

从交互维度到交互功能:构建数字图书馆交互评估理论模型^{*}

李月琳 梁娜 齐雪

摘要 数字图书馆评估研究从不同角度确立了大量的评估指标,但是评估指标的多元化也造成了共识性缺乏的弊端,削弱了评估成果的应用价值和应用绩效;特别是关于数字图书馆交互功能评估的研究,仍处于成果短缺、零散的状态。基于此,本研究以用户与数字图书馆的交互模型为理论基础,采用实验研究方法,探索用户与数字图书馆的交互过程,分析不同的交互维度、子维度及其相关指标对于数字图书馆交互功能评估的作用,最终获得数字图书馆交互功能评估的预测性子维度与关联性子维度,并据此构建了数字图书馆交互功能评估模型,为数字图书馆交互评估提供了理论框架,同时也为提升数字图书馆交互功能的设计提供借鉴。图1。表12。参考文献39。

关键词 数字图书馆评估 交互功能 理论模型

分类号 G252

From Interaction Dimensions to Interaction Function: Developing a Theoretical Model for Evaluating Interaction in Digital Libraries

LI Yuelin, LIANG Na & QI Xue

ABSTRACT

So far, many researchers have done research on digital library evaluation, mostly focusing on usability issues or conceptual framework for comprehensively evaluating digital libraries. These studies have produced many criteria or indicators for evaluation. However, the agreement on these criteria or indicators is still lacking. Such situation weakens the implications of these studies. In particular, much less studies focus on evaluating interaction function of digital libraries. Thus, this study attempts to identify the critical criteria or factors affecting users' evaluation of interaction function of digital libraries and make the evaluation more effective and efficient. For the purpose, the study explores user-digital library interaction in three dimensions, such as information, technology, and tasks, by which to examine the relationships between users' interaction with these dimensions and their evaluation of interaction function of digital libraries. An experiment was

^{*} 本文系国家社会科学基金项目“基于用户多维交互行为的数字图书馆可持续发展评估模型与实证研究”(编号:11BTQ009)和教育部新世纪优秀人才支持计划“个性化数字图书馆”的研究成果之一。(This article is an outcome of the project “Developing a Digital Library (DL) Evaluation to Support Sustainable Development of DL Based on User's Multi-dimensional Interactive Behavior” (No.11BTQ009) supported by National Social Science Foundation of China and “Personalization of Digital Library” supported by program for New Century Excellent Talents in University.)

通信作者:李月琳,Email:yuelinli70@163.com,ORCID:0000-0002-1496-6741 (Correspondence should be addressed to LI Yuelin,Email:yuelinli70@163.com,ORCID:0000-0002-1496-6741)

conducted to explore the issue. The study performed card sorting to identify four topics that potential participants were interested. Then, based on the topics, the study designed four simulated work task situations. In addition, the participants were required to bring a real work task to the experiment. Latin Square was used for the experimental design. During the experiment, the participants were required to think aloud. The study recruited 48 participants. CNKI, one of the most popular digital libraries, was selected to be the experimental system. During the experiment, the participants were required to finish the entry questionnaire, pre-search questionnaire, post-search questionnaire, and evaluation form. The study used SPSS to perform Pearson correlation analysis and Stepwise regression analysis. It reveals the factors that are significantly correlated to users' evaluation result of digital libraries, including information authenticity, comprehensiveness of collection, navigational function, and so on, labeled as "correlated sub-dimension". It also identifies the factors that could predict users' evaluation results of digital libraries, such as information appropriateness, link richness and validity, and rationality of page layout, and so forth, labeled as "predictive sub-dimension". Based on the results, the study develops a theoretical model for evaluating interaction function of digital libraries.

The originality of the study lies in that it identifies different sub-dimensions of each dimension in the integrated three-dimension model of user-digital library interaction and singles out the critical factors from those identified by previous studies. Furthermore, it identifies the correlated and predictive sub-dimensions that affect users' evaluation results of user-digital library interaction. The study then develops the theoretical model that could be used to evaluate interaction between users and digital libraries, as well as inform the interaction design of digital libraries.

The study only analyzes the data collected from the questionnaires and does not examine the experimental recordings. Also, the study only analyzes participants' perception of their interaction with the digital library. In fact, their search performance, such as precision or average precision, may be also necessary for in-depth analysis. In addition, the study asked the participants to bring a real work task for the experiment. However, for convenience of data analysis, the study treats all 48 real work tasks as the same one. This may bias the results. Future studies will analyze the recordings and further categorize the real work tasks based on their task characteristics for more robust results. 1 fig. 12 tabs. 39 refs.

KEY WORDS

Digital library evaluation. Interaction function. Theory model.

0 引言

数字图书馆是一种交互信息检索系统,是一系列交互元素的集合^[1],支持用户与系统的交互是数字图书馆开发与设计的重要方面^[2]。交互的成功是用户有效获取数字图书馆资源的重要保证,为此,应将交互功能纳入到数字图书馆设计和评估活动之中。

数字图书馆评估方法和路径多种多样,通常可以归纳为以系统为中心的评估和以用户为中心的评估。当前的数字图书馆评估仍是以系统为中心^[3],以用户为视角的评估相对短缺,尤其缺少针对数字图书馆交互功能的评估。毋庸置疑,数字图书馆支持的是用户与知识的互动^[4],从这一视角,评估应以用户而非系统、以交互而非界面为焦点^[5]。因此,本研究拟从交互的角度探讨数字图书馆的评估,此角度既能

了解系统功能的质量,又能揭示用户对数字图书馆不同维度的感知,提供了一种可以整合系统和用户两种途径的新视角。根据李月琳等的研究^[6],在检索过程中,用户与数字图书馆在技术维、信息维与任务维等不同维度上进行交互,从而影响用户与系统的交互绩效。然而,这些维度如何影响用户对数字图书馆交互功能的评估,尚待厘清。数字图书馆的交互功能体现在多个方面,就其信息检索系统的本质而言,搜索和浏览功能^[7]是最为重要的交互机制:数字图书馆的搜索界面和结果界面共同支持用户的搜索活动,与此同时,这两种界面还具有支持用户浏览以获取有用信息的功能,用户通过搜索和浏览实现与系统的交互。系统所提供的内容与技术也同时体现在这些界面上,并通过这些界面传递给用户,用户便是在不同的界面与层次上实现与系统的交互,获取有用信息,以支持工作任务的完成。因此,数字图书馆交互功能对用户搜索过程的支持往往决定着一个数字图书馆的发展前景。

基于上述问题,本研究以交互为中心,关注数字图书馆交互功能对用户搜索过程的支持,试图厘清用户与不同层面或维度的交互如何影响其对数字图书馆交互功能的评估,探索数字图书馆交互功能评估的可预测性,以此帮助数字图书馆完善交互评估体系,提升其交互功能的开发与设计。

1 文献回顾

1.1 数字图书馆评估研究

评估是数字图书馆非常重要的研究领域。美国数字图书馆创始计划一期和二期即包含了众多评估项目。此外,一些研究机构也组织了大量的数字图书馆评估项目,催生了一批与项目同名的评价指标,包括 DigiQUAL^[8]、E-Metrics^[9]等。我国的数字图书馆评估研究起步虽晚,但也开发了一些普适性的标准,如 WH/T 47—2012(图书馆数字资源统计规范)^[10]、

CALIS 数字资源评估指标体系^[11]等。上述评估指标体系,既有单独针对电子资源/数字资源或是数字服务的指标,也有全面考评数字图书馆质量的指标。

Saracevic 提出数字图书馆评估的理论框架,该框架由以用户为中心的社会、机构、个人,连接用户与系统的界面,及以系统为中心的工程、处理、内容层面构成^[1]。在此基础上,Zhang Y 构建了由内容、技术、界面、服务、用户、情境六个方面构成的数字图书馆整体评估指标体系,发现用户首先关注的是内容的可用性、服务的可持续性,其次关心内容和服务对于交互的影响,再次关心使用数字图书馆的绩效^[12]。Xie 的研究也表明可用性是用户最为关注的问题,其余依次为藏品质量、服务质量、系统运行效率、用户意见征求^[13]。

一些学者跳出既有数字图书馆概念界定的藩篱,从结构—功能的视角构建数字图书馆的概念模型,并设计数字图书馆评估指标。DELOS 模型^[14]、5S 模型^[15]、DiLEO^[16]都是被用于数字图书馆评估的概念模型。DELOS 小组将数字图书馆各个利益相关群体关注的问题抽象出来,形成数字图书馆域,包含用户、数据/藏品、系统/技术、使用四个要素,基于此,DELOS 小组建立了一套整合性的评价指标体系^[14]。5S 模型以流—结构—空间—情境—社会(Streams-Structures-Spaces-Scenarios-Societies)为基础,提炼出一个抽象的数字图书馆四元组模型:仓储、元数据目录、服务(索引、搜索、浏览)、社会^[15],研究人员据此构建了数字图书馆质量指标体系^[15]和 5SQual 质量评估工具^[17]。DiLEO 是数字图书馆评估领域的本体,揭示了数字图书馆评估领域的概念及其关系^[16]。Tsakonas 基于该本体的研究表明数字图书馆评估活动仍是以系统为中心,研究聚焦于绩效测评、效果及技术卓越性三个方面^[3]。

1.2 数字图书馆评估指标

1.2.1 用户和系统之间的交互——可用性

可用性被广泛用于数字图书馆评估。由于

可用性测试发生在人工控制的环境中,结果不具有代表性^[18],需要通过结合真实任务^[19]、实时交互反馈^[20]等加以克服。此外,尽管存在大量的可用性标准,但必须对其进行调整以适应数字图书馆的用户群^[21]。目前,数字图书馆可用性评估应当包含哪些要素并无定论。Nielsen^[22]提出的可学习性、效率、可记忆性、出错率、满意度的使用率最高。除此以外,易用性^[23-24]、有效性^[7,12]、一致性^[12]也有较高的使用率。可用性的评估通常基于用户对可用性界面的感知,一些学者专门研究数字图书馆用户界面评价,如黄晓斌等从网站结构、页面表现、人机交互、网络性能四个方面构建了界面评价细则^[25]。

1.2.2 用户与内容之间的交互——有用性

有用性评估包含对内容本身的评估和对用户与内容交互的评估。内容的评估分为性质、结构和管理三个方面^[18],包括数字对象、元数据、藏品、目录、完备性等^[26]。其中,数字对象是唯一的数字图书馆资源类型^[12],采用数字图书馆特定指标,如保真度^[27]、可访问性、针对性等^[26]进行评估;其他的则采用传统指标测度,如准确性、易理解性、清晰度、可读性等^[12]。考察用户与内容交互的指标借鉴了大量用户研究领域的成果,如信息需求、满意度影响因素等^[18]。

1.2.3 系统与内容之间的交互——绩效

绩效评估集中在软硬件技术和服务的评估。例如软硬件都适用的易用性、响应时间、成本等^[12,28],专门针对硬件技术的指标,如可靠性^[24]、效果^[29]、效率^[30]等,软件指标如相关性^[12];服务的测评集中于完整度、可靠性、响应性、可达性、有用性几个方面^[12]。另外,数字图书馆咨询服务已形成一个专门的领域,如Lankes识别出数字图书馆参考咨询服务的六个指标:满意度、成本、礼仪、准确性、重复用户、意识^[31],White建立了包括使命和目的、结构和责任、核心功能、质量控制四个领域的数字参考咨询框架体系^[32]。

1.3 数字图书馆交互研究

用户与系统的交互是数字图书馆的重要内容^[2],然而,这种交互是复杂的,其复杂性是由用户、信息资源、信息对象、任务复杂度共同决定的^[18]。Fuhr构建了数字图书馆交互三要素模型(The Interaction Triptych Model),描述了用户、系统、内容之间的交互。内容与系统之间的交互表现为绩效因素,如查全率、查准率、响应时间等;用户与系统之间的交互表现为可用性因素,如效果、满意度等;用户与内容的交互表现为有用性^[18]。GROW数字图书馆项目以交互性作为其概念框架和设计理念,Budhu认为交互发生在信息资源、藏品、情境三个层次,相互性、反馈、及时性等因素是影响交互的核心因素^[5]。Zhang X等发现用户的交互受数字图书馆搜索和浏览交互设计的影响,搜索方面过多的零命中、出错反馈、不正确的搜索结果显示等增强了用户的挫败感,导致不良交互绩效^[7]。

在使用数字图书馆时,用户通常需要完成特定的任务。在任务情境下,数字图书馆仅是实现任务目标的工具,而个人与数字图书馆的交互活动仅是实现目标的必经过程^[33]。Marchionini指出数字图书馆的设计和评估需以用户为中心,其一便是评价用户的信息需求以及由此产生的任务^[34]。任务导向的数字图书馆评估,应当考察任务类型、任务阶段、任务绩效等之间的差异,对组织层级的冲击、信息获取系统的类型和进程活动的类型等因素^[18]。

基于任务的信息检索是情报科学的重要研究领域之一,Li Y和Belkin提出任务的分面分类框架^[35],为界定和测度与信息搜寻相关的任务提供了理论基础。基于此框架,他们探讨了工作任务和用户交互信息行为之间的关系^[36]。李月琳等提出的数字图书馆交互三维模型表明,数字图书馆以技术为支撑向用户提供信息资源,在使用数字图书馆完成目标任务的过程中,用户与数字图书馆在信息、技术和任务三个层面/维度进行交互,用户是交互过程的核心^[6]。基于该模型开展的实验研究表明,信息

维度的“信息充分性”,技术维度的“栏目组织合理性”“导航清晰性”“界面易用性”,任务维度的“获得任务所需信息的信心程度”能预测用户与数字图书馆的交互绩效^[6]。

文献综述表明,以往的研究关注数字图书馆的整体评估,学者们见仁见智,观点不一,然而,研究结果的多样化也造成了共识性缺乏的弊端。这种共识性的缺乏往往影响研究结果的推广应用,削弱了评估研究领域的社会效益。另一方面,指标体系的繁杂也使评估研究变得复杂,且耗时耗力,影响了评估的效率。因而,识别关键的评估指标,简化数字图书馆交互评估体系也是本研究的目的之一。

2 理论框架

本研究借鉴以上学者的成果,总结以往研究中具有一定共识性的评估指标,细化李月琳等提出的三维模型中的各交互维度,明确子维度(见表1),并以此作为本研究的理论框架,深入探索哪些因素影响用户对数字图书馆交互功能的评估,以揭示可预测用户对数字图书馆交互功能评估的因素,识别影响用户评估数字图书馆交互功能的主要指标,构建有效的数字图书馆交互功能评估模型。

表1 用户与数字图书馆交互维度、子维度及解释

交互维度	子维度	解释/内涵
信息维度	可获得性	用户通过一定的检索途径查询后能够找到所需信息,并且采用一定的方法可以将所需内容下载或复制下来。
	准确性	检索出的内容能够准确反映用户想了解的内容。
	有用性	所获得的信息是有用信息,若获得的信息对用户无用,那么获得的信息就是无效的。
	适用性	信息内容适用,即能够与用户实际的信息需求相匹配,所获信息与实际需求具有一致性。
	易理解性	信息内容易于理解,内容本身不繁琐、不复杂,更有益于用户对于信息内容的掌握和使用。
	保真度	信息内容清晰、不模糊,保持原本的状态。
	馆藏全面性	满足不同用户对于信息内容的不同需求,馆藏尽可能多地覆盖不同类型的信息内容,以保证不同用户在搜索不同类型的信息时均能获得。
	完整性	信息内容是完整的,而不是片段或是部分截取。
技术维度	网站结构	网站的构成状况,体现为网站组织有序,菜单树合理,导航功能完备,符合浏览规律,联系方式灵活多样。
	页面表现	页面的特征,体现为页面排版布局合理,内容简洁易懂,图像色彩搭配协调,视觉清晰度强,多媒体使用合理,链接丰富有效,清晰标识,主题突出。
	视觉美感	用户对界面的感知,包括:导航条是否美观,浏览工具是否容易找到,馆藏列表区分度是否明显,已访问的链接与底色反差是否明显,收集用户反馈的链接是否容易发现,网站界面的色彩是否阴暗,网站是否能吸引用户。
	可用性	界面的可用性,体现在易学性、易用性、可记忆性、一致性、高效性、直观性、容错和处理能力。
	用户体验满意度	用户的体验感知,体现在简单性、舒适性、可控性、可视化界面、友好程度。

续表

交互维度	子维度	解释/内涵
任务维度	客观任务复杂度	用户完成工作任务路径的多寡,越多则越复杂。
	相互独立性	用户执行任务是否需要与小组成员进行合作。
	主观任务复杂度	完全由用户自己决定所执行的工作任务对本人来说是否复杂。
	主观任务困难度	完全由用户自己决定所执行的工作任务对本人来说是否困难。
	搜索任务困难度	搜索相关信息时感知的困难程度。
	搜索任务复杂度	搜索相关信息时感知的复杂程度。
	主题熟悉程度	由用户对任务所具有的原有知识所决定,即用户是否熟悉该任务的主题。
	搜索经历的丰富程度	是否具有相同的搜索经历,并且可以解决搜索中遇到的问题。
	获取任务所需信息的信心程度	用户对于获得完成任务所需信息是否具有足够的信心。
	完成任务的过程和方法的熟悉程度	用户对搜索方法和搜索过程知识的熟悉程度。

3 研究方法

本研究采用实验研究方法,将卡片分类、设计模拟仿真工作任务、观察及出声思考等方式相结合,探讨用户在不同交互维度上的感知与他们对数字图书馆交互功能评估之间的关系。

3.1 变量与测度

“用户与数字图书馆在不同层面的交互”的测量指标及操作化定义见表1。实验中的问卷设计基于表1对各指标的界定,进一步细化。如“可用性”,进一步细化为“易学性”“易用性”“可记忆性”“一致性”“高效性”等。问卷中的题项通常以简洁易懂的表述呈现,并采用7点Likert量表进行测量。用户对“数字图书馆交互功能评估”的测量是在用户完成搜索后,通过综合评估“系统的搜索和浏览功能设计(即系统的交互功能设计)对其搜索过程的支持程度”来完成,也采用7点Likert量表来测量。

3.2 实验设计

3.2.1 实验系统

本次实验所选取的数字图书馆为中国知网(CNKI),选择实验参加者平常使用CNKI的学生宿舍作为实验场所。他们使用统一的笔记本电脑完成实验任务,笔记本电脑装有屏幕录像专家v2013,全程记录实验参加者与数字图书馆的交互过程。

3.2.2 实验任务

模拟仿真工作任务情境是交互信息检索实验中备受青睐的设计模式^[37],但也因模拟任务与真实任务存在偏差而遭到质疑^[38]。因此,本次实验设计了四个模拟仿真工作任务,同时要求实验参加者携带一个真实工作任务参与实验,以增强研究结果的可靠性。任务主题选自社会热点及学生关心的话题,初始共设计了12个主题,采用方便抽样选取10名来自不同专业的数字图书馆用户,要求其采用卡片排序(Card Sorting)方法,按照检索意愿进行卡片排序,最终选定排在前四位的4个主题:微信与微博营销模式、雾霾的成因与防治、房价居高不下的原因、

大学生就业困难的原因。之后,分别编制解决问题(Decision/Solution Task)和智识任务(Intellectual Task)两类^[35]模拟仿真工作任务情境。

为避免实验中的学习效应(Learning Effects),本研究采用4×4拉丁方块(Latin Square)实验设计,对模拟仿真工作任务进行排序。同时,将真实工作任务安排在四项模拟任务之前或之后,并依次轮换。根据实验设计,每一轮实验需招募8名实验参加者参与实验。

3.2.3 实验参加者的招募

为增强研究结果的可靠性,本研究共招募48名(6×8)实验参加者,根据实验设计,实验参

加者及任务顺序如表2所示。考虑到不同的学历层次对搜索的影响,实验招募本科生、硕士生、博士生各16名,招募渠道为某高校BBS。48名实验参加者分别来自33个不同专业,男女比例分别为39.6%和60.4%。采用7点Likert量表对实验参加者使用CNKI的基本经验进行调查,发现其中59.6%经常使用数字图书馆,40.4%能够熟练使用CNKI,74.5%认为在CNKI的检索比较成功。可见,实验参加者总体来说具有使用数字图书馆的经验,尤其是使用CNKI的经验。

表2 用户与数字图书馆交互实验设计

实验参加者(P)	任务(T)顺序	实验参加者(P)	任务(T)顺序
P1,9,17,25,33,41	T1,2,3,4,5	P5,13,21,29,37,45	T5,1,2,3,4
P2,10,18,26,34,42	T2,3,4,1,5	P6,14,22,30,38,46	T5,2,3,4,1
P3,11,19,27,35,43	T3,4,1,2,5	P7,15,23,31,39,47	T5,3,4,1,2
P4,12,20,28,36,44	T4,1,2,3,5	P8,16,24,32,40,48	T5,4,1,2,3

3.2.4 实验流程

首先,被试者阅读和签署承诺书,填写基本信息问卷。基本信息问卷收集用户的人口统计学变量信息及使用CNKI及其他数字图书馆相关经验的数据。然后,阅读任务1,完成搜索前问卷。该问卷收集与任务维相关的数据。随后在CNKI中进行检索,检索的过程中需要进行出声思考(Thinking Aloud),实验中提供出声思考指南。完成检索之后填写搜索后问卷,该问卷收集与信息维相关的数据。每一项任务都要填写搜索前问卷和搜索后问卷。在完成所有的模拟仿真工作任务之后,填写整体性评估问卷;真实任务完成之后,也需填写一份整体性评估问卷。整体性评估问卷着重收集技术维及用户评估交互功能的相关数据。最后,在实验完成之后,就实验内容对实验参加者进行简短的访谈并支付实验报酬。每次实验耗时约1.5小时,实验全程采用屏幕录像专家软件摄录。

3.3 数据分析

研究采用SPSS 19.0进行数据分析,通过Pearson分析和Stepwise回归分析探讨用户在不同维度的交互与用户对CNKI交互功能评估之间的关系。齐雪的分析表明,本研究中的模拟仿真工作任务和真实的工作任务并没有显著影响用户的交互绩效及评估结果^[39],因而,本研究在数据分析的过程中对两类任务未作区分。

4 研究结果

4.1 信息维对交互功能评估的影响

Pearson相关分析的结果见表3。信息维的可获得性、准确性、易理解性、保真度、馆藏全面性、完整性、有用性及适用性与用户对CNKI交互功能的评估结果显著相关。

表3 Pearson 相关分析结果(信息维)

维度	子维度	R
信息维	可获得性	$r(96) = 0.404^{**}$
	准确性	$r(96) = 0.307^{**}$
	易理解性	$r(96) = 0.371^{**}$
	保真度	$r(96) = 0.404^{**}$
	馆藏全面性	$r(96) = 0.453^{**}$
	完整性	$r(96) = 0.399^{**}$
	有用性	$r(96) = 0.564^{**}$
	适用性	$r(96) = 0.582^{**}$

注: ** $p < 0.01$

Stepwise 多元回归表明适用性、易理解性、准确性及信息的可获得性可以预测用户对系统交互功能的评估。表4和表5分别显示回归模型和回归系数。

模型4的R方为44.7%，调整R方为42.3%，其解释程度较为理想。分析表明，残差成正态分布，满足方差齐性假设；Tolerance指标及VIF值显示满足非共线性假设；Durbin-Watson值(1.621)显示满足相邻观察值的独立性假设。可以得知该回归模型并未违反回归分析应满足

的相关假设。因此，该模型是成立的，即信息维的适用性、易理解性、准确性及可获得性显著影响并可预测用户对系统交互功能的评估结果。T检验表明(见表4)，各预测变量的p值均显著($p < 0.05$)。通过表5可知，适用性、易理解性及可获得性均可正向预测用户对CNKI交互功能的评估，而准确性则负向预测了用户对交互功能评估的结果，关于这一研究发现，在5.1中有详细阐述。其中，适用性是最为重要的预测指标。

表4 信息维与交互功能评估的回归模型

模型	R	R方	调整R方	df	残差	F	Sig.	Durbin-Watson
1	0.582 ^a	0.339	0.332	94	64.502	48.270	0.000	1.621
2	0.614 ^b	0.377	0.364	93	60.792	28.173	0.000	
3	0.645 ^c	0.416	0.397	92	57.001	21.856	0.000	
4	0.669 ^d	0.447	0.423	91	53.971	18.401	0.000	

注: a. 预测变量: (常量), 适用性; b. 预测变量: (常量), 适用性, 易理解性; c. 预测变量: (常量), 适用性, 易理解性, 准确性; d. 预测变量: (常量), 适用性, 易理解性, 准确性, 可获得性。

表5 信息维与交互功能评估的回归系数

模型		非标准化系数		标准系数	T	Sig.	共线性统计量	
		B	标准误差	试用版			容差	VIF
1	(常量)	1.394	0.517	0.582	6.948	0.008	1.000	1.000
	适用性	0.641	0.092	2.695		0.000		
2	(常量)	0.599	0.605	0.517 0.206	0.991	0.324	0.898 0.898	1.114 1.114
	适用性	0.569	0.095		5.982	0.000		
	易理解性	0.228	0.096		2.382	0.019		
3	(常量)	0.533	0.589	0.665 0.318 -0.290	0.904	0.368	0.596 0.695 0.462	1.679 1.438 2.163
	适用性	0.732	0.114		6.440	0.000		
	易理解性	0.352	0.106		3.328	0.001		
	准确性	-0.296	0.119		-2.474	0.015		
4	(常量)	0.516	0.577	0.644 0.279 -0.454 0.268	0.894	0.374	0.590 0.672 0.330 0.432	1.694 1.487 3.029 2.317
	适用性	0.708	0.112		6.345	0.000		
	易理解性	0.309	0.105		2.934	0.004		
	准确性	-0.463	0.138		-3.345	0.001		
	可获得性	0.243	0.108		2.260	0.026		

4.2 任务维对交互功能评估的影响

Pearson 相关性分析表明,搜索任务困难程度($r(96) = -0.258, p < 0.05$)与 CNKI 对搜索和浏览功能的交互支持程度呈显著负相关。

Stepwise 多元回归分析表明,任务维的各子维度中,仅有用户感知的搜索任务困难程度显著影响并可预测交互功能评估结果(见表6和表7)。

表6 任务维与交互功能评估的回归模型

模型	R	R 方	调整 R 方	df	残差	F	Sig.	Durbin-Watson
1	0.258 ^a	0.067	0.057	94	91.103	6.729	0.011	1.516

注:a.预测变量:(常量),搜索任务困难程度。

表7 任务维与交互功能评估的回归系数

模型		非标准化系数		标准系数	T	Sig.
		B	标准误差	试用版		
1	(常量)	6.288	0.530	-0.258	11.856	0.000
	搜索信息的困难程度	-0.340	0.153		-2.594	0.011

表6显示,模型1的R方值为6.7%,调整R方仅为5.7%,该模型解释能力偏弱。通过残差的正态分布及方差齐性检验,以及Durbin-Watson值(1.516)可以看出,该模型符合回归分析应满足的假设,因此搜索任务困难程度可以

预测用户对数字图书馆交互功能的评估,且B值为-0.340(见表7),表明搜索任务的困难程度负向影响用户对交互功能的评估,即用户搜索信息难度越大,对CNKI的交互功能评价越低;越小,则越好。

4.3 技术维对交互功能评估的影响

Pearson 相关性分析结果显示(见表 8), 导航功能完备、导航条是否容易找到、浏览工具是否容易找到、馆藏列表区分度、符合浏览规律、

页面排版布局合理、主题突出、链接丰富有效、易学性、高效性、网站组织有序化及可控性, 与用户对数字图书馆交互功能的评估结果显著相关, 且均为正向相关。

表 8 Pearson 相关分析结果(技术维)

维度	子维度	R
技术维	导航功能完备	$r(96) = 0.492^{**}$
	导航条是否容易找到	$r(96) = 0.318^{**}$
	浏览工具是否容易找到	$r(96) = 0.421^{**}$
	馆藏列表区分度	$r(96) = 0.227^{**}$
	符合浏览规律	$r(96) = 0.417^{**}$
	页面排版布局合理	$r(96) = 0.443^{**}$
	主题突出	$r(96) = 0.450^{**}$
	链接丰富有效	$r(96) = 0.541^{**}$
	易学性	$r(96) = 0.394^{**}$
	高效性	$r(96) = 0.224^{*}$
	网站组织有序化	$r(96) = 0.533^{**}$
	可控性	$r(96) = 0.324^{**}$

注: ** $p < 0.01$; * $p < 0.05$

从表 9 可以看出, 链接丰富有效、网页组织有序化、页面排版布局合理均进入回归模型, 模型 1—3 均显著($p < 0.01$)。其中, 模型 3 的解释

力最强, 链接丰富有效、网页组织有序化及页面排版布局合理解释了变异量的 44.2% (R 方值), 调整 R 方值为 42.4%, 模型的拟合度较好。

表 9 技术维与交互功能评估的回归模型

模型	R	R 方	调整 R 方	df	残差	F	Sig.	Durbin-Watson
1	0.541 ^a	0.292	0.285	94	69.091	38.821	0.000	1.669
2	0.632 ^b	0.400	0.387	93	58.570	31.007	0.000	
3	0.665 ^c	0.442	0.424	92	54.437	24.330	0.000	

注: a. 预测变量: (常量), 链接丰富有效; b. 预测变量: (常量), 链接丰富有效, 网页组织有序化; c. 预测变量: (常量), 链接丰富有效, 网页组织有序化, 页面排版布局合理。

表 10 显示, 模型 3 中不同指标的容差值介于 0.677 至 0.804 之间, VIF 值未大于评鉴指标值 10, 表明进入回归方程的自变量间没有线性重合的问题, 残差符合正态分布及方差齐性假设。Durbin-Watson 值为 1.669, 该模型满足回归

分析的相关假设。模型 3 的预测变量中, 链接丰富有效最为重要。同时, 各预测变量 T 检验结果均显著($p < 0.05$), 这些指标正向影响并可预测用户对 CNKI 交互功能的评估结果。

表 10 技术维与交互功能评估的回归系数

模型		非标准化系数		标准系数	T	Sig.	共线性统计量	
		B	标准误差	试用版			容差	VIF
1	(常量)	1.757	0.518	0.541	3.392	0.001	1.000	1.000
	链接丰富有效	0.691	0.111		6.231	0.000		
2	(常量)	0.880	0.525	0.380 0.366	1.675	0.097	0.806 0.806	1.240 1.240
	链接丰富有效	0.486	0.114		4.246	0.000		
	网页组织有序化	0.367	0.090		4.087	0.000		
3	(常量)	-0.020	0.612	0.367 0.266 0.232	-0.032	0.974	0.804 0.677 0.789	1.244 1.477 1.268
	链接丰富有效	0.469	0.111		4.223	0.000		
	网页组织有序化	0.266	0.095		2.806	0.006		
	页面排版布局合理	0.293	0.111		2.643	0.010		

4.4 各维度对交互功能评估的综合影响

以上部分独立分析了不同维度与用户对数字图书馆交互功能评估之间的关系,本部分综合考量不同维度的子维度与用户对交互功能评估之间的关系。Stepwise 多元回归分析的结果如表 11 和表 12 所示。

结果显示,按照指标对于变异性解释能力

的大小,依次进入回归模型的指标分别是适用性、链接丰富有效、页面排版布局合理、主题突出、搜索任务的困难程度、馆藏列表区分度。表 11 显示,模型 1—6 均显著,模型 2—6 的调整 R 方均超过了 40%。其中模型 6 的解释能力最强,共解释了 62.5%的变异性,该模型的调整 R 方为 59.9%,说明模型的拟合度较优。

表 11 综合维度与交互功能评估的回归模型

模型	R	R 方	调整 R 方	df	残差	F	Sig.	Durbin-Watson
1	0.582 ^a	0.339	0.332	94	64.502	48.270	0.000	1.839
2	0.706 ^b	0.499	0.488	93	48.955	46.230	0.000	
3	0.748 ^c	0.560	0.545	92	43.004	38.952	0.000	
4	0.766 ^d	0.587	0.569	91	40.273	32.298	0.000	
5	0.778 ^e	0.605	0.583	90	38.554	27.579	0.000	
6	0.790 ^f	0.625	0.599	89	36.639	24.690	0.000	

注:a.预测变量:(常量),适用性;b.预测变量:(常量),适用性,链接丰富有效;c.预测变量:(常量),适用性,链接丰富有效,页面排版布局合理;d.预测变量:(常量),适用性,链接丰富有效,页面排版布局合理,主题突出;e.预测变量:(常量),适用性,链接丰富有效,页面排版布局合理,主题突出,搜索任务困难程度;f.预测变量:(常量),适用性,链接丰富有效,页面排版布局合理,主题突出,搜索任务困难程度,馆藏列表区分度。

表 12 综合维度与交互功能评估的回归系数

模型		非标准化系数		标准系数	T	Sig.	共线性统计量	
		B	标准误差	试用版			容差	VIF
1	(常量)	1.394	0.517	0.582	2.695	0.008	1.000	1.000
	适用性	0.641	0.092		6.948	0.000		
2	(常量)	-0.368	0.557	0.471 0.414	-0.661	0.510	0.928 0.928	1.077 1.077
	适用性	0.519	0.084		6.185	0.000		
	链接丰富有效	0.530	0.097		5.435	0.000		
3	(常量)	-1.442	0.605	0.428 0.362 0.258	-2.383	0.019	0.902 0.891 0.913	1.109 1.122 1.096
	适用性	0.471	0.080		5.868	0.000		
	链接丰富有效	0.463	0.094		4.943	0.000		
	页面排版布局合理	0.327	0.092		3.568	0.001		
4	(常量)	-1.831	0.609	0.414 0.301 0.224 0.186	-3.006	0.003	0.897 0.797 0.879 0.805	1.115 1.255 1.138 1.242
	适用性	0.456	0.078		5.829	0.000		
	链接丰富有效	0.385	0.096		3.992	0.000		
	页面排版布局合理	0.284	0.091		3.119	0.002		
	主题突出	0.211	0.085		2.484	0.015		
5	(常量)	-0.932	0.749	0.365 0.292 0.244 0.199 -0.142	-1.245	0.216	0.797 0.794 0.862 0.799 0.872	1.254 1.260 1.160 1.252 1.147
	适用性	0.402	0.082		4.920	0.000		
	链接丰富有效	0.374	0.095		3.930	0.000		
	页面排版布局合理	0.308	0.090		3.414	0.001		
	主题突出	0.226	0.084		2.692	0.008		
	搜索任务困难程度	-0.187	0.093		-2.004	0.048		
6	(常量)	0.859	0.735	0.343 0.363 0.273 0.255 -0.166 -0.181	-1.169	0.246	0.782 0.660 0.829 0.710 0.851 0.597	1.279 1.516 1.206 1.408 1.175 1.676
	适用性	0.378	0.081		4.672	0.000		
	链接丰富有效	0.464	0.102		4.541	0.000		
	页面排版布局合理	0.346	0.090		3.835	0.000		
	主题突出	0.289	0.087		3.308	0.001		
	搜索任务困难程度	-0.218	0.093		-2.352	0.021		
	馆藏列表区分度	-0.160	0.074		-2.156	0.034		

容差值及 VIF 值表明进入回归方程的自变量间没有共线性问题,残差符合正态分布及方差齐性假设, Durbin-Watson 值(1.839)表明不存

在自相关。由此可以看出,该模型并没有违反回归分析应满足的相关假设。从回归系数来看,各变量的 T 检验均显著($p < 0.05$),除搜索任

务困难程度及馆藏列表区分度外,其余各预测变量均能正向预测用户对 CNKI 交互功能的评估,而搜索任务困难程度及馆藏列表区分度则负向预测用户对 CNKI 交互功能的评估。在这些预测变量中,适用性是最为重要的预测变量。

数据分析表明,显著影响用户对数字图书馆交互功能评估的子维度可分为两类:一类是关联性子维度,即通过 Pearson 关联分析发现的与用户评估结果相关的子维度;另一类是预测

性子维度,即通过 Stepwise 回归分析明确的可预测用户评估数字图书馆交互功能结果的子维度。据此,本研究构建了基于多维度用户交互的数字图书馆交互功能评估理论模型(见图 1)。由于“准确性”与“馆藏列表区分度”的关联性分析结果和回归结果不一致,但又对用户数字图书馆交互功能的评估存在一定程度的影响,故仅将其归入“关联性子维度”。

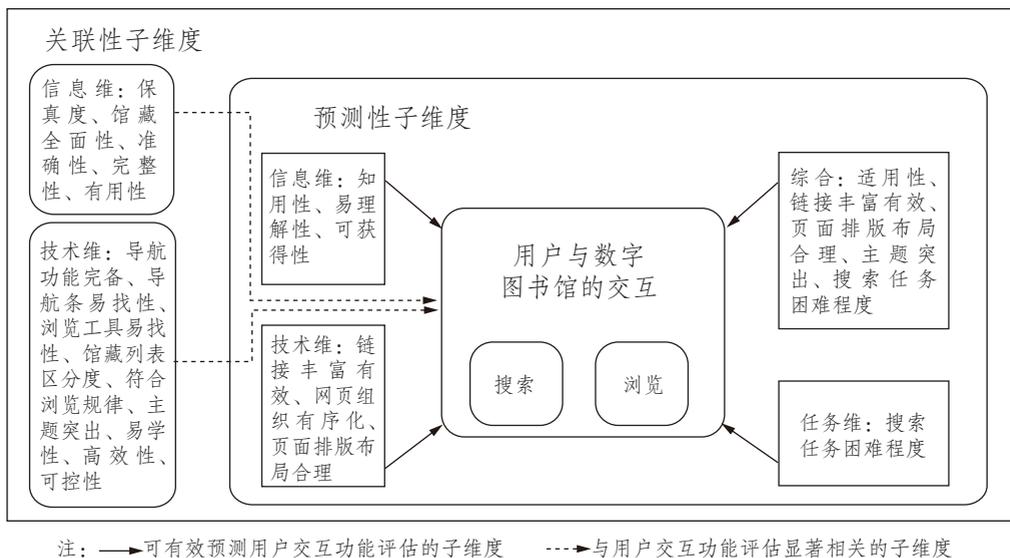


图 1 基于用户多维交互的数字图书馆交互功能评估理论模型

5 讨论与结论

本研究探讨交互维度与用户对数字图书馆交互功能评估之间的关系,构建了基于多维交互视角的数字图书馆交互功能评估理论模型。该模型从繁多的数字图书馆评估指标中明确了交互功能评估的关联性子维度和预测性子维度。前者集中在信息维和技术维,与用户对数字图书馆的交互功能评估结果显著正相关或负相关;后者则分布在不同维度,可预测用户评估数字图书馆交互功能的结果。本研究为评估数字图书馆的交互功能提供了简化的、可操作的理论模型。

相较于其他数字图书馆评估模型,如 DELOS 模型^[14]、5S 模型^[15]、DiLEO^[16] 这些更多针对数字图书馆综合评估的概念模型,本研究构建的模型更具针对性和可操作性,且聚焦于数字图书馆的交互功能。下面进一步讨论这些交互维度与用户评估数字图书馆交互功能之间的关系。

5.1 信息维、技术维与数字图书馆交互功能的评估

本研究的结果支持了 Xie^[13] 的研究发现,即用户重视数字图书馆的可用性。基于技术维度的回归分析表明,技术维度的回归模型能较好地预测用户对数字图书馆交互功能的评估。另外本研究也同时支持了 Zhang Y^[12] 的研究,表明

用户同样重视数字图书馆内容的有用性。在信息维度与技术维度上,回归模型均能有效预测用户对数字图书馆交互功能的评估。由此可见,对用户而言,数字图书馆内容层面的有用性和技术层面的可用性同等重要,信息维和技术维均是影响用户对其交互功能评估的重要方面。这一结果启示我们,数字图书馆具有良好的交互功能且要赢得用户,内容建设和可靠的技术支持缺一不可。

同时,本研究并未发现传统的可用性指标,如易用性、可记忆性、一致性等,显著影响用户对数字图书馆交互功能的评估。这一方面说明,尽管“可用性”的确是人机交互研究的重要方面,但是系统的“可用性”和“交互功能”存在概念上的区别,两者不能使用同样的指标来衡量和评估。另一方面也说明实验中所使用的CNKI已在一定的程度上满足传统的可用性指标,用户对这些指标的关注下降,而对其页面表现及网站结构的相关指标则更为关注。这一点从研究发现的关联性子维度和预测性子维度可见一斑,尤其是预测性子维度,基于技术维度的回归模型中明确的因素均与页面表现和网站结构有关,而在综合维度的回归模型中,与页面表现相关的因素也占了一半。可见,针对数字图书馆的交互研究应超越传统的可用性指标,更多关注页面表现与网站结构两个子维度,以更准确地评估数字图书馆系统的交互功能。

研究发现,信息的适用性和有用性与用户对数字图书馆的交互功能评估的相关性最强,因而,提供适用的信息,增强内容的有用性是数字图书馆内容建设的重要任务,也是提升数字图书馆交互功能的必经之路。回归分析表明,信息的适用性越强、越易于理解、可获得性越强,用户对数字图书馆交互功能的评估结果越好;适用性是最重要的预测变量。这启示我们,在数字图书馆资源建设过程中,更多考虑用户群体特点及其工作任务和情境是十分必要的,只有如此,才能提供满足用户需求的适用信息。

然而,研究发现,尽管信息的准确性与数字

图书馆的交互功能评估正向相关,但在回归模型中,准确性却是负向预测用户对数字图书馆交互功能的评估。导致这一结果的原因可能是由于模型中不同因素之间的交互,影响了准确性对数字图书馆交互功能评估的预测;此外,也可能存在一些潜在的、未被发现的因素或中介变量,影响了准确性与数字图书馆交互功能评估之间的关系。例如,用户如果在获得准确信息的过程中付出了过多的交互努力和承受了过重的认知负担,都可能影响用户对系统交互功能的评估,给数字图书馆带来负面的影响。再者,样本中的异常值也可能影响回归的结果。总之,就目前的研究结果,准确性不是一个很好的预测指标,值得进一步探讨。

综合各维度的回归分析发现,技术层面的页面表现是最为突出的子维度,其中的因素包括链接丰富有效、页面排版布局及主题突出,能有效预测用户对数字图书馆交互功能的评估。这一发现对数字图书馆的设计具有实践意义,即在开发和设计数字图书馆的过程中,开发和设计者应着重考虑这几项因素,从而有效提升数字图书馆的交互功能。在综合各维度的回归模型中,馆藏列表区分度负向预测数字图书馆的交互功能,与 Pearson 相关分析的结果相左,原因极有可能是中介变量的存在、不同因素之间的交互及异常值的存在。因而,馆藏列表区分度也不是一个很好的预测指标。

5.2 任务维与数字图书馆交互功能的评估

本研究将任务作为用户的交互维度之一加以考察,试图探究其对数字图书馆评估的影响,尤其是对交互功能评估的影响。相较于数字图书馆的内容和技术支持,任务对用户与数字图书馆的交互绩效影响较小。虽然在此前的研究中,也发现任务与交互绩效相关,但仅有“获得任务所需相关信息的信心程度”能有效预测用户的交互绩效^[6],本研究印证了这一结果。在数字图书馆的交互功能评估中,相比于信息维和技术维,任务维对预测用户对数字图书馆交

互功能的评估表现偏弱。虽考察了任务的诸多维度,但仅有“搜索任务困难度”与数字图书馆的交互功能显著相关,并能有效预测用户对数字图书馆交互功能的评估结果。尽管调整R方偏低,但从不同变量的残差分析来看,回归模型合理,模型F值显著,表明该要素是较好的预测指标。这一结果启示我们,该子维度在开发和设计数字图书馆交互功能时需引起重视,即应增强系统对不同困难程度任务的支持,尤其是高难度搜索任务的支持;适应并支持不同困难程度的搜索任务,可提升数字图书馆的个性化性能。从这一角度看,本研究对推进基于任务的个性化信息检索的研究具有一定意义。

更为重要的是,由于任务只在有限的方面影响用户与数字图书馆的交互绩效及对数字图书馆的评估,因而,在数字图书馆评估及交互信息检索系统评估过程中,我们确实可以摒弃评估实验中采用模拟仿真工作任务情境还是真实工作任务情境的争执^[6],采用任何一类任务或者是混合两类的任务类型。取而代之的应是重视“搜索任务困难度”这一因素,即在评估研究中,尤其是实验研究中,任务设计要考虑不同的工作任务是否引发不同困难程度的搜索任务,如果忽略了这一因素,采用困难程度相当的任务设计,评估的结果是不全面且存在偏差的。因而,本研究的发现为更科学地设计评估实验中的任务类型提供了实证依据。

5.3 研究启示、局限性与未来的研究

本研究基于以往的数字图书馆评估研究,从繁多的评估指标中析出与用户对数字图书馆交互功能评估相关的关联性子维度及预测性子

维度,明确了影响用户评估数字图书馆交互功能的重要子维度。这些子维度同时也是数字图书馆在满足基本的可用性要求之后,在设计交互功能时尤其需要重视的要素。本研究所构建的数字图书馆交互评估理论模型简化了数字图书馆交互功能的评估,模型所确定的关联性子维度和预测性子维度可为制定数字图书馆交互功能评估标准和开发数字图书馆交互功能评估工具提供参考。此外,本研究为数字图书馆交互功能的开发和设计提供了实证依据,明确了开发过程中应重点关注的因素,尤其是应重视预测性子维度所昭示的方面,为提升数字图书馆的交互功能提供了明确的方向。同时,本研究将任务这一情境要素融入数字图书馆评估研究中,不仅发现用户感知的搜索任务困难程度显著影响其对数字图书馆交互功能的评估,还为信息检索系统评估研究中如何更科学有效地设计任务提供了借鉴。

本研究具有一定的局限性,数据来源仅取自于实验过程中收集的问卷数据,并未分析视频资料。此外,由于要求每位实验参加者带来一项真实工作任务,这些任务的特性表现不一,而为数据处理的方便,本研究把这48项真实任务当作同一任务(即T5)来看待,这在一定程度上可能会影响研究结果。本研究中数字图书馆交互功能的评估仅以用户的感知作为测评的依据,未来的研究将进一步从系统和用户角度完善交互功能评估的各项指标,继续探讨用户与数字图书馆的交互,分析实验的视频资料,探究用户与数字图书馆交互的关键成功因素,为更精确地评估数字图书馆的交互功能,提升数字图书馆交互功能的设计提供理论支撑和实证基础。

参考文献

- [1] Saracevic T. Digital library evaluation: toward an evolution of concepts [J]. Library Trends, 2000, 49(2): 350-369.
- [2] Saracevic T. The stratified model of information retrieval interaction: extension and applications [J]. Proceedings of the American Society for Information Science, 1997(34): 313-327.
- [3] Tsakonas G, Mitrelis A, Papachristopoulos L, et al. An exploration of the digital library evaluation literature based

- on an ontological representation [J]. *Journal of the American Society for Information Science and Technology*, 2013, 64(9): 1914-1926.
- [4] Arms C R. Some observations on metadata and digital libraries [C/OL] // *Bicentennial Conference on Bibliographic Control for the New Millennium*, Washington, DC, November 15-17, 2000; *Confronting the Challenges of Networked Resources and the Web* [2015-06-23]. http://www.loc.gov/catdir/bibcontrol/arms_paper.html.
- [5] Budhu M, Coleman A. The design and evaluation of interactivities in a digital library [J/OL]. *D-Lib Magazine*, 2002, 8(11) [2015-06-04]. <http://www.dlib.org/dlib/november02/coleman/11coleman.html>.
- [6] 李月琳,肖雪,仝晓云. 数字图书馆中人机交互维度与用户交互绩效的关系研究[J]. *图书情报工作*, 2014(2): 38-46, 120. (Li Yuelin, Xiao Xue, Tong Xiaoyun. Relationships between dimensions of human computer interaction and users' interaction performance in digital libraries [J]. *Library and Information Service*, 2014(2): 38-46, 120.)
- [7] Zhang X M, Li Y L, Liu J J, et al. Effects of interaction design in digital libraries on user interactions[J]. *Journal of Documentation*, 2008, 64(3): 438 - 463.
- [8] Kyrillidou M, Giersch S. Developing the DigiQUAL protocol for digital library evaluation [C] // *ACM, Proceedings of the 5th ACM/IEEE-CS Joint Conference*. New York, 2005: 172-173.
- [9] Blixrud J C. Measures for electronic use: the ARL E-Metrics Project [C] // *Statistics in Practice- Measuring and Managing 2002*: 73-84.
- [10] 中华人民共和国文化部. 图书馆数字资源统计规范 (WH/T 47-2012) [S]. 北京: 国家图书馆出版社, 2013. (Ministry of Culture of the People's Republic of China. *Library Digital Resources Statistics (WH/T 47-2012)* [S]. Beijing: National Library of China Publishing House, 2013.)
- [11] 肖珑,李浩凌,徐成. CALIS 数字资源评估指标体系及其应用指南 [J]. *大学图书馆学报*, 2008(3): 2-8, 17. (Xiao Long, Li Haoling, Xu Cheng. Indicator system and its application guidelines for the evaluation of CALIS digital resources [J]. *Journal of Academic Libraries*, 2008(3): 2-8, 17.)
- [12] Zhang Y. Developing a holistic model for digital library evaluation [J]. *Journal of the American Society for Information Science and Technology*, 2010, 61(1): 88-110.
- [13] Xie Hong (Iris). Evaluation of digital libraries: criteria and problems from users' perspectives [J]. *Library and Information Science Research*, 2006(28): 433-452.
- [14] Fuhr N, Hansen P, Mabe M, et al. Digital libraries: a generic classification and evaluation scheme [C] // *Proceedings of the 5th European Conference on Research and Advanced Technology for Digital Libraries*. London, UK: Springer-Verlag, 2001: 187-199.
- [15] Goncalves M A, Fox E A, Watson L T, et al. Streams, structures, spaces, scenarios, societies (5S): a formal model for digital libraries[J]. *ACM Transactions on Information Systems*, 2004, 22(2): 270-312.
- [16] Tsakonas G, Papatheodorou C. An ontological representation of the digital library evaluation domain [J]. *Journal of the American Society for Information Science and Technology*, 2011, 62(8): 1577-1593.
- [17] Moreira B L, Gonçalves M A, Alberto H F, et al. Evaluating digital libraries with 5SQual [C] // *The 11th European Conference on Research and Advanced Technology for Digital Libraries*, 2007: 466-470.
- [18] Fuhr N, Tsakonas G, Aalberg T, et al. Evaluation of digital libraries [J]. *International Journal on Digital Libraries*, 2007, 8(1): 21-38.
- [19] Blandford A, Stelmaszewska H, Bryann - Kinns N. Use of multiple digital libraries: a case study [C] // *Proceedings of the first ACM/IEEE-CS joint conference on digital libraries*, 2001: 179-188.
- [20] Castillo J, Hartson H R, Hix, D. Remote usability evaluation: can users report their own critical incidents [C] // *CHI 1998 Conference Summary on Human Factors in Computing Systems*, 1998: 253-254.
- [21] Borgman C L, Leazer G H, Gilliland-Swetland A L, et al. Iterative design and evaluation of a geographic digital library for university students: a case study of the Alexandria Digital Earth Prototype (ADEPT) [G] // *Lecture Notes in Computer Science: Research and Advanced Technology for Digital Libraries*, 2001, 2163: 390-401.

- [22] Nielsen J. 可用性工程[M]. 刘正捷,等,译. 北京:机械工业出版社,2004:17-24. (Nielsen J. Usability engineering [M]. Liu Zhengjie, et al, trans. Beijing: China Machine Press, 2004: 17-24.)
- [23] Hill L L, Carver L, Larsgaard M, et al. Alexandria digital library: user evaluation studies and system design [J]. Journal of the American Society for Information Science and Technology, 2000, 51(3): 246-259.
- [24] Champeny L, Borgman C L, Leazer G H, et al. Developing a digital learning environment: an evaluation of design and implementation processes [C] // Proceedings of the 2004 Joint ACM/IEEE Conference on Digital Libraries, 2004: 37-46.
- [25] 黄晓斌, 卢琰. 论数字图书馆用户界面的评价[J]. 图书馆论坛, 2005(3): 16-19. (Huang Xiaobin, Lu Yan. On the evaluation of the digital library user interface [J]. Library Tribune, 2005(3): 16-19.)
- [26] Goncalves M A, Moreira B L, Fox E A, et al. What is a good digital library? A quality model for digital libraries [J]. Information Processing and Management, 2007(43): 1416-1437.
- [27] Kenney A R, Sharpe L H, Berger B. Illustrated book study: digital conversion requirements of printed illustration [C] // Proceedings of the Second European Conference: Research and Advanced Technology for Digital Libraries, 1998: 279-293.
- [28] Tsakonas G, Papatheodorou C. Exploring usefulness and usability in the evaluation of open access digital libraries [J]. Information Processing and Management, 2008(44): 1234-1250.
- [29] Jones S, Paynter G W. Automatic extraction of document key phrases for use in digital libraries: evaluation and application [J]. Journal of the American Society for Information Science and Technology, 2002, 53(8): 653-677.
- [30] Bosman F J M, Bruza P D, Van de Weide T P, et al. Documentation, cataloging, and query by navigation: a practical and sound approach [C] // Proceedings of the Second European Conference: Research and Advanced Technology for Digital Libraries, 1998: 459-478.
- [31] Lankes R D, Gross M, McClure C R. Cost, statistics, measures, and standards for digital reference services: a preliminary view [J]. Library Trends, 2003, 51(3): 401-413.
- [32] White M D. Digital reference services: framework for analysis and evaluation [J]. Library and Information Science Research, 2001, 23(3): 211-231.
- [33] Cousins S B. A task-oriented interface to a digital library [C] // CHI'96 Companion, Vancouver, BC Canada, 1996: 103-104.
- [34] Marchionini G, Plaisant C, Komlod A. The people in digital libraries: multifaceted approaches to assessing needs and impact [M] // Digital Library Use: Social Practice in Design and Evaluation. Cambridge: MIT, 2003: 119-160.
- [35] Li Y, Belkin N J. A faceted approach to conceptualizing task in information seeking [J]. Information Processing and Management, 2008, 44(6): 1822-1837.
- [36] Li Y, Belkin N J. An exploration of the relationships between work task and interactive information search behavior [J]. Journal of the American Society for Information Science and Technology, 2010, 61(9): 1771-1789.
- [37] Borlund P, Ingwersen P. The development of a method for the evaluation of interactive information retrieval systems [J]. Journal of Documentation, 1997, 53(3): 225-250.
- [38] Blomgren L, Vallo H, Bystrom K. Evaluation of an information system in an information seeking process [C] // Proceedings of ECDL. Heidelberg: Springer-Verlag Berlin, 2004: 57-68.
- [39] 齐雪. 数字图书馆交互功能评估研究[D]. 天津:南开大学, 2014. (Qi Xue. Developing an evaluation model for interaction in digital libraries [D]. Tianjin: Nankai University, 2014.)

李月琳 南开大学商学院信息资源管理系教授, 博士生导师。天津 300071。

梁娜 南开大学商学院信息资源管理系博士研究生。天津 300071。

齐雪 天津联怡科技有限公司中级工程师。天津 300384。

(收稿日期: 2015-09-22; 修回日期: 2015-10-09)