

大数据应用开发（Java）

职业技能等级标准

标准代码：510009

（2021年2.0版）

国信蓝桥教育科技（北京）股份有限公司 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 | 4 |
| 参考文献 | 12 |

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：国信蓝桥教育科技(北京)股份有限公司、工业和信息化部人才交流中心、泸州职业技术学院、重庆科创职业学院、山西大学、桂林电子科技大学。

本标准主要起草人：李建伟、倪光南、吴文虎、孔颖、景征骏、李秉璋、余萍、卜磊、茨木·俊秀、加藤·直树、郑未、张航、颜群、张崇杰、王树生、张建平、李伟、魏素荣、卢锡文、侯大成、李艳萍、单宝军、卢正才、曹小平。

声明：本标准的知识产权归属于国信蓝桥教育科技(北京)股份有限公司，未经国信蓝桥教育科技(北京)股份有限公司同意，不得印刷、销售。

1 范围

本标准规定了大数据应用开发（Java）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应用开发（Java）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 35295—2017 信息技术 大数据 术语

GB/T 22263.8—2010 物流公共信息平台应用开发指南 第 8 部分：软件开发管理

GB/T 36462—2018 面向组件的虚拟样机软件开发通用要求

GB/T 35274—2017 信息安全技术 大数据服务安全能力要求

GB/T 35589—2017 信息技术 大数据 技术参考模型

中国大百科全书·电子学与计算机

3 术语和定义

GB/T 35295—2017 界定的以及下列术语和定义适用于本标准。

3.1

大数据应用 big data application

运用编程手段或非编程手段完成的大数据系统（big data system）的全部或部分功能。

3.2

关系（型）数据库 relational database

数据按关系模型来组织的数据库。

3.3

数据采集 data acquisition

又称数据获取，是利用某种装置，从系统外部采集数据并输入到系统内部的一个接口。

3.4

数据预处理 data preprocessing

数据预处理是指在主要的处理以前对数据进行的一些处理。主要是清理异常值、纠正错误数据。

3.5

数据分析 data analysis

用适当的统计分析方法对收集来的大量数据进行分析,将它们加以汇总和理解并消化,以求最大化地开发数据的功能,发挥数据的作用。

3.6

数据可视化 data visualization

数据可视化旨在借助于图形化手段,清晰有效地传达与沟通信息。

3.7

NoSQL Not Only SQL

泛指非关系型的数据库。

3.8

前台系统 foreground system

相对于系统的中台和后台,一般指一个系统中与用户交互的部分。

3.9

功能测试 functional testing

是通过测试来检测每个功能是否都能正常使用,着眼于程序外部结构,不考虑内部逻辑结构,主要针对软件界面和软件功能进行测试。

3.10

测试用例 test case

为特定目标而开发的一组测试输入、执行条件和预期结果,其目标可以是测试某个程序路径或核实是否满足某个特定的需求。

4 适用院校专业

4.1 参照原版专业目录

中等职业学校:移动应用技术与服务、计算机应用、计算机网络技术、软件与信息服务

务。

高等职业学校：大数据技术与应用、云计算技术与应用、物联网应用技术、人工智能技术服务、计算机应用技术、计算机网络技术、软件技术、嵌入式技术与应用。

高等职业教育本科学校：大数据技术与应用、计算机应用工程、信息安全与管理、软件工程、网络工程。

应用型本科学校：数据科学与大数据技术、人工智能、计算机科学与技术、信息工程、信息安全、网络工程、软件工程、信息与计算科学。

4.2 参照新版职业教育专业目录

中等职业学校：移动应用技术与服务、计算机应用、计算机网络技术、软件与信息服务。

高等职业学校：大数据技术、云计算技术应用、物联网应用技术、人工智能技术应用、计算机应用技术、计算机网络技术、软件技术、嵌入式技术应用。

高等职业教育本科学校：大数据工程技术、云计算技术、人工智能工程技术、计算机应用工程、信息安全与管理、软件工程技术、网络工程技术。

应用型本科学校：数据科学与大数据技术、人工智能、计算机科学与技术、信息工程、信息安全、网络工程、软件工程、信息与计算科学。

5 面向职业岗位（群）

【大数据应用开发（Java）】（初级）：主要面向大数据公司、互联网公司、软件开发公司、软件运维公司、软件营销公司等 IT 类公司，以及政府机关、企事业单位的信息管理与服务部门，从事软件运行维护、软件外包服务、软件销售及售后服务等基础性工作。

【大数据应用开发（Java）】（中级）：主要面向大数据公司、互联网公司、软件开发公司、软件运维公司等 IT 类公司，以及政府机关、企事业单位的信息管理与服务部门，从事软件开发，软件测试、数据管理等工作。

【大数据应用开发（Java）】（高级）：主要面向大数据公司、互联网公司、软件开发公司、软件运维公司等 IT 类公司，以及政府机关、企事业单位的信息管理与服务部门，从事项目需求分析，系统框架设计，大数据分析等工作。

6 职业技能要求

6.1 职业技能等级划分

大数据应用开发（Java）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。

【大数据应用开发（Java）】（初级）：能根据应用系统的业务需求，运用面向过程和面向对象编程技术完成应用系统编程；能针对应用系统设计测试用例、准备测试数据，完成测试的执行与分析；能根据工作安排，完成与客户沟通、技术支持与服务等工作。

【大数据应用开发（Java）】（中级）：能根据应用系统的业务需求，运用 Java 高级 API 和高级机制完成应用系统编程并能修正基本的程序缺陷；能运用基本的数据结构和算法、多线程并发编程、SQL 优化等技术提升应用程序性能；能编程获取网络数据，并将其存储在关系数据库或非关系数据库中。

【大数据应用开发（Java）】（高级）：能根据应用系统的业务需求，综合运用 JavaEE 和框架技术，完成应用系统的前后端程序开发；对大数据应用平台有基本了解，能通过工具或编程手段汇集数据并管理数据，能对这些数据做基本的统计分析，完成可视化开发。

6.2 职业技能等级要求描述

表 1 大数据应用开发（Java）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. 大数据应用基础编程 | 1.1 代码编写环境搭建 | 1.1.1 根据开发团队的要求，正确安装配置 JDK； 1.1.2 根据 Java 程序工作机制，验证 Java 程序的编译和运行是否正确； 1.1.3 能使用 EditPlus 等基础工具编写程序源代码； 1.1.4 能利用 Eclipse 等常见集成开发环境创建工程项目，并按规范对文件命名。 |
| | 1.2 面向过程代码编写 | 1.2.1 能运用 Java 数据类型和运算符完成简单运算； 1.2.2 能熟练运用分支、循环等流程控制完成较复杂程序设计； 1.2.3 能运用 Java 中的“方法（Method）”完成代码块封装； 1.2.4 能熟练利用 JavaSE 的 String API 完成字 |

| | | |
|--------------|-----------------|--|
| | | <p>字符串存取和运算；</p> <p>1.2.5 能熟练运用数组存取数据。</p> |
| | 1.3 面向对象代码编写 | <p>1.3.1 理解类和对象机制，熟练运用 Java 的面向对象机制，用“类”的语法封装对象的行为和状态；</p> <p>1.3.2 能熟练运用继承和多态机制编写复用度高的 Java 代码；</p> <p>1.3.3 能熟练运用接口机制设计灵活的 Java 程序；</p> <p>1.3.4 能熟练运用 Java 语言的包和访问控制机制提升代码安全性。</p> |
| 2. 大数据应用产品测试 | 2.1 测试计划与测试环境搭建 | <p>2.1.1 能对被测大数据应用产品进行测试需求分析，协作制定测试计划，控制测试进度，把握测试风险；</p> <p>2.1.2 能根据被测大数据应用产品的特点和项目管理要求，选择质量管理软件；</p> <p>2.1.3 能根据测试需求分析说明文档，设计测试流程；</p> <p>2.1.4 能模拟被测大数据应用产品的运行环境，独立搭建测试环境；</p> |
| | 2.2 测试用例设计 | <p>2.2.1 能运用等价类划分法设计测试用例；</p> <p>2.2.2 能运用边界值分析法设计测试用例；</p> <p>2.2.3 能运用判定表驱动法设计测试用例；</p> <p>2.2.4 能运用因果图法设计测试用例；</p> <p>2.2.5 能运用场景法设计测试用例；</p> <p>2.2.6 能运用正交试验法设计测试用例；</p> <p>2.2.7 能运用错误推测法设计测试用例。</p> |
| | 2.3 测试大数据获取 | <p>2.3.1 针对大数据应用产品的特点，设计数据规模；</p> <p>2.3.2 能运用数据采集工具（非编程）获取测试数据；</p> <p>2.3.3 能根据大数据应用产品的测试要求，检查数据的完整性和一致性；</p> <p>2.3.4 能根据大数据应用产品的测试要求，清除无效数据；</p> <p>2.3.5 能根据大数据应用产品的测试要求，运用数据转换工具对数据进行再加工。</p> |
| | 2.4 测试执行与分析 | <p>2.4.1 能使用各种黑盒测试方法完成测试工作；</p> <p>2.4.2 能运用主流测试软件进行测试管理；</p> <p>2.4.3 能执行每个测试用例，提交缺陷报告，跟踪缺陷的处理流程，熟练使用缺陷管理工具；</p> |

| | | |
|------------|----------|--|
| | | <p>2.4.4 能对测试过程和被测大数据应用产品质量进行评估，完成并提交测试报告；</p> <p>2.4.5 能对测试过程和测试结果进行有效管理，并提出过程改进建议。</p> |
| 3. 技术支持与服务 | 3.1 客户沟通 | <p>3.1.1 能与团队成员协作，开展市场调研，与准客户进行需求沟通；</p> <p>3.1.2 能根据大数据应用产品市场定位，与团队协作开展客户心理分析和差异性分析；</p> <p>3.1.3 能与团队成员协作，采用多种沟通模式和沟通技巧，做好客户服务。</p> |
| | 3.2 售前支持 | <p>3.2.1 能编制大数据应用产品或解决方案的宣传材料和应用演示文稿；</p> <p>3.2.2 能配合销售端开展市场机会跟进、项目方案准备、技术方案宣讲和应用系统演示；</p> <p>3.2.3 能参与编制招标应标文档，配合销售端开展投标准备、标书讲解、用户答疑及现场投标工作。</p> |
| | 3.3 技术服务 | <p>3.3.1 能现场解决基本的大数据应用环境问题；</p> <p>3.3.2 能进行常见故障处理，独立编写并提交BUG报告；</p> <p>3.3.3 能及时收集并反馈客户的新需求；</p> <p>3.3.4 能根据产品使用说明书，培训客户。</p> |

表 2 大数据应用开发（Java）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. 大数据应用编程 | 1.1 Java 高级 API 编程 | <p>1.1.1 能对数据进行结构化和非结构化分析，熟练运用 List、Set、Map 等接口及其子类存取复杂数据对象；</p> <p>1.1.2 能熟练运用 JavaSE 中的 IO 包完成大数据文件的读写和输入输出控制；</p> <p>1.1.3 能运用 JavaSE 中的 IO 包完成大数据文本解析；</p> <p>1.1.4 能运用网络编程 API 获取网络数据或发送网络数据。</p> |
| | 1.2 Java 高级机制编程 | <p>1.2.1 能运用泛型机制编写更加灵活的 Java 程序；</p> <p>1.2.2 能运用 Java 反射机制监控 Java 程序运行时状态；</p> <p>1.2.3 能运用 Java 异常处理机制编写更健壮的 Java 程序；</p> <p>1.2.4 能运用 Java 注解机制提升代码可读性，提升代码健壮性；</p> <p>1.2.5 能运用 Java8 的新特性完成函数式编程，提升编程效率。</p> |

| | | |
|----------------|-----------------|--|
| | 1.3 代码调试与程序缺陷修正 | <p>1.3.1 能根据程序语法规则,独立完成代码语法的错误识别和修正;</p> <p>1.3.2 能根据软件功能需求,独立完成代码逻辑错误的识别和修正;</p> <p>1.3.3 能通过输入输出调试程序逻辑;</p> <p>1.3.4 能运用 JUnit 进行代码单步调试、断点调试;</p> <p>1.3.5 能独立进行异常处理调试。</p> |
| | 1.4 海量数据获取 | <p>1.4.1 能利用 Java 语言及相关工具包,发送网络请求,并获得响应;</p> <p>1.4.2 能利用 Java 语言完成文档解析,过滤出符合要求的数据;</p> <p>1.4.3 能将数据存储于关系型数据库或新型非关系型数据库中;</p> <p>1.4.4 能利用编程自动探索和发现新的数据,并自动持续地获取数据;</p> <p>1.4.5 在相关法律法规的规范下,合法获取数据,保护隐私数据。</p> |
| 2. 大数据应用程序性能优化 | 2.1 数据结构与算法应用 | <p>2.1.1 能利用线性、树形、图结构组织数据并用 Java 语言实现;</p> <p>2.1.2 能利用顺序存储和链式存储实现数据结构并用 Java 语言实现;</p> <p>2.1.3 能利用二分查找法提升在有序数据中查找数据的程序的性能;</p> <p>2.1.4 能利用选择排序等算法完成普通数据排序任务;</p> <p>2.1.5 能利用快速排序等算法完成更高要求的数据排序任务。</p> |
| | 2.2 多线程并发优势应用 | <p>2.2.1 能熟练使用 Java 的多线程 API 创建线程类;</p> <p>2.2.2 能有效控制线程的启动、终止和暂停;</p> <p>2.2.3 能在多个线程间共享数据并有效控制数据的同步访问;</p> <p>2.2.4 能设计有效的预防死锁的多线程协作程序;</p> <p>2.2.5 能利用多线程机制并发采集数据。</p> |
| | 2.3 SQL 优化 | <p>2.3.1 能编写简洁的 SQL 语句,避免 SQL 解析损耗;</p> <p>2.3.2 能有效利用索引提升数据检索性能;</p> <p>2.3.3 能有效利用内置函数减少数据访问次数,提升数据检索性能;</p> <p>2.3.4 能有效分析 SQL 性能,为后续优化提供依据。</p> |
| 3. 海量数 | 3.1 MySQL 数据 | 3.1.1 能独立安装 MySQL 数据库; |

| | | |
|-----|---------------------|---|
| 据管理 | 库运行管理 | 3.1.2 能运用 MySQL 客户端访问数据库服务器； 3.1.3 根据系统需求，配置数据库管理系统； 3.1.4 能对业务数据库进行备份、还原和迁移； 3.1.5 根据系统需求，控制用户访问权限。 |
| | 3.2 关系型数据库 SQL 编程 | 3.2.1 能编写 SQL 语句完成单表数据查询和排序； 3.2.2 能利用 MySQL 提供的内置单行函数简化数据查询和运算； 3.2.3 能利用转换函数完成数据格式的转换； 3.2.4 能编写 SQL 语句完成多表连接查询； 3.2.5 能编写 SQL 语句完成子查询； 3.2.6 能利用组函数完成分组统计； |
| | 3.3 MySQL 数据库数据对象管理 | 3.3.1 熟练运用数据操纵语言完成数据的新增、删除和修改； 3.3.2 根据业务需求，完成事务控制； 3.3.3 根据业务需求，完成表的维护，并有效设置数据约束； 3.3.4 根据业务需求，完成视图、序列和索引等数据库对象的管理； |
| | 3.4 NoSQL 数据库海量数据管理 | 3.4.1 能利用 Redis 作为数据缓存，提升数据存取效率，掌握其 Java API 的运用； 3.4.2 能利用 HBase 存储面向列的海量数据，并掌握其 Java API 的运用； 3.4.3 能利用 MongoDB 存储海量文档型数据，掌握其 Java API 的运用； 3.4.4 能利用 Neo4J 等图形数据库存储海量图结构的数据，掌握其 Java API 的运用。 |

表 3 大数据应用开发（Java）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. 大数据应用前台系统开发 | 1.1 大数据应用前台系统界面开发 | 1.1.1 能熟练使用 Tomcat 和 Eclipse 完成 JavaWeb 项目开发环境的搭建； 1.1.2 能使用 JSP 基础语法完成基本的动态页面开发； 1.1.3 能使用 EL 表达式和 JSTL 标签简化 JSP 代码； 1.1.4 能使用 JSP 和 JavaBean 技术实现数据库访问，动态呈现数据库内容。 |
| | 1.2 大数据应用前台系统服务端程序开发 | 1.2.1 能熟练使用 Servlet3.0 API 开发服务端程序； 1.2.2 能在 MVC 设计模式下编写服务端程序； 1.2.3 能利用数据库访问接口，实现数据持久化操作； 1.2.4 能在三层架构分层模式下完成数据库访问； |

| | | |
|--------------|-----------------|--|
| | | <p>1.2.5 综合利用 JSP、Servlet 系列 API、Apache Commons 系列 API、MVC 模式、三层架构模式实现数据分页呈现和文件上传下载；</p> <p>1.2.6 熟悉协同开发环境，利用版本控制工具（如 Git、SVN）管理代码版本；</p> <p>1.2.7 能使用 Eclipse 编写单元测试代码，对服务端程序进行调试。</p> |
| | 1.3 企业级框架应用 | <p>1.3.1 能熟练使用 Spring 应用程序框架作为对象容器，并在其支持下完成 AOP 编程；</p> <p>1.3.2 能熟练运用 MyBatis 框架完成 ORM，提升数据库访问质量；</p> <p>1.3.3 能使用 Spring MVC 框架优化 MVC 设计；</p> <p>1.3.4 能使用 Restful API 设计跨平台跨语言的数据接口；</p> <p>1.3.5 使用 Spring Boot 快速开发微服务；</p> <p>1.3.6 使用 Spring Cloud 整合微服务项目。</p> |
| 2. 大数据应用平台建设 | 2.1 大数据应用平台搭建 | <p>2.1.1 能根据大数据应用产品的运行要求，选择恰当的硬件环境；</p> <p>2.1.2 能根据大数据应用产品的运行要求，选择并安装操作系统（如 CentOS、Ubuntu）；</p> <p>2.1.3 能在操作系统上安装并正确配置 Hadoop、Hive、HBase、MySQL 等大数据分析平台所需的软件；</p> <p>2.1.4 能正确配置 Hadoop、Hive、HBase 等软件的集群环境。</p> |
| | 2.2 大数据应用平台运行维护 | <p>2.2.1 能通过 shell 脚本或者监控工具收集大数据应用平台主机信息，如内核、内存、负载、磁盘存储、进程数等；</p> <p>2.2.2 能通过 shell 脚本或者监控工具收集 Hadoop 等进程级应用的运行信息，如 JVM 堆内存、GC 时间、吞吐量等；</p> <p>2.2.3 能对采集的指标进行分析并识别出基本的平台运行风险；</p> <p>2.2.4 能解决大数据平台常见的环境配置问题和运行问题。</p> |
| | 2.3 海量数据汇集与管理 | <p>2.3.1 综合应用各种工具将旧有系统数据迁移到大数据平台；</p> <p>2.3.2 综合应用各种工具自动聚集分散的数据，如各服务器上的日志；</p> <p>2.3.3 能利用编程手段或管理工具对数据进行管理（上传、删除、查询等）；</p> <p>2.3.4 能利用编程手段或管理工具进行数据备份和恢复。</p> |
| 3. 大数据 | 3.1 大数据预处理 | 3.1.1 能利用编程方式检查数据的完整性和一 |

| | | |
|--------------|------------|---|
| 基础分析与可视化产品开发 | 理 | 致性； 3.1.2 能利用编程方式清除无效数据； 3.1.3 能利用编程方式填补缺失域； 3.1.4 能利用编程方式或数据转换工具对数据进行再加工，如连续数据离散化。 |
| | 3.2 数据统计分析 | 3.2.1 能根据大数据应用产品的需求完成数据的频数分析； 3.2.2 能根据需求完成数据的集中趋势分析； 3.2.3 能根据需求完成数据的离散程度分析； 3.2.4 能根据需求完成数据的分布分析。 |
| | 3.3 大数据可视化 | 3.3.1 能利用大数据平台提供的各种工具为用户提供海量数据环境下的直接查询功能； 3.3.2 能制作柱状图、饼图、折线图等基本图形表示数据挖掘结果； 3.3.3 能制作词云、热点图、三维图等图形表示数据挖掘结果； 3.3.4 撰写数据分析报告对数据挖掘的结果进行解释和评价，转换为能够最终被用户理解的知识。 |

参考文献

- [1] 《中华人民共和国职业分类大典》（2015年）
- [2] 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）
- [3] 《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）
- [4] 《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）
- [5] 《国家职业教育改革实施方案》（国发〔2019〕4号）
- [6] 《中华人民共和国高等教育法》（2018年修正本）
- [7] 《中华人民共和国标准化法》（2017年修正本）