

华为云培训服务建议书

智能数据



总目录

第一部分： 引言	1
1. 大纲概述	2
2. 华为云培训服务	3
第二部分： 华为云智能数据培训服务	4
1. 培训方案	5
2. 课程描述	6

第一部分：导言

1. 大纲概述

1.1 引言

随着云时代的到来，云生态加速演进，云计算产业的从业人员面临着全方位的挑战。华为云引领技术创新，融合技术能力，并针对不同技术领域从业者的学习和进阶需求，精心打造不同级别和内容的课程，致力于客户与合作伙伴团队和人才的同步增值，促进云时代人才生态的发展！

1.2 大纲目标

为提升华为云培训服务的实用性和有效性，满足客户和伙伴的云技能提升需求，2022年云学院培训大纲进一步加强新场景新技术的培训项目，基于模块化的课程设计，更加清晰和系统地构建华为云培训体系，推出通用云服务、人工智能、物联网、鲲鹏等技术领域内容，涉及总裁班、工程师培训、工程师高级培训、开发者培训、高级开发者培训、专家班等多个类别的课程。

1.3 大纲构成

培训大纲由两大部分组成：培训方案和课程描述。

在“培训方案”中，包含推荐学习路径和项目列表两部分：

1) “推荐学习路径”中，说明了某类培训的培训项目进阶、承接关系。培训项目是华为推荐的培训包，一般是针对某一类内容的集合，由一个或多个培训课程组成。

2) “项目列表”中，以列表的形式说明提供的培训项目。

在“课程描述”中，说明了该培训项目的培训对象、入学要求、培训目标、培训课程、培训方式以及培训时长，并提供了清晰的培训路径，便于了解培训内容的层次和递进关系。

2. 华为云培训服务

2.1 通用云服务培训

覆盖华为云的通用基础知识如计算、存储，网络、数据库、云容器及专业架构设计指导与实践，并开设基于 DevOps 的开发实战课程，确保业务上云后的正常运维和业务创新，培训项目结合企业战略、运营和技术发展的目标，能够有效提升人员能力与组织绩效。

2.2 产业云技术培训

聚焦人工智能，物联网，大数据，鲲鹏，CloudVR 的客户及合作伙伴的技能提升诉求，以几类产业的发展趋势与核心技术为基础，提供面向企业管理者，工程师，开发者，高校老师和学生的特定培训课程。

2.3 工具与平台: MOOC

MOOC (Massive Online Open Course, 大规模网络开放课程) 是在线学习中的一种，是一种性价比高、覆盖面大、容易获取的学习方法，学习者可以随时随地开始学习，并轻松地通过在线问答、论坛等在线形式进行经验交流。

学习形式为在线多媒体课程，课程以基础和普及类科目为主。

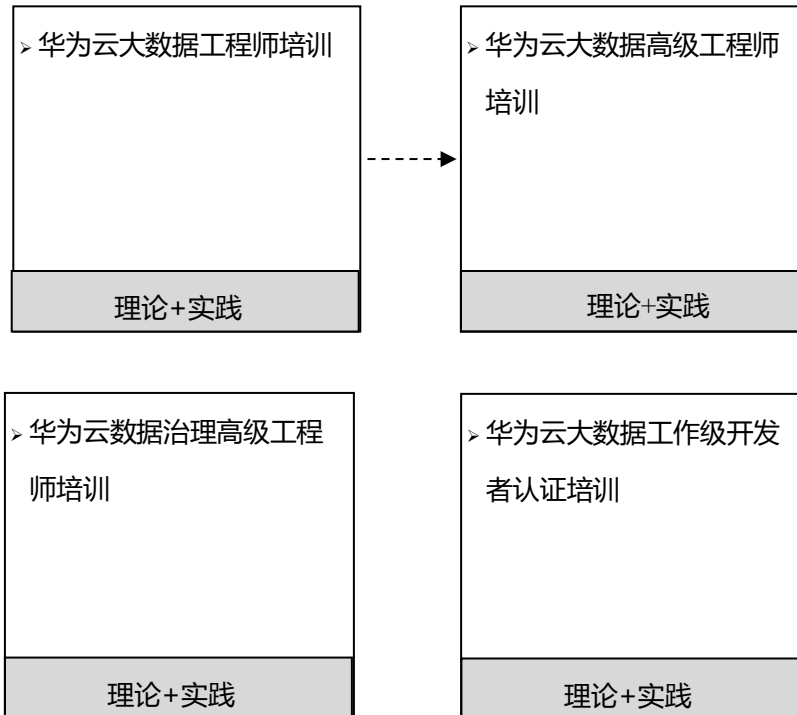
推荐通过在线学习的方式在培训面授之前对学习内容进行预习。

第二部分：华为云智能数据培训服务

1. 培训方案

1.1 智能数据培训方案

1.1.1 推荐学习路径



1.1.2 项目列表

序号	培训项目	培训形式	集中培训(天)	学员人数
1	华为云大数据工程师培训	理论+实践	5	20人以内
2	华为云大数据高级工程师培训	理论+实践	5	20人以内
3	华为云数据治理高级工程师培训	理论+实践	5	20人以内
4	华为云大数据工作级开发者认证培训	理论+实践	5	20人以内

2. 课程描述

2.1 华为大数据工程师培训

定位于培养系统理解大数据架构、技术原理，熟练掌握华为云鲲鹏大数据平台 MRS 服务方案，同时具备大数据平台部署和运维管理能力的工程师。

培训路径

1	大数据前沿趋势及发展趋势 华为鲲鹏大数据介绍	
	理论授课	0.5 天

2	大数据组件技术原理及应用	
	理论授课	4.25 天

3	华为大数据解决方案培训	
	理论授课、上机实习	0.25 天

培训对象

- 希望成为大数据工程师的人员
- 希望了解大数据的市场/销售/项目管理/技术支持工程师

入学要求

- 掌握网络技术基础
- 具备计算机 Linux 基础操作能力

培训目标

完成该项目培训后，您将能够：

- 掌握常用且重要的大数据组件技术原理与架构
- 掌握分布式文件系统 HDFS 的基础操作
- 掌握分布式数据库 HBase 客户端及表操作
- 掌握分布式数据仓库 Hive 的常用 HQL 语句查询

- 掌握分布式搜索引擎 Elasticsearch 进行数据的存储与检索
- 掌握 Flink 进行实时数据的处理等
- 了解典型场景的综合应用等
- 具备大数据售前、大数据项目管理、大数据开发等岗位所必备的知识 and 技能

培训内容

- 大数据发展趋势与鲲鹏大数据
 - ◆ 大数据时代
 - ◆ 大数据的应用领域
 - ◆ 大数据时代企业所面临的机遇和挑战
 - ◆ 华为鲲鹏解决方案
- HDFS 分布式文件系统与 Zookeeper
 - ◆ HDFS 概述及应用场景
 - ◆ HDFS 的系统架构
 - ◆ HDFS 关键特性
 - ◆ Zookeeper 相关概念
 - ◆ Zookeeper 使用方法
- Hive 分布式数据仓库
 - ◆ Hive 概述
 - ◆ Hive 功能与架构
 - ◆ Hive 基本操作
- HBase 技术原理
 - ◆ HBase 基本介绍
 - ◆ HBase 功能与架构
 - ◆ HBase 关键流程
 - ◆ HBase 突出特点
 - ◆ HBase 性能优化
 - ◆ HBase 常用 Shell 命令

- MapReduce 分布式离线批处理和 Yarn 资源协调
 - ◆ MapReduce 和 Yarn 基本介绍
 - ◆ MapReduce 和 Yarn 功能与架构
 - ◆ Yarn 的资源管理和任务调度
 - ◆ 增强特性
- Spark2x 基于内存的分布式计算
 - ◆ Spark 概述
 - ◆ Spark 数据结构
 - ◆ Spark 原理与架构
- Flink 流批一体分布式实时处理引擎
 - ◆ Flink 原理与技术架构
 - ◆ Flink 的 Time 与 Window
 - ◆ Flink 的 Watermark
 - ◆ Flink 的容错
- Flume 海量日志聚合
 - ◆ Flume 简介及架构
 - ◆ Flume 关键特性介绍
 - ◆ Flume 应用举例
- Loader 数据转换
 - ◆ Loader 简介
 - ◆ Loader 作业管理
- Kafka 分布式消息订阅系统
 - ◆ Kafka 简介
 - ◆ Kafka 架构与功能
 - ◆ Kafka 数据管理
- LDAP+Kerberos 安全认证
 - ◆ 统一身份认证管理

- ◆ 目录服务及 Ldap 基本原理介绍
- ◆ 单点登陆及 Kerberos 基本原理介绍
- ◆ 华为大数据安全认证场景架构
- ElasticSearch 分布式全文检索服务
 - ◆ ElasticSearch 简介
 - ◆ ElasticSearch 系统架构
 - ◆ ElasticSearch 关键特性
- Redis 内存数据库
 - ◆ Redis 应用场景与业务流程
 - ◆ Redis 特性及数据类型
 - ◆ Redis 的优化
 - ◆ Redis 应用案例
- 华为大数据解决方案
 - ◆ ICT 行业发展趋势概述
 - ◆ 华为大数据服务
 - ◆ 华为数据湖服务

培训时长

5 天

学员人数

最大 20 人

2.2 华为云大数据高级工程师培训

培训路径

1	大数据开发理论与实践培训	
	理论,实验	5天

培训目标

完成该课程后，您将能够：

- 掌握大数据开发的高级理论和实践方法
- 掌握华为云进行大数据开发的高级使用方法

培训内容

- 第一章 大数据应用开发总指导 (0.5 天)
 - ◆ 大数据主流技术
 - ◆ 大数据场景化解决方案
 - ◆ 大数据应用开发
- 第二章 大数据离线处理场景化解决方案 (1.5 天)
 - ◆ 离线批处理场景方案
 - ◆ 离线批处理框架的介绍、技术原理、参数属性、重要配置等，包括数据存储 HDFS，数据仓库 Hive，离线分析 SparkSQL，数据采集工具 Loader、Sqoop、Kettle 等
 - ◆ 离线批处理实际案例
- 第三章 大数据实时检索场景化解决方案 (1.5 天)
 - ◆ 实时检索场景方案
 - ◆ 分布式数据库介绍、技术原理、参数属性、重要配置等，包括 HBase、ElasticSearch、GraphBase 等
 - ◆ 实时检索实际案例
- 第四章 大数据实时流计算场景化解决方案 (1.5 天)
 - ◆ 实时流计算应用方案
 - ◆ 实时流计算组件的介绍、技术原理、参数属性、重要配置等，包括 Flume、Kafka、Flink、

SparkStreaming、Redis 等

- ◆ 实时流处理配置方案及成功案例

培训时长

5 天, 其中上机时长 3 天

学员人数

最大 20 人

2.3 华为云数据治理高级工程师培训

培训路径

1	华为云数据治理高级工程师培训	
	理论,实验	5天

培训对象

- 大数据工程师
- 数据治理工程师

入学要求

- 了解大数据
- 了解计算机基本知识

培训目标

完成该项目培训后，您将能够：

- 掌握 DGC 的架构
- 掌握 DGC 工作流使用方法
- 掌握数据治理的主要流程和方法

课程内容

- 华为 DAYU 数据治理方法论
 - ◆ 数字化发展趋势和挑战
 - ◆ 数据治理总体框架
 - ◆ 数据治理模块域
 - ◆ 数据治理各模块域之间的关系
- DGC 概述
 - ◆ DGC 产品架构
 - ◆ DGC 功能概述
 - ◆ DGC 应用场景

- ◆ DGC 与其他云服务
- 数据集成
 - ◆ 数据集成概述
 - ◆ CDM 集群操作与管理
 - ◆ DIS 集群操作与管理
- 规范设计
 - ◆ 规范设计概述
 - ◆ 审核中心
 - ◆ 运维中心
 - ◆ 配置中心
- 数据开发
 - ◆ 数据开发概述
 - ◆ 数据开发流程一览
 - ◆ 数据开发作业调度
 - ◆ 数据开发配置管理
- 数据服务
 - ◆ 数据服务概述
 - ◆ 开发 API
 - ◆ 调用 API
 - ◆ 运营管理
- 数据质量
 - ◆ 数据质量监控概述
 - ◆ 业务指标监控
 - ◆ 数据质量监控
- 数据资产
 - ◆ 数据资产概述
 - ◆ 元数据采集管理

- ◆ 数据地图管理
- ◆ 数据权限管理
- ◆ 数据安全治理
- DGC 综合实验

培训时长

5 天

学员人数

最大 20 人

2.4 华为云大数据工作级开发者认证培训

培训路径

1	华为云大数据工作级开发者认证培训	
	理论+实验	5.0 天

培训对象

- 大数据开发工程师
- 数据治理工程师

入学要求

- 了解大数据组件
- 了解数据治理理论
- 了解计算机基本知识

培训目标

完成该项目培训后，您将能够：

- 理解一站式大数据平台 MRS 的架构
- 掌握 MRS 常用大数据组件使用方法
- 理解数据湖治理中心 DGC 的架构
- 掌握 DGC 工作流使用方法
- 掌握华为数据湖探索服务 DLI 使用方法
- 了解华为大数据搬迁方案

课程内容

- 大数据挑战&发展趋势
 - ◆ 大数据发展现状
 - ◆ 大数据落地的挑战
 - ◆ 大数据的发展趋势

- 华为大数据解决方案
 - ◆ MRS 大数据服务简介
 - ◆ MRS 大数据服务优势介绍
 - ◆ MRS 大数据服务应用场景
 - ◆ 华为大数据解决方案
- 华为云 MRS 服务开发与应用
 - ◆ MRS 集群部署开通
 - ◆ 基础组件赋能实操 (HDFS/MR/YARN/Hive/Tez/HBase/Spark)
 - ◆ 数据接入组件赋能实操 (Flume/Sqoop/Kafka)
 - ◆ 高级组件赋能实操 (ClickHouse/Impala/Kudu/Flink)
 - ◆ MRS 特性介绍
- 大数据搬迁方案与实战
 - ◆ 数据迁移工具介绍
 - ◆ 云上大数据分层迁移方案
 - ◆ 数据校验方案
- 华为云 DGC 服务开发与应用
 - ◆ DGC 产品简介及开通
 - ◆ 华为数据治理方法论
 - ◆ DGC 数据集成
 - ◆ DGC 数据开发
 - ◆ 规范设计
 - ◆ 数据质量平台
 - ◆ DGC 数据资产
 - ◆ DGC 数据服务
- 华为云 DLI 服务开发与应用
 - ◆ DLI 产品简介及开通



- ◆ DLI 特性&架构介绍
- ◆ DLI 案例开发

培训时长

5 天

学员人数

最大 20 人