

人才培养全周期大数据驱动的医卫类职业教育

质量诊改方案的研究与应用研究报告

一、项目基本情况

本项目属于 2018 年广东省高等职业教育教学改革研究与实践项目，于 2019 年 8 月 21 日立项。本项目经过为期近三年的研究，取得了初步的成效，共公开发表相关论文六篇，建立了校本教育大数据平台，整合了包括智慧校园系统，一卡通系统，在线教学平台，在线直播平台等多个信息化系统，采集各个系统运行过程中的产生的各类过程性数据，并建立了大数据可视化系统，实现了数据的融合、交叉查询、动态筛选，初步实现了基于教育大数据的诊断与改进。项目经过 2019、2020、2021 三个年级近 1 万名学生的持续数据跟踪分析，初步实现了基于大数据的学校治理，教育教学、生活后勤等方面的宏观质量诊改，以及面向师生个人的个性化的成长记录，质量评价。

二、人才培养全周期大数据平台的构建

1. 大数据基础平台的构建

本项目的研究与实践，完全依赖于人才培养全周期大数据平台的建立。如何针对医卫类职业教育的特点，定制开发一套有针对性的基于大数据的教育质量诊断与改进平台（以下简称“大数据平台”），是本项目研究的难点之一。

项目研究初期，项目组经过多次的反复技术讨论和遴选，最终决定采用开源 JEECG 框架为基础，搭建大数据平台。Jeeccg 框架有完善的技术架构，前后端分离，能实现安全稳定的敏捷开发。本项目以

该框架为中心，采用框架提供的用户管理结构，实现了用户管理和统一身份认证。该框架前端对接学校企业微信和微信公众平台，实现了师生用户的手机微信绑定、手机微信认证，只需要通过手机微信即可登录和验证身份。框架后端数据库对接各个应用系统，通过中间数据库的形式与各个应用系统的基础数据库进行数据对接和数据同步。在此基础上，项目组定制开发了数据同步机制和数据清洗模块，从技术上实现了数据的采集、互联互通。

2. 数据可视化平台的构建

数据可视化系统是大数据平台的展示端，能够将错综复杂的数据自动生成可视化图表，实现大数据的展示、筛选、分类统计等功能。本项目研究初期选定了 Superset 开源数据可视化系统，作为本项目的可视化平台（<http://s.welearning.net.cn>），并构建了一系列的大数据可视化诊改看板。

看板		
Search		
Dashboard	Creator	Modified
2020年科研工作量统计大数据	李智高 网络教学中心	2 months ago
2021新生报到大数据	李智高 网络教学中心	7 months ago
2021级中职报名大数据	李智高 网络教学中心	10 months ago
2020年联盟联考成绩分析系统2	李智高 网络教学中心	1 year, 1 month ago
2020年联盟联考成绩分析系统	李智高 网络教学中心	1 year, 1 month ago
考生择校意愿问卷调查结果	李智高 网络教学中心	1 year, 6 months ago
2020级签到数据大屏	李智高 网络教学中心	1 year, 7 months ago
基础护理技术评价系统	李智高 网络教学中心	1 year, 9 months ago
高职扩招直播考勤大数据	李智高 网络教学中心	1 year, 9 months ago
2020毕业考试及补考大数据	李智高 网络教学中心	1 year, 10 months ago
2019随身课堂各课程各模块统计表	李智高 网络教学中心	2 years ago
2019高职扩招大数据	李智高 网络教学中心	2 years ago
单词达人大数据	李智高 网络教学中心	2 years ago
2017级实习生问卷调查	李智高 网络教学中心	2 years ago
随身课堂运营综合数据大屏	李智高 网络教学中心	2 years ago
第二轮课程联考试题数据	李智高 网络教学中心	2 years ago

经过一年多的实践，项目组发现 Superset 开源数据可视化系统，

作为一个开源的数据，可视化平台功能十分强大，可定制程度高，但是也有很多不足之处。1 是该平台作为国际主流的开源大数据可视化平台，主要的语言为英语，本地化语言汉化不足，导致在使用过程中大量界面出现英文按钮；2 是该平台作为国际开源大数据平台使用大量的国外开源插件，在国内没有独立的维护团队，从可靠性和安全性考虑，都需要另外寻找能够更好服务本地的数据可视化平台。

经过对国内同类型产品项目的遴选，项目组最终选定使用开源的 DataEase 平台，替代 Superset 平台，作为本项目的数据可视化平台（<http://10.16.10.211> 仅供内网访问）。该平台同样是开源的数据可视化平台由国内团队开发和维护，相对于前者有更好的本地化服务以及更适合国内技术发展趋势。经过将近 9 个月的部署实施和应用，项目组对该平台的用户界面、图表功能、数据统计分析等功能进行了全面的分析和研究，最终决定全面转向该平台。目前已经实现了该平台的用户管理与大数据平台对接，各类数据与大数据平台对接，并初步实现了智慧校园系统、在线教学平台，一卡通系统，在线互动直播平台等系统的数据可视化分析。





三、人才培养全周期大数据的数据整合

1. 数据标准的构建

(1) 数据字典

由于本项目涉及到的信息化系统繁多，各个系统在建立初期没有统一的数据字典，导致各个系统当中的数据不标准。本项目研究初期，首要完成的就是对校本数据字典的标准化建立。通过对数据字典的建立，将学校的基础数据，包括各个行政部门，岗位、二级学院、专业、班级、人才培养方案、教学计划、课程编码、各类建筑物、房间等基础数据全部统一编码。制定了各类过程性编码的结构规则，确定了，以学生学号、教职工教工号，为系统唯一索引字段。为后续的数据整合、查询分析打下基础。

(2) 系统用户架构

根据本校的实际情况，学校在数字校园、一卡通、企业微信、腾讯微校等多个系统分别都有用户架构，经过项目组的调研讨论研究，最终决定以企业微信的用户架构为基础，以大数据平台作为用户架构统一管理和统一认证平台，其他平台根据此架构进行调整，接入用户

管理，其他应用系统不再管理用户身份信息。

(3) 用户身份绑定

确定系统用户架构后，用户的数字校园系统，一卡通系统，企业微信，腾讯微笑等多个系统，全部需要与大数据平台的用户帐号进行身份的绑定。完成身份绑定后，用户通过上述的任意系统，通过手机进行扫码即可登录完成身份认证，实现各个应用系统的统一身份认证和单点登录。

(4) 应用数据接口

项目组经过研究分析各类应用系统的接口分成两个主要类型：一种是通过 API 进行数据对接，另外一种是直接通过数据库对接。

通过 API 进行数据交换，需要在大数据平台开发对应的 API 数据接口，例如：“腾讯会议”系统。通过 API 接口方式，能够安全稳定高效的获取相关应用数据，此类方式比较适合互联网应用和成熟的第三方应用。

直接通过数据库进行对接，常见的数据库，包括 MySQL（智慧校园、在线教学平台）、SQLServer（科研管理平台）、Oracle（一卡通平台）。从安全的角度，直接连接数据库方式并不可行，一般由应用开发商提供中间库的形式，将关键的数据同步到中间库，再从中间库对接到大数据中心数据库进行查询，以确保原始应用数据安全性。

2. 智慧校园数据采集

智慧校园系统由成都中教智慧信息科技有限公司提供，系统采取 BS 结构，数据为 MySQL，系统内置了包括用户管理、OA 行政管理、

人事、教务、学工、后勤、迎新、质量监控等一系列应用，同时系统还构建了数据中心，已经将各自应用系统的通用数据集中管理。通过对接中间库，目前实现了人事数据、学工数据、数字校园基础数据、教务、数字迎新等基础数据的对接。

3. 教育教学过程数据采集

学校早于 2007 年就建立了校本教学平台“随身课堂”，一直沿用至今。平台覆盖全校各专业所有相关课程，服务全校各专业师生。平台已开源 moodle 系统为基础二次开发，数据库为 MySQL。大数据平台通过与数据库对接，实现了教学平台数据、在线课程数据、课程运行数据、用户使用数据五个层面的数据分析，能详细跟踪用户的每一次点击，每一次交互操作。

4. 生活后勤数据采集

生活后勤数据主要包括四大部分，1 是一卡通消费系统、2 是智慧热水系统、3 是智慧电表系统、4 是校园网络系统。上述系统用户管理全部通过数据中心用户统一身份认证系统对接，所有用户全部以学号或工号为唯一的索引。通过一卡通系统可以获取用户每一次刷卡消费的准确信息，包括刷卡消费的时间、金额、消费位置等基本信息。通过智慧热水系统可以获取到每个用户使用热水的具体时间，数量，消费金额，时长等基本信息。通过智慧电表系统可以获取到各个宿舍用电的情况，包括每小时的用电情况，用电峰值、用电功率、用电量等基本信息。通过校园网络系统可以获取到每个用户的每天上网流量、上网时长以及上网时间。

四、基于人才培养全周期大数据的诊改



1. 学校管理大数据诊改

通过数据整合，实现了学校的人力资源、固定资产、教学资源大数据统计分析。

人力资源层面，可以通过各类筛选、快速查找、统计分析全校、各个二级学院的教职工队伍结构，性别比例，学历职称，各专业教师队伍构成、年龄情况、教师专业特长、继续教育，职称晋升等多维度大数据。通过过程性数据，年度数据对比，可快速诊断人力资源、教职工队伍架构性问题。

固定资产层面，可快速统计分析全校及各个二级学院相关固定资产实训教学设备、各年度学校及二级学院固定资产投入、生均经费投入、结合教务系统提供的质量评价数据，有效能量化投入产出比，对学校固定资产建设、项目投入有很好的诊断和决策依据。

教学资源层面，对全校多媒体课室座位、实训室工位、公共计算机实训中心、图书馆等设备设施的数据进行整体统计分析，根据各校区、各二级学院、各专业的学生数量，统计分析相关专业、相关校区的各类教学资源，为校区建设、专业建设、实训室建设提供诊断和决

策依据。

2. 教学质量大数据诊改

教学质量监控大数据主要依赖三个层面，1 是线下教学督导评价反馈；2 是在线教学平台运行数据；3 是在线互动直播平台运行数据。

通过三个层面的数据整合，可以对某一门课程的线下教学、线上教学的数据进行统计分析，了解教师线下教学情况，线上教学情况，师生互动情况，在线交流活跃程度，在线任务作业完成及批改情况，在线答疑频率，在线直播及互动情况。

通过数据统计分析和筛选，可以快速定位某一个专业、某一门课程、某一位老师整体教学情况及每一个单独教学单元的具体实施情况，对教育教学实现过程性常态化监控，系统实时对教学潜在如：作业完成率低、线上课程参与度、低直播课程参与度低、师生在线交流频率低等问题，进行数据统计分析和预警，实现教学质量的常态化、过程性诊断与改进。

3. 学生成长评价大数据诊改

学生成长评价数据主要包括了后勤生活数据及学习行为大数据两部分组成。

生活后勤数据包括了一卡通消费、用水、用电等数据。一卡通消费对于学生的消费行为习惯进行统计分析，能精准判断学校的贫困学生。通过智慧热水系统，智慧电表、校园网络管理系统数据，能够准确的掌握学生生活作息时间，网络及手机使用时间，对学生身心健康成长形成数据监控，对非正常数据表现进行及时预警和干预。

学习行为大数据，通过教务系统、实训管理平台、在线教学平台、在线互动直播平台采集的数据，可以对全校各专业各年级学生的学习表现进行宏观的评价和对比，也可以针对个人进行个性化的数据。实现学生学业水平终末性评价和课程学习过程性评价相结合。

结合上述的数据，利用消费、用水、用电、在线学习等数据的动态变化，还能对学生校园安全等问题进行统计分析和实时预警。

五、项目不足之处和后续研究方向

1. 不足之处

- (1) 由于学校新校区建设施工落后于预期进度，部分业务应用系统未能按照计划上线，未能实现智慧校园系统应用的全面数据对接。
- (2) 本项目基本达到预期目标，完成了应用系统的对接的基础技术平台和教学诊改框架。但是本项目接入的数据仍在不断扩充，对数据的清洗、关联、统计分析的信度、效度仍有较大的改进空间。
- (3) 本项目研究周期较短，仅有一届完整的学生数据，未能形成有效的横向对比数据。因此，尚未形成可实施的管理制度和方案。

2. 后续研究方向

(1) 进一步完善数据采集来源

结合学校新校区建设，进一步整合各类门禁系统、物联网等系统，在校内垂直领域，实现更深荣、更多维度的数据源头来采集。通过对生源数据采集，实习就业岗位的数据采集，横向拓宽数据源头。

(2) 利用新技术提升线下教学数据采集

结合常态化录播系统，利用人工智能、语音识别、图像识别等技

术,实现对线下课堂教学的质量评价和监控。通过对智能化教学模型、实训系统,实现实训教学的智能化评价。

(3) 形成可实施的管理制度和方案

再经过 1~2 年的实践,形成 2~3 届学生的完整人才培养全周期过程性数据,可实现横向数据对比分析,进一步完善数据评价模型和数据预警系统。在技术的支撑下,逐步形成可实施的管理制度,以技术为基础,以制度为保障,形成一套面向医卫类职业院校,基于教育大数据的人才培养全周期教育教学质量诊断与改进方案。