

文藻外語大學使用生成式 AI 參考指引

一、目的

隨著生成式 AI 技術的普及和發展，文藻外語大學(簡稱本校)教職員工生已開始廣泛將 AI 技術融入課程，用於學習和創作。生成式 AI 快速發展，對全球產官學研各界均產生深遠影響。

生成式 AI 模型旨在生成類似人類製作的新內容，例如大型語言模型，能夠根據目的和條件自動學習並產生多樣化內容。然而，其大量蒐集、學習與產出之資料可能涉及智慧財產權、人權或業務機密的風險。此外，由於生成結果受限於學習資料之品質與數量，可能出現真偽難辨或創造虛假資訊的情況。因此，對於生成式 AI 產出的資訊與潛在風險需進行客觀且專業的評估。

本指引旨在確保本校師生在使用生成式 AI 技術時，能正確且有效地生成資料，避免潛在風險和問題。並期望師生以負責的態度使用生成式 AI，促進其在教育領域的長期發展，充分發揮其教育價值。

師生在使用生成式 AI 工具時，應秉持以下原則：

1. 負責任與信任：以負責任與可信賴之態度操作，掌握自主權與控制權，避免衛生成資料全然信任。
2. 安全性與隱私：不揭露未公開之公務資訊，或分享個人隱私資訊。
3. 資料治理與問責：遵循資料安全、隱私保護及問責原則，並在教育領域中創造其最大價值。

二、生成式 AI 的風險

生成 AI 雖具潛力，但也伴隨多方面風險，主要以下幾類：

1. 錯誤信息與誤導

- **不正確的內容**：生成未經驗證或錯誤的資訊，可能導致誤導。
- **偏見**：若訓練數據存在文化、性別或種族偏見，生成內容可能放大偏見，對學生形成不公平或有害的觀點。
 - 範例：AI 生成的歷史資料片面描述特定群體的貢獻或經歷。
- **缺乏學術嚴謹性**：內容可能缺乏參考、數據支持或邏輯性，影響教育質量。

- 範例：學生使用 AI 工具生成科學報告，結果未經過驗證，導致誤導性成果。

2. 版權與侵權問題

- 內容抄襲：可能無意中複製訓練數據，侵犯版權。
- 藝術與創作爭議：生成內容是否為原創，或版權歸屬問題引發法律和道德爭議。
 - 範例：AI 生成的圖片與現有藝術作品過於相似，可能引起侵權訴訟。

3. 誤用與不當應用

- 虛假信息：生成技術可能被用於製作深偽視頻或虛假新聞，引發公眾恐慌。
- 詐騙：模擬他人語音或文本，用於社會工程攻擊或詐騙。
 - 範例：使用 AI 模擬名人語音進行詐騙電話。

4. 道德與社會影響

- 影響創作者生計：AI 可能取代人類創作者，威脅某些行業的就業。
- 過度依賴：削弱自主思考與創造的機會。
 - 範例：新聞機構大量使用 AI 撰寫文章，導致記者減少。

5. 安全與隱私問題

- 數據洩漏：生成內容可能洩露敏感數據。
- 攻擊性內容：內容可能包含暴力、仇恨或其他不適合公開的元素。
 - 範例：生成個人敏感信息，如地址或銀行帳號。

6. 學習過程的削弱

- 依賴性過強：學生可能忽視自主思考與研究能力培養。
- 降低批判性思維能力：直接提供結論而未引導學生分析或多角度思考。
 - 範例：學生用 AI 完成作業，但對作業涉及的核心概念缺乏真正的理解。

7. 抄襲與學術誠信問題

- 學術不端：學生可能利用 AI 工具完成作業、論文或考試內容，導致抄襲和誠信問題。
- 難以檢測抄襲：AI 生成內容形式原創，但並非學生自己的想法或創作，難以通過傳統抄襲檢測工具辨別。
 - 範例：學生使用 AI 撰寫論文，教師難以判斷是否為學生真實能力的表現。

8. 隱私與數據安全

- 「學生數據風險」：AI 工具可能收集學生的學習數據或個人信息，存在數據洩露或濫用的風險。在未經適當過濾的情況下，AI 可能生成不適合教育環境的內容。學生在使用 AI 工具時，意外提供了個人敏感信息，導致隱私洩露。

三、生成式 AI 管理規範

1. 本校教職員生使用 AI 進行生成之資訊，應就其風險進行客觀且專業之最終判斷，不得取代人之自主思維、創造力。製作機密文書應由業務承辦人親自撰寫，禁止使用生成式 AI，包括本校法令規章、研究計畫書、中長程發展計畫等。
2. 承辦人不得向生成式 AI 提供涉及本校教職員生等公務應保密、個人及未經本校同意公開之資訊，亦不得向生成式 AI 詢問可能涉及本校機密業務或個人資料之問題。但封閉式地端部署之生成式 AI 模型，在確認系統環境安全性後，得依文書或資訊機密等級分級使用。
3. 使用者應確保作品原創，避免抄襲。將 AI 生成的內容作為參考，進行修改和創新以體現個人想法。遵守著作權規定，正確引用來源，包括 AI 工具，並在學術創作中使用標準格式避免誤導讀者。
4. 使用 AI 訓練數據時，應確保來源合法且符合道德規範，尊重數據所有者的權利，並在獲得授權後使用。避免使用涉及隱私或可能造成不良影響的數據，並對 AI 內容的用途和應用負責，必要時取得原創者同意。
5. 使用 AI 生成內容時，應尊重不同文化與社會群體的需求，避免冒犯、歧視或引起爭議的內容。鼓勵包容與對話，尊重多元觀點。利用 AI 支持不同背景和需求的學習者，提供公平的教育機會，促進文化交流和教育多樣性，打造包容的學習環境。

6. 生成式 AI 應作為輔助工具，而非完全依賴的學習或創作方式。使用者需了解 AI 的優勢與限制，在適當時機合理使用，同時保持主動性與創造力，提升自身能力。對 AI 的輸出應批判性對待，避免盲目依賴，並對其內容進行審核和修改。在使用過程中，需理解 AI 的運作原理，保持獨立判斷，並在必要時尋求他人建議以確保內容的準確性與可靠性。