本文摘自Ficke, S. H.的〈[From Text to Tags: The Digital Humanities in an Introductory Literature Course](http://www.jstor.org/stable/44378548)〉，由余欣薏統整後撰文。

數位人文在大學課堂內能扮演什麼角色？這個問題不易回答，因為數位人文所涵蓋的範圍非常廣泛，故這篇文章以「文學研究入門」（Introduction to Literary Study）課程為例，介紹數位人文工具如何應用於文學作品的數位化與文本的閱讀及分析，透過實作讓學生感受到文本數位化過程與工具使用均會涉及到如何詮釋文本（an act of interpretation），進而影響文本的解讀。

多虧Google圖書（Google Books）與古騰堡計畫（Gutenberg），學生能於這些網站搜尋出版已久的書籍，本課程從數位化開始，學生學習如何合作將較舊且損壞程度較高的實體書轉化為數位文本（digitization），其所建置的數位文本與媒體資料會在坦雅．豪（Tonya Howe）博士所開發的Omeka網站上展示。

首先，學生分組後指定閱讀三本電子書，挑選重點主題與議題，決定所要輸入的後設資料（metadata）和書本的重要資訊，作為呈現於Omeka網站上的主要內容，此實作讓學生學到數位文本的呈現會受到所選定的內容與技術的運用而有所影響。除了文字內容，學生也嘗試運用光學文字辨識（OCR）技術把圖片內容轉換為文字檔，然而轉檔後錯誤仍需靠人工修正。

學生將段落轉換為機器可讀的文本後，接下來便是讓學生使用Voyant Tools中的文字雲工具之一—Cirrus進行分析。此工具允許使用者自行編輯停用詞列表（stop list），學生能透過編輯字詞，建立一個對分析更有幫助的文字雲。過程中，學生先觀察無編輯停用詞的文字雲，確認是否產出對分析有用的結果後，再次觀察停用詞編輯後的結果。有些學生確實透過文字雲發現先前所忽略的主題，或發現設想的主題是由不期待的詞組成。

下一步則練習使用Voyant Tools的其他工具，第一個是上下文工具（Context Tool），學生將文字雲與此工具結合，觀察所出現字詞的前後文。第二個是字詞趨勢圖（Words Trends Graph），此工具顯示各字詞於文本中出現的頻率，這有助於學生透過文本關鍵詞觀察一個主題或議題的變化。藉由Voyant Tools所提供的這三種文本分析工具即能讓學生感受到電腦如何讓擴展讀者與文本的互動方式，同時訓練學生的批判思考能力，洞察之前未能看到或被忽略的其他主題或不同議題。

最後，學生使用可延伸標記式語言（XML）對文本進行標記。XML標籤可用來標示文本格式和內容，儘管能自創標籤，但許多學者仍建議使用文本邊碼格式標準（TEI）進行標註，TEI為一套不斷發展的標籤標準。讓學生熟悉XML和標籤概念及TEI標籤的清單後，便能進行文本標記。文中提到學生常在標記過程中面臨到如何針對某個字或某句話進行標記，思考這樣的問題其實就是一種學習或練習詮釋文本的過程，雖然學生認知標記相對耗時，不過也於理解標記過的數位文本中獲得有趣的進展，例如有位學生將<制度>標籤應用於教堂，從而發現也可應用於奴隸一詞，這是他先前從未思考過的。學生能透過標記將文本結構化，並根據研究主題，在標註過程中塑造自身和其他人對文本不同的理解與詮釋。

透過文本數位化的過程讓大學部學生體驗數位化技術是建立於基本閱讀和詮釋的能力，文中建議使用短文本，如﹕短篇故事或詩歌，讓學生進行深度閱讀，孰悉文本之後再應用數位工具進行文本分析的實作，使其體驗數位工具能為文本分析帶來不同層面的閱讀，藉由分析結果創造新的知識。