

设计——网络阅卷技术应用的深入之道

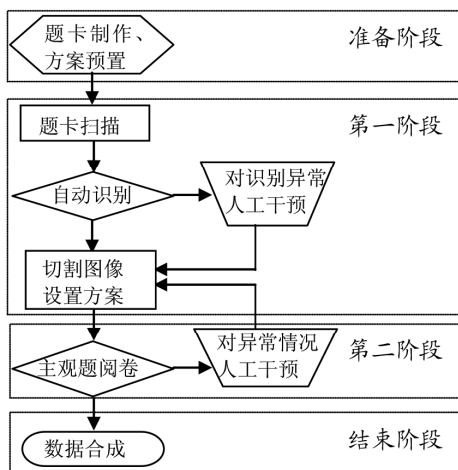
金坛市教育研训中心 王志忠

随着时代的发展,网络阅卷这一形式已经被各级教育考试机构在大、中招生考试中普遍采用。这种现代阅卷技术与传统手工阅卷技术相比,在提升阅卷质量方面表现出了无可比拟的优势,例如计算机控制之下,这样的阅卷不可能存在着客观题错批、漏批、结错分这类低级错误,而且能对主观题的给分过程进行科学、有效地动态管理。网络阅卷不仅能成倍提升阅卷的工作效率,在考试数据统计与分析方面更是表现出强大的优势。所以近年来,网络阅卷这一现代技术应用也逐渐普及到一般学校,应用于学校间的普通考试、测试。

尽管目前应用市场上的软件品牌有所不同,但其技术方法、应用方式是基本相同的。所谓的网络阅卷系统,是指以电子扫描技术和计算机网络技术为依托,将纸质的试卷或者答题卡扫描后获得电子图像后,实现客观题自动阅卷、主观题网上人工评卷,最后按要求自动输出各类数据报表的综合系统。综合来看,这套系统的工作流程大致分成四个阶段(以武汉丽升网络阅卷系统为例,如右图):

准备阶段,主要工作是答题卡(试卷)的编制与印刷,阅卷方案的预设。第一阶段,是整个流程中工作量最大的部分,需要先进行所有答题卡(试卷)的扫描,然后对图像进行自动识别和切割。通过识别可以自动获取每张答题卡的考号、客观题的选择项。切割就是按既定方案,把扫描得到的一整张图像中有关主观题部分切割成若干小图块,以便进行后续的主观题批阅。这其中还要对考号填涂错误、选择项填涂不清晰、图像不能被自动识别等大量异常情况进行人工干预。第二阶段的主要工作是组织老师通过计算机网络对主观题部分的题目进行批阅,对答题位置、图像异常的情况进行人工干预。结束阶段,要对数据进行合成,然后按要求导出各类数据、统计表格。

从上面的介绍中,大家都会感觉到网络阅卷系统仅功能强大,而且在技术方法也相当科学、严密,作为普通用户,只要能应用就可以了。然而通过多年的实践应用后,我们



认为一项完整的技术,应该是一种“系统”的大概念,而不单单是物质化的物件。所以“网络阅卷系统”也绝非仅指高速扫描仪、计算机及其相应软件,这些只是网络阅卷技术实现的工具,或者叫物化的“外壳”,技术需要在人的主导下,方能彰显其力量,这个系统绝对不能缺失主导者,而且这才是这套系统的核心与灵魂。

从这个角度上说,我们每个用户在与技术交互的过程中,必须展现出一种“能动的主体”的姿态,也就是说,我们在应用网络阅卷这类全新技术的过程,就必须经历初步尝试、熟练应用、反思改进,进而达到能根据自己的需求,通过精心设计,让技术为“我”所用。只有这样我们才能真正成为技术的统帅,而不是附庸。我们在对全市部分小学调研中发现,某年级小学数学的一类计算存在着普遍的问题,为了进一步调查、研究这个问题,在期末测试中,研究人员对这个问题进行思考,并有了初步判断,然后有意设计若干测试题,在阅卷过程中,单独设置成一个给分项目,并且要求阅卷老师在批阅过程中注意搜集典型的错误类型(保存图像、做相应标志)。在后续研究中,利用详尽的数据结合大量典型答题过程,进行科学、高效地研究与分析。从这个例子中,我们可以看出实现技术过程的核心是人,而要成为一个统帅技术的应用者,其关键又在于“设计”。

关于“设计”,在百度百科中是这样说的:“人类通过劳动改造世界,创造文明,创造物质财富和精神财富,而最基础、最主要的创造活动是造物。设计便是造物活动进行预先的计划,可以把任何造物活动的计划技术和计划过程理解为设计。”这样的定义很能反映设计在网络阅卷技术深入应用中的地位与作用。因为在网络阅卷技术应用过程中的设计也常常表现为计划技术和计划过程。计划技术,可以理解为对网络阅卷技术的应用更加适当;计划过程,可以理解为对网络阅卷过程的管控更加自如。

计划技术,对网络阅卷技术的应用更加适当

考试是各级各类学校教学活动中最为常见的活动,也是大家最为重视的环节,老师们对组织考试的每个环节都非常熟练,因此在传统的纸笔考试、手工阅卷的形式下,可以通过“设计”来改进此项工作的空间并不大,但是一旦采用网络阅卷形式,以计算机技术为核心的技术方法则整体介入,所以尽管对于主观题,仍然需要人工判分,但是整项工作的组织方法上,则发生了质的变化:人工阅卷过程中脱离了试卷这个“不可分割”的物化载体,取而代之的是在计算机技术管控之下的“题块”,这既是两种阅卷方式的一个重大不同之处,也是我们可以进行“设计”的着手之处。笔者愿意结合自己的实践,以“题块”的设计为例,谈谈对网络阅卷技术的适当应用。

所谓“题块”,是将答题卡扫描后获得的整幅图像,根据一定的目的,分割成若干小图像,然后在计算机数据库控制下,通过网络呈现给指定的批阅人客户端界面中。我们习惯把这样的小图像称为“题块”,它也相当于这套系统中的“细胞”。我们可以延续手工阅卷的习惯,根据题量、批阅人数等情况,确定每个人的批阅任务,从前往后,将每项

任务划成一块。但这样的思维还只是停留在手工阅卷时代,为此,我们有必要通过有意识地设计,将计划的优势更好地为我所用。例如还可以这样设计“题块”:

■ 块之分:有目的地细分。例如出于某种需要,在出题时就已经预设精确的监测题点,我们可以根据需要将这些题块划得很精细,这样就可以在后期获得精细数据。除了细分这个办法,还可以通过“粗中有细”来实现,即通过设置,让批阅者在一个大题块中给出其中各个小题的分值。例如,一次物理考试中,一共有五道填空题,前二道是力学部分的题,后三道是电学部分的题,于是我们可以让阅卷老师在批阅时分别各给出一个分值。通过这样的设计,我们在后期既可以获得整个填空题的数据,也可以获得前、后两类题目的数据,进而可以了解到在填空这个题型上学生的整体表现、每个模块上的掌握情况。

■ 块之合:有目的地组合。在网络阅卷技术背景下,不仅可以扫描图像的不同区域规划成相应题块,还可以根据一些特定的需要,将题块进行组合。例如,在一次小学数学考试中,测试学生计算能力的题目是第4题和第7题,我们为了从整体上了解学生的计算能力,就可以将这两题划成一个题块。块之合,也可以进行跨越式的拼合。例如,在完成划块后,分配阅卷任务时,我们发现某某老师特别善长批阅第1块和第11块,我们也可以让网络阅卷系统将这两块重新拼合成一个题块,从而达到既定目的。

■ 块之增:有目的地增设。一般情况下,一个图像区域设置成一个题块,但是也经常需要在同一区域进行多重判分。这在网络阅卷中也可以轻松实现,例如语文考试中,既要给学生的作文判分,也想增设作文书写得分点,这样的要求,就可以通过在一块中增设两个得分点,或者同一图像区域内再增设一个题块来实现。

在上述例举中,我们之所以能根据需要,自如地进行题块的分与合,是因为有网络阅卷系统背后计算机技术、数据库技术作支撑,是依靠自动化管理、大数据计算来实现的。

计划过程,对网络阅卷过程的管控更加自如

网络阅卷之所以能具有显著的优势,是因为其基于了计算机的自动化管理与管制。但是,我们如果仔细观察前面所列的“工作流程”,就会发现在整项流程中,至少有两处“对异常进行人工干预”。例如一次考试中,某位学生在答题时发现自己答案错了,就在超出答题区域的旁侧继续作答,这样就可能会产生两种问题,一是阅卷系统可能不能识别这一张答题卡(试卷),二是阅卷老师无法看到后续作答的图像。一旦出现这样的问题,都需要通过人工的干预才能完成整个流程,因此在实践中,网络阅卷的过程是否能顺利、高效地进行,不仅取决于阅卷系统软、硬件的性能,还取决于答题卡(试卷)是否符合系统要求等人为因素。我们可以对答题卡(试卷)的印制等环节进行严格的质量控制,但是却无法对学生使用答题卡(试卷)的过程作出过于刚性的要求,因此我们在答题

卡(试卷)的制作过程中,就必须进行充分地设计,通过计划学生答题的过程,来确保网络阅卷实施过程的顺畅。所谓的“计划过程”除了计划学生答题过程,还应包括计划阅卷老师的阅卷过程,因为在网络阅卷系统中,所有的过程都是由数据库的控制之下进行的,因为我们可以通过设计,充分利用数据对阅卷全过程进行动态地掌控。

■ 建立规范。促使学生的答题过程不无意外、少出意外,其根本要义在于引导他们有序地作答,因此我们在设计答题卡(试卷)时就必须形成一整套规范,以规范促成有序。例如,字体、字号在标题、正文、引用文、提示语等不同之处的统一、规范使用;答题区域用粗实线划定、不同题目(块)间用细实线区割、答题下划线用细虚线等等。将这一套规范完整地应用所有学科的考试中,就能使答题卡(试卷)的结构规范、布局明晰,以此来促成学生作答的规范、明晰。

■ 合理引导。少年儿童的年级特点决定了他们在考试过程中,会由于种种原因产生各种错误,因此在建立规范的同时,进行合理的引导就显得十分必要。例如,答题卡中的题号用大号黑体字显示,必要时加少量的题干作提示,就可以有效避免学生作答时“张冠李戴”。在语文等学科考试中,常会有一、两道选择题散布在卷面上,如果在答题处标示出“请用2B铅笔填涂选项”,就可以大大减少学生漏答的情况。对于各类作(画)图类型的题目,提示学生用一定标准的铅笔作答,就能避免学生反复擦改而导致卷面破损的问题。

■ 合情帮助。对于年级较低的学生,在设计答题卡(试卷)时还必须给每一道的作答区域留有余地,帮助他们顺利改错。在题块划定的过程中,要适当扩大区域,照顾到少部分“答过界”的学生。学生作答形成的其他各类错误、异常,要组织监考老师考后详细记录,然后在阅卷过程给予合规的弥补。

■ 创新形式。网络阅卷技术始用于高校、中学,因此在形式上,也一般都是“答题卡分离”,也就是在常规的试卷之外,再增加一张专门用于答题的“答题卡”。而随着这一技术的逐渐普及,在小学阶段使用时常会出现部分小学生习惯性地答案写在试卷上的严重问题。为此,我们将传统的试卷按网络阅卷的技术规划进行必要的调整,成功实现了“答题卡合一”,从而完全避免了上述弊病,也节省了系统运作的成本。

通过上面所列举的实例,我们可以发现,通过以人为本地设计,可以创设出更加友好的交互环境,在帮助学生顺利作答的同时,也确保了系统高效地运行。从我们的实践来看,充分利用技术优势,还可以通过计划老师阅卷的过程,来确保网络阅卷的质量。

■ 自评双评,各有所用。长期以来,如何确保阅卷的质量,一直是困扰我们的一个问题,很多情况下只能依赖阅卷老师的“自觉”。但是在网络阅卷技术中,我们就可以通过设计“自评率”来有效地观察到每一位阅卷老师的质量。例如,某一道填空题,我们在设置批阅方案时,将其“自评率”定为5%,也就是说,阅卷系统会在阅卷老师已经批阅过的题中,随机抽取5%的题目让这位老师再次批阅,然后比较两次给分的相似度。通

过这样的设计与应用,我们就可以较明晰地了解到每位老师的批阅质量。而对于作文等主观类题型,我们通过设置双评、超过预设分值再一次进行三评的形式,由系统自动进行控制。

■ 效率质量,动态掌控。在传统手工阅卷中,因为手段的缺乏、精力的限制,管理者根本无法对每一位阅卷老师的动态过程进行有效掌控。但是在网络阅卷技术环境下,原本不可能的事情就会变得轻而易举。在网络阅卷系统中,关于客户端设置一般都分了阅卷员、科组长、管理员等不同等级的用户角色。我们可以事先预设科组长、学科管理员等相应用户,然后在阅卷过程中,指定管理人员使用相应管理帐号登陆阅卷系统后,他们就能通过数据、图表详尽地了解到每一位阅卷老师的工作状态,例如:批阅的数量与速度,还有他给分的平均值以及方差(表现他的给分值的离散度,代表该老师给分的波动情况,例如在一些有一定自由裁量度的题目批阅中,如果打“保守分”,那么就会从方差这个值上体现出来)、吻合度(阅卷老师个人给分均分与本题均分之间的吻合程度,在一题多人批阅时,可以表现出个体与群体之间的吻合度,也可以通过阅卷时的给分曲线的吻合情况来观察)、有效度等项目,通过对这些数据的跟踪观察,我们就可以对阅卷实施的动态过程了然于心,掌握自如。

本文发表于《中小学信息技术教育》2014年第11期