

刻意规避学术不端软件检测论文的再审查

朱 银 周

《昆明理工大学学报(自然科学版)》编辑部,650500,昆明

摘 要 由于多种原因,学术不端行为检测软件存在自身的不足,只能作为纯文字复制检测工具,对纯文字复制行为做出分析界定,不能从根本上遏制所有学术不端行为。与学术期刊相关的主体人员不能过度依赖学术不端行为检测软件,还应通过多种人工方式判别深度的学术不端行为,才能最大限度杜绝学术不端行为的隐式泛滥。深度学术不端行为已成为近来的一种发展趋势,应该成为预防的重点。

关键词 学术论文;学术不端行为;学术不端文献检测系统(AMLC);学术期刊

Recheck of academic papers checked by academic misconduct literature check system//ZHU Yinzhou

Abstract Due to various reasons, Academic Misconduct Literature Check System (AMLC) has its own shortcomings; it can therefore only be used to detect and define pure text copy, but cannot be used to prevent all academic misconduct fundamentally. Personnel involved in academic journals should not rely solely on AMLC software, but also on various manual methods to detect academic misconduct of deeper levels and to avoid such implicit misconduct which has recently become a trend.

Keywords academic paper; academic misconduct; AMLC; academic journal

Author's address Editorial Department of Journal of Kunming University Science and Technology, 650500, Kunming, China

学术不端行为检测软件的使用,对论文中的纯文字复制行为的发现十分有效,在一定程度上遏制了初级的学术不端行为,为学术期刊编辑人员甄别学术不端行为的类别有重要的价值;但学术不端行为检测软件不是万能的,会发生误检测的情况^[1],能发现的不端行为也主要局限在文字复制方面,需要正确认识其功能。正因为检测软件存在自身的缺陷,就会被有的人所利用,经过刻意伪装的学术不端行为无法被检测软件识别,而过度依赖学术不端行为检测软件的编辑容易放过存在严重学术不端行为的“论文”,致使学术不端行为继续泛滥。

编辑不能仅仅通过软件检测来稿,复制比没有超过规定红线就鉴定为学术论文不存在学术不端行为,还应承担起更多的社会责任^[2]。实际上,检测结果仅仅只能作为参考^[3],还应该再经过严格的学术不端行为人工判断^[4],查找是否存在深度的学术不端行为,才算相对安全。赵蔚亦指出系统误判和漏判的情况^[5];赵茜认为在使用学术不端行为检测系统时要把握合理的“度”,建议人工和检测系统并用^[4];吉家友认为编辑不能简单、机

械地以软件数据和重复率高低来决定文章能否录用^[6];倪爱香认为应当用中国知网开发的“学术不端文献检测系统”和万方数据开发的“论文相似性检测系统”同时并查^[7];庞海波针对学术不端行为的认识误区提出防范策略^[8];张宏等^[9]和尤伟杰等^[10]提出通过用论文数据库检索、网络搜索引擎、防剽窃系统角度挖掘学术不端行为。笔者认为,编辑应该承担起主要责任,运用多种方法判别和剔除学术不端行为。

笔者依据自身使用学术不端文献检测系统的经验,列举学术不端行为检测软件的一些缺陷,以及针对这些缺陷将深度隐藏在稿件中的学术不端行为,通过人工判断方法来予以辨识。

1 学术不端行为检测软件的缺陷

中国知网编制的《学术期刊论文不端行为的界定标准》对剽窃、伪造、篡改、不当署名、一稿多投、重复发表和拆分发表等7类学术不端行为进行了细分和界定,而由其开发的学术不端行为检测软件正是以《学术期刊论文不端行为的界定标准》为基础,仅针对其中的观点剽窃、自我剽窃、一稿多投、过度引用、整体剽窃、重复发表和剽窃文字表述等学术不端行为进行检测。

除了以上7种学术不端行为以外,伪造、篡改、不当署名、剽窃图像、剽窃实验方法和剽窃未发表成果等亦属于不端学术行为之列,学术不端行为检测软件对此是无能为力的。更为可怕的是针对检测软件的自身缺陷而写成的“学术论文”,其中隐藏了深层次的学术不端行为而蒙蔽了检测软件,逃避了打击。检测软件功能仅仅限于对单语句与数据库比对后依据文字复制比例与位置做出分析与判断。

学术不端行为检测软件是存在缺陷的,这主要受制于计算机软件开发技术。软件开发者的目的主要甄别出机械式的抄袭和拼凑行为,提供各个学术期刊编辑部和各个大中专院校毕业论文的简单检测。至于更深层次的隐藏着学术不端行为的“学术论文”的检测,就需要人工来更进一步地甄别与认定。

2 学术不端行为的表现方式

2.1 文字变图表:围栏保护式抄袭 学术不端行为检测软件对图表几乎无法检测,这主要受制于计算机技

术及图表格式的多样化;但正因为存在这样的技术盲区,就容易被某些别有用心的人利用:对于存在学术不端行为的主要内容通过改写,变为表格或者图形方式进行表述,为存在学术不端行为的内容加以保护,形成了针对检测软件不可逾越的“围栏”。这样,学术不端行为检测软件就因为“无能为力”而放过存在一定学术不端行为的内容。这种行为多被在校的硕士研究生和博士研究生所采用。特别是图形,由于其格式多种多样,检测软件识别更加困难。

2.2 文字表达方式改写:意思式抄袭 众所周知,独立撰写一篇有一定学术价值的论文费时费力,需要长时间的技术和知识累积,需要付出艰辛的努力。存在于论文数据库中的学术论文已经经过很多人的订正、修改与学术价值认证,各种错误已经降低到最低程度。数据库中同类的论文就是改写和拼接的目标论文的模板,把已经发表的论文进行意思改写和拼接,就会显得事半功倍。对学术论文各种潜在利益的追逐,导致各行各业部分作者追求短期利益,就不得不追寻这条简单规则下的依葫芦画瓢与移花接木模式,成为短期内多产论文的作者。

表达意思相同,表述方式却千差万别,这就是针对检测软件的缺陷而提出的最简单的策略。检测软件是基于论文数据库为基准的语句比对式算法原理,因为比对论文文字没有雷同而认为不存在学术不端行为。这种方式最简单但又普遍存在于学术论文中。有时我们将同类论文放在一起比对,总有“千文一面”的雷同感觉,创新度严重不足。

2.3 外文译为中文:变脸式抄袭 游走在中文论文数据库之外还有大量的外文数据库,特别是小语种论文数据库,国内接触的人比较少,往往成为一个无法触及的死角。少数作者凭借自身熟悉外文和相关专业的双重优势,剽窃外文论文,翻译为中文,并加以改头换面,重新拼接,成为一篇自己的“论文”。这种模式在在读研究生中的比例较高,归国的留学生中也有一定的比例。短平快的方式就可以通过论文的语种变换变为自己的学术成果,获得论文背后潜在的利益。这种语言变换并重新拼接的“论文”,学术不端行为检测软件更是无能为力,只有通过专业人士比对才可以发现。

2.4 创新观点的变换表达:观点剽窃抄袭 创新是一篇学术论文的灵魂,这是原作者长期研究的结果,也正是编辑和审稿专家取舍一篇学术论文的主要依据;但对创新点的来源往往又难以追溯,就容易对学术论文造成误判。正因为如此,部分作者会剽窃原作者的局部创新点而又不标引,加以分割,变换描述方式,适当扩展,变成了另一篇文章。这种现象十分常见。这类

学术不端行为检测软件是无法发现的,更多的则依靠编辑和审稿专家来判定。

2.5 对老旧文献、专著和各类报告的剽窃:推陈出新式抄袭 学术不端行为检测软件是以各类数据库为基础开发的,只要数据库中不包含的内容,被检测的目标论文因为没有比对的“标准样本”而被视为安全。鉴于20世纪90年代以前所刊发的论文,部分没有进入论文数据库,也就无法成为学术不端行为检测软件比对的原始论文,对这类“论文”的学术不端行为,检测软件无法查出来。还有,大部分专著因涉及版权和其他问题而没有进入比对数据库,无法成为比对样板数据。还有各类无法进入数据库的报告和汇编材料,依然无法成为比对的原始样板。此外,还有各类统计资料,也没有进入学术不端行为检测软件的比对数据库,对其抄袭的内容也无法检测出来。

2.6 公式罗列:容易被忽略的抄袭 在一些基础类学科论文中,主要公式抄袭现象十分严重,很多并不指明出处,无法分辨公式的隶属关系。检测软件是无法识别数学公式的,这在于数学公式的复杂性:其中包含了多种运算符号、多级层次关系、多种字符,而且很多公式是用公式编辑器或者其他专业编辑软件编辑的,来源就很复杂,涉及兼容性问题,检测软件很难识别。对公式的学术不端行为的检测因为技术障碍不得不放弃,特别是以数学公式推理、论证为主的论文,几乎就无法检测到其中的学术不端行为;但这类论文中不等于不存在学术不端行为,恰恰相反,学术不端行为深藏其中。在这类学科领域要有所突破相当困难,合理引用已经成为这类论文首当其冲要解决的问题。

3 学术不端行为把关的主体人员及其职责

3.1 编辑和学生管理人员:多种方法共用识别学术不端行为 学术不端行为检测软件开发者的初衷是为学术期刊编辑和高校研究生学位论文的管理者提供一种学术不端行为检测工具。当部分作者深知这种工具软件的硬伤,加之学术期刊编辑和高校学生管理者过度依赖学术不端行为检测软件时,隐藏式的深层学术不端行为便更难发现,对学术论文背后利益的过度追逐,将会有学术不端行为继续泛滥的危险。

作为学生学位论文管理者,不能拒绝使用学术不端行为检测软件,更不能迷信学术不端行为检测软件。实际上,使用学术不端行为检测软件对学术论文进行检测,仅仅是学术论文学术不端检测的开始,而后续的甄别学术不端的进程却更为重要;但现实的情况是学位论文普遍偏长,学生人数众多,要想对深度的学术不端行为做出准确判断费时费力,打击学位论文中的学

术不端行为显得力不从心。很多学位论文带病又进入学位论文数据库,成为后续检测论文的标准库。

编辑不能将对所有的对学术论文的学术价值的判断,全部都交给审稿专家,特别是对论文是否存在隐藏式的学术不端的识别和判断,编辑应该承担起前期的责任。通过稿件的关键词和学科方向可以查询现有的论文数据库,找出同类和近似的论文,做一定的对比后,可以简单地进一步判断论文是否存在隐藏的学术不端行为。最直接可以利用的论文数据库,如中国知网、校园网和研究所的论文数据库。这里不得不提的是各个大中专院校图书馆的论文数据库,已经集成了国内外很多知名的各学科论文数据库,有的甚至达到100多个数据库,通过学科或关键词查询和搜索的方式可以找到相类似的论文,判断论文的深层次学术不端行为和学术价值,方便快捷。同时,在查询论文数据库的同时,编辑可将自己无法判断或已经发现而又无法准确界定的学术问题提交给审稿专家进一步甄别。这种方法简单、实用。此外,编辑还应从英文关键词的角度查询论文是否涉及国外英文论文的学术不端行为。

3.2 审稿专家:从学术创新角度遏制学术不端行为

审稿专家是学术论文价值最主要的判定者,更多是从创新角度判断学术论文的价值。学科前沿的审稿专家总能从学术价值角度对论文进行取舍,无论论文存在的学术不端行为以何种面目伪装,总难逃过专家的法眼;因此,寻找学科前沿的领军审稿专家尤为重要。编辑有必要完善审稿表,在审稿表中增加“是否存在学术不端行为”的专家填写项目,栏目之间要有内在的逻辑关联,填完相关项目就已经最大程度地剔除了深度学术不端行为。同时要求专家务必通过查询论文数据库对所评审的论文的学术价值做进一步的佐证,对论文创新性进行查新确认,避免仅凭个人印象和知识容量做出粗略判断就下录用与否的结论;因为有些学科知识更新很快,动态变化中的知识容易造成审稿专家误判。这样做对创新观点抄袭和剽窃更为直接和有效,也就能准确判断论文的学术价值。

3.3 编委参与的定稿会:对学术不端行为把关的最后一道闸门 编委成员大多本身就是某个学科的领军人物,编委参与的定稿会对学术不端行为的进一步发现与剔除起着至关重要的作用。无论采取何种方式的定稿会,都应将对应的稿件交给对应的编委从学术不端行为和创新价值方面进行进一步审查,编委将会从期刊发展和学科专家的角度对论文做出客观评价,不能省去这道流程。定稿会是编辑和编委共同参与的会议,是一种共同交流的会议,对发现论文中的学术不端

行为体现出一种集体智慧。

3.4 读者反馈机制的建立:最有效的发现学术不端行为的补救措施 学术论文发表之后,往往意味着一篇学术论文的编辑出版阶段完成,也就很少有编辑再去追究其学术不端行为;但读者例外,刊发后的论文,读者才是判别其价值的第一人。读者在利用文献的过程中,会将同类论文做比较,往往会发现论文中的很多学术不端行为,甚至严重到观点雷同、数据一致,仅仅是表达方式和顺序不一样而已。

这就提醒各期刊编辑部,有必要建立一套完善的读者反馈机制。可以借鉴网络论坛的某些做法,健全评价和举报机制。在论文发表之后,在网络版每篇论文题名之后或文末附上读者对“该文学术价值评价”和“学术不端行为举报”选项,将结果存储到采编系统中,前台采用屏蔽模式,定期从后台收集读者意见。这样,编辑能从评价和举报信息中得到学术价值的真实评判,也能知晓发表的论文是否存在深度的学术不端行为。同时,可以从读者评价中获得学科发展方向的热点、焦点问题,为将来选择同类学科稿件提供导向。对于存在学术不端行为较轻的论文,在论文数据库中做出明确标记,建议后续读者慎重对待这篇文章。对于发现有严重学术不端行为的论文,就需要做出公开“撤销发表”的处理。

4 参考文献

- [1] 谢文亮,李俊吉,张宜军. 期刊学术不端文献检测系统误检分析[J]. 中国科技期刊研究,2013,24(6):1126
- [2] 游苏宁,陈浩元. 科技学术期刊编辑应承担更多的社会责任[J]. 编辑学报,2006,18(2):81
- [3] 谭华,崔洁. 学术不端文献检测系统的使用建议[J]. 编辑学报,2010,22(2):153
- [4] 赵茜. 使用学术不端文献检测系统如何把握好合理使用的“度”[J]. 中国出版,2010(10):59
- [5] 赵蔚. 学术不端检测结果的修正标准初探:基于“文字复制比”与“文章抄袭率”的辨析[J]. 中国出版,2011(22):20
- [6] 吉家友. 学术不端文献检测系统数据分析[J]. 中国出版,2010(4):27
- [7] 倪爱香. 学术期刊编辑对学术不端行为的防范策略[J]. 中国科技期刊研究,2013,24(3):586
- [8] 庞海波. 科技学术期刊编辑对学术不端行为的认识误区与防范策略[J]. 编辑学报,2011,23(2):103
- [9] 张宏,程建霞,王小唯,等. 学术不端现象分析及期刊编辑应对策略[J]. 编辑学报,2010,22(1):52
- [10] 尤伟杰,李小萍. 学术不端论文的一般规律及应对策略[J]. 编辑学报,2009,21(2):151

(2014-08-28 收稿;2014-09-16 修回)