

# 人物與機構之知識圖譜分析工具發展與數位人文應用

陳志銘<sup>1</sup>、林俊佑<sup>2</sup>、張鐘<sup>3</sup>

<sup>1,2</sup> 國立政治大學圖書資訊與檔案學研究所

<sup>3</sup> 國立政治大學華人文化主體性研究中心

## 摘要

本研究旨在開發支援數位人文之「人物與機構之知識圖譜分析工具」，以輔助人文學者透過知識圖譜功能，更有成效與效率的探索人物傳記文本。此一工具不僅能輔助人文學者更清楚掌握人物與人物之間的關聯，更能透過人物與機構關係看到潛在於文本中的人物關係。實驗結果發現本研究開發之系統能有效提高受測對象對於文本探索的成效及探索效率，並且受測對象對於本系統的認知有用性及易用性呈現正面肯定態度。此外，從滯後序列分析與訪談資料顯示，透過系統提供的個人知識圖譜能準確地看到人物與人物，以及人物與機構之間的關係，並藉由一套有系統性的行為組合進行人物脈絡探索，以完成實驗所賦予的人物脈絡探索任務。

**關鍵詞：**數位人文、人物與機構關係、知識圖譜

## 壹、緒論

人文研究者過去的研究方式多半是以閱讀傳統文本方式進行，然而單純依靠人力進行紙本閱讀後再進行歸納的研究模式，在面對現今資訊爆炸的時代，已顯得不夠具有效率，也難以進行超過人工閱讀負荷量之基於大數據文本之探查研究。因此，人文研究者開始思考要如何運用資訊科技輔以人文研究，以探索傳統人文研究中所無法研究的現象與問題(Schreibman, 2008)。近幾年以來，數位人文的發展主要係受到數位典藏的影響，透過數位典藏資料庫輔助進行人文研究，已逐漸成為人文學者的重要研究模式(Hockey, 2004)。此外，在數位典藏內容的基礎上發展數位人文工具，亦扮演數位人文發展上的推波助瀾重要角色。數位人文工具的發展目前主要可以區分為文字與圖像兩個領域，其中探查文本內命名實體之間的關係，為基於文本分析之數位人文領域的重要研究議題。然而，大部分命名實體關係的研究，大都聚焦於探查人物彼此之間的實體關係，鮮少探討人物與機構之間的關係，而人物與機構往往有著密不可分的關係。從機構觀點所進行的命名實體關係文本脈絡探索，有助於人文學者以更豐富的探查面向從中獲得隱藏於人物與機構之間的潛在人物關係。

2012年 Google 釋出了語意網知識圖譜的技術，對於語言學和社會科學而言，具有極高的價值與意義，可針對知識進行具架構與視覺化的表述(Al-Khatib, Hou, & Wachsmuth, 2020)，以支援使用者在數位資源上進行更有效的探索，並且具有知識表達並視覺化的特性，更有利於知識的引導與吸收(Haslhofer, Isaac, & Simon, 2018)。知識圖譜由節點和邊所組成，節點為文本中的實體，例如人名、地

名和機構名；邊則是代表實體的屬性或實體之間的關係，例如朋友、夫妻、工作等關係。換句話說，知識圖譜能將繁複的資訊簡潔有效的彙整於一張圖中，並能以同時呈現實體與實體之間關係的宏觀遠讀視角給使用者進行關係脈絡解讀。

綜合以上所述，本研究基於《馬來西亞興化群賢錄》發展「人物與機構關係之知識圖譜分析工具」，使其能透過機器學習之自動名稱實體及關係識別技術，將人物與人物，以及人物與機構關係以知識圖譜的遠讀視角方式呈現，輔助人文學者進行此一傳記類文本之人物與機構關係脈絡解讀，亦提供閱讀文本介面，達到遠讀與近讀的交互參照功能，讓人文學者除了能探索人物與機構的關係外，亦能透過機構視角找到單看個人傳記所無法完整解讀之潛在人物關係。為了驗證此一工具支援數位人文研究的效益，本研究以人物與機構關係之人物脈絡探索成效、使用者行為歷程，以及科技接受度三個面向進行評估，並透過半結構式深度訪談輔助，以瞭解人文學者使用此一工具支援人物與機構關係脈絡解讀之數位人文研究的成效、感受與建議，並據此結果提出未來持續改善工具之具體建議。

## 貳、文獻探討

### 一、數位人文

數位人文(Digital Humanities)為數位技術與人文科學整合的學科，旨在探討利用數位資料庫和數位工具輔以人文領域研究，所產生的助益與影響。傳統的人文研究係透過閱讀大量紙本資料方式，以及人工自行歸納與分析方式進行，例如透過人工標註來源文本，或者單純於段落旁撰寫筆記，以輔助分析文本中的主題脈絡，此一過程相當耗費時間與人力。然而，隨著數位資料庫與數位人文工具的發展，已逐漸改變了人文學者對於資料的探查與分析方式，也可以從事過去人工閱讀分析，所無法探查的研究議題。Schreibman (2008)亦指出使用數位工具輔以進行數位人文研究，已成為重要的人文學研究方法與趨勢(Kirschenbaum, 2012)。

在數位人文研究領域中，近讀(close reading)與遠讀(distant reading)為文本分析的兩個主要面向，「近讀」為透過閱讀文本後，對文本進行的自我註解與歸納歷程(Fisher & Frey, 2012)；「遠讀」則為不仰賴人工閱讀的方式解讀資料，而是透過數位工具將文本內容進行摘要或抽象式概念的提取，並搭配資訊視覺化進行資料的解構與重組歷程(Jänicke et al., 2015)。數位人文工具的發展一直以來受到人文學界高度的關注(Bradley, 2003)，我國自2002年提出「數位典藏國家型科技計畫」起至今，已累積了大

量的數位典藏資料，然而透過數位典藏從事數位人文研究所面臨的問題並不是資料量的缺乏，而是在於面對如此龐大的數位化資料時，要如何進行有效且有效率的分析和應用，使其產生意義與價值(Rosenzweig, 2003)。因此，近年來數位人文研究工具或平台的發展百花齊放，Bingenheimer, Hung 與 Wiles(2011)等人利用地理資訊系統，將《高僧傳》傳中的內容進行呈現，建構出完整的時空平台，提供研究者以不同的角度來詮釋文本中的內容；Chen, Chang 與 Chen(2020)則基於「中國歷代人物傳記資料庫(CBDB)」發展史料人物關係圖工具，輔助人文學者解讀人物關係社會網絡。此外，亦有許多數位人文研究平台被發展出來，例如，國家圖書館與政治大學圖書館共同開發的「通用型古籍數位人文研究平台」(Chen & Chang, 2019)、國立臺灣大學圖書館的「台灣歷史數位圖書館」(Chen, Hsiang, Tu, & Wu, 2007)，以及法鼓山的「CBETA 數位研究平台」(Tu, Hung, & Lin, 2012)等。

綜合以上所述，近幾年來人文領域研究已慢慢演變成以數位人文工具來輔助人文學者進行研究，然而數位人文工具扮演的角色應為輔助，而非取代人文學者(Moretti, 2016)。本研究發展之「人物與機構關係之知識圖譜分析工具」，即為一個可以輔助人文學者進行傳記類文本之人物與人物，以及人物與機構關係解讀的數位人文工具。

## 二、知識圖譜

隨著文字探勘技術的逐漸成熟，將文字之間的關係轉換為實體間連結關係的知識表述需求，已越來越受到重視。知識圖譜(knowledge graph)被視為是實體與概念之間關係的呈現橋樑，近年來已被廣泛應用在各個領域，舉凡語義搜尋、智慧問答、自然語言，甚至物聯網皆能找到知識圖譜的應用(Pujara, Miao, Getoor, & Cohen, 2013)。知識圖譜是由語意網路(sematic web)、知識表示(knowledge representation)、本體論(ontology)以及自然語言處理(natural language processing)等技術發展而來，為人工智慧領域的重要應用。更具體的來說，知識圖譜是一種以節點和邊所組成的圖模型，節點為實體，例如人名、地名和機構名等，邊則可以為實體之間的關係或屬性，例如夫妻、同儕等，形成一個三元組的知識表述方法，用以敘述知識概念的連結關係(Singhal, 2012)。Tim Berners-Lee、James 與 Ora(2001)於1989年即提出建置以連結為中心的資訊系統(linked information system)，他認為相較於以樹狀結構為基礎的資訊組織方式，以節點及其連結的知識圖譜方式呈現資訊，會更適合現今網際網路這種開放式的資訊系統。

知識圖譜近年來也活躍於數位人文領域，特別是2020年數位典藏與數位人文國際研討會，便將知識圖譜列入為會議的探討主題，倡導如何利用語意網與人工智慧技術，發展更切合輔助人文學者的數位人文研究工具，知識圖譜具有動態連結資訊的特性，其結構化知識表述可讓資訊變得更為好用。在人物關係領域的研究中，知識圖譜應用亦是相當重要的發展方向，然而知識圖譜的建立相當耗費人力，因此預測實體間未知的關係或屬性，已成為了目前

讓知識圖譜的建立更快速的有效方法(Shen, Ding, Li, & Yang, 2020)。He, Yun 與 Lin(2019)提出了一種結合知識圖譜與深度學習的人物關係探勘方法，透過 DNN 與 BiGRU 神經網絡來識別人物命名實體，並提取人物之間的關係，再將人物關係的三元組數據儲存於 Neo4j 的圖形數據庫中，輔助使用者探查人物之間的關係脈絡。

綜合以上所述，知識圖譜應用於人物關係的研究已相當盛行，然而卻鮮少透過人物與機構關係的面向來尋找潛在人物關係。因此，本研究發展之「人物與機構關係之知識圖譜分析工具」，除了建構人物與人物之間的關係外，亦以人物與機構等實體關係建立知識圖譜，希望有助於人文學者透過人物與機構兩個面向，解讀傳記類文本人物關係，以不同的視角來進行文本脈絡探索，進而對傳記類文本有更深一層的理解。

## 參、系統設計

### 一、系統架構

本研究所開發的輔助傳記類文本脈絡解讀之「人物與機構關係之知識圖譜分析工具」係以《馬來西亞興化群賢錄》為分析之主要文本。此一工具具有兩個主要功能，其一為透過命名實體識別(named entity recognition, NER)技術將文本中的實體抽取出來，再將人物實體彼此之間的關係進行自動識別，目前可識別的人物關係包括親戚、夫妻、工作等八個關係；其二則可針對傳記中個別人物進行與之相關的機構名稱提取，將同屬於該機構的人物、對於機構的貢獻和事蹟列出，並且會連結原始文本以達到近讀的目的，建置此一功能的主要目的係希望能透過文本中的機構面向，輔助人文學者進一步掌握單看個別人物傳記所無法找到的潛在人物關係。本研究發展之「人物與機構關係之知識圖譜分析工具」將上述兩個功能整合起來，並以視覺化知識圖譜方式呈現，以方便人文學者輔以進行傳記類文本關係脈絡解讀，系統架構如圖1所示。以下針對系統架構及運作流程進行說明：

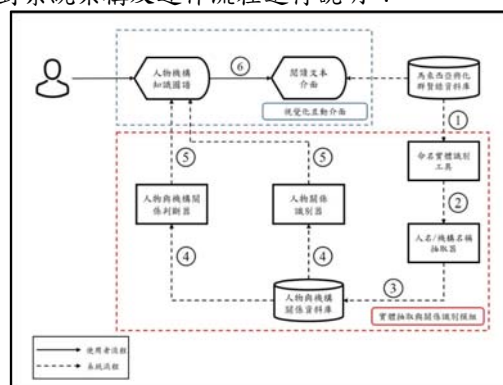


圖 1. 人物與機構關係之知識圖譜分析工具系統架構及運作流程圖

1. 將馬來西亞興化人物誌中的掃描檔及全文文檔以章節為單位儲存於「馬來西亞興化群賢錄資料庫」，除了提供資料給命名實體識別模組外，亦將資料傳給閱讀文本介面進行掃描

- 檔及全文文檔之顯示。
- 命名實體識別模組將文本的實體，舉凡人名、地名、官名和機構名等實體名稱識別出來。
  - 透過人名/機構名稱抽取器將人名及機構名進行提取，並且將其存入「人物與機構關係資料庫」中。
  - 從「人物與機構關係資料庫」提取資料進行人物與人物，以及人物與機構的關係識別。
  - 將「人物關係識別器」及「人物與機構關係判斷器」的結果彙整成「人物與機構知識圖譜」，以供人文學者進行人物與機構關係探查。
  - 人文學者從「人物與機構知識圖譜」，以及閱讀文本介面輔以進行遠讀與近讀結合的傳記類文本交互解讀。

## 二、系統元件

本研究發展之「人物與機構關係之知識圖譜分析工具」所採用的各項系統元件，說明如下：

### (一) 馬來西亞興化群賢錄資料庫

《馬來西亞興化群賢錄》係由「馬來亞大學馬來西亞華人研究中心」與「馬來西亞興安總會」合作進行之馬來西亞興化人物研究計畫成果，其記錄85位馬來西亞興化籍具有代表性的華人人物誌，這些人物來自於財政、經濟、軍事和文教等不同領域，反映了馬來西亞華人在當代史的各個時段發展脈絡。而「馬來西亞興化群賢錄資料庫」則是國立政治大學華人文化主體性研究中心參照《馬來西亞興化群賢錄》一書中條目的編排方式，將書中每一位人物分別獨立成一個條目，並以都柏林核心集作為主要後設資料欄位，所建立的一個華人人物誌資料庫。

### (二) 命名實體識別模組

命名實體識別(NER)，又稱為專名識別，是自然語言處理過程中一項基礎前置處理程序。命名實體識別一般是指將文本中較有意義的詞，例如人名、地名、官名、組織名、機構名和時間等進行抽取。本模組則為進行人物關係識別，以及人物與機構關係判別流程前的前處理。首先，將文本先透過BERT(Bidirectional Encoder Representations from Transformers)轉為語意向量後(Sun, Qiu, Xu, & Huang, 2019)，再進行人物實體命名辨識，所使用的NER技術係以BiLSTM-CRF為「命名實體模組」的訓練模型，流程如圖2所示。

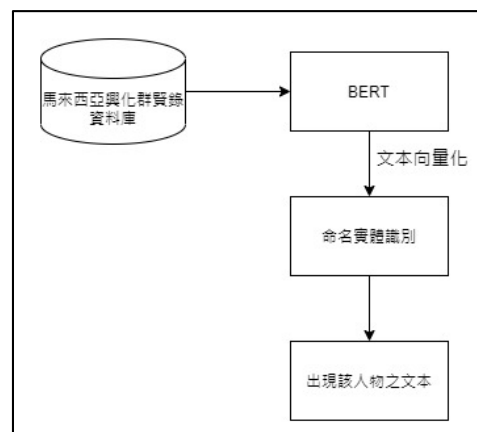


圖 2. 文本命名實體自動識別流程

### (三) 人物關係識別模組

本模組將從命名實體識別模組所提取出的人名實體，以成對方式導入至預先透過訓練資料訓練好的人物關係分類器，來進行人物關係識別。訓練人物關係分類器的資料係透過網路爬蟲，爬取維基及百度中的描述人物實體關係的句子來做為訓練文本。目前分類器將人物關係區分為夫妻、手足、親子、朋友、師生、親戚、工作、其他及無關係等共九類關係。分類的項目是與熟知「馬來西亞史」及「馬來西亞興化群賢錄」之人文學者討論後初步歸納。若識別結果非無關係類別但其關係未被分至分類器類別中，則將結果判為「其他」並存入「人物與機構關係資料庫」中，作為以視覺化方式呈現於人物與機構知識圖譜介面使用。

### (四) 人物與機構關係判別模組

本模組目的是抽取與目標人物有關聯的機構，透過整合與該機構有關之人物及他們對於該機構的貢獻和事蹟，提供人文學者有別於單純僅以人物關係作為解讀人物誌文本的方法。透過機構的面向來解讀人物誌文本的考量，是希望藉此功能輔助人文學者探索到更深一層的潛在人物實體關係。本研究所使用的命名實體識別模組在提取出實體後，自動給予實體一個標籤，例如人物會以PER標示，地區會以LOC標示，而本模組需要的機構則會以ORG來標示，透過這些標籤，我們便能得到與目標人物可能有關的機構。另外，本模組也會透過正規表達式的方法，以機構名稱進行文本篩選，提取提及該機構名稱的段落，作為輔助人文學者判斷該人物對於該機構貢獻的文本解讀之用。

## 三、系統介面與功能

本節說明本研究發展之「人物與機構關係之知識圖譜分析工具」的使用者介面與功能，包括「知識圖譜頁面」、「檢索篩選器功能介面」與「關係資訊頁面」等三個主要部分，以輔助人文學者解讀文本中人物與人物，以及人物與機構之間的脈絡關係，達到遠讀結合近讀的交互參照功能，整體使用者介面如圖3所示。各功能說明如下：

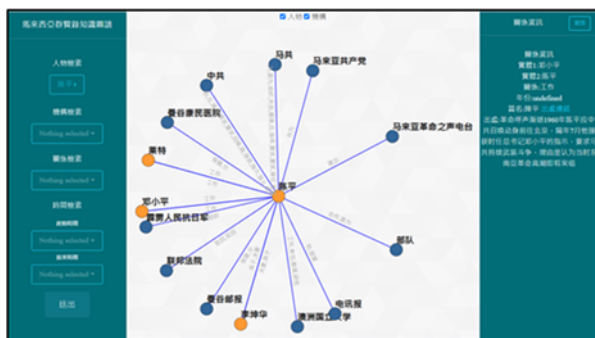


圖 3. 「人物與機構關係之知識圖譜分析工具」整體使用者介面

### (一) 知識圖譜頁面

「知識圖譜頁面」主要分為「人物與人物關係識別」、「人物與機構關係判斷」兩項功能，系統會透過這兩項功能建構一張視覺化的知識圖譜，可讓人文學者透過人物與人物關係或人物與機構關係的不同面向切入，以更多元的資訊呈現方法來輔助人文學者探索傳記類文本的人物與人物，以及人物與機構關係脈絡。

系統將命名實體識別技術所提取到的人物名進行關係識別後呈現在知識圖譜的連接線上，這些關係包含夫妻、工作和親子等，人文學者可以快速的瀏覽得知這些傳記類文本中主要人物之間的關係，如圖4所示。另外，人物與機構關係則是會顯示與該人物有關的機構名稱，一樣以連線的方式與人物節點連接。在知識圖譜中，橘色為人物節點，藍色為機構節點，系統也提供過濾器功能，避免當圖上資訊量過多時會造成使用者閱讀上的不便，如圖5所示。

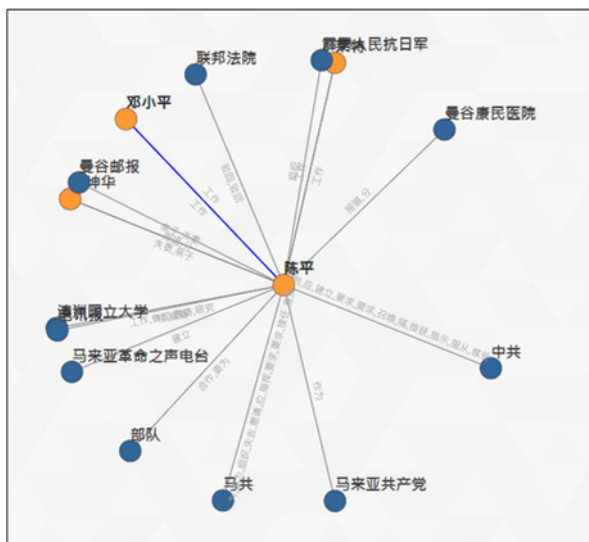


圖 4. 知識圖譜分析結果之使用者介面

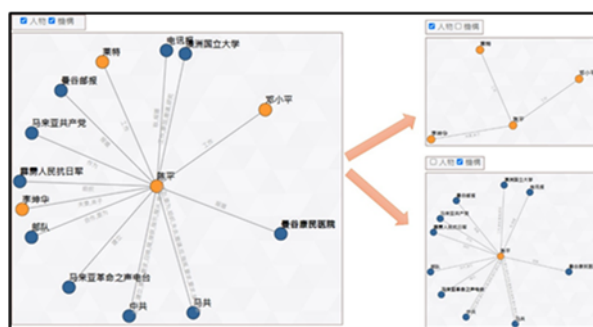


圖 5. 知識圖譜過濾器

### (二) 檢索篩選器功能介面

如圖6所示之「檢索篩選器功能介面」提供人文學者不同檢索條件，其中「人物檢索」、「機構檢索」和「關係檢索」能讓使用者以三個面向為切入點來產生知識圖譜，該介面亦提供多選功能，目的為方便使用者同時比較兩位人物或兩個機構間共同擁有哪些相關資訊。此外，此一介面亦包含時間元素，設計時考量到若人文學者將時間區段切得過於細，會導致過濾出來的資料量太少，因此提供以十年為區間的時間切割檢索，讓人文學者透過時間檢索功能，可以過濾出想觀看時間的知識圖譜。



圖 6. 檢索篩選器功能介面

### (三) 關係資訊頁面

「關係資訊頁面」係針對人文學者所點擊的知識圖節點或結點之間的連線所給予更進一步的補充資訊，如圖7所示。顯示的補充資訊包括實體名稱、實體關係、關係年份、關係出處篇章，以及關係出處段落等。其中實體名稱及實體關係為知識圖譜上所呈現之資訊，關係年份為該關係線所出現的年份，關係出處篇章除了提供出處的篇名外，亦可連結回該人物篇章進行閱讀。最後關係出處段落則為當初判斷該關係時所對應到的文本段落，上述這些呈現在此頁面的補充資訊，提供人文學者一套遠讀結合近讀的傳記類文本關係解讀模式。





圖 7. 機構資訊檢視介面示意圖

## 肆、研究方法

### 一、研究設計

為了驗證本研究發展之「人物與機構關係之知識圖譜分析工具」是否能有效輔助人文學者進行傳記類文本之人物與機構關係解讀，本研究採用準實驗研究法之對抗平衡設計(counterbalanced design)，比較使用有「人物與機構關係之知識圖譜分析工具」之「馬來西亞興化群賢錄資料庫」的實驗組實驗對象與使用無「人物與機構關係之知識圖譜分析工具」之「馬來西亞興化群賢錄資料庫」的控制組實驗對象，在解讀人物與機構關係成效、解讀人物脈絡探索效率、科技接受度及行為模式上是否具有顯著差異，以評估採用有無「人物與機構之知識圖譜分析工具」輔助人文學者進行數位人文研究之效益。

本研究於實驗結束後，邀請實驗組與控制組實驗對象填寫兩個系統使用過後的科技接受度問卷，探討兩組實驗對象對於兩個系統的整體科技接受度，以及科技接受度之「認知有用性」與「認知易用性」感受是否具有顯著的差異。同時進一步進行半結構式深度訪談，以瞭解實驗對象在使用兩組不同工具輔助解讀傳記類文本人物與機構關係脈絡之感受與建議，希望據此訪談補足量化分析上的不足。訪談內容包括「使用「人物與機構關係之知識圖譜分析工具」輔助解讀傳記類文本人物與機構關係脈絡上是否遇到困難？」、「對於「人物與機構關係之知識圖譜分析工具」的介面與功能設計是否能理解如何有效使用？」、「對於「人物與機構關係之知識圖譜分析工具」是否有改善的建議？」和「使用「人物與機構關係之知識圖譜分析工具」是否有助於解讀人物與機構關係脈絡？」等問題。

### 二、實驗流程

本研究以「馬來西亞興化群賢錄資料庫」為目標傳記類文本進行實驗設計，該傳記收錄了85位馬來西亞興化地區在政界、商界、文教、文藝界等不同領域的代表性華人人物。實驗對象會採用有無「人物與機構關係之知識圖譜分析工具」之「馬來

西亞興化群賢錄資料庫」輔助進行文本人物與機構關係脈絡解讀，整體實驗流程如圖8所示。實驗共分為四個階段，總計實驗時間為130分鐘，詳細說明如下：

實驗說明階段會先向實驗對象說明整個實驗的流程，並且在介紹完實驗目的與流程後填寫實驗同意，時長為10分鐘。為了避免實驗結果因為兩個不同實驗工具的使用先後順序而受到影響，本研究採用讓每位使用者交互使用兩個不同工具輔以傳記類文本人物與機構關係脈絡解讀之對抗平衡設計，故將實驗對象分為甲、乙兩組，如表1所示。在實驗實施階段一，會依據不同組別進行實驗工具的說明，包含工具中各功能的設計目的及其操作方法，時長為10分鐘。接著甲組的實驗對象會先使用具「人物與機構關係之知識圖譜分析工具」之「馬來西亞興化群賢錄資料庫」對經濟面向之文本進行探索，並且在探索的過程中填寫人物脈絡探索表，該探索表主要係針對文本中的潛在人物與機構關係進行內容的發想，時間為30分鐘；乙組的實驗對象則使用無「人物與機構關係之知識圖譜分析工具」之「馬來西亞興化群賢錄資料庫」進行人物與機構關係脈絡的探索，並針對文教面向進行探索。任務完成後邀請兩組實驗對象填寫科技接受度問卷與半結構訪談，時間合計20分鐘。接著，進行實驗實施階段二，此時甲、乙兩組所使用的工具和探索的任務面向(經濟/文教)會與實驗實施階段一相反，實驗對象會根據各自被分配到的任務面向進行人物與機構關係脈絡解讀、填寫人物脈絡探索表，時間共計30分鐘。任務完成後邀請兩組實驗對象填寫科技接受度問卷，時間為5分鐘。而兩組實驗對象於實驗實施階段一與實驗實施階段二的系統操作行為皆會透過 Google Analytics 進行紀錄。於任務完成後，邀請兩組實驗對象進行半結構訪談，每個受訪者的受訪時間為15分鐘，希望透過訪談質性資料的分析，彌補量化分析之不足。

表 1. 甲、乙組別實驗參與人數

組別	第一階段人物脈絡探索 (文教面向)	第二階段人物脈絡探索 (經濟面向)	受測人數
甲	有知識圖譜功能之人物與機構關係分析工具	無知識圖譜功能之人物與機構關係分析工具	6
乙	無知識圖譜功能之人物與機構關係分析工具	有知識圖譜功能之人物與機構關係分析工具	6

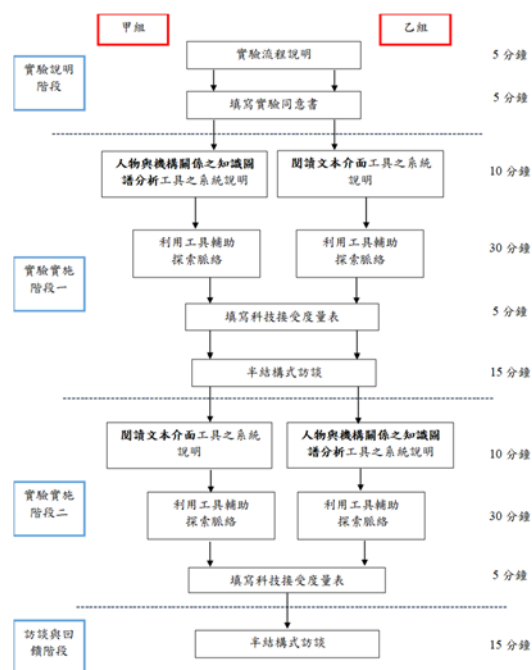


圖 8. 本研究之實驗流程

### 三、研究對象

本研究之實驗對象需要對「馬來西亞興化群賢錄」文本具有一定程度瞭解的具數位人文研究需求學者。考慮到研究成本、時間與空間等因素，本研究邀請對「馬來西亞興化群賢錄」或馬來西亞當地歷史具有一定認識馬來西亞籍華人12人為研究對象，其中包含碩士生(含畢業)4人、大學生(含畢業)8人；男性8人，女性4人；所有實驗對象年齡分布介於21歲到42歲之間。這些研究對象均對於採用數位工具輔以數位人文研究具有興趣。

### 四、研究工具

#### (一) 人物與機構關係之知識圖譜分析工具

本研究所發展之「人物與機構關係之知識圖譜分析工具」係透過命名實體識別(NER)技術進行人物與機構實體名稱識別前處理後，再以機器學習之關係識別技術將人物與人物之間的關係進行自動化的判別，再輔以提供使用者以機構的面向來瞭解整個文本的人物與機構關係脈絡。視覺化的部分則是透過提取不同種類的實體與關係信息來組成知識圖譜，提供使用者遠讀分析的視角，並且透過閱讀文本介面達到遠讀結合近讀的交互參照效果。

#### (二) 人物脈絡探索表

為紀錄受測對象於實驗中所探索到的人物與機構關係脈絡，本研究設計「文教觀點之人物脈絡探索表」與「經濟觀點之人物脈絡探索表」兩個人物與機構脈絡關係探索任務。讓實驗對象於工具使用過程中填寫這兩個表格，以紀錄其在特定任務面向下，「感興趣之實體」、「脈絡關係描述」、「重點事蹟」，以及「進一步重點整理」等資訊，並邀請熟知「馬來西亞興化群賢錄」之馬來西亞人文學者共同

訂定人物脈絡探索表之評分機制。

### (三) 科技接受度問卷

本研究參考 Hwang、Yang 與 Wang(2013)所編製之科技接受度問卷，並根據實驗需求些微調整問卷中的內容語句，以符合本研究調查使用本研究所發展「人物與機構關係之知識圖譜分析工具」之科技接受度需求。此一問卷採用李克特六點量表，其中包括兩大構面，分別為「認知有用性」8題，以及「認知易用性」3題，合計11題。在問卷信度方面，「認知有用性」構面之 Cronbach's  $\alpha$  值為0.95；「認知易用性」構面之 Cronbach's  $\alpha$  值為0.94，兩個構面皆達到良好的信度。此一問卷於實驗後邀請研究對象填寫，藉以瞭解實驗對象採用有無「人物與機構關係之知識圖譜分析工具」之「馬來西亞興化群賢錄資料庫」輔助解讀人物與機構關係的科技接受度。

### (四) Google Analytics

Google Analytics 是 Google 提供的免費網站系統分析服務，此一分析服務可以幫助網站和應用程式的擁有者，追蹤使用者使用網站和應用程式的操作行為。本研究採用 Google Analytics 技術，紀錄受試的人文學者使用有無「人物與機構關係之知識圖譜分析工具」之「馬來西亞興化群賢錄資料庫」輔助解讀人物與機構關係的點擊系統功能、輸入查詢和移動滑鼠等行為，再將紀錄的行為歷程進行滯後序列分析，探討實驗對象之有效解讀人物與機構關係的行為模式轉移。

### 伍、實驗結果與分析

本研究採用無母數統計法中的曼恩-惠尼 U 檢定 (Mann-Whitney U Test)，分別分析採用兩種不同工具輔以馬來西亞人物誌之人物脈絡探索成效、人物脈絡探索效率及科技接受度是否具有顯著的差異。敘述統計結果如表2所示。

表 2. 使用兩組不同系統進行人物脈絡探索之敘述統計

項目	組別			
	有知識圖譜功能之分析工具(n=12)		無知識圖譜功能之分析工具(n=12)	
	平均數	標準差	平均數	標準差
探索成效得分	19.17	7.49	12.50	5.96
探索所花時間	248.81(s)	98.617(s)	403.67(s)	242.719(s)
整體科技接受度	5.30	0.63	4.82	0.98
認知有用性	5.40	0.64	4.69	1.24
認知易用性	5.06	0.66	5.17	1.00

#### 一、兩系統輔以人物脈絡探索之成效差異分析

表3為使用不同工具之實驗對象於人物脈絡探索成效之曼恩-惠尼 U 檢定分析結果，結果顯示使用兩種不同工具實驗對象在人物脈絡探索成效得分(U=33.50,  $p=.026<.05$ )上具有顯著的差異，並且採用「有知識圖譜功能之人物與機構關係分析工具」

的實驗對象顯著優於採用「無知識圖譜功能之人物與機構關係分析工具」實驗對象。

表 3. 使用兩組不同系統於人物脈絡探索成效之曼恩-惠尼 U 檢定結果

項目	組別				曼恩-惠尼 U	顯著性(雙尾)
	有知識圖譜功能之分析工具		無知識圖譜功能之分析工具			
	(n=12)		(n=12)			
	等級 平均 數	等級 總和	等級 平均 數	等級 總和		
探索 成效 得分	15.71	188.50	9.29	111.50	33.50*	.026

\*  $p < .05$

## 二、兩系統輔以完成人物脈絡探索之效率差異分析

曼恩-惠尼 U 檢定結果如表4所示。結果顯示分別使用不同系統之實驗對象，在完成人物脈絡探索中找到人物關係的平均時間( $U=42.50$ ,  $p=.084 > .05$ )上，並未達顯著的差異。

表 4. 使用兩組不同系統於人物脈絡探索所花時間之曼恩-惠尼 U 檢定結果

項目	組別				曼恩-惠尼 U	顯著性 (雙尾)
	有知識圖譜功能之分析工具		無知識圖譜功能之分析工具			
	(n=12)		(n=12)			
	等級	等級	等級	等級		
	平均	總和	平均	總和		
數	數	數	數			
探索所花平均時間	10.04	120.50	14.96	179.50	42.50	.084

## 三、兩系統輔以人物脈絡探索之科技接受度差異分析

曼恩-惠尼 U 檢定結果如表5所示。結果顯示兩組實驗對象於整體科技接受度、認知有用性和認知易用性上皆無顯著的差異，表示實驗對象在輪流使用完兩個不同系統後，主觀感受兩個系統的整體科技接受度，以及認知有用性和認知易用性相當。此外，從敘述統計上的結果來看，實驗對象對於兩個不同系統在整體科技接受度、認知有用性及認知易用性上均呈現相當高之正面態度。

表 5. 使用兩組不同系統於人物脈絡探索之科技接受度之曼恩-惠尼 U 檢定結果

項目	組別				曼恩-惠尼 U	顯著性 (雙尾)
	有知識圖譜功能之分析工具		無知識圖譜功能之分析工具			
	(n=12)		(n=12)			
	等級平均數	等級總和	等級平均數	等級總和		
整體科技接受度	14.00	168.00	11.00	132.00	54.00	.297
認知有用性	14.42	173.00	10.58	127.00	49.00	.178
認知易用性	11.17	134.00	13.83	166.00	56.50	.345

## 四、兩系統輔以人物脈絡探索之使用者行為歷程分析

本研究透過 Google Analytics 的微歷程行為記錄器，紀錄實驗對象在操作兩種工具進行人物脈絡探索之各項系統工具操作行為，並以滯後序列分析(Lag Sequential Analysis)來分析具有顯著行為轉移之序列，以了解實驗對象進行人物脈絡探索之有效行為模式。實驗對象使用有無「知識圖譜功能之人物與機構關係分析工具」之系統操作行為序列轉移圖分別如圖9、圖10所示。圖中各節點表示使用者的操作行為，箭頭表示行為之間的轉移，箭頭上的數值則表示該行為轉移之 Z 值，若  $Z > 1.96$  即表示該行為轉移達顯著的水準，顯著水準超過7則以粗箭頭表示，行為轉移圖只繪製達顯著水準的行為轉移，以下針對使用兩個工具實驗對象之行為歷程轉移差異進行分析。

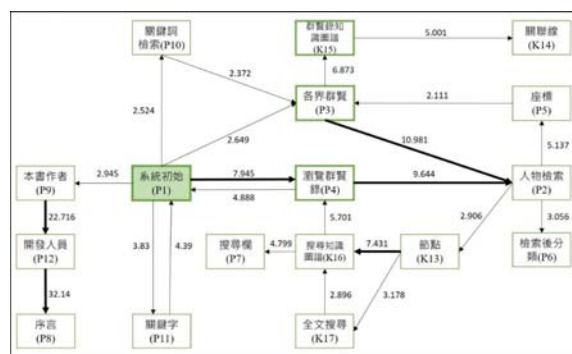


圖 9. 採用「有知識圖譜功能之人物與機構關係分析工具」實驗對象之顯著行為轉移序列圖



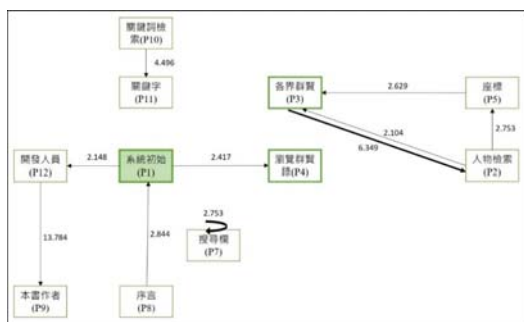


圖 10. 採用「無知識圖譜功能之人物與機構關係分析工具」實驗對象之顯著行為轉移序列圖

### (一) 採用不同系統實驗對象操作「各界群賢分頁」的偏好方式

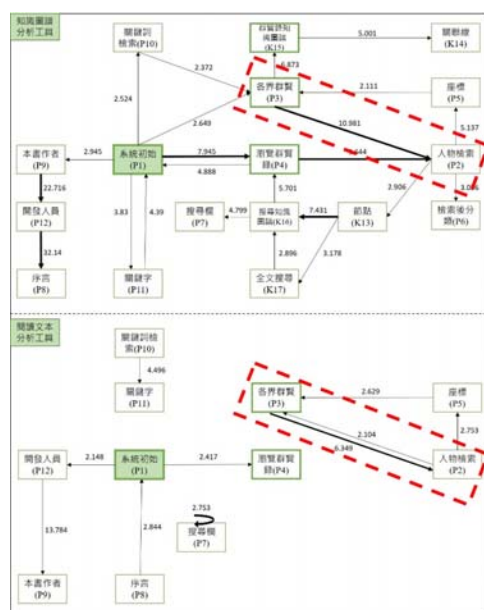


圖 11. 採用不同系統實驗對象之「各界群賢分頁」的偏好方式轉移序列比較圖

採用「有知識圖譜功能之人物與機構關係分析工具」之實驗對象在「點擊各界群賢」到「人物檢素」出現顯著的行為轉移( $Z=10.981>1.96$ ), 如圖11上半部所示。而採用「無知識圖譜功能之人物與機構關係分析工具」之實驗對象則除了在「點擊各界群賢」到「人物檢素」中出現顯著的行為轉移( $Z=2.104>1.96$ )外, 在「人物檢素」回到「點擊各界群賢」也出現顯著的行為轉移( $Z=6.349>1.96$ ), 如圖11下半部所示。本研究依據訪談結果推測實驗對象在使用「有知識圖譜功能之人物與機構關係分析工具」時, 在人物檢素介面中即能夠透過系統提供的個人知識圖譜準確地看到人物與人物, 以及人物與機構之間的關係, 進而完成學習單所賦予的人物脈絡探索任務。但是採用「無知識圖譜功能之人物與機構關係分析工具」之實驗對象在「點擊各界群賢」後, 只能單純透過文本閱讀方式進行人物脈絡探索任務, 因此較難找到符合預期所需的內容, 這也迫使實驗對象必須反覆的切換「各界群賢頁面」和「人

物檢素頁面」進行比對查找, 因此在「人物檢素」與「各界群賢」也出現了顯著的行為轉移( $Z=2.104>1.96$ )。

### (二) 採用不同系統實驗對象操作「瀏覽群賢錄分頁」的偏好方式

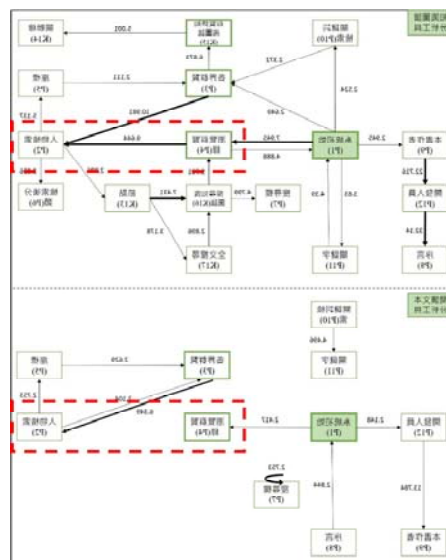


圖 12. 採用不同系統實驗對象之「各界群賢分頁」的偏好方式轉移序列比較圖

採用「有知識圖譜功能之人物與機構關係分析工具」之實驗對象在「點擊瀏覽群賢錄」與「人物檢素」中出現顯著的行為轉移( $Z=9.644>1.96$ ), 如圖12上半部所示。採用「無知識圖譜功能之人物與機構關係分析工具」之實驗對象在「點擊瀏覽群賢錄」與「人物檢素」中則並未出現顯著的行為轉移, 如圖12下半部所示。本研究透過訪談結果推測實驗對象在操作「有知識圖譜功能之人物與機構關係分析工具」時, 通常會透過「點擊瀏覽群賢錄」、「人物檢素」、「點擊節點」與「點擊搜尋知識圖譜」之行為組合完成所需之人物脈絡探索流程。但是採用「無知識圖譜功能之人物與機構關係分析工具」中的人物檢素頁面則並未有個人的知識圖譜, 可能因此導致實驗對象在點擊人物檢素後並未得到有用的資訊, 因而不以此一方式進行人物脈絡探索。

## 五、訪談資料分析

為了瞭解實驗對象對於兩個系統輔以人物脈絡探索上的看法、感受與建議。因此, 在實驗結束後邀請12名實驗對象進行半結構式深度訪談, 藉此得到實驗對象的使用感受與對於系統後續功能改善的建議, 以此彌補量化資料分析之不足, 並進行實驗結果之交互驗證。

首先, 多數實驗對象認為, 在對文本不熟悉的情況下, 相較於只能透過閱讀文本及使用基本檢素工具來探索人物脈絡, 知識圖譜分析工具不但是個很好的探索切入點, 亦有助於更快速的探索人物脈絡, 並加速瞭解整個文本的意涵。但是部分實驗對象也提到知識圖譜工具在操作上的困難點, 例如呈現資訊量有點過多以及判斷人物彼此關係與人物



與機構關係的精確性上需要進行改善等。接著，實驗對象認為「檢索篩選器」功能能讓他們針對欲探索的面向進行篩選，檢索上很直觀，且所呈現的內容一目了然。但同樣也提及，在對於文本不太熟悉的情況下，檢索器提供的分類有點太多。此外，在「過濾器功能」上，實驗對象提及能在檢索後進行人物或機構的過濾是非常實用的，因為在一次檢索多個人物或機構時，知識圖譜工具提供的資訊量可能會過於龐大，若能針對人物或機構進行過濾，不僅能降低資訊負荷，更能幫助他們聚焦於感興趣的內容。但部分實驗對象亦提及，若能在過濾器上新增更多過濾選項，例如，機構的結構分層(如青年團總會、麻六甲青年團、婦女團等)，將會更有幫助。最後，實驗對象認為知識圖譜工具提供了人物與機構的實體關係視覺化雖已足夠，但若增加其他實體識別的視覺化，如地名、官名或勳銜等，將會使得探索面向及過程更加多元與豐富。

## 陸、討論

本研究透過曼恩-惠尼 U 檢定 (Mann-Whitney U Test)，分別針對受測對象使用有無「知識圖譜之人物與機構關係分析工具」輔以進行人物脈絡探索之成效、人物脈絡探索之效率、科技接受度及行為模式進行分析。在人物脈絡探索成效顯示，採用「有知識圖譜功能之人物與機構關係分析工具」實驗對象完成之人物脈絡探索成效顯著高於採用「無知識圖譜功能之人物與機構關係分析工具」實驗對象，表示知識圖譜分析工具更能幫助實驗對象釐清文本中的人物脈絡關係，並提供其更快更有統整性的資料。而在探索人物脈絡的效率上，結果雖然顯示採用「有知識圖譜功能之人物與機構關係分析工具」的實驗對象在完成人物脈絡探索上的時間效率並無顯著優於採用「無知識圖譜功能之人物與機構關係分析工具」之實驗對象，但從敘述統計得知，採用「有知識圖譜功能之人物與機構關係分析工具」的實驗對象的平均檢索時間遠低於採用「無知識圖譜功能之人物與機構關係分析工具」的實驗對象，表示採用知識圖譜分析工具能更快速的掌握整體人物脈絡資訊，並且提供實驗對象更清晰的人物脈絡探索過程。

本研究發現受測對象分別使用有無「知識圖譜之人物與機構關係分析工具」，在整體科技接受度、認知有用性與認知易用性等面向上皆未達統計上的顯著差異。此一結果顯示知識圖譜功能的增加，並未造成使用者在系統操作上有不適應或不友善的地方，並且在敘述統計結果顯示，採用「有知識圖譜功能之人物與機構關係分析工具」實驗對象在整體科技接受度、認知有用性及認知易用性平均數皆遠高於李克特六點量表之中位數，表示實驗對象對於其輔以人物脈絡探索，均持相當高的肯定態度。

從滯後序列分析結果發現，採用「有知識圖譜功能之人物與機構關係分析工具」之實驗對象在「各界群賢」至「人物檢索」的行為達顯著轉移；而採用「無知識圖譜功能之人物與機構關係分析工具」實驗對象則是在「各界群賢」至「人物檢索」，以及「人物檢索」至「各界群賢」彼此來回之間出現顯著行為轉移。從訪談結果推測，「有知識圖譜功

能之人物與機構關係分析工具」能讓實驗對象準確的看到人物與人物，以及人物與機構的關係，並透過檢索機構面向的方式進而完成更為深入的人物脈絡探索。而採用「無知識圖譜功能之人物與機構關係分析工具」實驗對象在「點擊各界群賢」後，並沒有準確的資訊可輔助其快速查看該人物是否符合人物脈絡探索需求，因此只能單純透過文本閱讀方式進行，導致其需投注更多的時間在閱讀文本上，但是閱讀完後常常會發現沒有所需的人物脈絡內容，迫使其必須反覆地切換「各界群賢頁面」和「人物檢索頁面」。另外，採用「有知識圖譜功能之人物與機構關係分析工具」實驗對象在「點擊瀏覽群賢錄」與「人物檢索」中出現顯著行為轉移，而採用「無知識圖譜功能之人物與機構關係分析工具」則無。從訪談結果推，實驗對象在操作知識圖譜分析工具時能有一套較為有效率的探索方式，以完成一次人物脈絡探索流程，並重複之。但是「無知識圖譜功能之人物與機構關係分析工具」中並未有此類功能，因而造成實驗對象的平台使用行為較為零散，缺乏一套有效率的人物探索方式。由上述各項分析結果得知，發展知識圖譜之遠讀視覺化工具的同時，搭配鏈結回文本內文的近讀功能，能讓使用者能有效幫助使用者更深入的分析文本，該項結果與 Fisher 和 Frey(2012)的研究結果一致。

## 柒、結論與未來研究

本研究之研究目的為發展輔助人物與機構關係解讀之「人物與機構關係之知識圖譜分析工具」，以有效支援人文學者探索傳記類文本中人物與人物，以及人物與機構之間的關係脈絡。特別是藉由機構面向讓人文學者不只單純以人物關係進行遠讀探索，還能同時以人物與機構面向來探索隱藏在文本中的潛在人物關係，進而對傳記類文本之人物與機構關係進行更深一層之解讀。分析結果顯示，採用「具知識圖譜之人物與機構關係分析工具」有助於提升使用者進行人物脈絡探索的成效，幫助使用者更加精確的掌握文本中人物與人物，以及人物與機構之間的關係。在探索人物脈絡的效率上，使用兩系統完成人物脈絡探索的時間上未達顯著的差異，但是於敘述統計上發現採用「具知識圖譜之人物與機構關係分析工具」受測對象完成人物脈絡探索所花費的平均時間，遠少於採用「無知識圖譜之人物與機構關係分析工具」受測對象。表示知識圖譜分析工具不但提供人物脈絡探索的切入點，還能幫助使用者提升人物脈絡探索的效率。此外，不論是在整體科技接受度、認知有用性，以及認知易用性上，受測對象對於兩個系統輔以人物脈絡探索的適切性上均持正面肯定態度。從滯後序列分析與訪談資料顯示，大部分受測對象認為在使用「有知識圖譜功能之人物與機構關係分析工具」時，能夠透過系統提供的個人知識圖譜準確地看到人物與人物，以及人物與機構之間的關係，並藉由一套有系統性的行為組合進行人物脈絡探索，以完成實驗所賦予的人物脈絡探索任務。但是採用「無知識圖譜功能之人物與機構關係分析工具」之受測對象，因較難找到符合預期所需的內容，這也造成受測對象花費更多時間在尋找有用的人物資訊。

根據上述研究結論，本研究提出了未來研究方向。首先，本研究目前透過命名實體識別技術取出呈現於知識圖譜的實體資訊包括人物與機構，未來可以考慮引入時間與地名等其他重要實體至知識圖譜中，讓知識圖譜可以發揮更多輔助人物情境探索之效益。其次，亦可將更多南洋人物誌或名人傳記文本等典藏至此一平台，形成一個內容更為豐富的南洋人物與名人傳記資料庫，應更能展現知識圖譜輔助人物情境探索。

## 參考文獻

- Al-Khatib, K., Yufang, H., Henning, W., Charles, J., Francesca, B., & Benno, S. (2020). End-to-end argumentation knowledge graph construction. *Proceedings of the AAAI Conference on Artificial Intelligence*, 34, 7367-7374. Retrieved from <https://doi.org/10.1609/aaai.v34i05.6231>
- Berners-Lee, T., James, H., & Ora, L. (2001). The semantic web. *Scientific American*.
- Bingenheimer, M., Hung, J. J., & Wiles, S. (2011). Social network visualization from TEI data. *Literary and Linguistic Computing*, 26(3), 271-278.
- Bradley, J. R., & Conway, R. W. (2003). Managing cyclic inventories. *Production and Operations Management*, 12(4), 464-479. Retrieved from <https://doi.org/10.1111/j.1937-5956.2003.tb00215.x>
- Chen, C. M. & Chang, C. (2019). A Chinese ancient book digital humanities research platform to support digital humanities research. *The Electronic Library*, 37(2), 314-336.
- Chen, C. M., Chang, C., & Chen, Y. T. (2020). A character social network relationship map tool to facilitate digital humanities research. *Library Hi Tech*.
- Chen, S.-P., Hsiang, J., Tu, H.-C., & Wu, M. (2007). On building a full-text digital library of historical documents. In *International Conference on Asian Digital Libraries* (pp. 49-60). Springer.
- Fisher, D., & Frey, N. (2012). Student and teacher perspectives on a close reading protocol. *Literacy Research and Instruction*, 53, 25-49. Retrieved from <https://doi.org/10.1080/19388071.2013.818175>
- Haslhofer, B., Antoine, I., & Rainer, S. (2018). Knowledge graphs in the libraries and digital humanities domain. Retrieved from [https://doi.org/10.1007/978-3-319-63962-8\\_291-1](https://doi.org/10.1007/978-3-319-63962-8_291-1)
- He, Y., Yun, H., & Lin, L. (2019). The character relationship mining based on knowledge graph and deep learning, 2019 5th International Conference on Big Data Computing and Communications (BIGCOM), Qing Dao, China, 22-27. doi: 10.1109/BIGCOM.2019.00011
- Hockey, S. (2004). The history of humanities computing. In R. Siemens & S. Schreibman, (Eds.), *A Companion to Digital Humanities*. Retrieved from <http://www.digitalhumanities.org/companion/>
- Hwang, G. J., Yang, L. H., & Wang, S. Y. (2013). A concept map-embedded educational computer game for improving students' learning performance in natural science courses. *Computers & Education*, 69, 121-130.
- Jänicke, S., Franzini, G., Cheema, M. F., & Scheuermann, G. (2015). On close and distant reading in digital humanities, *A Survey and Future Challenges*, 21.
- Kirschenbaum, M. (2012). What is Digital Humanities and what's it doing in English departments? *Debates in the Digital Humanities*, 3.
- Moretti, G., Sprugnoli, R., Menini, S., & Tonelli, S. (2016). ALCIDE: Extracting and visualising content from large document collections to support humanities studies. *Knowledge-Based Systems*, 111, 100-112.
- Pujara, J., Hui, M., Lise, G., & William, C. (2013). Knowledge graph identification. *The Semantic Web-ISWC*, 542-57. *Lecture Notes in Computer Science*. Berlin, Heidelberg: Springer. Retrieved from [https://doi.org/10.1007/978-3-642-41335-3\\_34](https://doi.org/10.1007/978-3-642-41335-3_34)
- Rosenzweig, R. (2003). Scarcity or abundance? Preserving the past in a digital era. *The American Historical Review*, 108(3), 735-762.
- Schreibman, S., Siemens, R., & Unsworth, J. (2008). A companion to digital humanities. Retrieved from <http://www.digitalhumanities.org/companion/>
- Shen, Y., Ding, N., Zheng, H., Li, Y., & Yang, M. (2020). Modeling relation paths for knowledge graph completion. *IEEE Transactions on Knowledge and Data Engineering*. doi: 10.1109/TKDE.2020.2970044
- Singhal, A. (2012). Introducing the knowledge graph : Things, not string. *Official Blog (of Google)*. Retrieved from <https://blog.google/products/search/introducing-knowledge-graph-things-not/>
- Sun, C., Qiu, X., Xu, Y., & Huang, X. (2019). How to fine-tune BERT for text classification? *Chinese Computational Linguistics*, 194-206. Retrieved from [https://doi.org/10.1007/978-3-030-32381-3\\_16](https://doi.org/10.1007/978-3-030-32381-3_16)
- Tu, A., Hung, J.-J., & Lin, Y.-H. (2012). Building a text analysis platform for Chinese Buddhist text -An example based on CBETA and Tripitaka Catalog Projects. Presented at the PNC 2012 Annual Conference and Joint Meetings, UC Berkeley School of Information, USA. Retrieved from <http://joeyhung.info/publications/>