

本書學習地圖

什麼是 AI

AI 演算法可以處理資料來解決各種困難的問題。不同種類的问题需要用到不同的演算法。還能同時使用多種演算法來解決更複雜的問題。

搜尋基礎

無資訊的搜尋演算法是透過搜尋所有可能的路徑來找到最佳解，但運算成本相當昂貴。這件事強調了其他 AI 演算法需要更合用的資料結構。

進階搜尋

有資訊的搜尋演算法運用啟發法來搜尋更佳解，且可被用於須將其他代理視為因子的對抗問題中。

進化演算法

基因演算法運用了演化的概念將可能的解決方案編碼起來，並經過多個世代的演化來找出效能更好的解決方案。

群體智能：螞蟻

螞蟻最佳化是根據真實世界中螞蟻群的運作而來，相關概念可用於探索新路徑並記得過往已發現的較佳路徑。

機器學習

迴歸與分類演算法可學習資料中的樣式來針對變動數值或案例類別進行預測。良好機器學習模型的關鍵在於對於資料的充分理解與妥善準備。

群體智能：粒子

粒子群體最佳化是根據真實世界中鳥群的移動方式而來，相關概念可用於探索區域的解空間，並記下群體已發現的良好解。

強化學習搭配 Q- 學習

強化學期運用了試誤法搭配在環境中執行動作之後所收到的獎勵與懲罰，藉此學習為了達成某個目標所需的良好動作。

類神經網路

類神經網路可說是人腦與神經系統運作方式的鬆散建模結果。神經網路會接收、加權並處理訊號，最後根據在輸入訊號中所找到的關聯性來提供一個結果。



前言



本書前言意在說明科技的演進、對於自動化的需求，以及在使用 AI 塑造未來時，我們有責任作出符合道德倫理的決定。

對於科技與自動化的迷戀

綜觀歷史長河，我們始終渴望能夠解決各種問題，還要減少體力勞動與人力介入。我們一直為著生存而努力，也透過開發各種工具與任務自動化來保留自身能量。有些朋友可能抱持不同觀點，認為我們是藉由在解決問題或在創作文學、音樂與藝術之過程中尋求創新的美好意志，但本書並非要討論人類本身的哲學問題。本書提供了對於許多人工智慧（**artificial intelligence, AI**）方法的總覽，並可有效處理許多真實世界的問題。我們解決了許許多多困難的問題來讓生活更輕鬆、安全、健康、更滿足，當然也更快樂。所有你在過去歷史與現今全世界所看到的發展，當然也包含了 AI，都是為了滿足個人、群體與國家的各種需求。

為了塑造未來，我們就必須理解過往的一些關鍵里程碑。在多次技術革命中，人類的創新不但改變了生活方式，也造就了我們與世界的互動方式，以及我們對其的思考方式。我們不斷改良所使用的工具來做到這件事，也因此開啟了更多未來的可能性（圖 0.1）。

以下關於歷史與哲學的簡圖，單純是想幫助你建立對於科技與 AI 的最基礎理解，並在製作專案時鼓勵你做出負責任的決策。

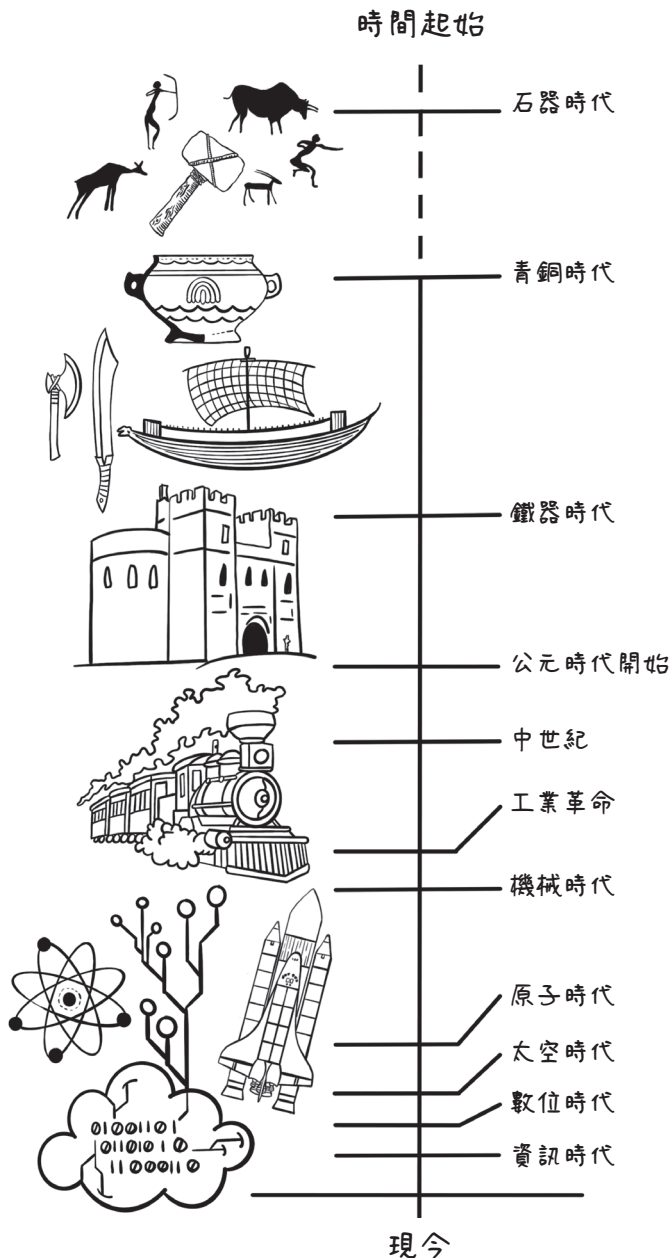


圖 0.1 科技演進在歷史長河中的簡易時間軸

在上述時間軸中，請注意較接近現今的里程碑變得愈來愈密集。過去的 30 年中，最顯著的進展就是微晶片的大幅改良、個人電腦大量普及、連網裝置遍地開花，以及產業數位化之後打破了實體疆界，並將世界連接了起來，這也正是為何 AI 成為一個值得追尋的可行領域。

- 網際網路已連通全世界，使得大量收集幾乎是所有事物的資料變為可行。
- 運算硬體的效能演進，使我們得以使用已往所收集的巨量資料來執行各種早期的演算法，同時還能在過程中探索新的演算法。
- 各個產業已看到進一步運用資料與演算法的重要性，希望能做出更優質的決策、解決更困難的問題、提供更好的解決方案，以及讓我們的生活更好，如同最初的人類所做的一樣。

雖然我們傾向於把科技發展的過程看為線性的，但回顧歷史，這個過程更像是指數型成長，未來也會是這樣（圖 0.2）。科技的進展一年比一年更快，雖然一直都要去學新的工具與技術，但支撐這一切的依然在於解決問題的基本原則。

本書彙整了一些有助於解決困難問題的最基本概念，但同時也希望能讓一些困難的概念變得更易理解。

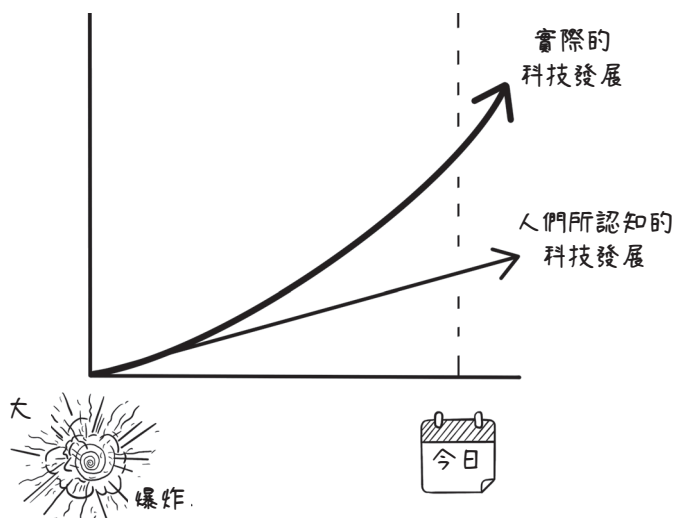


圖 0.2 人們所認知的科技發展與實際的科技發展

不同的人對於自動化的理解可說是大不相同。對技術人員來說，自動化可能是指透過程式腳本來讓軟體開發、部署與發佈更順暢無縫以及更少出錯。對工程師來說，則可能是提高工廠生產線的效率，使其產出更高或更少瑕疵。對農夫來說，又變成了透過自動拖拉機與灌溉系統搭配相關工具來讓作物產出最佳化。相較於人力介入所能達到的程度，自動化是透過減少人力需求來提高生產力或增加更多附加價值的絕妙解方（圖 0.3）。

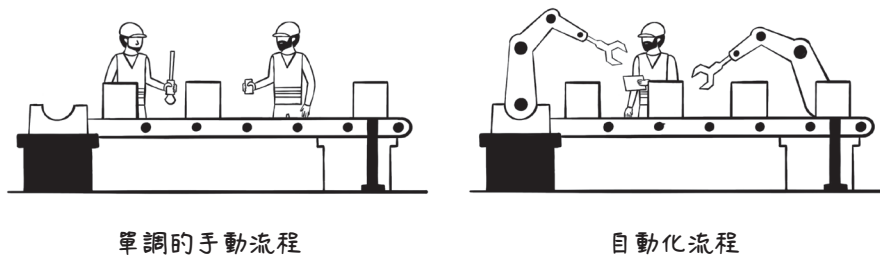


圖 0.3 手動流程與自動化流程

如果要想出一個反自動化的理由，最主要的理由是當這個任務需要在狀況中考量到多種看法、需要抽象化創意思考，或需要理解社交互動與人性本質時，我們人類可以把事情做得更好、不易出錯，精確率也更高。護理師不僅僅是完成工作而已，還會關心與照顧他們的病人。相關研究也顯示，照顧他人這類的人際互動是康復過程的關鍵因素。老師也不僅僅是傳遞知識而已，還要找到極富創意的方式來呈現知識、指導，並根據學生的能力、人格與興趣來引領他們。也就是說，藉由科技所達成的自動化有其一席之地，當然也有一塊是留給我們人類的。有了今時今日各種創新，透過科技所達成的自動化將成為所有職業的親密戰友。

道德規範、法律問題與我們自身的責任

你可能會好奇為什麼在技術書籍中會有一段講到道德與責任。好吧，隨著我們逐漸邁向一個生活方式與科技密不可分的世界時，創作某項科技的人所擁有的影響力實際上比它們自身所理解的來得更大。微小的貢獻就能造成巨大的連鎖效應。重點在於我們應該心存善意，並力求所做所為的產出不會造成其他傷害（圖 0.4）。

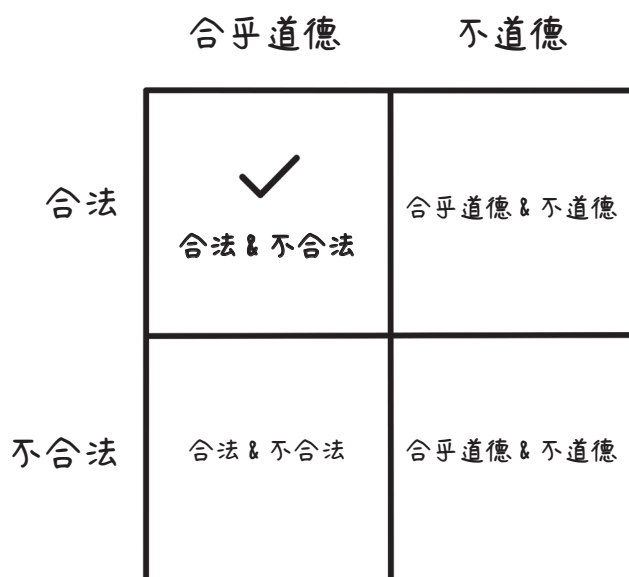


圖 0.4 致力讓科技應用既合法也合乎道德

意圖與影響：理解你的眼界與目標

在開發任何東西時，不管是新產品、服務或軟體，一定會面臨到的問題就是其背後的意圖。你所開發的軟體將會對世界帶來正面影響，還是你是心懷惡意的？你有思考過所開發的東西可能帶來更深遠的影響嗎？企業總是能找到能夠更賺錢更強大的方法，這也正是企業成長的意義所在。它們運用各種策略來找出最佳方法，諸如打敗對手、獲取更多顧客，以及變得更有影響力。也就是說，企業必須自問其出發點是否良善，不只是為了生存，更是要讓其顧客與整個社會更好。

許多有名的科學家、工程師與科技人士已同聲表達有必要管理 AI 的使用方式來避免誤用。即便是個人，我們也要負起道德上的義務來做那些對的事情，並建立強健的核心價值觀。當你被要求做一些違反自我原則的事情時，就有必要為這些原則大聲疾呼。

非預期用途：防範惡意使用

找出非預期用途並設法防範，是非常重要的。雖然這似乎很明顯也不難做到，但其他人將如何使用你的創作實際上是非常難以得知的，更難的是去預測這是否符合你的，以及組織的價值觀。

以 Peter Jensen 在 1915 年發明的擴音器為例。擴音器原本是叫做 Magnavox，最初是在舊金山對廣大群眾播放歌劇音樂，算是相當善意的科技用途。不過德國的納粹政權則有不同的想法：他們將擴音器放在公共空間，這樣一來所有人只能被迫聽到希特勒的演說與宣言。也由於這些長篇大論根本避無可避，人們就更容易受到希特勒的主張所影響，而在這時間點之後，納粹政權在德國就獲得了絕大多數的支持。這並非 Jensen 對這項發明所能預見的用途，但他對此已無能為力了。

時代不斷改變，我們對於所製作的東西的掌握度愈來愈高，尤其是軟體。要得知你所創造的科技會如何被使用依然是一件難以想像的事情，但幾乎能拍胸脯保證的是，不論後果是好是壞，一定會有某人找到一個你從未想過的方式來使用它。在這個事實之下，我們身為所工作的科技產業與組織中的專業人士，一定得想方設法來減緩各種惡意用途。

非預期偏誤：做出所有人都適用的方案

在製作 AI 系統時，我們會用到自身對於相關脈絡與領域的理解。我們也會使用各種演算法來找出資料中的樣式並據以回應。無法否認的是，偏誤就在我們四周。偏誤（bias）是指對於某人或某一群人的偏見，包括但不限於他們的性別、種族與信仰。這些偏誤中許多都是來自全世界各種社交互動、歷史事件以及文化與政治觀點中的突發行為。這些偏誤會影響到我們所收集的資料。由於 AI 演

算法正是要處理這些資料，會造成一個無可避免的問題：機器也將「學會」這些偏誤。

從技術觀點來說，我們當然有辦法讓系統完美運作，但到了最後，與這些系統互動的還是我們人類，因此把偏誤與偏見盡可能降到最低就是我們的責任了。我們所採用的演算法的表現，最多只能和提供給它的資料一樣好。理解資料與其所被使用的脈絡是對抗偏誤的第一步，而理解程度則能幫助你打造出更棒的解決方案，因為你對於所要處理的問題空間會更加熟練。提供良好平衡的資料並把偏誤降到最低，應該就能產出更好的解決方法。

法律、隱私與同意：理解核心價值的重要性

我們所做所為的法律觀點可說是重中之重。考量到社會的整體利益，法律規範了哪些事情是我們可以與不可以做的。由於許多法條是在電腦與網路之於生活相較於今日還不那麼重要時所訂定的，因此在如何開發相關科技以及該科技可被允許用來做哪些事情上，可能會有許多灰色地帶。也就是說在適應科技的快速創新方面，法律的變化相對是慢多了。

舉例來說，藉由在電腦、行動電話與其他裝置上的互動，我們幾乎無時無刻都在對自身的隱私做出讓步。我們正在把大量關於自身，有些甚至非常私密，的資訊傳送出去。這些資料是如何被處理與儲存的？在製作解決方案時應該要把這些事情考慮進去。人們應得以選擇要擷取哪些資料；如何處理以及如何儲存；資料的使用方式；以及誰有機會存取這些資料。在我的經驗中，人們幾乎都會接受那些會運用其資料來改良產品並在生活中加入更多價值的方案。最重要的是，當人們面對一個受到重視的選項時，他們會更樂意接受。

奇異點：探索未知

奇異點（*singularity*）的概念是我們做出了一款 AI，它的智能高到一個程度使得它能夠自我改良並擴充智能，至終到了一些階段，它成為了超級智慧。大家擔心的是，這個超出人類所能理解的影響層面可能會因為某些我們無法理解的原因，

使得人類文明遭到顛覆。某些人會擔心這樣的智能可能會將我們人類視為一種威脅；又有一些人的看法是，超級智能看待我們正如我們看待螞蟻，我們通常不會特別關心螞蟻或覺得牠們的生活方式與我們有什麼關係，但如果我們被牠們激怒的話，就會把牠們單獨處理掉。

不論這些假設對於未來的陳述是否正確，我們都必須對自己所做的決定負責並時時自省，因為它們至終會影響到個人、一群人甚至整個世界。

關於本書



本書運用了各種相關比喻、實用範例與視覺說明，希望能讓人工智慧演算法對於科技產業的一般人士來說，在理解、實作及解決問題等方面得以更平易近人。

本書是為誰所寫

本書針對各種理論以及數學證明提供了實用範例與視覺說明，寫給想要明白人工智慧相關概念與演算法的軟體開發者以及軟體產業人士。

只要對於電腦程式概念有基本理解，包含變數、資料型態、陣列、條件敘述、遞迴、類別與函式，有任何一種程式語言的經驗便已足夠；另外也適合具備基本數學觀念的讀者，例如資料變數、函數表示法以及把資料與函數繪製成圖表。

本書架構

本書包含了十個章節，各自聚焦在不同的人工智慧演算法或方法論。本書首先介紹了基礎演算法與相關概念來打好基礎，以便學習後續章節所介紹的各種複雜演算法。

- 第 1 章 —— 人工智慧的基本觀念，介紹了關於資料、問題類型、演算法的分類與範例，以及人工智慧演算法用途等等的基本常識與概念。
- 第 2 章 —— 搜尋演算法基礎，介紹了資料結構的核心觀念、簡易搜尋演算法的原理與其用途。
- 第 3 章 —— 智慧搜尋，延續先前的簡易搜尋演算法，並進一步介紹在找解上更有效率，以及可在競爭型環境中找解的搜尋演算法。
- 第 4 章 —— 進化演算法，深入介紹了基因演算法的運作原理，其中問題的解是藉由模仿自然界中的演化過程來迭代產生並改良。
- 第 5 章 —— 進階進化演算法，本章是基因演算法的延續，並談到關於如何調整演算法各步驟的進階概念，藉此來更有效地解決不同類型的問題。
- 第 6 章 —— 群體智慧：蟻群，本章談到了群體智能的基本觀念，並實際示範了蟻群最佳化演算法如何根據螞蟻的生活與工作方式來解決各種困難的問題。
- 第 7 章 —— 群體智慧：粒子，接續群體演算法並深入說明何謂最佳化問題，以及由於粒子群體最佳化方法可在大型搜尋空間中找到良好解，因此也可用於處理這類問題。
- 第 8 章 —— 機器學習，說明機器學習的工作流程，涵蓋資料準備、處理、建模、測試。談到了運用線性迴歸來解決迴歸問題，以及使用各種決策樹來處理分類問題。
- 第 9 章 —— 類神經網路，說明了在訓練、運用類神經網路來找出資料中的樣式並進行預測時，所需的基本觀念、邏輯性步驟以及數學計算過程；同時也會強調類神經網路在機器學習流程中所扮演的角色。
- 第 10 章 —— 使用 Q-學習進行強化學習，介紹了強化學習的重要觀念，從行為心理學開始，一路談到如何使用 Q-學習演算法讓代理學會其在環境中所做的決策品質好壞。

本書章節應從頭依序閱讀。相關的概念與理解會隨著章節一路往下而慢慢建立起來。讀完各章之後，參考本書 GitHub 所提供的 Python 程式碼來嘗試並獲得各演算法在實作上的實務性見解，這是相當有幫助的喔！

關於程式碼

本書整理了許多偽代碼，目的是為了專注於演算法背後所蘊含的基本觀念與邏輯思考，並確保不論喜歡哪一種程式語言，本書所提供的程式碼對大家來說都是可用的。偽代碼是在程式碼中說明相關指令的一種非正式方法。它希望能做到更易讀易懂；基本上就是對讀者更友善。

正因如此，本書所提到的所有演算法都在 GitHub 上提供了可執行的 Python 範例（<http://mng.bz/Vgr0>）。設定教學與相關註解都已在原始碼中說明了，以便在你的學習過程中給予幫助。本書建議的學習方式就是逐章閱讀，接著參考程式碼來加強對於相關演算法的理解。

Python 原始碼是作為實作演算法時的參考。這些範例是針對學習來最佳化，而非產品。本書程式碼只能當作教學輔助工具。在製作將會被商品化的專案時，建議使用既有的函式庫與框架，因為它們通常都已針對效能最佳化、經過良好測試並有完整的支援。

線上資源

本書程式原始碼：<http://mng.bz/Vgr0>

作者個人網站：<https://rhubans.com>



本章內容

- 已知關於人工智慧之定義
- 適用於人工智慧的基本觀念
- 於電腦科學與AI中的問題類型及屬性
- 簡介本書談到的AI演算法
- AI的實際應用

什麼是人工智慧？

所謂的智慧是一個謎，這個概念至今仍然沒有一個統一的定義。哲學家、心理學家、科學家與工程師對智慧是什麼，以及其出現的方式都有不同的看法。在周遭的自然環境中可以看到以人類的思維和行為方式來理解的智慧，例如一群分工合作的生物。一般來說，具備自主性與適應性的事物可視為具有智慧。自主性表示不需要持續的指令，而適應性則代表該事物會隨著環境或問題空間而改變行為。當我們在

觀察生物或機器時，會發現運作的核心要素是資料。所見所聞皆為資料，對於周圍一切事物之測量也全都是資料。我們使用、處理並根據這些資料做出決定，因此，對周邊資料的概念有基本的了解在理解人工智慧（AI）演算法上至關重要。

定義 AI

有些主張認為，我們根本不了解什麼是 AI，因為要定義智慧本身就很困難。薩爾瓦多·達利認為抱負為智慧的特徵之一，他曾說：「聰明卻缺乏抱負，如同失去雙翅的飛鳥。」愛因斯坦相信想像力是智慧的重要因素之一：「智慧的真正象徵不是知識，而是想像力。」而說過「智慧是適應轉變的能力」的霍金，重視的則是應變的能力。這三位偉人對智慧都有不同的見解。雖然還沒有關於智能的明確答案，但至少可以說，我們對智能的理解出自於人類作為優勢（且最聰明）物種的這個地位。

為了大家好，並忠於本書的應用實例，在此粗略把 AI 定義為能夠表現出「智能」行為的綜合系統。與其試圖定義某種事物是否為 AI，不如來談談它與 AI 的相似性。某些事物可能會表現出部分智能，因為它可以幫助我們解決難題並提供價值與效益。通常，模擬視覺、聽覺等自然感官的 AI 實作會被視為類 AI。能夠在適應新資料與環境的同時自主學習的解也是類 AI。

以下為具備 AI 性質的一些例子：

- 可以玩多種複雜遊戲的系統。
- 癌症腫瘤檢測系統。
- 基於少量輸入資訊便可生成藝術作品的系統。
- 自動駕駛汽車。

侯世達（Douglas Hofstadter）曾說：「所有未竟之事都是 AI。」在上述舉例中，自動駕駛汽車因為還沒有被完善，所以看起來非常先進聰明。這和不久前能夠計算加法的電腦被認為非常聰明，現在卻被視為稀鬆平常是一樣的概念。

追根究柢，*AI* 是一個模稜兩可的詞彙，對不同人、行業或學科來說都有著不同的意義。本書中的演算法從以前到現在都被歸類為 *AI* 演算法，是否能夠賦予 *AI* 一個明確的定義其實並不重要，重點是它們能夠用於解決難題。

認知到資料是 *AI* 演算法的核心

資料為執行近乎神蹟的美妙演算法提供了輸入資訊。如果資料選擇不當、代表性不足或缺失，演算法便無法正確運算，因此輸入資料的品質關乎到結果的好壞。這個世界充滿了各種資料，有些甚至以我們無法感知的形式存在。資料能夠以數值的方式呈現，像是北極當前的溫度、池塘中有幾條魚、或是以天數為單位的目前年齡。上述舉例皆涉及根據事實而獲取的準確數值，因此很難誤判。特定時間地點的溫度是絕對真實的，不會受到任何偏差的影響。這類資料被稱為**定量資料**。

資料也可以代表觀察值，例如花香或是某人對特定政治家政策的認同程度。這類資料被稱為**定性資料**，有時候不好解讀因為它不是絕對真實，而是某個人對真相的感知。圖 1.1 舉例了生活中常見的定量與定性資料。

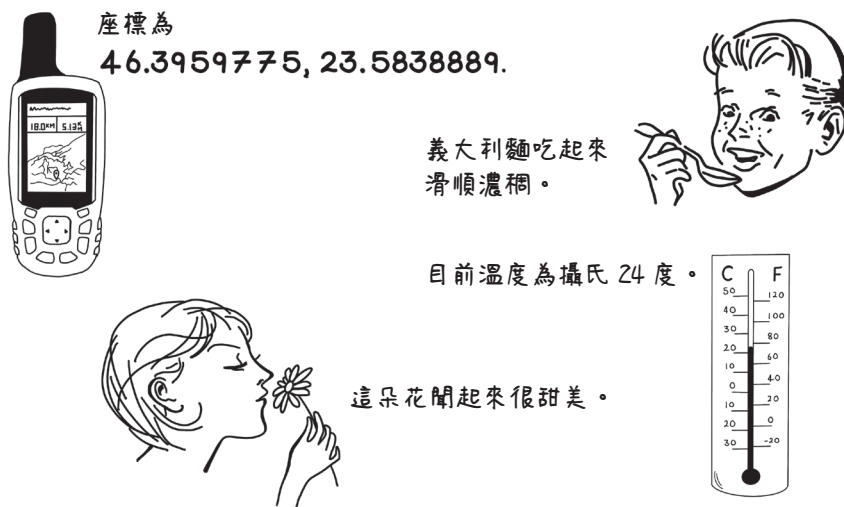


圖 1.1 生活中的資料

資料為事物的原始事實，因此記錄通常不存在偏見。然而事實上，資料是在特定脈絡與使用方式等前提下被人們蒐集、記錄與連結。基於資料而建立有意義的觀點以回答問題的行為稱為建立資訊。基於過往經驗來活用資訊，並且有意識地應用資訊則產生了知識。這便是我們嘗試要以 AI 演算法模擬出來的部分。

圖 1.2 說明了如何解讀定量與定性資料。時鐘、計算機或磅秤等標準化儀器通常用於測量定量資料，而氣味、聲音、味道、觸感與視覺，乃至我們的固執成見，通常用於建立定性資料。

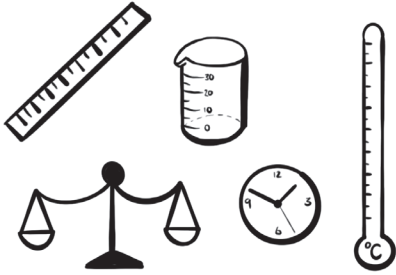
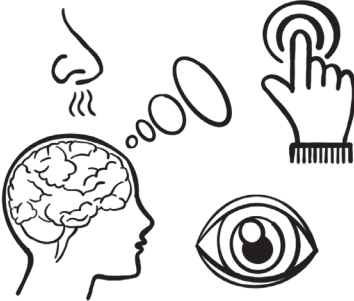


	定量	定性
儀器		
以卡布奇諾作為範例	 <ul style="list-style-type: none"> • 咖啡杯容量為 350ml • 溫度為 91°C • 總重量為 226 克 • 陶瓷杯 • 咖啡豆來自非洲 	 <ul style="list-style-type: none"> • 口感滑順濃稠 • 味道濃郁帶一點巧克力味 • 咖啡顏色金黃 • 杯子是白色的 • 香氣濃郁

圖 1.2 定量與定性資料對照表

不同的人會根據各自對該領域的理解和對世界的看法對資料、訊息和知識做出不同的解釋，這無庸置疑地會影響到解的品質——因此使得科學觀點在創造各種科技時變得額外重要。藉由遵循可重複的科學過程來獲取資料、實驗並精準報告，我們便可以確保在使用演算法處理資料時能得到更準確的結果與更優秀的解。

演算法如同食譜

現在我們對 AI 有了大概的定義，也明白資料的重要性了。由於本書將討論到數種 AI 演算法，因此清楚地理解演算法是什麼會很有幫助。演算法是指為了實現特定目標而提出的一連串指令與規則的設定。通常，演算法接收到輸入後，會在不同狀態下進行幾個有限的步驟以產生輸出。

即使是像閱讀這麼簡單的事情也可以用演算法來表示。以下是閱讀本書所涉及到的幾個步驟：

1. 找到「凡人也能懂的白話人工智慧演算法」一書。
2. 打開書本。
3. 仍有尚未閱讀的內容時：
 - a. 閱讀當前內容
 - b. 翻頁
 - c. 思考學到的事情
4. 思考如何將所學應用在現實世界中。

如以下圖 1.3 所示，演算法就像一份食譜。食材和用具為輸入，烹飪特定料理的步驟為指令，而最終的菜餚便是輸出。

口袋麵包的演算法

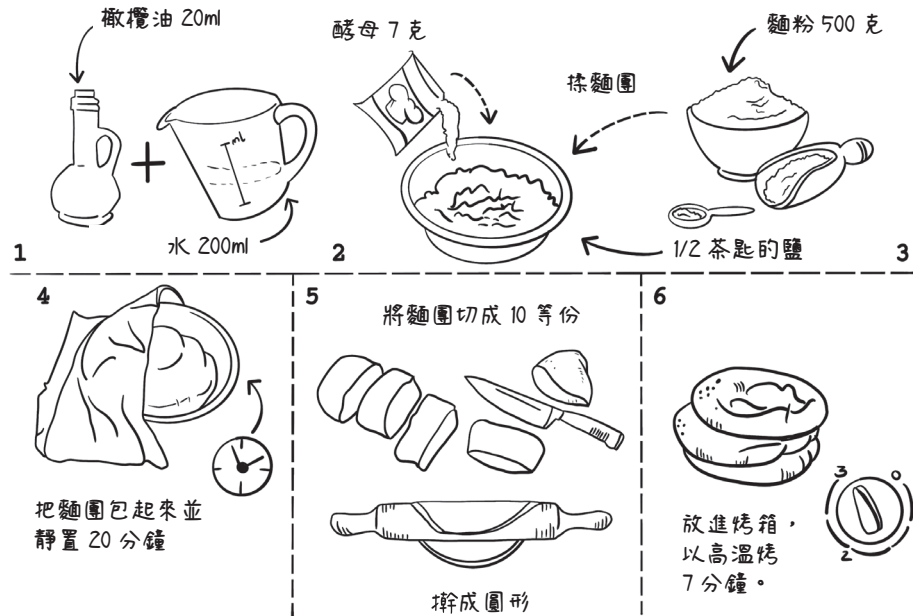


圖 1.3 演算法就像是食譜

演算法已被用於許多不同的解。例如，透過壓縮演算法便可以讓世界各地的人即時視訊聊天，藉由即時選路演算法讓我們可以透過地圖應用程式在城市中導航。即使只是一個簡單的「Hello World」程式，背後也涉及了許多不同的演算法，好將人類可讀的程式碼語言翻譯成機器可讀的程式碼以便在硬體上執行指令。只要仔細看，您可以在任何地方找到演算法的蹤跡。

圖 1.4 所描述的猜數字遊戲可以幫助您進一步瞭解本書將談及的演算法。電腦在指定範圍內隨機挑出一個數字，而玩家要試著猜出來。請注意這個演算法具有離散性的步驟，在進行下一個操作之前會先執行動作或做出決定。

基於我們對技術、資料、智能與演算法的理解：AI 演算法為一組指令，使用資料建立可表現出智能行為並解決難題的系統。

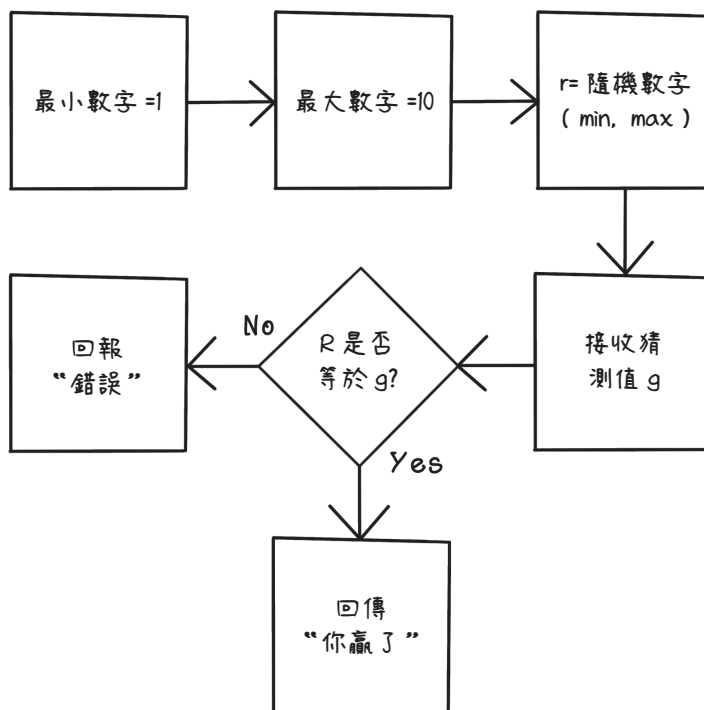


圖 1.4 猜數字遊戲演算法之流程

人工智慧簡史

回顧過去我們在 AI 中取得的進展，有助於了解如何結合舊技術和新想法以創造出新的解。AI 並非全新的概念，歷史中充滿了關於機器人和會自己「思考」的機器人的故事。回首過去，不難發現自己其實是站在巨人的肩膀上看世界。也許我們也都可以為人類共同的知識寶庫做出一點貢獻。

檢視過去的發展可以凸顯出理解 AI 基礎知識的重要性，幾十年前的演算法在現代 AI 實作中至關重要。本書將從有助於建立解決問題概念的基本演算法開始，然後慢慢進入更有趣且更現代的方法。

圖 1.5 並非所有的 AI 成果，這只是一小部分，還有許多其他的重大突破呢！

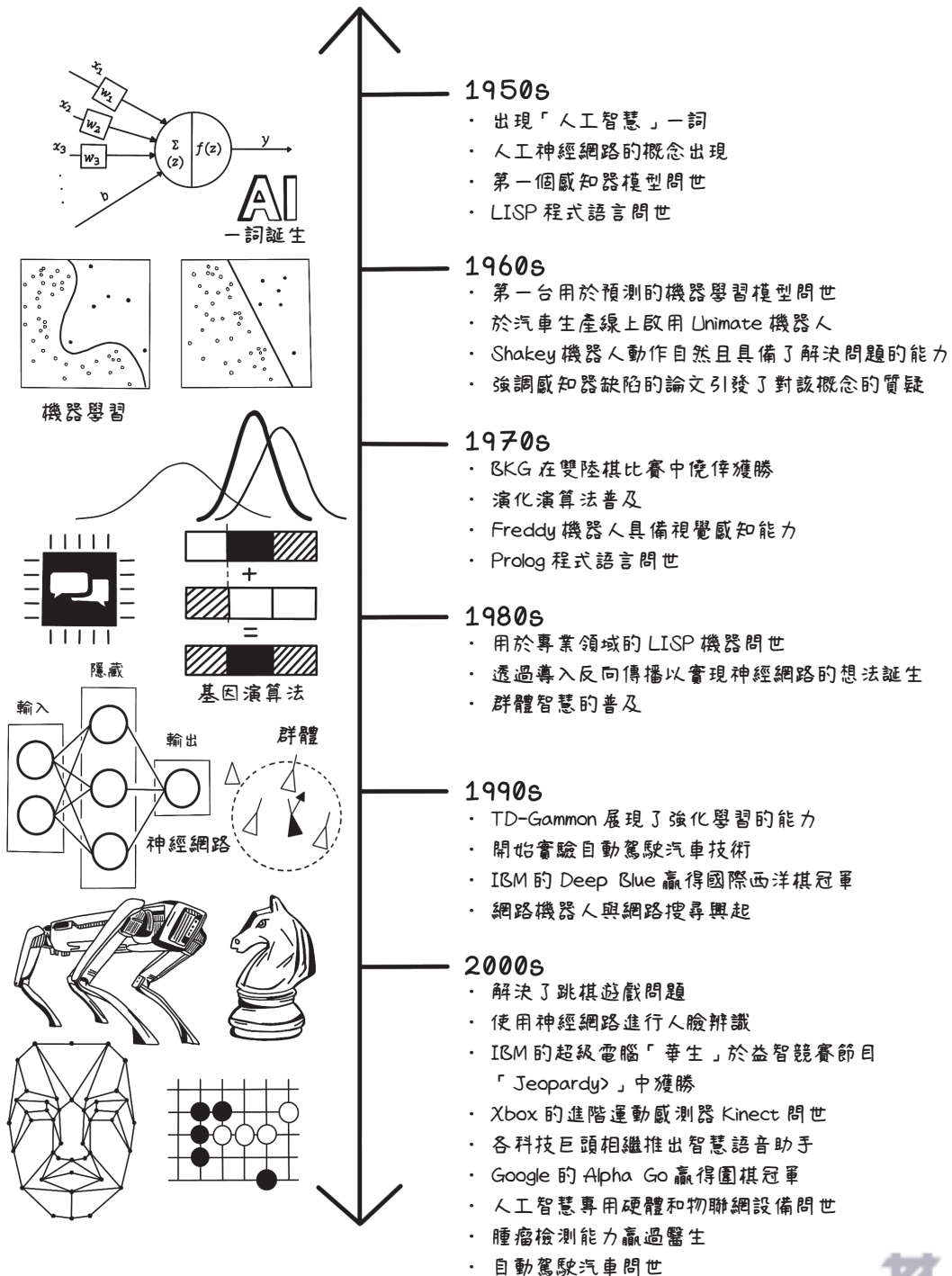


圖 1.5 AI 的演變