

互联网时代科技文献评价影响因素探讨

□ 孙卫 / 中国科学技术信息研究所 北京 100038
/ 科学出版社 北京 100717

摘要: 对于科技内容的评价方式, 仅从期刊文章被引的角度进行评判, 在互联网时代已经显得不足了。互联网时代发表科技观点的多样性、传播的广域性、传播的快速性使得科技内容被利用的因素有很多。使用者对文章观点的态度、作者和专家公开发表的科技内容的情况、使用者利用互联网进行发现与利用的情况等, 都可能对科技内容进行评价分析的因素。

关键词: 引文分析, 作者分析, 机构分析, 检索结果分析

DOI: 10.3772/j.issn.1673—2286.2012.06.005

概述

自从1995年美国试验型互联网逐步转变成公众互联网以来, 互联网信息服务成为了在互联网上最成功的应用之一。以美国谷歌(信息检索能力)、亚马逊公司(单种、册、件图书服务能力)等为代表的互联网服务(Internet Services)企业跻身进入世界500强企业, 而传统出版企业没有进入, 就是最好的证明^[1]之一。

数字图书馆伴随着互联网发展而快速发展。传统的图书馆的序化管理已经成功地被用于数字图书馆管理; 数字图书馆的种、册、件、篇、章、节、段的组织和服务能力远超过了传统图书馆以种、册、件为核心的组织和服务能力; 数字图书馆基于知识细分、工具细分组织和服务^[2,3]远远地超过了传统图书馆靠资深馆员的经验和方法进行组织和服务的范畴; 数字图书馆公有资源^[4,5](公开发表的学术论文、期刊、书籍、会议报告、专利等)正在从单个图书馆的镜像拥有变成了远程服务的购买。在过去的17年, 以互联网为基础的数字内容信息服务业极大地动摇了传统图书馆的定位和地位^[6]。

科技文献的传播与利用是科技发展中最关键的催化剂之一。而在互联网时代, 如何看待学术水平、如何评价科技文献的水平等成为科技文献评价分析理论需要进一步发展一个方向。在传统图书馆中, 学术内容主要是科技文献以种、册、件等物理实体作为传播与利用的方式方法, 所以文献的半衰期^[7]相对比较长, 利用科学引文利用和统计的方法是一个相对准确和被普遍认可的一种科技文献评价方法。传统的文献老化理论(半衰期)是上个世纪70年代建立的, 至今几乎没有出现互联网环境下对科技文献半衰期统计进行更深入的研究和形成公认的成果的报道。而在互联网时代, 数字类型的科技文献传播的效率和广域性远远超过了传统科技文献以物理实体为传播对象的效率和区域, 使得科技文献的半衰期不能简单地用单一因素来衡量^[8]。在互联网时代, 国际水平的专业会议的论文、发表人等信息完全公开了, 不同于传统的国际会议要等会议论文集出版到达图书馆后才知道相关信息。互联网时代科技工作者研究所需要的信息环境也发生了重要的变化^[9], 发表科技观点的方式更多, 实效性和针对性更强, 例如博客、个人网站、专业社区等。基于纳税人的国

家基金和很多民间基金用于科技研发项目所强调的信息公开的制度,使得很多项目的基本情况在互联网^[10,11]上也可以获得了。

传统文献计量理论主要是基于“权威”或者“经典”的文献资料来佐证科技发展的。可是由于纸媒载体形式和图书馆借阅传播方法的限制,很多科技发展初期的科技内容由于科研工作者不出名,而在初期发表时被这些“权威”或者“经典”的纸媒载体的编辑或者同行评议的专家所谢绝了,更谈不上进入图书馆借阅传播了。但是,愿意发表自己研究成果、方法、学术观点的科技工作者现在可以选择通过传统纸媒和互联网网媒进行更快的、更多样性的发表与传播了。什么是“权威”,一个研究成果被肯定,被广泛采用,创造这个成果的机构、个人就成为了“权威”。可是,对于一个新研究者研究的成果更多是下一次创造的机会,而这些机会在初期往往是得不到业界承认的。所以,所谓的“权威”仅限某一个专业下的某一个主题方向下的研究成果在一个时间阶段的代表者(机构、个人)。而“经典”表现在学术上的争论结束,已经得到认可,变成相对稳定的方法、技术、理论,成为学术研究基础的才能称之为“经典”。由此可见推动科技进步的不是依靠权威和经典,而是更多的科技研究者通过科学的方法或者大脑的想象并经过科学技术的应用进行证实和成为学术成果的新创造。美国全民科学素养主要强调几点^[12]:①建立科学的世界观。世界是可以理解的、科学想法是可以改变的、科学知识是经得起时间考验的、科学并非能提供完整的解答等。②掌握科学探究的方法。科学需要证据、科学是逻辑和想象的混合体、科学的价值是解释和预测能力、科学趋势的确认而避免偏误、科学不是权威。③了解与参与科学事业。科学是复杂的社会活动、科学被分成许多学科内容并且在不同的组织机构中进行研究与探索、在科学研究与探索中要遵循公共的伦理规则、科学工作者作为专家和公民要参与公共事务。所以,以科学素养的科学世界观看,经典是存在的(科学知识是经得起时间考验的),而以掌握科学探究的方法看,绝对的权威是不存在的(科学不是权威)。

所以,可以看到的事实是,传统文献的评价分析、多样性科技信息、科技文献被使用的状态信息的获取、人物信息的获取、项目信息的获取等,为如何多角度地进行科技文献评价分析具备了数据基础,探讨用什么信息和怎么用这些信息是数据建立的基础。

1 传统的科学引文索引统计分析——科技文献评价体系

科学引文索引(SCI)是由美国科学情报研究所(ISI)在上个世纪五十年代开始收集整理各种学术期刊的文章,研究与学科专业的相关性,在1963年正式出版SCI,逐步从年刊转变成季刊,1979年改为双月刊。1976年发布《期刊引用报告》,1997年,汤姆森整合了SCI、SSCI(社会科学引文索引)、AHCI(人文艺术引文索引)后正式推出了科学互联网(Web of Science)信息服务,变成每周更新^[13]。

该方法的优点就是利用已经发表的期刊文章进行累计和统计,通过篇名、作者、机构、地址、主题、摘要等构成了SCI,再利用每个期刊的文章总数作基础,被引过的文章做参考,利用统计学的角度计算形成了期刊引用报告(JCR)。由于这个方法是上个世纪50年代研究的,所能面临的公开的文献主要是书籍和期刊,这个SCI和JCR主要是针对期刊这类连续出版物的,所以积累统计相对比较稳定。

今天大量使用互联网传播和利用以后,这个方法出现了一些缺点。①时效性不好。早期的SCI是从年、季、双月逐步改变更新周期,是为了更好解决时效性问题。1997年汤姆森采用Web of Science以后,更新周期变成了每周,有了更好的时效性。而我们国家的四个相似的科学引文索引(北大、南大、中信所、科学院图书馆)都还没有做到真正的实时更新(事实上中国知网、万方数据提供的期刊服务是具有更好的时效性的),所以时效性问题很突出。②科技期刊收集不全,造成的公信力受到质疑。爱思维尔的Scopus具有来自5000个出版机构的18000种期刊文章^[14],汤姆森路透的Web of Science收录了10000种期刊的文章^[15],就存在收集学术期刊种类严重不同的问题。那么,图书馆情报专业追求的“全”和“准”的理念仅靠爱思维尔或者汤姆森的二次文献库是否也存在偏颇呢?③对于期刊引证报告(JCR)是否代表学术水平引发争论。JCR仅仅是一个角度来表征不同的期刊在一个相同指标体系下的对比情况。长期以来,重要的科技工作者在这些对比之下排名靠前的期刊发表文章,所以很多情况下期刊的权威性和学术的权威性是正相关的,这个是能被佐证一些已经发展起来的科技情况的,不能完全佐证科技上升趋势和苗头。

在互联网信息服务时代,研究的多样性、新的研究

者的非出名因素、可能代表着未来科技发展的趋势等更多的是利用网络环境来发表与传播,而非物理图书馆场所或者服务提供了。那么,研究从发现到形成完整的研究成果之间的这个上升期间,如何对于科技文献进行评判呢?

如果要利用科技文献对于作者的科研工作进行评价,需要关注几个因素:

(1) 作者引用别人的文章,或者别人引用作者的文章,不是简单地看文章名(传统是只看文章名和出处),而是需要从语义的角度去判断。文章的主要观点,引用者是完全同意(统计意义上完全同意的多,作者的学术观点得到肯定)、引用者不同意(统计意义上不同意的多,作者的学术观点引发了争鸣,学术观点分裂趋势的表现)、引用者不完全同意提出修正(统计意义上代表对学术观点的认可,但是要完善)。由此可见语义层面的分析才真正具备学术价值判断的意义。

(2) 解决时效性测量问题。如果不量度一篇文章从发表到被引用最短需要多少时间,那么很容易忽略传播和利用的效率因素。目前,主要学术期刊纸刊出版3-6个月后,互联网才提供服务。如果二者同步,这个量度将有助于判断学术观点的趋势和引领作用。

(3) 解决学术观点出现,引用率在时间周期上的分布关系。当一个学术观点出现,从没有被关注到被关注这个阶段,属于学术观点研究的上升期(趋势);而当这个学术观点在一个阶段被广泛关注,是学术观点研究的高峰期(热点);一个学术观点在出现高峰后的关注逐步降低,预示学术观点研究的分裂开始,或者学术研究即将结束。

(4) 解决互联网利用科技文献过程留下的事实数据测量问题。在互联网科技文献利用中,检索、点击、阅览、下载是有事实记录的,这个也和传统科技文献利用方式完全不同。

例如:用“数字图书馆”、“信息服务”在万方数据库中检索,见图1检索结果。

在图1中,符合检索条件的有1328篇文章(说明这些文章中出现过这两个关键词)与检索条件相关。在图1的左上部分是这1328篇文章按照中图分类法大类目分布的情况,左面的中部是按照年代分布的情况,左下部是按照刊分布的情况。右中部有题名、期刊名、卷(期)、哪一种核心期刊评价体系、作者、摘要等信息。检索者可以根据摘要来判断是否需要翻下一页检索结果浏览、阅读全文、下载全文等。



图1 检索结果

而检索、检索结果翻页、全文阅读、全文下载等的使用过程信息都会遗留在服务器端。而这些信息记录了使用者对于主题(关键词)、判断主题和摘要(检索翻页浏览)、主题内容的进一步探知(全文阅读)、收藏或者离线阅读价值(全文下载)等过程。

在技术角度上既然已经遗留了很多科技内容被利用过程的痕迹,而这些痕迹也是真正被利用过程的事实数据之一,再加上主流的科技文献的使用方式上大都采用远端集中服务的方式(不再本地的镜像内容服务器,而是购买集中的科技内容网络服务,直接使用信息服务商的网络化服务),那么分析服务器端的检索用关键词(主题、学术方向),检索命中(元数据与关键词的相关性,与文章的相关性),命中后的翻页快速阅读(摘要的阅读)、链接的点击与阅读(在线阅读)、下载到本地的保存再利用等数据是极有价值的。这个参照系的建立,对于科技文献评价分析将会增加从网络使用的角度进行分析(科学引文索引更多是从写的角度进行分析)的可能性,而不只是从写作的角度去评判了。

例如:对于第一作者华东师范大学范并思教授,在中国知网(CNKI)上用作者名检索,得到图2(范并思教授文章引用与下载一览表)。在这张图里,可以清晰地看到引用、下载的关系。序号3、5、17、18都是零引用,而下载量超过100次。序号2、6引用只有一次,下载量小于100次。由此可见,研究互联网利用科技文献的事实数据已经成为对于科技文献进行评价所需要重视的参照体系。

另一个问题就是入选期刊范围和数量的多少,会使得被引频数的计算变化很大。比如由科学技术文献

1	1	图书馆学的历史源头	阮冈纳	世纪中国图书馆学	2004/09	99	518
2	2	杜邦图书馆与档案信息公平——纪念《情报科学工作》创刊30周年	阮冈纳, 黄良	情报科学工作	2010/06	1	30
3	3	论公共图书馆的典藏服务	阮冈纳	图书馆	2011/03	133	
4	4	主题表在“2010国际图书馆学会图书馆学委员会青年论坛”上的发言——纪念50次的世界大会	阮冈纳	现代图书馆	2011/03	8	
5	5	图书馆学概念与内涵的关联思考	阮冈纳	图书馆建设	2011/10	131	
6	6	公共图书馆建设和深刻“图书馆之城”建设	阮冈纳	图书馆	2006/01	1	76
7	7	公共图书馆的社会角色	阮冈纳, 周洁	图书馆理论与实践	2010/02	12	316
8	8	上海图书馆学研究二十年和回顾的缘起——写在《上海图书馆学研究工作》出版二十年之际	阮冈纳	上海图书馆学报	2010/01	37	
9	9	图书馆学内涵研究	阮冈纳	图书与情报	2010/02	13	474
10	10	图书馆社会服务专著导读	阮冈纳	图书馆建设	2010/07	3	216
11	11	新世纪10年我国图书馆基础理论的内涵与梦想	阮冈纳	基础图书馆工作	2010/04	12	232
12	12	理解图书馆社会责任——中国图书馆学会会长年启应	阮冈纳	图书馆理论与实践	2010/03	2	262
13	13	图书馆学理论的发展——图书馆社会责任的思考	阮冈纳	图书馆研究	2010/03	6	264
14	14	吴军中的图书馆学理论发展研究	阮冈纳, 陈小著	图书馆学报	2009/01	2	404
15	15	在新的信息和技术环境中中国图书馆的抉择——2006年的中外图书馆界会议和会议研究	阮冈纳, 李淑平, 俞晓正, 于文忠, 肖梅, 郭发成, 初胜利, 刘洪, 黄良	中国图书馆学报	2009/03	8	1301
16	16	云计算图书馆管理面临挑战	陈小著, 阮冈纳	大学图书馆学报	2009/04	111	2265
17	17	公共图书馆与城市文化	阮冈纳	公共图书馆	2009/03	192	
18	18	叶圣陶和图书馆学理论	阮冈纳	图书与情报	2009/06	147	
19	19	云计算和图书馆学: 云计算基础维护	阮冈纳	图书馆工作	2009/21	33	1729
20	20	图书馆2.0构建的图书馆服务	阮冈纳, 陈小著	大学图书馆学报	2009/01	496	3016

图2 范并思教授文章引用与下载一览表 (2012年1月18日检索获得)

出版社出版的《中国高被引指数分析(2011年版)》一书中,范并思教授的“图书馆2.0构建新的数字图书馆服务”总被引频数是71^[16],而在图2中,这个总被引频数高达456。

这个检索结果和《中国高被引指数分析(2011年版)》结果的对比中,有过写作引用的,引用频次在高被引与高下载量时是正交相关的,可以间接认为高被引是有价值的,但是低被引和下载量就不一定是正交相关了,那说明只从写作被引探究利用率是存在瑕疵的。

这个事实也间接说明了入口期刊的总数对被引频数、总下载等的影响很大,利用互联网,实时性的变化是显著的。

2 人物和机构的相关信息统计分析——辅助加权体系

由于任何专业的活动,都是以人为主要执行者的,而且大部分的人具有对应的机构属性,所以,如何利用人物和机构分析,对于科技问题观察角度的补充也是多角度科技评价体系所需要考虑的。利用公开的信息源是可以获得以下相关信息的:

(1) 作者、专家的基本信息。这些基本信息在期刊、项目申请、个人简介中都是有披露的。有很多信息是需要经过长期的收集与整理的。比如, 经历过程, 是

可以从发表的早期论文、到每次可以公开收集到的论文中看到学历、职称、专业方向的变化。为什么这些基本信息的积累比某一个文章的作者、专家的片段的信息更为重要? 因为科学、技术、工程是一个严谨、连续的过程, 通过积累这些信息才能比较完整地了解、判断一个作者和专家的基本情况。最基本的信息有姓名、职称、机构、主题(间接判断专业方向)、学历、专业、联络方式等。

(2) 作者、专家的专业轨迹。通过对作者、专家等在国际专业会议与国内专业会议的投稿、录用、选定主旨发言等的积累,可以看出作者、专家对于某个专业方向的关注程度、研究的方向、别人的兴趣程度来判断其专业的轨迹和影响情况。而这些国际会议与国内会议的论文发表情况、主题发言情况绝大部分是可以从互联网上公开获得的。在传统观点中,更多的是判断学术期刊文献的质量和会议论文的质量的优劣。而对于专业工作者而言,更多的是科研方向、学术研究的方法等的交流和讨论,所以,在国际和国内的专业会议上,对于科技交流和传播是最为重要的,而不是写作的水平是最重要的。在这些论文中,可以积累几个信息,如主题(专业方向)、机构、分组情况、会议报告人、时间、地址等信息。要特别重视国际和国内专业会议的论文录取、主旨发言,这些是科技信息的重要传播和交流的方式,会比期刊更早披露研究的方向、观点。而学术期刊一般是结题和完整的学术活动后的总结。

(3) 作者、专家的项目的轨迹。国际项目(中国申请的欧盟、美国、日本、德国等国际合作项目)、国内国家级的项目(自然科学基金、社科基金、973、863等国家级的项目)的参加情况,可以反映出这些作者、专家在国际项目合作、国内项目合作的研究方向、研究时间等。科学、技术、工程等的科技工作者除了研究工作的论文以外,副高以上职称才有资格申请国际、国内合作项目,这个也反映了专业技术水准。可以通过期刊论文、学位论文、会议论文、项目申请网站等获得公开的作者、专家参加项目的信息。这也是科技信息评价中很重要的辅助判断因素之一。

(4) 作者、专家参加互联网专业社区的活跃程度。作者、专家的对互联网社区、专业讨论群的参与程度、社区中的主题(专业方向)、讨论等是衡量在同类主题、学术观点讨论与参与的态度。而普通的微博、博客等更多的是粉丝、朋友之间的关注,大部分都不是以专业的探讨为目的的。很多社区是国际化的,这类社区也

不同于本国的专业社区,所以也要注意积累社区的国际性、地域性等相关的信息。而作者、专家开的个人网站,把自己的讲稿、国际、国内会议论文、研究的项目等进行公开,都是科技信息辅助判断需要的角度之一。

利用上述4种因素建立引用关系的加权系数处理,可以更好地判断期刊论文、会议论文、学术论文研究成果的质量,更好地评判专业的水平和专业水平的可信度空间,更好地判断专家的相对水平。

例如:某作者关于自然语言处理的算法的一篇期刊论文投稿,期刊编辑部请专家帮助判断论文水平,分析处理过程是:

(1) 这篇论文通过同方CNKI和万方数据的论文相似性检测工具检测——相似率为0%。

(2) 通过二次文献数据库,检索作者的专业轨迹。

用同方CNKI、万方数据以作者名作为输入条件,列出该作者的所有的论文。然后,对于作者写过的论文主题进行统计分析,发现作者的论文主题多为文献研究和以图情专业为主,没有自然语言处理相关特别是算法相关的论文——作者在自然语言处理专业的专业水平受到质疑。

(3) 再检索自然语言处理相关论文,找到与这个论文研究相关的作者或者专家。

确定真正的同行评审专家,判断这篇论文的相关疑点,最后确定本文是在更改了语料库,去掉了复杂语料情况下,而得出了研究的算法比别人好的结论。

3 建立互联网时代科技信息评价体系基础——累计性数据库结构

为了建立互联网时代科技信息评价体系,首先要建立、完善、累计科技信息相关的数据库结构:

3.1 科技文献资料数据库结构

(1) 对于各种文献建立包含文章名称、文章唯一标识、摘要(有则需要)、主题词(有则需要)、分类(有则需要)、作者、作者唯一标识、作者机构、机构地址(通信方式)的基础稳定数据库;

(2) 建立文章附属属性,包含文章唯一标识(与基础数据中的文章唯一标识一致)、表明文章来源与时间(期刊、学位、国际会议(分录用、专业发言)、国内会议

(分录用、专业发言)、是否大会发言交流(仅限专业会议)、是否是项目论文(仅限论文中表明来自于项目)的附属属性数据库;

(3) 建立文章被利用动态数据结构,包含引用该文章的文章名字、引用文章的来源与时间、引用该文章时间、引用者相关信息、被引用的文章出现在文章的结构部分(综述部分、问题与解决问题部分、结论部分、致谢部分)、对于被引用文章的语义(同意、不同意、修正)等动态应用数据库。

该数据库不同于传统的科技引文索引,增加了很多附属和动态的语义关系。例如通过动态应用数据库可以计量出引用和原文之间的时间间隔;可以看出引用在文章结构中对于文章的作用;可以计算出对于引用的语态,同意与否。

3.2 作者、专家数据库结构

(1) 对于出现过的作者、专家建立包括作者名、作者唯一标识、作者出生年月、作者出生地址、作者简介等基础稳定数据库;

(2) 建立作者、专家附属属性,包含作者唯一标识(与基础稳定数据库中作者唯一标识一致)、学历与学历时间、学位与学位时间、职称与职称时间、服务机构与服务时间、项目名与项目时间、互联网社区名和建立时间、个人专业网站唯一标识和建立时间等附属属性数据库;

(3) 建立作者和专家与文献的动态关系,包含文章唯一标识、文章名称、被引用累加次数、结构累加次数(综述部分、问题与解决问题部分、结论部分、致谢部分)、引用语义总次数(同意、不同意、修正)、专业社区讨论发表观点的总次数等动态更新数据库。

该数据库的累积与建立,帮助判断作者、专家的专业水平等有很好的辅助作用。

4 结语与建议

对于科技引文的统计分析方法,能否增加语义分析?这个是一个计算机全文处理中对于语义逻辑判断的算法过程,实现这种对于学术观点的同意、反对、同意并修正的具体价值的判断可能还需要一个漫长的阶段。

而基于作者、专家为主的相关的会议论文、各种项目、各种基金、个人博客、专业社区、个人网站等信息的

收集与整理目前是基本可行的,需要建立数据规范、发现与采集模式、利用技术手段实现是下一步的重点工作之一。

基于互联网利用的,应该和中国知网、万方、重庆维普等集中信息服务商进行合作,获取这类被利用的信息用于使用过程的分析是有可能的。

研究一个科技文献评价体系,以科技文献引用(写

作角度)和科技文献应用(使用角度)为核心,以人物、机构、主题、时间相关性等为加权关系,建立一个新型的模型,是图情文献计量学理论发展的一个值得探索的方向。

利用十多年的数字图书馆的积累,利用技术,来进行科技文献更加准确、效率更高、更多角度的评判工作也是数字图书馆服务更加深入的一个具体体现。

参考文献

- [1] 2010年世界500强企业排名表[OL].[2011-02-20]. <http://www.it.com.cn/news/quanzi/2010/05/31/09/812731.html>.
- [2] 爱斯维尔面向研究团队服务[OL].[2011-02-20]. <http://www.info.scival.com/>.
- [3] 万方数据辅助决策支持系统[OL].[2011-02-20]. <http://stads.wanfangdata.com.cn/>.
- [4] 中国知网[OL].[2011-02-20]. <http://www.cnki.net/>.
- [5] 爱斯维尔信息服务[OL].[2011-02-20]. http://www.elsevier.com/wps/find/homepage.cws_home.
- [6] 1950年到2050年可能消失的行业[OL].[2011-02-20]. http://nowandnext.com/PDF/extinction_timeline.pdf.
- [7] 陈光.文献老化[OL].[2011-02-20]. <http://www.lis.ntu.edu.tw/~khchen/course/bm/bm2006/bm200612.pdf>.
- [8] 钟晶晶.游毅,索传军.新信息环境下数字文献老化趋势及影响因素新探[J].情报杂志,2011,30(12): 39-42.
- [9] JONES E. E-Science Talking Points for ARL Deans and Directors[OL].[2011-02-20]. <http://www.arl.org/bm~doc/e-science-talking-points.pdf>.
- [10] 国家自然科学基金[OL].[2011-02-20]. <http://www.nsf.gov.cn/Portal0/default152.htm>.
- [11] 美国自然科学基金[OL].[2011-02-20]. <http://www.nsf.gov/>.
- [12] 科学素养的意义及其在九年一贯自然与生活科技领域课程的启示[OL].[2011-02-20]. <http://163.32.133.206/htm/ceag2/htm/Science/f/f004/f004.htm>.
- [13] 激励发现、推动创新[OL].[2011-02-20]. <http://www.lib.whu.edu.cn/sybz/204-1.ppt>.
- [14] 爱斯维尔公司[OL].[2011-02-20]. <http://www.scopus.com/home>.
- [15] 汤姆森路透公司[OL].[2011-02-20]. <http://science.thomsonreuters.com.cn/productservices/webofscience/>.
- [16] 曾建勋.中国高被引指数分析[M].科学技术文献出版社,2011:578.

作者简介

孙卫,中国科学技术信息研究所,计算机高级工程师,中国计算机学会高级会员,北京通信学会理事,中国图书馆学会会员。目前主要在知识组织、知识挖掘、知识处理技术上进行研究及教学。科学出版社,从事数字出版相关工作。E-mail:sunwei0125@gmail.com

Study on the Evaluation Factors on the Scientific and Technological Documents in the Internet Era

Sun Wei / Institute of Scientific and Technical Information of China, Beijing, 100038
/ Science Press Ltd., Beijing, 100717

Abstract: In terms of the evaluating of scientific and technological document, it seems insufficient, in the internet era, to make a judgment of an article only from the point of view of how much it was cited. Due to the diversity of the scientific and technological views published, and the speed and broadness of information spread, many factors enable scientific and technological resources to be easily utilized nowadays. Thus, the attitude of the document user toward the article, the scientific and technological resources published by the writer or professionals, and the ability of the readers in finding relevant information through internet may become the assessment factors of the resources.

Keywords: Citation analysis, Author analysis, Organization analysis, Search results analysis

(收稿日期: 2012-02-20)