

# 浅谈学术搜索系统用户体验优化

□ 包冬梅 / 南京政治学院上海分院军事信息管理系 上海 200433

**摘要:** 文章从用户体验入手,就学术搜索系统如何优化用户体验展开探讨。文章从学术信息资源集合的指引导航、检索结果集合的组织与显示、个性化推荐与服务、用户协作优化等方面展开论述。

**关键词:** 用户体验, 用户体验优化, 学术搜索系统, Web2.0

DOI: 10.3772/j.issn.1673-2286.2010.09.012

## 引言

搜索引擎作为网络信息组织最主要的方式,已然成为互联网中最有价值 and 影响的领域。但是,目前搜索引擎遭遇的最大挑战,是如何根据每个用户的细微差别,为其提供最妥帖的搜索内容。

在Web2.0主导的网络信息环境下,搜索用户信息获取和利用的模式发生了深刻的变化,用户信息存取的特点和要求决定了未来搜索的模式和搜索引擎的发展方向。搜索引擎未来发展必须与用户体验相融合,为用户提供社区化、个性化、智能化的智能搜索<sup>[1]</sup>。

作为搜索引擎的细分和延伸,垂直搜索引擎是为了克服通用搜索引擎搜索范围庞杂、查询噪声大、搜索深度不够等弊病而发展起来的一种搜索引擎服务模式。垂直搜索引擎通常是针对某一领域或行业的专业搜索引擎。垂直搜索引擎和通用搜索引擎的最大区别是对信息进行了结构化信息抽取,经过深度加工处理后返回给用户,因而搜索往往更加专注、具体和深入。为此,

笔者认为垂直搜索引擎在优化用户体验,提高搜索性能,实现智能搜索方面具有更大的提升空间。其中针对学术用户的学术搜索引擎(即学术搜索系统)的优化尤其值得探讨。本文是笔者在对各类学术搜索系统使用体验的基础上所作的粗浅总结,以期传统学术搜索系统平台在功能优化过程中,以及研究型图书馆在对所拥有的大量数字学术资源作二次整合开发时关注学术用户搜索体验的提升。

## 1 学术搜索系统的用户体验

### 1.1 Web用户体验

用户体验(User Experience)是一种在用户使用一个产品(服务)的过程中建立起来的纯主观的心理感受<sup>[2]</sup>。一个界定明确的用户群体,其用户体验的共性可以指导系统的人机界面和功能设计;反之,以用户为导向的系统架构和设计又能为用户带来良好的用户体验。用户体验主要是来自用户和人

机界面的交互过程。

用户体验的概念从系统开发的最早期就开始进入整个流程,并贯穿始终。其目的就是保证:(1)对用户体验有正确的预估;(2)认识用户的真实期望和目的;(3)保证功能核心同人机界面之间的协调工作。

根据Jesse James Garrett的观点<sup>[3]</sup>,Web用户体验的要素可以从面向任务和面向信息两个角度来解构,但两者在系统架构设计中又是互为融合的,如图1所示。

具体就搜索引擎而言,用户

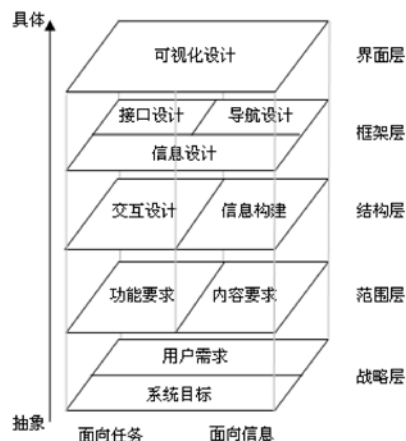


图1 Web用户体验要素

体验最重要的就是如何分析用户及其信息需求特征,确立系统目标,注重导航设计、交互设计、接口设计。

良好的搜索引擎用户体验,就是通过数据统计分析用户最经常发生的搜索行为(查询、点击、收藏、推荐),进而让用户具有经过最短的路径、花费最少的时间就能轻松方便地获取所需信息的搜索体验。

## 1.2 学术搜索系统用户体验的优化

学术搜索系统是面向学术用户,以学术信息资源为对象进行采集、组织(分析、索引)、提供搜索服务的一类搜索系统,包括学术搜索引擎、开放存取库、各大数据库出版商出版的学术数据库等,是学术用户从事学术科研活动的主要信息来源,是支撑学术科研工作不可或缺的工具。学术信息较一般的信息资源通常在内容上具有关联性、延续性特点,在结构上具备良好的格式规范。从事学术科研的科研院所的研究生、博士后、教授、专家学者等是这类系统的主要用户。

学术用户通常具有较高的信息素养,包括获取、利用、交流、共享的能力以及利用新技术的能力。该类用户群的信息需求大多具有主题明确、专业性强、理论层次高等特点,将权威性、时效性强的数字资源作为获取信息的首选<sup>[4]</sup>,并具有对特定主题不断跟踪和关注的特点。

根据学术资源的特点、学术用户的信息需求特征和信息利用行为特征,笔者认为学术搜索系统的用户体验优化可以从资源集合的浏览

导航、检索结果的组织与显示、个性化推荐与服务、用户协同优化等方面来进行。

## 2 浏览与导航优化

### 2.1 信息资源集合的多维导航

导航体系决定用户对信息资源集合的选择和搜寻策略。当信息集合覆盖多学科、多层次、多类型的巨量内容资源时,如何引导用户快速定位到目标子集合、找到所需要的信息内容就是导航组织所要解决的问题,也是体现信息集合可用性和易用性的一个重要方面。

信息集合常常被分解为多个维度,每一个维度就是信息集合的一个分面(或属性)。多维导航方式的选择必须综合考虑信息资源集合的特点、信息资源的构成、目标用户群等因素。一个优秀的学术搜索系统往往会对资源集合采用多维导航方式,让用户根据自身的搜索偏好和对搜索任务的解构任意选择自己的搜索路径起点,快速定位到最精确的目标子集合,展现出良好的搜索体验。

学术搜索系统可以根据学术用户的学科专业和 Information 需求层次、特点、喜好将信息集合进行分面分析,提供多维导航。

### 2.2 检索结果集合的后处理组织

如果说资源集合的导航组织主要用于限定和定位检索目标,那么对检索结果集合的聚类组织和排序显示则是对检索结果的进一步重组,有助于用户快速发现和定位到

准确的检索结果,减轻检索负担。对应答检索结果集的后处理,一方面有助于实现从多角度、多层次反映搜索结果与查询要求;另一方面改变了传统检索过程中不考虑用户检索偏好,缺少根据用户具体需求变化动态调整检索策略的弊病,进而提高了查准率和易用性<sup>[5]</sup>。

ISI Web of Knowledge 学术搜索系统根据学术用户搜索资源的普遍特点和需求将传统上线性、散列的检索结果多维化、计量化处理,通过对检索结果深入的细化分组(综合大类;主题范围;文档类型;作者等)和排序组织(按被引频次),为用户提供了从多个维度进一步精炼检索结果(Refine Results)的功能及对检索结果进行高级分析(Analyze Results)的功能,不仅迎合了学术用户多样化思维和利用的特点,而且实现了检索结果的快速筛选和准确定位,增强了用户搜索体验。CNKI也提供了类似的功能。

## 3 个性化用户体验优化搜索

### 3.1 建立 Alerts 跟踪提醒功能

学术搜索系统可以根据用户的使用行为,学习用户的学术兴趣和关注主题,为用户提供信息推送服务,在任何特定的时候(如用户保存的引文或期刊或检索式等发生了某种变化的时候)以各种方式将相关信息推送给用户。常用的 Alerts 方式有:以“某一检索式”为原点的跟踪提醒(类似于定题信息服务);基于引文系统的以“某一个检索结果(如某篇用户认为高度

相关或有用的学术文章”为原点的引文跟踪提醒；以“某种期刊”（我最关注的期刊）为原点的最新目次信息的推送。

目前，我国购买的大多数外文学术资源搜索平台都提供了较为完善的个性化提醒与推送服务，如 ISI Web of Knowledge、SAGE等。Alerts方式的专业针对性、指向性、发展性、及时性，契合了学术用户信息需求的特点，增强了学术用户系统、及时地获取学术信息的个性化体验。

### 3.2 概念查询扩展

基于概念的查询扩展可以通过两种途径来实现：一种是通过内置叙词表实现同义或相关扩展；一种是系统自动产生的推荐概念。用户通过使用查询扩展功能，选择、规范或扩展检索词，以修正和完善检索提问式。学术搜索系统的用户对象为学术研究用户，这类用户往往要求较高的查全率和查准率，而且学科专业性更为明显，因此搜索系统可以充分利用众多系统、规范的

专业叙词表辅助优化学术用户专业化的信息查询需求。

著名的理工专业学术数据库“INSPEC”在WOK平台的数据检索中提供了基于INSPEC Thesaurus的完善的受控搜索功能，对检索词从拼写、缩写歧义、专业词汇/同义词、标点符号表达等方面的差异通过词表进行规范控制<sup>[7]</sup>；医学搜索系统PubMed在检索式构造过程中调用MeSH词表，规范用户检索的规范性和学术性，为用户从学科专业层面提供了检索帮助，有助于提高检索效果。

Vivisimo Velocity企业级搜索平台则采用了词库词表和推荐概念的方式来提升用户的搜索体验（如图3所示）。学术搜索系统完全可以借鉴这一方式。



图2 ISI Web of Knowledge期刊最新目次的跟踪推送

## 4 用户体验协同优化

协同过滤 (Collaborative Filtering) 是个性化推荐系统中应用最为广泛的技术之一。常用的协同过滤技术可以分为：基于用户 (User-based) 的协同过滤；基于项目 (Item-based) 的协同过滤<sup>[9]</sup>。随着2.0理念的深入人心，2.0技术和应用的普及，“人件”（用户）这一要素已经成为系统架构的“中心”，协同优化过滤技术应用环境逐渐成熟。用户主动贡献信息已成为希望优化信息获取的必备条件。换句话说，用户的网络行为（贡献）越多，那么其所得的个性化信息收益越大。因此，基于用户智慧和参与的协同优化必然是未来搜索优化的方向。

学术搜索系统通过提供人与系统的2.0交互功能界面，一方面鼓励用户对检索结果的评价、添加

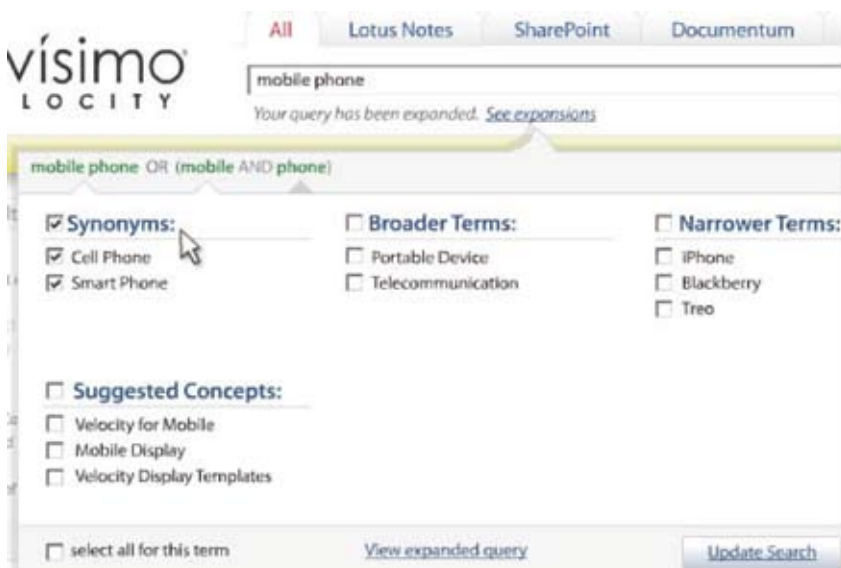


图3 Vivisimo Velocity概念搜索实例<sup>[8]</sup>

标签、评注、建立收藏等用户体验机制，为搜索打上信息价值的动态标签，最终实现搜索结果的协同优化；另一方面通过挖掘用户贡献的内容和用户行为，为用户建立学术网络，挖掘学术领域的关注热点等提高用户体验的功能，从而增强用户对搜索系统的粘性。

## 4.1 用户对检索结果的评价

搜索引擎对信息相关性的判断体现在检索结果列表的排序上，但由于各种商业因素的介入，搜索引擎自动评价排序检索结果的可信度一直饱受争议。

在Web2.0环境下，学术搜索系统可以借用掘客（Digg）<sup>①</sup>机制，相信学术用户的判断，把搜索结果的筛选权交给用户。从理论上讲（基于学术用户客观评价的假设），通过这种方式筛选出的检索结果可能才是最有价值的，最受用户关注的。

### 4.1.1 用户为检索结果投票

允许用户对搜索到的检索结果自主投票，推举有用结果，剔除不相关结果，改变检索结果的排序，过滤无关信息。通过用户的投票，后续的检索用户就会看到哪个或哪些结果被本领域的同行认为是最相关的。用户对检索结果相关性判断的自主性有助于克服搜索引擎按算法和其他盈利性因素决定检索结果排序的弊病，也有助于搜索引擎调整对检索结果的相关性评估。同时，在学术搜索结果中借用Wiki

机制，让学术用户推荐搜索结果之外的相关学术文档，不断完善检索结果。图4为Google（英文）检索结果用户投票推举（置顶、剔除和

推荐）示例。

系统还可以进一步为用户提供对相关结果的有用性评价机制（即对相关检索结果的有用性程度进行



图4 Google（英文）用户投票置顶、剔除和推荐检索结果示例<sup>[10]</sup>

评估)。通常可以将有用性分为多个等级，采用五星级评价方式，实现检索结果相关性的进一步量化。

### 4.1.2 用户添加标签

学术用户具有不同的知识结构和学科专业背景，但在搜索目的上具有很大的同质性——学术研究参考，这是保证Tag质量的前提。学术用户为学术资源添加的标签往往可以反映出研究者的认知和视角，因而Tag是从用户角度出发对学术资源进行组织的一种重要方式。Tag是对内容主题的提炼，用以帮助用户挖掘出其可能感兴趣的更多相关内容；Tag也是一种内容组织形式，Tag将多用户对不同主题核心内容的提炼精华，并按照一定的

规则有效地组织在一起，产生关联性；Tag也是一种检索方式，是同质用户群找到“组织”的最短路径。

具体采用的方式可以采用推荐关键词集，可以是一个预设词汇集，或预先定义词典的方式，供用户添加标签时参考。添加关键词标签有助于用户从该结果文档中辨析出最重要的概念，以便于其他检索用户能在检索结果排序和显示中看到这一信息，实现快速定位，节约搜索时间。

### 4.1.3 用户添加评注

系统允许用户为其认为有用的搜索结果添加评注，包括对某一检索结果的学术观点、创新性、学术参考价值等的评判，有助于其他用

<sup>①</sup> Digg是Web2.0的又一种具体表现形式。Digg的核心思想是发动大众进行新闻挖掘，Digg把新闻筛选的权利交给了网民，由网民民主投票来决定网站首页应该显示哪些新闻。Digg相信大众的眼光和评审，这样，筛选出的新闻往往是最有价值的，最受网民关注的新闻。

户通过阅读用户的评注快速了解检索结果的内容,为判断检索结果的相关度和有用性提供参考依据。此外,用户评注的另一功能是为具有共同学术关注点的用户之间提供一个相互交流讨论的平台。理论上,我们假设受评注相对较多的检索结果应当是受到较多用户关注的、相对较为重要的。因而通过评注的积聚,学术用户群体为检索结果添加评注最终必定有助于改进搜索系统检索结果的相关度和有用性排序,更契合学术用户群体的信息需要。

学术用户的专业背景和相对严肃的信息利用行为,有助于保证评注内容的有效性和可信度,从而保证学术搜索系统评注功能的合理、有效运行和检索结果的不断优化,最终实现用户体验的优化。

#### 4.1.4 用户收藏

用户收藏某一检索结果的行为包含着对该检索结果有用性的评价,因为非常有用才会收藏。搜索系统通过对用户收藏记录的分析可以判断用户的个性化信息需求,并将对某一结果的收藏用户和收藏次数推荐给兴趣相关的其他用户。

随着个人拥有数字化信息资源的不断增加,对搜索结果的有效序化组织和管理对个人学习、研究和工作效率的作用突显了出来。针对学术用户的信息搜寻行为的动机和目的,为用户提供对所收藏的检索结果的后续个性化管理是非常必要和重要的,即“Keep Found Things Found”(让找到的信息容易找到)。为此,学术搜索系统应为用户提供个性化的信息空间,提供在线文献管理工具(如Zotero<sup>[11]</sup>、CiteULike<sup>[12]</sup>、Connotea<sup>[13]</sup>、

NoteExpress<sup>[14]</sup>),保存有用的检索结果,允许对保存下来的网页和文本进行文本的高亮显示和插入批注,从而可以更好地辅助研究者进行参考文献的阅读和思考,也有利于提升搜索系统的用户粘度。

以上用户对检索结果的评价必须与用户个人在学术搜索系统中的帐号关联(或绑定),这样才能在优化个性化搜索体验的同时优化其他用户的搜索体验。

### 4.2 用户信息使用行为的挖掘

学术领域内存在着作者与作者、作者与机构、机构与机构、作者与主题间的学术关系。学术资源之间通过资源的元数据属性(主题词、作者、来源机构等)及学术引证关系的描述和揭示,可以建立起学术资源之间及学术对象(包括人和组织)之间的关联网络,实现特定学术领域内典型学术关系网络结构和演化过程的可视化分析,发现机构、科研团队和个人之间学术关系的特点和趋势。这些功能在引文学术搜索系统,如ISI Web of Knowledge、CiteSeer、CNKI中已经得到较为极致的体现。

从另一个层面上,通过对上述4.1用户贡献的内容和用户行为的挖掘,学术搜索系统可以计算、汇总并分析用户的群体行为结果,如用户关联、热门领域主题等,也可实现为用户建立学术关系网络,揭示和预测学术热点等功能,进一步提高用户体验的功能。

#### 4.2.1 挖掘学术关系网络

学术关系网包括学术资源之间

的资源关联网和学术用户之间的人际关联网。

学术搜索系统充分发掘标签独特的“聚合”、“关联”能力,通过共同标签的创建,共同学术资源的关注,通过资源之间的聚合关联,建立起学术资源创建者与学术用户之间、学术用户与学术用户的人际关联,并进而在资源利用层面上建立起学术资源之间动态的隐性关联网。Springer旗下的CiteULike<sup>[13]</sup>社会化学术标签网站(不仅仅提供在线文献管理功能),根据用户对某一资源标注的“标签”,为其推荐对同一资源标注过“标签”的其他用户及其所关注的相关资源,构建起用户之间的学术关联网,协助用户实现分享学术资源(进一步的学术资源关联)和开展广泛的学术交流;中文360doc<sup>[16]</sup>同样以标注、评论、收藏等用户的贡献或行为,经由资源建立起用户之间的关联网络。

学术搜索系统还可以通过从多个来源和文档中无缝挖掘、抽取个人学术数据信息(如专业研究领域方向、联系信息和最近使用的标签和共享活动等),并在搜索过程中将信息发送到关联主题的搜索结果中,让用户容易地在查找的过程中发现某主题或领域的专家,进一步丰富学术用户人际关系网。如读秀学术搜索的根据搜索主题提供“相关人物”、ISI Web of Knowledge提供的“与科学家自传信息的链接”(Author Biography)可帮助用户获取某学科领域的核心研究人员的基本信息及其个人站点、博客、研究成果等资源。

#### 4.2.2 发现和揭示学术前沿、热点

学术搜索系统基于标签的创建, 可以为用户提供图形化的指示图, 揭示学科专业领域近期最被关心的文档(可以从用户个人、用户群等层面组织), 从而让用户对学科领域的重点和方向有一个概要的了解, 及时把握研究前沿热点。

协同优化让用户记录下信息检索经验的同时, 开放自己的经验与他人分享, 系统通过对用户按兴趣模式聚类来增强用户间的协同与协作。发现需求相近的用户、借鉴彼此的检索经验、分享对方判别检验确定后的历史检索结果, 从而实现创建一个人与人之间因兴趣而相互协作的系统。

## 5 结语

学术搜索系统用户体验的优

化是一个整体系统的优化, 包括系统结构和功能组合的预先设计和安排。在Web2.0环境下, 学术搜索系统尤其要引进2.0理念与技术, 在系统架构中要充分考虑“人件”因素——搜索用户, 在资源组织阶段提供尽可能多的资源访问入口和路径; 在使用阶段增强用户的自主性, 挖掘用户的智慧, 调动其参与的积极性。根据用户的群体行为综合调整优化整个系统的功能组织结构, 并可根据用户的行为和自主发展的协作关系网络提供个性化信息搜索服务, 提供更为个性化、智能化的“专、精、深”学术信息服务, 增强用户搜索和体验的优化。当然, 对用户的评价与参与行为必须辅以正确的教育和引导, 才能真正发挥学术搜索系统利用Web2.0理

念和技术优化用户体验的效果。

学术搜索系统的用户体验不仅局限于单体自身功能的优化, 而且还要考虑到与构成整个学术资源生态环境中的其他学术资源系统之间的整合与关联。基于现有的学术资源网络生态环境和技术条件, 学术搜索系统应支持开放链接协议(如OpenURL协议), 与整个学术资源环境(大系统)整合关联, 透过简单的无缝链接的查询界面, 让查找资源的信息来源变得更为多元, 查找结果更为丰富。

当然本文的不足之处是对学术用户的信息获取、利用、交流、共享行为研究还不够, 没有实证分析, 这也正是本课题下一步努力和研究的内容。

### 参考文献

- [1] 王学松. Lucene+nutch搜索引擎开发[M]. 北京:人民邮电出版社,2008.
- [2] Sonic. 什么是用户体验[EB/OL]. [2010-04-02]. <http://www.burgeoncn.com/sonic/weblog/article.asp?id=20>.
- [3] GARRETT J J. The Elements of User Experience [EB/OL]. [2010-04-02]. <http://www.jjg.net/elements/-pdf/elements.pdf>.
- [4] 刘秀华. 学术用户的数字信息需求研究[J]. 图书情报工作,2008(9).
- [5] 包冬梅. 减缓网络用户检索负担的有效方法——可视化大规模应答结果集[C]// 戴维民,等. 网络环境下的信息组织——全国第四次情报检索语言发展方向研讨会论文集. 北京:北京图书馆出版社.
- [6] 李晓明,等. 搜索引擎——原理、技术与系统[M]. 北京:科学出版社,2005.
- [7] INSPEC [EB/OL]. [2010-04-02]. <http://isiknowledge.com/inspec>.
- [8] Vivisimo [EB/OL]. [2010-04-02]. <http://vivisimo.com/docs/overview>.
- [9] 白丽君,等. 协作过滤研究概述[J]. 电脑开发与应用,2002,15(11).
- [10] Google [OL]. [2010-04-02]. <http://www.google.com>.
- [11] Zotero [OL]. [2010-04-02]. <http://www.zotero.org>.
- [12] Connotea [OL]. [2010-04-02]. <http://www.connotea.org>.
- [13] Citeulike [OL]. [2010-04-02]. <http://www.citeulike.org>.
- [14] NoteExpress [OL]. [2010-04-02]. [http://www.reflib.org/index\\_chs.htm](http://www.reflib.org/index_chs.htm).
- [15] 360doc [OL]. [2010-04-02]. <http://www.360doc.com>.

### 作者简介

包冬梅 (1976-), 南京政治学院上海分院军事信息管理学系讲师, 研究方向: 信息组织与检索。通讯地址: 上海市四平路2575号 200433。E-mail: iambadongmei@163.com

### User Experience Optimization of Academic Search Systems

Bao Dongmei / Department of Information Management of Shanghai Branch, Nanjing Political Academy, Shanghai, 200433

Abstract: The article discusses how to optimize the user experience in academic search Systems. From the author's point of view, the optimization can be carried out from the following aspects: navigating academic database; organizing and displaying search results; personal recommendation and service; and user collaboration.

Keywords: User experience, User experience optimization, Academic search system, Web2.0

(收稿日期: 2010-04-16)