

數位人文研究

探索新現象，
建構新文化與認知



Gale Digital Scholar Lab

解決現代數位 研究者的困擾

不需 疲於學習與
了解大數據分析

不需 同時具備人
文與資訊技術

不需 耗時產出可
分析的文本數據

不需 管理與儲存
格式化機讀內容



當代數位人文研究者的最佳解決方案

數位學術研究是學術界正在迅速興起的創新研究方法，Gale Digital Scholar Lab 消除了進入數位學術研究領域的主要障礙：檢索、創建和組織數據集，研究者得以發掘未被發現的數據、測試理論、分析結果並獲得新的認知。



資料探勘技術 搭配易用工具
初學 - 高級研究員均適用



Gale 超過 60 年的資源投入
提供海量原始文獻 (OCR Text)

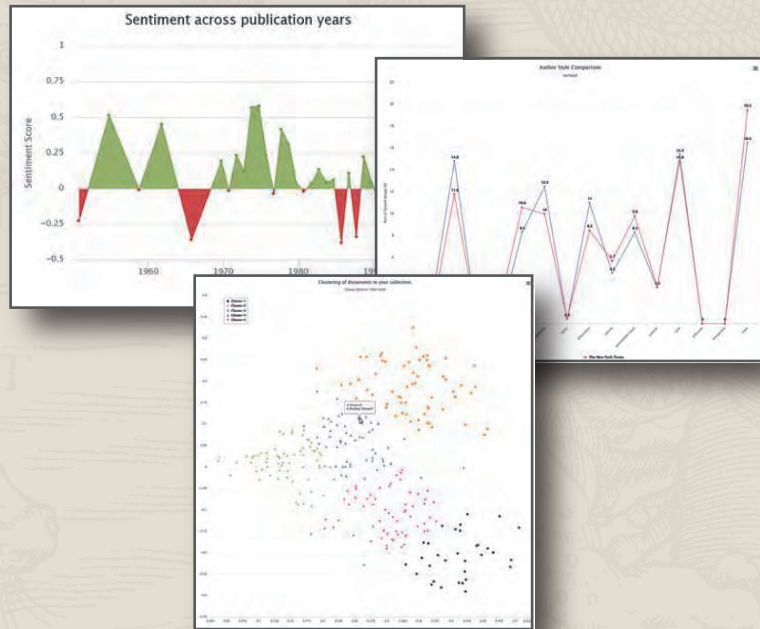


雲端研究環境
資料儲存空間不是問題

6 大分析工具 + 8 種原始文獻

分析工具(可客製設定值)

1. Topic Modelling 主題建模
2. Ngrams n 元語法
3. Clustering 集群分析
4. Parts-of-Speech Tagger 詞性標記
5. Sentiment Analysis 情感分析
6. Named Entity Recognition
命名實體辨識



原始文獻(OCR Text)

* 主題涵蓋：商業、政治、經濟、文學、藝術、電影、戲劇、文化、法律、社會、歷史、區域研究

- | | |
|---|---|
| ① 十八世紀經典古籍 (ECCO) | ⑤ 泰晤士報文學增刊 (Times Literary Supplement Centenary Archive) |
| ② 泰晤士報 (The Times Digital Archive) | ⑥ 現代法律之路：法學專論 (Legal Treatises) |
| ③ 近代經濟學 (Making of Modern World) | ⑦ 現代法律之路：審判 (Trials) |
| ④ 經濟學人 (The Economist Historical Archive) | ⑧ 現代法律之路：美國最高法院記錄與法庭檔 (U.S. Supreme Court Records and Briefs) |

★ 單位如欲加購 Gale 其他原始文獻，另有優惠 ★

應用實例



John O' Brien
University of Virginia

18 世紀很多詩歌都「隱藏」在報紙和雜誌中，研究者利用了伯尼典藏的 OCR 文本，檢測出報紙中的詩歌，並將其提取出來，這將讓詩歌語料庫數量成長至少 **3 倍** 以上。

Gale Source:
17th and 18th Century Burney Newspapers Collection (17-18 世紀伯尼報紙典藏)



Kat Gupta
University of Nottingham

Gupta 博士使用了來自《泰晤士報》數位典藏的 OCR 文本。她利用來自「語料庫語言學」領域的技術，透過分析報紙描述了解 1908-1914 年英國女性取得投票權的運動。

Gale Source:
The Times Digital Archive (泰晤士報)

