

“基于 CADAL 平台的资源共享与应用”国际研讨会综述

□ 范晨晓*

摘要 2017年6月15—16日,“基于CADAL平台的资源共享与应用”国际研讨会在浙江大学成功举办。与会专家围绕数字资源共享与应用主题开展了深入讨论,探讨了数字时代图书馆的挑战和机遇,分享了国内外在资源数字化、数字资源整合、合作共享、数字人文等方面的经验。会议不仅有技术上的创新,更有大量的实践案例,为如何建设更好的数字图书馆及提供数字资源服务提出了实践方向和创新思路。

关键词 数字化 数字资源整合 合作共享 数字人文 CADAL

分类号 G250.76

DOI 10.16603/j.issn1002-1027.2017.05.001

2017年6月15—16日,“基于CADAL平台的资源共享与应用”国际研讨会在浙江大学紫金港校区召开。会议由大学数字图书馆国际合作计划(China Academic Digital Associative Library,简称“CADAL”)项目管理中心、数字图书馆教育部工程研究中心、中国工程科技数据和知识技术研究中心联合主办。来自哈佛大学、牛津大学、北京大学、清华大学等国内外高校的100余位代表参加了会议。浙江大学副校长、图书馆馆长罗卫东,教育部高等学校图书馆情报工作指导委员会主任、北京大学图书馆馆长朱强,哈佛大学燕京图书馆馆长郑炯文,CADAL管理中心副主任庄越挺在开幕式上分别致辞。

CADAL作为国家投资建设的教育部“211工程”重点工程,由浙江大学联合国内外高等院校、科研机构共同承担。自2002年启动至今,共建立了8个数据中心、33个服务中心、2个数字化加工基地和40余个数字化加工中心,形成了全世界最大的资源数字化网络,建成的全文数据库总量达250万册(件),囊括中外文图书、音视频资料以及报刊论文等重要文献。罗卫东在会议开幕式上致辞指出,CADAL项目已取得丰硕成果,未来还将进一步与中国工程院及其他机构合作共建,共谋发展。

会议以CADAL资源为基础,以合作和共享为切入点,进行了跨文化、跨地区的学术交流与研讨,报告主题涵盖了数字资源组织揭示、合作共享、数字人文等热点问题。

1 当前数字资源建设背景的变化

信息化浪潮风起云涌、席卷全球。高校信息环境和知识服务方式都随之发生了急剧的变化,高校图书馆事业也因此进入一个新的发展阶段。

全国政协外事委员会主任、中国工程院潘云鹤院士作了题为“数字图书馆的智能化浪潮”报告。潘院士在报告中指出,当前数字资源建设背景巨变,图书馆面临众多挑战,因此,数字图书馆的建成只是一系列巨变的一个开端。它将随着信息基础设施升级、数据资源膨胀和人工智能技术的巨变而继续变化。计算机图灵奖获得者、卡内基梅隆大学计算机与机器人学教授、CADAL创始人之一的美方代表雷伊·雷蒂(Raj Reddy)作题为“为子孙后代保存我们的过去和现在”的报告。雷伊·雷蒂在报告中强调了数字保存的重要性与可能性,他强烈呼吁一个世界范围内的新版权保护规则的出现,并建议成立中国国家数字档案馆以保存国家文化遗产。

两位教授的思考和见解深邃独特、高瞻远瞩。

* 通讯作者:范晨晓,ORCID:0000-0002-7921-3644,cxfan@zju.edu.cn。

图书馆智能化浪潮和数字保存的重要性也让大家思索下一步数字图书馆应走向何方。

2 数字资源组织揭示

数字资源是数字图书馆建设的核心内容、立足之本,其重要性不言而喻。长期以来,图书馆都非常重视数字资源建设。会上多位国内外图书馆馆长分享了各自馆内数字资源组织揭示的经验。

2.1 特藏资源数字化

中国现存最早的私家藏书楼——天一阁,曾经为保护藏书楼而制定了极为森严的规定。很长一段时间内封闭甚严,几绝人迹。如今的图书馆自然不比古时藏书楼,但是如要将古籍珍本深藏馆中,且隔着山高水远,读者恐也不觉得十分便利。所幸图书馆人在努力,通过善本数字化、特藏电子化等将古籍珍本化身千百,传本扬学。

哈佛大学燕京图书馆馆长郑炯文作题为“CADAL 数字图书馆与哈佛燕京图书馆之合作及其他”的报告^[1]。哈佛燕京图书馆用 10 年时间完成了 4200 种 52000 册善本的数字化并免费上网公开。报告中还介绍了 2011 年至今已经做了 30 多项数字化项目,接下来还要进行更多数字化项目,如费吴生档案等。斯坦福大学东亚图书馆中国研究馆员薛昭慧作题为“从资源数字化到数字人文:斯坦福大学图书馆的实践”的报告,向大家展示了斯坦福大学东亚图书馆特藏资源的数字化情况,并与大家分享了在资源发现和分享上的创新,如一些开放源代码软件的使用,与第三方机构的合作经验等。

梵蒂冈图书馆东亚馆员余东女士作题为“梵蒂冈图书馆中国和远东地区古籍写本文献及其电子化现状”的报告^[2]。梵蒂冈图书馆的中国古籍文献的电子化工作在 2008 年正式展开,目前已基本完成。现在梵蒂冈图书馆电子化项目的网站上(<http://digi.vatlib.it/>),已有 6000 余种电子化的善本和写本供参阅。

厦门大学图书馆馆长萧德洪作题为“东南亚研究文献的访求”的报告^[3]。萧德洪馆长探讨了如何发扬传统优势和抓住时代机遇,以案例分享的方式来讨论东南亚地区研究相关文献的采集和访求的路径,进一步加强研究型馆藏的建设。清华大学图书馆窦天芳副馆长作题为“在图书馆里讲述清华故事——以‘清华印记’互动体验区建设为例”的报告^[4]。

“清华印记”互动体验空间包括数字学术、瀑布流图书借阅、历史长廊、数字人文等。该空间体现了数字化技术、视觉传达、图书馆服务三者有机结合的实施经验和建设模式。

2.2 从数字化到数据化

美国计算机科学家、图灵奖获得者吉姆·格雷(Jim Gray)提出了科学研究的“第四范式”。“第四范式不仅是研究方法的变化,更是人类思维方式的重大变化。在这种研究范式中,研究者面对浩如烟海的数据,不再抽取少量的样本进行分析,而是把所有数据作为一个整体,利用数据挖掘、计算、分析等技术,直接从数据中探寻所需要的信息、知识和智慧。”^[5]

杜克大学图书馆国际与地区研究部的中国研究馆员周珞从照片元数据的角度切入,作题为“建立一个国际历史照片数据库的元数据考虑”的报告^[6],介绍了现在两个主要国际数据的标准:一是欧洲数字图书馆的数据模型(Europeana Data Model, EDM)。二是北美主要采用的美国公共数字图书馆数据模型(Digital Public Library of America, DPLA)。CADAL 现在采用的是都柏林核心元素集(Dublin Core Element Set, DC)。她强调了数据的一致性和灵活性以便实现“数据链接”的可能。

牛津大学中国中心图书馆馆长邵玉书作题为“Serica, 牛津大学中文古籍数字化项目与未来的资源共享”的报告^[7]。邵玉书馆长向大家介绍了牛津大学博德利图书馆目前正在进行的中文编目和数字化项目(Cataloguing and Digitization Project of the Bodleian's Pre-modern Chinese Collections, 简称“Serica”),该项目对馆藏海内外的孤本、手稿进行数字化,以及元数据结构化等工作,并免费开放给公众。报告中还介绍了一些手稿数字化工具,如清单编辑器(IIF Manifest Editor),该编辑器可以编辑元数据,新建图像顺序等。

上海师范大学副教授孙红杰作题为“数字化艺术藏品的增值服务个案:基于 CADAL 资源的‘数字人文’中英合作项目”的报告^[8],孙红杰老师曾作为 CADAL 项目的“境外协调员”,在英国牛津大学开展“数字文化遗产”方面的中英合作研究。报告围绕艺术品浏览系统和中国音乐文物数据库介绍了已开展的数据关联试验及为数字化艺术藏品提供系统化增值服务的意义、潜能和途径。艺术品浏览系统是

荟萃 1 万余幅高清分辨率数字绘画和书法图像的数据库,以中国作品为主,涵盖了西晋以来各个历史时期、涉及各个艺术题材的艺术品资源。对于这些艺术品,CADAL 除了做整体的元数据标注之外,还对一些画作上的基本元数据进行优化。中国音乐文物数据库以 CADAL 数百万册电子图书中涉及到的中国传统音乐的海量图片、图书、乐谱资源为依托,对这些音乐文物的元数据进行标注。

3 数字资源合作共享

在数字资源日益丰富、读者需求不断更新的大背景下,图书馆间、数字资源项目间建立资源共享联盟是解决单个图书馆或项目资源匮乏及避免重复建设的有效途径,也是图书馆界的大势所趋。“合则强,孤则弱”,千年前的话同样适用于当今数字时代的图书馆。图书馆资源共建共享,是未来图书馆可持续发展的重要方向。

3.1 馆际合作

伯克利加州大学东亚图书馆馆长周欣平作题为“从收藏到链接:数字时代伯克利加州大学图书馆与亚洲及太平洋地区图书馆之间的合作与资源共享”的报告^[9]。伯克利加州大学东亚图书馆在馆藏数字化方面与国内外积极合作,如与台湾中央图书馆合作数字化扫描,与中华书局合作赵元任档案数字化项目,与日本国文学研究资料馆合作 2800 种日本江户时代写本目录和图录项目等。

哥伦比亚大学东亚图书馆馆长程健作题为“哥伦比亚大学图书馆的数字化资源多重分享模式”的报告,与大家分享了数个合作典型案例,如中国门神纸马收藏数字化项目、甲骨收藏 3D 数字化项目、顾维钧档案数字化项目等。

上海交通大学图书馆馆长陈进作题为“从资源共享平台建设到特色文化建设”的报告^[10]。陈进馆长的报告以上海交通大学参与 CADAL 建设的若干项目为例,包括特色资源数字化、数据中心建设、读者服务立体化协同工作平台建设等,详尽分析了资源共建共享体系为图书馆发展带来的深远影响。

3.2 联盟共享

机构知识库是一种有效和积极的知识保存、组织、管理机制,也是促使学术健康发展、自由开放的重要手段。北京大学图书馆朱强馆长作题为“高校机构知识库联盟的建设与发展”的报告。朱强馆长

在报告中回顾了 2011—2017 年中国高校机构知识库的发展,并介绍了中国高校机构知识库联盟进展,朱强馆长还特别就“2C—大数据合作”即中国高等教育文献保障系统(CALIS)和 CADAL 两者间的数据合作进行了阐述。

超星集团副总经理叶艳鸣作题为“‘联盟+’—互联网环境下的共享联盟服务推广”的报告^[11],他在报告中指出,以“共建、共知、共享”为核心的图书馆共享联盟是图书馆发展中的重要思想,但在发展过程中,“重建设、轻应用、约束多、传播差”的现象普遍存在。报告以“互联网+”的开放视角,探讨了图书馆共享联盟如何借助具有广泛社会影响力的互联网学术传播渠道扩大联盟的社会认同和共享服务的思路。

4 数字人文

数字人文是目前国际上的新兴学科和前沿研究领域,具有创新性强、多学科交叉、实用性突出的特点。近 10 余年来,北美、欧洲和亚洲的重要高校和研究机构,如哈佛大学、斯坦福大学、牛津大学等开始涉足并深入数字人文研究领域。国内外图书馆界也积极响应,展开数字人文研究。可以说,数字人文是目前图书馆为数字学术开展服务的一个重点领域。

浙江大学人文学院徐永明教授应邀为大会作题为“数字资源、人文数字地图与文史研究”的报告^[12],从人文学者的角度来思考数字人文。徐永明教授在数字人文研究方面已有不少学术产出。如使用中国历代人物传记资料库、中国历史地理信息系统、哈佛大学世界地图计划(World Map)等进行学术研究,对全元文作者、明清女性作者、云南清代作者的地理分布进行了可视化研究。

4.1 人文数据库或数据集的建设

北京大学图书馆朱本军和聂华在《跨界与融合:全球视野下的数字人文——首届北京大学“数字人文论坛”会议综述》一文中指出,“数字人文的内容主要集中在四个层次:一是人文数据库或数据集的建设。这一层次主要表现为两个方面:其一是将非数字的人文资料加工转化为数字内容……其二是对非结构化的数字文本内容按照某种使用目的进行规范化标注著录的数据集(Dataset)建设……”^[13]

一是将人文资源数字化,形成特色资源数据库,

方便学者进行史料或文本的深度挖掘。浙江大学求是特聘教授、蒋介石与近现代中国研究中心主任陈红民作题为“蒋介石资料数据库的建设构想”的报告^[14]。蒋介石资料数据库是浙江大学蒋介石研究中心与 CADAL 合作项目,力求做到资料完备、内容丰富、使用便捷、共建共享。国家图书馆出版社民国文献编辑室副主任、数字出版部主任助理李强作题为“国家图书馆出版社历史文献数字化的探索与实践”的报告^[15],李强副主任与大家分享了国图出版社的三个重要数据库,分别是中华再造善本数据库、民国时期文献总库·民国图书数据库和中国历史人文传记资源库。

二是将人文资源数据化,形成数据集,方便多学科学者进行交叉研究。中国历代人物传记资料库(China Biographical Database, CBDB)就是这一类型的数据集。北京大学历史学系、CBDB 北大小组项目经理杨光作题为“中国历代人物传记资料库(CBDB)的数据建设与网络基础设施建设”的报告^[16]。CBDB 作为一个关系型数据库,致力于将中国历史上的人物传记资料根据不同方面的内容进行分类和编码,把这些古籍当中的内容作为一些数据来存储到数据库当中。报告从数据建设与网络基础设施建设的实践出发,总结 CBDB 在数据库建设及其他数字人文项目合作等方面的经验,并对项目未来的发展方向、与其他数字人文项目的合作设想作了展望。

4.2 人文数字工具的开发使用

在文字识别、文本挖掘方面,哈佛大学东亚语言与文明系讲师、哈佛费正清中国研究中心博士后德龙作题为“数字人文与数字图书馆:中国历代文献的文字识别、群众外包及文本挖掘”的报告^[17]。报告介绍了中国哲学书电子化计划(<http://ctext.org/zhs>)中的主要技术,可归类为三种:一是前现代中文资料的文字识别技术(OCR),该系统已处理 2500 多万页资料,其结果已在网上公开;二是群众外包、公众参与,世界各地的用户可以参与数字化过程并积极协助内容的扩展;三是使用既可以实现与其他线上工具之间的整合,又能提供文本挖掘途径的开放式应用程序界面(API),从而有效地借用日益增长的资料库文本内容来达到数字人文研究和教学的目的。

在特藏文物数字化技术层面,浙江大学图书馆

技术总监徐建刚作题为“CADAL 浅刻文物特藏的数字化技术”的报告^[18]。为了真实记录和还原 CADAL 项目中的部分浅刻文物特藏,需要探索有别于传统扫描的数字化技术。浙江大学图书馆、哥伦比亚大学图书馆等成员单位经过调研选择了反射变换成像技术对甲骨等浅刻文物进行数字化加工和展示,获得了很好的效果。

5 结语

图书馆的数字化进程已历 20 余载,数字化、数据化对教学科研的影响日益深远,美国信息技术与创新基金会曾于 2013 年 11 月发布报告“数据创新 101:支持数据驱动创新的技术和政策介绍”^[19],强调数据的重要性,如果说石油是 20 世纪经济发展的“助推剂”,那么数据将成为 21 世纪经济发展的“助推剂”。文本数字化,元数据规范化,都是为知识服务的智能化鸣起的前奏。

通过这次会议,全球视野下的资源合作共享途径得以拓宽,特色馆藏数字化、数据化的经验得以分享,数字人文的内涵与实践得以丰富。学者们精彩的报告和交流也必将对数字时代的图书馆建设起到积极的推动作用。

参考文献

- 1 郑炯文. CADAL 数字图书馆与哈佛燕京图书馆之合作及其它[EB/OL]. [2017-07-10]. <http://www.cadal.cn/drsa2017/download.zh.html>.
- 2 余东. 梵蒂冈图书馆中国和远东地区古籍写本文献及其电子化现状[EB/OL]. [2017-07-10]. <http://www.cadal.cn/drsa2017/download.zh.html>.
- 3 萧德洪. 东南亚研究图书的访求[EB/OL]. [2017-07-10]. <http://www.cadal.cn/drsa2017/download.zh.html>.
- 4 窦天芳. 在图书馆里讲述清华故事——以“清华印记”互动体验区建设为例[EB/OL]. [2017-07-10]. <http://www.cadal.cn/drsa2017/download.zh.html>.
- 5 郑永晓. 加快“数字化”向“数据化”转变——“大数据”、“云计算”理论与古典文学研究[J]. 文学遗产, 2014(6): 141—148.
- 6 周路. 建立一个国际历史照片数据库的元数据考虑[EB/OL]. [2017-07-10]. <http://www.cadal.cn/drsa2017/download.zh.html>.
- 7 邵玉书. Serica, 牛津大学中文古籍数字化项目与未来的资源共享[EB/OL]. [2017-07-10]. <http://www.cadal.cn/drsa2017/download.zh.html>.
- 8 孙红杰. 数字化艺术藏品的增值服务个案:基于 CADAL 资源的“数字人文”中英合作项目[EB/OL]. [2017-07-10]. <http://www.cadal.cn/drsa2017/download.zh.html>.
- 9 周欣平. 从收藏到连接:数字时代伯克利加州大学图书馆与亚洲及太平洋地区图书馆之间的合作与资源共享[EB/OL]. [2017-07-10]. <http://www.cadal.cn/drsa2017/download.zh.html>.

- 07-10]. <http://www.cadal.cn/drsa2017/download.zh.html>.
- 10 陈进. 从资源共享平台建设到特色文化传播[EB/OL]. [2017-07-10]. <http://www.cadal.cn/drsa2017/download.zh.html>.
- 11 叶艳鸣. “联盟+”——互联网环境下的共享联盟服务推广[EB/OL]. [2017-07-10]. <http://www.cadal.cn/drsa2017/download.zh.html>.
- 12 徐永明. 数字资源、人文数字地图与文史研究[EB/OL]. [2017-07-10]. <http://www.cadal.cn/drsa2017/download.zh.html>.
- 13 朱本军, 聂华. 跨界与融合: 全球视野下的数字人文——首届北京大学“数字人文论坛”会议综述[J]. 大学图书馆学报, 2016(5): 16-21.
- 14 陈红民. 蒋介石资料数据库的建设构想[EB/OL]. [2017-07-10]. <http://www.cadal.cn/drsa2017/download.zh.html>.
- 15 李强. 国家图书馆出版社历史文献数字化的探索与实践[EB/OL]. [2017-07-10]. <http://www.cadal.cn/drsa2017/download.zh.html>.
- 16 杨光. 中国历代人物传记资料库(CBDB)的数据建设与网络基础设施建设[EB/OL]. [2017-07-10]. <http://www.cadal.cn/drsa2017/download.zh.html>.
- 17 德龙. 数字人文与数字图书馆: 中国历代文献的文字识别、群众外包及文本挖掘[EB/OL]. [2017-07-10]. <http://www.cadal.cn/drsa2017/download.zh.html>.
- 18 徐建刚. CADAL 浅刻文物特藏的数字化技术[EB/OL]. [2017-07-10]. <http://www.cadal.cn/drsa2017/download.zh.html>.
- 19 Castro D, Korte T. Data Innovation 101[EB/OL]. [2017-07-30]. <http://www.datainnovation.org/2013/11/data-innovation-101/>.
- 作者单位: 浙江大学图书馆, 杭州, 310027
收稿日期: 2017 年 8 月 21 日

Review of the International Conference on Digital Resources Sharing and Application Based on CADAL Service Platform

Fan Chenxiao

Abstract: June 15-17, 2017, the International Conference on Digital Resources Sharing and Application based on CADAL service platform was held in Zhejiang University. The presentations and discussions addressed focused on the challenges and opportunities of libraries in the digital era and shared the following topics: resource digitization, digital resource integration, interlibrary cooperation, Digital Humanities and etc. A great number of innovative technologies and practices have been shared during the conference to provide a practical direction and innovative ideas for the construction of better digital libraries and the provision of digital resources services.

Keywords: Resource Digitization; Digital Resource Integration; Cooperation and Sharing; Digital Humanities; CADAL

中山大学程焕文教授当选国际图书馆协会联合会管理委员会委员

2017 年 6 月 15 日, 国际图书馆协会联合会(International Federation of Library Associations and Institutions, IFLA) 秘书长 Gerald Leitner 先生发来电子贺信, 祝贺中山大学资讯管理学院教授、中山大学图书馆馆长、校长助理程焕文教授当选 IFLA 管理委员会委员, 任期两年(2017-2019 年)。

IFLA 成立于 1927 年, 拥有 140 多个国家和地区的 1350 个学会和机构会员, 是全球最具代表性的图书馆国际组织。IFLA 管理委员会由国际图联主席、秘书长、10 名竞选委员和 7 个委员会主席组成, 是 IFLA 的最高管理机构。中国图书馆学会作为发起会员之一, 自 1929 年开始参加 IFLA 世界图书馆大会。1993 年中国图书馆学会的代表第一次当选 IFLA 管理委员会委员, 其后中国先后有三位代表当选。2006 年中文成为其七种官方语言之一。2012 年后连续五年 IFLA 管理委员会设有中国代表, 在本次换届选举中, 经中国文化部和图书馆学会推荐, 程焕文教授作为亚洲唯一的竞选代表在激烈的竞选中成功当选, 为中国图书馆界争取到了应有的国际地位。

程焕文教授从事图书馆学教学、科研和管理工作 30 余年, 是国际知名的图书馆学家和图书馆专家, 现任国务院学位委员会全国图书情报专业学位研究生教育指导委员会委员、教育部高等学校图书馆学学科教学指导委员会副主任委员、教育部高等学校图书馆情报工作指导委员会副主任委员、广东高校图书馆情报工作指导委员会主任委员、中国图书馆学会副理事长、广东图书馆学会名誉理事长, 在国内外具有广泛的学术影响和专业影响力。作为国际图联管理委员会委员, 程焕文教授将在两年的任期内致力于促进国际图书馆界对中国图书馆事业的认知, 宣传中国图书馆事业, 加强中国图书馆界与世界各国图书馆的交流与合作, 共同推进人类图书馆事业的发展。

【信息来源: 中山大学新闻网: <http://news2.sysu.edu.cn/news01/150278.htm> (2017-06-16)】