

大数据治理

职业技能等级标准

标准代码：510063

(2021年2.0版)

成都云上天府大数据研究院有限公司 制定

2021年12月 发布

目 次

前言	1
1 范围	2
2 规范性引用文件	2
3 术语和定义	2
4 适用院校专业	3
5 面向职业岗位 (群)	4
6 职业技能要求	5
参考文献	13

前 言

本标准按照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。

本标准起草单位：成都云上天府大数据研究院有限公司、中国科学院虚拟经济与数据科学研究中心、中国科学院大数据挖掘与知识管理实验室、四川省中华职业教育社、川投信息产业集团有限公司、北京中科卓越未来教育科技有限公司、北京中科五极数据科技有限公司、晨越建设项目管理集团股份有限公司、联通大数据有限公司、新华中经信用管理有限公司、中核核电运行管理有限公司、国网四川省电力公司、中冶赛迪工程技术股份有限公司、大庆油田信息技术公司、中海油能源发展工程技术公司、北京协力筑成金融信息服务股份有限公司（36氩）、北京春雨天下软件有限公司、上海黑瞳信息技术有限公司、北京智数天下科技有限公司、博雅创智（天津）科技有限公司、北京光年无限科技有限公司（图灵机器人）、北京普开数据技术有限公司、成都汇誉通数据科技有限公司、中央财经大学管理科学与工程学院、常熟理工学院计算机科学与工程学院、苏州工业园区服务外包职业学院信息工程学院、乌鲁木齐职业大学、长春信息技术职业学院、天津市大学软件学院、四川大学锦城学院、成都电子信息学校、江苏省人工智能学会、天津市软件行业协会、天津市商务委员会信息中心。

本标准主要起草人：石勇、郭桂蓉、陈鲸、田英杰、赵景华、王波、王宏毅、赵越、龚俊中、寇纲、李建平、李爱华、蔡黎亚、赵宗萍、孟凡、张悦今、朱梅红、陈振松、李兴森、张强、张展、伏玉琛、陈永、衣马木艾山·阿布都力克木、郝文宁、王军、单杏花、李伟、宋雨伦、罗晓伊、刘旭嘉、吕爽、李方俊、曲晓慧、奚伟、曾柏毅、时维、陈阿古达木、刘鑫、俞志晨、蔡津津、徐厚东、刘建、王建、陈薛锦、张建勇、陆浩、王森、房伟、张云青。

声明：本标准的知识产权归属于成都云上天府大数据研究院有限公司，未经成都云上天府大数据研究院有限公司同意，不得印刷、销售。

1 范围

本标准规定了大数据治理职业技能等级对应的工作领域、工作任务及职业技能要求。

本标准适用于大数据治理职业技能培训、考核与评价，相关用人单位的人员聘用、培训与考核可参照使用。

2 规范性引用文件

下列文件对于本标准的应用是必不可少的。凡是注日期的引用文件，仅注日期的版本适用于本标准。凡是不注日期的引用文件，其最新版本适用于本标准。

GB/T 33745-2017 物联网 术语

GB/T 35274-2017 信息安全技术 大数据服务安全能力要求

GB/T 35295-2017 信息技术 大数据 术语

GB/T 34960.5-2018 信息技术服务 治理 第5部分:数据治理规范

GB/T 36073-2018 数据管理能力成熟度评估模型

GB/T 37721-2019 信息技术 大数据分析系统功能要求

GB/T 37722-2019 信息技术 大数据存储与处理系统功能要求

GB/T 37973-2019 信息安全技术 大数据安全管理指南

GB/T 37988-2019 信息安全技术 数据安全能力成熟度模型

GB/T 5271.1-2000 信息技术 词汇 第1部分:基本术语

GB/T 5271.17-2010 信息技术 词汇 第17部分:数据库

3 术语和定义

GB/T 5271.1-2000、GB/T 33745-2017、GB/T 37973-2019、GB/T 5271.17-2010 界定的以及下列术语和定义适用于本标准。

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。

[GB/T 5271.1—2000，定义01.01.02]

3.2

数据采集 data acquisition

利用数据采集工具，将各种来源和类型数据进行收集和集中的过程。

3.3

数据预处理 data preprocessing

在主要的分析处理之前对数据进行的一些处理。

3.4

数据挖掘 data mining

从大量的数据中通过算法搜索隐藏于其中信息的过程。

[GB/T 33745-2017，定义2.5.3]

3.5

数据可视化 data visualization

通过图形图像等形式对数据进行展示，以揭示数据特征的过程。

3.6

大数据 big data

具有体量巨大、种类多样、流动速度快、特征多变等特征，并且难以用传统数据体系结构和数据处理技术进行有效组织、存储、计算、分析和管理的数

据集。

[GB/T 37973-2019, 定义3.1]

3.7

大数据处理 big data processing

大数据的采集、存储、预处理、分析和挖掘、可视化等过程。

3.8

大数据治理 big data governance

把大数据作为组织、企业或者国家的战略资产进行管理，以及在管理流程中进行分析、控制和决策的活动集合。

3.9

数据标准 data standard

数据的命名、定义、结构和取值的规则。

3.10

元数据 metadata

关于数据或数据元素的数据（可能包括其数据描述），以及关于数据拥有权、存取路径、访问权和数据易变形的数据。

[GB/T 5271.17-2010, 定义17.06.05]

3.11

主数据 master data

组织中需要跨系统、跨部门进行共享的核心业务实体数据。

3.12

数据质量管理 data quality management

在数据全生命周期中的各个阶段，对组织中可能引发的各类数据质量问题进行识别、度量、监控、预警等一系列的管理活动。通过有效的数据质量管理控制，提高数据在各类应用中的价值。

3.13

大数据安全管理 big data security management

通过技术和管理手段来确保大数据环境下数据的保密性、完整性和可用性，以及分析并解决大数据活动对国家安全、社会影响、公共利益、个人的生命财产安全等造成的影响。

3.14

数据隐私保护 data privacy protection

对组织敏感的数据进行保护的措施。

3.15

数据融合 data fusion

一个整合多维度数据资源来产生比单个数据源更为连续的、准确的和有用的信息的过程。

3.16

数据共享管理 data sharing management

开展数据共享和交换，实现数据内外部价值的一系列活动。

4 适用院校专业

4.1 参照原版专业目录

中等职业学校：电子与信息技术、电子技术应用、计算机应用、软件与信息服务、移动应用技术与服务、通信运营服务、统计事务、客户信息服务、社

区公共事务管理。

高等职业学校：环境信息技术、城市信息化管理、电子信息工程技术、移动互联网应用技术、计算机应用技术、软件技术、软件与信息服务、电子商务技术、计算机信息管理、大数据技术与应用、人工智能技术服务、移动应用开发、财政、税务、互联网金融、信用管理、财务管理、会计、审计、会计信息管理、信息统计与分析、统计与会计核算、商务数据分析与应用、网络舆情监测。

高等职业教育本科学校：电子信息工程、计算机应用工程、软件工程、大数据技术与应用、金融管理、互联网金融、财务管理、会计。

应用型本科学校：经济统计学、数字经济、金融学、金融工程、信用管理、互联网金融、金融科技、数学与应用数学、信息与计算科学、统计学、应用统计学、电子信息工程、信息工程、人工智能、计算机科学与技术、软件工程、智能科学与技术、数据科学与大数据技术、信息管理与信息系统、大数据管理与应用、计算金融、公共事业管理、信息资源管理。

4.2 参照新版职业教育专业目录

中等职业学校：电子信息技术、电子技术应用、计算机应用、软件与信息服务、大数据技术应用、移动应用技术与服务、通信运营服务、统计事务、客户信息服务、社区公共事务管理。

高等职业学校：生态环境大数据技术、智慧城市管理技术、工业互联网应用、计量测试与应用技术、生物信息技术、电子信息工程技术、移动互联网应用技术、计算机应用技术、软件技术、大数据技术、人工智能技术应用、移动应用开发、健康大数据管理与服务、财税大数据应用、金融科技应用、信用管理、大数据与财务管理、大数据与会计、大数据与审计、会计信息管理、统计与大数据分析、统计与会计核算、市场调查与统计分析、商务数据分析与应用、网络舆情监测。

高等职业教育本科学校：电子信息工程技术、计算机应用工程、软件工程技术、大数据工程技术、人工智能工程技术、财税大数据应用、金融管理、金融科技应用、信用管理、大数据与财务管理、大数据与会计、大数据与审计、企业数字化管理。

应用型本科学校：经济统计学、数字经济、金融学、金融工程、信用管理、互联网金融、金融科技、数学与应用数学、信息与计算科学、统计学、应用统计学、电子信息工程、信息工程、人工智能、计算机科学与技术、软件工程、智能科学与技术、数据科学与大数据技术、信息管理与信息系统、大数据管理与应用、计算金融、公共事业管理、信息资源管理。

5 面向职业岗位（群）

【大数据治理】（初级）：主要面向企事业单位的大数据平台建设、系统开发、大数据运营、大数据服务等职业岗位，辅助完成大数据采集、大数据预处理、大数据可视化以及运用工具进行简单数据分析的工作。

【大数据治理】（中级）：主要面向企事业单位的大数据平台建设、系统开、大数据分析、大数据运营、大数据服务等职业岗位，主导完成大数据采集、大数据存储、大数据预处理、大数据可视化等工作，能够进行数据集成、数据分析和可视化，并完成基于数据分析结果的业务解释工作。

【大数据治理】（高级）：主要面向政府、企事业单位、大数据应用型企业的大数据平台建设、系统开发、大数据分析、大数据运营、大数据服务等相关职能部门及岗位，主导完成大数据预处理、大数据可视化、数据标准管理、

数据质量管理、数据安全、数据融合、数据共享等大数据治理工作。

6 职业技能要求

6.1 职业技能等级划分

大数据治理职业技能等级分为三个等级：初级、中级、高级，三个级别依次递进，高级别涵盖低级别职业技能要求。

【大数据治理】（初级）：能够根据业务情况，辅助进行大数据采集；能够根据大数据预处理需求，完成基本的大数据预处理工作；能够根据大数据可视化需求，运用数据分析和可视化工具，进行基本的分析，以及数据的静态和动态可视化展示；能够辅助进行数据标准的架构体系设计，能够进行元数据和主数据的创建、查询、维护等基本操作。

【大数据治理】（中级）：能够根据业务需求，针对复杂数据资源，编写定制化的数据采集代码并应用数据采集工具，完成数据采集；能够根据大数据预处理方案，进行基本的描述性统计分析，针对数据存在的缺失值、重复值、异常值、一致性问题，采用相应的方法进行处理；能够对不同数据源的数据实体进行识别，并进行数据集成；能够根据大数据可视化方案，采用基本的大数据分析语言、工具和算法，对加工后的数据进行分析，并完成对离线和在线数据的可视化；能够根据大数据治理需求，在企业、社会、政府等的大数据治理过程中，参与数据质量管理体系的建设。

【大数据治理】（高级）：能够根据业务需求，选定合适的大数据预处理工具和方法，进行多源异构数据的集成、归约和存储；能够根据大数据分析方案，选定合适的算法模型和编程语言，完成数据分析任务；能够根据大数据可视化方案，完成数据探索、数据分析过程和数据分析结果的可视化；能够根据业务需求，明确数据权限；能够针对敏感指标进行数据脱敏，保护数据安全；能够对多源异构数据进行融合；能够对大数据治理实施过程中产生的体系、方法、软件、模型等进行归纳整理，形成知识库；能够根据业务需要，进行大数据治理的材料编制。

6.2 职业技能等级要求描述

表 1 大数据治理职业技能等级要求（初级）

工作领域	工作任务	职业技能要求
1.大数据采集	1.1 大数据采集基础知识了解	<p>1.1.1 了解大数据的概念、来源和意义，了解大数据的特征和结构，了解大数据的表现形式和应用场景。</p> <p>1.1.2 熟悉大数据采集的概念，及其在大数据治理过程中的意义。</p> <p>1.1.3 熟悉大数据采集的数据来源、数据规模、数据类型和数据存储方式等。</p> <p>1.1.4 熟悉大数据采集的方式（数据录入、文本文件导入、数据库文件导入、网页爬虫等）及其基本原理。</p> <p>1.1.5 熟悉至少一种大数据采集通用工具（Python、Excel、Flume、Sqoop、八爪鱼等）的使用。</p>
	1.2 大数据采集方案制定	<p>1.2.1 能够根据业务情况，辅助进行大数据采集业务需求调研。</p> <p>1.2.2 能够根据数据采集需求，辅助进行数据资源的</p>

		<p>梳理。</p> <p>1.2.3 能够根据梳理好的数据资源，辅助确定数据采集的方法和工具。</p> <p>1.2.4 能够辅助编写详细的大数据采集方案。</p>
	1.3 大数据采集实施	<p>1.3.1 能够进行数据的快速录入。</p> <p>1.3.2 能够进行至少一种常用数据库（Access、Mysql、MongoDB、Oracle 等）的快速导入。</p> <p>1.3.3 能够基于录入和导入的数据，进行增、删、改、查等基本操作。</p> <p>1.3.4 能够掌握基本的数据查询技巧。</p> <p>1.3.5 能够按照数据标准进行数据内容的规范与处理。</p>
2.大数据预处理	2.1 大数据预处理基础知识了解	<p>2.1.1 了解大数据预处理的概念、环节，及其在大数据处理过程中的意义。</p> <p>2.1.2 熟悉数据清洗的概念和意义，并能够进行去重、空值处理等基本的数据清洗。</p> <p>2.1.3 熟悉数据集成的概念和意义，并能够进行数据合并、数据表关联等简单的数据集成。</p> <p>2.1.4 熟悉至少一种大数据预处理工具（Excel、Python、Sql、存储过程、Kettle 等）的使用。</p>
	2.2 大数据预处理方案制定	<p>2.2.1 能够理解大数据预处理需求。</p> <p>2.2.2 能够根据大数据预处理需求，梳理不同类型和结构的数据。</p> <p>2.2.3 能够针对至少一种类型和结构的数据，选定合适的大数据预处理工具和方法。</p> <p>2.2.4 能够辅助编写大数据预处理方案。</p>
	2.3 大数据预处理实施	<p>2.3.1 能够掌握至少一种类型数据文件（txt、csv、html、xml、xls 等）的加载。</p> <p>2.3.2 能够进行数据的排序、筛选和分类汇总。</p> <p>2.3.3 能够处理数据中存在的格式问题。</p> <p>2.3.4 能够处理数据中存在的 content 问题。</p> <p>2.3.5 能够进行数据列的合并与拆分。</p>
3.大数据可视化	3.1 大数据可视化基础知识了解	<p>3.1.1 了解大数据可视化的概念。</p> <p>3.1.2 了解可视化在大数据治理过程中的意义。</p> <p>3.1.3 了解大数据可视化的形式和应用。</p> <p>3.1.4 熟悉至少一种大数据可视化工具（Excel、Tableau、Python、R 等）的使用。</p>
	3.2 大数据可视化方案制定	<p>3.2.1 能够理解大数据可视化需求。</p> <p>3.2.2 能够进行基本的描述性和探索性数据分析。</p> <p>3.2.3 能够理解数据可视化的目标。</p> <p>3.2.4 能够辅助编写大数据可视化方案。</p>
	3.3 大数据可视化	<p>3.3.1 能够进行数据可视化的创建、编辑和布局等。</p> <p>3.3.2 能够使用大数据可视化工具，实现数据探索过程的图像化静态和动态展现。</p>

		<p>3.3.3 能够使用大数据可视化工具，实现数据分析结果的图像化静态和动态展现。</p> <p>3.3.4 能够辅助编写大数据可视化报告。</p>
4.大数据治理	4.1 大数据治理基础知识了解	<p>4.1.1 了解大数据治理的概念和关键要素。</p> <p>4.1.2 了解大数据治理的对象和内容。</p> <p>4.1.3 了解大数据治理的框架和核心准则。</p> <p>4.1.4 了解大数据治理的实施保障。</p>
	4.2 数据标准体系建设	<p>4.2.1 了解数据标准的定义和内容。</p> <p>4.2.2 了解数据标准管理的组织架构和设计流程。</p> <p>4.2.3 了解数据标准的制定和执行流程。</p> <p>4.2.4 能够根据数据标准设计的要求，进行数据标准化。</p> <p>4.2.5 能够根据业务需求，辅助进行数据标准的架构体系设计。</p> <p>4.2.6 能够在企业、社会、政府等的大数据治理过程中，参与数据标准体系的建设。</p>
	4.3 元数据管理	<p>4.3.1 了解元数据的概念、类型和应用。</p> <p>4.3.2 了解元数据管理的概念和意义。</p> <p>4.3.3 了解元数据的生成、获取和维护。</p> <p>4.3.4 了解元数据的存储、整合和共享。</p> <p>4.3.5 能够进行元数据的创建、查询和维护等基本操作。</p>
	4.4 主数据管理	<p>4.4.1 了解主数据的概念、类型和应用。</p> <p>4.4.2 了解主数据管理的概念和意义。</p> <p>4.4.3 了解主数据的获取和维护。</p> <p>4.4.4 了解主数据的存储、集成和共享。</p> <p>4.4.5 能够根据业务实施方案，进行主数据的创建、查询和维护等基本操作。</p>

表 2 大数据治理职业技能等级要求（中级）

工作领域	工作任务	职业技能要求
1. 大数据采集	1.1 大数据采集知识与技能掌握	<p>1.1.1 掌握大数据采集的概念，及其在大数据治理过程中的意义。</p> <p>1.1.2 掌握大数据采集过程中数据来源、数据规模、数据类型、数据存储方式等出现的难点。</p> <p>1.1.3 掌握三种以上大数据采集的方式（数据录入、文本文件导入、数据库文件导入、网页爬虫等）及其基本原理。</p> <p>1.1.4 掌握至少一种大数据采集通用工具（Python、Excel、Flume、Sqoop、八爪鱼等）的使用。</p>
	1.2 大数据	1.2.1 能够根据业务情况，进行大数据采集业务需求

	采集方案制定	<p>调研，明确数据采集需求。</p> <p>1.2.2 能够根据数据采集需求，进行数据资源的梳理。</p> <p>1.2.3 能够根据梳理好的数据资源，确定数据采集指标要求，并完成指标的说明及规范。</p> <p>1.2.4 能够根据梳理好的数据资源，确定数据采集的方式和工具。</p> <p>1.2.5 能够制定详细的大数据采集方案。</p>
	1.3 大数据采集实施	<p>1.3.1 能够根据大数据采集方案，组织实施大数据采集任务。</p> <p>1.3.2 针对复杂数据资源，能够编写定制化的数据采集代码并应用数据采集工具，完成数据采集。</p> <p>1.3.3 能够指导初级人员完成基本的数据采集工作。</p> <p>1.3.4 能够根据数据质量及安全保障规范要求，保障数据质量及安全。</p>
2. 大数据预处理	2.1 大数据预处理知识与技能掌握	<p>2.1.1 掌握大数据预处理的概念，及其在大数据治理过程中的意义。</p> <p>2.1.2 掌握大数据预处理的各个环节和内容。</p> <p>2.1.3 掌握大数据预处理各环节（数据清洗、数据集成、数据归约、数据变换等）的基本操作。</p> <p>2.1.4 掌握至少一种大数据开发的编程语言（Python、Java、存储过程等），解决基本的数据预处理问题。</p> <p>2.1.5 掌握至少一种大数据预处理通用工具（Excel、Python、Sql、存储过程、Kettle 等）的使用。</p>
	2.2 大数据预处理方案制定	<p>2.2.1 能够根据业务需求，初步明确大数据预处理需求。</p> <p>2.2.2 能够进行基本的描述性统计分析，了解数据整体情况，明确需预处理的数据。</p> <p>2.2.3 能够根据数据整体情况，初步选定合适的大数据预处理工具和方法。</p> <p>2.2.4 能够初步制定大数据预处理方案。</p>
	2.3 大数据预处理实施	<p>2.3.1 能够掌握不同类型数据文件（txt、csv、html、xml、xls 等）的加载。</p> <p>2.3.2 能够针对数据存在的缺失值、重复值、异常值、一致性问题，采用相应的方法进行处理。</p> <p>2.3.3 能够根据大数据预处理方案，对数据进行预处理。</p> <p>2.3.4 能够对不同数据源的数据实体进行识别，并进行数据集成。</p>
3. 大数据可视化	3.1 大数据可视化知识与技能掌握	<p>3.1.1 掌握大数据可视化的概念</p> <p>3.1.2 掌握可视化在大数据治理过程中的意义。</p> <p>3.1.3 掌握大数据可视化的形式和应用。</p> <p>3.1.4 掌握至少一种大数据可视化工具（Excel、</p>

		Tableau、Python、R 等) 的使用。
	3.2 大数据可视化方案制定	3.2.1 能够根据业务需求, 初步明确大数据可视化需求。 3.2.2 能够根据业务需求, 进行描述性和探索性数据分析。 3.2.3 能够根据数据整体情况和大数据可视化需求, 初步明确基本的数据可视化目标。 3.2.4 能够初步制定大数据可视化方案。
	3.3 大数据可视化	3.3.1 能够运用合适的可视化工具, 完成离线数据探索和分析结果的可视化。 3.3.2 能够运用合适的可视化工具, 完成在线数据探索和分析结果的可视化。 3.3.3 能够基于数据可视化结果进行基本的分析和解释。 3.3.4 能够初步编写大数据可视化报告。
4 大数据治理	4.1 大数据治理知识与技能掌握	4.1.1 掌握大数据治理的概念和关键要素。 4.1.2 掌握大数据治理的对象、内容、框架和核心准则。 4.1.3 熟悉至少一种大数据治理的相关工具 (ERWin、Excel、SVN、禅道等)、技术支撑和实施保障。 4.1.4 熟悉大数据治理的产业发展和应用场景。
	4.2 数据质量管理	4.2.1 掌握大数据质量管理的概念和框架。 4.2.2 掌握影响数据质量的因素。 4.2.3 掌握数据质量评价模型和体系。 4.2.4 掌握数据质量过程控制。 4.2.5 掌握至少一种数据质量管理工具 (Excel、存储过程、禅道等) 的使用。 4.2.6 能够在企业、社会、政府等的大数据治理过程中, 参与数据质量管理体系的建设。
	4.3 大数据治理材料编制	4.3.1 能够根据业务需求, 初步进行大数据治理的需求分析。 4.3.2 能够辅助制定数据标准管理的相关规范和制度。 4.3.3 能够辅助制定数据质量管理的相关规范和制度。 4.3.4 能够辅助进行大数据治理材料的编制。

表 3 大数据治理职业技能等级要求 (高级)

工作领域	工作任务	职业技能要求
1. 大数据预处理	1.1 大数据预处理知识	1.1.1 掌握大数据的概念、来源和意义, 掌握大数据的特征和结构, 掌握大数据的表现形式和应用场景。

	与技能掌握	<p>1.1.2 掌握描述性统计分析、数据预处理的原因和方法。</p> <p>1.1.3 掌握并能够实现数据清洗、数据集成、数据归约和数据变换。</p> <p>1.1.4 熟练掌握至少一种大数据开发的编程语言（Python、Java、存储过程等），解决基本的数据预处理问题。</p> <p>1.1.5 熟练掌握至少一种大数据预处理通用工具（Excel、Python、Sql、存储过程、Kettle 等）的使用。</p>
	1.2 大数据预处理方案制定	<p>1.2.1 能够采用有效的沟通方式，与客户进行交流，明确业务需求。</p> <p>1.2.2 能够根据业务需求，明确大数据预处理需求。</p> <p>1.2.3 能够进行数据统计分析，掌握数据整体情况。</p> <p>1.2.4 能够根据分析需求和数据整体情况，选定合适的大数据预处理工具和方法。</p> <p>1.2.5 能够制定详细的大数据预处理方案。</p>
	1.3 大数据预处理实施	<p>1.3.1 能够根据大数据预处理方案，组织实施大数据预处理任务。</p> <p>1.3.2 能够根据大数据预处理方案，进行数据描述性统计分析，发掘噪声、离群点等异常值，进行数据预处理。</p> <p>1.3.3 能够将多个数据源的数据进行集成和统一存储。</p> <p>1.3.4 能够对数据进行维归约、数量归约和数据压缩等处理。</p> <p>1.3.5 能够采用光滑、聚集、规范化和离散化等方法，将数据转化为更适于分析的形式。</p> <p>1.3.6 能够指导初级和中级的人员完成大数据预处理工作。</p>
2.大数据分析及可视化	2.1 大数据分析及可视化知识与技能掌握	<p>2.1.1 掌握数据挖掘的概念、过程、功能和应用领域。</p> <p>2.1.2 掌握关联规则、聚类、分类和预测等的基本概念。</p> <p>2.1.3 掌握相关算法的基本思想和优缺点，能够进行实际问题的分析。</p> <p>2.1.4 掌握至少一种主流数据挖掘工具（Excel、Python、SPSS、SAS、Weka、Matlab 等）的使用，能够应用工具进行数据分析及可视化。</p>
	2.2 大数据分析及可视化方案制定	<p>2.2.1 能够根据业务需求，明确大数据分析及可视化需求。</p> <p>2.2.2 能够根据大数据分析的需求，对数据分析对象进行梳理，明确数据分析的任务。</p> <p>2.2.3 能够根据数据分析的任务，明确数据分析方</p>

		<p>法。</p> <p>2.2.4 能够根据数据整体情况和数据分析任务，明确数据可视化的目标。</p> <p>2.2.5 能够制定详细的大数据分析与可视化方案。</p>
	2.3 大数据分析	<p>2.3.1 能够根据大数据分析方案，组织实施大数据分析任务。</p> <p>2.3.2 能够根据数据分析的任务和方法，明确合适的算法模型和编程语言。</p> <p>2.3.3 能够针对选定的算法模型完成代码编写，完成数据分析任务。</p> <p>2.3.4 能够结合业务需求，对数据分析的结果进行全面的评估。</p> <p>2.3.5 能够根据评估后的结果，向客户进行业务解释。</p> <p>2.3.6 能够编写详细的大数据分析报告。</p>
	2.4 大数据可视化	<p>2.4.1 能够根据大数据可视化方案，组织实施大数据可视化任务。</p> <p>2.4.2 能够根据数据可视化目标，明确合适的可视化语言和工具。</p> <p>2.4.3 能够进行数据探索、数据分析过程和数据分析结果的可视化。</p> <p>2.4.4 能够对数据可视化结果进行分析、解释和优化。</p> <p>2.4.5 能够编写详细的大数据可视化报告。</p>
3.大数据治理	3.1 大数据治理知识与技能掌握	<p>3.1.1 掌握大数据治理的主要方法及其在相关领域的应用。</p> <p>3.1.2 掌握至少一种大数据治理的相关工具（ERWin、Excel、SVN、禅道等）、技术支撑和实施保障。</p> <p>3.1.3 掌握领域知识的概念和表达。</p> <p>3.1.4 掌握 CRISP-DM 模型。</p> <p>3.1.5 能够指导初级、中级人员进行企业、社会、政府等的大数据治理实施。</p>
	3.2 大数据安全管理与隐私保护	<p>3.2.1 掌握大数据安全管理与隐私保护的概念，及其在大数据治理中的意义。</p> <p>3.2.2 掌握与大数据安全和隐私保护相关的法律法规、制度和标准。</p> <p>3.2.3 熟悉大数据安全防护技术，包括数据发布匿名保护、社交网络匿名保护、数据溯源技术和访问控制技术。</p> <p>3.2.4 熟悉大数据隐私保护技术，包括威胁发现技术、大数据认证技术、数据真实性分析技术、数据失真处理技术、数据加密技术和限制发布技术等。</p> <p>3.2.5 能够根据业务需求明确数据权限，规范不同用户对不同数据的权限范围。</p>

		3.2.6 能够识别数据中的敏感指标，针对敏感指标进行数据脱敏，保护数据安全。
	3.3 大数据融合与共享管理	3.3.1 掌握数据融合的概念，及其在大数据治理中的意义。 3.3.2 掌握数据共享管理的概念，及其在大数据治理中的意义。 3.3.3 了解对组织内部和组织外部的不同来源数据进行融合的基本方法。 3.3.4 了解对结构化、半结构化和非结构化数据进行融合的基本方法。 3.3.5 能够根据业务需求，明确数据共享的类型、方式、内容、对象和条件。
4.大数据治理实施与材料编制	4.1 大数据治理方案制定	4.1.1 能够根据业务情况，进行大数据治理的需求分析。 4.1.2 能够根据需求分析，编制大数据治理方案。包括大数据采集方案、大数据预处理方案、大数据分析方案、大数据可视化方案、数据标准管理方案、数据质量管理方案、数据安全方案、数据融合与共享管理方案等。 4.1.3 能够指导初级、中级人员完成相关方案的制定和编写。
	4.2 大数据处理与治理实施	4.2.1 能够根据大数据治理方案，进行数据采集、数据预处理、分析与可视化、数据标准管理、数据质量管理、数据安全方案、数据融合与共享管理等大数据治理工作。 4.2.2 能够与需求方进行有效沟通，解决大数据治理实施中出现的各类问题。 4.2.3 能够有效组织团队共同完成大数据治理实施工作。
	4.3 材料编制	4.3.1 能够对大数据治理实施过程中产生的体系、方法、软件、模型等进行归纳整理，形成知识库。 4.3.2 能够根据大数据治理的结果，撰写大数据治理技术报告。 4.3.3 能够根据业务需要，撰写大数据治理相关的标准、规范及制度。 4.3.4 能够指导初级、中级人员完成相关材料的编制。

参考文献

- [1] GB/T 35295-2017信息技术 大数据 术语
- [2] GB/T 37721-2019信息技术 大数据分析系统功能要求
- [3] GB/T 37722-2019信息技术 大数据存储与处理系统功能要求
- [4] GB/T 37973-2019信息安全技术 大数据安全管理指南
- [5] GB/T 34960.5-2018 信息技术服务 治理 第5部分:数据治理规范
- [6] GB/T 35274-2017信息安全技术 大数据服务安全能力要求
- [7] GB/T 37988-2019信息安全技术 数据安全能力成熟度模型
- [8] GB/T 36073-2018 数据管理能力成熟度评估模型
- [9] GB/T 5271.1-2000 信息技术 词汇 第1部分：基本术语
- [10] GB/T 5271.17-2010 信息技术 词汇 第17部分：数据库
- [11] GB/T 33745-2017 物联网 术语
- [12] 数据资产管理实践白皮书4.0
- [13] 中等职业学校专业教学标准
- [14] 高等职业学校专业教学标准
- [15] 普通高等学校本科专业类教学质量国家标准
- [16] 教育部关于印发《职业教育专业目录（2021年）》的通知（教职成〔2021〕2号）
- [17] 《教育部关于公布2019年度普通高等学校本科专业备案和审批结果的通知》（教高函〔2020〕2号）
- [18] 《教育部关于公布2020年度普通高等学校本科专业备案和审批结果的通知》（教高函〔2021〕1号）