

《数据湖应用与发展》

讲授专家：李勇

培训对象：数据管理（总监、经理、主管）、业务管理人员、数字化转型部门、产品及运营中心管理人员、技术中心管理人员等；

课程时间：1天（6小时）

课程背景：

对许多企业来说，数据变得越来越重要，数据的重要性在多个方面得到了凸显。事实上，企业正在用核心数据来改善自己的运营。本课程将首先介绍什么是数据，它与企业的相关性，以及企业如何利用这些数据以数字化的方式改变自己。为了利用数据，企业需要能存储海量数据的“仓库”，在当今时代，这些仓库不再被称为“数据仓库”，相反，它们被称为“数据湖”。本课程将会对数据湖的发展和應用做详细解析。

课程收益：

- 学习数据湖的基本概念；
- 掌握数据湖对企业的影响；
- 掌握数据湖的典型应用；
- 掌握数据湖的构建方法；

课程大纲：

单元	大纲	内容
单元一	什么是数据湖	<ol style="list-style-type: none">1. 解码数据湖2. 数据湖必清的两个概念3. 什么样的数据可以进数据湖4. 数据湖和数据中台，数据仓库，数据集市，数据治理等之间的关联关系和差异
单元二	数据湖对企业的影响	<ol style="list-style-type: none">1. 为什么企业要做数据湖2. 数据湖对管理和业务的价值3. 数据湖对企业不同角色的价值4. 企业应用数据湖前后的对比

单元三	典型的数据湖应用解析	<ol style="list-style-type: none">1. 零售行业数据湖应用解析2. 金融行业数据湖应用解析3. 制造行业数据湖应用解析4. 医疗行业数据湖应用解析5. 教育行业数据湖应用解析6. 轨道交通行业数据湖应用解析
单元四	数据湖是如何工作的	<ol style="list-style-type: none">1. 数据采集，汇聚2. 数据处理3. 数据分析4. 数据存储5. 数据价值统计
单元五	数据湖的构建方法	<ol style="list-style-type: none">1. 原始数据进入对象存储2. 优化原始数据文件3. 添加元数据工具4. 下游数据集成数据资产
单元六	避免数据沼泽	<ol style="list-style-type: none">1. 什么是数据沼泽2. 如何避免数据沼泽3. 数据湖合理运营的机制