
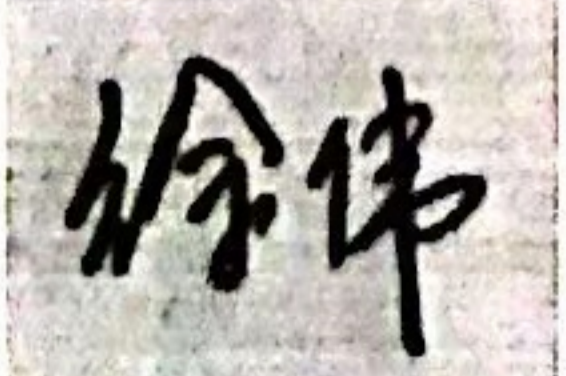
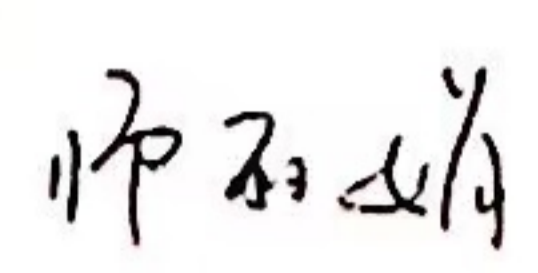
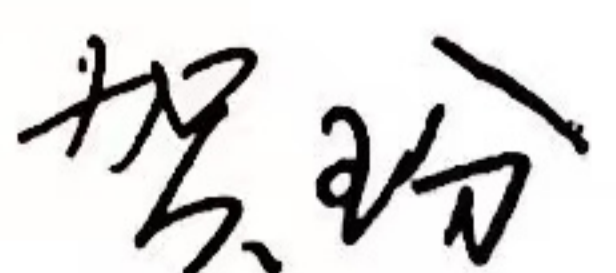

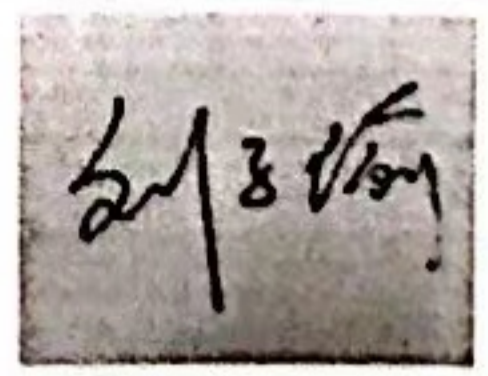


项目结题验收单

专家验收表（主持人所在单位组织 3-5 名专家对项目进行验收、自评。）

项目名称	资源发现系统功能特点分析及发展方向研究				
主持人	李雪原	职务/职称	副研究馆员		
所在单位	(加盖单位公章)				
专家意见	<div style="text-align: center;">  </div> <p>在网络学术资源发现与获取系统的架构和界面趋于“同质化”的情景下，基于用户视角开展用户体验度比较与评估资源发现系统，对优化资源发现中的检索服务具有重要参考价值，项目选题具有重要的现实意义。</p> <p>用户体验测试在一定程度上能够反映不同资源发现系统的优势与缺陷，发现系统的平台界面质量、信息质量、服务质量及用户满意度等。本项目选取代表性文献资源发现系统，围绕不同系统检索功能、资源获取便捷性、以及相关衍生功能进行用户角度的调研，通过比较得出用户体验报告，从用户体验的角度揭示发现系统的改进方向，提出优化完善建议，研究结果可为优化文献资源发现系统检索服务提供参考，为资源联盟建设者提供改进策略。</p> <p>研究逻辑清晰，数据详实，研究结论具有较高的参考价值。</p> <p>同意结题。</p>				
	(如需要可增加页数)				
专家签字					
职务/职称	副馆长/副研究馆员	研究馆员	副研究馆员	副研究馆员	副研究馆员



项目编号:

注:项目编号请查看立项通知,也可缺省

CALIS 全国农学文献信息中心研究项目 结题报告

项目名称: 资源发现系统功能特点分析及发展方向研究

项目关键词: 资源发现系统; 知识发现系统; primo; summon;
ebshost; 百度学术; 超星发现系统

项目单位(盖章): _____

通信地址: (详细地址
址含邮编) _____

项目主持人: 李雪原

联系电话: 010-62732748

电子邮件: Tsg019@cau.edu.cn

提交日期: 2023.3.10

资源发现系统功能特点分析及发展方向研究

摘要：文章以 primo、summon、ebSCOhost、百度学术、超星发现 5 种较为常用的发现系统为分析对象，通过平台检索测试与对比分析，总结归纳发现系统的功能特点，提出发现系统应该在以下几个方面进行重点建设：加强元数据的规范建设、提升文献信息的揭示度与获取途径、馆际互借与文献传递系统的嵌入、优化相关性排序、建立个性化功能模块、利用衍生数据推动精准服务。

关键词：资源发现系统；知识发现系统；primo；summon；ebSCOhost；百度学术；超星发现

1 研究背景、目的及意义

在资源发现系统出现之前，图书馆界为了更好的揭示文献信息，大多是通过手工编目或计算机编目。20 世纪 80 年代目录学进入到一个快速发展时期，90 年代中期，目录学系统的知识体系与书目著录的概念均已形成，在此基础之上，许多图书馆的主要任务就是对书目进行严谨、系统、规范的著录。许多高等院校图书馆、公共图书馆都对编目人员进行岗前培训，编目员需要通过严格的考试，取得编目员证书后执证上岗，可见当时图书馆界内对编目及著录规则给予的重视与投入。在经历了书目工具的编制、纸质文献手工卡片的编目、以及 90 年代初期兴起的计算机书目著录（OPAC）后，随着电子文献资源的增多，计算机技术、网络技术的快速发展，如何适应文献资源的激增、如何将快速涌现的各种类型、各种载体形式的文献资源进行科学、系统、规范的整合与揭示，为目录学提出了挑战。

面对新技术和方法的影响，很多目录学从业者出现了担忧和困惑，Wang（2020）提出了“应用目录学”，但依然不能阻止目录学研究的降温态势，而国外的目录学研究者 Breeding（2007）提出了“下一代目录学”的工程计划，元数据的概念也应运而生。基于元数据的蓬勃发展，各种资源整合模式与索引模式出现。世界上最大的书目记录数据库——WorldCat 联合目录数据库于 1971 年建库，到目前为止，共收录有 483 种语言总计达 4.9 亿多条书目信息，并且每年以 200 万条的速度增长，包含了 OCLC 近两万家成员馆的编目馆藏信息，其目录数据库每日更新，图书馆对文献目录的整合与揭示进入了崭新的快速发展阶段，但由于各种条件的限制，都无法让用户达到满意，直到 Google 学术搜索的出现，一站式的检索服务让用

户不在纠结“搜索应该从哪里开始”。在经历了资源导航、联邦搜索的探索后，图书馆借鉴了 Google 学术搜索的理念与模式，走上了资源发现系统之路。2009 年，Series Solution 公司发布了全球第一个网络资源发现系统——Summon。自此，基于元数据整合、网络爬虫技术、与索引技术的资源发现系统快速占领市场，截至 2022 年 1 月，summon 在国内的用户已达到 70 家以上。资源发现系统在国内外图书馆领域均已非常普遍，国内的 39 所“985”院校图书馆已全部配置资源发现系统，有的图书馆配置中外文两套发现系统。国内中文发现系统主要有超星发现、学知搜索、以及百度学术等。中外文发现系统不断的迭代、升级，为目录学的发展、以及全球范围内的文献传递事业做出了不可估量的贡献。

目前国外的资源发现系统主要由 EDS、primo、ebSCOhost、summon、worldcat 等，国内主要有 e 读、读秀、超星学术、指针搜索、百链云、Findplus、百度学术搜索等。相关研究主要有张毅、陈芳等人的高校图书馆发现软件使用测评的调研与分析；苏建华、胡玮等人的中文发现系统比较研究；以及李慧芳等人的资源发现系统与学术搜索引擎的比较研究。总体而言，相关研究具有现实性，对文献资源发现中的检索服务优化具有参考价值，但研究大多是从建设者角度出发，缺少从用户角度出发，对用户使用过程中的体验分析的较少。基于此种现状，本研究拟选取代表性文献资源发现系统，围绕其检索功能、资源获取便捷性、以及相关衍生功能进行用户角度的调研，比较分析后，得出用户体验报告，从用户体验的角度分析发现系统的改进方向，提出优化完善建议，为优化文献资源发现系统检索服务提供参考，为资源联盟建设者提供改进策略。

2 研究内容及方法（思路、方法、具体内容）

外文发现系统在中国用的最普遍与最好的是 primo、summon、EbscoHost、OCLC 的 World cat，号称四大发现系统，为更好分析比较发现系统的功能与特点，在这四大外文发现系统中选取 Primo、summon、Ebscohost 三个作为典型案例；中文发现系统选取百度学术搜索与国内使用最多的超星发现系统作为典型案例。为了深入了解各个发现系统平台的特点，对选取的 5 家发现系统进行了检索测试，测试的版本选取：summon 选取了北京大学的未名学术搜索、Primo 选取了清华大学的水木学术搜索、EDS 选取了武汉大学的珞珈学术搜索、超星发现系统选取了中国农业大学的测试版本；测试内容主要围绕以下三个方面：1 检索环节与界面揭示；2 信息质量与获取渠道；3 服务与反馈，分析结果见表 1。经过检索测试，发现五家发现系统各有特色，基本均可满足用户需要，为了更好的了解用户对这五家发现系统的

使用感受，以问卷的方式进行用户使用测试。

表 1 5 家中外文发现系统功能特点归纳

测试内容	Summon	Primo	Ebscohost	百度学术	超星发现
检索环节与界面揭示	有单一检索框与高级检索；分面检索与结果限定项设置得清楚合理。以电脑桌面版未名学术搜索为例，主界面从左到右依次有 3 栏：快速筛选区、搜索结果区、参考资料区。在这三栏上方，还有搜索框、高级检索按钮，以及收藏夹、菜单按钮、反馈入口等。	有单一检索框与高级检索；检索环节中可做纸本资源与电子资源区分；在“作者”栏目内，添加了“超星数图信息技术有限公司”，选取该限定项后，不仅可以查看馆藏的纸本书目数据，且可以直接查看超星电子书，可见其元数据层对纸本馆藏书目与电子资源作了映射匹配，方便用户在使用时对文献载体一目了然并随意选择。	有单一检索框与高级检索；检索界面可做简单定制；在进一步检索中发现，系统有个“相似图书”与“此作者撰写的其它书刊”链接，点击后，可根据现有书目进行扩展搜索，推荐出很多相似的图书与该作者撰写的其它资料，并且有浮动灵活的书目介绍框方便用户以书找书，非常有特色。	有单一检索框与高级检索；检索环节较前三家发现系统较简单；结果多以学术文章为主，分面检索多以时间、领域、期刊来源、机构、作者等为主。可一键切换到英文搜索，相关性排序可以被引量排序。	有单一检索框与高级检索；此外还有专业检索框，分面检索选项较多，有精炼检索、内容类型、年份、关键词等。可能是试用版原因，没有对馆藏优先排序，检索结果排位较前的是超星期刊、中国知网、维普、以及全国图书馆参考咨询服务平台的文献传递服务链接。可视化分析模块新颖实用，对全局掌握检索主题的脉络很有帮助，有相关知识点、第一作者与相关作者、相关机构、以及各种文献类型的发展趋势曲线、期刊刊种的分类统计、基金、地区的可视化统计分析。
信息质量与获取渠道	优先对馆藏了排序；文献信息与来源标识清楚。整合揭示的文献类型非常丰富，除图书、期刊、论文等常见文献类型外，还有图像、蓝图、影像、缩微资料、幻灯片、乐谱等。其检索文献信息的范围非常广，可以检索到域外多家开放性较好的图书馆馆藏资料及文献提供者的界面。在测试过程中，检索波兰一家数字图书馆的地图。	对馆藏做了优先排序；信息质量与信息来源以分面检索的形式揭示较清楚。随书光盘可直接链接到非书资源平台（书附光盘/音像）。在返回的检索结果中，每条文献的信息揭示中不仅包括馆藏信息、书目详细信息、链接、简介、相关阅读，还有一个问题反馈，这个功能方便用户帮助提升文献的元数据质量，相当于建立了一个数据反馈纠错机制。	在对检索结果勾选限定项后，信息来源、全文链接、文献传递标识清楚，但是信息丰富度大大减少。没有开放获取标识与渠道。提供许多辅助功能（工具），如电子邮件、保存、引用、导出、打印等。总体来讲，可能由于该发现系统过于全球化、国际化，其界面逻辑对信息素养要求较高。	信息多以专业的学术文章为主，获取渠道为有名的数据库商家、百度文库、开放获取的机构与出版社等。信息来源与全文链接标识清楚。最大的特点是对文章被引量及引文趋势揭示清楚。全网免费使用的同时，包含了少量收费项目如“单篇购买”。	信息多以专业学术文章、图书、会议论文、学位论文、专利为主。信息来源与全文链接标识清楚，并可以一键收藏、分享、及部分文献可以扫描阅读。在期刊导航中，实现了 13000 多种 OA 期刊的揭示服务，点击后可以直接链接到相应的数据库。在精炼检索中，有个“本单位学术成果”限定选项，勾选后，显示的主要是本机构的学术成果，如学位论文、本机构科研人员发表的文章、会议论文、发明专利等。有“看过本文的还看了”检索扩展服务。
服务与反馈	有问题反馈服务，用户留言后邮箱回复。	有新版搜索问题反馈。对新版发现系统没有醒目的推广与介绍。	与图书馆在线咨询服务合二为一。	没有在线咨询服 务，可由互助中心寻求帮助。	在“联系我们”中提供智能在线咨询服 务，同时可转人工。

2.1 问卷设计

资源发现系统是一种典型的、主流使用的信息系统，其使用效果与信息质量、系统质量、服务质量密切相关，参考陈颖的“基于相关性判据的学术信息检索系统成功模型实证分析”与赵钊等人的“基于 D&M 模型的信息系统成功评价研究”信息系统评价模型，考虑资源发现系统的特点，将信息质量、系统质量、服务质量、用户满意度、用户使用意愿作为主要测评方向，同时增加一些开放性问答题型，对五家资源发现系统进行测评。为保证测评质量，提前与图书馆专业人士、教师、研究生进行访谈与商榷，结合用户检索信息的行为习惯，对问卷的题项设置反复推敲调整，最后设计出问卷。问卷主要包括两部分，第一部分是考察样本的人口统计学特征以及使用偏好，如年龄、身份、教育程度及平时查找信息的入口习惯。第二部分采用 Likert5 分制考察用户在使用资源发现系统时对信息质量、系统质量、服务质量的体验感受。第三部分则是问答题，询问用户在检索时最看重哪一部分因素，并进行优先排序。

2.2 数据收集

问卷先是采用了在线问卷的形式，在进行了一周的预测试后，效果反响不佳，回收的预测试问卷结果显示，许多用户对资源发现系统不是很了解，这跟资源发现系统平时的推广度不够有关。为了保证用户在测试时遇到的问题能够得到及时解答，采取了线下问卷调研的形式，分别选取了中国农业大学、青岛大学、北京师范大学三所高校进行了调研，主要调查对象为本科生、研究生与教师，最后回收问卷 379 份，剔除作答不完整与无效问卷后，实际统计分析问卷数量为 338 份。

2.3 数据统计与分析

问卷调查样本的性别、身份及查找文献优先选择的检索入口特征，结果如表 2 所示。从被调查者的人口数据可看出：男女比例较均衡，占比分别为 46.45%与 53.55%。样本身份主要以研究生为主，占到 58.28%，其次是本科生，占比为 33.43%，最后是教师，仅占 8.28%，说明该调查研究生与本科生的参与积极性较高，教师的参与度欠缺。当询问查找资料优先使用哪个检索入口时，在样本中占比最高的是“知网、SCI 等数据库”，占到 50.59%，其次是“百度学术”，占比为 20.12%，而“图书馆书目查询系统”的占比为 15.09%，“图书馆提供的发现平台”的占比为 14.20%。由此可以看出，大部分研究人员在选择文献查找渠道时，并没有选择图书馆提供的检索入口，而是直接选择经常使用的数据库以及百度学术。

表 2 样本人口学统计分析

项目	类别	人数	比例
性别	男	157	46.45%
	女	181	53.55%
身份	本科生	113	33.43%
	研究生	197	58.28%
	教师	28	8.28%
查找资料时优先使用哪个检索入口?	图书馆书目查询系统	51	15.09%
	图书馆提供的发现平台	48	14.20%
	知网、SCI 等数据库	171	50.59%
	百度学术	68	20.12%

数据收集整理后, 利用 spssau 数据处理平台, 对数据的信度、效度进行检验, 得到表 3。总量表 Cronbach' α 系数为 0.915, 17 个潜变量的 Cronbach' α 系数均大于 0.8, KMO 值为 0.731, $P < 0.001$ 。根据统计学研究观点, Cronbach' α 系数在 0.8 以上说明样本数据信度良好, KMO 值大于 0.7 说明效度良好。此外, 因子载荷与 CR 值均大于 0.7, AVE 值大于 0.5, 说明聚合效度良好。

表 3 因子平均值、因子载荷、Cronbach' α 系数、CR、AVE 值

因子	指标	平均值	因子载荷	Cronbach' α 系数	CR	AVE
信息质量	INFQ1	4.310	0.809	0.920	0.888	0.666
	INFQ2	4.405	0.876			
	INFQ3	4.333	0.720			
	INFQ4	4.238	0.851			
系统质量	SYSQ1	4.238	0.873	0.867	0.882	0.653
	SYSQ2	4.381	0.767			
	SYSQ3	4.238	0.844			
	SYSQ4	4.095	0.740			
服务质量	SERQ1	3.476	0.814	0.853	0.882	0.714
	SERQ2	3.810	0.836			
	SERQ3	3.976	0.884			
用户满意度	USQ1	3.976	0.767	0.823	0.842	0.642
	USQ2	3.905	0.738			
	USQ3	4.238	0.890			
用户使用意愿	UIQ1	4.048	0.896	0.894	0.898	0.746
	UIQ2	4.286	0.844			
	UIQ3	4.048	0.850			

2.4 结果与讨论

(1) 从图 1 平均值对比图可以看出, SERQ1、SERQ2、SERQ3、USQ1、USQ2 项, 分值较低, 在 4 分以下, 其他分值均在 4 分以上。最低的分值是 3.48, 对应的指标是“个性化功能(检索历

史自动保存/订阅推送服务/可视化分析、社交分享等)”，倒数第二分值为 3.81，对应指标是“实时咨询服务”，第三低的分值为 3.98，对应指标是“人工智能服务(拼写检查、检索词智能修正等)”。三项最低分值的指标均是“服务质量”指标，在测评的五家发现系统内，用户对系统提供的服务满意度都比较低。但对于五家发现系统的信息质量、平台系统质量的打分均在 4 分以上，用户的使用意愿均值也在 4 分以上，这说明基于信息质量、发现系统质量，用户的使用意愿还是满强烈的，只是服务质量的欠缺，造成用户的满意度稍低。

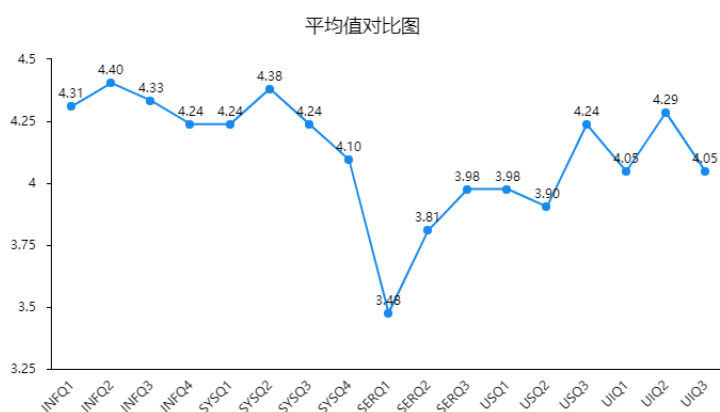


图 1 各指标平均值比较

(2) 在询问获取信息过程中哪个要素比较重要时，调查对象给出的答案排在前三位最多的分别是：全文获取度高、相关度排序合理、信息足够丰富。发现系统作为文献检索平台，除了尽可能准确全面揭示文献信息外，其重要的功能就是提供全文获取渠道，如果最终不能获得全文，用户的前期检索工作将失去很大意义，通过调研也可以看出，大部分科研用户很看重最后的全文获取这一步，所以怎样最大限度的实现这一功能，发挥平台效益，是平台未来发展的重中之重。此外，信息的丰富程度和检索结果的相关性也是用户比较在意的，有研究显示，人们在检索过程中往往只查看前 20 条检索内容，有些资源发现系统在相关文献结果排序过程中，侧重于自己拥有的资源与有商业合作的数据库资源，这种有倾向的排序难免会造成用户错过一些重要的资源。因此，资源发现系统怎样提升排序算法的合理性、科学性，是发现系统进一步需要解决的问题。

(3) 在询问最影响发现系统易用性的因素有哪些时，排在前三的选项是：界面干净明了，导航清晰；全文获取链接设置明显；有实时在线的参考咨询服务。对于检索界面干净明了，导航清晰的要求大部分发现系统都做得很好，而全文链接的设置上，在测评的 5 家发现系统中，超星发现链接了全国馆藏信息，并可以直接链接到 BALIS 参考服务平台。艾利贝斯也将旗下的 primo 发现系统与著名的图书馆馆际互借系统 RapidILL 进行技术上的对接整合，百度则是链

接各大数据库商，如果不在服务 IP 范围内，则需要付费获取全文。对于排序第三的“有实时在线参考咨询服务”这一项，有用户反应发现系统的宣传推广做得太少，用户对发现系统的使用、认知、熟练程度都很欠缺，在检索过程中也确实遇到过“检索失败”、“检索到 x 条结果但并未显示结果”、“检索方式不习惯”、“系统无纠错能力”等问题，这些问题的解决有赖于服务的加强，如“在平台中嵌入检索技巧讲解视频”、“平台本身具备一定的智能纠错(如拼写检查、检索词智能提示、修正、扩展等)”、以及设置“人工在线咨询服务”等。在测试过程中，超星发现在这一项上做得较好，有“人工在线咨询服务”，随时可以解答问题。

(4) 在询问“5 个检索发现系统中您优先给朋友推荐哪个”时，排序前三的是：Summon 发现系统、超星发现系统、百度学术搜索。从表 4 可以看出，在五家发现系统中，每家发现系统的侧重点与优势各不相同。Summon 发现系统在信息丰富度、相关性排序两方面分值最高；Primo 发现系统在信息描述与标引、全文在线揭示度分值最高；超星发现系统在个性化功能上分值最高，百度学术因为开放免费，简单易用，大部分用户表示“会继续使用此系统”、并“将系统作为检索信息的首选”。

表 4 五家发现系统测评指标分值比较

因子	指标	指标内容	平均值	Summon	Primo	ebscohost	超星	百度学术
信息 质量	INFQ1	信息类型的丰富度	4.310	4.468	4.341	3.951	4.411	4.379
	INFQ2	信息描述与标引全面清晰	4.405	4.303	4.581	4.356	4.455	4.33
	INFQ3	检索结果相关性排序准确合理	4.333	4.512	4.409	4.317	4.425	4.002
	INFQ4	信息全文的在线获取	4.238	4.418	4.423	4.312	4.261	3.776
系统 质量	SYSQ1	系统界面清晰明了	4.238	4.415	4.313	3.927	4.357	4.178
	SYSQ2	检索字段设置合理	4.381	4.439	4.418	4.146	4.435	4.467
	SYSQ3	检索结果分面导航功能	4.238	4.321	4.156	4.132	4.378	4.203
	SYSQ4	全文链接有效打开	4.095	4.61	4.293	4.012	4.151	3.409
服务 质量	SERQ1	个性化功能(检索历史自动保存/订阅推送服务/可视化分析、社交分享等)	3.476	3.231	3.158	3.083	4.013	3.895
	SERQ2	实时咨询服务	3.810	3.891	3.786	3.801	3.973	3.599
	SERQ3	人工智能服务(拼写检查、检索词智能修正等)	3.976	4.275	4.171	3.95	4.256	3.228
用户 满意 度	USQ1	检索过程方便顺利	3.976	4.031	4.029	3.921	4.015	3.884
	USQ2	检索过程轻松愉悦	3.905	3.979	4.053	3.823	3.872	3.798
	USQ3	对检索结果满意	4.238	4.271	4.268	4.213	4.235	4.203
用户	UIQ1	会继续使用此系统	4.048	3.902	3.75	3.683	3.925	4.98

使用 意愿	UIQ2	将系统作为检索信息的首选	4.286	4.331	4.251	4.173	4.252	4.423
	UIQ3	愿意参加系统的相关培训或活动	4.048	4.354	4.12	4.011	4.185	3.57
		总分		71.751	70.52	67.811	71.599	68.324

在调查问卷最后，很多积极的用户写下了这次测评中的使用感受：

对于 Summon 发现系统的开放评价主要有：信息丰富且使用方便，检索的内容很全面，相关度也较高，且具有非常丰富的筛选选项，效率很高，能直观的搜索到内容，也可以在线预览全文内容；检索界面清晰明了，信息导航清楚，层次感很好，不足点是有用户反应存在检索失败的情况。检索前必须进行身份认证，校外访问不方便、推广少，认知不够。

对于 Primo 发现系统的开放评价主要有：页面简洁、在菜单栏就能看到数据库页面，检索效率很高，搜索途径很丰富，不会出现链接卡死情况，很流畅；分类齐全、能很直观的搜索到检索内容，描述得也较为详细、但是检索内容不是很全面、总体使用感一般。检索前必须进行身份认证，校外访问不方便、推广少，认知不够。

对于 ebscohost 发现系统的开放评价主要有：文献检索的内容较少，但是整理出的数据库很全，分类清楚，全文获取链接非常醒目而且获取途径非常便捷，有拼写错误检查，相关度排序合理准确，但检索界面设计有待改进、检索时出现了“检索到×条结果”的字样，但并未显示结果，系统无纠错能力，个人觉得系统检索方式不习惯。检索前必须进行身份认证，校外访问不方便、推广少，认知不够。

对于超星发现系统的开放评价主要有：信息丰富度高，精炼检索高效、搜索结果直观清晰，资源来源丰富，使用顺畅，可以满足检索需要，页眉处有一个相关论文发表年份趋势图，非常有特色，而且可以更直观的感受相关研究热点的变化，并按照年份来筛选，让人耳目一新。除此之外检索的信息量非常大，检索全面，文献获取渠道丰富、相关度也很不错，筛选选项也非常多，检索设计风格也很简约，整体效率非常高、界面干净，信息导航清楚，可以凭账号密码登录，没有校内校外限制，感觉不错。

对于百度学术搜索的开放评价主要有：开放性大，使用方便，简单易上手，受众较广，实效性高，本人每次检索英文文献都会使用百度学术，因为它的 doi 号比较全，但是大多数情况下都要配合 scihub 使用，可以搜索到很多内容，可以从多个途径获取文献资源，但获取全文方式太过麻烦，且大部分文章都要收费，否则无法显示全文，检索界面不够层次分明，相关度一般。在没有发现系统时值得推荐，大众较常用该系统，doi 号较全，使用方便。

从上述评价中可以看出，发现系统与百度学术最大的不同是，无论哪种发现系统，在使用前都需要进行身份认证，受校内校外访问的限制，而百度则足够开放，但是在获取全文时，发现系统的优势呈现出来，可以直接获得全文或通过文献传递的方式获得。还有一个点是，发现系统普遍推广宣传得太少，用户普遍认知不够，这一点在调查时，用户多次多方面的反应，因此使用率、熟练度不高。但是通过该次测评，大部分用户承认，发现系统较百度学术与常用数据库，还是一个很好的选择的，通过此次测评与学习，以后会考虑把发现系统作为查找信息的主要入口。

3 结论与建议

3.1 加强元数据规范建设

全文的获取对资源发现平台来说固然重要，但文献资源的元数据建设才是首要重任。在对五家发现系统的测评过程中深刻体会到，元数据目录的精准、规范、全面而有序是用户检索及获取全文的基本保障，在竞争越来越激烈的市场环境下，商家难免会急于求成，盲目地追求元数据大而全，从而造成元数据缺少深层次标引，这会造成文献在检索中无法有效地聚合，宽泛的资源得不到深层聚类，反倒给用户带来信息冗杂的困扰。在 primo 资源发现系统中有一个用户反馈机制，该机制允许用户在检索过程中对元数据进行纠错与反馈，是对元数据维护工作的一个补充。这种机制值得在任何一款发现系统中进行推广与借鉴，否则，过多冗杂的元数据会将原本简单的发现变成复杂的发现。

3.2 提升文献信息的揭示度与获取途径

原文的获取渠道与途径是资源发现平台的又一项重要内容，它是在建立完善准确的书目数据库后，需要解决的内容建设。在测评的五家发现系统中，几乎每家发现平台都和拥有全文的数据库商进行了合作，大多是以数据库全文链接的形式来补充全文内容。需要注意的是，全文的链接一定要醒目、有效，维护的服务要及时。同时，对全文和图书的信息揭示也尤其重要，在调研的发现系统中，文献信息揭示度较好的是 primo 发现系统与超星发现系统，两家平台均可以图片形式查看到图书的目录页，超星发现同时在书目信息栏目下嵌入了读者推荐购买模块，增加了文献获取的途径。

3.3 馆际借阅及文献传递系统的嵌入

在测评的 5 家发现系统中，超星发现链接了全国馆藏信息，并可以直接链接到 BALIS 参考

服务平台。primo 发现系统与著名的图书馆馆际互借系统 RapidILL 进行技术上的对接整合，使用户可以在该馆际互借系统中提交文献资源申请单，与 RapidILL 类似的还有 Borrow Direct 馆际互借系统，也可以通过技术与发现系统进行整合使用。在信息时代的今天，“发现”是件上至国家层面的书目建设，下至每位用户对元数据进行纠错与反馈的事情，资源发现系统在完成元数据与全文获取建设后，急需发展的便是方便高效的文献共建共享服务，只有文献机构系统内的联盟合作、服务合作真正高效、有序地开展起来，才能为用户更好地实现“发现——获取”服务。

3.4 优化相关性排序

检索结果的相关性排序是非常复杂的内容，算法很多且不好平衡。同时，相关性排序的算法是发现系统的核心竞争力，很多产品不对外公开发布，因此，资源发现系统一直在结果排序上无法取得阶段性的进步。在测评的 5 家发现系统中，百度学术在检索结果上是按照文献引文量进行排序并呈现文献的引文脉络，同时推介一些热点文章。Primo 系统是通过用户的个性化信息定制或以数据的引用和使用统计情况来提升学术内容的检索质量，并申请了名为“ScholarRank”的技术专利，该方法还能让本地化资源得到相对较高的相关性评分值和优先排列序位。ebSCOHOST 系统的相关性排序方法以主题索引为优化原则，同时以全文和摘要信息支持更深层的索引，并通过对主题、篇名、刊名、关键词、摘要和全文等数据赋予不同的权重来实现检索结果的相关性评分和排序。Summon 系统主要通过对元数据或全文的词频以权重学习的方式支持相关性排序，但是由于不同资源类型的元数据存在分布的不均衡性，该系统比较偏向于优先发现报纸类的资源。

如何智能的识别出用户的检索初衷，需要计算机的深度学习，以及对文献的聚类分析。有向搜索的用户在检索过程中明确知道自己要查找的文献内容，即目标清晰明确；而无向搜索的用户在检索过程中是查找某一主题、学科或者研究对象的文献，并无清晰的题录信息(Tamar, 2013)。对无向搜索的用户来说，资源发现系统的结果排序就起到至关重要的作用。Boram & EunKyung (2016)指出同一条文献在不同的发现平台检索结果中其排序位置差异很大；有些资源发现系统在相关文献结果排序过程中，侧重于自己拥有的资源与有商业合作的数据库资源，这种有倾向的排序难免会造成用户错过一些重要的资源。因此，资源发现系统怎样提升排序算法的合理性、科学性，是进一步需要解决的问题。Ding et al. (2019)提出了基于用户日志分析进行的相关性排序算法，即统计用户的检索日志，分析出用户的学科、个性、背景后，根据该“用户画像”进行信息推介与排序。所以，为了更好的进行用户画像与信息相关度测算不仅需要收集

文献的共引共现信息(Wang, 2020, 16), 还需要进行用户的社交、科研圈子、科研学术交流的测度分析。

3.5 建立个性化功能模块

在测评的 5 家发现系统平台中, 有些平台是和本馆的“个人图书馆”进行了整合, 可以直接登录到“我的图书馆”。超星可以将文献分享到 QQ、新浪微博、微信, 百度学术除了分享到 QQ、新浪微博、微信外, 还可以链接到印象笔记与有道云笔记。而真正像 BiblioCore (BiblioCommons) 一样, 打造一个馆员推荐与评论功能模块的较少。或者像 Blacklight 与 VuFind 那样, 允许用户给书目添加 Tag, 建立自己的分类体系。这样个性化功能有利于图书馆进行“用户画像”的数据收集。“随着图书零售渠道的逐渐减少, 为读者推荐有用的而他们可能不知道的图书变得越来越重要。”BiblioCommons 联合创始人 Betn Jefferson 说: “图书馆员可以在图书书目中根据他们的专业知识做一些推荐, 这样能够帮助读者注意到更多他们可能感兴趣但自己又可能发现不了的图书” (Brandi, 2014)。在信息指数增长的今天, 仅仅靠图书领域专家的力量与商业的力量也不足以将文献信息进行有序的整合、聚类与共享, 只有打造这样个性化功能模块, 才能让文献的查找与获取变得更加轻松可行。

3.6 利用衍生数据推动精准服务

Andrew et al. (2013) 提出, 与发现资源相比, 评估资源与选择资源变得更为重要了, 因此, 资源发现系统的聚合功能与深层次的知识发现功能尤为重要。目前, 中外文发现系统的聚合功能更多的是在同一载体、同一著者、同一概念间的聚合, 而对于基于知识的数据关联和知识发现远远不够。百度学术是通过引文共引共现以及高被引文章向用户推送, 而超星则是通过可视化图谱聚合文献。相比而言, 外文发现系统的文献聚合功能更多是对元数据的依赖与去重整合, 分面检索的聚合功能大同小异, 主要有学科、作者、主题词等 (Peng & Zheng, 2016)。图书馆应挖掘利用发现系统的衍生数据, 如用户身份数据、检索行为数据、文献使用数据, 通过这些数据的利用与分析, 可以得出用户画像、文献聚合与共现情况, 以此数据来推动服务, 从而形成“用户产生数据、数据推动服务”的良性迭代。

4 项目成果 (发表的文章、开发的软件、取得的实践效果等)

文章《资源发现系统测评分析》待发表

5 参考文献

- [1] 刘洋.我国高校图书馆资源发现系统现状调查——以“985 工程”院校为例[J].河北科技图苑,2016,29(04):86-90+96.DOI:10.13897/j.cnki.hbkjty.2016.0113.
- [2] Chickering, F. W., & Yang, S. Q. (2014). Evaluation and Comparison of Discovery Tools: An Update. *Information Technology and Libraries*, 33(2), 5-30. <https://doi.org/10.6017/ital.v33i2.3471>
- [3] Yang, S.Q. and Wagner, K. (2010), "Evaluating and comparing discovery tools: how close are we towards next generation catalog?", *Library Hi Tech*, Vol. 28 No. 4, pp. 690-709. <https://doi.org/10.1108/07378831011096312>
- [4] 成颖.基于相关性判据的学术信息检索系统成功模型建构[J].现代图书情报技术,2011(09):46-53.
- [5] 赵钊, 孙伟新, 赵珊珊. 基于 D&M 模型的信息系统成功评价研究[C]国际信息系统协会中国分会. 国际信息系统协会中国分会, 2015.
- [6] Sabeh, Hala Najwan et al. "A Systematic Review of the DeLone and McLean Model of Information Systems Success in an E-Learning Context (2010 - 2020)." *IEEE Access* 9 (2021): 81210-81235.
- [7] Petter, S., DeLone, W. & McLean, E. Measuring information systems success: models, dimensions, measures, and interrelationships. *Eur J Inf Syst* 17, 236 - 263 (2008). <https://doi.org/10.1057/ejis.2008.15>
- [8] Clough, Paul D. and Paula Goodale. "Selecting Success Criteria: Experiences with an Academic Library Catalogue." *Conference and Labs of the Evaluation Forum* (2013).
- [9] 邓小昭等,《网络用户信息行为研究》,2010,科学出版社出版.
- [10] 窦天芳,姜爱蓉.资源发现系统功能分析及应用前景[J].图书情报工作,2012,56(07):38-43.
- [11] 杨薇,曾丽军.从“快传”(RapidILL)和“立借”(Borrow Direct)看馆际互借与文献传递服务体系的发展[J].大学图书馆学报,2018,36(04):18-23+44.DOI:10.16603/j.issn1002-1027.2018.04.003.
- [12] 黄静.变革中的文献传递服务:案例剖析与路向管窥[J].图书情报工作,2013,57(07):55-59.
- [13] Tamar, S. (2013).“From Search to Discovery”. *Bibliothek Forschung und Praxis*, 39(2), 212-224.<https://doi.org/10.1515/bfp-2015-0028>
- [14] 王连喜.知识发现系统的相关性排序与主题聚类功能问题探析[J].图书馆工作与研究,2015,(12):56-60.
- [15] Marshall. The Future of Library Resource Discovery. 2015.http://tefkos.comminfo.rutgers.edu/Courses/e553/Articles/Articles%20Sp15/NISO%20report%20ofuture_library_resource_discovery%202015.pdf
- [16] Boram & EunKyung (2016)Lee, B. and E. Chung, An Analysis of Web-scale Discovery Services From the Perspective of User's Relevance Judgment. *The Journal of Academic Librarianship*, 2016. 42(5): p. 529-534.<https://doi.org/10.1016/j.acalib.2016.06.016>.
- [17] 丁梦晓,毕强,许鹏程,李洁,牟冬梅.基于用户兴趣度量的知识发现服务精准推荐[J].图书情报工作,2019,63(03):21-29.DOI:10.13266/j.issn.0252-3116.2019.03.003.
- [18] 汪滢.基于用户日志分析的搜索引擎相关排序算法优化[J].电脑知识与技术,2020,16(18):99-101.

- [19] Brandi, Scardilli. 2014. BiblioCommons catalog-centric library operations. *Information Today*, 31(7), 23.
- [20] Andrew D. Asher, Lynda M. Duke, and Suzanne Wilson. 2013. Paths of discovery: Comparing the search effectiveness of EBSCO Discovery Service, Summon, Google Scholar, and Conventional Library Resources. *College & Research Libraries*, 9, 464-488.
- [21] <http://pdfs.semanticscholar.org/c4a1/08fd88e9d593c575285e9a341e5247b62901.pdf>
- [22] 彭佳, 郑巧英. 信息资源聚合与组织研究——以发现系统为例 [J]. *图书馆杂志*, 2016, 35(03):80-85. DOI:10.13663/j.cnki.lj.2016.03.014.
- [23] Delaney, T.G. and Richins, M. (2012), "RapidILL: an enhanced, low cost and low impact solution to interlending", *Interlending & Document Supply*, Vol. 40 No. 1, pp. 12-18. <https://doi.org/10.1108/02641611211214233>
- [24] Kronenfeld, M. (2015). Library resource discovery. *J Med Libr Assoc*, 103(4), 210-213. <https://doi.org/10.3163/1536-5050.103.4.011>
- [25] Kemperman, Suzanne Saskia, Bill Brembeck, Elizabeth W. Brown, Alexandra de Lange-van Oosten, Theodore Fons, Catherine Giffi, Noah Levin, Alistair Morrison, Carlen Ruschoff, Gregg A. Silvis, and Jabin White. 2014. Success Strategies for Electronic Content Discovery and Access: A Cross-Industry White Paper. Dublin, OH: OCLC. <http://www.oclc.org/content/dam/oclc/reports/data-quality/215233-SuccessStrategies.pdf>.
- [26] Stublely, P. and Kidd, T. (2002) Questionnaire surveys to discover academic staff and library staff perceptions of a National Union catalogue. *Journal of Documentation*, 58 (6). pp. 611-648. ISSN 0022-0418 <https://doi.org/10.1108/00220410210448183>
- [27] Lundrigan, C., Manuel, K., & Yan, M. (2015). "Pretty Rad" : Explorations in User Satisfaction with a Discovery Layer at Ryerson University. *College & Research Libraries*, 76(1), 43-62. <https://doi.org/10.5860/crl.76.1.43>