

The Methods of Subject Forecast Based on the Databases of ESI and SCIE: Taking Qingdao University for Example

Baoyan Wang, Shuling Kong, Lin Cao

Library of Qingdao University, Qingdao Shandong
Email: bywang@qdu.edu.cn

Received: Apr. 3rd, 2018; accepted: Apr. 17th, 2018; published: Apr. 23rd, 2018

Abstract

The key of the "Double World-classes" Development is that subjects are world-class. It requires more subjects to enter the world's top 1%, or even the top 1‰. How to discern and identify potential subjects and the precise prediction of potential subjects have been received great interest for the departments of scientific research and management in the universities. This would promote more disciplines theoretically into the top positions in the global ranking. Based on the data accumulation in the Web of Science database, combined with the successful prediction of chemistry of Qingdao University, this paper discusses and evaluates other subjects in Qingdao University by using the least squares method. The promotion of potential ESI subjects is a long-term system task. This method has the practical significance of its prediction and evaluation of potential subjects. It should be applicable to other universities and research institutes to carry out their dynamic monitoring and strategic adjustment of the potential subjects.

Keywords

ESI, Subject Forecast, Potential Subjects, Threshold, Access Rate of Subjects

基于ESI和SCI-E数据库的学科发展预测方法：以青岛大学为例

王宝燕, 孔淑玲, 曹琳

青岛大学图书馆, 山东 青岛
Email: bywang@qdu.edu.cn

收稿日期: 2018年4月3日; 录用日期: 2018年4月17日; 发布日期: 2018年4月23日

摘要

“双一流”的关键是学科世界一流。这就需要更多学科进入全球排名前1%、甚至前1%。如何发掘、识别潜力学科，同时对潜力学科的发展前景进行精准预测，从理论上指导更多学科进入全球排名前列，已经引起了各科研管理部门的重视。本文基于Web of Science数据库中的数据积累，利用最小二乘法，结合对青岛大学化学学科成功预测的基本方法，对其他潜力学科进行了识别与评测。推动ESI潜力学科的提升是一个长期的系统工程，本文方法基于数据的积累，具备识别和评估潜力学科，并对其预测的实践意义。该方法适用于其他高校及科研院所对本机构的潜力学科进行动态监测及策略调整。

关键词

ESI, 学科预测, 潜力学科, 门槛值, 学科准入率

Copyright © 2018 by authors and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

2015年11月国务院印发关于《统筹推进世界一流大学和一流学科建设总体方案》的通知(国发[2015]64号)[1]，各级政府[2][3][4]及高校[5]相继推出了适合本地区本单位“双一流”建设发展的政策方案。基本科学指标数据库(Essential Science Indicators、简称ESI)作为一个学科评价和跟踪学科发展趋势的工具，近年来受到国内外各高校和科研机构的重视。在《2012年学科评估指标体系》中ESI高被引论文(Highly Cited Papers)首次纳入大学评价指标。在一流学科建设方面相关学科进入ESI全球排名前1%、前1%等指标已广泛作为衡量高校学科进入国际先进行列的一个重要指标。因此，通过ESI跟踪学科发展动态，使科研管理人员及学者们了解某学科在世界所处的位置及未来的发展趋势尤为重要。目前，ESI的相关研究较多，但关于高校或科研机构利用ESI进行潜在学科预测方面的研究却不多，董政娥[6]等提出了“学科比重”预测方法，是用某机构某学科的被引频次与ESI相应学科的总被引频次的比值，ESI学科总被引频次基数大，导致结果间差距小；且在文后可以看出比值的大小并不能作为进入ESI前1%的依据，如东华大学化学学科的学科比重小于该校的数学、物理及计算机科学等学科，但东华大学化学学科早已进入ESI全球前1%的排名，且排名还比较靠前。由此可以看出利用学科比重用于预测不是很直观。任勇[7]等提出“学科ESI准入率”概念，其计算公式为 $Z_i = B_i/C_i$ ，其中i为ESI 22个学科中第i个学科， B_i 为SCI-E数据库中某机构对应的ESI学科归属期刊中所有发文的总被引频次， C_i 为该学科的ESI被引频次门槛值，但仅仅用了单次的更新数据。

本文以青岛大学(以下称“本机构”)为例，利用ESI数据库以及SCI-E数据库通过科学计量学的相关统计方法，在近三年的时间里，在ESI每个更新时段持续跟踪分析了本机构近十年数据，并加以对比分析。本文沿用“学科准入率 Z_i ”的概念， Z_i 值接近1，表明对应学科的总被引频次接近门槛值，该学科则越有可能进入ESI前1%。选取本机构各学科 Z_i 值排名靠前的几个学科作为潜力学科，对本机构最有潜力进入ESI全球前1%的学科化学、神经科学与行为及材料学提出预测；希望本方法能帮助其他科研机构掌握自身的学科动向，为科研管理与决策参考提供客观详实的数据支撑，通过更深入的统计分析，为

自身的潜力学科早日进入 ESI 提出相应建议。

2. 数据来源及统计预测方法

2.1. 数据来源

本文选取两组数据,一为 ESI 数据库中 22 个学科被引频次阈值(Citation Thresholds) (在 ESI 各学科中,进入该学科的所有机构最后一名的被引频次,如表 1 所示,因为阈值是一个动态变化值,表 1 中给出的是 2017 年 5 月份的更新数据);二为与 ESI 相对应学科的本机构总被引频次(未进入 ESI 前 1% 的学科无法利用 ESI 进行统计,故本文利用 SCI-E 数据库进行统计)。具体方法为:利用 ESI 数据库中期刊列表,导出 22 个学科的期刊列表,形成检索式,导入 SCI-E 数据库中,限定时间与 ESI 每个更新时段所涵盖数据范围相对应,限定文献类型 Article 和 Review 等条件,力求与 ESI 数据库的数据来源相一致,截止 2017 年 5 月 15 日的本机构相对应 22 个学科被引频次见表 1 所示。

Table 1. Thresholds for 22 subjects and citations of individual subjects and Zi values in Qingdao University (ESI Update 2017-05)

表 1. 22 个学科的阈值及青岛大学各学科被引频次及 Zi 值统计表(2017 年 5 月 ESI 更新数据)

学科	ESI 全球前 1% 学科阈值	被引频次(青岛大学)	学科准入率 Zi
临床医学	1855	9433	5.09
工程学	1876	6649	3.54
化学	5981	9724	1.63
神经与行为科学	4959	4590	0.93
材料学	3918	3438	0.88
药理学与毒理学	2864	1764	0.62
计算机科学	2837	1294	0.46
生物与生物化学	5079	2224	0.44
交叉学科	2087	1072	0.51
物理学	14029	3508	0.25
数学	3620	647	0.18
环境生态学	3319	491	0.15
免疫学	3844	665	0.17
农业科学	1692	276	0.16
动植物学	2236	316	0.14
分子生物与遗传学	10243	2395	0.23
社会科学	1142	123	0.11
微生物学	4421	242	0.055
地学	4795	201	0.042
精神病学/心理学	3464	68	0.02
经济学与商学	3549	14	0.0039
空间科学	27851	3	0.00011

2.2. 统计预测方法

本文沿用“学科准入率 Z_i ”的概念,利用 Z_i 值确定本机构的潜力学科。在前期的工作中,对本机构各学科在 ESI 的各个更新时段进行跟踪统计分析,并在此基础上,将本机构潜力学科的总被引频次与 ESI 全球前 1% 学科门槛值进行对比分析,数据格式如表 1 所示,将时间线引入形成曲线,避免了单次数据或少量时间节点所形成的不明确趋势等缺点。将两条曲线利用最小二乘法形成线性关系,并通过 Excel 的线性预测功能形成本机构该学科的学科预测趋势。

2.3. 学科预测实例

根据 ESI 的 2016 年 3 月份的更新数据显示,本机构化学学科首次进入全球排名前 1%。本文研究者自 2014 年 5 月份开始到 2015 年 11 月份针对每次 ESI 的更新时段,对本机构化学学科的学科准入率 Z_i 进行了统计计算,同时计算公式中的 B_i 值与 C_i 值形成了两条趋势渐进的曲线,利用最小二乘法,可以近似看做一条直线,如图 1 虚线所示。由图 1 可以看出,本机构化学学科总被引频次(B_i)的增长速度高于 ESI 化学学科门槛值(C_i)的增长速度,两者截止 2015 年 11 月 20 日,ESI 化学学科门槛值(C_i)为 5652,青岛大学化学学科的总被引频次(B_i)为 5494,通过两者的比值可以得出化学学科的学科准入率 Z_i 为 0.97,接近于 1。通过 Excel 的线性预测功能可以看出在接下来的更新周期,学科准入率将超过 1,即青岛大学化学学科的被引频将超过 ESI 化学学科全球前 1% 的门槛值,青岛大学化学学科将进入全球前 1% 行列。由图 2 中可以明确看出,截止统计日期 2016 年 3 月 1 日,青岛大学化学学科被引频次已经超过门槛值,学科准入率达到 1.13。并于 2016 年 3 月的更新周期进入 ESI 全球排名前 1%。由此可以看出通过引入时间线,构成两条变化曲线,可以用来预测学科的情况。

3. 数据统计及分析

3.1. 数据统计与计算

第一步,选取神经科学与行为、材料学等五个学科,采集了 ESI 数据库自 2014 年 5 月至 2017 年 5 月期间十七个更新时段五个学科的门槛值;第二步,利用 ESI 数据库中期刊列表所构成的检索式,在 SCI-E

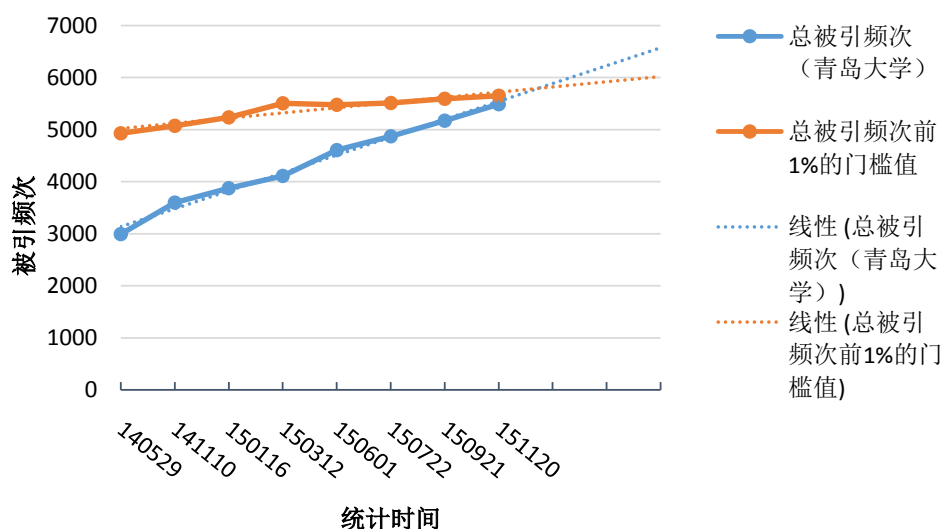


Figure 1. Comparison between the growth of threshold with citations of Chemistry in Qingdao University

图 1. 本机构化学学科总被引频次与门槛值增长对比图

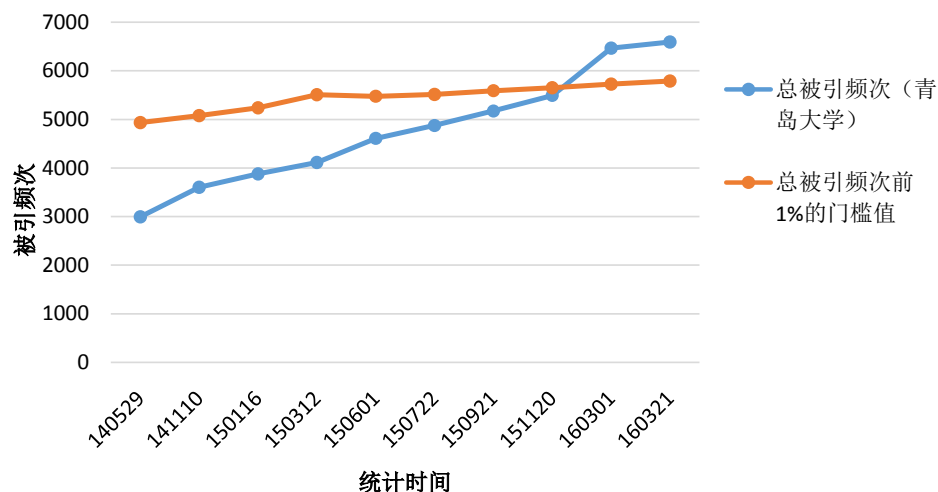


Figure 2. Comparison between the growth of threshold with citations of Chemistry in Qingdao University

图 2. 本机构化学学科总被引频次与阈值增长对比图

数据库中检索出五个学科的被引频次；第三步，利用学科准入率计算公式，计算出每个更新时段的 Z_i 值，构成曲线，其中 Z_i 值排名前两位神经科学与行为、材料学的总被引频次与阈值的对比图如图 3、图 4 中实线部分所示。

3.2. 趋势预测

图 3 中的神经科学与行为阈值随时间的变化曲线(a)与本机构该学科总被引频次随时间的变化曲线(b)，可以分别看做一条直线。因此可分别取一次函数 $y = kx + l$ (相对应阈值的变化曲线)与 $y' = k'x' + l'$ (相对应某机构的学科总被引频次的变化曲线)来近似的描述。利用最小二乘法，得到 a 拟合后的一次函数为 $y = 64.787x + 4030$ ，b 拟合后为 $y' = 200.83x' + 1170.8$ 。k 与 k' 分别为两个一次函数的斜率，a 的斜率为 64.787，b 的斜率为 200.83；斜率，亦称“角系数”，表示一条直线相对于横轴的倾斜程度。代入本应用中，可以分别看做阈值的增长速度，和某机构对应学科总被引频次的增长速度。由此可以看出，b 的增长速度高于 a 三倍多。

当 $y = y'$ 时，a 与 b 相交，通过前期的示例可以看出，当两条直线相交时可以看做为该学科进入 ESI 该学科全球前 1% 的节点。

Excel 中可以添加趋势预测同时可以选择不同的周期进行显示，如图 5 所示。

当趋势预测周期达到向前 4 个时，两条直线基本处于相交的位置。在统计数据中，均为 ESI 数据库更新周期(ESI 数据库每 2 个月更新一次)，而机构某学科总被引频次为通过 SCI-E 数据库(每日更新)统计获取，由于两者之间时间差造成的比值变化，神经科学与行为总被引频次将超过 ESI 前 1% 的阈值，这将预示本机构神经科学与行为一年后最有可能进入 ESI 全球排名前 1%。

如上同样的方法，分别得出材料学相对应的一次函数及斜率。由图 4 可以看出，材料学的被引频次增速高于阈值的增速。但是增速较缓，在第 6 个预测周期后，还是与阈值有一定的差距，这也预示材料学还需进一步提升论文质量，提高科研论文的影响力，为更快的进入 ESI 全球前 1% 做好更充分的准备。

4. 结论与建议

1) 由本文可以看出，通过计算学科准入率 Z_i 值，可以准确识别本机构的潜力学科。

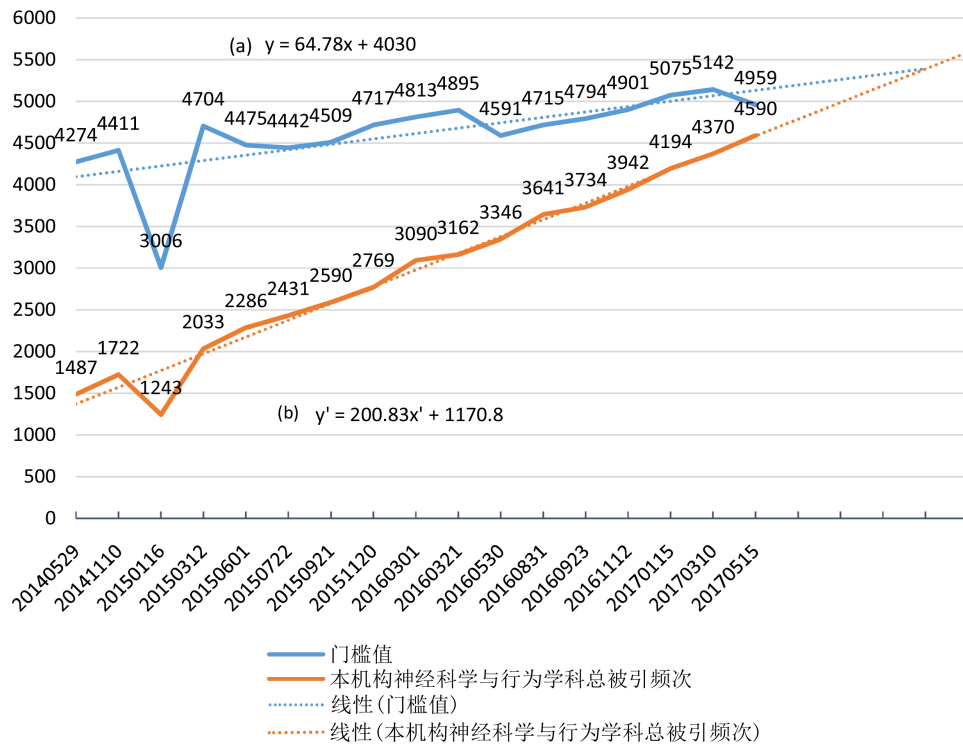


Figure 3. Trend forecasting and comparison between the growth of threshold with citations of neuroscience & behavior in Qingdao University

图 3. 神经科学与行为学科总被引频次与阈值对比及趋势预测图

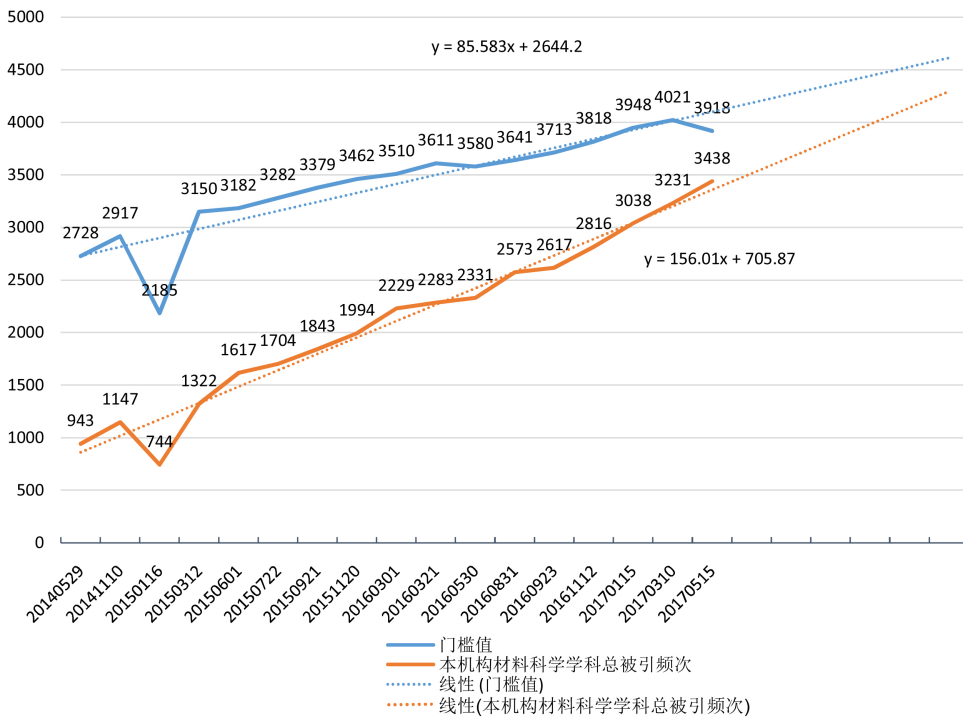


Figure 4. Trend forecasting and comparison between the growth of threshold with citations of material science in Qingdao University

图 4. 材料科学学科总被引频次与阈值对比及趋势预测图

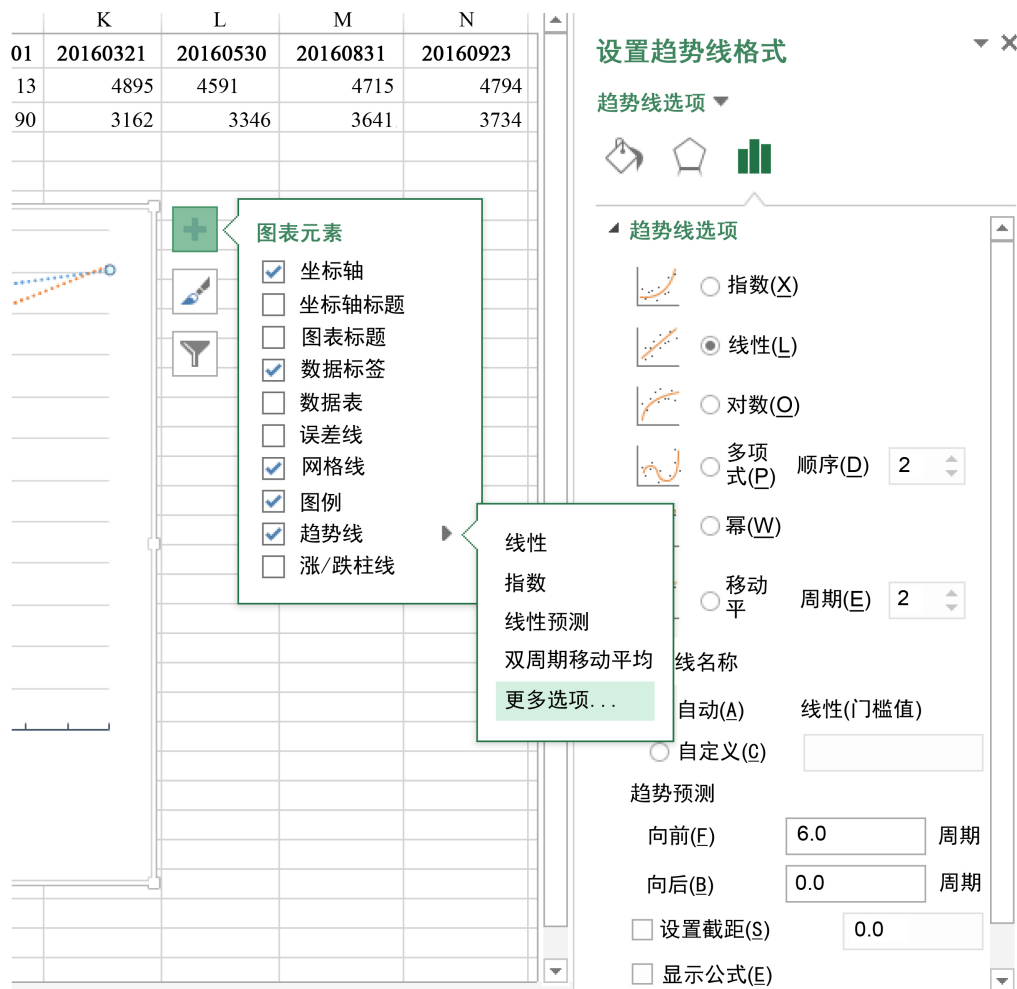


Figure 5. Schematic illustration of trend forecasting
图 5. 趋势预测示意图

2) 通过阈值与本机构潜力学科总被引频次的跟踪统计分析, 可以清晰的看到本机构潜力学科的发展势头, 明确该学科与 ESI 前 1% 对应学科的差距。

3) 通过拟合后的直线斜率可以很清楚的看到本机构潜力学科的被引频次增速, 以及对应学科阈值的增速, 从而可以很好的判断该学科是否有望计入 ESI 全球前 1%。有助于机构科研政策的制定, 多方位促进机构“双一流”建设。

参考文献

- [1] 国务院. 国务院关于印发统筹推进世界一流大学和一流学科建设总体方案的通知[EB/OL]. http://www.gov.cn/zhengce/content/2015-11/05/content_10269.htm, 2015-11-05.
- [2] 山东省人民政府. 山东省人民政府关于印发推进一流大学和一流学科建设方案的通知[EB/OL]. http://sdgb.shandong.gov.cn/art/2017/2/13/art_4563_5654.html, 2016-12-27.
- [3] 湖南省人民政府. 湖南省人民政府关于印发《湖南省全面推进一流大学与一流学科建设实施方案》的通知[EB/OL]. http://www.hunan.gov.cn/szf/hnzb/2017/2017nd4q/szfwj_98679/201703/t20170306_4701483.html, 2017-02-10.
- [4] 湖北省人民政府. 省人民政府关于推进一流大学和一流学科建设的实施意见[EB/OL]. http://gkml.hubei.gov.cn/auto5472/auto5473/201701/t20170106_936549.html, 2016-12-28.

-
- [5] 石河子大学. 关于印发《石河子大学一流学科建设实施方案》的通知[EB/OL].
<http://fgc.shzu.edu.cn/2017/0109/c1615a87065/page.htm>, 2016-12-14.
- [6] 董政娥, 陈惠兰. 图书馆基于 ESI 和 Incites 数据库支持高校学科科研评价的服务模式探讨[J]. 图书馆杂志, 2014, 33(11): 23-28.
- [7] 任勇, 林晖, 赵争光, 等. ESI 潜在学科预测评估研究-以安徽农业大学为例[J]. 数字图书馆论坛, 2016(7): 28-33.

知网检索的两种方式:

1. 打开知网页面 <http://kns.cnki.net/kns/brief/result.aspx?dbPrefix=WWJD>
下拉列表框选择: [ISSN], 输入期刊 ISSN: 2169-2556, 即可查询
2. 打开知网首页 <http://cnki.net/>
左侧“国际文献总库”进入, 输入文章标题, 即可查询

投稿请点击: <http://www.hanspub.org/Submission.aspx>
期刊邮箱: ass@hanspub.org