



T/MIITEC 006-2021

大数据产业人才岗位能力要求

Big Data Industrial Talents Competency Framework

2021-10-21 发布

2021-10-21 实施

工业和信息化部人才交流中心 发布

目 次

前 言.....	III
引 言.....	IV
1 范围.....	1
2 规范性引用文件.....	1
3 术语和定义.....	1
4 大数据主要方向及岗位.....	2
4.1 主要方向.....	2
4.2 主要岗位及职责.....	3
5 大数据产业人才岗位能力要素.....	4
6 大数据产业人才岗位能力要求.....	5
6.1 大数据预处理岗位能力要求.....	5
6.2 大数据标注岗位能力要求.....	7
6.3 大数据分析岗位能力要求.....	9
6.4 大数据产品开发岗位能力要求.....	11
6.5 大数据项目实施与运维岗位能力要求.....	14
6.6 大数据平台建设岗位能力要求.....	16
6.7 大数据安全岗位能力要求.....	18
6.8 大数据管理岗位能力要求.....	19
6.9 大数据运营与应用岗位能力要求.....	20
6.10 大数据咨询服务岗位能力要求.....	22
附录 A（资料性附录） 大数据产业人才岗位能力提升.....	24
A.1 大数据产业人才岗位能力提升内容.....	24
A.2 大数据产业人才岗位能力提升阶段和方式.....	24
A.3 大数据产业人才岗位能力提升活动供给类别.....	24
附录 B（资料性附录） 大数据产业人才岗位能力评价.....	26
B.1 大数据产业人才岗位能力评价方法.....	26
B.2 大数据产业人才岗位能力评价等级.....	26
B.3 大数据产业人才岗位能力等级评价权重.....	26
参考文献.....	27

前 言

本文件按照 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。

本标准由工业和信息化部人才交流中心提出并归口。

本标准起草单位：工业和信息化部人才交流中心、中联研究院有限公司、国家工业信息安全发展研究中心、电科云（北京）科技有限公司、中国电信学院、中电长城网际系统应用有限公司、北京仁略信息技术有限公司、北京国软工程咨询有限公司、北京百度网讯科技有限公司、中国联合网络通信有限公司软件研究院、中国航空制造技术研究院、北京平高清大科技发展有限公司、长城计算机软件与系统有限公司、北京纷扬科技有限责任公司、信科创智（北京）网络科技有限公司、中国联合网络通信有限公司北京市分公司、国基北盛（南京）科技发展有限公司、中国电力企业联合会科技开发服务中心、数据堂（北京）科技股份有限公司、北京深度搜索科技有限公司、阿尔特汽车技术股份有限公司、北京骛沃德文化传播有限公司、天津市电视技术研究所有限公司、天津市电子学会、北京大学工学院、北京化工大学、北京体育大学新闻与传播学院、云南师范大学、唐山工业职业技术学院、南京金陵科技学院网络与通信工程学院、重庆仙桃大数据与物联网职业培训学院、北京红猫教育科技有限公司、达内时代科技集团有限公司、北京课工场教育科技有限公司。

本标准主要起草人：李学林、色云峰、程宇、李利利、刘建军、孙万军、孙红菊、贾桂花、柳吉、刘巍、王鹏达、杨艳华、钟方正、梁晓、陈庆良、马洪军、刘倩、兰帅辉、夏升旺、许三利、薛大龙、张玺、吴建国、李岭涛、马冀、孙达明、杨扬、房建伟、周涛、张诚、朴乾、李保奎、崔伟杰、赵冲、贾娜、李杰、冯永亮、李勇、李鸥、戴荣、华玮茜、苏建平、赵志伟、回景旺、刘旭日、陈昊、张乐、孟凡丽、袁毅、韩露、王长乐。

本标准为首次制定。

引言

大数据是新一轮科技革命和产业变革的重要驱动力量，是构建数字社会和数字经济的核心基础，是塑造国家竞争力的重要抓手。“十三五”期间，我国确立大数据发展战略，将大数据作为基础性战略资源，全面实施促进大数据发展行动，加快推动数据资源共享开放和开发应用，助力产业转型升级和社会治理创新。“十四五”规划纲要对数字中国建设做出顶层设计，提出“打造数字经济新优势、加快数字社会建设步伐、提高数字政府建设水平、营造良好数字生态”。

自“十三五”时期起，基于海量数据资源和丰富应用场景的先天发展优势以及国家大数据战略支持，我国大数据生态体系逐步完善，大数据技术与产业得到长足发展，市场规模不断扩大，社会影响力不断提高。要实现大数据产业高质量发展，加快建设实体经济、科技创新、现代金融、人力资源协同发展现代产业体系，大数据产业人才是不可或缺的资源。加快培养发展一大批兼顾大数据技术和行业经验的复合型人才是我国从“数据大国”迈向“数据强国”的关键。但是，当前大数据人才队伍建设滞后产业发展，存在人才结构性短缺、人才质量良莠不齐等问题。因此，建立以产业需求为导向、岗位能力需求为基础的大数据产业人才标准是推进我国大数据产业人才队伍建设的重要举措之一。

本标准立足大数据产业技术体系及大数据企业实际岗位需求，围绕大数据数据预处理、数据标注、数据分析、产品开发、项目实施与运维、平台建设、数据安全、数据管理、运营与应用、咨询服务等10个方向，梳理出31个具体岗位的能力要求。该标准为加快大数据人才需求侧与人才培养供给侧有效对接提供支撑，为企业人才遴选招聘、院校大数据学科建设、社会培训、人才评价、人才服务提供指导参考。

大数据产业人才岗位能力要求

1 范围

本标准规定了大数据主要方向岗位能力要求。
本标准适用于指导各单位开展大数据领域人才培养、人才评价、人才招聘、人才引进等工作。

2 规范性引用文件

下列文件对于本文件的应用是必不可少的。
GB/T 35295-2017 《信息技术 大数据 术语》
GB/T 35589-2017 《信息技术 大数据 技术参考模型》
GB/T 36073-2018 《数据管理能力成熟度评估模型》
T/MIITEC 004-2020 《工业和信息化人才岗位能力评价通则》

3 术语和定义

下列术语和定义适用于本文件。

3.1

大数据 big data

具有体量巨大、来源多样、生成极快、且多变等特征并且难以用传统数据体系结构有效处理的包含大量数据集的数据。

[来源：GB/T 35295-2017 《信息技术 大数据 术语》，2.1.1]

3.2

非结构化数据 unstructured data

不具有预定义模型或未以预定义方式组织的数据。

[来源：GB/T 35295-2017，2.1.25]

3.3

大数据生存周期模型 lifecycle model for big data

用于描述大数据的“数据—信息—知识—价值”生存周期和指导大数据相关活动的模型;这些活动主要由收集、准备、分析和行动等阶段覆盖。

[来源: GB/T 35295-2017 《信息技术 大数据 术语》, 2.1.26]

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。

[来源: GB/T 35295-2017, 2.2.9]

3.5

数据预处理 data preprocessing

包括数据验证、清洗、标准化、格式化和存储。

[来源: GB/T 35589-2017 《信息技术 大数据 技术参考模型》, 7.3.3]

3.6

数据处理 data processing

数据操作的系统执行。

[来源: GB/T 35295-2017 《信息技术 大数据 术语》, 2.2.2]

3.7

数据管理 data management

在数据处理系统中, 提供对数据的访问, 执行或监视数据的存储, 以及控制输入输出操作等功能。

[来源: GB/T 35295-2017 《信息技术 大数据 术语》, 2.2.3]

3.8

大数据服务 big data service

基于大数据参考体系结构提供的数据服务。

[来源: GB/T 35295-2017, 2.1.15]

4 大数据主要方向及岗位

4.1 主要方向

根据大数据技术体系及对人才的需求，本标准围绕数据预处理、数据标注、数据分析、产品开发、项目实施与运维、平台建设、数据安全、数据管理、运营与应用、咨询服务等 10 个方向对人才的岗位能力进行说明。

4.2 主要岗位及职责

本标准主要涉及以下大数据岗位，具体如表1所示。

表 1 大数据各方向主要岗位及职责

序号	方向	岗位名称	岗位职责
01	数据预处理	数据采集工程师	负责大数据采集方案设计与开发，实现基于系统集成、日志、网络爬虫等的数据采集
02		数据清洗工程师	负责发现和处理数据异常，制定确保数据质量的流程和制度
03		数据存储工程师	负责设计和开发大数据存储系统，解决存储性能优化、容量规划
04	数据标注	标注采集工程师	负责合法、合规地采集人工智能算法所需特定结构化和非结构化数据
05		数据标注工程师	负责数据标注方案的执行，保证标注质量、数据质量和数量
06		标注质检工程师	负责数据质检和验收、质检人员培训考核、数据质量反馈与改善
07		标注管理工程师	负责数据采集、标注、质检项目方案制定与执行、质量把控和交付
08	数据分析	数据算法工程师	负责选择、组合或定制化设计适合问题模型的求解算法，并对算法进行调优、性能分析与泛化能力评价
09		数据挖掘工程师	负责利用算法从大量数据中搜索隐藏于其中的信息，提高大数据利用效率
10		数据分析工程师	负责数据统计分析、深度挖掘分析与业务预测，并形成分析报告
11		数据可视化工程师	负责开发数据可视化产品、输出数据可视化图表和报告
12	产品开发	数据产品架构工程师	负责大数据产品架构设计和优化
13		数据开发工程师	负责大数据产品技术方案设计与大数据应用开发
14		数据测试工程师	负责大数据产品质量控制活动
15		大数据产品经理	负责大数据产品需求、规划、设计、交付、产品持续迭代优化
16	项目实施与运维	大数据解决方案工程师	负责大数据项目售前技术方案的编写，协同制定项目整体实施方案
17		数据实施工程师	负责大数据项目现场部署、实施、跟进

18		数据运维工程师	负责大数据项目基础环境、硬件、软件、安全运维及运维管理
19	平台建设	数据平台架构师	负责企业级大数据平台的架构设计与建设方案制定，统筹大数据平台建设实施
20		数据平台开发工程师	负责基于企业级大数据平台进行平台软件定制开发
21		数据平台运维工程师	负责完成大数据平台的部署、监控、分析、调优
22	数据安全	数据安全架构工程师	负责制定大数据安全体系顶层规划与设计，设计组织架构和安全管理體系
23		数据安全评估工程师	负责分析、评估大数据中存在的威胁、漏洞及风险，并提出改进措施
24		数据安全运维工程师	负责大数据安全巡检、安全加固、脆弱性检查、渗透性测试应急保障
25	数据管理	数据治理工程师	负责制定大数据战略、组织结构、规章制度
26		数据管理工程师	负责大数据全生命周期管理
27	运营与应用	数据运营工程师	负责大数据平台、社区、生态、产品、数据、内容等具体的运营推广工作
28		数据应用工程师	负责垂直行业大数据价值挖掘与大数据应用开发和推广
29	咨询服务	大数据咨询师	负责提供大数据平台建设、数据全生命周期管理、业务数字化及数字业务化解决方案的咨询服务
30		大数据培训师	负责大数据培训方案设计、课程开发、人才培养
31		社群管理师	负责社群用户和交互数据的安全可追溯，打造符合定位的社群内容与产品

5 大数据产业人才岗位能力要素

本标准从综合能力、专业知识、技术技能、工程实践能力四个维度提出了大数据产业人才岗位能力要素。

表2 大数据产业人才岗位能力要素列表

维度	要素	说明
综合能力	软能力	指相应岗位人才为完成工作任务所应具备的行为特征和综合素质，包括学习追踪、沟通协调、需求与趋势分析、业务场景把握等技能
专业知识	基础知识	指相应岗位人才应掌握的通用知识，主要包括基本理论、相关标准与规范知识以及有关法律法规、安全、隐私等
	专业知识	指相应岗位人才完成工作任务所必备的知识，主要指与具体岗位要求相适应的理论知识、技术要求和操作规程等
技术技能	基本技能	指相应岗位人才为完成工作任务所应具备的对基础知识应用的水平以及熟练程度
	专业技能	指相应岗位人才为完成工作任务所应具备的对专业知识应用的水平以及对特殊工具使用的掌握
工程实践	经验	指相应岗位人才在实际工程与项目推进中应当具备的经验

6 大数据产业人才岗位能力要求

6.1 大数据预处理岗位能力要求

6.1.1 数据采集工程师

a) 综合能力

- 熟悉大数据采集汇聚技术体系及其发展趋势；
- 熟悉典型应用场景下大数据采集的解决方案和流程；
- 具备良好的沟通表达及团队合作能力；

b) 专业知识

- 熟悉典型文件采集、日志采集、消息采集、DPI采集原理及方案；
- 熟悉网络协议，如TCP/IP、FTP、HTTP等技术原理，了解COAP、MQTT等物联网协议技术原理；
- 熟悉HTML技术、正则表达式及爬虫原理；
- 熟悉多进程/多线程并发处理及容器化技术；
- 熟悉大数据采集相关的标准与规范知识、有关法律法规以及行业准则；

c) 技术技能

- 熟悉标准SQL语言，熟练至少一种常用关系型数据库，如MySQL、PostgreSQL等；
- 熟悉至少一种常用编程语言，如Python、GO、Java等，具备编程基本技能及良好的代码编写习惯；
- 熟悉网页爬取技术，熟练使用Scrapy等常见爬虫框架，并且可做二次开发；
- 熟悉ELKB日志采集分析方案，熟练使用如Logstash、Beats等常见日志采集工具；
- 熟悉如Kafka、Flume、RocketMQ、RabbitMQ等消息队列方案，为不同类型的大数据计算框架提供数据汇聚；
- 熟悉至少一种NoSQL数据库、分布式文件系统、ETL工具，如MongoDB、HDFS、Sqoop等；

d) 工程实践

- 具备日志分析、网页数据爬取、数据库数据采集、传感器或移动终端数据采集等典型业务场景下数据采集方案的设计和开发经验；
- 具备良好的项目方案及技术文档编制能力；
- 具备典型业务场景下数据采集系统的建设、实施、运维经验。

6.1.2 数据清洗工程师

a) 综合能力

- 熟悉数据治理体系、数据质量控制、数据稽核方法，以及行业发展趋势；
- 具备良好的沟通能力和表达能力，善于学习和归纳，能够独立开展业务调研、规范制定、流程设计、报告编写工作；
- 具备良好的协调能力和合作精神，善于配合开展数据采集、加工、存储、分析工作；

b) 专业知识

- 掌握统计学、应用数学相关知识，了解数据分析常用手段及其数学原理；
- 熟悉常见的数据清洗方法和流程；
- 熟悉常见的数据预处理方法，如数据转换、平滑处理、规范化处理和数据脱敏的方法和原则；

- 掌握管理学知识、工程统筹知识，熟悉人机协作计划制定的关键要素；
- 掌握常用数据库的管理和使用方法、常用数据库间数据传输技术，掌握标准SQL语言；
- 了解数据质量相关的标准与规范知识；

c) 技术技能

- 熟悉使用编程语言或清洗工具进行数据清洗，如Java、Python、Kettle等；
- 熟悉Hadoop生态组件，如Hive、HDFS、HBase、Spark、Storm、Flume、Kafka等；
- 熟悉标准SQL语言，能熟练使用和管理常用数据库；
- 熟悉至少一种主流NoSQL数据库，熟悉至少一种主流关系型数据库；
- 熟悉Linux，能熟练使用Shell、Python等脚本语言，并编写自动化脚本；
- 熟练运用统计方法进行数据统计；

d) 工程实践

- 具备数据仓库建模、设计ETL数据抽取流程、设计任务调度、制定维护作业计划、按照数据存储结构进行数据清洗的经验；
- 具备设计稽核点的经验，能合理设计稽核逻辑，完成完整性、一致性、及时性、准确性稽核；
- 具备相当的文档撰写能力，能够完成系统运维文档、维护作业计划、应急处理方案等文档的编写。

6.1.3 数据存储工程师

a) 综合能力

- 熟悉典型应用场景下大数据存储需求及问题解决的方案；
- 具备良好的学习能力，能够跟踪大数据存储技术发展趋势；
- 具备良好的沟通能力和表达能力；

b) 专业知识

- 熟悉大数据存储分类、存储方式、处理特性、数据库数据模型、数据库存储与处理架构等相关知识；
- 熟悉分布式文件系统HDFS基础知识、熟悉NoSQL数据库，如列数据库Hbase、文档型数据库MongoDB、键值数据库Redis等的相关技术原理和运行机制；
- 熟悉关系数据库概念、原理和应用场景；
- 了解云数据库的概念、基本原理和代表性产品的使用方法；
- 了解大数据安全存储方法；

c) 技术技能

- 熟悉常见分布式文件系统，如HDFS、FastDFS等系统的操作和使用；
- 熟悉列数据库HBase、文档型数据库MongoDB、键值数据库Redis等NoSQL数据库的部署及使用；
- 熟悉标准的SQL语言，熟练使用常用关系数据库如Oracle、MySQL等；
- 熟悉常用的数据库中间件如Hibernate、Mango、Cobar等的部署及应用；
- 熟悉一种主流编程语言、脚本语言、大数据相关编程语言，具备较强的编码能力和良好的代码编写习惯；
- 能够制定多副本存储，数据备份和恢复数据备份，故障恢复与迁移故障恢复方案；

d) 工程实践

- 具备相当的大数据存储需求分析与存储架构设计经验；
- 具备相当的大数据项目相关数据库部署经验。

6.2 大数据标注岗位能力要求

6.2.1 标注采集工程师

a) 综合能力

- 了解人工智能产业发展趋势；
- 了解数据标注典型业务场景和业务流程；
- 具备良好的沟通表达及团队合作能力；
- 熟悉人工智能数据相关的需求和定义，掌握解析、采集项目需求文档的能力；

b) 专业知识

- 熟悉数据采集作业的规范和操作流程；
- 了解常见数据类型采集所需的专业设备及型号；
- 掌握数据采集相关知识，了解数据标注工具组件功能；
- 掌握人工智能、深度学习、语音识别、计算机视觉、自然语言理解等基础知识和原理；
- 深入了解数据标注规范及质检标准；

c) 技术技能

- 掌握合法合规获取人工智能数据的渠道和方法；
- 掌握专业采集环境及工具（含硬件）的搭建和调配；
- 掌握复杂场景或专业领域的数据采集任务技能，且数据质量满足规范要求；
- 掌握特定需求下的数据采集场景搭建能力；
- 能够熟练应用数据采集软件，具备对于指定数据采集设备应用的快速学习能力；
- 能够独立完成采集工具部署以及采集脚本开发准备；

d) 工程实践

- 具备完成通用领域，不涉及专业采集设备的简单数据和复杂数据采集任务的能力；
- 具备完成复杂场景或专业领域的数据采集任务的能力，如3D点云采集、3D人脸采集等。

6.2.2 数据标注工程师

a) 综合能力

- 了解人工智能产业发展趋势；
- 了解数据标注典型业务场景和业务流程；
- 具备良好的沟通表达及团队合作能力；
- 熟悉人工智能数据相关的需求和定义，掌握解析、标注项目需求文档的能力；

b) 专业知识

- 掌握人工智能基础知识和原理；
- 熟悉数据标注作业的架构体系和操作流程；
- 熟悉数据标注工具组件功能；
- 了解数据标注的工作内容，使用范围及形式，熟悉各项标注类型；
- 掌握数据标注规范及质检标准；

c) 技术技能

- 具备解析标注项目需求文档的能力；
- 熟练使用人工智能数据标注工具和平台；
- 掌握复杂场景、专业领域的标注能力，且标注数据质量满足规范要求；
- 熟悉Windows、Linux等常见操作系统、常用办公软件；

——具备保密条件下的数据标注能力；

d) 工程实践

——具备完成通用领域简单数据标注任务的能力，如图像标框、视频切割标注、语音切分标注、中文朗读普通话转写标注和中文实体名标注等；

——具备完成通用领域复杂数据标注任务的能力，如人脸106点关键点标注、视频人体追踪标注、方言识别标注、文本相关性判断标注等；

——具备完成复杂场景或专业领域的标注任务的能力，如图文混合标注、3D点云标注、航空语音标注等。

6.2.3 标注质检工程师

a) 综合能力

——了解采集和标注数据的典型应用场景；

——掌握解析采集和标注项目需求文档的能力；

——具备总结提炼能力，能够将工作实践上升为理论与经验；

b) 专业知识

——掌握人工智能、深度学习、语音识别、计算机视觉、自然语言理解等基础知识和原理；

——熟悉质检审核作业操作规范和流程；

——掌握样本抽样检测知识，能够合理设计有效质检方案；

c) 技术技能

——熟练应用数据标注工具进行在线数据审核与修改；

——掌握各类数据格式，能够实现对数据交付技术参数的检查；

——掌握复杂场景和专业领域内的数据质量要求、评判标准和质检/验收流程，具有拟定验收标准的能力；

——能够完成项目质检，对已有的质量问题提供改善方案，并对验收结果负责；

d) 工程实践

——具备全类型数据标注结构审核与修改经验；

——具备根据项目需求，完成相应的质检任务并按照流程反馈质检结果的能力；

——具备完成复杂场景或专业领域的标注项目质检能力。

6.2.4 标注管理工程师

a) 综合能力

——了解人工智能产业发展趋势；

——了解数据标注典型业务场景和业务流程；

——具备大项目、项目集管理与沟通能力；

——能够建立清晰的项目数据集管理机制；

b) 专业知识

——掌握项目管理相关知识；

——了解数据定制服务的报价规则；

——了解工期管控和精准预测的方法；

——掌握人工智能数据生产工艺流程；

c) 技术技能

——熟悉项目环境，能够针对不同标注项目制定特定标注流程、配置对应工具模板；

- 熟悉项目范围管理，能够使用标注工具和平台执行在线任务分配；
- 具备项目风险管理能力，能够制定项目计划，管控项目资源，控制项目整体风险；
- 熟悉项目约束，能够把握项目进度、质量与成本间的平衡关系，合理控制项目变更；
- 熟悉项目文档管理，能够完成项目交付文档，汇报项目整体进展，总结项目实践经验；

d) 工程实践

- 具备丰富的数据采集、标注、质检项目经验；
- 熟悉项目操作规范和流程，具备组建与带领团队的组织协调经验，完成数据采集、标注任务，且数据质量满足规范要求。

6.3 大数据分析岗位要求

6.3.1 数据算法工程师

a) 综合能力

- 熟悉大数据算法演进及其演化趋势，具备典型业务场景下的算法应用能力；
- 具备优秀的抽象建模能力、逻辑思维能力，独立将现实问题转化为数学模型；
- 具备良好的沟通技巧和推动能力，能够正确理解业务需求，并能够选择合适的算法方案；

b) 专业知识

- 熟悉数据结构与算法，熟悉概率论、数理统计、线性代数、随机过程、最优化理论等相关基础知识；
- 熟悉折半、倍增、搜索、分治、贪心、暴力、构造等基础算法；
- 熟悉常见的机器学习和深度学习模型、原理和应用场景；
- 掌握大数据生命周期管理方法；
- 熟悉大数据平台基础知识及相应的大数据算法运行环境及其并行算法框架；

c) 技术技能

- 了解一种或多种主流深度学习框架，如TensorFlow、PyTorch、MXNet、Caffe、Theano等，具备根据实际应用场景设计算法和算法调优的能力；
- 熟悉MATLAB等仿真工具，熟悉数据分析工具SAS、SPSS等的使用；
- 熟悉Hadoop生态组件，如Hive、HDFS、HBase、Spark、Storm、Flume、Kafka等；
- 熟悉Oracle、MySQL、SQL Server等关系型数据库及相应的数据库中间件，熟悉标准SQL语言使用；
- 熟悉使用Java、C++、Scala、Python、R等编程语言，具备扎实可靠的编程能力及良好的代码编写习惯；
- 熟悉Hadoop、Spark、Storm、Flink等常见的大数据计算框架；

d) 工程实践

- 具备相当的大数据算法分析研发经验；
- 具备相当的大数据算法优化经验。

6.3.2 数据挖掘工程师

a) 综合能力

- 熟悉大数据下数据挖掘各方法体系及其应用领域的现状和发展趋势；
- 掌握典型的数据分析与数据挖掘的任务场景和问题解决过程；
- 掌握数据分析思维方法，善于沟通并充分理解需求；

——掌握良好的业务建模能力，能够协调数据团队、业务团队协同工作；

b) 专业知识

——熟悉数据结构与算法，熟悉概率论、数理统计、线性代数、随机过程、最优化理论等相关基础知识；

——掌握常见数据挖掘算法的原理和应用，如分类算法、聚类算法、回归算法、关联分析算法、特征工程算法；

——熟悉数据采样、数据透视、数据集成、缺失值处理、异常值检测、异常值处理、标准化等数据预处理方法；

——了解多层感知机、卷积神经网络、循环神经网络、生成式对抗网络等深度神经网络的原理和应用；

c) 技术技能

——熟练使用Pandas、Numpy、Matplotlib、Scikit-learn等数据处理与挖掘算法包；

——掌握数据库的原理和标准SQL语言的使用；

——掌握交互式数据挖掘分析工具Orange Datamining的使用；

——掌握大数据平台下的Spark SQL查询语句和Spark MLlib机器学习框架；

——了解至少一种深度学习算法框架，如Tensorflow、PyTorch、飞桨等；

d) 工程实践

——具备相当的工程应用能力，能够权衡并解决工程应用中的性能、安全、隐私、成本等问题；

——具备相当的文档撰写能力，能够撰写分析报告以及应用开发方案等。

6.3.3 数据分析工程师

a) 综合能力

——熟悉大数据分析技术体系及其发展趋势；

——熟悉典型应用场景下大数据分析的需求场景和解决方案；

——具备良好的沟通能力和表达能力，善于进行数据分析成果的推广和辅助决策；

b) 专业知识

——熟悉统计学、应用数学相关知识，了解数据分析常用手段及其数学原理；

——掌握数据结构、数据库、数据建模、数据挖掘、数据分析等知识；

——掌握数据分析常见分类、基本原理与常见数据分析方法；

——掌握基本数据分析方法和高级数据分析方法，如数据统计分析与特征描述、数据相关性分析、分类分析、回归分析、聚类分析、异常检测、离群点分析等；

——了解大数据分析相关的标准、规范知识以及有关法律法规；

c) 技术技能

——掌握常用于统计分析、交互式分析、批处理分析和实时流分析等的数据分析工具，如SAS、SPSS、Hive SQL/Spark SQL等；

——熟悉至少一种主流编程语言、脚本语言和大数据相关编程语言，如Java、Python、Scala等，具备编程基本技能及良好的代码编写习惯；

——熟悉主流NoSQL数据库，如Hbase、Kudu、Redis、MongoDB等，熟悉至少一种主流关系型数据库，如Oracle、SQL Server、MySQL等；

——熟悉SQL、HQL等数据库编程语言及相关管理工具，熟悉脚本语言开发和优化提升；

——熟悉Hadoop、Spark、Storm、Flink等常见的大数据计算框架；

d) 工程实践

- 具备较强的典型业务场景下的数据敏感性和分析能力，有能力选择适合的数据分析方法；
- 具备较强的大数据项目分析经验；
- 具备一定的大数据分析环境搭建经验。

6.3.4 数据可视化工程师

a) 综合能力

- 熟悉大数据下数据可视化相关技术和方法的应用领域和发展趋势；
- 掌握常见业务场景下的数据可视化方案；
- 掌握数据分析思维方法和一定的美学设计原则；
- 掌握利用数据可视化技术解决具体问题的方法；

b) 专业知识

- 掌握数据可视化的含义、作用和基本流程，熟悉数据可视化基本工具和数据可视化科学研究工具的类型和应用领域；
- 掌握OLAP方法的定义及应用，掌握数据多维分析方法的原理；
- 掌握常用的数据统计分析方法，掌握常用统计指标的含义及用法；
- 掌握商业场景数据分析与可视化理论和方法；
- 掌握Java服务端开发框架，包括Spring Boot、JPA、Spring Security等；
- 掌握JavaScript面向对象的思想以及MVVM的前端开发理论；

c) 技术技能

- 精通数据库的原理和标准SQL语言的使用；
- 精通一款商业级可视化分析工具或图表制作工具；
- 精通Hive、HBase、Spark SQL、MongoDB等NoSQL数据库应用与查询优化；
- 精通至少一门应用数据开发语言，如Java语言、Python语言等；
- 理解并掌握JavaScript语言核心技术，如DOM、BOM、Ajax、JSON等；
- 熟练常见前端框架及工具库，如Webpack、Vue.js、jQuery、RequireJs；
- 熟悉常见网页前端数据可视化技术，如Echarts、Highcharts、D3.js等；

d) 工程实践

- 具备相当的业务需求分析能力，能够发现数据可视化分析的目标问题；
- 具备相当的数据可视化技术应用能力，能够完成数据可视化分析过程；
- 具备相当的文档撰写能力，能够撰写数据分析报告，并提出业务问题的解决方案；
- 具备一定的产品设计能力，研究并持续改善产品的质量、性能和用户体验。

6.4 大数据产品开发岗位能力要求

6.4.1 数据产品架构工程师

a) 综合能力

- 具备较强的学习能力，能够跟进大数据产品架构技术发展趋势；
- 熟悉典型的大数据业务场景，掌握大数据产品架构方法；
- 具有责任感和团队意识，能够推动大数据产品开发落地；

b) 专业知识

- 熟悉分布式系统的基础知识和架构体系；
- 掌握大数据分布式架构理论、设计原则、常见架构等相关知识；

- 掌握大数据相关技术，熟悉大数据典型业务流程、应用方式，熟知大数据架构设计中部署架构、安全架构相关规范及设计实现的要求；
- 熟悉我国现行的大数据标准体系，熟悉大数据架构相关技术分类、技术标准，能够按照相关标准规范设计和实施大数据架构；
- 熟悉大数据主流技术，如大数据开发技术、分布式技术、数据仓库技术、缓存技术、大数据安全相关技术等，熟知相关技术的原理、特点、应用方式；

c) 技术技能

- 熟悉分布式技术如分布式消息、分布式计算、分布式存储、分布式监控、分布式版本控制、微服务及容器技术的使用场景和常用工具；
- 熟悉大数据产品微服务架构环境，如OpenStack、Docker、Kubernetes、Istio等；
- 掌握基于ToGAF架构方法展开业务架构、应用架构、技术架构、数据架构、部署架构、安全架构，且能够根据具体业务场景及需求设计并实施架构；
- 熟悉大数据架构所需的实时性、可靠性、一致性等核心指标及实现方案；
- 熟悉至少一种主流编程语言、脚本语言、大数据相关编程语言，如Java、Python、Scala等，具备较强的编码能力和良好的代码编写习惯；

d) 工程实践

- 具备较强的高并发、分布式环境下的架构需求分析与设计经验；
- 具备相当的大数据产品项目架构开发经验。

6.4.2 数据开发工程师

a) 综合能力

- 具备较强的学习能力，能够跟进大数据开发技术发展趋势；
- 能够熟悉典型的大数据业务场景，进而把握大数据产品开发的方向和目标；
- 具备良好的沟通协调能力，团队意识强；

b) 专业知识

- 熟悉大数据生命周期管理方法，包括大数据采集、预处理、存储、分析和可视化；
- 掌握Hadoop生态圈各功能组件基本概念、主要特点、运行机制；
- 掌握SQL、NoSQL数据库原理及应用场景；
- 熟悉分布式存储和大规模数据批量计算的实现原理；
- 熟悉分布式计算框架的核心思想和适用场景；
- 熟悉分布式资源管理框架在通用资源管理和调度方面的工作机制；
- 掌握大数据产品开发过程和方法；

c) 技术技能

- 熟悉常见分布式计算框架，如MapReduce、Flink、Storm、Spark/Spark Streaming等；
- 熟悉分布式数据存储相关技术，如HDFS、HBase、Hive等；
- 熟悉大数据采集环境，包括Flume、Kafka、Sqoop、RocketMQ、RabbitMQ等；
- 熟悉分布式资源管理与集群协调工具，如YARN、ZooKeeper等；
- 熟悉主流NoSQL数据库，如HBase、Kudu、Redis、MongoDB等，熟悉至少一种主流关系型数据库，如Oracle、SQL Server、MySQL等；
- 熟悉SQL、HQL等数据库编程语言及相关管理工具，熟悉脚本开发和优化提升；
- 熟悉至少一种主流编程语言、脚本语言、大数据相关编程语言，如Java、Python、Scala等，具备较强的编码能力和良好的代码编写习惯；
- 具备一定的大数据应用工程项目规划、设计、管理能力；

d) 工程实践

- 具备相当的大数据开发环境搭建经验和大数据项目开发经验；
- 具备一定的大数据项目技术管理经验。

6.4.3 数据测试工程师

a) 综合能力

- 具备较强的学习能力，能够跟进大数据测试技术发展趋势；
- 熟悉典型的大数据业务场景；
- 具有良好的逻辑思维能力、业务分析能力、沟通和表达能力；

b) 专业知识

- 熟悉大数据测试方案、测试计划、测试用例、测试报告等的编写方法，以及测试用例执行和测试结果分析的方法；
- 熟悉大数据系统测试常见的测试模型、测试类型、测试方法；
- 熟悉大数据生命周期管理方法，包括数据采集方法、数据清洗方法、数据存储架构、数据分析算法、数据可视化方法等；
- 熟悉大数据平台架构和运行机制；
- 熟悉数据质量保证和质量控制的基本方法、基本流程和工作产品；
- 了解设计和开发测试工具和自动化测试框架的方法；

c) 技术技能

- 能够搭建大数据测试环境，包括平台环境、架构环境、采集环境、存储环境、测试平台等；
- 能够根据测试方案和测试计划，编写和执行测试用例，输出测试报告，并不断优化测试过程、提高测试效率；
- 能够选择合适的测试方法和测试工具进行不同类型的测试，熟悉常用的Bug管理跟踪工具，熟悉主流性能测试工具，如JMeter等；
- 熟悉大数据自动化测试框架，具备自动化测试脚本编写能力；

d) 工程实践

- 具备相当的大数据测试环境搭建经验；
- 具备相当的大数据用例编写和执行经验；
- 具备相当的测试结果分析和测试报告编写经验。

6.4.4 大数据产品经理

a) 综合能力

- 熟悉大数据产品技术体系及大数据产品发展趋势，熟悉典型的大数据产品业务场景；
- 具备良好的沟通和协调能力，能够把握产品关键点和用户痛点，并能简明扼要的表达产品的价值点；
- 了解行业趋势，善于捕捉市场、技术、竞争者等各种变化对产品发展的影响，收集用户体验与市场需求，发掘确认创新或改进产品的机会；

b) 专业知识

- 掌握大数据及其产品的生命周期管理方法；
- 熟悉大数据产品技术环境，如大数据平台构成及运行机制、大数据存储环境和大数据计算环境；

- 熟悉大数据产品敏捷开发方法，如产品规划方法、产品设计方法、产品实现方法和产品测试方法，熟悉BRD/MRD/PRD/FSD等文档撰写；
- 了解大数据相关的标准与规范知识以及有关法律法规；

c) 技术技能

- 掌握需求调研、需求分析、产品策划、原型设计、Demo制作等所需的工具，如XMind、Axure、Visio等；
- 熟悉需求调研、需求分析，能够完成产品策划、原型设计、Demo制作、产品流程和交互设计，完成BRD/MRD/PRD/FSD等文档的撰写和管理；
- 熟悉常用大数据分析平台及相应数据分析工具，如Hadoop、Spark、Storm等；
- 熟悉标准SQL语言，了解常用编程语言Python、Scala、R等的一种；
- 了解常用的SQL和NoSQL数据库，如Oracle、HBase、MongoDB等；

d) 工程实践

- 具备相当的产品设计经验，能够开展业务调研、需求分析、产品设计等工作；
- 具备相当的大数据产品策划、设计、开发和应用管理经验。

6.5 大数据项目实施与运维岗位能力要求

6.5.1 大数据解决方案工程师

a) 综合能力

- 具备较强的分析和撰写方案的能力；
- 具备较强的沟通技巧和文案呈现能力；
- 熟悉大数据行业的应用场景，发掘整合行业生态应用；
- 熟悉大数据行业相关标准及生态系统的发展趋势；

b) 专业知识

- 熟悉国内外大数据安全政策法规及标准；
- 熟悉大数据平台的硬件选型标准及常用的组网架构；
- 熟悉数据库范式、数据仓库、商务智能等理论知识；
- 熟悉大数据常用的架构，了解主流的批流一体、湖仓一体方案；
- 熟悉产品原型设计的理论及人机交互设计的知识；
- 了解业务关系管理、业务营销、财务管理等IT服务营销相关知识；

c) 技术技能

- 掌握IT基础架构设计技能，包括当前主流的服务器、存储、网络、安全等产品；
- 熟悉常用关系型数据库与大数据平台的整合方案等；
- 熟悉批处理及流处理的大数据集群组建方案，如Hadoop、Spark、Storm、Flink等；
- 熟悉云原生环境下大数据与容器、K8S编排、存算分离的整合方案；
- 熟悉常用的数据治理工具及数据安全工具，如Atlas、Ranger、Sentry等；

d) 工程实践

- 具备政务、交通、旅游、教育等大数据行业项目售前支持、策划、调研与方案编写等的实践经验；
- 具备调研、梳理及分析客户痛点的能力，完成有针对性的行业解决方案规划及设计；
- 具备较强的办公软件操作能力，能有效撰写技术方案书，绘制网络拓扑、数据流图等。

6.5.2 数据实施工程师

a) 综合能力

- 了解大数据技术体系及其发展趋势，熟悉常见的大数据实施场景；
- 具备良好的沟通能力、表达能力和项目管理能力，能够有效推动项目落地；
- 具备较强的责任心和质量意识，实施过程能够满足系统设计要求；

b) 专业知识

- 熟悉大数据平台相关的基础知识和技术原理；
- 熟悉大数据项目实施的主要阶段、主要内容和关键活动，如个性化定制、系统部署、系统培训、系统验收和系统运维；
- 了解大数据项目生命周期的各个阶段；
- 了解大数据实施相关的标准、规范知识以及有关法律法规和项目管理知识领域；

c) 技术技能

- 熟悉主流版本的Linux和Windows Server等服务器版操作系统的部署和管理；
- 熟悉常见的大数据平台及生态组件，如Hadoop、HDFS、HBase、Spark、Storm、Flume、Kafka、Hive等；
- 熟悉主流的资源管理与调度系统、大数据采集系统、大数据存储平台、大数据计算环境的安装和部署；
- 熟悉常用思维导图、产品原型工具，如XMind、Axure、Visio等；
- 熟悉项目实施过程，能够根据项目需要开展定制开发、系统部署、系统培训、系统验收和系统运维等活动；

d) 工程实践

- 具备相当的大数据项目部署和实施经验，推动系统上线与验收工作；
- 具备相当的文档撰写能力，能够完成项目需求说明书和验收报告等文档的编写。

6.5.3 数据运维工程师

a) 综合能力

- 具备较强的学习能力，能够跟踪大数据运维技术发展趋势；
- 具备服务意识，能够与用户进行有效沟通，提升用户满意度；
- 具备较强的预防意识和责任心，能够发现大数据系统运行过程中存在的问题和隐患，并及时解决；

b) 专业知识

- 熟悉计算机硬件组成原理、操作系统原理、计算机网络、服务器硬件等基础知识；
- 熟悉大数据平台构架，具备大数据平台相关组件安装部署及日常运维的基础知识；
- 熟悉信息系统运维方法论和大数据系统运维相关知识，包括大数据系统安装部署、安全配置、事件处理、问题诊断、日志分析、性能调优等；
- 掌握大数据运维平台构成、运行机制和使用方法，掌握自动化运维方法和自动化运维脚本编写方法；
- 熟悉主流IT运维方法论和规范标准，如DevOps、ITIL等；

c) 技术技能

- 熟悉大数据系统运维技能，包括操作系统安装配置、大数据系统安装部署、安全配置、问题诊断与处理、性能调优等；
- 熟悉常见自动化运维语言或工具，如Python、Shell等；
- 熟练常见自动化运维配置管理工具、自动化运维监控管理工具，如Puppet、Zabbix等；

- 熟悉Hadoop主流生态组件的使用，熟悉主流分布式开发组件的安装部署及调优；
- 了解设计和开发自动化运维与监控的工具；

d) 工程实践

- 具备自动化运维工具的使用经验；
- 具备相当的系统问题发现与处置经验；
- 具备一定的运维管理经验。

6.6 大数据平台建设岗位能力要求

6.6.1 数据平台架构师

a) 综合能力

- 熟悉主流大数据平台技术体系以及发展趋势；
- 熟悉根据各类场景设计大数据平台建设的解决方案；
- 具备良好的沟通能力、组织协调能力；
- 具备较强的建设方案文档与项目文档撰写能力；

b) 专业知识

- 熟悉大数据平台生态体系各类组件的原理、开发方法、部署方法以及调优方法；
- 掌握大数据平台整体设计方法，包含数据采集与处理、离线计算、实时计算、作业调度、数据能力开放、平台安全管理等；
- 熟悉元数据管理、数据标准管理、数据质量管理、数据安全治理、数据生命周期管理等数据治理理论知识；
- 熟悉分布式系统模型、远程调用RPC、分布式协调与同步机制等分布式系统理论知识；
- 了解机器学习、深度学习、人工智能等数据挖掘相关理论知识；
- 了解Devops开发运维一体化、AIOps智能运维等高级运维体系建设理论知识；
- 了解信息系统安全管理、计算机网络安全等安全理论知识以及国内外数据安全政策法规与数据安全管理标准；

c) 技术技能

- 熟悉Hadoop、Spark、Storm、Flink等常见的大数据计算框架的原理、开发方法、部署方法以及调优方法；
- 熟悉Prometheus、Grafana、Zabbix、Ganglia、Ansible等主流大数据平台运维开发工具；
- 熟悉Java、Scala、Python等常用编程语言；
- 熟悉标准SQL语言以及HiveQL、Spark SQL、Flink SQL等语言，掌握数据库规划设计与管理方法；
- 了解云计算开发管理工具OpenStack与容器化管理工具Kubernetes；
- 了解一种或多种主流深度学习框架，如TensorFlow、PyTorch、MXNet、Caffe、Theano等；

d) 工程实践

- 具备丰富的企业级大数据平台顶层架构设计与工程建设实践经验；
- 具备数据分类管理、数据质量管理、数据生命周期管理等数据治理经验。

6.6.2 数据平台开发工程师

a) 综合能力

- 熟悉主流大数据平台组件的行业应用范围与发展趋势；
- 具备良好的沟通表达能力和较强的团队协作意识；

b) 专业知识

- 掌握程序设计基本知识，包括面向服务架构设计、高并发程序设计、面向对象程序设计、数据结构与算法设计模式等；
- 熟悉计算机科学与技术基础理论知识，包括计算机系统人机交互理论、操作系统原理、计算机网络、数据库管理等；
- 了解数据仓库理论，熟悉各类多维数据建模方法，了解数据治理知识体系内容；

c) 技术技能

- 精通至少一种常用编程语言，如Java、Scala、Python等
- 掌握SpringBoot、SpringCloud、Netty、Django等主流企业级应用开发框架；
- 熟悉Hadoop、Spark、Storm、Flink等常见的大数据计算框架；
- 熟悉JetBrain IDEA、JetBrain Pycharm、VSCode、Git、Jemter等程序开发与测试工具的使用方法；

d) 工程实践

- 具备丰富的企业级大数据平台应用开发实践经验；
- 具备高性能、高并发、高可靠应用软件设计与优化经验。

6.6.3 数据平台运维工程师

a) 综合能力

- 熟悉大数据平台的整体架构演进趋势，以及典型的大数据应用架构；
- 掌握常见的大数据平台应用场景，能够针对不同场景给出大数据平台选型建议；
- 掌握IT系统运维的核心理论，善于从逻辑出发解决系统运维问题；
- 掌握良好的沟通能力，善于与产品和研发进行协作，实现研发运维一体化；

b) 专业知识

- 掌握进程管理、存储管理、系统安全、虚拟化等Linux操作系统的原理；
- 精通路由交换、通信安全、网络服务等计算机网络知识；
- 掌握磁盘阵列、网络存储、分布式文件系统、分布式数据库等理论知识；
- 掌握程序语言编译、应用部署、服务监控、日志排查等方法；
- 熟悉评估和处理常见安全风险问题和安全事件的方法；
- 了解ITSS服务流程，并了解规划设计，部署实施、服务运营、持续改进、监督控制内容；

c) 技术技能

- 熟悉Linux日常运维和管理，具备基本Shell、Python脚本阅读能力，深入理解持续集成和持续交付；
- 掌握VPN的构建管理、防火墙的配置、路由表的配置等网络操作技能；
- 熟悉Java、Python、Node.js等编程语言的编译、构建、运行等运维技能；
- 熟悉Docker、Kubernetes的原理、使用、日常运维等；
- 掌握HDFS、Hive、HBase、MongoDB等分布式数据系统的安装与调试；
- 掌握Spark、Kafka、Flume、Redis等分布式处理系统的安装与调试；
- 熟悉常见开源软件、组件，如HAProxy、Nginx、Tomcat、Git、FTP等；

d) 工程实践

- 具备相当的IT资源规划能力，可以根据部署需求给出合理的资源规划；
- 具备相当的业务量分析能力，可以根据业务量确定合理平台架构和硬件需求；
- 具备相当的运维实践能力，可以独立完成大数据平台系统的搭建与运维；

——具备实际工作能力和实践经验，能够进行大数据平台部署、联调和管理。

6.7 大数据安全岗位能力要求

6.7.1 数据安全架构工程师

a) 综合能力

- 关注大数据相关安全技术，能够跟踪大数据安全架构技术发展趋势；
- 具备良好的沟通能力和表达能力，能够与大数据产品工程师、数据开发工程师进行有效沟通；
- 具有良好的逻辑思维，能够把握典型场景下存在的安全问题，并能够针对安全问题提出可行的解决方案；

b) 专业知识

- 掌握网络安全、数据安全相关的工作原理和技术方法；
- 掌握传统的大数据安全架构和零信任大数据安全架构方法，以及相应的安全防护模型和安全风险评估方法；
- 掌握大数据平台及其生态组件的安全技术；
- 熟悉常用的大数据安全防护技术、安全标准、安全规范以及相关的法律法规；

c) 技术技能

- 掌握常用的传统大数据安全技术防护方案和信任安全架构下的大数据安全技术防护方案；
- 掌握大数据产品及其生态组件安全技术、风险分析技术与安全评估技术；
- 熟悉安全运维支撑平台，如OSSIM等；

d) 工程实践

- 具备大数据安全架构设计能力，并能够编制相应的安全架构文档；
- 具备大数据安全架构问题分析能力，并能够针对问题提出解决方案。

6.7.2 数据安全评估工程师

a) 综合能力

- 关注大数据相关安全技术及其发展趋势，跟进最新的安全政策与法律法规研究；
- 具备良好的沟通能力和表达能力，能够与大数据产品工程师、数据开发工程师进行有效沟通；
- 具有良好的逻辑思维和快速解决问题的能力，能够在复杂的大数据场景中迅速发现、定位安全问题，并提出可行的解决方案；

b) 专业知识

- 掌握信息安全技术基础，了解主流的网络安全和数据安全设备；
- 熟悉网络安全法、等保、数据安全管理办法等法律法规及相关行业监管要求；
- 熟悉常见的安全攻防技术，熟悉常见漏洞原理、修复及降低风险的方法；
- 掌握渗透测试技术和大数据安全防护技术；
- 掌握安全评估模型和流程；

c) 技术技能

- 熟悉主流大数据安全技术，如行为监控、数据脱敏、权限管控、加解密算法等；
- 熟悉国内外主流安全检测系统与工具，如Burpsuit、AWVS、Nmap、Nessus、OpenVAS、Metasploit、Kali等；

- 了解大数据产品及其生态组件安全技术、风险分析技术与安全评估技术，包括但不限于Hive、HBase、Kafka、Flink、ES、Spark等；
- 掌握大数据风险分析与评估防范过程；

d) 工程实践

- 具备大数据安全评估能力，并能够编制相应的评估文档；
- 具备大数据安全问题分析能力，并能够提出整改方案。

6.7.3 数据安全运维工程师

a) 综合能力

- 具备良好的学习能力，能够跟踪大数据安全技术和安全运维技术发展趋势；
- 具备较强的安全意识和安全敏感性，能够发现大数据系统各环节存在的安全隐患和问题，并及时解决；
- 具备良好的沟通能力；

b) 专业知识

- 熟悉大数据安全原理以及大数据常见威胁、漏洞和风险；
- 掌握常用的大数据安全攻防方法，包括安全域划分、访问控制列表、渗透测试等；
- 熟悉大数据安全构架方法，包括传统的安全架构方法和零信任安全架构方法，以及相应的大数据安全防护体系；
- 掌握大数据安全运维规程，了解常见的安全设备及系统；
- 了解大数据安全运维相关的标准规范、法律法规；

c) 技术技能

- 熟悉常见的信创安全产品及相关主流技术；
- 熟悉服务器版操作系统的安全运维技术，如Linux、Windows Server等；
- 熟悉大数据平台及其生态组件的安全运维技术，如Hadoop、HDFS、HBase、Spark、Storm、Flume、Kafka、Hive等；
- 熟悉安全运维平台的部署与使用，如OSSIM；
- 熟悉常见的渗透测试工具，如OpenVAS、Metasploit等；

d) 工程实践

- 具备相当的安全运维经验，特别是常见运维平台和运维工具的使用；
- 具备一定的安全文档编写能力，能够完成安全日志整理及安全事件报告书的编写；
- 具备一定的开源文档阅读水平，能够按照文档指引处理故障。

6.8 大数据管理岗位能力要求

6.8.1 数据管理工程师

a) 综合能力

- 熟悉大数据管理技术体系及其发展趋势；
- 熟悉典型应用场景下大数据分析的需求场景和解决方案；
- 具备良好的沟通能力和表达能力，善于进行数据分析成果的推广和用于辅助决策；
- 能够独立开展业务调研、数据管理和报告编写工作；

b) 专业知识

- 熟悉常用SQL语言，熟练操作Mysql、Oracle等大型数据库；
- 熟悉数据仓库架构及原理，能独立进行数据仓库数据建模；
- 熟悉相关数据分析组件ODPS、ADS等；

- 熟悉存储过程、触发器等数据库开发；
- 了解大数据管理相关的标准与规范知识及相关法律法规；

c) 技术技能

- 掌握数据全生命周期管理要求；
- 掌握常用数据统计、审计等管理工具；
- 掌握监控数据库的性能能力；
- 具备处理数据库的备份恢复、应急恢复能力；

d) 工程实践

- 具备分析数据和批量处理数据的能力；
- 具备较强的典型业务场景下的数据敏感性和管理能力，能够使用适合的数据管理方法，运用数据辅助决策；
- 具备相当的文档撰写能力，能够完成数据需求说明书和数据管理报告等文档的编写。

6.8.2 数据治理工程师

a) 综合能力

- 熟悉大数据治理技术体系及其发展趋势；
- 熟悉典型应用场景下大数据治理的需求场景和问题解决方案；
- 具备良好的沟通能力和表达能力，善于进行数据治理成果的推广并将成果用于辅助决策；

b) 专业知识

- 掌握基本数据和高级数据的治理方法；
- 熟悉统计学、应用数学相关知识，了解数据分析常用手段及其数学原理；
- 熟悉数据治理目标、内容、架构等相关知识，熟悉数据战略的规划与实施方法；
- 熟悉数据治理体系，包括组构、制度、流程、绩效和标准等；
- 了解数据中台技术和微服务架构环境；
- 了解大数据治理相关的国内外治理模型、标准与规范知识；

c) 技术技能

- 熟悉常用的数据治理工具及其他相关工具，如Atlas、Ranger、Sentry等；
- 熟悉数据治理的常用工具和最佳实践，如数据指标、元数据、数据标准、数据质量、数据安全、数据资产目录等；
- 掌握数据的优化和管理工具，具备开展数据分析、数据评估、数据清洗、数据监控、错误预警等的能力；
- 具备针对数据组织流程进行优化和管理的能力，如完成制定目标、评估流程、优化计划、监督审核机制、评估改善效果等多个环节；

d) 工程实践

- 具备较强的典型业务场景下的数据敏感性和治理能力，能够使用适合的数据治理方法，运用数据辅助决策；
- 具备大数据治理及管理(元数据、数据标准、数据质量、数据安全、数据资产目录等)方案制定、实施及评估能力；
- 具备相当的文档撰写能力，能够完成数据需求说明书和数据治理报告等文档的编写。

6.9 大数据运营与应用岗位能力要求

6.9.1 数据运营工程师

a) 综合能力

- 了解大数据发展历史和现状，关注企业数字化转型发展趋势；
- 具备商业敏感性和洞察力、数据分析能力，对行业及市场敏感、有前瞻性；
- 具备良好的沟通技巧、执行力，善于整合内外资源和协同各类渠道，达成业务目标；

b) 专业知识

- 熟悉大数据、数据统计、数据挖掘、分类分析等基础知识；
- 掌握大数据生命周期管理方法及大数据架构方法；
- 掌握大数据市场营销、电子商务等基础知识；
- 了解大数据运营建设、管理体系构建方法、宣传推广方法和途径；
- 了解熟悉大数据隐私保护和安全防护技术；
- 熟悉大数据运营相关法律法规和标准规范；

c) 技术技能

- 能够根据业务需求，基于数据统计、数据挖掘、分类分析等工具进行数据分析并指导运营决策；
- 能够基于大数据从生态、品牌、渠道、社区、商家、产品、内容、市场等各方面展开运营优化；
- 掌握大数据运营方法和运营流程，协助优化运营流程，提高流程的效率，改进使用体验；

d) 工程实践

- 具备根据前台部门的业务需求提供数据服务支撑的能力和经历；
- 具备基于大数据产品提供运营支撑及运营分析并达成业务目标的能力和经历。

6.9.2 数据应用工程师

a) 综合能力

- 了解大数据发展历史和现状，关注企业数字化转型发展趋势；
- 具备良好的问题分析和处理能力；
- 熟悉大数据典型应用场景，基于对垂直行业业务和大数据技术的深刻理解，制定垂直行业大数据项目的解决方案，推广大数据应用；

b) 专业知识

- 熟悉大数据生命周期管理方法和大数据应用方法；
- 熟悉相关的业务知识及数据业务化方法，熟悉大数据可视化技术；
- 熟悉大数据应用环境，如大数据平台构成及运行机制、大数据存储环境和大数据计算环境等；
- 了解大数据技术平台及相关生态组件等的技术原理；

c) 技术技能

- 掌握大数据应用技术，包括业务架构技术、应用架构设计、数据架构设计；
- 熟悉大数据可视化框架和技术，包括Kibana、D3.js、ECharts等；
- 熟悉Hadoop生态组件，如Hadoop、Spark、Storm、Flume、Kafka、Hive等；
- 熟悉至少一种主流编程语言、脚本语言、大数据编程语言，如Python、Java、Scala等；
- 熟悉至少一种NoSQL数据库，熟悉至少一种主流关系型数据库；

d) 工程实践

- 具备电商、金融、健康医疗等领域大数据应用经验；
- 具备大数据应用展示推广经验。

6.10 大数据咨询服务岗位能力要求

6.10.1 大数据咨询师

a) 综合能力

- 熟悉行业发展趋势，善于捕捉市场、技术、竞争者等各种变化对行业发展的影响，发掘创新或改进产品的机会；
- 具备良好的逻辑思维能力、学习能力，能够通过数据发现问题、分析问题，并提出解决方案；
- 具备良好的沟通能力和表达能力；

b) 专业知识

- 熟悉大数据治理和大数据应用等知识，熟悉主流大数据技术体系；
- 掌握大数据生命周期管理方法；
- 熟悉大数据平台概念、构成及运行机制；
- 了解大数据相关的标准与规范知识以及有关法律法规；

c) 技术技能

- 具备数据治理、数据分析、数字化转型解决方案规划与设计的能力；
- 熟悉云计算、大数据相关产品和技术；
- 熟悉常用批处理分析平台和实时流分析平台等数据分析工具，如Hive/Spark SQL、Hadoop、Spark、Storm等；
- 熟悉至少一种主流编程语言、脚本语言、大数据相关编程语言；
- 熟悉至少一种主流NoSQL数据库，熟悉至少一种主流关系型数据库；

d) 工程实践

- 具备数据挖掘、数据仓库或商业智能项目管理经验；
- 熟悉大数据典型应用场景和解决方案，具备方案汇报、论证能力以及良好的文档撰写能力；
- 具备大数据建设项目咨询经验，能够完成数据需求说明书和数据分析报告等文档的编写。

6.10.2 大数据培训师

a) 综合能力

- 具备良好的学习能力，善于快速学习大数据知识与技术，并进行转化、归纳、总结；
- 具备良好的逻辑思维能力，能够输出结构化人才培养方案及课程体系；
- 具备良好的沟通交流能力，善于与需求方、学员进行沟通交流；
- 熟悉大数据产业发展趋势、所在行业知识与应用案例、大数据主流技术体系；
- 熟悉大数据技术人才成长规律；

b) 专业知识

- 熟悉统计学、应用数学、计算机基础、大数据领域技术原理与应用；
- 掌握大数据课程开发方法；
- 掌握大数据教学设计原理；
- 了解教育学、心理学、社会学等方面的知识；

c) 技术技能

- 掌握需求调研工具，如问卷法、访谈法等；
- 掌握常用的培训方案设计与课程开发工具，如ADDIE、PPT等相关软件；

- 熟悉常用的授课技巧，如TTT、SPOT引导技术；
- 掌握大数据领域实战教学能力，如Python编程教学、机器学习实操、大数据平台实操演示等；
- 熟悉常用教学评估方法，如柯氏四级评估方法；

d) 工程实践

- 具备大数据相关培训需求下的方案设计与实施经验；
- 具备一定时间量的教学实践经验，能够独立负责课堂；
- 具备相当的文档撰写能力，能够完成培训需求说明书、咨询报告和教学方案等文档的编写；
- 具备一定的大数据培训相关的咨询经验。

6.10.3 社群数据管理师

a) 综合能力

- 了解大数据发展历史和现状，关注企业数字营销下私域社群流量的用户大数据的发展趋势；
- 熟悉社群管理的典型应用场景，对私域流量运营的战略决策和目标制定有着深刻理解；
- 具备社群管理数据分析和能力，善于从数据中发现和定位问题，并调整优化；
- 具备良好的沟通表达及团队合作能力；

b) 专业知识

- 掌握现行社群管理相关标准、规范及相关法律法规；
- 熟悉新媒体社交产品规则，掌握私域社群用户的大数据管理的方法，并确保数据的安全；
- 掌握内容运营、用户运营的知识和方法，通过数据标签实现精准用户运营；

c) 技术技能

- 掌握企业微信及主流私域运营和数据统计和管理工具的应用；
- 熟悉图文编辑、视频剪辑软件的应用，如PS、PR、剪映等；
- 掌握社群运营绩效衡量、用户画像分析、产品服务数据等数据分析方法；
- 掌握从社群定位、标准作业程序到团队协同的综合管理方法，维持社群内容的合规性，确保所有社群用户和交互数据的安全可追溯；

d) 工程实践

- 具备社群管理整体规划、设计及实施经验；
- 具备典型业务场景下社群建设、精准用户引流、用户留存及转化的数据统计及效果评估经验；
- 具备私域社群用户运营全流程及数据安全运营经验；
- 具备社群管理运营复盘分析和持续改进的能力。

附 录 A
(资料性附录)
大数据产业人才岗位能力提升

A. 1 大数据产业人才岗位能力提升内容

岗位能力提升内容应包括：

- a) 软技能等相关综合能力提升；
- b) 基础知识、专业知识等相关知识提升；
- c) 基本技能、专业技能等相关技术技能提升；
- d) 基于项目经验的工程实践能力提升。

A. 2 大数据产业人才岗位能力提升阶段和方式

大数据产业人才岗位能力提升分为岗前提升和在岗提升两个阶段，这两个阶段共同构成大数据相关岗位产业人才不同阶段和能力水平的终身教育体系。

- a) 岗前提升方式，包括：
 - 1) 理论教学；
 - 2) 理论与实践一体化教学；
 - 3) 项目实训、企业实习等方式。
- b) 在岗提升方式，包括：
 - 1) 内部在岗培训；
 - 2) 外部脱岗培训；
 - 3) 项目实践或导师辅导等。

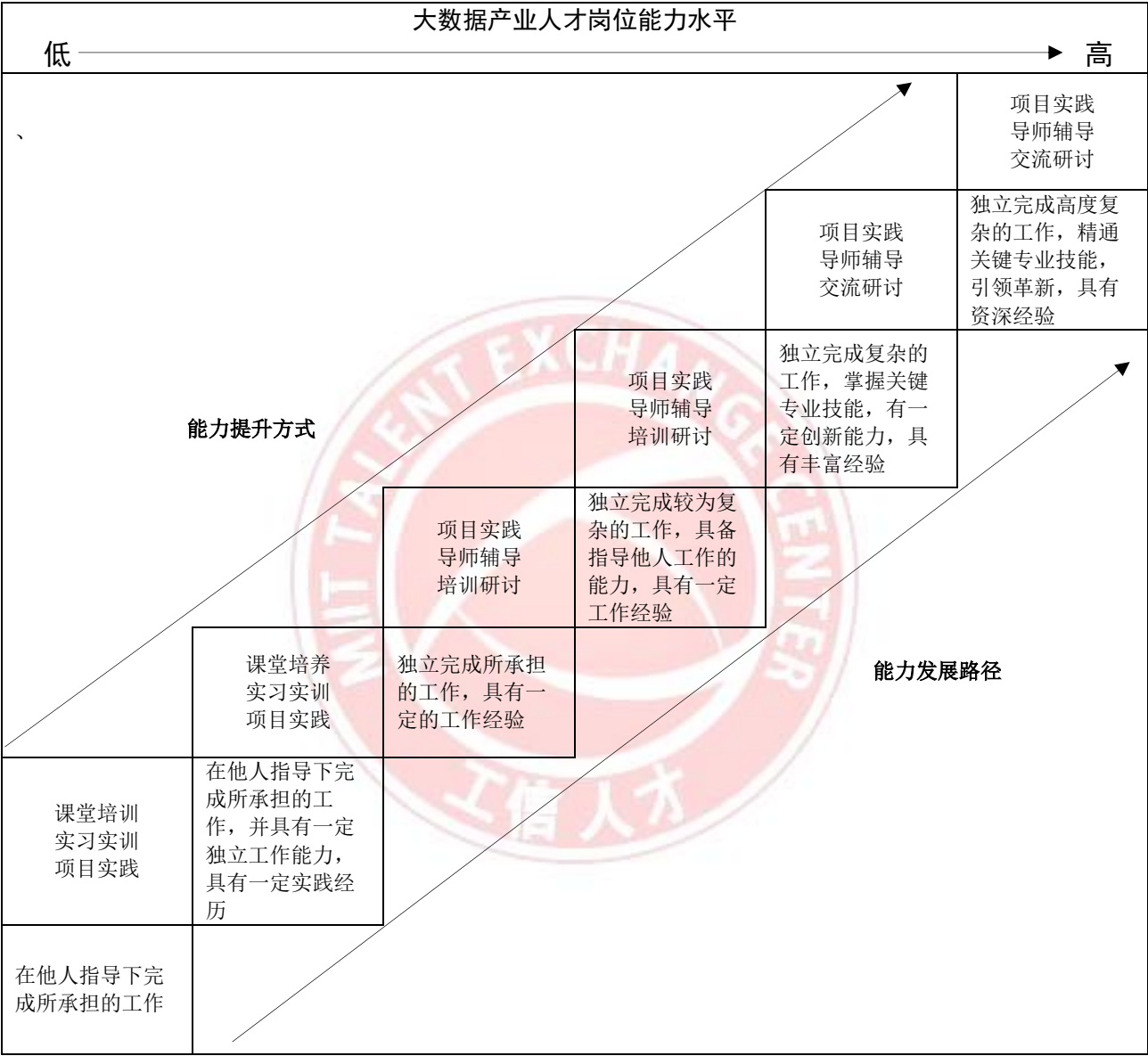
A. 3 大数据产业人才岗位能力提升活动供给类别

大数据产业人才岗位能力提升活动供给包括：

- a) 教育、培训机构培养：符合要求的各级教育机构（普通高校、中等和高等职业院校等）及培训机构应根据大数据产业人才各岗位能力要求，制定人才能力提升方案，为大数据产业及企业培养合格的从业人员，满足个人发展需要；
- b) 企业培养：企业结合业务发展需要，应根据大数据产业人才各岗位能力要求有针对性、有计划地实施岗位能力提升计划，满足个人发展需要，增强企业竞争力；
- c) 个人培养：从业人员根据个人发展计划，做好职业规划与岗位定位，对标大数据产业人才岗位能力要求，不断积累提高综合能力，积累专业知识、技术技能和工程实践经验。

大数据产业人才岗位能力提升路径见图A. 1。

图A. 1 大数据产业人才岗位能力提升路径



附 录 B
(资料性附录)
大数据产业人才岗位能力评价

B.1 大数据产业人才岗位能力评价方法

对从业人员进行评价和定级，评价结果可以作为大数据产业人才能力胜任、职业发展等活动的依据。评价方式包括：

- a) 综合能力主要通过笔试或答辩等方式进行评价；
- b) 专业知识主要通过笔试考核的方式进行评价；
- c) 技术技能主要通过实验考核方式进行评价；
- d) 工程实践主要通过成果评价方式进行评价。

B.2 大数据产业人才岗位能力评价等级

大数据产业人才岗位能力评价等级可以分为初、中、高级三级，能力分为9等。

a) 初级（1—3级）：在他人指导下完成所承担的工作，并具有一定独立工作能力，具有一定实践经历；

b) 中级（4—6级）：独立完成较为复杂的工作，具备指导他人工作的能力，具有一定工作经验；

c) 高级（7—9级）：独立完成高度复杂的工作，精通关键专业技能，引领革新，具有资深经验。

B.3 大数据产业人才岗位能力等级评价权重

大数据产业人才岗位能力等级评价权重表如下：

大数据产业人才岗位能力等级评价权重表

评价维度		专业知识	技术技能	工程实践/综合能力
岗位等级		评价分值权重		
高级	9 级	20%	30%	50%
	8 级			
	7 级			
中级	6 级	50%	25%	25%
	5 级			
	4 级			
初级	3 级	70%	25%	5%
	2 级			
	1 级			
备注		评价总分满分为 100 分，由综合能力、专业知识、技术技能、工程实践四项评价维度的权重总分所得。		

参 考 文 献

- [1] GB/T 38673-2020 《信息技术 大数据 大数据系统基本要求》
 - [2] GB/T 38672-2020 《信息技术 大数据 接口基本要求》
 - [3] GB/T 37721-2019 《信息技术 大数据分析系统功能要求》
 - [4] GB/T 38675-2020 《信息技术 大数据 计算系统通用要求》
 - [5] GB/T 38633-2020 《信息技术 大数据 系统运维和管理功能要求》
-

