



协同信息检索行为研究

RESEARCH ON THE COLLABORATIVE
INFORMATION SEEKING BEHAVIOR

吴 丹 著

国家社会科学基金项目最终成果

协同信息检索行为研究

吴 丹 著



國家圖書館出版社
National Library of China Publishing House

图书在版编目(CIP)数据

协同信息检索行为研究/吴丹著. --北京:国家图书馆出版社,2015.9
ISBN 978-7-5013-5666-9

I. ①协… II. ①吴… III. ①情报检索—研究 IV. ①G252.7

中国版本图书馆 CIP 数据核字(2015)第 204147 号

书 名 协同信息检索行为研究
著 者 吴 丹 著
责任编辑 高 爽

出 版 国家图书馆出版社(100034 北京市西城区文津街 7 号)
(原书目文献出版社 北京图书馆出版社)

发 行 010-66114536 66126153 66151313 66175620
66121706(传真),66126156(门市部)

E-mail btsfxb@nlc.gov.cn(邮购)

Website www.nlcpress.com →投稿中心

经 销 新华书店

印 装 北京华艺斋古籍印务有限责任公司

版 次 2015 年 9 月第 1 版 2015 年 9 月第 1 次印刷

开 本 710×1000(毫米) 1/16

印 张 19

字 数 360 千字

书 号 ISBN 978-7-5013-5666-9
定 价 70.00 元

前言

长期以来,信息检索的单用户研究一直是该领域的研究热点。近年来,随着 Web2.0 和计算机支持的协同工作技术的发展,协同信息查寻与检索(Collaborative Information Seeking & Retrieval, CIS&R)的研究已逐渐成为信息科学领域新的研究热点,特别是有关协同信息检索领域的用户行为研究更成为当前信息检索领域的重要研究课题。由于协同信息检索领域用户研究的复杂性,其研究内容涉及用户个人特征(如性别、年龄、职业)、用户的心理和认知、信息需求、查寻动机、检索策略、结果评价等方面。随着检索任务复杂性的增加,跨学界领域合作的开展,协同技术的不断开发,社群环境下人们交流与协作的逐渐频繁,科学研究困难性的加剧,突破时空地域限制的合作增多,人们对协同信息检索系统的需求迫切增加,各个领域的协同信息检索行为也越来越多地开展起来。因此,在对协同信息检索的用户行为进行研究时,为全面考虑各种因素,研究者需从多维角度并采用多种研究方法对其进行分析研究,通过研究分析用户协同信息检索过程中的行为特征、认知特点、情感取向等来进一步开发协同检索系统,协同信息检索用户行为研究逐渐成为学界所关注的热点。

协同信息检索的用户行为研究并不是研究简单的检索行为活动,它包含了各个层面的复杂因素:工作任务、上下文情境、用户的认知和情感因素等。对于该领域的研究调查要涉及多个要素,还需全面考虑影响用户协同信息检索行为的各个因素及其之间的相互关系。本书在总结有关协同信息检索行为的相关概念、研究视角、理论基础、影响因素、行为模型、典型系统等理论研究的基础上,提出了一个协同信息检索行为模型及若干假设。模型主要围绕三大影响因素、七项协同信息检索行为,以及不同的协同信息检索情境展开。然后设计了三个实验分别检验一般情境下的小组因素和任务因素、一般情境下的个人因素和任务因素,以及科研合作情境下的小组因素。三个实验除了包含不同情境,检验了三大影响因素对七项协同信息检索行为的影响外,还涉及不同人数构成的小组、特定情境下的协同信息检索行为特点,以及检验了各因素对模型外其他方面的影响。

通过研究我们发现,用户的协同信息检索行为首先是与情境密切相关的。由于一般情境下的社会化搜索任务是多元化的,因此任务是影响协同

信息检索的重要因素。并且在一般情境下,当前的协同检索工具基本能满足用户需求。而特殊情境往往是一些真实的工作场景,用户会表现出一些特有的行为特点。以科研合作情境为例,协同学术信息检索行为的产生集中体现了学术环境的影响,沟通交流是协同学术信息检索行为区别于个人信息检索行为的重要体现。

其次,用户的协同信息检索行为不同于个人信息检索行为,主要体现在检索方式、认知方式、协同方式、编辑方式、检索效果、交流分享、情感状态几个方面。其中,检索方式主要包括检索行为、浏览网页行为、推荐行为、使用图片行为、加书签行为、产生重复检索式行为;认知方式则包括情境认知、任务认知和协同认知;协同方式有按内容分工、按角色分工、动态分工策略、静态分工策略。

最重要的是,个人因素、小组因素和任务因素均会在不同程度上对协同信息检索行为产生影响。其中,个人协同能力对全部七项行为均产生影响,对检索方式的影响在于推荐行为和重复检索式行为上,对认知方式的影响在于情境认知和协同认知上。小组类型对检索方式、编辑方式、协同方式、认知方式四项行为产生影响,对检索方式的影响仅在推荐行为上,对认知方式的影响仅在任务认知上。小组知识结构对除编辑行为之外的六项行为均产生影响,对检索方式的影响在于重复检索式行为和书签行为,对认知方式的影响在于情境认知和协同认知上。任务类型对除编辑行为和协同方式之外的五项行为产生影响,对检索方式的影响在于网页浏览行为、检索行为和使用图片行为上,对认知方式的影响在于任务认知上面。任务难度对检索方式、编辑方式、协同方式、情感状态四项行为产生影响,对检索方式的影响在于检索行为和浏览网页行为上。

本书采用的数据收集方法主要包括:①模拟实验法。实验一和实验二是通过二人组和三人组的模拟协同信息检索实验,让小组成员在规定时间内协作完成模拟的任务,以此来研究不同协同能力、不同小组类型、不同类型的和不同难度的检索任务对用户协同信息检索行为的影响。②结构化行为日志法。实验三是真实场景下为期三个月的长期研究,结构化行为日志主要反映各成员在研究初期的合作行为,如在寒假期间是否进行过相关学术行为、利用哪些学术资源以及情绪的变化等。③系统日志收集法。三个实验都用到了 Coagmento 系统。通过对揭示各小组在完成不同任务时的具体行为和检索活动的系统日志数据的收集,可以探究他们协同信息检索行为的特征。④问卷调查法。调查问卷主要调查用户的背景信息、用户的认知情况以及实验用户对系统的评价等。⑤半结构化访谈法。通过对实

验者进行单独或小组采访,可以从个人和团队整体的视角上深入探讨各用户的检索行为、情绪变化、总体感受以及对实验系统的建议等主观方面的内容。

本书采用的数据分析方法主要包括:①比较分析法。比较了影响协同信息检索行为的各影响因素,提出我们要研究的三大因素;比较了各协同信息检索行为相关模型,选择在 C2S 模型的基础上提出本研究所需的协同信息检索行为模型框架;比较了几个典型的协同信息检索系统,选择 Coagmento 系统作为三个实验的研究平台。②统计分析法。采用了多项统计分析方法来分析实验数据,如单因素方差分析法(Analysis of Variance, ANOVA)、双因素方差分析法(Two-way Analysis of Variance, Two-way ANOVA)等分析点击流数据和问卷调查数据。③情感分析法。对日志中有关情感的数据进行了分析,通过提取聊天记录中的情感词汇,进行词频和情感粒度分析,并辅之以知网(HowNet)的词表进行检验与二次分析,以此来分析个体用户的情感特征。④内容分析法。针对定性数据,利用 ATLAS.ti 7 质性分析软件,并结合 Strauss 和 Glaser 等提出的扎根理论(Grounded Theory)研究方法对访谈数据进行开放编码、主轴编码和聚类编码分析,分析用户协同检索过程中的行为特征。

全书共分为十章:

第一章,协同信息检索行为相关概念。从信息检索中的协同和信息检索与协同的关系出发引入研究的缘起和研究目的,并介绍协同信息检索的相关概念和协同信息检索行为的定义与特点。

第二章,协同信息检索行为研究综述。对协同信息检索行为研究进行全面的国内外文献综述。

第三章,协同信息检索行为相关理论基础。介绍了协同机制、计算机支持的协同理论和支持协同信息检索的技术。

第四章,影响协同信息检索行为的因素。对已有关于协同信息检索行为影响因素的研究进行归纳、分类和总结,找出还有哪些因素值得研究。

第五章,协同信息检索行为模型框架的提出。对有关协同信息检索行为模型进行分类比较,基于 C2S 模型提出协同信息检索行为模型研究框架和相关假设。

第六章,协同信息检索系统比较及实验平台选择。在介绍协同信息检索系统概念和分类的基础上,对典型的协同信息检索系统进行比较分析,选出适合本研究的系统作为实验平台。

第七章,一般情境下小组因素和任务因素分析:二人组协同信息检索

实验一。设计一个由两名用户构成一个小组的实验,用于检验一般情境下小组类型和任务难度对协同信息检索行为的影响。

第八章,一般情境下个人因素和任务因素分析:三人组协同信息检索实验二。设计一个由三名用户构成一个小组的实验,用于检验一般情境下个人协同能力和任务类型对协同信息检索行为的影响。

第九章,科研合作情境下小组因素分析:多人组协同信息检索实验三。设计一个真实科研项目申报的团队协同学术信息检索实验,用于检验科研合作情境下小组知识结构对协同信息检索行为的影响。

第十章,协同信息检索行为模型修正与讨论。根据三个实验结果,对所提出的协同信息检索行为模型相关假设进行验证,并修正模型,同时对模型的三要素进行讨论。另外,基于实验结果分析协同信息检索活动中的协同机制,并提出对协同信息检索系统的建议。最后,总结本研究的结论,对未来研究工作进行展望。

本书研究的学术创新主要体现在以下四点:

第一,已有的协同信息检索行为研究中有关基础理论,特别是协同机制的研究较少,而合作是协同信息检索区别于个人信息检索的关键,对协同定义、原则、过程等的研究能更好地加强协同信息检索行为的理论建设。所以本书对协同机制、计算机支持的协同理论和支持协同信息检索的技术等相关理论进行探究,梳理协同信息检索行为理论体系,弥补当前研究的不足。

第二,现有研究对不同用户类型的协同信息检索行为研究较少。协同能力的强弱与用户的协同检索行为和活动直接相关,而按专业和兴趣对小组分类也是一个崭新的研究视角。尽管现有文献对于任务与信息行为之间的关联性研究较多,但大多处于概略性描述,实践研究较少,并且任务类型大多只涉及事实类任务,关于特定任务类型对协同检索行为产生了怎样具体影响的研究几乎没有。因此,任务类型和任务难度的差异也是重要的研究视角。总的来说,本书所提出的几个协同信息检索行为影响因素是较新颖的,丰富了协同信息检索行为理论研究。

第三,以往的协同学术信息检索研究主要集中在行为本身,研究比较表层。大多以学生完成课堂作业为实验对象,多是真实场景的调研,是对现实情况的翔实记录,研究方法单一,总结出影响因素,却忽视了对这些因素影响方式、影响程度的分析。本书所进行的协同学术信息检索研究是以真实的科研立项为情境,将真实场景和可控实验相结合,在真实场景观察中嵌套实验设计,多途径、多角度地收集描述协同检索过程中的用户数据,

包括外在行为表现和内在认知与情感态度。重点探讨协同学术信息检索过程中关键因素及其影响机制,识别相互关系,在一定程度上弥补了之前研究的分散性,使协同学术信息检索行为研究更全面,更体系化。

第四,现有研究所进行的大多是以两个小组成员为一组的用户协同实验。有研究发现,对于共同完成协同检索任务的小组,80.7%是两个小组成员的合作,只有19.3%是在三个或以上成员的合作,且基于协同检索系统进行的三人及以上为小组的实验更少。其原因有可能是缺乏支持多成员交互的工具和系统,且实验成本更高、可控性更差。而本书在同一个模型框架下,根据不同的情境和影响因素,设计了二人组、三人组、多人组的用户实验。这既能体现对不同影响因素考察的针对性,又能体现对协同信息检索行为影响因素考察的全面性。因此,本书所提出的模型和相关实验检验是具有创新性的,将通过协同检索行为和效果的深层次分析来丰富这个领域的研究文献,促进协同信息检索行为研究的理论发展。

本书研究的学术价值体现在:①对协同信息检索行为的相关理论基础进行了梳理,深入探究了协同机制、计算机支持的协同理论和支持协同信息检索的技术等内容,为协同信息检索行为后续研究提供了理论支撑。②从影响因素的视角所构建的协同信息检索行为模型分析可以丰富该领域的文献,为将来的研究提供一些理论依据和支持,开创了新的研究视角,具有重要的理论意义。③本研究最终验证了用户协同能力、小组类型、小组知识结构、任务类型、任务难度、不同情境对协同信息检索行为是有影响的,扩展了研究的视角,并注重了用户行为研究中多因素的考量以及各个因素之间的相互关系,促进协同信息检索行为研究的进一步发展。④本书所进行的实验较为全面,突破二人小组实验的模式,涵盖二人组、三人组、多人组实验模式;既有模拟任务检索实验,又有基于真实场景的实验;实验任务设计有严格分类,体现了不同难度,设置了评价标准;研究方法上也采用了多种数据收集方法和数据分析方法。这些实验研究经验将丰富协同信息检索实验研究方法,为以后用户行为实践研究提供一些借鉴。

本书是国家社会科学基金后期资助项目“协同信息检索行为研究”的成果。我的研究生余文婷、向雪、邱瑾参与了本项目研究。其中,余文婷进行了多人组协同信息检索实验三的实施,并参与撰写了第三章、第五章、第六章和第九章;向雪进行了二人组协同信息检索实验一的实施,并参与撰写了第四章和第七章;邱瑾进行了三人组协同信息检索实验二的实施,并参与撰写了第二章和第八章。

本书得以完成,还要感谢全国哲学社会科学规划办公室及图书馆、情

报与文献学学科规划评审组各位专家的大力支持,特别是黄长著研究员的指导与帮助;感谢武汉大学信息管理学院各位领导和老师的帮助,特别是马费成教授和陈传夫教授的指导与推荐;感谢北京大学信息管理系赖茂生教授的推荐;感谢国家图书馆出版社高爽老师的辛勤付出;感谢所有本书参考文献作者;感谢所有参与实验的同学们。

限于能力和水平,本书还有许多不足之处,有很多问题还有待进一步深入研究。敬请各位专家、学者、同仁、读者批评指正。

国家社科基金后期资助项目

出版说明

后期资助项目是国家社科基金项目主要类别之一,旨在鼓励广大人文社会科学工作者潜心治学,扎实研究,多出优秀成果,进一步发挥国家社科基金在繁荣发展哲学社会科学中的示范引导作用。后期资助项目主要资助已基本完成且尚未出版的人文社会科学基础研究的优秀学术成果,以资助学术专著为主,也资助少量学术价值较高的资料汇编和学术含量较高的工具书。为扩大后期资助项目的学术影响,促进成果转化,全国哲学社会科学规划办公室按照“统一设计、统一标识、统一版式、形成系列”的总体要求,组织出版国家社科基金后期资助项目成果。

全国哲学社会科学规划办公室

2014年7月

目 录

前 言	(1)
第一章 协同信息检索行为相关概念	(1)
第一节 信息检索中的协同	(1)
第二节 信息检索与协同的关系	(3)
第三节 协同信息检索相关概念	(3)
第四节 协同信息检索行为	(5)
第二章 协同信息检索行为研究综述	(9)
第一节 协同信息检索行为研究方法	(9)
第二节 协同信息检索行为研究发展阶段	(21)
第三节 协同信息检索行为研究内容	(23)
第四节 协同信息检索行为实践研究领域	(26)
第五节 协同信息检索行为研究特点与趋势	(29)
第三章 协同信息检索行为相关理论基础	(37)
第一节 协同机制	(37)
第二节 计算机支持的协同理论	(39)
第三节 支持协同信息检索的技术	(45)
第四章 影响协同信息检索行为的因素	(52)
第一节 任务因素	(52)
第二节 场景因素	(54)
第三节 用户群因素	(60)
第四节 协同信息检索行为影响因素小结	(63)
第五章 协同信息检索行为模型框架的提出	(67)
第一节 协同信息检索行为相关理论模型	(67)
第二节 基于 C2S 模型的研究框架构建	(80)
第三节 相关假设的提出	(81)

第六章 协同信息检索系统比较及实验平台选择	(87)
第一节 协同信息检索系统	(87)
第二节 典型协同信息检索系统	(89)
第二节 协同信息检索系统比较	(91)
第三节 系统评估	(96)
第四节 协同信息检索系统比较小结	(98)
第五节 本研究实验平台选择	(99)
第七章 一般情境下小组因素和任务因素分析:二人组协同	
信息检索实验一	(104)
第一节 实验设计	(104)
第二节 实验结果分析	(109)
第三节 实验结果讨论	(141)
第四节 实验一结论	(142)
第八章 一般情境下个人因素和任务因素分析:三人组协同	
信息检索实验二	(144)
第一节 实验设计	(144)
第二节 实验结果分析	(150)
第三节 实验结果讨论	(175)
第四节 实验二结论	(178)
第九章 科研合作情境下小组因素分析:多人组协同	
信息检索实验三	(180)
第一节 实验设计	(181)
第二节 实验结果分析	(183)
第三节 实验结果讨论	(226)
第四节 实验三结论	(227)
第十章 协同信息检索行为模型修正与讨论	(230)
第一节 协同信息检索行为模型修正	(230)
第二节 协同信息检索情境讨论	(235)
第三节 用户协同信息检索行为讨论	(236)

第四节	协同信息检索行为影响因素讨论	(240)
第五节	协同信息检索中的协同机制	(242)
第六节	协同信息检索系统评价与建议	(243)
第七节	研究结论与展望	(245)
附录 1	实验一调查问卷	(250)
附录 2	实验一访谈大纲	(260)
附录 3	实验一各个任务的评价标准	(261)
附录 4	实验二调查问卷	(262)
附录 5	实验二访谈大纲	(269)
附录 6	实验二各个任务的评价标准	(270)
附录 7	实验三背景问卷	(275)
附录 8	实验三结构化行为日志	(280)
附录 9	实验三实验问卷	(282)
附录 10	实验三访谈大纲	(286)

第一章 协同信息检索行为相关概念

很长时间以来,信息检索一直多注重单个用户层面上的研究。然而在现代社会中,越来越多的工作、活动是由一群有着共同目标的人一起参与,合作完成的。有研究表明,在很多搜索任务中,包括和他人的合作,如计划旅行、组织社会活动或者完成家庭作业等,人们有着协同检索的需求^[1]。信息检索中的协同现象日益突出,研究对象应该由个人层面扩展到群体层面。随着 Web2.0 及 CSCW(Computer Supported Cooperative Work,计算机支持的协同工作)等新技术的发展,也使协同信息查寻与检索(Collaborative Information Seeking & Retrieval, CIS&R)研究逐渐成为信息科学领域新的研究热点,特别是有关协同信息检索领域的用户行为研究更成为当前信息检索领域的重要研究课题。Ingwersen 和 Järvelin^[2]认为,信息检索嵌套在信息查寻行为的情境之内,其本身则处于与工作相关或日常的工作任务情形和情境之中,交互和合作是信息查寻和利用过程的重要部分。一些信息技术,如团队电子日历、即时交流通信工具、电子公告栏、团队项目管理工具等逐步应用到协同信息检索活动中,提高了团队合作的效率。但是面对信息量的剧增和信息质量的良莠不齐,在如何让用户快速找到信息,以及让用户之间能就查寻体验进行沟通等方面,传统的搜索引擎、检索系统以及检索技术局限明显,面临巨大的难题。而且现在的网络搜索工具是为单用户独立工作而设计的,支持多用户共同检索的工具较少。为了解决这些问题,研究者们将协同的概念引进到信息检索中,以期解决在检索过程中遇到的问题,辅助检索活动的有效开展。

第一节 信息检索中的协同

协同对人们来说并不是一个陌生的概念,它随着社会的出现和进步而不断发展。当面临困难或者复杂任务时,人们会本能地选择协同^[3]。比

如,各级教育中常出现的小组作业,需要学生协调完成;在学术环境中,越来越多的科研项目需要来自不同领域、不同机构甚至不同地区的多位学者共同完成;一些企业会根据客户及其需求的不同设置针对性的项目小组,这种工作小组已成为一种普遍的工作模式。又比如,大多数搜索引擎都利用用户点击的链接或者其他查寻行为来影响他人网页搜索结果的排序或是检索词推荐,从而提高搜索体验。虽然这种情境下用户并不是主动帮助他人,甚至没有意识到他们的检索行为将会影响他人的检索结果,但实际上也实现了一种隐性的协同。可以看出,协同常见于生活中很多场景。在此类场景中还经常包括信息查寻、检索、分享和整合等行为。但很长一段时期里,人们极少将其看作协同信息检索行为。此外,传统的信息检索以单用户为中心,被认作是个人行为 and 单纯的信息传递过程,所以协同信息检索的相关研究和支持较少。

近年来,随着时代的发展,特别在当前网络环境和大数据背景下,协同信息检索行为逐渐成为一个前沿、热点的新研究领域,从理论研究走向实践研究。这反映了信息检索从单用户发展为多用户协同的趋势,同时也和协同信息技术的发展进步密切相关。在数字化网络环境下,网络内容从某种程度上来看是社会性的,被人们创造和分享。因此信息检索实际上成为一个社会化意义建构和分享的过程,而且共享检索结果和共享过程也是复杂问题解决的群体智能形成过程,是社会记忆过程^[4]。另外,协同信息检索行为有其自身的特点。首先,考虑到任务的复杂性和难度,协同信息检索项目通常会经历多个阶段,持续时间长,逐步推进。其次,在协同信息检索中,所有的参与者都有着共同的目标和共同的利益,而这些目标能将他们紧密联系,实现互惠互赢。最后,在整个过程中各方互动交流频繁,互动性极强。虽然国内外都有学者开展协同信息检索相关研究,但总体来说这个领域的研究不算太多,也谈不上成体系。人们对协同信息检索行为,如协同原理、行为特点等知之甚少。所以,协同信息检索行为是极具研究价值的领域。

本书旨在对协同信息检索的不同情境、不同行为方式和影响因素进行分析,从而揭示协同信息检索行为机制并构建行为模型,实现协同信息检索行为的综合研究。具体而言,本书首先介绍协同信息检索行为相关概念,梳理现有各相关领域研究脉络,并总结协同信息检索行为理论基础,完善协同信息检索行为理论研究。随后,本书在分析影响协同信息检索行为因素的基础上,选择借鉴 C2S 模型,提出了本研究的协同信息检索行为模型框架。最后,在 Coagmento 系统平台上进行实验研究。通过三个不同情境下不同类型的实验对此模型框架进行了检验与修正。

第二节 信息检索与协同的关系

信息检索广义上是寻找信息用以满足信息需求的过程,不仅仅包括检索和搜索,还包括浏览、分享、评价、整合信息以及意义建构(sensemaking)等。而协同是多方为达到互惠互赢的共同目标而一同工作的过程。这两个看似隶属不同领域的概念之间却存在着内在相关性,并相互促进。

一方面,由于任务复杂程度的增加和个人知识结构与能力的局限性导致个人很难胜任各种类型的信息检索任务,而协同可以帮助人们解决个人力量很难处理的困难或复杂问题,完成此类信息检索活动。比如,规划全家旅游时会涉及不同的家庭成员。他们需要考虑出行时间、出行方式、旅行地点、游览景点、食宿等多个因素,检索并综合大量信息,较为复杂。为完成共有互惠目标(即家庭旅行),他们需要共同参与,协调各方意见,最终完成旅游方案。

另一方面,身处信息时代,人们的生活深受信息影响。很多情景下,人们都需要检索信息以解决各种问题,协同也不例外。信息检索能支持协同,因为在协同过程中经常也需要信息检索,并且在很多协同活动中信息检索是最主要的行为。比如,在完成科学项目时,会涉及多个步骤,如成员分工、资料搜集、论文撰写等,而信息检索是资料搜集中十分重要的内容。有效的信息检索能提高资料搜集的效率和效果,为后续工作的完成奠定坚实基础。

综上,不论是信息检索活动还是协同活动,都与协同信息检索紧密相关,很难将这两种行为完全区分。

第三节 协同信息检索相关概念

协同信息查寻与检索,即 Collaborative Information Seeking & Retrieval, CIS&R(本书简称协同信息检索)。协同信息检索行为作为协同信息行为的一种,是信息检索由个人层面发展到群体层面的现象。协同信息检索是一个跨学科领域,涉及信息检索、人机交互和 CSCW 等领域的研究。然而,在不同的领域中所用到的术语表达不尽相同,如社会搜索(social searching)、协同探索式搜索(collaborative exploratory search)、共同浏览

(co-browsing)、协同信息行为(collaborative information behavior)等^[5]。关于协同信息检索的研究最早可以追溯到1968年,由Taylor开创了协同信息检索领域^[6]。半个世纪以来,协同信息检索研究已有了较大的发展。

由于每个学科的研究侧重点不同,协同信息检索在这些学科中的定义也各不相同,没有一个统一的定论。在对于协同信息检索内涵的认识上,不同的学者基于不同的研究视角给出了不同的定义,主要包括基于用户行为过程的定义和基于技术视角的定义。

(1)基于用户行为过程的定义。Dumais^[7]和Hust^[8]等人将其理解为任何共同解决信息问题的协同信息行为。Hertzum^[9]认为协同信息检索是用户信息寻求与协作背景的合体。Naderi^[10]等人把协同信息检索定义为一种社会化的信息检索方法,通过记录用户与系统的交互行为来为以后的用户提供更好地检索服务。

(2)基于技术视角的定义。Foley和Smeaton^[11]将其理解为在信息检索过程中应用不同的技术来支持的协同活动。Krömer^[12]等人认为协同网络搜索(Collaborative Web Search,CWS)是一种查寻者之间通过隐性协作来进行知识交换的查寻模式。Agarwal等^[13]认为协同信息检索是通过利用协作用户的专业知识和经验来帮助用户满足他们信息需求的一组相关技术。

基于用户行为过程的协同信息检索更多关注用户的行为特征研究,在各种认知、情境下探讨用户的协作行为,重视对协同信息检索中影响用户行为的各种因素的分析。基于技术角度的定义更强调技术对协同检索与信息共享的支持性,将研究局限在组织、技术范围内。尽管各学者所给出的定义不同,然而他们都会在定义中强调信息检索过程中的合作。对于这种合作,有学者认为劳动分工和知识共享是它的两个关键部分。一方面降低重复劳动,提高效率,另一方面能积极影响其他的合作者^[14]。也有学者将协作分为基于任务和基于目标两种。两者的区别在于,前者建立在显性合作意图上,而后者在潜在的共同信息需求上分享知识,是隐性的合作意图^[15]。从上文中也可以总结出协同信息检索的动机,即需求、分工和技能的多样化^[16]。

本书更倾向于采用Foster^[17]的定义:“在信息查寻和检索过程中使个人走向合作的系统和实践研究。”该定义结合了用户视角和技术视角,强调以系统和实践为中心的用户间的协同查寻和检索信息。

研究者们除了讨论协同信息检索的环境、动机和方法外,也构建了相关的系统,范围也涉及信息检索、信息收集和 Information sharing,有的还扩展到了信息合作及协同意义建构方面,在学术、工商业、军事和医学领域均有应用。