

关于高校学位论文查重制度的反思

——从学术伦理教育的角度看

柴伟佳

桂林医科大学 广西壮族自治区桂林市 541199

摘要: 目前中国高校的学位论文查重制度并不能有效防止学术不端,也不利于对学生的学术伦理教育。为了进一步促进学术伦理建设,建议各高校在学位论文撰写前就指导学生如何正确地进行学术引用,而在对已完成论文进行查重时需要区分不同的专业与研究领域来设置重复率的限制,在计算文字重复率时需要区分不同类型的文本,并应该增加人工审核程序来避免合格的学位论文不能通过查重检测以及不合格的论文可以通过查重检测两种情况。

关键词: 学术伦理;学位论文;论文查重;知网;教育;

1 前言

自从以“中国知网”为代表的各种论文数据库平台开始提供论文查重服务后,越来越多的高校也选择与知网等平台合作,使用这些平台对本校毕业生的论文进行查重,希望可以避免传统的人工审核不完善的问题。理论上讲,通过计算机审核可以迅速和有效地发现论文中的学术不端情况,有效减少学位论文的学术不端,完善高校的学术伦理建设。但是,事实真的如此吗?

2. 当前学位论文查重系统的运行模式——以“知网”为例

2.1 查重系统与算法介绍

根据“中国知网”官网对其学术不端检测系统^[1]的介绍,该系统的服务内容包括三方面:文献抄袭检测,文献保护,文献比对。那么该系统是否可以直接断定某论文是否出现了学术不端呢?知网对此没有给出直接的回答,只是说该功能可“可快速、准确、高效的检测文献中的文字复制情况。为发现抄袭与剽窃、伪造、篡改、不当署名、一稿多投等学术不端行为提供科学、准确的线索和依据。”

在查重系统的“免责声明”^[2]中,知网明确表示:“检测系统根据您上传的文献自动生成检测结果,不代表中国知网赞成检测结果或立场”。在关于查重系统的“常见问题解答”^[3]中,关于重复率多少才算抄袭或者学术不端,知网的回答是“系统提供的百分比只是描述检测文献中重合文字所占的比例大小程度,并非给定特定检测文献的抄袭严重程

度。是否认定为抄袭等,需要用户根据具体文献的内容及考核的侧重点和考核标准进行判定。”也就是说知网明确提出该系统检测出来的重复率相关数据并不是最终判断,最终还需要用户人工进行判断。

知网之所以会有这种自己只提供数据,但不做关于抄袭的认定的言论,当然是有客观原因的。一是本身查重系统存在一定的局限性。它通常不能把抄袭的内容与仅仅是重复的内容完全区分开。比如说,知网自己也承认它的查重系统也会将论文的致谢内容纳入文字比对范围,而除了致谢内容外,一些常识性内容、经典诗词谚语、常见的一些法律法规政策文件、一些政府领导人的公开讲话等都是许多论文会共同使用的,而且有时出于巧合不同作者在某些内容方面的表述比较类似。然后论文写作本身也常常需要原文引用一些其他文献的内容,这些内容都会被算入重复率,但是它们本质上都不是抄袭或学术不端。二是,关于什么是学术不端、什么是抄袭或剽窃,目前学术界本身并没有一个统一认可的标准,因此知网自己也不可能提出按照某一套主观的标准来进行认定,只能把这个涉及价值判断的问题留给用户自己。

2.2 高校学位论文查重制度介绍

目前大多数高校对学位论文进行查重时会规定,本科生的学位论文文字重复率不能高于某个百分比,硕士或博士学位论文的文字重复率不能高于某个百分比,然后每个同学大概有两次进行论文查重的机会,如果两次查重都不能通过,就不能参加学位论文答辩,因此也就不能顺利拿到学位。通

常,各个高校并不会区分不同的专业,根据不同的专业或学科的特点来设定重复率百分比。同时,这个重复率百分比通常是指总百分比,也就是所有重复文字的比例,而不区分是经过引用的重复文字还是未经引用的重复文字。

3. 当前的学位论文查重工作中出现的问题

3.1 学术不端标准被错误地等同于重复率标准

由于各个高校在学术规范的要求方面只是规定了论文的文字重复率的百分比,而几乎没有其他的要求,这就导致很多学生将文字重复率等同于学术伦理的唯一标准。很多学生误以为只要查重通过了,论文就不存在学术规范问题,而如果查重没有过,就必然存在学术规范问题。或者认为一篇论文的文字重复率越低,文章的质量就越高。但是这些观点其实都是不正确的。

首先,学术不端的类型有很多种,剽窃只是其中的一种类型,并非不存在剽窃就不存在学术不端。而抄袭也只是剽窃的其中一种类型,并非不存在抄袭就不算是剽窃。因为如果使用了他人的观点、创意或某个想法,却没有使用他人的文字,这种情况下也算剽窃却不算是通常意义上的抄袭。其次,重复的文字未必都是抄袭,因为只要是根据学术规范进行了正确引用的文字就不算是抄袭。而高校规定的论文百分比并不区分是正确引用的文字还是未经引用的文字,就会导致那些合乎学术规范的文字也被看作是抄袭的文字而被纳入到重复率计算中,这也就使得学生不能了解什么是正确的引用规范。最后,高校所规定的文字重复率一刀切,未考虑不同学科、不同的主题与不同类型的论文的特点,全部按照同一个标准衡量,不具有合理性。所以论文的重复率高低与其论文的质量或学术不端之间没有必然联系。并非论文重复率越低论文的质量就越高。

3.2 查重系统的学术不端检测可被轻松欺骗

查重系统初衷是进行学术不端的检测,但是现实生活中,论文的重复率往往并不能真正反映论文的学术不端的情况或严重的程度。就算大多数高校能够改革查重制度,对进行了正确引用的文字与未正确引用文字进行区分,还存在着其他的问题。

现实生活中,学生们已经掌握了一些欺骗或应付查重系统的办法,或者可以通过购买淘宝等网上商家的服务对其论文进行特定的“修改”或“编辑”来人为地降低重复率,来逃过查重系统的检测。常见的办法有:抄袭那些在查重系

统数据库之外的文献(一些外文文献、书籍、网络文章、线下会议论文等)的文字、将某些文字的语法或字词进行同义替换、将文字转为图表的形式等。这些办法可以让那些本来不合格的论文轻易逃过查重系统的检测,也间接地促使某些学生学会了更多进行抄袭的方法。同时也让那些担心查重不能通过而主动提前去查重的同学不必要地耗费了大量时间和精力与金钱。

3.3 重复率设置不合理导致学位论文质量下降

学术规范的建设的目标是促进学术的发展。但是目前反而本末倒置了。因为学生为了达到学校所设置的重复率标准,可能将本应该引用或撰写的文字删掉,导致文章内容不完整。或者为了避免文字相似而故意修改文字导致难以理解甚至出现语法错误,将文字转为图片或图表导致阅读困难,篇幅增加。或者在写作时尽量自己想自己写而未能充分使用已有的前人的研究成果。或者,使用了他人的成果却没有标注,导致读者无法找到原始文献。甚至有的学生会撰写两个版本的学位论文,一个作为查重用,一个作为最终答辩用,也有学生为了降低查重率主动找第三方机构进行检测,而这种行为又可能导致论文被平台盗用,如果是涉密论文则后果更为严重。

4 解决办法与建议

4.1 将人工审核而非查重结果作为最终标准

通过上文对知网的“学术不端检测系统”的介绍,我们可以看得出来,知网本身并没有声称该系统检测结果可以直接给出是否存在抄袭或学术不端的结论,而是需要用户本身根据检测结果并结合其他信息与规范来进行判断。所以直接以查重得到的重复率来判断学生论文是否存在抄袭或其他问题是一种非常简单粗暴极易出错的做法,正确的做法应该是由指导老师以及其他相关工作人员进行人工审核进行最终的判断。

4.2 科学和有针对性地设置重复率相关要求

如果高校对于论文的文字的重复率必须要设置一个标准,也不应该一刀切,而应该科学和有针对性。第一,重复率相关要求应该区分有经过引用的文字的重复率与未经引用的文字的重复率。同时,也要区分论文里直接原文引用的文字与间接引用的文字,所有原文引用的文字不应该纳入到重复率的计算中。但如果是非原文引用或间接引用,即只是在脚注、尾注或参考文献里注明了所引用的文献,但是文本

的表述方式与原文献没有实质性的不同，这也算是学术不端（这可能也是知网为何将所有间接引用部分也纳入重复率计算中的一个原因），所以这种间接引用的文字理论上应该是纳入重复率的计算中。第二，应该针对不同的专业与学科的特点来有针对性的设置不同的重复率标准。第三，如果是文献综述类的论文，或者是论文中的文献综述部分，不应该设置与其他论文或文本同等的重复率标准。第四，应该针对不同的学历的学位论文设置不同的重复率标准，尤其是针对本科与专科学位论文不应设置过高的重复率标准，甚至应该取消文字重复率的要求。因为根据教育部的规定，本科生的论文本身不需要具有创新性。既然本科生的学位论文本身创新性不大，那么其论文的内容必然就会有很大一部分是对已有研究的引用，所以其总文字重复率比较高是非常正常的现象。

4.3 取消某些专业不必要的学位论文考核

在国内，大多数本科生的专业与几乎所有的研究生专业都采取毕业论文的方式进行最终考核。但是其实不一定所有专业都适合以论文形式进行毕业考核。很多人误以为所有本科专业都应该以毕业论文的形式来考核学生，其实《中华人民共和国学位条例》规定本科毕业生只要成绩优良，达到一定的学术水平者即可授予学士学位，并没有要求所有本科毕业生都要写学位论文。

21 世纪教育研究院副院长熊丙奇提出^[4]，在发达国家，高校对本科毕业生与应用型的硕士研究生通常并不要求撰写学位论文，只需通过规定的课程拿到学分即可毕业。而国内要求几乎所有的本科生与硕士生写论文可能会有不良后果。很多学生对于毕业论文基本上都是简单应付一下，甚至去造假。因此，我们应该取消一部分的本科甚至研究生专业的不必要的学位论文考核，而采取其他方式进行毕业考核。

4.4 建立长效的学术伦理教育与管理体制

大多数高校只在毕业论文开题时才对学生简单介绍一下甚至没有介绍有关学术规范的内容，所以学生通常在写论文的时候通常根本不了解如何合乎学术规范进行撰写。因此，高校应该在学生毕业论文开题前（最好是刚刚上大学时）

就做好学术伦理规范（而不仅仅是格式规范）的教育，例如在平时的课堂论文考核中就要重视学术伦理规范。避免学生由于不懂学术规范或者不懂学术引用的格式所引起的重复率过高，否则真正到了论文撰写和答辩的阶段就因为时间有限等原因难以改正了。

学者顾曼^[5]认为，查重所出现的问题根本原因是高校未能在真正落实“严出”，他提出“如果大学严进严出不再是口号，而是引入社会评价体系，让升学率、毕业率不再成为高校制定政策、教师指导学生时的顾虑，高校更多地体现对社会负责，培养的学生未来发展好、用人单位喜欢，不仅毕业论文的查重问题，我觉得现阶段很多教学问题都能得到解决”。

5 总结

学位论文查重乱象反映了目前高校在学术规范与学术论文教育方面的很多问题。我们真正需要的是建立一个学术规范的长远机制。假设高校与用人单位对学术伦理高度重视，任何有学术不端的人都会被取消聘用或学位，还会有人这样做吗？论文抄袭貌似是写作者个人的品德问题，实则反映的却是社会对学术不端或不诚信问题的轻视。只有一个真正珍视和鼓励诚信的社会和高校才能在根本上杜绝学术不端。

参考文献

- [1] 中国知网. 关于学术不端检测系统 [EB/OL]. <https://check.cnki.net/Article/about/2009/03/25.html>, 2009-03-25
- [2] 中国知网: 免责声明 [EB/OL]. <https://check.cnki.net/Article/about/2010/05/142.html>, 2009-03-25
- [3] 中国知网: 常见问题解答 [EB/OL]. <https://check.cnki.net/Article/faq/Index.html>, 2009-03-25
- [4] 福州高校某系回应毕业论文查重“几乎都不达标”: 从严要求 [EB/OL]. https://www.thepaper.cn/newsDetail_forward_1658138, 2017-04-09
- [5] 《中国科学报》(2019-07-24 第 1 版 要闻) [EB/OL]. <https://news.sciencenet.cn/htmlnews/2019/7/428656.shtml>, 2019-07-24