

# 图书馆数字特藏建设的紧迫性与发展路径研究

王彦力 杨新涯 冉蔚然

**摘 要** 在国家文化数字化战略的背景下,图书馆的数字化转型对数字特藏建设提出了新的要求。本文系统阐释数字特藏概念的内涵和外延,按内容和形式将数字特藏划分为七种类型:实体特藏数字版本、数字灰色文献、多媒体资源、公开的互联网内容、个人社交网络文书、科学数据、NFT 数字藏品。在调研国内外实体特藏数字化和原生数字特藏建设现状的基础上,总结图书馆数字特藏建设的不足:对数字特藏的整体重视程度和认识不够,没有形成图书馆行业完整、规范、普遍的数字特藏管理与服务模式,缺乏长期可持续的入藏计划和技术保障。建议:图书馆应以“当下即历史”的理念,以专题形式收集互联网内容、多媒体文献、科学数据和个人社交网络文档等未来可能成为特藏的原生数字内容;应制订完善的数字特藏建设规划,研发采集系统和管理平台,研制收集、整理和服务标准体系;重点要树立特藏观念、进行方法革新,依托特藏进行服务创新,加强数字特藏的版权控制与隐私保护等。图 1。参考文献 22。

**关键词** 图书馆 数字特藏 馆藏建设 数字化转型 文化数字化

**分类号** G253

## The Urgency and Development Path of Digital Special Collections Construction in Library

WANG Yanli, YANG Xinya & RAN Weiran

### ABSTRACT

Under the national cultural digitization strategy, digital transformation of library puts forward new requirements for the construction of digital special collections. At present, the library special collection resources are still mainly physical special collections, and the construction of digital special collections in most libraries is insufficient. The paper explains the connotation and extension of digital special collections, and divides it into seven types according to content and form: digital version of physical special collections, digital grey literature, multimedia resources, publicly available Internet content, personal social networking documents, scientific data and NFT digital collections.

After investigating the digitization of physical special collection and the construction of born digital special collection at home and abroad, the paper summarizes the shortcomings of digital special collection construction: the insufficient emphasis on the digital special collections; there is no comprehensive, standardized and general digital special collections management and service model in the library industry; the lack of long-term sustainable collection plan and technology support. Then, it explains the urgent need of libraries to strengthen the construction of digital special collections.

The paper suggests that governments at all levels should be encouraged to strengthen policy-level planning and

通信作者:王彦力, Email: wangyanli@cqu.edu.cn, ORCID: 0000-0002-8251-9089 (Correspondence should be addressed to WANG Yanli, Email: wangyanli@cqu.edu.cn, ORCID: 0000-0002-8251-9089)

support, and libraries need to accelerate the digitization of physical special collections, determine the construction of digital special collections with a focus on thematic construction, and advocate a personal submission system to implement the construction of digital special collections. And it proposes that a comprehensive digital special collections plan should be formulated, priority should be given to the construction of a digital special collections management system, adaptive changes to library management and service processes and entire sequential process and standards for digital special collections should be developed, suggesting that the concepts and methods of special collections construction in libraries in the process of digital transformation need to keep pace with the times, further promote the innovation of digital special collections services, and the relevant copyright control and privacy protection need to be properly addressed. 1 fig. 22 refs.

#### KEY WORDS

Library. Digital special collections. Collection construction. Digital transformation. Culture digitisation.

## 0 引言

2022年5月,中共中央办公厅和国务院办公厅印发《关于推进实施国家文化数字化战略的意见》,明确到“十四五”时期末,我国基本建成文化数字化基础设施和服务平台,形成线上线下融合互动、立体覆盖的文化服务供给体系,逐步实现国家文化大数据体系,其重点任务包括增强公共文化数字内容的供给能力,提升公共文化服务数字化水平等。同年10月,党的二十大报告再次强调“实施国家文化数字化战略,健全现代公共文化服务体系”。2023年2月27日,中共中央、国务院印发《数字中国建设整体布局规划》,提出“推进文化数字化发展,深入实施国家文化数字化战略,建设国家文化大数据体系,形成中华文化数据库”。在此背景下,图书馆应发挥资源优势、体系优势、专业优势和人才优势,在文化资源整合和计算利用、文化基础设施建设和提升服务、文化标准规范体系建设等方面落实国家文化数字化战略,其中数字资源建设是数字化转型的关键,也是图书馆向中华文化数据库输入的核心内容。

图书馆是保存人类文化和记忆的机构,收藏人类不同历史时期创造的各种载体的文献。当前,随着数字化转型的不断深入,越来越多的

文献以数字形式产生,通过不同载体和渠道传播。但是,目前图书馆入藏的数字资源仍以图书、期刊、学位论文、标准和专利等类型为主,对于网络文献、非正式出版物等数字灰色文献的收集整理和保存关注不够,对于大多数图书馆而言,馆藏数字资源还不够丰富,值得长期保存的特藏更是稀少。着眼未来,应以“当下即历史”的发展策略,重视数字时代和未来可能成为特藏的数字资源建设,逐步建立图书馆数字特藏体系。这里的“数字特藏”概念,是“数字”和“特藏”两个词汇的交集,“数字”为特藏的表现形式,“特藏”指文献资源的内涵价值。“数字特藏”概念的提出,以及对图书馆主动规划、寻访、整理、保存数字特藏等相关业务流程的梳理,对于加强实体特藏的数字化进程,加大力度入藏互联网内容、多媒体资源、科学数据等原生数字资源,具有理论和现实意义,也有助于图书馆在数字化转型和文化数字化建设中,与时俱进,充分发挥基本职能。

## 1 数字特藏的内涵与外延

早在2011年,我国就有学者提出了“数字特藏”的概念,即图书馆通过自建或者参与建设的具有自己独特风格的数字资源馆藏<sup>[1]</sup>。经过十余年的发展,图书馆已经逐渐认识到数字特

藏建设的重要性,但对于数字特藏的内涵与外延尚缺乏全面的认识和系统的阐释。本文在探讨“数字特藏”建设的紧迫性和发展路径之前,首先对“数字特藏”的概念进行系统梳理,以期为图书馆数字特藏建设实践提供指导。

### 1.1 数字特藏的概念

美国研究图书馆协会(ARL)和研究收藏委员会(RCC)将图书馆特藏定义为“以多种形式存在的馆藏(如稀有书籍、手稿和档案,以及许多其他类型的资源),是国家智力资本的重要组成部分;具有珍稀性、独特性、历史性的特点;可以支持教育机构的教学和研究任务,为国家和国际学术提供不可或缺的资源支撑”<sup>[2]</sup>。区别于普通馆藏,图书馆对特色馆藏的收、存、管、用等实行较为特殊的业务工作方式。在图书馆日常管理与服务业务中,特藏资源可根据载体形式分为“实体特藏”和“数字特藏”两部分。

“数字特藏”是以数字化为载体、具有严格版权控制的特色文献,由纸质特藏的数字化和具有长期保存价值的原生数字资源两部分组成,是数字时代图书馆发挥保存人类文化和记忆这一基本职能的重要馆藏对象。数字特藏的“数字”为表现形式,决定其具有特殊的建设、存储、使用等管理规范,“特藏”为内涵意蕴,表明其具有长期保存价值的内容信息。

### 1.2 数字特藏的主要类型

数字特藏的分类有多种方式,可以根据内容、载体形式、目标成果等进行分类,本文主要依据内容和形式进行分类。

(1) 实体特藏的数字版本。古籍、民国文献、其他珍贵文献等实体特藏的数字化成果是数字特藏的重要组成部分,其中古籍及民国文献的数字版本占比最高,国内数字特藏引起广泛关注的典型案例多源于此。如国家图书馆牵头并指导的中华古籍保护计划普查全国各级各类图书馆的珍贵古籍,综合性古籍特藏数字资源发布共享平台“中华古籍资源库”是该计划的

重要成果,截至2023年初,全国累计发布古籍及特藏文献影像资源达13万部(件)<sup>[3]</sup>。

(2) 数字灰色文献。灰色文献一般指由各级政府部门、学术机构、工商业界等推出的各类电子和印刷形式的资料<sup>[4]</sup>,多为非正式出版发行的文献,因其信息广泛、内容新颖、所涉种类繁多且各具特色而具有非常高的收藏价值。数字灰色文献则是非公开出版发行的原生数字文献,主要包括各类建设方案、宣传页、产品说明书、PPT课件、资料汇编、网络办公的公文、未经同行评议或编辑审校的预印本论文等,这是数字特藏中最具有特色的部分,数量大、类别多、保存价值较高。

(3) 多媒体资源。此类内容在原生数字特藏建设中起步较早,伴随数码相机、数码摄像机、智能手机的出现而大量产生,包括机构和个人的数码图片、音频、视频。此类数字特藏容易收集、表现直观,但对储存空间需求较大。图书馆可确定某一主题后逐步入藏并自成体系,如广州图书馆自2013年开始入藏纪录片,2019年与广州电视台合作共建共享广州本土影像资源已达9万小时,并在广州图书馆建立起了广州纪录片研究展示中心<sup>[5]</sup>。

(4) 公开的互联网内容。主要包括网页、新媒体内容、信息公开的相关内容等,其特征也是数量庞大,分布散乱,内容更新和技术迭代快,但也因此,图书馆的记忆职能可以得到更广泛、有效的发挥。图书馆可根据馆藏基础和特色方向制订保存目标和行动计划,如新冠疫情出现以来全球多个图书馆都开展相关网页内容的收集、整理、展示和利用。

(5) 个人社交网络文书。与公开的互联网内容相反,个人社交网络文书属于私密内容,建设者多、资源量大且更具独特性。该类资源类似于传统时代的书信、手稿等,目前对其收集、整理、利用的意识不强,案例鲜见。由于网络平台 and 个人的双重原因,散失或不能永久保存的可能性更大。需注意的是此类资源的版权和隐私保护问题,最好由个人和运营商授权并主动提交邮件、聊天记录等,并能自行设置封存期。

(6) 科学数据。2018 年国务院颁发《科学数据管理办法》,说明国家高度重视科学数据的管理维护。科学数据包括各研究领域开展基础研究、应用研究、试验开发等产生的数据,以及通过观测监测、考察调查、检验检测等方式取得并用于科学研究活动的原始数据及其衍生数据,应该是各高校图书馆重点建设的数字特藏,如复旦大学图书馆承担管理了多个科学数据、城市数据项目。

(7) NFT 数字藏品。非同质化代币 (Non-Fungible Token, NFT) 随区块链技术而兴起,是对应特定物品生成的唯一数字凭证,拥有数字版权并据此实现发、买、收、用,具有一定的金融属性。一些机构和图书馆已经围绕特藏内容进行 NFT 数字藏品发行,但国内还不能自行交易。图书馆应积极跟踪 NFT 数字藏品的发展,探索管理与服务方法。

## 2 图书馆数字特藏建设现状及不足

21 世纪初,在计算机、存储和网络等信息技术的推动下,国内外图书馆高度关注实体特藏的数字化,普遍认为这是保护特藏的最有效方法,逐渐建立起相应的标准、整理方法论和流程规范。各大图书馆的实体特藏逐步完成数字化加工,这些实体特藏的数字化版本构成了最初的数字特藏。而 3G 通信和 Web 2.0 技术使互联网进入个人时代,每个人都成为知识的生产者,人类的知识数量呈爆发式增长,新型知识载体层出不穷,数字特藏类型多样。对此,图书馆界给予了高度关注,开展了大量的实践,以搜集、整理和保存数字特藏。

### 2.1 实体特藏的数字化

目前国外图书馆数字特藏建设的实践主要是将实体特藏数字化,即“特藏的数字馆藏”(Special collections—Digital collections),这方面有许多成功案例。美国国会图书馆“美国记忆”项目按主题划分,汇聚同一主题分藏在不同机

构的文献,数字化后上传,包括书籍、印刷品、照片、绘画、手稿、地图、乐谱、音频及视频文件等,按照原始格式、题材或创建者、收集者、捐赠者,分成 100 多个主题集合向公众提供免费访问。典型的主题如国会图书馆藏亚伯拉罕·林肯的手稿计 4.5 万余份,其中有林肯第一次和第二次就职演说稿、《解放宣言》初稿等,超过半数已数字化并在网站上提供在线查询<sup>[6]</sup>;鲍灵格林州立大学通过“元数据模式文本编码倡议 TEI”(Text Encoding Initiative)在线提供第一次世界大战士兵的信件,为利用元数据整合开展图书馆特藏联合服务提供借鉴<sup>[7]</sup>;纽约州立大学斯蒂芬·卢斯图书馆收藏 1700 年以来关于海事历史的文献和文档,开展稀有海洋资源历史资料的整理与研究,并对这些珍贵资源进行了数字化<sup>[8]</sup>。法国于 2005 年 4 月组织欧洲 19 个国家联合创建数字图书馆,希望在数字图书馆建设方面形成统一的欧盟体系,最初命名为“European Digital Library”,后更名为“Europeana”,建设目标是通过构建统一的网络平台来整合欧洲具有代表性的文化遗产资源,为人们了解欧洲历史文化提供一站式浏览与检索服务,实现欧洲数字文化资源更广泛的传播与共享,截至 2023 年 10 月,已有来自欧洲的 4 000 个机构参与,集聚了 5 500 万条元数据<sup>[9]</sup>。日本国立国会图书馆从 2001 年起就对其特藏进行数字化,截至 2022 年完成约 311 万件,并将其中的 152 万件面向日本注册用户公开服务,同时计划未来五年是数字化转型推动期,将扩充国家数字化信息基础,向读者提供全方位的信息资源业务<sup>[10]</sup>。

国内图书馆数字特藏的建设也源自实体特藏的数字化,相关工作起步于数字图书馆建设。1997 年国家图书馆牵头的“中国试验型数字式图书馆项目”的重点就是文献数字化,该项目为中国数字图书馆工程奠定了基础<sup>[11]</sup>。2001 年,教育部启动大学数字图书馆国际合作计划 (CADAL),联合各高校图书馆开展馆藏资源数字化,并实现共建共享<sup>[12]</sup>。根据 CADAL 后台数



据显示,截至2022年12月,CADAL入库总量为2 888 088册(件),在线量为2 720 420册(件),其中古籍和民国文献等实体特藏文献的数字件达574 399件、满铁文献和侨批文献等特藏资源的数字件达68 769件。国家图书馆有许多特藏文献,目前已建成包括甲骨世界、敦煌遗珍、宋人文集、西夏碎金、赵城金藏、碑帖菁华、数字方志、年画撷英、前尘旧影等在内的特藏数据库并开放服务。上海图书馆特藏数字化工作成果丰硕,截至2023年11月,馆藏盛宣怀档案17.8万件已全部完成数字化,其中15.7万件数字资源在“盛宣怀档案知识库”中公开服务<sup>[13]</sup>,馆藏家谱3万余种、已数字化1.5万余种,馆藏古籍13万余种、已数字化2.3万余种。中山大学图书馆是目前海内外收藏徽州民间历史文献最多的机构<sup>[14]</sup>,截至2023年11月,中山大学徽州文书已数字化约104万页。该馆另有碑帖数字资源约4万张、古籍数字资源约1 000种。

专题特藏建设、海外特藏的数字化回归、实体特藏的数字化技术与方法等,是图书馆早期数字特藏工作的主要关注点,但实践的重心基本上都是围绕实体特藏的数字化展开的。

## 2.2 原生数字特藏建设情况

在原生数字资源收藏方面,一些图书馆已逐步开展互联网内容的收集整理和保存工作。如美国国家数字图书馆的数字馆藏中,有1 073个关于网页抓取的原生数字内容的主题,包括奥巴马在任期间的白宫推特网页<sup>[15]</sup>。高校图书馆则重点关注科学数据管理,如伦敦大学、牛津大学、布里斯托大学等高校的图书馆针对不同类型科学数据提供两项以上的数据存储服务<sup>[16]</sup>。悉尼大学图书馆数字特藏特别设立针对COVID-19期间的数据交存板块,专门用于记录悉尼大学教职员工和学生在新冠疫情期间的经历以及教育工作发生的变化;截至2023年11月底共收集到55个图像合集、24个短信合集、21个音视频合集<sup>[17]</sup>。普林斯顿大学图书馆数字特藏中的Archive-It项目,收集自2015年开始普林

斯顿大学网站的存档版本,并提供此集中所有网站的浏览和全文搜索服务<sup>[18]</sup>。

国内也有一些大型图书馆对原生数字资源收藏给予了高度关注。2019年4月,国家图书馆启动互联网信息战略保存项目,国内首家互联网信息战略保存基地落户新浪,新浪网发布的新闻和微博上公开发布的千亿级博文数据都在此项目得以保存<sup>[19]</sup>。2019年,上海图书馆与阅文集团达成网络文学专藏战略合作,启动全国首个网络文学专藏计划,首批十部作品入藏<sup>[20]</sup>。2022年12月,中国国家版本馆接收腾讯、阅文、腾讯音乐娱乐三家互联网企业的网络文学、视频、游戏、数字文保及数字音乐5大类45项网络数字版本入藏,宣告“中国国家版本馆数字版本入藏计划”全面启动<sup>[21]</sup>。

通过对国内外图书馆的实践调研发现,目前图书馆的数字特藏建设与服务已取得一定程度的进展,主要集中在实体特藏的数字化、专题资源库项目建设、特定网页抓取与长期保存等方面,以及少数原生数字资源的非系统性、非持续性入藏。国内外数字特藏建设还存在三个方面的不足:第一,业内对于数字特藏的整体重视程度和认识不够,一些图书馆已经意识到数字特藏的重要性并开始尝试建设与推进,如国家图书馆和上海图书馆等,但还没有形成普遍的行业认知;第二,已有的实践案例通常以项目为导向,还没有形成图书馆行业完整、规范、普遍的数字特藏管理与服务模式;第三,缺乏长期可持续的入藏计划与技术保障,与此同时,由于数字时代的内容与技术环境迭代比以前快了数倍甚至数十倍,大量有价值的原生数字资源被行业忽视,被迭代的数字内容也在快速消失。这三个方面的不足,折射出目前图书馆在资源建设中存在类型缺失、系统性和持续性不够等问题。在国家文化数字化战略背景下,图书馆应在收集整理实体特藏文献的同时,加强对原生数字特藏资源的建设和开发利用,这是当前图书馆文献资源建设十分紧迫的任务。

### 3 图书馆数字特藏建设的紧迫性及其发展路径

#### 3.1 数字特藏建设的紧迫性

从目前数字特藏的工作进展来看,图书馆已经以项目为导向开展数字特藏的收集、整理工作,但远远没有得到足够重视。主要原因,一是对“特藏”的认知不充分,仅局限在实体的古籍、民国文献、书画拓片、珍稀文献等方面,围绕这些资源开展保护和服务;二是即便已经认识到原生数字特藏的重要性,但如何建设、如何管理和利用,没有有效的制度和技术方法,原生数字资源每时每刻都在增长,也因政策、网络运营公司和个人等诸多原因,大量有长期保存价值的数字资源在快速消失。

数字特藏建设的紧迫性主要体现在以下四个方面。

(1)目前仍有大量的实体特藏并未完成数字化工作。实体特藏保护包括提供恒温恒湿的保存环境、抢救性修复等多重手段,但只能延缓资源的消逝而不能保证其顺畅利用与永久延续。数字化是实体特藏再生性保护最有效的手段之一,也是实体特藏开放、利用、传播的重要手段,因此全国应该统筹规划,在文献调研的基础上,有序完成现有实体特藏的数字化工作。

(2)多数原生数字特藏具有唯一性,具有重要的长期保存价值,而图书馆却还没有及时入藏。特藏资源重点就在于“特殊”,多数资源都是独一无二的,如某知名学者的电子邮件或聊天记录,类似之前的名人书信,某摄影师长期对于某个特定主题的拍摄而形成的完整图像资料等,图书馆都应在得到授权后及时整理入藏。

(3)原生数字资源更新迭代较快,更易散失,要及时从中发现特藏价值,收集整理、开发利用并传承保护。数字时代发展极其迅速,数字内容的生成、更改、删除、复制比传统载体时代更加容易,不及时保存就等同于消失。同时,网络信息管理政策、网络平台的经营状况、技术

稳定性等,都直接影响着原生数字资源的保存时限。只有迅速抓取保存、编目著录、在线流通,同时构建起长期保存系统,定期备份,才能保障这些原生数字资源的永久存储利用。

(4)标准规范是数字特藏建设的前提,图书馆急需研制数字特藏收集、整理与服务的标准体系。数字特藏究其根本是数字资源,具有极强的分散性,且当下图书馆的多种管理标准,也对数字特藏的管理信息系统与标准体系建设提出了挑战,一些图书馆有想法但无法推动。因此每个图书馆在建设之初就需要严谨规划,根据自身特色明确本馆数字特藏的建设范围、发展方向、特色专长等,遵循相关的标准开展工作并长期坚持。

数字特藏建设的紧迫性体现在每一件目标藏品上,很多数字资源的存在时长几无定数,图书馆应尽快议定数字特藏的建设计划并付诸行动,逐步构建具有各馆特色的数字特藏体系,这将是图书馆在数字时代的重要使命。

#### 3.2 数字特藏的建设路径

数字资源建设主要有资源收割、采购、读者交存三种渠道,一定程度上也适用于数字特藏建设。除了实体特藏对应的数字版本外,其余类别的数字特藏资源基本都来自图书馆外。图书馆需要从难以计数的数字内容中,确定本馆的数字特藏入藏范围、类型、专题特色和预期规模等策略,同时构建长期可持续的数字特藏建设的保障机制,确保资源稳定增长的持续性。当前重点是需要尽快起步,在实践中探索和积累数字特藏的建设经验。

##### 3.2.1 进一步加强政策层面的统筹规划与支持

通过多年实践,数字特藏建设具有了一定的行业共识和规范,但宏观政策层面的规划与支持尚有欠缺。数字特藏作为文化大数据体系中的重要组成部分,需要政策层面的统筹规划、制度建设与项目支持。在统筹规划方面,可将数字特藏建设纳入行业、地区、学校的文化事业建设范畴中,规划分层设立的集中式云存储基

地、重点实验室的数据基础设施、文旅融合的创新服务项目等;在制度建设方面,应重点建立健全数字特藏的交存机制,要求部分原生数字资源必须提供给图书馆、档案馆进行长期保存,对个体交存给予鼓励和支持,同时完善法律制度,对以保存为主的资源的版权管理进行适当豁免等;在项目支持方面,应予以人财物等基础条件的稳定支持,策划相关项目,如中国国家版本馆作为国家重大文化工程,写进“十四五”规划,通过政策的指导支持建成,成为履行国家版本资源保藏传承职责的中华文化种子基因库<sup>[22]</sup>。统筹规划、制度建设、项目支持等多措并举,数字特藏的发展才能获得源源不断的生命力。

### 3.2.2 尽快完成实体特藏的数字化及标引工作

当前实体特藏数字化后生成的数字特藏资源占比极大,这依赖于几十年来古籍、民国文献等数字化工程的实施。但是实体特藏数字化工作仍然需要保持高度的积极性和稳定的进度,一方面尚有难以计数的实体特藏并未数字化,另一方面因为技术、标准控制和存储空间等问题,早期部分数字化特藏的图像像素低、成像质量差、标引不够全面,因此图书馆可规划利用新技术重新进行数字化加工,建议不低于1200dpi,同时应用3D、光学、VR等技术,生成高度还原的数字特藏版本。新的数字特藏资源还需完成更为精细的深度标引,为后续打造知识库、知识图谱等奠定基础。

### 3.2.3 确定以专题建设为主的数字特藏建设

图书馆资源建设的源头是“采访”,目前这项业务更依赖“采”,导致数据库商、馆配商很大程度上左右了文献资源体系建设,但数字特藏更需要“访”,意识上要从被动采购转变为主动寻访。数字特藏有着数字资源的典型特点,分散性极强,只有将目光放在更长远的历史环境中,主动开展征集、交存、自动收割等工作,才能从海量杂乱的数字环境中寻得值得收藏的资源。其次,因数量巨大,每个图书馆的数字特藏规划都应以专题建设为主,评估本馆建设能力后,根据所处地域、地区文化、历史文化、机构、

学科、人物、技术等因素,明确本馆数字特藏的专题目标、收藏范围。公共图书馆应更多面向城市、社会读者、公共服务,收集对象以本地域历史记忆和地方文化为主;高校图书馆则应聚焦于学校的学科发展,收藏对象以学术沿革、学术成果的积累、学术脉络的承袭以及学术思想传承等为主。确定数字特藏的专题目标后,再采取“化整为零”的思路,用多个专题支撑全馆数字特藏架构体系,整合多媒体、互联网内容等多种形式的数字资源,资源之间相互支撑印证,充分利用信息技术构建管理、展示和利用平台,提升读者体验感。

### 3.2.4 研发专用数字特藏采集系统,倡导个人交存

数字特藏资源包括前文所提到的七个类别,还应根据数字特藏建设的需求和信息技术的发展进行增加,从发展趋势来看,建设目标应以原生数字特藏为主。原生数字特藏具有独特的网络环境、应用平台、存储格式和使用方式,因此建议针对不同类型的数字特藏研发专用的、易用的、支持多终端和多来源的采集系统,这是数字特藏资源建设成效的关键。同时推行“读者+图书馆”的资源共建模式,加强服务营销,才能快速积累起丰富的数字特藏。这些专用采集系统应倡导以个人为主体的主动参与,不受时间、地点和经济等因素影响,随时随地可实现个人交存。

### 3.2.5 案例:重庆大学图书馆的“数字特藏”建设实践

为了保存校园文化记忆,扩充多媒体类型的数字特藏,重庆大学图书馆专门研发了三个微信小程序:①2019年启用的“+馆藏”系统,读者可实现图片、音频、视频的主动交存,但必须按照图书馆著录标准和流程规范提交;②因特藏著录较为严格和繁琐,读者积极性受到一定影响,2020年又启动“时间胶囊”系统,简化读者需要填写的著录信息,增加学生志愿者和馆员,审核和完善标准著录信息,截止到2023年10月底这两个系统已经拥有数字特藏超过21万件;③2020年再启动“青春记忆纪念册”小程序,针

对毕业生群体,通过各种举措让毕业生留存在大学期间的照片、成绩册、奖状等资源,截至2023年10月底已有2 800多名毕业生建立了自己的电子纪念册。

与此同时,图书馆要求各类阅读推广、文化展览等活动结束后交存数字版本到系统平台。通过这些实践发现,数字特藏的资源建设必须要有“众筹”功能,以图书馆“记忆恒久”的特点来吸引读者参与。未来面向个人的文档、电子邮件、社交网络记录等更需要读者授权并主动交存,而专用程序的设计便捷性、功能丰富性和营销宣传的效能,将很大程度影响数字特藏的建设效果。

#### 4 图书馆数字特藏的整理与服务方案

数字特藏收集完成后仍有大量工作,整理、编目、管理、利用等每一个流程都需要完善的标准和实施方案,包括制订数字特藏的编目与著录规则,建立数字特藏整理与服务流程,搭建专门的信息资源管理与服务平台等,从而构建起完善的数字特藏整理和服务体系。

##### 4.1 在传统文献整理的方法论基础上延展数字特藏的整理流程

关于实体特藏资源的数字化,目前图书馆已建立起较为完善的标准与规范体系,原生数字特藏建设也应传承相关方法论,既保证资源整理与揭示的合理性,又保证各个图书馆之间的数据质量控制和数据交互。据此,数字特藏需要建立严格的“采访、审核、转码、著录、典藏、流通”整序流程,并根据数字技术的迭代及时拓展,实现对数字特藏的科学管理与服务。具体内容包括:每件数字特藏都必须有唯一的标准数字馆藏号,此馆藏号纳入图书馆整体馆藏体系;私密数字特藏需注意加密保管,实行严格的版权控制,界定流通范围;利用新的信息技术突破实体特藏资源的标引局限,对具有全文的数字特藏应侧重于内容更加细粒度的深度标引,

深挖文本内涵,揭示其知识元素和潜在的数据关联,为信息可视化奠定基础;在原有的著录标准基础上,扩充与图片、音视频等多媒体数字特藏相适应的GIS、EXIF等信息,在实施过程中充分利用人工智能技术进行元数据自动抽取等。总之,在坚持使用传统文献整理基本方法的前提下,数字特藏整序更需要体现载体适应性,为图书馆之间资源共享和读者服务奠定基础。

##### 4.2 制订完善的数字特藏整理与服务流程

图书馆对数字特藏进行收集、整序后,还需建立科学的管理与服务流程,才能确保数字资源建设与服务的长期可持续性发展。①应对图书馆现有的管理体制进行重构,整合技术部门和特藏部门,工作流程和服务流程应区别于传统图书馆的相关业务;②针对数字特藏七个类别各自的特点,分别制定采访、整理、系统支持、网络服务的全流程标准与规范,根据各自的特色梳理各层级、各环节的标准规范;③业务流程应充分关联本馆的文献管理、电子资源管理、实体特藏资源管理、单篇订购管理等系统,纳入统一的图书馆搜索系统和移动阅读平台,使之更加便捷地为读者服务;④制订严格的数据备份政策,以及长期保存的数据校验机制,确保珍贵特藏的安全性和有效性;⑤数字特藏的元数据存储应采用都柏林核心元素集作为标准,有利于互联网的数据交互,也遵循了图书馆的行业规范。由此逐渐形成完整的数字特藏生态体系。

##### 4.3 研发数字特藏管理平台并逐步拓展功能

基于以上调研认知与初步尝试,本文认为当前图书馆可以采用“以管理信息系统建设为核心”发展思路,利用数字特藏管理系统来实现数字特藏的规范化管理。这个系统是图书馆实现数字特藏相关资源管理和服务的基础,因数字特藏类型的多样性和文献整理的复杂性,它必定是一个复杂的、综合性的管理信息系统。本文提出数字特藏管理系统的软件架构,如图1所示。



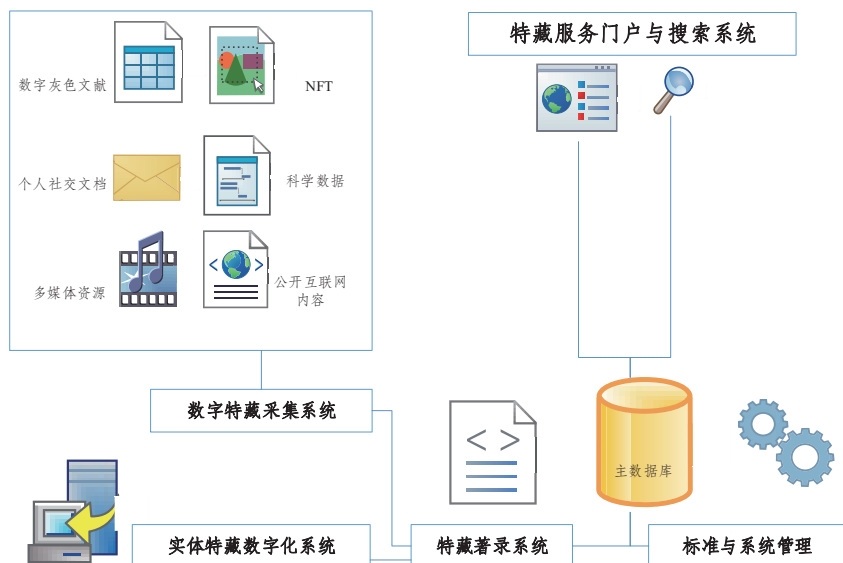


图1 数字特藏管理系统的软件架构

根据图书馆的长期保存和服务要求,数字特藏管理系统应将数字特藏资源的采集、审核、加工、描述、保存、发布、展示等功能集为一体,并需要运用媒体处理、人工智能辅助著录、人工智能审核、大数据和云技术等方面的信息技术,提供规范、标准和安全的数字特藏全流程管理。馆员完成收集、整理后进入管理系统,编目馆员在系统中依次进行分类、著录、标引、典藏流程,系统同时满足统计、回溯、查询、盘存的管理功能和浏览查阅、申请查阅、文献传递、反馈交流等服务功能,通过系统中的采集子系统、著录子系统、服务门户与搜索子系统、标准与系统管理体系,实现数字特藏资源从“收”到“用”的全面信息化管理。

系统的服务功能重点应面向网络读者,注重用户体验,打造一系列功能,包括数字特藏交存功能、地方文化传承与守护功能、科学研究和教学支持功能,以及信息素养和全民阅读等主题的特色资源展示功能等。能够根据用户行为提供个性化服务,且各个子系统的用户行为数据、分析统计数据等均纳入系统中进行集中管理。信息化流程的管控可以让整个数字特藏的

收集、整理和服务变得规范化、高效率,也能让资源得到逐渐积累和丰富,稳步推进图书馆数字特藏的建设。

## 5 图书馆数字特藏建设需要破解的主要问题与展望

尽管部分图书馆数字特藏建设起步颇早且佼佼者频现,但整体上依然呈现出明显的“失衡”状态,对于优秀者,已难以望其项背,后者者有心却显乏力,更多图书馆则是尚未起步。根据目前的研究与实践情况,图书馆数字特藏建设亟待解决三个方面的问题。

### 5.1 数字化转型过程中的特藏观念和方法革新

智慧图书馆建设以全面资源数字化和全面信息管理为基础,利用大数据、人工智能、区块链等信息技术构建面向用户的个性化、人性化和主动化服务体系,全面实现图书馆的数字化转型,当然也包括图书馆特藏建设与服务的数字化转型。在这个过程中,特藏资源建设的观念和方法必须转变为以数字资源为主。依靠历



史传承获得的特藏,鲜有图书馆有能力构建起实体特藏体系,倡导“数字特藏”会让更多的图书馆找到发展的新机遇。与普通资源相比,特藏资源建设需要有坐冷板凳、打持久战的态度,数字特藏资源同样如此。一方面因数字特藏建设的紧迫性,需要迅速行动,另一方面由于其数量大、类型多,资源建设进程必定复杂缓慢,不能有急功近利心态,一些数字特藏需要十年甚至更长时间才能显现其价值,必须按部就班地开展规划、建设、整理等工作。特藏整理与服务的方法也需要革新,数字特藏需要利用资源众筹、人工智能辅助著录与审核、互联网内容抓取、专用系统的资源收割接口、基于区块链的版权管理等新方法,这些领域尚需进一步研究和探索,并且要不断适应信息技术的发展。

## 5.2 依托数字特藏的图书馆服务创新

有了收集和整理的方法论,以及规范化的流程和标准体系,在充分的需求调研和数据分析的基础上,图书馆可以据此构建起数字特藏的管理与服务信息系统。当图书馆积累起丰富的数字特藏资源,就可以在长期保存、浏览、下载等基础功能之外,根据数字特藏的特征,针对文化建设、文旅融合、元宇宙建设、科学研究等实现服务功能延伸。比如元宇宙视域下的虚拟展厅构建,因数字特藏丰富且多样,可以利用元宇宙和人工智能技术,将数字特藏的展示虚拟化、个性化、交互化,提供多种展示模板和展示场景,降低现场展览在时间、空间、经济、人工方面的限制;又如支撑各类数字记忆项目的建设,数字记忆项目的建设需要资源在人物、时间、地点、事件、物品等多维度支撑记忆,数字特藏具有丰富的多媒体资源,完全符合数字记忆项目的建设需求;还可以重点开展数字特藏的文创产品推广,如定期抽取本馆重要的数字特藏进行 NFT 发行,有效拓展数字特藏的受众面。

## 5.3 数字特藏的版权控制与隐私保护

数字图书馆建设没能很好地解决数字资源

的版权问题,多数数字资源只有使用权而没有版权,不能赠送、去重、升值等,使图书馆在管理和服务领域处于被动。由于数字化复制的便利性极易给图书馆造成资产损失,侧重于数字版本的数字特藏更需要高度重视数字资产确权问题。可以通过区块链技术先对本馆的数字特藏进行确权,以长期保存为主要目的;再启动技术保护机制,对读者服务过程中的资源流转进行有效跟踪。另外,数字特藏的一部分资源会涉及隐私问题,如个人社交网络文档涉及个人隐私,科学数据涉及科研从业人员或科研机构的隐私,数字灰色文献及多媒体资源涉及个体或机构的隐私等。要积极争取个人和机构的授权文献,对于未取得授权的文献,图书馆要有放五十年或者更长时间直至版权保护失效的思想和行动准备,对于这部分特藏,初期只能进行整合保存而不纳入服务体系。数字资产确权、数字特藏管理系统的严格操作与保护机制、个人隐私的判定与对应管理制度、从业人员的道德建设与规范等,值得深入研究和探索。

展望数字特藏未来的发展,必须密切跟踪信息技术的进步和用户需求的改变。在用户不再缺乏信息资源而缺乏能够真正满足其需求的信息资源的时代,图书馆数字特藏的管理和服务必须适应所在城市的文化发展需求、所在大学的重点学术方向需求、个人的教育学习需求等。资源内容将更加多元化,数量更加庞大;资源整理将严重依赖人工智能的自动收集、标引和存储,同时具备完善的版权保护制度和相应的技术保障手段;资源服务会朝着主动化和沉浸式方向发展,图书馆基于用户行为研究和数据分析,主动推送相关的特藏资源或提供定制化的利用体验,以便更好地理解用户的需求,优化特藏资源的收藏等;VR/AR 等特藏资源虚拟展览大行其道,提高读者的阅读体验感。总而言之,数字特藏是图书馆未来发展的重要方向,将呈现资源丰富化、管理智能化和服务智慧化等特征,图书馆的重要性和影响力也必将随之得到提升。

## 6 结语

整个社会都在积极推进数字化转型和高质量发展,图书馆也必须开拓创新,充分发挥自己的基本职能,主动保存和保护在快速转型期的各类数字资源,特别是原生数字资源。同时积极研究和探索数字特藏的收藏、整理和服务的

全新方法论,构建具有标准化流程体系的管理信息系统,逐步扩大数字特藏的收藏范围,探索数字特藏的长期保存与保护机制等,由此逐步建设起面向未来的、具有各自特色的数字特藏体系。数字特藏建设必将成为图书馆践行国家文化数字化战略的重要举措,扩大文化资源供给范围,提升文化自信,促进图书馆逐步实现数字化转型。

**致谢:**本文系国家社会科学基金后期资助项目“我国电子图书发展史”(项目编号:21FTQB006)的研究成果。

## 参考文献

- [1] 李书宁,吕岩彦,杨春燕,等. 985高校图书馆数字特藏建设现状调查与分析[J]. 图书馆杂志,2011,30(8):58-63. (LI S N, LV Y Y, YANG C Y, et al. The investigation and analysis of digital special collection of 985 project university libraries[J]. Library Journal, 2011, 30(8): 58-63. )
- [2] HWEITT J A, PANITCH J M. The ARL special collections initiative[J]. Library Trends, 2003, 52(1): 151-171.
- [3] 年旭. 古籍数字化,让更多人共享经典魅力[N]. 光明日报,2023-02-03(02). (NIAN X. The digitalization of ancient books allows more people to share the charm of classics[N]. Guangming Daily, 2023-02-03(02). )
- [4] FARACE D J, FRANTZEN J. Perspectives on the design and transfer of scientific and technical information[C]//Third International Conference on Grey Literature. Luxembourg: 1997.
- [5] 广州图书馆. 纪录片[EB/OL]. [2023-11-28]. <https://www.gzlib.org.cn/fwjlp/index.jhtml>. (Guangzhou Library. Documentary[EB/OL]. [2023-11-28]. <https://www.gzlib.org.cn/fwjlp/index.jhtml>. )
- [6] Library of Congress. Digital collections[EB/OL]. [2023-02-28]. <https://www.loc.gov/collections>.
- [7] HERTENSTEIN E M, RABINE J L, SWEET M C. On the frontlines of richer metadata: technical services and TEI[J]. Technical Services Quarterly, 2018, 35(4): 353-366.
- [8] CONSTANTIA C. Digitizing rare maritime collections[C]//IS&T Archiving Conference 2013. Springfield: IS&T, 2013: 93-97.
- [9] Europeana PRO. Our mission[EB/OL]. [2023-11-29]. <https://pro.europeana.eu/about-us/mission>.
- [10] National Diet Library, Japan. About us[EB/OL]. [2023-02-21]. [https://www.ndl.go.jp/zh/aboutus/vision\\_ndl.html](https://www.ndl.go.jp/zh/aboutus/vision_ndl.html).
- [11] 孙承鉴,刘刚. 中国试验型数字式图书馆的探索与实践[J]. 数据分析与知识发现, 2001(6): 3-6. (SUN C J, LIU G. Exploration and practices in China pilot digital library project[J]. Data Analysis and Knowledge Discovery, 2001(6): 3-6. )
- [12] 金佳丽,薛霏,黄晨. 学术数字图书馆二十年:从数字化合作到数字知识服务联盟[J]. 中国图书馆学报, 2022, 48(2): 25-35. (JIN J L, XUE F, HUANG C. Academic digital libraries over 20 years: from digital collaboration to digital knowledge service alliance[J]. Journal of Library Science in China, 2022, 48(2): 25-35. )
- [13] 上海图书馆. 盛宣怀档案知识库[EB/OL]. [2023-12-09]. <https://sd.library.sh.cn/sd/home/index>.

- (Shanghai Library. Knowledge base of SHENG Xuanhuai archive [EB/OL]. [2023-12-09]. <https://sd.library.sh.cn/sd/home/index.> )
- [14] 王蕾, 申斌. 徽州民间历史文献整理方法研究——以中山大学图书馆馆藏为例[J]. 图书馆论坛, 2014(4): 120-126. (WANG L, SHEN B. Research on scholarly editing of Huizhou folk historical literature: case study of collections in Sun Yat-sen University Libraries[J]. Library Tribune, 2014(4): 120-126. )
- [15] Digital Public Library of America. 1 073 results for webpage[EB/OL]. [2023-02-21]. <https://dp.la/search?q=webpage>.
- [16] 王友转. 英国图书馆科研数据管理服务调研分析[J]. 新世纪图书馆, 2020(11): 71-76. (WANG Y Z. Investigation and analysis of scientific research data management services in British libraries[J]. New Century Library, 2020(11): 71-76. )
- [17] The University of Sydney Library. The University of Sydney during COVID-19[EB/OL]. [2023-11-30]. <https://digital.library.sydney.edu.au/nodes/view/8374>.
- [18] Princeton University Library Special Collections. Archive-It[EB/OL]. [2023-11-30]. <https://library.princeton.edu/special-collections/databases/archive-it>.
- [19] 国家图书馆研究院. 国家图书馆启动互联网信息战略保存项目[J]. 国家图书馆学报, 2019(3): 24. (The National Library of China Research Institute. The National Library of China launches the Internet information strategy preservation project[J]. Journal of the National Library of China, 2019(3): 24. )
- [20] 罗昕. 上海网文又放“大招”: 设立全国首个网络文学专藏库[EB/OL]. (2019-08-29) [2023-02-28]. <http://www.chinawriter.com.cn/GB/n1/2019/0829/c403994-31325898.html>. (LUO X. Shanghai's online literature shows another "big move": set up the first online literature collection in China[EB/OL]. (2019-08-29) [2023-02-28]. <http://www.chinawriter.com.cn/GB/n1/2019/0829/c403994-31325898.html>.)
- [21] “中国国家版本馆首批网络数字版本入藏仪式”举办[EB/OL]. (2022-12-29) [2023-02-28]. <http://www.xinhuanet.com/tech/20221229/1114b2ccfe184f5d859a4a119e729e6b/c.html>. (The first batch of online digital version collection ceremony of the China National Archives of Publications and Culture was held[EB/OL]. (2022-12-29) [2023-02-28]. <http://www.xinhuanet.com/tech/20221229/1114b2ccfe184f5d859a4a119e729e6b/c.html>.)
- [22] 薛鹏, 杨文佳. 中国国家版本馆开馆 新时代标志性文化传世工程[EB/OL]. (2022-07-31) [2023-02-28]. [http://www.news.cn/politics/2022-07/31/c\\_1128878199.htm](http://www.news.cn/politics/2022-07/31/c_1128878199.htm). (XUE P, YANG W J. The opening of the China National Archives of Publications and Culture is a landmark cultural inheritance project in the new era[EB/OL]. (2022-07-31) [2023-02-28]. [http://www.news.cn/politics/2022-07/31/c\\_1128878199.htm](http://www.news.cn/politics/2022-07/31/c_1128878199.htm).)

**王彦力** 重庆大学图书馆特藏部主任, 重庆大学人文社科高等研究院副研究馆员, 硕士生导师。重庆 400044。

**杨新涯** 重庆大学宣传部常务副部长, 重庆大学人文社科高等研究院研究馆员, 博士生导师。重庆 400044。

**冉蔚然** 重庆大学图书馆馆员。重庆 400044。

(收稿日期: 2022-12-05; 修回日期: 2023-02-28)