

「人工智慧AI」無人艦艇 軍事倫理問題初探

陳建源 上校

提 要：

- 一、隨著高科技與人工智慧(AI)的發展，AI在艦艇上的應用越來越普遍，引起有識之士對AI無人艦艇逐漸擴散的重視，並開始研究制訂AI倫理守則之可能性。
- 二、如果放任AI無人艦艇發展下去，後果可能不堪設想，但是以目前的科技，尚無法植入AI倫理守則，因此，國際間開始有了停止研發AI武器的呼籲。
- 三、指揮官應在軍事倫理的基礎上，確保軍事武力被負責任地使用，如果完全自動化的AI無人艦艇出現，將會面臨自主開火權、戰爭遊戲化與軍事必要性等軍事倫理問題。
- 四、戰爭是暴力的最大化使用，AI武器的軍備競賽最終恐會導致人類自我毀滅，我國是否應配合國際間的呼籲，停止有關AI武器研發呢？這是一個值得深思的問題。

關鍵詞：AI、AI武器、軍事倫理、未來戰爭

壹、前言

目前許多國家正積極地開發各式各樣的人工智慧(AI)武器，包括：無人艦艇、無人機、無人坦克、和自走武器等等，但完全不需要人類指揮的機器人，技術還沒有成熟到可以上戰場。雖然如此，但在運作上類似無人艦艇的AI武器，已經廣泛被運用在戰場上，並且引發激烈的爭論。本文以引發許多爭議的AI無人艦艇為例，討論發展AI武器所帶來的問題，並藉此讓我們進一步思考：萬一

未來具備AI的武器真的上戰場，可能會引起哪些軍事倫理問題？

所有的文明社會都以某種形式管制使用武裝力量、殺戮以及其他形式的有意圖的暴力。若沒有這樣的管制，社會本身就無法以任何穩定的狀態而存在。除了自衛外，法律與倫理禁止使用暴力，但軍隊例外；軍隊運用武力被認為是正當的，即使到了因為保家衛國的需要而殺人的程度。而這樣的正當性使得軍事倫理引起特別的關注，這種正當性必須與軍隊所服務的目的，及其所保護的國

家有關¹，也就是說，在發展AI武器的同時，也必須相對應的發展植入軍事倫理守則的方法，才不會導致濫殺無辜或增加平民的傷亡。

我國目前面臨少子化的情況下，以AI無人艦艇取代傳統的人力，除了可以發揮我國科技的優勢，解決兵力不足的問題，其實也是我國在建軍備戰上一個可以發揮「不對稱」戰力的選項。但是由於AI武器的危險性太高，所以國軍應該審慎看待發展AI武器時所可能遇到的問題，尤其是戰略性武器的發展，更不可能交由AI全權處理。但是在共軍積極從事AI無人艦艇研發的情況下，為了國家安全，我國是否應該積極投入AI武器的研發？或者應該順應國際間的要求，放棄研發AI武器？這是一個值得深入探討的問題，也是撰寫本文的主要目的。

貳、人工智慧(AI)發展

人工智慧(Artificial Intelligence，簡稱AI)是計算機科學領域的部分範疇，意指讓機器具備和人類一樣的思考邏輯與行為模式，發展過程包括學習(大量讀取資訊、並判斷何時使用該資訊)、感知、推理(利用已知資訊做出結論)、自我校正，以及如何操縱或移動物品。AI發展的領域包括：語音識別(Speech Recognition)、電腦視覺

(Computer Vision)與專業系統(Expert Systems)²，而AI武器即是將人工智慧應用於武器系統上，具有指揮高效化、打擊精確化、操作自動化、行為智慧化等特點的智慧型武器。21世紀最先進的軍隊，主要靠的是尖端科技，由AI武器組成的高科技部隊，正在取代由人組成的軍隊，而指揮作戰的將軍也會有越來越多關鍵決定，必須仰賴AI來處理³，有血有肉的戰士除了容易受到恐懼、饑餓和疲勞的影響外，思考及行動的速度也越來越無法趕上現代戰爭的步調⁴，在戰爭節奏越來越快速的未來戰場上，將AI應用於武器裝備將會越來越普遍，以下將就AI無人艦艇的發展及其危險性逐一探討。

一、AI無人艦艇的發展

舉凡無人艦艇、無人機、無人坦克等，AI在軍事領域的應用越來越廣泛，無人船艦可從遠端控制，執行巡邏、礦脈探測、偵測海中潛藏水雷等任務。未來若將這項技術用於軍事用途，將可避免兵力傷亡、折損，長時間的海上任務也能免除人員、物資後勤補給成本，船隻載重也能因此減輕⁵。英國的勞斯萊斯公司2017年公布了搭配AI技術的無人船艦原型概念，預計船長60公尺、能航行3,500浬，配有電力推進系統，能不中斷航行100天(如圖一)；而美國國防高等研究計畫署(DARPA)近期將名為「海洋獵人」(Sea

註1：Anthony E. Hartle著，莫大華、段復初、郭雪真譯，《軍事決策中的道德議題》(Moral Issues in Military Decision Making)(臺北：政戰學校，2005年)，頁188。

註2：李彥宏等著，《智能革命：迎接AI時代的社會、經濟與文化變革》(臺北：遠見天下文化，2017年)，頁76。

註3：Yuval Noah Harari著，林俊宏譯，《人類大命運：從智人到神人》(Homo Deus: A Brief History of Tomorrow)(臺北：遠見天下文化，2017年)，頁348。

註4：同註3，頁350。

註5：數位時代，〈勞斯萊斯跨足無人船艦，免除人力、後勤補給，軍事任務將更安全〉，科技新報，2017年9月15日，<https://technews.tw/2017/09/15/rolls-royce-autonomous-ship-navy/>，檢索日期：2018年3月12日。



圖一：英國勞斯萊斯無人船艦

資料來源：<http://technews.tw/2017/09/15/rolls-royce-autonomous-ship-navy/>，檢索日期：2018年5月17日。



圖二：美「海洋獵人」反潛無人船艦

資料來源：<https://defpost.com/leidos-awarded-contract-sea-hunter-ii-actuv-autonomous-vessel/>，檢索日期：2018年2月6日。

Hunter)反潛無人船艦(ACTUV，如圖二)正式移交給海軍研究實驗室，可望在今(2018)年內服役⁶。

美國和俄羅斯一直在軍備武器上不斷的進行競爭，美國海軍近期和國內最大的軍事武器製造公司簽訂無人潛航器「殺人鯨」的研發工作(如圖三)，美軍打算花費4,320萬美元打造出最新的海底武器「殺人鯨」，一口氣要建造五艘這樣的超級武器，這種無人潛航器可以攜帶大量的作戰武器，同時可以



圖三：美無人潛航器「殺人鯨」

資料來源：<https://lockheedmartin.com/en-us/products/orca-extra-large-unmanned-underwater-vehicle-xluuv.html>，檢索日期：2018年5月17日。

潛伏在海底很長時間，對於所有國家的航艦或者軍艦來說都是很大的威脅。另一方面，在美國剛剛決定研發無人潛航器「殺人鯨」時，俄羅斯最新武器就亮相，俄羅斯的「地位-6」無人潛艇可以攜帶核武器進行軍事打擊，該無人潛艇可以到達海底1,000公尺處，一次航行可達到10,000公里以上，對航艦來說是致命威脅。俄羅斯研發的此一最新無人潛艇自然是為了對抗美國的航空母艦，要知道這些無人潛艇可以最大安全距離接近敵方航艦，給予致命一擊，未來這些無人潛艇也將會成為主要的海上作戰武器⁷。

美俄兩國是世界上數一數二的軍事大國，雙方在各個方面都在相互較量。像俄國總統普丁今(2018)年2月在國情咨文中低調提到的神秘武器「波塞冬」(Poseidon)核魚雷將會在10年內交付俄國海軍使用(如圖四)，因為這是普丁在2018-2027年國家武裝計畫中提出的一部分，因此一定會在計畫結束前

註6：陳建鈞，〈可望年內服役 DARPA 移交海軍新式反潛無人艦「海洋獵人」〉，智慧機器人網，2018年2月5日，<http://www.limitlessiq.com/news/post/view/id/3477/>，檢索日期：2018年2月7日。

註7：軍事，〈美俄之爭，美國無人潛航器「殺人鯨」剛研發，俄羅斯最新武器就亮相〉，每日頭條，2017年10月27日，<https://kknews.cc/zh-hk/military/62k8gqm.html>，檢索日期：2018年3月12日。



圖四：俄羅斯「波塞冬」核魚雷

資料來源：<http://www.chinatimes.com/realtime-news/20180515004554-260417>，檢索日期：2018年5月17日。



圖六：中共「海翼1000型」水下無人載具

資料來源：<http://mil.firefox.sina.com/18/0220/09/BJ7YU5FT-WH1DG6VK.html>，檢索日期：2018年5月17日。



圖五：中共「D3000型」匿蹤無人海洋戰鬥艇

資料來源：<http://www.chinatimes.com/realtime-news/20170926003250-260417>，檢索日期：2018年2月6日。

，正式交付給海軍⁸。

中共海軍與海警也引進多款智慧或自主型無人水面載具，執行爭議水域平戰時巡弋任務。2017年9月，「中國航天科技集團」就公布一款D3000型匿蹤無人海洋戰鬥艇，可執行反潛與水面作戰任務(如圖五)，該公司另外的B850及C1500則可執行高速海上巡邏／攔截任務及反潛作戰；中共另挹注發展

船艦AI控制系統，將開始整合無人水面載具和海軍艦艇同步作戰，海軍已將自主水下無人載具列為重點研究項目，並為核子潛艇的電腦系統進行AI化，賦予潛艇決策、偵測周遭環境及戰術應用等能力，共軍海上作戰的實力將因此提升至不同的層次⁹；而「中國科學院瀋陽自動化研究所」開發的「海翼」載具，已先後於南海執行多次任務。最新型的「海翼1000型」(如圖六)近期才完成連續3個月的深海探測與外來潛艦偵測驗證¹⁰，中共海軍希望運用是類系統強化反潛戰力，抵消美國傳統水下優勢，並運用水下無人載具建立「水下長城」監控網絡。

二、AI武器的發展

AI武器系統現在正處於發展尖端，並且具有非常真實的潛力，會對無辜的人民造成重大傷害，以及造成全球不穩定。韓國已經在使用致命的AI武器，三星公司的SGR-A1哨

註8：陳建鈞，〈戰鬥民族太強大！俄研發新款無人潛艇「替代者」能「擬態」欺敵〉，智慧機器人網，2017年3月27日，<http://www.limitlessiq.com/news/post/view/id/267/>，檢索日期：2018年3月20日。

註9：陳建鈞，〈戰場 AI 化！中國要打造人工智慧核子潛艇 可協助偵測環境、制定戰術〉，智慧機器人網，2018年2月6日，<http://www.limitlessiq.com/news/post/view/id/3495/>，檢索日期：2018年2月7日。

註10：黃文啟，〈AI、軍事革新 中共打造未來戰力〉，《青年日報》，2018年2月1日，<https://www.ydn.com.tw/News/275767>，檢索日期：2018年3月21日。

兵在技術上有能力自主射擊，它擁有監視、語音識別和跟蹤能力，且可用機槍或榴彈發射器射擊；英國航太系統公司(BAE Systems plc)開發的塔拉尼斯無人機(Taranis drone)能夠攜帶空對空和空對地彈藥，並具有充分的自主權。俄羅斯、美國和其他國家正在開發可以遠程控制或自主操作的機器人坦克，像俄國Uran-9無人戰鬥地面車輛的自主版本，以及改裝成自主系統的常規坦克¹¹。目前世界上已有70多個國家軍隊在發展無人化系統平台，美軍已裝備的無人機達7,000多架，在伊拉克、阿富汗戰場上投入運用的地面輪式(或履帶式)機器人超過12,000個¹²。美軍未來將實現地面機器人占地面兵力的三分之一，艦載型X-47B無人機將占艦載機總量的三分之一，進一步推進有人平台與無人平台之間的協同編組演訓¹³。

2016年7月，美國海軍陸戰隊測試模組化先進武裝機器人系統(MAARS)，利用感測器和攝影機安裝於AI控制持槍機器人；美國陸軍研製的陸軍全球軍事指揮控制系統，目前已經裝備陸軍航空部隊運輸直升機，可使直升機駕駛員與前線士兵保持聯絡，並指揮地面部隊；俄羅斯戰略導彈部隊正在研製的狼式-2移動式機器人系統使用履帶式底盤，可在5公里範圍內通過無線電頻道控制，運用熱顯像儀、彈道電腦、雷射測距儀和陀

螺穩定器，以保證射擊精準度，能夠在時速35公里的情況下擊中目標。以色列研製的多戈(Dogo)自動武裝戰術作戰機器人，配備一個標準格洛克26型9毫米口徑手槍，堪稱小而強的AI武器。

鑒於AI武器的危險性，2017年8月21日特斯拉汽車公司(Tesla)首席執行官馬斯克(Elon Musk)和其他領先AI專家共116人，共同呼籲聯合國在全球禁止使用AI武器，包括無人機、坦克和機關槍等。馬斯克和谷歌(Google)旗下的AI公司迪普麥(Deep Mind)共同創辦人蘇雷曼(Mustafa Suleyman)等科技名人簽署一封公開信，信中寫道：「一旦這個潘朵拉的盒子被打開，將很難關閉。」要求聯合國防止目前正在進行的AI武器軍備競賽，這種軍備競賽有可能在火藥和核子武器之後引發「第三次戰爭革命」¹⁴。致命的自主武器一旦發展，將使武裝衝突規模比以往任何時候都要大，而且發展速度之快，超出人類能理解的程度。這些AI武器可能成為恐怖分子用做對一般尋常百姓使用的武器，畢竟「自主武器(Autonomous Weapons)在道德上是錯誤的」。專家希望將AI武器增列到聯合國禁用武器清單中，其中包括化學武器和故意致盲的雷射武器。然在AI的本質中，是否就隱含著它們最終失控的可能性？這正是目前人們對AI關注的熱門話題，用馬斯克

註11：夏雨，〈馬斯克聯手115專家 籲全球禁止殺人機器人〉，大紀元，2017年8月22日，<http://www.epochtimes.com/b5/17/8/21/n9551675.htm>，檢索日期：2018年3月12日。

註12：天下，〈人工智慧武器，人類是在打造未來還是在召喚惡魔？〉，ifuun，2017年2月12日，<http://www.ifuun.com/a20172121137536/>，檢索日期：2018年1月31日。

註13：〈The U.S. military will have more robots than humans by 2025〉，Mint Press News by Whitney Webb，2018年2月19日，<https://mronline.org/2018/02/20/the-u-s-military-will-have-more-robots-than-humans-by-2025/>，檢索日期：2018年5月17日。

註14：BBC，〈Killer robots: Experts warn of 'third revolution in warfare'〉，BBCNEWS，2017年8月21日，<http://www.bbc.com/news/technology-40995835>，檢索日期：2018年1月31日。

附表：名人、學者對人工智慧未來發展的看法彙整表

名人、學者	看法	理由
發明家、未來學家、同時是谷歌工程部門主管庫爾茨魏爾(Ray Kurzweil)	非常樂觀	相信人工智慧在2029年前，可達到人類智慧的高度，由於科技有協助治病和清理環境的潛力，「在控制箇中危險的情況下，我們有實現這種希望的道德必要性」。
暢銷作家、理論物理學家及未來學家加來道雄(Michio Kaku)	普通樂觀	以較長遠及務實的觀點來看，人工智慧是世紀末問題，如果屆時人類找不到較好辦法約束胡作非為的人工智慧機器人，那麼乾脆「在它們大腦置入晶片，關閉他們」。
企業家、慈善家及微軟公司創辦人比爾·蓋茲(Bill Gates)	有擔憂	低度人工智慧可成為替代勞工的工具，但唯恐數十年後出現的「超級智慧」系統，會「厲害到令人擔憂」。
理論物理學家、作家、黑洞物理學先驅霍金(Stephen Hawking)	有擔憂	人工智慧兼具奇蹟和災難兩種可能性，這是「人類史上頭等大事」，但也可能是「最後」的大事，「除非我們學會如何避免風險」。
企業家、太空探索科技公司(SpaceX)創辦人及特斯拉執行長馬斯克	擔憂	人工智慧是「我們最大的生存威脅」，無異「召喚魔鬼」。

資料來源：<http://www.chinatimes.com/realtimenews/20160309006996-260412>，檢索日期：2018年5月2日。

的話來說，AI武器正變成比核彈更危險，且征服世界指日可待¹⁵。

尖端科技的應用最早是從軍事領域開始，戰爭是涵蓋多面向的綜合比較，而AI武器的加入將加快各國軍事變革步伐。觀察兩次工業革命和兩次世界大戰的過程，我們會發現科技和戰爭存在著必然的關係，科技會觸發戰爭，戰爭又反過來推動科技的發展。現在各國在武器裝備資訊化、智慧化發展方面都取得長足進步，各種精密定位、精確打擊、精準評估武器系統層出不窮，但是尚未能將「人」從武器系統中分離出來，大部分的操作依然需要人力完成。AI與武器裝備相結合就意味著以後從搜索發現目標、到威脅評估、到鎖定摧毀、再到效果評估，這一系列過程完全不需要人類參與，機器就能幫我們作決定，做到作戰無人化發展¹⁶。但「兵者

，國之大事也」，就我國目前的科技發展狀況而言，情報的蒐集、處理、判斷與運用還是必須交由人來做決定，尤其在我國國情與目前兩岸關係緊張的狀況下，尚須由最高指揮官下達政治判斷，不適合貿然交由AI來處理複雜的兩岸關係，因為AI武器導致戰爭爆發的危險性實在太高了。

三、AI武器的危險性

(一)AI將為人們的生活帶來許多改變，但並非所有人都對此抱持樂觀看法，知名科學家、企業家和未來學家對人工智慧的看法分歧，有人認為是好的發展，也有人認為將帶來災難，時代雜誌(TIME)就曾列出多位名人對人工智慧未來發展的看法¹⁷(彙整如附表)。除了馬斯克、霍金這些科技大老外，俄羅斯總統普丁(Putin)最近也少見的提出他對AI的憂慮，特別是對AI近期的發展。普丁

註15：同註2，頁21。

註16：朱啟超、王婧凌、李大光，〈人工智慧叩開智慧化戰爭大門〉，中國軍網，2017年1月23日，http://www.81.cn/jmy-wyl/2017-01/23/content_7463310.htm，檢索日期：2018年3月13日。

註17：中央社，〈造福人類或災難 人工智慧利弊互見〉，中時電子報，2016年3月9日，<http://www.chinatimes.com/realtimenews/20160309006996-260412>，檢索日期：2018年5月2日。

認為這項技術的崛起，讓難以預測的機會和威脅提高，一旦有人在AI領域獲得突破發展並成為業界領導者，同時也將成為這個世界的統治者，為了避免這種情況，普丁承諾俄羅斯將和其他國家分享關於AI的知識，「如果有人獲得壟斷地位，情況將會變得十分嚴峻。」然隨著AI技術持續發展，未來的戰爭形態也會有所改變，屆時AI控制的AI武器將成為國家之間的戰爭代理工具，「當一方的AI武器被另一方摧毀，除了投降之外別無選擇」¹⁸。

(二)現今的AI已經發展到一個階段，在未來幾年內，自主武器將能完全派上戰場，然而帶來的風險也很高；自主武器被形容是繼槍砲、核武發明後，戰爭形態的第三次大變革。2015年4月，聯合國在日內瓦舉辦一場會議，討論未來武器、軍備等問題，包括所謂的「殺人機器人」(Killer Robots)；儘管英國抗議團體反對聲浪不斷，甚至發起「停止製造殺人機器人」活動，英國政府仍不顧反彈聲浪，反對一項有關於發展自主武器的禁令¹⁹。雖然不少人認為AI武器可以降低戰場上的人員傷亡，但由於無人操作化，自主武器可輕易引發更多戰爭，造成更多傷亡損害²⁰；科學家們也擔憂若是各國政府競

相投入AI武器的研究，將開啟當年核武時代的軍備競賽。然而與核子武器不同的是，自主武器沒有那麼昂貴、原料取得也相對容易，因此自主武器未來將隨處可見，而AI武器能夠在無人控制的情況下瞄準並射擊目標，將比核子武器來的更難以監控與掌握。

AI武器也將是人類存在以來最大的威脅，若是發展到極致，AI武器可能終結人類文明。霍金相信未來的科技將會出現自我意識，並超越及取代人類，而人類受限於緩慢的生物進化過程，將無法與AI匹敵。我們現在使用的AI，已經證明其功能的強大，針對未來科技的走向，很明顯地AI武器將成為明日的AK-47自動步槍；但是對人類而言，現在關鍵的問題是，我們要開啟這場全球AI武器軍備競賽，還是選擇阻止它？我們做出的選擇將決定我們的未來²¹。

(三)再從軍事作戰方面來看AI的應用，尤其當這些機器人被賦予某種程度的自動功能時²²，AI武器的危險之處在於軟體工程師可以很輕鬆地給它重新程式設計，使其成為濫殺無辜的機器。AI武器所蘊含的危險，確實不是那種常規的、由人類操控的武器可以比擬的。因此要求禁止AI武器的呼籲，也理應獲得大家的聲援。我國雖然不是聯合國的

註18：Nana Ho，〈未來戰爭就看無人機！普丁：AI 技術領導者將統治世界〉，科技新報，2017年9月4日，<https://technews.tw/2017/09/04/putin-believe-ai-could-lead-to-global-monopolies/>，檢索日期：2018年3月14日。

註19：〈UK opposes international ban on developing 'killer robots'〉，The Guardian，2015年4月13日，<https://www.theguardian.com/politics/2015/apr/13/uk-opposes-international-ban-on-developing-killer-robots>，檢索日期：2018年1月31日。

註20：Future of Life，〈Autonomous Weapons: an Open Letter from AI & Robotics Researchers〉，Future of Life Institute，2015年7月28日，<https://futureoflife.org/open-letter-autonomous-weapons>，檢索日期：2018年1月31日。

註21：江玟，〈霍金等千名科學家連署 反對人工智慧投入軍備競賽〉，風傳媒，2015年7月30日，<http://www.storm.mg/article/59318>，檢索日期：2018年3月14日。

註22：Jason Palmer，〈Call for debate on killer robots〉，BBCNEWS，2009年8月3日，<http://news.bbc.co.uk/2/hi/technology/8182003.stm>，檢索日期：2018年3月5日。

會員國，但一向主動遵守國際法與國際間的慣例，以維持我國在國際間的正面形象，而在面對AI武器發展的問題上，我國也應該主動配合國際間停止發展AI武器研發的呼籲，因為AI武器不僅危險，而且容易擴散，一旦落入恐怖分子的手中，對人類的危害就更大了。

四、AI武器擴散風險

1925年簽訂的國際條約禁止在戰爭中使用化學武器和生化武器²³，而中空彈(Hollow Point Bullet)也早在1899年就被禁用，因為這些武器會導致極大的、不必要的痛苦；它們特別容易導致平民傷亡，比如公眾可能會吸入毒氣，醫生在嘗試取出中空彈時可能會受傷。這些武器很容易給無辜者造成痛苦和死亡，所以它們遭到了禁用，是否有某種類型的AI武器同樣也應該被禁用呢？答案是肯定的。如果一種AI武器可以很便宜、很容易地改造成無差別殺戮，那麼我們就應該制定禁止它的國際公約，因為這樣的機器與放射性金屬在本質上沒有什麼不同，它們都可以和平利用，造福人類。但是我們必須小心翼翼地對它們加以管控，因為它們很容易被改造成毀滅性武器；不同的是，對AI機器進行改造，將其用於破壞性目的，遠比把核電站的反應堆改成核武容易得多。我們禁止AI武器，不是因為它們全都具有不道德的屬性，而是因為人們可以利用在網站上找到的資

料，輕易就把AI武器改裝成猙獰的嗜血怪物；只需一段簡單的代碼，就可以把很多人工智慧產品轉變成不亞於化學武器、生化武器的殺人機器²⁴。

南韓知名學府科學技術院(KAIST)最近成為全球AI專家的杯葛對象，逾50名AI專家發表連署信，要求該校停止研發AI致命武器，否則將全面拒絕與該大學的合作計畫。連署信上寫著：「如果研發成功，自主武器(Autonomous Weapons)將讓戰爭開打速度加快、戰爭規模也會遠大於以往，更可能成為恐怖攻擊武器。」連署信要求學校立刻停止研發由AI控制的致命自主武器，否則所有參與連署、50多名來自全球30國的AI專家將拒絕造訪KAIST，這些專家也不再和來自該校的學術人員進行交流²⁵。雖然AI將為人們的生活帶來許多改變，但AI武器將是人類存在以來最大的威脅，畢竟武器的發展，應該是避戰而非養戰，AI武器可能終結人類文明的擔憂，應該要有相對的因應措施，這也就是AI倫理研究計畫的由來。

參、AI武器倫理守則

一、阿西莫夫的機器人三定律

從1942年開始，著名的美籍俄羅斯裔科幻作家艾薩克·阿西莫夫在《驚奇科幻小說》(Astounding Science Fiction)上提出了機器人的倫理問題。今天，阿西莫夫被世人

註23：UN，〈1925 Geneva Protocol〉，UNODA，1928年2月8日，<https://www.un.org/disarmament/wmd/bio/1925-geneva-protocol/>，檢索日期：2018年1月31日。

註24：Kathy，〈比機器人還危險！禁止人工智慧武器迫在眉睫〉，騰訊科技，2015年8月29日，<http://tech.qq.com/a/20150829/008808.htm>，檢索日期：2018年3月13日。

註25：桂家齊，〈南韓知名學府研發人工智慧武器 全球AI專家聯手杯葛〉，《蘋果日報》，2018年4月5日，<https://tw.appledaily.com/new/realtime/20180405/1328850/>，檢索日期：2018年5月2日。

視為機械倫理學(Roboethics)之父，這門哲學學科催生了許多相關的研討會和無數的書籍，在美國尤其如此。阿西莫夫在1942年出版的小說《轉圈圈》(Runaround)中描述了機器人三定律(就像電影「機械公敵」的場景一樣)，在未來，AI機器人被人類廣泛使用，機器人在人類生活中占據了十分重要的位置，因為機器人三定律的限制，人類對機器人充滿信任，甚至已經被視為一個家庭的重要成員，他提出：所有的機器人程式都必須考慮到這三條定律，目的是為了約束這些自主機器的行為，讓它們服從以保護人類為目的的強制性道德準則²⁶。第一定律：機器人不得傷害人類，或坐視人類受到傷害；第二定律：除非違背第一法則，否則機器人必須服從人類的命令；第三定律：在不違背第一及第二法則下，機器人必須保護自己。此後，阿西莫夫又添加了一條定律：第零定律：機器人不得傷害人類族群，或因不作為使人類族群受到傷害²⁷。

遵守這些定律時引發的困局是阿西莫夫所著的500本書籍的主題，他嚴肅地思考了這些定律的意義，並想像了這些定律被採納時發生的情形；然而時至今日，阿西莫夫關於機械倫理的設想並未如願。如今的機器人

並沒有安裝強制它們服從機器人定律的機制或指令，為什麼會這樣呢？現代社會中廣泛使用的機器人的工作非常專業化，它們不具備與人類用通俗語言交流的能力，我們無法要求它們遵守阿西莫夫的機器人定律，是因為它們(指機器人)完全不能理解這些定律的意義。如果想要應用機器人定律，智慧系統必須具備邏輯分析和推理能力，並且能夠預測行為所導致的後果，以判斷人類下達的指令是否和任何一條定律相違悖。

二、AI倫理研究計畫

AI學會正在籌建倫理委員會，谷歌在收購迪普麥(Deep Mind)²⁸的AI開發公司時，也在自己公司內部設立了倫理委員會，這一領域對於社會的必要性正變得越來越顯著。倫理委員會的相關人員，應該必須是各個領域的專家，因為AI本身就跨越了眾多不同領域，所以這個委員會需要具備全面科學知識的人，當然也需要社會學和經濟學等方面的專家²⁹。中共2017年也在其《新一代AI發展規劃》中提出，建立AI法律法規、倫理規範和政策體系，形成AI安全評估和管控能力³⁰。未來，通過對AI相關法律、倫理和社會問題的深入探討，為智慧社會劃出法律和倫理道德的邊界，讓AI服務人類社會³¹，這也是世

註26：Piers Bizony，〈Asimov's Three Laws of Robotics〉，IET，2015年6月15日，<https://eandt.theiet.org/content/articles/2015/06/asimovs-three-laws-of-robotics/>，檢索日期：2018年2月1日。

註27：adminadam，〈The Four Laws of Robotics〉，Thriventnotes，2010年1月2日，<http://www.thriventnotes.com/the-four-laws-of-robotics/>，檢索日期：2018年2月1日。

註28：Deep Mind係2012年成立於倫敦的一家人工智能領域的新興企業，主要業務包括機器學習、深度學習和模仿腦細胞功能的神經網絡等技術研發，2014年1月被谷歌收購。

註29：松尾豐、鹽野誠著，陸貝旋譯，《大智能時代—智能科技如何改變人類的經濟、社會與生活》(北京：機械工業，2015年)，頁182。

註30：中共國務院，〈國務院關於印發新一代人工智能發展規劃的通知〉，中共中央人民政府，2017年7月20日，http://www.gov.cn/zhengce/content/2017-07/20/content_5211996.htm，檢索日期：2018年2月1日。

註31：科貓，〈智慧時代到來，細數人工智能那點事兒〉，網易，2017年11月2日，<http://mp.163.com/v2/article/detail/D28DOKEJ0511L5EH.html>，檢索日期：2018年2月1日。

界範圍內的一項共識。麻省理工學院媒體實驗室和哈佛大學伯克曼·克萊因互聯網與社會研究中心合作推出了AI倫理研究計畫；微軟、谷歌等巨頭也因AI的發展風險而成立了AI倫理委員會，而越來越多的機器人專家呼籲在機器人和自動化系統上安裝道德黑盒子，以紀錄機器的決定與行為。人們已經意識到，AI的發展應該以人類社會的穩定與福祉為前提³²。

為AI機器人設定倫理和規則，看來也是必要的。2017年1月5至8日在美國加州舉行的「Beneficial AI 2017」會議上，1980名來自產業、學術和研究界的相關人士，共同制定了「23條原則」，旨在確保AI的健全發展，人人都能從AI受益。這23條原則分為研究問題、倫理價值觀和長期問題三個部分³³，其中在倫理價值觀方面，23條原則中提出多項願景與呼籲，包括AI系統的安全與故障透明度(如果AI系統造成傷害，應該可以確定原因)；由機器人做出的任何涉及司法決策的行動，都要透明、可解釋、供主管人權機構審核；設計者與使用者要擔負起道德影響責任；要確保高度自治的AI系統，其目標和行為與人類的價值觀一致，並且尊重人類的尊嚴、權利、自由和文化多樣性；要能保障個人隱私，而且不僅AI系統能夠分析、利用人類產生的資料，人類也應該有權獲取、管理和控制自身產生的資料；不得不合理限

制人類真實或感知到的自由；AI帶來的利益，應當是普惠人類的；由人類選擇如何與是否委託AI系統，以完成人類選擇的目標；AI的力量，應當尊重和改善社會健全發展所需的社會和公民進程，而不是顛覆這種進程；避免發展AI軍備競賽等³⁴。

三、目前科技尚無法植入軍事倫理守則

美國iRobot公司的創始人羅德尼·布魯克斯(Rodney Brooks)曾就科技植入軍事倫理守則問題直白地表達了自己的觀點。iRobot是一家設計、製造和銷售機器人的公司，該公司製造的PackBot機器人大量用於福島核電站和阿富汗，媒體採訪他時，問起iRobot生產的機器人為何沒有設置為自動服從阿西莫夫定律，他回答道：「理由很簡單，我們根本沒法把阿西莫夫的定律寫入它們的系統中」。雖然未來的AI系統可能會隨著科技的進步而發展出智力，並且AI系統中負責自主抉擇的組件也會獲得性能提升，但這種機器智力和人類仍然不可同日而語，也就是說，科幻作家當初讓機器人理解和服從阿西莫夫定律的希望，現在已經變成了一件異想天開的事。因為目前的AI尚無法透過比較生命的價值做出判斷。就如邁可·桑德爾(Michael J. Sandel)說過的「一人死他人活」³⁵的情況，相信未來的AI武器也可能會遇到相同情況；當遇到這種情況時是否會根據事先輸入的演算法做出選擇？就像以前的

註32：陳聽雨，〈人工智慧，以法律和倫理為界〉，新華網，2017年8月23日，http://news.xinhuanet.com/tech/2017-08/23/c_1121525753.htm，檢索日期：2018年3月21日。

註33：李彥宏等著，《智能革命：迎接AI時代的社會、經濟與文化變革》(臺北：遠見天下文化，2017年)，頁378。

註34：同註33，頁380。

註35：Michael J. Sandel著，樂為良譯，《正義：一場思辨之旅》(Justice: What's the Right Thing to Do?)(臺北：雅言文化，2011年)，頁28。

阿西莫夫機器人三定律一樣。比如說，遇到事故緊急狀況的時候，如果要選擇殺死一個人還是殺死兩個人？那麼從人數上來說，犧牲者當然是越少越好；如果是要救一個老人還是一個孩童，是不是應該先救孩子？如果是一個男人和一個女人，是不是應該救女人？真的到了這種時候，也許我們不得不讓AI武器對各種狀況進行分析，然後優先考慮在屬性上與設定值相近的人³⁶。但是按照目前的科技發展，尚無法於AI武器中植入這種軍事倫理守則，所以國際間的有識之士才會站出來呼籲，停止AI武器的研發。

而科學和哲學又迫使我們重新思考。我們亟須對目前已經出現的(殺人或不殺人)機器人進行公正的評估，並採取相應的手段；這些繁複的評估任務包含了許多難解的倫理和哲學謎題，而其中一些迫切需要答案。哲學家、倫理學家、邏輯學家、軍方人士和軟體工程師應當展開討論，並達成一致的意見，如果在AI的武器形態運用上，可以設定某種特定的族群或規範、或在設定中能導入深度學習技術，那或許就有可能解決和改善問題，也較能避免無法收拾的情況發生。

肆、AI軍事倫理探討

針對軍事倫理的定義，美國學者佛廷(Nicholas Fotion)與依利費士羅姆(Gerard Elfstrom)曾在1986年合著《軍事倫理學：

平時與戰時的指針》(Military Ethics: Guidelines for Peace and War)一書認為：軍事倫理學在於釐清由於動用武裝力量所衍生的道德問題，解釋這些問題彼此的關係，並試圖解析期間的頭緒³⁷。高階指揮官應在軍事倫理的基礎上，確保軍事武力被負責任地使用³⁸，伴隨著各國對AI武器的趨之若鶩，AI武器還將帶來更深刻的問題，且波及軍事倫理等多個方面。以下分別就AI武器的自主開火權、戰爭遊戲化與軍事必要性等相關問題分別說明。

一、自主開火權

如果完全自動化的AI武器出現，這將澈底挑戰人類的軍事倫理底線：機器是否有權力決定人類生死？因此，AI武器的應用從《國際人道法》角度來說是無法接受的，它不但會增加意外傷亡率，而且有可能將人類置於機器統治的恐懼中。有時人們提出是否開火或何時開火，應由人來決定這樣的觀點，假設空中有兩架無人機正在交戰。其中一架需要先收到遠方碉堡操作人員的許可，才能開火，另一架則能完全自主開火，此時何者勝算大？當然是能完全自主開火的那架勝算大³⁹；所以未來的AI武器一定會擁有自主開火權。因此，命令下達必須有嚴密的機制考核、評估和管制，而解決之道必然會是進一步提升AI武器的運算法與深度學習能力。

二、戰爭遊戲化

註36：同註29，頁125。

註37：Nicholas Fotion & Gerard Elfstrom, *Military Ethics: Guidelines for Peace and War*(Boston: Routledge & Kegan Paul, 1986), pp.1-5。

註38：莫大華、段復初、郭雪真合譯，《美軍軍事倫理譯文集：理論、教學與資源》(臺北：國防大學政戰學院，2007年)，頁57。

註39：同註2，頁350。

AI武器也引發了一些軍事倫理問題。它切斷了公眾與正在進行的戰爭的聯繫，而且這種脫節還在加深，軍事倫理學家承認，AI武器可以把戰爭變成電動遊戲，造成平民傷亡。在英文中，空拍機也俗稱無人機(Drones)，但其實無人機的正式名稱為「無人飛行載具」(Unmanned Aerial Vehicles)；關於名稱問題，有論者認為常用來稱呼各種機器的技術性詞彙，其實沒有表達出這些機器的真正意義，以及它們廣大的影響層面；如果拿「全自動化武器系統」與「殺戮機器人」這兩個名稱來比較，後者可能會引發我們的各種聯想，以及因此而來的厭惡感；而前者雖然貌似中性，只單純地描述機器的功能，卻淡化了這功能其實指的是殺戮和毀滅，這種感覺上的距離，正是問題的重點所在：機器人的存在，常常讓我們以非直接的方式去感覺，遠距遙控的無人機的運作，正揭露出間接的感覺方式所引發的問題。

透過操作員在地面上遠距遙控，無人機已經廣泛被用來作戰，它們也是目前美、英各國正在發展的AI武器的藍本。遠距互動技術，確實在一定程度上傳達遠方的狀態給使用者知曉，我們很容易從廣泛使用的電話或是遠距視訊，瞭解遠距互動技術帶來的好處。不過，遠距活動通常只能傳遞極為有限的視覺和聽覺資訊；在此情況下，我們賴以溝通的其他感官經驗能力，在遠距活動中幾乎都派不上用場。遠距技術的限制，在戰場上

產生了嚴重的後果，一般認為，無人機在所謂「反恐」戰爭的應用上，只會針對極少數特定的對象，而且造成的平民傷亡很少。但事實上，在阿富汗及敘利亞的戰場上，無人機已經造成數千人死亡，其中有許多都是平民，當無人機的操作者按下射擊鈕那一瞬間，感覺到的不是血肉橫飛的震撼，而只是螢幕上的人物的姿態改變而已。遠距無線技術，會讓戰爭變成一場無關參與者痛癢的電腦遊戲，武裝人員可以藉此技術遠離戰場的危險，但這卻是以更多無辜平民的傷亡為代價。

未來，上戰場的是不太會有人關心其「死活」的AI武器，那麼，殺戮會不會因為不影響戰爭參與者的感覺，而變得更肆無忌憚？這正是遠距遙控無人機的使用，所給予我們的警示⁴⁰。人類在遠端操控武裝無人機進行攻擊時，操縱者在距離攻擊目標很遠的地方，命令武裝無人機實施轟炸，這種情況下，遠程操控者的罪惡感會是什麼樣的感覺？也許反而會對身體與責任的割裂這件事本身產生罪惡感。例如以色列製造的Harop武裝無人機具有自動選擇打擊目標的功能，自動化水準已經相當高了，但是在國際法領域，武裝無人機的使用還是一個盲區。一般來說，我們認為使用武裝無人機即使沒有違反國際法，也是與海牙陸戰法規和慣例公約中的馬爾頓斯條款⁴¹相牴觸的。這是一個十分現實的問題，個人認為人類有必要針對無人介入的AI武器進行明確的討論，當濫殺無辜的

註40：林宗德，〈機器人倫理學專題報導(一)：無人機的警示〉，科技大觀園，2014年10月1日，<https://v1.scitechvista.nat.gov.tw/zh-tw/Feature/C/0/13/10/1/1740.htm>，檢索日期：2018年3月6日。

註41：海牙陸戰法規和慣例公約中的馬爾頓斯條款(The Martens Clause)：認為即使國家之間沒有締結條約，戰爭時也應當從人道及倫理角度充分考慮相關人員的利益。

事件發生時，由誰來負責？⁴²尤其是在濫殺無辜的事件發生前，誰來預防和管控是最重要的部分，畢竟武器是沒有生命和是非判斷能力的。

三、軍事必要性

瞭解任何遭受AI武器戰爭影響的傷害，都必須符合軍事必要性是一件很重要的事，然而這種論調經常遭到誤解。因為軍事必要性並不能將指揮官或士兵在達成目標，所必須採取的任何與所有的行動所造成的傷害正當化，任意或放縱AI武器所造成的傷害違反了軍事必要性。畢竟軍事必要性必須從戰時正義的限制中來加以理解；而戰時正義是提供關於什麼可以被傷害，與被傷害程度的標準，並不允許為了軍事目的而違反這些標準；換句話說，AI武器的使用必須是一個軍事的理由所造成的正當傷害，且傷害必須符合戰時正義的標準⁴³。孟子曰：「生，亦我所欲也；義，亦我所欲也。二者不可得兼，舍生而取義者也。」⁴⁴像孟子這種東方的倫理思想也是近年來西方所推崇和學習的，在現代知識如此普及的狀況下，中外倫理哲學形成雙方交流的管道，其中也包含了軍事倫理這一重要的課題。

戰爭的主導者是戰場指揮官，武器只是一種工具，而如何運用AI武器，完全掌握在指揮者的身上，所以真正的源頭是指揮者的倫理觀念，以及組織架構的嚴密編組，能夠發揮極大的相互監察督導之責，如若要落實

這個觀念，自應從最基礎的軍事教育做起了。

伍、結語

AI武器除了目前的科技發展尚無法植入軍事倫理守則外，另一個原因是人類製造AI武器的目的就是為了殺戮。人類製造用來殺戮同類的武裝無人機已經有數十年的時間，而且只要AI武器、自主性和殺戮這三個特性一同出現，情況就變得十分棘手。拿埋在地雷為例子，我們可以說它具有自主性，因為當它覺察到四周出現物體時會決定爆炸，而這個決定並不需要人類的許可。許多導彈和炸彈起爆的唯一條件是，它們敏感的引信裝置覺察到自己將要和目標或者地面發生接觸。例如中共已經發展出偵查打擊一體化的武裝無人機，再加上臉部識別系統，利用這種具有AI的自主軍事武器，把它們派遣到地面戰場上或空中去搜索敵人，無須獲得人類許可，它們就能自己辨認並消滅目標。就算我們假定在戰鬥決策環中有人類存在，如果不對這些機器自主性的深層概念進行調整的話，它們還是可以撇開人類獨自執行任務。

有些人認為，AI武器的精確性與無人化是否將使戰爭逐漸變得更加人道化與仁慈化了呢？絕對不是，AI武器絲毫不能掩飾戰爭的殘酷性，外科手術式的精準打擊也不能改變戰爭暴力的本質；相反地，AI武器將使戰爭的場面更加慘烈。克勞塞維茲早就指出，

註42：同註29，頁122。

註43：Bill Rhodes著，莫大華、段復初、郭雪真譯，《軍事倫理學簡論：一本參考手冊》(An Introduction to Military Ethics: A Reference Handbook)(臺北：國防大學政戰學院，2012年)，頁149。

註44：高思謙著，《中外倫理哲學比較研究》(臺北：中央文物，1983年)，頁58。

火藥的發明與武器的不斷改進已經充分地證明，文明程度的提高絲毫沒有改變戰爭概念所固有的消滅敵人的傾向，戰爭從來就是殘酷、暴力和具破壞性的，戰爭形式上的文明，改變不了戰爭的暴力本質，也不能說明戰爭具有仁慈性，因為戰爭本身就不是什麼仁慈的行為，如果說流血的屠殺是殘酷、可怕的，那麼這只能使我們更加嚴肅地看待AI武器⁴⁵。所以，國際間停止研發AI武器的呼籲，並成立AI植入倫理守則的研究，為我國的建軍備戰方向上，提供了一條可以發展的和平道路。

個人認為，在兩岸關係緊張的狀況下，中共積極投入AI武器的研發，我國照理應該急起直追，善用我國中科院AI科技發展的優勢，運用AI解決少子化所面臨的兵源不足問題，像英國一樣不顧國際反對的聲浪，自力研發AI武器，以加強國防戰力。雖然我國106年「國防報告書」亦有提出研發無人載具、智慧水雷與微型艦艇的「不對稱作戰」

發展戰略，但就AI武器的自主研發而言，需要投入大量的時間與資源；就敵我態勢而言，共軍目前已占優勢，我們是否要長期投入，並與共軍進行AI武器軍備的競賽；就國際關係而言，外交戰略的層次高於國防戰略，既然AI武器的危險性等同於核生化武器一樣恐怖，那我國是否應該配合國際間的呼籲，維持愛好和平的形象，停止AI武器的研發？抑或是必須為了維護國家安全，不顧國際上有識之士的呼籲，要求中科院全力投入AI武器與無人艦艇的研發，發揮「不對稱作戰」的戰力優勢呢？這個值得深思的問題，確實值得決策者發揮智慧與勇氣，為我國軍建軍備戰指出一條明確的發展方向。

作者簡介：

陳建源上校，政校82年班、政戰學院博士班102年班；曾任連、營輔導長、遠朋班聯絡官、學員生大隊長，現任國防大學政戰學院政治學系助理教授。

註45：劉淑萍、趙楓主編，《現代軍事倫理學概論》（北京：國防工業，2005年），頁170。

老軍艦的故事

永新軍艦 MSC-159



永新軍艦為一掃雷艦，在美海軍服役時之編號為MSC-302，係由美新澤西州多徹斯特造船公司承建，民國53年2月29日下水。民國54年3月5日由首任艦長蘇曾敬少校率領官兵前往美國賓夕凡尼亞州費城海軍第四軍區接收，由駐美武官池孟彬少將及美方第一軍區司令史廷遜上校共同主持命名及升旗典禮，同年10月9日返抵左營正式服役。

永新軍艦成軍服役後曾參加掃雷作戰，兩棲作戰，中美混合訓練、漢光、自強等演習，主要武器為二座雙管20公厘砲。該艦設計之主要任務為近岸掃雷，為求在掃除磁性水雷時，能確保本身之安全，故其艦體均為木質構成。但其他所有裝備及機件，均以非磁性之金屬製成。(取材自老軍艦的故事)