# 大数据开发工程师需掌握技术

大数据开发行业在我国是朝阳行业，目前正处于成长期，对大数据人才的需求量巨大，而且不同公司对大数据工作的要求也不尽相同，有的强调数据库编程、有的突出应用数学和统计学知识等，因此，也衍生出了多个大数据开发方向。

想成为大数据开发工程师，需要从整体上了解大数据开发服务，以下是比较全面的大数据开发平台系统。



对于大数据的学习，可以从以下五个阶段进行着手：

**第一阶段：JavaSE+MySql+Linux**

**学习内容：**Java 语言入门 → OOP 编程 → Java 常用Api、集合 → IO/NIO → Java 实用技术 → Mysql 数据库 → 阶段项目实战 → Linux 基础 → shell 编程；

**学习目标：**学习java语言，掌握java程序编写、面向对象程序开发，掌握MySql体系结构及核心编程技术，打好 Linux 基础，为后续学习提供良好的语言基础。

**第二阶段：Hadoop 与生态系统**

**学习内容：**Hadoop → MapReduce → Avro → Hive → HBase → Zookeeper → Flume → Kafka → Sqoop → Pig；

**学习目标：**掌握大数据学习基石Hadoop、数据串行化系统与技术、数据的统计分析、分布式集群、流行的队列、数据迁移、大数据平台分析等。

**第三阶段：Storm 与Spark 及其生态圈**

**学习内容：**Storm → Scala → Spark → Spark SQL → Spark Streaming → Spark 机器学习；

**学习目标：**让大家拥有完整项目开发思路和架构设计，掌握从数据采集到实时计算到数据存储再到前台展示，所有工作一个人搞定!并可以从架构的层次站在架构师的角度去完成一个项目。

**第四阶段：其他**

**学习内容：**Mahout 机器学习 → R 语言→ Python；

**学习目标：**机器学习领域经典算法的实现，熟练使用 R语法和统计思维，可以基于具体问题建立数学模型，掌握python技术与数据分析，将数据结果以可视化的直观方式展示给目标用户。

**第五阶段：项目实战、技术综合运用**

**学习内容：**某手机公司bug 系统 → 传统广告怎么用大数据 → 类互联网电商网站 → 网站日志收集清洗系统 → 网站流量统计分析系统；

**学习目标：**具备企业级大型完整项目开发能力，综合运用大数据分析知识，完成数据分析、收集、展示的完整流程。

国内大数据行业还处于起步阶段，大数据人才资源匮乏，大数据开发具有一定进入门槛，对于想从事大数据行业的技术人员来说是一大机遇!