出國報告（出國類別：會議）

赴馬來西亞參加「AI 素養與選舉」工作坊出國報告

服務機關：中央選舉委員會

職稱姓名：葉高級分析師志成

派赴國家：馬來西亞

出國期間：113 年 11 月 17 日至 11 月 21 日

報告日期：114 年 1 月 20 日

目錄

壹、 緣起

國際民主及選舉協助機構（International Foundation for Electoral Systems）IDEA 亞太區主任 Tamang 女士邀請本會派高階主管於 113 年 11 月 18 日至 20 日馬來西亞吉隆坡艾美酒店參加「人工智慧素養與選舉」工作坊。該工作坊係探討民主等多元角度探索人工智慧之積極性作用，人工智慧、倫理議題及其在選舉管理中之運用與優勢，及潛在危害之防制策略與解決之道，並討論對弱勢團體的影響，在民主原則及法規框架下之運作方式，當前在選舉過程中之人工智慧融入選舉管理之機會，以及如何以民主為指導原則，實施人工智慧解決方案，並提供許多跨界、跨國選舉管理機關交流學習機會，並從民主價值觀、多元角度探討人工智慧的積極作用與潛在危害之解決策略。本次會議 IDEA 限本會 1 位參加，由本會綜合規劃處葉高級分析師志成代表參加。

本次工作坊 IDEA 邀請，由 Microsoft 和 Open AI 協辦，這是 Microsoft 民主前進計畫培訓研討會的第一次，後續將複製此活動，在歐洲、非洲、拉丁美洲進行交流。工作坊邀請許多國家選舉管理機關代表、民間社會組織、媒體代表，本次工作坊豐富的跨界交流機會，促進了不同領域知識的交融與創新，共同探討人工智慧、倫理議題及在選舉管理中之運用與防範潛在危害。

AI 素養與選舉工作坊 合影

貳、 工作坊課程

會議為期三天 113 年 11 月 18 日至 20 日，本次會議係由 IDEA 主辦， 3 天會議課程（如附錄）均在馬來西亞吉隆坡艾美酒店進行，議題課程如下：

一、AI 基礎介紹（11 月 18 日 9:45-11:00）

二、AI 倫理與人權（11 月 18 日 11:30-15:00）

三、AI 與社交媒體（11 月 18 日 15:30-17:00）

四、AI 選舉管理（11 月 19 日 9:00-13:00）

五、外部行為者與 AI（11 月 19 日 14:00-18:00）

六、AI 立法與法規（11 月 20 日 9:00-11:30）

每個議題 IDEA 都有指派講者來引導，IDEA 組建一個團隊講解相關議題，並致力於保護選舉安全，保護健康的資訊流通，促進民主和公民參與，基於關注選舉面臨的挑戰，尤其是網路 AI 危害、虛假訊息，本工作坊主要講者是牛津大學社會數據科學的博士候選人 Prathm Juneja，專長是關於全球 AI 收集運用，並協助 IDEA 與 Microsoft 設計本課程。課程說明如下：

一、 第 1 場：AI 基礎介紹

主講人：Mr. Prathm Juneja（IDEA，牛津大學社會數據科學的博士候選人）

課程摘要：

講師牛津大學社會數據科學的博士候選人曾經和 IDEA 合作撰寫了一份關於 AI 與選舉管理的報告，涵蓋了從其他國家如何使用 AI 到選舉管理機構如何利用 AI 改進選舉過程等各方面。因此，本次研討會正是基於這份報告發展出來的。AI 已經在全球各地影響著選舉，它的影響有很多，且這個領域的研究相對有限，目前仍然缺乏深入的討論，因此舉辦本次工作坊的原因。

AI 無處不在，大家在新聞中、工作中都經常聽到 AI 的討論，AI 可能影響人類所有領域，本項課程目標是要討論 AI 如何影響選舉。AI 的研究始於 1950 年代，GPT 模型進一步加速了 AI 的發展。目前，AI 可以生成文本、圖像、影音，甚至模仿人聲和音樂；它也具分析識別模式，處理複雜數據，並適合執行許多涉及數據識別的重複性任務。課程中進一步討論生成式 AI 與區分式 AI，以及它們在不同應用具體作為。目前常見的大型語言模型首先聯想到的是 OpenAI 的 ChatGPT 或 Facebook 的 LLaMA 模型。

以 ChatGPT 為例，訓練過程始於大規模的數據集，比如所有被數位化的圖書或整個維基百科餵給一個 AI 神經網路並訓練它，如果回答錯誤就調整其參數，模型持續更新並改進，讓它變得更加準確。語言模型不僅可以生成文字，也可以生成其它內容。然而，它們也存在所謂的「幻覺」問題，指模型可能生成虛假的或不準確的資訊。

二、 第 2 場：AI 倫理與人權

主講人：Ms. Alia Yofira（印尼安全網路實驗室）

課程摘要：

許多外部行為者已經廣泛利用人工智慧來影響選舉。因此，需要全面了解人工智慧的運作方式，AI 應用發展應確保其符合倫理並尊重人權。人工智慧常見偏見，通常可以分為三類：

（一） 數據偏見：包括數據資料無法充分代表所有群體，無法反映多樣性。

實例：某些面部識別系統因主要訓練於白人數據集，而在處理其他種族的面部特徵時表現較差。

國際語言模型訓練資料以英文為主，中文部分以簡體中文為

主。

（二） 算法偏見：模型過程中放大了數據的不平等。

（三） 部署偏見：即使人工智慧系統經過訓練，在實際應用中也可能出現偏見。

在選舉管理和公民社會組織中，人工智慧被各種行為者使用，但可能對邊緣化群體造成更多傷害。以下是人工智慧對人權與選舉的幾個關鍵風險：

（一） 種族與宗教少數群體：如選民數位身份認證系統因 AI 偏見，可能導致無法正確識別少數群體，致無法投票，影響其政治參與權及引發歧視。

（二） 女性與性別少數群體：AI 偏見可能體現在性別刻板印象，尤其女性政治人物容易成為深偽影像攻擊騷擾目標。

（三） 經濟弱勢群體：經濟落差可能限制其參與選舉的機會。

（四） 虛假資訊：如候選人利用 AI 進行誤導性宣傳。

據統計 96%的深偽影音涉及非自願的私密影像，女性與性別少數群體在此類問題上承受了更多影響，我們需要從法律與監管的角度正視這些問題。另外，人工智慧大幅消耗能源，對全球資源提取及環境問題亦須重視。

三、 第 3 場：AI 與社交媒體

主講人：Dr. Gayathry Venkiteswaran（馬來西亞諾丁漢大學）

課程摘要：

（一） 虛假資訊的有效傳播需要滿足以下三個條件：

1. 供應：需要大量的虛假資訊。
2. 傳播與放大：虛假資訊需要通過各種方式廣泛傳播。

3. 受眾接受度：需要有人願意相信這些虛假資訊。

（二） 虛假資訊的主要形式：

1. 文本型虛假資訊。
2. 音訊深度偽造。
3. 影音深度偽造。
4. 圖像偽造。

虛假資訊是本次研討會重點及各國選舉機關所憂慮，隨著科技進步，尤其是人工智慧技術的快速發展，AI 在選舉過程中的應用也日益增多。選舉活動不再僅僅依賴於傳統的競選手段，結合 AI 社交媒體如 Facebook、Twitter、Instagram、X 等，為候選人及其支持者提供了前所未有的機會，也成為選民獲取信息和表達意見的主要平台，透過社群媒體平台進行的政治宣傳、數據分析及精準推銷，已經成為影響選舉結果的重要因素之一。AI 技術可以對大量社交媒體數據進行深度分析，從中識別出關鍵話題、情感走向和民意趨勢。情感分析技術及 AI 聊天機器人能夠精確地判斷選民對特定候選人或議題的態度，並且依此調整宣傳策略，提前制定出有效的競選對策。

四、 第 4 場：AI 選舉管理

主講人：Mr. Prathm Juneja（IDEA，牛津大學社會數據科學的博士候選人）

課程摘要：

（一） 選舉的人工智慧應用：

1. 社交媒體監測及管理：AI 或許可以幫助監測這些社交媒體通訊，然而獲取社交媒體平台的數據相當困難，若只能獲取部分選舉活動的數據，可能導致監管偏向某一特定陣營，這將嚴重損害選舉公平性，同時也需考慮言論自由的影響。
2. 虛假資訊監測：技術上結合大型語言模型和圖神經網路來檢測以識別虛假資訊，目前的研究主要集中於英語，但仍面臨非常困難技術挑戰，尤其是社交媒體平台數據取得。
3. 使用光學字符識別（OCR）和光學標記識別（OMR）技術：這些 AI 技術可用於提高選舉數據處理的準確性，但仍需進一步的測試。

（二） AI 對選務影響：

1. 生成式 AI 會產生錯誤訊息、虛假資訊。
2. 易遭入侵誤用，獲得選民的資訊。
3. AI 在數據分析方面的應用，尤其是利用大數據和機器學習技術，能夠對選民的行為進行深入研究。AI 技術可以根據選民的年齡、性別、地理位置、興趣以及以往的投票行為等資料，精確推測出其可能的選擇傾向。AI 分析技術能夠精確地判斷選民對特定候選人或議題的態度，並且依此調整宣傳策略。這使得候選人能提前制定出有效的競選對策。
4. AI 深度學習技術，已經可以生成極為真實的虛假新聞和影音（例如 Deepfake 技術）。這些技術可以創造出偽造的候選人發

言或行為，並通過社交媒體迅速傳播，誤導選民，甚至抹黑對手。AI 生成的虛假資訊不僅能夠快速散播，還能夠根據受眾的反應進行動態調整，極大地提高其有效性和影響力，尤其是女性與邊緣群體從政傷害。

5. 聊天機器人蒐集選民個人化訊息：聊天機器人可以自動與選民互動，回應問題，提供定制化的候選人資訊，甚至進行心理輿論的引導。這些系統不僅可以提供選民認知，還能分析選民的需求，從而生成有針對選民個性化宣傳。
6. 精準廣告推送：AI 技術可以對大量社交媒體數據進行深度分析，從中識別出關鍵話題、情感走向和民意趨勢。也可以利用深度分析資訊，候選人可以將特定訊息或廣告精準地推送給相對應的選民群體，從而提高宣傳效果並對選舉結果產生影響。
7. 社群、AI、聊天機器人競選經費難管制。
8. 外國介選：例如有報導俄羅斯用 AI 製作的深偽影音，賀錦麗在尚比亞打獵。
9. 資訊安全：AI 工具能幫助惡意行為者實施系統攻擊，AI 幫助駭客更容易製造新病毒，更容易產生釣魚或社交工程資訊，且更難分辨真偽，使相關攻擊變得更容易成功。

（三） AI 虛假深偽防治制作法：

1. 性影像處理中心 NCII 平台合作：親密影像通知平台業者移除違法影像。
2. 拍攝照片時自動產生內容憑證，辨識是否篡改，目前研究將這些技術嵌入相機。
3. 提高公眾意識

（1） 教育公眾瞭解人工智慧可能被濫用，未經驗證內容應持懷疑

態度。

（2） 鼓勵平臺標記人工智慧生成內容，以增強透明度。

4. 利用 AI 對抗 AI

（1） 開發工具檢測並標記虛假資訊。

（2） 開發全面的檢測工具實務上非常困難。

5. 加強多方合作：政府、科技公司與民間社會需要協作，共同建立強有力的規範和審查機制。

6. 提供準確資訊：通過可信管道傳播正確的選舉資訊，以對抗虛假資訊的影響。

五、 第 5 場：「外部行為者與 AI」

主講人：Mr. Prathm Juneja（IDEA，牛津大學社會數據科學的博士候選人）

課程摘要：

本課程請部分國家選舉管理機構分享AI使用或困境：

（一） 菲律賓：

1. 2022 年選舉官員和新員工點擊了釣魚郵件，造成資料洩露的風險。
2. 外國干涉：選舉機關難查核誰資助濫用 AI 行為者，有可能國外特定團體資金。
3. 規劃制定「深偽責任與透明法」要求影音媒體中披露相關內容。
4. 2022 年制定「公平選舉應用規範」，規範社交媒體和網路付費廣告。
5. 計劃成立新的工作小組「選舉真相、公正與正義工作小組」，進一步推動選舉中的透明與公正。

（二） 印尼：

1. 已經有 AI 生成虛假資訊，例如影音或虛擬角色被用來傳播誤導性訊息，或者修改人物形象。
2. AI 在高風險領域（如選舉）中被大量使用，但相關法規尚未出現。
3. 選民數據曾外洩。
4. 外來人士可能影響本地選舉，例如外地選民的名單被利用，導致不居住在該州的人決定該州的選舉結果，這對民主機制構成了嚴重操控。
5. 候選人非法使用他人個人數據進行登記曾引發重大爭議。
6. 目前缺乏數位服務規範，來規定數位平台、社交媒體公司。

（三） 韓國：計劃引入 AI 選舉法平台，彙集選舉法規、最高法院裁決、規範和案例解釋，透過整合這些信息，AI 可建立一個集中式知識庫，隨時提供詳細答案參考。

（四） 斯里蘭卡：計劃引入選民登記系統、選舉結果的統計與宣傳民眾知悉選舉結果。

（五） 巴基斯坦：計劃引入 AI 結合圖像處理，對選民圖像進行比對，確保選民名單正確性。

六、 第 6 場：AI 立法與法規

主講人：Ms. Alia Yofira（印尼安全網路實驗室）

Mr. Sebastian Becker（IDEA）

課程摘要：

AI 深度學習技術，已經可以生成極為真實的虛假新聞和影音，通過社交媒體迅速傳播，誤導選民，甚至抹黑對手，極大地提高選舉影響力。是以 AI 立法更顯重要，AI 技術在選舉中的應用雖然提高了宣傳效率，

但也帶來了不少道德問題。首先，過度利用 AI 進行選民情感操控，可能會引發對選舉公平性的質疑。選民的選擇應該基於真實信息和理性分析，而非受到無形的 AI 數據操作引導。其次，虛假資訊和深度偽造技術的廣泛使用，也可能加劇社會分裂，造成不必要的政治對立。

目前，對於 AI 在選舉中應用的法律監管仍處於初步階段。在某些國家，社交平台的選舉干預已經成為政策重點，並開始對虛假新聞、濫用個人數據等問題進行立法。然而，由於 AI 技術的快速發展，現有的法律框架很難有效應對新型的政治操控行為。對 AI 在選舉中的監管，需要政府和社會各界的密切合作，確保選舉的公平性與透明度。

AI 技術的進步使得選舉過程中可以進行更加精準且高效影響與操作，但這同時也帶來了對選舉公正性和選民意願的挑戰。為了保障民主制度的運行，必須對 AI 技術在選舉中的應用進行更加嚴格的監管，避免其被不當利用。因此，我們必須加強 AI 立法作為：

（一） AI 應用的法律監管，確保選舉過程中的數據使用不會被濫用。

（二） 推動透明度與公正性：要求候選人公開其使用 AI 進行宣傳的方式及其數據來源。

（三） 加強選民教育：提高公眾對 AI 技術及其可能帶來的影響的認識，防範被虛假資訊所誤導。

現行的法律框架，如人權法、反歧視法及隱私保護法，也可應用於解決與 AI 有關的挑戰。探討現有法律工具如何幫助應對 AI 帶來的問題。以國際人權法為例，AI 可能影響表達自由、隱私權等基本權利，可以應用於對抗虛假資訊或限制言論的問題。 2016 年以來，全球各地均有呼應 AI 立法，我們已看到部分國家法規試圖應對 AI 帶來的挑戰例如歐盟、巴西等。另外從立法方向，中國的 AI 政策目標是促進國內 AI 的發展，歐洲的 AI 政策則專注於實現 AI 的善用與普及，美國的政策

旨在加速領導地位的提升。

參、 心得及建議

一、 人工智慧回應與臺灣文化或價值觀不符：目前國際語言模型訓練資料以英文為主，中文部分以簡體中文為主，我國民眾使用國際通用人工智慧，會產生回應簡體中文文化之情形，並可能使兒童或年輕族群因人工智慧回應簡體中文文化所誤導，為保護本國繁體中文及文化永續發展，我國政府自 112 年 4 月起推動 TAIDE 大型繁體中文語言模型，讓我國人民可以獲取正確臺灣文化、選務資訊與價值觀，本會亦已規劃導入 TAIDE 大型繁體中文語言模型。

二、 人工智慧因數據偏見，影響女性、少數群體、經濟弱勢選舉公平性及人權：本次工作坊 AI 倫理與人權課程，強調 AI 對選舉業務及社會發展深遠影響， AI 選舉偏見極可能對對選舉公平性及人權產生重大傷害，尤其是女性、少數群體、經濟弱勢群體參與者，國外選民制度大都採登記制，影響較大，我國雖採戶籍制，但仍應重視 AI 對選舉公平性及人權影響。

三、 虛假訊息或深偽影音對選舉危害：AI 虛假資訊及深偽影音影響選舉結果，是各國選舉機關最關注且難以管理議題，利用 AI 開發全面的檢測工具，實務上非常困難且難以全面防制，但身為選務管理機關仍有必要結合第 3 方事實查核中心及各社群平台，儘可能杜絕虛假訊息或深偽影音，或者利用 AI 技術及可信管道傳播正確的選舉資訊，以對抗虛假資訊的影響。

四、 社群經費透明化管理、虛假訊息或深偽影音利用社群影響選舉結果：從馬斯克 X 平台對 2024 美國大選選舉結果之影響，預期未來選舉將著重於社群競選，尤其是虛假訊息或深偽影音利用社群影響選舉結果，候選人競選經費必須管制與透明化，並杜絕外國勢力利用社群介選，惟相關防制作為仍須仰賴各大社群自制及托播廣告透明化，

惟近日（114 年 1 月 7 日）社群媒體巨擘 Meta 宣布將結束在美國的第 3 方事實查核計畫，將使政治虛假資訊利用社群介選難以管控。

五、 強化選務資訊安全及選民個資保護將更加重要：AI 聊天機器人及社群蒐集選民個人化訊息，利用 AI 技術對大量社交媒體數據進行深度分析、民意趨勢，進行精準廣告推送，從而對選舉結果產生重大影響。利用 AI 產製新病毒、或者更難分辨真偽之社交工程電子郵件或網站，從而竊取選民個資，造成選民投票意願遭誤導，難以分辨理想候選人，影響選舉公平性，因此強化選務資訊安全及選民個資保護將更加重要。

六、 立法防止 AI 成為不正當競爭的工具，避免顛覆選舉公平及人類的未來：選舉作為民主制度的基石，必須保證其過程的公正性，AI 易遭有心人惡意運用，成為不正當競爭的工具，為維護選民的真實意願與選擇權，AI 立法限制有心人惡意運用，還可規範 AI 改變您我生活，且不會顛覆人類的未來。此外，我國 112 年 6 月 12 日修正總統副總統選舉罷免法第 47 條及公職人員選舉罷免法第 51 條規定，規範擬參選人、候選人 AI 深偽影音處理流程，即是體現 AI 立法第 1 步。

七、 其他國家選舉管理機構分享 AI 使用或困境，作為本會選務 AI 資訊規劃借鏡：其他國家選舉管理機構分享為本次會議帶來了豐富的經驗和啟發，鑑於 AI 在全球範圍內的快速發展，其對社會、經濟以及政治的影響也越來越深遠參與，瞭解其他國家選舉管理機構對人工智慧於選務方面之應用及趨勢，及各國防制 AI 選務虛假訊息作法，可作為本會選務 AI 資訊化規劃及因應選務虛假訊息作法參考。

最後，本工作坊研討讓本會瞭解未來選舉過程中，AI 技術已展現出選務虛假資訊傳播、選民行為分析、精準廣告推送、資訊安全及選民個保護等

相關議題，也瞭解 AI 技術的應用及其背後的挑戰，並思考如何在推動 AI 的同時，同步避免 AI 對選務的危害，借鏡各國選舉管理機構對 AI 的管理，捍衛我國民主選舉制度。本會於工作坊結束後，並於 113 年 12 月 19 日至 12 月 20 日舉行之「113 年選務發展及 AI 應用研習會議」中，特別於會議研習課程中，增列馬來西亞 AI 素養與選舉工作坊心得分享，將本次出國心得及知識分享於本會及直轄市、縣（市）選舉委員會主管，共同集思廣義推動 AI 於選務運用及思考防範 AI 所產生選舉危害，共創永續選舉正義與公平。

附錄、工作坊課程表

AI Literacy for Electoral Actors Workshop

International IDEA AI Literacy for Electoral Actors Workshop

Implemented by International IDEA and funded by Microsoft and Open AI

18-20 November 2024

Le Méridien Hotel in Kuala Lumpur, Malaysia

DAY 1: Monday, 18 November 2024	
9:00 – 09:45	Welcome remarks & Presentation of the Workshop
09:45 – 11:00	Module 1: Basics of AI *The session covers the basic technical details of modern AI systems, providing an overview of Large Language Models (LLMs) and transformer models, including their functionalities. It also focuses on the current applications of AI and key issues related to modern AI systems.*
11:00 – 11:30	Coffee Break

11:30 – 13:00	Module 2: AI Ethics & Human Rights *The session explores the key principles of ethical AI, common biases in AI systems, and their sources, providing examples in each area. It examines how the development and use of AI systems can impact human rights and perpetuate and amplify existing gender inequalities.*
13:00 – 14:00	Lunch
14:00 – 15:00	Module 2 (cont.): AI Ethics & Human Rights
15:00 – 15:30	Coffee Break
15:30 – 17:00	Module 3: AI & Social Media *The session provides an overview of the basics of how social media platforms operate, their user demographics, and the mechanisms of information dissemination. It also addresses key topics at the intersection of elections and technology.*

DAY 2 – Tuesday, 19 November 2024	
09:00 – 11:00	Module 4: AI & Electoral Management *Attendees will be able to scope and evaluate which AI use cases for election administration are most useful for their specific election environments and understand the potential harms & benefits of those use cases, as well as ways to mitigate some of those harms.*
11:00 – 11:30	Coffee Break
11:30 – 13:00	Module 4 (cont.): AI & Electoral Management
13:00 – 14:00	Lunch
14:00 – 15:30	Module 5: External Actors & AI *The session explores the impact of AI on elections through content generation, with a focus on approaches sensitive to gender and marginalized communities. Participants will gain skills to identify and address disinformation, misinformation, and hate speech, understanding how they are created and evaluating their significance within their country's context.*

15:30 – 16:00	Coffee Break
16:00 – 18:00	Module 5 (cont.): External Actors & AI

DAY 3 – Wednesday, 20 November 2024	
9:00-11:30	Module 6: Legislation and Regulation *The session covers the broader regulatory landscape around AI, specifically in the context of elections, to help participants understand how emerging regulations may impact their work.*
11:30 – 11:45	Coffee Break
11:45 – 13:00	Concluding Remarks
13:00 – 14:00	Lunch