



圖書館學 情報學 理論與實踐系列叢書

辛希孟題



JIGOU ZHISHIKU DE JIANSHE YU
FUWU TUIGUANG

机构知识库的建设与服务推广



《图书情报工作》杂志社 编



海洋出版社

名家视点 第6辑

机构知识库的建设与服务推广

《图书情报工作》杂志社 编

海洋出版社

2015年·北京

图书在版编目 (CIP) 数据

机构知识库的建设与服务推广/图书情报工作杂志社编. —北京: 海洋出版社, 2015. 5

ISBN 978 - 7 - 5027 - 9136 - 0

I. ①机… II. ①图… III. ①学术机构 - 知识库 - 文集
IV. ①G311 - 53

中国版本图书馆 CIP 数据核字 (2015) 第 076914 号



责任编辑: 杨海萍

责任印制: 赵麟苏

海洋出版社 出版发行

<http://www.oceanpress.com.cn>

北京市海淀区大慧寺路 8 号 邮编: 100081

北京旺都印务有限公司印刷 新华书店北京发行所经销

2015 年 5 月第 1 版 2015 年 5 月第 1 次印刷

开本: 787 mm × 1092 mm 1/16 印张: 20.25

字数: 348 千字 定价: 48.00 元

发行部: 62132549 邮购部: 68038093 总编室: 62114335

海洋版图书印、装错误可随时退换

《名家视点丛书》编委会

主 任：初景利

委 员：易 飞 杜杏叶 徐 健 王传清
王善军 刘远颖 魏 蕊 胡 芳
袁贺菊 王 瑜 邹中才 贾 茹
刘 超

序

由《图书情报工作》编辑部编选的《名家视点：图书馆学情报学理论与实践丛书》第6辑即将由海洋出版社出版发行与广大读者见面。这是一件值得高兴的事情。从期刊的角度，这是编者从大量的已经发表的文章中精心挑选出来的专题文章，虽然均在本刊发表过，但以专题的形式集中出版，是期刊内容与论文内容的一种增值，体现期刊价值的再利用；对作者而言，这是另外一种传播途径，增强研究成果再次被阅读、被利用的机会，实现论文再次得到关注和充分利用；对读者而言，通过专辑而阅读到多篇同一专题的文章，可以高效率地了解 and 跟踪该领域的研究进展，深化对该领域的认识，对于开展深度的研究或应用到实践工作奠定良好的基础。

本专辑共有4册。第一册是《机构知识库的建设与服务推广》，共收录32篇文章，涉及到机构知识库从基本概念、政策、技术、应用、服务的各个方面，也基本涵盖了机构知识库建设与服务的各个方面的问题，也是有关机构知识库国内重要作者研究成果的大汇聚。机构知识库作为开放获取的重要内容和学术机构自主知识资产的管理与服务系统，是知识管理的重要体现形式，也是图书馆业务与服务新的增长点，具有良好的发展前景和战略意义。对图书馆而言，开发、管理、维护机构知识库并提供基于机构知识库分析的情报分析与科研布局咨询，对图书馆业务与服务的转型发展具有十分重要的意义。

第二册是《移动图书馆服务的现状与未来》共收录37篇文章，涉及移动互联网用户阅读行为、移动图书馆服务模式、移动图书馆服务质量控制、国外移动图书馆服务实践进展、移动图书馆需求与评估等方面。移动图书馆服务在国内图书馆界研究成果不少，学界和业界也高度认同，但由于收到诸多因素的制约，实践上的发展并不够普及和深入。随着移动互联网技术的发展和相关设施的普及，移动图书馆建设仍然是一个值得重视并加大投入的一个领域，其发展前景将十分广阔。

第三册是《馆藏资源聚合研究与实践进展》共收录32篇文章，涉及馆藏聚合模式、数字资源语义关联、关联数据与本体、协同推荐、知识图谱、面向下一代的知识整合检索等。馆藏资源聚合是一个前沿性命题，也是图书馆从资源建设走向基于资源的挖掘与服务的必然过程。这些方面的研究对于深

度地利用馆藏数字资源，实现馆藏资源价值的最大化，具有十分重要的现实意义和应用前景。

第四册是《知识网络研究的进展与创新》共收录 31 篇文章，涉及科研合作的网络分析、共词分析、主题演化分析、学科知识结构探测、研究热点聚类、科研合作网络等，体现了学界业界对这些领域的最新探索和应用性研究成果。为科研提供深度的前沿热点揭示和发现服务，对图书馆服务能力的提升具有重大的意义。图书馆（特别是大学图书馆和专业图书馆）需要加大这一领域的研究、研发和应用的投入，加快图书馆向知识服务的转变。

虽然本专辑的这 4 本书只是从《图书情报工作》近年来发表的文章精选出来的，但也可基本上代表国内学界对相关问题的最新研究成果和图书馆界实践上的探索与创新，具有学术上的引领和实践上的示范作用。尽管研究者还不够多，研究水平也还有待提升，实践应用也处于探索阶段，但也能显示作者们对这些领域的贡献以及潜在的广泛应用价值。

期待这些研究成果能通过这一专辑的出版，对推动国内的学术和实践产生应有的作用，引起更多的图书情报机构的重视，引发更多的研究人员的后续研究，并不断走向深化。在此也感谢所有作者的智慧和贡献，感谢海洋出版社的倾心出版，感谢编辑部同仁所付出的努力。

初景利

《图书情报工作》杂志社社长、主编
中国科学院文献情报中心教授，博士，博士生导师
2015 年 4 月 23 日 于北京中关村

目 次

综 合 篇

机构知识库的政策、功能和支撑机制分析	张晓林(3)
机构知识库发展趋势探析	陈 和(14)
中美图书馆咨询知识库比较研究	王 毅 罗 军(23)
中国科学院系统与高等学校机构知识库建设比较研究	徐红玉 李爱国(32)
机构知识库联盟发展现状及关键问题分析	曾 苏 马建霞 祝忠明(44)
日本机构知识库发展与现状研究	陈枝清 徐 婷(53)
日本机构知识库存缴后印本论文的著作权策略研究	朱莲花 牟建波(60)
机构知识库建设中存缴和发布已发表作品的法理透析 ——梳理中国科学院研究所机构知识库主要著作权疑虑	周玲玲(68)
从心理学角度谈如何促进机构知识库资源建设	蔡 屏(75)
以人为本 科学构建学者知识库	何继红(84)
期刊出版商版权协议对我国机构知识库发展的影响	于佳亮 马建霞 吴新年(92)
国内机构知识库研究文献的可视化分析	奉国和 吴敬学(101)

建 设 篇

机构知识库建设模式研究	邓 君(115)
机构库共享机制研究	谢 琴 王 军 赵伯兴(125)
公众议题知识库的多层本体设计	张鹏翼 周 妍 袁兴福(135)

分类主题词表的计算机自动编制

- 兼论用于自动分类的知识库的改进 顾 颖 何 琳(151)
- OAI 资源库更新频率的定量分析和同步模型的建立 ... 翟中会 石 蕾(161)
- 汉语框架网络知识库的语义角色特征识别 ... 贾君枝 赵文娟 王东元(170)
- 创作共用协议在机构知识库建设中的应用与意义 刘玉婷 马建霞(179)

应 用 篇

中国科学院机构知识库建设推广与服务

- 张冬荣 祝忠明 李 麟 王 丽(189)

国家科学图书馆咨询知识库的研究与实践

- 李 玲 姚大鹏 魏 韧 张杰龙 范 炜(201)

西安交通大学机构知识库数据共享集成研究与实践

- 魏青山 张雪蕾 陈楠楠 邵晶(211)

高校机构知识库与用户的互动关联策略研究

- 傅 俏 卢章平 盈江燕 袁 润(223)

以用户需求为导向的高校机构知识库自存储服务机制研究 胡海燕(234)

基于文献数据规律的机构知识库数据转换模型研究

- 侯瑞芳 李玲 陈嘉勇 肖 明(244)

基于本体知识库的模糊信息检索研究 俞扬信(253)

科研人员对 OA 知识库的认知程度和使用现状分析 ... 李 武 卢振波(262)

评 价 篇

国内外机构库评价研究现状述评 袁顺波 华薇娜 马学良(275)

机构库评价的关键问题研究 马学良(282)

基于成熟度视角的机构库评价研究 袁顺波(289)

Altmetrics 指标在机构知识库中的应用研究 ... 邱均平 张心源 董 克(296)

开放存取知识库的网络计量排名和评价研究 崔宇红(308)

综 合 篇

机构知识库的政策、功能和支撑机制分析

张晓林

(中国科学院国家科学图书馆 北京 100080)

机构知识库 (Institutional Repositories, IR) 作为机构知识管理的重要部分, 正日益成为研究型图书馆的重要职责之一^[1]。但是, IR 不仅是技术平台建设, 更多地涉及复杂的目标、政策、功能和支撑机制。本文结合国家科学图书馆的实际需要, 提出机构知识库建设的相关政策与机制的基本分析。

1 机构知识库目标分析

机构知识库是机构知识管理的重要机制^[2-3]。根据定位的不同, 机构知识库的建设目标可以分为三个层次^[4-5]:

- 机构知识资产管理。主要任务和目标是对本机构已经形成的显性的知识产品本身的收集、组织和长期保存。
- 机构知识传播管理。主要目标扩展到知识内容在内外传播, 将知识资产转化为生产力, 关注对显性或隐性知识的及时捕获、转化、传播以及对知识内容传播状况的审计, 支持机构最大限度利用自己所产生的知识内容。
- 机构知识能力管理。主要目标进一步扩展到对机构的知识需求分析 (“需要什么知识”, 例如可存储关于需求的知识, 结合领域或任务知识地图的对照分析)、知识能力分析 (“拥有什么知识”, 可按主题和时间构造机构知识地图)、知识关系分析 (“机构内外知识内容间存在什么关系、如何利用这些关系产生新知识或新能力的关系”, 例如知识合作关系分析)、知识资产应用分析 (“拥有的知识是否得到利用”, 可按浏览检索、或应用程度分析)、知识需求与知识能力匹配分析 (“需要与拥有的差距”, 可进行领域知识地图与机构知识地图的对照分析)、知识能力竞争力分析 (与竞争对手知识地图的对照分析) 等, 支持对真实知识状态及知识缺陷的鉴别, 促进机构提升自身的知识能力。

机构知识库建设应从发展机构知识能力出发, 保证实现对本机构知识资产

的收集和长期保存，积极建设对知识内容进行捕获、转化、传播、利用和审计的能力，逐步建设包括知识内容分析、关系分析和能力审计在内的能力。

2 知识内容类型

机构知识库覆盖的知识内容，一般指本机构成员学生和合作者产生的知识内容。

2.1 作者范围

- 机构成员作品。即本机构员工、学生及直接合作者所创作的知识内容，其中直接合作者指在本机构资助下或以机构名义参与有关工作的其他人员，例如项目聘任人员、委托研究人员、参加本机构资助项目的合作研究人员等。

- 其他合作者作品。例如本机构牵头合作项目的参加者、本机构主办承办会议的参加者等生产的知识内容。

- 合作作品。本机构作者与其他作者合作产生的知识内容，包括在合作课题中、或在合作者机构支持课题中产生的知识内容。

- 本机构出版物。例如学术期刊、专著或其他出版物，其中包含机构成员和非机构成员的作品。

2.2 内容形式

- “文献型”作品，例如文件、论文、报告、图书等文献。

- “非文献型”知识产品，例如 PPT、图片、课件、音视频资料、软件等多媒体或复合文献。

- “数据型”知识内容，包括数据库、数据集等。

- “隐性”知识内容，即那些隐藏在个人或部门的工作流程、方法、习惯等中的知识，尚未转化为正式资料，但可记录和转化成为显性、公共的知识。

- 个人知识内容，属于机构成员的但对本机构学术、专业、管理等有关的内容，可记录、收集和转化成为显性、公共的知识。

2.3 内容性质

具有普遍科学意义的知识内容（论文、科研报告、专著等）；有竞争意义的知识内容（例如涉及竞争性产品、方法、技术等）；内部管理性质的知识内容（例如内部工作计划、预算、评价、规范、程序等）；带有个人私密信息的知识内容（例如应聘招聘、学位学习和答辩成绩、课题申请人信息等）。

2.4 发布形式

正式发布或发表的作品，例如已发布的各种正式文件、在期刊和论文集上

发表的论文等。“相对公开使用”但尚未发表的作品，例如已在学术会议、课题网站、个人博客、社区论坛、公开讲座等上发布，或已在一定范围“交付使用”（例如呈交、批准、通过等）且无专门传播限制的报告、论文、演示文档、访谈等；未公开发布或发表的内容，指已经完成但尚未发表或发布的作品，也可指相应的初稿、讨论稿、征求意见稿等，或个人认为有意义的已经形成文字或其他媒体表现形式的知识内容。

2.5 权属性质

根据著作权法及其实施条例等^[6]，知识内容区分为：

2.5.1 法人作品 由机构（法人或非法人组织，下同）主持，代表本机构意志创作，并由本机构承担责任的作品，本机构视为作者，例如机构的文件、工作报告、年度报告、工作规范等。法人作品的著作权属于法人。

2.5.2 职务作品 机构成员根据机构下达的指示创作与本机构工作业务有关的作品是职务作品，尤其是利用机构专门提供的物质技术条件来创作，并由机构承担责任的工程设计图、产品设计图、地图、计算机软件、数据集等作品，由机构组织人员进行创作，提供创作条件并承担责任的百科全书、辞书、教材、大型摄影画册等编辑作品，或者法律、行政法规规定或合同约定著作权由机构享有的职务作品。对于职务作品，作者只享有署名权，著作权的其他权利由本机构享有。

2.5.3 其他职务作品 机构成员在机构物质技术条件支持下、为完成机构工作任务所创作的其他作品（例如学位论文和课题研究报告），除合同另有规定者外，著作权由作者享有，但机构有权在其业务范围内优先使用。

2.5.4 课题作品 由机构成员在机构或公共经费支持下承担课题所完成的学术论文、科研报告、专著等作品，除合同另有约定外，著作权属于作者。但作品中涉及到的工程设计、产品设计图纸及其说明、计算机软件、地图等作品，如果是利用本机构的物质条件创作、并由本机构承担责任的，其著作权归属本机构。

2.5.5 机构出版的汇集作品 机构出版的学术刊物、会议录、汇编文集和其他汇编性质的作品，属于机构主持、由机构承担管理责任的作品。除合同另有规定者，其中完全属于机构业务工作、在机构提供物质技术条件下产生的内容，应视同作者仅享有署名权的职务作品；由机构成员创作、属于作者完成机构工作任务或在机构课题支持下完成的作品，属于作者享有著作权的职务作品或课题作品。

2.5.6 其他作品 机构成员利用其他条件所完成的学术论文、科研报

告、专著等作品，除法律、合同另有约定外，著作权属于作者。

3 服务功能设计

机构知识库应在快速实现强大的机构知识资产管理的同时，积极试验、不断充实知识传播管理和知识能力管理的相应功能。

3.1 收集

应建立相关制度、方法和技术支持高效方便的收集。

- 在制度上，要确定什么人的什么作品在什么时间以什么方式进行存缴^[7-8]。应规定（Require，下同）立即存缴法人作品；应规定机构成员（含直接合作者，下同）立即存缴正式发表的职务作品和课题作品（包括合作作品），并建立相应的保障激励制度（例如作为评价依据）以及必要的传播保护政策；应规定机构成员立即存缴“相对公开使用”的职务作品和课题作品（包括合作作品），并建立相应的激励与约束制度（作为评价依据、优先权属证据和知识能力证据）以及必要的传播保护政策；应规定立即存缴以机构及所属部门名义出版的汇集作品，但其中由非机构成员完成的非合作作品，应由作者选择传播方式。应要求（Request，下同）机构成员及时存缴其他作品，并建立相应的鼓励制度（作为优先权属证据和知识能力证据）以及相应的传播保护政策；应鼓励（Recommend，下同）其他合作者及时存缴已发表的和“相对公开使用”的作品，并建立相应的激励与约束制度以及必要的传播保护政策；应逐步建立对隐性知识进行记录、收集和转化的机制，同时建立相应的鉴别、权属管理、传播管理和鼓励政策；应通过多种服务鼓励存缴（例如提供浏览统计证明）；可建立集中授权机制，例如机构成员授权机构对自己正式发表或“相对公开使用”的作品进行收集和存缴。

- 在方法上，应建立可非常方便使用的存缴接口，支持连续上载、批量上载，应建立从其他内容系统（学位论文系统、期刊编辑出版系统、ARP 系统、学术会议平台、研究生教育管理系统等）获得作品或元数据的接口；应逐步建立自动抽取元数据的方法。应逐步建立对非数字作品进行数字化、转换、封装和载入的方法，应逐步建立对隐性知识进行记录、收集、转化、鉴别及权属管理的方法。

- 在技术上，应支持机构成员以常见格式上载作品；应提供将常见格式转换为保存格式的机制。应建立将用户格式转换为可保护的传播格式的方便机制，应提供将用户格式和保存格式、传播格式链接的内部机制。

3.2 组织

应建立相关技术手段来支持灵活多样的内容组织。

- 从内容组织角度，应支持按照个人、部门（年级）、课题（课程）、领域、主题、类型、出版或发布渠道等多个维度关联组织所收集的作品；应支持同一内容的多类相关作品（例如某一学位论文与其他论文、数据集、软件等）的关联组织；应支持同一人员、部门或课题内容多个作品的逻辑组织；应支持多个个人、部门和课题间通过合作作者、合作单位、资助单位、资助课题或项目等的关联组织。可建立具有灵活关联标引功能的元数据格式，并在标引和检索时提供相应功能。

- 应逐步支持利用复合对象格式来关联或封装相互关联的多个作品。先期可将一般对象格式作为长期复合对象格式的一个可扩展的特例格式。

3.3 传播

促进知识内容传播是机构知识库的重要任务，需要建立积极的传播政策、必要的传播保护机制和灵活多样的传播技术。

传播政策应综合考虑促进机构知识能力发展、保障公众获取和保护作者权益的需要，规定什么人的什么作品在存缴后什么时间以什么形式传播。

对于具有普遍科学意义的知识内容，应尽量采用开放获取方式广泛传播。对于具有竞争意义的知识内容，可以有合理的传播保护措施，但应避免对没有实质竞争意义的知识内容过分保护，避免对已经失去竞争意义的知识内容继续保护，而且应树立和宣传“传播就是竞争”（获得优先权属证明）的意识；对于具有内部管理性质的知识内容，如果包含具有普遍意义的学术或技术知识和可以提供广泛借鉴的战略、政策、制度、方法，只要不影响竞争和不涉及个人隐私，应尽量广泛传播，而且公共机构有义务公开自己的战略、政策、制度和方法，保证透明性和可公共问责；对于涉及个人隐私的内容，则应采取有效的保护措施。

从传播科学、服务公众、宣传机构、提高可公共问责性角度，应规定机构成员已公开发表的职务作品和课题作品在存缴时立即实行开放获取；应规定机构成员“相对公开使用”的职务作品和课题作品在存缴后立即实行开放获取或在有限保护期后实行开放获取（保护期内也可向指定范围开放传播）；应规定机构出版的汇集作品在存缴后立即实行开放获取或在有限保护期后实行开放获取（其中由非机构成员完成的非合作作品，应事先征得作者同意）；应鼓励机构成员的其他作品尽可能实行开放获取或在有限保护期后实行开放获取；应鼓励其他合作者已公开发表或“相对公开使用”作品尽量尽早实行开放获

取。无论作品是否有传播保护措施，除非有法定保密要求，应全面开放所有保存作品的元数据。

应提供多种形式的检索浏览能力，包括对个人、部门、课题、类型等不同维度的集成浏览能力，提供利用各种关系关联检索的能力。

应提供机构知识库元数据开放收割接口^[9]和开放检索接口，支持第三方对机构知识内容的揭示和对可开放获取内容的检索，建立帮助获得非开放获取内容的获取机制及咨询服务。

3.4 长期保存

机构知识库是知识资产长期保存的基础，但它还必须在存缴、组织和保存转移等方面的政策和技术上为长期保存提供具体支持。

- 从存缴角度，应规定作者存缴原始内容如果有传播保护措施一例如只发布 PDF 版本一则可由知识库系统自动转换。如果原始对象为非常用格式，应由知识库系统自动转换为常用格式。应要求作者存缴完整对象。应提供将不同时间提供的但实质上属于同一完整对象的内容关联起来的机制。知识库元数据应包含存缴对象的技术信息、权属信息和其他与长期保存密切相关的信息。应建立已存缴作品的撤出管理机制，对规定存缴的内容，未经机构知识库管理者同意不能撤出；对要求和鼓励存缴的内容，应通过必要的传播保护和激励措施鼓励人们长期保存，并建立撤消存缴登记制度，使得机构知识库掌握和更新保存情况。

- 从对象组织的角度，应积极采用开放数字对象格式来组织存缴内容，支持存缴对象的结构化组织、迁移、输出等；应逐步支持复合对象格式，关联组织或封装涉及相同人员、课题、部门等的多个作品。

- 从对象保管的角度，应建立所存缴内容的数据管理机制，通过自动、定期的存储库审计功能，确认所存缴内容是否得到妥善保存；审计功能应能按照个人、部门、课题等分别进行，以方便理解的方式呈现和报送审计结果；应建立所保存内容的定期备份机制；应建立所保存内容的转移输出功能，按照开放封装格式进行转移，应能按照个人、部门、课题等分别进行转移。

3.5 扩展服务

应建立相关机制来支持基于保存内容的丰富服务。

- 从知识能力宣传的角度，可利用知识库所保存内容建立虚拟的个人、课题、部门等的知识目录（Portfolio），集中展现相应对象所产生的知识成果。应提供知识目录的定制功能，提供知识目录相互关联浏览的能力。

- 从知识能力审计的角度，应支持对存缴内容按照时间、部门、课题和

个人等统计的功能；应支持对所保存内容被不同范围用户从不同方式使用（例如浏览题录或摘要、下载全文）的统计功能。应支持利用所保存内容建立知识地图的功能。统计结果应能正式输出以支持宣传和评价的需要。

- 从知识能力评价的角度，应逐步支持对机构的知识需求分析、知识能力分析、知识关系分析、知识资产应用分析、知识需求与知识能力匹配分析、知识能力竞争力分析等。这些可以利用外部工具，分析与评价结果应保存到机构知识库。

3.6 互操作服务

应建立相关机制来支持机构知识库与外部知识系统的有机链接，促进机构知识资产的充分传播，促进外部相关知识的发现和利用。

- 从关联利用外部相关知识的角度，应建立从机构知识库内特定知识内容直接检索外部知识系统乃至集成机构知识库内外资源的功能。应嵌入帮助机构成员获取外部知识内容的服务，应嵌入为检索外部知识内容提供咨询帮助的服务。

- 从提升知识管理能力的角度，应建立机构知识库与科研管理系统、办公管理系统、机构信息发布平台、学位论文系统、期刊编辑出版管理系统、学术会议平台、教育管理系统等双向内容交换和验证的机制，保障知识内容的及时捕获、可靠记载和与管理流程的紧密结合。

4 权益保护政策

机构知识库所保存内容涉及机构、机构成员和合作者的作品及权益，公共机构还必须考虑国家和公众的权益。因此，需要建立相应的权益保护机制，保证知识内容被可靠收集，知识产权得到妥善保护，知识管理能规范进行，消除对竞争性使用和不合理使用的顾虑，规避管理和法律风险，保证可持续运行。

4.1 传播保护机制

由于各个方面对不同性质的作品拥有不同的权利，应针对性地制定相应的传播保护政策，妥善处理个人与机构、保护与传播的关系。

在需要进行传播保护时，应首先考虑非限制性的传播保护措施，例如采用创作共享版权许可（Creative Commons Licenses, CC）^[10]，支持开放传播。或将原始文档转换为不可修改及不可摘取片段的 PDF 文件开放传播。

在竞争性需要或出版者限制下需要采取限制性传播保护措施时，应建立“例外限制”政策，即默认是开放传播，只在必要时才能过专门选择设置传播限制，避免自动进行笼统的限制；应确定“有限限制”和“最小限制”原则，