

数字版权视域下的数字人文应用平台 构建模式研究*

欧阳剑^{1,2}

(1. 上海外国语大学图书馆, 上海 201620; 2. 上海外国语大学数字学术中心, 上海 201620)

摘要: 数字人文应用平台是数字人文基础设施的组成部分, 其构建模式的选择直接影响人文数据及研究工具的使用方式, 相关研究对数字人文开发与应用具有重要意义。数字版权是目前数字人文应用平台构建模式的主要影响因素, 本文对数字人文应用平台模式进行论述与归纳, 并对数据开放型、数据封闭型、混合型3种主要的数字人文应用平台构建模式特点进行分析, 论述了3种模式应用平台的应用场景, 从数字人文应用平台构建模式选择实践来看, 混合型数字人文应用平台模式比较符合目前的使用现状。

关键词: 数字版权; 数字人文; 应用平台; 数字人文基础设施; 应用平台模式

中图分类号: G215 **DOI:** 10.3772/j.issn.1673-2286.2022.02.009

引文格式: 欧阳剑. 数字版权视域下的数字人文应用平台构建模式研究[J]. 数字图书馆论坛, 2022 (2): 60-66.

数字人文是一个将现代计算机和网络技术深入应用于传统的人文研究与教学的新型跨学科研究领域, 它的产生与发展得益于数字技术的进步及其在科学领域的普及应用, 近年来, 传统文史哲等人文学科和社会科学都在不同程度上开展了数字人文研究的探索。数字人文研究的兴起给人文学者带来了新的研究视角与思维模式, 也为传统的人文研究带来了新的方法、工具和平台。数据、研究方法、工具与平台是数字人文研究的重要组成部分, 数据是数字人文研究的基础, 基础数据库很大程度上缓解了人文研究数据缺乏的状况, 而数字人文研究方法主要通过应用工具及平台来实现, 应用工具及平台体现出数字人文研究的基本理论、方法与技术等。随着人文数据建设越来越受到重视, 基于数据融合的多维度数据应用平台应时而生, 传统数字化资源的建设也逐渐由资源库向数字研究平台转变, 以不断满足人文学者提出的辅助其研究的新需求^[1]。

近年来, 各种数字人文应用平台建设此起彼伏, 由于数字人文应用平台构建的理念不同, 出现了不同的应

用平台构建模式, 与传统应用平台不同, 数字人文应用平台集数据、方法及工具为一体, 其构建模式的选择也会给人文数据及研究工具的使用方式带来影响。因此, 数字人文应用平台模式及构建研究对数字人文资源的开发与利用具有重要意义, 同时对我国数字人文基础设施的建设模式选择也具有实际的参考价值。

1 国内外数字人文应用平台建设现状

在美术馆、图书馆、档案馆与博物馆 (GLAM) 领域, “基础设施”的提法由来已久, 数字人文基础设施是一种支持人文学科研究活动的基础设施, 是数字环境下开展人文研究的基本条件, 包括与研究主题相关的文献、数据、软件工具、学术交流和出版的公用设施及相关服务等^[2], 通过平台化的架构为人文学者提供各类量化分析工具与可视化数据服务。数字人文基础设施对于学者的研究与实践至关重要, 数字人文基础设施作为一种技术平台, 将工具、服务、资源和方法用于

* 本研究得到国家社会科学基金项目“图书馆古籍文献的数字人文开发与应用模式研究” (编号: 17XTQ003) 资助。

数字研究之中,通过标准化协议将异构的数字人文基础设施相互关联成本地节点,有助于在不同基础设施系统之上建立一个包容性的资源获取网络,从而实现人文数据的开放获取,由此可见,数字人文基础设施具有高连接性、标准化及易访问的特点^[3]。

数字人文基础设施的建设为数字环境下开展人文研究提供必备的基本条件,相关项目的建设如火如荼,涌现出一批优秀的数字人文基础设施。上海图书馆历史人文大数据平台以关联数据的方式向外公开发布了上海图书馆数字人文项目所组织的基础知识库(人、地、时、事、物)、文献知识库(家谱、手稿档案、古籍等)、本体词表,以及数字人文项目建设过程中所用到的各种数据清洗和转换工具^[4]。哈佛大学东亚语言文明系教授包弼德(K. B. Peter)负责的中国历代人物传记资料库(China Biographical Database, CBDB)项目提供了多个涵盖人物信息的关联数据集^[5],已成为历史领域数字人文研究的典范。类似的还有中国台湾地区“中研院”创建的基础地理信息系统“中华文明之时空基础架构”(Chinese Civilization in Time and Space, CCTS)^[6],该系统以中国地图为基础底图,整合了“中研院”的汉籍电子文献系统、清代粮价资料库、明清地方志联合目录资料库等专业应用系统或资料库。此外,还有陕西师范大学出版总社、首都师范大学张萍教授和西安云图电子信息有限公司合作打造的丝绸之路历史地理信息开放平台^[7],王兆鹏教授主持的“唐宋文学编年地图”项目^[8],等等。这些数字人文基础设施的建设极大地促进了数字人文的发展,丰富了人文学科研究的基础数据。

数字人文基础设施与传统资源库的本质差别在于:数字人文基础设施嵌入了数字人文学者需要的数据,提供人文学者需要的研究及分析手段、算法、工具,实现基础数据与人文学者的无缝连接,数字人文应用平台是数字人文基础设施的重要组成部分。近年来,数字人文应用平台也不断涌现,其典型代表有中国台湾地区“中研院”的数位人文研究平台^[9]、中国台湾大学的DocuSky数位人文学术研究平台^[10]、复旦大学历史地理研究中心的数字禹贡^[11]、Gale数字学术实验室研发的基于云服务的历史文献分析平台^[12]、HathiTrust研究中心(HathiTrust Research Center, HTRC)的“数据胶囊”(Data Capsule)项目等,一些专有数据库厂商如Jstor Constellate、ProQuest(TDM studio)等开发了文本和数据挖掘工具,使人文学者可以使用词频统计、

术语提炼及主题模型等工具对专有数据库的数据进行研究,进一步促进了数字人文研究的发展,给普通人文学者的研究带来了极大的便利。数字人文研究平台的建立使人文学者不必再完全依赖于信息技术专家,为人文学者提供了一个友好互动的数字人文研究环境,提供了个性化的服务、协同合作的机制以及开放的资源,很大程度上克服了人文数据匮乏的困难,使得人文学者自由地融合数字人文技术与方法进行相关人文研究,加速人文研究进程。

2 数字人文应用平台建设中的数字版权风险分析

作为数字人文基础设施的核心,数字人文应用平台与传统的数字图书馆服务平台的显著区别在于:数字人文应用平台以文本化、数字化的数据为主,并辅以相关的数字化研究工具等,为人文学者提供一个数字化的研究环境。总的来说,数字人文应用平台以数字化、数据化、文本化为主要方式对各种类型人文资料进行组织与重构,将图书馆、档案馆、博物馆、文化遗产机构等分散的数字化馆藏资源进行整合作为应用平台的基础数据,通过平台化的架构为人文学者提供各类量化分析工具与可视化数据服务。数字人文应用平台是在传统资源库的基础上发展而来的,保留着诸多传统资源库的特征,特别是在人文数据的组成方面更是以传统的特藏资源为基础,通过对传统文献进行标注、融合和重组等工作使传统资源转化为数据形式,从而实现人文数据服务。

数字人文中的人文数据建设主要有人文数据复原与人文数据重构两种形式^[13],即通过对传统文献资源转录、改编、转换、重组以及发布等一系列加工处理,并实现人文数据之间的关联,从而构建完整、权威的人文数据集,人文数据建设不仅包含数字化,还包含文本、图像、音视频的多角度、颗粒化深度标引与元数据描述、数据化、数据融合、知识关联等工作,而这些工作与著作权法中的使用权、复制权、改编权等密切相关。当属于著作权法允许的少量、适当引用他人作品,且在自主加工成数据的过程中投入智力劳动这一情况时,一般不涉及著作权问题,版权风险相对较低。然而著作权法对原始的、非结构化的数据的版权保护比较薄弱,因为一些原始数据可能不符合“原创性作品”的创造性要求,而通过对原始数据进行开发或加工产生

的分析数据或汇编数据才应当被视为著作权客体,现有的法律对数据的版权保护还不完善,在有些情况下,法院愿意将版权保护范围扩大到涉及数据且具有足够创造性的作品^[14],因此,在人文数据建设的过程中很容易因为版权不明确而产生纠纷。

数字人文应用服务平台是以“数据化”为主要方式对各种类型人文资料进行组织与揭示,其中的人文数据更多来自文献原始数据或汇编数据。人文数据除了来自传统的特藏资源外还有其他来源渠道,如自建特色数据、研究机构的开放数据以及数字人文中的众包数据等,这些来源的数据大部分具有明确的数字版权。尽管一些发布的数据已申明遵从知识共享许可协议,但使用者依然需要遵守作者的意愿,其中就涉及是否需要署名、是否可以被商用、能否修改后重新发布等问题。

与数字人文基础设施类似,数字人文应用平台也具有高连接性、标准化及易访问的特点,数字人文应用平台往往涉及数据开放获取及共享,在数据开放获取及共享的过程中也会涉及版权问题,可能会为数字人文应用服务平台带来侵权风险。由此可见,人文数据的创建与组织、数据使用方式、数据安全等成为数字人文应用平台建设中影响数字版权的重要因素。

数字人文应用平台对数字人文研究的实现具有重要意义,数字人文应用平台是数据版权的最后把关“人”,数据使用的合法性、安全性等都通过应用服务平台实现。数据版权与人文数据开放理念存在一定的矛盾,作为研究者来说,希望方便、快捷地获取自己需要的人文数据,实现内外部数据的整合与关联访问,得到“一站式”的数据服务,但这种高度开放性的数据服务增加了数据侵权的风险;而对数字人文应用服务管理者来说,数据版权则是主要考虑的因素。因此,双方利益平衡是数字人文应用平台版权保护的根基。

3 数字版权视域下的国内外数字人文应用平台构建模式

数字人文应用平台的建设旨在实现数据开放共享,为人文学者构建一个数字研究环境,使其专注于学术问题研究。人文数据已成为人文学科研究的关键生产要素和基础资源,数据也是数字人文基础设施建设的重要内容,因此数字版权成为各数字人文应用平台构建的核心影响因素,进而使得各平台数据的开放及使用出现差异。根据数据开放程度,数字人文应用平台

构建模式主要分为开放型、封闭型和混合型。

3.1 开放型平台

数字人文应用平台作为数字人文基础设施的重要组成部分,起到数据基础设施的作用,数据开放共享成为数字人文应用平台的重要特征之一,数据开放型数字人文应用平台成为最基本的构建模式。数据开放型数字人文应用平台以无版权或被授权的数据开放利用为目的,实现数据的广泛共享。首先,用户可以根据自己的需求将平台数据进行组合、重构、浏览及下载;其次,平台的数据对外开放,或提供相应的数据API接口供其他用户及平台使用;除此之外,也允许导入及调用其他平台的数据,连接及共享其他平台的数据,并利用平台工具进行阅读、分析及可视化利用。

目前,诸多数字人文应用平台呈现出数据开放型模式特征。中国台湾大学数位人文研究中心研发的DocuSky数位人文学术研究平台是典型的数据开放型平台,该平台基于个人化资料库构建及数字工具分析理念,致力于开放链接技术支持人文学者上传各种不同来源、格式相异的文本数据,并且学者可以利用平台所提供的各种数字工具满足自己的个性化需求。在DocuSky平台中,学者无须求助专业技术人员就可以自主构建个人的多功能云端数据库,可以自由选用丰富的工具与材料,在上传个人文本与权威文档的同时,还可以导入其他开放资料库的文本,如中国哲学书电子化计划Ctext^[15]、日本京都大学的汉籍文本Kanripo^[16]、中国台湾地区“中研院”史语所汉籍文本等权威文档等。

数据开放型平台侧重数字人文研究数据及工具的集成与共享,主要提供各种数据处理及分析工具,通过定义一套标准化的API数据访问接口或关联数据技术进行链接,具有高度的数据开放性及易访问性。同时,由于数据开放型平台可以引入外部数据进行分析,使得平台的研究方法及工具也具有极高的共享性,这都为学者的研究带来了极大的便利,因而数据开放型平台具有良好的开放性、公共性和可持续性。

3.2 封闭型平台

从版权角度来说,有的数字人文应用平台又具有一定的封闭性及排他性——既不开放数据,也不接收外部数据,这种数字人文应用平台形成一种数据封闭

型模式。数据封闭型数字人文应用平台是一种基于数字版权保护机制的封闭平台,其将版权(私有)数据封装到一个可控制访问权限的系统中,从而实现数据版权保护。数据封闭型平台通过软件、硬件系统来隔离用户与数据之间的直接联系,通过基于策略接口和网络访问的控制限制他人对版权数据的非法操作与访问,用户可以选择具有自定义功能的模块或工具用于数据分析。相对于数据开放型数字人文应用平台,数据封闭型平台侧重数据版权保护的运作模式,并以平台数据及分析工具为研究者构建一个数字学术研究环境。

数据封闭型数字人文应用平台模式是一种较好的平台数据版权解决方案,并在实践中被广泛应用。随着数字研究环境的发展,内容分析研究逐渐兴起,文本内容挖掘策略较好地解决了版权数据使用限制。HathiTrust是美国一个长期保存数字资源的公共平台,HathiTrust数字图书馆拥有1 500多万卷/册资源,但有900多万卷/册因受到版权限制,而不能被HathiTrust成员馆用户公开使用。为解决此问题,HathiTrust研究中心(HathiTrust Research Center, HTRC)一直希望创建一套能让这些受版权限制的资源更加开放地被学者使用的模式,为了满足数字人文的需要,HTRC提出了“非消费型研究”(non-consumptive research)服务理念^[17]。基于这种服务理念,HTRC为版权数据开发了封闭的应用环境^[18-19],这既为学者访问这些受版权保护资源开辟了新的访问方式,又尊重了版权限制。所谓“数据胶囊”,简单来说就是将受版权保护的数据通过技术手段封装起来,避免用户直接读取原始全文,再基于特定的算法对封装的原始全文进行计算分析,并返回分析结果供研究人员使用,这种运作机制打破了数字版权的屏障,满足了研究人员的数据需求,又较好地保护了原始数据的版权。

数据胶囊服务为研究人员提供虚拟计算及分析功能,研究人员可以根据需要选择适合需求的分析工具及分析数据集进行分析。数据胶囊支持研究人员灵活配置运算环境,将不同的运算机制嵌入数据胶囊中,用户可以在该运算机制下生成分析结果,目前数据胶囊支持越来越多的分析类型,并且内置各种通用的分析工具。此外,HTRC与Google实验室合作开发了“Bookworm”文本分析和可视化工具^[20],旨在帮助学者应对大规模的HathiTrust文本数据所带来的挑战,在尊重版权的基础上充分服务学者。Bookworm以一种简单而强大的方式对数字化文本库中的语言使用趋势进行可视化,通过一

组强化的基于内容和元数据的特性来支持数据的多面“切片和切块”,从而帮助学者更好地建立他们的工作集。HathiTrust与Bookworm可视化工具可供研究人员绘制HathiTrust语料库中的单词趋势图,并通过书目元数据对其进行搜索。Gale数字学术实验室也将Gale 1.7亿页、跨越500多年的原始档案文字识别数据与文本构建用于研究的语料库,并将挖掘和可视化工具整合在一个平台之下,为学者提供了可对历史、文化、社会、政治等众多领域数据进行分析与挖掘的人文计算工具^[21]。

数据封闭型数字人文应用平台侧重人文数据的版权保护,对数据的合理使用进行了严格的规范,在未被授权的情况下用户不能直接访问平台的数据,用户可以通过分析及挖掘工具获得分析及处理结果。数据封闭型数字人文应用平台对用户来说具有单向性,大多只能分析平台提供的数据集,不能链接分析平台外部的数据,因此平台分析工具及数据共享功能差,不利于人文数据及研究工具的开放及共享。

3.3 混合型平台

在实践中,数字人文应用平台中的数据组成形式呈多样化,既包含开放共享数据,也包含版权数据,因而需要一种兼顾这两种形式的混合模式数字人文应用平台。混合型平台介于数据封闭型与数据开放型之间,这种平台在以开放共享为主导模式的基础上对一些受保护的数据进行封装,借鉴“数据胶囊”式建设理念构建数据分析环境,通过建立合理的数据应用机制保护版权(私有)数据,从而以实现数据版权(私有)保护与数据利用之间的利益平衡,对于无版权限制的数据则开放共享,同时研究工具也可对应用平台的所有数据进行分析。

混合型数字人文应用平台注重数据的共享,以笔者开发的中国古籍基础数据应用平台^[22]为例,数据版权风险是古籍数字人文应用服务体系构建面临的现实挑战之一,也是影响古籍数字人文健康发展的关键因素。从理论上来说,古籍文献本身已超过50年的版权保护期,使用上无版权限制,但古籍数字人文应用平台的古籍数据更多来自古籍的标点、注释、匡正、补遗等整理作品,依据我国《著作权法》第十二条规定,“改编、翻译、注释、整理已有作品而产生的作品,其著作权由改编、翻译、注释、整理人享有”,按此项规定,如果整理的新作品具有独创性,则理应受著作权保护^[23]。此

外,古籍数据及知识的整理是一件费时、费力的工作,整理者对古籍数据及知识的整理付出了智力劳动,因而拥有所有权,也具有版权。就目前来说,应用于古籍数字人文平台的数据主要来源于传统数字资源的数据化,因此古籍数字人文应用平台需要具有完善的版权(私有)数据保护机制,从而对版权及私有数据进行保护。同时,古籍数字人文应用平台的开放数据服务对人文数据的互联互通具有重要意义,是推动人文研究创新的重要基础,需要考虑研究者的使用需求,需要选择合适的古籍数字人文应用平台构建模式,从而促进古籍数据在更大范围内共建共享。

综上所述,为了寻求古籍数据保护与数据开放之间的平衡关系,中国古籍基础数据应用平台采用混合型模式构建,混合型古籍数字人文应用平台中的数据来源渠道多元化,混合模式既对含有版权的数据进行了保护,又为人文学者提供了相对开放的数据服务应用环境,同时也实现了其他数据的开放与共享,有力地支持

了人文研究。

混合型古籍数字人文应用平台主要由数据中台、算法中台等组成(见图1)。数字人文研究加速了人文学科数据驱动型研究的发展,数据建设是数字人文的重要内容,通过建立古籍数据中台,聚合和治理跨域数据,从而驱动数据服务的转型,满足人文学者研究需求。古籍数据中台的核心思想是数据共享与数据版权保护,古籍数据中台作为一个数据共享的核心,在混合型古籍数字人文应用平台构建的过程中对有版权(私有)的古籍数据进行封装,通过将图像、文本、数据、知识等进行结构化处理,进而聚合跨域多源数据,解决人文研究中面临的古籍数据孤岛问题,建立数据挖掘、数据管理、数据利用与共享等机制,最终满足人文学者研究的多维数据需求,实现不同古籍应用场景的人文研究。古籍数据中台存储受保护的版权(私有)数据,也存储开放共享数据,从而实现数据版权(私有)与数据利用之间的利益平衡。

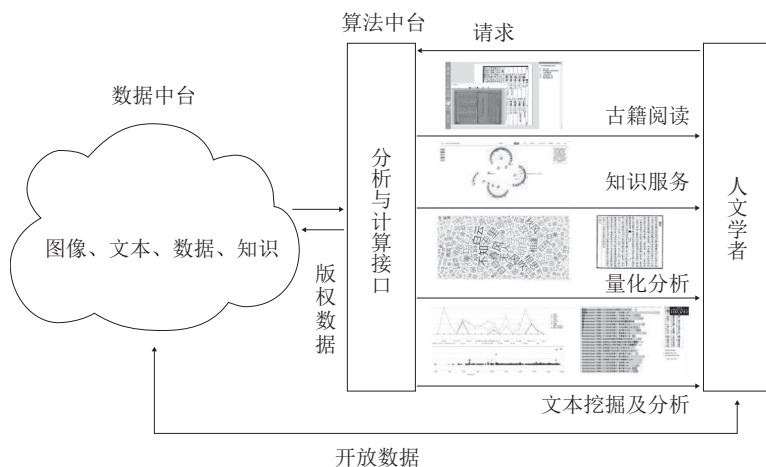


图1 混合模式的古籍数字人文应用平台框架

算法中台不但为人文学者提供各种古籍研究场景的分析方法及工具,也为人文计算提供高性能硬件计算服务,为人文学者构建一个流畅、稳定、可扩展的研究环境,提供更加个性化的服务,增强用户体验。算法中台将用户与易于使用的工具相连接,实现分析方法工具化、平台化,其接收用户的计算分析及数据调度请求,通过内置的文本分析和可视化工具等进行数据分析,并向用户返回计算及分析结果,而计算及分析与数据调度都由算法中台完成,用户根据自己的研究需要配置相应参数并选择对应功能模块即可完成分析及计算,为人文学者实现数据与应用的无缝对接。人文学者

既可以通过应用分析与计算接口对平台受保护的图像、文本、数据、知识进行间接访问,系统把分析及计算的结果返回给人文学者,避免了人文学者直接访问平台的版权(私有)数据;人文学者也可以直接访问平台的开放数据,分析与计算接口是联系古籍数据与人文学者的桥梁。

3.4 数字人文应用平台模式比较

数据开放型、数据封闭型及混合型数字人文应用平台模式各自的特点不同(见表1),侧重不同的应用场

景。从数据的开放性来说，数据开放型数字人文应用平台的开放程度最好，也是人文学者比较喜欢的模式，学者可以共享平台数据，人文学者参与程度高，但也因为数据的保护性差，学者上传个人数据的意愿降低，也限制了版权数据的发布与共享，从而限制了平台的数据类型和数量。数据封闭型数字人文应用平台则更强调对数据的保护，数据的开放程度不够，难以对数据进行共享，实践中也缺乏人文学者的广泛参与，目前，商业性数字人文应用平台多为数据封闭型模式，大多采用付费方式供用户使用。而混合型数字人文应用平台模式则兼顾前两种的优点，既可以对版权（私有）数据进行保护，又可以实现广泛的数据共享；既满足学者的需要，也符合数据管理者的需求；既能吸引人文学者的广泛参与，也对商业性数据参与具有吸引力。但从平台实现角度来看，混合型数字人文应用平台构建复杂程度也较前两种高。

表1 平台模式比较

平台模式	共享性	版权保护程度	开放性	学者参与度
数据开放型	高	低	高	高
数据封闭型	低	高	低	低
混合型	高	高	高	高

4 结语

数字人文应用平台是开展数字人文服务的重要组成部分，也是数字人文重要的基础设施，近年来各种数字人文应用平台建设不断涌现，从现有数字人文应用平台构建模式来看，数字版权是目前数字人文应用平台构建模式的主要影响因素，数字版权的授权与获取直接影响着人文数据及研究工具的使用方式与使用范围，也影响着数字人文应用平台构建模式的选择。目前的数字人文应用平台可分为数据开放型、数据封闭型、混合型3种，不同的构建模式其侧重点不同，适用于不同的应用场景，数据开放型侧重数据及工具共享，数据封闭型则侧重数字版权的保护，而混合型既对版权（私有）数据进行了保护，又给传统的人文研究带来数据、工具的共享。

从数字人文应用平台构建实践来看，目前数字人文应用平台的数据有3种类型：①以传统数字资源为基础，通过数据化的形式把传统数字资源转化为人文数据，平台对这类数据具有明确的数字版权；②人文学者

在数字人文的实践中加工、整理的数 据，这类数据属于私有数据，加工即整理者拥有所有权；③无版权的公共数据。由此可见，混合型数字人文应用平台是解决目前数字版权问题的首选构建模式。

本文从数字人文的人文数据版权角度出发对数字人文应用平台模式进行了研究，由于人文数据的组成复杂，其版权也存在特殊性与复杂性，对于版权风险的论述还不够具体，有待进一步对引发版权纠纷的因素进行进一步阐述，并提出应对措施。

参考文献

- [1] 卢彤, 李明杰. 中文古籍数字化成果辅助人文学术研究功能的调查 [J]. 图书与情报, 2019 (1): 70-79.
- [2] 刘炜, 谢蓉, 张磊, 等. 面向人文研究的国家数据基础设施建设 [J]. 中国图书馆学报, 2016, 42 (5): 29-39.
- [3] PAWLICKA-DEGER U. Infrastructuring digital humanities: On relational infrastructure and global reconfiguration of the field [J/OL]. Digital Scholarship in the Humanities, 2021 [2022-01-01]. <https://doi.org/10.1093/lc/fqab086>.
- [4] 历史人文大数据平台 [EB/OL]. [2021-10-15]. <https://dhc.library.sh.cn/>.
- [5] China Biographical Database Project (CBDB) [EB/OL]. [2022-01-29]. <http://projects.iq.harvard.edu/cbdb/home>.
- [6] 中华文明之时空基础架构 [EB/OL]. [2022-01-29]. <https://ccts.ascc.net>.
- [7] 丝绸之路历史地理信息开放平台 [EB/OL]. [2022-01-29]. <http://www.srhgis.cn/>.
- [8] 诗词地理 [EB/OL]. [2021-10-12]. <https://www.sou-yun.cn/PoemGeo.aspx>.
- [9] “中央研究院”数位人文研究平台 [EB/OL]. [2021-09-09]. <http://dh.ascdc.sinica.edu.tw/member/index.html>.
- [10] DocuSky数位人文学术研究平台 [EB/OL]. [2021-09-16]. <http://docusky.org.tw/DocuSky/ds-01.home.html>.
- [11] 数字禹贡 [EB/OL]. [2021-09-16]. <http://hgis.fudan.edu.cn/>.
- [12] Gale数字学术实验室 [EB/OL]. [2021-09-16]. <https://www.gale.com/cn/primary-sources/digital-scholar-lab>.
- [13] 欧阳剑, 彭松林, 李臻. 数字人文背景下图书馆人文数据组织与重构 [J]. 图书情报工作, 2019, 63 (11): 15-24.
- [14] U.S. Copyright Protections for Market Data [EB/OL]. [2021-10-20]. <https://www.natlawreview.com/article/us-copyright-protections-market-data>.

- [15] 中国哲学书电子化计划 [EB/OL]. [2022-01-31]. <https://ctext.org/>.
- [16] Kanripo 漢籍リポジトリ [EB/OL]. [2022-01-31]. <https://www.kanripo.org/>.
- [17] Non-Consumptive Use Research Policy [EB/OL]. [2021-10-16]. https://www.hathitrust.org/htrc_ncup.
- [18] MURDOCK J, JETT J, COLE T, et al. Towards Publishing Secure Capsule-Based Analysis [C] //2017 ACM/IEEE Joint Conference on Digital Libraries (JCDL). 2017.
- [19] HathiTrust Research Center Data Capsule Terms of Use [EB/OL]. [2021-10-18]. https://www.hathitrust.org/htrc_dc_tou.
- [20] Exploring the Billions and Billions of Words in the HathiTrust Corpus with Bookworm: HathiTrust + Bookworm Project [EB/OL]. [2021-10-23]. <https://ischool.illinois.edu/research/projects/hathitrust-bookworm-project>.
- [21] Gale Scholar系列讲座——Gale数字学术实验室：一个新的人文研究视角 [EB/OL]. [2021-10-23]. <http://www.library.fudan.edu.cn/2020/1116/c952a161836/page.htm>.
- [22] 中国古籍基础数据应用平台 [EB/OL]. [2021-10-25]. <http://121.201.35.124:88>.
- [23] 赵江龙. 馆藏古籍数字化版权保护问题及解决对策 [J]. 内蒙古科技与经济, 2015 (8): 141, 143.

作者简介

欧阳剑, 男, 1976年生, 博士, 研究员, 研究方向: 信息资源管理、数字人文, E-mail: oyjjj@163.com。

Construction Model of Digital Humanities Application Platform under the Perspective of Digital Copyright

OUYANG Jian^{1,2}

(1. Library of Shanghai International Studies University, Shanghai 201620, P. R. China; 2. Digital Scholarship Center of Shanghai International Studies University, Shanghai 201620, P. R. China)

Abstract: The digital humanities application platform is an integral part of the digital humanities infrastructure, and the choice of its construction mode directly affects the way humanities data and research tools are used, and related research is of great significance to digital humanities development and application. The study of digital humanities application platform construction mode is of great significance to digital humanities development and application. Digital copyright is the main influencing factor of the current digital humanities application platform construction model. This paper discusses and summarizes the digital humanities application platform models, and analyzes the characteristics of three major digital humanities application platform construction models: open data, closed data, and hybrid. The paper analyzes the characteristics of three major digital humanities application platform construction modes, and discusses the application scenarios of three modes of application platforms. From the perspective of digital humanities application platform construction mode selection practice, the hybrid digital humanities application platform mode is more in line with the current usage status.

Keywords: Digital Copyright; Digital Humanities; Application Platform; Digital Humanities Infrastructure; Application Platform Model

(收稿日期: 2022-02-04)