

# 文献资源发现系统分面检索功能比较研究\*

林鑫<sup>1,2</sup> 吴茜<sup>2</sup>

(1.中国科学技术信息研究所,北京 100038; 2.华中师范大学信息管理学院,武汉 430079)

**摘要:**分面检索是文献资源发现系统的重要服务,其功能设计是影响用户体验的关键因素。为了解文献资源发现系统分面检索功能设计现状、存在的主要问题,选取8个代表性发现系统为对象,从分面体系框架及焦点设置、分面及焦点排序策略、分面体系展现策略和交互机制4个方面进行调研与比较分析。结果显示,各系统的分面检索功能设计区别显著,存在一些共性问题亟待解决。建议推进分面体系的规范化、系统化与语义化,采用动静结合策略优化分面及焦点排序效果,实施分面体系展现的智能化动态控制,提升分面交互的灵活性与可理解性。

**关键词:**文献资源发现;资源发现系统;分面检索;功能设计

**中图分类号:** G252.7

**DOI:** 10.3772/j.issn.1673-2286.2019.09.003

2009年Serials Solutions公司发布全球第一个网络级文献资源发现系统Summon,自此之后,文献资源发现系统迅速在全球范围内得到普及<sup>[1]</sup>。在各文献资源发现系统提供的服务中,分面检索一直是非常重要的文献资源发现服务。与传统的关键词检索系统仅关注查全查准和检索结果的排序不同,功能设计是影响分面检索服务体验的关键部分<sup>[2]</sup>。基于此,如何从功能设计角度优化用户体验一直是文献资源分面检索研究与实践关注的重要主题。

相关研究主要从两个方面入手,一是围绕文献资源分面检索中的具体功能设计进行优化方案设计。包括分面体系的扩充完善,如胡昌平等<sup>[3]</sup>将副主题词作为分面加入期刊论文分面检索体系,陈雪<sup>[4]</sup>将图书标签、用户评分、图书热度作为分面加入图书分面检索体系,Zhang等<sup>[5]</sup>提出应将格式、作品集加入分面体系,Fafalios等<sup>[6]</sup>采用访谈法确定专利分面检索中最有用的9个分面(应用者、发明者、发布时间等);已有分面或交互机制的完善,如田宁<sup>[7]</sup>提出推进检索结果的可视化展现,Ben-Yitzhak等<sup>[8]</sup>建议给出分面取值的质量信息以更好地支持用户决策(比如作者分面中给出每个作者的畅销书数量)。二是对已有文献资源分面检

索实践进行调研,归纳总结最佳实践或分析当前存在的问题并提出改进建议。典型研究包括:张福俊等<sup>[9]</sup>以Trove、DPLA、Europeana、WDL和文津搜索为对象,崔明等<sup>[10]</sup>以百度学术和Summon为对象,许丽媛等<sup>[11]</sup>以Elsevier、Springer、CNKI、Primo、Summon和EDS为对象,从检索方式、分面设置进行调研和比较分析,并提出了优化建议;郑雯译<sup>[12]</sup>对EDS、Summon、Primo 3个分面检索系统的分面设置,以及文献类型、主题、数据库和出版日期4个分面的焦点设置进行比较分析;Hall<sup>[13]</sup>于2012年分别对美国100家高校图书馆和100家公共图书馆的OPAC系统的分面设置情况进行调研,包括上线分面检索系统的情况,以及分面的个数、常见分面、分面的排序等,并提出主题分面的设置需要进一步优化以解决信息过载问题。

总体而言,相关研究具有现实性,对文献资源发现中的分面检索服务优化具有参考价值,但研究的全面性、系统性存在一定不足,一是分面检索功能相关研究集中于个别功能的优化,或者分面设置介绍、部分分面的焦点设置,对分面及焦点的排序、展现控制机制、交互机制等方面关注不足;二是相关调研对象在选取的代表性上有所欠缺,大多数研究集中于Summon、Primo、

\*本研究得到博士后基金“文献资源发现服务体系及实现路径研究”(编号:2018M641445)资助。

EDS、超星发现等系统,没有对当前的主流学术搜索平台和公益性发现系统进行全面分析。基于此,本研究拟选取代表性文献资源发现系统,围绕其分面检索功能设计进行全面、系统地调研与比较分析,梳理当前功能设计的现状及共性问题,并提出优化完善建议,为文献资源发现中的分面检索服务体验优化提供参考。

## 1 调研设计

在文献资源发现系统建设实践中,虽然同时存在综合性和针对具体文献类型的专门性文献资源发现系统,但鉴于综合性文献资源发现系统应用更广泛且支持单一文献类型的资源发现,因此为便于横向比较分析,本次调研与比较分析聚焦于综合性文献资源发现系统。

### 1.1 调研对象选取

调研对象选取时,一方面要求必须是当前的主流系统,从而保证其能够得到有效的升级、维护,反映文献资源发现系统分面检索功能设计的最新动向;另一方面要求样本的整体结构合理,能较全面地反映文献资源发现系统的全貌。依据这两个原则,最终选择百度学术、超星发现、文津搜索、E读、Find plus、Summon、EDS和Primo 8个发现系统作为调研对象,从而既涵盖了主流的商业性发现系统,也涵盖了影响较大的公益性发现系统;既涵盖了国外引进的全球流行的发现系统,也涵盖了本土研发的知名发现系统。需要说明的是,机构在应用Find plus、Summon、EDS和Primo发现系统时大多进行了个性化设置和二次开发,基于此,进一步选择了武汉大学珞珈学术搜索、中国海洋大学行之远学术搜索、中山大学智慧搜索和北京师范大学木铎搜索作为对象。

### 1.2 调研策略

从国内外分面检索功能相关研究来看,分面检索功能设计较核心的模块包括分面体系设置、分面及焦点排序、交互界面及机制设计。基于此,调研与比较分析从这4个方面展开,具体为:①分面体系框架及焦点设置,即有哪些分面,以及每个分面有哪些焦点(focus)或如何产生焦点,这是最基础的功能;②分面及焦点排序策

略,即多个分面之间如何排序,同一个分面内部各个焦点如何排序;③分面体系展现控制策略,重点关注分面体系何时展现及何时隐藏,默认状态下分面体系如何展现;④分面体系交互机制,即用户如何通过与分面体系的交互对检索结果进行探索、筛选。

鉴于无法全面获悉各文献资源发现系统分面检索功能设计的资料,因此只能通过其外在表现对其进行功能设计策略推断,为尽可能全面地分析各发现系统的分面检索功能设计,在调研实施中采取以下3种方法:

①调研中以综合检索入口下的功能为主,兼顾面向专门文献类型的分面检索功能设计,如图书、学位论文、期刊论文等;②使用覆盖不同学科领域、不同元数据字段的多个关键词作为初始检索词进行检索;③对于检索结果,采用多种分面及焦点组合方式进行交互,观察其交互机制。

## 2 文献资源发现系统分面检索功能比较及问题

在对8个文献资源发现系统充分调研的基础上,对分面检索功能设计思路进行归纳总结与比较分析,并在此基础上分析其存在的共性问题,下文将从分面体系框架及焦点设置、分面及焦点排序策略、分面体系展现控制策略、分面体系交互机制4个方面进行分析。

### 2.1 分面体系框架及焦点设置

分面体系是影响分面检索系统用户体验的基础和关键要素,也是本次文献资源发现系统分面检索功能比较关注的重点。比较分析中,重点关注了分面体系框架和焦点的设置策略。

(1)分面设置及特点。在分面设置上,8个发现系统设置了年份、语言、资源类型、关键词等21个不同的分面(调查过程中,对于焦点相同或相似,仅名称不同的分面合并为一个),各发现系统的分面设置如表1所示。

可以看出,文献资源发现系统在分面设置上有3个特征。①各发现系统在分面设置上区别较为显著。8个系统均涵盖的分面有年份和资源类型2个,占全部分面的9.52%;而中图分类、学位授予时间等有1个或2个系统涵盖的分面数量为9个,占全部分面的42.86%。②结合文献资源类型进行分面设置。如核心期刊、期刊分面是关于期刊文献的,学位级别、导师等分面是关于学位论文

表1 文献资源发现系统常用分面设置及排序比较

发现系统	百度学术	超星发现	文津搜索	E读	Find plus		Summon	EDS		Primo
					中文	外文		中文	外文	
显示范围限定	1	1	-	1	1	1	1	1	1	1
年份	2	4	2	2	4	2	3	2	2	10
语言	-	2	4	3	-	-	6	6	8	9
资源类型	7	3	1	4	3	-	2	3	3	2
关键词	6	5	-	-	5	-	5	-	-	5
期刊	9	10	-	-	-	-	-	5	6	11
作者	8	8	3	-	2	-	-	-	-	6
作者单位	10	9	-	-	-	-	-	-	-	16
学科	3	6	-	5	6	-	4	4	4	15
收藏机构	-	-	-	7	-	-	7	8	10	4
来源数据库	-	-	5	6	-	3	-	9	11	8
获取方式	5	-	-	-	-	-	-	-	-	-
基金	-	12	-	-	-	-	-	-	-	-
核心期刊	4	7	-	-	-	-	-	-	-	-
馆藏位置	-	-	-	-	-	-	-	7	9	3
导师	-	-	-	-	-	-	-	-	-	7
中图分类号	-	-	-	-	-	-	-	-	-	13
学位级别	-	-	-	-	-	-	-	-	-	14
出版机构	-	-	-	-	-	-	-	-	5	12
学位授予时间	-	-	-	-	-	-	-	-	-	17
国家地区	-	11	-	-	-	-	-	-	7	-

注：①数字表示该分面在发现系统分面体系中的顺序，“-”表示不含该分面；②显示限定范围，该分面各系统设计不一，一般用来对检索结果进行范围限定，如检索信息是否可全文访问、结果为电子资源或馆藏目录；③资源类型，指检索结果的文献类型，一般分为期刊、图书、报纸；④期刊，收录检索结果的相关期刊名称；⑤获取方式，获取文献的具体方式，分为免费下载、付费下载或者登录查看；⑥核心期刊，期刊索引名称或机构发表的核心期刊目录名称

的，但绝大部分发现系统更关注期刊论文，对其他类型资源的个性化特征考虑不足。③从分面内容角度看，各发现系统均涵盖了文献的形式特征和内容特征，并且在形式特征利用上较为充分，不但数量丰富，而且对部分形式特征进行了二次加工。如将论文的来源期刊进一步按照被索引数据库的收录情况作为分面进行提供。

(2) 焦点设置策略。常见的焦点设置策略：一是已有元数据抽取，如学术论文的作者；二是基于元数据的元数据抽取，如学术论文所属期刊被索引情况；三是根据分面体系构建的需要自行设置，将时间分面分为“2018年以来”“2017年以来”“2016年以来”等。从调研情况看，绝大多数分面都采用了直接从元数据中抽取的策略，收藏机构、馆藏位置、核心期刊这些少量分面采用了第二种策略，仅有年份采用了自行设置策略。需要指出的是，对于同一个分面，各发现系统可能采用不同的分面设置策略，如对于年份分面，文津搜

索、E读、Find plus采用了元数据抽取策略；百度学术、Summon、EDS、Primo采用的是自行设置策略；而超星发现则是采用的元数据抽取和自行设置策略并用。

(3) 分面体系框架及焦点设置中存在的主要问题。从当前实践情况看，无论是分面体系框架还是焦点设置均存在较突出的问题，在引发信息过载的同时难以有效支持用户对检索结果的快速过滤和探索，存在的主要问题包括以下方面。第一，部分焦点数量较多的分面缺乏合理的亚面体系构建，导致信息过载严重，难以有效发挥作用。如关键词（主题词）分面，当前仅是文献关键词（主题词）的罗列，缺乏系统的分面化改造。第二，分面体系框架不够完善，缺乏对文献资源内容、质量的全面揭示。如揭示图书的适用对象和用途，将能够帮助用户进行高效决策。第三，对焦点的语义信息揭示不足，导致分面利用困难，既包括对同义词、多义词焦点的不当处理（如同名作者、同义关键词等），也包括

对焦点背后含义揭示的不足(如用户常常通过作者判断文献的权威性,但焦点设置中常常是作者的罗列,并未区分哪些是权威学者)。

## 2.2 分面及焦点排序策略

合理的分面及焦点排序策略对减少分面体系带来的信息过载、提升分面体系的利用效率具有重要作用,分面与焦点是分面体系中的两个不同层次,故而分别从两个方面对各分面系统进行排序策略比较,并进而分析存在的共性问题。

(1) 分面排序策略。常见的排序策略可以分为固定顺序和动态排序两类,两者的区别在于对于不同的检索词,分面间的顺序是否会动态调整。从调研结果看,所有被调研的发现系统均采用固定顺序的排序策略,以综合检索入口为例,各系统的分面顺序见表1。显然,各发现系统仅在少量分面的排序位置上达成了一致,如显示范围限定分面应该是一个排序比较靠前的分面,来源数据库分面排序应靠后一些;但在分面的整体顺序上其认识远未达成一致,甚至相距甚远。又如百度学术将分面“学科”放到了第3位,而武汉大学珞珈学术搜索(Find plus)将其排到了最后。

(2) 焦点排序策略。与分面排序策略相近,焦点也存在固定顺序和动态策略。调研发现,8个发现系统均采用了多种焦点排序策略,对于一半以上的发现系统均涵盖的10个热门分面,其排序策略分别为:显示范围限定分面均采用了人工指定顺序策略;关键词、期刊、作者、收藏机构和来源数据库5个分面均采用了动态排序策略;年份分面,部分发现系统采用了时间顺序策略,部分采用了人工指定顺序策略;语言、资源类型和学科

3个分面均是部分发现系统采用人工指定顺序策略,部分采用动态排序策略。总体来说,发现系统在焦点排序策略上有4个特点:①所有的发现系统均综合采用了固定顺序和动态排序相结合的焦点排序策略,而且大多数是以动态分面排序策略为主;②对于同一个分面,不同系统可能采用不同的分面排序策略;③采用固定排序策略时,针对不同的分面会采用差异化的策略;④当采用动态排序策略时,所有发现系统均采用了按相关文献的数量降序的排序策略。

(3) 分面及焦点排序策略存在的问题。分面排序中,对于只有少量分面的文献资源发现系统,采用固定顺序具有一定合理性;但分面数量较多时,依然采取纯固定顺序的排序策略,则可能导致分面查找困难,利用效率不高。同时,这些文献资源发现系统在分面顺序设置上未进行充分调研与分析,导致系统之间对同一个分面的排序差异很大,与用户利用习惯不符;加之分面均纵向展现于页面左侧,对用户体验影响较大。

焦点排序上,各文献资源发现系统均根据分面特点采用了固定顺序和动态排序相结合的策略,但在动态排序上均采用按照相关结果数排序的策略,未考虑分面自身的特征,从用户需求出发进行排序策略设计,导致一些分面的焦点排序效果不佳(如期刊分面中,载文最多的期刊常与领域权威的几种期刊难以重合)。

## 2.3 分面体系展现控制策略

除排序策略外,合理的分面展现控制策略也可以大幅优化用户体验,降低信息过载风险。围绕这一问题,本次比较分析从分面展现与否、无结果焦点展现、单一焦点的分面展现、焦点相关结果数量展现、分面展现数

表2 发现系统分面展现控制策略比较

系统	是否动态控制分面展现	是否展现无结果焦点	单一焦点分面是否隐藏	相关结果数量是否展现	展现分面上限	分面默认焦点上限
百度学术	是	是	否	部分展现	9	3
超星发现	是	否	否	是	12	14
文津搜索	是	否	否	否	5	5
E读	是	否	否	是	7	5
Find plus	是	否	否	是	6	5
Summon	是	否	否	是	7	5
EDS	是	否	否	是	11	6
Primo	是	否	部分隐藏	是	17	5

量、焦点展现数量6个方面进行比较分析,如表2所示。

整体来说,各文献资源发现系统的分面体系控制策略相似性较高。检索结果为非空时,均会展现分面体系;除百度学术以外,其他7个系统结果数为零的焦点会被隐藏显示;在焦点相关结果数量的展现上,文津搜索默认状态下不展现,而在结果数量少于50万时,可以勾选显示结果数量的复选框进行展现,其他7个系统中,除了百度学术的部分分面外,均会展现焦点相关的结果数量。然而,在分面和焦点的默认展现数量方面,尽管半数发现系统的分面数量为5~7个,多数发现系统的默认展现焦点数量在5个左右,但其波动较大。如文津搜索展现5个分面,而Primo可能多达17个分面;百度学术将默认展现的焦点数量控制在3个以内,而超星发现可能多达14个。此外,调查还发现,对排序靠后的分面,默认状态下Summon和EDS只展现了分面名称,未展现焦点,而其他6个系统对所有分面均在默认状态下展现了焦点。

当前,尽管所有的文献资源发现系统均对分面体系展现进行了动态控制,但智能化程度依然不高。①在分

面体系展现与否的控制上,仅从有无检索结果角度进行控制,而未从分面效用角度进行更智能化的控制。如检索结果较少时无须展现分面体系,无结果焦点、只有1个焦点的分面都无须展现等。②在每个分面的焦点展现数量控制上较为机械。一方面除Summon和EDS外,其他6个文献资源发现系统均在默认状态下对所有分面展现焦点,从而引发一定程度的信息过载,如超星发现系统默认展现的焦点数量高达近百个;另一方面,在默认状态下的焦点展现数量控制上应该灵活一些,尽量在不造成信息过载的前提下,将绝大部分用户可能选择的焦点全部展现出来。

## 2.4 分面体系交互机制

分面检索与关键词检索的典型区别在于用户可以通过与检索结果的继续交互实现对结果的过滤和探索,因此分面交互机制对用户体验具有重要影响。8个发现系统的分面体系交互机制比较见表3。

表3 发现系统分面交互机制比较

发现系统	同一分面内		多分面 焦点逻辑 关系	单次检 索可选 焦点数	通过分面体系进行交互后		
	是否支持 焦点多选	焦点间 逻辑			各分面 变化情况	已选焦点 删除方式	焦点 添加方式
百度学术	否	-	与	单个	当前分面隐藏其他焦点,其他分面刷新	每次删除1个焦点	已选分面无法再选,其他分面点击焦点即可
超星发现	是	或	与	多个	当前分面不变,其他分面刷新	一体化,并支持同时操作多个焦点	
文津搜索	是	或	与	单个	当前分面不变,其他分面刷新	每次删除1个焦点	点击新焦点即可
E读	是	或	与	单个	当前分面不变,其他分面刷新	每次删除1个焦点	点击新焦点即可
Find plus	是	或	与	单个	部分分面仅刷新其他分面,部分全部刷新	每次删除1个焦点	点击新焦点即可
Summon	是	或	与	单个	当前分面不变,其他分面刷新	每次删除1个焦点	点击新焦点即可
EDS	是	或	与	单个	所选焦点置顶于该分面内,其余焦点下移,其他分面刷新	每次删除1个焦点	点击新焦点即可
Primo	是	或	与	多个	若所选分面焦点互斥,则隐藏该分面并刷新其他分面;否则根据新结果刷新分面	每次删除1个焦点	点击新焦点即可

可以看出,尽管各发现系统在分面交互机制设计上不尽相同,但其存在较高的相似性。第一,除百度学术外,其他7个系统均支持用户在同一分面内选择多个焦点,而且其逻辑关系均是焦点间进行“或”运算。第二,超星发现和Primo支持用户一次选择多个焦点,其交互方式均是用户在通过复选框选择焦点后,再点击按钮对结果进行过滤;其他6个发现系统仅支持用户一次选择1个焦点,其交互方式也一致,均是点击焦点

后直接对结果进行过滤。第三,对于不同分面的焦点,各个系统的交互策略一致,均是在焦点间进行“与”运算。第四,当用户选择焦点发起结果过滤请求后,各发现系统均会根据新的检索结果刷新其他分面,而且所选焦点对应分面的处理上差异较大,包括当前分面保持不变、隐藏当前分面、根据当前检索结果刷新当前分面等。第五,用户选择焦点发起结果过滤请求后,超星发现实现了删除已选焦点和选择新焦点操作的一体

化,即通过对复选框操作后重新发起过滤请求实现;其他发现系统则通过在结果页删除相应条件实现已选焦点的删除,其中百度学术的同一分面无法再选其他新焦点,其他分面新焦点的选择与首次选择操作相同,除此的其他系统在选择新焦点的操作上都与首次选择操作相同。

分面体系交互机制方面的问题主要集中于单个分面的交互机制。显然,类似百度学术不支持单个分面内选择多个焦点的做法是不可取的,而对于支持选择多个焦点的发现系统,还存在交互机制不够灵活、可理解性不足的问题。对于前者,单个分面的多个焦点之间,用户除了进行“或”运算需求外,现在也存在“与”运算和“非”运算需求,如检索同时包含两个关键词的文献,所有的文献资源发现系统的分面检索功能均无法满足这类需求。对于后者,为支持单个分面的多焦点选择,文献资源发现系统均采取保持当前分面不变(除非分面内焦点互斥),刷新其他分面的策略,这既导致用户难以理解为何当前分面未随着焦点的选择而进行刷新,也导致用户无法利用当前分面对当前检索结果进行过滤和探索。

### 3 文献资源发现系统分面检索功能优化建议

通过调研和比较分析可以发现,8个国内外典型的文献资源发现系统在分面检索功能设计上差异显著,这从一定程度上也体现出其发展仍不成熟;在功能设计的各个方面还发现了诸多需要优化之处。针对这些问题,需要从完善分面体系设置、优化分面及焦点排序策略,推进分面展现的智能控制、提升分面体系交互的灵活性和可理解性4个方面进行全面优化。

#### 3.1 推进分面体系的规范化、系统化与语义化

良好的分面体系构建是保障分面检索用户体验的基础和关键,针对当前存在的问题,需要从分面词汇的规范化,以及分面体系内容框架的系统化与语义化推进等方面进行分面体系完善。第一,借助已有的控制词表、规范文档等资源,以及同义词挖掘、语义相似度计算、名称消歧等自然语言处理技术进行焦点词表的规范化;第二,从用户认知框架出发,按照文献类型、学科领域完善文献资源分面体系框架,加强对文献内容特

征、质量特征的揭示,辅助用户更好地进行检索结果的探索与过滤;第三,对文献主题(关键词)、来源期刊、作者等分面进行基于语义的分面化改造,实现其内部焦点的系统化,降低用户查找焦点的成本。具体而言,既需要利用焦点间的语义关联,如主题词(关键词)间丰富的语义关系;也需要利用焦点自身的属性信息,如来源期刊所属学科、被索引数据库收录情况、权威性 etc 等语义信息。

#### 3.2 动静结合优化分面及焦点排序效果

鉴于文献资源的分面及部分分面的焦点数量较多,因此适宜采用固定顺序与动态排序相结合的策略。第一,对于静态排序部分,需要结合分面特点区分哪些适合固定顺序,并进行广泛、深入的用户调研,形成适合大多数用户使用习惯的分面排序。除了需要针对多文献类型的综合检索场景下进行处理外,还需要结合不同文献类型的特征,进行单一文献类型场景下的分面排序设置。第二,在分面的动态排序策略设计中,需要以用户效用最大化为出发点,综合考虑检索结果的分布、用户行为习惯特征两类因素。第三,焦点动态排序中,需要改变仅考虑相关结果数量的思路,从用户视角进行差异化排序策略设计,以改进用户体验。如作者、刊名、作者单位等分面,可以按其领域权威性排序,主题词分面可以将检索词与焦点词间的语义关系考虑进去。

#### 3.3 实现分面体系展现的智能化动态控制

在分面体系展现控制上,各发现系统采用的是资源驱动的机械策略,即只要结果对应的分面体系非空就对其进行展现,导致信息过载及用户认知负担的不断加剧。为解决这些问题,需要借鉴相关研究成果<sup>[14-15]</sup>,推进基于用户效用的智能控制。第一,结合用户查询意图识别与检索结果数量及分布,动态判断是否需要展现分面体系。只有在特定的查询意图下且检索结果数量较多、分布较为分散时,分面体系才能起到快速过滤结果或帮助用户进行检索结果进一步探索的作用,这才需要展现分面体系,否则,分面体系的展现反倒会给用户带来干扰。第二,结合用户查询意图、检索结果分布动态判断相关分面的展现样式,一方面,对于不同重要程度的分面采用不同的展现样式,比如对于只有单个焦点对应的结果非空时,隐藏该分面;对于不重要的分面默

认状态下不展现其焦点。另一方面,对于重要分面,需要在满足阈值限制的条件下,动态调整展现的焦点数量,减少不重要焦点引发的信息过载问题。第三,对于多层结构的分面,需要根据检索结果分布进行层级的动态调整,以尽可能实现分面的扁平化。如一个亚面相关的文献集中于其中1个或少量几个焦点上时,则在分面体系中剔除该亚面,而将其焦点上移,以减少分面体系的层级。

### 3.4 提升分面交互的灵活性与可理解性

用户通过分面体系与检索结果的交互需求可以抽象为焦点间的“与”“或”“非”逻辑运算,但当前仅支持分面内焦点的“或”运算和分面间焦点的“与”运算,且存在部分发现系统不支持分面内选择多个焦点的情形,因此交互机制上存在灵活性不足问题。基于此,在交互机制设计上应进一步提升其灵活性,支持更加丰富的逻辑运算模式。同时,对可理解性的提升也是改善交互体验的重要方面,一方面需要全面推进分面筛选后根据检索结果刷新分面体系的设计模式;另一方面可以借鉴可视化技术优化交互机制,如借鉴Kajiyama等<sup>[16]</sup>所提出的设计方案,以立方体结构进行分面体系呈现,在结果上方呈现分面名称的同时,以颜色加以区分检索中使用到的分面,在结果外围用同色方框显示检索词并在其左右显示临近焦点,帮助用户更加便捷地了解检索词和相关焦点,提高交互性。

## 4 结语

围绕分面检索功能设计,选取Primo、Summon、百度学术、文津搜索等8个国内外文献资源发现系统为对象,从分面体系框架及焦点设置、分面及焦点排序策略、展现控制和交互机制4个方面进行了比较分析。结果发现,各发现系统在分面体系框架上彼此差异较大,规范性不足,同时普遍存在资源内容特征、质量特征揭示不足的问题;在焦点设置上常用元数据抽取的生成策略,这在降低实现成本的同时,也引发了信息过载问题;分面的排序均采用了固定顺序的策略,但分面顺序确定的随意性较强,用户体验不理想;焦点的排序采用了动态与静态相结合的策略,但动态排序策略普遍采用了按焦点相关结果数量降序的机制,缺乏对各分面个性化特征的考虑;在展现控制上普遍缺乏结合用户

查询意图与检索结果的动态调整机制,交互机制上灵活性不足,仅支持分面内的“或”和分面间的“与”运算,并且部分设计可能带来可理解性问题。针对这些问题,在分面检索功能设计实践中,需要充分遵循用户的认知规律,推进分面体系的规范化、系统化与语义化,动静结合优化分面及焦点排序效果,实施分面体系展现的智能化动态控制,提升分面交互的灵活性与可理解性。

### 参考文献

- [1] 张毅. “211工程”高校图书馆发现系统调研与思考[J]. 数字图书馆论坛, 2018(12): 41-48.
- [2] TRAPIDO I. Library discovery products: discovering user expectations through failure analysis[J]. Information Technology and Libraries, 2016, 35(3): 9-26.
- [3] 胡昌平, 林鑫, 陈果. 科技文献副主题词抽取及其在分面检索中的应用[J]. 情报学报, 2014(8): 837-845.
- [4] 陈雪. 基于豆瓣图书信息资源整合的OPAC分面检索优化[J]. 图书馆学研究, 2016(24): 47-54.
- [5] ZHANG H, DURBIN M, DUNN J, et al. Faceted search for heterogeneous digital collections[C]//Acm/ieee-Cs Joint Conference on Digital Libraries. ACM, 2012: 425-426.
- [6] FAFALIOS P, SALAMPASIS M, TZITZIKAS Y. Exploratory Patent Search with Faceted Search and Configurable Entity Mining[C]//Proceedings of the 1st International Workshop on Integrating IR technologies for Professional Search (ECIR 2013 Workshop), 2013.
- [7] 田宁. 读者服务视角下的资源发现系统可视化服务研究[J]. 图书馆学研究, 2014(17): 71-75.
- [8] BEN-YITZHAK O, GOLBANDI N, HAR'EL N, et al. Beyond Basic Faceted Search[C]//International Conference on Web Search and Web Data Mining, WSDM 2008, Palo Alto: ACM, 2008: 33-44.
- [9] 张福俊, 高雪, 周秀霞. 国内外数字资源发现平台比较研究——以Trove、DPLA、Europeana、WDL和文津搜索为例[J]. 国家图书馆学刊, 2018, 27(1): 86-96.
- [10] 崔明, 王振斌. 百度学术搜索和Summon学术发现的比较研究[J]. 图书馆杂志, 2018, 37(4): 74-80.
- [11] 许丽媛, 吴振新, 谢靖. 图书馆主流资源发现平台分析[J]. 图书馆学研究, 2015(8): 33-38.
- [12] 郑雯译. 国外发现系统分面过滤功能比较——EDS、Summon、

- Primo [J]. 大学图书情报学刊, 2016, 34 (1) : 26-30.
- [13] HALL C E. Facet-based library catalogs: A survey of the landscape [J]. Proceedings of the American Society for Information Science & Technology, 2012, 48 (1) : 1-8.
- [14] KHARLAMOV E, GIACOMELLI L, SHERKHONOV E, et al. SemFacet: Making Hard Faceted Search Easier [C] //ACM on Conference on Information and Knowledge Management. ACM, 2017: 2475-2478.
- [15] 胡昌平, 林鑫. 科技文献检索中基于主题词表分面化改造的分面构建 [J]. 情报学报, 2015, 34 (8) : 875-884.
- [16] KAJIYAMA T, SATOH S. A video navigation interface using multi-faceted search hierarchies [C] //ACM Multimedia Systems Conference. ACM, 2013: 136-140.

## 作者简介

林鑫, 男, 1987年生, 博士, 副教授, 研究方向: 数字信息资源管理与服务, E-mail: xinlin@mail.ccnu.edu.cn。  
吴茜, 女, 1995年生, 硕士研究生, 研究方向: 数字信息资源管理与服务。

### Comparative Study on Faceted Search Function of Literature Resource Discovery System

LIN Xin<sup>1,2</sup> WU Qian<sup>2</sup>

( 1. Institute of Scientific and Technical Information of China, Beijing 100038, China; 2. School of Information management of Central China Normal University, Wuhan 430079, China )

Abstract: Faceted search is an important service in literature resource discovery system, and the function design is the key factor that influencing user experience. In order to understand analyses the status of the function design of the literature resource discovery system, as well as the existing problems, this paper selects eight representative discovery systems as objects, and carries out the investigation and comparison from the following four aspects: the framework and focus setting of the faceted system, the sorting strategy of the facet and focus, the presentation strategy of the faceted system and the interaction mechanism. The results show that the design of the faceted search function of each system is quite different, and there are some common problems that need to be solved. To solve these problems, it is recommended to promote the standardization, systemization and semantization of the faceted system, combine fixed with dynamic ranking strategy to optimize the facets and focus sorting effect, implement the intelligent dynamic control of the faceted system and enhance the flexibility and comprehensibility of faceted interaction.

Keywords: Literature Resource Discovery; Resource Discovery System; Faceted Search; Function Design

(收稿日期: 2019-08-25)