

面向数字人文的图书馆众包平台构建研究*

——以上海图书馆历史文献众包平台为例

姚啸华¹ 贺晨芝¹ 徐孝娟² 全石峰¹ (1 上海图书馆 2 安徽大学管理学院)

摘要 本文从众包运营框架和流程入手,结合国内外众包实践对面向数字人文的图书馆众包平台的作用、构建模式和功能模块等关键问题进行了探讨,并以上海图书馆历史文献众包平台的实践为例,对其目标定位、构建模式、功能特色以及试运营情况等进行了具体分析,总结其构建经验并展望未来实际运营中可能存在的问题,以为图书馆在众包平台构建与数字人文服务方面提供一定参考借鉴。

关键词 数字人文 图书馆 众包平台 案例分析

DOI: 10.13663/j.cnki.lj.2020.06.015

Construction of Crowdsourcing Platform of the Library for Digital Humanities: A Case Study of the Crowdsourcing Platform for Digital Historical Documents of Shanghai Library

Yao Xiaohua¹, He Chenzhi¹, Xu Xiaojuan², Quan Shifeng¹

(1 Shanghai Library; 2 School of Management, Anhui University)

Abstract This paper discusses the problems about the role, construction mode and function module of the crowdsourcing platform of the library for digital humanities through the normal operation framework, process and the crowdsourcing projects at home and abroad. And this paper construction of the crowdsourcing platform through a case study of the one for digital historical documents of Shanghai Library from various perspectives, including building purpose, construction mode, functional features and operational test. The paper sums up its experience and foresees problems in the future operation of the platform in order to provide reference for other libraries constructing the crowdsourcing platform for digital humanities.

Keywords Digital humanities, Library, Crowdsourcing platform, Case analysis

0 引言

数字人文被认为是图书馆重要的发展趋势之一^[1]。近年来国内外图书馆积极投身数字人文的研究与实践中,扮演着重要的角色,其中很重要的一部分工作是对馆藏文献资源的数字化和再开发,为数字人文研究提供基础设施支持。与传统人文研究相比,数字人文研究对基础设施需求的显著特征在于对文献内容细粒度的需求^[2],这就要求图书馆需要从对文献资源的数字化向文献内容的文本化转变,从文献外部特征的组织揭示向文献内容深度加工挖掘转变。然而,由于一些资源如手稿、图片等,往

往体量巨大又具有复杂性和专业性等特性,使得仅仅靠技术手段去实现资源的大规模文本化和深度加工挖掘几乎是不可能的。随着 Web2.0 技术的兴起,图书馆 2.0 环境的形成,都为应对上述难题提供了技术保障和新的理念,为了应对这一难题,图书馆开始探索借助大众智慧,采用众包模式,试图走出一条馆藏资源建设开发的新路。国内外图书馆如 CADAL 数字图书馆、广州图书馆^[3]、澳大利亚图书馆、美

* 本文系国家社会科学青年项目“数字人文视角下特藏资源建设的众包模式驱动因素与激励设计研究”(项目编号:17CTQ001)的研究成果之一。

国会图书馆、伦敦大学学院图书馆等都已经开展了一些众包实践。从这些案例中不难发现,众包平台的构建是众包项目中至关重要的一环,但目前对图书馆众包的研究还主要以概念辨析和案例介绍分析为主,把众包平台构建作为重点内容进行具体探讨的较少。笔者在分析面向数字人文众包项目特征的基础上,从众包运营框架和流程入手结合国内外众包实践对面向数字人文的众包平台作用、构建模式和功能模块等关键问题进行了深入探讨,并以上海图书馆历史文献众包平台的实践为例,对其目标定位、构建模式、功能特色以及试运营情况等进行了具体分析,总结其构建经验并展望未来实际运营中可能存在的问题,以期为图书馆在众包平台构建与数字人文服务方面提供一定参考借鉴。

1 面向数字人文的众包平台概述

面向数字人文的众包项目区别于商业的众包以及其他一般的网络互动或用户内容生成(UGC)项目,一般具有以下特征^[4]:(1)针对一个可以被定义的人文方向或人文问题;(2)对原始资料的价值具有转化或增值的可能;(3)任务可以被定义和拆分;(4)众包活动对于不同量级的数据和不同层次的人群具有可扩展性。虽然面向数字人文的众包项目有上述固有特征,但在众包平台作用、构建模式和基本功能模块上与一般众包平台仍有相通之处,笔者从众包的运营框架入手,结合国内外案例,对这些众包平台的关键问题进行归纳总结。

1.1 众包运营框架与众包平台作用

众包概念由杰夫·豪(Jeff Howe)在2006年首次提出^[5],指一个公司或机构把过去由员工执行的工作任务,以自由自愿的形式,外包给非特定的(通常是大型的)大众网络的做法^[6],其核心思想就是利用群体的智慧和力量完成个人或机构无法或难以完成的任务^[7]。这一思想在商业领域得到应用后,也开始受到其他领域包括数字人文领域学者的关注。众包项目不论是在商业领域还是其他领域中的运作,其运营框架是基本一致的。一般认为一个众包项目运营框架由3个主体构成,包括发包方、

接包方和众包平台^[8-11]。也有研究者在3个主体基础上增加第三方机构作为主体^[12],并将数据管理、激励培训等职能从众包平台中剥离,但鉴于第三方机构的主体职能最终还是需要通过众包平台加以实现的,所以笔者认为不论是3主体还是4主体框架,众包平台都处于核心地位,发挥重要的作用。

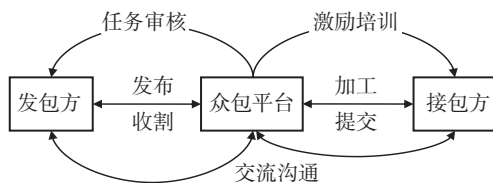


图1 一般众包的运营框架

以3主体的框架为例(如图1所示),可以发现众包平台在发包方与接包方中间起到一个连接中介的角色,具体作用可以概括为:

(1)提供发包方与接包方参与众包的空间。众包主要构建于网络之上,需要一个网络上的虚拟空间承载发包方发包、接包方加工等众包行为。

(2)成为发包方与接包方、接包方与接包方之间沟通的桥梁。众包项目要顺利开展,主体间的沟通必不可少,特别是发包方与接包方的交流更为重要,众包平台在发包方与接包方之间起到关键的沟通桥梁的作用。

(3)承担众包任务审核与接包方激励培训等管理职能。随着众包平台的开放发展,众包平台需要根据自身定位,对发包方提供的众包任务进行审核,对任务发布进行有效的管理。同时,也要根据任务的差异做好接包方的培训与激励,吸引更多志愿者参与到众包项目并有效地完成众包任务。

(4)降低发包方与接包方信息交换、协作加工的成本。众包平台最终是要起中介的作用,对于发包方来说众包平台可以减少其寻找大量合适的任务志愿者的成本,对于接包方来说可以降低寻找到自己感兴趣并适合自己参与的众包任务的成本。

1.2 众包平台的构建模式

近年来,众包项目在数字人文领域已有不少实践案例,图书馆在其中扮演重要角色,一

般图书馆就是项目的发起者与平台构建者。这些面向数字人文的众包平台构建模式不尽相同，经过调研，笔者将这些众包平台的构建模式归纳为嵌入式构建、第三方平台构建和独立构建三种模式。这些模式各具特点，适用的众包内容也不尽相同（如表1所示）。

嵌入式构建模式是指众包平台嵌入图书馆原有资源服务平台或系统中，没有明显的众包平台边界，这种构建模式使得众包加工与内容服务融为一体，但无法同时适用于多种类型的资源，每种类型资源的众包都需重新开发平台，也很难发挥众包平台的中介作用。这种构建模式的代表有澳大利亚图书馆的报纸文本化众包项目^[13]和CADAL描述信息修订众包项目^[14]等。

第三方平台构建模式是指图书馆通过直接入驻第三方平台建立众包项目，这种构建模式开发成本相对较低，直接利用第三方平台本身强大的社会交互功能，对于照片、图片等资源的评注比较有利，但这一构建模式所适用的众包内容相对单一，功能也受限于第三方平台本身可提供的功能，众包平台中介性也很难体现。其代表有大英图书馆^[15]、美国国会图书馆^[16]等在Flickr平台上建立的图片众包项目等。

独立构建模式是指图书馆专门针对众包项目建立一个平台或网站，与本身的服务平台有

一定联系又保持相对独立，这一模式具有更好的扩展性，但该模式需要重新开发一个完整的平台，成本相对较高。此外，该模式更偏重众包加工功能，众包内容的服务可能受到一定影响。该模式的代表有英国伦敦大学学院图书馆的边沁手稿文本化项目^[17]和美国纽约公共图书馆的菜单摘抄项目^[18]等。

1.3 众包平台的基本功能模块

在众包项目运行过程中，众包平台要发挥其作用主要是通过具体功能模块实现的，运行过程中每个环节需要体现众包平台的不同功能模块。

一个众包项目从启动到加工完成再到提供内容服务，一般可以分为五步闭环过程，分别是众包需求提出、组包发包、接包加工提交、数据回收整合以及内容服务，经过整合的内容可以在专门的服务平台上提供服务，在服务过程中可能产生新的需求，开启新的众包项目，这样就形成了众包加工服务的闭环过程（如图2所示）。

根据众包项目的一般过程可以归纳出众包平台所需构建的基本功能模块，包括组包发包模块、加工模块、回收整合模块和管理模块。

组包发包模块主要功能包括资源和相关元数据导入功能、组织任务包以及发布任务等功能。

表1 众包平台构建模式归纳

构建模式	代表项目	众包内容	特点	不足
嵌入式构建	澳大利亚图书馆的报纸文本化项目和CADAL描述信息修订众包项目	文本化和元数据校对为主	众包加工与内容服务一体化	无法同时适用于多种类型的资源；中介性不足
第三方平台构建	大英图书馆、美国国会图书馆等照片评注项目	照片、图片评注为主	构建成本相对较低；社会交互功能强	众包内容单一，功能受限；中介性不足
独立构建模式	英国伦敦大学学院图书馆的边沁手稿文本化项目和美国纽约公共图书馆的菜单摘抄项目	文本化、元数据校对、照片图片评注等皆可	平台边界明显，相对独立；扩展性较强	开发成本相对较高；内容服务不足

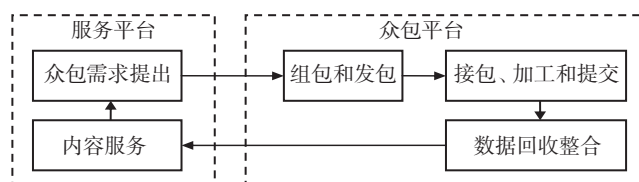


图2 众包项目一般过程

加工模块主要是为接包方提供在线浏览任务、选取任务、加工及提交任务的功能,对于不同的资源类型和加工要求,加工模块的功能设置会有相应变化。

回收整合模块主要功能是将众包加工完成的数据进行质量控制,整合已完成加工的数据,形成服务内容。

管理模块贯穿众包过程始终,主要功能包括面向发包方提供对用户权限、激励管理、用户行为、数据统计等管理功能。

2 面向数字人文的上海图书馆历史文献众包平台实践

2.1 概述

近年来,上海图书馆在数字人文方面有不少实践,取得一定成果,如基于关联数据技术建设的家谱数据库、名人手稿档案库和盛宣怀档案知识库等服务平台。然而要进一步推动数字人文资源的建设,深入挖掘这些档案手稿资源,将其文本化是目前首要的挑战。目前主要的文本化途径是OCR(Optical Character Recognition,光学字符识别),但对于书写笔迹凌乱、风格迥异的手稿识别率并不理想。随着众包理念的深入人心,国外也涌现了不少成功案例,相关的系统开发和管理技术日益成熟,上海图书馆的档案手稿文本化开始转向众包模式,在明确众包平台定位并选定构建模式后,自2016年开始尝试搭建起,目前上海图书馆已完成上海图书馆历史文献众包平台(以下简称“上图众包平台”)[9]二期的构建,并于2018年10月末开始进行试运营,平台功能得到进一步优化,中介性开放性得到进一步提升,阶段性目标基本实现。

2.2 上图众包平台目标定位与构建模式选择

构建上图众包平台主要就是要解决海量数字人文资源建设中文本化的问题,从短期来看,就是要建立一个以档案、手稿等手写资源元数据著录和全文文本化为主题的众包平台,为深度揭示利用这部分数字人文资源提供保障;从长远来看,为了使平台成为数字人文重要基础设施之一,平台还将面向更广阔的应用场景,承载更加丰富、不同来源的资源以及多

种类型的众包加工任务,一方面要打破上海图书馆馆藏的限制,吸引外部各种来源的数字人文众包任务发布;另一方面要优化用户体验,吸引广大用户到平台参与众包,充分体现众包平台的中介特性,连通资源建设与服务,发挥数字人文基础设施的保障作用。

以上述众包平台的目标定位为基点来选择众包平台的构建模式,嵌入式构建模式和第三方构建模式是无法适应的。首先从资源角度来看,由于定位于多种文献资源和多样的加工类型,如文字的抄录、图片的标注、“人地时事”的深度标引等,嵌入式构建模式和第三方构建模式更适合相对单一的加工类型,无法适应这样的定位。

其次,从平台中介性的角度来说,想要发挥众包平台的中介作用,不单为自身馆藏的数字人文的标引服务,还要面向广大用户大众,提供众包任务发布功能,不仅吸引加工用户,也要吸引发布用户,使平台有一个双向集聚的效应,那么相对明晰的平台边界是必要的。嵌入式构建模式和第三方构建模式都无法很好地体现平台的中介性。

综合考虑目标定位、资源特点和成本等因素,最终上图众包平台选择了独立构建的模式来构建一个相对独立、能够面向数字人文、充分体现众包特点和优势的平台。

2.3 上图众包平台整体架构与功能模块特色

2.3.1 整体架构

上图众包平台整体架构主要由用户平台、管理平台、基础应用平台和支撑平台组成(如图3所示)。用户平台主要提供发布者组包发包功能、接包者加工交互功能以及审核者的审核功能等;管理平台主要提供全流程管理功能,包括数据管理、用户管理和意见管理等;基础应用平台和支撑平台主要对整个平台底层提供基础的技术支撑,包括内容缓存、负载均衡、服务器、数据库和操作系统等的部署与搭建。

2.3.2 平台功能模块

从众包流程来看,上图众包平台功能模块主要由组包发包、加工、回收整合和平台管理功能模块构成。

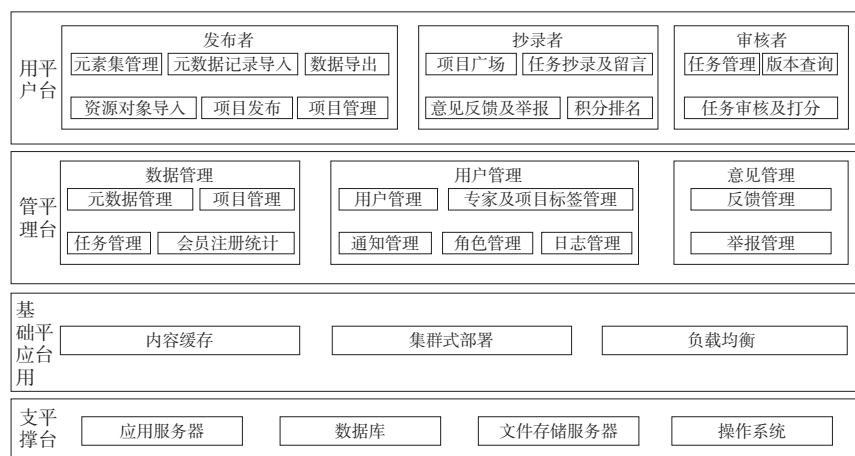


图3 上图众包平台整体架构图

元素集管理	添加			
		英文名	中文名	操作选项
	元数据记录导入	id	唯一标识	
	资源对象导入	filelocation	文件路径	
	我的发布	title	题名	
	任务审核及打分	content	内容	
	数据导出	amount	数量	 
	所有项目	annotations	附注	 
	所有任务	attachment	附件	 
	用户管理	contributor	捐赠者	 
角色管理		creator	责任者	 
日志管理				
注册用户统计				
标签管理				
我的抄录				
我的积分				

图4 平台元素集管理界面

2.3.2.1 组包发包功能模块

组包发包功能模块由文献上传、任务发布和任务管理3个子模块构成。以盛宣怀档案手稿抄录众包项目为例，文献上传子模块主要包括元素集设置、元数据导入以及图片资源关联上传等功能。其中，元素集设置功能需要具有灵活性供发包方选择或创建，这样才能使多种众包加工类型得以实现（如图4所示）；任务发布子模块主要是根据主题并以项目为单位将文献打包发布，提供全文抄录以及元数据字段抄录两种众包形式；任务管理子模块允许发包方设置奖励分值、任务难度、抄录次数、截止

时间等任务相关要素（如图5所示）。

2.3.2.2 加工功能模块

以盛宣怀档案手稿抄录众包项目为例，提供接包方对全文和元数据进行文本化及标注修改的功能为例。平台主要提供用户领取任务、抄录标引和保存提交的功能。二期平台进一步优化加工界面，提供分屏功能，使接包方加工操作体验更好。在加工过程中优化交互功能，用户可通过留言方式与平台管理者和发包方沟通。

2.3.2.3 回收整合功能模块

该模块主要实现用户提交数据的审核、打

分、退回、历史版本查看及数据入库等功能。审核功能主要通过人工审核完成,专家可以在审核列表中选择待审核的任务锁定并审核,也可将任务转发给其他专家。专家在审核过程中对不符合任务要求的数据可以选择退回或零分通过。退回任务需给出不足之处,以使用户修改。专家还可跟踪查看退回任务的历史版本,以便根据用户的留言反馈指导用户修改(如图6所示)。零分通过也将视为任务关闭,可由发包方重新发包形成新的任务。审核完成的数据

可由发包方导出利用。

2.3.2.4 平台管理功能模块

管理功能模块贯穿整个众包加工过程,实现数据管理、用户管理和激励管理等功能。数据管理主要对元数据、项目、任务等数据进行管理;用户管理主要对用户权限、角色、日志等进行管理;激励管理主要对激励机制进行设置,根据项目难度系数、加工数量以及专家打分3个因素给予对应积分,并采用了排行榜概念,将用户按积分排序,以此激励用户参与众

图5 平台发包设置界面

查看历史版本			
版本时间	元素数据	全文数据	留言
2020-05-25 20:31:21	题名: 汉冶萍公司辛亥军兴以来损失数目清摺 责任者: 汉冶萍公司收支所 地点: 辛亥革命; 财产; 汉冶萍公司; 汉阳铁厂; 大冶铁矿 日期公元: [1912年] 日期年号:	第1页 共15页 (甲)有形损失 一 漢陽鐵廠 一 損失頭號生鐵 噸 每噸 計洋例 合元三千五百五十九 二 損失二號生鐵 噸 每噸 計洋例 合元三千五百五十九 三 損失馬鋼軋 噸 每噸 計洋例 合元六千八百四十四 四 損失馬鋼軋 噸 每噸 計洋例 合元四千一百五十九 五 損失鋼板 噸 每噸 計洋例 合元三千五百五十九 六 損失鋼釘 噸 每噸 計洋例 合元一百一十五 七 損失魚尾板 噸 每噸 計洋例 合元九十二 第2页 共15页 一 漢陽鐵廠 一 損失頭號生鐵 噸 每噸 計洋例 合元三千五百五十九 二 損失二號生鐵 噸 每噸 計洋例 合元三千五百五十九 三 損失馬鋼軋 噸 每噸 計洋例 合元六千八百四十四 四 損失馬鋼軋 噸 每噸 計洋例 合元四千一百五十九 五 損失鋼板 噸 每噸 計洋例 合元三千五百五十九 六 損失鋼釘 噸 每噸 計洋例 合元一百一十五 七 損失魚尾板 噸 每噸 計洋例 合元九十二 第3页 共15页 (甲)有形损失 一 漢陽鐵廠 一 損失頭號生鐵 噸 每噸 計洋例 合元三千五百五十九 二 損失二號生鐵 噸 每噸 計洋例 合元三千五百五十九 三 損失馬鋼軋 噸 每噸 計洋例 合元六千八百四十四 四 損失馬鋼軋 噸 每噸 計洋例 合元四千一百五十九 五 損失鋼板 噸 每噸 計洋例 合元三千五百五十九 六 損失鋼釘 噸 每噸 計洋例 合元一百一十五 七 損失魚尾板 噸 每噸 計洋例 合元九十二 八 損失焦炭	退回: 抄录不完整, 请认真抄录, 重新提交, 谢谢。 91149554 2020-05-25 *“噸”前的字不太认识, 请专家指教, 谢谢!
2020-05-25	题名: 汉冶萍公司辛亥军兴以来损失数目清摺 责任者: 汉冶萍公司收支所	第1页 共15页 (甲)有形损失 一 漢陽鐵廠 一 損失頭號生鐵 噸 每噸 計洋例 合元三千五百五十九 二 損失二號生鐵 噸 每噸 計洋例 合元三千五百五十九 三 損失馬鋼軋 噸 每噸 計洋例 合元六千八百四十四 四 損失馬鋼軋 噸 每噸 計洋例 合元四千一百五十九 五 損失鋼板 噸 每噸 計洋例 合元三千五百五十九 六 損失鋼釘 噸 每噸 計洋例 合元一百一十五 七 損失魚尾板 噸 每噸 計洋例 合元九十二 八 損失焦炭	

图6 历史版本审核界面

包加工。

2.3.2.5 平台功能模块特色

上图众包平台的功能特色主要体现在:

(1) 平台开放, 支持不同机构或个人发包。上图众包平台二期, 对发包机制和相应功能进行了改革。允许组织或个人在平台上发布任务, 从而打破了平台相对封闭、发包方单一的状态。更多数字人文众包项目的引入有助于吸引更多不同层次、不同兴趣的用户加入数字人文的众包项目, 使上图众包平台更好地发挥中介的作用, 为整个数字人文基础资源建设做出贡献。

(2) 支持不同众包加工类型, 灵活的元素集管理。上图众包平台摆脱仅支持单一众包加工类型的局限, 现已支持元数据抄录和全文抄录两种加工类型。在组包发包阶段提供元素集管理、积分设置、任务认领数量等方面的设置, 特别是元素集管理提供发包方创建并灵活配置元素集的功能, 该功能保障了多种加工类型得以实现, 也给予发包方在任务设计上一定的调整空间, 从而丰富众包项目的内容和形式, 这也体现了上图众包平台的自由开放性。

(3) 引入多人模式, 完善质量保障机制。上图众包平台一期, 一个任务仅允许一人接包, 由于用户良莠不齐, 质量难以保证, 审核成本较高。平台二期引入多人模式, 以众包抄录任务为例, 多人可抄录同一手稿档案, 这样最后众包结果可以由多份抄录内容综合得到, 从而提高抄录准确性。由于引入多人模式, 审核功能也进一步优化, 审核专家可以对历史版本进行比对, 以便查看抄录者修改情况并给予指导。

2.4 上图众包平台试运营情况

上图众包平台二期对一期进行了一定改进, 优化了平台设计与功能, 2018年10月末开始试运营。试运营首先要解决缺乏初始用户的问题, 也就是平台的冷启动问题, 这也是众包项目启动初期普遍要面临的问题。为此, 上图众包平台通过与南京大学联合举办抄录竞赛的方式解决冷启动的问题, 竞赛历时3个月, 吸引文学、历史学、美术学、计算机等多个专业来自不同高校的60余名学生及社会人士参

赛, 累计完成任务225项。通过竞赛所吸引的这部分用户, 对于平台冷启动至关重要, 他们一定程度上成了“种子用户”, 可以为平台推广, 吸引更多用户提供帮助, 截止到2019年7月, 平台注册用户增至346人。但也应看到由于试运营并没有将冷启动范围面向全社会, 吸引到的用户还有一定局限, 主要来自高校的师生, 来源相对单一。通过平台冷启动过程, 也测试了众包平台运营的整个流程和功能模块。从试运营情况来看, 上图众包平台二期实现了包括档案、手稿等手写资源元数据和全文文本化等多类主题任务发布的功能, 也提供馆外不同发包主体发包的功能, 并引入多人模式保障众包加工质量。总的来说, 上图众包平台二期基本完成了平台构建的短期目标, 也在开放性和中介性等方面做了一定探索, 为实现平台的远期目标打下坚实的基础。

3 结论与展望

笔者根据众包运营框架和过程, 对面向数字人文的图书馆众包平台在国内外实践的情况进行考察, 归纳出目前面向数字人文众包平台主要的构建模式以及基本的功能模块, 并通过上图众包平台的实践加以具体分析, 得到以下有益经验:

首先, 构建众包平台之前, 明确构建的目标定位十分必要。众包平台构建模式与功能设置都与平台的目标定位密切相关。所以, 众包平台设计与开发者必须根据自身情况先明确平台目标定位, 只有这样才能选择适合平台自身发展的构建模式并设置相应的功能模块。

第二, 平台运营之初, 冷启动问题是无法回避的, 如何找到初始的“种子用户”值得众包平台运营方好好思考。上图众包平台采用同高校合作竞赛的方式来解决冷启动问题, 这一方式未必适合所有众包平台, 不同的众包平台还需要根据自身涉及的资源类型和面向的用户对象类型等因素进行综合考虑, 进行方案设计, 来解决平台的冷启动问题。

第三, 从众包运营的框架来看, 众包平台处于核心位置, 是连接众包各方的桥梁, 所以众包平台需要着力提升自身的中介性和开放

性。上图众包平台通过优化发包功能和机制,支持不同机构或个人发包,并支持元素集灵活管理,使发布不同加工类型的众包任务成为可能。这都是为提升众包平台的中介性和开放性而做出的有益探索,有助于未来平台引入更丰富多样的众包任务,更好地吸引不同用户群体参与众包,发挥其数字人文基础设施的作用。

众包平台要发挥群体智慧的优势,最终必将面向广阔的社会大众,小范围的试运营很多平台构建的问题可能还难以完全凸显出来,比

如如何细分众包项目以适应不同喜好和能力的用户,如何优化排行榜、设计游戏机制激励用户更多参与众包,如何优化平台社交功能促进各方交流,如何完善众包任务回收机制提高众包加工质量和效率,如何扩大平台开放度吸引更多来自不同机构或个人发包者发布多样的任务,如何同其他数字人文服务平台协同融合等,这些问题涉及众包平台运营的各个环节,关乎平台的长远发展,还有待学者和图书馆同仁一同研究探索。

参考文献

- [1] 杜宗明. 数字人文环境下的图书馆角色定位与实践路径[J]. 农业图书情报学刊, 2017, 29(10): 103-106.
- [2] 刘伟, 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] 颜运梅. 众包在国内古籍数据库建设中的应用研究[J]. 图书馆研究, 2016, 46(5): 30-34.
- [4] Hedges M, Dunn S. Crowd-Sourcing Scoping Study: Engaging the Crowd with Humanities Research[J]. International Journal of Humanities & Arts Computing, 2012.
- [5] HOWE J. The rise of crowdsourcing[J]. Wired, 2006, 14(6): 176-183.
- [6] 张利斌, 钟复平, 涂慧. 众包问题研究综述[J]. 科技进步与对策, 2012, 29(6): 154-160.
- [7] 李书宁, 曾姗. 国外图书馆数字馆藏众包建设实践调查与分析[J]. 图书情报工作, 2014, 58(23): 83-90.
- [8] 夏恩君, 赵轩维, 李森. 国外众包研究现状和趋势[J]. 技术经济, 2015, 34(1): 28-36.
- [9] 庞建刚, 刘志迎. 科研众包参与主体及流程的特殊性[J]. 中国科技论坛, 2015(12): 16-21.
- [10] Zhao Y, Zhu Q. Evaluation on Crowdsourcing Research: Current Status and Future Direction[J]. Information Systems Frontiers, 2014, 16(3): 417-434.
- [11] 顾姝姝, 陈曦. 众包平台研究综述与众包平台绩效影响机制构建[J]. 科技进步与对策, 2017, 34(22): 153-160.
- [12] 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.
- [13] Australian Newspaper Digitisation Program [EB/OL]. [2019-02-17]. <https://www.nla.gov.au/content/newspaper-digitisation-program>.
- [14] CADAL [EB/OL]. [2019-02-17]. <http://www.cadal.zju.edu.cn/index>.
- [15] Flickr archive [EB/OL]. [2019-02-17]. <https://www.flickr.com/photos/britishlibrary/>.
- [16] Library of Congress Photos on Flickr [EB/OL]. [2019-02-17]. http://www.loc.gov/rr/print/flickr_pilot.html.
- [17] TRANSCRIBE BENTHAM A PARTICIPATORY INITIATIVE [EB/OL]. [2019-02-17]. <http://blogs.ucl.ac.uk/transcribe-bentham/>.
- [18] What's on the menu? [EB/OL]. [2019-02-17]. <http://menus.nypl.org/>.
- [19] 上海图书馆众包系统 [EB/OL]. [2019-02-17]. <http://zb.library.sh.cn/>.

姚啸华 上海图书馆(上海科学技术情报研究所), 馆员。研究方向: 数字人文。作者贡献: 选题、论文撰写。E-mail: xhyao@libnet.sh.cn 上海 200031
贺晨芝 女, 上海图书馆(上海科学技术情报研究所), 助理工程师。研究方向: 数字人文。作者贡献: 提供资料、修改论文。上海 200031
徐孝娟 女, 安徽大学管理学院, 副教授。研究方向: 网络信息资源管理。作者贡献: 论文修改。安徽合肥 233042
全石峰 上海图书馆(上海科学技术情报研究所), 高级工程师。研究方向: 图书馆信息技术。作者贡献: 论文修改。上海 200031

(收稿日期: 2019-03-01 修回日期: 2020-02-04)