·现代教育技术·

大数据环境下高校学生的心理预警机制构建

何小波　陈方方
重庆医科大学信息中心　400016
通信作者：陈方方，Email：494182371@qq.com

【摘要】 将高校学生心理“大数据”与心理专家库数据结合，利用数据挖掘算法提取特征属性，从稳定的算法模型中得出预警结果。此方法构建的心理预警平台，不仅能实时监测学生心理动态，而且能预测其心理行为趋势，弥补传统预警方式的滞后性及准确率低的不足。心理预警平台提供数据支撑，心理预警制度提供运行保障，两者结合形成的心理预警机制，能够为教育工作者提供科学依据，最终达到利用大数据技术提高高校学生心理健康的目标。

【关键词】 大数据；　行为数据；　数据挖掘；　心理预警
【中图分类号】 R33
基金项目：2015 年重庆市教委高校人文社会科学研究项目（15SKG023）
DOI: 10.3760/cma.j.cn116021-20200702-00524

Construction of psychological warning mechanism for college students in big data environment

He Xiaobo, Chen Fangfang
Information Center of Chongqing Medical University, Chongqing 400016, China
Corresponding author: Chen Fangfang, Email: 494182371@qq.com

【**Abstract**】 Combining the psychological "big data" of college students with the data of the psychological expert database, using data mining algorithm to extract characteristic attributes, deriving the warning results from a stable algorithm model, the psychological early warning platform constructed by this method can not only monitor students' psychological dynamics in real time, but also predict their psychological behavior trend, which makes up for the lag and low accuracy of traditional early warning methods. The psychological early warning platform provides data support and the psychological early warning system provides operational guarantee. The psychological early warning mechanism formed by the combination of the two can provide scientific basis for educators and finally achieve the goal of using big data technology to improve the mental health of college students.

【**Key words**】 Big data;　Behavioral data;　Data mining;　Psychological warning
Fund program: University Humanities and Social Sciences Project of Chongqing Municipal Education Commission in 2015 (15SKG023)
DOI: 10.3760/cma.j.cn116021-20200702-00524

目前，大数据的应用已深入到社会的各行各业，影响着人们的生活方式、个人思想与行为等。高校作为知识密集、思想活跃、最易接受新鲜事物的前沿阵地，学校的工作方法、教育管理模式等也受到了较大影响。如何将大数据技术的先进思想和理念应用到高校心理健康教育工作中，是高校教育工作者面临的新课题。

高校自开展信息化建设以来，建立了各类业务子系统，例如：教务管理系统、学生综合管理系统、人事管理系统、一卡通管理系统、科研管理系统等。这些系统经过长年的运行，累积了大量的数据。各高校将这些宝贵的资源整合起来，建立了自己的大数

据平台,其中与学生生活、学习有关的行为数据组成学生数据中心。学生在入学时进行的心理普查数据以及在读期间的心理数据,往往掌握在心理健康教育工作部门。这种通过专业心理评测或心理咨询师描述性记录获取的数据构成了学生心理数据。这两类数据之间看似没有简单、直观的因果关系,却存在着某种必然联系。因此,将这两类数据整合成为学生心理"大数据",并对其进行深入分析,才能在学生出现心理危机之前进行预测,使高校教师提前对学生心理问题进行疏导,减少因学生心理问题而导致的事故的发生。

高校学生的心理预警机制由心理预警平台及心理预警制度两方面构成。具体、有效、可操作的心理预警机制能够帮助教育工作者及时发现潜在的危机,进而采取相应措施,对学生进行心理危机干预,将危机的不良影响控制在最小范围内。

1　高校学生心理"大数据"

高校心理教育部门掌握着学生的基本信息、定期日常表现信息、与家长沟通等数据,高校的学工部、教务处、相关学院等部门重点掌握学生的考试成绩、校园卡消费及在图书馆的使用情况等数据。这些在一定程度上反映了学生日常表现、兴趣爱好、性格特点等[1]。因此,大数据时代下,大学生心理预警的数据来源不仅包括通过日常的心理普查、心理排查以及门诊问询等常规方式得到的心理数据,而且应该包括学生在学校期间与学习、生活相关的行为数据。

1.1　数据来源

1.1.1　学生心理数据

高校在开展心理健康调研工作时,为客观地掌握学生的心理健康状况,学校教育工作者会开展形式多样的心理工作,并产生大量数据。

学生心理数据来源,从不同角度出发,来源各不相同,具有多样化的特点。从数据的获取途径来说,数据来源主要有心理测评、心理约谈、心理咨询、心理普查等量表数据;从数据的产生主体来说,数据主要来源于高校学生、学生家长、教育工作者、心理咨询师及心理教师等;从数据来源的组织来说,数据来源主要有学院、社团、宿舍、心理咨询中心、家庭等组织团体。综上可知,心理数据的来源较多。但从目前来看,数据还是主要来源于学校的心理普查或测评、心理咨询师、心理健康咨询中心等方面,没有注重从其他来源去收集数据、整合数据。这也是学生心理预警研究面临的新课题。

1.1.2　学生行为数据

高校学生身处在校园环境中,与学习、生活相关的成绩数据、个人资料、考勤数据、上网行为数据、一卡通消费等数据蕴藏着巨大价值。

学生的行为数据量相对更大。高校学生使用网络媒体及智能终端越来越普遍,在使用中都会产生大量的文字、图片、视频数据。这些数据每天通过上网行为管理设备被记录下来,形成了海量的数据。这些数据是学生生活情况、思想状况、行为现况的现实反映。对这些数据进行分析,不仅能总结一些普遍性的规律,而且能发现危机的个例情况,对提升高校心理健康教育工作有着积极的意义。

在规律性问题的发现上,通过整合、汇总学生参加心理课程的时间、内容、课时数信息及参加校园文化活动的类别、周期频度、参加人数信息,可以发现学生的关注点及兴趣变化。这样可以帮助教育工作者及时调整心理教育课程授课的内容,扩充其范围,使其更丰富。同时,对受众度高的文体活动,可以提高开展频度,从而提高校园文化活动的影响力与参与度。此外,统计分析高校学生在图书馆借阅心理方面的图书信息以及心理方面电子文献的检索、下载信息,有助于了解学生心理方面的共性问题,为开展普遍的心理指导工作提供数据支持。

在个例问题的发现上,也具有十分重要的意义。比如,学生的上网行为主要体现在微信、QQ、论坛、微博等社交平台以及使用的软件和浏览的网页等。对这些行为进行跟踪和记录,能够为学生的心理研究提供最真实、全面的数据[2]。而上网行为管理设备是对学生行为轨迹记录的重要工具,其中主要包括访问的网页、网络上的发帖、传送的邮件及文件、参与的游戏、观看的影音、下载的资源等。设置如"自杀""厌世""抑郁"等关键词,结合搜索频度分析这些结果数据;重点关注搜索频度高的个别学生,有针对性地开展心理疏导工作,以化解其心理危机。再如,学生公寓区的门禁数据,如发现个别学生长期晚归或者夜不归宿,可以发出预警,提示心理教育工作者,排查存在于这些学生中的心理健康隐患,并给予适当的帮助,减少因心理问题给自己及他人造成的伤害。

1.2　数据类型

高校学生心理"大数据",有结构化和非结构化

数据两种。结构化的数据主要包括:心理健康工作者提供的数值型信息资源(如心理教育课程成绩、心理健康测量等),高校其他部门提供的教学、科研、财务信息及信息化主管部门提供的上网行为数据等。非结构化数据主要包括:心理健康工作者提供的文本、音视频、图片资源。其中,文本型资源主要来自心理课程作业、心理咨询记录、深度访谈记录和心理危机处理记录等;音视频型资源主要来自于学生心理活动、心理咨询、心理危机处理等[3];图片型资源主要来自于心理健康教育活动。从目前来看,非结构化的文本信息资源由于掌握在心理教育部门,获取较为容易且内容较为丰富,对心理健康工作者帮助较大,应用的较多;而对于高校管理部门提供的结构化行为数据,由于不太容易分析,应用的稍弱些。但这些数据价值较大,高校应重视这部分数据的收集,并将这些数据应用于学生的心理预防及干预工作中。

1.3　使用现状

心理数据与行为数据目前处于"孤岛"状态,学生心理数据主要掌握在心理教育部门,而行为数据则由学校的其他管理部门掌握。这导致学生心理数据与行为数据之间发生了割裂,无法共享。心理教育工作者努力通过心理普查、心理量表等方式获取到的大学生心理健康状况数据,仅来源于本部门,资料有限,只能在有限的方面得到应用。另外,学校的其他管理部门,因获取的学生心理数据中可公开的部分较少,也无法全面掌握学生的心理动态。在大数据的背景下,单一来源的数据应用价值非常有限,并且存在片面性的不足。高校只有将这些跨部门的数据,通过一定的逻辑规则进行整合,才能揭示其中的规律,进而指导心理教育工作者开展有针对性且有效的心理教育工作。

2　心理预警机制的构建

大数据背景下,充分利用信息技术构建高校学生心理预警平台,不仅能够及时、全面地获取学生信息;还能深入挖掘出与其心理状态相关的信息,及时帮助学校发现有心理问题的学生,并采取一定的措施帮助他们。

2.1　目前心理预警机制的不足

高校学生心理预警机制能够增强高校心理危机管控的实时性和有效性,从而创新高校学生心理健康教育工作[4]。近几年,高校虽然加强了对大学生心理危机预警的关注,但仍存在以下问题。

2.1.1　排查指标之间关联性差,预警效果具有片面性

高校学生心理预警常规的数据来源主要有心理量表、心理普查及门诊问询。这几个指标在学生心理预警工作中是相互独立的,每项排查指标间没有相互关联;并且这些常规的排查指标与学生行为数据也是相互独立的,在进行心理预警时没有综合考虑。这就影响了对问题学生进行心理危机识别的准确性,预警的结果存在片面性。

2.1.2　心理危机预警的时效性差

高校心理健康教育工作的有效性很大程度上取决于超前性。只有超前地用心理预警机制对学生的心理健康状况作出评判及预测,才能及时发现问题并采取相应的措施,以减少学生心理问题的出现及心理危机状况的发生。然而,现在心理危机预警工作的开展往往是通过相关教师的观察,发现学生的心理危机,然后上报给相应心理教育部门进行心理干预。这种人工发现问题、处理问题的方式所花费的时间在一定程度上影响了心理干预的时效性。这样的监控模式实时性差,而且不能进行互动,对学生的隐性心理问题无法及时发现。

2.1.3　无法动态监控心理变化

目前国内高校大都是采用在线答题的方式进行心理普查,且一般在学生入学时或入学后的一段时间定期开展,普查的频率有限。这样的方式不仅耗费了大量的人力、物力,在短期内不能重复进行,而且普查的数据也只是一些静态数据。而人的心理状态是随着环境、年龄等因素动态变化的,尤其是某些隐性的心理问题只有在特定的环境下才会产生。因此,静态记录的学生心理数据严重滞后于心理状态变化,无法全面地记录其心理变化轨迹,影响了心理预警的时效性。

2.2　心理预警平台的构建

高校学生心理预警平台的建立流程如图 1 所示。

2.2.1　收集数据

大数据背景下进行心理预警的前提是收集尽可能多的学生生活、学习及心理方面的信息,并将这些



图 1　心理预警平台的建立流程

信息进行整合,形成心理预警数据库。在进行数据收集时,将每一类数据项称为一个数据标签,这些数据标签如图 2 所示。

通过图 2 的数据采集方式,可以更加全面地采集大学生心理数据,有效地弥补了传统问卷调查存在的数据量小、分析样本不全面的缺点;对大学生心理状态做到了全面动态的掌握,为促进大学生的心理健康提供了数据支持。但是这种大范围的采集数据具有零散化、动态化、筛选难度大等问题,因此要做数据清洗、数据建模和数据挖掘工作。

2.2.2　清洗数据

由于数据来源于不同的部门,数据格式及质量存在较大差异,因此有必要重新对数据进行审查和校验,删除重复数据、纠正存在错误的数据。数据清洗的目的是过滤那些不符合要求的数据,将过滤的结果交给相应的业务部门,确认是否过滤掉或是修正之后再进行抽取。

2.2.3　数据模型的构建

高校信息技术工作者对预警平台的实现方式是熟悉的,但他们对需求可能是模糊的。因此,数据模型的构建需要各部门、各级组织的协同配合。心理教育工作者及学生管理工作者,依据工作经验的积累提供规则的描述;信息技术工作者依据一定的数据挖掘算法,进行代码的实现。

2.2.4　提供数据服务

通过数据挖掘技术,对存在于各个系统中的隐含的、有意义的数据进行分析和处理,发现新的关联关系,实现对学生数据的关联分析、概念描述、偏差检测,从而实现对学生心理趋势和心理行为进行自动预测[5-7];然后把数据挖掘得到的结果提供给授权的用户。

通过对数据进行采集、清洗、挖掘和分析,从繁杂多样的数据中获取到隐含的、潜在的和未知的有用信息[8-10],可以帮助心理健康教育工作者知道学生所思、所需、所行,提高其工作的科学性、预见性,从而更加精准地服务于学生, 然后开展相应工作。因此,它更应该是一个数据的平台、一个学习与交流的知识站点。通过将原始的、简单的、散落的、无关联的数据资源整合成知识与行动,服务于某些特定使用者;将心理健康教育服务体制中的被动服务转变为主动服务,实现数据向知识的转变、向行动的转变;最终服务于高校、服务于学生,使得心理健康教育能够将理论、技术和行动进行最完美的结合。

而本文提出的“大数据环境下大学生心理预警平台”,就是充分利用大数据的关联分析、分类、预测等特点,通过对高校学生进行数据的收集、实时跟踪和分析,最大程度规避研究对象在测试过程中受到干扰;从而得出学生行为与心理健康的关系,达到能够分级别的预警;一旦发现异常行为,可以立刻通过平台通知到相关的部门和个人,并提供有针对性的建议,帮助学校及时地进行干预,促进学生的心理健康。同时也为校级领导的决策分析提供有效的数据

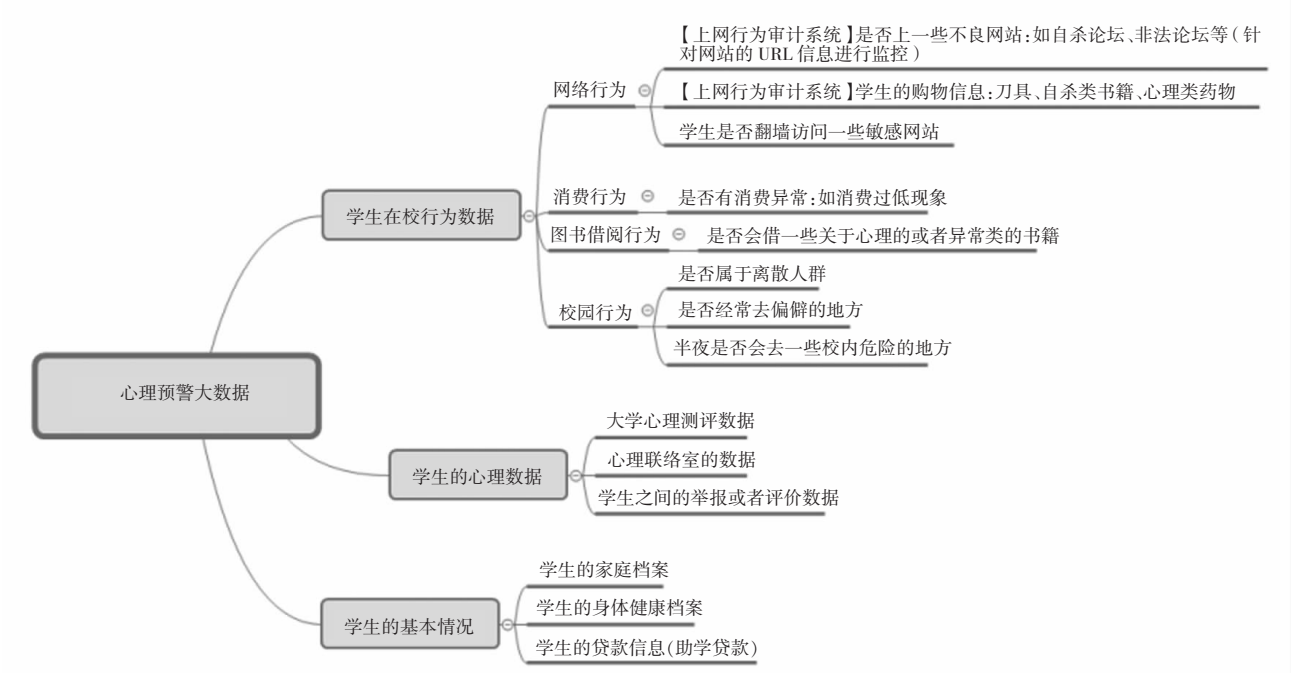


图 2　心理数据标签

服务,帮助教育部门准确、可视化地得到潜在危险个体,及时排查和干预不稳定因素和风险。

2.3　心理预警制度的保障

在大数据背景下,工作制度的健全是心理预警工作能有效开展的前提。它主要包括预警工作组的建立、数据的收集制度及数据的保密制度。

心理预警平台的有效运作,不能仅靠学校的心理健康教育部门,更需要学生处、教务处、财务处等多部门的协同配合,需要心理教育部门的专职教师、学生处教师、教务处教师、学生干部、医院心理咨询师等多群团人员的配合。高校应设立专门的工作小组,做好总的协调工作,明确工作职责。

为保证数据收集的全面性和准确性,各部门要明确数据"谁产生,谁维护"的制度,这样才能清洗掉"脏数据",使数据顺利流转。同时,也可采用主动填报的方式来提高数据收集效率,尤其对于那些没有数据来源的数据,这种收集方式非常有效。

强化数据保密意识,注重隐私数据的安全。大数据环境下,学生在校期间网络访问情况、一卡通消费情况、学习成绩等数据涉及学生的个人隐私,尤其要重点保护。预警平台要做好权限控制,只有授权的用户才可访问相应资源;并且授权用户要加强数据保密意识,保证在合法的前提下开展心理预警工作,保护好学生隐私。

3　心理预警平台的模型建立与实验结果分析

重庆医科大学以数据中心为基础,按照我国教育大数据的建设规范对业务系统、网络系统的相关数据进行采集、挖掘、分析,利用大数据技术,构建了心理预警平台。该平台的数据来源于学生行为数据与心理专家库数据。将行为数据与专家库数据结合,构建数据模型,利用数据挖掘算法,得出预警结果,是研究者研究的重点。目前研究者完成了学生画像与孤僻预警两个模块。

3.1　孤僻预警的模型建立

第一步:选择特征属性。在本研究中,研究者从心理和行为两方面提取孤僻人员的特征属性。特征属性的来源:①心理特征。重庆医科大学利用自身医学优势,由附属医院心理专家,根据工作经验并利用专业的心理分析工具提取出孤僻人员的心理特征。这些特征值存放在心理专家库中,供研究者分析研究。②行为特征。学生的在校行为,主要通过一卡通数据展现。一方面,以班级为单位来抽取学生的生活及学习时间数据。选取的特征属性有早餐时间、中餐时间、晚餐时间、就寝门禁时间及上网时长。另一方面,是否有就餐同步的行为及是否有借阅同步的行为也作为行为特征记录。

第二步:特征标准化。本研究的目的是要从心理和行为两个方面分析出具有孤僻性格的人群。哪些学生是孤僻性格首先是不知道的,因此在具体的算法分析时,只能考虑用无分类标记的分类算法进行分析。这类算法也称之为无监督学习,主要通过聚类方法得到。

第三步:聚类算法进行聚类。聚类算法是指事先不知道样本的类别,通过某种方法,把相似的样本归为一类,而不相似的样本划分到不同类别的过程。根据实际,适合混合型数据类型的算法只有 k-means 算法和 k-prototype 算法。通过建模,两种聚类算法得到的孤僻人群分类结果是一致的。这说明这个聚类结果是稳定可靠的。

3.2　孤僻预警模型的可行性验证

重庆医科大学的一卡通系统有学生消费数据分析模块,从中抽取出学生就餐社交关系见图 3。图 3 中一卡通得到的数据,与孤僻预警模型得到的结果数据比较,二者是基本一致的。因此,此模型的建立是可行的。

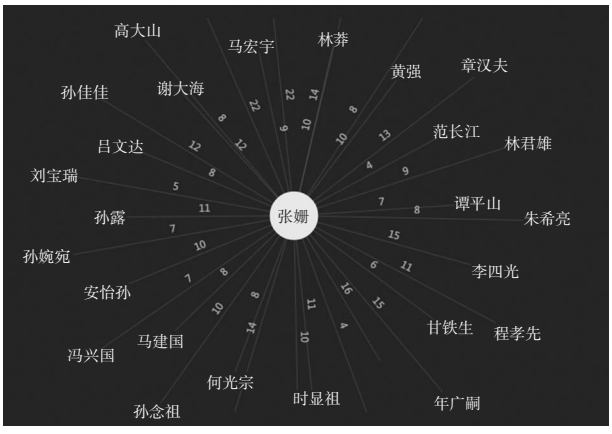


图 3　不活跃学生社交关系图

4　心理预警平台的实施

心理预警平台模型建立后,研究者通过软件开发,构建孤僻预警功能。通过对相关学生、相关学生专业平均水平的同吃饭、同购物等数据以图和表的形式进行展示及预警,并且对特定的人员授权查看,以此来帮助教育工作者开展工作。图 4 为平台的孤

僻人员预警页面，该生 6 个月内与同专业的同学同吃饭、同购物的次数是 0。这时平台就会有预警提示，教育工作者可结合此学生的日常生活、学习情况对其进行关注。

通过对在校学生的轨迹记录（门禁、一卡通、上网、考勤、WiFi 连接等）实时监控，分析学生的在校情况，对疑似不在校的学生进行识别和预警。教育工作者可根据预警信息对疑似不在校学生及早处理，根据不在校的不同情况对学生进行分类管理（图 5）。

沉迷游戏分析是对学生的每天上网时长与排名、每天游戏时长与集中游戏时间段、累计挂科门数、图书馆借阅次数与排名、宿舍门禁记录等各方面情况进行大数据分析，展示出学校各个学院或专业沉迷游戏的人群（图 6）。

晚寝晚归预警是从学生未住宿、晚寝晚归、宿舍



序号	类别	同去食堂吃饭（次）	同去超市购物（次）	同出入图书馆（次）	同出入宿舍（次）	同出入学校（次）
1	该生	0	0	0	0	0
2	同年级同专业平均	46.48	0.91	0	0	0
3	同年级同专业最高	404	18	0	0	0

图 4　孤僻人员预警

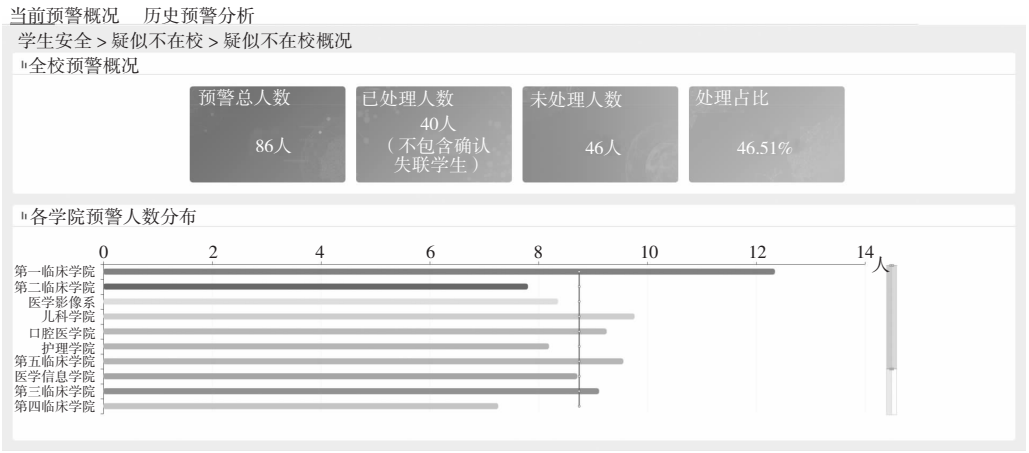


图 5　疑似不在校预警



图 6　沉迷游戏预警



图 7　晚归预警

住宿人数等方面进行预警;同时将预警情况及时通过短信、微信、邮件等方式推送给教师与学生,便于教育工作者及时了解学生的情况并进行及时干预,消除学生管理存在的隐患(图 7)。

未来,研究者将结合学生的上网行为、图书借阅、专家库等数据,构建更多模型,提高预警的准确性。

5　结语

大数据背景下,高校学生心理预警工作除了常规的心理排查、门诊问询等方式,还可以结合学生的行为数据进行数据挖掘。心理预警平台的构建,能够挖掘出潜在的、隐含的和未知的有价值信息,能大大提高高校心理健康教育工作的科学性和预见性。同时,在工作开展中,要注重学生隐私的保护,保证数据安全,使心理预警工作科学有序地进行。

在这种时代背景下,针对高校学生的心理健康教育工作,在工作提升和创新改进方面的空间变得更大,工作平台也更加广阔。通过构建大数据背景下的大学生心理预警机制,将学生的心理数据和每天的学习、生活等行为数据结合,进行全面的数据收集和分析,从生活习性、行为规律、心智状况等方面对学生进行一个画像描述。这不仅有效地解决了传统预警工作存在的被动性、受限性以及滞后性等问题,还能帮助学校及时地发现和识别出大学生中各种潜在或现实的危机因素,进而能够及时主动地采取专业措施进行有效干预。

利益冲突　所有作者均声明不存在利益冲突

作者贡献声明　何小波:项目实施、凝练主要观点、撰写论文;陈方方:数据收集、平台搭建、协助成文

参考文献

[1] 张家明, 王纯静. 基于大数据技术的大学生心理危机预警研究[J]. 教育与职业, 2015(30): 75-77. DOI:10.3969/j.issn.1004-3985.2015.30.025.

[2] 卢小杰, 黄道斌, 叶明全. 大数据四层结构的学生心理健康促进模式研究[J]. 电脑知识与技术, 2017, 13(13): 1-2. DOI: 10.14004/j.cnki.ckt.2017.1285.

[3] 安哲锋, 张鹏. 大数据背景下高校学生心理健康数据分析及反馈体系构建研究[J]. 思想教育研究, 2016, 9(9): 93-96.

[4] 徐文明, 梁芹生. 大数据技术在心理健康教育中的运用与研究[J]. 嘉应学院学报, 2017, 35(4): 96-100. DOI: 10.3969/j.issn.1006-642X.2017.04.019.

[5] 刘同明. 数据挖掘技术与应用[M]. 北京: 国防工业出版社, 2001: 34-179.

[6] 杨利. 基于大数据的数据挖掘算法再大学生心理危机预警中的研究[J]. 数字技术与应用, 2017, 16(3): 147-148. DOI: 10.3969/j.issn.1007-9416.2017.03.082.

[7] 刘冰, 董小玉. 基于大数据技术的学生心理危机预警机制构建研究[J]. 大众科技, 2017, 19(4): 122-124. DOI: 10.3969/j.issn.1008-1151.2017.04.046.

[8] 张凯, 应奕彬, 赵泽宇, 等. 面向对象的个人数据中心规划与实施[J]. 华东师范大学学报(自然科学版), 2015(z1): 266-273. DOI: 10.3969/j.issn.1000-5641.2015.z1.043.

[9] 赵泽宇, 张凯, 宓詠. 高校一站式信息服务[J]. 科研信息化技术与应用(中英文), 2012, 3(3): 52-59.

[10] 宓詠, 赵泽宇. 大数据创新智慧校园服务[J]. 中国教育信息化·基础教育, 2013, 24(12): 3-7.

(收稿日期:2020-07-02)
(本文编辑:张学颖)