



# 用户信息搜索中的学习行为及过程探究<sup>\*</sup>

□刘畅 宋筱璇 杨子傲

**摘要** 用户信息搜寻的目的是为了学习新知识或完成某项任务,以往研究中学者主要关注搜索行为,但是对于用户如何使用搜索到的信息进行知识积累以完成学习任务的相关研究仍比较缺乏。对用户学习行为和过程加以探索,深入理解用户信息搜索过程中知识学习的发生、演进特征,有助于我们更好地设计支持用户学习任务下的信息搜索系统。本研究采用实验法,被试者在实验室环境下完成四个与学习相关的搜索任务,在搜索过程中要求被试者在电脑的记事本文件中记录和整理自己认为对完成任务有用的信息或答案。这种设计可使用户的学习过程通过其信息记录行为外显化,进而通过分析用户的信息记录行为特征,揭示用户的学习行为特征。在数据分析中,依据不同阶段用户记录完成度的推进情况,将他们的记录策略分为三种不同的模式:早期记录型、平均记录型、后期记录型。研究还探索了影响用户记录策略的因素,发现用户的自信程度、话题熟悉度、任务类型以及任务难度都会显著影响用户采取何种记录策略。在不同记录策略与搜索交互、任务产出的关系分析中发现,平均记录策略作为记录行为随搜索过程逐步推进的一般模式,学习产出的质量相对更好,因此相对而言更有效率。这些发现有助于深入理解用户的信息使用行为和学习发生过程,启发支持学习相关搜索的信息系统的设计与优化。

**关键词** 学习行为 记录策略 记录效果 学习型搜索

**分类号** G251

**DOI** 10.16603/j.issn1002-1027.2019.04.005

## 1 引言

随着知识经济时代的到来,知识已经成为一种重要的生产要素和社会资源。现代计算机科学技术也朝着服务人类学习、促进知识的增长与利用的方向发展。图书馆尤其是高校图书馆也在经历着巨大转变,大家普遍认为高校图书馆正在塑造学习和研究的未来<sup>[1-2]</sup>,图书馆不仅是用户获取信息资源的场所,更是支持用户学习和研究的主力,图书馆的信息搜索系统和信息服务都是为了支持高校师生的学习和研究。图书情报学自起源之时,就致力于帮助人们进行信息的存储、组织、查询和使用,通过实现从信息到知识的转化,辅助人们的决策与学习。信息爆炸的背景使得搜索工具在信息的识别和访问方面扮演着不可或缺的角色,信息搜索已经成为现代社会人们获取知识过程中一个必不可少的步骤。但

是当人们通过搜索系统获得信息之后,还需要对所获得的信息进行评估、理解和使用,生成学习产出。当前在信息行为研究领域,学者提出“搜索即学习”(Search as Learning),强调将信息搜索的过程看作用户自我学习的认知过程。目前的搜索系统在用户查找事实型信息方面表现较好,但对于查找过程中知识获取和输出方面仍处于初探阶段。因此本文选取用户搜索过程中的学习行为和过程作为研究对象,分析该过程的特点和影响因素及效果,以便对用户利用搜索系统进行学习、获取知识的全过程有更全面深入的理解,启发各类搜索系统从支持学习的角度进行设计和优化。

已有的信息行为研究,主要关注用户获取信息过程中的搜索行为、浏览行为等特征,如用户检索式的构造、重构行为,在搜索结果页、网页内容页的点

<sup>\*</sup> 本文系国家社会科学基金项目一般项目“学习型搜索中用户交互行为与学习效果关系研究”(项目编号:18BTQ090)和国家自然科学基金青年科学基金项目“基于用户检索行为和搜索任务情境的个性化信息检索系统研究”(项目编号:71303015)研究成果之一。

通讯作者:宋筱璇(ORCID:0000-0002-6589-4071),邮箱:songxiaoxuan@pku.edu.cn。



击、浏览行为等<sup>[3-6]</sup>。研究发现用户的信息搜索行为会受到诸多因素的影响,如用户个人特征、信息源的特征、搜索任务的特征等<sup>[7-9]</sup>,对用户搜索行为的分析有助于优化图书馆网站设计<sup>[10-11]</sup>。然而,搜索和浏览等行为仅是用户完成任务过程中的部分行为,用户在获得信息后,如何使用信息——对所获取的信息进行思考、理解,将信息与个人已有的知识进行整合,整理、表达、输出学习到的知识,这是当前评估搜索系统服务效能中更加需要关注的<sup>[12-13]</sup>。相对于用户的搜索行为,用户的学习行为和过程主要发生于用户的认知层面,具有内隐特征,较难进行观察和分析。本研究采用实验法,设计了与学习相关的实验搜索任务,要求用户在搜索过程中记录、撰写相应的笔记或文档来对信息进行整理和理解,以便完成学习任务。通过这种设计,使得用户的学习过程通过其信息记录行为外显化,通过分析用户的信息记录和整合过程,进而分析用户的学习过程。对用户搜索过程中何时开始记录及如何完成记录的探究,有助于理解用户在搜索过程中学习发生的过程,也会对支持学习型搜索的信息系统的设计与优化有所启发。基于此,本研究选择从信息记录行为的视角对用户搜索中的学习行为及过程展开深入研究,试图分析不同用户在搜索过程中记录行为的差别,探索导致这种差别的原因,以及这些差别对其学习效果的影响。

## 2 相关研究

### 2.1 学习型搜索

“搜索即学习(Search as Learning)”是将搜索过程概念化为用户通过多种搜索活动,对相关信息进行分析、评估和使用,创造出新知识的学习过程。“搜索即学习”这一概念不仅关注搜索输出结果(Search Outputs),更强调搜索过程中发生的学习,即用户的知识结构的变化,以及搜索任务的学习产出(Learning Outcomes)<sup>[12]</sup>。在“搜索即学习”这一主题下,衍生出对一种搜索任务类型的关注——“学习型搜索任务(Learning Related Search Task)”<sup>[12]</sup>。学习型搜索任务指的是用户在学习情境中,以学习为目的,利用搜索系统对信息进行收集、分析、评估和使用等,最终完成该学习任务的搜索活动。虽然学习型搜索这种提法提出的时间并不长<sup>[14]</sup>,但是对于非事实型任务的搜索行为分析是近十年来的研究

热点,如探索式搜索。探索式搜索区别于事实查找的搜索在于,除了需要查找基本的事实和信息,探索式搜索更侧重于知识的建构、概念的理解、观点的阐释和对比或数据概念的聚合等几方面<sup>[15]</sup>。学者认为学习型搜索与探索式搜索有很紧密的关联,区别是学习型搜索强调以学习为目的,需要对搜索到的信息有更加深入的理解,并重视通过搜索过程的知识建构情况及其产出的学习效果评估搜索系统的服务效能<sup>[12]</sup>。已有研究发现,在探索式搜索中,用户交互行为的特征和表现与事实型搜索有较大差异。例如袁红和李秋<sup>[16]</sup>的研究发现对于提问应答式搜索,用户倾向于将问题本身作为检索词;而在探索式搜索中,任务越复杂,用户越倾向使用检索词扩展进行搜索,且检索词数量越多,词频越低,词间的联系更不紧密。同时,探索式搜索中浏览网页的数量也比提问应答式更多,链接深度更深。张云秋等人<sup>[17]</sup>发现在探索式搜索中,随着搜索主题难度的增加,页面的平均浏览时间减少,更难集中在某一网页上进行精读。张云秋等人<sup>[18]</sup>也尝试对探索式搜索过程中用户知识结构的改变进行分析,通过被试者在搜索前后绘制的任务知识概念图,发现在探索式搜索任务中,搜索者的知识结构会有较大改变,搜索者在搜索过程中试图构建新的、复杂的和更为专业的知识结构。

从交互行为的角度看,学习型搜索中的交互活动类型更加复杂与多样。用户在完成任务的过程中,不仅仅要进行信息的搜索,还需要完成信息的阅读、理解、选择、比较等多种活动<sup>[19]</sup>。李秀英(S. Y. Rieh)等学者<sup>[12]</sup>提出在学习型搜索过程中,应该关注用户对于信息的分析、汇集、评价、使用等直接反映用户学习的活动。用户的记录行为是对已获得的信息的加工和使用行为,是用户对所获得的信息进行分析、理解,并将该信息与个人已有知识结构进行整合的关键点,在一定程度上显性代表了用户在学习型搜索过程中,学习行为的发生及过程特征。因此,在学习型搜索过程中,研究用户记录行为的相关特征很有必要。2017年在德国举办的信息技术前沿论坛——达堡研讨会(Dagstuhl Seminar)中,与会者针对学习型搜索的讨论,提出了该领域未来发展需要解决的8个主要研究问题:1)理解搜索作为人类的学习过程:在搜索过程中,学习是何时以及如何发生的? 2)学习过程和学习情境(教育相关的,工作



相关的)的关系是什么? 3)如何借鉴认知心理学、教育学的相关理论构建搜索交互过程中的学习、知识获取、心智表达等模型? 4)在信息系统交互层和算法层分别设立什么样的功能和干预能够增强搜索过程中用户的学习? 5)在搜索和与搜索相关的活动中,什么是有效的、可扩展的学习信号? 如何衡量和应用这些信号? 6)在搜索过程中如何测量学习效果和学习产出? 7)如何快速搜索,以节省更多的时间用于学习? 8)如何让搜索系统变得更加具备会话型特征,以增强搜索过程中的学习?<sup>[20]</sup>。本研究的目的是对第一个议题提出的问题展开探索,具体分析不同用户在搜索过程中如何展开记录活动,对记录活动的过程和效果以及搜索情境对记录活动的影响展开分析。

## 2.2 搜索过程中的学习行为

在已有的对搜索过程中交互行为的研究中,主要关注用户的搜索行为,例如与检索式相关的行为和策略<sup>[4,6]</sup>,与搜索结果页相关的交互行为等<sup>[5]</sup>,但对于搜索过程中学习行为的相关探索还比较缺乏。其中主要的原因是学习过程是一种认知过程,发生在用户的内隐知识层面,很难观察和研究,但是也有一些研究通过用户在内容页面上的阅读和浏览行为来推断用户的认知过程和信息使用行为的特征及其影响因素。例如学者金正玄(J. Kim)<sup>[3]</sup>发现在探索式搜索中,用户感知到的任务难度越大,在搜索过程中浏览的内容页数量越多。张云秋等人<sup>[17]</sup>进一步研究发现,用户内容页的平均浏览时间减少,更难集中在某一网页上进行精读。也有研究通过搜索过程中的记录行为来分析用户的信息使用和学习过程,但目前这种研究仍然很有限,如刘婧婧和贝尔金(J. Liu 和 N. J. Belkin)<sup>[21]</sup>关注任务过程中的搜索行为和记录行为,并探索了影响这两类行为的因素,发现用户的话题熟悉度对记录效率有积极的正向作用。刘涵蕊和刘畅<sup>[13]</sup>分析了用户的认知风格和话题熟悉度对其搜索过程中的搜索、阅读和记录行为的影响,发现用户认知风格主要影响用户的搜索行为,而搜索前的话题熟悉度主要影响用户的阅读和记录行为。波特哈斯特和哈根(M. Potthast 和 M. Hagen)等人<sup>[22-23]</sup>探索了针对不同的主题,文章撰写者如何利用 Clueweb09 搜索以完成写作的过程,并根据撰写者每次编辑后已完成的文本的长度,识别出两类不同的文本记录策略:积累型(build up)和精炼型

(boil down)。其中采用积累型策略的撰写者在整个写作过程中不断增加文章的长度,而采用精炼型策略的撰写者在写作前期先大量积累素材,记录的文本长度呈现快速增长,接着才对文本进行重新组织,缩短文章的长度。他们还比较了写作过程不同阶段中文章撰写者搜索行为上的差异。他们将任务总时间划分为三个阶段:早期、中期、后期,结果发现用户在早期阶段构造了更多的检索式,并有更多的点击活动发生,而用户的记录活动更主要地集中在后期阶段。然而,该文没有对何种因素可能影响用户的记录策略做进一步的探索,也没有分析记录策略与用户学习成果或任务产出之间的关系。本研究将借鉴这个研究的分析思路,将搜索任务总时间分为若干阶段,尝试识别不同用户在搜索过程中的记录策略,并着重分析哪些因素会影响用户记录策略的选择,以及用户的记录策略与记录效果之间的关系。

## 3 研究设计

### 3.1 用户实验

本研究采用实验法,来自北京大学的40名本科生参与实验,其中男生20人,女生20人。他们来自人文类、社会科学类、自然科学类、工科类和医学类等不同院系,且分布于不同年级:大一5人,大二11人,大三15人,大四9人。每位被试者被单独邀请到实验室,需要在实验室的台式电脑上搜索指定的4个搜索任务。实验开始前,被试者首先填写一份个人背景信息问卷。在每个任务开始前,被试者需要阅读该任务的任务描述,并填写一份搜索前问卷,对该任务下被试者的话题熟悉度、完成任务的自信程度等进行量表评估。搜索过程中,被试者需要记录和整理自己认为合理的答案,并将其记录在电脑的记事本文件中。被试者提前被告知,他们撰写的记事本文档内容将在实验完成后被评估,质量最好的文档将会得到额外的奖励。在完成实验任务或任务时间用尽后,搜索停止,被试者填写搜索后问卷对该任务中的表现和体验进行评价。直至4个搜索任务全部完成后,被试者填写实验后问卷,评价实验全过程的体验,实验结束。实验中利用 Morae Recorder 在后台记录用户的搜索交互行为。

#### 3.1.1 实验任务设计

本实验借鉴了吴宛青(Wan-Ching Wu)等人<sup>[24]</sup>



的方法,从学习认知的复杂度的角度设计搜索任务。吴宛青(Wan-Ching Wu)等人根据<sup>[25]</sup>的学习认知过程的前5个维度设计搜索任务,发现随着维度的升高,完成任务所需的认知和努力都随之增加<sup>[24]</sup>。其中,根据“记忆”维度设计的搜索任务是让搜索用户查找若干个事实的具体内容,与学者之前提出的“事实型搜索任务类型”非常接近。考虑到本研究会对被试者的搜索时间有严格限制,因此选择了认知过程的前两个维度——记忆和理解,并参考前人的研究设计了两种任务类型:(1)事实型信息查找任务(Fact Finding,简称FF):搜索特定、具体的事实和信  
息;(2)理解型信息查找任务(Information Understanding,简称IU):搜索过程中需要理解和总结找到的信息来获取知识。每种任务类型各有两个不同主题的搜索任务,共计4个搜索任务,具体描述如下:

FF-1:听说印度婚礼很有趣也有很多的民俗特点,请你尝试着搜索印度婚礼的风俗、特点和流程,比如婚礼前的准备、新郎新娘的服饰、手绘的特点以及婚宴食物的准备等印度婚礼中特别的环节?

FF-2:你的一个高中同学说他被隐翅虫咬了,感觉皮肤发痒,痒的地方用手挠,结果还化脓了。你听说之后很担心,就想上网查一下什么是隐翅虫,有没有毒,遇到隐翅虫应该怎么办?如果被隐翅虫咬了应该怎么治疗?

IU-1:你的侄子正在考虑加入学校足球队,你的大部分亲戚均对这个主意感到赞同。但你认为足球是一项危险的运动,并且你担心会有潜在的健康风险。因此你需要知道,长期踢足球是否对青少年的健康有危害?

IU-2:多立克柱式是古希腊建筑中很有特色的一种建筑形式。请尝试搜索多立克柱式的一般特点和典型的代表作品,并查找多立克柱式是否对中国建筑产生影响,有哪些表现?

### 3.1.2 实验中的时间限制

为了探究时间限制对用户搜索过程中记录策略的影响,实验中设计了两种时间情境:有时间限制(Time Constraint,简称TC)和没有时间限制(No Time Constraint,简称NTC)。在NTC情况下,当被试者认为自己已经查找到并记录了足够的信息时,可以随时停止搜索。而在TC情况下,被试者搜索每个任务的时间为5分钟。通过预实验的实验结

果发现,被试者一般需要花费10分钟左右来完成任务,因此在本实验中,将TC情境的时间控制在较合理同时又能触发一定压力的5分钟。

针对任务类型(FF和IU)和时间限制(TC和NTC)两个因素,本实验采用2X2因素组内设计,即每一位被试者在实验中需要在两种不同时间压力的情况下完成两类搜索任务。时间限制和搜索任务的次序采用拉丁方进行充分轮换,即被试者中的一半先在TC情况下搜索两类任务,然后在NTC情况下搜索另两类任务;另外一半则相反,先进行NTC情况下的另两类任务,再进行TC情况下的两类任务。在实验的中场时间,即被试者们完成两个搜索任务后,可进行5分钟左右的短暂休息。休息结束后,改变时间限制的情况,被试者们继续在改变后的时间情况下完成剩余的两个搜索任务。

### 3.2 相关变量

本研究的实验一共有40个被试者参与,每个被试者搜索4个任务,总计收集了160个搜索会话。但实验数据中有3个被试者的问卷数据丢失,另外有2个会话的交互行为日志数据丢失。去掉这些缺失数据,本研究的有效分析数据包含146个搜索会话。在分析过程中,本研究主要测量和分析了情境变量、交互行为变量和记录效果变量与用户记录策略之间的关系。下面逐一介绍各变量的具体内容和测量方法。

#### 3.2.1 情境变量

为了探索用户的记录策略会受到哪些情境因素的影响,本实验中主要控制和测量了两大类的情境变量:个人特征相关变量和搜索任务特征相关变量。用户个人特征变量主要包括用户在搜索前的话题熟悉度和搜索前的自信程度(这两个变量均通过搜索前问卷李克特五点量表测量);搜索任务特征相关变量主要包括任务类型(事实型FF和理解型IU)、时间限制(有时间限制TC和无时间限制NTC)、任务难度(通过搜索后问卷李克特五点量表测量)。这些情境变量的描述和具体分组如表1所示。

表1 情境变量及其测量分组方式

情境变量	变量测量及分组方式
个人特征	话题熟悉度:熟悉度低(1-2)N=113;熟悉度高(3-5)N=33
	自信程度:自信程度低(1-3)N=64;自信程度高(4-5)N=82



情境变量	变量测量及分组方式
任务特征	任务类型: FF N=71; IU N=75
	时间限制: TC N=75; NTC N=71
	任务难度: 任务难度低(1-3)N=89; 任务难度高(4-5)N=57

### 3.2.2 交互行为变量

本研究还考量了用户的记录策略与其交互行为变量之间的关系,旨在分析用户的记录策略与其搜索行为和记录行为之间的联系,目的是探究用户的交互行为变量能否从一定程度上反映其记录策略。本研究中的交互行为变量包括用户记录过程中和搜索过程中的键盘活动、鼠标点击活动、鼠标滚轮活动等交互活动的次数。

- (1) 记录过程中交互活动次数。
- (2) 记录过程中键盘活动次数。
- (3) 记录过程中鼠标点击活动次数。
- (4) 记录过程中鼠标滚轮活动次数。
- (5) 搜索过程中交互活动次数。
- (6) 搜索过程中键盘活动次数。
- (7) 搜索过程中鼠标点击活动次数。
- (8) 搜索过程中鼠标滚轮活动次数。

### 3.2.3 记录效果变量

在用户记录的效果评估方面,我们主要参考了<sup>[21]</sup>对用户搜索产出的评估方式,选取了包括文档中记录的段落数、句子数、词数、分面数等变量来对用户搜索过程中的记录文档进行分析。这四个指标中,段落数和句子数通过段落符号和句号进行划分计算;字数则通过中文字数和非中文单词的词数的总和计算得到;分面数的计算参考了宋筱璇,刘畅此前用过的方法<sup>[26]</sup>,由两名研究人员对被试者记录的文档内容进行人工编码,并通过一致性检验,旨在统计用户在记录中主要从哪些方面对搜索主题进行了记录和整理。

## 4 分析结果

### 4.1 用户记录策略特征

为了明晰学习型搜索过程中用户的记录策略特征,本文提出以“记录策略曲线”来刻画用户记录的完成度随任务总时间的推进情况。如图 1 所示,曲线 W0-W1-W2-W3-W4 就是一条完整的用户记录策略曲线。该曲线的 X 轴为用户完成该任务

所用的总时间,为了统一不同时长的搜索会话,所有的搜索会话都被平均分成四个阶段,分别是 25%、50%、75%、100%,以分析用户的记录完成情况是如何随着任务总时间的这四个阶段的推进而累积的,因此该曲线的 Y 轴为当前时间阶段下用户的记录完成度(1)。

记录完成度 = 用户已经记录的时间 / 该任务下记录总时间 ..... (1)

本研究截取了用户在搜索全过程中与记事本文档交互的所有时间,作为用户完成任务的记录总时间。记录完成度为在当前时间阶段下,用户已经与记事本文档交互的时间占该任务中与记事本文档交互总时间的百分比。

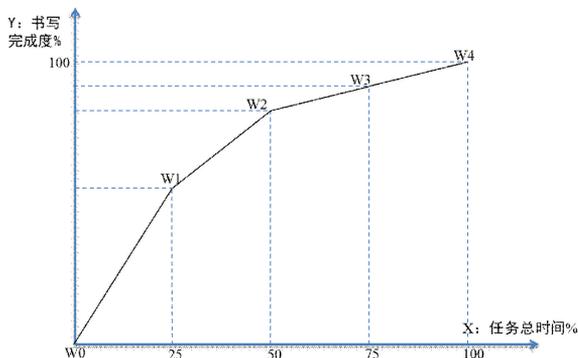


图 1 记录策略曲线示例

进一步地,本研究在“记录策略曲线”的基础上,计算了“记录策略面积”,以便比较用户在任务过程中的记录策略特征的差异。记录策略面积计算的是记录策略曲线与 X 轴之间的面积值(2),体现了用户的记录活动是如何随着任务总时间而推进的。

记录策略面积 =  $\int_0^1 f(x)$  注:  $f(x)$  表示用户该任务中的记录策略曲线 ..... (2)

整体而言,本研究分析的 146 个任务会话中,用户的记录策略面积分布在 [0.125, 0.875] 之间,平均值为 0.436,方差为 0.137。这个分布表明,对于不同的任务会话,用户的记录活动随任务总时间的推进有很大的不同:在一些会话中,用户倾向于在较早的任务阶段开始相应的记录活动,在任务的前期,记录完成度已经达到一个较高的值,这类会话的记录策略面积相对较大;在一些会话中,用户倾向于在相对靠后的任务阶段开始相应的记录活动,记录完成度在任务后期有明显上升,这类会话的记录策略面积相对较小;还有部分会话,用户的记录活动伴随着任

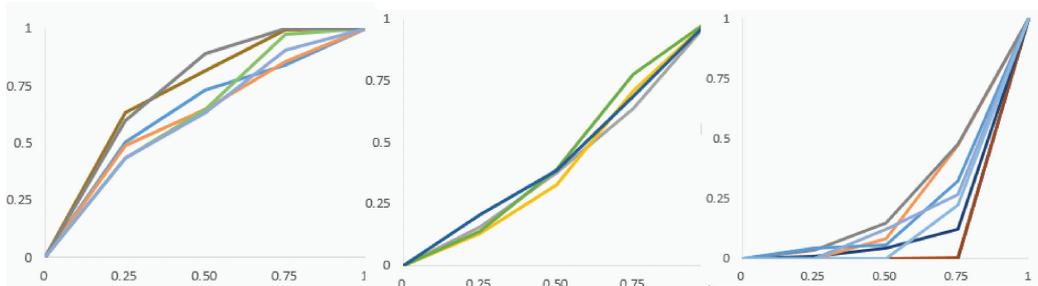


图2 部分记录策略曲线实例(左:早期记录型;中:平均记录型;右:后期记录型)

务总时间阶段逐渐推进,记录完成度呈现逐步上升的趋势,这类会话的记录策略面积大小居中。基于用户记录策略曲线的以上特征,本研究中依据记录策略面积的三分位点,将用户在学习型搜索过程中的记录策略划分为三种类型,分别为:早期记录型(N=49;面积大小∈(0.484, 0.875])、平均记录型(N=48;面积大小∈[0.370, 0.484])、后期记录型(N=49;面积大小∈[0.125, 0.370))。各种类型的记录策略曲线实例如图2所示。

#### 4.2 个人特征对记录策略的影响

在识别出用户记录策略的三种类型后,我们接着探究了哪些因素会影响用户的记录策略类型。该部分主要验证记录策略是否会受到用户个人特征的影响,包括话题熟悉度和搜索前的自信程度。由于该组变量与用户的记录策略均属于定类变量,在检验方法上采用卡方检验。

首先,结果显示用户搜索前的自信程度显著影响他们的记录策略( $X^2 = 21.9, p = 0.001$ )。具体如图3所示,搜索前自信程度较高的用户更倾向于采用后期记录型策略(41.5%),而自信程度较低的用户仅有21.9%选择后期记录型策略;搜索前自信程度较低的用户比起自信程度较高的用户更倾向于采用早期记录型策略(42.2% vs. 26.8%)。这个结果显示了用户搜索前的自信程度越低,他们越倾向于更早地开始记录活动。

其次本研究发现,用户搜索前的话题熟悉度会显著影响他们的记录策略( $X^2 = 19.19, p = 0.004$ )。搜索前话题熟悉度较低的用户,其记录策略则更倾向于采取早期记录型(37.2% vs. 21.2%);而搜索前话题熟悉度较高的用户,则更倾向于采取后期记录型(42.4% vs. 30.1%)。

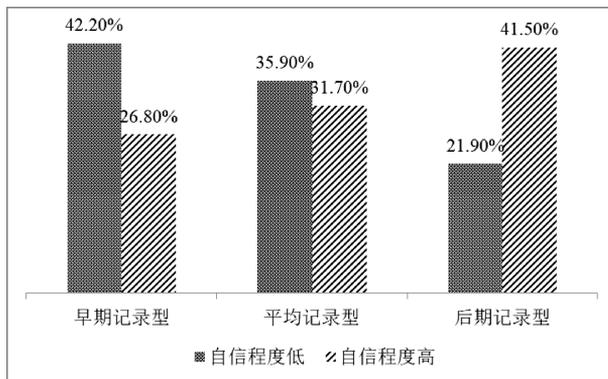


图3 用户自信程度对记录策略的影响

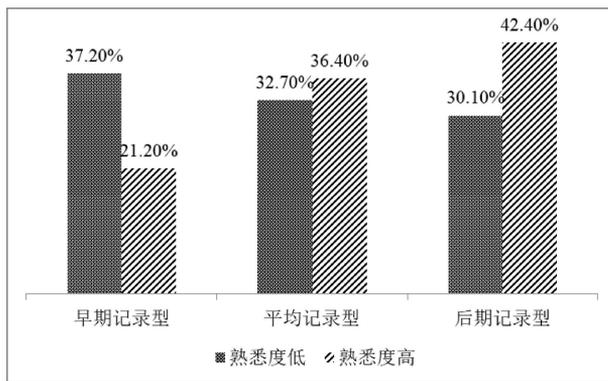


图4 用户的话题熟悉程度对记录策略的影响

#### 4.3 任务特征对记录策略的影响

在任务特征方面,本文重点关注任务类型、任务难度和时间限制对用户记录策略的影响。由于该组变量与用户的记录策略均属于定类变量,在检验方法上采用卡方检验。

结果显示,任务类型显著影响用户的记录策略( $X^2 = 13.388, p = 0.004$ )。如图5所示,面对理解型信息查找任务(IU),用户更倾向于选择后期记录型策略(42.3% vs. 25%)。这可能是由于在用户完成IU任务的过程中,在记录活动开始之前,需要进行更多的信息收集,并在此基础上开展理解、分析、比较等认知活动;而面对事实型信息查找任务

(FF),由于认知复杂度较低,用户喜欢在信息获取的同时进行相应的记录活动。

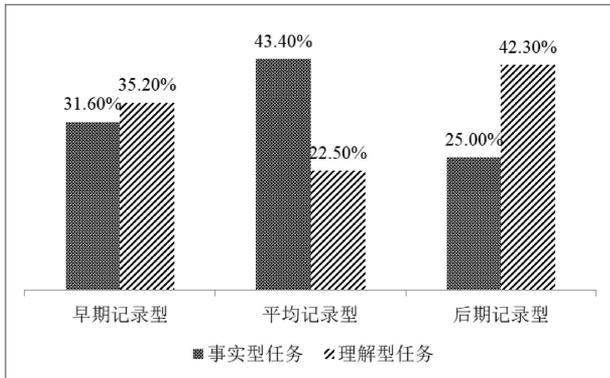


图 5 任务类型对记录策略的影响

在任务难度方面,任务难度显著影响用户的记录策略( $X^2 = 25.216, p < 0.0001$ )。如图 6 所示,在相对简单的任务中,用户更倾向于选择平均记录策略(42.7%),而对于相对困难的任務,只有 19.3%的用户选择了该类型。这个结果显示平均记录策略是用户在学习型搜索过程中的一般记录模式,但当用户感知到搜索任务变得更为困难时,可能会打破这种循序渐进的记录模式,在搜索任务的前、后阶段做更多的记录努力,即转变为早期或后期的记录策略。

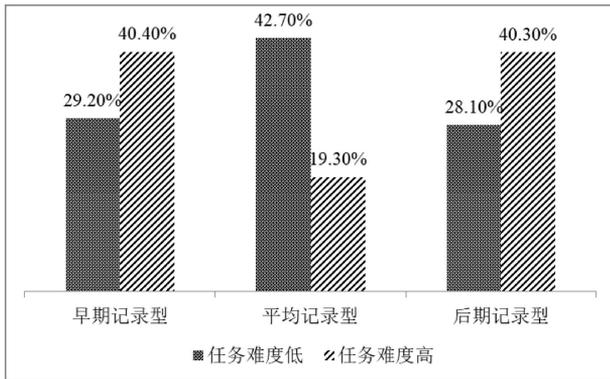


图 6 任务难度对记录策略的影响

在时间限制方面,本研究比较了有无时间限制(TC 和 NTC)这两种类型的时间情境,尝试探索在学习型搜索过程中,用户是否会调整其记录策略以适应 TC 和 NTC 两种不同的时间情境。但是,结果显示,时间限制对于用户记录策略没有显著的作用( $X^2 = 3.04, p = 0.39$ )。

#### 4.4 用户记录策略与搜索交互行为的关系

为了探索学习型搜索过程中,用户的搜索行为是否能在一定程度上反映出他们所采用的记录策

略,本文进一步分析了用户的记录策略与搜索交互行为之间的关系。由于已有研究发现,用户的搜索交互行为的数量在是否有时间限制的情境下存在很大的差异<sup>[27]</sup>,因此,在本文的分析中,分别对 TC 和 NTC 两种时间情境下,采取不同记录策略的用户的搜索交互行为变量进行比较,主要采用 Kruskal-Willis 秩和检验进行统计检验。

研究发现,当用户处于 TC 的时间情境中时,不同记录策略的会话在搜索交互行为上没有显著差异;然而当用户没有时间限制时(NTC),用户会话中的记录策略与过程中的交互行为显著相关,具体表现在搜索过程中交互活动次数、键盘敲击次数、鼠标点击次数等方面。如表 2 所示,采用平均记录型的用户在搜索过程中的交互活动次数、键盘敲击次数以及鼠标点击次数上,显著少于早期记录型和后期记录型用户。该结果显示,当没有时间限制时,采用平均记录型的会话的任务搜索过程相对简单,相较于早期和后期记录策略的会话,在采取平均记录策略的会话中,用户花费更少的交互努力来完成当前任务。这个结果从一定程度上说明,平均记录策略是相对高效的搜索与记录交互的模式。

表 2 无时间限制下(NTC)三种记录策略的用户交互行为比较(均值)

搜索交互变量	早期记录型	平均记录型	后期记录型	<i>p</i>
写文档过程中交互活动次数	278	227	317	0.213
写文档过程中键盘活动次数	130	141	221	0.265
写文档过程中鼠标点击活动次数	37.9	28	28.3	0.382
写文档过程中鼠标滚轮活动次数	109.6	58	68.1	0.181
搜索过程中交互活动次数	1145	786	1193	0.012 *
搜索过程中键盘活动次数	329	146	262	0.042 *
搜索过程中鼠标点击活动次数	161	119	161	0.008 *
搜索过程中鼠标滚轮活动次数	651	526	769	0.073

注:标\*为  $p < 0.05$  的情况



#### 4.5 用户记录策略与记录效果的关系

在学习型搜索中,除了关注搜索过程中的交互行为,也要关注用户搜索中的记录效果,在一定程度上,记录效果是搜索效果的一种表现。本研究关注用户记录的文档中的内容,包括记录的段落数、句子数、字数、分面数等等,采用 Kruskal-Willis 秩和检验分析方法,分析用户不同的记录策略是否影响用户学习型搜索的记录效果。

结果如表 3 所示,用户的记录策略显著影响任务输出文档中的句子数量( $p=0.005$ )。采取早期和平均记录策略的会话,输出文档中的句子数量较多,平均 36 个左右;而采取后期记录策略的会话,在输出文档中的句子数量最少,平均约 23 个。此外,尽管三种不同类型的记录策略,在输出文档的段落数、字数和分面数上没有显著差异,但从均值上仍可以大致看出,采取后期记录策略的会话输出的段落数、句子数、字数以及分面数总是三种策略中最少的;采取平均记录型的会话产出的句子数、字数都居中,但是产出的段落数和分面数却是最多的。总体上,从任务产出的角度来看,用户在会话中相对早地开始记录活动有利于记录更多的内容,后期记录策略记录的内容相对较少。如果结合上一节提到的交互次数的变量,可以看出平均记录策略花费的交互次数最少,但是记录产出却最多,相对记录效果最好,也说明平均记录策略是三种策略中比较高效和有效的记录策略。

表 3 三种记录策略的记录效果比较(均值)

记录效果变量	早期记录型	平均记录型	后期记录型	$p$
文档中的段落数	10.59	13.49	8.53	0.071
文档中的句子数	36.39	36.02	23.3	0.005 *
文档中的字数	1747	1693	1160	0.182
文档中的分面数	3.67	3.96	3.55	0.067

注:标\*为  $p < 0.05$  的情况

## 5 讨论

本研究从用户记录策略的角度,对搜索过程中用户的学习行为的特征及其影响因素进行探索,依据不同任务时间点用户记录完成度的推进情况,将用户搜索会话的记录策略分为三种不同的模式:早期记录型、平均记录型、后期记录型。研究的目的是识别哪些情境变量会影响用户记录策略的选择,研

究发现,用户的自信程度、话题熟悉度、任务类型以及任务难度都显著影响用户采取何种记录策略。具体而言,用户搜索前的自信程度越低、话题熟悉程度越低,越倾向于较早地开始记录活动;自信程度越高,话题熟悉程度越高,越倾向于较晚地开始记录活动。这是比较能够理解的,因为越不自信的用户或者越是认为自己对该搜索主题不熟悉的用户,越希望通过搜索较好地完成任务,也会相对较早地开始学习活动。另外,任务类型也会影响用户的记录策略选择。在事实型信息查找任务中,用户倾向于采取平均记录策略;而在理解型信息查找任务中,用户倾向于采取后期记录策略。当用户感知搜索任务难度一般时,倾向于使用平均记录策略,而当他们感知难度提升时,记录策略将更多地转变为早期或后期记录型。总体而言,当任务不需要较多的认知负担,或搜索任务难度不高时,用户倾向于选择平均型的记录策略。库尔斯奥(C. C. Kuhlthau)的信息查询过程模型(ISP)<sup>[28]</sup>提出,用户在搜索过程中的行为转变与认知、情感的变化是紧密联系的。在本研究中,任务类型(事实型和理解型)的差异体现在用户的认知复杂度上<sup>[24]</sup>,任务难度的感知也是用户认知的一种体现。由此可知,平均记录策略反映的是用户普遍的正常的记录模式,但在用户搜索过程中,当认知复杂度增加或感知难度增加时,用户的记录策略都可能发生相应的变化,由平均记录模式转变为早期或后期的记录模式。但是本研究没有发现时间限制对用户记录策略产生显著影响。与之前的研究<sup>[29]</sup>发现时间限制并没有改变用户在搜索、阅读和记录上的时间比例的结论有一致性,这可能是因为记录行为与人们的认知习惯紧密结合,不太容易在短期内做出较大改变。这个结论也有待后续研究进一步验证。

除此之外,本研究还分析了不同记录策略与搜索交互、记录效果之间的关系,结果发现在有时间限制的条件中,不同记录策略下的搜索交互行为没有差异;但是在没有时间限制的条件下,采用平均记录型的会话的交互活动次数、鼠标点击活动次数、滚轮活动次数等都显著少于其他两种类型,这一发现说明,平均记录型是一种更具有效率的记录模式,在该模式下,用户完成同样的任务,但花费了更少的搜索努力。在记录效果方面,尽管采用早期记录策略的会话在输出文档中的句子数多于其他两种记录策略



的会话,但是采用平均记录策略的会话却能在输出文档的内容中覆盖最多的分面。设立分面的覆盖情况这个指标,试图通过理解内容来评估该搜索的任务产出效果,很多关注学习型搜索中知识产出评价的研究也都使用了这个指标<sup>[26,30]</sup>,比单纯计算句子数量更具有意义。因此这个结果也说明,平均型记录策略是一种相对高效、有效的记录策略。

## 6 结论

本研究采用实验法,从研究用户记录策略的角度,探索学习型搜索过程中用户的学习行为相关特征,发现依据用户的记录完成度随任务总时间阶段的推进情况,用户的记录策略可划分为三种类型:早期记录型、平均记录型和后期记录型。进一步研究发现,用户采取不同的记录策略类型受到用户个人特征和搜索情境的影响。搜索前自信程度较高的用户、对搜索话题更熟悉的用户,搜索理解型任务的用户,都倾向于选择后期记录策略;对于事实型信息查找任务和当用户感知搜索任务较为简单时,他们更可能采用平均记录策略;当任务的难度较大时,在搜索过程中,用户对信息的记录更可能调整为早期或后期记录型。本研究还发现三种记录策略中,采用平均记录策略的用户的记录行为随搜索过程逐步推进,并且这种记录策略更有效率,在学习产出质量上也相对更好。在搜索过程中,特别是与学习相关的搜索过程中,用户的记录策略是一个重要的行为特征,有助于深入理解用户的信息使用行为和学习发生的过程,对于系统感知用户的学习特征,识别学习中的困难并提供相应的帮助,具有重要的意义,值得进一步深入研究。另外,目前大多数搜索系统都仅提供搜索服务功能,还需在用户搜索过程中的信息使用和笔记撰写等环节提升服务,帮助用户在搜索过程中积累知识。尤其是在现实学习工作中,用户可能会通过一段时间的搜索,不断累积知识,来完成相关的学习任务,如撰写学期论文或毕业论文的过程可能会持续若干天或更长时间,更需要有辅助记录和学习的功能来更好地提高效率。同时通过这种辅助软件记录行为,也可以更好地在自然状态下了解用户的学习进度,分析搜索行为和学习过程之间的关系,以便优化搜索系统,为用户完成学习型搜索任务提供服务和帮助。

## 参考文献

- 1 Shaw, C. University libraries are shaping the future of learning and research [EB/OL]. <https://www.theguardian.com/higher-education-network/blog/2013/aug/06/university-libraries-learning-shapes-design>, 2018-07-06.
- 2 Fallin L. Beyond books: the concept of the academic library as learning space[J]. *New Library World*, 2016, 117(5/6): 308-320.
- 3 Kim J. Task difficulty in information searching behavior: expected difficulty and experienced difficulty[C]//Proceedings of the 5th ACM/IEEE-CS joint conference on Digital libraries. New York, NY, USA: ACM, 2007:383-383.
- 4 White R W, Dumais S T, Teevan J. Characterizing the influence of domain expertise on web search behavior[C]// Proceedings of the Second ACM International Conference on Web Search and Data Mining. New York, NY, USA: ACM, 2009:132-141.
- 5 Jiang J, He D, Allan J. Searching, browsing, and clicking in a search session: changes in user behavior by task and over time [C]//Proceedings of the 37th international ACM SIGIR conference on Research & development in information retrieval. New York, NY, USA: ACM, 2014:607-616.
- 6 Monchoux S, Amadieu F, Chevalier A, et al. Query strategies during information searching: Effects of prior domain knowledge and complexity of the information problems to be solved[J]. *Information Processing & Management*, 2015, 51(5):557-569.
- 7 丁韧. 网络信息搜索行为影响因素研究:基于高校学生的实证[J]. *图书情报工作*, 2012, 56(6):42-45.
- 8 石长翼. 高校学生学术信息搜索行为特征及影响因素研究[D]. 北京大学, 2011.
- 9 查先进, 张晋朝, 严亚兰. 微博环境下用户学术信息搜寻行为影响因素研究——信息质量和信源可信度双路径视角[J]. *中国图书馆学报*, 2015, 41(3):71-86.
- 10 姜婷婷, 陈舜昌, 高慧琴. 大学图书馆 OPAC 系统用户信息搜寻路径的可视化分析[J]. *大学图书馆学报*, 2017, 35(1):63-71.
- 11 周义刚, 聂华, 韦成府, 吴越, 张轶雯, 张乃帅. 新信息环境下用户需求驱动的图书馆门户设计与实现——以北京大学图书馆为例[J]. *大学图书馆学报*, 2014, 32(1):71-77.
- 12 Rieh S Y, Collins-Thompson K, Hansen P, et al. Towards searching as a learning process: A review of current perspectives and future directions [J]. *Journal of Information Science*, 2016, 42(1):19-34.
- 13 刘涵蕊, 刘畅. 认知风格与话题熟悉度对学习型任务下搜索交互行为的影响研究[J]. *情报理论与实践*, 2018, 41(04):56-62.
- 14 Rieh S Y, Gwizdka J, Freund L, et al. Searching as learning: Novel measures for information interaction research[J]. *Proceedings of the American Society for Information Science & Technology*, 2014, 51(1):1-4.
- 15 Marchionini G. Exploratory search: From finding to understanding[J]. *Communications of the ACM*, 2006, 49(4):41-46.



- 16 袁红,李秋.搜索任务和搜索能力对用户探索式搜索行为的影响研究[J].图书情报工作,2015,59(15):94-105.
- 17 张云秋,安文秀,于双成.探索式搜索中用户认知的实验研究[J].情报理论与实践,2013,36(6):73-77.
- 18 张云秋,安文秀,冯佳.探索式信息搜索行为研究[J].图书情报工作,2012,56(14):67-72.
- 19 Vakkari P. Task-based information searching [J]. Annual Review of Information Science & Technology, 2003, 37(1):413 - 464.
- 20 Collins-Thompson K, Hansen P, Hauff C. Search as Learning (Dagstuhl Seminar 17092)[C]//Dagstuhl Reports. Schloss Dagstuhl-Leibniz-Zentrum fuer Informatik, 2017, 7(2).
- 21 Liu J, Belkin N J. Searching vs. writing: Factors affecting information use task performance[C]// Proceedings of the American Society for Information Science & Technology, 2012, 49(1):1 - 10.
- 22 Potthast M, Hagen M, Völske M, Et Al. Crowdsourcing Interaction Logs to Understand Text Reuse from the Web[C]// Meeting of the Association for Computational Linguistics, 2013: 1212-1221.
- 23 Hagen M, Potthast M, Gomoll J, et al. How Writers Search: Analyzing the Search and Writing Logs of Non-fictional Essays [C]// Proceedings of the 2016 ACM on Conference on Human Information Interaction and Retrieval. New York, NY, USA: ACM, 2016:193-202.
- 24 Wu W, Kelly D, Edwards A, Et Al. Grannies, Tanning Beds, Tattoos and NASCAR: Evaluation of search tasks with varying levels of cognitive complexity[C]// Proceedings of the 4th Information Interaction in Context Symposium. New York, NY, USA: ACM, 2012: 254-257.
- 25 Anderson L W, Krathwohl D A. A taxonomy for learning, teaching and assessing: A revision of Bloom's taxonomy of educational objectives[M]. New York: Longman, 2001: 67-68.
- 26 宋筱璇,刘畅.搜索前后用户知识水平的评估及其变化情况析[J].图书情报工作,2018(2):108-116.
- 27 刘畅.时间限制对用户搜索交互行为的影响及其预测[J].情报学报,2016,35(02):127-136.
- 28 Kuhlthau C C. Inside the search process: Information seeking from the user's perspective[J]. Journal of the Association for Information Science & Technology, 1991, 42(5): 361-371.
- 29 刘畅,张璐.时间限制和搜索任务类型对搜索体验的影响分析[J].现代图书情报技术,2015,31(9):1-8.
- 30 Wilson M J, Wilson M L. A comparison of techniques for measuring sensemaking and learning within participant-generated summaries[J]. Journal of the American society for information science & technology, 2013, 64(2):291-306.

作者单位:北京大学信息管理系,北京,100871

收稿日期:2018年8月28日

## Characterizing and Exploring Users' Learning Behaviors and Process during Information Search

Liu Chang Song Xiaoxuan Yang Ziao

**Abstract:** Users search information in order to learn new things or to complete a task. Previous studies mainly focused on users' information search behavior, and it has not been extensively examined how users would use the information they have obtained for accumulating knowledge in order to accomplish learning-related tasks. In this study, we conducted a controlled user experiment, assigned four learning-related tasks for participants to search and asked them to write down notes on the computer during search. Users' learning process was manifested through their writing behavior, so that users' learning behavior characteristics were revealed by analyzing their writing behavior characteristics. Based on their writing process along with searching, their search sessions were divided into three types: early writing strategy, average writing strategy and late writing strategy. After characterizing three writing strategies, we further examined which factors were related to searchers' writing strategies. It was found that searchers' writing strategies were influenced by users' pre-confidence, topic familiarity, type and difficulty of the task. The examination of search interactions and task outcome indicated that "average writing strategy" was the normal pattern and also the most efficient strategy with best quality outcome. These findings help understand searchers' information use behavior and learning process while searching, and have implications on search system design to support learning-related search task accomplishment.

**Keywords:** Learning Behavior; Writing Strategy; Writing Outcome; Learning-Related Search