

**数据治理大作业**

|  |  |
| --- | --- |
| **课程名称：** | **数据治理** |
| **学生姓名：** | **刘博严** |
| **学生学号：** | **202030390308** |
| **学生专业：** | **电子商务** |
| **开课学期：** |  |

**电子商务系**

**2023年 6月**

识别主数据管理的驱动因素和需求

在大学的各种系统当中，识别主数据管理的驱动因素和需求有着以下几点：

满足学校的数据需求：学校中多个业务领域需要访问相同的数据集，并且他们都相信这些数据集是完整的、最新的、一致的。主数据通常是这些数据集的基础。

管理数据质量：数据的不一致、质量问题和差异均会导致决策错误或丧失机会。主数据管理通过使用统一的标识来定义对组织至关重要的实体，以降低风险。

管理数据集成的成本：在没有主数据的情况下，将新数据源集成到一个已经很复杂的环境中成本会更高，这减少了因对关键实体的定义和识别方式的变化而产生的额外成本。比如学校的各种系统，竞赛系统，综测系统与教务系统的集成。

降低风险：主数据简化了数据共享架构，从而降低了与复杂环境相关的成本和风险。就是数据的解耦。

管理参考数据的驱动力也是如此。集中管理的参考数据使组织能够：管理参考数据的驱动力也是如此。集中管理的参考数据使组织能够：

通过使用一致的参考数据，满足多个项目的数据需求，降低数据整合的风险和成本。

评估和评价数据源

评估和评价不同数据源的数据质量和适用性，以确保数据来源准确、及时、完整和可靠。评估的内容包括从数据准确性、数据完整性、数据安全性、数据可用性以及数据更新和维护等5个角度。同时还需要评价数据的来源和采集方式，以提高数据源的适用性。

就我所知的教务系统来说，教务系统的数据准确性做到不好，培养树和培养计划不一致，有的课程根本没有但却出现在培养树上，而且有些学过的课程没法归入必修课中。



学分与课程数量不一致。

数据完整性:在查询成绩的时候，只能查询到总成绩，如果想看平时分和期末考试各考了多少分，是不能查看的，那为什么还要查看详情。

数据安全性:很差，在寒假的时候，通过特定的url，可以查询到全校所有人的学习成绩，可能还有没暴露出来的修改接口，这样的话华工学生的学习成绩的可信度就很受影响，整个学校的声誉也会受到影响，会影响到成绩的合法性，影响学生的保研和出国。

数据可用性:低，没有便捷的搜索功能，而且在期末周访问量大的时候会发生内部服务器错误，说明高可用性没有做好。而且很难搜索到想要的数据。

数据更新与维护:差，转专业的学生很多系统都没有跟着转过去，还是维持原专业的信息，在关系数据库理论中，这是没有做到第三范式的表现。系统过于混乱，系统之间的依赖也是一团乱麻。

识别主数据

在教务系统当中，识别主数据是用于唯一标识和区分不同数据实体的关键属性。以下是大学教务系统中常见的识别主数据：

学生学号：每个学生在教务系统中都会被分配一个唯一的学号，该学号用于区分不同的学生。

教师工号：每位教师在教务系统中都拥有一个独特的工号，用于识别和管理教师信息。

课程代码：每门课程都会有一个独立的课程代码，用于唯一标识和区分不同的课程。

班级代码：每个班级都会被分配一个独特的班级代码，用于准确识别和管理不同的班级。

考试编号：每次考试都会有一个唯一的考试编号，用于区分不同的考试和记录成绩信息。

学院代码：每个学院都拥有一个独特的代码，用于标识和区分不同的学院。

专业代码:每个专业拥有一个独立的代码，用来区分不同的专业。

通过识别主数据，教务系统能够准确地识别和管理各项教务信息，确保数据的准确性和一致性，并支持学校进行学生管理、课程安排、成绩记录等工作。

定义主数据架构

需要定义主数据的架构，包括数据模型、数据字典、元数据、代码表和规则等方面的内容。主数据的整合、规范化和管理需要结合现有系统、数据库和业务需求制定主数据管理策略。主数据架构要满足学校教学数据的特征和需求，为主数据的管理提供技术与方法的支持。

主数据的架构可以分为以下几个层次：

教务管理层级：该层级包括学院、专业和课程等信息。学院是指学校的二级组织单位，可以有多个学院。每个学院下面可以有若干专业，而每个专业又可以开设多门课程。这一层级的数据描述了学校的教学组织结构和课程设置情况。

学生管理层级：该层级包括学生信息、学籍管理和班级管理等。学生信息包括学生的个人信息、入学年份、学籍状态、学历层次等。学籍管理涉及到学生的入学、转专业、休学、复学、毕业等学籍变动情况。班级管理包括班级信息、班级人数、班主任等。

教师管理层级：该层级包括教师信息、授课管理和考试管理等。教师信息包括教师的个人信息、聘用关系和任课情况等。授课管理涉及到每位教师所授课程和对应的开课班级等。考试管理包括考试科目、考试时间、考场安排以及学生成绩管理等。

成绩管理层级：该层级包括学生考试成绩和绩点计算等。每门课程的成绩数据记录学生的得分情况，而绩点计算则基于成绩数据计算学生的绩点，并提供相关查询和统计功能。

教务服务层级：该层级包括与学生、教师和管理员相关的服务功能，例如选课功能、教务通知、成绩查询、课表查询、考试安排等。这些功能通过系统界面提供给用户使用，方便他们进行教务管理操作和获取信息。

建模主数据和相关参考数据

对主数据和相关参考数据建模，要考虑到其特点、属性和业务规则，建立数据元素的定义，建立数据元素间的关系模型和数据流模型，以保证主数据的可管理性。建模主数据也需要参考现有的标准和模型，确保主数据的现存的技术标准和法律规定等。

建模主数据和相关参考数据

在教务系统中，建模主数据和相关参考数据是用于描述和组织教务系统各个方面的关键数据元素。以下是大学教务系统中常见的主数据和相关参考数据：

学生主数据：包括学生的基本信息，如学生姓名、学号、性别、年龄、联系方式等。这些数据被用于识别和区分不同的学生，并建立他们与其他数据的关联。

教师主数据：涵盖教师的基本信息，例如教师姓名、工号、职称、所属部门、联系方式等。这些数据用于管理教师的人事信息以及与他们相关的教学和课程安排。

课程主数据：包括课程的基本信息，如课程名称、课程代码、学分、课程描述等。这些数据用于管理大学的课程目录和课程计划，以及为学生提供选课和课程安排的依据。

班级主数据：涉及班级的信息，如班级名称、班级代码、年级、专业等。这些数据用于组织学生的集体学习和管理班级的教学活动。考试成绩主数据：记录学生在各门课程考试中的成绩信息，包括课程名称、学生学号、考试成绩等。这些数据用于评估学生的学业表现和计算绩点。

学院主数据：涵盖大学的各个学院或学部的信息，如名称、代码、系主任等。这些数据用于组织和管理不同学院的教学和行政工作。

此外，相关参考数据是指为主数据提供补充信息和参考值的数据。例如：

教室参考数据：包括教室的编号、容量、位置等信息，用于安排课程和教学场地。

时间表参考数据：对应于学期的时间表，包括每门课程的上课时间和上课地点等信息，用于编排课程和学生的课程安排。

专业参考数据：涵盖不同专业的信息，包括专业名称、代码、培养方案等，用于帮助学生进行专业选修和培养计划。

通过建模主数据和相关参考数据，大学教务系统能够有效管理和利用重要的教务信息，支持学校的教学、课程安排、学生管理等各项工作。

定义教务元数据架构

教务元数据架构是指用于描述和组织大学或学校教务系统中的数据元素、属性和关系的结构。它提供了对教务系统中数据的统一定义和标准化，以便支持数据的共享、查询和管理。教务元数据架构通常包括以下组成部分：

实体：教务系统中的实际对象，如学生、教师、课程、班级等。每个实体都具有独特的属性来描述其特征和属性。

属性：用于描述实体的特征、属性或特性，例如学生的姓名、学号、性别、年龄等。属性提供了对实体更详细的描述和分类。

关系：表示不同实体之间的关联和连接，例如学生与课程之间的选课关系、教师与班级之间的授课关系等。关系描述了实体之间的交互和依赖关系。

数据类型：定义属性的数据类型，如字符串、整数、日期等。数据类型确保属性具有适当的格式和规范，以便正确地存储和处理数据。

数据约束：定义数据的有效性和一致性规则，包括唯一性约束、范围约束、引用完整性约束等。数据约束确保输入的数据符合预期的规范，并保持数据的完整性和准确性。

元数据：描述教务系统中各个数据元素和属性的信息，如名称、定义、类型、长度、取值范围等。元数据提供了对数据的描述和解释，帮助理解和使用数据。

提出主数据管理职责和维护流程建议

教务系统的主数据管理职责和维护流程对于确保教务系统数据的准确性和一致性非常重要。以下是一些建议的主数据管理职责和维护流程：

主数据管理职责：

确定负责主数据管理的部门或人员，明确其职责和权限。

确定教务系统中的关键主数据对象，如学生、教师、课程等，并定义其相关属性。

确保主数据的唯一性和标识性，例如通过分配唯一的ID或代码来区别不同的数据实体。

负责协调和管理主数据的录入、修改和删除过程，确保数据的完整性和正确性。

监测和解决主数据中的重复、冲突和错误，及时进行数据清理和修正。

维护流程建议：

设立数据管理规范和流程，明确数据维护的标准和要求。

建立数据录入和修改的审核机制，确保数据的合法性和准确性。

定期进行数据备份和恢复，以防止数据丢失或损坏。

建立数据访问权限控制机制，限制数据的访问和修改权限，确保数据安全性。

建立数据更新和同步机制，确保各个子系统或模块的数据一致性。

建立数据质量监控机制，定期进行数据质量评估和验证。

通过严格的主数据管理职责和维护流程，可以保证教务系统中的主数据的准确性、一致性和完整性，并提高教务管理的效率和质量。

需要制定主数据管理职责和维护流程标准，制定数据质量标准和评估方法，分配责任，明确教务数据的管理流程，并确保流程的执行和监控。同时，主数据管理需要与运营团队、IT部门和数据管理人员等协作，实现对教务数据的全流程管理。

主数据治理不是一蹴而就的过程，需要建立一套科学合理的教务数据管理方案，不断完善流程和方法，提高数据管理和维护的能力与水平，从而更好地应用和利用数据，助力学校教学管理。