

丝路科技知识服务系统 “一带一路”专题数据库的构建与实践*

李娟 张雪蕾 陈楠楠 杨峰
(西安交通大学图书馆, 西安 710049)

摘要: 本文梳理了国内“一带一路”专题数据库现状, 介绍了西安交通大学丝路科技知识服务系统建设目标、框架, 重点介绍系统“一带一路”专题数据库资源建设实践, 包括多元化服务对象的资源建设规划制订、高质量资源筛选、多途径资源的采集加工与揭示、使用效果评估等, 并就“一带一路”专题数据库资源建设提出思考。

关键词: 一带一路; 专题数据库; 资源建设

中图分类号: G253

DOI: 10.3772/j.issn.1673-2286.2019.02.006

1 “一带一路”专题数据库现状

“一带一路”倡议(以下简称“倡议”)自提出以来, 取得了一系列显著成果。用户对与之相关的特色文献信息资源的需求也呈激增状态。高校图书馆作为文献资源中心, 亟待构建起较为完善的“一带一路”知识服务体系, 以更好地服务教学科研, 满足不同服务主体(如研究学者、留学生、企业从业人员等)的信息需求。与此同时, “一带一路”信息资源内容呈多元化趋势, 涵盖新闻、学术动态、学术期刊、会议、专利等, 涉及多个学科领域, 因此图书馆常规采购的数据库与学科文献已不能满足多元化服务对象的实际需求, 需要通过构建新的知识服务平台将相关信息资源进行整合、补充与完善。

目前, 政府、高校、科研院所和企业等纷纷利用自身特色资源优势构建“一带一路”专题数据库, 为“一带一路”发展提供知识服务。优质的信息资源对于推动“一带一路”倡议具有重大作用, 同时也对资源选取、采集与深度挖掘提出了挑战。国内众多的“一带一路”专题数据库依据各自功能定位和资源优势, 配置了不同内容的核心信息资源。表1梳理了部分较为知名的“一带一路”专题数据库, 并对其功能定位与核心资源进行介绍。

可以看出, 专题数据库的服务对象可满足多领域用户需求, 如中国一带一路网、“一带一路”研究与决策支撑平台, 涉及领域广泛、资源类型多样, 能够为各级政府部门战略决策、政策制定提供支持。同时, 专题数据库特色内容侧重分明, 如列国志数据库重点关注国家基础信息, 新华丝路侧重于经济贸易, 但尚未有关于“一带一路”沿线国家和地区深入、全面和系统的专题数据库, 特别是对沿线国家和地区的历史文化、工业经济、人口环境、国情咨文、政策法规等的研究和认识都较为模糊。要实现倡议提出的“五通”愿景, 就必须发挥高校人力资源、学术资源等优势, 立足自身特色和学科优势, 兼顾“一带一路”沿线国家和地区发展及国内研究需要, 构建特色数字资源, 全方位地提供沿线国家和地区有关信息资源, 剖析沿线国家潜在机遇与挑战, 为我国国家战略实施提供人才支撑、交流服务和智力支持。

西安交通大学作为“新丝绸之路大学联盟”的发起者, 推动了“新丝绸之路经济带”沿线国家和地区大学在校际交流、人才培养、科研合作、文化沟通、政策研究、医疗服务等方面的交流与合作。2015年, 中国工程院、西安交通大学、西安高新区三方共建丝路学院, 旨在为国家“一带一路”倡议的实施培养国际化、专业化、

*本研究得到中国工程院“中国工程科技知识中心建设项目”子项目“丝路科技知识服务系统”(编号: CKCEST-2018-2-6)和陕西省图书馆学会课题“高校图书馆‘一带一路’数字资源建设策略研究——以西安交通大学图书馆为例”(编号: 181015)资助。

表1 部分“一带一路”专题数据库内容梳理

“一带一路”专题数据库	依托机构	功能定位	核心资源
列国志数据库 ^[1]	社科文献出版社	国别、国际问题研究资讯平台	155个国家和地区的政治、经济、历史等方面信息
一带一路数据库 ^[2]	社科文献出版社	基础资料库、学术理论库、投资指南库	社科文献出版社出版的“一带一路”相关研究文献
“一带一路”研究与决策支撑平台 ^[3]	国研网	提供决策支持、学术研究参考，以及引领中国企业“走出去”	政府机构发布的最新研究成果、案例、招标投标投资等
新华丝路 ^[4]	新华通讯社	商品贸易和投资服务平台	投资和融资环境与风险分析，研究成果、最新动态、合作信息查询等
中国一带一路网 ^[5]	国家信息中心	国内外有关的投资贸易、税收检验等政策法规	政策发布解读、基础数据、投资指南、案例查询
EBSCO一带一路资源中心数据库 ^[6]	EBSCO Industries, Inc	基于EBSCO数据库文献信息资源整合	沿线国家的贸易、商业、经营、投资、文学等学科文献

实用性人才^[7]。通过一系列行动和政策保障，西安交通大学承接了“中国工程科技知识中心建设项目”子项目“丝路科技知识服务系统”，该系统基于“一带一路”沿线国家和地区知识共享及人才培养的需求建设，知识服务将作为重要功能模块为国内从业人员、“一带一路”沿线国家和地区技术人员及留学生提供资源信息。

2 丝路科技知识服务系统“一带一路”专题数据库建设实践

2.1 丝路科技知识服务系统概述

丝路科技知识服务系统（以下简称“系统”）是在联合国教科文组织框架协议下，中国工程院委托西安交通大学建设的国际工程科技知识中心。系统通过各类科技资源汇集，建成知识共享平台，有效地凝聚丝路工程科技力量，衔接“新丝绸之路大学联盟”盟校教育、教学实践，沟通国家政策与院校改革等，为丝路沿线国家相关研究、人才培养、创新创业提供知识服务。其中，“一带一路”专题数据库作为系统建设重点，支撑整个系统知识图谱、个性化推荐等特色资源服务。

借助系统建设契机，西安交通大学图书馆积极参与系统专题数据库建设工作，通过数据挖掘和信息整合构建专题数据库。系统专题数据库与“一带一路”专题数据库相较，其共同点在于提供的资源能够满足“一带一路”沿线国家和地区知识共享及人才培养需求；不同点在于系统所建“一带一路”专题数据库是依托西安交通大学图书馆建设，图书馆在资源采购甄选和信息整合方面更加专业和系统，提供的特色数据资源覆

盖范围更为广泛，其服务对象更多，因此资源更新速度及时；资源以英文为主，也提供俄语、阿拉伯语等小语种；内容除“一带一路”相关资源外，也包括大量高质量教育资源等。系统专题数据包括国情咨文数据库、政策法规数据库、历史文化数据库、人口环境数据库、工业经济数据库、教育科技数据库。数字资源建于2016年3月—2018年9月，建成的数据库共包含各类特色资源230余万条，资源类型涵盖期刊、会议、图书、专利、国家报告、学位论文、视频等。

2.2 “一带一路”专题数据库建设实践

2.2.1 面向不同用户开展需求分析，制定数字资源建设规划

由于服务对象对资源需求各异，有必要规划数字资源建设，调研数字资源类型、分布和流动状况，最终确定系统专题数据库资源建设内容和类型。本文将专题数据库的服务对象分为高校、科研院所或政府机关的研究人员，“一带一路”沿线国家和地区留学生，以及企业单位从业人员，通过线上、线下调研了解各类服务对象的信息资源需求，将主要信息资源需求整理见表2。

通过需求分析与主题数据库数字资源内容对接，同时对数字资源建设应达到何种水平作出规定，如数量目标、质量目标和特色化目标等内容。数量目标，即一定时期内数字资源增量或学科覆盖率；质量目标，即对采集的数字资源广度、深度和新度提出明确要求；特色化目标，即构建何种特色数字资源，根据目标制定出科学、切实可行的数字资源建设规划。

表2 服务对象资源需求分析

服务对象	资源需求	对接的六大主题数据库
高校、科研院所或政府机关的研究人员	国家政策法规和战略理论	政策法规数据库
	沿线国家政治动态	国情咨文数据库
	沿线国家历史和文化	历史文化数据库
	工业经济发展、贸易投资环境等	工业经济数据库
	沿线国家发展指数、环境污染与治理等	人口环境数据库
“一带一路”沿线国家和地区留学生	免费获取最新科学技术成果	教育科技数据库
企业单位从业人员	工业经济发展、贸易投资环境等	工业经济数据库
	国家关于投资方面的政策法规	政策法规数据库

2.2.2 筛选数字资源内容, 涵盖多样化资源类型

图书馆在筛选数字资源时着重强调资源的广度、深度、新度和特色性, 通过提供不同知识应用满足个性

化需求, 提供多语言支持和扩展。服务对象的资源需求各异, 因此承载信息资源类型也应非常丰富, 须从不同渠道获取, 从不同来源实现资源提取与整合。表3为六大特色数据库的内容、资源类型及采集建设方式。

表3 专题数据库下设六大特色数据库情况

资源名称	简介	资源类型	采集建设方式
国情咨文	《国家报告》针对某个国家或地区的主要事件进行全面深入的分析	国家报告、期刊等	信息抽取、自建数据库、网上免费资源获取
政策法规	多项技术实施标准, 以及沿线国家政策法规、国际关系等	标准、期刊等	同上
工业经济	沿线国家的经济现状、政策与国际贸易、企业基础信息等	SWOT、期刊等	同上
人口环境	沿线国家人口概况, 环境污染、治理等	期刊、会议等	同上
教育科技	“新丝绸之路大学联盟”盟校发表的科技文献、开放获取资源	期刊、会议等	同上
历史文化	沿线国家区域政治、经济、文化、发展策略等	电子图书	购买

2.2.3 多渠道开展资源获取, 保障高质量数据资源

除遴选已购买电子资源外, 图书馆应充分发掘和利用现有专题数据库, 重视开放性网站资源信息整理, 完善专题数据库建设。目前, 系统专题数据库的数据获取方式主要包括: 一是对已购数据库资源进行深度挖掘; 二是通过自建数据库采集; 三是通过外部数字资源采集, 主要包括开放获取资源和互联网资源; 四是适当购置“一带一路”特色数字资源。数据通过API、XML推送、OAI协议、Z39.50协议等方式, 从不同数据源采集, 具体采集方式见图1。同时, 从专业数据库中获取数据字典, 将数字资源按照数据字典进行数据转换, 对转换后的数据进行加工、标引和整合, 最后在丝路科技服务平台进行揭示。

(1) 已购数据库采集。西安交通大学图书馆已购置了国内外电子资源110个平台约311个子库, 中英文全文

电子期刊近60万种, 电子图书150万余种^[8]。包括Elsevier出版社全文电子期刊、EBSCO学术与商业信息数据库、剑桥大学出版社电子期刊及“一带一路”专题图书等资源。已购数据库均有独有资源, 并形成了较为完善的体系, 各自具有优势。因此, 在已购高质量优质数字资源基础上, 利用数据挖掘技术或与数据库商合作, 对资源进行充分利用与挖掘, 在建设特色资源的同时, 通过多途径的揭示与发现, 提高馆藏数据资源利用率。

(2) 图书馆自建特色数据库采集。国内众多高校图书馆根据学校建设情况、学科特色、区域特点及自身资源建设等优势, 建立了内容丰富、广泛的自建特色库, 如硕博学位论文数据库、学科导航数据库、机构知识库、学者库等。这些自建库学科针对性强, 学术利用价值较高, 数据著录格式规范, 元数据质量较高。如西安交通大学机构知识库^[9]自2009年建设以来, 累计保存、管理本机构期刊、会议论文、专著、学位论文等各类

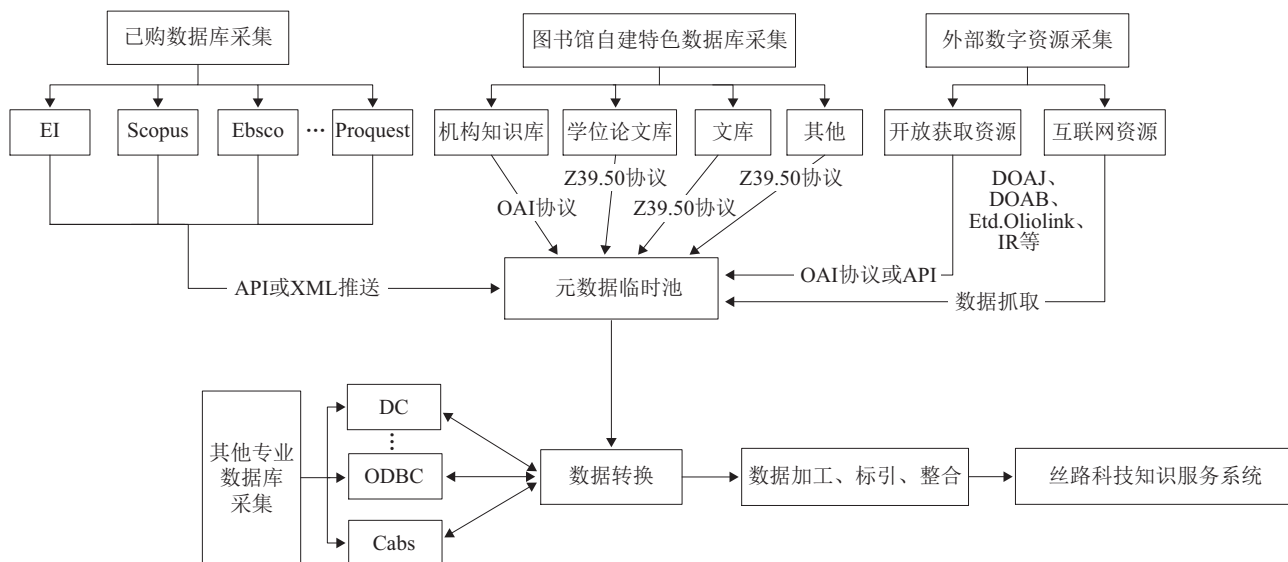


图1 数字资源的采集来源及方式

型成果产出32万余条，可遴选出相关的学术文献作为专题数据库数字资源建设的重要补充。此外，学位论文库、文库及本校其他特色自建数据库也均为专题数据库提供了特色资源。

(3) 外部数字资源采集。对互联网资源，尤其对“一带一路”专题网站进行充分发掘，如表1中提及的专题数据库和北京大学“一带一路”数据分析平台等。在系统专题数据库建设过程中，搜集整合此类专题平台的可用信息，通过专题网站提供的英文摘要、来源出处等线索寻求原始出处，将这类资源纳入专题数据库。

除互联网资源外，开放获取资源也是重要组成部分。开放获取运动促进了科学信息的分享、传播利用和知识创新，有利于沿线国家免费获取全文。目前，越来越多的机构愿意开放共享，比较有影响力的如DOAB^[10]、DOAJ^[11]等，允许通过OAI协议获取，用户可下载元数据；如英国剑桥大学机构典藏库提供部分全文下载；Etd.Oliolink收录联盟大学约9.3万篇博硕士论文的免费数据，还提供专利、报告等资源^[12]。上述开放获取资源学术价值与元数据质量较高，并提供免费使用的接口或协议。教育科技数据库与开放获取资源相结合，从众多开放获取资源中优选学术价值较高、适合“一带一路”沿线国家和地区各类用户的教育科技类资源。

所有获取的数字资源均由图书馆负责内容筛选和质量评价，优先遴选权威出版社电子期刊和电子书，如历史文化数据库主要收集的资源是以iGroup Publishing和

EBSCO出版商关于“一带一路”沿线国家和地区相关的区域政治、文化和发展策略为主的电子书。此外，图书馆在专题数据库中，还对元数据进行标引与加工，通过提供高质量的元数据，方便用户通过浏览元数据信息就可以基本判断信息资源是否符合需求。

2.2.4 数字资源加工与整合揭示

专题数据库按照不同资源类型制定自定义数据字典，按照统一的标准规范加工创建标准元数据，保证入库后元数据质量，为后期数据复用、揭示与传播奠定基础。对于异构数据库资源，依据统一标准，采用物理或逻辑方式格式转换与处理，形成元数据；对于外部互联网资源，根据相应组织体系进行清洗、加工、标引等。处理后的数据资源统一存放于数据仓储，最终在丝路科技服务平台专题数据库中展示，提供导航、检索与个性化信息服务。资源整合逻辑架构见图2。

2.2.5 使用效果评估

资源使用效果评估是资源建设的重要环节，资源是否被合理和充分使用、投入产出比等是资源建设质量评估的主要内容。专题数据库对已购数据库中资源使用效果评估，主要通过SHSUI协议收割数据库商基于COUNTER标准的数据使用量统计报告；对自建特色数据库、开放获取资源及网络资源，借助西安交通大

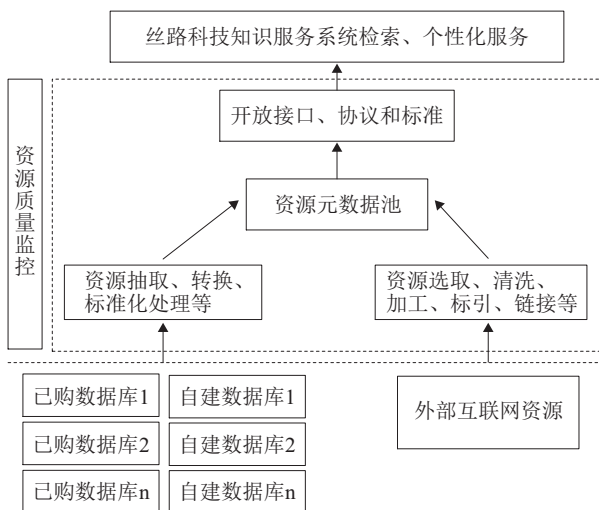


图2 数字资源整合逻辑架构

学电子资源利用与统计系统对资源进行评估,通过访问量、下载量、被拒访问量等评估资源使用效果,通过价格模型评估单篇资源成本,为下一期采购经费使用提供依据。合理使用方面,利用电子资源流量访问控制系统^[13]设置下载量阈值,对用户使用电子资源行为进行预警提醒,避免过量下载行为,对恶意超量下载行为利用黑名单机制限制违规用户继续使用。

3 “一带一路”专题数据库建设的思考

3.1 知识产权的问题

数字资源的元数据理应免费开放,但仍应要求用户引用时必须提供元数据URL;至于全文,对于不同数据源作品应分级开放,且提供多种全文获取路径。此外,专著或报告等的著作权也是不容忽视的问题;非PDF格式的数字资源(如流媒体、XLS等),合理使用量的界定并不明确;部分国外数据库,出于对资源保护或其他目的采用动态链接技术,因此无法为用户提供一个长期有效的访问地址,只能提供数据库全库地址供用户使用,从便捷性角度来讲用户体验感较差。

3.2 数据来源的选择

在资源建设过程中,可选择购买学术价值较高的数据源库,如EBSCO、Elsevier、Scopus等优质数据库,将其作为教育科技资源建设的重点数据源;与沿

线国家相关的历史文化资源,可选择Burma Research Society(缅甸研究学会)、Nordic Institute of Asian Studies(北欧亚研究所)等知名机构的出版物作为数据源;对于政策法规、人口环境、工业经济等内容,可从Gale、ABI、IEEE等数据库资源遴选,并将高质量的开放获取资源如DOAJ、DOAB、机构知识库等作为各类学术资源的补充。沿线国家的风土人情、人文历史等领域的音频视频、拓片、博物馆相关文物展示、手稿等资源较少,在未来可扩大此类数据源采集范围。同时,在数据源选择时应尽量避免资源内容重复,合理利用采购经费,降低检索及下载成本,通过评估试用数据库来制定采购决策。

3.3 重点关注引进小语种资源

目前“一带一路”专题数据库以中文为主,资源建设极少覆盖沿线国家主要语种。纵观我国学术界和高等教育界对现有信息的涉猎和积淀,对沿线国家和地区的研究十分缺失和薄弱。长期以来,教育和研究围绕在欧美等发达国家和地区^[14],且几乎将绝大多数精力聚焦美国,因此对沿线各国、各民族的全方位了解极为不足。以往的忽略和缺失,导致倡议推进过程中智力支持乏力,也使许多快速推进的经济和外交行为受到诸多未知因素的制约。在“一带一路”专题数据库建设时,需要特别重视小语种原版资料的引进,从沿线国家的机构知识库或学科库中获取所需要的开放资源,进而整合到相关数据库中。

3.4 资源共建共享

同“新丝绸之路大学联盟”盟校开展合作,将盟校优势资源整合揭示,如优势专业信息、学科资源、线上学习资源等,通过共建共享实现信息资源互通。此外,还可以与同区域“一带一路”研究机构(如西北大学丝绸之路研究所)建立合作,深入了解用户资源建设需求,补充完善及确定资源建设内容,使所建资源真正为用户所需所用,从而提高资源利用率来开展细致的知识服务。

参考文献

[1] 列国志数据库[EB/OL]. [2018-07-10]. <https://www.lieguozhi>.

- com/.
- [2] 一带一路数据库 [EB/OL]. [2018-07-12]. https://www.ydylcn.com/skwx_ydyl/sublibrary?SiteID=1&ID=8721.
- [3] “一带一路”研究与决策支撑平台 [EB/OL]. [2018-07-12]. <http://ydyl.drcnet.com.cn/www/ydyl/>.
- [4] 新华丝路 [EB/OL]. [2018-06-10]. <http://silkroad.news.cn/>.
- [5] 中国一带一路网 [EB/OL]. [2018-06-10]. <https://www.yidaiyilu.gov.cn/>.
- [6] 图书馆开通EBSCO“一带一路资源中心”数据库试用 [EB/OL]. [2018-06-10]. http://www.lib.pku.edu.cn/portal/cn/news/0000001366?qt-search_quicktab=2.
- [7] 西安交大与西安市联合培养“一带一路”国际人才 [EB/OL]. [2018-07-10]. http://www.sohu.com/a/194094419_99983297.
- [8] 图书馆简介 [EB/OL]. [2018-07-10]. <http://www.lib.xjtu.edu.cn/gk/tsgjs/tsgjj.htm>.
- [9] 西安交通大学机构知识门户 [EB/OL]. [2018-08-10]. <http://ir.xjtu.edu.cn>.
- [10] DOAB [EB/OL]. [2018-07-12]. <https://www.doabooks.org/doab>.
- [11] DOAJ [EB/OL]. [2018-06-12]. <https://doaj.org/>.
- [12] OhioLink [EB/OL]. [2018-6-12]. <https://etd.ohiolink.edu>.
- [13] 邵晶, 阎晓弟, 周琴, 等. 电子资源流量控制需求分析及其解决方案 [J]. 大学图书馆学报, 2012, 30 (4): 11-13.
- [14] “一带一路”中国高校大有可为 [EB/OL]. [2018-06-10]. http://www.jyb.cn/zgjyb/201705/t20170504_627977.html.

作者简介

李娟, 女, 1970年生, 硕士, 副研究馆员, 研究方向: 数字资源整合与发现、图书馆应用系统、数字资源建设, E-mail: xiaomi70@qq.com。
张雪蕾, 女, 1988年生, 硕士, 馆员, 研究方向: 机构知识库、电子资源流量控制。
陈楠楠, 女, 1975年生, 硕士, 馆员, 研究方向: 查收查引、竞争情报分析。
杨峰, 男, 1968年生, 硕士, 副研究馆员, 研究方向: 文献资源建设、图书馆集成系统。

Construction and Practice of the Belt and Road Special Database of Silk Road Science and Technology Knowledge Service System

LI Juan ZHANG XueLei CHEN NanNan YANG Feng
(Xi'an Jiaotong University Library, Xi'an 710049, China)

Abstract: This paper summarizes and analyzes the status of the topical knowledge service guarantee of the Belt and Road development in China, and introduces the construction goals, framework, and focuses on special database construction practice of Silk Road Science and Technology Knowledge Service System of Xi'an Jiaotong University, including the formulation of resource construction planning for different objects, the screening of high-quality resources, the collection and processing of multi-channel resources and the evaluation, the evaluation of the use effect, etc.. It also puts forward some thoughts on the construction of the special resources of the Belt and Road.

Keywords: Belt and Road; Special Database; Resource Construction

(收稿日期: 2018-12-18)