

學術出版數字化：未完成的轉型

——以期刊與知網關係為中心*

桑海

[提 要] 知網的目標和定位在發展變化,而其理想化的目標和數據庫的基本屬性卻貫穿其中。知網主導的數字出版是一種數據庫出版,形成了將期刊收入全文數據庫並進行網絡傳播的知網模式。知網模式既是特定環境的產物,也是知網基於數據庫邏輯的選擇。數據庫出版和紙本出版的不同框架,使知網與期刊之間存在視差,造成了數據庫出版的雙重性和雙主體現象。但由於知網和期刊都難以成為數字學術出版真正的主體,導致了數字期刊主體事實上的缺位,知網模式中的數字期刊更像是紙本期刊在數據庫中的鏡像。只有找到能激發數字時代學生產和學術交流潛能的新出版模式,數字學術出版的主體才會誕生,學術出版數字化轉型才能真正完成。

[關鍵詞] 學術出版數字化 學術期刊 知網模式 數據庫出版 學術交流 知識共享

[中圖分類號] G237 [文獻標識碼] A [文章編號] 0874 - 1824 (2023) 03 - 0135 - 11

引言：從“知網模式”說起

有關學術出版數字化轉型的探討方興未艾,學術期刊在目前的學術出版中處於核心位置,因而相關研究尤為“多姿多彩”。^①近年來,朱劍反思了中國學術期刊數字化轉型問題,其中對“知網模式”的批判引人矚目。所謂知網模式,指的是知網在其“光盤版電子雜誌社”時期設計並實行的一套數字出版和傳播的新模式。^②在這種模式中,期刊編輯部只需在期刊排版定稿後將電子版提交知網,以單篇文本形式匯入知網數據庫及數字平台,就可以實現期刊的數字出版。

知網模式出現已經近 30 年,那麼中國期刊數字化轉型到底有沒有完成呢?有人對此持肯定態度,認為知網“中國知識資源總庫”的初步建成“意味著我國學術期刊數字化、網絡化建設的目標基本實現”。^③朱劍雖然承認知網模式是當前中國學術期刊數字化的主流模式,甚至是唯一模式,但對其基本持否定態度,因為它看似數字時代的新事物,但卻是一種面向過去的設計,使紙本時代構建的學術期刊體制得以延續,所以只能是一種過渡性產物,“走出知網模式才能贏得未來”。^④知網模

* 本文係清華大學人文學院劉石教授主持的國家社科基金重大項目“基於大數據技術的古代文學經典文本分析與研究”(項目號:18ZDA238)的階段性成果。

式的出現有其特定歷史環境,即中國期刊不成體系、分散布局的刊號配置方式,就算沒有知網,這一模式的翻版也會出現。因此,中國期刊的數字化轉型,並不是簡單地將數字技術引入出版領域就能實現,而是一場整個行業的革命,需要對圍繞期刊的一系列關係進行重構,因而“轉型仍然在路上”。^⑤這種批判無疑極具啟示性,但仍存在一些令人困惑之處,比如:如何理解不斷發展變化中的知網?知網數據庫與出版之間有著怎樣的關係?知網模式中的數字出版主體是誰?中國學術出版數字化路在何方?本文將帶著這些疑問,對中國學術出版數字化轉型再做一番審視。

一、知網是什麼

2019年2月,北京電影學院博士、演員翟天臨在直播互動時問出“知網是什麼”,隨後他的博士學位以及博士後錄用通知書的真實性受到質疑,並牽連出一場學術不端調查。這一事件說明知網已經深深嵌入中國學術界,以至於不知道知網的人甚至沒有在學術界生存的資格。然而,要認真回答“知網是什麼”這個問題,也許並沒有那麼簡單。知網的發展大致沿著內容和平台兩條線索進行:在平台方面,從光盤版發展為網絡版,再從網絡版發展成CNKI的一系列數據庫和工具,進而走向國際化、智能化和大數據;在內容方面,則是從精選期刊到中國期刊,再到中國學術文獻,進而擴展為世界文獻。知網一直在變,因應國家規劃、國際動態的變化,從小到大一步步提出新的目標。

(一)光盤時期:產業模式初建

知網的創業,緣起於在新技術與學術出版之間發現的商機,數據庫本身就是一種科研與產業的結合。中國數據庫建設始於1980年代,^⑥在1983年清華大學就曾組織全國360多所高校研建“中國高等學校學報數據庫”。^⑦信息檢索也隨之開展,^⑧1990年代初出現了以光盤儲存和檢索信息的嘗試,當時只有文獻題錄檢索,還無法直接獲取全文。1995年,時任北京清華信息公司總經理的王明亮敏銳地看到光盤文獻數據庫蘊含的商機,提出以“版面顯示+全文檢索技術”為基礎的“中國學術期刊(光盤版)”(以下簡稱“光盤版”)設想,並被列入清華大學科研和產業化開發項目。

知網在光盤時期初步構建起數據庫的產業模式:上游是數據源——期刊編輯部;下游是用戶——圖書館等圖情機構;而自己作為數據庫運營商居於中心。1996年初,光盤版實現期刊全文檢索,又陸續以電子書號和電子刊號出版,光盤版電子雜誌社也隨之成立。在短暫的光盤時代,知網和期刊的關係比較和諧,多數學術期刊並不怎麼排斥這種新事物,因為學術期刊更看重影響力——論文文本的傳播,而非銷量——紙本期刊的傳播,光盤版似乎只是增加了一種傳播渠道。更深層的原因在於,知網最初的設計思路就是傳統與數字出版的融合運營,具有“融合出版”^⑨的色彩,基本不觸動現有出版體制和利益格局,在新技術與舊體制之間形成了共謀。1997年,光盤版與圖書館等機構合作,在全國建立300多家檢索諮詢站,由於既解決了紙本期刊佔用空間的問題,又能集中採購和便捷查詢,還可以向用戶收取查詢費用來增加收入,也受到圖書館界歡迎。

(二)網絡時期:數據庫出版的成熟

光盤是數據庫發展史中的一種過渡性介質,很快被互聯網所替代,憑藉光盤版時期的積累,知網順利實現了網絡化轉型,其數據庫出版也得以成熟。從國際趨勢上看,在知網開始創業的1995年,光盤數據庫已經到了產業生命週期的末期,光盤版剛剛發布就面臨網絡熱潮興起,向互聯網轉型已經迫在眉睫。1999年,“中國期刊網”開通,初步實現了從光盤數據庫到網絡數據庫的轉變。

如果說光盤版是期刊數字化的產物,那麼中國期刊網則是期刊網絡化的產物。儘管人們習慣於把網絡化視為數字化的一部分,但互聯網對數據庫的影響比數字化更為深刻。因為數字化是一

種虛擬化,只有與互聯網結合起來,其優勢才能被釋放出來。這使得數據可以實現空間上的瞬時流動,數據庫中的內容可以便捷地呈現在每個人面前,真正成為一種可以與紙本期刊抗衡的“出版”。從光盤版向網絡版轉變的過程中,知網的產品性質發生了一些重要的變化,比如由精選來源期刊轉變為能收盡收;呈現方式也由整本期刊變為單篇文獻,諸多原刊似乎被一個包羅萬象的數字大刊所替代。這或許是知網在轉型過程中,基於網絡數據庫邏輯所作出的合理選擇,也意味著知網數據庫出版的逐漸成熟。“中國期刊網”開通後,隨著原有的光盤檢索諮詢站升級為鏡像站點,知網與圖書館的關係也發生了微妙變化。圖書館漸漸從知網的合作者變成用戶,而知網在為圖書館提供數字化轉型解決方案的同時,也具有了某種跨界競爭的潛能,“本身也相當於一個專業文獻的數字圖書館”。^⑩這與期刊界所擔憂的知網成為一本包羅萬象的大刊,似有異曲同工之處。

(三) 平台時期:中國知識基礎設施

提出“中國知識基礎設施”的目標後,知網開始在期刊數據庫基礎上向縱橫兩個方向擴展,逐漸由網絡數據庫發展為網絡平台。1998年,世界銀行提出國家知識基礎設施的概念,次年中國政府也提出建設國家層面“知識基礎建設工程”的意向。知識基礎設施是一個新名詞,但並非新事物,隨著1665年《學者雜誌》、《哲學匯刊》等第一批學術期刊創辦,就逐漸形成了以學術發表為中心的知識基礎設施,“出版商、同行評議、書目信息、索引和摘要服務、信息檢索系統以及期刊影響因子等評價指標都是知識基礎設施的組成部分”。^⑪當然,這是一種印刷時代的知識基礎設施,在互聯網時代則需要新的知識基礎設施。有學者將知識基礎設施定義為“由人、構件和組織機構三種要素組成的文件網絡”,^⑫而數據庫的本質就是文件網絡,所以知識基礎設施很容易被想象成以數據庫為基礎的網絡平台。這對於數據庫行業無疑是一個巨大的機遇,知網敏銳地發現其期刊數據庫符合知識基礎設施的方向,便順勢提出“中國知識基礎設施”(CNKI)概念,並將其作為戰略目標和理想願景。2003年,“中國知識資源總庫”建設工程啟動,中國期刊網也更名為中國知網。^⑬

知網把CNKI的目標表述為:“通過建設一個國家級的可供知識生產、傳播、擴散和開發利用的知識信息資源系統與交互式網絡平台,實現知識信息的大規模整合與網絡化出版。使任何人在任何時候、任何地方都可以獲得所需知識,使知識信息在整合傳播、智能推送、網絡出版和互動應用的過程中不斷地增值,使知識資源系統成為高度社會化共享的文獻檢索、數字化學習與研究、知識發現與知識創新的工具。”^⑭其中幾個關鍵詞值得注意:“共享”,知識信息資源在全社會範圍高度共享;“獲得”,使任何人在任何時候、任何地方都可以獲得所需知識,知識可以自由流動;“交互”,建立知識生產、傳播、擴散和開發利用的互動平台;“增值”,使知識信息在整合與流動中提升價值,造福人類。由此可以勾勒出一幅令人憧憬的未來圖景,不乏理想主義氣息甚至知識共產主義色彩。

“中國知識資源總庫”是知網學術文獻的總樞紐和CNKI的數據庫基礎,後來的一系列平台都依此建立。因為有了CNKI的宏大目標,知網已經不滿足於期刊數據庫的定位,而要縱向打通知識生產、傳播、擴散與利用全過程,橫向擴展到博、碩士學位論文、報紙、會議論文、科學文獻、專利、圖書等不同類型文獻的數據庫。由於目標定位的變化,期刊數據庫由唯一產品變成了諸多學術文獻中的一種,知網也陸續推出採編系統、學術不端檢測系統、自動排版校對系統、期刊影響力排行榜等工具,呈現向出版流程各環節滲透之勢。由於規模化效應及國家工程的加持,知網在與期刊的互動過程中佔據了更為主動的位置,由其主導的“知網模式”也愈加牢固。

(四) 走向國際化與智能化:世界知識大數據

在全球化語境和複雜的國際形勢下,知網面對強大的國際競爭者,又提出了“世界知識大數

據”的設想。這一目標可以看作 CNKI 的國際化擴展,同時也包含了大數據和人工智能的技術意識,以及提升中國學術國際話語權的現實考量。進入 21 世紀後,隨著國際交流的加強,學術用戶開始不滿足於中國本土知識,希望獲得更多國際學術資源。知網因應這種變化,開始引進海外文獻資源,如與施普林格合作,整合雙方數據庫在知網發布。與此同時,知網也在推動中國學術期刊數據庫的國際化推廣,從 2010 年開始建設中國學術文獻全球營銷平台。在 2015 年前後,知網因應全球化趨勢和國際競爭形勢,確立了開發移動知識服務平台“全球學術快報”和建設“世界知識大數據”(WKBD)的目標,並從學術文獻拓展到覆蓋科學、社會、政務等領域,以期最終建成具有創新支撐意義的世界大數據。到 2021 年底,“世界知識大數據”已涵蓋中國大陸文獻近 1.8 億篇,境外文獻 3.4 億篇,並以平均每分鐘更新 166 篇、平均每小時更新 1 萬篇的速度不斷膨脹。知網在這些文獻基礎上,運用知識化標注、結構化與碎片化處理、知識挖掘等技術手段,形成了全文級、段落級和知識元級三個層次的大數據。全球學術快報是基於世界知識大數據搭建的平台,涵蓋全球學術期刊 7 萬種。在這一階段,知網也逐漸從數據庫出版商向知識服務商轉型。

(五)理想化目標與數據庫屬性

在知網發展變化的過程中,理想化的目標與數據庫的屬性一以貫之。從創業初期的產研結合,到網絡時期數據庫的包羅萬象,再到中國知識基礎設施和世界知識大數據,知網都為自己設立了某種理想化的目標。這種理想化的目標有正反兩方面的影響:一方面,建構了既有體制與國際趨勢的平衡點,並以此確定產業方向,使企業發展獲得外部支持和內部動力;另一方面,目標陳義過高,在實踐中難免會與眼前利益發生衝突。知網在歷史上曾遭到不少批評,^⑤近年來更是因為著作權和反壟斷等事件被推上輿論的風口浪尖。一個以知識共享和交流為理想目標的企業,卻頻頻與“侵權”、“壟斷”、“牟利”等負面評價發生關聯,無疑是令人遺憾的,這既是其數據庫出版模式趨於集中化的自然後果,也是理想化目標和商業化運作之間矛盾的體現。另外,目前知網的核心產品和經營基礎仍然是“中國學術期刊網絡出版總庫”及其擴展版本“中國學術文獻網絡出版總庫”,那些令人眼花繚亂的諸多知識庫、工具和平台等,都是從這一基礎上衍生而來。因此,在理解知網時,應當看到其最基本的屬性仍是數據庫,而其最重要的數據來源就是中國學術期刊。

二、理解數據庫出版

(一)在數據庫與出版之間

在知網數據庫與出版之間,究竟有著怎樣的關係?如果從印刷時代的出版觀念——將作品經過編輯、印刷、發行等環節實現傳播來看,光盤版或知網或許很難被視為出版。但“出版”這一概念也是在不斷游移和變化的,印刷時代的出版概念已經趨於終結,而數字時代的出版概念仍在形成之中。按當下較有共識的看法,出版是一種將作品通過任何方式公之於眾的行為。邁克斯·巴斯卡爾認為在出版的中心有兩個活動——審核和推廣,因為它們解決了“將什麼公之於眾”的問題。他把出版描述成以審核和推廣為核心的“內容機器”,這是一個頗有德勒茲色彩的比喻,突出了出版在本質上是一種連接和運動,具有流動性、開放性和多元性的特質。在審核與推廣兩者中,審核偏重於內容,而推廣偏重於運營。相對而言,審核在出版活動中更具有內在性,因為如果內容未經審核而推廣,出版就變成了媒介。只要審核這一要素存在,編輯就會持續發揮作用,即便學術期刊消亡,編輯仍然有一席之地。^⑥當然,一旦審核由算法完成,編輯功能也有可能被人工智能取代,這樣出版就會從“內容機器”變成“內容算法”。^⑦隨著人工智能特別是 ChatGPT 等大模型為代表的生成

式 AI 不斷推進,出版可能還需要再重新定義,不過以“內容機器”來理解出版,仍然符合數字出版的現狀。

巴斯卡爾還強調“框架”在出版中的作用,所謂框架可以理解為媒介、渠道等要素的組合,內容在其中得以呈現和傳播。如期刊就可以看作由紙張、印刷、文本、學術規則、發行渠道等因素構成的框架。數字內容並非絕對自由,只有在計算機硬件、代碼、文本標記語言、傳輸協議、顯示機制等構成的新框架中才能傳輸和顯示。在數字時代,出版的框架沒有消失,只是從舊框架轉向了新框架,而框架的打造者就成為新型出版商。在紙本出版物與數據庫之間,並沒有不可逾越的鴻溝。例如,1928年出版的《牛津英語詞典》已是一部規模宏大的出版物,在多次修訂重編的過程中規模又不斷擴大。1980年代,《牛津英語詞典》在IBM幫助下逐漸數字化,從2000年起在網上發布電子版並持續更新。如今的《牛津英語詞典》已經不再是書,而完全演變成了數據庫。數字化內容通過屏幕、代碼、數據庫軟件、信息管理工具、服務器以及顯示器重新構造,獲得了數據庫這一新的出版框架。^⑧牛津大學出版社也隨之從圖書出版商變成了數據庫出版商。框架具有某種主觀性和互動性的特質,它總是與人們看待世界的方式及媒介經驗有關,正如伽達默爾所揭示的,意義建立在讀者與文本的互動之上,框架建構起來之後,各種關係才能夠為參與者所理解,並在其中發生互動。

知網是學術期刊數字出版領域的框架建構者。中國期刊體制陳舊,布局分散,規模小,技術薄弱,在數字化趨勢面前無所適從,於是外來的數據庫商知網按其邏輯搭建了新的框架,並建立起媒體模式和商業模式,即知網模式。因此知網模式也可以理解為,用數據庫出版的方式實現期刊數字出版。在這一新框架中,推廣是知網的優勢自不待言,審核環節毫無疑問也是存在的,而且還包括紙本期刊與知網的雙重審核。知網建立起一個與期刊編輯部相對應的審核系統,有嚴格的編輯和審讀流程,會對原刊文獻進行檢查與加工,並在頁面上增加水印,甚至還會因政審或學術不端等原因剔除個別文獻。知網的審核不僅限於文本,還包括在元數據和全文數據等層面上的質量控制,其質量指標詳細而嚴格,如規定元數據“中文摘要”項的著錄文字錯誤率應小於0.1%,版面型全文數據庫的版面錯誤率應小於0.05%。^⑨因此,知網的數據庫顯然是一種出版,而且被很多人視為中國數字學術出版的主流形式。有學者指出,“中國學術期刊的數字化可分為兩種模式:一是以內容為主導的數字出版模式,二是以技術提供商和數據庫運營商為主導的數字出版模式,前者起作用的主體是期刊出版單位,後者是技術提供商和數據庫運營商。高校學術期刊大多採取的是第二種模式,借中國知網數據庫跨平台進行數字化傳播”。^⑩也有出版教材把數據庫出版作為數字期刊出版兩大類之一:“目前,數字期刊出版主要表現為網絡數字期刊出版和期刊在線數據庫出版兩大門類。”^⑪

(二)數據庫出版的邏輯

在知網發展過程中,一些方向選擇、設計理念、產品標準和技術開發起到了重要作用,我們可以嘗試從中探尋知網的數據庫出版邏輯。

1. 全文數據庫:數字鏡像的誕生

在王明亮對光盤版的最初設計中,全文數據庫是一個極具前瞻性的選擇,符合國際數據庫發展的潮流。以1980年MDC米德數據中心的NEXIS為開端,全文數據庫已經在各種文字型數據庫中異軍突起。“桌面出版”使很多文本誕生時就已經是數字形態,無需進行數字轉換,獲取全文資源變得更加便利,進一步促進了全文數據庫的發展。全文數據庫幾乎複製了紙本文獻的所有內容,在數據庫中建立起其數字鏡像——不但內容與紙本一致,連版式也可以藉由PDF之類頁面圖像軟件復現。隨之發生的是一種閱讀方式的轉變,即從紙本閱讀轉向數字閱讀。由於學術文獻的特殊性

質,即更偏重信息獲取而非閱讀體驗,這一轉變發生得更迅速和普遍。大多學術讀者真正關心的並不是文獻的具體類型、來源和存儲方式,而是其內容,知網所提供的全文數據已經可以滿足其所需。在光盤版時期,數據庫主要用於文獻檢索,藉助打印實現全文閱讀;而到了網絡時期,數據庫中的論文可以在任何時間地點下載和在線閱讀,全文數據的威力充分地發揮出來。在某種意義上,正是因為這種全文數據庫的方式,使紙本期刊在傳播中地位日趨式微,這也加重了期刊界對於出版主體的困惑。全文數據庫不但可以取代紙本期刊的閱讀功能,還能實現規模效應和數據關聯,支持檢索和深度開發,成為亞里士多德“整體大於部分之和”命題的現代例證。因此全文數據庫雖然看似紙本期刊系統的數字鏡像,但這只是其多個維度之一,潛在的價值遠不限於此。網絡時代有一條所謂“第22條軍規”——必須免費發布一些內容才能吸引讀者。知網的文獻題錄免費向全社會提供檢索服務,而獲取全文則要收費,這也印證了全文數據庫是知網的價值重心所在。

2. 查全查準:印刷穩定性的重建

在數字化、網絡化時代,已發布的文本是可編輯的,這意味著印刷穩定性的終結,而現代知識是建立在印刷穩定性基礎上的。讀者雖然更在意文獻的內容而非載體,但對文獻可靠性有很高的要求,這成為數字化出版需要面對的重要問題。為將印刷穩定性遷移到數字版本上,知網做了一系列的努力,例如:在邏輯鏈條上,數字文本作為紙本期刊的數字鏡像,其可靠性得到紙本期刊的背書,而紙本期刊的可靠性來自以刊號為核心的期刊管理體制、期刊評價體系、同行評議制度、期刊品牌及編輯的學術眼光等;在形式上,以PDF、CAJ等格式保留了紙本文獻的排版方式,具有仿真模擬的風格;在程序上,知網文獻上線前還要經過二次審核和數據標引,強調數據的標準化,這也增加了其穩定感。此外,“中國知識基礎設施”的宏大目標,“國家級”工程的形象,都加強了知網數據庫主流、官方的性質。知網對印刷穩定性的重建,還體現在“查全查準”之上。在“中國學術期刊網絡出版總庫”設計目標中,第一條就是“權威性全文文獻檢索工具”,其主要特徵是“期刊品種收錄齊全、文獻內容收錄完整、出版更新及時、信息加工規範、檢索功能先進,整個系統具有較高的查全查準性能”。²³知網從光盤版向網絡版轉型過程中,之所以從精選變為能收盡收,就是基於這一邏輯。也因為有了如此龐大的文獻數據,知網開發的學術不端檢測系統等產品才會受到普遍的重視。有些單位在認證學術論文時,甚至把知網可檢索看作比紙本期刊更可靠的依據,這種數字版本與印刷版本之間地位的翻轉耐人尋味,同時也證明知網已將印刷穩定性成功地遷移到了數據庫和互聯網上。

3. 元數據:標準化與關聯性

知網的元數據結構是參照都柏林核心元數據集(Dublin Core,以下簡稱DS)規範制定的。巧的是,DS出現的1995年正是知網創業開始之年。元數據是“描述數據模式或結構的數據”,²⁴藉由元數據,知網把文獻轉化成標準化和結構化的數據,並在不同顆粒度的數據之間建立關聯。從元數據的視角,我們可以更深刻地理解數據庫與期刊文獻之間的關係。數據庫中的期刊文獻,看起來只是紙本文獻的數字鏡像,實質上卻是一種由數據和程序生成的物體,許熠稱之為“數碼物”。所謂數碼物,“指形成於屏幕上或隱藏於電腦程序後端的物體,它們由受結構或方案管理的數據與元數據組成”。²⁵數字文獻這一數碼物就可以用元數據來描述和定位。CNKI數據庫採用的基礎元數據有15項:“題名、作者、主體、描述、出版者、時間特徵、資源類型、格式、標識符、來源、語種、覆蓋範圍、權限、全文文本。”²⁶值得注意的是,在這一元數據框架中,對DC的元數據標準進行了調整,將創建者和其他責任者合併為作者項,定義為“資源的創建者或與創建有關者”。²⁷基於全文數據庫產品的特點,為更好地描述和記錄文獻的全文文本,CNKI元數據擴充了DS所沒有的“全文”元素。²⁸

因此,儘管 CNKI 元數據與 DS 一樣,都包含 15 個元素,但其中還是有細微的差異,“全文”元素是其最大的特色。通過元數據的建構,被期刊視為自己生命的全文內容,在知網數據庫中變成了諸多元素中的一個。因此數據庫出版並不只是文本等內容的出版,而是包括內容在內的一系列數據的出版,其中包含了比內容更多的信息。元數據的作用,在知網節等技術中得到了更充分的表現。

4. 知網節:知識網絡與文獻宇宙

知網節可以簡單地理解成知識網絡中的節點:“以一篇文獻作為節點,提供文獻的題錄摘要以及知識元、參考文獻、引證文獻、相似文獻、讀者推薦文獻和作者與機構等擴展信息,支持擴展信息的鏈接和分類導航,通過概念相關、事實相關等方法揭示知識之間的關聯關係,具有支持知識獲取、學習、發現和管理的強大功能,是單篇文獻各種擴展信息的入口匯聚點和相關知識信息的鏈接點。”^②知網節的原理與 SCI 相仿,SCI“引文索引”打破了以往數據庫按主題或分類檢索文獻的常規做法,以一篇文獻節點,通過其引用和被引情況發現相關文獻並建立文獻網絡。^③知網節也是將參考文獻和引證文獻等鏈接到一起,形成內容關聯的知識網絡。在互聯網中,文獻通過超鏈接、元數據等方式實現關聯,建立起一個文獻宇宙。數據庫中的期刊與單篇文獻,其實只是文獻信息在不同顆粒度的呈現。期刊拆散為單篇文獻,使之可以自由流動,實現豐富的關係組合。期刊是少量文獻的單一關係組合,知網節則可以在海量文獻之間建立無窮的組合。而進入比文獻更細顆粒度的知識元層次,則可能帶來更多的知識發現,“採用智能搜索、智能聚類和分類等技術,實現了文獻內容的直接揭示和微觀整合”。^④知網節本身也是一種媒介,藉由知識內容的相關性與讀者檢索行為的相關性,文獻就能實現有效傳播,讀者可以藉由知網節的鏈接在文獻宇宙中漫遊。在這一過程中,人與人之間、知識與知識之間不斷建立起新的關係。正如許煜所說:“在某種程度上,可以說內容不是數碼物的關鍵問題;真正重要的是關係。在整個數碼物網絡中,它同時也是一個關係網絡。”^⑤知網節的視角,有助於傳統期刊人走出狹隘的內容中心主義,理解關係在數據庫出版中的生成性。

(三) 知網數據庫出版的兩重性

數據庫出版和紙本出版的不同框架,導致了知網與紙本期刊之間的存在視差。印刷是一種模擬技術,其內容與媒介並沒有明顯地分離。在數字化時代初期,克勞德·香農提出消息與渠道幾乎無關的信息理論,馬歇爾·麥克盧漢則認為一種媒介的內容是另一種媒介。他們的主張看似相反,一個把媒介看作幾乎可以忽略的“啞管道”,另一個則幾乎把媒介看成一切,但同樣都在強調媒介與內容的分離。隨著數字化的深入,尤其是互聯網的出現,這種將信息與渠道、媒介與內容分離的思想漸漸被人們理解。傳統出版者對數字出版的理解更多是香農式的,出版是一種內容產業,只是傳播載體由印刷品變成了數字。而在數據庫等科技公司看來,媒介遠比內容重要,內容只是一種數據資源,關鍵是在數據庫規模化和標準化的基礎上建立關聯,由此實現整體傳播和增值服務。他們的媒介觀可能是麥克盧漢式的:期刊數據庫是一種媒介,其內容是期刊;而期刊也是一種媒介,其內容是文章;文章還是一種媒介,其內容是碎片式的知識元。在知網看來,自己所從事的出版是作為數據庫整體的出版,期刊等出版物是其數據源,無需過多關心其具體內容;而在期刊看來,知網所做的幾乎算不上出版,而只是一種傳播,最多是二次文獻的數字出版,因為自己仍然牢牢掌握著對文本編輯審核的權力。這樣數據庫出版就具有了雙重性,即在數據庫層面的整體出版和文本內容層面的出版。然而,由於知網的數據庫是全文數據庫,所以知網數據庫出版的數據恰好是紙本期刊出版的內容,或者說這是對紙本期刊內容的數據庫出版——一種對於出版的出版,因此這兩個層面常常容易混為一談。從技術或媒介角度看,這是兩個不同層面的出版主體對同一內容的不同形式分

別進行的出版行為，兩種出版不能混同；但從版權的角度看，內容的版權歸屬是同一的，這又使兩種出版無法割裂。知網試圖基於前一種視角建構自己出版主體的地位，卻因版權的制約而陷於困境，這也使知網數據庫出版的主體問題更加令人困惑。

三、數字出版主體的困惑

(一) 數字出版定義中的主體分歧

有人曾這樣定義數據庫出版：“運用數據庫技術為主的多種技術，藉助計算機或者類似設備，將零散的獨立作品、數據或其他材料經過系統或有序的編排，存儲在磁、光等介質中，並最終形成可通過計算機或類似設備單獨加以獲取的數據庫產品的出版活動。”^②這是一種強調介質的定義，對數據庫出版的主體並未作限定。在數字出版的定義中，就已經出現了關於介質與主體的分歧，國內主流定義傾向於強調出版介質或技術的作用，^③而虛化了出版主體。這意味著把主體的位置留給能“運用技術”的人，如果傳統出版商不能“運用技術”，能“運用技術”的人就會趁虛而入。這或許與學者們觀察到的數字出版現狀有關。但也有學者特別強調出版的主體，認為數字出版是內容提供商將著作權人的作品數字化，經過對內容的選擇和編輯加工，再複製和傳送到多種載體上。^④這種定義突出了內容提供商的主體身份，並將內容審核置於數字出版活動的中心，隱含了對傳統出版商角色延續性的期待。有學者認為，在知網模式中傳統學術期刊和知網兩個出版主體並存，知網充當了數字出版的唯一主體，期刊雖是紙本期刊的出版主體，卻被排除在數字出版之外。^⑤

(二) 國際數據庫出版中的主體

若要理解知網在數據庫出版中的主體特質，有必要將其放入國內外數據庫出版的比較視野中考察。有研究者把國際數據庫發展史分為三個階段：一是由政府主導、以磁帶和磁盤數據庫為主的階段（1951~1980年）；二是由營利性組織主導、以光盤數據庫為主的階段（1980~1995年）；三是國際性數據庫出版商主導、網絡數據庫為主的階段（1995年以後）。^⑥從介質角度，可以將這三個階段命名為前光盤時期、光盤時期和網絡時期。國內數據庫發展與此類似，但時間上有一定滯後性。在前光盤時期，數據庫的生產與服務分離，數據庫生產者通常將數據庫出售給數據庫服務商，由數據庫服務商加工處理後發行或運營。到了光盤時期，商業性數據庫日趨增加，越來越多數據庫生產者傾向於直接發行或提供服務，形成了生產與服務合一的數據庫集成商。而傳統出版機構以及其他組織將積累的內容資源提供給集成商，成為數據庫的內容提供商。隨著數字技術和互聯網的迅猛發展，部分擁有內容資源的傳統出版商發生了數字覺醒。許多國際出版集團在保持原有出版方式的同時，紛紛試水數字出版，其中也包括數據庫出版，一些大型跨國出版傳媒集團轉型為數據庫出版商。愛思唯爾就是集內容提供商和數據庫集成商身份於一體的典型，目前已建立起集內容生產、傳播、數據檢索為一體的大型數字化內容資源平台，旗下擁有多家出版社和期刊社，如《柳葉刀》、《細胞》以及 Science Direct 數字期刊平台，還將海量的自有內容資源集成為 Scopus 等數據庫。

(三) 知網：有缺陷的數據庫出版主體

在中國，知網、萬方和維普等數據庫出版巨頭都是有技術背景的科技公司，傳統出版機構在數據庫出版領域的參與度不高，這造成了內容提供商與集成商的分離——科技公司並不擁有內容資源的版權，只是集成傳統出版機構的內容；傳統學術出版商尚未發生數字覺醒，或者是受限於環境無法施展，未能演化出愛思唯爾那樣的數字出版巨頭，只能扮演數據提供商的角色。國際大型數據庫出版商是真正意義上的“內容機器”，同時擁有印刷時代出版業的核心資源——有著作權的內

容,和數字時代的核心資源——先進數字技術,因此可以從源頭上爭取高水平作者,開發多種形式的數字出版和增值服務,在全球市場上形成強大的優勢。與此形成對照的是,由於缺乏自有內容,中國數據庫出版商在數據庫出版中的主體地位是有缺陷的,只能算是瘸腿的“內容機器”。常為著作權所困就是其典型症候之一,著作權是印刷時代遺留下來的問題,因為知網遠離內容生產過程,不與論文作者直接產生聯繫,加之作者未能分享到應得的利潤,因此糾紛不斷,這甚至成為知網的“阿基琉斯之踵”。數據庫出版商與內容作者和最終用戶之間存在隔閡,數字出版的生產和傳播鏈條沒有真正打通,而不能真正參與到內容生產過程中的出版主體,其主體性總是可疑的。

(四)數字期刊:無主體的出版

儘管不夠充分,但知網作為數據庫出版主體的身份還是明確的,期刊在紙本出版中的主體地位也沒有動搖。但數據庫中的數字期刊出版的主體究竟是誰?光盤版是一種有刊號的數字出版物,其載體是有形的,更新和流通方式與期刊出版也有相似之處。由於全文數據庫與紙本期刊的內容完全相同,又具備了數字期刊的形式,因而常被看作數字期刊出版的主體。知網並不否認期刊編輯部的出版主體地位,在其數據庫設計目標中也並未包含數字期刊出版,^⑤但這不意味著知網未曾作出成為數字期刊出版主體的努力。事實上,經過一系列的產品設計和模式構建,知網曾經非常接近這一主體位置。但由於知網沒有真正的自有內容,而修訂後的著作權法,又使其出版主體地位難以獲得體制的承認,以致在最後一步功墮垂成。對於數字期刊出版主體這一身份,期刊編輯部卻是敬謝不敏,因為畢竟自己似乎並沒有作出什麼改變,如果這樣就完成了數字化轉型,成功未免來得太容易;同時也基本沒有得到數字化的紅利,幾乎還是數字出版的檻外人。可見,無論知網還是期刊的主觀意願如何,都没能成為得到公認的數字期刊出版主體。於是,數據庫中的數字期刊事實上成了一種沒有主體的數碼物,期刊和知網分別負責其內容審核和數字傳播,但都不覺得自己能真正掌控它們。在知網的數據庫出版中,看似有兩個主體,但實際上主體的位置卻是虛置的。從這個意義上,甚至可以說在知網模式中,數字期刊並不存在,在數據庫中看到的大約只是紙本期刊的鏡像。

四、結語:如何走出主體困境

在知網模式中,數據庫中的數字期刊並沒有得到公認的真正主體。那麼,如何才能解決數字期刊出版的主體困境?作為數據庫出版主體的知網和作為紙本期刊出版主體的期刊,分別有各自的優勢:知網擁有技術和數據資源,使知識可以在數據平台上流動和關聯;期刊擁有刊號、品牌、編輯力、作者資源和學術敏感度,與知識生產的現場保持著密切的互動。同時二者也有各自的局限:知網並沒有真正參與學術生產,與學者缺少互動,其產品模式也未脫光盤和互聯網 1.0 時期遺留下來的數字化早期思維;期刊則固守印刷時代的思維模式,未能充分理解新技術為出版帶來的無限可能性。真正的數字學術期刊,呼喚一種新型的主體,目前的知網和學術期刊都難以勝任這一角色。

有學者提倡的“關係性生存”對打破數字學術出版的主體困境或許有所啟示。在西方思想史上,所謂主體是以自我為中心建立起來的有界存在,而關係性生存則建立在對這種主體的反思之上,強調走出自我而融入關係的流動之中,進而創造新的生成物:“一種新的對人類行為的解釋正在出場,這種解釋將聯合行動的匯流(co-active confluence)置於關切的中心”,“生成性關係過程刺激了意義的擴張和流動,它們可能成為最終決定我們未來的福祉的關鍵”。^⑥如果知網和學術期刊困守在各自的主體視域之內,一旦由刊號、法律、政策、語種、國界等構成的行業屏障發生變化,就難免出現生存的危機。只有擺脫自我中心的視域限制,充分利用各自的優勢,聯合創造能激發新知識

和新型知識形式的新型出版平台,新主體才有可能在學術與技術、內容與媒介的碰撞交融中誕生。

作為數據庫型出版主體的知網應該意識到,在數字學術出版中最重要的價值取向是促進學術流動。^①在數字化的世界中,開放和連接會產生更大價值,限制性做法會破壞知識流動和創造性。如果數據庫出版商人為設置各種限制,將文本鎖在人工孤島,就會危害整個學術生態系統。因此,包括開放獲取(Open Access, OA)、開源代碼(Open Source, OS)在內的開放學術成為一種趨勢。在知網針對反壟斷處罰提出的整改措施中,有一條“強化平台公益屬性”,這或許意味著今後知網有可能會開闢某種 OA 學術出版平台,甚至其海量數據也有可能與學術和管理機構實現數據共享。“善的經濟是互利,不是競爭。競爭帶來寡佔與壟斷,不利全體社會共同利益”,^②若知網能夠在互利與競爭、企業效益與社會利益之間的關係上進行積極反思和調適,向國企的社會責任和自己的理想目標回歸,或許會打開更廣闊的發展空間。在數字出版領域,知網應當開發更多知識共享與流動的方式,在學術生產前沿與學者和各類知識創造者充分互動,促進開放學術和數據學術的發展,形成新思想和新方法的匯流,進而激發人文社科領域知識生產方式和學術交流模式的創新。

作為內容型出版主體的學術期刊則應看到,在關係的時代,“內容為王”的口號已經不再適用,內容在出版中的地位將由關係接手。“這並不意味著我們不需要內容,而是形式使內容呈現為不同的關係。因此,內容的主要功能是形成用於產生關係的資源。”^③比如,在元數據中,文本不過是眾多元素中的一種,內容變成了數據的一項。而在知網節中,藉由元數據的內在關聯性,可以在文本與文本之間、人與人之間建立多元的關係網絡,內容的重要性在於其可發現和可連接,能夠成為建立關係的節點。數字出版的內容本身也在變化,比如數字人文發表的成果可能並非文本,而是數據集、軟件或網站,甚至是面向機器而非人類的語義出版。這些都對印刷時代的學術出版概念形成了挑戰。值得欣慰的是,在知網模式之外,期刊也在積極嘗試以微信公眾號等方式,探索數字出版新的可能性。期刊應當把內容當做連接和激發各種關係的行動而非固定之物,因應技術和學術潮流的變化,想象和創造新的出版框架和模式,讓“內容機器”充分運作。

總而言之,知網和期刊需要走出各自局限,在相互理解的基礎上建立良性的互動與合作,^④在生成性的關係中共同摸索新型數字出版模式。從更長遠的眼光看來,在數據庫與期刊等不同出版主體之間以某種方式進行統合與重構,或許是大勢所趨。唯其如此,中國數字學術期刊出版的主體才會真正誕生,中國學術出版的全球競爭力也才有望隨之提升。當然,這涉及一些體制上的難題,需要有關部門和學術界、出版界共同努力,改變以期刊體系和評價體系為中心的學術生態——真正意義上的中國知識基礎設施。只有各方聯合行動,才有可能使我們的知識基礎設施實現“全社會知識信息資源共享”^⑤的目標。

①姬建敏:《新世紀二十年中國高校社科學術期刊研究熱點、問題與進路》,澳門:《澳門理工學報》,2022年第3期。

②朱劍:《中國知網與入編期刊及其作者關係十論——從趙德馨教授訴中國知網侵權案說起》,北京:《清華大學學報》,2022年第2期。

③⑩涂佳琪等:《中國知網歷史與發展研究》,廣州:

《圖書館論壇》,2019年第9期。

④朱劍:《構建互聯網時代學術傳播的新秩序——以高校學術期刊發展戰略為中心》,武漢:《武漢大學學報》,2016年第2期。

⑤⑮朱劍:《在路上:中國學術傳播數字化轉型三十年回望——以學術期刊的轉型為中心》,重慶:《理論與改革》,2022年第4期。

⑥1980年就建立了化工文獻庫,參見王軍:《我國數據庫產業發展戰略研究》,鄭州:《河南圖書館學刊》,2005年第1期。

⑦②⑨徐菊:《商業性文獻數據庫的營銷策略研究》,上海:華東師範大學碩士學位論文,2008年。

⑧以中國科學院文獻情報中心為例,1976年10月就試驗了QT-11計算機文獻檢索系統;1987年就開展了國際聯機檢索終端。中國科學院文獻情報中心網站,http://www.las.cas.cn/about/lasyg/

⑨“融合出版”強調傳統出版與新興出版的融合。參見《全國科技名詞委組織召開融合出版概念及定義審定會議》,全國科技名詞審定委員會網站,http://www.cnterm.cn/xwdt/tpxw/202201/t20220114_678525.html,2022-01-14。

⑩⑫克莉絲汀·L.伯格曼:《大數據、小數據、無數據:網絡世界的數據學術》,北京:機械工業出版,2017年,第31頁;第27頁。

⑬此後中國知網或CNKI逐漸成為這家企業的通稱,為敘述方便,本文有時也用“知網”作為光盤版及中國期刊網的通稱。

⑭⑲⑳㉑㉒《〈中國學術期刊網絡出版總庫〉術語》,《中國學術期刊(光盤版)電子雜誌社清華同方知網(北京)技術有限公司企業標準》(Q/CNKI101—2006),2006年。

⑮比如光盤版學術規範的制定推廣、期刊影響力排行榜的發布等,都曾引發期刊界的批評。

⑯桑海:《學術編輯:學術傳播新模式的靈魂——以“中國高校系列專業期刊”和“域出版學術平台”為例》,北京:《傳媒》,2016年第19期。

⑰⑱邁克斯·巴斯卡爾:《內容之王:出版業的顛覆與重生》,趙丹、梁嘉馨譯,北京:機械工業出版社,2017年,第245~246頁;第103頁。

⑲參見《〈中國學術期刊網絡出版總庫〉元數據結構及數據質量》,《中國學術期刊(光盤版)電子雜誌社清華同方知網(北京)技術有限公司企業標準》(Q/CNKI103—2006),2006年;《〈中國學術期刊網絡出版總庫〉全文數據質量》,《中國學術期刊(光盤版)電子雜誌社清華同方知網(北京)技術有限公司企業標準》(Q/CNKI104—2006),2006年。

⑳武京閔:《學術期刊數字出版模式探析》,澳門:《澳

門理工學報》,2014年第3期。

㉑萬安倫、呂建生:《數字出版導論》,北京:北京師範大學出版社,2021年,第96頁。

㉒⑳《〈中國學術期刊網絡出版總庫〉總則》,《中國學術期刊(光盤版)電子雜誌社清華同方知網(北京)技術有限公司企業標準》(Q/CNKI100—2006),2006年。

㉓㉔㉕許煜:《論數碼物的存在》,李婉楠譯,上海:上海人民出版社,2019年,第141頁;第141頁;第1頁。

㉖㉗《〈中國學術文獻網絡出版總庫〉產品標準》,《中國學術期刊(光盤版)電子雜誌社清華同方知網(北京)技術有限公司企業標準》(Q/CNKI125—2012),2012年。

㉘元素又稱為“項”,是對數據對象某一特徵的描述,構成元數據記錄的基本單位。DC的元數據集DC-MES本身也具有可擴展性和可修改性。

㉙㉚羅曼:《數據庫出版發展研究》,武漢:武漢理工大學碩士學位論文,2013年,第17頁。

㉛參見謝新洲:《數字出版技術》,北京:北京大學出版社,2002年;徐麗芳:《數字出版:概念與形態》,北京:《出版發行研究》,2005年第7期。

㉜祁庭林:《傳統出版該如何應對數字出版的挑戰》,太原:《編輯之友》,2007年第4期。

㉝參見《〈中國學術期刊網絡出版總庫〉總則》。

㉞肯尼思·J.格根:《關係性存在:超越自我與共同體》,楊莉萍譯,上海:上海教育出版社,2017年,第63頁。

㉟參見桑海:《平台重構與交流復興——媒介變革為學術期刊帶來新契機》,澳門:《澳門理工學報》,2018年第4期。

㊱何日生:《善經濟:經濟的利他思想與實踐》,北京:宗教文化出版社,2021年,第30頁。

㊲例如部分高校學報與知網合作創辦的“中國高校學報系列專業期刊”,參見桑海:《從網刊到新平台——“中國高校系列專業期刊”及其升級轉型》,江蘇蘇州:《蘇州教育學院學報》,2020年第1期。

作者簡介:桑海,《澳門理工學報》編輯,澳門理工大學講師,博士。

[責任編輯 劉澤生]