

麦肯锡发布

《大数据：创新、竞争和生产力的下一个前沿》报告

麦肯锡全球研究院于 2011 年 5 月发布了《大数据：创新、竞争和生产力的下一个前沿》报告。该报告系统阐述了大数据概念，详细列举了大数据的核心技术，深入分析了大数据在不同行业的应用，明确提出了政府和企业决策者应对大数据发展的策略。麦肯锡认为只要给予适当的政策支持，“大数据”将促进生产力增长并推动创新。

1. 大数据及其作用概览

麦肯锡认为，“大数据”是指其大小超出了典型数据库软件的采集、储存、管理和分析等能力的数据集。该定义有两方面内涵：一是符合大数据标准的数据集大小是变化的，会随着时间推移、技术进步而增长；二是不同部门符合大数据标准的数据集大小会存在差别。目前，大数据的一般范围是从几个 TB 到数个 PB（数千 TB）。

大数据可以发挥重要的经济作用，不但有利于私人商业活动，也有利于国民经济和民生。数据可以为世界经济创造重要价值，提高企业和公共部门的生产率和竞争力，并为消费者创造大量的经济剩余。麦肯锡研究了“大数据”中尚未开发的巨大价值。例如，充分利用“大数据”的零售商将能够将营业利润率提高 60% 以上。“大数据”在公共领域也有极大潜力可挖。如果美国医疗保健行业有效利用“大数据”，就能把成本降低 8% 左右，从而每年创造出 3,000 多亿美元的产值。在欧洲发达国家，如果政府利用“大数据”提高运作效率，那么将节省至少 1,000 亿欧元（约合 1490 亿美元）的成本。而利用个人位置数据（personal location data）提供的服务将可以创造 6,000 亿美元的消费者经济剩余（consumer surplus）。

2. 大数据的关键技术

1) 可用于大数据分析的关键技术

可用于大数据分析的关键技术源于统计学和计算机科学等多个学科，其中关于分析新数据集方法的研究仍在继续。需要注意的是，A/B 测试、回归分析等技术也可应用于小数据集分析。

麦肯锡认为，可用于大数据分析的关键技术主要包括 A/B 测试、关联规则挖掘、分类、数据聚类、众包、数据融合和集成、数据挖掘、集成学习、遗传算法、机器学习、自然语言处理、神经网络、神经分析、优化、模式识别、预测模型、回归、情绪分析、信号处理、空间分析、统计、监

督式学习、无监督式学习、模拟、时间序列分析、时间序列预测模型、可视化技术等。

2) 专门用于处理大数据的关键技术

麦肯锡认为，专门用于整合、处理、管理和分析大数据的关键技术主要包括 Big Table、商业智能、云计算、Cassandra、数据仓库、数据集市、分布式系统、Dynamo、GFS、Hadoop、HBase、MapReduce、Mashup、元数据、非关系型数据库、关系型数据库、R 语言、结构化数据、非结构化数据、半结构化数据、SQL、流处理、可视化技术等。

3) 可视化技术

麦肯锡认为，可视化技术是大数据应用的重点之一，目前主要包括标签云、Clustergram、历史流、空间信息流等技术和应用。

3. 利用大数据的方法

麦肯锡对医疗保健、零售、公共领域、制造、个人位置数据等领域进行了重点分析，提出了利用“大数据”的五种方法：

1) 以时效性更高的方式向用户提供“大数据”。在公共领域，跨部门提供“大数据”能大幅减少检索与处理时间。在制造业，集成来自研发、工程、制造单元的数据可以实现并行工程，缩短产品投放市场的时间。

2) 通过开展数据分析和实验寻找变化因素并改善产品性能。由于越来越多的交易数据都以数字形式存在，各机构可以收集有关产品或用户的更加精确和详尽的数据。

3) 区分用户群，提供个性化服务。“大数据”能帮助机构对用户群进行更加细化的区分，并针对用户的不同需求提供更加个性化的服务。这是营销和危机管理方面常用的方法，但也可以为公共领域等带来变革。

4) 利用自动化算法支持或替代人工决策。复杂分析能极大改善决策效果，降低风险，并挖掘出其他方法无法发现的宝贵信息。此类复杂分析可用于税务机构、零售商等。

5) 商业模式、产品与服务创新。制造商正在利用产品使用过程中获得的数据来改善下一代产品开发，以及提供创新性售后服务。实时位置数据的兴起带来了一系列基于位置的移动服务，例如导航和人物跟踪。

4. 政策制定者应对大数据的策略

大数据不仅是企业竞争和增长的引擎，而且对于提高发达国家和发展中国家的生产率、创新能力和整体竞争力都有着重要作用。政策制定者需要认识到利用大数据可以刺激经济的下一波增长。为帮助企业获取大数据收益，政策制定者需要从如下 6 方面制定相关政策：

1) 加强大数据人才培养

政府可以采取多种措施增加大数据相关人才供给，包括实施教育培养计划、消除从其它国家地区引进人才的障碍等。除此之外，政府应该创造激励措施并对企业管理者进行数据分析技术培训。

2) 制定奖励措施促进数据共享

政府在创造数据共享和交易的有效市场方面可以发挥重要作用，包括制定知识产权方面规则、制定鼓励数据共享的奖励措施、强制要求收集并公开国企财务数据、以及面向公众开放和共享政府部门活动和项目信息等。

3) 制定平衡数据使用与数据安全保护的政策

政策制定者需要加强制定并执行关于商业和个人数据隐私的方针和法律，并通过强大的法律阻挡黑客和其它袭击。当然，政府、非营利组织和私人部门需要开发教育项目，以便公众理解哪些个人信息是可以获取的，如何使用、怎样使用，以及个人是否允许这种使用。

4) 建立有效的促进创新的知识产权框架

在大数据时代，数据价值链中的创新将不断出现，更好地产生和获取数据的技术也将出现。这些创新需要建立有效的知识产权保护体系，促进数据创造价值、数据共享和整合。

5) 克服技术障碍并加速关键技术研发

政策制定者加强制定 IT 工具或数据资源池的标准和指南，鼓励存在缺口的重要领域关键技术研发，推动行业标准制定机构制定覆盖 IT 工具和数据类型标准，并给予资金支持、税收支出和减免、金融支持等激励支持大数据研究。

6) 确保信息通信技术基础设施投资

政策制定者应该使基础设施成为大数据发展的重要组成部分。很多国家对扩建基础设施制定了专门的激励措施。例如，美国政府出台了一系列货币奖励措施，鼓励宽带建设（如农村宽带工程）和实施电子医疗记录。

总之，政策制定者在人才、研发、基础设施和培育创新等关键领域能发挥重要作用，促进企业从大数据中获取最大收益。但是保持企业和公众间权利的平衡是个艰巨的任务，政府在赋予企业在更大范围使用数据以获取潜在收益同时，要减轻公众对隐私和个人信息安全的担忧。

唐川 刘晓 检索 田倩飞 编译自

http://www.mckinsey.com/mgi/publications/big_data/pdfs/MGI_big_data_full_report.pdf