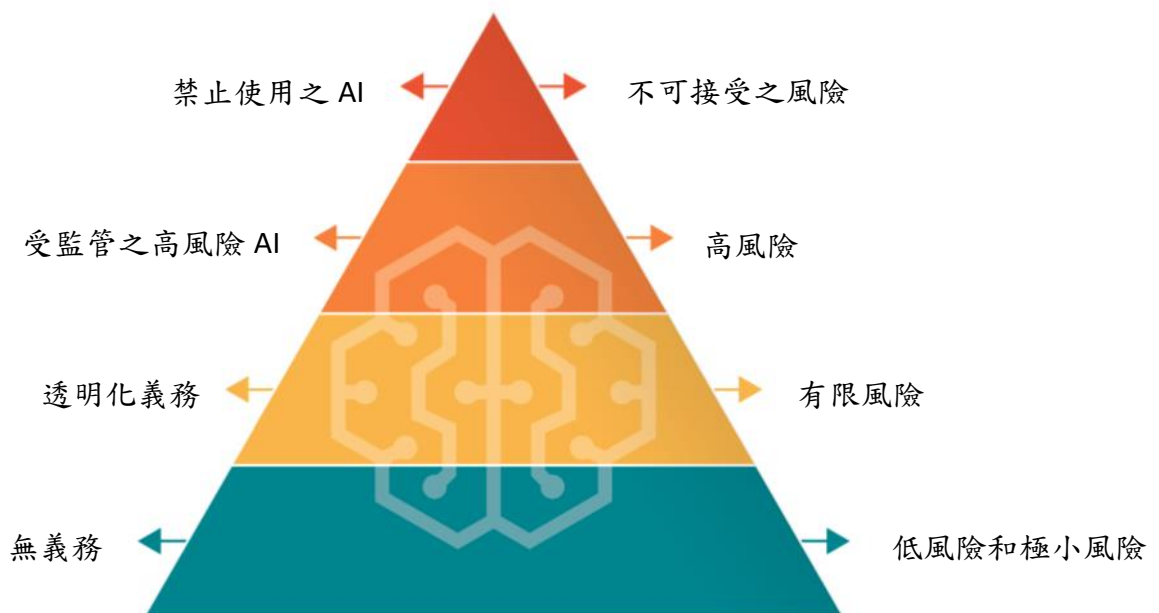
為達成上述目標，該草案採均衡且合比例之水平式（horizontal）監理方式，將監理措施限於因應 AI 相關風險和問題之最小必要範圍內，以免不當限制或阻礙 AI 技術之發展，或不當增加 AI 解決方案投

⁴ European Commission, Proposal for a Regulation of the European Parliament and of the Council Laying Down Harmonised Rules on Artificial Intelligence (Artificial Intelligence Act) and Amending Certain Union Legislative Acts (21.4.2021) (hereinafter the “AI Act Proposal”), COM(2021) 206 final, 3.

入市場之成本。

該草案對「人工智慧」提出技術中性（technology-neutral）且具前瞻性（future-proof）之定義，且採用風險基礎（risk-based）之監理模式，依 AI 系統之風險層級，區分不同監理密度（如下圖所示）。草案並提出促進創新及企業自律之規範，並規劃歐盟及會員國層面之治理機制與執法安排。該草案亦關注個資保護議題，並要求 AI 系統之開發與使用須遵循歐盟個資法「一般資料保護規則」（General Data Protection Regulation, GDPR）之規範。

圖 1 AI 系統風險層級示意圖⁵



⁵ 參考來源：European Parliament Briefing: Artificial Intelligence Act (14/1/2022), <https://digital-strategy.ec.europa.eu/en/policies/regulatory-framework-ai> (accessed 30 August 2022).

二、AI 法草案之監理制度設計

（一） 人工智慧之定義

依該草案第 3 條第 1 項，人工智慧系統（AI system）係指使用該草案附件 I 中所列的一種或多種技術和方法開發之軟體，並且該軟體對於給定的一組人類定義的目標，可生成影響與其互動環境之內容、預測、建議或決定等產出（output）⁶。

該草案附件 I 所列技術包括機器學習等，依該草案第 4 條，歐盟執委會得依市場和技術發展，對附件 I 進行修正，增列性質類似的其他技術。

（二） 該草案之適用範圍

依該草案第 2 條，該草案之適用對象包括：

- 1、 將 AI 系統投放市場或投入使用的供應商，無論其是否在歐盟境內設有據點。
- 2、 歐盟領域內 AI 系統的使用者。
- 3、 位於第三國的 AI 系統的供應商和使用者的，但以其 AI 系統之產出係用於歐盟境內為限。

⁶ AI Act Proposal, Article 3(1), ‘artificial intelligence system’ (AI system) means software that is developed with one or more of the techniques and approaches listed in Annex I and can, for a given set of human-defined objectives, generate outputs such as content, predictions, recommendations, or decisions influencing the environments they interact with.

(三) 風險基礎之 AI 監理框架

考量 AI 技術對歐盟價值觀、基本權利及原則的風險和影響，該草案將 AI 技術可能導致之風險分為四個層級，並分別適用不同監理密度與監理方法：

1、不可接受之風險

某些 AI 作業將會對公民的安全、日常生活和基本權利顯具威脅，故該草案第 5 條予以明令禁止。此類 AI 系統包括：

- (1) 透過影響人類潛意識，操控人類行為，且可能對人類身心造成傷害。
- (2) 利用人類年齡、身心障礙等方面之弱點，操縱其行為，且可能對人類身心造成傷害。
- (3) 直接或間接被公務機關用於社會評量(social scoring)目的，且可能對特定個人或族群造成不公正之影響。
- (4) 在公共場所為執法目的執行即時遠程生物特徵辨識，但在特定例外情形下有絕對使用必要者除外。例外得使用即時遠程生物特徵辨識系統之情形包括：在緊急情形下搜尋特定受害人(例如失蹤兒童)，防範急迫之恐怖主義威脅；偵查、定位、確定或起訴嚴重刑事犯罪之犯罪嫌疑人等，但必須對使用之時間、地理範圍和搜尋資料庫進行適當限制。

2、高風險

該草案第 6 條將高風險 AI 分為兩類：

(1) 對於附件 II 所列歐盟健康和安安全法規所規範之產品，AI 系統為該等產品或該等產品之安全部件。

(2) 附件 III 所列領域之 AI 系統。附件 III 目前包括 8 類領域，未來歐盟執委會可作修正：

A. 對自然人之生物特徵識別（identification）與歸類（categorisation）；

B. 關鍵基礎設施之管理與運作；

C. 教育和職業訓練；

D. 聘僱、員工管理；

E. 重要私部門服務和公共服務之取得；

F. 刑事執法；

G. 移民、難民和邊境控管；

H. 司法和民主進程。

前述兩類高風險 AI 系統皆須符合該草案第 8 條至第 15 條之要求，包括：

(1) 覆蓋 AI 系統完整生命週期的風險評估系統。

(2) 以高品質之資料集訓練、驗證和測試 AI 系統，防範系統歧視

(bias)。

- (3) 訂定並及時更新 AI 系統技術文件。
- (4) AI 系統各項作業應由自動日誌記錄(log)，以利控管和追溯。
- (5) 向使用者充分揭露資訊，協助使用者理解系統產出並適當使用系統。
- (6) AI 系統使用期間，應由自然人實施有效監督。
- (7) 充分確保 AI 系統之準確性、穩健性和安全性。

此外，依該草案第 16 條，前述兩類高風險 AI 之提供者亦有義務執行符合性評估 (conformity assessment)，並將其 AI 系統登錄於歐盟執委會所設立之統一資料庫。

3、有限風險

該草案第 52 條規定，有限風險之 AI 系統須負透明化義務：

- (1) 自然人互動之 AI 系統須讓使用者知悉其正與 AI 系統互動。
- (2) 情緒辨識系統或生物特種歸類系統須讓使用者知悉該系統正在運作中。
- (3) 深偽 (deep fake) 性質之圖像、影像或錄音編輯系統須向使用者明確揭露相關圖像、影響或錄音係人工生成或編輯而得。

4、低風險和極小風險

該草案允許自由開發和使用低風險和極小風險之 AI 系統。

但此類 AI 系統之提供者得依第 69 條訂定自律性行為守則，或自願遵循該草案關於有限風險或高風險 AI 系統之規範。

(四) AI 法之執法機制

該草案第 56 條至第 59 條規定 AI 法之執法機制：

- 1、在歐盟層面，該草案提出設立「歐洲人工智慧委員會」(European Artificial Intelligence Board)，由各會員國及歐盟執委會分別指派代表組成，協助促進 AI 法規範之一致適用，增進會員國主管機關間之合作。
- 2、在會員國層面，各會員國將須設立和指派權責機關負責 AI 法之落實，並從權責機關中，確定一國家監管機關，負責監管高風險 AI 之遵循狀況。

另該草案第 71 條規定 AI 法之罰則。依違反程度不同，違反行為最高可處 3,000 萬歐元或相當於前一年度全球總營業額 6% 之行政罰鍰。

三、AI 法草案與歐盟 GDPR 之適用關係

AI 系統之開發、訓練、測試、使用等與個人資料之蒐集、處理和利用密切相關。AI 程式一方面在開發過程中，可能利用個人資料建立演算法模型，另一方面在使用過程中，可能將其演算法用於分析處理

個人資料，從而得出對個人有影響之產出。

歐盟執委會於提出該草案時明確表示，該草案不影響歐盟個資法 GDPR 之適用⁷。據此，因 AI 系統開發、訓練、測試、使用等目的而蒐集、處理和利用個人資料，須遵守 GDPR 之規範。

該草案之部分條文明確配合 GDPR 要求。例如，該草案第 29 條第 6 項要求高風險 AI 系統之使用者應依 GDPR 之規範，對該 AI 系統之使用執行個資保護影響評估（Data Protection Impact Assessment，DPIA）。

該草案亦包含補充 GDPR 之規範。例如，該草案第 10 條第 5 項允許為 AI 系統歧視監測、偵測和校正目的蒐集、處理和利用特種個人資料，從而補充 GDPR 第 9 條第 1 項關於特種個資利用原則禁止之規範。

四、歐盟理事會對 AI 法草案之審修

2022 年 12 月 6 日，歐盟理事會（Council of the European Union）表示，其已對 AI 法草案達成一致立場（General Approach，以下稱理事會版本），並公布對 AI 法草案的修正版本，其修正重點摘要如下。

⁷ AI Act proposal, Explanatory Memorandum, 4.

（一） 人工智慧之定義

理事會版本調整「人工智慧系統」之定義，刪除原 AI 法草案附件 I，改以描述性方式，將之調整為「具備自主性 (autonomy) 要素，基於機器和（或）人類提供之資料與輸入內容 (input)，利用機器學習和（或）基於邏輯/知識之方法推測實現特定目的之方式，並形成系統生成 (system-generated) 之內容、預測、建議或決定等產出 (output)，從而影響與其互動之環境」。

（二） AI 法草案之適用範圍

理事會版本調整了草案第 2 條所列之適用對象，在原本「在歐盟領域內將 AI 系統投放市場或投入使用的提供者」、「歐盟領域內 AI 系統的使用者」及「AI 系統之產出於歐盟境內使用之第三國 AI 系統的提供者和使用者」基礎上，增加下列三類對象：

- 1、 AI 系統之進口商和經銷商。
- 2、 以自身名義或商標，將 AI 系統與其自身產品一併投放市場或投入使用之產品生產者。
- 3、 提供者在歐盟境內設立據點之授權代表。

（三） 風險導向之 AI 監理框架

理事會版本仍採風險導向之 AI 監理框架，對草案原本「不可接

受之風險」、「高風險」、「有限風險」及低風險和極小風險」四項分類之內容或監理措施加以調整，並在該四項分類之基礎上，增加「通用性 AI 系統（general purpose AI system）」之規範。

1、不可接受之風險

對於草案第 5 條規範之不可接受風險之 AI，理事會版本將作出如下調整：

- (1) 直接或間接用於社會評量（social scoring）目的，且可能對特定個人或族群造成不公正之影響之 AI，皆屬不可接受之風險，而不以「公務機關所用者」為限。
- (2) 在公共場所為執法目的執行即時遠程生物特徵辨識原則仍屬不可接受之風險，但其例外情形作如下調整：其一，將防範急迫威脅之保護範圍，擴張及於「關鍵基礎設施」；其二，將「嚴重刑事犯罪」之範圍，除原本 2002/584/JHA 號理事會框架決定（Council Framework Decision 2002/584/JHA）第 2(2) 條所列之犯罪類型外（且該犯罪在相關會員國之最重刑罰為三年以上有期徒刑），亦納入會員國法律規定之其他犯罪，但以其最重刑罰為五年以上有期徒刑者為限。

2、高風險

對於本草案第 6 條及附件 II、附件 III 所規範之高風險 AI 系

統之範圍，理事會版本作如下調整：

- (1) 對於附件 II 所列歐盟健康和安安全法規所規範之產品，AI 系統為該等產品或該等產品之安全部件時，如「投放市場或投入使用前須經第三方符合性評估（conformity assessment）」，方屬高風險 AI。
- (2) 附件 III 所列領域之 AI 系統，原則仍屬高風險，然若就相關行動或決策而言，該系統之產出純屬附隨性質（purely accessory），且因此不致（not likely）對健康、安全或基本權利構成嚴重風險，則不屬高風險 AI 系統。
- (3) 對於草案附件 III 所列之 8 類領域，理事會版本將「對自然人之生物特徵識別與歸類」調整為「生物特徵」，並對各領域內高風險 AI 系統之內容酌作調整。例如，對於「關鍵基礎設施」領域，除原本之道路交通以及水、瓦斯、暖氣及電力供應外，亦納入「關鍵數位基礎設施管理及運作中作為安全部件的 AI 系統」。

對於高風險 AI 系統應遵循之要求，理事會版本維持草案第 8 條至第 15 條之監理框架，並酌作調整。例如，對使用者的資訊揭露方面，在草案第 13(3)(b)(i)條要求揭露「預定目的」之基礎上，理事會版本進一步要求所揭露之目的包括「該預定使用該高風險

AI 系統之具體位置 (geographical)、行為 (behavioural) 或功能 (functional) 設定」。

3、有限風險

對於草案第 52 條關於有限風險 AI 系統之透明化義務，理事會版本仍然維持透明化義務之基本框架，即與自然人互動之 AI 系統須讓使用者知悉其正與 AI 系統互動；情緒辨識系統或生物特種歸類系統須讓使用者知悉該系統正在運作中；以及深偽性質之圖像、影像或錄音編輯系統須向使用者明確揭露相關圖像、影響或錄音係人工生成或編輯而得，但對於透明化義務之程度與例外作如下調整：

- (1) 理事會版本仍允許為刑事偵防目的而豁免本條之透明化義務，但增訂「須有保障第三人權利與自由之適當安全維護措施 (safeguard)」之限制。
- (2) 對於深偽性質之編輯系統，基於言論自由或藝術與文學自由而主張豁免透明化義務時，須以所涉內容顯屬創造性 (creative)、諷刺性 (satirical)、藝術性或虛構性著作或活動 (programme) 為限。
- (3) 對於透明化義務之落實，理事會版本要求以明確、顯著之方式提供相關資訊，至遲於自然人首次與該 AI 系統互動

或接觸時提供。

4、低風險和極小風險

對於低風險和極小風險之 AI 系統，理事會版本維持草案規範框架，允許此類 AI 系統自由開發和使用，且鼓勵依第 69 條訂定自律性行為守則，或自願遵循本草案關於有限風險或高風險 AI 系統之規範。理事會版本在草案「AI 系統提供者」之基礎上，並將「AI 系統之使用者」亦納入自願性行為守則之主體範圍。

5、通用性 AI 系統

相較於歐盟執委會之草案，理事會版本增訂「通用性 AI 系統」之規範。依第 3(1b)條，「通用性 AI 系統」係指「意在提供圖像或語音辨識、聲音或影像生成、規律偵測、回應提問、翻譯等通用功能，可用於多種情境或與多種其他 AI 系統整合之 AI 系統，而無論其以何種方式投放市場或投入使用」。

依理事會版本第 4b 條，通用性 AI 系統如可作為高風險 AI 系統或其組成部分，則應遵守高風險 AI 之大多數要求，包括第 8 條至 15 條之各項監理措施，以及第 16 條之符合性評估義務。

依理事會版本第 4c 條，若通用性 AI 系統之提供者已於該系統之使用說明或其他所附資訊中，明確排除高風險用途，則將無

須依第 4b 條遵守高風險 AI 系統之相關規範。但此等排除須基於善意作出，且若通用性 AI 系統之提供者獲知該系統於市場上遭濫用之情事，應採取必要且符合比例之因應措施。

（四） AI 法之執法機制

理事會版本對草案第 56 條至第 59 條規定之執法機制亦有調整，包括在歐盟層面，調整「歐洲人工智慧委員會」之組成結構。委員會僅由各會員國指派之代表組成，歐洲個人資料保護監察人（European Data Protection Supervisor，EDPS）和歐盟執委會得參加會議，但無投票權。

此外，針對草案第 71 條規定之罰則，理事會版本在維持最高 3,000 萬歐元或相當於前一年度全球總營業額 6% 之行政罰鍰上限的同時，針對中小企業（small and medium enterprises，SMEs）增訂特別規範，將其裁罰上限降至前一年度全球總營業額之 3%。

第三節 英國 AI 與資料保護指引

2020 年 7 月，英國個資保護主管機關資訊委員辦公室（Information Commissioner's Office, ICO）提出「AI 與資料保護指引（Guidance on AI and data protection）⁸」，說明 AI 開發、訓練、測試、利用等作業依 GDPR 等法規落實個資保護之最佳實務。

該指引並未就「人工智慧」作出明確定義，而是從不同領域之視角，對人工智慧作出描述：

- (1) 從 AI 研究之角度，「人工智慧」係指「使用非人類系統從經驗中學習，並模仿人類智慧活動」之各種方法。
- (2) 從資料保護之角度，「人工智慧」係指「能夠從事通常需人類智慧方可為之的計算機系統的相關理論和實務發展」。

該指引主要包括四個部分，分別為：

- (1) AI 系統資料保護之課責性要求與治理，包括 DPIA。
- (2) AI 系統個資處理之合法性、公平性和透明性，包括法律依據、AI 系統成效之評量與改善，以及降低歧視。
- (3) AI 系統之資料安全和資料最小化義務。
- (4) AI 系統與資料主體權利，包括與個資自動化處理相關之權利。

⁸ UK ICO, Guidance on AI and data protection, <https://ico.org.uk/for-organisations/guide-to-data-protection/key-dp-themes/guidance-on-artificial-intelligence-and-data-protection/> (accessed 30 August 2022).

該指引採取風險基礎之方法，針對以上四項議題，評估 AI 系統對個人權利與自由之風險，並採取適當措施因應該等風險。該指引並提供「AI 與資料保護工具組」(AI and data protection risk toolkit)，以自檢表單之形式，協助組織識別和管理 AI 所涉之資料保護風險和法遵挑戰。

一、AI 系統之課責性

AI 系統可顯著提升組織之效率與能力，但也可能給個人之權利與自由造成風險，並使組織面臨法令遵循方面的挑戰。AI 系統之個資保護影響依具體應用案例 (use case)、使用之對象、法規要求，以及社會、文化和政治因素而不同。為管理 AI 系統造成之風險，組織應：

- (1) 評估 AI 使用所造成之風險。
- (2) 確定風險因應措施。
- (3) 明確因應措施對 AI 使用之影響。

(一) AI 系統之 DPIA

1、DPIA 之必要性判斷

雖然並非所有 AI 系統皆可能造成權利與自由之高風險，但依 GDPR 第 35 條第 3 項第 a 款，若 AI 系統有如下情形之一，則須於 AI 系統開始處理個資前，執行 DPIA：

- (1) 關於自然人之系統性及大規模的個人特質評估，而該評估是基於自動處理，包含剖析 (profiling)，且基於該評估作成關於該自然人之法律效果或其他重大影響該自然人之決定；
- (2) 處理大規模之特種個人資料，前科及犯罪相關之個人資料；
- (3) 大規模系統性監控公共區域。

2、DPIA 之內容與執行方式

AI 系統之 DPIA 應至少包含如下內容：

- (1) 對 AI 系統個資處理活動之系統性描述，包括資料流和自動化決策對個人產生影響的階段。
- (2) 對 AI 系統資料處理必要性和合比例性之分析。
- (3) 所識別之對個人權利與自由之風險。
- (4) 風險因應措施。
- (5) 擬採取之額外措施及剩餘風險。

(二) AI 系統之控管者與處理者

AI 系統之開發和使用通常有多個組織共同參與，若一個或多個組織對下列事項作出決定，則其為該 AI 系統之控管者：

- (1) 是否蒐集個人資料；
- (2) 所蒐集之個人資料類別；
- (3) 個人資料之利用目的；

- (4) 所涉及之資料主體；
- (5) 資料保存期限；
- (6) 如何回應資料主體之權利行使請求。

若 AI 開發或使用過程中，某一組織並非為自己之目的處理個人資料，而是依他人之指示行事，則該組織通常為處理者。資料處理者仍可能依契約就技術事項作出決定，例如：

- (1) 所使用之 IT 系統和方法；
- (2) 如何儲存資料；
- (3) 資料安全維護措施；
- (4) 如何檢索、傳輸或刪除資料。

二、AI 系統個資處理之合法性、公平性與透明性

（一）AI 系統個資處理之合法性

AI 系統之開發和部署涉及為不同目的、以不同方式處理不同個人資料。個資處理之合法性原則要求組織區分各項不同個資處理作業，並分別確認其處理目的及法律依據。

1、同意

若組織與資料主體間存在直接關係，則歐盟 GDPR 第 6 條

第 1 項第 a 款之「同意」(consent) 可作為 AI 系統處理個人資料

料之法律依據。然而，組織必須確保同意係自主、具體、知情、明確⁹，且由資料主體以積極行為為之。

2、公共任務

若組織在履行公共職務或符合公益之法定任務的過程中使用 AI，例如公務機關使用 AI 提供公共服務，則歐盟 GDPR 第 6 條第 1 項第 e 款之「公共任務」(public task)¹⁰可作為在必要範圍內處理個資之法律依據。

如 AI 系統涉及處理特種個人資料，則除具備適當法律依據外，亦須遵守 GDPR 關於特種個資利用之規範。

如 AI 系統涉及純自動化決策 (solely automated decision-making)，則須遵守 GDPR 第 22 條之要求，僅得於訂立或履行契約所必要、經法律授權、當事人明確同意時方可為之。

(二) AI 系統個資處理之公平性

若組織使用 AI 系統推測關於個人之資訊，則個資處理之公平性原則要求該組織確保：

⁹ GDPR, Article 4(11).

¹⁰ 依 GDPR 第 6 條第 1 項第 e 款，「執行符合公共利益之職務或行使控管者享有之公權力所必須」係蒐集、處理或利用個人資料之法律依據之一。英國 ICO 將該款法律依據之兩種情形合稱為「公共任務」，參見：UK ICO, Guide to the UK General Data Protection Regulation (UK GDPR), Public Task, <https://ico.org.uk/for-organisations/guide-to-data-protection/guide-to-the-general-data-protection-regulation-gdpr/lawful-basis-for-processing/public-task/> (accessed 2 December 2022).

1、該 AI 系統具統計上之準確性 (statistical accuracy)，並避免偏見和歧視。

2、考量該等作業對個人合理期待之影響。

「統計上之準確性」之概念與個資法之「資料正確性 (data accuracy)」原則不同。許多情形下，AI 系統之產出結果係對個人之「推測」，而非事實性資訊。「統計上之準確性」係反映 AI 系統產出之推測之準確性。個資法之資料正確性原則並不要求 AI 系統統計上之準確性達 100%，但組織應確保 AI 系統之推測結果符合統計規律。

適當確定 AI 系統統計上之準確性，須衡量如下兩類錯誤：

1、偽陽性錯誤 (第 I 型錯誤)，即 AI 系統誤將非匹配對象標記為「匹配」。

2、偽陰性錯誤 (第 II 型錯誤)，即 AI 系統誤將匹配對象標記為「非匹配」。

AI 系統之歧視，可能有多種原因，例如訓練資料之內容代表性不均衡、訓練資料反映過往之歧視、開發者之文化觀念、模型之部署方式等。相關因應方式包括：

1、對於不均衡之資料，透過增加代表性不足之資料、刪減代表性過多之資料，增強均衡性。

- 2、對於反映過往歧視之資料，修改資料內容、變更學習方式，或在訓練後調整模型等。

（三） AI 系統個資處理之透明性

個資處理之透明性原則要求該組織對其 AI 系統之個資處理活動保持透明。然英國 ICO 另就 AI 系統透明性議題制定單獨指引¹¹，故本指引對此未作深入說明。

三、AI 系統之資料安全與資料最小化義務

（一） AI 系統之資料安全

- 1、AI 系統通常採機器學習技術，需使用大量訓練資料集與測試資料集，且所涉資料可能有多種不同來源、以不同格式儲存於不同位置，可能增加資料之追溯與管理之難度。
- 2、AI 系統之技術人員應對個資之移動與儲存保留紀錄，以利採取適當安全維護措施，並評估該等措施之有效性。組織並應就 AI 系統保留明確之稽核軌跡。
- 3、含有個資之檔案如不再需要，則應儘速刪除。

¹¹ UK ICO, Explaining decisions made with AI, <https://ico.org.uk/for-organisations/guide-to-data-protection/key-dp-themes/explaining-decisions-made-with-ai/> (accessed 30 August 2022).

- 4、依資料處理對個人造成風險之可能性與嚴重性，組織可能需要在從原始檔案中提取訓練資料、在內部或外部分享訓練資料前，對個資進行去識別化處理。
- 5、AI 系統可能面臨「模型逆推攻擊」(model inversion attack)、成員推論攻擊 (membership inference attack) 等攻擊，從而使攻擊者得以推知 AI 系統訓練資料所含之個資。因此，AI 系統之訓練者應評估其模型是否包含可直接或間接識別特定個人之資料、遭攻擊時資料被揭露之風險，並採取因應措施。

(二) AI 系統之資料最小化義務

AI 系統之開發通常需要大量資料，因而可能為個資法之資料最小化原則帶來挑戰。AI 系統之開發與部署過程中，組織應採取適當措施，落實資料最小化原則：

- 1、自 AI 系統設計階段即充分考慮資料最小化要求。如使用第三方提供之 AI 系統，則應於採購之盡職調查中，納入資料最小化評估考量。
- 2、充分瞭解 AI 系統之資料處理流程。
- 3、適當衡平資料最小化要求與統計上之準確性要求。若使用較少資料即可實現足夠準確性，則不應處理更多資料。

- 4、評估使用隱私強化技術（privacy-enhancing technology），例如擾動（perturbation）或增加雜訊（noise）、合成資料（synthetic data）、聯合學習（federated learning）等。

四、AI 系統與資料主體（當事人）權利

AI 系統之開發與部署，仍可能涉及個人資料。雖然 AI 系統需將資料從原始形式處理為模型訓練、測試等所需之適當形式，且可能不含身分識別資訊、聯絡資訊等，但若經處理之資料能夠將特定個人與其他人相區別（single out），則仍屬個人資料。資料主體有權依個資法對該等資料行使權利。

- 1、組織須向資料主體告知個資用於 AI 系統訓練之相關資訊，並確保以公平且透明之方式處理個資。
- 2、對於 AI 系統訓練資料集內之個資，更正權、刪除權等仍有適用；
- 3、對於 AI 系統之訓練資料集，通常不適用資料可攜權，然用於製作資料集之原始資料，仍有資料可攜權之適用。
- 4、若組織就其掌握之資料，無從識別特定資料主體，則不適用 GDPR 第 15 條至第 20 條規定之資料主體權利（近用權、更正權、刪除權、限制處理權、資料可攜權）。然而，若資料主體提供實現識別性之補充資訊，組織不得拒絕。

5、組織得以資料主權權利行使請求「顯無理由」(manifestly unfounded)或過度(excessive)為由，拒絕滿足其請求。但不得僅以 AI 系統較難滿足該請求為由，拒絕滿足其權利行使請求。

五、AI 與資料保護工具組

該指引提供「AI 與資料保護工具組」，從 AI 系統之業務需求與設計、資料獲取與準備、訓練與測試、部署與監測四個階段，以自檢表單之形式，協助組織識別和管理 AI 所涉之資料保護風險和法遵挑戰。該工具組之主要內容，可摘要做成如下檢核表：

表 1 AI 系統資料保護檢核表 (英國 ICO)

AI 系統資料保護檢核表		
1. 業務需求與設計階段		
1.1	DPIA 之執行	<u>須</u> 對高風險之資料處理活動執行 DPIA。
		<u>須</u> 於 DPIA 執行過程中適當徵求可能受影響之個人的意見。
		於識別出無法降低之高風險時， <u>須</u> 諮詢主管機關意見。

		於 DPIA 執行過程中， <u>應</u> 諮詢組織內部資訊、法務等相關同仁之意見。
1.2	個資保護權責分配	<u>應</u> 指定由某一高層人員負責落實課責性原則。
		<u>應</u> 訂定 AI 系統個資保護之執执行程序。
1.3	AI 系統之資料處理紀錄	<u>須</u> 以隱私權聲明等形式，事前向資料主體告知 AI 系統資料處理之目的。
		<u>應</u> 考慮執行 AI 系統資料流盤點。
1.4	評估 AI 系統之公平性風險並採取因應措施	初始評估 <u>應</u> 以資料處理對個人之預期影響為重點。
		風險識別過程中 <u>應</u> 與利益相關方互動。
1.5	評估 AI 系統之可解釋性與透明性	<u>須</u> 於資料蒐集時即向資料主體告知隱私相關資訊。如為間接蒐集，則取得資料後合理期間內告知。
		<u>應</u> 依資料處理之領域、背景等，評估 AI 系統解釋之內容和範圍。

		<u>應</u> 評估 AI 系統產出之易懂性，以確定 AI 系統解釋之詳盡程度。
1.6	評估 AI 系統之安全風險並採取因應措施	<u>須</u> 考量 AI 系統與現有系統整合之安全風險，並在設計和開發階段即採取風險控管措施。
		<u>應</u> 考量安全事故通報和因應程序。
		<u>得</u> 考慮向專家諮詢安全維護之最新技術。
1.7	記錄 AI 系統訓練資料並評估其準確性 (accurate)、適足性 (adequate)、相關性 (relevant) 與目的限制遵循性	<u>應</u> 於隱私權聲明中詳述資料保存期間、留存必要性評估之執行週期。
		<u>應</u> 確保定期執行資料必要性之內部評估。
		<u>得</u> 評估使用隱私強化技術 (privacy-enhancing technology)。
1.8	資料主體權利行使紀錄	<u>應</u> 對 AI 系統所處理之個資制定索引，以利依資料主體之請求檢索相關資料。

		<u>應</u> 對開發者進行教育訓練，使其瞭解系統須具備適當功能，以供組織回應資料主體權利行使請求。
		<u>得</u> 執行使用者測試，取得其對隱私相關資訊之回饋意見。
1.9	評估人為干預之時間點、執行者、決策考量要素	對於 GDPR 第 22 條規範之純自動化決策， <u>須</u> 確保設立有效人為干預機制。
		<u>須</u> 確保前述人為干預機制中，審核人員具備否決自動化決策之權限，且人為決策時考量自動化決策以外之其他要素。
		<u>應</u> 確保審核人員能夠干預自動化決策，並就其決策時所見資訊保持紀錄。
2. 資料獲取與準備階段		
2.1	識別並記錄個資處理之法律依據	<u>須</u> 識別個資處理之法律依據。如涉及特種個資，並識別處理特種個資之例外事由。

		<u>須</u> 於隱私權聲明中列明資料處理之目的及相應法律依據。
		如所涉法律依據有「必要性」要求， <u>須</u> 評估處理與所涉目的間符合比例。
2.2	評估訓練資料集之 代 表 性 (representative)、 可靠性與相關性	<u>應</u> 考量降低偏見之適當技術措施，例如重新加權（re-weighting）。
		<u>應</u> 考量過往歧視之風險
		<u>得</u> 研究 AI 系統可能影響之對象，並識別歧視之風險
2.3	評估可能涉及歧視、因而不用於決策之資料來源或特徵，並作成紀錄	<u>得</u> 考量蒐集此等特徵，用作歧視分析。 歧視分析係一獨立處理目的，應具備獨立之法律依據與處理條件。
2.4	評估為歧視分析目的所蒐集之保護特	<u>須</u> 就歧視分識別法律依據與處理條件。
		<u>須</u> 確保資料主體知悉其資料被用於歧視分析。

	徵 ¹² 資料，並作成紀錄	<u>應</u> 考慮是否需處理額外資料、是否為所掌握資料貼標籤（label）等。
2.5	記錄資料貼標籤之明確標準與責任機制	<u>應</u> 制定貼標之標準，且該標準應易懂、對各標籤進行適當描述、提供各標籤之例示，並包含極端情形之處理方式。
		<u>應</u> 製作貼標與註記（annotation）之訓練手冊。
2.6	評估所用模型擬合過度之可能性，並作成紀錄	<u>須</u> 移除可能導致擬合過度(overfitting)之特徵。
		<u>應</u> 檢測模型成效數據（例如精確率『Precision』、召回率『Recall』），以確定擬合過度問題之根源。
2.7	評估所處理之個資及其處理目的，並透過隱私相關資訊	<u>須</u> 以便於取得且易於理解之方式，使用清晰簡明之語言，向資料主體告知個資處理活動

¹² 保護特徵（protected characteristics）係英國《2010年平等法（Equality Act 2010）》規定之概念。依該法第4條，保護特徵包括：年齡、身心障礙、性別重置、婚姻及伴侶關係（civil partnership）、懷孕與生育狀況、宗教或信仰、性別、性取向。

	揭露	<u>須</u> 告知資料處理之目的、法律依據，以及所涉個資之類別。
		<u>應</u> 測試隱私相關資訊對於 AI 系統之使用對象而言是否便於取得且易於理解。
2.8	於提取和分享訓練資料前，對個資作去識別化處理	<u>應</u> 評估使用符合用例需求之隱私強化技術。
3. 訓練與測試階段		
3.1	評估當前處理目的與原始目的是否不同	若資料之利用或揭露目的與原始目的不同， <u>須</u> 確保為新目的利用個資之公平性、合法性與透明性。
		<u>須</u> 更新隱私相關資訊及相關文件，增列新的處理目的。
		<u>須</u> 考量所依據之同意是否足以涵蓋新目的。
3.2	評估 AI 系統能否以	<u>應</u> 考量是否需就特定族群蒐集更多資

	公平且準確之方式 處理多種不同族群 之資料	料，以提升結果準確性。
		就蒐集更多資料（以提升結果準確性） 或不蒐集更多資料（以避免對該等族群 其他權利與自由之不利影響）之決定， <u>應</u> 進行評估並說明正當理由。
3.3	測試AI系統是否就 保護特徵不同之個 人，產出類似結果	<u>應</u> 評量各類錯誤（例如偽陽性、偽陰性 等）。
		<u>應</u> 考量是否測試資料集適足性。
		<u>應</u> 記錄模型在統計上之準確性方面的 限制。
3.4	評估AI系統之可解 釋性，並考量使用 輔助工具向受影響 之個人解釋AI系統	若以純自動化決策作出對個人有法律 或類似重大影響之決策， <u>須</u> 就決策邏輯 提供有意義之資訊。
		<u>應</u> 針對受影響之個人，測試解釋之有效 性。
		解釋之設計， <u>應</u> 滿足不同階段與AI系

		統互動之需要（例如首次與 AI 系統互動者、使用自動化處理之結果支援關鍵決策者）。
3.5	嚴格管控訓練資料、訓練程式碼、部署程式碼之存取權限，並作成紀錄	<u>應</u> 考慮權限最小原則（least privilege）之要求，定期檢視及取消不必要之權限，並保留權限變更紀錄。
3.6	稽核訓練與測試階段個資之移動與儲存，並保留軌跡	<u>應</u>建立並及時更新 AI 系統清冊。 <u>應</u>製作安全程序文件（包括事故通報與處理程序），並確保 AI 系統開發和部署之各參與者能自由取得該文件。
3.7	評估所使用軟體之安全風險並採取因應措施	<u>應</u>對軟體之安全性執行內部或委外測試。 若軟體係委外維護，<u>得</u>訂閱安全建議示警並確保落實修補/升級程序。
3.8	再次評估資料對 AI 系統訓練和測試之	<u>須</u> 考量訓練和測試過程中對資料之複製。

	必要性、適足性和相關性，刪除不必要之資料。	<u>應</u> 考量資料最小化和統計準確性之平衡。
3.9	設計並執行審核人員教育訓練	對於 GDPR 第 22 條之純自動化決策， <u>須</u> 確保 AI 系統進行人為干預之有效性。
4. 部署與監測階段		
4.1	設計定期測試機制並作成紀錄	如利用個人資料執行測試， <u>須</u> 告知資料主體。
		<u>應</u> 確立模型漂移(model drift)之指標，以及模型漂移引發後續審查與測試之標準。
		<u>應</u> 保留模型之各個版本，以利追蹤重大漂移之來源。
4.2	設計補救及回饋機制，以供個人對 AI 系統之解釋提出意	<u>得</u> 積極與個人溝通，以利改善 AI 系統之解釋。

	見	
4.3	界定降低安全風險之技術性與組織性措施	<u>應</u> 評估模型可解釋性與安全事故間之衡平。
		<u>應</u> 積極監測 AI 系統表現，調查異常狀況。
		<u>應</u> 導入異常狀況即時偵測技術。
		<u>得</u> 禁止匿名使用 AI 系統，要求使用者皆表明身分。
4.4	設計模型成效監測機制。若發生模型漂移，評估並移除不適足或不相關之訓練資料	<u>須</u> 定期評估模型漂移，並於必要時以新資料重新訓練模型。
		<u>應</u> 依處理之性質、範圍、背景、目的和風險等，確立重新訓練之啟動標準。
4.5	設計人為干預有效性之確保機制	<u>得</u> 定期測試審核人員能夠識別故意作出之不準確決策。
		<u>得</u> 將審核人員否決之決策及否決原因

		作成紀錄。
4.6	評估 AI 系統之目的變更，確保變更符合法律要求	若資料之利用或揭露目的與原始目的不同， <u>須</u> 確保為新目的利用個資之公平性、合法性與透明性。
		<u>應</u> 定期檢視處理作業、文件和隱私權聲明，以確認現行處理目的未超出原始目的。

第四節 美國人工智慧權利法案藍圖

2022 年 10 月，美國白宮科技政策辦公室（White House Office of Science and Technology Policy）發布「人工智慧權利法案藍圖：令自動化系統為美國人民服務（Blueprint for an AI Bill of Rights: Making Automated Systems Work for the American People）」之政策文件（以下稱「人工智慧權利法案藍圖」）¹³。

該藍圖針對「自動化系統（Automated System）」之設計、使用與部署，提出下列五大應遵守之原則，以期保護民眾在人工智慧時代之權益¹⁴：

- 1、安全有效之系統（Safe and Effective Systems）；
- 2、演算法歧視之保護措施（Algorithmic Discrimination Protections）；
- 3、資料隱私（Data Privacy）；
- 4、告知與解釋（Notice and Explanation）；
- 5、人為替代、考量及備案（Human Alternatives, Consideration, And Fallback）。

¹³ White House Office of Science and Technology Policy, The Blueprint for an AI Bill of Rights: Making Automated Systems Work for the American People (October 2022) (hereinafter “The Blueprint for an AI Bill of Rights”), <https://www.whitehouse.gov/wp-content/uploads/2022/10/Blueprint-for-an-AI-Bill-of-Rights.pdf> (accessed 04 January 2023).

¹⁴ The Blueprint for an AI Bill of Rights, 4.

該藍圖並提出「從原則到實踐：人工智慧權利法案藍圖技術指引」
(From Principles To Practice: A Technical Companion to The
Blueprint for An AI Bill Of Rights)，詳述各項原則之重要性 (Why
This Principle Is Important)、期待達成之遵循效果 (What Should
Be Expected of Automated Systems)、落實期待效果之相關例示
(How These Principles Can Move Into Practice)，為各項原則之遵
循與落實提供方向性指引。

該藍圖所列五大原則之主要內容摘要如下：

一、安全有效之系統

自動化系統應符合下列期待，以保護民眾免受不安全或無效系統
之傷害：

1、持續主動保護民眾免受自動化系統傷害

(1) 諮詢：在自動化系統之設計、應用、部署、取得、維護等

各個階段均應徵求民眾意見。諮詢應直接面向受影響之不
同族群，並且考慮該系統對其所特有、或對其而言格外嚴
重之疑慮或損害。

(2) 測試：測試系統在部署前應接受大量測試，且其測試應儘
可能遵循相關領域之最佳實務做法。

(3) 風險識別與因應：自動化系統於部署前，應持續主動識別

並因應其潛在風險，尤其是該系統對民眾權利、機會以及近用管道（access）之潛在影響。

(4) 持續監測：自動化系統應設有持續性監測程序（包括重新校准程序），以確保其性能之維持。監測之範圍應同時包括技術系統部分及人為作業部分。

(5) 明確的組織監督：開發或應用自動化系統之組織應當建立明確的治理架構與程序。

2、避免利用不適當、低品質或無關資料，以及該等資料再利用所致之複合型傷害（compounded harm）

(1) 相關且高品質之資料：構建、評估或部署自動化系統所用之資料應當具備相關性、高品質，且符合利用之特定需求。資料之屬性（attribute）與來源之正當化理由亦應予以記錄。

(2) 對衍生資料來源之追蹤與評估：透過演算法從其他資料所衍生之資料，應予以識別與追蹤。

(3) 敏感領域（sensitive domain）資料之再利用限制：刑事司法、財務、就業、住房等敏感領域資料之再利用，應受額外檢視，以確保安全性與有效性。

3、證明該系統的安全性和有效性

(1) 獨立評估：自動化系統之設計應提供獨立評估之可行性。

獨立評估執行人員應能夠取得進入系統以及獲取相關資料樣本之權限。

(2) 報告：開發或應用自動化系統之組織應定期更新其報告，

報告內容應包括系統概述、資料利用狀況描述、公眾諮詢結果、風險識別及因應措施、性能測試結果、持續監測程序、獨立評估程序及結果等。報告應當使用簡明易懂之語言，並使用機器可讀之格式。

二、演算法歧視之保護措施

自動化系統應以公平方式設計和應用，以保護民眾免受演算法歧視：

1、持續主動保護民眾免受自動化系統傷害

(1) 設計中主動執行公平性評估（equity assessment）：在自動化系統之設計或取得階段，應進行主動執行公平性評估，評估範圍應儘可能納入各類弱勢族群。評估可對系統進行定性與定量評價。

(2) 具代表性之穩健資料：用於系統開發或評估之資料對預期部署情境所涉族群有充分代表性，且應具充分穩健性，以識別及協助因應偏見和潛在傷害。

- (3) 防範代理屬性 (proxy) 所致歧視：在自動化系統之設計、開發或部署中，與人口特徵 (demographic feature) 高度相關的屬性 (即代理屬性) 易造成演算法歧視。因此，特徵資料與資料屬性間之對應關係，識別並移除代理屬性。
- (4) 於設計、開發與部署階段確保可及性 (accessibility)：系統之設計、開發與部署，應確保其對於身心障礙人士之可及性。
- (5) 差異性評估 (disparity assessment)：應對自動化系統進行廣泛測試，評估系統之組成部分是否會對不同族群造成差別待遇。
- (6) 差異性因應：若所識別出的差別待遇可能導致歧視或實質損害，應採取措施予以因應。
- (7) 持續監測及緩解：系統應進行持續監測，以便識別部署後所發生之演算法歧視。

2、證明該系統足以保護民眾免受演算法歧視

- (1) 獨立評估：應用或管理自動化系統之組織應允許對其系統可能導致之演算法歧視進行獨立評估。
- (2) 報告：開發或應用自動化系統的組織應提供之演算法影響評估報告，其內容至少應包括：諮詢之結果、設計階段的

公平性評估、可及性設計及測試、差異性評估、剩餘差別待遇之紀錄，以及因應措施及其評估。演算法影響評估結果應儘可能予以公開。報告應當使用簡明易懂之語言，並使用機器可讀之格式。

三、資料隱私

自動化系統應透過內建之保護措施，保護民眾免受資料濫用行為之影響，保障民眾對其資料利用之自主性：

1、隱私保護設計（by design）與預設（by default）

- (1) 隱私設計與預設：自動化系統之完整開發週期內，應持續評估隱私風險，並採取適當之技術性和政策性風險因應措施。
- (2) 資料蒐集與使用案例範圍限制：資料蒐集應限於經明確界定之特定目的範圍內，且應限於最小範圍。為特定目的蒐集之資料，非經另行隱私風險評估及採取因應措施，不得用於其他情境。
- (3) 風險識別與因應：蒐集、利用、揭露或儲存敏感資料之組織應主動識別相關風險，並採取因應措施。
- (4) 保護隱私之安全措施：構建、利用或管理自動化系統之組織應遵循隱私與安全保護之最佳實務做法，保護資料及詮

釋資料 (metadata) 之揭露限於經同意之特定使用案例。

2、保護民眾免受恣意監控

(1) 強化對監控行為之監督：監控系統應持續接受強化監督，

且監督內容至少應包括對潛在侵害之評估。

(2) 有限且符合比例之監控：非基於實現正當目的之絕對必要、

且與需求成比例，不得進行監控。

(3) 限定監控範圍，以保護權利與民主價值：民主權利與自由

不應受監控系統威脅，亦不應受自動化系統輔助之騷擾行

為限制。

3、向民眾提供適當且有意義之同意、資料近用及管控機制

(1) 針對特定利用之同意：資料蒐集者或自動化系統徵求同意

時，應針對限定利用情境、利用時間及具體利用者的特定

同意。

(2) 以簡明直接方式徵求同意：向使用者徵求同意時，應使用

簡明易懂之語言，以利使用者理解利用之情景、時長及利

用者。

(3) 資料之存取與更正：使用者之資料經自動化系統蒐集、利

用、揭露或儲存者，應有權存取該等資料，並於必要時更

正其資料。

- (4) 撤回同意與資料刪除：組織應在法律規定範圍內，允許使用者撤回利用其資料之同意，並在撤回同意後及時從系統中移除該資料。
- (5) 自動化系統支援：設計、開發和部署自動化系統之組織應建立並維護相關功能，允許個人利用其自身之自動化系統作出同意、存取資料，並管控複雜資料系統中之決策。

4、證明資料隱私及使用者控制權受到保護

- (1) 獨立評估：應用或管理自動化系統之組織應允許對其資料政策之獨立評估。
- (2) 報告：當民眾想要瞭解其資料在系統中的利用狀況時，開發該系統之組織應儘速就資料蒐集或儲存狀況提供報告。

5、為敏感領域相關資料提供額外保護

- (1) 僅限必要功能：敏感資料原則僅可用於其所屬領域絕對必要之功能，或為行政目的所必要之功能。
- (2) 倫理審查與利用禁止：敏感資料之利用如可能限制權利、機會或近用管道，應經受倫理審查。特定情形下，倫理審查可能判定縱經同意，該資料亦不得利用或揭露。
- (3) 資料品質：在敏感領域，組織更應維持資料之品質。
- (4) 限制敏感資料及衍生資料之存取：敏感資料及其衍生資料

不應被經資料中介或其他契約而販售、揭露或公開。

- (5) 報告：蒐集、利用、儲存或揭露敏感資料之組織應定期公開敏感資料報告，包括資料安全疏失或事故、倫理審查結果等。

四、告知與解釋

自動化系統應符合下列期待，以確保民眾知悉自動化系統正在運作中，且瞭解該系統如何以及為何協助產出對個人造成影響之結果：

1、提供明確、及時、易於理解且易於取得之告知與解釋

- (1) 語言簡明、易於取得之文件：組織應公開揭露系統描述文件，該文件應以簡明易懂之語言，描述系統之作用方式及其各組成部分之決策方式。
- (2) 課責性：告知應明確揭露負責設計或利用系統各組成部分之組織。
- (3) 及時更新：自動化系統啟用或對使用者造成影響前，應向使用者踐行告知。告知應及時更新。
- (4) 簡潔且明確：應對告知和解釋加以評估，確保使用自動化系統或受其影響的使用者易於取得該告知和解釋，能夠快速閱讀、理解，並據此採取行動。

2、解釋自動化系統如何以及為何作出某項決策或採取某項行

為

- (1) 合於目的：對自動化系統之解釋應合於使用者使用該解釋之目的。
- (2) 合於溝通對象：對自動化系統之解釋應符合溝通對象之特徵。
- (3) 合於風險等級：對自動化系統之解釋內容應符合該系統之風險等級，高風險或須嚴格監督之系統，應內建預先解釋風險之機制。
- (4) 有效：自動化系統提供之解釋應準確呈現影響決策之各項要素。

3、證明告知與解釋相關之保護措施

- (1) 報告：應針對前述各項要素提供摘要報告，內容包括負責組織、系統目標和使用案例、受影響之民眾、解釋有效性和可及性評估、風險評估等。

五、人為替代、評估及備案（fallback）

自動化系統應符合下列期待，以確保民眾有權選擇退出，且在遇到問題時，獲得及時人為評估與救濟：

- 1、提供便利機制，供民眾拒絕使用自動化系統，而以人為方案替代

- (1) 簡潔、明確且易於取得之告知與說明：應以簡明且易於取得之方式，向受自動化系統影響之個人告知其有權選擇退出。
- (2) 適時提供人為替代方案：若自動化系統係獲得權利、機會或近用管道之一部分，則應提供及時之人為替代方案。
- (3) 及時且無負擔之人為替代方案：退出程序及人為替代方案皆應及時，且不得造成過重負擔。

2、針對自動化系統失效、出現故障或使用者異議時，透過備案和申訴機制，提供及時人為考量和救濟

- (1) 符合比例：人為評估和備案機制之可用性，應與該自動化系統對權利、機會或近用管道之影響成比例。
- (2) 可及性：檔明、紙本、電話等各類人為評估和備案機制，應易於取得和使用。
- (3) 便利性：與自動化系統相比，人為評估和備案機制不得造成不合理負擔。
- (4) 公平性：人為評估和備案機制之結果應與自動化系統相當，且應對弱勢族群提供公平近用管道。
- (5) 及時性：人為評估和備案機制應及時可用。對於有時效性之系統，其替代機制應即時可用，甚至在損害發生前即已

提供。

(6) 有效性：人為評估和備案機制應可有效推翻原始決策，並執行新的決策。

(7) 妥適維護：在自動化系統使用期間，人為評估和備案機制應獲妥適維護。

3、透過教育訓練、評估和監督機制，糾正自動化偏見，確保人為制度有效

(1) 教育訓練與評估：對自動化系統進行管理、互動或解釋其決策之人員，皆應獲得適當訓練。

(2) 監督：應透過治理架構，對人為系統之有效性及潛在偏見進行有效監督。

4、針對敏感領域強化人為監督和安全維護

(1) 限定資料及其推論之範圍：敏感領域之自動化系統應受人為監督，確保其所用各項資料或資料屬性皆適於特定使用案例。

(2) 合於特定情境：敏感領域之自動化系統應受人為監督，確保其符合特定使用案例及實際部署情境。

(3) 高風險決策須經人為評估：自動化系統非經人為評估，不得直接於高風險情境進行決策。

(4) 檢視系統之有意義的管道：自動化系統之設計者、開發者和部署者應適當豁免保密要求，以利其系統在敏感領域受到有意義之監督。

5、證明人為替代、評估及備案機制之可及性

(1) 報告：應對人為替代方案之可及性、及時性和有效性進行報告，並定時公開。報告應包括對人為替代方案及時性、所造成額外負擔之評估，以及人為替代方案使用狀況之統計資料分析。

第五節 日本：以人為本的人工智慧社會原則

2019 年 3 月，日本內閣府訂定「以人為本的人工智慧社會原則」(Social Principles of Human-centric AI)¹⁵，提出 AI 應遵守下列 7 項社會原則：

一、以人為本原則

AI 之使用不得侵害憲法與國際標準所保障之基本人權。AI 之開發、使用與部署應拓展人類之能力，使人類之不同成員得以追求其福祉。使用 AI 之社會宜採取措施促進 AI 之合理使用，防範人類過度依賴 AI 或濫用 AI 操縱他人意思決定。

- (1) AI 可拓展人類能力和創造性。AI 除代替人力勞動外，並可作為輔助人類之先進工具。
- (2) 人類須自主判斷和決定如何使用 AI。
- (3) AI 應具使用者友善特質，防範與「資訊弱勢」和「技術弱勢」族群間之數位鴻溝。

二、認知教育原則

AI 相關之決策者和管理者應準確理解 AI，以及 AI 在社會中合理

¹⁵ Japan Cabinet, Social Principles of Human-Centric AI, <https://www.cas.go.jp/jp/seisaku/jinkouchinou/pdf/humancentricai.pdf> (accessed 30 August 2022).

使用所涉之知識與倫理。AI 之使用者應對 AI 有一般認知，並就 AI 之使用獲得充分教育訓練。AI 之開發者除掌握 AI 技術面之知識外，並應理解 AI 在社會中的使用方式，以及掌握社會科學和倫理相關知識。

- (1) 為各年齡段之民眾提供廣泛的教育機會。
- (2) 構建提升民眾 AI、數理、資料科學基礎素養之教育系統。
- (3) 轉變現行單向、統一之教育方式，靈活運用 AI 改善教育方式，構建交互式教育環境。
- (4) 不以強制方式增加學校負擔，透過與私人企業和民眾之獨立合作實現教育環境轉變。

三、隱私保護原則

使用 AI 時，除依資料之重要性和敏感性處理個人資料外，並應更加謹慎，以免個人資料之揭露或使用損害個人權益：

- (1) AI 相關之個人資料利用（包括政府之利用），均不得侵害個人自由、尊嚴和平等。
- (2) 確保個人資料使用行為的正確性、正當性，使本人能夠實質性參與隱私管理，與並從 AI 使用中獲益。
- (3) 個人資料之利用和保護應相均衡。

四、確保安全原則

AI 之使用可能帶來新的安全風險，因此，社會應衡平 AI 相關之利益於損害，並努力提升社會整體安全性和永續性。

- (1) 促進 AI 相關研究和開發之廣度與深度，例如正確評估 AI 風險、降低風險之措施等。
- (2) 重視 AI 使用之永續性，不依賴單一或少數種類之 AI。

五、公平競爭原則

- (1) 特定國家或企業縱使集中掌握 AI 相關資源，亦不得利用其優勢地位不當蒐集資料，或從事侵害他國主權或不當市場競爭之行為。
- (2) AI 之使用不得導致財富和社會影響不公平地流向特定利害關係人。

六、公平、課責與透明原則

在已準備好使用 AI 之社會中（AI-Ready Society），須確保決策之公平性與透明性，結果之課責性，以及對技術之信任，以免民眾因其個人背景受歧視，或其人格尊嚴受侵害。

- (1) 於 AI 之設計理念中，不因種族、性別、國籍、年齡、政治立場、宗教等被不當地差別對待民眾。
- (2) 適當解釋 AI 之適用、AI 資料之蒐集與利用的、確保 AI 運作結果適當性等。

- (3) 透過開放式對話使民眾理解 AI 提案並作出判斷。
- (4) 採取措施確保對 AI 及其相關資料和演算法之信賴。

七、創新原則

- (1) 為實現持續創新，應促進 AI 發展之國際化、多元化及產官學合作。
- (2) 促進大專院校、研究機構與企業間之平等合作及人才流動。
- (3) 推廣 AI 品質與可靠性評估方法。除 AI 工程相關研究外，並推動 AI 倫理、經濟等學門之確立與發展。
- (4) 在確保隱私和安全之前提下，促進資料和資源共享。
- (5) 調整政府監理機制，消除對 AI 發展之阻礙，促進 AI 技術之社會應用。

第六節 新加坡：人工智慧治理模範框架

2019 年 1 月，新加坡資訊通信媒體發展局（Infocomm Media Development Authority, IMDA）和個人資料保護委員會（Personal Data Protection Commission, PDPC）聯合發布第一版人工智慧治理模範框架（Model Artificial Intelligence Governance Framework, first edition）。2020 年 1 月，新加坡 IMDA 和 PDPC 對該框架內容加以調整，形成第二版人工智慧治理模範框架¹⁶。

該模範框架包括四部分，分別為：

- (1) 內部治理架構與措施；
- (2) 決定 AI 輔助決策（AI-augmented decision-making）中的人為參與程度；
- (3) 作業管理；
- (4) 利害關係人互動與溝通。

一、內部治理架構與措施

組織之內部治理架構和措施有助於確保對 AI 使用之有效監督。該模範框架以「內部治理架構與措施」部分說明組織如何調整現有治理架構或措施、增設新的治理架構或措施，以利納入演算法決策相關

¹⁶ Singapore IMDA & PDPC, Model Artificial Intelligence Governance Framework, Second Edition (2020), <https://www.pdpc.gov.sg/-/media/files/pdpc/pdf-files/resource-for-organisation/ai/sgmodelaigovframework2.ashx> (accessed 2 December 2022).

之價值、風險和責任。

組織宜考量執行下列內部治理架構之發展作為：

(一) 為 AI 部署明定角色和責任

就 AI 部署之各個階段、各項活動，組織宜制定適當人員和部門負責。AI 部署相關之關鍵角色和責任包括：

- 1、 利用既有風險管理框架，實施風險管控措施，以利：
 - (1) 評估和管理 AI 部署之相關風險，包括對個人之不利影響；
 - (2) 確定 AI 輔助決策中的人為參與程度；
 - (3) 管理 AI 模型訓練和選擇過程。
- 2、 維護、監測、記錄並檢視已部署之 AI 模型，必要時採取補救措施。
- 3、 評估與利害關係人之溝通管道和互動狀況。
- 4、 確保 AI 系統相關同仁接受適當教育訓練。

(二) 風險管理和內控

組織宜就 AI 部署相關風險執行嚴格風險管理和內控，相關措施例如：

- 1、 合理確保用以訓練 AI 之資料集符合預定目的之需要，並評估和管理準確性不足或偏見之風險。
- 2、 建立監測和通報程序，確保上層管理者瞭解 AI 部署之成效及相

關問題。

- 3、確保 AI 事務相關人員變動時的適當交接。
- 4、組織架構或關鍵人員發生重大變動時，檢視內部治理架構和措施。
- 5、定期檢視內部管理架構和措施之相關性和有效性。

二、決定 AI 輔助決策中的人為參與程度

該模範框架於「決定 AI 輔助決策中的人為參與程度」部分，說明組織如何確定 AI 應用之可接受風險，並確定 AI 決策中適當之人為參與程度。

於 AI 部署前，組織應確定其 AI 應用之商業目的，並將該目的與 AI 決策之風險相衡量。

AI 決策過程中人為監督之程度可分為三個層級：

（一） 人為控制（human-in-the-loop）

人類充分掌控決策過程，AI 僅提供建議。若無人類積極行為（例如給出指示），則無法做成決策。

（二） 無人控制（human-out-of-the-loop）

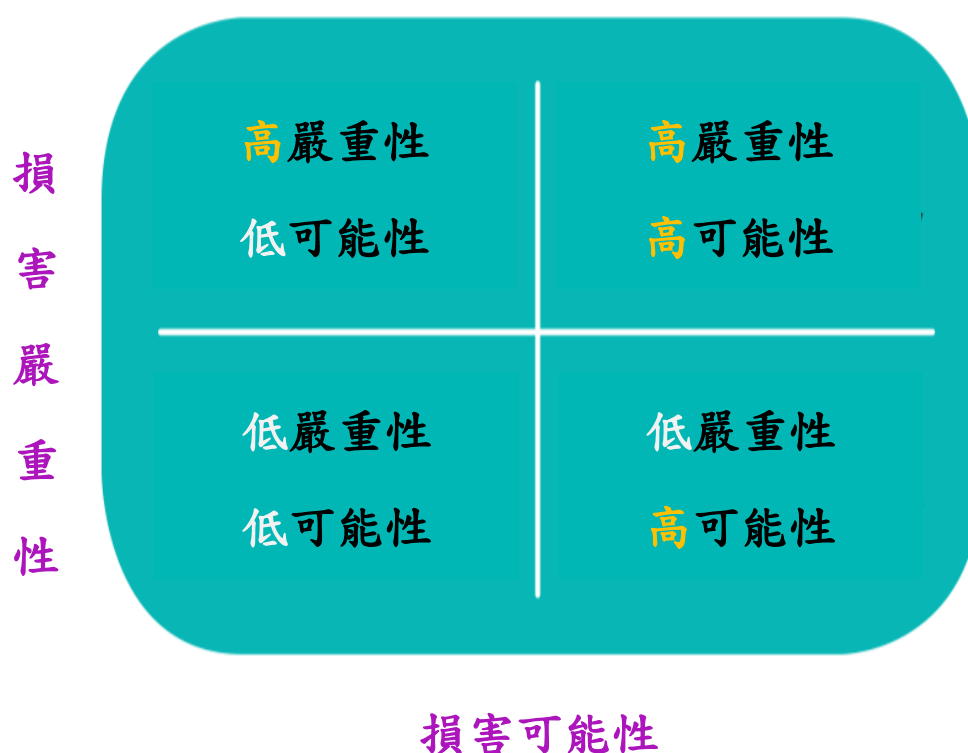
AI 系統完全控制決策過程，人類無法否決 AI 決策。

（三）人為干預（human-over-the-loop, human-on-the-loop）

人類處於監督角色，有權在特定情形下介入和推翻 AI 決策。

該模範框架並提出 AI 決策對個人之損害矩陣模型，以協助組織確定人為參與之適當程度，「損害」之內涵、可能性及嚴重性皆應依具體個案確定。此外，該指引明確指出，雖然損害可能性及損害嚴重性雖屬影響人為參與程度之重要影響因素，然在確定適當人為參與程度時，不應僅考量此兩項因素，而應依個案情境，綜合考量損害之性質、可逆性（reversibility）、人為介入之可行性等。

圖 2 AI 決策對個人之損害矩陣模型¹⁷（新加坡 IMDA & PDPC）



¹⁷ 參考來源：Singapore IMDA & PDPC, Model Artificial Intelligence Governance Framework, second edition (2020), 31.

三、作業管理

組織宜採風險基礎之方式管理 AI 相關作業。AI 模型之開發、選擇和維護（包括資料管理）將涉及作業管理議題。AI 開發與部署之流程可分為三個階段：

圖 3 AI 開發與部署之三階段及作業流程¹⁸（新加坡 IMDA & PDPC）

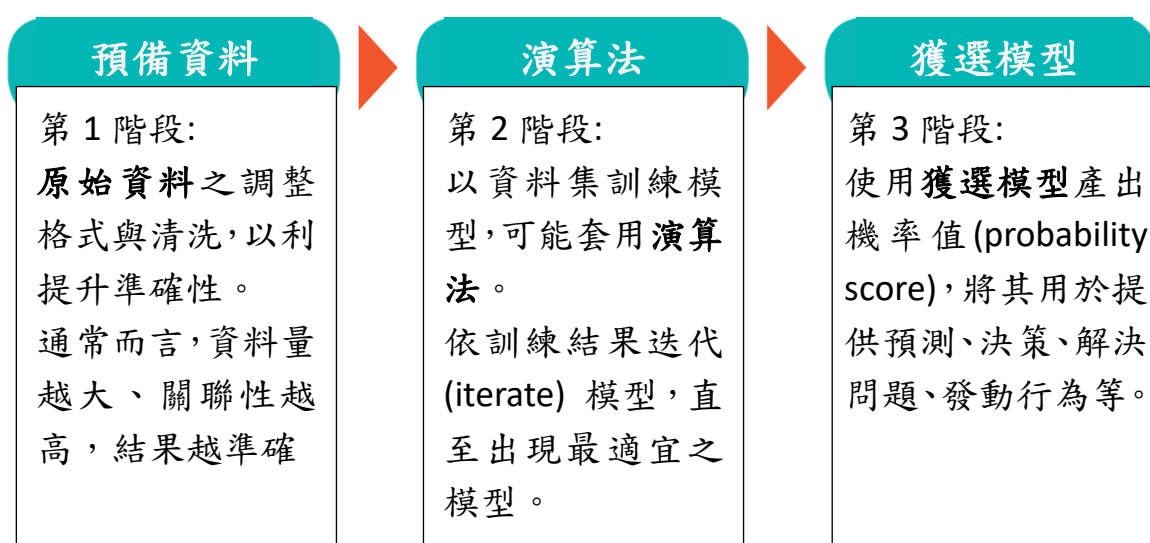
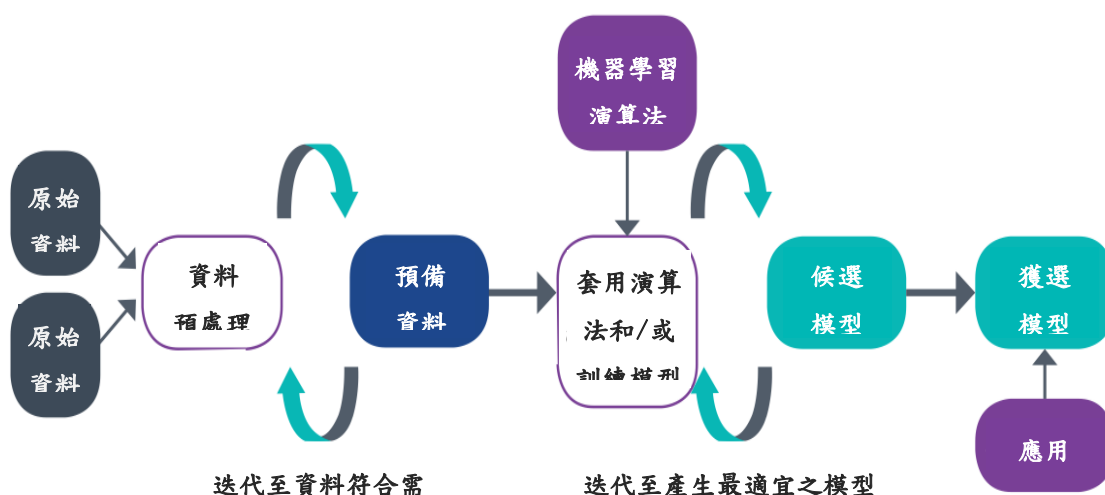


圖 4 AI 開發與部署之作業流程¹⁹（新加坡 IMDA & PDPC）



¹⁸ 參考來源：Singapore IMDA & PDPC, Model Artificial Intelligence Governance Framework, second edition (2020), 35.

¹⁹ 參考來源：Singapore IMDA & PDPC, Model Artificial Intelligence Governance Framework, second edition (2020), 35.

（一） AI 開發之資料

為確保 AI 應用之有效性，組織內相關部門應合作落實良好資料管理，例如：

- 1、理解資料之歷程（lineage），包括資料之來源、蒐集方式、組織內處理和流動方式、正確性維護方式等。
- 2、確保資料品質，包括資料集之準確性（accuracy）、全面性（completeness）、真實性（veracity）、新舊程度、相關性（relevance）、完整性（integrity）、可用性（usability）、人為干預程度（例如人為篩選、貼標、編輯等）。
- 3、降低固有偏見（bias），包括降低篩選偏見（selection bias）和評量偏見（measurement bias）。
- 4、於訓練、測試和驗證階段，分別使用不同資料集。
- 5、定期檢視和更新資料集。

（二） 演算法和模型

AI 系統可能透過 AI 模型中的演算法提供眾多功能。就下列面向採取措施，可提升 AI 模型中演算法之透明性：

- 1、透過解釋 AI 模型中演算法之作用原理、決策過程中如何使用對模型產出結果等，實現可解釋性（explainability）。

- 2、若難以實現可解釋性，則透過記錄 AI 模型產出結果展現可重複性（repeatability）。
- 3、提升 AI 系統在無效輸入或壓力環境條件下正常運作之能力，實現穩健性（robustness）。
- 4、建立內部制度，定期對 AI 模型進行調適，確保已部署之模型符合消費者行為之演變。
- 5、落實可追蹤性（traceability），提升透明性並協助瞭解模型運作狀況。
- 6、透過外部獨立驗證，落實可再現性（reproducibility）。
- 7、接受內部或外部人員稽核，落實可稽核性（auditability）。

四、利害關係人互動與溝通

組織可採取下列措施，協助取得利害關係人對 AI 部署之信任：

- （一） AI 之整體說明：就 AI 之使用提供整體說明，包括 AI 是什麼、AI 如何用於消費者相關之決策、使用 AI 之優勢、決定使用 AI 之原因、風險因應措施、AI 在決策中的作用等。
- （二） AI 之解釋政策：訂定 AI 解釋政策，規定向個人解釋 AI 之內容和時機，以利確保溝通內容之一致性。
- （三） 依溝通對象、目的和背景，提供有意義的資訊。
- （四） 若消費者與 AI 互動，依消費者之資訊需求，確定資訊提供之適當方式（如下圖所示）。

(五) 評估提供拒絕使用 AI 產品或服務之選項。

(六) 向消費者提供下列兩種溝通管道：

1、供消費者提供意見或提出詢問之回饋管道 (feedback channel)；

2、供消費者檢視影響其權益之 AI 決策之決策審核管道 (decision review channels)。

圖 5 消費者與 AI 互動之資訊提供²⁰ (新加坡 IMDA & PDPC)



²⁰ 參考來源：Singapore IMDA & PDPC, Model Artificial Intelligence Governance Framework, second edition (2020), 56.

第七節 香港－開發及使用人工智能道德標準指引

香港個人資料私隱專員公署（Office of the Privacy Commissioner for Personal Data, PCPD）於2021年8月發布「開發及使用人工智能道德標準指引」，協助機構在開發及使用人工智慧²¹時能有效遵守香港個人資料私隱條例之規範。

香港PCPD認為，人工智慧透過節省人力、改善運作效率、優化資源分配、提供個人化服務以及引發新構想，為不同行業帶來龐大機遇和利益，惟因人工智慧的潛能乃透過大數據實現，過程常涉及個人資料之使用，尤其新一代的人工智慧之「智能」是透過利用複雜的機器學習演算法分析大量訓練數據而獲得，形同持續測試個人資料保護法規的各項規範，加以個人資料經由人工智慧系統分析後，當事人之權利、自由或利益可能因此受到影響（例如由人工智慧作成自動化決策），其間亦涉及道德倫理之風險。

據此，香港PCPD從「制定人工智慧策略及治理」、「進行風險評估及人為監督」、「實行人工智慧模型的開發與管理及人工智慧系統的管理」、「促進與利害關係人的溝通及交流」等四面向提出管理循環模型，引此建構符合道德標準的人工智慧開發及使用工作流程。

²¹ 考量香港與臺灣用語差異，本研究將PCPD該指引內容調整為臺灣習慣用語。例如，「人工智能」酌調為「人工智慧」。

一、人工智慧策略及治理

(一) 人工智慧策略

香港 PCPD 認為²²，機構應制定人工智慧策略，以展示高級管理層通過合乎道德標準的方式開發及使用人工智慧的決心。該策略可包含下列要素：

- (1) 訂明使用人工智慧系統的業務目標，例如計劃利用人工智慧系統幫助解決哪些問題；
- (2) 界定人工智慧系統在機構的科技生態系統中提供的功能；
- (3) 參考上述介紹的人工智慧的道德原則，制定適用於機構在開發及使用人工智慧方面的道德原則；
- (4) 列明人工智慧系統的可接受用途，以及哪些用途是不容許的。機構可就人工智慧的應用採取交通燈機制²³；
- (5) 確保人工智慧系統的使用符合機構的願景、使命及價值；
- (6) 就如何合乎道德標準地設計、開發及使用人工智慧制定具體的內部政策和程序，包括制度化的決策過程和上報準則；
- (7) 定期把人工智慧策略、政策和程序傳達予所有相關人士，包括內

²² 香港，個人資料私隱專員公署，*開發及使用人工智慧道德標準指引*，2021 年，8 月，頁 12。

²³ 在交通燈機制下，人工智慧的實際應用會分為紅、黃、綠三個類別。紅色類別是被禁止採用的人工智慧，因其所涉及的風險太高。綠色類別是低風險的人工智慧，這類別的人工智慧在採用前無須進行嚴格的風險評估。黃色類別包含沒有在紅色或綠色類別出現的所有其他情況。要決定黃色類別內的人工智慧應否獲准許使用，應先進行嚴格的風險評估。轉引自香港，個人資料私隱專員公署，*開發及使用人工智慧道德標準指引*，註 12，2021 年，8 月，頁 13。

部各級職員和外部利害關係人（如適當），例如業務夥伴。

（二）治理架構

香港 PCPD 認為²⁴，開發及使用人工智慧需要不同領域的專業知識，例如電腦工程、數據科學、網路安全、用戶體驗設計、法律與合規、公共關係等，因此機構應建立具足夠資源、專業知識和決策權的內部治理架構，可包含下列要素：

- (1) 人工智慧治理委員會（或類似組織）²⁵監督人工智慧由開發、使用以至終止的整個生命週期；
- (2) 就人工智慧的開發及使用為不同部門或人員訂明清晰的角色及責任；
- (3) 在財政和人力上有足夠的資源開發及使用人工智慧；
- (4) 就開發及使用人工智慧建立有效的內部通報機制，例如通報系統故障或提出有關資料保障或道德標準上的關注，以便人工智慧管治委員會能恰當地作出監察。

²⁴ 香港，個人資料私隱專員公署，*開發及使用人工智慧道德標準指引*，2021 年，8 月，頁 13。

²⁵ 香港 PDPC 建議，高級管理層參與以及跨專業領域合作應是人工智慧治理委員會最重要的特質。機構應成立包含不同技能和觀點的跨部門團隊，包括業務運作人員、系統分析師、系統架構師、數據科學家、網路安全專家、法律及合規專業人員、人力資源人員、客戶服務人員等。機構應指派高級管理人員（例如行政總裁、資訊總監、私隱總監或類似職位）領導該跨部門團隊。另作為選擇性措施，人工智慧治理委員會可向外部專家尋求人工智能及道德標準方面的獨立意見。如果某項目規模龐大、影響廣泛、備受注目，其道德價值有可能受到挑戰，機構亦可成立另一個人工智慧道德委員會以進行獨立檢視。香港，個人資料私隱專員公署，*開發及使用人工智慧道德標準指引*，2021 年，8 月，頁 14。

（三）培訓及加強認識

香港 PCPD 指出²⁶，機構為妥善執行人工智慧策略與治理，應為相關人員提供足夠的培訓，確保具備適當之知識與技能，例如：

- (1) 為系統分析師、系統架構師及數據科學家提供有關遵從法律、規例、內部政策，以及網路安全風險方面的培訓；
- (2) 為法律及合規專業人員和人工智慧使用者（包括業務運作人員）提供人工智慧科技的培訓；
- (3) 為負責監督人工智慧系統決策的審查人員提供培訓，使其能在人工智慧系統所作的決定中查找並糾正不義的偏見、非法的歧視和錯誤。

二、風險評估及人為監督

（一）風險因素

從人工智慧的個人資料保護與道德層面而言，香港 PCPD 建議的風險評估因素包含²⁷：

- (1) 用來訓練人工智慧模型的資料或輸入人工智慧系統用作決策的資料的准許用途。
- (2) 用於訓練人工智慧模型或人工智慧系統運作的資料（尤其是個人

²⁶ 香港，個人資料私隱專員公署，*開發及使用人工智慧道德標準指引*，2021 年，8 月，頁 15。

²⁷ 香港，個人資料私隱專員公署，*開發及使用人工智慧道德標準指引*，2021 年，8 月，頁 18。

資料)的數量。

- (3) 所涉資料的敏感程度。一般被視為較敏感的資料包括生物辨識資料、健康資料，以及弱勢群體（例如兒童）的個人資料。
- (4) 所涉資料的質素（品質），當中須考慮其來源、可靠性、真實性、準確性、一致性、完整性、相關性及可用性。
- (5) 在開發或使用人工智慧時的個人資料安全，當中應考慮到個人資料如何在機構的科技生態系統以及人工智慧系統中轉移。
- (6) 隱私風險（例如過度蒐集、濫用或外洩個人資料）出現的可能性及其潛在傷害的嚴重程度。
- (7) 人工智慧系統對受影響個人及社會大眾的潛在影響（包括益處和傷害）。
- (8) 人工智慧系統的影響出現之可能性，以及其嚴重程度和持續時間。
- (9) 降低風險的緩減措施之足夠程度。

（二）決定人為監督的程度

香港 PCPD 認為²⁸，機構以「風險為本」²⁹的方式（risk-based approach）開發及使用人工智慧，只有在剩餘風險降至可接受的水平，

²⁸ 香港，個人資料私隱專員公署，*開發及使用人工智慧道德標準指引*，2021 年，8 月，頁 20。

²⁹ 香港 PCPD 開發及使用人工智慧道德標準指引原文如此。

且與人工智慧帶來的益處相稱時，才可接受開發及使用人工智慧。其中，人為監督即是減低風險的主要措施：

(1) 高風險的人工智慧系統應以「人在環中」³⁰(human-in-the-loop)

方式進行人為監督。人類行動者在決策過程中保留控制權，以防止人工智慧出錯或作出不當決定。

(2) 沒有真正風險或低風險的人工智慧系統可採取「人在環外」³¹

(human-out-of-the-loop) 方式進行人為監督。人工智慧系統可以在沒有人為介入下作出決定，以達致完全自動化決策。

(3) 如「人在環中」及「人在環外」兩種方式均不適合，例如當風險

程度不可忽視，而「人在環中」的方式又不能符合成本效益或不可行，機構可考慮「人為管控」(human-in-command) 方式。人類行動者監督人工智慧系統的運作，在有需要時才介入。

三、人工智慧模型的開發及人工智慧系統的管理³²

(一) 數據（資料）準備

人工智慧需要資料以找出規律、作出推斷、建議或決策，此階段應確保資料合法（個人資料保護），並具備良好品質，以確保人工智慧系統的公平性。

³⁰ 香港 PCPD 開發及使用人工智慧道德標準指引原文如此。

³¹ 香港 PCPD 開發及使用人工智慧道德標準指引原文如此。

³² 香港，個人資料私隱專員公署，*開發及使用人工智慧道德標準指引*，2021 年，8 月，頁 22。

（二） 模型開發

備妥資料後，機構可利用機器學習演算法分析訓練資料，以開發人工智慧模型。香港 PCPD 建議機構應了解不同機器學習演算法的特點，從而選擇符合需要之演算法。其中應考慮者包含人工智慧分析結果所欲達成之準確程度與可解釋程度。

（三） 管理與監察

香港 PCPD 強調，由於人工智慧系統的風險因素會隨時間改變，影響人工智慧系統的可靠性、穩健性與安全性，因此機構應依風險等級高低程度，持續監察人工智慧系統。

四、與利害關係人的溝通及交流

香港 PCPD 認為，機構應該讓利害關係人知悉人工智慧的使用，並展示機構的人工智慧管理決心，有效地與利害關係人溝通交流以獲得信任。

五、自我核對清單

在該指引附錄 A，香港 PCPD 將指引內建議的人工智慧開發及使用管理循環措施列為自我核對清單供機構參考，本研究亦整理提出如下：

表 2 人工智慧開發使用自我評估核對清單（香港 PCPD）

人工智慧策略及治理			
	問題	答案 (有/沒有)	所需的進一步行動
1	貴機構在開發及使用人工智慧前有沒有制定人工智慧策略？		
2	貴機構有沒有就合乎道德標準地設計、開發及使用人工智慧制定具體的內部政策和程序？		
3	貴機構有沒有成立人工智慧治理委員會（或類似組織）以監督人工智慧系統由開發、使用，以至終止的整個生命週期？		
4	人工智慧治理委員會（或類似組織）有沒有： - 來自不同專業領域及部門的成員就人工智慧的開發及使用互相合作？ - 高級管理人員監督其運作？		
5	貴機構有沒有為開發及使用人工智慧的人員訂明清晰的角色及責任？		
6	貴機構有沒有在財政和人力上提供足夠的資源以開發及使用人工智慧？		
7	貴機構有沒有為開發及使用人工智慧的人員提供與其職位相關的培訓？		
8	貴機構有沒有定期為所有相關人員安排活動，加強他們對使用人工智慧的認識？		
風險評估及人為監督			

	問題	答案 (有/沒有)	所需的進一步行動
1	貴機構有沒有在開發及使用人工智慧之前進行風險評估？		
2	貴機構的風險評估有沒有考慮個人資料私隱的風險及人工智慧系統的道德方面的影響？		
3	風險評估結果有沒有由人工智慧治理委員會（或類似組織）檢視及認可？		
4	貴機構有沒有在考慮過人工智慧系統的風險概況後，對人工智慧系統訂立適當程度的人為監督及其他緩減風險的措施？		
人工智能模型的開發及人工智能系統的管理			
	問題	答案 (有/沒有)	所需的進一步行動
準備數據			
1	貴機構有沒有採取步驟把使用的個人資料減至最少及確保遵從《私隱條例》的規定（例如使用匿名化或合成數據；了解個人資料的來源及准許的用途；檢查個人資料的準確性等）？		
2	貴機構在使用數據前，有沒有採取步驟確保數據的可靠性、真實性、準確性、一致性、完整性、相關性、公平性及可用性？		
人工智能模型的開發			
3	貴機構在使用機器學習演算法之前，有沒有評估其特點？		
4	貴機構有沒有對人工智能模型		

	進行嚴格測試，確保模型可靠、穩健及公平？		
5	貴機構有沒有制定足夠的風險緩減措施，包括人為監督，以應付使用人工智慧系統時可能出現的錯誤或故障？		
6	貴機構有沒有制定足夠的保安措施，保障人工智慧系統免受攻擊？		
7	貴機構有沒有制定應變計畫，在有需要時暫停人工智慧系統及啟動後備解決方案？		
管理與監察			
8	貴機構有沒有對人工智慧系統的數據處理、風險評估、設計、開發、測試及使用妥善地記錄存檔？		
9	貴機構有沒有制定計劃當人工智慧系統在功能或運作上有重大改變，或規管或科技環境出現重大變化時，重新評估人工智慧的風險？		
10	貴機構有沒有定期檢視、調校及再訓練人工智慧模型？		
11	貴機構有沒有根據評估的風險程度對人工智慧系統訂立適當程度的人為監督？		
12	貴機構有沒有為人工智慧系統的使用者提供運作上的支援及反饋途徑？		
13	貴機構有沒有在人工智慧系統由開發、使用、監察以至終止的整個生命週期實施適當的保安		

	措施？		
14	貴機構有沒有計劃定期評估宏觀的科技環境，以識別與機構現時的科技生態系統的差距？		
15	貴機構有沒有定期進行內部審核，確保人工智慧的開發及使用遵從內部政策的規定？		
與利害關係人的溝通及交流			
	問題	答案 (有/沒有)	所需的進一步行動
1	貴機構有沒有清楚明確地向個別的消費者披露機構使用人工智慧系統？		
2	貴機構有沒有告知個別的消費者在產品或服務中使用人工智慧系統的目的、益處及效果？		
3	貴機構有沒有披露人工智慧系統的風險評估結果(如適合)？		
4	貴機構有沒有向個人提供退出使用人工智慧的途徑(如可能)？		
5	貴機構有沒有向個人提供途徑讓他們改正不準確之處、作出反饋、尋求解釋以及要求人為介入(如可能)？		
6	與利害關係人的通訊有沒有以淺白語言、清楚易明的方式書寫？		

第八節 法規及指引文獻比較分析

本研究共研析七部法規、指引文獻，依功能定位及內容重點，各研析文獻比較如下：

表 3 AI 法規及指引比較表

功能/內容 法規/指引	綜合性 AI 專法	原則性 AI 倫理指引	操作性 AI 倫理指引	AI 個資保護 指引
歐盟可信賴人工智慧倫理指引		√		
歐盟人工智慧法 (草案)	√			
英國 AI 與資料保護 指引				√
美國人工智慧權利 法案藍圖		√		
日本以人為本的人 工智慧社會原則		√		
新加坡人工智慧治 理模範框架			√	
香港開發及使用人 工智能道德標準指 引			√	

由上述法規與指引文獻可知，現行國際實務尚未正式通過施行人工智慧法律，各國人工智慧而各國的人工智慧指引大抵由「資料治理」角度強調「課責性」，建議機關從制度面自上而下推動人工智慧的開發及使用政策，並於過程中確保遵循個人資料保護法律，適當遵守「透

明性」原則以顧及當事人權利，且應隨時留意人工智慧應用過程的倫理（公平、準確、不歧視等）與安全。綜合比較前述法規與指引，可得出以下觀察：

1. 人工智慧之治理策略與基本原則

風險基礎之方式（risk-based approach）可謂人工智慧治理之通用方案，歐盟人工智慧法草案、英國 AI 與資料保護指引、新加坡人工智慧治理模範框架皆明文採用風險基礎之方式。其中，歐盟人工智慧法草案將風險基礎之思維引入立法，區別風險層級高低而有不同之監理機制與合規義務，應值借鏡。

人工智慧治理之基本原則方面，前述各法規、指引呈現出高度一致性。人工智慧應尊重人類自主，並遵守合法性、公平性、安全性、穩健性、透明性、可解釋性、課責性、尊重隱私及個資保護等基本原則。

2. 人工智慧之管理制度與組織

本研究所研析的兩部操作性 AI 倫理指引，即新加坡人工智慧治理模範框架和香港開發及使用人工智能道德標準指引，皆注重人工智慧開發及應用中之管理制度與管理組織。其中新加坡指引從作業角度，強調 AI 開發和應用過程中，於組織內部明定相關部門之角色和責任，並建立風險管理和內控機制。香港 PCPD 則提出更制度化之建議，認

為組織宜就人工智慧之開發及應用制定整體策略，明確人工智慧之倫理原則、可接受使用方式等，並於組織面設立人工智慧治理委員會等組織，負責監督人工智能由開發、應用以至終止使用之完整生命週期。香港 PCPD 之建議，有助於組織構建系統化之人工智慧治理制度，從而確保組織開發或應用之各項水平，皆合於較高倫理標準。

3. 人工智慧系統之風險評估與人為監督

風險基礎之人工智慧治理模式下，人工智慧之治理措施應與其風險水平相符，因此，事前風險評估乃人工智慧開發及應用管理之重要舉措。歐盟可信賴人工智慧倫理指引、歐盟 AI 法草案、美國人工智慧權利法案藍圖、新加坡人工智慧治理模範框架和香港開發及使用人工智能道德標準指引皆不同程度地就風險評估提供規範或指引。因歐盟（及英國）個人資料保護法制中，已對個資保護影響評估（DPIA）制度作出明確規範，英國 ICO 於其指引中，對人工智慧系統是否及如何執行 DPIA 作出較詳細說明。

人為監督係防範人工智慧決策對人類權益不當侵害之關鍵之一。歐盟可信賴人工智慧倫理指引、歐盟 AI 法草案、美國人工智慧權利法案藍圖、新加坡人工智慧治理模範框架、香港開發及使用人工智能道德標準指引及英國 AI 與資料保護指引皆明文要求人工智慧系統應有適當之人為監督機制。整體而言，人為監督之方式應與人工智慧系

統之風險相符，人工智慧系統對人類之風險越高，人為監督之密度亦應隨之升高。惟具體個案中人工智慧應用之風險水平及適當之人為監督方式，應依該應用之功能、運作方式與環境、所涉損害之可能性、嚴重性、性質、可逆性等因素綜合判斷。因此，除歐盟 AI 法草案針對「遠程生物特徵辨識」要求採「人為控制」(human-in-the-loop)方式落實人為監督外，前述七部法規或指引中，未見其他針對「具體人工智慧應用」明定人為監督方法之規範。相較之下，各國的人工智慧指引（及歐盟、英國個人資料保護法制），多係基於人工智慧應用之「風險水平」區分不同人為監督密度與方法，而具體人工智慧應用之風險水平，則由組織依個案狀況，綜合考量多種因素加以判定。

4. 人工智慧系統之透明性與利害關係人溝通

人工智慧系統之透明性、演算法可解釋性乃實現「可信賴人工智慧」之重要基礎，因此，本研究所研析之七部法規或指引皆強調透明性原則。其中，新加坡人工智慧治理模範框架和香港開發及使用人工智能道德標準指引皆建議組織向當事人提供人工智慧之一般性資訊（例如人工智慧之功能、特性、優勢、應用方式等），並對人工智慧之決策提供解釋。英國 AI 與資料保護指引則從個人資料保護法制之角度，要求人工智慧系統對個人資料之蒐集、處理及利用應具透明性，並提供相關實務建議。

第四章 蒐集之案例彙整

北市府 1999 熱線目前提供「市政諮詢」、「轉接來電」、「告知」、「接受陳情」、「受理派工」五大類服務，係民眾取得北市府市政服務之單一窗口之一。本研究以 1999 熱線於市政服務中之功能定位為基礎，蒐集彙整國內外政府及民間數位轉型及導入 AI 之案例。

第一節 官方專線服務導入 AI 案例

一、比利時（佛拉蒙地區）Flemish Infoline 1700

比利時佛拉蒙區的 Flemish Infoline 1700 以自然語言處理技術自動分類民眾電話進線問題，整合性編譯、記錄該次對話，並提供建議答案，以此 AI 技術優化進線服務³³。

具體而言，該計畫旨在更快速的識別並分類進線問題，並將歷次回覆再次收斂整合，以供未來無論是以口頭或書面方式，更迅速、正確、自動的提供回覆人員建議答案。

為此，該計畫廣泛應用自然語言處理、機器學習、深度學習及語言技術（語音轉文字 Speech-to-Text, STT 及文字轉語音 Text-to-Speech, TTS）等 AI 模型。

³³ 1700 supported by Artificial Intelligence, <https://www.innovatieveoverheidsopdrachten.be/en/projects/1700-supported-artificial-intelligence> (accessed 2 December 2022).

二、愛爾蘭（稅務局）VDA（Virtual Digital Agent）

2018 年初，愛爾蘭稅務局（Revenue Commissioners）開始研究以人工智慧提供使用電話通訊服務的居民更有效率的使用經驗。當局執行示範專案（Pilot Project），檢視以自然語言處理技術為基礎之人工智慧應用，是否能夠對居民提供更好的服務、降低成本並增加效率。當局進而使用設計於處理納稅人為完稅（tax clearance）事務而進線的電話之 VDA（語音機器人）。

該 VDA 採自然語言處理技術，將民眾電話進線之語音轉化為文字，對文字內容進行分析並形成回應文本，再將回應文本轉化為語言，於電話中告知進線民眾。示範專案執行期間，該 VDA 表現出高度自動化效能，近 50% 的進線電話完全由該 VDA 處理，僅有 10% 的進線因 AI 無法理解進線需求而轉由人工處理³⁴。

該示範專案結束後，愛爾蘭稅務局並未立即將其推廣應用，而是繼續研擬適於大規模應用之安全身分驗證模型³⁵。但該專案表明，人工智慧影響可有效提升服務品質並降低服務成本。

³⁴ Revenue Chatbot, Our Public Service, <https://www.ops.gov.ie/actions/innovating-for-our-future/innovation/emerging-technology/artificial-intelligence/revenue-chatbot/> (accessed 2 December 2022).

³⁵ Ibid.

三、西班牙（中央政府）060 專線

060 專線於 2006 年開始提供服務，作為西班牙國民針對各類國家服務的中央聯繫端點，為國民提供資訊並連結至不同的公家服務。

該服務使用自然語言處理技術，回答居民常見的問題，並將正式的官方用語轉化為提供給居民的非正式用語³⁶。

四、我國（勞動部）多元智能電服中心

勞動部於民國 91 年起設立電話服務中心，利用電腦與電話整合（Computer and Telephone Integration, CTI）之技術與功能，提供民眾諮詢服務。長久以來因民眾電話諮詢問題複雜，新進電服員學習使用原有工具書需花費相當多時間，為強化組織知識分享，累積智能客服 Q&A 知識庫，勞動部遂於 108 年起，逐步建置電服中心知識管理系統，為導入人工智慧應用建置智能客服預作準備³⁷。

至 109 年，勞動部考量因行動裝置普及，民眾期待政府諮詢平台能更便利、更接近生活；且新冠疫情（COVID-19）期間，電服中心每天平均新增 3 萬 3 千通民眾來電³⁸。傳統電話服務中心由真人提供服

³⁶ Salvador Estevan Martínez & Doaa Samy, Natural Language and AI Solutions for Citizen Services Evolution (2019), [https://joinup.ec.europa.eu/sites/default/files/inline-files/08-2%20Salvador Estevan Martinez%20and%20Doaa Samy 0.pdf](https://joinup.ec.europa.eu/sites/default/files/inline-files/08-2%20Salvador%20Estevan%20Martinez%20and%20Doaa%20Samy%200.pdf)(accessed 30 August 2022).

³⁷ 勞動部，108 年度「本部電話服務中心知識管理系統建置、導入、擴充暨維護服務案」，徵求廠商企劃書，頁 3。

³⁸ 丁玉珍、林子倫，人工智慧提升政府公共治理的應用潛力探討，檔案半年刊，109 年 12 月第 19 卷第 2 期，頁 33。

務，人員需要長時間訓練才能精準回答問題，難期在短時間內擴充人力，也無法提供 24 小時不間斷服務，導致部分未能於上班時間來電的民眾，無法得到即時的答覆。此外，原本以人工方式蒐集、彙整電話諮詢服務問答資料，電話服務員也需花費大量時間記憶資料夾路徑及相關勞動領域知識。

勞動部因此啟動建置跨裝置的諮詢服務平台及智能客服系統與真人文字客服系統，介接國人常用的溝通管道，使民眾可以透過問答方式進行諮詢獲得解答，提升自助服務的比率，由傳統電話客服一對一的服務模式，升級為一對多的互動服務模式。且智能客服系統介接知識管理系統，創造「人機協作」機制，讓智能客服回答大宗且簡易的問題，難以處理的問題再由人工接手處理³⁹。110 年，勞動部電話服務中心正式更名為「勞動部多元智能電服中心」

勞動部將電話服務中心逐步轉型為智能客服中心過程中，應用人工智慧之具體作法如下⁴⁰：

- 1、由勞動部同仁（領域知識專家）建立標準問答知識庫，利用資訊擷取（information extraction）、資訊分類（information categorization）與資訊檢索（information retrieval）技術，定

³⁹ 勞動部，109 年度「本部電話服務中心智能客服系統與智能分析平台建置、導入、擴充暨維護案」，徵求廠商企劃書，頁 3。

⁴⁰ 丁玉珍、林子倫，人工智慧提升政府公共治理的應用潛力探討，檔案半年刊，109 年 12 月第 19 卷第 2 期，頁 34。

時、自動、跨裝置的檢索勞動法規及最新發布或修正之法規命令等資料並自動文件分類，有效節省人工資料彙整及搜尋的時間。

2、搜尋知識庫知識點頻率作為推播（push notification）依據，減輕電話服務員工作負荷；推播也可以針對民眾關心議題、熱門議題或民眾曾經諮詢問題，主動提供資訊以提高公民服務品質。

3、Google 於 2018 年釋出自然語言處理的預訓練模型 BERT（Bidirectional Encoder Representations from Transformers）作為意圖辨識的基礎，並以勞動領域問答知識進行微調（fine-tune）訓練問答系統，透過遷移式學習（transfer learning）的方式來增加機器對於自然語言的理解能力，藉此提高智能客服系統「問題－答案」的正確率。

4、利用「人機協作」機制，提供智能文字客服與真人文字客服兩種服務方式，由機器人負責第一線回應，人工監督及修正回應內容，機器人亦可輔助人工查詢問題；文字客服可以自由切換成智能文字客服或真人文字客服，使線上的客服人員可以快速瞭解民眾問題及適時提供正確的答案，進一步使傳統一個電話服務員單次只能接聽一通電話諮詢服務的「一對一」的服務模式，轉變成可以提供「一對多」的諮詢服務模式，提高一般諮詢電話服務效率；當智能文字客服與真人文字客服均無法解決民眾諮詢服務的問題

題之後，最後才由真人電話諮詢提供服務，分散傳統電話服務中心真人電話諮詢總量，而且簡單問題由機器人回答，複雜問題才由真人回答，可提高電話諮詢服務品質，也可隨時因應新興議題諮詢服務大增的需求。未來透過機器學習自動記錄電話服務員的績效，針對電話諮詢服務的類別進行分類（classification）或分群（clustering），免除現行電話服務員需逐筆登錄的重複工作；並利用過往使用真人客服流量的資料，對不同日期和時間電話進線量進行統計分析，以找出最有效率的人員配置數量，再搭配居家值機、全時與部分工時的策略，提高績效評估與營運管理。

在上述目標之下，勞動部電話諮詢服務知識庫透過「人機協作」，快速增加「問題－答案」的資料量來提升智能客服系統回答問題的覆蓋率和正確率⁴¹。

此外，電話服務過程如能即時掌握民眾的情緒反應，有助於強化機關對於民眾滿意度的評量，並據以改善服務內容與流程。傳統語音客服所蒐集的資料較難進行分析與再利用，例如諮詢電話只能進行大分類（例如申請聘僱外勞之相關規定、外國專業人士、工時等），無法分析產生大量重複問題的原因、民眾的痛點等，智能文字客服及真

⁴¹ 丁玉珍、林子倫，人工智慧提升政府公共治理的應用潛力探討，檔案半年刊，109 年 12 月第 19 卷第 2 期，頁 34。

人文字客服可以將大量結構化資料（客服系統紀錄）與非結構化資料（民眾的詢問和對話文字）保留下來，再加上利用 STT（語音轉文字）技術蒐集語音客服的資料，利用自然語言處理進行選定目標、對話解析（dialogue parsing）、斷詞處理（word segmentation）、關鍵詞識別（keyword spotting）、詞彙關聯（word association）、主題模型（topic model）、語句情感分析（sentence-based sentiment classification）、字彙的情緒分析（lexicon-based sentiment analysis）、資料視覺化等步驟，從民眾詢問和對話資料的文字採礦（text mining）與意見採礦（opinion mining），可以提取民眾的主觀訊息或意見、辨識出正面與負面等情緒，瞭解民眾對於政策的意見⁴²。

輿情分析方面，智能分析平臺進一步串接內外部系統進行整合，建立自動化的資料處理流程，自動統計各諮詢平臺之業務類別及陳情內容，即時搜集資料後進行有效數據分析，掌握民眾熱門諮詢議題及民眾的痛點，提供政府決策參考及建立政策回饋機制，降低關鍵決策的偏差（bias），以優化決策品質⁴³。

⁴² 丁玉珍、林子倫，人工智慧提升政府公共治理的應用潛力探討，檔案半年刊，109 年 12 月第 19 卷第 2 期，頁 35。

⁴³ 丁玉珍、林子倫，人工智慧提升政府公共治理的應用潛力探討，檔案半年刊，109 年 12 月第 19 卷第 2 期，頁 35。

五、我國中國信託商業銀行 AI 語音客服⁴⁴

中國信託商業銀行（中信銀行）客服每月來電量約 100 萬通，為快速提供解答並優化客戶體驗，於 109 年底將「自然語言對話服務」導入客服中心語音系統，打造出全國金融業首創的 AI 語音客服。

中信銀行組成跨單位專案團隊，導入大數據及 AI 金融科技應用，由資深客服同仁擔任 AI 訓練師，分析上萬句對談內容，進行逾百次客戶訪談蒐集回饋，並擷取逾 300 小時的服務音檔，反覆訓練語意模型，以此打造高度擬真的 AI 語音客服。

六、我國長庚醫院 AI 智能客服⁴⁵

長庚醫院考量傳統民眾撥打電話語音掛號，須自行瀏覽門診表，通過多層選單才能完成，過程中可能因選項輸入錯誤或忘記代碼導致失敗，歷時 2 年開發，於 111 年推出智能客服系統，以擬人化對答方式，協助民眾文字或電話語音掛號、查詢門診，甚至主動傳送簡訊或撥電話給等待住院或預約手術的病人，確認報到及術前回覆事項。

該系統落地長庚醫院，以降低民眾個資外洩風險，讓民眾透過長庚醫院官網、App，或是撥打單一市話專線，在 AI 文字客服或語音

⁴⁴ 見自由時報，AI 語音取代真人客服 0800 用講的嘛ㄟ通，<https://market.ltn.com.tw/article/10582>，最後到訪為 111 年 12 月 10 日。

⁴⁵ 見 iThome，長庚醫院啟用 AI 智能客服，與民眾對話協助查詢門診、預約掛號、住院等資訊，<https://www.ithome.com.tw/news/154482>，最後到訪為 111 年 12 月 10 日。

客服的協助之下，查詢或預約門診掛號、病床，以及衛教諮詢。

第二節 語音辨識 AI 案例

一、愛沙尼亞（國家議會）Hans

Hans 為一使用自然語言處理技術之演說識別系統，可編譯愛沙尼亞議會中的發言⁴⁶。由於總理要求要在會議後1小時內發布逐字稿，需要投入大量人力完成，因此愛沙尼亞政府導入該系統並規劃以機器可讀的格式產出檔案，成為開放資料供其他系統使用。

二、德國（聯邦訊息技術中心）C-19

為因應新冠疫情爆發後的大量諮詢需求，德國聯邦訊息技術中心（Informationstechnikzentrum Bund, ITZBund）導入使用 C-19 聊天機器人⁴⁷，該 AI 主要用於語言識別，但也可處理對話。政府先將既有常見問答輸入成為對話，再由 C-19 提供預先擬定的制式回答，並可提供相關的其他來源資訊。

三、德國（聯邦移民及難民辦公室）難民語言偵測

德國聯邦移民及難民辦公室（Bundesamt für Migration und

⁴⁶ The Riigikogu has a new system for making verbatim reports, <https://www.riigikogu.ee/en/press-releases/others/the-riigikogu-has-a-new-system-for-making-verbatim-reports/> (accessed 2 December 2022).

⁴⁷ Chatbot C-19, <https://c19.bundesbots.de/> (accessed 2 December 2022).

Flüchtlinge, BAMF) 使用聲音辨識軟體識別沒有文件的移民來源國家⁴⁸，目標是要偵測自稱來自敘利亞的移民使用的阿拉伯語以提升獲得政治庇護的機會。

該自動辨識軟體將搭配其他已經使用的技術一起建立該移民的身分。在做出決定前，該移民之語言和其他文件都將經檢視。如有必要，官員也會徵詢專家的意見。

不過，曾有語言專家質疑自動辨識軟體的正確性，因為語言及用詞時常變動，可能降低效率與正確性。

四、我國法務部語音辨識系統⁴⁹

法務部於 110 年 2 月於各檢察機關啟用新的語音辨識系統，使用 AI 語音辨識技術，以大量會議錄音檔訓練，並搭配司法用語詞彙，讓系統能辨識專業的司法用語。

透過此語音辨識系統，不僅可快速產生一般會議的文字紀錄，甚至成為檢察機關偵辦案件的幫手，將蒐集各式影音證據，利用語音辨識快速產生逐字稿，讓檢察官能夠快速勘驗案件相關證據，不需再聽完完整的語音資料，快速找到可能的線索。

⁴⁸ 見 www.dw.com/en/automatic-speech-analysis-software-used-to-verify-refugees-dialects/a-37980819，最後到訪為 111 年 12 月 10 日。

⁴⁹ 見 iThome，用 AI 改善機關作業效率，法務部靠 IT 再造為轉型打底，<https://www.ithome.com.tw/people/147486>，最後到訪為 111 年 12 月 10 日。

第三節 AI 線上聊天機器人案例

一、英國（英國駕照及行照監理局）聊天機器人

英國駕照及行照監理局（Driver and Vehicle Licensing Agency, DVLA）保有英國 8,000 萬份駕照及行照資料，並設有英國政府內部最大的話務中心，工作人員總數達 1,200 人，每月電話進線量達 110 萬通、線上對話諮詢達 10 萬次。於 2018 年起，DLVA 與外部廠商合作開發線上聊天機器人，使用自然語言處理技術，透過選單（menu）與使用者互動，接收並回應使用者之詢問⁵⁰。若使用者要求諮詢人工服務人員，則聊天機器人會將對話轉由人工服務。

該聊天機器人開發期間，為因應 DLVA 話務中心內部人員關於安全或服務品質之顧慮，開發團隊與法律及安全專家商討風險因應措施，取得資料安全部門主管之核可，自開發初期即向利害關係人演示聊天機器人作業模型，並採用新功能逐步上線策略，以管理使用量突然大增之狀況。該聊天機器人上線後，開發團隊繼續監測其表現，並評估其改善方式，包括分析使用者之回饋，並據此調整聊天機器人之風格和內容。

該聊天機器人顯著降低了 DLVA 話務中心工作人員之負擔，其可

⁵⁰ UK Government Digital Service, How the DVLA deployed a chatbot and improved customer engagement, <https://www.gov.uk/government/case-studies/how-the-dvla-deployed-a-chatbot-and-improved-customer-engagement> (accessed 2 December 2022).

獨立因應 25%-30%之民眾諮詢，並將各諮詢之平均處理期間從 8 分鐘縮短至 2.5 分鐘。2020 年，該專案被英國中央數位和數據辦公室（Central Digital and Data Office, CCDO）列入「聊天機器人和線上對話工具應用指引」（Using chatbots and webchat tools）之示範應用之一⁵¹。

二、奧地利（聯邦數位及經濟事務部）Mona

奧地利聯邦數位及經濟事務部（Bundesministerium für Digitalisierung und Wirtschaftsstandort, BMDW）的 Mona 聊天機器人使用自然語言處理技術，回答企業關於新冠疫情下與經濟相關的所有資訊，例如短期勞工、遠距工作等勞工法律事宜⁵²。該服務於政府網站及行動應用程式提供。

三、丹麥（腓特烈堡自治市）Kiri

Kiri 聊天機器人使用自然語言處理技術，作為數位助理角色於網站上即時回答市民對於當地政府服務的問題⁵³。目前 Kiri 僅能回答例如護照、駕照服務等領域問題，還無法處理個案提問。

⁵¹ UK Central Digital and Data Office, Guidance: Using chatbots and webchat tools, <https://www.gov.uk/guidance/using-chatbots-and-webchat-tools> (accessed 2 December 2022).

⁵² Corona-Krise - Schramböck: Chatbot „Mona“ als zentraler Infopoint für Unternehmen am USP, <https://www.bmaw.gv.at/Presse/Archiv/Pressemeldungen-BMDW/2020/April-2020/Corona-Krise-Chatbot-USP-.html> (accessed 2 December 2022).

⁵³ Roskilde kommune consistently resolves 92% of citizen queries with conversational AI, <https://www.boost.ai/case-studies/kommune-kiri-roskilde-kommune> (accessed 2 December 2022).

四、愛沙尼亞（統計局）Iti

愛沙尼亞統計局（Statistikaamet）Iti 線上聊天機器人使用自然語言處理技術，作為虛擬助理角色掌握愛沙尼亞人口、薪資級距、平均薪資總額、消費者物價變動等資料⁵⁴。如果 Iti 沒有答案，便將需求依照內容重新導回該資料提供者的客戶服務專人，或該統計資訊的使用者。Iti 的知識來自於經年累月的線上聊天提問輸入，每次提問都讓 Iti 變得更有「智慧」。

五、芬蘭（奧盧市）OuloBot

OuloBot 聊天機器人使用自然語言處理技術，奧盧市計畫將 OuloBot 打造為全年無休的虛擬助理，提供日常生活的所有事務資訊，包含旅遊、企業與市政服務。如果無法回答使用者的問題，在上班時段將會由奧盧市服務人員處理⁵⁵。

六、羅馬尼亞（克盧日－納波卡市）Antonia

2018 年，克盧日－納波卡（Cluj-Napoca）市政廳因應針對公共服務所需的 64 種表格，使用名為 Antonia 的人工智慧系統。該系統包含聊天機器人，可回答關於如何填寫表格的問題。在該技術的第二

⁵⁴ Statistics Estonia's virtual assistant Iti helps both users of statistics and data providers, <https://www.stat.ee/en/uudised/news-release-2019-042> (accessed 2 December 2022).

⁵⁵ Virtual Assistant OuloBot, <https://www.ouka.fi/oulu/oulubot/english> (accessed 2 December 2022).

階段，政府規劃 Antonia 將製作為全像投影⁵⁶。

七、新加坡（微軟）對話式 AI 技術解決方案

微軟為知名人工智慧技術解決方案之一，對話式 AI 技術（conversational AI）為其廣泛應用之解決方案。新加坡資訊通信媒體發展局（IMDA）和個人資料保護委員會（PDPC）在聯合發布「人工智慧治理模範框架」之同時，也發布人工智慧使用案例集（Compendium of AI Use Cases）⁵⁷，就選定個案分析其如何呼應人工智慧治理模範框架之指引。微軟之對話式 AI 技術解決方案入選案例集第 2 卷⁵⁸，係落實新加坡人工智慧治理模範框架之實例之一。

微軟於對話式 AI 技術解決方案之開發與供應過程中，透過下列措施促進人工智慧之良善治理（good governance）：

1. 負責任人工智慧之治理架構

微軟將公平、安全可靠、隱私保障、包容性、透明性和課責性列為人工智慧之核心原則，並設立工程和研發部門 AI、倫理和影響委員會（AI, Ethics, and Effects in Engineering and Research (Aether)

⁵⁶ Antonia, The First Virtual Public Servant In Romania, <https://www.romaniajournal.ro/society-people/antonia-the-first-virtual-public-servant-in-romania/> (accessed 2 December 2022).

⁵⁷ Singapore IMDA & PDPC, Compendium of AI Use Cases, Practical Illustrations of the Model AI Governance Framework (2020), <https://www.pdpc.gov.sg/-/media/Files/PDPC/PDF-Files/Resource-for-Organisation/AI/SGAIGovUseCases.pdf> (accessed 2 December 2022); Compendium of AI Use Cases (Volumn 2), Practical Illustrations of the Model AI Governance Framework (2020), <https://file.go.gov.sg/ai-gov-use-cases-2.pdf> (accessed 2 December 2022).

⁵⁸ Singapore IMDA & PDPC, Compendium of AI Use Cases (Volumn 2), Practical Illustrations of the Model AI Governance Framework (2020), 25-30.

Committee) 和負責任 AI 辦公室 (Office of Responsible AI)。Aether 委員會由技術、研究、法律等不同部門人員組成，負責就 AI 相關之關鍵挑戰向管理層提供建議，而負責任 AI 辦公室依 Aether 委員會建議落實公司內部治理及公共政策事務。

2. 選擇人工智慧之適當人為監督水平

於人工智慧應用開發過程中，微軟依具體使用案例 (use case) 之風險水平及對當事人之影響，確定適當之人為監督水平。例如，在保險理賠情境中，若對話式 AI 技術被用於篩選及回應消費者簡單提問，複雜問題將交由人工處理，則人為干預 (human-over-the-loop) 係適當方式，然若對話式 AI 技術係用於實質分析消費者理賠請求，則應採人為控制 (human-in-the-loop) 之監督方式。

微軟透過提供技術文件、非技術性指引、實例演示、訓練資料等方式，向其客戶解釋負責任之 AI 應用方式，並透過契約要求客戶不得以違法方式應用 AI、以高風險方式應用 AI 時應搭配適當安全維護措施。

3. 提升人工智慧之公平性

為防範人工智慧之偏見，微軟基於經驗，總結出識別和降低偏見之多種方式，例如持續評估訓練資料之代表性和品質，釐清資料流和相關特徵；採用機器學習及關鍵字過濾技術，確保 AI 應用適當偵測

和回應使用者輸入之敏感或不當資訊；使用 AI 公平性檢核表表（AI Fairness Checklist）⁵⁹評估偏見等。

4. 確保人工智慧之透明性

微軟針對人工智慧技術之功能和限制，向各類利害關係人（例如企業客戶、監理機關、於 AI 直接互動之使用者、第三方等）分別提供有意義之資訊，以此提升人工智慧之透明性，贏取利害關係人之信賴。其中，微軟強調應向直接使用者明確告知其正與人工智慧應用互動。微軟並鼓勵事前向使用者揭露個人資料蒐集與利用相關資訊並取得同意。相關措施包括向使用者提供隱私設定頁面、個人資料存取/刪除按鈕等。

八、我國高雄市政府 1999 智能客服萬事通⁶⁰

高雄市政府 1999 萬事通智能客服系統，整合市民、環保、交通、地政、衛生、社福與地方稅務資訊，以 AI 文字聊天機器人形式為市民提供 24 小時單一窗口服務。

運用 AI 人工智慧技術，民眾可透過電腦或手機搜開啟機器人智能交談對話模式，就各項問題提供互動式諮詢。如此便可不受地域時間限制，彌補在非上班時間無法服務民眾的缺口。

⁵⁹ Microsoft, AI Fairness Checklist, <https://query.prod.cms.rt.microsoft.com/cms/api/am/binary/RE4t6dA> (accessed 2 December 2022).

⁶⁰ <https://robot.kcg.gov.tw/wise/webchat/default/?t=oC56ls9rDQCvE6pSN%2B9i0w%3D%3D>，最後到訪為 111 年 12 月 10 日。

第四節 案例比較分析

由本研究目前彙整之他國案例來看，各國公私部門於服務中導入人工智慧應用以提升服務品質與效率已所在多見。從功能面而言，本研究研析案例比較如表 4 所示。由該比較表可知，本研究蒐集之對 1999 熱線較具參考價值之案例中，人工智慧應用多具資訊提供、輔助政府作業之功能，以期實現降低人工作業負擔、提升服務效率之目的，而多無涉對民眾使用公共服務之實質決策。

從法規遵循面而言，本研究經蒐集評估公開資料，發現各國導入人工智慧應用過程中，較注重技術面之方案評估，而較少見執行專門性適法性評估，此或因國際上尚未通過施行正式的人工智慧相關法律，各機關均回歸個人資料保護及其他法律適用，並遵守特定人工智慧指引，且與 1999 熱線相類似之人工智慧應用多具資訊提供、輔助政府作業之功能，並無實質決策之故。

本章節所研究之案例雖與前文第三章所研究之法規或指引之來源國家有所重疊，然因法規制定進程、指引之內容或效力、公開資料可獲得性等因素，大多數案例對法規或指引之遵循狀況較難具體研判。但新加坡微軟對話式 AI 技術解決方案係落實新加坡人工智慧治理模範框架之實例之一，透過該案例，可較明確地感受個案中如何落實負責任之治理架構、選擇適當人為監督水平、確保透明性等。

表 4 實務案例比較表

功能或技術 實務案例		與使用者 直接 互動	人為監督 方式	技術*			
				NLP	ML	STT	TTS
服務專線	比利時 Flemish Infoline	是	人為介入 回應進線提問	√	√	√	√
	愛爾蘭 VDA	是	人為介入 回應進線提問	√	√	√	√
	西班牙 060 專線	是	人為介入 回應進線提問	√	√	√	
	我國勞動部 多元智能電服中心	是	人為介入 回應進線提問	√	√		
語音辨識	愛沙尼亞 Hans	否	人為確認 產出內容	√	√	√	
	德國 C-19	是	使用者另行聯繫有關機關	√	√		
	德國難民語言偵測	否	產出結果僅供參考，由工作人員決策	√	√	√	
聊天機器人	英國 DVLA 聊天機器人	是	人為介入 回應線上提問	√	√		
	奧地利 Mona	是	使用者另行聯繫工作人員	√	√		
	丹麥 Kiri	是	使用者另行聯繫工作人員	√	√		
	愛沙尼亞 Iti	是	人為介入 回應線上提問	√	√		
	芬蘭 OuloBot	是	人為介入 回應線上提問	√	√		
	羅馬尼亞 Antonia	是	使用者另行聯繫工作人員	√	√		
	新加坡微軟對話式 AI 技術解決方案	是	視使用案例 需求而定	√	√	√	√

* NLP：自然語言處理技術；ML：機器學習技術；

STT：語音轉文字技術；TTS 文字轉語音技術。

第五章 臺北市民當家熱線現況分析

第一節 1999 臺北市民當家熱線簡介

北市府自民國 94 年起設立 1999 臺北市民熱線，作為市政大樓電話單一窗口，復於民國 97 年將該熱線升級優化為「1999 臺北市民當家熱線」，以「一個號碼、全面服務」為目標，構建整合性電子化政府服務平台，向市民提供「諮詢服務」、「轉接服務」、「告知電話」、「陳情服務」、「派工服務」五大類服務。為配合國家發展委員會服務型智慧政府 2.0 推動計畫，致力推動數位轉型，現擬於 1999 熱線導入人工智慧優化市民服務，以精進政府服務體驗，提升政府服務效率，強化政民互信。初階規劃由人工智慧以語音轉文字方式輔助話務人員記錄民眾來電、分析諮詢問題與建議答案，另對文本內容加以關鍵字分析。

個人自主控制個人資料之資訊隱私權乃我國憲法第 22 條所保障之人民基本權利，我國並以個人資料保護法（以下稱個資法）規範個人資料之蒐集、處理及利用。北市府規劃以人工智慧此一新興技術協助回應民眾於 1999 熱線之進線，必將以該技術蒐集、處理及利用民眾個人資料，從而涉及我國個資法中各項規範之遵循及技術性因素對民眾權益之影響，而有事前評估法規遵循風險及權利保障措施之必要。

為此，本研究針對 1999 熱線⁶¹導入人工智慧後所涉之個人資料蒐集、處理及利用，執行個資保護影響評估（Data Protection Impact Assessment, DPIA），以檢視本服務人工智慧應用方案之合目的性及必要性，評估對我國個人資料保護法各項規範之遵循狀況，識別本服務導入人工智慧後對民眾個人資料的潛在風險，並據以採取適當降低風險的措施。

第二節 個資保護影響評估（DPIA）

一、 評估程序

本次個資保護影響評估採人員訪談、文件審閱、操作體驗等方式進行，由研究團隊審查北市府提出之 1999 熱線標準作業流程、人工智慧應用導入規劃說明、與外部廠商之個人資料保護附加條款等，並訪談北市府人員及透過電子通訊方式溝通 1999 熱線服務模式涉及的個人資料作業流程與安全維護細節，再以使用者身分親自進線使用 1999 熱線服務，以確認 1999 熱線服務流程。

本次評估之法規要求事項乃依我國個人資料保護法提出，針對該法規範的各項蒐集、處理、利用個人資料之原則、要件與義務，檢核

⁶¹ 北市府單一陳情服務系統現已發展出「線上單一陳情系統」、「手語視訊服務系統」、「Hello Taipei APP」、「台北通 APP」等多媒介服務模式，然考量 1999 市民熱線在各項陳情管道中人力需求最為密集、自單一陳情系統上線以來至 111 年 11 月底止，網路管道受理案件佔比為 45.7%，1999 管道為 46.2%，其他（書信、臨櫃等）管道佔 8.1%。且北市府人工智慧導入規劃以 1999 熱線為重心，故本評估標的以電話「1999 臺北市民當家熱線」為限，尚不包括線上、手語視訊、APP 等管道。

1999 熱線服務模式涉及之個人資料作業流程的法規遵循程度，並就剩餘風險提供意見。

個人資料作業流程之個資保護影響評估以法規遵循為重心，並受資安技術影響。本次評估除於法規遵循之「個人資料的安全維護面向」檢視 1999 熱線已採行之技術性安全維護措施外，並基於資料遭竊取、洩漏、竄改或其他侵害視角，適度涵蓋可立即識別的資安技術風險，對剩餘風險提供意見。惟本評估報告就資安技術措施之風險識別均以北市府工作人員對安全技術之說明作出判斷，研究團隊並未在技術方面稽核北市府說明與實際情形的一致性，併此敘明。

二、 背景說明與資料流

1. 1999 熱線服務方式

1.1 1999 熱線目前提供「諮詢服務」、「轉接服務」、「告知電話」、「陳情服務」、「派工服務」五大類服務，進線需求皆由人工使用客製化資訊系統處理。具體而言，民眾以手機、市內電話或台北通 APP 內建網路電話功能撥打 1999 後，由話務人員接聽其進線，並依民眾電話中所述內容，判斷其需求為市政諮詢、電話轉接、告知、陳情案件或派工案件，並依相關標準作業流程處理：

1.1.1 如為市政諮詢案件，由接線之話務人員查詢市府常見問答資料

庫等資訊，即時回覆民眾提問。如無法即時回覆，必要時，則請民眾留下聯絡方式，由話務人員查明相關資訊後回覆。進線及回電中與話務人員之對話皆全程錄音，並將通話資訊記錄於話務系統。

1.1.2 如為轉接來電案件，由接線之話務人員透過話務系統轉接至北市府內相關業務單位或承辦人員。進線民眾與話務人員通話將進行錄音，進線資訊記錄於話務系統，與業務單位人員之通話原則不予錄音。

1.1.3 如進線民眾所詢問屬中央權責單位或外縣市政府範圍，則由接線之話務人員告知民眾業務單位、名稱及聯絡方式，請民眾改撥。

1.1.4 如為陳情案件，由接線之話務人員執行陳情人資料確認程序後，將陳情內容登錄於臺北市單一陳情系統，並由陳情系統產出案件編號及案件密碼以簡訊方式發送至陳情人手機。如陳情人日後來電詢問陳情處理進度，則由話務人員請陳情人提供案件編號及查詢密碼後，於單一陳情系統前端查詢後向民眾告知進度。進線民眾與話務人員通話將進行錄音，進線資訊記錄於話務系統。

1.1.5 如為派工案件，由接線之話務人員執行陳情人資料確認程序後，

將陳情內容登錄於 1999 派工系統，並由派工系統產出案件編號及案件密碼以簡訊方式發送至陳情人手機，依陳情人陳情內容，亦可能同時於臺北市單一陳情系統登錄該案件。如陳情人日後來電詢問陳情處理進度，則由話務人員請陳情人提供案件編號及查詢密碼後，於單一陳情系統前端查詢後向民眾告知進度。進線民眾與話務人員通話將進行錄音，進線資訊記錄於話務系統。

- 1.2 依北市府之初步規劃，將於話務系統導入人工智慧語音辨識功能，將進線民眾所述內容即時轉化為文字，並對文字稿執行關鍵字分析，以減輕話務人員記錄負擔，提升記錄準確性，協助話務人員理解進線民眾需求。因該功能主要作為話務人員之輔助工具，故該功能導入後，前述服務流程應無需明顯調整。如北市府將人工智慧用於進線分流、簡單回覆民眾提問等，將涉服務流程之調整，然因其屬長期規畫，現階段尚難確定服務流程之具體調整方式。

2. 1999 熱線蒐集、處理及利用個人資料之目的

- 2.1 依 1999 熱線之功能與服務方式，其蒐集、處理及利用個人資料之目的可區分兩類，即「市政熱線服務之提供與管理」與「市政服務品質提升與優化」。

2.1.1 市政熱線服務之提供與管理：1999 熱線之核心功能係接聽民眾進線，依民眾所述需求答覆市政諮詢、轉接負責機關、受理陳情及派工處理通報事件，並提供民眾查詢處理進度及結果之機制。為提供該等服務，北市府基於行政程序法等法規要求以及提供服務之現實可行性，建設、管理及維護資訊系統，作為 1999 熱線服務之基礎設施，並以該資訊系統接聽民眾進線，蒐集進線民眾姓名、聯絡方式、來電需求等個人資料，做成系統內紀錄，提供權責機關存取該紀錄並為後續處理。

2.1.2 市政服務品質提升與優化：1999 熱線係北市府行政服務之一部分，由北市府投入人力、設備、場地、資金等資源執行營運及維護。為提升資源利用效率並改善服務品質，1999 熱線可能對民眾進線時間、頻率、內容、處理狀況等加以分析，用以評量服務成效、研判民眾需求、調整服務流程、作為施政參考等。

2.2 人工智慧應用之規劃、開發、導入與執行階段，1999 熱線蒐集、處理及利用個人資料，將增加「人工智慧應用之訓練與改善」之目的。1999 熱線規劃導入人工智慧協助記錄民眾進線內容、理解民眾需求、乃至簡單回應進線需求等。該人工智慧技術之設計、開發、測試與部署，將須由北市府提供民眾進線資料作為基礎素材。該人工智慧技術實際投入應用後，亦須藉由處理民眾進線而

修補缺陷與精進功能。

3. 個人資料範圍

3.1 綜合觀察話務系統、陳情系統、派工系統內相關資料欄位及 1999

熱線作業各項規範，1999 熱線所蒐集、處理及利用之個人資料，

可分為如下幾類：

3.1.1 進線民眾之基本資料，例如：姓名、電話（系統可自動記錄進

線電話，亦可由進線人自行提供其他聯絡電話）、性別等，必要時亦會蒐集聯絡地址、電子郵件信箱、傳真號碼、身分證字號等。

3.1.2 進線民眾服務需求相關紀錄，例如案件分類、案件來源、案件

主旨、案件內容、派工地點、回覆方式、服務小結、話務人員於備註中記錄與進線人之對話內容，例如所述市政諮詢問題、電話轉接之目標機關/部門、陳情所涉相對機關/個人等。依進線人來電內容，可能涉及進線民眾以外他人之個人資料，例如進線人之送達代收人、陳情所涉相對人等；亦可能涉及個資法第 6 條所稱之特種個人資料，例如健康檢查資料、犯罪前科資料等。

3.1.3 1999 熱線依進線提供服務之紀錄，例如通話錄音、立案/分案紀

錄、滿意度調查紀錄，以及系統自動記錄之受理管道、小結編號/案件編號、案件密碼、受理時間、結束通話時間、回電紀錄等。

3.1.4 話務人員對進線人之分析、判斷相關資訊，例如評估進線人進線內容及次數，並對顯有濫用風險之進線人採取因應措施。

3.2 人工智慧應用之規劃、導入與執行階段，1999 熱線所涉之個人資料範圍是否變化，將須視所涉人工智慧應用之性質與功能而定。例如，導入人工智慧語音辨識功能後，因該功能主要作為話務人員之輔助工具，除其將進線人通話內容轉錄而得之文字稿外，1999 熱線所涉個人資料範圍應無基本維持不變。然如導入簡單回覆進線提問等更具獨立性之人工智慧應用，1999 熱線所涉之個人資料範圍可能會相應擴大。例如，進線民眾依人工智慧應用之提示操作，可能需提供更多資訊。

4. 個人資料流程

4.1 資料蒐集

4.1.1 民眾以手機、市話或網路電話撥打 1999 後，可依互動式語音系統（Interactive Voice Response，IVR）提示，自行輸入電話分機直接聯絡北市府內承辦單位，或選擇諮詢方向和語言後由話

務人員接聽進線。民眾進線接通後，1999 熱線即由系統開始電話號碼等個人資料蒐集作業，然 IVR 提示內容尚未包含個人資料蒐集告知相關內容。

4.1.2 話務人員接聽後，依 1999 熱線標準作業流程與進線民眾互動，並進一步蒐集進線民眾本人及其他自然人之個人資料，包括進線民眾稱呼、需求細節等，視民眾之進線需求，亦可能蒐集姓名、聯絡地址、電子郵件信箱、送達代收人、陳情對象等。依通話之進行及話務人員操作，1999 熱線系統將自動生成通話起訖時間等紀錄。若為陳情案件或派工案件，進線民眾須提供真實姓名及有效聯絡方式方可立案，並將由系統自動生成案件編號及案件密碼。

4.1.3 導入人工智慧語音辨識功能後，在法規關於民眾陳情之規範或要求未變之情形下，1999 熱線向進線人蒐集之資料範圍應不致變化，僅人工智慧應用運作過程中，可能因語音分析及識別而產生必要紀錄，例如進線人語音轉譯而得之文字稿、語音識別處理作業起止時間紀錄、話務人員編輯紀錄等。

4.2 資料處理

4.2.1 1999 熱線蒐集之個人資料，包括話務人員對進線人所述內容之紀錄、系統自動產生之進線資訊等，皆依標準作業流程，記錄

於話務系統、陳情系統或派工系統。進線人與話務人員之通話亦將由系統全程錄音，初步保存於話務系統，並由委外廠商定期轉存至資料庫中。依據北市府研究發展考核委員會(研考會)、資訊局公告之「研考會保有及管理個人資料之項目彙整表」、「資訊局保有個人資料檔案公開項目彙整表」⁶²，1999 話務電話紀錄、1999 話務客戶管理、單一陳情系統及 1999 派工系統所含之個人資料，係由北市府資訊局保有，而民眾來電 1999 之通話錄音檔，係由北市府研考會保有。

4.2.2 1999 熱線導入人工智慧語音辨識功能後，進線人與話務人員之通話內容，將由人工智慧應用分析並處理為轉譯文字稿，並對文字稿中關鍵字進行標示，供話務人員參閱、核對、修改，以及以自動或手動方式將該文字內容轉貼至系統相關欄位。如未來導入之人工智慧應對系統架構及作業流程有實質調整，例如以雲端服務方式提供語音辨識，則資料之處理方式、儲存位置等將相應改變。

4.3 資料利用

4.3.1 內部利用：

⁶² 研考會保有及管理個人資料之項目彙整表請見：
https://rdec.gov.taipei/News_Content.aspx?n=F5702E09C9B011E9&sms=40C86180686507FA&s=B03182725078223D；資訊局保有個人資料檔案公開項目彙整表請見：
<https://doit.gov.taipei/cp.aspx?n=39C8F94C564144D4>。

- 4.3.1.1 1999 熱線所蒐集之個人資料，首先用於對民眾需求之判斷分類，並依標準作業流程回應民眾所提需求，例如：依民眾所述問題協助立案；在陳情/派工案件立案過程中回撥確認身分；依民眾提供之資訊查詢並告知陳情/派工案件處理結果；依民眾要求撤銷案件；以及於無法及時回覆民眾需求，必要時請民眾留下聯絡方式、由話務人員查明後回撥等。
- 4.3.1.2 1999 熱線並會將所蒐集之個人資料用於服務管理，例如：於系統內分派/改派案件予權責機關；依標準作業流程，就進線人所提案件製作各類存查紀錄等。
- 4.3.1.3 1999 熱線亦可能將所蒐集之個人資料用於分析服務使用狀況、改善服務品質、調整服務制度、供作施政參考等。
- 4.3.1.4 人工智慧語音辨識應用導入過程中，1999 熱線可能會將所蒐集之個人資料用於測試廠商提供之人工智慧應用的準確度、辨識率等，以利評估廠商履約品質。該人工智慧應用導入後，1999 熱線會將所蒐集之個人資料交由該應用進行辨識、關鍵字分析等利用。
- 4.3.1.5 北市府規劃採購外部廠商之技術及資訊服務，並導入 1999 熱線作業流程。該人工智慧之開發與導入過程中，北市府將委託廠商處理和利用 1999 熱線所蒐集之個人資料。例如，導

入方案評估階段，將需委由廠商利用錄音資料作概念驗證（proof of concept, POC）。於人工智慧語音辨識應用開發過程中，廠商將需把錄音資料預處理為適當形式，用以訓練/調適聲學模型，以及將系統內話務人員對錄音內容之紀錄輔助訓練，以及將進線錄音及系統內進線內容紀錄用於測試人工智慧效果等。

4.3.2 外部利用

4.3.2.1 1999 熱線登載於陳情系統、派工系統中之個人資料，將供權責單位依其內容派員處理。1999 熱線陳情系統預設不提供陳情人之聯絡資料予權責單位承辦人員，僅限陳情人要求權責單位以電話或書面回覆處理結果時，方向權責單位承辦人員提供聯絡資料。

4.3.2.2 1999 熱線亦可能依申請將資料對外提供與其他機關，例如：配合檢調機關案件偵辦需要，依刑事訴訟法、通訊保障及監察法等提供錄音資料；依北市府內其他機關業務辦理需要，經簽核程序後於必要範圍內提供錄音資料；或依民意代表請求，提供陳情案件清冊等資料。

4.4 資料之安全維護

4.4.1 北市府就 1999 熱線訂有「資訊安全管理標準作業流程」，規定

1999 熱線服務提供及話務中心運作過程中個人資料保護及資訊安全防護之各項作業原則，具體包括人員教育訓練、實體及環境安全、作業安全、系統安全四大部分，並就安全事故之因應訂定標準程序。依該標準作業流程，1999 熱線之各項資料、文件及作業系統，原則僅限於話務中心作業場所內，依作業權限存取，且非經許可不得攜出話務中心之外。話務中心電腦硬體及軟體採常態性防護，且電腦內所存文件資料，非經許可不得以移動式儲存裝置轉存。

4.4.2 1999 熱線以整體話務服務委外方式營運，並由北市府研考會督導管理，故話務中心的話務人員於提供 1999 熱線服務範圍內，皆屬受北市府研考會委託而蒐集、處理及利用個人資料。前述資訊安全管理標準作業流程直接適用於話務中心話務人員，並訂有懲戒規定。此外，北市府與各參與服務之話務人員個人亦簽訂保密協議。

4.4.3 1999 熱線擬導入外部廠商開發及導入人工智慧應用，此一過程中會將錄音、系統內進線內容紀錄將委由廠商用於人工智慧訓練和測試。據北市府提供資料及北市府網站公開資訊，參與人工智慧開發導入專案之委外廠商皆簽署個人資料保護條款，且參與專案之廠商員工，亦以個人身分簽署保密同意書及切結書，

約定資料保密事宜。在目前已執行之小範圍內部測試中，北市府亦要求廠商派執行專人至 1999 熱線話務中心內部進行相關作業，由北市府工作人員協同廠商人員設置地端系統及工具，該內部測試所產出之語音辨識結果僅限於話務中心內使用及呈現，並要求專案團隊於測試完成後即刪除錄音檔案⁶³。

⁶³ 參見 111 年 4 月 15 日臺北市政府研究發展考核委員會新聞稿，錄音檔轉文字合作內部測試增進 1999 話務人員服務效率，
https://1999.gov.taipei/News_Content.aspx?n=56435265EBCBF5BE&sms=EB0622377E126BF0&s=706744107896516A。

三、 法規遵循風險識別

1. 誠信原則

個人資料保護法 第 5 條

個人資料之蒐集、處理或利用，應尊重當事人之權益，依誠實及信用方法為之，不得逾越特定目的之必要範圍，並應與蒐集之目的具有正當合理之關聯。

1.1 1999 熱線以「一個號碼，全面服務」為目標，向民眾提供「諮詢」服務、「轉接服務」、「告知電話」、「陳情服務」、「派工服務」五大類服務，在民眾自願進線提出需求的基礎上，依據行政程序法等法規要求，基於「履行法定職務之必要」等法律依據，向民眾提供「市政諮詢」、「轉接來電」、「接受陳情」，記錄民眾進線中提供之資訊，並轉達予權責機關，以此滿足民眾的需求。北市府（1999 熱線）因此建置相關專門性、客製化之資訊系統，以標準化方式製作及保存民眾進線紀錄，並提供權責機關存取權限，以利確保服務品質，強化服務管理。1999 熱線之系統設計亦考量個資當事人隱私保護，以「進線人聯絡資訊不提供予權責機關」為默認設定，應係尊重個資當事人（民眾）之權益，並符合誠實及信用之作為。

1.2 1999 熱線規劃導入人工智慧應用，初階由人工智慧以語音轉文字

方式輔助話務人員記錄民眾來電，長期規劃包括由人工智慧執行進線分流，並初步回應民眾簡單提問等，旨在以數位技術優化服務流程、精進政府服務體驗，亦符合我國數位轉型之政策取向，於該目的之必要範圍內，以合於個人資料應認亦屬係尊重個資當事人（民眾）之權益，並符合誠實及信用之方法。

2. 目的限制原則

個人資料保護法 第 5 條

個人資料之蒐集、處理或利用，應尊重當事人之權益，依誠實及信用方法為之，不得逾越特定目的之必要範圍，並應與蒐集之目的具有正當合理之關聯。

2.1 資料蒐集

2.1.1 在資料蒐集階段，民眾使用 1999 熱線服務時，如提出陳情或派工需求，須提供真實姓名與聯絡電話，方可完成立案。此等身分識別資料係依行政程序法第 173 條、臺北市政府及所屬各機關處理人民陳情案件注意事項第 5 點與第 12 點而蒐集，且為及時向陳情人確認重要陳情內容之疑義、告知處理結果所必須，應未逾越陳情案件處理等特定目的之必要範圍。

2.1.2 此外，1999 熱線蒐集民眾進線時之通話內容，以及來電號碼、時間等。就陳情或派工服務而言，蒐集此等資料係依行政程序

法第 169 條、臺北市政府及所屬各機關處理人民陳情案件注意事項第 5 點規定為之，且為迅速、確實及有效處理人民陳情案件所必須；就市政諮詢、轉接來電服務而言，蒐集此等資料係 1999 熱線管理及改善服務所必須，且與 1999 熱線服務之技術特性相符。據此，1999 熱線蒐集該等資料，應未逾越陳情及其他市政服務之提供、1999 熱線資訊系統管理等特定目的之必要範圍。

2.1.3 1999 熱線導入人工智慧語音辨識功能，係為提升陳情服務之品質與效率，且導入該功能後，向民眾直接蒐集之身分識別資料、進線通話資料範圍應不致變化，系統因導入該功能而額外生成之紀錄，亦限於語音辨識之技術需求範圍內，故應認亦未逾越陳情及其他市政服務之提供、1999 熱線資訊系統管理等特定目的之必要範圍。

2.2 資料處理

2.2.1 1999 熱線將民眾陳情、派工等進線內容、時間、來電號碼等做成紀錄，並於提供服務期間內持續保存，以供權責機關存取及進線民眾查閱處理進度，符合服務提供及管理之現實需要，應未逾越特定目的之必要範圍。當次服務結束後，服務紀錄於系統內繼續保存相當時間，以滿足進線民眾查閱處理結果、機關

評估服務成效等需求，亦未逾越資料蒐集特定目的之必要範圍。

2.2.2 1999 熱線導入人工智慧語音辨識功能後，該功能產出之進線人通話文字內容及關鍵字標示，將提供給話務人員檢視、編輯、更正等，以利減輕話務人員記錄負擔、協助話務人員理解進線人需求，提升紀錄內容精準度，應未逾越資料蒐集特定目的之必要範圍。

2.3 資料利用

2.3.1 資料利用階段，1999 熱線就民眾進線所述需求，依標準作業流程對資料作各項利用，例如將資料提供與權責機關用於陳情處理或派工作業、對民眾回撥確認身分、分析服務提供成效等，應未逾越資料蒐集特定目的之必要範圍。

2.3.2 1999 熱線應請求而提供相應資料予外部單位，例如檢調機關、北市府內其他機關業務、民意代表等，因該等對外提供已超出 1999 熱線服務提供及管理之範圍，且多具偶然性與不確定性，故屬資料蒐集特定目的範圍外之利用，將於後文「一般/特種個資的蒐集處理及利用」、「當事人同意」等小節具體分析其適法性。

2.3.3 1999 熱線於人工智慧語音辨識應用開發及導入過程中，將委由廠商利用進線錄音、系統內進線內容紀錄等，以進行人工智慧

概念驗證、訓練和測試。因該等驗證、訓練及測試與社會通念認知下政府職務執行之常規作業尚有落差，故除確有紀錄表明資料蒐集之時已具「人工智慧應用之訓練與改善」目的外，宜認其亦屬資料蒐集特定目的範圍外之利用，其適法性將於後文「一般/特種個資的蒐集處理及利用」、「當事人同意」等小節具體評估。然若 1999 熱線明確將人工智慧應用之訓練與改善作為資料蒐集之特定目的之一，則此等將成為特定目的必要範圍內之利用，惟應適當確定其蒐集之法律依據、踐行告知等。

3. 正當合理關聯

個人資料保護法 第 5 條

個人資料之蒐集、處理或利用，應尊重當事人之權益，依誠實及信用方法為之，不得逾越特定目的之必要範圍，並應與蒐集之目的具有正當合理之關聯。

3.1 1999 熱線為受理民眾之進線需求並提供相應服務，依法規關於陳情處理之具體規範，以及將陳情內容傳達權責機關、提供民眾查閱及通知等現實需求，而蒐集、處理及利用民眾身分識別資料（姓名、聯絡電話）與進線內容紀錄，均與蒐集資料之目的具有正當合理關聯。

3.2 1999 熱線導入人工智慧語音辨識功能後，以該功能蒐集、處理及

利用民眾進線內容紀錄資料，與該等資料之蒐集目的亦具有正當合理關聯。

3.3 1999 熱線為藉助數位技術優化服務流程、精進政府服務體驗，規劃導入人工智慧應用，而於其開發與導入過程中，將進線錄音、系統內進線內容紀錄提供予廠商作訓練與測試用途，委託廠商處理及利用等資料，縱依蒐集目的之範圍判斷，構成對個資於蒐集之特定目的範圍外利用，然與 1999 熱線數位轉型之目的間，應認具正當合理關聯，惟其應符合個資法規定之其他適法要件，例如適當之法律依據等，將於後文「一般/特種個資的蒐集處理及利用」、「當事人同意」等小節具體討論。

4. 一般個資的蒐集處理及利用

個人資料保護法 第 15 條

公務機關對個人資料之蒐集或處理，除第六條第一項所規定資料外，應有特定目的，並符合下列情形之一者：

- 一、執行法定職務必要範圍內。
- 二、經當事人同意。
- 三、對當事人權益無侵害。

個人資料保護法 第 16 條

公務機關對個人資料之利用，除第六條第一項所規定資料外，應

於執行法定職務必要範圍內為之，並與蒐集之特定目的相符。但有下列情形之一者，得為特定目的外之利用：

- 一、法律明文規定。
- 二、為維護國家安全或增進公共利益所必要。
- 三、為免除當事人之生命、身體、自由或財產上之危險。
- 四、為防止他人權益之重大危害。
- 五、公務機關或學術研究機構基於公共利益為統計或學術研究而有必要，且資料經過提供者處理後或經蒐集者依其揭露方式無從識別特定之當事人。
- 六、有利於當事人權益。
- 七、經當事人同意。

4.1 資料之蒐集與處理

4.1.1 1999 熱線所蒐集和處理之個人資料，絕大部分非屬個資法第 6

條第 1 項所稱之特種個人資料。1999 熱線蒐集和處理個人資料之目的，首先在於「市政熱線服務之提供與管理」與「市政服務品質提升與優化」，於此兩項目的之必要範圍內蒐集和處理個人資料，與北市府內相關機關之法定職務有密切關聯。

4.1.2 依個資法施行細則第 10 條，個資法第 15 條第 1 款所稱之「法

定職務」係指「下列法規中所定公務機關之職務：一、法律、法律授權之命令。二、自治條例。三、法律或自治條例授權之自治規則。四、法律或中央法規授權之委辦規則」。依我國司法

實務及個資法主管機關（法務部、國家發展委員會）之見解，個資法所稱之「法定職務」，除規範公務機關行為的作用法外，亦包含依組織法所劃定之權限在內，公務機關基於組織法賦予之任務，得在目的限制的必要範圍內蒐集、處理或利用個人資料⁶⁴。

4.1.3 北市府依行政程序法第 170 條之授權，訂定臺北市政府及所屬各機關處理人民陳情案件注意事項，於該注意事項第 5 點明定當事人以電話方式言詞陳情之處理程序。北市府亦依臺北市政府組織自治條例第 6 條之授權，訂定臺北市政府研究發展考核委員會組織規程，於該組織規程第 3 條第 4 款、第 6 款分別規定，北市府研考會內設服務精進組，掌理「服務品質提升及為民服務規劃考核等事項」，並設話務管理組，掌理「1999 市民熱線督導管理及提升話務服務品質等工作」。北市府並依臺北市政府組織自治條例第 6 條之授權，訂定臺北市政府資訊局組織規程，於該組織規程第 3 條第 3 款明文規定，北市府資訊局內設系統研發中心，掌理「共同政務資訊系統、共同服務資訊系統、共通系統發展平臺等事項之規劃、推動及管理」。

⁶⁴ 最高行政法院 106 年度判字第 54 號判決，法務部 106 年 8 月 7 日法律字第 10603510070 號、107 年 7 月 13 日法律字第 10703510240 號、107 年 10 月 19 日法律字第 10703515980 號、108 年 7 月 12 日法律字第 10803509450 號函釋，國家發展委員會 108 年 1 月 22 日發法字第 1080000958 號函釋參照。

4.1.4 因此，1999 熱線以資訊系統接聽民眾進線，提供諮詢、陳情等各項服務，並據此提升市政服務品質，應屬北市府研考會與資訊局之法定職務。北市府建設 1999 熱線，基於陳情及其他市政服務之提供、1999 熱線資訊系統管理等特定目的，於市政諮詢、電話轉接、受理陳情及派工等服務之提供及管理必要範圍內，以資訊系統接聽民眾自願進線電話，蒐集、處理及利用進線民眾及陳情所涉其他自然人之個人資料，應屬北市府研考會與資訊局執行法定職務所必要，係以個資法第 15 條第 1 款「執行法定職務必要範圍內」為法律依據。

4.1.5 1999 熱線導入人工智慧語音辨識功能後，因該功能主要作為 1999 熱線內部作業之輔助性措施，對 1999 熱線向進線人蒐集之資料範圍較無影響，故其所涉之個人資料處理及利用，應亦以個資法第 15 條第 1 款「執行法定職務必要範圍內」為法律依據。

4.1.6 若 1999 熱線以「人工智慧應用之訓練與改善」作為個人資料之蒐集目的之一，因其人工智慧應用之開發、訓練、測試與改善等，其最終目的在於落實 1999 熱線服務轉型，優化 1999 熱線服務流程，提升市政服務品質，且客觀上有協助北市府法定職務執行之效，故應有主張以個資法第 15 條第 1 款「執行法定職

務必要範圍內」為法律依據之空間。

4.1.7 然依個資法主管機關之見解，基於個資法第 15 條第 1 款之「執行法定職務必要範圍」蒐集、處理及利用個人資料時，應選擇對當事人權益唯一或最小之侵害方式⁶⁵。考量人工智慧應用之開發並非受理及回應民眾進線需求所必須，且其訓練與測試通常較難為進線民眾所預見，1999 熱線如主張人工智慧之訓練與改善屬「執行法定職務必要範圍」，並進而依個資法第 8 條第 2 項主張免為向當事人踐行告知，可能引發當事人對該利用必要性及透明性之疑慮。為降低此等疑慮，1999 熱線可考慮向當事人適當告知個人資料用於人工智慧訓練與測試之情形。1999 熱線亦可改以個資法第 15 條第 2 款「經當事人同意」為法律依據，更加強化對當事人資料自主權之保障。

4.2 資料之利用

4.2.1 1999 熱線與蒐集個人資料後，依標準作業流程對資料作各項利用，例如將資料提供與權責機關用於陳情處理或派工作業、對民眾回撥確認身分、分析服務提供成效等，應屬執行法定職務必要範圍內，且與資料蒐集之特定目的相符，符合個資法第 16 條本文之規範。

⁶⁵ 國家發展委員會 110 年 6 月 9 日發法字第 1102000884 號函釋參照

4.2.2 1999 熱線導入人工智慧語音辨識功能後，因該功能主要作為 1999 熱線內部作業之輔助性措施，故其所涉之資料利用，亦應屬執行法定職務必要範圍內，且與資料蒐集之特定目的相符，合於個資法第 16 條本文之規範。

4.2.3 如前文「目的限制原則」小節第 2.3.2 點所述，1999 熱線將通話錄音、系統內紀錄等個人資料提供予外部單位，屬蒐集之特定目的外利用個人資料。其是否具備個資法第 16 條但書所列之合法事由，應視對外提供之對象與原因分別判斷。例如，依刑事訴訟法通訊保障及監察法等向檢調機關提供錄音資料，應係以第 16 條第 1 款「法律明文規定」為依據；基於行政協助之必要，遵循規定程序將錄音或案件清冊提供予北市府內其他機關或依民意代表，或可以第 16 條第 2 款「增進公共利益所必要」為依據。

4.2.4 如前文「目的限制原則」小節第 2.3.3 點所述，就現階段而言，1999 熱線委由廠商將通話錄音、系統內紀錄等個人資料用於人工智慧語音辨識應用之訓練與測試，視資料蒐集目的之內容與範圍，可能構成蒐集之特定目的外利用個人資料。因該等目的外利用係為落實 1999 熱線數位轉型，精進政府服務體驗，提升公共資源利用效率，同時減輕話務人員之作業負擔，故應有主

張以個資法第 16 條第 2 款之「增進公共利益所必要」為依據之空間。然個資法主管機關曾有見解認為，個資法第 16 條第 2 款規定之「增進公共利益」，係指「社會不特定多數人可以分享之利益而言，倘資料目的外利用僅係有利於公務機關執行業務，則與該條所稱公共利益之概念尚有差異」⁶⁶。就 1999 熱線而言，導入人工智慧應用固然有利於增進服務品質優化、服務效率提升之公益，然如特定人工智慧應用以便利話務中心運作為核心功能，增進公益之效果較間接，則可能落入前述實務見解所稱「僅係有利於公務機關執行業務」之範圍。針對此類人工智慧應用實例，可能較宜評估是否以個資法第 16 條第 7 款之「當事人同意」為法律依據。

5. 特種個資的蒐集處理及利用

個人資料保護法 第 6 條

I. 有關病歷、醫療、基因、性生活、健康檢查及犯罪前科之個人資料，不得蒐集、處理或利用。但有下列情形之一者，不在此限：

一、法律明文規定。

二、公務機關執行法定職務或非公務機關履行法定義務必要範圍內，且事前或事後有適當安全維護措施。

⁶⁶ 國家發展委員會 108 年 7 月 4 日發法字第 1082001104 號函釋參照。

三、當事人自行公開或其他已合法公開之個人資料。

四、公務機關或學術研究機構基於醫療、衛生或犯罪預防之目的，為統計或學術研究而有必要，且資料經過提供者處理後或經蒐集者依其揭露方式無從識別特定之當事人。

五、為協助公務機關執行法定職務或非公務機關履行法定義務必要範圍內，且事前或事後有適當安全維護措施。

六、經當事人書面同意。但逾越特定目的之必要範圍或其他法律另有限制不得僅依當事人書面同意蒐集、處理或利用，或其同意違反其意願者，不在此限。

II. 依前項規定蒐集、處理或利用個人資料，準用第八條、第九條規定；其中前項第六款之書面同意，準用第七條第一項、第二項及第四項規定，並以書面為之。

5.1 1999 熱線由進線民眾提供之個人資料中，通常不會包含個資法第 6 條規定之特種個人資料。

5.2 然因 1999 熱線所蒐集資料大部分為進線民眾於通話中主動提供，故不排除特定市政諮詢、陳情或派工案件將涉及健康檢查資料、犯罪前科等特種個人資料。因進線民眾主動提供該等資料，通常係因該等資料與諮詢問題、陳情或派工需求緊密相關，故 1999 熱線蒐集、處理及利用該等特種個人資料，應係服務提供及管理之所必要，應以個資法第 6 條第 1 項第 2 款「公務機關執行法定職務必要範圍內」為法律依據，且應依該項之提示性規定，於事前

或事後採行適當安全維護措施⁶⁷。

5.3 1999 熱線導入人工智慧語音辨識功能後，因該功能主要作為 1999 熱線內部作業之輔助性措施，應不致影響 1999 熱線所蒐集特種個人資料之範圍與數量，故該功能對特種個人資料之處理與利用，應得以「公務機關執行法定職務必要範圍內」為法律依據。

5.4 1999 熱線如以「人工智慧應用之訓練與改善」作為一般個人資料之蒐集目的之一，則有主張以「執行法定職務必要範圍內」為法律依據之空間，然為強化當事人資料自主權，亦可考量以「當事人同意」為法律依據，業如前文「一般個資的蒐集處理及利用」小節所分析。如所蒐集之資料蒐集涉及特種個人資料，亦可採類似分析，認其有主張以「公務機關執行法定職務必要範圍內」為法律依據之空間。

5.5 然因特種個人資料對當事人而言之私密程度更高，故從個資法第 5 條比例原則之角度，於個案中亦可能較宜以個資法第 6 條第 1 項第 6 款之「經當事人書面同意」為依據。因 1999 熱線設有服務網站並已與台北通 APP 串接，故此「書面」方式應得以「電子文件」為之，惟應符合電子簽章法第 4 條第 2 項「依法令規定應以書面為之者，如其內容可完整呈現，並可於日後取出供查驗者，

⁶⁷ 國家發展委員會 107 年 11 月 7 日發法字第 1072002035 號函參照。

經相對人同意，得以電子文件為之」之規定。如 1999 熱線考量取得「書面同意」於現實面較困難，則宜規劃資料分類制度，將包含特種個人資料之通話錄音及系統內紀錄排除於人工智慧應用訓練之外。

5.6 1999 熱線現階段將通話錄音、系統內紀錄等個人資料委由外部廠商用於人工智慧應用之訓練與測試，視資料蒐集目的之內容與範圍，可能構成對個人資料之目的外利用，業如前文「目的限制原則」小節所述。然個資法第 6 條並未區分特種個人資料之蒐集目的「範圍內」與「範圍外」利用之法律依據，故該等目的外利用之法律依據，應與前文第 5.4 點、5.5 點之分析相同，即應有主張以「執行法定職務必要範圍內」為法律依據之空間，然於具體個案中，亦可考量以「當事人書面同意」為法律依據。

6. 告知義務

個人資料保護法 第 8 條

- I. 公務機關或非公務機關依第十五條或第十九條規定向當事人蒐集個人資料時，應明確告知當事人下列事項：
- 一、公務機關或非公務機關名稱。
 - 二、蒐集之目的。
 - 三、個人資料之類別。

四、個人資料利用之期間、地區、對象及方式。

五、當事人依第三條規定得行使之權利及方式。

六、當事人得自由選擇提供個人資料時，不提供將對其權益之影響。

II. 有下列情形之一者，得免為前項之告知：

一、依法律規定得免告知。

二、個人資料之蒐集係公務機關執行法定職務或非公務機關履行法定義務所必要。

三、告知將妨害公務機關執行法定職務。

四、告知將妨害公共利益。

五、當事人明知應告知之內容。

六、個人資料之蒐集非基於營利之目的，且對當事人顯無不利之影響。

個人資料保護法 第 9 條

I. 公務機關或非公務機關依第十五條或第十九條規定蒐集非由當事人提供之個人資料，應於處理或利用前，向當事人告知個人資料來源及前條第一項第一款至第五款所列事項。

II. 有下列情形之一者，得免為前項之告知：

一、有前條第二項所列各款情形之一。

二、當事人自行公開或其他已合法公開之個人資料。

三、不能向當事人或其法定代理人為告知。

四、基於公共利益為統計或學術研究之目的而有必要，且該

資料須經提供者處理後或蒐集者依其揭露方式，無從識別特定當事人者為限。

五、大眾傳播業者基於新聞報導之公益目的而蒐集個人資料。

III. 第一項之告知，得於首次對當事人為利用時併同為之。

個人資料保護法 第 6 條

II. 依前項規定蒐集、處理或利用個人資料，準用第八條、第九條規定；其中前項第六款之書面同意，準用第七條第一項、第二項及第四項規定，並以書面為之。

6.1 告知之踐行

6.1.1 1999 熱線以通話方式提供服務，於進線接通前之 IVR 提示內容中，尚未納入個資蒐集告知相關內容。1999 熱線之各項標準作業流程中，尚無向進線民眾踐行告知之流程或例稿。1999 熱線設置「1999 市民當家熱線服務網」，提供 1999 熱線服務說明、FAQ 常見問答、系統維護公告等資訊，該網站「隱私權及資訊安全政策」內容尚未涵蓋 1999 熱線所涉個人資料之蒐集、處理及利用。據此，1999 熱線直接向個資當事人（進線民眾）蒐集個資時，應尚未告知蒐集個人資料之目的、利用個人資料之對象及方式、當事人權利等個資法第 8 條第 1 項規定應告知之事項。

6.1.2 1999 熱線除向進線民眾直接蒐集其個人資料外，亦有間接蒐集個人資料之情事（例如自進線民眾蒐集送達代收人、陳情對象之個人資料）。就 1999 熱線之各項標準作業流程及 1999 市民當家熱線服務網隱私權及資訊安全政策觀察，亦未見依個資法第 9 條第 1 項、第 3 項向個資當事人踐行告知。

6.2 免為告知之事由

6.2.1 依個資法第 8 條第 2 項第 2 款，「個人資料之蒐集係公務機關執行法定職務所必要」時，蒐集之公務機關得「免為告知」同條第 1 項之應告知事項。1999 熱線於市政諮詢、電話轉接、受理陳情及派工等服務之提供及管理必要範圍內，以資訊系統接聽民眾進線，蒐集進線民眾個人資料，應認屬北市府研考會與資訊局「執行法定職務所必要」，業如前文第「一般個資的蒐集處理及利用」小節所分析。因此，1999 熱線應得免依個資法第 8 條第 1 項向民眾踐行告知。

6.2.2 於 1999 熱線間接蒐集個人資料之情形，依個資法第 9 條第 2 項第 1 款援用同法第 8 條第 2 項之規範，「個人資料之蒐集係公務機關執行法定職務所必要」時，間接蒐集個人資料亦得免為告知，且其「法定職務」之範圍解釋，與前文直接蒐集個人資料情景下相同。據此，1999 熱線於服務提供及管理之必要範圍內，

間接蒐集進線民眾個人資料，應認屬北市府研考會與資訊局「執行法定職務所必要」，而得免依個資法第 9 條第 1 項、第 3 項向民眾踐行告知。

6.3 綜合前述，1999 熱線於市政諮詢、電話轉接、受理陳情及派工案件之服務必要範圍內，蒐集進線民眾個人資料，應可援用「個人資料之蒐集係公務機關執行法定職務所必要」，作為免於踐行告知之依據。1999 熱線導入人工智慧語音辨識功能後，因該功能主要作為 1999 熱線內部作業之輔助性措施，故其所涉之資料處理與利用，應仍屬「個人資料之蒐集係公務機關執行法定職務所必要」之範圍，從而據此得免為告知。

6.4 如前文「一般個資的蒐集處理及利用」、「特種個資的蒐集處理及利用」兩小節所分析，若 1999 熱線以「人工智慧應用之訓練與改善」作為個人資料之蒐集目的之一，其蒐集雖有主張「執行法定職務必要範圍」從而免為踐行告知之空間，然因該等目的與社會通念下政府職務之執行較有落差，且較難為當事人事前預見，且人工智慧應用之範圍與功能多樣，對當事人權益之影響程度不一，故從保障當事人對其個人資料之知情與自主控制之視角，1999 熱線宜考量於服務流程中適當納入個人資料蒐集之告知，包括資料將用於「人工智慧應用之訓練與改善」目的、所涉資料類別、利

用之期間、地區、對象和方式，以及當事人得行使之權利等。

6.5 個資法並未要求公務機關或非公務機關就個人資料之目的外利用向當事人踐行告知，然若目的外利用係以「當事人同意」為依據，則向當事人告知目的外利用相關資訊，係取得有效同意之前提。1999 熱線於現階段將個人資料用於人工智慧應用之訓練、測試、廠商概念驗證等，如構成目的外利用，有主張「增進公共利益所必要」（一般個人資料）或「執行法定職務必要範圍」（特種個人資料）之空間，此時，1999 熱線應無須將目的外利用之情事告知當事人，惟從增進個人資料利用透明度、保障當事人資料自主權之角度，1999 熱線亦可考量以網站公告等方式，適當向當事人為揭露。如於具體個案中，1999 熱線認較宜依據「當事人同意」將個人資料於蒐集之特定目的外利用，則需向當事人告知相關資訊，具體告知內容與方式，將於後文「當事人同意」小節具體討論。

6.6 此外，1999 熱線將個人資料提供予外部單位或廠商，如該提供行為構成資料之目的外利用，則資料之接收者屬自 1999 熱線間接蒐集該等資料，其告知義務及免為告知之事由，應依個資法第 9 條獨立判斷。

7. 當事人同意

個人資料保護法 第 7 條

- I. 第十五條第二款及第十九條第一項第五款所稱同意，指當事人經蒐集者告知本法所定應告知事項後，所為允許之意思表示。
- II. 第十六條第七款、第二十條第一項第六款所稱同意，指當事人經蒐集者明確告知特定目的外之其他利用目的、範圍及同意與否對其權益之影響後，單獨所為之意思表示。
- III. 公務機關或非公務機關明確告知當事人第八條第一項各款應告知事項時，當事人如未表示拒絕，並已提供其個人資料者，推定當事人已依第十五條第二款、第十九條第一項第五款之規定表示同意。
- IV. 蒐集者就本法所稱經當事人同意之事實，應負舉證責任。

個人資料保護法 第 6 條

- II. 依前項規定蒐集、處理或利用個人資料，準用第八條、第九條規定；其中前項第六款之書面同意，準用第七條第一項、第二項及第四項規定，並以書面為之。

7.1 北市府建設 1999 熱線，基於陳情及其他市政服務之提供、1999 熱線資訊系統管理等特定目的，於市政諮詢、電話轉接、受理陳情及派工等服務之提供及管理必要範圍內，以資訊系統接聽民眾自願進線電話，蒐集、處理及利用進線民眾及陳情所涉其他自然人之個人資料，應以個資法第 15 條第 1 款「執行法定職務必要範圍

內」為法律依據（如前文「一般個資的蒐集處理及利用」、「特種個資的蒐集處理及利用」兩小節所分析），而不涉及個資法第 15 條第 2 款之「當事人同意」。

7.2 如前文「一般個資的蒐集處理及利用」「特種個資的蒐集處理及利用」等小節所分析，1999 熱線導入人工智慧語音辨識功能後，因該功能主要作為 1999 熱線內部作業之輔助性措施，故其所涉之資料處理與利用，應仍屬「個人資料之蒐集係公務機關執行法定職務所必要」之範圍。以「人工智慧應用之訓練與改善」作為個人資料之蒐集目的之一，亦有主張屬「執行法定職務必要範圍內」之空間。而不涉及個資法第 15 條第 2 款之「當事人同意」。

7.3 然而，若 1999 熱線經衡酌判斷，認基於「人工智慧應用之訓練與改善」目的蒐集、處理及利用民眾個人資料，較宜以「當事人（書面）同意」為依據，則宜依第 8 條、第 9 條向當事人告知法定事項，以符個資法第 7 條關於知情同意之規定。且此時，1999 熱線就「人工智慧應用之訓練與改善」之目的依法踐行告知後，宜徵求當事人允許之意思表示（例如口頭說明「我同意」，或依 IVR 提示按鍵選擇「我同意」），依個資法第 7 條第 1 款取得當事人同意。考量 1999 熱線以提供市政服務為核心功能，民眾乃基於使用市政服務之需求進線並亦提供其個人資料，故 1999 熱線不宜援用

第 7 條第 3 款「推定同意」之規定，以避免捆綁同意之疑慮⁶⁸。

7.4 此外，如前文「一般個資的蒐集處理及利用」小節所分析，1999 熱線現階段將進線通話錄音、系統內進線內容紀錄委託外部廠商用於人工智慧概念驗證、訓練和測試，可能構成個人資料之目的外利用，且有主張個資法第 16 條第 2 款「增進公共利益所必要」之空間。惟於具體個案中，若 1999 熱線認較宜以個資法第 16 條第 7 款之「當事人同意」作為目的外利用之依據，則需依個資法第 7 條第 2 款，在將該等資料提用於驗證、訓練和測試前，明確向個人資料當事人告知該委託利用之目的、範圍，以及同意與否對其權益之影響，並取得個人資料當事人對該等目的外利用之單獨同意。

8. 當事人權利

個人資料保護法 第 3 條

當事人就其個人資料依本法規定行使之下列權利，不得預先拋棄或以特約限制之：

- 一、查詢或請求閱覽。
- 二、請求製給複製本。
- 三、請求補充或更正。
- 四、請求停止蒐集、處理或利用。

⁶⁸ 國家發展委員會 107 年 11 月 21 日發法字第 1072002136 號函釋參照。

五、請求刪除。

個人資料保護法 第 10 條

公務機關或非公務機關應依當事人之請求，就其蒐集之個人資料，答覆查詢、提供閱覽或製給複製本。但有下列情形之一者，不在此限：

- 一、妨害國家安全、外交及軍事機密、整體經濟利益或其他國家重大利益。
- 二、妨害公務機關執行法定職務。
- 三、妨害該蒐集機關或第三人之重大利益。

個人資料保護法 第 11 條

- I. 公務機關或非公務機關應維護個人資料之正確，並應主動或依當事人之請求更正或補充之。
- II. 個人資料正確性有爭議者，應主動或依當事人之請求停止處理或利用。但因執行職務或業務所必須，或經當事人書面同意，並經註明其爭議者，不在此限。
- III. 個人資料蒐集之特定目的消失或期限屆滿時，應主動或依當事人之請求，刪除、停止處理或利用該個人資料。但因執行職務或業務所必須或經當事人書面同意者，不在此限。
- IV. 違反本法規定蒐集、處理或利用個人資料者，應主動或依當事人之請求，刪除、停止蒐集、處理或利用該個人資料。
- V. 因可歸責於公務機關或非公務機關之事由，未為更正或補充之個人資料，應於更正或補充後，通知曾提供利用之對象。

- 8.1 1999 熱線以電話方式提供服務。依各項標準作業流程，進線民眾之陳情或派工需求經系統立案後，可再次撥打 1999 熱線，提供案件編號和查詢密碼，由話務人員協助查詢後告知處理進度。1999 熱線僅接受同時提供案件編號和查詢密碼進行查詢，如查詢民眾僅提供姓名、電話、案件關鍵字等，則話務人員無法協助查詢。此外，若民眾再次索取案件編號或密碼，僅可由話務人員透過「系統重發簡訊」方式提供，且民眾須提供反映案件的時間、地點、內容、當時所留陳情人全名及聯繫手機以供核對，核對通過後方可由系統再次發送包含案件編號和查詢密碼之簡訊。1999 熱線僅以前述方式提供案件處理進度、案件編號/密碼之查詢，而不提供陳情人姓名、電話等身分識別資料之查詢。
- 8.2 進線民眾反映之需求於系統立案後，1999 熱線即不接受民眾再行來電變更案件所涉陳情人姓名、電話、電子郵件信箱等資料。如陳情民眾再次進線，反映已立案之陳情內容有遺漏或錯誤，則由話務人員以「承前案」方式再行立案，以補充或更正前案。
- 8.3 除上述查詢、補充、更正資料陳情案件之方式外，1999 熱線亦於 1999 市民當家熱線服務網「FAQ 常見問答」中說明⁶⁹，民眾得以

⁶⁹ 民眾是否可申請調閱 1999 錄音檔案？處理時間為多久？處理結果如何取得？，
https://www.gov.taipei/News_Content.aspx?n=EEC70A4186D4C828&sms=87415A8B9CE81B16&s=82DA34A1B60CF4E0。

書面申請並檢附身分證明之方式，依政府資訊公開法向北市府研考會話務管理組申請調閱 1999 熱線通話錄音。

8.4 經檢視 1999 市民當家熱線服務網之隱私權及資訊安全政策、FAQ

常見問答等，未見關於當事人依個資法對行使權利之方式有明確說明，且 1999 熱線各項標準作業流程中，亦未見接獲當事人權利行使請求時之具體作業規範。然保有 1999 熱線通話錄音檔之北市府研考會，以及保有 1999 熱線話務紀錄、陳情系統及派工系統內所含資料之資訊局已分別訂定「臺北市政府研究發展考核委員會個人資料保護管理要點」和「臺北市政府資訊局個人資料保護管理要點」，此兩項規範皆規定當事人權利之行使程序，包括身分確認、駁回權利行使請求之事由、查詢閱覽之收費、提供查詢時對他人個人資料之保護等⁷⁰。北市府研考會並於其網站「個人資料保護專區」列明「圖資管理組」為其受理當事人權利行使請求之窗口，並提供聯絡電話。

8.5 綜合以上觀察，1999 熱線應已提供當事人依個資法行使權利之管道。考量電話方式提供服務之技術特性，1999 熱線依當事人請求內容而區分不同權利行使方式（於通話中直接提供查詢、以書面申請查詢等），應屬合理。1999 熱線對於系統內已立案之案件，

⁷⁰ 臺北市政府研究發展考核委員會個人資料保護管理要點，第 29 點至第 31 點；臺北市政府資訊局個人資料保護管理要點，第 27 點至第 29 點。

即不允許變更姓名、電話、電子郵件信箱等資料，係出於避免案件資料洩露之考量，若於此等資料登載確有錯誤之情形，可採「承前案」方式再行立案予以更正補充，則應合於個資法第 11 條第 1 項維護個人資料正確之要求。然除查詢、補充、更正資料陳情案件之權利行使請求可由話務人員於通話中即時受理外，1999 熱線似無統一之受理管道，而是依當事人所請求之內容（研考會保有之通話錄音或資訊局保有之系統內其他紀錄）與依據（個資法或政府資訊公開法），設定不同受理機關，恐有增加民眾行使當事人權利負擔之虞。

8.6 1999 熱線導入人工智慧語音辨識功能後，因該功能係協助提升服務效率之內部輔助措施，1999 熱線對外服務流程無需調整，故當事人權利行使方式應可繼續沿用。惟回應當事人權利行使請求時，應將該功能亦納入考慮範圍。例如，若系統內留存該功能轉譯之當事人通話內容文字稿，則當事人就系統內紀錄請求製給複製本時，該文字稿應一併提供。然若未來導入簡單回覆進線提問等更具獨立性之人工智慧應用，1999 熱線之對外服務流程將與現行流程不同，當事人權利行使方式亦可能有調整之必要。

8.7 1999 熱線若以「人工智慧之訓練與改善」作為個人資料蒐集、處理及利用的目的之一，個資法第 15 條第 1 款之「執行法定職務必

要範圍內」為法律依據，並遵循個資法之其他規範，則該蒐集目的存續期間，當事人將難以依個資法第 11 條第 3 項、第 4 項之規範，拒絕將其個人資料用於人工智慧之訓練與改善。惟依憲法法庭之見解，因符合特定要件而允許未獲當事人同意而經蒐集、處理及利用之個資，當事人仍具事後控制權，且其事後控制權之內涵包括請求刪除、停止利用或限制利用個資之權利⁷¹。1999 熱線基於執行法定職務之必要，依法不經當事人同意，為人工智慧之訓練與改善目的蒐集、處理及利用進線人通話錄音等個人資料，與 1999 熱線服務之核心內容尚有差距，且進線人出於使用市政服務之需求，較難事前自主決定是否提供資料，故 1999 熱線宜依前述憲法法庭判決之意旨，提供當事人請求停止為「人工智慧之訓練與改善」目的處理和利用其個人資料之管道與程序。

- 8.8 若 1999 熱線經衡酌判斷，認基於「人工智慧應用之訓練與改善」目的蒐集、處理及利用民眾個人資料，較宜以「當事人（書面）同意」為依據，則需依個資法第 7 條取得當事人知情同意，業如前文「一般/特種個資的蒐集處理及利用」、「當事人同意」等小節所述。我國個資法雖未明文規範當事人「撤回同意」之權利，然實務中，個資法主管機關承認當事人撤回同意之可能性，並認為

⁷¹ 憲法法庭 111 年憲判字第 13 號判決，第 32 段，第 69~71 段。

基於當事人同意而合法蒐集、處理之個人資料，「若當事人事後撤回其同意，則自其撤回時起，如蒐集個人資料之特定目的或要件已不存在」，則應依個資法第 11 條第 3 項主動或依當事人請求，刪除、停止處理或利用該等個人資料⁷²。據此，如 1999 熱線基於當事人同意將個人資料用於人工智慧應用之訓練與改善，則於當事人主張撤回同意並據此請求刪除或停止處理利用個人資料時，需依個資法第 11 條第 3 項回應其請求。

9. 資料保存期限

個人資料保護法 第 11 條

III. 個人資料蒐集之特定目的消失或期限屆滿時，應主動或依當事人之請求，刪除、停止處理或利用該個人資料。但因執行職務或業務所必須或經當事人書面同意者，不在此限。

9.1 1999 熱線於當次服務期間，於系統內保存民眾當次進線所涉之個人資料，應屬資料蒐集特定目的之範圍內，符合個資法第 11 條第 3 項之要求。

9.2 服務結束後，服務紀錄於系統內繼續保存合理期間，以滿足進線民眾查閱處理結果、機關評估服務成效等需求，亦屬資料蒐集特定目的之必要範圍，符合個資法第 11 條第 3 項之要求。

⁷² 法務部 106 年 11 月 10 日法律字第 10603512680 號函釋參照。

9.3 1999 熱線相關個人資料分別由北市府研考會、資訊局保有。查此兩機關分別訂定之個人資料保護管理要點，皆要求定期查明個人資料依法令之保存期間，要求於法令無明文規定之情形下，視執行業務必要性及合理性，「於年度檔案分及保存年限表明定，或訂定經告知當事人之合理保存期間」。該等管理要點並規範蒐集之特定目的消失或期限屆滿時個人資料之處置程序，且對資料之刪除作業訂定標準⁷³。據此，1999 熱線之管理與維護機關應已就個人資料之保存期限訂定規則，以符合個資法第 11 條第 3 項之要求。

9.4 1999 熱線導入人工智慧語音辨識功能後，因該功能係協助提升服務效率之內部輔助措施，所產出之當事人通話內容文字稿等，應屬 1999 熱線市政服務所涉個人資料之一部，依前述保存期限加以保存。

9.5 若 1999 熱線係基於「當事人同意」而將個人資料用於「人工智慧應用之訓練與改善」，宜依個資法主管機關之見解，認可當事人撤回同意之可能性，且當事人撤回同意時，需依個資法第 11 條第 3 項回應當事人請求，業如前文「當事人權利」一章所分析。據此，1999 熱線如接獲當事人撤回同意之請求，宜依個資法判斷是否應

⁷³ 臺北市政府研究發展考核委員會個人資料保護管理要點，第 21 點、第 24 點，臺北市政府資訊局個人資料保護管理要點，第 18 點、第 21 點。

刪除、停止處理或利用相關個人資料。

10. 個人資料的安全維護

個人資料保護法 第 18 條

公務機關保有個人資料檔案者，應指定專人辦理安全維護事項，防止個人資料被竊取、竄改、毀損、滅失或洩漏。

個人資料保護法施行細則 第 12 條

- I. 本法第六條第一項但書第二款及第五款所稱適當安全維護措施、第十八條所稱安全維護事項、第十九條第一項第二款及第二十七條第一項所稱適當之安全措施，指公務機關或非公務機關為防止個人資料被竊取、竄改、毀損、滅失或洩漏，採取技術上及組織上之措施。
- II. 前項措施，得包括下列事項，並以與所欲達成之個人資料保護目的間，具有適當比例為原則：
 - 一、配置管理之人員及相當資源。
 - 二、界定個人資料之範圍。
 - 三、個人資料之風險評估及管理機制。
 - 四、事故之預防、通報及應變機制。
 - 五、個人資料蒐集、處理及利用之內部管理程序。
 - 六、資料安全管理及人員管理。
 - 七、認知宣導及教育訓練。
 - 八、設備安全管理。

九、資料安全稽核機制。

十、使用紀錄、軌跡資料及證據保存。

十一、個人資料安全維護之整體持續改善。

10.1 北市府就 1999 熱線訂有「資訊安全管理標準作業流程」，規定

1999 熱線服務提供及話務中心運作過程中個人資料保護及資訊

安全防護之各項作業規範，明列人員教育訓練、實體及環境安全、

作業安全、系統安全四大面向之具體措施，並就安全事故之因應

訂定標準程序，符合個資法施行細則第 12 條之要求。

10.2 1999 熱線相關個人資料分別由北市府研考會、資訊局保有。查

此兩機關分別訂定之個人資料保護管理要點，皆設立機關內個人

資料保護之專責組織⁷⁴，並規定個人資料定期盤點、風險評估、紀

錄或證據保存、資料處理和利用流程監測、定期教育訓練、年度

稽核、個資事件因應與通知通報等機制⁷⁵，符合個資法第 18 條及

施行細則第 12 條之要求。

10.3 1999 熱線導入人工智慧語音辨識功能後，因該功能係在現有資

訊系統內部增設，並不改變資訊系統整體之架構與作業模式，對

1999 熱線之系統安全性較無影響，故前述安全維護措施應可繼續

⁷⁴ 臺北市政府研究發展考核委員會個人資料保護管理要點，第 4 點（設立個人資料保護管理執行小組），臺北市政府資訊局個人資料保護管理要點，第 3 點（設立資訊與個資安全管理委員會）。

⁷⁵ 臺北市政府研究發展考核委員會個人資料保護管理要點，第 36~45 點，臺北市政府資訊局個人資料保護管理要點，第 34~43 點。

沿用，惟安全維護措施之落實，宜就該功能作具體考量。然若未來導入之人工智慧應用將影響系統安全性，例如雲端服務方式提供語音辨識等，則前述安全維護措施則可能有調整之必要。

10.4 1999 熱線於人工智慧語音辨識應用開發過程中，將進線通話錄音、系統內進線內容紀錄委託外部廠商用於人工智慧訓練和測試，將在一定程度上升高資料安全性風險，應就該等利用強化資料安全維護措施。在目前已執行之小範圍內部測試（即概念驗證）中，北市府除與廠商及廠商人員簽訂資料保護或保密協議外，並要求廠商僅限於話務中心內部、北市府工作人員陪同下處理及利用該等個人資料，應認屬強化資料安全維護之適足措施。又北市府委託廠商處理和利用該資料，亦涉廠商於受託處理利用期間，依北市府監督而採取安全維護措施，詳如後文「委外監督」小節所分析。

11. 委外監督

個人資料保護法 第 4 條

受公務機關或非公務機關委託蒐集、處理或利用個人資料者，於本法適用範圍內，視同委託機關。

個人資料保護法施行細則 第 8 條

- I. 委託他人蒐集、處理或利用個人資料時，委託機關應對受託者為適當之監督。
- II. 前項監督至少應包含下列事項：
 - 一、預定蒐集、處理或利用個人資料之範圍、類別、特定目的及其期間。
 - 二、受託者就第十二條第二項採取之措施。
 - 三、有複委託者，其約定之受託者。
 - 四、受託者或其受僱人違反本法、其他個人資料保護法律或其法規命令時，應向委託機關通知之事項及採行之補救措施。
 - 五、委託機關如對受託者有保留指示者，其保留指示之事項。
 - 六、委託關係終止或解除時，個人資料載體之返還，及受託者履行委託契約以儲存方式而持有之個人資料之刪除。
- III. 第一項之監督，委託機關應定期確認受託者執行之狀況，並將確認結果記錄之。
- IV. 受託者僅得於委託機關指示之範圍內，蒐集、處理或利用個人資料。受託者認委託機關之指示有違反本法、其他個人資料保護法律或其法規命令者，應立即通知委託機關。

11.1 1999 熱線以整體話務服務委外方式營運，並由北市府研考會督導管理，故話務中心的話務人員於提供 1999 熱線服務範圍內，皆屬受北市府研考會委託而蒐集、處理及利用個人資料。依個資法第 4 條規定，廠商視同北市府研考會，即廠商之行為將視同北市府研考會之行為，若廠商違反個資法規定致民眾之個人資料受到

侵害，將由北市府研考會承擔責任。

11.2 北市府就 1999 熱線訂之「資訊安全管理標準作業流程」，內含資料之利用與攜出限制、人員異動時權限取消及資料刪除等規範等，直接適用於話務中心話務人員，並設有遵循監督及懲戒規範。此外，北市府與話務中心委外廠商之採購契約並訂有個資保護約定及外洩罰則，北市府亦與話務中心工作人員個人間亦簽訂保密協議，應認已依個資法施行細則第 8 條規定，對廠商為適當之監督。

11.3 1999 熱線導入人工智慧語音辨識功能後，因該功能係資訊系統內增設之話務人員輔助工具，並未實質改變現有受託者（話務人員）蒐集處理利用資料之範圍，亦未引入其他受委託蒐集、處理或利用個人資料者，故前述委外監督措施應可繼續沿用，惟其執行宜就該功能作具體考量。然若未來導入之人工智慧應用將引入其他受託者或實質變更現有委託範圍，例如因導入雲端作業而委託雲端廠商儲存資料等，則前述委外監督措施則可能有調整之必要。

11.4 北市府擬導入外部廠商開發及導入人工智慧應用，且將因此委由廠商將錄音、系統內進線內容紀錄用於人工智慧概念驗證、訓練和測試。就 1999 熱線現階段規劃中之人工智慧應用導入方案⁷⁶

⁷⁶ 例如 1999 導入語音辨識評估專案，<https://smartcity.taipei/projdetail/275>；北市府 1999 心聲分析規劃專案，<https://smartcity.taipei/projdetail/288>。

而言，由公開資料及北市府與廠商間契約觀察，現行規劃方案係在廠商已初步開發之人工智慧應用模型基礎上，依 1999 熱線進線之常見內容與通話方式加以調適，最終產出符合 1999 熱線需求之人工智慧應用。外部廠商可處理和利用之 1999 熱線相關個人資料之範圍，以及處理和利用之目的、方法等，皆由北市府確定；廠商開發過程中，須在北市府提供之受控環境中，依北市府人員之監督處理利用該等個人資料；且開發結束後，廠商不得保存所接收和利用之個人資料。據此，外部廠商係非受北市府委託而處理及利用與 1999 熱線相關之個人資料。外部廠商雖可於人工智慧訓練和測試過程中，依其專業知識決定訓練過程中處理和利用個人資料之具體技術方法，然其技術方法之最終目的，仍係滿足北市府之指示，故廠商並非獨立蒐集、處理及利用 1999 熱線相關個人資料，併予敘明。

- 11.5 北市府於目前已執行之小範圍內部測試中，除與廠商及廠商人員簽訂資料保護或保密協議外，並要求廠商僅限於話務中心內部、北市府工作人員陪同下處理及利用該等個人資料，應認已於個案中對廠商為必要、適當之監督。如未來北市府繼續委託廠商為人工智慧應用訓練和測試目的而蒐集、處理或利用個人資料，亦應依委託之內容與範圍，依個資法施行細則第 8 條規定，對受託廠

商為適當之監督，包含事前的契約約定權利義務、事中確認（稽核）廠商遵守契約或各項法遵與資安義務要求，以及事後（契約終止後）的資料返還或刪除。

四、資安技術風險識別

資訊安全風險並非本次評估之重點事項，本報告僅基於資料遭竊取、洩漏、竄改或其他侵害之視角，對可立即識別之資安技術風險為簡要評估。

1. 資料儲存風險

1.1 1999 熱線相關個人資料集中儲存於話務系統、陳情系統與派工系統。由北市府資訊局主責管理與維護。北市府就其所屬資訊系統之安全防護，訂有臺北市政府資通安全管理規定、臺北市政府資通系統安全作業指引、臺北市政府資通訊資產及電子資料安全作業指引等規範。1999 熱線各項資訊系統亦須遵守該等規範之資安防護要求。

1.2 1999 熱線未來導入人工智慧應用後，如涉及雲端儲存等委外資料儲存方式，將須遵循臺北市政府資通訊業務委外作業指引規範，包括要求業者提供資安國際認證證明⁷⁷。

⁷⁷ 臺北市政府資通訊業務委外作業指引，第 9 點。

2. 資料存取風險

- 2.1 為避免 1999 熱線相關個人資料遭話務人員不當存取，1999 熱線資訊安全管理標準作業流程對話務人員行為訂有嚴格控管規範，包括話務人員僅能使用話務中心傳統桌機登入各項系統；非經北市府研考會許可，不得將存取系統內之個人資料、錄音檔案；話務中心內禁止攜入個人電子通訊設備且禁止拍照等。1999 熱線各項服務作業流程亦包含個人資料保護之規範，例如留電回覆標準作業流程要求話務人員將「留下民眾電話或個資的紙本及電子檔資料皆需確實銷毀或使用檔案加密」，以保護資料不致外洩。
- 2.2 為避免人工智慧委外廠商不當處理或利用 1999 熱線相關個人資料，1999 熱線於目前已執行之小範圍內部測試中，要求廠商親至北市府控管之實體環境，由北市府人員將錄音檔案以儲存媒介交付廠商，於現場使用廠商軟硬體設備執行人工智慧分析，事後歸還錄音檔案並不得下載攜出。
- 2.3 1999 熱線未來人工智慧應用之後續開發過程中及導入人工智慧應用後，如因此新增其他受委託蒐集、處理或利用個人資料之廠商，宜參照前述控管密度，並遵守臺北市政府資通訊業務委外作業指引，管理委外廠商之資料存取行為。

3. 身分冒用風險

3.1 1999 熱線以電話方式提供服務，民眾無須註冊帳號，進線即可實現陳情或派工案件立案、查詢等，且服務過程以姓名與手機號碼作為主要身分確認方式，故存在身分冒用之風險。

3.2 為降低冒名查詢案件之風險，陳情與派工程序中，如進線人擬查詢已立案之案件，1999 熱線僅接受以案件編號與查詢密碼查詢進度，不接受以姓名、電話、關鍵字等資訊查詢；且查詢內容僅限案件之處理進度，而不包括陳情人姓名、電話等資料。若陳情人遺失案件編號或查詢密碼，則須提供陳情當時提供之姓名、聯絡電話、陳情內容供系統比對，比對通過方可由系統再次載有發送案件編號和查詢密碼之簡訊。

3.3 為降低冒名變更案件之風險，陳情與派工程序中，如進線人擬撤銷已立案之案件，則須提供同時提供案件編號、查詢密碼、陳情時所留姓名和聯絡電話，並說明撤銷原因。

3.4 1999 熱線未來導入人工智慧應用後，如該人工智慧應用將與進線人直接互動，視人工智慧應用之服務內容，可能調整或補充身分冒用防範措施。

五、剩餘風險及因應建議

1. 法規遵循剩餘風險

1.1 經研究團隊檢視，1999 熱線於「市政熱線服務之提供與管理」與

「市政服務品質提升與優化」之目的範圍內，以當前資訊系統接聽民眾進線，提供諮詢、陳情等各項服務，並據此提升市政服務品質，所涉之個人資料蒐集、處理及利用，原則上均與我國個資法規定相符。惟除查詢、補充、更正資料陳情案件之權利行使請求可由話務人員於通話中即時受理外，1999 熱線似無統一之受理管道，而是依當事人所請求之內容與依據，設定不同受理機關（研考會或資訊局），恐有增加民眾行使當事人權利負擔之虞。

1.2 如 1999 熱線導入人工智慧語音辨識應用，因該應用係協助話務

人員提升服務效率之內部輔助工具，1999 熱線資料蒐集、處理和利用流程應無實質變化，故其適法性評價應維持相同。然如所導入之人工智慧應用將實質變更 1999 熱線個人資料流程（例如以雲端服務方式提供服務、由人工智慧應用直接與進線人互動等），則個人資料之安全維護、委外監督、當事人權利行使方式等面向將須另行評價。

1.3 若 1999 熱線以「人工智慧應用之訓練與改善」作為個人資料的蒐

集目的之一，應有主張個資法第 15 條第 1 款、第 6 條第 1 項第 2 款「執行法定職務必要範圍內」之空間，並得因此援用同法第 8 條第 2 項第 2 款規定之「公務機關執行法定職務所必要」而免為向當事人踐行告知。然因該等目的與社會通念下政府職務之執行較有落差，且較難為當事人事前預見，故從保障當事人對其個人資料之知情與自主控制之視角，1999 熱線仍宜考量於服務流程中納入個人資料蒐集之告知，包括資料將用於「人工智慧應用之訓練與改善」目的、所涉資料類別、利用之期間、地區、對象和方式，以及當事人得行使之權利等。考量熱線服務通話時長等因素之限制，1999 熱線可採用分層告知方式，於 IVR 系統踐行第一層告知，敘明 1999 熱線將蒐集、處理及利用進線人之個人資料，並請進線人參閱 1999 熱線官方網站取得更多資訊；於 1999 熱線官方網站踐行第二層告知，完整揭露個人資料保護法要求之告知事項。1999 熱線並宜依憲法法庭判決意旨，提供當事人請求停止為「人工智慧之訓練與改善」目的處理和利用其個人資料之管道與程序，以保障當事人對其個人資料之事後控制權。

- 1.4 如 1999 熱線經衡酌判斷，認為「人工智慧應用之訓練與改善」目的蒐集、處理及利用進線人通話錄音等個人資料，較宜以「當事人（書面）同意」為法律依據，則需依個資法第 8 條、第 9 條向

當事人告知法定事項，並依同法第 7 條第 1 款徵求當事人允許之積極意思表示，取得當事人知情同意。

1.5 1999 熱線人工智慧應用之規劃、開發、導入與執行階段，如委由外部廠商將現行已蒐集之進線通話錄音、系統內紀錄等個人資料用於人工智慧應用之概念驗證、訓練、測試等，視資料原始蒐集目的之內容與範圍，可能屬蒐集之特定目的外利用，並應有主張以個資法第 16 條第 2 款之「增進公共利益所必要」（一般個人資料）或第 6 條第 1 項第 2 款「執行法定職務必要範圍」（特種個人資料）之空間。此時，依個資法之現行規範，1999 熱線應無須將目的外利用之情事告知予當事人。惟從增進個人資料利用透明度、保障當事人資料自主權之立場，1999 熱線亦可考量以網站公告等方式，適當向當事人揭露資料目的外利用於人工智慧應用訓練與測試之範圍、資料保護措施等。又此情形亦涉及委託廠商處理及利用個人資料，依個資法第 4 條規定，廠商之行為視為北市府之行為，北市府並須依個資法施行細則第 8 條對廠商為必要、適當之監督，並應考量訓練過程的個人資料安全維護事項（個資法第 18 條，個資法施行細則第 12 條）。

1.6 然如於具體個案中，1999 熱線認較宜依據「當事人（書面）同意」將個人資料於蒐集之特定目的外，用於「人工智慧應用之訓練與

改善」，則就一般個人資料而言，需依個資法第 7 條第 2 項，明確向個人資料當事人告知提供予廠商之利用目的、範圍，以及同意與否對其權益之影響，並取得個人資料當事人對該等目的外利用之單獨同意；就特種個人資料而言，需依個資法第 6 條第 1 項第 6 款取得當事人書面同意。

2. 資安技術剩餘風險

2.1 1999 熱線服務仍有冒用身分查詢案件之風險，例如案件編號、查詢密碼被他人知悉等。如發生冒用身分查詢案件，因 1999 熱線僅可查詢案件處理進度，不會提供陳情人個人資料，故對陳情人之侵害較低，應屬可接受之風險。

2.2 1999 熱線服務雖有冒用身分撤銷案件之風險。如發生冒用身分撤銷案件，則陳情人之案件將無法獲得有效處理，將對陳情人權益之侵害程度中等。然陳情人通常亦會自行追蹤陳情進度，冒名撤案之情事有較高機會在較短時間內被發覺，陳情人可透過再次陳情方式補救，故陳情人權益損害之發生可能性較低，應屬可接受之風險。1999 熱線亦可於撤案後，向陳情人寄送撤案通知簡訊，以降低冒名撤案造成陳情人權益損害之機率。

2.3 1999 熱線儲存之個人資料雖訂有較嚴密之資料安全維護措施，然

仍有遭外部惡意入侵或內部人員（含委外廠商人員）不法或不當存取之風險，此風險應透過北市府資安制度之有效落實及持續改善而控管。

2.4 1999 熱線因人工智慧應用之規劃、開發、導入與執行，所面臨之資安技術風險可能升高。此風險應透過北市府資安制度之有效落實以及與廠商間契約之履約管理而控管。

第三節 臺北市民當家熱線導入人工智慧應用評估小結

經本研究檢視及評估，1999 熱線於「市政熱線服務之提供與管理」與「市政服務品質提升與優化」之目的範圍內，以當前資訊系統接聽民眾進線，提供諮詢、陳情等各項服務，並據此提升市政服務品質，所涉之個人資料蒐集、處理及利用，原則上均與我國個資法規定相符。

如 1999 熱線導入人工智慧語音辨識應用，因該應用係協助話務人員提升服務效率之內部輔助工具，1999 熱線資料蒐集、處理和利用流程應無實質變化，故其適法性評價應維持相同。然如所導入之人工智慧應用將實質變更 1999 熱線個人資料流程（例如以雲端服務方式提供服務、由人工智慧應用直接與進線人互動等），則個人資料之安全維護、委外監督、當事人權利行使方式等面向須另行評價。

若 1999 熱線以「人工智慧應用之訓練與改善」作為個人資料的蒐集目的之一，應有主張個資法第 15 條第 1 款、第 6 條第 1 項第 2 款「執行法定職務必要範圍內」之空間，並得因此援用同法第 8 條第 2 項第 2 款規定之「公務機關執行法定職務所必要」而免為向當事人踐行告知。然因該等目的與社會通念下政府職務之執行較有落差，且較難為當事人事前預見，故從保障當事人對其個人資料之知情與自主控制之視角，1999 熱線仍宜考量於服務流程中納入個人資料蒐集之

告知，包括資料將用於「人工智慧應用之訓練與改善」目的、所涉資料類別、利用之期間、地區、對象和方式，以及當事人得行使之權利等。考量熱線服務通話時長等因素之限制，1999 熱線可採用分層告知方式，於 IVR 系統踐行第一層告知，敘明 1999 熱線將蒐集、處理及利用進線人之個人資料，並請進線人參閱 1999 熱線官方網站取得更多資訊；於 1999 熱線官方網站踐行第二層告知，完整揭露個人資料保護法要求之告知事項。1999 熱線並宜依憲法法庭判決意旨，提供當事人請求停止為「人工智慧之訓練與改善」目的處理和利用其個人資料之管道與程序，以保障當事人對其個人資料之事後控制權。

如 1999 熱線經衡酌判斷，認為「人工智慧應用之訓練與改善」目的蒐集、處理及利用進線人通話錄音等個人資料，較宜以「當事人（書面）同意」為法律依據，則需依個資法第 8 條、第 9 條向當事人告知法定事項，並依同法第 7 條第 1 款徵求當事人允許之積極意思表示，取得當事人知情同意。

1999 熱線人工智慧應用之規劃、開發、導入與執行階段，如將現行已蒐集之進線通話錄音、系統內紀錄等個人資料提供予外部廠商用於人工智慧應用之概念驗證、訓練、測試等，視資料原始蒐集目的之內容與範圍，可能屬蒐集之特定目的外利用，並應有主張以個資法第 16 條第 2 款之「增進公共利益所必要」（一般個人資料）或第 6 條第

1 項第 2 款「執行法定職務必要範圍」(特種個人資料)之空間。此時，依個資法之現行規範，1999 熱線應無須將目的外利用之情事告知予當事人。惟從增進個人資料利用透明度、保障當事人資料自主權之立場，1999 熱線亦可考量以網站公告等方式，適當向當事人揭露資料目的外利用於人工智慧應用訓練與測試之範圍、資料保護措施等。又此情形亦涉及委託廠商處理及利用個人資料，依個資法第 4 條規定，廠商之行為視為北市府之行為，北市府並須依個資法施行細則第 8 條對廠商為必要、適當之監督，並應考量訓練過程的個人資料安全維護事項(個資法第 18 條，個資法施行細則第 12 條)。

然如於具體個案中，1999 熱線認為較宜依據「當事人(書面)同意」將個人資料於蒐集之特定目的外，用於「人工智慧應用之訓練與改善」，則就一般個人資料而言，需依個資法第 7 條第 2 項，明確向個人資料當事人告知提供予廠商之利用目的、範圍，以及同意與否對其權益之影響，並取得個人資料當事人對該等目的外利用之單獨同意；就特種個人資料而言，需依個資法第 6 條第 1 項第 6 款取得當事人書面同意。

1999 熱線儲存之個人資料雖訂有較嚴密之資料安全維護措施，然仍有遭外部惡意入侵或內部人員(含委外廠商人員)不法或不當存取之風險，此風險應透過北市府資安制度之有效落實及持續改善而控管。

1999 熱線因人工智慧應用之規劃、開發、導入與執行，所面臨之資安技術風險可能升高。此風險應透過北市府資安制度之有效落實以及與廠商間契約之履約管理而控管。

第六章 指引草案與焦點座談

第一節 臺北市政府開發及應用人工智慧作業指引草案之提出

本研究依據對歐盟、英國、美國、日本、新加坡與香港為法規、文獻之研析，參酌國際實務之見，就制度面之流程及注意事項，提出「臺北市政府開發及應用人工智慧作業指引」草案。

本草案正文包括前言、基本原則與法規框架、管理制度與組織、影響評估與人為監督、透明性與個資當事人權利、安全維護、實務管理、採購與委外管理、其他注意事項九個章節，向北市府各機關說明人工智慧系統可能涉及之法令遵循及當事人權益影響，提示構建人工智慧系統開發及應用制度、管理人工智慧系統開發及應用全流程時之重點考量事項，並彙整國際上相關之最佳實務做法，供各機關參考遵循。

本研究並提出「人工智慧系統開發及應用檢核表」，作為本草案之附錄。檢核表將指引草案之正文內容轉化為具體評估項目，供北市府各機關於開發及應用具體人工智慧系統時，以較直觀之方式評量及檢核該系統法令遵循及民眾權益保障狀況。

第二節 焦點座談

本研究於提出前述指引草案後，於 111 年 11 月 29 日邀集相關產

官學研代表參與焦點座談，針對本研究提出之個資影響評估報告、前述指引草案與建議方向，就 1999 臺北市民當家熱線導入人工智慧應用之適法、倫理等疑義，就法規面、流程面與制度面提供意見。

一、焦點座談議程

1999 臺北市民當家熱線數位轉型與個資保護對策之研究委託案 焦點座談		
日期及時間	主持人	地點
111 年 11 月 29 日(二) 下午 14:00-17:00	計畫主持人 葉奇鑫 律師	臺北市政府市政大樓 1201 會議室
時間	議程	
10 分鐘	1. 主席/長官致辭	
20 分鐘	2. 研究團隊報告	
80 分鐘	3. 與會者討論	
10 分鐘	4. 其他對本案之建議	
討論主題		
1. 人工智慧系統開發及應用之基本原則與管理制度		
2. 人工智慧系統之影響評估與人為監督		
3. 人工智慧系統之實務運作與管理		
4. 人工智慧系統之透明性與當事人權利		
5. 人工智慧系統之安全維護與委外監督		
6. 其他意見與建議		

座談專家名單		
產業別	受訪者代號	任職機構
產	專家 A	美斯潔股份有限公司
官	專家 B	國家發展委員會法制協調中心（個人名義）
學	專家 C	東吳大學法學院 東吳大學人工智慧應用研究中心
研	專家 D	工研院技術移轉與法律中心
	專家 E	資策會科法所

二、焦點座談討論議題大綱

1999 臺北市民當家熱線數位轉型與個資保護對策之研究委託案

焦點座談討論議題大綱

1. 人工智慧系統開發及應用之基本原則與管理制度

- 1.1 指引草案所列之人工智慧系統開發及應用各項基本原則，其範圍與內容以何者為宜？
- 1.2 指引草案所列之人工智慧系統開發及應用法規框架，其範圍與內容以何者為宜？
- 1.3 機關對於人工智慧系統之開發及應用，於整體層面和個案層面，分別宜如何確立管理制度？
- 1.4 機關對於人工智慧系統之開發及應用，宜如何建立管理組織？

2. 人工智慧系統之影響評估與人為監督

2.1 人工智慧系統之事前影響評估，其步驟、風險考量要素和風險因應方法，以何者為宜？

2.2 人工智慧系統宜如何規劃與評估其人為監督機制？

3. 人工智慧系統之實務運作與管理

3.1 人工智慧系統之方案設計階段，機關宜如何納入行政倫理與個資保護之考量與設計？

3.2 機關為人工智慧系統之開發與應用準備資料時，宜採取何種個資保護作為？

3.3 機關於人工智慧演算法訓練和測試階段，宜採取何種作為防範對當事人權益之風險？

3.4 機關於人工智慧演算法部署、使用與監測階段，宜採取何種作為防範對當事人權益之風險？

4. 人工智慧系統之透明性與當事人權利

4.1 機關宜採取何種作為，落實人工智慧系統之透明性和可解釋性，促進民眾對人工智慧系統開發及應用之知情與瞭解？

4.2 機關宜採取何種作為，實現當事人能夠針對人工智慧系統行使其個資法上之各項權利？

5. 人工智慧系統之安全維護與委外監督

5.1 機關宜採取何種作為，強化人工智慧系統之安全維護？

5.2 機關如自外部廠商採購人工智慧系統，或委託外部廠商開發或運作人工智慧系統，宜對廠商採取何種管理措施？

6. 其他意見與建議

- 6.1 指引草案對人工智慧系統及其行政倫理與個資保護影響之簡介，其範圍與內容以何者為宜？
- 6.2 為落實尊重當事人權益、提升人工智慧系統可信賴性，指引草案宜納入其他何種優良實務做法？
- 6.3 指引草案所列「人工智慧系統開發及應用檢核表」，其內容與範圍以何者為宜？
- 6.4 座談專家是否有其他意見或建議？

三、焦點座談家意見彙整紀錄

1999 臺北市民當家熱線數位轉型與個資保護對策之研究委託案 專家意見彙整紀錄

(以下按專家發言順序臚列)

專家 B

- 1、現行關於 AI 議題本來即無專法，應該要統合適用法規。演算法和演算能力比較偏技術面，重點是資料應用，涉及個人資料。參考歐盟、歐洲執委會的 AI 法規架構，在還沒有特別法之前，都回到個資法。
- 2、指引對市政府的需求應該是先考慮情境，依照情境再回頭調整指引。未來應該增加範本、範例，針對不同業務開發、應用 AI 系統，應該要如何適用指引或調整。

- 3、委員會除對外溝通外，對內也要衡平法制單位和業務單位。委員會的組成、條件，甚至次級還有諮詢委員會。
- 4、DPIA 最難的是一開始的「識別」，未來要搭配施行細則或指引的引導。第一階段整體由業務單位描述 AI 系統的核心，資料量、資料種類、資料主體（包含數量）、人工參與決策程度。第二階段對於現有資料處理的系統描述，第三階段利害關係人（特別是弱勢族群、受委託者），第四階段是評估使用 AI 是不是妥當？當事人的期待是什麼。
- 5、若採購已開發完成的 AI，廠商的條款裡可能有模糊空間，例如廠商與機關共享資料、廠商有權使用資料等。荷蘭政府採購微軟產品，DPIA 可參考。
- 6、不管是既有產品或開發產品，要約定好資料所有權，不能僅出現廠商制式化的隱私權宣告，就處理掉個資問題

專家 A

- 1、措辭修改，「宜」如何應該考量機關能力困難度，但政府機關遇到「宜」就不會做，所以有些宜應該改成「應」。因為很多系統一旦開發，如果出錯，之後很難再調整。應該在前期開發的時候就嚴格要求。或者雖然是「宜」，但案件數夠多或金額夠大後，有足夠經驗，就改成「應」。原則是「應」，但書滿足特定條件，經過特

定程序後排除規範。

- 2、AI 系統如果使用資料訓練模型，當事人要求刪除資料，可能無法達成。在法律上個資的定義是不是還適用？在 AI 系統內刪除個資的定義是什麼需要定義，或是定義進到模型後的資料即不具識別性。機關才敢使用。
- 3、技術安全風險要留意，只列特定風險，機關就只會注意這些風險。應該要定期更新。

專家 C

- 1、管理組織及權責部分，可考量納入監管或稽核機制。
- 2、研考會可以參考勞動部研究計畫。例如廠商要有特定 ISO 認證，才能被委外 AI。
- 3、語音轉文字後續將文字做其他處理，在個資法上是不是對於再使用能夠透過檢驗。機關是不是可以主張機關有這些資料，就可以大數據用於訓練人工智慧，宜一併評估。
- 4、跨單位的資料傳送、共用，也宜納入評估。
- 5、建議於指引中定義低中高風險的區別。

專家 D

- 1、指引和檢核表對行政機關要求較多，有沒有可能區分對待。例如判斷為中風險或低風險後，有沒有什麼義務可以不用作？應考量

讓行政機關知道可以怎麼操作，給公務機關一些彈性，讓公務機關知道在什麼情境下可以有一些選擇。

- 2、如區分風險有難度，可參考美國做法，區分企業大小。
- 3、委外實務管理上大多請受託廠商出自評表、承諾書。如果指引針對中高風險，可能要搭配執行實地稽核

專家 E

- 1、可參考歐盟做法，做風險類型區別。甚至可從高風險類型開始執行即可。例如接觸到某些資料，金融、健康、國安等。針對高風險系統關注事項列出，可參考歐盟。
- 2、治理委員會是核可與監督角色，應該可加入社會溝通、社會對話功能，甚至有教育單位，對外（對民眾）教育推廣學習平台，再轉化為完整的溝通平台。追求指引的法規化。
- 3、如指引鎖定中高風險 AI 系統，可適當納入機關替代機制，例如「因現有技術等因素限制，得採其他替代方式」。
- 4、德國 AI 實務管理，聯邦政府的人工智慧戰略，要政府在導入措施時，要確保實驗空間，可以考慮資料沙盒。參考英國做法，但這需要個資主管機關畫清界線。如果放在政府情境下有沒有機會試行人工智慧沙盒，評估人工智慧使用效益。

第三節 臺北市政府開發及應用人工智慧作業指引草案全文

本研究參酌焦點座談專家意見，對「臺北市政府開發及應用人工智慧作業指引」及其所附「人工智慧系統開發及應用檢核表」加以修正，形成草案修正版本，全文如下。

臺北市政府開發及應用人工智慧作業指引（草案）

一、前言

（一）臺北市政府（以下簡稱本府）為協助本府各機關（構）（以下簡稱各機關）掌握開發及應用人工智慧系統之重點考量事項，並保障民眾權益，特訂定本指引。

（二）人工智慧係指以電腦程式或其他非人類系統模仿人類學習、推理、解決問題、決策等智慧活動之技術或方法。人工智慧技術包括自然語言處理（natural language processing）、機器學習（machine learning）、知識表示（knowledge representation）等技術。人工智慧系統可用於產出內容、預測、建議或決策等，常見應用包括人臉辨識、語音辨識、聊天機器人、自動化決策等。人工智慧透過演算法產出模型（model），並選擇最適宜的模型部署於系統中。人工智慧系統之開發，通常包括資料蒐集與準備、演算法訓練與測試、模型選擇與驗證三個階段。人工

智慧模型之訓練與建構須使用大量資料組成之資料集 (dataset)。

(三) 人工智慧系統之開發及應用，有利於節約人力、改善效率、優化資源分配、提供個人化服務、促進創新等。然人工智慧模型之訓練、建構、部署及運作，通常涉及個人資料之蒐集、處理及利用，因而可能對隱私權及個人資料保護造成衝擊。此外，人工智慧系統對個人資料加以分析與評估，並據此作出推測、判斷等，可能影響個資當事人之權利、自由或利益。因此，如人工智慧遭不當開發或應用，可能損及人民之人性尊嚴、自主權、平等權等權益，亦可能損害公共利益及政府機關之公信力。因此，人工智慧系統之開發及應用，應遵循個人資料保護法規範，並應兼顧所有利害關係人之權利、自由與利益。

二、 人工智慧系統開發及應用之基本原則與法規框架

(一) 各機關開發及應用人工智慧系統時，應遵守下列基本原則：

1. 合法性與倫理性：各機關開發及應用人工智慧系統，應遵循一切相關法令規範、社會倫理原則與價值準則。
2. 公平性：各機關開發及應用人工智慧系統，應確保平等與公正之利益分配與成本負擔，並保護個人及群體免受偏見、歧視或

污名化對待。若人工智慧系統區別對待不同個人或群體，須有正當理由且符合比例原則。

3. 安全性：各機關應避免人工智慧系統對人類造成傷害；各機關所開發及應用之人工智慧系統並應採用穩健技術且確保安全，具備相當容錯能力，防禦惡意攻擊或濫用。
4. 透明性：各機關就其人工智慧系統之開發及應用作業、人工智慧系統之運作，以及所涉之個人資料蒐集、處理或利用，應確保公開透明。各機關並應盡可能確保人工智慧系統之產出結果具備可解釋性，以利受影響之利害關係人質疑或挑戰其產出結果。
5. 課責性：各機關應適當評估及因應人工智慧系統所涉風險，對其開發及應用行為負責。
6. 尊重民眾權益：各機關所開發及應用之人工智慧系統，應以人類為中心，尊重人類之自主權及其他權益，不得脅迫、欺騙或操縱人類。

(二) 各機關開發及應用人工智慧系統時，應依相關法令規範進行評估，擇定適當之政策或措施，包括：

1. 遵守憲法關於基本權利保障之規範，保障人民之人性尊嚴、各項自由權、平等權、資訊隱私權等。
2. 依行政程序法第 4 條之規範，受法律及一般法律原則之拘束，並遵守同法第 7 條比例原則之要求：
 - (1) 採取人工智慧之開發或應用方案應有助於目的之達成（適當性）；
 - (2) 選擇對人民權益損害最少之方式（必要性或侵害最小性）；
 - (3) 造成之損害不得與欲達成目的之利益顯失均衡（衡量性或狹義比例原則）。
3. 確保人工智慧系統所涉之個人資料蒐集、處理或利用符合個人資料保護法之規範，尊重個資當事人之權益，依誠實及信用方法為之，不得逾越特定目的之必要範圍，與蒐集之目的具有正當合理之關聯。
4. 依資通安全管理法及相關子法規之規範，確保人工智慧系統及其所涉資訊之機密性、完整性及可用性，防範其遭受未經授權之存取、使用、控制、洩漏、破壞、竄改、銷毀或其他侵害。
5. 遵守機關業務之其他法規，防範人工智慧系統不當侵害人民權利、自由或利益。

(三) 各機關開發及應用人工智慧系統時，應依所涉人工智慧系統之風險，評估並執行本指引所列措施：

1. 對於高風險人工系統，應落實本指引所列各項措施。下列人工智慧系統，應視為具高風險：

(1) 大規模蒐集、處理或利用個人資料保護法第 6 條第 1 項所規定之資料（以下稱特種個人資料）；

(2) 大規模執行生物特徵辨識（**recognition**），例如人臉辨識、指紋辨識等；

(3) 大規模自動化分析或評估自然人之個人特徵，例如工作或學業表現、健康或經濟狀況、個人偏好或興趣、信用、行為、位置等，並據此做成決定；

(4) 實質影響公共安全、本機關法定職務之執行，或能源、水資源、交通、通訊、金融、醫療等領域關鍵設施之安全或運作；

(5) 實質影響刑事或行政執法活動之過程或結果；

(6) 評估或分析民眾獲取或享有本機關公共服務或福利之資格、條件或方式。

2. 對於未達高風險之人工智慧系統，應參酌本指引所列各項措施，落實相關法令規範要求。

3. 本指引發布機關得綜合考量技術發展及法規變動等因素，公告修正本項第一款所列高風險人工智慧系統範圍。

三、 人工智慧系統開發及應用之管理制度與組織

- (一) 各機關針對人工智慧系統之開發及應用，應就下列事項訂定相關管理制度：

1. 人工智慧系統開發及應用之整體政策，其內容應包括：

- (1) 本機關開發及應用人工智慧系統之目標，以及人工智慧系統與本機關業務執行中之功能定位；
- (2) 本機關開發及應用人工智慧系統應遵循之基本原則；
- (3) 人工智慧系統開發及應用之管理組織及權責；
- (4) 確保人工智慧系統合法性、倫理性、公平性、安全性、透明性及課責性，防範不當侵害人民權利、自由或利益之措施。

2. 開發或應用特定人工智慧系統之專案文件，其內容應包括：

- (1) 該人工智慧系統之使用案例（use case）描述，包括該人工智慧系統之功能、開發流程、應用對象、權限管理措施、績效評估指標等；
- (2) 使用該人工智慧系統之標準作業程序（SOP）；

(3) 該人工智慧系統之影響評估結果；

(4) 與利害關係人溝通該人工智慧系統開發或應用事宜之措施；

(5) 該人工智慧系統開發或應用之監督制度；

(6) 應用該人工智慧系統之定期教育訓練制度。

(二) 各機關針對人工智慧系統之開發及應用，應就下列事項訂定相關管理組織及其權責：

1. 於本機關最高管理階層內設置人工智慧治理委員會（或類似功能之組織），負責人工智慧開發及應用之核可與監督，協調平衡機關內各單位之作業，統籌推行機關與外部之社會溝通與對話。
2. 明定機關內各單位於人工智慧系統開發及應用過程中之權責，例如：

(1) 就本機關開發或應用之各人工智慧系統，指定權責單位就該人工智慧系統向人工智慧治理委員會報告；

(2) 由資訊單位協助權責單位設計、開發、監測及維護該人工智慧系統；

(3) 由法制單位協助權責單位確保該人工智慧系統符合法令規範及執行影響評估；

(4) 由相關業務單位依管理制度使用該人工智慧系統。

四、 人工智慧系統開發及應用之影響評估與人為監督

(一) 於開發或應用人工智慧系統前，或對應用中的人工智慧系統進行重大變更前，應先行識別人工智慧系統之作業流程與資料流，並全面檢視其與欲達成之目的間之必要性及合比例性，評估其對相關法令之遵循性，鑑別對人民權利、自由和利益可能產生之風險等因素，並規劃採行適當之風險控管或緩解措施。風險評估結果應報人工智慧委員會（或類似功能之組織）核備。

1. 識別人工智慧系統之作業流程與資料流，應考量下列因素：

(1) 該人工智慧系統之功能、組成部分、開發流程、應用流程、管理措施等；

(2) 該人工智慧系統之資料來源、蒐集方式與目的、保存方式與期限、在機關內外之流動與提供狀況等。

2. 檢視人工智慧系統與欲達成之目的間之必要性及合比例性，應考量下列因素：

(1) 所欲達成之目的及其與社會公益之關聯；

(2) 為實現所欲達成之目的，該人工智慧系統係對人民權益影響最小之措施；

(3) 開發及應用該系統之預期成效及其與當事人權益影響程度之均衡性。

3. 評估人工智慧系統對相關法令之遵循性，應考量下列因素：

(1) 訓練、驗證或測試人工智慧演算法、構建或操作人工智慧系統所涉之個人資料蒐集、處理或利用，是否合於個人資料保護法及相關子法規；

(2) 人工智慧系統之設計、構建、操作或維護，是否合於資通安全管理法及相關子法規；

(3) 人工智慧系統之開發或應用，是否合於機關業務之其他法令規範。

4. 鑑別人工智慧系統對人民權利、自由或利益可能產生之風險，宜考量下列因素：

(1) 所採用之技術方案對人民資訊隱私之影響程度及發生可能性，包括蒐集、處理及利用資料是否確有必要、當事人對其個人資料使用狀況之知悉與控制、資料安全風險、資料委託第三方蒐集、處理或利用之風險等；

(2) 所採用之技術方案對人民平等權之影響程度及發生可能性，包括辨識結果之準確性等；

(3) 所採用之技術方案對人民人性尊嚴、表現自由等基本權之影響程度及發生可能性；

(4) 所採用之技術方案對人民其他權益之影響程度及發生可能性。

5. 各機關應依據所鑑別之風險，預先規劃足以避免或降低風險之因應措施。

6. 各機關執行上述影響評估，宜徵詢機關內外利害關係人（例如受影響之人民、機關工作人員、受託機關或業者）之意見。

7. 各機關應將影響評估結果與決策作成紀錄，並宜適當公開評估之主要發現或結論。

8. 各機關經評估後仍開發、應用或變更人工智慧系統者，應持續監測並定期再次評估人工智慧系統之風險及因應措施之有效性。

(二) 人工智慧系統模仿人類智慧活動，其作業應受人為監督，且其產出結果應確保準確性，以確保使用者可就人工智慧系統之產出作成知情自主判斷。

1. 各機關應依人工智慧系統之風險及影響，採行人為監督人工智慧系統之適當程度與方式。

(1) 高風險之人工智慧系統，原則應以「人為控制」(human-in-the-loop) 方式執行人為監督，由人類充分掌控決策過程，人工智慧僅提供建議。若無人類積極行為(例如給予指示)，則無法做成決策。

(2) 惟對於高風險人工智慧系統之具體應用情境，得綜合評估對當事人權益之影響、當事人對人工智慧產出結果之自主控制程度、機關作業之現實需求等因素，例外以「人為干預」(human-in-command)方式執行人為監督，由人類監督人工智慧系統的運作，並有權在特定情形下介入和推翻人工智慧系統之決策。

(3) 機關應確保人為監督之執行人員充分理解人工智慧系統之風險與局限，獨立研判人工智慧系統之產出並作出自主決策(包括否決人工智慧系統之建議或決策)，避免過度依賴自動化決策，防範自動化偏見(automation bias)。

2. 人工智慧系統資料集內之資料品質、不同群體代表性等因素，可能影響人工智慧系統產出之準確性，進而造成偏見。各機關應考量以下列措施降低人工智慧系統造成偏見之風險：

- (1) 依人工智慧系統之作業性質、範圍、背景與目的，以及對個人權利與自由之影響，適當設定產出標準；
- (2) 確保人工智慧系統資料集內之資料品質及不同群體之適當代表性；
- (3) 建立辨識準確率監測及通報機制；
- (4) 對人工智慧系統之使用者執行適當訓練，防範自動化偏見。

五、 人工智慧系統之透明性與個資當事人權利

(一) 各機關應採取下列措施，提升其人工智慧系統之開發及應用作業之透明性與可解釋性：

1. 因人工智慧系統之開發或應用而蒐集、處理或利用個人資料者，於蒐集個人資料前，應向個資當事人告知個人資料保護法第 8 條第 1 項各款或第 9 條第 1 項應告知事項。告知內容並應包括適當申訴管道等資訊。縱有個人資料保護法第 8 條第 2 項或第 9 條第 2 項規定例外「得」免除告知義務之具體情形，考量人工智慧系統對個人當事人權利、自由或利益之風險，仍宜儘量告知前述事項。
2. 如將已蒐集之個人資料為蒐集特定目的外之利用，若以個人資料保護法第 16 條但書第 7 款「經當事人同意」為目的外利用之

依據，應依個人資料保護法第 7 條第 2 項於事前將該利用行為之目的、範圍及同意與否對其權益之影響告知個資當事人；若依個人資料保護法第 16 條但書其他各款為目的外利用個人資料之依據，仍宜儘量告知前述事項。

3. 各機關應向民眾揭露其人工智慧系統應用狀況，包括人工智慧技術之簡介、本機關應用之人工智慧系統、使用人工智慧系統之益處、決定使用人工智慧系統之原因、風險因應措施、人工智慧系統在本機關業務活動中之作用等。

4. 各機關應對人工智慧系統自動化決策或輔助決策提供解釋，包括：

(1) 人工智慧系統參與決策之方式、範圍及程度，以及人為監督之方式與程度；

(2) 人工智慧系統決策所涉邏輯之有意義資訊，包括決策所使用之資料內容及其相關性與必要性、決策之主要考量要素、決策所依據之標準等；

(3) 人工智慧系統決策對個人權益之重要影響和預估後果。

5. 若因現有技術等因素限制，機關難以滿足人工智慧系統決策邏輯之可解釋性，則仍應記錄人工模型與同樣情境下產出結果之一致性，展現人工智慧決策結果之重複性。
 6. 前述告知或揭露資訊應使用清晰簡明之語言，以便於當事人取得且易於理解之方式為之。
 7. 各機關宜評估人工智慧系統各類利害關係人之特徵與需求，並參酌與該等利害關係人溝通之目的、背景、互動方式等因素，相應訂定適宜之告知或資訊揭露方案。其方案內容宜包括：
 - (1) 本項第 1 款至第 5 款告知或資訊揭露之具體方式，包括考量利害關係人與人工智慧系統互動之具體情形，採分層方式為之；
 - (2) 對利害關係人應用人工智慧系統時，使其適當知悉人工智慧系統運作狀況及可行替代方案之具體方式；
 - (3) 便利利害關係人表達意見並及時取得協助之具體管道。
- (二) 各機關開發和應用人工智慧系統時，應採取下列措施尊重與保障人民權益：
1. 因人工智慧系統之開發或應用而蒐集、處理或利用個人資料者，應提供當事人查詢、更正、補充個人資料之管道（個人資料保

護法第 10 條、第 11 條第 1 項及第 2 項)；且於特定目的消失或期限屆滿時，原則應主動或依個資當事人之請求，刪除、停止處理或利用該個人資料（個人資料保護法第 11 條第 3 項）。協助個資當事人行使其權利之措施應包括：

(1) 對人工智慧系統所蒐集、處理或利用之個人資料執行清查盤點；

(2) 確保人工智慧系統具備適當功能，足使本機關回應個資當事人之權利行使請求。

2. 以個人資料保護法第 6 條第 1 項第 6 款或第 15 條第 2 款「當事人同意」作為個人資料蒐集、處理之依據者，如個資當事人不同意機關為人工智慧系統之開發或應用而蒐集、處理或利用其個人資料，機關不得以影響個資當事人權益之方式迫使其同意，並應依其情形提供其他合理之替代方案。個資當事人未表示意見者，除符合個人資料保護法第 7 條第 3 項者外，應視為不同意。個資當事人並得於事後撤回其同意。

3. 若人工智慧系統之自動化決策結果將對人民產生法律效果或將影響重大權益，應提供受影響之人民說明、異議或請求人為介入之管道。

4. 機關就其所開發或利用之人工智慧系統，宜參酌其專案性質、開發及應用背景、系統功能、所涉個人資料內容、蒐集處理或利用個人資料之法律依據等，提供當事人請求該人工智慧系統停止利用其個人資料之機制。

六、 人工智慧系統之安全維護

(一) 各機關開發及應用人工智慧系統時，應依個人資料保護法、資通安全管理法等法令規範，採取組織性和技術性措施，確保人工智慧系統在預定使用期限內穩健可靠運作，防禦惡意攻擊或濫用，避免或減低對人民權益之意外傷害。相關措施應包括：

1. 對人工智慧系統執行嚴格測試，提升系統可靠性、穩健性及公平性，包括測試人工智慧系統在無效輸入或壓力環境條件下正常運作之能力；
2. 建立人工智慧系統使用者權限管理及身分驗證制度；
3. 嚴格管控人工智慧系統演算法訓練資料、訓練程式碼、部署程式碼之存取權限，並作成紀錄；
4. 建立自動日誌機制，記錄人工智慧系統各項使用、存取、變更等活動；
5. 對人工智慧系統所利用、儲存或傳輸之資料予以加密；

6. 建立事故偵測、通報和應變程序；

7. 確保人工智慧系統及資料之實體安全。

- (二) 各機關應識別、評量及有效因應人工智慧系統因其技術特性而面臨之特殊安全風險，包括操控資料集、故意引發模型錯誤產出、利用模型缺陷之風險，例如資料汙染（data poisoning）、模型逆推攻擊（model inversion attack）、成員推論攻擊（membership inference attack）、對抗例（adversarial examples）、白箱推論（white-box inference）等。各機關並應及時掌握人工智慧系統面臨之新型技術威脅及其因應方式。

七、人工智慧系統開發及應用之實務管理

- (一) 各機關於開發或應用人工智慧系統前，應清楚識別該人工智慧系統之應用情境與流程，採行人工智慧系統之規格方案，並對該方案執行風險評估及採取相應之風險因應措施。
- (二) 各機關於準備用以訓練、測試或支援人工智慧演算法運作之資料時，應確保資料來源及利用之合法性，確認資料具備良好品質，評估資料集內資料之代表性、可靠性與相關性，採取降低偏見及安全風險之適當技術措施，並盡可能對個人資料作去識別化處理。

- (三) 各機關應依所採行之規格方案，訓練與測試人工智慧模型，評估人工智慧模型之公平性、準確性、安全性、可解釋性、適當人為監督程度和方法等，並選擇最適宜之人工智慧模型。
- (四) 各機關完成人工智慧系統部署後，應確保對人工智慧系統之人為監督，持續監測模型運作狀況及成效，並定期以新資料調校及重新訓練人工智慧模型。

八、 人工智慧系統之採購與委外管理

- (一) 各機關如自外部採購已開發完成之人工智慧解決方案，或委託外部廠商開發、運作或維護人工智慧系統，應透過契約約定等方式，確保廠商提供之產品或服務符合相關法令及本機關人工智慧系統管理制度之要求。
- (二) 各機關對於廠商開發完成之人工智慧系統，應確認廠商訓練及測試資料來源合法、品質良好且具備適當代表性，並由內部或外部測試驗證其系統之公平性、準確性、安全性、可解釋性、具備適當人為監督等。
- (三) 各機關委託廠商運作或維護人工智慧系統，應對受託廠商予以適當監督，確保廠商嚴格遵守機關之指示，對人工智慧系統採

取適當安全維護和監測措施，及時向機關報告並協助因應系統事故等，並適時以實地稽核等方式確認廠商之遵循狀況。

九、 其他注意事項

(一) 各機關應對人工智慧系統之開發及應用作成充分紀錄，例如：

1. 人工智慧模型訓練所用資料來源、貼標 (labelling)、演算法之紀錄；
2. 人工智慧系統各項決策之輸入及輸出資料之紀錄；
3. 人為監督結果及其依據之紀錄。

(二) 各機關應對人工智慧系統執行定期稽核，確認人工智慧系統符合相關法令及本機關人工智慧系統管理制度之要求。

各機關應對參與人工智慧系統開發及應用之相關人員提供教育訓練，使其具備適當知識與技能。

指引草案附錄：

人工智慧系統開發及應用檢核表

人工智慧系統開發及應用檢核表						
1. 管理制度與組織			完全符合	部分符合	尚未符合	不適用
1.1	管理制度	訂定人工智慧系統開發及應用之整體政策。	☐	☐	☐	☐
		訂定開發或應用特定人工智慧系統之專案文件。	☐	☐	☐	☐
1.2	管理組織	於本機關最高管理階層內設置人工智慧治理委員會（或類似功能之組織），負責人工智慧開發及應用之核可與監督。	☐	☐	☐	☐
		明定機關內各單位於人工智慧系統開發及應用過程中之權責，例如： ☐ 該人工智慧系統之權責單位； ☐ 資訊單位之分工與責任； ☐ 法制單位之分工與責任； ☐ 相關業務單位之分工與責任。	☐	☐	☐	☐
2. 影響評估與人為監督			完全符合	部分符合	尚未符合	不適用
2.1	影響評估	於開發或應用人工智慧系統前，或對應用中的人工智慧系統進行重大變更前，執行影響評估。*	☐	☐	☐	☐
		識別人工智慧系統之作業流程與資料流： ☐ 系統功能與組成部分； ☐ 系統開發流程； ☐ 系統應用流程； ☐ 系統管理措施； ☐ 所用資料來源； ☐ 資料蒐集方式與目的； ☐ 資料保存方式與期限； ☐ 資料在機關內外之流動與提供狀況。	☐	☐	☐	☐
		評估人工智慧系統與欲達成之目的間之必要性及合比例性： ☐ 所欲達成之目的及社會公益相關性；	☐	☐	☐	☐

		☐ 該系統對達成該目的之必要性； ☐ 該系統之預期成效與當事人權益影響程度之均衡性。				
		評估人工智慧系統對人民權利、自由或利益可能產生之風險： ☐ 所影響之權利、自由或利益； ☐ 影響之程度； ☐ 影響發生之可能性。	☐	☐	☐	☐
		依所鑑別之風險，規劃避免或降低風險之因應措施。	☐	☐	☐	☐
		執行影響評估時，徵詢機關內外利害關係人之意見。	☐	☐	☐	☐
		適當公開影響評估之主要發現或結論。	☐	☐	☐	☐
		持續監測並定期再次評估人工智慧系統之風險及因應措施之有效性	☐	☐	☐	☐
2.2	人為監督	依人工智慧系統之風險及影響，確定對人工智慧系統進行人為監督之適當程度與方式： ☐ 原則：高風險系統應採人為控制（human-in-the-loop）方式； ☐ 例外：高風險系統經個案情形綜合評估，例外得採人為干預（human-in-command）方式； ☐ 人為監督之執行人員獨立評估與決策，防範自動化偏見。	☐	☐	☐	☐
		採取措施確保人工智慧系統之準確性，防範偏見，例如： ☐ 就系統產出設定適當標準； ☐ 確保所用資料之品質及適當代表性； ☐ 建立辨識準確率監測及通報機制； ☐ 防範使用者自動化偏見。	☐	☐	☐	☐
3. 透明性與個資當事人權利			完全符合	部分符合	尚未符合	不適用
3.1	透明性	就人工智慧系統所涉之個人資料蒐集、處理與利用，向個資當事人告知法定事項： ☐ 本機關名稱； ☐ 蒐集之目的； ☐ 個人資料之類別；	☐	☐	☐	☐

	☐ 個人資料利用之期間、地區、對象及方式； ☐ 當事人依個人資料保護法得行使之權利及方式； ☐ （直接蒐集資料時）當事人得自由選擇提供個人資料時，不提供將對其權益之影響； ☐ （間接蒐集資料時）個人資料之來源； ☐ （目的外利用已蒐集之資料時）特定目的外之其他利用目的、範圍及同意與否對其權益之影響。				
	向民眾揭露本機關人工智慧系統應用狀況： ☐ 人工智慧技術之簡介； ☐ 本機關應用之人工智慧系統； ☐ 使用人工智慧系統之益處； ☐ 決定使用人工智慧系統之原因； ☐ 風險因應措施； ☐ 人工智慧系統在本機關業務活動中的作用。	☐	☐	☐	☐
	對人工智慧系統自動化決策或輔助決策提供解釋： ☐ 人工智慧系統參與決策之方式、範圍及程度； ☐ 人為監督之方式與程度； ☐ 人工智慧決策所使用之資料內容及其相關性與必要性； ☐ 人工智慧決策之主要考量要素及所依據之標準； ☐ 人工智慧系統決策對個人權益之重要影響和預估後果。	☐	☐	☐	☐
	採取適宜之利害關係人資訊告知或揭露方案： ☐ 考量溝通對象之特徵與需求，以及溝通之目的、背景、互動方式等，確定告知之適當方式； ☐ 對溝通對象應用人工智慧系統時，採取適當措施使其知悉人工智慧系統運作狀況及可行替代方案；	☐	☐	☐	☐

		☐ 提供便利溝通對象表達意見並及時取得協助之具體管道。				
3.2	當事人權利	就人工智慧系統所涉之個人資料蒐集、處理與利用，提供個資當事行使其權利之管道。*	☐	☐	☐	☐
		採取措施協助個資當事人行使其權利： ☐ 清查盤點人工智慧系統蒐集、處理或利用之個人資料； ☐ 確保人工智慧系統具備回應個資當事人請求之適當功能。	☐	☐	☐	☐
		依個資當事人同意使用人工智慧系統時，採取措施確保有效同意：* ☐ 對不同意之個資當事人提供合理之替代方案； ☐ 提供個資當事人事後撤回其同意之管道。	☐	☐	☐	☐
		人工智慧系統之自動化決策結果有法律效果或影響重大權益者，應提供受影響者予以說明、異議或請求人為介入之管道。*	☐	☐	☐	☐
		參酌人工智慧系統之性質、背景、功能、所涉個人資料內容、蒐集處理或利用個人資料之法律依據等，提供當事人請求該人工智慧系統停止利用其個人資料之適當機制。	☐	☐	☐	☐
4. 安全維護			完全符合	部分符合	尚未符合	不適用
4.1	資訊安全措施	依個人資料保護法、資通安全管理法等法令規範，採取組織性和技術性措施，確保人工智慧系統穩健運作：* ☐ 嚴格測試系統可靠性、穩健性及公平性，包括在無效輸入或壓力環境條件下正常運作之能力； ☐ 建立使用者權限管理及身分驗證制度； ☐ 嚴格管控系統組成部分及資料之存取； ☐ 建立自動日誌機制； ☐ 系統利用、儲存或傳輸資料之加密保護； ☐ 建立事故偵測、通報和應變程序； ☐ 確保人工智慧系統及資料之實體安全。	☐	☐	☐	☐

4.2	防範人工智慧系統 專有風險	識別、評量及有效因應人工智慧系統因其技術特性而面臨之特殊安全風險。	☐	☐	☐	☐
5. 開發及應用之實務管理			完全符合	部分符合	尚未符合	不適用
5.1	方案設計	清楚識別該人工智慧系統之應用情境與流程，確定人工智慧系統之規格方案。	☐	☐	☐	☐
		對人工智慧系統方案執行風險評估及採取相應風險因應措施	☐	☐	☐	☐
5.2	資料準備	應確保資料來源及利用之合法性。*	☐	☐	☐	☐
		確認資料具備良好品質，評估資料集內資料之代表性、可靠性與相關性。				☐
5.3	模型訓練與測試	對人工智慧模型進行訓練與測試，評估人工智慧模型之公平性、準確性、安全性、可解釋性、適當人為監督程度和方法等，選擇最適宜之人工智慧模型	☐	☐	☐	☐
5.4	部署與監測	確保對人工智慧系統之人為監督。	☐	☐	☐	☐
		持續監測模型運作狀況及成效，定期以新資料調校及重新訓練人工智慧模型。	☐	☐	☐	☐
6. 採購與委外管理			完全符合	部分符合	尚未符合	不適用
6.1	採購管理	確認廠商訓練及測試資料來源合法、品質良好且具備適當代表性。	☐	☐	☐	☐
		執行內部或外部測試，驗證其系統之公平性、準確性、安全性、可解釋性、具備適當人為監督等。	☐	☐	☐	☐
6.2	委外管理	確保受託廠商嚴格遵守機關之指示。*	☐	☐	☐	☐
		確保受託廠商對人工智慧系統採取適當安全維護和監測措施。*	☐	☐	☐	☐
		確保受託廠商及時向機關報告並協助因應系統事故。*	☐	☐	☐	☐
7. 其他			完全符合	部分符合	尚未符合	不適用
7.1	紀錄保持	對人工智慧系統之開發及應用保持充分紀錄。	☐	☐	☐	☐

7.2	定期稽核	對人工智慧系統執行定期稽核，確認人工智慧系統符合相關法令及本機關人工智慧系統管理制度之要求。	☐	☐	☐	☐
7.3	教育訓練	對參與人工智慧系統開發及應用之相關人員提供教育訓練，使其具備適當知識與技能。	☐	☐	☐	☐
檢核結果彙整（各項勾選總數）						
註： * 標註之欄位評估結果如有「尚未符合」者，於人工智慧系統啟動開發或應用前，應完成改善。						

第七章 結論與建議

第一節 立即可行之建議

一、法規面

在個人資料保護法規遵循方面，就現行北市府規劃對於 1999 臺北市民當家熱線採用之人工智慧應用為「語音轉文字及文本內容關鍵字分析」而言，由於該措施之目的係為輔助話務人員更有效率服務進線市民，或識別服務品質滿意度，此行為應可主張屬於個人資料保護法第 16 條本文「於執行法定職務必要範圍內為之，並與蒐集之特定目的相符」之目的內利用（縱使遭認定為目的外利用個資行為，尚有主張符合個人資料保護法第 16 條但書第 2 款後段「為增進公共利益所必要」之空間）；且於蒐集個資階段，亦可主張個人資料保護法第 8 條第 2 項第 2 款「公務機關執行法定職務所必要」事由而例外免除「告知義務」。

然由保障當事人對其個人資料之知情與自主控制之視角，1999 熱線於導入人工智慧應用前，宜考量向當事人踐行告知，包括當事人之個人資料將用於「人工智慧應用之訓練與改善」目的、所涉資料類別、利用之期間、地區、對象和方式，以及當事人得行使之權利等。

此外，如針對擬導入人工智慧應用，北市府經衡酌判斷，認較宜

以人（書面）同意」為法律依據，則需依個資法第 8 條、第 9 條向當事人告知法定事項，並依同法第 7 條第 1 款徵求當事人允許之積極意思表示，取得當事人知情同意。

又北市府於人工智慧應用驗證過程中，由廠商提供相關工具，將市民進線通話錄音樣本置於獨立儲存裝置中，供廠商於話務中心內部測試轉為文字檔案，此行為涉及委託他人處理或利用個人資料，依個人資料保護法第 4 條規定，受託者（廠商）就受託行為，於個人資料保護法範圍內視同委託機關（北市府研考會等機關），是該機關應依個人資料保護法施行細則第 8 條之要求，對廠商為必要、適當之監督，並應考量訓練過程的個人資料安全維護事項（個人資料保護法第 18 條、施行細則第 12 條）。於驗證概念階段，北市府要求廠商親至其控管之實體環境，由北市府將錄音檔案以儲存媒介交付廠商，於現場使用廠商軟硬體設備執行人工智慧分析，事後歸還錄音檔案並不得下載攜出，應可認為已於個案對廠商為必要、適當之監督。而日後之導入階段之對廠商監督之具體措施，應視日後委託之目的、範圍等條件而定。

二、流程面

如前所述，由於北市府目前對於 1999 臺北市民當家熱線之人工智慧應用僅於評估規劃「語音轉文字及文本內容關鍵字分析」，以輔

助話務人員更有效率服務進線市民，或識別服務品質滿意度之階段，應有認定為「執行法定職務所必要」之空間，而毋需「取得當事人同意」，亦得免除蒐集個資時的「告知義務」。

惟考量日後人工智慧應用之目的與情境發展，參酌國際法規實務在人工智慧倫理層面對於「透明性」與「當事人控制權」（包含前述我國憲法機關對於當事人資訊隱私權中的事後控制權）的要求趨勢，北市府未來如有其他使用人工智慧之目的（例如人工智慧自動化決策，或變更原始蒐集目的而為新目的利用個人資料訓練或使用人工智慧）時，仍宜考量：

- 1、以「階層」方式於通話中達到「告知法定資訊」的效果（例如搭配 IVR 系統的使用，於市民進線時語音告知蒐集通話內容之目的與利用方式等重要事項，再告知市民可至網站取得完整隱私權政策，或再以簡訊寄送完整隱私權政策之網址連結）。
- 2、若有「取得當事人同意」始得合法利用資料之需求，除考量由話務人員於通話中徵求同意之外，亦可搭配 IVR 系統，由市民於通話時（例如於該次進線服務結束後）或於簡訊中按下特定符號或回復覆特定數字以表達同意。此外，由於 1999 市民當家熱線與北市府其他系統連結，可由市民於網站查詢案件進度，因此尚可考量於網站中在市民查詢案件操作過程徵求同意。

3、在所述與市民溝通過程中，所揭露資訊包含市民拒絕（或撤回同意）北市府利用其個人資料訓練人工智慧，或以人工智慧分析其個人資料而作成自動化決策，或其他變更原始蒐集目的之利用（於人工智慧）行為之條件與程序。

三、 制度面

本研究依據對歐盟、英國、美國、日本、新加坡與香港為法規、文獻之研析，參酌國際實務之見，就制度面之流程及注意事項，提出「臺北市政府開發及應用人工智慧作業指引」草案，並提供「人工智慧系統開發及應用檢核表」作為附件。

本研究以北市府初階規劃之語音轉文字及文本內容關鍵字分析人工智慧應用為對象，基於 1999 熱線之運作現狀，利用草案所附之「人工智慧系統開發及應用檢核表」進行評估實作，結果如附錄一所示。北市府亦可於明確人工智慧應用之具體方案、對 1999 熱線作業流程加以調整後，再行以該檢核表加以評估，以利掌握其對法規遵循、人工智慧治理及當事人權益之具體影響。

第二節 長期性建議

一、法規面

對於 1999 臺北市民當家熱線採用之人工智慧應用為「語音轉文字及文本內容關鍵字分析」而言，北市府對其所涉之個人資料蒐集、處理或利用，應有主張屬於「執行法定職務必要範圍內」而免於徵求同意或向當事人為告知之合法空間，業如前述。

然就市民對於利用其個人資料（通話檔案）訓練人工智慧之拒絕權而言，參照憲法法庭 111 年憲判字第 13 號判決意旨：「依前述對事後控制權之說明，就健保署因辦理健保業務而合法蒐集健保資料中個人健保資料，並提供公務機關或學術研究機構原始蒐集目的外利用此一限制個人資訊隱私權之行為，當事人之停止利用權應仍受憲法第 22 條規定保障。但由相關法制整體觀察，卻未見立法者依所保護個人資訊隱私權與利用目的間法益輕重及手段之必要性，而為適當權衡及區分，一律未許當事人得請求停止利用，停止利用應遵循之相關程序亦無規定，亦即欠缺當事人得請求停止利用之相關規定，顯然於個人資訊隱私權之保護有所不足」（判決理由第 71 段）可知，縱是「特定目的仍存在」，市民對其個人資料之事後控制權似仍須獲得保障。據此，北市府宜針對人工智慧應用之拒絕條件與程序訂定規範，適當供當事人（市民）行使其受憲法保障之資訊隱私權。

二、流程面

北市府現階段規劃之 1999 臺北市民當家熱線「語音轉文字及文本內容關鍵字分析」人工智慧應用，對於該熱線作業流程應無實質影響，業如前述。惟考量日後人工智慧應用之目的與情境發展，參酌國際法規實務在人工智慧倫理層面對於「透明性」與「當事人控制權」（包含前述我國憲法機關對於當事人資訊隱私權中的事後控制權）的要求趨勢，北市府未來如有其他使用人工智慧之目的（例如人工智慧自動化決策，或變更原始蒐集目的而為新目的利用個人資料訓練或使用人工智慧）時，仍宜考量：

- 1、於開發或應用人工智慧系統前，或對應用中的人工智慧系統進行重大變更前，執行影響評估，先行識別人工智慧系統之作業流程與資料流，並全面檢視其與欲達成之目的間之必要性及合比例性，評估其對相關法令之遵循性，鑑別對人民權利、自由和利益可能產生之風險等因素，並規劃採行適當之風險控管或緩解措施。
- 2、依人工智慧系統之風險及影響，確定對人工智慧系統進行人為監督之適當程度與方式，如涉高風險之人工智慧系統，原則應以「人為控制」方式執行人為監督。
- 3、針對人工智慧系統與市民進行溝通，除向市民告知其拒絕（或撤回同意）北市府利用其個人資料訓練人工智慧、或以人工智慧分

析其個人資料而作成自動化決策、或其他變更原始蒐集目的之利用（於人工智慧）行為之條件與程序外，並就北市府人工智慧系統之應用狀況、各項自動化決策或輔助決策之作成方式及影響等向市民提供解釋。

三、制度面

因現行國際實務尚未正式通過施行人工智慧法律，各國政府多採指引方式提供公私部門開發或使用人工智慧系統的優良實務建議。本研究亦建議現階段宜以指引方式作為北市府人工智慧應用合規之控管檢核機制。據此，本研究提出「臺北市政府開發及應用人工智慧作業指引」草案從人工智慧系統開發及應用之基本原則與法規框架、管理制度與組織、影響評估與人為監督、透明性與個資當事人權利、安全維護、事務管理、委外管理等面向，為北市府提出人工智慧應用的制度面建議，並提供「人工智慧系統開發及應用檢核表」作為指引附錄。

北市府可於對該草案內容加以審修後，正式發布該指引，供北市府各機關遵循，以利強化於開發及應用人工智慧系統時之法令遵循及良好治理。北市府未來並宜綜合考量技術發展及法規變動等因素，適時修正該指引所列高風險人工智慧系統範圍。

第八章 參考文獻

法律規範

- 中華民國個人資料保護法。
- 中華民國個人資料保護法施行細則。
- 臺北市政府研究發展考核委員會個人資料保護管理要點。
- 臺北市政府資訊局個人資料保護管理要點。
- Regulation 2016/679/EU of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (GDPR).
- Singapore Personal Data Protection Act 2012 (amended 2021).
- United Kingdom Data Protection Act 2018 (amended 2021).
- United Kingdom Equality Act 2010.

法院判決

- 中華民國憲法法庭 111 年憲判字第 13 號判決。
- 中華民國最高行政法院 106 年度判字第 54 號判決。

實務指引與工作文件

- 法務部 106 年 8 月 7 日法律字第 10603510070 號函。
- 法務部 107 年 7 月 13 日法律字第 10703510240 號函。
- 法務部 107 年 10 月 19 日法律字第 10703515980 號函。
- 法務部 108 年 7 月 12 日法律字第 10803509450 號函。

- 勞動部，109 年度「本部電話服務中心智能客服系統與智能分析平台建置、導入、擴充暨維護案」，徵求廠商企劃書。
- 國家發展委員會 107 年 11 月 7 日發法字第 1072002035 號函。
- 國家發展委員會 107 年 11 月 21 日發法字第 1072002136 號函。
- 國家發展委員會 108 年 1 月 22 日發法字第 1080000958 號函。
- 國家發展委員會 108 年 7 月 4 日發法字第 1082001104 號函。
- 國家發展委員會 108 年 11 月 1 日發法字第 1080021162 號函。
- 國家發展委員會 110 年 6 月 9 日發法字第 1102000884 號函。
- 國家發展委員會，服務型智慧政府 2.0 推動計畫(110 年至 114 年)。
- 臺北市政府資通訊業務委外作業指引。
- 臺北市政府資訊局個人資料保護管理要點。
- 臺北市政府研究發展考核委員會個人資料保護管理要點。
- 香港，個人資料私隱專員公署，開發及使用人工智慧道德標準指引（2021 年）。
- European Commission, Proposal for a Regulation of the European Parliament and of the Council Laying Down Harmonised Rules on Artificial Intelligence (Artificial Intelligence Act) and Amending Certain Union Legislative Acts (21.4.2021) (hereinafter the “AI Act Proposal”), COM(2021) 206 final.
- European Union High-Level Expert Group on Artificial Intelligence, Ethics Guidelines for Trustworthy AI (2019).

- Japan Cabinet, Social Principles of Human-Centric AI.
- Singapore IMDA & PDPC, Model Artificial Intelligence Governance Framework, Second Edition (2020).
- Singapore IMDA & PDPC, Compendium of AI Use Cases, Practical Illustrations of the Model AI Governance Framework (2020).
- Singapore IMDA & PDPC, Compendium of AI Use Cases (Volumn 2), Practical Illustrations of the Model AI Governance Framework (2020).
- UK ICO, Guidance on AI and data protection.
- UK Central Digital and Data Office, Guidance: Using chatbots and webchat tools.
- UK ICO, Guide to the UK General Data Protection Regulation (UK GDPR).
- UK ICO, Explaining decisions made with AI.

學術研究資料

- 丁玉珍、林子倫，人工智慧提升政府公共治理的應用潛力探討，檔案半年刊，109 年 12 月第 19 卷第 2 期。
- Berryhill, J., et al. (2019), "Hello, World: Artificial intelligence and its use in the public sector", *OECD Working Papers on Public Governance*, No. 36, OECD Publishing, Paris, <https://doi.org/10.1787/726fd39d-en>.
- European Commission, Joint Research Centre (JRC) (2021): Selected AI cases in the public sector. European Commission, Joint Research Centre (JRC) [Dataset] PID: <http://data.europa.eu/89h/7342ea15-fd4f-4184-9603-98bd87d8239a>.

網路資料

- 臺北市政府資訊局(111 年 9 月 15 日),《臺北市政府資訊局保有個人資料檔案公開項目彙整表》,載於:
<https://doit.gov.taipei/cp.aspx?n=39C8F94C564144D4>。
- 臺北市政府研究發展考核委員(111 年 12 月 2 日),《臺北市政府研究發展考核委員會保有及管理個人資料之項目彙整表》,載於:
https://rdec.gov.taipei/News_Content.aspx?n=F5702E09C9B011E9&sms=40C86180686507FA&s=B03182725078223D。
- 臺北市政府研究發展考核委員(111 年 4 月 15 日),《臺北市政府研究發展考核委員會新聞稿:錄音檔轉文字合作內部測試 增進 1999 話務人員服務效率》,載於:
https://1999.gov.taipei/News_Content.aspx?n=56435265EBCBF5BE&sms=EB0622377E126BF0&s=706744107896516A。
- 臺北市政府常見問答(110 年 8 月 20 日),《民眾是否可申請調閱 1999 錄音檔案?處理時間為多久?處理結果如何取得?》,載於:
https://www.gov.taipei/News_Content.aspx?n=EEC70A4186D4C828&sms=87415A8B9CE81B16&s=82DA34A1B60CF4E0。
- 臺北市政府 Smart City 官方網站,《1999 導入語音辨識評估專案》,載於:<https://smartcity.taipei/projdetail/275>。
- 臺北市政府 Smart City 官方網站,《北市府 1999 心聲分析規劃專案》,載於:<https://smartcity.taipei/projdetail/288>。
- Romania Journal,*Antonia, The First Virtual Public Servant In Romania*, (Apr 20, 2018),
<https://www.romaniajournal.ro/society-people/antonia-the-first-virtual-public-servant-in-romania/>.
- *Chatbot C-19*, (9. Januar 2023), <https://c19.bundesbots.de/>.
- Chaire Intelligence artificielle responsable à l'échelle mondiale,*Comparing European and Canadian AI Regulation*, (November 2021),

<https://www.lco-cdo.org/wp-content/uploads/2021/12/Comparing-European-and-Canadian-AI-Regulation-Final-November-2021.pdf>.

- *Corona-Krise - Schramböck: Chatbot „Mona“ als zentraler Infopoint für Unternehmen am USP*, (April 2020), <https://www.bmaw.gv.at/Presse/Archiv/Pressemeldungen-BMDW/2020/April-2020/Corona-Krise-Chatbot-USP-.html>.
- JRC Publications Repository, *EU research “Artificial Intelligence in public services”*, (2020), <https://publications.jrc.ec.europa.eu/repository/handle/JRC120399>.
- JRC Publications Repository, *EU research “National strategies on Artificial Intelligence: A European perspective”*, (2021), <https://publications.jrc.ec.europa.eu/repository/handle/JRC122684>.
- Lorena HERNANDEZ, *EU research “AI governance in the public sector”*, (17/04/2020), <https://joinup.ec.europa.eu/collection/elis-e-european-location-interoperability-solutions-e-government/document/paper-ai-governance-public-sector-three- Tales-frontiers-automated-decision-making-democratic>.
- Jaana Leikas, Aditya Johri, Marko Latvanen, Nina Wessberg, Antti Hahto, *Governing Ethical AI Transformation: A Case Study of AuroraAI*, (10 February 2022), <https://www.frontiersin.org/articles/10.3389/frai.2022.836557/full>.
- M. Madaio, L. Stark, J. W. Vaughan, H. Wallach, *Microsoft, AI Fairness Checklist*, (2020), https://query.prod.cms.rt.microsoft.com/cms/api/am/binary/R_E4t6dA.
- Our Public Service, *Revenue Chatbot*, <https://www.ops.gov.ie/actions/innovating-for-our->

[future/innovation/emerging-technology/artificial-intelligence/revenue-chatbot/](#).

- *Roskilde kommune consistently resolves 92% of citizen queries with conversational AI*,(2022), <https://www.boost.ai/case-studies/kommune-kiri-roskilde-kommune>.
- Salvador Estevan Martínez & Doaa Samy, Natural Language and AI Solutions for Citizen Services Evolution (2019), https://joinup.ec.europa.eu/sites/default/files/inline-files/08-2%20Salvador_Estevan_Martinez%20and%20Doaa_Samy_0.pdf.
- *Statistics Estonia's virtual assistant Iti helps both users of statistics and data providers*,(17. April 2019), <https://www.stat.ee/en/uudised/news-release-2019-042>.
- *The Riigikogu has a new system for making verbatim reports*,(14.09.2020), <https://www.riigikogu.ee/en/press-releases/others/the-riigikogu-has-a-new-system-for-making-verbatim-reports/>.
- Veli-Matti Keloneva, Sirpa Myllymäki, Jarmo Rossi, Marjo-Riitta Lalli, *Virtual Assistant OuluBot*, <https://www.ouka.fi/oulu/oulubot/english>.
- Government Digital Service ,*UK Government Digital Service, How the DVLA deployed a chatbot and improved customer engagement*,(3 April 2020), <https://www.gov.uk/government/case-studies/how-the-dvla-deployed-a-chatbot-and-improved-customer-engagement>.
- Information Flanders ,*1700 supported by Artificial Intelligence*, <https://www.innovatieveoverheidsopdrachten.be/en/projects/1700-supported-artificial-intelligence>.

附錄一：1999 熱線檢核實作結果

人工智慧系統開發及應用檢核表

人工智慧系統開發及應用檢核表						
1. 管理制度與組織			完全符合	部分符合	尚未符合	不適用
1.1	管理制度	訂定人工智慧系統開發及應用之整體政策。	☐	☐	☐	☒
		訂定開發或應用特定人工智慧系統之專案文件。	☐	☐	☐	☒
1.2	管理組織	於本機關最高管理階層內設置人工智慧治理委員會（或類似功能之組織），負責人工智慧開發及應用之核可與監督。	☐	☐	☐	☒
		明定機關內各單位於人工智慧系統開發及應用過程中之權責，例如： ☐ 該人工智慧系統之權責單位； ☐ 資訊單位之分工與責任； ☐ 法制單位之分工與責任； ☐ 相關業務單位之分工與責任。	☐	☐	☐	☒
2. 影響評估與人為監督			完全符合	部分符合	尚未符合	不適用
2.1	影響評估	於開發或應用人工智慧系統前，或對應用中的人工智慧系統進行重大變更前，執行影響評估。*	☐	☐	☐	☒
		識別人工智慧系統之作業流程與資料流： ☐ 系統功能與組成部分； ☐ 系統開發流程； ☐ 系統應用流程； ☐ 系統管理措施； ☐ 所用資料來源； ☐ 資料蒐集方式與目的； ☐ 資料保存方式與期限； ☐ 資料在機關內外之流動與提供狀況。	☐	☐	☐	☒
		評估人工智慧系統與欲達成之目的間之必要性及合比例性： ☐ 所欲達成之目的及社會公益相關性；	☐	☐	☐	☒

		☐ 該系統對達成該目的之必要性； ☐ 該系統之預期成效與當事人權益影響程度之均衡性。				
		評估人工智慧系統對人民權利、自由或利益可能產生之風險： ☐ 所影響之權利、自由或利益； ☐ 影響之程度； ☐ 影響發生之可能性。	☐	☐	☐	☒
		依所鑑別之風險，規劃避免或降低風險之因應措施。	☐	☐	☐	☒
		執行影響評估時，徵詢機關內外利害關係人之意見。	☐	☐	☐	☒
		適當公開影響評估之主要發現或結論。	☐	☐	☐	☒
		持續監測並定期再次評估人工智慧系統之風險及因應措施之有效性	☐	☐	☐	☒
		2.2	人為監督	依人工智慧系統之風險及影響，確定對人工智慧系統進行人為監督之適當程度與方式： ☐ 原則：高風險系統應採人為控制（human-in-the-loop）方式； ☐ 例外：高風險系統經個案情形綜合評估，例外得採人為干預（human-in-command）方式； ☐ 人為監督之執行人員獨立評估與決策，防範自動化偏見。	☐	☐
採取措施確保人工智慧系統之準確性，防範偏見，例如： ☐ 就系統產出設定適當標準； ☐ 確保所用資料之品質及適當代表性； ☐ 建立辨識準確率監測及通報機制； ☐ 防範使用者自動化偏見。	☐			☐	☐	☒
3. 透明性與個資當事人權利			完全符合	部分符合	尚未符合	不適用
3.1	透明性	就人工智慧系統所涉之個人資料蒐集、處理與利用，向個資當事人告知法定事項： ☐ 本機關名稱； ☐ 蒐集之目的； ☐ 個人資料之類別；	☐	☐	☐	☒

	☐ 個人資料利用之期間、地區、對象及方式； ☐ 當事人依個人資料保護法得行使之權利及方式； ☐ （直接蒐集資料時）當事人得自由選擇提供個人資料時，不提供將對其權益之影響； ☐ （間接蒐集資料時）個人資料之來源； ☐ （目的外利用已蒐集之資料時）特定目的外之其他利用目的、範圍及同意與否對其權益之影響。				
	向民眾揭露本機關人工智慧系統應用狀況： ☐ 人工智慧技術之簡介； ☐ 本機關應用之人工智慧系統； ☐ 使用人工智慧系統之益處； ☐ 決定使用人工智慧系統之原因； ☐ 風險因應措施； ☐ 人工智慧系統在本機關業務活動中的作用。	☐	☐	☐	☒
	對人工智慧系統自動化決策或輔助決策提供解釋： ☐ 人工智慧系統參與決策之方式、範圍及程度； ☐ 人為監督之方式與程度； ☐ 人工智慧決策所使用之資料內容及其相關性與必要性； ☐ 人工智慧決策之主要考量要素及所依據之標準； ☐ 人工智慧系統決策對個人權益之重要影響和預估後果	☐	☐	☐	☒
	採取適宜之利害關係人資訊告知或揭露方案： ☐ 考量溝通對象之特徵與需求，以及溝通之目的、背景、互動方式等，確定告知之適當方式； ☐ 對溝通對象應用人工智慧系統時，採取適當措施使其知悉人工智慧系統運作狀況及可行替代方案；	☐	☐	☐	☒

		☐ 提供便利溝通對象表達意見並及時取得協助之具體管道。				
3.2	當事人權利	就人工智慧系統所涉之個人資料蒐集、處理與利用，提供個資當事行使其權利之管道。 *	☒	☐	☐	☐
		採取措施協助個資當事人行使其權利： ☐ 清查盤點人工智慧系統蒐集、處理或利用之個人資料； ☐ 確保人工智慧系統具備回應個資當事人請求之適當功能。	☐	☐	☐	☐
		依個資當事人同意使用人工智慧系統時，採取措施確保有效同意： * ☐ 對不同意之個資當事人提供合理之替代方案； ☐ 提供個資當事人事後撤回其同意之管道。	☐	☐	☐	☒
		人工智慧系統之自動化決策結果有法律效果或影響重大權益者，應提供受影響者予以說明、異議或請求人為介入之管道。 *	☐	☐	☐	☒
		參酌人工智慧系統之性質、背景、功能、所涉個人資料內容、蒐集處理或利用個人資料之法律依據等，提供當事人請求該人工智慧系統停止利用其個人資料之適當機制。	☐	☐	☐	☒
4. 安全維護			完全符合	部分符合	尚未符合	不適用
4.1	資訊安全措施	依個人資料保護法、資通安全管理法等法令規範，採取組織性和技術性措施，確保人工智慧系統穩健運作： * ☐ 嚴格測試系統可靠性、穩健性及公平性，包括在無效輸入或壓力環境條件下正常運作之能力； ☐ 建立使用者權限管理及身分驗證制度； ☐ 嚴格管控系統組成部分及資料之存取； ☐ 建立自動日誌機制； ☐ 系統利用、儲存或傳輸資料之加密保護； ☐ 建立事故偵測、通報和應變程序； ☐ 確保人工智慧系統及資料之實體安全。	☒	☐	☐	☐

4.2	防範人工智慧系統專有風險	識別、評量及有效因應人工智慧系統因其技術特性而面臨之特殊安全風險。	☐	☐	☐	☒
5. 開發及應用之實務管理			完全符合	部分符合	尚未符合	不適用
5.1	方案設計	清楚識別該人工智慧系統之應用情境與流程，確定人工智慧系統之規格方案。	☐	☐	☐	☒
		對人工智慧系統方案執行風險評估及採取相應風險因應措施	☐	☐	☐	☒
5.2	資料準備	應確保資料來源及利用之合法性。*	☒	☐	☐	☐
		確認資料具備良好品質，評估資料集內資料之代表性、可靠性與相關性。	☐	☐	☐	☒
5.3	模型訓練與測試	對人工智慧模型進行訓練與測試，評估人工智慧模型之公平性、準確性、安全性、可解釋性、適當人為監督程度和方法等，選擇最適宜之人工智慧模型	☐	☐	☐	☒
5.4	部署與監測	確保對人工智慧系統之人為監督。	☐	☐	☐	☒
		持續監測模型運作狀況及成效，定期以新資料調校及重新訓練人工智慧模型。	☐	☐	☐	☒
6. 採購與委外管理			完全符合	部分符合	尚未符合	不適用
6.1	採購管理	確認廠商訓練及測試資料來源合法、品質良好且具備適當代表性。	☐	☐	☐	☒
		執行內部或外部測試，驗證其系統之公平性、準確性、安全性、可解釋性、具備適當人為監督等。	☐	☐	☐	☒
6.2	委外管理	確保受託廠商嚴格遵守機關之指示。*	☒	☐	☐	☐
		確保受託廠商對人工智慧系統採取適當安全維護和監測措施。*	☒	☐	☐	☐
		確保受託廠商及時向機關報告並協助因應系統事故。*	☒	☐	☐	☐
7. 其他			完全符合	部分符合	尚未符合	不適用
7.1	紀錄保持	對人工智慧系統之開發及應用保持充分紀錄。	☐	☐	☐	☒

7.2	定期稽核	對人工智慧系統執行定期稽核，確認人工智慧系統符合相關法令及本機關人工智慧系統管理制度之要求。	☐	☐	☐	☒
7.3	教育訓練	對參與人工智慧系統開發及應用之相關人員提供教育訓練，使其具備適當知識與技能。	☐	☐	☐	☒
檢核結果彙整（各項勾選總數）			6	0	0	33
註： * 標註之欄位評估結果如有「尚未符合」者，於人工智慧系統啟動開發或應用前，應完成改善。						

附錄二：啟動會議記錄

「1999 臺北市民當家熱線數位轉型與個資保護對策之研究」
啟動會議紀錄

時 間：111年06月30日（星期四）下午2時00分

地 點： Google meet線上會議室

主 持 人：研考會陳宗亨組長

出席人員：研考會 張堯雯約聘企劃師、許家碩助理研究員、胡婷舜

助理研究員、蔡衣林組員

達文西個資暨高科技法律事務所 葉奇鑫計畫主持人、王

慕民協同主持人、孔德濤研究助理、海靖怡研究助理

紀 錄：孔德濤研究助理

會議結論：

一、 本次研究之主題，沒有包含草案規劃，係聚焦於導入AI的指引，也希望在專案中團隊可以考量以下事項：

1. 市政府研考會在今年訂定個資保護管理要點，屬於通案性的規定，針對機關對個資蒐集的過程，以及個資處理利用的管理程序、當事人的權利行使、個人資料風險安全評估等。
2. 市政府制定的個資相關自我查核檢核表，其內容是針對個

資保護、資安風險評估及個資法基本原理原則。

3. 1999專線的个人資料告知聲明設計。
4. 目前對蒐集資料（包含民眾個資、音檔、對話內容等）的保存期限規劃。
5. 當事人權利行使的方式，希望能與行政程序及成本取得平衡。
6. 台北通APP中已經有網路電話的功能，可以直接帶入會員的个人資料，跟傳統的電話服務過程有區別，所以希望知道在導入技術的過程中如何避免個資保護的相關風險，以發揮系統最好的功能。

二、 團隊對於市府提供之相關研究資料，請遵守契約書所訂定條之保密條款，勿對外揭露散佈資料。

三、 依據本次討論，各項主題焦點座談及訪談的時程都沒問題，
，團隊可以用市政府的名義邀請及聯繫與會人，但主要是由
團隊負責聯繫事務。目前話務中心使用的系統皆係由市政府
資訊局建置及維護，建議將資訊局納為訪談對象。

四、 八月適逢本市議會民政部門質詢及城市博覽會開幕式，專案
行程較多，希望團隊實地參訪市政府的時間能予避開。

五、 為利訪談邀約，請研展組協助在正式的DPIA訪談之前，建立

溝通管道俾供市府及團隊意見交流。

- 六、 有關資料交換部分，非敏感性、機密性的資料會先放置於公務雲端硬碟或TAPEION等公務管道供團隊下載；若屬敏感性、機密性資料會以光碟或隨身碟方式給予團隊。

附件：

「1999 臺北市民當家熱線數位轉型與個資保護對策之研究」期初工作會議簽到表

時間：111 年 6 月 30 日（星期二）14 時 0 分

地點：視訊會議

主席：研考會陳宗亨組長

紀錄：胡婷舜助理研究員

單 位	職 稱	姓 名	簽 名
臺北市政府研考會話務管理組	組長	陳宗亨	TAIPEION 簽到補 13:57:45
臺北市政府研考會話務管理組	約聘企劃師	張堯雯	台北通簽到補 13:57:40
臺北市政府研考會話務管理組	助理研究員	許家碩	台北通簽到補 13:57:49
臺北市政府研考會研究發展組	助理研究員	胡婷舜	台北通簽到補 13:57:35
臺北市政府研考會圖資管理組	組員	蔡衣林	台北通簽到補 13:59:24
達文西個資暨高科技法律事務所	計畫主持人	葉奇鑫	台北通簽到 13:30:33
達文西個資暨高科技法律事務所	協同主持人	王慕民	台北通簽到 13:43:13
達文西個資暨高科技法律事務所	研究助理	海靖怡	台北通簽到 13:51:04
達文西個資暨高科技法律事務所	研究助理	孔德浩	台北通簽到 13:54:28

會議代碼:111619086

附錄三：期中報告審查會議紀錄

「1999 臺北市民當家熱線數位轉型與個資保護對策之研究」

委託研究案

期中報告審查會議紀錄

壹、 時間：中華民國 111 年 10 月 4 日(星期二)上午 10 時

貳、 地點：視訊會議

參、 主持人：臺北市政府研究發展考核委員會黃嫻嬪專門委員

肆、 出(列)席人員：(詳簽到簿)

專家學者：

長慧法律事務所

資訊工業策進會科技法律研究所

資訊工業策進會科技法律研究所

明日科技法律事務所

黃于珊審查委員

顧振豪審查委員

王自雄審查委員

王珣瑩審查委員

市府單位：

臺北市政府研考會本部

臺北市政府研考會話務管理組

臺北市政府研考會研究發展組

臺北市政府研考會研究發展組

臺北市政府研考會話務管理組

臺北市政府研考會話務管理組

臺北市政府研考會圖資管理組

臺北市政府資訊局數據治理中心

黃嫻嬪專門委員

陳宗亨組長

林潔瑜組長

胡婷舜助理研究員

張堯雯約聘企劃師

許家碩助理研究員

楊詩涵聘用人員

陳卉荊約僱人員

研究團隊：

達文西個資暨高科技法律事務所

達文西個資暨高科技法律事務所

達文西個資暨高科技法律事務所

理

葉奇鑫計畫主持人

王慕民協同主持人

海靖怡研究助

伍、 主席致詞：(略)

陸、計畫主持人報告：(略)

柒、討論與建議：

一、王自雄委員

(一) 話務服務與 AI 為 1999 很重要的委外系統，建議團隊針對前開不同的委外項目，分別提出市府與廠商間個資法遵及委外關係適用的指引。

(二) 團隊於報告中有獲致香港、英國等資料保護檢核表單，惟就個人觀察，政府機關願意勾選檢核表的情境多限於客觀事實有無，如涉及上位法律原則或需要實質判斷時，行政單位是否有意願或能量勾選檢核表，可能需要相關制度設計或配套機制。

二、王珣瑩委員

(一) 「蒐集國內外政府或民間數位轉型相關案例」係為本計畫需求之一，目前報告內容較著重於國外政府導入 AI 案例而較少國內或民間部分，建議團隊於報告內容增補挑選案例的標準，以及欠缺民間或國內政府機關案例的原因，俾使讀者了解蒐集案例的參考價值。

(二) 有關第三章國外案例部分，建議團隊補充目前蒐集的案例是中央或地方，以及其應用的成功與否，俾於後續參

考利用。

- (三) 目前報告以「功能」區分國外案例，建議團隊整理後以表格方式呈現，例如橫軸、縱軸比較，橫軸為功能、縱軸為各國來做區分。
- (四) 關於第四章蒐集英國、亞洲及新加坡等國的相關法規，建議團隊進一步查詢該國應用的實例及使用情境。
- (五) 有關本研究開發人工智慧指引草案部分，宜先思考臺北市自訂指引的定位，及該指引與個資法及中央或地方相關法規的相互關係。

三、顧振豪委員

- (一) 關於各國文獻整理部分，建請團隊說明各國內容指引之特點及挑選原因，並以表格分析的方式處理，利於呈現國外執行重點。
- (二) 報告第 76 頁提及尚欠缺各國公部門對於人工智慧的個資保護評估及相關具體資訊，建議團隊於期末報告完整增補。
- (三) 目前團隊初步分析，1999 係執行法定職務所必要，依法不用取得當事人同意、免除告知義務，惟是否有其他國家相關案例，給予當事人事前選擇機制，建議團隊後續

與市府討論斟酌。

(四) 建議團隊提出 1999 與廠商間個資委外關係的操作指引，

俾供委辦單位運用。

四、黃于珊委員

(一) 建議團隊在國外資料研究後，對照 1999 的數位化部分，

將國外資料整理落實到本專案提供建議。

(二) 團隊初判 1999 基於執行法定職務必要範圍，依法不用取

得當事人同意、免除告知義務，惟為避免民眾觀感與認

知落差，希望團隊多蒐集國內外實務見解充實上開論述。

五、臺北市政府研考會話務管理組

(一) 案例彙整及文獻分析，希望團隊可以列出各國比較表，

例如歸納共同優點，或是分析各國特色。

(二) 本案評估標的是各種人工智慧及技術導入，希望團隊可

以多方評估導入風險，例如：關鍵字去判斷民眾進線反

映，思考及推演時不限於語音轉文字。

(三) 本案用字用語須斟酌，例如「全息投影」。而部分內容提

到的系統名稱並非準確，建議改成本案以 1999 電話服務

及各種輔助技術為評估標的。

(四) 提供操作評估表，未來會有其他技術，指引部分希望可

以涵蓋未來各種應用。

六、臺北市政府研考會研究發展組

- (一) 請予確認報告全文是否遵守學術文獻的引用規範，例如第三章案例彙整(P12 起)未加註引用資料的網站或出處，且勿使用對岸用語(例如 P15 最後一行「全息投影」)。
- (二) 本案委託標的係以研究報告為主，惟期中報告的附件扮演著很重要的核心角色，包含 1999 個案背景分析、本國法制適用與探討，建議將附件修改成本文第五章「臺北市民當家熱線現況分析」，俾於承上(文獻探討)啟下(調查分析)，增加報告的完整性；另因附件內容牽涉較多話務中心內部流程（例如 VIP、非理性等），請團隊會後洽話務組確認。
- (三) 本案預期成果有三(提出法規制度或服務流程再造建議、評估新服務的衝擊與風險影響等)，請團隊收斂研究資料，扣合標規構思其中可能產出的具體可操作性工具(例如評估檢核表或操作手冊指引)。
- (四) 請團隊持續蒐集國內外人工智慧法制導入的實務案例，如仍查無相關資料，請於報告內容的研究限制中敘明。
- (五) 其餘報告內容、錯漏字和統一文字用語等如後附錄，請

團隊修正。

捌、 會議結論

經審查委員討論結果，期中報告審查通過，請研究團隊依委員及委託機關的意見修正，併請團隊會後洽話務組確認報告中有關話務中心流程的正確性。

散會：上午 11 時 50 分。

「1999 臺北市民當家熱線數位轉型與個資保護對策之研究」
期中報告審查會議簽到表

時間：111 年 10 月 4 日（星期二）10 時 0 分

地點：視訊會議

主席：黃嫻嫻專門委員

紀錄：胡婷舜助理研究員

單 位	職 稱	姓 名	簽 名
臺北市政府研考會本部	專門委員	黃嫻嫻	台北通簽到 09:43:10
臺北市政府研考會話務管理組	組長	陳宗亨	TAIPEION 簽到 09:56:43
臺北市政府研考會研究發展組	組長	林潔瑜	台北通簽到 09:41:03
臺北市政府研考會話務管理組	約聘企劃師	張堯雯	台北通簽到 09:58:04
臺北市政府研考會話務管理組	助理研究員	許家碩	台北通簽到 09:58:12
臺北市政府研考會研究發展組	助理研究員	胡婷舜	台北通簽到 09:04:11
臺北市政府研考會研究發展組	約聘企劃師	曾國俊	台北通簽到 10:02:47
臺北市政府資訊局數據治理中心	約僱人員	陳卉荊	台北通簽到 09:19:10
臺北市政府研考會圖資管理組	聘用人員	楊詩涵	台北通簽到 09:51:26
長慧法律事務所律師	審查委員	黃于珊	台北通簽到 09:50:29
資訊工業策進會科技法律研究所	審查委員	顧振豪	台北通簽到 10:05:10
資訊工業策進會科技法律研究所	審查委員	王自雄	台北通簽到 10:05:05
明日科技法律事務所	審查委員	王琍瑩	台北通簽到 09:58:11
研究團隊(達文西個資暨高科技法律事務所)	計畫主持人	葉奇鑫	台北通簽到 09:55:14
研究團隊(達文西個資暨高科技法律事務所)	協同主持人	王慕民	台北通簽到 09:52:50
研究團隊(達文西個資暨高科技法律事務所)	研究助理	海靖怡	台北通簽到 09:48:51

附錄四：期中審查意見回應說明

「1999 臺北市民當家熱線數位轉型與個資保護對策之研究」

委託研究案

期中報告審查意見回應說明

委員	題號	審查意見	回應說明
王自雄委員	一	話務服務與 AI 為 1999 很重要的委外系統，建議團隊針對前開不同的委外項目，分別提出市府與廠商間個資法遵及委外關係適用的指引。	本研究依委員建議，提出「臺北市政府開發及應用人工智慧作業指引」草案，於草案第八點提出北市府採購與委外管理人工智慧系統之最佳實務。
	二	團隊於報告中有獲致香港、英國等資料保護檢核表單，惟就個人觀察，政府機關願意勾選檢核表的情境多限於客觀事實有無，如涉及上位法律原則或需要實質判斷時，行政單位是否有意願或能量勾選檢核表，可能需要相關制度設計或配套機制。	本研究依委員建議，提出「臺北市政府開發及應用人工智慧作業指引」草案，並提供「人工智慧系統開發及應用檢核表」作為指引附件，以「完全符合」、「部分符合」、「尚未符合」三類級距，供北市府各機關自我檢核。
王珣瑩委員	一	「蒐集國內外政府或民間數位轉型相關案例」係為本計畫需求之一，目前報告內容較著重於國外政府導入 AI 案例而較少國內或民間部分，建議團隊於報告內容增補挑選案例的標準，以及欠缺民間或國內政府機關案例的原因，俾使讀者了解蒐集案例的參考價值。	本研究依委員建議，於期末報告第四章補充我國勞動部多元智能電服中心、新加坡微軟對話式 AI 技術解決方案之案例，以實現案例之多元比較。
	二	有關第三章國外案例部分，建議團隊補充目前蒐集的案例是中央或地方，以及其應用的成功與否，俾於後續參考利用。	本研究依委員建議，於期末報告第四章補充各案例之層級（中央政府機關或地方政府），並就可得之資訊，補充案例係示範專案（例如愛爾蘭稅務局 VDA）或成功導入（例如德國 C19、奧地利 Mona）之相關內容。

委員	題號	審查意見	回應說明
	三	目前報告以「功能」區分國外案例，建議團隊整理後以表格方式呈現，例如橫軸、縱軸比較，橫軸為功能、縱軸為各國來做區分。	本研究依委員建議，於期末報告第六章第一節第二小節，補充「實務案例比較表」，從功能、與使用者互動狀況、人為監督方式、所涉技術等方面，以更直觀之方式呈現所研析案例之分布狀況。
	四	關於第四章蒐集英國、亞洲及新加坡等國的相關法規，建議團隊進一步查詢該國應用的實例及使用情境。	本研究依委員建議，於期末報告第四章補充英國駕照及行照監理局聊天機器人、新加坡微軟對話式 AI 技術解決方案之案例。
	五	有關本研究開發人工智慧指引草案部分，宜先思考臺北市自訂指引的定位，及該指引與個資法及中央或地方相關法規的相互關係。	本研究提出「臺北市政府開發及應用人工智慧作業指引」草案，向臺北市政府各機關提示開發或應用人工智慧系統之重點考量事項，應係臺北市政府落實數位轉型之內部參考文件，不致直接對行政相對人產生影響。
顧振豪委員	一	關於各國文獻整理部分，建請團隊說明各國內容指引之特點及挑選原因，並以表格分析的方式處理，利於呈現國外執行重點。	本研究依委員建議，於期末報告第六章第一節第一小節，補充「AI 法規及指引比較表」，以更直觀之方式呈現各國法規或指引位階比較，並對各指引內容之共同重點予以分析。
	二	報告第 76 頁提及尚欠缺各國公部門對於人工智慧的個資保護評估及相關具體資訊，建議團隊於期末報告完整增補。	經本研究後續蒐集資訊，除部分案例可知其事前經安全性及適法性評估（例如英國駕照及行照監理局聊天機器人）外，未見其他案例有個資保護影響評估之明確資訊。本報告將此狀況補充於期末報告第一章第四節「研究範圍與限制」部分。
	三	目前團隊初步分析，1999 係執行法定職務所必要，依法不用取得當事人同意、免除告知義務，惟是否有其他國家相關案例，給予當事人事前選擇	經本研究後續蒐集資訊，就所研析之官方專線、線上聊天機器人案例而言，當事人通常可自由選擇是否使用，如不同意使用，可要求人工服務

委員	題號	審查意見	回應說明
		機制，建議團隊後續與市府討論斟酌。	或改用其他溝通方式，惟就人工智慧技術之導入，未見提供當事人事前選擇之公開資訊。
	四	建議團隊提出 1999 與廠商間個資委外關係的操作指引，俾供委辦單位運用。	本研究依委員建議，提出「臺北市政府開發及應用人工智慧作業指引」草案，於草案第八點提出北市府採購與委外管理人工智慧系統之最佳實務。
黃于珊委員	一	建請團隊在國外資料研究後，對照 1999 的數位化部分，將國外資料整理落實到本專案提供建議。	本研究依委員建議，提出「臺北市政府開發及應用人工智慧作業指引」草案，參酌期末報告第六章之比較研究發現及各國外指引內容，從基本原則與法規框架、管理制度與組織、影響評估與人為監督、透明性與個資當事人權利、安全維護、事務管理、委外管理等面向，為北市府提出人工智慧應用的制度面建議。
	二	團隊初判 1999 基於執行法定職務必要範圍，依法不用取得當事人同意、免除告知義務，惟為避免民眾觀感與認知落差，希望團隊多蒐集國內外實務見解充實上開論述。	本研究依委員建議，審酌調整相關論述，配合「臺北市政府開發及應用人工智慧作業指引」草案第五點內容，提示北市府留意民眾認知落差，提升人工智慧系統開發及應用之透明性。
臺北市政府研考會話務管	一	案例彙整及文獻分析，希望團隊可以列出各國比較表，例如歸納共同優點，或是分析各國特色。	本研究依委員建議，於期末報告第六章第一節補充「AI 法規及指引比較表」及「實務案例比較表」，以更直觀之方式呈現各國法規或指引位階比較，並對各國案例之功能重點、與使用者互動狀況、人為監督方式、所涉技術等要素加以對比。
	二	本案評估標的是各種人工智慧及技術導入，希望團隊可以多方評估導入風險，例如：關鍵字去判斷民眾進線反映，思考及推演時不限於語音轉文	本研究依委員建議，提出「臺北市政府開發及應用人工智慧作業指引」草案，於較廣泛地考量不同應用情境之風險（例如，於第二點舉例說明高風

委員	題號	審查意見	回應說明
理組		字。	險應用)，並依不同風險情形提出相應因應風險之最佳實務建議。
	三	本案用字用語須斟酌，例如「全息投影」。而部分內容提到的系統名稱並非準確，建議改成本案以 1999 電話服務及各種輔助技術為評估標的。	本研究依委員建議，檢視並調整全文用字用語。
	四	提供操作評估表，未來會有其他技術，指引部分希望可以涵蓋未來各種應用。	本研究依委員建議，提出「臺北市政府開發及應用人工智慧作業指引」草案，並提供「人工智慧系統開發及應用檢核表」作為指引附件，於較廣泛地考量不同應用情境之風險（例如，於第二點舉例說明高風險應用），並依不同風險情形提供相應因應建議。
臺北市 政府 研考會 研究 發展 組	一	請予確認報告全文是否遵守學術文獻的引用規範，例如第三章案例彙整(P12 起)未加註引用資料的網站或出處，且勿使用對岸用語(例如 P15 最後一行「全息投影」)。	本研究依委員建議，補充引註來源，並檢視及調整全文用字用語。
	二	本案委託標的係以研究報告為主，惟期中報告的附件扮演著很重要的核心角色，包含 1999 個案背景分析、本國法制適用與探討，建議將附件修改成本文第五章「臺北市民當家熱線現況分析」，俾於承上(文獻探討)啟下(調查分析)，增加報告的完整性；另因附件內容牽涉較多話務中心內部流程(例如 VIP、非理性等)，請團隊會後洽話務組確認。	本研究依委員建議，將期中報告附件 DPIA 內容修改為期末報告第五章，並依市府意見酌修其細部內容。
	三	本案預期成果有三(提出法規制度或服務流程再造建議、評估新服務的衝擊與風險影響等)，請團隊收斂研究資料，扣合標規構思其中可能產出的	本研究依委員建議，提出「臺北市政府開發及應用人工智慧作業指引」草案，向北市府提示最佳實務做法，該草案並提供「人工智慧系統開發及應

委員	題號	審查意見	回應說明
		具體可操作性工具(例如評估檢核表或操作手冊指引)。	用檢核表」作為附件，提供北市府自我檢核之工具。
	四	請團隊持續蒐集國內外人工智慧法制導入的實務案例，如仍查無相關資料，請於報告內容的研究限制中敘明。	本研究依委員建議後續蒐集資訊，除補充實務案例相關內容(例如英國駕照及行照監理局聊天機器人、新加坡微軟對話式 AI 技術解決方案)外，亦於期末報告第一章第四節「研究範圍與限制」中補充說明本研究資料蒐集之限制。
	五	其餘報告內容、錯漏字和統一文字用語等如後附錄，請團隊修正。	本研究依委員建議，檢視並調整全文用字用語。

附錄五：期末報告審查會議紀錄

「1999 臺北市民當家熱線數位轉型與個資保護對策之研究」

委託研究案

期末報告審查會議紀錄

壹、時間：中華民國 112年1月4日(星期三)上午10時

貳、地點：視訊會議

參、主持人：臺北市政府研究發展考核委員會黃嫻嬪專門委員

肆、出(列)席人員：(詳簽到簿)

專家學者：

長慧法律事務所

黃于珊審查委員

資訊工業策進會科技法律研究所

顧振豪審查委員

資訊工業策進會科技法律研究所

王自雄審查委員

明日科技法律事務所

王珣瑩審查委員

市府單位：

臺北市政府研考會本部

黃嫻嬪專門委員

臺北市政府研考會話務管理組

陳宗亨組長

臺北市政府研考會研究發展組

林潔瑜組長

臺北市政府研考會研究發展組

胡婷舜助理研究員

臺北市政府研考會話務管理組

張堯雯約聘企劃師

臺北市政府研考會話務管理組

許家碩助理研究員

研究團隊：

達文西個資暨高科技法律事務所

葉奇鑫計畫主持人

達文西個資暨高科技法律事務所

王慕民協同主持人

達文西個資暨高科技法律事務所

李蘊家研究助理

伍、主席致詞：(略)

陸、計畫主持人報告：(略)

柒、討論與建議：

一、黃于珊委員

(一) 2022年10月中，美國白宮提出人工智慧權利法草案藍圖，因美國乃 AI 發展大國，建議於法規草案蒐集中將此規範納入。

(二) 報告附件一所提出之人工智慧作業指引，於第166頁明訂

6項為高風險項目，但欠缺開放性，建議補充授權市府針對高風險項目進行增修之條文。

(三) 報告第175頁針對以當事人同意為個人資料蒐集處理利用依據之情形，認為當事人若不同意可選擇退出，但未論及非以當事人同意作為依據時之當事人退出權，建議補充。

(四) 報告附件一提供「人工智慧系統開發及應用檢核表」，建議研究團隊後續與北市府討論，將1999市民熱線之人工智慧應用套入該檢核表進行分析。

二、顧振豪委員

(一) 據觀察，國際對人工智慧之相關規畫及草案之訂定等，皆著重於透明性及與利害關係人溝通，除報告第150頁提到之依個資法向當事人踐行告知外，建議團隊後續研究是否有社會溝通之其他建議方式做社會溝通。

三、王自雄委員

(一) 歐盟 AI 法之重點係建立人工智慧產品之審查標準，與本研究案重點應有區別，本研究案係以人工智慧應用之風險控管為核心，指引草案為政府對外公開政策之一，因此，除檢核表之外，如何強化與利害關係人之溝通，應為較為重要之部分。

(二) 建議團隊對檢核表增列評級表，供公務機關之同仁針對檢核表逐項填覆後，以總分揭示風險及提醒應注意之程度，以利確定後續應對作為，或針對人工智慧應用作類型化處理，將本研究案之效益最大化。

四、王珣瑩委員

(一) 報告第141頁、第145頁分別是法規指引及案例研究之比較分析，英國和新加坡兩國，同時都有研究該國指引和案例，但較缺少所指引及案例間在適用面的對照，或許是因較難找到公開資料，這部分頗為可惜，建議團隊補充說明。

(二) 建議與實際需求端即市府同仁進行交流，因實務上電話進線時，較難於短時間內完成個資告知並取得同意，甚至進線民眾可選擇略過該步驟，直接進入問答部分。故當實際發生此情形時，法律效力應如何認定？此部分牽涉

到是否合規，應提出明確見解。

五、臺北市政府研考會話務管理組

- (一) 報告裡有提到法務部、勞動局等案例，不知該單位於執行作業時有無遇過法律評估上之困境，實務上要取得內部過程不容易，建請團隊以現有檢核表或草案，利用府內資料進行評估測試，以使結果更有參考性。
- (二) 若電話客服個資蒐集等皆須民眾同意，須個別取得同意，此在進線語音設計及執行上恐有困難，且電話服務節奏緊湊，僅是於語音中供民眾選擇同意與否，若民眾未做選擇，可否視為有告知且民眾已同意？建請提供明確見解。
- (三) 民眾使用同意後續撤回機制，之前已有在設想，故較有把握。

六、臺北市政府研考會研究發展組

- (一) 請確認報告全文是否遵守學術倫理規範，例如 P3第3行「Lanx 提出…」、圖1系統風險層級示意圖應註記引用出處，P189焦點座談應隱匿發言之專家學者姓名以代碼呈現(例如:專家 A)，第八章標題應為參考文獻而非參考資源，並依學術研究報告格式製作(如 APA 格式或芝加哥格式)。
- (二) 報告請依本府委託研究要點規定格式製作，增補研究摘要(包括研究目的、過程、重要發現、主要建議及政策意涵)，第七章標題請修改為「結論與建議」，其建議事項分為「立即可行之建議」及「長期性建議」，附錄增補歷次工作會議紀錄。
- (三) 第二章研究方法建議與第三、四、五、六章撰寫順序一致，第一節建議改為法規及文獻分析，第二節為案例彙整，第三節為個資保護影響評估，第四節為焦點座談。
- (四) 因焦點團體討論內容係為本案研究成果之一，建議第六章研究發現呈現「臺北市政府開發及應用人工智慧作業指引(草案)」以及焦點座談調查(含施作方式、訪綱及以主題式的摘述各與會專家的意見)，原第六章內容回歸於第三~五各章節中，另立小節，以利整體閱讀。
- (五) 其餘報告內容、錯漏字和統一文字用語等如後附錄，請團隊修正。

捌、 會議結論

經審查委員討論結果，期末報告審查通過，請研究團隊依委員及委託機關之意見修正，併請團隊於112年2月2日交付修正版本。

散會：上午11時15分。

「1999 臺北市民當家熱線數位轉型與個資保護對策之研究」

期末報告書研考會審查意見

壹、報告內容

- 一、第四章第二節法務部語音辨識系統及第三節勞動部及高雄市政府 1999 智能客服，屬現行已有在推行之數位轉型政策，本研究主旨在探討數位轉型與個資保護對策，故建議團隊應於這兩個部分進一步探討法務部、勞動局及高雄市政府在推動語音辨識及智能客服時的法源依據及個資保護政策，以供本府未來推動時的參考。
- 二、第五章部分內容會有疑慮，建議修改為較中性的說法，例如：P107 4.1.7「或將引發是否對當事人權益侵害最小之疑慮」；P108 4.2.4「僅係有利於公務機關執行業務」（實際上話務中心加速服務及維護品質亦是在服務大眾，亦為增進公共利益），並請刪除參考文獻中內部文件部分，例如：P122 陳情服務標準作業流程。
- 三、P99 2.1.1 臺北市政府及所屬各機關處理人民陳情案件注意事項第 13 點次內容已於 111 年 8 月 4 日函頒修正為第 12 點，請確認。
- 四、P126 當初有請團隊協助提供保存期限，惟目前研究內容係直接說明本府已就個人資料之保存期限訂定規則，此部分請協助說明；9.2「服務結束後，服務紀錄於系統內繼續保存合理期間，以滿足進線民眾查閱處理結果、機關評估服務成效等需求，亦未屬資料蒐集特定目的之必要範圍，符合個資法第 11 條第 3 項之要求」，文意矛盾，應修正為「…亦屬資料蒐集特定目的之必要範圍…」，請確認。

貳、文書修正(依報告書頁碼)：

- 一、部分內容字體不一，例如 P15 第一段第一行「4 月 21 日」（多空格請一併刪除）、P17(一)第一行「人工智慧」、倒數第二~五行

「使用者」、P18 第 1 點第一~二行「日常生活」、「明令禁止」、「嫌疑人」、「進行適當限制」、P21「行政罰」、「一方面」、P22「從而」、P45 2019 年 3「月」、P49 第一段第一行「1 月」，請再審視全文校正。

二、P5 最後一段第三行建議修改為「是『故』有關人工智慧應用…」，使語意通順。

三、P11 第一節第二行，重要方「面」應為重要方「向」，同段第四行「關鍵目標之一」後多一個空格。

四、P12 第 4 點第十二行「盡」量應為「儘」量，及第十三行「盡」可能應為「儘」可能。

五、P16 第二行風險導向「risk-based」較常翻譯為風險基礎，及 P61 倒數第三行「機構以『風險為本』的方式(risk-based approach)」較常翻譯為風險基礎方法，請確認。

六、P17 為避免與後續本研究提出之草案混淆，建議將 AI 法草案改為「該」草案，而非「本」草案。

七、請統一法條引用方式，如 P17(一)草案「第 3(1)條」、「P22 第 29(6)」、「第 10(5)」、「第 9(1)」、P24 倒數第四行「第 35(3)(a)條」、P26 GDPR「第 6 條第 1 項第 a 款」。

八、P23 第二~三行「AI 與個資保護指引…」漏下引號。

九、P18 第(4)點第四行，應為失「蹤」兒童。

十、P20 遵循性自評(conformity assessment)較常翻譯為符合性評估，請再確認。

十一、P22 倒數第 4 行「個資保護影響評估 (DPIA)」，請補上英文全稱。

十二、P23 提出「AI 與資料保護指引 (Guidance on AI and data Protection)」，遺漏下引號。

十三、P24 一(2)「確定風險因應措施；和」及 P25(二)(5)「資料保存期限；以及」分號後出現贅字，請刪除；此外，全文各分點後所使用的標點符號有分號(如 P24)、句號(如 P31)，亦有

未加標點符號(如 P49)，請統一用法。

十四、P38 表格最右邊第三格(例如精確率(Precision)、召回率(Recall)標點符號請調整為(例如精確率『Precision』、召回率『Recall』)。

十五、P49 標題「人工智慧管理模範框架」與同頁下文及 P81 之「人工智慧治理模範框架」翻譯不一致，請修正。

十六、P53 圖三有兩種圖示(AI 開發與部署之三階段)，請釐清修正。

十七、P55 第四點項下 1~6 項目符號階層錯誤，請改正。

十八、P58 當頁註 18 第二行「風險太 高」多空格，文末引用應為「註 12, 頁 13」。

十九、P59 當頁腳註 20 文末「頁 16」應為「頁 14」，請修正。

二十、P60 第一行末「便應為相關…」建議調整為「應為相關…」，使語意通順。

二十一、有關(human-in-the-loop)之翻譯，有翻為「人為控制」(P51、144)、「人在環中」(P62)、「人為干預」(P82)，請統一；另(human-out-of-the-loop)亦有翻譯不一致之情形(P62、82)請調整。

二十二、P71、89 答「復」應為答「覆」，P86、87、88、134 回「復」應為回「覆」，請修正。

二十三、P75 第二段最後一行「…以此大造…」應為打。

二十四、P78 英國 8,000 萬、工作人員總數達 1,200 人，請加上千分位為 8,000 萬、1,200 人。

二十五、P81 第一段第二行「新加坡」重複。

二十六、P83 第二段前面少空兩格。

二十七、P84、86、98 五大服務名稱建議與 1999 服務實施計畫一致，調整成「諮詢服務」、「轉接服務」、「告知電話」、「陳情服務」、「派工服務」。

二十八、P85 第一點第三行、P88「規畫」應為「規劃」，當頁註「受理陳情案件約占陳情案件總量約 50%」請調整數據及統計時

間為「自單一陳情系統上線以來至 111 年 11 月底止，網路管道受理案件占比為 45.7%，1999 管道為 46.2%，其他（書信、臨櫃等）管道占 8.1%。」

二十九、P91 倒數第四行(IVR)請補上英文全稱。

三十、P129 小標 10.21999 請多空格修改為 10.2 1999 以利閱讀，同頁其他小標隨同調整。

三十一、P145 表 4 請調整為同一頁以利閱讀。

「1999 臺北市民當家熱線數位轉型與個資保護對策之研究」

期末報告審查會議簽到簿

時間：112 年 1 月 4 日（星期三）10 時 0 分

地點：視訊會議

主席：黃嫻嫻專門委員

單 位	職 稱	姓 名	簽 名
臺北市政府研考會本部	專門委員	黃嫻嫻	台北通簽到 09:40:30
長慧法律事務所律師	審查委員	黃于珊	台北通簽到 09:45:25
明日科技法律事務所	審查委員	王琍瑩	台北通簽到 09:53:30
資訊工業策進會科技法律研究所	審查委員	顧振豪	台北通簽到 09:55:30
資訊工業策進會科技法律研究所	審查委員	王自雄	台北通簽到 10:04:20
臺北市政府研考會研究發展組	組長	林潔瑜	台北通簽到 09:56:13
臺北市政府研考會話務管理組	組長	陳宗亨	TAIPEION 簽到 09:49:05
臺北市政府研考會話務管理組	約聘企劃師	張堯雯	台北通簽到 09:53:28
臺北市政府研考會研究發展組	助理研究員	胡婷舜	台北通簽到 09:54:27
臺北市政府研考會話務管理組	助理研究員	許家碩	台北通簽到 09:57:08
研究團隊(達文西個資暨高科技法律事務所)	計畫主持人	葉奇鑫	台北通簽到 09:55:58
研究團隊(達文西個資暨高科技法律事務所)	協同主持人	王慕民	台北通簽到 09:51:46
研究團隊(達文西個資暨高科技法律事務所)	研究助理	李蘊家	台北通簽到 09:53:54

附錄六：期末報告審查意見回應說明

「1999 臺北市民當家熱線數位轉型與個資保護對策之研究」

委託研究案

期末報告審查意見回應說明

委員	題號	審查意見	回應說明
黃于珊委員	一	2022年10月中，美國白宮提出人工智慧權利法草案藍圖，因美國乃 AI 發展大國，建議於法規草案蒐集集中將此規範納入。	本研究依委員建議，於總結報告第三章第四節，補充該藍圖相關研究發現。
	二	報告附件一所提出之人工智慧作業指引，於第166頁明訂6項為高風險項目，但欠缺開放性，建議補充授權市府針對高風險項目進行增修之條文。	本研究依委員建議，於總結報告第六章第三節，「臺北市政府開發及應用人工智慧作業指引」草案第二點第三項第三款，針對高風險項目之開放性補充建議。
	三	報告第175頁針對以當事人同意為個人資料蒐集處理利用依據之情形，認為當事人若不同意可選擇退出，但未論及非以當事人同意作為依據時之當事人退出權，建議補充。	本研究依委員建議，於總結報告第六章第三節，「臺北市政府開發及應用人工智慧作業指引」草案第五點第二項第四款，補充當事人退出機制之建議。
	四	報告附件一提供「人工智慧系統開發及應用檢核表」，建議研究團隊後續與北市府討論，將1999市民熱線之人工智慧應用套入該檢核表進行分析。	本研究依委員建議，對1999熱線現階段擬導入之人工智慧應用，使用「人工智慧系統開發及應用檢核表」進行評估實作，並將檢核結果列於總結報告附錄一。
顧振豪委	一	據觀察，國際對人工智慧之相關規劃及草案之訂定等，皆著重於透明性及與利害關係人溝通，除報告第150頁提到依個資法向當	本研究依委員建議，於總結報告第六章第三節，「臺北市政府開發及應用人工智慧作業指引」草案第五點第一項第七款，補充強

委員	題號	審查意見	回應說明
員		事人踐行告知外，建議團隊後續研究是否有其他建議方式增強社會溝通。	化人工智慧相關社會溝通之建議。
王自雄委員	一	歐盟 AI 法之重點係建立人工智慧產品之審查標準，與本研究案重點應有區別，本研究案係以人工智慧應用之風險控管為核心，指引草案為政府對外公開政策之一，因此，除檢核表之外，如何強化與利害關係人之溝通，應為較重要之部分。	本研究依委員建議，於總結報告第六章第三節，「臺北市政府開發及應用人工智慧作業指引」草案第五點第二項第七款，補充強化人工智慧相關社會溝通之建議，並於第三章第二節第四目，補充歐盟理事會對 AI 法草案之最新修正要點。
	二	建議團隊對檢核表增列評級表，供公務機關之同仁針對檢核表逐項填覆後，以總分揭示風險及提醒應注意之程度，以利確定後續應對作為，或針對人工智慧應用作類型化處理，將本研究案之效益最大化。	本研究依委員建議，以「適法性」為風險劃分依據，於總結報告第六章第三節，「臺北市政府開發及應用人工智慧作業指引」草案檢核表增列*標註之適法性項目，並建議該等項目如有經評估有尚未符合者，應於人工智慧系統啟動開發或應用前完成改善。
王珣瑩委員	一	報告第141頁、第145頁均有臚列英國和新加坡的法規指引及案例研究之比較分析，但較缺少各該指引及案例在適用面的對照，或許是因較難找到公開資料，這部分頗為可惜，建議團隊補充說明。	本研究依委員建議，於總結報告第四章第四節，補充說明新加坡案例對新加坡 PDPC 指引之執行狀況，並補充研究限制之說明。
	二	建議與實際需求端(即市府同仁)進行交流，因實務上電話進線時，較難於短時間內完成個資告知並取得同意，甚至進線民眾可選擇略過該步驟直接進入問答，故當實際發生此情形時，法	本研究依委員建議，於總結報告第六章第三節，「臺北市政府開發及應用人工智慧作業指引」草案第五點第二項第二款，補充當事人未表示同意時之處理建議。

委員	題號	審查意見	回應說明
		律效力應如何認定?此部分牽涉到是否合規，應提出明確見解。	
臺北市 政府 研考 會 話 務 管 理 組	一	報告裡有提到法務部、勞動局等案例，不知該單位於執行作業時有無遇過法律評估上之困境，若實務上要取得內部過程不容易，建請團隊以現有檢核表或草案，利用府內資料進行評估測試，以使結果更有參考性。	本研究依委員建議，對1999熱線現階段擬導入之人工智慧應用，使用「人工智慧系統開發及應用檢核表」進行評估實作，並將檢核結果列於總結報告附錄一。
	二	若電話客服個資蒐集時皆須取得民眾個別同意，此在進線語音設計及執行上恐有困難，且電話服務節奏緊湊，僅於語音中供民眾選擇同意與否，若民眾未做選擇，可否視為有告知且民眾已同意?建請提供明確見解。	本研究依委員建議，於總結報告第六章第三節，「臺北市政府開發及應用人工智慧作業指引」草案第五點第二項第二款，補充當事人未表示同意時之處理建議。
	三	依民眾表達同意之後續撤回機制，之前已有在設想，應可設計出具可行性之方案。	感謝說明。
臺北市 政府 研考 會 研 究 發 展 組	一	請確認報告全文是否遵守學術倫理規範，例如 P3第3行「Lanx 提出…」、圖1系統風險層級示意圖應註記引用出處，P189焦點座談應隱匿發言之專家學者姓名以代碼呈現(例如:專家 A)，第八章標題應為參考文獻而非參考資源，並依學術研究報告格式製作(如 APA 格式或芝加哥格式)。	本研究依委員建議，檢視、補充參考來源，並調整引註格式。
	二	報告請依本府委託研究要點規定格式製作，增補研究摘要(包括研究目的、過程、重要發現、主要建議及政策意涵)，第七章	本研究依委員建議，補充中英文研究摘要、歷次工作會議紀錄，調整第七章之標題及建議呈現方式。

委員	題號	審查意見	回應說明
		標題請修改為「結論與建議」，其建議事項分為「立即可行之建議」及「長期性建議」，附錄增補歷次工作會議紀錄。	
	三	第二章研究方法建議與第三、四、五、六章撰寫順序一致，第一節建議改為法規及文獻分析，第二節為案例彙整，第三節為個資保護影響評估，第四節為焦點座談。	本研究依委員建議，調整第二章各項內容順序，與第三~六章撰寫順序一致。
	四	因焦點團體討論內容係為本案研究成果之一，建議第六章研究發現呈現「臺北市政府開發及應用人工智慧作業指引(草案)」以及焦點座談調查(含施作方式、訪綱及以主題式的摘述各與會專家的意見)，原第六章內容回歸於第三~五各章節中，另立小節，以利整體閱讀。	本研究依委員建議，調整第六章內容，分三節呈現指引草案之提出以及焦點座談內容，並將原第六章內容經適當調整後，列作第三章第八節、第四章第四節，以及第五章第三節。
	五	其餘報告內容、錯漏字和統一文字用語等如後附錄，請團隊修正。	本研究依委員建議，檢視並調整全文用字用語。

發行人：俞振華

發行所：臺北市政府研究發展考核委員會

地址：臺北市市府路 1 號 11 樓東南區

電話：02-27208889

影印所：達文西個資暨高科技法律事務所

地址：臺北市中正區衡陽路 51 號 6 樓之 9

電話：02-2367-0902

中華民國 112 年 2 月

初版 25 本