

※「數位時代的出土文獻」專輯※

淺談數位化時代的出土文獻 及相關資源資料庫

單育辰*

在計算機迅猛發展的幾十年間，絕大多數學者拋棄了紙上作業；與此相應，電子資料也被研究者越來越多地利用。現今時代，學者如果不能全面掌握數位化資源，在研究中自會處於劣勢。本文準備在此介紹一下出土文獻電子資料庫。不過，目前電子資源數量很多，筆者無法全部試用，現僅依筆者所知、所見、所用者分甲骨文、金文、簡帛、傳世文獻四個方面做一簡單介紹。有些舊有的網絡資源，以前常被使用，但如今似已失效，在本文中亦摒棄不談。

一、甲骨文

香港中文大學中國文化研究所中國古籍研究中心製作的「漢達文庫」¹中之「甲骨文」資料庫，收錄七種最主要的大型甲骨著錄，共收卜辭五萬餘片，須在線檢索，付費使用，此資料庫釋文後來被出版為《甲骨文校釋總集》²。目前此資料庫的釋文及綴合正由研究者重新審定，以後將推出新版。

社會科學文獻出版社於二〇〇八年出版「甲骨文合集 Unicode 釋文」、「東京大

* 單育辰，吉林大學古籍研究所講師。

¹ 網址：<http://www.chant.org/default.asp>。

² 曹錦炎、沈建華：《甲骨文校釋總集》（上海：上海辭書出版社，2006年），全書二十冊。該書釋文比起舊時的甲骨文釋文如《甲骨文合集釋文》、《殷墟刻辭摹釋總集》來說，略有進步，但進步並不大，並且近來綴合成果突飛猛進，此書由於時間限制，無法吸收。在文字考釋方面，對當時已經出現的古文字考釋的一些很好的意見沒有吸收，而是採用了比較通行（但是錯誤）的說法。

學東洋文化研究所藏甲骨文字 Unicode 釋文」、「英國所藏甲骨集 Unicode 釋文」、「懷特氏等藏甲骨文集 Unicode 釋文」、「蘇、德、美、日所見甲骨集 Unicode 釋文」、「花園莊東地甲骨 Unicode 釋文」、「甲骨文合集補編 Unicode 釋文」這七種甲骨文資料的光碟，其來源為「瀚堂典藏」所開發的「出土文獻數據庫」之「甲骨文分庫」。

中央研究院歷史語言研究所所有甲骨文數位典藏³，「主要處理對象是藏於史語所的甲骨文拓片，約有四萬餘件，可分為考古發掘甲骨之拓片及購藏所得甲骨之拓片兩類」。目前只能檢索拓片，須在線檢索，申請使用。

華東師範大學中國文字研究與應用中心製作「花園莊東地甲骨檢索系統」⁴，須在線檢索，免費使用。

二、金 文

香港中文大學中國文化研究所中國古籍研究中心製作的「漢達文庫」中之「金文」資料庫，須在線檢索，付費使用，資料庫釋文來源於《殷周金文集成釋文》，其書已由香港中文大學中國文化研究所出版⁵。

廣西教育出版社發行的「金文資料庫」單機版（版本 1.0，出版時間 2003 年），為華東師範大學中國文字研究與應用中心製作，收錄了《殷周金文集成》及此後所出的一些銘文，使用尚稱方便，但釋文中體現的學術觀點較為陳舊，不能反映出版時學術界金文研究的水準⁶。

中央研究院史語所金文工作室製作的「殷周金文暨青銅器資料庫」⁷，收錄

³ 網址：<http://rub.ihp.sinica.edu.tw/~oracle/home.htm>。

⁴ 網址：<http://www.wenzi.cn/huadong/index.htm>。

⁵ 中國社會科學院考古研究所編：《殷周金文集成釋文》（香港：香港中文大學中國文化研究所，2001 年），全書共六冊。該書釋文有一定水準，但比起突飛猛進的古文字考釋來說，不少釋文又顯得過時，並且，對當時已經出現的古文字考釋的一些很好的意見並沒有吸取，而是採用了比較通行（但是錯誤）的說法。

⁶ 此資料庫的紙本即華東師範大學中國文字研究與應用中心編：《金文引得（殷商西周卷）》（南寧：廣西教育出版社，2001 年）；及華東師範大學中國文字研究與應用中心編：《金文引得（春秋戰國卷）》（南寧：廣西教育出版社，2002 年）。

⁷ 網址：<http://www.ihp.sinica.edu.tw/~bronze/>。

《殷周金文集成》、《新收殷周青銅器銘文暨器影彙編》中的青銅器，須在線檢索，申請使用。

中央研究院的「青銅器拓片數位典藏」⁸，收錄了歷史語言研究所收藏的青銅器拓片，數量超過一萬件，且其中一些拓片不見著錄，須在線檢索，申請使用。

陝西省考古研究所吳鎮烽領銜的課題組與西安廣才科技公司製作的「商周金文資料通鑒」單機版，收錄金文拓片一萬六千餘張，資料較全面，且經常更新。

三、簡 帛

香港中文大學圖書館製作的「郭店楚簡資料庫」⁹，資料來源於《郭店楚簡研究·第一卷·文字編》¹⁰，在線檢索，免費使用。

中央研究院歷史語言研究所文物圖像研究室製作的「簡帛金石資料庫」¹¹，收錄郭店、包山、曾侯乙、睡虎地秦簡、馬王堆帛書、銀雀山漢簡、居延漢簡、敦煌漢簡、兩漢鏡銘集錄、漢代石刻集成、漢印文字徵及補遺等六十餘種資料，在線檢索，免費使用。

香港大學中文系、武漢大學簡帛研究中心製作的香港大學中文學院「楚系文字字形、辭例數據庫」，收錄郭店、上博一、上博二、包山、望山、九店、長臺關、曾侯乙等十二種楚簡資料¹²，在線檢索，免費使用。此外，武漢大學簡帛研究中心亦有「中國古代簡帛字形、辭例數據庫」¹³，其程序與前者一致，但多收錄有上博三至上博七、新蔡葛陵等楚簡，還有睡虎地秦簡、張家山漢簡，數量比前者多，在線檢索，「中國簡帛文獻信息國際合作計畫」成員可免費使用。

上海教育出版社發行的「戰國楚文獻檢索系統」（版本 1.0，出版時間 2003 年），為華東師範大學中國文字研究與應用中心製作，收錄郭店、上博一、上博

⁸ 網址：http://ndweb.iis.sinica.edu.tw/rub_public/System/Bronze/Search/Search.jsp。

⁹ 網址：<http://bamboo.lib.cuhk.edu.hk/>。

¹⁰ 本書為張光裕主編：《郭店楚簡研究·第一卷·文字編》（臺北：藝文印書館，1999 年）。因該書在《郭店楚墓竹簡》剛出版後即出版，故現在來看，有不少字詞考釋已經過時。

¹¹ 網址：<http://saturn.ihp.sinica.edu.tw/~wenwu/bib.htm>。

¹² 網址：<http://characters.chinese.hku.hk/index.zx.php>。

¹³ 網址：<http://www.bsm-whu.org/zxcl/>。

二、包山、九店、曾侯乙、子彈庫等十二種楚簡帛書以及楚系金文、璽印、貨幣中的資料。後來，華東師範大學中國文字研究與應用中心又推出「戰國楚簡帛文字典型形體檢索系統」¹⁴，是在「戰國楚文獻檢索系統」基礎上增補而成的，比原來多收入上博三至上博五、新蔡葛陵等楚簡，在線檢索，免費使用。

香港中文大學中國文化研究所中國古籍研究中心開發製作的「漢達文庫」中之「竹簡帛書」資料庫收錄睡虎地秦簡、馬王堆帛書、銀雀山漢簡、武威漢簡等十二種出土文獻資料，在線檢索，付費使用。

復旦大學出土文獻與古文字研究中心製作「上博簡字詞全編資料庫」及「清華簡字形辭例數據庫」單機版，使用方便，但尚在測試中，流傳不廣。

中央研究院歷史語言研究所製作的「漢代簡牘數位典藏」¹⁵，收錄中央研究院歷史語言研究所收藏的漢簡，在線檢索，申請使用。

香港中文大學圖書館製作的「走馬樓三國吳簡・嘉禾吏民田家莧資料庫」，在線檢索，免費使用¹⁶。

依上所見，出土文獻的數位化研發有重複建設之處，不過，對於研究者來說，使用這些資料庫不嫌其複，它們收入的材料、圖像及釋文往往互異，研究者若能多加參照，正可以相輔相成。

四、傳世文獻

與出土文獻密切相關的是先秦兩漢傳世典籍，它們與出土文獻密切相關，研究者對這些資料庫的需求更大，使用更為頻繁，故此我們在此也做一介紹，因傳世文獻資料庫為數眾多，我們這裏僅擇要列舉一些有代表性者。

香港中文大學中國文化研究所中國古籍研究中心開發製作的「漢達文庫」中有「先秦兩漢」及「魏晉南北朝」資料庫，基本上把先秦兩漢、魏晉南北朝的所有古籍悉數收錄，加以標點，校勘精良，檢索較便利，但須在線檢索，付費使用。另，「漢達文庫」又有「類書」資料庫，收錄自魏晉六朝起以迄明清所有類書文獻。

¹⁴ 網址：<http://www.wenzi.cn/shuzi/czjs.asp>。

¹⁵ 網址：<http://rub.ihp.sinica.edu.tw/~woodslip/index.htm>。

¹⁶ 網址：<http://rhorse.lib.cuhk.edu.hk/>。

陝西師範大學歷史文化學院袁林、張宇等人開發「漢籍全文檢索系統」，有單機版，筆者所見最高版本為第四版，收錄從先秦到明清的常用古籍，本數據庫所用為簡體字，且校對有一些問題，有缺字現象。對於經常檢索先秦兩漢文獻的學者來說，不如「漢達文庫」方便。此檢索庫現已有網路版¹⁷，處於測試階段，不支援單字符檢索，免費使用。

中央研究院漢籍電子文獻瀚典全文檢索系統中的「漢籍電子文獻資料庫」¹⁸，累計收錄歷代典籍四百六十餘種，在線檢索，免費使用。

北京國學時代文化傳播股份有限公司製作的「國學寶典」¹⁹，有單機版，目前版本為 v9.0，但售價甚昂；除此之外，還有在線檢索版，可以免費使用。另，該公司還推出「中國歷代基本典籍庫」，其中的「先秦兩漢魏晉南北朝卷」對於古文字專業比較有用。

北京時代瀚堂科技有限公司開發的「瀚堂典藏」資料庫，須在線檢索，付費使用²⁰。

北京愛如生數位化技術研究中心製作的「中國基本古籍庫」，號稱收錄古籍總量相當於三部《四庫全書》，但售價甚昂。其公司製作的「斷代史料系列」之「上古秦漢史料庫」（2009年），對古文字專業來說較有用。其公司另有「搜神」和「典海」兩個數據庫網路版²¹，似正在調試。

此外，《四庫全書》有香港迪志文化出版有限公司製作的「《文淵閣四庫全書》電子版」單機全文檢索版（目前最新版本是 3.0 版）；濟南匯文科技開發中心製作的《四庫全書》單機原文電子版（武漢大學出版社 1998 年出版，只有圖像）。

《古今圖書集成》有廣西金海灣音像出版社和廣西師範大學出版社（只有圖像，1999 年）；臺北故宮博物院、東吳大學研製、漢珍數位圖書公司（2004 年）等發布的幾種單機版本。

《四部叢刊》有北京書同文數位化技術有限公司製作的單機版（2002 年、

¹⁷ 網址：<http://guoxue.zgwww.com/search.php>。

¹⁸ 網址：<http://hanchi.ihp.sinica.edu.tw/ihp/hanji.htm>。

¹⁹ 網址：<http://www.gxbd.com/>。

²⁰ 網址：<http://www.hytung.cn/Default.aspx>。

²¹ 網址：<http://www.er07.com/>。

2009年），另有網路版²²，須付費檢索。

《中國歷史地圖集》有中國地圖出版社製作的單機版。

以上僅就筆者所知而列，由此可見中國開發的出土文獻與漢籍資源已非少數，雖然由於各種原因，有些研發者沒有取得預期的經濟效益，但他們對於中國古代文史研究的貢獻，值得我們深深地景仰，他們的努力與成就，必將永載史冊。

當然，我們依據現有的情況，也可以提出一些建議：

（一）建議以後所有出土文獻資料必須附有以光碟、隨身碟或其他材質儲存的電子圖像，現在對出土文獻的研究越來越深入，僅僅靠拓片、黑白圖像來研究已經不夠用了。甲骨文不僅需要拓片，還需要彩照，並且甲骨文背面的拓片、彩照也是非常重要的；金文需要拓片、彩照、器形圖像並存；竹簡不僅要有彩照，還要有背面的彩照。並且，這些出土文獻既需要等比例照片，更需要大於原大數倍的高清圖版。如果全憑紙質印刷，一定會增加巨額成本，不僅出版社難以承擔如此高昂的費用，也很少有讀者能承受得起。更重要的是，以紙張為媒介發表的圖像，經過數次轉手，必不如電子版清晰。在數位化如此發達的時代，還單純地以紙張為唯一發表的載體，實為不智之舉。就以竹簡來說，發表者將來在紙質媒介上可以僅刊登竹簡原大正面彩照，但在光碟中，則應附有竹簡原大正反面彩圖及數倍於原大的竹簡正反面彩圖。據稱中國國家圖書館將出版《國家圖書館藏甲骨文全編》，把館內所藏的甲骨文悉數收錄，不但將出版紙本書，還附以所有圖像的數位化光碟²³。在國圖的「甲骨世界」這個資料庫裏，已經把國圖甲骨陸續上網，現在公布了四千餘件甲骨文的拓片及彩照²⁴，圖像十分清晰。如果出土文獻資料所有者都能以此形式來發表的話，無疑是古文字研究者的福音。

（二）在線檢索的資料庫並不方便，一個是不少研究者上網網速並不會太快，檢索一個詞會花費不少時間；另一個是研究者不見得隨時隨地能上網。還有一個重要的原因是資料庫開發者如果以後無力承擔租賃或維護伺服器的費用，則線上檢索資料庫必將隨之消亡，現在很多舊時常用的線上服務已經不能使用，就是一個很好

²² 網址：<http://guji.unihan.com.cn/Permit/>。

²³ 見〈國家古籍整理出版「十一五」（2006-2010年）重點規劃〉，《古籍整理出版情況簡報》，2006年第4期。

²⁴ 網址：<http://res4.nlc.gov.cn/home/index.trsf?channelid=10>。

的例子。值得慶幸的是，有不少門類的出土文獻及傳世文獻都有單機檢索版，但甲骨文仍然缺乏可以利用的單機檢索軟體，這對於甲骨文研究者來說，是個非常大的缺憾。現今很多出土文獻的數位化工程都能得到國家資金的資助，我們設想，如果有研究機構花費一些精力，找專業人士把甲骨文檢索製成單機版，以較低的價格出售，恐怕不是什麼難事。

（三）現在仍缺乏戰國文字的全面綜合的數位化檢索，也缺乏戰國文字的分國分域數位化檢索，以後出土文獻的數位化工作可以朝這個方向努力。

（四）《續修四庫全書》、《四庫全書存目叢書》、《四庫禁燬書叢刊》、《四庫未收書輯刊》還缺乏可全文檢索的電子版。當然，這些書內容極多，且書中字跡遠不如《四庫全書》工整，尤其是《存目》、《禁燬》這些書，漫漶不清之處不在少數，把它們轉化為文本，在技術上實在令人生畏，但希望國家能出資扶植一些數位技術公司，即使不靠技術而靠人力，也爭取在較短的時間內，把它們全部文本化，真真正正地為子孫留下一筆寶貴的文化遺產。

關於出土文獻及相關資源的電子資料庫我們就介紹到這裏，我們對軟體的開發、建構很不瞭解，本文僅僅從個人角度對出土文獻的數位化提出一些建議，其中不妥之處，在所難免，還望大家多多加以批評。

本文得到了馮勝君先生、程少軒先生的指教，特此致謝！