

※ 「數位時代的出土文獻」專輯 ※

「數位時代的出土文獻」專輯弁言

范麗梅*

數位科技的迅速擴展，在二十一世紀的今天，已滲透、參與、激盪了現代學術的各個領域。就中國學術的人文學科而言，重製文獻資料，以利議題的提出、考證與論述，這已是研究中不可缺少的重要環節。似乎有無窮無盡的為了滿足於不同需求的電子資料庫，在學者之間或互聯網上不斷被提及、散布與流傳。若嘗試追溯這股潮流，或許可以聯想到上世紀初，燕京大學洪業先生與多位學者共同編纂的「引得」系列，這項工作將數十種重要的古籍進行校勘、整理與條例化處理，大大拓展了使用的便利性、有效性與全面性¹。如同其《引得說》指出的，在「引得」系列的構想實現以前，傳統學術已出現不少為了滿足個別治學需求的工具書²。誠然，包括字書、韻書、辭書、類書、目錄等工具性質的著作，在過去已是層出不窮，反映的是歷代學者面對各種文獻資料所持有的基本態度與方式。從最廣泛的角度看來，如何有效地處理數量龐大的信息資訊，或許一直以來就是人類在擁有、保存、整理、歸納——當然也包括捨棄、破壞、禁絕——文獻資料時，不斷反覆出現的一個主旋律。數位科技發皇進步、震耳欲聾的聲響，若置於此歷史脈絡中，或可說是此旋律在新時代的最新變奏。

中國學術在上世紀的另一重要發展，是過去百年出土文獻的陸續問世。此嶄新局面濫觴自十九世紀末，延續至二十一世紀蔚為大觀。尤其是上古的出土古文字文獻，提供了因時代邈絕而難以替換的第一手材料，直接刺激學者去關注新發現所促

* 范麗梅，本所助研究員。

¹ 此後，香港中文大學劉殿爵先生引入數位技術，在先秦兩漢部分延續了洪業先生的工作。

² 洪業：《引得說》（北平：哈佛燕京學社引得編纂處，1932年）。

發的許多新學問³。而這些都建立在新材料應如何被準確地辨識、理解，如何被有效地納入自有傳承而又與現時交融的既有的知識架構中，又如何促使新的知識架構繼起而建立？回答這些問題，是現代學者研究出土文獻所面臨的艱鉅考驗。

出土文獻的數位化，是以上兩大趨勢的交集。近幾十年來，香港、臺灣、大陸、日本、歐美等多個地區，已有專家學者陸續建立多種出土文獻的數位資料庫，這些努力顯示了上述問題的急迫性。具體而言，即是如何有效地結合出土文獻與數位科技，使具備第一手材料價值的出土文獻，能夠通過數位化的重製處理，增加其使用的便利性、有效性與全面性。同時，還期許可能經過各種數位化處理過程的刺激，開發出不同以往的研究方法與學術議題，以締造新時代創新、紮實的學術成就⁴。

本專輯「數位時代的出土文獻」正是基於以上關懷所提出的。一共邀請了華東師範大學中國文字研究與應用中心、中央研究院歷史語言研究所、香港中文大學中國古籍研究中心、復旦大學出土文獻與古文字研究中心、武漢大學簡帛研究中心、吉林大學古籍研究所等六個單位的學者作為代表，提供各單位相關出土文獻數位化或具體資料庫的完整訊息。六篇論文說明如下：

（一）張再興〈互聯網時代出土文獻數據庫建設的思考與實踐〉：本文提出徹底數位化的原則，開發各種類型的古文字字體，包含原形、隸定字形、偏旁構件的字形等各個層次的字形，並建立「字形關係樹」以連繫眾多的異體字形。同時，為便於檢索，還開發了「古文字字形輸入編碼」，編製「古文字輸入法」軟體。此外還製作「古文字字形檢索文檔」，以進行偏旁檢索。在字形以外，又開發具全文檢索功能的「出土文獻釋文資料庫」，此包括隸定、釋字、破讀三個層次的釋文，並據此續製合乎字形、詞彙、語法三種屬性的語料庫。以上這幾個重要的建置構想貫徹於其金文資料庫、青銅器斷代庫等多個資料庫。

（二）洪一梅〈甲骨學研究學術性系統發展芻議：理論與實踐〉：本文旨在面向二十一世紀的甲骨學研究上，擬定一個學術性系統的發展芻議。其中詳細說明了所建置具交互連結的「考古資料數位典藏資料庫」與「歷史語言研究所藏甲骨文拓

³ 參王國維：〈最近二三十年中中國新發現之學問〉，謝維揚、房鑫亮主編：《王國維全集》（杭州：浙江教育出版社，2009年）。

⁴ 關於出土文獻與數位科技結合使用的關注，可參考裘錫圭：〈推動古文字學發展的當務之急〉，《學術史與方法學的省思：中央研究院歷史語言研究所七十周年研討會論文集》（臺北：中央研究院歷史語言研究所出版品編輯委員會，2000年）。

片資料庫」，包括其功能與限制。又對現行甲骨資料庫的建置與困難提出說明，所提出的各項重要問題，可作為今後各資料庫建立者相互交流改善的目標。此外，更重要的是從社會網絡理論與資訊分享概念來說明甲骨學研究在新世代的轉化。

（三）何志華〈香港中文大學中國古籍研究中心出土文獻數據化研究之回顧與前瞻〉：如上文已指出的，「漢達文庫」為開發較早的資料庫。而在出土文獻研究方面，設計有「甲骨文資料庫」、「金文資料庫」、「竹簡書庫資料庫」三種。依據不同種類、性質的出土文獻設計適應性強的檢索功能。檢索結果皆提供釋文對照功能，使用方便。其資料庫開發較早，但已展開合乎全面、準確與檢視方便三大構建方針的重建計畫。

（四）程少軒〈「上博簡字詞全編資料庫」的設計和製作〉：本文說明「上博簡字詞全編資料庫」的設計和製作，並總結過程中的經驗教訓。此資料庫可供檢索上博楚簡的字形、詞彙、辭例和相關論著。資料庫包括一個由單字字形到全文段落的數據庫；含有字形圖片、索引、造字，或釋文文本與論著文檔的數據庫附屬文件，並據此設計字形、詞義、辭例和論著四個相對獨立的檢索程序。在總結經驗教訓中提出了開放數據庫修改功能與權限的新觀念，可在短時間內確保資料庫的學術品質。

（五）魯家亮、李靜〈數位時代的簡帛研究〉：本文說明其單位著力於網路交流平臺、數據庫、紅外線成像技術應用三方面的工作。在交流平臺方面，有計畫地邀請世界各地的學者共同參與。開發了學術動態、國際信息合作、文庫、論壇等多種平臺，大大突破了學術交流上的時、空限制，促使出土文獻的研究迅速擴展與深化。在簡帛字形、辭例資料庫方面，其中在字形上，設計對任一單字或偏旁兩種方式進行檢索，檢索結果以字形的圖片形式出現。此外還可進行單字辭例的檢索，檢索結果則以簡牘文本形式出現。與其他單位相較之下，其單位在紅外線成像技術應用方面也取得了較多的經驗。

（六）單育辰〈淺談數位化時代的出土文獻及相關資源資料庫〉：本文整理了包括甲骨文、金文、簡帛、傳世文獻等與出土文獻相關的大部分電子資料庫，為相關領域學者提供基礎的資料庫資訊。另外，又以使用者或研究者的立場提出了若干建議，包括提供電子圖像、單機檢索，以及戰國文字的綜合與分類的數位化檢索等。

總得而言，就各單位資料庫的建立情況來看，已整理或未來陸續整理的新材料

日趨多樣化，包括多種載體的文獻，例如甲骨、金文、簡牘、石刻、玉版、帛書、古璽、古陶、古幣等等。其中還包括單一載體文獻資料的詳盡蒐羅，例如甲骨方面，包括以《甲骨文合集》為首的各種重要圖錄，簡牘方面也涵蓋楚簡、秦簡、漢簡等大批材料。此外，更計劃擴及到與出土文獻研究相關的其他材料，包括《說文解字》，魏代的三體石經以及宋代《汗簡》、《古文四聲韻》等年代較晚，但來源有據的傳世材料。

出土文獻資料庫的建設工作，在相當程度上依據目前學界研究的認識或共識；但若反向考慮，資料庫是否也有可能推進或修訂，甚至徹底改變研究的現狀？以下僅以使用者的角度從「字形」、「字詞對應」、「語境」等三個方面提出一些感想，向各位撰稿人以及學界同仁請教。（下文提及各單位的研究成果，僅屬舉例，不周之處，敬請海涵。）

首先就「字形」層面而言，一個漢字字形的豐富訊息應如何據實反映並方便獲取？所謂字形，包括書體線條的變化趨勢、字形筆劃的差異結構、偏旁部件的置換組合等多個層面。這些訊息固然可以由單字圖版的連接如實提供，然而若就資料庫設計的角度而言，如何輸入、如何檢索提取，卻是一個可以提出的問題。事實上，這與過去學者討論如何「隸定」轉寫古文字資料，所接觸的許多問題有關。最主要的是涉及標準化的問題：資料庫在輸入各個字形時，應採取何種隸定標準？對於原整理者所提供的隸定字形，即使有明確理由認為有誤或不精確的，是否仍應為了資料的完整性、一致性而予以保留？抑或另外提出新的釋文，作為隨後資料庫操作的基礎？深入地說，所謂隸定問題牽涉到每一個字形獨具的歷時、歷地發展，無法在短時間內一一釐清，也難以徹底解決以引導資料庫的製作。然而可認識到的是上古出土文獻絕大多數是手寫本，而這種手寫特質才是出土文獻相對於印刷版本的最大特色。

回應以上問題，在有限範圍內，增加分層檢索的方法，或是逐步建立字形隸定的統一標準等，是目前各單位努力的方向。例如在分層檢索上，由武漢大學發軔，各個單位的資料庫都陸續設計了「偏旁檢索」的功能。除了如其所言：「適應了有的古文字只能認識一部分、而不能認識全字的情形。」此一設計理念也相當切合古文字與文獻考釋的實際需要，尤其是能夠較全面而系統地得到有關偏旁可能對應的多個不同字詞。與此相關，復旦大學的上博簡數據庫設有「字頭數據庫」，就字形分析其全部構件，進而提供部首或構件的索引。再者，華東師範大學在古文字字庫

中也開發了「偏旁構件的字形」，編製「古文字字形檢索文檔」，這些都是為進行偏旁檢索而設置的。

而在建立字形隸定統一標準方面，中研院長期以來不斷推動「漢字構形資料庫」的建置，同時其甲骨文文庫也有甲骨文文字與隸定字的對照。再者，華東師範大學提出了徹底數位化的原則，開發「古文字字形輸入編碼」，編製相應的「古文字輸入法」軟體，使得字形的檢索有儘量窮盡的可能。此外，復旦大學為論文檢索程序提供一個創新的概念，即預先使用一個資料庫製作者所提供的統一標準的隸定或釋文，以其中的字形作為檢索的字詞，則能有效檢索到相關論文，其中對文字隸定還設訂了嚴格的體例與標準，確保進一步的檢索準確率。在上述各單位付出的努力之餘，可以提出的是這個標準能否得到資料庫建立者與古文字學家共同合作，進而發展跨單位、統一的全字形隸定總表，以及與此總表相連接的輸入法，似乎是可以再擴大思考的問題⁵。

此外，可以特別提及的是香港中文大學所構擬研發的手寫輸入檢索。設想手寫一個古文字，輸入檢索後，若能因此獲取一系列相關的字形，這個檢索成果所扮演的角色，必然不僅止於輔助研究而已，其勢必能引導研究並開展字形認識層面的視野。更進一步說，倘若書體線條相近、字形筆劃部分相似，或偏旁部件相近的各個字形，都能通過數位處理而進行某種程度的歸納，那麼對於總結出某些演變規律相近的系列而言，資料庫所能提供的裨益，將很可能引起新課題的誕生⁶。

其次就「字詞對應」的層面而言，此牽涉到古文字與古漢語的記錄問題，是出土文獻研究最基礎，但卻最糾結的問題。復旦大學開發了有關上博楚簡的「字詞對應數據表」，同時還設計字詞關係的多功能檢索，「反映與某字或某詞密切相關的字詞關係網絡」等，已將「字詞對應」概念帶入資料庫的建置中。與此相關，在判

⁵ 2011年復旦大學由裘錫圭先生帶領參與「中華字庫」的研發工程，其成果將對出土文獻的字形隸定起莫大的統合效益。相關訊息參鍾馨：〈「中華字庫」工程研發工作啟動大會在京召開〉，復旦大學出土文獻與古文字研究中心網站，2011年8月3日，網址：http://www.gwz.fudan.edu.cn/srcShow_NewsStyle.asp?Src_ID=1612。

⁶ 與此相關，常見於漢代簡牘的所謂「習字」材料，字跡多半十分笨拙，幾不成字。這類不規範的書寫材料也考驗著目前對於字形認識的標準、定義與實踐。除一般字形外，出土文獻中還有一些諸如墨鈎、墨丁等符號的書寫，對於判斷資料的體例與分類很有幫助，但一般不被資料庫收錄。

斷何字應對應何詞時，上古音是不可缺乏的考慮因素，相關訊息的輸入可為研究者提供不少方便。事實上，已有單位注意到古文字上古音標示的需要，華東師範大學的「字形屬性庫」已採用王力古音系統的標記，而復旦大學的上博簡數據庫也標明一字的上古聲紐與韻部⁷。考慮到「字詞對應」的設計概念，若能將語音相同或相近的詞彙，以及其對應的字形，經由檢索連繫起來，應可以在研究上帶來更大的方便，甚至是發現意想不到的結果。此外，在「字詞對應」的概念下，除了上古音之外，更大膽的構想是能否再加入以詞義作為檢索單位的設計，在某種程度上將詞義接近或相同的詞彙連貫在一起？

與「字詞對應」相關的，還有各種引文、文句互見的情況。這類對照材料不僅是提供異文、異句的珍貴語言文字資料，同時還展現了寫本文獻歷時、歷地，以及不同傳抄者的側面，是文獻傳承、思想文化等各個領域不可多得的材料，涉及了出土文獻應如何結合傳世文獻來進行研究的複雜問題。這類文句若能通過數位技術加以繫連，則能有效地整合相對完備的訊息。關於這一點，香港中文大學在簡帛資料庫的設計上，若有傳世版本可資對照的文獻，則以不同顏色字體增補，或設校改紀錄，此外還提供互見文獻、相合文獻之比對等，以此展示其中異同。並且更進一步出版多種詳盡蒐集引據資料的彙編工具書，見於其「先秦兩漢典籍引經典資料彙編」系列。在此基礎上，或許可以提出以下的問題：是否有可能通過數位技術對文獻中文句互見的情況加以定義、蒐集，使得出土與傳世文獻中文句互見的部分也能夠通過檢索的方式儘量獲取？

其三是關於「語境」的問題，過往古文字研究的重點多在單文只字上，相應的工具書便是大家普遍使用的文字編（與數量相對較少的逐字索引）。隨著出土材料

⁷ 上古音研究近年來出現不少新的發展，包括通過漢語親屬語言與方言的比較研究所獲得的許多新的假說，除了 William H. Baxter（白一平）、Laurent Sagart（沙加爾）兩人陸續發表的成果，還有鄭張尚芳：《上古音系》（上海：上海教育出版社，2003年）、Axel Schuessler（許思萊）：*ABC Etymological Dictionary of Old Chinese*（Honolulu: University of Hawai'i Press, 2007）等專書，其中提供了相對完整的古音構擬，應可考慮納入出土文獻的資料庫中。承程少軒先生提示，由上海高校比較語言學 E-學院所主辦的東方語言學網站（<http://www.eastling.org>），已設有上古音檢索系統，收錄了高本漢、李方桂、王力、白一平、鄭張尚芳、潘悟雲等學者的上古音構擬。除此之外，臺灣大學中國文學系和中研院資訊科學研究所共同研發的「漢字古今音資料庫」，目前也已開放線上使用（<http://xiaoxue.iis.sinica.edu.tw/ccr/>）。上述這些資源為出土文獻資料庫在此方面的進一步擴展，提供一定的基礎。

的日益豐富，文獻類的材料逐漸增多，詮釋的層次、不同理解之間的差異也越來越複雜。古文字所在文獻中的「語境」問題，已成為文字考釋與研究的重大挑戰，也關連著此材料進一步闡釋的思想、文化與其他歷史的議題。所謂「語境」指的可能是一根竹簡、一片布帛，或同一篇文字、同一書手所寫、同一座墓所出的文字資料，也有可能是研究者基於語言、思想、文化與其他歷史背景所做的進一步繫連，涉及不同詮釋脈絡的鋪陳，直接影響出土材料性質的判斷。這些不同的脈絡應如何呈顯，以供研究判斷的需要，已是一個有必要深思的問題。

針對以上問題，多個單位已做出了相當的回應，在檢索結果上，快速連結相關資料所在的原文獻全圖版，以提供不失真的上下文句脈絡。香港中文大學就甲骨、金文、簡帛皆能於檢索結果中附有原拓片或原全圖版，以供確認。此外，中研院史語所設計的一系列資料庫，除了此處介紹的「甲骨文拓片資料庫」，尚有「殷周金文暨青銅器資料庫」，都盡量提供原物的拓片和影像，使得必備的相關訊息沒有遺漏。另一方面，華東師範大學開發的「出土文獻釋文資料庫」，設計有釋字釋文與破讀釋文等層次，並開始注意到原始文獻與各家釋讀等多層次、多方位的關連對應；又陸續出版《古文字考釋提要總覽》及配套的網絡檢索數據庫，皆有助於此問題更精細化的思考。復旦大學數據庫有「論著目錄表」連接上博簡字詞考釋相關的論文，同時其釋文要求充分吸收學術界最新的研究成果，也是此工作的基礎。

總得而言，上述這些成果無疑顯示了出土文獻這門學科的逐漸發展，雖然距離完備整合的程度還須相當的努力，但許多單位、學者付出辛勞，已開啟了邁向成功的步伐。這些成果也無疑證明了傳統學術可以因為新的發現和新的技術、新的思考角度而開創出嶄新的視野。可以期待的是，隨著數位科技重製的便捷與迅速，考古文物作為第一手的學術文獻，將能超越以往只能前往文物所在地進行研究的限制，包括文獻文字的清晰與無限放大功能、龐雜繁複相連疊出資料的有效提取、學術資訊的傳播與交流，都將引發革命性的學術議題、方法與規範。數位技科可將大量資料重整，不僅影響資料的駕馭，調整研究的方法，未來對學術研究的貢獻，將隨著數位科技的不斷改進而日益擴大，無可限量。