

●曹 梅^{1,2}, 朱学芳¹
(1. 南京大学 信息管理系, 江苏 南京 210093 2. 南京师范大学 教育技术系, 江苏 南京 210097)

用户信息行为的研究方法体系初探^{*}

摘 要: 本文从研究方法学的角度, 初步构建了用户信息行为的研究方法体系, 揭示其基本特征和走向, 并详细阐述了各类实用数据采集技术和分析方法。
关键词: 用户; 信息行为; 研究方法

Abstract The paper constructs a tentative user information behavior research method system from the perspective of research methodology. Its basic features and trend are discussed. Various data acquisition and analysis methods are expounded.
Keywords user information behavior research method

胡昌平、乔欢等在《信息服务与用户》中专门介绍了信息服务与用户研究的基本方法, 包括直接调查与间接调查方法、统计测量分析、抽样方法、比较分析法、相关分析法、回归分析法、德尔菲法、马尔科夫分析法等, 并通过实例分析揭示各类方法在情报学用户研究中的具体应用方式^[1]。胡岷撰文介绍了当前在国外及我国台湾地区, 研究人员在进行用户检索行为研究时惯常使用的研究方法, 包括: 问卷调查法、访谈法、小组讨论法、实验法、观察法、出声思维法和记录分析法等 7 种。胡岷认为, 用户调查是用户研究中最重要研究方法^[2]。笔者认为, 这些研究方法的归纳缺乏一定的科学体系, 有必要从研究方法学角度, 对用户信息行为研究方法的体系进行整体架构。

1 研究方法体系及其特征

1.1 研究方法体系

从研究方法学角度而言, 研究方法包括研究策略、数据采集、分析方法三大部分^[3]。研究策略是一种对内容研究而言的整体性方法, 典型的研究策略有调查策略、定性策略、案例或行动研究策略以及实验策略等; 典型的数据采集方法包括: 结构化问卷调查、访谈、采访、小组讨论、出声思维等定量或定性的方法; 在分析方法上, 有统计分析、归纳性内容分析、数学分析等典型方法。见表 1。

就用户信息行为研究方法而言, 本文在相关文献分析

表 1 用户信息行为研究方法体系

研究策略	数据采集	分析方法
调查策略	结构化问卷 半结构化访谈	统计测量分析: 描述性统计 回归分析 比较分析 相关分析 归纳性内容分析: 话语分析 关键事件分析 数学方法: 马尔科夫分析法
定性策略	主题采访 观察法 小组讨论 时间线访谈 出声思维法 用户小组日志	
实验策略	标准实验 准实验设计	
网络策略	系统日志文件 网络工具采集日志 社会标签	

基础上, 尝试给出了一个研究方法体系。其中, 在教育和社会学中经常采用的“案例与行动研究”在信息行为研究中应用较少, 未纳入该体系; 而“网络策略”则因其特色应用而纳入进来, 一些以计算机系统日志、网络日志采集工具、网络链接、社会标签等为代表的新数据采集技术正逐步应用到信息行为研究中来。

1.2 “走向定性”的特征

总体而言, 20 世纪 80 年代信息搜寻理论和元理论方面的发展 (如意义建构法) 使研究策略和数据采集方法的重点发生了必要的转变^[3], 呈现“走向定性”的特征。

1960—1985 年实验性信息搜寻研究明显的特点是使用基于结构化问卷和访谈进行数据采集的定量调查, 有

^{*} 本文为 2008 年教育部人文社会科学项目的系列研究成果之一, 项目编号: 08JA870009。
• 情报理论与实践 •
© 1994-2012 China Academic Journal Electronic Publishing House. All rights reserved. http://www.cnki.net — 37 —

关调查方面的研究性文章占 80% 以上^[3]。结构化的数据收集方法支持定量的数据分析, 但需要对现象的研究进行预先假设, 如果在数据采集、研究现象的关键维度方面有所忽略, 结果也许是致命的。整个信息检索行为并非简单的线性行为活动, 它涵盖有复杂的认知和情感因素在其中, 研究的目的在于尝试解释人们所面对真实信息情景的复杂性, 并就影响用户检索行为的各因素及其相互关系予以深入的分析。Wilson 认为定性方法特别适合于引发信息查寻行为的需求研究^[3]。半结构化或无限制的采访被推荐用于数据采集, 例如, 用户的独白、对话或访谈记录经文字的转译后可捕捉用户反应情景, 这些话语资料经研究者逐一编码、归类便可实现对所观察现象的整合。因此, 用户信息行为研究中越来越多地采用深入访谈、小组讨论、出声思维、观测、用户小组以及日志等定性方法, 定性方法比调查表需要付出更多努力, 但更能详尽分析信息需求。但是, “走向定性” 不是不作分析的借口, 相反, 综合应用定量和定性两种方法是必要的选择。

1.3 描述性研究和解释性研究

在信息行为研究中, 描述性研究多, 而解释性的研究少。定量调查的研究结果往往是描述性的, 能够确定观察到的现象的特征或者现象之间的相关性, 但是一般找不出因果关系。Vakkari 提出, 信息查寻的定性方法通常不以解释为目的, 而是以理解行动者及其行动特征为目的。他还指出, 需要仔细考察信息查寻现象的构件 (或变量) 之间的关系及其所支持的行动^[3]。因此, 在信息行为的研究中, 对情境、情景、任务过程、行动者、信息、查寻过程、来源、系统以及信息利用关系的了解既需要描述性的研究方法, 包括定性方法和定量调查, 也需要解释性的研究方法, 比如: 现场和受控的实验设计。

2 数据采集技术

2.1 结构化调查问卷

从已有用户研究来看, 作为数据收集的手段和行为分析的基础, 用户调查已成为其最重要的研究方法。用户的普及得益于它们的易用性和廉价, 并且支持定量分析。在用户信息行为的调查策略中, 常用的数据采集方法有: 结构化调查问卷、访谈、小组讨论等; 调查内容则围绕对用户个人特征 (如性别、年龄、专业、认知类型等)、信息需求、信息查寻动机、查寻方式、查寻策略、查寻困难、结果评价等方面展开。一般地, 问卷发放有传统的邮寄、当面递送等人工方式, 现在也可以利用网络这一方便快捷的交流工具, 将问卷放到网上或经由网络交流工具如 QQ 等让网络用户自由填写。不过一些研究者认为, 仅依靠电子方式的调查会导致错误的结论^[2], 最好将人工发放和

电子发放结合起来。

有时候调查问卷与实验法同时进行。一般研究者常会把它分成两部分, 一部分放在查寻前, 调查内容一般包括对用户个人特征及查寻意图等背景资料, 用户只有填完这部分问卷才可以进行信息查寻; 一部分放于行为后, 用以调查用户行为过程的体验和对查寻结果的满意度及评价, 等等。

相关研究案例很多, 如 Ellis 和 Haugan 使用半结构采访考察一家国际石油天然气公司内研究人员的信息查寻行为^[4]; 胡昌平等采用结构化问卷, 对科研人员信息查询行为调查分析, 揭示网络环境下高校科研人员一些信息行为特征^[5]。

2.2 访谈

访谈法也是用户研究中最常用的调查方法, 有结构化访谈和主题访谈两种方式。该方法更注重研究目的的需要而不是样本的代表性, 其调查对象常常是那些具备研究所要求的经历, 特别是高层的难以征集的有某种经验或特性、并愿意接受采访的人。问卷调查主要收集行为表面以及行为者人口统计特征方面的数据, 而深入的访谈则可以让研究者获得现象背后的定性数据, 比如人们对检索结果的预期值和实际评价, 以避免量化数据不足解释影响行为发生和变化的深层因素。林丽娟认为: 针对特定用户进行的深入访谈, 可以详细记录用户的人口特征、认知形态、学习使用过程中思路的变化与概念形成及知识建构的方式^[2]。

访谈也可分信息查寻前和查寻后。查寻前的访谈 (开放式和非结构化的) 用来了解用户的需求、过去的查寻经验和检索行为以及自我评价等, 而查寻后的访谈主要问及用户对本次查寻的感觉和他们是否找到需要的信息, 等等。后继访谈也可用来确定用户是否按意愿行事。对访问者来说提问可以帮助他们判断隐藏在某一行为背后的意图。

2.3 小组讨论

小组讨论法是一种定性研究方法, 其主要特点是强调小组内成员间的互动, 通过集体访谈来了解参与者的想法及其原因, 从而使得研究者可以接触到参与者的观念、态度和意见, 等等^[2]。小组讨论法一般先由研究人员征集志愿者, 在确定对象符合自己的研究条件后, 比如是否近期内有过图像查寻行为, 将其分组。讨论开始时由研究人员指定讨论的话题, 之后便由参加者自由讨论。同访谈法一样, 这种方法一般都配以录音供研究人员分析用。小组讨论法有助于探索 and 发现那些人们很少了解的论题, 它可以发现问题的新视点, 调查参与者行为发生的背景。例如, Esposito 等人利用小组讨论法去获取大学生使用某些网站

的理由以及他们希望系统做出哪些改进等信息^[6]。

2.4 时间线访谈

戴文 (Dervin) 创造了一种时间线 (Time Line) 访问调查法^[2], 此方法注重与应答者相关的、过去发生的关键事件, 也称关键事件研究法。采访应答者提出问题或忧虑的状况, 该状况再沿时间回溯上去, 形成一种情形的再创造, 可以发现阻碍进程的信息缺口以及所寻求的帮助和行动措施。时间线访问法一般使用开放式问题, 分三部分: ①过去一个月中遇到的问题或忧虑关心之事; ②处理某一特别问题的详细过程; ③被访人将如何处理 4 个假设的状况。这种方法的长处在于过去真实而具体的情节有助于了解真实的信息问题和查寻过程, 缺点在于是一种回溯性的, 可能会影响应答者对过去事件的报告。Hewins 对关键事件研究法在医学信息学方面的应用进行了综述^[7]。

2.5 观察法

观察法用在信息行为的研究中, 指研究人员在信息查寻、检索现场或直接或利用特定的辅助工具对用户的行为过程进行观察并记录测量的过程, 可基于真实场景的观察, 也可在基于模拟场景的观察, 后者在信息行为研究中较多采用, 一般采用行为录像的方式录制被试搜索行为全过程的视频, 然后对视频回放、编码分析。观测法一般配合实验法使用。例如, Goodnum 等设计用户搜索实验, 观察用户网络图像查寻过程中的状态转移, 发现行为模式和查寻策略^[8]; 朱明泉等采用实验室观测法记录被试完成不同类型的信息搜索任务, 进行数据分析探讨用户与网络交互的行为模型^[9]。观测法既可使研究人员把观察的事实同研究对象所处的情境联系起来, 又能对查寻过程进行直接、真实、具体的了解。然而观测法有一个挑战, 就是研究人员要想仔细观察用户的动作步骤而不为人所知是很困难的。

2.6 出声思维法

出声思维法是指用户在检索过程中, 大声将其进行的行动和思考过程用语言表达出来, 所有的独白被录音, 之后研究者再根据录音做进一步的定性分析。整个过程中研究者只能以观察者伫立一旁, 除提示用户出声外不可与他们交谈。出声思维法可以帮助研究者精确详细地了解用户伴随行为变化的思维过程。例如, Yang 的研究中的出声思考协议包括要求调查对象描述他们在解决给定问题时的想法。通过对口头数据的记录和分析, 找出受试者在问题解决过程中的目标导向和信息查寻模式^[10]; 王佩玲等借助屏幕抓取软件和对用户出声思考的录音, 创建一种过程跟踪方法, 采用统计测量和话语分析方法, 研究用户—网络交互的查寻行为模型^[11]。这种方法的不足之处在于, 首先用户由于要去考虑表达词汇, 思维多少会受影响, 严

重的可能会中断甚至改变原先的思维进程; 其次, 用户可能会在出声的同时有意识调整自己的行为, 存在行为典型性减弱的风险。出声思维实验之前必须有一个训练过程。

2.7 实验法

实验可以定义为一种研究环境: 研究人员在这个环境中能够准确地说明或“控制”在研究中普遍存在的各种条件, 它是检验因果关系假设的最适宜的技术。用户查寻行为研究中一个重要内容就是尝试确定各种影响行为的因素, 而这往往依赖于现场或受控的实验设计。实验法并不一定基于严格的实验室环境, 在行为研究领域更多采用的是“准实验”设计, 准实验设计是在自然的信息查寻环境下利用自然发生的事件作为自变量来设计实验的。常见的程序是: 根据研究目的创设信息查寻的环境, 如特定的检索系统或范围、给定不同的查寻任务等, 跟踪用户的查寻过程和结果, 从各个不同维度来分析比较, 结合用户个人特征、查寻任务的分析, 找出各因素间的因果或相关关系。例如, Isenberg 基于模拟用户实验分析专家和新手在解决问题过程和质量之间的区别, 采用出声思维和话语分析方法^[12]; 黄丽霞通过实验, 探讨网络环境下对不同类型的问题人们选择的第一检索策略^[13]; Fukumoto 基于开放任务/封闭任务的图像查寻的比较研究^[14]。

2.8 网络数据采集

利用网络技术来采集反映用户信息行为的各类数据, 包括: 特定的信息检索系统内嵌的日志文件, 利用网络监控工具来采集某个系统的用户行为日志, 网络社会标签集等各种形式。一些系统不仅保存了系统的响应, 同时也保存了用户在键盘和鼠标的的所有输入。它可以提供用户做什么而不是说什么的数据, 数据的生成完全不受研究的影响, 同时数据收集也不受研究者左右。这种方法能提供详细的用户与系统的交互信息, 获得一些群体特征和面向系统的特征。例如, Jansen 和 Spink 等人对 Excite 搜索引擎的用户分析^[15]; 王建勇等对“天网”中英文搜索引擎查寻日志的分析^[16]; 沙勇忠等使用网络计量工具 WebTrends 对某科研机构网站的 Web 日志文件处理, 提取有关科研人员信息行为的量化指标进行实证分析^[17]。这种方法的缺陷也很明显: 不能将检索与特定的用户联系起来, 无法反映用户特性与检索行为之间的关系, 从而无法获知用户检索的真实意图和对检索结果的满意度。所以很多研究者在实际研究中往往还会结合其他方法。值得注意的是, Web 2.0 环境下社会网络标签成为新的体现用户信息需求和行为特征的网络数据, 得到很多研究者的关注。例如, Yoon 采集并分析 Flickr 网站的图像标签集, 揭示用户标签所呈现的概念分布及其语义空间, 为构建面向用户的图像分类体系寻求有效途径^[18]。

3 分析方法

很多论著中对各种数据分析方法都有阐述, 本文出于体系完整考虑, 对用户信息行为研究中的主流并且较具有特色的分析方法稍作介绍。

3.1 统计测量分析

利用社会学中的统计测量方法获取用户特征资料。在有关问题中, 对这些特征资料稍加归纳、整理和分析便可以得出直观而明确的结论。其中, 描述性统计、比较分析、相关分析、回归分析等是其中常用的分析思路。

3.2 归纳性内容分析

归纳性内容分析是使用归纳方法发展出的关于现象理论的定性方法学^[3], 体现扎根理论的具体应用。特别适合于从数据中衍生出理论而非证实既有理论, 即通过对数据的系统分析或“编码”, 揭示出它的特性和关系。对于现实生活中信息问题状态的识别, 配合半结构化访谈、时间线访谈、主题采访、出声思维法、网络日志文件等数据采集而进行归纳性内容分析是有效的组合。它是最重要的定性分析方法。

话语记录分析是归纳性内容分析中的一种特例, 即对用户的话语文本进行归纳, 可能是谈话, 是对信息查寻的关键事件的回顾, 是关于信息行为的时间线访谈结果, 也可能是认知的出声思考等。

3.3 马尔科夫分析

马尔科夫在用数学方法分析布朗运动的过程中发现了后来被称之为马尔科夫链的随机现象, 创立了马尔科夫分析理论。在信息行为研究中, 常用马尔科夫链理论来分析用户查寻信息的连续过程, 发现一些稳定的行动模式或其他信息查寻的规律等。例如, Chen和Cooper利用马尔科夫分析法, 研究网络信息系统中的用户信息利用模式^[19]。

4 结束语

本文对用户信息行为研究中的各种实用研究方法进行了系统归纳; 构建了一个初步的方法体系, 并揭示了该领域方法体系的走向和特征; 研究策略中的核心部分之“数据采集技术”得到了详细介绍并辅以研究个案解释, 希望能够为用户信息行为的研究提供方法上的参考。□

参考文献

- [1] 胡昌平, 乔欢. 信息服务与用户 [M]. 武汉: 武汉大学出版社, 2001: 66-148.
- [2] 胡岷. 用户行为研究常用调查方法略述 [J]. 成都理工大学学报: 社会科学版, 2005, 13 (3): 54-57.
- [3] INGERSEN P, JARVELIN K. 转折: 在情境中集成信息查寻与检索 [M]. 张新民, 等译. 北京: 科学技术文献出版

社, 2007: 80.

- [4] ELLIS D, HAUGAN M. Modeling the information seeking patterns of engineers and research scientists in an industrial environment [J]. Journal of Documentation, 1997, 53 (4): 384-403.
 - [5] 胡昌平, 贺娜, 张俊娜. 网络环境下高校科研人员信息查询行为的调查与分析 [J]. 情报理论与实践, 2008 (2): 223-225.
 - [6] ESPOSITO D, et al University students' perceptions of the Internet: an exploratory study [J]. The Journal of Academic Librarianship, 1999, 25 (6): 456-461.
 - [7] HEWNS E T. Information needs and use studies [C] // WILLIAMS M E (Ed.). ARIST25, 1990: 147-172.
 - [8] GOODRUM A, et al A state transition analysis of image search pattern on the Web [C]. CNR 2003 LNCS2728: 281-290.
 - [9] 朱明泉, 张智君, 任衍具. 互联网信息搜索用户行为模型的探索性研究 [J]. 浙江大学学报: 理学版, 2006 (7).
 - [10] YANG S. Information seeking as problem-solving using a qualitative approach to uncover the novice learners information-seeking process in a perseus hypertext system [J]. Library & Information Science Research, 1997 (19): 71-92.
 - [11] WANG P L, et al User's interaction with world wide Web resources: an exploratory study using a holistic approach [J]. Information Processing & Management, 2000, 36 (2): 229-251.
 - [12] ISENBERG D J. Thinking and managing: a verbal protocol analysis of managerial problem solving [J]. Academy of Management Journal, 1986, 29 (4): 775-788.
 - [13] 黄丽霞. WWW环境下信息检索策略实验调查研究 [J]. 中国图书馆学报, 2003 (3): 54-56.
 - [14] FUKUMOTO T. An analysis of image retrieval behavior for metadata type image database [J]. Information Processing and Management, 2006, 42 (3): 723-728.
 - [15] JANSEN B J, SPINK A, SARACEV I C T. Real life: real user and real needs: a study and analysis of user queries on the Web [J]. Information Processing and Management, 2000, 36 (2): 207.
 - [16] 王建勇. 海量Web搜索引擎系统中用户行为的分布特征及其启示 [J]. 中国科学, 2001 (8).
 - [17] 沙勇忠, 阎劲松, 苏云. 网络环境下科研人员的信息行为分析 [J]. 情报科学, 2006 (4): 485-491.
 - [18] YOON J. Towards a user-oriented thesaurus for non-domain-specific image collections [J]. Information Processing and Management, 2009, 45 (4): 452-468.
 - [19] CHEN H, COOPER M D. Stochastic modeling of usage pattern in a Web-based information system [J]. Journal of the American Society for Information Science and Technology, 2002, 51 (9): 841-857.
- 作者简介: 曹梅, 女, 1977年生, 博士生, 讲师。
朱学芳, 男, 1962年生, 教授, 博士生导师。
收稿日期: 2009-09-16