

## 國際人工智慧公約——創設全球人工智慧治理架構

鄒鶴慈 編譯

### 摘要

由歐洲理事會所主導之首個具法律約束力的國際人工智慧公約，旨在填補人工智慧技術快速發展下可能造成之法律真空，並確保各國使用人工智慧系統活動符合國際人權法之義務。該公約內容具有彈性，使各國得依據自身情況基於善意原則履行公約內容，但也必須確保人工智慧系統侵犯人權時，能採取有效的救濟措施。該公約提供一個全球架構，以因應人工智慧對人權、民主及法治所構成的風險，並致力於協調不同國家法規的挑戰，進一步促進締約方在人工智慧治理的合作，幫助全球建立一致的法律架構及標準。

(取材資料：Rhys Dipshan, *International AI Treaty Signed by US, UK, and EU Leaves Much to Be Determined*, AMERICAN LAWYER MEDIA (Sept. 11, 2024), <https://www.law.com/legaltechnews/2024/09/11/international-ai-treaty-signed-by-u-s-u-k-eu-leaves-much-to-be-determined/>.)

今 (2024) 年 9 月 5 日，以色列、美國、英國、歐盟及其他多個歐洲國家簽署了一項由歐洲理事會提出的人工智慧公約<sup>1</sup>。其正式名稱為《歐洲理事會人工智慧與人權、民主及法治架構公約 (Council of Europe Framework Convention on Artificial Intelligence and Human Rights, Democracy and the Rule of Law)》<sup>2</sup> (以下簡稱公約)，旨在促進人工智慧治理<sup>3</sup>。該公約被視為「首個具有法律約束力的國際人工智慧公約」，由 46 個歐洲理事會成員國、歐盟及其他 11 個非歐盟成員國，包括澳洲、加拿大、以色列、日本、墨西哥與美國等共同談判而成<sup>4</sup>。此公約包含參與國家制定或維持相關措施，如法律、規則或行政措施<sup>5</sup>。鑑於制定一套全球之人工智慧法律架構面臨極大挑戰，該公約展現了其在適應不同國家與文化需求上的可塑性。

<sup>1</sup> *Council of Europe Opens First Ever Global Treaty on AI for Signature*, COUNCIL OF EUR. (Sept. 5, 2024), <https://www.coe.int/en/web/portal/-/council-of-europe-opens-first-ever-global-treaty-on-ai-for-signature>.

<sup>2</sup> Council of Europe Framework Convention on Artificial Intelligence and Human Rights, Democracy and the Rule of Law, Sept. 5, 2024 C.E.T.S. No. 225 [hereinafter The Framework Convention].

<sup>3</sup> *What Does the World's First International AI Treaty Include?*, AI MAGAZINE (Sept. 5, 2024), <https://aimagazine.com/articles/what-does-the-worlds-first-international-ai-treaty-include>.

<sup>4</sup> *Council of Europe Opens First Ever Global Treaty on AI for Signature*, COUNCIL OF EUR., *supra* note 1.

<sup>5</sup> The Framework Convention, *supra* note 2, art. 1.2.

以下先介紹國際人工智慧公約之性質，接著說明該公約之重點規範，並分析其所面臨之挑戰與影響，最後作一結論。

## 壹、國際人工智慧公約之特性

該公約與歐盟全面且具體規範的「人工智慧法 (AI Act)」截然不同<sup>6</sup>，此公約本質上是國家必須遵守的一些上位原則，但如何確實地達成並執行這些原則，以及這些原則之適用範圍，主要取決於各國的自主裁量權<sup>7</sup>。本公約設有廣泛的豁免條款，如國家安全利益、國防及研發階段之豁免，但目前缺乏具體的執行機制。儘管其條款之設計提供各國相當大的自主裁量權，使各國法規的統一性相對降低，仍有論者認為該公約以人權為中心之主張，為未來全球人工智慧之規範標準奠定了基礎，建立了人工智慧管制的共同語言<sup>8</sup>。對歐洲來說，這進一步推動了「布魯塞爾效應 (The Brussels Effect)」<sup>9</sup>，使其他國家採用與歐盟一致之人工智慧監管及風險管理原則。

## 貳、國際人工智慧公約之介紹

### 一、人工智慧系統之定義

該公約對何謂人工智慧目前已有定義，這個定義追隨了歐盟「人工智慧法」中的定義，而歐盟「人工智慧法」又借鑒了經濟合作暨發展組織 (Organisation for Economic Co-operation and Development, OECD) 的定義<sup>10</sup>。依據 OECD 之規定<sup>11</sup>，「人工智慧系統 (artificial intelligence system)」是指一種基於機器 (machine-based) 的系統，該系統為了明確或隱含的目標，根據其所接收的輸入資料推斷如何生成輸出之結果，例如未來預測、內容生成、偏好建議或決策。

在日常生活中所使用之科技產品經常會運用人工智慧，使得產品使用更加便利，如語音辨識、圖片生成、導航系統等，皆是人工智慧應用之實例。

<sup>6</sup> Regulation (EU) 2024/1689, 2024 O.J. (L series).

<sup>7</sup> Council of Europe, Explanatory Report to the Council of Europe Framework Convention on Artificial Intelligence and Human Rights, Democracy and the Rule of Law art. 4, para. 38, Sept. 5, 2024, C.E.T.S. No.225 [hereinafter The Explanatory Report].

<sup>8</sup> *Id.* para. 5.

<sup>9</sup> Anu Bradford, *The Brussels Effect*, 107 NW. U. L. REV. 1 (2012); 關於布魯塞爾效應詳參：莊東霖，歐盟人工智慧法對國際準則之影響，經貿法訊，331 期，頁 13，2024 年 6 月 11 日，<https://tradelaw.nccu.edu.tw/epaper/no331/3.pdf>。

<sup>10</sup> The Explanatory Report, *supra* note 7, art. 2.

<sup>11</sup> Org. for Econ. Coop. & Dev. [OECD], *Recommendation of the Council on Artificial Intelligence*, OECD Doc. OECD/LEGAL/0449 [hereinafter OECD AI] (May 3, 2024).

## 二、國際人工智慧公約之規範

本公約之規範共八章，分別為總則、一般義務、與人工智慧系統生命週期內的活動相關之原則、救濟措施、評估與減緩風險及不利影響、公約的實施、後續機制與合作、以及最終條款，以下將統整公約重點內容，並列點說明。

### (一) 適用範圍之規範及豁免

#### 1. 適用範圍之規範

本公約的範圍涵蓋人工智慧系統生命週期 (AI system lifecycle) 內可能干擾人權、民主及法治的活動<sup>12</sup>，適用於公部門或代表其行事之私部門在人工智慧系統生命週期內所進行的活動<sup>13</sup>，同時亦責成各國解決私部門使用該系統所產生的風險與影響<sup>14</sup>。

#### 2. 適用範圍之豁免

該公約不適用涉及國家安全利益或任何與國防有關的人工智慧系統<sup>15</sup>，亦不適用於尚未可供使用的人工智慧系統之研發活動，除非該人工智慧系統對人權、民主及法治構成風險<sup>16</sup>。然而，此公約呼籲各締約方在適當情況下建立監理沙盒，以便在主管機關的監督下開發、實驗與測試人工智慧系統<sup>17</sup>。

### (二) 保護人權之義務與救濟措施

締約方應採取或維持相關措施以確保人工智慧系統的運用符合國際法及其國內法所規定保護人權的義務<sup>18</sup>，並在使用人工智慧時尊重人類尊嚴及個人自主權<sup>19</sup>，且保證人工智慧系統不被用來破壞民主制度與民主程序的完整性、獨立性及有效性的措施，包括權力分立原則、尊重司法獨立與訴諸司法的權利<sup>20</sup>，

<sup>12</sup> 「人工智慧系統生命週期」分為四個階段：(1)設計、資料與模型 (design, data and models) 階段；(2)驗證並確認 (verification and validation) 其有效性之階段；(3)部署階段；(4)作業及監控 (operation and monitoring) 階段，上述階段通常是以迭代方式進行的，且不一定是按順序進行，並在作業與監控階段的任一時間點，皆可能做出放棄使用人工智慧系統的決定。OECD AI, *supra* note 11 (May 3, 2024).

<sup>13</sup> The Framework Convention, *supra* note 2, art. 3.1(a).

<sup>14</sup> *Id.* art. 3.1(b).

<sup>15</sup> *Id.* arts. 3.2, 3.4.

<sup>16</sup> *Id.* art. 3.3.

<sup>17</sup> *Id.* art. 13.

<sup>18</sup> *Id.* art. 4.

<sup>19</sup> *Id.* art. 7.

<sup>20</sup> *Id.* art. 5.1.

並考慮對人權、民主及法治的不利影響<sup>21</sup>，以實現公平、公正及平等的結果<sup>22</sup>。締約方應在國際義務所要求的範圍內，並依照其國內法律制度，確保能提供可行且有效的救濟措施，以應對人工智慧系統造成的侵犯人權行為<sup>23</sup>。

### (三) 公約之風險管理原則與實施

締約方必須針對人工智慧系統的特定情境與風險制定透明度與監督要求<sup>24</sup>，並要求各國識別、評估、預防與調和人工智慧對人權、法治及民主影響的風險，但具體做法由各國自行決定，賦予其一定靈活性<sup>25</sup>。例如，該公約指出，各國措施應根據具體情況進行分級與區分，並在適當情況下考慮人工智慧的應用背景、預期用途及其潛在影響，而後納入相關利害關係人之觀點<sup>26</sup>。基於現有人權之保障，本公約的任何內容均不得解釋為限制、減損或以其他方式影響締約方相關法律或任何其他相關國際協定，亦不得影響締約方給予比本公約規定更廣泛的保護措施的可能性<sup>27</sup>。

### (四) 後續機制與合作

本公約由締約方代表組成締約方大會<sup>28</sup>，各方應定期諮商以促進公約的有效實施與執行<sup>29</sup>，並鼓勵締約方酌情協助非締約方依公約之條款行事，藉此促進公約在國際場域之實施，並使該些國家最終加入公約<sup>30</sup>。此外，簽署條約的國家必須在兩年內定期向締約方大會通報其遵守公約的情況<sup>31</sup>，各締約方也必須指定一個或多個、獨立與公正的機構來監督自身之法遵情況<sup>32</sup>。

## 參、國際人工智慧公約面臨挑戰及影響

各國就如何解決人工智慧帶來的風險和挑戰達成共識，一直是一大挑戰。為實現此一目標，公約所展現之廣泛、可調整的彈性或將是凝聚共識之必要條件。只要就人工智慧相關的基本原則達成共識，即使公約涵蓋範圍廣泛，其仍

---

<sup>21</sup> *Id.* art. 13.

<sup>22</sup> *Id.* art. 10.2.

<sup>23</sup> *Id.* art. 14.1.

<sup>24</sup> *Id.* art. 8.

<sup>25</sup> The Explanatory Report, *supra* note 7, para. 16.

<sup>26</sup> The Framework Convention, *supra* note 2, art. 16.2.

<sup>27</sup> *Id.* art. 22.

<sup>28</sup> *Id.* art. 23.

<sup>29</sup> *Id.* art. 23.2(a).

<sup>30</sup> *Id.* art. 25.1.

<sup>31</sup> *Id.* art. 24.

<sup>32</sup> *Id.* arts. 26.1, 26.2.

然可以成為全球人工智慧對話之基礎。各國在達成共識後根據自身的法律體系來實施相關規範，不需要完全遵循歐盟「人工智慧法」之條文內容來執行以符合該法之規範。但無庸置疑地，該公約與「一般資料保護規則 (General Data Protection Regulation, GDPR)」與歐盟「人工智慧法」等歐盟現行法律一樣<sup>33</sup>，均將人權議題視為關注的核心焦點，皆是許多歐洲法律架構之基石。

#### 肆、結論

目前雖有眾多國家簽署該公約，但其國會均未批准。在其國會批准公約後，這些國家必須即時向歐洲理事會提交一份聲明<sup>34</sup>，說明他們將如何履行公約義務<sup>35</sup>，以及如何將其適用於私部門的人工智慧系統活動<sup>36</sup>。而國際人工智慧公約提供了國際間監管人工智慧系統之藍圖，標示著國際社會在規範人工智慧方面的里程碑，也強調在科技創新的同時，保護人權及民主價值之必要性。各國如何利用公約之架構制定更加具體且有效的法規，以應對人工智慧所帶來的挑戰，將是全球人工智慧治理的關鍵。

---

<sup>33</sup> Regulation (EU) 2016/679, 2016 O.J. (L119) 1.

<sup>34</sup> The Explanatory Report, *supra* note 7, para. 31.

<sup>35</sup> The Framework Convention, *supra* note 2, art. 3.1(b).

<sup>36</sup> The Explanatory Report, *supra* note 7, para. 30.