

基于学科联盟热点推荐平台的热点评价与信息推送研究*

徐 敏 杨应全 陈祖琴

重庆科技学院图书馆 重庆 401331

[摘要] 介绍学科联盟的概念和形成方式、热点推荐平台的概念,提出发挥学科联盟的优势为热点推荐平台的建立作出贡献的建议。详细论述基于学科联盟的热点推荐平台中的热点评价和信息推送过程,认为这两个过程的好坏直接影响到用户对平台的认可度。

[关键词] 学科联盟 热点评价 信息推送

[分类号] G203

Research of Hotspot Evaluation and Information Push in Recommended Disciplines Focus Platform Based on Discipline Alliance

Xu Min Yang Yingquan Chen Zuqin

Library of Chongqing University of Science & Technology, Chongqing 401331

[Abstract] This paper introduces the concept and process of discipline alliance, the concept of recommended discipline focus platform and proposes that it can contribute the advantages of discipline alliance to the building of the platform. Then the authors elaborate the processes of hotspot evaluation and information push of the platform, whose qualities directly influence the user's recognition to the platform.

[Keywords] discipline alliance hotspot evaluation information push

1 引言

在科学技术飞速发展的过程中,“信息爆炸”不仅对我们日常生活中信息的有效获取产生困扰,也对在科研工作中的许多研究者造成了“学术信息爆炸”的困扰,学术研究的大范围交叉以及研究深度的提升使得越来越多的科研工作者不能及时有效地了解研究领域的发展情况,许多科研人员不知道如何搜集和学习本研究领域的知识,从而形成了“学术信息障碍”,同时由于学术信息没有做到有效地共建共享,使得实力较强的科研单位无法帮助实力较弱的科研单位,形成了“学术鸿沟”。因此本文作为中国冶金教育学会教育科学研究计划课题“学科发展热点推荐平台的实施模式研究”的成果之一,主要研究一个基于学科联盟的热点推荐平台的热点评价和信息推送的过程,这是因为热点评价的效果和信息推送的过程直接影响用户的使用体验度。

2 基于学科联盟的热点推荐平台

2.1 学科联盟

中华人民共和国教育部截止 2010 年 12 月 30 日的统计数据表明全国高等教育学校(机构)中研究生培养机构有 796 所,普通高校有 2 305 所,成人高等学校有 384 所,民办的其他高等教育机构有 812 所^[1]。普通高等学校设置的专业达到了 1 759 门(其中研究生教育的专业有 623 种,普通高等教育的专业有 646 种,成人高等教育的专业有 490 种)^[2]。从以上数据可以看出中国高等教育的普及面是非常广的,总计 4 297 个高等教育的学校机构涉及了 1 759 门学科的教学研究工作。但是拥有研究生培养机构的仅有 796 所,说明虽然高等教育的广度很大,但是真正有研究实力的学校机构只占了总比例中的少数,这种情况加深了“学术鸿沟”。而“学科联盟”这个概念的提出可以很好地解决“学术信息爆炸”和“学术鸿沟”的问题,它是指根

* 本文系中国冶金教育学会教育科学研究计划课题“学科发展热点推荐平台的实施模式研究”(项目编号:YGT09030)研究成果之一。

据国内专业以及研究方向的设置情况进行细分,以每个专业或研究方向的强实力科研单位为领头,联合本专业或研究方向的其他科研单位形成本专业或研究方向的联盟,再和与本专业或研究方向平行学科的联盟一起组成上一级的学科联盟,直至形成符合学科体系的学科联盟体。以冶金学科为例(见图 1),实心五角星标注的北京科技大学冶金与生态工程学院、昆明理工大学冶金与能源工程学院、中南大学粉末冶金研究所分别作为钢铁冶金领域、真空冶金领域和粉末冶金领域的领头与本研究领域的其他科研单位组成不同的联盟,再与其他冶金研究领域的联盟组成上一级的冶金学科联盟,然后与其他学科或研究方向的联盟最终构成了学科联盟体。

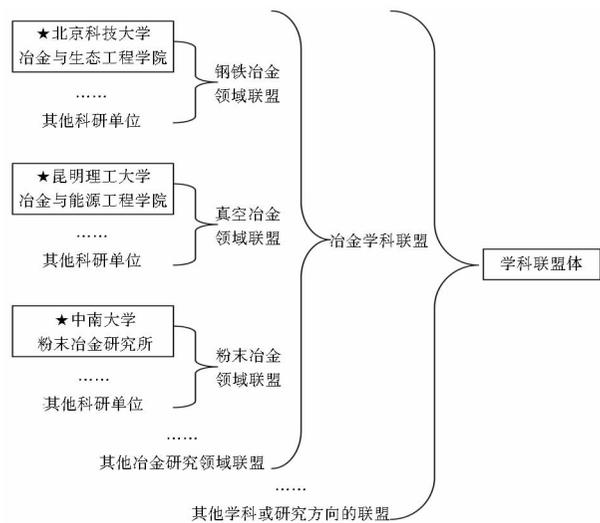


图 1 学科联盟体构成

2.2 平台概念

学科发展热点推荐平台是指针对不同学科不同研究方向形成不同的栏目,栏目的基本设置与分级参照教育部的专业设置,用户可以通过访问平台进入不同的栏目寻找或定制热点信息或提交学术问题。学科联盟和热点平台的结合有以下几点好处:

第一,两者的结合让学科联盟体的建立有了学术资源共建共享支持,每个学科下的不同研究方向的强实力研究单位可以针对本研究领域的特点为栏目板块的建立提出有特色的意见;平台因为有了学科联盟的支持,其学科栏目内容和资源的数量和质量都得到了保证。

第二,两者的结合让学科联盟体可以借助平台由强实力科研单位向弱实力科研单位进行学术传导,消除由于地域差异、资源差异和学术马太效应等因素导

致的学术鸿沟。

第三,两者的结合让学科联盟体中的成员利用自身对学科了解的优势制定出符合学科研究特点的热点评价体系,以期使平台生成的热点更加客观科学更符合学科,研究人员需求。

第四,两者的结合让平台的用户不仅可以浏览定制学科研究领域的热点信息,还可以在线提交科研工作中的疑惑,由学科联盟体的专家解答,此外系统还会根据用户注册时的学术背景信息以及其在平台上的浏览历史记录找出用户现阶段所属的研究领域供学科联盟体专家参考。

3 热点评价

热点评价是本平台的主要特色,即面向不同学科研究领域的科研人员,定期评价出不同学科研究领域的前沿和热点,科研用户可以通过访问相关学科领域的栏目了解学科研究的发展方向,为其科研工作的创新提供帮助。热点评价过程中最关键的是评价指标、评价权重和评价词提取这三部分。首先平台的爬虫程序会在所有网络学术资源库的特定页面中根据评价指标采集特定信息至平台的指标信息索引库中,然后根据评价指标体系中的指标权重对数据进行处理运算得出每篇论文的热度,再利用自动切词技术对热度最高的一些论文的标题、关键词和摘要等信息进行分词并计算词频,得出专业的热点词汇,这些热点词汇和热点论文会根据学科研究方向分布至不同栏目,对于一些交叉型或复合型研究的学术成果,会在其涉及的多个学科中都会体现,但是在热点评价中会因为评价指标的不同或研究的新颖程度排列顺序有所不同。

3.1 评价指标

评价指标是指对学术研究成果的若干个单项指标所反映的学术内容进行评价,通过搜集整理分析这些指标并汇总结果,提取出不同学科研究领域的热点信息。评价指标主要分为两大部分,第一部分为期刊的评价指标,包括了期刊的级别、主办单位、被引次数、影响因子、立即指数、发文量、被引半衰期、平均引文率、期刊他引率等;第二部分为学术论文的评价指标,包括了论文的发表者、发表者单位、所属期刊、被引频次、他引率、学术网站下载数等。以上这些是热点评价中的基础指标,而学科联盟体中各个专业研究领域的专家除了可以根据本领域的研究特点调整以上这些基础指标外,还可以增加有专业研究特色的新指标,从而形成

符合专业特点的特色评价指标。

以上评价指标中的大部分都是数据型指标,来源于平台中的爬虫程序对网络上的学术资源库网站的相关页面的信息抓取,这些数据型指标可以很容易加入到热点评价的运算过程中,但是平台在此还采集了类似期刊的级别、期刊的主办单位、论文的发表者、发表者单位等文字型指标,之所以这么做是考虑到首先某研究领域的发展方向和前沿热点的学术成果往往最先最集中发表在本学科期刊级别中的核心期刊和权威期刊上;其次学科强实力单位主办的期刊中反映的也大多数是本学科的最新研究成果;再次,每个领域专家的学术论文也会体现其最新研究成果,所以一篇论文的发表者如果是某研究方向的著名专家,其论文的参考价值将是非常大的;最后,一个领域的著名专家背后都有一个强大的学术团队,成员里可能包括了其学生、同事或其他科研搭档,而学术团队的体现主要是通过论文中的发表者单位来实现,即一个强实力学科的单位或著名专家所在单位的作者们的学术研究往往具有很强的连续性、关联性和新颖性,所以也很具有参考价值。通过这些分析发现文字型指标也同样具有很高的评价作用,所以可以将文字型指标进行量化,为其赋上数值型数据加入到评价运算的过程中,例如可以针对SCI(SSCI)收录、EI收录、ISTP收录、权威核心、CSCSI收录、一般核心、一般期刊等级别给予不同的值加入到评价过程中。

3.2 评价权重

评价指标是热点评价的基础,没有它就无法衡量研究成果的重要性,但是仅靠评价指标是无法形成科学客观的评价过程,所以对于评价指标中的每一个元素需要赋予不同的权重值,而学科联盟体中的专家正是可以针对本研究领域的特征调整指标的权重值。举个例子,冶金工业学科中同为核心期刊的《粉末冶金技术》和《稀有金属与硬质合金》在不同研究领域中被专家所赋予的权重值就是不同的。在粉末冶金的研究方向下,《粉末冶金技术》拥有很高的权重值,而《稀有金属与硬质合金》则在金属冶炼和强度测试的研究方向下拥有很高的权重值。这种同属于一个大专业的同期刊级别的期刊在不同研究方向下权重值不同的现象正是来源于学科联盟体中不同领域的专家为了体现本研究领域特点而设置的特色评价权重体系。

3.3 评价词提取

平台推荐的热点主要是两部分:第一部分是爬虫程序抓取网络上的学术资源库的论文页面信息,经过

评价过程得出不同学科研究领域热度最高的一些论文,供用户浏览论文概况并可以通过超链接跳至相应的学术资源库网站的论文页面付费下载,这里的评价词提取指的就是论文的标题、作者、单位、摘要等不涉及全文的关键信息,这个评价词提取过程的前提是已经对抓取的论文的指标进行了热度评价,自动挑选出热度最高的一些论文;第二部分是建立在第一部分基础上,对其进行更深一步的数据挖掘,即除了推荐给用户高热度的文章之外,还采用搜索引擎中的自动切词技术将采集的文章的标题、摘要进行分词统计,然后计算标题、摘要、关键词中实词的绝对词频和相对词频,再与该学科领域内统计文章中该实词的总词频进行比较就可以得出较为客观的热点词汇了。例如用户进入冶金工业学科的粉末冶金研究方向的栏目中,可以一目了然地看到按热度高低排序的文章列表,还有就是粉末冶金研究方向的最新热点词汇,用户在进行科研工作时就多关注这些词汇所代表的知识点。

4 信息推送

热点评价的重要性体现在可以通过提取学科研究热点来辅助用户的论文撰写、课题申报等科研工作,而信息推送则能够提高用户体验,增加用户的忠诚度和依赖度。

4.1 推送方式

4.1.1 用户定制信息 平台热点信息的推送主要分为主动推送和被动推送,其中用户通过定制信息获得热点推送的方式为被动推送,即满足用户需求的目标性推送,而根据用户浏览记录推送热点信息和根据用户背景信息推送热点信息的方式为主动推送。平台用户分为访客用户和注册用户,访客用户指未在本平台注册的用户,他们在浏览页面信息的同时可以通过站外电子邮件的形式定制所需的栏目信息,他们是潜在的注册用户,在有效及时的信息推送过程中,他们会有一部分在平台上进行注册转化为注册用户并邀请其他人员加入注册用户的队伍,而注册用户除了可以在平台系统中利用站内邮件、外部电子邮件、RSS和客户端等形式定制自己需要的栏目信息之外,还拥有自己的个性化面板,利用 Ajax 技术可以很容易实现用户在登入的主页中定制不同栏目,调整其顺序,改变其外观布局等功能。

4.1.2 根据用户浏览记录推送 用户在平台上访问和定制信息的过程是被记录下来的,通过定期对浏览

记录的分析可以发现用户不同阶段的兴趣点,从而达到在用户自己定制信息之外主动推送其可能关注的栏目的目的,再通过对用户访问主动推送栏目的频率进行评估,动态调整推送策略,提高用户使用体验。

4.1.3 根据用户背景信息推送 平台的服务对象主要是科研用户,他们在平台注册时会被建议填写自己的学术背景信息,包括了用户教育经历(毕业院校、专业和研究方向)、科研成果(参与项目、项目级别、发表论文等信息)、指导老师、工作单位、所在岗位、感兴趣的研究方向等,这些信息对于系统挖掘用户的学术趋向是非常重要的,通过以上信息系统能够挖掘出与用户关联性较强的一些潜在学科热点方向,帮助用户从事交叉型或复合型学科的研究工作。

4.1.4 推送信息的评价 信息的推送是为了满足用户的科研需求,用户在不同阶段进行科研的方向是不同的,用户的科研兴趣也是随着时间的推移在变化的,所以推送的信息不能够一成不变,系统会通过用户对用户的操作记录进行定期评估,查看某些推荐栏目是不是访问量过少或长时间不访问,为用户作出提示,让用户自主决定是否修改自己的定制策略。此外用户还可以对于推送的信息的质量和有效性作出自己的评价,即在相应的热点论文和热点词汇上进行投票并作出文字性的评价,系统会考虑用户的投票结果动态调整热点信息的排序和内容,而对热点信息的文字性评价也会为其他用户的科研提供参考,可以让其他用户了解热点信息的支持度、内容性问题等方面。

4.2 推送形式

4.2.1 RSS RSS 是英文 Really Simple Syndication 的简写,被称为聚合内容,是在线共享内容并快速获取时效性比较强的信息的一种简易方式。它的特点是个性化聚合多来源信息、信息发布时效性高成本低、信息针对性强无不相关信息,热点推荐平台的 RSS 输出能够让用户迅速了解栏目内容的最新更新。热点推荐平台的 RSS 阅读器主要分成二种,第一种是运行在计算机桌面上的应用程序,通过平台上所订阅栏目的信息发送,能自动、定时地更新热点,现成的软件有 Awasu、FeedDemon 和 RSSReader 等;第二种是在线的 Online RSS,用户不需要安装任何软件,并且可以保存阅读状

态,推荐和收藏常关注的热点。

4.2.2 电子邮件 利用电子邮件推送信息是现在 Web 资源类和门户类网站应用最普遍也是最有效的方式之一,因为该方式推送热点不论对象是注册用户还是访客用户只需要在定制热点信息时制定投递的电子邮箱即可。

4.2.3 站内邮件 站内邮件是针对注册用户,在其每次登陆平台之后,友好提示其关注的栏目发生了哪些信息变更,并能通过站内邮件中的超链接一键访问平台的相关栏目。

4.2.4 短信平台 在手机非常普及的今天,越来越多的网站架设了手机短信平台,通过短信形式传递给用户信息,而在图书馆领域,关于手机图书馆的研究论文已经很多了,有的图书馆也开始尝试利用手机图书馆的形式,提醒读者图书的借阅归还以及新书推荐等,平台的信息推送也可以通过定期向注册用户发送其关注的栏目的热点更新信息,方便科研人员随时跟踪学科进展。

5 结 语

学术鸿沟和学术信息爆炸给许多科研工作者造成了很大的困难,联合学科专业形成学科联盟体能够让专业内部的科研单位联系更加紧密,学科发展热点推荐平台的设想就是为学科联盟提供学术资源共建共享的平台,充分发挥学科联盟的合作优势,以热点评价和热点推送的形式帮助广大科研人员进行学术研究,平台中的热点评价和信息推送的重要性不言而喻,本文主要从学科联盟热点平台的热点评价和信息推送的实施方法上进行了论述,希望可以为更进一步的研究起到较好的参考作用。

参考文献:

- [1] 中华人民共和国教育部高等学校(机构)数. [2010-01-11]. <http://www.moe.edu.cn/publicfiles/business/htmlfiles/moe/s4960/201012/113595.html>.
- [2] 中华人民共和国教育部普通高等学校设置专业数. [2010-01-11]. <http://www.moe.edu.cn/publicfiles/business/htmlfiles/moe/s4960/201012/113592.html>.

【作者简介】徐 敏,男,1982 年生,助理馆员,发表论文数篇。

杨应全,男,1968 年生,研究馆员,发表论文 40 余篇。

陈祖琴,女,1981 年生,馆员,发表论文数篇。