

如何成為「人」：缺陷及其經驗作為對人工智能之啟發—以自動駕駛技術為例

“Becoming Human” : the distinction of “de-fault/default” as an inspiration for artificial intelligence – cases from self-driving cars

(初稿請暫勿引用)

劉育成

南華大學應用社會學系助理教授

ycliu15@gmail.com

中文摘要：

以「失誤 (fault)」概念為基礎的「缺陷」，及對其之經驗，或許是「人之所以為人」的關鍵元素。無論是有機體或自然等概念，在某種意義上來說，都是有缺陷的。反過來說，一個想像上的完美身體或許並非是自然的。在某種意義上，這樣的身體是脫離於自然狀態的。「缺陷」構成了社會之所以能夠存在、發展與變遷的基礎之一。「成為一個人」意謂著一個如何面對缺陷、處理缺陷的過程。

「缺陷」是人類經驗的重要來源之一，而「缺陷的經驗」是「成為一個人」的重要基礎。本研究將延伸技術哲學家 Bernard Stiegler 的觀點，將人類因為艾比米修斯的「失誤」而得以「成為人」的論述，置於對人類智能 (human intelligence, HI)、人工智能 (artificial intelligence, AI)，以及當前發展中的機器學習 (machine learning) 與深度學習 (deep learning) 之探討與啟發上。在此一不斷地透過技術以及技術的外部化、客體化等過程，人類嘗試彌補其「失誤」，且藉由技術將失誤予以外部化而區別自身與其他物種（包括自然），也就是藉由技術來肯定「成為人」的存在意義與價值。這個在一開始就註定要不斷地「去除原初之失誤 (original de-fault)」或者是不斷地再次確認或再進入這個「原初的預設 (original default)」的辯證過程，或許有助於我們理解人工智能之發展所面臨的問題：怎樣的人工智能可以讓更像人、「成為人」，甚至是超越人類智能？從形式理論觀點來看，任何形式的產生都是一組區別在該區別的其中一邊不斷複

製該區別的產物。若將 Stiegler 所玩弄的雙關「去一失誤／預設 (de-fault/default)」當作是一組區別，那麼人類智能乃是在前者中複製該組區別自身，而目前的人工智能發展，以及稍後的機器學習與深度學習等，則是在後者之中複製自身。本研究也將以人工智能在自動駕駛技術中的應用為例，以及在後半部分中延伸而提出「例示／窮盡 (illustration/exhaustion)」概念，以作為對人工智能乃至於對人類智能的內涵提供不同的理解與貢獻。

關鍵字：人類智能、人工智能、機器學習、深度學習、去失誤／預設

Keywords: human intelligence, artificial intelligence, machine learning, deep learning, de-fault/default

一、「像人一樣思考 (think like humans)」？

在討論人工智能之前，我們或許應該從何謂人類智能 (human intelligence) 出發，並且探究何謂「像人一樣思考」。¹「像人一樣思考」這句話指涉了一個不同於人的對象物：有個東西不是人，但希望可以具有人的思考能力或思考方式。換句話說，其已經預設了一組區別：人與非人。此外，該論述更期待可以讓非人也像人。人工智能概念的提出與發展，則預設了這組區別，以此為前提而想像了一種可以像人一樣思考的機器或演算法。一般認為，1956 年夏季在達特茅斯舉辦為期六週的「關於人工智能的夏季研究計畫 (Summer Research Project on Artificial Intelligence)」會議，被視為人工智能研究領域的開端 (Bostrom, 2014)。Nils Nilsson 在回顧人工智能發展歷史上，則認為有三場對其影響重大的聚會，除了 1956 年的達特茅斯會議之外，尚有前一年 (1955) 在洛杉磯舉辦之 Western

¹ 本研究將 artificial intelligence 翻譯為「人工智能」而非大多數所見之「人工智慧」，主要原因在於 intelligence 與 wisdom 無論是在英文或中文（智能 vs. 智慧）均有不同的指涉內涵，但 wisdom 在中文裡經常譯為智慧而非智能。在中文的語用學中，智慧也與智能不同。在教育部編國語辭典修訂本中，智能指的是「智識與才能」，而智慧則是「分析、判斷、創造、思考的能力；聰明才智」等。因此，智能一般被定義為「具有最適切、有效及有效率地解決問題之知識與才能（技能）」。²相對地，智慧則不僅止於此，更包含創造、思考等能力。儘管人工智慧一詞在台灣語境中較常出現，但本研究認為區別兩者仍有其意義。這也涉及到我們如何理解人工智能的內涵與應用。「像人一樣思考」是指要跟人一樣具有智慧，還是只需要有解決問題的技能，這或許不僅值得思考，此一區別似乎也指出的是，「人之所以為人」以及「人有別於非人（包括自然界、動植物、技術物等）」兩者在人與非人之互動上具有的辯證關係。

Joint Computer Conference 中關於「學習機器 (Learning Machines)」的場次，以及 1958 年在英國的國家物理實驗室 (National Physical Laboratory) 舉辦的「思考過程的機器化 (Mechanization of Thought Processes)」論壇 (Nilsson, 2009, p. 73)。根據 Nils，1955 年關於機器學習場次中所發表的論文主要有兩個方向，其一是對神經系統的模仿，其二對心智層級的模仿。儘管兩者有可能是殊途同歸，亦即都是在了解大腦的運作，但實則是在不同層次上對大腦進行描述 (Nilsson, 2009, pp. 76-77)。對人工智能的定義，或許首要出現在 John McCarthy 提給洛克斐勒基金會尋求贊助會議的計畫中。他將當前的人工智能問題定義為「使機器以能夠被稱之為智能的方式運作，就如同人類也是如此運作一般」 (Nilsson, 2009, p. 77)。這個定義明確區別於過去關於自動化理論 (automata theory) 的研究，以及避免與模控論 (cybernetics) 連結在一起。換句話說，不只是自動化與控制問題，人工智能關注的是人類智能的形式、內容與運作，而非單純的自動化運作或者是控制與反饋的問題 (Nilsson, 2009, p. 78)。第三場聚會是關於思考過程的機器化問題，包括人工思考、特徵與模式辨認、學習、機器語言翻譯、生物學、自動編程、工業計劃與辦公室的機器化等，主要代表人物有 Minsky、McCarthy、Ashby、Selfridge 以及 McCulloch 等 (Nilsson, 2009, pp. 81-82)。

達特茅斯會議中大多數的參與者，對模仿較高層次的人類思想相當有興趣，也就是關於人類如何解決問題 (how human solve problems) 的過程。換句話說，關於人工智能的討論，在一定程度上是回到何謂人類智能的問題上。**關於人類智能的討論，我想先從 Howard Gardner 在《心智的框架》一書中對智能的可能內涵之討論出發**，他提出所謂「多重智能 (multiple intelligences)」概念，後為教育及認知心理學所沿用並發展之。Gardner 對智能 (intelligence) 的定義是：「智能是解決問題的能力、創造產物的能力，而且這些問題或產物在一個或更多文化場景中是有價值的 (An intelligence is the ability to solve problems, or to create products, that are valued within one or more cultural settings)」 (Gardner, 2011, p. xxviii)。不同於過去從先驗的 (a priori) 觀點來發展關於智能的論述，Gardner 則是從經驗出發，認為一般性的「智能就是一組解決問題的技巧—使個人能夠解決真正的問題或困難，並且創造有效的

產物（結果）—以及發現或創造問題的潛能」（Gardner, 2011, pp. 64-65）。因此，Gardner 將智能連結於文化脈絡，文化在各種不同智能的出現上扮演重要角色，且在不同的文化中有不同的智能需求發展與展現。在初期研究中，Gardner 區分了六種智能，後增加為八種：語言智能、音樂智能、邏輯—數學智能、空間智能、身體—動覺智能、人際智能、內省智能，以及自然主義智能(Ibid., p. 77)。如果從 Gardner 所定義的智能以及其所區別的多重智能來看，我們或可發現人工智能所謂的「像人一樣思考」，指涉的也是這些不同面向及其混合或綜合體。

如果人類智能的出現與發展無法脫離於文化，那麼人工智能的研究不僅需要考量的是一般性的智能概念，若要讓其「像人一樣思考」，似乎也得讓人工智能能夠具有在不同文化中學習的能力。根據 Gardner，文化在人類智能的發展上扮演之角色是，文化動員了這些基本的訊息處理能力—這些核心的智能—以及形塑它們以獲得各自之目標(Gardner, 2011, p. 341)。這與人工智能的概念與技術發展是不同的。或許先值得一提的是，這也呼應 Bostrom 對人類智能與人工智能之關係的討論中所提到的「沙盒 (sandbox)」現象，亦即相對於人類智能的終極目標或價值如果是自我存續，那麼人工智能的終極目標及其價值可能為何的問題(Bostrom, 2014, pp. 146-147)。沙盒現象的吊詭在於，人類智能希望打造出對其自身來說是安全的人工智能，因此，人類智能總是努力地確保其安全無虞。換句話說，人類智能總是期待在一個人工智能—或 Bostrom 的「超智慧 (superintelligence)」—受到控制的環境下—亦即「沙盒」—對其進行觀察。在此過程中憑藉的仍是人類智能的「經驗」，因此這個經驗也限制了人工智能得以被觀察或獲得發展的內容：

在沙盒中表現得體是友善人工智能與不友善人工智能共通的趨同工具目標。一個智能充足的不友善人工智能會察覺到，如果它一開始表現出友好的態度而被放出沙盒，它的終極目標才有可能實現。等到我們就算發現也沒差的時候—也就是說，當人工智能已經夠強大，而且人類的反抗已經無效時—它才會開始展現它不友善的本質 (Ibid., p. 147)。

相對地，人類智能的發展或其目的，並非在確保人類智能對其自身而言是安全無虞的，也因此發展過程中，我們透過各種失誤、錯誤等來學習。人類智能不需要在其發展之初，即需要經過一個篩選的過程，以讓友善的人類智能勝出並且獲得發展。相反地，人類智能的友善／不友善這組區別是在與其自身環境的互動過程中，關於該區別的內涵才獲得形塑。然而，人工智能的發展卻似乎正好相反。Bostrom 指出人工智能在發展過程中，友善／不友善這組區別已經在操作上（先驗上）被用來觀察人工智能自身，在確保其必須具有友善的結果之基礎上，人工智能的學習跟人類智能的學習過程便有著完全不同的內涵。

Andy Clark 在對人類認知強化 (cognitive enhancement) 的討論中提到，大腦如何透過一個特定形式的整體問題解決組織 (a specific form of overall problem-solving organization) 來運作並處理從外界而來的訊息 (Clark, 2013, p. 323)。在其稍早關於人類智能與賽伯格 (cyborgs) 的討論中，將賽伯格視為是我們自身生物學本質的偽裝想像 (a disguised vision of our own biological nature)，一如其書名所示：我們是「天生的賽伯格 (natural-born cyborgs)」 (Clark, 2003, pp. 5-6)。對他而言，所謂的人類大腦以及對人類智能具有的獨特特徵的最好解釋是，「其有能力與非生物學上的建造物、支柱物與輔助物建立深層且複雜的關係 (their ability to enter into deep and complex relationships with nonbiological constructs, props, and aids)」 (Ibid., p. 5)。這些「非生物學上的建造物、支柱物與輔助物」指涉對象包括技術物與人造物等，或許也包含了文化、社會、政治與經濟制度等。綜上所述，人類智能的發展及其特徵，或許與人類智能與其自身之環境的互動有關，文化也是在其環境之中。據此，反過來說，如果我們期待人工智能「像人一樣思考」，那麼是否也表示其需要具有與其自身之環境建立「深層且複雜的關係」之能力。如同 Clark 所指出的，人類智能是一種「embodied intelligence」：

Humans are never disembodied intelligences; work in telepresence, virtual reality, and telerobotics, far from bolstering any mistaken vision of detached, bodiless

intelligence, simply underlines the crucial importance of touch, motion, and intervention. In all the cases we have examined, what matters are the complex feedback loops that connect action-commands, bodily motions, environmental effects, and multisensory perceptual inputs. It is the two-way flow of influence between brain, body, and world that matters, and on the basis of which we construct (and constantly re-construct) our sense of self, potential, and presence. The biological skin-bag has no special significance here. It is the flow that counts (Clark, 2003, p. 114).

據此，有關人類智能以及何謂「像人一樣思考」的討論，便可從人類智能與其環境之互動出發，Clark 描述其為一種「流動 (flow)」²。然而，Clark 並未進一步探究「流動」的可能內涵為何。從技術哲學的角度觀之，就如同前述的沙盒現象，當代對人工智能的期待仍是以人類為中心的 (anthropocentric) 思考方向，只有在人類智能得以控制的情況下，才有所謂的人工智能。也因此，1960 年代以降的人工智能，乃至於近年來較屬於狹義人工智能的機器學習 (machine learning) 與深度學習 (deep learning) 等科學與科技發展，大多是以最佳化、最適切、有效及有效率等為檢驗人工智能是否成功的原則，而非對 Bostrom 所謂的「人工常態智能 (artificial general intelligence)」，或者是 Nilsson 稱之為「強人工智慧 (strong AI)」有更多的關注或進展 (Bostrom, 2014, p. 18)。換句話說，當前的人工智能由於各種既有的限制，大多偏向以支援人類思考的「弱人工智能 (weak AI)」方向發展，例如現下流行的聊天機器人 (Chatbot)、智慧型手機語音助理 (蘋果公司的 Siri、Google 的 Google Assistant，或者是 Amazon 的 Alexa 等)，以及高頻率運算交易程式 (algorithmic high-frequency traders programs) 等²。然而，Bostrom 卻也理解所謂的「學習能力」，對人工常態智能而言，「並非是事後的添加延伸，而應該是核心設計的必要

² 高頻率運算交易程式被廣泛地用在全球金融市場，其主要用途是為大型投資公司自動化股票交易系統，包括將投資管理人的買賣執行指令予以自動化，或者是用來適應多變的市場狀況，進行複雜的交易策略。高頻率運算交易程式得以在非常短地時間內回應，其基於理性邏輯的運算也終究導致了 2010 年的美國股票市場大崩盤 (Bostrom, 2014, pp. 16-17)。

特徵」。此種學習能力包括對不確定性與機率資訊的有效處理。這也讓後來的人工智能研究轉向腦神經研究與生物遺傳學研究。此外，Bostrom 也提到，人工智能不見得要與人類的心智相類似，相反地，人工智能可能會「有著與生物智能徹底相異的認知架構，且在發展初期就會有截然不同的認知強弱項」，且「人工智能的目標系統可能會和人類的大相逕庭」(Bostrom, 2014, p. 29)。

若是回到 Gardner 將人類智能發展與文化相連結，以及 Clark 將人類智能的發展視為一種由自身與其環境（包括文化）之間各種流動的討論，關於人工智能是否能夠「像人一樣思考」的可能性，或許便得將這兩個論述納入考量。每一種文化的展現及其內涵，相對於人類智能整體而言，總是局部的（part and partial）。透過文化，我們僅看得到人類智能的某個或某些面向，而非其全貌。**文化的角色、發展及其內容，乃是人類透過技術將其自身之缺乏予以外部化（externalization）的產物之一。**文化是技術的，同時也是外在於人的。人在文化之中且受其影響，但人卻不屬於文化本身。從 Stiegler 的技術人類學觀點而言，人類將其自身原初之缺乏（original absence）的外部化，乃是人之所以能夠「成為人（becoming human）」的重要元素。接下來的第二節便從技術哲學一主要延伸自 Bernard Stiegler 的技術人類學一觀點來探討缺陷及其經驗，以及在第三節中討論由此所衍生出來的「去一失誤／預設（de-fault/default）」這組區別如何重構人類智能與人工智能在起源與發展上的不同路徑，並且以自動駕駛技術為例說明。最後在代結論的部分，我將提出以「例示／窮盡（illustration/exhaustion）」觀點作為對人工智能發展的反思與建議。

二、缺陷及對其之經驗作為「成為人」的基礎

在生命教育中經常提到的一句話是，「生命的缺陷乃是上帝化了妝的祝福」（鄒川雄、周平，2018）。儘管這句話具有較強烈的基督教信仰基礎，但從生物學角度觀之，沒有一個生命體是完美的、沒有缺陷的，至少，所有的生物體都會面臨死亡的這個事實，可以說是最大也是最無可避免的「缺陷」之一。然而，為什麼死亡會被視為一種缺陷？站在人類智能的立場，我們將死亡視為是生命體的

終極命運，在科學研究上，我們也將其視為一種「缺陷」，而非欣然接受自然的安排。因此，在人類歷史上，對於死亡的延續不僅在各種文化或神話中都可窺見一二，或者有各種大多數都是對於長命的祝福，以及對長生不老的期盼。甚至我們透過科學與科技來改善可能影響死亡的因素，或者強化延長生命的可能性。在希臘神話中，人類的缺陷被認為是從一個失誤 (fault) 而來，一個眾神在創造人類與其他生物時所犯下的錯誤。在創世之初，艾比米修斯 (Epimetheus) 被交辦分配給每一個生物體一種本能，但他最後卻忘了要分配給人類。也因此，人類便不像其他生物體，一出生便具有某種得以生存的能力。原初的人類沒有任何本能，這是所謂艾比米修斯的「失誤」，Stiegler 稱之為「原初的失誤 (original fault)」 (Stiegler, 1998)。這個失誤的根源是遺忘，也因此，艾比米修斯的智慧又被稱之為是一種後見之明，也就是所謂的「傳統」。epimetheia 這個字指的是傳承 (heritage)，以及透過學習、教學、關注及理解某事，因此，所謂的傳統就是，「生成於一個已經在那裡且只不過是技術性的失誤之中」 (Stiegler, 1998, p. 206)。儘管艾比米修斯忘記給予人類任何本能，但其失誤本身卻正好變成了人類的本質，亦即，一種缺乏 (absence) 的狀態。這個缺乏的狀態需要被填補，以彌補艾比米修斯因為遺忘所造成的失誤。人類的本質從此一缺乏狀態出發，卻讓人類具有將失誤予以外部化的可能性。另一位希臘神普羅米修斯 (Prometheus)，在知道艾比米修斯的失誤後，由於其憐憫並且想保護人類（因為人類是普羅米修斯用土與水捏製出來的），因此從眾神處盜了火與各種技術給予人類，因而成為人類填補這個缺乏狀態的開端。這個故事的啟發通常除了描述人類用火與技術的起源之外，或許更重要的意義是，人類因為這個缺乏狀態且為了填補這個狀態而開啟的人類本質從 de-fault（自然意義上）到 default（技術意義上）的轉變過程。這個轉變也構成了現代性中對科學與科技發展的想望，以及各種版本的科技烏托邦或者是科技樂觀主義。

普羅米修斯的行為在眾神眼中也是一種失誤，因為他盜竊了火與技術給與人類。因此，宙斯將其綁在高加索山上，並且每天派一隻老鷹去吃掉他的肝，但隔天他的肝會重新生長出來，如此日覆一日地承受被啄食肝臟的痛苦。普羅米修斯的失誤作為對艾比米修斯之失誤的彌補，讓人類得以不因為後者的失誤而無法生存。艾比米修斯的失誤，因為普羅米修斯的失誤而成為人類最重要也是根本的缺

陷，且普羅米修的失誤更是讓艾比米修斯之失誤得以外部化（externalization）的契機，而此一失誤的外部化，乃成為人類透過技術重新發現自身的開端。所謂的「重新發現」指的是，人類只有在使用技術之後，才得以認識到自身作為人類一而非其他物種一而存在。因此，Stiegler 說，「人類的存有就是外在本身（外在於自身）」，而普羅米修斯給予了人類「將其放在自身之外」的禮物。這個禮物毋寧指的就是「技術」。因此，技術就是讓人類能夠「將自身放在自身之外」。換句話說，也就是讓人類具有可以與自身（原本作為自然之一部分）建立距離，並且透過技術的協助而能夠將此距離予以客體化或客觀化的能力，甚至讓其成為是一種必須性，因為此一能力乃是人之所以能「成為人」的重要元素。這也是「輔具（prosthesis）」這個字的意思，pros-thesis 就是被放在前面的東西，也就是被放在外面的東西。假如那個「被放在外面的東西」其實就是那個將其放在外面之東西外部化的產物，那麼此存有就是外在本身，同時也是外在於自身。此也正符應了馬克思對資本主義作為一種技術的觀察，其不僅外加於人類，更是將人類異化於其自身的觀察。

艾比米修斯的這個「原初失誤」因為普羅米修斯的「失誤」而得以被彌補，此一雙重失誤因而構成人類的本質或預設（default）：或者，原初的去一失誤（original de-fault）。就此而言，「成為人」乃是一個不斷去一失誤的過程，也因此成為「成為人」的預設（default）。Stiegler 指出，包括語言在內，諸如發明、發現、洞見、想像等，都是此一 original de-fault 的特徵。De-fault 指的是「存有的缺陷（a flaw of being）」。「所有用來 de-fault 的東西都是外在於人的，因此是技術的（非自然的），也就是具有輔助性、是人造的。」「成為人」意謂著，人具有某種「人的性質」。然而，在 Stiegler 對人與其他物種、人與其環境之互動的討論中指出，「人性沒有性質，沒有預定，其必須發明、實現、製造出性質，並且也沒有任何暗示說，被製造出來的這些性質將會實現人性，或者這些性質將會變成是人類的性質，因為這些性質或許更會變成的是技術的性質（p. 193-194）」何謂人性（humanity）？如何理解技術與人性的關係？如果人性沒有固定的內容（預設），如同艾比米修斯的失誤所指出的，那麼我們如何理解人性？或者，因為遺忘所給予人類的預設就是，「人性的預設（default）就是不斷地去失誤（de-fault）」，我們可以說，人性本身就是技術的。人性的內容取決於技術的

內容或性質。然而，反過來說卻不一定成立，這些被製造出來的性質並無法用來完全地實現或展現人性，或告訴我們人性的真正內容，或其預設。相對地，我們只能從這些不斷透過技術所製造出來的性質中，窺見或接近人性的內容，永無止息—就像普羅米修斯的懲罰一樣，日復一日地在痛苦與恢復痛苦的循環中，證明自身依舊是那個已經與原初人類（being-in-fault）分離開來的「人（being-in-de-fault）」。

「成為人」乃是人類將其自身的那個原初失誤透過技術予以外部化的後果。這些技術包括傳統、信仰、文化，甚至是科學知識等。原初失誤並不會讓人類意識到自身與其他物種的不同，原初失誤作為一個遺忘的產物，其仍舊可以是一種屬於自然界中的本能。「成為人」的關鍵在於原初失誤透過技術予以外部化的過程。原初失誤的外部化，意謂著人類獲得或發展出去除這個原初失誤的可能性，也就是填補這些失誤所造成的缺乏狀態。然而，要能夠看到這個原初失誤並且更重要的是將其「當作是」失誤而需要去除，其首要條件即是創造出與其之距離，也就是將其放到人類自身之外，不再將其視為自身及自然的一部分。這個外部化過程也是人類將自身予以去自然化（denaturalization）的過程。人類本質（原初失誤）的去自然化，得以讓人類作為人類而出現，而與動物有所區別。但也正是去自然化，也就是透過技術或技術物來彌補原初之失誤，讓人類成為非人類而消失在自然之中。這樣的弔詭或許已然是預設在其本質之中，也就是那個缺乏狀態。「讓人類成為非人類而消失」意思是，人類在去自然化的過程中，將其自身予以客體化或外部化，在客體化後便可在客觀上予以消除。這個經由將人類予以客體化而出現的非人類，乃是一種非自然意義下的人類，這種非自然意義下的人類是技術的，或者是與技術相結合的（賽伯格，cyborgs）。這樣的結合讓屬於自然那部份的人類消失，留下的是人與技術之結合的後人類（posthuman）。從這個角度來看後人類，或許可以提問的是，後人類是不是非人類？如果是，那麼後人類事實上意味著人類的消失，消失在與技術的結合過程之中。去自然化讓人類得以作為與其他物種或自然之區別而「現身」在歷史或世界發展之中，但這樣的現身卻也標示了屬於自然的人類的消失，人類在歷史上的現身是一種技術意義上的現身（technologically present），而非自然意義上的現身。「一出現即消逝」或者，「以消逝為前提的現身」，不僅作為「成為人」的重要元素，此一

元素也符應於技術在人類歷史與社會中的發展傾向，「轉瞬即逝」一無論是時間、訊息、事件一無一不是具現了人類從一開始就具備的傾向或本質。人類的現身，必須依賴於其自身的消逝，消逝於技術將其外部化的過程中：「人類在其本質的去自然化中，乃是其自身之消失（Man is his disappearance in the denaturalization of his essence.）」（Stiegler, 1998, p. 121）。

在自然狀態，人類不會感受到何謂死亡，因為死亡是作為自然的一部份。在許多神話中，人死後並未消失，而是進入到自然或者以各種「存有」形式繼續存在。隨著科學、科技與各種技術物的介入，將人的死亡與自然切割開來，如今死後也不再是自然的一部分，「死了就是死了，那些讓人「成為人」的事物都不會留下，也不會回到自然」。技術介入帶來了將死亡從自然切割出去，變成是與人類自身不同的東西之可能性，也就是前述 Stiegler 所謂的外部化，或者客體化、對象化等過程。人類開始可以將死亡當作是外在於自身的東西來經驗、觀看或理解一藉由技術。這也是 Stiegler 所探討的可完美性（perfectibility）所具有的虛擬性（virtuality），一旦這個虛擬性得以透過技術或技術物而獲得實現（actualize），人也就得以跟自然與其他物種區別開來，自然狀態下的人也就因為技術或技術物的介入而逐漸消失。這也是 Stiegler 在《技術與時間》一書中所要呈現的主要概念。然而，他大概沒有想到的是，現代科技發展讓這一天更快速的到來。他說：

As long as the savage is not in the disequilibrium of freedom, as long as his perfectibility remains virtual and does not perturb the originary play between nature and his own nature, as long as his virtuality does not become real, that is *technical*, he does not have the feeling of death and does not anticipate: he is not *in time* (Stiegler, 1998, p. 122).

據此，原初的人類並不會對死亡產生畏懼。當人類開始害怕死亡時，他便受到了前述那種「自身的缺席（an absence in himself）」的損害（影響），因為其已經與那個原初之預設的（自然）自身分離了。Stiegler說，原初人類害怕的不

是死亡這件事，而是苦痛 (pain) (Ibid., p. 123)。因此，「原初的人類並非是會死的」(Stiegler, 1998, p. 126)。這句話的意思是，原初人類並未意識到死亡本身是一種需要被獨立出來、被關注、被害怕的事物：死亡是自然的一部分。換句話說：死亡尚未被外部化，死亡仍然是作為自然的人類的一部分。對死亡的意識、恐懼，亦即一種改變對自然時間的看法，取決於「形勢 (circumstances)」(Ibid.)。人類在面對各種形勢時，事實上是在面對人類自身的缺陷 (或缺乏)，也就是艾比米修斯的失誤。當人類被野獸包圍，或者是需要可以填飽肚子的東西時，人類便需要與這些形勢產生互動。形勢這個字的英文是circumstance，其字源指的是encircle、encompass，也就是包圍、圍繞之意。人類被某個情境所包圍或圍繞，因而使得人類必須回應或採取應變行為。沒有形勢，人類也將不需要回應，形勢開了一條路，或者作為以下之前提：讓沒有任何本能的人類，得以或必須依靠技術或技術物的介入來回應這些圍繞在其周圍的各種情況。也因為形勢的多變、存在及對其之回應，人類得以或必須脫離其原初之預設，因而建立與原初之預設的距離。這個距離經由技術 (物) 來彌補之，也就是Stiegler所謂的「縫隙」。這個縫隙 (距離) 因為科技發展而越來越大或越遠，這也是人類所面臨之形勢，橋接或彌補的方式也只能是透過技術 (物) 而行之。因此，circumstance提供或建立了兩種距離，或者使兩種距離得以可能，一種是人與其原初預設的距離，一種是人與其所面對之事物的距離。這個「圍繞」所造成或提供的「與原初預設之分離 (距離)」之可能性，或許也是形式理論的基礎，也或許是人類一作為人類而非自然中的動物一開始作為觀察者的起點。人類區別自身與其環境的可能性，乃是建立在人類與其原初預設之分離的基礎上。這也是Stiegler所謂的，內環境 (原初預設作為在人類自身內的環境) 變成了外環境 (技術作為人類自身之外的環境)。

De-fault 是一個雙重失誤的產物：艾比米修斯的失誤 (忘記給予人類一種性質或本能)，再加上普羅米修斯的失誤 (從眾神處盜火與技術給人類)，因而產生了「去一失誤」(Stiegler, 1998, p. 194)。這個「去失誤」也成為人類的預設狀態 (default)。這個雙重失誤或許作為一種辯證關係而出現，構成了「人之所以為人」的內涵。這邊的「人」意指的已經是那個與自身之缺乏性質建立了距離的「技術人」。與自身的缺乏性質建立距離，才能夠讓自己變成為是技術人，也才

能夠知道「為什麼自己是人」，而不是其他物種。換句話說，如果是處在一種沒有與自身之缺乏性質建立距離的狀態（無距離狀態），那麼人依舊是動物或自然的一部分。然而，因為普羅米修斯的失誤，將火跟技術給了人類，讓人類與自身之缺乏性質得以建立了距離，也正是因為這個距離，讓人類獲得關於技術的智慧與技術，用來彌補自身之缺乏性質，從而脫離動物或自然，成為凡人一知道何謂死亡、會對死亡產生畏懼與害怕的凡人。因此，技術將原本作為在場的失誤 (fault as presence)，轉變為作為缺乏的失誤 (fault as absence)，去失誤 (de-fault) 也就轉變成了預設 (default)。原本艾比米修斯的失誤，並非會造成一種缺乏的狀態，如果仍然是自然的一部分，與其他動物一般跟自然和諧共存，那麼這個失誤也不會變成是一種缺乏，失誤就只是失誤。但普羅米修斯將技術給了人類，讓這個原本只是在自然上的失誤，具有了非自然的意義。人類與自然分開了，原本的失誤也變成為一種缺乏，在人類與自然拉開距離的過程中。這樣的缺乏，便只能透過非屬於自然的技術來填補，而非是回到自然本身，在與自然的互動中找到出口。據此，普羅米修斯的失誤一盜火給人類一為人類帶來了原初的去失誤，也讓人類自此必須要不斷地透過外部化（或技術）來與那個原初預設劃定界線。

缺陷一艾比米修斯的失誤一在不斷透過技術予以彌補的過程中，出現了Stiegler所謂的從虛擬的可完美性 (virtual perfectibility) 到實際的可完美性 (actual perfectibility) 之轉化。人類的可完美性具有的只是一種虛擬性，因為遺忘所帶來的失誤，最終透過外部化過程，將其放到人類自身一自然一之外來處理。後者即是Stiegler所謂實際的可完美性概念。實際的可完美性，乃是在意識到自身之缺乏時才得以出現，並且會嘗試透過技術來達成。原初的可完美性之所以是虛擬的，是因為原初的失誤並未脫離人類自身，人類也沒有與其建立距離（脫離自身或外部化）。實際的可完美性則發生在外部化過程之後，在人類與其自身之原初失誤建立了距離，也因此，人類逐漸認為似乎可以透過技術達至一種完美狀態。³缺陷及其經驗的外部化，是人類用來不斷辨認出自身作為人類之基礎或證明。缺陷經驗的外部化，也帶給人類原初失誤本身便具有之潛力，亦即，提供或建立與動物或自然之距離，然後再透過技術來彌補該距離的可能性。從「去

³ 近年來發展出一項新的科技一仿生手，也就是讓失去手的人能藉此恢復功能，其功能甚至包括各種感知，也就是人類身體所具有的能力。試用者指出，仿生手也讓他感覺到自己是完整的。

一失誤」出發，我們更可能接近何謂「成為人」的內涵：「去失誤或許是使人類與自身之原初失誤建立距離，而讓「（技術意義下的）人類」一而非動物或原初的人類、自然的人類一得以出現的關鍵元素。若是回到人工智能本身的發展及運作，其並未有這樣的一種從內環境到外環境、從原初失誤到去失誤等過程，換句話說，人工智能與其自身是沒有距離，其無法成為觀察者而觀看自身，也無法具有反身性地對自身的觀察進行觀察，也因此無法有所謂的「像人一樣思考」而出現的可能性。人工智能是否能夠思考？答案應該是肯定的，但其可能更接近的是「像人工智能一樣思考」，而不一定是「像人一樣思考」。這也回應了Bostrom對人工智能的發展所提出的觀點：我們不應將其擬人化，「陷入人類的框架，都會萌生毫無根據的期待」（Bostrom, 2014, p. 116）。下節中便以自動駕駛技術為例，延伸「去失誤／預設」這組區別的內涵，作為重新理解人工智能的另種可能性。

三、「去一失誤／預設」作為重新理解人工智能發展的可能性—以自動駕駛技術為例

近年來的自動駕駛技術，不僅已經成為一個熱門的議題，同時也是一個對於未來移動的想像。全世界有許多企業與政府均投入相關技術與產業的研發與發展，例如 Google、Apple、Tesla、BMW、Uber 等。自動駕駛技術是以人工智能科技為基礎，企圖實現無人駕駛的各種運輸載具，包括汽車、公車、飛機、船隻等。2016 年，美國國家公路交通安全管理局（National Highway Traffic Safety Administration, NHTSA）依據美國汽車工程師協會（Society of Automotive Engineers, SAE）所定義的不同自動化程度，將自動駕駛分為五個等級：

等級 0：無自動駕駛。完全手動操作。

等級 1：弱駕駛輔助。人類負責主要駕駛任務，但機器可單一輔助方向或保持速度，例如定速巡航。

等級 2：部分自動駕駛。人類負責主要駕駛任務，但機器可根據環境因素進行自動操作，例如自適應巡航。

等級 3：有條件全自動。系統完成所有操作，人類做部分選擇回應。

等級 4：高度自動化。系統完成所有操作，人類不一定要做選擇回應。

等級 5：完全自動化。系統完成所有操作，人類不用做選擇回應，在所有情況下皆可正常行駛。

(資料來源：<https://technews.tw/2017/01/18/what-is-autopilot/>)

自動駕駛技術的原理基本上就是將汽車視為一種機器人，透過各種感測器感測路況與周邊情況，傳輸到處理器，處理器根據人工智能對情況進行判斷，接著通知電機系統，電機系統根據信號操控機械裝置，進而控制車輛進行各種動作。⁴基本上就是感測技術、定位技術與決策技術（人工智能）三者結合的產物。⁵感測技術以影像偵測辨識技術等視覺科技為最重要的一環，也是目前較廣泛的應用範疇，因此目前的無人車均配備有攝影機或雷射感測器，其目的就是要讓無人車可以「看見」並且可以據此進行判斷與動作。⁶這也是人工智能技術非常仰賴 Nvidia 這類的影像處理晶片技術大廠的主要原因之一。Nvidia 自動駕駛部門總監 Danny Shapiro 便指出，人工智能的精準程度，在很大程度上取決於「極端的」計算能力與大量代碼，以及強大的電腦及影像處理元件，「人工智能需要決定前面的物體是一個人、車輛、消防栓還是什麼別的東西」。⁷在同一篇文章中也提到，Uber 高級工程經理、美國卡耐基梅隆大學研究教授 Jeff Schneider 指出，汽車交通事故大多肇因於兩類錯誤，其一是認知錯誤，其二是決策錯誤。他認為，認知錯誤可以透過感測器、雷達、攝影機、Lidar（遠端感應系統）等技術來避免，這些資料再透過人工智能系統進行計算，最後形成決策。此外，他也提到，提高精確度的方式之一就是確保自動駕駛車輛不會搞錯各種不確定、模糊不清的情況。自動駕駛技術涉及到的同時是對外部環境與車輛內部資訊的搜集、判讀、動作與監控，並且兩者的結合與互動或許是自動駕駛技術是否能夠成功的關鍵。然而，這其中仰賴的人工智能技術，仍有許多問題有待解決，最重要的仍是在面對道

⁴ 資料來源：

<https://www.stockfeel.com.tw/%E7%84%A1%E4%BA%BA%E9%A7%95%E9%A7%9B%E5%88%86%E5%9B%9B%E5%80%8B%E9%9A%8E%E6%AE%B5%EF%BC%8C%E6%88%91%E5%80%91%E7%8F%BE%E5%9C%A8%E5%88%B0%E5%93%AA%E5%80%8B%E9%9A%8E%E6%AE%B5%E4%BA%86%EF%BC%9F/>

⁵ 資料來源：<https://www.techbang.com/posts/49294-is-your-exclusive-driving-the-car-the-car-a-get-operation-principle-and-development-of-unmanned-vehicles>

⁶ 例如：<https://www.bnext.com.tw/article/42885/how-self-driving-work>

⁷ 資料來源：<http://www.knowledgeatwharton.com.cn/zh-hant/article/9327/>

路上複雜情況時，能否做出精確判斷與動作的問題。這部分也因為人工智能並未能夠像人類智能一樣，透過經驗、情感、個人認知等對真實路況進行判斷，因此對外部資訊的搜集與判讀之需求便更加重要。

自動駕駛技術在高速道路上的試驗較為成功，原因是高速道路上的相對而言較為單純許多。一般道路上的情況更為複雜，不只是交通號誌、行人、自行車等各種車輛，還有許多突發性的情況，例如自行車或行人經常有穿越道路的行為舉動，或者道路邊可能有障礙物等。這些也都需要仰賴外部資訊的搜集以及對其之判讀的人工智能技術。儘管近年來全世界對自動駕駛技術的未來都抱持樂觀看法，但近日在美國亞利桑那州所發生的一起 Uber 自駕車撞死橫越馬路的行人事件，或許也凸顯了一些過去未受到關注的問題。⁸事件發生後，Uber 也立即停止其在北美的測試計畫，待問題釐清後再議。在稍後公布的現場影片中，我們則可以看到行人突然出現在車輛行經道路上，且當下駕駛並未接管自動駕駛，因此造成遺憾。在事後的報導中，除了探討自動駕駛車輛似乎未偵測到異狀而進行減速，以及駕駛當下的分心行為之外，也有報導指稱可能是行人的錯，而非自駕車的問題。⁹無論行人是否突然出現，至少可以肯定的是，自動駕駛技術在面對如此突如其來的情況時，依舊難以快速回應。自動駕駛車輛在測試期間，便已經有不少問題出現，這些問題大多被歸因於與人類駕駛的魯莽行為有關。¹⁰道路情況的複雜性，不僅考驗人工智能技術，同時也在逐漸形塑人類智能與人工智能如何共存的問題。隨著人工智能的發展，人類智能似乎越來越被視為是對其之阻礙，而非是與之合作的對象。這在理解人工智能與人類智能之差異上，似乎不僅是道德上的議題，更是一個科技哲學上的問題。

就前述自動駕駛的相關技術為例，這些技術的首要目標在於廣泛且儘可能詳細的資料搜集，以及對這些資料進行精準便是與判讀的能力。換句話說，**技術的目的在「窮盡 (exhaust)」關於自身與其環境中儘可能的現象。其邏輯假定，假如搜集的資料越齊全、完整，就越有可能進行更精確的解讀，以便有更正確的回**

⁸ 資料來源：<https://tw.appledaily.com/new/realtime/20180322/1319658/>

⁹ 資料來源：<https://buzzorange.com/techorange/2018/03/22/temp-police-release-uber-car-crash/>

¹⁰ 資料來源：<https://chinese.engadget.com/2016/09/26/googles-self-driving-car-is-the-victim-in-a-serious-crash/>

應與動作。換句話說，「窮盡」乃是科學以及當前的人工智能科技所依賴仍舊是最重要的原理原則之一。人工智能科技的運行要有更精準的解讀，仍需要搜集儘可能詳細的資料以供處理。在交通號誌的辨識、行人與車輛的行徑與動作的解讀等，都還是在一個「窮盡所有可能性」的選項下來發展。此一透過窮盡以作為理性與客觀之基礎，或許也構成了人工智能乃至於現代科技的道德角色內涵。其背後的語言似乎是，如果窮盡一切後仍有失誤，那麼就是人類智能的問題，而非技術的問題。因此，對自動駕駛技術而言，**窮盡其預設的內涵，而非去失誤 (de-fault)**，去失誤反而成為其環境。這是說，窮盡是作為解決去失誤的基礎，在操作上則是其主要方式。如果用形式理論觀點來表示，就是「**預設／去失誤 (default/de-fault)**」這組區別或形式 (form)。人工智能在努力邁向「像人一樣思考」的道路上，更可能用來說明這個形式本身的內涵，儘管技術可能都具有或多或少的類似性質。Default 除了譯為「預設」之外，經常也會看到有「默認」之譯。這組區別或許可用來理解人工智能系統的形成。這意思是說，人工智能系統在形成自身的運作上，乃是使用了這組區別，並且在預設的這一端複製這組區別，以此而使其得以持續地維持其預設之運作，也就是透過窮盡作為與其環境之間的溝通。就此而言，去失誤乃為其預設之環境，前者所具有的複雜性便可能遠高於後者。人工智能系統以窮盡作為其預設，而對來自其環境中的 de-fault 進行化約，以使人工智能系統得以持續運作。也就是說，以自動駕駛技術來看，其所使用的人工智能科技必須要首要能夠辨識並理解周遭的事物，以及這些事物及其動作代表的可能意義。然而，不只是事物本身，還有這些事物的運動，以及事物彼此之間、事物的運動彼此之間所可能具有的關係，都必須要能夠為人工智能科技所觀察到並且予以解讀，否則其回應將會遇到一定程度的困難或阻礙，甚至造成誤判。對人工智能而言，窮盡作為其「預設」，可能是一組相對其環境而言簡單或單純的「程式」，這組程式可以在一定程度上對其環境進行窮盡的運作——亦即化約複雜性，儘管大多數情況並非能夠如其所願，但「去失誤」作為其環境，其複雜性或許便已經遠高於人工智能系統的「預設」本身。

人類智能則不同於人工智能，如前面所述，去失誤乃為其預設，其表現為「去失誤／預設 (de-fault/default)」，預設反為去失誤的環境。就 default 是一種缺乏狀態而言，de-fault 的運作是期待能夠補充或彌補這個缺乏狀態，而非要窮

盡它。這或許也是人類智能至今仍具有諸如創造力、想像力等人工智能尚未能夠具有的能力之原因。從 Stiegler 的觀點來看，being-in-de-fault 是發明、發現、創造、想像的來源，正因為人類沒有被指定的性質，所以人類反而具有跳脫其僅作為自然之一部分的潛力。人類智能面對的是原初之缺乏、缺陷以及對其之經驗，就人類智能而言，default 作為其環境反而提供了豐富的可能性，亦即具有較 default 高的複雜性。人類智能的學習與認知世界—彌補其缺乏狀態—之過程，更像是 Garfinkel 的索引式表達 (indexical expression) (Garfinkel, 1992[1967], pp. 4-6)。相較於索引式表達，人工智慧技術已經可以非常輕易地完成客觀式表達—理性且客觀的資料搜集與分析，但對於前者卻仍持續在嘗試發展出有效的方式，主要原因在其在相當程度上是主觀的，每個人的意識索引系統或許都不同，索引式表達是構成社會成員對某個場景或脈絡之解明的重要元素，社會成員對場景的解明呈現或多或少的差異性，也是因為每個人所使用的索引都有些許差異。因此，同樣都是過馬路的場景，不同的參與者因為攫取的索引不同，也會導致其對該場景的解明有差異，進而引導其對場景的回應有所差異。

人類智能並非是以清晰、精確、理性、客觀等方式來獲取關於其環境(default)的資訊，而毋寧是一種索引式的，必須要依賴線索的、相對來說是主觀的方式。如果我們以神經科學哲學研究的粗淺論述來看，大腦神經元的運作對人類智能（部分展現為意識）之生成的最可能貢獻，或許是其生產了「模糊性」這樣的狀態。模糊性是神經元運作過程中所突現出來的性質。模糊性可能是意識生成的基礎，或者其可能也可以用來解釋主觀經驗的內涵。大腦提供了記憶的功能，不論是長期或短期記憶。無論是哪一種記憶，在某種程度上，我們應該可以同意，這些記憶都不是精確的。大腦運用各種方式來記錄，但每一種記錄都只是一種索引式的紀錄，只有記憶的部分或片段被記錄下來，而非所有細節。心理學研究也指出，即便是再精確的記憶都具有一定程度的模糊性，也就是都有著不那麼精確的部分。這些部分或許即是意識運作的基礎。相反地，人工智能所依賴的「窮盡」以及電子計算機的記憶等，都是以精確、無遺漏、鉅細彌遺為基礎，以至於其無法從中茁生出模糊性。若再回到自動駕駛技術，以辨識物件而言，大腦的辨識運作經常呈現的是一種模糊性操作，大腦綜合了許多資訊，例如物件的顏色、紋理、形狀、出現的位置，以及這些特徵與記憶的連結，以及彼此間的互動關係等，但

這些連結並非都是確定的，而是在一堆可能的連結中，模糊性地比對，然後產生出可能是正確的回應。我們不需要精確辨認出物件的紋理、顏色、形狀，使用了以模糊性為基礎而建立的各種連結，在這些連結中找到最可能的答案。所謂在意識層次上的「確定」，諸如「我確定這是一件白色背心，且我知道這是我的衣服」，可能也只是一種大腦運作出來暫時性的結果，這些暫時性的運作產物也就是以模糊性為基礎所建立的各種可能連結。

這個預設狀態就如同其另一半「去失誤」一樣，只是辯證過程的一部分。如果將「預設」當作是系統運作上的真值，也就是被標示出來的那一端，那麼無論是強人工智能或者是深度學習（deep learning），就可能會朝向將可完美性視為是實際的（actual），而非虛擬的（virtual）。「去失誤」作為「預設」的環境，前者的複雜性要大於後者。如果從複雜性觀點來看，系統藉由自身之運作以化約來自環境的複雜性。在此運作中，系統是簡化了環境的複雜性，卻無法處理或涵蓋其環境中所有的複雜性。假如環境的複雜性總是高於系統之複雜性此一假設是可接受的，那麼無論系統如何運作，或提高其自身之複雜性，在實際上都無法處理環境中的所有複雜性。如此一來，人工智能若只是在發展一種理解（化約）其環境複雜性的運算系統，那麼其目的可能永遠無法達成。然而，如果將辯證的另一端「（去）失誤」視為系統運作的真值，或許情況會有不同，而這也是人工智能科技發展上正在發展的一塊。

四、代結論：「例示／窮盡」作為對人工智能技術之反思

綜上所述，人工智能是在虛擬世界中看真實世界，而相反地，人類智能卻是在真實世界中創造虛擬性，例如符號系統、語言、制度等。在人工智能的眼中一假如其真的看得到的話，真實世界只是一堆數字跟代碼。相對地，感官經驗的數位化資訊在人類智能的眼中，也只是一堆數字跟代碼。人類智能不知道那是什麼，也無法從那些代碼中聞到味道或者看到東西。例如，我們無法從代碼中知道這棵樹的樣子、形狀或者是顏色，甚至是這棵樹給我們的感覺、情緒等。在人類智能的眼中，我們是以「如其所是」的方式來觀看世界，相反地，在人工智能眼中，

乃是用「如其所不是」的方式來看世界。也因此，當後者想要「進入」或「看到」真實世界的企圖時，其就必須要「窮盡」這些「如其所不是」的事物。只有在窮盡之後，才有可能或認為其是所理解的「如其所是」的世界。然而，人類智能的如其所是，並非是先窮盡才知悉，而是一種「例示 (illustration)」的方式。我們透過如其所是來建構關於這個世界的知識，透過如其所是來知道何謂如其所不是，而非相反。這兩個路徑，「窮盡／例示」，或許正指出了人工智能的盲點之一。

從自動駕駛技術與人工智能之應用來看，「窮盡」具有兩種意涵，其一是其需要儘可能的透過各種感測器搜集環境數據，儘可能全面且精確地辨識這些來自環境中的訊號，換句話說就是窮盡其可能性。其二是其也需要「窮盡自身 (exhaust itself)」，亦即，窮盡其運算法本身所可能提供的判讀與回饋的選項。人工智能的精確性與可信度，必須要結合兩者，並且在兩者中取得一定程度的平衡，才有可能更好地應用在日常生活中，或者可以跟人類智能並存，而非僅是其輔助：



Pic. 1: Learning process of HI and AI (2018).

自動駕駛技術是可能的，即便如 Level 5 的自動駕駛汽車，也就是全自動駕駛、無需人工干預、沒有方向盤、油門或汽車後視鏡等的完全自動駕駛汽車，都是可以預期會有實現的一天。然而，我們也可以輕而易舉地就理解到，前述的自動駕駛都是在受到控制的環境中——也就是對人工智能而非人類智能友善的環境——獲得實現。這些交通工具基本上不需要與人類智能互動，不需要處理因為人類智能所可能導致的各種情況，也因此可以避免這些可能無法解讀的風險。換句話說，人工智能科技若仍是以「窮盡」為其技術邏輯，那麼距離要能夠「像人一樣思考」的人類智能水準，可能還有很長的路要走。此外，這個實現更可能會是以犧牲人類智能為代價，才有可能達到。換句話說，我們需要創造出來的可能是一個沒有人類智能或者人類智能最小化的世界或日常環境。如此一來，人工智能便可暢行

無阻，而且可以完美的執行。這樣的世界或日常生活，有可能不再是人類智能優先的世界，而是人類智能如何適應人工智能的世界，而非相反。機器出錯，我們會責怪人類，而非機器為何無法適應人類智能。例如，為什麼你不遵守號誌、不走斑馬線、不依照規定來行動等。因為這些地方才有人工智能可以準確執行的餘地，我們必須遵守號誌、走斑馬線、依照能夠順利讓人工智能運作的規則來行動，如此一來，人工智能才不會出錯、才不會撞上我們、才不會無法辨識我們的行為等。

我們（人類智能）允許「認錯人」這件事，是人類智能之所以為人類智能的關鍵元素，也是人類智能學習的開端。相對地，我們並未容許人工智能也能夠有跟人類智能一樣的錯誤。當我們期待開發出能夠並駕齊驅或甚至是超越人類智能的人工智能時，我們並未同意人工智能也可以犯下跟人類智能相同的錯誤，或至少在實際的運作上，我們乃是努力地期待或想要解決人工智能總是不那麼精確的部分或總是會出錯的結果，例如自駕車上路後遇到的各種問題，在對各種情況的判斷中，我們不容許其像人類智能一樣擁有犯錯的權利，例如對道路上人類行為的判讀，總是在避免造成「對人類的傷害」，但這樣的「避免」卻是在「不允許犯錯」的基礎上，而非是在「允許犯錯，但可以更小心」的基礎上，而後者即是人類智能在學習上與人工智能的差異，讓其可以「更精確地」完成某些事。換句話說，我們對人工智慧的開發與期待上，在形式上受到人類智能的限制。我們期待人工智能可以如人類智能一樣，具有解決複雜問題、抽象思考、創造性的能力，但我們卻不容許其有犯錯的權利。當我們想要儘可能地避免錯誤發生時，或許也不可避免地再次落入「窮盡」的思維模式。儘管「窮盡」的本質便是在提供所有可能性，並且從中給予正確、有效、有效率的判斷，但也正是此一窮盡，在很大程度上脫離了人類智能的本質，亦即，對缺陷或錯誤的理解及允許，或許才是其獨特之處。由此也可以看出來，過去六十年人工智慧的研究與發展，逐漸朝向模仿自然或仿生的方向前進，例如神經網絡及遺傳演算法等。我們依舊認為，如果可以理解大腦神經網絡或生物體的運作，便可以透過科技製造或模擬出一個一模一樣的東西。然而，延伸 Stiegler 的角度來看，如果缺陷經驗是「成為人」的基礎，也正是其雙關「去一失誤／預設」在人類這邊是不斷地在前者複製該區別，「成為人」的預設便是一種「不斷去失誤」的過程，也就是「不斷地去失誤」本

身就是人類智慧的目的。因此，一但失誤可以真正地被去除，那麼或許也代表人類智慧的終結——人類將不再能夠「成為人」，而是其他東西 (something else)。

參考書目：

Bostrom, N. (2014). *Superintelligence: Paths, Dangers, Strategies*. Oxford, UK: Oxford University Press.

Clark, A. (2003). *Natural-Born Cyborgs: Minds, Technologies, and the Future of Human Intelligence*. Oxford & New York: Oxford University Press.

Clark, A. (2013). Re-Inventing Ourselves: The Plasticity of Embodiment, Sensing, and Mind. In M. More & N. Vita-More (Eds.), *The Transhumanist Reader*. West Sussex, UK: Wiley-Blackwell.

Gardner, H. (2011). *Frams of Mind: The Theory of Multiple Intelligences*. New York: Basic Books.

Garfinkel, H. (1992[1967]). *Studies in Ethnomethodology*. Cambridge: Polity Press.

Nilsson, N. (2009). *The Quest for Artificial Intelligence: A History of Ideas and Achievements*. Cambridge: Cambridge University Press.

Stiegler, B. (1998). *Technics and Time, 1* (R. Beardsworth & G. Collins, Trans.). Stanford, California: Stanford University Press.