

關於中文

A character could not be read from the file stream using the provided code page. The reason for this could be that a column break cuts a multi-byte character into two pieces.

代表 SSIS 在讀取檔案時，**檔案的編碼 (Code Page)** 設定與實際檔案內容不符，或是分隔符號將多位元組 (如 UTF-8 漢字) 切成兩半，導致資料無法正確解析。這常見於處理含有特殊字元或多語言 (如中文、日文、韓文等) 的檔案。

修復步驟

1. 確認檔案實際編碼

- 用 Notepad++、VS Code 或其他編輯器開啟檔案，檢查實際編碼 (如 UTF-8、ANSI、Big5、GB2312 等)。
- 若檔案是 UTF-8，請確認有無 **BOM (Byte Order Mark)**，部分舊版 SSIS 需 BOM 才能正確辨識 UTF-8。

2. 正確設定 Flat File Connection Manager 的 Code Page

- 在 SSIS 的 Flat File Connection Manager 屬性中，將「Code Page」選擇為正確的編碼：
 - **65001**：代表 UTF-8
 - **950**：繁體中文 Big5
 - **936**：簡體中文 GB2312
 - **1252**：西歐語系 (Latin1)
- 若檔案為 Unicode (UTF-16)，請選「Unicode」而非「UTF-8」。

3. 避免欄位分隔符切割多位元組字元

- 確認 Flat File Connection Manager 的欄位分隔符設定正確，**不要讓分隔符出現在多位元組字元中間**。
- 若是固定長度 (Fixed Width) 檔案，請確保每個欄位的長度設定正確，否則容易將一個中文字切成兩半。

4. 欄位型態設定為 Unicode

- 在 Flat File Connection Manager 的「Advanced」頁籤，將文字欄位的 DataType 設為 **Unicode string [DT_WSTR]**，以避免因型態不符導致的轉碼錯誤。

5. 嘗試將整行先匯入一個大欄位再做拆解 (進階)

- 若仍有問題，可將每一行先匯入一個大欄位 (如 DT_WSTR(4000))，再用 SQL 或 SSIS Script Component 拆解。

常見補救方法

- **重新存檔**：用支援 UTF-8 with BOM 的編輯器重新存檔。
- **避免特殊字元**：檢查檔案是否有不完整的多位元組字元，或在分隔符前後有亂碼。
- **Script Component 處理**：若遇到難以解決的編碼問題，可用 Script Component 進行自訂解碼或清理。

小結

1. 確認檔案實際編碼，並在 SSIS 設定相同 Code Page。
2. 確保分隔符不會切割多位元組字元。
3. 將欄位型態設為 **Unicode (DT_WSTR)**。
4. 必要時整行匯入再拆解。

這樣可避免「A character could not be read from the file stream using the provided code page」這類編碼錯誤。