

图书馆数字人文众包项目实践*

贺晨芝, 张 磊

摘 要 众包模式与数字人文研究的结合在促进数字人文研究深化和突破的同时, 也起到了提升全民科学素养的作用, 营造了大众和数字人文研究机构的双赢局面, 成为数字人文研究的发展趋势。上海图书馆近年来致力于将众包理念应用于各数字人文项目。文章以上海图书馆各众包项目为例, 从建设需求、用户对象、众包形态等方面介绍数字人文众包项目建设理念; 基于系统/平台使用现状, 为众包在图书馆数字人文中的应用提出建议。

关键词 图书馆 数字人文 众包 上海图书馆

引用本文格式 贺晨芝, 张磊. 图书馆数字人文众包项目实践[J]. 图书馆论坛, 2020, 40 (5): 3-9.

Library Digital Humanities Crowdsourcing Projects

HE Chenzhi, ZHANG Lei

Abstract The combination of crowdsourcing and digital humanities research, which not only promotes the furtherance and breakthroughs in digital humanities research, but also plays a role in improving the scientific literacy of the people, has created a win-win situation for the public and digital humanities research institutions, thus becoming one of the inevitable trends in the development of digital humanities research. Shanghai Library has committed to applying crowdsourcing to various digital humanities projects in recent years. Taking the crowdsourcing projects of Shanghai Library as examples, this paper introduces the construction concept from the aspects of construction demand, user objects and crowdsourcing forms, and makes some suggestions on the application of crowdsourcing in library digital humanities based on current situation of system/platform usage.

Keywords library; digital humanities; crowdsourcing; Shanghai Library

0 引言

随着新技术的渗透, 图书馆除提供纸质文献保存与服务外, 还尝试将信息技术融入传统服务, 提供全面、精准、智能的知识服务。数字人文作为一门将关联数据、文本分析、机器学习、UGC(用户生成内容)、数据可视化等新兴技术应用于传统人文研究的计算与人文交叉学科, 契合

图书馆发展方向, 是图书馆发展趋势之一^[1]。在数字人文研究中, 图书馆因拥有丰富的馆藏资源和元数据, 以及其知识收集、存储、服务的职能, 主要承担资源的管理和组织角色, 为人文研究提供大规模高质量的数据^[2]。

“众包”(crowd sourcing)概念由美国《连线》杂志记者杰夫·豪(Jeff Howe)于2006年提出^[3], 核心思想是利用用户的知识盈余来共创价

*本文系国家社会科学基金项目“面向数字人文研究的图书馆开放数据体系构建与服务模式设计研究”(项目编号: 18BTQ027)研究成果。

值。随着Web2.0技术的发展,基于用户贡献的众包理念被应用于各领域,数字人文领域也在不断尝试各类众包应用模式。Terras Melissa指出,众包模式在数字人文中的应用,在利用民众知识盈余来建立可应用于数字人文研究的数据集的同时,将大众融入人文研究领域,为人文学科与公众建立了更紧密的关联^{[4][31]}。赵宇翔提出,众包模式在数字人文中的应用不仅能够促进数字人文研究的深化和突破,也能提高民众的人文素养^[5]。图书馆数字人文众包项目是请大众为机构工作,更是为用户提供深入参与公众记忆建设的机会^[6]。用户和数字人文研究机构的双赢局面使众包成为数字人文研究发展的趋势。

上海图书馆(以下简称“上图”)自2016年开始尝试在数字人文项目中融入众包模式,试图利用共创价值理念,吸纳用户盈余知识,以较低成本收集资源,实现资源文本化,从而为用户提供更丰富精准的知识服务,加强人文知识大众化和普及化。通过分析已有数字人文众包项目,结合本馆实际需求,上图有选择性地开发建设数字人文众包项目,先后推出历史文献众包平台、验证码项目,并在家谱、上海年华之从武康路出发、文化名人手稿等数字人文项目中加入在线识谱、上传家谱、上传照片、标识留言等功能,不断尝试将众包模式应用于数字人文项目,取得了初步成效。本文以上图数字人文众包项目为例,从建设需求、用户对象、众包形态3个方面入手,结合应用现状分析,为图书馆数字人文众包项目建设提供参考。

1 文献调研

1.1 众包案例

澳大利亚国家图书馆2008年推出澳大利亚报纸数字化项目,项目开发的Trove网站为民众提供了对澳大利亚报纸OCR文本化结果进行校对的平台,是最早的具有广泛影响力的数字人文众包实践项目^[7]。随后,一些图书馆和大学纷纷推出各类数字人文众包项目,包括:卡耐基梅隆大学的“reCAPTCHA”系统以OCR软件无法识

别的文字扫描图代替原来的验证码图片,借助用户登录进行文字识别,为纽约时报、Google图书等完成文字识别工作^[8];芬兰国家图书馆digitalkoot项目中的Mole Hunt游戏吸引用户在玩游戏的同时,对影像图片文本化结果进行核对,已完成超过650万次校验工作^[9];爱尔兰美奴斯大学“Letters 1916-1923”项目向用户征集信件,并为用户提供信件内容的抄录平台,2018年该项目将由大众捐献、抄录的3,000余封信件建设成知识库供用户查询^[10];美国国会图书馆“By the People”平台邀请大众标引、抄录、审核图书馆馆藏资源^[11];中华书局的数字化整理平台为在线用户提供共同参与古籍审校的工作机会^[12];伦敦大学学院发起的边沁手稿抄录项目将已有边沁手稿发布于平台,通过收集用户抄录的手稿全文数据,建设可全文搜索的边沁文集知识库^[13]。

1.2 分类分析

学者对诸多数字人文众包项目进行综合分析,提出相关分类方法。分类角度主要有建设需求、用户对象、众包形态3个方向。这3个方向恰好是建设数字人文众包项目需要重点思考的内容。

数字人文众包项目建设综合考虑了数字人文研究的需求及众包模式的特色。刘炜等指出,数字人文研究的内容是基于数据的、细粒度的、可分析的文本资源,是数字人文项目建设的基础之一^[14]。数字人文众包项目以聚集大众智慧为数字人文研究提供可分析资源为目标。Oomen等借鉴新西兰国家图书馆有关电子产物生命周期“描述、管理、揭示、利用/再利用、创造”的概念,将数字人文众包项目从内容上分为:文本纠错/抄录与分类(对资源进行文本化、分类等描述工作);内容补充(对已有资源进行注解、知识补充);共策(非专业大众与专家共同策划展示内容);资源征集(向公众收集主题资源);众筹(筹集资金购买资源或加工资源)^[15]。Carletti等将该分类进一步归纳为已有资源扩充深化(如抄录、标引、分类、校对)和新资源收集,分别针对有明确主题的数字人文项目及筹建中的数字人文项目^[16]。

在用户方面,边沁手稿抄录项目负责人

Terras Melissa指出,与一般商业化众包模式不同,人文领域的众包项目并非面向任意大众群体,而是针对具有一定人文领域知识、极具热情并愿意贡献知识的群体^[47]。Claude等将众包按用户动机分为3类:自愿、有偿、无参与者^[17]。自愿类项目吸引用户自愿开展众包工作,是数字人文众包项目中最常见的类型。有偿类项目是指用户在完成众包任务后会获得一定报酬,如亚马逊的土耳其机器人。无参与者类项目是令用户在无意识中开展众包工作,如验证码项目(reCAPTCHA)。同样从用户维度来分类,Bonney等的分析侧重点为用户与项目组织方的关系,将数字人文众包项目分为贡献、协作、共创3类^[18]。Simon在这一分类的基础上再添加“自主创建”类,特指由机构提供数据资源、用户利用资源自主开发的数字人文项目^[19]。

众包在数字人文项目中的应用旨在为用户提供提供一个可以贡献盈余知识的平台和入口,多以网站平台形式出现,包括自建平台、使用第三方平台、以功能形式融入网站,以及游戏等。Andro等将数字人文众包项目从形态上分为3类:显性式、游戏式、隐性式^[20]。显性式众包泛指传统的任务发布及领取模式。游戏式众包是指将众包任务融入游戏,令用户在娱乐中开展文本化纠错等众包工作。隐性式众包特指验证码项目(reCAPTCHA),取其没有明显的形态、巧妙融入众包任务之意。

通过对照自身的需求以及上述各维度众包分类,上图分析建设需求、定位用户群体、选择众包形态,创建了符合上图特色的数字人文众包项目。部分项目已投入使用,取得了初步成效。

2 上海图书馆众包项目

2.1 建设需求

上图2014年开始探索数字人文项目,依托各类特色馆藏资源(如家谱、盛宣怀档案、上海年华、手稿)建设了一系列特色主题文献知识库及服务平台,取得了较好的成效,也遇到了可分析资源不足瓶颈。上图数字人文项目主要是基于各类特色馆藏资源的元数据分析、关联与展示,而对

于蕴藏最多知识信息的资源文本内容,由于资源数量庞大且具有专业性,文本化的时间成本与资金成本对图书馆都具有很大压力,从而造成文本资源不足,无法进行深入挖掘与分析。此外,图书馆对民众愿意分享的民间资源,由于没有便捷的收集入口,无法进行有效征集,错失了丰富馆藏的机会。基于以上问题,众包在数字人文项目中的应用需求可以分为两大类:影像图片文本化、资源征集。

上图拥有大量特色馆藏资源,大部分已完成了数字化、影像化工作。目前市场上主流的图片文本化技术OCR(光学字符识别)对印刷品中的中文识别效率非常高,但手稿、古籍、家谱等由于书写笔迹凌乱、风格迥异、形式复杂,OCR识别正确率低下,文本化成本极高。因此,影像图片文本化是上图数字人文众包的最大需求。面对这一问题,上图根据不同的抄录需求,分别建设了提供元数据深度标引和全文著录的“历史文献众包平台”,将难度较低的抄录内容切割成单字作为验证码的“验证码项目”,以及专门针对家谱世系图形式的家谱抄录“在线识谱”功能。

众包在数字人文项目资源拓展上的助力主要表现为:向公众收集各类资源,从数量和形式上丰富馆藏。该类需求主要出现在已确定的数字人文项目中,针对某一特定主题,向公众征集资源,对征集到的资源进行再加工、展示,从而更好地服务大众。针对该类需求,上图在“家谱知识服务平台”增加“上传家谱、在线修谱”功能,向公众征集家谱资源;在“上海年华之从武康路出发”(简称“武康路项目”)中开发分享空间,提供上传照片入口,收集武康路老照片、用户自拍照片;在“名人手稿档案库”(又称“文化名人手稿项目”)中设计“标注、留言”功能,获取用户对手稿的见解,收集手稿背后的故事。

2.2 用户对象

上图作为非盈利机构,无法开展有偿的众包项目,主要依靠用户自愿或无意识地参与众包,因此众包需求明确,项目多为完全由机构指定的“用户贡献”类型。此外,为了解、吸纳大众对

数字人文的想法与创意，上图连续多年推出开放数据开发竞赛，为参赛团队提供海量数据，由参赛者自主开发数字人文项目，属于赋予用户最大自由度的“自主创建”众包类型。

不同主题的数字人文众包项目，其众包需求和用户群体各有不同。历史文献众包平台作为一个文本抄录及元数据深度标引平台，目前推出的任务主要针对上图特藏档案资源，需要用户在繁体字识别、手写字迹辨认、人文历史背景上具有一定的知识储备；在冷启动阶段，上图与高校合作举办文献招录竞赛，邀请历史、中文、艺术等专业的大学生参与平台的试用。家谱作为上图最早推出的数字人文项目，已聚集比较稳定的用户群体，这一成熟的用户群体也是在线识谱、家谱上传、在线修谱等众包任务的主要目标群体。武康路项目旨在通过用户收集更多建筑及马路照片，因此将主要用户群定位为武康路游客，该项目与“老洋房阅读之旅——行走武康”活动合作，供游客使用。文化名人手稿项目中的标注及留言功能，主要面向具有专业知识或历史渊源的用户，希望用户能够提供手稿背后的故事以及关于手稿的独到见解，依赖他们的知识积累，丰富上图已有资源。验证码(项目)嵌入上图主页和上图数字人文项目登录界面，面向上图所有线上用户。开放数据开发竞赛具有一定的开发技能要求，目标用户主要为具有开发能力、对上图特藏数字资源感兴趣、有独到见解的群体。

在数字人文众包项目群体定位时发现数字人文具有较强的领域性。不同主题的众包项目，其目标群体需要具有相应的人文领域知识，对用户的人文素养要求较高，在推广使用上面临的难度要高于一般众包项目。为此，需进行定点推广，并配以具有吸引力的激励机制。

2.3 众包形态

参照显性式、游戏式、隐性式等众包形态特征，上图选择了显性式与隐性式这两种较易实现的众包形式。游戏式众包需要进行专业的游戏化设计，开发成本高，且市场上的游戏式众包多用在文本识别纠错及资源分类上，上图尚无这方面

需求。对于显性式众包，上图分别开发了综合性开放众包平台以及嵌入已有数字人文项目的众包功能来满足不同的需求；对于隐性式众包，上图借鉴reCAPTCHA项目，开发了验证码项目。

2.3.1 综合性开放众包平台

文本资料是数字人文研究的必要素材，影像图片文本化是通用于所有数字人文项目的基本需求。面对包括古籍书本、创作手稿、信函电报等在内的种类繁多的文献文本化需求，上图建设了综合性开放众包平台——“历史文献众包平台”(见图1)，在支持用户对文献影像图片进行全文抄录的同时，还允许其对文献中的人、地、时、事等的内容特征进行深度标引。馆藏文献在收集、数字化的过程中由馆员进行元数据著录，丰富的元数据信息可以帮助用户更好地理解文献。由于不同类型文献的元数据字段组成不同，平台特别建立了灵活的元素集管理机制，允许发包方在导入待抄录文献图片的同时，导入相匹配的元数据信息并显示在抄录界面上，方便用户抄录。



图1 上海图书馆历史文献众包平台主页

开放性是该平台的一大特色。文本化资源的匮乏不仅是上图面临的问题，更是业界普遍存在的现象。对此，上图选择建设一个能够支持不同机构及个人发包的开放性平台，将发包过程功能化、流程化，向可信赖的机构、个人开放；再结合可发布不同元数据结构类型文献的特色，为暂无独立开发众包抄录平台计划但想进行众包尝试的个人或机构提供便捷的任务发布渠道，也不失为一种避免重复建设平台而造成资源浪费的方式。

2.3.2 嵌入式众包功能

对已有数字人文项目，考虑到其众包需求主要针对某一主题，且网站已聚拢了相当数量的用户群体，上图以在已有平台基础上增加新功能的方式进行任务众包。(1)家谱二期项目中的在线识谱，是专门针对家谱这一特殊形态文献的抄录平台。家谱中的世系图形式多变，蕴含了大量的关系信息，普通的抄录平台无法满足这一需求。因此，上图特意设计能够揭示家谱知识的在线识谱功能，并能可视化展示识谱结果。(2)武康路项目需要大量关于建筑及马路的照片，通过在项目中加入“分享空间”的方式，让用户在分享观赏感受的同时丰富上图的馆藏资源。(3)文化名人手稿项目的众包重点是收集用户对手稿的标注。手稿的珍贵之处除了其蕴藏的内容外，其展示形式也值得研究。该项目融入标注功能，并为用户提供圈划功能(见图2)，令用户能够准确表达对手稿每一处的见解，这是简单的文字描述所不能替代的。



图2 文化名人手稿项目用户圈画、标注、展现界面

2.3.3 隐形式众包工具

验证码项目借鉴了卡耐基梅隆大学reCAPTCHA项目的理念，在验证码中融入待识别文字，将繁复的文献抄录众包任务分划为单字识别微任务，让用户在登录过程中输入图片内的文字，无意识地完成文本化抄录。这种去任务化、碎片化的众包形态能最大化地利用大众智慧，令网络用户在日常操作中“主动”贡献认知盈余。对发包方而言，需把待识别的文字图片导入验证码项目操作平台，配置验证码形式，然后通过调用接口将验证码嵌入网站(见图3)。



图3 上海图书馆门户网站登陆界面

3 使用现状

在上图多个数字人文众包项目中，历史文献众包平台、验证码项目、武康路微站已投入使用，并进行了适当推广；家谱、手稿等项目仅在平台中添加众包功能，尚未正式推广使用。下文从质量保障机制、用户激励机制、平台/系统使用情况等方面探讨上图众包项目的管理、使用和成效。

3.1 质量保障机制

数字人文众包项目旨在为数字人文研究提供可分析的数据，因此众包所得数据需在质量上得到保障。上图众包项目中的抄录任务(如历史文献众包平台、家谱在线识谱功能)采用专家审核制度进行质量管控。其中，历史文献众包平台还支持多人抄录，即同一份任务由多人分头开展，通过综合多份抄录结果，进一步保证抄录质量。验证码项目则是每次在验证码中放入多个字，其中包含至少1个已知值的字用于判断抄录结果正确与否，当待抄录字被抄录为同一值的次数超过设置的阈值时，判定该值为正确答案。通过后期人工抽检，验证码项目准确率介于80%~90%；准确率可通过调高阈值来提升，但相应地会降低抄录效率。对资源扩充类众包任务，以手稿为例，由于尚未找到合适的判断内容准确性的审核机制，仅将标注内容以弹幕形式展示在手稿中以丰富手稿内容，暂不纳入可研究的范围。现有质量保障机制对人力投入要求较高，未来或可

考虑引入自动对比、机器学习等元素以提高审核效率；亦可借鉴国外项目经验，增加纠错类众包任务，以游戏化形态展示，让用户来审核抄录结果。

3.2 用户激励机制

验证码项目作为用户无意识参与的众包形式，无需激励措施。与现有数字人文项目相结合的众包功能由于项目前期已具有一定的用户积累，且原有系统没有用户激励需求，因此尚未匹配相应的激励机制。唯有历史文献众包平台，作为一个从零开始的众包项目，发布内容难度高，在开发过程中设计了积分排行榜进行用户激励；该项目冷启动阶段开展的抄录竞赛设有奖金及证书，以激励大家参赛。当用户积分达到一定额度后，管理员会视其专业水准赋予专家头衔及相应的任务审核权限，以兹鼓励。

3.3 平台/系统使用情况

上图与南京大学联合举办“文化遗产数字化”竞赛，对上图发布的“盛宣怀档案”主题资源进行抄录比拼，作为历史文献众包平台的冷启动。共有来自南京大学、复旦大学等多间高校的60余人参赛，历时3个月，完成抄录任务200余件。竞赛结束后，由于没有后继推广活动，平台用户活跃度明显下降，每月抄录任务仅5~10份。目前平台用户171人，共计完成任务346项。

验证码项目应用于上图主页(<https://library.sh.cn/#/index>)和各数字人文平台的登录界面。以上图首页登录界面为例，2018年11月-2019年7月，用户登录达50余万次，输入字数152,315个(用户多次刷新验证码均失败后，会转入常规验证码登录方式)，其中正确验证了109,228字，约为网站用户登录量20%。在使用中，部分用户反馈验证码难度不合理；目前设定3次刷新后出现常规验证码，刷新率为40.92%。

“上海年华之从武康路出发”与上海“老洋房阅读之旅——行走武康”活动项目合作，作为武康路的导览平台，活动期间(2018年6-11月)网站浏览量达1万余次。但由于平台缺乏指向性引导，鲜有用户使用“分享空间”上传照片。

4 思考与建议

上图形式多样的众包项目为数字人文研究提供了更多的可分析数据，历史文献众包平台丰富了上图馆藏资源中的文本化数据，验证码项目已对家谱、古籍、信函等多种类型的资源完成抄录。在建设使用过程中，遇到缺乏活跃用户、质量保障成本过高等问题。结合上图数字人文众包项目建设经验，对众包在数字人文领域中的应用提出4个建议：

(1)加强机构间的合作力度，扩大众包平台用户群体。通过与南京大学的合作，上图历史文献众包平台成功启动，并完成了一定数量的抄录任务。该平台具有开放性，可以供其他有需求的高校或机构发布任务、开展活动。而此类机构本身就具有一定的用户群，与其建立稳定的合作关系，能为上图众包平台带来大量活跃的用户。

(2)定向推广众包项目，提高用户活跃度、粘度。由于数字人文众包内容具有特殊性，参与群体需要具备相关人文领域的知识，而不同主题的众包项目面向的群体也有所不同。因此，项目应在特定的用户群体中进行推广，如在已有数字人文项目中加入众包功能或实现跳转引导，以形成或加入相关人文主题社区，获取更多活跃、有粘合度的使用者。

(3)细化众包任务颗粒度，增强任务可完成性。同为抄录类众包项目，验证码项目将任务碎片化，让每一个登陆用户无意识地参与抄录工作，抄录量远远高于其他众包平台或项目。而在众包领域，现有的所有游戏式众包项目皆为细粒度的众包任务。由此可见，细粒度、简易化的任务呈现形态更丰富、有趣，更容易为大众所接受。

(4)引入自动审核功能，优化质量保障机制。为了在保障质量的前提下降低审核成本，建议引入自动化审核功能。比如，历史文献众包平台支持多人抄录同一份任务，未来可考虑引入机器自动对比合并功能，减轻专家工作量，提高审核效率；对验证码项目而言，可考虑将前期成果数据作为机器学习训练集，用以提高OCR(文字识别)

准确率；未来可在审核前先进行自动识别，将识别结果作为参考答案供专家审核时作对比。

5 结语

众包是图书馆开展数字人文研究的必经之路。上图结合国内外已有项目的经验，根据自身研究需求，针对不同用户群体，尝试开发上海图书馆历史文献众包平台、验证码项目、家谱在线识谱、名人手稿标注留言等不同形态、不同内容的数字人文众包应用及功能。经过观察和分析各项目的测试和使用情况，发现若要作为成熟应用向大众推广，将面临使用群体小众及质量保障成本高昂两大阻碍。本文结合所有项目的使用现状，借鉴国内外数字人文众包的应用经验，提出4点建议，作为上图数字人文众包项目的改进方向：加强机构间的合作力度，扩大众包平台用户群体；定向推广众包项目，提高用户活跃度、粘度；细化众包任务颗粒度，增强任务可完成性；引入自动审核功能，优化质量保障机制。

参考文献

- [1] 张舵, 吴跃伟. 国外图书馆支持数字人文的实践及启示[J]. 图书馆杂志, 2014, 33 (8): 47-52.
- [2] 夏翠娟, 张磊, 贺晨芝. 面向知识服务的图书馆数字人文项目建设: 方法、流程与技术[J]. 图书馆论坛, 2018, 38 (1): 1-9.
- [3] Howe J. The Rise of Crowdsourcing [EB/OL]. (2016-06-01) [2019-06-12]. <https://www.wired.com/2006/06/crowds/>.
- [4] Terras Melissa. Crowdsourcing in the Digital Humanities[M]// A New Companion to Digital Humanities. New Jersey: John Wiley & Sons, 2015.
- [5] 赵宇翔. 科研众包视角下公众科学项目刍议: 概念解析、模式探索及学科机遇[J]. 中国图书馆学报, 2017, 43 (5): 42-56.
- [6] Trevor Owens. Crowdsourcing Cultural Heritage: The Objectives Are Upside Down[EB/OL]. (2012-03-10) [2019-07-27]. <http://www.trevorowens.org/2012/03/crowdsourcing-cultural-heritage-the-objectives-are-upside-down/>.
- [7] Trove [EB/OL]. [2019-08-16]. <https://trove.nla.gov.au/>.
- [8] Von Ahn L, Maurer B, Mcmillen C, et al. reCAPTCHA: Human-Based Character Recognition via Web Security Measures[J]. Science, 2008, 321 (5895): 1465-1468.
- [9] Hunt the Mole[EB/OL]. [2019-08-16]. <https://now.libraries.fi/mole.html>.
- [10] Letters 1916-1923 [EB/OL]. [2019-08-16]. <http://letters1916.maynoothuniversity.ie/>.
- [11] By the People [EB/OL]. [2019-08-16]. <https://crowd.loc.gov/>.
- [12] 古籍整理工作平台[EB/OL]. [2019-08-16]. <http://auto-edit.ancientbooks.cn/DocZhengli/>.
- [13] Transcribe Bentham [EB/OL]. [2019-08-16]. <https://blogs.ucl.ac.uk/transcribe-bentham/>.
- [14] 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.
- [15] Oomen J, Aroyo L. Crowdsourcing in the Cultural Heritage Domain: Opportunities and Challenges[C]// International Conference on Communities & Technologies. [S.l.]: ACM, 2011.
- [16] Carletti L, Giannachi G, Price D, et al. Digital Humanities and Crowdsourcing: An Exploration [EB/OL]. (2013-02-05) [2019-12-25]. <https://mw2013.museumsandtheweb.com/paper/digital-humanities-and-crowdsourcing-an-exploration-4/>.
- [17] Claude Guittard, Eric Schenk. Le Crowdsourcing : Une typologies des pratiques d' externalisation vers la foul [EB/OL]. [2019-12-25]. <https://www.strategie-aims.com/events/conferences/2-xixeme-conference-de-l-aims/communications/25-le-crowd-sourcing-une-typologie-des-pratiques-d-externalisation-vers-la-foule/download>.
- [18] Bonney R, Ballard H, Jordan R, et al. Public Participation in Scientific Research: Defining the Field and Assessing Its Potential for Informal Science Education: A CAISE Inquiry Group Report [EB/OL]. [2019-12-25]. <https://files.eric.ed.gov/fulltext/ED519688.pdf>.
- [19] Simon N. The Participatory Museum[J]. Curator the Museum Journal, 2011, 54 (1): 111-115.
- [20] Andro M, Saleh I. Bibliothèques numériques et crowdsourcing: Une synthèse de la littérature académique et professionnelle internationale sur le sujet[C]. Fès: CI-DE'17, 2014.

作者简介 贺晨芝 (通信作者, ORCID: 0000-0001-6967-8854, czhe@libnet.sh.cn), 上海图书馆系统网络中心工程师; 张磊, 上海图书馆系统网络中心正高级工程师。

收稿日期 2019-12-01

(责任编辑: 何燕)