



开放科学  
(资源服务)  
标识码  
(OSID)

# 多模态信息搜索行为实验研究

朱佳雨<sup>1</sup> 李志义<sup>2</sup> 王振芳<sup>3</sup>

1. 中国科学院庐山植物园 九江 332900;
2. 华南师范大学经济与管理学院 广州 510006;
3. 福建师范大学文学院 福州 350007

**摘要:** [目的/意义] 研究多模态信息搜索特征有利于发现用户在图文音频信息搜索过程中的特征,进一步丰富多模态信息搜索以及信息行为的理论研究,为多模态信息系统建设和多模态信息素养提升提供依据。[方法/过程] 本研究依托经典的信息搜索行为理论,通过实验法和统计分析法,分析受过高等教育群体的多模态信息搜索行为特征,提出多模态信息系统开发与设计和多模态信息素养提升建议。[结果/结论] 通过实验研究,发现目前多模态信息搜索在准确性上表现较好,而完整性上则表现较差。在需求阶段需求理解上比较费时,普遍存在重复确认现象。在搜索方式上仍以文字搜索和跨平台搜索为主,且搜索策略调整次数较多,在搜索模态上转变较频繁。在信息收集上集中于百科类网页,特别是百度百科占了绝大多数,而社交媒体和电商应用软件也成为用户的多模态信息来源之一。

**关键词:** 多模态信息; 信息搜索; 信息行为

**中图分类号:** G35

## Experimental Research on Multi-modal Information Search Behavior

ZHU Jiayu<sup>1</sup> LI Zhiyi<sup>2</sup> WANG Zhenfang<sup>3</sup>

1. Lushan Botanical Garden, Chinese Academy of Sciences, Jiujiang 332900, China;
2. School of Economics and Management, South China Normal University, Guangzhou 510006, China;
3. College of Literature, Fujian Normal University, Fuzhou 350007, China

**Abstract:** [Objective/Significance] Research on the characteristics of multimodal information search is helpful to discover the characteristics of users in the process of audio-visual information search, and further enrich the multimodal information. The theoretical research of information search and information behavior provides the basis for the construction of multimodal information system and the improvement of multimodal information literacy. [Method/process] This study is based on Based on the classical information search behavior theory, this paper analyzes the multi-modal information search behavior characteristics

**基金项目** 广东省哲学社会科学规划 2023 年度学科共建项目《面向图书馆智慧参考咨询服务的多模态知识融合研究》(批准号:GD23XTS09)。

**作者简介** 朱佳雨(1993-), 硕士研究生, 主要从事信息搜索、学术评价和科技情报等研究。李志义(1968-), 副教授, 硕士, 硕士生导师, 主要从事 Web 挖掘研究, E-mail: Leeds@senu.edu.cn。王振芳(2000-), 江西九江人, 硕士, 研究方向为中国古代科技史和中国古代文学。

**引用格式** 朱佳雨, 李志义, 王振芳. 多模态信息搜索行为实验研究 [J]. 情报工程, 2023, 9(4): 54-67.

of the group with higher education through experimental method and statistical analysis, and puts forward multi-modal. The development and design of the state information system and the suggestion of improving multimodal information literacy. [Result/Conclusion] Through experimental research, the accuracy of current multimodal information search is shown in the table above. The present is better, but the integrity of the performance is poor. It takes a long time to understand the requirements in the requirement stage, and there is a common phenomenon of duplicate confirmation. The search method is still text search. The cross-platform search is mainly based on cable and cross-platform search, and the search strategy is adjusted more frequently, and the search mode changes more frequently. Focus on encyclopedic web pages for information collection, especially. Baidu Encyclopedia accounts for the vast majority, while social media and e-commerce applications have also become one of the multimodal sources of information for users.

**Keywords:** Multimodal Information; Information Search; Information Behavior

## 引言

信息搜索行为链接着被搜索信息与信息系统的主客体两端，在图片扫描检索、语音识别检索等技术发展完善之前，通常以文字描述进行搜索，至今仍有大量人使用这种搜索方式，无论搜索文字、图片还是音乐，文字语义描述不可避免的一点就是必须克服“语义鸿沟”，这一问题在图片、音乐这些跨模态搜索中尤其突出。早在2005年，温超等<sup>[1]</sup>就关注过图像检索中的语义差异问题，王靖豪等也认为目前网络环境中的多模态信息缺乏信息层次结构之间的关联，缺少多种模态数据的信息融合，致使相关研究大多依赖于某一种模态数据的信息<sup>[2]</sup>，从而影响搜索的查全率和查准率。对此，胡文婷<sup>[3]</sup>利用眼动跟踪技术记录人浏览图像时的眼动行为，证实眼动行为可表达语义，克服文字与图片之间的语义边界，网站利用这一点，将可以更高效地推送用户需要的图片。秦静、林鸿飞等<sup>[4]</sup>提出了一种基于示例语义的音乐检索模型，使用语义相关标签来标记未知，将音乐

映射到“语义空间”满足语义音乐检索的需求。邵曦、郁青玲<sup>[5]</sup>分析了文本特征和音乐内容特征潜在的统计关系，通过子空间映射解决了不同模态之间的特征异构问题，再根据欧氏距离的大小衡量两者的相关性，从而实现了音乐跨模态搜索。由于用户的知识背景不同，所以对文字、图片、音乐等信息特征的提炼和归纳能力有所不同，使得每个检索者面临的语义差异问题也有相应的不同。但随着技术发展，网站可以直接扫描图片，匹配相同或相似的图片，音乐也可以通过语音识别，查找或匹配相似信息，王培培等<sup>[6]</sup>提出一种基于语句特征的音乐哼唱快速检索技术，将音乐数据库和用户提供的哼唱片段，按自然停顿方式划分音乐语句，使用BDTW算法对音乐语句片段进行音高相似性计算，能够减少用户的检索时间。这些技术的突破意味着不再需要用户对被搜索信息进行特征提炼，绕开了“语义鸿沟”这个问题，使单模态信息的查准率和查全率得到了较好解决。

随着信息技术的深化与推广，各种文字、图像和视频等互动分享网站日新月异，催生了

大量相同语义而有着多种表现方式的信息——多模态信息。多模态信息搜索问题便成为了学者关注的重点,学者 Li<sup>[7]</sup> 等研究了多模态信息搜索行为影响因素,该研究基于 TAM 理论模型,认为多模态信息搜索行为与多模态信息的搜索意愿直接相关,而搜索意愿与多模态信息搜索有用性和容易度感知相关,且用户的信息检索能力及信息环境直接影响多模态信息的多模态信息搜索有用性和容易度感知。邱瑾、吴丹<sup>[8]</sup> 进行了用户协同信息检索行为实验及结果分析,发现任务类型不同,用户“浏览网页”“检索”和“使用图片”行为差距较大,导航类任务涉及专业知识,用户需要输入更多的检索式以便对其有一个全面的认识。对于专业性任务,许多专业人士已经做了一些总结和分析,可能一个图片就包含了大量有用的信息,事务类任务解决的是实践问题,用户在检索过程中思考、探索的时间较多,而浏览页面和检索次数都较少。检索任务类型不同对用户选择的检索式也会有较大影响。

综上,不难发现用户的信息需求已呈现出由慢到快、由少到多、由单一到多样的变化,信息搜索行为也逐渐从单一模态信息搜索向多模态信息搜索转移。而当前国内学者缺乏对多模态信息搜索行为特征的挖掘,使多模态信息搜索行为研究存在空白,为提高多模态信息搜索效率,有必要坚持以人为本的研究理念,以用户为中心,围绕用户在信息技术、网络技术发展进程中突现的日益增长的多模态信息需求,研究其搜索多模态信息行为特征。因此,本文通过设计多模态搜索任务,研究用户的多模态信息行为特征,将有利于丰富信息搜索理论,

为实现多模态信息系统建设、提升多模态信息素养提供依据。

## 1 研究设计

### 1.1 多模态信息搜索行为分析指标

对于信息搜索行为研究,基于用户信息搜索过程的,比较著名的有 Willson 模型、Kuhlthau 模型和 Choo 三阶段模型<sup>[9]</sup>。其中 Willson 模型重点对影响搜索需求到搜索行为发生整个过程的因素进行了较全面的分析。Kuhlthau 模型重点对信息搜索全过程中的用户认知影响进行了分析<sup>[10]</sup>。Choo 模型将信息搜索活动分为信息需要、信息搜索和信息利用 3 个过程,并检验了认知、情感和情境因素对这些过程的影响<sup>[11]</sup>。国内学者甘利人将用户的多模态信息搜索行为过程划分为信息定义、信息搜索和信息收集三个阶段,并根据不同阶段的具体行为表现归纳用户的行为特征<sup>[12]</sup>。对此本文在借鉴前人研究成果的基础上,将多模态信息搜索行为分成三个阶段进行分析,并确定了多模态信息搜索行为特征分析指标,如表 1 所示。

### 1.2 实验任务与流程

在搜索任务划分上,有依据搜索路径将搜索任务分为水平性任务、简单性任务和垂直性任务<sup>[20]</sup>;或根据任务的多个属性将信息任务划分为事实性、解释性和探索性任务<sup>[21]</sup>。Brod-er<sup>[22]</sup> 则将搜索任务划分为导航型任务、事实型任务和事务性任务。孙丽<sup>[23]</sup> 根据任务的实际应用情景将网络健康信息搜寻任务类型划分为开放式任务和封闭式任务。

表 1 多模态信息搜索行为分析指标

行为阶段	维度	变量	变量来源
任务定义	需求理解	搜索任务阅读时长	李法运(2003) <sup>[13]</sup> ; 丁莉(2018) <sup>[14]</sup> ;
	需求确认	重复阅读次数	
信息搜索	选择搜索途径	搜索路径	吴丹等(2014) <sup>[15]</sup> ; 张云秋(2014) <sup>[16]</sup> ; 朱明泉(2006) <sup>[17]</sup> ; Allison(2002) <sup>[18]</sup> ; 罗赛峰(2015) <sup>[19]</sup> ;
	选择搜索方式	搜索形式	
	浏览搜索结果	浏览时长、网页浏览次数	
	调整搜索策略	途径、搜索词、转换情况	
信息收集	提取搜索结果	网站类型、网站数量	甘利人等(2007) <sup>[12]</sup>

考虑多模态信息搜索不仅搜索模态多,搜索路径也多,因此本文按照不同模态进行划分,分别从文本、图片和音频三个搜索模态设计实验任务,三类实验任务定义如表 2。每个被试者均需依次完成三个任务,在执行任务的过程中可以根据实际情况运用一种或多种模态进行搜索。

表 2 任务类型定义

任务类型	任务定义
文本搜索	指用户根据文本描述内容,输入关键字、自然语言等搜索相同主题的不同模态信息
图片搜索	指用户根据信息需求,向系统直接输入图片或图片的语义描述搜索相同主题的不同模态信息
音频搜索	指用户根据信息需求,向系统直接输入一段声音或音频的语义描述搜索相同主题的不同模态信息

任务 1 (文本搜索任务): 假设你看完某个电影后便喜欢上了李安导演。你想通过上网搜索了解更多的关于他的信息,现请你把他的基本信息、照片和他的相关视频通过网络搜索整理出来。(参考用时 8 分钟)

任务 2 (图片搜索任务): 请你根据图片 1,

通过网络搜索,将图片名称、作者的基本信息以及其他相似图片搜索出来。(参考用时 10 分钟)

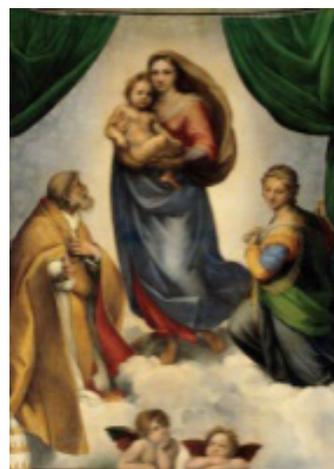


图 1 实验图例

任务 3 (音频搜索任务): 请你根据实验员播放的音频,通过网络搜索,将音频名称、作者的基本信息和及其他相似音乐搜索出来。(参考用时 12 分钟)

### 1.3 实验流程

实验流程主要包括被试者招募、实验平台安排和实验过程设计等。

在被试者招募方面,笔者根据研究目的,并考虑实验任务需具备一定知识储备,因此将

招募对象设为受过高等教育的群体。通过微博、微信等发布有偿招募实验志愿者公告，并根据报名情况筛选符合研究条件的被试者。周晓英等<sup>[24]</sup>在研究互联网信息搜索用户行为模型时的样本数量为32人。吴丹等<sup>[25]</sup>在研究移动搜索策略时选取了30名大学生为研究对象。因此本研究最终招募了45名被试者。

在实验平台设计方面，本研究考虑到实验过程的连续性，实验数据的完整性，且为了避免用户实验受陌生机型及应用的影响，故将本研究的实验平台设定为被试者自带的智能移动手机。实验前统一安装录屏软件，用于录制多模态信息搜索全过程。考虑到录制过程会涉及到隐私信息，如搜索历史等，故在实验开始前会温馨提示被试者整理手机可能会涉及到的隐私信息。

在实验前，组织者向受试者讲解实验过程和注意事项。实验顺序依次为任务1、2、3，每个任务流程为：阅读题目-实施搜索-保存搜索结果。每项搜索任务完成后，被试者需将每个任务的搜索结果收集整理发给实验负责人，以此类推，直至三个任务全部完成。在实验过程中，为使被试者端正态度提高搜索效率，每项搜索任务提供参考用时，且实验员将进行计时操作，但为了体现实验的完整性，故即使被试者用时超过参考用时也不会被打断。实验完成后，评委会根据网络信息搜索常用指标“准确性”“完整性”<sup>[26]</sup>对被试者的搜索结果进行评分，并以此奖励被试者。准确性表示搜索目标与搜索任务的匹配度，分值为5。完整性指搜索到的信息在模态、数量以及任务完成的情况，分值为5。每项任务的分值为10，三个任

务总分为30分。

## 2 结果与讨论

本研究共有45名符合研究条件的志愿者参与。其中女性被试者为26人，占比约57.78%，男性被试者为19人，占比约42.22%。被试者学历大专有1人，本科有19人，研究生学历25人（含硕士研究生23人，博士研究生2人），群体涉及有社会在职人员与在校学生，专业涉及生物工程、农学、法学、金融学、理论物理、光学工程、中国现当代文学、情报学和图书馆学等。删除画质不清、实验不完整视频，最终获得有效实验数据41份。41份实验数据共约691分48秒，减去程序加载、网页缓冲和其他与任务搜索无关时间等，最终具有统计分析意义时间约606分52秒。

### 2.1 多模态信息搜索结果分析

本研究邀请了来自图情领域与非图情领域的三名学者为评委，对实验者的搜索结果进行评分，用三者评分的平均分作为被试者多模态信息搜索结果的最终得分，以此分析调查者的多模态信息搜索的效率，其中包括了多模态信息搜索的准确性、完整性。如表3。

表3 多模态信息搜索任务完成情况

任务 \ 指标	准确性 (5分)	完整性 (5分)	总分	平均分
文本搜索	193	179	372	9.07
图片搜索	163	168	331	8.07
音频搜索	175	165	340	8.29
总分	531	512	1043	25.43

数据显示,从整体上看,多模态信息搜索准确性较好,在完整性上表现较差。从不同模态看,文本信息搜索的整体完成情况最好,准确性和完整性平均分超9分;音频搜索整体评分居中;图片搜索的整体平均分最低,且准确性分数最低。说明文本为搜索模态搜索多模态信息,仍是目前效率最高的搜索方式;图片搜索的准确性则有待提高。

## 2.2 多模态信息搜索行为分析

图2反映的是用户的多模态信息搜索各个阶段人均耗时。数据显示,多模态的搜索任务人均耗时约983秒,结合录屏,文本搜索单人最高耗时645秒,图片搜索单人最高耗时923

秒,音频搜索单人最高耗时690秒,这三人所在专业均属生物类。从不同阶段上看,搜索阶段耗时最长,其次是信息收集阶段,定义阶段耗时最短。从不同搜索模态上看,文本搜索耗时最短,图片搜索耗时最长,音频搜索居中。结合录屏可知,用户向信息系统输入搜索需求时,文本输入速度普遍较快且系统识别的准确率和完整性较高;输入图片或音频时,信息系统在识别音频或图片时需要较长时间,且识别失败率较高,迫使被试者重新检索,从而一定程度上增加耗时。说明当前阶段用户的多模态信息搜索速度,特别是图片搜索和音频搜索,受信息系统的响应速度和识别精准度影响较大。

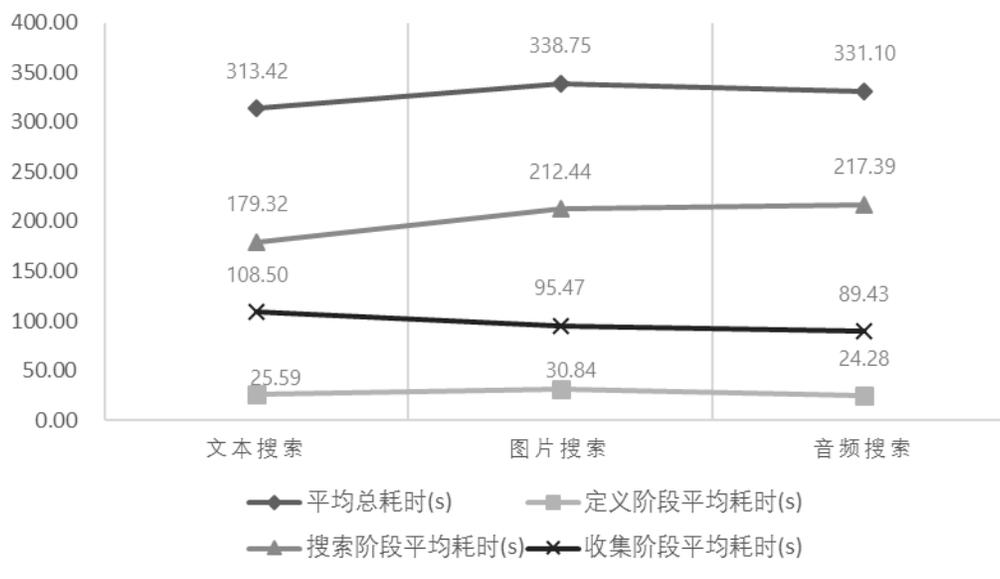


图2 多模态信息搜索各阶段人均耗时

### (1) 多模态信息需求定义阶段

在信息搜索过程中,用户需理解所找到的信息,从而确定自己的需求<sup>[27]</sup>。图3统计的是被试者理解搜索任务的时长,即从打开任务文档到

打开搜索软件这一时间段的时长。由图可知,图片搜索需求理解平均耗时最长达30.84秒,音频搜索平均耗时最短,文本搜索居中。结合录屏,被试者在进行图片搜索时,用户除理解搜索任务

外，还会对图片细节进行放大确认，在细节确认过程中耗时较长。对于音频搜索，大部分被试者

在听到音频后便第一时间打开音乐软件进行识别，在选择搜索方法与途径上反应迅速。



图3 信息定义阶段需求理解耗时

表4反映的是被试者确认搜索需求情况，即重复阅读任务题目的次数。实验显示，大部分被试者在确定搜索任务需求时需重复确认，最高重复阅读次数多达5次。音频搜索重复阅读任务题目次数最少，且出现较多不需要重复阅读的情况，而文本搜索需求确认次数最多。

表4 信息定义阶段需求重复确定次数

编号	文本搜索	图片搜索	音频搜索
重复阅读总次数	68	54	51
人均重复阅读次数	1.66	1.32	1.24

综上，在多模态信息搜索需求定义阶段，

用户对音频搜索需求的理解能力较强，在确认和选择音频搜索方法与途径上比较迅速；而对于图片搜索，虽然大部分用户依赖图片识别工具，但由于图片细节较多，用户在对图片细节捕捉上需耗费一定精力与时间。而文本搜索符合大部分用户现阶段的阅读习惯，所以用户在此阶段无需花费过多的时间。

(2) 多模态信息搜索阶段

信息搜索是多模态信息搜索的核心阶段，具体行为表现有：选择搜索途径、选择检索方式、筛选搜索结果和搜索策略调整等。

① 搜索路径

搜索途径选择是用户进行多模态信息搜索

动作的第一步，本文从用户所选的搜索路径来分析用户搜索不同模态信息的行为特征。图4显示，目前多模态信息搜索选择最多的搜索方式为搜索引擎，其次是浏览器。结合实验，发现被试者在进行文本搜索时大都选择浏览器，而在图片搜索时则倾向于具有图片识别功能的搜索引擎，其次是电商类APP如淘宝等；音频

搜索倾向于专业APP如网易云等，其次是搜索引擎。由此可以说明搜索引擎是当前多模态信息搜索的主要途径，但用户会根据搜索模态选择多种且具有针对性的方式，这与前人的研究结果保持一致，无论哪种类型的搜索问题都有大部分大学生选择搜索引擎作为信息检索的第一工具<sup>[28]</sup>。

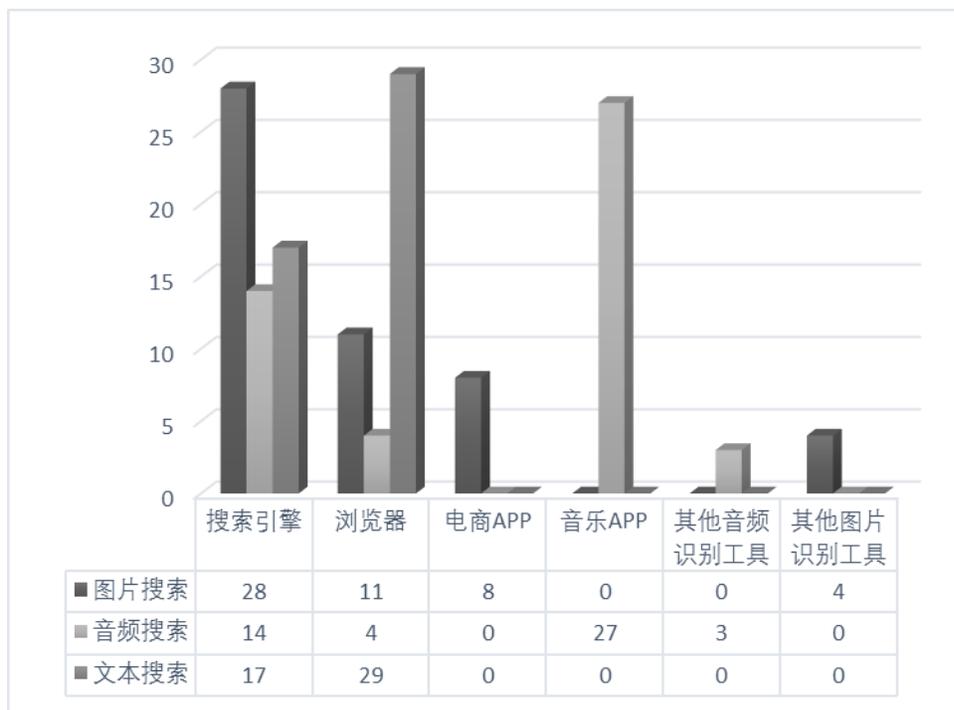


图4 搜索路径

## ②选择搜索方式

构造检索式是信息需求的外化过程，是用户的认知灵活性与知识迁移、应用和创新能力的具体表现。根据图5可知，短语搜索多模态信息是搜索过程中使用最多的检索方式，图片搜索和音频搜索除了依赖专业的图片识别和音频识别工具外，均使用文本搜索辅助搜索。此外，值得注意的是语音输入仅是个别用户选择。说明目前关键词搜索仍是多模态信息搜索的主流，而图片识别

和音频识别也成为了重要的搜索方法之一。

图6反映的是用文本、图片和音频作为搜索模态占全部搜索方式的比例及搜索词从实验题材中提取的情况。数据显示，被试者使用图片识别和音频识别作为搜索方式各占9.77%，使用文本搜索的高达80.46%。结合录屏，发现在音频搜索任务中，检索关键词取词来源大部分取自歌词本身，而图片搜索和文本搜索部分搜索词源于自身对搜索任务的认知，并在此基

基础上思考并提出新的检索词。说明在完成信息搜索任务的过程中,音频搜索仅依靠直观理解,并没有做出太多个人的判断与评价,缺乏对检

索词的深度加工,而图片搜索和文本搜索则相反,这一结论进一步验证了信息搜索受用户的学习认知投入的影响<sup>[29]</sup>。

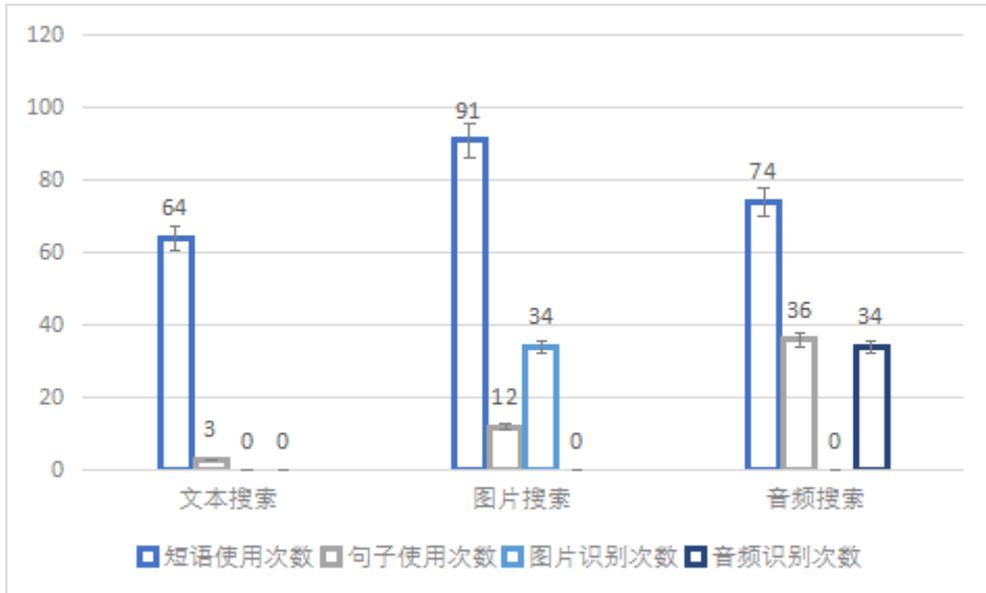


图 5 短语、句子、图片和音频识别使用频次

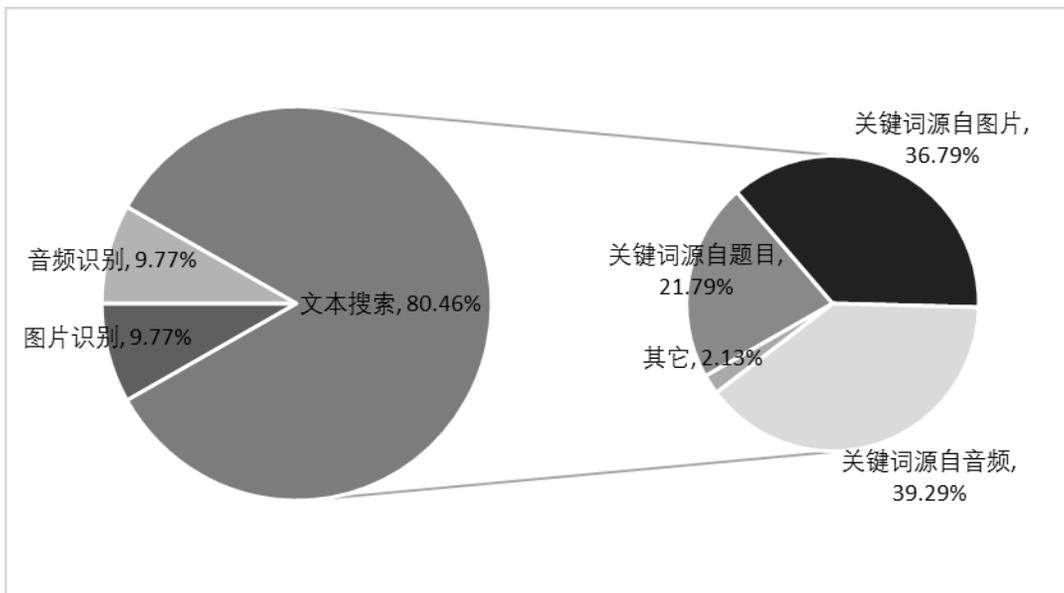


图 6 搜索方式比例

### ③筛选搜索结果

信息筛选是指用户根据信息系统反馈的信息查找和筛选目标信息，具体行为表现有网页浏览。表5数据显示，网页浏览占整个多模态信息搜索阶段的耗时一半以上，其中音频搜索中的网页浏览耗时最长，在整个搜索过程的耗时占比也最大，文本搜索则相反。结合实验发现，音频搜索用户点击浏览网页的数量最多，人均点击量约为7.34页；网页点击浏览数量最少的是文本搜索，人均点击量约为5.9页。由此反映用户及信息系统均对文本信息搜索需求捕捉较精准，反馈效率高，而图片和音频信息需求捕捉和信息反馈的准确性和时效有待提高，这也说明信息系统在满足用户偏好可视化和娱乐性强的新媒体信息需求上<sup>[30]</sup>，还需做进一步改进和提升。

表5 网页浏览耗时

任务 浏览耗时(s)	任务		
	文本搜索	图片搜索	音频搜索
网页浏览总耗时	6635	8710	8913
网页平均浏览耗时	27.42	29.23	29.61
人均网页浏览耗时	161.83	212.44	217.39
占总搜索耗时比	58.81%	71.42%	73.09%

### ④搜索策略调整

相较于搜索方法，搜索策略立足于整体，强调对整个搜索过程的计划与变通<sup>[31]</sup>。表6数据显示，大部分被试者在进行多模态信息搜索过程中都需要调整搜索策略，不需要调整或调整次数3次及以上的是少数。结合录屏，发现多模态信息搜索策略调整主要是平台切换，如文本搜索主要从浏览器转为百度

搜索引擎，而图片和音频搜索则主要从专业识别工具转为浏览器或百度搜索引擎。另外，图片搜索和音频搜索在搜索策略调整上比较频繁且主观性较强，文本搜索则相反。从搜索模态上看，目前图片搜索和音频搜索则存在较频繁的图文、音文相互转换的情况，而文字搜索则没有这种情况。

综上，可以说明当前被试者的多模态搜索仍以跨平台搜索为主，大部分多模态信息搜索与收集无法集中在一个平台内完成。与此前研究成果中的发现一致：信息搜索行为在不同任务环节存在一定差异，尤其是检索次数、检索工具等<sup>[32]</sup>。

表6 多模态信息搜索策略调整情况

文本搜索		图片搜索		音频搜索	
调整次数	频数	调整次数	频数	调整次数	频数
0	5	0	4	0	7
1	27	1	29	1	27
2	8	2	5	2	3
3次及以上	1	3次及以上	3	3次及以上	3

### (3) 多模态信息收集阶段

图7反映的是多模态信息搜索结果的信息来源情况。数据显示，用户的多模态信息收集渠道比较集中，百科类网页特别是百度百科占了绝大多数，且百度系列网页是绝大部分被试者的多模态信息来源。说明百度作为搜索引擎的代表，在多模态信息融合方面具有较大的优势，普遍被用户所接受。值得注意的是，新兴的社交媒体如豆瓣和应用软件

如淘宝、京东等电商网站也逐渐成为了用户的多模态信息来源，这一结果验证了社交网

络逐渐成为除搜索引擎与数据库之外的主要信息来源渠道<sup>[33]</sup>。

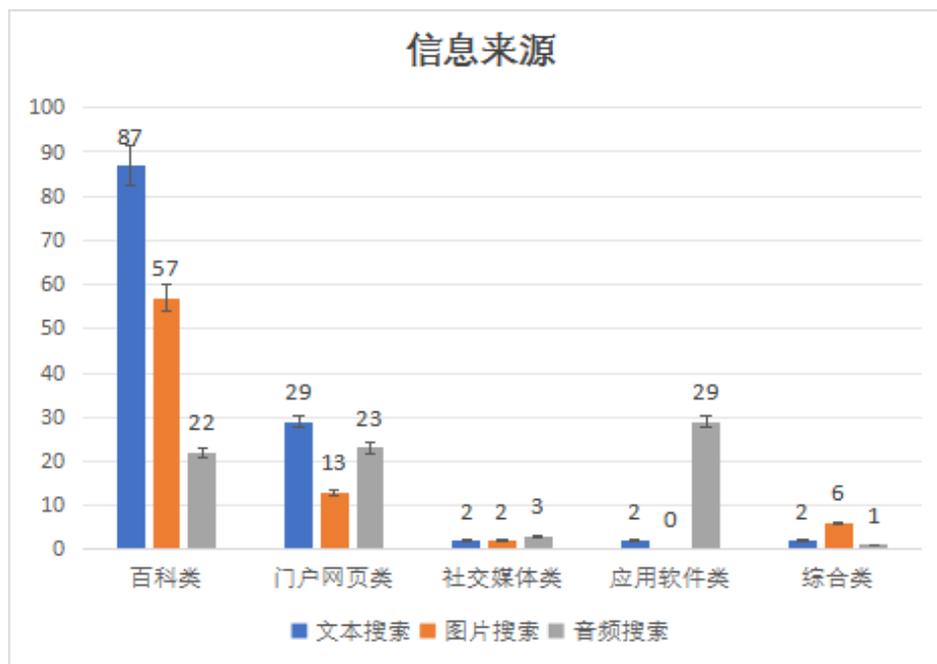


图7 多模态信息收集来源

综上，纵观整个多模态信息搜索过程，可以发现多模态信息搜索在准确性上表现较好，而完整性上则表现较差。在需求阶段需求理解上比较费时，普遍存在重复确认现象。在搜索方式上仍以文字搜索和跨平台搜索为主，且搜索策略调整次数较多，在搜索模态上转变较频繁。在信息收集上集中于百科类网页，特别是百度百科占了绝大多数，而社交媒体和电商应用软件也成为用户的多模态信息来源之一。

### 3 研究建议

通过本研究，发现当前的信息系统还无法兼顾文本、图片和音频信息搜索的查全率和查准率，致使用户在搜索多模态信息时，需跨平台、

换模态和多次重复搜索，且搜索结果还不完整。在内容上也无法满足用户的可视化、娱乐性和富含文化底蕴的信息需求。此外，实验还发现用户本身对多模态信息搜索也存在认知不足，包含对信息的多模态表现形式、获取方式和信息整合等认知，具体表现在需求确定过程中重复确认搜索任务的次数多，搜索策略调整次数多，在搜索模态上转换频繁。对此本研究从多模态信息系统设计与开发与用户信息素养教育方面提出建议。

#### 3.1 多模态信息系统开发与设计

在多模态信息搜索系统开发与设计方面，本研究认为，要兼顾查全率和查准率，需在检

索页面设计方面,探索图像、视频、文本、语音等多模态输入的高效感知与综合理解手段<sup>[34]</sup>,提供文字输入的同时,增加语音、图片和音频输入功能,放置显眼区域,强化引导用户用多模态进行检索,扩大检索面。此外,针对用户的跨屏信息搜索需求也应提供更好的跨屏检索服务<sup>[35]</sup>。在多模态信息融合方面,建立多模态信息语义系统,将语义相同的不同模态信息进行多维度关联,扩充信息系统的不同模态的信息资源,以提升整个系统的查全率和查准率。在多模态信息呈现方面,改变以文本信息为主的传统网页呈现形式,呈现兼顾可视化、娱乐性和文化底蕴的多模态信息,具体可通过语义关联将文本信息、图片信息和音频信息深度融合并分类呈现,努力做到多模态信息一站式检索,需求一站式满足。

### 3.2 多模态信息素养提升

在数字化媒体盛行的新时代,信息素养是学生适应当下、迎接未来所必须具备的能力素养之一<sup>[36]</sup>。因此在多模态信息素养培养方面,学校应根据多模态信息特征,努力帮助学生避免因信息表现形式多而导致信息迷航,跳出因精准推荐导致的信息茧房。重点围绕帮助学生明确问题对象-“要何种信息做什么”;明确信息搜索目标-“满足什么样的需求”;明确时间-解决问题的时间节点;明确信息内容-信息的表现方式(文本、图片或音频等模态);明确信息搜索策略(不同搜索模态的不同搜索路径、搜索方法和信息源等)。增设多模态信息搜索、加工利用教学课程,培养相关能力。一方面,针对不同信息对象灵活开展各种模态信息的相

关检索策略、语言和路径等系统的理论指导教育,培养学生多模态搜索信息的意识,引导学生善于利用知识整合需求信息、构造检索式、多途径搜索。另一方面,学校需在理论指导上辅以合作基地进行实践教学,例如建立以文字为主的中国知网数据库、图片为主的淘宝网电商平台、以视频为主的抖音短视频开发商等校企合作链,定期安排学生到合作基地进行学习,加强学生对图片和音频信息的特征挖掘、信息融合和检索技术等方面的认识与实践。

### 3.3 加强多模态信息搜索行为伦理与法规引导

多模态信息涉及到了许多知识产权、设计专利和个人隐私等信息,规范多模态信息行为的重要性不言而喻。规范多模态信息行为,从源头上看是要用户形成信息自觉,实现则从培养信息用户的信息伦理意识着手。要使用户在主观上尊重知识产权,保护隐私信息,抵制不良信息,在行动上杜绝网络谣言、信息暴力和网络诈骗等行为。重点是完善相关的法律法规,特别是知识产权和个人信息方面的法律,增加多模态信息违法成本,将法律监管覆盖多模态信息的产生、加工和传播等全过程。建立多模态信息诚信黑名单制度,一旦发现有相关不良信息记录的用户或平台立即发出警告,严重者将其列入信息诚信黑名单、封锁相关账号并追究其相关法律责任,限制其对多模态信息发表、下载、转发和保存的权限。

## 4 结束语

本文依据信息搜索行为阶段理论,分析多

模态信息搜索行为特征,通过实验研究,发现多模态信息搜索在准确性上表现较好,而完整性上则表现较差;对多模态信息搜索在需求阶段需求理解上比较费时,普遍存在重复确认现象。在搜索阶段仍以文字搜索和跨平台搜索为主,且搜索策略调整次数较多,在搜索模态上转变较频繁。在信息收集阶段,收集渠道集中于百科类网页特别是百度百科占了绝大多数,而社交媒体和电商应用软件也成为用户的多模态信息来源之一。

本研究也存在一定的局限性。一方面人工对录屏数据进行统计分析可能存在一定误差,个人理解差异也可能对分析结果产生一定的影响;另一方面研究群体覆盖面不够广,分析指标也需要进一步丰富。未来可针对以上不足,扩大样本,采用多种统计分析方法,提高研究的科学性和全面性。

## 参考文献

- [1] 温超,耿国华.基于内容图像检索中的“语义鸿沟”问题[J].西北大学学报(自然科学版),2005(5):46-50.
- [2] 王靖豪,刘箴,刘婷婷,等.基于多层次特征融合注意力网络的多模态情感分析[J].中文信息学报,2022,36(10):145-154.
- [3] 胡文婷.眼动行为的语义表达及其在图像检索中的应用研究[D].南京:南京大学,2016.
- [4] 秦静,林鸿飞,徐博.基于示例语义的音乐检索模型[J].山东大学学报(理学版),2017,52(6):40-48.
- [5] 邵曦,郁青玲.基于典型相关的音乐跨模态检索[J].计算机技术与发展,2015,25(7):76-81.
- [6] 王培培,杨晓春,王斌,等.基于语句特征的音乐哼唱快速检索技术[J].东北大学学报(自然科学版),2017,38(3):315-319.
- [7] LI Z, ZHU J, LI X. Factors influencing the behavior of multi-modal information search[J]. Library Hi Tech, 2022, 40(5): 1459-1482.
- [8] 邱瑾,吴丹.用户协同信息检索行为与系统评价研究——以任务类型和协同能力为视角[J].现代图书情报技术,2012,224(9):62-68.
- [9] WILSON T D. Models in information behavior research[J]. Journal of Documentation, 1999, 5(3): 249-270.
- [10] KUHLTHAU C C. Accommodating the use's information search process: challenges for information retrieval system designers[J]. Bulletin of the American Society for Information Science and Technology, 1999, 25(3): 12-16.
- [11] CHOO C W. Closing the cognitive gaps: How people process information[C]//Financial Times of London, March. 1999, 3-22.
- [12] 甘利人,岑咏华,李恒.基于三阶段过程的信息搜索影响因素分析[J].图书情报工作,2007,51(2):59-62.
- [13] 李法运.网络用户信息检索行为研究[J].中国图书馆学报,2003(2):64-79.
- [14] 丁莉.少儿网络信息搜索行为研究—以百度为例[D].武汉:华中师范大学,2018.
- [15] 吴丹,蔡卫萍.青少年网络信息查询行为影响因素实证分析[J].图书情报工作,2014,58(19):61-67.
- [16] 张云秋,李玉玲,王洪媛.基于日志与认知分析的探索式医学搜索行为研究[J].图书情报工作,2014,58(11):36-42.
- [17] 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(4):475-480.
- [18] ALLISON D, ELIZABETH F, LESHELL H, et al. How children search for information on the Internet[J]. Computer Education, 2002, 10(2): 17-19.
- [19] 罗赛峰.儿童信息检索行为若干问题的讨论[J].图书馆杂志,2015,34(8):7470-7477.
- [20] KIM K S, ALLEN B. Cognitive and Task Influences

- on Web Searching Behavior[J]. Journal of the Association for Information Science & Technology, 2002, 53(2): 109-119.
- [21] HSIEH C C, CHENG W H, CHANG C H, et al. Photo Navigator[A]. Proceedings of the 16th ACM international conference on Multimedia[C]. Vancouver, 2008: 419-428.
- [22] BRODER A. A Taxonomy of Web Search[J]. ACM SIGIR Forum, 2002, 36(2): 2-10.
- [23] 孙丽. 任务类型对网络健康信息搜寻行为的影响及其预测模型研究[D]. 长春: 吉林大学, 2015.
- [24] 周晓英, 蔡文娟. 大学生网络健康信息搜寻行为模式及影响因素[J]. 情报资料工作, 2014, 199(4): 50-55.
- [25] 吴丹, 梁少博, 冉爱华. 大学生用户移动搜索策略研究[J]. 中国图书馆学报, 2016, 42(3): 55-73.
- [26] STOKER D, COOKE A. Evaluation of Networked Information Sources[C]. Essen Symposium, 1994: 287-312.
- [27] ZHANG P Y, SOERGEL D. Cognitive mechanisms in sensemaking: a qualitative user study[J]. Journal of the Association for Information Science and Technology, 2020, 71(2): 158-171.
- [28] 陆溯. 大学生网络信息搜索行为实证研究——基于搜索引擎的利用[J]. 图书馆理论与实践, 2018(1): 79-82.
- [29] 孙晓宁, 季馥春, 刘思琦. 信息搜索用户的学习投入影响研究: 基于 Kolb 学习风格与认知弹性理论[J]. 情报学报, 2023, 42(1): 90-102.
- [30] 张平, 张卫良. 大学生新媒体使用行为及信息选择特点的实证分析——兼论对主流意识形态教育的启示[J]. 深圳职业技术学院学报, 2019, 18(2): 62-69.
- [31] SAVOLAINEN R. Information seeking and searching strategies as plans and patterns of action[J]. Journal of Documentation, 2016, 72(6): 1154-1180.
- [32] 王志红, 曹树金. 工作任务视角下大学生信息搜索与利用行为研究: 以撰写课程论文为例[J]. 情报理论与实践, 2022, 45(8): 38-46.
- [33] 张海涛, 张泉慧, 魏萍, 等. 网络用户信息检索行为研究进展[J]. 情报科学, 2020, 38(5): 169-176.
- [34] 李学龙. 多模态认知计算[J]. 中国科学: 信息科学, 2023, 53(1): 1-32.
- [35] 贾明霞, 周前钱, 徐跃权. 同一任务情境下大学生跨屏信息搜索影响因素研究[J]. 大学图书情报学刊, 2022, 40(6): 42-54.
- [36] 马会仁, 朱钰. 大数据时代小学生信息素养培养策略探究[J]. 电脑知识与技术, 2022, 18(30): 178-180.