

在 AI 時代談談新聞真實性跟事實查核的議題

文 / 李怡志 (政治大學新聞系助理教授)

新聞媒體的目的在於報導真相，但生成式人工智慧出現後，因為可以快速產製圖片、影片與文字，對目前營收已經相當困難的新聞業，產生了不同層面的影響。從2025年開始可以觀察到，台灣的商業電視台已經越來越習慣在新聞中使用生成式工具來產製示意影片與聲音，而且多採取逼真的風格、搭配反差很低而且字級很小的揭露，藉此讓不在場的媒體記者也能表現出在場的樣子。

最使我感到衝擊的例子，是中天電視在柯文哲被羈押的期間，建立了柯文哲的聲音模型，模擬柯文哲的聲音念公開信。我聽到時心中一驚，想說柯文哲應該還在看守所，何時放出來了？後來才在電視螢幕上反差很低、不容易辨識的文字中看到是由AI生成——如果我不是研究這個領域的研究者、沒有刻意去注意，而且又使用螢幕小的手機觀看時，應該就不會注意到AI生成的揭露，這樣我認知的真實就被AI內容再建構了。

「示意圖」與真實現場之間的距離

目前新聞媒體經常利用生成式圖像再現「示意圖」，但「示意圖」有多重意涵，這些意涵一般人可能比較少思考、或者無法體會與實際影像之間的差異。首先，示意圖應是基於記者目擊的結果，只是記者的目擊並沒有或無法取得影像紀錄（例如法庭新聞），但生成圖像時，記者可以將自己的記憶與生成內容相互比對，確保影像與記憶趨近，這種作法通常是最接近真實影像的作法。

但現在台灣媒體常見的作法是以下幾種：一、記者並沒有親臨現場，依靠目擊者的描述後用AI再現，但示意圖並沒有經過目擊者確認。二，記者不但沒有親臨現場，也找不到目擊者，只靠轉述或紀錄（警方筆錄或判決書）生成示意，如此一來記者已經與真實的現場相隔非常多層，也不可能有人能確認是否跟真實接近。如果整個媒體都沒有人看過現場、或請目擊者確認與真實是否吻合，就不適合使用逼真的模式並且將影像動態化，試圖讓閱聽人相信這些AI生成的影像與真實之間有指涉關係。

使用生成式圖像來指涉記者未曾目擊的現場有許多風險，特別是讓AI生成細節時，經常會忽略細節對閱聽人認知的影響。許多研究都發現，人類偏好膚色較淺而討厭膚色深的人（Colorism），而AI模型又容易將負面提示詞的人物繪製成膚色較深的人，所以新聞媒體有可能在沒有意圖的情況下，誤導了閱聽人對新聞當事人的感知。其他包含表情、服裝、體態、環境等細節，也都有類似的效果，本文就不細述。



為了避免閱聽人誤信示意影像為真，媒體應該讓閱聽人能夠清楚辨識內容並非真實。首先，應該要有高反差、顯著的標示，並且即便在手機上觀看都要能清晰可見。其次，如果記者並未在現場目擊，也無法查證AI生成圖像與真實之間的落差，應該要採用抽象的風格，並且降低彩度、減少細節。最後，如果媒體並沒有目擊現場，在生成圖像時應該誠實標記該影像生成是基於轉述或其他記錄，而不是只透過少量的文字資料，就放大成有高度細節並且擬真的示意動態影像。

生成式AI + 認知偏誤 = 錯假訊息傳播力

台灣人對參與政治比較冷感，但很喜歡看政治新聞，特別是每兩年一次的選舉往往成為各種虛假訊息傳播舞台之一；而有了生成式AI後，製作政治類的虛假訊息成本變得更低、傳播速度更快。我最近的觀察發現，不論任何陣營的支持者都已經開始製作競爭對手的影片，這些影片主要的目的是發表評論，但評論可能基於未完全查證的小道消息，甚至是蓄意製造的謠言。由於使用AI生成影像後十分逼真，反對者因為認知偏誤——不論教育程度如何——都有可能第一時間就轉傳，或甚至明知為假也蓄意轉傳。可以預料到年底的投票日之前，大量這樣的內容會讓候選人完全沒有回應或反擊的能力，並且還會進而再影響新聞媒體的報導內容。

網路虛假內容被媒體跟進的案例越來越多，例如民視新聞在

2025年就曾在新聞中誤用AI生成的緬甸地震新聞、2026年又在京都殺人案中，誤信網路傳言誤植嫌疑人國籍。原本應該報導真實的新聞媒體，不應該變成AI加上演算法的放大器。

虛假訊息批量生產的年代，媒體需要進階守門力

在生成式AI的年代，媒體更應該減少報導瑣碎、單一的現象，多關注、報導現象背後的可能路徑與影響；同時，也要避免搶快——減緩報導速度，不要在一開始跟進，畢竟跟進得愈快就愈沒有時間查證，虛假訊息就愈多。最後，媒體也需要嚴守基本查證的能力，確定網路訊息背後的來源，來源不明的不要跟。現在處於收入與時間壓力下的媒體記者，只要看到社群平台如Threads或IG上有什麼貼文，即便看起來可能是才剛開的帳號，也會跟進，而忽略背後貼文者的身分有可能是商業操作或政治操弄。媒體應該要對此建立基本的規範，如果記者無法查證貼文者的真實身分，就要更多的交互查證才能引用。

在這個虛假資訊可以被AI自動化批量生產的年代，新聞記者如果還要做好這一行，只有傳統的採訪、寫稿、拍攝的能力已經完全不夠，還需要花更多時間理解社群的原理、演算法、廣告機制等等，並在這個基礎上建構自己的資料收集網絡。最好在不同的社交媒體平台上，都要有來自不同立場、且消息可信的來源，避免追著單一面向的訊息跑；甚至可以開設專屬工作用的帳號，

避免演算法被個人原本的立場或態度影響。

這幾年個人觀察，真正關心新聞品質、正確性的閱聽人不如一般學者樂觀，許多閱聽人並不在意新聞的真偽。因此，希望看到真實內容的讀者應該更有意識地用金錢支持媒體，畢竟媒體如果沒有好的收入，就只能降低成本、更依賴廣告與演算法。閱聽人如果都在免費的平台看免費的內容，只會讓有其他動機的單位趁虛而入，以生成式AI灌入越來越多虛假內容。

