

二十一世纪前十年数字图书馆技术研究 and 系统建设的9个思潮

□ 张智雄 刘建华 / 中国科学院国家科学图书馆 北京 100190

洪娜 许德山 徐健 / 中国科学院国家科学图书馆 北京 100190

/ 中国科学院研究生院 北京 100049

摘要：文章分析了二十一世纪前十年数字图书馆技术研究和系统建设的历程，总结了这十年来在数字图书馆技术研究和系统建设中的9个思潮：数字化和数字资源研究与建设思潮、门户研究与建设思潮、集成和互操作研究与建设思潮、“下一代数字图书馆”研究与建设思潮、数字仓储研究与建设思潮、数字保存研究与建设思潮、本体和语义网研究与建设思潮、Web2.0思潮、开放存取与开源思潮。

关键词：数字图书馆，技术研究，系统建设，历程回顾

DOI: 10.3772/j.issn.1673-2286.2010.03.004

世纪之交之时，全球数字图书馆技术的研究正处于一个承上启下的时代。从数字图书馆研究和建设的项目资源来看，美国的DL-I已经结束，DL-II正在稳步跟进；欧洲已经进入了其第5框架实施期间，正在为构建用户友好的信息社会而努力；而英国的eLib计划也已结束，RSLP（The Research Support Libraries Programme）刚刚启步。也正是在2000年，两个数字图书馆会议ACM Digital Libraries Conference和IEEE的Conference on Advances in Digital Libraries宣布合并，发起新的ACM/IEEE CS Joint Conference on Digital Libraries（JCDL）。这些都预示着数字图书馆的技术研究和系统建设将在二十一世纪迎来大发展的阶段。

回顾二十一世纪00年代数字图书馆研究和建设的历程，可以深刻感受到这十年来数字图书馆技术

研究和系统建设呈现出一种活跃繁荣、与时俱进的景象。笔者总结这十年来数字图书馆技术研究和系统建设的历程，期望借用“思潮”这一可能不十分严肃的用语来反映这十年数字图书馆技术研究和系统建设过程中出现的核心问题、技术理念、项目实践和系统建设情况。这十年来数字图书馆技术研究和系统建设的9个思潮可总结为：数字化和数字资源研究与建设思潮、门户研究与建设思潮、集成和互操作研究与建设思潮、“下一代数字图书馆”研究与建设思潮、数字仓储研究与建设思潮、数字保存研究与建设思潮、本体和语义网研究与建设思潮、Web2.0思潮、开放存取与开源思潮。现分述如下。

1 数字化和数字资源研究与建设思潮

按照Abby Smith的观点，将主要精力投入到数字化项目中的数字图书馆，被称为“第一代数字图书馆”^[1]。持同样观点的还有Harvard大学的Dale Flecker^[2]。

数字化和数字资源建设是数字图书馆建设的基础。与传统图书馆一样，数字图书馆需要收集、选择、描述、保存各类数字资源，并提供相关服务实现资源存取和利用。最显著的差别体现在，即使在数字图书馆建设的初期，由于使用了数字化方式，数字内容能够被更多、更广泛的人们所获取和利用的能力就已经得到了空前的提升。

在美国数字图书馆计划进行的第一个研究阶段（DL-I时代（1994-1998）），如何实现印本资源、模拟信号的数字化，如何实现各类数字资源的建设和管理成为了数字图书馆研究的重要主题。例如在1994年，美国国会图书馆发

起了国家数字图书馆项目,并且以American Memory历史资源作为这一项目的标志性建设内容^[3];1995年,在NSF和University of Michigan Digital Library Project的资助下^[4],University of Michigan发起了JSTOR (Journal Storage Project)项目^[5]。

二十一世纪00年代,随着几个大的数字化和数字资源建设项目的出现,数字化和数字化资源的研究和建设出现了两次高潮。

第一次数字化和数字化资源建设的高潮出现在20世纪90年代末和21世纪00年代初。这一时期,出现了由美国Carnegie Mellon University和中国、印度的相关机构共同发起的百万图书计划项目(The Million Book Project)^[6],又名全球图书馆(the Universal Library)^[7],这一项目扫描包括中文、英文、泰卢固文、阿拉伯文在内的图书,并利用OCR(光学字符识别)技术实现文字的识别以实现对这些图书的全文检索,并基于Web提供这些图书的免费服务。

与此相对应的是欧盟第五框架计划在2001年启动的被称为“欧洲知识网关”的“欧洲图书馆”项目(The European Library: Gateway to Europe's Knowledge)^[8],这一项目由芬兰、德国、意大利(佛罗伦萨)、意大利(罗马)、荷兰、葡萄牙、斯洛文尼亚、瑞士和英国的9个国家图书馆发起,其目标是促进欧洲国家图书馆的图书、期刊、音像资料及其他资源的元数据和数字内容可以被广泛地检索和传递。

除了图书文献之外,在人类文化和历史遗产方面有重要影响的还有Global Memory Net项目^[9]。这一项目启动于2000年,由美国国家自然科学基金的国际数字图书馆项目

资助,目标在于构建人类文化历史遗产的图像资源库。

除上述较大的几个项目之外,很多著名的高校也在这一阶段发起了数字图书馆项目,如University of California在1997年发起了California Digital Library(CDL)项目^[10],Harvard大学在1998年发起了图书馆数字计划(Library Digital Initiative,LDI)^[11]。

第二次数字化和数字资源建设的高潮出现在2004年10月Google发起Google Print项目之后。同年12月Google将这一项目扩展为Google Print Library Project项目,并与University of Michigan、Harvard(Harvard University Library)、Stanford(Green Library)、Oxford(Bodleian Library)和New York Public Library等5家重要图书馆签署合作协议,将在十年内,实现1500万卷图书的数字化,并通过Google Books提供服务。这一项目后来又吸引了很多学术图书馆。2009年10月,Google宣布已经扫描了1000多万卷图书。

欧洲以欧洲图书馆(TEL)^[12]和Europeana^[13]的建设掀起了数字化和数字资源建设的高潮。2005年欧洲图书馆(TEL)作为一个联网服务开始投入运行,随着后续项目的投入,其数字内容的范围不断增长。同年欧盟发布了欧洲信息社会i2010计划(European Information Society i2010 Initiative)^[14],其中一项重要的功能就是推动欧洲数字竞争力的提高。2008年被认为是欧洲数字图书馆的Europeana在i2010计划的eContentplus项目支持下开始上网服务。Europeana最初被称为欧洲数字图书馆网络(European digital library network – EDLnet),它吸收

了欧洲的100多家机构参加,所建设的数字资源不仅包括图书馆藏书,还包括了博物馆、档案馆和其他人类和历史文化遗产。到目前为止,Europeana数字资源包括了600万条数字内容,内容类型涉及图像、文本、音频、视频等多种形式。

在Google发起其Google Book项目之后,微软也在2006年后期发起过类似的名为“Live Search Books”^[15]的项目,但由于种种原因于2008年放弃此项目,其数字化的馆藏成为Internet Archives的一个部分^[16]。Google Book目前的竞争对手还包括法国国家图书馆的Gallica 2数字化图书项目(Gallica 2 Digitised Book Project),以及由Indiana University和University of Michigan来领导和构建的HathiTrust数字图书馆。这两个项目都于2008年正式提供服务。

全球数字化的浪潮直接推动了全球数字图书馆项目的建设。2009年4月,联合国教科文组织(UNESCO)^[17]、美国国会图书馆及其他图书馆发布了公开免费利用的全球数字图书馆(WDL)^[18]网站,这一数字图书馆以高质量的、反映UNESCO成员国家的文化历史遗产为特色。

上述两次数字化和数字资源研究和建设的思潮,从技术上来讲,推动了OCR、DRM(数字版权管理)、内容管理、数字对象、数字仓储等技术的研究和发展。特别是第二次数字化建设的大潮,迫切需要加快大规模数字化的速度、提高大规模数字化的质量、完善对传统文献的字符识别。因此,00年代后期,数字图书馆界还出现了很多解决大规模数字化问题的项目和系统,如欧洲的IMPACT项目的目

标就是解决大规模数字化中的相关问题,而由SUB Göttingen开发的Goobi软件系统则是一个数字化过程中对工作流进行管理的工具。

2 门户研究与建设思潮

在现实数字图书馆的建设中,客观地存在着资源和系统分散、独立存在的情况。如何将各类分布资源和服务按照相互关系进行有机组织,实现信息资源、信息服务和信息系统的整合,提供集成连贯的信息环境是数字图书馆建设的一个重要要求。随着互联网门户技术的发展,在数字图书馆界也出现了利用“门户”来解决上述资源和系统整合集成的思潮。

从技术上讲,这十年的数字图书馆研究和建设中,也有两种类型的门户存在。第一种门户被称为信息门户(Information Gateway)、学术门户(Scholars Portal)或主题门户(Subject Gateway),主要侧重于从信息资源整合的角度来构建门户;而另一种门户主要从技术框架和系统集成的角度出发,目标在于通过门户(Portal)系统将各个独立存在的系统有机整合起来,提供一个集成连贯的用户环境,以方便用户在同一平台之上对多个系统的使用。

早在20世纪90年代末,在英国Electronic Libraries(eLib)项目和欧盟第4框架项目的支持下,英国和欧洲都构建了许多信息门户,目标在于发现、描述和组织Web上存在着的高质量的信息资源,构建相关的浏览检索服务,使这些资源可以被用户使用。

进入21世纪之后,门户研究

和建设的热潮不减。英国2000年发起了资源发现网络(Resource Discovery Network, RDN)项目^[19],将原先建立的有关生物医学、工程计算、人文、物理和社会科学的5个信息门户集成,形成一个统一的RDN hub^[20]。2006年,RDN Hub进一步发展,形成了目前拥有数百万条经过领域专家认真挑选的、重要Internet学术资源的Intute信息门户^[21]。

与英国RDN项目同期的项目还有:欧盟第五框架支持的Renardus项目(欧洲学术主题网关服务,Academic Subject Gateway Service Europe)^[22,23];美国研究图书馆(ARL)于2002年发起的Scholars Portal Project项目^[24];德国2003年开始在线服务的vascodaportal项目^[25]。

为了实现对信息门户中的资源进行有效的收集、整理、加工、管理并提供服务,University of Wisconsin-Madison的The Internet Scout Project项目于2000年在Andrew W. Mellon Foundation的资助下,开发了开源的Scout Portal Toolkit(SPT)工具软件^[26,27]。这一工具被应用到很多信息门户的建设之中。另一个被广泛应用于信息门户建设的软件系统是由Library of the University of California, Riverside开发的iVia^[28,29]。与SPT相比,它更强调网络内容爬行、下载和索引的功能。

MyLibrary系统被认为是一种具有个性化功能的信息门户系统。Cornell University Library在2000年推出了MyLibrary服务^[30],而由The North Carolina State University Library于2000年发布的MyLibrary@NCState^[31]开源软件也被很多图书馆

所应用。

真正意义上的门户(Portal)是能够将各个独立存在的系统有机整合起来的一种软件系统框架。与前面所说的信息门户不一样,Portal技术实现单点登录,集成许多其他系统,只需一次登录,就可以存取多个系统;实现内容聚合,在一个页面上聚合来自其他多个来源的内容和信息;实现信息的有效传递,将适当的信息传递给适当的用户;构建面向用户的视图,使用户可以定制自己的门户界面;实现科研资源的集成和导航。在JISC的推动下,英国发起了Portals Programme^[32,33],促使很多英国的图书馆、大学和科研机构纷纷采用Portal技术构建机构门户。

为了支持机构门户的构建,一些门户软件系统也被开发了出来。这些软件要求支持Portlet技术,特别是JSR 168和WSRP标准协议,能够方便地提供单点登录、内容定制和个性化功能。在机构门户的构建中,被广泛使用的门户软件系统有uPortal^[34]和Liferay^[35],被广泛使用的单点登录软件(SSO)为Shibboleth系统^[36]。

3 集成和互操作研究与建设思潮

在上世纪90年代末,影响至今的很多数字图书馆互操作协议已经开始孕育。例如1998到2000年间Herbert Van de Sompel及其同事完成了SFX的研究;1999年10月,在图书馆和信息资源委员会(CLIR)、数字图书馆联盟(DLF)、学术出版和学术资源联盟(SPARC)、研究图书馆协会(ARL)等的支持下,参加“统一预印本服务”

会议的E-Print建设机构发起了OAI计划^[37,38]；同样在1999年10月的法兰克福图书展销会上，基于DOI（Digital Object Identifier）和CrossRef^[39]的期刊全文（和参考文献）链接系统开始出现。这些事件预示着在21世纪，数字图书馆的建设将会迎来一个数字图书馆集成和互操作技术与系统建设的高潮。

2001年，OAI-PMH（The Open Archives Initiative Protocol for Metadata Harvesting）1.0版本发布。OpenURL^[40]在数字图书馆联盟、ICOLC和Ex Libris的稳步推动下，被迅速应用于学术信息环境之中。2000年Ex Libris推出了SFX Server的Beta版本并在很多大学图书馆中进行了应用测试；2004年OpenURL成为美国国家标准^[41]。CrossRef也于2000年6月开始正式投入运行^[42]。

数字图书馆其他相关的数据标准和互操作协议也在21世纪得到了大发展。例如在检索协议方面，2000年有相关机构开始从事SRW/SRU/CQL研究，到2002年，推出了实验性质的SRW/SRU/CQL 1.0版本。2004年2月第一个正式的1.1版本发布，1.2版本在2007发布，而到2009年7月，OASIS Search Web Services Technical Committee^[43]推出了SRU 2.0和CQL 2.0草案^[44]。在数字参考咨询方面，2004年由NISO Standards Committee AZ发布了在数字参考咨询系统间交换问题答案数据，以实现协作咨询的问题/答案交换协议（QATP）^[45]。

近年值得关注的数字图书馆互操作协议是开放存档计划的对象重用和交换协议（Open Archives Initiative Object Reuse and Exchange，OAI-ORE）^[46]，这一协

议的目标在于定义对Web资源聚合体进行描述和交换用的标准。OAI-ORE项目于2006年10月发起，2008年10月发布了OAI-ORE规范的v1.0版本。

在这十年中，还出现了很多数字图书馆的数据和编码标准，这也推动了数字图书馆的集成和互操作的研究和建设^[47]。如2001年，DLF提出了元数据编码和传输标准（METS）^[48]；2003年PREMIS开始保存元数据的设计，并于2005年5月发布了PREMIS保存元数据字典^[49]。

为了实现数字图书馆之间集成和互操作，除了标准协议之外，一些数字图书馆组织（如数字图书馆联盟DLF）提出需要达成数字图书馆系统框架的共识，构建一个数字图书馆系统的参考模型，以实现各种数字图书馆项目之间的协作和互操作。在这十年间，数字图书馆界提出了两个典型的数字图书馆系统模型。一个为利用组件方式构建数字图书馆资源和服务的OCKHAM框架^[50,51,52,53]；另一个为2007年DELOS提出的DELOS数字图书馆参考模型（Digital Library Reference Model）^[54]。前者曾在2002年的ECDL会议上做过广泛的讨论^[55]，并得到了美国自然科学基金的支持，很多成果已经在2007年以源码的形式进行了发布；而后者仍然是目前ECDL和JCDL会议上讨论的热点。

4 “下一代数字图书馆”研究与建设思潮

随着网络技术的进步，特别是00年代中期Web Services（包括Semantic Web Services）、P2P、

GRID技术的盛行、SOA框架的兴起，以及数字图书馆研究和建设密切相关的eScience的崛起，很多数字图书馆研究者提出了“下一代数字图书馆框架（Next Generation DL Architectures）”的理念^[56,57,58]。持此理念的研究者认为，下一代数字图书馆框架的目标是将原先集中的、中心控制的系统转换成为一种基于P2P网络和SOA（Service-oriented Architectures）框架，可以根据数字图书馆的资源和服务的需求，灵活集成装配的架构。根据他们的理念，下一代数字图书馆框架是一种分散的、灵活分布的、可按需组装的架构，支持数字图书馆项目的协作共建。这一框架最大的特点有两个：（1）可以按照一定的授权方式实现成员间数字内容、服务和计算资源的共享；（2）可以根据需要动态地聚集相关的共享资源，以实现数字图书馆的按需构建。

2006年开始出现的Delos-DLMS^[59]是DELOS推出的下一代数字图书馆管理系统的原型系统，这一系统的核心是构建大规模支撑的、可信赖的中间件系统环境OSIRIS（Open Service Infrastructure for Reliable and Integrated process Support）^[60]，通过它可以实现分布式数字服务的调用和组装。这一原型系统封装了ISIS数字图书馆服务，可以实现多媒体信息管理。它还封装了DELOS成员开发的多种系统工具，可以实现文本、音视频的检索、可视化等功能^[61]。

与DelosDLMS相似的还有2002年以来就一直在完善中的OpenDLlib系统^[62,63]。为了扩展OpenDLlib的存储和处理能力，OpenDLlib于2006年利用gLite网格框架（gLite Grid

infrastructure) 增强了处理海量信息的能力^[64]。

基于P2P (peer-to-peer) 架构的典型应用是欧盟第6框架支持的BRICKS (Building Resources for Integrated Cultural Knowledge Services) 项目^[65,66], 在这一项目中, 利用了基于P2P的分布式框架来实现文化遗产领域内知识和资源的共享。

欧盟第6框架的另外一个项目DILIGENT (A Digital Library Infrastructure on Grid Enabled Technology) ^[67,68]将网格技术和数字图书馆技术有机集成, 其目标在于构建一个基于Grid的数字图书馆的实验平台, 以促进相关的e-Science团体可以按一种安全、协调、动态和有效的方式来实现数字资源和知识的共享协作。DILIGENT项目开发的软件系统gCube^[69]目前已经成为一个被广泛应用的、基于网格的虚拟团队建设环境。

在数字图书馆基础框架方面, 近两年来, 随着云计算技术的不断发展, 基于云计算构建数字图书馆或在数字图书馆中利用云计算技术也成为热门的话题^[70]。基于Apache hadoop^[71]的开放云计算平台, 被广泛地应用于数字图书馆的资源收集、信息存储和海量计算之中。如何利用好Apache Hadoop开放云计算平台, 值得引起国内同行的关注。

5 数字仓储研究与建设思潮

“仓储 (Repository)” 虽然在20世纪90年代后期在数字图书馆界有一些应用, 但在20世纪可能谁都没有想到, 在跨世纪之后的十年,

“Repository” 这一词在数字图书馆界竟然是这样的走红。

其实在20世纪90年代, 被后来称为“Digital Repository”的数字内容管理系统就已经出现, 如Los Alamos National Laboratory的arXiv.org^[72], 英国University of Southampton的CogPrints^[73]和存储电子硕博论文的NDLTD^[74] (The Networked Digital Library of Theses and Dissertations) 等, 当时它们被称为E-print Archive。

OAI计划的提出, 对“Repository”进行了明确的定义, 并提供了互操作的标准规范, 客观上对Repository热潮的形成起到了推动作用。进入二十一世纪之后, 遵循OAI协议的开源机构仓储软件纷纷推出。例如, 2000年University of Southampton推出Eprints系统^[75], 2001年Cornell University发布了Fedora (Flexible Extensible Digital Object Repository Architecture) 系统^[76], 2002年Hewlett-Packard和MIT开发出了Dspace软件系统^[77]。正是这些数字仓储构建工具的出现, 使得数字图书馆界在2000年代初期迅速掀起“数字仓储”建设的热潮。机构仓储 (Institutional Repository) 和主题仓储 (Subject Repository) 成了“数字仓储”的建设主体, 一些研究机构也将其构建的数字仓储称为eScholarship^[78] (如CDL) 或Digital Commons (如DigitalCommons@ILR) ^[79]。

为了让各个机构建设的开放数字仓储可以方便地查询和利用, University of Southampton于2005年建立了开放存取仓储登记系统ROAR^[80], 目前已经有1645个数字仓储在此进行了登记。2006年University of Nottingham大学构建了

开放存储仓储目录OpenDOAR^[81], 可供用户查询其收集的1593个数字仓储。

一些国家和地区建立起了数字仓储建设联盟, 以推动数字仓储系统的建设。如澳大利亚2003年开始建设ARROW项目 (Australian Research Repositories Online to the World) ^[82]来实现对澳大利亚相关机构仓储内容的集中收割和发现服务。荷兰于2006年构建了NARCI门户^[83], 实现对所有荷兰大学、科研机构中数字仓储内容的开放存取。欧盟支持的DRIVER (Digital Repository Infrastructure Vision for European Research) 项目^[84], 目前已经发展成为一个国际间的联盟, 通过其系统可以实现33个国家、249个仓储内容的查询。OCLC^[85]的OAster服务则实现了1100家数字仓储中2300万条元数据的收割和索引, 被认为是数字仓储内容的联合目录。

数字图书馆界在这十年间还积极推动数字仓储在教学、科研、实验等各个方面的深入应用。这方面的主要项目有: 德国Max Planck Society和FIZ Karlsruhe的eSciDoc项目^[86]通过数字仓储建设构建一个eResearch环境, 以实现全球范围的、可跨领域的交流协作。英国JISC R4L (Repository for the laboratory) 项目^[87]在推动数字仓储在实验室数据处理方面起到积极作用。而英国JISC X4L (Exchange for Learning) 项目^[88]的一个重要内容是构建JORUM学习对象仓储服务^[89]。最近University of Hull在RepoMMan、REMAP的基础之上, 开始了一个名为Hydra的国际合作项目, 其目标就是构建一个数字仓储支撑的学术工作平台 (Scholars'

Workbench)^[90]。

在数字仓储的互操作方面,除了OAI-PMH和OAI-ORE之外,另一个值得关注的数字仓储研究点是Deposit API,目前正由英国的SWORD(Simple Web service Offering Repository Deposit)项目推动^[91]。

6 数字保存研究与建设思潮

Digital Curation、Digital Archival和Digital Preservation显然是21世纪数字图书馆技术研究和系统建设中居主导地位的学科。数字保存的研究发端于20世纪90年代末,风行于21世纪的00年代。在20世纪90年代末,开始出现了英国的Cedars(CURL Exemplars in Digital Archives)项目^[92]和欧洲的NEDLIB(Networked European Deposit LIBrary)项目^[93],两者都开始于1998年。

2002年美国空间数据系统咨询委员会CCSDS发布的Reference Model for an Open Archival Information System(OAIS)是数字保存领域的权威性文件,它提出的OAIS成为所有数字保存系统建设必须参考的标准^[94]。

数字图书馆界有影响的数字保存项目和系统基本上都出现在21世纪。如荷兰国家图书馆(Koninklijke Bibliotheek, KB)在2002年12月建成了e-Depot^[95]数字保存项目的核心系统:数字信息存档系统(Digital Information and Archiving System, DIAS)。美国的National Digital Information Infrastructure and Preservation Program(NDIIPP)^[96]项目于2000

年12月开始,目前已经形成了由130家机构参加的覆盖美国全国的数字保存网络。

目前支持多方参与的、成熟应用的系统除DIAS系统之外,还有LOCKSS和Portico。Stanford University Libraries的LOCKSS(Lots of Copies Keep Stuff Safe)系统^[97]在2000年推出beta版本,并于2000至2002年间在全世界的50多个图书馆中进行了测试,目前约有400多个出版社的2700种期刊在LOCKSS系统上进行了保存。2002年ITHAKA^[98]创建的Portico^[99]在2005、2006年期间得到了飞速发展,目前已经成为全球领先的数字保存服务机构,共有91家出版社、657个图书馆成为其成员。Portico已经保存了8,207种期刊,1,940本电子书,共计14,918,003个数字单元^[100]。

有些机构基于开源软件系统开发了其数字保存系统。如California Digital Library(CDL)于2005年推出了Digital Preservation Repository(DPR)^[101]。2007年12月Florida Center for Library Automation^[102]以开放源码的方式推出了DAITSS(Dark Archive in the Sunshine State)^[103]数字保存仓储系统,目前这一系统还在不断升级完善之中。

在这10年中,为了推动数字保存工作的开展,一些国家的研究机构和组织纷纷组建专门的数字保存研究和建设机构,构建数字保存联盟,定期召开一系列已然形成品牌的数字保存会议。这些数字保存的研究和交流活动,更促进了这十年数字保存的盛行。典型的例子有:英国在2001年构建了数字保存联盟DPC(The Digital Preservation Coalition)^[104]于2004年组建了数

字保存中心DCC(Digital Curation Centre)^[105]。欧盟在第6框架期间构建了Digital Preservation Europe(DPE)协作网^[106]。德国在2003年构建Nestor(Network of Expertise in Long-term STorage of Digital Resources)^[107]数字保存能力网络。2003年7月12个Web存档机构构建了国际互联网保存联盟(International Internet Preservation Consortium, IIPC)^[108]。在学术交流方面,目前与数字保存相关的会议有:国际数字对象保存会议iPRES(International Conference on Preservation of Digital Objects)^[109]、国际数字保存会议(International Digital Curation Conference)^[110]、DigCCurr会议^[111]。这些会议都形成了定期召开的惯例。

7 本体和语义网研究与建设思潮

知识组织是数字图书馆的一项重要内容。进入二十一世纪之后,以强调概念关联为核心的一些知识组织体系如词网(WordNet)、概念图、主题地图、本体等开始成为研究和应用的热点。这一时期图书馆传统的知识组织体系也在朝着这个方面前进。如2000年7月推出的杜威十进分类法网络版^[112],2001年推出的美国国会图书馆分类法(LCC)的网络版^[113],以及美国医学图书馆推出的2003年一体化医学语言系统(UMLS)^[114],都加强了概念的关联处理。

2001年5月,Web之父Tim Berners-Lee在《科学美国人》杂志上发表了“The Semantic Web”一文,正式提出了语义网的概念。语义Web的框架中最重要的两项

技术是RDF(S)^[115]和OWL^[116], W3C分别于2004年和2008年推出了SKOS标准和RDF(S)的查询语言SPARQL^[117],促进了语义网技术在数字图书馆知识组织中的应用^[118]。

一般认为Ontology是指一种“形式化的,对于共享概念体系的明确而又详细的说明”^[119]。从其具体形式来看,Ontology是一种对领域知识进行规范化的抽象与描述的工具,在构建过程中清楚和明确地描述了领域内的概念及其之间的各种关系,与分类表、术语表、叙词表等词表相比,Ontology支持更完整、更精确的领域模型,可以用来解决语义异构的问题,以不断提高数字图书馆信息整合的深度和效果,迈向知识组织和知识服务。

21世纪00年代,语义Web技术和基于Ontology的知识组织得到了数字图书馆界的特别关注。一些数字图书馆领域的重要会议如JCDL2006^[120]、JCDL2008^[121]等都将语义数字图书馆作为一个独立的主题进行讨论,希望利用语义Web相关技术来解决数字图书馆面临的各种问题。2007年国际数字图书馆与语义Web大会^[122]第一次召开,讨论数字图书馆与语义Web技术的结合问题。2009年该会议召开第二次会议,语义数字图书馆再次成为会议讨论的热点。与此同时,业界的研究成果和原型系统也不断出现。2005年英国爱尔兰企业数字研究中心开发出的JeromeDL系统是一个语义数字图书馆引擎,利用语言Web技术和社会化网络技术来提升数字图书馆的浏览和搜索能力^[123]。

近年来,有人将语义Web和内容挖掘分析为主的数字图书馆称为第二代数字图书馆^[124]。英国和新西兰合作开发的音乐分析网络化环

境(NEMA)^[125]、美国梅隆基金会提供赞助的一个数字化平台MONK项目^[126]、英国曼彻斯特大学和南安普敦大学合作的一个科研项目MyExperiment^[127]、欧洲委员会资助的NESTER项目^[128]、以及美国伊利诺伊大学的Meandre平台^[129]都被认为是具有第二代数字图书馆性质的新型数字图书馆。

目前被应用得较为广泛的、基于语义网络构建学术社区的系统是Vivo^[130],由美国康奈尔大学开发。

8 Web 2.0思潮

在上世纪90年代末期,Wiki、Blog等交互性协同技术其实已经出现,但直到2005年,在Tim O'Reilly^[131]提出Web2.0的概念之后,这些技术才成为大众化的应用。当前Web2.0已经成为一个强有力的社会网络(Social Networking)平台,大量的网络合作项目开始出现,例如:YouTube^[132],Flickr^[133],Wikipedia^[134]等等,它们深刻地影响着网络世界的发展,也对数字图书馆在00年代的发展带来强有力的冲击。

Web2.0带给数字图书馆研究和建设者的一个理念是:数字图书馆用户的身份已不仅仅是内容的消费者,而且也是内容的创建和组织者。数字图书馆界传统的知识组织方式也随之而改变,社会标注(Social Annotations)成为了数字内容组织和揭示的一种重要方式。通过允许用户对资源进行标注操作,不仅能够吸引更多用户,也提供了捕获分散信息片段的新途径。Digital Library for Earth System Education(DLESE)在2003年开发了OpenDLib系统^[135],使用

用户对数字图书馆馆藏的标注来制定相关资源内容的有效教育策略。Pennsylvania大学于2006年进行的PennTags project^[136]项目允许用户对社会标签(Social Bookmarking)、OPAC条目、Web Urls和期刊文章数据库条目进行标注。由Andrew W. Mellon基金于2009年资助的Open Annotation Collaboration(OAC)^[137,138]项目力图建立一个交互的标注环境,它能够使研究者使用多种标注客户端来分享他们的标注,不管对于完全开放的环境还是在实际的或虚拟的组织边界内。

借助Web2.0技术开展知识内容关联,实现对知识内容的深度组织和利用是数字图书馆越来越关注的另一个方面。Web的发明人Tim Berners-Lee于2006年提出了Linked Data概念,它提供了一种为Web所用的含义更丰富的联结方式,将网络中的超文本链接(文献到文献)推进到超数据的联结(内容资源的联结)层面^[139]。W3C于2007年启动了关联开放数据项目LOD(Linking Open Data),截止到2009年7月,LOD云图中已经有100多个开放数据集,包括4.7亿个RDF三元组,以及1.4亿个RDF内部链接。而且,关联数据正在迅速发展,包括MIT、HCLS、BBC、IEEE、国会图书馆、英美政府等大学、研究机构、媒体服务机构、传统出版机构、图书馆以及政府部门等纷纷加入到关联数据出版行列。

伴随着Web2.0和社会化网络的兴起,以不同数据源和Web服务的混合与集成为主要特点的Mashup的提出为数字图书馆服务集成提供了良好范例。Mashup始于Paul Rademacher在2005年所做的工作,他在Housingmap.com网站建

设中,创造性地将在线房屋租赁网站Craiglist.org的数据源依据位置信息标识于Google Maps地图中^[140],由此Mashup获得广泛的认知。众多的网站如Flickr.com、Amazon.com、Google Maps等纷纷以数据源(Feeds)、API、Web Services等方式提供可编程扩展,鼓励第三方开发者与用户以创新的方式重组这些已有资源,构建Mashup作品。数字图书馆界也积极尝试将Mashup应用于其服务集成。英国的萨塞克斯大学2007年开始的SPLASH(Student Personal Learning and Social Homepage)项目通过开发提供用户使用的Portal或mashup网络服务,提供一种更加个性化的学习体验。MIT图书馆和MIT人工智能实验室共同主持SIMILE项目对地图、时间序列、散列表、交互式图表中展示数据方面展开了研究。由瑞士苏黎世联邦理工学院的计算机科学学院的信息与通信系统小组(ICS)主持的JOpera项目提出把可视化语言应用于服务集成。

网络虚拟信息环境的兴起也为数字图书馆界提供了新的展示空间和服务渠道。林登实验室于2003年向公众开放了Second Life虚拟空间^[141]。用户是这个空间以及空间中生命的构思者、创造者。作为一种互联网2.0的尝试性应用,“Second Life”给网民和组织提供了一个能够自由发挥的网络平台。2006年,路透社在这个虚拟空间内开设了一个数字新闻分社^[142],成为现实中最新一家承认这个空间的重要性的机构,对这一虚拟空间带来了相当大的影响。还有许多机构出于同样的原因已经开始使用“Second Life”,包括美国联邦政府的国家海洋和大气管理局,BBC英国广播

公司等。

图书馆界利用“Second Life”虚拟空间的尝试也早已开始。目前,全球已有大约400多个图书馆在这个虚拟空间中建立了服务社区。2006年4月,联盟图书馆系统(Alliance Library System, ALS)与联机图书馆活动(Online Programming for All Libraries, OPAL)宣布,他们将一起合作在“Second Life”中建一个图书馆。11月,该图书馆首次亮相,名称为“Cybrary City”^[143]。2007年1月,著名图书馆自动化系统厂商SirsiDynix在2007美国图书馆协会仲冬会议上宣布将赞助联盟图书馆系统的“Second life”计划——“信息岛”及“青少年信息岛”^[144]。

Web2.0的许多思想也被借鉴到Library2.0的建设中。如JISC于2008年资助了“Towards implementation of Library 2.0 and the e-Framework”^[145]项目,目标是在当前信息环境下,解决如何有效集成和利用Web2.0技术构建Library 2.0的问题。

9 开放存取和开源思潮

20世纪90年代初期,由于出版商对期刊利润的最大化追逐,学术交流和知识传播的壁垒日趋严重,出现了“学术期刊危机”的局面。全球图书馆积极加入到了开放获取(Open Access, OA)的运动之中。第一个国际性的开放存取声明是2002年发布的“布达佩斯开放获取计划”^[146],随后两个重要文献是2003年6月的“Bethesda开放获取出版声明”^[147]和2003年10月的“开放获取科学和人文知识柏林

宣言”^[148]。

21世纪00年代,OA运动也是数字图书馆研究和建设的一个特色。在2004年至今的7次柏林大会上^[149],数字图书馆界都从OA指南、OA系统构建、OA在各个领域的实践、OA面临的法律问题等方面深化和实践OA运动。

“开源(Open Source)”是数字图书馆21世纪00年代的另一个“开放”思潮。早在1999年的ARL/OCLC的战略问题论坛上,80多名研究图书馆的领导人就提出了图书馆与开源软件的关系原则^[150]。随后在2000年,美国图书馆和信息技术协会LITA(美国图书馆协会ALA的分支机构)就成立了开放源码兴趣小组,推动开放源码软件在图书馆中的应用,2003年年初,其技术专家更是将开放源码软件作为当时图书馆技术领域的7大发展趋势之一^[151]。几乎与LITA同步,数字图书馆联盟DLF针对开放源码也陆续召开了相关专家会议及研讨会,从在图书馆领域开发和应用开放源码软件的行动纲领、采取的具体措施等方面展开了讨论^[152]。与前两者类似,图书馆电子信息协会eIFL也一致致力于推进开源软件在图书馆中的应用,如专门增加FOSS栏目、提供与数字图书馆开放源码软件相关的资源和信息、调查数字图书馆开源软件的应用情况、提出eIFL的开放源码软件建议等^[153,154],而旨在支持对开源集成图书馆系统进行评价和迁移的eIFL-FOSS ILS项目极大地促进了其成员馆对开源软件的利用^[155]。当然,在支持和鼓励开源软件在图书馆的应用方面,IFLA也表现得相当积极,在其多次的IFLA WLIC大会上,都有开放源码软件的报告和发言。这些报告不仅关注

欧美等国家的开源软件利用,也关注推动亚非等发展中国家图书馆中开源软件的应用^[156,157,158]。此外,D-lib^[159]、Oss4Lib^[160]和Code4Lib^[161]等也分别通过出版一系列数字图书馆开源软件相关的文章,设计和实现一系列专为图书馆应用的开源软件等来推动数字图书馆开源软件开发、共享和应用。

目前在数字图书馆的研究和建设中,已经有各种类型的开源软件可供利用,如检索领域中的Solr^[162]和Nutch^[163],仓储建设中的E-Prints、Fedora和DSpace,信息抽取中的GATE^[164]等。一些重要的数字图书馆项目,也基于开源软件系统构建,如美国国家科学数字图书馆(NSDL)系统基于Fedora框架^[165]、国际Internet保存联盟(IIPC)利用Heritrix、NutchWAX等构建其Web存档系统^[166]等。

近来,JISC融合了OA、Open

Resource、Open Source和Open Standards,进一步提出了“Open Technologies”的概念,并公布了其在Open Technologies方面资助的核心项目、计划等,旨在进一步鼓励和推动开放的公众可获取的能力^[167]。而由New America Foundation资助的Open Technology Initiative项目也希望通过探究Open Technologies在社会和经济方面的影响,增强Open Technologies在创新方面的推动作用^[168]。

10 结语

回顾二十一世纪前十年数字图书馆技术研究和系统建设的历程,充分印证了事物发展变化的规律是“螺旋式上升,波浪式前进”。

从现象上来看,数字图书馆的技术研究的确存在着随时间的发展变化、高低起伏的特点。当数字化

和数字资源建设被认为在二十世纪已经被解决了的时候,谁都不会想到在二十一世纪00年代的后期,全球又掀起了一个数字化和数字资源建设的新高潮。

从更本质的方面来讲,尽管有这么多的思潮,数字图书馆技术研究和系统建设所要解决的核心问题应当还是相对稳定的。知识的组织和集成可以从本体和语义网的角度来研究,而Web 2.0的社会化标签则是解决问题的另一个方向;这十年的前期和中期,人们提出了基于网格构建下一代数字图书馆,而这十年后期,人们更偏向于利用云计算来构建数字图书馆的技术体系。解决思路可能不同,但要解决的问题还是相对统一的。

由此可以看到,这十年中涌现的9个思潮也仅仅是解决数字图书馆核心技术问题的历史长河中的一个片段。

参考文献

- [1] SMITH A. Strategies for Building Digitized Collections. D-Lib Magazine, 2001, 9[EB/OL]. [2010-03-05]. <http://www.diglib.org/pubs/dl094/dl094.htm>.
- [2] FLECKER D. Harvard's Library Digital Initiative: Building a First Generation Digital Library Infrastructure. D-Lib Magazine, 2000, 11[EB/OL]. [2010-03-05]. <http://www.dlib.org/dlib/november00/flecker/11flecker.html>.
- [3] American Memory[EB/OL]. [2010-03-05]. <http://memory.loc.gov/ammem/about/index.html>.
- [4] UMDL[EB/OL]. [2010-03-05]. <http://www.si.umich.edu/UMDL/intro.html>.
- [5] JSTOR[EB/OL]. [2010-03-05]. <http://en.wikipedia.org/wiki/JSTOR>.
- [6] The Million Book Project[EB/OL]. [2010-03-05]. http://en.wikipedia.org/wiki/Million_Book_Project.
- [7] The Universal Library[EB/OL]. [2010-03-05]. <http://www.ulib.org/>.
- [8] The European Library[EB/OL]. [2010-03-05]. <http://search.theeuropeanlibrary.org/portal/en/index.html>.
- [9] Global Memory Net[EB/OL]. [2010-03-05]. <http://www.memorynet.org/about/>.
- [10] FLECKER D. Harvard's Library Digital Initiative: Building a First Generation Digital Library Infrastructure. D-Lib Magazine, 2000, 6[EB/OL]. [2010-03-05]. <http://www.dlib.org/dlib/november00/flecker/11flecker.html>.
- [11] LDI Overview[EB/OL]. [2010-03-05]. <http://hul.harvard.edu/ois/ldi/>.
- [12] The European Library[EB/OL]. [2010-03-05]. http://www.theeuropeanlibrary.org/portal/organisation/about_us/aboutus_en.htm.
- [13] Europeana[EB/OL]. [2010-03-05]. <http://www.europeana.eu/portal/aboutus.html>.
- [14] i2010 - A European Information Society for growth and employment[EB/OL]. [2010-03-05]. http://ec.europa.eu/information_society/eeurope/i2010/index_en.htm.
- [15] Live_Search_Books[EB/OL]. [2010-03-05]. http://en.wikipedia.org/wiki/Live_Search_Books.
- [16] Internet Archive[EB/OL]. [2010-03-05]. <http://www.archive.org/>.
- [17] UNESCO[EB/OL]. [2010-03-05]. http://portal.unesco.org/ci/en/ev.php-URL_ID=28549&URL_DO=DO_TOPIC&URL_SECTION=201.html.
- [18] World Digital Library[EB/OL]. [2010-03-05]. <http://www.wdl.org/en/about/background.html>.
- [19] Resource Discovery Network for Further Education[EB/OL]. [2010-03-05]. <http://www.jisc.ac.uk/whatwedo/programmes/x41/rdnfe.aspx>.
- [20] Subject Portal[EB/OL]. [2010-03-05]. <http://www.portal.ac.uk/spp/>.
- [21] Intute[EB/OL]. [2010-03-05]. <http://www.intute.ac.uk/history.html>.
- [22] Renardus[EB/OL]. [2010-03-05]. <http://www.ukoln.ac.uk/metadata/renardus/>.
- [23] HEERY R, CARPENTER L, DAY M. Renardus Project Developments and the Wider Digital Library Context. D-Lib Magazine, 2001, 4[EB/OL]. [2010-03-05]. <http://www.dlib.org/dlib/april01/heery/04heery.html>.

- [24] Scholars Portal[EB/OL].[2010-03-05].<http://www.arl.org/resources/pubs/portals/scholarsportal.shtml>.
- [25] vascoda-portal[EB/OL].[2010-03-05].<http://www.vascoda.de/>.
- [26] The Scout Portal Toolkit (SPT)[EB/OL].[2010-03-05].<http://scout.wisc.edu/Projects/SPT/index.php>.
- [27] ALMASY E,SLEASMAN D,BOWER R.Software for Building a Full-Featured Discipline-Based Web Portal. D-Lib Magazine,2002,11[EB/OL].[2010-03-05].
<http://www.dlib.org/dlib/november02/almasy/11almasy.html>.
- [28] iVia[EB/OL].[2010-03-05].<http://ivia.ucr.edu/>.
- [29] MITCHELL S,MOONEY M,MASON J,etc. iVia Open Source Virtual Library System. D-Lib Magazine,2003,1[EB/OL].[2010-03-05]. <http://www.dlib.org/dlib/january03/mitchell/01mitchell.html>.
- [30] MyLibrary Personalized Electronic Services in the Cornell University Library. D-Lib Magazine,2000,4[EB/OL].[2010-03-05].<http://www.dlib.org/dlib/april00/mistlebauer/04mistlebauer.html>.
- [31] MORGAN K. Source Code Released for MyLibrary@NCState. D-Lib Magazine,2000,4[EB/OL].[2010-03-05]. <http://www.dlib.org/dlib/april00/04inbrief.html#MORGAN>.
- [32] JISC Information Environment Architecture[EB/OL].[2010-03-05]. <http://www.ukoln.ac.uk/distributed-systems/jisc-ie/arch/faq/portal/>.
- [33] Portals Programme[EB/OL].[2010-03-05]. <http://www.jisc.ac.uk/whatwedo/programmes/portals.aspx>.
- [34] uPortal[EB/OL].[2010-03-05].<http://www.jasig.org/uportal>.
- [35] Liferay[EB/OL].[2010-03-05].<http://www.liferay.com/>.
- [36] Shibboleth®[EB/OL].[2010-03-05].<http://shibboleth.internet2.edu/>.
- [37] VAN DE SOMPEL H.The Santa Fe Convention of the Open Archives Initiative. D-Lib Magazine,2000,2[EB/OL].[2010-03-05].<http://www.dlib.org/dlib/february00/vandesompeel-oai/02vandesompeel-oai.html>.
- [38] The Open Archives Initiative Protocol for Metadata Harvesting[EB/OL].[2010-03-05].<http://www.openarchives.org/OAI/1.0/openarchivesprotocol.htm>.
- [39] CrossRef[EB/OL].[2010-03-05].<http://www.crossref.org/01company/02history.html>.
- [40] VAN DE SOMPEL H.Open Linking in the Scholarly Information Environment Using the OpenURL Framework. D-Lib Magazine,2001,3[EB/OL].[2010-03-05].
<http://www.dlib.org/dlib/march01/vandesompeel/03vandesompeel.html>.
- [41] ANSI/NISO Z39.88 - The OpenURL Framework for Context-Sensitive Services[EB/OL].[2010-03-05].http://www.niso.org/kst/reports/standards?step=2&project_key=d5320409c5160be4697dc046613f71b9a773cd9e.
- [42] CrossRef[EB/OL].[2010-03-05].<http://www.crossref.org/01company/02history.html>.
- [43] DENENBERG R.Search Web Services - The OASIS SWS Technical Committee Work[EB/OL].[2010-03-05].<http://www.dlib.org/dlib/january09/denenberg/01denenberg.html>.
- [44] SRU / CQL / ZeeRex[EB/OL].[2010-03-05]. <http://www.loc.gov/standards/sru/>.
- [45] NISO Networked Reference[EB/OL].[2010-03-05]. <http://www.loc.gov/standards/netref/>.
- [46] Open Archives Initiative,Object Reuse and Exchange[EB/OL].[2010-03-05]. <http://www.openarchives.org/ore/>.
- [47] Digital Library Standards and Practices[EB/OL].[2010-03-05]. <http://www.diglib.org/standards.html#endorsed>.
- [48] METs[EB/OL].[2010-03-05]. <http://www.loc.gov/standards/mets/>.
- [49] Introduction and Supporting Materials from PREMIS Data Dictionary for Preservation Metadata[EB/OL].[2010-03-05]. <http://www.loc.gov/standards/premis/v2/premis-report-2-0.pdf>.
- [50] OCKHAM Initiative[EB/OL].[2010-03-05].<http://www.ockham.org/>.
- [51]MORGAN E L,FRUMKIN J,FOX E A.The OCKHAM Initiative - Building Component-based Digital Library Services and Collections. D-Lib Magazine,2004,11[EB/OL].[2010-03-05].<http://www.dlib.org/dlib/november04/11inbrief.html>.
- [52] XIANG X,MORGAN E L.Exploiting "Light-weight" Protocols and Open Source Tools to Implement Digital Library Collections and Services.D-Lib Magazine,2005,10[EB/OL].[2010-03-05]. <http://www.dlib.org/dlib/october05/morgan/10morgan.html>.
- [53] OCKHAM: A framework and forum for evolving digital library architectures[EB/OL].[2010-03-05].<http://www.diglib.org/architectures/ockham.htm>.
- [54] DELOS[EB/OL].[2010-03-05]. http://www.delos.info/index.php?option=com_content&task=view&id=345.
- [55] ECDL 2002 Panels[EB/OL].[2010-03-05]. <http://ecd2002.iei.pi.cnr.it/program/panels.html>.
- [56] Approaches for Large Scale Digital Library Infrastructures by Example, Tutorial at ECDL2007.(Abstract)[EB/OL].[2010-03-05].<http://www.ecdl2007.org/tutorials.php>.
- [57] RISSE T.Approaches for Large Scale Digital Library, Tutorial at RCDL 2007[EB/OL].[2010-03-05]. http://www.l3s.de/~risse/RCDL2007_tutorial.pdf.
- [58] Building Digital Libraries On-demand by Sharing Content, Services and Computing Resource, Tutorial at ECDL2007(Abstract)[EB/OL].[2010-03-05].<http://www.ecdl2007.org/tutorials.php>.
- [59] SCHEK H-J,SCHULDT H. DelosDLMS - Infrastructure for the Next Generation of Digital Library Management Systems. ERCIM News ,2006,6[EB/OL].[2010-03-05]. http://www.ercim.eu/publication/Ercim_News/enw66/schek.html.
- [60] The OSIRIS Process Support Middleware and the ISIS Process-Based Digital Library Application[EB/OL].[2010-03-05]. http://dbis.cs.unibas.ch/delos_website/jpa2/osiris.pdf.
- [61] The DELOS Architecture Work Package[EB/OL].[2010-03-05]. http://dbis.cs.unibas.ch/delos_website/.
- [62] OpenDLib[EB/OL].[2010-03-05].<http://www.opendlib.com/home.html>.
- [63] CASTELLI D,PAGANO P,THANOS C. OpenDLib: an infrastructure for new generation digital libraries, International Journal on Digital Libraries,2004,8[EB/OL].[2010-03-05].<http://springerlink.metapress.com/content/gdytpq8q3br0ubtu/?p=4e668c9082fb4c5d8a94bbadc648f031&pi=2>.
- [64] CANDELA L,CASTELLI D,PAGANO P,SIMI M. OpenDLibG: Extending OpenDLib by Exploiting a gLite Grid Infrastructure .Research and Advanced Technology for Digital Libraries,2006,1-13[EB/OL].[2010-03-05].<http://springerlink.metapress.com/content/lt147r211w68074/?p=4e668c9082fb4c5d8a94bbadc648f031&pi=0>.
- [65] HASLHOFFER B,KNEŽEVIĆ P. The BRICKS Digital Library Infrastructure.Semantic Digital Libraries, 2009[EB/OL].[2010-03-05].<http://www.springerlink.com/content/uw35466419356165/>.
- [66] BRICKS: A Digital Library Management System for Cultural Heritage[EB/OL].[2010-03-05].http://www.ercim.eu/publication/Ercim_News/enw61/meghini.html.
- [67] DILIGENT[EB/OL].[2010-03-05].<http://diligent.ercim.eu/>.
- [68] CASTELLI D.DILIGENT: A Digital Library Infrastructure on Grid Enabled Technology . ERCIM News ,2004,7[EB/OL].[2010-03-05].http://www.ercim.eu/publication/Ercim_News/enw59/castelli.html.
- [69] gCube Framework[EB/OL].[2010-03-05]. <http://www.gcube-system.org/>.
- [70] MORRIS C M.Cloud Computing, Big Data, and Open Access at EDUCAUSE 2009 ,D-Lib Magazine.2010,1[EB/OL].[2010-03-05]. <http://www.dlib.org/dlib/january10/morris/01morris.html>.

- [71] Apache Hadoop[EB/OL].[2010-03-05]. <http://hadoop.apache.org/>.
- [72] arXiv.org[EB/OL].[2010-03-05]. <http://arxiv.org/>.
- [73] CogPrints[EB/OL].[2010-03-05]. <http://cogprints.org/>.
- [74] NDLTD[EB/OL].[2010-03-05]. <http://www.ndltd.org/>.
- [75] Eprints[EB/OL].[2010-03-05]. <http://www.eprints.org/>.
- [76] Fedora[EB/OL].[2010-03-05]. <http://fedora-commons.org/>.
- [77] Dspace[EB/OL].[2010-03-05]. <http://dspace.org/>.
- [78] eScholarship[EB/OL].[2010-03-05]. <http://escholarship.org/>.
- [79] Digital Commons[EB/OL].[2010-03-05]. <http://digitalcommons.ilr.cornell.edu/>.
- [80] ROAR[EB/OL].[2010-03-05]. <http://roar.eprints.org/information.html>.
- [81] OpenDOAR[EB/OL].[2010-03-05]. <http://www.openoar.org/>.
- [82] Australian Research Repositories Online to the World[EB/OL].[2010-03-05]. <http://www.arrow.edu.au/about/>.
- [83] National Academic Research and Collaborations Information System[EB/OL].[2010-03-05]. <http://www.narcis.info/>.
- [84] DRIVER[EB/OL].[2010-03-05]. <http://www.driver-repository.eu/>.
- [85] OCLC[EB/OL].[2010-03-05]. <http://www.oclc.org/oaister/>.
- [86] eSciDoc[EB/OL].[2010-03-05]. <https://www.escidoc.org/>.
- [87] R4L[EB/OL].[2010-03-05]. <http://www.jisc.ac.uk/whatwedo/programmes/digitalrepositories2005/r4l.aspx>.
- [88] X4L[EB/OL].[2010-03-05]. <http://www.jisc.ac.uk/whatwedo/programmes/x4l.aspx>.
- [89] JORUM[EB/OL].[2010-03-05]. <http://www.jorum.ac.uk/>.
- [90] Towards a Repository-enabled Scholar's Workbench: RepoMMan, REMAP and Hydra[EB/OL].[2010-03-05]. <http://www.dlib.org/dlib/may09/green/05green.html>.
- [91] SWORD[EB/OL].[2010-03-05]. http://www.ukoln.ac.uk/repositories/digirep/index/SWORD_Project.
- [92] Cedars(Curl Exemplars in Digital Archives)[EB/OL].[2010-03-05]. <http://www.webarchive.org.uk/ukwa/target/99695/>.
- [93] NEDLIB(Networked European Deposit Library)[EB/OL].[2010-03-05]. <http://nedlib.kb.nl/>.
- [94] CCSDS (Consultative Committee for Space Data Systems) (2002). Reference Model for an Open Archival Information System (OAIS). Blue Book, Issue 1. Washington, DC (US): CCSDS Secretariat, January 2002. Technical report. CCSDS 650.0-B-1. Recommendation for Space Data System Standards[EB/OL].[2010-03-05]. <http://public.ccsds.org/publications/archive/650x0b1.pdf>.
- [95] e-Depot[EB/OL].[2010-03-05]. <http://www.kb.nl/hrd/dd/index-en.html>.
- [96] NDIIPP[EB/OL].[2010-03-05]. <http://www.digitalpreservation.gov/>.
- [97] LOCKSS[EB/OL].[2010-03-05]. <http://lockss.stanford.edu/lockss/Home>.
- [98] ITHAKA[EB/OL].[2010-03-05]. <http://www.ithaka.org/>.
- [99] Portico[EB/OL].[2010-03-05]. <http://www.portico.org/>.
- [100] Portico Archive Facts&Figures[EB/OL].[2010-03-05]. <http://www.portico.org/digital-preservation/the-archive-content-access/archive-facts-figures/#page-1554>.
- [101] Digital Preservation Repository (DPR)[EB/OL].[2010-03-05]. <http://www.cdlib.org/services/uc3/dpr.html>.
- [102] FCLA[EB/OL].[2010-03-05]. <http://www.fcla.edu/digitalArchive/>.
- [103] DAITSS[EB/OL].[2010-03-05]. <http://daitss.fcla.edu/>.
- [104] DPC[EB/OL].[2010-03-05]. <http://www.dpconline.org/>.
- [105] DCC[EB/OL].[2010-03-05]. <http://www.dcc.ac.uk/>.
- [106] DPE[EB/OL].[2010-03-05]. <http://digitalpreservationeurope.eu/>.
- [107] Nestor[EB/OL].[2010-03-05]. <http://www.langzeitarchivierung.de/eng/index.htm>.
- [108] IIPC[EB/OL].[2010-03-05]. <http://www.netpreserve.org/about/index.php>.
- [109] iPRES[EB/OL].[2010-03-05]. <http://rdd.sub.uni-goettingen.de/conferences/ipres/ipres-en.html>.
- [110] International Digital Curation Conference[EB/OL].[2010-03-05]. <http://www.dcc.ac.uk/events/conferences>.
- [111] DigCCurr[EB/OL].[2010-03-05]. <http://ils.unc.edu/digccurr/index.html>.
- [112] WebDewey[EB/OL].[2010-03-06]. <http://www.oclc.org/dewey/overview/default.htm>.
- [113] Library of Congress Classification Online[EB/OL].[2010-03-06]. <http://www.loc.gov/catdir/cps/lcco/>.
- [114] UMLS[EB/OL].[2010-03-06]. <http://www.nlm.nih.gov/research/umls/>.
- [115] FRANK MANOLA,ERIC MILLER. RDF Primer[EB/OL]. (2004-02-10)[2010-03-06]. <http://www.w3.org/TR/2004/REC-rdf-primer-20040210/>.
- [116] FRANK VAN HARMELEN,DEBORAH L.MCGUINNESS. OWL Web Ontology Language Overview[EB/OL]. (2004-02-10)[2010-03-06]. <http://www.w3.org/TR/2004/REC-owl-features-20040210/>.
- [117] SPARQL[EB/OL].[2010-03-06]. <http://www.w3.org/TR/rdf-sparql-query/>.
- [118] SKOS[EB/OL].[2010-03-06]. <http://www.w3.org/2006/07/SWD/SKOS/reference/20090315/implementation.html>.
- [119] GRUBER T. "A translation approach to portable ontology specifications". Knowledge Acquisition[EB/OL].[2010-03-06]. <http://tomgruber.org/writing/ontolingua-kaj-1993.pdf>.
- [120] JCDL2006[EB/OL].[2010-03-06]. <http://www.jcdl.org/archived-conf-sites/jcdl2006/>.
- [121] JCDL2008[EB/OL].[2010-03-06]. <http://www.jcdl2008.org/tutorials/tutorial2.html>.
- [122] ICSD[EB/OL].[2010-03-06]. <http://www.icsd-conference.org/>.
- [123] JeromeDL[EB/OL].[2010-03-06]. <http://www.jeromedl.org/>.
- [124] Workshop On Integrating Digital Library Content with Computational Tools and Services, ACM/IEEE Joint Conference on Digital Libraries (JCDL 2009), [C].[2010-03-06]. <http://www.music-ir.org/sgdl-workshop/sgdl-cfp.pdf>.
- [125] NEMA[EB/OL].[2010-03-06]. <http://nema.lis.uiuc.edu>.
- [126] MONK[EB/OL].[2010-03-06]. <http://monkproject.org>.
- [127] MyExperiment[EB/OL].[2010-03-06]. <http://myexperiment.org>.
- [128] NESTER[EB/OL].[2010-03-06]. <http://www.nester-ru.eu/>.
- [129] Meandre[EB/OL].[2010-03-06]. <http://seasr.org/meandre>.
- [130] VIVO[EB/OL].[2010-03-06]. [http://en.wikipedia.org/wiki/VIVO_\(Network\)](http://en.wikipedia.org/wiki/VIVO_(Network)).
- [131] O'REILLY, T. what is Web 2.0 Design Patterns and Business Models for the Next Generation of Software[EB/OL]. (2007)[2010-03-06]. <http://www.oreillynet.com/pub/a/oreilly/tim/news/2005/09/30/what-is-web-20.html>.

- [132] YouTube[EB/OL].[2010-03-06]. <http://www.youtube.com>.
- [133] Flickr[EB/OL].[2010-03-06]. http://www.flickr.com/learn_more.gne.
- [134] Wikipedia[EB/OL].[2010-03-06]. <http://www.wikipedia.org/>.
- [135] AGOSTI M, FERRO N. Annotations: Enriching a Digital Library. Research and Advanced Technology for Digital Libraries. Proceedings of the European Conference on Digital Libraries (ECDL 2003)[EB/OL].[2010-03-06]. <http://www.springerlink.com/content/66xwbxtw78yww4aq/>.
- [136] DAY A. Using Social Bookmarks in an Academic Setting – PennTags[EB/OL]. (2006)[2010-03-06]. <http://units.sla.org/division/dche/2006/day.pdf>.
- [137] Coalition for Networked Information- An Introduction&Program Plan 2009-2010[EB/OL]. (2009-2010). [2010-03-06]. <http://www.cni.org/program/2009-2010/2009-2010-program.pdf>.
- [138] Open Annotation Collaboration[EB/OL].[2010-03-06]. <http://www.cni.org/tfms/2009a.spring/abstracts/presentations/cni-open-cole.pdf>.
- [139] BURNERS-LEE T. Linked Data[EB/OL].[2010-03-06]. <http://www.w3.org/DesignIssues/LinkedData.html>.
- [140] Housingmap[EB/OL].[2010-03-06]. <http://www.housingmaps.com/>.
- [141] Secondlife[EB/OL].[2010-03-06]. <http://secondlife.com/>.
- [142] Hidigg[EB/OL].[2010-03-06]. <http://www.hidigg.com/game/onlinegames/200806/26-13093.html>.
- [143] Cybrary City[EB/OL].[2010-03-06]. <http://jintan.wordpress.com/2006/11/12/cybrary-city-in-secondlife/>.
- [144] Sirsi Dynix Sponsors Alliance Library Second Life Library and Teen second Life library[EB/OL].[2010-03-06]. <http://infoisland.org/2007/01/20/sirsi-dynix-sponsors-alliance-library-second-life-library-and-teen-second-life-library/>.
- [145] Towards implementation of Library 2.0 and the e-Framework[EB/OL].[2010-03-06]. <http://www.jisc.ac.uk/whatwedo/programmes/resourcediscovery/tile.aspx>.
- [146] Budapest Open Access Initiative[EB/OL].[2010-03-06]. <http://www.soros.org/openaccess/read.shtml>.
- [147] Bethesda Statement on Open Access Publishing[EB/OL].[2010-03-06]. <http://www.earlham.edu/~peters/fos/bethesda.htm>.
- [148] Berlin Declaration on Open Access to Knowledge in the Sciences and Humanities[EB/OL].[2010-03-06]. <http://oa.mpg.de/openaccess-berlin/berlindeclaration.html>.
- [149] Berlin7[EB/OL].[2010-03-06]. <http://www.berlin7.org/>.
- [150] The Keystone Principles. ARL Bimonthly Report, 1999(207)[EB/OL].[2010-03-06]. <http://www.arl.org/newsltr/207/keystone.html>.
- [151] ALA .THE TOP TRENDS[EB/OL].(2003)[2010-03-06]. <http://www.ala.org/ala/lita/litaresources/toptechrends/midwinter2003.htm>.
- [152] DLF Forums[EB/OL].[2010-03-06]. <http://www.diglib.org/forums/fall2008/>.
- [153] eIFL.net - eIFL - Free and Open Source Software[EB/OL].[2010-03-06]. <http://www.eifl.net/cps/sections/services/eifl-foss>.
- [154] eIFL.net—FOSS for Libraries[EB/OL].[2010-03-06]. <http://www.eifl.net/cps/sections/services/eifl-foss/foss-for-libraries>.
- [155] eIFL-FOSS ILS[EB/OL].[2010-03-06]. <http://www.eifl.net/cps/sections/services/eifl-foss/ils>.
- [156] AFIFI M. Shibboleth: An Open-Source Solution To Resource Sharing, World Library and Information Congress: 70th IFLA General Conference and Council, 22-27 August 2004, Buenos Aires, Argentina. [C/OL] [2010-03-06]. <http://www.ifla.org/IV/ifla70/papers/055e-Afifi.pdf>.
- [157] Information Technology Section Conferences [C/OL].Oslo, Norway 2005, [2010-03-06]. <http://www.ifla.org/VII/s21/21/Conferences.htm>.
- [158] Open Source[EB/OL].[2010-03-06]. <http://www.ifla.org/en/taxonomy/term/161>.
- [159] D-Lib Magazine, 2003,1[EB/OL].[2010-03-06]. <http://www.dlib.org/dlib/january03/01contents.html>.
- [160] oss4lib[EB/OL].[2010-03-06]. <http://www.oss4lib.org/>.
- [161] code4lib[EB/OL].[2010-03-06]. <http://www.code4lib.org/>.
- [162] Solr[EB/OL].[2010-03-06]. <http://lucene.apache.org/solr/>.
- [163] Nutch[EB/OL].[2010-03-06]. <http://lucene.apache.org/nutch/>.
- [164] GATE[EB/OL].[2010-03-06]. <http://gate.ac.uk/>.
- [165] Carl Lagoze. NSDL - A national digital library using RDF and Fedora Web Services[EB/OL].[2010-03-06]. <http://seminar.deff.dk/presentations/lagoze.ppt>.
- [166] International internet preservation consortium[EB/OL].[2010-03-06]. <http://netpreserve.org/>.
- [167] Open Technologies[EB/OL].[2010-03-06]. <http://www.jisc.ac.uk/whatwedo/topics/opentechnologies.aspx>.
- [168] Open Technologies Initial[EB/OL].[2010-03-06]. <http://opentechinstitute.org/>.

作者简介

张智雄 (1971-), 男, 中国科学院国家科学图书馆, 研究馆员, 博士生导师, 发文90余篇。通讯地址: 北京市海淀区中关村北四环西路33号, 中国科学院国家科学图书馆 100190。E-mail: zhangzhx@mail.las.ac.cn

洪娜 (1980-), 女, 中国科学院国家科学图书馆, 在读博士研究生。通讯地址同上

许德山 (1979-), 男, 中国科学院国家科学图书馆, 在读博士研究生, 通讯地址同上

刘建华 (1984-), 女, 供职于中国科学院国家科学图书馆, 通讯地址同上

徐健 (1977-), 男, 中国科学院国家科学图书馆, 在读博士研究生, 通讯地址同上

Nine Waves of Technology Research and System Development in Digital Library at the First Decade of 21st Century

Zhang Zhixiong, Liu Jianhua / National Science Library, Chinese Academy of Sciences, Beijing, 100190

Hong Na, Xu Deshan, Xu Jian / National Science Library, Chinese Academy of Sciences, Beijing, 100190

/ Graduate School of the Chinese Academy of Sciences, Beijing 100049

Abstract: In this paper, the authors analyze the technology research and systems development in digital library in the first decade of 21 century, and conclude there are 9 waves of research and development in digital library study during the past ten years, which include research on digitalization and digital resource, research on portal, research on integration and interoperation, research on “The Next Generation Digital Library”, research on digital repository, research on digital preservation, research on ontology and semantic web, research on Web 2.0, research on open access and open source.

Keywords: Digital library, Technical research, System construction, Development retrospect

(收稿日期: 2010-02-28)