

ChinaXiv预印本服务平台构建*

王颖¹, 张智雄^{1,2}, 钱力¹, 叶志飞¹, 王玉菊¹, 黄金霞¹, 刘静羽¹, 陈雪飞¹

(1.中国科学院文献情报中心, 北京 100190; 2.中国科学院武汉文献情报中心, 武汉 430071)

摘要: 为满足保护作者首发权, 促进中国优秀科研成果率先保存并服务于中国科技界的需要, 中国科学院科技论文预发布平台 (ChinaXiv) 开展项目建设, 旨在鼓励科研人员第一时间在公共预印本平台上公布科研成果, 通过该平台听取同行的评议和建议, 促进更大范围的学术交流, 并推动科研成果的开放获取。本文以工作实践为基础, 从建设背景、系统建设及服务、领域拓展与培育、组织管理机制、宣传推广模式、建设现状以及未来发展等多个方面介绍ChinaXiv如何构建国际通行模式规范运营的预印本服务平台, 为其他预印本平台建设提供参考。

关键词: ChinaXiv; 预印本; 开放获取; 学术交流

中图分类号: G258

DOI: 10.3772/j.issn.1673-2286.2017.10.004

随着开放获取、开放研究、开放数据、开放创新的不断发展, 重要科技大国陆续推出国家开放科技政策, 推动公益经费资助的科研成果更方便地开放获取。目前, 基于领域 (主题) 的论文预印本系统和已发表论文的开放存档系统日益得到关注, 一些重要的领域 (主题) 数字仓储平台得到广大科研团体、科技期刊界、文献数据库出版商和数字图书馆界的认同, 与传统基于期刊的成果发布方式形成有效互补, 成为被科技界认可的一种新型学术交流平台。如量子物理学家Ginsparg于1991年创建的arXiv开启了电子预印本时代, 目前, arXiv已经成为物理学、数学、计算机科学、计量生物学、统计学等领域预印本存缴和发布的权威平台^[1], Perelman对庞加莱猜想的证明成果就发布在arXiv平台。SSRN (the Social Science Research Network) 致力于促进社会科学研究在全球范围内的迅速传播^[2], 鼓励科研人员将早期科研成果通过提交摘要和研究论文在全世界发布, 成果涉及领域包括财经、会计、法律、经济、管理等。作为一个免费的社会科学论文存储开放平台, SSRN受到广大学术研究者与爱好者的欢迎与支持。2013年, 非营利机构美国冷泉港实验室启动BioRxiv^[3], 旨在将其打造成生物学领域专有的预印本

文献库。与传统出版模式不同, 作者在写作完成后, 先将草稿预公开供同行审阅并提出更改建议, 而后再提交至学术期刊进行出版。此外, 一些国内外预印本平台, 如心理学领域的PsyArXiv、农业领域的AgrXiv、工程领域的engrXiv以及化学领域的ChemRxiv等也先后被推出。

2016年, 中国科学院启动“中国科学院科技论文预发布平台 (ChinaXiv)”项目, 由中国科学院传播局组织实施, 中国科学院发展规划局提出具体指导, 中国科学院文献情报中心承担建设, 相关研究所和期刊编辑部共同参与。经过一年多建设, ChinaXiv已在科技界产生一定影响, 得到广大科研人员和学生的关注, 同时获得一些知名科学家的大力支持。

1 ChinaXiv建设背景

1.1 建设需求

预发布平台的建设是促进中国优秀科研成果率先保存并服务于中国科技界的需要。目前, 我国最新的科研成果多优先在国外发表, 即使是开放的预印本也存储

* 本研究得到中国科学院2016年度传播项目“中国科学院科技论文预发布平台” (编号: 院1617) 资助。

在国外预发布平台,不能及时有效地被我国科研人员使用。将中国优秀科研成果率先保存并服务于中国科技界,对保障我国优秀科研成果的首发权认定,促进我国优秀科研成果的知识产权保护,提高我国科研成果可见度和竞争力等都具有重要意义。

预发布平台的建设是推动中国科研成果开放获取的需要。对科研人员、科研机构、政府及其科研资助部门而言,保证科研成果的公共可获得性,已经成为一项重大的科学和社会责任。建设本国的开放式、支持先发布后发表的知识库,实现科研论文的开放预发表和已发表论文的开放存档,将有效地推动科研成果的开放获取。

预发布平台的建设是促进我国相关学术领域开放学术交流机制形成的需要。通过开放评论等即时交互的学术交流机制,促进交流与合作,激发创新思维,有利于产出更多的优秀科研成果。通过相关机制的建设,协调好科研人员、期刊、学协会间的关系,搭建三者间的桥梁,加速和改进科学传播,解决资源协调和合作问题,有助于形成良好的学术交流生态环境。

预发布平台的建设是有效保护作者首发权和监控学术不端行为的需要。由于科研工作存在一定的竞争机制,为保护作者首发权,需要防止科研人员的学术成果被窃取或盗用。预发布平台通过公开发布、开放评论等方式,避免科研成果在同行评议中被评审专家剽窃,让相关领域的科研成果能在较大范围内被科学界公开监管、广泛交流,以此来监控学术不端行为,构建公开公正的学术交流环境。

1.2 目标与定位

ChinaXiv作为中国科学院科技期刊“十三五”规划和中国科学院科技期刊改革与发展率先行动计划的一个重要部分,以中国科学院为基础,面向全国科研人员接收中英文科学论文的预印本存缴和已发表科学论文的开放存档服务。致力于构建一种新型的学界自治的科研成果交流和共享平台,与传统基于期刊的成果发布方式形成有效互补。对鼓励科研人员公平竞争,保障优秀科研成果首发权的认定,推动科研成果的开放获取,促进更大范围的学术交流,构建公开公正的学术交流环境,助力科技期刊和科研出版的集约化、国际化发展有重要意义。

ChinaXiv的建设目标包括:(1)建立“开放存缴+质

量检查+学术交流+分发传播+开放评阅”的服务机制,通过可靠的首发权登记、即时开放获取、丰富的作者和读者服务、有效的期刊服务,最终建成国内科研人员在本土的可靠、规范、权威的论文预印本首发和快速传播平台,实现科研论文预印本的开放存档和发布;(2)建立有效的读者和期刊服务机制、组织管理机制、宣传推广和可持续运行机制,嵌入学术交流工具,最终将ChinaXiv建成被中国科技界广泛认可和在国际上有一定影响力的新型学术交流平台;(3)实现与中国科学院科技期刊的合作支撑,以合作期刊的模式建立与期刊的关联,将优秀的预印本资源纳入中国科学院期刊采编系统,有效支撑中国科学院科技期刊建设工作,提供多种便捷服务,帮助科研人员发表和传播其科研成果,帮助科技期刊出版商第一时间获得优质科研论文。

2 平台建设及服务

ChinaXiv服务平台的基本要求是支持中英文版论文存缴,支持多种论文提交方式,通过自动检查和关联匹配辅助论文管理人员对提交的论文进行审查管理,保证论文质量;提供多维度的论文浏览,让用户随时跟踪关注的主题和文章定制。ChinaXiv基于开放获取的理念,深入整理分析预印本的资源类型及特点,根据科研人员需求和预印本平台实际的服务方式、服务功能和服务分布要求,制定具有通用性、开放性、可扩展、模块化的总体建设方案,方便平台和第三方系统进行有效集成,最大程度发挥预印本平台的资源与服务优势,整体框架如图1所示。

(1)面向作者建立论文提交系统、提交审查管理平台、自动检查和关联引擎。论文提交系统要支持多种提交手段,如邮箱提交、在线批量提交,解决作者身份确认、授权许可确认、诚信承诺确认等细节问题;提交审查管理平台支持论文自动合规审查、自动重复性检查、自动剽窃内容检查、自动作者识别和关联作者识别、自动分配唯一标识等,方便质量管理人员进行论文审查操作。

(2)在预发布论文数据基础上建立领域仓储,面向所有自然科学领域开放,接收全领域的中英文预印本论文。此外,建立相关领域的学协会、期刊、出版机构等预印本子平台。ChinaXiv平台作为公共服务平台,论文无需订购付费,直接访问可随意检索、浏览、获取所需科研成果;为用户提供论文评论等服务功能,支持在线学术交流及论文评阅。

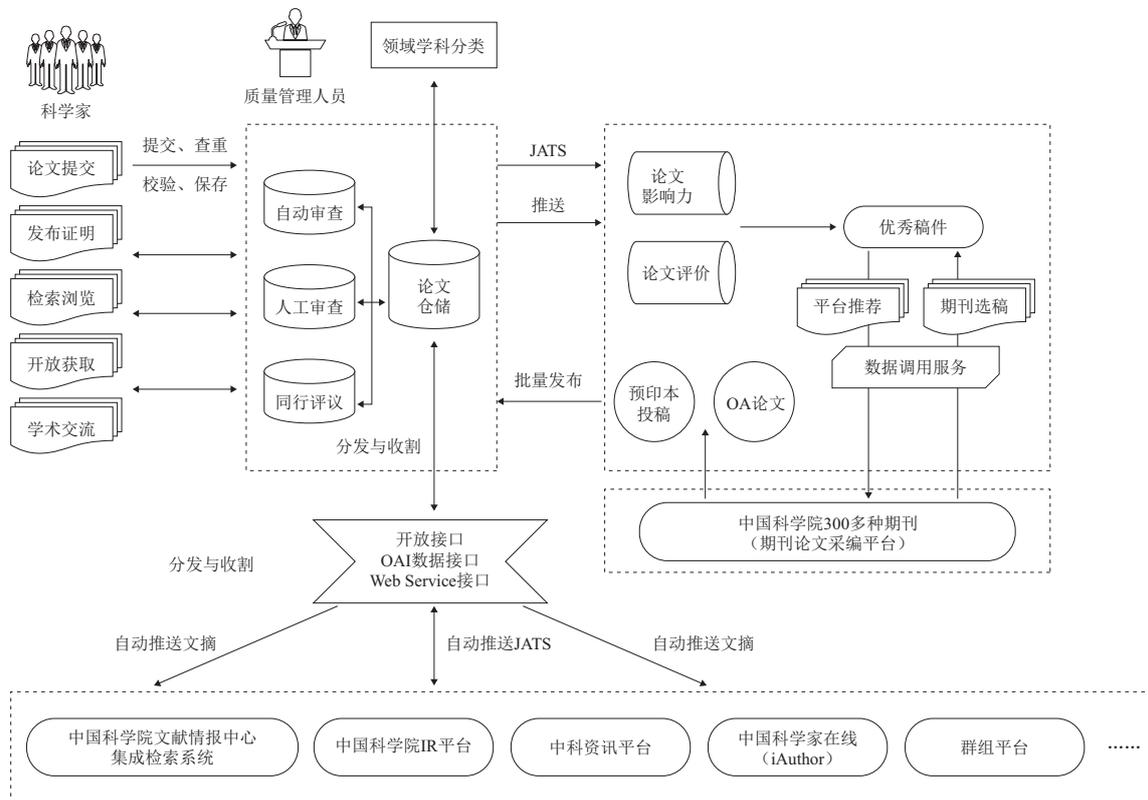


图1 ChinaXiv建设框架

(3) 建立期刊合作机制，为预印本论文与期刊进行技术对接。ChinaXiv为合作期刊分别建立对应期刊预印本子平台，并根据论文的影响力和评价向期刊推荐优秀论文。各期刊可对其子平台进行自定义设置和论文管理。

(4) 面向第三方系统建立开放服务接口、摄入接口工具、分发接口工具和基础设施公共接口。ChinaXiv作为开放仓储平台，基于OAI协议提供有效的OAI开放接口，能够将ChinaXiv平台数据及应用嵌入第三方平台（如中国科学院文献情报中心集成检索平台、中科资讯平台等）。ChinaXiv充分集成并采用研究基础设施公共接口（如作者唯一标识符），以有效利用公共开放获取资源。

3 领域开拓与培育

基于预印本的领域（主题）仓储已经成为一种重要的学术交流渠道，因此ChinaXiv将建设领域仓储作为重点环节，而领域的开拓和培育对ChinaXiv的建设和发展至关重要。如arXiv的成功与其以高能物理领域为开局不无关系，现在arXiv已经成为物理学、数学、

计算机科学、计量生物学、统计学等领域的重要学术交流基础设施，这些领域的科研人员对预印本认可度普遍较高^[4]，以此为基础开展预印本领域仓储建设能吸引一定的受众群体。

我国是天文学、生物学、数学、材料科学领域的研究大国，在开放获取、知识共享方面有很好的学术传统，而中国科学院在这些领域具备良好的学科基础。因此，ChinaXiv在建设初期集中开拓和培育物理学、计算机科学、天文学、生物学、数学、材料科学等领域的论文存缴。在此基础上，ChinaXiv充分考虑这些领域的科学家、学协会、期刊/出版机构、主要研究科学团队、主要科研人员的基础和主要相关政策来开始工作，建立起与上述领域相关的学协会、期刊/出版机构、科技媒体、重要科研团队的交流合作机制。

在发展建设过程中，ChinaXiv在上述领域基础上陆续开拓了空间科学、工程热物理学、海洋科学、能源科学等领域。随着宣传与推广工作的开展，科研人员对于预印本的认可度不断提升，一些科研人员也表达了对于首发权认定和学术交流的需求。因此，ChinaXiv现已面向自然科学领域接收中英文科技论文预印本存缴，通过不断积累作者提供的数据和开放数据，逐步建立

可靠权威的领域仓储。

4 组织管理机制

在组织管理机制建设上, ChinaXiv参照arXiv模式采用国际通行的预发布平台管理机制^[5-6](见图2)。在管理咨询层面, 由中国科学院和国内知名高校、图书馆的25位专家组成ChinaXiv管理委员会、科学咨询委员会和用户咨询委员会。管理委员会指导预发布平台的

运行, 科学咨询委员会指导仓储的学术领域规划, 用户咨询委员会指导ChinaXiv的服务和应用。同时, 在管理委员会下组建运营工作组, 专项负责对平台系统功能和数据建设的常规运营。运营工作组又细分为系统运维组、论文质量管理组、政策研究组、用户服务组, 负责对预发布平台系统构建、平台保障, 论文质量跟踪, 研究解决预发布平台在推广中遇到的政策和法律相关问题、平台各方权益问题, 同时负责开展相关领域用户培训和推广宣传工作。

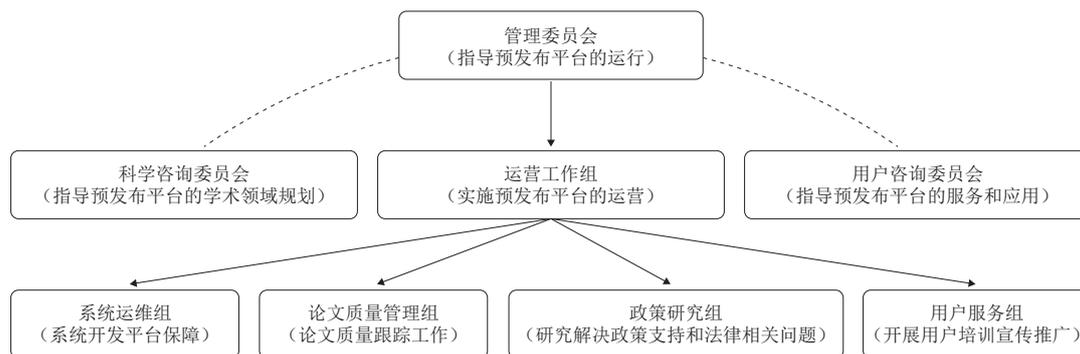


图2 ChinaXiv的组织机制

在业务组织和服务支持机制方面, ChinaXiv安排运营工作组相关人员支持预发平台的业务工作和服务工作。具体而言, 系统运维组需对平台功能建设和数据建设进行常规运营维护, 定期更新和维护版本建设, 对系统的硬件运行环境和运行安全进行定期监测和维护, 做好数据备份工作及系统推广后数据量激增的应对预案等工作。论文质量管理组对用户提交的论文进行质量审核, 协助科研人员上传论文至ChinaXiv平台, 保障科研人员论文的即时发布, 并对平台数据进行有效地组织、监控和管理, 保证数据准确有效。

在论文质量管理机制方面, 由论文质量管理组保障预发布论文的质量, 杜绝非科技论文的发布。审核团队包括初审人员2—3名, 终审人员1名, 负责预发布论文的及时审核与发布。基本的质量控制机制包括论文准入审核(研究者身份识别、研究机构识别、是否在同领域有同行评审论文)及准入论文的格式、重复性、剽窃性和研究性审核, 保证论文的学术性。当论文的研究性、科学性、作者研究身份等无法识别时, 启用专家评审流程。组建领域评审团队, 由2—3位专家提供审核意见, 由审核团队终审人员作出是否符合收录范围的判断。此外, ChinaXiv通过即时公开、发布、开放评论等方式, 让平台发布的科研成果能在较大范围内被科学

界公开监管, 广泛交流, 以此来监控学术不端行为。

5 宣传推广模式

我国当前科技工作对预发布平台的认识还处于初级阶段, 有效调动作者积极性对平台的建设作用巨大, 因此要加强预发布平台的宣传推广力度。项目组制定了三层宣传推广模式(见图3), 在底层通过优化系统平台使用功能和加强与其他平台的对接, 方便用户操作; 通过多种渠道获得科研人员对ChinaXiv的认同; 通过招募推广大使等办法发动科研人员力量, 开展推广宣传。在中层通过内部发力与外部联合共同开展立体化、多形式的宣传推广。内部发力具体推广方式包括制作宣传单、宣传海报、宣传视频、Newsletter等, 并通过微信宣传、电子邮件投递、直接发放张贴、重要视频网站播放等方式长期或定期投放; 自动组织策划大型活动, 如每季度举办一次大型主题活动, 每月持续开展面向研究所、期刊、学协会、高校的培训与宣传活动, 制定“百所行”“高校巡回宣传”计划, 直接深入科研一线进行宣传演讲, 鼓励科研人员分享科研成果; 从几个重点领域扩展至自然科学全领域, 扩大平台受众群体。外部联合, 一方面联合研究所、期刊、学协会、高校, 通

过签订子任务书或合作协议等方式，建立合作关系；另一方面，由合作方开展或配合开展推广宣传活动，采取嵌入式宣传途径，如借助合作方网站、专版、邮件列表等宣传渠道，参与合作方学术活动或会议等。在同行合作方面，积极地与其他预发布平台（如arXiv、bioXiv、SSRN等）洽谈合作，努力实现合作共赢。此外，积极地

与用户咨询委员会、科学咨询委员会成员交流，发挥科研人员学术影响力，获取用户服务建议，提升平台影响力。在上层政策机制方面，努力争取国家部委出台预印本相关政策，积极组织和策划国家预印本平台重大项目，联合学协会、科学技术协会提高学术界预印本认知度，打造国家层面预印本平台。

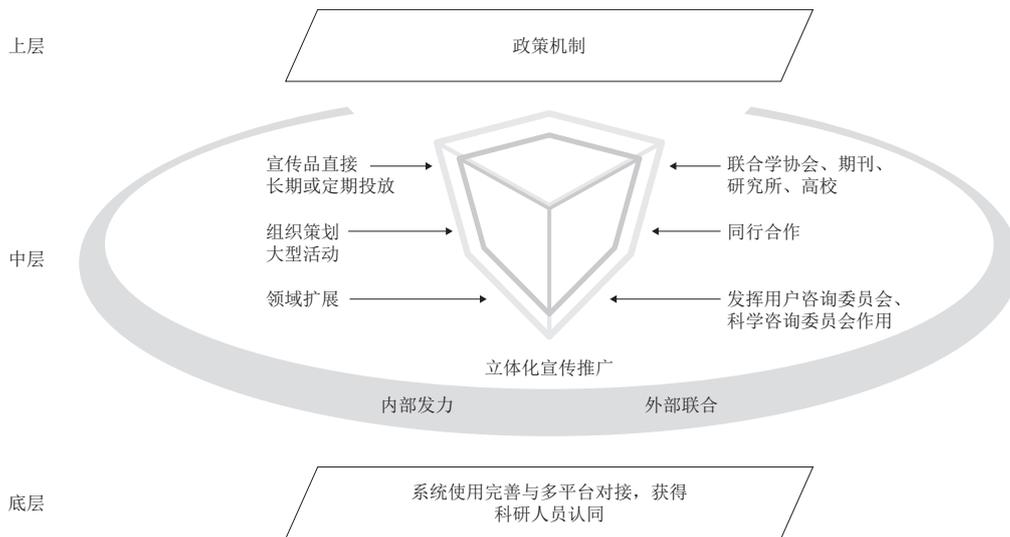


图3 立体化宣传推广模式

6 平台建设现状

6.1 平台使用情况

ChinaXiv平台提供一站式查询检索及全文下载服务，目前接收论文6 689篇，审核通过3 964篇；如图4所示，主要分布在物理学、生物学、材料科学、地球科学领域，其中物理学领域1 548篇（占比39.0%），生物学领域882篇（占比22.2%）。平台累计用户43 841个，其中有2 449名科研人员进行了论文存缴；中国科学院院内有54家研究所的科研人员在ChinaXiv平台提交了论文，院外共242家机构的科研人员在ChinaXiv平台进行论文提交，如开普半岛科技大学、美国杜克大学、美国纽约州立大学、劳伦斯伯克利国家实验室等机构的人员均在平台提交了论文。截至目前，平台累计总浏览量126 197次、总下载量62 208次，大幅提高预印本论文的被发现率和被引用率。

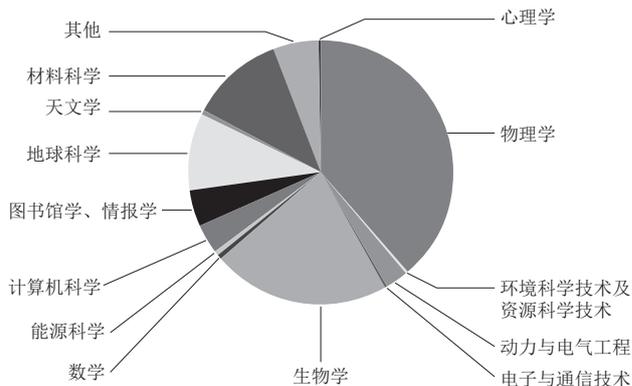


图4 各领域存缴论文情况

6.2 平台推广合作进展

目前，已经有中国科学院高能物理所、国家空间科

学中心等15个研究所通过子项目正式加入ChinaXiv建设。合作期刊包括《中国科学院院刊》《金属学报》《生态学报》《工程热物理学报》《中国生物工程杂志》《古脊椎动物学报》等11种。在推广合作中定期组织交流会，以便合作的研究所、期刊间相互借鉴经验，推动共同发展，促进合作机制的建立。

7 未来发展

在项目组、合作研究所、合作期刊、学协会的共同

努力下, ChinaXiv的建设工作已取得一定成效。下一阶段, 项目组将从以下两个方面继续推动ChinaXiv的发展与建设。(1) 积极探索可持续发展机制。一方面, 进一步加强与期刊的合作, 促进中国科学院期刊资源的有机集成、整合和关联, 同时积极拓展与学协会、高校的合作, 通过积极协调、多方合作积极争取和扩大服务范围, 提升科研人员对预印本的认识度, 促进科研人员学术交流; 另一方面, 通过有用的平台、简便的系统、优质的服务, 逐步建成被国内科技界广泛认可、在国际上有一定影响力、权威的新型学术交流平台, 更加突显ChinaXiv在国际科研生态中的作用, 通过支撑国家、研究机构的战略需要, 争取公益性经费的支持。(2) 以ChinaXiv为基础, 项目组将致力于建设中国科学院科技论文开放获取平台。基于中国科学院文献情报中心现有平台和工具(如iSwitch、IR Grid、GoOA、iAuthor等)建立中国科学院科技论文开放获取平台, 支持科技工作者开放获取国内科研人员的科技论文预印本、中国科学院科技期刊论文, 及科研人员在海外发表的论文、机构仓储重要科研成果、国外高水平开放获取期刊论文等资源, 并建设科学社交平台, 提供科研人员间便捷

的交流共享服务。

致谢: 本文撰写基于ChinaXiv项目成果, 项目建设与发展得到中国科学院传播局、中国科学院发展规划局、中国科学院文献情报中心、合作研究所、合作期刊、合作学协会等机构的大力支持, 由于篇幅所限无法一一列出, 在此, 谨对参与项目规划和建设的各位领导, ChinaXiv管理委员会、科学咨询委员会、用户咨询委员会、运营工作组的各位成员, 及对ChinaXiv提出宝贵意见的各位专家和老师等表示感谢。

参考文献

- [1] GINSPARG P.As we may read[J].The Journal of Neuroscience,2006, 26(38):9606-9608.
- [2] SSRN(the Social Science Research Network)[EB/OL].[2017-09-15].https://www.ssrn.com.
- [3] BioRxiv[EB/OL].[2017-09-15].https://www.biorxiv.org/about-biorxiv.
- [4] 张智雄, 顾立平, 张晓林, 等. 组建中国ArXiv服务工作组, 促进我国有效参与OA活动[J]. 图书情报工作, 2013, 57(1):55-59.
- [5] arXiv Scientific Advisory Board[EB/OL].[2017-07-10].https://arxiv.org/help/scientific_ad_board.
- [6] Member Advisory Board[EB/OL].[2017-07-10].https://confluence.cornell.edu/display/arxivpub/Member+Advisory+Board.

作者简介

王颖, 女, 1982年生, 博士, 馆员, 研究方向: 本体、知识组织。
 张智雄, 男, 1971年生, 博士, 研究员, 博士生导师, 通讯作者, 研究方向: 信息系统和智能信息处理, E-mail: zhangzhx@mail.las.ac.cn。
 钱力, 男, 1980年生, 博士, 副研究馆员, 硕士生导师, 研究方向: 大数据技术和智能信息处理。
 叶志飞, 男, 1984年生, 硕士, 馆员, 研究方向: 智能信息处理。
 王玉莉, 女, 1980年生, 硕士, 馆员, 研究方向: 智能信息处理。
 黄金霞, 女, 1972年生, 博士, 副研究馆员, 硕士生导师, 研究方向: 开放资源建设。
 刘静羽, 女, 1989年生, 硕士, 馆员, 研究方向: 开放资源建设。
 陈雪飞, 女, 1987年生, 硕士, 馆员, 研究方向: 开放资源建设。

The Construction of ChinaXiv Preprint Service Platform

WANG Ying¹, ZHANG ZhiXiong^{1,2}, QIAN Li¹, YE ZhiFei¹, WANG YuJu¹, HUANG JinXia¹, LIU JingYu¹, CHEN XueFei¹
 (1.National Science Library, Chinese Academy Science, Beijing 100190, China; 2.Wuhan Library, Chinese Academy of Sciences, Wuhan 430071, China)

Abstract: To protect the priority right of authors and to promote the excellent scientific research articles first preserved and serve for other researchers in our country, Scientific Article Preprint Platform of Chinese Academy Science (ChinaXiv) is proposed to build. It aims to encourage researchers to share their articles on public preprints platform in the first time, and receive peer reviews and suggestions for a wider range of academic communication. Based on the practice and strategic plan of ChinaXiv construction, this paper describes how to build a preprint platform following the international standard operation regulation, including construction background, system construction and service, domain development and cultivation, organization and management mechanism, propagandizing and popularizing mode, development status and future. This paper can provide references for other preprint repositories.

Keywords: ChinaXiv; Preprint; Open Access; Academic Communication

(收稿日期: 2017-10-10)