

CET | 中教科
cnedutech.com

在数据治理实践中探究 教育大数据标准化需求

2021年12月25日

张洪斌

目录

CONTENTS



背景



研究教育大数据的主要视角及其影响



对教育大数据标准化的需求



一、背景

信息孤岛在学校信息化系统中普遍存在，根据20多所高职、中职院的情况统计，学校一般至少采用来源于五个以上企业开发的业务系统，由我公司承担信息化项目的某高职院校，同时运行近20个业务系统，系统提供商超过10家。

信息孤岛的主要危害

01

影响数据的有效性、一致性

02

难以进行跨部门的全局性数
据分析和统计

2021年4月起，江苏中教科信息技术有限公司实施数据治理项目，在多年开发、运维学校信息系统基础上，确定将数据治理作为主要业务方向。

目前，项目已经完成80%。



数据治理的实践过程迫切需要数据标准化

在实施数据治理项目中，“构建数据体系--建立数据模型”是核心环节，成为项目实施过程中最艰难的部分，最主要原因是缺乏数据标准化的支持。





必须充分认识教育大数据体系的内在规律，按主数据、业务数据、时序数据、分析数据分类处理数据，按数据域组织数据。

01



数据标准化是数据治理最重要的保障条件，许多问题源于缺乏标准化支持；

02



在技术上，充分利用技术元数据是确保软件实施数据治理过程的关键。

03

二、研究教育大数据的主要视角及其影响

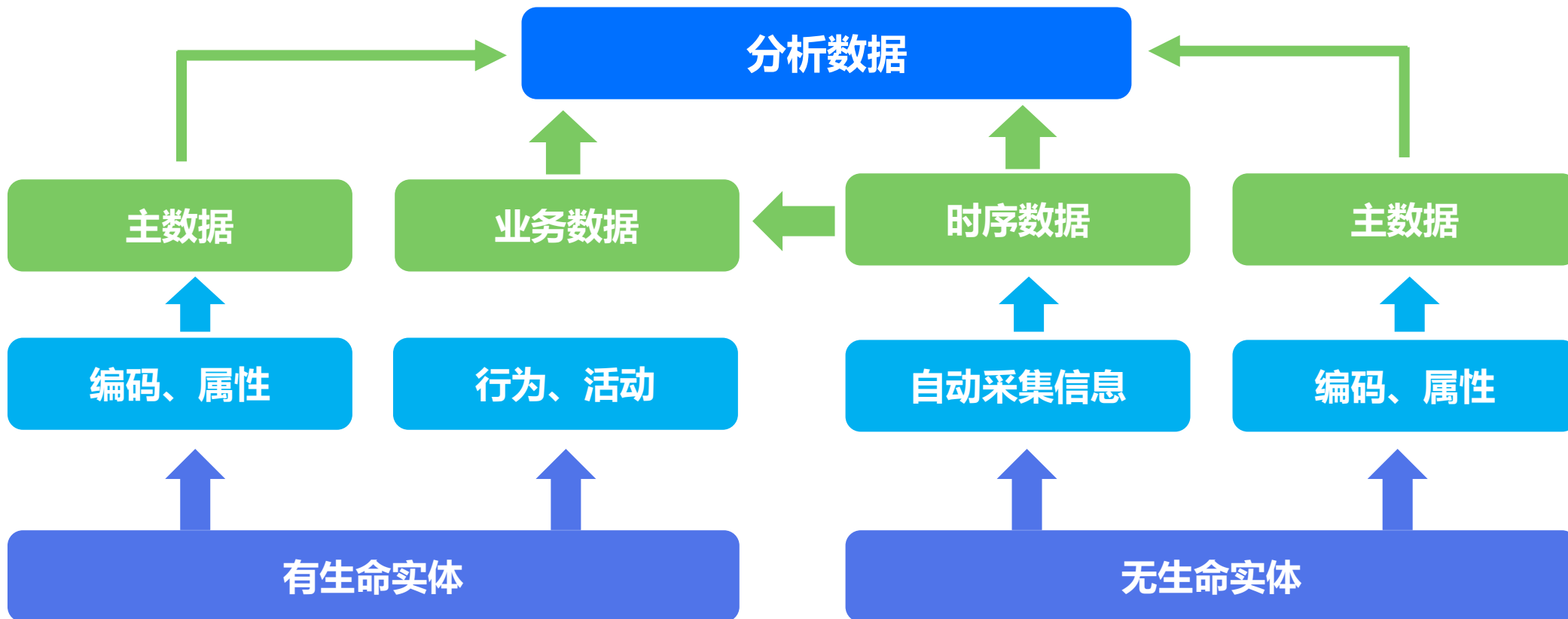
1、逻辑视角——研究者视角

(1) 数据分类：继承工业大数据的数据分类

类别	子类
元数据	业务元数据
	技术元数据
	操作元数据
主数据	参考数据
	基本信息
	编码
	业务数据
	时序数据
	分析数据

1、逻辑视角

大数据中不同类型数据之间的逻辑关系



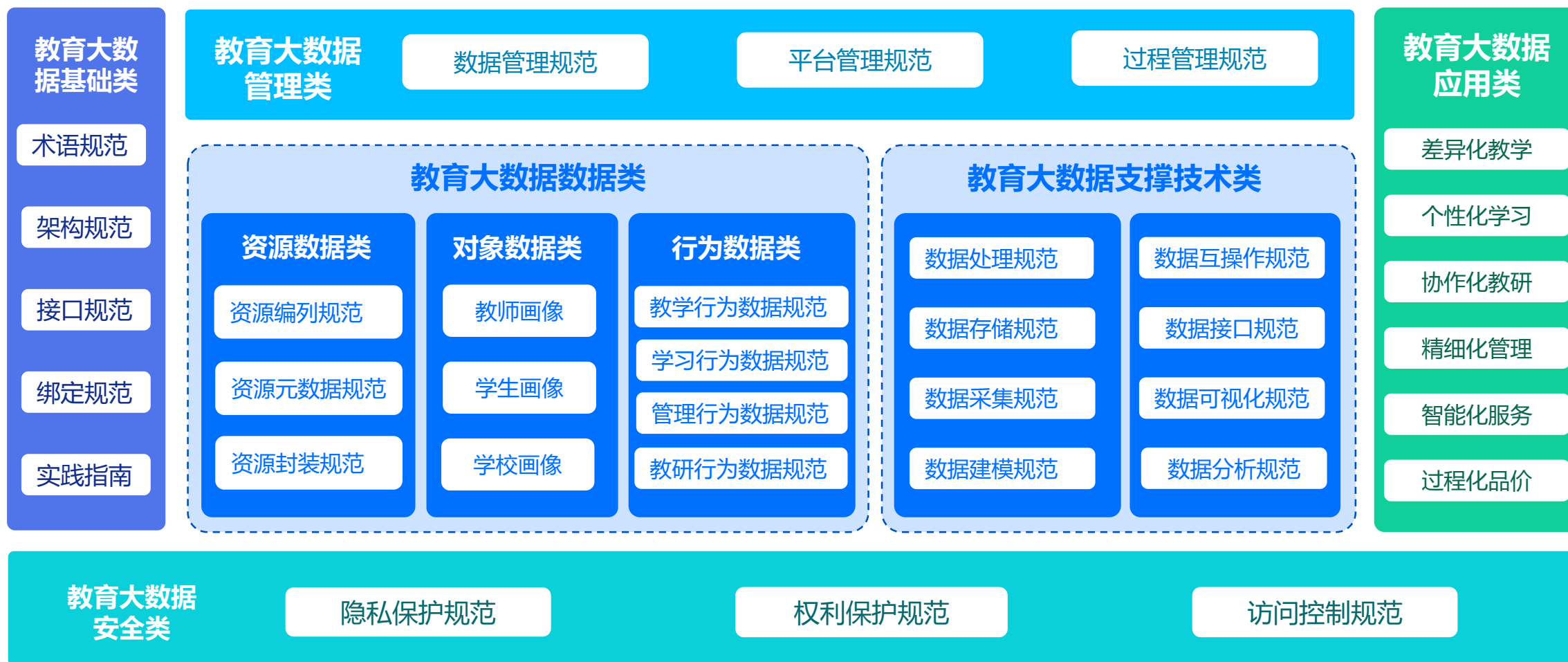
1、逻辑视角

教育大数据中的数据域

类型	实体和数据域
人员	学生、校友、教师、工勤人员、行政人员、家长、企业人员等
机构	学校、院系、职能部门、专业、专业群、班级、教研室、科室、企业、事业单位、政府部门等
教育方案和专项工作	课程、实验实训、校企合作、招生、迎新、就业、学生奖助、社团、创新创业、教师培训、科研、教研、顶岗实习、财务、党建、纪检监察、后勤服务、宿舍管理、资产管理、食堂等
教学资源 and 办学条件	场地(教学、文化、体育、公共(含绿化、停车场等))、建筑物、房间(教室、实验实训室、图书馆、宿舍、礼堂、体育馆、食堂、办公室等)、仪器设备、图书资料、智能传感器(摄像机、照相机、温度、湿度、空气、亮度、水表、电表、一卡通、区域入侵等)

1、逻辑视角

教育大数据标准体系框架（来源于国标）



1、逻辑视角

举例：源于《GB/T 38667-2020 信息技术 大数据 数据分类指南》

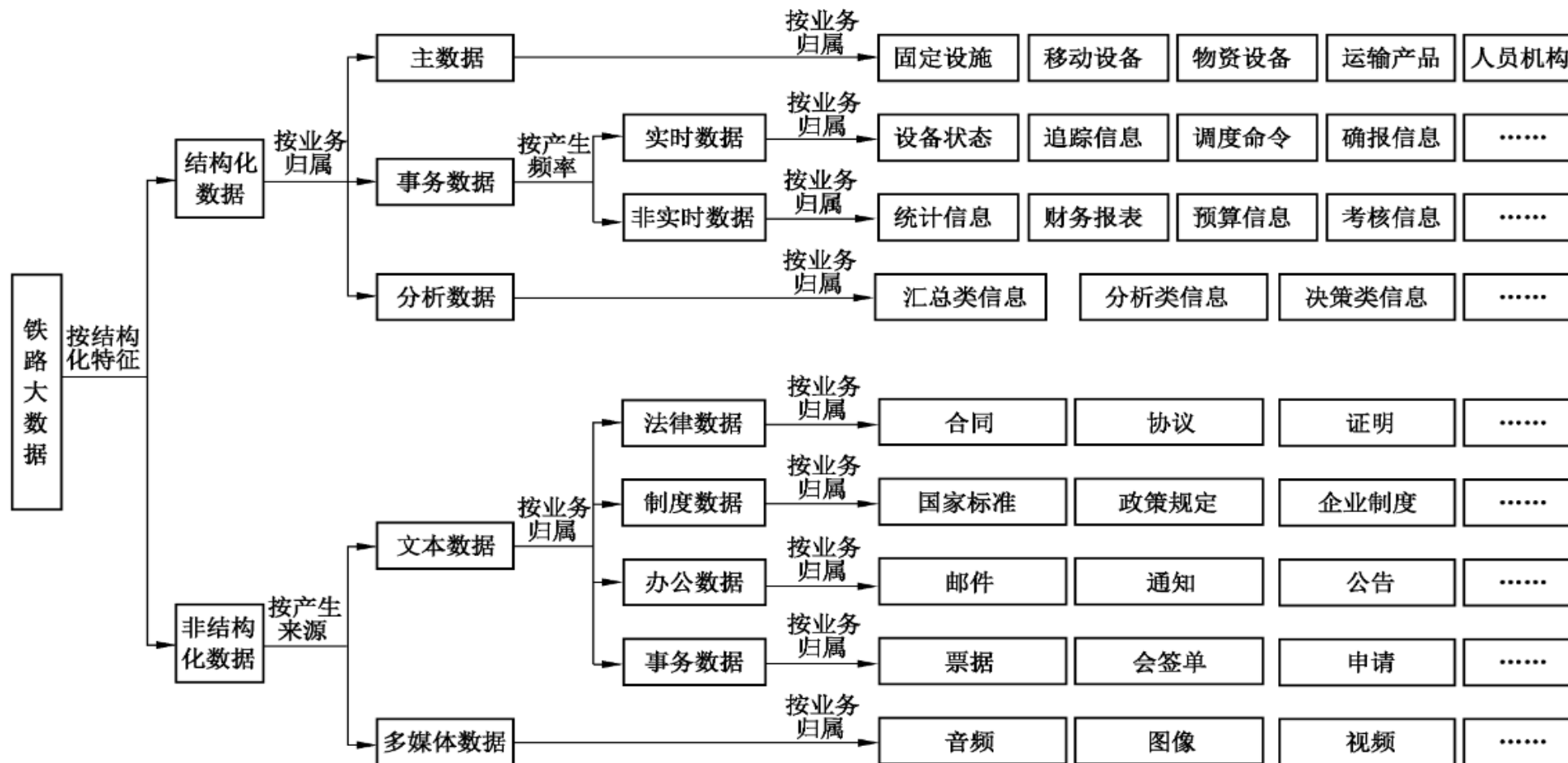


图 A.1 铁路大数据线分类方法示例

1、逻辑视角

相对于工业大数据，教育大数据具有以下主要特征：

数据域关系复杂

人是教育大数据主要实体，教育教学活动相关数据域具有特殊性，导致教育大数据数据域复杂

数据对象

结构化数据为主

时序数据

智能设备迅速增加，时序数据需转换处理后才有实用价值

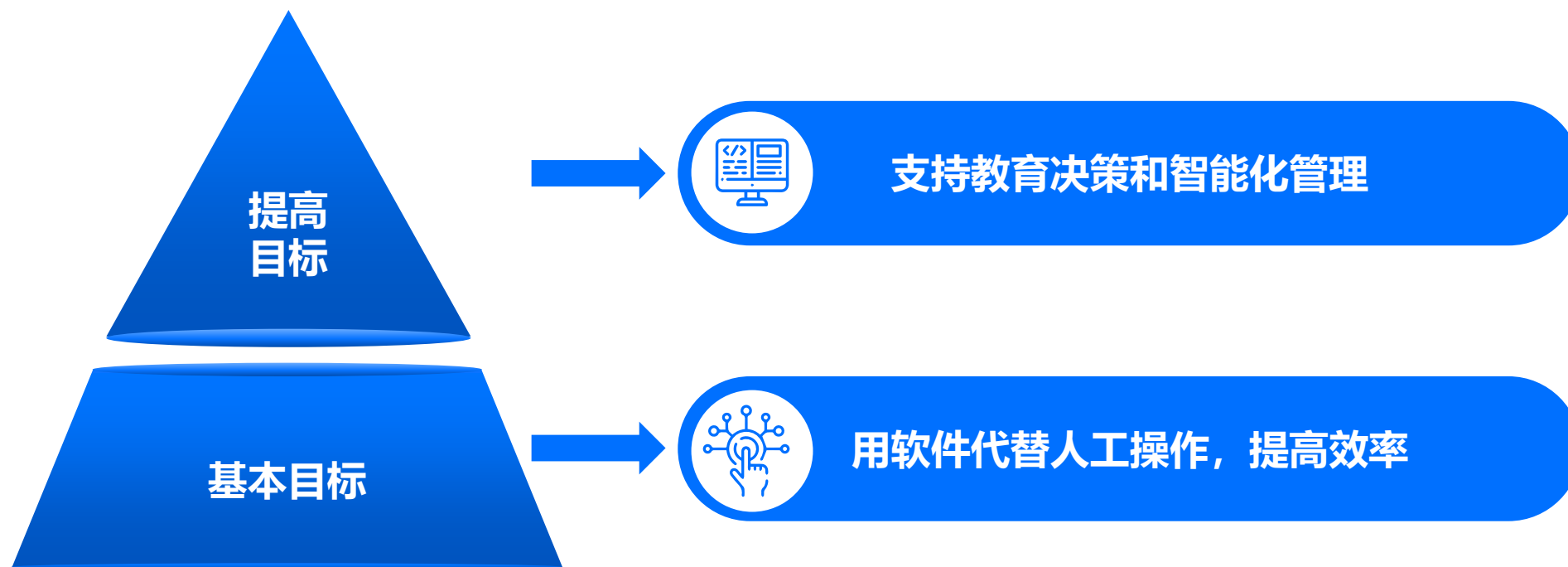
分析数据

目前主要按学期（学年）统计，对实时的动态分析和人工智能算法分析数据的需求迅速增加。



2、应用者视角

教育大数据的主要应用者是教育主管部门、学校、教师、学生（学生家长）和企业等。



2、应用者视角

逻辑思路：



视角

指标

功能

报表

流程

时序数据

分析数据

业务数据

主数据

3、开发者视角

软件企业按需求分析、概要设计、详细设计、测试、部署、运维的基本步骤开发系统，从开发者视角看教育大数据系统：



最看重软件功能和软件技术框架



满足于按学校业务流程要求开发系统，熟悉主数据和业务数据，并能够按算法要求计算分析数据



热衷采用智能部件和技术，但对其在教学活动中的实际作用缺乏认识；



对分析数据缺乏理解。

4、视角差异导致的问题

01

开发者和应用者对于教育大数据认识程度严重不对称，在确定系统需求期间就缺乏共同语言

02

有一定比例的教育大数据系统项目失败或远没有达到用户要求，反复修改；

03

上级教育主管部门和学校按条块提出项目需求，分期投入，各个业务系统之间分别设置主数据信息，导致信息孤岛普遍存在；

04

难以开发高质量的、充分发挥教育大数据优势的信息系统。



三、对教育大数据标准化的需求

1、标准化是统一开发者、应用者认识的关键

因为缺乏必要的标准：

01

信息化企业和教育主管部门、学校的沟通交流缺乏统一的参照系和权威性；

02

在实施数据治理过程中，在建立数据仓库时没有统一的数据编码和数据字段标准，步履维艰。

2、数据治理对标准的需求

现有与教育大数据相关标准

标准分类	现行	废止	被代替
学习、教育和培训	47		
大数据	16		
信息安全技术	255	13	20
软件开发和系统	158	16	6
信息技术服务	35		
合计	511	29	26

现有教育大数据标准

标准分类	数量
电子书包和课本	6
课程	5
数据规范	19
系统结构	7
虚拟实验	8
学习资源	2
合计	47

2、数据治理对标准的需求

需要增加的标准：

01

教育评价指标
体系的标准

02

课程教学活
动数据标准

03

教育教学专
项工作的数
据标准



2、数据治理对标准的需求

(1) 教育评价指标体系的标准

A、标准分两大类：

- 按数据域制定评价指标体系参考标准；
- 制定跨数据域评价指标体系参考标准。

B、教育评价指标体系标准的四个要素：

- 评价对象；
- 指标项树结构：指标项分为叶结点指标项和父结点指标项；
- 指标项权重；
- 与指标项对应的采集数据方法和算法（含分值确定算法）。

C、制定标准的基本方法：

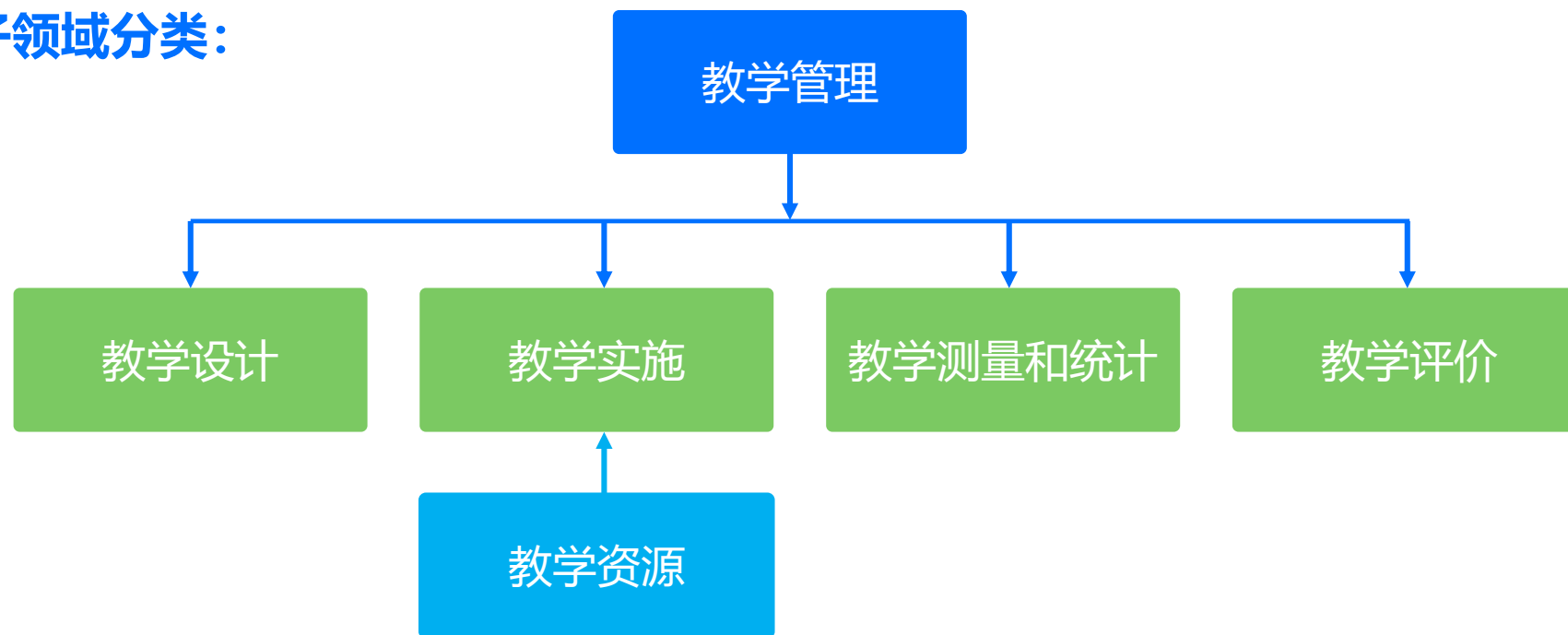
- CIPP法+与学校协商

2、数据治理对标准的需求

(2) 课程教学数据标准:

亟待建立能够覆盖传统教学、网络课程、精品课程、资源共享课程、在线课程等课程教学种类的课
程数据标准。

课程相关子领域分类:



2、数据治理对标准的需求

(2) 课程教学数据标准:

序号	分类	包含内容
1	教学设计	课程标准、教案、实验实训方案
2	教学资源	教学场地、教材(含电子化教材)、实验实训教材(含指导书)、参考书籍、辅助资料、授课PPT、网上资源(含视频)、教学设备、网上资源(含视频)
3	教学实施	教师授课方法手段、学生听课、自学和操作、师生交流、问卷调查、作业及批改、测验成绩、测验考试成绩、学生作品
4	教学测量与统计	教学环境数据测量、师生教学行为数据测量、测量数据统计分析
5	教学管理	课程编码、排课、教学检查、督导
6	教学评价	课程教学指标体系、课程指标体系指标项计算和生成

2、数据治理对标准的需求

(3) 专项工作的数据标准：

专项工作(如：**校企合作、招生、迎新、就业、学生奖助、学生社团、创新创业、教师培训、科研、教研、顶岗实习**等)的数据标准。

专项工作数据标准的要素：

- A、工作步骤及其顺序；
- B、描述各个步骤结果的数据表和字段；
- C、各个步骤的操作用户，要求可以在学校机构设置不同时灵活设置。

2、数据治理对标准的需求

专项工作的数据标准框架举例（顶岗实习）：

序号	步骤	要素	数据表字段
1	实习单位	基本信息、实习岗位、实习条件和实习待遇	企业信息(名称、代码、行业、规模、注册资金、地域(顶岗实习城市、县区))、实习岗位名称、实习岗位简要说明、实习薪资、保险、住宿条件等；操作的机构和用户
2	安排方案	学生顶岗实习岗位、预定实习待遇、校内指导教师	顶岗实习学生、编号、实习单位、实习地域、实习岗位、校内指导教师、预定实习薪资、预定保险、预定住宿条件、计划派遣时间；操作的机构和用户
3	实习派遣	学生到实习岗位时间、企业指导教师、实习待遇	学生、编号、企业名称、企业代码、实现地域、实际到企业时间、实际实习岗位、企业指导教师、实际实习薪资、购买保险、实际住宿条件；操作的机构和用户
4	岗位变动	岗位变动情况	学生、编号、原实习企业、原实习岗位、原薪资、原住宿条件，变动后实习企业、变动后实习岗位、变动后薪资、变动后住宿条件，岗位变动原因；操作的机构和用户
5	师生交流	学生问题、校内指导教师指导意见	学生问题(学生、编号、问题)，教师指导意见(教师、编号、指导意见)；操作的机构和用户
6	实习过程	学生实习过程主要工作情况	学生、编号、企业名称、企业编号、岗位名称、岗位工作简述、开始时间、结束时间、薪资、住宿；操作的机构和用户
7	顶岗实习成绩和评价	学生成绩、学生对顶岗实习评价、学校评价顶岗实习工作、企业评价学生、企业评价学校	学生顶岗实习成绩(学生、编号、成绩(百分制、等级制))；学生顶岗实习满意度(以班级、专业、院系、学校和在一个企业实习学生为口径统计)；学校评价顶岗实习工作(以班级、专业、院系和学校为统计口径)；企业评价学生(以在一个企业实习学生为口径)；企业对学校满意度(企业名称、编号、满意度(百分制、等级制))；操作的机构和用户。

CET | 中教科
cnedutech.com

谢谢

T H A N K S