# 大数据概述

由于新技术，设备和类似的社交网站通信装置的出现，人类产生的数据量每年都在迅速增长。美国从一开始的时候到2003年产生的数据量为5十亿千兆字节。如果以堆放的数据磁盘的形式，它可以填补整个足球场。在2011年创建相同数据量只需要两天，在2013年该速率仍在每十分钟极大地增长。虽然生产的所有这些信息是有意义的，处理起来有用的，但是它被忽略了。

## **一、什么是大数据？**

大数据是不能用传统的计算技术处理的大型数据集的集合。它不是一个单一的技术或工具，而是涉及的业务和技术的许多领域。

## **二、大数据类型**

大数据包括通过不同的设备和应用程序所产生的数据。下面给出的是一些在数据的框架下的领域。

* 黑匣子数据：这是直升机，飞机，喷气机的一个组成部分，它捕获飞行机组的声音，麦克风和耳机的录音，以及飞机的性能信息。
* 社会化媒体数据：社会化媒体，如Facebook和Twitter保持信息发布的数百万世界各地的人的意见观点。
* 证券交易所数据：交易所数据保存有关的“买入”和“卖出”，客户由不同的公司所占的份额决定的信息。
* 电网数据：电网数据保持相对于基站所消耗的特定节点的信息。
* 交通运输数据：交通数据包括车辆的型号，容量，距离和可用性。
* 搜索引擎数据：搜索引擎获取大量来自不同数据库中的数据。



因此，大数据包括体积庞大，高流速和可扩展的各种数据。它的数据为三种类型。

结构化数据：关系数据。

半结构化数据：XML数据。

非结构化数据：Word, PDF, 文本，媒体日志。

## **三、大数据的好处**

通过保留了社交网络如Facebook的信息，市场营销机构了解可以他们的活动，促销等广告媒介的响应。

利用信息计划生产在社会化媒体一样喜好并让消费者对产品的认知，产品企业和零售企业。

使用关于患者以前的病历资料，医院提供更好的和快速的服务。

## **四、大数据的挑战**

大数据相关的主要挑战如下：

* 采集数据
* 策展
* 存储
* 搜索
* 分享
* 传输
* 分析
* 展示

为了实现上述挑战，企业通常需要企业级服务器的帮助。