

文章编号:2095-0365(2012)03-0096-04

互联网环境下学术期刊防范 学术不端行为技术研究

谢文亮¹, 张宜军²

(1. 广东商学院学报编辑部, 广东 广州 510320;
2. 沈阳农业大学 出版部, 辽宁 沈阳 110161)

摘要:利用互联网环境,从技术的角度出发,对一稿多投、抄袭行为、剽窃行为以及捏造行为等学术不端行为的检测与防范进行研究,提出了各种操作上可行的方法。实践证明,运用这些方法,可有效地防范有学术不端行为的文章在学术期刊上发表。

关键词:学术期刊;学术不端;检测技术

中图分类号: **文献标识码:**A

一、引言

学术本来是一个崇高的名词,但由于各种原因,学术论文现在成为“抄袭”和“作假”的代名词,这是让人痛心疾首的事。学术期刊编辑部有责任和义务对发表在本刊上的文章作层层把关,防范有学术不端行为的文章的发表。

当前,期刊在学术不端行为的防范上一般从制度、管理、和技术上进行研究。文献[1]提出树立责任意识和品牌意识,杜绝人情稿、关系稿和学术泡沫,坚持“小同行”送审,发挥审稿专家在防范学术不端方面的作用;文献[2]提出了完善期刊的管理制度,强化期刊和编辑的连带责任;文献[3]出版后“随访”的方法,变被动为主动;文献[4-5]从技术的角度,提出用中国知网研发的学术不端检测系统来防范学术不端行为。在互联网还未普及的年代,要检测文章,需要花费很大的精力,资料和文献的寻找无从下手,检测结果也极不全面。随着互联网推广,各种各样的学术论文数据库的建立和发展壮大,大量文献资料的在线共享,以及学术不端行为检测技术的研究,使期刊编辑部能尽可能地检测出论文所潜在的学术不端行为。然而,学术不端行为检测系统也不是万能,还

要结合各种各样的检测手段,最大可能地检测出有学术不端行为的文章。

二、学术不端行为分类

(一)一稿多投

一稿多投是一篇文章同时往多个期刊投稿,造成两个或多个期刊同时发表一篇同样的文章,是一种恶劣的行为。其变种是,将题目和内容作了极少的修改,发表在不同的期刊,该行为也属一稿多投的范畴。

(二)抄袭行为

一字不少地抄袭是极罕见的,现在主要的抄袭行为是段抄和句抄。段抄,是将一篇文章中一整段文字照搬过来。由于互联网发达,信息资源丰富,不少作者写文章东复制一段,西复制一段,就构成了一篇论文。句抄与段抄本质上是一样,不过作者为了避免被人发现,在每段中间添加了一些自己的文字。

(三)剽窃行为

盗用别人思想、算法或结论。学术论文讲究

收稿日期:2012-08-28

作者简介:谢文亮(1979-),男,技术编辑,研究方向:编辑学。

创新和研究结论,如果没有研究成果或创新性的东西,该论文是没有发表意义的。经常有作者在没有任何研究成果的情况下,为了发表论文,盗用别人的思想或算法或结论,占为己有,在参考文献中刻意地回避引用参考文献,特别是英文参考文献。

(四)捏造行为

捏造行为指的是写论文时为了方便,没有经过考证,捏造出一些不存在或不准确的东西,如无效的文献引用、造假数据夸大研究成果等。学术论文是建立在别人研究的结果上,因此,优秀的学术论文必须有一定数量的参考文献。不少学术期刊规定必须有15篇以上的参考文献,对于没有作研究的作者,很难凑够15篇文献,因此只能随便复制几篇内容相似的参考文献,滥竽充数。在研究成果上,一般要有数据作佐证,才能证明所研究的成果比前人更进一步。有些作者在没有经过实践的情况下,通过总结前人的数据,对数据稍作修改,就当作研究成果写进论文中。

三、学术期刊反学术不端行为技术

学术期刊仅仅从制度和管理上反学术不端行为还不够,必须深入研究反学术不端行为的技术,作到“知己知彼,有的放矢”。

(一)一稿多投检测

一稿多投的检测方法有以下几种:

(1)利用当前强大的搜索引擎——百度和谷歌,特别是谷歌来检测。这是一种比较粗略的检测方法。如果作者对一稿多投的文章不作任何修改或只作细微的修改,那么该检测方法是很有有效的。其方法分为两个步骤:

第一,输入的搜索内容为:文章名+作者。如果检测出有同名作者和文章名,则可判定文章已经发表过,属一稿多投,检测结束;如果没有发现同名作者和文章名,则进入第二步检测。

第二,输入搜索内容为:随机抽取一整段文字。如果在百度和谷歌的浏览器上显示有80%以上的红色字,则可判定该文章有一稿多投的嫌疑。

(2)利用中国知网研发的学术不端文献检测系统检测。该系统的数据库包含了中国学术期刊网络出版总库、中国博士学位论文全文数据库、中

国优秀硕士学位论文全文数据库、中国重要会议论文全文数据库、中国重要报纸全文数据库、中国专利全文数据库、互联网资源和英文数据库(涵盖期刊、博硕、会议的英文数据以及德国 Springer、英国 Taylor&Francis 期刊数据库等),历年来中国大部分的学术期刊发表的论文,报纸上刊登的报道,申请的专利,和各大高校的研究生优秀毕业论文等都有收录。登陆学术不端文献检测系统检测帐号,上传所需要检测的论文后,系统自动检测,检测完后自动产生检测报告。如果文章是一稿多投,系统将其已经发表的刊物名称和发表时间等信息显示出来。

除了中国知网的学术不端文献检测系统之外,北京万方数据股份有限公司研发的论文相似性检测系统的检测功能也比较强大。随着中国知网和万方数据纷纷推出优先出版,一篇论文要经过两个系统的检测才能起到效果。

(3)社会监督方法。如果一稿多投的稿件,被一个期刊所录用,在文章尚未出版的情况下,已经在另一个期刊出版,但该期刊未及时将电子文档上传给各大数据库的情况下,单靠检测是不能得出结果。最好的办法是社会监督。当前,互联网在公开监督初显威力,本文第一作者进行的一项调查表明,国内的各大期刊中,尚未有一家期刊利用互联网对预发表的文章进行网络公示。具体实行起来所需要的技术并不复杂,有以下几种方法:一是在本期刊的官方网站上设立公示栏,栏目上在明显的位置上设立投诉电话,投诉 E-Mail,和投诉联系人,将下一期将要发表的文章作者名,文章名,文章摘要,文章关键字等信息在公示栏上公布出来。设置适当的公示日期,一般期限10天为宜。如果过了公示日期,未有投诉,即可正式发表。二是利用QQ群、微博、博客等定时发布预发表的文章信息,与国内各大期刊编辑信息互动。虽然互联网公示方法并不能做到百分之百发现一稿多稿,但随着各大期刊的信息共享程度越来越高,这种公示监督方法势将成为一种趋势。

(二)抄袭行为检测

抄袭行为比较容易检测,将整篇文章上传到中国知网的学术不端文献检测系统,检测结果中,会用不同颜色图例将结果显示出来。其颜色含义说明如表1所示。

表1 学术不端检测系统中颜色的意义

颜色抄袭性质	重合字数
青色无问题记录	无
黄色轻度抄袭	大于1000字或重合百分比小于40%
橙色中度抄袭	大于5000字或重合百分比在40%~50%
红色重度抄袭	大于10000字或重合百分比大于50%

当然,有时候单看颜色也不能说明作者就有抄袭行为,因为有时作者会引用别人的文字或是引用作者自己曾经写过的文章。这种情况下,学术不端文献检测系统提供了百分比数据供参考,一是总文字复制比,另一个是去除引用文字复制比。一般来说,去除引用文字复制比是重要的参考数据。另外,学术不端文献检测系统还提供了疑似问题文章与数据库中已有文章的内容对比参照,将有相同文字的文章用两个窗口一起摊开,相同文字的内容用红色显示,方便检测者判断。

如果是从网上的论坛或博客上面复制别人的文字而写成文章,这种情况下,论坛或博客上的内容一般是没有公开发表的,也就是说中国知网的数据库中是检测不到的,那么,就必须利用百度和谷歌两大搜索引擎来检测。其方法是,从文章中输入随机抽取一小段文字进行搜索。如果是该段文字是从网上复制过来的,浏览器上字体会显示红色。

(三)剽窃行为检测

盗用别人思想、算法、实验数据以及研究结论等剽窃行为比较隐蔽,不像抄袭那样可进行文本匹配检测。虽难以检测,但也是有规律可循。盗用的情形有两种,一种是盗用中文学术文献;另一种是直接翻译外国学术文献,将其作为自己的研究成果。检测方法如下:对该思想或算法或结论作的内容大致作个核心压缩,然后在百度、谷歌或学术不端检测系统作检测,如果有类似的结果出现,就证明了该结果是盗用别人的研究成果。

如果在中文检测不出结果,但该研究结果又是该论文的核心内容,则要考虑是否是直接从外文学术文献中翻译过来的。方法是将核心压缩出来的内容翻译成英文,语法和表达式并不重要,但对其中的关键词和数据翻译必须准确,翻译完成后,复制到谷歌(英文版)上搜索。该方法的准确率要看责任编辑的总结能力以及英

语翻译能力。

(四)捏造行为检测

参考文献的真实性检测对提高期刊文章质量是至关重要的,本着方便读者扩展阅读、对读者负责、防范不实参考文献对读者的误导,责任编辑有必要对文后参考文献作一个把关。能够对文章的全部参考文献的真实性进行检测是最理想的,但是一篇论文一般有15个以上的参考文献,一个个验证是需要很大的工作量。一般的方法是,从检测者的判断能力角度,将文献引用分为重要文献引用和一般文献引用。从重要文献引用中随机抽取1/3的参考文献,然后下载该文献电子版,进行验证。

随着互联网和数据库技术的发展,国内外在参考文献管理系统的研究上也取得一定的成果,研发了专业的参考文献检测系统。国际上知名的参考文献管理软件包括Endnote(SCI汤姆森路透)、Mendeley,国内有单机版NoteExpress和西安三才研发的网络版NoteFirst。将参考文献输入检测系统,系统自动将参考文献与权威数据库进行数据比较,逐条验证数据的准确性,并提供修改意见。

在数据造假方面,传统的方法是选择该学科的一个专家,将文章送审,让学科专家验证数据真实性。互联网环境下,可以在百度的“知道”栏目上发布提问,或在博客、微博、论坛上发起讨论,经一番交锋,集众人之智慧,数据的真实性就自然水落石出。

四、结语

实践经验表明,运用以上的方法,对于有学术不端行为的文章,检测准确率达90%以上,有效地防范有学术不端行为的文章在学术期刊上发表。

随着互联网技术的进一步发展,云计算能力的提高,数据库功能的增加,资源共享的扩大,检测技术也会继续向前推进。作为学术期刊的责任编辑,必须学会利用互联网、网络搜索技术和数据库技术,结合在工作中的实践经验,研究和总结更多的检测方法,在审稿过程中牢牢把关,尽力避免录用有学术不端行为的稿件,以体现学术期刊遵守学术道德的使命。

参考文献:

- [1]李新根,徐用吉. 学术期刊编辑如何防范学术不端行为[J]. 科技与出版,2010(8):64-66.
- [2]刘英. 学术期刊对学术不端行为的防治对策——以学术出版者控制职能为视角[J]. 编辑之友,2011(4):48-50.
- [3]张儒祥,丁茂平. 医学期刊责任编辑在防止学术不端论文出版中的作为[J]. 编辑学报 2010(6):519-520.
- [4]李刚,陈京香,王健梅,等. 学术不端行为分析及编辑部预防对策措施[J]. 编辑之友 2011(8):77-78.
- [5]吴均,江润林,张晓琴. 利用学术不端检测系统研究科技论文中存在的问题[J]中国科技期刊研究:2010,21(5):636-639.

Research on Techniques for Journal in Preventing

Academic Misconduct in the Internet Era

XIE Wen-liang¹, ZHANG Yi-jun²

(1. Editorial Office of Journal of Guangdong University of Business Studies, Guangzhou 510320, Guangdong Province, China;

2. Publishing Department, Shenyang Agricultural University, Shenyang 110161, Liaoning Province, China)

Abstract: From aspects of technologies in the Internet environment, this article explores the possible ways of anti academic journal misconduct, such as the same paper contributed for more than one journal, copying, theft and counterfeit behavior, etc. Some suggestions were given for preventing academic misconduct at journal. Experimental result shows that by adopting these methods, papers with academic misconduct can be effectively prevented from being published.

Key words: academic journal; academic misconduct; detection technologies.

(责任编辑 杨继成)

(上接第 60 页)

参考文献:

- [1]自媒体时代[EB/OL]. <http://baike.baidu.com/view/7029241.htm>.
- [2][俄]巴赫金. 诗学与访谈 [M]. 白春仁,顾亚玲,译. 石家庄:河北教育出版社,1998.
- [3][俄]巴赫金. 巴赫金全集·拉伯雷的创作与中世纪和文艺复兴时期的民间文化[M]. 李兆林,夏忠宪,译. 石家庄:河北教育出版社,1998.
- [4][美]威廉·J. 米切尔. 比特之城:空间. 场所. 信息高速公路[M]. 范海燕,译. 上海:三联书店,1999.
- [5][俄]巴赫金. 巴赫金文选[M]. 北京:中国社会科学出版社,1996.
- [6][日]北冈诚司. 巴赫金:对话与狂欢[M]. 魏炫,译. 石家庄:河北教育出版社,2002.
- [7][俄]巴赫金. 陀思妥耶夫斯基诗学问题[M]. 上海:三联书店,1988:178.

Study on "Self-media era" from the Perspective of Bakhtin's Carnavalesque Theory

WANG Qi-yu

(Hebei Normal University, Shijiazhuang 050024, China)

Abstract: Thanks to the rapid development of IT, traditional mass media are replaced by new high-technological media like Internet. With the diversification of media forms and the new communicating paradigm, in the new high-technological media with the characteristic of massive, interactive and trans-time-and-space information, the "person" himself becomes the main subject of spreading and individual expression is fully reflected. In addition, currently, "grassroot" culture stands out on the internet and blossoms to a popular carnival feast. At this point, the interaction between people, society and the world has entered a new era—"self-media era", an era to express self-purpose through the new media. We try to expand and develop Bakhtin's theory, and discusses this social focus grass-roots carnival under the background of "self-media era".

Key words: Bakhtin; carnivalesque theory; self-media era; popular culture

(责任编辑 王丽娟)