对于“大数据”（Big data）研究机构[Gartner](https://baike.baidu.com/item/Gartner" \t "_blank)给出了这样的定义。“大数据”是需要新处理模式才能具有更强的[决策力](https://baike.baidu.com/item/%E5%86%B3%E7%AD%96%E5%8A%9B" \t "_blank)、[洞察](https://baike.baidu.com/item/%E6%B4%9E%E5%AF%9F)发现力和流程优化能力来适应海量、高增长率和多样化的信息[资产](https://baike.baidu.com/item/%E8%B5%84%E4%BA%A7)。

[麦肯锡](https://baike.baidu.com/item/%E9%BA%A6%E8%82%AF%E9%94%A1)全球研究所给出的定义是：一种规模大到在获取、[存储](https://baike.baidu.com/item/%E5%AD%98%E5%82%A8)、管理、分析方面大大超出了传统[数据库](https://baike.baidu.com/item/%E6%95%B0%E6%8D%AE%E5%BA%93)软件工具能力范围的数据集合，具有海量的数据规模、快速的数据流转、多样的数据类型和价值密度低四大特征。

[大数据](https://baike.baidu.com/item/%E5%A4%A7%E6%95%B0%E6%8D%AE/1356941)技术的战略意义不在于掌握庞大的[数据](https://baike.baidu.com/item/%E6%95%B0%E6%8D%AE/33305)信息，而在于对这些含有意义的数据进行专业化处理。换而言之，如果把大数据比作一种产业，那么这种产业实现盈利的关键，在于[提高](https://baike.baidu.com/item/%E6%8F%90%E9%AB%98)对数据的“加工能力”，通过“[加工](https://baike.baidu.com/item/%E5%8A%A0%E5%B7%A5/7353178)”实现数据的“[增值](https://baike.baidu.com/item/%E5%A2%9E%E5%80%BC)”。

从技术上看，大数据与[云计算](https://baike.baidu.com/item/%E4%BA%91%E8%AE%A1%E7%AE%97" \t "_blank)的关系就像一枚[硬币](https://baike.baidu.com/item/%E7%A1%AC%E5%B8%81/2613)的正反面一样密不可分。大数据必然无法用单台的[计算机](https://baike.baidu.com/item/%E8%AE%A1%E7%AE%97%E6%9C%BA)进行处理，必须采用[分布式架构](https://baike.baidu.com/item/%E5%88%86%E5%B8%83%E5%BC%8F%E6%9E%B6%E6%9E%84)。它的特色在于对海量数据进行分布式[数据挖掘](https://baike.baidu.com/item/%E6%95%B0%E6%8D%AE%E6%8C%96%E6%8E%98)。但它必须依托云计算的[分布式处理](https://baike.baidu.com/item/%E5%88%86%E5%B8%83%E5%BC%8F%E5%A4%84%E7%90%86)、[分布式数据库](https://baike.baidu.com/item/%E5%88%86%E5%B8%83%E5%BC%8F%E6%95%B0%E6%8D%AE%E5%BA%93)和[云存储](https://baike.baidu.com/item/%E4%BA%91%E5%AD%98%E5%82%A8)、[虚拟化技术](https://baike.baidu.com/item/%E8%99%9A%E6%8B%9F%E5%8C%96%E6%8A%80%E6%9C%AF)。

随着[云时代](https://baike.baidu.com/item/%E4%BA%91%E6%97%B6%E4%BB%A3" \t "_blank)的来临，大数据（[Big data](https://baike.baidu.com/item/Big%20data)）也吸引了越来越多的关注。分析师团队认为，大数据（Big data）通常用来形容一个公司创造的大量[非结构化数据](https://baike.baidu.com/item/%E9%9D%9E%E7%BB%93%E6%9E%84%E5%8C%96%E6%95%B0%E6%8D%AE" \t "_blank)和[半结构化数据](https://baike.baidu.com/item/%E5%8D%8A%E7%BB%93%E6%9E%84%E5%8C%96%E6%95%B0%E6%8D%AE)，这些数据在下载到关系型数据库用于分析时会花费过多时间和金钱。大数据分析常和[云计算](https://baike.baidu.com/item/%E4%BA%91%E8%AE%A1%E7%AE%97)联系到一起，因为实时的大型数据集分析需要像MapReduce一样的框架来向数十、数百或甚至数千的电脑分配工作。

大数据需要特殊的技术，以有效地处理大量的容忍经过时间内的数据。适用于大数据的技术，包括大规模[并行处理](https://baike.baidu.com/item/%E5%B9%B6%E8%A1%8C%E5%A4%84%E7%90%86)（MPP）[数据库](https://baike.baidu.com/item/%E6%95%B0%E6%8D%AE%E5%BA%93)、数据挖掘、[分布式文件系统](https://baike.baidu.com/item/%E5%88%86%E5%B8%83%E5%BC%8F%E6%96%87%E4%BB%B6%E7%B3%BB%E7%BB%9F)、[分布式数据库](https://baike.baidu.com/item/%E5%88%86%E5%B8%83%E5%BC%8F%E6%95%B0%E6%8D%AE%E5%BA%93)、[云计算平台](https://baike.baidu.com/item/%E4%BA%91%E8%AE%A1%E7%AE%97%E5%B9%B3%E5%8F%B0)、互联网和可扩展的[存储系统](https://baike.baidu.com/item/%E5%AD%98%E5%82%A8%E7%B3%BB%E7%BB%9F)。