

中国大数据发展调查报告 (2017年)

中国信息通信研究院

2017年3月



版权声明

本调查报告版权属于中国信息通信研究院，并受法律保护。转载、摘编或利用其它方式使用本调查报告文字或者观点的，应注明“来源：中国信息通信研究院”。违反上述声明者，本院将追究其相关法律责任。

报告目录

前言	III
观点摘要	1
一、调查背景	3
(一) 调查方法及样本	3
(二) 报告术语界定	4
二、大数据市场规模	5
(一) 市场总体规模	5
(二) 细分领域规模	5
三、大数据应用	6
(一) 大数据应用现状	6
(二) 大数据应用场景	8
(三) 大数据应用带来的效果	8
(四) 大数据应用的主要障碍	9
(五) 未来大数据投入趋势	9
四、大数据资源	10
(一) 数据规模	10
(二) 数据来源	10
(三) 数据类型	11
五、大数据平台建设模式	12
(一) 建设模式占比	12
(二) 自建大数据平台模式	12
(三) 采购公共云服务模式	14
六、政策需求和资源需求	15
(一) 政策需求	15
(二) 资源需求	15

前言

随着大数据政策环境和技术手段的不断完善，大数据行业应用持续升温，中国企业级大数据市场进入快速发展时期。互联网、电信、金融等开始实际部署大数据平台并付诸实践，带动软件、硬件和服务市场快速发展。为进一步掌握中国大数据发展和应用情况，中国信息通信研究院开展了2016—2017年度中国大数据发展状况的调查。本次活动采用现场访问、电话访谈和在线调查相结合的方式，对中国大数据市场规模、应用现状、数据资源情况、平台建设情况、政策和数据需求等问题进行了调研，共回收有效问卷1,465份。本报告以调查数据为基础，结合行业专家的访谈，力争详实客观的反映中国大数据发展现状，为政府和企业了解中国大数据发展状况和制定相关决策提供参考，也为广大关注大数据的从业人员、专家学者和研究机构提供真实可信的大数据发展报告。

本报告的数据搜集工作得到了相关政府部门、大数据解决方案提供商、企业用户以及社会各界的大力支持。我们在此表示衷心的感谢！同时也对接受大数据调查访问的企业朋友表示最诚挚的谢意！

中国信息通信研究院

2017年3月27日

观点摘要

① 大数据市场规模

2016年中国大数据市场规模为168.0亿元，增速达到45%；预计2017-2020年增速保持在30%以上。

② 大数据应用

大部分企业均已意识到数据分析对企业发展的重要性。

59.2%的受访企业已经成立了数据分析相关部门；同时，27.3%的企业正在计划成立数据部门。

大数据应用得到进一步普及。

35.1%的受访企业已经应用了大数据，与2015年调查结果相比有所提高；同时，22.9%的企业计划将在未来1年内应用大数据，发展前景较为乐观。

企业应用大数据所带来的主要效果包括实现智能决策、提升运营效率和改善风险管理。

其中应用大数据实现智能决策的企业占比最高(55.8%)，其次为提升了运营效率(48.2%)以及更好的管理风险(25.7%)。

企业将进一步加大在大数据领域的投入。

一半以上的受访企业未来计划加大对大数据的投入，其中近20%的企业预计投入增长在50%以上。

③ 大数据资源

企业数据资源仍以内部数据为主。

包括内部生产数据(49.0%)、客户/用户数据(47.8%)以及内部经营管理数据(42.4%)。

企业数据资源量占比最重的为非结构化数据。

66.1%的企业表示非结构化数据的比例在70%以上，22.0%的企业非结构化数据比例为50% - 70%。

4 大数据平台建设模式

企业更倾向本地部署大数据平台。

64.9%的受访企业选择本地部署大数据平台，其中倾向选择自建大数据平台的企业占比达到48.6%。

Hadoop是最受企业欢迎的大大数据平台架构。

接近四成的受访企业倾向选择Hadoop作为大数据平台的技术架构；其次，Spark近两年发展迅速，22.4%的企业倾向选择Spark作为大数据平台的技术架构。

服务器数量在100台以上的企业增加。

自建大数据平台的企业中，31.0%的企业用于大数据平台的服务器数量在100台以上，相比2015年显著提高。

在线数据库服务、在线大数据分析工具和通用IaaS是企业通过公共云服务部署大数据应用时的优先选择。

购买在线数据库服务是企业选择最多的方式，占比43.3%；其次是在线大数据分析工具以及通用IaaS，分别占比37.1%和32.5%。

5 政策需求和资源需求

完善行业标准、健全法律法规对推动大数据发展至关重要。

完善行业标准、健全法律法规是2016年新增选项，也是企业关注度最高的选项，可见大数据行业标准亟需进一步完善；其次为加强个人信息保护（41.2%）和开放更多政府信息公开信息资源（35.1%）。

企业对数据资源的需求更加强烈。

其中企业工商信息依旧是需求最大的数据资源，占比达到49.0%；其次为社会保障数据（33.9%）、劳动就业数据（32.7%）以及市政管理数据（29.4%）。

一、调查背景

（一）调查方法及样本

1、调查方法

本次报告信息收集采用现场访问、电话访谈和在线调查相结合的方式，收集到 1,465 份大数据用户的有效问卷。其中，通过现场访问收集问卷 635 份，通过电话访谈收集问卷 467 份，通过在线调查收集问卷 363 份。

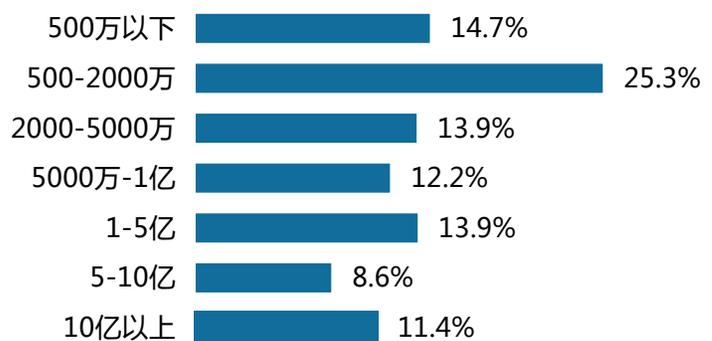
2、样本描述

参与调查企业规模：以中小企业为主，年营业额在 1 亿元以下的企业占总体样本的 66.1%。

参与调查企业所在区域：覆盖华北、东北、华东、中南、西南、西北以及海外地区。

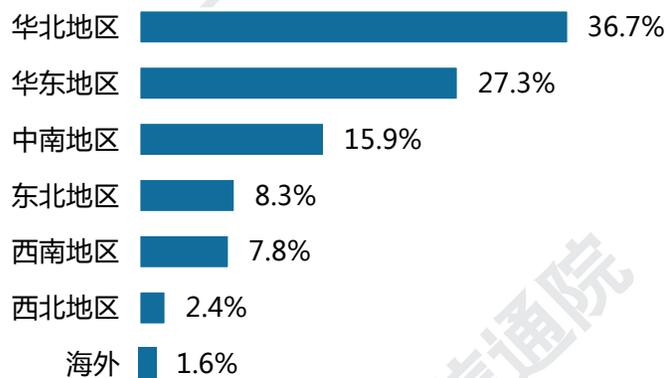
参与调查企业所在行业：包括互联网、软件及信息技术服务、电信、金融、制造、交通及物流、能源、医疗、政府、科教文卫以及农林牧渔等行业。

图1 企业营业额分布 (N=1,465)



数据来源：中国信息通信研究院

图2 企业区域分布 (N=1,465)



数据来源：中国信息通信研究院

（二）报告术语界定

大数据产业：本报告所称大数据产业，包含大数据直接相关的软件、硬件及专业服务。大数据市场规模的计算口径，相应的包含大数据软件、硬件及服务的直接产值，不含因大数据应用产生的附加值。

大数据软件：本报告所称大数据软件，主要指以Hadoop、Spark等技术架构为基础的新型分布式平台软件，以及基于这类平台软件的应用软件。

大数据服务：本报告所称大数据服务，主要包括大数据分析公共云服务及大数据系统相关的咨询、设计和系统集成服务等。

大数据硬件：本报告所称大数据硬件，是企业购买的用于支撑企业内部大数据相关业务的硬件基础设施，主要包括服务器、网络设备和存储设备等。特别说明，大数据系统一般使用通用硬件，技术上并无特殊性，但同时也包含部分经过专门优化的一体机设备。

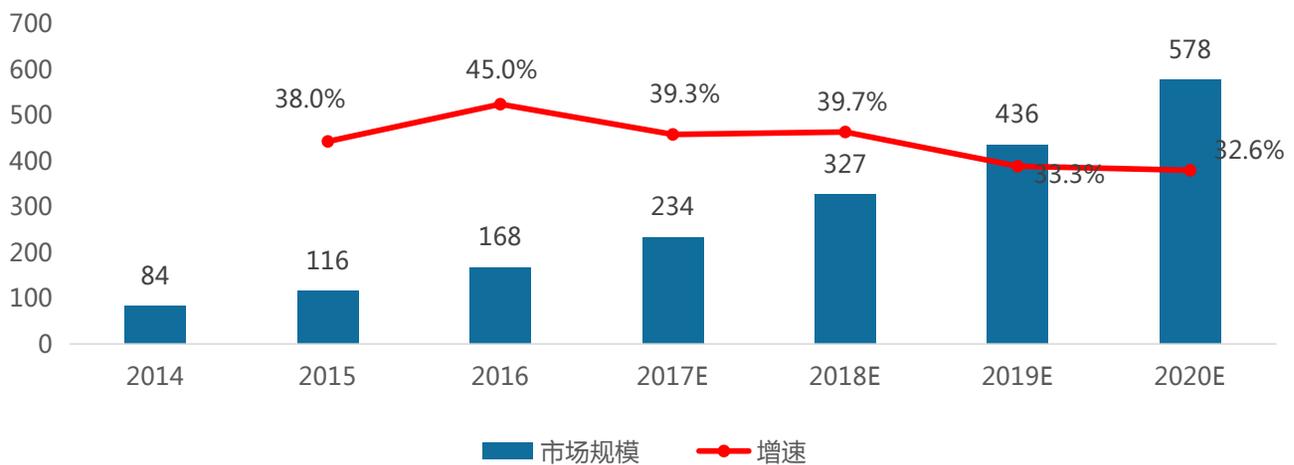
二、大数据市场规模

（一）市场总体规模

目前大数据产业的统计口径尚未建立，对于中国大数据产业的规模，各个研究机构均采取间接方法估算。中国信息通信研究院结合对大数据相关企业的调研测算，2016年中国大数据核心产业的市场规模约为168亿元，较2015年增速达45%。

随着国家政策激励以及大数据应用模式的逐步成熟，未来几年中国大数据市场仍将保持快速增长，预计到2020年中国大数据市场规模将达到578亿元。

图3 中国大数据市场规模及增速(单位：亿元)

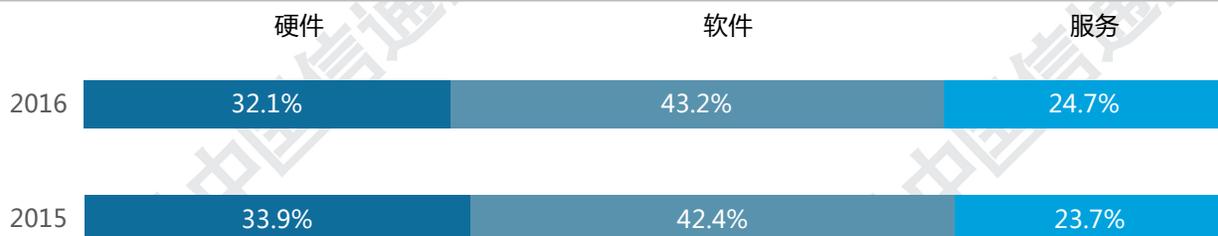


数据来源：中国信息通信研究院

（二）细分领域规模

从大数据市场细分领域来看，近年来，软件和服务在市场整体规模中的比重呈上升趋势，硬件比重逐年减少，整体变化趋势较为平缓。与2015年相比，硬件市场规模为53.9亿元，占比下降1.8%；软件市场规模为72.6亿元，占比提高0.8%；大数据服务市场规模41.5亿元，占比提高1.0%。

图4 中国大数据细分市场构成



数据来源：中国信息通信研究院

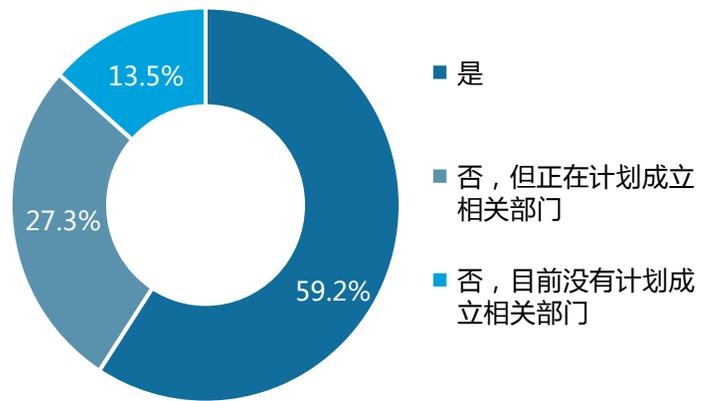
三、大数据应用

（一）大数据应用现状

2016年，国家加大了对大数据应用的推动力度，批复了京津冀等7个国家级大数据综合试验区和超过10个大数据国家工程实验室；同时，针对医疗、交通等行业，有关部门均出台了关于大数据发展的指导意见，继续推动大数据在各垂直领域的应用进程。

近六成企业已成立数据分析相关部门。调查发现，59.2%的受访企业已经成立了数据分析相关部门；同时，27.3%的企业正在计划成立数据部门，绝大部分企业均已意识到数据分析对企业发展的重要性。

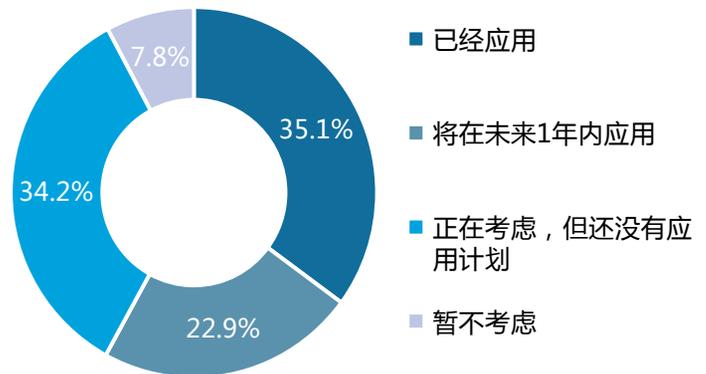
图5 企业数据分析相关部门建设情况 (N=1,465)



数据来源：中国信息通信研究院

超过 1/3 的企业已应用大数据。调查发现，35.1%的受访企业已经应用了大数据，与2015年调查结果相比有所提高，大数据应用得到进一步普及；同时，22.9%的企业计划将在未来1年内应用大数据，发展前景较为乐观。

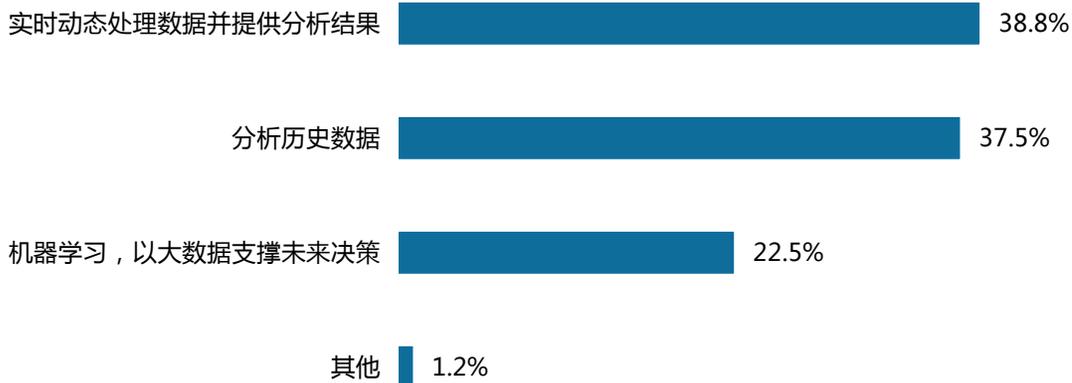
图6 企业大数据应用状况 (N=1,465)



数据来源：中国信息通信研究院

使用实时动态大数据做分析的企业最多。已经应用大数据的企业中，38.8%的企业选择实时动态数据进行处理和分析；其次为历史数据分析，占比为37.5%；另外有22.5%的企业选择了通过机器学习的方式辅助决策。

图7 企业数据分析方式 (N=514)



数据来源：中国信息通信研究院

使用图形图表和智能可视化展现方式的企业越来越多。传统的数据报表依然是企业选择使用最多的数据展现方式，占比为67.5%；而图形图表和智能可视化展现方式被越来越多的企业选择和使用，选择比例分别为58.8%和50.0%。

图8 企业数据分析结果的展现方式 (N=514)

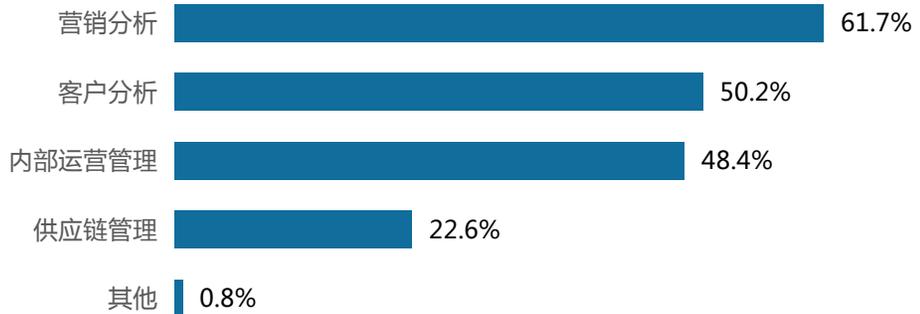


数据来源：中国信息通信研究院

（二）大数据应用场景

营销分析、客户分析和内部运营管理是大数据应用最广泛的三个领域。调查发现，超过60%的企业将大数据应用于营销分析；其次，50.2%的企业将大数据应用于客户分析；另外，将近50%的企业将大数据应用于内部运营管理。相比之下大数据分析在企业供应链管理等方面的应用比例还有待提升。

图9 企业大数据应用场景 (N=514)

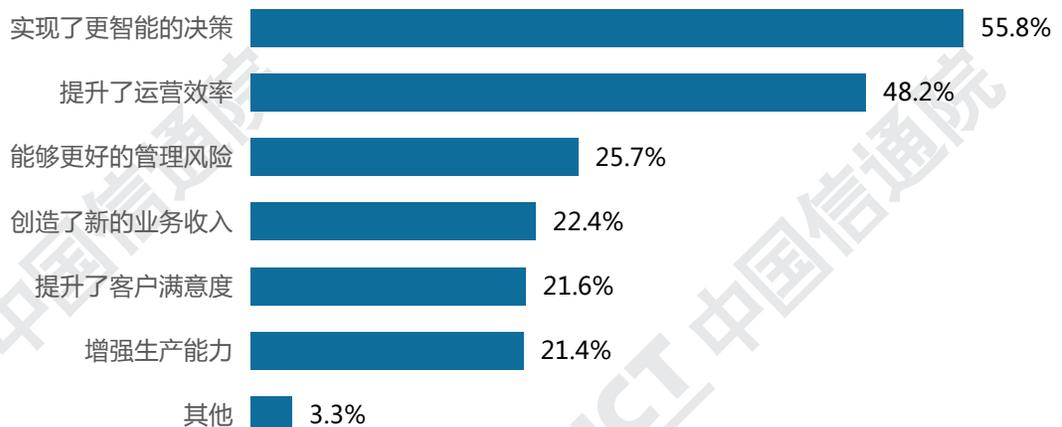


数据来源：中国信息通信研究院

（三）大数据应用带来的效果

大数据应用为企业带来的最明显效果是实现了智能决策和提升了运营效率。调查发现，应用大数据后实现智能决策的企业占比最高，达到55.8%。机器学习和认知计算等技术的发展进一步推动了大数据对企业决策的支撑。其次，48.2%的企业表示应用大数据后提升了运营效率；其他效果还包括：更好的管理风险（25.7%）、创造了新的业务收入（22.4%）、提升了客户满意度（21.6%）以及增强生产能力（21.4%）。

图10 企业应用大数据带来的效果 (N=514)

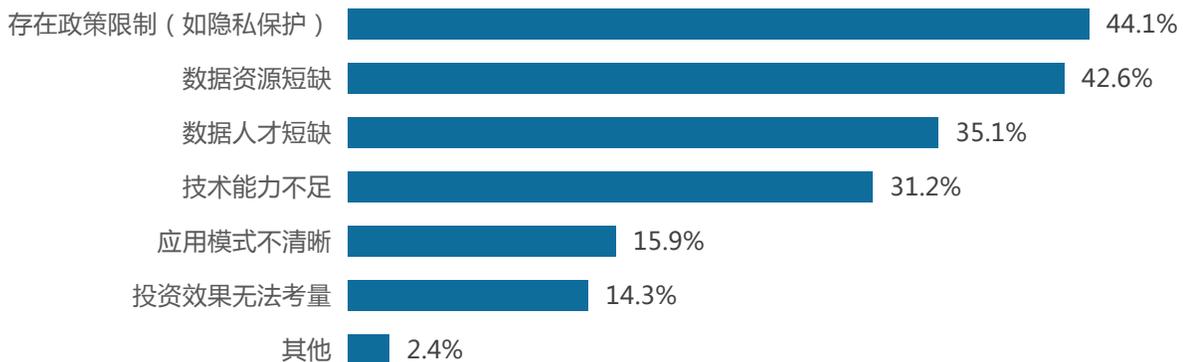


数据来源：中国信息通信研究院

（四）大数据应用的主要障碍

政策限制和数据资源短缺是企业推进大数据应用时最突出的障碍。44.1%的受访企业认为隐私保护等政策限制是制约企业大数据发展的首要障碍，企业隐私保护、数据权属等大数据相关政策尚需完善；其次为数据资源短缺(42.6%)，进一步引导企业推动数据的开放，加强数据共享至关重要；另外，超过1/3的受访企业认为数据人才短缺是企业发展大数据所面临的主要问题之一，大数据专业人才仍有待进一步补充。

图11 制约企业大数据发展主要因素 (N=1,465)

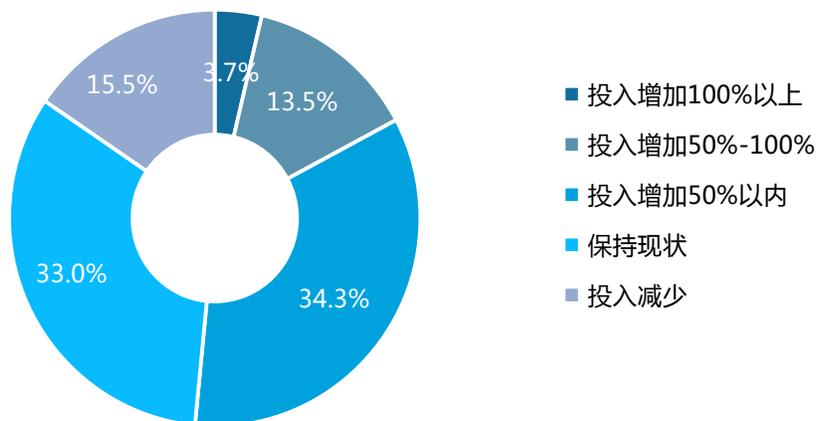


数据来源：中国信息通信研究院

（五）未来大数据投入趋势

企业将进一步加大在大数据领域的投入。随着大数据技术的逐步成熟以及国家政策的大力推进，受访企业普遍看好大数据的发展前景，一半以上的受访企业未来计划加大对大数据的投入，其中近20%的企业预计投入增长在50%以上。

图12 企业预计未来大数据投入趋势 (N=1,465)



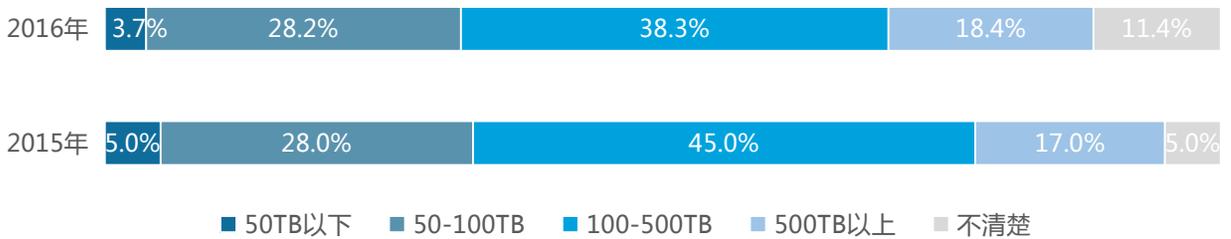
数据来源：中国信息通信研究院

四、大数据资源

（一）数据规模

企业数据资源规模增长迅速。调查发现，目前约70%的企业拥有的数据资源总量在50-500TB之间。与2015年相比，企业数据资源总量呈增长趋势，2016年数据量在500TB以上企业占比为18.4%，较2015年增加1.4%；数据资源在50TB以下的企业占比由2015年的5.0%下降到2016年的3.7%。

图13 企业数据资源总量 (N=1,465)

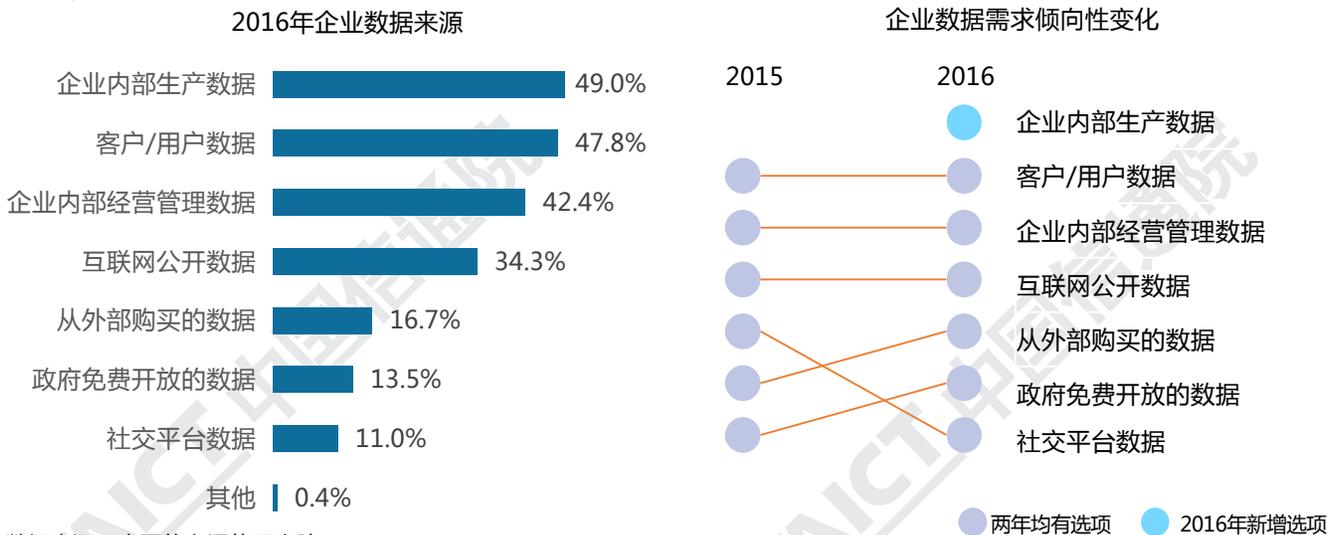


数据来源：中国信息通信研究院

（二）数据来源

企业数据资源主要来自于企业内部。调查发现，企业数据主要来源包括内部生产数据(49.0%)、客户/用户数据(47.8%)、内部经营管理数据(42.4%)。与2015年相比，企业对外部购买的数据和政府免费开放的数据需求越来越多。但外部购买的数据(16.7%)和政府免费开放的数据(13.5%)的应用比例仍处于较低水平。

图14 企业数据来源情况 (N=1,465)

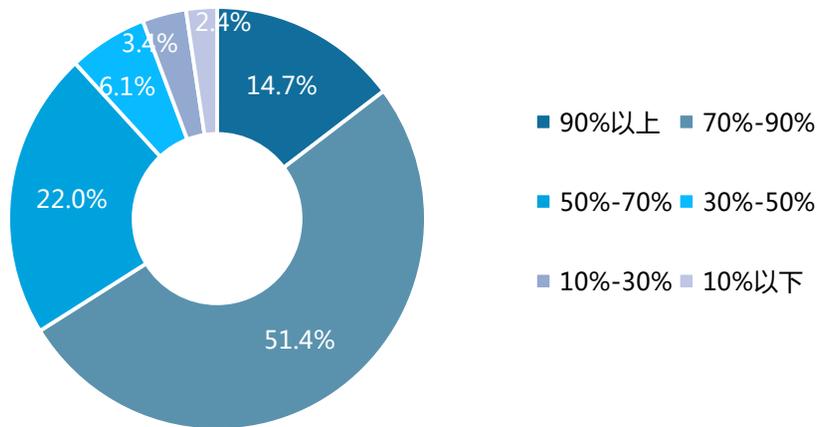


数据来源：中国信息通信研究院

（三）数据类型

企业绝大部分数据资源为非结构化数据。66.1%的受访企业表示非结构化数据的比例在70%以上。22.0%的企业非结构化数据比例为50% - 70%。根据访谈，在企业新增数据中非结构化数据的增速远高于结构化数据。

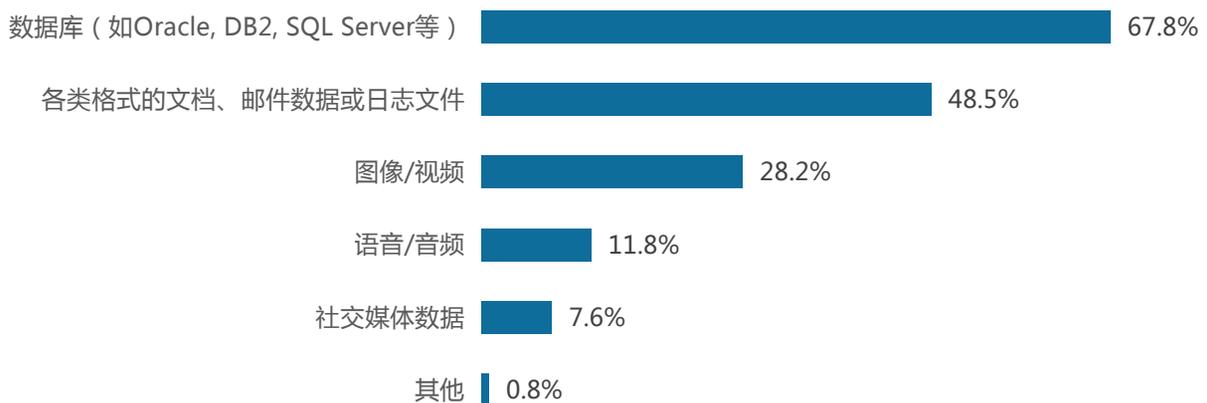
图15 企业非结构化数据比例 (N=1,465)



数据来源：中国信息通信研究院

但企业非结构化数据的分析和利用仍有待提高。67.8%的受访企业表示目前仍是基于结构化数据为基础的数据库进行大数据分析。仅有28.2%和11.8%的企业对图像/视频、语音/音频等非结构化数据进行了分析。随着大数据技术和处理能力的提升，未来企业对非结构化数据的利用率将有望进一步得到提升。

图16 企业各类数据分析利用情况 (N=1,465)



数据来源：中国信息通信研究院

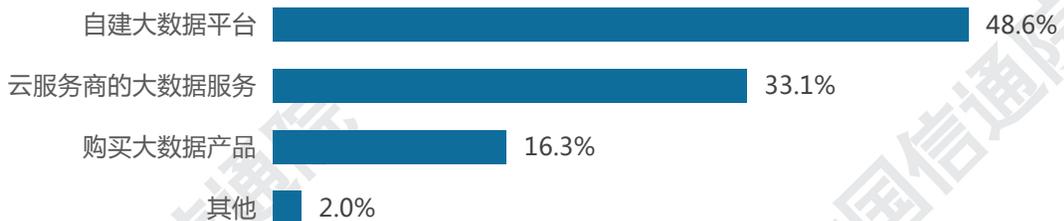
五、大数据平台建设模式

（一）建设模式占比

本次调研根据大数据平台建设模式分为本地部署大数据平台和采购公共云服务模式。本地部署大数据平台包括：1、购买大数据相关产品或解决方案，在本地建设大数据相关基础设施、存储和计算平台、分析工具和应用展示等；2、采购大数据服务提供商的集成产品，如大数据一体机。采购公共云服务指的是购买某一家大数据云服务提供商所提供的在线大数据服务。

企业更倾向本地部署大数据平台。64.9%的受访企业选择本地部署大数据平台，其中倾向选择自建大数据平台的企业占比达到48.6%。另有33.1%的企业倾向通过采购云服务的方式应用大数据，随着以云计算为基础的大数据应用逐渐成熟，未来该比例有望进一步提高。

图17 企业大数据平台建设模式占比 (N=1,465)



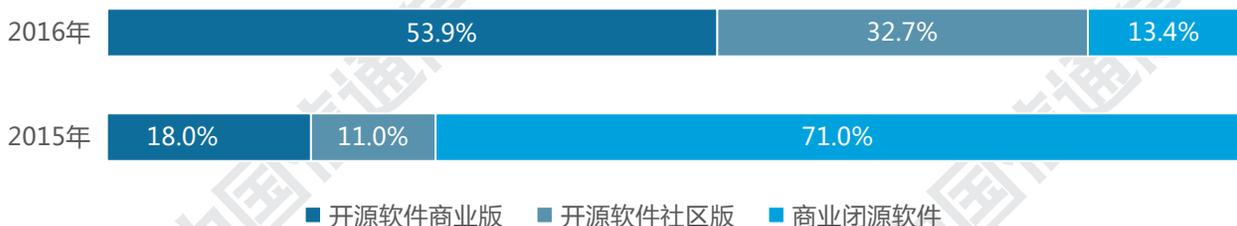
数据来源：中国信息通信研究院

（二）自建大数据平台模式

1、软件类型选择

企业对开源软件的接受度大幅提高。相比于商业闭源软件，企业更青睐开源软件，86.6%的受访企业倾向使用开源软件搭建大数据平台，与2015年相比大幅提高。其中，受访企业更青睐开源软件商业版，53.9%的企业选择了商业版开源软件，相比社区版高21.2%。

图18 自建大数据平台软件类型选择 (N=1,465)

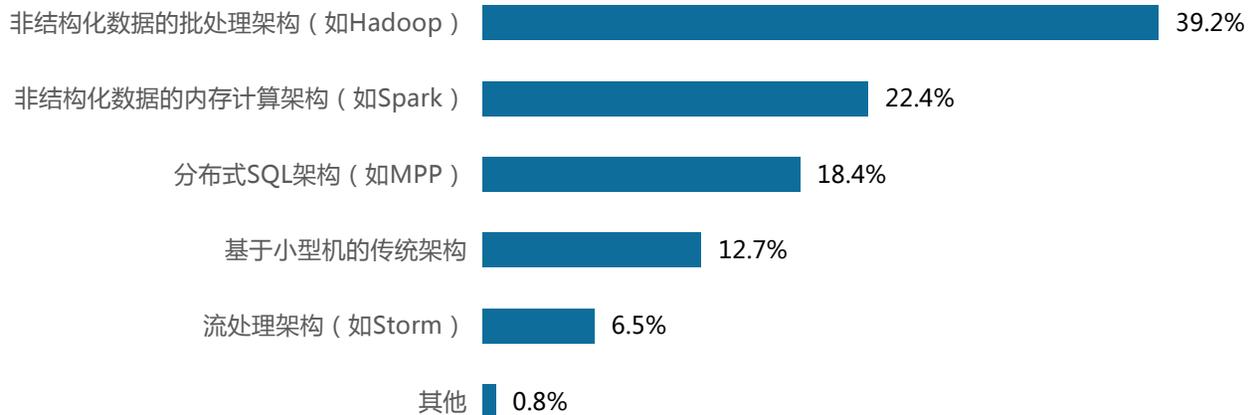


数据来源：中国信息通信研究院

2、平台架构选择

Hadoop 是最受企业欢迎的大数据平台架构。调查发现，接近四成的受访企业倾向选择 Hadoop 作为大数据平台的技术架构；其次，Spark 近两年发展迅速，22.4% 的企业倾向选择 Spark 作为大数据平台的技术架构。

图19 自建大数据平台技术架构选择 (N=1,465)

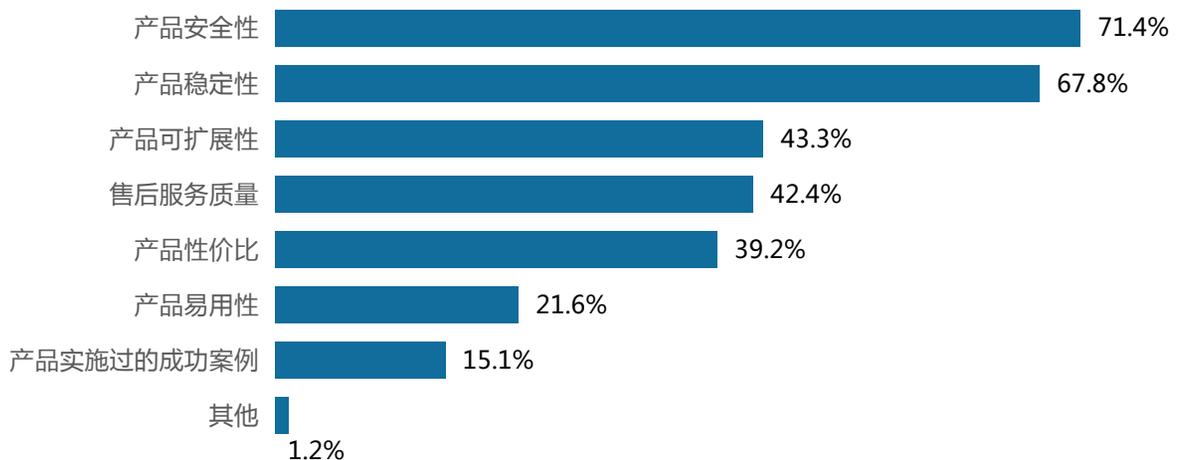


数据来源：中国信息通信研究院

3、选择大数据软件的考虑因素

企业选购大数据软件产品时对安全性、稳定性和可扩展性最为关注。调查发现，受访企业在选择大数据软件产品时对安全性的关注度最高，达到 71.4%；其次是产品的稳定性 (67.8%)、产品可扩展性 (43.3%)。

图20 自建大数据平台软件选择的考虑因素 (N=1,465)

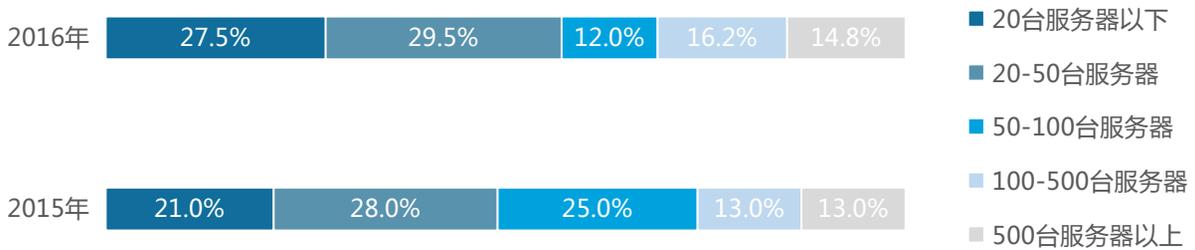


数据来源：中国信息通信研究院

4、服务器规模

服务器数量在 100 台以上的企业增加。调查发现，31.0% 的受访企业用于大数据平台的服务器数量在 100 台以上，相比 2015 年显著提高。由此看出，企业在加大大数据领域的投入。

图21 自建大数据平台的服务器数量 (N=291)

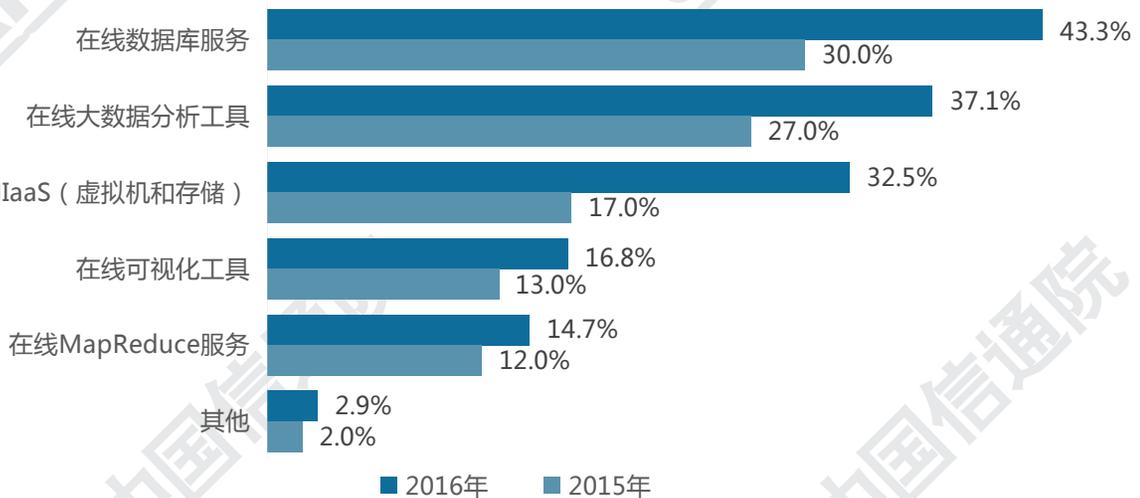


数据来源：中国信息通信研究院

（三）采购公共云服务模式

在线数据库服务、在线大数据分析工具和通用 IaaS 是企业通过公共云服务部署大数据应用时的优先选择。调查发现，43.3% 的受访企业采购云服务时倾向选择在线数据库服务；其次是在线大数据分析工具以及通用 IaaS，分别占比 37.1% 和 32.5%。随着企业大数据应用的逐渐深入，与 2015 年相比，企业购买各项云服务产品的比例显著提高。

图22 企业购买云服务产品类型选择 (N=1,465)



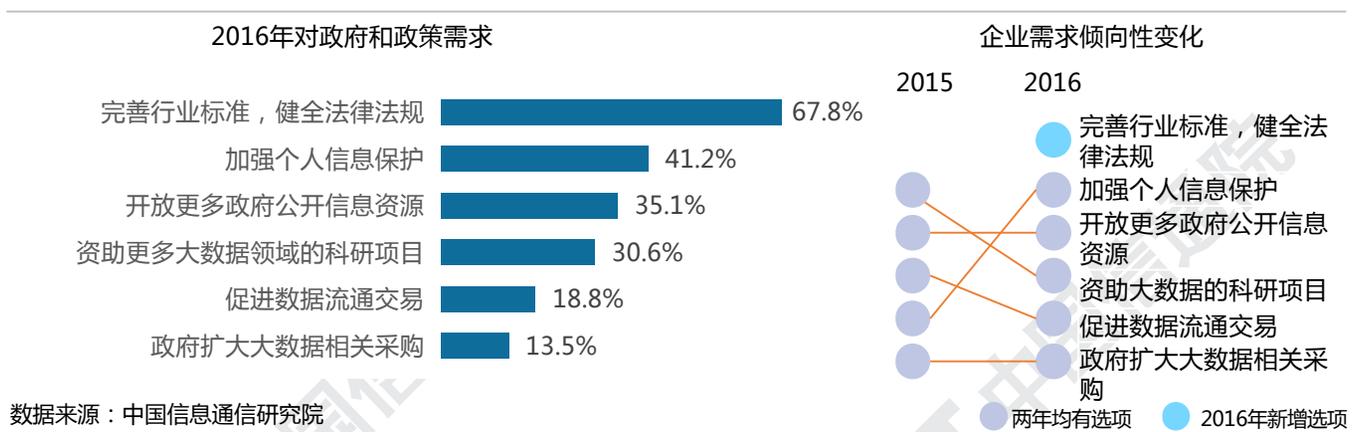
数据来源：中国信息通信研究院

六、政策需求和资源需求

（一）政策需求

完善行业标准、健全法律法规对推动大数据发展至关重要。在企业对大数据发展的政策需求调查中，完善行业标准、健全法律法规是2016年新增选项，也是企业关注度最高的选项，可见大数据行业标准亟需进一步完善；其次为加强个人信息保护（41.2%）和开放更多政府公开信息资源（35.1%）；其他需求还包括资助更多大数据领域的科研项目（30.6%）、促进数据流通交易（18.8%）以及扩大大数据相关采购（13.5%）。

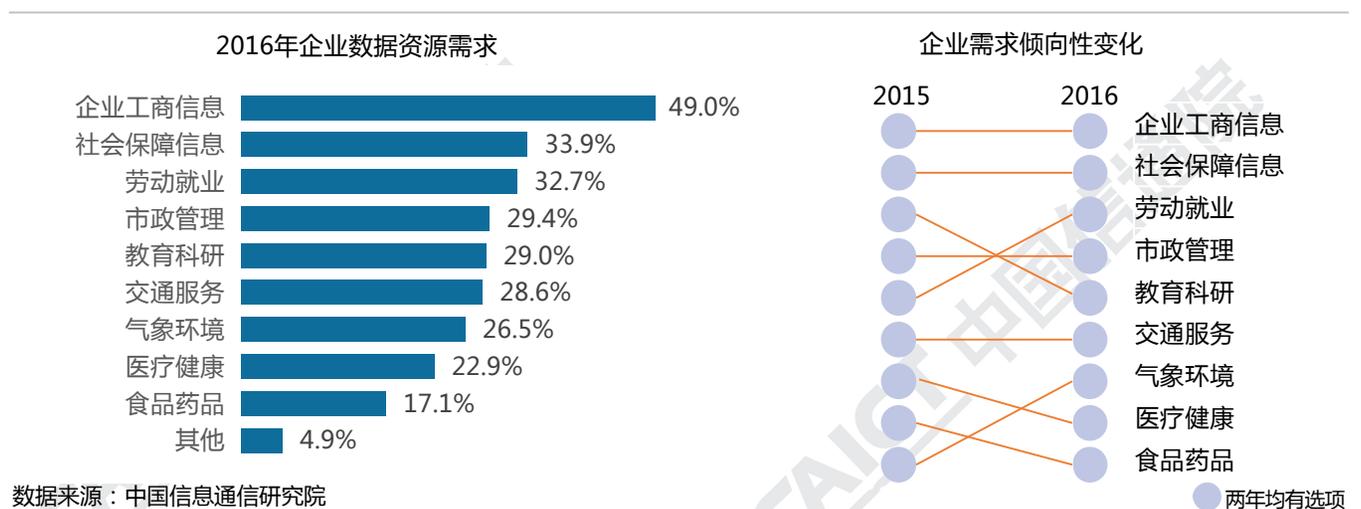
图23 企业对政府和政策的需求 (N=1,465)



（二）资源需求

企业对数据资源的需求愈加强烈。其中企业工商信息依旧是需求最大的数据资源，占比达到49.0%；其次为社会保障数据（33.9%）、劳动就业数据（32.7%）以及市政管理数据（29.4%）。相比于2015年，企业对劳动就业、气象环境等数据资源的需求比例进一步上升。

图24 企业对数据资源的需求 (N=1,465)



编后语

中国信息通信研究院已经连续两年开展中国大数据发展情况的市场调查工作，并发布《中国大数据发展调查报告》，在行业内也产生了重要影响。随着大数据技术的发展和大数据应用的深入，我院将持续跟踪大数据发展情况。欢迎大数据领域广大企业、研究机构、专家学者和从业人员与报告编写组联系（发邮件至：liuhan@caict.ac.cn），提出您感兴趣的调研问题，反馈您的宝贵意见和建议，帮助我们持续提升报告的针对性和实用价值。

感谢您对大数据产业的支持！

中国信息通信研究院

2017 年 3 月



关注我们

中国信息通信研究院

地 址：北京市海淀区花园北路52号

邮政编码：100191

联系电话：010-62304839

传 真：010-62304980

