

基于 ECDL 的数字图书馆技术结构及其演化研究 *

钱 力 张晓林 张智雄

摘 要 运用 Carrot²开源工具对 ECDL/TPDL 会议在 2007—2011 五年间发表论文的主题进行聚类分析,绘制出数字图书馆技术主题分布结构,并对由此反映出的数字图书馆发展趋势进行分析。结果显示:数字图书馆的发展模式与时俱进,逐步成为数字社会的表现形式;数字资源长期保存的内容得以丰富,范式发生变化;多领域协作的虚拟工作环境正在形成;知识组织体系不断完善,知识服务能力得以增强;用户体验的满意度和用户认可不断提高。通过对以上发展趋势的分析,揭示出可能产生重大影响的新技术与新应用,为我国数字图书馆的建设提供借鉴。图 1。表 3。参考文献 51。

关键词 数字图书馆 技术演化 ECDL 协同工作环境 数字资源长期保存 知识服务 用户行为

分类号 G250.76

Research of Structure and Evolution of Digital Library Technology Based on ECDL

Qian Li ,Zhang Xiaolin & Zhang Zhixiong

ABSTRACT By using Carrot² (an Open Source Search Results Clustering Engine) to conduct cluster analysis of subject of the papers in ECDL/TPDL Conferences from 2007 to 2011, a distribution structure of technical subjects in digital library is produced and an analysis of the growing trends of digital library is executed. The results show that with the development pattern of digital library advancing with the times, the digital library gradually becomes a manifestation of digital society; the content and the paradigm of long-term storage of digital resources have been enriched and changed; a multidisciplinary collaborative and virtual working setting is taking shape; knowledge organization system is constantly improving and knowledge service capability has been enhanced; satisfactory user experience and user acceptance have been increased. By analyzing above-mentioned growing trends, it discloses new technologies and applications that may exert major influences, and thus provides reference for digital library construction in China. 1 fig. 3 tabs. 51 refs.

KEY WORDS Digital library. Technology evolution. ECDL. Collaborative work environment. Digital resource preservation. Knowledge service. User behavior.

1 前言

欧洲数字图书馆会议 (European Conference on Digital Libraries, ECDL)^①是与数字图书馆联合会

(Joint Conference on Digital Library, JCDL)齐名的数字图书馆领域的国际权威学术会议。会议作者来自世界各国,尤其是在数字图书馆研究与应用居于前列的欧美国家,全面交流和研讨数字图书馆领域的前沿问题和先进实践。通过对 ECDL/TPDL 会

* 本研究系国家自然科学基金项目“网络科技信息中爆发主题的监测与分析方法研究”(项目编号:09BTQ035)的研究成果之一。

通讯作者:钱力,Email: qianl@mail. las. ac. cn

① ECDL 从 2011 年第 15 届更名为数字图书馆理论与实践会议(Theory and Practice of Digital Library, TPDL)

议 2007—2011 五年间发表论文的主题进行系统分析,有助于把握国际数字图书馆领域的研究重点和发展趋势,揭示可能产生重大影响的高新技术新应用。作者利用 Carrot²①开源工具分析聚类了 ECDL/TPDL 五年间的主题,并就其中的数字图书馆发展模式、数字资源长期保存、协同工作环境研究、知识组织体系建设和用户行为研究等五个重点方面的重要技术及实例进行了分析,希望为我国数字图书馆建设提供借鉴。

2 ECDL/TPDL 2007—2011 关注主题的分布结构

作者下载了 ECDL/TPDL 从 2007 至 2011 的五次会议全部论文,利用 Carrot²工具对其主题词进行聚类分析,绘制出这五年的主题分布结构图(见图 1)。

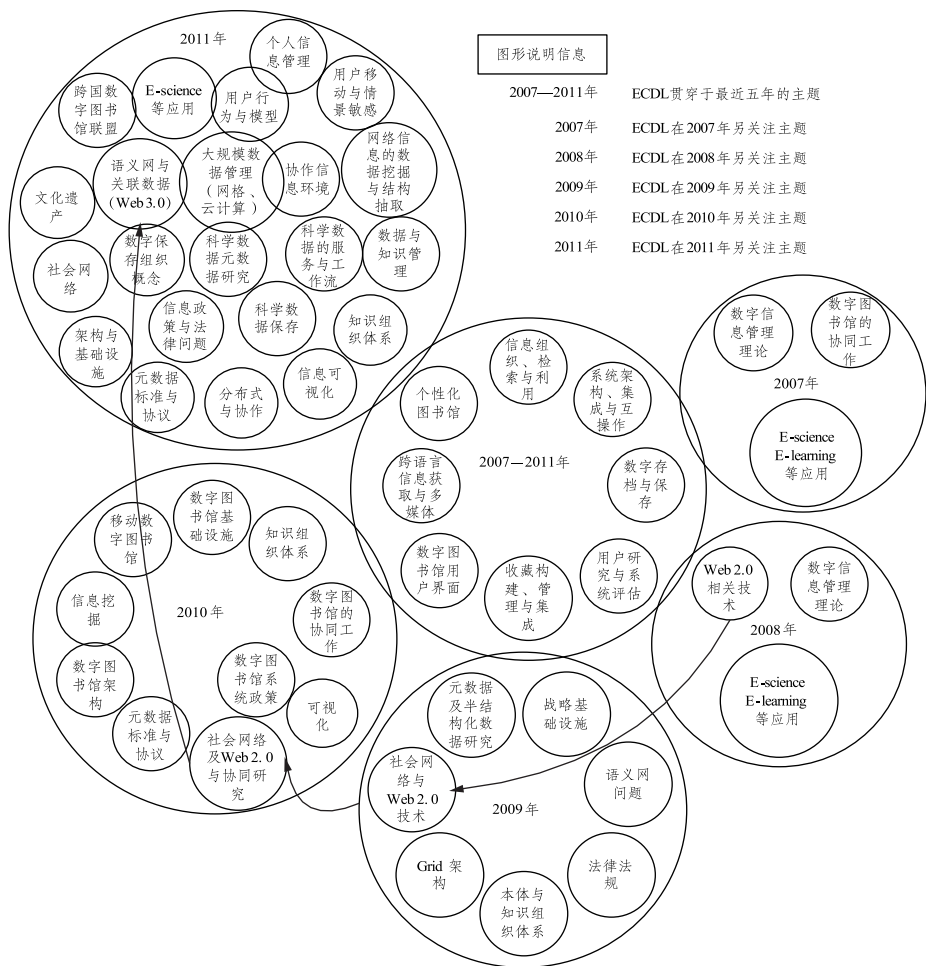


图 1 ECDL 2007—2011 年关注的主题分布结构及变化

① <http://search.carrot2.org/stable/search>

2.1 关注研究主题逐年增多、关注点逐年丰富

图 1 用中间圆列出了 2007—2011 年间的共同主题,并用其他圆分别列出了各年的专门关注主题。由图 1 可知,每年的研究主题都非常丰富,若干主题持续出现,研究主题逐年增多,分布结构图的范围越来越大。特别是 2011 年,出现从 Web 2.0 到 Web 3.0 的跨越,底层技术不断深入,中间资源层面则向研究数据和数字人文扩展,应用层面出现了愈加细化的应用与用户研究,形成了数字图书馆发展的新场景。

2.2 基本数字图书馆服务技术走向成熟

以图 1 中箭头方向示意可以发现,数字资源建设、系统构架、信息组织与检索、数字资源长期保存、用户界面研究、个性化图书馆、跨语言和多媒体信息检索、系统评估等成为贯穿数字图书馆发展的主题,并逐步走向成熟。知识本体、语义网、Web 2.0 等内容逐步出现,科学数据管理、关联数据、可视化、数字人文等成为新热点。协同工作环境主题虽然出现较早,但在后几年才开始具体化,并逐步深化,嵌入到用户的协同工作之中。

3 数字图书馆建设与发展模式

数字图书馆的概念出现很早,随着需求的变化和应用的发展,人们对“数字图书馆”模式与架构的认识不断深化。这方面的研究主题从最初的系统架构设计、集成技术应用,到后来的数字图书馆战略基础设施建设、协同工作环境以及用户应用环境研究。这既显现了这一主题的根基作用,又展示了数字图书馆概念与时俱进,逐步将数字图书馆作为数字社会和数字化工作生活的重要表现形式,并提出下一代泛在交流的数字图书馆范式等新概念。

3.1 数字图书馆架构与技术

数字图书馆架构研究逐步从系统平台视角扩展到数字图书馆发展整体架构,“以用户为中心”,引入新理念、新技术、新方法和新的服务模式,增强用户体验,例如 Web 2.0^[1] 相关技术的应用和 Web 3.0 技术的产生。表 1 描述了数字图书馆架构设计理念和应用模式的变化。

表 1 数字图书馆架构设计发展典型实例

年份	系统架构	具体描述
2007	The Bleek and Lloyd Collection	面向整个社会都在生产知识内容的发展趋势,未来每个角落都可能是可利用的数字资源,而实际形态的数据库很可能消失。Hussein suleman ^[2] 提出这样一个理念:“没有数据库的数字图书馆”,即通过 XML + XSLT + XHTML 技术构建一个基于终端的数字图书馆服务,目前已经应用在文化遗产数字资产的获取与保存方面。
2008	Gcube	“可扩展的虚拟数字图书馆生成器” ^[3] 架构的提出,以面向服务的插件机制为设计模式,通过组配方式搭建一个虚拟数字图书馆环境。
2009	E 环境下组件服务	Fabio Simeoni ^[4] 等人提出满足 E 环境的“数字图书馆组件服务”框架设计理念,实现与用户行为无缝连接。
2010	Probado	Probado ^[5] 主要目标是非文本文档数字服务功能的设计、开发与发布,同时,此工程提出了一套通用参考框架与协议,支持分布式非文本文档仓储联合服务。其中,主要面对的挑战是文档表达、索引以及基于内容的获取、组织与文档保存。

续表

年份	系统架构	具体描述
2011	UGC、CloudBooks、DL2.0	①Nicola Aloia 等来自多个国家的专家共同探讨“为欧洲数字图书馆设计的用户生成内容(UGC)服务模式” ^[6] ,支持欧洲用户协同工作,以多种方式为数字图书馆内容做出贡献,例如服务评述、批注、个人文化遗产收集与分享等,最终发布为公众所利用的服务; ②Hussein Suleman ^[7] 提出建设浏览器中的数字图书馆服务概念,实现在浏览器中进行操作计算,以增强用户体验,提高数字图书馆服务水平,打破原有依赖服务器的服务范式; ③Jennifer Pearson ^[8] 提出了支持在多种设备上阅读的服务框架(即 CloudBooks); ④Monika Akbar ^[9] 等人提出 DL2.0(下一代数字图书馆)作为整合社会知识和数字图书馆内容的泛在交流平台。

互操作机制一直被认为是数字图书馆跨区域、跨领域服务的重点。Kathleen Menzies^[10]重点介绍了“英国高等教育机构的信息系统互操作机制”,不再把互操作性作为单一的实现标准,而是作为一种起中介支持作用的“翻译问题”,通过一个中间共享数据交换平台,将相应服务链接起来。Nicola Ferro 等人^[11]提出了一个“共享文档描述元数据的方法体系”,实现在不同领域服务之间的无缝连接,另外 NDL^[12]项目工程,应用一种标准的“轻量级信息描述对象 XML 框架”(LIDO),同样可以规范描述博物馆资源对象,从而保证了可对图书馆与博物馆之间不同的数字对象进行检索、浏览和链接;文档批注^[13]方式的引入,创建了协作环境,丰富了数字图书馆中异构资源的集成,从而提高了互操作水平。而 Thomas Gotttron^[14] 提出一个“文档词云图”的概念,人与机器间互操作可通过视觉方式来完成,使用户更快、更容易理解内容。

3.2 数字图书馆是当前数字社会的一种表现形式

在 ECDL2007 中,Hussein Suleman^[2] 提出“没有数据库的数字图书馆”,即通过 XML + XSLT + XHTML 技术,充分利用社会化的知识内容,构建一个有用、能用的数字图书馆终端服务。而 EC-DL2009 中,Gary Machionini^[15] 以“数字图书馆与数字社会”为题做主旨报告,指出数字图书馆不仅已成为人们日常生活中不可或缺的工具,而且将成为大家参与、相互协作的社会知识服务平台。每一个

人都是知识的分享者,更是知识的生产者。因此,数字图书馆更要加强对标注、相互评价、标签云及平衡用户模式与用户体验的研究。优秀传统文化的数字服务,在 2011 年 TPD L 会议中也作为专题进行探讨。

3.3 下一代数字图书馆发展模式——以 DL2.0 为代表的协作研究社群环境

在 TPD L2011 大会中,以 Monika Akbar^[9] 为代表的 8 位专家,教育界知识传播与获取角度,提出了 DL2.0,即下一代数字图书馆的泛在交流平台。它整合社会知识和数字图书馆,利用用户与资源之间动态交互和虚拟社区,鼓励用户贡献、整合知识与资源,而且还扩展了内容管理、批注与评价、动态服务等,如内容的个性化和协同工作环境的创建。Adam Sofronijevic^[16] 以欧洲数字图书馆联盟的调研分析为基础,提出了以协作研究环境为基础,实现在阅读过程中的协作、互操作和参与,其核心理念类似于 DL2.0。Clifford Lynch^[17] 以“数字图书馆的未来:理念的演化”为题,从支撑科研环境建设角度,提出下一代数字图书馆将是一个科学知识管理环境,集成学术关系、学术社区和事实。

4 数字资源长期保存

ECDL 长期关注数字资源长期保存问题,对保存系统、方法、工具以及资源类型等多个方面进行

过探讨。在 2011 年 TPDL 会议中,针对文化遗产的数字保存服务受到了更为广泛的关注,从协作保存技术方法、元数据标准框架和语义相关性等角度予以深化。具体表现为数字资源的保存服务、战略技术工具与 Web Archive 三个方面。

4.1 数字资产的“保存服务与 workflow”

Carole Goble^[18]在 ECDL2008 会议中做了数字资产长期保存的“保存服务与 workflow”前沿报告,提出数字保存不仅仅是数字资产的管理、清洗、校验、批注、评价和可信赖保存的固化过程,而且应致力于面向未来的数字资产利用和增值,成为集保存服务、数字资产、工具以及代码脚本于一体的 workflow 场景,因此科学研究数字化流程的封装与数字资产、出版同样重要,也应该被保存。Silvia Wong^[19]指出,90%的航空发动机设计是变型设计,所利用到的 70%的信息都是以前的设计流程信息, workflow 场景的保存起到关键作用。所以,保存服务与 workflow,将扮演阅读出版物、处理数据和保存程序中的核心角色,是学术成果出版、保存、引用、组织、利用和存档等服务集成的新的表现形式。Carole Goble 还演示了 myGrid Taverna^① workflow 平台,目前该团队创建的 myExperiment 科学家社会网络和虚拟研究环境鼓励社团进行 workflow 的共享和保存。

4.2 长期保存战略及技术工具

数字资源长期保存已经成为许多国家和研究机构的战略计划,而且面对当前海量、多元化、分布的数字资源,长期保存范式已经发生了变化。2002

年实施的 NDIP 项目在 2009 年时又提出了创新计划^[20],来支持分散的 NDIP 资源收集。而 Mushashu Lumpa 等人^[21]针对 DSpace 与 LOCKSS 两个仓储平台,发挥各自优势、探索集成创新,在 DSpace 与 LOCKSS 之间创建一种通讯机制,无缝链接,解决了 DSpace 不支持由 LOCKSS 提供的高分布式复制功能的缺点。针对小型研究机构面临技术缺陷、资金匮乏、人员配置不够等问题,Hoppla 存档系统^[22]提供了一套解决方案。由亚历山大图书馆开发的 DAR 数字资产仓储集成平台^[23],有效管理数字资产的整个生命周期,建立了从生产、摄入、元数据管理、保存、出版和传播等一系列应用服务机制,也是一种选择。

4.3 Web Archive

ECDL 会议中经常涉及 web archive 问题。Web 存档主要的挑战是抓取、组织、保存,并能在未来提供离线浏览服务。基于这一点,Daniel Gomes^[24]等在 2011 年对 Web 存档计划进行了调查,列举了 1996 年之后启动的 42 个 Web 存档计划项目。这些项目大部分集中在发达国家和地区。该研究强调,Web 已经成为科研、教育的基础存在机制,但由于 Web 信息在线发布周期很短暂,80%以上的 Web 页面会在一年后更新或消失,所以 Web 存档对于长期保存意义重大。ECDL 针对 Web Archive 的数据收割、获取、网络保存以及随时间演化等技术问题,进行广泛的讨论与研究(见表 2)。

表 2 ECDL 研究 Web Archive 的几个关键问题

问 题	探讨解决方法
判断页面更新问题	Myriam Ben Saad ^[25] 等人提出了一种新的方法来检测页面版本的重要变化,即采用可视化网页段落,代表网页的语义块,使用 Vi-DIFF 监测段落变化的算法,对语义块的变化进行计算,进而决定页面是否具有重要变更。

① <http://www.mygrid.org.uk>

续表

问 题	探讨解决方法
数字内容一致性问题	Marc Spaniol ^[26] 等人针对 Web 存档中的数字内容一致性问题,采用可视化分析方式,将网络的演化行为进行可视化揭示。
网页连续性问题	英国国家档案局针对 Web Archive 内容的连续性问题,采用存档所有数字内容的方式 ^[27] ,不再考虑主题或者相关性,同时开发相应工具来自动维护运行版本与存档版本的内容链接关系,以保证现有版本不存在的情况下,系统会自动定位到 Web 存档版本中,支持 Web 页面的持久访问。 Myriam Ben Saad 等人 ^[28] 针对内容连续性,提出了通过面向连续性的 Web 抓取和导航算法,使用基于 Pattern 的先验解决方案,改善存档的连续性,而且又使用基于模式的新颖浏览方式作为后验解决方案,提高 Web 存档的连续性性能。
监测演化问题	Nina ^[29] 等人对跨越 200 年时间的文档进行监测,通过术语概念图来发现术语演化。

5 虚拟协同工作环境研究

随着网络技术和社群、社区与社交网络的普及,数字图书馆的应用环境也面临重大转变,逐步提出了将数字图书馆建设成为支持科研和教育的应用场所的要求。同时,各种科研信息的开放性不断扩大,e-Science 工程等也提出了开放性需

求^[30],搭建基于数字内容体系协作研究的基础设施已经迫在眉睫。ECDL 最近五年对此给予了高度关注,一些项目组已经创建了相对比较成熟的研究平台,为最终实现多领域协作的知识发现虚拟研究环境,形成基于 Web 的科学研究集成研究平台,实现科学数据开放共享与复用,实现用户研究平台的动态组合和虚拟实验室的动态构建等打下了根基。ECDL 历年在这方面的一些示例如表 3 所示。

表 3 虚拟协同工作环境研究资源列表

项目名称	项目需求背景	项目实施特点	相关工具或软件
基于共享阅读列表实现协作科研 ScholarLynk ^[31]	①学术研究需要综合调研、大量资料和信息资源,从中发现新的结论; ②研究人员在社群中交流、共享各自检索和阅读的文献,因此阅读列表成为一个代表学术领域的工作实体。	①将研究工作进行流程化描述,包括调查、分析、评估与知识创造等阶段,通过创建研究阅读列表同他人协作共享,形成相互交互意见的协作机制; ②拓宽个人或团队的研究知识点,提高个人或团队的研究水平,也保障了科学研究的突破性进展。	citeULike, Zotero,Connotea
虚拟研究环境的用户场景构建 D4Science-II EU ^[32]	①虚拟研究环境是基于网络的工作环境,满足分散用户的需求; ②用户场景真实体现研究人员需要做什么、能够做什么、怎样做,提高研究效能。	①动态构建用户界面,以面向组件的策略来动态选择、增加用户界面,增强虚拟社区之间的互操作; ②集成界面、服务、数据等应用服务。	gCube

续表

项目名称	项目需求背景	项目实施特点	相关工具或软件
群组协作工作平台 Cheshire	富交互性环境应用到数字图书馆中,支持学生、研究人员以群组、社区的形式协同完成工作,弥补自己知识点的盲区。	①Christian Wartena ^[33] 提出的协作标签概念与方法体系,通过终端用户对信息进行标签化; ②Alison Babau ^[34] 通过创建一个虚拟研究环境(VRE)作为知识发现平台,实现领域内的知识组织与识别,例如对人以及其他对象的识别; ③Cheshire ^[35] 支持“分布式批注系统协作研究”,无论文件在什么位置或是什么格式,将批注以 XML 文件与原文件关联并单独存储,多个文件可共享一个批注; ④Rainer Simon ^[36] 开发了多媒体批注框架,即 Yuma 平台。	Cheshire、Yuma
数字图书馆协作新范式 Reading2.0	针对用户在阅读过程中不断变化需求与操作行为进行信息组织与描述,反映用户的真实需求。	①用户的期望是一个协作、互动的环境; ②图书馆员与用户之间的协作问题; ③图书馆馆员之间的协作问题。	Enterpriser2.0

6 知识组织体系建设

2007—2010 年间,Web 2.0 相关技术的应用逐步成熟,ECDL 关注的研究主题也逐年增多。而在 2011 年,语义网与关联数据受到追捧,很可能演化成未来数字图书馆领域的研究热点,以顺应知识发现与知识服务的需求。在知识组织体系建设中,叙词表、分类体系、本体与词汇数据库等方面的研究更加深入,许多项目都尝试为某一领域的语义结构进行建模,最终提供智能检索、知识发现、语言映射以及语义网等知识服务。

6.1 知识组织体系建设和维护

在现实信息环境中,往往有太多的语料很难准确快速地收集和规范并纳入知识组织体系,特别是当前新概念、新词出现频率加快,更是增加了难度,尤其是要将“新出现”的语义概念与已有概念对齐(alignment)。欧洲数字图书馆联盟采取“软模块解决方案”^[37]来克服这个难题,通过各个来源词表中概念之间和元数据元素结构之间的两种对齐

方式,来创建一个语义参考网络,实现自动语义对齐,但是为了保证知识组织体系的质量,大规模的人工对齐方式仍然是一个重要环节。Christian Mader^[38]提出将简单知识组织系统(SKOS)作为当前可控网络词汇的出现标准,即判断一个词是否属于新词,例如 AGROVOC, EuroVoc, GEMET, and STW,依据 SKOS 自动判断新发现的词汇是否是新词汇,如果是新词汇即加入 SKOS,否则进行词汇的规范。

在知识组织体系的发布与维护方面, Armando Stellato 等人^[39]开发了 VocBench 平台,提供知识组织体系的扩展、更新和版本维护等功能。它具有以下特点:多家研究机构进行协同化管理;遵循 W3C 框架标准;开源工具;对外部数据进行收割,自动加入 VocBench 平台的知识组织体系。

6.2 知识组织体系与知识服务

知识组织体系建设的一个重要目标是提供知识服务,为未来的语义网建立基础。尽管目前取得的效果仍然有限,但 Felix Boteram^[40]提出,在当前信息高度异构的环境下,知识组织系统集功能、表

达以及相互关系于一体,将这些知识单元相互联系,才是进行信息检索与知识探索最有效的策略。Marjorie M. K. Hlava^[41]等人使用知识组织体系进行情报研究与分析,通过 KOS 或者语料库分别对 IEEE、Xplore 等几个数据库中的十年文献进行概念抽取,分析这些数据库的影响力以及发展趋势。

同时利用社会标签,扩展知识组织体系,增强知识服务能力。Denise A. D. Bedford 等^[42]提出将“社会标签”增加到知识组织体系中。具体过程如下:首先初始化一定体量的“社会标签”概念进入知识组织体系,然后用户根据知识组织体系给出的标签概念的排序,选择作为自己的 Social Tags,并可根据知识发现的内容,创建新的“社会标签概念”来扩充知识组织体系,实现知识组织体系的自我完善与扩充。

7 用户行为研究与 Web 2.0 技术

数字图书馆设计与建设的目标就是为用户服务,能够得到用户的认可服务成功的考核关键。了解用户的需求、操作特性、知觉特征、认知心理等,同时利用 Web 2.0 等先进技术,模拟用户的操作行为,增强用户体验,保证所提供的服务符合用户的习惯、经验与期待。ECDL 会议从 2007 到 2011 五年间,一直把用户行为研究作为关注焦点。

7.1 用户场景设计——分面与协作服务

在 2007 年 ECDL 会议中,Michal Tvarůzek 等人提出一种“分面浏览自适应的方法体系”^[43],支持检索结果的动态分面,一定程度上实现了知识分类。国际上的 summon、primo 及 google 等机构提供的发现服务中,都将分面作为它们的主要支撑技术。上海交通大学图书馆、北京大学图书馆以及西安交通大学等均引进上述发现服务中一种,完成了门户的更新。在体验完善方面, Peter Brusilovsky^[44]等人提出利用社会导航^[45]来提高数字图书馆服务。

由上述数字图书馆的发展看出,以协作服务场景主导的设计可能是下一代数字图书馆的重要

特征。而目前以 wikisearch^[46]为代表的协作编辑服务已经得到用户好评,其以可视化的形式揭示知识信息,提高了用户认知能力,支持形成一个内容再利用、再扩展的知识平台。Hamed Alhoori 等人^[47]在协作服务方面进行了一个调查研究,主要在协作和社会使用、社会参考管理人员两个方面,针对当前实践和动态学术活动,将定性采访方法和定量在线调查方法相结合进行数据收集,然后跟踪、比对分析,从中发现动态信息需求、用户行为和研究人员的学术活动,来帮助与改善研究人员的活动和数字图书馆的服务水平。

7.2 个人信息管理——数据恢复角度

个人信息管理是获取、组织、维护、检索和使用个人使用的文档、网页和 Email 等信息的机制,关键是要支持和帮助用户在正确的地方,以正确的格式,发现正确的信息,进行个性化组织。Anna Zaccchi 等人^[48]主要从用户习惯、服务机制的角度来加强个人信息管理,首先分析识别研究用户所处工作环境中的相关活动、资源文档的存储环境,然后决定对资源组织、命名,以及如何进行有效配置。

Rudolf^[49]等人从个人信息丢失的角度,研究如何进行恢复,提出一种新的个人信息灾难恢复方法,通过分析个人信息(如社区、邮件、博客、评论等)中所包含相关人员的上下文信息,从网络空间抽取信息,经组合处理后重新生成用户个人信息。

7.3 电子阅读行为研究——电子阅读桌面

技术的革新使电子阅读的模式更加丰富,操作更加人性化。Jennifer Pearson^[50]等人提出“电子阅读桌面”概念,通过设计一个文档阅读界面,模拟物理阅读环境,比如做笔记批注、折叠文件角或者插入书签等行为,形成读者与阅读桌面的互操作性。Moiria Norrie^[51]认为纸张在使用中所具有的启示性能够保证它将来继续在工作、家庭和公共场所使用。而且,随着信息技术的发展,许多研究工程已经将纸张与电子媒体、服务进行集成整合,使得触摸与电子纸张技术的混合构成取得巨大成功,例如电子白板、电子桌面等,已经实现了人的电子执行

行为更接近于现实中的翻页、取书、选择书中的某一部分以及进行批注等行为。

8 小结

ECDL 所关注的问题反映了数字图书馆领域的发展趋势。许多早期讨论的概念逐步实现了大规模应用,例如 2008 年关注的检索结果分面模式,已为许多数据库商以及图书馆服务机构所应用。综观 ECDL2007—2011 年,从信息检索技术的探索到知识组织及知识服务模式的探索,从单一数字资源的长期保存到跨地域、跨领域的协作保存,从用

户的被动服务模式到主动提供知识服务,从单一服务机制到任务驱动 workflow 服务机制,从封闭研究环境到协作虚拟研究环境,充分体现了数字图书馆服务范式的变化。当前,伴随着科学研究第四范式发展,知识创造方法与技术面临新的突破,例如以关联数据为代表的数据整合和集成服务模式,促进结构化数据信息的协作式产生和重复性使用,可以使图书馆、档案馆、博物馆数据、政府数据、研究数据等多种数据形态无缝结合,从而为用户提供融汇、集成的知识发现、应用服务,同时也促进了基于 e-Research 的协作科研信息环境的发展。这无疑是数字图书馆发展的一个机会。

参考文献:

- [1] Dimitris Gavrili. Enhancing library services with Web 2.0 functionalities[C]//Research and Advanced Technology for Digital Libraries, 2008; 148 – 159.
- [2] Hussein Suleman. Digital libraries without databases: The Bleek and Lloyd collection[C]// Research and Advanced Technology for Digital Libraries, 2007; 392 – 403.
- [3] Assante M, Candela L, et al. An extensible virtual digital libraries generator[C]// Research and Advanced Technology for Digital Libraries, 2008; 122 – 134.
- [4] Fabio Simeoni. Functional adaptivity for digital library services in e-infrastructures: The gCube approach[G]// Research and Advanced Technology for Digital Libraries, 2009; 51 – 62.
- [5] Berndt R, Blumel I, Clausen M, et al. The PROBADO project – approach and lessons learned in building a digital library system for heterogeneous non-textual documents[C]//Research and Advanced Technology for Digital Libraries, 2010; 376 – 383.
- [6] Nicola Aloia, Cesare Concordia, Anne Marie van Gerwen, et al. Design, implementation and evaluation of a user generated content service for Europeana[C]// Research and Advanced Technology for Digital Libraries, 2011; 477 – 482.
- [7] Hussein Suleman. In-browser digital library services[C]// Research and Advanced Technology for Digital Libraries, 2007; 462 – 465.
- [8] Jennifer Pearson, George Buchanan. CloudBooks: An infrastructure for reading on multiple devices[C]// Research and Advanced Technology for Digital Libraries, 2011; 488 – 492.
- [9] Monika Akbar, Weiguo Fan, et al. Digital library 2.0 for educational resources [C]// Research and Advanced Technology for Digital Libraries, 2011; 89 – 100.
- [10] Kathleen Menzies, Duncan Birrell, et al. New evidence on the interoperability of information systems within UK higher education institutions[C]// Research and Advanced Technology for Digital Libraries, 2010; 104 – 115.
- [11] Nicola Ferro. A methodology for sharing archival descriptive metadata in a distributed environment[C]// Research and Advanced Technology for Digital Libraries, 2008; 268 – 279.
- [12] Riitta Autere. Towards cross-organizational interoperability: The LIDO XML schema as a national level integration tool for the National Digital Library of Finland[C]//Research and Advanced Technology for Digital Libraries, 2011; 62 – 68.
- [13] Maristella Agosti, Nicola Ferro. Annotations: A way to interoperability in DL[C]// Research and Advanced Technology for Digital Libraries, 2008; 291 – 295.

- [14] Thomas Gottron. Document word clouds: Visualising web documents as tag clouds to aid users in relevance decisions [C]//Research and Advanced Technology for Digital Libraries, 2009: 94–105.
- [15] Gary M. Digital libraries as Phenotypes for digital societies[C]//Research and Advanced Technology for Digital Libraries, 2009: 1.
- [16] Adam Sofronijevic. New paradigm of library collaboration[C]//Research and Advanced Technology for Digital Libraries, 2011: 519–522.
- [17] Clifford Lynch. The futures of digital libraries: The evolution of an idea[C]//Research and Advanced Technology for Digital Libraries, 2011: 2.
- [18] Carole Goble. Curating services and workflows[EB/OL]. [2012–05–12]. <http://www.ecdl2008.org/documents/CaroleGobles.pdf>.
- [19] Sylvia C Wong. On a service-oriented approach for an engineering knowledge desktop[EB/OL]. [2012–05–12]. <http://eprints.soton.ac.uk/262184/1/pp050-wong.pdf>.
- [20] Laura E. Campbell. Recollection: Integrating data through access[C]//Research and Advanced Technology for Digital Libraries, 2009: 396–397.
- [21] Mushashu Lumpa. Interconnecting DSpace and LOCKSS[C]//Research and Advanced Technology for Digital Libraries, 2011: 493–496.
- [22] Stephan Strodl, Florian Motlik, Andreas Rawlber. Hoppla-digital preservation support for small institutions[C]// Research and Advanced Technology for Digital Libraries, 2009: 493–494.
- [23] Youssef Mikhail, et al. DAR: Institutional repository integration in action[C]//Research and Advanced Technology for Digital Libraries, 2011: 348–359.
- [24] Daniel Gomes. A survey on web archiving initiatives[C]//Research and Advanced Technology for Digital Libraries, 2011: 408–420.
- [25] Myriam B S, Stéphane G, Zeynep P. A novel web archiving approach based on visual pages analysis[C]//IWA2009, 2009: 10–19.
- [26] Marc S, Arturas M, Dimitar D, et al. Catch me if you can: Visual analysis of coherence defects in web archiving[C]//IWA2009, 2009: 27–37.
- [27] Amanda S, Brian O, Gabriel V. Past and present: Using the UK government web archive to bridge the continuity gap [C]//IWA2009, 2009: 38–42.
- [28] Myriam Ben Saad, et al. Coherence-oriented crawling and navigation using patterns for web archives [C]// Research and Advanced Technology for Digital Libraries, 2011: 421–433.
- [29] Nina T, Sukriti R, Thomas R. First results on detecting term evolutions[C]//IWA2009, 2009: 50–55.
- [30] Jenny Fry, Ralph Schroeder, et al. Open science in e-science: Contingency or policy?[J]. Journal of Documentation, 2009, 65(1): 6–32.
- [31] Gabriella Kazai. Architecture for a collaborative research environment based on reading list sharing[C]//Research and Advanced Technology for Digital Libraries, 2010: 294–306.
- [32] Massimiliano Assante. An approach to virtual research environment user interfaces dynamic construction[C]//Research and Advanced Technology for Digital Libraries, 2011: 101–109.
- [33] Christian Wartena. Automatic classification of social tags: A case study[C]//Research and Advanced Technology for Digital Libraries, 2010: 176–183.
- [34] Alison Babeu. Named entity identification and cyber infrastructure[EB/OL]. [2012–05–12]. <http://www.perseus.tufts.edu/~ababeu/ecdl2007.pdf>.
- [35] Fabio Corubolo, Paul B Watry. Location and format independent distributed annotations for collaborative research[EB/OL]. [2012–05–12]. <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.102.131&rep=rep1&type=pdf>.

- [36] Rainer Simon. The YUMA media annotation framework[C]//Research and Advanced Technology for Digital Libraries, 2011: 434–437.
- [37] Antoine Isaac, Jacco van Ossenbruggen, et al. Europeana and semantic alignment of vocabularies[EB/OL]. [2012–05–12] <https://www.comp.glam.ac.uk/pages/research/hypermedia/nkos/nkos2011/presentations/Isaac-NKOS11.pdf>.
- [38] Christian Mader, Bernhard Haslhofer. Quality criteria for controlled Web vocabularies[C]//International Conference on Theory and Practice of Digital Libraries, 2011, NKOS Workshop.
- [39] Armando Stellato. A collaborative framework for managing and publishing KOS[EB/OL]. [2012–05–12]. <http://www.fao.org/docrep/article/am814e.pdf>.
- [40] Felix Boteram. Information retrieval and knowledge exploration in distributed knowledge organization systems[EB/OL]. [2012–05–12]. <https://www.comp.glam.ac.uk/pages/research/hypermedia/nkos/nkos2010/abstracts/boteram.pdf>.
- [41] Marjorie M K Hlava. Using KOS as a basis for text analytics and trend forecasting[EB/OL]. [2012–05–12]. <https://www.comp.glam.ac.uk/pages/research/hypermedia/nkos/nkos2010/abstracts/hlava.pdf>.
- [42] Denise Bedford. Selecting and weighting semantically discovered concepts as social tags[EB/OL]. [2012–05–12]. <https://www.comp.glam.ac.uk/pages/research/hypermedia/nkos/nkos2011/abstracts/Bedford.pdf>.
- [43] Michal Tvarůzek. Personalized faceted browsing for digital libraries[C]//Research and Advanced Technology for Digital Libraries, 2007: 485–488.
- [44] Peter Brusilovsky. Enhancing digital libraries with social navigation: The case of ensemble[C]//Research and Advanced Technology for Digital Libraries, 2010: 116–123.
- [45] Dieberger A, Dourish P. Social navigation: Techniques for building more usable systems[J]. ACM interactions, 2000, 7(6): 38–40.
- [46] Elaine G Toms. WikiSearch—From access to use[C]//Research and Advanced Technology for Digital Libraries, 2009: 27–38.
- [47] Hamed Alhoori. Understanding the dynamic scholarly research needs and behavior as applied to social reference management[C]//Research and Advanced Technology for Digital Libraries, 2011: 169–178.
- [48] Anna Zacchi, Frank Shipman. Personal environment management[C]//Research and Advanced Technology for Digital Libraries, 2007: 345–356.
- [49] Rudolf M, Robert N, Andreas R. Data recovery from distributed personal repositories[C]//Research and Advanced Technology for Digital Libraries, 2009: 455–458.
- [50] J Pearson. The reading desk: Supporting lightweight note-taking in digital documents[C]//Research and Advanced Technology for Digital Libraries, 2011: 438.
- [51] Moira C. Norrie. Paper, pen and touch[C]//Research and Advanced Technology for Digital Libraries, 2011: 1.

钱 力 中国科学院国家科学图书馆 2011 级博士研究生。

通讯地址: 北京北四环西路 33 号。邮编: 100190。

张晓林 中国科学院国家科学图书馆馆长。通讯地址同上。

张智雄 中国科学院国家科学图书馆研究馆员。通讯地址同上。

(收稿日期: 2012–06–06)