

# “可用性”研究的文献计量与引文可视化分析

杜 华

( 安阳师范学院教育信息技术与传媒学院, 河南安阳 455000 )

**摘 要:** 利用HistCite软件对可用性研究文献进行了统计分析,理清了过去30多年国际可用性研究文献的时间分布,可用性研究的活跃地区、研究机构及载文期刊;通过分析引文编年图中的核心文献,揭示了可用性研究的发展历程和演进路径,并在此过程中列举了可用性研究的热点、前沿,预测了该领域的发展趋势。

**关键词:** 可用性; 量化分析; 可视化分析; 引文编年图; 研究热点; 研究趋势; HistCite

中图分类号: G350

文献标识码: A

DOI: 10.3772/j.issn.1674-1544.2013.04.014

## Quantification of Usability Research Literatures and the Analysis of Citation Visualization

Du Hua

(AnYang Normal University Institute of Education Information Technology and Communication, Anyang 455000)

**Abstract:** Firstly, this study untangles the temporal distribution, active region, research institution and publishing periodicals through quantitative analysis of past 31 years' research literatures about usability. Then, the author makes historiographs using Histcite. Finally, the article analyses the major literature and development track of usability research, meanwhile, forecasts the development tendency in this area.

**Keywords:** usability, quantitative analysis, visualization analysis, algorithmic historiography, research hotspot, research trends, Histcite

### 1 引言

“可用性”概念,由上世纪80年代中期出现的“对用户友好”的口号演变而来。可用性被称为“以用户为中心进行设计”的概念,指产品对学习者的有效、易学、高效、好记、少错和令人满意的程度,即用户能否用产品完成相应任务,效率如何,主观感受怎样。实际上,可用性是从用户角度所看到的产品质量,是交互式IT产品/系统的重要质量指标,更是产品竞争力的核心。近年来,随着多媒体计算机的快速发展和人们对软件产品质量日益关注,有关可用性的研究正逐渐成为人机交互和软件

设计等领域的热点。

本文以WOS(Web of Science)数据库中的两个子库SCI-E和SSCI(SCI-E指科学引文索引扩展版,SSCI为社会科学引文索引)为文献来源,检索策略为“主题=Usability”。鉴于2013年的数据并不完整,故时间跨度设置为2013年之前的所有年份。结果共检索到符合条件的文献记录1885篇(检索日期为2013年3月19日),文献数据集出版时间分布为1982-2012年。设置输出记录方式为“全记录(包含引用的参考文献)”,将结果保存为纯文本格式。

本研究采用引文图谱可视化分析方法,借助HistCite软件实现。HistCite(History of Cite)意为为引

**作者简介:** 杜华(1980—),女,安阳师范学院讲师,研究方向:信息技术教育应用。

**基金项目:** 河南省教育厅人文社会科学2012年重点项目“中国大学视频公开课的可用性研究”(2012-ZD-003);河南省教育厅科学技术研究重点项目“知识可视化技术支持的科普教育策略研究”(13A880020)。

**收稿日期:** 2013年5月22日。

文历史,或叫引文图谱分析软件。该软件系SCI发明人、美国著名情报学家和科学计量学家尤金·加菲尔德(Eugene Garfield)开发,能够用图示方式展示某一领域不同文献之间的关系,可以快速绘制出一个领域的发展历史,定位出该领域的重要文献以及最新的重要文献<sup>[1]</sup>。

将生成的纯文本格式数据导入HistCite软件,对1982–2012年可用性研究领域发表文献的时间、国家、研究机构、载文期刊的分布和被引频次、引文编年图等指标进行统计分析,结果见以下各节。

## 2 文献发文量

### (1) 时间分布

按年产量统计可用性研究文献的时间分布情况,结果如图1所示。由图1可以看出,WOS中对可用性进行学术研究有记载的文献始于1982年,过去30多年以可用性为主题研究的发文量总体呈现逐年递增的趋势。1994年以前,研究文献每年发文量都在10篇以内,1994–2002年维持在百篇以内,2003–2011年差不多每年都有一百多篇,2012年达到一个高峰,有211篇的学术文献在这一年发表。研究文献量的增长在一定程度上说明大家对此研究领域的关注度越来越高。

### (2) 国家(地区)分布

对研究可用性的国家(地区)进行分析,可以找出该领域研究的活跃地区。在HistCite中将文献按国家统计可以发现,过去31年内以可用性为主题的1885篇研究文献分布于68个国家(地区),其中未署国家(地区)的文献达314篇,文献量排名前15的国家(地区)如表1所示(其中,TLCS: Total Local Citation Score为本数据集的总被引次数, TGLS: Total Global Citation Score为WOS中的总被引次数)。无论是发文量还是被引次数,美国都稳居第一,远超过其他国家,其发文量甚至超过其后5个国家发文量的总和,占总发文量的28.17%。英国和德国分居二、三位,它们的发文量虽然也都过百,但依然远不及美国,前15个国家发文量共计1455篇,占总发文量的77.19%。中国排第十五位,发文量和被引次数都较低。

### (3) 研究机构分布

进一步对可用性文献的研究机构进行分析,发文量靠前的如表2所示(其中未署名研究机构的有123篇)。就发文量而言,美国的马里兰大学(Maryland Univ)和加拿大的多伦多大学(Toronto Univ)旗鼓相当,美国的印第安纳大学位列第三,IBM公司紧随其后。美国的普渡大学(Purdue Univ)和丹

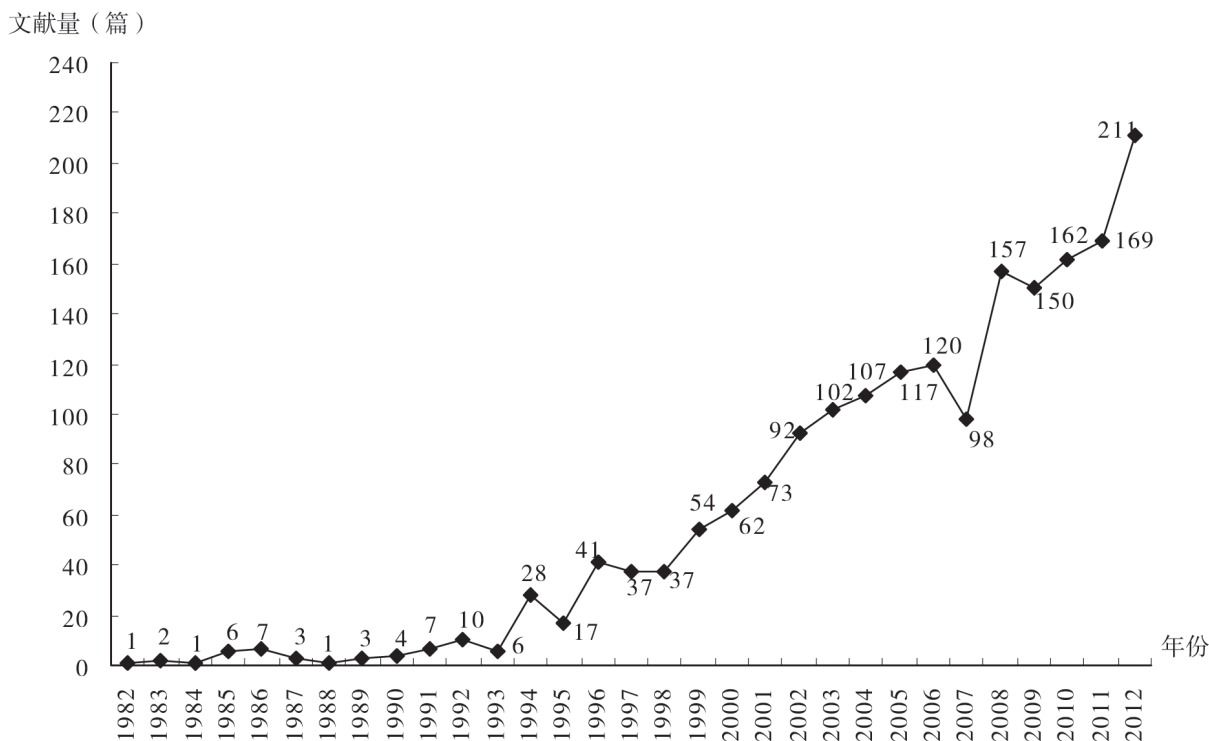


图1 可用性研究文献的年度时间分布

麦的哥本哈根大学 (Copenhagen Univ) 并列, 发文量同为 17 篇。美国密歇根大学 (Michigan Univ) 发文量排在德克萨斯大学 (Texas Univ) 之后, 与匹茨堡大学 (Pittsburgh Univ) 和荷兰的特温特大学 (Twente Univ) 均为 14 篇, 3 所研究机构并列。就单篇文献的被引次数 (TLCS/N) 而言, IBM 公司最高 (其 TLCS 值也最高), 这说明 IBM 公司研究学者的文献得到了可用性领域内更多学者的肯定和关注, 马里兰大学的 TGCS 值最高 (其 TGCS/N 也最高), 这说明马里兰大学学者发表的文献受到了与可用性相关研究领域最多的引用和转载。

#### (4) 载文期刊分布

1982–2012 年间共有 757 种期刊发表过以可用性为主题的文献, “文献量 (N)  $\geq 24$ ” 的期刊共计

10 种, 结果见表 3。10 种期刊合计载文量为 418 篇, 占总量的 22.18%。其中, 《Behaviour & Information Technology》载文量最高, 为 82 篇, 占总发文量的 4.35%, 该期刊的 TLCS 和 TGCS 值也都最高, 这说明在可用性研究领域, 该杂志所发文献不仅发文量最大, 受认可程度最高, 且得到其他相关研究领域专家引用和关注最多。

### 3 高影响力文献统计

LCS (Local Citation Score) 是某篇文献在当前数据集 (也即下载的 1885 篇文献) 中被引用次数, 它反映了某文献在某领域的受关注程度, 通过文献的 LCS 可以快速定位一个领域的经典文献。为了找出可用性研究领域的高影响力重要文献, 利

表 1 可用性研究文献量前 15 名国家分布

| 序号 | 国家/地区                | N (篇) | 百分比 (%) | TLCS | TGCS | TLCS/N | TGCS/N |
|----|----------------------|-------|---------|------|------|--------|--------|
| 1  | USA (美国)             | 531   | 28.17   | 889  | 5459 | 1.67   | 10.28  |
| 2  | UK (英国)              | 193   | 10.24   | 162  | 1436 | 0.84   | 7.44   |
| 3  | Germany (德国)         | 106   | 5.62    | 98   | 1026 | 0.92   | 9.68   |
| 4  | Canada (加拿大)         | 85    | 4.51    | 133  | 918  | 1.56   | 10.80  |
| 5  | Netherlands (荷兰)     | 82    | 4.35    | 51   | 494  | 0.62   | 6.02   |
| 6  | Spain (西班牙)          | 59    | 3.13    | 42   | 375  | 0.71   | 6.36   |
| 7  | Denmark (丹麦)         | 55    | 2.92    | 140  | 573  | 2.55   | 10.42  |
| 8  | South Korea (韩国)     | 51    | 2.71    | 56   | 410  | 1.10   | 8.04   |
| 9  | Turkey (土耳其)         | 51    | 2.71    | 12   | 456  | 0.24   | 8.94   |
| 10 | Japan (日本)           | 49    | 2.60    | 3    | 99   | 0.06   | 2.02   |
| 11 | Sweden (瑞典)          | 43    | 2.28    | 54   | 388  | 1.26   | 9.02   |
| 12 | Australia (澳大利亚)     | 42    | 2.23    | 12   | 147  | 0.29   | 3.50   |
| 13 | Italy (意大利)          | 42    | 2.23    | 42   | 242  | 1.00   | 5.76   |
| 14 | Finland (芬兰)         | 35    | 1.86    | 18   | 196  | 0.51   | 5.60   |
| 15 | Peoples R China (中国) | 31    | 1.64    | 8    | 52   | 0.26   | 1.68   |

表 2 文献量 N  $\geq 14$  的研究机构分布

| 序号 | 机构 (国家)              | N (篇) | 百分比 (%) | TLCS | TGCS | TLCS/N | TGCS/N |
|----|----------------------|-------|---------|------|------|--------|--------|
| 1  | Maryland Univ (美国)   | 27    | 1.43    | 86   | 837  | 3.19   | 31.00  |
| 2  | Toronto Univ (加拿大)   | 26    | 1.38    | 32   | 186  | 1.23   | 7.15   |
| 3  | Indiana Univ (美国)    | 22    | 1.17    | 24   | 230  | 1.09   | 10.45  |
| 4  | IBM Corp (美国)        | 18    | 0.95    | 98   | 352  | 5.44   | 19.56  |
| 5  | Purdue Univ (美国)     | 17    | 0.90    | 33   | 137  | 1.94   | 8.06   |
| 6  | Copenhagen Univ (丹麦) | 17    | 0.90    | 59   | 267  | 3.47   | 15.71  |
| 7  | Edinburgh Univ (英国)  | 16    | 0.85    | 11   | 44   | 0.69   | 2.75   |
| 8  | Texas Univ (美国)      | 15    | 0.80    | 56   | 189  | 3.73   | 12.60  |
| 9  | Michigan Univ (美国)   | 14    | 0.74    | 5    | 89   | 0.36   | 6.36   |
| 10 | Pittsburgh Univ (美国) | 14    | 0.74    | 1    | 26   | 0.07   | 1.86   |
| 11 | Twente Univ (荷兰)     | 14    | 0.74    | 2    | 94   | 0.14   | 6.71   |

表3 文献量 $N \geq 24$ 的载文期刊分布统计

| 序号 | 载文期刊                                                     | N (篇) | 百分比 (%) | TLCS | TGCS |
|----|----------------------------------------------------------|-------|---------|------|------|
| 1  | Behaviour & Information Technology                       | 82    | 4.35    | 218  | 862  |
| 2  | International Journal of Human-Computer Interaction      | 59    | 3.13    | 211  | 781  |
| 3  | Interacting with Computers                               | 58    | 3.08    | 103  | 705  |
| 4  | International Journal of Human-Computer Studies          | 41    | 2.18    | 107  | 512  |
| 5  | Technical Communication                                  | 40    | 2.12    | 15   | 90   |
| 6  | Ergonomics                                               | 32    | 1.70    | 27   | 134  |
| 7  | Work-A Journal of Prevention Assessment & Rehabilitation | 32    | 1.70    | 0    | 2    |
| 8  | Communications of The Acn                                | 25    | 1.33    | 85   | 675  |
| 9  | Journal of the American Medical Informatics Association  | 25    | 1.33    | 26   | 74   |
| 10 | International Journal of Psychology                      | 24    | 1.27    | 11   | 30   |

用HistCite的统计功能,将检索结果按LCS降序排列,其中 $LCS \geq 22$ 的17篇文献如表4所示。其中,GCS (Global Citation Score)表示文献在WOS中的被引用次数。

由表4可知,在所列17篇文献中,发表在《Behaviour & Information Technology》上的有4篇,《International Journal of Human-Computer Interaction》上的文献为3篇,这在一定程度上佐证了表3中对文献载文期刊分布情况的统计分析。每一篇研究文献都有一个编号,见表4中的文献编号列,这在引文编年分析部分会用到。

#### 4 引文编年分析

利用HistCite的Graph Maker功能,以LCS Count

为条件,设定节点数 (Limit) 为62 (鉴于时间跨度为31年,故选择31的倍数62作为引文编年图中能出现的节点数),绘制可用性研究文献的引文编年图,以期找出可用性研究领域的核心文献、研究演进路径、热点,并对该领域的研究前沿和发展趋势作出预测,结果如图2所示。

在可用性研究文献引文编年图中,出现最早的节点7来自1985年(图2中的节点7,节点号与表4中文献编号相对应)。该节点是由Gould和Lewis共同撰写的《为了可用性而设计:关键原则与设计师需要考虑的事情》。在文中,两位作者从理论和实证两个角度展开论述:在理论上,描述了生产有用且易用的计算机系统需要遵循的三条系统设计原则:提前并持续关注用户,基于实证的使用测量,

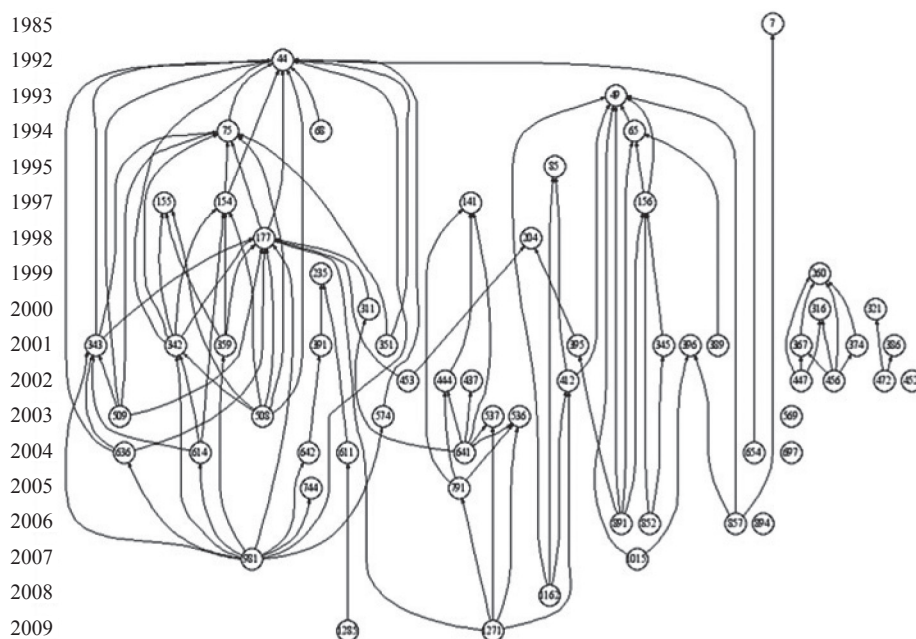


图2 可用性研究文献引文编年图



表4 当前数据集被引频次(LCS)≥22的文献

| 文献编号 | 题名                                                                                                         | 作者                        | 刊名(年,卷(期):页)                                                                   | LCS | GCS |
|------|------------------------------------------------------------------------------------------------------------|---------------------------|--------------------------------------------------------------------------------|-----|-----|
| 44   | Refining the test phase of usability evaluation: how many subjects is enough                               | Virzi RA                  | Human Factors.<br>1992, 34 (4): 457-468                                        | 56  | 157 |
| 177  | Damaged merchandise? A review of experiments that compare usability evaluation methods                     | Gray WD<br>Salzman MC     | Human-Computer Interaction.<br>1998, 13 (3): 203-261                           | 49  | 139 |
| 7    | Designing for usability: key principles and what designers think                                           | Gould JD<br>Lewis C       | Communications of the Acm.<br>1985, 28 (3): 300-311                            | 37  | 279 |
| 85   | IBM computer usability satisfaction questionnaires-psychometric evaluation and instructions for use        | Lewis JR                  | International Journal of Human-Computer.<br>Interaction. 1995, 7 (1): 57-78    | 35  | 209 |
| 857  | Current practice in measuring usability: challenges to usability studies and research                      | Hornbaek K                | International Journal of Human-Computer.<br>Studies. 2006, 64 (2): 79-102      | 34  | 129 |
| 343  | The evaluator effect: a chilling fact about usability evaluation methods                                   | Hertzum M<br>Jacobsen NE  | International Journal of Human-Computer.<br>Interaction. 2001, 13 (4): 421-443 | 29  | 68  |
| 453  | Assessing a firm's web presence: a heuristic evaluation procedure for the measurement of usability         | Agarwal R<br>Venkatesh V  | Information Systems Research.<br>2002, 13 (2): 168-186                         | 29  | 188 |
| 452  | Web site usability, design, and performance metrics                                                        | Palmer JW                 | Information Systems Research.<br>2002, 13 (2): 151-167                         | 27  | 304 |
| 641  | Cognitive and usability engineering methods for the evaluation of clinical information systems             | Kushniruk AW<br>Patel VL  | Journal of Biomedical Informatics.<br>2004, 37 (1): 56-76                      | 27  | 115 |
| 342  | Criteria for evaluating usability evaluation methods                                                       | Hartson HR<br>Andre RS    | International Journal of Human-Computer.<br>Interaction. 2001, 13 (4): 373-410 | 25  | 50  |
| 636  | Comparative usability evaluation                                                                           | Molich R<br>Ede MR        | Behaviour & Information Technology.<br>2004, 23 (1): 65-74                     | 25  | 39  |
| 49   | Sumi: the software usability measurement inventory                                                         | Kirakowski J<br>Corbett M | British Journal of Educational Technology.<br>1993, 24 (3): 210-212            | 23  | 97  |
| 154  | Tracking the effectiveness of usability evaluation methods                                                 | John BE<br>Marks SJ       | Behaviour & Information Technology.<br>1997, 16 (4/5): 188-202                 | 23  | 47  |
| 156  | A proposed index of usability: a method for comparing the relative usability of different software systems | Lin HX<br>Choong YY       | Behaviour & Information Technology.<br>1997, 16 (4/5): 267-278                 | 23  | 74  |
| 75   | Sample sizes for usability studies: additional considerations                                              | Lewis JR                  | Human Factors. 1994, 36 (2): 368-378                                           | 22  | 43  |
| 155  | Comparison of evaluation methods using structured usability problem reports                                | Lavery D<br>Cockton G     | Behaviour & Information Technology.<br>1997, 16 (4-5): 246-266                 | 22  | 38  |
| 611  | The interplay of beauty, goodness, and usability in interactive products                                   | Hassenzahl M              | Human-Computer Interaction.<br>2004, 19 (4): 319-349                           | 22  | 128 |

迭代设计。在实证方面,作者用数据表明他们提出的设计原则并不总是依靠设计师的直觉,而是有证据支持的,并且提供了上述设计原则成功应用于实践的例子<sup>[2]</sup>。7号文献的LCS值为37,排在所有可用性研究文献的第三位。

对引文编年图的进一步分析可以发现可用性研究在1992年出现一个重要节点44。它是由美国通用电话电子公司实验室Virzi RA教授1992年8月发表在《Human Factors》上的名为《精炼可用性评估的测试阶段:多少被试者就足够了》的文章。Virzi RA教授认为,友好的用户界面设计和低廉的测试成本有助于提高产品的开发生存周期。针对可用性

问题,他通过3个实验报告来说明在产品测试评估阶段需要多少被试者参与就足够了的问题,研究有3点基本发现:80%的可用性问题是由最初的4~5个被试者发现的;后面的被试者似乎越来越难给出新信息;最初的几个被试者有可能发现最严重的可用性问题<sup>[3]</sup>。由于该文献的研究成果是基于实验报告,来自产品可用性评估的实践,得出的结论自然比较有说服力,再加上其发表时间较早,故其被后来者争相引用也就不足为奇,这也致使该文的LCS值高达56,成为表4所列17篇文献中最高的。

按照时间顺序梳理,在随后的1993、1994和1995年又分别出现一个重要节点。1993年出现的节

点为49号文献,该文由Kirakowski J和Corbett M共同撰写,1993年9月发表在《英国教育技术杂志》上,名为《软件可用性测量目录》。软件可用性测量目录是从用户使用满意程度及遇到的问题角度调查软件可用性的问卷,它是MUSiC(计算领域可用性标准度量项目的一部分)。软件可用性测量目录由爱尔兰科克学院大学人因学研究小组开发,Jurek Kirakowski正是该研究小组的主管。SUMI采用国际化的50个题目问卷,针对每个题目,用户可以回答:同意、不确定、不同意。SUMI从以下方面评价用户对软件可用性的看法:易学性,即用户觉得自己能够开始使用软件并学习新特性的安心程度;效率,即用户觉得软件辅助自己工作的功效;喜欢程度,即用户对软件的一般情绪反应,例如:是否喜欢;用户控制,即用户觉得自己控制软件(而非被软件控制)的程度;对用户的帮助,即用户觉得软件辅助自己使用的程度。SUMI要求至少要有10个用户参与测试<sup>[4]</sup>。

1994年的节点为75号研究文献。该文由Lewis JR独著,名为《可用性研究的样本量:额外注意事项》。Lewis首先回顾了1992年Virzi的研究发现,也就是上面提到的44号文献,在此基础上,他利用一个独立的可用性研究来验证Virzi的3个基本发现,结果显示第二个基本发现,也就是后面的被试者似乎越来越不容易提供新信息是完全正确的,但第一个基本发现并不完全正确,第三个基本发现根本就没有证据支持。Lewis的研究发现,随着样本量增大,收益在递减。如果32%~42%的数据存在错误,那么最初的4~5个被试者就可以发现产品存在的80%的可用性问题,但如果数据的出错率低于32%,那么要发现产品可用性问题的80%就需要超过5个被试者参与。将发现的可用性问题按严重性分类,并未发现其与被发现速率之间存在关联。研究数据显示在给定了数据出错率平均估算值的前提下,二项式概率可以为预测问题发现曲线提供一个很好的模型。最后,作者通过一个经济学模型估算了不同情况下的投资回报率,数据显示具有最大投资回报率的样本量大小范围深受测试产品数据出错率的影响<sup>[5]</sup>。

1995年的节点85号文献也是由Lewis JR独著,名为《IBM电脑可用性满意度调查表:心理测量评估与使用说明书》。在文中,作者描述了IBM当时正进行的主观可用性测量研究,该研究的焦点

是用心理测量的方法开发并评估测量用户对系统可用性满意度的问卷。文章主要目标有两个:讨论IBM用于测量用户对电脑系统可用性满意度的4套问卷的心理特征;提供问卷、实施和评分说明。可用性实践者能很有信心地使用这些问卷帮助其测量用户对电脑系统可用性的满意度<sup>[6]</sup>。

1997年出现了3个重要节点,分别为154、155和156。154号文献是John BE和Marks SJ撰写的名为《追踪可用性评估方法的有效性》的文章。作者通过一个案例追踪分析了使用6种不同的可用性评估方法方法预测的可用性问题,6种可用性评估方法分别是权益分析、认知过程走查法、GOMS(目标、操作、方法、选择规则)、启发式评估、用户行为标注和简单阅读说明书。案例通过比较不同方法预言的用户测试结果评定其预测能力,通过计算引起执行代码改变问题的数量评定其说服力,通过用户测试新系统的结果评定设计修改的有效性。结论是,上述可用性评估方法并不像人机交互领域想要的那样有效果<sup>[7]</sup>。该文献借鉴了之前44号和75号文献的部分结论。节点155是名为《用于建构可用性问题报告评估方法的比较》的文章。在文中,作者回顾了前几年人机交互研究提出的用于预测交互系统潜在可用性问题的几种评价方法,在此基础上对这些评估方法的有效性提出了质疑。文章指出,由于这些评估方法缺乏严密性和普遍性,因此,用它们预测的与有经验的用户通过实际测试发现的可用性问题是否一致需要特别注意。该文作者提出了一个新的可用性问题评估的报告框架,用于改善评估方法预测的与有经验用户测试发现的可用性问题的一致性,并指出了这种报告格式适用的设计研究情境<sup>[8]</sup>。节点156由节点49和65发展而来,是名为《一个被推荐的可用性指标:一种用于比较不同软件系统相对可用性的方法》的文章。该文由美国普渡大学(Purdue Univ)教授撰写。作者指出,可用性正在成为一个越来越重要的软件标准,但是当前的可用性测量方法不是难以应用就是过于依赖评估者的专业知识。文章以人类信息加工理论为基础,确定了与软件可用性有关的8个需注意的人为因素,它们分别为兼容性、一致性、适应性、易学性、最少动作、最小记忆负载、知觉限制和用户指导。普渡大学的可用性测试问卷(PUTQ)正是来自上述8个因素和人类信息加工的3阶段。一个旨在测试PUTQ有效性的实验结果表明PUTQ和用户交

互满意度调查问卷 (QUIS) 存在高度相关。另外, PUTQ 发现了在两个试验性的交互系统之间用户绩效的差异, 但 QUIS 却未能发现<sup>[9]</sup>。

177 是 1998 年出现的重要节点, 由 44 和 75 号节点发展而来。该节点文献 LCS 值为 49, 排在所有可用性文献的第二位。名为《被损坏的商品? 有关可用性评估方法比较实验的综述》, 由 Gray WD 和 Salzman MC 合著。文章指出, 在人机交互领域, 界面设计已成为研究者和实践者的一个核心主题, 但实验设计尚未引起大家足够的重视。然而, 在某种程度上, 对界面设计可靠与有效的指导却往往依赖于实验设计。在文中, 作者通过比较检验了可用性评估方法的五个实验设计, 它们都对人机交互思想和实践有着重要的影响。结果显示, 实验设计中的一个很小的问题都有可能導致软件致命性的错误。实验的方法可作为一个强有力的工具, 不仅可以帮助选择可用性评估方法, 而且可以用于说明人机交互的其他事项。要想获得想要的结果, 就需要加强并密切关注实验设计<sup>[10]</sup>。

2001 年有 342 和 343 两个重要节点。342 是名为《用于评价可用性评估方法的标准》的文章。在文中, 作者指出, 可用性是软件系统质量的重要指标, 目前有关可用性评估的方法很多, 但却较少有使用者能够理解每一种方法的功能和局限, 哪一种方法更有效? 应如何使用, 又该用于什么场合等。然而, 由于缺少统一标准, 又无法对不同的评估方法进行可靠地评估和比较。基于此, 文章对可用性评估方法的要素、比较准则和绩效衡量等进行了讨论, 这些在比较可用性评估方法的研究中是非常有用的<sup>[11]</sup>。342 节点文献借鉴了 44、75、154、155 和 177 号文献的研究成果。343 号文献由 Hertzum M 和 Jacobsen NE 合著, 名为《评估员的影响: 一个关于可用性评估方法令人寒心的事实》。在文中, 作者认为计算机专业人员需要强健易用的可用性评估方法帮助其系统提升计算机人工制品的可用性, 但是目前使用最为广泛的 3 种可用性评估方法: 认知过程走查法、启发式评估和出声思维研究针对相同的系统, 如果是不同的人员进行评估, 即使是采用同一种方法, 得到可用性问题的答案都不一样, 这充分说明在可用性问题的评估过程中, 评估人员对结果的影响。作者在分析了 11 个案例的基础上, 得出以下 3 点结论: 模糊的目标分析导致任务场景的多变性; 评价过程的不清楚使得评价缺少抓手; 不明确的问题

界定标准使得任何事情都可能被当作可用性问题。为了尽量减少评估人员对结果的影响, 最简单的方法就是要让尽可能多的人员参与到可用性的评估当中<sup>[12]</sup>。

距今最近的、LCS 值又较高的节点是 857。该节点文献 2006 年 2 月发表在国际人机研究杂志上, 由丹麦哥本哈根大学教授 Kasper Hornbaek 独著, 名为《可用性测量的最新实践: 可用性研究面临的挑战》。文章指出, 如何测量可用性是人机交互研究和用户界面评价中的一个重要问题, 作者通过分析发表在人机交互核心期刊和会议上的 180 篇有关可用性测量论文并将其分类, 回顾了有关可用性测量的最新实践, 区分了与测量有关的几个问题, 包括这些研究是否确实在测量可用性? 测量能否覆盖足够大范围? 研究者是如何推理的? 结果是否达到了可用性测量的目的? 在已有的研究中, 对可用性测量的选择和推理大都缺乏一个有效可靠的, 类似在用户界面研究中使用的对可用性权威性的解释。基于上述文献综述, 作者讨论了可用性研究与如何进行可用性测量研究所面临的挑战, 主要包括依据经验区分和比较可用性的主观与客观测量, 集中开发和利用学习与记忆测量工具, 研究长期使用与可用性, 扩展使用后满意度测量的范围, 验证并标准化使用的主观满意度问卷, 使用微观和宏观任务及对应测量工具验证其与可用性测量的相互关系等等<sup>[13]</sup>。表 4 中的其他文献这里就不再一一赘述。

## 5 总结

(1) 从引文编年图中可以看出, 以“可用性”为主题的研究虽然出现时间不久, 但其相关研究发展较快, 并且出现了较为经典的综述性文章。

(2) 近几年高被引文献表明, 研究主要集中在对“可用性”评估、测量方法与标准的讨论与比较, 以及不同的评估方法、样本量大小和评估人员对评估结果的影响等。

(3) 目前可用性研究领域尚未形成一套最具权威性评估与测量方法及标准。不同“可用性”评估方法的适用场合、如何实施, 以及与之相应的个案研究或典型实践很有可能成为未来的研究热点, 而对权威性可用性评估方法、过程的提炼, 对评估人员的专业培训等也极有可能成为未来研究的主要趋势之一。

(下转第 110 页)



- [11] 秦明,赵伯飞,龙建成.学科生态系统模型的建构和解读[J].西安电子科技大学学报:社会科学版,2007,17(4):153-158.
- [12] 李爱彬,王鹏,赵文闯.基于监理制的“211工程”重点学科建设项目管理模式探讨[J].学位与研究生教育,2009(10):26-30.
- [13] 李爱彬,王鹏.高校“211工程”重点学科建设项目监理模型设计[J].科技管理研究,2011(16):108-111.
- [14] 王鹏,李爱彬.基于全寿命周期理念的高校重点学科建设项目监理角色探讨[J].学位与研究生教育,2010(10):52-56.
- [15] 郑凌莺,胡守忠,唐幼纯.基于BSC和DHP方法的地方高校学科评价[J].科技管理研究,2011(21):55-57.
- [16] 陈晶.高校学科建设绩效评价研究[D].武汉:武汉理工大学,2010.
- [17] 梁传杰.学科绩效评价方法的构建与应用[J].高教发展与评估,2010,4(26):36-46.
- [18] 谈毅,纓允桓.政府科技计划绩效评价理论基础与模式比较[J].科学学研究,2004(2):150-156.
- [19] 李丽亚,宋扬.科技项目绩效考评方法初探[J].中国科技论坛,2004(6):110-111.
- [20] 唐云锋.论我国科技政策评估体系中存在的问题[J].中国科技论坛,2004(4):106-110.
- [21] 戴国庆,李丽亚.国外科技项目绩效考评研究与借鉴[J].中国科技论坛,2005(5):45-49.
- [22] 杨建顺.国家科研项目监理的法学研究[J].河南省政法管理干部学院学报,2005(1):5-10.
- [23] 张艳丽.国家科研项目建立系统的构建[J].学术交流,2004(11):133-137.
- [24] 张艳丽.国家科研项目监理与系统工程[J].科学管理研究,2005(1):50-55.
- [25] 姚亦佳,李小燕.国外科研项目管理借鉴[J].政策与管理,2002(5):40-41.
- [26] 刘涛.国家科技项目监理制的有关问题探讨[J].科技进步与对策,2004(2):28-29.
- [27] 徐华,安琴.科技型中小企业技术创新基金项目监理模式研究[J].科学学与科学技术管理,2008,2:33-36.

(上接第87页)

## 参考文献

- [1] Garfield E. Historiographic Mapping of Knowledge Domains Literature[J]. Journal of Information Science, 2004,30(2):119-145.
- [2] Gould JD, Lewis C. Designing for Usability: Key Principles and What Designers Think[J]. Communications of the Acm, 1985,28(3):300-311.
- [3] Virzi RA. Refining the Test Phase of Usability Evaluation—How Many Subjects Is Enough[J]. Human Factors, 1992,34(4):457-468.
- [4] Kirakowski J, Corbett M. Sumi—the Software Usability Measurement Inventory[J]. British Journal of Educational Technology, 1993,24(3):210-212.
- [5] Lewis JR. Sample Sizes for Usability Studies—Additional Considerations[J]. Human Factors, 1994,36(2):368-378.
- [6] Lewis JR. IBM Computer Usability Satisfaction Questionnaires—Psychometric Evaluation and Instructions for Use[J]. International Journal of Human-Computer Interaction, 1995,7(1):57-78.
- [7] John BE, Marks SJ. Tracking the Effectiveness of Usability Evaluation Methods[J]. Behaviour & Information Technology, 1997,16(4/5):188-202.
- [8] Lavery D, Cockton G. Comparison of Evaluation Methods Using Structured Usability Problem Reports [J]. Behaviour & Information Technology, 1997, 16(4/5): 246-266.
- [9] Lin HX, Choong YY. A Proposed Index of Usability: A Method for Comparing the Relative Usability of Different Software Systems[J]. Behaviour & Information Technology, 1997,16(4/5):267-278.
- [10] Gray WD, Salzman MC. Damaged Merchandise? A Review of Experiments That Compare Usability Evaluation Methods[J]. Human-Computer Interaction, 1998,13(3):203-261.
- [11] Hartson HR, Andre RS. Criteria for Evaluating Usability Evaluation Methods [J]. International Journal of Human-Computer Interaction, 2001,13(4):373-410.
- [12] Hertzum M, Jacobsen NE. The Evaluator Effect: A Chilling Fact About Usability Evaluation Methods[J]. International Journal of Human-Computer Interaction. 2001,13(4):421-443.
- [13] Hornbaek K. Current Practice in Measuring Usability: Challenges to Usability Studies and Research[J]. International Journal of Human-Computer Studies, 2006, 64(2):79-102.