

开源机构库软件可用性评估方法的探讨

洪 梅 马建霞

(中国科学院国家科学图书馆兰州分馆 兰州 730000)

【摘要】 总结可用性内涵,简单介绍国外有关开源机构库软件可用性评估的案例,根据机构库建设的目的和意义,提出一套可用性运用于开源机构库软件评估的方法。

【关键词】 可用性评估 开源机构库软件 **【分类号】** G250.7

Usability Evaluation on Open – source Institutional Repository – building Software

Hong Mei Ma Jianxia

(Lanzhou Branch of National Science Library, CAS, Lanzhou 730000, China)

【Abstract】 This paper summarizes the definition of usability, and briefly introduces some foreign cases of usability evaluation on open – source institutional repository – building software. Then the usability method is proposed to evaluate open – source software based on the purpose and significance of building institutional repository.

【Keywords】 Usability evaluation Open – source institutional repository – building software

1 引言

近年来,国外机构库蓬勃发展,在国内,上海、香港和台湾等地举办的学术会议也开始热烈讨论机构库,多家高校和科研机构,如中国科学院国家科学图书馆、清华大学、北京大学、厦门大学、浙江大学等,都正在尝试着构建机构库。由于开源软件比商业软件具有成本低、质量好、透明度高、能更好满足机构需求等优点,因此机构库的建设多采用开源软件。

在机构库建设之初,需要对多种开源软件进行评估,以找出更适合本机构的软件。可用性作为系统的一项重要质量指标,是为了解决用户与设备、软件、工具等人工开发系统之间的关系而提出来的,涉及到用户与系统交互的许多方面。对于系统的用户而言,可用性关系到系统的性能是否满足用户的需要,流程是否符合用户的习惯,效果是否达到用户的期望;对于系统的工作人员而言,可用性关系到工作的效率和系统存在的意义;对于开发者而言,可用性直接决定着系统开发的成败。因此对机构库的各种开源软件的可用性进行研究和评估,对改进和完善机构库的功能,提高机构库的质量,激励用户更多更好地使用机构库,提高机构库的影响力,促进机构库

健康持续地发展,都具有极为重要的意义。

2 可用性的内涵

可用性研究的先驱和领导者雅各布·尼尔森(Jakob Nielsen)认为,可用性包括以下要素^[1]:

- (1)易学性:产品是否易于学习;
- (2)交互效率:即用户使用产品完成具体任务的效率;
- (3)易记性:用户搁置某产品一段时间后是否仍然记得如何操作;
- (4)出错频率和严重性:操作错误出现频率的高低及其严重程度如何;
- (5)用户满意度:用户对产品是否满意?

尼尔森认为,产品在每个要素上都达到很好的水平,才具有高可用性。

国际标准化组织针对可用性作出了范围更加广泛而确切的定义。其中引用最多的是“ISO 9241 – 11 即对带有显示终端(VDTs)的办公室工作的人机工程需求——第十一部分:可用性指导”^[2],可用性是指当用户在特定的环境中使用产品完成具体任务时,交互过程的有效性(Effectiveness)、交互效率(Efficiency)和用户满意度(Satisfaction)。其中,有效性是指用户完成特定任务时所达到的准确程度(Accuracy)和完整程度(Completeness),效率是指用户完成特定任务的准确程度和完整程度与所用

收稿日期:2007 – 10 – 15

收修改稿日期:2007 – 10 – 26

资源(如时间、精力)之间的比率,满意度指用户在使用产品过程中具有的主观满意和接受程度。

该定义中可用性的3个方面紧密结合相互补充,全面而简洁地回答了用户能否快捷而有效地实现其目的或完成任务的情况;在使用该产品过程中相关资源耗费的情况;总体感觉情况。这和尼尔森认为的5个要素也是相辅相成的。

通过以上分析,用图1表示可用性的几个维度如下:

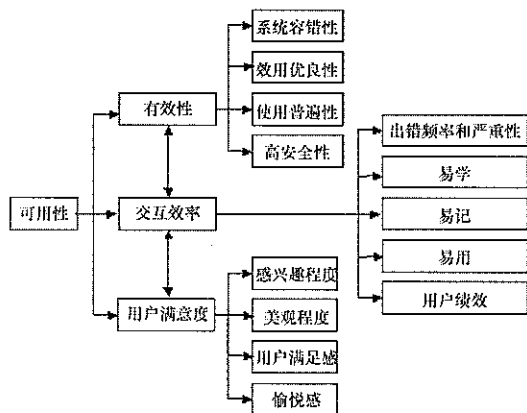


图1 可用性的维度

可用性的概念与通常所说的方便使用这一概念是不相同的。对一个产品而言,它没有固有的可用性属性,只有特定人员在特定背景中使用才有性能,评估可用性也不能离开人员孤立进行研究。

就建设机构库的开源软件而言,其涉及到的人员可以分为两个层次:一是使用开源软件进行机构库建设的技术人员;二是使用构建好的机构库平台的终端用户。对技术人员而言,机构库开源软件的可用性就是开源软件本身是否有利于更加方便有效地搭建出一个满意的服务平台,以满足机构需求,鼓励终端用户更好地使用,保证机构库的长期运营,更多地体现在可用性的有效性这个维度上。而交互效率和用户满意度这两个维度则是通过终端用户在使用构建好的机构库平台,完成登录、自存档、检索、浏览、获取等操作的过程中,客观的效率和主观的感受来表现的。

3 国外开源机构库软件可用性评估的案例简介

3.1 密歇根大学(University of Michigan)^[3]

在正式部署机构库之前,密歇根大学邀请信息学院(School of Information)可用性测试班的学生担当测试员,综合运用启发式评估,词汇分析,对用户直接测试等一系列方法,对DSpace的改进版本Deep Blue的可用性进行了三轮测试。第一轮测试的对象是从图书馆人员中选出

的专家,他们很熟悉机构库、元数据、自存档、搜索、浏览等概念。第二轮测试的对象是大学生和研究生,测试的内容是检索Deep Blue中存放的文献。第三轮测试的对象是机构人员,测试的内容是提交文献的界面。后两轮测试同时进行。测试结果表明,Deep Blue版本建立的机构库在搜索、浏览、提交、界面等方面都表现出很好的可用性,所有的测试对象都完成了任务,只需要消除一些细节问题,从而增强已有系统功能。

3.2 澳大利亚国立大学(Australian National University)^[4]

为了测试DSpace和Eprints这两个最常见的开源机构库软件的用户界面可用性,澳大利亚国立大学同时运用这两者部署了机构库,且保证其中具有相同的文献。在该研究中,首先制定了一系列任务,运用启发式评估,得出假定:用户在使用DSpace时,完成任务需要花的时间更少,出错率更低,用户满意度更高。再选取了18个大学生作为测试人员,分为两组,每组9人,分别就DSpace和Eprints对用户进行可用性测试,其结果很好地证实了先前的假定。从而提出了7个评估用户界面的指导方针,在构建机构库时,用这些指导方针可以有效地提高用户界面可用性。

3.3 俄勒冈州立大学(Oregon State University)^[5]

在该大学正式部署DSpace作为机构库之前,考虑到DSpace提交过程步骤繁多,在本地化提交界面的帮助文档后,基于可用性研究理论,对DSpace提交电子论的过程进行了可用性测试。具体操作是选择不同领域和年龄的本科生和研究生作为测试人员,要求他们完成提交“PDF”格式的电子论文。通过分析他们的行为,得出结论:虽然DSpace的提交界面还存在一些小问题,但是无论是本科生还是研究生,都可以由校园认证系统登录,在本地化后的帮助文档的指导下,很好地完成了提交任务;可用性测试既简单又实用。

3.4 怀卡托大学(University of Waikato)^[6]

该研究是在新西兰怀卡托大学的可用性实验室中,要求23个四、五年级的计算机专业的学生对开源软件Greenstone进行可用性测试,他们均能熟练地在Greenstone中进行检索。具体任务是借助3种Greenstone手册(安装手册、用户手册、开发者手册),用提供的两种方式创建专题栏目,其操作过程由录像录制下来。然后以问卷形式了解他们操作过程中的实际感受。通过分析录像和问卷,反映出软件和手册之间存在的 inconsistency 问题,也折射出开源软件的开发影响到软件产品的可用性,从4个角度分析了出现可用性问题的根本性原因,提出还是要

从实际用户的角度看待可用性问题的。

3.5 其他

罗彻斯特大学(University of Rochester)^[7]通过运用网站的标准和美工设计员进行可用性测试,把默认的DSpace的界面进行了重新设计。

康奈尔大学(Cornell University)^[8]在选择机构库模式时,把对人员和终端用户的可用性测试作为选择的指标之一。

2006年,在世界三大国际数字图书馆会议之一的JC-DL会^[9]上,指南一“个性化DSpace”提出需要改进DSpace以增强提交者和用户的可用性;2007年,该会议在指南一中继续并深化了“个性化DSpace”,仍提及DSpace的可用性问题^[10]。

4 可用性运用于开源机构库软件评估的方法

可用性评估的主要目的是通过使用评估方法对产品、系统等评估对象的可用性进行定性和定量分析。可用性评估的方法有多种,每种可用性评估方法都各有优缺点,可分别使用,也可结合使用。

对于开源机构库软件的可用性评估,按照有无用户参与可分为两类:没有用户参与的可用性评估和有用户参与的可用性评估。前者是后者的前序,具体又分为3种方法:比较评价、认知走查法、启发式评估;后者是前者的进一步深入。

4.1 没有用户参与的可用性评估

开展用户参与的可用性评估需要相应的成本。通过以下3个简单的方法,在没有用户参与的情况下评估开源软件,作为用户参与评估的前序,获得相关信息,以事先确定用户参与评估中需要调查什么,提高研究效率,节省时间和金钱。

(1) 比较评价

在机构库多个开源软件中,选择具有一定影响力的开源软件,如由MIT和惠普公司合作设计开发的DSpace、由南安普顿大学开发的Eprints、由弗吉尼亚大学和康奈尔大学共同开发的Fedora(Flexible Extensible Digital Object and Repository Architecture,基于灵活、可扩展的数字对象和知识库构建)、由新西兰的怀卡托大学的新西兰数字图书馆工程(New Zealand Digital Library Project)开发的Greenstone等,都是可用于建设机构库的开源软件的代表。由于各种开源软件在设计理念、体系结构、采用标准、实现技术、系统规模以及数据的组织方式、发布方式等很多方面都不尽相同,在实际选择软件时,可以从以下几个方面综合对比:

①体系结构及其数据模型;

②技术细节:支持OAI-PMH的版本;开放源代码协议;最新的发

布时间、版本号;支持的操作系统;编程语言;所需的数据库、服务器;安装的数量规模等;

③内容管理:对象和元数据的各种操作、内容的可移植性、文件格式的识别等;

④用户界面:可否修改界面外观;是否支持多语言界面;终端用户是否有自己的文件夹;可否自定义页面页脚;是否有相关论坛等;

⑤搜索功能:是否可以按作者、题名、主题词、发布日期或其他来进行浏览、检索;是否可以按相关度分类搜索的结果等;

⑥存档管理:系统是否分配标识符、CNRI手柄;是否定义数字保存战略;是否检查数据完整性等。

(2) 认知走查法(Cognitive Walkthrough, CW)

认知走查法是一种常用的可用性评估方法。应用于机构库的可用性评估同样是由一个或一组评估专家,根据一系列既定任务来评测用户界面并评估其易学易记易用程度。该方法一般来说适用于开发设计的初期,当然也可以用于测试运行阶段。对开源机构库软件进行认知走查评估可用性时,评估专家可以是技术开发者、一般的图书馆员、界面工程师等等。评估对象可以使用开源软件本身的系统界面。

具体操作的主要步骤是:让评估专家把自己想象为使用开源软件的典型终端用户,以一个局外人的角度浏览整个机构库并进行观察。首先,选择典型的界面任务,并为每一任务确定一个或多个正确的操作序列(Sequence of Actions),然后构建一个“可信的故事(Credible Story)”来解释为什么用户会选择正确的操作步骤完成任务,或者构建一个“失败的故事”解释为什么用户没有按照正确的操作步骤进行下去,从而列出用户在完成任务的过程中在什么方面会出现问题并提供解释。

构建“故事”的根据是在走查过程中对以下4个问题的回答^[11]:

①用户会努力获得正确的结果吗?这一点所关注的问题是用户的意图如何,思路是否清晰,如果用户不试图获求正确操作所要求的步骤,那么他们不可能选择正确的操作。

②用户会注意到正确的操作步骤或方式吗?对于直接的显而易见的操作如菜单和按钮,一般来说,用户使用起来不会有什么问题,但是像双击这样的操作,用户就不容易觉察到。

③用户会把正确的操作步骤与努力获得的结果联系起来吗?即便用户努力获求正确的结果,也能看到可用的正确操作,但是如果用户不知道在系统界面的何处可以进行该操作,依然会使任务的完成受到影响。

④在正确的操作之后,用户是否能得到合适的反馈?这一问题关注的是界面能否提供正确的及时的反馈,由于没有得到反馈,用户可能会取消正确的操作而去进行他们所认为的更为“合适的”操作。

(3) 启发式评估(Heuristic Evaluation, HE)

由Nielsen和Molich在1990年提出的启发式评估^[12]是最早使用的经典可用性评估方法之一。这是一种专家评审法,其一般思想是评价者们根据通用的可用性原则和经验,分别进行独立评判来找出系统潜在的可用性问题,从而构建一

个对系统的评价方案并试图找出具体的解决方案。

启发式评估应用于机构库的可用性评估时,最好是在其设计阶段,通过对开源软件默认系统界面进行评估,更容易改进大量出现的可用性问题,以不断完善机构库的各项功能和服务。

专业的评价者和公认的可用性原则是启发式评估中不可缺少的两大成分:

①关于专业的评价者人数的确定,根据 Nielsen 的实证研究^[13],人数越多,发现的可用性问题越多,一个评估人员可发现 35% 的可用性问题,5 个可发现大约 75% 的可用性问题。但是在超过 5 人以后,随着人数的增加,投入收益比率降低;

②Nielsen 提出了 10 个启发式规则是评估中最常用的可用性原则^[14]:系统状态的可视性;系统与用户的认知习惯相符合;用户控制和自由使用系统;系统设计的一致性和标准化;预防错误发生;依赖识别而非记忆;使用的灵活性及有效性;美观精炼的设计;帮助用户识别、诊断和从错误中恢复;帮助文件的设置。

因此,在进行开源软件的可用性评估时,只需要用 3-5 个评估者。评估者根据可用性原则或者自己的经验、习惯浏览界面至少两次,一次集中于运用开源软件的整个流程,另一次集中于个别的对话提示成分,独立对系统界面和功能进行评估。通过评估者单独完成书面报告或者评估者在评估过程中作出声报告,观察员辅以现场记录,来对评价结果进行记录和整理,最后将评估出来的可用性问题进行汇总,共同讨论,对出现的问题情况进行排序,并提出相应的建议。

4.2 有用户参与的可用性评估

通过没有用户参与的可用性评估,发现了关于开源软件运用于机构库还存在的许多可用性问题,但可能有一些问题不很确定,这就需要用户参与,进行进一步评估。有用户参与的可用性评估,对于了解真实用户如何使用系统有着无法替代的作用,由于用户真正地去执行实际的任务,因此评估能够更好地揭示与可用性相关的实际问题,发掘与设计人员预期的使用方式异同的地方,这些都是没有用户参与的可用性评估不能发现的问题。

由于评估对象是开源软件,主要评估的是用开源软件建立机构库中存在的可用性问题,而建立机构库的目的在于通过收集、组织、存储、管理本机构成员在工作过程中创建的智力成果,提供开放访问,促进平等使用和共享,以提高学术成果的可见度及学术机构的声望和影响,在实施可用性评估的各个环节时,都要围绕这个目的进行。

(1) 选择有代表性的用户

选择一定数量的本机构成员作为评估者,在具体选择时,要充分考虑到人员的不同背景、不同专业方向,对计算机的操作熟练与否,英语水平程度,合理分布各个层次的人员,使评

估者更具有代表性。

(2) 设计测试任务

设计测试任务的基本原则是所选择的任务尽可能地代表真正用户的使用。根据机构实际建立机构库的目的,围绕终端用户可进行的登录、自存档、检索、浏览、获取等操作来设计用户任务,对各个开源软件的用户界面、交互、功能等方面进行评估。

在具体实践时,有 3 种方式来设计测试任务:

①事先分配一个指定的任务:对任务作出细致而明确的表述,如“请在机构库中检索并下载文献 ABC”;

②动态地选择或修改任务:事先准备一系列任务,确定复杂程度递增的问题,在测试中可跳过任务或随时增加任务以更好地发现可用性问题;

③不设计任务:构建一个真实的用户环境,让用户实际使用。

(3) 确定测试地点

①可用性工程实验室:实验室的设置可以使测试准备更容易,并提供了理想的记录设备;但是可能使评估者感到压抑;而且很多地方可能没有这样的实验室

②办公室:进行测试准备比较容易,但是办公室可能不适合用作可用性实验室,需要做很多重新布置的工作。

③用户工作地点:在用户熟悉的环境中进行测试,可以更真实地体现用户使用操作的实际情况,但是可能出现许多和开源软件可用性无关的干扰因素,如网速缓慢,电脑内存过小等等。

在进行开源软件的可用性测试时,可以选择办公室,将其布置成为一个简易的可用性工程实验室,既和用户办公环境相似,使用户不会觉得紧张,又可在相同条件下测试各个开源软件,避免一些主观的条件影响开源软件的使用。

(4) 记录数据

在测试过程中,有多种方法用来记录测试数据^[15]:

①直接观察法:直接观察用户的任务执行过程,并及时记录下其行为和任务完成情况。

②边想边说法(Think Aloud):在任务执行过程中,让用户随时大声说出操作思路的一种方法。

③录音录像记录法:将用户的任务完成过程录制下来以备回顾和作进一步分析。

④问卷法:在测试结束后,以问卷的形式量化用户满意度,了解其主观感受。

根据实际的条件,尽量使用多种方法,以保证数据更加全面真实。

(5) 分析结果

根据记录的测试数据,一般来说,包括定量数据如任务完成时间等和定性数据如操作过程等,综合两类结果进行分析对比,以了解各个开源软件在导航、标识、交互、搜索、内容组织等可用性方面的优劣势。

5 结 语

前期研究越多意味着后续工作越少。通过以上几种

简单易行的方法,初步形成对各个开源软件的可用性总体认识,基于建立机构库的目的来设计各种用户的测试任务,分析了解各个开源软件运用于机构库的优劣。在机构实际建立机构库时,更好地选择适合机构需求的开源软件,再进行二次开发,从而节省人力财力。

参考文献:

- [1] Nielsen J. Usability Engineering [M]. Boston: Academic Press, 1993.
- [2] ISO/DIS 9241—11, Guidance on Usability [S], 1997.
- [3] Ottaviani J. University of Michigan DSpace (AKA Deep Blue) Usability Studies: Summary Findings [EB/OL]. (2006 - 06 - 02). [2007 - 09 - 25]. http://deepblue.lib.umich.edu/bitstream/2027.42/40249/1/Deep_Blue%28DSpace%29_usability_summary.pdf.
- [4] Kim J. Finding Documents in a Digital Institutional Repository: DSpace and ePrints [EB/OL]. [2007 - 09 - 25]. http://eprints.relis.org/archive/00005189/01/Kim_Finding.pdf.
- [5] Boock M. Improving DSpace@ OSU with a Usability Study of the ET/D Submission Process [EB/OL]. (2005 - 10 - 30). [2007 - 09 - 27]. <http://www.ariadne.ac.uk/issue45/boock/>.
- [6] Nichols D M, Thomson K, Yeates K A. Usability and Open - source Software Development [EB/OL]. [2007 - 09 - 24]. <http://www.cs.waikato.ac.nz/~daven/docs/oss.html>.
- [7] Gibbons S. UR Research; University of Rochester's Institutional Repository [EB/OL]. [2007 - 09 - 20]. <http://www.lita.org/ala/lita/litamembership/litaigs/emergingtechnol/URResearch2005Annual.ppt>.
- [8] Rieger O Y. Successful Digital Repositories; Selecting an Model Institution Repository [EB/OL]. [2007 - 09 - 15]. http://ecommons.library.cornell.edu/bitstream/1813/3727/1/Rieger_digital_repositories.pdf.
- [9] JCDL 2006: Opening Information Horizons [EB/OL]. [2007 - 09 - 25]. <http://jcdl2006.org/program/tutorials/morning>.
- [10] JCDL 2007: Building & Sustaining the Digital Environment [EB/OL]. [2007 - 09 - 25]. <http://www.jcdl2007.org/tutorials/>.
- [11] 周荣刚,张侃. 界面可用性评价之认知过程走查法 [J]. 人类工效学, 2003 (9): 30 - 33.
- [12] Nielsen J, Mack R. Usability Inspection Methods [M]. New York: John Wiley and Sons, 1994: 25 - 62.
- [13] Nielsen J. Usability Engineering [M]. San Francisco: Morgan Kaufmann, 1994: 115 - 164.
- [14] Nielsen J. Heuristic evaluation [EB/OL]. [2007 - 09 - 23]. http://www.useit.com/papers/heuristic/heuristic_list.html.
- [15] 吴昌旭,张侃. 人 - 计算机界面可用性评价方法 [J]. 心理科学, 2001, 24 (6): 727 - 728.

(作者 E-mail: holly0880@163.com)

欢迎订阅 2008 年《现代图书情报技术》(月刊)

《现代图书情报技术》杂志是由中国科学院国家科学图书馆主办的学术性、信息管理技术类专业期刊。1980 年创刊,原名《计算机与图书馆》,1985 年更名为《现代图书情报技术》,是国内图书馆学、情报学领域唯一一份技术性刊物,入选北大核心期刊要目总览,并被多次授予“中国图书馆学优秀期刊”荣誉称号。

(1) 期刊定位:面向国内信息技术领域的科研人员,跨图书馆学、情报学、信息科学等几大学科,以报道信息技术的研发与应用为主体,倡导原创性科研论文,同时兼顾应用实践型文章。

(2) 栏目设置:“数字图书馆”、“知识组织与知识管理”、“情报分析与研究”、“应用实践”、“动态”等一系列固定栏目以及“特邀专栏”、“专题”、“企业技术之窗”等不定期栏目。

月刊:国际通行 16 开版本

国内邮发代号:82 - 421

地址:北京中关村北四环西路 33 号(100080)

E-mail: jishu@mail.las.ac.cn

定价:56 元/期,全年定价:672 元

国外邮发代号:M4345

电话/传真:010 - 82624938

网址:<http://www.infotech.ac.cn>