人工智慧基本法草案總說明

人工智慧技術近年發展快速，被世界普遍認為可為整體產業與社會活動帶來廣泛之經濟及社會效益，並為我國企業及國家發展提供關鍵之競爭優勢。在氣候變遷、環境、醫療、金融、交通、內政、農業、公共服務等對民眾具廣泛影響力之領域中，更亟需積極採用人工智慧技術以推動數位轉型與永續發展。

人工智慧技術雖帶來經濟及社會效益，同時也可能對個人或社會帶來風險或影響。鑑於人工智慧技術創新之速度及可能面臨之挑戰，全球主要國家皆致力在不妨礙技術發展下，尋求建立人工智慧之治理方針與原則。經濟合作暨發展組織（Organisation for Economic Cooperation and Development，OECD）於二〇一九年五月通過「人工智慧建議書」（OECD Recommendation on Artificial Intelligence），提出基本價值原則，並給予各國政策制定者相關建議；同年歐盟發布「可信賴AI倫理準則」（Ethics Guidelines for Trustworthy AI），確保人工智慧發展所需之共同倫理原則。於此之後，如歐盟於二〇二一年提出「人工智慧法」（Artificial Intelligence Act），二〇二四年通過審議，韓國於二〇二四年通過「人工智慧發展及建立信賴基盤之基本法」，美國於二〇二二年發布「AI權利法案藍圖」（Blueprint for an AI Bill of Rights），英國於二〇二四年發布「生成式人工智慧治理框架」（Generative AI Framework）、新加坡於二〇二四年發布「生成式AI治理架構」（Model AI Governance Framework for Generative AI），皆著重於建立人工智慧技術發展之原則並建立大眾信任。

綜合考量各國做法並契合我國國情，我國人工智慧發展之基本法律應以鼓勵創新為主軸，同時兼顧人民權益保障，從基本原則、政府推動重點等構面提出基本價值、治理原則及施政方針，透過確立人工智慧研發與應用之基本原則，定錨國家重要發展方向，在促進產業發展與風險管理間取得平衡，進而提升國家競爭力，爰擬具「人工智慧基本法」草案，其要點如下：

一、本法之立法目的。（草案第一條）

二、本法所稱人工智慧定義。（草案第二條）

三、人工智慧研發與應用之基本原則。（草案第三條）

四、政府應推動人工智慧研發、應用及基礎建設。（草案第四條）

五、各目的事業主管機關得針對人工智慧創新產品或服務，建立或完備人工智慧創新實驗環境。（草案第五條）

六、政府得以公私協力方式，與民間合作，推動人工智慧創新運用，並應推動人工智慧相關之國際合作。（草案第六條）

七、政府應持續推動人工智慧教育。（草案第七條）

八、政府應避免人工智慧之應用造成違法情事，相關機關得提供或建議評估驗證之工具或方法。（草案第八條）

九、數位發展部應推動人工智慧風險分類框架，各目的事業主管機關應視需要，循風險分類框架訂定以風險為基礎之層級管理規範。（草案第九條）

十、政府得透過法令或指引建立標準、驗證、溯源或問責機制，強化人工智慧決策之可驗證性及人為可控性。（草案第十條）

十一、政府應就高風險人工智慧之應用，明確其責任歸屬及歸責條件，並建立其救濟、補償或保險機制。（草案第十一條）

十二、政府應保障勞工權益，就人工智慧利用所致之失業者應輔導就業。（草案第十二條）

十三、政府於人工智慧研發及應用過程應保護個人資料。（草案第十三條）

十四、政府應提升人工智慧使用資料之可利用性、品質與數量，以利人工智慧訓練及產出結果維繫國家多元文化價值及維護智慧財產權。（草案第十四條）

十五、政府公務使用人工智慧之應遵行事項。（草案第十五條）

十六、政府應於本法施行後依本法規定檢討及調整主管之職掌、業務、法規或訂定指引，及法規訂修前之解釋及適用方式。（草案第十六條）

人工智慧基本法草案

條文	說明
第一條　為促進以人為本之人工智慧研發與應用，保障國民人格尊嚴及權利，提升國民生活福祉、維護國家主權、安全及文化價值，增進永續發展及國家競爭力，特制定本法。	一、本法之立法目的。 二、人工智慧為攸關國家發展之科技，為積極研發與應用人工智慧，強化與深耕以人為本（human-centered）之人工智慧技術，促進技術應用與產業發展，同時保障憲法規定之人格尊嚴及國民權利，包括生命、身體、健康、安全等相關權利，以期人工智慧可回應人文與社會發展所需，邁向社會永續發展。因此，研發與應用人工智慧之同時，有賴於制定具有指標與引導性原則之法律，以作為發展人工智慧之規範與促進應用之法源基礎，爰制定本法。
第二條　本法所稱人工智慧，指具自主運行能力之系統，該系統透過輸入或感測，經由機器學習及演算法，可為明確或隱含之目標實現預測、內容、建議或決策等影響實體或虛擬環境之產出。	參考美國國家AI創新法案（National AI Initiative Act of 二〇二〇）美國法典（U.S. Code）第九四〇一章、國際標準化組織（ISO）及國際電工委員會（IEC）聯合制定技術規範（ISO/IEC）四二〇〇一：二〇二三人工智慧管理系統、美國國家標準暨技術研究院（National Institute of Standards and Technology，NIST）AI風險管理框架（Artificial Intelligence Risk Management Framework），以及歐盟人工智慧法（Artificial Intelligence Act）對於人工智慧系統之定義，人工智慧必須被設計為具備不同程度之自主運行能力（AI systems are designed to operate with varying levels of autonomy），透過輸入（input）或感測（sensing），經過機器學習（machine-learning）及演算法（algorithms），可為明確（explicit）或隱含（implicit）之特定目的（objectives）實現諸如預測、內容、建議或決策（such as predictions,content, recommendations,or

	decisions）等影響實體或虛擬環境之產出，與其他軟體系統有別，爰於本條定明人工智慧之定義。
第三條　人工智慧之研發與應用，應在兼顧社會公益及數位平權之前提下，發展良善治理，並遵循下列原則： 一、永續性：應兼顧社會公平及環境永續，降低可能之數位落差，使國民適應人工智慧帶來之變革。 二、人類自主性：應支持人類自主權，尊重人格權等個人基本權利與文化價值，並允許人類監督，落實以人為本並尊重法治及民主價值觀。 三、隱私保護及資料治理：應妥善保護個人資料隱私，避免資料外洩風險，並採用資料最小化原則；在符合憲法隱私權保障之前提下，促進非敏感資料之開放及再利用。 四、安全性：人工智慧研發與應用過程，應建立資安防護措施，防範安全威脅及攻擊，確保其系統之穩健性及安全性。 五、透明及可解釋性：人工智慧之產出應做適當資訊揭露或標記，以利評估可能風險，並瞭解對相關權益之影響，進而提升人工智慧可信任度。 六、公平性：人工智慧研發與應用過程中，應盡可能避免演算法產生偏差及歧視等風險，不應對特定群體造成歧視之結果。 七、可問責性：確保人工智慧研發與應用過程中不同角色承擔相應之責任，包含內部治理責任及外部社會責任。	一、我國發展人工智慧應衡平創新發展與可能風險，以回應國內人文及社會所需。爰參考國際協議及各國相關政策方針、法規或行政命令，訂定具有指標與引導功能之基礎準則，以作為人工智慧研發與應用之基本原則。 二、人工智慧研發與應用應兼顧社會公平與環境、經濟之協調發展，以追求對人類及地球有益之結果，從而促進永續發展（sustainable development），爰參考 G7 廣島 AI 國際行動規範（Hiroshima Process Code of Conduct for Organizations Developing Advanced AI Systems），於第一款定明永續性原則。 三、人工智慧研發與應用應在人工智慧系統之整個生命週期中尊重法治、人權及民主價值觀，為此，參考經濟合作暨發展組織（OECD）二〇一九年公布之人工智慧建議書（OECD Recommendation on Artificial Intelligence），於第二款定明人類自主性原則，應支持人類自主權（Human Autonomy），並尊重人格權（含姓名、肖像、聲音）等個人基本權利與文化價值，確保以人為本之基本價值。 四、人工智慧發展仰賴大量資料，惟資料之蒐集、處理及利用，能否確保資料安全與個人資料隱私，係目前人工智慧發展最多討論與疑慮之議題。爰參考美國二〇二二年 AI 權利法案藍圖（Blueprint for an AI Bill of Rights），於第三款定明隱私保護及

	資料治理原則，人工智慧研發與應用，應妥善保護個人資料，避免資料外洩風險，並採用資料最小化原則，而所謂資料最小化原則（data minimization），係指各階段蒐集之個人資料，皆須適當且具相關性，並僅止於符合資料處理目的所需之程度。同時，在符合憲法隱私權保障之前提下，促進非敏感（非個人或機敏）資料之開放及再利用。 五、人工智慧研發與應用應確保系統穩健性（robustness）與安全性，爰參考美國二〇二二年AI權利法案藍圖及新加坡二〇二四年生成式AI治理架構（Model AI Governance Framework for Generative AI），於第四款定明安全性原則，以防範人工智慧有關安全威脅與攻擊。 六、人工智慧所生成之決策對於利害關係人有重大影響，須保障決策過程之公正性。人工智慧研發與應用階段，應致力權衡決策生成之準確性，並提升可讓使用者及受影響者理解其影響及決策過程之可解釋性，兼顧使用者及受影響者權益。爰參考歐盟二〇一九年可信賴AI倫理準則（Ethics Guidelines for Trustworthy AI），於第五款定明透明及可解釋性（Transparency and Explainability）之原則。 七、人工智慧研發與應用須公平、完善，且演算法應避免產生偏差或歧視之結果，爰參考美國二〇二二年AI權利法案藍圖，於第六款定明公平性原則（Fairness），強調應重視社會多元包容，避免產生偏差與歧視等風險。 八、研發與應用人工智慧應致力於建立

	人工智慧應用負責機制，以維護社會公益。爰參考新加坡二〇二四年生成式 AI 治理架構（Model AI Governance Framework for Generative AI）有關對於人工智慧開發運用之生命週期中，應確保不同角色（如開發者、部署者、最終使用者等）能承擔相應之責任等精神，於第七款定明可問責性原則（Accountability）。
第四條　政府應積極推動人工智慧研發、應用及基礎建設，妥善規劃資源整體配置，並辦理人工智慧相關產業之補助、委託、出資、獎勵、輔導，或提供租稅、金融等財政優惠措施。	人工智慧研發與應用涉及領域甚廣，其整體資源規劃，應由政府各機關依其業務職掌負責辦理，爰參考產業創新條例第九條及科學技術基本法第六條，定明政府應推動人工智慧研發、應用及基礎設施，及可運用之推動方式。
第五條　為促進人工智慧技術創新及永續發展，各目的事業主管機關得針對人工智慧創新產品或服務，建立或完備人工智慧研發及應用服務之創新實驗環境。	參考歐盟人工智慧法，鼓勵其會員國政府建立人工智慧實驗沙盒制度（Regulatory Sandbox），提供受控環境，以促進人工智慧之創新，使其於投放市場或投入使用之前，可於有限時間開發、測試及驗證。爰定明各目的事業主管機關得建立或完備有關人工智慧研發及應用服務之創新實驗環境，進一步使人民受益於人工智慧創新科技。
第六條　政府得以公私協力方式，與民間合作，推動人工智慧創新運用。 政府應致力推動人工智慧相關之國際合作，促進人才、技術與設施之國際交流及利用，並參與國際共同開發及研究。	一、考量人工智慧應用與發展事務涵蓋範圍廣泛，故於第一項定明政府得與民間合作推動人工智慧創新運用。 二、參考科學技術基本法第二十一條，於第二項定明政府應積極推動人工智慧國際合作、接軌國際，並參與國際共同開發及研究。
第七條　為加強國民對人工智慧之關心與認識，政府應持續推動各級學校、產業、社會及公務機關（構）之人工智慧教育，以提升國民人工智慧之素養。	為落實二〇二三年行政院科技顧問會議結論全面推動人工智慧素養教育，爰參考科學技術基本法第二十二條，定明政府應推動各級學校、產業、社會及公務機關（構）之人工智慧教育，以提升國民人

	工智慧之素養。
第八條　政府應避免人工智慧之應用，造成人民生命、身體、自由或財產、社會秩序、國家安全、生態環境之損害，或出現利益衝突、偏差、歧視、廣告不實、資訊誤導或造假等而構成違法之情事。 數位發展部及其他相關機關得提供或建議評估驗證之工具或方法，以利各目的事業主管機關辦理前項事項。	一、為保障人民權益，於第一項定明政府應避免人工智慧之應用造成人民生命、身體、自由或財產、社會秩序、國家安全、生態環境損害，或出現利益衝突、偏差、歧視、廣告不實、資訊誤導或造假等問題，違反如兒童及少年福利與權益保障法、公平交易法、消費者保護法及個人資料保護法等相關法令之情事。 二、所稱造成人民生命、身體（含身心健康）等之損害或出現偏差、歧視等，悉依相關法令認定是否造成法令所保障之權利受損或有無其他違法之情形。 三、為利各目的事業主管機關辦理第一項業務，數位發展部及其他相關機關得提供或建議國內外評估驗證之工具或方法，爰為第二項規定。
第九條　數位發展部應參考國際標準或規範，推動與國際介接之人工智慧風險分類框架。 各目的事業主管機關應視人工智慧應用風險管理之需要，循前項風險分類框架，訂定以風險為基礎之層級管理規範。有前條第一項所列之情事者，應依法令限制或禁止之。	一、為促進人工智慧穩健及安全發展，爰參考國際標準或規範，例如美國NIST AI 風險管理框架，採取以風險管理為基礎，由數位發展部推動人工智慧風險分類框架，提供跨部門之指導原則，供其他目的事業主管機關據以識別潛在風險類別進而有效應對，爰訂定第一項規定。 二、為使人工智慧應用朝良善方向發展，各目的事業主管機關就所涉領域內之人工智慧應用，應依第十六條規定，依本法規定檢討制（訂）定、修正相關法令，並循第一項風險分類框架識別、評估風險後，視風險管理之需要，訂定層級管理規範；如為無風險或低風險之人工智慧應用，則無須訂定管理規範。又人工智慧應用有第八條第一項所列情事，各目的事業主管機關應依其主管法令，包括既有作用法或後續配合人

	工智慧應用訂定之法令，予以限制或禁止，爰為第二項規定。 三、另所稱層級管理規範，係指因應人工智慧應用所涉風險高低，採行不同程度之管理措施，例如強化資訊揭露，要求提供較為詳盡之解釋文件及可驗證性資訊；透明度標示（與使用者互動時提供清晰資訊）、對模型訓練資料及決策結果進行公平度測試，或進一步採行審核、許可等措施。 四、如人工智慧應用對人民生命、身體、財產等造成損害，或對社會秩序產生重大危害，依現行技術手段，仍無法有效管理或降低該應用風險者，目的事業主管機關應依其法令予以限制或禁止其應用。
第十條　政府得透過法令或指引建立標準、驗證、溯源或問責機制，於促進人工智慧研發與應用之同時，以風險管理為基礎，評估潛在弱點及可能濫用之情形，強化人工智慧決策之可驗證性及人為可控性，以提升人工智慧應用之可信任度。	一、為提升人工智慧應用之可信任度，政府得透過具有拘束力之法令（如法律、法規命令或行政規則）或不具拘束力之行政指導（如技術應用指引）之方式，推動安全標準（包括團體標準、國家標準、國際標準）、驗證、透明可解釋之溯源或問責機制，協助研發、部署或應用人工智慧者納入風險管理制度，針對人工智慧決策過程或內容結果進行驗證，並確保能由人類有效監督。 二、又「可驗證性」（verifiability），參考韓國「人工智慧發展及建立信賴基盤之基本法」第三十條規定，係指能夠對人工智慧決策過程或內容結果進行驗證，從而保障公平性與責任歸屬；「人為可控性」（human oversight），參考歐盟人工智慧法第十四條及聯合國教科文組織「人工智慧倫理建議書」，則指能夠由人類有效監督，人工智慧不應取代人類

	道德判斷或責任，必須保留人類最終控制權；可驗證性提供透明可解釋之基礎，人為可控性則確保人類主導地位，兩者均係為提升人工智慧應用之可信任度。綜上，爰為本條規定。
第十一條　政府應就高風險人工智慧之應用，明確其責任歸屬及歸責條件，並建立其救濟、補償或保險機制。 人工智慧之研發，於實際應用前，不適用前項規定。但其於實際環境測試，或運用研發成果提供產品、服務時，不在此限。	一、高風險人工智慧之應用，係依據潛在風險及影響程度判斷之，如於特定關鍵領域應用時可能造成人民基本權利、生命安全、財產保障或社會秩序之嚴重危害。為確保人工智慧之安全性，政府應針對高風險人工智慧應用所可能產生之損害風險，明確其責任歸屬及歸責條件，並建立救濟、補償或保險機制，爰訂定第一項規定。 二、為避免影響學術研究自由及產業前端研發，爰參考歐盟人工智慧法第二條第八項規定，人工智慧投入市場前之任何研究、測試或開發活動僅須根據適用之歐盟法律進行，除於真實世界測試外，不適用人工智慧法。爰訂定第二項，定明人工智慧技術開發與研究，於尚未實際應用階段，不適用第一項有關責任相關規範，及建立救濟、補償或保險機制之規定。惟仍應遵循本法其他規定，與既有研發相關之法令及學術倫理規範。倘已於實際環境測試，或運用研發成果提供產品、服務者，仍應遵守第一項規定及其他各目的事業主管機關既有或後續訂定之法令。
第十二條　政府為因應人工智慧發展，應避免技能落差，並確保勞動者之職業安全衛生、勞資關係、職場友善環境及相關勞動權益。 政府應就人工智慧利用所致之失業者，依其工作能力予以輔導就業。	一、因應人工智慧發展，為避免勞動者於需使用及應用人工智慧技術從事及執行該職務工作時，欠缺人工智慧相關技能，並確保勞動者之權益，包含職業安全衛生、勞資關係及職場友善環境等，爰為第一項規定。

	二、為因應人工智慧利用造成之失業情事，爰於第二項定明政府應就該等失業者，依其工作能力予以輔導就業。
第十三條　個人資料保護主管機關應協助各目的事業主管機關，於人工智慧研發及應用過程，避免不必要之個人資料蒐集、處理或利用，並應促進個人資料保護納入預設及設計之相關措施或機制，以維護當事人權益。	為避免個人資料外洩風險及蒐集過多不必要之敏感資訊，爰參考美國二〇二二年 AI 權利法案藍圖（Blueprint for an AI Bill of Rights），為本條規定。個人資料保護主管機關應協助各目的事業主管機關，配合其業管法令建立個人資料保護納入預設及設計之相關措施或機制（data protection by design and by default），例如數位發展部訂定之隱私強化技術應用指引等，以維護人民權益。
第十四條　政府應建立資料開放、共享及再利用機制，提升人工智慧使用資料之可利用性，並定期檢視與調整相關法令及規範。 政府應致力提升我國人工智慧使用資料之品質與數量，以利人工智慧訓練及產出結果維繫國家多元文化價值及維護智慧財產權。	一、資料為人工智慧發展之重要元素，政府有必要促進人工智慧之創新與產業發展得以取得高品質之資料，爰參考歐盟人工智慧法有關支援高品質資料近用之規定，於第一項定明政府應建立資料開放、共享及再利用機制，就相關規範定期檢視並為必要調整，俾利人工智慧發展所需。 二、為利人工智慧訓練及產出結果維繫國家多元文化價值，避免影響弱勢、多元族群權益及人民之智慧財產權，爰於第二項定明政府應致力推動之事項，以完善我國資料治理機制。
第十五條　政府使用人工智慧執行業務或提供服務，應進行風險評估，規劃風險因應措施。 政府應依使用人工智慧之業務性質，訂定使用規範或內控管理機制。	一、考量政府各機關使用人工智慧協助執行業務或提供服務，有助於行政效率之提升，且應參酌第九條風險分類及管理規範進行風險評估與規劃因應措施，爰參考英國二〇二四年「生成式人工智慧治理框架」（Generative AI Framework），於第一項定明政府使用人工智慧執行公務應進行風險評估及規劃風險因應措

	施。 二、為促使政府各機關依一致之認知及原則使用人工智慧，爰於第二項定明政府應依使用人工智慧之業務性質，訂定使用規範或內控管理機制。各機關之人工智慧應用情形，由行政院統籌協調、盤點之。
第十六條　政府應於本法施行後依本法規定檢討及調整所主管之職掌、業務並制(訂)定、修正、廢止法規或指引，以落實本法之目的。 前項法規制(訂)定、修正或廢止前，既有法規未有規定者，於符合第三條基本原則之前提下，以促進新技術及服務之提供為原則，由中央目的事業主管機關會商中央數位發展主管機關，依本法規定解釋、適用之。	一、為落實本法，確保人工智慧技術之有效推動發展，爰參酌教育基本法第十六條、通訊傳播基本法第十六條第一項、原住民族基本法第三十四條第一項、海洋基本法第十六條第一項，為第一項規定，以利行政院統籌各部會檢討現行法規與相關機制措施。 二、依第一項規定應訂修或廢止之相關法規，於未完成法定程序前，為使相關事務能符合本法規定，完善相關法規之解釋及適用，爰參考海洋基本法第十六條第二項、韓國國家資訊化架構法第十七條、澳洲二〇二三年安全且負責任之AI政策討論書（Safe and Responsible AI in Australia Discussion Paper），為第二項規定。
第十七條　本法施行日期，由行政院定之。	為利本法施行相關事宜之準備，爰規定本法施行日期由行政院定之。