

# 国家科技基础条件平台网站影响力评价

张显董诚

(中国科学技术信息研究所, 北京 100038)

**摘要:** 通过用户视角, 从推广宣传的角度出发, 设定搜索引擎中结果排名、平台官网的外链数量、平台的学术论文贡献情况等指标, 探究了国家科技基础条件平台网站的影响力, 得出搜索引擎影响力与论文产出贡献率显著正相关等相关结论, 并提出相关的评价建议。

**关键词:** 国家科技基础条件平台; 网站; 影响力; 科研贡献率; 搜索引擎; 评价

中图分类号: G311

文献标识码: A

DOI: 10.3772/j.issn.1674-1544.2014.03.013

## Study on the Influence of National Science & Technology Infrastructure Website

Zhang Xian, Dong Cheng

(Institute of Scientific and Technical Information of China, Beijing 100038)

**Abstract:** In this study, as a third party we based on the user's real need, start from the perspective of propagandizing and popularizing, and set a series of indexes such as the rank of the website in search engine result, the number of incoming links, the academic contribution to study the influence of national science & technology infrastructure website. Finally, we came to the conclusion that website's influence power and academic output were significantly positive correlation, and give some suggestions such as open network commissioner.

**Keywords:** national science & technology infrastructure, website, influence, academic contribution, search engine, evaluation

### 1 引言

从互联网发展至今, 国内外关于网站影响力评价有不少的研究。1995年Allison woodruff通过爬虫工具收集了大量网页, 分析页面的链接, 筛选出入链最高的网站, 通过分析与实证表明, 网页的入链可以作为网站影响力的一个指标<sup>[1]</sup>。1998年Peter ingwersen提出网络影响因子(Web-

IF), 他指出Web-IF是“在某一时间, 来源于外部和自身内部指向特定国家或网站的页数与该国家或网站中网页页数之比”引起和很大关注<sup>[2]</sup>。2002年, 邱均平通过对42种工程类中文期刊网站的入链和Web-IF进行计量分析, 发现这两项指标与期刊影响因子有统计意义的相关关系<sup>[3]</sup>。2007年曾荷通过建立4个评价指标: 一般访问量、深度访问量、可见度、链接指标来对电子政

**作者简介:** 张显\*(1991-), 男, 中国科学技术信息研究所硕士研究生, 研究方向: 科技资源共享。董诚(1970-), 男, 中国科学技术信息研究所研究员, 研究方向: 科技资源管理。

**收稿日期:** 2013年12月25日。

务信息资源网站影响力指数进行测评<sup>[4]</sup>。2009年陈斯杰通过对影响力内涵的探究，构建了包含外部变量和内生变量的影响力概念模型，其中外部变量包含信息质量、系统响应、服务可获得性、交互性等；内生变量包含感知特性、可靠性、认可度、忠诚度及影响力<sup>[5]</sup>。2013年许东惠、吕志先等在国家科技基础条件平台绩效概念界定的基础上，对平台运行服务绩效评价方法和影响因子进行了系统分析，并构建了国家科技基础条件平台运行服务绩效考核的指标体系<sup>[6]</sup>。

可以看出，国内外关于网站影响力的研究，定量的方法主要通过网站的入链以及Web-IF等与网站链接相关的指标来衡量，定性的方法则一般集中在资源的形式、内容建设、组织机构等方面，因此，评价主要集中在运行管理和服务绩效等方面。

本研究在收集国家科技基础条件平台网络调查数据的基础上，从普通用户的角度出发，对23家科技基础条件平台网站的搜索引擎排名和分布情况进行分析，以期揭示出不同平台的网络可见性及其学术论文贡献度，方便、完善平台网站影响力的评估。

## 2 评价指标体系

根据最小努力原则，人们的各种社会活动，总是想以最小的代价获得最大的收益<sup>[7]</sup>。在Web2.0时代更是如此，搜索引擎成为了用户的主要信息获取渠道。所以，此次调查指标都建立在搜索引擎的基础上，从平台网站的可见性和易见性两方面考虑，分别设立了平台网站搜索引擎可见性和搜索结果TOP30分布等二级指标；同时通过知网学术搜索来了解各个平台的学术产出贡献度，探寻其与网站影响力之间的关系，并辅助佐证指标设定的有效性。根据美国最早的搜索引擎营销专业服务商iProspect (www.iprospect.com) 和市场研究公司Jupiter Research联合调查：互联网用户使用搜索引擎越来越没有耐心，越来越多的互联网用户仅关注搜索结果第一页的内容，如果第一页面没有满意的结果，立刻变换关键词或

者更换搜索引擎重新进行检索，所以考虑官网出现排名情况时，只考虑第一页的结果；同时，高达90%的搜索者只查看搜索结果前三页，所以进行结果页面分析时，只考虑前三页也就是TOP30的结果<sup>[8]</sup>。在考虑学术产出影响力的时候，由于用户引用习惯的不同，我们认为只要该平台名字在论文中任何地方被提及，都算做该平台对该论文有贡献，所以我们选择具有全文检索的知网数据库和检索比较全的谷歌学术进行检索。具体指标如表1所示。

## 3 评价方法

### 3.1 评价原则

(1)以网站评价为起点和抓手。网站是科技资源信息对外揭示的主要的、最快捷的手段，因此，本评估的指标设计、评价过程安排等应紧密结合网站特点进行设计和部署。

(2)以用户需求为导向。科技基础条件平台是为了满足用户需求，本次所进行的平台网站影响力评价就是从用户角度出发的。

(3)以第三方立场为视角。本次所有调查和评价的参与者都是独立于国家科技基础条件平台的第三方普通用户，能够更加客观地研究和实施评价。评价基于网络客观数据，从公众用户的角度出发，符合科技基础条件平台的相关规定。

(4)指标设计应尽量简单，易于量化和获取。避免过于理论或者学术化，或者与科技资源信息开放共享无关的指标。

### 3.2 权重的设定

本研究中，主要采用层次分析法确定指标的权重。层次分析法是对人的主观判断作定量描述的一种方法，尤其适用于多目标的定性为主的决策<sup>[9]</sup>。使用这种定性分析和定量分析相结合的决策分析方法，我们首先要把评估的目标分解成若干层次，建立起有序的梯阶层次结构（即层次结构模型），从而使人的经验和判断能用数量形式加以表达和处理。

### 3.3 数据获取方式

(1)主要采用人工采集的方法，从搜索引擎

表1 国家科技基础条件平台网站影响力评价指标体系

一级指标	二级指标	三级指标	评分细则	权重/%
A搜索引擎影响力	A <sub>1</sub> 平台官网搜索引擎可见性	A <sub>11</sub> 官网在搜索结果中排名	以平台名称为关键词在搜索引擎中精确检索, 第一页结果中, 从下往上, 数官网出现的位置为其得分, 官网越靠前, 得分越高	18.25
		A <sub>12</sub> 指向平台官网的连接数量	使用open site explorer获取各个平台的外链主站数	11.75
	A <sub>2</sub> 搜索结果TOP30内容分布	A <sub>21</sub> 官方网站	统计搜索结果中官方网站(包含子页面)出现的次数	8.25
		A <sub>22</sub> 新闻类网站	统计搜索结果中新闻报道网站出现次数	4.25
		A <sub>23</sub> 文档类网站	统计搜索结果中文档类网站出现次数(如知网、万方、豆丁、百度文库等)	3.5
		A <sub>24</sub> 社交类网站	统计搜索结果中社交类网站出现次数(如贴吧、微博、论坛等)	1.75
		A <sub>25</sub> 百科类网站	统计搜索结果中百科类网站出现次数(如百度百科, 维基百科等)	7.25
		A <sub>26</sub> 机构类网站	统计搜索结果其他专业机构网站出现次数	3.25
		A <sub>27</sub> 导航链接类网站	统计搜索结果中导航类网站出现次数	1.75
B学术产出贡献度	B <sub>1</sub> 学术产出提及度	B <sub>11</sub> 知网学术论文提及度	在知网全文检索里以各平台名称为关键词进行精确检索, 统计检索结果次数	24.75
		B <sub>12</sub> 谷歌学术提及度	谷歌学术里, 以各平台名称为关键词检索, 统计检索结果	15.25

入口调查数据, 调查内容按照评价指标进行分解, 进而转为对各平台网站在搜索引擎中的收录情况和科研引用情况进行调查。在此过程中, 部分数据通过抽样方式获得, 如在所有检索结果中, 仅挑选前3页的内容进行分类分析。

(2) 部分指标通过专用网站工具获得。如对于各平台网站的外部链接情况, 使用open site explorer获取。

### 3.4 信息的处理方法

(1) 信息量化方法。该方法的主要功能是将定性信息进行量化。在本研究中, 指标都是定量指标。

(2) 信息的预处理方法。通过信息的预处理可以将不同种类、不同量纲的信息进行规范化处理, 使之具有统一的标准。若各种信息不按统一的标准进行预处理, 那么, 有些信息可能会被其他的信息淹没, 致使评价结果没有意义。本研究的信息预处理采用无量纲化和相对化技术。

(3) 信息的计算方法。本研究采用线性加权模型<sup>[10]</sup>, 将经过预处理的信息进行综合计算, 并得出评价结果。计算公式如下:

$$r(X_i) = \sum_{j=1}^m W_{ij} X_{ij} \quad (1)$$

$$r(X) = \sum_{i=1}^n [W_i r(X_i)] = \sum_{i=1}^n W_i \sum_{j=1}^m W_{ij} X_{ij} \quad (2)$$

其中,  $X_{ij}$ 为底层指标的评价值,  $W_{ij}$ 为底层指标权重,  $W_i$ 为上一级指标权重。将每个底层指标进行分别评价, 并将评价量化, 再将评价与表述该指标相对重要性的权重系数相乘, 同时对上一级指标求和, 得到上一级指标的评价值。直至达到顶层指标, 从而得到综合评价结果。

## 4 评价结果

### 4.1 总体评价

根据上述指标体系, 通过调查和统计计算得到国家科技基础条件平台网站公众影响力的评价结果。总体得分情况如表2所示, 分段统计结果如图1所示。

通过表1、图1可以看出, 从搜索引擎的入口出发, 各个平台的影响力表现整体一般, 高分区90~100的没有, 80以上的也不多。原因可能是, 各个平台都把工作中心放在了资源的建设

表2 国家科技基础条件平台得分情况汇总表

排名	平台名称	总分	搜索引擎影响力	学术产出影响力
1	北京离子探针中心	87.271	47.271	40.000
2	气象科学数据共享中心	80.690	43.770	36.920
3	中国数字科技馆	70.905	51.145	19.761
4	农业科学数据共享中心	70.231	50.059	20.172
5	地震科学数据共享中心	69.091	49.159	19.933
6	国家农作物种质资源平台	65.736	47.725	18.011
7	国家标准文献共享服务平台	64.535	51.560	12.976
8	地球系统科学数据共享平台	62.708	46.494	16.214
9	家养动物种质资源平台	62.081	44.309	17.772
10	国家标准物质资源共享平台	61.849	47.317	14.532
11	国家材料环境腐蚀野外科学观测研究平台	60.907	42.617	18.290
12	国家微生物资源平台	60.143	37.679	22.464
13	水产种质资源平台	58.783	41.570	17.213
14	人口与健康科学数据共享平台	58.438	41.994	16.445
15	国家实验细胞资源共享平台	55.864	38.893	16.971
16	国家大型科学仪器中心	55.326	34.643	20.683
17	科技文献共享平台	50.185	18.247	31.938
18	国家林木种质资源平台	49.521	36.161	13.360
19	国家生态系统观测研究网络	44.350	28.911	15.440
20	国家计量基标准资源共享基地	43.813	36.314	7.499
21	标本资源共享平台	43.141	25.485	17.656
22	林业科学数据平台	43.006	30.543	12.463
23	中国应急分析测试平台	42.303	34.729	7.574

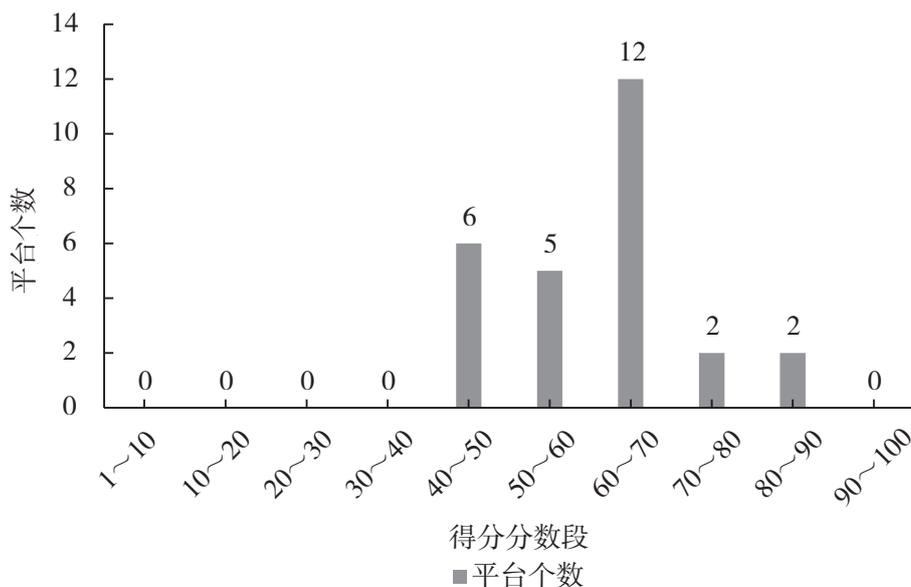


图1 国家科技基础条件平台网站影响力得分分段统计结果

和平台管理的完善上，而忽略了平台的推广，从而降低了平台的搜索引擎可见性和影响力。

针对调查对象得分情况，我们做出二级指标

平均得分以及得分率，结果如图2所示。

由图2可以看出，搜索结果TOP30分布得分率和学术产出影响力得分率比较低。这从一定程

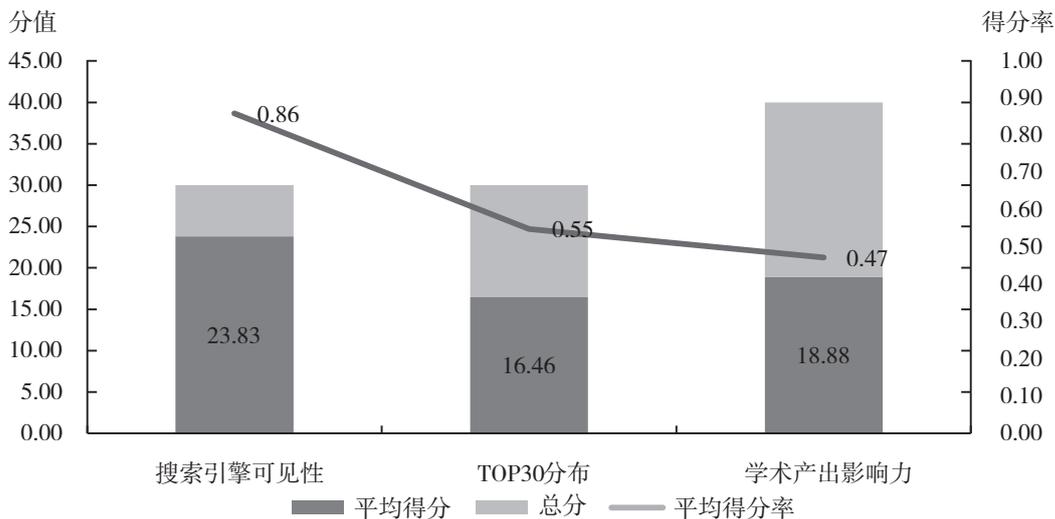


图2 国家科技基础条件平台网站影响力二级指标平均得分及得分率

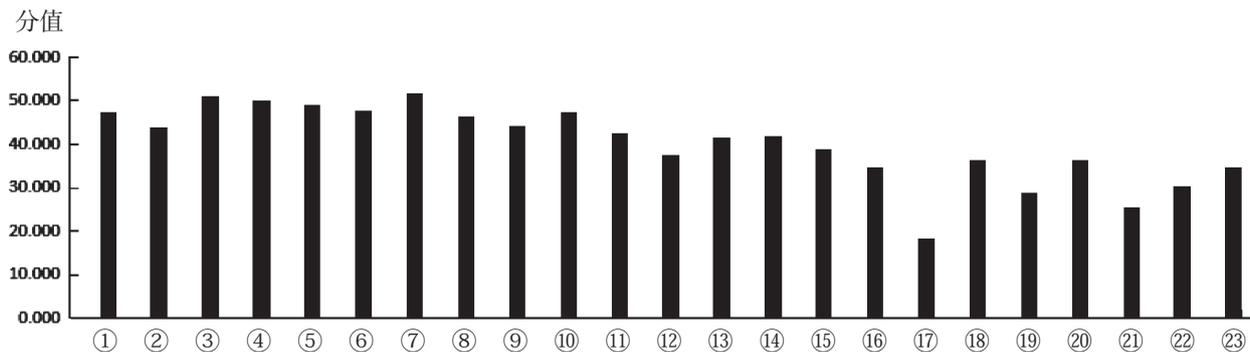


图3 国家科技基础条件平台网站搜索引擎影响力得分图

①—北京离子探针中心；②—气象科学数据共享中心（中国气象科学数据共享中心服务网）；③—中国数字科技馆；④—农业科学数据共享中心；⑤—地震科学数据共享中心；⑥—国家农作物种质资源平台（中国农作物种质信息网）；⑦—国家标准文献共享服务平台；⑧—地球系统科学数据共享平台；⑨—家养动物种质资源平台；⑩—国家标准物质资源共享平台；⑪—国家材料环境腐蚀野外科学观测研究平台（中国腐蚀与防护网）；⑫—国家微生物资源平台；⑬—水产种质资源平台；⑭—人口与健康科学数据共享平台；⑮—国家实验细胞资源共享平台；⑯—国家大型科学仪器中心；⑰—科技文献共享平台；⑱—国家森林（含竹藤花卉）种质资源平台；⑲—国家生态系统观测研究网络；⑳—国家计量基准（物理部分）资源共享基地；㉑—标本资源共享平台；㉒—林业科学数据平台；㉓—中国应急分析测试平台。

度上说明各个平台在搜索结果优化和学术产出影响力方面仍有改善的空间。

#### 4.2 搜索引擎影响力

搜索引擎影响力的各个平台得分情况如图3所示。图3展现了各个平台搜索引擎影响力的得分情况，反映的是用户通过搜索引擎入口，找到平台官方网站和获得相关信息的难易程度，以及搜索结果TOP30的分布情况。这从一个侧面反映了各个平台网站在搜索引擎方面的影响力。从结果来看，各个平台得分比较均衡，虽有高低之

分，但没有非常显著的差异。

通过图2可以发现，一级指标搜索引擎影响力下的二级指标TOP30分布得分率比较低。通过分析该二级指标下的三级指标的平均得分与得分率，得到结果如图4所示。

通过图4可以发现，社交类网站、百科类网站、导航类网站平均得分率低于60%，而这一类网站则主要是以用户互动为主，有明显的Web2.0特点。可见，各个平台在和用户互动以及用户之间互动方面的重视程度不够。

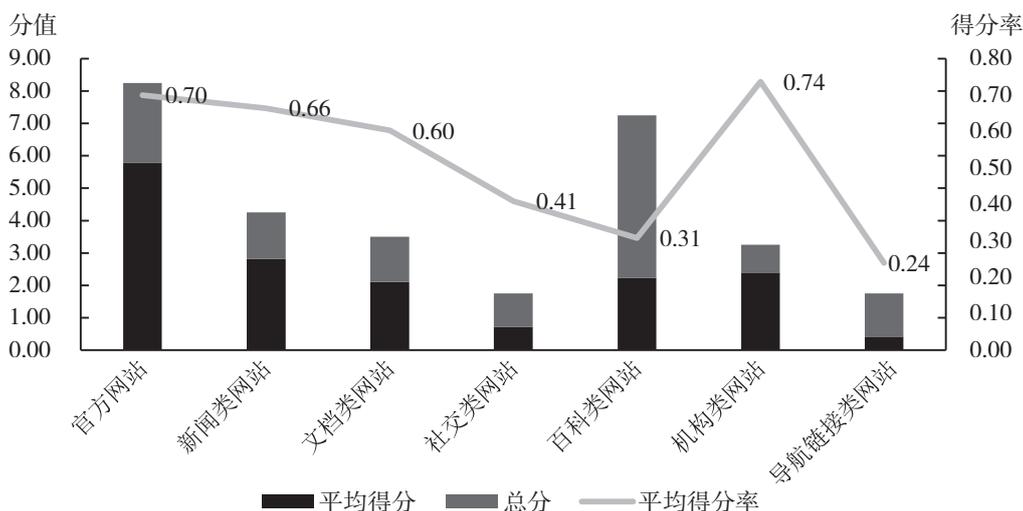


图4 国家科技基础条件平台搜索结果TOP30分布三级指标平均得分及得分率率

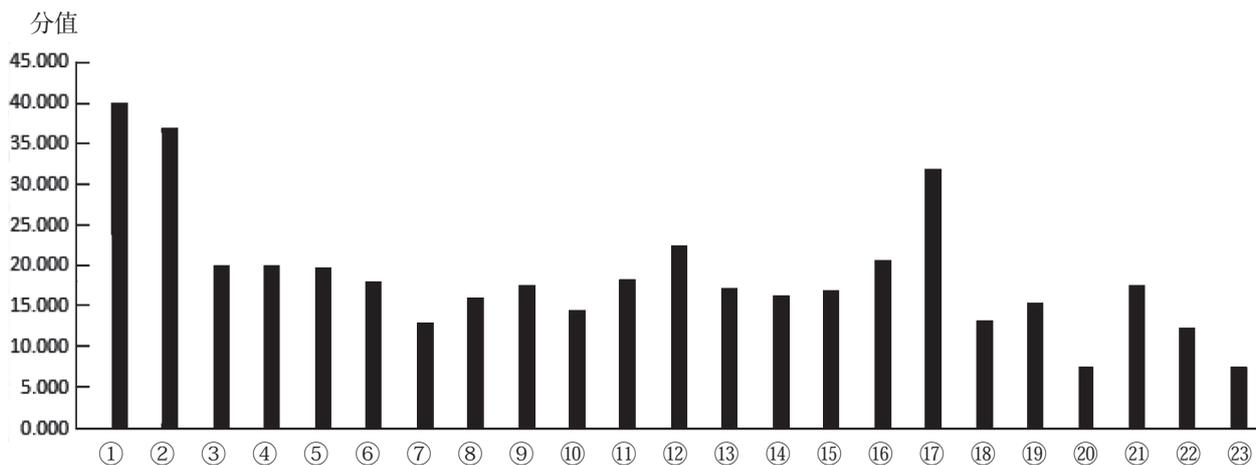


图5 国家科技基础条件平台科研贡献性评价得分

①—北京离子探针中心；②—气象科学数据共享中心（中国气象科学数据共享中心服务网）；③—中国数字科技馆；④—农业科学数据共享中心；⑤—地震科学数据共享中心；⑥—国家农作物种质资源平台（中国农作物种质信息网）；⑦—国家标准文献共享服务平台；⑧—地球系统科学数据共享平台；⑨—家养动物种质资源平台；⑩—国家标准物质资源共享平台；⑪—国家材料环境腐蚀野外科学观测研究平台（中国腐蚀与防护网）；⑫—国家微生物资源平台；⑬—水产种质资源平台；⑭—人口与健康科学数据共享平台；⑮—国家实验细胞资源共享平台；⑯—国家大型科学仪器中心；⑰—科技文献共享平台；⑱—国家森林（含竹藤花卉）种质资源平台；⑲—国家生态系统观测研究网络；⑳—国家计量基准（物理部分）资源共享基地；㉑—标本资源共享平台；㉒—林业科学数据平台；㉓—中国应急分析测试平台。

### 4.3 科研贡献性

我们认为只要某一篇文章中出现了某平台的名字，就算该平台为该文章贡献一次，而不去深入追究该文章是否真的使用了该平台的资源。这样，我们选取知网的全文检索和谷歌学术，精确检索每个平台的名称，得出每个平台的数据值，在经过打分加权计算，得到最终结果，如图5所示。

通过图5可以看出，各个平台的科研贡献性

差距很大，高低分差比较大。当然，这跟不同平台的性质有关，也涉及学术诚信问题。在这里，我们不作进一步讨论。

考虑到搜索引擎影响力和科研贡献性之间的关系，我们在SPSS中对这两个变量做了相关性检验，得到结果如表3所示。

通过表3可以看出，搜索引擎影响力和科研贡献性显著正相关，说明科技基础条件平台网站

表3 国家科技基础条件平台网站搜索引擎影响力和科研贡献性相关性结果表

			搜索引擎可见性	学术论文贡献度
Kendall 的 tau_b	搜索引擎可见性	相关系数	1.000	0.364*
		Sig. (双侧)	0	0.015
		N	23	23
	学术论文贡献度	相关系数	0.364*	1.000
		Sig. (双侧)	0.015	0
		N	23	23
Spearman 的 rho	搜索引擎可见性	相关系数	1.000	0.505*
		Sig. (双侧)	0	0.014
		N	23	23
	学术论文贡献度	相关系数	0.505*	1.000
		Sig. (双侧)	0.014	0
		N	23	23

注：\*为在置信度（双侧）为 0.05 时，相关性是显著的。

在搜索引擎里的排名以及搜索结果优化等等会影响到该平台资源的利用，进而影响到科研产出，同时也佐证了我们指标设定的合理性。

## 5 结论及建议

### 5.1 结论

本次评价在理论上，继承和发展了以往的评价体系，采取了传统的连接数量方面的指标，并在其基础上做了补充，加入搜索引擎排名以及搜索结果分布方面的指标，并且考虑到了学术方面的影响力，加设了相应的指标；摒弃了晦涩难懂的技术指标，而从普通用户的角度出发，具有一定的实践指导价值；以用户为中心是核心思路，可以扩展到其他领域。整体来看，就本次评价而言，有以下几个方面关键结论：（1）国家科技基础条件平台网站影响力表现一般，平台推广、宣传工作仍需加强。（2）平台学术论文贡献率和平台网站搜索引擎影响力呈显著性正相关关系。学术论文贡献方面表现好的平台，同样网站的搜索引擎影响力也比较大，优化平台网站将是提高科研产出的另一途径。（3）平台网站涉及社交类网站、导航类网站、百科类网站的程度比较低。平台网站在向 Web2.0 转型中不成功，仍然是以单方的资源提供，另一方的资源接受，平台与用户、

用户与用户间的互动较差。

本研究也有不足的地方，如搜索引擎选择方面的用户差异性没有考虑到，学术产出影响力方面指标仍需要进一步优化等。今后的研究将会集中在各个平台用户粘性以及平台科研贡献方面。

### 5.2 建议

针对研究中发现的问题，结合目前我国国家科技基础条件平台考核评价状况，在改进和完善影响力评价进而提升服务绩效方面，提出以下几点建议。

（1）针对不同的用户，如政府决策者、科研人员、普通公众等进行分类调查，制定符合不同用户使用习惯的个性化影响力评估考核体系，有针对性地完善对不同用户提供的服务。

（2）将网站影响力评价作为国家科技基础条件平台绩效考核的主要内容之一。通过丰富完善已有的平台考核指标体系，来加强平台承建单位对网站优化和开放共享的重视，提高平台资源的利用率。

（3）分析用户访问记录，了解用户习惯，针对用户偏好，不断调整优化相应指标权重，完善影响力评价体系，使其更加符合用户的真实需求，进而提升平台资源的利用率和服务绩效。

在不断完善和改进各平台网站影响力评价的同时,各平台承担单位也可以通过其他手段来增加其平台网站影响力,如设立网络专员、开设有针对性的社交网络账户或者社交平台、建立与搜索引擎的合作等。

在数据时代,资源异常丰富,同时也增加了筛选和查找的难度。这就要求资源拥有方扩大影响力,更加方便、快捷地向用户提供优质资源。网站这一信息传播通道将会变得更加重要。我们相信,在资源不断建设和影响力不断加强的背景下,我国科技事业必将得到更加健康全面的发展。

### 参考文献

- [1] Allison Woodruff. An Investigation of Documents from the World Wide Web[DB/OL]. [2014-01-17]. <http://db.cs.berkeley.edu/papers/www5-docs/www5-submit/Overview.html>.
- [2] Peter Ingwersen. The Calculation of Web Impact Factors[J]. *Journal of Documentation*, 1998,54(2):236-243.

- [3] 邱均平,安璐.中文期刊影响因子与网络影响因子和外部链接数的关系研究[J].*情报学报*,2003,22(4):398-402.
- [4] 曾荷.电子政务信息资源的网站影响力评价研究[D].上海:华东师范大学,2007.
- [6] 陈斯杰.基于用户视角的科技信息服务网站影响力评估研究[D].南京:南京理工大学,2009.
- [6] 许东惠,吕先志,袁伟,等.国家科技基础条件平台运行服务绩效考核指标体系研究[J].*中国基础科学*,2013(1):40-43.
- [7] 百度百科.最小努力原则[EB/OL].[2014-01-17]. <http://baike.baidu.com/link?url=-aEYqWXZZEbBXkN6CC-gPTCYN0rGIPO9Qrye4BpXSYIJ60PbwnIhoegJhimeD9M8qLDmbTpMVpcnpS8YevLkhl>.
- [8] 新竞争力.多数互联网用户仅关注搜索结果第一页的内容[EB/OL]. [2014-01-17]. <http://www.jingzhengli.cn/baogao/f20060413.htm>.
- [9] 朱茵,孟志勇,阚叔愚,等.用层次分析法计算权重[J].*北方交通大学学报*,1999,23(5):119-122.
- [10] 赵炜,魏玲,许军,等.线性加权评价与聚类分析理论及应用[J].*纯粹数学与应用数学*,2002,18(2):121-125.

(上接第67页)

## 6 结语

地学文献书目数据资源建设和运行,基本实现了服务国土资源部和地质调查局、面向全国的共享服务体系,不仅在中国地质调查局所属单位产生了广泛的影响,而且,随着文献资源工作的不断推进,用户对地学文献资源需求的加大及文献共享服务体系的发展,地勘行业、地学类高校具有特色收藏单位的加入,将不断扩大地学文献资源共享平台的服务范围。

地学文献是专业图书馆藏书体系中的独特资源,具有鲜明的特点且数量稀少。共享平台的建设应该是成员馆相互依赖、相互促进、共同发展。在中国地质调查局属单位地学文献资源基础上,联合更多的地勘行业图书馆、高校地学类图书馆等文献资源单位,逐步构建起全国地学文献资源保障体系。这个服务体系应该是:在线、机读、动态、可不断扩充和完善、供科研人员及公众共同使用的地学文献资源保障中心<sup>[8]</sup>。

### 参考文献

- [1] 马翠凤,蔡秀华,梁世莲,等.地勘行业文献信息联机联合编目共享服务体系建设[J].*国土资源科技管理*,2012,29(5):95-100.
- [2] 喻爽爽.CALIS联机合作编目服务体系[J].*数字图书馆论坛*,2013(1):37-46.
- [3] 秦聿昌,薛慧彬.分报告一:美国国会图书馆考察报告[J].*数字图书馆论坛*,2011(1):2-15.
- [4] 顾晓华.国家地学文献信息资源有效利用的思考[J].*地质与资源*,2011,20(1):70-74.
- [5] 周津慧,初景利.分报告七:OCLC考察报告[J].*数字图书馆论坛*,2011(1):61-70.
- [6] 谢琴芳,喻爽爽,刘素清.检索控制的全方位实现——CALIS联机规范控制系统[J].*大学图书馆学报*,2004(4):48-52.
- [7] 袁永旭,贺培风,王秀平,等.山西省医学科技文献信息资源与服务平台建设机制探索与研究[J].*医学信息学杂志*,2009,30(11):50-53.
- [8] 胡广翔,庞莹.中国社会科学院图书馆联机联合目录系统结构特点及功能[J].*情报资料工作*,2001(4):77,45.