

## 人工智能與人類地位

上周本欄提到天主教教宗良十四的《通諭》，題目是 *Manifica Humanitas*（偉大的人類）”（發表於 5 月 25 日）。裏面提出要抗拒“技術主義霸權”（Technocratic Tyranny）。是一個非常嚴厲的批評。Tyranny 這個字，一般用來形容“暴政”，是不擇手段危害人類的殘酷暴君。《通諭》引起了世界上廣泛的反應。

《通諭》長 200 頁，一萬多字。這裏且冒昧節錄梵蒂岡新聞網的幾段話。教宗在《通諭》的發布會上疾呼，“切莫把人工智能當成「掌控、排斥或死亡的工具」”。他呼籲世人“不可用人工智能來作惡”。

《通諭》的發布會，教宗與高科技專家同席，其中包括美國 Anthropic 公司聯合創始人 Christopher Olah。Olah 不久前拒絕與美國政府作軍事合作。Olah 的在場，佐證了教宗的話：“在這人工智能時代（與科學家）攜手前行、聆聽、發聲，一起為世人找出道路。”《通諭》



■ 教宗良十四世接見 Anthropic 創始人 Christopher Olah.

說：“這些新事物今天觸及「我們生活的許多領域」，對各種決定產生影響，而且正在徹底改變戰爭進行的方式。”這種戰爭方式，最近以色列和美國利用科技，跨境集體殺戮另一個國家的領導人，就是一個明證。

教宗說，“《通諭》的許多素材、反省和指示都源自於「聆聽」。”教宗如此解釋：「我聆聽了許多科學家和工程師，他們誠摯熱情地從事科技工作，能藉以減輕巨大的傷痛；我聆聽了很多政治領袖和官員，他們努力不懈地尋求公正的規則；我聆聽了家長和老師，他們由衷關切年輕世代的未來。」

### 不可用人工智能來作惡

“同時，也有許多對於自動化武器系統的擔憂之聲傳到教宗那裡去。他也聽到很多人擔心「由於偏見和不義對數據基礎的染指」，算法可能會阻礙一些人得到「醫療協助、工作和安全」。除了這些聲音之外，「在做決策之際，無法發聲者的沉默」也令人擔憂。教宗提醒道，「決策可能會造成新的排斥和苦難形式」。

有鑒於此，教宗再三強調，“不可用人工智能來作惡”。教宗知道這句話說得很重，但他「刻意選擇這麼說，因為此時此刻需要一些能吸引注意力的話語來喚醒良心，為全人類指出當走的道路」。

…「關於科技的決定絕對是與良心和責任密不可分的」。他也談到裁軍和平：「和平不只是沒有戰爭，更是正義伸張。然而，當科技削弱我們的批判意識時，和平本身就會陷入危險。無論如何，光是解除武裝仍有所不足，我們還必須進行建設。」

教宗又說：「唯有懷著全面的視野，才能將人工智能導向公共福祉。唯有齊心協力，設計系統和受系統影響者、富裕國家和貧困國家、機構和個人、權力中心和偏遠地區同舟共濟，我們才能建設未來。這不是少數人獨享的前途，而是整個人類大家庭的將來。」

他說：這正是天主教宣揚的「愛的文明」。出於這個動機，教會「謙遜且坦率地」渴望參與有關人工智能的對話：「我們既不掌握技術性的答案，也不試圖取代專家。但我們攜帶有關人類的智慧，這是我們當下急切需要的：每個人都是獨一無二、無可取代的，是自由且有良心的智慧體，能尋求天主、為彼此服務，並且照料我們共同的家園。」

### 取人首級於千里之外

很多人以為，教宗就是“修行”而已。筆者看到過梵蒂岡教廷裏面的教士，充滿著各方面的自然科學家和社會科學家、哲學家。筆者在哈佛的時候，曾經在一所耶穌會教士聚居的一棟房子吃晚飯，裏面有 22 位耶穌會的教士在念博士，念什麼學科的都有；他們說：在波士頓地區，就有同樣的 15 棟。這只是波士頓，可見一斑。

筆者不是天主教徒，但是上面不厭其詳的的節錄，可以引起多方面的思考。

第一、軍事。提到不要利用科技作惡，也提到科技引起的戰爭新形式。這是輿論上提得很少的。從伊朗一下子喪失一批領導人，除了“跨境”、“殺戮”這些不可忍受的戰爭行為，也要提防核武器以外，新的侵略會是怎樣發生的。發證在伊朗、委內瑞拉的事情，“取人首級於千里之外”，可以發生在任何一个國家。

第二、國際。技術主義霸權，也是國際關係中的一種現象。有論者認為，大約兩百年前，霸權在海洋，船艦炮利就成了國際霸權的象徵。於是出現了荷蘭、英國這些世界性的殖民者。二戰以來，核武器，成了霸權的象徵。君不見

媒體常常在數落每個國家有多少枚核彈頭、可帶核彈的遠程導彈！現在又有了 AI 衍生的新武器。與此同時，國際之間，特別是強國之間，科技的競爭，已經變成了科技的抵制。因此有史丹福 HAI 今年三月提出的：各國應該有各自發展 AI 的獨立“主權”（sovereignty），“擺脫研發商和美國政府”。（見 3 月 六日本欄）

### 從技術主義的迷茫中醒覺

第三、政策。《通諭》在不少地方提到決策與選擇。反對技術主義霸權，馬上有 AI 發展商，就因為擔心政府於是出面管制 AI 的發展，而反對《通諭》的警告。這也許是許多國家的政府都沒有納入議程的項目。但是另一方面，AI 的發展的確會有點瘋狂，沒有邊界。既超出了人類生活的實際需求，也罔顧科技發展引起的社會問題。這裏面既有研發者的價值觀，但很明顯也有強烈的商業利益在驅動。人類社會需要的是什麼？只有少數的科學家（例如醫學創新）在考慮。例如 Fairplay 的呼籲（5 月 8 日本欄）、各國的禁制未成年人使用社交沒媒體（5 月 15 本欄），都是微弱得很的聲音或措施。AI 的發展，猶如一批脫韁的野馬，還會繼續狂奔。

第四、如上周本欄所述，技術掛帥，是全球性的現象。《通諭》提出的其實是一項全球性的警世宣言：從技術主義的迷茫中醒覺，重新認識人性的重要性，恢復人的尊嚴。這種宣言的提出，說明情況已經發展到極端危險的地步，接受這種警示的，就完全不應該停留在一個宗教裏面。雖然，可以想見，在 AI 的威力無限擴大的時候，人類要佔主導地位，需要一定的努力。

第五、在教育而言，教學“以學生為中心”，提出已久，在中國（甚至華人社會）大概沒有多少反對的聲音。在生成式人工智能（ChatGPT）出現以後，“賦

能學生”的提法更是甚囂塵上。賦能學生，也就是讓學生擁有學習的主動權，世界上流行的“讓學生成為主動的學習者”、“self-regulated learning”的提法，也非常普遍。這些都是善意的、積極的想法。具體來說，卻有很多不同的理解。下周再議。