

国外数字图书馆评价研究与实践述评

王咏梅

【摘要】本文介绍了国外数字图书馆评价的最新发展动态,着重评述了国外数字图书馆在评价对象、标准、框架等方面的研究与实践。

【关键词】数字图书馆 评价

Abstract: By introducing the new trends of digital library evaluation abroad, the paper reviews the researches and practices on construct, criteria, framework and so on of digital library evaluation.

Key words: digital library evaluation

从20世纪90年代起,数字图书馆的研究及实践在全球蓬勃发展,但与其相关的数字图书馆评价却显得相对滞后。美国RUTGERS大学教授萨拉塞维克(TEFKO SARACEVIC)分析其中可能原由时提到:(1)数字图书馆发展尚未成熟,评价为时过早;(2)评价的例证不足,无法进行全面性评价;(3)人们对数字图书馆这一新生事物的兴趣受到抑制;(4)数字图书馆评价十分复杂,利用原有传统图书馆的知识无法进行评价。也就是说,“现有数字图书馆评价理论不足以指导实际的评价工作”^[1]。

莫切尼尼(Gary Marchionini)曾评述到:评价数字图书馆有点像评价一桩婚姻。婚姻是否美满,不仅看一方,还要看双方,最后还要看双方匹配如何。数字图书馆也一样,它是集使命、技术、实体图书馆文化、功能及计算机和通信文化为一体的复杂系统,评价其成功与否,也不能只是片面评价^[2]。尽管存在种种困难,但人们已开始认识到数字图书馆评价的重要性。近年来国外有关数字图书馆评价研究及实践的报道日显增多,其研究焦点主要集中在评价对象、评价框架、评价标准及评价方法四个方面。

1 关于评价对象的研究与实践

评价对象,也就是要评价的具体内容。由于目前对数字图书馆的理解和界定还不统一,要确定数字图书馆的具体内容并不容易。因此,国外数字图书馆评价研究实践多集中在将数字图书馆作为一整体进行评价或是对数字图书馆的某些过程进行评价两个方面。

1.1 对数字图书馆作为整体进行评价

由于数字图书馆定义不一,用户群体各不相同,建设的最终目标也大相径庭,整体评价则是根据不同数字图书馆建设类型,在建设目标的框架下,将其作

为一个特定的整体进行评价。在此,本文将数字图书馆按建设目标分为研发型和实践型。

1.1.1 研发型

研发型数字图书馆的建设目的以研发为主。其评价的主要目的是通过评价进一步认识数字图书馆这一事物及其影响。因此涉及的领域也较为广泛。其主要代表是美国Persues数字图书馆建设项目(Persues Digital Library)。美国Persues是以研发为主的有关古典人文学数字图书馆项目,其建设目标是向最为广大的用户群提供大规模的人文学研究的原始资料。项目评价与项目建设同时启动(1989),并不断发展完善。项目将评价定位为研发的一部分,而非真正的对最终产品的检测。评估组领导人莫切尼尼(Gary Marchionini)认为,数字图书馆是实体图书馆的延伸和增长。可以借鉴传统图书馆的评价指标,结合数字技术的评价指标,以此作为评价数字图书馆新服务、新产品和新功能的起点^[3]。该评价框架的最大特点在于评价对象与目标相交织,并以此确定一整套探讨性提问,提出数据收集的变量以及问题的分析方法。项目确定四个评价对象:学生、教师、技术系统及馆藏。此外,项目还确定了以“观察、访谈、数据分析及学习分析”四种收集数据方法。

1.1.2 实践型

实践型的项目注重实际用途和实际可操作性。因此评价重点也就放在用户使用上。比如英国电子图书馆项目(elib)的中心思想是“实践性”。其研究的重点不是新技术开发,而是如何将信息和网络技术应用于高校图书馆工作实践^[4]。elib三期的HyLiFe项目的评价所确定的主要对象之一就是项目界面绩效评价,从技术和可用性两个方面入手,主要通过最初用户需

求分析、问卷调查、系统交互记录分析和用户跟踪调查四个渠道来收集数据,支持开发以用户为中心的界面^[5]。

1.2 对数字图书馆系统过程进行评价

过程评价是抛开数字图书馆这一特定的整体,仅对与其相关的系统过程进行评价。当然,由于各项项目的目标与评价目的并不相同,因此选择的过程对象也不同,但主要包括以下几项:

(1) 评价数字图书馆使用不同描述,如专业术语、语意描述、主题句抽取等;(2) 评价不同的工具,如链接、视频搜索、以用户为中心的界面、图像检索的方法、服务器负载均衡等;(3) 评价服务,如宣传数字图书馆系统、数字参考服务等;(4) 评价计划;(5) 研究数字图书馆及服务的相关用户行为,如用户使用模式、用户眼中的数字图书馆、用户偏爱的搜索方式等。

萨拉塞维克教授指出:最具挑战是关于用户行为评价。因为关于用户行为的涉及面极广,研究的界线在哪里?是否只研究用户如何使用系统^[6]?当然某些用户行为,如系统使用的难易程度等问题,许多评价都有包括,但是否还涉及其它方面?这些问题都有待我们进一步探讨。

2 关于评价框架的研究与实践

评价框架包括评价方法、方向、角度、水平和目的。由于数字图书馆是一个集社会、机构及技术为一体的复杂系统。评价不可能包括所有方面,因此只能从评价的目的出发,从不同角度,采用不同的评价方法进行评价。现将国外在这一方面的研究与实践概况如下:

2.1 以系统为中心

这是最为常用的评价方法。通常包括系统性能评价。主要是对系统特性、设计以及技术方面的效用和效益进行评价。通过评价改进设计和运作。

2.2 以人中心

这也是一种广泛采用的方法。包括有关信息需求行为的研究,如信息检索、浏览、查询以及完成工作的性能。这种方法的目的是为了揭示用户行为、需求以及用户解决问题途径,并以此为基点,改进系统。

2.3 以可用性为中心

这一方法主要是从用户的角度出发,评价系统的不同功能特征。这是一种介于以系统为中心和以人中心的评价之间的方法,属于人机交互研究领域。

2.4 人文学方法

主要通过研究生活方式、文化、数字图书馆环境,评价数字图书馆影响度。此外,一些评价还触及到研究数字图书馆不同利益权人(Stakeholders)以及他们的文化与数字图书馆的关系。这一方法主要为了揭

示不同背景的用户使用数字图书馆的情况。

2.5 经济学方法

评价包括了成本、成本效益、经济价值及影响度。这种评价曾在数字图书馆建设初期,如美国 PEAK 项目的评价对象就包含了价格模型、收入和费用等经济因素^[7],但现在却甚少项目触及这一方面。

评价的水平则包括了从微观的特征到宏观的影响度评价。如美国开放式视频数字图书馆项目(The open videl digital library)评价就着眼于研究图像代理的识别速度这一微观层面^[8];而 Perseus 项目及英国电子图书馆的评价都看重项目的影响度,即宏观层面的评估。

3 关于评价标准的研究与实践

评价标准是评价活动的核心,对评价结果具有直接的影响。评价结论的不同,最主要的原因,是评价标准的不同。虽然上世纪 50 年代,曾提出针对信息检索系统评价的“适用性”的基本标准,但显然不适用于复杂的数字图书馆系统。对此,不同的项目都提出各自不同评价标准,主要集中在四个方面:可用性、系统性能、使用、人文及其他。如下表所示:

表 1 数字图书馆评价标准体系

可用性标准	信息资源	存取性、可获取性
		清晰度
		复杂度
		信息内容/量
		透明度
		可理解度
		适当度
		覆盖面
		质量/精确度
		有效性/可靠性
	过程(检索、浏览、获取等)	学习能力
		用户努力
		方便性/难易程度
		失败情况
		操作指南
		完成情况
		解析的难易程度
		对结果的信赖度
		出错率
	文档格式	吸引力/持续使用度
		稳定性
		标识描述性
		信息传递性
	总体评估	满意度
		成功率
		适用性
		影响测度及价值
		用户感受的好坏程度
		障碍
		偏爱程度
		学习程度

系统性能标准	技术性能	反应时间
		运行时间/速度
		功能/容量
	程序/运算性能	适用性
		聚类性
		相似性
		功能性
		灵活性
		友好程度
		出错率
		最优化
		逻辑判定
		路径长度
		点击通过率
		检索时间
	系统总体标准	可维护性
		可量测性
		协同工作能力
		共享性
使用标准		成本
		使用模式
		使用材料
		使用统计
		使用人、使用对象、使用时间
人文及其它标准		使用目的
		不同用户群体
		用户群体的背景及文化
		语言
		学习
		影响测度

体,许多方法都适用于其评价,但每种方法都有其优缺点,因此实际上,并不存在着一个最佳的评价方法^[9]。

由于数字图书馆仍处于成长阶段,随之发展起来的相关理论与实践也尚未成熟。尤其是数字图书馆评价更是处于起步阶段,完整的、系统化的评价体系尚未建成。而我国的数字图书馆的评价的研究与实践,虽然在一些学术刊物上已出现了一些理论性评述,但在实践方面依旧是一空白点。本文通过介绍国外数字图书馆评价的最新发展动态,着重评述了国外数字图书馆在评价对象、标准、框架等方面的研究与实践,以期对我国在这一领域的探索有所启发。

注释

- [1] [3] Saracevic, T. Digital Library Evaluation: Toward an Evaluation of Concepts. *Library Trends*, 2000, 49 (2): 350 ~ 369
- [2] Gray Marchionini. Evaluating Digital Libraries: A Longitudinal and Multifaceted View. *Library Trends*, 2001, 49 (2): 304 ~ 333
- [4] 王咏梅. 英国复合图书馆项目及贡献. *情报杂志*, 2004 (10): 90 ~ 92
- [5] 王咏梅. 关于英国 HyllFe 复合图书馆评介. *图书馆杂志*, 2004 (12): 57 ~ 58
- [6] [9] Saracevic, T. Evaluation of digital libraries: An overview. Presentation at the DELOS WP7 Workshop on the Evaluation of Digital Libraries, 4 - 5 October 2004 <http://www.scils.rutgers.edu/~tefko/articles.htm> [2005-07-05]
- [7] Mackie - Mason, J. K, Riveros, J. F. A Report on the Peak Experiment Usage and economic behavior, *D - Lib magazine*, 1999, 5 (7/8). <http://www.dlib.org/dlib/july99/mackie-mason/07mackie-mason.html> [2005-07-05]
- [8] Gray Marchionini, Gray Geisler. The Open Video Digital Library. *D - lib magazine*, 2002, 8 (12). <http://www.dlib.org/dlib/december02/marchionini/12marchionini.html> [2005-12-25]

4 关于评价方法论的研究与实践

至今,有关数字图书馆的评价的理论方法尚未形成,但大部分评价都采用定量与定性相结合的方法,唯一没有涉及到的是定性的历史研究法,主要包括以下几种方法:(1)调查,包括直接提问与在线调查;(2)结构式访谈;(3)个案调研;(4)社会调查;(5)观察跟踪;(6)系统交互记录;(7)实验方法;(8)记录分析;(9)使用分析;(10)文件、会议以及分析(人文学范畴);(11)经济分析。

SARACEVIC 教授认为:数字图书馆是个复杂的整

参考文献

- 1 乔欢. 数字图书馆评价研究. *国家图书馆学刊*, 2004 (3)
- 2 刘兹恒, 王海娟. 关于数字图书馆评价的若干问题探讨. *四川图书馆学报*, 2004 (3)
- 3 赵凯威, 张桂玲. 国外数字图书馆评价研究综述. *情报科学*, 2003 (5)
- 4 区颖薇, 沈平. 关于数字图书馆评价体系的思考. *图书馆学研究*, 2004 (2)

王咏梅 福州大学图书馆。