

T/JNBDA

济南市大数据协会团体标准

T/JNBDA 0002-2025

数据分析实训平台建设规范

Standards for Data Analysis Training Platform Construction

2025 - 11 - 11 发布

2025 - 11 - 11 实施

济南市大数据协会发布

目 录

前 言	II
引 言	III
1 范围	1
2 规范性引用文件	1
3 术语和定义	1
3.1 数据	1
3.2 结构化数据	1
3.3 非结构化数据	1
3.4 半结构化数据	1
3.5 数据分析实训平台	2
3.6 实训案例	2
4 缩略语	2
5 总体要求	2
6 平台架构	2
7 功能要求	3
7.1 数据治理	3
7.2 数据中台	4
7.3 数据应用	4
7.4 教学实训	5

前 言

本规范由济南市大数据协会提出并归口，按照 GB/T 1.1—2020《标准化工作导则 第1部分：标准化文件的结构和起草规则》的规定起草。

请注意本文件的某些内容可能涉及专利。本文件的发布机构不承担识别专利的责任。

本规范由山东利诚大数据有限公司、山东劳动职业技术学院、山东远盾网络技术股份有限公司联合起草。

本规范主要参与单位：山东普诺勤数字科技有限公司、山东土地数字科技集团有限公司、华唐能源科技（山东）有限公司、济南新物联自控科技有限公司、莱芜职业技术学院、鲁担（山东）数据科技有限公司、山东北天极能源科技有限公司、山东海看新媒体研究院有限公司、泰山财产保险股份有限公司。

本文件主要起草人：耿俊、倪小兵、高冲、王黎明、于宁、曹虓、司佳、李华、占稳、李彦、郭振、张敏、王浩、张浩。

引 言

为贯彻落实国家关于大数据产业发展和职业教育实训基地建设的有关政策精神，推动数据分析人才培养与产业对接，特制定本规范。本规范由济南市大数据协会牵头组织，山东利诚大数据有限公司、山东劳动职业技术学院、山东远盾网络技术股份有限公司、山东普诺勤数字科技有限公司、山东土地数字科技集团有限公司、华唐能源科技（山东）有限公司、济南新物联自控科技有限公司、莱芜职业技术学院、鲁担（山东）数据科技有限公司、山东北天极能源科技有限公司、山东海看新媒体研究院有限公司、泰山财产保险股份有限公司等单位共同参与，旨在为数据分析实训平台的建设、管理、运行与评价提供技术依据和规范指导，促进实训平台的标准化、智能化与可持续发展。

数据分析实训平台建设规范

1 范围

本文件规定了数据分析实训平台相关术语和定义、功能要求、性能要求、安全要求、教学支持及验收评价等基本规则和要求。

本文件适用于高等院校、培训机构及企业建设数据分析类实训平台的规划、设计、开发、验收与运行管理。

2 规范性引用文件

下列文件中的内容通过文中的规范性引用而构成本文件必不可少的条款。其中，注日期的引用文件，仅该日期对应的版本适用于本规范；不注日期的引用文件，其最新版本（包括所有的修改单）适用于本规范。

GB/T 35274 信息安全技术 网络安全等级保护基本要求

GB/T 35295 信息安全 大数据 术语

GB/T 25000.10-2016 系统与软件工程 系统与软件质量要求和评价（SQuaRE） 第10部分：系统与软件质量模型

GB/T 37721-2019 信息技术 大数据分析系统功能要求

GB/T 5271.1-2000 信息技术 词汇 基本术语

3 术语和定义

下面术语和定义适用于本规范。

3.1 数据 data

信息的可再解释的形式化表示，以适用于通信、解释或处理。

注：可以通过人工或自动手段处理数据。

[来源：GB/T 5271.1-2000，定义 01.01.02]

3.2 结构化数据 structured data

存储在数据库里，可以用二维表结构表示的数据。

[来源：GB/T 37721-2019，定义 3.2]

3.3 非结构化数据 unstructured data

除了结构化数据之外的没有明确结构约束的数据。

[来源：GB/T 37721-2019，定义 3.3]

3.4 半结构化数据 Semi-structured Data

指介于结构化数据和非结构化数据之间，具有一定的结构化特征，但不完全符合结构化特征的数据。

3.5 数据分析实训平台 Data Analysis Training Platform

面向数据分析教学与实训需求，集成数据采集、处理、分析、可视化及项目管理等功能的一体化软件与硬件环境。

3.6 实训案例

基于真实或模拟业务场景设计的数据分析任务或项目，用于教学与实操训练。

4 缩略语

下列缩略语适用于本文件。

API：应用程序编程接口（Application Programming Interface）

ETL：数据提取、转换、加载（Extract, Transform, Load）

BI：商业智能（Business Intelligence）

SaaS：软件即服务（Software as a Service）

5 总体要求

本规范主要从以下 4 个方面对数据分析实训平台系统功能做出如下要求：

- a) 应支持多用户并发实训；
- b) 应提供统一身份认证与权限管理；
- c) 应支持多种数据源接入与集成；
- d) 应具备数据清洗、分析、可视化等功能，支撑开展教学管理、项目实训、考核评价等事项。

6 平台架构

宜包括数据治理、数据中台、数据应用、教学实训等层级，宜支持容器化、SaaS 化等现代信息技术架构，如图 1 所示。

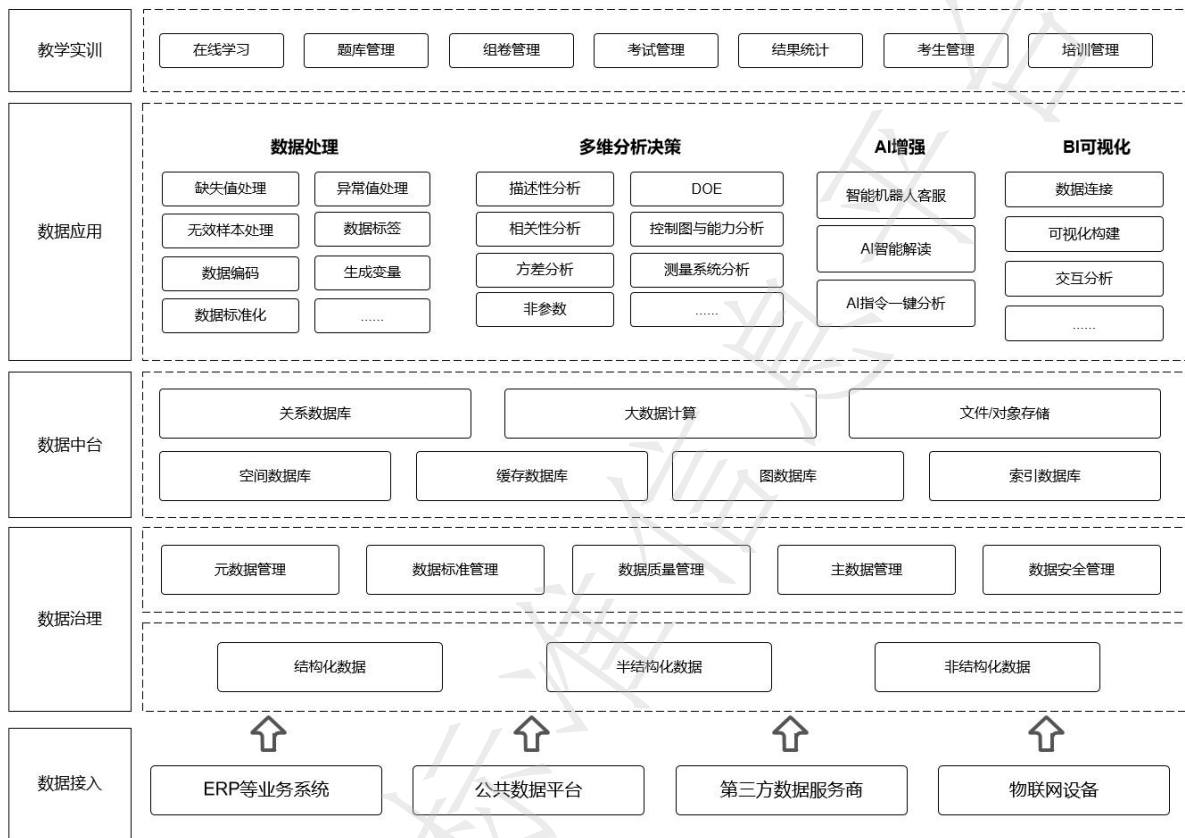


图1 数据分析实训平台系统架构图

7 功能要求

7.1 数据治理

为保障数据的规范性、准确性、唯一性、一致性、时效性和可访问性，对数据接入、元数据管理、数据质量等数据治理内容进行规范。

7.1.1 数据接入

应构建数据接入体系，包含数据本地上传、在线表格填写、数据库直连、数据中台汇聚等。实现数据的统一汇聚、存储，提升数据的存取效率。

7.1.2 元数据管理

元数据管理是数据治理的核心组成部分，通过数据中台，实现对元数据的全生命周期管理，以支持数据的可发现性、可理解性和可复用性。元数据管理的活动包括元数据识别、元模型构建、元数据创建整合、元数据变更管理和元数据应用等。通过数据中台进行元数据管理，提升数据资产的可信度与一致

性，降低数据冗余与歧义，增强数据溯源能力，提高数据检索与使用效率，推动企业级数据资产的沉淀与价值转化。

7.1.3 数据标准管理

数据标准管理是数据治理的关键环节，指通过统一规则对核心数据的定义、格式、质量及使用规范进行全局性约束，旨在消除数据孤岛、提升数据一致性并支撑跨系统协同。数据标准管理的主要活动包括数据标准分类设计、数据标准制定、数据标准发布、数据标准应用和数据标准维护。通过数据中台进行数据标准管理，可以确保数据在采集、存储、流转过程中的语义一致性与技术规范性，支撑跨部门、跨系统数据共享与互操作性，并满足行业监管要求。

7.1.4 数据质量管理

数据质量管理是数据治理的核心能力，旨在通过系统化方法确保数据在全生命周期中满足准确性、完整性、一致性、及时性等质量要求，从而为决策与运营提供可靠依据。数据质量管理的活动包括数据质量需求确认、数据质量检查、数据质量分析和数据质量提升。通过数据中台进行数据质量管理，可以保障数据在业务场景中的可信度，降低数据修复与返工成本，构建“预防-检测-纠正”的持续改进机制，形成数据质量治理闭环。

7.1.5 主数据管理

主数据管理是数据治理的关键领域，聚焦于对企业核心业务实体（如客户、供应商、产品、组织等）的高价值、跨系统共享数据进行统一管理，确保其在全企业范围内具有唯一性、准确性和一致性，支撑业务流程高效协同。主数据管理的活动宜包括主数据识别、主数据标准制定、主数据采集、主数据清洗、主数据集成、主数据应用管理。通过数据中台进行主数据管理，可以建立跨部门统一业务视图，避免因数据孤岛导致的运营低效，满足监管对主数据审计追溯的要求。

7.2 数据中台

数据中台能够直连企业业务系统，通过数据库直连或 ETL 工具抽取关系型数据库中的表数据，并可以连接电脑本地文件，实现多源异构数据的全量汇聚、标准化治理，构建统一数据资源池，确保底层数据的完整性、一致性和安全性。

7.3 数据应用

数据应用宜提供数据处理、数据分析、AI 增强、BI 可视化等功能。

7.3.1 数据处理

平台的数据处理方法通过异常值识别与修正、缺失值填充或删除、无效样本过滤等步骤确保数据完整性与可靠性，借助数据标签化、编码转换和衍生变量构建提升信息表达的有效性与分析维度，最终通

过标准化处理统一数据尺度，消除量纲差异影响，通过数据处理，为后续建模与分析提供高一致性、低噪声的优质数据基础，保障数据价值的精准挖掘与业务决策的科学性。

数据处理标准如下：

- a) 异常值处理：基于统计方法和业务规则识别并处理异常数据。
- b) 缺失值处理：采用插值、删除或预测填充等方法处理缺失值。
- c) 无效样本过滤：根据预设规则剔除不符合要求的数据样本。
- d) 数据标签：为数据添加语义标签增强可理解性。
- e) 数据编码：统一不同来源数据的编码格式。
- f) 生成变量：创建新的分析变量。
- g) 数据标准化：消除量纲差异影响。

7.3.2 数据分析

数据分析宜包含以下功能：

- a) 描述性分析（数据概览、正态性检验、频数分析、分类汇总等）；
- b) 差异性分析（单样本 T 检验、单样本 Z 检验、配对样本 T 检验、双样本 T 检验、单比率、双比率、单方差、双方差单因子方差分析、双因子方差分析、等方差检验、卡方拟合优度检验、卡方检验等）；
- c) 非参数检验（单样本符号、单样本 Wilcoxon、Mann-Whitney、游程检验等）；
- d) 相关回归分析（相关性分析、回归、逐步回归、拟合线图）；
- e) 时间序列（综合自回归移动平均等）；
- f) 机器学习（决策树、随机森林、AdaBoost、SVM 及 XGBoost 等）。

7.3.3 AI 增强

宜创建 AI 智能机器人、AI 智能解读、AI 指令一键分析，满足将模型输出自动转化为业务语言，降低数据分析门槛，使非专业用户也能快速完成数据清洗、数据分析、指标构建的全链路过程。

7.3.4 BI 可视化

支持商务数据分析，依托直观的拖拽式操作界面，无需编码即可快速搭建动态仪表盘。

- a) 支持用户将数据治理平台分析结果中的图表，直接拖入可视化画布，或者用户仅需选择分析结果描述，系统自动匹配最佳图表类型，并实时联动数据源更新，确保分析结果与业务动态同步。
- b) 支持多层次自定义配置：从图表样式（颜色、标签、坐标轴）到看板布局，均可按需求灵活调整；结合预设模板库与个性化组件（如动态筛选器、KPI 指标卡），可快速打造适配营销、运营、财务等场景的专属看板。
- c) 支持穿透式数据分析：BI 可视化看板的穿透式数据分析，可实现指标逐层下钻，快速定位问题和原因。

7.4 教学实训

平台宜具备在线学习、组卷、考试、分析、防作弊等功能，覆盖练测考评全场景。

7.4.1 在线学习

支持利用互联网技术，以时空分离为特征，建有学习资源和/或学习系统，配有服务人员和/或服务产品的学习服务。

[来源：ISO 29994；2021, 3.1. 有修改]

7.4.2 题库管理

- a) 支持无限层级管理目录、按知识点分类管理试题；
- b) 支持单选题、多选题、判断题、填空题、简答题；
- c) 支持通过 Excel 等模板一键批量快速导入试题；
- d) 支持批量导题查重及题库比对查重。

7.4.3 组卷管理

- a) 支持无限层级管理试卷目录、按业务场景管理试卷库；
- b) 支持人工选题组卷、固定抽题组卷、随机试题组卷；
- c) 支持按知识点、题型、难易程度等多条件精准抽题；
- d) 支持按考试时间限时、按固定时长限时；
- e) 支持启用/不启用合格分。

7.4.4 考试管理

- a) 支持正式考试、模拟考试，模拟考试考生可无限次模拟考试，回看历史考试结果；
- b) 支持考试结束自动交卷、答案自动缓存。

7.4.5 结果统计

支持按总分/总用时对个人成绩排名。

7.4.6 考生管理

- a) 组织架构管理：支持无限层级组织架构，解决部门或班级多的问题；
- b) 报名管理：考试报名审核成功后自动创建考生所属组织与人员账号；
- c) 考生账号：账号绑定手机号，支持账号密码/手机号验证码登录；
- d) 考生个人后台：报名审核成功后自动创建考生所属组织与人员账号。

7.4.7 培训管理

- a) 支持无限层级管理课件目录、按知识点分类管理课件资料；

- b) 支持管理培训课程，课程可关联课件资料、考试；
- c) 支持设置参与培训课程的学员名单，统计学员学习情况。

全国团体标准信息平台